Molecular characterization of Swiss Ceratopogonidae (Diptera) and evaluation of real-time PCR assays for the identification of Culicoides biting midges

Wenk, C E; Kaufmann, C; Schaffner, F; Mathis, A

Abstract: Biting midges of the genus Culicoides (Diptera, Ceratopogonidae) are vectors of several viruses of veterinary relevance, and they can cause insect bite hypersensitivity. As the morphological identification of these tiny insects is a difficult task in many cases, alternative approaches are expedient. With the aim to develop real-time PCRs, we determined partial mitochondrial cytochrome oxidase I gene (mt COI) sequences from 380 Culicoides midges representing three regions of Switzerland, namely the Alps, Midland north of the Alps (Atlantic climate), and South of the Alps (Mediterranean climate). The same region was also sequenced from non-biting midges of the genera Atrichopogon, Brachypogon, Dasyhelea, Forcipomyia and Serromyia. A total of 21 Culicoides species were identified by morphology. Sequence variability (haplotypes) was observed in all species. For each of C. grisescens and C. obsoletus, a novel cryptic species was identified. Whereas all individuals of C. grisescens and of the cryptic C. obsoletus species (O2) originated only from Alpine sites, the known C. obsoletus (O1) species was found in all three regions. Further, a sister taxon to C. pulicaris was identified based on the mt COI sequences and named Culicoides sp. Alignments of available mtCOI sequences from Ceratopogonidae (GenBank, this study) were used to design real-time PCR primers and probes to distinguish C. chiopterus, C. deltus, C. dewulfi, C. grisescens (including the cryptic species), C. imicola, C. lupicaris, C. obsoletus O1, C. obsoletus O2, C. pulicaris, C. scoticus and Culicoides sp. Specificities of primers and probes was tested with cloned targets representing 1 to 4 haplotypes of 18 Culicoides spp. and 1 haplotype each from 4 other Ceratopogonidae. No cross-reactivity was observed when plasmid template representing 5 × 10(6) gene copies was tested, but it was evident (Ct values  30) in few instances when plasmid template representing 5 × 10(9) gene copies was utilized, the latter corresponding to the total gene copy number (as determined in this study) in 20 insects. The sensitivities of two assays (C. imicola, C. grisescens) were tested by spiking single insects into pools of 99 or 999, randomly selected non-target Ceratopogonidae (with approx. 90% Culicoides specimens). In the pools of 100, Ct values were in the range of those obtained with single insects when employing 1% of the isolated DNA, whereas the sensitivity with the pools of 1000 was low, presumably due to the low DNA concentrations obtained with a protocol that seems inadequate for these larger pools. Thus, the assays as described are applicable for the specific identification of biting midges in small pools. Primers and probes of this study were devised to be suitable for multiplexed assays but these evaluations await to be performed.

DOI: https://doi.org/10.1016/j.vetpar.2011.08.034
Molecular characterisation of Swiss Ceratopogonidae (Diptera) and evaluation of real-time PCR assays for the identification of *Culicoides* biting midges

Claudia E. Wenk, Christian Kaufmann, Francis Schaffner, Alexander Mathis*

Institute of Parasitology, University of Zurich, Winterthurerstr. 266a, 8057 Zurich, Switzerland.

* Corresponding author. Tel.: +41 44 635 85 01; fax: +41 44 635 89 07; E-mail address: alexander.mathis@uzh.ch (Alexander Mathis).
ABSTRACT

Biting midges of the genus *Culicoides* (Diptera, Ceratopogonidae) are vectors of several viruses of veterinary relevance, and they can cause insect bite hypersensitivity. As the morphological identification of these tiny insects is a difficult task in many cases, alternative approaches are expedient. With the aim to develop real-time PCRs, we determined partial mitochondrial cytochrome oxidase I gene (mt COI) sequences from 380 *Culicoides* midges representing three regions of Switzerland, namely the Alps, Midland north of the Alps (Atlantic climate), and South of the Alps (Mediterranean climate). The same region was also sequenced from non-biting midges of the genera *Atrichopogon*, *Brachypogon*, *Dasyhelea*, *Forcipomyia* and *Serromyia*. A total of 21 *Culicoides* species were identified by morphology. Sequence variability (haplotypes) was observed in all species. For each of *C. grisescens* and *C. obsoletus*, a novel cryptic species was identified. Whereas all individuals of *C. grisescens* and of the cryptic *C. obsoletus* species (O2) originated only from Alpine sites, the known *C. obsoletus* (O1) species was found in all three regions. Further, a sister taxon to *C. pulicaris* was identified based on the mt COI sequences and named *Culicoides* sp. Alignments of available mtCOI sequences from Ceratopogonidae (GenBank, this study) were used to design real-time PCR primers and probes to distinguish *C. chiopterus*, *C. deltus*, *C. dewulfi*, *C. grisescens* (including the cryptic species), *C. imicola*, *C. lupicaris*, *C. obsoletus* O1, *C. obsoletus* O2, *C. pulicaris*, *C. scoticus* and *Culicoides* sp. Specificities of primers and probes was tested with cloned targets representing 1 to 4 haplotypes of 18 *Culicoides* spp. and 1 haplotype each from 4 other Ceratopogonidae. No cross-reactivity was observed when plasmid template representing $5 \times 10^6$ gene copies was tested, but it was evident (Ct values $\leq 30$) in few instances when plasmid template representing $5 \times 10^9$ gene copies was utilized, the latter corresponding to the total gene copy number (as determined in this study) in 20 insects. The sensitivities of two assays (*C. imicola*, *C. grisescens*) were tested by spiking single insects into pools of 99 or 999, randomly selected non-target Ceratopogonidae (with approx. 90% *Culicoides* specimens). In the pools of 100, Ct values were in the range of
those obtained with single insects when employing 1% of the isolated DNA, whereas the sensitivity with the pools of 1000 was low, presumably due to the low DNA concentrations obtained with a protocol that seems inadequate for these larger pools. Thus, the assays as described are applicable for the specific identification of biting midges in small pools. Primers and probes of this study were devised to be suitable for multiplexed assays but these evaluations await to be performed.

**Keywords:**

*Culicoides*, vector identification, barcode, real-time PCR, specificity, cryptic species.
1. Introduction

In northern Europe, biting midges (Diptera, Ceratopogonidae: *Culicoides* spp.) came to the fore as the putative vectors responsible for the unexpected and explosive spread of the bluetongue virus (BTV) that had been introduced there for the first time in 2006 (Carpenter et al., 2009). A number of Palearctic species of *Culicoides* have been incriminated as BTV vectors based on virus isolations or detections by real-time PCR from field-caught midges, by host preferences and vector competence studies (summarized in (Pages et al., 2009). Biting midges are the biological vectors of several other viruses which are of relevance for Europe at present: The African horse sickness virus has repeatedly been observed and eradicated in Southern Europe, and new incursions are considered feasible (Mellor and Hamblin, 2004; Zimmerli et al., 2010). Similarly, the epizootic hemorrhagic disease virus which has a wide distribution in the world but has never been observed in Europe has recently expanded to countries neighboring Europe (Anonymous, 2009; Paweska et al., 2005; Yadin et al., 2008). Another virus, most probably transmitted by biting midges, is the Toggenburg orbivirus which has recently been discovered in Switzerland causing mild bluetongue-like symptoms in sheep and clinically unremarkable infections in goats (Chaignat et al., 2009; Hofmann et al., 2008). In addition to this newly recognized role as vectors in Europe, biting midges are a well-known nuisance pest in many parts of the continent and they can cause insect bite hypersensitivity (named colloquially as ‘sweet itch’) particularly in Equids. Candidate allergens have been characterized from the saliva of model species which can be reared in the laboratory (*C. sonorensis*, *C. nubeculosus*) (Hellberg et al., 2009; Langner et al., 2009; Schaffartzik et al., 2010; Wilson et al., 2008) but the species that cause the clinical symptoms under field conditions are not known (Sloet van Oldruitenborgh-Oosterbaan et al., 2009; van der Rijt et al., 2008).

Little data in terms of the *Culicoides* fauna composition is available for many regions of Europe. In Switzerland, there are 35 established (indigenous) valid species based on published data (Merz et al., 2006). However, several new species have recently been
reported for the first time from Switzerland, and a number of specimens could not
unequivocally be identified by morphology (Cagienard et al., 2006; Casati et al., 2009;
Kaufmann et al., 2009). In comparison, 51 species have been listed for north-eastern France,
a region which has thoroughly been studied (Delécolle, 1985); Delécolle, personal
communication). In order to study the significance of the various *Culicoides* species with
regard to their role as vectors or elicitors of allergy, a comprehensive knowledge of midge
fauna as well as tools for their rapid and easy identification are required. Identification at
present is primarily carried out using morphological features. The observation of wing
patterns allows the classification of the insects into vector-relevant groups such as Obsoletus
and Pulicaris and other *Culicoides* spp. (Goffredo and Meiswinkel, 2004). Further, certain
species can be identified based on wing pattern while others require microscopic analysis of
slide-mounted parts of the bodies (Campbell and Pelham-Clinton, 1960; Delécolle, 1985).
Morphological identification can therefore be a time-consuming procedure and is known to be
a very difficult task in many cases even for expert taxonomists (Meiswinkel et al., 2008) due
to faint characteristics or intraspecific variability (see Pages et al., 2009). Furthermore, the
existence of cryptic species, i.e. morphologically similar midges that are genetically
distinguishable, has recently been described (Pages et al., 2009).
Several PCR-based tests have been developed for the identification of *Culicoides* spp.
targeting the ribosomal RNA genes internal transcribed spacer 1 or 2 (ITS-1, ITS-2) or the
mitochondrial cytochrome oxidase I gene (mt COI, barcode region; (Waugh, 2007). The initial
focus of these approaches was to detect *C. imicola* in pools of insects using a conventional
PCR targeting ITS-1 (Cêtre-Sossah et al., 2004) which was further developed to real-time
quantitative PCRs using SYBR green detection (Cêtre-Sossah et al., 2008; Monaco et al.,
2010). Conventional and multiplexed ITS-1or ITS-2 PCRs were further developed for
identifying individual insects from the Obsoletus group (Gomulski et al., 2005; Mathieu et al.,
2011; Mathieu et al., 2007; Stephan et al., 2009). Relatively few data are available from
these genetic loci, and considerable ITS-1 variation within one species has been
demonstrated (Ritchie et al., 2004). In contrast, sequence information available from the mt COI locus demonstrates low intra-specific divergence (haplotypes) and high inter-specific differences (Augot et al., 2010; Dallas et al., 2003; Linton et al., 2002; Pages et al., 2009). Several conventional and multiplexed PCRs targeting this mt COI locus have been published focusing on the detection of a number of *Culicoides* spp. from the Obsoletus and the Pulicaris groups (Balczun et al., 2009; Nolan et al., 2007; Pages et al., 2009; Schwenkenbecher et al., 2009).

Finally, another molecular technique (matrix-assisted laser desorption/ionization time of flight mass spectrometry; MALDI-TOF MS) has very recently proven its potential for rapid, simple and cost-effective characterization and identification of biting midges (Kaufmann et al., 2011).

The aim of this present study was to develop real-time PCRs for the identification of important biting midges of Switzerland.

### 2. Materials and methods

#### 2.1. Insects

Insects were trapped in Switzerland according to the criteria described (Goffredo and Meiswinkel, 2004) with Onderstepoort UV-light suction traps in the framework of a national entomological monitoring programme (Kaufmann et al., 2009). For the present study, the origin of the insects was classified as Alps (altitude 1000 – 2200 m; 7 trapping sites), Midland (north of the Alps, altitude 400 – 700 m, Atlantic climate; 13 trapping sites) and South (south of the Alps, altitude < 400 m; climate influenced by the Mediterranean Sea; 1 trapping site). Specimens were randomly selected from different catches, and the number of individuals investigated per trapping area roughly corresponds to the trapping effort. *Culicoides imicola* specimens originated from Corsica (France). Prior to morphological identification and DNA extraction, the midges were kept in 70% EtOH at 4 °C.
2.2. Morphological identification

Using a stereo microscope, the *Culicoides* spp. were separated from the other insects and sorted into Obsoletus group, Pulicaris group and other *Culicoides* spp. based on wing morphology (Goffredo and Meiswinkel, 2004). Species identification was achieved by microscopic analyses of wing patterns and by the observation of body parts (head, legs, wings and spermatheca of females or complete genitalia of males) of most specimens mounted on slides according to (Delécolle, 1985). The remains of the abdomen and the thorax were stored in 2 ml round-bottom Eppendorf tubes (Schönenbuch, Switzerland) at -20 °C for DNA isolation. Non-biting midges were identified to genus level according to (Wirth et al., 1974)).

2.3. DNA isolation

Body parts (see above) or pools of midges were ground in 180 µl Tris-EDTA buffer (pH 8.4) using a mixer mill (Retsch®, MM 300) with one (for single insects) or two (for pools of insects) steel bead(s) (3 mm diameter) at 30 Hz for 1 min twice (three times for pooled midges) with an in-between chill down step on ice. The homogenate was then incubated in a heating block for 5 min at 95°C, and total DNA was isolated using the Qiamp DNA mini kit (Qiagen, Hombrechtikon, Switzerland) according to the manufacturer’s instructions. DNA was eluted in 55 µl and was stored at -20 °C until further use.

2.4. Conventional PCRs, cloning, sequencing and data analyses

Part (585 bp) of the mitochondrial cytochrome oxidase subunit I gene (mt COI) was amplified with the primers C1-J-1718_mod (5’-GGWGGRTTTGGWAAYTGAYTAG-3’), modified from a primer described earlier and incorporating degenerate positions (Dallas et al., 2003), and with the new primer CW1_R (5’-AGHWCCAAAAAGTTCYTTTTCC-3’) designed to be insect-specific. The reaction volume of 50 µl consisted of 25 µl of the master mix (6 mM MgCl₂) from the multiplex PCR kit (Qiagen), each primer at a concentration of 1 µM and 10 µl
template DNA. Amplifications were done in an automatic thermal cycler (DNA engine, MJ Research, Bio-Rad Laboratories, Basel, Switzerland) with a profile including a HotStarTaq DNA polymerase activation step (95 °C for 15 min) and 40 cycles at 95 °C for 30 s, annealing at 50°C for 30 s and extension at 72 °C for 60 s. A final elongation step at 72 °C for 10 min was included. PCRs with allegedly Culicoides genus-specific primers (PanCul F / PanCul R targeting the rDNA ITS1 region; genF7 and COIR targeting the mt COI) were done as described (Cêtre-Sossah et al., 2004; Schwenkenbecher et al., 2009).

Cloning of amplicons was achieved using the Topo TA cloning-vector pCR 2.1 (Invitrogen, Carlsbad, CA) according to the manufacturer's manual. The plasmids were purified using the Qiaprep spin miniprep kit (Qiagen) following the manufacturer's instructions. DNA concentration was measured using a Nanodrop photometer (NanoDrop products, Wilmington, USA), the plasmids diluted to stock solutions of 10⁵ copies/µl and stored at -20 °C. Glycerol stocks of transfected bacteria in glycerol (Sigma, Buchs, Switzerland) are kept at -80 °C.

Sequencing of amplicons, either directly after purification with the minelute PCR purification kit (Qiagen) or after cloning, was done by a private company (Synergene GmbH, Schlieren, Switzerland).

Dendrograms were inferred from sequences aligned with ClustalW and from Neighbor-Joining (NJ) analyses using the software MEGA, version 4.1 (Tamura et al., 2007), with default settings.

2.5. Primer and probe design for species-specific real-time PCR assays

Partial mt COI sequences determined in this study and corresponding ones retrieved from GenBank were aligned using Multalin (Corpet, 1988). If necessary, the sequences were manually adjusted. In a stepwise approach, consensus sequences of the species were identified, and appropriate primers and probes were designed using the software Primer Express 1.5TM. If required, the primers were designed degenerated to account for haplotype
variability and/or manually designed, and their annealing temperature was calculated using
the program PerlPrimer. The primers and probes were tested for the formation of dimers and
hairpins using AutoDimerv1 software (default settings). Oligonucleotides that had $\Delta G$
values higher than -5 kcal/mol at 37 °C were redesigned. Primers and probes were ordered at
Microsynth AG (Balgach, Switzerland), except for the minor groove binding (MGB) probes
which were synthesized by Applied Biosystems (Rotkreuz, Switzerland). The ordinary
Taqman probes were PAGE purified and the MGB probes were HPLC purified.

2.6. Real-time PCRs, diagnostic parameters

PCRs were done in duplicates in 25 µl volumes including 12.5 µl iQ multiplex powermix (Bio-
Rad Laboratories, Basel, Switzerland), 2.5 µl of each primer and probe and 5 µl template
DNA. The reactions were run in an iCycler (Bio-Rad) in 96 well plates using the following
cycling conditions: 95 °C for 3 min (hotstart), 35 cycles at 95 °C for 15 s and 60 °C for 1 min.
All results were expressed in Ct values at a fixed threshold (default settings). The optimal
concentrations of primers (50 nM, 300 nM or 900 nM) and probes (50 nM, 100 nM or 200
nM) were determined. The specificities of the primers and probes were tested with $5 \times 10^6$
and $5 \times 10^9$ copies of the cloned target of other Culicoides spp. and other Ceratopogonidae
as template DNA (see table 2). The number of mt COI copies per individual insect was
calculated for 4 Culicoides species (C. scoticus, C. dewulfi, C. imicola and C. pulicaris)
based on standard curves obtained with three concentrations ($5 \times 10^6$, $5 \times 10^4$ and $5 \times 10^3$
copies) of the corresponding cloned target sequence. The sensitivity of two assays was
evaluated by analyzing five pools each of single target individuals (C. imicola or C.
grisescens) added to 99 or 999 randomly selected other Culicoides spp. and approx. 10%
non-biting midges before DNA extraction.

3. Results

3.1. Morphological and genetic identification of midges
A total of 380 *Culicoides* midges from three regions of Switzerland (Alps, Midland north of the Alps with Atlantic climate, South of the Alps with Mediterranean climate) were identified by microscopic analyses of mounted specimens and by determining partial sequences (465 – 541 bp) of the mt COI gene. In addition, 21 insects belonging to 5 genera of non-biting midges and *C. imicola* specimens from Corsica (France) were treated alike (Table 1). The biting midges (*Culicoides* spp.) could be assigned to 21 established species with a preponderance of members of the Obsoletus group, in particular *C. scoticus* (Table 1). For each *C. grisescens* and *C. obsoletus*, a novel cryptic species was identified, i.e. morphologically indistinguishable midges with distinctly different mt COI sequences (Figs. 1A and B; designation according to the suggested nomenclature for cryptic species of (Pages et al., 2009), as *C. grisescens* G1 and G2; *C. obsoletus* O1 and O2, respectively, for the established and the new cryptic species). Translational analyses of the sequences from *C. obsoletus* O1 and O2 using the invertebrate mitochondrial code revealed functional genes differing by 1-2 amino acids (aa) among the total 176 aa. Sequence analyses of a second locus (rDNA ITS1) showed low intra-species variability of the cloned amplicons from two individuals each of *C. obsoletus* O1 and O2 and a 2.8% inter-species diversity, placing the two species in different clusters in a dendrogram (not shown).

In addition, a genetically separate cluster of sequences was identified as a sister taxon to *C. pulicaris* (Fig. 1B). As the careful morphological analyses of these insects yielded different species identifications, including *C. lupicaris* (most specimens), *C. lupicaris/pulicaris* intermediate phenotype, atypical *C. pulicaris* and *C. deltus*, the midges from this cluster were not considered a cryptic species but were tentatively named *Culicoides* sp. The 12 haplotypes of the 34 individuals investigated divide into two branches (Fig 1B), but these do not reflect the morphological heterogeneity observed, as e.g. the 4 midges identified as *C. deltus* are found on both branches (nos. 360-363).

Around half of the species were collected from only one of the three regions, e.g. *C. deltus*, *C. grisescens* (both G1 and G2) and *C. obsoletus* O2 were only found in the Alps, whereas
C. dewulfi and C. lupicaris were identified from Midland traps only. Other species (C. chiopterus, C. obsoletus O1, C. pulicaris and C. scoticus) were present in all three climate regions. The genetic variability of important species (Obsoletus and Pulicaris groups, C. dewulfi) and of the new Culicoides sp. (Table 1A) was further analyzed. A certain degree of intra-species sequence variability (haplotypes) was observed in all species (Figs. 1A and B), e.g. 10 haplotypes were determined for the 20 specimens of C. lupicaris and 11 haplotypes for the 104 C. scoticus insects (for number of insects investigated see table 1). Within the species that occurred in all three investigated areas (C. chiopterus, C. obsoletus O1, C. pulicaris and C. scoticus), no obvious clustering of haplotypes according to the geographic origin was observed. The sequence data are available in GenBank under accession numbers HQ824371 to HQ824525.

3.2. Development of real-time PCRs

Eleven real-time PCR assays targeting the mt COI gene were developed for the specific identification of Culicoides (Tables 1A, 2): C. chiopterus, C. deltus, C. dewulfi, C. grisescens (for both G1 and G2), C. imicola, C. lupicaris, C. obsoletus O1, C. obsoletus O2, C. pulicaris, C. scoticus, Culicoides sp. Primers and probes were designed in silico by considering all sequence information available (GenBank, this study) for Ceratopogonidae. The optimal concentrations for primers and probes were assessed (table 2), and the specificities of the assays were evaluated by PCRs with cloned targets (n=40) of 1 to 4 haplotypes from 18 Culicoides spp. and of 1 haplotype each of 4 other Ceratopogonidae which are regularly found in the UV-light traps (Table 2). Hence, when testing 5 x 10^8 targets of the appropriate species, Ct values between 18.6 and 20 were observed, and all tests with targets from other species were negative. Using 1000-fold higher concentrations of target molecules (5 x 10^9) yielded Ct values of 8.5 - 10.5 for the proper targets, and cross-reactivity (Ct values ≤ 30) was observed in a few instances (Table 2).
The number of targets (mitochondrions) in single insects was evaluated for C. dewulfi, C. imicola, C. pulicaris and C. scoticus, based on standard curves obtained with the respective cloned sequences. Hence, $3.4 \times 10^7$ to $2.8 \times 10^8$ copies per single biting midge were calculated.

Finally, the sensitivities of the two assays for detection of C. imicola and C. grisescens were tested with spiked pools. DNA concentrations obtained from these pools ranged from 107 to 130 ng/µl (100 insects/pool), but were as low as 14 to 39 ng/µl for the larger pools (1000 insects). Real-time PCRs with 5 µl DNA solutions from the pools of 100 mostly were negative. These tests were repeated with 0.5 µl DNA solutions (i.e. 5 µl of a 1:10 dilution).

Hence, one C. imicola or C. grisescens specimen in total 100 midges was detectable with Ct values between 20.8 - 22.6 (C. imicola) or 21.9 - 23.9 (C. grisescens); these values being between 22.9 - 28.4 or 26.1 - 30.9, respectively, when investigating undiluted DNA from the pools containing 1000 insects.

4. Discussion

4.1. Real-time PCRs

Culicoides specimens usually are pre-sorted by morphological features into Obsoletus group, Pulicaris group and other Culicoides spp. The very initial aim of the present study was to develop a triplex real-time PCR to accomplish the same task, with the intrinsic added potential value of also gaining quantitative estimates of the composition of insect pools.

However, the extensive analyses of huge numbers of mt COI sequences from Culicoides spp., either retrieved from GenBank or determined during this study, revealed that this posed an unsolvable challenge for us. Hence, we focused on developing real-time PCR assays for the identification of specific Culicoides biting midges (Tables 1A, 2). For the first time, also corresponding sequences of non-biting midges, which are usual by-catches in light traps, were determined and taken into consideration when designing the Culicoides primers and probes. The necessity of this approach is illustrated by the fact that primers that were
described as being specific for the genus *Culicoides* also amplify DNA from the non-biting midges investigated in this study (see Table 1; results not shown). Hence, the primers genF7/COIR (Schwenkenbecher et al., 2009) targeting the mt COI have identical or highly similar sequences to the respective gene sequences of these non-biting midges, and the expected amplicon of 104 bp was obtained when performing that particular PCR assay e.g. with DNA from *Forsipomyia* sp. Further, supposedly genus-specific primers targeting the 18S and 5.8S rRNA genes and flanking the ITS1 (Cêtre-Sossah et al., 2004) readily produced amplicons in the range of 300 – 550 bp with DNA from the non-biting midges.

The existence of genetic variants (haplotypes) was found in all investigated species. For some of the target midges, degenerate primers had to be designed to account for this sequence heterogeneity at those gene sequence regions which were identified as suitable markers for a species (by reasonably differing from all other available sequences). Further, as the target gene is rather rich in A/T, minor groove binding probes were chosen in some instances to allow for a high specificity of shorter oligonucleotides (Yao et al., 2006).

Specificities of the assays were tested with two concentrations of cloned targets of 22 other midge species, with several haplotypes being tested for some species. Hence, no cross reactivities of the assays were observed when investigating $5 \times 10^6$ targets, whereas the 1000 fold higher concentrations yielded weak reactions in some instances. These cross-reactions, which interestingly mostly occurred only with a single haplotype of a species, were rather unexpected, as primers and probes were designed with several mismatches to non-target sequences, particularly at the 3’ end and in the middle, respectively. For example the probe of *C. scoticus* has 6 mismatches (length of probe: 27nt) and both primers 2 mismatches with the sequence of the cloned haplotype of *C. chiopterus* with which a Ct value of 26 was obtained in the real-time PCR spiked with $5 \times 10^9$ targets (Table 2). However, the number of non-target DNA used in these assays ($5 \times 10^9$) was very high, corresponding to the total number of target genes of at least 20 such non-target midges (which contain between $3.4 \times 10^7$ to $2.8 \times 10^8$ copies of the gene). As outlined below, the total
DNA of this number of insects in a single assay might be too high and cause inhibition of the reaction. Hence, these cross reactivities might be of little concern for practical purposes, and it must be stressed that such evaluations of other published PCR assays were done with DNA aliquots obtained from single biting midges of mostly a limited number only of non-target *Culicoides* species.

Real-time PCRs with DNA from pools of 100 midges for the detection of a single spiked target species was successful when employing 1% of the DNA but were hampered when using 10%, probably due to PCR inhibition. By using 1% of the DNA, approx. $10^6$ targets of the spiked individual are present in the reactions. The Ct values in these assays were around 21 for the detection of *C. imicola* which very well match the value of around 19 - 20 when testing $5 \times 10^6$ cloned targets (Table 2). Hence, this assay is highly sensitive and specific for the detection of a single midge in pools of at least 100. In contrast, DNA isolation from pools of 1000 midges yielded low concentrations, and the protocol applied needs to be improved. However, investigating larger pools of midges might in a few cases give false-positive results, e.g. when using the *C. chiopterus* assay with DNA obtained from pools containing huge numbers of *C. scoticus* (Table 2). Thus, our approach is applicable for the specific identification of midges in pools of less than 1000 midges.

The primers and probes of this study were devised to be suitable for multiplexed assays, i.e. all were analyzed for the formation of dimers with any other oligonucleotide and redesigned if necessary. The fluorophores of the probes (Table 2) were chosen to allow triplex assays e.g. a ‘bluetongue panel’ targeting the major vector species *C. imicola*, *C. obsoletus* O1 and *C. scoticus*, a ‘Pulicaris group panel’ covering *C. pulicaris*, *C. lupicaris*, and *C. grisescens* (both G1 and G2), or an ‘alpine panel’ including *C. deltus*, *C. obsoletus* O2 and *Culicoides* sp. which nearly exclusively have been collected at higher altitudes. However, these investigations have not yet been accomplished, and other combinations with the single assays described (Table 2) are feasible, possibly after adjusting the fluorophores.
4.2. Cryptic species/Culicoides sp.

The analyses of partial mt COI sequences confirmed earlier findings (Augot et al., 2010; Calvo et al., 2009; Dallas et al., 2003; Linton et al., 2002; Pages et al., 2009; Pages and Sarto, 2005) that this locus displays low intra-specific variation and considerable inter-specific variation in *Culicoides* spp. These analyses also revealed the presence of two new cryptic species, morphologically indistinguishable from *C. grisescens* and *C. obsoletus*, respectively, but differing by 10-11% at nucleotide level of the mt COI locus. Such cryptic species have recently been described for the first time in the genus *Culicoides* in three (*C. fagineus*, *C. newsteadi* and *C. pulicaris*) of five investigated species (Pages et al., 2009) by analyzing the same genetic locus as addressed in our study. An alternative explanation of this observed genetic variability could be the existence of mitochondrial DNA sequences that integrated in the nuclear genome (nuclear mtDNA, ‘NUMT’) and evolved as pseudogenes. Such NUMT elements have been identified in some insects (Sunnucks and Hales, 1996; Zhang and Hewitt, 1996), but recent analyses of insect genomes have revealed that in general they are scarce. The two hitherto recognized exceptions among insects are the honey bee *Apis mellifera* and the yellow fever mosquito *Aedes aegypti* which have plenty of NUMTs (summarized in (Black and Bernhardt, 2009). Translational analyses of the mt COI sequences of the cryptic species *C. grisescens* G2 and *C. obsoletus* O2 confirmed that these novel sequences encode for proteins (i.e. no frame shifts or stop codons present). In addition, analyses of *C. obsoletus* O2 at a second locus (rDNA ITS) confirmed its distinctness from and relatedness to *C. obsoletus* O1. Further, real-time PCRs with primers and probes targeting *C. obsoletus* O1 were negative on DNA from 2 individuals identified as *C. obsoletus* O2, and vice versa (not shown). Finally, mass spectrometric analyses by MALDI-TOF allow to differentiate between *C. grisescens* G1 and G2 (own unpublished data; *C. obsoletus* not investigated).

A puzzling finding was the variability of the morphological species designation of specimens which genetically clustered as sister taxon of *C. pulicaris* and which were named *Culicoides*
sp. For the morphological identification of midges species belonging to the Pulicaris group, the decisive criterion using currently available identification keys (Delécolle, 1985) is the wing pattern, all other features considered being non-discriminative. Intra-specific morphological variation has been observed in various species (literature compiled in (Pages et al., 2009), and obviously midges of the newly described species (designated *Culicoides* sp.) display variability in wing patterns causing uncertain identifications. Closer morphometric analyses of these items might allow devising a more precise key with reliable parameters. A similar approach has recently been described for the differentiation of *C. obsoletus* and *C. scoticus* females which were considered undistinguishable (Augot et al., 2010). A cryptic species of *C. pulicaris*, designated *C. pulicaris* P3, has recently been identified (Pages et al., 2009). These specimens were morphologically not discriminable from *C. pulicaris* and their mt COI sequences considerably differ (not shown) from those of the *Culicoides* sp. identified in Switzerland. Hence, based on the genetic data and on the morphological ambivalence, *Culicoides* sp. indeed seems to be a novel species.

4.3. Geographical distribution of species/haplotypes

Several *Culicoides* species have only been found in one of the three investigated climatic regions of Switzerland (Alps, midland north of the Alps with Atlantic climate, region south of the Alps with Mediterranean climate), but, clearly, the number of investigated specimens is too low to draw conclusions on their definitive geographic distribution. Interestingly, both discovered cryptic species (*C. grisescens* G2, *C. obsoletus* O2) were only sampled in the Alpine region; from the new species *Culicoides* sp. all but one specimen also originated from higher altitudes. It remains to be elucidated whether these species are truly adapted to this climate.

Specimens from 4 species (*C. chiopterus*, *C. obsoletus* O1, *C. pulicaris* and *C. scoticus*) were collected in all three climatic regions, and no clustering of haplotypes according to the geographic origin was obvious. Thus, the alpine crest seems not to be a barrier for the
dispersal of these midges. Further, co-clustering of mt COI sequences with those from Swiss midges is observed for sequences from Spain (C. pulicaris, C. scoticus), Great Britain (C. pulicaris, C. grisescens G1, C. scoticus, C. chiopterus, C. dewulfi) and from Bulgaria (C. obsoletus O1) (Figs. 1A and B), indicating a lack of barriers for these species in Europe. In contrast, Spanish C. obsoletus O1 and C. lupicaris are placed on separate branches, supported by high bootstrap values (Fig. 1). Therefore, the population genetic structure of the investigated midge species might differ across Europe, and this is of ecological and epidemiological significance, particularly with regard to vector competence.

Acknowledgements

We kindly thank Jeannine Hauri for her excellent technical support in the lab and all the farmers for carefully and reliably running the light traps. Furthermore, we are indebted to Dr. J.-C. Delécolle (Université Louis Pasteur, Strasbourg) for his inestimable assistance in the morphological identification of many of our trapped Ceratopogonidae biting midges and for providing the C. imicola specimens. This study was funded by the Swiss Federal Veterinary Office (project No. 1.08.10) and represents the doctoral thesis of Claudia E. Wenk, veterinarian.

References


*Culicoides obsoletus* s.s. and *Culicoides scoticus* in Southwest Germany identified by

(NUMTs) in the *Aedes aegypti* genome. Insect Mol. Biol. 18, 705-713.


Calvo, J.H., Calvete, C., Martinez-Royo, A., Estrada, R., Miranda, M.A., Borras, D., Sarto,
mitochondrial cytochrome c oxidase subunit I gene indicate northward expanding

*Culicoides* Latreille (Diptera, Ceratopogonidae). Proc. R. Soc. Edinburgh, Section B
67, 181-302.

Carpenter, S., Wilson, A., Mellor, P.S., 2009. *Culicoides* and the emergence of bluetongue

An investigation on the *Culicoides* species composition at seven sites in southern

Cêtre-Sossah, C., Baldet, T., Delécolle, J.-C., Mathieu, B., Perrin, A., Grillet, C., Albina, E.,
2004. Molecular detection of *Culicoides* *spp.* and *Culicoides imicola*, the principal
vector of bluetongue (BT) and African horse sickness (AHS) in Africa and Europe.

Cêtre-Sossah, C., Mathieu, B., Setier-Rio, M.L., Grillet, C., Baldet, T., Delécolle, J.-C.,
Albina, E., 2008. Development and evaluation of a real-time quantitative PCR assay
for *Culicoides imicola*, one of the main vectors of bluetongue (BT) and African horse


Figure 1. Dendrograms inferred from partial mitochondrial COI gene sequences of *Culicoides* spp. from A. the Obsoletus group (C. obsoletus, C. scoticus, C. chiopterus) and C. dewulfi and B. the Pulicaris group (C. pulicaris, C. lupicaris, C. deltus, C. grisescens) and *Culicoides* sp. Depicted are all haplotypes identified in the three investigated regions in Switzerland (Alps, Midland north of the Alps, South of the Alps). The number of isolates analyzed per taxonomic unit is given in Table 1. Included are single sequences from *C. imicola* (C. imi), *Forcipomyia* sp. (Forc. sp.), *C. pulicaris* (C. pul) (Fig 1A), *C. obsoletus* (C. obs) (Fig 1B) and corresponding *Culicoides* GenBank entries if available. Sequences were aligned with ClustalW and the tree was deduced from Neighbour-Joining (NJ) analyses using the software MEGA, version 4.1 (Tamura et al., 2007). Bootstrap values are given on the nodes. Sequences from GenBank are supplemented with the country origin of the midge (BG Bulgaria, ES Spain, GB Great Britain).
Fig. 1A

C. obsoletus O1

C. obsoletus O2

C. scoticus

C. chiopterus

C. dewulfi
Table 1: Trapping regions and number of midges (Ceratopogonidae) investigated. *Culicoides*
identification was based on morphological and/or genetic (partial sequence of mt COI gene) analyses.

<table>
<thead>
<tr>
<th>Species</th>
<th>Trapping regions in Switzerland&lt;sup&gt;1&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Alps</td>
</tr>
<tr>
<td>A. <em>Culicoides</em> for which real-time PCRs were evaluated in this study&lt;sup&gt;2&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td><em>C. chiopterus</em></td>
<td>2</td>
</tr>
<tr>
<td><em>C. deltus</em></td>
<td>7</td>
</tr>
<tr>
<td><em>C. dewulfi</em></td>
<td>0</td>
</tr>
<tr>
<td><em>C. grisescens G1</em></td>
<td>7</td>
</tr>
<tr>
<td><em>C. grisescens G2</em></td>
<td>12</td>
</tr>
<tr>
<td><em>C. lupicaris</em></td>
<td>0</td>
</tr>
<tr>
<td><em>C. obsoletus O1</em></td>
<td>7</td>
</tr>
<tr>
<td><em>C. obsoletus O2</em></td>
<td>9</td>
</tr>
<tr>
<td><em>C. pulicaris</em></td>
<td>2</td>
</tr>
<tr>
<td><em>C. scoticus</em></td>
<td>5</td>
</tr>
<tr>
<td><em>Culicoides</em> sp.</td>
<td>33</td>
</tr>
</tbody>
</table>

B. Other *Culicoides* spp.

| *C. brunnicans*                       | 3    | 0       | 0     |
| *C. circumscriptus*                   | 0    | 4       | 4     |
| *C. comosioculatus*                   | 4    | 2       | 0     |
| *C. duddinstoni*                      | 0    | 1       | 0     |
| *C. fascipennis*                      | 5    | 1       | 0     |
| *C. festivipennis*                    | 0    | 2       | 6     |
| *C. furcillatus*                      | 0    | 3       | 0     |
| *C. jurensis*                         | 3    | 0       | 0     |
| *C. kibunensis*                       | 0    | 0       | 1     |
| *C. pallidicornis*                    | 0    | 3       | 7     |
| *C. punctatus*                        | 0    | 8       | 1     |
| *C. reconditus*                       | 1    | 0       | 0     |
| *C. segnis*                           | 1    | 0       | 0     |

C. Non-biting Ceratopogonidae

| *Atrichopogon* sp.                    | 1    | 0       | 0     |
| *Brachypogon* spp.                    | 2    | 0       | 0     |
| *Dasyhelea* spp.                      | 0    | 2       | 0     |
| *Forcipomyia* spp.                    | 0    | 15      | 0     |
| *Serromyia* sp.                       | 0    | 1       | 0     |

<sup>1</sup>Alps (altitude 1000 – 2200 m), Midland (north of the Alps, altitude 400 – 700 m, Atlantic climate) and South (south of the Alps, altitude < 400 m; climate influenced by the Mediterranean Sea).
In addition, a real-time PCR was designed and evaluated for *C. imicola* of which individuals originating from Corsica (France) were available (kindly provided by J.-C. Delécolle).

PCR assay designed to be specific for both *C. grisescens* G1 and G2.
Table 2: Sequences of primers and probes (5’ - 3’) for real-time PCR assays, and cross-reactivities observed.

<table>
<thead>
<tr>
<th>Attempted specificity</th>
<th>Primers and probes (optimal concentration in nM)</th>
<th>Cross-reactivity with other Ceratopogonidae(^1) (Ct value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C. chiopterus</td>
<td>chi_F AGGTATAGTTCTATTTTAGGGGCT (300)</td>
<td>C. imicola (28), C. scoticus (30)</td>
</tr>
<tr>
<td></td>
<td>chi_R AATGATAAAAGRAGTAAAATGGCAATKAGA (300)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>chi_P FAM_CATATTAAATACGGTTTAAGGAATAAC-NFQ-MGB</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C. deltus</td>
<td>del_F TGGAACTGGATGAACCGTA (300)</td>
<td>Culicoides sp. (^4) (29)</td>
</tr>
<tr>
<td></td>
<td>del_R AGAAGAAATYCTCTGCTAAATCTTAGT (300)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>del_P Cy5 -TCACGCTGGGCCCTGAGTAGATTGA-BHQ-2 (50)</td>
<td></td>
</tr>
<tr>
<td>C. dewulfi</td>
<td>dew_F ATGCCGGAGCCTCG (300)</td>
<td>None</td>
</tr>
<tr>
<td></td>
<td>dew_R GGGATTTTTCGTTAATATTCTCTTTTGT (300)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>dew_P Cy5 -TCACGCTGGGCCCTGAGTAGATTGA-BHQ-2 (50)</td>
<td></td>
</tr>
<tr>
<td>C. grisescens G1 + G2</td>
<td>gri_F CMYTWCATYTWGGAGGTATTTAT (300)</td>
<td>Culicoides sp. (^4) (30)</td>
</tr>
<tr>
<td></td>
<td>gri_R CTAARACTGGRAGWRGAWARAGGAAAG (900)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>gri_P FAM_TGGAATACCTTTGTGCAATAC-NFQ-MGB (200)</td>
<td></td>
</tr>
<tr>
<td>C. imicola</td>
<td>imi_F TCCTCGAATAAAATAATAGTTTTGATATTA (300)</td>
<td>None</td>
</tr>
<tr>
<td></td>
<td>imi_R ACATTTGTGTAATGGAGGA (300)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>imi_P FAM_CCATCTATTACTCTCTCTTTATATAAG-NFQ-MGB</td>
<td></td>
</tr>
<tr>
<td>C. lupicaris</td>
<td>lup_F AATTCTCTATTCAGGTGCTGTG (300)</td>
<td>None</td>
</tr>
<tr>
<td></td>
<td>lup_R GCCAAAACTGTGAAAGGAAAATGAAATAT (300)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>lup_P ROX -ATGCGATCTAATGGAGTTTATCGACCAGTACC-BHQ-2 (100)</td>
<td></td>
</tr>
<tr>
<td>C. obsoletus O1</td>
<td>obs1_F GAAAAYGGAGCAACC (50)</td>
<td>C. scoticus (^4) (29)</td>
</tr>
<tr>
<td></td>
<td>obs1_R GAAAATAAGCATAATCCAAG (300)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>obs1_P VIC -TGCAAGGAGATTTAGTGAAGAG-GNFQ-MGB</td>
<td></td>
</tr>
<tr>
<td>C. obsoletus O2</td>
<td>obs2_F GGAGCCGTAATTATTTTAAACC (50)</td>
<td>C. scoticus (^4) (27/28), C. chiopterus (^4) (30)</td>
</tr>
<tr>
<td></td>
<td>obs2_R CTGCTAATACAGGAAAGATAGTGG (50)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>obs2_P FAM-TGGAATAATCTACTTTTGAGTTTATTTGATCAGT-BHQ-2 (100)</td>
<td></td>
</tr>
<tr>
<td>C. pulicaris</td>
<td>pul_F CGGAAGCTATTTGAGGTATG (300)</td>
<td>None</td>
</tr>
<tr>
<td></td>
<td>pul_R AATGGTTGTCAGTTAAATGYAGA (300)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>pul_P Cy5 -ACTGTTACATCTCTCTCTCCACTAGAGGACC-BHQ-2 (100)</td>
<td></td>
</tr>
<tr>
<td>C. scoticus</td>
<td>sco_F CCCCCCCATCTAGCGA (300)</td>
<td>C. chiopterus (^4) (26)</td>
</tr>
<tr>
<td></td>
<td>sco_R GCTAATACCTGCTAAATGYAGA (300)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>sco_P ROX -TGCTCTCAGCTGCAAGGCTACATTGA-BHQ-2 (50)</td>
<td></td>
</tr>
<tr>
<td>Culicoides sp.</td>
<td>Csp_F AACGGAGCCGGTACC (300)</td>
<td>None</td>
</tr>
<tr>
<td></td>
<td>Csp_R YCCGAGTAAAGAAATACCG (300)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Csp_P HEX-TGCAATATTTTCTCATGGCCCGAGCAGTATTGA-BHQ-1 (100)</td>
<td></td>
</tr>
</tbody>
</table>
1) Ct value ≤ 30 cycles when tested with 5 x 10^9 cloned targets. No cross-reactivity (i.e. Ct value ≤ 40 cycles) at all was observed when testing 5 x 10^6 targets. For comparison: Ct values with proper sequences were 8.5-10.5 (5 x 10^9 targets) and 18.6-20 (5 x 10^6 targets).

2) Species and numbers of haplotypes tested: C. chiopterus: 3; C. circumscriptus: 2; C. comosioculatus: 1; C. deltus: 2; C. dewulfi: 1; C. fascipennis: 1; C. festivipennis: 1; C. grisescens G1: 2; C. grisescens G2: 1; C. imicola: 1; C. lupicaris: 1; C. obsoletus O1: 4; C. obsoletus O2: 3; C. pallidicornis: 1; C. pulicaris: 2; C. punctatus: 2; C. scoticus: 3; Culicoides sp.: 4; one haplotype each of other Ceratopogonidae: Atrichopogon sp.; Brachypogon sp.; Dasyhelea sp.; Forcipomyia sp.; Serromyia sp.

3) MGB: minor groove binding probe.

4) Only one haplotype of the indicated Culicoides species positive; for number of haplotypes tested per species see footnote 2 above.