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

Authors: C. Espinosa del Alba1, A. Mondoni2, E. Fernández-Pascual1 & B. Jiménez-Alfaro1.

Affiliations:

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

2- 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 also 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 (307 words)

Seed oil content, one of the crucial energy reserves in angiosperms, is important for a wide arrange of biological processes that in turn influence plant fitness. Thus, seed oil is considered to be an adaptive trait with potential trade-offs with other seed traits and marked latitudinal patterns. Additionally, seed oil composition, namely the relative proportions of saturated and unsaturated fatty acids, is also supposed to be key for micro and macroevolutionary biogeographical patterns. However, oil data availability for wild species is very limited and few studies have attempted to address the potential biological and ecological correlates of seed oil and seed oil composition across environmental gradients.

Our study aims to explore these relationships with a unique data set of 47 different alpine species from the Cantabrian Mountains (NW Spain) across regional and local environmental gradients. Seed oil content and composition analyses were done with fresh seeds collected from the field, being able to distinguish up to 26 different fatty acids. The other measured seed traits incorporated into the study were seed mass, seed longevity and germination timing. Regional and local environmental 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 of oil content and composition significantly influencing seed longevity. These results are important to better understand the potential impacts on soil seed banks and adapt conservation strategies for seed banks and/or restoration programs. Interestingly, and against expectations, we did not find patterns corroborating regional or local environmental gradients driving seed oil content or oil composition. The lack of significant patterns might be an indication of a strong phylogenetic signal for oil content. This is a significant advance to provide a more complete picture on oil content and fatty acid composition correlates for wild species.

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

## 1. Introduction (1552 words)

Seeds act as reservoirs of energy 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 (Bewley *et al.*, 2013) 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) (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 or proteins on a per g basis of dry weight (Levin, 1974; Baud and Lepiniec, 2010). Consequently, plants mostly rely on lipids (hereafter “seed oil” for simplicity) stored in the seed for energy in the first life stages [8]. 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 (Ellis, 2006). FAs with one or more double bonds are referred to as unsaturated (UFAs) and FAs without double bonds are referred to as saturated (SFAs) (Ellis, 2006). The relative proportion of specific FAs as well as the relative proportion of UFA and SFA is what we will refer to as “oil composition”. The most abundant FAs found in seeds are the unsaturated oleic acid (18:1n-9, OLA), linoleic acid (18:2n-6, LA) and a-linolenic acid (18:3n3, ALA) and the saturated palmitic (16:0, PA) and stearic (18:0, SA) acids (Voelker and Kinney, 2001; Ellis, 2006; Baud and Lepiniec, 2010). SFAs and UFAs have differential biochemical properties (Lehninger, Nelson and Cox, 1993) that influence the differential seed responses to stimuli. SFAs have a higher melting point than UFAs, (Knothe and Dunn 2009 read) (Sanyal and Decocq, 2016) which influences the availability of the energy stored. On a per-carbon basis, UFAs cost more to produce and yield less energy when oxidized than SFAs (Lehninger, Nelson and Cox, 1993; Linder, 2000). Moreover, the storage of UFAs needs the storage of antioxidant molecules to prevent damage due to the higher potential for oxidative damage (Benson 1990, check sattler 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 highly varies in angiosperms (Voelker and Kinney, 2001) and, contrary to the expectation, many species synthesise a very low amount of SFAs (Linder, 2000). UFAs' biochemical properties increase the unpalatability, serving as a defence against herbivory, and also work as anti-freezing compounds (Linder, 2000).

Seed oil content and composition influence a wide range of biological processes and functions, and as such is crucial for wild plant adaptation (Levin, 1974). Consequently, selection could be acting simultaneously on multiple seed traits. Nevertheless, and despite its potential importance, few studies addressed these questions with the same species pool. 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 (Thedolou et al 2012 add ref)(Ellis, 2006). 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 small oily seed can release as much energy as a starchy seed that is twice as heavy (Bretagnolle *et al.*, 2016). Such a relationship suggests that oil synthesis is energetically costly and could be an advantage 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 (Hoekstra, 2005) and longevity, meaning the ability of seeds to remain viable over certain storage periods. For energy release, lipids need to be oxidated, which in turn generates free radicals and reactive oxygen species (ROS) (Bailly, 2004 read), that can cause detrimental effects on membrane integrity (Priestley and Leopold, 1979 read) (Kranner et al., 2002 read). Particularly, the oxidation of UFAs is highly contributing to free radicals’ production and subsequent attacks on other macromolecules (Benson 1990). To prevent the deleterious effect of lipid oxidation, 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 read; Falk & Munn\_e-Bosch, 2010 read). 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 are associated with shorter longevity (Hoekstra, 2005). Thus, not only oil content but also oil composition affects the storage behaviour of seeds (Walters et al. 2004 read; Volk et al. 2006 read; Walters et al. 2015 read). However, in general, the correlation between seed oil content and longevity has been described as weak (Nagel and Börner, 2010) Priestley et al., 1985 read; Walters et al., 2005 read) and even no significant relationships have been reported (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 lacking for the vast majority of wild species. ADD about germination speed here??

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 that seed oil content to be significantly higher in tropical plants compared to temperate plants (Sanyal and Decocq, 2016), probably because 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 higher proportions of UFAs (oleic and eicosenoic) with increasing latitude (Sanyal and Decocq, 2016) in concordance with previous studies (Linder, 2000; Sanyal and Linder, 2013), with the relative proportions of SFAs decreasing by 0.1% for each degree of latitude (Sanyal and Linder, 2013). Additionally, due to UFAs and SFAs different melting points, at lower latitudes with higher temperatures, seeds with higher proportions of SFA would be favoured because they would have more energy for growth without delaying or slowing germination (Sanyal and Decocq, 2016). At higher latitudes, and thus cooler temperatures, seeds that have a higher proportion of UFAs may germinate earlier and/or more (Linder, 2000). Rich UFA seeds could then germinate faster and earlier than rich SFA seeds in colder conditions, providing a competitive advantage where cold temperature regulates seed germination (Linder, 2000). Consequently, seed oil content and oil composition seem to be under strong selection by temperature at both micro- and macro-evolutionary levels (Sanyal and Linder, 2013; Sanyal and Decocq, 2016).

The goal of the study is 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 alpine environment is characterised by harsh environmental conditions and strong microclimatic gradients even at short distances (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 for successful regeneration. Alpine species are generally known for having small seeds (REF), delayed germination phenology (Fernández-Pascual *et al.*, 2021; Espinosa del Alba, Fernández-Pascual and Jiménez-Alfaro, 2024) and short longevity (Mondoni *et al.*, 2011a) which we expect to correlate with oil content and composition. We also expect to find micro and macroscale oil patterns related to temperature gradients along an altitudinal gradient and in concordance with the results previously reported in the latitudinal gradients (Sanyal and Linder, 2013; Sanyal and Decocq, 2016). At a more regional scale, we expect that strict alpine species (specialists living only in higher altitudes) will have less oil content than more generalist or strict lowland generalists’ species due to the lower temperatures experienced in alpine areas. In addition, we also tested the expectation of a negative correlation between seed mass and oil content. At a local scale, we aim to explore seed oil content and composition relationship to three other seed traits: seed mass, seed longevity (p50, time to decrease initial seed viability by 50%) and germination timing (T50, time to reach 50% of germination). We expect (1) that higher seed mass will correlate with less oil content and higher UFA/SFA ratio, (2) that lower longevity will correlate with higher oil content and a higher UFA/SFA ratio and (3) that faster germination (lower T50) will correlate with higher oil content and higher UFA/SFA ratio. Lastly, and also at the local scale, we aim to explore the relationship between oil content/composition with the local ecological optimum of the species.

## 2. Materials and Methodology (1581 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 (Jiménez-Alfaro *et al.*, 2021). 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 mostly dominated by *Poaceae* and *Cyperaceae* and the main lifeforms 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.

### Species data

We established eight sampling sites, 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 present in the area. At each sampling site, we collected floristic and community composition data for 119 alpine grassland species. We collected enough seeds to measure and analyse seed oil content and composition from 47 species (200 mg of dry seeds), the FAs data obtained were not known or published before. The collected species were classified between strict alpine and generalist according to their ecology. We identified plant specialists as those significantly associated with the target vegetation type (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 1800m a.s.l., most of them characteristic of subalpine or nitrophilous habitats and those were specified as a generalist. Seed collection permits were only necessary and 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 preliminary explored oil content variation within seed lots in five species, analysing three subsamples from each (also to ensure high precision of methodology), and we could not detect any statistically significant differences between subsamples. Our preliminary results were also consistent with the analysis of 360 accessions of Arabidopsis thaliana which contained identical FAs with only a slight variation in the relative proportion [9,10 read]. 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 to 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 oil content absolute value on a dry-weight base 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 of another 33 species native to Europe, classified as either generalist or strict lowland, congeneric of the species present we analysed, from SID database (2023). Our local species oil data in addition to the data obtained from SID database (2023) was used for regional-scale analysis, containing oil content and seed mass data for 80 species (29 strict alpine, 31 generalists and 20 strict lowland) from 19 different plant families. For local scale analysis, we used only our local species oil data, containing seed oil content and oil composition data for 47 species from 19 families.

### Seed traits data

Seed mass data was calculated for all local species averaging the weight of five replicates of 50 dry seeds. To estimate germination timing we used the T50 trait, calculated as the time (in days) to reach 50% germination, obtained from a published phenology germination experiment, done 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 to alpine species by (Mondoni *et al.*, 2011b), adapting only the number of seeds to adjust to 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 consisted of 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). For those species which required a cold stratification period (*Saxifraga oppositifolia*), after ageing period, were put in a refrigerator at 5ºC and darkness for a month before starting 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 state of the embryo. We considered seeds with white and firm embryos viable, i.e. potentially germinable (Baskin and Baskin, 2014) and we 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 are the days of ageing, Ki is the initial viability, 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 species' ecological preferences in the local community we measured the microenvironmental gradients establishing 20 additional plots (1m2) for each sampling site, five in each cardinal direction with a 10 m separation where we registered the relative abundances of all vascular species (following the methodology Jiménez-Alfaro *et al.*, 2024). In total, we established 160 vegetation plots (8 sites x 20 plots) where we also buried an iButton datalogger, buried at 5 cm deep, 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 across 11 months. From the environmental data recorded, we calculated a 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 deep 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 temperature were > -0.5 ºC (Zhang *et al.*, 2005). Then we use the bioclimatic indices to calculate species' ecological preferences averaging the climatic variables of the plots where the species was present and weighted by its coverage (only considering those plots where the species had more than 10% of relative coverage). Assuming that species would have more coverage in those plots with climatic conditions closer to their preferences.

### 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 ecology (strict alpine, generalist or strict lowland) as well as their local ecological preferences (GDD, FDD and Snow). In all models, we used weakly informative priors, 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. From the resulting posterior distributions, we calculated mean parameter estimates and 95% credible intervals (CI). 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 (1033 words)

### Altitudinal patterns at the regional scale

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 large 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.

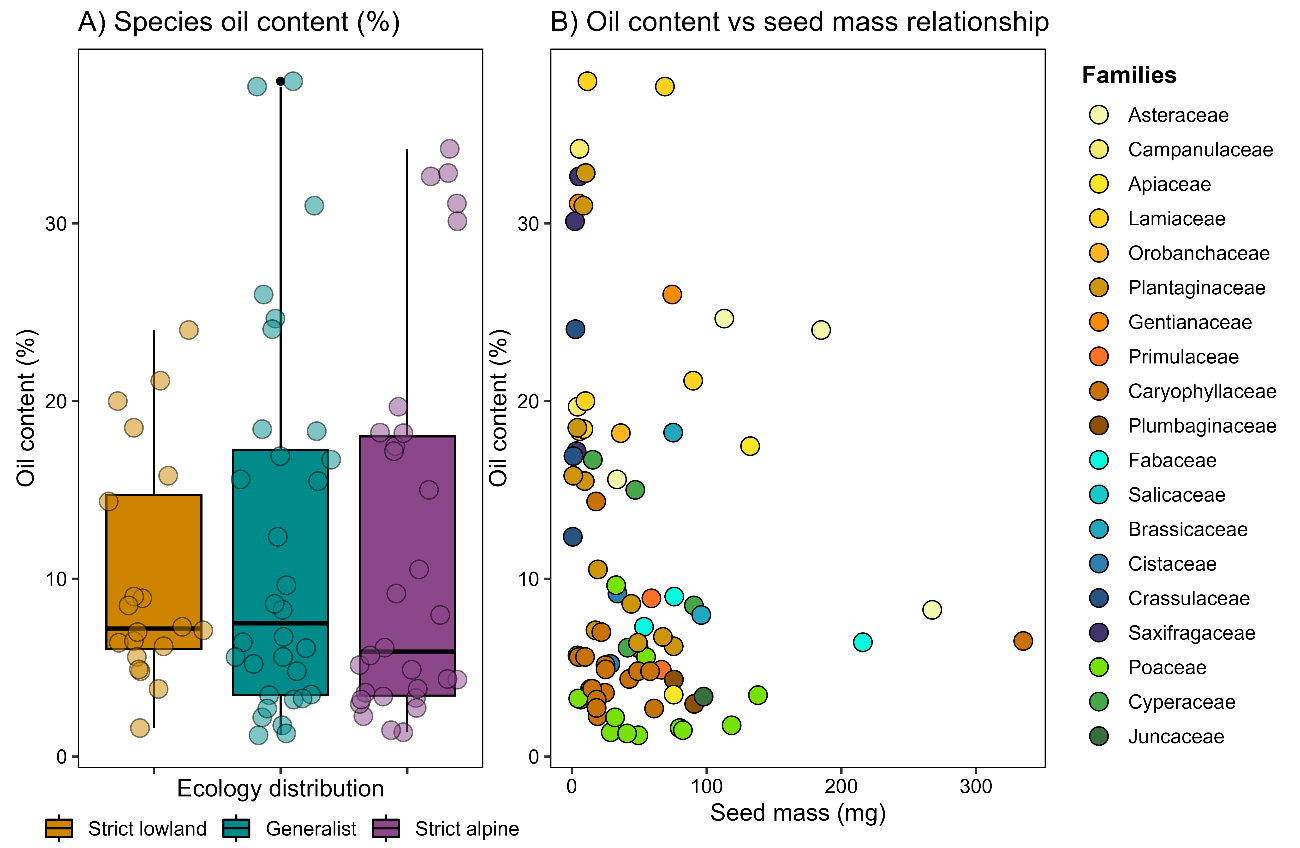


Figure 2. Regional oil content patterns (n=80 species). A) Seed oil content (%) between regional altitudinal distribution showed by different colours. B) Seed oil content (%) and seed mass (mg) non-significant relationship (MCMC-GLMMs), colours by families.

### Local 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, LA, 42.4%±2.4, mean ± se), oleic acid (C18:1n-9c, OLA, 22.6%±1.9) and alpha-linolenic acid (C18:3n3, ALA, 15.1%±2.7); and saturated palmitic (C16:0, PA, 10.8%±0.8) (details in supplementary Table S2). These four represent 94.8% of seed oil content in average (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 Caryophillaceae 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). In general, the Poaceae, Salicaceae and Cistaceae had low values of the UFA/SFA ratio, indicating higher synthesis SFAs, while 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 FAME’s with >3% relative proportion and with 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). UFA/SFA ratio and C16:0 FAMEs were contributing most in PC1 (26.4% and 18.6% respectively) and C20:1n9 and C22:1n9 are the ones contributing the most in the PC2 (21.9% and 16.3% respectively) (Figures 3D and 3E). Additionally, PCA revealed that oil content percentage is not highly correlated with any specific FA type (details in supplementary Table S5).

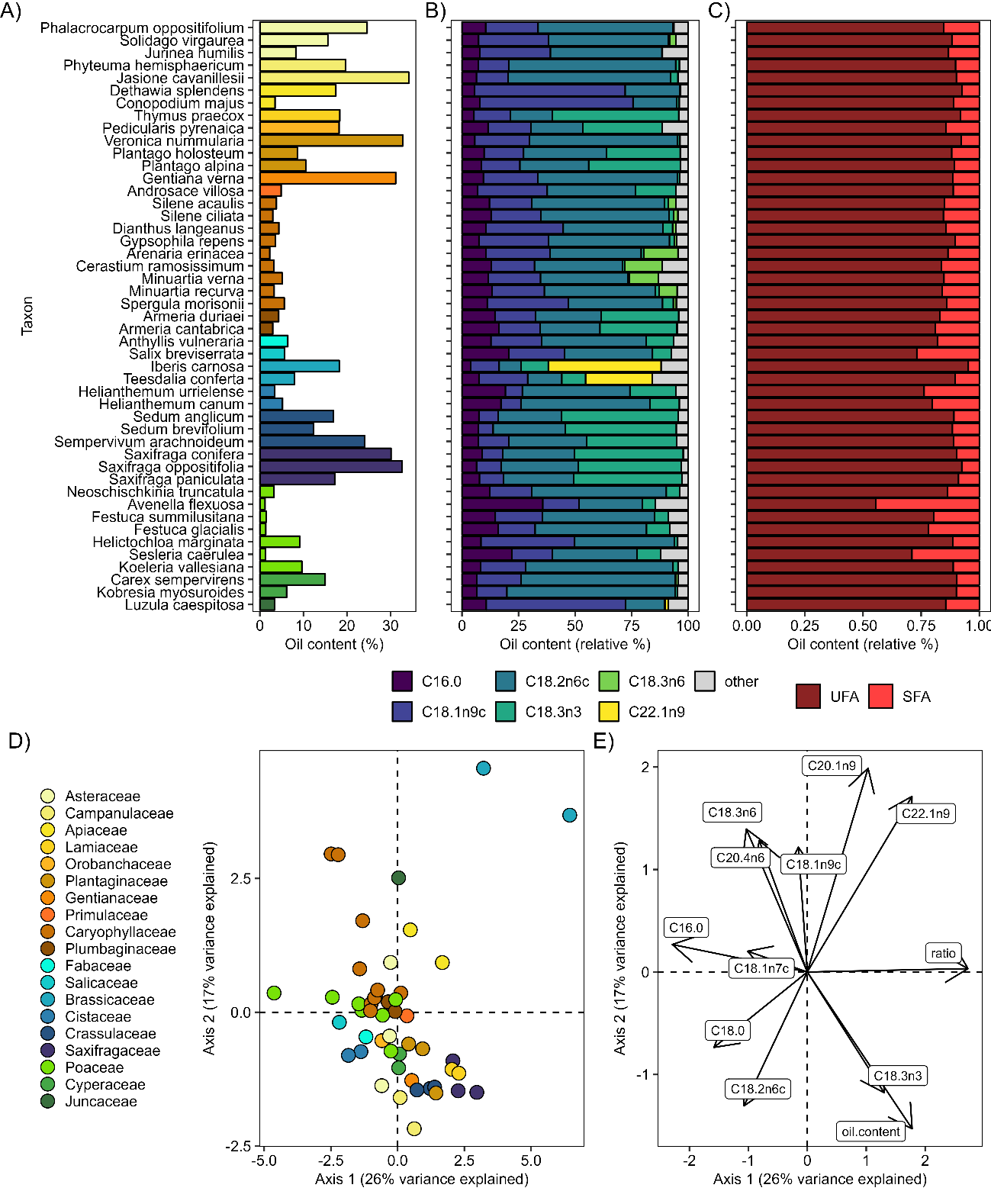


Figure 3. Local species oil content and composition data exploration (n=47). A) Oil content (%) per species, colours representing families as panel D. B) Seed oil composition of FAME’s with more than 10% of relative abundance on average across species. C) Seed oil composition per species with FAME’s 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 of oil content in small seeds, but its variability decreased in large seeds which at the same time had lower oil content. Although a general trend was visible, the relationship was between seed mass and oil content 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, firstly using the raw germination data from the artificial ageing protocol (with Multinomial MCMC-GLMM) but also using the p50 value (obtained from the probit analysis and Gaussian MCMC-GLMM). 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 signal 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).

#### Germination timing

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 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.

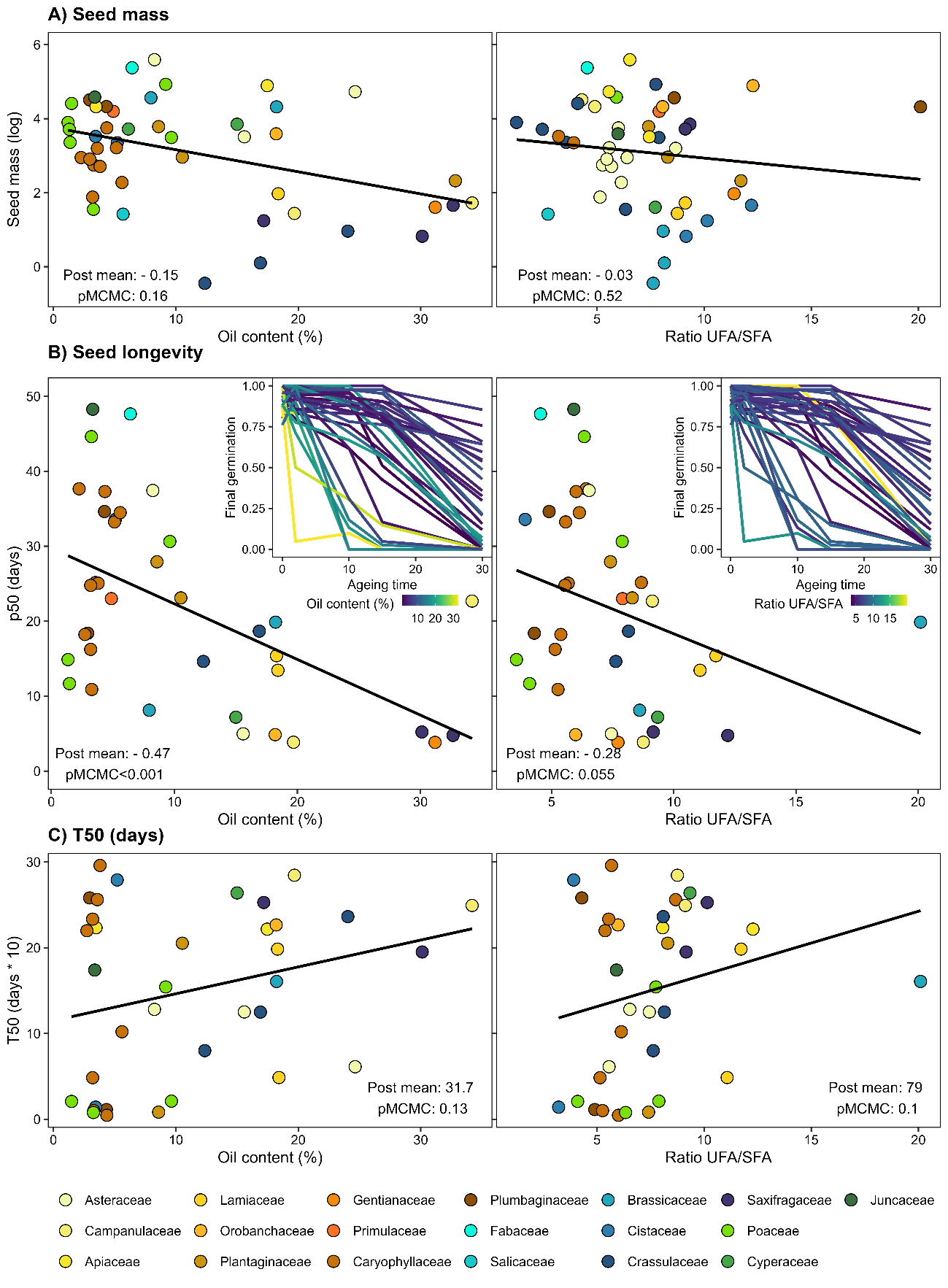


Figure 4. Seed oil content and Ratio UFA/SFA biological trade-offs. A) Correlation between seed mass and oil content) and ratio UFA/SFA, significances from MCMC-GLMM models (gaussian family, n=47) with all variables log-transformed. B) P50: time for viability to drop to 50%, based on probit analysis (GENSTAT software) and the plot inset with raw germination curves across artificially accelerated ageing protocol in the lab, colour represents 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 ratios. GDD values ranged from 650 to 2295°C with an average of 1421°C (se=67.9), data showed a negative trend with species in warmer conditions showing less oil content and lower ratio. FDD values ranged from 0.05 to 170°C with a mean of 31.8°C (se=4.7), data showed a positive trend but 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, data showed a positive trend where species with a preference for more snowy sites show higher oil content and UFA/SFA ratio. Model results details can be checked in supplementary Table S4.

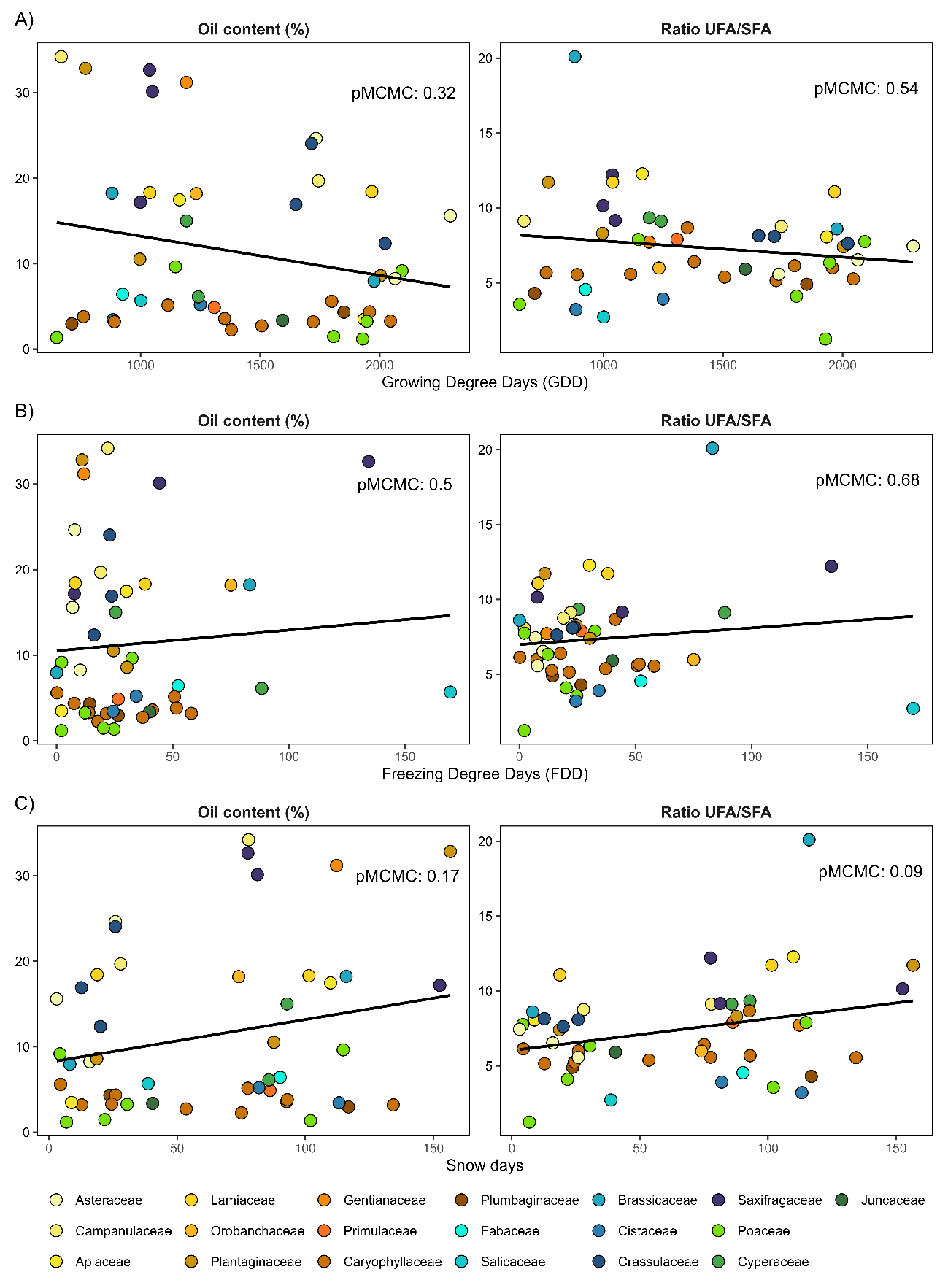


Fig 5. Seed oil content ecological trade-offs, oil content and ratio UFA/SFA log-transformed for MCMC-GLMM models (family=Gaussian, n=47). A) Relationship with GDD (Growing Degree Days °C); B) Relationship with FDD (Freezing Degree Days °C) and C) Relationship with Snow days.

## 4.Discussion (969 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 patterns corroborating macro or microevolutionary processes driving oil content and composition previously reported in the literature (Linder, 2000; Sanyal and Decocq, 2016) neither at the regional nor local scale. We did find strong indications that, in alpine species, seed oil content and composition is significantly influencing seed longevity, having potential impacts on soil seed banks and conservation strategies for seed bank managers or restoration programs.

Within our data set, we found no high correlations (>0.7) between total oil content and other specific FAs, against Sanyal and Decocq (2016) who found a significant negative correlation between seed oil content and palmitic (C16:0) and linoleic acids (C18:2n6) and positive correlation with oleic (C18:1n9), arachidic (C20:0) and eicosenoic (C20:1n9). The differences put into consideration the difficulty to describe global patterns when broad-scale data is lacking. 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 oil values range (1-38%), more similar to the one showed by weed species in France (Bretagnolle *et al.*, 2016), but still in the lower range compared to global records as expected due to the herbaceous nature of all the species included in the study (Levin, 1974). The oil composition was also similar to the ones reported by (Bretagnolle *et al.*, 2016), except for erucic (C22:1n9) for the Brassicaceae family. From other commercial Brassicaceae species, previous reports also show that is rich in elongated acyl chains from C20 to C24 (Baud and Lepiniec, 2010). The high levels of erucic are surprising due to the “high” melting point (33.5°C) (Sanyal and Linder, 2013) that would difficult energy release, however, erucic acid has also been correlated with lower seed predation (Jong *et al.*, 2016).

Against our expectations, we found no significant differences along the altitudinal gradient at the regional scale and also did not find any significant relationship with species' local ecological optimal. None of the local microclimatic gradients studied (GDD, FDD and snow) showed any significant relationship with oil content or oil composition. However, snow days showed a more pronounced relationship with the UFA/SFA ratio, showing that species leaving in more snowy sites tended to have higher UFA, which potentially could be consistent with the anti-freezing properties that UFA can provide. Nevertheless, these results seem to limit Sanyal and Linder's claims (Sanyal and Linder, 2013; Sanyal and Decocq, 2016) of strong selection at macro and microevolutionary levels constraining seed oil patterns. The lack of significant patterns might be an indication of a strong phylogenetic signal for oil content. Yet, it is difficult to provide a more complete picture due to the limited information on oil content and fatty acid composition available for wild species (Levin, 1974).

We found a huge 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). With both our regional and local species pool we found a negative correlation between seed mass and seed oil content, however non-significant. The same trend has been described in the literature for other plant datasets (Nagel and Börner, 2010; Hamilton *et al.*, 2013; Bretagnolle *et al.*, 2016) however not always significant (Finkelstein & Grubb 2002 read). Additionally, in our local alpine species pool, we found a negative trend, less pronounced, between seed mass and UFA/SFA ratio, however, the trend is contradictory with the expectation of small seeds storing more SFA, i.e. lower ratio values (Linder, 2000). 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 both seed oil content and composition. Alpine seeds, generally known for their short longevity (Mondoni *et al.*, 2011b), showed a consistent decrease in seed longevity with increasing oil content and with increasing UFA/SFA ratio, meaning more unsaturated FAs. 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 are also associated with shorter longevity (Hoekstra, 2005). Our results are consistent with the adverse effects generated during lipid oxidation previously described due to free radicals and ROS (Bailly, 2004 read Priestley and Leopold, 1979 read,Kranner et al., 2002 read), particularly, the oxidation of UFAs (Benson 1990).

Based on the results found by Sanyal and Decocq (2016) and Linder (2000), we predicted that alpine species, adapted to germinate under colder temperatures, would show faster germination with increasing oil content and increasing UFA/SFA ratio. Surprisingly, the non-significant positive trend between oil content and oil composition with T50 contradicted our initial expectations. We observed the opposite trend with species with higher oil and higher UFA germinating later. The lack of patterns is not completely surprising, as has been reported in other studies (Hamilton 2012 read) while others have indeed show that positive correlation (Gardarin 2011 read). A plausible explanation could be that the seed dormancy constraints present in alpine species (REF) drive germination patterns stronger than oil content and oil composition.

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 collection campaign. Nevertheless, we are confident to have collected fully mature seeds, following ENSCONET protocol (2009) and have applied a consistent methodology for all experiments and analyses.

## 5.Conclusions