frontiers in Forests and Global Change

TYPE Original Research
PUBLISHED 16 March 2026
DOI 10.3389/ffgc.2026.1755489

OPEN ACCESS

EDITED BY
Milica Zlatkovic,
University of Novi Sad, Serbia

REVIEWED BY
Giorgio Maresi,
Fondazione Edmund Mach, Italy
Jan Stejskal,
Czech University of Life Sciences Prague, Czechia

*CORRESPONDENCE
Patrick Sherwood
patrick.sherwood@slu.se

RECEIVED 27 November 2025
REVISED 01 March 2026
ACCEPTED 04 March 2026
PUBLISHED 16 March 2026

CITATION
Sherwood P (2026) Infrared spectroscopy enables rapid identification of Scots pine resistant to *Diplodia sapinea*.
Front. For. Glob. Change 9:1755489.
doi: 10.3389/ffgc.2026.1755489

Infrared spectroscopy enables rapid identification of Scots pine resistant to *Diplodia sapinea*

Patrick Sherwood*

Southern Swedish Forest Research Centre, Swedish University of Agricultural Sciences, Alnarp, Sweden

Many forest pathogens have become increasingly damaging as climate-driven stress intensifies disease outbreaks. A prominent example is *Diplodia sapinea,* which is causing more frequent and destructive shoot blight and dieback, particularly across northern Europe. Developing resistant planting stock is a priority, but screening currently depends on destructive inoculation assays and multi-year field trials, making the process slow and difficult to scale. This study evaluated whether a handheld Fourier-transform infrared (FT-IR) spectrometer that provides rapid and non-destructive sampling can identify constitutive chemical signatures associated with relative resistance prior to infection. FT-IR spectra were collected from needles, shoots, and stem phloem from Swedish pine families across two experimental years, followed by artificial inoculations to phenotype lesion development. A combined sparse partial least squares discriminant analysis and support vector machine workflow achieved 65–81% cross-validated accuracy in distinguishing resistant from susceptible trees, with shoot spectra consistently producing the strongest models. Principal component analysis indicated clear chemical differences between years, yet resistance-associated patterns were stable across tissues. The most frequently selected wavenumbers grouped into four biochemical domains: cell-wall polysaccharides, cellulose/hemicellulose bonding, phenolics and proteins, and aliphatic lipids/cutin, corresponding to known conifer defense pathways. These findings demonstrate that portable FT-IR spectroscopy can capture biologically meaningful variation in constitutive defense chemistry and offers a rapid, scalable approach for resistance phenotyping in Scots pine breeding programs.

KEYWORDS
biochemical fingerprint, forest pathology, FT-IR, phenotyping, tree health

Introduction

Forests worldwide are experiencing unprecedented disease pressure driven by a combination of biological invasions, intensifying climate stress, and shifts in pathogen behavior (Guégan et al., 2023; Ramsfield et al., 2016). Newly introduced pests continue to expand into previously unaffected regions, while many long-established pathogens are becoming more damaging as warming temperatures, altered precipitation patterns, and drought stress reduce host resilience. Prime examples include outbreaks of damaging canker and dieback diseases caused by latent fungi in the Family Botryosphaeriaceae (Desprez-Loustau et al., 2006; Phillips et al., 2013).

Diplodia sapinea is a prominent member of the Botryosphaeriaceae, causing shoot blight and canker pathogen on pines, where it can infect all tissues and life stages of susceptible species (Wingfield et al., 2024). It can also infect other conifers and even some hardwoods

(Zlatković et al., 2017). Symptoms include killed shoots and branches, blue staining of the wood that lowers timber value, and when severe, even tree mortality. Long considered an opportunistic fungus, *D. sapinea* is capable of persisting as a latent endophyte in asymptomatic pine shoots (Flowers et al., 2001; Flowers et al., 2006), becoming pathogenic when trees are weakened by drought, heat stress, or other environmental challenges. *Diplodia sapinea* is becoming increasingly important in northern Europe (Brodde et al., 2019), where reports of tip blight, shoot dieback, and stem cankers caused by *D. sapinea* have increased steadily over the decade (Müller et al., 2019; Oliva et al., 2013). Given the projected increase in climatic extremes, identifying pine genotypes with higher natural tolerance or resistance to *D. sapinea* should be a priority for forest health and breeding programs.

Exploiting natural host resistance remains the most sustainable long-term strategy for mitigating disease caused by both endemic and invasive pathogens (Showalter et al., 2018). Durable quantitative resistance, typically underpinned by complex chemical and physiological traits, is especially valuable because it tends to be less vulnerable to pathogen adaptation than single major resistance genes. In pine trees, resistance to *D. sapinea* has been linked to coordinated shifts in primary and secondary metabolism, including alterations in phenolic pathways and terpene-based defenses (Ghosh et al., 2024; Wallis et al., 2008). However, translating these defense traits into operational breeding has proven challenging. Traditional resistance screening relies on destructive inoculation assays followed by multi-year field trials, making the process slow, labor-intensive, and difficult to scale (Sniezko and Koch, 2017; Sniezko and Liu, 2021). The inability to rapidly and accurately identify resistant individuals remains one of the principal bottlenecks in developing disease-resilient planting stock.

There is therefore a growing need for rapid, non-destructive phenotyping tools capable of detecting biochemical signatures associated with host resistance. Vibrational spectroscopy, including Fourier-transform infrared (FT-IR), near-infrared (NIR), and Raman spectroscopy, has emerged as a promising solution for capturing broad chemical fingerprints of plant tissues. Because induced and constitutive defenses are largely chemical in nature, infrared spectra often capture subtle but biologically meaningful differences in cell wall composition, phenolics, lipids, and proteins that contribute to resistance expression. When coupled with chemometric or machine-learning approaches, infrared spectroscopy offers the possibility of assessing complex biochemical traits relevant to resistance.

Infrared-based phenotyping has already been successfully applied to several forest pathosystems, including resistance to ash dieback in *Fraxinus excelsior* (Bonello et al., 2025; Villari et al., 2018), Dutch elm disease in *Ulmus minor* (Martin et al., 2008), *D. sapinea* in *Pinus nigra* (Conrad et al., 2020), and fusiform rust in *Pinus taeda* (Lim-Hing et al., 2025). Collectively, these studies demonstrate that infrared spectroscopy can reveal resistance-associated chemistry long before, offering a promising path toward high-throughput resistance screening in tree breeding programs.

Most prior work has relied on benchtop FT-IR systems and extensive sample preparation, which limit field deployment and operational scalability. Handheld spectrometers are becoming more widely available, and portable NIR devices have been tested in forestry applications; however, handheld FT-IR remains largely unexplored in forest pathology. NIR primarily captures broad overtones and combination bands of C–H, O–H, and N–H groups, allowing deeper penetration into tissues, but providing comparatively coarse chemical resolution (Cozzolino, 2014). In contrast, FT-IR detects fundamental vibrational absorptions and therefore produces more information-rich spectra with stronger and more distinct bands for many defense related molecules, including cell-wall polysaccharides, lignin-associated aromatics, lipids, and proteins (Türker-Kaya and Huck, 2017). Thus, portable FT-IR spectroscopy may provide a powerful, yet underutilized, tool for rapid tree resistance phenotyping.

In this study, we used a handheld FT-IR spectrometer to characterize constitutive biochemical profiles of multiple Scots pine families prior to artificial inoculation of elongating shoots with *D. sapinea* in 2023 and 2024. Spectra were collected from intact needles, shoots, and stem phloem to evaluate tissue-specific and family-level chemical variation. Trees were then phenotyped for relative resistance based on lesion development following inoculation, and chemometric analyses were applied to determine whether FT-IR spectra could differentiate the most resistant and most susceptible individuals. This approach allowed us to assess the consistency of phenotypic ranking across tissues and years, and to evaluate the potential of portable mid-infrared spectroscopy as a tool for rapid resistance screening in Scots pine breeding programs.

Materials and methods

Inoculations and phenotyping

Four-year-old potted *Pinus sylvestris* seedlings representing 34 half-sib families from the Swedish Scots pine breeding program (Skogforsk) were used in this study. Families originated from one of four regions: Ekebo, Brunsberg, Albjershus, and Drögsnäs, although only a single family was available from Drögsnäs. Only visually healthy trees showing uniform vigor and comparable shoot development were included in the experiment. A total of 192 trees were inoculated in 2023 and 185 in 2024. A full list of families and their regional origins is provided in Supplementary Table S1. Trees were arranged in a completely randomized design in and conditioned to the greenhouse for 20 and 27 days prior to experimentation in 2023 and 2024, respectively. Trees received regular watering through conditioning and experimentation. Environmental fluctuations in the greenhouse were not logged but excessive temperatures and humidity were controlled by automated venting.

Elongating shoots were inoculated following Sherwood and Bonello (2016). Briefly, a shallow wound was made at a needle fascicle insertion point at the midpoint of each shoot. An agar plug colonized by *D. sapinea* (isolate CBS 117911; Westerdijk Fungal Biodiversity Institute, Utrecht, Netherlands), maintained on potato dextrose agar (BD Difco, USA), was placed over the wound and sealed with Parafilm. Each tree received three inoculations: the dominant apical leader and two side shoots from the uppermost branch whorl. Inoculations were performed from 26–30 June 2023 and 4–8 July 2024.

Lesion lengths were measured 11 days post-inoculation in 2023 and 14 days post-inoculation in 2024 by carefully removing the outer bark to expose the necrotic area. Lesion lengths were recorded separately for the leader shoot ("Leader") and the two side shoots ("Shoot"), and also averaged ("All") to represent whole-tree response. Differences in lesion length among tissues were evaluated using Kruskal Wallis and Dunn's test. The association between leader and shoot lesion lengths was assessed using Spearman's rank correlation (ρ), supported by a scatter plot to visualize the relationship.

To assign relative resistance categories, lesion lengths within each tissue × year combination were divided into quartiles: trees in the lowest quartile were classified as resistant, those in the upper quartile as susceptible, and those between these thresholds as intermediate. Differences in lesion length between resistant and susceptible trees within each tissue and year were tested using a one-sided Wilcoxon rank-sum test. To account for non-independence among offspring from the same half-sib family, lesion length data were analyzed using linear mixed-effects models. Mean lesion length per tree ("All") was log-transformed [$\log(x + 1)$] to reduce right skewness and improve model residuals. Models were fitted with year and seed origin (region) included as fixed effects, and family included as a random intercept to account for clustering of trees within families. Models were fitted using restricted maximum likelihood, and inference on fixed effects was based on Satterthwaite-approximated degrees of freedom. Analyses were conducted using the packages lme4 and lmerTest.

FT-IR data collection and processing

Immediately prior to inoculation, samples of main-stem phloem, a separate elongating shoot from the uppermost whorl, and four current-year needles were collected for spectral analysis. A 6-mm plug of bark and phloem was removed 5 cm above the soil line using a sterilized cork borer; the xylem-facing surface of this plug was scanned at three positions. The shoot sample was excised, stripped of needles, and scanned three times along its length. Needles were plucked at random from the upper crown and scanned directly without regard to adaxial or abaxial orientation.

Trees were scanned in sequential order corresponding to their randomized greenhouse placement, such that scan order was independent of family and region. All samples were scanned using an Agilent 4300 handheld FT-IR spectrometer (Agilent Technologies, USA) equipped with a diamond attenuated total reflectance (D-ATR) interface mounted on an Agilent 4300 sampling stand. The sampling stage applies a controlled clamping force to ensure consistent contact between the sample and the ATR crystal (see Supplementary Figure S1). Spectra were collected over the full mid-infrared range (4000–650 cm^{-1}) in MicroLab software using 32 co-added scans per measurement, a resolution of 4 cm^{-1}, Happ–Genzel apodization, no zero-filling, and Mertz phase correction. A new background spectrum was acquired approximately every 6 min, and the D-ATR crystal was cleaned with absolute ethanol between samples.

The chemometric and machine-learning analyses followed a sequential workflow consisting of (i) spectral preprocessing, (ii) exploratory principal components analysis (PCA), (iii) feature selection using sparse partial least squares discriminant analysis (sPLS-DA), and (iv) supervised classification on the reduced dataset (i.e., sPLS-DA selected wavenumbers) using support vector machine (SVM).

For spectral preprocessing, raw spectra were visually inspected to confirm adequate signal intensity and baseline stability. Spectra were collated and trimmed to include only wavenumbers 3,150–2,700 cm^{-1} and 1899–677 cm^{-1} to remove noisy regions, uninformative areas, and known CO_2 interference peaks, leaving 899 wavenumbers in total. Outlier spectra showing abnormally low absorbance peaks in the fingerprint region of 700–1,600 cm^{-1}, indicative of poor crystal contact, were removed. In total, 20 spectra were trimmed (two needle spectra and 18 shoot spectra) representing 0.5% of the total spectra collected. For each tree and tissue type, the three replicate scans were averaged to generate a single representative spectrum, which was then used for all preprocessing, PCA, and modeling steps. Averaged spectra were second-order derivatized using the prep.savgol function (width = 11, porder = 3, dorder = 2), and these derivative spectra were used directly in subsequent chemometric analyses. Grand mean spectra for each tissue type, calculated as the mean of the tree-level averaged spectra before and after derivatization, are shown in Supplementary Figure S2.

PCA was used to visualize major sources of chemical variation in the second-derivative spectra. Analyses were conducted using the prcomp function in R. Two complementary PCA plots were generated. To identify informative wavenumbers and evaluate predictive ability, a combined sPLS-DA + SVM model workflow was implemented. sPLS-DA feature selection was performed using the mixOmics R package for each combination of year (2023, 2024, both years combined) predictive tissue (needle, shoot, phloem) and phenotyping tissue (Leader, Shoots, All). Intermediate phenotypes were excluded, and models were trained only on resistant and susceptible trees. sPLS-DA was run with three latent components. The number of variables to retain (keepX) was optimized via M-fold cross-validation (up to five folds, depending on class sizes) using balanced error rate as the selection criterion. The most frequently occurring wavenumbers from models with the highest accuracies were used for SVM classification models with a linear kernel and cost = 0.01 were trained. For each tissue × year × phenotype combination data were split into 70% training and 30% testing using stratified sampling. Classification performance was evaluated using the training accuracy, test accuracy, test area under the receiver operator characteristic curve (ROC AUC), a 5-fold cross-validated (CV) accuracy, and CV ROC AUC, with PCA recalculated within each fold to avoid information leakage. All analyses were performed in RStudio 2025.09.1 Build 401 (Posit, USA). A large language model (ChatGPT v5.1, OpenAI) was used to assist with drafting and troubleshooting R code; all code, outputs, and interpretations were manually checked and validated.

Results

Lesion lengths were largely consistent between the tissues and years (Figure 1), with only the leader shoots in 2024 having significantly smaller lesions compared to the lesions of all other tissues from both years. See Supplementary Table S2 for lesion length tests of normality and homogeneity of variance. Although the lesion lengths were similar, there was only a moderate ($\rho = 0.52$), yet highly significant ($p < 0.001$), correlation between leader and shoot lesion lengths (Supplementary Figure S3), showing that trees with longer leader lesions tended to also develop longer shoot lesions.

When trees were split into quartiles for assigning relative resistance, phenotypes were mostly similar when comparing the average of all lesion lengths to the side shoot lesions only (Figure 2). When comparing leader shoots to all shoots, there are considerably more fluctuations in the spread. In both 2023 and 2024, one tree that was classified as susceptible based on the average of all lesion lengths is classified as resistant when using the leader lesion length. No matter the year or tissues used for phenotype assignment, trees classified as resistant always had significantly smaller lesions compared to the susceptible trees (Figure 2, see Supplementary Table S3 for Wilcox rank-sum tests), demonstrating that trees selected for modeling have intrinsic differences in their relative symptoms.

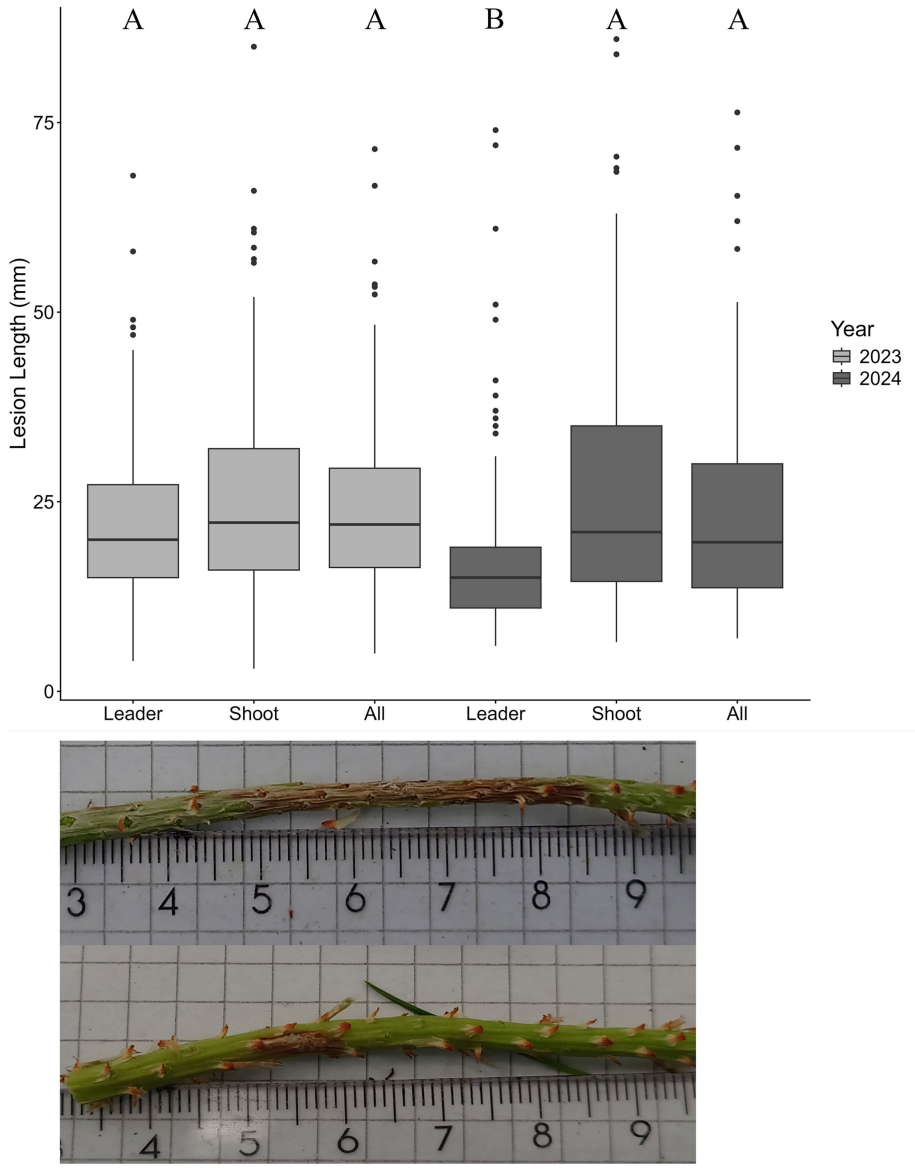

FIGURE 1
Top: Box plots of lesions lengths of *Pinus sylvestris* seedlings inoculated with *Diplodia sapinea*. Inoculations were done on the dominant apical shoot (Leader) and two random side shoots of the top branch whorl (Shoot). "Shoot" lesion lengths are the averages of the two lesions from the side shoots for the given tree, while "All" lesion lengths are the averages of the lesion lengths from the two side shoots and the leader shoot on the tree. Significance letters denote statistically different lesion lengths as determined by Dunn's test at the $\alpha < 0.05$ level. In 2023, $n = 192$ trees and lesions were measured 11 days after inoculation; in 2024, $n = 185$ trees and lesions were measured 14 days after inoculation. Bottom: pictures of example lesions from trees in 2023.

Linear mixed-effects models accounting for non-independence among half-sib offspring indicated significant phenotypic variation in lesion length among families (see Supplementary Table S4 for model results and Supplementary Figure S4 for lesion length box plots by family). Family-level differences accounted for approximately 18–19% of total phenotypic variation in log-transformed lesion length. In contrast, lesion length did not differ significantly between years (F_1, 365 = 1.14, $p = 0.29$) or among seed origins (regions) after accounting for family structure (F_3, 34.7 = 1.22, $p = 0.32$).

PCA of the second-derivative FT-IR spectra (Figure 3) revealed strong separation among tissue types and between experimental years along the first two principal components, with year-related separation most evident in needle and shoot tissues. In contrast, resistant, intermediate, and susceptible phenotypes did not form clearly distinct clusters, indicating that resistance-associated chemical differences were subtle relative to the dominant variation associated with tissue type and year.

Using sPLS-DA (model summary in Supplementary Table S5) to reduce the full FT-IR spectra from 899 wavenumbers to a subset of 189 informative wavenumbers still enabled the SVM models to achieve relatively high predictive performance across years and tissues (Table 1). Despite the substantial dimensionality reduction, many models still achieved testing accuracies between 0.60–0.89, demonstrating that a reduced set of wavenumbers retains substantial discriminative power for predicting resistance phenotypes.

The best performing models based on cross validation (CV) accuracy and area under curve (CV AUC) were for the phenotyped tissue all and Shoot using shoot spectra in 2023, with CV

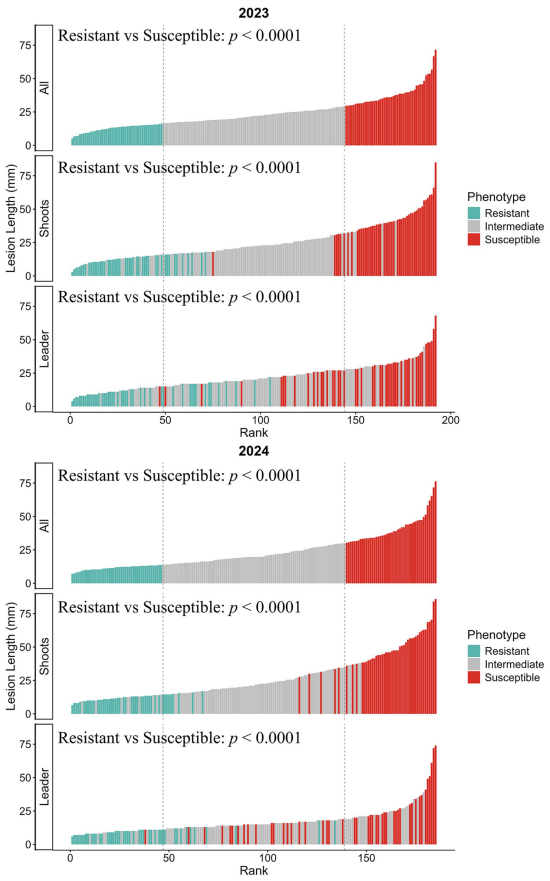

FIGURE 2
Bar plots of lesion lengths from each tissue type used to generate the assigned phenotypes for each experimental year. Inoculations were done on the dominant apical shoot (Leader) and two random side shoots of the top branch whorl (Shoot). "All" lesion lengths are the averages of the lesion lengths from the two side shoots and the leader shoot on the tree. Dashed vertical lines demarcate the first and fourth quartiles for a given year. Trees with lesion lengths in the lowest quartile (n = 48 in 2023; n = 46 in 2024) were considered resistant. Trees with lesion lengths in the middle quartiles were considered intermediate. Trees with lesion lengths in the highest quartile were considered susceptible. For every tissue and year, trees selected as resistant had significantly smaller lesions compared to the susceptible trees, as determined by Wilcox rank-sum test. Coloring of the bars in all plots corresponds to that tree's assigned phenotype using the "All" lesion lengths for a given year to indicate how the assigned phenotype varied by phenotyping tissue.

accuracies of 80 and 81%, respectively. The best performing model in 2024 was for phenotyped tissue all using shoot spectra, but it only achieved 69% CV accuracy. When combing the years, phenotyped tissue all was still the best performing model when using either shoots or phloem tissue (both 71% accurate). CV accuracy was emphasized over single train–test accuracy, because CV provides a more reliable estimate of model generalizability and reduces the influence of the particular 70:30 split, which can be substantial in datasets of this size. Using CV accuracy and CV AUC, shoot tissue was nearly always the best predictive tissue regardless of year or phenotyping tissue. Phloem and needle spectra sometimes achieved accuracies similar to those obtained using shoot tissues but often performed worse. These results suggest that constitutive chemistry captured in shoot tissues carries a strong and stable signal of relative resistance to *D. sapinea*, even under reduced feature sets and across multiple years.

Models built from the 2023 dataset consistently outperformed those from 2024, with higher test and CV accuracies and AUC values, while models trained on the combined 2023–2024 dataset showed intermediate performance. Patterns across phenotyped tissues were generally consistent, with All and Shoot models tending to perform better than those based on Leader phenotyping alone. Altogether, FT-IR and chemometrics can predict the relative resistance of shoots to *D. sapinea* with good accuracy when using shoot spectra, though models for leader shoot specifically are less accurate.

The most frequently selected wavenumbers by sPLS-DA could generally be divided into four biochemical groupings (Table 2). The strongest cluster occurred within 1,000–1,050 cm^{-1}, a region dominated by cell-wall polysaccharide vibrations, including O–H and C–O modes characteristic of cellulose, hemicelluloses, and matrix sugars. A second, narrower cluster around 1,140–1,160 cm^{-1} highlighted additional cellulose- and hemicellulose-associated C–O–C and C–O vibrations, indicating that variation in carbohydrate structure contributed substantially to resistance differentiation. The third major grouping, spanning 1,570–1,680 cm^{-1}, corresponded to the protein amide II/I envelope and aromatic lignin/phenolic C=C vibrations, suggesting that both structural phenolics and stress-associated protein chemistry were important discriminators of resistance class. Finally, a high-wavenumber cluster between 2,950–3,140 cm^{-1} represented aliphatic CH_2/CH_3 stretching associated with lipids, waxes, and cutin/suberin, indicating that variation in surface or membrane-related aliphatics also contributed to model separation.

Discussion

This study demonstrates that FT-IR spectroscopy combined with chemometric modeling can predict relative resistance of Scots pine to *D. sapinea* with promising accuracy (65–81% CV), using models trained exclusively on constitutive chemistry. Our results are consistent and even slightly better compared to other IR-based studies for predicting resistance levels in conifers. For example Conrad et al. (2020) achieved 64–67% CV accuracies for classifying *D. sapinea* resistance in *Pinus nigra*, and Lim-Hing et al. (2025) achieved 68 and 65% testing accuracies using portable NIR and benchtop FT-IR, respectively. Bonello et al. (2025) had very low balanced error rates (~7–18%) using sPLS-DA for classifying resistance in *Fraxinus excelsior* to the ash dieback (ADB) pathogen using a portable NIR spectrometer, while both Conrad et al. (2014) and Villari et al. (2018) achieved nearly 100% accuracies for resistance phenotyping in the ash-ADB and *Quercus agrifolia–Phytophthora ramorum* pathosystems using soft independent modeling of class analogy. None of the aforementioned studies used a portable FT-IR spectrometer scanning intact tissue and instead used phenolic metabolite extracts dried on the crystal interface. The success of our portable intact tissue is an important result for practical resistance screening: breeders require rapid, non-destructive, and pre-inoculation methods, and constitutive traits offer the most feasible target for large-scale operational phenotyping. Although induced biochemical responses are known to play a major role in pine defense (Eyles et al., 2010; Wallis et al., 2008), our results show that substantial predictive signal resides in the constitutive chemical makeup of uninfected tissues, indicating that baseline chemical differences alone can capture meaningful variation in susceptibility.

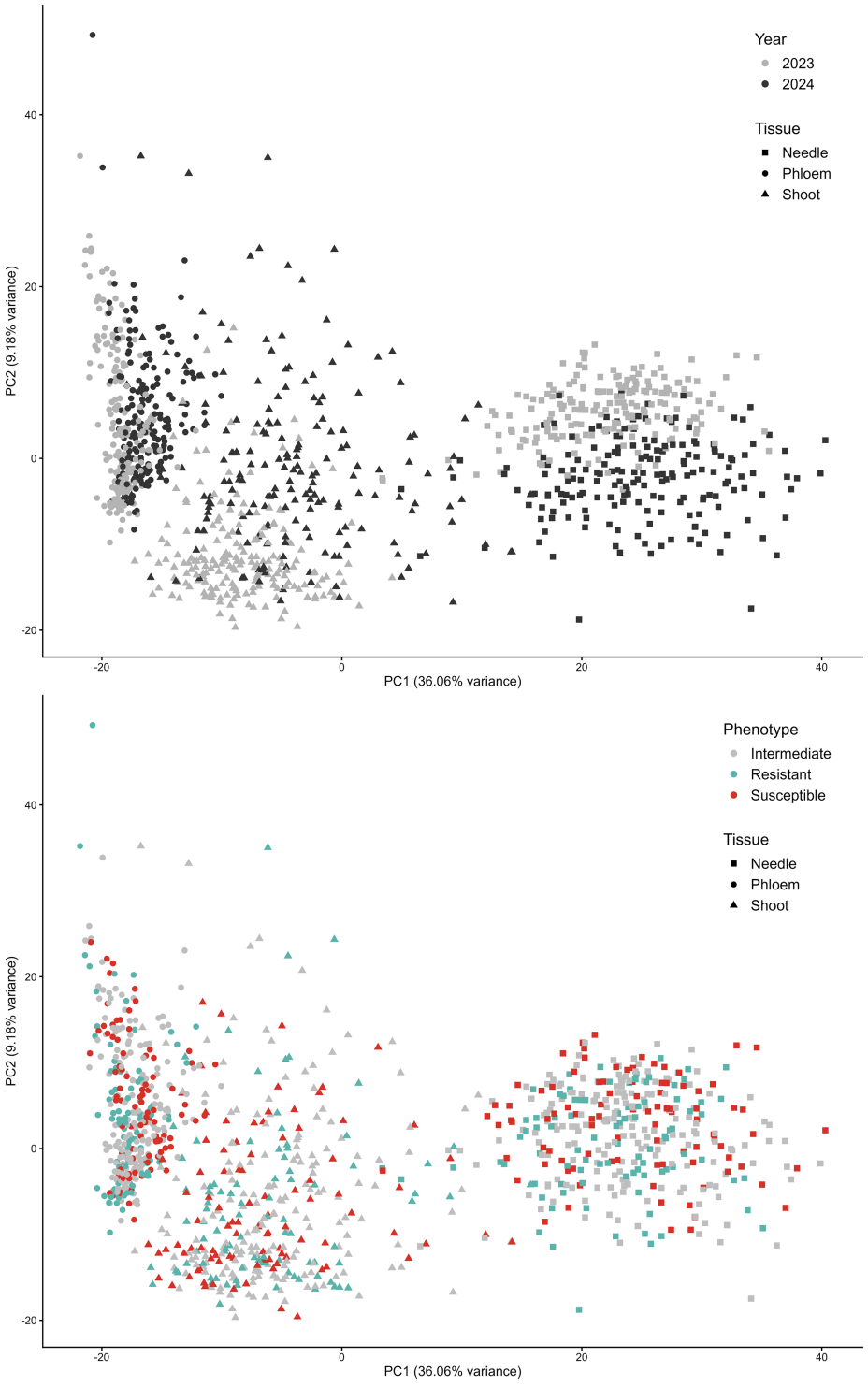

FIGURE 3
PCA of second-derivative FT-IR spectra illustrating strong separation among tissue types (needle, shoot, phloem) and experimental year using the first two principal components (top), but no clear separation of phenotypes (bottom).

A consistent outcome across analyses was that shoot tissues provided the strongest and most stable predictive performance. This aligns well with biological expectations, as the apical and lateral shoots were the tissues in which lesion development, and therefore resistance, was quantified. Since the predictive chemistry and the phenotyped trait originate from the same organ, shoots likely reflect the most relevant physiological context for interaction with *D. sapinea*. Previous IR-based studies in other pathosystems have found phloem to be the strongest predictor (Lim-Hing et al., 2025; Villari et al., 2018), but our results show that in the Scots pine–*D. sapinea* system, phloem was generally slightly less accurate than shoots, particularly in 2023. This study is also the first time shoot tissue has been examined for predictive ability. Needles consistently yielded the weakest models, suggesting that although needle chemistry is easy to sample with minimal

TABLE 1 Summary of SVM model performance using sPLS-DA–reduced FT-IR spectral features to predict relative resistance phenotypes (resistant vs. susceptible) in Scots pine following artificial inoculation with *Diplodia sapinea*.

Year	Phenotyped tissue[a]	Predictive tissue[b]	Train accuracy	Test accuracy	Test AUC	CV accuracy	CV AUC
2023	All	Needle	0.93	0.75	0.80	0.68	0.73
		Phloem	0.97	0.71	0.88	0.77	0.87
		Shoot	0.93	0.61	0.76	**0.80**	0.86
	Leader	Needle	0.93	0.75	0.82	0.69	0.73
		Phloem	0.94	0.57	0.59	0.61	0.64
		Shoot	0.91	0.89	0.92	**0.73**	0.80
	Shoots	Needle	0.93	0.54	0.51	0.75	0.81
		Phloem	0.96	0.71	0.85	0.74	0.82
		Shoot	1.00	0.71	0.83	**0.81**	0.86
2024	All	Needle	0.86	0.73	0.73	0.63	0.62
		Phloem	0.88	0.46	0.51	0.62	0.69
		Shoot	0.89	0.58	0.59	**0.69**	0.70
	Leader	Needle	0.91	0.46	0.53	0.56	0.64
		Phloem	0.86	0.81	0.73	0.48	0.56
		Shoot	0.91	0.58	0.67	**0.66**	0.73
	Shoots	Needle	0.89	0.62	0.69	**0.65**	0.73
		Phloem	0.91	0.62	0.75	0.61	0.69
		Shoot	0.88	0.73	0.83	0.60	0.68
2023 and 2024	All	Needle	0.85	0.66	0.69	0.57	0.60
		Phloem	0.90	0.57	0.61	**0.71**	0.71
		Shoot	0.90	0.60	0.73	**0.71**	0.78
	Leader	Needle	0.86	0.64	0.66	0.55	0.59
		Phloem	0.85	0.59	0.57	0.64	0.68
		Shoot	0.86	0.68	0.74	**0.67**	0.75
	Shoots	Needle	0.82	0.64	0.69	0.58	0.66
		Phloem	0.88	0.66	0.68	**0.68**	0.74
		Shoot	0.89	0.69	0.76	**0.68**	0.75

[a]Phenotyped tissue refers to the tissue from which lesion lengths were measured and used to assign relative resistance phenotypes ("resistant" or "susceptible"). All = average of the leader lesion length and the two side-shoot lesion lengths.
[b]Predictive tissue refers to the tissue scanned with the FT-IR spectrometer, whose spectra were used as predictors in the chemometric models.
The highest CV accuracies within each year × phenotyping tissue × predictive tissue group are shown in bold.

impact on the tree, it does not sufficiently capture metabolite profiles that govern shoot-level resistance in this pathosystem. From an applied perspective, shoots therefore represent the best compromise between predictive power and sampling feasibility, particularly for young greenhouse-grown trees where removing phloem tissue may be too destructive.

Models based on Leader phenotypes were consistently less accurate than those based on Shoots phenotype or the combined all phenotype. A key methodological factor likely contributing to this pattern is that FT-IR spectra were collected exclusively from non-leader tissues, as the leader shoot could not be both scanned and inoculated. Consequently, leader resistance was predicted using chemical information from lateral shoots or other tissues, creating an inherent tissue mismatch that is expected to reduce predictive accuracy relative to models in which the predictive and phenotyped tissues are the same. In addition, leader and lateral shoots are known to differ developmentally in pines. In pines, much of the leader shoot consists of growth units preformed in the previous year, whereas lateral shoots develop during the current growing season under distinct developmental and hormonal regimes (Hover et al., 2017). These differences translate into contrasting carbon allocation patterns and sensitivities to environmental conditions, with leader growth responding differently to nutrient and water availability than other shoot tissues (Axelsson and Axelsson, 1986; Dewar et al., 1994). Such differences can influence tissue anatomy and chemical composition and may further limit the transferability of chemical information between leader and side shoots. Consistent with this interpretation, leader lesions were slightly smaller than those on lateral shoots (significantly so in 2024) and lesion lengths between leaders and shoots were only moderately correlated. Together, these observations suggest that leader and lateral shoots are not strictly interchangeable with respect to their chemical–pathogen interactions. As a result, resistance phenotypes derived from leader lesions are more difficult to model when chemical predictors are obtained from non-leader tissues. From a practical perspective, these results emphasize the importance of aligning FT-IR sampling with the

TABLE 2 Biochemical groupings of the most important wavenumbers from the sPLS-DA model.

Group	Approx. range (cm^{-1})	Tentative assignment[a]
1	1,000–1,050	O–H and C–OH stretching and C–O ring modes of cell-wall polysaccharides (cellulose, hemicellulose, arabinan, galactan).
2	1,140–1,160	C–O–C stretch: cellulose; symmetric bonding of aliphatic CH_2, OH, or C–O stretch of various groups: cell wall polysaccharide.
3	1,570–1,680	Amide II and Amide I envelope (protein); aromatic C=C vibrations of lignin; phenolic C=C.
4	2,950–3,140	Symmetric and asymmetric CH_2/CH_3 stretching of aliphatic chains (lipids, waxes, cutin/suberin).

[a]Assignments inferred from the following references: Dorado et al. (2001), Kačuráková et al. (2000), Pandey and Pitman (2003), Schulz and Baranska (2007), Sene et al. (1994), and Türker-Kaya and Huck (2017).

specific tissue and biological process underlying the resistance phenotype when developing predictive chemotyping tools.

A strong year effect was evident in both the PCA of spectra and in model performance: models trained on the 2023 dataset were consistently more accurate than those from 2024, while models trained on the combined dataset performed intermediately. Although the inoculation dates differed only slightly, pine shoot physiology can change as the season progresses, including shifts in lignification, carbohydrate dynamics, and resin-related chemistry (Mutke et al., 2003). However, the relatively small difference in calendar timing makes it more likely that broader interannual factors like differences in tree vigor, greenhouse microclimate, or developmental conditions during bud formation were responsible for the distinct spectral profiles and reduced lesion development observed in 2024. The clear year-level separation in PCA suggests that these baseline chemical differences were large enough to reduce cross-year model generalizability, emphasizing the importance of accounting for interannual chemical variation when developing FT-IR–based phenotyping tools. More broadly, seasonal and phenological transitions are expected to influence constitutive chemical profiles (Näsholm and Ericsson, 1990; Pomeroy et al., 1970) and, consequently, resistance phenotyping based on FT-IR spectra. In the present study, sampling was intentionally constrained to the shoot elongation period, which coincides with the biologically most relevant window for *D. sapinea* infection (Flowers et al., 2001), when elongating shoots are relatively supple and less lignified and artificial inoculations are most successful. Determining how resistance-related chemical signatures vary across seasonal transitions, and identifying optimal phenotyping windows, represents an important direction for future work; however, the objective here was not to capture the full seasonal dynamics of pine chemistry, but to evaluate whether resistance-related phenotypes can be detected during the ecologically relevant infection period.

Biochemical interpretability of the models is limited, but the wavenumbers identified as being most important for model performance can provide insight into the mechanisms underlying relative resistance. The wavenumbers most frequently selected by sPLS-DA belonged to four major spectral regions representing cell-wall polysaccharides (1000–1,050 cm^{-1} and 1,140–1,160 cm^{-1}), phenolics and proteins (1570–1,680 cm^{-1}), and aliphatic lipids/cutin (2950–3,140 cm^{-1}). These chemical domains overlap strongly with known defense-related pathways in conifers, including structural reinforcement, phenolic-mediated antifungal activity, and cuticular or suberized barriers and partly conform to previous studies (Conrad et al., 2020; Lim-Hing et al., 2025). The consistency of these biochemical groupings across models supports the conclusion that resistance to *D. sapinea* has a robust constitutive chemical signature that FT-IR can reliably detect.

Although significant differences in lesion length were detected among families, this study was not designed to provide a definitive test of genetic differentiation or to quantify heritable resistance. Nevertheless, the presence of family-level differences suggests that inherent variation in susceptibility to *D. sapinea* may exist within Swedish Scots pine material. Genetic differences in susceptibility to *D. sapinea* have been documented previously in pine species and populations (Gerhold et al., 1994; Swart et al., 1996; Terhonen et al., 2025), supporting the plausibility of a genetic component underlying the observed phenotypic patterns. More targeted experiments using balanced family designs, replicated environments, and quantitative genetic frameworks would be required to rigorously assess genetic resistance and its interaction with environmental stress in Swedish Scots pine populations. Importantly, however, the objective of the present study was not to resolve the genetic basis of resistance, but to evaluate whether resistance-related phenotypes can be identified directly at the individual-tree level using constitutive chemical fingerprints. From an applied perspective, this distinction is critical; a phenotyping approach that operates independently of family structure or provenance allows resistant individuals to be identified without prior genetic information, pedigree records, or population-level inference. Such an approach permits rapid screening of large numbers of individuals where resistance may arise from multiple genetic and physiological pathways rather than from discrete, easily heritable traits.

Disease development in this study was induced using an artificial inoculation method in which trees were wounded and supplied with a relatively large inoculum load. Although this approach is well established for *D. sapinea* and provides a consistent basis for comparative phenotyping, it does not replicate the pathogen's dominant infection strategy in nature. Under field conditions, *D. sapinea* commonly persists as a latent endophyte and becomes pathogenic primarily when host trees experience abiotic stress, most notably drought or hail damage, at which point pre-existing internal inoculum rapidly colonizes host tissues (Brodde et al., 2023; Wingfield et al., 2024). In contrast, the experimental design used here imposed an immediate and synchronized infection pressure that likely favors the expression of preformed, constitutive defenses over stress-induced responses. Consequently, the resistance phenotypes identified in this study should be interpreted as relative resistance under acute infection conditions, rather than as comprehensive resistance across the full spectrum of stress-mediated disease scenarios. How these phenotypes would perform under drought-induced susceptibility remains unknown, as studies explicitly quantifying tree-to-tree variation in *D. sapinea* symptom severity under controlled drought conditions are rare. Thus, while constitutive chemistry captured by FT-IR provides a robust and operational signal for early-stage screening, it may not fully predict disease outcomes under prolonged or compounding abiotic stress.

A spore-based inoculation system (Oostlander et al., 2023), which more closely mimics natural infection processes by enabling the

establishment of latent infections prior to the application of abiotic or other stresses, may help bridge this gap in future work, although current protocols still produce low and inconsistent latent infection rates. At present, there is also no practical way to identify latently infected but asymptomatic trees without destructive molecular or culturing assays, making it difficult to assess whether trees classified as resistant or susceptible here also differ in their propensity to establish latent infections. More gradual, stress-mediated disease systems, such as ash dieback, where phenotypes emerge through multi-year field assessments, may ultimately provide greater ecological realism. Nonetheless, for acute damage scenarios such as hail-triggered outbreaks, where infection behaves more like a wound pathogen, the rapid phenotyping approach demonstrated here may still capture biologically meaningful variation in constitutive defense chemistry that translates into real differences in disease severity.

Conclusion

FT-IR spectroscopy offers a rapid, high-throughput, and biologically interpretable approach for predicting relative resistance of Scots pine to *D. sapinea*. Shoot tissues provide the strongest predictive signal, leader tissues behave differently and are more difficult to model, and substantial year-to-year variation highlights the importance of phenological and environmental context. Together, these results support the integration of FT-IR–based chemotyping into early-stage breeding pipelines, providing a practical tool for accelerating resistance screening and improving long-term forest health management.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found at: https://zenodo.org/records/17732891.

Author contributions

PS: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Visualization, Writing – original draft, Writing – review & editing.

Funding

The author(s) declared that financial support was received for this work and/or its publication. This work was supported by Formas, the Swedish Research Council for Sustainable Development [grant number 2020-01795]; and Föreningen Skogsträdsförädling [grant number 22-491].

Acknowledgments

PS thanks AV Sherwood for reviewing the draft manuscript before submission.

Conflict of interest

The author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declared that Generative AI was used in the creation of this manuscript. ChatGPT v5.1 was used to help write R code for statistical tests and plotting.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/ffgc.2026.1755489/full#supplementary-material

References

Axelsson, E., and Axelsson, B. (1986). Changes in carbon allocation patterns in spruce and pine trees following irrigation and fertilization. *Tree Physiol.* 2, 189–204. doi: 10.1093/treephys/2.1-2-3.189

Bonello, P., Conrad, A. O., Sadiković, D., Liziniewicz, M., and Cleary, M. (2025). Point-of-care diagnostics and resistance phenotyping to combat ash dieback. *Front. Forests Glob. Change* 8:1588428. doi: 10.3389/ffgc.2025.1588428

Brodde, L., Adamson, K., Julio Camarero, J., Castaño, C., Drenkhan, R., Lehtijärvi, A., et al. (2019). *Diplodia* tip blight on its way to the north: drivers of disease emergence in northern Europe. *Front. Plant Sci.* 9:1818. doi: 10.3389/fpls.2018.01818

Brodde, L., Stein Åslund, M., Elfstrand, M., Oliva, J., Wågström, K., and Stenlid, J. (2023). *Diplodia sapinea* as a contributing factor in the crown dieback of scots

pine (*Pinus sylvestris*) after a severe drought. *For. Ecol. Manag.* 549:121436. doi: 10.1016/j.foreco.2023.121436

Conrad, A. O., Rodriguez-Saona, L. E., McPherson, B. A., Wood, D. L., and Bonello, P. (2014). Identification of *Quercus agrifolia* (coast live oak) resistant to the invasive pathogen *Phytophthora ramorum* in native stands using Fourier-transform infrared (FT-IR) spectroscopy. *Front. Plant Sci.* 5:521. doi: 10.3389/fpls.2014.00521

Conrad, A. O., Villari, C., Sherwood, P., and Bonello, P. (2020). Phenotyping Austrian pine for resistance using Fourier-transform infrared spectroscopy. *Arboric. Urban For.* 46, 276–286. doi: 10.48044/jauf.2020.020

Cozzolino, D. (2014). Use of infrared spectroscopy for in-field measurement and phenotyping of plant properties: instrumentation, data analysis, and examples. *Appl. Spectrosc. Rev.* 49, 564–584. doi: 10.1080/05704928.2013.878720

Desprez-Loustau, M.-L., Marçais, B., Nageleisen, L.-M., Piou, D., and Vannini, A. (2006). Interactive effects of drought and pathogens in forest trees. *Ann. For. Sci.* 63, 597–612. doi: 10.1051/forest:2006040

Dewar, R. C., Ludlow, A. R., and Dougherty, P. M. (1994). Environmental influences on carbon allocation in pines. *Ecol. Bull.* 43, 92–101.

Dorado, J., Almendros, G., Field, J. A., and Sierra-Alvarez, R. (2001). Infrared spectroscopy analysis of hemp (*Cannabis sativa*) after selective delignification by *Bjerkandera* sp. at different nitrogen levels. Enzym. Microb. Technol. 28:6, 550–559. doi:10.1016/S0141-0229(00)00363-X

Eyles, A., Bonello, P., Ganley, R., and Mohammed, C. (2010). Induced resistance to pests and pathogens in trees. *New Phytol.* 185, 893–908. doi: 10.1111/j.1469-8137.2009.03127.x

Flowers, J. L., Hartman, J. R., and Vaillancourt, L. J. (2006). Histology of *Diplodia pinea* in diseased and latently infected *Pinus nigra* shoots. *Forest Pathol.* 36, 447–459. doi: 10.1111/j.1439-0329.2006.00473.x

Flowers, J., Nuckles, E., Hartman, J., and Vaillancourt, L. J. (2001). Latent infection of Austrian and scots pine tissues by *Sphaeropsis sapinea*. *Plant Dis.* 85, 1107–1112. doi: 10.1094/PDIS.2001.85.10.1107

Gerhold, H. D., Rhodes, H. L., and Wenner, N. G. (1994). Screening *Pinus sylvestris* for resistance to *Sphaeropsis sapinea*. *Silvae Genet.* 43, 333–338.

Ghosh, S. K., Ishangulyyeva, G., Erbilgin, N., and Bonello, P. (2024). Terpenoids are involved in the expression of systemic-induced resistance in Austrian pine. *Plant Cell Environ.* 47, 2206–2227. doi: 10.1111/pce.14875

Guégan, J.-F., de Thoisy, B., Gomez-Gallego, M., and Jactel, H. (2023). World forests, global change, and emerging pests and pathogens. *Curr. Opin. Environ. Sustain.* 61:101266. doi: 10.1016/j.cosust.2023.101266

Hover, A., Buissart, F., Caraglio, Y., Heinz, C., Pailler, F., Ramel, M., et al. (2017). Growth phenology in *Pinus halepensis* Mill.: apical shoot bud content and shoot elongation. *Ann. For. Sci.* 74:39. doi: 10.1007/s13595-017-0637-y

Kačuráková, M., Capek, P., Sasinková, V., Wellner, N., and Ebringerová, A. (2000). FT-IR study of plant cell wall model compounds: pectic polysaccharides and hemicelluloses. *Carbohydr. Polym.* 43, 195–203. doi: 10.1016/S0144-8617(00)00151-X

Lim-Hing, S., Conrad, A. O., Montes, C. R., Gandhi, K. J. K., Payn, K. G., Walker, T. D., et al. (2025). Near-infrared spectroscopy as a high-throughput phenotyping method for fusiform rust resistance in loblolly pine. *Plant Phenomics* 7:100066. doi: 10.1016/j.plaphe.2025.100066

Martin, J. A., Solla, A., Coimbra, M. A., and Gil, L. (2008). Metabolic fingerprinting allows discrimination between *Ulmus pumila* and *U. Minor*, and between *U. minor* clones of different susceptibility to Dutch elm disease. *Forest Pathol.* 38, 244–256. doi: 10.1111/j.1439-0329.2007.00542.x

Müller, M. M., Hantula, J., Wingfield, M., and Drenkhan, R. (2019). *Diplodia sapinea* found on scots pine in Finland. *Forest Pathol.* 49:e12483. doi: 10.1111/efp.12483

Mutke, S., Gordo, J., Climent, J., and Gil, L. (2003). Shoot growth and phenology modelling of grafted stone pine (*Pinus pinea* L.) in inner Spain. *Ann. For. Sci.* 60, 527–537. doi: 10.1051/forest:2003046

Näsholm, T., and Ericsson, A. (1990). Seasonal changes in amino acids, protein and total nitrogen in needles of fertilized scots pine trees. *Tree Physiol.* 6, 267–281. doi: 10.1093/treephys/6.3.267

Oliva, J., Boberg, J., and Stenlid, J. (2013). First report of *Sphaeropsis sapinea* on scots pine (*Pinus sylvestris*) and Austrian pine (*P. nigra*) in Sweden. *New Dis. Rep.* 27:23. doi: 10.5197/j.2044-0588.2013.027.023

Oostlander, A. G., Brodde, L., von Bargen, M., Leiterholt, M., Trautmann, D., Enderle, R., et al. (2023). A reliable and simple method for the production of viable Pycnidiospores of the pine pathogen *Diplodia sapinea* and a spore-based infection assay on Scots pine. *Plant Dis.* 107, 3370–3377. doi: 10.1094/PDIS-01-23-0107-RE

Pandey, K. K., and Pitman, A. J. (2003). Ftir studies of the changes in wood chemistry following decay by brown-rot and white-rot fungi. *Int. Biodeterior. Biodegrad.* 52, 151–160. doi: 10.1016/S0964-8305(03)00052-0

Phillips, A. J. L., Alves, A., Abdollahzadeh, J., Slippers, B., Wingfield, M. J., Groenewald, J. Z., et al. (2013). The Botryosphaeriaceae: genera and species known from culture. *Stud. Mycol.* 76, 51–167. doi: 10.3114/sim0021

Pomeroy, M. K., Siminovitch, D., and Wightman, F. (1970). Seasonal biochemical changes in the living bark and needles of red pine (*Pinus resinosa*) in relation to adaptation to freezing. *Can. J. Bot.* 48, 953–967. doi: 10.1139/b70-134

Ramsfield, T. D., Bentz, B. J., Faccoli, M., Jactel, H., and Brockerhoff, E. G. (2016). Forest health in a changing world: effects of globalization and climate change on forest insect and pathogen impacts. *Forestry* 89, 245–252. doi: 10.1093/forestry/cpw018

Schulz, H., and Baranska, M. (2007). Identification and quantification of valuable plant substances by IR and Raman spectroscopy. *Vib. Spectrosc.* 43, 13–25. doi: 10.1016/j.vibspec.2006.06.001

Sene, C., McCann, M. C., Wilson, R. H., and Grinter, R. (1994). Fourier-transform Raman and Fourier-transform infrared spectroscopy (an investigation of five higher plant cell walls and their components). *Plant Physiol.* 106, 1623–1631. doi: 10.1104/pp.106.4.1623

Sherwood, P., and Bonello, P. (2016). Testing the systemic induced resistance hypothesis with Austrian pine and *Diplodia sapinea*. *Physiol. Mol. Plant Pathol.* 94, 118–125. doi: 10.1016/j.pmpp.2016.06.002

Showalter, D. N., Raffa, K. F., Sniezko, R. A., Herms, D. A., Liebhold, A. M., Smith, J. A., et al. (2018). Strategic development of tree resistance against forest pathogen and insect invasions in defense-free space. *Front. Ecol. Evol.* 6:124. doi: 10.3389/fevo.2018.00124

Sniezko, R. A., and Koch, J. (2017). Breeding trees resistant to insects and diseases: putting theory into application. *Biol. Invasions* 19, 3377–3400. doi: 10.1007/s10530-017-1482-5

Sniezko, R. A., and Liu, J.-J. (2021). Prospects for developing durable resistance in populations of forest trees. *New For.* 54, 751–767. doi: 10.1007/s11056-021-09898-3

Swart, W. J., Donald, D. G. M., and Theron, J. M. (1996). Screening of *pinus radiata* progenies for resistance to *sphaeropsis sapinea*. *S. Afr. For. J.* 175, 15–18. doi: 10.1080/00382167.1996.9629887

Terhonen, E., Kujala, S., Pyhäjärvi, T., and Sutela, S. (2025). Genetic variation of resistance in scots pine as possible solution against *Diplodia sapinea*. *Silva Fenn.* 59:25028. doi: 10.14214/sf.25028

Türker-Kaya, S., and Huck, C. W. (2017). A review of mid-infrared and near-infrared imaging: principles, concepts and applications in plant tissue analysis. *Molecules* 22:1. doi: 10.3390/molecules22010168

Villari, C., Dowkiw, A., Enderle, R., Ghasemkhani, M., Kirisits, T., Kjaer, E. D., et al. (2018). Advanced spectroscopy-based phenotyping offers a potential solution to the ash dieback epidemic. *Sci. Rep.* 8:9. doi: 10.1038/s41598-018-35770-0

Wallis, C., Eyles, A., Chorbadjian, R., McSpadden Gardener, B., Hansen, R., Cipollini, D., et al. (2008). Systemic induction of phloem secondary metabolism and its relationship to resistance to a canker pathogen in Austrian pine. *New Phytol.* 177, 767–778. doi: 10.1111/j.1469-8137.2007.02307.x

Wingfield, M. J., Slippers, B., Barnes, I., Duong, T. A., and Wingfield, B. D. (2024). The pine pathogen *Diplodia sapinea*: expanding frontiers. *Curr. For. Rep.* 11:2. doi: 10.1007/s40725-024-00236-2

Zlatković, M., Keča, N., Wingfield, M. J., Jami, F., and Slippers, B. (2017). New and unexpected host associations for *Diplodia sapinea* in the Western Balkans. *For. Pathol.* 47:e12328. doi: 10.1111/efp.12328