Available online at www.sciencedirect.com
ScienceDirect
Journal of Microbiological Methods 69 (2007) 249–255
Journal of Microbiological Methods
www.elsevier.com/locate/jmicmeth

A method for comparison of growth media in objective identification of *Penicillium* based on multi-spectral imaging

Line H. Clemmensen [a,*], Michael E. Hansen [b], Jens C. Frisvad [b], Bjarne K. Ersbøll [a]

[a] *Technical University of Denmark, Informatics and Mathematical Modelling, Denmark*
[b] *Technical University of Denmark, BioCentrum, Denmark*

Received 25 October 2006; received in revised form 14 December 2006; accepted 18 December 2006
Available online 3 January 2007

Abstract

We consider the problems of using excessive growth media for identification and performing objective identification of fungi at the species level. We propose a method for choosing the subset of growth media, which provides the best discrimination between several fungal species. Furthermore, we propose the use of multi-spectral imaging as a means of objective identification.

Three species of the fungal genus Penicillium are subject to classification. To obtain an objective classification we use multi-spectral images. Previously, RGB images have proven useful for the purpose. We use multi-spectral bands as they provide additional information about the chemistry of the fungal colonies. In this study three media [Czapek yeast extract agar (CYA), oatmeal agar (OAT), and yeast extract sucrose agar (YES)] have been compared on their ability to discriminate between the three species. We propose a statistical method to test which medium or combination of media gives the best discrimination.

Statistical tests indicate that YES combined with CYA is the best choice of media in this case. However, for the objective identification one medium is sufficient to discriminate between the species. Statistical tests show that there are significant differences between the species on all individual media, and that these differences are largest on YES. The objective identification has been performed solely by means of digital image analysis. The features obtained from the image analysis merely correspond to macro-morphological features. The species have been classified using only 3–4 of the spectral bands with a 100% correct classification rate using both leave-one-out cross-validation and test set validation.

Keywords: Choice of growth media; Filamentous fungi; Macro-morphology; Multi-spectral images; Objective identification; *Penicillium*

1. Introduction

Penicillium is one of the most important fungal genera. Many species are used for production of important drugs, such as penicillin and compactin, while others are used for food fermentations, such as *Penicillium roqueforti* for blue mold cheeses, *P. camemberti* for white mold cheeses, and *P. nalgiovense* for mold fermented salami (Frisvad and Samson, 2004). Many other species can cause food and feed spoilage and deterioration of other stored materials. Several of these species can furthermore produce mycotoxins in those foods and feeds and in indoor construction materials (Samson et al., 2004; Rand et al., 2005).

Correct classification and identification at the species level is of paramount importance in order to be able to predict positive characteristics of fungi used in biotechnology or in order to predict and prevent mycotoxin formation. However, it has often been reported that most of the features used for classification and identification of fungi are highly subjective, especially macro-morphological features (Raper and Thom, 1949; Pitt, 1979a; Ramirez, 1982; Samson and Frisvad, 1993). Often minute differences in spore color, colony texture, diffusible pigments, and guttation droplets are used to identify new isolates to species level, occasionally leading to misidentifications (Frisvad et al., 2006). Colony colors, especially the different green colors of the conidia *en masse* of *Penicillium*, is often used as an important criterion for species identity and these colors can be determined using comparison to color handbooks, an inherently subjective

* Corresponding author. Informatics and Mathematical Modelling, Building 321, Technical University of Denmark, DK-2800 Lyngby, Denmark.
E-mail addresses: lhc@imm.dtu.dk (L.H. Clemmensen), meh@biocentrum.dtu.dk (M.E. Hansen), jcf@biocentrum.dtu.dk (J.C. Frisvad), be@imm.dtu.dk (B.K. Ersbøll).

0167-7012/$ - see front matter © 2007 Published by Elsevier B.V.
doi:10.1016/j.mimet.2006.12.020

measurement (Christensen et al., 1994). Conidium color is also a criterion that is dependent on growth medium, length of incubation, illumination of the colonies and other environmental factors, and even within a single fungal colony, colors may vary from the center to the margin of the colony (Raper and Thom, 1949). Clearly, objective methods are needed concerning macromorphology and some attempts at this have been reported (Dorge et al., 2000; Hansen and Carstensen, 2003, 2004; Hansen et al., 2003). For the usual subjective recordings of the features, several diagnostic media are used in *Penicillium* taxonomy (Raper and Thom, 1949; Pitt, 1979a; Ramirez, 1982; Frisvad and Samson, 2004), but it is time consuming and expensive to use a large number of identification media, so new methods are needed in order to choose the one or maybe two media, that are optimal for accurate identifications.

Three media [Czapek yeast extract agar (CYA) (Pitt, 1979b), oatmeal agar (OAT)(Samson et al., 2000), and yeast extract sucrose agar (YES)(Samson et al., 2000)] were statistically tested using multi-spectral imaging (18 spectral bands) to determine which medium (or media) contributed information necessary for the objective identification of the fungal species.

Due to the large interest in the *Penicillium* genus there is a substantial knowledge of the species and well identified isolates exist which gives an accurate ground truth for the classification. Three species of the *Penicillium* genus are investigated here: *P. polonicum* (pol), *P. venetum* (ven), and *P. melanoconidium* (mel). The three species are all in the section Viridicata (Frisvad and Samson, 2004), but where two of them belong to the series *Viridicata* the third one belongs to the series *Corymbifera*.

P. melanoconidium inhabits grains such as wheat, rye, oat, rice, and barley. Hence, it is most commonly found in cereals. It is one of the *Penicillium* species that has the most pure green colors *en masse* in the genus and is of the series *Viridicata* (Frisvad and Samson, 2004).

P. polonicum is a common mold on dry-cured meat products. Also, it inhabits wheat, barley, rice, rye, oat, rice, corn, peanuts, onions, and vegetable field soil. It is typically the *Penicillium* species with the largest amount of blue in the conidium color *en masse* and is of the series Viridicata (Frisvad and Samson, 2004).

P. venetum is commonly found in soil decaying with vegetation such as onions and flower bulbs and is therefore ecologically different from the cereal-borne members of the *Viridicata* series. It has blue–green conidia en masse and is of the series *Corymbifera* (Frisvad et al., 2004).

The striking color difference between *P. melanoconidium* and *P. polonicum* is illustrated in a rare color photograph in the 1949 monograph on *Penicillium* by Raper and Thom (page 428a). Superficially, *P. polonicum* and *P. venetum* could look like they are the most closely related, but it is in fact *P. polonicum* and *P. melanoconidium* that are the most closely related (Samson et al., 2004).

Furthermore, all species produce different mycotoxins. The natural products produced by the three species are reviewed by Frisvad et al. (2004). Hence, an objective method that can separate these three important species, and allow identification based on objective image analysis, is highly desirable. Additionally, it is of interest which species are separated best on the media to see whether the discrimination is based on macromorphological features or not.

During the identification it is desirable to identify the medium or combination of media on which the species are best discriminated.

Table 1
IBT numbers of the twelve *Penicillium* isolates

Isolates	Species		
	P. mel	*P. pol*	*P. ven*
a	3445	22,439	23,039
b	21,534	15,982	21,549
c	3443	14,320	16,215
d	10,031	11,383	16,308

2. Materials and methods

In this section we describe the experimental design, the acquisition of digital images of the fungal colonies, and how to obtain relevant information from the images.

2.1. Incubation of isolates

For each species 4 isolates were chosen that represent a wide geographical range. The fungal isolates were obtained from the IBT Culture Collection held at BioCentrum-DTU, Technical University of Denmark. The IBT numbers of the species are listed in Table 1.

The isolates were inoculated as three point cultures on three different media: CYA, YES, and OAT, and with three replica on each medium. The inoculation was performed in 9 cm petri dishes containing one of the three growth substrates. The total number of samples were thus 108.

After incubation in complete darkness for 7 days at 25 °C, the cultures reach their stationary phase and are able to produce secondary metabolites. At this stage the colonies have grown into three circular objects within the petri dish and the fungal colonies are digitized.

2.2. Digitalization

The Penicillium samples were digitized at the Department of Informatics and Mathematical Modelling (IMM), Technical University of Denmark (DTU) with a very accurate digital camera system called Videometer Lab[1]. The Videometer Lab is illustrated in Fig. 1, it consists of an integrating sphere illumination (an Ulbricht sphere) combined with a carefully calibrated digital camera. The sphere has a diameter of 36 cm. The inside of the sphere is covered with a matte titanium paint to create optimum light conditions. Diodes at the rim of the sphere provide an even and diffuse illumination of the sample. The Videometer camera system generates images with a radiometric resolution of 8 bits/pixel for each spectral channel. The result is a floating-point image of 18 spectral bands with a spatial

[1] http://www.videometer.com.

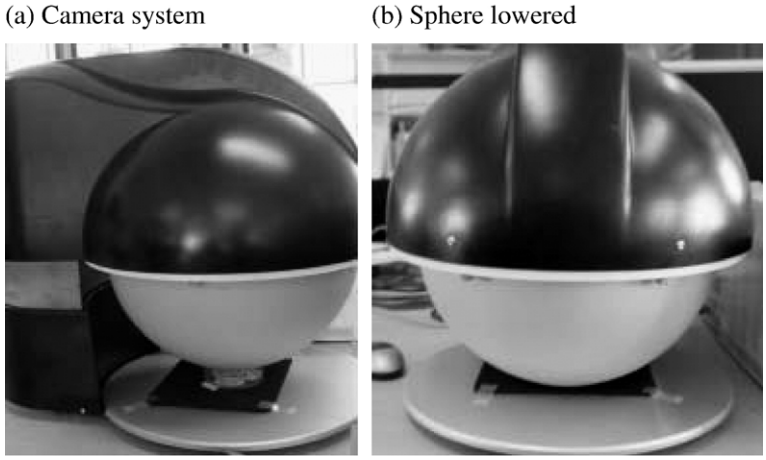


Fig. 1. The camera system from the side and from the front. In the image of the front the sphere has been lowered before taking the image.

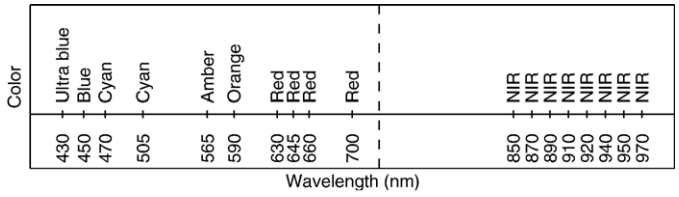
resolution of 960×1280 pixels. The spectral bands are illustrated in Table 2 and Fig. 2.

One of the samples was contaminated, hence 107 of the 108 samples were digitized. In order to avoid complications in the following statistical analyses due to missing values, an average of the features of the two remaining replica was used to replace the contaminated sample.

2.3. Image analysis

The fungal colonies were segmented from the digital images using the Histogram Pursuit method (Gomez et al., 2006). Histogram Pursuit looks for linear projections with bi- or multimodality in data and in this way partitions data into two or more subgroups. That is, by thresholding between modes in histograms of one-dimensional projections of the images, the images were labelled into several subgroups. Subsequently, the subgroups were identified as colony or background and masks of the colonies were obtained. Once the colonies were segmented from the background and the petri dish, features were extracted. The segmentations were performed on the 6th spectral band as this band gave a good contrast between the fungal colonies and the growth substrates. Fig. 3 illustrates one of the samples, the HP segmentation, and the resulting mask. In line with Dörge et al. (2000), the following features were extracted: The 1st, 5th, 10th, 30th, 50th, 70th, 90th, 95th, and 99th percentiles, the mean, the standard deviation, and the maximal intensity. The features were extracted both for the original spectra, the difference of the spectra, and the pairwise products of the spectra.

Table 2
The wavelengths of the eighteen spectral bands and a description of the colors corresponding to these wavelengths



Some of the features gave zero for all samples and were therefore disregarded. In total 3754 features were kept.

2.4. Methods of data analysis

2.4.1. Principal component analysis

Principal component analysis (PCA) is a data decomposition technique used to reduce the number of dimensions in data, a necessity during this study since the number of features was large relative to the number of observations.

The principal components (PCs) of a multivariate data set are mutually orthogonal linear combinations of the original variables. The PCs are composed to explain the largest possible amount of variance in data, hence, the first PC has the largest variance, the second PC the second largest variance and so forth (Jolliffe, 1986). Since the variance depends on the scale of the original variables, the data is standardized to have zero mean and variance one. The directions of the principal components are the eigenvectors of the data dispersion matrix. In many cases, the first few PCs explain most of the variance of data and are therefore sufficient for further analyses.

2.4.2. Multivariate analysis of variance

The multivariate analysis of variance verifies the validity of the experiment, as well as the reproducibility.

Three-sided multivariate analysis of variance (Rencher, 2002) tests the significance of the effects: Medium (M), species (S), and isolate (I). Species with respect to medium and isolate with respect to medium are so-called cross classifications whereas species, medium, and isolate with respect to repetition (R), and species with respect to isolate are hierarchical (or nested) classifications. That is:

$$\begin{array}{ll} S \times M & S \supset I \\ I \times M & S \supset R \\ & I \supset R \\ & M \supset R \end{array}$$

The medium and species factors are considered deterministic, while the repetition factor is considered random. It can be

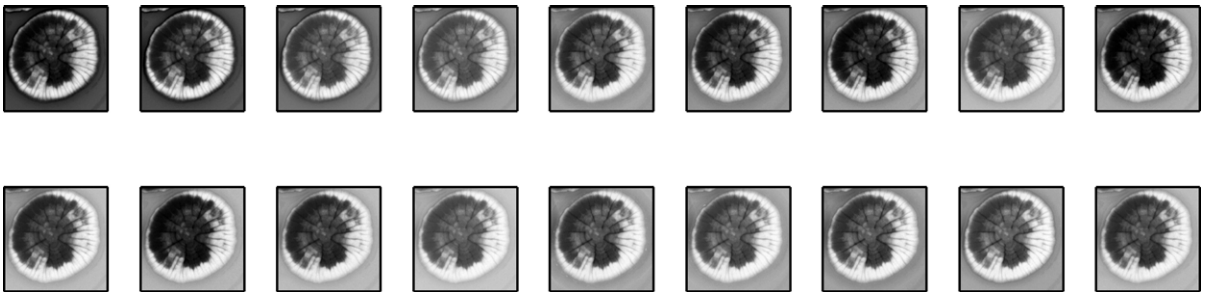


Fig. 2. One of the fungal colonies illustrated for the 18 spectral bands in increasing order reading from left to right.

argued whether the isolate factor is indeed deterministic or stochastic. As the isolates were chosen to represent a large geographic region and can be reproduced, they could be regarded deterministic (Model (1)). On the other hand, if another laboratory was to reproduce the experiment the isolates might not be the same and the factor could then be regarded as stochastic (Model (2)).

Two models were therefore investigated: A deterministic model where the isolate effect is deterministic (Model (1)), and a mixed model where the isolate effect is stochastic (Model (2)). The models state that the variables can be decomposed into the mean of all observations plus a contribution for each of the effects and their interactions.

The models are

$$\overrightarrow{X}_{kljv} = \overrightarrow{\mu} + \overrightarrow{m}_k + \overrightarrow{s}_l + \overrightarrow{ms}_{kl} + i(\overrightarrow{s})_{j(l)} + m\overrightarrow{i}(s)_{kj(l)} + R(\overrightarrow{msi})_{v(klj)} \quad (1)$$

and

$$\overrightarrow{X}_{kljv} = \overrightarrow{\mu} + \overrightarrow{m}_k + \overrightarrow{s}_l + \overrightarrow{ms}_{kl} + I(\overrightarrow{s})_{j(l)} + m\overrightarrow{I}(s)_{kj(l)} + R(\overrightarrow{msI})_{v(klj)}, \quad (2)$$

where X_{kljv} is an observation consisting of the two first PCs with

$k = 1, \ldots, 3$(medium)
$l = 1, \ldots, 3$(species)
$i = 1, \ldots, 4$(isolate (within species))
$v = 1, \ldots, 3$(repetition (within medium, species, and isolate)).

Conventionally, a term containing only lower case letters is a deterministic effect, while a term containing a capital letter is stochastic. Note, that the interaction between the stochastic isolate effect and the deterministic medium effect in Model (2) therefore is a stochastic term.

2.4.3. Statistical tests for media

Mahalanobis distance, Hotelling's T^2 test, and the test of additional information were used to determine which medium and/or combination of media were best with respect to discrimination between species.

Hotelling's T^2 tests (Rencher, 2002), testing for differences in means between species, were conducted. This test uses Mahalanobis distance as a measure of the distance between the species (Rencher, 2002). The null-hypothesis of the test is that the means of the three species are equal. Furthermore, tests were conducted to determine the significance of additional information provided by each medium to the discrimination (Rencher, 2002). The null-hypothesis states that the last q variables do not contribute to a better discrimination. The tests were conducted at a 5% level of significance.

2.4.4. Discriminant analysis

To classify the three species we used Bayes solution of discriminant analysis (Rencher, 2002; Hastie et al., 2001). The Bayes' rule chooses the population where the observed posterior probability is largest, i.e. the population a given observation is more likely to be observed in. Linear functions were used to discriminate between the populations. This implies that we assumed that the three populations have equal dispersion.

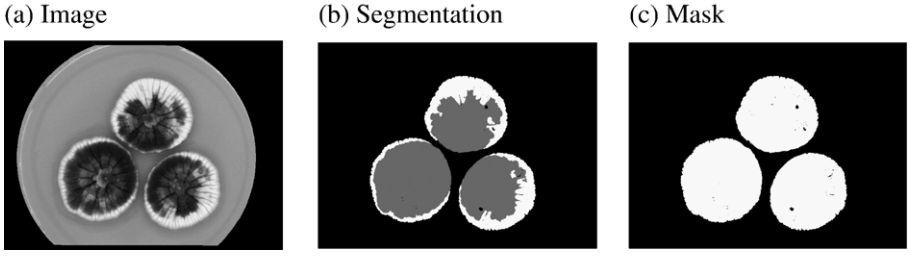


Fig. 3. The 6th spectral band (590 nm, orange) of *P. melanoconidium* isolate a, 1st repetition on YES, its corresponding segmentation based on HP, and its corresponding mask.

2.4.5. Wilks' lambda and forward selection

To discriminate between the classes, we wanted the within group deviation to be small compared to that between groups. One way of accomplishing this is to minimize

$$\Lambda = \frac{\det(\mathbf{W})}{\det(\mathbf{T})},$$

which is also called Wilks' Λ, where **W** is the within group deviation and **T** is the total deviation, i.e. the within group plus the between group deviation.

Forward selection was used to select features to include in the discriminant analysis (Rencher, 2002; Hastie et al., 2001). The feature with the most statistically significant Λ was selected first, secondly the feature with the most statistically significant partial Λ and so forth.

2.4.6. Cross- and test set validation

To validate the results obtained in the discriminant analysis both *leave-one-out cross-validation* and *test set validation* was used (Hastie et al., 2001; Duda et al., 2001). In leave-one-out cross-validation one observation is left out of the analysis and hereby classified based on the model obtained with the remaining observations. This procedure is conducted for all observations. In test set validation, data is split in two and the test part is classified based on the model obtained with the training set.

3. Results and discussion

3.1. Multivariate analysis of variance

The analysis of variance verified the validity and the reproducibility of the experiment. The two first principal components PC1 and PC2 were examined and in this case explained 56% of the original variance in data. Only two variables could be included in the analyses due to the degrees of freedom from the repetitions. The results of tests of each term in the two models were summed up in Tables 3 and 4.

For Model (1) all main effects and all interactions were strongly significant. Consequently, there was a statistical difference between media, between species, and between isolates. As well as some species favored one medium more than another ($ms_{kl} \neq 0$) and some isolate(s) within a species favored one medium more than another ($mi(s)_{kj(l)} \neq 0$).

For Model (2) the two most interesting main effects: Media and species were strongly significant, while their interaction was not quite significant. Hence, there was still a difference between media and between species, but it was no longer significant that one species favored one medium compared to another. The main effect of isolates was not significant in Model (2), i.e. there was no significant difference between the isolates within a species. The interaction between the media and the isolates within species was significant, i.e. some isolate within a species favored one medium more than another.

When isolate was considered deterministic all effects were significantly different from zero (Model (1)). When isolate, on the other hand, was considered stochastic we accepted the null-hypothesis that the variance of the isolates was zero (Model (2)). This leads us to conclude that the experiment can be reproduced both using the same isolates (Model (1)), and using different isolates (Model (2)).

The interest was now focused on determining which medium or combination of media were the better choice. Since there was a significant difference between media, the observations from each medium were henceforth either considered separately or as additional features to the same observations, i.e. there were 36 observations with three times the number of features.

Table 3
Summary of the multivariate tests based on Model (1)

H_0/variables	PC1 & PC2
$m_k=0$	*R*(1%)
$s_l=0$	*R*(1%)
$ms_{kl}=0$	*R*(1%)
$i(s)_{j\ (l)}=0$	*R*(1%)
$mi(s)_{kj(l)}=0$	*R*(1%)

The actual *p*-value of each test is given in parentheses.
A indicates the null-hypothesis (H_0) was accepted and *R* that H_0 was rejected.

Table 4
Summary of the multivariate tests based on Model (2)

H_0/variables	PC1 & PC2
$m_k=0$	*R*(1%)
$s_l=0$	*R*(1%)
$ms_{kl}=0$	*A*(8%)
$\sigma^2_{I\ (s)}=0$	*A*(18%)
$\sigma^2_{MI\ (s)}=0$	*R*(1%)

The actual *p*-value of each test is given in parentheses.
A indicates the null-hypothesis (H_0) was accepted and *R* that H_0 was rejected.

3.2. Tests for media

In these analyses the principal components were chosen to represent data. This was done in order to obtain the best description of the variance in data in as few dimensions as possible. Respectively, 87% and 92% of the variance in data was explained by the first 7 and the first 11 principal components. The multi-variable distances between species on the

Table 5
Mahalanobis distances between the species for each of the individual media and combinations of media

Medium/Distance	mel–pol	mel–ven	pol–ven
YES	216 (<1%)	53 (<1%)	140 (<1%)
OAT	73 (<1%)	54 (<1%)	19 (<1%)
CYA	41 (<1%)	35 (<1%)	21 (<1%)
YES & OAT	1582 (<1%)	142 (<1%)	480 (<1%)
YES & CYA	763 (<1%)	356 (<1%)	1410 (<1%)
OAT & CYA	217 (<1%)	642 (<1%)	195 (<1%)
YES & OAT & CYA	3710 (2%)	4440 (1%)	3669 (2%)

The calculations were based on the first seven PCs. For the features of the edge and fungi in one. In parentheses are given the *p*-values of Hotelling's T^2-test of the null-hypothesis that the means of the two species are equal.

different media were generally largest on the YES medium compared to the other two media, cf. Table 5. According to Hotelling's T^2 test, the species were significantly different from each other (the distances between species were significantly different from zero). In addition to this, the distances between *P. melanoconidium* and *P. polonicum* were larger than those between *P. polonicum* and *P. venetum*, implying that the largest variances were caused by macro-morphological features rather then genetic relations. As one might expect, using combinations of 2 or all 3 media gave increasing distances. It seems that the best choice of a combination of media depends on which two species are desired discriminated.

In order to address the problem of determining whether a medium actually contributed to the objective discrimination of species, each medium was regarded as additional information to the same observation. I.e. one observation had q PCs belonging to each of the three media. Tests of the null-hypothesis that a medium does not contribute to the discrimination compared to one or two media were conducted. This corresponds to the null-hypothesis that the q PCs belonging to that medium do not contribute to the discrimination. Tests of information were conducted both for discrimination between all three species and between two species at a time. When all species were regarded and the tests were based on the first 11 PCs all media contributed significantly to the discrimination. The tests where two species were regarded at a time are summed up in Table 6.

The results show that it can be assumed that using three media compared to two did not provide additional information to the discrimination between species. Hence, two media should be sufficient. In addition to this, it can be assumed that OAT and YES do not provide additional information to one another in terms of discrimination between *P. melanoconidium* and *P. venetum*. OAT does not provide additional information to YES with regard to discrimination between *P. polonicum* and *P. venetum*. CYA does not provide additional information to OAT with regard to discrimination between *P. melanoconidium* and *P. polonicum*. Consequently, YES and CYA should be the best choice of media.

Table 6
P-values for tests of the null-hypothesis that the variables belonging to the test medium do not contribute to the discrimination

Base media	Distance	Test medium		
		YES	OAT	CYA
YES & OAT & CYA	mel–pol	20%	58%	89%
YES & OAT	mel–pol	<1%	<1%	-
YES & CYA	mel–pol	<1%	–	5%
OAT & CYA	mel–pol	–	2%	10%
YES & OAT & CYA	mel–ven	44%	26%	12%
YES & OAT	mel–ven	17%	16%	-
YES & CYA	mel–ven	<1%	-	<1%
OAT & CYA	mel–ven	–	<1%	<1%
YES & OAT & CYA	pol–ven	18%	84%	40%
YES & OAT	pol–ven	<1%	6%	-
YES & CYA	pol–ven	<1%	-	<1%
OAT & CYA	pol–ven	–	<1%	<1%

Two species are regarded at a time and the tests were based on the first 7 PCs.

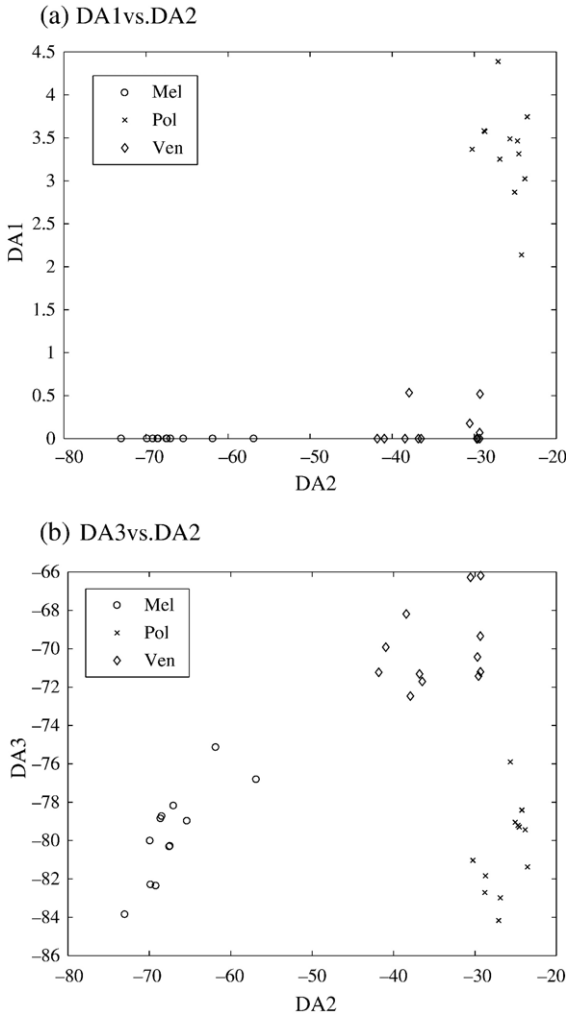


Fig. 4. Scatter plots of DA1 versus DA2, the variables chosen using leave-one-out cross validation, and DA3 versus DA2, the variables chosen using test set validation. Only the YES medium was used here.

We performed objective classification using only YES and using the combination of YES and CYA.

3.3. Classification

The same results were obtained on the YES medium as on the combination of YES and CYA. Forward selection based on Wilks' Lambda was used to select the most discriminatory variables. With both leave-one-out cross-validation and test set validation, 2 variables were sufficient to obtain a 100% correct classification. In Fig. 4 the observations are plotted for the

Table 7
The three variables selected according to Wilks' Λ in the discriminant analysis on the YES medium

Var	DA1	DA2	DA3
Image	Difference	Difference	Difference
Parameter	99th percentile	30th percentile	10th percentile
Bands (nm)	cyan & NIR (470&940)	ultra blue & red (430&645)	ultra blue & NIR (430&870)

chosen variables. Note, that the same variables were chosen if CYA was also included in the analysis. The variables are described in Table 7.

On CYA or OAT the results were not as good as on YES. They yielded, respectively, 2 and 1 misclassifications when 10 variables were selected and leave-one-out cross-validation was used.

4. Conclusion and future work

In this paper we have suggested and assessed a sequence of multivariate statistical methods to detect which medium or combination of media is most suitable to discriminate between different species of fungi.

Multivariate analyses of variance implied that the experiment can be reproduced with similar results both with the same and with different isolates within the three species.

According to the statistical tests of information, three media did not provide additional information compared to two media. Furthermore, the best choice of a pair of media was YES and CYA. However, in practice, one medium was sufficient and YES, on which the Mahalanobis distances between species in general were largest, gave the best results.

An objective identification was performed based solely on image analysis. The use of several spectral bands is new and valuable in comparison with the method of Dörge et al. (2000) as it provides additional information about the chemistry of the fungal colonies. The objective characterization of the fungi is an alternative to identification based on molecular methods (Samson et al., 2004). The three species: *P. melanoconidium*, *P. polonium*, and *P. venetum* were identified objectively with 100% correctness from just four spectral bands and two variables on the YES medium.

It is of interest to validate the method for testing the media on a larger study, possibly of another fungal genus, with more isolates and more media. Subsequently, make an even larger study, including more isolates only on the best media, and hereby validating the method of objective identification.

References

Christensen, M., Miller, S.L., Tuthill, D., 1994. Color standards — a review and evaluation in relation to Penicillium taxonomy. Mycol. Res 98, 635–644.

Dörge, T., Carstensen, J.M., Frisvad, J.C., 2000. Direct identification of pure *Penicillium* species using image analysis. J.Microbiol. Methods 41, 121–133.

Duda, R.O., Hart, P.E., Stork, D.G., 2001. Pattern Classification. John Wiley & Sons.

Frisvad, J.C., Samson, R.A., 2004. Polyphasic taxonomy of *Penicillium* subgenus *Penicillium*: a guide to identification of food and air-borne terverticilliate penicillia and their mycotoxins. Stud. Mycol 49, 1–173.

Frisvad, J.C., Smedsgaard, J., Larsen, T.O., Samson, R.A., 2004. Mycotoxins, drugs and other extrolites produced by species in *Penicillium* subgenus *Penicillium*. Stud. Mycol 49, 201–242.

Frisvad, J.C., Nielsen, K.F., Samson, R.A., 2006. Recommendations concerning the chronic problem of misidentification of mycotoxigenic fungi associated with foods and feeds. Adv. Exp. Med. Biol 571, 33–46.

Gomez, D.D., Clemmensen, L.H., Ersbøll, B., Carstensen, J.M., 2006. Precise acquisition and unsupervised segmentation of multi-spectral images. Special Issue on Advances in Vision Algorithms and Systems Beyond the Visible Spectrum for Computer Vision and Image Understanding Accepted.

Hansen, M.E., Carstensen, J.M., 2003. Color-based image retrieval from high-similarity image databases. Lecture Notes in Computer Science 2749, 1089–1105 IMM, DTU.

Hansen, M.E., Carstensen, J.M., 2004. Density-based retrieval from high-similarity image databases. Pattern Recognit 37, 2155–2164.

Hansen, M.E., Lund, E., Carstensen, J.M., 2003. Visual clone identification of *Penicillium* species using image analysis. J. Microbiol. Methods 52, 221–229.

Hastie, T., Tibshirani, R., Friedman, J., 2001. The Elements of Statistical Learning. Springer.

Jolliffe, I.T., 1986. Principal Component Analysis. Springer, New York.

Pitt, J.I., 1979a. The genus *Penicillium* and its Telemorphic States *Eupenicillium* and *Talaromyces*. Academic Press.

Pitt, J.I., 1979b. The Genus *Penicillium* and its Teleomorphic States Eupencillium and Taleromyces. Academic Press, London.

Ramirez, C., 1982. Manual and Atlas of the Penicillia. Elsevier, Amsterdam.

Rand, T.G., Giles, S., Flemming, J., Miller, J.D., Puniani, E., 2005. Inflammatory and cytotoxic responses in mouse lungs exposed to purified toxins from building isolated *Penicillium brevicompactum* dierckx and P. chrysogenum. Thom. Toxicol. Sci. 87, 213–222.

Raper, K.B., Thom, C., 1949. Manual of the Penicillia. Williams & Wilkins.

Rencher, A.C., 2002. Methods of Multivariate Analysis. John Wiley & Sons, New York.

Samson, R.A., Frisvad, J.C., 1993. New taxonomic approaches for identification of food-borne fungi. Int. Biodeterior. Biodegrad. 32, 99–116.

Samson, R.A., Hoekstra, E.S., Frisvad, J.C., Filtenborg, O., 2000. Introduction to Food and Airborne fungi6th edition. Centraalbureau voor Schimmelcultures, Utrecht.

Samson, R.A., Seifert, K.A., Kuijper, A.F.A., Houbraken, J.A.M.P., Frisvad, J.C., 2004. Phylogenetic analysis of *Penicillium* subgenus using partial β-tubulin sequences. Stud. Mycol 49, 175–200.