DOI:10.1021/jf9013235

JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY
ARTICLE

Classification of Wild and Farmed Salmon Using Bayesian Belief Networks and Gas Chromatography-Derived Fatty Acid Distributions

David E. Axelson,[†] Inger B. Standal,[‡,§] Iciar Martinez,[‡,∥] and Marit Aursand*[,‡,§]

[†]MRi_Consulting, 8 Wilmot Street, Kingston, Ontario, Canada K7L 4 V1, [‡]SINTEF Fisheries and Aquaculture, 7465 Trondheim, Norway, [§]Department of Biotechnology, Norwegian University of Natural Science and Technology, 7491 Trondheim, Norway, and [∥]Department of Marine Biotechnology, Norwegian College of Fishery Science, University of Tromsø, 9037 Tromsø, Norway

In this study, we present the use of Bayesian Belief Networks (BBN) for the classification of wild versus farmed Atlantic salmon (*Salmo salar* L.). Using a data set of 131 salmon samples from several geographical origins and the gas chromatography-derived distributions of 12 fatty acids (FAs), a Bayesian Belief Network was constructed, ultimately using only the three most important FAs (16:1n-7, 18:2n-6, and 22:5n-3). The training data set yielded a prediction error of 0% (68/68 farmed; 20/20 wild correct) while the validation data set prediction error was 4.65% (32/32 farmed; 9/11 wild correct). Different randomly chosen validation sets yielded similar prediction accuracies. This model was then applied to 30 market (store-bought) samples where predictions were compared with the product labels.

KEYWORDS: Bayesian Belief Network; authentication; Atlantic salmon; *Salmo salar* L.; fatty acids; wild; farmed

INTRODUCTION

In the European Union, the common organization of the markets in fishery and aquaculture products comes under Council Regulation (EC) 104/2000. In October 2002, Commission Regulation (EC) 2065/2001 was adopted that details the labeling, packaging, and traceability requirements for fishery and aquaculture products. The information includes specification of the commercial designation and scientific name, method of production of species, and the area in which fish were caught. There is a clear trend in the international market to labeling products with information about composition and quality. This, together with the increasing production and consumption of fish products including salmon, both from farmed and wild fish has led to an increasing demand for effective standardized analytical methods for authentication of fish products.

A distinction in terms of quality and, notably, price between the wild-fished and farmed products leaves open a real possibility of fraud. Wild salmon, however, is still perceived by many to be superior eating compared to farmed salmon, and because of the much restricted availability compared to the farmed fish, wild salmon typically commands a price 2 to 3 times that of the farmed equivalent. With such a price difference, there is a temptation to mislabel farmed fish as wild.

Different species of fish exhibit characteristic fatty acid (FA) profiles (*1*). The FA profiles examined on methyl esters by gas chromatography (GC) have been used as natural markers for stock identification through the analysis of heart, muscle, or brain tissue for different fish species (*2*). The FA composition of the muscle lipid of all fish varies according to the season (*3*), geographical location of the catch (*4*), diet and feeding (*5*,*6*), size (*7*,*8*), sex (*9*), state of their reproductive cycle (*10*−*13*), temperature (*14*), age, maturity, and salinity. In addition, the FA composition of commercial fishfeed is usually very different from the naturally available food sources in aquatic environments; therefore, the tissue FA composition has been used widely to discriminate farmed and wild-caught fish. FA compositions have been used to discriminate wild and cultured largemouth bass, black crappies, and white crappies (*15*), striped bass and hybrid striped bass (*16*), carp and rainbow trout (*17*), red drum (*18*), eight species of seawater fishes (*19*), gilthead sea bream (*20*), sweet smelt (*21*), sea bass (*22*), sea bream (*23*), and salmon (*24*−*28*).

Since September 2001, a European consortium of five partners from France, Italy, the United Kingdom, and Norway has been working to develop a validated method to enable official laboratories to discriminate between wild and farmed salmon and farming conditions (geographical origin). The analytical methodologies involved in this project (COFAWS; Confirmation of the Origin of Farmed and Wild Salmon and Other fish) included stable isotope analysis by SNIF-NMR (site specific natural isotope fractionation studied by nuclear magnetic resonance spectroscopy) and IRMS (isotope ratio mass spectrometry) of the fish oil, water from the fish, and other parts of the fish (*29*); ^{1}H and ^{13}C NMR profiling (*27*,*30*); and determination of FA content by GC. Statistical approaches to data evaluation included analysis of variance, correlation analysis, principal component

*Corresponding author. Tel: + 47 48200158. Fax: + 47 93270701. E-mail: Marit.Aursand@sintef.no.

 Published on Web 08/05/2009

analysis, quadratic discriminant analysis, probabilistic and Kohonen neural networks, Support Vector Machines, and partial least-squares discriminant analysis (*27*,*30*). The aim of the present study was to test the possibility of using only GC in combination with Bayesian Belief Network analysis, as a validated method to enable discrimination between farmed and wild salmon and to verify the origin of market samples. This approach has shown its great potential to identify the farm origin of farmed fish and possibly also escaped farmed fish from wild fish (*31*).

There are many potential advantages of the BBN approach as a classifier (*32*, *33*). Most significantly, one can calculate explicit probabilities for a sample belonging to a given class. Prior knowledge can be combined with observed data and the input may be discrete, categorical, or continuous. One can also find the variables with the highest impact possibly allowing one to learn about causal relationships. In practice, much of the complexity of setting up a network model can be mitigated by the fact that raw data can automatically be used to define model parameters. Since the network involves all variables from all samples, it can readily handle incomplete data sets; in this regard, the method is particularly useful for making probabilistic inferences about models characterized by inherent complexity and uncertainty. Finally, one can create a standalone, executable classification program that allows even casual users to implement an appropriately trained network.

MATERIALS AND METHODS

Fish Samples. Wild Atlantic salmon (*Salmon salar* L.) were obtained from Norway, Scotland, and Canada. Farmed Atlantic salmon were obtained from two different Norwegian, Scottish, Irish, Tasmanian, and Canadian farms. Fish from feeding trials run at North Atlantic Fisheries College, Port Arthur, Scalloway, Shetland, U.K., are included in the data set. Market samples were collected from supermarkets in Italy, the United Kingdom, and Norway, although the results noted here relate only to the market samples obtained from Norway (for which GC data were available).

Lipid Extraction. Lipid extraction from white muscle of salmon was performed according to a modified Bligh and Dyer procedure (*29*, *34*).

Gas Chromatography. *Preparation and Analysis of FA Methyl Esters* (*FAMEs*). The lipids were first transesterified with boron trifluoride-methanol and 0.5 M methanolic sodium hydroxide, and then the FA methyl esters (FAMEs) were extracted into hexane (AOCS Method CE 2-66). An internal standard 21:0 methyl ester (purity >99%, Nu-Chek. Prep. Inc.) was added to the extracted sample prior to methylation. This standard was chosen because it fitted the GC conditions best. FAMEs were analyzed on a Fison 8160 (Fisons Instruments S.pA. Milan, Italy) capillary gas chromatograph equipped with capillary cold on-column injector, a fused silica capillary column, and Omegawax 320 (30 m, 0.32 mm id, 0.25 μm film thickness; Supelco Inc., Bellefonte, PA) connected to a flame ionization detector (FID). The FID was connected to a computer implemented with Chrom-card for Windows 1.21 software. The gas chromatograph was provided with an AS800 autosampler. The oven temperature was increased from 80 to 180 °C at 25 °C min^{-1} and held for 2 min. Then the temperature of the oven was increased by 2.5 °C min^{-1} to 205 °C (held for 8 min) and up to 215 °C min^{-1} and held for 3 min. The temperature of the detector was 250 °C. Hydrogen was used as carrier gas at a flow rate of 1.6 mL min^{-1}. FAMEs were identified by the comparison of their retention times with those of a reference solution (Nu-Chek-Prep, Elysian, MN) chromatographed at identical GC conditions. FAs measured included 18:2n-6, 16:1n-7, 22:6n-3, 18:1n-7, 14:0, 22:5n-3, 16:0, 18:0, 22:1n-11, 20:1n-9, 20:5n-3, and 18:1n-9.

Statistical Analysis. All calculations were performed using Netica v4.02 (Norsys Software Corp., Vancouver B.C. Canada). Eighty-eight samples (approximately two-thirds) were randomly chosen and assigned to the training data set, and the remaining 43 samples were assigned to the validation data. The validation data set was not, in any way, used for the generation of the model.

A Bayesian net (*32*, *33*) is a graph-based model for representing probabilistic relationships among random variables. One way of expressing Bayes' rule in the present context includes a hypothesis (to which class the sample belongs), past experience, and evidence, (the detailed FA component analysis):

$$P(H|E,c) = \frac{P(H|c) \times P(E|H,c)}{P(E|c)}$$

where we can update our belief in hypothesis H (the class or fish species) given the additional evidence E, and the background context (past experience), c. The left-hand term, $P(H|E,c)$ is called the *posterior probability*, or the probability of hypothesis H after considering the effect of the evidence E on past experience c. This corresponds to determining the probability that the sample belongs to one of the possible classes (wild, farmed), given its FA composition. The term $P(H|c)$ is called the *a-priori probability* of H given c alone. This is the probability of any node being in one state or another without current evidence. This value is determined from the training data set. The term $P(E|H,c)$ is called the *likelihood* and gives the probability of the evidence assuming that the hypothesis H and the background information c are true. Finally, the last term $P(E|c)$ is independent of H and can be regarded as a normalizing factor. In the present context, we wish to determine the probability that a given sample characterized by a certain FA distribution belongs to the wild or farmed category

If one makes the simplified assumption that the attributes are conditionally independent (*35*), we have a Naïve Bayes Classifier (the model applied here), and the following equation applies:

$$P(H|C_i) = \prod_{n}^{k=1} P(h_k|C_i)$$

where the product of occurrence, for example, of 2 elements h1 and h2, given the current class is C, is the product of the probabilities of each element taken separately, given the same class, i.e., $P([h1,h2],C) = P(h1,C) \times P(h2,C)$. This greatly reduces the computation cost, and once the probability $P(H|C_i)$ is known, one assigns H to the class with maximum $P(H|C_i) \times P(C_i)$.

A Bayesian network consists of (1) nodes that represent the random variables, where each node has states (a set of probable values for each variable); (2) directed edges (arrows) that connect the nodes (representing dependencies where the absence of arrows indicates independence; (3) a conditional probability table (CPT) associated with each node (prior probability; probability of any node in the Bayesian belief network being in one state or another without current evidence); and (4) a directed acyclic graph (DAG) where the graph represents independence relationships between variables. In our Bayesian models, there is a path to the class node (species) from every evidence node (FA component). The class node is called the child of every evidence node, and the evidence nodes are called parents of the class node.

Probabilities on some nodes are affected by the state of other nodes. A conditional probability is stated mathematically as $P(x|p1, p2, ..., pn)$, i.e., the probability of variable X in state x given parent P1 in state $p1$, parent P2 in state $p2$, ..., and parent Pn in state pn. That is, for each parent and each possible state of that parent, there is a row in the CPT that describes the likelihood that the child node will be in some state. This table is derived from the training data set. However, each node of a Bayesian classifier must have a finite number of states; therefore, a state of a node representing a continuous valued attribute must be associated with a subrange of the possible values of that attribute. Finding those subranges is called discretizing the attribute because it allows a finite number of states to be associated with selected, discrete subranges of the possible values. We have chosen to create 10 discrete subranges for each component for our analyses as shown in **Figure 1** for the three selected FAs.

Once the Bayesian classifier has been built, classifications of unknown samples can be predicted by specifying the state of each node associated with an attribute whose value is known for the new entity (instantiating the evidence nodes). This is accomplished by changing the probabilities stored at these nodes so that the current state has a probability of 1, and all of the other states have a probability of 0 (i.e., the actual measured FA value for a given component is entered into the calculation) When instantiation is complete, the probabilities of the states of the class node can be recomputed using Bayes' rule. Inference is the process of instantiating

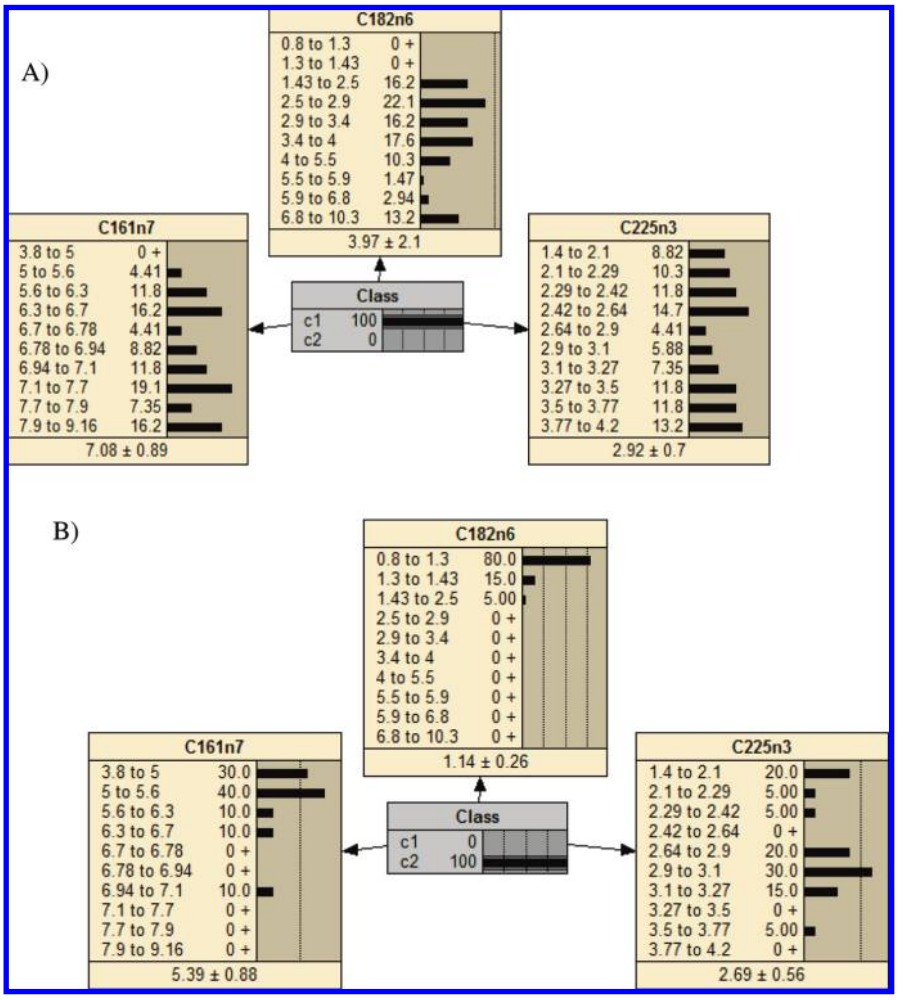

Figure 1. Schematic representation of Bayesian Belief Network for classification of wild vs farmed salmon. (**A**) Fatty acid distribution for all farmed fish only (c1) and (**B**) fatty acid distribution for all wild fish only (c2).

the known evidence nodes and recalculating the probability distribution of the class node, with the result being a new probability distribution for the class node representing the probabilities that the test sample is in each class.

For feature selection, the mutual information (*36*) between each attribute and the class attribute is measured. Mutual information measures the strength of the correlation between the values of the attribute and the values of the class. Mutual information quantifies the distance between the joint distribution of two discrete random variables X and Y and what the joint distribution would be if X and Y were truly independent. Mutual information is a measure of dependence in the following sense: $I(X; Y) = 0$ if and only if X and Y are independent random variables.

While there are many probabilities which must be specified, once the structure of the network is specified, all of these probabilities can be learned from a set of examples from the population being modeled (training set).

RESULTS AND DISCUSSION

A BBN was constructed using data for all 12 FAs measured. The network was then optimized and the relative importance (sensitivity analysis) of all variables calculated. Since using only three of the most important variables, namely, FAs C18:2n6, C16:1n7, and C22:5n6 (**Figure 1**), resulted in almost identical prediction accuracies compared to those using all 12 FAs, we decided to use this simplified model. As indicated in the legend, **Figure 1** also shows the distributions of the FA contents when one assumes that the class is farmed (**Figure 1A**) or wild (**Figure 1B**). Horizontal bars denote relative percentage of values of the given FA for all samples of the indicated class (wild vs farmed). Values at the bottom are the mean and standard deviations for the given FA. There are significant variations in the distributions according to the classification; the wild fish are distinguished from the farmed fish by having lower amounts of 18:2n-6 (linoleic acid) and 16:1n-7 (palmitoleic acid).

There are obvious significant variations in most FAs when comparing farmed and wild salmon. It has been shown that Atlantic salmon absorb FAs selectively during digestion of the diet. Johnsen et al. (*37*) found this to be the case for 18 diets, on the basis of fish oils from five different fish species with differences in FA composition. Low relative amounts of the polyunsaturated acids 20:5n-3, 22:5n-3, and 22:6n-3 in the feces indicated that they were efficiently absorbed, while for saturated and the monounsaturated FAs the degree of absorption decreased with increasing chain length (e.g., C24:1n-9 is poorly absorbed, while C16:1n-7 is almost as well absorbed as the polyunsaturated acids). In a short-term feeding study (*38*), the FA profile of the plasma of Atlantic salmon was found to reflect the diets in three groups, namely, diets of 100% fish oil, 100% rapeseed oil, and a 1:1 blend of the two, differed stepwise with respect to all major groups of FAs. It is clear that the FA composition of fish tissues is primarily affected by the dietary FA composition (*39*, *40*). However, incorporation of FAs into fish tissue may be affected by various metabolic factors such as preferential incorporation, oxidation, lipogenic activity, or FA elongation and desaturation processes (*40*, *41*), which, in turn, may be influenced by growth stage, culture system,

Table 1. Conditional Probability Table for the Three Fatty Acids Used for Classification (16:1n-7, 18:2n-6, and 22:5n-3) of c1 (Farmed Fish) and c2 (Wild Fish)

16:1n-7	3.9 to 5	5 to 5.6	5.6 to 6.3	6.3 to 6.7	6.7 to 6.78	6.78 to 6.94	6.94 to 7.1	7.1 to 7.7	7.7 to 7.9	7.9 to 9.16
c1	1.47e-05	4.412	11.765	16.176	4.412	8.824	11.765	19.118	7.353	16.176
c2	30.000	40.000	10.000	10.000	5.00e-05	5.00e-05	10.000	5.00e-05	5.00e-05	5.00e-05
18:2n-6	0.8 to 1.3	1.3 to 1.43	1.43 to 2.5	2.5 to 2.9	2.9 to 3.4	3.4 to 4	4 to 5.5	5.5 to 5.9	5.9 to 6.8	6.8 to 10.3
c1	1.47e-05	1.47e-05	16.176	22.069	16.176	17.647	10.294	1.471	2.941	13.235
c2	80.000	15.000	5.000	5.00e-05	5.00e-05	5.00e-05	5.00e-05	5.00e-05	5.00e-05	5.00e-05
22:5n-3	1.4 to 2.1	2.1 to 2.29	2.29 to 2.42	2.42 to 2.64	2.64 to 2.9	2.9 to 3.1	3.1 to 3.27	3.27 to 3.5	3.5 to 3.77	3.77 to 4.2
c1	8.824	10.294	11.765	14.706	4.412	5.882	7.353	11.765	11.765	13.235
c2	20.000	5.000	5.000	5.00e-05	20.000	30.000	15.000	5.00e-05	5.000	5.00e-05

Table 2. Sensitivity of Class Due to a Finding at Another Node

node	mutual info
C18:2n6	0.68876
C16:1n7	0.22172
C22:5n3	0.13799
C22:1n11	0.13146
C18:1n7	0.08747
C14:0	0.06300
C22:6n3	0.05612
C20:5n3	0.03090
C16:0	0.02955
C18:1n9	0.02635
C20:1n9	0.02148
C18:0	0.01650

and general environmental conditions (*40*). The possible complexity of these relationships with FA distributions also implies that the classification methods applied must be capable of dealing with them in a robust way.

Table 1 gives the conditional probabilities for the three variables used for classification, which provide the quantitative connection between the two classes (wild vs farmed) and the actual ranges of values of the FAs. Each column corresponds to the given fatty acid concentration range shown in the heading. Each row corresponds to the category (c1, farmed or c2, wild) and each number is a conditional probability. The most important probabilities for each class can be noted, demonstrating the systematic differences between classes.

Using all samples and all variables, the BBN model correctly classified 100 of 100 farmed samples and all 30 of 31 wild salmon, for an error rate of 0.76%. Variable subset selection was then applied to find the minimum number of FA components that resulted in similarly accurate classification predictions. Employing only the top 3 variables (16:1n-7, 18:2n-6, and 22:5n-3) (**Table 2**), the BBN model correctly predicted the class of 100 of the 100 farmed fish and 29 out of 31 wild fish for an error rate of 1.53%. The market sample predictions were the same as that for the model using all variables. The use of all samples does not necessarily reflect the ability of the model to predict samples that were not included in the development of the classifier. One of the options for validation is to extract a subset of samples and hold them out during the development of the model, and subsequently make predictions of these samples using the model in question. Therefore, to better estimate the robustness of the BBN model, we divided the samples into two sets: a training set with 88 samples and a validation set with 43 samples. The samples were randomly chosen. Using the top 3 variables again, 68 out of 68 farmed and 20 of 20 wild fish in the training set were correctly predicted (error rate of 0%). For the validation set, 32 out of 32 farmed and 9 out of 11 wild salmon were correctly predicted for an error rate of 4.65%.

Table 3. Prediction of Sample Class of Market Samples (Norway)[a]

store	market label	prediction correct
S1	F	Y
S2	F	Y
S3	F	Y
S4	F	Y
S5	F	Y
R1	F	Y
R2	F	Y
R3	F	Y
R4	F	Y
R5	F	Y
U1	F	Y
U2	F	Y
U3	F	Y
U4	F	Y
U5	F	Y
M1	W	Y
M2	W	Y
M3	W	Y
M4	W	Y
M5	W	Y
Ra1	F	Y
Ra2	F	Y
Ra3	F	Y
Ra4	F	Y
Ra5	F	Y
Ma1	F	Y
Ma2	F	Y
Ma3	F	N
Ma4	F	Y
Ma5	F	N

[a] W, wild; F, farmed; Y, yes (correct prediction); N, no (incorrect prediction).

To further examine the robustness of the defined model, different numbers of randomly chosen training and validation data were also tested. Splitting the data almost in half with 66 training samples and 65 validation samples, and using the top 3 variables again, 51/51 farmed and 13/15 wild in the training set were correctly predicted. For the validation set, 49/49 farmed and 16/16 wild were correctly predicted.

One of the primary goals of this classification procedure is to identify market samples that may not be correctly identified regarding their production method. **Table 3** summarizes the predictions for all of the market samples, and we note that the majority of samples appear to be correctly labeled. These results are also largely consistent with a related study involving market sample authentication (*27*). For the set of 30 market samples, 25 of them were labeled as farmed and 5 as wild. The 5 samples labeled wild were predicted to be wild, while 2 of the 25 labeled as farmed were actually predicted to be wild by the model (**Table 3**).

The two fish labeled as farmed but predicted to be wild displayed relatively low levels of 18:2n-6 (below 1.3% in contrast to the reference farmed fish, all with levels above 2%). Since the FA profile of the fish triacyl glycerols is a mirror of its diet (*42*), this indicates that these two fish have been given a feed with low 18:2n-6 content prior to slaughter. A second possibility of this miclassification could be that the reference set of wild fish accidently contained escaped farmed fish (or wild fish feeding around farms), which would lead to an erroneous model, but because of the low level of 18:2n-6 of the reference samples of wild fish (all below 1.5%), this is not the most likely explanation. The five market samples (M1−M5) classified as wild here (corresponding to their original market labels) were classified as farmed by corresponding ^{13}C NMR analysis (*30*). The origins of these differences are being investigated. Our results agree with Megdal et al. (*43*) in that 18:2n-6 is a good marker for our farmed fish, but the actual distribution of 18:2n-6 concentrations in our samples varies from those reported by Megdal et al. We must be aware of possible changes in the FA profile as the feed composition is improved and alternative ingredients are selected by feed manufacturers that, optimally, should be more environmentally friendly, cheaper, and induce a more natural FA profile in the fish (for example, by using oil from *Calanus finmarchicus* (*44*)). Moreover, the FA composition of the feed used during the period prior to slaughter is usually modified to highly resemble that of wild fish, with increased EPA and DHA profiles: Torstensen (*45*, *46*) examined the effect of replacing dietary fish (capelin) oil with increasing amounts of rapeseed oil and olive oil in a 42 week feeding trial in seawater in Norway using postsmolt fish and showed that after a washout period of 1788 day degrees when the salmon diets used contained only fish oil, only the fish that had been fed the highest amount of rapeseed oil (75% and 100% of the oil source) retained higher levels of C18:1n-9 and C18:2n-6.

In conclusion, the BBN method provides a realistic alternative method for effective classification and authentication of wild and farmed salmon where explicit probabilities are being calculated for each class. A robust classification model should encompass a sufficient number of relevant fatty acids to ensure, as much as possible, that observed FA variations can be handled. Models must be continually updated as databases are expanded. The more tools available for authentication, the more likely that such consumer fraud will be minimized. This method also has potential applications for similar investigations involving FA distributions for species authentication and even the assessment of the origin of fish that may have escaped from farms and are subsequently caught in open waters (*31*).

ACKNOWLEDGMENT

We thank Pheroze Jungalwalla and Tassal Ltd., Tasmania, for supplying samples of cultivated salmon.

LITERATURE CITED

(1) Ackman, R. G. Fish Lipids. In *Advances in Fish Science and Technology*. Connell, J. J., Ed.; Fishing New Books: Farnham, U.K., 1988; pp 86−103.

(2) Grahl-Nielsen, O. Fatty Acid Profiles As Natural Marks for Stock Identification. In *Stock Identification Methods; Applications in Fishery Science*. Cadrin, S. X., Friedland, K. D., Waldman, J. R., Eds.; Elsevier Academic Press: San Diego, CA, 2005; pp 247−272.

(3) Uysal, K.; Aksoylar, M. Y. Seasonal variations in fatty acid composition and the n-6/-3 fatty acid ratio of pikeperch (*Sander lucioperca*) muscle lipids. *Ecol. Food Nutr.* **2005**, *44*, 23–35.

(4) Çelik, M.; Diler, A.; Küçükgülmez, A. A comparison of the proximate compositions and fatty acid profiles of zander (*Sander lucioperca*) from two different regions and climatic conditions. *Food Chem.* **2005**, *92*, 637–641.

(5) Cai, Z.; Curtis, L. R. Effects of diet and temperature on food consumption, growth rate and tissue fatty acid composition of triploid grass carp. *Aquaculture* **1990**, *88*, 313–327.

(6) Gallagher, M. L.; Paramore, L.; Alves, D.; Rulfson, R. A. Comparison of phospholipid and fatty acid composition of wild and cultured striped bass eggs. *J. Fish Biol.* **1998**, *52*, 1218–1228.

(7) Agren, J. P.; Muje, O.; Hanninen, J.; Herranen, J.; Penttila, I. Seasonal variations of lipid fatty acids of Boreal fresh-water fish species. *Comp. Biochem. Physiol.* **1987**, *88*, 905–909.

(8) Silva, S. S. D.; Gunasekera, R. M.; Collins, R.; Ingram, B. A.; Austn, C. M. Changes in the fatty acid profile of the Australian short fin eel in relation to development. *J. Fish Biol.* **1997**, *50*, 992–998.

(9) Luzzana, U.; Serrini, G.; Morretti, V. M.; Grimaldi, P.; Paleari, M. A.; Valfre, F. Seasonal variations in fat content and fatty acid composition of male and female coregonid 'bondella' from Lake Maggiore and landlocked shad from Lake Como (Northern Italy). *J. Fish Biol.* **1996**, *48*, 352–366.

(10) Almansa, E.; Martian, M. V.; Cejas, J. R.; Badi, P.; Jerez, S.; Lorenzo, A. Lipid and fatty acid composition of female gilt-head sea bream during their reproductive cycle: effects of a diet lacking n - 3 HUFA. *J. Fish Biol.* **2001**, *59*, 267–286.

(11) Shirai, N.; Suzuki, H.; Toukairin, S.; Wada, S. Spawning and season affect lipid content and fatty acid composition of ovary and liver in Japanese catfish (*Silurus asotus*). *Comp. Biochem. Physiol. B.* **2001**, *129*, 185–195.

(12) Shirai, N.; Suzuki, H.; Tokairin, S.; Ehara, H.; Wada, S. Dietary and seasonal effects on the dorsal meat lipid composition of Japanese (*Silurus asotus*) and Thai catfish (*Clarias macrocephalus* and hybrid *Clarias macrocephalus* and *Clarias galipinus*). *Comp. Biochem. Physiol., A* **2002**, *132*, 609–619.

(13) Uysal, K.; Yerlikaya, A.; Aksoylar, M. Y.; Yontem, M.; Ulupinar, M. Variations in fatty acids composition of pikeperch (*Sander lucioperca*) liver with respect to gonad maturation. *Ecol. Freshwater Fish* **2006**, *15*, 441–445.

(14) Uysal, K.; Bülbül, M.; Dönmez, M.; Seckin, A. K. Changes in some components of the muscle lipids of three freshwater fish species under natural extreme cold and temperate conditions. *Fish Physiol. Biochem.* **2008**, *34*, 455–463.

(15) Tritt, K. L.; O'bara, C. J.; Wells, M. J. M. Chemometric discrimination among wild and cultured age-0 largemouth bass, black crappies, and white crappies based on fatty acid composition. *J. Agric. Food Chem.* **2005**, *53*, 5304–5312.

(16) Jahncke, M. L.; Seaborn, G. T.; Smith, T. I. J. Marine forensics program: Development of a biochemical method to distinguish wild from cultured fish: Final report. *NOAA Technology Mem. NMFS-SEFSC-301*, **1992**

(17) Suzuki, H.; Okazaki, K.; Hayakawa, S.; Wada, S.; Tamura, S. Influence of commercial dietary fatty acids on polyunsaturated fatty acids of cultured freshwater fish and comparison with those of wild fish of the same species. *J. Agric. Food Chem.* **1986**, *34*, 58–60.

(18) Villarreal, B. W.; Rosenblum, P. M.; Fries, L. T. Fatty acid profiles in red drum muscle: Comparison between wild and cultured fish. *Trans. Am. Fish. Soc.* **1994**, *123*, 194–203.

(19) Moon, S.-K.; Choi, B.-D.; Jeong, B.-Y. Comparison of lipid classes and fatty acid compositions among eight species of wild and cultured seawater fishes. *J. Fish. Sci. Technol.* **2000**, *3*, 118–125.

(20) Grigorakis, K.; Alexis, M. N.; Taylor, K. D. A.; Hole, M. Comparison of wild and cultured gilthead sea bream (Sparus aurata); composition, appearance, and seasonal variations. *Int. J. Food Sci. Technol.* **2002**, *37*, 477–484.

(21) Jeong, B.-Y.; Moon, S.-K.; Jeong, W.-G.; Ohshima, T. Lipid classes and fatty acid compositions of wild and cultured sweet smelt Plecoglossus altiVelis muscles and eggs in Korea. *Fish. Sci.* **2000**, *66*, 716–724.

(22) Alasalvar, C.; Taylor, K. D. A.; Zubcov, E.; Shahidi, F.; Alexis, M. Differentiation of cultured and wild sea bass (Dicentrarchus labrax): Total lipid content, fatty acid, and trace mineral composition. *Food Chem.* **2002**, *79*, 145–150.

(23) Rueda, F. M.; Hernandez, M. D.; Aguado, F.; Garcia, B.; Martinez, F. J. Differences in tissue fatty acid composition between reared and

wild sharpsnout sea bream, Diplodus puntazzo. *Br. J. Nutr.* **2001**, *86*, 617–622.
(24) Ackman, R. G.; Takeuchi, T. Comparison of fatty acids and lipids of smolting hatchery-fed and wild Atlantic salmon, *Salmo salar*. *Lipids* **1986**, *21*, 117–120.
(25) Aursand, M.; Axelson, D. Origin Recognition of Wild and Farmed Salmon (Norway and Scotland) Using 13C NMR Spectroscopy in Combination with Pattern Recognition Techniques. In *Magnetic Resonance in Food Science: A View to the Future*; Webb, G. A., Belton, P. S., Gil, A. M., Delgadillo, I., Eds.; RSC Books: London, U.K., 2001; pp 227−231.
(26) Aursand, M.; Mabon, F.; Martin, G. J. Characterization of farmed and wild salmon (*Salmo salar*). *J. Am. Oil Chem. Soc.* **2000**, *77*, 659–666.
(27) Masoum, S.; Malabat, C.; Jalali-Heravi, M.; Guillou, C.; Rezzi, S.; Rutledge, D. N. Application of support vector machines to 1H NMR data of fish oils: methodology for the confirmation of wild and farmed salmon and their origins. *Anal. Bioanal. Chem.* **2007**, *387*, 1499–1510.
(28) Bergstrom, E. Effect of natural and artificial diets on seasonal changes in fatty acid composition and total body lipid content of wild and hatchery-reared Atlantic salmon (*Salmo salar* L.). *Aquaculture* **1989**, *82*, 205–217.
(29) Thomas, F.; Jamin, E.; Wietzerbin, K.; Guerin, R.; Lees, M.; Morvan, E.; Billault, I.; Derrien, S.; Rojas, J. M.; Serra, F.; Guillou, C.; Aursand, M.; McEvoy, L.; Prael, A.; Robins, R. J. Determination of origin of Atlantic salmon (*Salmo salar*): The use of multiprobe and multielement isotopic analyses in combination with fatty acid composition to assess wild or farmed origin. *J. Agric. Food Chem.* **2008**, *56*, 989–997.
(30) Aursand, M.; Standal, I. B.; Praël, A.; McEvoy, L.; Irvine, J.; Axelson, D. E. ^{13}C-NMR pattern recognition techniques for the classification of Atlantic salmon (*Salmo salar* L.) according to their wild, farmed and geographical origin. *J. Agric. Food Chem.* **2009**, *57*, 3444–3451.
(31) Martinez, I.; Standal, I. B.; Aursand, M.; Axelson, D. E.; Finstad, B. Identification of the farm origin of salmon by using fatty acid and high resolution ^{13}C nuclear magnetic resonance profiling. *Food Chem.* **2009**, *116*, 766–773.
(32) Pearl, J. *Probabilistic Reasoning for Intelligent Systems*; Morgan Kaufmann: San Francisco, CA, 1988.
(33) Heckerman, D.; Geiger, D.; Chickering, D. Learning Bayesian networks: The combination of knowledge and statistical data. *Mach. Learning* **1995**, *20*, 197–243.
(34) Bligh, E.; Dyer, W. A rapid method of total lipid extraction and purification. *Can. J. Biochem. Physiol.* **1959**, *37*, 911–917.
(35) Domingos, P.; Pazzani, M. On the optimality of the simple Bayesian classifier under zero-one loss. *Mach. Learning* **1997**, *29*, 103–130.
(36) Peng, H.; Long, F.; Ding, C. Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Trans. Pattern Anal. Mach. Intelligence* **2005**, *27*, 1226–1238.
(37) Johnsen, R. I.; Grahl-Nielsen, O.; Roem, A. Relative absorption of fatty acids by Atlantic salmon *Salmo salar* from different diets, as evaluated by multivariate statistics. *Aquaculture Nutr.* **2000**, *6*, 255–261.
(38) Seierstad, S. L. *The Effect on Fish and Human Health of Replacing Marine Oils by Vegetable Oils in Feeds of Atlantic Salmon*. Ph.D. Thesis, Norwegian School of Veterinary Science, Oslo, Norway, **2008**.
(39) Turchini, G. M.; Metasti, T.; Frøyland, L.; Orban, E.; Caprino, F.; Moretti, V. M.; Valfrè, F. Effects of alternative dietary lipid sources on performance, tissue chemical composition, mitochon-drial fatty acid oxidation capabilities and sensory characteristics in brown trout (*Salmo trutta* L.). *Aquaculture* **2003**, *225*, 251–267.
(40) Sargent, J. R.; Tocher, D. R.; Bell, J. G. The Lipids. In *Fish Nutrition*, 3rd ed.; Halver, J. E., Hardy, R. W., Eds.; Elsevier Academic Press: San Diego, CA, 2002; pp 181−257.
(41) Turchini, G. M.; Francis, D. S.; De Silva, S. S. Fatty acid metabolism in the freshwater fish Murray cod (*Maccullochella peelii peelii*) deduced by the whole-body fatty acid balance method. *Comp. Biochem. Physiol.* **2006**, *144*, 110–118.
(42) Martinez, I. Authenticity Assessment Based on Other Principles: Analysis of Lipids, Stable Isotopes and Trace Elements. In *Fishery Products: Quality, Safety and Authenticity*; Oehlenschläger, J., Rehbein, H., Eds.; Blackwell Publishing Ltd.: Cambridge, MA, 2009; pp 388−410.
(43) Megdal, P. A.; Craft, N. A.; Handelman, G. J. A simplified method to distinguish farmed (*Salmo salar*) from wild salmon: Fatty acid ratios versus astaxanthin chiral isomers. *Lipids* **2009**, *44*, 569–576.
(44) Olsen, R. E.; Henderson, R. J.; Sountama, J.; Hemre, G.; Ringo, E.; Melle, W.; Tocher, D. R. Atlantic salmon, *Salmo salar*, utilizes wax ester-rich oil from *Calanus finmarchicus* effectively. *Aquaculture* **2004**, *240*, 433–449.
(45) Torstensen, B. E.; Froyland, L.; Lie, Ø. Replacing dietary fish oil with increasing levels of rapeseed oil and olive oil: Effects on Atlantic salmon (*Salmo salar* L.) tissue and lipoprotein lipid composition and lipogenic enzyme activities. *Aquaculture Nutr.* **2004**, *10*, 175–192.
(46) Torstensen, B. E.; Froyland, L.; Ørnsrud, R.; Lie, Ø. Tailoring of a cardioprotective muscle fatty acid composition of Atlantic salmon (*Salmo salar*) fed vegetable oils. *Food Chem.* **2004**, *87*, 567–580.

Received May 4, 2009. Revised manuscript received July 10, 2009. Accepted July 11, 2009. This work received financial support through the shared-cost RTD project COFAWS (Confirmation of the Origin of Farmed and Wild Salmon and Other Fish, Contract G6RD-CT-2001-00512) funded by the European Community under the 'Competitive and Sustainable Growth' Programme (1998−2002), for which the other partners were Eurofins Scientific (Nantes, France), Université de Nantes (Nantes, France), and Joint Research Centre (Ispra, Italy), and a project financed by the Norwegian Research Council (NFR-project 146932/130).