**Prospective title:** Biological and ecological correlates of seed oil content in alpine species

Authors: C. Espinosa del Alba1,2, A. Mondoni3, E. Fernández-Pascual1,2 & B. Jiménez-Alfaro1,2.

Affiliations:

1-Biodiversity Research Institute (IMIB), University of Oviedo-CSIC-Principality of Asturias, Mieres, Asturias, 33600 Spain.

2-Department of Organismal and systems Biology, University of Oviedo, Oviedo, Asturias, 33071 Spain

3-Dipartimento di Scienze della Terra e dell’Ambiente, University of Pavia, 27100 Pavia, Italy.

**Correspondence author**: Clara Espinosa del Alba. Email: clara.esdal@hotmail.es

**ORCID**

Clara Espinosa del Alba. https://orcid.org/0000-0001-8634-5808

Andrea Mondoni. <https://orcid.org/0000-0002-4605-6304>

Eduardo Fernández-Pascual. <https://orcid.org/0000-0002-4743-9577>

Borja Jiménez-Alfaro. <https://orcid.org/0000-0001-6601-9597>

**Acknowledgements**

We are grateful to the Analysis Service Unit facilities of ICTAN for the analysis of Chromatography. We want to thank Giovanni Marchese and Francesco Cortese for their assistance in laboratory work. We also thank the Picos de Europa National for support in field sampling. This research was funded by the Spanish Research Agency (AEI/10.13039/501100011033), SEEDALP project: Seed ecological spectrum of alpine plant communities (MCI-20-PID2019-108636GA-I00); MCI-21-PRE2020-092874: Ayudas para Contratos Predoctorales para la Formación de Doctores – Formación personal investigador - FPI - del Ministerio de Ciencia e Innovación and SV-23-GIJON-JBA: Ayuntamiento de Gijón/Xixón - Jardín Botánico Atlántico 2023-2026.

**Data availability statement**

Raw data and R script for analysis are available in GitHub, once the manuscript is accepted Github repository will be public and stored in Zenodo.

**Author contributions**

Borja Jiménez-Alfaro obtained the funding. Eduardo Fernández-Pascual, Borja Jiménez-Alfaro and Clara Espinosa del Alba conceived the idea, designed the methodology and collected field sampling. Clara Espinosa del Alba and Andrea Mondoni performed laboratory experiments and analysed the data. Clara Espinosa del Alba led the writing of the manuscript. All authors contributed critically to the drafts and gave final approval for publication.

Abbreviations:

FA: Fatty Acids; UFA: Unsaturated Fatty Acids; SFA: Saturated Fatty Acids

## Abstract (299 words)

Seed oil content, one of the crucial energy reserves in angiosperms, is important for a wide range of biological processes that influence plant fitness. Thus, seed oil can be considered an adaptive trait with potential trade-offs involving other seed traits and displaying marked latitudinal patterns. Additionally, seed oil composition, namely the relative proportions of saturated and unsaturated fatty acids, is also believed to be key for micro- and macro-evolutionary biogeographical patterns. However, the availability of oil data for wild species is minimal. Few studies have addressed the potential biological and ecological correlates of seed oil content and composition.

Our study aims to explore these relationships with a unique data set of 47 different grassland alpine species from the Cantabrian Mountains (NW Spain) across regional altitudinal gradients and local microenvironmental gradients. Seed oil content and composition analyses were conducted with fresh seeds collected from the field, successfully detecting 26 fatty acids. The other measured seed traits incorporated into the study were seed mass, seed longevity, and germination timing. Species local ecological optimum along microenvironmental gradients were calculated based on 160 inventories in the study area.

Seed mass and germination timing were not significantly correlated with oil content or oil composition, but we did find strong indications that oil content and composition significantly influence seed longevity. These results are essential for a better understanding the potential impacts on soil seed banks and for adapting conservation strategies for seed banks and/or restoration programmes. Interestingly, and against expectations, we did not find patterns corroborating regional or local gradients driving changes in seed oil content or composition. The lack of significant patterns might indicate strong phylogenetic constraints on oil content. This is a meaningful advance in providing a more complete picture of the correlates of oil content and oil composition for wild species.

Keywords: Seed oil content; Seed oil composition; Germination; Longevity; Alpine species; altitudinal gradient; microenvironmental gradients.

## 1. Introduction (1297 words)

Seeds act as energy reservoirs, accumulating three main macromolecules: proteins, lipids and carbohydrates (Levin, 1974; Baud and Lepiniec, 2010). The quantity and quality of these reserves are directly related to plant fitness, as they influence germination success (Linder, 2000), seedling establishment (Graham, 2008) and seed viability (Hoekstra, 2005). In angiosperms, the predominant reserve forms are lipids and carbohydrates, the firsts majorly in the form of triglycerides (an ester of glycerol plus three fatty acids (FA) (Harwood, 1980; Voelker and Kinney, 2001), and the latter accumulated as starch (Levin, 1974). Carbons in FAs are highly reduced, and through lipid oxidation, they release more than twice as much energy as the oxidation of starch on a per g basis of dry weight (Levin, 1974; Baud and Lepiniec, 2010). Consequently, plants mainly rely on lipids (hereafter “seed oil”) for energy mobilization during germination (Harwood, 1980). Seed oil is mainly stored in the cotyledons and endosperm (Ellis, 2006) but also in the radicle and hypocotyl (Li *et al.*, 2006). Although seed oil content variation is found within and among genera of the same family (Levin, 1974; Bretagnolle *et al.*, 2016), it is also highly constrained by phylogeny and subject to evolutionary change (Levin, 1974).

In seeds, most FAs range from 10 to 22 carbons in length, and the carbons may be joined by single or double bonds, referred to as saturated (SFAs) and unsaturated fatty acids (UFAs), respectively (Ellis, 2006). The relative proportion of specific FAs and the relative proportions of UFA and SFA constitute what we will refer to as “oil composition”. The most abundant FAs found in seeds are the unsaturated oleic acid (18:1n-9), linoleic acid (18:2n-6), alpha-linolenic acid (18:3n-3), and saturated palmitic (16:0) and stearic (18:0) acids (Voelker and Kinney, 2001; Ellis, 2006; Baud and Lepiniec, 2010). SFAs and UFAs possess distinct biochemical properties (Lehninger, Nelson and Cox, 1993), influencing seed response to stimuli. SFAs have a higher melting point than UFAs (Knothe and Dunn, 2009; Sanyal and Decocq, 2016), affecting energy storage availability, especially at low temperatures. On a per-carbon basis, UFAs are more expensive to produce and yield less energy when oxidised than SFAs (Lehninger, Nelson and Cox, 1993; Linder, 2000). Moreover, storing UFAs requires antioxidant molecules to prevent damage due to the higher potential for oxidative damage (Rael et al., 2004). Hence, a maximal energy storage strategy for seeds should maximise SFA storage instead of UFA (Linder, 2000). However, the relative abundance of UFAs and SFAs varies significantly in angiosperms (Voelker and Kinney, 2001), and contrary to expectations, many species synthesise a very low amount of SFAs (Linder, 2000). Additionally, the biochemical properties of UFAs increase their unpalatability, serve as a defence against herbivory, and also function as anti-freezing compounds (Linder, 2000); see Figure 1 for a summary table.

Seed oil content and composition influence a wide range of biological processes and functions and, as such, are crucial for wild plant adaptation (Levin, 1974; Sanyal and Decocq, 2016). Consequently, selection could be acting simultaneously on multiple seed traits (Sanyal and Decocq, 2016). There is a consensus in the literature that seed mass and oil content are negatively correlated, with seed oil content contributing up to 60% of seed mass weight (Ellis, 2006; Theodoulou and Eastmond, 2012). It is also known that seed energy and oil content are positively correlated (Levin, 1974; Bretagnolle et al., 2016). Large seeds generally store less oil than small seeds; however, a tiny oily seed can release as much energy as a larger, starchy seed (Bretagnolle et al., 2016). Such a relationship suggests that oil synthesis is energetically costly and could be advantageous only for small seeds, which can store more energy in a smaller volume (Bretagnolle et al., 2016). Previous studies have also claimed that seed oil content influences seed viability and longevity (Hoekstra, 2005), meaning the ability of seeds to remain viable over specific storage periods. Secondary lipid oxidation products, needed for energy release, can cause detrimental effects on other macromolecules (Graham, 2008), likely due to the generation of ROS (Sattler et al., 2004). Remarkably, the oxidation of UFAs contributes significantly to producing free radicals and subsequent attacks (Priestley and Leopold, 1979). To prevent the deleterious effects, oily seeds also store antioxidants, and supporting findings show a positive correlation between antioxidant levels and the relative proportion of UFAs (Sattler et al., 2004). Accordingly, previous studies have found oily seeds to be more sensitive to ageing (Nagel and Börner, 2010; Neto et al., 2019), and detailed studies, including oil composition, have also shown that a higher proportion of UFAs is associated with shorter longevity (Hoekstra, 2005). Thus, not only oil content but also oil composition affects the storage behaviour of seeds (Volk et al., 2006). However, in general, the correlation between seed oil content and longevity has been described as weak (Nagel and Börner, 2010), and even no significant relationships have been reported (Walters, Wheeler and Grotenhuis, 2005; Probert, Daws and Hay, 2009; Gardarin et al., 2010). Further investigations on the effects of seed oil (content and composition) on longevity are desirable (Pritchard and Dickie, 2003) but are lacking for the vast majority of wild species.

In 1974, Levin (1974) found significant seed oil content variations depending on habitat type (oil increased with woodiness and shade tolerance) and life form (lower oil in herbs than in shrubs and trees), but no geographical differences between temperate, subtropical and tropical regions. However, a more recent study found a significant latitudinal gradient with higher seed oil content in tropical plants compared to temperate plants (Sanyal and Decocq, 2016), probably because, in the tropics, seeds need higher energy to survive the first life stages, with high competition and/or low illumination (Levin, 1974). In the same study, they also found significant oil composition gradients, with SFAs decreasing by 0.1% for each degree of latitude (Sanyal and Decocq, 2016), in concordance with previous studies (Linder, 2000). The different FAs' melting points could explain this pattern. At higher latitudes (cooler temperatures), seeds with higher UFA proportions, which have a lower melting point, can access the energy stored and germinate earlier, providing a competitive advantage (Linder, 2000). At lower latitudes (higher temperatures), seeds do not have that temperature limitation to access higher SFA energy resources with higher melting points (Sanyal and Decocq, 2016). Higher growing temperatures have also been correlated with higher oil content in several cultivars (Piper and Boote, 1999) and higher SFA accumulation (Sanyal and Linder, 2013). Consequently, seed oil content and composition seem to be under strong selection by temperature at both micro- and macro-evolutionary levels (Sanyal and Linder, 2013; Sanyal and Decocq, 2016).

Imagen que contiene Diagrama

Descripción generada automáticamente

Figure 1. A) Latitudinal oil content and UFA/SFA ratio patterns described in the literature by Linder (2000) and Sanyal and Decocq (2016) derived from temperature and energy constraints. World map image by Jose Carlos García Lopez licensed under CC BY-SA 3.0. B) Altitudinal oil content and UFA/SFA ratio patterns expected in our study. C) Fatty acids biochemical properties.

The study aims to explore the patterns of seed oil content and composition in alpine plants and understand their biological and ecological correlates in an alpine environment. This kind of data is barely available for alpine species and, to our knowledge, has never been researched. The temperate alpine environment is characterised by extreme cold conditions and strong microenvironmental gradients (Scherrer and Körner, 2011) that physiologically limit plant regeneration (Körner, 2021). Alpine species have evolved under these circumstances and thus adapted, including the oil content and composition, to maximise the chances of successful regeneration. Alpine species are generally known for having average-sized seeds (Körner, 2021), delayed germination phenology due to physiological dormancy and cold temperatures (Fernández-Pascual et al., 2021; Espinosa del Alba, Fernández-Pascual and Jiménez-Alfaro, 2024), and short longevity (Mondoni *et al.*, 2011). We expect seed mass, longevity, and earliness of germination to correlate with oil content and composition; see Table 1 for detailed expectations. We also expect to find regional and local oil patterns related to temperature along an altitudinal gradient and in concordance with the results previously reported in the latitudinal gradients (Sanyal and Linder, 2013; Sanyal and Decocq, 2016); see Figure 1 and Table 1 for detailed expectations.

Table 1. Study expectations relationships between seed oil content and composition with biological and ecological correlates.

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | **Oil content** | **Ratio UFA/SFA** |
| Biological correlates | |  |  |
|  | Seed mass | Negative | Negative |
| Seed longevity | Negative | Negative |
| Earliness of germination | Positive | Positive |
| Ecological correlates | |  |  |
| Regional altitudinal gradient | | Lower in strict alpine | Not tested |
| Local ecological optimum | |  |  |
|  | Growing Degree Days | Positive | Negative |
|  | Freezing Degree Days | Negative | Positive |
|  | Snow Days | Positive | Positive |

## 2. Materials and Methodology (1626 words)

### Study system

We focused on herbaceous alpine species from grassland communities in the Cantabrian Mountains (north-western Spain), a transitional mountain hub between Eurosiberian and Mediterranean regions in southern Europe (García-Gutiérrez *et al.*, 2018). These grassland communities are continuously distributed along the mountain range, occupying reduced areas above the treeline and around mountaintops between 1750 and 2500 m a.s.l. Grassland communities are predominantly dominated by Poaceae and Cyperaceae, and the main life forms are Hemicryptophytes and Chamaephytes. Climatic conditions follow a north-south temperature and precipitation gradient, with colder and wetter conditions on northern slopes compared to warmer and drier conditions on southern slopes. Mean annual temperatures range from 2.5 to 4.5 ºC and mean summer precipitation range from 160 to 260mm (Values extracted from Chelsa 2.1 bio1 and bio17, Karger *et al.*, 2017).

### Species data

We established eight sampling sites above 1900 m a.s.l., four on the northern slopes and four on the southern slopes of the Cantabrian Mountains, ensuring representation of the flora from the two biogeographical regions in the area (Eurosiberian and Mediterranean) (García-Gutiérrez et al., 2018). We collected floristic and community composition data for all vascular plants at each sampling site. We collected enough seeds to measure and analyse seed oil content and composition from 47 species (200 mg of dry seeds) out of the 119 initially recorded species; the FA data obtained were unknown or unpublished. The collected species were classified as strict alpine or generalist according to their altitudinal distribution. We identified plant specialists as those significantly associated with alpine grasslands, using the Indicator Values (IndVal) in the indicspecies R package (De Cáceres and Legendre, 2009). The calculations were based on 12,000 vegetation plots of grasslands stored in the SIVIM database for the Cantabrian Mixed Forests ecoregion (paper classification). From the preliminary list of indicator species for the studied vegetation, we removed species with median elevation values below 1800 m a.s.l., most of which are characteristic of subalpine or nitrophilous habitats, and those were specified as generalists. Seed collection permits were granted for the sampling sites inside Picos de Europa National Park: expedient numbers CO/09/203/2021 and CO/09/192/2023.

### Oil content and composition

We preliminarily explored oil content variation within seed lots in five species, analysing three subsamples from each (also to ensure high precision of the methodology), and we could not detect any statistically significant differences within the seed lots. Consequently, a single sample for each species was analysed; thus, within-species seed oil variation was not explored. We sent the samples to an external analytical laboratory (USTA-CSIC), where they used a gas chromatographer with a Flame Ionization Detector (Agilent 7820A, EZChrom Elite software). Samples were manually ground with liquid nitrogen; then, fatty acids were transformed into fatty acid methyl esters (FAMEs) following the procedure described by Lee et al. (2012) section 2.2.4 (supplementary Table S1). As a result, we obtained the absolute value of the oil content on a dry-weight basis and the percentage of each FAME type relative to the total oil content. The analysis identified 26 different FAME types in the studied species (see supplementary Table S2). Additionally, we retrieved oil and seed mass information for another 33 species native to Europe, congeneric with the species we collected, from the SID database (2023) and classified as either generalist or strict lowland. Our local species oil data, in addition to the data obtained from the SID database (2023), were used for regional altitudinal gradients, containing oil content and seed mass data for 80 species (29 strict alpines, 31 generalists and 20 strict lowlands) from 19 different plant families. We used only our local species oil data for local scale analysis, containing seed oil content and oil composition data for 47 species from 19 families.

### Seed traits data

Seed mass was calculated for all local species, averaging the weight of five replicates of 50 dry seeds. To estimate the earliness of germination, we used the T50 trait, calculated as the time (in days) to reach 50% germination, obtained from a published phenology germination experiment with species from the same study area (Espinosa del Alba, Fernández-Pascual and Jiménez-Alfaro, 2024) and is available for 36 species. Longevity data was experimentally calculated for 35 species in the laboratory. We applied a standard comparative longevity protocol (Probert, Daws and Hay, 2009; Davies *et al.*, 2016) optimised for short-lived species (Davies *et al.*, 2016) and successfully applied it to alpine species by (Mondoni et al., 2011), adjusting the number of seeds to match availability. The artificially accelerated ageing protocol allows the measurement of the p50 value (amount of time for seed viability to drop to 50%). The initial RH value was 30-35% (Hygropalm 3 display unit; Rotronic Instrument UK Ltd, Crawley,UK). Before the start of the ageing protocol, all species were tested to ensure germination above 85% and species from the Cistaceae family were physically scarified with sandpaper. Seed samples (200 seeds/ species) were first rehydrated to 47% RH at 20ºC for five days in a non-saturated LiCl solution (Hay *et al.*, 2008) in crystal Petri dishes and kept in a 300 x 300 x 130 mm sealed electric enclosure box (Ensto UK Ltd, Southampton, UK) before moving the Petri dishes into the ageing conditions. The ageing conditions comprised a temperature of 45°C, 60% RH and darkness. A subsample of 42 seeds was withdrawn after 2, 10, 15 and 30 days in ageing conditions and consequently sowed in Petri dishes 1% agar with GA3 following Kew standard procedure (Davies, Sacco and Newton, 2015). After the ageing period, those species requiring a cold stratification period (Saxifraga oppositifolia) were put in a refrigerator at 5ºC and in darkness for a month before germination experiments. After ageing, germination conditions for the species were set with a 12/12 h photoperiod and two alternating temperatures: warmer conditions at 22-12ºC or 15/5ºC for those germinations which required colder temperatures (see supplementary Table S3). After sowing, the seeds were checked once a week for four weeks. Germination was scored when there was a visible root (>1.5 mm) and then removed from the Petri dishes. After the end of the 28-day germination test, the ungerminated seeds were cut-tested under the binocular stereoscope to visually assess the embryo's state. We considered seeds with white and firm embryos viable, i.e. potentially germinable (Baskin and Baskin, 2014) and removed empty or infected seeds from further analysis.

We applied the viability equations developed by seed bank managers

*v = Ki – (p/σ)*

where v is viability in NED (Normal equivalent deviates), p is the days of ageing, Ki is the initial viability, and σ is the standard deviation of the distribution of deaths over time. Using GENSTAT software (REF), which applies this equation, we calculated the p50, i.e. the time for viability to drop below 50%.

### Local ecological optimum

To calculate local species' ecological optimum, we measured the microenvironmental gradients, establishing 20 additional plots (1 m²) for each sampling site, five in each cardinal direction with a 10 m separation (following the methodology of Jiménez-Alfaro et al., 2024). In total, we established 160 vegetation plots (8 sites × 20 plots), where we also buried an iButton datalogger, at a depth of 5 cm, in each vegetation plot (Thermochron, iButton, Newbury, UK; accuracy: ± 0.5 °C from -10 °C to +65 °C, resolution: 0.5 °C, records every four hours), which recorded temperatures over 11 months. From the microenvironmental data recorded, we calculated three bioclimatic indices per plot following Jiménez-Alfaro et al. (2024). (1) Growing Degree Days (GDD) as the sum of daily mean temperatures for days in which the soil mean temperature at five cm was above 5 ºC (Körner, 2021); (2) Freezing Degree Days (FDD) as the sum of daily mean temperatures for days in which the mean temperature was below 0 ºC (Choler, 2018); and (3) days with snow cover, based on the period in which the maximum temperatures were < 0.5 ºC and the minimum temperatures were > -0.5 ºC (Zhang et al., 2005). Then, we used the bioclimatic indices to calculate the species' local ecological optimum, averaging the climatic variables of the plots where the species was present and weighting them by their coverage (only considering those plots where the species had more than 10% relative coverage), assuming species would have more coverage in those plots with climatic conditions closer to their local ecological optimum.

### Statistical analysis

We use GENSTAT software to apply a Probit analysis (Hay, Mead and Bloomberg, 2014) and calculate p50 values.

We analyzed the data by fitting Markov Chain Monte Carlo generalized linear mixed models (MCMCglmm) with Bayesian estimation using the R package MCMCGLMM (Hadfield, 2010). To model raw germination scores from the longevity experiment, we used binomial MCMCglmms (family = multinomial2), while for the rest of the traits, we scaled the values and used Gaussian MCMCglmms (family = Gaussian). For the seed oil-seed mass relationship, total oil content (in percentage) and UFA/SFA ratio (both log-transformed) were the response variables, and the explanatory variables were seed mass (log-transformed). For the other seed traits, p50 (square-root transformation) and T50 were set as response variables, while oil content and UFA/SFA ratio were used as explanatory variables (log-transformed). To test regional and local environmental gradients, total oil content (in percentage) and UFA/SFA ratio (both log-transformed) were the response variables and their altitudinal distribution (strict alpine, generalist or strict lowland) as well as their local ecological optimum (GDD, FDD and Snow). We used weakly informative priors in all models, with parameter-expanded priors for the random effects. Each model was run for 1,000,000 iterations, with an initial burn of 100,000 and a thinning interval of 100. We calculated mean parameter estimates and 95% credible intervals (CI) from the resulting posterior distributions. Phylogeny was included using a reconstructed tree (Supporting information Figure S1 and Figure S2 for regional and local species pool, respectively), created with V.PHYLOMAKER R package (Jin and Qian, 2019) using phylosignal and phylobase R packages (Keck *et al.*, 2016; R Hackathon et al., 2020).

## 3.Results (1066 words)

### Regional altitudinal patterns

We found no significant oil content differences between strict alpine, generalist and strict lowland species (Figure 2A) from our regional dataset (n=80); model details can be checked in supplementary Table S4. Oil content values ranged from 1.19% to 38%, with a mean value of 11.5%. Additionally, we found no significant correlation between oil content (%) and seed mass (mg) (posterior mean=-0.064, pMCMC=0.29). We observed a considerable variation of seed oil content in smaller seeds, while variation was reduced with increasing seed mass (Figure 2B). Seed mass values ranged from 0.64 mg to 335 mg, with a mean value of 51 mg.

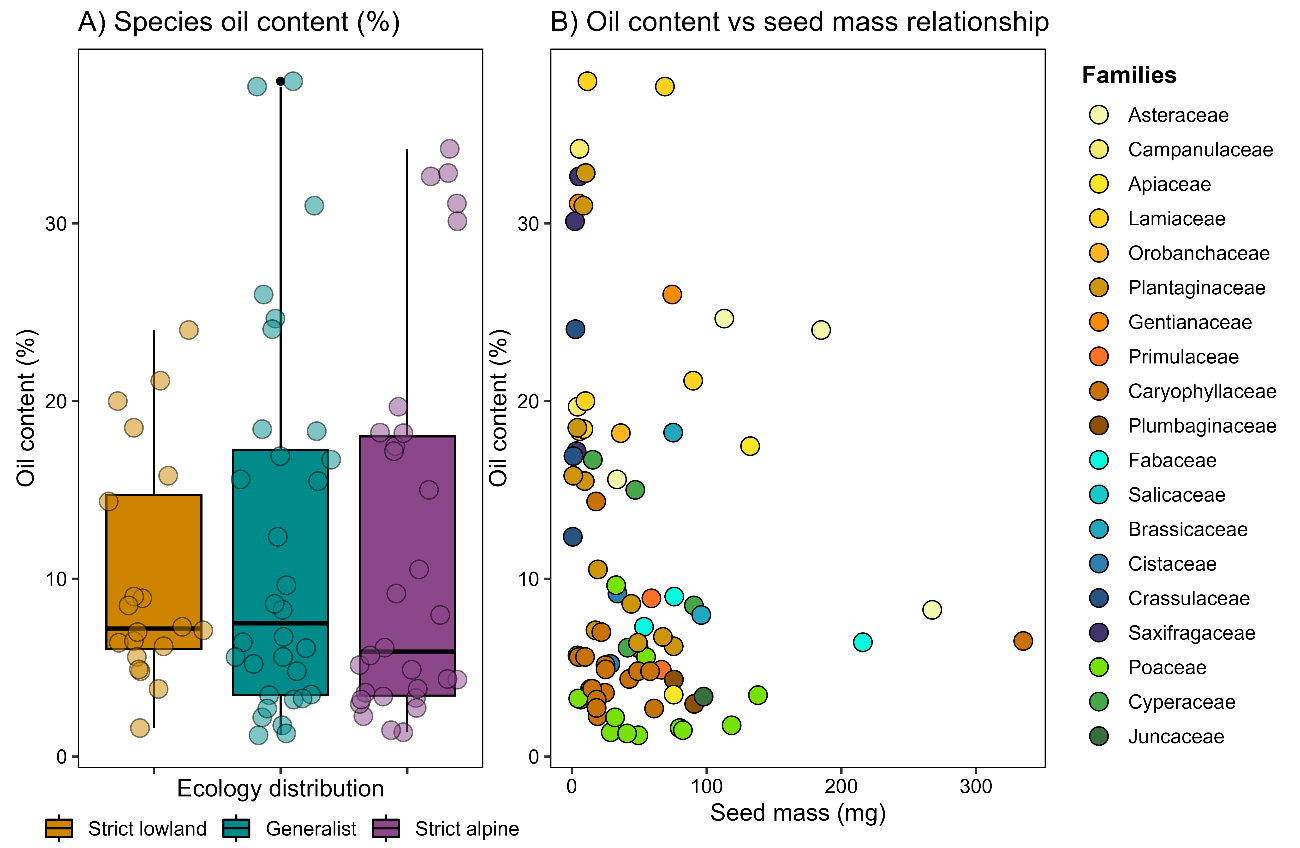


Figure 2. Regional oil content patterns (n=80 species). A) Seed oil content (%) between regional altitudinal distribution showed by different colours. B) Seed oil content (%) and seed mass (mg) non-significant relationship (MCMC-GLMMs), colours by families.

### Alpine seed oil content and oil composition patterns

Within our local species pool (n=47), oil content varies from 1.2% in Avenella flexuosa (Poaceae) to 34.2% in Jasione cavanillesii (Campanulaceae) (Figure 3A). The most abundant FAs in alpine seeds were the unsaturated linoleic acid (C18:2n-6c, 42.4%±2.4, mean ± se), oleic acid (C18:1n-9c, 22.6%±1.9), and alpha-linolenic acid (C18:3n3, 15.1%±2.7); as well as saturated palmitic (C16:0, 10.8%±0.8) (details in supplementary Table S2). These four average 94.8% of seed oil content (se=23). Additionally, in our data set, erucic acid (C22:1n9) also had high values only in Brassicaceae species and gamma-linolenic acid (C18:3n6) in some Caryophyllaceae species (Figure 3B). The mean frequency of SFA is 14.2%±0.95, and the mean ratio between UFA and SFA is 7.2 (se=0.5), ranging from 1.3 in A. flexuosa up to 20.1 in Iberis carnosa (Brassicaceae). The Poaceae, Salicaceae, and Cistaceae families generally had low UFA/SFA ratio values, indicating higher SFAs synthesis. At the same time, the Brassicaceae and Lamiaceae families were characterised by higher ratio values, showing a tendency to synthesise more UFAs (Figure 3C).

We restricted the exploratory PCA to those FAMEs with >3% relative proportion and correlations below 0.7 to reduce dimensionality. The multivariate analysis showed relatively low explained variation within the first two axes (26.1% and 16.7%, respectively). The UFA/SFA ratio and C16:0 FAMEs contributed the most in PC1 (26.4% and 18.6%, respectively), while C20:1n9 and C22:1n9 contributed the most in PC2 (21.9% and 16.3%, respectively) (Figures 3D and 3E). Additionally, PCA revealed that the oil content percentage is not highly correlated with any specific FA type (details in supplementary Table S5).

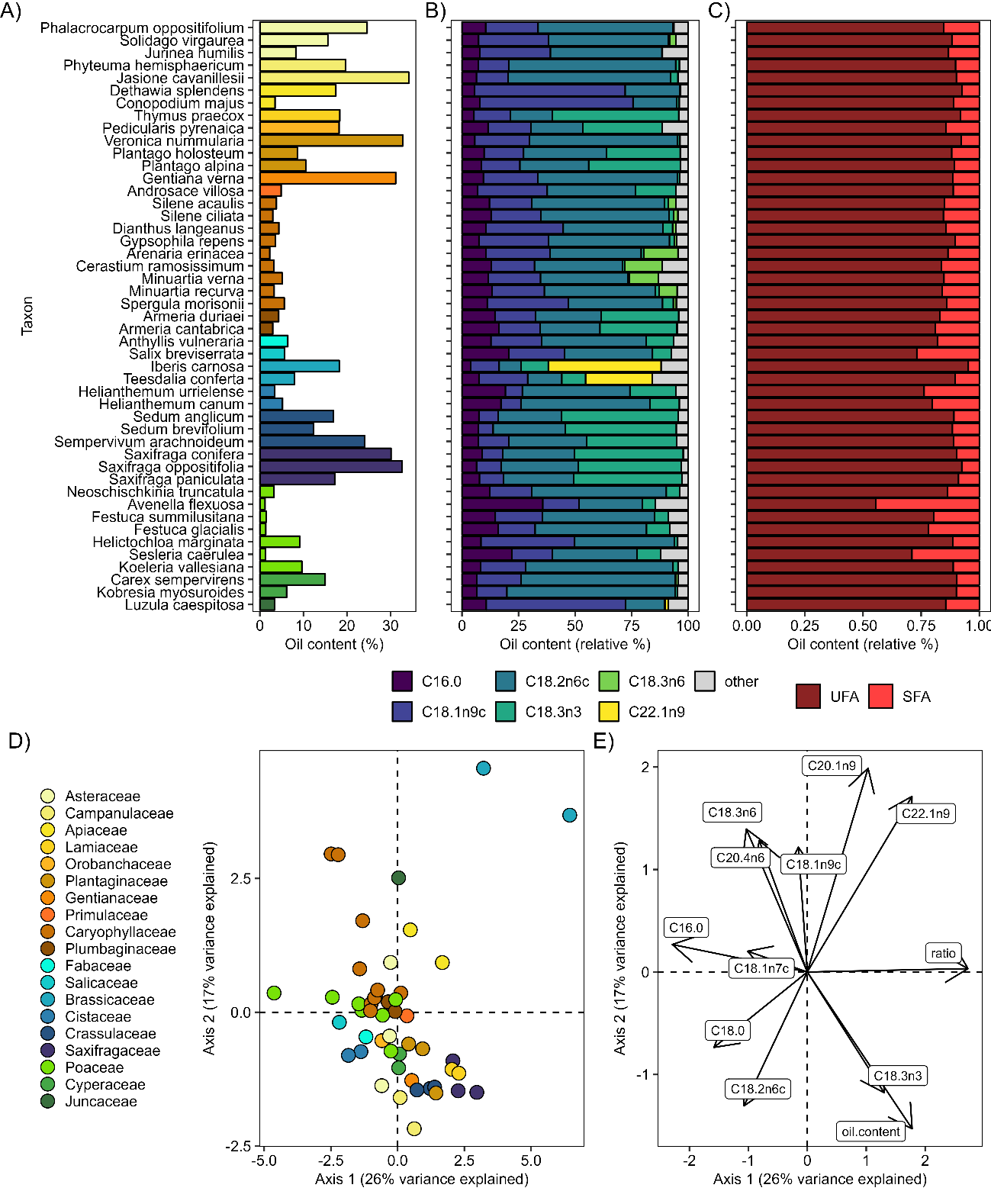


Figure 3. Local species oil content and composition data exploration (n=47). A) Oil content (%) per species, colours representing families as panel D. B) Seed oil composition of FAMEs with more than 10% relative proportion on average across species. C) Seed oil composition per species with FAME’s divided between Unsaturated Fatty Acids (UFA) and Saturated Fatty Acids (SFA). D) Exploratory PCA with species as points and E) variables directions and contributions.

### Biological correlates

#### Seed mass

Seed mass values ranged from 0.6 mg (Sedum brevifolium) to 268 mg (Jurinea humilis), with a mean of 45.4 mg (se=7.84). We observed a notable variation in oil content in small seeds, but its variability decreased in large seeds, which had lower oil content. Although a general trend was visible, the relationship between seed mass and oil content was not significant (posterior mean=-0.15, pMCMC=0.16, Figure 4A, left panel). Oil composition, i.e. UFA/SFA ratio, was also not significantly correlated with seed mass (posterior mean=-0.03, pMCMC=0.5, Figure 4A right panel).

#### Seed longevity

Seed longevity was analysed using two different approximations. First, the raw germination data from the artificial ageing protocol (with Multinomial MCMC-GLMM) and the p50 value (obtained from the probit analysis and Gaussian MCMC-GLMM). The p50 values ranged from 3 to 47 days, with a mean of 21 days (se=2). Results were consistent using both approximations, although raw germination scores showed higher statistical power. Higher oil content significantly reduced seed longevity (germination curves posterior mean=-0.47, pMCMC<0.001; p50 posterior mean=-0.83, pMCMC=0.019 Figure 4B left panel). The relationship for the ratio was only marginally significant; seeds with a higher UFA/SFA ratio (i.e. more unsaturated fatty acids) showed a decrease in longevity (germination curves posterior mean=-0.28, pMCMC=0.055; p50 posterior mean=-0.84, pMCMC=0.35 fig 4B right panel).

#### Earliness of germination

T50 values ranged from 4 to 295 days, with a mean of 150 days (se=16.6). We did not find any significant relationship between T50 and oil content (posterior mean=31.7, pMCMC=0.13, Figure 4C left panel) or the UFA/SFA ratio (posterior mean=79, pMCMC=0.1, Figure 4C right panel). Despite the non-significance, we observed a trend of seeds with higher oil content and more UFA experiencing delayed germination.

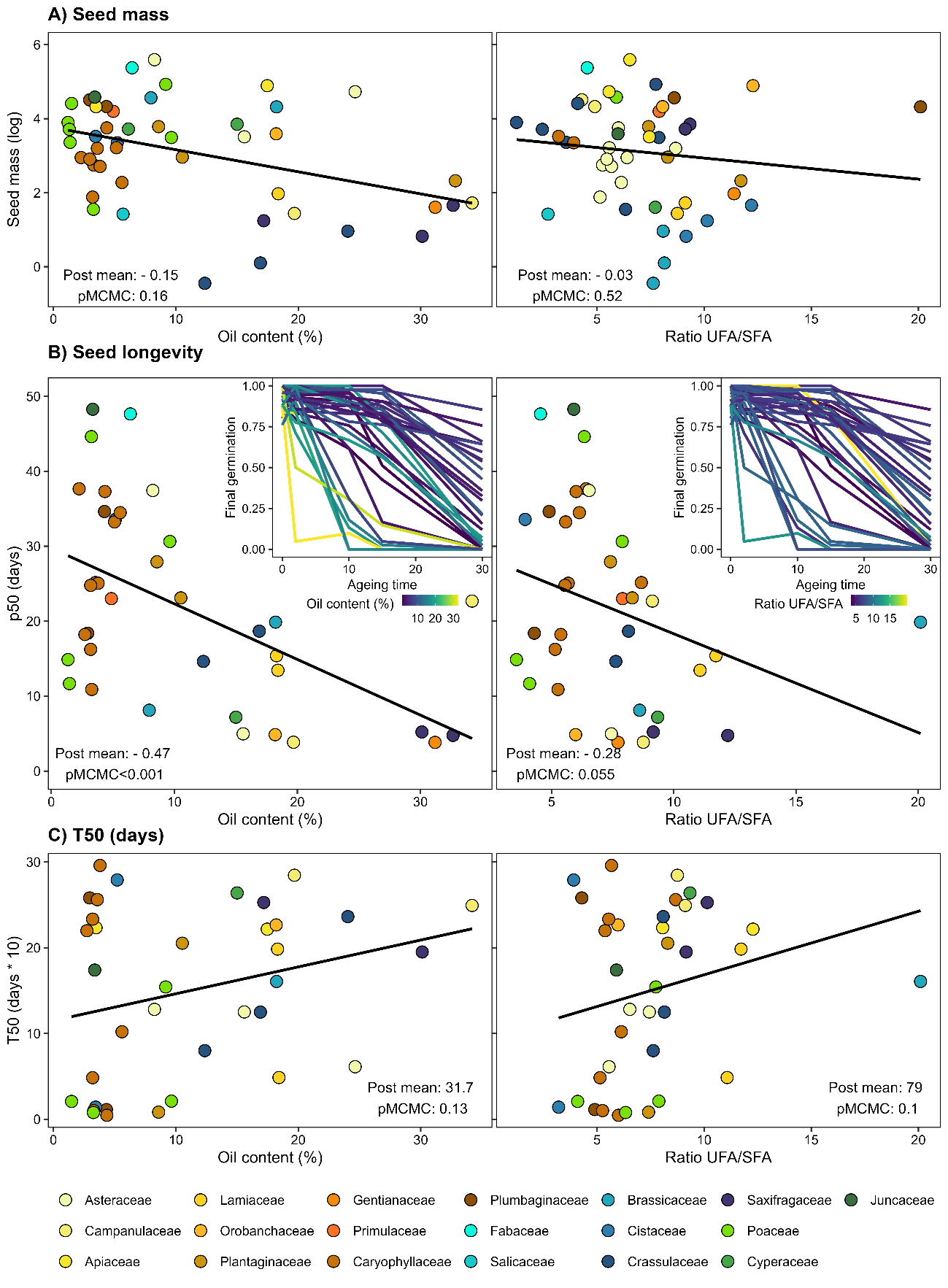


Figure 4. Seed oil content and ratio UFA/SFA biological trade-offs. A) Correlation between seed mass, oil content, and ratio UFA/SFA, significances from MCMC-GLMM models (Gaussian family, n=47) with all variables log-transformed. B) P50: time for viability to drop to 50%, based on probit analysis (GENSTAT software), and the plot inset showing raw germination curves across an artificially accelerated ageing protocol in the lab. The colour represents the oil content percentage (left panel) and UFA/SFA ratio (right panel). Significances from MCMC-GLMM (multinomial family, n=35) with oil content and ratio log-transformed. C) Germination timing, T50: time to reach 50% germination, depending on oil content and ratio UFA/SFA. Significances from MCMC-GLMM (n=36), both explanatory variables log-transformed. Colours represent the different families.

### Local ecological optimums

We found no significant correlations between species' local optimal ecological conditions (GDD, FDD, and snow days; see Figure 5) and oil content or UFA/SFA ratio. GDD values ranged from 650 to 2295 °C, averaging 1421 °C (se = 67.9). Data showed a negative trend, with species in warmer conditions displaying less oil content and a lower ratio. FDD values ranged from 0.05 to 170 °C, with a mean of 31.8 °C (se = 4.7). The data indicated a positive trend but was likely driven by a few species with more extreme oil content and composition values. Snow values ranged from 3 to 157 days, with an average of 62 days with snow; the data demonstrated a positive trend where species with a preference for snowier sites showed higher oil content and UFA/SFA ratios. The details of the model results can be found in supplementary Table S4.

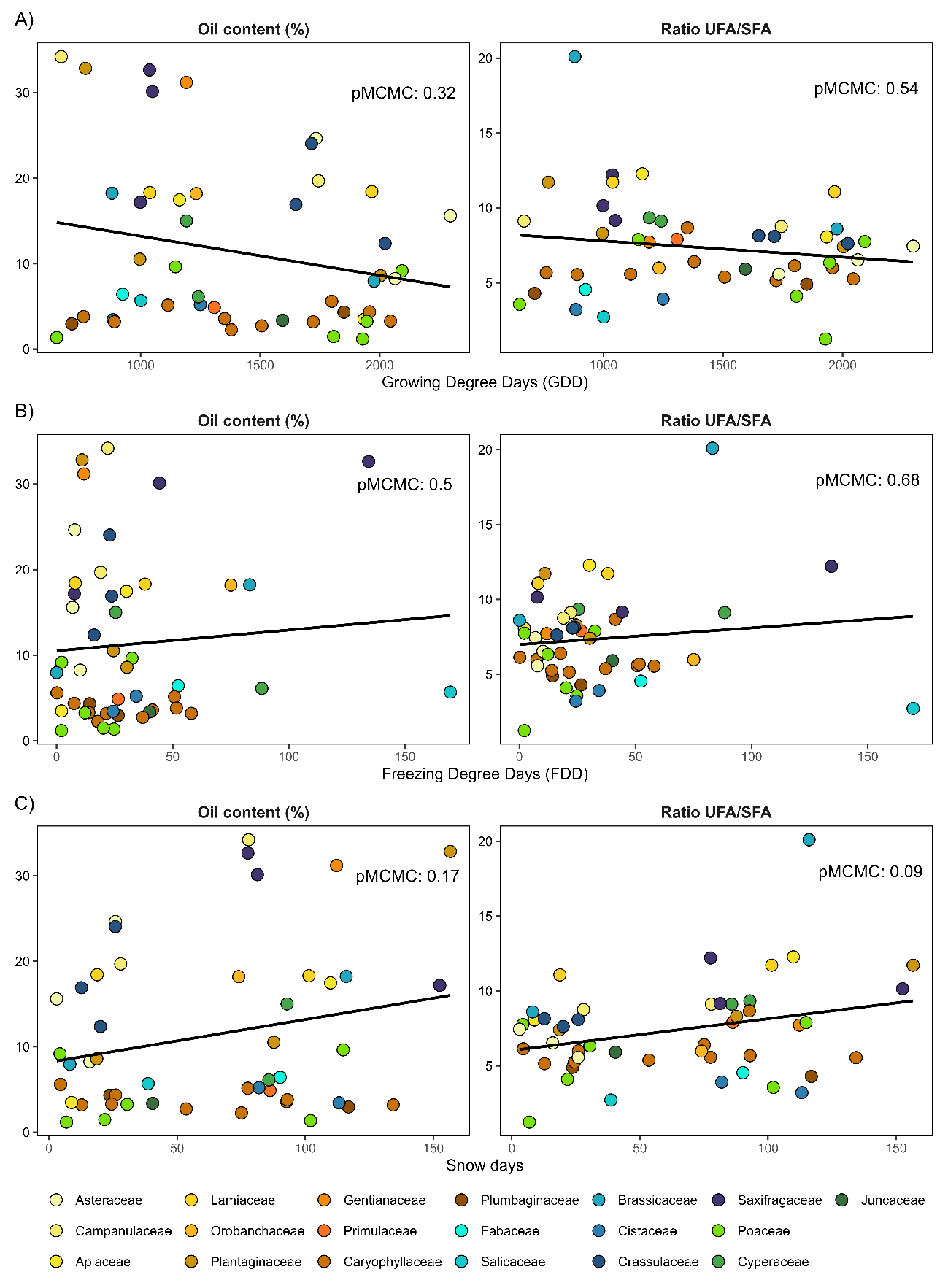


Fig 5. Seed oil content ecological trade-offs, oil content and ratio UFA/SFA log-transformed for MCMC-GLMM models (family=Gaussian, n=47). A) Relationship with GDD (Growing Degree Days °C); B) Relationship with FDD (Freezing Degree Days °C) and C) Relationship with Snow days.

## 4. Discussion (1269 words)

Our study aims to explore biological and ecological correlates of seed oil content and seed oil composition with a unique data set of 47 alpine species. Interestingly, we did not find temperature-related patterns corroborating macro- or microevolutionary processes driving oil content and composition at the regional or local scale. We did find strong indications that, in alpine species, seed oil content and composition significantly influence seed longevity, potentially impacting soil seed banks and conservation strategies for seed bank managers or restoration programmes.

Within our data set, we found no high correlations (>0.7) between total oil content and other specific FAs, in contrast to Sanyal and Decocq (2016) who found several significant negative correlations between specific FAs. While oil content is known to vary from 1% in Musa paradisiaca to 76% in Chrysobalanus icaco (Matthäus, 2012), our alpine species show a more constricted range of oil values (1-38%), which is similar to the one exhibited by weed species in France (Bretagnolle et al., 2016). However, it remains in the lower range compared to global records. Lower oil contents were expected due to the herbaceous nature of all the species included in the study (Levin, 1974). The oil composition was also comparable to what Bretagnolle et al. (2016) reported, except for erucic acid (C22:1n9) in the Brassicaceae family. Previous reports from other commercial Brassicaceae species confirmed high proportions of elongated acyl chains from C20 to C24 (Baud and Lepiniec, 2010). The high levels of erucic acid are surprising due to its “high” melting point (33.5°C) (Sanyal and Linder, 2013), which would hinder energy release; however, erucic acid has also been correlated with lower seed predation (Jong et al., 2016).

We could not corroborate the negative correlation between seed mass and seed oil content previously described in the literature (Nagel and Börner, 2010; Hamilton et al., 2013; Bretagnolle et al., 2016). A plausible explanation for the lack of significance is that the studies reporting a significant relationship include different vegetation types, namely herbs, shrubs, and trees. In contrast, our study focused solely on the herbaceous type. The same non-significant trend has been reported in another study that concentrated on trees (Finkelstein and Grubb, 2002). Nevertheless, more data and within vegetation-type analyses are necessary to confirm this idea. Additionally, in our local alpine species pool, we found a negative trend, less pronounced, between seed mass and UFA/SFA ratio. However, the trend contradicts the expectation of small seeds storing more SFA, i.e. lower ratio values (Linder, 2000). We found a considerable variation of oil content in small, although such variation strongly decreases towards low oil values in large seeds, corroborating the results reported by Bretagnolle *et al.* (2016). Nevertheless, there appears to be a strong phylogenetic trend (Levin, 1974), with species from the same family mostly grouped together.

In concordance with our expectations, we found a significant negative relationship between seed longevity and seed oil content and composition. Alpine seeds, generally known for their short longevity (Mondoni et al., 2011), showed a consistent decrease in seed longevity with increasing oil content and UFA/SFA ratio, indicating more UFAs. The results corroborate the hypothesis that oily seeds have lower longevity (Nagel and Börner, 2010; Neto et al., 2019) and that a higher proportion of UFAs is associated with shorter longevity (Hoekstra, 2005). Our results are consistent with previously described adverse effects observed in seed longevity, possibly due to lipid oxidation (Bailly et al., 1998; Tammela et al., 2003) and contrast with other studies reporting no relationship between seed longevity and oil content (Walters, Wheeler and Grotenhuis, 2005; Probert, Daws and Hay, 2009; Gardarin et al., 2010). With the current biodiversity crisis, plant conservation is a global priority, especially ex-situ conservation in seed banks. Having the necessary knowledge and appropriate methods to store and maintain seed viability in these facilities is vital. The results presented here suggest that, for alpine species, seed oil content can be an important factor to consider. Measuring seed oil content and composition is a destructive yet fast method that can be used to separate those alpine species with particularly low seed longevity. The data compiled in this study will be available in the GitHub repository and accessible to all.

Based on the results found by Sanyal and Decocq (2016) and Linder (2000), we predicted that alpine species, adapted to live and germinate under colder temperatures, would show earlier germination with increasing oil content and an increasing UFA/SFA ratio. Surprisingly, the negative relationship between oil content and the earliness of germination contradicted our initial expectations. The lack of patterns with oil content is not completely surprising in wild species, as has been reported in rainforest species in Australia (Hamilton et al., 2013), whereas weeds (Gardarin, Dürr, and Colbach, 2011) and crops (Gu et al., 2019; Cheng et al., 2024) have shown that positive expected correlation. Observing this pattern between wild and cultivated species, one could say that crop selection has effectively favoured varieties exhibiting both high oil content and early germination; however, there is an important oil data gap in wild species. We also observed the opposite expected trend with species with higher UFA germinating later against Linder (2000). A plausible explanation could be that alpine species' physiological seed dormancy constraints (Schwienbacher et al., 2011) drive germination patterns more strongly than oil content and composition.

Against our expectations, we found no significant differences along the altitudinal gradient at the regional scale and did not find any significant relationship with the species' local ecological optimum. Previous research in a few cultivars also failed to detect a significant positive correlation between seed oil and elevation (Dierig *et al.*, 2006; Cai *et al.*, 2012). These results might indicate a low effect of elevation in seed oil content (Dierig *et al.*, 2006), although there are reports of a positive correlation between oil content and mean temperatures (Piper and Boote, 1999). Parallelly, Cai *et al.* (2012) did find a positive correlation between elevation and percentage of UFA within a single species grown in a broad altitudinal range. None of the local microenvironmental gradients studied (GDD, FDD and snow) showed any significant relationship with oil content or composition. However, snow days demonstrated a more pronounced relationship with the UFA/SFA ratio, indicating that species living in more snowy sites tended to have higher UFA, which could potentially be consistent with the anti-freezing properties that UFA can provide (Linder, 2000). The lack of regional and local patterns seems to limit Sanyal and Linder's claims (Sanyal and Linder, 2013; Sanyal and Decocq, 2016) of strong temperature selection at macro and microevolutionary levels constraining seed oil patterns. These results might indicate a strong phylogenetic signal for oil content. However, providing a more complete picture is difficult due to the limited information on oil content and fatty acid composition available for wild species (Levin, 1974).

Although the patterns presented in the study provide new insights into the understudied topic of seed oil and composition and their trade-offs in wild species, we must assume some caveats and limitations in our study. For the complete seed oil and seed traits dataset, analyses and measurements were done in two or three batches, corresponding with the seed availability after each year's collection campaign. Nevertheless, we are confident that we have collected fully mature seeds, following the ENSCONET protocol (2009), and have applied a consistent methodology for all experiments and analyses.