



This is an electronic reprint of the original article. This reprint may differ from the original in pagination and typographic detail.

Ruiz-Jimenez, Jose; Raskala, Sanni; Tanskanen, Ville; Aattela, Elisa; Salkinoja-Salonen, Mirja; Hartonen, Kari; Riekkola, Marja Liisa

Evaluation of VOCs from fungal strains, building insulation materials and indoor air by solid phase microextraction arrow, thermal desorption–gas chromatography-mass spectrometry and machine learning approaches

Published in: Environmental Research

DOI: 10.1016/j.envres.2023.115494

Published: 01/05/2023

Document Version Publisher's PDF, also known as Version of record

Published under the following license: CC BY-NC

Please cite the original version:

Ruiz-Jimenez, J., Řaskala, S., Tanskanen, V., Aattela, E., Salkinoja-Salonen, M., Hartonen, K., & Riekkola, M. L. (2023). Evaluation of VOCs from fungal strains, building insulation materials and indoor air by solid phase microextraction arrow, thermal desorption–gas chromatography-mass spectrometry and machine learning approaches. *Environmental Research*, *224*, Article 115494. https://doi.org/10.1016/j.envres.2023.115494

This material is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of the repository collections is not permitted, except that material may be duplicated by you for your research use or educational purposes in electronic or print form. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone who is not an authorised user.

EISEVIER



# **Environmental Research**



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

# Evaluation of VOCs from fungal strains, building insulation materials and indoor air by solid phase microextraction arrow, thermal desorption–gas chromatography-mass spectrometry and machine learning approaches

Jose Ruiz-Jimenez<sup>a,d,\*</sup>, Sanni Raskala<sup>a</sup>, Ville Tanskanen<sup>a</sup>, Elisa Aattela<sup>c</sup>, Mirja Salkinoja-Salonen<sup>a,b</sup>, Kari Hartonen<sup>a,d</sup>, Marja-Liisa Riekkola<sup>a,d</sup>

<sup>a</sup> University of Helsinki, Department of Chemistry, P.O. Box 55, FI-00014, Finland

<sup>b</sup> Aalto University, Department of Electrical Engineering and Automation, P.O. Box 11000, FI-00076, Aalto, Finland

<sup>c</sup> Elisa Aattela Oy, FI-33100, Tampere, Finland

<sup>d</sup> Institute for Atmospheric and Earth System Research / Chemistry, P.O. Box 55, FI-00014, University of Helsinki, Finland

ARTICLE INFO

Handling Editor: Jose L Domingo

Keywords: Indoor air Building insulation material Fungal emissions Solid phase microextraction arrow Machine learning

### ABSTRACT

Solid phase microextraction Arrow and thermal desorption-gas chromatography-mass spectrometry allowed the collection and evaluation of volatile organic compounds (VOCs) emitted by fungal cultures from building insulation materials and in indoor air. Principal component analysis, linear discriminant analysis and supported vector machine were used for visualization and statistical assessment of differences between samples. In addition, a screening tool based on the soft independent modelling of class analogies (SIMCA) was developed for identification of fungal contamination of indoor air. Ten different fungal strains, incubated under ambient and microaerophilic conditions, were analyzed for time period ranging from 5 to 29 days after inoculation resulting in a total of 140 samples. In addition, the effect of additives on the fungal growing media was studied. The total number of compounds and concentration values were used for the evaluation of the results. Clear differences were observed for VOC profiles emitted by different fungal strains by exploiting long chain alcohols (3-octanol, 1-hexanol and 2-octen-1-ol) and sesquiterpenes (farnesene, cuprene). The analysis of glass-wool and cellulose based building insulation materials (3 samples) gave clear differences, mainly for oxygenated compounds (ethyl acetate and hexanal) and benzenoids (benzaldehyde). Moreover, the comparison of indoor air and insulation materials collected from a house with fungal indoor air problems indicated that 42% of the VOCs were found in both samples. The analysis of 52 indoor air samples demonstrated clear differences in their VOC profiles, especially for hydrocarbons, and between control (44 samples) and indoor air problem houses (8 samples). Finally, the SIMCA model enabled to recognize differences between control and fungi contaminated houses with a prediction capacity over 84%.

### 1. Introduction

Recent studies have reported that humans spend 90% of their time indoors (Schweizer et al., 2007). Volatile organic compounds (VOCs) present in indoor air gas phase and particles can be absorbed by lungs affecting the human health or at least the quality of life. (Gallon et al., 2020; Kim et al., 2011). These VOCs can be emitted into the indoor air by a wide variety of sources being mainly building materials, but also due to anthropogenic and biogenic activities. Biogenic compounds are emitted by microbes (Ruiz-Jimenez et al., 2022).

Bacteria, fungi and all microbes in general are ubiquitously present in the indoor air. Some of them are even able to grow on the building material surfaces, using them as nutrients and releasing volatile, semiand non-volatile compounds into the indoor air (Salonen et al., 2007). The volatiles are commonly called as microbial volatile organic compounds (MVOCs) (Korpi et al., 2009). Large variations on the microbial emission profiles are caused by the external environmental conditions such as temperature and relative humidity (Portnoy et al., 2004). Some of these compounds are toxic or hazardous for humans even at tiny concentrations, when the toxicity of many others depends on their

https://doi.org/10.1016/j.envres.2023.115494

Received 24 November 2022; Received in revised form 27 January 2023; Accepted 12 February 2023 Available online 18 February 2023

<sup>\*</sup> Corresponding author. University of Helsinki, Department of Chemistry, P.O. Box 55, FI-00014, Finland. *E-mail address:* jose.ruiz-jimenez@helsinki.fi (J. Ruiz-Jimenez).

<sup>0013-9351/© 2023</sup> The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC license (http://creativecommons.org/licenses/by-nc/4.0/).

### concentrations (Brown et al., 2013).

Individual MVOCs have attracted an interest for the identification of the different species or strains present in the air in the recent years (Korpi et al., 2009). Multiple databases are available in the literature containing detailed information of the MVOCs emitted by a wide variety of microbes (Lemfack et al., 2018). However, it should be emphasized that most of these individual compounds can be also emitted by many other sources, limiting their utilization as markers of microbes (Korpi et al., 2009; Ortíz-Castro et al., 2009). More complex approaches, based on MVOC fingerprints containing multiple compounds, might partially overcome this problem (Vergara-Fernández et al., 2018).

The different techniques and approaches used for indoor air MVOCs sampling and analysis are similar to those used for the analysis of building material emissions. Active sampling techniques based on the use of an external pump or vacuum system to move the air through the collection media, filter or sorbent traps are frequently employed for the collection of MVOCs (Ruiz-Jimenez et al., 2022). Sampling tubes packed with commercial sorbents Tenax TA and Carbograph 5TD, have been widely used due to their wide collection capacity and good sampling efficiency in presence of humidity in the air (Ahmed et al., 2018). However, passive sampling techniques, frequently used for the analysis of emissions from building materials and/or anthropogenic activities, are rarely used for the biogenic samples (Adamová et al., 2019). Some of the MVOCs are emitted at relatively low concentrations. This requires long sampling times to collect amounts high enough for a successful analysis. For this reason, miniaturized air sampling techniques, commonly employed in outdoor air analysis, are potential alternatives due to their clear advantages when compared their performance over that of conventional ones (Lan et al., 2020).

The MVOCs analysis usually involves thermal desorption (TD) before gas chromatographic (GC) analysis and mass spectrometry (MS) for the identification and quantification of the compounds but other detectors should not be discarded (Ruiz-Jimenez et al., 2022). The lack of commercial standard compounds for a reliable identification of all the MVOCs in untargeted analysis requires multiparameter identification approaches. The most common one involves the identification of the compounds by comparison of the GC retention times and mass spectra obtained from the different peaks and those provided by a database (Betancourt et al., 2013; Lee et al., 2015). In addition, more complex approaches, based on high resolution proton transfer reaction mass spectrometry, have also been used for the elucidation of a relatively large number of MVOCs in indoor air samples (Misztal et al., 2018).

The evaluation of the indoor air problems requires the collection and identification of the microbial species present in the houses and/or public premises. Vacuum surface sampling with commercial micro vacuum cassettes allows the collection of fungal spores that can be identified by classical techniques such as visual examination of agar cultures and strains or by more advanced ones based on DNA analysis (Park et al., 2018).

Supervised and non-supervised pattern recognition methods for the visualization and statistical evaluation of analytical data have been also widely exploited in the recent years (Tzanakou, 2017). They include classical statistical algorithms such as principal component analysis (PCA) or cluster analysis (CA) for the visualization of sample differences or machine learning (ML) algorithms for the statistical evaluation of the differences observed. ML includes relatively simple and more sophisticated algorithms, such as linear discriminant analysis (LDA), support vector machines (SVM), and neural networks (NN) (Multia et al., 2020; Liangsupree et al., 2021). The classification capacity of soft independent modelling of class analogies (SIMCA) to identify samples without belonging to the model, or belonging to a single class or overlapping classes should be emphasized (Pomerantsev and Rodionova, 2020). A careful interpretation of the different trends observed by SIMCA, for samples not included in the model development, might be of great interest in the analysis of natural samples.

The main objective of this research was to evaluate VOC emissions

from fungal cultures, insulation building materials and indoor air samples using SPME Arrow for the collection of the VOCs followed by TD-GC-MS analysis. The statistical analysis, based on classical and ML algorithms, enabled the visualization and evaluation of sample differences, the potential discrimination between samples and the potential indoor air quality evaluation. Fungal cultures, building insulation materials and natural indoor air samples were used as samples. The effect of important variables, such as the growing media, the incubation atmosphere and the time after inoculation on the fungal VOC emission profiles, was clarified. In addition, the potential use of these profiles for the identification of the different strains was also evaluated. Variations in VOC profiles from control houses and from those having contaminated indoor air were also assessed. Furthermore, potential differences on the emissions, based on building insulation material composition, were elucidated and compared with the emissions from fungi contaminated houses. Finally, a SIMCA based screening tool for the recognition differences between control and contaminated houses with fungi was developed.

### 2. Experimental

### 2.1. Instruments and apparatus

A PAL RTC auto-sampler controlled using PAL Cycle Composer software (CTC analytics AG, Zwingen Switzerland) was used for the collection and analysis of VOCs emitted from fungal cultures and building insulation materials. The combination of software and hardware allows the automated performance of all the SPME Arrow steps, avoiding all the contact between the analyst and the potential hazard or toxic compounds emitted by the samples.

A 6890 N gas chromatograph (Agilent Technologies, Palo Alto, USA) coupled to an Agilent 5975C quadrupole mass spectrometer was used for the separation and detection of the VOCs. The chromatograph was furnished with a GL Sciences (Eindhoven, The Netherlands) InertCap for Amines (30 m  $\times$  0.25 mm i. d., no information was provided for the film thickness) fused silica column coupled with a 3-m deactivated retention gap (i.d. 0.53 mm, Agilent).

Eight commercial SPME Arrows (CTC analytics), four coated with polydimethylsiloxane-divinylbenzene (PDMS-DVB) and four with carbon wide range (Carbon WR)-PDMS, were used for VOCs sampling, both in-vitro and natural samples studies.

#### 2.2. Samples

Twenty-seven 20 mL glass vials were filled with 5 mL of malt agar from NutriSelect®Plus (Merck KGaA, Darmstadt, Germany). Boric acid and borax at 2 mg mL-1 from Sigma-Aldrich (Darmstadt, Germany) were added to two vials to evaluate the potential effect of common antifungal additives on the fungal metabolism. The vials were seeded, at least in duplicate, with a selection of fungi commonly found in houses with indoor air problems. The list of fungi and the growing media can be found in Table S1. After inoculation, the vials partially closed with a PTFE/silicone septum screw-cap (both from Phenomenex, Torrance, California, USA) were divided into two different incubators, ambient and microaerophilic atmospheres, placed under a proper ventilation system. The latter atmosphere was achieved using an Oxoid CO<sub>2</sub> Gen Sachet (Thermo Fisher Scientific, Massachusetts, USA). The microaerophilic bag was replaced with a new one weekly to ensure the proper CO<sub>2</sub> concentration on the incubator atmosphere. The fungal cultures were split into both incubators, at least one vial from each fungal culture was placed in each of them. In addition, a vial containing 5 mL of malt agar but non-seeded were introduced into the incubators to check the substrate emissions and potential cross contamination. Cultures were incubated at room temperature for 5 days before the first analysis. Finally, the vials were closed before analysis and reopened once returned to the incubator. All materials and solutions were sterilized

before use to avoid undesirable contamination.

Three different building insulation materials, two glass wool- and one cellulose-based, were collected from a house, located in southern Finland, with indoor air quality problems. These building materials were collected during summer 2019 from three different rooms of the house right away after gypsum board removal. The samples were stored at -20 °C until analysis. Approximately 0.2 g of the building insulation materials were accurately weighted and placed into glass vials, similar to those previously described for the preparation of the fungal culture inoculums. The total volume of the sample in the vial was under one third of the vial total volume. Empty vials, previously sterilized, were used as blank samples to remove the contribution of the laboratory indoor air on the emission results.

Indoor air VOCs were collected from 24 houses at southern and eastern Finland regions during summer 2019 and summer/autumn 2020. Sampling places include private houses/apartments and public premises used as control samples (22 houses/apartments, 44 samples) and others with indoor air problems reported (3 houses, 8 samples). Houses with indoor air problems were defined as those that produce health problems (diseases or symptoms) in their inhabitants. These problems appeared repeatedly when the person goes into the house, and they were harmless or disappeared when the person was away from the house. It should be considered that these problems might be caused by the exposure of the inhabitants to multiple chemical substances at very low concentration or the presence in the house of fungi, emitting toxic compounds to the indoor air. The later were frequently confirmed using other test such as trained dogs for fungal detection. On the other side, the term reference houses include those used regularly by their inhabitants, for a period of at least 1 year, without any health problems associated to the use of the house inhabitation.

Indoor air samples were collected at least from two different rooms in each sampling place. Blank correction was achieved by analysis of SPME Arrow, previously conditioned, located in the sampling place but not exposed to the indoor air.

### 2.3. Sampling procedures

In all the cases, the SPME Arrow sorbents were conditioned before sampling for 10 min at 250  $^{\circ}$ C using the auto-sampler conditioning module. In addition, no internal standard was used to avoid potential contamination of the fungal cultures with other bacteria and fungi possibly present in the indoor air.

In the case of the fungal cultures, the SPME Arrow systems were exposed to the headspace of the vial at 30 °C for 30 min. Fungal cultures were analyzed in triplicate twice a week for a 4 weeks period to evaluate the effect of the incubation time after seeding on the fungal VOC emissions (5, 8, 12, 15, 19, 22, 26 and 29 days after inoculation). The caps of the vials were replaced, under a sterile atmosphere, from all the vials after analysis, thereby preventing from leaks and contamination. The building insulation materials were analyzed also in triplicate at 30 °C using the same method.

In the case of the indoor air samples, preconditioned SPME Arrows were stored at -20 °C in closed plastic zip bags until used at home by non-trained analysts following detailed instructions. The SPME Arrow samplers were exposed to the indoor air for a period ranged between 15 and 120 min depending on the samples, with the exception of those used as blanks. Storage time between preconditioning, collection and analysis was less than 48 h.

### 2.4. Gas chromatography mass spectrometry analysis

In all the cases, the adsorbed compounds on the SPME Arrow were injected into the GC by thermal desorption at 250 °C for 1 min using the chromatograph inlet operated in splitless mode. The oven temperature program was the following: the initial temperature (40 °C) was kept for 2 min and then ramped up to 250 °C at 20 °C min<sup>-1</sup> and held there for 5

min. Total run time was 17.5 min. Transfer line, ion source and quadrupole temperatures were 250 °C, 230 °C and 150 °C, respectively. Electron impact ion source was operated at 70 eV and the scan range was 15–350 m/z. Helium (99.996%, AGA, Espoo, Finland) was used as a carrier gas with flow rate of 1.2 mL min<sup>-1</sup>.

### 2.5. Data processing and statistical analysis

The procedure used for chromatogram alignment, peak detection and integration; extraction of mass spectra and tentative identification of the aligned peaks was carried out using the methodology developed in our previous research. Mzmine2 (version 2.53) was used for the development of all the steps of data processing, with exception of the peak identification. Experimental mass spectra and retention index values were manually compared with those provided by NIST 20 database for the tentative identification of the compounds (Lan et al., 2021; Pusfitasari et al., 2022).

An R 3.5.1 software was used for the development of the statistical analysis. The approach used for the quantitation/semiquantitation of the VOCs was similar to that developed in previous research with some modifications (Pusfitasari et al., 2022). Model compounds were classified by CA into three different groups, according to their mass spectra similarities and differences. These groups were used for both the development of partial least squares regression (PLSR) equations for the different coating materials as described in our previous publication (Kopperi et al., 2013) and the development of LDA models for the classification of the identified VOCs into these groups.

The visualization of the potential differences between the VOCs profiles obtained from the studies and the identification of the key compounds to explain these differences was achieved by PCA, a nonsupervised pattern recognition approach. These differences were confirmed by the use of basic and more advanced supervised pattern recognition techniques such as linear LDA and SVM. These techniques were also used for the differentiation of the different fungal strains used in the experiments based on their VOC emissions and the identification of houses and premises with indoor air problems (Liangsupree et al., 2021). In all the cases, training and validation sets containing samples randomly selected were established for the development and validation of the models. These groups of samples contain 75% and 25% of the samples under study for training and validation, respectively. The selection of the most appropriate models for the evaluation of the samples was based on the number of samples available for the calculations. Finally, the use of VOC profiles, specific to certain fungi, as potential biomarkers for the identification of houses and public premises contaminated with fungi were evaluated by SIMCA.

# 3. Results and discussion

Volatile organic compounds emitted by 10 different fungal strains were evaluated. Similarities and/or differences between the fungal cultures and additional studies for the detailed evaluation of the incubation time after inoculation, the atmosphere and growing media will be discussed. Machine learning models, developed for the discrimination between fungal strains, will be described. In addition, differences in the indoor air VOC profiles from control and houses/public premises with indoor air problems will be elucidated. The VOC emissions from insulation materials and indoor air samples collected from houses with indoor air problems will be also evaluated. Finally, VOC fungal emission profiles were used as data input for the development of screening tools for the discrimination between clean and fungal contaminated houses.

Unfortunately, the use of different fungal strains, growing substrate, incubation conditions, experimental setup and analysis time after inoculation hinder the quantitative comparison of the results achieved in this research with those provided by previous studies in the literature. However, microbial VOC 3.0 database allowed the comparison of the VOC identified in this research with those found in previous studies

using the same fungal strains (Lemfack et al., 2018).

3.1. Evaluation of the volatile organic compound emissions from fungal strains

### 3.1.1. Effect of the incubation atmosphere

Ten fungal strains were seeded in duplicate in vials partially filled with malt agar. These vials were divided into two different sets, containing one replicate from each strain, and subsequently incubated in two different  $CO_2$  concentrations (ambient and microaerophilic). The caps of the vials were partially opened to allow the equilibration between the gas phase of the vials and the incubator. Blank samples were analyzed to subtracts potential cross contamination. The VOCs emitted by the fungi to the headspace of the vials were analyzed for a period ranged between 5 and 29 days after inoculation. The total number of analyzed samples was 140. The results of the study were evaluated in terms of number of VOCs emitted and their concentrations.

As can be seen in Fig. 1 A, the number of VOCs emitted by the fungal cultures incubated under ambient conditions is larger in comparison with those incubated under microaerophilic ones, with some exceptions. Larger number of VOCs were emitted under microaerophilic atmosphere in comparison with the ambient one for *HJS11* and *HKIP12*. In addition, similar number of compounds was emitted by *Chaetomium globosum ABCD* in both atmospheres.

The evaluation based on the sum of the VOC concentrations, calculated for the different groups of compounds, gave similar trends (Fig. 1B and C) to that previously described for the number of compounds. The VOC concentrations were larger for the fungal cultures incubated under an ambient atmosphere than the microaerophilic one (0 vs 0 results were not considered), with some exceptions. Larger concentrations were found in the VOC emitted from microaerophilic cultures than in the ambient ones for *HKIP12* (hydrocarbons, oxygenated compounds and compounds with a lipid-like substructure), *HSJ11* (oxygenated compounds and compounds with a lipid-like substructure), Aspergillus versicolor 9775/K1 and Aspergillus westerdijkiae PP2 (hydrocarbons and oxygenated compounds), *Chaetomium globosum ABCD* (compounds with a lipid-like substructure); and *Aspergillus niger Hambi 495* (hydrocarbons).

Principal component analysis was used for the visualization of the differences between VOC emissions provided by fungal cultures incubated under different oxygen concentrations. Individual PCA models were developed for the different fungal strains under study. The detailed evaluation of the score plots proved clear differences between the cultures incubated under ambient and microaerophilic conditions (Fig. S1 A). In addition, the analysis of the loading plots revealed that certain compound groups, including oxygenated and heterocyclic compounds, benzenoids, hydrocarbons and compounds with a lipid like substructure, can explain differences in the VOC emissions. Finally, clear differences for the concentration values were observed in multiple fungal strains for several individual compounds, such as acetone, 3-methyl-1-butanol, 2methoxy-furan, 5.7-octadien-4-one, 1-hexanol, 3-octanone, 3-octanol and  $\alpha$ -farnese (Fig. S1 B). All these compounds, with the exception of 2methoxy-furan and 5,7-octadien-4-one, were already reported as VOCs emitted by fungi. However, other compounds with similar structures such as 2,6-dimethyl-5,7-octadien-4-one or furan derivatives have been frequently reported in the literature (Lemfack et al., 2018).

These differences were confirmed by the LDA, a simple supervised pattern recognition approach. The selection of this algorithm was based on the limited number of samples under study, ranging from 8 for *Chaetomium cochliodes OT7* to 19 for *Chaetomium globosum ABCD*. Furthermore, no LDA model was developed for *HJS11* due to its very small number of samples. The prediction accuracy, for samples not included in the model development, ranged from 64.7% for *Aspergillus niger Hambi* 495 to 100% for *Aspergillus calidoustus MH34*.



Fig. 1. Evaluation of the volatile organic compound emissions from fungal cultures as a function of A) the number of compounds and B) the sum of the average concentration values for fungal cultures incubated under ambient (RSD values were ranged between 1.3 and 13.1% for oxygenated compounds from *Chaetomium globosum ABCD* (sampling day 1) and benzenoids from *Aspergillus calidostus MH34* (sampling day 3) respectively) and C) microaerophilic atmospheres (RSD values were ranged between 2.3 and 12.7% for oxygenated compounds from *HJS11* (sampling day 1) and lipids from *Aspergillus versicolor GAS/226* (sampling day 19) respectively). Samples were analyzed in triplicate. Number of compounds and concentration values were blank subtracted. Label lipids includes compounds with a lipid like substructure.

|  | Hydrocarbons -           | 1  |     |      |      |       |      |    |     |         |
|--|--------------------------|--|-----|------|------|-------|------|----|-----|---------|
| Aspergillus versicolor 9775/K1   | Benzenoids -             | 1  |     |      |      |       |      |    |     |         |
|  | Lipids -                 |  |     |      |      |       |      |    |     |         |
| (Malt agar)  | Heterocyclic compounds - | 1  |     |      |      |       |      |    |     |         |
|  | Oxygenated compounds -   | 1  |     |      |      |       |      |    |     |         |
| HKIP12   | Benzenoids               | 3  |     |      |      |       |      |    |     |         |
| 111/11/12  | Lipids -                 | 3  |     |      |      |       |      |    |     |         |
| (Maltagar)   | Heterocyclic compounds - | 1  |     |      |      |       |      |    |     |         |
| (  | Oxygenated compounds -   | 1  |     |      |      |       |      |    |     |         |
|  | Hydrocarbons -           | 3  |     |      |      |       |      |    |     |         |
| HJS11  | Benzenoids -             | 3  |     |      |      |       |      |    |     |         |
|  | Lipids -                 | 1  |     |      |      |       |      |    |     |         |
| (Malt agar)  | Heterocyclic compounds - | 1  | _   |      |      |       |      |    |     |         |
| 27. www.53.  | Oxygenated compounds -   | 1  |     |      |      |       |      |    |     |         |
| Dessileren estisti Mio   | Benzenoids -             | 3  |     |      |      |       |      |    |     |         |
| Paecilomyces variotii Mo9  | Lipids -                 | 1  |     |      |      |       |      |    |     |         |
| (Maltagar)   | Heterocyclic compounds - | 1  |     |      |      |       |      |    |     |         |
| (marcagar)   | Oxygenated compounds -   | 1  |     |      |      |       |      |    |     |         |
| and the second s | Hydrocarbons -           | 3  |     |      |      |       |      |    |     |         |
| Chaetomium globosum ABCD   | Benzenoids -             | 3  |     |      |      |       |      |    |     |         |
| 3  | Lipids -                 | 1  |     |      |      |       |      |    |     |         |
| (Malt agar)  | Heterocyclic compounds - |  |     |      |      |       |      |    |     |         |
| Selfer of Sec  | Oxygenated compounds -   |  |     |      |      |       |      |    |     |         |
| Asservillus solideustus MUDA   | Banzanoida               | 1  |     |      |      |       |      |    |     |         |
| Aspergilius calidoustus MH34   | Lipide -                 |  | _   |      |      |       |      |    |     |         |
| <b>198</b> 14  | Heterocyclic compounds - |  |     |      |      |       |      |    |     | -       |
| (Malt agar)  | Oxygenated compounds -   | 1  |     |      |      |       |      |    |     |         |
|  | Hydrocarbons -           | -  | _   |      |      | _     |      |    |     | _       |
| Asperaillus westerdiikige PD2  | Benzenoids -             | 1  |     |      |      |       |      |    |     |         |
| Aspergillus westeruijkide FFZ  | Lipids -                 | -  |     |      |      |       |      |    |     |         |
| (Malt agar)  | Heterocyclic compounds - |  |     |      |      |       |      |    |     |         |
|  | Oxygenated compounds -   | 1  |     |      |      |       |      |    |     |         |
|  | Hydrocarbons -           | 1  |     |      |      |       |      |    |     |         |
| Chaetomium cochliodes OT7  | Benzenolds -             | 1  |     |      |      |       |      |    |     |         |
| (Maltagar)   | Heterocyclic compounds   | 3  |     |      |      |       |      |    |     |         |
| (mair agar)  | Oxygenated compounds -   | 3  |     |      | 1    |       |      |    |     |         |
| in the second se | Hydrocarbons -           | 1  |     | _    |      |       |      |    |     | _       |
| Asperaillus versicolor GAS/226   | Benzenoids -             | 1  |     |      |      |       |      |    |     |         |
| Aspergillus versicolor 0A3/220   | Lipids -                 | 1  |     |      |      |       |      |    |     | _       |
| (Malt agar)  | Heterocyclic compounds - | 1  |     |      |      |       |      |    |     |         |
| (  | Oxygenated compounds -   | 3  |     |      |      |       |      |    |     |         |
|  | Hydrocarbons -           |  |     |      |      |       |      |    |     |         |
| Aspergillus niger Hambi 495  | Benzenoids -             | 1  |     |      |      |       |      |    |     |         |
|  | Lipids -                 | 1  |     |      |      |       |      |    |     | _       |
| (Malt agar)  | Ovvgenated compounds -   | 1  |     | -    |      |       |      |    |     |         |
|  | Hydrocarbone             | 3  |     |      |      |       |      |    |     |         |
| Asperaillus piger Hambi 405  | Benzenoids -             | 1  |     |      |      |       |      |    |     |         |
| Aspergillus niger nambi 495  | Lipids -                 | 1  |     |      |      |       |      |    |     |         |
| (Mark  | Heterocyclic compounds - | 1  |     |      |      |       |      |    |     |         |
| (Mait agar + Borax)  | Oxygenated compounds -   | 1  |     |      |      | (     |      |    | [[] |         |
|  |                          | 5-   | 8-  | 2-   | 5-   | - 6   | 2-   | -9 | -6  | e       |
|  |                          |  |     | -    | -    | -     | 2    | 2  | 2   | rag     |
|  |                          |  |     |      | Sam  | plin  | ng d | ay |     | Ne      |
|  |                          | $\Sigma$ Concentration (up m <sup>-3</sup> ) |     |      |      |       |      |    |     |         |
|  |                          | -  | 200 | ncer | mail | λη (μ | y m  | )  |     | 1.1.1.1 |
|  |                          |  |     |      |      |       |      |    |     |         |
|  | OE+0                     | 3E+2   |     | 58   | +2   |       | 8E+2 |    | 1   | E+3 2   |

Fig. 1. (continued).

## 3.1.2. Effect of the incubation time after inoculation

The detailed analysis of the VOCs emitted by the fungal cultures during the time period from 5 to 29 days after inoculation allowed the evaluation of this parameter on the fungal emissions in terms of number of compounds and concentrations.

In the first case, the number of emitted compounds decreased with the incubation time after the first analysis (5 days from inoculation). As can be seen in Fig. 1A, some VOCs emitted by fungi were still observed at least 26 days after inoculation in the case of *Aspergillus niger Hambi 495*, *Aspergillus calidoustus MH34*, *Aspergillus westerdijkiae PP2*, *Paecilomyces variotii Mö9* and *Aspergillus versicolor 9775*. However, no–VOC emissions were seen after 19 days after inoculation for the rest of the fungal cultures.

In the second case, the detailed evaluation of the concentration values (Fig. 1 B and C) enabled the identification of different trends as a function of the type of compounds and the growing atmosphere, ambient or microaerophilic. Concentration values decreased with the incubation time after the first analysis (5 days from inoculation) for all the fungal cultures that grew under microaerophilic conditions. A different trend, maxima concentration values were achieved between 8 and 22 days after inoculation and then decrease, was observed for oxygenated compounds (*Aspergillus niger Hambi 495, Chaetomium cochliodes OT7, Aspergillus westerdijkiae PP2, Aspergillus calidoustus MH34, HJS11, HKIP12 and Aspergillus versicolor 9775/K1*); and compounds with a lipid-like substructure (*Aspergillus versicolor GAS/226 and Aspergillus versicolor 9775/K1*).

| Benzenoids<br>Lipids<br>Heterocyclic compounds<br>Oxygenated compounds<br>Hydrocarbons<br>Benzenoids<br>Lipids<br>Heterocyclic compounds<br>Oxygenated compounds<br>Hydrocarbons<br>Benzenoids<br>Lipids<br>Heterocyclic compounds<br>Oxygenated compounds |  |  |  |  |  |  |  |
|--|--|--|--|--|--|--|--|
| Lipids<br>Heterocyclic compounds<br>Oxygenated compounds<br>Hydrocarbons<br>Benzenoids<br>Lipids<br>Heterocyclic compounds<br>Oxygenated compounds<br>Hydrocarbons<br>Benzenoids<br>Lipids<br>Heterocyclic compounds<br>Oxygenated compounds               |  |  |  |  |  |  |  |
| Heterocyclic compounds<br>Oxygenated compounds<br>Hydrocarbons<br>Benzenoids<br>Lipids<br>Heterocyclic compounds<br>Oxygenated compounds<br>Hydrocarbons<br>Benzenoids<br>Lipids<br>Heterocyclic compounds<br>Oxygenated compounds                         |  |  |  |  |  |  |  |
| Oxygenated compounds<br>Hydrocarbons<br>Benzenoids<br>Lipids<br>Heterocyclic compounds<br>Oxygenated compounds<br>Hydrocarbons<br>Benzenoids<br>Lipids<br>Heterocyclic compounds<br>Oxygenated compounds   |  |  |  |  |  |  |  |
| Hydrocarbons<br>Benzenoids<br>Lipids<br>Heterocyclic compounds<br>Oxygenated compounds<br>Hydrocarbons<br>Benzenoids<br>Lipids<br>Heterocyclic compounds<br>Oxygenated compounds   |  |  |  |  |  |  |  |
| Benzenoids<br>Lipids<br>Heterocyclic compounds<br>Oxygenated compounds<br>Hydrocarbons<br>Benzenoids<br>Lipids<br>Heterocyclic compounds<br>Oxygenated compounds   |  |  |  |  |  |  |  |
| Lipids -<br>Lipids -<br>Heterocyclic compounds<br>Oxygenated compounds -<br>Hydrocarbons<br>Benzenoids<br>Lipids -<br>Lipids -<br>Heterocyclic compounds<br>Oxygenated compounds -   |  |  |  |  |  |  |  |
| Heterocyclic compounds<br>Oxygenated compounds<br>Hydrocarbons<br>Benzenoids<br>Lipids<br>Heterocyclic compounds<br>Oxygenated compounds   |  |  |  |  |  |  |  |
| Oxygenated compounds<br>Hydrocarbons<br>Benzenoids<br>Lipids<br>Heterocyclic compounds<br>Oxygenated compounds   |  |  |  |  |  |  |  |
| Hydrocarbons<br>Benzenoids<br>Lipids<br>Heterocyclic compounds<br>Oxygenated compounds   |  |  |  |  |  |  |  |
| Benzenoids<br>Enzenoids<br>Lipids<br>Heterocyclic compounds<br>Oxygenated compounds  |  |  |  |  |  |  |  |
| Lipids -<br>Lipids -<br>Heterocyclic compounds -<br>Oxygenated compounds -   |  |  |  |  |  |  |  |
| Lipids –<br>Heterocyclic compounds –<br>Oxygenated compounds –   |  |  |  |  |  |  |  |
| Anticologic Compounds -<br>Oxygenated compounds -  |  |  |  |  |  |  |  |
| Oxygenated compounds –   |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
| Hydrocarbons -   |  |  |  |  |  |  |  |
| Benzenoids -   |  |  |  |  |  |  |  |
| Lipids –   |  |  |  |  |  |  |  |
| Heterocyclic compounds -   |  |  |  |  |  |  |  |
| Oxygenated compounds   |  |  |  |  |  |  |  |
| Hydrocarbons -   |  |  |  |  |  |  |  |
| Benzenoids   |  |  |  |  |  |  |  |
| Lipide   |  |  |  |  |  |  |  |
| Lipius 7   |  |  |  |  |  |  |  |
| Ownershall compounds -   | _  |  |  |  |  |  |  |
| Oxygenated compounds -   |  |  |  |  |  |  |  |
| Hydrocarbons -   |  |  |  |  |  |  |  |
| Benzenoids –   |  |  |  |  |  |  |  |
| Lipids –   |  |  |  |  |  |  |  |
| Heterocyclic compounds -   |  |  |  |  |  |  |  |
| Oxygenated compounds -   |  |  |  |  |  |  |  |
| Hydrocarbons -   |  |  |  |  |  |  |  |
| Benzenoids -   |  |  |  |  |  |  |  |
| Lipids -   |  |  |  |  |  |  |  |
| Heterocyclic compounds   |  |  |  |  |  |  |  |
| Oxygenated compounds   | _  | 8  |  |  |  |  |  |
| Oxygenated compounds -   |  |  |  |  |  |  |  |
| Barranoida 1   |  |  |  |  |  |  |  |
| Benzenoids -   |  |  |  |  |  |  |  |
| Lipids -   |  |  |  |  |  |  |  |
| Heterocyclic compounds -   |  |  |  |  |  |  |  |
| Oxygenated compounds –   |  |  |  |  |  |  |  |
| Hydrocarbons -   |  |  |  |  |  |  |  |
| Benzenoids -   |  |  |  |  |  |  |  |
| Lipids -   |  |  |  |  |  |  |  |
| Heterocyclic compounds -   |  |  |  |  |  |  |  |
| Oxygenated compounds   |  |  |  |  |  |  |  |
| Hydrocarbons -   |  |  |  |  |  |  |  |
| Renzenoide   |  |  |  |  |  |  |  |
| Lipida 1   |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
| Peterocyclic compounds -   |  |  |  |  |  |  |  |
| Oxygenated compounds -   |  |  |  |  |  |  |  |
| Hydrocarbons -   |  |  |  |  |  |  |  |
| Benzenoids -   |  |  |  |  |  |  |  |
| Lipids –   |  |  |  |  |  |  |  |
| Heterocyclic compounds -   |  |  |  |  |  |  |  |
| Oxygenated compounds -   |  |  |  |  |  |  |  |
|  | 1 1  |  |  | -  |  | -  | -  |
|  | 4) 00  | 12   | 10   | 16   | 26   | 29   | an   |
|  |  | S  | amr  | lina   | dav  |  | E.C.   |
|  |  | <u> </u>   |  |  | aug  |  | d.   |
|  |  |  |  |  |  |  | 2  |
|  | 5.0  |  | tratics  | . (  | -31  |  | AV   |
|  | ΣС   | oncen  | tratior  | ι (μg r  | n <sup>-3</sup> )  |  | Av   |
|  | ΣС   | oncen  | tratior  | ם (µg r  | n <sup>-3</sup> )  |  | Av   |
|  | Lipids -<br>Heterocyclic compounds -<br>Oxygenated compounds -<br>Lipids -<br>Lipids -<br>Heterocyclic compounds -<br>Cygenated compounds -<br>Lipids -<br>Hydrocarbons -<br>Benzenoids -<br>Lipids -<br>Heterocyclic compounds -<br>Cygenated compounds -<br>Lipids -<br>Hydrocarbons -<br>Benzenoids -<br>Lipids -<br>Lipids -<br>Heterocyclic compounds -<br>Cygenated compounds -<br>Lipids -<br>Lipids -<br>Lipids -<br>Lipids -<br>Lipids -<br>Lipids -<br>Heterocyclic compounds -<br>Cygenated compounds -<br>Lipids -<br>Lipids -<br>Lipids -<br>Hydrocarbons -<br>Benzenoids -<br>Lipids -<br>Lipids -<br>Hydrocarbons -<br>Benzenoids - | Lipids -<br>Heterocyclic compounds -<br>Hydrocarbons -<br>Benzenoids -<br>Lipids -<br>Heterocyclic compounds -<br>Oxygenated compounds -<br>Hydrocarbons -<br>Benzenoids -<br>Lipids -<br>Heterocyclic compounds -<br>Oxygenated compounds -<br>Hydrocarbons -<br>Benzenoids -<br>Lipids -<br>Heterocyclic compounds -<br>Oxygenated compounds -<br>Oxygenated compounds -<br>Oxygenated compounds -<br>Lipids -<br>Heterocyclic compounds -<br>Lipids -<br>Hydrocarbons -<br>Benzenoids -<br>Lipids -<br>Heterocyclic compounds -<br>Oxygenated compounds -<br>Dipids -<br>Heterocyclic compounds -<br>Dipids -<br>Hydrocarbons -<br>Benzenoids -<br>Lipids -<br>Hydrocarbons -<br>Benzenoids -<br>Lipids -<br>Heterocyclic compounds -<br>Oxygenated compounds -<br>Dipids -<br>Heterocyclic compounds -<br>Dipids -<br>Hydrocarbons -<br>Benzenoids -<br>Lipids -<br>Heterocyclic compounds -<br>Dipids -<br>Hydrocarbons -<br>Benzenoids -<br>Lipids - | Lipids -<br>Heterocyclic compounds -<br>Benzenoids -<br>Lipids -<br>Hydrocarbons -<br>Benzenoids -<br>Lipids -<br>Heterocyclic compounds -<br>Oxygenated compounds -<br>Lipids -<br>Heterocyclic compounds -<br>Lipids -<br>Heterocyclic compounds -<br>Lipids -<br>Hydrocarbons -<br>Benzenoids -<br>Lipids -<br>Hydrocarbons -<br>Benzenoids -<br>Lipids -<br>Heterocyclic compounds -<br>Oxygenated compounds -<br>Lipids -<br>Heterocyclic compounds -<br>Dygenated compounds -<br>Lipids -<br>Hydrocarbons -<br>Benzenoids -<br>Hydrocarbons -<br>Benzenoids -<br>Benzenoids -<br>Hydrocarbons -<br>Benzenoids -<br>Hydrocarbons -<br>Benzenoids -<br>H | Lipids -<br>Heterocyclic compounds -<br>Benzenoids -<br>Lipids -<br>Heterocyclic compounds -<br>Benzenoids -<br>Lipids -<br>Hydrocarbons -<br>Benzenoids -<br>Be | Lipids -<br>Heterocyclic compounds -<br>Benzenoids -<br>Lipids -<br>Heterocyclic compounds -<br>Oxygenated compounds -<br>Hydrocarbons -<br>Benzenoids -<br>Lipids -<br>Heterocyclic compounds -<br>Oxygenated compounds -<br>Oxygenated compounds -<br>Hydrocarbons -<br>Benzenoids -<br>Lipids -<br>Heterocyclic compounds -<br>Oxygenated compounds -<br>Dipids -<br>Heterocyclic compounds -<br>Oxygenated compounds -<br>Dipids -<br>Hydrocarbons -<br>Benzenoids -<br>Lipids -<br>Heterocyclic compounds -<br>Oxygenated compounds -<br>Dipids -<br>Hydrocarbons -<br>Benzenoids -<br>Lipids -<br>Heterocyclic compounds -<br>Oxygenated compounds -<br>Dipids -<br>Hydrocarbons -<br>Benzenoids -<br>Lipids -<br>Heterocyclic compounds -<br>Oxygenated compounds -<br>Dipids -<br>Heterocyclic compounds -<br>Dipids -<br>Hydrocarbons -<br>Benzenoids -<br>Lipids -<br>Heterocyclic compounds -<br>Oxygenated compounds -<br>Dipids -<br>Hydrocarbons -<br>Benzenoids -<br>Lipids -<br>Heterocyclic compounds -<br>Oxygenated compounds -<br>Oxygenated compounds -<br>Dipids -<br>Hydrocarbons -<br>Benzenoids -<br>Lipids -<br>Heterocyclic compounds -<br>Oxygenated compounds -<br>Dipids -<br>Hydrocarbons -<br>Benzenoids -<br>Lipids -<br>Hy | Lipids -<br>Heterocyclic compounds -<br>Lipids -<br>Lipids -<br>Hydrocarbons -<br>Benzenoids | Lipids -<br>Heterocyclic compounds -<br>Lipids -<br>Lipids -<br>Lipids -<br>Hydrocarbons -<br>Benzenoids -<br>Lipids -<br>Heterocyclic compounds -<br>Lipids -<br>Hydrocarbons -<br>Benzenoids |

Fig. 1. (continued).

The differences in the fungal strains' metabolism and the nutrient availability can explain the results and trends (Calvo et al., 2002). The fungal cultures are able to intake nutrients just from the surface of the malt agar. Once those are finished, nutrient availability will depend on their diffusion through the agar until complete depletion (Sipiczki et al., 1998). In addition, it can be assumed that nutrient availability on natural samples is also limited. The proper interpretation of the results might be useful for the development of screening tools for the discrimination of houses and public premises contaminated by fungi.

### 3.1.3. Effect of the growing media additives

To evaluate the effect of the growing media additives on the fungal VOCs emission profiles, four vials were filled with 5 mL of malt agar, a

two of them seeded with *Aspergillus niger Hambi 495* as usual (Group 1) and the other two after addition of 2000  $\mu$ g mL<sup>-1</sup> of boric acid (Group 2). Antifungal properties of boric acid are well known and it is widely used as additive in building material manufacturing (De Seta et al., 2009). Both groups of *Aspergillus niger Hambi 495* samples were incubated under ambient and microaerophilic conditions. The evaluation of fungal emission profiles for VOCs, analyzed for a period ranging from 5 to 29 days after inoculation, were based on the number of compounds and on concentration values (Fig. 1 A to C).

Similar results were obtained for the number of compounds in fungal cultures incubated under ambient conditions. However, the number of identified VOCs was slightly larger for fungal cultures that grew after addition of boric acid to the malt agar in the vials incubated under microaerophilic atmosphere. In addition, the detailed evaluation of the VOC emission profiles, based on the functional groups present in the VOC molecules, gave clear differences. Those fungal cultures with boric acid addition, provided the largest concentration values.

Principal component analysis was used for the visualization of the differences between VOC emission profiles provided by the fungal cultures. The evaluation of the score plots revealed clear differences between the fungal cultures independently of the addition of antifungal additives. However, the analysis of the loading plots resulted in the concentration values of three compounds, 3-octanone, furan and styrene, that were statistically different for the samples under study (Fig. S2 B). These statistical differences between samples were confirmed by the LDA, due to the small number of samples. The prediction accuracy for samples not included in the model development was over 78%.

# 3.1.4. Comparison of volatile organic compounds emissions from fungal cultures

Volatile organic compounds emitted by 10 fungal cultures, half incubated under ambient and the rest under microaerophilic conditions, were analyzed for a time period ranging between 5 and 29 days after inoculation. The total number of analyzed samples were 140 and the number of compounds identified from each culture and the average concentration values were used as response factors for the potential discrimination between the fungal cultures under study (Fig. 1 A to C). The compounds analyzed from the blank vials were not considered as response.

Clear differences were observed in the VOC profiles achieved from the different fungal cultures under study. Just 2 VOCs were identified in *HJS11* and 26 compounds in *Aspergillus calidoustus MH34*. This fungal strain grown under ambient conditions gave also the highest VOC concentration values. However, the lowest concentration values were observed for *Chaetomium cochliodes OT7* grew under microaerophilic conditions.

The classification of the VOCs into groups based on their molecular structures revealed even more clear differences between fungal strains. For most of them, the highest concentration values were obtained for the oxygenated compounds with some exceptions, such as benzenoids (ambient) and lipids (microaerophilic) for *Aspergillus versicolor GAS/*226, lipids for *Aspergillus calidoustus MH34* (both ambient and microaerophilic), hydrocarbons for *Paecilomyces variotii Mö9* (both ambient and microaerophilic); and compounds with a lipid like substructure for *Aspergillus versicolor 9775* (ambient).

The differences, in the number of compounds and their concentrations, were exploited for the differentiation of fungal strains by a twostep approach. In the first one, PCA was used for the visualization of the potential differences between samples and the identification of marker compounds. In the second, SVM was used to confirm these differences and to evaluate its applicability for the identification/classification of the different fungal strains based on the VOC emission profiles.

The PCA scores plot for the visualization of the differences between fungal strains allowed the clear differentiation of at least one fungal strain in each PCA, making possible the development of a new and simpler PCA model after removal of samples previously identified. At least 6 PCA models, up to 8 for samples incubated under microaerophilic atmosphere, were required for the proper visualization of the differences between all the strains under study (Fig. S3 A).

In all the cases, PCA loading plots for the different PCs were used for the identification of the compounds responsible for the differences between the fungal strains, following the procedure described in our previous research (Liangsupree et al., 2021). The detailed evaluation of these plots showed that a combination of at least 23 compounds might be useful to distinguish the different fungal strains under study (Fig. S3 B). Most of these compounds (10) had a lipid like substructure, such as long chain alcohols (2-methyl-3-octanol, 3-octanol, 1-hexanol and 2-octen-1-ol) and sesquiterpenes (farnesene, cuprene, sesquisabene, chamigrene, bergamotene and valencene). Oxygenated compounds were the second group with the highest number of components (6). Short chain alcohols (3-methyl-1-butanol, and 2-methyl-1-butanol), ketones (acetone, 3-pentanone and 3-octanone) and organic acids (acetic acid) belonged to this group. Five hydrocarbons were included in the next group, including aliphatic and aromatic compounds (ethylene, undecane, 3,4-dimethyl-heptane, 1-3-dimethyl-benzene, and 1-ethyl-3-vinyl-adamantane). Finally, two heterocyclic compounds (2-(2-propenyl)-furan, and 2-methyl-furan) formed the last group. All these compounds, with the exception of sesquisabene, 3-pentanone, 3, 4-dimethyl-heptane, 1-3-dimethyl-benzene, 1-ethyl-3-vinyl-adamant ane and 2-(2-propenyl)-furan were already reported as VOCs emitted by fungi (Lemfack et al., 2018). However, it should be clarified that other compounds with similar structures to those ones have been frequently reported in the literature as emitted by fungi.

The potential use of the VOC emission profiles for the differentiation between fungal cultures was confirmed by the supervised pattern recognition learning algorithms. Poor prediction accuracy is expected for simple machine learning algorithms such as LDA, due to the relatively large number of samples and compounds. In this way, SVM should be considered as the most adequate algorithm to obtain reliable results. The samples were divided into two groups as a function of the growing atmosphere for the development of the SVM models. The results are clear in the case of the samples grew under ambient conditions. The percentage of samples, not included in the model, correctly classified was over 80%. However, the percentage of samples properly classified decreased to 67% in the case of the fungal cultures grown under microaerophilic conditions. This could be explained by the limited number of compounds emitted by the fungi under these conditions. Anyway, these results confirm the differences on the VOCs emitted by the fungal cultures and their potential identification capability.

# 3.2. Evaluation of volatile organic compounds profiles from indoor air samples

Volatile organic compounds from 52 indoor air samples, collected by SPME Arrow from control (44 samples) and previously indoor air problems reported houses (8 samples), were analyzed by TD–GC–MS. The results were evaluated in terms of number of compounds and concentrations. Potential differences between control and problematic samples were visualized by PCA and SVM, allowing the potential elucidation of marker compounds for indoor air problems.

Two hundred and forty-five compounds were identified after analysis of indoor air samples collected from houses with previous reported indoor air problems and reference ones. From these compounds, none of them was identified in all the samples under study. In addition, the number of compounds identified in a half of the samples was smaller than 10%. Finally, the percentage of compounds identified in a single sample was close to 16%. These results indicated the complexity of the indoor air samples, involving the contribution of multiple emission sources such as building materials, biogenic and anthropogenic activities. Moreover, the potential effect of the outdoor air should not be discarded in houses and premises with non-existent or inefficient air filtration systems (Geiss et al., 2011).

As can be seen in Fig. 2, the number of VOCs identified in the indoor air collected from houses with previously reported indoor air problems was clearly higher than that in reference houses. In addition, only 20% of VOCs identified were the same in the samples collected from reference and problematic houses. More differences were found when concentration values were compared after classification of the identified compounds into groups according to the different functional groups present in the molecules. Highest concentrations were obtained for oxygenated compounds and hydrocarbons in problematic houses. For example, the concentration of hydrocarbons were 8 times higher in problematic houses than in the reference ones. Lipids and benzenoids were more common in reference houses.

Principal component analysis was used for the visualization of the



**Fig. 2.** Evaluation of volatile organic compounds profiles from indoor air samples as a function of A) the number of compounds and B) the sum of the average concentration values (n = 3). The number of compounds belonging to each functional group is expressed in brackets. Label lipids includes compounds with a lipid like substructure. Houses with previously reported indoor air problems ( $\_$ ) and reference houses ( $\_$ ). Standard deviation values are included as error bars (n = 3).

differences between VOC emission profiles achieved from indoor air samples collected from control and problematic houses (Fig. S4). As expected from the results described in the previous paragraph, PCA score plot demonstrates clear differences between control and problematic houses indicating that a combination of at least 9 compounds might be useful to distinguish control and problematic houses (Fig. S4). These included compounds with a lipid like structure (hexanoic acid, limonene and  $\beta$ -phellandrene), benzenoids (benzaldehvde and toluene), hydrocarbons (ethylene and 3-methyl-nonane) and oxygenated compounds (ethyl acetate and 1,1-dimethoxy-propane). All these compounds with the exception of 3-methyl-nonane and 1,1-dimethoxypropane were already reported as VOCs emitted by fungi (Lemfack et al., 2018). In addition, some of these compounds can be considered as potential biomarkers to show differences between control and problematic houses. These statistical differences between the samples were confirmed by the use of SVM. In this case, the prediction accuracy for samples not included in the model development was over 91%. The large number of samples correctly classified confirms the clear dissimilarity observed for the VOC profiles of the indoor air samples collected from control and problematic houses.

# 3.3. Evaluation of emission of volatile organic compounds from insulation materials

Cellulose and glass-wool based insulation materials were collected from house with indoor air problem. The samples were collected and analyzed as described in the experimental section. The VOCs emitted by the different materials were evaluated in terms of number of identified compounds and concentration values. 3.3.1. Comparison of emissions from glass wool and cellulose insulation materials

As can be seen in Fig. 3 A, clear differences between VOC emission profiles were found in terms of number of compounds and concentrations. The number of VOCs emitted by glass-wool based insulation materials was clearly smaller than those from the cellulose-based ones. From all the compounds, 42% was emitted by both materials. Also, the concentration values were largest for all the groups of compounds from the cellulose-based insulation materials with the exception of the heterocyclic compounds.

Principal component analysis allowed the visualization of the differences between VOC that were emitted by glass-wool and cellulose based insulation materials (Fig. S5) and the score plots proved clear differences between both insulation materials. In addition, the analysis of the loading plots indicated differences on the VOC emission profiles, explained by the concentration values for ethyl acetate and hexanal (oxygenated compounds), toluene and benzaldehyde (benzenoids) (Fig. S5). Although the number of samples were limited, LDA could be used for the confirmation of these differences. The prediction accuracy for the differentiation between glass-wool and cellulose based insulation material samples not included in the model development was close to 87%. It should be emphasized that all the cellulose-based samples were classified correctly.

# 3.3.2. Comparison of indoor air and material emissions from houses contaminated with fungi

It is widely known that the fungi can use insulation materials as potential nutrient sources and growing media in houses affected by water damage (Jiang et al., 2022). Accordingly, the analysis of the indoor air samples and the insulation materials collected from a house





**Fig. 3.** Evaluation of volatile organic compounds emissions from insulation materials. Comparison of A) glass-wool and cellulose based insulation materials and B) indoor air and material emissions from houses contaminated with fungi. Label lipids includes compounds with a lipid like substructure. In brackets the number of compounds is belonging to the following ratios: indoor air samples (H)/indoor air samples and cellulose based insulation material (HC)/cellulose-based insulation material (C); indoor air samples (H)/indoor air samples and glass-wool based insulation material (HG)/glass-wool based isolation material (G). Glass-wool based insulation materials (\_\_\_), cellulose-based insulation materials (\_\_\_) and indoor air samples (\_\_). Standard deviation values are included as error bars (n = 3).

with indoor air problems was carried out to allow the potential identification of fungal activity markers among the VOCs. Because it was not possible to study the direct correlation between these compounds and the compounds directly emitted by the fungi, additional analysis of previously sterilized insulation materials was required to identify those VOCs that emitted by non-contaminated insulation materials from the emission profiles. Unfortunately, it was not possible to compare the concentration values of VOCs of indoor air and insulation materials due to the different conditions used in the analysis.

Higher number of VOCs in the indoor air samples compared to the VOCs from the insulation alone, demonstrated the complexity of the indoor air. Theoretically, all compounds in the insulation material might be also present in the indoor air samples. Compounds with a lipid like structure, oxygenated compounds and benzenoids were found in both indoor air and insulation materials (Fig. 3 B).

# 3.4. Development of screening tools for the discrimination of fungal contaminated samples

As discussed in the previous sections, the VOCs emitted by fungi and those found in the indoor air samples can be exploited for the identification of fungal strains and for the differentiation between houses of indoor air problems and control houses. In addition, when the VOC results obtained for indoor air samples and for the insulation materials (contaminated with fungi) were compared, the same compounds could be found in both samples. This allows the utilization of machine learning algorithms for the discrimination between control and fungal contaminated samples. However, different factors should then be considered before the selection and development of an appropriate machine learning algorithm. In general terms, simultaneous collection of gas phase, aerosol particles and/or dust are required for the development of the statistical models based on the VOC emission profiles, easily elucidated from gas phase samples. Yet, the identification of the fungal strains present in the houses requires the analysis of the aerosol or dust samples by classical identification techniques such as visual examination of agar cultures and strains or by more advanced ones based on DNA analysis. In this study, the statistical models were developed using the VOCs emitted by fungal cultures as variables. The VOCs identified from the indoor air samples were used for the discrimination between control and contaminated houses. The datasets from different sources might have a clear effect on the prediction capacity of the developed models. In addition, the limited number of fungal strains, used for model development, should also be considered as a limiting factor.

Finally, it should be noted that houses and premises might be contaminated by fungal strains not considered in this study or even by a mixture of multiple fungal species. In addition, the presence of houses with indoor air problems not related with fungal activity should not be discarded. In this way, the use of conventional machine learning approaches for classification of the samples into a single class might be problematic to provide reliable results. However, most of these problems can be partially overcome by the particular prediction capabilities of SIMCA that allows the classification of samples into no-belonging to the model, belonging to a single class or to overlapping classes.

In this work, the SIMCA model was developed using 140 VOC profiles obtained from 10 different fungal strains grew under 2 different atmospheres (ambient and microaerophilic) and malt agar as substrate. Effect of borax as additive was tested only with one strain. The great diversity of growing conditions, selected for the samples in the model development, allowed the simulation of fungal growing conditions in real samples. 52 indoor air samples were collected from control (44 samples) and houses with previously reported indoor air problems (8 samples). From problematic samples, 5 were collected from houses with fungal problems (2 houses) and the other 3 from a house with indoor air problems, not related to fungal activity. All the samples were collected and analyzed as described in the experimental section.

The validation of the SIMCA models involved two different

approaches. First the prediction capacity of the developed model was evaluated by the prediction of samples not used in the model development. Samples were randomly divided into training and validation sets containing 75% and 25% of the samples, respectively. Secondly, SIMCA model prediction capacity for fungal strains not included in the model development was evaluated using an approach developed in our previous research (Kopperi et al., 2013). Briefly, the VOC profiles from the fungal cultures were carefully divided into 10 training and validation sets. All the samples from the same fungal strain were included into a single validation set and the rest of the samples in the training set were used for the development of 10 new SIMCA models.

The validation of the SIMCA model using the first approach gave a prediction accuracy over 84% considering all the correct responses into a single class or multiple classes. However, this value dropped to 62% correct classification if single group was considered. On the other side, the results achieved for the validation of the SIMCA models using the second approach proved that none of them could be classified into a single class. In all the cases, the samples were classified into multiple classes (Aspergillus versicolor GAS/226, Chaetomium cochliodes OT7, Aspergillus westerdijkiae PP2, HJS11 and Aspergillus versicolor 9775/K1) in a typical SIMCA overlapping classes response. This response for samples not used in model development might indicate potential contamination of the samples by fungi. The average prediction capacity of the SIMCA for the classification of groups of samples not used in the model development was over 93%, indicating the potential applicability of this SIMCA model as screening tool for the discrimination of fungal contaminated samples.

The SIMCA model applied to the analysis of natural indoor air samples (Table 1) proved that none of the control samples was classified as contaminated ones. In addition, 80% of the samples from houses contaminated with fungi were classified properly. As expected, all the samples were classified as contaminated with *Aspergillus versicolor GAS/*226, Chaetomium cochliodes OT7, Aspergillus westerdijkiae PP2, HJS11 and Aspergillus versicolor 9775/K1. However, it should be noted that

Table 1

Application of the SIMCA model to the analysis of natural indoor air samples, model validation and prediction results.

|                                      | Ν  | Contaminated | Single<br>Class | Multiple<br>class | No<br>Contaminated |
|--------------------------------------|----|--------------|-----------------|-------------------|--------------------|
| Aspergillus niger<br>Hambi 495       | 19 | 17           | 10              | 7                 | 2                  |
| Aspergillus<br>versicolor<br>GAS/226 | 13 | 13           | 10              | 3                 | 3                  |
| Chaetomium<br>cochliodes OT7         | 9  | 8            | 7               | 1                 | 1                  |
| Aspergillus<br>calidoustus<br>MH34   | 19 | 16           | 8               | 8                 | 3                  |
| Chaetomium<br>globosum<br>ABCD       | 20 | 19           | 11              | 8                 | 1                  |
| Aspergillus<br>westerdijkiae<br>PP2  | 10 | 10           | 8               | 2                 | 0                  |
| Paecilomyces<br>variotii Mö9         | 12 | 11           | 11              | 0                 | 1                  |
| HJS11                                | 17 | 15           | 0               | 15                | 2                  |
| HKIP12                               | 9  | 9            | 9               | 0                 | 0                  |
| Aspergillus<br>versicolor<br>9775/K1 | 12 | 12           | 12              | 0                 | 0                  |
| Reference<br>houses                  | 44 | 0            | 0               | 0                 | 44                 |
| Problematic<br>houses                | 8  | 4            | 0               | 4                 | 4                  |
| Houses<br>contaminated<br>with fungi | 5  | 4            | 0               | 4                 | 1                  |

samples obtained from a house with indoor air problems, not related to fungal contamination, were classified as non-contaminated. In addition, one false negative result was achieved for the sample collected far away from the rooms with fungal problems where the ventilation system was working properly.

Overall, the SIMCA model seems to be a promising screening tool to recognize differences between control and fungi contaminated houses. However, as described earlier, the proper identification of the fungal strains was not possible.

### 4. Conclusions

The applicability of the SPME Arrow for the collection of VOCs emitted by fungi and insulation materials, and by indoor air was demonstrated in this research. In addition, a wide range of machine learning techniques were exploited for the visualization and statistical differentiation between the samples.

The VOC emission profiles allowed the differentiation of the fungal strains. However, it should be emphasized that fungal emission profiles were clearly influenced by a wide range of external variables including the incubation atmosphere, the incubation time after inoculation and additives used in the growing media.

The comparison of the emissions caused by indoor air and insulation materials collected from houses contaminated with fungi, displayed the presence of a relatively large number of VOCs in both samples. These compounds were considered as fungal markers in indoor air samples. Unfortunately, it was not possible to make the discrimination between the compounds emitted by the insulation material and those by the fungi present in the samples.

Clear differences were observed between indoor air samples collected from reference houses and from those with indoor air problems. These differences enabled the development of SIMCA based screening tool using VOC fungal emission profiles for the discrimination between control houses and fungi contaminated houses. It is noteworthy that it is not unfortunately possible to use totally different samples for model development and analysis for the identification of the individual fungal strains. However, the results achieved for the discrimination between the control and contaminated houses with fungi form an excellent platform for the future development of a indoor air screening tool.

#### Credit authors statement

Jose Ruiz-Jimenez: Conceptualization, Supervision, Investigation, Visualization, Writing – original draft. Sanni Raskala: Investigation, Formal analysis, Visualization. Ville Tanskanen: Investigation, Formal analysis, Visualization. Elisa Aattela Investigation, Visualization. Mirja Salkinoja-Salonen: Conceptualization, Supervision, Visualization. Kari Hartonen: Conceptualization, Supervision, Reviewing and Editing. Marja-Liisa Riekkola: Conceptualization, Supervision, Funding acquisition, Reviewing and Editing.

### Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Jose Ruiz-Jimenez, Marja-Liisa Riekkola reports financial support was provided by Jane and Aatos Erkko Foundation. Jose Ruiz-Jimenez, Marja-Liisa riekkola reports financial support was provided by Academy of Finland.

# Data availability

Data will be made available on request.

### Acknowledgements

Financial support was obtained from the Jane and Aatos Erkko Foundation and ACCC flagship "Finnish Research Flagship funded by the Academy of Finland (grant no. 337549). We thank Eka Dian Pusfitasari, Meri Sillanpää, Maria Andersson, Giorgia Demaria, Hangzen Lan, Nicola Zanca and Ilmari Heiskanen for their cooperation and valuable help.

### Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.envres.2023.115494.

#### References

- Adamová, T., Hradecký, J., Prajer, M., 2019. VOC emissions from spruce strands and hemp shive: in search for a low emission raw material for bio-based construction materials. Materials 12 (12), 2026.
- Ahmed, W.M., Geranios, P., White, I.R., Lawal, O., Nijsen, T.M., Bromley, M.J., Goodacre, R., Read, N.D., Fowler, S.J., 2018. Development of an adaptable headspace sampling method for metabolic profiling of the fungal volatome. Analyst 143 (17), 4155–4162.
- Betancourt, D.A., Krebs, K., Moore, S.A., Martin, S.M., 2013. Microbial volatile organic compound emissions from Stachybotrys chartarumgrowing on gypsum wallboard and ceiling tile. BMC Microbiol. 13 (1), 1–10.
- Brown, V.M., Crump, D.R., Harrison, P.T., 2013. Assessing and controlling risks from the emission of organic chemicals from construction products into indoor environments. Environ. Sci.: Process. Impacts 15 (12), 2164–2177.
- Calvo, A.M., Wilson, R.A., Bok, J.W., Keller, N.P., 2002. Relationship between secondary metabolism and fungal development. Microbiol. Mol. Biol. Rev. 66 (3), 447–459.
- De Seta, F., Schmidt, M., Vu, B., Essmann, M., Larsen, B., 2009. Antifungal mechanisms supporting boric acid therapy of Candida vaginitis. J. Antimicrob. Chemother. 63 (2), 325–336.
- Gallon, V., Le Cann, P., Sanchez, M., Dematteo, C., Le Bot, B., 2020. Emissions of VOCs, SVOCs, and mold during the construction process: contribution to indoor air quality and future occupants' exposure. Indoor Air 30 (4), 691–710.
- Geiss, O., Giannopoulos, G., Tirendi, S., Barrero-Moreno, J., Larsen, B.R., Kotzias, D., 2011. The AIRMEX study-VOC measurements in public buildings and schools/ kindergartens in eleven European cities: statistical analysis of the data. Atmos. Environ. 45 (22), 3676–3684.
- Jiang, L., Pettitt, T.R., Buenfeld, N., Smith, S.R., 2022. A critical review of the physiological, ecological, physical and chemical factors influencing the microbial degradation of concrete by fungi. Build. Environ. 108925.
- Kim, K.-H., Jahan, S.A., Lee, J.-T., 2011. Exposure to formaldehyde and its potential human health hazards. J. Environ. Sci. Health, Part C 29 (4), 277–299.
- Kopperi, M., Ruiz-Jiménez, J., Hukkinen, J.I., Riekkola, M.-L., 2013. New way to quantify multiple steroidal compounds in wastewater by comprehensive twodimensional gas chromatography-time-of-flight mass spectrometry. Anal. Chim. Acta 761, 217–226.
- Korpi, A., Järnberg, J., Pasanen, A.-L., 2009. Microbial volatile organic compounds. Crit. Rev. Toxicol. 39 (2), 139–193.
- Lan, H., Hartonen, K., Riekkola, M.-L., 2020. Miniaturised air sampling techniques for analysis of volatile organic compounds in air. TrAC, Trends Anal. Chem. 126, 115873.
- Lan, H., Ruiz-Jimenez, J., Leleev, Y., Demaria, G., Jussila, M., Hartonen, K., Riekkola, M.-L., 2021. Quantitative analysis and spatial and temporal distribution of volatile organic compounds in atmospheric air by utilizing drone with miniaturized samplers. Chemosphere 282, 131024.
- Lee, S., Hung, R., Yap, M., Bennett, J.W., 2015. Age matters: the effects of volatile organic compounds emitted by Trichoderma atroviride on plant growth. Arch. Microbiol. 197 (5), 723–727.
- Lemfack, M.C., Gohlke, B.-J., Toguem, S.M.T., Preissner, S., Piechulla, B., Preissner, R., 2018. mVOC 2.0: a database of microbial volatiles. Nucl. Acids Res. 46 (D1), D1261–D1265.
- Liangsupree, T., Multia, E., Saarinen, J., Ruiz-Jimenez, J., Kemell, M., Riekkola, M.-L., 2021. Raman spectroscopy combined with comprehensive gas chromatography for label-free characterization of plasma-derived extracellular vesicle subpopulations. Anal. Biochem. 647, 114672.
- Misztal, P.K., Lymperopoulou, D.S., Adams, R.I., Scott, R.A., Lindow, S.E., Bruns, T., Taylor, J.W., Uehling, J., Bonito, G., Vilgalys, R., 2018. Emission factors of microbial volatile organic compounds from environmental bacteria and fungi. Environ. Sci. Technol. 52 (15), 8272–8282.
- Multia, E., Liangsupree, T., Jussila, M., Ruiz-Jimenez, J., Kemell, M., Riekkola, M.-L., 2020. Automated on-line isolation and fractionation system for nanosized biomacromolecules from human plasma. Anal. Chem. 92 (19), 13058–13065.
- Ortíz-Castro, R., Contreras-Cornejo, H.A., Macías-Rodríguez, L., López-Bucio, J., 2009. The role of microbial signals in plant growth and development. Plant Signal. Behav. 4 (8), 701–712.
- Park, J.H., Sulyok, M., Lemons, A.R., Green, B.J., Cox-Ganser, J.M., 2018. Characterization of fungi in office dust: comparing results of microbial secondary metabolites, fungal internal transcribed spacer region sequencing, viable culture and other microbial indices. Indoor Air 28 (5), 708–720.

#### J. Ruiz-Jimenez et al.

Pomerantsev, A.L., Rodionova, O.Y., 2020. Popular decision rules in SIMCA: critical review. J. Chemometr. 34 (8), e3250.

Portnoy, J.M., Barnes, C.S., Kennedy, K., 2004. Sampling for indoor fungi. J. Allergy Clin. Immunol. 113 (2), 189–198.

- Pusfitasari, E.D., Ruiz-Jimenez, J., Heiskanen, I., Jussila, M., Hartonen, K., Riekkola, M.-L., 2022. Aerial drone furnished with miniaturized versatile air sampling systems for selective collection of nitrogen containing compounds in boreal forest. Sci. Total Environ. 808, 152011.
- Ruiz-Jimenez, J., Heiskanen, I., Tanskanen, V., Hartonen, K., Riekkola, M.-L., 2022. Analysis of indoor air emissions: from building materials to biogenic and anthropogenic activities. J. Chromatogr. Open 100041.
- Salonen, H., Lappalainen, S., Lindroos, O., Harju, R., Reijula, K., 2007. Fungi and bacteria in mould-damaged and non-damaged office environments in a subarctic climate. Atmos. Environ. 41 (32), 6797–6807.
- Schweizer, C., Edwards, R.D., Bayer-Oglesby, L., Gauderman, W.J., Ilacqua, V., Juhani Jantunen, M., Lai, H.K., Nieuwenhuijsen, M., Künzli, N., 2007. Indoor time-microenvironment-activity patterns in seven regions of Europe. J. Expo. Sci. Environ. Epidemiol. 17 (2), 170–181.
- Sipiczki, M., Takeo, K., Yamaguchi, M., Yoshida, S., Miklos, I., 1998. Environmentally controlled dimorphic cycle in a fission yeast. Microbiology 144 (5), 1319–1330.

Tzanakou, E.M., 2017. Supervised and Unsupervised Pattern Recognition: Feature Extraction and Computational Intelligence. CRC press.

Vergara-Fernández, A., Revah, S., Moreno-Casas, P., Scott, F., 2018. Biofiltration of volatile organic compounds using fungi and its conceptual and mathematical modeling. Biotechnol. Adv. 36 (4), 1079–1093.