Taking a Bite out of Diversity – Taxonomy and systematics of biting midges

Jonas Strandberg
Taking a Bite out of Diversity

Taxonomy and systematics of biting midges

Jonas Strandberg
To Frida, Vilhelm, Lova and Elvira.
Abstract

The biting midges (family Ceratopogonidae) is one of the most species rich amongst the biting flies (Diptera) and has been recorded from most parts of the world. The species are mostly known for their capability to act as vectors for several important diseases, which have helped in shaping the focus to one of its genera, *Culicoides* Latreille, 1809.

This thesis gives an overview of the knowledge of the Swedish diversity, in the first paper (**paper I**) with a closer look at the species of *Dasyhelea* Kieffer, 1911 where all twenty species found in Sweden are presented with their associated localities, and two new species are described. In the second paper (**paper II**) the biting midge diversity of Sweden is presented based on specimens collected from several localities. All these individuals were barcoded using the mitochondrial cytochrome oxidase I gene (COI). The analysis included 773 specimens that were assigned into 214 barcoding clusters (BINs) and sorted into 164 groups based on their morphology. The third paper (**paper III**) broadens the scale were the evolutionary relationships within the family are investigated by applying five protein coding genes (COI, CAD, TPI, AATS and PGD) and specimens from different parts of the World. The analysis recovers Ceratopogonini, *Forcipomyia* Meigen, 1818 and *Bezzia* Kieffer, 1899 as paraphyletic and *Palpomyia* Meigen, 1818 polyphyletic. In the last and fourth paper (**paper IV**) the family is used as a model organism together with Hymenoptera for an alternative analysis method for reducing the impact of saturation and long-branch attraction using non-synonymous coding (e.g. Degen1) on only parts of a dataset. The effectiveness of the method is compared to the removal of the faster evolving third codon position. The result yields a higher number of supported nodes as well as a higher median of support for the method as well as an ability to reduce long-branch attraction artifacts.

**Keywords:** Ceratopogonidae, *Dasyhelea*, barcoding, COI, phylogeny Sweden, *Forcipomyia*, *Bezzia*, *Palpomyia*, Degen1.
List of papers

This thesis is based on the following papers, referred to by their Roman numerals in the text:


III. Strandberg J & Johanson KA. Evolutionary relationships among higher taxa of biting midges (Diptera: Ceratopogonidae) re-evaluated, based on molecular data of five protein-coding genes. Manuscript.


The published paper is open access and are herein reproduced under the permission of the creative commons by attribution license (CC-BY)
# Contents

Introduction .......................................................................................................................... 11  
  Collecting and preparations ............................................................................................ 13  
  Diversity ............................................................................................................................ 14  
  Aim of the thesis ............................................................................................................... 15  

Summary of papers .............................................................................................................. 17  
  Paper I ............................................................................................................................... 17  
  Paper II ............................................................................................................................. 17  
  Paper III ............................................................................................................................ 18  
  Paper IV ............................................................................................................................ 18  

Discussion and further research .......................................................................................... 19  

References ........................................................................................................................... 21  

Acknowledgements ............................................................................................................. 25  

Svensk Sammanfattning ..................................................................................................... 27
Introduction

The fact that an accurate number of the species inhabiting our planet has not been established is disheartening and at the same time exciting. It is a challenge for taxonomists as the constant loss of diversity means that some species may never be described before their time is at an end. Doom and Gloom aside, this uncertainty also means that there most likely are much more to discover. There will always be a need for taxonomists, both to understand and discern this diversity which at the present only cover a fraction of the organisms in the world (Hammond 1992; Hawksworth & Kalin-Arroyo 1995), and to describe and record our surroundings more efficiently. Due to centralization of societies into towns and cities fewer and fewer people visit natural habitats minimizing encounters with wild life. Some encounters with nature might not be very pleasant such as biting flies spreading diseases and being an overall nuisance to both man and beast alike. Still, even these irritating creatures are vital to understand rather than eradicate; and to understand something it is often best to start with defining what is in front of you. Carl von Linné (1707-1778) did this when he described the first species of biting midges in the 18th century. The description is short, mentioning only hyaline wings with spots and that the species bite leave a colored puncturing.

Since the first species named by Linné (1758) the family Ceratopogonidae now contains more than 6,200 described species distributed over 111 genera (Borkent 2015) and is recorded from all biogeographical regions. A realistic number of extant species was estimated to more than twice of today’s known diversity (Borkent & Spinelli 2007) and is the most species diverse monophyletic group of biting Diptera (Adler 2009).

Ceratopogonidae is one of eight families constituting the infraorder Culicomorpha. The other ones being the non-biting midges (Chironomidae), black flies (Simuliidae), solitary midges (Thaumalidae), mosquitoes (Culicidae), phantom midges (Chaoboridae), frog-biting midges (Corethrellidae) and meniscus midges (Dixidae). The biting midges closely resemble the non-biting midges, which are regarded as the closest relative, particularly due to the presence of a pair of plumose antennae in the males. The biting midges are however, easily distinguished from the non-biting ones by having biting mouthparts, shorter front legs and unique wing venation (Fig 1).
The family is divided into four extant subfamilies: Leptoconopinae Noè, 1907, Dasyheleinae Lenz, 1934, Forcipomyiinae Lenz, 1934, Ceratopogoninae Newman, 1834, and Lebaneseculicoidinae Borkent 2000 including only extinct species. Species in the Ceratopogoninae are further assigned into the tribes Culicoidini, Ceratopogonini, Heteromyiini, Johannsenomyiini, Sphaeromiini s. str and s. lat., Palpomyiini and Stenoxenini (Borkent 2015).

Larval habitats are diverse but require a certain level of humidity to survive and are therefore found in a wide range of biotopes such as damp earth, moss, mud, tree holes, in plants, rock pools, wet meadows, peat bogs, rivers, ponds and lakes and in fresh, brackish as well as saline water (Boorman 1997, Wirth 1978, Mogi & Yong 1992, Disney & Wirth 1982; Szadziewski 1983). But a few species can be ascribed a more terrestrial habitat in the form of animal dung, rotten fungi, ants nests, rotting wood and plants, wet bark and tree sap (Szadziewski et al. 1997, Waugh & Wirth 1976, Graves & Graves 1985, Vattier 1964). Generally, old lineages are restricted to smaller and temporal habitats compared to more recently evolved lineages such as Heteromyiini, Sphaeromiini, Palpomyiini and Stenoxenini, that with some exceptions are to be found in permanent aquatic habitats such as rivers and lakes (Borkent 2005).

Both males and females feed on nectar and/or honeydew, and the females of most species also require intake of extra protein for the maturation of the eggs. The evolutionary older taxa are either blood-feeding on vertebrates hosts ranging from lizards, birds and mammals (including humans) to turtles, frogs and toads (Borkent 2005) (genera Leptoconops Skuse, 1889, Forcipomyia (Lasiohelea) Kieffer, 1921 and Culicoides) or ectoparasites on large insects such as dragonflies (Anisoptera), stick insects (Phasmatida), blister beetles (Meloidae), damselflies (Zygoptera), alderflies (Megaloptera), lacewingflies (Planipennia), butterflies and moths (Lepidoptera) and their larvae, crane flies (Tipulidae), mosquitos (Culicidae); and non-insect arthropods like harvestmen (Opiliones), spiders (Araneae) as well as their prey (klepto-

Fig. 1 Ceratopogonidae (top) and Chironomidae (bottom).
parasitism) (genera *Forcipomyia* and *Atrichopogon* Kieffer, 1906) and larvae of sawflies (Symphya) (Gad 1951, Macfie 1932, Mayer 1937, Seguy 1941, Tokunaga 1939, Clastrier & Legrand 1991, Downes & Smith 1969, Marshall *et al.* 2015, Wirth 1956 and references therein). The females of *Dasyhelea* Kieffer, 1911 however, have reduced mouthparts and do not take blood meals. In the more evolutionary recent lineages (e.g. Ceratopogoninae, excluding *Culicoides*) most females are predators on other Nematocera or small mayflies (Downes 1978, Downes & Wirth 1981). There is also presence of cannibalism in some tribes (Palpomyiini) where the female devours the male during mating, and females have been observed with male genitalia still attached in copula (Downes 1978 and references therein).

![Fig. 2 Malaise trap at lake Bästeträsk, Gotland.](image)

**Collecting and preparations**

The specimens available for the papers in this thesis have mainly been collected using Malaise traps (Fig. 2), and in some instances also sweeping a net along vegetation or just above the ground in muddy or moist areas. These techniques only catch adults. To collect immature in the form of larvae or pupae other techniques must be implemented, such as putting mud or debris into a bucket of water and letting the pupae float up to the surface. These pupae can then be put on moist paper in small containers to then hatch into adults. Raising larvae into adult can be done, but is a difficult task. Biting midges are small in size and to be able to be determined, they must be macerated, a process leaving only the exoskeleton, and then carefully dissected. One leg from each pair, head, wings and abdomen in females and genitalia in
males are separated from the body and subsequently mounted on a microscope slide in separate drops of a medium (Euparal) that hardens over time. When the body parts have been place it is covered by a thin slip of glass. The preparations are permanent and will be placed in a collection, typically the Swedish Museum of Natural History in Stockholm.

Diversity
Linnaeus (1758) named *Culicoides pulicaris* (originally in the genus *Culex*), which represents the first scientific species description of Ceratopogonidae in history. Since then the most important contributions to the knowledge of the Swedish (and World) fauna were given by Zetterstedt (1850, 1852 and 1855), Kieffer (1914, 1916), Lundström (1916), Mayer (1940), Krzywinski (1996), Dominiak & Szadziewski (2010), Nielsen *et al.* (2010) and Szadziewski *et al.* (2013). These contributions have resulted in records of 105 species of Ceratopogonidae, but estimations of the size of the Swedish fauna are given to 130 (Hedström 1994) or 170 species (Rehnberg & Brodin 2010), the latter authors suggested a potential increase of 60% of the Swedish Ceratopogonidae fauna in the future. Recently three species of *Culicoides* (Ander *et al.* 2012, Kirkeby & Dominiak 2014, Nielsen *et al.* 2015), 15 species of *Dasyhelea* of which two new to science (*Paper I*: Strandberg and Johanson 2015), and three species in two genera (*Kolenohelea* Meillon & Wirth, 1981 and *Stilobezzia* Kieffer, 1911) (Strandberg & Stigenberg
2015) have been added to the Swedish fauna. These reports combined resulted in 126 species. **Paper II** in this thesis was based on specimens sampled throughout Sweden (Fig. 3), using mainly Malaise traps, and resulted in recognition of 164 morphological different species distributed over 214 barcoding clusters (BINs), of which many are new to Sweden. At present the Swedish fauna holds approximately 160 species, which will likely increase as undetermined specimens are identified and could surpass previous estimation (Rehnberg & Brodin 2010). In Europe 565 species have been recorded, and the country with highest number of reported species is Germany (256 species) closely followed by France (241 species) and Poland (211 species). Being only a tenth in area compared to Sweden, Estonia has 184 species reported (Szadziewski et al. 2013) (Fig 4).

![European diversity of Ceratopogonidae species modified to represent species per area (km²/1000) of each country. The species per area for Sweden is based on the number of reported species at the start of this PhD project.](image)

**Fig. 4** European diversity of Ceratopogonidae species modified to represent species per area (km²/1000) of each country. The species per area for Sweden is based on the number of reported species at the start of this PhD project.

**Aim of the thesis**

The aim of this thesis was to increase the knowledge about Swedish diversity of Ceratopogonidae and to clarify important taxonomic questions regarding the evolutionary relationships within the family. This was accomplished through collecting, sorting and determination of specimens by using morphology and DNA data (**papers I and II**). The evolutionary relationships
were based on DNA data from multiple molecular markers and resulted in phylogenetic hypotheses (papers III and IV).
Summary of papers

Paper I

In this paper two new species of Dasyhelea Kieffer, 1911 is described for the first time, and includes also a report of fifteen species new to Sweden. The findings is a result of intensive collecting from various localities. Dasyhelea gothlandica and D. dominiakae both belonging in the subgenus Dicryptolescena Enderlein, 1936, where collected in Bästeträsk (Gotland) and Limhamns kalkbrott, (Skåne), respectively. The new species to Sweden are shortly presented with localities, distribution data and biological notes. From these species two new subgenera, Prokempia Kieffer, 1913 and Sebessia Remm, 1979, are reported from Sweden for the first time. The new species added to the Swedish fauna now sets the known species of the genus from the country to twenty.

Paper II
Strandberg J & Johanson KA. Barcoding of Swedish biting midges. Manuscript

The second paper is an expansion of the findings presented in paper I and includes material from other genera acquired through Malaise trap samples from different localities in Sweden. Of these, 773 individuals were sequenced for the mitochondrial cytochrome oxidase I gene (COI) also known as the barcode region, and were found to represent 164 morphological species in 19 genera, 5 tribes and 3 subfamilies. Neighbor-joining analysis of the species recovered 214 barcoding clusters (BINs), of which each cluster potentially represents a distinct species. The findings indicate that the diversity of biting midges in Sweden is far much larger than what currently is known.
Paper III

Strandberg J & Johanson KA. Evolutionary relationships among higher taxa of biting midges (Diptera: Ceratopogonidae) re-evaluated, based on molecular data of five protein-coding genes. *Manuscript*

Previous hypotheses on the evolutionary history of the earliest lineages within Ceratopogonidae were traditionally based on morphological characters of adults and juveniles. In this paper phylogenetic analyses are carried out based on some of the specimens from papers I and II along with specimens from French Guiana, Madagascar, Vietnam, Canada, Turkey, Costa Rica representing genera not present in Europe. The data is molecular sequences of five protein-coding genes, carbamoylphosphate synthetase (CAD), triose-phosphate isomerase (TPI), alanyl tRNA synthetase (AATS), phosphogluconate dehydrogenase (PGD) and cytochrome oxidase subunit I (COI). Approximately 100 species representing 32 out of the 111 known extant genera were included, representing all extant subfamilies and tribes, except Sphaeromiini s. lat. The result raises questions about the monophyly of Ceratopogonini, Palpomyiini, Palpomyia, Bezzia and Forcipomyia.

Paper IV

Strandberg J, Bukontaite R & Malm T. Partial degeneration - a solution to LBA and saturation problems. *Manuscript*

The fourth paper was written due to issues of analyzing a dataset spanning over several hierarchical levels and evolutionary age. Degen1 coding has been implemented on whole data sets in other papers with success. This paper aims to test this approach with a limit to only part of the dataset. Common problems when studying evolutionary relationships are substitution saturation and compositional heterogeneity that may result in erroneous relationships and long-branch attraction artifacts (LBA). We tested the effectiveness of this method to reducing these problems and how it compares to another well-used method, being excluding from analysis the fast-evolving third codon position in protein coding genes. We tested two different datasets, covering both the orders Diptera and Hymenoptera.

Both methods of dataset alternation yielded similar and more likely trees, but our method of partial degeneration results in higher number of supported nodes, a higher median support and manages to resolve LBA artifacts in a data set that is poorly sampled or suffers from substitution saturation or compositional heterogeneity.
Discussion and further research

Through the included papers the diversity of biting midges is presented with many new species previously not reported from Sweden. The highest increase of species is found in *Dasyhelea, Forcipomyia* and *Atrichopogon*. Combining the new reports from *papers I and II* approximately 50 species has been found that are new to the Swedish fauna. Initially, as this project started, the number of recorded species in the country was just below one hundred species, and since 2010 this figure has nearly doubled. Many specimens from *paper II* are still in the need of determination and some are singleton individuals that most likely are species new to science.

A big obstacle in the process of investigating the fauna is the lack of inclusive determination literature. There are no keys for the Swedish or Nordic fauna and the keys that do exist for parts of Europe are often only valid for one genera or a subset of one genus. Creating such tools would be of high priority when venturing forward in the field of Ceratopogonidae. The other part of this thesis is about the evolutionary relationships among taxa in the family. In the thesis the largest molecular analysis of the family is presented (*paper III*). Previous attempts to resolve the phylogeny of higher taxa are almost exclusively based on morphology, and with ambiguous results. An exception is the phylogeny based on pupal data by Borkent (2014) who showed a higher resolution than earlier works. The phylogeny presented in *paper III* expresses many similarities with the paper by Borkent, but raises also many questions regarding the established taxonomy of many genera and also the tribus Ceratopogonini. This subfamily is the most numerous in genera and contains the predatory feeding species. In our analyses only less than one-third of the genera classified in the subfamily are included and increasing representation by more genera is important for fully understanding the evolution of Ceratopogoninae. Also, more in-depth analyses of the respective genera would be necessary to further increase our understanding of the evolution of the family. The analysis performed to recover a phylogeny that is based on protein-coding genes often comes with its own set of issues. The idea of an alternative method for data alteration (*paper IV*) originated when a subset of the data of paper III was analyzed and erroneous pairing of taxa was recovered and when the well-known method of reducing this artefact lead to reduction of support for nodes at the terminal ends of the tree. The method’s ability to comprize a wider range of datasets should be tested through simulations as our findings, although looking promising, is only
based on a small sample and a more thorough statistic comparison should be made.

The genus most frequently investigated is *Culicoides*. The genus contains many species that bite humans, leaving irritating marks from their incision. Some also spread diseases that have the potential to ruin farmers or make you seriously sick (Borkent 2005), and it is understandable that the research focus is on these animals. Still, it is also important to shed light on the other almost 5000 species currently known in the world. One area that is especially neglected is the ecology of immature stages. The information accumulated in paper II is a big step towards this part of the biting midges life cycle due to abilities to safely determine juveniles by using DNA data. With the larval preferences for aquatic and semiaquatic habitat the species in the family could potentially be used as bioindicators and for monitoring health of habitats. We have now made a major step towards making this possible by identifying juveniles of a high number of species via barcoding.

This set of papers is a contribution to the understanding of a very awesome but interesting group of insects.
References


Borkent A 2014. The pupae of the biting midges of the world (Diptera: Ceratopogonidae), with a generic key and analysis of the phylogenetic relationships between genera. Zootaxa. 3879 (1): 001-327. doi: 10.11646/zootaxa.3879.1.1

Borkent A 2015. World species of biting midges (Diptera: Ceratopogonidae).


Acknowledgements

Through the years I have encountered a lot of nice people and I would like to thank you all. Some have been more important than others and deserves maybe a little more gratitude.

For giving me the opportunity to do this along with a lot of advice and inspiring ideas I would like to thank my supervisor, Kjell Arne Johanson. After talking to you I always leave with a head full of ideas, if only I could remember them long enough to implement them. Doing research comes with a lot of problems, often technical, and Tobias Malm has been the “1337 haXxorZ” I have bothered with most of my issues. Thank you for that and not to forget, all the great movie ideas that are bound to win a lot of awards if they would ever be made. Rasa Bukontaite, how have you survived all these years being in the same room as me? You have been an inspiration (and target) for crazy philosophical conundrums, strange business arrangements and suspicious advice. Still waiting for you to sign that contract so we can get started with the “Epic Crocodile Run”. To Julia Stigenberg, always ready for a mock fight, a prank or the food projects instigated in the lunchroom. You are a great person!

To all my colleagues at the museum, Sandra, Hege, Gunvi, Niklas, Yngve, Tin, Johannes, Mattias, Rasmus, Emma, Sibylle, Jeanette. You all make it a very interesting and fun place to work at.

To Frida, my loving wife, who has let me spend all these late nights and long hours at the museum. You have been a balancing point in my existence these years, insisting that working is not everything in life as well as giving a fresh perspective on things. I could not have done this without you. To my little monsters: Vilhelm, Lova and Elvira. You have not made things easier at work, on the contrary, but you are what I long for when at work. I love you all.

To my parents, Nina and Göran. You have been a great help and immensely supportive during these years. You guys have made my life a lot easier and I cannot thank you enough.

During the initial time as a PhD student Dr. Art Borkent and his wife Annette demonstrated their fantastic hospitality during my visit at Salmon Arms, British Columbia, Canada. Furthermore, I want to express my gratitude to Dr. Ryszard Szadziewski for sharing his knowledge of Ceratopogonidae and all the fantastic food during my visits in Gdansk, Poland. A great
thank goes to the Swedish Taxonomy Initiative (the Swedish University of Agricultural Sciences) for the financial support of this project.
Svensk Sammanfattning


En annan fråga som denna avhandling försöker besvara är de evolutionära släktskapen i familjen. I **paper III** används några av de djur från papper I och II samt djur från andra delar av världen för att kunna inkludera släkten som annars saknas i Sverige. Fem protein kodande gener (COI, CAD, TPI, PGD och AATS) användes för att analysera släktskapet av 100 individer från 32 släkten. Resultaten ifrågasätte monofylin hos Ceratopogonini, och Palpomyiini samt släktena Palpomyia, Bezzia och Forcipomyia.

Sista pappret (**IV**) testar en analys metod som reducerar förvillande mönster i sekvenser vid analys samt hur effektiv metoden är jämfört med en annan vanlig metod där snabbt evolverande delar av sekvensen tas bort. Denna metod använder sig av degenerering av synonyma säten på, men i motsats till tidigare arbeten tillämpas det här bara på en del av data setet. Två olika data set med Hymenoptera och Diptera baserade på protein kodande gener. Resultaten visar att en delvis degenerering av data seten är mer effektiv än den metod där delar har exkluderats. Resultaten pekar också på att vanligt förekommande problem som kan ge upphov till felaktiga tolkningar av släktskap kan undvikas.

Arbetena i den här avhandlingen ger tillsammans en förbättrad bild av svidknottens diversitet i Sverige samt hur deras släktskap har utvecklats och kunna ligga till grund för fortsatta studier för att öka kunskapen om denna väldigt intressanta grupp.