

Characterization of Pollen by Vibrational Spectroscopy

BORIS ZIMMERMANN

Department of Organic Chemistry and Biochemistry, Ruđer Bošković Institute, Bijenička 54, 10000 Zagreb, Croatia

Classification, discrimination, and biochemical assignment of vibrational spectra of pollen samples belonging to 43 different species of the order *Pinales* has been made using three different vibrational techniques. The comparative study of transmission (KBr pellet) and attenuated total reflection (ATR) Fourier transform infrared (FT-IR) and FT-Raman spectroscopies was based on substantial variability of pollen grain size, shape, and relative biochemical composition. Depending on the penetration depth of the probe light, vibrational techniques acquire predominant information either on pollen grain walls (FT-Raman and ATR-FT-IR) or intracellular material (transmission FT-IR). Compared with the other two methods, transmission FT-IR obtains more comprehensive information and as a result achieves superior spectral identification and discrimination of pollen. The results strongly indicate that biochemical similarities of pollen grains belonging to the same plant genus or family lead to similar features in corresponding vibrational spectra. The exploitation of that property in aerobiological monitoring was demonstrated by simple and rapid pollen identification based on relatively small spectral libraries, with the same (or better) taxonomic resolution as that provided by optical microscopy. Therefore, the clear correlation between vibrational spectra and pollen grain morphology, biochemistry, and taxonomy is obtained, while successful pollen identification illustrates the practicability of such an approach in environmental studies.

Index Headings: Fourier transform infrared spectroscopy; FT-IR spectroscopy; FT-Raman spectroscopy; Attenuated total reflection; ATR; Pollen; Aeroallergen; Identification; Classification.

INTRODUCTION

Identification of pollen grains, the prime focus of palynology, is incorporated in various scientific fields, from plant taxonomy, phenology, aeroallergens monitoring, forensics, and honey analysis to paleobotany and paleoecology. It was established in the early 20th century as a scientific method based on microscopic analysis of morphological features of pollen grains. Despite the tremendous technological advances of the last hundred years, palynological studies are still heavily dependant on optical (light) microscopy, although considerable advance in resolution was gained by the introduction of the scanning electron microscope (SEM).¹ For that reason current pollen identification requires laborious visual recognition of pollen from each pollen taxon by a qualified specialist with specific knowledge of pollen morphology. This shortcoming is especially noticeable in aerobiological monitoring of allergenic pollen grains that cause seasonal allergic rhinitis (hay fever) and asthma. During pollination season these airway diseases affect roughly a quarter of the population, causing multibillion costs in medical treatments and reduced work productivity.^{2–5} Data on aeroallergens are required to track and manage inhalant allergies through symptom correlation and timely hay fever forecasts for avoidance and prevention. What is more, qualitative and quantitative analysis of airborne pollen can provide valuable phenological data in studies of environmental and climate change (for example, monitoring of invasive

species or annual change of forest biodiversity). An automated black-box approach with an inexpensive and rapid identification method is therefore highly desirable.

In recent years there have been a growing number of experimental works dedicated to an alternative and more objective pollen identification method based on vibrational spectra of pollen samples.^{6–14} Vibrational spectra are already used daily in identification of different mixtures (pharmaceuticals, drugs, paintings, coatings, explosives, etc.) through the well-developed application of spectral libraries,¹⁵ and an analogous procedure was implemented on pollen analysis. As opposed to morphological characterization of pollen grains with microscopy, vibrational spectroscopy offers a bias-free approach based on chemical characterization via identifiable spectral features. Vibrational pollen measurements ranged from various Raman and Fourier transform infrared (FT-IR) spectroscopies, including Raman and FT-IR microspectroscopy,^{6,9–11} diffuse reflection FT-IR (DRIFTS),⁸ and single-reflection attenuated total reflection FT-IR (SR-ATR).⁷ Because of their ability to objectively identify a single pollen grain microspectroscopies deserve special notation among these techniques.

The main practical requirement of a useful spectroscopic pollen identification method is that spectral features that characterize pollen samples of one specific taxon must simultaneously be distinguishable from other taxon possibilities. The previous studies clearly demonstrated that vibrational spectra of pollen can be used for discrimination between species that belong to the same genus, something that cannot be achieved by microscopic analysis.^{6,8,9} However, it must be stated that, due to the rather small number of recorded samples (approximately ten pollen taxa per study) and quite limited number of different species within the same genus (two to four species), it is far too early to make the general conclusion that vibrational spectroscopy can reliably identify any pollen sample with taxonomic resolution at the species level. For example, microspectroscopy studies had indicated that intra-genus and intra-species spectral variability can significantly differ from one pollen genera to another.^{6,9} The more pressing issue than the exact level of taxonomic resolution of vibrational spectroscopy is its applicability in real-world measurements. In order to establish itself as a viable alternative to optical microscopy, spectroscopy-based pollen identification must be simple, rapid, and economical. Unfortunately, although microspectroscopic techniques have obtained the most promising results, they achieved this by performing time-consuming and impractical measurements on rather expensive instruments, and their applicability is, therefore, limited to a small number of highly valuable samples.^{8,9} Spectroscopic measurement of macroscopic pollen samples, on the other hand, is still an option for rapid everyday pollen analyses. In fact, due to complete digitalization of the obtained spectra, data analysis can be carried out at a different physical location than data acquisition (as opposed to optical microscopy). Regarding the

Received 10 June 2010; accepted 17 September 2010.
E-mail: bzimmer@irb.hr.

TABLE I. The plant taxonomy of the analyzed pollen samples (*Pinales* order).

Family	Genus	Species	Name
Cupressaceae	<i>Platycladus</i>	<i>orientalis</i>	Chinese Arborvitae
	<i>Chamaecyparis</i>	<i>lawsoniana</i>	Lawson's Cypress
		<i>obtusata</i>	Hinoki Cypress
	<i>Juniperus</i>	<i>chinensis</i>	Chinese Juniper
		<i>sabina</i>	Savin Juniper
		<i>phoenicea</i>	Phoenicean Juniper
		<i>oxycedrus</i>	Prickly Juniper
		<i>communis</i>	Common Juniper ^a
		<i>virginiana</i>	Eastern Juniper ^a
	<i>Calocedrus</i>	<i>decurrens</i>	California Incense Cedar
	<i>Cupressus</i>	<i>sempervirens</i>	Mediterranean Cypress
		<i>x leylandii</i>	Leyland Cypress
	<i>Cunninghamia</i>	<i>lanceolata</i>	China Fir
Taxaceae	<i>Thujopsis</i>	<i>dolabrata</i>	Thujopsis ^a
	<i>Taxus</i>	<i>baccata</i>	European Yew
Cephalotaxaceae	<i>Cephalotaxus</i>	<i>harringtonia</i>	Japanese Plum Yew
	<i>Torreya</i>	<i>californica</i>	California Torreya
Pinaceae	<i>Tsuga</i>	<i>canadensis</i>	Eastern Hemlock
	<i>Abies</i>	<i>koreana</i>	Korean Fir
		<i>pinsapo</i>	Spanish Fir
		<i>cephalonica</i>	Greek Fir
	<i>Picea</i>	<i>abies</i>	Norway Spruce
		<i>asperata</i>	Dragon Spruce
		<i>omorika</i>	Serbian Spruce
		<i>orientalis</i>	Caucasian Spruce
		<i>pungens</i>	Blue Spruce
		<i>smithiana</i>	Morinda Spruce ^a
	<i>Pinus</i>	<i>pinaster</i>	Stone Pine
		<i>mugo</i>	Mountain Pine
		<i>syvestris</i>	Scots Pine
		<i>tabulaeformis</i>	Chinese Red Pine
		<i>banksiana</i>	Jack Pine
		<i>pinaster</i>	Maritime Pine
		<i>densiflora</i>	Japanese Red Pine
		<i>nigra</i>	European Black Pine
		<i>ponderosa</i>	Ponderosa Pine
		<i>resinosa</i>	Red Pine
		<i>wallichiana</i>	Himalayan pine
		<i>bungeana</i>	Lacebark Pine
		<i>strobus</i>	Eastern White Pine
		<i>peuce</i>	Macedonian Pine
		<i>heldreichii</i>	Bosnian Pine
		<i>neritfolius</i>	Brown Pine
Podocarpaceae	<i>Podocarpus</i>		

^a Recorded in the transmission FT-IR measurements only.

vibrational techniques, it seems that FT-IR has more potential than Raman, mainly because strong fluorescence in Raman measurements (with visible excitation lasers) considerably prolongs spectral acquisition.^{6,10,14} All things considered, FT-IR and FT-Raman (with near-infrared laser excitation) spectrometers are the only prospective alternative to optical microscopy, though firm correlation between laboratory experiments and real-world measurements has not yet been established.

The following report describes the detailed vibrational measurements by means of transmission (KBr pellets), reflection FT-IR (SR-ATR), and FT-Raman spectroscopies (Nd:YAG laser source) of pollen samples belonging to 43 different species of the order *Pinales* (Table I). The methods are accommodated for powder samples and therefore enable straightforward and rapid spectral acquisition (pollen samples were measured as collected in nature, without prior drying or freezing). Furthermore, because these methods measure the same phenomenon, vibrational frequencies of samples, the acquired data offers better biochemical characterization of

pollen samples as well as double verification and reexamination, thus making the measurement more error-tolerant. The pollen samples for this research were chosen with the aim of obtaining taxonomic resolution of the method based on a statistically significant number of different taxa and their good distribution and representation at the family and genus level. Pollen of the plant order *Pinales* offers substantial variability in grain size, shape, and relative biochemical composition and is therefore an appropriate system for testing and verifying the mentioned vibrational methods. The study was focused on defining the appropriate criteria for spectral analysis and identification of pollen samples. This research obtains clear correlation between vibrational spectra, pollen biochemistry, and plant taxonomy. It also offers a broader perspective into pollen spectral variations between closely related species (within conifer families and genera). Lastly, the successful pollen analysis illustrates the practicability of such an approach in environmental studies and monitoring of aeroallergens.

MATERIALS AND METHODS

Materials. Samples of pollen belonging to 43 different species were collected during their pollination season at two facilities of the University of Zagreb, the Botanical Garden of the Faculty of Science and the Botanical Garden "Fran Kušan" of the Faculty of Pharmacy and Biochemistry. Pollen samples were stored at +5 °C (without drying or storing in a desiccator) and IR spectra of their KBr sample pellets were recorded within 48 h. The same technique was used for measurements of stability and reproducibility of pollen samples of *Pinus sylvestris*, *Pinus ponderosa*, *Abies koreana*, *Taxus baccata*, *Cupressus sempervirens*, *Tsuga canadensis*, and *Picea abies*. These samples were considered representative for the entire collection and their spectra were recorded every few months. Since no significant changes in spectral appearances were detected during the period of ten months (when compared to the samples recorded within 24 h of collection) it can be concluded that the outer wall of pollen grains indeed provides sufficient chemical stability of the grain composition that vibrational spectral features remain almost constant over a substantial time frame.¹⁶ ATR and Raman spectra of pollen samples were recorded approximately two months after their collection in nature. No less than three spectra per employed vibrational technique were recorded for every sample (species).

Vibrational Measurements. The IR spectra were recorded at a resolution of 4 cm⁻¹ on an ABB Bomem MB102 single-beam FT-IR spectrometer, equipped with CsI optics and a DTGS detector. The KBr sample pellets were prepared by mixing ~1 mg of a sample with approximately 100 mg of KBr using a pestle and mortar. The IR spectra of the sample pellets were recorded with a total of 10 scans, and transmission spectra thus obtained are termed raw spectra. In general, a transmission IR pollen spectrum of good quality is easily obtained with even relatively small numbers of interferograms, which enables more than satisfactory signal-to-noise ratio. A sample-free setup was used to obtain the spectral background, as opposed to the KBr background used in previous transmission measurements.⁸ Because even carefully prepared pellets made only of KBr are not optically reproducible due to transparency variations, the KBr background can degrade spectral data by affecting the spectrum baseline, causing it to change from constant offset and slope to various curvatures.

The ATR-IR spectra were recorded with a total of 30 scans using a horizontal single-reflection ATR diamond prism with 45° angle of incidence. Each spectrum was recorded as the ratio of the sample spectrum to the spectrum of the empty ATR plate.

The Raman spectra in backscattering geometry were recorded on an FT-Raman FRA 106/S model, coupled with a Bruker Equinox 55 IR spectrometer, equipped with a Nd:YAG laser (1064 nm, 9394 cm⁻¹, and germanium detector cooled with liquid nitrogen). The spectra were recorded at a resolution of 4 cm⁻¹ with a total of 64 scans using Blackman-Harris 4-term apodization and with a laser power of 400 mW.

Data Treatment. The spectral data treatment, the hierarchical cluster analysis (HCA), the principal component analysis (PCA), and the spectral search were performed using programs written for Matlab 6.5 (The MathWorks, Inc., Natick, MA). In order to avoid personal bias and over manipulation of the spectral data, only the simplest baseline correction, assuming linear baseline, was carried out. The baseline-corrected spectra were obtained by subtracting a straight line defined by the points $\alpha(\tilde{\nu}_1)$ and $\alpha(\tilde{\nu}_2)$, where $\tilde{\nu}_1$ and $\tilde{\nu}_2$ are the limits of the analyzed spectral region (1900 and 405 cm⁻¹ in the transmittance IR spectra, 1800 and 800 cm⁻¹ in the ATR IR spectra, and 1800 and 700 cm⁻¹ in the Raman spectra). The spectra were smoothed using the Savitzky-Golay algorithm (11-point moving second-degree polynomial). The normalized spectra were obtained by applying unit area normalization.

Lastly, the first-derivative spectra were obtained in order to enhance spectral variations.¹⁷ The first-derivative spectra were obtained as the first derivative of absorbance (or Raman intensity) as a function of wavenumber. The described spectral data treatment that involves baseline correction, smoothing, normalization, and derivation is standard for all spectral search procedures¹⁵ and was in some form implemented in all previous vibrational measurements of pollen. Such data analysis can be fully automated and obtained with a low-budget PC in a matter of seconds.

For the additional analyses at the level of family and genera the complete spectral data set (*Pinales*) was separated into smaller subsets. These data subsets are termed according to plant taxonomy (*Pinacea* (excluding *Tsuga canadensis* spectral data), *Pinus*, and *Juniperus* subsets).

Evaluation of similarities between the first-derivative spectra (three spectra per species) was performed by using PCA and HCA. HCA and PCA are methods that allow unbiased (i.e., unsupervised) classification of spectral data. HCA is highly descriptive categorical method that provides strict classification of a sample (pollen spectrum) in a particular class (cluster). The HCA plots (dendrograms) were calculated with Euclidian distance measure and either Ward (for *Pinales* data set) or average distance clustering algorithm (for *Pinus* and *Juniperus* subsets).

Principal components analysis of spectral data set can reduce the number of original spectral variables (wavenumber dimensions) to a smaller number of a new set of variables (principal components). Graphical representations of correlations between samples, principal components, and wavenumbers allow visual detection of groups of related samples (in this case, based on plant taxonomy) and consequent identification of major spectral features that are causing this differentiation.

The reduced spectral data that retained only spectral regions of highly diagnostic value were termed "diagnostic spectra".

The data reduction was based on the PCAs for the *Pinales* data set and the *Pinaceae* subset. The PCAs for all three vibrational techniques revealed that differentiation between *Cupressaceae*/*Cephalotaxaceae*/*Taxaceae* and *Pinaceae*/*Podocarpaceae* samples can be based on PC 1 vector (PCAs of *Pinales* data sets), while differentiation between *Pinus*, *Picea*, and *Abies* samples can be based on PC 1 and PC 2 vectors (PCAs of *Pinaceae* data sets). The resulting PC plots (as a function of wavenumber) were analyzed visually in order to detect spectral regions with high positive or negative factor loadings. Only these spectral regions were retained from the original spectral data set while all other "redundant" data was excluded. The reduced data were obtained from the derivative spectra by keeping the following spectral regions: (1) transmittance FT-IR: 1780–1680 and 1200–800 cm⁻¹; (2) ATR FT-IR: 1770–1720, 1620–1580, 1530–1490, 1230–1070, and 860–810 cm⁻¹; and (3) FT-Raman: 1680–1560, 1520–1410, and 1240–1130 cm⁻¹. The qualitative analysis of pollen was based on root mean square error (RMSE) analysis of these diagnostic spectra. The spectral data (three spectra per species) was separated into five taxonomic groups: (1) *Cupressaceae* (14 species in transmittance FT-IR, 11 species in ATR-FT-IR and FT-Raman), (2) *Cephalotaxaceae* and *Taxaceae* (three species), (3) *Abies* (three species), (4) *Picea* (six species in transmittance FT-IR, five species in ATR-FT-IR and FT-Raman), and (5) *Pinus* (15 species). The spectral library containing approximately 20% of randomly selected spectra from each taxonomic group was created (in total 21 spectra: nine spectra of *Pinus*, three *Picea*, one *Abies*, one *Cephalotaxaceae* or *Taxaceae*, and seven *Cupressaceae*). The remaining 80% of spectra were treated as "unknown" (test spectra) for verification of the spectral method (this approach was repeated ten times per spectral method). The spectral search designated a test spectrum into one of the five taxonomic groups by comparing a test spectrum with the library spectra (i.e., calculating RMSE). The spectral search was termed "confident" if the ratio between the smallest RMSE and the second smallest RMSE that belongs to some other taxonomic group was less than 0.7; otherwise the result was termed as "questionable".

The spectral identification of the pollen samples of *Pinus nigra* and *Pinus sylvestris* (the samples of each species were collected from three different plants) was based on the RMSE analysis of the derivative transmission IR spectra (spectral region: 1900–405 cm⁻¹). Three KBr pellets were recorded for each sample, and after the data treatment (baseline correction, Savitzky-Golay smoothing, and normalization), the average of three spectra was calculated (one average spectrum per sample) and its first derivative was obtained. The spectral library containing randomly selected spectra (one per species) was created, and the remaining spectra were treated as the test spectra (this approach was repeated for all nine variations of the spectral library). The spectral search designated a test spectrum into either the *Pinus nigra* or *Pinus sylvestris* category.

RESULTS AND DISCUSSION

Microscopy measurements of pollen grains seldom result in identification of the sample species. Due to negligible morphological differences between pollen samples of the related species (even when morphological features are obtained via SEM), pollen grains are rarely identified to the species

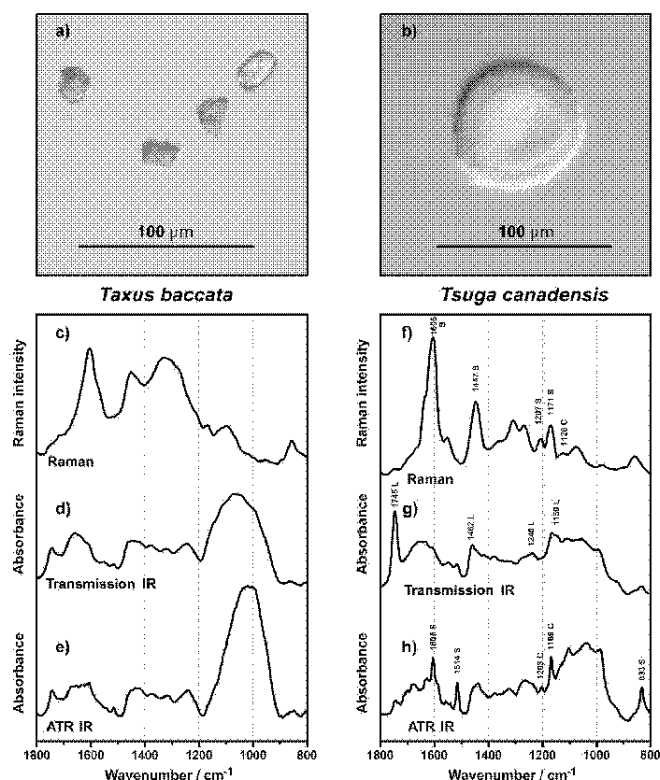


FIG. 1. Optical microscope images of (a) *Taxus baccata* and (b) *Tsuga canadensis* pollen grains. Vibrational spectra of *Taxus baccata* pollen: (c) Raman, (d) transmission IR, and (e) ATR-IR spectra; vibrational spectra of *Tsuga canadensis* pollen: (f) Raman, (g) transmission IR, and (h) ATR-IR spectra (S = sporopollenin, C = cellulose, L = lipids vibrational bands).

level, and some may have to be placed in groups of multiple genera or families. For example, in the conventional hay fever forecasts, a pollen grain belonging to the order *Pinales* is classified into either the *Pinaceae* group or the *Cupressaceae/Taxaceae* group. Therefore, this research was focused on obtaining simple and rapid spectral identification of pollen samples with equal or better level of classification.

Comparative Study of Vibrational Methods. The spectroscopic and the morphological features of the typical pollen grains belonging to the order *Pinales* are shown in Fig. 1. Although all three spectroscopic techniques measure analogous physical phenomenon, there are some fundamental differences in spectral acquisition that lead to distinct results. For example, in transmission FT-IR measurements a whole grain interacts with IR radiation, while in the case of ATR measurements, depending on the penetration depth of the IR light, an external part of a grain is in greater interaction with the light than the grain interior. Bearing in mind that this penetration depth is in the 0.5 to 5 μm range, and that the diameter of pollen grains can vary from less than 10 μm to more than 100 μm , with substantial biochemical variations throughout the radii,⁶ it is perfectly clear that the obtained spectra from these two techniques can not be alike. As expected, better agreement between transmission and ATR spectra is gained for smaller pollen grains, such as *Taxus baccata* (European yew) where the light penetrates proportionally deeper than for larger grains (Figs. 1d and 1e). Such results are characteristic for the pollen samples belonging to *Cupressaceae* and *Cephalotaxaceae* families as well. Contrary to that, large discrepancies are

noticeable in the spectra of the pollen samples belonging to the *Pinaceae* family, whose grains are seldom smaller than 70 μm . These spectral differences are evident comparing the transmission and the ATR spectra of *Tsuga canadensis* (eastern hemlock), with especially noticeable extreme disparity of the carbonyl band intensity (1745 cm^{-1}) (Figs. 1g and 1h). It must be emphasized that commercially available ATR correction algorithms that can convert ATR into transmission IR spectra cannot be implemented in this case since those methods can cope only with light penetration depth variations through a homogeneous sample and not with biochemically heterogeneous samples such as pollen grain.

The application of FT-Raman spectroscopy with near-infrared (NIR, 1064 nm) laser excitation (Nd:YAG laser source) resulted in good quality spectra devoid of the marked fluorescence background that hampered previous Raman measurements of pollen with excitation in the visible region (Figs. 1c and 1f).^{6,10} Consequently, the FT-Raman measurements were considerably faster (up to 20 times) than the mentioned Raman measurements, which required extended photo-bleaching of the specimen prior to the spectral acquisition. Unfortunately, there is a lack of some strong vibrational bands in the Raman spectra that can be attributed specifically to intracellular lipids, as is the case with the carbonyl band at 1745 cm^{-1} in the IR (the strong CH_2 deformation band at $1440\text{--}1460\text{ cm}^{-1}$ is common not only in Raman spectra of lipids but also in spectra of proteins and carbohydrates). Moreover, analogous to the ATR technique, an external part of a grain is excessively represented in a Raman spectrum due to a greater probability of multiplicative scattering in the grain interior. For all these reasons the Raman measurement has disproportional bias toward pollen grain wall.

Although the correlation between vibrational spectral features (band intensity and peak wavenumber) and chemical composition (functional group) is quite straightforward, for extremely complicated biochemical systems such as pollen grain this task is unattainable. Similar to other biological systems, pollen grains are composed of water, lipids, proteins, nucleic acids, and carbohydrates. To make things worse, grains have a tough outer wall (exine) made of sporopollenin, a complex biopolymer containing phenylpropanoids, phenolics, fatty acids, and carotenoids.^{18–21} For that reason the vibrational bands can only be tentatively assigned to some basic biochemical groups of compounds,^{22–25} and even that should be made with great caution. Approximate chemical maps that reveal biochemical heterogeneity within pollen grain were obtained recently by spatially resolved Raman microspectroscopy measurements.⁶

The Raman and the ATR spectra of *Tsuga canadensis* pollen clearly show dominant profiles of vibrational bands associated with exine (sporopollenin) and intine (an inner thin cellulose grain wall) (Figs. 1f and 1h). The sporopollenin bands (ATR IR: 1605 , 1512 , and 833 cm^{-1} ; Raman: 1605 , 1447 , 1207 , and 1171 cm^{-1}) can be associated with the vibrations of aromatic rings, while cellulose bands (ATR IR: 1203 , 1169 , and 895 cm^{-1} ; Raman: 1126 cm^{-1}) can be associated with the vibrations of pyranose units of the cellulose skeleton. These bands have less pronounced relative intensity in the transmission spectra where bands associated with intracellular lipids dominate, such as the carbonyl band at 1745 cm^{-1} (Fig. 1g). Therefore, by measuring vibrational spectra of relatively small (*Cupressaceae*) and large pollen grains (*Pinaceae*) this

research offers a comparative study between the two complementary vibrational spectroscopies (FT-Raman and FT-IR), as well as between the two FT-IR techniques that measure different parts of grains (superficial part with ATR and the whole grain with transmission FT-IR).

The pollen samples of the order *Pinales* offer one additional feature that can be experimentally measured. Besides considerable size variations of grains within this plant order there are also considerable morphological differences. Pollen grains of the *Pinaceae* and *Podocarpaceae* families are described as saccate pollen due to a large hollow projection (saccus) from the central body of pollen grain composed only by the exine. *Abies* (fir), *Picea* (spruce), and *Pinus* (pine) grains have two sacci, giving them the characteristic “Mickey Mouse” silhouette in the equatorial view of optical microscopy (Figs. 2a and 2b). *Tsuga* (hemlock) pollen has only a single saccus that encircles the central body of the pollen grain (Fig. 1b). The presence of saccus, therefore increases the relative amount of sporopollenin in *Pinaceae* and *Podocarpaceae* pollen grains and for that reason the vibrational bands associated with the sporopollenin dominate to a larger extent in their ATR spectra than is the case with *Cupressaceae* spectra.

Since the two FT-IR techniques measure either the superficial part of pollen grain (ATR) or the whole grain (transmission) it is quite interesting to notice how this can affect the results. Of the three measured *Abies* taxa, the transmission IR spectrum of *Abies koreana* (Korean fir) (Fig. 2c) is more closely related to the spectrum of *Tsuga canadensis* (Fig. 1g) than to the two *Abies* taxa, mostly owing to the strong carbonyl stretching vibration. A closer look at the differential spectrum (the difference between *Abies koreana* and *Abies pinsapo* transmission spectra) that highlights the spectral differences between these two species reveals that principal spectral variations can be attributed to the difference in relative quantity of intracellular lipids (Figs. 2c, 2e, and 2g). The lipids are characterized by the vibrational bands at 1747 (C=O st), 1462 (CH₂ def), 1236 (C–O st), and 1150 cm⁻¹ (C–O st). When comparing the ATR spectra however, the three *Abies* pollen samples have such similar spectral features that the ones belonging to *Abies koreana* and *Abies pinsapo* (spanish fir) cannot even be discriminated (Figs. 2d, 2f, and 2g). The reason for such poor distinction when one applies the ATR-FT-IR technique lies in the mentioned inability of the IR light to penetrate deep enough for considerable interaction with intracellular lipids, and since the superficial biochemicals do not offer any specific signals for each *Abies* species, the exact identification can not be made. The situation with ATR spectroscopy is in this case comparable to optical microscopy, in which morphological features of various *Abies* pollen grains are not specific enough to allow taxonomic resolution below the level of genus (Figs. 2a and 2b). On the other hand, morphological and ATR spectral differences of pollen grains between *Tsuga* and *Abies* genus are quite obvious (Figs. 1b, 2a, and 2b).

The measurements of pollen samples have shown that their ATR-IR spectra are of a good quality and of better spectral reproducibility than the transmission IR data. Because of sample preparation in KBr matrix (KBr sample pellet), spectral reproducibility of the transmission IR spectra is slightly worse than is the case with the ATR spectra. Due to structural imperfections and geometry variations of the sample pellets, the optical properties of the pellets significantly vary even if the

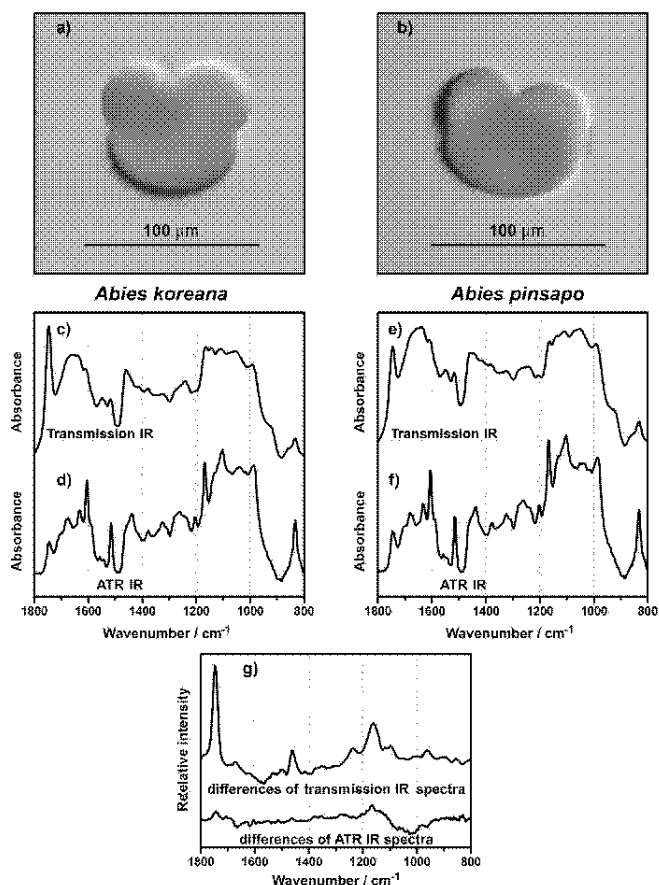


FIG. 2. Optical microscope images of (a) *Abies koreana* and (b) *Abies pinsapo* pollen grains. Vibrational spectra of *Abies koreana* pollen: (c) transmission IR and (d) ATR-IR spectra; vibrational spectra of *Abies pinsapo* pollen: (e) transmission IR and (f) ATR-IR spectra; (g) spectral differences of *Abies koreana* and *Abies pinsapo* transmission and ATR-IR spectra.

quantity of KBr and pollen is kept constant. This all causes noticeable baseline variations in the raw infrared spectra, resulting in a somewhat inferior spectral reproducibility. An additional source of concern is the presence of water in the sample pellet due to hygroscopicity of KBr, which results in superposition of water and sample spectra. These problems can be successfully tackled with consistent sample preparation and careful data treatment (constant KBr quantity, baseline correction, and use of derivative spectra). On the other hand, spectral reproducibility of the Raman spectra is significantly inferior when compared with the IR measurements.

Multivariate Analysis of Pollen Spectra. Both multivariate methods (HCA and PCA) provide rapid examination of the recorded data and their comparison with the taxonomy of the *Pinales* order. The HCA of the spectral data obtained with FT-IR methods results in clean partition into two groups: (1) *Cupressaceae*, *Cephalotaxaceae*, and *Taxaceae*, and (2) *Pinaceae* and *Podocarpaceae* (Fig. 3 and Fig. S1, available online in the Supplemental Material). A closer look at the HCA results reveals that the pollen samples are also rather well separated at the genus level. Although this is not that obvious in the case of the *Cupressaceae* family due to a quite small number of measured species for each genus (only the *Juniperus* genus had more than two measured species), it is undeniable that *Pinaceae* samples are clustered according to their genus in

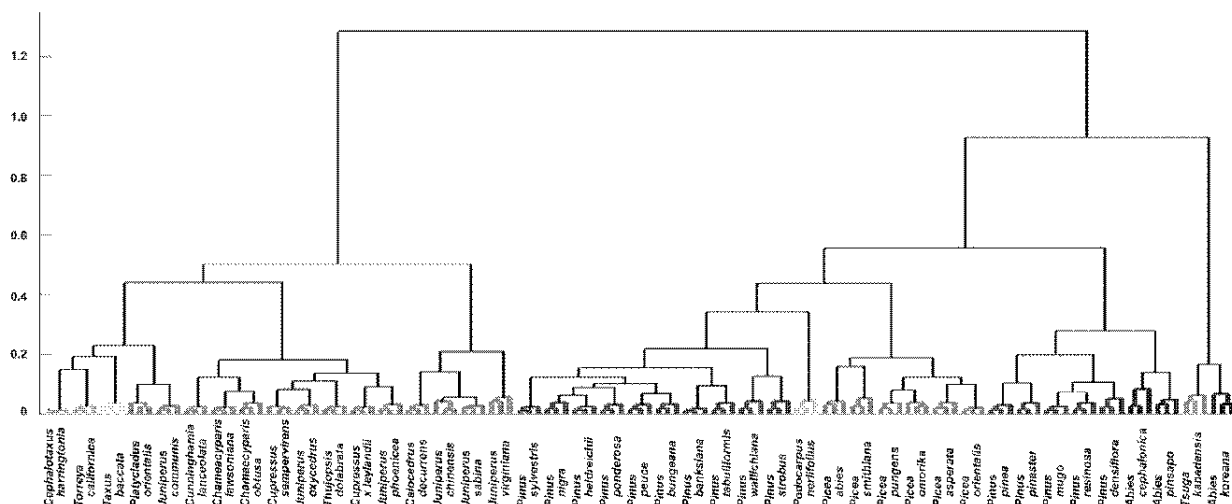


FIG. 3. The hierarchical cluster analysis dendrogram of the transmission IR spectra of the complete spectral data set (Ward distance clustering algorithm).

the *Pinus*, *Picea*, and *Abies* groups. The spectral data is generally of good reproducibility with distinct spectral features for each species; thus, it is theoretically possible to identify pollen species by its IR spectrum.

The previously reported study conducted on a Raman microspectrometer already revealed that there is no general rule for intra-genus spectral differentiation (for example, while the three measured *Acer* (maple) species are distinguishable from one another, the same does not apply for the two *Quercus* (oak) species).⁶ It should be emphasized that this analysis was based on the entire Raman spectral signature (380–1700 cm^{-1}), whereas the use of possibly more informative spectral sub-regions was not investigated. Taking into account that some plant genera (such as *Quercus*) can contain several hundred species, it is unrealistic to expect that each and every species will possess some characteristic spectrum. It can be anticipated that some more closely related species share identical spectral features, especially considering that a number of plant species are the result of hybridization. Such tendency is visible in the already mentioned example of the transmission IR spectra belonging to pollen of the three *Abies* species. Figure 2g depicts substantial spectral differences between the spectra of Korean and Spanish fir, and analogous results are obtained for differences between Korean and Greek fir. However, when the spectra of the two European fir species are compared, the spectral differences are smaller and comparable with the spectral variation within the sample.

The HCAs on *Juniperus* (six species) and *Pinus* (15 species) data subsets were conducted in order to obtain rough quantification of these intra-genus spectral variations. Compared to some previous studies,^{6,9} the number of measured species in each genera was statistically more significant (approximately 10% of the total number of species belonging to these genera). The resulting dendrogram from the HCA on the *Juniperus* subset reveals clustering in accordance with plant taxonomy (Fig. 4). There are two clusters affiliated with two sections: the *Juniperus* section (*J. oxycedrus*, *J. phoenicea*, and *J. communis*) and the *Sabina* section (*J. chinensis*, *J. sabina*, and *J. virginiana*). It is apparent for this later group that the spectral differences between the two Asian species (*J. chinensis* and *J. sabina*) are much smaller then between either

of the two and the American species (*J. virginiana*). In fact, the IR spectra of the two Asian species are more or less undistinguishable and therefore the exact identification of species is improbable, especially taking into account that the hybrid between the two species is found in the wild where habitats of the two species overlap. These minute spectral differences are emphasized when the clustering algorithm is changed from Ward distance (two separate clusters, Fig. 3) to average distance (one cluster due to the outlier of *J. chinensis*, Fig. 4). The *Pinus* pollen samples show similar behavior (Fig. S2, Supplemental Material), with clustering of closely related species (*Picea* section: *P. pinea* and *P. pinaster*; *Quinquefoliae* section: *P. wallichiana* and *P. strobus*). However, spectral similarity of pollen belonging to *P. peuce* and *P. bungeana* (i.e., different subgenera) is a clear remainder that even some phylogenetically more distinct species can share similar spectral features. To conclude, the tendencies within the fir, juniper, and pine genera have undoubtedly shown that intra-

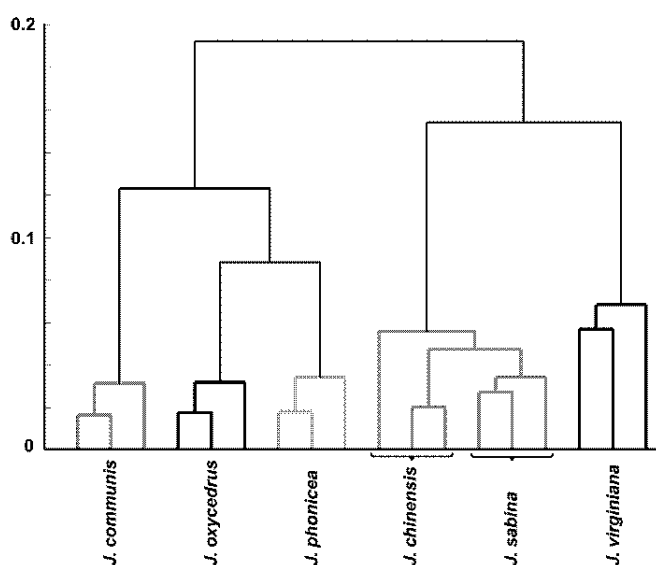


FIG. 4. The hierarchical cluster analysis dendrogram of the transmission IR spectra of the *Juniperus* subset (average distance clustering algorithm).

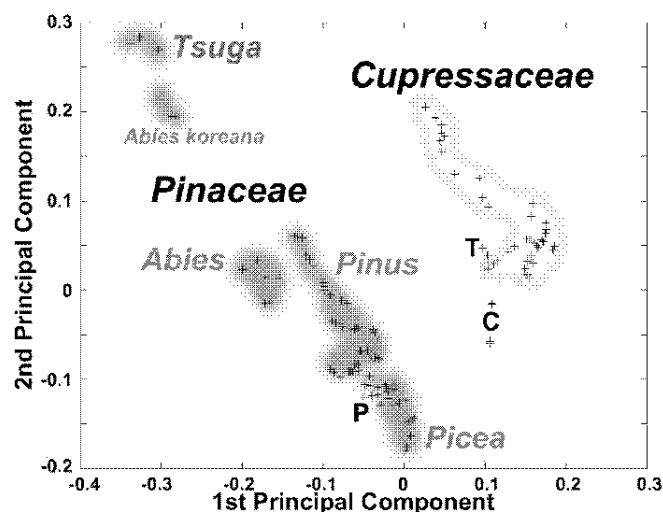


Fig. 5. The principal component analysis of the complete transmission IR spectral data set (*Pinales* order) with tentative depiction of plant families: *Pinaceae* (gray shaded), *Cupressaceae* (light gray outlined), *Cephalotaxaceae* (C), *Taxaceae* (T), and *Podocarpaceae* (P). The percent variances for the first five PCs are 43.48, 31.45, 7.98, 4.07, and 3.10.

genus spectral variations of pollen are not consistent and that they can vary significantly from one to another pollen genus depending on the species involved in the measurement. The rule of thumb in pollen identification is that closely related species have similar and in some cases even indistinguishable spectra, which is perfectly understandable and expected.

The Raman spectra, as opposed to the IR, have rather poor spectral information, mostly due to noticeable spectral noise and proportionately wider vibrational bands than is the case in IR spectra. As a result, it is hard to extract unambiguous answers from the results of the HCA of the Raman data, though some common grouping at the family level is definitely present (Fig. S3, Supplemental Material).

The PCA of transmission IR spectral data reveals the same pattern presented by the HCA, that is, separation into the two groups comprising either *Cupressaceae*/*Cephalotaxaceae*/*Taxaceae* or *Pinaceae*/*Podocarpaceae* samples with further subclustering of samples belonging to *Pinus*, *Picea*, and *Abies* genus (Fig. 5). Examination of the principal components reveals that such separation is mainly caused by spectral variations of the vibrational bands associated with lipids (intracellular material), sporopollenin (exine), and cellulose (intine). Analogous PCA of the spectral data comprising only *Pinaceae* samples reveals that the principal difference between *Pinus*, *Picea*, and *Abies* genus are the relative intensities of bands associated with lipids (Figs. 6 and 7), with the addition of a small influence of the amide I band (intracellular protein material).

The PCA of the ATR and the Raman data are in agreement with the analogous HCA results. The ATR spectral data confirmed the previously published conclusion that this technique offers reproducible and specific spectra,⁷ enabling general taxonomic resolution at the level of species (Fig. S4, Supplemental Material). The data for the *Pinaceae* family is once again grouped based on plant genus, albeit the group boundaries are not as clear-cut as for the transmission data. The Raman spectra provide excellent separation at the family level but fail to provide the clear sub-separation at the level of genus

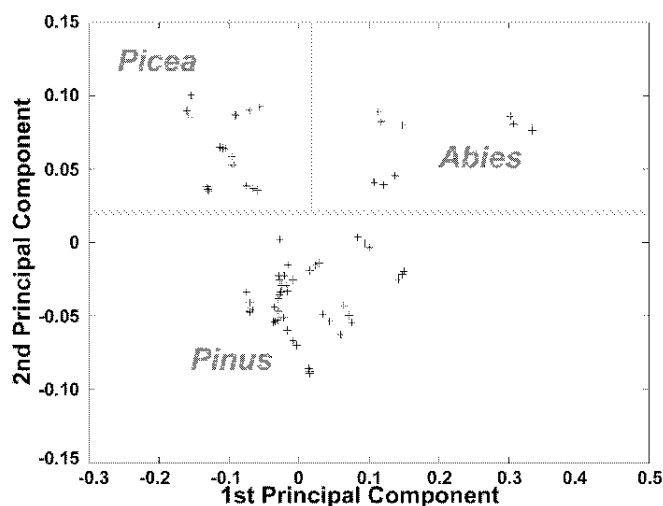


Fig. 6. The principal component analysis of the *Pinaceae* transmission IR spectral data subset with tentative depiction of plant genera. The percent variances for the first five PCs are 69.54, 13.65, 5.23, 2.61, and 1.80.

that is characteristic for the FT-IR measurements (Fig. S5, Supplemental Material). The reason is the inability of the FT-Raman method to obtain (for identification purpose) any relevant biochemical information other than the already mentioned information on sporopollenin (and to some minor extent cellulose). In contrast, the ATR-FT-IR identification is based primarily on sporopollenin, cellulose, and (albeit to a smaller extent) lipid vibrational bands. The shortcoming of insufficient IR light penetration depth is compensated in ATR-FT-IR with superb reproducibility and more than adequate signal-to-noise ratio.

An intriguing fact is that the analysis of both the ATR and the Raman data pointed to distinguishable spectral features for *Cunninghamia lanceolata* when compared with the rest of the *Cupressaceae* family, something that does not occur in the transmission measurement. The reason is the size of *Cunninghamia lanceolata* pollen grains, which is significantly larger than the typical *Cupressaceae* family grain (Fig. S6, Supplemental Material). Therefore, the ATR and the Raman spectra of this species have relatively intensive bands associated with sporopollenin, as opposed to other *Cupressaceae* samples. Because the transmission FT-IR method obtains information of the whole grain, this difference is not nearly as prominent in the analysis of the transmission data. However, the general conclusion for spectral discrepancy of this genus can not be given without additional measurements of some other pollen samples belonging to the *Cunninghamia* genus. In any case, morphological and spectral features of *Cunninghamia lanceolata* pollen are in perfect agreement and therefore present convincing proof for hereby presented spectral assignment and theoretical arguments.

Spectral Search. The vibrational spectroscopy research of pollen is predominantly focused on spectral identification of pollen species.^{6–10} However, that level of taxonomic resolution is often unnecessary and, taking into account the level of resolution required in standard aerobiological monitoring, it is generally more important to obtain reliable resolution at the level of genus. The PCA results indicate that each pollen spectrum has some common features connected to a specific plant genus and family, and these properties could be

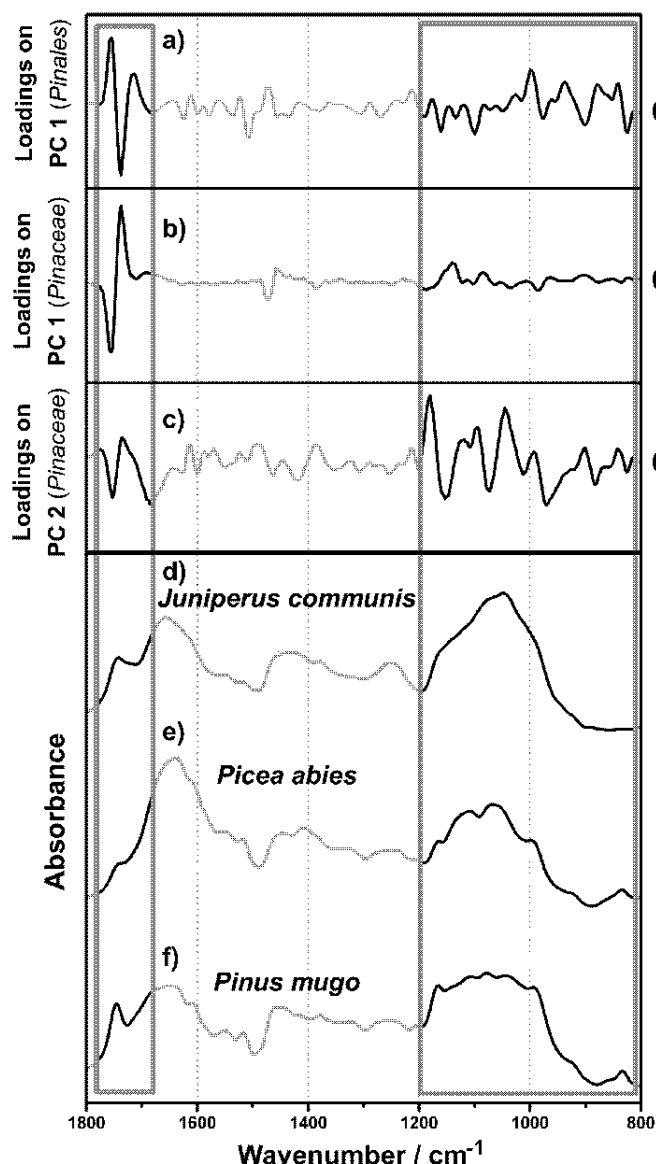


Fig. 7. Loadings plot on (a) the first principal component of the PCA of the *Pinaceae* transmission IR data set, and (b) the first and (c) second principal components of the PCA of the *Pinaceae* transmission IR data subset. Transmission IR spectra of (d) *Juniperus communis*, (e) *Picea abies*, and (f) *Pinus mugo* pollen. The spectral regions of high diagnostic value for spectral search are highlighted.

capitalized upon in quantitative and qualitative analyses of complex pollen mixtures. Therefore, the working hypothesis stated that a spectral library comprising minimal spectral data on certain pollen group (genus or family) is sufficient for a reliable identification of all pollens belonging to that group. The principal goal of spectral search was identification of pollen family based on small spectral libraries and, in the case of successful results, concomitant genus identification.

Because the detailed examination of the PCA results highlighted some spectral regions of highly diagnostic value (Fig. 7), the spectral search was implemented on the diagnostic spectra that retained only a third of the original data (i.e., reduced data obtained from the derivative spectra by keeping the diagnostic spectral regions). The spectral data was

separated into five taxonomic groups and then randomly grouped into either the spectral library or test spectra, as described in the Experimental section. It should be emphasized that the library kept the original sample ratio of the taxonomic groups and therefore was biased toward the groups with more recorded samples (each library contained only one *Abies* spectrum as opposed to three *Picea* and nine *Pinus* spectra). Generally, the testing conditions were quite severe given that the spectral libraries on average contained less than 50% of the tested pollen species (and less than 20% of the recorded spectra). The spectral search was based on the RMSE analysis of diagnostic spectra that valued not only if the prediction is true or false but also the confidence of each prediction (the ratio of the RMSE results).

The results of the spectral search showed that the level of identification attained in conventional hay fever forecasts, i.e., separation between the *Pinaceae* group and the *Cupressaceae/Taxaceae/Cephalotaxaceae* group, is easily reached with all vibrational methods (Table II). The spectral separation between these two groups can provide almost 100% accuracy of identification. The only discrepancy is connected with erroneous identification of *Cunninghamia lanceolata* pollen with ATR-FT-IR and FT-Raman methods. The reason for that is the already-mentioned difference in the pollen grain size between *Cunninghamia lanceolata* and the average *Cupressaceae* specimen. With the principal goal of the research achieved, the spectral search tested the possibility of separation at the level of genus within pollen spectra belonging to the *Pinaceae* family. The transmission FT-IR method accomplished excellent identification with 100% accurate genus identification, except in the case of the *Abies* genus. Because each spectral library contained only one *Abies* spectrum, and given the previously discussed spectral differences between *Abies koreana* and the other two *Abies* species, it is understandable why the spectral search made errors. However, the spectral search obtained completely reliable identification of all *Picea* and *Pinus* samples. It should be stated, however, that the considered dataset is too unbalanced to obtain a definitely reliable estimation of the true error rate of the classifier. Therefore, a further estimation of classification accuracy on a more balanced, independent dataset is desirable.

The spectral search based on ATR-FT-IR measurements also results in a similar level of exceptional prediction. However, in this case discrepancy is not connected with one single cause, as was the case with *Abies koreana* in the transmission measurements, and identification errors were distributed quite evenly within each genus. Concerning the Raman spectra, the spectral search only confirmed the indications based on the HCA and the PCA of the Raman data that the FT-Raman method can not provide reliable identification of pollen genus.

The reason for such different level of prediction when comparing these three spectroscopic methods lies in the dissimilarity of information acquisition. As previously stated, the principal difference between the pollen grains belonging to the *Pinaceae* group and the *Cupressaceae/Taxaceae/Cephalotaxaceae* group is the relative amount of sporopollenin. Because all three vibrational methods obtain accurate spectral information on the relative quantity of sporopollenin, they all provide correct differentiation between these two pollen groups. However, to obtain identification at the level of genus a spectroscopic method must acquire information on the intracellular material. Although that goal is beyond the reach of

TABLE II. Pollen classification.^a

	<i>Pinaceae/Cupressaceae</i> ^b				<i>Pinus/Abies/Picea</i>			
	True		False		True		False	
	TC	TQ	FQ	FC	TC	TQ	FQ	FC
Transmission FT-IR	100	–	–	–	91	5	2	2
ATR-FT-IR	99	1	–	–	79	18	3	–
FT-Raman	97	2	1	–	56	23	10	11

^a Correctness of classification (T: true, F: false) and confidence of prediction (C: confident, Q: questionable). Numerical values refer to percentage of classified spectra. Results are average values of spectral searches based on ten separate spectral libraries per vibrational method.

^b Classification group name is identical to the corresponding plant family or genus except for the *Cupressaceae* group, which also contained three species belonging to the *Taxaceae* and *Cephalotaxaceae* families.

FT-Raman where the pollen samples of the *Pinaceae* family are concerned, this method could still provide good results for some smaller pollen grains.

These results strongly indicate that pollen spectra contain some common features connected to a specific plant family and genus. By focusing the data analysis on small spectral regions of high diagnostic value one can obtain more accurate analytical information than when dealing with the entire region. These diagnostic spectra could also be valuable in obtaining preliminary biological classification of ambiguous pollen samples, hence enabling rapid classification methods in various environmental studies. Naturally, for that purpose the spectral library should contain optimum spectral variability (and not unbiased random spectra as was the case in this study) by way of spectrum selection and calibration.^{26,27}

The discrimination at the species level was performed for *Pinus nigra* and *Pinus sylvestris* samples only. These samples were collected from three different plants (for each species), and the correct spectral differentiation of these pollen samples, based on RMSE analysis of their transmission IR spectra, showed that accurate discrimination of pollen spectra of species belonging to the same genus can be obtained. It should be noted that this result is identical to the one obtained with the same technique (transmission IR measurement of KBr sample pellets) on the pollen samples of the *Citrus* genus.⁸ Unfortunately, due to the fact that the majority of other studied species were represented with only one sample, more thorough analysis at that taxonomic level was not possible. However, taking into consideration the complexity of real-world samples, the taxonomic resolution at the level of species is somewhat unreasonable and redundant. The sufficient taxonomic resolution for pollen identification in a majority of environmental studies is at the level of genera, and this study clearly showed that spectral similarities between related species enable simple identification of correct family and/or genus.

CONCLUSION

The acquisition of IR and Raman spectra of the same pollen specimens and in similar conditions presented an opportunity of making a thorough analysis on the complete set of vibrational data. Depending on the penetration depth of the probe light (infrared light or laser excitation), vibrational spectra acquire different information on the biochemical composition of pollen grains. FT-Raman and ATR-FT-IR techniques predominantly obtain spectral information on the superficial structures of pollen grains, such as exine (spore-pollenin) and inine (cellulose) grain walls. The transmission

FT-IR method obtains more comprehensive information on pollen grains, including not only the spectral fingerprint of the grain wall, but also information on intracellular biochemicals such as lipids and proteins. As a result, the transmission FT-IR method achieves superior spectral identification of pollen over the ATR-FT-IR or FT-Raman techniques. The theoretical taxonomic resolution of FT-IR methods is at the level of species and therefore is better than that obtained with optical microscopy.

Although biochemical similarities of pollen grains that belong to the same plant genus or family indirectly lead to similar overall features in the corresponding vibrational spectra, the results of HCA, PCA, and spectral search strongly indicate that pollen analysis can be based on the use of diagnostic spectral sub-regions. These discriminating spectral regions result from the performed PCA for the targeted pollen group, either order or family depending on the targeted level of taxonomic resolution. The exploitation of that property in aerobiological monitoring was demonstrated by simple and rapid pollen identification based on relatively small spectral libraries, with the same (or even better) taxonomic resolution than that provided by optical microscopy. There are strong indications that the expected loss in qualitative and quantitative accuracy (when compared with microscopy) can be more than compensated by faster, more economical, and bias-free measurement that requires only basic technical knowledge. Consequently, these results encourage application of FT-IR spectroscopy in environmental studies, such as forest management, monitoring of aeroallergens and invasive species, and phenology.

ACKNOWLEDGMENTS

The author thanks V. Stamenković, D. Mihelj, B. Juretić (Faculty of Science, University of Zagreb), and D. Kremer (Faculty of Pharmacy and Biochemistry, University of Zagreb) for help and advice on collecting botanical samples, and I. Hrga (Institute of Public Health "Dr. Andrija Štampar") and B. Mitić (Faculty of Science, University of Zagreb) for advice regarding aeroallergen pollen monitoring and identification. The author thanks A. Hozić Zimmermann for optical microscope images and G. Baranović, S. Geček, and J. Klanjšček (Ruđer Bošković Institute) for discussions and advice. This work was supported by the Ministry of Education, Sciences and Sports of the Republic of Croatia (project no. 098-0982904-2927).

SUPPLEMENTAL MATERIAL

Supplemental Material mentioned in the text, including Figs. S1 through S6, are available online at <http://www.s-a-s.org>, in the electronic version of the Journal.

1. M. Hesse, H. Halbritter, R. Zetter, M. Weber, R. Buchner, A. Frosch-

- Radivo, and S. Ulrich, *Pollen Terminology* (Springer, Wien, 2009), pp. 11–13.
2. C. Janson, J. Anto, P. Burney, S. Chinn, R. de Marco, J. Heinrich, D. Jarvis, N. Kuenzli, B. Leynaert, C. Luczynska, F. Neukirch, C. Svanes, J. Sunyer, and M. Wyst, *Eur. Respir. J.* **18**, 598 (2001).
3. A. A. Cruz, T. Popov, R. Pawankar, I. Anenisi-Maesano, W. Fokkens, J. Kemp, K. Ohta, D. Price, and J. Bousquet, *Allergy* **62**, 1 (2007).
4. D. V. Wallace, M. S. Dykewicz, D. I. Bernstein, J. Blessing-Moore, L. Cox, D. A. Khan, D. M. Lang, R. A. Nicklas, J. Oppenheimer, J. M. Portnoy, C. C. Randolph, D. Schuller, S. L. Spector, and S. A. Tilles, *J. Allergy Clin. Immunol.* **122**, S1 (2008).
5. J. Peccia, D. K. Milton, T. Reponen, and J. Hill, *Environ. Sci. Technol.* **42**, 4631 (2008).
6. F. Schulte, J. Lingott, U. Panne, and J. Kneipp, *Anal. Chem.* **80**, 9551 (2008).
7. E. Gottardini, S. Rossi, F. Cristofolini, and L. Benedetti, *Aerobiologia* **23**, 211 (2007).
8. C. S. Pappas, P. A. Tarantilis, P. C. Harizanis, and M. G. Polissiou, *Appl. Spectrosc.* **57**, 23 (2003).
9. R. Dell'Anna, P. Lazzeri, M. Frisanco, F. Monti, F. Malvezzi Campeggi, E. Gottardini, and M. Bersani, *Anal. Bioanal. Chem.* **394**, 1443 (2009).
10. N. P. Ivleva, R. Niessner, and U. Panne, *Anal. Bioanal. Chem.* **381**, 261 (2005).
11. A. R. Boyain-Goitia, D. C. S. Beddows, B. C. Griffiths, and H. H. Telle, *Appl. Opt.* **42**, 6119 (2003).
12. R. Manoharan, E. Ghiamati, K. A. Britton, W. H. Nelson, and J. F. Sperry, *Appl. Spectrosc.* **45**, 307 (1991).
13. A. Sengupta, M. L. Laucks, and E. J. Davis, *Appl. Spectrosc.* **59**, 1016 (2005).
14. F. Schulte, J. Mäder, L. W. Kroh, U. Panne, and J. Kneipp, *Anal. Chem.* **81**, 8426 (2009).
15. S. R. Lowry, "Automated spectral searching in infrared, Raman and near-infrared spectroscopy", in *Handbook of Vibrational Spectroscopy*, J. M. Chalmers and P. R. Griffiths, Eds. (Wiley, Chichester, 2002), vol. 3, pp. 1948–1961.
16. D. G. J. L. van Bilsen, F. A. Hoekstra, L. M. Crowe, and J. H. Crowe, *Plant Physiol.* **104**, 1193 (1994).
17. G. Talsky, *Derivative Spectrophotometry* (Wiley-VCH, Weinheim, 1994), p. 146.
18. W. J. Guilford, D. M. Schneider, J. Labovitz, and S. J. Opella, *Plant Physiol.* **86**, 134 (1988).
19. M. Kawase and M. Takahashi, *Grana* **34**, 242 (1995).
20. E. Domínguez, J. A. Mercado, M. A. Quesada, and A. Heredia, *Grana* **37**, 93 (1998).
21. E. Domínguez, J. A. Mercado, and M. A. Quesada, *Heredia, Sex. Plant Reprod.* **12**, 171 (1999).
22. J. De Gelder, K. De Gussem, P. Vandenabeele, and L. Moens, *J. Raman Spectrosc.* **38**, 1133 (2007).
23. H. Fabian and W. Mantele, "Infrared spectroscopy of proteins", in *Handbook of Vibrational Spectroscopy*, J. M. Chalmers and P. R. Griffiths, Eds. (Wiley, Chichester, 2002), vol. 5, p. 3400.
24. K. Brandenburg and U. Seydel, "Vibrational spectroscopy of carbohydrates and glycoconjugates", in *Handbook of Vibrational Spectroscopy*, J. M. Chalmers and P. R. Griffiths, Eds. (Wiley, Chichester, 2002), vol. 5, pp. 3484–3486.
25. K. Maquelin, L. P. Choo-Smith, C. Kirschner, N. A. Ngo-Thi, D. Naumann, and G. J. Puppels, "Vibrational spectroscopic studies of microorganisms", in *Handbook of Vibrational Spectroscopy*, J. M. Chalmers and P. R. Griffiths, Eds. (Wiley, Chichester, 2002), vol. 5, pp. 3311–3314.
26. T. Isaksson and A. H. Aastveit, "Classification methods", in *Handbook of Vibrational Spectroscopy*, J. M. Chalmers and P. R. Griffiths, Eds. (Wiley, Chichester, 2002), vol. 3, pp. 2107–2114.
27. J. W. Einax, H. W. Zwaninger, and S. Geiß, *Chemometrics in Environmental Analysis* (Wiley-VCH, Weinheim, 1997), Chap. 5, pp. 153–189.

Supplementary Material for:

Characterization of pollen by vibrational spectroscopy

Boris Zimmermann

Department of Organic Chemistry and Biochemistry,
Ruđer Bošković Institute, Bijenička 54, 10002 Zagreb, Croatia
bzimmer@irb.hr

<u>Table of Contents</u>	<u>Page</u>
Hierarchical cluster analysis of pollen vibrational spectra	S2
Principal component analysis of pollen vibrational spectra	S3
Optical microscope images and IR spectra of <i>Cunninghamia lanceolata</i> and <i>Cupressus sempervirens</i> pollen	S4

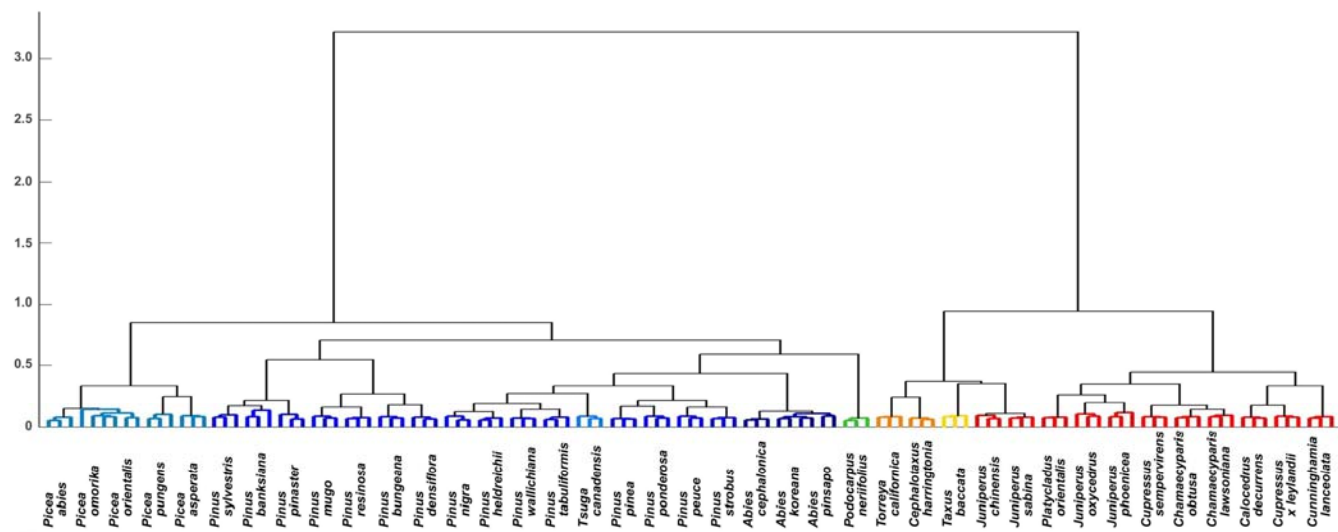


Fig. S1 The hierarchical cluster analysis dendrogram of ATR IR spectra

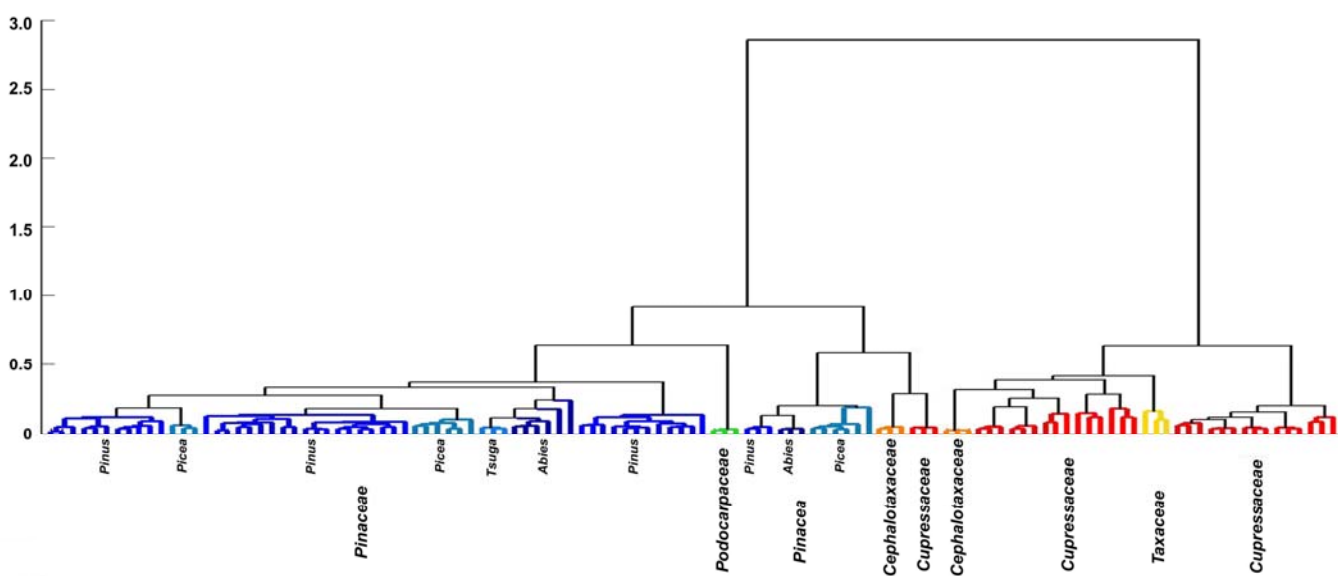


Fig. S2 The hierarchical cluster analysis dendrogram of Raman spectra

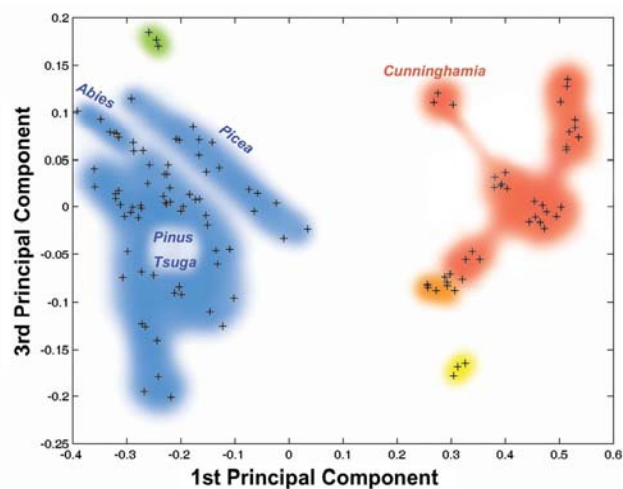
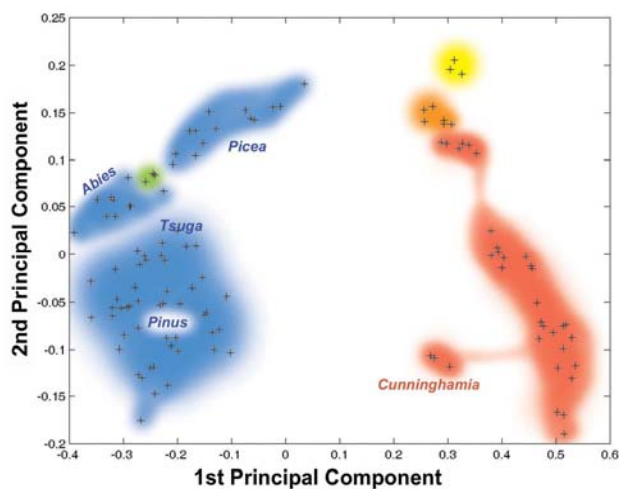


Fig. S3 The principal component analysis of ATR IR spectra with tentative depiction of plant families: *Pinaceae* (blue), *Cupressaceae* (red), *Cephalotaxaceae* (orange), *Taxaceae* (yellow) and *Podocarpaceae* (green)

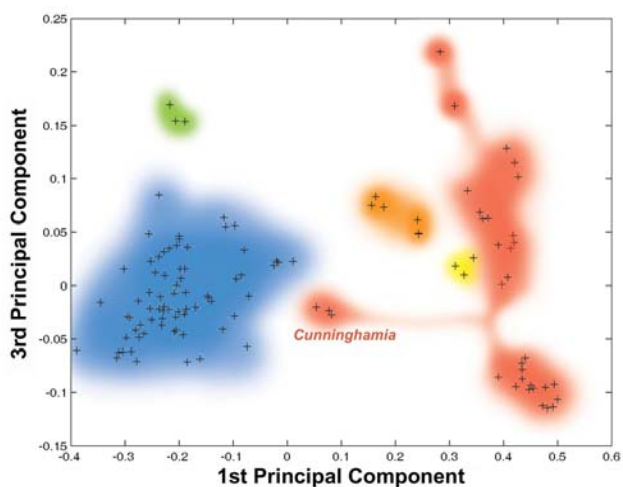
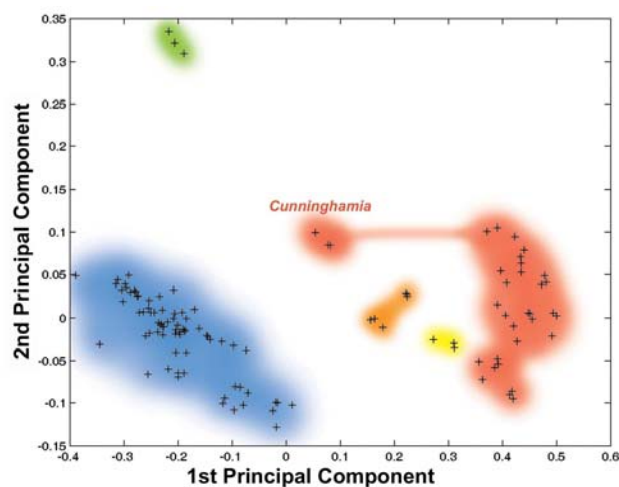


Fig. S4 The principal component analysis of Raman spectra with tentative depiction of plant families: *Pinaceae* (blue), *Cupressaceae* (red), *Cephalotaxaceae* (orange), *Taxaceae* (yellow) and *Podocarpaceae* (green)

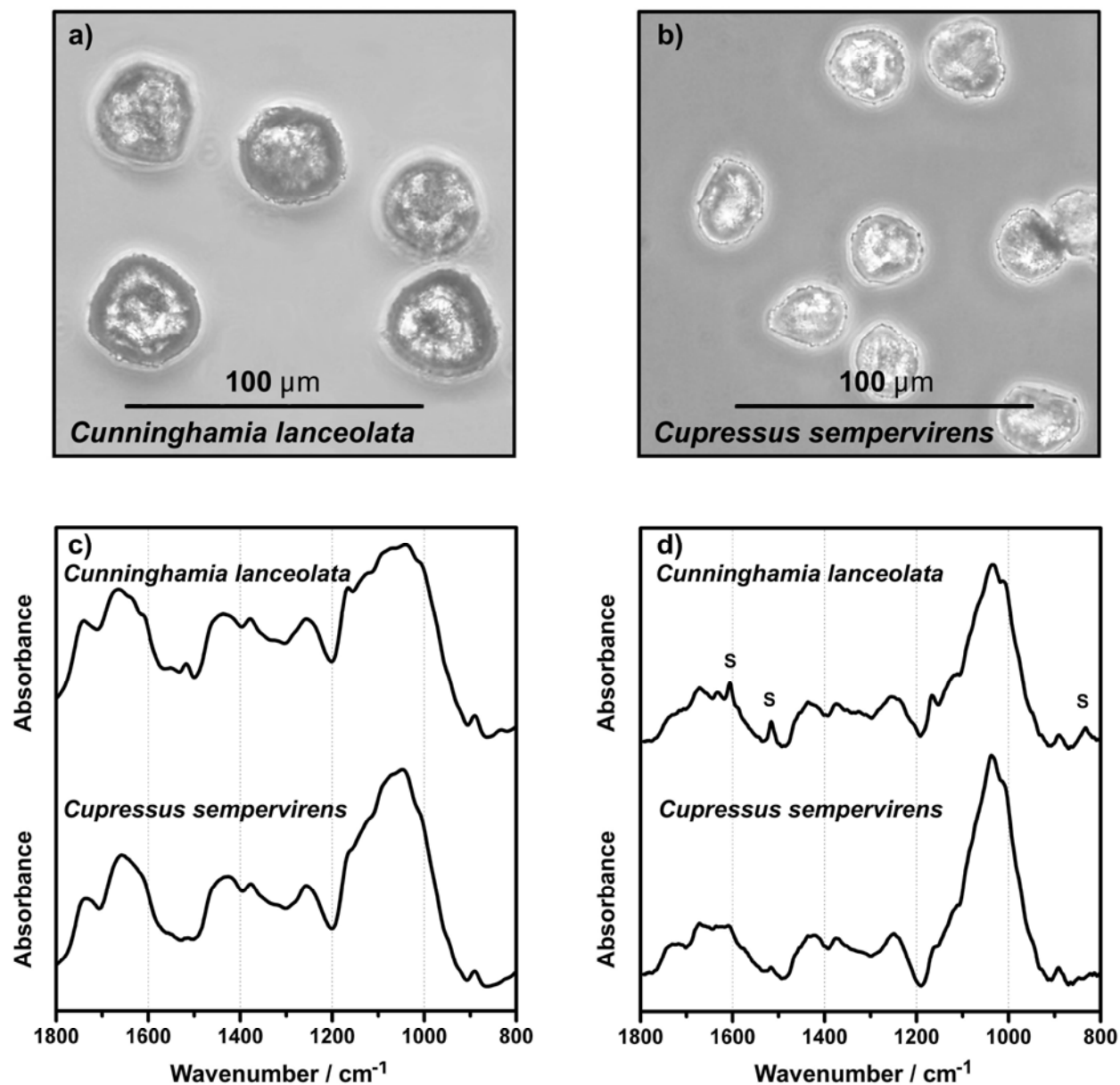
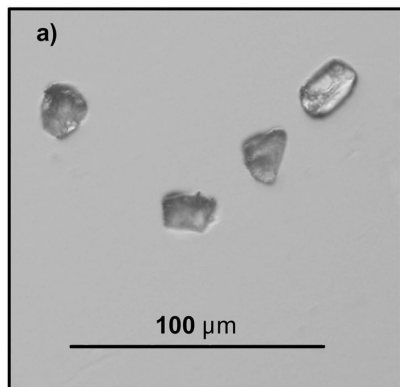
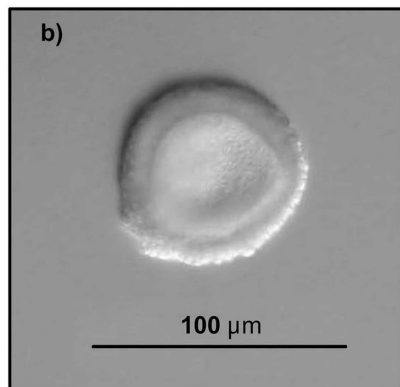


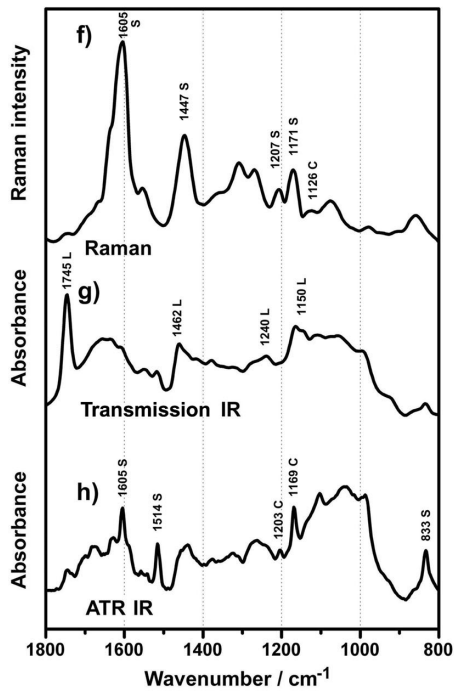
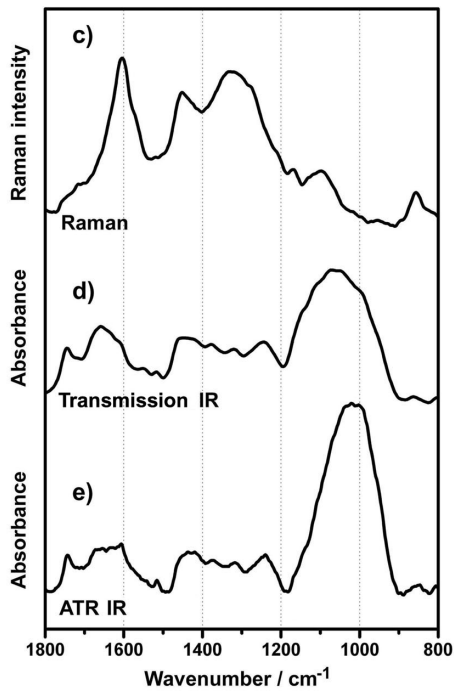
Fig. S5 Optical microscope images of a) *Cunninghamhamia lanceolata* and b) *Cupressus sempervirens* pollen grains; c) transmittance IR spectra of pollen, and d) ATR IR spectra of pollen (S = sporopollenin vibrational bands)

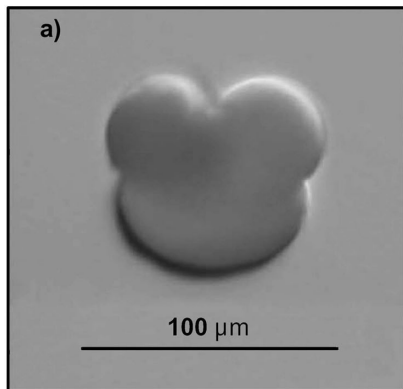


Taxus baccata

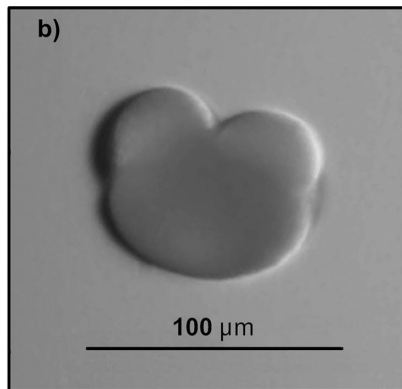


Tsuga canadensis





Abies koreana



Abies pinsapo

