Norwegian University
of Life Sciences

Master's Thesis 2019 30 ECTS
Faculty of Environmental Sciences and Natural Resource Management

Seasonal variation in chemical composition of *Chamerion angustifolium* pollen as measured by Fourier-Transform Infrared Spectroscopy

Simen Arne Kirkhorn
Natural Resource Management

Acknowledgements

This thesis has been a cooperation between two faculties; The Faculty of Science and Technology (REALTEK) and Faculty of Environmental Sciences and Natural Resource Management (MINA). The pollen analysis took place at REALTEK, during fall 2018. This thesis will be the last credits that will result in my master's degree in natural resource management.

First of all, I would like to thank my supervisors; professor Mikael Ohlson (MINA) and researcher Boris Zimmermann (REALTEK). Thank you, Mikael, for introducing me to this thesis, being available, helpful, for funding of really cool electron microscope images, and for always being encouraging through the whole process.
Thank you, Boris, for being patient, giving thorough feedback, and introducing me to and helping me understand very useful data analysis software. Thank you also for conducting the field work. Thank you both for interesting discussions along the way.

I also want to thank doctoral research fellow Simona Dzurendová and other colleagues at REALTEK for helping me with microscopy and lab work. In addition, I would like to thank senior engineer Hilde Kolstad for helping me with scanning electron microscope images.

Finally, I would like to thank my girlfriend, family and friends for general support through the whole process.

Norwegian University of Life Sciences

Ås, May 14th, 2019

_____

Simen Arne Kirkhorn

Abstract

The quality of pollen is of fundamental importance for male reproductive success in spermatophytes, but our knowledge as to whether pollen quality is plastic and varies throughout the flowering season is limited. Pollen quality can be estimated indirectly through information about the chemical composition of pollen.

In this study, pollen from *Chamerion angustifolium* was collected from a natural population over the course of a flowering season, lasting 48 days. The chemical composition of pollen was analyzed by Fourier-transform infrared (FTIR) spectroscopy.

Pollen was found to be chemically different throughout the flowering season, revealing a clear correlation between chemical spectra and pollination day. Additionally, pollen size decreased through the flowering season and pollen chemistry differed between two populations.

The indicated plasticity of pollen chemistry throughout the season is important, as it is a major prerequisite for seasonal changes in pollen quality. Whether changed pollen chemistry over the season results in altered pollen quality is still unknown. Additional studies should validate this, by combining FTIR spectroscopy measurements of pollen with measurements of pollen quality parameters.

Sammendrag

Pollenkvalitet er fundamentalt viktig for befruktningsprosessen hos frøplanter, men kunnskap om hvorvidt pollenets kvalitet er plastisk og endrer seg gjennom blomstringssesongen er begrenset. Pollenkvalitet kan estimeres indirekte ved å kjenne til pollenets kjemiske sammensetning.

I denne studien ble pollen fra geitrams (*Chamerion angustifolium*) samlet fra en naturlig populasjon, i løpet av en blomstringssesong på 48 dager. Pollenets kjemiske sammensetning ble analysert ved bruk av Fourier-transformert infrarød (FTIR) spektroskopi.

Resultatene viste at pollen er kjemisk forskjellig gjennom blomstringssesongen, og viser en klar korrelasjon mellom pollenets kjemiske spekter og pollineringsdato. I tillegg indikerer resultatene at pollenets størrelse reduseres i løpet av blomstringssesongen, og at pollenets kjemiske sammensetning er forskjellig mellom to populasjoner.

Plastisitet i pollenkjemi gjennom sesongen er viktig, ettersom det er en sentral forutsetning for sesongmessig variasjon i pollenkvalitet. Hvorvidt endret pollenkjemi gjennom sesongen medfører endret pollenkvalitet er fortsatt uvisst. Videre studier bør validere dette ved å kombinere FTIR-spektroskopimålinger av pollen, med målinger av pollenkvalitet-parametere.

Contents

1 Introduction

Seed plants (spermatophytes) comprise the vast majority of land plants with over 350 000 species. Although many seed plant species are able to reproduce asexually, they largely depend on sexual reproduction in order to survive and disperse. Fertilization enables genetically diverse offspring, which is necessary for the plants to make evolutionary responses in an ever-changing environment. Together with ovules, pollen plays a fundamental role in the fertilization process, as it contains the male gametes that donate parental genetic traits to the offspring. Despite the important role of pollen in fertilization, studies on how plant fitness is affected by environmental factors is largely based on measurements of female function (Schaeffer et al. 2013). In their meta-analysis of 96 studies measuring male function, Schaeffer et al. (2013) suggested that the effect of environmental factors on plant reproduction could be misrepresented by measuring female function and not male function.

The development of pollen grains starts with meiotic division from the diploid anthers of the mother plant, which results in a tetrad of haploid microspores. The microspores are first polarized and form a vegetative and generative cell (pollen mitosis I), before the generative cells elongates and divide into two sperm cells in mature pollen grains (pollen mitosis II) (Rutley & Twell 2015). Fully developed pollen grains consist of a pollen wall, with an inner layer containing cellulose (intine), and a tough outer layer (exine) containing sporopollenin, which are also found in spores (Halbritter et al. 2018, pp. 24-25). As pollen grains develop, they are provisioned with nutrient storage products and enzymes from specialized tapetal cells in the anthers of the mother plant. Through pollination, pollen is transported to receptive stigmas, where pollen rehydrate, germinate, and form pollen tubes by metabolizing the stored nutrients and enzymes (Delph et al. 1997; Johannsson & Stephenson 1998; Stanley & Linskens 1974, p. 15). Stored nutrients also aid the pollen tube growth from the stigma through the pistil into the ovary (Rutley & Twell 2015).

In addition to nutrient reserves, pollen transcriptome is of fundamental importance for successful reproduction. The stored transcripts of mature pollen grains and its sperm cells are unique to the parent plant and are involved with germination and pollen tube growth (Rutley & Twell 2015). Transcripts are also involved with communication and signaling with the

pistil sporophyte and female gametophyte, which guide pollen tube growth (Palanivelu & Tsukamoto 2012). Consequently, there is no straightforward association between large nutrient reserves in pollen and fertilization success, as pollen with favorable genetic composition might fertilize ovules first (Marshall & Ellstrand 1986).

Male reproductive success is influenced by several factors (Schaeffer et al. 2013; Williams & Mazer 2016), of which pollen quantity and quality are of fundamental importance (Zimmermann et al. 2017). Pollen quantity relate to how much pollen is available for dispersal and can potentially reach stigmas, whereas pollen quality refers to the viability, germination capacity and growth rate of pollen tubes (Lau & Stephenson 1993; Williams & Mazer 2016). These parameters of pollen quality are closely related to the chemical composition of pollen grains, which again is related to nutrient and energy reserves. Energy and nutrient content are therefore an important components of pollen quality (Zimmermann et al. 2017). Carbohydrates and lipids serve as principal nutrients of pollen grains (Pacini 1996), and these compounds do also have specialized functions (Zimmermann 2018). For example, lipids play crucial roles for both pollen germination and pollen tube growth (Rodríguez-García et al. 2003). In addition, enzymes from protein are important during pollen tube growth (Roulston et al. 2000), while cytoplasmic carbohydrates help prevent heat stress (Vesprini et al. 2002) and desiccation (Pacini 1996; Speranza et al. 1997). All nutrients also contain important structural components (Zimmermann 2018). The chemical composition of pollen can therefore be used as an indication or proxy of pollen quality.

Species-specific stigma-ovule distances are key-determinants of the amount of resources a given pollen species need in order to grow a pollen tube that is long enough to fertilize an ovule (Roulston et al. 2000; Williams & Mazer 2016). Pollination strategy is another selective agent that influences the chemical composition of pollen grains (Baker & Baker 1979; Roulston et al. 2000; Zimmermann & Kohler 2014). Insects often use pollen for nutrition in addition to nectar (Roulston et al. 2000; Wäckers et al. 2007), hence pollen from many insect-pollinated (entomophilous) plant species contain more lipids and protein (Hanley et al. 2008; Zimmermann & Kohler 2014) compared to self-pollinated (autogamous) or wind pollinated (anemophilous) species, that often contain more carbohydrates (Baker & Baker 1979; Zimmermann & Kohler 2014). Since production of protein and lipids are more costly than production of starch (carbohydrate), selection would favor pollen with more starch for plants with non-rewarding pollen, and less starch (more protein /lipids) for entomophilous species

(Baker & Baker 1979). As different families and species differs with regards to pollination strategy and stigma-ovule length, the chemical composition of pollen grains is therefore typically species- and/or family specific (Baker & Baker 1979; Zimmermann 2010; Zimmermann & Kohler 2014).

Environmental conditions influence the chemical composition and quality of pollen, and if changing environmental conditions affect pollen quality negatively, it may cause reduced fertilization, which in turn can have considerable ecological effects. It is therefore essential to know whether the chemical composition of pollen is genetically fixed or plastic in response to environmental change. Several studies have shown that environmental stress can affect the chemical composition of pollen (Lau & Stephenson 1994; Stanley & Linskens 1974; Van Herpen 1981; Van Herpen 1986; Zimmermann et al. 2017), and pollen performance (Delph et al. 1997; Elgersma et al. 1989; Lau & Stephenson 1994; Lau & Stephenson 1993; Quesada et al. 1995; Van Herpen & Linskens 1981). For instance, pollen grains are sensitive to heat stress (Pham et al. 2015; Schaeffer et al. 2013; Van Herpen & Linskens 1981), particularly during development (Bokszczanin et al. 2013; Johannsson & Stephenson 1998; Rieu et al. 2017). Most of the studies have in common that they are conducted under artificial conditions indoor, in order to isolate the effect of different environmental variables. Even though these studies have been done indoors, they indicate that pollen quality and performance should also be affected by environmental conditions under natural outdoor conditions.

Many plants flowers sequentially, i.e. flowering and pollination occurring over several weeks. It is thought that for co-existing plant species, sequential flowering has evolved in order to reduce fitness reductions caused by competition for pollinators (Forcella & Wood 1986). Due to the long flowering periods, sequentially flowering plants are good candidates for studying the effect of seasonal variation on pollen plasticity as sequentially flowering plants will experience larger variations in both biotic and abiotic factors during flowering. Different groups of pollinators show season-specific variation in their abundance, often caused by flower abundance at the community level (Ramrez 2006). Temperature, precipitation and moisture levels vary considerably over a season, and are expected to show larger variations due to changing climate. Temperature is expected to increase on average, and flowering commence earlier. As climatic conditions are expected to become increasingly more variable, phenotypic plasticity of plants have been suggested to be an important factor in plant survival and reproduction, along with adaptive ability (Anderson et al. 2012).

Environmental conditions mentioned may vary considerably over the course of a flowering season, hence seasonal variation in pollen quality is best assessed if pollen is collected from sequential flowering plants, with continuous pollen production. Neither of these conditions are met by previous research, due to the study of too short pollination periods (Zimmermann et al. 2017), or discontinuous pollination (Bağcıoğlu et al. 2017). As far as I know, there are no studies that have measured how pollen quality may change over the season in a sequentially flowering plant that flowers over an entire growing season. There is thus a need for more knowledge about how, and to what extent pollen quality may change over the season in sequentially flowering plants.

In this study, pollen was collected from *Chamerion angustifolium subsp. angustifolium,* a species that meets these conditions. It has sequential flowering with continuous pollen production, lasting for several weeks during the summer season, normally from June to September (Myerscough & Whitehead 1966). *C angustifolium* is a herbaceous perennial forb, naturally occurring throughout vast areas in the northern hemisphere, in North-America, Europe, and Asia (Myerscough 1980). Due to the long flowering season, seed dispersal occurs from late July to late September (Myerscough 1980), and therefore flowering, fertilization and seed dispersal typically occur simultaneously in *C. angustifolium*.

The standard analysis of pollen focus on identifying the morphological structure of pollen grains and is applied in a number of scientific fields such as taxonomy, phenology, forensics, paleoecology and to make pollen forecasts (Zimmermann 2010). This category of pollen analysis is largely dependent on optical microscopy and scanning electron microscope (SEM) (Hesse et al. 2009, pp. 51-52) and do not give information concerning the chemical composition of pollen grains (Bağcıoğlu et al. 2015; Bağcıoğlu et al. 2017). Analyses regarding chemical composition of pollen grains (Piffanelli et al. 1998; Roulston et al. 2000; Speranza et al. 1997; Van Herpen 1981) has been studied to a lesser extent as the methodologies require complex preparation of samples, which is both time demanding, labor-intensive and expensive (Bağcıoğlu et al. 2015; Zimmermann et al. 2017; Zimmermann 2018). Moreover, measurements of each chemical group, i.e. carbohydrates, lipids and proteins require its own method, allowing only one group of chemicals to be studied at a time. In order to find the effects environmental stress has on pollen, much data is needed, and this demands a more efficient method (Bağcıoğlu et al. 2017).

Fourier-transform infrared (FTIR) spectroscopy is an alternative efficient method that can be used to analyze the complete chemical composition of pollen all in the same measurement. This method has had a breakthrough in recent years (Bağcıoğlu et al. 2015; Bağcıoğlu et al. 2017; Gottardini et al. 2007; Jiang et al. 2015; Lahlali et al. 2014; Pappas et al. 2003; Zimmermann 2010; Zimmermann & Kohler 2014; Zimmermann et al. 2015; Zimmermann et al. 2017). In FTIR spectra, vibrational frequencies of chemical bonds are directly related to functional groups that make up biochemical compounds in pollen. Therefore, the FTIR spectra of pollen can provide information on major biochemical components, such as carbohydrates, lipids and proteins (Zimmermann et al. 2017). In this way, FTIR provides an overview of the chemical composition of pollen with fast and economical measurements, since it often does not require any chemical pretreatment of the samples. FTIR spectra of pollen can be considered as a biological fingerprint for phenotyping of pollen since a number of studies have shown that the spectra can be used for chemical characterization and identification with respect to environmental stress and the phylogenetic affiliation of pollen (Bağcıoğlu et al. 2015; Lahlali et al. 2014; Zimmermann & Kohler 2014; Zimmermann 2018). *C. angustifolium* is a suitable species for FTIR-analysis, as it produces sufficient pollen for FTIR analysis (>0.5 mg (Zimmermann et al. 2015)).

The aim of this study was to examine whether the chemical composition of *C.* angustifolium pollen varies in response to pollination day, by collecting pollen throughout the flowering season and perform FTIR measurements. An additional aim was to analyze seasonal variations in pollen size (pollen mass and diameter of pollen grains). Pollen was also collected and analyzed from another population to uncover potential population specific effects of the study population.

The hypotheses were:

1) Pollen size (mass and diameter) will be smaller late in the flowering season due to diminishing resources.
2) Pollen chemistry will change over the course of the flowering season, which can be observed by collection of pollen samples over several weeks with varying environmental conditions.
3) Pollen chemistry will differ between the study population and the comparison population.

2 Materials and methods

2.1 Study area

The study population was located in Ås, Akershus county, at (6622360.96, 263614.81, UTM zone 33, EUREF 89), in an urban area, at 124 m.a.s.l. For comparison, a smaller amount of pollen was collected from another population, which was located in Kittilbu, Oppland county, at (6117480,99535, UTM zone 33, EUREF 89). This population was located roadside on wasteland at approximately 800 m.a.s.l. in sub-alpine vegetation below the upper forest boundary. Figure 1 shows pictures of both the study population, and the Kittilbu population. The bedrock of the Ås study population consists of gneiss and granite and the superficial deposits are dominated by marine deposits, with thick moraine and eskers. The bedrock of the Kittilbu population consist of schist and metasandstone, with moraine and peatland as superficial deposits (NGU 2019). The average daily temperature of the flowering season (48 days) was 15.5 °C, and the average daily precipitation was 3.3 mm (160 mm in total), both within the normal range. However, the month of July was considered a drought period, as the precipitation was 44 mm, compared to the average of 81 mm. The weather data in Ås are retrieved from BIOKLIM field station, 1.7 km from the study site (Wolff et al. 2018).

Figure 1: The study population in Ås on the left panel, the Kittilbu population on the right. Photo: Boris Zimmermann (left panel), Mikael Ohlson (right panel).

2.2 Study Species

C. angustifolium is protandrous, i.e. stamens ripe and shed pollen before pistils open, a common strategy for cross-pollination. *C. angustifolium* is considered a generalist in terms of pollinator attraction (Kooi et al. 2016), being pollinated by a wide range of insects, e.g. *Hymenoptera, Lepidoptera, Coleoptera and Diptera* (Myerscough 1980). Although self-compatible, experiments shows that *C. angustifolium* rarely are able to produce seeds, without insects or wind (Mosquin 1966). Since *C. angustifolium* is a species that propagates vegetatively, shoots observed in the field are often ramets of the same genet (Myerscough 1980). As none of the ramets in the present study was genetically tested due to budget limitations, it remains unknown whether ramets are genetically different.

2.3 Sampling of pollen

Seven ramets were selected for the study, and the stamens from each ramet were collected in the flowering season of 2017, locally from 11th of July to 27th of August. The population was observed every morning and afternoon each day of the flowering season, for sampling of newly opened flowers. All stamens were collected in the same developmental phase, i.e. anthesis start, when flowers open and before pollen was exposed on the anther and before stigmas were developed. This is shown as stage B in Figure 2. Height measurements of individual flowers, inflorescences and whole ramets was carried out during the field work. The sampled stamens were then kept in Eppendorf tubes at room temperature for 24 hours, to desiccate. The stamens were shook to separate pollen from the anthers and filaments, and the pollen weight was recorded. Subsequently, the pollen was stored in Eppendorf tubes, at -18 °C. The dry weight of ramets were measured in the end of the flowering season. At Kittilbu, three stamens were collected from four different ramets, for 12 pollen samples in total. Stamens were collected in the afternoon on 17th of August 2017. The pollen samples were left to desiccate indoor, and pollen was separated from stamens and stored at -18 °C correspondingly to the method used for the study population in Ås.

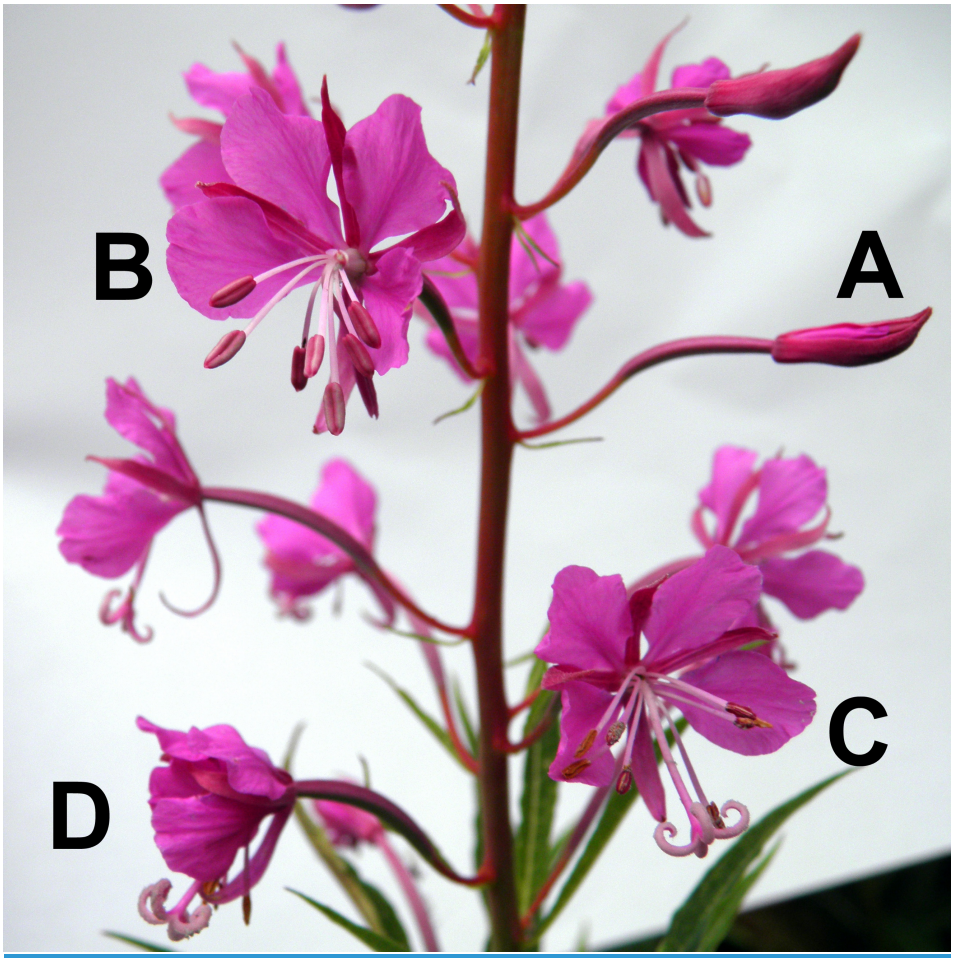


Figure 2: Developmental stages of *C. angustifolium* flowers. A: closed flowers. B: Fresh and closed stamens before pollen desiccation, and closed stigma. C: Stamens releasing pollen, and stigma just opened. D: dry stamens where pollen has been shed, and fully opened stigma covered with pollen. Photo: Boris Zimmermann.

2.4 Pollen size (microscopy measurements)

The structure of the pollen grains can be seen in the SEM images in Figure 3, and in the diameter measurements from optical microscopy in Figure 4. The SEM-images were taken at the Imaging Centre at The Norwegian University of Life Sciences (Ås), with help from Hilde Kolstad. Pollen grains have two different morphological conditions, either dry or rehydrated. Pollen grains change condition due to changed osmotic pressure in the cytoplasm resulting in folding or unfolding of the pollen wall, also known as the harmomegathic effect. This effect protect the pollen from severe desiccation during pollen presentation and dispersal (Halbritter et al. 2018, p. 57). When pollen grains are dehydrated, the pollen wall fold and form sunken furrows, as seen in Figure 3. When hydrated, they expand and have a more inflated appearance (Heslop-Harrison 1979) (Figure 4). Pollen grains of *C. angustifolium* contain viscin threads (Hesse et al. 2009; Myerscough 1980), visible as long thin threads (Figure 3). They arise from the outer part of the pollen wall (exine) and contain sporopollenin. They connect tetrads or

single pollen grains and facilitate the aggregation of pollen (Hesse 1983). *C. angustifolium* also contain pollenkitt, and both these structures help spreading more pollen grains with pollinators (Hesse 1981).

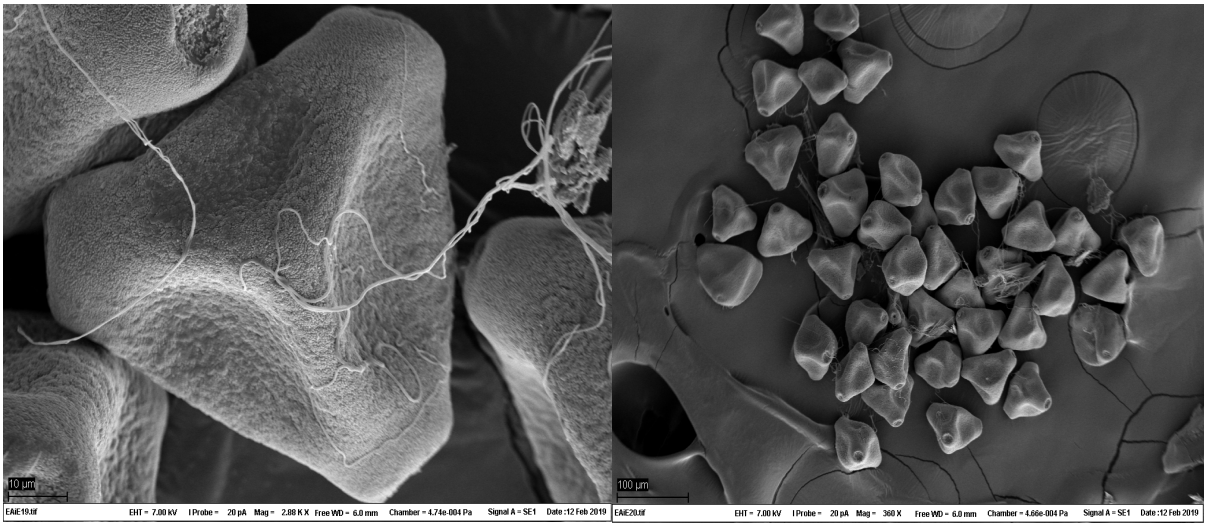


Figure 3: SEM-images of *C. angustifolium* pollen grains. Scale 10 µm (left panel), and 100 µm (right panel). Pollen grains are visibly sunken, due to dehydration.

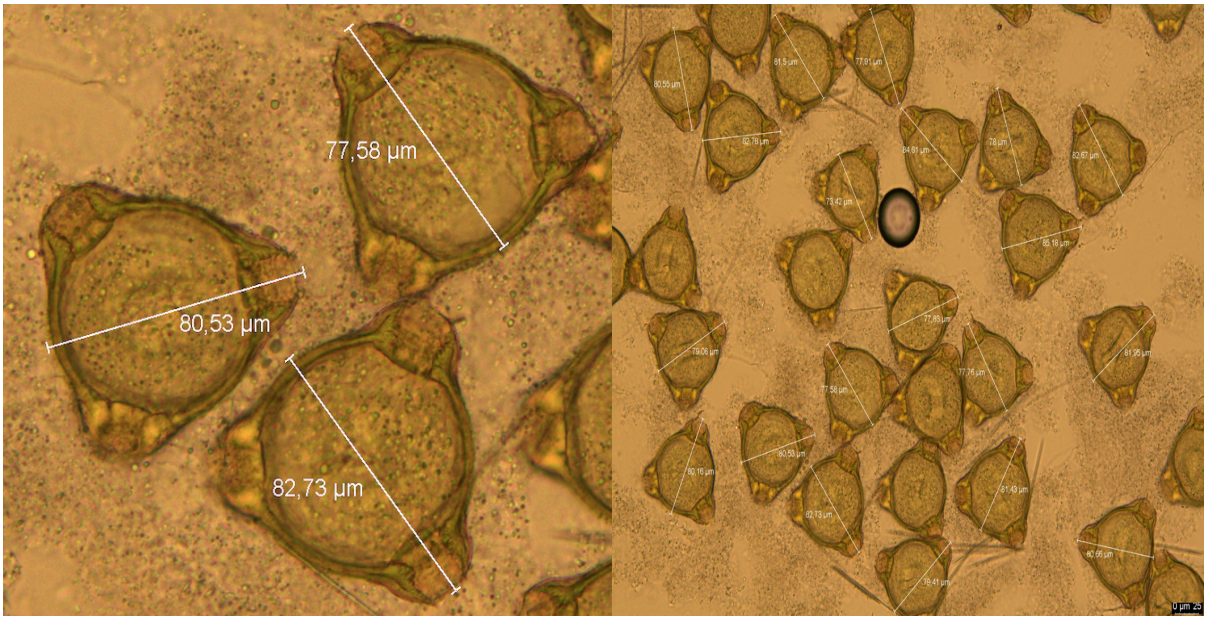


Figure 4: Optical microscopy picture of KOH-hydrated pollen grains, with diameter measurements.

The seasonally first five, middle five, and last five pollen samples (15 in total) of each ramets flowering sequence (ramet specific flowering season) were selected for diameter measurements, to enable analysis on seasonal variation of the diameter of pollen grains. Variation between ramets was also analyzed. In addition to this, five pollen samples from the Kittilbu population were also measured. For each pollen sample, 20 individual pollen grains were measured. Before each measurement, approximately 0.5-1 mg of pollen were transferred with a spatula to microscope slides and rehydrated with a 2,5 % KOH solution. After

approximately one or two minutes, the pollen was hydrated and ready for measurements. Each pollen grain was measured once with ×20 objective. Only successfully hydrated pollen grains with a polar view, (i.e. seen as triangular shape in Figure 4) were measured, hence no measurements in equatorial view were made. Pollen grains of *C. angustifolium* consists of three apertures (i.e. triporate pollen), which are located approximately just as far apart from each other (Halbritter 2017). Pollen grains were measured from the aperture that was furthest away from the opposite side (Figure 4), as pollen size should be indicated by the largest diameter measurement possible (Hesse et al. 2009 ,p. 17).

2.5 Data analysis of pollen size

To assess whether pollen mass and pollen size (diameter) varied between ramets and throughout the flowering season, statistical tests were conducted in Rstudio 1.1.419, USA. . The flowering season was split into three seasonal parts from dates 11.07-26.07 (early), 27.07-11.08 (mid), and from 12.08-27.08 (late). It should be noted that the number of pollen producing flowers declined towards the end of the flowering season, so that the number of samples differed between the three seasonal parts (Table 1).

Table 1: Number of pollen samples for each seasonal part (pollen mass), and number of pollen grains for each seasonal part (pollen diameter).

Seasonal part	Pollen mass	Pollen diameter
Early	286	1341
Mid	131	537
Late	53	421

Seasonal variation of pollen mass was also examined by pollination day, i.e. day of anthesis start. Shapiro tests and Q-Q plots were used to check if data were normally distributed, and the Levene's and the Bartlett test were used to check if data had equal variance. If the assumption of equal variance were met ($p > 0.05$), a parametric one-way ANOVA test was conducted to check if differences were statistically significant. If the assumption of equal variance was not met, or if normal transformation failed, then a non-parametric Kruskal-Wallis test is an alternative to one-way ANOVA to check whether differences are statistically significant. To check which groups (ramets or seasonal parts) within the dataset were statistically different, a Tukey test (parametric) or pairwise Wilcoxon rank sum test (non-parametric) was conducted.

2.6 Measurement of infrared spectra

A Vertex 70 FTIR spectrometer (Bruker Optik GmbH, Germany) with single reflectance attenuated total reflectance (SR-ATR) was used to record spectra. 32 scans were used to record the infrared spectra from ATR, using the horizontal SR-ATR diamond prism with 45° angle of incidence on a High Temperature Golden gate ATR Mk II (Specac, United Kingdom). Spectrum from each pollen sample was recorded, with background spectrum (empty ATR plate) being recorded between each sample measurement in order to minimize the effect of changed air conditions in the laboratory. Spectra that revealed noise or was considerably different, had several technical replicates. The OPUS/LAB software (Bruker Optik GmbH, Germany) was used for data acquisition and instrument control. Approximately 1 mg of pollen were transferred to the ATR plate for each measurement, and all pollen samples were measured once. Since a minimum of 0.3 mg was needed to acquire a high-quality FTIR spectrum, pollen samples with less than 0.3 mg were not measured. The spectra were obtained in the frequency range between 4000-600 cm^{-1}, with 4 cm^{-1} spectral resolution.

In ATR-FTIR measurements, the infrared light is directed at an angle where the light is totally reflected. Accompanied with high refraction, this causes the evanescent wave to reach into the sample and create an evanescent wave (Barth 2007). Penetration depth of the evanescent wave is in the range of 0.5-5 μm, and it depends on IR wavelength as well as on the chemical and optical properties of the sample. Due to the relatively small penetration depth, the chemical information obtained by ATR-FTIR is biased towards pollen grain wall chemistry. Because of the short penetration depth and the fact that the biochemical variation is large within pollen grains (Zimmermann 2010), it is essential that the sample is in good contact with the crystal. This is ensured by applying pressure with a built-in clamp device on the spectrometer.

2.7 Spectral pre-processing and data analysis of chemical spectra

Raw spectra from vibrational spectroscopy are often affected by random measurement noise, CO_2 and water in the experimental setting, and unwanted physical variations which can have interfering effects on the chemical information of the spectra. Although these problems can be minimized by proper sample preparation and procedures, preprocessing methods are nevertheless important to correct for these problems (Afseth & Kohler 2012). Before preprocessing, the spectral region from 1800 to 800 cm^{-1} was selected as this region contain characteristic bands for pollen grains (Bağcıoğlu et al. 2015; Gottardini et al. 2007; Pappas et al. 2003; Zimmermann 2010; Zimmermann & Kohler 2014; Zimmermann et al. 2015). The

preprocessing method can affect what chemical information is highlighted in the vibrational spectra of pollen. As opposed to non-derivated data that show broad spectral features associated with carbohydrates and proteins, derivated data will highlight narrow spectral features such as those associated with sporopollenins and lipids (Zimmermann & Kohler 2014). In order to reduce the risk of ignoring important chemical information, original data from FTIR was therefore preprocessed in two procedures, one with and one without conversion to second-derivative spectra.

The first preprocessing procedure is based on non-derivated data and is only preprocessed by multiplicative signal correction (MSC) in MATLAB (The MathWorks, Natick, USA). MSC is a model based preprocessing method that normalize spectra by using a reference spectra (Geladi et al. 1985). The method is used to separate and remove physical interferent information from relevant chemical information and exclude baseline effects (Martens et al. 1983, cited in Afseth & Kohler, 2012). In the second preprocessing procedure, the raw spectra was transformed into second derivate form with the Savitzky-Golay algorithm (Savitzky & Golay 1964), using a window size of 13 and a polynomial order of two, followed up by MSC with Orange 3.18.0 data mining toolbox for Python (University of Ljubljana, Slovenia). In all cases, an average spectrum was used as the reference in MSC. The algorithm settings were chosen based on the noise level and peak-parameters (Zimmermann & Kohler 2013). Calculating second derivates of spectra resolve overlaying band signals, reduce background effects and remove baseline effects caused by scattering and light source variations (Zimmermann & Kohler 2013). It should also be noted that carbohydrates are overestimated in FTIR-ATR measurements, as this method emphases lower wavenumbers (Misfjord 2014).

Principal component analysis (PCA) was used to examine biochemical differences between pollen samples. PCA transforms the original set of variables into a group of new variables, so-called principal components (PCs). The percentage of explained variance determines the ordering of the new principal components, hence PC1 account for the most variation, PC2 the second most variation, etc. This process reduces the dimensionality of the data, without substantial loss of information (Sarmiento et al. 2011). FTIR data from biological samples are highly colinear, because spectral bands comprise many variables, and bands are often associated with the same biochemicals (Berdahl 2014). PCA is a good tool for collinear data, because the resulting PCs from a PCA are able to show patterns of the co-variance between the original set of variables (Martens & Martens 2001, p. 109), that are otherwise hidden when

visually inspecting the raw spectra (Sarmiento et al. 2011). These patterns can be visually inspected in a score plot, showing how different individual samples differ in their PC scores. Each PC also has loading plots, that illuminates how the PCs are associated with the original spectral data. The loadings can thus be used to find how spectral regions contribute to their respective PC (Kohler et al. 2008). As different spectral bands and regions are related to different biomolecules (Zimmermann & Kohler 2014), the loadings could show different ratios in the chemical composition of pollen. Therefore, the chemical information retrieved from loading plots can help explain the sample pattern in the score plots (Kohler et al. 2008). The dataset was grouped according to ramet membership and in groups consisting of the first five, middle five and last five pollen samples in the flowering season for each ramet. This was done in order to inspect potential variation in chemical composition of pollen samples between different ramets and throughout the flowering season. The number of components were selected to cover at least 95.0 % of the variation. PCA was conducted using the Orange 3.18.0 software.

Following the PCA analysis, a classification analysis was conducted to further explore biochemical differences. The complete dataset of 427 spectra belonging to different pollen samples were split into a training set of 287 spectra, and a validation set of 140 spectra, i.e. a 67/33-split. This division of data was chosen to balance the variance of parameter estimates and the variance of the performance statistic. Both preprocessing procedures were used, and grouped ramet-wise. The data was also grouped into seasonal parts in the same way as for PCA, to inspect potential chemical variation in pollination sequence throughout the flowering season. This dataset consisted of a training dataset of 105 spectra, and validation set of 39 spectra (also 67/33-split).

Four different set of learning algorithms were used; k-nearest neighbor (k-NN), support vector machine (SVM), artificial neural network (ANN), and random forest (RF). The k-NN parameters were the following: 3 number of neighbors, eucledian distance with uniform weight. The SVM parameters were the following: SVM type with cost of 1 and a regression loss epsilon value of 0.10, a sigmoid kernel (tanh(g x*y+c)) with g=auto and c=1,00, optimization parameters with a numerical tolerance of 0,0010, and a iteration limit of 100. ANN were set to include 500 neurons in hidden layers, with rectified linear unit activation function and Adam optimization algorithm, regularization value of 0,0001, with 300 maximum iterations. ANN consists of a multi-layer perceptron (MLP) algorithm (several algorithms) with backpropagation and is able to learn both linear and non-linear models. For random forest, the

number of trees were 1000, with five attributes considered at each split, and subsets smaller than 3 were not split. To control for potential overfitting, the data were randomized before being grouped and processed by the same algorithms with identical settings. Correct classification of a large proportion of randomized samples would indicate overfitting of the analysis that is not randomized. The classification analysis was conducted using the Orange 3.18.0 software.

Additionally, partial least squares regression (PLSR) analysis was used to evaluate the relation between spectral data and pollination day (i.e. anthesis start) and the relation between spectral data and pollen mass. In PLSR, the x and y variables (spectral data and pollination day/pollen mass respectively) are projected into sets of latent factors (scores of x and y), in a way that maximizes the covariance between them (Martens & Næs 1992). X-scores are used to predict Y-scores, and the Y-scores are used to make predictions for the response. The purpose of PLSR is to obtain the latent factors that is explaining most of the variation in the response (Randall 1995). Cross-validation consisting of 10 segments was used to determine the optimal number of components (PLSR factors) for the calibrated models. The resulting calibrated models thus had the highest predicted R^2-values possible, without using too many factors.
The root mean squared error (RMSE) and determination coefficients (R^2) representing predictive ability were used to evaluate the strength of the models. Similarly, to PCA, loading plots for the regression coefficients in PLSR can be used to identify spectral bands or regions that are affecting the models (Correa et al. 2012) (i.e. to obtain chemical information relevant to the reference parameter, such as pollination day or pollen mass). PLSR analysis were conducted in Unscrambler X 10.5 (CAMO software, Oslo, Norway).

3 Results

3.1 Pollen production success

The pollen production differed between ramets, and all ramets produced flowers containing sufficient amount of pollen for FTIR measurements. Most ramets had few or none undeveloped flowers, i.e. flowers that did not open, or open flowers that did not produce pollen (Table 2). Figure 5 shows both a normal pollen producing stamen, and a stamen that did not produce pollen. Ramet C stands out with 10 undeveloped/closed flowers, and 3 open flowers without pollen. Ramet C was the only ramet that produced side ramets (side shoots). It produced five side ramets which were analyzed together. On side-ramets, 23 out of 50 flowers were

undeveloped or closed. In total, 415 out of 470 flowers produced enough pollen for FTIR measurements. All 12 samples from the Kittilbu population produced sufficient pollen for FTIR measurements. The flowering period differed between ramets (Table 4), and ramets differed in terms of flowering frequency (Appendix 1).

Table 2: Flower production success for different ramets. Ramet F had two undeveloped flowers, and two were damaged by storm. For side-ramets, the number of pollen samples is the combined number for all five side-ramets.

Ramet	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**Side-ramets**	**K**
Open flowers without pollen		2	3			3			
Undeveloped flower/closed	4		10			4*	6	23	
Flowers containing sufficient amount of pollen for FT-IR measurements	51	62	84	90	33	45	23	27	12

Table 3: Height measurements for ramets, inflorescences and flowers.

Ramet	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**Side-ramets**
Ramet height	179	156	195	184	120	164	161	
Inflorescence height	40.9	29.6	42.1	39.8	14.7	16.6	7.4	26.2
Average height per flower	0.74	0.46	0.43	0.44	0,45	0.32	0.26	0.52

Table 4: Flowering period (dates in 2017) and total flower production of different ramets of the study population.

Ramet	**Flowering start**	**Flowering end**	**Number of flowers**
All (Ås population)	**11.07**	**27.08**	**470**
A	11.07	25.08	55
B	11.07	05.08	64
C	13.07	27.08	97
D	13.07	27.08	90
E	19.07	05.08	33
F	22.07	15.08	52
G	04.08	16.08	29
Side-ramets	20.07	08.08	50

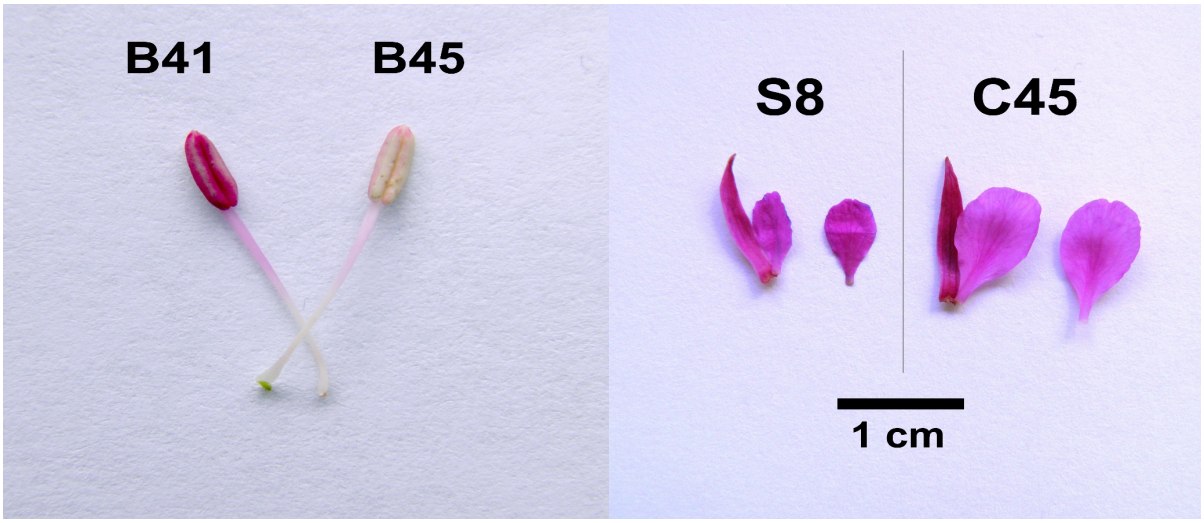


Figure 5: Left panel: a stamen with normal pollen production from sample B41, and a stamen without pollen (B45). Right panel: length of petals and sepals from two samples belonging to ramet C, and one of its side-ramets. Both S8 and C45 were collected on the same date (23.07.2017). Photo: Boris Zimmermann.

3.2 Pollen mass

The pollen mass differed between ramets, and in terms of total pollen production, ramet D produced most pollen, and ramet G produced the least (Table 5). If ramet C is considered with its side-ramets as one individual ramet, then ramet C produced most pollen, with 266.6 mg pollen. Ramet A produced on average most pollen per flower, while the side-ramets belonging to ramet C produced the least relative amount of pollen per flower. The side-ramets produced smaller petals, compared to other ramets (Figure 5).

Table 5: Pollen mass of ramets, average pollen mass per flower, and standard deviance.

Ramet	A	B	C	D	E	F	G	S
Total pollen mass (mg)	177.7	137.6	212.1	233.2	64.1	82.9	46.8	54.5
Average pollen mass per flower (mg)	3.2	2.2	2.2	2.6	1.9	1.6	1.6	1.1
Standard deviance (mg)	1.2	0.8	1.1	0.6	0.5	0.9	0.9	1.1

Pollen mass was found to differ among ramets, with high statistical significance (Figure 6). A pairwise Wilcoxon rank sum test revealed that 22 out of 28 interactions between ramets were significant. As seen in in Figure 6, ramet A produced significantly more pollen per flower on average, compared to all other ramets. The side-ramets produced significantly less pollen per flower, compared to all other ramets except for G.

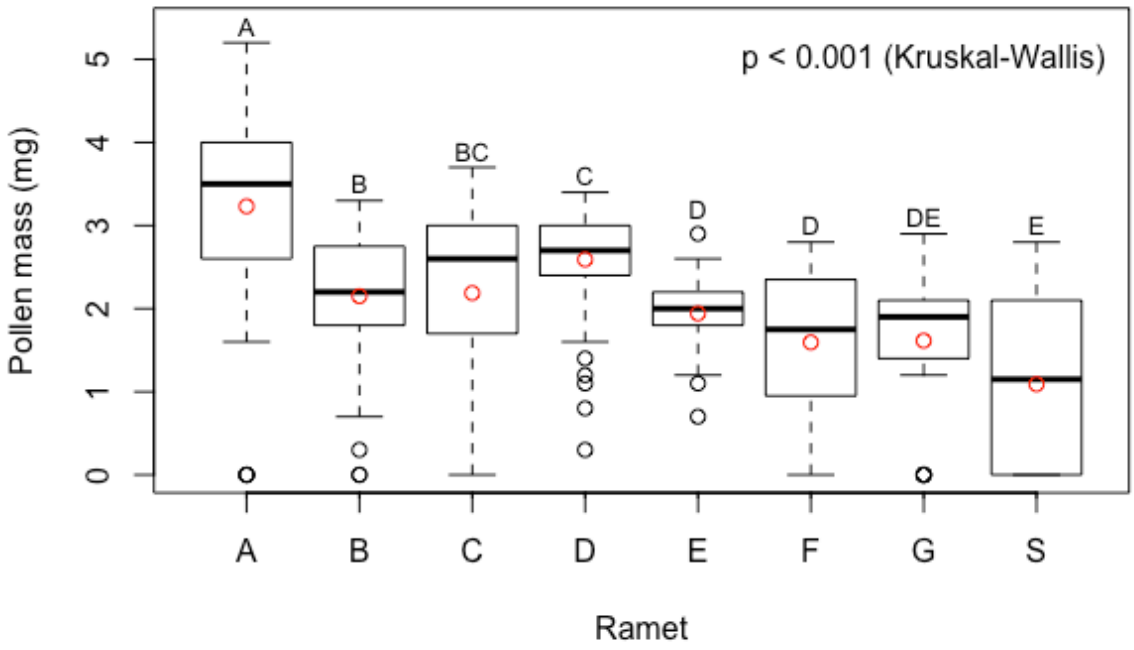


Figure 6: Boxplot showing pollen mass of each ramet. The boxplot shows the median, lower and upper quartiles (25 % and 75 %), lower and upper adjacent values (horizontal line on whiskers), and outliers. Adjacent values are the most extreme values equal to, or within 1.5 x inter quartile range (IQR) from the nearest quartile, where IQR is the difference between the upper and lower quartiles. All observations that fall outside 1.5 x IQR, are outliers (Rstudio standard settings (RStudioTeam 2016)). Red circles show mean value. Boxplots topped with different letters indicate significant differences ($P < 0.05$, Pairwise Wilcoxon rank sum test).

The seasonal differences of pollen mass were found to be statistically significant, decreasing throughout the season (Figure 7). Seasonal differences were also investigated by plotting pollination day (i.e. Anthesis start) against pollen mass for different ramets, as seen in Figure 8. Ramet B, D, F and G showed a statistically significant correlation between pollination day and pollen mass, with decreasing pollen mass throughout the flowering season.

All analyses of pollen mass were conducted using the complete dataset from the study population, including closed flowers and open flowers without pollen. As most of the pollen samples without pollen belonged to ramet C and its side-ramets, their boxplots in particular would be higher in Figure 6, if those pollen samples were not included. As for seasonal variation, another analysis was conducted without closed flowers and open flowers without pollen. This was done in order to know whether pollen mass per flower actually decrease on average throughout the season, or if the reductions in average pollen mass per flower is

caused merely by flowers without pollen. These results are shown in Appendix 2 and 3 and reveal that pollen mass is not significantly different between the midseason and late in the season. The average mean pollen mass in the first part of the season is 2.6 mg and in the late season 1.9 mg. This reduction in mean average pollen mass per flower is less than when samples without pollen is included, where the mean average pollen mass per flower decreases from 2.3 mg to 1.4 mg. As for individual ramets, the results are fairly similar, with ramet B, D, F and G still having a significant decrease in pollen mass as the season progress. Overall, Appendix 2 and 3 shows that a certain proportion of the seasonal reduction in pollen mass throughout the season are caused by samples with no pollen, but it is still decreasing nevertheless.

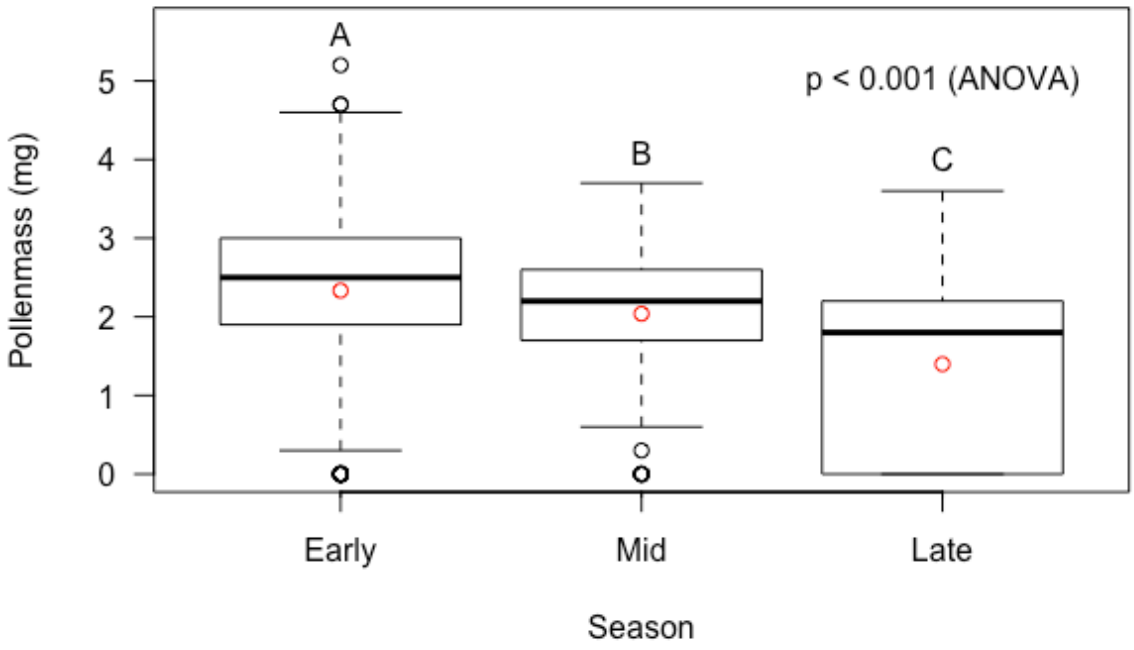


Figure 7: Boxplot showing pollen mass throughout the flowering season for all pollen samples. Boxplot parameters explained in Figure 6. Boxplots topped with different letters indicate significant differences ($P < 0.05$, Tukey).

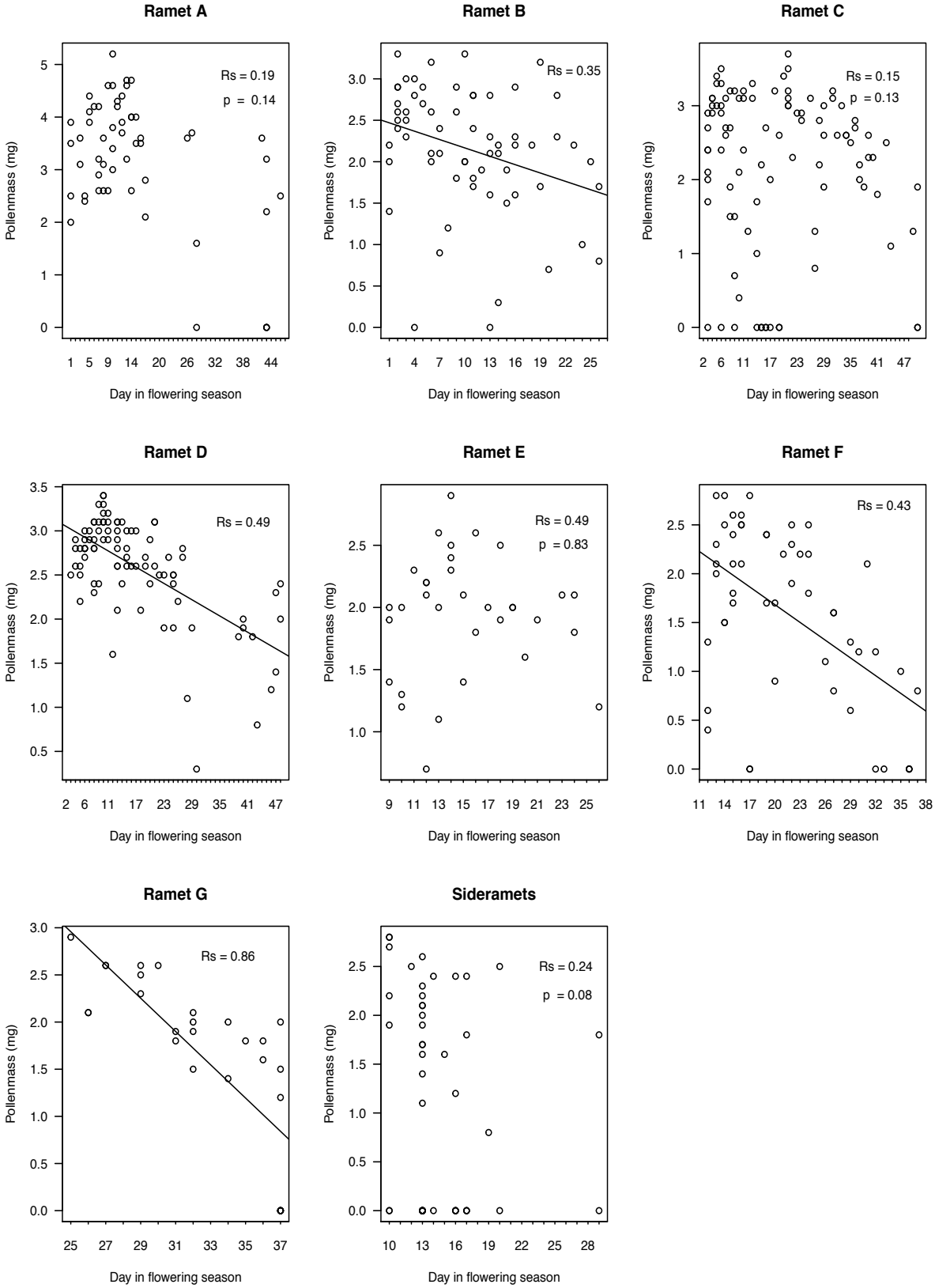


Figure 8: Scatterplots of pollination day on the x-axis, i.e. the day in the flowering season where a flower would open, and pollen mass on the y-axis. Pollination day are population specific, with day 1: 11.07.2017 and the last day (day 48): 27.08.2017. Lines indicate significant correlation, $P < 0.05$. Spearman rank correlation coefficient is reported as Rs. Note different scale on the vertical axis for different ramets.

3.3 Pollen diameter

As with pollen mass, the diameter of pollen differed among ramets, and a pairwise Wilcoxon test showed that 19 out of 36 ramet interactions were statistically different. 14 of those, are between ramet A and ramets B-S, and ramet K and ramets B-S, as seen in Figure 9. Results for ramet K are based on fewer diameter measurements than the study population (see section 2.4).

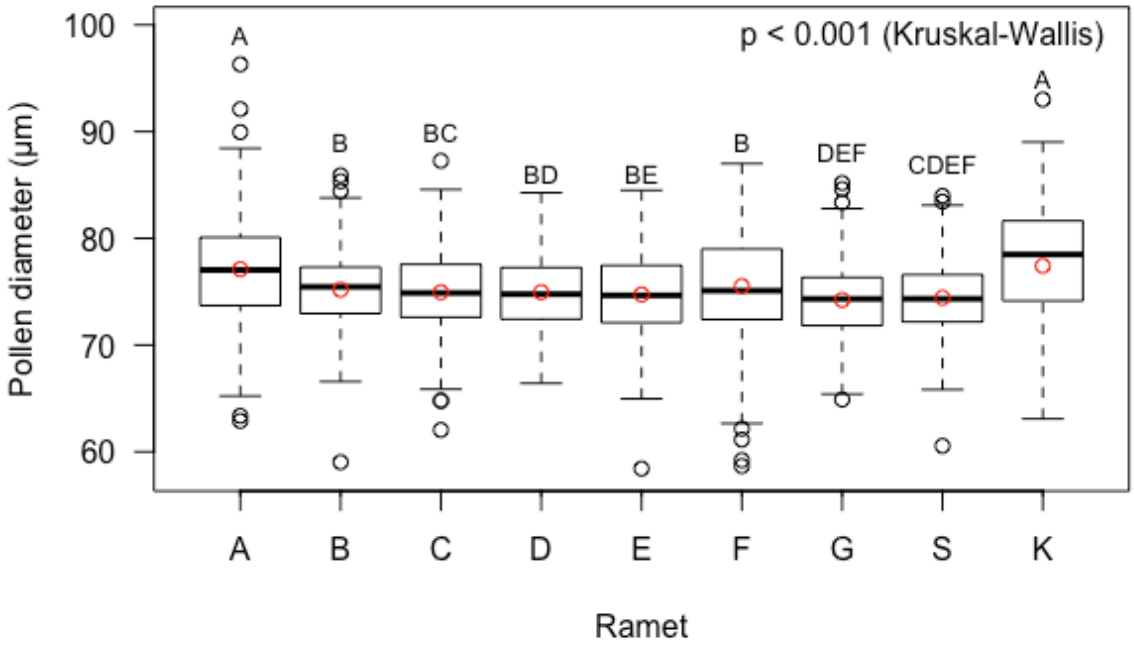


Figure 9: Boxplot showing pollen diameter for ramets. Boxplot parameters explained in Figure 6. Boxplots topped with different letters indicate significant differences (P < 0.05, Pairwise Wilcoxon rank sum test).

Pollen diameter was significantly different between the seasonal parts and decreasing as the flowering season progressed (Figure 10).

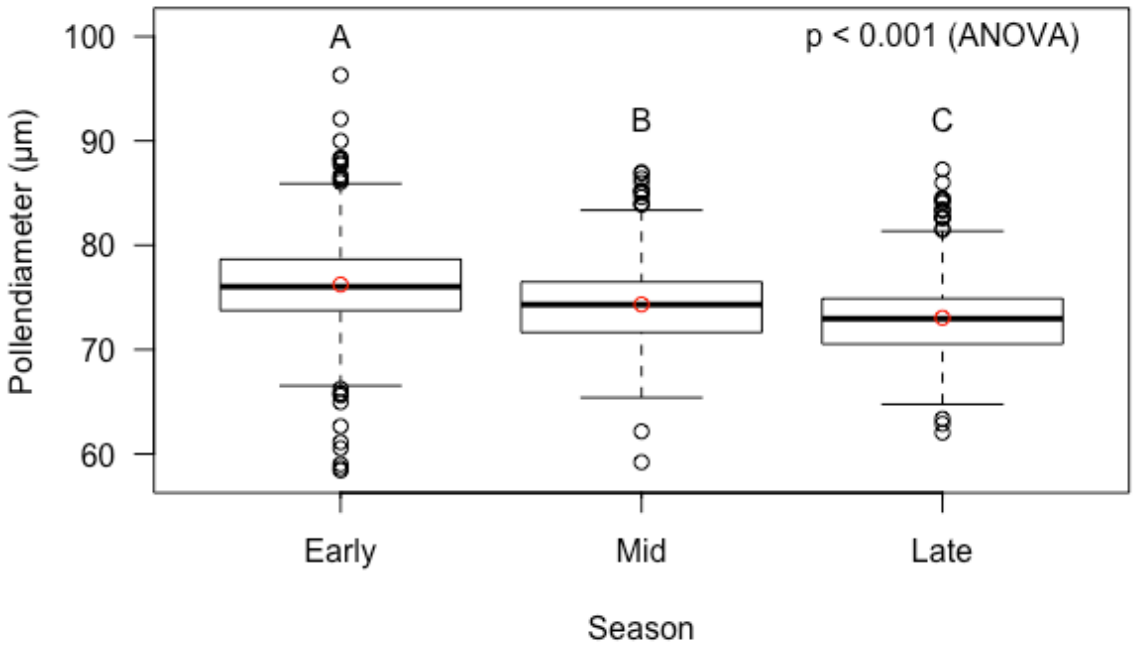


Figure 10: Boxplot showing pollen diameter throughout the flowering season. Boxplot parameters explained in Figure 6. Mean values were the following; 76.2 µm (Early), 74.3 µm (Mid) and 73.0 (Late). Boxplots topped with different letters indicate significant differences (P < 0.05, Tukey).

Finally, seasonal differences in pollen diameter for each individual ramet were investigated (i.e. pollination sequence), by grouping the five first, middle five and last five pollen samples. Figure 11 shows that only ramets A, B and D show statistically different pollen diameter between the seasonal parts when decreasing throughout the season. It should be noted however, that some ramets flowered throughout most of the flowering season of the study population, while others flowered only for a few days (e.g. ramet G and side-ramets, see Table 4 and Appendix 1) and is therefore not expected to show clear relationships.

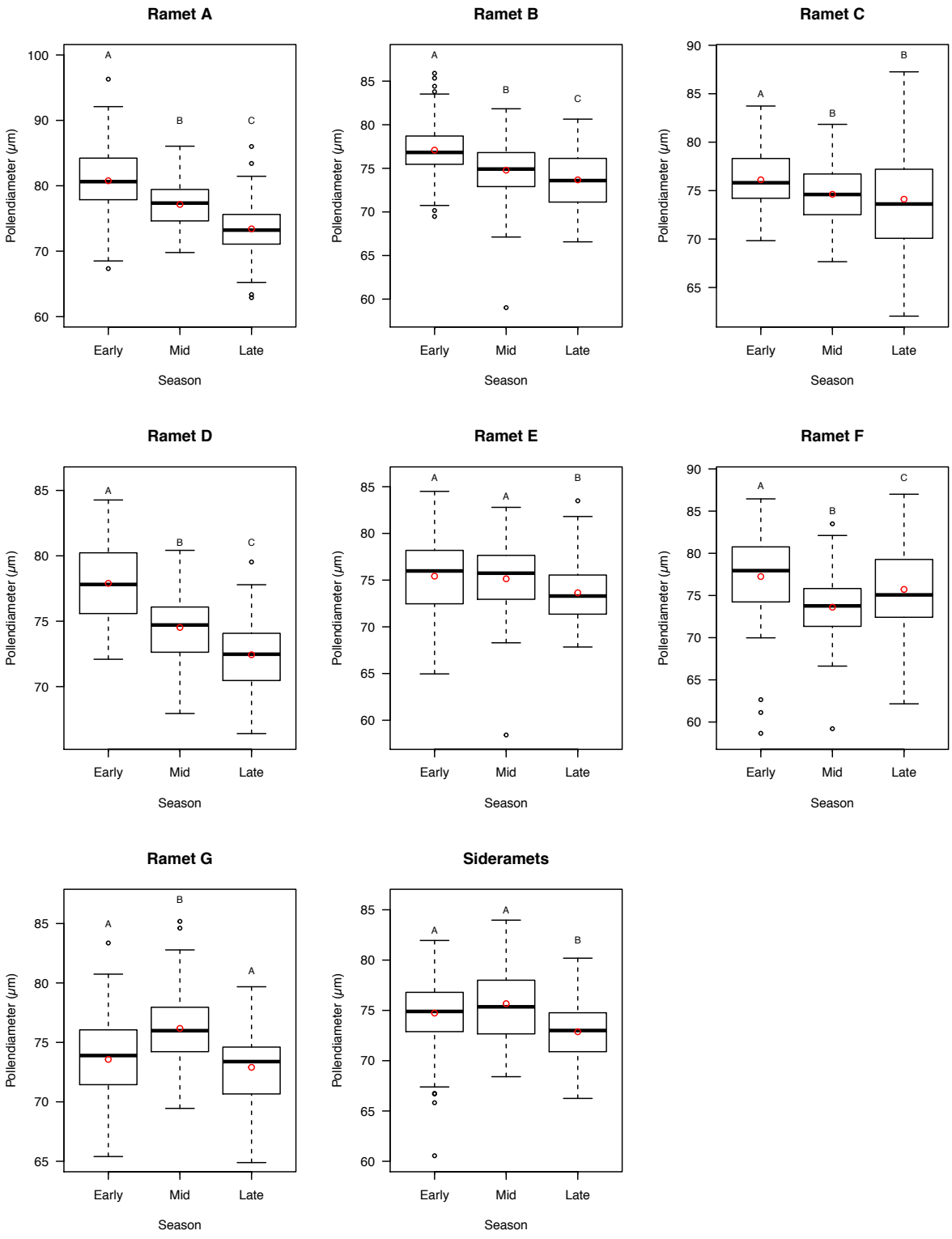


Figure 11: Seasonal differences in pollen diameter for each individual ramet, i.e. showing the effect of pollination sequence. Boxplot parameters explained in Figure 6. Different letters indicate significant differences ($P < 0.05$, Tukey) for ramets B, D and G, ($P < 0.05$, Pairwise Wilcoxon rank sum test) for ramets A, C, E, F and S. Note different scales on the vertical axis for different different ramets.

3.4 FTIR results

The vibrational spectrum of a representative pollen sample shows peaks at spectral bands that are associated with different group of chemical compounds (Figure 12). It should be noted that vibration types are absorbed at different frequencies (spectral bands), e.g. C=O stretch are found in different chemical compounds; triglycerides, free fatty acids, and proteins, but are assigned to functional group depending on the frequency in which it is absorbed (Bağcıoğlu 2016 ,p. 13). The peak at 1745 cm^{-1} is associated with C=O stretching vibrations and is characteristic for lipids. The spectral region between 900-1200 cm^{-1} contains the vibrational bands that are associated with carbohydrates. These vibrational bands include the C-O-C stretch, C-OH stretch, C-O-H deformation, C-O-C deformation, and the pyranose and furanose ring vibrations. The vibrational bands that are associated with proteins are found in 1700-1630 cm^{-1} range (C=O stretching, amide I), and at the 1550-1520 cm^{-1} range (C-N stretching and NH deformation, amide II) (Bağcıoğlu et al. 2015; Gottardini et al. 2007; Pappas et al. 2003; Zimmermann 2010; Zimmermann & Kohler 2014; Zimmermann et al. 2015)

In addition to carbohydrates, lipids, protein, water and nucleic acids, pollen also contain sporopollenins in the outer pollen wall (exine) (Zimmermann 2010). Sporopollenins are a resilient group of biopolymers based upon phenylpropanoid building blocks (Blokker et al. 2006), and thus some of the bands in their vibrational spectra are related to aromatic rings (Zimmermann 2010; Zimmermann & Kohler 2014). Table 6 provide an overview of what functional groups and chemical components the different spectral bands represent. Signals designated in Table 6 in which no range is given should be referred to as narrow spectral regions and can vary by a few wavelengths.

Table 6: Spectral bands characteristic for FTIR, and associated chemical structures (Bağcıoğlu 2016, pp. 14-16; Bağcıoğlu et al. 2017).

Compound	Spectral band (cm^{-1})	Vibration type	Biochemical group
Lipids	1745	C=O str	Triglycerides
	1705	C=O str	Free fatty acids
	1240	P=O asymmetrical str	Phospholipids
	1090	P=O symmetrical str	
Proteins	1700-1630	C=O str (Amide I)	Amino acids
	1550-1520	NH deformation and C-N str (Amide II)	Amino acids
Carbohydrates	1200-900 (region)	C-C str, C-O str, C-O-H deformation, C-O-C deformation, pyranose and furanose rings	
	1107		Pyranose
	1055		
	1028		
	1076		Amylose
	995		
Sporopollenin	1605	Aromatic ring vibrations	Phenylpropanoids
	1515		
	1171		
	853		
	833		
	816		

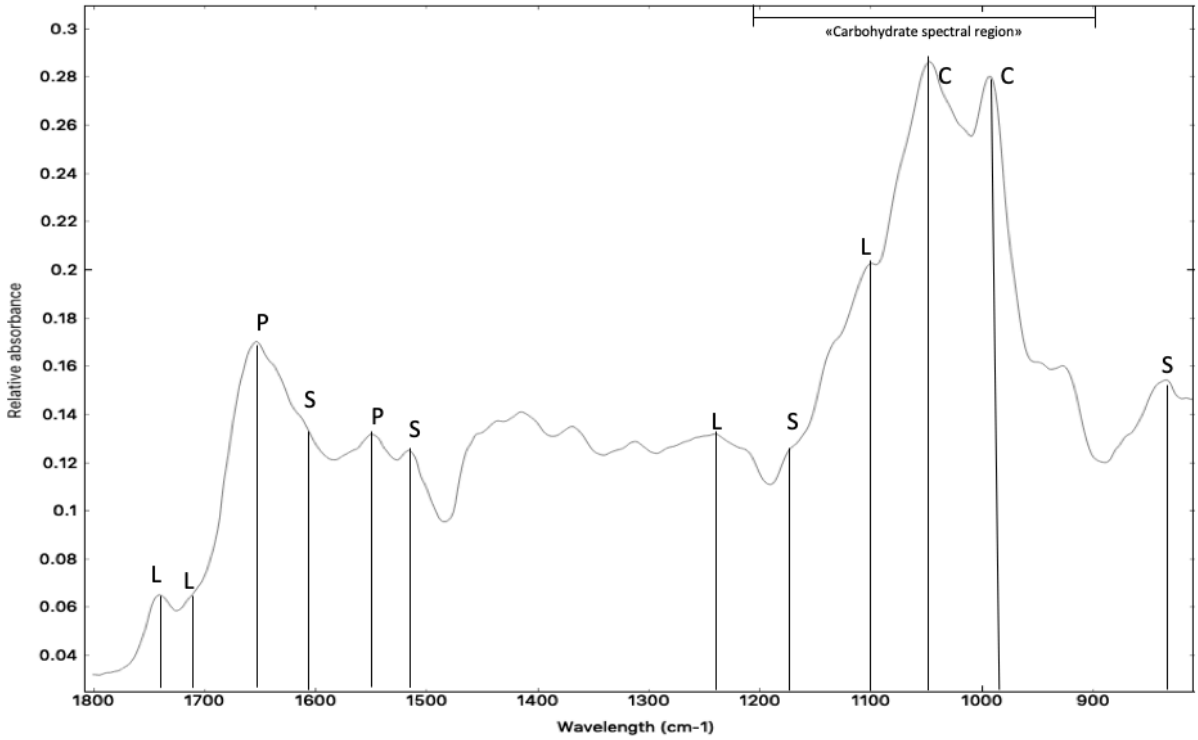


Figure 12: FTIR vibrational spectra from one representative sample of *C. angustifolium*. X axis show wavelengths in the spectral region from 1800 to 800 cm^{-1}. Y axis denote relative absorption. Characteristic spectral bands (Table 6) are marked, and related to lipids (L), proteins (P), Sporopollenin (S) and Carbohydrates (C).

3.5 Principal component analysis

The PCA results did not show any pattern of clustering or deviance in either preprocessing procedure, with regards to seasonal differences. As for ramet differences, the most noticeable result is the relatively high PC1 scores of ramet D, and high PC2 scores of samples belonging to the Kittilbu population (Figure 13). The loading plot for PC1 in Figure 14, shows negative loadings in the "carbohydrate" spectral region at 1048 and 995 cm^{-1}, and positive loading for proteins at 1672 cm^{-1} and lipids at 1707 and 1090 cm^{-1} bands. Furthermore, positive loadings at 1607,1510,1173 and 840 cm^{-1} indicate relatively lower nutrient content or relatively higher sporopollenin content, as these bands are characteristic for sporopollenin. When interpreting the loading plots together with PC scores, spectral features of the loading plots can explain chemical differences between pollen samples in the score plots. It must be emphasized that the PCA plots can be related only to relative ratio of chemical components, and not absolute values representing actual content of these chemical components. This means that ramet D with high PC1 scores (Figure 13) have a high ratio of proteins and lipids to carbohydrates (Figure 14), but whether the actual protein and lipid content is high or carbohydrate content is low, remains unknown. Similarly, high peaks at sporopollenin bands for the PC1 loading does not tell whether sporopollenin content is high or nutrient content low, but only that the sporopollenin-to- nutrient ratio is high.

The loading plot for PC2 (Figure 15) shows high values in the "carbohydrate" spectral region, both negative (990 cm^{-1}) and positive (1107,1079 and 1028 cm^{-1}). As the positive loadings at 1107 and 1028 cm^{-1} are associated with cellulose (Table 6), the PC2 loading plot could indicate that the Kittilbu population have a higher relative cellulose content compared to the study population. Amylose bands at 1079 and 990 cm^{-1} are more difficult to interpret as they represent both the highest and lowest loadings. They suggest however, that the relative amylose content of the Kittilbu population differs from the study population. For ramet differences, the second preprocessing procedure with second derivatives in addition to MSC, showed analogous results as the first preprocessing procedure for all PCs and loading plots.

In order to visualize the effect of PC1 more clearly, pollen samples with the high and low PC1 scores that had similar PC2 scores was selected (Figure 16). The same method was used for PC2 as well (Figure 17). Figure 16 show the same trend as the PC1 loading plot but also illustrates how the chemical spectra of pollen with low PC1 score looks like, which is most

common in the study population. As for PC2, Figure 17 shows another high value associated with cellulose at 1057 cm^{-1}, compared to the PC2 loading plot.

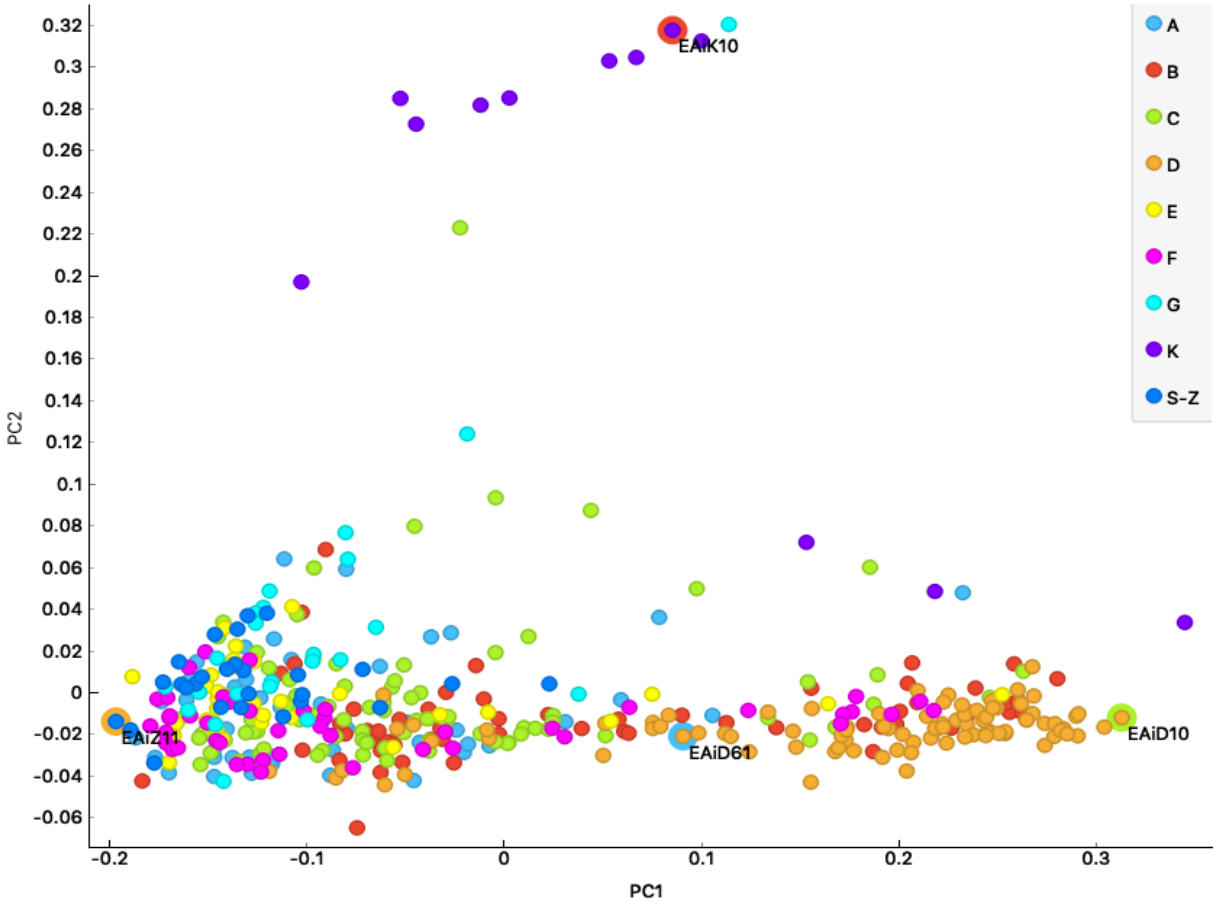


Figure 13: PCA score plot of the FTIR data. PC1 are shown on the x-axis, and the PC2 on the y-axis. Color codes denote ramet membership, where S-Z: side-ramets, and K: Kittilbu population. Explained variance (PC1: 80.1, PC2:9.9). Marked pollen samples were selected for comparison spectra's in Figure 16 and 17.

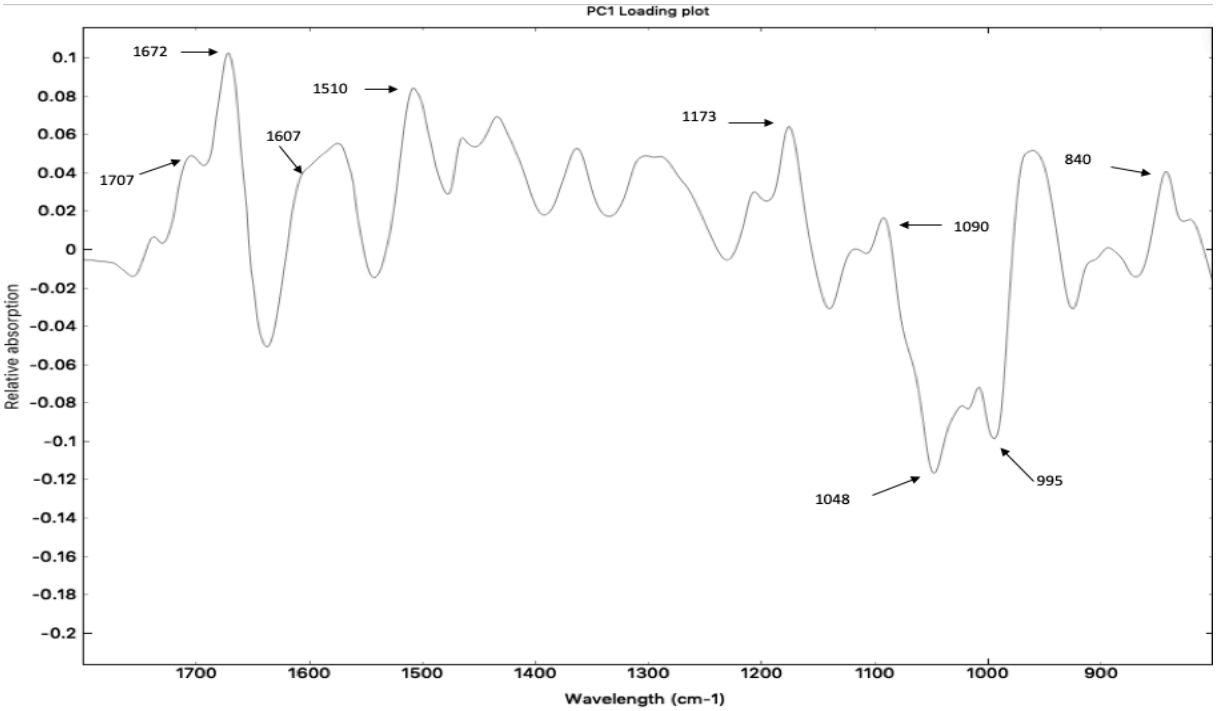


Figure 14: PC1 loading plot (80.1% explained variance), with wavelength on the x-axis, and relative absorbance on the y-axis.

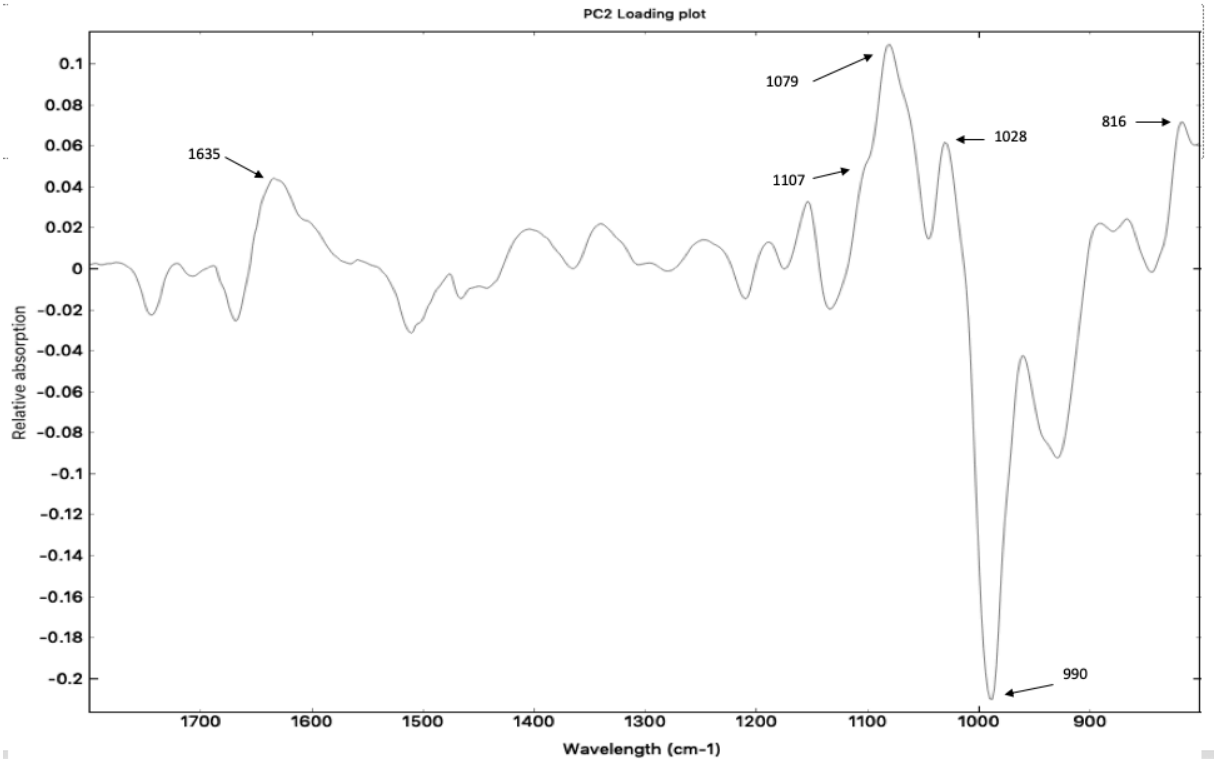


Figure 15: PC2 loading plot (9.9 % explained variance), with wavelength on the x-axis, and relative absorbance on the y-axis.

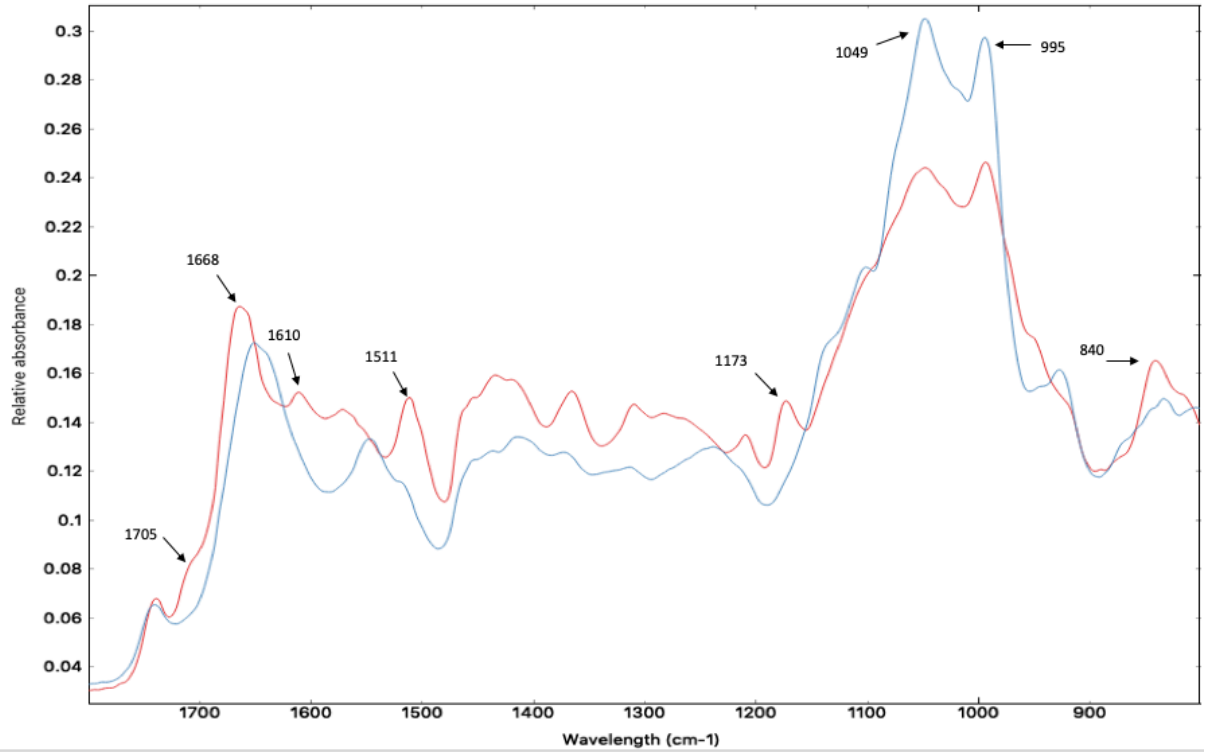


Figure 16: Spectral bands for PC1, with wavelength on the x-axis, and relative absorbance on the y-axis. Red spectra show high PC1 scores, whereas blue spectra show low PC1 scores. Spectral bands characteristic for functional groups marked.

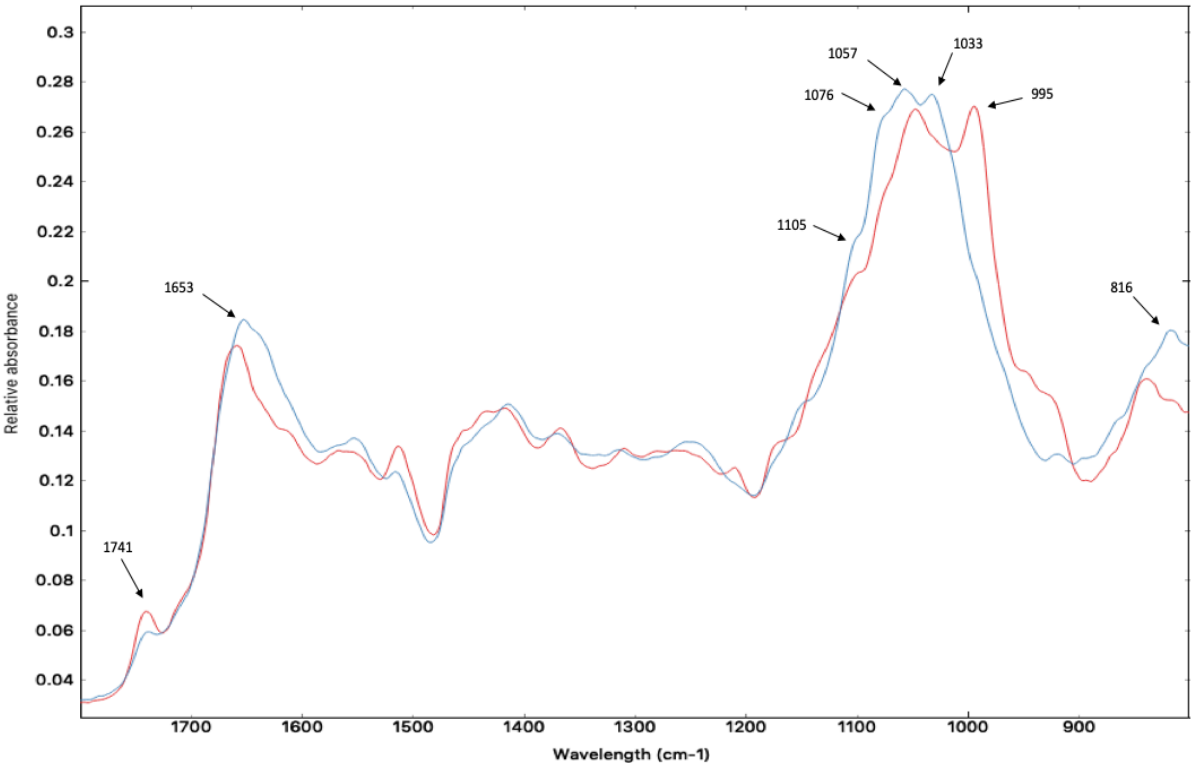


Figure 17: Spectral bands for PC2, with wavelengths on the x-axis, and relative absorption on the y-axis. Blue spectra representing high PC2 score, red spectra show low/normal PC2 score. Spectral bands characteristic for functional groups marked.

The Kittilbu population also stands out in PC3 and PC4 score plots and loading plots, found in Appendix 4-7. The loading plot for PC3 in Appendix 5 show clear negative loadings for 1547 and 1640 cm^{-1} which are associated with proteins. Furthermore, positive loadings are found at 1173 and 1020 cm^{-1} which are associated with sporopollenin and carbohydrate content, respectively. The Kittilbu population is also markedly different for PC4, with great variation of both negative and positive scores (Appendix 6). The accompanying loading plot in Appendix 7 indicate positive loadings for spectral bands associated with protein and negative loadings for bands in the carbohydrate spectral region. This points to opposite results of the PC2 loading plot (Figure 15), but since most Kittilbu samples have negative PC4 scores and not positive, PC2 results are still relevant. Overall, score plots and loading plots of PC3 and PC4 shows similarly to PC2, that the Kittilbu population differs from the study population in terms of higher carbohydrate-to-protein ratio, and that the Kittilbu population also differs slightly in terms of relative sporopollenin content.

3.6 Classification analysis

The results of the classification analyses show that ANN algorithms were able to classify both individual ramets, and seasonal parts with high success rates. Classification for ramet identification of spectral data by ANN are shown in Figure 18 and are based on a training dataset of 140 samples. Success rates for ramet identification are shown in the confusion matrix in Table 7 (blue highlight). The confusion matrix also highlights percentage of misclassified samples (red highlight), e.g. for ramet A, 88.2 % are classified correctly, while 5.9 % of the A samples were misclassified as B, and the same percentage misclassified as C. The classification results in Figure 18 and Table 7 are obtained by resampling several times, where particularly ramet A and D had consistently higher classification rates, compared to the other ramets. Ramet B, G, side-ramets (S-Z) and K showed large variation in classification success. Especially for K this is expected, since it only had 4 samples in the validation set. Randomization resulted in very low classification success, indicating that the original classification models are not overfitted. As the validation dataset is relatively large (1/3 of all samples, at n=140), the results are robust and indicate that the ANN algorithms are able to distinguish between ramets based on chemical differences. The other learning algorithms tested, were not able to differentiate between ramets as well as ANN, although random forest classification had similar misclassification rate.

The classification analysis was also conducted for seasonal variation of spectral data, and as with ramet identification, ANN algorithms was the only learning method capable of successful classification. Randomization resulted in low classification success, thus demonstrating that the original classification is not overfit. The results are shown in Figure 19 and illustrates that there are chemical differences regarding the pollination sequence of ramets.

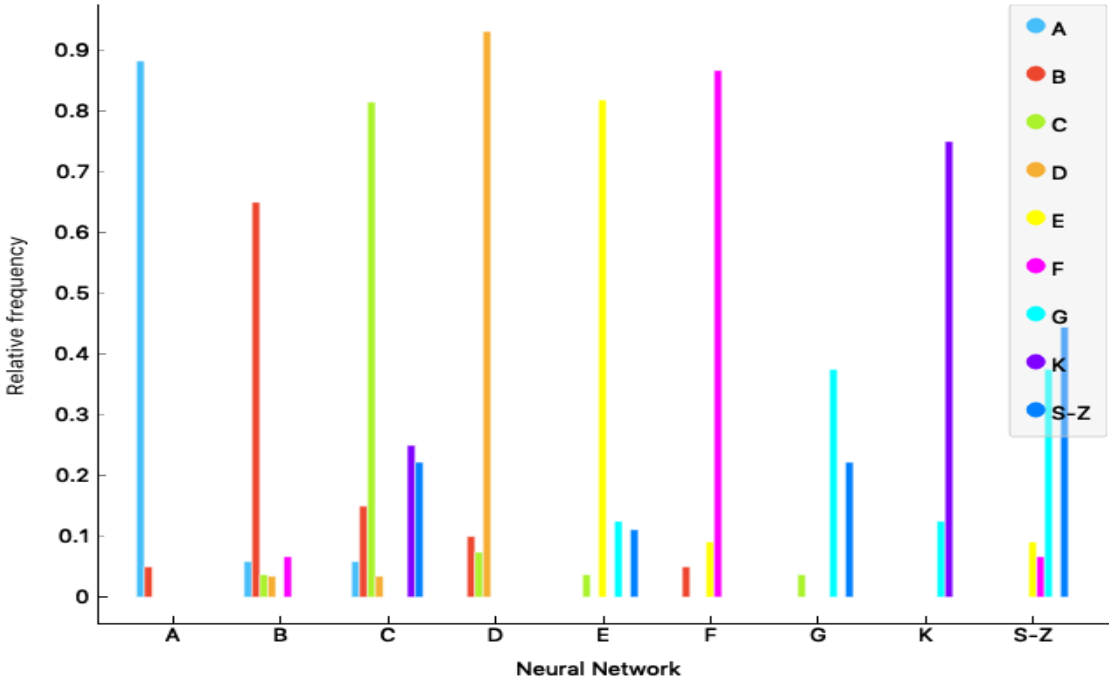


Figure 18: Classification results of ramet differences by the ANN learning algorithms. Color codes denote ramet membership, and the y-axis indicate classification success.

Table 7: Confusion matrix on ramet differences by the ANN learning algorithms. Values with blue background indicate percentage of correctly classified samples from the validation dataset, n=140. Values with red background highlight misclassification. Σ refer to the number of samples from different ramets that are used in the validation set, and corresponds to one third of the number of samples in the original dataset. This proportion is similar for all ramets.

		Predicted									
		A	**B**	**C**	**D**	**E**	**F**	**G**	**K**	**S-Z**	**∑**
Actual	**A**	88.2 %	5.9 %	5.9 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	**17**
	B	5.0 %	65.0 %	15.0 %	10.0 %	0.0 %	5.0 %	0.0 %	0.0 %	0.0 %	**20**
	C	0.0 %	3.7 %	81.5 %	7.4 %	3.7 %	0.0 %	3.7 %	0.0 %	0.0 %	**27**
	D	0.0 %	3.4 %	3.4 %	93.1 %	0.0 %	0.0 %	0.0 %	0.0 %	0.0 %	**29**
	E	0.0 %	0.0 %	0.0 %	0.0 %	81.8 %	9.1 %	0.0 %	0.0 %	9.1 %	**11**
	F	0.0 %	6.7 %	0.0 %	0.0 %	0.0 %	86.7 %	0.0 %	0.0 %	6.7 %	**15**
	G	0.0 %	0.0 %	0.0 %	0.0 %	12.5 %	0.0 %	37.5 %	12.5 %	37.5 %	**8**
	K	0.0 %	0.0 %	25.0 %	0.0 %	0.0 %	0.0 %	0.0 %	75.0 %	0.0 %	**4**
	S-Z	0.0 %	0.0 %	22.2 %	0.0 %	11.1 %	0.0 %	22.2 %	0.0 %	44.4 %	**9**
	∑	**16**	**17**	**30**	**31**	**12**	**15**	**6**	**4**	**9**	**140**

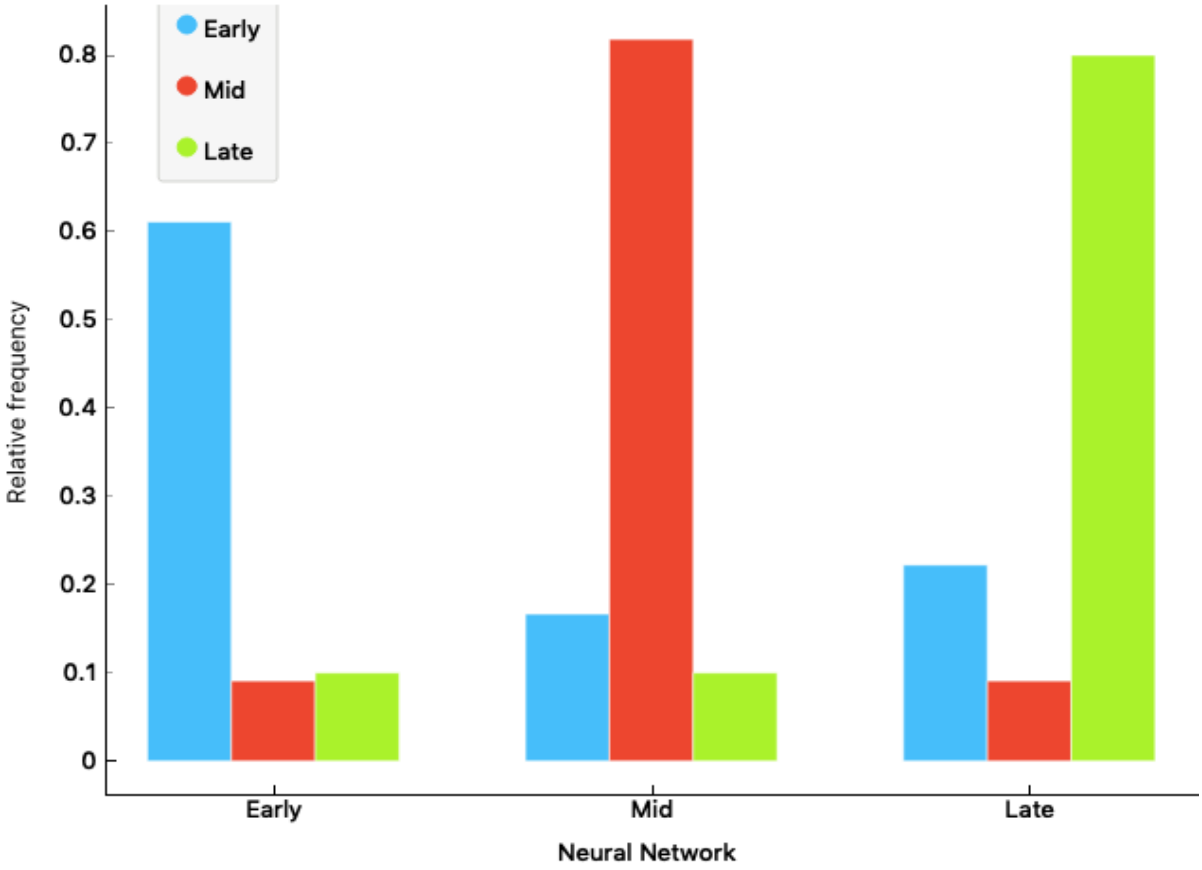


Figure 19: Classification success of seasonal variation by ANN algorithms. Color codes denote Seasonal part, and the y-axis indicate classification success.

Table 8: Confusion matrix of seasonal variation by ANN algorithms. Values with blue background indicate percentage of correctly classified samples from the validation dataset, n=39. Values with red background highlight misclassification.

		Predicted			
		Early	**Mid**	**Late**	**∑**
Actual	**Early**	61.1 %	16.7 %	22.2 %	**18**
	Mid	9.1 %	81.8 %	9.1 %	**11**
	Late	10.0 %	10.0 %	80.0 %	**10**
	∑	**13**	**13**	**13**	**39**

3.7 Partial least squares regression

Partial least squares regression (PLSR) results shows that the chemical composition of pollen is correlated with the day in the flowering season (Table 9 and Figure 20), flowering sequence (Appendix 8 and 9), and pollen mass (Table 10 and Figure 21). Tables 9, 10 and Appendix 8 show R^2 (correlation coefficients) of the fitted model and predicted R^2 after a 10 segmented cross-validation, and root mean squared error (RMSE) of the model before and after cross-

validation. For pollination day and pollination sequence, both Table 9 and Appendix 8 show the number of pollen samples analyzed by PLSR, pollination period in days and the number of days pollination occurred. The PLSR results in Table 9, 10 and Appendix 8 shows results for both all the ramets in the study population and individual ramets. It should be noted that the row "all" include all ramets, n =415, and that only one out of the five side-ramets (S, n=10) could be analyzed separately due to small number of pollen samples belonging to the other side-ramets. The Kittilbu population was found to be outliers compared to the study population, and thus removed from the rest of the analysis.

The predicted $R^2 = 0.51$ for pollination day of all ramets (Table 9) indicate that pollen chemistry shares 51 % of the variation in pollination day. Predicted RMSE is 7.38 for all ramets, i.e. the model is able to predict pollination day with chemical spectra, with a mean error of roughly 7 days. When comparing this to the pollination period of the study population which is 48 days, the model seems to offer some predictive ability. As can be seen in Table 9 10 and Appendix 8, PLSR results from individual ramets shows that R^2 and RMSE differs between them. Large discrepancies between model and prediction after cross-validation, indicate overfitting for some of the ramets models.

For pollination sequence of the population, the predicted $R^2 = 0.43$ and predicted RMSE is 7.65, thus showing comparable but slightly weaker predictive ability, compared to pollination day. This is expected however, as pollination day of the study population and pollination sequence of individual ramets are highly correlated (Table 4, Appendix 1). The spectra were also correlated to pollen mass of the population, with predicted $R^2 = 0.35$ and predicted RMSE = 0.63. Table 10 shows the average pollen mass for different ramets, plus standard deviance, minimum and maximum values. R^2 and RMSE differed considerable between ramets, as seen in Table 9 and 10.

Table 9: The PLSR coefficient of determination (R^2) and root mean squared error (RMSE) for pollination day based on chemical spectra. m: model, p: prediction. $A_{opt:}$ optimal number of components. S: side-ramet S, n=10 (not all side-ramets).

Ramet	Number of pollen samples analyzed by FT-IR	Pollination period (days)	Days with pollination	R^2 (m)	R^2 (p)	RMSE (m)	RMSE (p)	A_{opt}
All	415	48	46	0.68	0.51	5.96	7.38	13
A	51	46	22	0.72	0.50	5.59	7.70	6
B	62	26	24	0.64	0.46	4.18	5.26	4
C	84	46	39	0.75	0.70	2.67	6.79	12
D	90	46	35	0.96	0.68	2.30	6.66	13
E	33	18	16	0.92	0.54	1.24	3.40	7
F	47	25	22	0.81	0.39	3.03	5.71	5
G	23	13	11	0.99	0.42	0.23	2.85	11
S	10	7	3	0.85	0.29	0.72	1.79	3

Table 10: The PLSR coefficient of determination (R^2) and Root mean squared error (RMSE) for pollen mass based on chemical spectra, and data on pollen mass (min, max, average, Sd: standard deviation), m: model, p: prediction. $A_{opt:}$ optimal number of components. S: side-ramet S, n=10 (not all side-ramets).

Ramet	Min	Max	Average	Sd	R^2 (m)	R^2 (p)	RMSE (m)	RMSE (p)	A_{opt}
All	0.3	5.2	2.4	0.8	0.46	0.35	0.58	0.63	8
A	1.6	5.2	3.5	0.8	0.10	0.05	0.76	0.82	1
B	0.3	3.3	2.2	0.7	0.09	NA	0.62	0.67	1
C	0.4	3.7	2.5	0.7	0.78	0.31	0.32	0.59	9
D	0.3	3.4	2.6	0.6	0.69	0.18	0.31	0.51	7
E	0.7	2.9	1.9	0.5	0.16	0,00	0.43	0.49	1
F	0.4	2.8	1.8	0.7	0.74	0.73	0.32	0.34	1
G	1.2	2.9	2.0	0.4	0.60	0.49	0.27	0.32	2
S	1.2	2.8	2.0	0.6	0.76	0.40	0.26	0.47	2

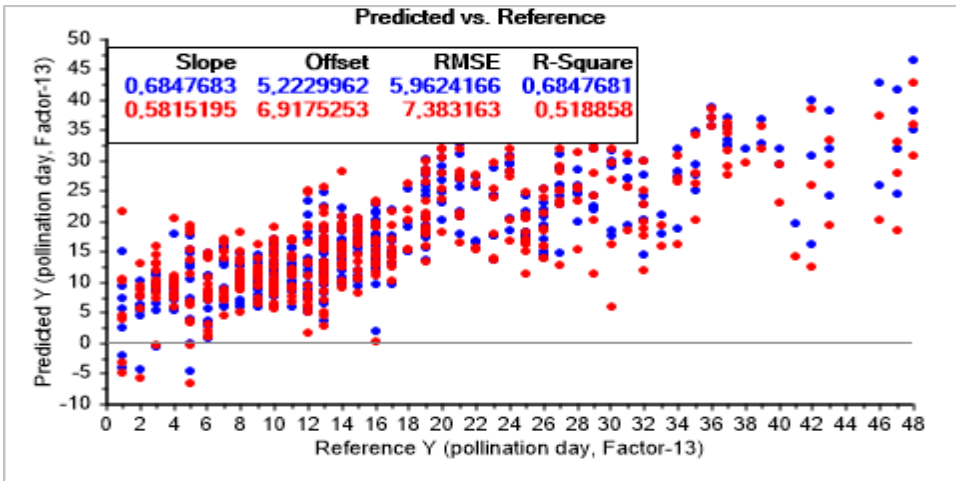


Figure 20: PLSR response plot of pollination day for all pollen samples. Actual pollination day is shown on the x-axis and predicted pollination day on y-axis. Blue dots indicate model fit, whereas red dots indicate model performance after cross-validation.

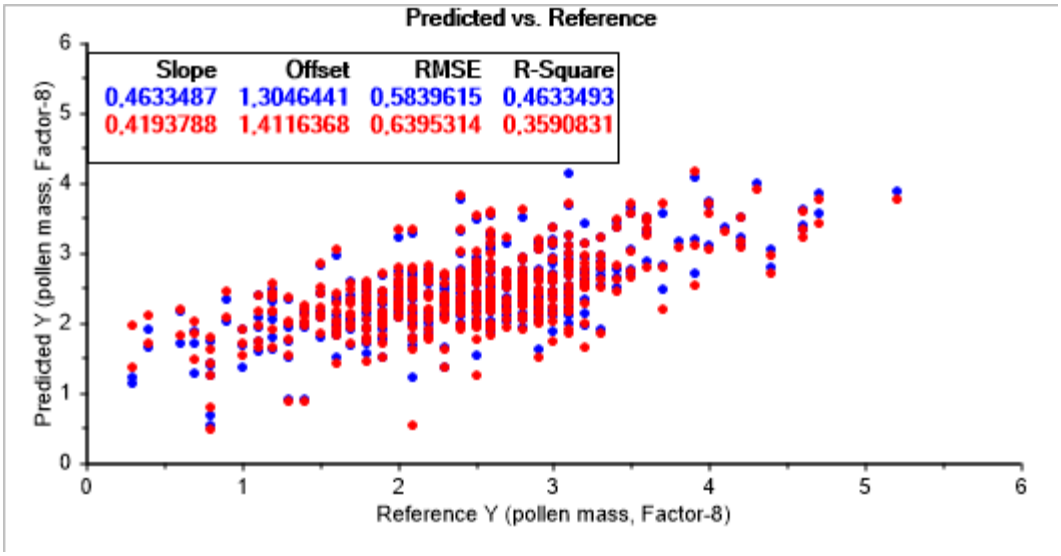


Figure 21: PLSR response plot of pollen mass for all pollen samples. Actual pollen mass is shown on the x-axis and predicted pollen mass on the y-axis. Blue dots indicate model fit, whereas red dots indicate model performance after cross-validation.

4 Discussion

4.1 Pollen size

The results clearly show that in terms of pollen production, ramet C and D stands out with markedly higher production than other ramets (Table 5). This could be due to the fact that ramet C and D flowered for the longest duration of time (Table 4). Additionally, the observed differences in pollen production among ramets could partially be attributed to seasonal differences because the different ramets flowered at different dates (Table 4, Appendix 1). Ramet A, B, C and D all flowered earlier in the season and produced more pollen than the rest of the ramets.

As for pollen mass produced per flower, ramet A produced markedly more than all other ramets (Figure 6). This cannot be attributed to seasonal difference, as ramet A and B had roughly similar flowering periods (Table 4). Ramet A flowered longer, but with very little pollen production after ramet B stopped flowering (Appendix 1).

Although another analysis found that some of the seasonal variation in pollen mass per flower is caused by flowers without pollen (Appendix 2 and 3), pollen mass is still dropping throughout the season. The results show clearly that for the study population as a whole, pollen mass per flower decrease throughout the flowering season (Figure 7). However, this trend is not uniform for all ramets as Figure 8 show that this only applies to ramets B, D, F and G.

Although pollen diameter was also found to be statistically different among ramets, most of the differences are due to ramet A and the Kittilbu comparison population, which stands out from the rest. Figure 10 show that pollen diameter decreases over the flowering season. However, as with pollen mass, the trend is not uniform when analyzing each ramet separately (Figure 11). Only ramet A, B and D shows a clear trend, with all seasonal parts being significantly different and shrinking as the season progress. Nevertheless, all mean values for the late seasonal part are lower than other seasonal parts except for ramet F (Figure 11), indicating that there is a trend of seasonal reductions in pollen diameter for most of the population. Ramet C illustrates this, where the middle and late seasonal part are not significantly different, but the latter have the lowest mean. Overall, although the results do not show an unambiguous trend for all ramets, both pollen mass and pollen size are generally decreasing over the course of the flowering season. This indicates that pollen grain size goes down in the study population, but whether pollen quantity in terms of pollen grain number change, remains unknown.

Pollen quantity (number of pollen grains) is fundamentally important in reproduction, because pollen limitation is more likely to be prevented with high pollen production (Knight et al. 2005), and when pollen competition is higher, genetically favorable pollen is more likely to be selected (Charlesworth et al. 2005; Stephenson et al. 2003). Although pollen quantity should contribute positively to male reproductive success, it does not necessarily have a profound impact alone. *C. angustifolium* have low pollen-ovule ratios (P/O) as compared to other species with equivalent pollination strategy (Cruden & Jensen 1979). Cruden and Jensen (1979) argues that this is due to efficient pollination caused by large stigmas and numerous long viscin threads. The stigmas are large compared to the pollen attaching area on the pollinator, causing better pollen receipt. Viscin threads facilitate clumping of pollen grains, allowing more pollen to be transported. With this in mind, the drop in mean pollen mass of the study population from 2.3 mg early in the season to 1.4 mg late in the season (Figure 7), might not by itself cause considerable fitness reductions.

Some studies have suggested that there could be a trade-off between pollen size and number, by finding inverse relationship between them (Mione & Anderson 1992; Sarkissian & Harder 2001), whereas other studies have not found this relationship (Aguilar et al. 2002; Jürgens et al. 2012). Although pollen mass and diameter are both decreasing throughout the flowering season in this study, pollen mass are reduced the most in relative terms. This could represent a trade-off, where *C. angustifolium* prioritize pollen size over number.

It is, however, reasonable to believe that in addition to pollen quantity, pollen size is of importance for male reproductive success, since survival, germination and pollen tube growth all depend on material and energy reserves in the pollen grain (Williams & Mazer 2016). Larger pollen grains are expected to have higher fertilization success due to the higher storage capacity of materials and energy (Baker & Baker 1979). Pollen grains of larger size are also shown to outperform smaller grains during sexual competition in the style, and sire more seeds (McCallum & Chang 2016).

Large pollen grain size and large energy reserves could be particularly important for *C. angustifolium* because of its long style (stigma-ovule distance) of approximately 10 mm during anthesis (Myerscough 1980). Baker and Baker (1979) compared 147 species with style lengths of 10 mm or longer, to 843 species of shorter style lengths and found that mean diameter was significantly larger for longer styles, irrespective of chemical composition. Moreover, Roulston et al. (2000) found pollen volume and style length to be significantly correlated for more than 50 plant families. However, although nutrients play a critical role in early stages of pollen tube growth, pollen receive energy from stylar tissue as well in later stages (Baker & Baker 1979; Stephenson et al. 2003). This energy is acquired through genetic interactions between the gamete in pollen and the pistil (Stephenson et al. 2003). Differential provisioning of pollen grains by stylar tissue represent "female choice", in which preferable male gametes are provided more nutrients and are thus more likely to be successful in reproduction, under pollen limitation (Bernasconi et al. 2004; Marshall & Ellstrand 1986). The effect of female choice is particularly prevalent for distinguishing between cross- or self-pollen (Charlesworth et al. 2005). In contrast to Baker and Baker (1979), Cruden and Lyon (1985) therefore suggested that the distance that the pollen tube have to grow in order to reach stylar tissue (stigma depth) was functionally related to pollen size as opposed to style length. Nevertheless, pollen grains of larger size are more likely to lead to higher male fertilization success (Baker & Baker 1979; McCallum & Chang 2016)

Even small differences in pollen diameter can severely affect pollen viability. Kelly et al. (2002) found that the mean difference between viable and inviable pollen grains in *Mimulus guttatus* was 13 μm (41.9 μm and 28.9 μm). Despite overlapping between the groups, they found a critical distinction at 35 μm, where most pollen grains > 35 μm were viable whereas pollen grains < 35 were not. Their results showed that at some parts of the distribution, a difference of less than 5 μm could considerably affect viability. The implications of this should not be

overstated though, as the study is only somewhat comparable. *C. angustifolium* have different ecology, pollen grains are generally more than twice the size (Halbritter 2017), and the difference in pollen grain diameter between the early and late seasonal part is only 3.2 μm (Figure 10). A change in 3.2 μm is percentagewise a much smaller difference for *C. angustifolium* (4%), compared to *M. guttatus*. Nevertheless, as the pollen size in this study was larger early in the flowering season, early onset of flowering could be beneficial for male reproductive success.

4.2 Pollen chemistry

The classification of ramet membership shows that pollen chemistry varies between ramets. However due to the nature of ANN algorithms, we cannot connect spectral (i.e. chemical) information with the classification, hence it is not known how the ramets differ in their chemical properties. It is possible that ANN is detecting chemical differences among genetically different pollen grains. During development, formation of pollen depends both on sporophytic and gametophytic genome, and, since individual pollen grains are genetically unique, they are also chemically different (Piffanelli et al. 1998). However, since pollen were measured in bulk in this study, chemical differences will be averaged, so that genetic differences among individual pollen grains does not contradict the classification result.

The PCA results shows that ramet D and the Kittilbu population K are chemically different from the other ramets. The fact that ramet D have on average higher PC1 scores than the other ramets in the study population, could suggest that it is genetically different. Ramet D is also arguably the ramet with highest total fitness, as all 90 pollen samples produced sufficient pollen with little deviance. Ramet C with side-ramets produced had more continuous pollen production through the season and higher pollen production, but 27 % of pollen samples failed (Table 2). The notion that different ramets of the study population could be genetically different is probable also due to the classification results, differences in total pollen production, pollen production per flower and pollen size differences. However, as no genetic test was conducted, this remains unknown.

The uncertain implications of the classification results, the fact that the PC1 scores are gradual, and show that ramet D also have low PC1 scores for some samples, and that ramet B, C and F also contain samples with high PC1 scores, suggest oppositely that individual ramets might not be genetically different.

Most ramets except for ramet D have low PC1 scores, and thus a relatively higher ratio of carbohydrates to lipids and proteins. Carbohydrates in form of cytoplasmic disaccharides and sucrose play an important role in hydration of pollen grains. Pollen needs to be partly dehydrated, in order to absorb or lose water without detrimental dehydration. Hydrated or non-hydrated pollen grains are both more vulnerable to water loss resulting in inviable pollen. Partly dehydrated pollen has higher concentrations of cytoplasmic disaccharides and sucrose and is therefore more likely to survive pollen transport and germination. The fact that the earliest part of the flowering season was a drought period (Wolff et al. 2018), could be one reason why several ramets produced more carbohydrates (Pacini 1996). However, in their study of 990 species, Baker and Baker (1979) found that starch containing pollen is associated with larger pollen grains. Since pollen grains of *C. angustifolium* are relatively large (58-96 μm; own results, but also see Halbritter (2017), the higher peaks at spectral bands associated with amylose (Figure 15), could be due to the sheer size of pollen grains.

The observed differences in chemical composition between the study population and the Kittilbu comparison, is consistent with other studies (Bağcıoğlu et al. 2017; Misfjord 2014; Zimmermann et al. 2017). The Kittilbu population differs from the study population for PC2, PC3 and PC4, particularly PC2. The relative cellulose content is higher, and the amylose composition is different. Overall, carbohydrates play a central role in the alpine Kittilbu population. Additionally, score plots of PC1, PC2 and PC4 all show separation in the Kittilbu population, where pollen samples EAiK01, EAiK02, EAiK03 form clusters apart from the other pollen samples. This is expected because these pollen samples belong to a ramet collected approximately 10-15 m apart from the other ramets of the Kittilbu population and is therefore likely genetically different. These three pollen samples from Kittilbu is also more similar to the study population, as indicated by the three lowest PC2 scores of the Kittilbu population in Figure 13, although slightly high.

The classification of seasonal variation shows a seasonal change in pollen chemistry. However, and similar to the ramet classification, the chemical differences behind this differentiation are not known. The PLSR results indicate that there is a good correlation (predicted R^2=0.51) between the chemical spectra of pollen and pollination day, during the pollination period of 48 days. The PLSR response plot in Figure 20 clearly show a trend going in one direction as the season progress. The PLSR response plot do provide information regarding what parts of the

chemical spectra that are causing the observed trend, through regression coefficients. However, due to model complexity, the regression coefficient plots investigated were difficult to interpret. The fact that pollen chemistry changed with regards to the pollination sequence of ramets, with $R^2 = 0.43$ (Appendix 9), might suggest that the sequence itself are causing changes in chemical composition of *C. angustifolium* pollen. Environmental conditions do change over the course of a flowering season, but this was not measured in this study. Additional studies could test sequentially flowering plants, under controlled environmental conditions, to validate if pollen chemistry change with pollination sequence, irrespective of environmental conditions.

Whether the observed seasonal differences in chemical spectra is related to lipids, proteins, carbohydrates, sporopollenin or combinations of these, remains unknown. It is plausible, however, that pollen quality decrease throughout the flowering season since other pollen fitness components (pollen mass and size) decrease as the season progress. This is also probable when considering the relationship between the chemical spectra and pollen mass (Figure 21).

Several studies have found changed pollen chemistry under different environmental conditions (Lau & Stephenson 1994; Stanley & Linskens 1974; Van Herpen 1981; Van Herpen 1986; Zimmermann et al. 2017) and changed pollen chemistry between seasons (Bağcıoğlu et al. 2017; Zimmermann & Kohler 2014). As environmental conditions change within a flowering season, the indicated changes in chemical composition of *C. angustifolium* pollen found in this study is expected. Bağcıoğlu et al. (2017) found considerable intra-seasonal in the chemical composition of *Corylus avellana* pollen, collected over more than five weeks at four different sampling dates. However, the authors suggested that the variation in pollen chemistry was not due to seasonal variation, as the plants collected likely were coming from different populations. Furthermore, Zimmermann et al. (2017) found variability in the chemical composition of pollen from one individual of *Festuca ovina*, collected over 11 days. The variability detected was only slightly higher than technical replicates though, and substantially lower than the variability in the population with corresponding environmental conditions. However, this finding is only partially comparable as the *F. ovina* pollen analyzed was collected from a population with fixed environmental conditions over a considerably shorter pollination period. *F. ovina* also differs fundamentally from *C. angustifolium* in ecological characteristics, e.g. by being wind-pollinated.

5 Concluding remarks and conclusion

These remarks are based on re-examination of the hypothesis in the introduction.

1) The hypothesis that pollen size (mass and diameter) will be smaller late in the flowering season, is partially supported as both pollen mass and pollen size were clearly reduced as the season progressed. Of the two size parameters at the population level, pollen mass was reduced most in relative terms. Although the reduction in pollen size was significant, it is percentagewise small (4 %). Additionally, as these trends were not uniform and not valid for all ramets, additional studies of supplementary populations are needed to firmly show that seasonal variations in pollen size is a characteristic trait in the reproductive biology of *C. angustifolium*.
2) Changes in pollen chemistry throughout the flowering season is supported.
   Results obtained from classification analyses cannot provide information regarding the difference in pollen chemistry, however some broad indications are provided by classification and PLSR analyses. There is a strong correlation between chemical spectra and pollination day (R^2= 0.51). However, since we do not know what parts of the chemical spectra that are causing the correlation, this result opens for further and more detailed chemical studies to find out more about the causality behind the correlation.
3) Different chemical composition of pollen of the study population and the comparison population, is supported. They differ mainly in terms of carbohydrate composition, and relative carbohydrate content to proteins and lipids.

To the best of my knowledge, this is the first study that describes both pollen size and pollen chemistry of a sequentially flowering plant (*C. angustifolium)* over the course of an entire flowering season. The results indicate that chemical composition of pollen and pollen size are plastic over the course of a flowering season. Together with adaptive ability, phenotypic plasticity of plants is important in an ever-changing environment, particularly with ongoing climate change. Future perspectives could focus on combining FTIR spectroscopy measurements of pollen with pollen quality measurements. This would enable us to validate if, and to what extent, the seasonal variation in pollen chemistry detected by FTIR spectroscopy is related to pollen quality and male reproductive success.

6 References

Afseth, N. K. & Kohler, A. (2012). Extended multiplicative signal correction in vibrational spectroscopy, a tutorial. *Chemometrics and Intelligent Laboratory Systems*, 117: 92-99.

Aguilar, R., Bernardello, G. & Galetto, L. (2002). Pollen–pistil relationships and pollen size-number trade-off in species of the tribe Lycieae (Solanaceae). *Journal of Plant Research*, 115 (5): 335-340.

Anderson, J. T., Inouye, D. W., McKinney, A. M., Colautti, R. I. & Mitchell-Olds, T. (2012). Phenotypic plasticity and adaptive evolution contribute to advancing flowering phenology in response to climate change. *Proceedings of the Royal Society B- Biological sciences*, 279 (1743): 3843-3852.

Bağcıoğlu, M., Zimmermann, B. & Kohler, A. (2015). A Multiscale Vibrational Spectroscopic Approach for Identification and Biochemical Characterization of Pollen. *PLoS ONE*, 10 (9): 1-19.

Bağcıoğlu, M. (2016). *Multiscale vibrational spectroscopy of pollen*. PhD. Ås,Norway: Norwegian University of Life Sciences.

Bağcıoğlu, M., Kohler, A., Seifert, S., Kneipp, J. & Zimmermann, B. (2017). Monitoring of plant–environment interactions by high-throughput FTIR spectroscopy of pollen. *Methods in Ecology and Evolution*, 8: 870-880.

Baker, H. G. & Baker, I. (1979). Starch in angiosperm pollen grains and its evolutionary significance. *American Journal of Botany*, 66 (5): 591-600.

Barth, A. (2007). Infrared spectroscopy of proteins. *Biochimica Et Biophysica Acta-Bioenergetics*, 1767 (9): 1073-1101.

Berdahl, M. H. (2014). *Pollen analysis by FTIR spectroscopy: a feasibility study for an automated method*. Master's thesis. Ås: Norwegian University of Life Sciences.

Bernasconi, G., Ashman, T. L., Birkhead, T. R., Bishop, J. D. D., Grossniklaus, U., Kubli, E., Marshall, D. L., Schmid, B., Skogsmyr, I., Snook, R. R., et al. (2004). Evolutionary ecology of the prezygotic stage. *Science*, 303 (5660): 971-975.

Blokker, P., Boelen, P., Broekman, R. A. & Rozema, J. (2006). The occurrence of p-coumaric acid and ferulic acid in fossil plant materials and their use as UV-proxy. *Plant Ecology*, 182 (1): 197-207.

Bokszczanin, K. L., Efragkostefanakis, S., Ebostan, H., Ebovy, A., Echaturvedi, P., Chiusano, M. L., Efiron, N., Eiannacone, R., Ejegadeesan, S., Eklaczynskid, K., et al. (2013). Perspectives on deciphering mechanisms underlying plant heat stress response and thermotolerance. *Frontiers in Plant Science*, 4 (315): 1-20.

Charlesworth, D., Vekemans, X., Castric, V. & Glémin, S. (2005). Plant self-incompatibility systems: a molecular evolutionary perspective. *New Phytologist*, 168 (1): 61-69.

Correa, E., Sletta, H., Ellis, D., Hoel, S., Ertesvåg, H., Ellingsen, T., Valla, S. & Goodacre, R. (2012). Rapid reagentless quantification of alginate biosynthesis in Pseudomonas fluorescens bacteria mutants using FT-IR spectroscopy coupled to multivariate partial least squares regression. *Analytical and Bioanalytical Chemistry*, 403 (9): 2591-2599.

Cruden, R. W. & Jensen, K. G. (1979). Viscin threads, pollination efficiency and low pollen-ovule ratios. *American Journal of Botany*, 66 (08): 875-879.

Cruden, R. W. & Lyon, D. L. (1985). Correlations Among Stigma Depth, Style Length, and Pollen Grain Size: Do They Reflect Function or Phylogeny? *Botanical Gazette*, 146 (1): 143-149.

Delph, L. F., Johannsson, M. H. & Stephenson, A. G. (1997). How environmental factors affect pollen performance: Ecological and evolutionary perspectives. *Ecology*, 78 (6): 1632-1639.

Elgersma, A., Stephenson, A. & Nijs, A. (1989). Effects of genotype and temperature on pollen tube growth in perennial ryegrass (*Lolium perenne* L.). *Sexual Plant Reproduction*, 2 (4): 225-230.

Forcella, F. & Wood, H. (1986). Sequential flowering of thistles (Cynareae, Asteraceae) in southern Australia [weeds]. *Australian Journal of Botany*, 34 (4): 455-461.

Geladi, P., MacDougall, D. & Martens, H. (1985). Linearization and scatter-correction for near-infrared reflectance spectra of meat. *Applied spectroscopy*, 39 (3): 491-500.

Gottardini, E., Rossi, S., Cristofolini, F. & Benedetti, L. (2007). Use of Fourier transform infrared (FT-IR) spectroscopy as a tool for pollen identification. *International Journal of Aerobiology - including the online journal `Physical Aerobiology'*, 23 (3): 211-219.

Halbritter, H. (2017). *Epilobium angustifolium*. PalDat - A palynological database. Available at: https://www.paldat.org/pub/Epilobium_angustifolium/302941;jsessionid=E6197EA8663B38CB459754C0BDE6E96F (accessed: 20.04.2019).

Halbritter, H., Ulrich, S., Grímsson, F., Weber, M., Zetter, R., Hesse, M., Buchner, R., Svojtka, M. & Frosch-Radivo, A. (2018). *Illustrated Pollen Terminology*. 2nd ed. Cham: Springer International Publishing.

Hanley, M. E., Franco, M., Pichon, S., Darvill, B. & Goulson, D. (2008). Breeding system, pollinator choice and variation in pollen quality in British herbaceous plants. *Functional Ecology*, 22 (4): 592-598.

Heslop-Harrison, J. (1979). An interpretation of the hydrodynamics of pollen. *American Journal of Botany*, 66 (6): 737-743.

Hesse, M. (1981). Pollenkitt and viscin threads: their role in cementing pollen grains. *Grana*, 20 (3): 145-152.

Hesse, M. (1983). Dissimilar pollen tetrad development in Ericaceae and Onagraceae causes family-specific viscin thread configuration. *Plant Systematics and Evolution*, 143 (1): 163-165.

Hesse, M., Halbritter, H., Weber, M., Buchner, R., Frosch-Radivo, A., Ulrich, S. & Zetter, R. (2009). *Pollen terminology: an illustrated handbook*. 1st ed. Wien: Springer Science & Business Media.

Jiang, Y., Lahlali, R., Karunakaran, C., Kumar, S., Davis, A. R. & Bueckert, R. A. (2015). Seed set, pollen morphology and pollen surface composition response to heat stress in field pea. *Plant, Cell & Environment*, 38 (11): 2387-2397.

Johannsson, M. H. & Stephenson, A. G. (1998). Effects of temperature during microsporogenesis on pollen performance in *Cucurbita pepo* L. (Cucurbitaceae). *International Journal of Plant Sciences*, 159 (4): 616-626.

Jürgens, A., Witt, T. & Gottsberger, G. (2012). Pollen grain size variation in Caryophylloideae: a mixed strategy for pollen deposition along styles with long stigmatic areas? *Plant Systematics and Evolution*, 298 (1): 9-24.

Kelly, J. K., Rasch, A. & Kalisz, S. (2002). A method to estimate pollen viability from pollen size variation. *American Journal of Botany*, 89 (6): 1021-1023.

Knight, T. M., Steets, J. A., Vamosi, J. C., Mazer, S. J., Burd, M., Campbell, D. R., Dudash, M. R., Johnston, M. O., Mitchell, R. J. & Ashman, T.-L. (2005). Pollen limitation of plant reproduction: pattern and process. *Annual review of ecology, evolution, and systematics*, 36: 467-497.

Kohler, A., Hanafi, M., Bertrand, D., Quannari, E. M., Janbu, A. O., Møretrø, T., Naterstad, K. & Martens, H. (2008). Interpreting Several Types Of Measurements In Bioscience. In Lasch, P. & Kneipp, J. (eds) *Biomedical Vibrational Spectroscopy*, pp. 333-356. Hoboken, NJ, USA: John Wiley & Sons.

Kooi, C. J., Pen, I., Staal, M., Stavenga, D. G. & Elzenga, J. T. M. (2016). Competition for pollinators and intra-communal spectral dissimilarity of flowers. *Plant Biology*, 18 (1): 56-62.

Lahlali, R., Jiang, Y., Kumar, S., Karunakaran, C., Liu, X., Borondics, F., Hallin, E. & Bueckert, R. (2014). ATR-FTIR spectroscopy reveals involvement of lipids and proteins of intact pea pollen grains to heat stress tolerance. *Frontiers in plant science*, 5 (747): 1-10.

Lau, T.-C. & Stephenson, A. (1994). Effects of soil phosphorus on pollen production, pollen size, pollen phosphorus content, and the ability to sire seeds in *Cucurbita pepo* (Cucurbitaceae). *Sexual Plant Reproduction*, 7 (4): 215-220.

Lau, T. C. & Stephenson, A. G. (1993). Effects of soil nitrogen on pollen production, pollen grain size, and pollen performance in *Cucurbita pepo* (Cucurbitaceae). *American Journal of Botany*, 80 (7): 763-768.

Marshall, D. L. & Ellstrand, N. C. (1986). Sexual Selection in *Raphanus sativus*: Experimental Data on Nonrandom Fertilization, Maternal Choice, and Consequences of Multiple Paternity. *The American Naturalist*, 127 (4): 446-461.

Martens, H., Jensen, S. A. & Geladi, P. (1983). *Multivariate linearity transformations for near infrared reflectance spectroscopy*. Nordic Symposium on Applied Statistics, Stavanger, Norway: Stokkland forlag. 205-234 pp.

Martens, H. & Næs, T. (1992). *Multivariate calibration.* Chichester: Wiley.

Martens, H. & Martens, M. (2001). *Multivariate analysis of quality: an introduction*. Chichester: Wiley.

McCallum, B. & Chang, S. M. (2016). Pollen competition in style: Effects of pollen size on siring success in the hermaphroditic common morning glory, *Ipomoea purpurea*. *American Journal of Botany*, 103 (3): 460-470.

Mione, T. & Anderson, G. J. (1992). Pollen-ovule ratios and breeding system evolution in *Solanum section Basarthrum* (Solanaceae). *American journal of botany*, 79 (3): 279-287.

Misfjord, K. (2014). *Pollen chemical composition determined by infrared spectroscopy: species identification and environmental effects*. Master's Thesis. Ås: Norwegian University of Life Sciences.

Mosquin, T. (1966). A new taxonomy for *Epilobium angustifolium* L. (Onagraceae). *Brittonia*, 18 (2): 167-188.

Myerscough, P. J. & Whitehead, F. H. (1966). Comparative biology of *Tussilago farfara* L., *Chamaenerion angustifolium* L. Scop., *Epilobium montanum* L. and *Epilobium adenocaulon* Hausskn. *New Phytologist*, 65 (2): 192-210.

Myerscough, P. J. (1980). Biological flora of the British Isles. 148. List Br. Vasc. Pl. (1958) No. 255/1: *Epilobium angustifolium* L. (*Chamaenerion angustifolium* (L.) Scop.). *Journal of Ecology*, 68 (3): 1047-1074.

NGU. (2019). *Norges Geologiske Undersøkelse: Berggrunn N250, løsmasser*. Available at: http://geo.ngu.no/kart/berggrunn_mobil/ (accessed: 03.05.2019).

Pacini, E. (1996). Types and meaning of pollen carbohydrate reserves. *Sexual Plant Reproduction*, 9 (6): 362-366.

Palanivelu, R. & Tsukamoto, T. (2012). Pathfinding in angiosperm reproduction: pollen tube guidance by pistils ensures successful double fertilization. *Wiley Interdisciplinary Reviews: Developmental Biology*, 1 (1): 96-113.

Pappas, C. S., Tarantilis, P. A., Harizanis, P. C. & Polissiou, M. G. (2003). New Method for Pollen Identification by FT-IR Spectroscopy. *Applied Spectroscopy*, 57 (1): 23-27.

Pham, V. T., Herrero, M. & Hormaza, J. I. (2015). Effect of temperature on pollen germination and pollen tube growth in longan (*Dimocarpus longan* Lour.). *Scientia Horticulturae*, 197: 470-475.

Piffanelli, P., Ross, J. H. & Murphy, D. (1998). Biogenesis and function of the lipidic structures of pollen grains. *Sexual plant reproduction*, 11 (2): 65-80.

Quesada, M., Bollman, K. & Stephenson, A. G. (1995). Leaf Damage Decreases Pollen Production and Hinders Pollen Performance in *Cucurbita texana*. *Ecology*, 76 (2): 437-443.

Ramrez, N. (2006). Temporal variation of pollination classes in a tropical Venezuelan plain: the importance of habitats and life forms. *Botany*, 84 (3): 443-452.

Randall, T. (1995). An introduction to partial least squares regression. In *Proceedings of the twentieth annual SAS users group international conference*, pp. 1250-1257. Orlando: SAS Institute Inc Cary.

Rieu, I., Twell, D. & Firon, N. (2017). Pollen Development at High Temperature: From Acclimation to Collapse. *Plant physiology*, 173 (4): 1967.

Rodríguez-García, M. I., M'Rani-Alaoui, M. & Fernández, M. C. (2003). Behavior of storage lipids during development and germination of olive (*Olea europaea* L .) pollen. *Protoplasma*, 221 (3): 237-244.

Roulston, T. A. H., Cane, J. H. & Buchmann, S. L. (2000). What governs protein content of pollen: Pollinator preferences, pollen-pistil interactions, or phylogeny? *Ecological Monographs*, 70 (4): 617-643.

RStudioTeam. (2016). RStudio: Integrated Development for R. Boston, MA: RStudio.Inc.

Rutley, N. & Twell, D. (2015). A decade of pollen transcriptomics. *Plant Reproduction*, 28 (2): 73-89.

Sarkissian, T. S. & Harder, L. D. (2001). Direct and indirect responses to selection on pollen size in *Brassica rapa* L. *Journal of Evolutionary Biology*, 14 (3): 456-468.

Sarmiento, A., Pérez-Alonso, M., Olivares, M., Castro, K., Martínez-Arkarazo, I., Fernández, L. & Madariaga, J. (2011). Classification and identification of organic binding media in artworks by means of Fourier transform infrared spectroscopy and principal component analysis. *Analytical and Bioanalytical Chemistry*, 399 (10): 3601-3611.

Savitzky, A. & Golay, M. J. E. (1964). Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Analytical Chemistry*, 36 (8): 1627-1639.

Schaeffer, R. N., Manson, J. S. & Irwin, R. E. (2013). Effects of abiotic factors and species interactions on estimates of male plant function: a meta-analysis. *Ecology Letters*, 16: 399-408.

Speranza, A., Calzoni, G. & Pacini, E. (1997). Occurrence of mono-or disaccharides and polysaccharide reserves in mature pollen grains. *Sexual Plant Reproduction*, 10 (2): 110-115.

Stanley, R. G. & Linskens, H. F. (1974). *Pollen: biology biochemistry management*. Berlin: Springer.

Stephenson, A. G., Travers, S. E., Mena-Ali, J. I. & Winsor, J. A. (2003). Pollen performance before and during the autotrophicheterotrophic transition of pollen tube growth. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 358 (1434): 1009-1018.

Van Herpen, M. (1981). Effect of season, age and temperature on the protein pattern of pollen and styles in *Petunia hybrida*. *Acta botanica neerlandica*, 30 (4): 277-287.

Van Herpen, M. & Linskens, H. (1981). Effect of season, plant age and temperature during plant growth on compatible and incompatible pollen tube growth in *Petunia hybrida*. *Acta Botanica Neerlandica*, 30 (3): 209-218.

Van Herpen, M. (1986). Biochemical alterations in the sexual partners resulting from environmental conditions before pollination regulate processes after pollination. In Mulcahy, D. L., Mulcahy, G. B. & Ottaviano, E. (eds) *Biotechnology and Ecology of Pollen: Proceedings of the international conference on the biotechnology and ecology of pollen, 9-11 July, 1985, University of Massachusetts, Amherst, MA, USA*, pp. 131-133. New York: Springer.

Vesprini, J. L., Nepi, M., Cresti, L., Guarnieri, M. & Pacini, E. (2002). Changes in cytoplasmic carbohydrate content during Helleborus pollen presentation. *Grana*, 41 (1): 16-20.

Williams, J. H. & Mazer, S. J. (2016). Pollen—Tiny and ephemeral but not forgotten: New ideas on their ecology and evolution. *American Journal of Botany*, 103 (3): 365-374.

Wolff, M., Thue-Hansen, V. & Grimenes, A. A. (2018). *Meteorologiske data for Ås 2017*. Ås: Norges miljø- og biovitenskapelige universitet, Fakultet for realfag og teknologi. Available at: https://www.nmbu.no/fakultet/realtek/laboratorier/bioklim/meteorologiske-data (accessed: 04.02.2019).

Wäckers, F. L., Romeis, J. & van Rijn, P. C. J. (2007). Nectar and Pollen Feeding by Insect Herbivores and Implications for Multitrophic Interactions. *Annual Review of Entomology*, 52 (1): 301-323.

Zimmermann, B. (2010). Characterization of Pollen by Vibrational Spectroscopy. *Applied Spectroscopy*, 64 (12): 1364-1373.

Zimmermann, B. & Kohler, A. (2013). Optimizing Savitzky–Golay Parameters for Improving Spectral Resolution and Quantification in Infrared Spectroscopy. *Applied Spectroscopy*, 67 (8): 892-902.

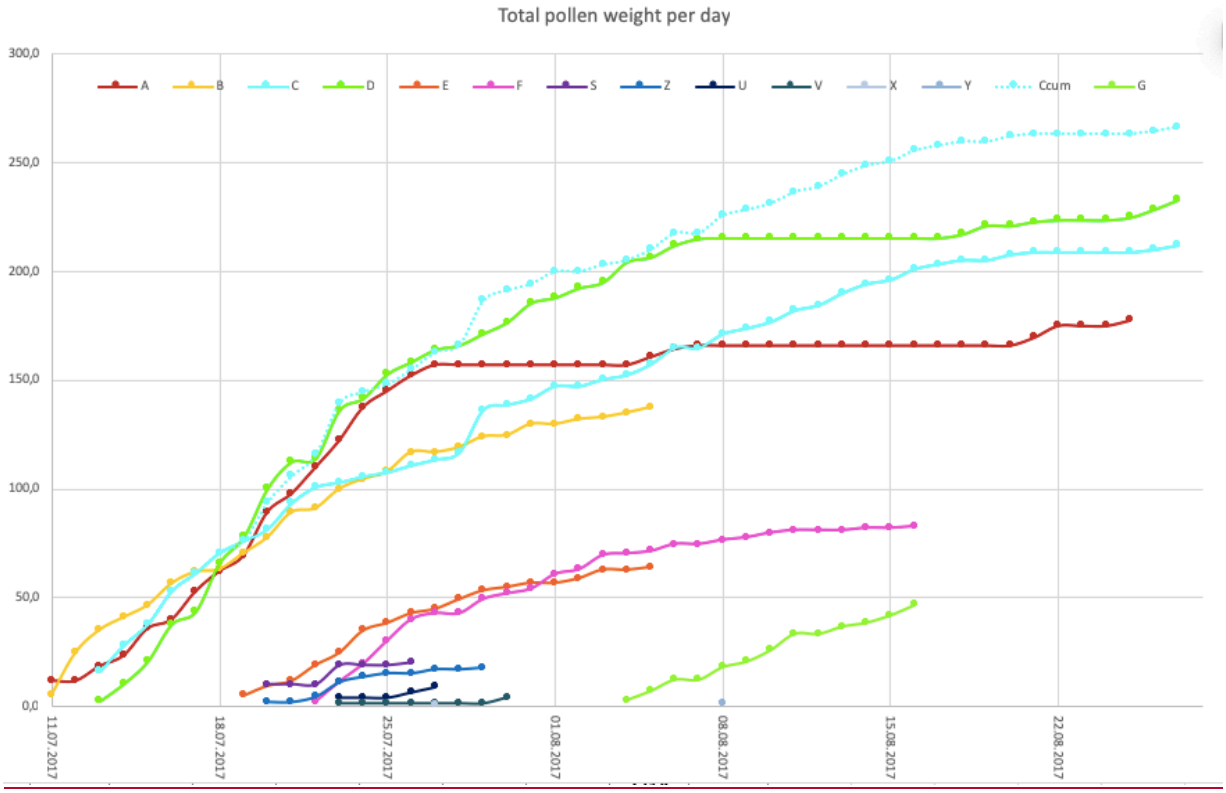
Zimmermann, B. & Kohler, A. (2014). Infrared spectroscopy of pollen identifies plant species and genus as well as environmental conditions. *PLoS ONE*, 9 (4): 1-12.

Zimmermann, B., Tkalčec, Z., Mešić, A. & Kohler, A. (2015). Characterizing aeroallergens by infrared spectroscopy of fungal spores and pollen. *PLoS ONE*, 10 (4): 1-22.

Zimmermann, B., Bağcıoğlu, M., Tafinstseva, V., Kohler, A., Ohlson, M. & Fjellheim, S. (2017). A high-throughput FTIR spectroscopy approach to assess adaptive variation in the chemical composition of pollen. *Ecology and Evolution*, 7 (24): 10839-10849.

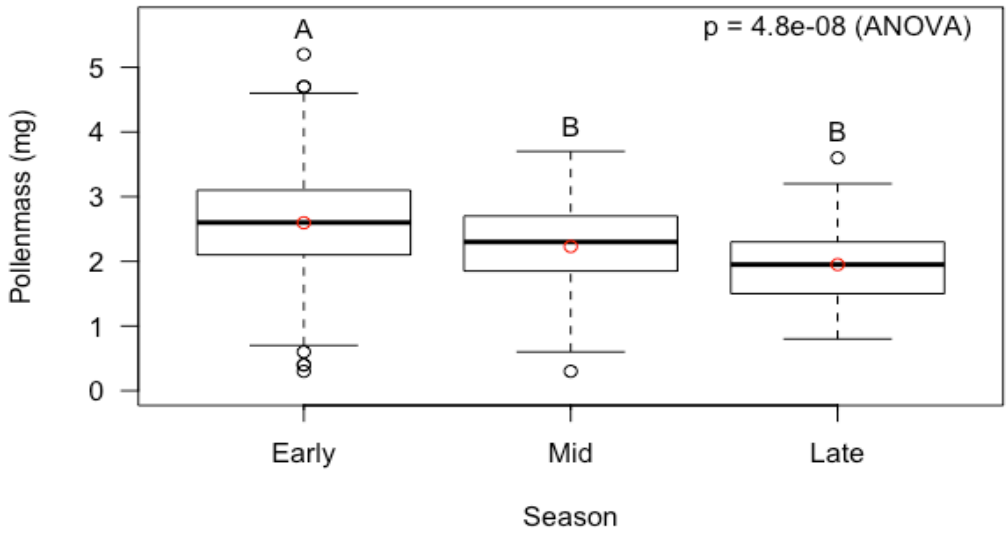
Zimmermann, B. (2018). Chemical characterization and identification of Pinaceae pollen by infrared microspectroscopy. *Planta*, 247 (1): 171-180.

Appendix 1



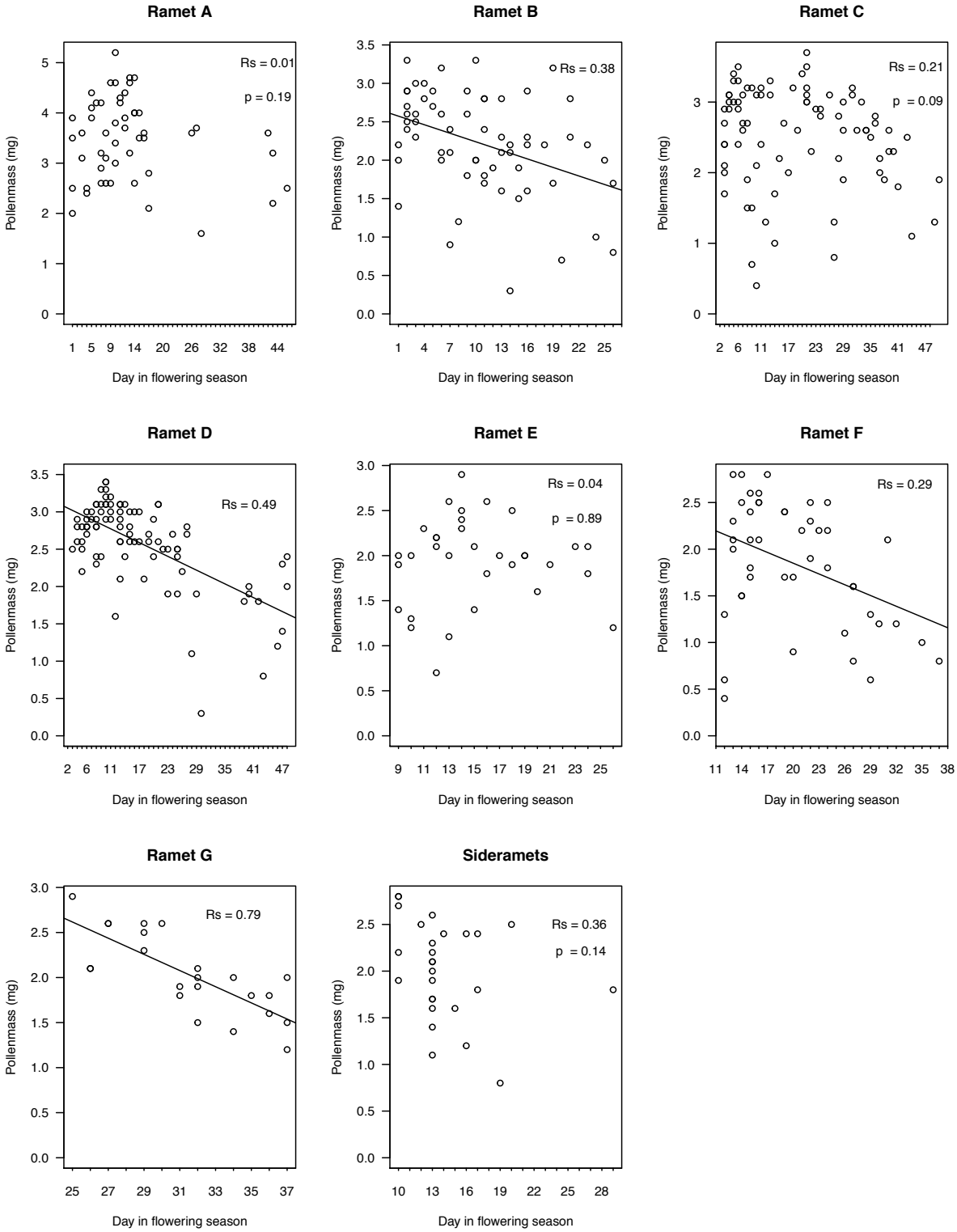
Pollen production over the flowering season, with dates on the x-axis and accumulated pollen mass on the y-axis. Ramets S, Z, U, V, X and Y are side-ramets.

Appendix 2



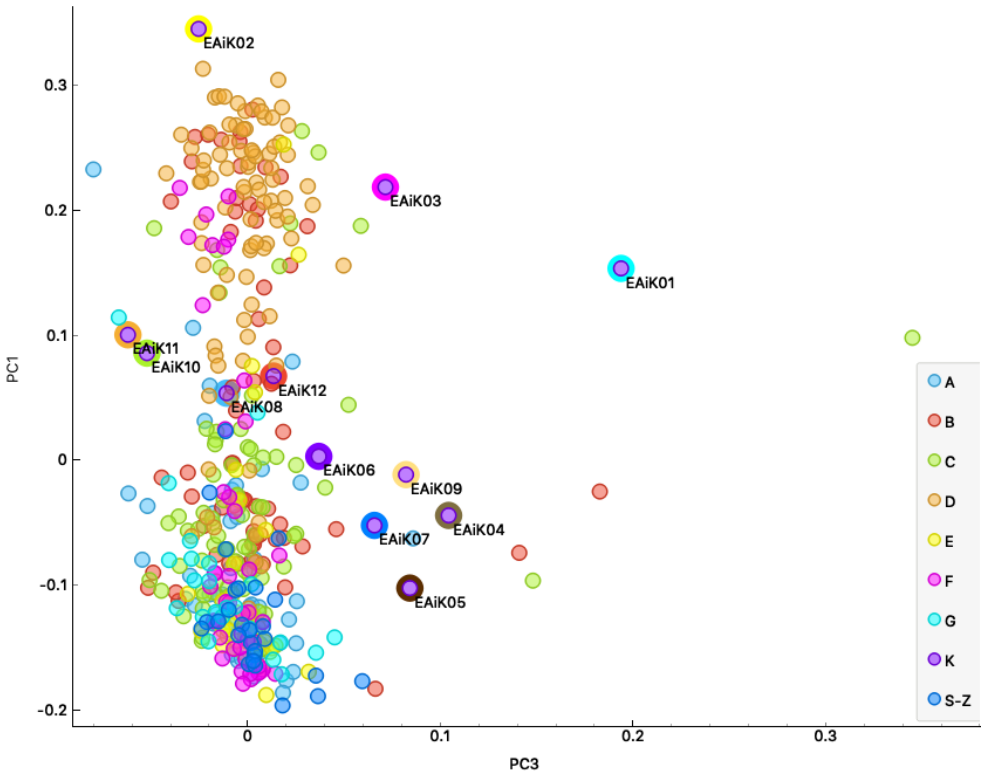
Boxplot showing pollen mass throughout the flowering season for all pollen samples. Red circles show mean value. Boxplots topped with different letters indicate significant differences ($P < 0.05$, Tukey). Only pollen samples > 0.0 mg included.

Appendix 3



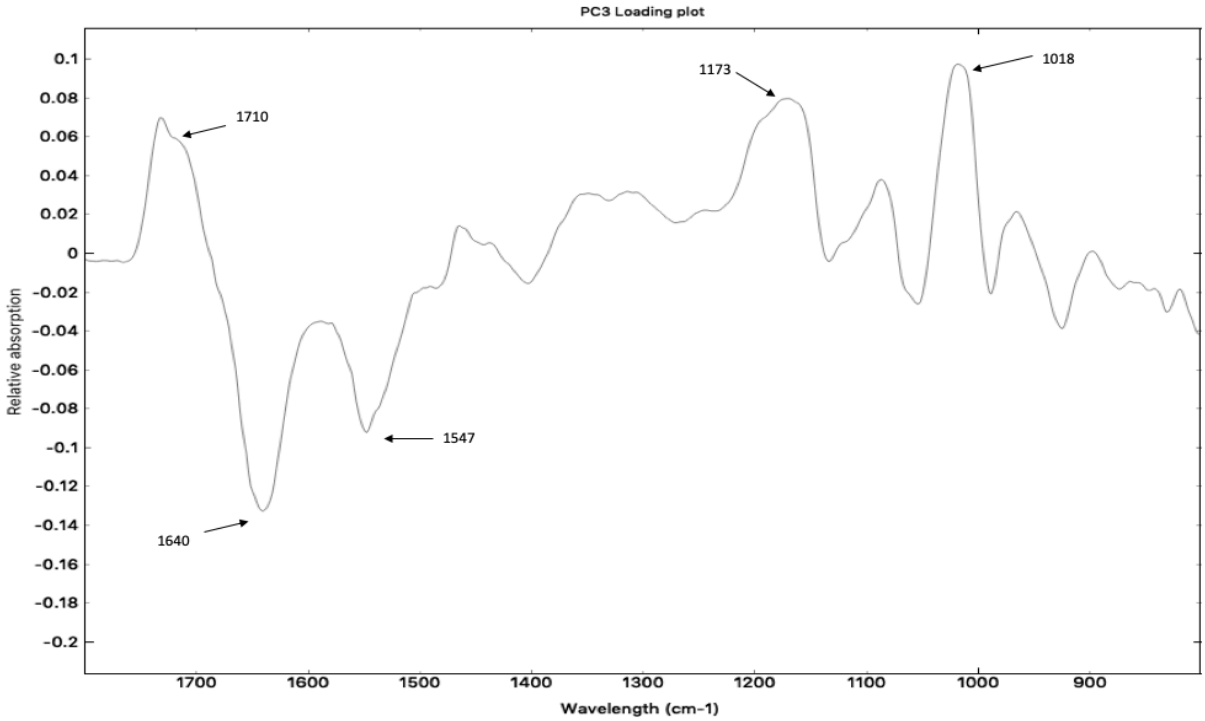
Scatterplots of pollination day on the x-axis, i.e. the day in the flowering season where a flower would open, and pollen mass on the y-axis. Pollination day are population specific, with day 1: 11.07.2017 and the last day (day 48): 27.08.2017. Lines indicate significant correlation, $P < 0.05$. Spearman rank correlation coefficients are reported as Rs. Only pollen samples > 0.0 mg included. Note different scales on the vertical axis for different different ramets.

Appendix 4



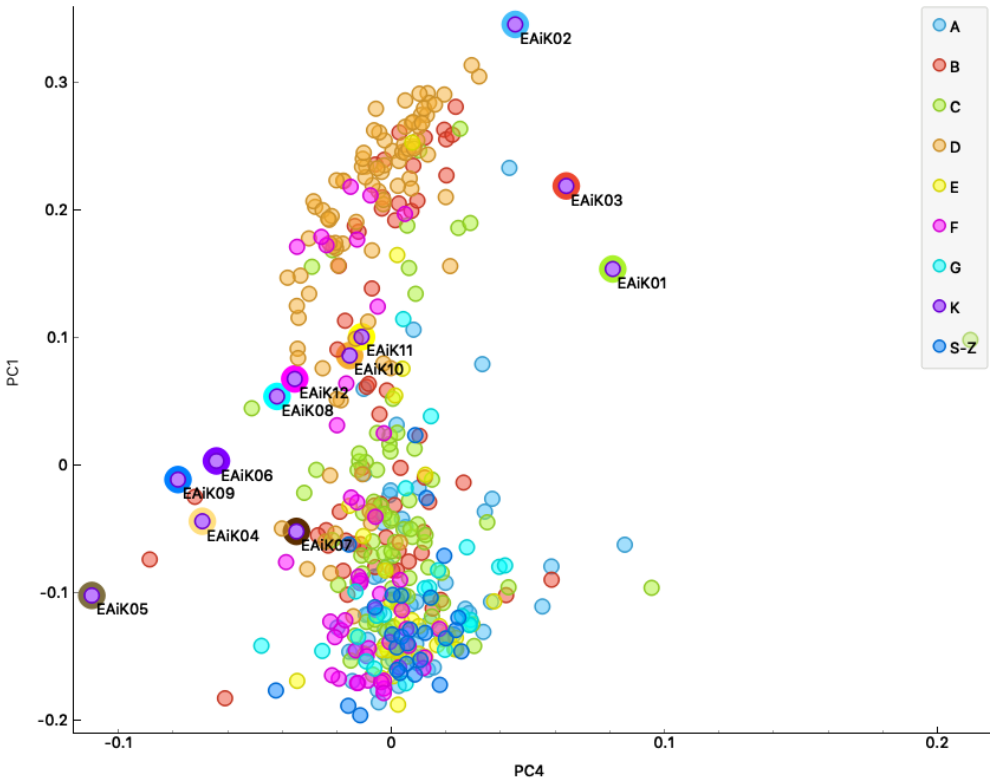
PC score plot, with PC3 on the x-axis, and PC1 on the y-axis. All samples from the Kittilbu population marked.

Appendix 5



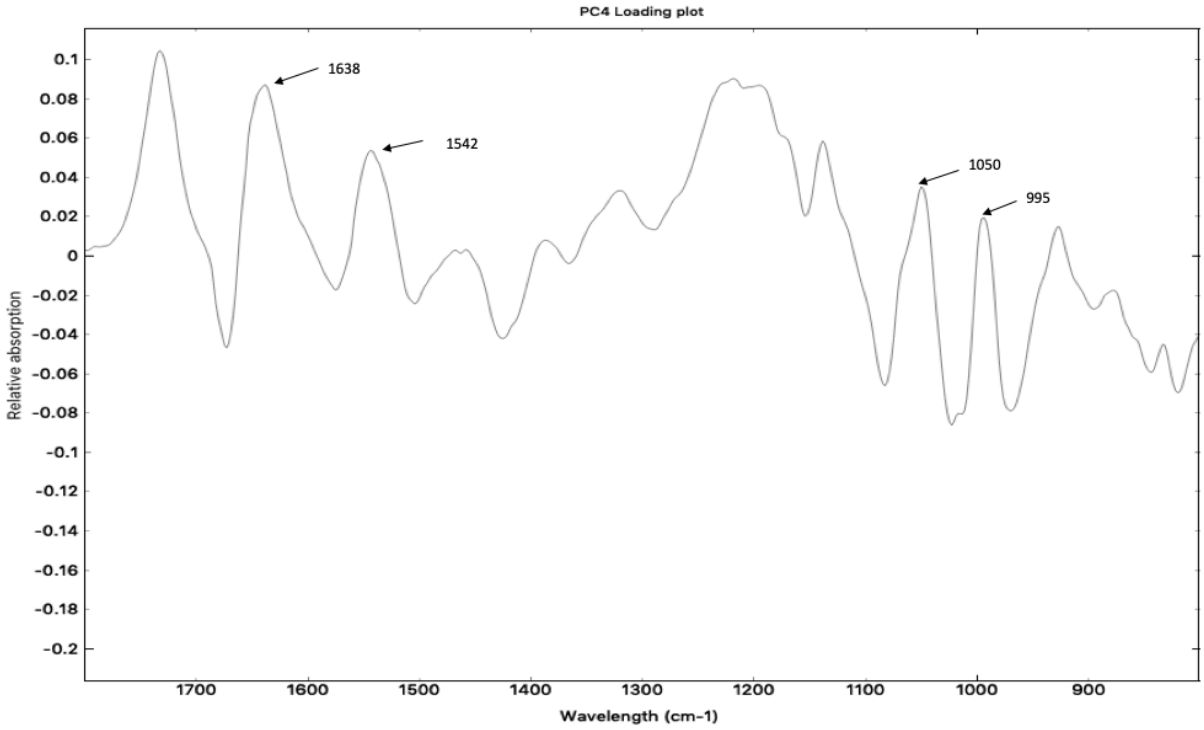
PC3 loading plot of FTIR spectra (explained variation: 3,8%). The x-axis shows wavelengths, and the y-axis show relative absorption.

Appendix 6



PC score plot, with PC4 on the x-axis, and PC1 on the y-axis. All samples from the Kittilbu population marked.

Appendix 7



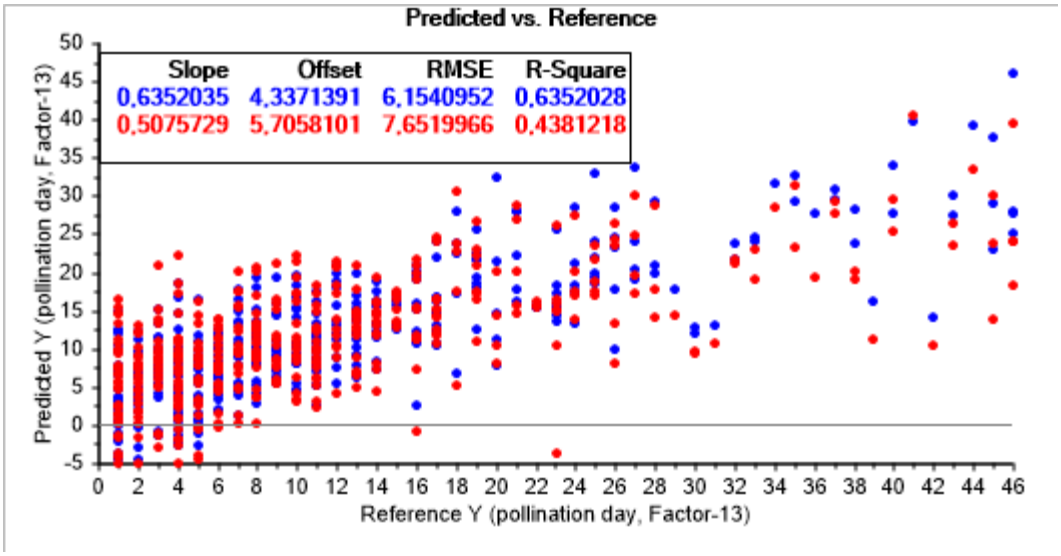
PC4 loading plot of FTIR spectra (explained variation: 2.1 %). The x-axis shows wavelengths, and the y-axis show relative absorption.

Appendix 8

The PLSR coefficient of determination (R^2) and root mean squared error (RMSE) for pollination sequence based on chemical spectra. m: model, p: prediction. A_{opt}: optimal number of components. S: side-ramet S, n=10 (not all side-ramets).

Ramet	Number of pollen samples analyzed by FT-IR	Pollination period (days)	Days with pollination	R^2 (m)	R^2 (p)	RMSE (m)	RMSE (p)	A_{opt}
All	415	48	46	0.63	0.43	6.15	7.65	13
A	51	46	22	0.70	0.50	5.78	7.52	5
B	62	26	24	0.64	0.48	4.18	5.09	4
C	84	46	39	0.77	0.59	5.90	8.00	7
D	90	46	35	0.94	0.68	2.84	6.74	12
E	33	18	16	0.87	0.49	1.64	3.46	6
F	47	25	22	0.94	0.44	1.69	5.40	8
G	23	13	11	0.99	0.29	0.23	3.23	11
S	10	7	3	0.85	0.29	0.72	0.79	3

Appendix 9



PLSR response plot of pollination sequence for all pollen samples. Actual pollination sequence is shown on the x-axis and predicted pollination sequence on the y-axis. Blue dots indicate model fit, whereas red dots indicate model performance after cross-validation.

Norges miljø- og biovitenskapelige universitet
Noregs miljø- og biovitskapelege universitet
Norwegian University of Life Sciences

Postboks 5003
NO-1432 Ås
Norway