Food Chemistry 367 (2022) 130760

Contents lists available at ScienceDirect

Food Chemistry

journal homepage: www.elsevier.com/locate/foodchem

Evaluation of volatile metabolites as potential markers to predict naturally-aged seed vigour by coupling rapid analytical profiling techniques with chemometrics

Tingting Zhang [a], Charfedinne Ayed [b], Ian D. Fisk [b,c], Tong Pan [a], Jianhua Wang [a], Ni Yang [b,*], Qun Sun [a,*]

[a] Department of Plant Genetics &Breeding and Seed Science, College of Agronomy and Biotechnology, China Agricultural University/Beijing Key Laboratory of Crop Genetic Improvement/The Innovation Center (Beijing) of Crop Seed Sciences Ministry of Agriculture, Beijing 100193, PR China
[b] Division of Food, Nutrition and Dietetics, School of Biosciences, University of Nottingham, Sutton Bonington Campus, Loughborough LE12 5RD, United Kingdom
[c] The University of Adelaide, North Terrace, Adelaide, South Australia, Australia

ARTICLE INFO

Keywords:
Seed vigour
Natural ageing
Artificial ageing
Gas chromatography – ion mobility spectrometry
Chemometrics

ABSTRACT

Rapid volatile detection methods for seed vigour rely heavily on artificial ageing (AA), however the comparability of volatile organic compounds (VOCs) to natural ageing (NA) and practicability of the detection models were not well known. In this study, VOCs between AA and NA sweet corn seeds were compared and Partial Least Squares Regression (PLS-R) models were constructed based on AA to predict the seed vigour of NA. A total of 33 VOCs were identified, among which aldehydes showed the highest consistency between NA and AA. Furthermore, a AS-PLS-R model with variable importance in projection (VIP > 1) and Pearson Correlation Coefficient (r > 0.9) algorithms, which was built on 3 volatile markers: benzaldehyde monomer, *n*-nonanal, 1-butanol monomer, achieved the best performance (R^2p of 0.901 and RMSEP of 0.050). Therefore, coupling Gas Chromatography- Ion Mobility Spectrometry (GC-IMS) with chemometrics can be an effective way to monitor and predict stored seeds vigour.

1. Introduction

Sweet corn (*Zea mays,* saccharata Sturt.) is one of the most popular vegetables all over the globe owing to its good taste and high nutritional value. It contains several phytochemicals, as well as dietary fibre, protein, fats, starch and vitamin C (Coşkun, Yalçın, & Özarslan, 2006). Moreover, the seed-eating quality, i.e. nutritional value, have been proven to be further enhanced by sprouting/ germination (Xiang et al., 2017). However, during seed aging, fat, protein, starch and other substances are degraded. This will cause taste deterioration and reduction of their value. Hence, a reliable future supply of sweet corn seeds with high quality (vigour) is important for the global food industry.

Unlike field corns, sweet corn is naturally produced by recessive mutations in one or several characteristic genes of the endosperm of corn seed that control the conversion of sugar to starch (Singh, Langyan, & Yadava, 2014). The high soluble sugar content and low starch content in sweet corn seeds have a high economic value, however they are also more susceptible to ageing, than field corn seeds during storage (Wang et al., 2020). During sweet corn seed storage, seed vigour is lost and ultimately seed viability degrades over storage time (Walters, 1998). Therefore, real-time monitoring and prediction of the longevity of sweet corn seeds during storage is vitally important for food companies and farmers.

Artificial ageing (AA) typically uses a combination of high ambient temperature and high relative humidity conditions to simulate natural ageing (NA). A few studies (Xu et al., 2015; Yan et al., 2018; Neto, Custodio, & Takaki, 2001) investigated the physiological, biochemical and structural alterations that occur in seeds during ageing, and the insights gained from AA were then used for longevity prediction and mechanistic investigations of NA. However, many researchers are sceptical about the consistency of metabolic mechanisms of seeds during these two different ageing processes (Neto, Custodio, & Takaki, 2001; Priestley & Leopold, 1983). It was found that the alterations of chemical substances in soybean seeds induced by AA were not the same as those induced by NA (Priestley & Leopold, 1983). Another study (Rajagopal & Sen-Mandi, 1992) also observed that the acid phosphatase activity of AA

* Corresponding authors.
E-mail addresses: Ni.Yang@nottingham.ac.uk (N. Yang), sunqun@cau.edu.cn (Q. Sun).

https://doi.org/10.1016/j.foodchem.2021.130760
Received 31 January 2021; Received in revised form 8 July 2021; Accepted 1 August 2021
Available online 3 August 2021
0308-8146/© 2021 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

rice seeds was higher than that of NA. Subsequently, Neto, Custodio, and Takaki (2001) showed the degradation pattern of proteins in NA and AA bean seeds and concluded that they are probably driven by different physiological mechanisms.

The volatile organic compounds (VOCs) released by seeds during ageing are proposed to be the major by-products of catabolism, such as lipid peroxidation, Maillard reactions and alcoholic fermentation (Colville et al., 2012). Zhang et al. (1993) were the first to discover 11 VOCs released from carrot seeds, including propylene, methanol, acetaldehyde, butene, ethanol, acetone, isopropanol, isobutanal, acetic acid and ethyl acetate. A correlation between volatile emission and the vigour of seeds during NA and AA has been widely investigated. Specifically, Zhang et al. (1993) and Colville et al. (2012) observed that the abundance of the VOCs generated from seeds increased with the time of NA (23 °C) and AA (50 °C). Hailstones and Smith (1989) heated soybean seeds to 130 °C for 2 h and reported that the amounts of thermally-derived aldehydes in soybeans increased with the decrease of seed vigour. Secondary products of lipid peroxidation and sugar hydrolysis can participate in Amadori and Maillard reactions, which will promote seed deterioration by non-enzymatic modification of proteins and enzymes (Narayana & Sun, 2000). Sun and Leopold (1995) also demonstrated that the Maillard reaction was driven in soybean seeds during NA and AA. Many VOCs emitted by seeds are reactive and may be toxic to seeds, perpetuating reactions will accelerate seed deterioration, this has been shown for both acetaldehyde and ethanol (Mira et al., 2016; Zhang et al., 1995). Therefore, the VOCs that can indicate the types of chemical reactions occurring in seeds are potentially the causes and effects of seed vigour loss. The analysis of volatile molecules provides a feasible approach for evaluating the chemical reactions that occur in seeds and developing biochemical or molecular markers for seed quality loss. However, only a few niche studies have investigated in this area, and the similarities and differences of volatile metabolites released by sweet corn seeds during NA and AA still have yet been fully explained.

Gas chromatography–ion mobility spectrometry (GC-IMS) and electronic nose (E-nose) techniques have been widely utilised for volatile analysis as rapid profiling tools, due to their fast detection speed, convenient operation and portability. These two techniques were used to understand the loss in seed/ kernel quality by detecting and characterising the broad range of complex VOCs (Ghasemi-Varnamkhasti et al., 2018; Gu & Wang, 2019; Ying, Liu, Hui, & Fu, 2015; Zhang et al., 2020b). However, most studies have used accelerated ageing protocols to build generic predictions and have not fully considered the limitations of these approaches.

In summary, the similarities and differences of volatile metabolites released by sweet corn seeds during NA and AA still have yet to be fully explained and there is no effective model that can be used to fully predict the natural ageing process. Therefore, this study aimed to: i) evaluate the similarities and differences between the volatile metabolites of naturally and artificially aged seeds; ii) compare two rapid volatile profiling techniques (E-nose and GC-IMS) and evaluate their ability to predict the vigour of naturally aged seeds; iii) investigate the performance of PLS-R models based on AA data sets for predicting the vigour of seeds during NA; iv) identify key marker compounds using variable importance in projection (VIP) and Pearson correlation coefficient (PCC) from the optimum prediction models.

2. Materials and methods

2.1. Seed samples

Sweet corn seeds (37.5 kg), Lvsechaoren, were purchased from the Hezhiyuan Seed Company (Weifang, Shandong, China) and their moisture content and raw germination percentage (GP) was evaluated to ensure seed stability and viability, both moisture content (11.0% wet weight basis) and GP (85.5%) were within specification limits and were deemed suitable for the study. Seeds were divided into three batches.

- Control samples (1.5 kg) were immediately sealed in a polyethylene bag and stored at −20 °C.
- Natural ageing (NA) samples (18 kg) were sub-divided into 12 sub-samples. Each sub-sample weighed 1.5 kg and was stored in cotton bags at ambient room temperature, which was regulated between 17 ~ 28 °C and 30 ~ 60% RH.
- Artificial ageing (AA) samples (18 kg) were sub-divided into 12 sub-samples. Each sub-sample weighed 1.5 kg and was stored in nylon mesh bags. Samples were then placed in an electric oven (45 °C) and suspended over distilled water inside sealed glass containers (100 % RH) (Muasya, Simiyu, Muui, Rao, Dulloo, & Gohole, 2009; Zhang et al., 2017).

For NA, every month, one sub-sample was taken and GP evaluated. This was continued until the GP of the seed batch reduced to 75% and then samples were taken every half month once the GP was below 75%. For AA, every 16 h, one sub-sample was taken and GP evaluated. This was continued until the GP of the seed batch reduced to 75% and then samples were taken every 8 h once the GP was below 75%. The isolated sub-samples were dried in an incubator at 23 ± 1 °C until the moisture content returned to the original value of 11.0% using a PM-8188-A grain moisture analyser (KETT Corp., Japan). Physically damaged seeds were manually removed and 200 seeds were then randomly selected from every sub-sample for germination assessment, any remaining seeds were immediately sealed in a polyethylene bag and stored at −20 °C. When the GP of the naturally and artificially aged seeds reduced to 75% and 65%, the corresponding sub-samples were retained for further experiments.

2.2. Germination assessment

Germination assessment was conducted according to the standard germination test (Ilbi, Kavak, & Eser, 2009) using four replicates of 50 sweet corn seeds (about 7.1 g/ replicate). Briefly, seeds were first surface sterilized by immersing them into 0.5% sodium hypochlorite for 5 min and rinsed with distilled water for 1 min. Then the seeds were placed between two moist germination papers (Anchor Corp., America, 25 × 38 cm). Papers were rolled and kept at 25 ± 1 °C with continuous light germination cabinets for seven days. Seeds exceeding 2 mm in both plumule and radicle were evaluated as germinated seeds (seedlings). Seedlings with spiral formed hypocotyls, stunted primary root and without radicle protrusion were considered as abnormal seedlings and eliminated. The rest of the seedlings were normally germinated seedlings (Zhang et al., 2020a). GP was calculated as follows:

$$GP\ (\%) = \frac{number\ of\ normally\ germinated\ seedlings}{total\ number\ of\ seeds} \times 100\%$$

2.3. Electronic nose (E-nose) analysis

Volatiles were detected by an E-nose device (PEN3; AIRSENSE Company, Germany), which consists of ten metal oxide sensors: 1) W1C (very sensitive to aromatic compounds; Toluene, 10 ppm); 2) W5S (broad range of sensitivity, reacts with nitrogen oxides; NO_2, 1 ppm); 3) W3C (Sensitive to ammonia; Benzene, 10 ppm); 4) W6S (Selective sensitivity to mainly hydrogen; H_2, 100 ppb); 5) W5C (Sensitive to alkanes, aromatic compounds, less polar compounds; Propane, 1 ppm); 6) W1S (Sensitive to methane. Broad range; CH_4, 100 ppm); 7) W1W (Detects inorganic sulfur compounds and sensitive to many terpenes and sulfur-containing organic compounds; H_2S, 1 ppm); 8) W2S (Detects alcohols and partially aromatic compounds; CO, 100 ppm); 9) W2W (Detects aromatic compounds, inorganic sulfur and organic compounds; H_2S, 1 ppm); 10) W3S (Reacts with methane at high concentrations; CH_4, 10 CH_4, 100 ppm). Before detection, volatiles were collected by putting a 15 g seed sample into a 50 mL sealed vial at 25 °C for 30 min. The AA data set was used as the calibration set and contained 90

artificially aged seed samples, with 30 samples in each vigour groups (85%, 75%, 65%). The NA data set was used as a prediction set and contained 45 naturally aged seed samples with 15 samples in each vigour group. The detection procedure was based on previously reported methods (Gu & Wang, 2019; Wang et al., 2019) and adjusted slightly. The detection parameters were as follows: 400 mL/min gas flow rate (bring volatiles to the sensors), 101 s acquisition time, 200 s flush time (exhaust volatiles left by the previous sample).

Representative E-nose response signals of a seed sample (NA 65%) are shown in **Fig. S1**. Each curve represents the change in the reported conductivity of each sensor over the acquisition period. The response data of the sensors is represented as a G / G0 ratio (G = conductance of ten sensors for the sample gas, and G0 = conductance of ten sensors for clean air).

2.4. GC–IMS analysis

The GC-IMS Flavourspec® instrument (G.A.S. Dortmund Company, Germany) was equipped with an auto-sampler unit, a syringe, a heated splitless injector and a radioactive ionization source for headspace (HS) analysis. The detection processes of HS-GC–IMS were conducted as described by Zhang et al. (2020b) and adjusted slightly according to sweet corn seeds characteristics. 3 g seeds were put into a 20 mL headspace vial and then incubated at 40 °C for 15 min. Subsequently, 500 μL of headspace was injected into the injector at a temperature of 85 °C. After injection, the sample was driven into a FS-SE-54-CB capillary column (15 m × 0.53 mm) at 60 °C by nitrogen (99.99% purity) at flow program as follows: 2 mL/min for 2 min, 10 mL/min for 8 min, 100 mL/min for 10 min, 150 mL/min for 5 min. The analytes were ionized using the 3H ionization source of 300 MBq activity with a positive ion mode in the ionization chamber and then driven to a 9.8 cm drift tube with a constant voltage (5 kV) at 45 °C.

Each spectrum was an average of 12 scans. Each vigour group were determined with five repetitions (about 20 seeds/ repetition). The retention index (RI) of volatile metabolites was calculated using C4-C9 *N*-ketones (Sinopharm Chemical Reagent Beijing Co., Ltd, China) as external references. Metabolites were identified by comparing RI and drift time (Dt, the time required for ions to reach the collector through the drift tube, milliseconds) of standard in the GC–IMS library (G.A.S., Dortmund, Germany). The GC-IMS fingerprint analysis was conducted by comparing GC retention time and IMS drift time. Due to the high concentration of ethanol and its serious tailing in the chromatogram, it was not selected for analysis.

2.5. Statistical analysis

Initially, principal component analysis (PCA) and hierarchical cluster analysis (HCA) as unsupervised multivariate methods were used i) to reduce the dimensionality of GC-IMS data; ii) to identify differences between seed sub-samples; iii) to investigate the possible clustering of volatile molecules and obtain a general overview of the volatile metabolites.

In order to improve signal-to-noise ratio, principal pre-treatments, autoscaling (AS) and standard normal variate (SNV), were then used to pre-process the data collected from E-nose and GC-IMS.

Afterwards, partial least squares regression (PLS-R) as a supervised multivariate regression analysis was used to explore the relationship between multiple dependent predictors X-matrix (E-nose signal data or GC–MS volatile data) and a response Y-variable (seed vigour). Specifically, the raw and pre-processed AA sub-samples data (E-nose or GC–MS) were assigned to calibration sets to establish PLS-R models for seed vigour prediction. Leave-one-out cross validation was utilized to internally verify the performance of these models. The models built based on the AA dataset in this study were referred to as AA-based models for simplicity. The performance of each AA-based model was evaluated by the coefficient of determination (R^2) of calibration dataset (R^2c), cross-validation dataset (R^2cv) and the root mean square error (RMSE) of calibration (RMSEC) and cross validation (RMSECV). The R^2 and RMSE were related to accuracy and precision, respectively. Good accuracy and precision indicate that the R^2 value is close to 1 and the RMSE is close to 0 (Gan et al., 2016; Zaragozá et al., 2013).

Consequently, the optimal AA-based models were selected based on higher R^2, lower RMSE, and smaller differences among R^2c and R^2cv. After screening and obtaining the optimal AA-based models, the data of the NA sub-samples were divided into prediction sets to external validate the practicality/ robustness of these ideal AA-based models. The criteria for this testing practicality of these AA-based models were based on the coefficient of determination of the prediction dataset (R^2p) and the root mean square error of prediction (RMSEP). Therefore, the strong practicability/ robustness models with the potential for further industrial application, should have higher R^2 (R^2c, R^2cv and R^2p) and lower RMSE (RMSEC, RMSECV and RMSEP).

In addition, Pearson correlation coefficient (PCC) analysis was used to evaluate the correlation of volatile metabolites between NA and AA. Variable importance in projection (VIP) was used to evaluate the importance of each volatile based on the PLS-R analysis results to explore the potential key metabolites. Considering the particularity of data sets (calibration sets from AA sub-samples and prediction sets from NA sub-samples), potential maker compounds was identified by VIP and VIP combined with PCC, respectively. The schematic flow from data collection to final predictions was presented in Fig. S2.

The volatile metabolites were assessed by one-way analysis of variance (ANOVA) with Duncan's test on SPSS 22.0 (IBM, New York, USA) (Liu et al., 2019). The values were presented as mean ± standard deviation of five independent determinations (Table S1). PCA was implemented using Excel XLSTAT software (2018 version, Addinsoft) (Liu et al., 2019). The processes of data pre-processing and PLS-R modelling were developed using MATLAB R2014a (The MathWorks, Natick, MA, USA) (Gu & Wang, 2019).

3. Results and discussion

3.1. Impact of natural and artificial ageing on seed vigour

Seeds continue to age under storage or high temperature and high relative humidity conditions, and the corresponding seed vigour decreases gradually. Generally, a notable characteristic change in seed vigour during ageing (NA and AA) is the reverse S-shaped survival curve (Mira et al., 2016). This curve includes a three-stage process: i) Phase I starts with a plateau phase (also called an initial asymptomatic stage); ii) Phase II follows by a rapid decreasing phase; iii) a slow decrease phase completes Phase III (Yin et al., 2017). The loss of seed vigour can be observed in both AA and NA samples over time, this is presented in Fig. 1. Germination percentage (GP) of seeds in NA and AA both fluctuate strongly between 85% and 75% (NA: 74.5%, AA: 76.0%) during Phase I. During phase II, seed vigour for both NA and AA drops rapidly from 74.5% to 54.0% and 76.0% to 55.5%, respectively. Seed vigour (GP) in Phase III was less than 50%, which has no practical significance to be predicted therefore for study we focused on the first and second phases of seed ageing.

Simple linear regression (SLR) analysis was conducted based on each phase to calculate the relationship between ageing time (x) and GP of seeds (y) under NA and AA, respectively. As shown in Fig. 1, GP for Phase I was defined by an SLR equation for NA ($y = -0.019x + 83.263$, $R^2 = 0.231$) and for AA ($y = -3.098x + 84.481$, $R^2 = 0.448$). The intercept in NA (83.263) was close to that in AA (84.481), but the slope difference between NA (-0.019) and AA (-3.098) suggested that loss of GP might occur 160 × faster in AA conditions during Phase I. Accordingly, the storage life of seeds with the same GP (>80.0%) can be easily predicted through multiplying the time of artificially aged seeds by the value by about 100. However, the SLR equation for Phase II obtained from NA data had different intercepts and slopes from these by AA, and

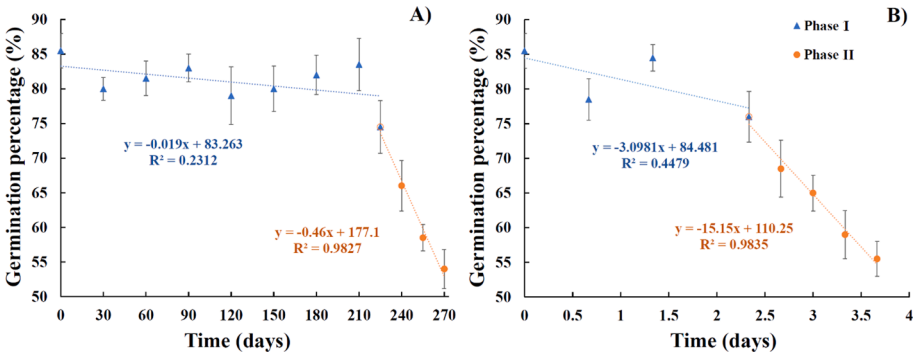

Fig. 1. Germination percentage of sweet corn seeds over time, treated by natural ageing (A) and artificial ageing (B). Linear regression for each phase (Phase I, GP > 75% and Phase II, GP < 75%) are shown in blue and red respectively. GP: germination percentage. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

no clear correlation between NA and AA can be established during Phase II.

The transformation from the plateau Phase I to the rapid decreasing Phase II is known as the critical node, which is important for seed conservation and highly concerned in the research of seed physiology and biochemistry (Yin et al., 2017). In this study, the critical node was approximately 75% for both NA and AA (Fig. 1), but it was about 225 d for the NA samples and about 2.3 d (56 h) under AA conditions (Fig. 1). Samples with GP of approximately 75% (NA: 74.5%, AA: 76.0%) were, therefore, selected as key samples for further analysis in addition to the raw samples (85%) in Phase I and samples in Phase II with a GP of approximately 65%. Samples in the same vigour group were statistically identical and samples in different vigour groups, which were chosen to represent different phases of loss of seed vigour were significantly different (*p-value* < 0.05), and therefore could be used for further analysis.

3.2. Evaluation of natural and artificial ageing by E-nose

The selected samples that represent the different vigour groups were firstly evaluated by a broad spectrum, low cost, easy to use E-nose. Partial least squares regression (PLS-R) models combined with standard normalized variate (SNV), and autoscaling (AS) were then built to predict NA based on AA (Table 2**A**). The SNV-PLS-R model was obtained from the calibration set (R^2_c = 0.686, RMSEC = 0.045) with cross-validation (R^2_{cv} = 0.615, RMSECV = 0.050). However, the performance of the model was low (R^2_p = 0.082, RMSEP = 0.081) when applied to the prediction set. Therefore, the model developed using the E-nose technology might not be able to predict vigour changes in sweet corn seeds during storage.

3.3. Evaluation of natural and artificial ageing by GC-IMS

3.3.1. Comprehensive flavour characterization of sweet corn seeds

A total of 33 VOCs, including 5 alcohols, 8 ketones, 10 esters, 7 aldehydes, 2 terpenes and 1 sulphur compounds, were identified in the ageing and control sweet corn seeds (Table S1 and Fig. 2**A**). Among these chemicals, 1-butanol is likely derived from lipid peroxidation and its relative abundance was found to increase with the decrease of seed vigour during NA and AA. Lee et al. (2000) also reported that the

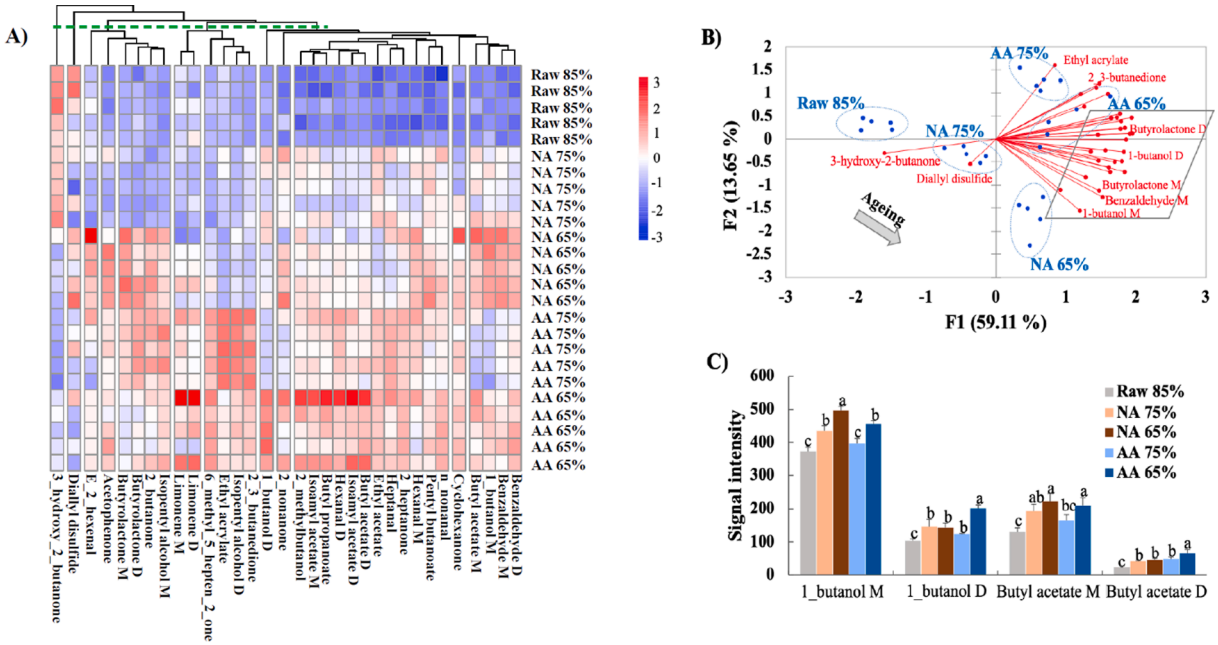

Fig. 2. Cluster heatmap analysis of VOC (A), PCA analysis of VOC (B), and monomeric and dimeric compounds (C) in sweet corn seeds with different vigour groups. Raw 85%: raw seed samples with 85% germination percentage; NA 75% and NA 65%: naturally aged seed samples with 75% and 65% germination percentage; AA 75% and AA 65%: artificially aged seed samples with 75% and 65% germination percentage. M: monomer; D: dimer. Different letters above the bar chart represented the same molecule indicate significant differences between groups of samples (p < 0.05).

amount of butanol in aged seeds was higher than non-aged soybeans and snap beans. On the contrary, 3-hydroxy-2-butanone was present in the highest relative abundance in the raw seeds and then decreased during both NA and AA. Additionally, 3-hydroxy-2-butanone was the most abundant followed by ethyl acetate in NA seeds; whereas ethyl acetate was predominant followed by 3-hydroxy-2-butanone in AA seeds. These results indicated that the changes of volatile metabolites in NA and AA seeds were not consistent. Due to the high proton affinity or signal of the analytes, the ions (monomers) may form dimers or trimers when they transferred into the drift tube (Li et al., 2020), and these are further analysed later (Section 3.3.3).

To explore the similarities and differences of the VOCs in NA and AA, volatile metabolites in seed samples were visualised and analysed by hierarchically clustered heatmap (Fig. 2**A**). The compounds were clustered into different groups, and the relative abundance of the compound in group 1 (3-hydroxy-2-butanone) reduced with ageing, while group 2 (diallyl disulfide) showed no significant difference before and after ageing ($p > 0.05$). Thirty-one of 33 compounds were clustered into three groups (group 3, 4 and 5), and the VOC content released from 65% GP of NA and AA increased by 42.1% and 57.8% respectively when compared to raw seeds. The increase in concentration of most volatiles in the seeds during ageing is consistent with the previously reported results (Colville et al., 2012; Zhang et al., 1993). However, it should be noted that the changes in VOC for group 3 and 4 compounds was different for the two sample groups (NA and AA). Hence, the VOCs in group 3 and 4 would not suitable as AA predictive markers of NA. This result may explain why the E-nose technology model was not able to predict natural ageing, as the E-nose technology detects mixed volatile compounds instead of the specific volatile molecules that relevant for natural ageing.

Intriguingly, sixteen of the nineteen compounds in group 5 that present similar variation between NA and AA belong to the chemical classes of aldehydes, alcohols and esters. This indicates that aldehydes (6/7), alcohols (3/5) and esters (7/10) have higher consistency between NA and AA than ketones (2/6), terpenes (0/2) and sulphur compounds (0/1). Whether these compounds are suitable for the construction of seed vigour prediction models and even highly important for the optimal model (biomarkers) still needs further data mining and analysis.

3.3.2. Principle component analysis

PCA was performed to reduce the GC-IMS data dimension, discriminate the VOC characteristics between the vigour groups and display the structure of the data. As shown in Fig. 2**B**, the first two principal components (PCs) of the PCA scatter plot explained 72.76% of the total variance, including 59.11% by PC1 and 13.65% by PC2, which suggested that the two PCs contain most of the information about the VOCs. Different vigour groups were clearly divided into distinct areas, suggesting that differences of volatile metabolites existed among different vigour groups. The variation of volatile metabolites in AA from 85% to 75% was greater than that from 75% to 65%, but the variation of the VOCs in NA was relatively smaller. During both ageing systems (NA and AA), more VOCs especially esters (such as butyrolactone, butyl acetate and pentyl butanoate) and aldehydes (such as benzaldehyde, hexanal and *n*-nonanal), were produced (marked by the quadrilateral in Fig. 2). Some molecules were found to be located in specific regions, such as 3-hydroxy-2-butanone and diallyl disulfide in NA, ethyl acrylate and 2, 3-butanedione in AA, indicating that critical VOCs of different ageing methods may differ. Moreover, molecules located in the region between NA and AA (such as butyrolactone M, Fig. 2**B**) had different variation trends in NA and AA (Fig. 2**A**). These may further indicate that many molecules are not suitable for predicting the vigour of seeds during NA based on AA data. Therefore, filtering interference information and selecting effective information for modelling NA from AA is essential.

3.3.3. Monomeric and dimeric compounds

Some monomers (M) and dimers (D) of the same molecules were divided into different groups (isopentyl alcohol, Fig. 2**A**) or presented in different quadrants (butyrolactone, Fig. 2**B**), and this indicated that the trends between monomer and dimer of the same molecule may be different even under the same ageing method. For example, compared with dimers, the trends of some monomers, like 1-butanol M and butyl acetate M, presented similar trend between 65% and 75% GP for both NA and AA (Fig. 2**C**). These may indicate that the monomers and dimers could lead to different performances in seed vigour prediction. Therefore, the data collected from GC-IMS technology were divided into three data sets: All (including all compounds), M (without dimeric compounds) and D (without monomeric compounds).

3.3.4. Calibration and optimisation of models with artificial ageing data

Two separate PLS-R models were established based on the VOCs of data sets from NA and AA, respectively (Table 1**A**). The NA-PLS-R and AA-PLS-R models both presented good performance with high R^2 values ($R^2 > 0.9$) and low RMSE values (RMSE < 0.03), demonstrating that strong quantitative relationships were existed between volatile molecules of sweet corn seeds and their vigour in the same ageing method. Then, two variables selection strategies based on a one-way univariate analysis (ANOVA), were used to screen effective variables from the models (bold) with the best performance in each dataset (Fig. 3**A-H**). Generally, volatile metabolites with higher values of variable importance in projection (VIP) indicated the discriminatory potential biomarkers among different seed vigour groups. One-way ANOVA was utilized to further verify the significance of the discriminatory molecules (Cuevas, Moreno-Rojas, Arroyo, Daza, & Ruiz-Moreno, 2016). Therefore, the first strategy was to select metabolites with VIP > 1 and $p < 0.05$ (Fig. 3**A-F**) for building simpler VIP-PLS-R models (Table 1**B**). Compared with the models built based on all compounds, these simpler models using fewer variables also performed well. The automatic scaling (AS) pre-treatment performed best in the simpler models, but not in the full-compound models. This may indicate that the AS is an effective pre-processing method to well equalize the magnitude among the volatile metabolites when the number of volatile metabolites is reduced. Three substances with the highest VIP values in the optimal models were then selected as input data to establish the simplest VIP-PLS-R models for verifying their discrimination ability in NA and AA, respectively (Table 1**C**). VIP-AS-PLS-R model based on D data set ($R^2c = 0.973$, RMSEC = 0.013; $R^2cv = 0.961$, RMSECV = 0.016) and VIP-AS-PLS-R model based on M data set ($R^2c = 0.963$, RMSEC = 0.016; $R^2cv = 0.952$, RMSECV = 0.018) presented the best performance in AA and NA, respectively.

3.3.5. External validation of optimal models with natural ageing data

In order to simulate the routine laboratory operations of seed production plants, PLS-R models were constructed using data collected from artificially aged seeds to predict the vigour of seeds during storage. Moreover, for the sake of the prediction accuracy of PLS-R models, the potential marker compounds in the same data set (M and D) of NA and AA were merged. The merged compounds from AA were used as input data to build PLS-R models. Accordingly, the merged compounds from NA were used as prediction data to external validate the practical performance of the established models (Table 2**B**). The data in these models was pre-processed by AS method. Unfortunately, although the models exhibited good performance on the artificial ageing data (calibration and cross-validation sets), their performance on the natural ageing data (prediction sets) were not ideal ($R^2 < 0.85$). This result can be explained by the fact that the above models were constructed independently by data from each ageing method, which may lead to the neglect of the substances that are strongly correlated between NA and AA but were not critical in modelling when filtering the potential marker compounds.

Therefore, the second strategy was introduced, combining Pearson correlation coefficient (PCC) and VIP, to explore the relationship of VOCs between natural ageing and artificial ageing. 15 molecules with the values of the correlation coefficient above 0.8 ($r > 0.8$) and $p < 0.05$ (one-way ANOVA) were selected (Fig. 4) to establish PCC-PLS-R models

Table 1
Performances of the PLS-R models based on AA data and NA data, respectively, with A) All variables, variables selected by B) VIP > 1, C) top three VIP values and D) r > 0.8 of PCC.

	Ageing methods	Data sets	VOCs selection methods	No.	Pre-treatments	LV	Calibration		Cross-Validation	
							R^2	RMSEC	R^2	RMSECV
A)	AA	All	None	33	None	3	0.935	0.021	0.903	0.026
					AS	**3**	**0.984**	**0.010**	**0.957**	**0.017**
					SNV	3	0.944	0.019	0.914	0.025
		M		25	None	3	0.957	0.017	0.918	0.024
					AS	**3**	**0.978**	**0.012**	**0.936**	**0.021**
					SNV	3	0.954	0.017	0.912	0.025
		D		25	None	7	0.999	0.002	0.914	0.025
					AS	**3**	**0.984**	**0.010**	**0.943**	**0.020**
					SNV	5	0.982	0.011	0.883	0.029
	NA	All		33	None	3	0.951	0.018	0.927	0.022
					AS	5	0.999	0.003	0.938	0.021
					SNV	**3**	**0.965**	**0.015**	**0.942**	**0.020**
		M		25	None	4	0.970	0.014	0.944	0.019
					AS	**2**	**0.975**	**0.013**	**0.958**	**0.017**
					SNV	4	0.974	0.013	0.949	0.019
		D		25	**None**	**4**	**0.981**	**0.011**	**0.959**	**0.017**
					AS	1	0.953	0.018	0.937	0.020
					SNV	4	0.982	0.011	0.953	0.018
B)	AA	All	VIP	19	None	4	0.943	0.019	0.805	0.038
					AS	**2**	**0.984**	**0.010**	**0.965**	**0.015**
					SNV	3	0.895	0.027	0.704	0.045
		M		14	None	4	0.955	0.017	0.872	0.029
					AS	**2**	**0.969**	**0.014**	**0.926**	**0.022**
					SNV	3	0.865	0.030	0.601	0.052
		D		14	None	4	0.931	0.021	0.738	0.045
					AS	**2**	**0.987**	**0.009**	**0.962**	**0.016**
					SNV	3	0.843	0.032	0.631	0.051
	NA	All		5	None	4	0.960	0.016	0.944	0.019
					AS	**2**	**0.977**	**0.012**	**0.959**	**0.017**
					SNV	3	0.871	0.029	0.786	0.038
		M		15	None	2	0.955	0.017	0.942	0.020
					AS	**2**	**0.978**	**0.012**	**0.961**	**0.016**
					SNV	3	0.952	0.018	0.933	0.021
		D		4	None	2	0.719	0.043	0.652	0.048
					AS	**3**	**0.950**	**0.018**	**0.892**	**0.027**
					SNV	2	0.732	0.042	0.673	0.047
C)	AA	All	Top three VIP	3 (i)	AS	2	0.956	0.017	0.942	0.020
		M		3 (ii)	AS	1	0.960	0.016	0.947	0.019
		D		**3 (iii)**	**AS**	**2**	**0.973**	**0.013**	**0.961**	**0.016**
	NA	All		3 (iv)	AS	1	0.960	0.016	0.940	0.020
		M		**3 (v)**	**AS**	**1**	**0.963**	**0.016**	**0.952**	**0.018**
		D		3 (vi)	AS	1	0.929	0.022	0.898	0.026
D)	AA	All	PCC	15	AS	4	0.988	0.009	0.948	0.019
	NA				AS	2	0.973	0.013	0.953	0.018

(i) Benzaldehyde D, 1-butanol D, 1-butanol M.
(ii) Benzaldehyde M, Butyl acetate M, 1-butanol M.
(iii) *n*-nonanal, Benzaldehyde D, 1-butanol D.
(iv) 3-hydroxy-2-butanone, Benzaldehyde M, Ethyl acetate.
(v) Hexanal M, Benzaldehyde M, 1_butanol M.
(vi) 3-hydroxy-2-butanone, Ethyl acetate, hexanal D.
The models shown in bold indicate the best performance among the groups.
Abbreviations: Natural ageing (NA); Artificial ageing (AA); Number of variables (No.); Latent variable (LV); Partial least squares regression (PLS-R); Root mean square error of calibration (RMSEC); Root mean square error of cross-validation (RMSECV); Root mean square error of prediction (RMSEP); Standard normal variate (SNV); Autoscaling (AS); Variable importance in projection (VIP); Pearson correlation coefficient (PCC); Monomer (M); Dimer (D).

(Table 1**D**). Models based on the data of NA and AA both performed well. Subsequently, two approaches were used to further screen the potential marker compounds (Fig. 3 G-H). One approach was to select all molecules with VIP > 1 and r > 0.9. The other approach was to select the molecules that both exist in NA and AA, with VIP > 1 and r > 0.9. The AA data of selected compounds was used as the input data to build PLS-R models (Table 2**C)**. The corresponding NA data as the prediction set was applied to external validate the practicality of the models. These models, based on the compounds considering the correlation coefficient and VIP values, both presented better performance for NA prediction (R^2p > 0.85) compared to the optimal VIP-PLS-R models using the merged - molecules (Table 2**B, R2p** = 0.807). Among them, the model using 4 compounds (Benzaldehyde M, Benzaldehyde D, *n*-Nonanal, 1-butanol M) achieved the best performance in NA data prediction with R^2p of 0.893 and RMSEP of 0.041. Nevertheless, the model using only three molecules (Benzaldehyde M, Benzaldehyde D, 1-butanol M) also predicted well (R^2p = 0.871, RMSEP = 0.054) and was cost-effective. Therefore, benzaldehyde M, benzaldehyde D, 1_butanol M were initially identified as the potential marker compounds.

Furthermore, considering that many traditional laboratories only have GC–MS equipment, that is, dimeric compounds cannot be detected. Meanwhile, we found that the *n*-nonanal presented the highest correlation coefficient value in the molecules of VIP > 1. The benzaldehyde D was then replaced by *n*-nonanal and a new simplest regression model was reconstructed (Table 2**D**). In all cases, the PCC-VIP-PLS-R model in Table 2**D** performed best in NA prediction, with R^2p value of 0.901 and

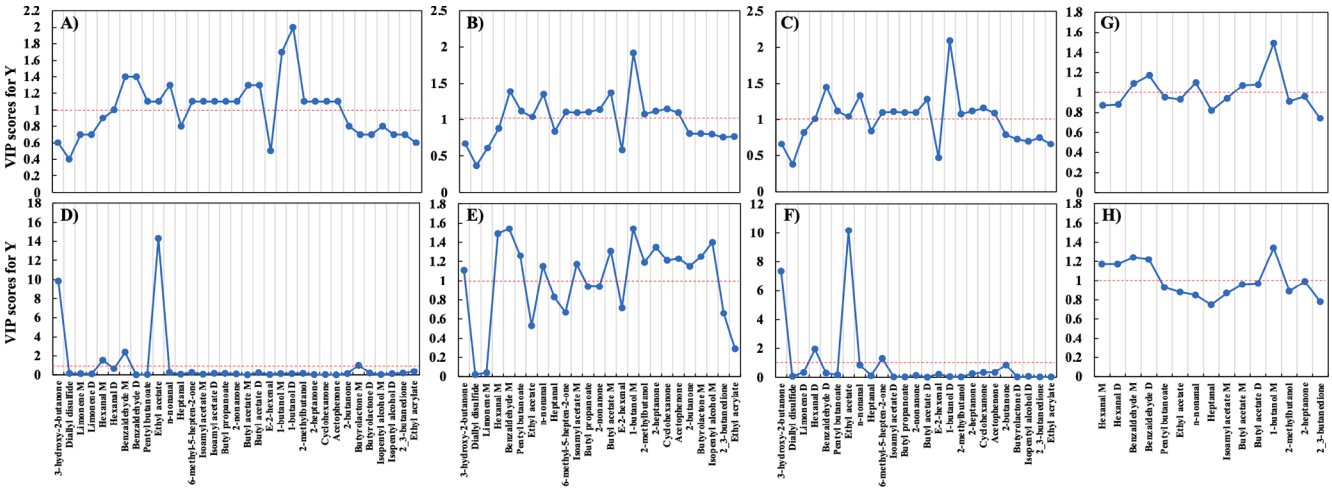

Fig. 3. VIP plots of the NA-ALL-SNV (A), NA-M−AS (B), NA-D-None (C), AA-ALL-AS (D); AA-M−AS (E); AA-D-AS (F), AA-PCC-AS (G) and NA-PCC-AS (H) PLS regression models. VIP: variable importance in projection; PCC: Pearson correlation coefficient; SNV: standard normal variate; AS: autoscaling; PLS: partial least squares regression; NA: natural ageing; AA: artificial ageing; M: monomer; D: dimer. The abscissas of the plots in the same column are consistent. The compound names in plots A-F were linked with the number of variables in (A) models (bold) of table 1. Correspondingly, plots G-H were linked with (D) models of table 1.

Table 2
Performances of the PLS-R models for vigour prediction in naturally aged seeds by A) E-nose data and GC-IMS data with B) VIP, C) and D) PCC-VIP selection methods.

	Techniques	VOC selection methods	No.	Pre-treatments	LV	Calibration (AA)		Cross -Validation		Prediction (NA)	
						R^2	RMSEC	R^2	RMSE CV	R^2	RMSEP
A)	E-nose	None	Ten sensors	None	4	0.577	0.053	0.533	0.055	0.102	0.094
				AS	3	0.597	0.051	0.549	0.054	0.061	0.077
				SNV	5	0.686	0.045	0.615	0.050	0.082	0.081
B)	GC-IMS	VIP	**4 (i)**	**AS**	**3**	**0.963**	**0.016**	**0.927**	**0.022**	**0.807**	**0.075**
			6 (ii)	AS	1	0.907	0.025	0.881	0.028	0.496	0.055
C)	GC-IMS	PCC-VIP	4 (iii)	AS	2	0.955	0.017	0.919	0.023	0.893	0.041
			3 (iv)	AS	2	0.954	0.018	0.922	0.023	0.871	0.054
D)	GC-IMS	PCC-VIP	**3 (v)**	**AS**	**2**	**0.947**	**0.019**	**0.906**	**0.025**	**0.901**	**0.050**

(i) Hexanal M, Benzaldehyde M, 1_butanol M, Butyl acetate M.
(ii) 3-hydroxy-2-butanone, Ethyl acetate, hexanal D, *n*-nonanal, Benzaldehyde D, 1-butanol D.
(iii) Benzaldehyde M, Benzaldehyde D, *n*-nonanal, 1-butanol M.
(iv) Benzaldehyde M, Benzaldehyde D, 1-butanol M.
(v) Benzaldehyde M, *n*-nonanal, 1-butanol M.
The models shown in bold indicate the best performance among the groups.
Abbreviations: Natural ageing (NA); Artificial ageing (AA); Number of variables (No.); Latent variable (LV); Partial least squares regression (PLS-R); Root mean square error of calibration (RMSEC); Root mean square error of cross validation (RMSECV); Root mean square error of prediction (RMSEP); Standard normal variate (SNV); Autoscaling (AS); Variable importance in projection (VIP); Pearson correlation coefficient (PCC); Monomer (M); Dimer (D).

RMSEP of 0.050. Therefore, benzaldehyde M, *n*-nonanal and 1_butanol M were eventually identified as potential key VOC markers, which were able to predict the vigour of sweet corn seeds during storage based on the data collected from artificial ageing method. Moreover, these key markers also existed in the five group of the heatmap (Fig. 2**A**), where compounds changed consistently in NA and AA, demonstrating the reliability of the results from PCC-VIP-PLS-R model.

The previous literature of the VOCs mainly based on the rich-oil seeds, such as soybean and cotton. The study on seeds with high soluble sugar content has not been reported. Hence, it is the first time to explore the relationship between the VOCs in naturally and artificially aged sweet corn seeds, and the VOC markers for stored seed vigour prediction.

Nonanal as a saturated fatty aldehyde is generated owing to the lipid peroxidation and has also been detected in many seeds (Adhikary, Mukherjee, & Barik, 2015; Widjaja, Craske, & Wootton, 1996). Lipid peroxidation could occur through auto-oxidation in the seeds with low moisture content. Sweet corn seeds contain high levels of free fatty acid, such as linoleic acid and linolenic acid (Ratcliff et al., 1993), which can be oxidized into hydroperoxides and further degraded into ketones, aldehydes and alcohols (Sun & Leopold, 1995). In our study, a wide range of aldehydes were detected by GC-IMS (Table S1 and Fig. 2**A**). Among these, hexanal has been widely studied and considered to be a potential marker molecule of lipid peroxidation in seeds during storage (Colville et al, 2012; Mira et al., 2016). Besides, nonanal and hexanal had been proved to inhibit the seed germinability of pigweed, green foxtail and soybean (French & Leather, 1979; Gardner, Dornbos Jr, & Desjardins, 1990). *N*-nonanal and hexanal identified in our experiment were firstly clustered into the same group (Fig. 2**A**), indicating that the variation trends of their content are roughly the same; this is, the amount of *n*-nonanal and hexanal increased in seeds during NA and AA, which were consistent with the results of many studies (Hailstones & Smith, 1989; Mira et al., 2016). However, we found interesting phenomena, which have not been reported before. Compared with hexanal, the correlation of *n*-nonanal in seeds between NA and AA was higher (Fig. 4). Moreover, *n*-nonanal presented a higher variable importance value than hexanal in AA-based model for seed vigour prediction (Fig. 3**C**). These eventually led to *n*-nonanal being selected as a key molecule for predicting the vigour of seeds during NA based on the optimal AA-based model.

Sun and Leopold (1995) investigated the Maillard reaction in soybean seeds during NA and AA and demonstrated that although the content of reducing sugars in soybean seeds was only trace amounts, the

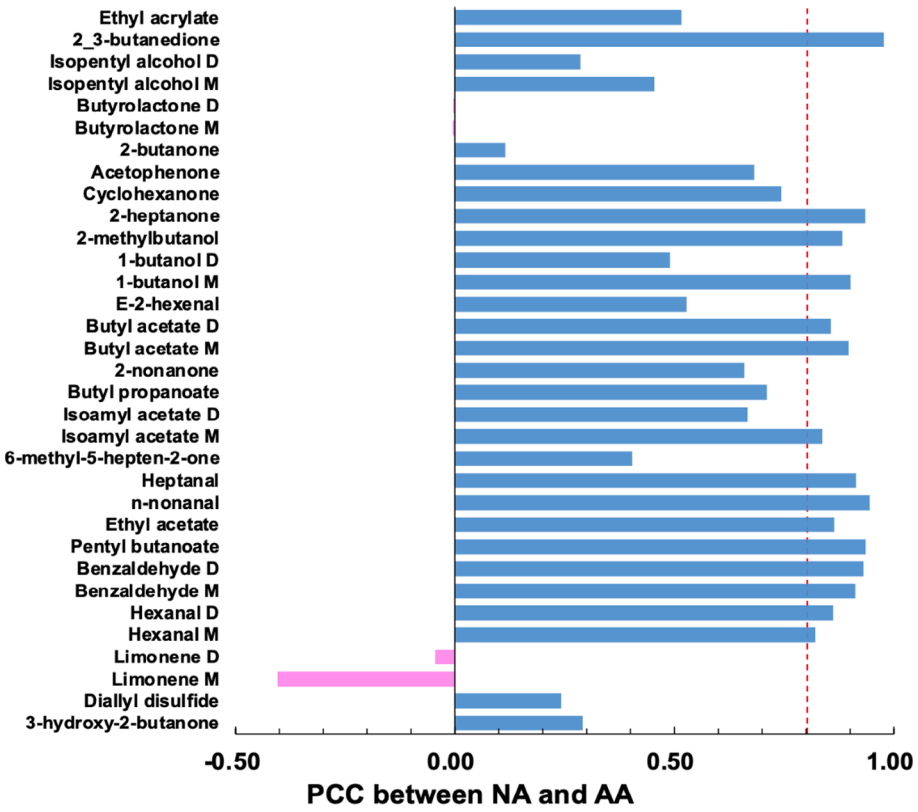

Fig. 4. Pearson correlation coefficient analysis between NA and AA of volatile compounds from sweet corn seeds. NA: natural ageing; AA: artificial ageing; M: monomer; D: dimer.

sugar degradation or hydrolysis in seeds during ageing would increase the formation of reducing sugar and drive the Maillard reaction. Sweet corn seed contains a high level of carbohydrates, such as sucrose, glucose and fructose (Reyes, Varseveld, & Kuhn, 1982). Therefore, in our study, many Maillard products were identified (Table S1**).** Some of them, such as benzaldehyde and 2-heptanone, were clustered into the same group and presented similar changes between NA and AA (Fig. 2**A**). Like *n*-nonanal, benzaldehyde was finally chosen as one of the biomarkers to predict the vigour of seeds during storage.

Notably, among the identified VOCs, aldehydes, alcohols and esters exhibited higher consistency between NA and AA than that of ketones, terpenes and sulphur compounds (Table S1 and Fig. 2**A**). Meanwhile, among the seven aldehydes, six compounds exhibited high correlation coefficient between NA and AA ($r > 0.8$, Fig. 4). Furthermore, two of the three biomarkers also belong to aldehydes (Table 2**D**). Hence, we suggest that aldehydes might be an ideal starting point for seeds with high sugar content, like sweet corn seeds, to study the consistency of metabolic mechanisms and longevity prediction during different ageing processes.

In addition, our results showed that the GC-IMS fingerprint is superior to E-nose in monitoring the vigour of sweet corn seeds (Table 2). This is mainly due to the fact that the E-nose, which consists of an array of metal oxide semiconductor sensors, will evaluate all of the volatile components within a sample. However, according to the VOCs presented by cluster heatmap in Fig. 2**A**, different compounds in the same family, such as 3-hydroxy-2-butanone and 2-heptanone, released from naturally and artificially aged seeds may change differently during natural ageing when compared to artificial ageing. These substances will therefore adversely affect the ability of an AA-based PLS-R model to predict the quality of seeds during NA. Unlike E-nose, GC-IMS combines high separation capacity (GC) and fast detection with high molecular specificity (IMS), and low detection threshold at ppbv or pptv levels (Wang, Chen, & Sun, 2020). These features of the GC-IMS directly improved the practical ability of the AA-based models to predict seed vigour during natural ageing, this will have significant practical applications in the food industry. The VOCs with the same changing trends in NA and AA were mainly alcohols, esters and aldehydes. Hence, in further work, electronic chemical sensors should be developed with high molecular specificity to these compounds, this would therefore offer a portable, fast, cheap and precise detection tool that would specifically monitor predicted quality of stored sweet corn seeds.

Whilst, we have achieved encouraging results for the samples tested, a wider range of seed storage environments, varieties and production methods should also be considered for future applications. These differences may impact the volatile compounds associated with the raw seeds and also impact the development of the volatile profile during ageing.

4. Conclusions

In this study, the VOCs characterisation of naturally and artificially aged seeds was investigated by E-nose and GC-IMS. The combination of VOC fingerprint obtained with GC-IMS, and intelligent data mining approaches were demonstrated to be the strongest approach to monitor and predict the quality of stored sweet corn seeds from data of AA seeds.

Particularly, nineteen of thirty-three identified VOC profiling presented a certain level of consistency in the variation trends between NA and AA based on the hierarchically clustered heatmap. Meanwhile, aldehydes showed the highest consistency between NA and AA. After analysing molecular information by PCA and one-way ANOVA, PLS-R models based on full variables of NA and AA for seed vigour prediction were developed, respectively. To obtained the optimal PLS-R models, effective variables were screened by Strategy 1 (VIP > 1 & $p < 0.05$) and Strategy 2 (VIP > 1 & $r > 0.9$ & $p < 0.05$), respectively. Ultimately, three compounds (benzaldehyde monomer, *n*-nonanal and 1-butanol monomer) selected by Strategy 2 were identified as primary

markers of sweet corn seed ageing and were used as key variables in a methodology of AA-based model (GC-IMS-PCC-VIP-PLS-R) with autoscaling algorithm, which approach was shown to offer the best prediction of loss of seed vigour during NA. E-nose was also evaluated, as an alternative rapid detection technology, but due to its low level of molecular specificity, it was unable to effectively predict the vigour changes of seeds during NA. GC-IMS seems to be a good tool to provide specific fingerprints of volatile changes in seeds during ageing. Moreover, this quick and portable technique can be easily adapted for routine analysis in production facilities and will offer a faster prediction of seed vigour when compared to the traditional germination tests. The application of these findings will, therefore, reduce the sale of low-vigour seeds, which ultimately will enhance food quality for the consumers.

CRediT authorship contribution statement

Tingting Zhang: Investigation, Writing – original draft, Writing - review & editing. **A. Charfedinne:** Writing - review & editing, Supervision. **Ian D. Fisk:** Writing - review & editing, Supervision. **Tong Pan:** Writing - review & editing, Supervision. **Jianhua Wang:** Writing - review & editing. **Ni Yang:** Writing - review & editing, Supervision. **Qun Sun:** Writing - review & editing, Supervision.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by the National Key Research and Development Project of the 13th Five-year Plan (2018YFD0100903).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.foodchem.2021.130760.

References

Adhikary, P., Mukherjee, A., & Barik, A. (2015). Attraction of *Callosobruchus maculatus* (F.)(Coleoptera: Bruchidae) to four varieties of *Lathyrus sativus* L. seed volatiles. *Bulletin of Entomological Research, 105*(2), 187–201.

Colville, L., Bradley, E. L., Lloyd, A. S., Pritchard, H. W., Castle, L., & Kranner, I. (2012). Volatile fingerprints of seeds of four species indicate the involvement of alcoholic fermentation, lipid peroxidation, and Maillard reactions in seed deterioration during ageing and desiccation stress. *Journal of Experimental Botany, 63*(18), 6519–6530.

Coşkun, M. B., Yalçın, İ., & Özarslan, C. (2006). Physical properties of sweet corn seed (*Zea mays* saccharata Sturt.). *Journal of Food Engineering, 74*(4), 523–528.

Cuevas, F. J., Moreno-Rojas, J. M., Arroyo, F., Daza, A., & Ruiz-Moreno, M. J. (2016). Effect of management (organic vs conventional) on volatile profiles of six plum cultivars (*Prunus salicina* Lindl.). A chemometric approach for varietal classification and determination of potential markers. *Food Chemistry, 199*, 479–484.

French, R. C., & Leather, G. R. (1979). Screening of nonanal and related volatile flavor compounds on the germination of 18 species of weed seed. *Journal of Agricultural and Food Chemistry, 27*(4), 828–832.

Gan, H. H., Yan, B., Linforth, R. S. T., & Fisk, I. D. (2016). Development and validation of an APCI-MS/GC–MS approach for the classification and prediction of Cheddar cheese maturity. *Food Chemistry, 190*, 442–447.

Gardner, H. W., Dornbos, D. L., & Desjardins, A. E. (1990). Hexanal, trans-2-hexenal, and trans-2-nonenal inhibit soybean, Glycine max, seed germination. *Journal of Agricultural and Food Chemistry, 38*(6), 1316–1320.

Ghasemi-Varnamkhasti, M., Amiri, Z. S., Tohidi, M., Dowlati, M., Mohtasebi, S. S., Silva, A. C., et al. (2018). Differentiation of cumin seeds using a metal-oxide based gas sensor array in tandem with chemometric tools. *Talanta, 176*, 221–226.

Gu, S., & Wang, J. (2019). In *Rapid detection and classification of fungi contamination in rice grains based on electronic nose with chemometrics* (p. (p. 1):). American Society of Agricultural and Biological Engineers.

Hailstones, M. D., & Smith, M. T. (1989). Thermally-derived volatile aldehydes in relation to seed viability in soybean seeds. *Seed Science and Technology, 17*(3), 649–658.

Ilbi, H., Kavak, S., & Eser, B. (2009). Cool germination test can be an alternative vigour test for maize. *Seed Science and Technology, 37*(2), 516–519.

Lee, P.-C., Taylor, A. G., Zhang, M., & Esashi, Y. (2000). Volatile compounds and accumulation of acetaldehyde-protein adducts in relation to seed quality and storage conditions. *Journal of New Seeds, 2*(1), 59–76.

Li, X., Wang, K.e., Yang, R., Dong, Y., & Lin, S. (2020). Mechanism of aroma compounds changes from sea cucumber peptide powders (SCPPs) under different storage conditions. *Food Research International, 128*, 108757. https://doi.org/10.1016/j.foodres.2019.108757.

Liu, C., Yang, N.i., Yang, Q., Ayed, C., Linforth, R., & Fisk, I. D. (2019). Enhancing Robusta coffee aroma by modifying flavour precursors in the green coffee bean. *Food chemistry, 281*, 8–17.

Mira, S., Hill, L. M., González-Benito, M. E., Ibáñez, M. A., & Walters, C. (2016). Volatile emission in dry seeds as a way to probe chemical reactions during initial asymptomatic deterioration. *Journal of Experimental Botany, 67*(6), 1783–1793.

Muasya, R., Simiyu, J., Muui, C., Rao, N., Dulloo, M., & Gohole, L. (2009). Association of Official Seed Analysts Society of Commercial Seed Technologists. *Seed Technology, 31*(2), 134–143.

Narayana Murthy, U., & Sun, W. Q. (2000). Protein modification by Amadori and Maillard reactions during seed storage: Roles of sugar hydrolysis and lipid peroxidation. *Journal of Experimental Botany, 51*(348), 1221–1228.

Neto, N. M., Custodio, C., & Takaki, M. (2001). Evaluation of naturally and artificially aged seeds of Phaseolus vulgaris L. *Seed Science and Technology, 29*(1), 137–150.

Priestley, D. A., & Leopold, A. C. (1983). Lipid changes during natural aging of soybean seeds. *Physiologia Plantarum, 59*(3), 467–470.

Ratcliff, S. L., Wilson, D. O., Knott, E. A., & Mohan, S. K. (1993). Free Fatty Acids in Shrunken-2 Sweet Corn Seed. *Crop Science, 33*(4), 871–873.

Rajagopal, A., & Sen, M. S. (1992). Studies on acid and alkaline phosphatases in aged rice embryos. *Seed Science and Technology, 20*, 215–222.

Reyes, F. G. R., Varseveld, G. W., & Kuhn, M. C. (1982). Sugar composition and flavor quality of high sugar (shrunken) and normal sweet corn. *Journal of Food Science, 47*(3), 753–755.

Singh, I., Langyan, S., & Yadava, P. (2014). Sweet Corn and Corn-Based Sweeteners. *Sugar Tech, 16*(2), 144–149.

Sun, W. Q., & Leopold, A. C. (1995). The Maillard reaction and oxidative stress during aging of soybean seeds. *Physiologia Plantarum, 94*(1), 94–104.

Walters, C. (1998). Understanding the mechanisms and kinetics of seed aging. *Seed Science Research, 8*(2), 223–244.

Wang, Q., Li, L., Ding, W., Zhang, D., Wang, J., Reed, K., et al. (2019). Adulterant identification in mutton by electronic nose and gas chromatography-mass spectrometer. *Food control, 98*, 431–438.

Wang, S., Chen, H., & Sun, B. (2020). Recent progress in food flavor analysis using gas chromatography–ion mobility spectrometry (GC–IMS). *Food Chemistry, 315*, 126158. https://doi.org/10.1016/j.foodchem.2019.126158.

Wang, Y., Peng, Y., Zhuang, Q., & Zhao, X. (2020). Feasibility analysis of NIR for detecting sweet corn seeds vigor. *Journal of Cereal Science, 93*, 102977. https://doi.org/10.1016/j.jcs.2020.102977.

Widjaja, R., Craske, J. D., & Wootton, M. (1996). Comparative studies on volatile components of non-fragrant and fragrant rices. *Journal of the Science of Food and Agriculture, 70*(2), 151–161.

Xiang, N., Guo, X., Liu, F., Li, Q., Hu, J., & Brennan, C. S. (2017). Effect of light-and dark-germination on the phenolic biosynthesis, phytochemical profiles, and antioxidant activities in sweet corn (*Zea mays* L.) sprouts. *International journal of molecular sciences, 18*(6), 1246.

Xu, H., Wei, Y., Zhu, Y., Lian, L., Xie, H., Cai, Q., et al. (2015). Antisense suppression of *LOX3* gene expression in rice endosperm enhances seed longevity. *Plant Biotechnology Journal, 13*(4), 526–539.

Yan, S., Huang, W., Gao, J., Fu, H., & Liu, J. (2018). Comparative metabolomic analysis of seed metabolites associated with seed storability in rice (*Oryza sativa* L.) during natural aging. *Plant Physiology and Biochemistry, 127*, 590–598.

Yin, G., Xin, X., Fu, S., An, M., Wu, S., Chen, X., et al. (2017). Proteomic and carbonylation profile analysis at the critical node of seed ageing in *Oryza sativa*. *Scientific reports, 7*(1). https://doi.org/10.1038/srep40611.

Ying, X., Liu, W., Hui, G., & Fu, J. (2015). E-nose based rapid prediction of early mouldy grain using probabilistic neural networks. *Bioengineered, 6*(4), 222–226.

Zaragozá, P., Fuentes, A., Fernández-Segovia, I., Vivancos, J.-L., Rizo, A., Ros-Lis, J. V., et al. (2013). Evaluation of sea bream (*Sparus aurata*) shelf life using an optoelectronic nose. *Food Chemistry, 138*(2-3), 1374–1380.

Zhang, M., Liu, Y., Torii, I., Sasaki, H., & Esashi, Y. (1993). Evolution of volatile compounds by seeds during storage periods. *Seed Science and Technology, 21*(2), 359–373.

Zhang, M., Yajima, H., Umezawa, Y., Nakagawa, Y., & Esashi, Y. (1995). GC-MS identification of volatile compounds evolved by dry seeds in relation to storage conditions. *Seed Science and Technology, 23*(1), 59–68.

Zhang, T. T., Fan, S. X., Xiang, Y. Y., Zhang, S. J., Wang, J. H., & Sun, Q. (2020a). *Non-destructive analysis of germination percentage, germination energy and simple vigour index on wheat seeds during storage by Vis/NIR and SWIR hyperspectral imaging* (p. 118488). Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy.

Zhang, X., Dai, Z., Fan, X., Liu, M., Ma, J., Shang, W., et al. (2020b). A study on volatile metabolites screening by HS-SPME-GC-MS and HS-GC-IMS for discrimination and characterization of white and yellowed rice. *Cereal Chemistry, 97*(2), 496–504.

Zhang, W. J., Zhang, P. L., Wang, Y., Qian, C. J., & Du, W. L. (2017). Comparative analysis of the effects of artificial ageing on seed vigour in maize types with inverted giant embryos and normal embryos. Seed Science and Technology, 45(3), 691-707.