Food Research International
August 2020, Volume 134, Pages 109254 (11p.)
https://doi.org/10.1016/j.foodres.2020.109254
https://archimer.ifremer.fr/doc/00624/73590/

Use of random forest methodology to link aroma profiles to volatile compounds: application to enzymatic hydrolysis of Atlantic salmon (*Salmo salar*) by-products combined with Maillard reactions

Cardinal Mireille [1, *], Chaussy Marianne [1], Donnay-Moreno Claire [1], Cornet Josiane [1], Rannou Cecile [2], Fillonneau Catherine [2], Prost Carole [2], Baron Regis [1], Courcoux Philippe [3, 4]

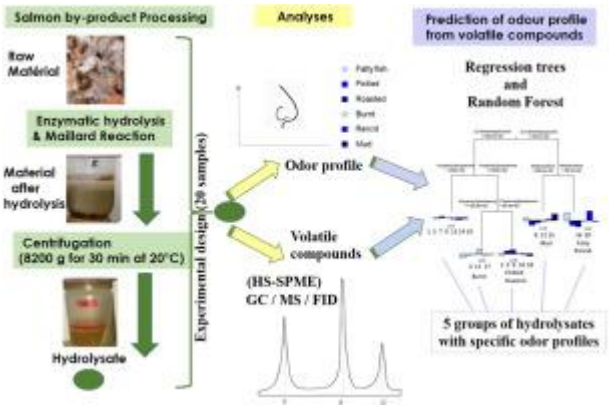
[1] Ifremer, laboratoire EM3B, rue de l'île d'Yeu, 44311 Nantes Cedex, France
[2] Oniris, UMR CNRS 6144 GEPEA, groupe Flaveur, Nantes, France
[3] Oniris, StatSC, rue de la Géraudière, 44322 Nantes, France
[4] INRA USC 1381, 44322 Nantes, France

* Corresponding author : Mireille Cardinal, email address : cardinal@ifremer.fr

Abstract :

To use salmon protein hydrolysates as food ingredients and to mask the fish odor, Maillard reactions were associated with enzymatic production of hydrolysates. The study explored an original approach based on regression trees (RT) and random forest (RF) methodologies to predict hydrolysate odor profiles from volatile compounds. An experimental design with four factors: enzyme/substrate ratio, quantity of xylose, hydrolysis and cooking times was used to create a range of enzymatic hydrolysates. Twenty samples were submitted to a trained panel for sensory descriptions of odor. Hydrolysate volatile compounds were extracted by means of Headspace Solid Phase MicroExtraction (HS-SPME) and analyzed using gas chromatography/mass spectrometry (GC-MS). The results showed that RT and RF methodologies can be useful tools for predicting an entire sensory profile from volatile compounds. Four main volatile compounds made it possible to separate hydrolysates into five groups according to their specific sensory profile. 2,5-dimethylpyrazine, 1-hydroxy-2-propanone and 3-hydroxy-2-pentanone were identified as the main predictors of the roasted odor, whereas methanethiol was associated with a mud odor. These results also suggest the appropriate process conditions for obtaining a typical roasted odor.

Graphical abstract



Highlights

►Regression Trees and Random Forests methodology : a tool to predict a whole sensory profile. ► Four main volatile compounds identified in the final regression tree made it possible to separate hydrolysates into five groups. ► Prediction results may be sensitive to sensory measurements variability. ► Appropriate process conditions combining hydrolysis parameters and Maillard Reaction lead to specific roasted odor.

Keywords : sensory characteristics, volatile compounds, HS-SPME/GC-MS, regression tree, random forest, hydrolysate, Maillard reactions

1. Introduction

Today, using available resources has become a matter of major concern in all the sectors of activity. This is particularly true in the context of the fishing industry, which produces considerable quantities of by-products such as heads, viscera, skin, backbones, cutoffs and blood. The waste may represent 65 % of the initial material in the case of the tuna canning industry and a similar situation can be observed with farmed salmon. Although using major waste as fishmeal (Refstie, Olli, & Standal , 2004; Nguyen, Pérez-Gálvez, & Bergé, 2012) is a

widespread practice, other applications can play a part in reducing this waste while offering higher added value. Applications include recovery of long-chain polyunsaturated fatty acids (de Oliveira et al. 2017), using bioactive compounds that are beneficial for human health (Charoenphun, Youravong, & Cheirsilp, 2013), and developing cosmetic products (Venkatesan, Anil, Kim, &. Shim, 2017).

According to the FAO (2014), the need to gain approval from the regulatory authorities for the specific health claims of nutraceuticals and health supplements may be a serious obstacle to their development and they therefore consider that using the by-products from fish processing directly as food, or indirectly as food by producing feed ingredients, is a more realistic solution. Enzymatic hydrolysis has been studied extensively for over 30 years (Ravallec-Ple, Gilmartin, Van Wormhoudt, & Le Gak, 2001; Halim, Yusof, &. Sarbon, 2016) and appears to be an efficient means of recovering valuable components, such as proteins, from marine biomass (Sathivel *et al*., 2003; Nguyen *et al*. 2011). In addition, developing cost efficient industrial food grade protease has made it possible to produce new kinds of protein hydrolysate for different applications (Aspevik, Egede-Nissen, &. Oterhals, 2016). In the case of fish protein hydrolysates (FPH), while their functional properties and nutritional value have been recognized as good, their use as food ingredients can be limited by the fish flavor that persists even after processing (Sylla, Bergé, Prost, Musabyemariya, & Seydi, 2009).

To reduce or mask the natural fish odor in the products, one of the solutions could be to promote the Maillard reaction (MR) during production of the hydrolysate by adding sugar to the by-product (Kouakou et al., 2014; Zhao, Shen, Guo, Wu, & Dai, 2016). The MR is a complex series of chemical interactions that occurs during the processing between the lysine amino group in peptides or proteins and the carbonyl group of reducing sugars. This reaction leads to a variety of intermediates and brown products such as melanoidins, which play an important role in the aroma, taste and color of processed foods (Machiels & Istasse, 2002). The MR can

add a pleasant flavor to the food through the development of roasted notes and therefore play a role in consumer acceptability. Temperature, time, pH, and water activity are all factors strongly that are involved in MR (Ajandouz, Tchiakpe, Dalle Ore, Benajiba, & Puigserver, 2006), but are also known to influence hydrolysate characteristics (Molla & Hovannisyan, 2011; Prabha, Narikimelli, Infanshia Sajini, & Vincent, 2013). Producing fish hydrolysates with aromatic notes such as a caramelized odor for human food applications is therefore challenging.

The main purpose of this work was to better understand the relationships between volatile compounds and the odor properties of hydrolysates in order to identify the main compounds potentially involved in sensory perceptions. To achieve this aim, an experimental design methodology was used to create a range of samples thanks to variation in four factors: enzyme/substrate ratio, hydrolysis time, quantity of sugar and cooking time. These parameters were chosen as being representative of the main parameters involved in hydrolysis conditions and controlled at the industrial scale. Parameter levels were set according to previous results (Kouakou et al., 2014). After an hydrolysis step associated with Maillard reactions, the hydrolysates were submitted to a panel for sensory description and gas chromatography was used to quantify the volatile compounds. In line with the work carried out by Vigneau, Courcoux & Symoneaux (2018), we assumed that the random forest methodology could be applied to both link an entire sensory profile to volatile compounds, and identify the importance of these compounds in sample sensory characteristics. This study was oriented towards the relationships between volatile compounds and sensory profiles and will not include other results on the chemical characteristics of hydrolysates.

2. Materials and methods

2.1. Raw material and additives

Salmon by-products (backbones from the filleting process) were provided by the company Copalis (Boulogne/Mer, France) from fish processing plants. One hundred and fifty kg of by-products were roughly ground and frozen at -20°C by Copalis and transferred to the laboratory by refrigerated transportation. On arrival, the raw material was divided into four kg samples and stored in plastic bags at -20°C until hydrolysis processing.

The enzymes used for the hydrolysis was provided in liquid form by Novozymes AS (Bagsvaerd, Denmark). Novozym® F.M.2.4 L (EC number: 3.4.21.62) is a bacterial serine endopeptidase (subtilisin) prepared from a strain of *Bacillus lichenformis*. This enzyme was developed to hydrolyze food proteins. It also satisfies the purity requirements for food-grade enzymes, as set by the Joint FAO/WHO Expert Committee on Food Additives (JECFA) and the Food Chemicals Codex (FCC). The optimal working conditions for Novozym® F.M.2.4 L are reported to be a pH between 7 and 9 and a temperature between 30 and 65°C. Novozym® F.M.2.4 L has a declared activity of 2.4 Anson Units (AU) g^{-1} (Novo Nordisk AS). To protect against oxidation of the hydrolysates, a commercial mixture of natural antioxidants, Naturox (tocopherols and rosemary) from the company Jan Dekker International, was used at a level of 250 mg per kg of raw material. After preliminary experiments (Kouakou, 2012), this level was chosen as the minimum content needed to significantly reduce lipid oxidation in hydrolysates. Of all the sugars available for promoting flavor generation (Ames, Guy, & Kipping, 2001), D-xylose was chosen for the good reactivity of pentose, as well as for economic reasons. Xylose was provided by Danisco (Denmark).

The standards used to identify the volatile compounds were purchased on Sigma-Aldrich. The following purity was specified: pentane (≥99.0%), hexane (≥97.0%), dodecane (99.0%), methylbenzene (99.9%), ethylbenzene (≥99.0%), styrene (≥99.5%), benzaldehyde (≥99.0%), acetaldehyde (≥99.5%), propanal (97%), 2-methylpropanal (99.0%), butanal (≥99.0%), 2-methylbutanal (95.0%), 3-methylbutanal (97.0%), hexanal (98.0%), heptanal (95.0%), ethanol

(≥99.8%), 1-propanol (≥99.9%), 1-penten-3-ol (99.0%), (E)-2-penten-1-ol (95.0%), (Z)-2-penten-1-ol (95.0%), 2-butanone (99.0%), 2,3-butanedione (97.0%), 1-hydroxy-2-propanone (90.0%), ethyl acetate (99.5%), acetic acid (≥99.9%), 3-methylbutanoic acid (99.0%), 2-methylfuran (99.0%), furfural (99.0%), 2-methylpyrazine (99.0%), 2,5-dimethylpyrazine (≥99.9%), dimethyl disulfide (98.0%), methional (≥97.0%), 3-methyl-1-butanol (≥99.9%), γ-butyrolactone (99.0%), 2-acetylthiazole (99.0%) and 2-furanmethanol (99.0%). Two other standards were used: 2-methyl-1-propanol (Merck, 99.0%) and 2-propanone (Riedel de Haën, ≥99.9%).

2.2. Experimental design

Four processing variables were investigated using the response surface methodology (RSM) and a randomized three level-four factor Composite Draper-Lin design (Statgraphics Centurion XV.II, Statpoint, Herndon, USA). The three levels chosen for the selected factors were enzyme/substrate ratio (E/S) (0.1, 0.25, and 0.4 %) (w/w), hydrolysis time (HT) (10, 50, 90 min), sugar (xylose) concentration (X) (2, 6, 10 $g.kg^{-1}$) and cooking time (CT) (30, 60, 90 min). A total of nineteen experiments was required. Results from a previous study (Kouakou et al., 2014) have shown that adding 10g of sugar to 1 kg of by-product was enough to develop roasted notes during enzymatic hydrolysis. This level was thus set as the high level in the experimental design in order to limit any possible residual sugar in the hydrolysate. A sample produced without added sugar and in hydrolysis conditions set at the highest level for each factor (E/S, 0.4; hydrolysis time, 90 min; cooking time, 90 min) was introduced as a supplementary sample to illustrate a non-Maillard reaction sample (Table 1). Each sample was the result of one production. The central point of the experimental design was repeated three times (samples 5, 10, and 17) in order to test the repeatability of the productions.

2.3 Enzymatic hydrolysis

Frozen minced by-products were thawed at 4°C for 15 hours. For each design experiment, 4 kg of salmon by-products were ground with antioxidant at a knife rotation speed of 1000 rpm, in a Roboqbo Qb8-3 reactor (capacity of 8 liters) (Bentivoglio, Italy). The reactor had a double jacket to make thermal exchanges (heating or cooling) possible, and thus reached the optimal temperature of 40°C for the enzyme within 5 minutes. Hydrolysis was started at a speed of 300 rpm by adding the enzyme. pH was not controlled in order to stay close to industrial conditions. Once the hydrolysis time had elapsed, xylose was added just before stopping the hydrolysis reaction by heating the product to 95°C for 30 to 90 min, depending on the experimental design (cooking time factor). The choice of cooking conditions (time and temperature) was defined in order to favor the Maillard reactions while at the same time providing sufficient inactivation time for enzyme activity in agreement with regulatory obligation. Once this step had been completed, the temperature of the reactor was adjusted down to a temperature of 40°C and the hydrolyzed product was removed through a sieve to eliminate the bones. This product was then centrifuged at 8200 g for 30 min at 20°C in a Beckman coulter to separate and collect the aqueous fraction. In this paper, the word hydrolysate will refer to this fraction. All the samples were stored at -80°C for further sensory and biochemical analyses.

2.4. Sensory evaluation

The sensory analysis was carried out with sixteen panelists (12 females, 4 males, between 32 and 65 years old) from an internal panel at IFREMER. They already had experience in salmon hydrolysate evaluation and had received training in the quantification of descriptors for 1h twice

a week over a three-month period (Cardinal, Baron, Kouakou, Prost, & Courcoux, 2014), but received further training before starting this experiment. During preliminary screening on process parameters, the following steps were proposed: - a sorting task on odor perception with 21 hydrolysates, - a discussion session with the whole panel in order to find a consensus on the main discriminative odors; this discussion was based on the results of the sorting as well as the list of descriptors previously used, -a scoring session where panelists were invited to test 6 samples illustrating the main characteristics of the hydrolysates in order to share a consensual intensity level for each attribute, - two profiling sessions to check the panel's discriminative power and the agreement between panelists and the whole panel. From the initial twenty-one panelists, sixteen were selected for their ability to recognize the selected odors, and for the good correlation between their individual scores and mean panel sensory scores. They were invited then to carry out a quantitative descriptive analysis (Stone & Sidel, 2004) on the sensory characteristics of salmon hydrolysates from the experimental design. The hydrolysates were presented in plastic flasks wrapped in aluminum foil in accordance with the conditions described by Kouakou et al. (2014). Using a continuous scale from 0 to 10, the panelists had to score the six following odor descriptors: fatty fish, pickled (like pickled anchovies), roasted, burnt, rancid and mud (sulfur notes). Twenty samples were scored in two sessions. Sample presentation was balanced according to factor levels in order to have the range of variation for each processing factor within each session. The tests were performed in individual booths equipped with computers using data acquisition software (Fizz, Biosystems, Couternon, France) under white lighting and at ambient temperature (20°C).

2.5. Volatile compounds

The procedure for analyzing the volatile compounds was adapted from Kouakou et al. (2014).

2.5.1. Extraction of the volatile compounds by Headspace Solid Phase MicroExtraction (HS-SPME).

Five ml of hydrolysate were placed in a 20-mL glass vial closed with a screw top and equipped with a Teflon septum. The sample was equilibrated for 60 min at 40°C. The extraction of the volatile compounds was performed using a Carboxen/PDMS fibre (85 µm, 1 cm, Carboxen/PDMS StableFlex, Supelco, Sigma-Aldrich Chimie, Lyon, France) for 15 min at 40°C. Analyses were performed in triplicate on each hydrolysate.

2.5.2. Gas chromatography / Mass spectrometry / FID

The apparatus used was a gas chromatograph (Agilent 7890A, Wilmington, DE, USA) equipped with a flame ionization detector (FID) and coupled to a mass spectrometer (electronic impact source, Agilent 5975CNetwork, Wilmington, DE, USA). The inlet temperature was 260°C, the FID detector temperature 250°C and the MS detector temperature 280°C. The carrier gas was helium and the pressure was 150 kPa. The splitless mode was used for the injection, and the desorption time was 7 min. The capillary column was a DB-WAX (30 m, 0.25 mm, 0.5 µm, J&W Scientific, Folsom, CA). The program used was 40°C for 10 min, ramped up to 240°C at 7°C/min then equilibrium at 240°C for 3 min. Effluent from the end of the GC was split 1/1 between the MS and FID. Peaks were integrated with MSD Chemstation software (Agilent Technologies). Mass spectra were recorded in electron impact mode (70 eV) between 33 and 300 m/z mass range at a scan rate of 2.7 scan.s^{-1}.

The volatile compounds were identified according to 3 criteria: comparison with the literature of their Kovats retention index, comparison of their mass spectra with those of the Wiley 6 library, and comparison of their retention index with those of the corresponding standards when the standard was available. The semi-quantified results were obtained from the FID chromatogram and expressed as a peak area. The results obtained are only semi-quantitative in order to compare the samples, but do not reflect the exact quantity of each volatile compound

present in the hydrolysate. Analyses were performed in triplicate on each hydrolysate which means that for each volatile compound and each hydrolysate, the mean relative peak area is obtained from 3 values.

2.6. Statistical analysis

A standardized Principal Component Analysis (PCA) was performed on the mean of the panel score for each product and each sensory descriptor to highlight the main odor characteristics of the products. The link between volatile organic compounds and sensory perception of the products was investigated using regression trees and random forest methodologies. Regression trees (RT) belong to recursive partitioning techniques and their aim is to predict a quantitative response from a set of quantitative predictors (Breiman, Friedman, Olshen, & Stone, 1984). In our case, the response was a sensory attribute; the panel mean score and predictors were the volatile compounds. A regression tree can be considered as a set of decision rules created by recursively splitting the whole set of products into subsets by maximizing the homogeneity of the two resulting nodes. Random forests (RF) were introduced by Breiman (2001) and consist in a large number of regression trees, randomly generated by resampling the training dataset in order to improve the predictive accuracy of individual trees. Random forests make it possible to compute the Variable Importance measure (VI) which quantifies the role played by each variable in predicting the response. The confidence intervals of these importance measures were obtained by repeating the RF on the same learning set. This technique is a simple tool for selecting predictors with a significant effect on the response. The regression tree based on this selection of compounds can be considered to be more robust than the one built on the complete set of predictors.

One of the main features of the random forest methodology is the robustness of the predictions, obtained thanks to the construction principle of the forests: bagging (bootstrap aggregating).

Each decision tree for a random forest is created from the training set by a doubly randomized process: the bootstrapping of the individuals (random resampling with replacement of products in our case) and the random selection of variables at each node of the trees (at each node, the best volatile compound is chosen among a third of all the compounds). One single decision tree has a tendency to overfit and the bagging process leads to an improvement in the predictive performance. The samples that are not selected for a given tree (the Out-of-Bag or OOB samples) may be used as a validation step or the solution. The computation of the Variable Importance is based on the mean decrease in accuracy among all trees for the Out-of-Bag samples when the values of the given variable are randomly permuted. Out-of-Bag samples play the role of validation set without having to divide the data-set into calibration and validation sets. In addition, the length of a decision tree (the number of leaves) is obtained by minimizing the error of prediction generally obtained by LOO (leave one out) cross-validation step.

This type of machine learning techniques has recently been used in many fields, including sensory studies (Gomez-Meire, Campos, Falqué, Díaz, & Fdez-Riverola , 2014; Brillante et al., 2015, Vigneau et al. 2018), demonstrating its accuracy and robustness even in the case of non-linear relationships, interactions between predictors or high correlations among a set of predictors. In addition, regression trees may be considered as a technique for supervised clustering, providing decision rules and giving a simple interpretation of the link between response and predictors.

As the sensory profile of products is composed of several sensory attributes, we considered a multivariate generalization of the RT and RF methodologies. Introduced by De'ath (2002) in the field of ecology, multivariate regression trees and random forests have been developed for predicting a multivariate response. In this case, the splitting rule was based on the minimization

of the inertia in the child nodes. In our study, each node in the multivariate regression tree was described by means of the sensory profile of the individuals belonging to this node.

Multivariate regression trees and random forests were carried out using language R 3.5.1 (R Core Team, 2018) and the R packages mvpart (De'ath, 2014) and randomForestSRC (Ishwaran & Kogalur , 2019).

3. Results and discussion

3.1. Odor characteristics of the hydrolysates

The first plane of the principal component analysis (PCA) with standardization performed on the means of the sensory scores of each hydrolysate and each descriptor, accounted for 73.2% of the total variance (Fig. 1a). The first axis (54.9% of total variance) was mainly created by the roasted, pickled, rancid and fat criteria (Fig. 1b) and made a clear separation possible between one group of samples associated with a roasted and pickled odor and three samples: 15, 20 and 16. These samples were distributed according to their main odor characteristic, mud for sample 15, fat fish and rancid for numbers 16 and 20. The medium position of samples 8 and 12 on this first axis reflected intermediate sensory characteristics. The second axis (18.3% of total variance) added specific information through the 'burnt' descriptor that particularly differentiated samples 4 and 11. The three replicated samples, 5, 10 and 17, presented similar profiles and were close on this sensory map; a clustering analysis performed on the principal components of PCA confirmed that these samples were grouped in the same class of products (not shown).

A first general approach suggested that the sample separation could not be explained only by the level of sugar added to the by-product, but also by the specific process conditions associated. While most hydrolysates on the right side of the sensory map were produced with the lowest level of sugar or without sugar, an exception can be seen with sample 16. In this case, all the

factor conditions, including the sugar factor, were set at the high level except the cooking time (CT) set at the lowest level. Sample 20, the only sample with no sugar added, presented similar characteristics to sample 16, the highest scores for fatty fish and rancid odors. Although hydrolysis conditions, such as a long hydrolysis time associated with a high enzyme/substrate ratio, seemed favorable for producing small peptides and therefore for making reactions possible between amino groups in peptides or proteins and reducing sugar, the results showed that the hydrolysates were mainly characterized by odors illustrating an oxidation reaction. The absence of sugar (sample 20) or a too short cooking time (sample 16) could explain these results. Sample 15, characterized by a mud odor, was processed at the lowest level for each of the four factors. Samples 8 and 12 had similar characteristics but at a lower intensity than sample 15. The same level of xylose (2 $g.kg^{-1}$) in the three samples could suggest either the need to add a sufficient quantity of sugar in the reaction mixture to favor a roasted aroma, and/or the importance of combining other factors such as E/S, HT and CT at a required level for each of them to prevent or mask the formation of sulfur notes (Farmer, Mottram & Whitfield, 1989). It was likely that these hydrolysis conditions were not conducive to developing Maillard reactions and their related aroma. In the case of the two samples separated on the second axis, samples 4 and 11, the only common processing condition for these two samples produced with medium or high levels of sugar was the low level of E/S (0.1%). This low level could result in lower enzyme activity and therefore a lower production of peptides with different sizes. Li, Zhong, Yokoyama, Shoemaker, Zhu, & Xia (2013) mentioned in their study that rice protein hydrolysates with a higher degree of hydrolysis were found to have more pyrazines such as 2,5-dimethyl-pyrazine or methyl-pyrazine. The formation of these compounds from α-amino acids, along with reducing sugars such as xylose could therefore be reduced as the hydrolysis conditions were not favorable for small peptide production.

3.2. Volatile compounds in the hydrolysates

- Identification of volatile compounds

A total of 44 volatile compounds was identified in the hydrolysates (Table 2). The chemical compounds belonged to various chemical classes such as aldehydes (7), ketones (7), alcohols (6), benzene compounds (4), alkanes (3), sulfur compounds (3) and others (14). Most of the compounds were identified in the 20 hydrolysates, with variation only in their quantity (Table 3).

Carbonyl compounds, **aldehydes** and **ketones** were the most abundant volatile compounds in the hydrolysates. Aldehydes are generated via two main formation pathways: lipid oxidation and Maillard reaction. Aliphatic aldehydes, such as hexanal, heptanal or nonanal, are mainly derived from the lipid oxidation occurring in fish flesh (Varlet, Prost, & Sérot, 2007).. The second pathway for producing aldehydes is through Strecker degradation, which occurs during the Maillard reaction (Varlet et al., 2007; Xu et al., 2018).. Aldehydes are one of the most important odor-active compounds because of their low odor threshold values (Peinado, Koutsidis & Ames, 2016a). They may produce desirable aromas (roasty, malty, cocoa, nutty) and undesirable aromas (green, rancid, oxidized) (Giri, Osako, Okamoto, & Ohshima, 2010). Like aldehydes, ketones can be formed through lipid oxidation and the Maillard reaction (Peinado, Miles, & Koutsidis, 2016b). Most of the ketones identified are associated with buttery or creamy aromas on the one hand or ethereal, solvent aromas on the other.

Alcohols were the second most abundant compounds. Alcohols can be formed by secondary decomposition of the hydroperoxides in fatty acids, or by enzymatic peroxidation of the n-3 and n-6 polyunsaturated fatty acids present in fish flesh (Peinado et al., 2016b). Alcohols have various odor thresholds, meaning that they contribute in different ways to the overall aroma.

Alcohols are associated with alcoholic and green odors. The amount of 1-penten-3-ol seems to be related to the amount of oil in the product (Peinado et al., 2016a, 2016b).

Benzene compounds. Benzene compounds are not significant potent odorants. Only benzaldehyde has a relatively low odor threshold (350-3500ppb in water, (Leffingwell, 2019). Most probably, benzaldehyde could be produced through the Maillard reaction, but it could also be generated by oxidation or photochemical degradation of toluene, or other hydrocarbons (Varlet et al., 2007).

Sulfur compounds, such as dimethyl disulfide and dimethyl trisulfide, are generally associated with a deterioration of the material because of their strong unpleasant odor and low detection threshold (Peinado et al., 2016a). These compounds may originate in the raw material or be generated during the fermentation process from the free, peptidic and proteinic sulfur amino acids in fish flesh (Peinado et al., 2016a).

Furans and pyrazines are generated through the Maillard reaction. Their odor is associated with empyreumatic aromas such as toasty, cocoa, nutty, chocolate and caramel. These compounds are formed mainly when hydrolysates are heated.

- Semi-quantification of the volatile compounds

Quantitatively, significant amounts of carbonyl compounds (aldehydes and ketones) and alcohols were present (expressed in relative peak area/g of product). Carbonyl compounds are generally odor-active compounds contributing to the overall odor of the food product (Varlet et al., 2007). Alcohols have slightly lower odor thresholds than carbonyl compounds, depending on their nature and quantity. In comparison, furans, pyrazines and sulfur compounds are present in relatively low quantities, but generally have low odor thresholds. These compounds are thus particularly important for the overall aroma of the product.

The hydrolysate containing the highest quantity of volatile compounds was sample 18. This hydrolysate was obtained by applying the highest level of xylose concentration, hydrolysis time

and cooking time. These parameters seemed to have a particular impact on the production of volatile compounds, especially those generated during the Maillard reaction. Sample 18 contained the highest quantity of furans and pyrazines. In the literature, these compounds are known to be odor–active and responsible for roasted and burnt odors. On the contrary, sample 15 was the one with the lowest total quantity of volatile compounds. This sample was obtained with the lowest level of all the parameters involving the production of few volatile compounds. A direct relationship between the nature and quantity of volatile compounds produced, and the process parameters applied was observed.

In more detail, the most represented volatile compounds in all the hydrolysates were 3-methylbutanal, ethanol, 2-propanone + 2-methylpropanal and 1-penten-3-ol. 3-methylbutanal is associated with malty, ethereal, aldehydic, chocolate and fatty odors (www.thegoodscentscompany.com). 2-propanone + 2-methylpropanal are described respectively as ethereal, solvent, apple and aldehydic, floral, and green (www.thegoodscentscompany.com). Considering the two alcohols, 1-penten-3-ol is described as green, vegetable, tropical and fruity whereas ethanol is perceived as alcoholic, ethereal and medical (www.thegoodscentscompany.com).

3.3. Predicting sensory characteristics from volatile compounds

The importance of volatile compounds as predictors of the main odor characteristics of enzymatic hydrolysates is presented in Fig.2. The importance measure quantifies the contribution of each volatile compound to the prediction of the sensory profile. The confidence interval for each importance value was obtained by repeating 50 random forests. A compound was therefore significantly more important if the lower limit of its confidence interval was greater than zero. Of all the volatile compounds identified, eleven contributed significantly to the sensory profile prediction: methanethiol, 2,5-dimethylpyrazine, 1-hydroxy-2-propanone,

propanone, 2-methyl-1-propanol, furfural, 2-methylfuran, 2,3-pentanedione, hexanal, dodecane and 3-hydroxy-2-pentanone. The odor description of these compounds ranged from cabbage and garlic for methanethiol, to green, herbal and fatty for hexanal and included roasted, caramel, butter, wood, truffle or ethereal notes for the other compounds. These eleven compounds were selected to build the optimal regression tree for predicting hydrolysate sensory profiles (Fig. 3). A regression tree is built by recursively splitting the set of products into two groups by choosing, at each node, the most discriminant predictor (a volatile compound) and the appropriate threshold. This technique leads to a supervised clustering of the whole set of products. Therefore, the optimal tree is the best clustering of samples for predicting the sensory profile from the volatile composition.

Specific odors produced during Maillard reactions, and especially roasted odors, have been identified as potentially interesting notes for food applications. The first compound which played a part in splitting the initial 20 samples into 2 groups was 2,5- dimethylpyrazine at a threshold value of $56x10^3$ peak area/g of product. Five hydrolysates with a 2,5-dimethylpyrazine value below this threshold were grouped together. A mud odor was the characteristic for three of them when methanethiol level was higher than $29.9x10^3$ peak area/g of product, and the two samples left had fat and rancid notes for a level of methanethiol below this threshold. Methanethiol was not identified among the highly abundant volatile compounds, but was selected in the random forest procedure as a discriminative compound for sensory prediction. The low odor threshold (0.02 ppb) of this compound originated from the breakdown of sulfur-containing amino acids such as cysteine or methionine (Varlet & Fernandez, 2010), which could explain its importance on the sensory characteristics of the hydrolysates. 2,5-dimethylpyrazine was described as cocoa, roasted nuts, roast beef, woody, grass, medical. This compound was known to be produced through the Maillard reaction. Its odor threshold is

relatively high (800-1 800 ppb). Both these compounds were identified as odor-active compounds possibly impacting the roasted odor of a food product.

The fifteen remaining samples, with a level of 2, 5- dimethylpyrazine higher than the $56.2x10^3$ peak area/g of product, were first separated according to the level of the compound 1-hydroxy-2-propanone. One sub-set of seven samples with no specific characteristics was identified when the level of this compound was less than $494x10^3$ peak area/g of product, and a group of eight samples when the level was greater. This latter group was finally divided into two sub-sets depending on their 3-hydroxy-2-pentanone content. A group of three samples, with a burnt odor, appeared when the level of this compound was higher than $60x10^3$ peak area/g of product. When the level of 3-hydroxy-2-pentanone was below $60x10^3$, the five samples left presented specific roasted and pickled notes.

Three compounds: 2,5-dimethylpyrazine, 1-hydroxy-2-propanone and 3-hydroxy-2-pentanone were identified as playing a part in empyreumatic aromas. Considering the formation pathway of 2,5-dimethylpyrazine, as well as its odor description, it is hardly surprising that a higher amount of this compound will enhance the Maillard notes. But the relative ratio between the three compounds may have an influence on the nature of the sensory characteristics, either roasted, burnt or neutral. The main groups of products identified through the regression tree were in line with previous sensory results with a few slight variations. The two groups with specific notes, either mud or fatty and rancid were clearly separate from the others. The odor activity of the volatile compounds selected in the regression tree was confirmed in a second step through olfactometry measurements. Regarding the three replicated samples (5, 10 and 17), they were distributed into three different groups. All these groups had a common threshold for 2,5-dimethylpyrazine, greater than $56.2x10^3$ and only small level differences on 1-hydroxy-2-propanone and 3-hydroxy-2-pentanone were detected. It is therefore likely that the variability in sensory measurements, and especially the pickle odor, could explain this result.

To mask potential fishy odors through the production of roasted notes, the results suggest finding processing conditions that make it possible to combine the presence of 2,5-dimethylpyrazine and 1-hydroxy-2-propanone, while limiting the level of 3-hydroxy-2-pentanone to avoid the burnt characteristic. Sensory results have shown that perception of roasted notes increased with the cooking time and sugar level, thus confirming that the Maillard reaction setting was driven by sugar content and a sufficient period at high temperature. However, controlling these factors did not seem to be enough. The low level of the E/S ratio (0.1), combined with a too short hydrolysis time, could lead to burnt or mud odors, depending on the cooking time used rather than a roasted odor, even when there was a high sugar content. A low E/S ratio or a short hydrolysis time may affect the hydrolysis reaction by reducing the number of peptide bonds broken and by therefore reducing the potential generation of certain Maillard reaction compounds. We can suppose that the cooking time used can then control the nature of the compounds formed, either for caramelization products with a long cooking time, or sulfur compounds with a short cooking time.

Moreover, for further application of these results, a complementary study will be needed to investigate taste perception and the possible effects on bitterness or other characteristics of process parameters such as a long heating time at 95°C.

Conclusion

This study based on experimental design methodology confirmed previous results on the advantages of coupling Maillard reactions and enzymatic hydrolysis as a way of producing hydrolysates with a range of aromatic properties making it possible to mask initial fish odors. Results suggest some appropriate process conditions such as level of sugar, E/S ratio combined with hydrolysis time for obtaining a typical roasted note. One of the main conclusions of the study concerns the use of RT and RF methodologies to predict, for one of a first times, a whole

odor profile from volatile compounds. The results show that four main volatile compounds contribute to separate hydrolysates into five groups according to their specific sensory characteristics. Three of them, 2,5-dimethylpyrazine, 1-hydroxy-2-propanone and 3-hydroxy-2-pentanone are mainly involved in the perception of roasted notes while methanethiol is associated with a mud odor. The distribution of the three replicates in different sensory groups in the final regression tree probably reflects higher variability in sensory measurements compared to instrumental analysis, and reminds us of the importance of the choice of sensory descriptors used in profiling. In order to consolidate the results obtained, it may be necessary to add to the RF analysis replicated samples obtained from the same production batch, as well as new samples produced from salmon by-products of other origin (plant, country), or samples hydrolyzed with different enzymes that have an influence on the volatile compounds of the hydrolysates. However, once these considerations have been integrated, the results obtained in this study, which follow up on the works of Vigneau et al. (2018), suggest that a multivariate version of regression trees and random forest methodologies may be a useful tool in practice for establishing the main relationships between sensory perception and major volatile compounds.

Acknowledgements

The authors are grateful for the financial support from FUI 18 (Fonds Unique Interministériel). We would also like to warmly thank all the panelists from the trained sensory panel at Ifremer.

References

Ames, J. M., Guy, R. C. E., & Kipping, G. J. (2001). Effect of pH, temperature, and moisture on the formation of volatile compounds in glycine/glucose model systems. J. Agri. Food Chem. 49 (9): 4315–4323

Ajandouz, E.H., Tchiakpe, l.S., Dalle Ore, F., Benajiba, A., & Puigserver, A. (2006) Effects of pH on Caramelization and Maillard Reaction Kinetics in Fructose-Lysine Model Systems. *Journal of Food Science*, *66 (7)*, 926-931.

Aspevik, T., Egede-Nissen, H., & Oterhals, A. (2016). A Systematic Approach to the Comparison of Cost Efficiency of Endopeptidases for the Hydrolysis of Atlantic Salmon (Salmo salar) By-Products. *Food Technology and Biotechnology*, 54(4), 421-431.

Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). Classification and Regression Trees. New-York, Chapman & Hall.

Breiman, L. (2001). Random Forests. *Machine Learning*, *45*, 5-32.

Brillante, L., Gaiotti, F., Lovat, L., Vincenzi, S., Giacosa, S., Torchio, F., Segade, S. R., Rolle, L., & Tomasi, D. (2015). Investigating the use of gradient boosting machine, random forest and their ensemble to predict skin flavonoid content from berry physical-mecahnical characteristics in wine grapes. *Computers and Electronics in Agriculture, 117*, 186-193.

Cardinal, M., Baron, R., Kouakou, C., Prost, C., & Courcoux, P. (2014). Comparative value of a sorting procedure and quantitative descriptive analysis to investigate the influence of processing parameters: case study of hydrolysate production from salmon by-products. *Journal of Sensory Studies, 29(2)*, 159-170.

Charoenphun, N., Youravong, W. & Cheirsilp, B. (2013). Determination of reaction kinetics of hydrolysis of tilapia (*Oreochromis niloticus*) protein for manipulating production of bioactive peptides with antioxidant activity, angiotensin-I-converting enzyme inhibitory activity and Ca-binding properties. *International Journal of Food Science & Technology***,** 48(2), 419-428.

De'ath G. (2002). Multivariate regression trees: a new technique for modelling species-environment relationships. *Ecology*, *83*(4), 1105-1117.

De'ath, G. (2014). mvpart: Multivariate partitioning. Version 1. 6-2. R package.

FAO (2014). The state of World Fisheries and Aquaculture. Opportunities and Challenges**,** 169-172.

Farmer, L. J., Mottram, D. S., & Whitfield, F. B. (1989). Volatile compounds produced in maillard reactions involving cysteine, ribose and phospholipid. *Journal of the Science of Food and Agriculture*, *49(3)*, 347-368.

Giri, A., Osako, K., Okamoto, A. & Ohshima, T. (2010) Olfactometric characterization of aroma active compounds in fermented fish paste in comparison with fish sauce, fermented soy paste and sauce products. *Food Research International*, *43*, 1027-1040

Gómez-Meire, S., Campos, C., Falqué, E., Díaz, F. & Fdez-Riverola, F. (2014). Assuring the authenticity of northwest Spain white wine varieties using machine learning techniques. *Food Research International*, *60*, 230-240.

Halim, N. R. A., Yusof, H. M., & Sarbon, N. M. (2016). Functional and bioactive properties of fish protein hydolysates and peptides: A comprehensive review. *Trends in Food Science & Technology***,** 51, 24-33.

Ishwaran, H. & Kogalur, U.B. (2019). Random Forests for Survival, Regression, and Classification (RF-SRC), Version 2. 8-0. CRAN R package.

Kouakou, C. (2012). Etude du potentiel aromatique d'hydrolysats marins : Application aux co-produits de saumon (Thèse de doctorat). Nantes (France) : Université de Nantes – Angers - Le Mans.

Kouakou, C., Bergé, J. P., Baron, R., Lethuaut, L., Prost, C., & Cardinal, M. (2014). Odor Modification in Salmon Hydrolysates Using the Maillard Reaction. *Journal of Aquatic Food Product Technology***,** *23(5),* 453-467.

Leffingwell (2019). http://www.leffingwell.com/odorthre.htm

Li, Y., Zhong, F., Ji, W., Yokoyama, W., Shoemaker, C. F., Zhu, S., & Xia, W.S. (2013). Functional properties of Maillard reaction products of rice protein hydrolysates with mono-, oligo- and polysaccharides. *Food Hydrocolloids. 30*(1), pp.53-60

Machiels, D., & Istasse, L. (2002). Maillard reaction: importance and applications in food chemistry. *Annales de Médecine Vétérinaire*, *146 (6)*, 347–352.

Molla, A.E., & Hovannisyan, H. G. (2011) Optimization of enzymatic hydrolysis of visceral waste proteins of beluga *Huso huso* using Protamex. *International Aquatic Research*, *3,* 93-99.

Nguyen, H. T. M., Sylla, K. S. B., Randriamahatody, Z., Donnay-Moreno, C., Moreau, J., Luyen, T. T., & Berge, J-P. (2011). Enzymatic Hydrolysis of Yellowfin Tuna (Thunnus albacares) By-Products Using Protamex Protease. *Food Technology and Biotechnology*, *49(1)*, 48-55.

Nguyen, H. T. M., Pérez-Gálvez, R., Bergé, J-P. (2012). Effect of diets containing tuna head hydrolysates on the survival and growth of shrimp *Penaeus vannamei*. *Aquaculture, (324-325)*, 127-134.

de Oliveira, D., Licodiedoff, S. Furigo, A. Ninow, J. L. Bork, J. A. Podesta, R. Block J. M. & Waszczynskyj, N. (2017). Enzymatic extraction of oil from yellowfin tuna (*Thunnus albacares*) by-products: a comparison with other extraction methods. *International Journal of Food Science and Technology,* 52(3), 699-705.

Peinado, I., Koutsidis G. & Ames, J. (2016a). Production of seafood flavour formulations from enzymatic hydrolysates of fish by-products. *Lwt-Food Science and Technology*, *66*, 444-452.

Peinado, I., Miles, W. and Koutsidis, G. (2016b) Odour characteristics of seafood flavor formulations produced with fish by-products incorporating EPA, DHA and fish oil. *Food Chemistry*, *212*, 612-619.

Prabha, J, Narikimelli, A., Infanshia Sajini., M., & Vincent, S. (2013) Optimization for autolysis assisted production of fish protein hydrolysate from underutilized fish *Pellona ditchela*. *International Journal of Scientific & Engineering Research, 4,(12),* 1863-1869.

Ravallec-Ple, R., Gilmartin, L., Van Wormhoudt, A., & Le Gak, Y. (2001, July). Influence of the experimental conditions on the hydrolysis process in fish hydrolysates. Conference at 9th European Congress on Biotechnology (ECB9), Brussels, Belgium. *Engineering and Manufacturing for Biotechnology*, 4, 51-58.

Refstie, S., Olli, J. J., & Standal, H. (2004). Feed intake, growth and protein utilization by post-smolt Alantic salmon (Salmo salar) in response to graded levels of fish protein hydrolysate in the diet. *Aquaculture, 239*, 331–349.

Sathivel, S., Bechtel, P. J., Babbitt, J., Smiley, S., Crapo, C., Reppond, K. D., & Prinyawiwatkul, W. (2003). Biochemical and functional properties of herring (*Clupea harengus*) by product hydrolysates. *Journal of Food Science*, 68(7), 2196-2200.

Stone, H., & Sidel, J.L. (2004). Sensory Evaluation Practices, 3rd edition. Elsevier Academic Press, Amsterdam.

Sylla, K. S. B., Bergé, J.P., Prost, C., Musabyemariya, B., & Seydi, Mg (2009). Sensory and aromatic characteristic of tongue sole by products hydrolysates (*Cynoglossus senegalensis*). *Microbiologie et Hygiène Alimentaire, 21(60)*, 35-43.

Varlet, V., Prost, C., & Sérot, T. (2007) Volatile aldehydes in smoked fish: Analysis methods, occurrence and mechanisms of formation. *Food Chemistry*, *105*, 1536-1556.

Varlet, V., & Fernandez, X. (2010). Review. Sulfur-containing volatile compounds in seafood: Occurrence, odorant properties and mechanisms of formation. *Food Science and Technology International, 16*(6), 463–503.

Venkatesan, J., Anil, S., Kim, S. K., & Shim, M. S. (2017). Marine Fish Proteins and Peptides for Cosmeceuticals: A Review. *Marine Drugs*, *15*(5).

Vigneau, E., Courcoux, P., Symoneaux, R. (2018) Random forests: A machine learning methodology to highlight the volatile organic compounds involved in olfactory perception. *Food Quality and Preference*, 68, 135-145.

Xu, Y., Li, L., Mac Regenstein, J., Gao, P., Zang, J., Xia, W. & Jiang, Q. (2018) The contribution of autochthonous microflora on free fatty acids release and flavor development in low-salt fermented fish. *Food Chemistry*, *256*, 259-267.

Zhao, Q. L., Shen, Q., Guo, R., Wu, J. J., & Dai, Z. Y. (2016). Characterization of Flavor Properties from Fish (*Collichthys niveatus*) Through Enzymatic Hydrolysis and the Maillard Reaction. *Journal of Aquatic Food Product Technology*, *25(4),* 482-495.

Figure Caption

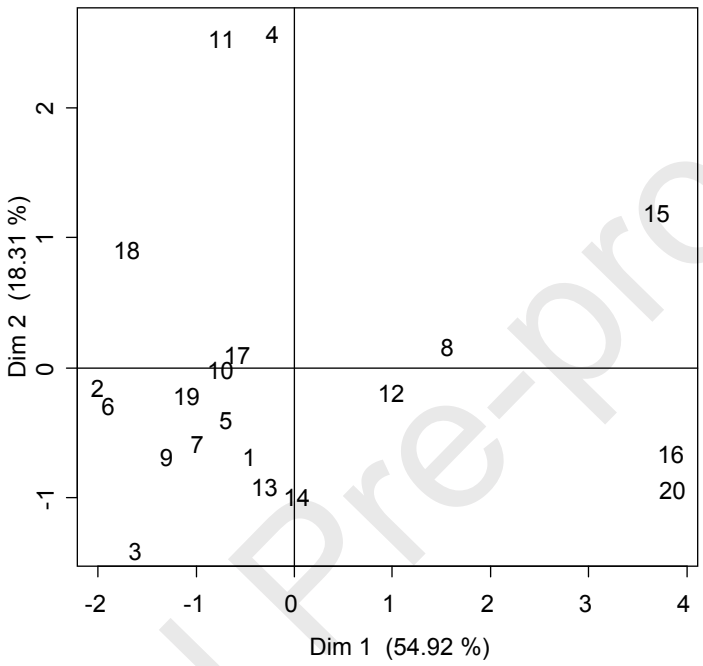
Fig 1. (a) Representation of salmon hydrolysates on the first two dimensions of Principal Component Analysis (PCA) from profiling data.
(b) Projection of sensory descriptors in the first plane of PCA

Fig.2 Variable importance of the 44 volatile compounds in sensory descriptors of odor. Confidence intervals (95%) of the importance of compounds were obtained with 50 random forests of 1000 trees.

Fig.3 Regression tree for prediction of all sensory descriptors from volatile compounds
Legend: Number (n) of samples for each group defined by a specific sensory profile with sample reference number

Fig 1.

(a)



(b)

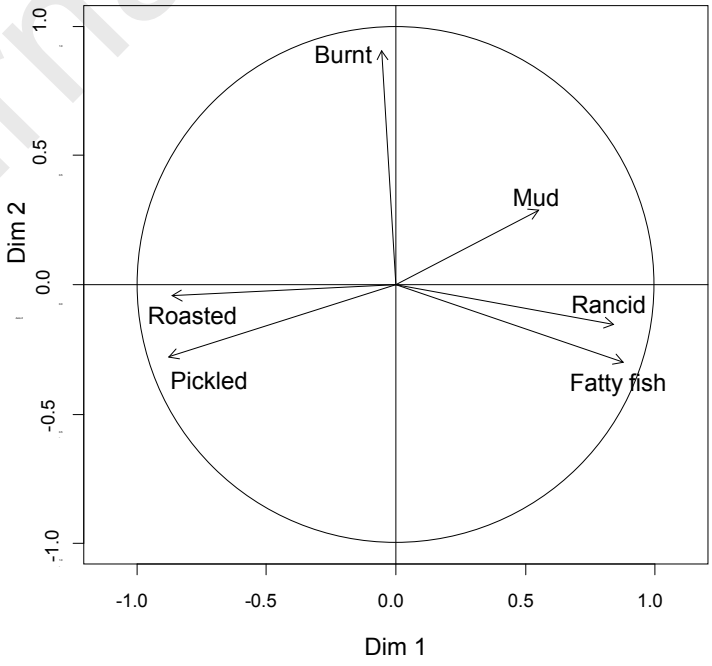


Fig.2

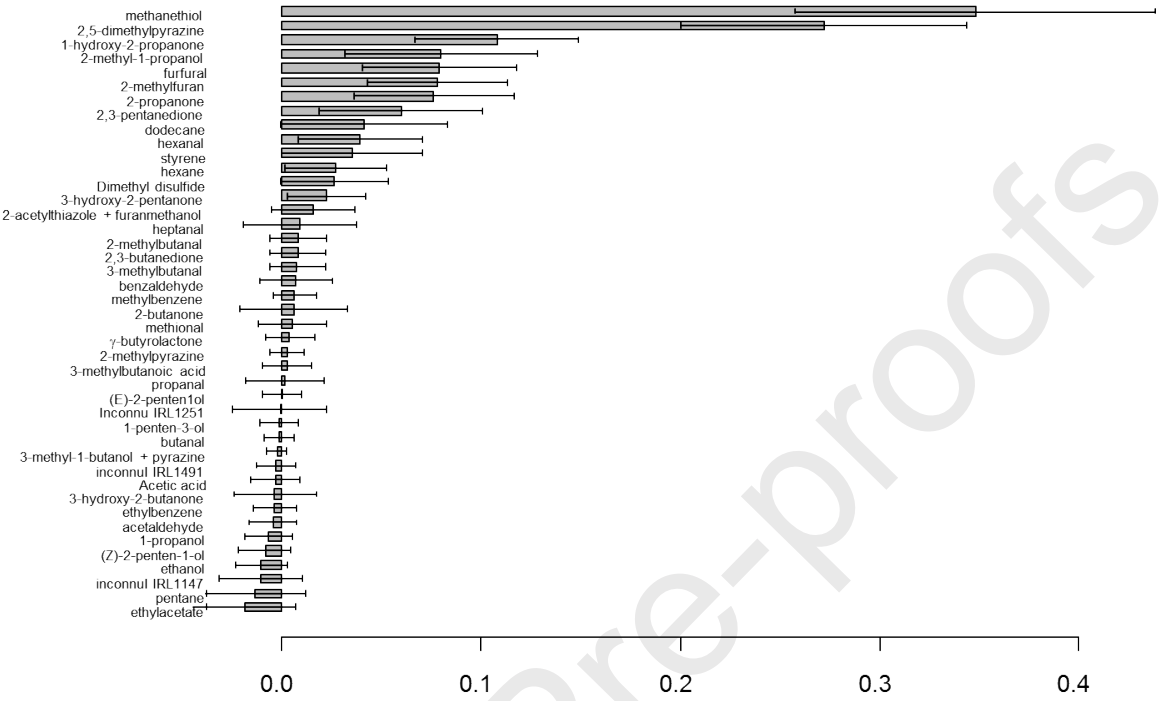


Fig.3

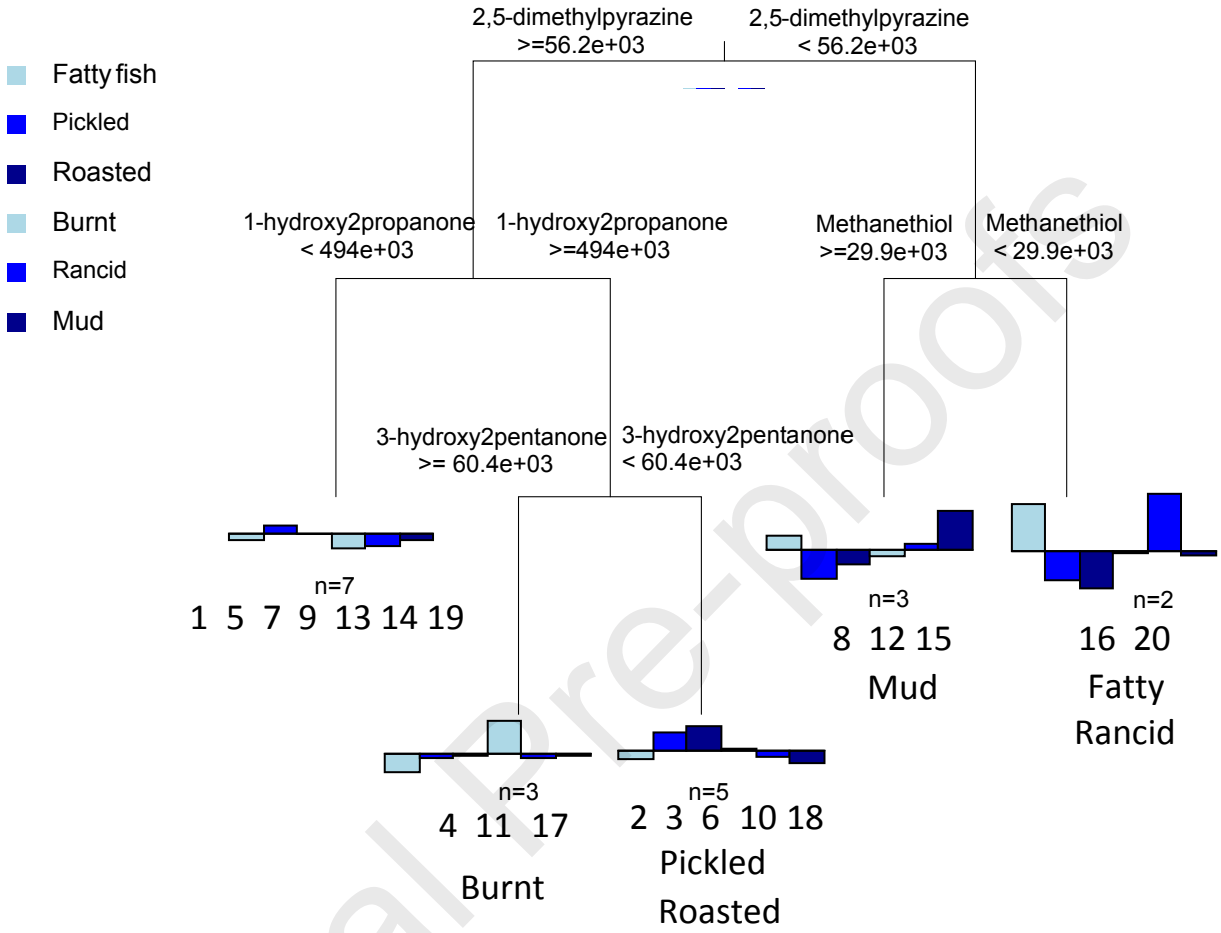


Table 1. Factor levels for the experimental design

	Independent factors			
Run	E/S	X	HT	CT
1	0.4	2	90	90
2	0.25	10	50	60
3	0.25	6	90	60
4	0.1	10	10	90
5	0.25	6	50	60
6	0.25	6	50	90
7	0.25	6	10	60
8	0.1	2	90	30
9	0.4	6	50	60
10	0.25	6	50	60
11	0.1	6	50	60
12	0.25	2	50	60
13	0.25	6	50	30
14	0.4	10	10	30
15	0.1	2	10	30
16	0.4	10	90	30
17	0.25	6	50	60
18	0.1	10	90	90
19	0.4	2	10	90
20 (extra sample)	0.4	0	90	90

Independent factors E/S, X, HT, CT represent the Enzyme/Substrate ratio ($g.100g^{-1}$), Xylose concentration ($g.kg^{-1}$), Hydrolysis Time at 40°C (min) and Cooking Time at 95°C (min) respectively

Table 2: Volatile compounds identified in the hydrolysates.

Volatile compound	CAS number	RI[a]	Identification[b]	Odour threshold[c]	Compound origin [d]	Odour description[e]
Alkanes						
Pentane	109-66-0	500	MS, RI, Std			
Hexane	110-54-3	600	MS, RI, Std			
Dodecane	112-40-3	1198	MS, RI, Std			
Benzene compounds						
Methylbenzene	108-88-3	977	MS, RI, Std		MR[5]	sweet
Ethylbenzene	100-41-4	1138	MS, RI, Std			
Styrene	100-42-5	1268	MS, RI, Std	730		sweet balsam floral plastic
Benzaldehyde	100-52-7	1541	MS, RI, Std	350-3 500	MR[5]	strong sharp sweet bitter almond cherry
Aldehydes						
Acetaldehyde	75-07-0	702	MS, RI, Std	15-120	MR	pungent ethereal aldehydic fruity
Propanal	123-38-6	794	MS, RI, Std	9.5-37	LO	earthy alcohol wine whiskey cocoa nutty
2-methylpropanal	78-84-2	794	MS, RI, Std	0.1-2.3	MR[1]	fresh aldehydic floral green
butanal	123-72-8	869	MS, RI, Std	9-37.3		pungent cocoa musty green malty bready
2-methylbutanal	96-17-3	909	MS, RI, Std	1	MR[1]	musty cocoa coffee nutty
3-methylbutanal	590-86-3	913	MS, RI, Std	0.2-2	MR[1]	ethereal aldehydic chocolate peach fatty
Hexanal	66-25-1	1095	MS, RI, Std	4.5-5	LO[1]	fresh green fatty aldehydic grass leafy fruity sweaty
Heptanal	111-71-7	1196	MS, RI, Std	3	LO[2]	fresh aldehydic fatty green herbal wine-lee ozone
Alcohols						
ethanol	64-17-5	935	MS, RI, Std	100 000	F[3], LO[3]	strong alcoholic ethereal medical
1-propanol	71-23-8	1060	MS, RI, Std	9 000		alcoholic fermented fusel musty
2-methyl-1-propanol	78-83-1	1121	MS, RI, Std	7 000		ethereal winey
1-penten-3-ol	616-25-1	1180	MS, RI, Std	400	LO[4]	pungent horseradish green vegetable tropical fruity
(E)-2-penten-1-ol	1576-96-1	1326	MS, RI, Std			mushroom
(Z)-2-penten-1-ol	1576-95-0	1334	MS, RI, Std			green plastic ethereal fruity
Ketones						
2-propanone	67-64-1	814	MS, RI, Std	500 000		solvent ethereal apple pear
2-butanone	78-93-3	900	MS, RI, Std	50 000		acetone-like ethereal fruity camphor
2,3-butanedione	431-03-8	977	MS, RI, Std	2.3-6.5		strong butter sweet creamy pungent caramel
2,3-pentanedione	600-14-6	1076	MS, RI, Std			pungent sweet butter creamy caramel nutty cheese
3-hydroxy-2-butanone	513-86-0	1297	MS, RI	800		sweet buttery creamy dairy milky fatty

Volatile compound	CAS number	RI[a]	Identification[b]	Odour threshold[c]	Compound origin [d]	Odour description[e]
1-hydroxy-2-propanone	116-09-6	1312	MS, RI, Std			pungent sweet caramellic ethereal
3-hydroxy-2-pentanone	3142-66-3	1355	MS, RI			herbal truffle
Acids and esters						
Ethyl acetate	141-78-6	883	MS, RI, Std	5-5 000		ethereal fruity sweet weedy green
Acetic acid	64-19-7	1452	MS, RI, Std		F[1]	sharp pungent sour vinegar
3-methylbutanoic acid	503-74-2	1682	MS, RI, Std	120-700	F[1]	sour stinky feet sweaty cheese tropical
Furans						
2-methylfuran	534-22-5	863	MS, RI, Std		MR	ethereal acetone chocolate
Furfural	98-01-1	1471	MS, RI, Std	3 000-23 000	MR[3]	sweet woody almond fragrant baked bread
Pyrazines						
2-methylpyrazine	109-08-0	1281	MS, RI, Std	60-105 000	MR	nutty cocoa roasted chocolate peanut green
2,5-dimethylpyrazine	123-32-0	1339	MS, RI, Std	800-1 800	MR[4]	cocoa roasted nuts roast beef woody grass medical
Sulfur compounds						
Methanethiol	74-93-1	676	MS, RI	0.02		decomposing cabbage garlic
Dimethyl disulfide	624-92-0	1085	MS, RI, Std	0.16-12	M[3], F[3]	sulfurous vegetable cabbage onion
Methional	3268-49-3	1465	MS, RI, Std	0.2	MR[2,3]	musty potato tomato earthy vegetable creamy
Others						
Unknown LRI 1147						
3-methyl-1-butanol + pyrazine	123-51-3	1228	MS, RI, Std MS, RI	250-300 +	F[3]	fusel oil alcoholic whiskey fruity banana pungent sweet corn like roasted hazelnut barly
Unknown LRI 1251						
Unknown LRI 1491						
g-butyrolactone	96-48-0	1655	MS, RI, Std			creamy oily fatty caramel
2-acetylthizaole + 2-furanmethanol	24295-03-2 98-00-0	1670	MS, RI, Std MS, RI, Std			nutty popcorn roasted peanuts hazelnut alcoholic chemical musty sweet caramel bread coffee

[a]RI: Retention Index (RI) calculated on a DB-WAX column

[b]Methods of identification of the volatile compounds : RI: Comparison of the retention index calculated with the literature, MS: comparison of the mass spectra of the compound with a database, Std : comparison of the retention index of the volatile compound with that of the corresponding standard

[c]Odour threshold expressed in parts per billion (http://www.leffingwell.com/odorthre.htm)

[d]Compound origin :LO: lipid oxidation, MR: Maillard reaction, F: fermentation, M: marine, O: other

[1]Peinado et al. (2016) LWT 66:444-452, [2]Varlet et al. (2007) Food Chemistry 1536-1556, [3]Giri et al., Food Res Int 43:1027-1040, [4]Peinado et al. (2016b) Food Chem 212:612-619, [5]Chung et al. (2002)

[d] www.thegoodscentscompany.com

Table 3: Relative quantity of the volatile compounds of the fish hydrolysates expressed in relative peak are per gram of product. Means are obtained from 3 measures. Standard deviation (SD) is specified for all the samples.

Volatile compound	**Mean relative peak area / g of product (x10^3) ± SD**									
	1	2	3	4	5	6	7	8	9	10
Alkanes										
Pentane	154 ± 24	111 ± 10	122 ± 19	68 ± 10	96 ± 6	68 ± 12	43 ± 1	57 ± 5	177 ± 49	84 ± 21
Hexane	45 ± 5	28 ± 6	24 ± 2	19 ± 1	23 ± 6	25 ± 6	20 ± 4	15 ± 2	30 ± 8	17 ± 3
Dodecane	64 ± 2	62 ± 21	50 ± 9	79 ± 15	71 ± 3	53 ± 9	89 ± 6	85 ± 12	52 ± 10	101 ± 5
Benzene compounds										
Methylbenzene	85 ± 19	-	89 ± 8	16 ± 2	18 ± 1	21 ± 6	52 ± 8	71 ± 3	32 ± 3	23 ± 2
Ethylbenzene	122 ± 1	17 ± 9	71 ± 5	45 ± 6	40 ± 4	39 ± 8	82 ± 6	97 ± 4	73 ± 3	68 ± 6
Styrene	252 ± 38	198 ± 41	180 ± 5	199 ± 13	223 ± 11	164 ± 16	172 ± 2	205 ± 20	240 ± 50	216 ± 12
Benzaldehyde	57 ± 5	67 ± 4	78 ± 10	60 ± 8	57 ± 3	71 ± 5	58 ± 10	54 ± 3	59 ± 3	60 ± 6
Aldehydes										
Acetaldehyde	299 ± 9	346 ± 36	307 ± 9	359 ± 15	336 ± 3	305 ± 19	442 ± 15	244 ± 2	335 ± 13	374 ± 23
Propanal	1107 ± 169	440 ± 22	563 ± 23	425 ± 19	602 ± 15	370 ± 23	911 ± 52	1095 ± 31	633 ± 23	864 ± 37
Butanal	342 ± 39	172 ± 38	188 ± 25	200 ± 47	180 ± 25	169 ± 15	260 ± 20	238 ± 3	241 ± 9	225 ± 23
2-methylbutanal	1040 ± 30	1249 ± 172	1487 ± 183	886 ± 102	1013 ± 75	1471 ± 64	648 ± 73	802 ± 13	999 ± 72	1011 ± 72
3-methylbutanal	6935 ± 410	6551 ± 109	7877 ± 329	2288 ± 158	5549 ± 106	6220 ± 161	2929 ± 195	5355 ± 23	5938 ± 214	5461 ± 86
Hexanal	397 ± 32	152 ± 5	214 ± 23	104 ± 14	223 ± 20	149 ± 6	229 ± 32	375 ± 21	251 ± 19	252 ± 1
Heptanal	55 ± 5	23 ± 4	27 ± 2	24 ± 4	29 ± 4	23 ± 3	39 ± 5	36 ± 5	33 ± 3	35 ± 7
Alcohols										
Ethanol	6512 ± 393	6511 ± 387	6183 ± 268	6311 ± 198	7935 ± 78	7553 ± 462	7178 ± 750	7338 ± 484	7455 ± 463	6033 ± 349
1-propanol	259 ± 38	152 ± 7	181 ± 17	185 ± 21	189 ± 22	164 ± 43	186 ± 24	187 ± 21	172 ± 21	194 ± 15
2-methyl-1-propanol	47 ± 4	35 ± 4	32 ± 6	39 ± 1	37 ± 4	34 ± 6	54 ± 2	36 ± 2	48 ± 9	37 ± 6
1-penten-3-ol	3636 ± 316	2021 ± 46	1873 ± 53	2580 ± 53	2269 ± 108	1424 ± 46	3387 ± 115	2681 ± 125	2385 ± 87	3059 ± 75
(E)-2-penten-1-ol	151 ± 12	78 ± 6	54 ± 2	105 ± 3	82 ± 4	45 ± 2	149 ± 15	99 ± 6	96 ± 11	123 ± 9
(Z)-2-penten-1-ol	193 ± 28	101 ± 4	121 ± 10	142 ± 11	122 ± 12	86 ± 6	146 ± 6	232 ± 16	126 ± 15	169 ± 10
Ketones										
2-propanone + 2-methylpropanal	6041 ± 799	7095 ± 398	5341 ± 334	8189 ± 559	5474 ± 117	5700 ± 402	7413 ± 608	2619 ± 304	6147 ± 353	6107 ± 322
2-butanone	691 ± 49	577 ± 30	703 ± 29	638 ± 29	602 ± 66	476 ± 11	638 ± 21	408 ± 34	604 ± 17	609 ± 28
2,3-butanedione	1103 ± 26	733 ± 63	852 ± 80	654 ± 66	767 ± 18	860 ± 60	836 ± 113	801 ± 58	879 ± 167	721 ± 9
2,3-pentanedione	95 ± 7	76 ± 18	95 ± 16	57 ± 7	97 ± 17	76 ± 11	64 ± 4	212 ± 35	89 ± 12	115 ± 8
3-hydroxy-2-butanone	1674 ± 198	1638 ± 118	1169 ± 63	2151 ± 54	1824 ± 57	2086 ± 51	1280 ± 28	1584 ± 3	1500 ± 68	1757 ± 60
1-hydroxy-2-propanone	348 ± 35	630 ± 42	553 ± 47	1305 ± 45	494 ± 25	882 ± 29	470 ± 11	111 ± 9	424 ± 31	494 ± 32
3-hydroxy-2-pentanone	56 ± 2	44 ± 3	49 ± 2	61 ± 3	50 ± 1	43 ± 3	57 ± 3	59 ± 2	37 ± 3	54 ± 5

Volatile compound	Mean relative peak area / g of product (x10^3) ± SD									
	1	2	3	4	5	6	7	8	9	10
Acids and esters										
Ethyl acetate	122 ± 26	91 ± 21	80 ± 2	102 ± 17	108 ± 10	96 ± 15	120 ± 7	102 ± 1	101 ± 2	99 ± 5
Acetic acid	599 ± 160	422 ± 19	494 ± 129	758 ± 194	601 ± 75	778 ± 304	574 ± 71	463 ± 123	464 ± 112	693 ± 83
3-methylbutanoic acid	46 ± 7	41 ±11	43 ± 1	20 ± 3	35 ± 6	44 ± 3	33 ± 2	40 ± 4	38 ± 2	41 ± 3
Furans										
2-methylfuran	46 ± 27	71 ± 10	72 ± 7	162 ± 19	59 ± 12	94 ± 6	64 ± 27	-	68 ± 8	52 ± 9
Furfural	103 ± 6	111 ± 12	94 ± 1	107 ± 9	87 ± 1	83 ± 3	83 ± 3	39 ± 3	91 ± 4	91 ± 4
Pyrazines										
2-methylpyrazine	26 ± 5	54 ± 13	34 ± 4	69 ± 2	32 ± 7	60 ± 11	25 ± 3	32 ± 6	32 ± 5	29 ± 6
2,5-dimethylpyrazine	75 ± 24	85 ± 9	133 ± 7	85 ± 8	91 ± 20	152 ± 34	78 ± 6	-	77 ± 4	93 ± 10
Sulfur compounds										
Methanethiol	41 ± 6	43 ± 8	39 ± 4	40 ± 3	43 ± 2	46 ± 1	44 ± 1	33 ± 2	47 ± 4	53 ± 8
Dimethyl disulfide	40 ± 6	118 ± 18	88 ± 9	124 ± 2	77 ± 10	88 ± 3	72 ± 12	33 ± 5	110 ± 36	71 ± 3
Methional	46 ± 1	55 ± 3	54 ± 6	48 ± 4	48 ± 2	45 ± 1	42 ± 1	41 ± 3	43 ± 1	48 ± 3
Others										
Unknown LRI 1147	32 ± 2	27 ± 3	23 ± 3	16 ± 5	17 ± 1	14 ± 3	18 ± 2	16 ± 2	15 ± 3	21 ± 2
3-methyl-1-butanol + pyrazine	96 ± 4	122 ± 24	86 ± 3	128 ± 14	135 ± 11	142 ± 16	94 ± 4	99 ± 5	108 ± 6	96 ± 2
Unknown LRI 1251	71 ± 2	36 ± 1	36 ± 5	44 ± 7	55 ± 12	58 ± 11	30 ± 4	27 ± 5	47 ± 5	67 ± 12
Unknown LRI 1491	93 ± 4	67 ± 9	67 ± 14	70 ± 9	75 ± 9	87 ± 12	79 ± 12	94 ± 7	72 ± 11	80 ± 8
g-butyrolactone	28 ± 3	27 ± 4	22 ± 4	28 ± 5	22 ± 2	27 ± 5	25 ± 5	21 ± 4	22 ± 3	23 ± 2
2-acetylthizaole + 2-furanmethanol	39 ± 4	37 ±1	31 ± 4	42 ± 1	26 ± 5	49 ± 3	19 ± 1	12 ± 1	27 ± 4	26 ± 2

Volatile compound	Mean relative peak area / g of product (x10^3) ± SD									
	11	12	13	14	15	16	17	18	19	20
Alkanes										
Pentane	66 ± 11	73 ± 5	166 ± 12	141 ± 28	373 ± 53	107 ± 8	194 ± 38	89 ± 24	124 ± 17	187 ± 30
Hexane	17 ± 1	28 ± 7	34 ± 4	35 ± 8	99 ± 16	42 ± 7	41 ± 9	17 ± 4	16 ± 2	27 ± 6
Dodecane	87 ± 6	100 ± 23	42 ± 8	97 ± 24	50 ± 5	123 ± 15	86 ± 10	92 ± 15	93 ± 8	118 ± 18
Benzene compounds										
Methylbenzene	22 ± 3	41 ± 12	27 ± 4	45 ± 7	35 ± 2	45 ± 4	17 ± 1	25 ± 2	28 ± 3	31 ± 6
Ethylbenzene	64 ± 7	91 ± 4	73 ± 6	153 ± 20	42 ± 5	80 ± 6	36 ± 2	98 ± 8	88 ± 4	80 ± 6
Styrene	215 ± 1	263 ± 37	115 ± 10	197 ± 45	210 ± 3	284 ± 32	190 ± 11	183 ± 51	259 ± 27	295 ± 17
Benzaldehyde	58 ± 7	56 ± 5	57 ± 10	60 ± 7	50 ± 3	63 ± 2	55 ± 5	73 ± 7	55 ± 8	56 ± 3
Aldehydes										
Acetaldehyde	414 ± 13	305 ± 7	296 ± 37	458 ± 22	409 ± 15	377 ± 3	374 ± 14	399 ± 8	406 ± 10	334 ± 22
Propanal	1106 ± 45	1070 ± 76	955 ± 84	1461 ± 44	1250 ± 24	941 ± 26	1022 ± 85	510 ± 26	1126 ± 84	1508 ± 215
Butanal	307 ± 11	269 ± 21	230 ± 30	309 ± 6	232 ± 25	233 ± 5	272 ± 21	238 ± 18	342 ± 11	375 ± 7
2-methylbutanal	992 ± 91	653 ± 46	554 ± 20	555 ± 55	381 ± 27	998 ± 46	970 ± 20	2459 ± 69	523 ± 35	647 ± 11
3-methylbutanal	4720 ± 143	4662 ± 135	4054 ± 626	3174 ± 87	2220 ± 245	5922 ± 492	5453 ± 221	7391 ± 206	2960 ± 97	5599 ± 123
Hexanal	297 ± 37	377 ± 36	331 ± 22	316 ± 43	351 ± 26	329 ± 12	315 ± 18	130 ± 9	325 ± 32	738 ± 57
Heptanal	38 ± 6	41 ± 4	43 ± 6	38 ± 6	50 ± 5	47 ± 2	40 ± 6	32 ± 7	57 ± 3	81 ± 8
Alcohols										
Ethanol	7824 ± 586	7497 ± 386	6772 ± 608	7139 ± 367	6595 ± 405	6524 ± 425	6085 ± 755	6464 ± 308	6817 ± 538	6295 ± 415
1-propanol	234 ± 41	199 ± 16	210 ± 22	213 ± 29	149 ± 11	203 ± 22	240 ± 12	218 ± 24	211 ± 29	207 ± 59
2-methyl-1-propanol	45 ± 6	45 ± 8	28 ± 5	47 ± 2	53 ± 8	64 ± 4	48 ± 5	45 ± 2	40 ± 1	73 ± 8
1-penten-3-ol	3551 ± 386	2821 ± 197	2353 ± 41	3185 ± 92	2202 ± 81	2408 ± 56	3275 ± 171	3305 ± 122	4231 ± 21	4432 ± 141
(E)-2-penten-1-ol	137 ± 14	111 ± 11	81 ± 4	129 ± 4	94 ± 4	85 ± 5	132 ± 8	120 ± 7	199 ± 2	188 ± 16
(Z)-2-penten-1-ol	180 ± 8	154 ± 33	167 ± 7	176 ± 14	112 ± 10	143 ± 9	197 ± 10	206 ± 13	198 ± 5	319 ± 17
Ketones										
2-propanone + 2-methylpropanal	5866 ± 479	3832 ± 87	2896 ± 260	5027 ± 123	3021 ± 81	4473 ± 429	5817 ± 518	11064 ± 208	5998 ± 273	2638 ± 203
2-butanone	609 ± 44	458 ± 17	415 ± 27	542 ± 25	282 ± 38	560 ± 28	623 ± 62	838 ± 7	597 ± 5	468 ± 37
2,3-butanedione	786 ± 137	1027 ± 62	1051 ± 83	822 ± 14	903 ± 78	925 ± 125	710 ± 25	716 ± 79	828 ± 66	843 ± 16
2,3-pentanedione	113 ± 21	146 ± 22	170 ± 32	157 ± 29	123 ± 13	180 ± 25	111 ± 13	76 ± 14	68 ± 12	124 ± 24
3-hydroxy-2-butanone	1630 ± 81	2107 ± 64	2042 ± 49	1755 ± 62	1849 ± 37	2190 ± 74	2316 ± 139	1323 ± 63	1717 ± 19	1717 ± 89
1-hydroxy-2-propanone	502 ± 32	211 ± 8	212 ± 4	273 ± 15	135 ± 5	314 ± 12	552 ± 35	1123 ± 77	295 ± 9	113 ± 6
3-hydroxy-2-pentanone	85 ± 12	55 ± 5	76 ± 4	64 ± 4	56 ± 1	55 ± 2	67 ± 3	60 ± 5	55 ± 2	61 ± 2

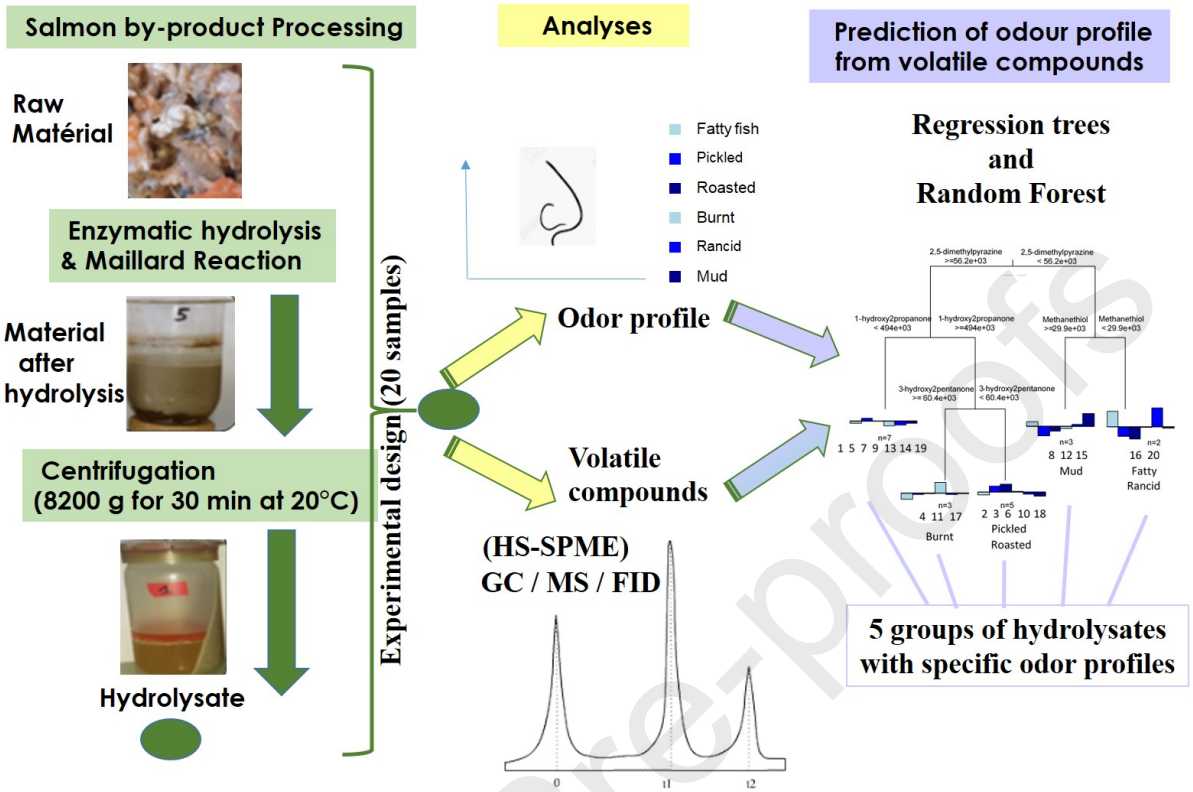
Volatile compound	Mean relative peak area / g of product (x10^3) ± SD									
	11	12	13	14	15	16	17	18	19	20
Acids and esters										
Ethyl acetate	131 ± 4	103 ± 3	90 ± 13	126 ± 6	86 ± 29	118 ± 21	105 ± 15	110 ± 8	125 ± 7	76 ± 15
Acetic acid	587 ± 330	371 ± 86	404 ± 45	449 ± 87	671 ± 128	477 ± 84	636 ± 134	907 ± 116	455 ± 23	787 ± 200
3-methylbutanoic acid	38 ± 7	30 ± 4	28 ± 3	29 ± 2	23 ± 6	39 ± 4	33 ± 2	45 ± 1	23 ± 1	32 ± 1
Furans										
2-methylfuran	67 ± 7	-	50 ± 9	54 ± 7	-	66 ± 8	71 ± 12	192 ± 32	61 ± 10	-
Furfural	82 ± 8	62 ± 2	51 ± 5	83 ± 9	39 ± 2	64 ± 3	96 ± 10	150 ± 5	82 ± 3	48 ± 5
Pyrazines										
2-methylpyrazine	43 ± 7	33 ± 13	74 ± 4	26 ± 3	29 ± 7	51 ± 3	31 ± 2	61 ± 9	47 ± 8	23 ± 3
2,5-dimethylpyrazine	85 ± 13	45 ± 12	58 ± 9	62 ± 10	24 ± 3	55 ± 8	110 ± 5	159 ± 17	71 ± 12	19 ± 2
Sulfur compounds										
Methanethiol	42 ± 5	32 ± 1	40 ± 10	45 ± 9	33 ± 1	27 ± 5	44 ± 2	40 ± 2	42 ± 3	19 ± 1
Dimethyl disulfide	66 ± 10	31 ± 6	40 ± 7	67 ± 1	33 ± 5	71 ± 12	62 ± 8	140 ± 22	48 ± 5	19 ± 5
Methional	49 ± 3	37 ± 2	41 ± 2	38 ± 4	32 ± 2	51 ± 1	49 ± 2	68 ± 4	36 ± 1	43 ± 5
Others										
Unknown LRI 1147	41 ± 4	15 ± 3	14 ± 3	30 ± 9	14 ± 2	16 ± 1	21 ± 3	27 ± 4	21 ± 3	30 ± 2
3-methyl-1-butanol + pyrazine	108 ± 10	115 ± 10	93 ± 7	101 ± 10	104 ± 1	99 ± 5	106 ± 5	89 ± 4	90 ± 3	100 ± 4
Unknown LRI 1251	45 ± 5	67 ± 5	151 ± 48	48 ± 8	42 ± 5	143 ± 12	77 ± 17	64 ± 6	63 ± 8	97 ± 9
Unknown LRI 1491	81 ± 14	89 ± 8	90 ± 23	78 ± 12	84 ± 15	75 ± 6	94 ± 8	83 ± 10	82 ± 7	96 ± 7
g-butyrolactone	26 ± 3	26 ± 3	22 ± 1	22 ± 3	21 ± 2	25 ± 2	24 ± 1	30 ± 1	23 ± 3	27 ± 6
2-acetylthizaole + 2-furanmethanol	27 ± 4	18 ± 1	14 ± 3	13 ± 1	8 ± 2	23 ± 4	26 ± 2	68 ± 3	27 ± 4	42 ± 4

Writing of the article: MCa, CR and PC
Experimental design conception: RB, MCa, MCh
Technical organization of the study, hydrolysate preparation and analysis: MCh; CDM
Sensory analysis and data analysis: JC and MCa
HS-SPME/GC-MS and data analysis: CF, CR, CP
Regression tree and Randon forest methodology, data treatment: PC
Supervision of the research project: RB
Manuscript review: RB, CP

Declarations of interest: none

Authors declared no conflict of interest. The authors of this manuscript received no payment or benefit, directly or indirectly, from any organization.

Graphical abstract



Highlights

- Regression Trees and Random Forests methodology : a tool to predict a whole sensory profile
- Four main volatile compounds identified in the final regression tree made it possible to separate hydrolysates into five groups
- Prediction results may be sensitive to sensory measurements variability
- Appropriate process conditions combining hydrolysis parameters and Maillard Reaction lead to specific roasted odor