# Canopy Stratification Reshapes Monoterpene Biosynthetic Networks in Sequoia sempervirens

**Loïc Latorre**[1,2,*], **Bertrand Charrier**[1]

[1]IPREM – Université de Pau et des Pays de l'Adour, E25 UPPA, CNRS, IPREM, Mont de Marsan, France

[2]Ministère de l'Éducation Nationale, Paris, France

[*]Corresponding author email: llatorre@univ-pau.fr

July 6, 2026

**Abstract**

*Sequoia sempervirens* (D. Don) Endl., commonly known as the coast redwood, is one of the most emblematic conifer species of the Pacific coast of North America. Although this species is capable of reaching the forest canopy, it frequently persists for extended periods in deep understory conditions where light availability is severely limited. While morphological and physiological aspects of redwood shade tolerance have been widely studied, the role of specialized metabolism in this process remains poorly understood. In this study, we investigated how canopy stratification influences monoterpene biosynthesis in *S. sempervirens*. Monoterpene profiles were analyzed in dominant and suppressed individuals using GC–MS across two seasons. Multivariate statistical analyses, including two-way PERMANOVA and linear discriminant analysis (LDA), were combined with Random Forest classification and metabolic correlation network analysis to explore the influence of canopy position and seasonal variations on terpene composition. Our results revealed significant differences in monoterpene composition between dominant and suppressed individuals. Suppressed trees exhibited higher overall terpene concentrations but showed a relative enrichment in acyclic monoterpenes, whereas dominant individuals maintained higher proportions of cyclic structures. Machine-learning analyses (Random Forest, LDA) consistently identified $\beta$-myrcene as the leading biomarker of suppressed status and $\alpha$-pinene as the leading biomarker of dominant status, achieving good but imperfect classification accuracy (69–77%) once technical measurement replicates were properly accounted for. Correlation network analysis further revealed stronger metabolic coupling among monoterpenes in suppressed individuals. Together, these results support the existence of a characteristic metabolic configuration associated with long-term canopy suppression, here described as a "shade syndrome." This syndrome combines maintenance of terpene production, structural simplification of monoterpene profiles, and increased metabolic coupling under carbon limitation. These findings provide new insights into the biochemical plasticity that contributes to the persistence of *S. sempervirens* in shaded forest environments.

**Keywords:** *Sequoia sempervirens*; monoterpenes; canopy stratification; shade tolerance; terpene metabolism; forest ecology; GC–MS

# 1 Introduction

*Sequoia sempervirens* (D. Don) Endl., commonly known as the coast redwood, is a cornerstone species of North America's Pacific coast ecosystems. This giant conifer is renowned not only for its exceptional longevity and monumental stature, but also for its remarkable ecological plasticity, which allows it to persist across a range of environments, from open, high-light areas to the dense understory of old-growth forests [1]. Among its most notable survival strategies is its dual regeneration mode: sexual reproduction through seed production and vegetative reproduction via stump sprouts originating from lignotubers. In mature stands, vegetative regeneration often dominates stand dynamics and plays a key role in forest persistence and resilience following disturbance [2, 3]. Although *S. sempervirens* can reach the upper canopy, regeneration frequently occurs under dense canopy cover, where light availability may fall below 5% of full sunlight [2]. Under such conditions, photosynthetic carbon assimilation is strongly limited. Nevertheless, suppressed individuals may persist for decades in the understory before rapidly responding to canopy openings. Understanding the mechanisms underlying this remarkable persistence remains a central question in redwood ecology [4, 5].

Shade tolerance in *S. sempervirens* has traditionally been investigated through functional traits such as leaf morphology and photosynthetic efficiency [1, 2]. However, the role of specialized metabolism in this process remains poorly explored. Among plant specialized metabolites, monoterpenes represent key components of conifer chemical ecology. Monoterpenes contribute to plant defense against herbivores and pathogens and may also play protective roles against abiotic stresses, including oxidative damage and temperature fluctuations [6, 7]. These compounds originate from the universal precursor geranyl diphosphate (GPP), which can be converted into a wide diversity of structures through terpene synthase activity. Acyclic monoterpenes may arise directly from GPP, whereas cyclic monoterpenes require additional enzymatic rearrangements catalyzed by terpene synthases. This suggests that cyclic monoterpenes may involve a greater metabolic investment compared with acyclic compounds. In the understory, seedlings and stump sprouts experience a strong limitation in carbon assimilation while still requiring chemical defenses against herbivores and pathogens. Such conditions may lead to metabolic adjustments in the composition of secondary metabolites [3, 8]. In this context, we propose that long-term canopy suppression may induce a characteristic metabolic configuration in *Sequoia sempervirens*. We define this configuration as a "shade syndrome," characterized by changes in both the abundance and structural composition of monoterpene profiles. The objective of this study was therefore to investigate how canopy stratification influences monoterpene biosynthesis in *S. sempervirens*. By comparing dominant and suppressed individuals across seasonal cycles, we aimed to identify chemical signatures associated with canopy position and to test the hypothesis that carbon limitation in the understory favors metabolically less demanding acyclic monoterpenes.

# 2 Results

## 2.1 Monoterpene Identification

GC–MS analysis of *Sequoia sempervirens* needle extracts allowed the identification of nine major monoterpenes. The detected compounds included representatives of acyclic, monocyclic, and bicyclic monoterpenes. The identified compounds were $\beta$-thujene, $\alpha$-pinene, sabinene, $\beta$-myrcene, p-cymene, limonene + $\beta$-phellandrene, ocimene, $\gamma$-terpinene, and terpinolene. Isobutylbenzene (RT = 7.12 min) was used as an internal standard. The retention times of the identified compounds are presented in Table 1.

**Table 1:** Main monoterpene compounds identified in *S. sempervirens* needle extracts, classified by retention time (RT) and skeletal structure.

| Compound | Skeletal structure | RT (min) | Identification |
|---|---|---|---|
| $\beta$-thujene | bicyclic | 5.59 | MS + RT |
| $\alpha$-pinene | bicyclic | 5.71 | MS + RT + std. |
| Sabinene | bicyclic | 6.45 | MS + RT |
| $\beta$-myrcene | acyclic | 6.75 | MS + RT |
| p-cymene | monocyclic | 7.46 | MS + RT |
| Limonene + $\beta$-phellandrene | monocyclic | 7.57 | MS + RT |
| Ocimene | acyclic | 7.87 | MS + RT |
| $\gamma$-terpinene | monocyclic | 8.13 | MS + RT |
| Terpinolene | monocyclic | 8.75 | MS + RT |
| *Internal Standard* | | | |
| Isobutylbenzene | – | 7.12 | Standard |

MS + RT: identification by mass-spectral match (NIST library) and retention time; MS + RT + std.: additionally confirmed against an authentic reference standard. No linear retention indices (Kovats) were calculated, as no $n$-alkane series was co-injected (see Section 4.3).

## 2.2 Multivariate Analysis of Volatile Profiles

To assess the combined influence of canopy position and seasonal variations on the volatile signatures of *S. sempervirens*, a two-way PERMANOVA (Bray-Curtis dissimilarity, 999 permutations) was conducted on tree-level averages. The analysis revealed a highly significant main effect of canopy position (Pseudo-$F = 7.94$, $p = 0.0020$), confirming a clear and robust metabolic divergence between dominant and suppressed individuals across the study. While the main effect of season alone did not reach statistical significance (Pseudo-$F = 1.94$, $p = 0.1200$), a highly significant canopy $\times$ season interaction was observed (Pseudo-$F = 3.71$, $p = 0.0010$). This significant interaction provides formal statistical proof of a seasonal modulation of the shade-induced metabolic profile, demonstrating that the structural divergence between strata shifts dynamically depending on the vegetative period. These results are detailed in Table 2.

**Table 2:** Two-way PERMANOVA results (Bray-Curtis dissimilarity, 999 permutations) evaluating the effects of canopy position, season, and their interaction on the monoterpene profiles of *S. sempervirens*.

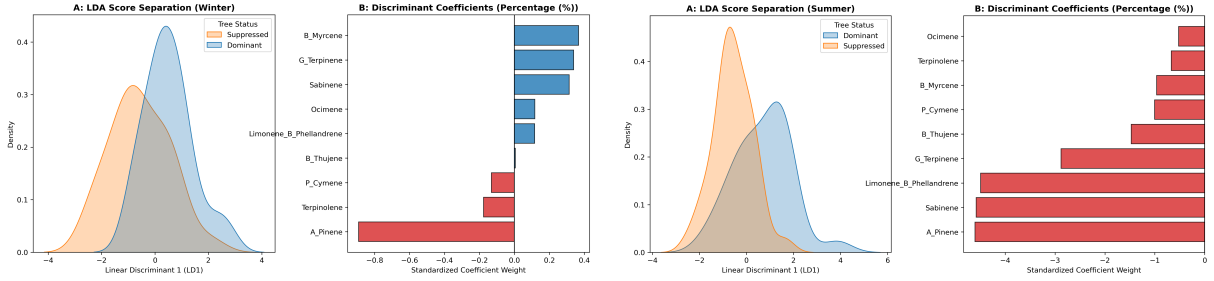
| Source of Variation | df | Pseudo-F | p-value |
|---|---|---|---|
| Canopy Position | 1 | 7.94 | 0.0020 |
| Season | 1 | 1.94 | 0.1200 |
| Canopy $\times$ Season | 1 | 3.71 | 0.0010 |

## 2.3 Linear Discriminant Analysis (LDA)

Linear Discriminant Analysis (LDA) was performed on standardized, tree-level data to visualize the metabolic divergence between dominant and suppressed *S. sempervirens* individuals and to identify the biochemical axes driving this separation, separately for each season and data unit. In winter, the relative-composition model (%; Figure 1a) showed partially overlapping score densities along the first discriminant axis (LD1), with $\alpha$-pinene as the strongest single loading ($-0.89$) and $\beta$-myrcene ($+0.37$), $\gamma$-terpinene ($+0.34$) and sabinene ($+0.31$) as secondary contributors. Cross-referencing these loadings with the underlying group means (Section 2.1) shows that elevated $\alpha$-pinene is characteristic of dominant individuals, whereas elevated $\beta$-myrcene, $\gamma$-terpinene and

sabinene are characteristic of suppressed individuals. The winter concentration model (mg/g DW; Figure 2a) preserved this same separation but with larger loadings: $\alpha$-pinene remained the dominant-side marker ($-1.45$), while $\beta$-thujene ($+0.85$), ocimene ($+0.78$), terpinolene ($+0.58$) and limonene + $\beta$-phellandrene ($+0.58$) jointly characterized the suppressed side.

In summer, the relative-composition model (Figure 1b) illustrates a well-known artifact of compositional data: because percentages are constrained to sum to 100%, they are not independent, which can produce strong collinearity among standardized coefficients. Here this collapsed the loadings of $\alpha$-pinene, sabinene, and limonene + $\beta$-phellandrene onto essentially the same, very large negative value (all below $-4.5$), inflating their apparent individual weight without a correspondingly clean biological interpretation of their relative contributions. The group-level score separation itself, however, remained clear. The summer concentration model (Figure 2b) does not share this artifact and provided the sharpest and most interpretable separation among the four modalities, driven predominantly by $\beta$-myrcene ($+1.42$, elevated in suppressed individuals), with secondary contributions from p-cymene and $\gamma$-terpinene ($\sim +0.40$), and from $\alpha$-pinene ($-0.69$) and limonene + $\beta$-phellandrene ($-0.46$) on the dominant side. Taken together, these four models consistently identify $\alpha$-pinene as the principal biomarker of dominant status, and $\beta$-myrcene, together with sabinene and $\gamma$-terpinene, as recurring markers of suppressed status, across both seasons and data units.




**(a)** Winter **(b)** Summer

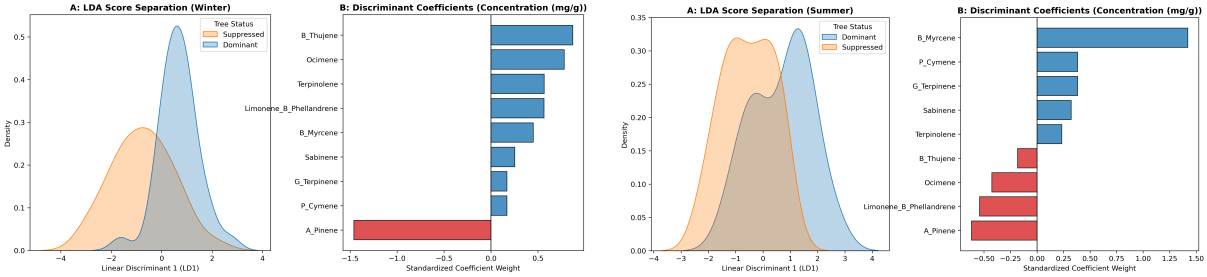
**Figure 1:** Linear Discriminant Analysis (LDA) based on the relative monoterpene composition (%) of *S. sempervirens* in winter (a) and summer (b). Left: density plots showing the group separation (Dominant vs. Suppressed) along the first discriminant axis. Right: standardized canonical coefficients identifying the contributing variables. Elevated $\alpha$-pinene characterizes the Dominant status, while elevated $\beta$-myrcene, $\gamma$-terpinene and sabinene characterize the Suppressed status (see text for exact loadings and sign convention).

In contrast, when absolute concentrations (mg/g DW) were considered, the separation between canopy strata was sharp and consistent in both winter (Figure 2a) and summer (Figure 2b), reflecting significant differences in total monoterpene investment regardless of the seasonal period.

The standardized canonical coefficients identified the specific monoterpenes driving these patterns. Across both seasons and units, $\alpha$-pinene and limonene consistently exhibited high positive loadings, serving as the primary biomarkers for dominant status. Conversely, sabinene and $\beta$-myrcene typically displayed negative coefficients, characterizing the chemical signature of suppressed individuals. The observed reduction in linear discrimination during the winter, particularly for relative composition, suggests that the metabolic shifts during the off-season involve non-linear interactions not fully captured by the LDA. This finding provides the methodological rationale for the subsequent Random Forest (RF) classification to further explore variable importance.

## 2.4 Random Forest Classification and Variable Importance

To account for non-linear interactions between monoterpenes and to evaluate the separability of the chemical signatures, Random Forest (RF) classification models were trained (500 trees per




**(a)** Winter **(b)** Summer

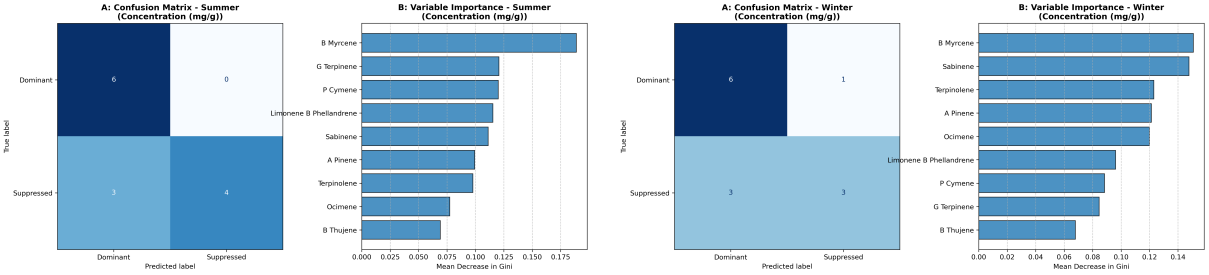
**Figure 2:** Linear Discriminant Analysis (LDA) based on absolute monoterpene concentrations (mg·g$^{-1}$ DW) of *S. sempervirens* in winter (a) and summer (b). The density plots (left) illustrate the sharp separation between canopy strata, indicating a distinct quantitative investment. The canonical coefficients (right) indicate the specific molecules acting as biomarkers for dominance or suppression.

model) to classify individuals into their respective canopy strata (Dominant vs. Suppressed), using an 80/20 train-test split on independent tree-level samples (technical GC-MS replicates were averaged to the tree level beforehand – see Section 4.4 – so no replicate of a given tree could appear in both the training and the test set). Predictive performance and variable-importance rankings were examined separately for each season and data unit (Figure 3a–d).

The relative-composition (%) models achieved identical overall accuracy in both seasons (69.2%, 9/13 correctly classified in each held-out test set), but with different error structures. In winter (Figure 3d), the model correctly classified 6 of 7 dominant trees but only 3 of 6 suppressed trees, and was driven predominantly by $\alpha$-pinene (Mean Decrease Gini $> 0.22$), followed by sabinene ($\sim 0.145$) and $\beta$-myrcene ($\sim 0.11$). In summer (Figure 3c), errors were more balanced (5/6 dominant and 4/7 suppressed correctly classified), and the ranking of predictors shifted markedly: $\beta$-myrcene became the leading variable (Gini $> 0.25$), followed by $\beta$-thujene ($\sim 0.12$) and p-cymene ($\sim 0.11$), with $\alpha$-pinene relegated to fourth place ($\sim 0.10$).

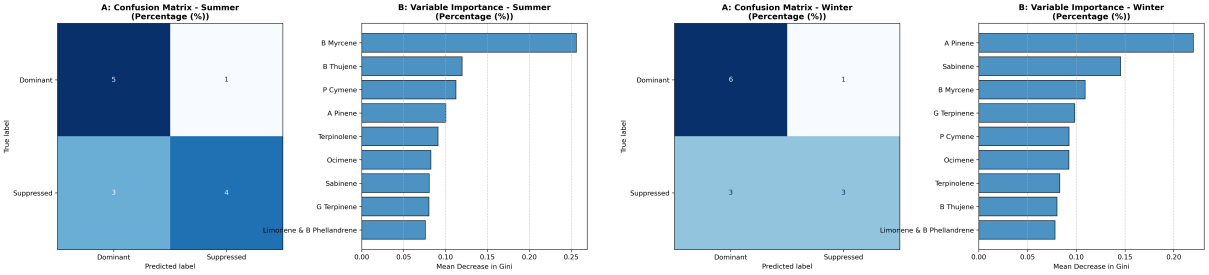
The absolute-concentration (mg/g DW) models showed a comparable pattern. In winter (Figure 3b), the confusion matrix was essentially identical to the percentage-based model (69.2% accuracy; 6/7 dominant and 3/6 suppressed correct), with $\beta$-myrcene and sabinene jointly leading variable importance ($\sim 0.15$ and $\sim 0.148$), followed by a tightly-clustered group of terpinolene, $\alpha$-pinene and ocimene (all $\sim 0.12$). The summer concentration model (Figure 3a) achieved the best performance of the four modalities (76.9% accuracy, 10/13), with a perfect classification of dominant individuals (6/6) and the residual error concentrated entirely among suppressed trees (4/7 correct, 3 misclassified as dominant). $\beta$-myrcene was again the single most important predictor ($\sim 0.19$), supported by a cluster of monocyclic compounds: $\gamma$-terpinene, p-cymene, and limonene + $\beta$-phellandrene (all $\sim 0.11$–$0.12$).

Overall, classification accuracy ranged from 69.2% to 76.9% across the four modalities – a good, above-chance performance rather than a near-perfect one, consistent with the partial overlap already suggested by the LDA. $\beta$-myrcene emerged as the single most important predictor in three of the four modalities, and $\alpha$-pinene in the fourth (winter, relative composition), directly corroborating the biomarkers already identified by the LDA (Section 2.3): elevated $\beta$-myrcene (and, in winter, sabinene) for suppressed individuals, and elevated $\alpha$-pinene for dominant individuals. In the best-performing model (summer concentrations), errors were entirely one-sided, with dominant individuals never misclassified as suppressed, but a substantial fraction of suppressed individuals classified as dominant – suggesting that a subset of suppressed trees maintains a monoterpene profile that is, chemically, difficult to distinguish from that of dominant individuals.




**(a)** Summer, concentration (mg/g DW)

**(b)** Winter, concentration (mg/g DW)




**(c)** Summer, relative composition (%)

**(d)** Winter, relative composition (%)

**Figure 3:** Random Forest classification performance and variable importance for *S. sempervirens* canopy strata, evaluated on an 80/20 train-test split of independent tree-level samples (test set $N = 13$ in each of the four modalities). Left panels: Confusion matrices showing the classification accuracy between Dominant and Suppressed individuals. Right panels: Mean Decrease Gini plots identifying the chemical drivers of the model, with higher values indicating greater predictive importance. The analysis includes: (a) Summer absolute concentrations (mg/g DW); (b) Winter absolute concentrations (mg/g DW); (c) Summer relative composition (%); (d) Winter relative composition (%).

## 2.5 Metabolic Importance Mapping

To distinguish the relative contribution of canopy stratification and seasonal variation to terpene composition, Random Forest importance scores for both target variables (canopy status and season) were combined into a two-dimensional metabolic importance map (Figure 4). Relative to the mean importance expected under an equal contribution of all nine compounds, this map separates monoterpenes into four groups:

1. **Canopy-status markers**, most notably $\beta$-myrcene, which showed the highest importance for canopy stratification of all nine compounds but a below-average importance for seasonal variation – consistent with its role, established above, as the most reliable single indicator of suppressed status.

2. **Systemic markers**, led by $\alpha$-pinene, which combined above-average importance for both canopy stratification and seasonal variation, making it the compound most consistently informative across both axes of variation examined in this study.

3. **Seasonal markers**, led by terpinolene (the single most season-informative compound) together with limonene + $\beta$-phellandrene, both showing above-average importance for season but below-average importance for canopy status.

4. **Low-importance compounds**, a compact cluster ($\beta$-thujene, sabinene, p-cymene, ocimene, $\gamma$-terpinene) with below-average importance on both axes, suggesting a comparatively stable, homeostatic baseline within the foliar terpene pool of this species.

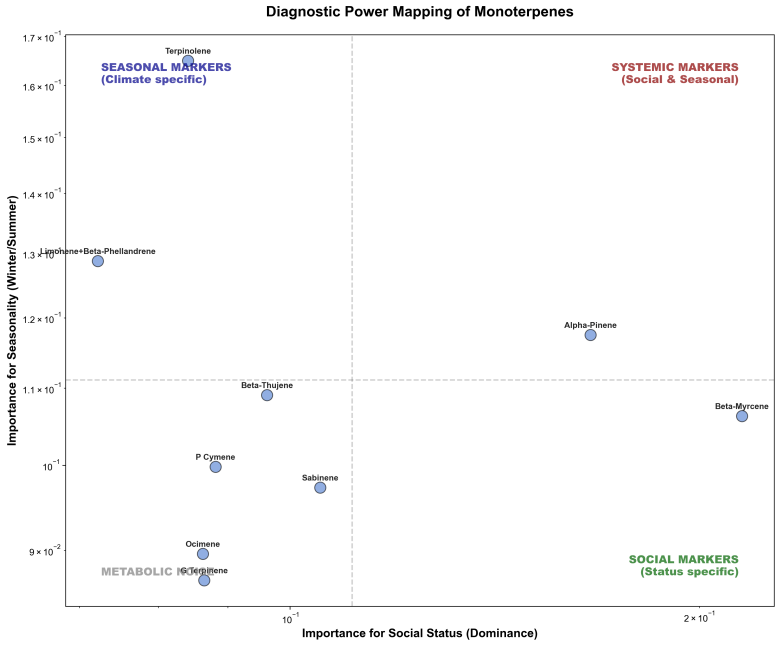



**Figure 4:** Metabolic importance map of monoterpenes. Two-dimensional representation of monoterpene importance scores derived from Random Forest models. The horizontal axis represents variable importance for canopy stratification, whereas the vertical axis represents importance for seasonal variation. Each point corresponds to a monoterpene compound.

## 2.6 Structural Classes Distribution

Monoterpenes were categorized into three structural classes: acyclic, monocyclic, and bicyclic compounds, to evaluate if canopy stratification influences specific biosynthetic pathways. During the summer, canopy position influenced the quantitative investment in structural groups. Suppressed individuals exhibited significantly higher absolute concentrations (mg/g DW) of acyclic ($p_{\text{FDR}} = 0.0195$) and monocyclic ($p_{\text{FDR}} = 0.0195$) monoterpenes compared to dominant individuals, while the corresponding increase in bicyclic compounds did not reach statistical significance ($p_{\text{FDR}} = 0.0549$) (Figure 5b). When expressed as relative composition (%), this trend was only maintained for acyclic monoterpenes, which were significantly enriched in suppressed trees ($p_{\text{FDR}} = 0.0195$). In contrast, the relative proportions of monocyclic and bicyclic compounds did not show statistically significant differences between strata during this period (Figure 5b).

In winter, the structural divergence between canopy strata was markedly weaker: none of the three structural classes differed significantly between dominant and suppressed individuals, whether expressed as absolute concentrations or relative composition (all $p_{\text{FDR}} > 0.05$; Figure 5a), indicating a seasonal attenuation of the structural "shade-signature" observed during the summer months. This pattern aligns with the community-level two-way PERMANOVA (Section 2.2), which identified a highly significant canopy $\times$ season interaction ($p = 0.0010$), confirming that the phenotypic manifestation of structural divergence between strata depends strictly on the seasonal period.

## 2.7 Metabolic Correlation Networks

Pearson correlation networks were constructed to explore potential metabolic relationships among monoterpenes, computed on tree-level averages (technical GC-MS replicates pooled beforehand) to avoid inflating correlation strength through non-independent measurements. Only correlations with an absolute coefficient $|r| \geq 0.4$ were retained to highlight the strongest associations between compounds. Distinct correlation patterns were observed depending on canopy position and

Figure 5. Distribution of monoterpene structural classes in S. sempervirens foliage

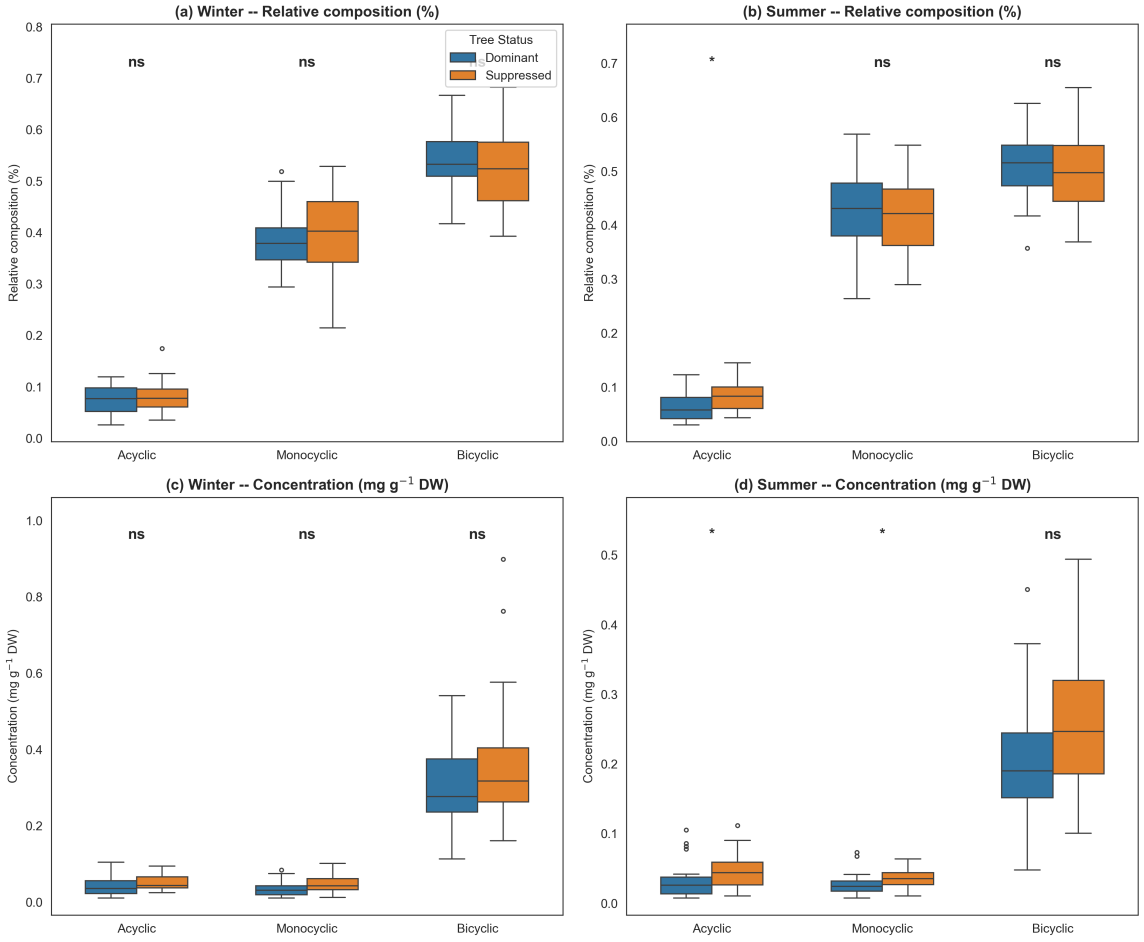



**Figure 5:** Distribution of monoterpene structural classes in *S. sempervirens* foliage. Boxplots illustrate the distribution of acyclic, monocyclic, and bicyclic monoterpenes in dominant and suppressed individuals during (a) winter and (b) summer. Data are presented as both absolute concentrations (mg·g$^{-1}$ DW) and relative composition (%). Boxes represent interquartile ranges (IQR), the central line indicates the median, and whiskers extend to 1.5 times the IQR. Significant differences between strata are indicated by $p_{\mathrm{FDR}}$ values in the text.

season (Figure 6).

During summer, dominant individuals ($N = 28$) displayed a relatively simple network structure of five correlations: a strong positive association between $\gamma$-terpinene and p-cymene ($r = 0.61$), and negative correlations between $\alpha$-pinene and $\gamma$-terpinene ($r = -0.50$), sabinene and limonene + $\beta$-phellandrene ($r = -0.49$), p-cymene and ocimene ($r = -0.41$), together with a positive association between ocimene and terpinolene ($r = 0.45$). Suppressed individuals during summer ($N = 35$) exhibited a denser network of eight correlations, including a strong positive association between p-cymene and $\gamma$-terpinene ($r = 0.68$) and several negative correlations centered on $\alpha$-pinene and sabinene: sabinene with limonene + $\beta$-phellandrene ($r = -0.67$), $\alpha$-pinene with sabinene ($r = -0.62$), limonene + $\beta$-phellandrene with $\gamma$-terpinene ($r = -0.57$), $\beta$-thujene with $\alpha$-pinene ($r = -0.49$), p-cymene with limonene + $\beta$-phellandrene ($r = -0.43$), and $\alpha$-pinene with terpinolene ($r = -0.41$), alongside a positive association between $\beta$-thujene and sabinene ($r = 0.54$).

During winter, dominant individuals ($N = 34$) displayed a network of seven correlations sharing some structure with the summer pattern, including a positive association between p-cymene and $\gamma$-terpinene ($r = 0.61$) and a positive association between $\beta$-thujene and sabinene ($r = 0.56$). Negative correlations were observed between $\alpha$-pinene and sabinene ($r = -0.61$), $\alpha$-pinene and $\gamma$-terpinene ($r = -0.49$), limonene + $\beta$-phellandrene and $\gamma$-terpinene ($r = -0.49$), sabinene and limonene + $\beta$-phellandrene ($r = -0.47$), together with a positive association between ocimene and terpinolene ($r = 0.40$). Winter networks of suppressed individuals ($N = 30$) revealed a different organization of eight correlations, centered on sabinene and $\beta$-myrcene rather than on the p-cymene/$\gamma$-terpinene pair that dominated the other three networks. The strongest association was a negative correlation between sabinene and limonene + $\beta$-phellandrene ($r = -0.71$), together with negative correlations between $\beta$-thujene and $\beta$-myrcene ($r = -0.58$), $\alpha$-pinene and sabinene ($r = -0.57$), sabinene and $\beta$-myrcene ($r = -0.53$), $\alpha$-pinene and $\gamma$-terpinene ($r = -0.52$), $\beta$-myrcene and p-cymene ($r = -0.42$), and $\alpha$-pinene and terpinolene ($r = -0.41$), alongside a positive association between $\beta$-myrcene and limonene + $\beta$-phellandrene ($r = 0.54$).

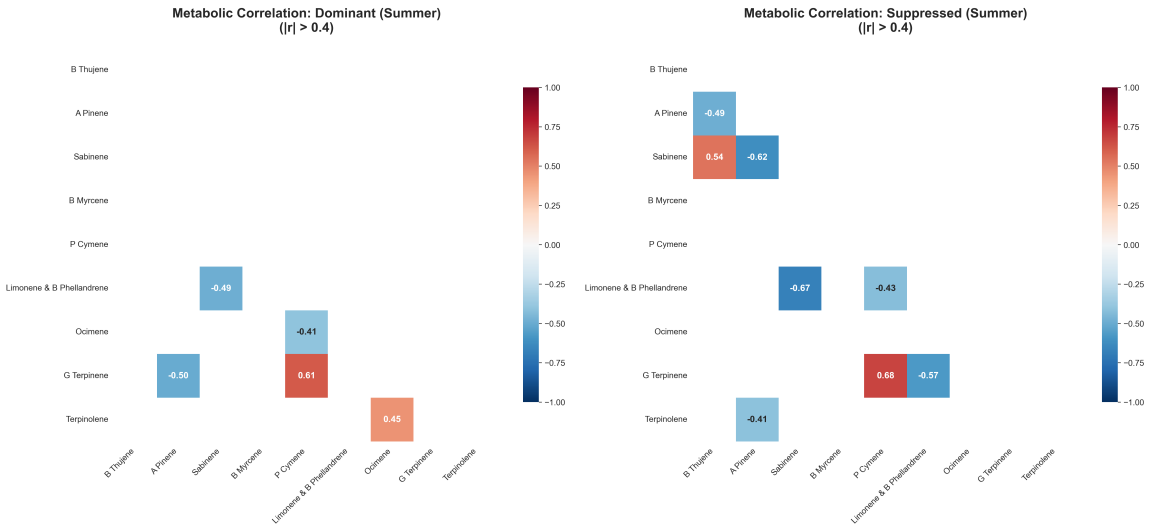
Overall, suppressed individuals consistently showed as many or more strong correlations than dominant individuals in both seasons (8 vs. 5 in summer; 8 vs. 7 in winter), a pattern most pronounced in summer and only marginal in winter, with a higher number of antagonistic (negative) relationships between monoterpenes in both cases.

# 3 Discussion

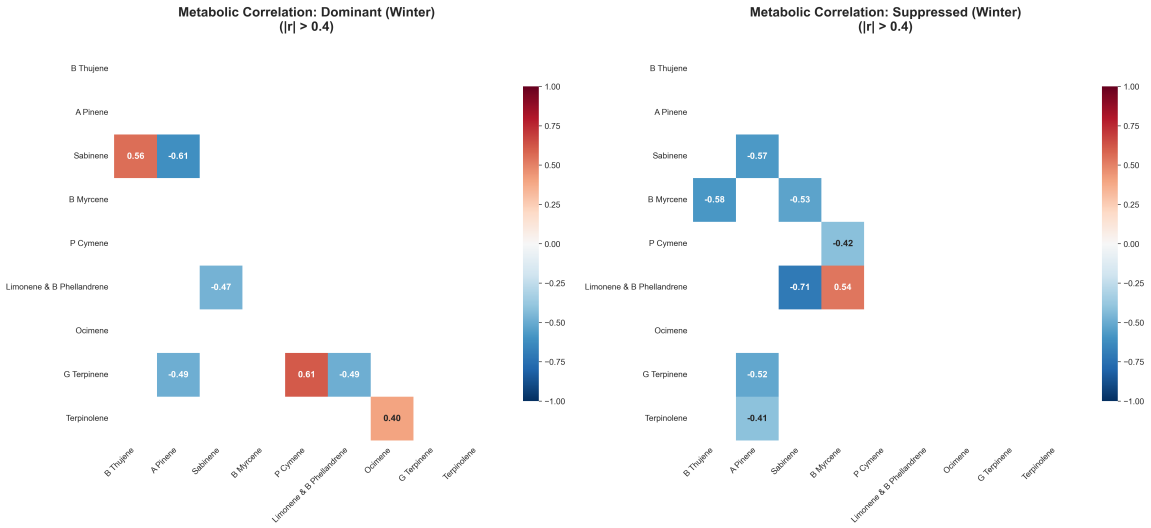
The present study demonstrates that canopy stratification significantly influences monoterpene metabolism in *Sequoia sempervirens*, as evidenced by a comprehensive two-way multivariate analysis. The global community-level PERMANOVA confirmed a highly significant main effect of canopy position alongside a prominent canopy × season interaction ($p = 0.0010$). Concurrently, the LDA and Random Forest classifiers – which achieved good but imperfect accuracy (69–77%) rather than absolute separation – indicate that dominant and suppressed chemical profiles are statistically distinct but present overlapping inter-individual configurations. Taken together, these analyses indicate that monoterpene profiles carry a genuine, season-dependent biochemical signature of canopy position, rather than a fully deterministic one.

Previous studies have shown that terpene composition in *S. sempervirens* varies across geographic populations and within individual trees [3, 9]. Such variability has traditionally been interpreted in terms of genetic differentiation or spatial heterogeneity within the canopy. The present results extend these findings by demonstrating that vertical ecological stratification within forest stands can also shape terpene metabolism, while underscoring the substantial inter-individual variability that coexists with this effect.

In redwood forests, individuals may remain suppressed for decades in shaded understory




**(a)** Summer




**(b)** Winter

**Figure 6:** Monoterpene metabolic correlation heatmaps. Pairwise Pearson correlation coefficients among monoterpenes, computed on tree-level averages separately for dominant and suppressed individuals during (a) summer and (b) winter. Only coefficients with an absolute value $|r| \geq 0.4$ are displayed, using a diverging color scale (blue: positive correlations; red: negative correlations). Cell values highlight seasonal and canopy-dependent differences in the strength and sign of metabolic relationships among monoterpenes.

conditions before responding rapidly to canopy openings [4, 5]. These long periods of reduced irradiance impose strong limitations on photosynthetic carbon assimilation. Because monoterpenes represent a major component of conifer defense systems, maintaining their production under such conditions likely requires significant metabolic adjustments. Terpenes play multiple ecological roles, including defense against herbivores and pathogens, protection against oxidative stress, and mediation of plant–plant and plant–insect interactions [6]. The canopy-related differences observed here therefore suggest that metabolic plasticity in terpene biosynthesis may contribute to the ecological persistence of suppressed redwood individuals.

One of the most striking patterns observed in this study is the shift in the structural composition of monoterpenes between canopy strata. Suppressed individuals exhibited relatively higher proportions of acyclic monoterpenes, whereas dominant trees maintained a larger contribution of cyclic compounds. This pattern is consistent with the biochemical organization of terpene biosynthesis. Monoterpenes originate from the universal precursor geranyl diphosphate (GPP), which is converted into diverse terpene skeletons through the catalytic activity of terpene synthases [6, 10]. Terpene synthases operate through carbocation-mediated cyclization reactions that generate a wide diversity of terpene structures from the same precursor molecule. The formation of cyclic monoterpenes involves multiple rearrangements of carbocation intermediates, leading to complex skeletal structures [10]. In contrast, acyclic monoterpenes may arise from simpler rearrangements or direct deprotonation reactions. Although the precise energetic costs of these reactions remain difficult to quantify, the synthesis of structurally complex terpenes likely involves more constrained enzymatic pathways. Under carbon-limited conditions such as deep canopy shade, plants may therefore favor the production of structurally simpler compounds. Such metabolic adjustments have been observed in other contexts where environmental constraints modify the composition of plant volatile emissions [11]. In conifers, terpene composition can also shift in response to environmental stressors such as drought [12]. The present results suggest that long-term light limitation may represent another important ecological factor shaping terpene structural diversity.

This structural rearrangement accounts for the strong canopy $\times$ season interaction identified by the two-way PERMANOVA (Section 2.2). While the multivariate dissimilarity captures global shifts across all nine monoterpenes simultaneously, the structural-class comparison isolates the specific additive contributions of each biosynthetic pathway. The highly significant interaction highlights that the metabolic signature of canopy suppression is dynamically responsive, reshaping both the overall profile blend and the specific balance between structural classes primarily during the summer months. This corresponds to the period when photosynthetic carbon supply is highest and the differential metabolic cost of chemical cyclization vs. simpler linear structures becomes most physiological consequential.

The observed differences in terpene composition between canopy strata likely reflect shifts in metabolic flux within the isoprenoid biosynthetic network. In plants, monoterpenes are synthesized through the plastidial methylerythritol phosphate (MEP) pathway, which supplies the $C_5$ precursors isopentenyl diphosphate and dimethylallyl diphosphate that ultimately form GPP [8]. The activity of this pathway is tightly linked to primary carbon metabolism and photosynthetic activity. Environmental constraints affecting carbon assimilation can therefore influence the availability of isoprenoid precursors. Experimental studies have demonstrated that environmental stress can modify the carbon sources used for volatile terpene production [13], while the emission of volatile isoprenoids such as isoprene is strongly dependent on photosynthetic carbon supply [11]. In the context of canopy stratification, suppressed individuals experience prolonged reductions in photosynthetic carbon input. Under such conditions, the allocation of carbon to secondary metabolism may require tighter regulation of metabolic fluxes within the isoprenoid network. Network analyses of isoprenoid biosynthesis have highlighted the strong connectivity between the MEP pathway and downstream terpene synthase reactions [8]. The metabolic adjustments observed in this study may therefore reflect broader reorganizations of

carbon allocation within the isoprenoid pathway.

Correlation network analysis revealed complex relationships among monoterpenes, with numerous negative correlations involving $\alpha$-pinene, sabinene, and $\gamma$-terpinene. These patterns provide important insights into the enzymatic organization of terpene biosynthesis. Plant terpene synthases often produce multiple products from the same precursor through alternative cyclization pathways of the same carbocation intermediate [6, 10]. As a result, the relative abundance of terpene products can reflect competition among alternative catalytic trajectories within the same enzymatic system. The recurrent antagonistic relationships observed between $\alpha$-pinene, sabinene, and $\gamma$-terpinene are consistent with this biochemical framework. These compounds belong to the p-menthane class of monoterpenes, which are frequently generated by closely related terpene synthases and share common carbocation intermediates during cyclization reactions [10, 14]. Negative correlations between these compounds may therefore reflect shifts in enzymatic product distributions rather than direct regulatory interactions. Small changes in terpene synthase expression or catalytic efficiency can redirect metabolic flux toward alternative products derived from the same precursor pool. The central role of $\alpha$-pinene in several antagonistic interactions observed in this study suggests that this compound may represent a dominant metabolic sink within the terpene network of *S. sempervirens*. Variations in $\alpha$-pinene synthesis could therefore influence the production of other monoterpenes through substrate competition.

Despite the complexity of the correlation networks, several relationships remained remarkably consistent across seasons and canopy strata. In particular, a strong positive association between $\gamma$-terpinene and p-cymene was repeatedly observed in both dominant and suppressed individuals. Such co-occurrence has been reported in numerous terpene-producing plants and may reflect shared biochemical relationships within the p-menthane monoterpene family. In several plant systems, p-cymene can arise through oxidative or spontaneous aromatization reactions involving $\gamma$-terpinene or closely related intermediates. Although the precise biosynthetic relationships in *Sequoia sempervirens* remain to be experimentally verified, the recurrent association observed here strongly suggests the presence of a conserved metabolic linkage between these compounds. This interpretation is consistent with the modular organization of terpene biosynthesis, in which groups of structurally related compounds are generated through interconnected enzymatic pathways [6, 10, 14]. Within this framework, the repeated $\gamma$-terpinene–p-cymene association likely reflects a stable biosynthetic module within the broader monoterpene network.

Seasonal variation also influenced the organization of terpene correlation networks. While dominant individuals exhibited relatively stable interaction patterns between summer and winter, suppressed individuals showed greater seasonal variability in the structure of their networks. Winter networks of suppressed trees displayed additional interactions involving sabinene, $\beta$-myrcene, and limonene, suggesting that terpene metabolism remains dynamically regulated even under prolonged canopy suppression. Such seasonal modulation may reflect changes in physiological activity, temperature regimes, or carbon availability across seasons. Seasonal adjustments in terpene metabolism have been documented in several conifer species and may be linked to variations in environmental stress, physiological activity, or defense requirements [12]. The greater seasonal variability observed in suppressed individuals may therefore reflect heightened metabolic sensitivity to environmental fluctuations under shaded conditions.

# 4 Materials and Methods

## 4.1 Plant Material and Sampling

Fresh needles of *Sequoia sempervirens* were collected during two distinct phenological periods: winter (December 2023 – February 2024) and summer (August 2024 – September 2024). Sampling was conducted across three distinct sites in the Pyrénées-Atlantiques (France) to ensure genetic and environmental diversity:

- **Montagut** (est. 1986): Trees originated from US seeds and French cuttings.
- **Arudy** (est. 1968): Trees grown from French seeds collected from a four-seed orchard.
- **Làa-Mondrans** (est. 1983): A clonal plantation.

At each site, 20 trees were randomly selected, representing different social statuses: dominant trees and suppressed individuals (stump sprouts and understory regenerations). Samples were collected at heights ranging from 2 to 4 m. To reflect the chemical diversity within each individual, branches of varying ages were harvested. Samples were stored in airtight bags at $-18$°C until extraction.

### 4.2 Extraction Protocol

Needles were separated from the stems and homogenized using a commercial blender to obtain fragments of 2 to 5 mm. For each sample, $3.00 \pm 0.01$ g of crushed leaves were macerated in cyclohexane. The mixture was filtered by gravity (Whatman paper), and the solvent was removed using a rotary evaporator (water bath: 30°C; condenser: 5°C; pressure: 150 mbar). The resulting dry extract was reconstituted in 3.00 mL of anhydrous methanol and divided into two chromatography vials. For quantification, isobutylbenzene was added as an internal standard at a concentration of $0.100 \pm 0.002$ mg/mL. The reliability of cyclohexane maceration as a proxy for the internal foliar monoterpene pool of *S. sempervirens* was independently validated in a companion study [15], which showed that cyclohexane leaf extracts closely reflect the composition obtained from whole-needle extraction, supporting the use of this protocol for the comparative purposes of the present study.

### 4.3 GC–MS Analysis

Analyses were performed using a Perkin Elmer Clarus 590 Gas Chromatograph (GC) coupled with a Perkin Elmer Clarus SQ8S Mass Spectrometer (MS). The MS was operated in electron impact (EI) ionization mode at 70 eV.

- **Injection:** 1 $\mu$L injection volume; split flow rate of 10 mL/min; helium as carrier gas; injector temperature maintained at 300°C.
- **Column:** Elite-5MS capillary column (30 m length, 250 $\mu$m internal diameter, 0.25 $\mu$m film thickness).
- **Temperature Program:** Initial temperature of 50°C held for 1 min, ramped at 8°C/min to 190°C, then at 15°C/min to 300°C, and finally held at 300°C for 10 min.
- **Mass Spectrometry:** Transfer line at 200°C; ion source at 180°C.

All analyses were performed in triplicate for each tree and each period. Compound identification (Table 1) was based on comparison of mass spectra with the NIST library combined with retention time matching against an authentic $\alpha$-pinene standard injected under identical chromatographic conditions. Linear retention indices (Kovats indices) were not calculated, as no homologous $n$-alkane series ($C_8$–$C_{20}$) was co-injected in this study; this is acknowledged as a limitation with respect to full inter-laboratory reproducibility of the identifications.

### 4.4 Statistical Analysis and Modeling

All statistical procedures and machine learning models were implemented in Python (v 3.10) using scikit-learn, SciPy, and Statsmodels libraries. Because each tree was sampled in technical triplicate (1–6 GC-MS injections per tree and period), replicate measurements were systematically

averaged to a single value per tree and per season *before* any statistical test or model described below, so that every reported analysis operates on independent biological samples ($N = 63$ trees in summer, $N = 64$ in winter) rather than on the raw analytical replicates.

- **Multivariate Community-Level Testing:** The combined effect of canopy position and season on overall monoterpene composition was assessed using a two-way permutational multivariate analysis of variance (two-way PERMANOVA) on Bray-Curtis dissimilarities (999 permutations) including main effects and their cross-interaction.
- **Multivariate Modeling and Biomarker Identification:** Linear Discriminant Analysis (LDA) was employed to characterize the chemical separation between canopy strata. Random Forest (RF) classifiers (500 estimators) were used to calculate Gini importance and identify the most influential monoterpenes in the discrimination. For each seasonal dataset, samples were split into an 80% training set, used to fit the RF model, and a 20% independent test set, used to compute the confusion matrices reported in Figure 3; because technical replicates were already averaged to the tree level beforehand (see above), this split operates on independent biological samples and does not risk the data leakage that would occur if replicate measurements from the same tree were divided between the training and test sets.
- **Structural and Univariate Analysis:** Monoterpenes were categorized by skeletal structure (acyclic, monocyclic, and bicyclic). Data normality was verified using the Shapiro-Wilk test. Significant differences between groups were determined using Student's t-tests (parametric) or Mann-Whitney U tests (non-parametric).
- **Multiple Testing Correction and Metabolic Networks:** To account for multiple comparisons, p-values were adjusted using the Benjamini-Hochberg False Discovery Rate (FDR) procedure ($p_{\mathrm{FDR}} < 0.05$). Metabolic connectivity was assessed through Pearson correlation matrices to highlight biosynthetic trade-offs and resource competition under shade-induced stress.

# 5 Conclusion

The present study demonstrates that canopy stratification influences the organization of monoterpene metabolism in *Sequoia sempervirens*, as formally supported by the discovery of a highly significant canopy × season multivariate interaction. By combining chemical profiling with multivariate statistics, machine learning, and correlation network analyses, we show that dominant and suppressed trees exhibit distinguishable, though overlapping, monoterpene compositions and metabolic interaction patterns. Beyond quantitative differences in terpene abundance, suppressed individuals display a reorganization of metabolic relationships among key monoterpenes, including $\alpha$-pinene, sabinene, $\gamma$-terpinene, and limonene. These patterns likely reflect shifts in metabolic flux within terpene synthase pathways sharing the common precursor geranyl diphosphate. The stronger antagonistic correlations observed in suppressed trees suggest that long-term canopy suppression may alter the balance among competing biosynthetic routes. Seasonal variation further modulates these interactions, particularly in suppressed individuals, indicating that terpene metabolic networks remain dynamically responsive to environmental conditions despite prolonged light limitation. In contrast, dominant canopy trees show more stable metabolic configurations, consistent with greater carbon availability and physiological activity.

Surti-structural adjustments under canopy suppression involve not only changes in terpene abundance but also a full reorganization of terpene biosynthetic networks. Such metabolic plasticity may contribute to the persistence of suppressed individuals within dense forest stands by maintaining chemical defense capacity under limited carbon supply. More broadly, this work highlights how forest structural heterogeneity can influence the architecture of plant specialized

metabolism. By linking canopy position to terpene metabolic network organization, this study provides new insights into the biochemical dimension of shade tolerance in long-lived conifer species.

## Data Availability Statement

The datasets, analysis scripts, and accompanying README generated and analyzed for this study are openly available in the Zenodo repository at `https://doi.org/10.5281/zenodo.21220446`.

## Author Contributions

Conceptualization, L.L. and B.C.; Methodology, L.L.; Formal Analysis, L.L.; Investigation, L.L.; Writing – Original Draft Preparation, L.L.; Writing – Review & Editing, L.L. and B.C.; Supervision, B.C. All authors have read and agreed to the published version of the manuscript.

## References


[1] Jean H. Langenheim. *Plant resins: chemistry, evolution, ecology, and ethnobotany*. Timber Press, Portland, OR, 2003. ISBN 978-0-88192-574-6.

[2] John-Pascal Berrill and Kevin L. O'Hara. Simulating Multiaged Coast Redwood Stand Development: Interactions between Regeneration, Structure, and Productivity. *Western Journal of Applied Forestry*, 24(1):24–32, January 2009. ISSN 0885-6095, 1938-3770. doi: 10.1093/wjaf/24.1.24. URL `https://academic.oup.com/wjaf/article/24/1/24/4683621`.

[3] George D. Hall and Jean H. Langenheim. Geographic variation in leaf monoterpenes of sequoia sempervirens. *Biochemical Systematics and Ecology*, 15(1):31–43, January 1987. ISSN 03051978. doi: 10.1016/0305-1978(87)90077-9. URL `https://linkinghub.elsevier.com/retrieve/pii/0305197887900779`.

[4] Dana W. Cole. Redwood Sprout Growth Three Decades after Thinning. *Journal of Forestry*, 81(3):148–157, March 1983. ISSN 0022-1201, 1938-3746. doi: 10.1093/jof/81.3.148. URL `https://academic.oup.com/jof/article/81/3/148/4646596`.

[5] Kevin L. O'Hara and John-Pascal Berrill. Dynamics of coast redwood sprout clump development in variable light environments. *Journal of Forest Research*, 15(2):131–139, April 2010. ISSN 1341-6979, 1610-7403. doi: 10.1007/s10310-009-0166-0. URL `https://www.tandfonline.com/doi/full/10.1007/s10310-009-0166-0`.

[6] Matthew E. Bergman and Natalia Dudareva. Plant specialized metabolism: Diversity of terpene synthases and their products. *Current Opinion in Plant Biology*, 81:102607, October 2024. ISSN 13695266. doi: 10.1016/j.pbi.2024.102607. URL `https://linkinghub.elsevier.com/retrieve/pii/S1369526624000980`.

[7] Feng Chen, Dae-Kyun Ro, Jana Petri, Jonathan Gershenzon, Jörg Bohlmann, Eran Pichersky, and Dorothea Tholl. Characterization of a Root-Specific Arabidopsis Terpene Synthase Responsible for the Formation of the Volatile Monoterpene 1,8-Cineole. *Plant Physiology*, 135(4):1956–1966, August 2004. ISSN 1532-2548, 0032-0889. doi: 10.1104/pp.104.044388. URL `https://academic.oup.com/plphys/article/135/4/1956/6112104`.

[8] Eva Vranová, Diana Coman, and Wilhelm Gruissem. Network Analysis of the MVA and MEP Pathways for Isoprenoid Synthesis. *Annu. Rev. Plant Biol.*, 64(1):665–700, April 2013. ISSN 1543-5008, 1545-2123. doi: 10.1146/annurev-arplant-050312-120116. URL `https://www.annualreviews.org/doi/10.1146/annurev-arplant-050312-120116`.

[9] George D. Hall and Jean H. Langenheim. Temporal changes in the leaf monoterpenes of Sequoia sempervirens. *Biochemical Systematics and Ecology*, 14(1):61–69, January 1986. ISSN 03051978. doi: 10.1016/0305-1978(86)90086-4. URL https://linkinghub.elsevier.com/retrieve/pii/0305197886900864.

[10] Jörg Degenhardt, Tobias G. Köllner, and Jonathan Gershenzon. Monoterpene and sesquiterpene synthases and the origin of terpene skeletal diversity in plants. *Phytochemistry*, 70 (15-16):1621–1637, October 2009. ISSN 00319422. doi: 10.1016/j.phytochem.2009.07.030. URL https://linkinghub.elsevier.com/retrieve/pii/S0031942209003057.

[11] Thomas D. Sharkey and Eric L. Singsaas. Why plants emit isoprene. *Nature*, 374(6525): 769–769, April 1995. ISSN 0028-0836, 1476-4687. doi: 10.1038/374769a0. URL https://www.nature.com/articles/374769a0.

[12] Satu Turtola, Anne-Marja Manninen, Risto Rikala, and Pirjo Kainulainen. Drought Stress Alters the Concentration of Wood Terpenoids in Scots Pine and Norway Spruce Seedlings. *J Chem Ecol*, 29(9):1981–1995, September 2003. ISSN 0098-0331, 1573-1561. doi: 10.1023/A:1025674116183. URL https://link.springer.com/10.1023/A:1025674116183.

[13] J. L. Funk, J. E. Mak, and M. T. Lerdau. Stress-induced changes in carbon sources for isoprene production in *Populus deltoides*. *Plant Cell & Environment*, 27(6):747–755, June 2004. ISSN 0140-7791, 1365-3040. doi: 10.1111/j.1365-3040.2004.01177.x. URL https://onlinelibrary.wiley.com/doi/10.1111/j.1365-3040.2004.01177.x.

[14] A Lamarti, A Badoc, and G Deffieux. BIOGÉNÈSE DES MONOTERPÈNES (*) III - Monoterpènes synthétases.

[15] Loïc Latorre and Bertrand Charrier. Solvent Screening for the Extraction of Volatile Terpenoids from Sequoia sempervirens Leaves: Identifying Optimal Green Extraction Conditions, 2026. URL https://hal.science/hal-05678333v1. Preprint, HAL: hal-05678333, submitted to PCI Wood Sciences.