New Phytologist

European oak chemical diversity – from ecotypes to herbivore resistance

Marko Bertić[1], Hilke Schroeder[2], Birgit Kersten[2], Matthias Fladung[2], Franziska Orgel[2], Franz Buegger[3], Jörg-Peter Schnitzler[1] and Andrea Ghirardo[1]

[1]Research Unit Environmental Simulation (EUS), Institute of Biochemical Plant Pathology, Helmholtz Zentrum München, Ingolstädter Landstr. 1, 85764 Neuherberg, Germany; [2]Thünen-Institute of Forest Genetics, Sieker Landstrasse 2, 22927 Grosshansdorf, Germany; [3]Institute of Biochemical Plant Pathology (BIOP), Helmholtz Zentrum München, Ingolstädter Landstr. 1, 85764 Neuherberg, Germany

Author for correspondence:
Andrea Ghirardo
Email: andrea.ghirardo@helmholtz-muenchen.de

Received: *21 May 2021*
Accepted: *5 July 2021*

New Phytologist (2021) **232:** 818–834
doi: 10.1111/nph.17608

Key words: adaptation, biomarker, chemotype, eco-metabolomics, ecotype, flavonoids, *Quercus robur*, *Tortrix viridana*.

Summary

- Climate change is increasing insect pressure and forcing plants to adapt. Although chemotypic differentiation and phenotypic plasticity in spatially separated tree populations are known for decades, understanding their importance in herbivory resistance across forests remains challenging.
- We studied four oak forest stands in Germany using nontarget metabolomics, elemental analysis, and chemometrics and mapped the leaf metabolome of herbivore-resistant (T-) and herbivore-susceptible (S-) European oaks (*Quercus robur*) to *Tortrix viridana*, an oak pest that causes severe forest defoliation.
- Among the detected metabolites, we identified reliable metabolic biomarkers to distinguish S- and T-oak trees. Chemotypic differentiation resulted in metabolic shifts of primary and secondary leaf metabolism. Across forests, T-oaks allocate resources towards constitutive chemical defense enriched of polyphenolic compounds, e.g. the flavonoids kaempferol, kaempferol and quercetin glucosides, while S-oaks towards growth-promoting substances such as carbohydrates and amino-acid derivatives.
- This extensive work across natural forests shows that oaks' resistance and susceptibility to herbivory are linked to growth-defense trade-offs of leaf metabolism. The discovery of biomarkers and the developed predictive model pave the way to understand *Quercus robur*'s susceptibility to herbivore attack and to support forest management, contributing to the preservation of oak forests in Europe.

Introduction

Forest ecosystems are currently exposed to opposing forces of climate change. For example, global warming positively affects plant growth (Saxe *et al.*, 2001; Pretzsch *et al.*, 2014) and extends tree growth periods but also increases herbivory pressure on forest ecosystems. Milder winters improve survival of overwintering insects (Pureswaran *et al.*, 2018), and a warmer climate accelerates insects' development and increases voltinism, the number of generation cycles per year (Hamann *et al.*, 2021). The formation of more insects per growing season has negative consequences for plant fitness and growth (Bebber *et al.*, 2013; Bacon *et al.*, 2014).

Overall, growth and reproductive allocations support plants improving inter-species competition, while investments in defense mechanisms help plants to resist, e.g. herbivorous pressure, maximizing plant fitness by reducing damage and improving survival. However, resistant phenotypes come at a cost: resource allocation towards chemical defense reduces growth by diverting energy and metabolic precursors from processes such as vegetative tissue expansion, biomass yield, and seed production. This is predicted in the classical growth-differentiation-balance (GDB) hypothesis developed by Herms & Mattson (1992). It describes the physiological case for a mutually exclusive allocation of limited resources to one or the other function – growth vs differentiation (including defense) to maximize the plant's fitness. Thereby, the plant's dilemma between growth or defense has important ecological consequences. At the intraspecific level, it is expected that both growth and defense increase with resource availability (van Noordwijk & de Jong, 1986; Agrawal, 2020). Furthermore, individuals with greater access to limited resources can allocate larger amounts of resources to growth and defense than conspecifics who have limited access as a consequence of phenotypic plasticity (Agrawal, 2020).

In addition to phenotypic plasticity, plants' adaptation to changing environmental conditions and herbivore pressure is based on intraspecific genetic variation at individual and population levels. Through the formation of chemotypes, it provides the potential for adaptation under natural selection. Chemotypic

differentiation in plant populations is influenced by various environmental factors, pathogen and herbivore pressure, and their interactions (Levin, 1976; Bradshaw, 1984; van Tienderen, 1992; Linhart & Grant, 1996; Galloway & Fenster, 2000; Montalvo & Ellstrand, 2000; Etterson, 2004). As a response, spatially dispersed populations tend to adapt to specific ecological niches and evolve genotypic and metabolic patterns, the so-called 'ecotypes' (Kollmann & Bañuelos, 2004; Kleessen *et al.*, 2012; Nagler *et al.*, 2018; Salomé-Abarca *et al.*, 2020). Because phenotypic differences among members of a population driven by genetic differentiation lead to adaptation through natural selection (Alberto *et al.*, 2013), plant phenotypes with increased resistance to herbivores may become selected over generations under higher herbivory pressure (Geber & Griffen, 2003; Schemske *et al.*, 2009). However, plants that invest in growth/reproduction are more competitive, especially during periods in the absence of abiotic or biotic stresses. To date, there is still scarce evidence on the prioritization of defensive vs growth strategies in naturally occurring tree populations. To study the susceptibility of our forests to the effects of climate change, we need to better understand the phenotypic variation in relation to herbivory pressure and chemodiversity within a forest tree population and how they occur among members of a population and among populations in natural forests. The ecological consequences of intraspecific chemodiversity and the impacts on the interactions of plants with their biotic environment remain largely unknown (Müller *et al.*, 2020). A valuable tool to understand plant–herbivory interactions is the identification of biomarkers, metabolites that are characteristic to the susceptibility or resistance of plants to insects.

Here we analyzed the interactions between phenotype, chemotype and herbivore resistance in the European oak (*Quercus robur* L.) by the development of metabolic biomarkers. *Quercus robur* is a widespread, ecologically, and economically important long-lived forest tree species (Yela & Lawton, 1997; Ehmcke & Grosser, 2014; Annighöfer *et al.*, 2015). Native to Europe, it is cultivated in the temperate regions of Asia and North America (Eaton *et al.*, 2016) and has been described as tolerant to high temperatures and drought. European oaks are characterized by a high genetic diversity, which helps species' adaptation under changing environmental conditions (Müller-Starck *et al.*, 1993; Bresson *et al.*, 2011). In previous works, we investigated grafted trees from an oak population grown under controlled conditions and identified two naturally occurring, insect-susceptible (S-) and insect-resistant (T-) European oak phenotypes (Ghirardo *et al.*, 2012; Kersten *et al.*, 2013). In the forests, the scions' mother plants are differently defoliated by the green oak leaf roller (*Tortrix viridana* L.), an herbivorous pest that can cause severe damage in oak forests during outbreak years (Hunter, 1990; Hartmann & Blank, 1992). Metabolomic and transcriptomic analyses of plants grown in a phytochamber let us hypothesize that S-oaks might prioritize the biosynthesis of metabolites related to growth processes, while T-oaks to constitutive defense (Kersten *et al.*, 2013).

In this article, we tested the growth-defense prioritization hypothesis of resistant and susceptible oaks in natural forests by developing metabolic biomarkers that can distinguish T-oaks and S-oaks irrespective of the oak populations in metabolically contrasting forests (ecotypes). For this, we analyzed the leaf metabolome of T- and S-oaks in four geographically separated forest sites of European oaks in North Rhine-Westphalia, Germany, with respect to geolocation and resistance to the insect *T. viridana*. We used a nontargeted metabolomics approach to study the oak's strategy in terms of prioritization towards growth or defense and related to tree-specific oaks' resistance/susceptibility inventory data collected over the last 24 yr at the four sites. Nontargeted metabolomics is a useful tool to understand plant–environmental interactions (Kuzina *et al.*, 2009; Peñuelas & Sardans, 2009), and its use has recently increased to address ecometabolomic questions (Sardans *et al.*, 2011; Rivas-Ubach *et al.*, 2013, 2019; D. M. Allevato *et al.*, 2019).

Here we show that (1) each oak population is metabolically a different ecotype; (2) each ecotype community has conserved both resistant and susceptible phenotypes in regards to *T. viridana* infestation; (3) metabolic biomarkers and chemometrics can be used to develop a prediction model capable of classifying S- or T-oak phenotypes independently of their geographical origins. These results pave the way to study the European oak's susceptibility to herbivore attacks and may contribute to the long-term conservation of European oaks, an ecologically important tree species grown in temperate latitudes of Europe.

Materials and Methods

Plant material

We investigated the dark-adapted leaf metabolome of 67 European oak trees (*Q. robur* L.) from four mixed forest stands near the cities of Borken, Everswinkel, Münster, and Warendorf (data set BEMW) in North Rhine-Westphalia, Germany (Fig. 1a; Table 1) (Schroeder & Degen, 2008). The range of tree ages is 150–185 yr. Oaks were phenotypically classified as insect-resistant (T-oaks) or as insect-susceptible (S-oaks) by monitoring from 1994 to 2018 the defoliation rates that occurred as a consequence of outbreaks years of the specialized herbivore *T. viridana* L. (Lepidoptera, Tortricidae). The degree of herbivory defoliation has been visually estimated as percentage of defoliation using standard images (Evers, 2004). In outbreak years, all trees with a foliage loss of > 90% were classified as S-oaks (41 trees), and those of < 60% as T-oaks (26 trees). Exceptionally for Warendorf, we classified as potential T-oaks those trees that had the lowest (< 80%) percentage of defoliation, as the overall insect defoliations were much more severe.

From 21 to 24 May 2019, we sampled sun-exposed leaves of the tree crowns (Fig. 1b) by taking down branches of the trees using an arborist throw-line launcher (Youngentob *et al.*, 2016). This method was necessary for reaching the small top branches of the 20–40 m spreading crown of tall oak trees. For each tree, we collected three leaves from different branches and inserted them in aluminum foil bags. The three leaves per tree were combined for analysis to account for metabolic variation within the canopy. To minimize the influence of developmental-dependent changes (Riipi *et al.*, 2004), all leaves were of the same growth stage #5

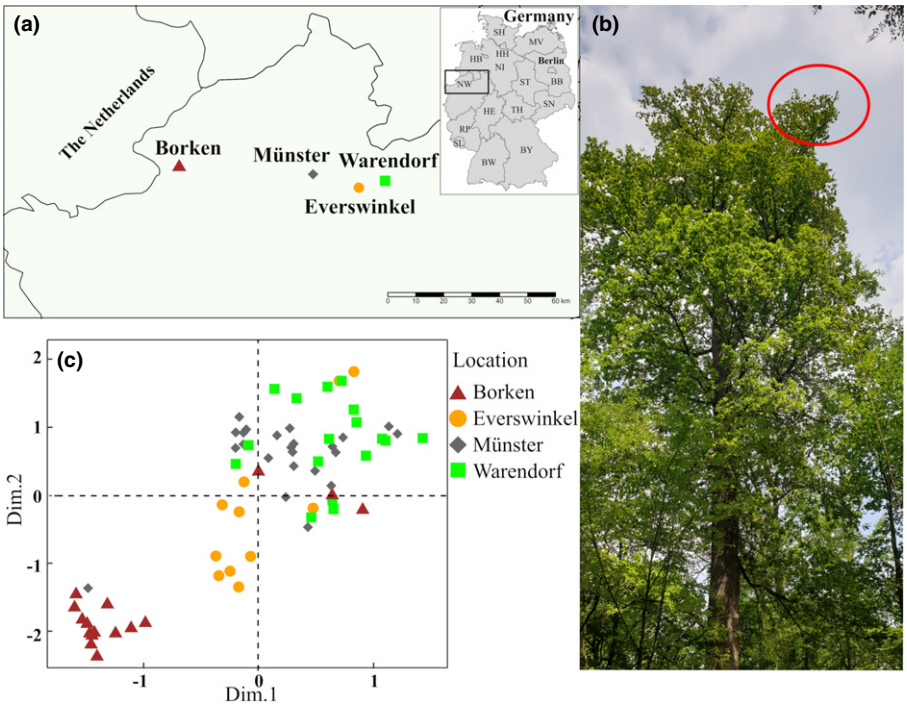

Fig. 1 (a) Map of the four forest sites studied (Borken, geographic coordinates (latitude, longitude): (52.021493, 6.938061), Everswinkel (51.907194, 7.884056), Münster (51.906404, 7.751286) and Warendorf (51.834023, 7.891956)) in North Rhine-Westphalia, Germany drawn with R package 'TMAP' (Tennekes, 2018). The site Borken is *c.* 50 km further from the other three sites, located in a radius of 20 km. (b) A representative herbivore-resistant (T-) oak and typical sampling branch (red circle) at the top of the canopy directly exposed to the sunlight. (c) Unsupervised uniform manifold approximation and projection (UMAP) with R package 'UMAP' (McInnes *et al*., 2018) of all mass features measured by nontargeted metabolomics indicates differences in the overall leaf metabolic pattern of oaks from the four forest stands.

Table 1 Environmental factors and forest characteristics of the four forest sites in North Rhine-Westphalia, Germany, during the sampling campaign in 2019.

	Borken	Everswinkel	Münster	Warendorf
Number of trees	26	12	24	18
Coordinates (lat., long.)	(52.021493, 6.938061)	(51.907194, 7.884056)	(51.906404, 7.751286)	(51.834023, 7.891956)
T (°C) $\pm \sigma$	23.6 ± 2.2	12.9 ± 1.6	19.0 ± 2.1	14.6 ± 1.9
Relative humidity (%) $\pm \sigma$	34.7 ± 10.2	75.5 ± 6.7	49.4 ± 13.2	90.8 ± 10.1
Weather	Sunny	Cloudy	Sunny	Cloudy
Soil	Brown earth podsol	Alluvial soil	Alluvial soil	Brown earth from marl and calcareous gravel
Age	150–180	150–180	156	185
Quercus robur percent forest composition	NA/mixed forest	NA/mixed forest	60/(*Fagus sylvatica*, *Carpinus betulus*, *Faxinus excelsior*)	70/(*Q. petraea*, *Fagus sylvatica*)

The temperature and humidity values are shown as the mean value and standard deviation (σ). NA, not available.

(Supporting Information Fig. S1). We randomly sampled T- and S-oaks between 10:00 and 17:00 h. Due to difficulties in performing the flash-freezing of samples collected by the throw-line launcher method, in the availability of liquid-nitrogen (N_2) in the forests, and sensitivity of some metabolites to changes of light condition unavoidable in the nature, we dark-adapted the sampled leaves for 60–80 min inside a polystyrene box before deep-freezing in dry ice. Such procedure helps to reduce photosynthetic-dependent metabolic variation (Dyson *et al*., 2015) and lower bias in the analysis. Laboratory analysis indicated that the metabolome of dark-adapted leaves differs from those sampled under light, although most of the compounds differently expressed in T- and S-oaks are end-products (e.g. tannins, sugars, lipids, amino acids) remain unchanged (see also validation later).

Complementary and validation analyses were performed on grafted oaks originating from a different oak population of a forest *c.* 300 km away from the forest stands in BEMW near the village of Asbeck (data set Asbeck) in North Rhine-Westphalia (Schröder, 2010; Kersten *et al*., 2013). Branches of these oaks were grafted in 2008 and were previously subjected to metabolic and molecular studies under controlled environmental conditions (Ghirardo *et al*., 2012). Since 2008, the grafted oaks have been growing in a common garden under similar climate and fertilization conditions for the last 12 yr at Thünen-Institute of Forest Genetics, Grosshansdorf, Germany. Leaves from 10 ramets each of five oaks and 50 individuals (20 S- and 30 T-oaks) were collected. Because the Asbeck data-set contains S- and T-oaks growth independently to climate conditions, oak population, sampling time, insect feeding, and sampling procedure (dark-

adapted vs light conditions and liquid-N_2 vs dry-ice), we use this data to evaluate the prediction model performance of the here newly developed statistical approach based on dark-adapted metabolome of BEMW field samples (see later). For this, larvae have been set on grafted trees at 17:30 h and leaves from these trees have been sampled 19 h later by flash-freezing with liquid-N_2. Samples were stored at −80°C until metabolomics analysis.

Nontargeted metabolomics analysis

All leaf material was homogenized under cryogenic condition using mortar, pestle, and liquid-N_2 to a fine powder and then freeze-dried at −50°C under the vacuum condition of 0.040 mbar (Alpha 1-4 LDplus, Christ, Osterrode, Germany). For extraction, 500 µl of cold (5°C) methanol : 2-propanol : water (1 : 1 : 1, v/v/v) extraction solvent mixture were added to 25 mg of dried leaf powder containing 50 µl of internal standard (IS) mixture (0.028 µmol ml^{-1} of magnolol, rosmarinic acid, 3,4-dihydromandelic acid, and 3′,4′-dihydroxyacetophenone, Table S1). Samples were mixed for 1 min inside a 2 ml polypropylene tube and sonicated in an ultrasonic bath for 10 min at 5°C. The solution was centrifuged at 9.3 ***g*** for 10 min at 5°C, and 400 µl of supernatant was recovered. The extraction was repeated with the rest of the precipitate, resulting in 800 µl of supernatant. The supernatant was dried by SpeedVac (Univapo 150H, Uniequip, Planegg, Germany), and the residue was dissolved in 350 µl of 50% (v/v) acetonitrile in water. The solution was mixed for 1 min and centrifuged at 9.3 ***g*** for 10 min at 5°C, and 300 µl of supernatant was transferred in 350 µl amber glass vials. The chemicals (LCMS hyper grade) methanol/water were purchased from Merck (Darmstadt, Germany) and 2-propanol/acetonitrile from Honeywell (Puchheim, Germany).

Nontargeted metabolomics analysis was performed following Ghirardo *et al.* (2020) and Hemmler *et al.* (2018), using an ultra-performance liquid chromatography (UPLC) ultra-high resolution (UHR) tandem quadrupole/time-of-flight (QqToF) mass spectrometry (MS). The instrument is comprised of an Ultimate 3000RS UPLC (Thermo Fisher, Bremen, Germany), a Bruker Impact II (QqToF) and an Apollo II electrospray ionization (ESI) source (Bruker Daltonic, Bremen, Germany). Separation of nonpolar and polar metabolites was achieved using a reversed-phase liquid chromatography (RPLC) column (C_{18}, Acquity BEH (Waters, Eschborn, Germany), 150 mm × 2.1 mm, 1.7 µm) and a hydrophilic interaction liquid chromatography (HILIC) column (Acquity BEH Amide, Waters, 100 mm × 2.1 mm, 1.7 µm), respectively. Each sample was separately analyzed with RPLC, and HILIC columns with MS operated in both positive (+) and negative (−) ESI modes. For details on chromatography and MS parameters see Methods S1.

Data processing and compound identification

The liquid chromatography-tandem mass spectrometry (LC-MS/MS) data were analyzed using METABOSCAPE 4.0 (Bruker) to perform the post-acquisition peak-peaking, alignments, isotope filtering, and peak-grouping based on peak-area correlation (Domingo-Almenara *et al.*, 2018). Detailed parameter settings are listed in Table S2. The software merged automatically the data obtained from the ± measurement modes and returns processed data in a result table. Results from RP and HILIC analyses were merged manually. The table contains the exact mass, chemical formula and peak area of compounds if the expected adducts and isotopologues co-occurred at $r^2 > 0.7$. For the most abundant compounds, we confirmed the annotation by standards and therefore achieved the annotation level 1 (Sumner *et al.*, 2007; Reisdorph *et al.*, 2020). For other compounds, the annotation was putative and achieved using MS/MS spectra, when available, matched by METABOSCAPE to the libraries HMDB (http://www.hmdb.ca/) (Wishart *et al.*, 2009), GNPS (Global Natural Product Social Molecular Networking) (https://gnps.ucsd.edu/ProteoSAFe/static/gnps-splash.jsp), MONA (Mass Bank of North America), Vaniya/Fiehn Natural Products Library; Fiehn HILIC; RESPECT (http://spectra.psc.riken.jp) (Sawada *et al.*, 2012); LC-MS/MS Spectra (https://mona.fiehnlab.ucdavis.edu/downloads). Mass features without MS/MS spectra were tentatively annotated on MS^1 level using 5.0 mDa tolerance for the precursor mass with an in-house built R code (https://osf.io/s9d2j/?view_only=733a0c1a9e444f669d44c6eaad44f253). Chemical class classification of compounds was achieved by the 'multidimensional stoichiometric compound classification' (MSCC) approach according to the elemental ratio compositions (Rivas-Ubach *et al.*, 2018).

Data processing included the replacement of missing values with the average area value from all samples for the corresponding mass feature (Denkert *et al.*, 2006). Data were normalized by IS (Sysi-Aho *et al.*, 2007) mixture, composed of four plant metabolites that spread through the whole analytical mass and RT range (Table S1) but not detectable in pure oak leaf extracts. Finally, peak areas were normalized to dry leaf weight. The final data table (Table S3) was the input of the statistical analysis.

Elemental and stable isotope analyses

For the determination of the content of macroelements (Ca, K, Na, S, P), 90.00 mg of freeze-dried leaf powder was extracted as described by Schramel *et al.* (1980) and further analyzed by inductively coupled plasma optical emission spectrometry (ICP-OES) (Schramel, 1983).

Carbon (C) and nitrogen (N) contents and stable isotope signature of $\delta^{13}C$ and $\delta^{15}N$ were measured by isotope ratio mass spectrometry (IRMS, delta-V Advantage, Thermo Fisher, Dreieich, Germany) coupled to an elemental analyzer (Euro EA, Eurovector, Milan, Italy) as described in Methods S2.

Multivariate data analysis

Before multivariate analyses, data were always logarithmically (log_{10}) transformed, centered, and Pareto scaled (van den Berg *et al.*, 2006; Eriksson *et al.*, 2013).

Descriptive analysis The principal component analysis (PCA) and orthogonal partial least square regression discriminant analysis (OPLS-DA) were performed using the SIMCA-P v.13.0.3.0 (Umetrics, Umeå, Sweden). Uniform manifold approximation and projection (UMAP) and cluster analysis (CA) were made with R software (R Development Core Team, 2019) using packages 'UMAP' (McInnes *et al.*, 2018) and 'HEATMAP' (Gu *et al.*, 2016) respectively. The unsupervised PCA and UMAP methods were used to describe the metabolic patterns and detect potential outliers. Individual trees were scores, and the normalized peak areas of mass features were the loadings in PCA, UMAP and OPLS-DA. To visualize metabolic differences between both four locations and the S- and T-oak phenotypes, respectively, CA (based on Euclidean distances) was performed on forest-specific and phenotype-specific discriminant mass features. Hypergeometric tests were performed using the function 'phyper' in R v.3.6.0 (R Development Core Team, 2019).

Discriminant analysis for biomarker discovery and classification of oak phenotypes OPLS-DA was used to found forest-specific discriminant mass features on BEMW oaks. The discriminant model was computed using four *Y*-variables that corresponded to the four forests BEMW and assigning a binary discriminating variable codex to their class (Ghirardo *et al.*, 2005). The dimensions of the data matrix analyzed were 67 × 10 206 (trees × normalized peak areas to IS and leaf dry-weight), representing the individual trees (from the four groups of forest sites) and the mass features, respectively. Discriminant metabolites were defined as the loadings of the significant OPLS-DA (cross-validated ANOVA < 0.05; Eriksson *et al.*, 2008) that passed the following criteria: (1) VIP values > 2 (Cocchi *et al.*, 2018); (2) adjusted *P*-value < 0.05 (*t*-test and Benjamini–Hochberg correction (Benjamini & Hochberg, 1995)). The adj-*P*-value was computed for all six possible forest site comparisons, resulting in six values for each mass feature.

To build a multivariate prediction model that allows classifying the European oaks in T- and S-oaks on the base of the metabolic profiles, we deployed OPLS-DA, as MS data with partial least squares regression can be used for biomarker discovery and classification purposes (Ghirardo *et al.*, 2005; Wiklund *et al.*, 2008; Boccard & Rutledge, 2013). We computed the prediction model using 50 trees (training set) randomly selected from the dataset BEMW and kept the remaining nine trees (internal validation set, ratio training : validation = 80 : 20) to test the ability of the model to predict the oak phenotype. Such validation is termed 'internal' as the data originate from the same oak populations. The dimensions of the training set data matrix were 50 × 10 206 (trees × normalized peak areas to IS and leaf dry-weight). The model was computed using two *Y*-variables representing the T- and S-oak phenotypes. To identify reliable biomarkers, we maximized the regression sum of squares (R^2Y), the prediction sum of squares (Q^2Y), and minimizing the root mean square error of estimation (RMSEE) and RMSEcv (root mean square error of cross-validation) by reducing the number of *X*-variables to those that have VIP > 1.0, relative averaged abundances of T-/S- metabolites with $\log_2$ (T/S) ratio of < −0.5 or > 0.5, significant changes of individual metabolite (*t*-test and Benjamini–Hochberg correction) and group separation (CV-ANOVA) between T-/S-oaks at adj-*P*-value < 0.05. The resulting compounds were defined as phenotypic biomarker candidates. Furthermore, we defined borders between S- and T-oak phenotypes (0–0.4 T-oaks, 0.4–0.6 unclassified, 0.6–1.0 S-oaks) and introduced the range near the border as 'unclassified'.

Model validation To evaluate the model's robustness (Andersen *et al.*, 2006; Broadhurst & Kell, 2006; Worley & Powers, 2013) in predicting the oak phenotype, the OPLS-DA model was tested with the independent dataset Asbeck, a procedure termed 'external validation'.

Results

Eco-metabolomic analysis reveals specific *in situ* metabolic signatures of four oak forest sites

We studied the leaf metabolome of *Q. robur* growing in four geographically separated forest stands of North Rhine-Westphalia, Germany (Tables 1, S3; Fig. 1a). An unsupervised UMAP analysis of 10 206 metabolomic-related mass features detected in this study suggested differences in the overall metabolic pattern of the forest stands, as indicated by the first two dimensions (Fig. 1c). We disentangled the discriminant mass features and studied in details the 100 most location-correlated metabolites that showed significant differences (OPLS-DA, CV-ANOVA, $P < 0.05$) in the metabolic profiles of the oaks at the four forest sites (Fig. 2). Hierarchical clustering and heatmaps illustrate the changes in the relative abundance of metabolites related to the metabolisms of proteins, carbohydrates, lipids, nucleotides, and secondary metabolites (Fig. 3). Although these 100 metabolites were present in leaves at all four sites, the metabolic profile of *Q. robur* leaves from the four forest sites differed in terms of relative abundance in the primary metabolisms of protein, lipids, carbohydrates, and nucleotides or to the secondary metabolisms of condensed tannins (CTs) and flavonoids. We referred to these metabolites hereafter as site-discriminant metabolites. The forest stand Borken, which is geographically most distinct to the others (Fig. 1a), displayed the most different chemical composition in lipids and secondary metabolites (Fig. 2). All forest stands, located in a *c.* 50 km radius, are growing under similar climate conditions. Complementary analysis on eight oaks (two from each forest, with two replicates of each tree) indicated that in respect to tree-to-tree variation, the within-tree variation accounts from 0.52% to 11.25% (Table S4). Therefore, the analysis indicated the presence of four *Q. robur* ecotypes, exhibiting a distinguishable chemical profile.

Because specific patterns in the metabolomes may be attributable to resource limitations in nutrient availability, we analyzed the macroelement composition (Ca, K, Na, S, N, P, and C) and the stable isotope signature of C and N ($\delta^{13}C$ and $\delta^{15}N$) of the leaves. In general, the contents of all analyzed macro

elements and isotopic signature poorly correlated with geolocation, except for the overall C and N content and their stable isotope signature (Fig. 2c–f). The $\delta^{15}N$ were lowest in S-oak leaves from Warendorf and $\delta^{13}C$ highest in leaves from the forest stand in Münster (Fig. S2). Taken together, our analysis showed that the four forest stands are composed of four corresponding *Q. robur* populations that possess chemically different leaf metabolite profiles independent from nutrient availability.

Outbreaks of *Tortrix viridana* unveils the resistant oak phenotype in all four-forest sites

The periodic outbreaks of the herbivore *T. viridana* cause infestations and severe *Q. robur* defoliation (Schröder, 2010; Ghirardo *et al.*, 2012). However, at all investigated sites, some of the European oak trees become much less defoliated during *T. viridana* outbreaks, compared to almost entirely defoliated neighboring oaks. Analysis of the degree of defoliation showed a clear insect-resistant phenotype (T-oaks) compared to an almost entirely defoliated susceptible phenotype (S-oaks) (Fig. 4). Across the four forest sites, 66 S-oaks and 44 T-oaks have been monitored for 24 yr, and differences in the canopy defoliation were always observed consistently different and statistically significant (two-sample *t*-test, $P < 0.05$). The percentage of S-oaks in the sample set differs from forest to forest and ranges from 40.7% at the stand near Borken to 88.0% at the stand near Warendorf.

Trade-offs of growth and defense-related compounds in the herbivory-resistance of the two oak phenotypes across forests

Knowing that the four metabolically different forest stands are composed of the two different T- and S-oak phenotypes, we question whether the two phenotypes' metabolic fingerprint is consistent across the four forests. CA and UMAP showed that the chemical differences in S- and T-oaks are consistent across the four forests (Figs 5, S3; Table S3). Because compound identification in metabolomics study is still in its infancy (Domingo-Almenara *et al.*, 2018), we used the Van Krevelen diagram in combination with MSCC to classify all those detected mass features that could be assigned to a chemical formula but not to a specific chemical compound (Fig. 5b). This comprehensive analysis showed that in all the forest samples, carbohydrates and amino acids were overrepresented in S-oak, whereas flavonoids and their glucosides were overrepresented in T-oaks (adj. *P*-value < 0.05, hypergeometric test, Table S5). In turn, this statistical result suggested a constitutive strategy of S-oaks to invest plant resources towards growth, and T-oak leaves in defensive compounds against herbivorous feeding.

Fig. 6 shows the top 10 most abundant metabolites discovered using our approach for each chemical class (amino acids and protein-related, sugars, lipids and secondary compounds). Among them, the defensive compound kaempferol, kaempferol-3-*O*-glucoside, quercetin-3-glucuronide, quercetin-3-*O*-malonylglucoside, quinic acid and pipecolinic acid showed significant increased levels in T-oaks together with the levels of some NSC (nonstructural carbohydrates) such as maltose, glucose and arabinose and the amino acid tryptophan (Fig. 6, adj-*P* values < 0.05). By contrast, S-oaks contained higher levels of the sugar glucose 1,6-bisphosphate, catechin (component of proanthocyanidins), the putatively identified hydrolyzable tannin 1,2,3,6-tetrakis-*O*-galloyl-β-D-glucose, and the precursor of gallotannins 1-*O*-galloyl-β-D-glucose (syn. glucogallin) (adj-*P* values < 0.05) (Fig. 6). Levels of quercetin was found unchanged.

Therefore, nontargeted metabolomic analysis correlated metabolites to the resistant phenotype that are involved in plant defense mechanisms.

Combining metabolomics and chemometrics predicts *Tortrix viridana*-resistant oak phenotype

The discovery of herbivory-resistant biomarkers among ecotypes is a key for developing a prediction model that can classify the T- and S-phenotypes of European oak. Based on VIP values, $\log_2$ (T/S) ratio, and *P*-values (see the Materials and Methods section), we selected 17 metabolites and tested them as biomarkers to discriminate oaks in *T. viridana*-resistant trees. We built a chemometric prediction model based on OPLS-DA, which was highly robust ($R^2Y = 0.86$, $Q^2Y = 0.81$, *P*-values $= 1.08 \times 10^{-15}$, CV-ANOVA) (Fig. 7a,b; Table 2). Then, we first used the model to test its ability to classify a subset of nine oaks from the four forest stands in North Rhine-Westphalia that were not used in building the model, a procedure termed as internal model validation. The prediction model's accuracy was 100% for both S- and T-oaks (Fig. 7c). This initial analysis demonstrated that the model could predict the oak phenotype from the same tree communities used to calibrate the model. To validate the robustness of our model, we performed an external validation (Anderssen *et al.*, 2006; Broadhurst & Kell, 2006; Worley & Powers, 2013): we used the 50 oaks from the Asbeck population growing under different environmental conditions and sampled differently than those BEMW oaks used in building the model (see the Materials and Methods section). The prediction model's accuracy was 100% for both S-oaks and T-oaks (Fig. 7b).

Discussion

Eco-metabolomic analysis of *in situ* samples reveals the existence of ecotypes

When exposed to different environments (E), spatially separated populations of a plant species can adapt and, according to the gene pool (G; genotype) present in the population, can form a variety of locally adapted chemotypes often exhibiting a different phenotype (P) according to the equation $G \times E = P$ (Sultan, 1987; Tack *et al.*, 2012; E. Allevato *et al.*, 2019). In addition, environmentally induced epigenetic modifications of chromatin structure influence gene expression, plant phenotype and contribute to plant adaptation (Rasmann *et al.*, 2012; Schmid *et al.*, 2018; Thiebaut *et al.*, 2019).

In our work, individual sites strongly correlated with several metabolite-related mass features that could be distinguished from

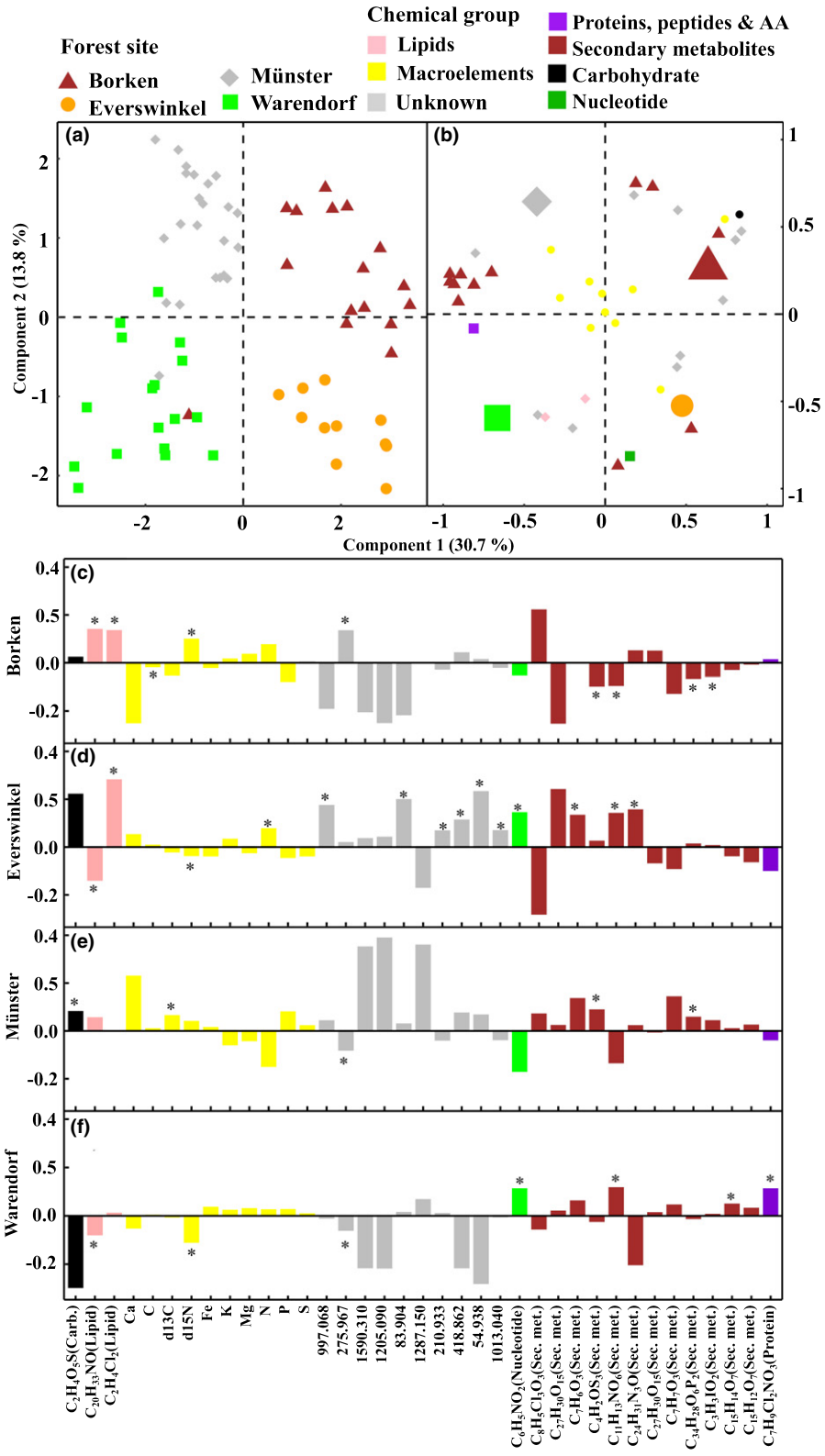

Fig. 2 Ecometabolomics and orthogonal partial least square regression discriminant analysis (OPLS-DA) of leaf oaks collected from four geographically separated forest sites (BEMW) reveal the presence of four ecotypes and show their correlations to the main chemical groups. (a) Scores, (b) X- (small circles) and Y- (large symbols) loadings of OPLS-DA computed from the metabolic composition of 27 site-specific metabolites (that significantly change for at least four comparisons (adj. *P*-value < 0.05)). The first two components describe the data, with 44.5% of explained variance. (c–f) Scaled (to 1) and centered correlation coefficient plots of OPLS-DA, showing the relationship between the *X*- and the *Y*-variables for the predictive components. OPLS model fitness: R^2X(cum) = 0.655, R^2Y(cum) = 0.714, Q^2Y (cum) = 0.625 using three predictive components. RMSEE/RMSEcv (location): 0.24/0.27 (Borken), 0.20/0.22 (Everswinkel), 0.28/0.31 (Münster) and 0.23/0.25 (Warendorf). CV-ANOVA, $P = 9.15 \times 10^{-28}$. The mass feature significant for the OPLS-DA model is noted with an asterisk (*). Data: *X*-variables, discriminant metabolite-related mass features, macroelements, and stable isotopic signature of nitrogen (N) and carbon (C). *Y*-variables: forest sites: Borken, Everswinkel, Münster, Warendorf.

their chemical profiles. This agrees with the concept of complex metabolomic networks driven by adaptation to local environments (Andrew *et al.*, 2010; Nagler *et al.*, 2018; D. M. Allevato *et al.*, 2019; Salomé-Abarca *et al.*, 2020), although it may also arise from long-distance gene flow and random genetic drift (Kremer *et al.*, 2012).

Chemotype variations at the sites are due to the interaction between the environment (including herbivory pressure) and the

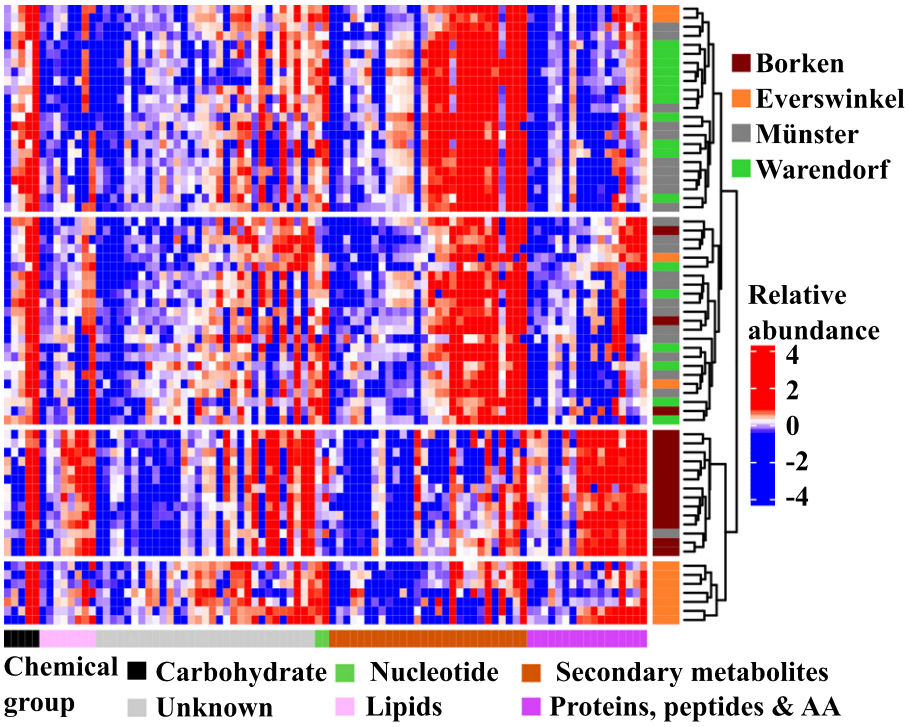

Fig. 3 Clustered heatmap of site-specific mass features of 67 oak trees (BEMW). Data depicts the different metabolic patterns for each forest site (Borken, Everswinkel, Münster, Warendorf, all in North Rhine-Westphalia, Germany) and reveals the four forest ecotypes. All mass features were discriminant for ecotypes separation in orthogonal partial least square regression discriminant analysis (OPLS-DA) (VIP > 2; CV-ANOVA < 0.05; individual mass feature adj. *P*-values < 0.05). Values are relative abundances, logarithmically transformed, and Pareto scaled with centering (van den Berg *et al*., 2006; Eriksson *et al*., 2013). Metabolites with higher values have more importance for the sample.

genetic pool (Mitchell-Olds & Schmitt, 2006; Andrew *et al*., 2010). Natural regeneration via seeds at each site suggests that individuals may be closely related, possibly explaining similar chemical patterns, an assumption that requires future genomic analyses. Genetic differences are dependent on ecotype distances, as showed in single nucleotide polymorphism (SNP) genotyping studies of three *Arabidopsis thaliana* populations in Austria compared to outside European populations (Nagler *et al*., 2018). Although we observed the most different geolocation-correlated chemo-ecotype in the forest of Borken compared to the other forest sites, further genetic studies are necessary to prove the genetic dependence on ecotype. By analyzing the macroelements and stable isotopes in the leaves, we observed shifts of N and C content and isotope signatures. These results suggest the influence of local environmental conditions and soil nutrient supply on ecotype formation. Variation in foliar $\delta^{15}N$ reflects the variation in soil N availability and is affected by local differences in N fixation and uptake (Craine *et al*., 2015). Lower N contents found in oak leaves in Münster and lower $\delta^{15}N$ in Warendorf samples may reflect a somehow lower N availability in the soils. Nevertheless, it must be noted that higher $\delta^{13}C$ in Münster and the trend of higher $\delta^{13}C$ in Warendorf might be related to their higher proportion of S-oaks, compared to the two other forest sites. Leaves of S-oaks were more consumed by herbivory, resulting in a relatively higher proportion of leaf ribs than those of T-oaks, and leaf ribs are usually characterized by higher $\delta^{13}C$ signature compared to intercostal tissue (Schleser, 1990). However, compared to the relative metabolites' abundances, correlations between macroelements and ecotypes were much less remarkable. Our analysis showed that the four ecotypes could be described by the relative abundance of 100 metabolites of the primary and secondary metabolisms, differently expressed in European oaks at the four forest stands. It should be noted that the top 10 most abundant metabolites (per chemical class) detected using our analytical approach in oak leaf extracts, such as the precursors of hydrolyzable tannins (HTs) are neither discriminant for forest site nor oak phenotype. Together with comparable elemental compositions, this is an indication that general nutrient availability in the forest was similar among the four ecotypes. One reason of such typical metabolic patterns might be the plant chemical diversity, and the outbreak of the insect *T. viridana* is the driver of the susceptible vs resistant oak phenotypes.

Changes in concentration and composition of flavonoids in leaves may reflect genetic drifts or local adaptation to environmental factors, such as spectral composition and intensity of the solar radiation (Ryan *et al*., 2001; Azuma *et al*., 2012), temperature (Goh *et al*., 2016), and biotic stress that generally triggers the formation of polyphenolic compounds (Treutter, 2005; Miranda *et al*., 2007; Koskimäki *et al*., 2009). The chemical diversity of secondary compounds is as diverse as their biophysical and biological functions in plants. They range from photoprotective pigments such anthocyanins and other flavonoids (Steyn *et al*., 2002; Koes *et al*., 2005; Agati *et al*., 2013), to defense substances such as proanthocyanidins and other tannins (Jaakola & Hohtola, 2010; Marsh *et al*., 2020), to quenchers of reactive oxygen species (Winkel-Shirley, 2002; Bailey-Serres & Mittler, 2006; Agati & Tattini, 2010). The wide functional range is also reflected in the complex transcriptional regulation of their biosynthesis. Our analysis at the one-time point in June 2019 represents an aggregated snapshot where the secondary metabolites from the group of flavonoids and flavonoid-like molecules play an important role in describing the metabolic differences of

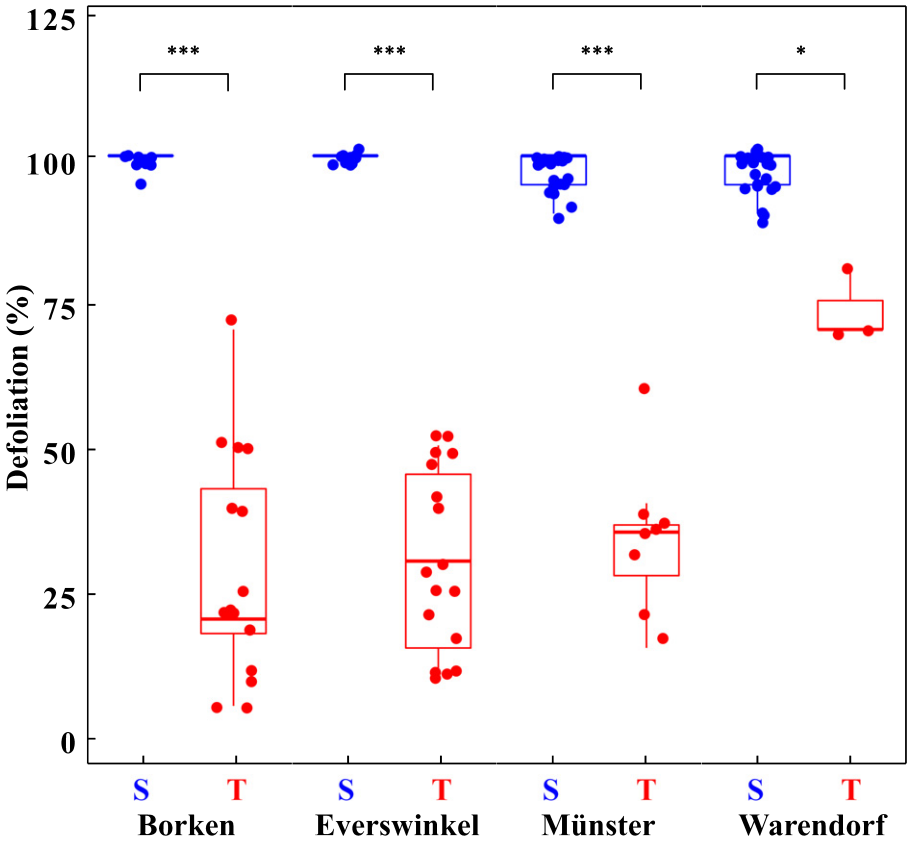

Fig. 4 Box plots of defoliation levels (percent of tree canopy defoliation) during the outbreak years of *Tortrix viridana* for four forest sites (BEMW) define the existence of two oak phenotypes naturally occurring in nature independently of tree origin and location. The trees with defoliation levels above 90% are defined as S-oaks, and trees with defoliation levels equal to or below 60% are defined as T-oaks (exception Warendorf, see main text). Lines in boxes indicate the median, the bottom and top of each box denotes the first and third quartile, respectively, and whiskers denote the 5th and 95th percentiles. Significant differences are noted as: ***, adj. *P*-values < 0.001; *, adj. *P*-values < 0.05.

the four forest sites, suggesting different plant responses to site-specific environmental influences. This can be an explanation. Additionally, long-distance gene flow and random genetic drift in the local European oak populations likely contributed to the observed metabolite patterns (Kremer *et al.*, 2012). To which extend will be the subject of future genomic studies.

Trade-offs of primary and secondary metabolisms in the resistance of *Quercus robur* to *Tortrix viridana* and potential ecological implications

More than 10 yr ago, we started to study the mechanisms of the susceptibility/tolerance of European oaks under controlled conditions by using grafted plants originating from one single forest (Ghirardo *et al.*, 2012). Compared to S-oaks, T-oak leaves showed upon herbivory feeding different induced responses of volatile organic compounds (VOCs). The major differences were related to the terpene emission profiles (Ghirardo *et al.*, 2012). VOCs are key chemical cues for oviposition site choice during the host–plant recognition between *T. viridana* and *Q. robur*. The enriched sesquiterpene content in the VOC blend of T-oaks vs more monoterpene content of S-oaks could be used by adult females of *T. viridana* as chemical cues to locate S-oaks and avoid the higher phenolic content of T-oaks (Ghirardo *et al.*, 2012). In this way, the insects benefit from the better quality of the leaves of S-oaks, as the larvae forced to feed on leaves from T-oaks need to consume more leaf material than when feeding on leaves of S-oaks to reach the same pupal weight (Ghirardo *et al.*, 2012). Given the long lifespan and slow reproduction rates of European oaks compared to those of *T. viridana*, it is likely that some members of a tree community invest a conspicuous amount of resources towards defense against fast evolving herbivory to improve community fitness.

Here, the detected differences in oak leaves' nonvolatile metabolomes across the four forest stands enabled us to disentangle the metabolic fingerprint of resistant oaks. Positively correlated to the resistant phenotype were several metabolites of the secondary metabolism. Some major flavonoids (e.g. kaempferol, kaempferol and quercetin glucosides) – known for their roles in plant resistance (Treutter, 2005) – were higher in T-oaks. Feeding assays previously demonstrated the consequences of the increased defensive molecules of T-oaks' leaf metabolome on herbivory deterrence. *Tortrix viridana* larvae need to consume more leaf biomass from T-oaks to achieve the same larval weight as comparable larvae that fed on S-leaves only (Ghirardo *et al.*, 2012). We do not know to date whether the larger amounts of flavonoids or some other unidentified secondary metabolites were responsible for the limited herbivory fitness on T-oak compared to S-oak dietaries. For example, the antifeedants polyphenolic HTs (gallotannins and ellagitannins) and CTs might also be involved (Feeny, 1970; Anstett *et al.*, 2019). Although HTs and CTs are abundant in leaves of various *Quercus* species (Salminen *et al.*, 2004), not all end-products of HT and CT metabolism were identified or detected in our nontargeted approach. Detection of some HTs and CTs is complicated by their larger molecular weight, as large polymers are not easily separated by high-

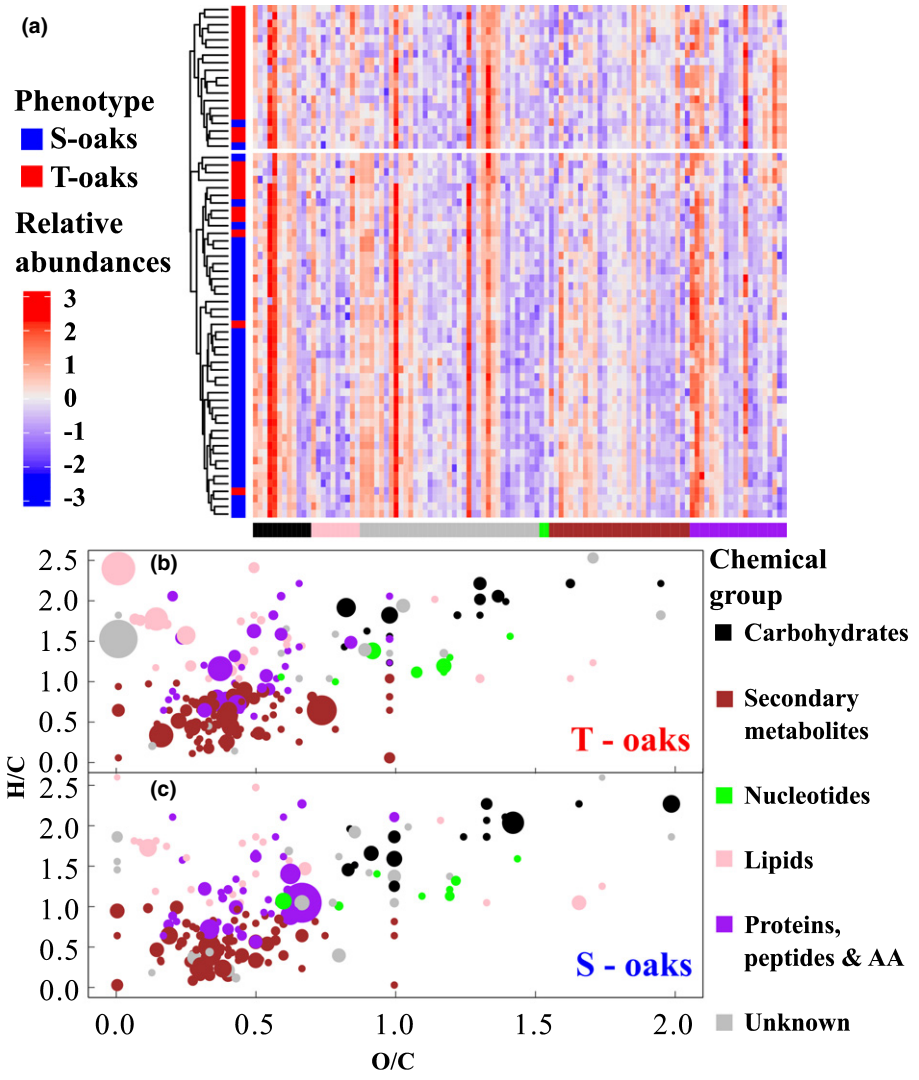

Fig. 5 (a) Clustered heatmap of phenotype-specific mass features. Data depicts the different metabolic patterns of resistant (T-) and susceptible (S-) oak phenotypes (two major clusters) independently of the ecotypes. (b, c) According to assigned chemical formulas, the Van Krevelen diagram combined with multidimensional stoichiometric compound classification (MSCC) classifies all nonannotated mass features. All mass features in (a–c) were discriminant for phenotype separation in orthogonal partial least square regression discriminant analysis (OPLS-DA) (VIP > 1.0; CV-ANOVA < 0.05; individual mass feature *P*-values < 0.05; $\log_2$ (T/S) ratio of < −0.5 or > 0.5). Values are relative abundances, logarithmically transformed, and Pareto scaled with centering (van den Berg *et al*., 2006; Eriksson *et al*., 2013).

performance liquid chromatography (HPLC) (Schofield *et al*., 2001) and the measured mass range is limited (in our analysis, *m/z* 20–2000). It is also likely that HTs and CTs were not fully extracted from leaf materials as they require specific extraction procedures and the use of enzymatic steps or acid-catalyzed hydrolysis to breakdown the polymers into analyzable monomers (Mueller-Harvey, 2001; Schofield *et al*., 2001). Indeed, among the most prominent detected metabolites, we mostly found low-molecular-weight compounds associated with the metabolism of HTs and CTs, such as the gallotannin precursor glucogallin, the components of the CTs catechin and epicatechin, and the ellagitannin derivative ellagic acid. Further targeted analyses are therefore needed to study HTs and CTs in the sensitive and resistant oaks.

Another option for herbivory resistance might be based on other defensive mechanisms, e.g. protein inhibitors (Fürstenberg-Hägg *et al*., 2013). However, the crucial molecular mechanisms that enable resistant oaks to withstand increasing herbivory pressure remain to be fully elucidated and require the incorporation of further biochemical and genetic methods.

Plants tend to adopt different defense strategies against abiotic and biotic stresses and to balance between the costs of investing in defense compound production (secondary metabolites) or to growth that allows plant competition (Herms & Mattson, 1992; Lerdau & Coley, 2002). Our field survey of 67 resistant/susceptible European oaks across four different forests showed that, regardless of geolocation-dependent differences in the oak metabolome pattern, there is a general pattern of constitutively upregulated defense metabolism in T-oak leaves vs constitutively upregulated growth-related metabolism in S-oak leaves. This metabolic shift is in line with the GDB theory (Herms & Mattson, 1992). Independent of nutrient availability, it is remarkable to note that carbohydrates were overrepresented in S-oaks (Fig. 6). NSCs are generally important in growth/differentiation (Aspinwall *et al*., 2011; Woodruff & Meinzer, 2011) and play a role in osmoregulation, supporting physiological functioning in contrasting severe drought episodes (Sevanto *et al*., 2014; Hartmann, 2015). Because the frequency and intensity of drought episodes are projected to increase in the future (Spinoni *et al*., 2018; Hari *et al*., 2020), S-oaks' strategy to invest in NSCs

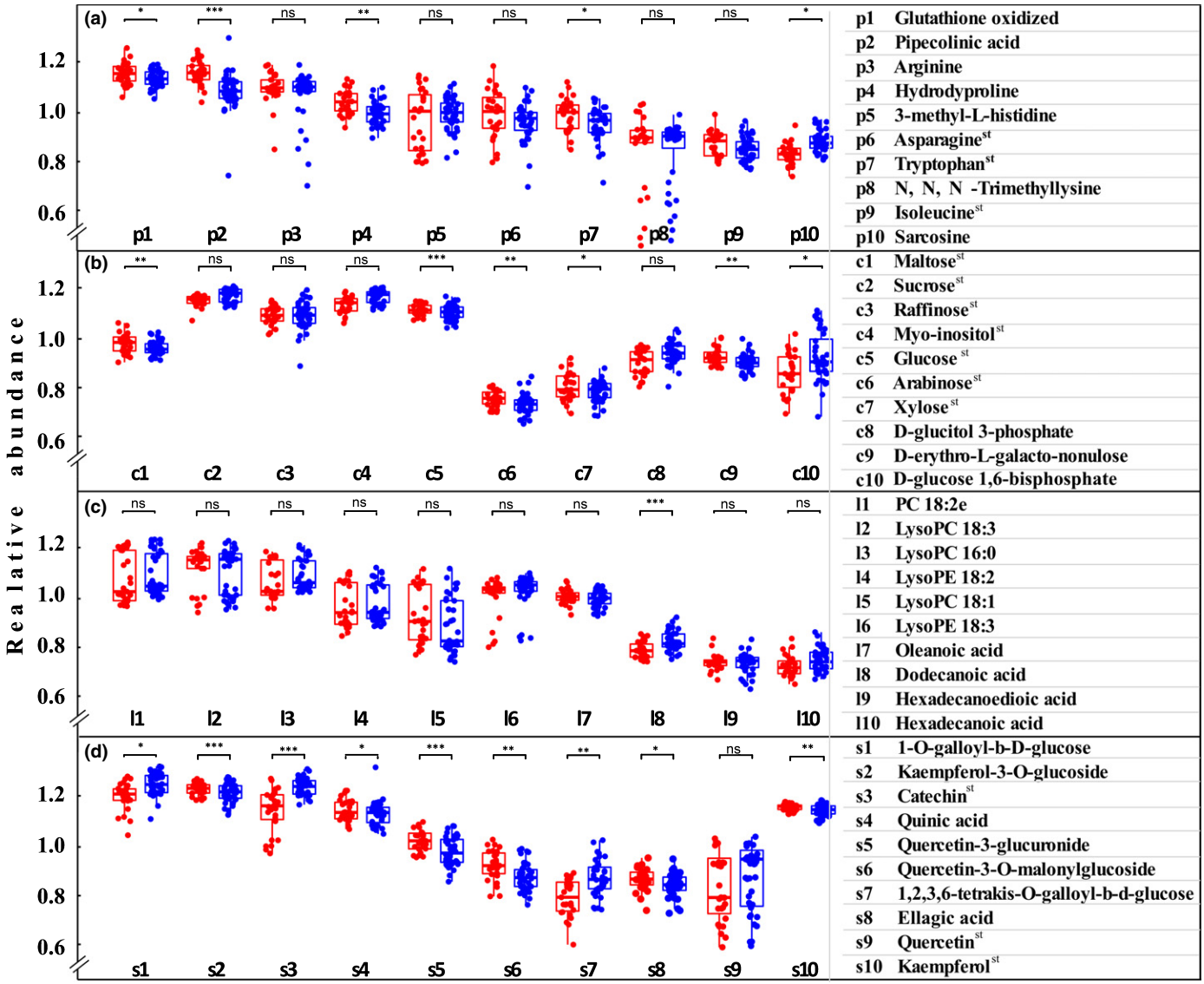

Fig. 6 Comparison of the 10 most abundant detected metabolites in resistant (T-) and susceptible (S-) oak phenotypes for four major chemical classes: (a) proteins, peptides and amino acid (AA); (b) carbohydrates; (c) lipids; (d) secondary metabolites. Metabolites marked with [st] are identified with pure standards. Metabolites levels are shown as logarithmically transformed chromatographic areas. Significant differences are noted as: ***, adj. *P*-values < 0.001; **, adj. *P*-values < 0.01; *, adj. *P*-values < 0.05; ns, not significant.

might be beneficial under the effects of climate change. Consistent with the GDB theory, the strong correlation of NSCs and amino acid derivatives to S-oaks was in concert with the negative correlation of detected defense molecules, suggesting trade-offs of resource allocations towards growth or defense in plants. Despite being the most mature of the plant defense hypotheses in evolutionary ecology, direct testing of the GDB is difficult (Stamp, 2004). Fig. 8 summarizes the metabolic differences of two oak phenotypes that support the GDB theory based on field data. Remarkably, the overrepresentation of constitutive carbon-based defense compounds (polyphenols) and the underrepresentation of NSCs in T-oak leaves (Fig. 8) is not correlated to general increases of leaf nutrient and resource availability, as it would be predicted by the carbon-nutrient balance hypothesis (Bryant *et al.*, 1983; Lerdau & Coley, 2002; Agrawal, 2020).

Despite the potential advantages of S- and T-oaks for the consequences of regional climate change, the co-occurrence of both susceptible and resistant phenotypes in all different forest ecotypes and the presence of different chemotypes within the same forest stand is remarkable. These mixtures of trees suggest that natural European oak communities are formed by oaks that constitutively follow the strategies of 'growth' or 'defense'. Our current data support the idea that the European oak population embraces a strategy to preserve both phenotypes in the same forest stand. Because species' survival may be assured if the community has a high intraspecific genetic variation (Norberg *et al.*, 2001), prioritizing the allocation of resources to either growth or defense in members of a plant community, when inherited and shared within a population, seems a successful strategy to improve the overall community fitness. Such a 'community'

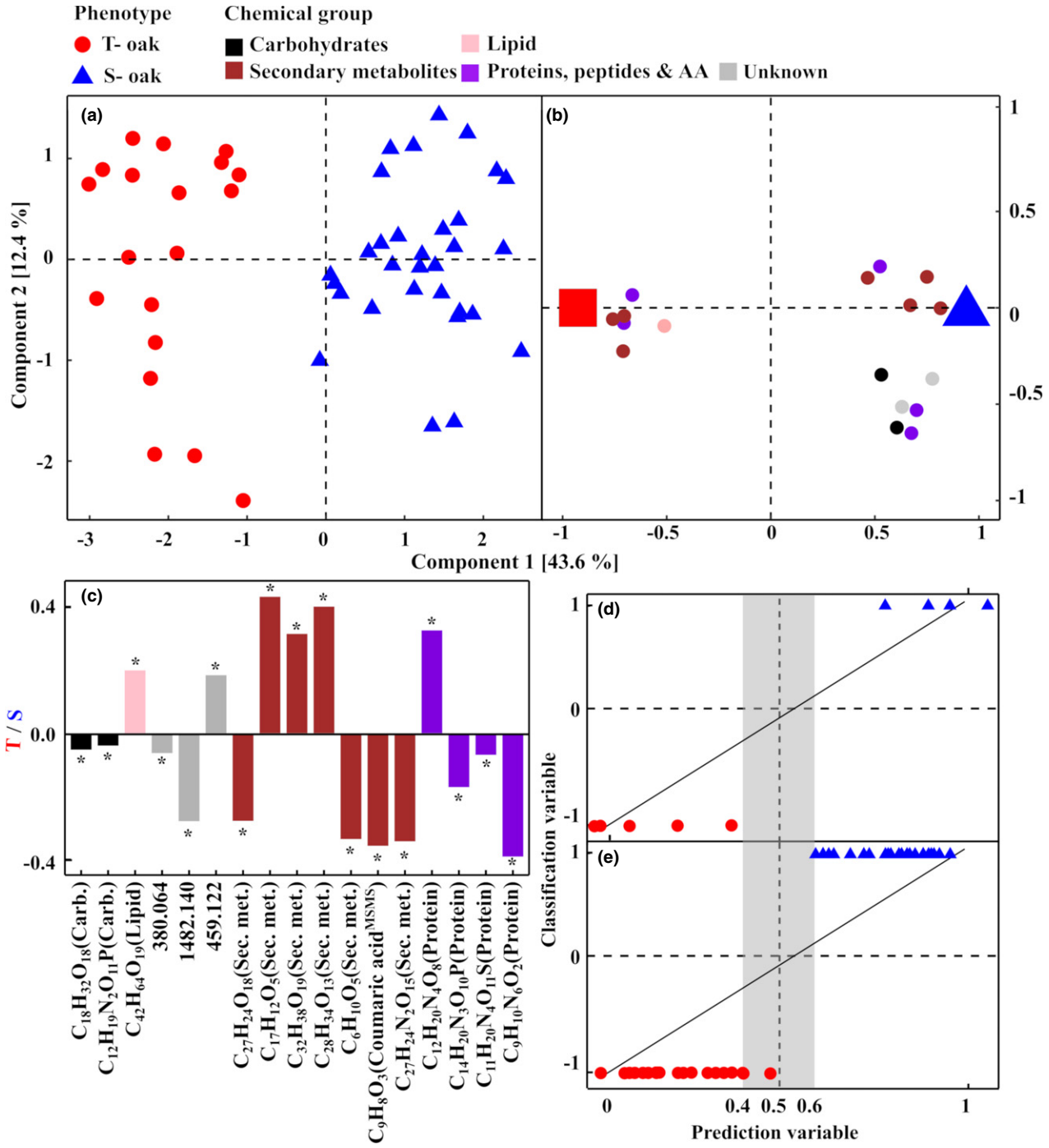

Fig. 7 Orthogonal partial least square regression discriminant analysis (OPLS-DA) of discriminant mass features for resistant (T-) and susceptible (S-) oak phenotypes. (a) Scores, (b) *X*- and *Y*-loadings of OPLS-DA computed from the metabolic composition of 17 phenotypic biomarker candidates shows (a) the separation of two oak phenotypes, and (b) different relation of main chemical groups to a specific phenotype. The data is described by the first two components with 56.0% of explained variance. (c) Scaled and centered correlation coefficient plot of 17 biomarker candidates shows that T-oaks are positively correlated to secondary compounds and negatively correlated to primary metabolism compounds (carbohydrates and proteins). OPLS model fitness: $R^2X(\text{cum}) = 0.560$, $R^2Y(\text{cum}) = 0.858$, $Q^2Y(\text{cum}) = 0.810$ using two predictive components. RMSEE (root mean square error of estimation) = 0.19 and RMSEcv (root mean square error of cross-validation) of 0.21; CV-ANOVA, $P = 1.08 \times 10^{-15}$. OPLS-DA model showing the linear relationship between observed *Y*-variables (known phenotypes) and model-predicted *Y*-variables for the (d) internal (BEMW data set) and (e) external (Asbesk data set) validation analysis. (d) Predictions of the oak phenotype for a (d) randomly selected subset of nine oaks from BEMW (accuracy of 100% for both S- and T-oaks) and (e) model-independent subset of 50 oaks from a different oak population (Asbeck) (accuracy of 100% for both S-oaks and T- oak). Oaks with $Y_{pred} < 0.4$ and > 0.6 are classified by the model as T-oaks, and S-oaks, respectively. The gray zone (Y_{pred} values between 0.4–0.6) denotes the unclassified region (see Material and Methods section). $Y_{var} = 1$, S-oak; $Y_{var} = 0$, T-oak. The mass feature significant for the OPLS-DA model is noted with an asterisk (*).

Table 2 Orthogonal partial least square regression discriminant analysis (OPLS-DA) model fitness.

OPLS-DA prediction model parameters	
PCs	2
R^2X (cum)	0.560
R^2Y(cum)	0.858
Q^2Y(cum)	0.810
CV – ANOVA	
P-value	1.08×10^{-15}
RMSEE	0.19
RMSEcv	0.21

Misclassification table			
	S-oaks	T-oaks	Correctly classified
S-oaks	20	0	100%
T-oaks	0	30	100%
Fisher's probability	2.1×10^{-14}		

PCs, number of predictive components; R^2X(cum), R^2Y(cum), cumulative regression sum of squares; Q^2Y(cum), prediction sum of squares; RMSEE, root mean square error of estimation; RMSEcv, root mean square error of cross-validation; S-oak; susceptible (S-) oak; T-oak, resistant (T-) oak.

strategy would allow plant species adaptation and ensure long-term survival. Future genomic studies will be essential to assess the composition of susceptible and resistant phenotypes in oak forests.

Biomarker development to assist tree nurseries with early oak seedlings phenotype selection for building stronger forests

To counteract increasing herbivory pressure, phenotype selection has been employed in tree breeding nurseries with great success (Naidoo *et al.*, 2019). Forest restoration aims to increase the number of resistant trees to the point where the population will be self-sustain with preserved genetic diversity (Sniezko & Koch, 2017). We aimed to create a reliable and robust model for the phenotype selection of young herbivory-resistant trees in tree nurseries. Remarkably, the analysis of those few defensive-related compounds in European oak leaves is fully sufficient to differentiate the T-oaks from the S-oaks. When assessing OPLS-DA prediction model parameters, Q^2Y and R^2Y values together with CV-ANOVA *P*-value are the best indicators of model performances (Eriksson *et al.*, 2008; Westerhuis *et al.*, 2008). The model could predict an independent external test set of a different European oak population growing under completely different environmental conditions (see details in Methods section). Also, validation samples were flash-frozen under light, compared to dark-adapted leaves in the field, indicating that the chosen biomarkers are not quickly degraded neither sensitive to quick change of light conditions. The here developed prediction model may be implemented to help tree nurseries to select T-oaks among young trees.

European oak is widely distributed all across Europe (Eaton *et al.*, 2016), and it is the second most common deciduous tree

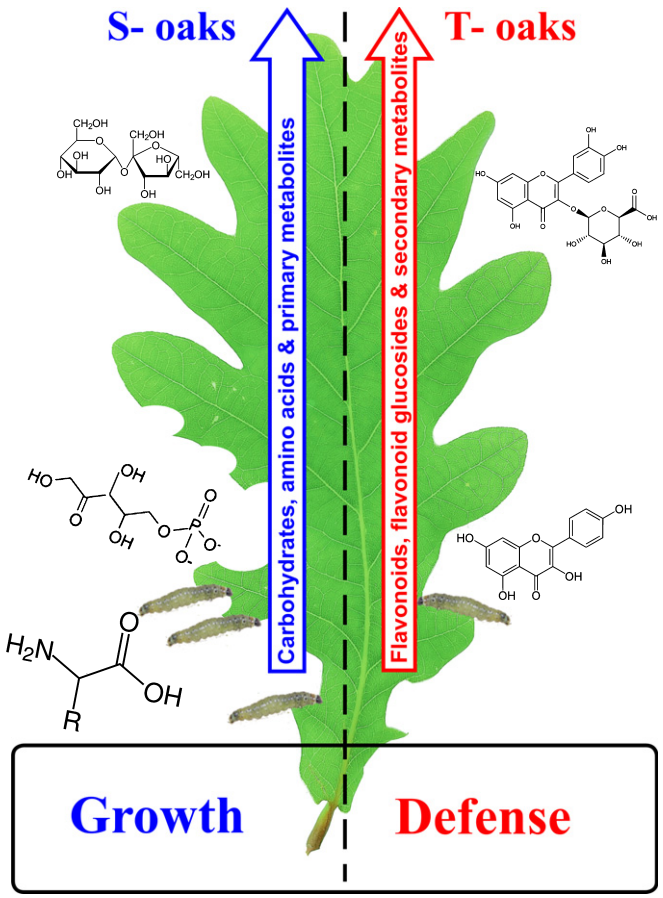

Fig. 8 Schematic overview of the leaf metabolome characteristic of resistant (T-) and susceptible (S-) oaks in combination with the proposed plant strategy for the allocation of available resources. The S-oak leaves are enriched in nutritive compounds of the primary metabolism (carbohydrates, amino acids) that sustain growth (left, in blue). The T-oak leaves are enriched in secondary compounds that help plants in chemical defense (right, in red). *Tortrix viridana* prefers feeding on S-oaks, causing higher defoliation rates (symbolized by the number of larvae and leaf damage). The growth-defense trade-offs of leaf metabolism in S- and T-oaks allowed the determination of herbivory-resistant oak phenotype using orthogonal partial least square regression discriminant analysis (OPLS-DA) and metabolomics.

species in Germany after beech (Polley *et al.*, 2014). It has the highest biodiversity of the native tree species on all trophic levels (Yela & Lawton, 1997). The European oak situation in Germany has been critical for decades, caused by clearcutting, pathogen infestation, and insect outbreaks (Führer, 1998) without noticeable improvement (Hartmann & Blank, 1992). Another substantial threat is the predicted insect range expansion and invasion of new insect pest species, with unknown consequences (Sturrock *et al.*, 2011; Pureswaran *et al.*, 2018). In addition, impending climate change in global warming and prolonged drought episodes may act as a primary stressor for European oak. It has been noted that deciduous oak forests were the least persistent to the projected global change (Merlin *et al.*, 2015; Acácio *et al.*, 2017; Madrigal-González *et al.*, 2017). This can lead to indirect consequences such as new herbivorous insects that will most likely migrate from Southern and Eastern Europe and become potential

feeding pests for European oak stands in Germany (Delb, 2012). Concerns are increasing that the genetic adaptation of long-living species to rapid ongoing global change may not be quick enough (Burrows *et al.*, 2011; Dawson *et al.*, 2011; Hoffmann & Sgrò, 2011; Duputié *et al.*, 2015) as the genetic adaptations of long-lived species are slow (Savolainen *et al.*, 2004). Therefore, European oak as a long-living tree species deploys more phenotypic plasticity than through microevolution as a response to the rapid environmental changes (Chevin *et al.*, 2013; Franks *et al.*, 2014).

Moreover, the adaptation through phenotypic plasticity is not fully efficient, requiring further evolutionary changes to avoid maladaptation (Gienapp *et al.*, 2013). One way to tackle forest survival under global change is through forest management (Noss, 2001; Bolte *et al.*, 2009; D'Amato *et al.*, 2011), and the selection of appropriate herbivory-resistant phenotypes might be crucial to improve forest fitness under severe herbivory outbreaks in Europe (Saxe *et al.*, 2001). Using the cost-effective, highly reliable metabolic-based phenotypic biomarkers developed herein, it might be possible to identify oak phenotypes during early seedling and plant cultivations. In particular, this strategy may support nurseries in the early selection of an appropriate proportion of two oak phenotypes from different proveniences that can be used in managing the forest composition during afforestation activities of threatened forests under climate-driven, enhanced herbivory pressure. Additionally, the further development of biomarkers, including robust diagnostic genetic markers combined with an extensive screening of oak forest, will pave the way to study Central European forest's susceptibility to future insect outbreaks and support forest management. Thereby, it will contribute to the long-term conservation of European oaks, an ecologically important tree species grown in temperate latitudes of Europe.

Acknowledgements

This project is financed by the Fachagentur Nachwachsende Rohstoffe eV (FNR) in the program 'Waldklimafonds' (28W-B-4-113-01/2) funded by the Federal Ministry of Food and Agriculture and Federal Ministry for Environment, Nature Conservation and Nuclear Safety, Germany. The sample set from North Rhine-Westphalia was obtained with the help from technical assistants and gardeners from the Thünen-Institute of Forest Genetics: Susanne Jelkmann, Matthias Wellern, Dieter Boedecker, and Maike Voss. The authors are thankful to Ina Zimmer (EUS) for her excellent work with sample preparation and Bernhard Michalke's platform for the elemental analysis service.

Author contributions

AG, HS, BK, MF and J-PS designed the research. MB conducted metabolomics and all data analyses. HS, FO and BK collected and processed field survey data. FB performed stable isotope analyses. AG, MB and J-PS wrote the manuscript. All authors read, corrected, and approved the manuscript.

ORCID

Marko Bertić https://orcid.org/0000-0001-9016-0902
Matthias Fladung https://orcid.org/0000-0001-9301-8581
Andrea Ghirardo https://orcid.org/0000-0003-1973-4007
Birgit Kersten https://orcid.org/0000-0001-9900-9133
Jörg-Peter Schnitzler https://orcid.org/0000-0002-9825-867X
Hilke Schroeder https://orcid.org/0000-0002-0908-3397

Data availability

All UPLC-UHR-QqToF-MS raw data have been deposited in the open science framework data repository and can be accessed following the link https://osf.io/s9d2j/?view_only=733a0c1a9e444f669d44c6eaad44f253.

References

Acácio V, Dias FS, Catry FX, Rocha M, Moreira F. 2017. Landscape dynamics in Mediterranean oak forests under global change: understanding the role of anthropogenic and environmental drivers across forest types. *Global Change Biology* **23**: 1199–1217.

Agati G, Brunetti C, Di Ferdinando M, Ferrini F, Pollastri S, Tattini M. 2013. Functional roles of flavonoids in photoprotection: new evidence, lessons from the past. *Plant Physiology and Biochemistry* **72**: 35–45.

Agati G, Tattini M. 2010. Multiple functional roles of flavonoids in photoprotection. *New Phytologist* **186**: 786–793.

Agrawal AA. 2020. A scale-dependent framework for trade-offs, syndromes, and specialization in organismal biology. *Ecology* **101**: e02924.

Alberto FJ, Aitken SN, Alía R, González-Martínez SC, Hänninen H, Kremer A, Lefèvre F, Lenormand T, Yeaman S, Whetten R *et al.* 2013. Potential for evolutionary responses to climate change – evidence from tree populations. *Global Change Biology* **19**: 1645–1661.

Allevato DM, Kiyota E, Mazzafera P, Nixon KC. 2019. Ecometabolomic analysis of wild populations of *Pilocarpus pennatifolius* (Rutaceae) using unimodal analyses. *Frontiers in Plant Science* **10**: 258.

Allevato E, Saulino L, Cesarano G, Chirico GB, D'Urso G, Falanga Bolognesi S, Rita A, Rossi S, Saracino A, Bonanomi G. 2019. Canopy damage by spring frost in European beech along the Apennines: effect of latitude, altitude and aspect. *Remote Sensing of Environment* **225**: 431–440.

Anderssen E, Dyrstad K, Westad F, Martens H. 2006. Reducing over-optimism in variable selection by cross-model validation. *Chemometrics and Intelligent Laboratory Systems* **84**: 69–74.

Andrew RL, Wallis IR, Harwood CE, Foley WJ. 2010. Genetic and environmental contributions to variation and population divergence in a broad-spectrum foliar defence of *Eucalyptus tricarpa*. *Annals of Botany* **105**: 707–717.

Annighöfer P, Beckschäfer P, Vor T, Ammer C. 2015. Regeneration patterns of European oak species (*Quercus petraea* (Matt.) Liebl., *Quercus robur* L.) in dependence of environment and neighborhood. *PLoS ONE* **10**: e0134935.

Anstett DN, Cheval I, D'Souza C, Salminen JP, Johnson MTJ. 2019. Ellagitannins from the Onagraceae decrease the performance of generalist and specialist herbivores. *Journal of Chemical Ecology* **45**: 86–94.

Aspinwall MJ, King JS, Booker FL, McKeand SE. 2011. Genetic effects on total phenolics, condensed tannins and non-structural carbohydrates in loblolly pine (*Pinus taeda* L.) needles. *Tree Physiology* **31**: 831–842.

Azuma A, Yakushiji H, Koshita Y, Kobayashi S. 2012. Flavonoid biosynthesis-related genes in grape skin are differentially regulated by temperature and light conditions. *Planta* **236**: 1067–1080.

Bacon SJ, Aebi A, Calanca P, Bacher S. 2014. Quarantine arthropod invasions in Europe: the role of climate, hosts and propagule pressure. *Diversity and Distributions* **20**: 84–94.

Bailey-Serres J, Mittler R. 2006. The roles of reactive oxygen species in plant cells. *Plant Physiology* **141**: 311.

Bebber DP, Ramotowski MAT, Gurr SJ. 2013. Crop pests and pathogens move polewards in a warming world. *Nature Climate Change* **3**: 985–988.

Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)* **57**: 289–300.

van den Berg RA, Hoefsloot HCJ, Westerhuis JA, Smilde AK, van der Werf MJ. 2006. Centering, scaling, and transformations: improving the biological information content of metabolomics data. *BMC Genomics* **7**: 142.

Boccard J, Rutledge DN. 2013. A consensus orthogonal partial least squares discriminant analysis (OPLS-DA) strategy for multiblock omics data fusion. *Analytica Chimica Acta* **769**: 30–39.

Bolte A, Ammer C, Löf M, Madsen P, Nabuurs G-J, Schall P, Spathelf P, Rock J. 2009. Adaptive forest management in central Europe: climate change impacts, strategies and integrative concept. *Scandinavian Journal of Forest Research* **24**: 473–482.

Bradshaw AD. 1984. Ecological significance of genetic variation between populations. In: Dirzo R, Sarukhan J, eds. *Perspectives on plant population ecology*. Sunderland, MA, USA: Sinauer Associates, 213–228.

Bresson CC, Vitasse Y, Kremer A, Delzon S. 2011. To what extent is altitudinal variation of functional traits driven by genetic adaptation in European oak and beech? *Tree Physiology* **31**: 1164–1174.

Broadhurst DI, Kell DB. 2006. Statistical strategies for avoiding false discoveries in metabolomics and related experiments. *Metabolomics* **2**: 171–196.

Bryant JP, Chapin FS, Klein DR. 1983. Carbon/nutrient balance of boreal plants in relation to vertebrate herbivory. *Oikos* **40**: 357–368.

Burrows MT, Schoeman DS, Buckley LB, Moore P, Poloczanska ES, Brander KM, Brown C, Bruno JF, Duarte CM, Halpern BS *et al.* 2011. The pace of shifting climate in marine and terrestrial ecosystems. *Science* **334**: 652.

Chevin L-M, Collins S, Lefèvre F. 2013. Phenotypic plasticity and evolutionary demographic responses to climate change: taking theory out to the field. *Functional Ecology* **27**: 967–979.

Cocchi M, Biancolillo A, Marini F. 2018. Chemometric methods for classification and feature selection. In: Jaumot J, Bedia C, Tauler R, eds. *Comprehensive analytical chemistry*. Amsterdam, the Netherlands: Elsevier, 265–299.

Corwin JA, Kliebenstein DJ. 2017. Quantitative resistance: more than just perception of a pathogen. *Plant Cell* **29**: 655–665.

Craine JM, Brookshire ENJ, Cramer MD, Hasselquist NJ, Koba K, Marin-Spiotta E, Wang L. 2015. Ecological interpretations of nitrogen isotope ratios of terrestrial plants and soils. *Plant and Soil* **396**: 1–26.

D'Amato AW, Bradford JB, Fraver S, Palik BJ. 2011. Forest management for mitigation and adaptation to climate change: insights from long-term silviculture experiments. *Forest Ecology and Management* **262**: 803–816.

Dawson TP, Jackson ST, House JI, Prentice IC, Mace GM. 2011. Beyond predictions: biodiversity conservation in a changing climate. *Science* **332**: 53.

Delb H. 2012. Eichenschädlinge im Klimawandel in Südwestdeutschland. *FVA-Einblick* **2**: 11–14.

Denkert C, Budczies J, Kind T, Weichert W, Tablack P, Sehouli J, Niesporek S, Könsgen D, Dietel M, Fiehn O. 2006. Mass spectrometry-based metabolic profiling reveals different metabolite patterns in invasive ovarian carcinomas and ovarian borderline tumors. *Cancer Research* **66**: 10795.

Domingo-Almenara X, Montenegro-Burke JR, Benton HP, Siuzdak G. 2018. Annotation: a computational solution for streamlining metabolomics analysis. *Analytical Chemistry* **90**: 480–489.

Duputié A, Rutschmann A, Ronce O, Chuine I. 2015. Phenological plasticity will not help all species adapt to climate change. *Global Change Biology* **21**: 3062–3073.

Dyson BC, Allwood JW, Feil R, Xu Y, Miller M, Bowsher CG, Goodacre R, Lunn JE, Johnson GN. 2015. Acclimation of metabolism to light in *Arabidopsis thaliana*: the glucose 6-phosphate/phosphate translocator GPT2 directs metabolic acclimation. *Plant, Cell & Environment* **38**: 1404–1417.

Eaton E, Caudullo G, Oliveira S, de Rigo D. 2016. *Quercus robur* and *Quercus petraea* in Europe: distribution, habitat, usage and threats. In: San-MiguelAyanz J, de Rigo D, Caudullo G, Houston Durrant T, Mauri A, eds. *European atlas of forest tree species*. Luxembourg City, Luxembourg: The Publications Office of the European Union, e01c6df.

Ehmcke G, Grosser D. 2014. Das Holz der Eiche-Eigenschaften und Verwendung. *LWF Wissen* **75**: 53–64.

Eriksson L, Byrne T, Johansson E, Trygg J, Vikström C. 2013. *Multi-and megavariate data analysis basic principles and applications*. Malmö, Sweden: MKS Umetrics AB.

Eriksson L, Trygg J, Wold S. 2008. CV-ANOVA for significance testing of PLS and OPLS® models. *Journal of Chemometrics* **22**: 594–600.

Etterson JR. 2004. Evolutionary potential of *Chamaecrista fasciculata* in relation to climate change. I. Clinal patterns of selection along an environmental gradient in the great plains. *Evolution* **58**: 1446–1456.

Evers J. 2004. *Waldbäume: Bilderserien zur Einschätzung von Kronenverlichtungen bei waldbäumen*. Kassel, Germany: Verlag M. Faste.

Feeny P. 1970. Seasonal changes in oak leaf tannins and nutrients as a cause of spring feeding by winter moth caterpillars. *Ecology* **51**: 565–581.

Franks SJ, Weber JJ, Aitken SN. 2014. Evolutionary and plastic responses to climate change in terrestrial plant populations. *Evolutionary Applications* **7**: 123–139.

Führer E. 1998. *Oak decline in Central Europe: a synopsis of hypotheses. Proceedings of population dynamics, impacts, and integrated management of forest defoliating insects*. USDA Forest Service General Technical Report NE-247. Washington, DC, USA: USDA Forest Service, 7–24.

Fürstenberg-Hägg J, Zagrobelny M, Bak S. 2013. Plant defense against insect herbivores. *International Journal of Molecular Sciences* **14**: 10242–10297.

Galloway LF, Fenster CB. 2000. Population differentiation in an annual legume: local adaptation. *Evolution* **54**: 1173–1181.

Geber MA, Griffen LR. 2003. Inheritance and natural selection on functional traits. *International Journal of Plant Sciences* **164**: 21–42.

Ghirardo A, Fochi V, Lange B, Witting M, Schnitzler J-P, Perotto S, Balestrini R. 2020. Metabolomic adjustments in the orchid mycorrhizal fungus *Tulasnella calospora* during symbiosis with *Serapias vomeracea*. *New Phytologist* **228**: 1939–1952.

Ghirardo A, Heller W, Fladung M, Schnitzler J-P, Schroeder H. 2012. Function of defensive volatiles in pedunculate oak (*Quercus robur*) is tricked by the moth *Tortrix viridana*. *Plant, Cell & Environment* **35**: 2192–2207.

Ghirardo A, Sørensen HA, Petersen M, Jacobsen S, Søndergaard I. 2005. Early prediction of wheat quality: analysis during grain development using mass spectrometry and multivariate data analysis. *Rapid Communications in Mass Spectrometry* **19**: 525–532.

Gienapp P, Lof M, Reed TE, McNamara J, Verhulst S, Visser ME. 2013. Predicting demographically sustainable rates of adaptation: can great tit breeding time keep pace with climate change? *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* **368**: 20120289.

Goh HH, Khairudin K, Sukiran NA, Normah MN, Baharum SN. 2016. Metabolite profiling reveals temperature effects on the VOCs and flavonoids of different plant populations. *Plant Biology* **18**: 130–139.

Gu Z, Eils R, Schlesner M. 2016. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* **32**: 2847–2849.

Hamann E, Blevins C, Franks SJ, Jameel MI, Anderson JT. 2021. Climate change alters plant–herbivore interactions. *New Phytologist* **229**: 1894–1910.

Hari V, Rakovec O, Markonis Y, Hanel M, Kumar R. 2020. Increased future occurrences of the exceptional 2018–2019 Central European drought under global warming. *Scientific Reports* **10**: 12207.

Hartmann H. 2015. Carbon starvation during drought-induced tree mortality – are we chasing a myth? *Journal of Plant Hydraulics* **2**: e005.

Hartmann G, Blank R. 1992. Winter frost, insect defoliation and *Agrilus biguttatus* Fabr. as causal factors of oak decline in northern Germany. *Forst und Holz* **15**: 443–452.

Hemmler D, Heinzmann SS, Wöhr K, Schmitt-Kopplin P, Witting M. 2018. Tandem HILIC-RP liquid chromatography for increased polarity coverage in food analysis. *Electrophoresis* **39**: 1645–1653.

Herms DA, Mattson WJ. 1992. The dilemma of plants: to grow or defend. *The Quarterly Review of Biology* **67**: 283–335.

Hoffmann AA, Sgrò CM. 2011. Climate change and evolutionary adaptation. *Nature* **470**: 479–485.

Hunter MD. 1990. Differential susceptibility to variable plant phenology and its role in competition between two insect herbivores on oak. *Ecological Entomology* **15**: 401–408.

Jaakola L, Hohtola A. 2010. Effect of latitude on flavonoid biosynthesis in plants. *Plant, Cell & Environment* **33**: 1239–1247.

Kersten B, Ghirardo A, Schnitzler J-P, Kanawati B, Schmitt-Kopplin P, Fladung M, Schroeder H. 2013. Integrated transcriptomics and metabolomics decipher differences in the resistance of pedunculate oak to the herbivore *Tortrix viridana* L. *BMC Genomics* **14**: 737.

Kleessen S, Antonio C, Sulpice R, Laitinen R, Fernie AR, Stitt M, Nikoloski Z. 2012. Structured patterns in geographic variability of metabolic phenotypes in *Arabidopsis thaliana*. *Nature Communications* **3**: 1319.

Koes R, Verweij W, Quattrocchio F. 2005. Flavonoids: a colorful model for the regulation and evolution of biochemical pathways. *Trends in Plant Science* **10**: 236–242.

Kollmann J, Bañuelos MJ. 2004. Latitudinal trends in growth and phenology of the invasive alien plant impatiens *Glandulifera* (*Balsaminaceae*). *Diversity and Distributions* **10**: 377–385.

Koskimäki JJ, Hokkanen J, Jaakola L, Suorsa M, Tolonen A, Mattila S, Pirttilä AM, Hohtola A. 2009. Flavonoid biosynthesis and degradation play a role in early defence responses of bilberry (*Vaccinium myrtillus*) against biotic stress. *European Journal of Plant Pathology* **125**: 629–640.

Kremer A, Ronce O, Robledo-Arnuncio JJ, Guillaume F, Bohrer G, Nathan R, Bridle JR, Gomulkiewicz R, Klein EK, Ritland K *et al.* 2012. Long-distance gene flow and adaptation of forest trees to rapid climate change. *Ecology Letters* **15**: 378–392.

Kuzina V, Ekstrøm CT, Andersen SB, Nielsen JK, Olsen CE, Bak S. 2009. Identification of defence compounds in *Barbarea vulgaris* against the herbivore *Phyllotreta nemorum* by an ecometabolomic approach. *Plant Physiology* **151**: 1977–1990.

Lerdau M, Coley PD. 2002. Benefits of the carbon-nutrient balance hypothesis. *Oikos* **98**: 534–536.

Levin DA. 1976. The chemical defences of plants to pathogens and herbivores. *Annual Review of Ecology and Systematics* **7**: 121–159.

Linhart YB, Grant MC. 1996. Evolutionary significance of local genetic differentiation in plants. *Annual Review of Ecology and Systematics* **27**: 237–277.

Madrigal-González J, Ruiz-Benito P, Ratcliffe S, Rigling A, Wirth C, Zimmermann NE, Zweifel R, Zavala MA. 2017. Competition drives oak species distribution and functioning in Europe: Implications under global change. In: Gil-Pelegrín E, Peguero-Pina JJ, Sancho-Knapik D, eds. *Oaks physiological ecology. Exploring the functional diversity of genus* Quercus *L.* Cham, Switzerland: Springer International, 513–538.

Marsh KJ, Wallis IR, Kulheim C, Clark R, Nicolle D, Foley WJ, Salminen J-P. 2020. New approaches to tannin analysis of leaves can be used to explain in vitro biological activities associated with herbivore defence. *New Phytologist* **225**: 488–498.

McInnes L, Healy J, Melville J. 2018. Umap: uniform manifold approximation and projection for dimension reduction. *arXiv preprint arXiv:1802.03426*.

Merlin M, Perot T, Perret S, Korboulewsky N, Vallet P. 2015. Effects of stand composition and tree size on resistance and resilience to drought in sessile oak and Scots pine. *Forest Ecology and Management* **339**: 22–33.

Miranda M, Ralph SG, Mellway R, White R, Heath MC, Bohlmann J, Constabel CP. 2007. The transcriptional response of hybrid poplar (*Populus trichocarpa* x *P. deltoides*) to infection by *Melampsora medusae* leaf rust involves induction of flavonoid pathway genes leading to the accumulation of proanthocyanidins. *Molecular Plant–Microbe Interactions* **20**: 816–831.

Mitchell-Olds T, Schmitt J. 2006. Genetic mechanisms and evolutionary significance of natural variation in Arabidopsis. *Nature* **441**: 947–952.

Montalvo AM, Ellstrand NC. 2000. Transplantation of the subshrub *Lotus scoparius*: testing the home-site advantage hypothesis. *Conservation Biology* **14**: 1034–1045.

Mueller-Harvey I. 2001. Analysis of hydrolysable tannins. *Animal Feed Science and Technology* **91**: 3–20.

Müller C, Bräutigam A, Eilers E, Junker R, Schnitzler J-P, Steppuhn A, Unsicker S, van Dam N, Weisser W, Wittmann M. 2020. Ecology and evolution of intraspecific chemodiversity of plants. *Research Ideas and Outcomes* **6**: 1–31.

Müller-Starck G, Herzog S, Hattemer HH. 1993. Intra- and interpopulational genetic variation in juvenile populations of *Quercus robur* L and *Quercus petraea* Liebl. *Annals of Forest Science* **50**: 233s–244s.

Nagler M, Nägele T, Gilli C, Fragner L, Korte A, Platzer A, Farlow A, Nordborg M, Weckwerth W. 2018. Eco-metabolomics and metabolic modeling: making the leap from model systems in the lab to native populations in the field. *Frontiers in Plant Science* **9**: 1556.

Naidoo S, Slippers B, Plett JM, Coles D, Oates CN. 2019. The road to resistance in forest trees. *Frontiers in Plant Science* **10**: 273.

van Noordwijk AJ, de Jong G. 1986. Acquisition and allocation of resources: their influence on variation in life history tactics. *The American Naturalist* **128**: 137–142.

Norberg J, Swaney DP, Dushoff J, Lin J, Casagrandi R, Levin SA. 2001. Phenotypic diversity and ecosystem functioning in changing environments: a theoretical framework. *Proceedings of the National Academy of Sciences, USA* **98**: 11376.

Noss RF. 2001. Beyond Kyoto: forest management in a time of rapid climate change. *Conservation Biology* **15**: 578–590.

Peñuelas J, Sardans J. 2009. Ecological metabolomics. *Chemistry and Ecology* **25**: 305–309.

Polley H, Hennig P, Kroiher F, Marks A, Riedel T, Schmidt U, Schwitzgebel F, Stauber T. 2014. *Der Wald in Deutschland: Ausgewählte Ergebnisse der dritten Bundeswaldinventur*. Berlin, Germany: Bundesministeriums für Ernährung und Landwirtschaft.

Pretzsch H, Biber P, Schütze G, Uhl E, Rötzer T. 2014. Forest stand growth dynamics in Central Europe have accelerated since 1870. *Nature Communications* **5**: 4967.

Pureswaran DS, Roques A, Battisti A. 2018. Forest insects and climate change. *Current Forestry Reports* **4**: 35–50.

R Core Team. 2019. *A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. [WWW document] URL www.r-project.org.

Rasmann S, De Vos M, Casteel CL, Tian D, Halitschke R, Sun JY, Agrawal AA, Felton GW, Jander G. 2012. Herbivory in the previous generation primes plants for enhanced insect resistance. *Plant Physiology* **158**: 854–863.

Reisdorph NA, Walmsley S, Reisdorph R. 2020. A perspective and framework for developing sample type specific databases for LC/MS-based clinical metabolomics. *Metabolites* **10**: 8.

Riipi M, Haukioja E, Lempa K, Ossipov V, Ossipova S, Pihlaja K. 2004. Ranking of individual mountain birch trees in terms of leaf chemistry: seasonal and annual variation. *Chemoecology* **14**: 31–43.

Rivas-Ubach A, Liu Y, Bianchi TS, Tolić N, Jansson C, Paša-Tolić L. 2018. Moving beyond the van Krevelen diagram: a new stoichiometric approach for compound classification in organisms. *Analytical Chemistry* **90**: 6152–6160.

Rivas-Ubach A, Peñuelas J, Hódar JA, Oravec M, Paša-Tolić L, Urban O, Sardans J. 2019. We are what we eat: a stoichiometric and ecometabolomic study of caterpillars feeding on two pine subspecies of *Pinus sylvestris*. *International Journal of Molecular Sciences* **20**: 59.

Rivas-Ubach A, Pérez-Trujillo M, Sardans J, Gargallo-Garriga A, Parella T, Peñuelas J. 2013. Ecometabolomics: optimized NMR-based method. *Methods in Ecology and Evolution* **4**: 464–473.

Ryan KG, Swinny EE, Winefield C, Markham KR. 2001. Flavonoids and UV photoprotection in *Arabidopsis* mutants. *Zeitschrift für Naturforschung C* **56**: 745–754.

Salminen J-P, Roslin T, Karonen M, Sinkkonen J, Pihlaja K, Pulkkinen P. 2004. Seasonal variation in the content of hydrolysable tannins, flavonoid glycosides and proanthocyanidins in oak leaves. *Journal of Chemical Ecology* **30**: 1693–1711.

Salomé-Abarca LF, Mandrone M, Sanna C, Poli F, van der Hondel CAMJJ, Klinkhamer PGL, Choi YH. 2020. Metabolic variation in *Cistus monspeliensis* L. ecotypes correlated to their plant–fungal interactions. *Phytochemistry* **176**: 112402.

Sardans J, Peñuelas J, Rivas-Ubach A. 2011. Ecological metabolomics: overview of current developments and future challenges. *Chemoecology* **21**: 191–225.

Savolainen O, Bokma F, García-Gil R, Komulainen P, Repo T. 2004. Genetic variation in cessation of growth and frost hardiness and consequences for

adaptation of *Pinus sylvestris* to climatic changes. *Forest Ecology and Management* **197**: 79–89.

Sawada Y, Nakabayashi R, Yamada Y, Suzuki M, Sato M, Sakata A, Akiyama K, Sakurai T, Matsuda F, Aoki T *et al.* 2012. RIKEN tandem mass spectral database (ReSpect) for phytochemicals: a plant-specific MS/MS-based data resource and database. *Phytochemistry* **82**: 38–45.

Saxe H, Cannell MGR, Johnsen Ø, Ryan MG, Vourlitis G. 2001. Tree and forest functioning in response to global warming. *New Phytologist* **149**: 369–399.

Schemske DW, Mittelbach GG, Cornell HV, Sobel JM, Roy K. 2009. Is there a latitudinal gradient in the importance of biotic interactions? *Annual Review of Ecology, Evolution, and Systematics* **40**: 245–269.

Schleser GH. 1990. Investigations of the $\delta^{13}C$ pattern in leaves of *Fagus sylvatica* L. *Journal of Experimental Botany* **41**: 565–572.

Schmid MW, Heichinger C, Coman Schmid D, Guthörl D, Gagliardini V, Bruggmann R, Aluri S, Aquino C, Schmid B, Turnbull LA *et al.* 2018. Contribution of epigenetic variation to adaptation in Arabidopsis. *Nature Communications* **9**: 4446.

Schofield P, Mbugua DM, Pell AN. 2001. Analysis of condensed tannins: a review. *Animal Feed Science and Technology* **91**: 21–40.

Schramel P. 1983. Consideration of inductively coupled plasma spectroscopy for trace element analysis in the bio-medical and environmental fields. *Spectrochimica Acta Part B: Atomic Spectroscopy* **38**: 199–206.

Schramel P, Wolf A, Seif R, Klose BJ. 1980. Eine neue Apparatur zur Druckveraschung von biologischem Material. *Fresenius' Zeitschrift für analytische Chemie* **302**: 62–64.

Schröder H. 2010. Sommerveredelung bei Eichen–eine Erfolgsgeschichte. *AFZ-Der Wald* **5**: 16–17.

Schroeder H, Degen B. 2008. Genetic structure of the green oak leaf roller (*Tortrix viridana* L.) and one of its hosts, *Quercus robur* L. *Forest Ecology and Management* **256**: 1270–1279.

Sevanto S, McDowell NG, Dickman LT, Pangle R, Pockman WT. 2014. How do trees die? A test of the hydraulic failure and carbon starvation hypotheses. *Plant, Cell & Environment* **37**: 153–161.

Sniezko RA, Koch J. 2017. Breeding trees resistant to insects and diseases: putting theory into application. *Biological Invasions* **19**: 3377–3400.

Spinoni J, Vogt JV, Naumann G, Barbosa P, Dosio A. 2018. Will drought events become more frequent and severe in Europe? *International Journal of Climatology* **38**: 1718–1736.

Stamp N. 2004. Can the growth–differentiation balance hypothesis be tested rigorously? *Oikos* **107**: 439–448.

Steyn WJ, Wand SJE, Holcroft DM, Jacobs G. 2002. Anthocyanins in vegetative tissues: a proposed unified function in photoprotection. *New Phytologist* **155**: 349–361.

Sturrock RN, Frankel SJ, Brown AV, Hennon PE, Kliejunas JT, Lewis KJ, Worrall JJ, Woods AJ 2011. Climate change and forest diseases. *Plant Pathology* **60**: 133–149.

Sultan SE. 1987. Evolutionary implications of phenotypic plasticity in plants. In: Hecht MK, Wallace B, Prance GT, eds. *Evolutionary biology, vol. 21*. Boston, MA, USA: Springer, 127–178.

Sumner LW, Amberg A, Barrett D, Beale MH, Beger R, Daykin CA, Fan T-M, Fiehn O, Goodacre R, Griffin JL *et al.* 2007. Proposed minimum reporting standards for chemical analysis. *Metabolomics* **3**: 211–221.

Sysi-Aho M, Katajamaa M, Yetukuri L, Orešič M. 2007. Normalization method for metabolomics data using optimal selection of multiple internal standards. *BMC Bioinformatics* **8**: 93.

Tack AJM, Johnson MTJ, Roslin T. 2012. Sizing up community genetics: it's a matter of scale. *Oikos* **121**: 481–488.

Tennekes M. 2018. Tmap: Thematic maps in R. *Journal of Statistical Software* **84**: 1–39.

Thiebaut F, Hemerly AS, Ferreira PCG. 2019. A role for epigenetic regulation in the adaptation and stress responses of non-model plants. *Frontiers in Plant Science* **10**: 1–7.

van Tienderen PH. 1992. Variation in a population of *Plantago lanceolata* along a topographical gradient. *Oikos* **64**: 560–572.

Treutter D. 2005. Significance of flavonoids in plant resistance and enhancement of their biosynthesis. *Plant Biology* **7**: 581–591.

Westerhuis JA, Hoefsloot HCJ, Smit S, Vis DJ, Smilde AK, van Velzen EJJ, van Duijnhoven JPM, van Dorsten FA. 2008. Assessment of PLSDA cross validation. *Metabolomics* **4**: 81–89.

Wiklund S, Johansson E, Sjöström L, Mellerowicz EJ, Edlund U, Shockcor JP, Gottfries J, Moritz T, Trygg J. 2008. Visualization of GC/TOF-MS-based metabolomics data for identification of biochemically interesting compounds using OPLS class models. *Analytical Chemistry* **80**: 115–122.

Winkel-Shirley B. 2002. Biosynthesis of flavonoids and effects of stress. *Current Opinion in Plant Biology* **5**: 218–223.

Wishart Ds, Knox C, Guo Ac, Eisner R, Young N, Gautam B, Hau Dd, Psychogios N, Dong E, Bouatra S *et al.* 2009. HMDB: a knowledgebase for the human metabolome. *Nucleic Acids Research* **37**(Suppl. 1): D603–D610.

Woodruff DR, Meinzer FC. 2011. Water stress, shoot growth and storage of non-structural carbohydrates along a tree height gradient in a tall conifer. *Plant, Cell & Environment* **34**: 1920–1930.

Worley B, Powers R. 2013. Multivariate analysis in metabolomics. *Current Metabolomics* **1**: 92–107.

Yela JL, Lawton JH. 1997. Insect herbivore loads on native and introduced plants: a preliminary study. *Entomologia Experimentalis et Applicata* **85**: 275–279.

Youngentob KN, Zdenek C, van Gorsel E. 2016. A simple and effective method to collect leaves and seeds from tall trees. *Methods in Ecology and Evolution* **7**: 1119–1123.

Supporting Information

Additional Supporting Information may be found online in the Supporting Information section at the end of the article.

Fig. S1 Classification of the *Quercus robur* leaf developmental stages.

Fig. S2 Elemental analyses and isotope patterns of nitrogen (N) and carbon (C).

Fig. S3 Uniform manifold approximation and projection (UMAP) of nontargeted metabolomics.

Methods S1 Details on chromatography and mass spectrometry (MS) parameters.

Methods S2 Stable isotope analysis of carbon (C) and nitrogen (N).

Table S1 List of internal standards.

Table S2 METABOSCAPE 4.0 parameters.

Table S3 Metabolomic data matrix of samples from four forest stands in North Rhine-Westphalia and Asbesk.

Table S4 Metabolomic tree-to-tree variation and within-tree variation.

Table S5 Hypergeometric test results.

Please note: Wiley Blackwell are not responsible for the content or functionality of any Supporting Information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.