Master's thesis

NTNU
Norwegian University of Science and Technology
Faculty of Natural Sciences
Department of Biology

Morgan Amundsen

Testing species boundaries in Scandinavian specimens of *Nomada flavoguttata* (Kirby, 1802) using an integrative approach

Master's thesis in Nordic Master's Biology – Biodiversity and Systematics
Supervisor: Frode Ødegaard
Co-supervisor: Torbjørn Ekrem
May 2023

NTNU
Norwegian University of Science and Technology

Morgan Amundsen

Testing species boundaries in Scandinavian specimens of *Nomada flavoguttata* (Kirby, 1802) using an integrative approach

Master's thesis in Nordic Master's Biology – Biodiversity and Systematics
Supervisor: Frode Ødegaard
Co-supervisor: Torbjørn Ekrem
May 2023

Norwegian University of Science and Technology
Faculty of Natural Sciences
Department of Biology

NTNU
Norwegian University of Science and Technology

Testing species boundaries in Scandinavian specimens of *Nomada flavoguttata* (Kirby, 1802) using an integrative approach

Sammendrag

Kleptoparasittiske bier utgjør omtrent 13% av alle verdens biearter. Deres spesialiserte økologi gjør at de er gode indikatorer på tilstanden i terrestriske økosystemer. *Nomada* er den største slekten av kleptoparasitiske bier. Tidligere strekkodede *Nomada flavoguttata* individer fra Skandinavia blir delt inn i tre godt understøttede klader. I dette prosjektet testet jeg artsavgrensningen mellom disse tre kladene. Ved å bruke en integrert tilnærming så jeg på molekylær, morfologisk, og økologisk data. Fem gensegmenter (COI, 16S, 28S, ITS-2, og EF1-α) ble sekvensert og analysert ved bruk av flere metoder for artsavgrensning (GMYC, PTP, BPP, og ASAP). I tillegg testet jeg for *Wolbachia*, en endosymbiontisk mikroorganisme som i noen tilfeller manipulerer de genetiske linjene til leddyr. Resultatene fra de forskjellige metodene for molekylær artsavgrensning var motstridene, enten en eller tre arter ble foreslått. Ingen av de undersøkte morfologiske karakterene viste seg å være egnet til å sikkert skille kladene. Hvorvidt den genetiske variasjonen i *Nomada flavoguttata* er tilknyttet vertsart forblir uløst. Det ble oppdaget en veldig lav infeksjonsrate (2.6%) av *Wolbachia* hos *N.flavoguttata*. Bare ett individ var infisert og det ble plassert sammen med de uinfiserte individene, noe som tyder på at *Wolbachia* ikke har påvirket det genealogiske slektskapet indikert av mtDNA hos *Nomada flavoguttata*.

Abstract

Cleptoparasitic bees comprise approximately 13% of all the world's bee species. Their specialized ecology makes them good indicators for assessing the status of terrestrial ecosystems. *Nomada* is the largest genus of cleptoparasitic bees. Previously barcoded *Nomada flavoguttata* specimens from Scandinavia are placed in three well-supported clades. In this project, I tested the species boundaries of these three clades using an integrative approach, looking at molecular, morphological, and ecological data. Five gene segments (COI, 16S, 28S, ITS-2, and EF1-α) were sequenced and analyzed using several methods for species delimitation (GMYC, PTP, BPP, and ASAP). In addition, I screened for *Wolbachia*, an endosymbiotic microorganism known to occasionally manipulate arthropod genealogical lineages. The results from the different methods for molecular species delimiting were conflicting, suggesting either one or three species. None of the morphological characters investigated proved to be good at confidently separating the clades. Whether or not the genetic diversity of *N. flavoguttata* is linked to host association remains unsolved. A very low infection frequency (2.6%) of *Wolbachia* was detected in *N. flavoguttata*. Only one specimen was infected, and it was placed alongside the uninfected specimens, suggesting that *Wolbachia* has not affected the genealogical relationships indicated by mtDNA in *N. flavoguttata*.

Contents

List of Figures

List of Tables

1.0 Introduction

Bees are a species-rich group with approximately 20,000 species described worldwide (Sann et al. 2018). They exist on every continent except Antarctica (Orr et al. 2021). Bees are divided into seven families, where Apidae is the largest, with more than 5600 species described (Cardinal et al. 2010). Bees are key units in natural terrestrial ecosystems. The vital role they play in the pollination of plants is crucial for the health of the ecosystems (Potts et al. 2010). In addition to the pollination services they provide in natural ecosystems, they also provide essential pollination services worth billions of US dollars in agricultural crops (Garibaldi et al. 2013; Klein et al. 2007; Losey & Vaughan 2006). Loss in bee diversity may trigger a cascade of declines in wild plant communities (Biesmeijer et al. 2006; Pauw 2007).

The bees show great variance in morphology and life history strategies. Based on life history strategies, bees can be divided into three main groups, solitary, social, and brood parasites (Danforth et al. 2019). One of the more specialized life history strategies among bees is brood parasitism (also known as cleptoparasitism), a form of parasitism where the parasite lays its eggs in the nest cells of the host species (Litman et al. 2013; Thorogood et al. 2019). The parasite larva then kills the host egg/larva and feeds on the food provided for the host larva. This behavior resembles the behavior of cuckoos, and bees with this behavior are commonly referred to as cuckoo bees (Litman 2019). About 13 % of the world's bee species are brood parasites (Danforth et al. 2019). Most cuckoo bees have few host species, usually between two and five. In Europe, approximately 25% of the cuckoo bees parasitize only one host species (Bogusch et al. 2006; Habermannová et al. 2013).

Nomada is the largest genus of brood-parasitic bees, with approximately 850 species described worldwide (Falk 2015). *Nomada* is divided into 16 species groups based on morphology (Alexander 1994). A recent phylogenetic study of *Nomada* recognizes 14 of the 16 species groups as monophyletic (Odanaka et al. 2022). Most *Nomada* species are cleptoparasites of bee species from the genus *Andrena*. Some are parasites on other genera representing at least five different families, giving *Nomada*, as a genus, a very wide range of hosts (Lim et al. 2022; Michener 2007). They usually parasitize one or several closely related species (Tengo & Bergstrom 1977).

Many studies show that wild bee populations worldwide are declining at an alarming rate (e.g. Biesmeijer et al. 2006; Cameron et al. 2011). As a result, the need for wild bee conservation has received increased attention (Drossart & Gérard 2020). Cleptoparasitic bees depend on healthy populations of their hosts. Therefore, among bees, they are the first to respond to ecosystem disturbances. This makes them good indicators for assessing the status of ecosystems (Sheffield et al. 2013).

To be able to conserve the bees, we need a better taxonomic and distributional understanding of them (Gonzalez et al. 2013). Despite more than 250 years of taxonomical work on bees, new species are frequently discovered even in well-studied faunas, such as in North/Central Europe (e.g. Litman et al. 2021; Praz et al. 2022). Many genera of bees contain cryptic diversity, which has gone undetected using traditional morphological methods. Recently, this has led to the description of many new species (Pauly et al. 2015; Praz et al. 2022; Wood 2021).

The taxonomy of *Nomada* is challenging as the genus includes many morphologically similar species. In poorly studied areas, they are often not identified to species (Mitai & Tadauchi 2007). One of the issues, especially in poorly studied areas, is that many *Nomada* species are described based on one gender only (Mitai & Tadauchi 2007). This is challenging since many of the species are sexually dimorphic. In addition, the coloration in many species can be very variable, making identification more difficult (Droege et al. 2010). The addition of mtDNA has helped to deal with these problems. However, in some cases, mtDNA does not differentiate between morphologically and ecologically distinct species (Falk et al. 2022). Despite all the challenges, there have been many recent changes in the taxonomy of *Nomada*, including the description of new species (Arturo Roig 2009; Proshchalykin & Lelej 2010; Smit 2018), suggestion of new synonymies (Droege et al. 2010), or elevating a form to a valid species (Falk et al. 2022).

Nomada flavoguttata (Fig. 1) is the smallest of the about 20 *Nomada* species registered in Norway. It is a common yet easily overlooked species due to its size. In Norway apart from a few old records in northern Norway, it is found north to Trøndelag. The absence of newer records north of Trøndelag is most likely due to being overlooked rather than not being present (Stenløkk 2011). In the rest of the world, *N. flavoguttata* is found in all of Europe, North Africa, and Asia, from Turkey to Japan (Smit 2018). It is bivoltine in most of its range,

with spring and summer generations. In the northern part and upland areas, it is usually univoltine with one generation in May-June (Falk 2015). Five host species are recorded for *N. flavoguttata*, all in subgenus *Andrena* (*Micrandrena*) (Smit 2018). All the host species are registered in Norway: *Andrena falsifica, A. minutula, A. minutuloides, A. semilaevis,* and *A. subopaca*. Both *A. falsifica* and *A. minutuloides* are rare species, *A. minutuloides* with only one record in Norway and *A. falsifica* with six records in the last ten years (Artskart 2023). The three others are common; they all have 60+ records in the last ten years (Artskart 2023). *Nomada flavoguttata* is sexually dimorphic, with the females having red markings that the males lack, especially on the top and sides of the thorax. The size of the genders is similar, ranging from 5-8 mm (Smit 2018). In Norway, it is usually easily recognizable due to its small size. In other parts of Europe and Asia, it can be confused with several species, especially *N. castellana,* which is very similar. However, *N. castellana* has a yellow labrum, while most *N. flavoguttata* specimens have a dark labrum. In addition, the males of *N. castellana* have distinct tubercles on the underside of antennal segments 4-12, which barely, if at all, are visible in *N. flavoguttata* (Falk 2015).

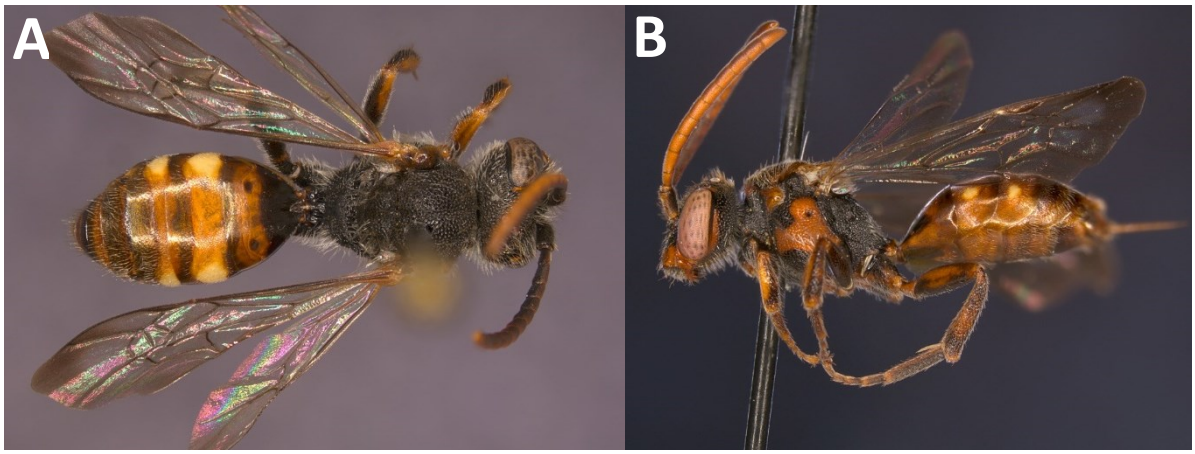


Figure 1. Pictures of Nomada flavoguttata. **A)** Male in dorsal view. **B)** Female in lateral view.

What defines a species has been debated for decades, and more than 30 species concepts have been proposed (Zachos 2016). However, there is widespread agreement that species are separately evolving metapopulation lineages (De Queiroz 2007; Zachos 2018). This primary concept cannot be used in species delimitation. Thus, secondary concepts are needed to practically recognize species (Zachos 2018).

Traditionally, taxonomists have identified species based on morphological characters alone. This can often be a difficult and time-consuming task, requiring experts in taxonomy

(Packer et al. 2009). A more modern approach to species identification is using DNA sequences in the identification process (Hebert et al. 2003). After the description of the "universal" primers to amplify the mitochondrial cytochrome c oxidase subunit I (COI) in 1994 (Folmer et al. 1994) and the introduction of DNA barcoding by Hebert et al. (2003), DNA barcoding has become a very popular supplementary tool for identifying species (Ajmal Ali et al. 2014; Fišer Pečnikar & Buzan 2014). In many cases, DNA barcoding allows us to identify species that are difficult to determine just by morphology (Janzen et al. 2005). DNA barcoding has led to the discovery of cryptic diversity, which morphological studies have not detected (Dincă et al. 2011). However, especially in alpha taxonomy, morphological studies are critical to ensure proper taxonomical descriptions (Beheregaray & Caccone 2007; Schlick-Steiner et al. 2007). One factor that may weaken results inferred from mtDNA sequences regarding species delimitation or species identification is endosymbiotic microorganisms (Hurst & Jiggins 2005). *Wolbachia* is a genus of endosymbiotic bacteria known to parasitize many arthropods (Duron et al. 2008). Recent estimates suggest that 40-66% of all arthropod species are infected by *Wolbachia* (Hilgenboecker et al. 2008; Weinert et al. 2015; Zug & Hammerstein 2012). In some insect groups the estimates are even higher, up to around 80% in Lepidoptera (Ahmed et al. 2015). *Wolbachia* infections may reduce mtDNA's genetic diversity compared to nuclear DNA's diversity from the same specimens (Ballard et al. 1996; Jiggins 2003). Many bee species, including several *Nomada* species, have tested positive for *Wolbachia* infections (Gerth et al. 2011; Gerth et al. 2013; Queffelec et al. 2022).

In recent years, there has been an increase in the number of methods for delimiting species (Burbrink & Ruane 2021; Carstens et al. 2013; Puillandre et al. 2021). Each method has a significant failure rate alone, but combined, the failure rate becomes minute (Dayrat 2005). Therefore, species delimitation should be based on multiple methods rather than one, i.e., integrative taxonomy (Schlick-Steiner et al. 2010; Zamani et al. 2022). In some cases, there is a mismatch between methods, e.g., morphological and DNA-based methods (Trewick 2008), which emphasizes the importance of using more than one method. Integrative taxonomy has become a popular approach, and newer studies regarding species delimitation often use an integrative approach (e.g. Moraes et al. 2021; H. Zhang et al. 2021; Aarvik et al. 2022).

Nomada flavoguttata (Kirby, 1802) is interesting to investigate for several reasons. Five host species have been recorded for the species (Smit 2018). The DNA-barcode samples from

Norway and Sweden place *N. flavoguttata* specimens in three well-supported clades (Fig. 2). The sequences are distinctly separated from other *Nomada* species. In many insect species, the intraspecific variation in the COI gene is relatively high. About a quarter out of more than 60,000 species have intraspecific variation higher than 3% (H. Zhang & Bu 2022). The intraspecific variation of the COI-barcoded *N. flavoguttata* specimens is a maximum of about 1,8%, and the mean distance between the clades is about 1,5%. In several other *Nomada* species, the interspecific COI variation is very low, and sometimes species cannot be distinguished using COI sequences (Falk et al. 2022). Based on this, it is worth investigating the species boundaries of *N. flavoguttata* using an integrative approach.

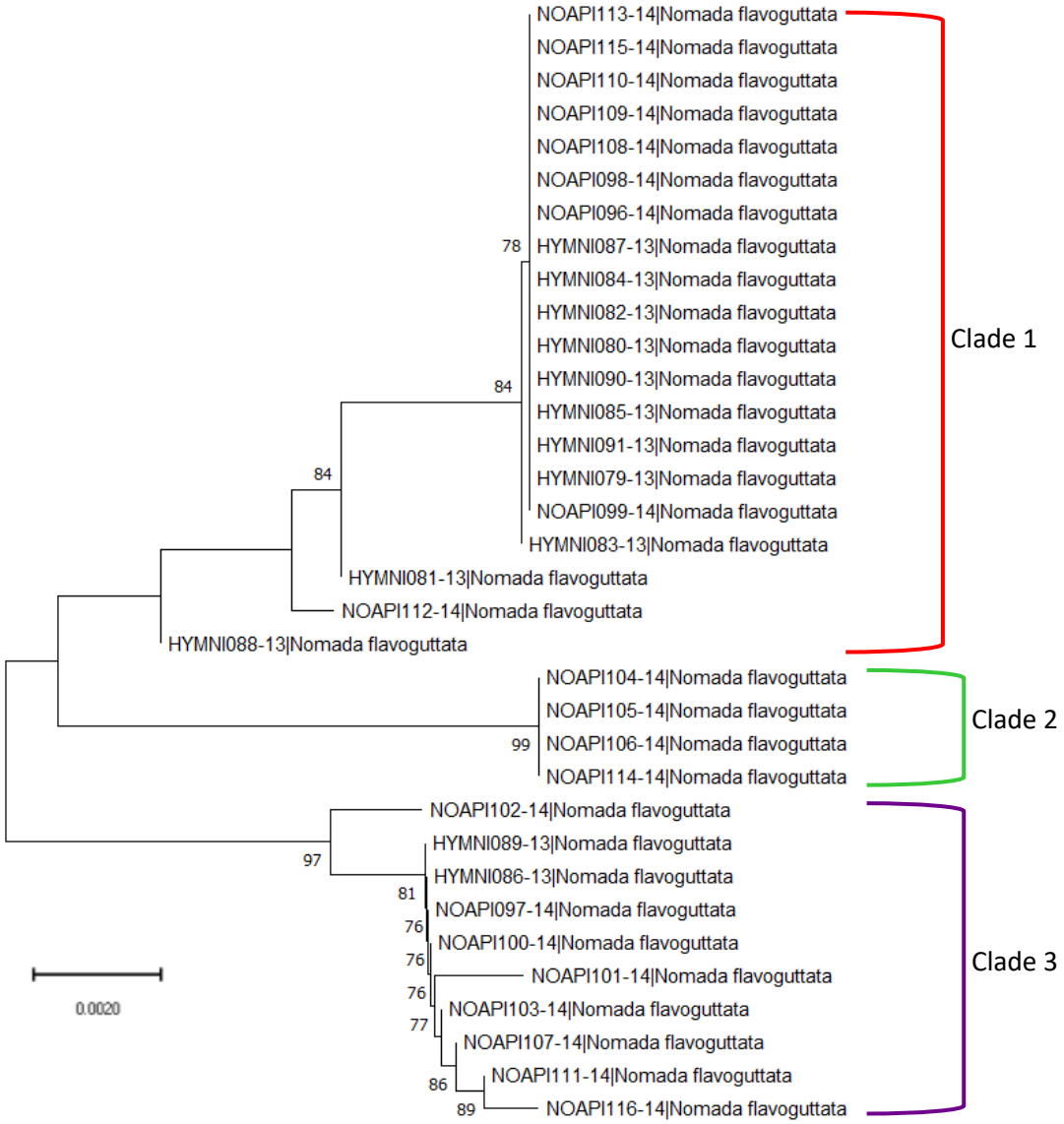


Figure 2. Neighbor joining tree constructed from COI of *N. flavoguttata* specimens barcoded prior to the present study. Created in MEGA 11 (Tamura et al. 2021), bootstrap values higher than 75 are shown.

Morphology has been a central part of taxonomy since the beginning of the discipline. The inclusion of morphometric analysis has grown in popularity in many organisms. It is suggested that morphometrics may be less prone to errors than the interpretation of qualitative characters (Csősz et al. 2021). I will look at morphology in this study, including some morphometric measurements.

Ecology on its own is usually not suitable for distinguishing species, but in many cases, it is a good supplementary tool (Schlick-Steiner et al. 2010). As cleptoparasitic bee species usually are very host-specific, host association may aid in the identification of species (Falk et al. 2022). I will investigate if there are differences between *N. flavoguttata* specimens associated with different *Andrena* species.

As a result of the growing popularity and success of DNA barcoding, many methods to delimit species based on DNA sequences have been made (Luo et al. 2018; Puillandre et al. 2021). Some of the widely used methods for molecular species delimitation Include generalized mixed Yule-coalescent (GMYC) (Pons et al. 2006), Poisson tree processes (PTP) (J. Zhang et al. 2013), Assemble species by automatic partitioning (ASAP) (Puillandre et al. 2021), and Bayesian Phylogenetics and Phylogeography (BPP) (Rannala & Yang 2003; Yang 2015).

The molecular methods can be divided into two main categories. The exploratory ones, which are methods that can propose new species delimitations. They are most often used on single-locus datasets. All the methods mentioned above, except BPP, fall within this category (Puillandre et al. 2021). Both GMYC and PTP use a phylogenetic tree to delimit species. They differ in that GMYC uses a tree with branch lengths proportional to time, while PTP uses a tree with branch lengths proportional to genetic differences. In that way, PTP is less demanding, as constructing ultrametric trees under the assumption of relaxed molecular clocks can require a lot of computing power for large datasets (Luo et al. 2018). A very computing-efficient method that only requires an alignment of sequences as input is ASAP. This method uses pairwise distances to cluster sequences, avoiding the computing power needed for phylogenetic reconstruction (Puillandre et al. 2021).

The other main category contains methods that, rather than propose species hypotheses, test the likelihood of competing species hypotheses. These methods are generally designed

to work with multi-locus data and therefore require a lot of computing power. One of the most used methods in this category is BPP (Puillandre et al. 2021).

1.1 Aim

This study aims to investigate and test species boundaries in Scandinavian specimens of *Nomada flavoguttata* using an integrative approach. The following hypotheses were formed at the onset of the project:

(i) Within *Nomada flavoguttata,* there is more than one species.
(ii) The genetic diversity within *Nomada flavoguttata* is linked to host preference.
(iii) *Wolbachia* does not affect the genetic diversity of mtDNA in *Nomada flavoguttata*.

2.0 Method

2.1 Fieldwork

This study was conducted in the southern part of Norway. The specimens collected in this study were collected from 9th May – 29th June 2022. Fieldwork was primarily conducted during sunny periods when bees are known to be most active. *Nomada flavoguttata* specimens were sampled from 14 localities (Fig. 3). The sampling sites were found in situ by looking for suitable habitats. A site was abandoned if no *N. flavoguttata* or *Andrena* (*Micrandrena*) were found after a while of searching.

The specimens were collected using a sweep net and stored in 96% ethanol to preserve DNA. *Nomada flavoguttata* is quite small and anonymous, which makes it difficult to spot in the field. Therefore, sweeping the vegetation proved to be the most efficient method. The goal was to collect *N. flavoguttata* and a series of *Andrena* from the subgenus *Micrandrena* from each locality.

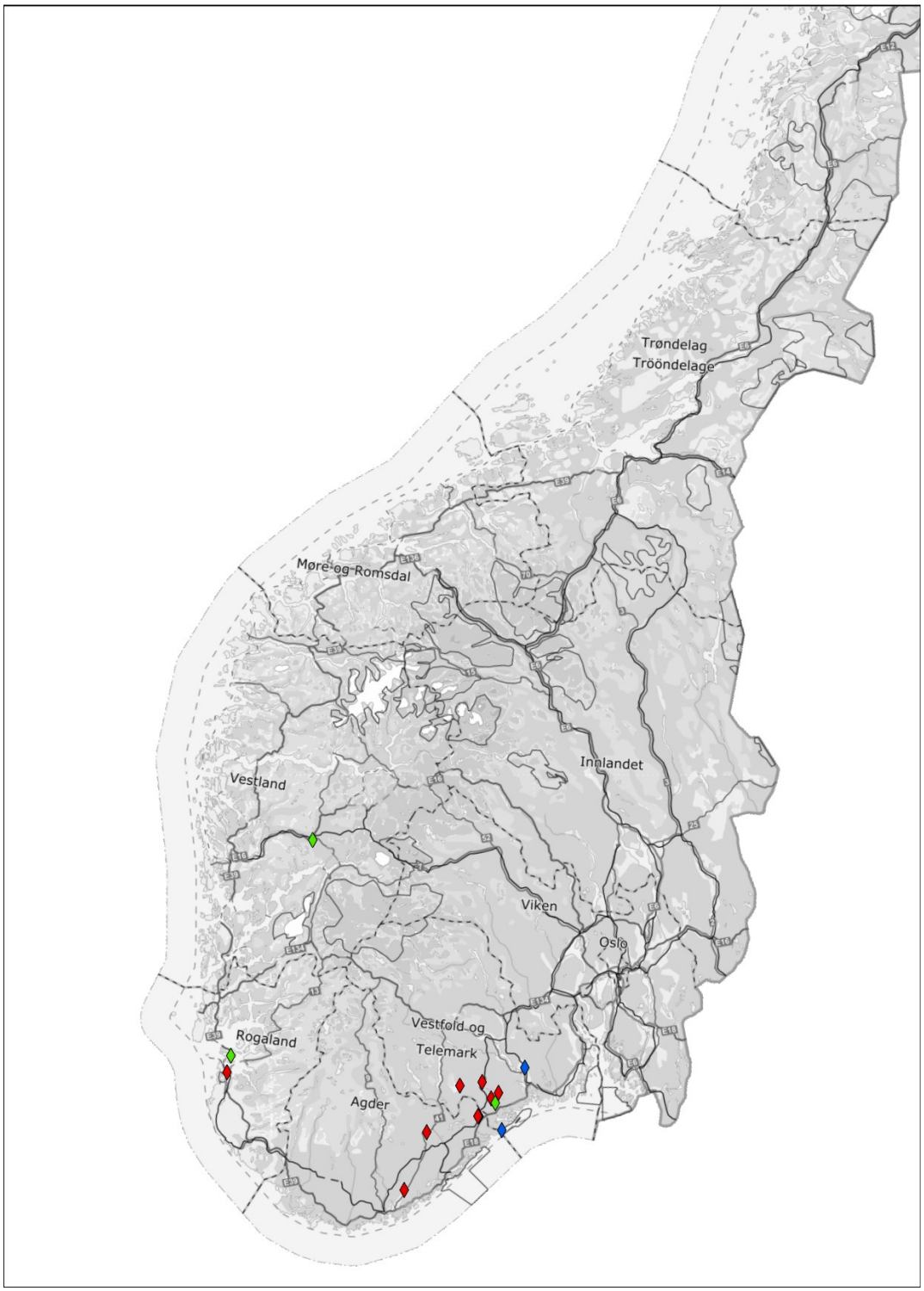


Figure 3. Sampling localities in present study. Blue = *N. flavoguttata* + *A. semilaevis*. Green = *N. flavoguttata* + *A. subopaca*. Red = *N. flavoguttata* with no, or only a few *Andrena*. Map created in ArcGIS pro.

2.2 Obtaining DNA sequences

DNA was extracted using the DNeasy Blood & Tissue kit from Qiagen following the recommended protocol, except that the lysis was carried out overnight in a thermo block

shaking at 700 rpm. Initially, whole specimens were used for non-destructive DNA extraction. Pinned specimens were pulled off the pin and placed directly in the lysis buffer. Ethanol-preserved specimens were dried on filter paper at room temperature for some minutes before they were punctured a few times using a small new/sterile insect pin. After lysis, the specimens were cleaned two times in water, followed by two times in ethanol, then remounted on insect pins.

I noticed that the lysis degraded the specimens by obscuring color and ruining pilosity. Therefore, DNA was extracted using three legs on additional specimens. The first samples extracted using three legs were eluted with 100 µl AE buffer. Due to low DNA concentration, the rest of the specimens were eluted with 50 µl two times, with an extra incubation at 56 °C for five minutes before each centrifuging.

Five DNA markers were included in this study: the mitochondrial COI (Folmer et al. 1994), 16S (Costa et al. 2003), and the nuclear EF1-α, 28S (Cardinal et al. 2010) and ITS-2 (Ji et al. 2003) (Table 1). Before the sampling season, these markers were tested on eight museum specimens already barcoded with COI. The DNA was extracted as described above, using the whole specimens. Initially, all the markers were tested on eight specimens sampled in this study. Due to preliminary results, I decided to continue with COI, 16S, and EF1-α for all my samples.

Table 1. Primers that were used in this study. Dir = Direction, F = Forward, R = Reverse.

DNA marker	Primer name	Dir	Sequence (5'-3')	Source
COI	Lco1490	F	GGT CAA CAA ATC ATA AAG ATA TTG G	(Folmer et al. 1994)
COI	Hco2198	R	TAA ACT TCA GGG TGA CCA AAA AAT CA	(Folmer et al. 1994)
EF1-α	HaF2For1	F	GGG YAA AGG WTC CTT CAA RTA TGC	(Cardinal et al. 2010)
EF1-α	F2-rev1	R	AAT CAG CAG CAC CTT TAG GTG G	(Cardinal et al. 2010)
ITS-2	NG 02955	F	ATG AAC ATC GAC ATT TCG AAC GCA CAT	(Ji et al. 2003)
ITS-2	AB 052895	R	TTC TTT TCC TCC GCT TAG TAA TAT GCT TAA	(Ji et al. 2003)
16s	LR13943F	F	CAC CTG TTT ATC AAA AAC AT	(Costa et al. 2003)
16s	LR13392R	R	CGT CGA TTT GAA CTC AAA TC	(Costa et al. 2003)
28s	A-28S-For	F	CCC CCT GAA TTT AAG CAT AT	(Cardinal et al. 2010)
28s	Mar28Srev	R	TAG TTC ACC ATC TTT CGG GTC CC	(Cardinal et al. 2010)
28s	Bel28S	F	AGA GAG AGT TCA AGA GTA CGT G	(Cardinal et al. 2010)
28s	28SD4Rev	R	GTT ACA CAC TCC TTA GCG GA	(Cardinal et al. 2010)
WSP	wspF	F	GTC CAA TAR STG ATG ARG AAA C	(Sazama et al. 2017)
WSP	wspR	R	CYG CAC CAA YAG YRC TRT AAA	(Sazama et al. 2017)

The PCR reactions were prepared in a volume of 20 µl, consisting of 2 µl dNTP, 2 µl 10X Buffer taql, 2 µl of 2 µM primer, 0.1 µl Takara EX Taq HS, 1 µl DNA, and remaining volume Hyclone Biology Grade Water. The DNA amount used was 3 µl for the specimens extracted using legs. For some samples, a 25 µl PCR reaction was prepared with 12.5 µl Multiplex PCR master mix 2x, 2.5 µl of 2 µM primer, 2 µl DNA, and 8 µl H_2O. The latter method yielded rather poor results for some of the nuclear markers and on the mitochondrial markers from the samples extracted using legs. Therefore, the 20 µl PCR reaction with Takara EX Taq HS was used for the most part.

The PCR setup for all the reactions started with 95°C for 5 mins and ended with 72°C for 10 mins (28S & EF1-α) or 5 mins (COI, ITS-2, 16S). For COI, the program included five cycles of (94°C 30 sec, 45°C 30 sec, 72°C 1 min) followed by 35 cycles of (94°C 30 sec, 51°C 30 sec, 72°C 1 min). For 28S and EF1-α, 35 cycles of (94°C 1 min, 54°C (58°C for 28S) 1 min, 72°C 1.5 min) (Cardinal et al. 2010). For 16S, 35 cycles of (94°C 1 min, 42°C 1.5 min, 64°C 1.5 min) (Costa et al. 2003). For ITS2, 45 cycles of (95 °C 20 sec, 62°C 40 sec, 72°C 20 sec) (Ji et al. 2003).

N. flavoguttata were tested for presence of *Wolbachia* using 20 µl PCR reactions consisting of 10 µl Multiplex PCR master mix 2x, 2 µl of 2 µM primer, 4 µl DNA, and 4 µl H_2O. The primers used were wspF and wspR, and the PCR protocol was as in Sazama et al. (2017).

Gel electrophoresis on a 2% agarose gel was used to test the PCR product (Fig. 4), generally at 90 volts, for about one hour. Successful PCR products were cleaned by adding 1 µl Exo-SAP-IT™ and incubating the mixture in a thermocycler for 15 mins at 37°C, followed by 15 mins at 80°C. After cleaning, the PCR products were sent to Eurofins for bi-directional sequencing using BigDye termination.

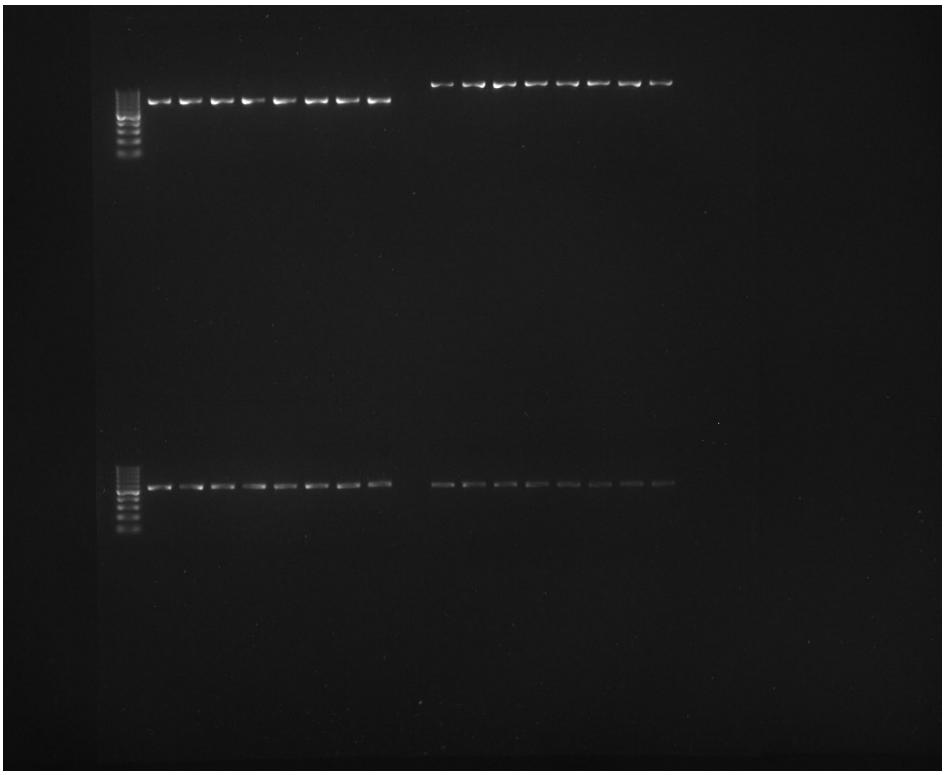


Figure 4. Example of PCR product visualized by gel electrophoresis. Top left (before gap) showing samples with COI, top right EF1-α, bottom 16S.

2.3 Sequence editing

The sequences were trimmed and edited using the programs Pregap4 and Gap4 from the Staden Package (Bonfield et al. 1995; Staden 1996). As Gap4 did not allow the special characters for ambiguous sites in the consensus sequence, the sites were left as a "- ", and the positions were noted. After editing, the ambiguous sites were located again in the fasta file, appearing as an "n", and replaced there. The sequences were aligned using the MUSCLE algorithm in MEGA11 (Tamura et al. 2021). After alignment, the primer sequences were trimmed off on both ends. For EF1-α, the primer sequence was not included in the alignment, and the ends were then trimmed off using the chromatograms. A combined dataset was made from the three DNA markers, starting with COI, followed by 16S and EF1-α, respectively. Only specimens with successful sequences from all the markers were included in the combined dataset.

2.4 Analyses

Maximum likelihood (ML) trees were constructed in MEGA 11 using partial deletion (>95%) and 500 bootstrap replications. The best substitution model was found using the "Find Best DNA/Protein Models (ML)" function (Hall 2013). The best models were Tamura 3-parameter

with gamma distribution (COI, combined dataset), Tamura 3-parameter (16S), and Kimura 2-parameter (EF1- α)(Kimura 1980; Tamura 1992).

TCE haplotype networks for COI and 16S were constructed using the program PopART (Clement et al. 2000; Clement et al. 2002; Leigh et al. 2015).

Ultrametric trees were constructed for each gene fragment using BEAST v2.7.3 (Bouckaert et al. 2019). Default settings were used, but the prior was changed to Coalescent Constant Population and the substitution model as described below. The number of generations was set to 10 million, sampling every 1000 trees. Due to low effective sample size (ESS) values, 50 million generations were run on the combined dataset, sampling every 5000 trees. Due to an obvious overestimate of species from the COI tree in GMYC, another COI tree was created, using only 2-3 sequences from each haplotype and including four additional outgroup species. Additional outgroups were added because GMYC may perform worse when few species are involved (Talavera et al. 2013). The new COI tree was run for 50 million generations for the same reason as the combined dataset. As not all the substitution models suggested by MEGA were available in BEAST, the best substitution models to use in BEAST were found using jModelTest and BIC calculations (Posada 2008; Schwarz 1978). The models with the lowest BIC scores were TIM3+I (COI, combined dataset), TPM3uf (16s), and TrNef (EF1- α) (Kimura 1981; Posada 2003; Tamura & Nei 1993). The SSM package for BEAUti was used to apply the models (Bouckaert & Xie 2017). Two individual runs were performed and combined with a 10% burn-in using the program LogCombiner. Tracer v1.7.2 was used for MCMC convergence diagnostics (Rambaut et al. 2018). Maximum clade credibility (MCC) trees were constructed in TreeAnnotator 2.7.3 using mean node height.

A single threshold GMYC analysis was conducted in R, using the RStudio GUI and the splits package (Ezard et al. 2009; Posit team 2023; R Core Team 2023). The ultrametric trees created in BEAST, and associated programs, were used in the GMYC analysis. The analysis was performed for each gene fragment, and on the combined dataset. A PTP and mPTP analysis were conducted for each gene fragment at https://species.h-its.org/ptp/ and https://mptp.h-its.org/#/tree, respectively (Kapli et al. 2017; J. Zhang et al. 2013). Maximum likelihood trees, consisting of single sequences of unique haplotypes with 3-5 outgroup species, were used as input trees in the PTP and mPTP analyses. Sequences from additional outgroup species were downloaded from GenBank based on similarity and availability. An

ASAP analysis for each gene fragment was performed at https://bioinfo.mnhn.fr/abi/public/asap/ using the Kimura 2 parameter substitution model (Puillandre et al. 2021). The command line version of BPP (Flouri et al. 2018; Yang 2015) was used following the tutorial in Flouri et al. (2020). Both (A10) and (A11) analyses were performed, running up to 500 000 MCMC generations with a burn-in of 10% (Rannala & Yang 2013, 2017; Yang & Rannala 2010, 2014).

2.5 Morphology

Pictures for morphometric analyses were taken with a camera attached to a Leica DM 500 stereomicroscope. The morphometric measurements were conducted in the LAS X software. The number of specimens included in the different measurements varied, as some of the characters were damaged/missing in some of the specimens. Morphological characters were selected based on what I found variable and what has been used on *Nomada* previously. A total of 17 characters were measured (Table 2). The pygidial plate notch depth measurements were excluded from the analysis, as they seemed very random and caused difficulties in the analyses. The pygidial plate angle was also excluded from the analysis, as it is an angle rather than a length measurement. The characters were measured as shown in (Figs. 5-6). When possible, measurements were done on both sides, and the mean value was used. Characters were measured when a measurement's start and end point was in focus simultaneously. An exception was the measurements in (Fig. 5B), which were measured when the top edge of each eye was in focus simultaneously as illustrated by the green arrows. Due to very few female specimens, measurements were only done on males. In addition to the characters measured, qualitative characters, e.g., coloration, were investigated to see if any could be used to separate the clades. A principal component analysis (PCA) was performed in RStudio using the package "stats" (R Core Team 2023). The R-package and the tutorial in Baur and Leuenberger (2020) were used to perform a shape PCA (Baur et al. 2014; Baur & Leuenberger 2011). Both PCA and shape PCA were visualized using the R package ggplot2 (Wickham 2016). LDA ratio extractor was used to find the best ratios to discriminate the clades (Baur et al. 2014; Baur & Leuenberger 2011, 2020). Only specimens where I had measurements from all characters were included in the PCAs and LDA extractor. This included 19 specimens where eight, six, and five were from clade 1, clade 2, and clade 3, respectively, measurements available in (Appx. 1). The isometric size (isosize) was extracted from the shape PCA analysis and used as a measurement for size. Isosize is a

way to assess overall body size using several characters. It is especially useful when no single character is a good measurement for size (Coelho et al. 2017).

Table 2. Characters that were measured and used in the PCA and LDA analyses.

Character	Description
GMW	Gonocoxite max width
GSW	Gonocoxite smallest width, below max width
GL	Gonocoxite length
PW300	Width of pygidial plate 300 µm from tip
PW100	Width of pygidial plate 100 µm from tip
PA	Sloping angle of the pygidial plate, calculated from PW300 & PW100
PND	Pygidial plate notch depth
PTT	Distance between pygidial plate tips
A3	Antennal segment 3 directly from below
A4	Antennal segment 4 directly from below
A5	Antennal segment 5 directly from below
OP	Distance between posterior ocelli
OE	Shortest distance between eye margin and posterior ocelli
OA	Distance between posterior and anterior ocelli
FW	Face width at antennal height
FL	Face length, from anterior ocellus to anterior margin of clypeus
Costa	Length of costa of the fore wing from the humeral plate to the start of pterostigma

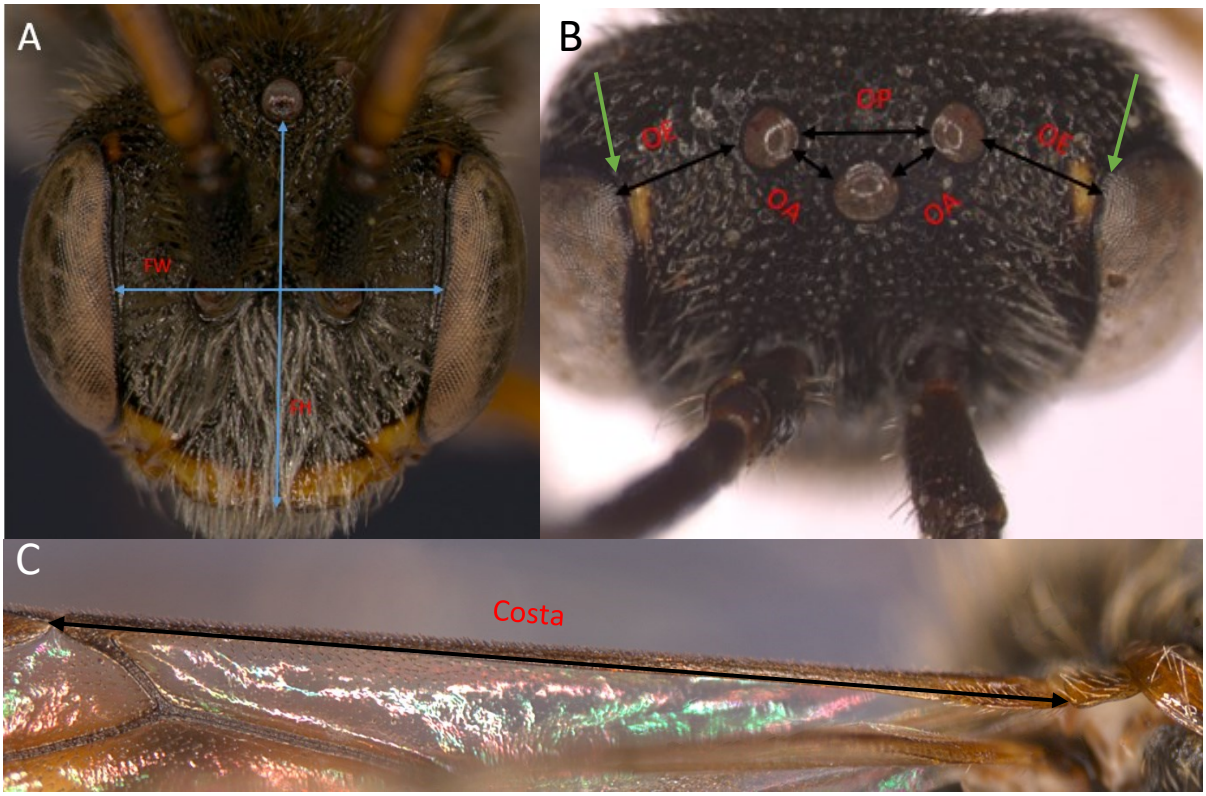


Figure 5. A) Face seen directly from the front. FW = Face width at antennal height, FH = Face height between anterior ocellus and anterior margin of clypeus. **B)** Head seen directly from above. OE = Shortest distance between eye margin and posterior ocelli, OP = Distance between posterior ocelli, OA = Distance between posterior and anterior ocelli. The measurements were done when the top edge of each eye, shown by the green arrows. were in focus simultaneously. **C)** Forewing seen from below, measurement of costa done from the humeral plate to the start of pterostigma.

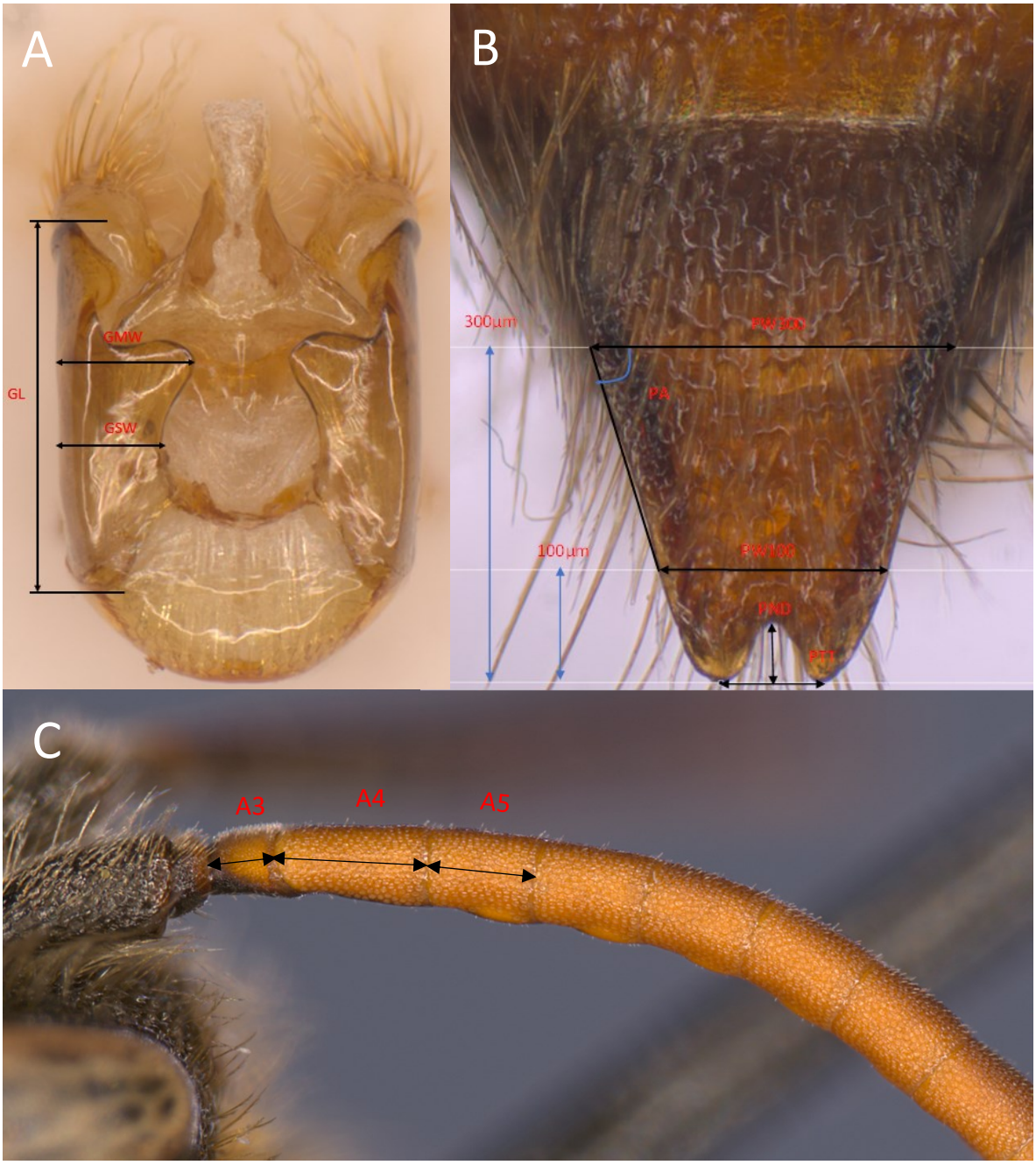


Figure 6. A) Male genitalia capsule viewed from the dorsal side. GL = Gonocoxite length, GMW = Gonocoxite max width, GSW = Gonocoxite smallest width, below max width. **B)** Pygidial plate viewed directly from above. PW300 = Width of pygidial plate 300 µm from tip, PW100 = Width of pygidial plate 100 µm from tip, PA = Sloping angle of pygidial plate, PND = Pygidial plate notch depth, PTT = Distance between pygidial plate tips. **C)** Antenna viewed directly from below. A3-A5 = Antennal segment 3-5.

3.0 Results

3.1 Data collection

I collected 40 *Nomada flavoguttata* specimens from 13 localities (Table 3). Out of the specimens, 29 were males, and 11 were females. A minimum of four *N. flavoguttata* specimens and a series of *Andrena* (*Micrandrena*) were found at five of the localities. Only one species of Andrena was found in four of these, while the last locality was dominated by one species (12 *A. semilaevis* + 2 *A.* subopaca). From two of the localities where I also found *Andrena* (*Micrandrena*), I only found males of *N. flavoguttata*. In addition to the 40 *N. flavoguttata* specimens, I found one specimen initially thought to be an abnormal *N. flavoguttata.* However, the DNA barcoding revealed that it was a small abnormal *Nomada cf. panzeri*. All the female specimens collected in this study were assigned to clade 1, and from the museum material, I only had one female from clade 3 and zero from clade 2.

Table 3. Sampling localities, number of samples, and associated host. County: VT = Vestfold & Telemark, RO = Rogaland, VE = Vestland, AG = Agder.

Locality name	**County**	**Municipality**	**Date**	**Lat**	**Long**	***N. flavoguttata***	***Andrena* Host**
Kurdøla	VT	Kragerø	23,26.V.22	58,966	9,258	7 (6M, 1F)	14 *subopaca*
Litla Stokkavatnet	RO	Stavanger	31.V.22	58,968	5,690	2F	NA
Rennesøy, Dale	RO	Stavanger	31.V.22	59,084	5,710	7M	14 *subopaca*
Vossavangen, Gjernesmoen	VE	Voss	01.VI.22	60,616	6,415	6M	9 *subopaca*
Gjerpen	VT	Skien	06.VI.22	59,223	9,610	4 (2M, 2F)	12 *semilaevis* + 2 *subopaca*
Nordskogen, Hestånmonen	AG	Åmli	10.VI.22	58,725	8,394	3M	1 *subopaca*
Vestre Mollestad, Saga	AG	Birkenes	10.VI.22	58,317	8,177	2 (1M, 1F)	NA
Henseid, Springheim	VT	Drangedal	12.VI.22	59,037	9,289	1M	NA
Stangnes, Furumyr	VT	Kragerø	13.VI.22	58,785	9,374	4 (3M, 1F)	15 *semilaevis*
Gjerstad, Kufjell	AG	Gjerstad	20.VI.22	58,866	9,057	1F	NA
Gjerstad, Vik	AG	Gjerstad	20.VI.22	58,866	9,045	1F	3 *subopaca*
Prestestranda, Stemmenåsen	VT	Drangedal	22.VI.22	59,101	9,060	1F	NA
Gautefall, Kittilsbu	VT	Drangedal	29.VI.22	59,063	8,771	1F	NA

3.2 DNA sequences

The older museum specimens did not give any sequences for any of the nuclear markers. For the initial eight specimens collected in this study, the reverse read of ITS-2 was generally of

bad quality, making the sequences less reliable. The editing and alignment of 28S were, in general, quite straightforward. However, 28S had very little variation and was therefore excluded from further specimens.

Out of 38 samples, 36 yielded successful COI sequences. Together with the 34 already barcoded specimens, I had a total of 70 COI sequences for *N. flavoguttata*. After trimming off the primer sequences, the alignment for COI had a sequence length of 658 base pairs (bp), including 24 variable sites and 19 parsimony informative sites (12 variable sites and 10 parsimony informative sites when only looking at the 36 sequences from samples collected in this study). For 16S, all 38 samples were successful. The sequence length was 515 bp, with 6 variable sites and 4 parsimony informative sites. For EF1-α, 36 of the samples were successful. The sequence length was 948 bp, with 8 variable sites and 5 parsimony informative sites. The combined dataset consisted of 34 sequences, with a length of 2121 bp, including 26 variable sites and 19 parsimony informative sites.

3.3 Phylogenetic trees

After partial deletion, a total of 623 (COI), 489 (16S), 920 (EF1- α), and 2010 (combined) bp were used to construct the ML trees.

The original three clades from (Fig. 1) remain as separate clades in the COI tree (Fig. 7). Overall, the tree for COI is relatively well supported, especially the division of the three main clades. Three of the specimens collected in this study were assigned to clade 2, while the rest were assigned to clade 1. Within clade 1, there are three subclades. None of the specimens sampled in this study were assigned to clade 3. In the 16S tree (Fig. 8), clade 1 from the COI tree is no longer monophyletic, as clade 2 is placed within clade 1. The overall genetic variation is lower than in COI, and the support is weak. In the tree resulting from analyses of the EF1-α sequences (Fig. 9), clades 1 and 2 are not separated. The tree structure reflects the low genetic divergence in the ingroup, and most clades have low support, except for one moderately supported subclade of four specimens. In the tree constructed using the combined dataset, clade 1 is again not monophyletic (Fig. 10). There is a subclade consisting of two specimens from clade 1 outside of the clade which contains clade 2 and the remaining clade 1 specimens.

Figure 7. Maximum likelihood tree inferred from COI, 500 bootstrap replicates. Bootstrap values above 50 are shown. ★ = Positive for Wolbachia.

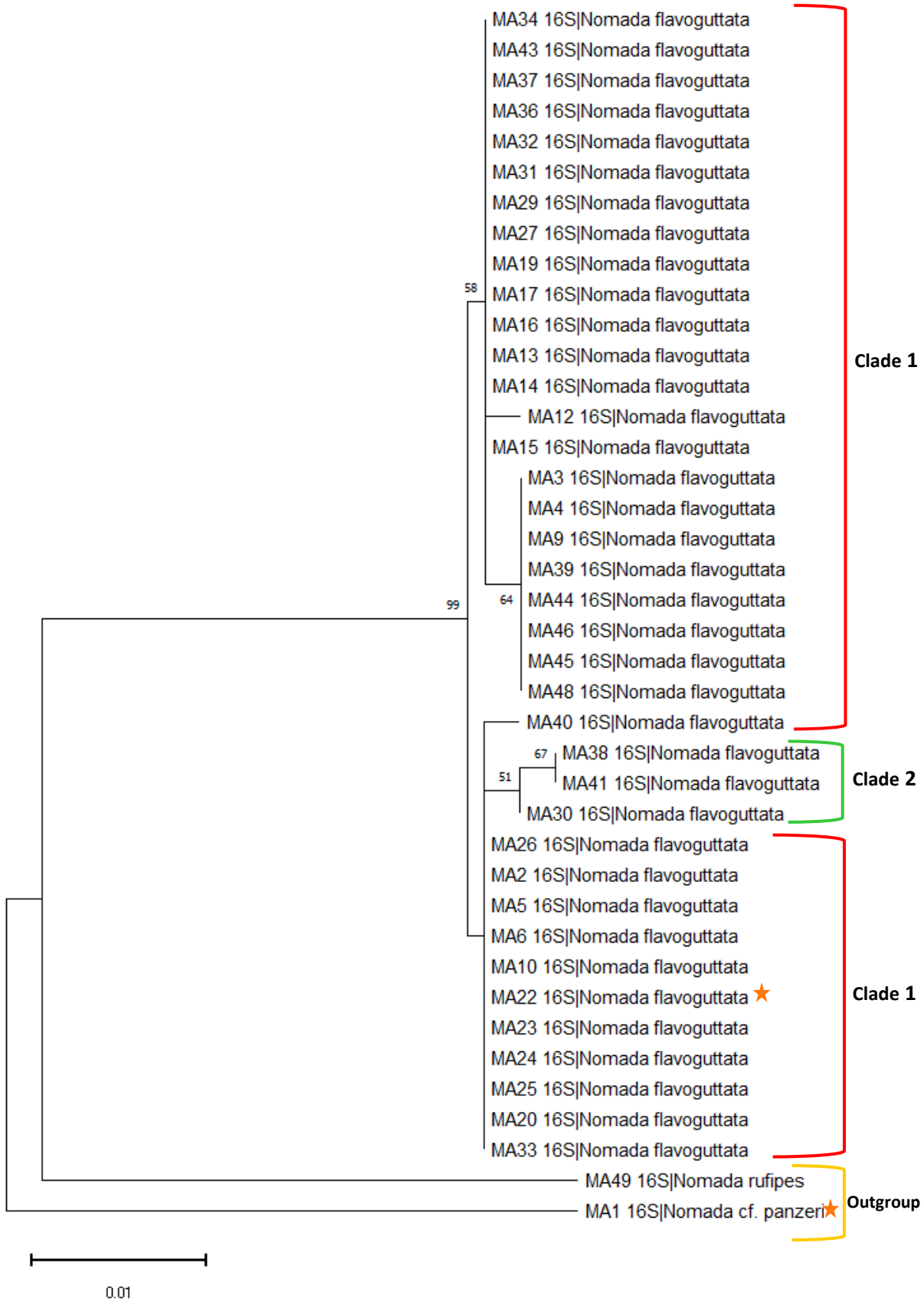


Figure 8. Maximum likelihood tree inferred from 16S, 500 bootstrap replicates. Bootstrap values above 50 are shown.
★ = Positive for Wolbachia.

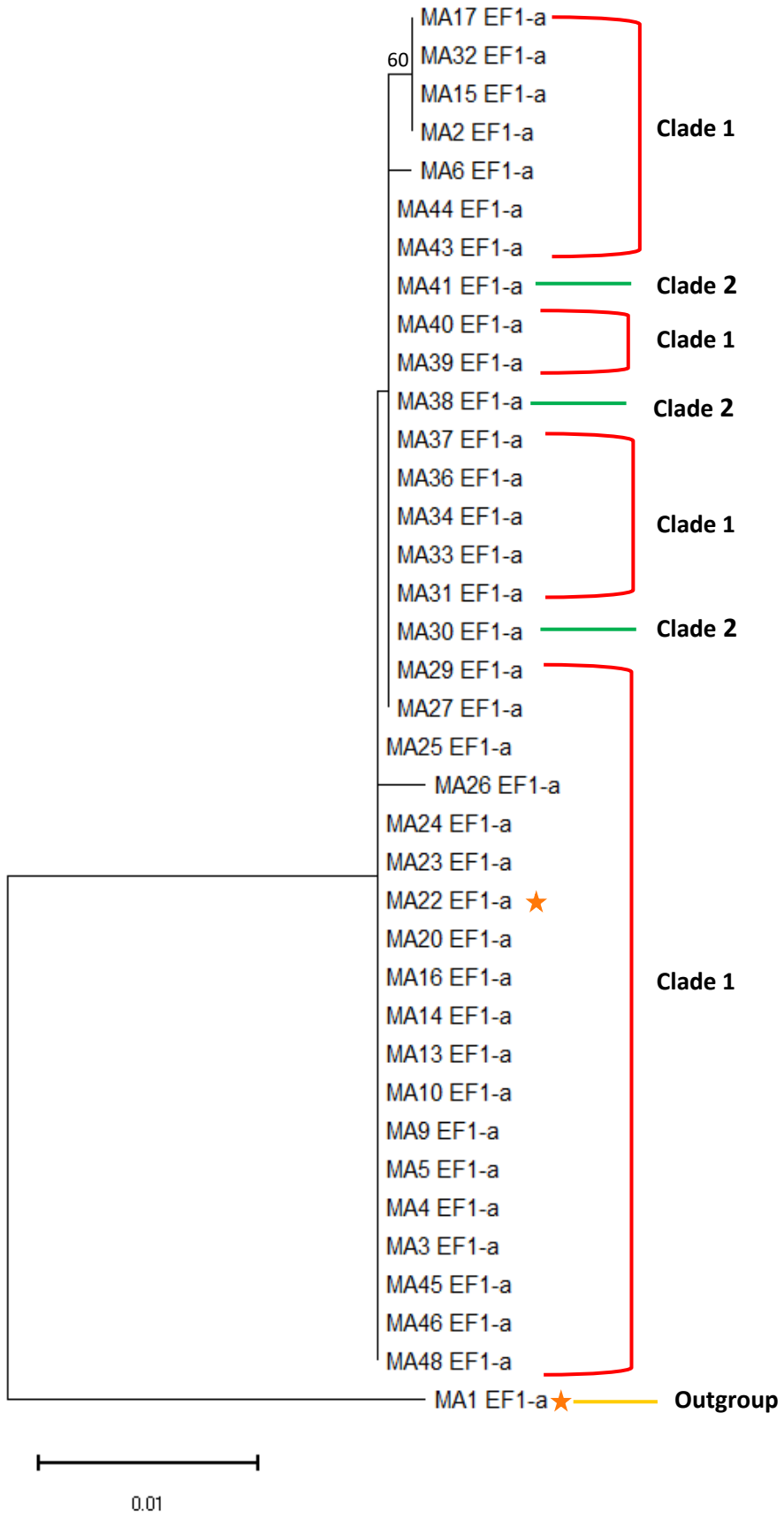


Figure 9. Maximum likelihood tree inferred from EF1-α, 500 bootstrap replicates. Bootstrap values above 50 are shown. ★ = Positive for Wolbachia.

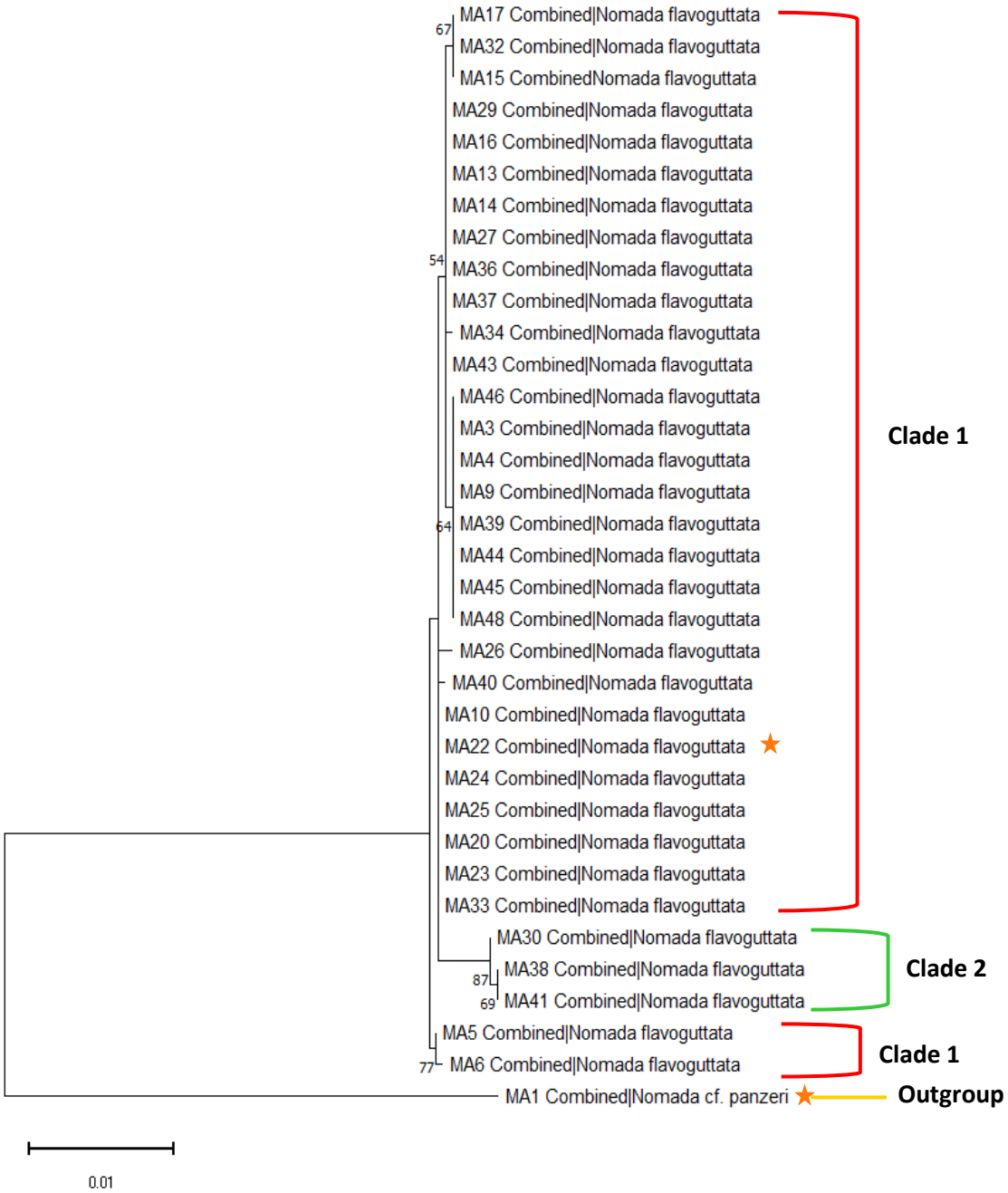


Figure 10. Maximum likelihood tree inferred from the combined dataset, 500 bootstrap replicates. Bootstrap values above 50 are shown.★ = Positive for Wolbachia.

3.4 Haplotype network

The COI haplotype network (Fig. 11) shows that the number of substitutions between the clades are few. However, there are at least three substitutions between each clade. All the specimens collected in association with *Andrena subopaca* were placed in clade 1, while those collected with *A. semilaevis* were placed both in clade 1 and clade 2. None of the specimens collected in this study were placed in clade 3. In the 16S network (Fig. 12), I am missing data from clade 3. The two other clades are still separated, but only with a minimum of one substitution.

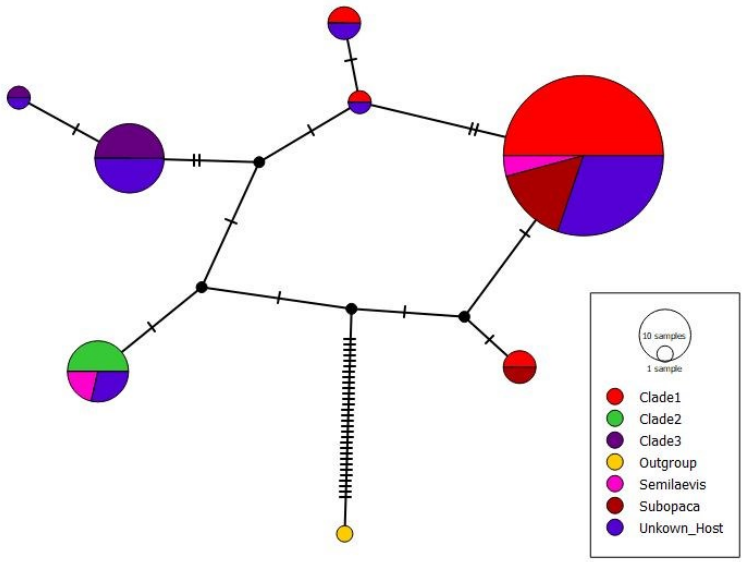


Figure 11. TCE haplotype network for COI sequences. The top half of the circle indicates clade, while the bottom half indicates host association. The outgroup species is *N. cf. panzeri*.

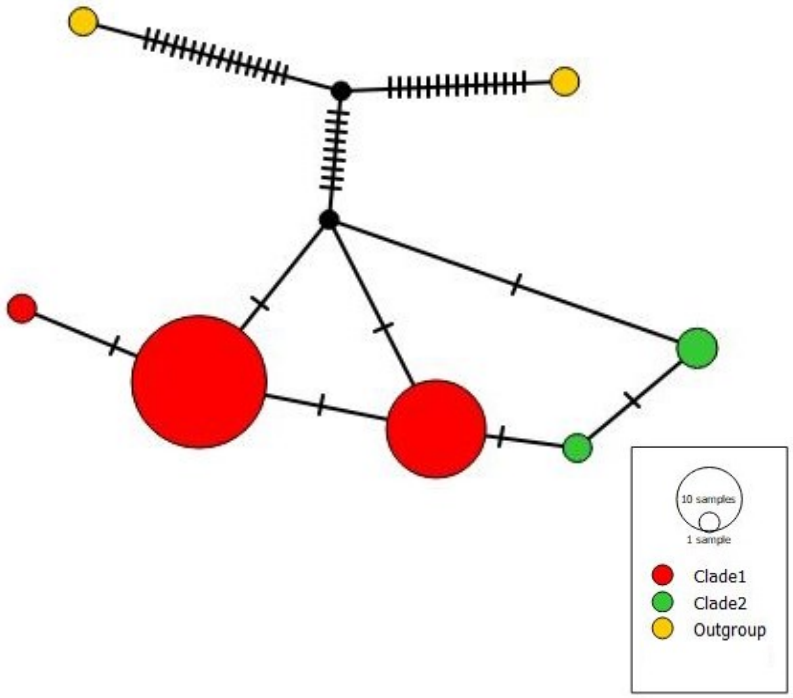


Figure 12. TCE Haplotype network for 16S sequences. The outgroup species are *N. cf. panzeri* & *N. rufipes*.

3.5 Species delimitation

Depending on which method was used and what gene fragment, *N. flavoguttata* was delimited into a various number of species (Table 4). In both versions of PTP and in the ASAP analysis, the result was always one species for the single markers. In the combined dataset, the results from PTP and ASAP were one and two species, respectively. The two species ASAP proposed represent clade 1 and clade 2. In the GMYC analysis on the first dataset for COI, which included all sequences and only one outgroup, *N. flavoguttata* was delimited into 19 entities. The second dataset from COI, and the combined dataset, delimited each of the three clades into their own entities. In GMYC, the other single markers, 16S and EF1- α, delimited *N. flavoguttata* to one and seven entities, respectively. The BPP analysis always had the highest posterior probability for the three clades being three species.

Table 4. Molecular delimitation methods and the number of species delimited. Numbers represent the number of species delimited when outgroups are excluded. * When using an outgroup species more closely related to N. flavoguttata, ASAP suggested, on the COI dataset, that the three clades are three different species.

	GMYC	PTP (mPTP)	BPP	ASAP
COI	3	1(1)	-	1(3)*
16S	1	1(1)	-	1
EF1- α	7	1(1)	-	1
Combined	3	1(1)	3	2

3.6 Morphology

The PCA of the 15 characters shows no apparent separation of the three clades (Fig. 13). In the shape PCA, clade 1 still overlaps the other two clades to a degree but is somewhat separated (Fig. 14). Clades 2 and 3 are still clustered together. Clade 1 differs from the others mainly along shape PC2, where it has a slightly higher value. Shape PC2 was moderately correlated with isosize with a correlation coefficient magnitude of 0.58. The isosize of specimens from clade 1 is, in general, larger than those from clade 2 and clade 3 (Fig. 15).

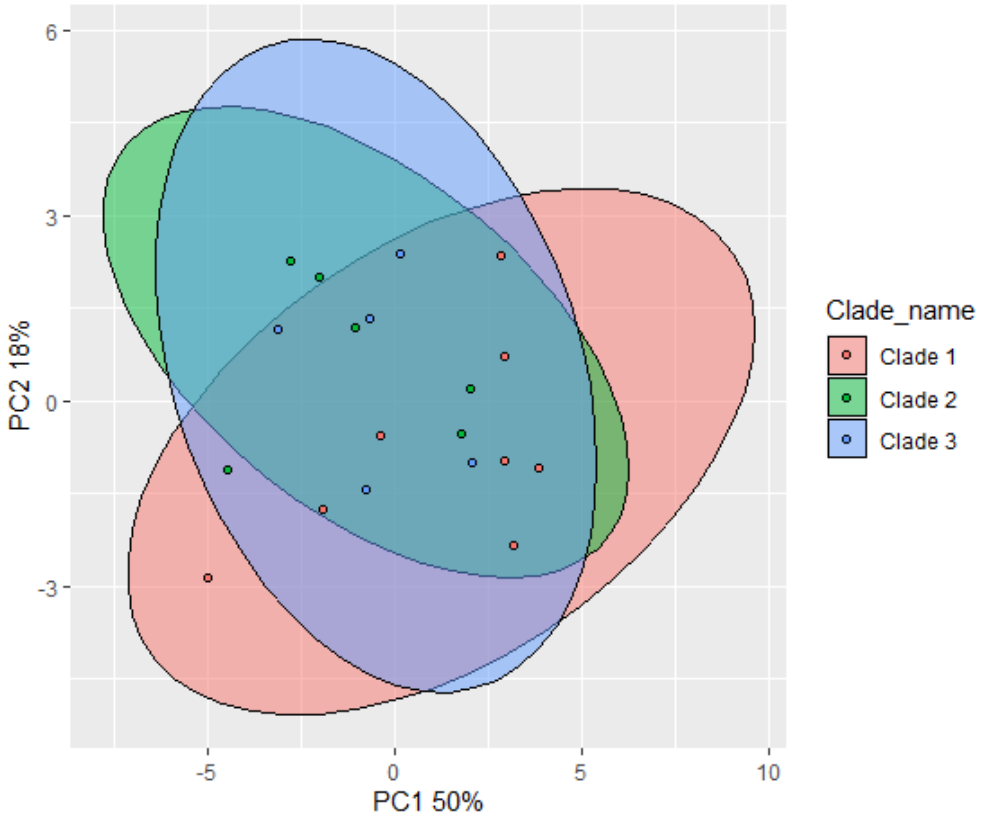


Figure 13. PCA of 15 characters from 19 male specimens, showing a 95 % confidence interval ellipse.

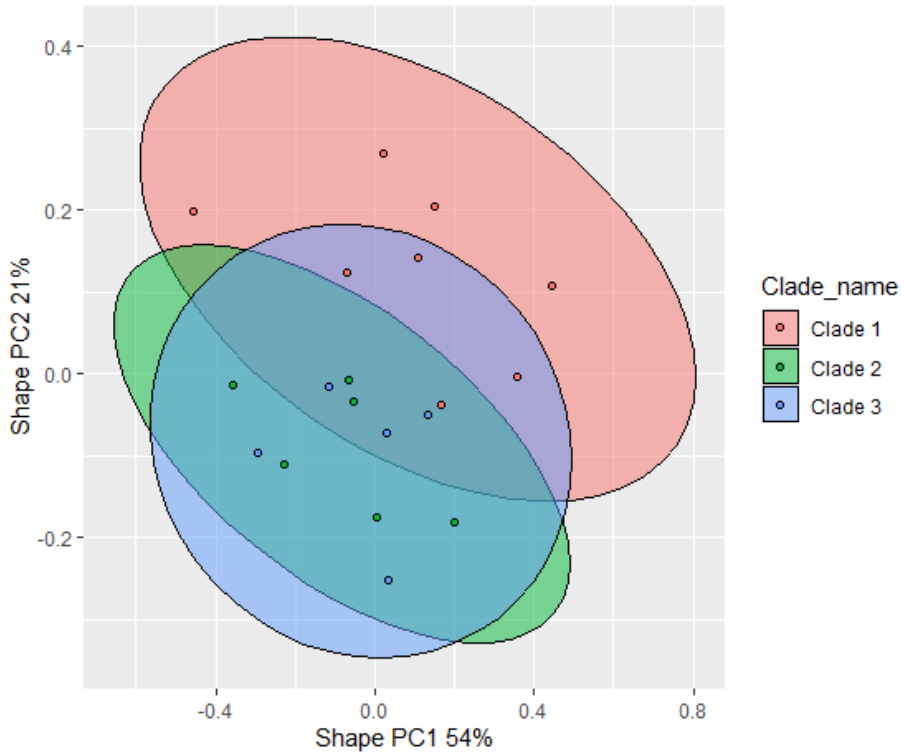


Figure 14. Shape PCA of 15 characters from 19 male specimens, showing a 95 % confidence interval ellipse.

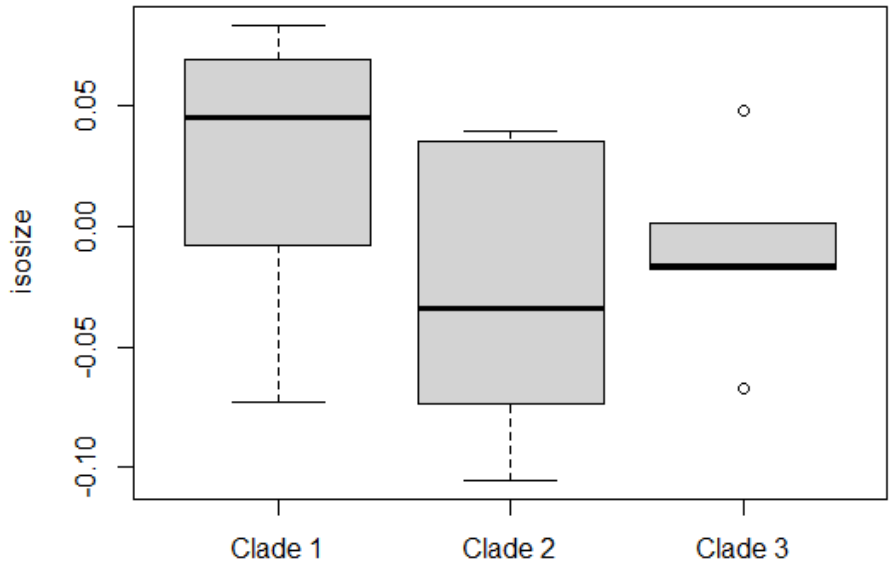


Figure 15. Boxplot comparing the isosize between the three clades.

When excluding all the pygidial plate characters, the two best ratios from the LDA extractor separate the three clades (Fig. 16). Clade 1 is mainly separated from clade 2 by ratio 1 (GMW/OP), while it is separated from clade 3 by both ratios. Clades 2 and 3 are mainly separated by ratio 2 (FL/Costa). The two characters GMW and OP, which were included in the best ratio, had quite different correlations to isosize with a correlation coefficient magnitude of 0.73 and 0.27, respectively. The two characters, FL and Costa, had a similar correlation to isosize with a correlation coefficient magnitude of 0.80 and 0.87, respectively.

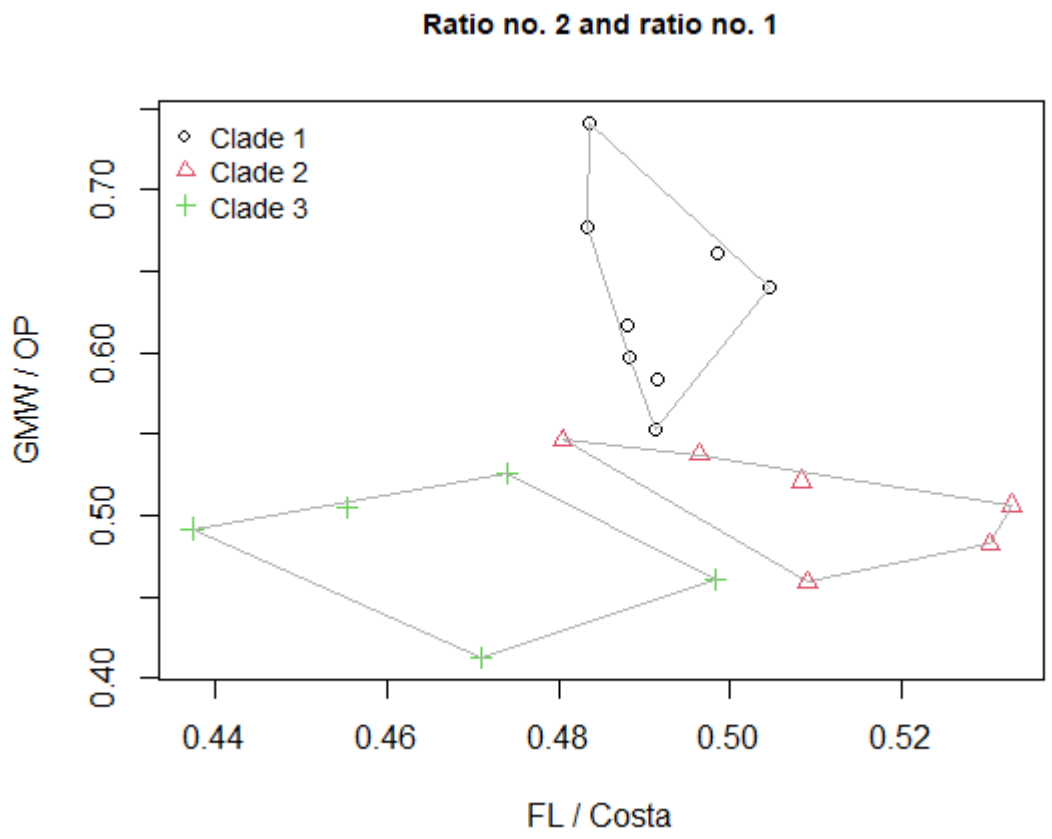


Figure 16. LDA ratio extractor showing the two best ratios for discriminating the three clades. Extracted from 13 characters when all the pygidial plate characters were excluded.

All the females investigated from clade 1 had less than half of the clypeus darkened (Fig. 17A), while the one female specimen from clade 3 had more than half of the clypeus darkened (Fig. 17B). All seven males from clade 2 had a more or less lightened posterior margin of the labrum (Fig. 17C). In contrast, the posterior margin of the labrum was completely black (Fig. 17D) in 19/25 and 6/8 specimens from clade 1 and 3, respectively. The males from clade 2 had a broad yellow anterior band on the clypeus (Fig. 17E), while the center of the anterior band on the clypeus was darkened (Fig. 17F) in 15/24 and 7/8 specimens from clade 1 and 3, respectively. The coloration of tergite 1 was variable, from extensively black to half black with two obvious black dots. All three clades had representatives with obvious black dots and representatives with more extensively black tergite 1.

The pygidial plate slope angle and notch depth were quite variable (Figs. 18 A, B). The angle varied from 53 to 66 degrees, and the notch depth ranged from 10 μm to 41 μm. No pattern, regarding clade, was detected for either character PA or PND. A character often used to identify *Nomada* species is the arrangement of spines on the hind tibia. This character seemed very variable in *N. flavoguttata*, with some specimens having shorter, more robust spines (Fig. 18C) and some having longer and thinner spines (Fig. 18D).

Specimens from clade 3 had, in general, less obvious tyloids on the antennal segments than specimens from clade 1 and 2. However, some clade 3 specimens also had obvious tyloids; on the other hand, some clade 1 and 2 specimens had less obvious tyloids.

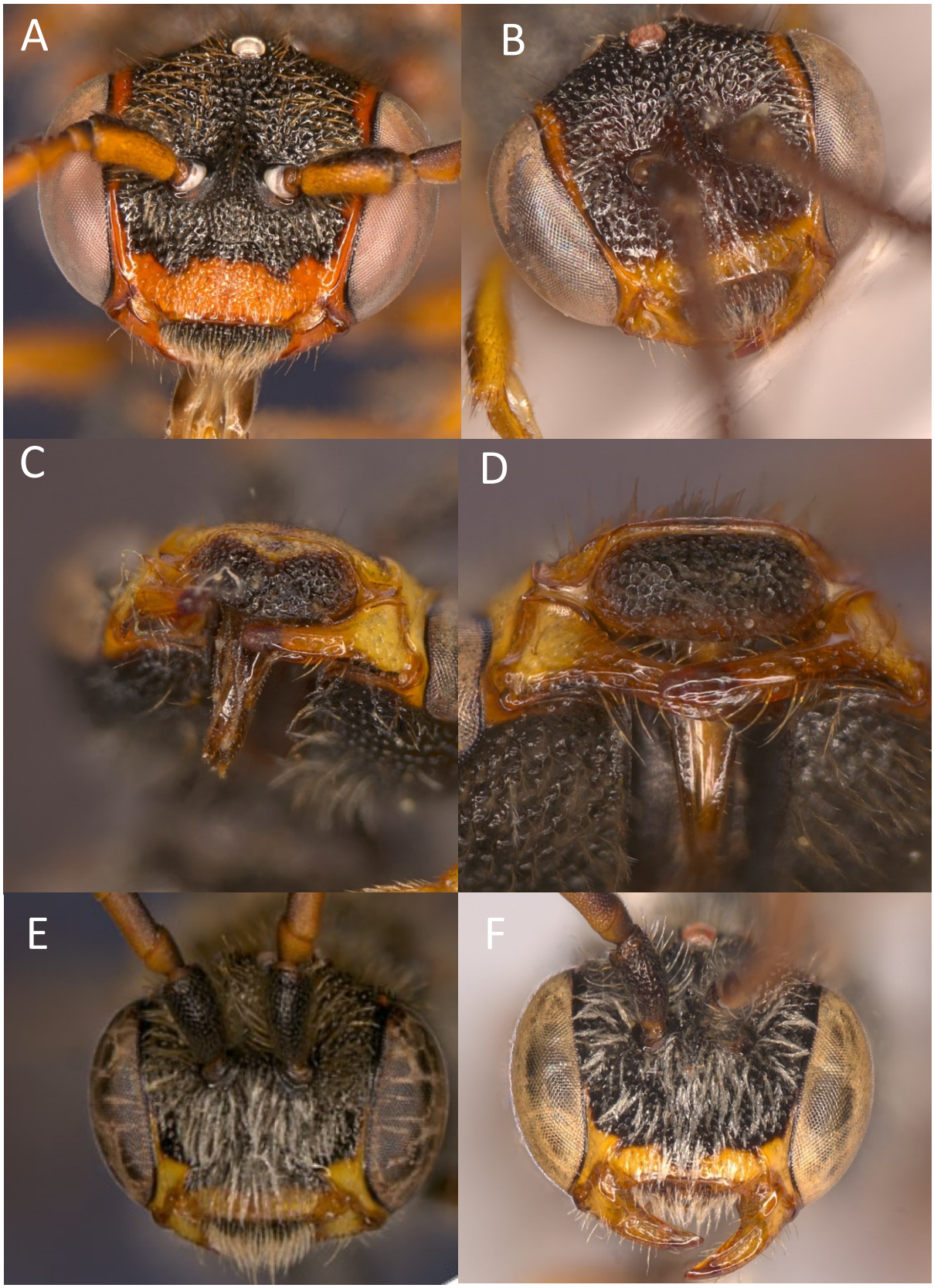


Figure 17. Coloration differences in facial characters **A)** Face of female from clade 1. **B)** Face of female from clade 3. Notice the lighter clypeus of the clade 1 female. **C)** Labrum of a male from clade 2 showing a lightened posterior margin. **D)** Labrum of a male from clade 1 showing a completely black posterior margin. **E)** Male from clade 2 showing a broad yellow band anterior on clypeus. **F)** Male from clade 3 showing a dark patch on the yellow band on clypeus.

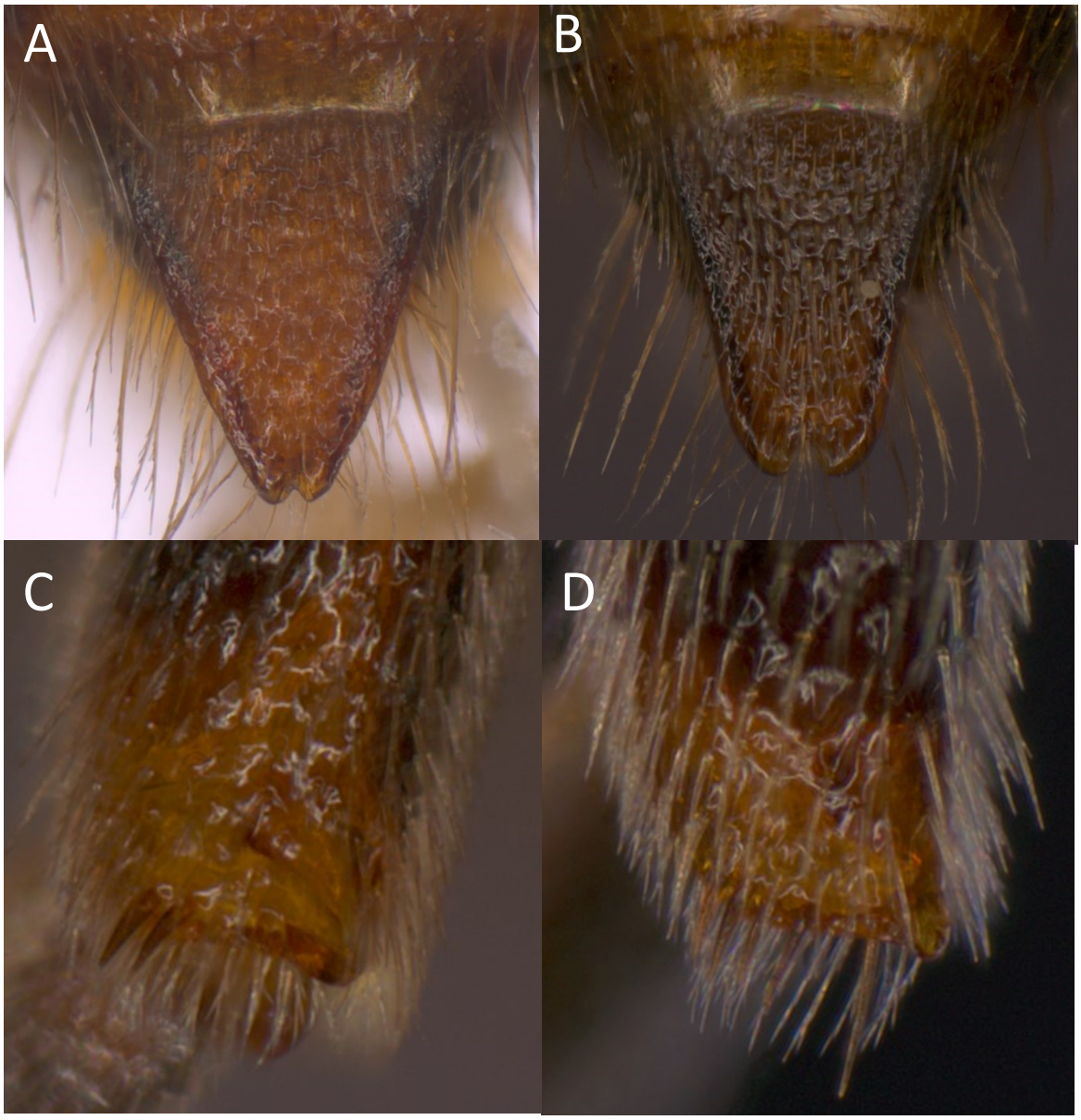


Figure 18. Pictures showing differences in pygidial plate angle and arrangement of hind tibial spines **A)** Showing a quite pointy pygidial plate (53 °) with a very shallow notch in the end. **B)** Showing a less pointy pygidial plate (66 °) with a relatively deep notch in the end. **C)** Hind tibia of a male from clade 1, showing short robust spines. **D)** Hind tibia of a male from clade 1, showing longer and thinner spines.

3.7 *Wolbachia*

I tested 39 *N. flavoguttata* specimens, and the *N. cf. panzeri* for *Wolbachia*. One *N. flavoguttata* and *N. cf. panzeri* tested positive, giving *N. flavoguttata* an infection rate of 2.6%. The final *Wolbachia* sequences from the two specimens were 538 bp long and identical. The sequences were validated as *Wolbachia* with a search in BLAST. The best match was 100% with a *Wolbachia* sequence isolated from *Nomada flava* in Gerth et al. (2013).

4.0 Discussion

4.1 Phylogeny

Only in the COI tree do the three proposed clades remain monophyletic. In both the combined and the 16S tree, clade 1 is paraphyletic, while in EF1-α, none of the clades are recognizable. I only had information on clade 3 in the COI dataset. Thus, where specimens assigned to clade 3 in the COI tree would end up in the 16S or EF1-α tree is difficult to say. In some invertebrate groups, the substitution rate of nuclear genes is considerably slower than in mtDNA (Allio et al. 2017), which could explain the low genetic variation in EF1-α. *Nomada cf. panzeri* might not be the best outgroup species, as it is not very closely related to *N. flavoguttata*. When looking at COI sequences, the two closest related *Nomada* species to *N. flavoguttata* are *N. conjungens* and *N. castellana*. Neither of these are present in Norway, so I did not have them available. In addition, very few 16S and EF1-α sequences from *Nomada* species were available in GenBank, so those available were used as additional outgroups.

4.2 Species delimitation

The number of species delimited varied between the methods and the dataset used. One to nineteen species were proposed, and the two most frequent scenarios were that *N. flavoguttata* is either one or three species. Both PTP analyses and ASAP always resulted in *N. flavoguttata* being one species, except from ASAP on the combined dataset. ASAP delimited clades 1 and 2 as separate species on the combined dataset. I did not have any specimens from clade 3 in this dataset. For COI, ASAP also had a low asap-score for all the clades being separate species. As ASAP work with barcode gaps, having more closely related outgroups is better, as a distantly related outgroup species may create an artificial high gap distance. This was tested by adding the closest related species found in GenBank, *N. castellana.* This resulted in the lowest asap-score to group *N. flavoguttata* and *N. castellana* as a single species. On the contrary, when only *N. castellana* was kept as outgroup species, the lowest score suggested that the three *N. flavoguttata* clades should be separate species. These conflicting results, based only on changes in the outgroup species, indicate that when few species are involved, ASAP may be sensitive to how related the outgroup is to the ingroup.

The number of entities from GMYC was rather variable. The initial 19 entities for COI and 7 entities for EF1-α is an obvious overestimate, while one entity from 16S and three from the second COI tree and the combined dataset are more reasonable. A factor that may have led

to the vast overestimation is that I only had one outgroup species in the initial datasets. Using species-poor data sets may lead to overestimation (Dellicour & Flot 2015). It is suggested to avoid running GMYC on fewer than five species. When interested in a few species only, additional outgroup species should be added to keep the number of species at a level where GMYC works reasonably well (Talavera et al. 2013). Another issue with GMYC is that it has quite a high misidentification ratio, about 20%. Therefore, it is suggested that GMYC entities should be treated as potential species rather than actual species (Talavera et al. 2013).

The BPP analyses were worryingly confident in the three clades being three species, almost always having a posterior probability of 1. This is unexpected and seems unlikely as the genetic difference is not very high between the clades. Further splitting clade 1 into two clades gave a posterior probability of 0.65 for clade 1 consisting of two species. This also suggests that my dataset might not work well with BPP. As the older museum specimens yielded no results from the nuclear markers, I had minimal data on other markers than COI. Since BPP is primarily built for working with multiple loci, it may struggle to give reliable outputs when the input data is mainly from one locus.

4.3 Ecology

Some species cannot be distinguished using DNA sequences, which is also the case in some *Nomada* species (Falk et al. 2022). Other hymenopterans with similar lifestyles as *Nomada* may also be difficult to determine using DNA sequences or morphology (e.g. Soon et al. 2021). The reason behind low interspecific genetic differences may be recent radiation of the species involved (Falk et al. 2022; Kuhlmann et al. 2007). Cleptoparasitic species and their hosts are in a continuous arms race (Castillo et al. 2022). This often forces the brood parasites to only parasite one or a few related species. The idea that host preferences or host shifts may lead to speciation is more than 150 years old but has been strengthened recently (Forbes et al. 2017).

Interestingly, all *N. flavoguttata* specimens sampled in association with *A. subopaca* are placed in clade 1 in the COI tree, while *N. flavoguttata* specimens sampled with *A. semilaevis* are placed in both clade 1 and clade 2. Whether the observed genetic variation and grouping are linked with host species preference is difficult to say based on my data, and additional data is required. It would be interesting to see where *N. flavoguttata* specimens collected in

association with the third common *Andrena* host (*A. minutula*) would end up in relation to the other specimens. If they are placed in clade 3, that will strengthen the hypotheses that the genetic differences are related to host preferences. Five of the ten specimens placed in clade 3 have been collected from localities where *A. minutula* has been recorded. The remaining five have been collected from localities near known records of *A. minutula*. Although this is interesting, it does not confirm anything since the other host species could also be present at those localities.

At each locality, all specimens were found in a relatively small area. From one of the localities, they were collected in an area only a few square meters large, showing that they can be extremely local. In one of the localities, both *Andrena semilaevis* and *A. subopaca* were found, which makes it difficult to conclude which species the *N. flavoguttata* specimens at that locality are associated with. This could also be the case at other localities, as the other host species may have gone undetected. Even if most, or all, of the *Andrena* specimens were one species at a locality, that does not necessarily reflect the actual composition of the *Andrena* species or the host association at that locality. Several factors may cause one or the other species to be more abundant during sampling. The daily foraging pattern of bees may change depending on available flower resources (Vaudo et al. 2014). Although both *Andrena semilaevis* and *A. subopaca* are generalist feeders, they may have foraging preferences making them more numerous on different days or times of the day at a given locality. In addition, a mass hatching of one species at the time of sampling could cause an overrepresentation of that species. Rearing *N. flavoguttata* from cells with known hosts or catching females interested in nest cells would be the safest way to assess host association. However, finding nest cells of *Andrena* (*Micrandrena*) is not easy, as it can be hidden among grasses.

Almost all *N. flavoguttata* specimens caught at localities where both *N. flavoguttata* and *Andrena* were collected were males. *Nomada* males are known to “swarm” as a courtship behavior (Schindler et al. 2018). Therefore, they can quite easily be caught in numbers if they are present in a locality. The females do not aggregate the same way as males, making it more difficult to catch many of them at a locality. The males fly for a relatively short period, while the females continue flying for some time longer. A way to possibly catch more females would be to revisit the localities where I found many males, as those localities

should also have a decent population of females. By revisiting localities, it would also be possible to resample *Andrena* specimens to check if the species composition is the same, thus strengthening the information about host species.

Despite being relatively easy to find, sampling enough *N. flavoguttata* specimens in one season proved to be difficult. The flight period is relatively short, making the sampling season very short. In addition, sampling is very weather dependent since they are primarily active during sunny days. Although common in many habitats, several sampling trips yielded no *N. flavoguttata* specimens.

4.4 Morphology

Unfortunately, the presumably non-destructive DNA extraction using whole specimens was quite destructive. The coloration became very difficult to interpret afterward. Also, most of the pilosity was either lost or laid flat against the body of the specimens. This made some of the specimens difficult to include in the morphological analysis.

Nomada flavoguttata is known to vary in coloration, especially between the first and second generations (Falk 2015; Smit 2018). It was difficult to compare the females with only one female from clade 3 and none from clade 2. In addition, the female from clade 3 is almost certainly a second-generation female, as it was caught 10th of August 2012. The difference between the coloration of clypeus in the female from clade 3 compared to the females from clade 1 may therefore be generation differences. However, if additional first-generation specimens assigned to clade 3 would show the same pattern, it would be worth looking closer at. In general, the males from clade 2 had a lighter clypeus and a lighter posterior margin of labrum than those from the two other clades. However, there were too many exceptions for it to be good characters to separate the clades. In other *Nomada* species, the coloration varies as well. Therefore, coloration is not always reliable for discriminating between *Nomada* species (Falk et al. 2022).

The arrangement of spines on the hind tibia is often used to identify *Nomada* species. Falk (2015) states that the spines on the hind tibia of *N. flavoguttata* are very variable. This is also my perception; their number, length, and thickness varied (Figs. 19C, D). Thus, the arrangement of spines on the hind tibia is not a very good feature for identifying *N. flavoguttata*.

Generally, the specimens from clade 1 were larger than those from the two other clades (Fig. 15). In some cases, different body characters do not vary proportionally to the overall size. This is known as allometry (Nakagawa et al. 2017; Pélabon et al. 2014). The two characters that made up the best ratio were GMW and OP (Fig. 16). The two characters had a very different correlation to isosize. GMW had a moderate to strong correlation with isosize, while OP had a low correlation with isosize. Therefore, the size will strongly affect this ratio, and the clear separation of clade 1 from the other two clades using this ratio may very well be due to size differences.

Insufficient sampling sizes may lead to errors, especially in estimates of mean shapes (Cardini & Elton 2007). I had very little material available from clades 2 and 3. Ideally, I should have had 30+ specimens from each clade to avoid statistical errors which may occur when working with a small number of samples. Especially since *N. flavoguttata* seems to be quite variable in many characters, a larger dataset would be required to say anything with confidence.

Many of the characters investigated were relatively variable. Even though some characters showed tendencies to be related to a clade, there were always exceptions. Therefore, none of the characters investigated proved to be good at confidently separating the clades.

4.5 *Wolbachia*

Several *Nomada* species are known to be hosts of *Wolbachia* (Gerth et al. 2013). To my knowledge, this is the first time *Wolbachia* has been recorded from *N. flavoguttata*. In most cases, the intraspecific infection frequency of *Wolbachia* is either very low (<10%) or very high (>90%) (Hilgenboecker et al. 2008). The present study detected a very low infection frequency of 2.6% in *N. flavoguttata*. The infection rates of *Wolbachia* may vary among populations of the same species (Gora et al. 2020). The *N. flavoguttata* specimen that tested positive for *Wolbachia* was from a locality where five other specimens were also tested. This suggests that even infected populations of *N. flavoguttata* may have a low infection frequency.

The *Wolbachia* sequences isolated from *N. flavoguttata* and *N. cf. panzeri* were identical. They were also identical to a *Wolbachia* sequence isolated from *N. flava* by Gerth et al. (2013). The presence of identical *Wolbachia* strains among different species may suggest horizontal transmission. Physical contact is a prerequisite for horizontal transmission of

Wolbachia (Gerth et al. 2013). It is suggested that shared food sources are ways for horizontal transmission of *Wolbachia* (Ahmed et al. 2016). *Nomada flavoguttata* is a generalist feeder known to visit flowers from many different plant families (Falk 2015). In addition, the host species of *N. flavoguttata* are common in a wide variety of habitat types. This means there is a high possibility for physical interactions between *N. flavoguttata* and many other bee/insect species. Similar *Wolbachia* strains can also result from recombination of different *Wolbachia* strains. In some cases, recombination is very frequent. Thus, interpreting the origin of *Wolbachia* strains may be difficult (Jiggins et al. 2001; Reuter & Keller 2003).

The low infection frequency and the fact that the infected specimen is placed alongside the uninfected specimens suggest that *Wolbachia* has not affected the genealogical relationships indicated by mtDNA in *N. flavoguttata.*

5 Conclusion

Only the mitochondrial marker COI recognized the three proposed clades as monophyletic. The molecular species delimiting methods had conflicting results, delimiting the three clades to one or three species. Some morphological characteristics showed tendencies to be related to certain clades. However, there were always exceptions. Thus, I did not find any good morphological characters for separating the clades. Based on the data in this study, I can neither confirm nor refute that the genetic variation is linked to host preferences. The low infection frequency of 2.6% and the fact that the infected specimen is placed alongside the uninfected specimens suggest that *Wolbachia* has not affected the genealogical relationships indicated by mtDNA in *N. flavoguttata.* I have no strong evidence for recommending any changes in the taxonomy of *N. flavoguttata*, and it should still be regarded as one species.

As my dataset was relatively small, it was difficult to draw conclusions. Further studies should be conducted, focusing on sampling more females of *N. flavoguttata*, and sampling from localities where *A. minutula* is present.

6 Acknowledgement

I want to thank my supervisors, Frode Ødegaard and Torbjørn Ekrem, for valuable feedback while writing this thesis. I also want to thank Mohsen Falahati for all the help he provided in the molecular labs.

References

Ahmed, M. Z., Araujo-Jnr, E. V., Welch, J. J., & Kawahara, A. Y. (2015). *Wolbachia* in butterflies and moths: geographic structure in infection frequency. *Front Zool, 12*(1), 16-16. doi:10.1186/s12983-015-0107-z

Ahmed, M. Z., Breinholt, J. W., & Kawahara, A. Y. (2016). Evidence for common horizontal transmission of *Wolbachia* among butterflies and moths. *BMC Evol Biol, 16*(1), 118-118. doi:10.1186/s12862-016-0660-x

Ajmal Ali, M., Gyulai, G., Higvégi, N., Kerti, B., Al Hemaid, F. M. A., Pandey, A. K., & Lee, J. (2014). The changing epitome of species identification – DNA barcoding. *Saudi Journal of Biological Sciences, 21*(3), 204-231. doi:10.1016/j.sjbs.2014.03.003

Alexander, B. A. (1994). Species-groups and cladistic analysis of the cleptoparastic [sic] bee genus *Nomada* (Hymenoptera: Apoidea). *The University of Kansas science bulletin, 55*, 175-236. doi:10.5962/bhl.part.776

Allio, R., Donega, S., Galtier, N., & Nabholz, B. (2017). Large variation in the ratio of mitochondrial to nuclear mutation rate across animals: Implications for genetic diversity and the use of mitochondrial DNA as a molecular marker. *Mol Biol Evol, 34*(11), 2762-2772. doi:10.1093/molbev/msx197

Artskart. (2023). Retrieved from https://artskart.artsdatabanken.no/.

Arturo Roig, A. (2009). A revision of the bee genus *Nomada* in Argentina (Hymenoptera, Apidae, Nomadinae). *Revista del Museo Argentino de Ciencias Naturales (1999), 11*(2), 221-241.

Ballard, J. W. O., Hatzidakis, J., Karr, T. L., & Kreitman, M. (1996). Reduced Variation in *Drosophila simulans* Mitochondrial DNA. *Genetics, 144*(4), 1519-1528. doi:10.1093/genetics/144.4.1519

Baur, H., Kranz-Baltensperger, Y., Cruaud, A., Rasplus, J. Y., Timokhov, A. V., & Gokhman, V. E. (2014). Morphometric analysis and taxonomic revision of *Anisopteromalus* Ruschka (Hymenoptera: Chalcidoidea: Pteromalidae) – an integrative approach. *Syst Entomol, 39*(4), 691-709. doi:10.1111/syen.12081

Baur, H., & Leuenberger, C. (2011). Analysis of Ratios in Multivariate Morphometry. *Syst Biol, 60*(6), 813-825. doi:10.1093/sysbio/syr061

Baur, H., & Leuenberger, C. (2020). Multivariate Ratio Analysis (MRA): R-scripts and tutorials for calculating Shape PCA, Ratio Spectra and LDA Ratio Extractor (1.05). Zenodo. doi:10.5281/zenodo.3890195.

Beheregaray, L. B., & Caccone, A. (2007). Cryptic biodiversity in a changing world. *J Biol, 6*(4), 9-9. doi:10.1186/jbiol60

Biesmeijer, J. C., Robertson, S. P., Reemer, M., Ohlemuller, R., Edwards, M., Peeters, T. M. J., . . . Kunin, W. E. (2006). Parallel Declines in Pollinators and Insect-Pollinated Plants in Britain and the Netherlands. *Science, 313*(5785), 351-354. doi:10.1126/science.1127863

Bogusch, P., Kratochvíl, L., & Straka, J. (2006). Generalist cuckoo bees (Hymenoptera: Apoidea: *Sphecodes*) are species-specialist at the individual level. *Behavioral ecology and sociobiology, 60*(3), 422-429. doi:10.1007/s00265-006-0182-4

Bonfield, J. K., Smith, K. F., & Staden, R. (1995). A new DNA sequence assembly program. *Nucleic Acids Res, 23*(24), 4992-4999. doi:10.1093/nar/23.24.4992

Bouckaert, R., Vaughan, T. G., Barido-Sottani, J., Duchêne, S., Fourment, M., Gavryushkina, A., . . . Drummond, A. J. (2019). BEAST 2.5: An advanced software platform for

Bayesian evolutionary analysis. *PLoS Comput Biol, 15*(4), e1006650-e1006650. doi:10.1371/journal.pcbi.1006650

Bouckaert, R., & Xie, D. (2017). BEAST2-Dev/substmodels: Standard Nucleotide Substitution Models v1.0.1 (v1.0.1). Zenodo. doi:10.5281/zenodo.995740

Burbrink, F. T., & Ruane, S. (2021). Contemporary Philosophy and Methods for Studying Speciation and Delimiting Species. *Copeia, 109*(3), 874-894. doi:10.1643/h2020073

Cameron, S. A., Lozier, J. D., Strange, J. P., Koch, J. B., Cordes, N., Solter, L. F., & Griswold, T. L. (2011). Patterns of widespread decline in North American bumble bees. *Proc Natl Acad Sci U S A, 108*(2), 662-667. doi:10.1073/pnas.1014743108

Cardinal, S., Straka, J., & Danforth, B. N. (2010). Comprehensive phylogeny of apid bees reveals the evolutionary origins and antiquity of cleptoparasitism. *Proc Natl Acad Sci U S A, 107*(37), 16207-16211. doi:10.1073/pnas.1006299107

Cardini, A., & Elton, S. (2007). Sample size and sampling error in geometric morphometric studies of size and shape. *Zoomorphology, 126*(2), 121-134. doi:10.1007/s00435-007-0036-2

Carstens, B. C., Pelletier, T. A., Reid, N. M., & Satler, J. D. (2013). How to fail at species delimitation. *Mol Ecol, 22*(17), 4369-4383. doi:10.1111/mec.12413

Castillo, R., Wurdack, M., Pauli, T., Keller, A., Feldhaar, H., Polidori, C., . . . Schmitt, T. (2022). Evidence for a chemical arms race between cuckoo wasps of the genus *Hedychrum* and their distantly related host apoid wasps. *BMC Ecol Evol, 22*(1), 138-138. doi:10.1186/s12862-022-02093-8

Clement, M., Posada, D., & Crandall, K. A. (2000). TCS: a computer program to estimate gene genealogies. *Mol Ecol, 9*(10), 1657-1659. doi:10.1046/j.1365-294x.2000.01020.x

Clement, M., Snell, Q., Walke, P., Posada, D., & Crandall, K. (2002). TCS: Estimating gene genealogies. *Parallel and Distributed Processing Symposium, International Proceedings*, 2, 184. doi:10.1109/IPDPS.2002.1016585

Coelho, P., Kaliontzopoulou, A., Rasko, M., & van der Meijden, A. (2017). A 'striking' relationship: scorpion defensive behaviour and its relation to morphology and performance. *Functional ecology, 31*(7), 1390-1404. doi:10.1111/1365-2435.12855

Costa, M. A., Del Lama, M. A., Melo, G. A. R., & Sheppard, W. S. (2003). Molecular phylogeny of the stingless bees (Apidae, Apinae, Meliponini) inferred from mitochondrial 16S rDNA sequences. *Apidologie, 34*(1), 73-84. doi:10.1051/apido:2002051

Csősz, S., Seifert, B., Mikó, I., Boudinot, B. E., Borowiec, M. L., Fisher, B. L., . . . Elek, Z. (2021). Insect morphometry is reproducible under average investigation standards. *Ecol Evol, 11*(1), 547-559. doi:10.1002/ece3.7075

Danforth, B. N., Minckley, R. L., Neff, J. L., & Fawcett, F. (2019). *The Solitary Bees: Biology, Evolution, Conservation*. Princeton: Princeton University Press.

Dayrat, B. (2005). Towards integrative taxonomy. *Biological journal of the Linnean Society, 85*(3), 407-415. doi:10.1111/j.1095-8312.2005.00503.x

De Queiroz, K. (2007). Species Concepts and Species Delimitation. *Syst Biol, 56*(6), 879-886. doi:10.1080/10635150701701083

Dellicour, S., & Flot, J.-F. (2015). Delimiting Species-Poor Data Sets using Single Molecular Markers: A Study of Barcode Gaps, Haplowebs and GMYC. *Syst Biol, 64*(6), 900-908. doi:10.1093/sysbio/syu130

Dincă, V., Lukhtanov, V. A., Talavera, G., & Vila, R. (2011). Unexpected layers of cryptic diversity in wood white Leptidea butterflies. *Nat Commun, 2*(1), 324-324. doi:10.1038/ncomms1329

Droege, S., Rightmyer, M. G., Sheffield, C. S., & Brady, S. G. (2010). New synonymies in the bee genus *Nomada* from North America (Hymenoptera: Apidae). *Zootaxa, 2661*(1), 1-32. doi:10.11646/zootaxa.2661.1.1

Drossart, M., & Gérard, M. (2020). Beyond the Decline of Wild Bees: Optimizing Conservation Measures and Bringing Together the Actors. *Insects (Basel, Switzerland), 11*(9), 649. doi:10.3390/insects11090649

Duron, O., Bouchon, D., Boutin, S., Bellamy, L., Zhou, L., Engelstädter, J., & Hurst, G. D. (2008). The diversity of reproductive parasites among arthropods: *Wolbachia* do not walk alone. *BMC Biol, 6*(1), 27-27. doi:10.1186/1741-7007-6-27

Ezard, T., Fujisawa, T., & Barraclough, T. G. (2009). SPLITS: SPecies' LImits by Threshold Statistics. R package version 1.0-20.

Falk, S. (2015). *Field Guide to the Bees of Great Britain and Ireland*. London: Bloomsbury Publishing Plc.

Falk, S., Johansson, N., & Paxton, R. J. (2022). DNA and morphological characterisation of the Bilberry Mining Bee *Nomada glabella* sensu Stöckhert nec Thomson in Britain with discussion of the remaining variation within *N. panzeri*. *British Journal of Entomology & Natural History, 35*(2), 91-111.

Fišer Pečnikar, Ž., & Buzan, E. V. (2014). 20 years since the introduction of DNA barcoding: from theory to application. *Journal of applied genetics, 55*(1), 43-52. doi:10.1007/s13353-013-0180-y

Flouri, T., Jiao, X., Rannala, B., & Yang, Z. (2018). Species Tree Inference with BPP Using Genomic Sequences and the Multispecies Coalescent. *Mol Biol Evol, 35*(10), 2585-2593. doi:10.1093/molbev/msy147

Flouri, T., Rannala, B., & Yang, Z. (2020). A Tutorial on the Use of BPP for Species Tree Estimation and Species Delimitation. In C. Scornavacca, F. Delsuc, & N. Galtier (Eds.), *Phylogenetics in the Genomic Era* (pp. Chapter 5.6 pp. 1-16).

Folmer, O., Black, M., Hoeh, W., Lutz, R., & Vrijenhoek, R. (1994). DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Mol Mar Biol Biotechnol, 3*(5), 294-299.

Forbes, A. A., Devine, S. N., Hippee, A. C., Tvedte, E. S., Ward, A. K. G., Widmayer, H. A., & Wilson, C. J. (2017). Revisiting the particular role of host shifts in initiating insect speciation. *Evolution, 71*(5), 1126-1137. doi:10.1111/evo.13164

Garibaldi, L. A., Steffan-Dewenter, I., Bartomeus, I., Benjamin, F., Boreux, V., Cariveau, D., . . . Afik, O. (2013). Wild Pollinators Enhance Fruit Set of Crops Regardless of Honey Bee Abundance. *Science, 339*(6127), 1608-1611. doi:10.1126/science.1230200

Gerth, M., GeißLer, A., & Bleidorn, C. (2011). *Wolbachia* infections in bees (Anthophila) and possible implications for DNA barcoding. *Systematics and biodiversity, 9*(4), 319-327. doi:10.1080/14772000.2011.627953

Gerth, M., Röthe, J., & Bleidorn, C. (2013). Tracing horizontal *Wolbachia* movements among bees (Anthophila): a combined approach using multilocus sequence typing data and host phylogeny. *Mol Ecol, 22*(24), 6149-6162. doi:10.1111/mec.12549

Gonzalez, V. H., Griswold, T., & Engel, M. S. (2013). Obtaining a better taxonomic understanding of native bees: where do we start?: Revising species in systematic melittology. *Systematic entomology, 38*(4), 645-653. doi:10.1111/syen.12029

Gora, N. V., Serga, S. V., Maistrenko, O. M., Ślęzak-Parnikoza, A., Parnikoza, I. Y., Tarasiuk, A. N., . . . Kozeretska, I. A. (2020). Climate Factors and *Wolbachia* Infection Frequencies

in Natural Populations of *Drosophila melanogaster*. *Cytology and genetics, 54*(3), 189-198. doi:10.3103/S0095452720030044

Habermannová, J., Bogusch, P., & Straka, J. (2013). Flexible host choice and common host switches in the evolution of generalist and specialist cuckoo bees (Anthophila: *Sphecodes*). *PLoS One, 8*(5), e64537-e64537. doi:10.1371/journal.pone.0064537

Hall, B. G. (2013). Building phylogenetic trees from molecular data with MEGA. *Mol Biol Evol, 30*(5), 1229-1235. doi:10.1093/molbev/mst012

Hebert, P. D. N., Cywinska, A., Ball, S. L., & deWaard, J. R. (2003). Biological identifications through DNA barcodes. *Proc Biol Sci, 270*(1512), 313-321. doi:10.1098/rspb.2002.2218

Hilgenboecker, K., Hammerstein, P., Schlattmann, P., Telschow, A., & Werren, J. H. (2008). How many species are infected with *Wolbachia*? - a statistical analysis of current data. *FEMS Microbiol Lett, 281*(2), 215-220. doi:10.1111/j.1574-6968.2008.01110.x

Hurst, G. D. D., & Jiggins, F. M. (2005). Problems with mitochondrial DNA as a marker in population, phylogeographic and phylogenetic studies: the effects of inherited symbionts. *PROC R SOC B, 272*(1572), 1525-1534. doi:10.1098/rspb.2005.3056

Janzen, D. H., Hajibabaei, M., Burns, J. M., Hallwachs, W., Remigio, E., & Hebert, P. D. N. (2005). Wedding biodiversity inventory of a large and complex Lepidoptera fauna with DNA barcoding. *Philos Trans R Soc Lond B Biol Sci, 360*(1462), 1835-1845. doi:10.1098/rstb.2005.1715

Ji, Y. J., Zhang, D. X., & He, L. J. (2003). Evolutionary conservation and versatility of a new set of primers for amplifying the ribosomal internal transcribed spacer regions in insects and other invertebrates. *Molecular ecology notes, 3*(4), 581-585. doi:10.1046/j.1471-8286.2003.00519.x

Jiggins, F. M. (2003). Male-killing *Wolbachia* and mitochondrial DNA: selective sweeps, hybrid introgression and parasite population dynamics. *Genetics, 164*(1), 5-12. doi:10.1093/genetics/164.1.5

Jiggins, F. M., Schulenburg, J. H. G. v. d., Hurst, G. D. D., & Majerus, M. E. N. (2001). Recombination confounds interpretations of *Wolbachia* evolution. *Proc Biol Sci, 268*(1474), 1423-1427. doi:10.1098/rspb.2001.1656

Kapli, P., Lutteropp, S., Zhang, J., Kobert, K., Pavlidis, P., Stamatakis, A., & Flouri, T. (2017). Multi-rate Poisson tree processes for single-locus species delimitation under maximum likelihood and Markov chain Monte Carlo. *Bioinformatics, 33*(11), 1630-1638. doi:10.1093/bioinformatics/btx025

Kimura, M. (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J Mol Evol, 16*(2), 111-120. doi:10.1007/BF01731581

Kimura, M. (1981). Estimation of evolutionary distances between homologous nucleotide sequences. *Proc Natl Acad Sci U S A, 78*(1), 454-458. doi:10.1073/pnas.78.1.454

Klein, A.-M., Vaissière, B. E., Cane, J. H., Steffan-Dewenter, I., Cunningham, S. A., Kremen, C., & Tscharntke, T. (2007). Importance of pollinators in changing landscapes for world crops. *PROC R SOC B, 274*(1608), 303-313. doi:10.1098/rspb.2006.3721

Kuhlmann, M., Else, G. R., Dawson, A., & Quicke, D. L. J. (2007). Molecular, biogeographical and phenological evidence for the existence of three western European sibling species in the *Colletes succinctus* group (Hymenoptera: Apidae). *Organisms diversity & evolution, 7*(2), 155-165. doi:10.1016/j.ode.2006.04.001

Leigh, J. W., Bryant, D., & Nakagawa, S. (2015). popart: full-feature software for haplotype network construction. *Methods in ecology and evolution, 6*(9), 1110-1116. doi:10.1111/2041-210X.12410

Lim, K., Lee, S., Orr, M., & Lee, S. (2022). Harrison's rule corroborated for the body size of cleptoparasitic cuckoo bees (Hymenoptera: Apidae: Nomadinae) and their hosts. *Scientific reports, 12*(1), 10984-10984. doi:10.1038/s41598-022-14938-9

Litman, J. R. (2019). Under the radar: detection avoidance in brood parasitic bees. *Philos Trans R Soc Lond B Biol Sci, 374*(1769), 20180196-20180196. doi:10.1098/rstb.2018.0196

Litman, J. R., Fateryga, A. V., Griswold, T. L., Aubert, M., Proshchalykin, M. Y., Le Divelec, R., . . . Praz, C. J. (2021). Paraphyly and low levels of genetic divergence in morphologically distinct taxa: revision of the *Pseudoanthidium scapulare* complex of carder bees (Apoidea: Megachilidae: Anthidiini). *Zoological journal of the Linnean Society*. doi:10.1093/zoolinnean/zlab062

Litman, J. R., Praz, C. J., Danforth, B. N., Griswold, T. L., & Cardinal, S. (2013). Origins, evolution, and diversification of cleptoparasitic lineages in long-tongued bees. *Evolution, 67*(10), 2982-2998. doi:10.1111/evo.12161

Losey, J. E., & Vaughan, M. (2006). The Economic Value of Ecological Services Provided by Insects. *BioScience, 56*(4), 311-323. doi:10.1641/0006-3568(2006)56[311:TEVOES]2.0.CO2

Luo, A., Ling, C., Ho, S. Y. W., & Zhu, C.-D. (2018). Comparison of Methods for Molecular Species Delimitation Across a Range of Speciation Scenarios. *Syst Biol, 67*(5), 830-846. doi:10.1093/sysbio/syy011

Michener, C. D. (2007). *The Bees of the World*. Baltimore: Baltimore: Johns Hopkins University Press.

Mitai, K., & Tadauchi, O. (2007). Taxonomic Study of the Japanese Species of the *Nomada ruficornis* Species Group (Hymenoptera, Apidae) with Remarks on Japanese Fauna of the Genus *Nomada*. *Esakia : occasional papers of the Hikosan Biological Laboratory in Entomology, 47*(47), 25-167. doi:10.5109/8324

Moraes, S. S., Montebello, Y., Stanton, M. A., Yamaguchi, L. F., Kato, M. J., & Freitas, A. V. L. (2021). Description of three new species of Geometridae (Lepidoptera) using species delimitation in an integrative taxonomy approach for a cryptic species complex. *PeerJ (San Francisco, CA), 9*, e11304-e11304. doi:10.7717/peerj.11304

Nakagawa, S., Kar, F., O'Dea, R. E., Pick, J. L., & Lagisz, M. (2017). Divide and conquer? Size adjustment with allometry and intermediate outcomes. *BMC Biol, 15*(1), 107-107. doi:10.1186/s12915-017-0448-5

Odanaka, K. A., Branstetter, M. G., Tobin, K. B., & Rehan, S. M. (2022). Phylogenomics and historical biogeography of the cleptoparasitic bee genus *Nomada* (Hymenoptera: Apidae) using ultraconserved elements. *Mol Phylogenet Evol, 170*, 107453-107453. doi:10.1016/j.ympev.2022.107453

Orr, M. C., Hughes, A. C., Chesters, D., Pickering, J., Zhu, C.-D., & Ascher, J. S. (2021). Global Patterns and Drivers of Bee Distribution. *Curr Biol, 31*(3), 451-458.e454. doi:10.1016/j.cub.2020.10.053

Packer, L., Gibbs, J., Sheffield, C., & Hanner, R. (2009). DNA barcoding and the mediocrity of morphology. *Mol Ecol Resour, 9*(s1), 42-50. doi:10.1111/j.1755-0998.2009.02631.x

Pauly, A., Devalez, J., Sonet, G., Nagy, Z. T., & Boevé, J.-L. (2015). DNA barcoding and male genital morphology reveal five new cryptic species in the West Palearctic bee

Seladonia smaragdula (Vachal, 1895) (Hymenoptera: Apoidea: Halictidae). *Zootaxa, 4034*(2), 257-290. doi:10.11646/zootaxa.4034.2.2
Pauw, A. (2007). Collapse of a Pollination Web in Small Conservation Areas. *Ecology, 88*(7), 1759-1769. doi:10.1890/06-1383.1
Pélabon, C., Firmat, C., Bolstad, G. H., Voje, K. L., Houle, D., Cassara, J., . . . Hansen, T. F. (2014). Evolution of morphological allometry. *Ann. N.Y. Acad. Sci, 1320*(1), 58-75. doi:10.1111/nyas.12470
Pons, J., Barraclough, T. G., Gomez-Zurita, J., Cardoso, A., Duran, D. P., Hazell, S., . . . Hedin, M. (2006). Sequence-Based Species Delimitation for the DNA Taxonomy of Undescribed Insects. *Syst Biol, 55*(4), 595-609. doi:10.1080/10635150600852011
Posada, D. (2003). Using MODELTEST and PAUP* to select a model of nucleotide substitution. *Current Protocols in Bioinformatics, Chapter 6*. doi:10.1002/0471250953.bi0605s00
Posada, D. (2008). jModelTest: Phylogenetic Model Averaging. *Mol Biol Evol, 25*(7), 1253-1256. doi:10.1093/molbev/msn083
Posit team. (2023). RStudio: Integrated Development Environment for R. Posit Software, PBC, Boston, MA. URL http://www.posit.co/.
Potts, S. G., Biesmeijer, J. C., Kremen, C., Neumann, P., Schweiger, O., & Kunin, W. E. (2010). Global pollinator declines: trends, impacts and drivers. *Trends Ecol Evol, 25*(6), 345-353. doi:10.1016/j.tree.2010.01.007
Praz, C., Genoud, D., Vaucher, K., Bénon, D., Monks, J., & Wood, T. J. (2022). Unexpected levels of cryptic diversity in European bees of the genus *Andrena* subgenus Taeniandrena : implications for conservation. *Journal of Hymenoptera research*(1), 375. doi:10.3897/jhr.91.82761
Proshchalykin, M., & Lelej, A. (2010). Review of the *Nomada roberjeotiana* species-group (Hymenoptera: Apidae) of Russia, with description of new species. *Zootaxa, 2335*(1), 1-15. doi:10.11646/zootaxa.2335.1.1
Puillandre, N., Brouillet, S., & Achaz, G. (2021). ASAP: assemble species by automatic partitioning. *Mol Ecol Resour, 21*(2), 609-620. doi:10.1111/1755-0998.13281
Queffelec, J., Postma, A., Allison, J. D., & Slippers, B. (2022). Remnants of horizontal transfers of *Wolbachia* genes in a *Wolbachia*-free woodwasp. *BMC Ecol Evol, 22*(1), 36-36. doi:10.1186/s12862-022-01995-x
R Core Team. (2023). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.
Rambaut, A., Drummond, A. J., Xie, D., Baele, G., & Suchard, M. A. (2018). Posterior Summarization in Bayesian Phylogenetics Using Tracer 1.7. *Syst Biol, 67*(5), 901-904. doi:10.1093/sysbio/syy032
Rannala, B., & Yang, Z. (2003). Bayes Estimation of Species Divergence Times and Ancestral Population Sizes Using DNA Sequences From Multiple Loci. *Genetics, 164*(4), 1645-1656. doi:10.1093/genetics/164.4.1645
Rannala, B., & Yang, Z. (2013). Improved reversible jump algorithms for Bayesian species delimitation. *Genetics, 194*(1), 245-253. doi:10.1534/genetics.112.149039
Rannala, B., & Yang, Z. (2017). Efficient Bayesian Species Tree Inference under the Multispecies Coalescent. *Syst Biol, 66*(5), 823-842. doi:10.1093/sysbio/syw119
Reuter, M., & Keller, L. (2003). High levels of multiple *Wolbachia* infection and recombination in the ant Formica exsecta. *Mol Biol Evol, 20*(5), 748-753. doi:10.1093/molbev/msg082

Sann, M., Niehuis, O., Peters, R. S., Mayer, C., Kozlov, A., Podsiadlowski, L., . . . Ohl, M. (2018). Phylogenomic analysis of Apoidea sheds new light on the sister group of bees. *BMC Evol Biol, 18*(1), 71-71. doi:10.1186/s12862-018-1155-8

Sazama, E. J., Bosch, M. J., Shouldis, C. S., Ouellette, S. P., & Wesner, J. S. (2017). Incidence of *Wolbachia* in aquatic insects. *Ecol Evol, 7*(4), 1165-1169. doi:10.1002/ece3.2742

Schindler, M., Hofmann, M. M., Wittmann, D., & Renner, S. S. (2018). Courtship behaviour in the genus *Nomada* – antennal grabbing and possible transfer of male secretions. *Journal of Hymenoptera research, 65*(65), 47-59. doi:10.3897/jhr.65.24947

Schlick-Steiner, B. C., Seifert, B., Stauffer, C., Christian, E., Crozier, R. H., & Steiner, F. M. (2007). Without morphology, cryptic species stay in taxonomic crypsis following discovery. *Trends Ecol Evol, 22*(8), 391-392. doi:10.1016/j.tree.2007.05.004

Schlick-Steiner, B. C., Steiner, F. M., Seifert, B., Stauffer, C., Christian, E., & Crozier, R. H. (2010). Integrative Taxonomy: A Multisource Approach to Exploring Biodiversity. *Annu Rev Entomol, 55*(1), 421-438. doi:10.1146/annurev-ento-112408-085432

Schwarz, G. (1978). Estimating the Dimension of a Model. *The Annals of statistics, 6*(2), 461-464. doi:10.1214/aos/1176344136

Sheffield, C. S., Pindar, A., Packer, L., & Kevan, P. G. (2013). The potential of cleptoparasitic bees as indicator taxa for assessing bee communities. *Apidologie, 44*(5), 501-510. doi:10.1007/s13592-013-0200-2

Smit, J. (2018). Identification key to the European species of the bee genus *Nomada* Scopoli, 1770 (Hymenoptera: Apidae), Including 23 new species. *Entomofauna, M3*, 1-253.

Soon, V., Castillo-Cajas, R. F., Johansson, N., Paukkunen, J., Rosa, P., Ødegaard, F., . . . Niehuis, O. (2021). Cuticular Hydrocarbon Profile Analyses Help Clarify the Species Identity of Dry-Mounted Cuckoo Wasps (Hymenoptera: Chrysididae), Including Type Material, and Reveal Evidence for a Cryptic Species. *Insect systematics and diversity, 5*(1). doi:10.1093/isd/ixab002

Staden, R. (1996). Staden sequence analysis package. *Mol Biotechnol, 5*(3), 233-241. doi:10.1007/BF02900361

Stenløkk, J. A. (2011). On the genus *Nomada* Scopoli, 1770 (Hymenoptera, Apoidae) in Norway. *Norwegian journal of entomology (Oslo, Norway : 1999), 58*(1), 44-72.

Talavera, G., Dincă, V., Vila, R., & Paradis, E. (2013). Factors affecting species delimitations with the GMYC model: insights from a butterfly survey. *Methods in ecology and evolution, 4*(12), 1101-1110. doi:10.1111/2041-210X.12107

Tamura, K. (1992). Estimation of the number of nucleotide substitutions when there are strong transition-transversion and G + C-content biases. *Molecular Biology and Evolution, 9*(4), 678-687. doi:10.1093/oxfordjournals.molbev.a040752

Tamura, K., & Nei, M. (1993). Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular Biology and Evolution, 10*(3), 512–526. doi:10.1093/oxfordjournals.molbev.a040023

Tamura, K., Stecher, G., & Kumar, S. (2021). MEGA11: Molecular Evolutionary Genetics Analysis Version 11. *Mol Biol Evol, 38*(7), 3022-3027. doi:10.1093/molbev/msab120

Tengo, J., & Bergstrom, G. (1977). Cleptoparasitism and Odor Mimetism in Bees: Do *Nomada* Males Imitate the Odor of *Andrena* Females? *Science, 196*(4294), 1117-1119. doi:10.1126/science.196.4294.1117

Thorogood, R., Spottiswoode, C. N., Portugal, S. J., & Gloag, R. (2019). The coevolutionary biology of brood parasitism: a call for integration. *Philos Trans R Soc Lond B Biol Sci, 374*(1769), 20180190-20180190. doi:10.1098/rstb.2018.0190

Trewick, S. A. (2008). DNA Barcoding is not enough: mismatch of taxonomy and genealogy in New Zealand grasshoppers (Orthoptera: Acrididae). *Cladistics, 24*(2), 240-254. doi:10.1111/j.1096-0031.2007.00174.x

Vaudo, A. D., Patch, H. M., Mortensen, D. A., Grozinger, C. M., & Tooker, J. F. (2014). Bumble bees exhibit daily behavioral patterns in pollen foraging. *Arthropod-plant interactions, 8*(4), 273-283. doi:10.1007/s11829-014-9312-5

Weinert, L. A., Araujo-Jnr, E. V., Ahmed, M. Z., & Welch, J. J. (2015). The incidence of bacterial endosymbionts in terrestrial arthropods. *Proc Biol Sci, 282*(1807), 20150249-20150249. doi:10.1098/rspb.2015.0249

Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York. ISBN 978-3-319-24277-4, https://ggplot2.tidyverse.org.

Wood, T. J. (2021). Revision of the *Andrena* (Hymenoptera: Andrenidae) fauna of Bulgaria and North Macedonia with description of three new species. *Belgian Journal of Entomology, 117*, 1-39.

Yang, Z. (2015). The BPP program for species tree estimation and species delimitation. *Acta Zoologica Sinica, 61*(5), 854-865. doi:10.1093/czoolo/61.5.854

Yang, Z., & Rannala, B. (2010). Bayesian species delimitation using multilocus sequence data. *Proc Natl Acad Sci U S A, 107*(20), 9264-9269. doi:10.1073/pnas.0913022107

Yang, Z., & Rannala, B. (2014). Unguided species delimitation using DNA sequence data from multiple Loci. *Mol Biol Evol, 31*(12), 3125-3135. doi:10.1093/molbev/msu279

Zachos, F. E. (2016). *Species Concepts in Biology : Historical Development, Theoretical Foundations and Practical Relevance*(1st 2016. ed.).

Zachos, F. E. (2018). (New) Species concepts, species delimitation and the inherent limitations of taxonomy. *J Genet, 97*(4), 811-815. doi:10.1007/s12041-018-0965-1

Zamani, A., Dal Pos, D., Fric, Z. F., Orfinger, A. B., Scherz, M. D., Bartoňová, A. S., & Gante, H. F. (2022). The future of zoological taxonomy is integrative, not minimalist. *Systematics and biodiversity, 20*(1), 1-14. doi:10.1080/14772000.2022.2063964

Zhang, H., & Bu, W. (2022). Exploring Large-Scale Patterns of Genetic Variation in the COI Gene among Insecta: Implications for DNA Barcoding and Threshold-Based Species Delimitation Studies. *Insects, 13*(5), 425. doi:10.3390/insects13050425

Zhang, H., Ning, X., Yu, X., & Bu, W.-J. (2021). Integrative species delimitation based on COI, ITS, and morphological evidence illustrates a unique evolutionary history of the genus *Paracercion* (Odonata: Coenagrionidae). *PeerJ (San Francisco, CA), 9*, e11459-e11459. doi:10.7717/peerj.11459

Zhang, J., Kapli, P., Pavlidis, P., & Stamatakis, A. (2013). A general species delimitation method with applications to phylogenetic placements. *Bioinformatics, 29*(22), 2869-2876. doi:10.1093/bioinformatics/btt499

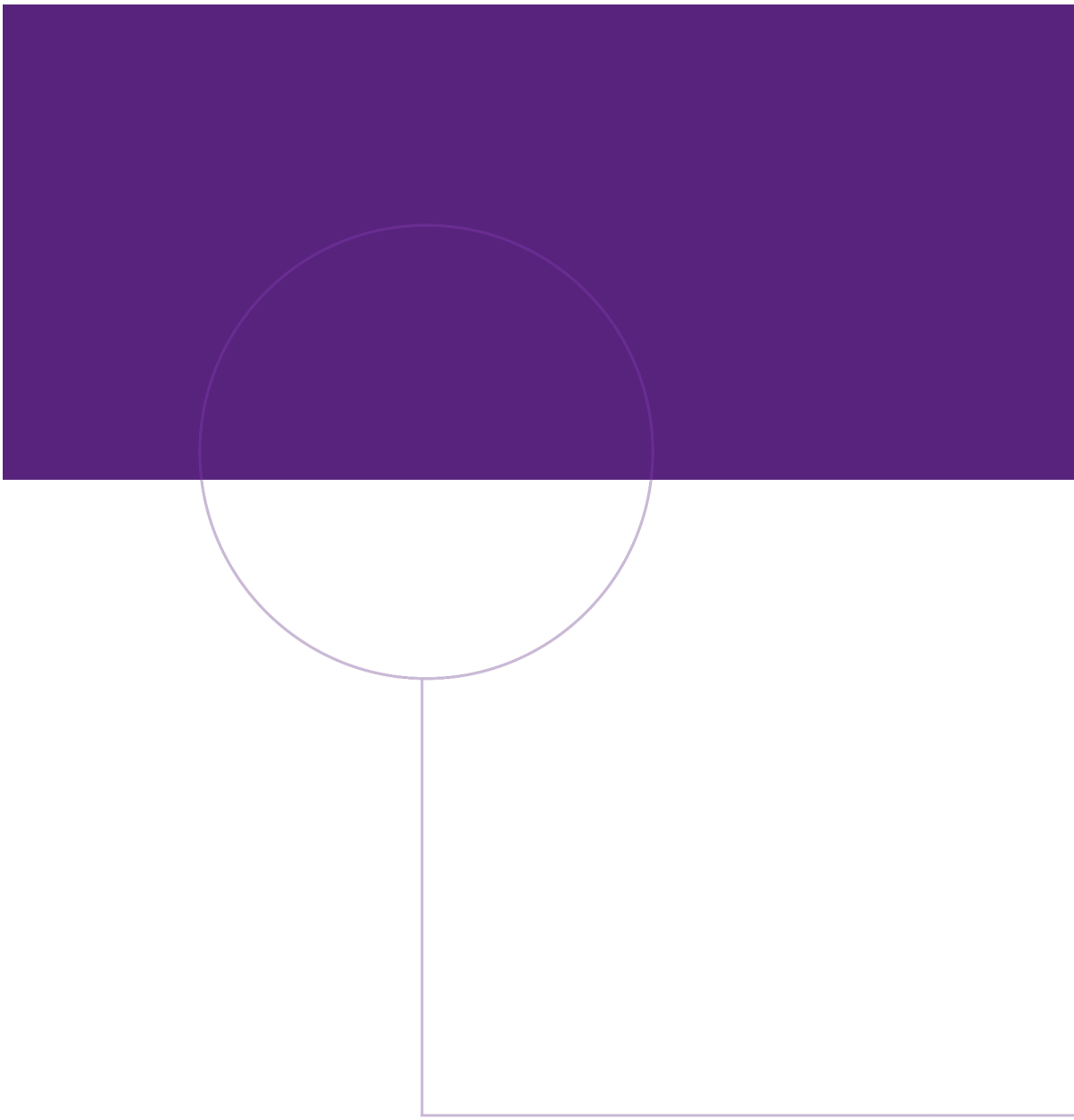
Zug, R., & Hammerstein, P. (2012). Still a host of hosts for *Wolbachia*: analysis of recent data suggests that 40% of terrestrial arthropod species are infected. *PLoS One, 7*(6), e38544-e38544. doi:10.1371/journal.pone.0038544

Aarvik, L., Huemer, P., Berggren, K., Lee, K. M., & Mutanen, M. (2022). Integrative taxonomy reveals overlooked cryptic diversity in the conifer feeding *Batrachedra pinicolella* (Zeller, 1839) (Lepidoptera, Batrachedridae). *ZooKeys, 1085*, 165-182. doi:10.3897/zookeys.1085.76853

Appendix

Appendix 1. Measurements from the specimens where I had measurements for all characters. ID, orange = clade 1, green = clade 2, purple = clade 3. All characters were measured in µm.

ID	GMW	GL	GSW	FL	FW	OP	OE	OA	A3	A4	A5	PW300	PW100	PND	PTT	Costa
MA5	159	377	107	1133	1014	266	321	111	130	314	227	305	188	36	81	2320
MA6	162	338	113	1054	918	245	290	107	120	294	197	305	195	28	83	2114
MA14	195	390	167	1247	1151	288	378	126	145	351	236	331	200	40	67	2580
MA16	199	400	139	1285	1163	311	390	130	135	324	227	313	194	39	61	2546
MA17	176	396	150	1246	1108	285	360	133	139	340	238	341	196	31	72	2553
MA25	162	389	116	1145	1019	293	350	130	138	307	226	316	193	32	71	2330
MA34	192	385	146	1262	1120	259	358	127	155	323	238	334	215	40	71	2610
MA37	185	426	162	1229	1086	317	356	132	138	340	240	313	162	17	42	2500
NOAPI106	165	387	136	1246	1145	307	369	130	133	314	236	329	207	10	58	2510
NOAPI105	160	367	130	1205	1121	307	369	132	146	324	240	335	210	11	57	2370
NOAPI104	152	353	114	1151	1047	315	332	136	119	297	226	305	175	23	47	2170
MA30	164	384	119	1082	998	300	305	134	125	286	206	283	174	12	40	2252
MA38	123	383	97	1119	993	243	305	100	127	269	216	306	189	21	60	2100
MA41	147	391	113	1186	1108	320	326	130	123	294	214	305	198	41	60	2330
NOAPI101	160	382	128	1170	1100	317	360	130	139	324	242	349	222	31	62	2570
NOAPI097	155	368	125	1080	971	315	286	132	141	311	215	347	208	38	67	2470
NOAPI111	128	393	113	1194	1118	310	373	132	125	281	202	309	192	36	62	2535
NOAPI102	150	361	121	1103	1007	285	325	125	126	282	204	275	180	10	51	2327
NOAPI107	157	370	122	1231	1144	341	315	125	130	294	246	308	171	34	46	2470



NTNU

Norwegian University of
Science and Technology