Identification of field-caught *Culicoides* biting midges using matrix-assisted laser desorption/ionization time of flight mass spectrometry

CHRISTIAN KAUFMANN[1], FRANCIS SCHAFFNER[1], DOMINIK ZIEGLER[2], VALENTIN PFLÜGER[2] *and* ALEXANDER MATHIS[1]*

[1] *Vector Entomology Unit, Institute of Parasitology, Vetsuisse Faculty, University of Zürich, Zürich, Switzerland*
[2] *Mabritec SA, Riehen, Switzerland*

SUMMARY

Culicoides biting midges are of great importance as vectors of pathogens and elicitors of allergy. As an alternative for the identification of these tiny insects, matrix-assisted laser desorption/ionization time of flight mass spectrometry (MALDI-TOF MS) was evaluated. Protein mass fingerprints were determined for 4–5 field-caught reference (genetically confirmed) individuals of 12 *Culicoides* species from Switzerland, *C. imicola* from France, laboratory-reared *C. nubeculosus* and a non-biting midge. Reproducibility and accuracy of the database was tested in a validation study by analysing 108 mostly field-caught target *Culicoides* midges and 3 specimens from a non-target species. A reference database of biomarker mass sets containing between 24 and 38 masses for the different species could be established. Automated database-based identification was achieved for 101 of the 108 specimens. The remaining 7 midges required manual full comparison with the reference spectra yielding correct identification for 6 specimens and an ambiguous result for the seventh individual. Specimens of the non-target species did not yield identification. Protein profiling by MALDI-TOF, which is compatible with morphological and genetic identification of specimens, can be used as an alternative, quick and inexpensive tool to accurately identify *Culicoides* biting midges collected in the field.

Key words: *Culicoides*, biting midge, MALDI-TOF MS, identification, insect, vector.

INTRODUCTION

Culicoides biting midges are the proven or suspected biological vectors for a number of viruses, including bluetongue virus, African horse sickness virus, epizootic haemorrhagic disease virus and Toggenburg orbivirus (Mellor *et al*. 2000; Mellor and Hamblin, 2004; Paweska *et al*. 2005; Meiswinkel *et al*. 2007; Carpenter *et al*. 2009; Chaignat *et al*. 2009). In addition, they are a nuisance pest and can cause insect bite hypersensitivity, particularly in Equids (Hellberg *et al*. 2009; Sloet van Oldruitenborgh-Oosterbaan *et al*. 2009). In order to study the significance of the various *Culicoides* species with regard to their role as vectors or elicitors of allergy, tools for their rapid and easy identification are required. The morphological identification of these tiny (1–3 mm) insects to species level is very difficult in many cases (Goffredo and Meiswinkel, 2004; Meiswinkel *et al*. 2008; www.culicoides.net) requiring the time-consuming analyses of slide-mounted microscopical insect preparations (Delécolle, 1985). Hence, in large

entomological surveys, trapped midges are grossly separated based on morphological features of wing patterns into Obsoletus group, Pulicaris group and other *Culicoides* spp. (Goffredo and Meiswinkel, 2004). Several PCR-based tests have been developed for the specific identification of a number of species, in single- or multiplexed assays both in conventional and in real-time PCR formats (compiled by Kaufmann *et al*. 2011 and Wenk *et al*. 2011).

Matrix-assisted laser desorption/ionization time of flight mass spectrometry (MALDI-TOF MS) has emerged as an alternative technique for species identification. This proteomic approach has come of age for the high throughput, accurate and reproducible identification of clinically relevant microorganisms (bacteria, yeasts, filamentous fungi) at low cost and minimal preparation time (Mellmann *et al*. 2009; Santos *et al*. 2010; Sauer and Kliem, 2010; Stevenson *et al*. 2010; van Veen *et al*. 2010). This technique has been described for the identification of metazoans, namely fish species (Mazzeo *et al*. 2008), plants (lentil varieties, Caprioli *et al*. 2010) and insects (*Drosophila* fruit flies, Campbell, 2005; Feltens *et al*. 2010; aphid species, Perera *et al*. 2005). In a recent study (Kaufmann *et al*. 2011), we have demonstrated the suitability of MALDI-TOF MS analysis to reproducibly identify biomarker masses from laboratory-reared *C. nubeculosus*,

* Corresponding author: Vector Entomology Unit, Institute of Parasitology, Vetsuisse Faculty, University of Zürich, Winterthurerstrasse 266a, CH-8057 Zürich, Switzerland. Tel: +41 44 635 85 01. Fax: +41 44 635 89 07. E-mail: alexander.mathis@uzh.ch

independent of the insects' sex, age and the duration of their storage in 70% ethanol up to 102 days. The only prerequisite was to remove the abdomen because host blood in female midges blurred the spectra in the first few days after the bloodmeal. Further, no differences were found in the spectra from thoraxes only or from thoraxes with head, legs and wings, leaving these parts for other purposes, e.g. morphological and genetic identification.

With this study, we aimed at establishing a reference database of biomarker masses of field-caught biting midges and to evaluate this MALDI-TOF-based identification system in a blind study.

MATERIALS AND METHODS

Insect origins and morphological and genetic identification

Ceratopogonid midges were caught between 2008 and 2010 on 10 cattle farms in 3 regions of Switzerland (Alps, Midland north of the Alps with Atlantic climate, South of the Alps with Mediterranean climate; Kaufmann *et al*. 2009) using Onderstepoort UV-light suction traps as described (Venter and Meiswinkel, 1994). Whole insects were stored between 2 and up to 24 months in 70% ethanol at room temperature. *Culicoides* spp. were identified under a stereo-microscope, based on wing pattern and morphological features (Delécolle, 1985). *Culicoides imicola* specimens originated from Corsica (France), and *C. nubeculosus* (initially provided by the Institute for Animal Health, Pirbright, UK) were obtained from our in-house colony maintained according to Boorman (1974), with the exceptions that larvae were fed with pulverized Tetramin® only and that blood feeding was achieved through a Nescofilm®-membrane with fresh heparinized sheep blood. The non-biting midges of another ceratopogonid genus, *Forcipomyia*, were selected from one of the trapping sites from the Midland (Dittingen/BL; Kaufmann *et al*. 2009), and individuals of the predominant morphotype of this genus were used for the study.

Individuals were dissected using a stereo-microscope as described (Kaufmann *et al*. 2011). From the abdomens, DNA was isolated with a kit (Qiamp DNA mini kit, Qiagen, Hildesheim, Germany) according to the manufacturer's instruction after mechanical homogenization as described (Wenk *et al*. 2011).

Genetic species confirmation was done by PCR/ sequencing of 585 bp of the mitochondrial cytochrome oxidase subunit I gene (mt COI). The primers C1-J-1718-mod (5′-GGW GGR TTT GGW AAY TGA YTA G-3′) and C1-N-2191-mod (5′-AGH WCC AAA AGT TTC YTT TTT CC-3′) were modified from previously described primers (Dallas *et al*. 2003) and were devised to be specific for

insects, by considering the corresponding sequences of *Aedes aegypti*, *D. melanogaster*, *C. arakwae*, *C. dewulfi* and *Homo sapiens* (Wenk *et al*. 2011). In addition, species conformation at a different locus (rDNA internal transcribed spacer 1, ITS1) according to Cêtre-Sossah and colleagues (2004) was occasionally performed. Direct sequencing of the amplicons was performed by a private company (Synergene, Schlieren, Switzerland).

Sample preparation and MALDI-TOF MS parameters

Thoraxes with head, wings and legs were manually homogenized in 20 μl of formic acid (10%) and 5 μl thereof were mixed with 7·5 μl of SA (saturated solution of sinapic acid in 60% acetonitrile, 40% H_2O, 0·3% trifluoroacetic acid; Sigma-Aldrich, Buchs, Switzerland). One μl was spotted within a maximum of 2 h onto a steel target plate in quadruplet for reference spectra or in duplicate in the validation study, and air-dried at room temperature. Protein mass fingerprints were obtained using a MALDI-TOF Mass Spectrometry Axima™ Confidence machine (Shimadzu-Biotech Corp., Kyoto, Japan) as described earlier (Stephan *et al*. 2010) with the threshold apex peak detection being performed in the dynamic threshold type setting using an offset to 0·020 mV and a threshold response factor of 1·2. All target plates were externally calibrated by *Escherichia coli* DH5α as the reference strain.

Peak matrix generation for unsupervised cluster analysis

Generated protein mass fingerprints were analysed with SARAMIS software (Spectral Archive and Microbial Identification System, AnagnosTec, Potsdam-Golm, Germany). Binary matrix was generated using the SARAMIS SuperSpectrum tool and exported to a text file. Intensity and error columns were removed and average spectra were generated, eliminating masses with a presence lower than 50% within the 4 replicates. The adapted binary matrix was imported into PAST freeware, and multivariate cluster analysis was performed using the paired group dice algorithm. The generated dendrogram was exported in nexus file format, and the dendrogram illustration was performed by the FigTree freeware.

Generation of MALDI-TOF MS biomarker mass sets

Protein mass fingerprints were determined in quadruplets from the 69 reference *Culicoides* biting midges covering 14 species and one non-biting midge of the genus *Forcipomyia*. Five genetically confirmed specimens (only 4 for a cryptic species of *C. grisescens* denominated *C. grisescens* II; Wenk *et al*. 2011) of the following species were included (origin: A: Alps,

M: Midland north of the Alps with Atlantic climate, S: South of the Alps with Mediterranean climate): (a) Obsoletus group: *C. chiopterus* (M and S); *C. obsoletus* (M and S); *C. scoticus* (M); (b) Pulicaris group: *C. deltus* (A); *C. grisescens* I (A); *C. grisescens* II (A); *C. lupicaris* (M); *C. pulicaris* (M); *C. punctatus* (M); (c) other *Culicoides* spp.: *C. circumscriptus* (S); *C. dewulfi* (M); *C. festivipennis* (M and S); *C. imicola* (Corsica, France); *C. nubeculosus* (from in-house colony, only 2 specimens genetically analysed).

Only reads with at least 50 peaks were further included. The peak lists were imported into SARAMIS, trimmed to a mass range of 2–20 kDa, and peaks with a relative intensity below 1% were removed. Peak lists were binned and average bio-marker masses were calculated using the SARAMIS SuperSpectrum tool with an error of 800 ppm. The specificity of these potential biomarker masses was determined by comparison against the whole SARAMIS spectral archive and additional Mabritec-owned spectral data sets including more than 90 000 spectra covering >2700 different species of various taxa (mostly bacteria, but also fungi, eukaryotic cell lines, and a few insects; Kaufmann *et al*. 2011). In accordance with the SARAMIS user guidelines, the threshold for identification was set at 75% biomarker matches based on the reference data set.

Validation study

For the validation of the biomarker mass sets, 111 *Culicoides* biting midges were analysed in duplicate (blind study). These included midges caught in the field in the 3 regions of Switzerland and morpho-logically identified as belonging to one of the species for which a biomarker mass set was generated (*C. chiopterus*, n = 3; *C. circumscriptus*, n = 3; *C. deltus*, n = 5; *C. dewulfi*, n = 2; *C. festivipennis*, n = 3; *C. grisescens*, n = 2; *C. lupicaris*, n = 2; *C. obsoletus* (males), n = 5; *C. obsoletus/scoticus* (females), n = 37; *C. pulicaris*, n = 5; *C. punctatus*, n = 22; *C. scoticus* (males), n = 8). In addition, *C. imicola* (n = 1); *C. nubeculosus* (n = 10) from the colony and, as a nontarget species, field-caught *C. pallidicornis* (n = 3) were included. All specimens from the validation study were genetically analysed.

Generated mass fingerprints were imported into SARAMS software for automated identification against >3400 biomarker mass sets, including 16 insect species-specific sets and, if required, for manual full spectra comparison against the 69 reference *Culicoides* biting midges.

RESULTS

Creating a reference database of biomarker masses

MALDI-TOF MS reference spectra were deter-mined for 12 field-caught *Culicoides* biting midge species from Switzerland, including a cryptic species of *C. grisescens* (denominated *C. grisescens* II), *C. imicola* from France, from laboratory-reared *C. nubeculosus* as well as from 1 non-biting midge of the ceratopogonid genus *Forcipomyia*. In general, 5 individuals per species (total 74 specimens), morphologically determined and confirmed by genetic analyses, were used to develop an MS reference database. The protein profiles of these *Culicoides* biting midges revealed between 50 and 169 peaks in the mass range of 2 to 20 kDa with an average of 99 peaks per spectrum. In Fig. 1, profiles of individuals belonging to 6 biting midge species (i.e. *C. obsoletus*, *C. scoticus*, *C. dewulfi*, *C. imicola*, *C. pulicaris*, and *C. punctatus*) are shown for the range between 5 and 10 kDa.

The total mass spectra of all 69 reference biting midges and of 1 *Forcipomyia* specimen as outgroup were used to generate a dendrogram (Fig. 2). Thus, all individuals of a species clustered on discrete branches. In particular, species that are indistin-guishable by morphology (i.e. females of *C. obsoletus/scoticus*, the cryptic species *C. grisescens* I/II) are clearly distinguished by MALDI-TOF MS.

Based on this clear-cut clustering of the species, specific biomarker mass sets, containing between 24 and 38 masses for the different species (Fig. 1, Table 1), could be generated using SARAMIS and be used for automated *Culicoides* species identification. The above mentioned morphologically similar species had only few common biomarker masses, namely 4 (from 26 and 38, respectively, of *C. obsoletus* and *C. scoticus*) or 7 (from 25 and 30 of *C. grisescens* I/II) (Table 1).

Validation study

The reproducibility and accuracy of the reference database were tested in a second step in a validation study. Analyses against the biomarker mass sets of all reference species were performed for 2 profiles per individual. From 90 of the 108 specimens investi-gated, both profiles allowed for an automated identification. Eleven individuals yielded only 1 profile suitable for identification by automated analysis, whereas the replicate had fewer than 75% biomarker matches, requiring manual full spectra comparison which in all these cases confirmed the result of the automated classification. From 7 midges, both profiles were of lower quality (displaying fewer masses and lower intensities), requiring manual full spectra comparison which in 5 of these cases consistently identified the species. In another case, only 1 of the duplicates yielded an analysable profile. The seventh of these specimens, morpho-logically and genetically identified as *C. deltus*, resulted in spectra which most closely resembled those derived of the reference *C. deltus* specimens
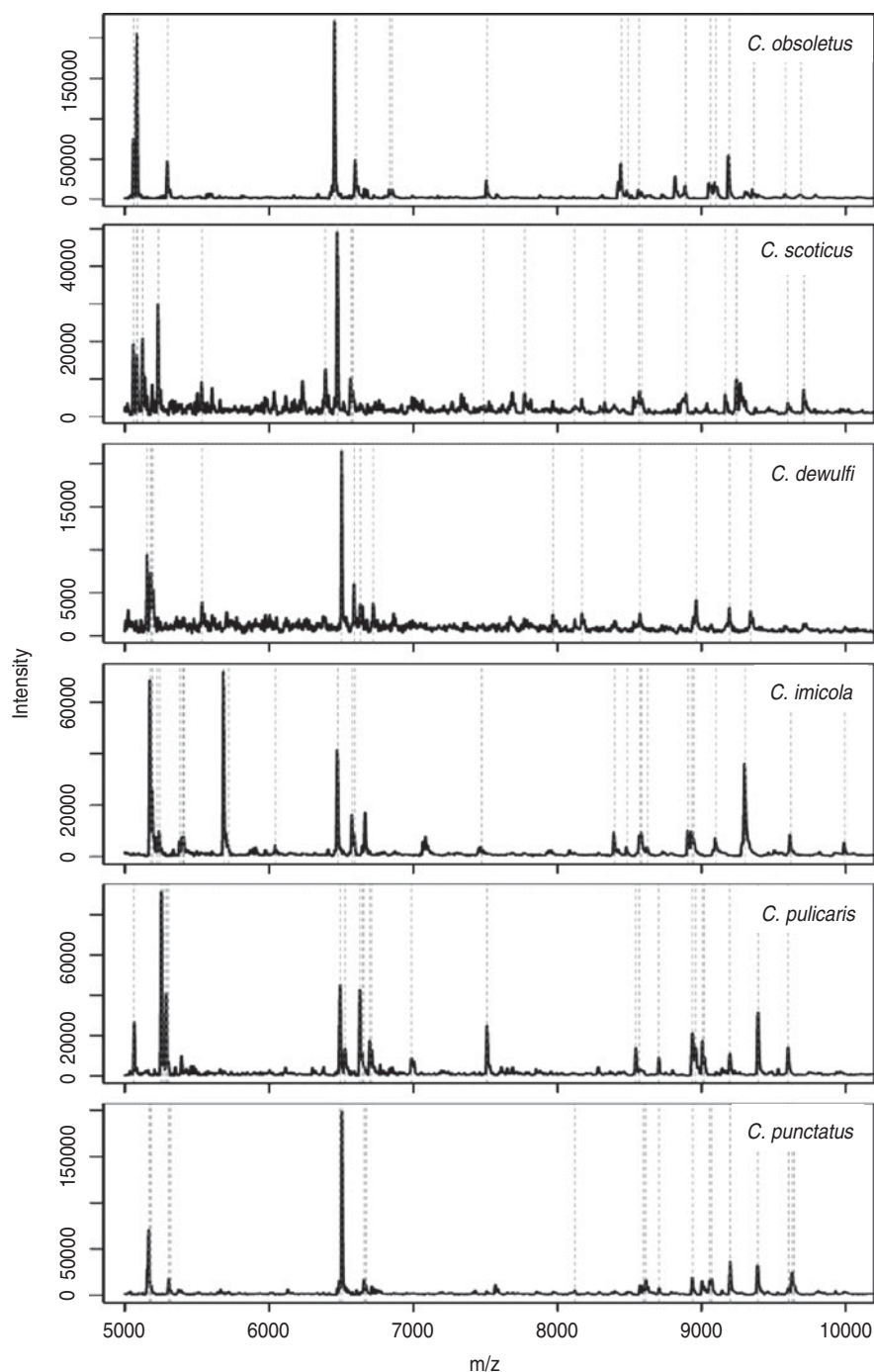
Fig. 1. Matrix-assisted laser desorption/ionization time of flight (MALDI-TOF) mass spectra in the range of 5–10 kDa of field-caught *Culicoides* insects. Biomarker masses as eventually determined by SARAMIS are highlighted by dashed lines.

in one and of *C. grisescens* I in the other read. Finally, all 3 specimens of *C. pallidicornis* (non-target species) did not yield identification neither by automated biomarker identification nor by manual comparison against the spectra of all reference specimens.

Except for 1 specimen with an ambiguous result (see above), all 107 biting midges from the evaluation phase were correctly identified by MALDI-TOF MS, either by automated biomarker identification or by manual full spectra comparison. Initially,

discrepancies of the identification by morphology and MALDI-TOF MS were obtained with 8 individuals. Genetic analyses of these insects were always in accordance with the MALDI-TOF MS results indicating that the morphological identification was inaccurate. Thus, 2 of the 5 specimens morphologically identified as *C. deltus* in fact were *C. pulicaris*, 3 of 5 *C. pulicaris* were *C. punctatus* (2) and *C. lupicaris* (1), respectively, 2 *C. grisescens* turned out to be *C. deltus* and a specimen of *C. lupicaris* was misidentified as *C. punctatus*.
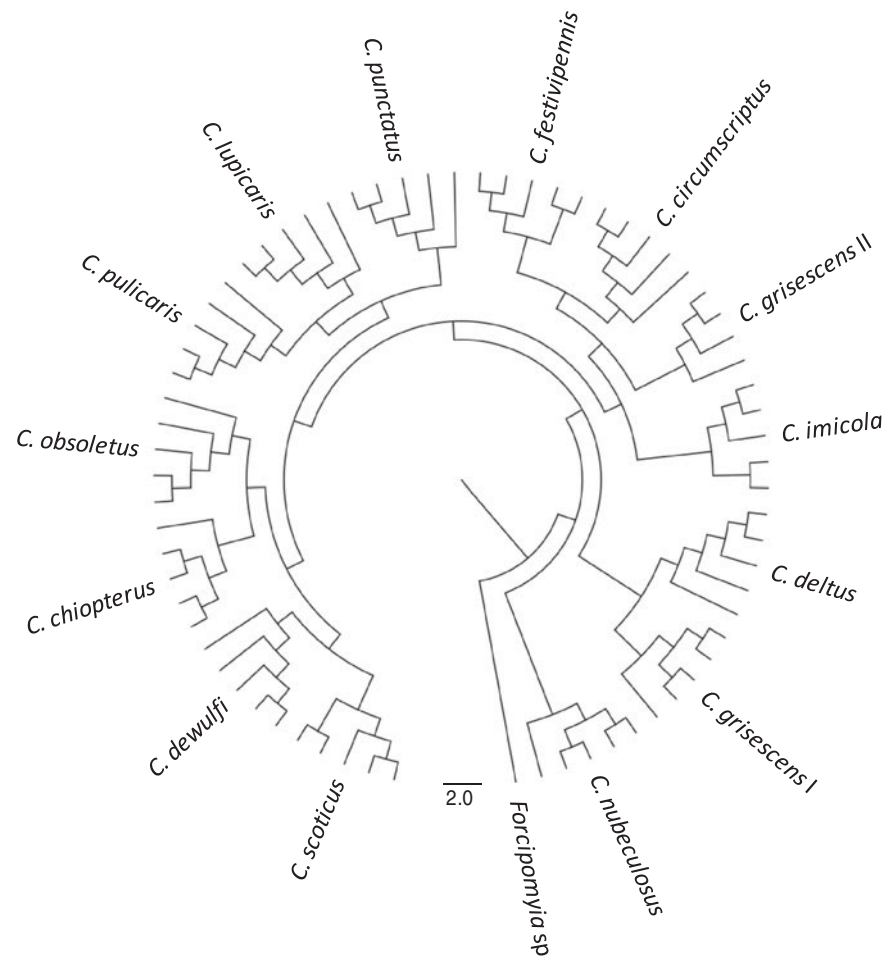
Fig. 2. Dendrogram of matrix-assisted laser desorption/ionization time of flight (MALDI-TOF) mass spectra of thoraxes from (4-) 5 individuals of 13 species of field-caught *Culicoides* biting midges and laboratory-reared *C. nubeculosus*. One individual of a *Forcipomyia* sp. is included as outgroup.

DISCUSSION

MALDI-TOF MS using homogenates of whole insects has been shown in proof-of-principle studies by Campbell (2005) and Perera and colleagues (2005) to be suitable for differentiation of the few species they investigated (fruit flies, aphids). Feltens and colleagues (2010) applied this proteomic approach to discriminate 13 species of *Drosophila*, and the suitability of MALDI-TOF MS to characterize a haematophagous insect (*C. nubeculosus*) has recently been demonstrated (Kaufmann *et al.* 2011). All these investigations relied on laboratory-reared insects. The study presented here is the first evaluation of MALDI-TOF MS for species identification of field-caught haematophagous insects. These included the *Culicoides* species from the Obsoletus and Pulicaris group that are the incriminated biological vectors of the bluetongue virus in Europe, but also *C. imicola*, the main vector in Africa, as well as other European species.

Using a very crude sample preparation method (mechanical grinding in 10% formic acid) it was possible to obtain specific and discrete biomarker mass sets for the 14 *Culicoides* spp. investigated, and only few biomarker masses were shared by different species, even among closely related ones such as *C. grisescens* and a cryptic species (denominated *C. grisescens* II). Thus, it seems feasible that this approach can be expanded to include other *Culicoides* spp. Indeed, *C. pallidicornis* specimens, included in the validation phase as non-target organism, were not recognized by comparing with the spectra of these 14 species, and a unique biomarker mass set could also be calculated for *C. pallidicornis* (not shown).

Also using a minimal sample preparation approach (mechanical insect homogenization in water), the discrimination of sibling species of *Drosophila* by MALDI-TOF MS analyses has been reported (Campbell, 2005). In contrast, when focusing on the analyses of proteins/peptides purified by chromatography after mechanical homogenization under protein-denaturing conditions (6 M urea), the differentiation of closely related and laboratory-reared *Drosophila* spp. (i.e. *D. miranda, D. pseudoobscura*, belonging to the Obscura species group) was not possible (Feltens *et al.* 2010). Between 168 and 390 protein and peptide peaks in the range of 1·8 to

Table 1. MALDI-TOF MS reference biomarker masses for 14 *Culicoides* spp. and *Forcipomyia* sp.

Biomarker mass (Da)	C. chiopterus	C. circumscriptus	C. deltus	C. dewulfi	C. festivipennis	C. grisescens I	C. grisescens II	C. imicola	C. lupicaris	C. nubeculosus	C. obsoletus	C. pulicaris	C. punctatus	C. scoticus	Forcipomyia sp.
2527											X				
2541											X				
2581						X							X		
2616									X						
2624												X			
2646		X													
2801						X									
2831				X						X					
2844															X
2904				X											
3034														X	
3191															X
3200		X													
3227											X				
3234								X						X	
3243												X			
3252	X			X									X		
3258							X								
3271			X												
3312															
3433						X									
3547														X	
3587						X									
3686		X				X									
3716										X					
3755				X								X			
3761		X													
3818		X				X								X	
3830															X
3962						X	X								
4018			X			X								X	
4031														X	
4039														X	
4070				X											
4086				X											
4133														X	
4174		X												X	
4187															X
4247						X									
4261						X									
4287						X									
4304	X														
4392															X
4442					X										
4453								X							
4467				X			X					X			
4483		X													
4495				X											
4514												X			
4530										X					
4548		X									X				
4562															
4579														X	
4592									X	X					
4650								X							
4664				X						X					
4678										X					
4684						X									
4692			X			X			X			X			
4706		X	X												
4716						X	X								
4730				X											
4742									X						X
4793	X														
4814								X					X		
4819													X		
4881														X	
4909				X											
4930														X	
4949				X											
4965				X											
5049	X														
5063											X			X	
5087											X			X	
5127										X				X	
5142										X					
5153				X	X										
5167													X		
5182			X					X							
5192			X				X			X					
5205			X							X					
5218							X								
5231														X	

Table 1 (*Cont.*)

Biomarker mass (Da)	C. chiopterus	C. circumscriptus	C. deltus	C. dewulfi	C. festivipennis	C. grisescens I	C. grisescens II	C. imicola	C. lupicaris	C. nubeculosus	C. obsoletus	C. pulicaris	C. punctatus	C. scoticus	Forcipomyia sp.
5241								X	X						
5254			X												
5271			X				X					X			
5288															
5302	X	X							X			X	X		
5320	X								X			X			
5336															
5348										X					X
5362					X										X
5383								X							
5401								X							
5413															
5442															
5451					X				X						
5466					X										
5494												X			
5535				X										X	
5591			X			X									
5607			X			X				X					
5637			X							X					
5670										X					
5687								X							
5720								X							
5738					X										
5777		X													
5833						X									
5878						X									
5903															X
6003						X				X					
6033															
6044								X							
6133					X										
6247															X
6360										X					
6376												X			X
6389															X
6406		X													
6459											X				
6476					X	X		X						X	X
6492									X			X			
6504	X			X								X	X		
6519							X								
6533			X						X			X			X
6551		X							X						
6569			X											X	
6579								X							
6594			X	X	X	X									X
6602															
6615	X				X										
6630			X	X								X			
6640										X		X			
6653			X			X				X					X
6669					X					X			X		
6685					X										
6699												X			
6716										X					
6728				X											
6743		X													
6762		X													
6797										X					
6818	X														
6837	X					X				X					
6858						X	X			X		X			
6986												X			
7360						X	X								
7373		X													
7415			X												
7434			X							X					
7452															
7475								X							
7490									X					X	
7512												X			
7628							X								
7706															X
7771														X	
7854									X						
7869									X						
7969				X											
8119													X	X	
8169				X											
8207										X					
8305															
8331		X												X	

Table 1 (*Cont.*)

Biomarker mass (Da)	C. chiopterus	C. circumscriptus	C. deltus	C. dewulfi	C. festivipennis	C. grisescens I	C. grisescens II	C. imicola	C. lupicaris	C. nubeculosus	C. obsoletus	C. pulicaris	C. punctatus	C. scoticus	Forcipomyia sp.
8396								X							
8413										X					
8444											X				
8485											X				
8509															
8525															
8542												X			
8554									X						
8567	X	X		X				X			X			X	
8591	X														
8598	X														
8614	X		X				X						X		
8628	X						X								
8668					X										
8706		X	X		X				X				X		
8720							X								
8789										X					
8805										X					
8851		X													
8876															
8889		X												X	
8913		X													
8937	X		X			X	X	X	X			X	X		
8947															
8963		X		X			X					X			
8977		X													
9012												X			
9023							X								
9042			X	X					X						
9055															
9076	X												X		
9099								X			X				
9123															
9166	X													X	
9184									X						
9193	X									X		X		X	
9208										X					
9223	X														
9240			X						X					X	
9256															X
9295								X							
9339															X
9360										X	X				
9371					X										
9387						X		X				X		X	
9408		X			X										
9459			X												
9471															
9495															X
9576						X					X				
9595						X		X	X			X	X	X	
9615		X													
9227													X		
9641															
9693											X				
9708		X			X										
9932															
9992								X							
10299									X						
10630	X														
10763											X				
10790															
10801								X							
10834									X						
10874					X			X	X						
11329						X		X				X	X	X	X
11368															X
12227												X			
12263						X								X	
12273		X													
12353	X													X	
12494															
12520					X										
12723										X					
12825															
12901															X
12954															
13345												X			
13469														X	
16704									X						
19379											X				
19549								X							
Total	**25**	**27**	**28**	**31**	**24**	**30**	**25**	**34**	**34**	**33**	**26**	**36**	**26**	**38**	**28**

15 kDa, (average 236 peaks per spectrum) were identified in that study (present study: 50 to 169 peaks in the mass range of 2 to 20 kDa with an average peak number of 99), but spectral heterogeneity within species and a limited number of species-specific peaks were noted. Feltens and colleagues (2010) identified 18 potential protein markers by electrospray ionization MS/MS, revealing that most of them stem from muscle tissue like myosin heavy chain, troponin, tropomyosin, flightin or from mitochondria-like ATP synthase subunits alpha and beta. In our study, mainly the muscle-containing thorax (and the head) were used for protein extraction because potential bloodmeal traces in the abdomen were shown to impair the MALDI spectra quality (Kaufmann *et al*. 2011). Thus, the spectra that we generated have a lower average datacount as compared to the study of Feltens and colleagues (2010) but they cover most of the potential discriminative and reproducible muscle tissue and mitochondrial proteins.

Although no emphasis was laid on phylogenetic analyses, the dendrogram shown in Fig. 2 reasonably reflects the current comprehension of *Culicoides* kinship. Thus, the species of the Obsoletus group (*C*. *chiopterus*, *C*. *obsoletus*, *C*. *scoticus*) cluster together, also including *C*. *dewulfi*, which recently was suggested not to be a member of this group, according to genetic analyses at 3 loci (Schwenkenbecher *et al*. 2009). The recognized species of the Pulicaris group are grouped in 2 clusters (*C*. *lupicaris*, *C*. *pulicaris*, *C*. *punctatus*; *C*. *deltus*, *C*. *grisescens* I). Interestingly, the cryptic species denominated *C*. *grisescens* II, which is morphologically indistinguishable from the type species *C*. *grisescens* I but differs by approximately 10% at the barcode locus mt COX I gene (Wenk *et al*. 2011), is not part of one of the Pulicaris group clusters. Thus, more in-depth investigations using MALDI-TOF MS, as described (Feltens *et al*. 2010), might be of value to address taxonomic and phylogenetic questions in the genus *Culicoides*.

In a previous study with *C*. *nubeculosus* (Kaufmann *et al*. 2011), male and female specimens had very similar spectra allowing derivation of a set of biomarker masses valid for both sexes. Similar spectra of males and females of *D*. *melanogaster* using a crude sample preparation were also reported (Campbell, 2005) whereas a differentiation of the two sexes of this fruit fly species was achieved based on mass analyses of purified proteins/peptides (Feltens *et al*. 2010). In the present study, individuals of both sexes of all but 2 species were included either in the reference specimens or in those used in the validation study, and no discrimination between males and females was observed (data not shown). Only females were available from *C*. *deltus* and *C*. *imicola*. However, this is of minor importance because only females are haematophagous and thus

act as vectors and because the commonly used UV-traps primarily capture this gender. Further, only males of the rare species *C*. *pallidicornis*, which was used as the non-target species in the validation phase, were available. Considering the findings with the other species. i.e. gender does not matter with regard to identification by MALDI-TOF patterns, the eventually calculated biomarker mass sets of these male specimens most probably are also valid to accurately identify females of this species.

The insects used in this work were stored in 70% EtOH which is considered suitable for subsequent analyses by MALDI-TOF MS (Campbell, 2005; Feltens *et al*. 2010), and it was experimentally proven that such a storage up to around 3 months (102 days) yielded stable mass spectrum results (Kaufmann *et al*. 2011). In the present study, it was observed that freshly caught species generated more data counts (peaks) than species which were kept for a longer period (> 3 months) in 70% ethanol after trapping (data not shown). Nevertheless, the biomarker mass sets compiled from the data of the reference specimens allowed reliable identification in the blind study midges kept over 2 years in 70% ethanol (at 4° C) (not shown).

Reproducibility and accuracy of the references were tested in a blind study of 111 morphologically identified biting midges mainly caught in the field all over Switzerland. Thus, a correct and unambiguous result was obtained in 110 of the 111 insects tested. Initial discrepancies between morphology on one side and mass analysis/genotyping on the other side were encountered in 8 individuals. All these morphologically misidentified specimens belong to species from the Pulicaris group which are notoriously hard to distinguish with certainty as they are known to display natural variation in morphology as well as overlapping characters. Thus, the accuracy of the morphological characters employed in identification keys might be re-evaluated according to MALDI-TOF MS or genetic identifications. Further, long storage in ethanol might have blurred morphological features such as the wing patterns. Morphological misidentifications as identified by genetic data also were observed when selecting specimens for creating the reference database, but these individuals were excluded.

The application of MALDI-TOF MS for the identification of insects has just begun, and further investigations will generate more biomarker mass sets which preferably should be united in a single database. For *Culicoides*, a relatively low number of very discrete biomarker masses were identified (24–38) as compared to bacteria which usually are characterized by 30–50 such masses. In addition, it is not uncommon that more than 1 biomarker mass set is available in databases for a single species, e.g. there are >40 sets currently available for *Staphylococcus aureus* in the SARAMIS library of Anagnostec.

Larger datasets with *Culicoides* spp. might reveal different sets for populations of a species from different origins, although our limited data revealed that individuals of 3 species (*C. chiopterus*, *C. festivipennis*, *C. obsoletus*) that were collected at either site of the Alpine crest could be integrated in single sets each.

In conclusion, the presented results demonstrate that protein profiling by MALDI-TOF can be used as an alternative, quick and inexpensive tool to accurately identify *Culicoides* biting midges collected in the field, including cryptic species. MALDI-TOF can be used complimentarily to other identification techniques (morphology, genetics) as only part of the insect need to be investigated. This approach has the potential to become the method of choice for a centralized, robust and high throughput screening of midge populations in connection with the surveillance of eventually emerging midge-transmitted agents.

REFERENCES

Boorman, J. (1974). The maintenance of laboratory colonies of *Culicoides variipennis* (Coq.), *C. nubeculosus* (Mg.) and *C. riethi* Kiff. (Diptera, Ceratopogonidae). *Bulletin of Entomological Research* **64**, 371–377.

Campbell, P. M. (2005). Species differentiation of insects and other multicellular organisms using matrix-assisted laser desorption/ionization time of flight mass spectrometry protein profiling. *Systematic Entomology* **30**, 186–190.

Caprioli, G., Cristalli, G., Ragazzi, E., Molin, L., Ricciutelli, M., Sagratini, G., Seraglia, R., Zuo, Y. and Vittori, S. (2010). A preliminary matrix-assisted laser desorption/ionization time-of-flight approach for the characterization of Italian lentil varieties. *Rapid Communications in Mass Spectrometry* **24**, 2843–2848.

Carpenter, S., Wilson, A. and Mellor, P. S. (2009). *Culicoides* and the emergence of bluetongue virus in northern Europe. *Trends in Microbiology* **17**, 172–178.

Cêtre-Sossah, C., Baldet, T., Delécolle, J. C., Mathieu, B., Perrin, A., Grillet, C. and 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. *Veterinary Research* **35**, 325–337.

Chaignat, V., Worwa, G., Scherrer, N., Hilbe, M., Ehrensperger, F., Batten, C., Cortyen, M., Hofmann, M. and Thuer, B. (2009). Toggenburg Orbivirus, a new bluetongue virus: initial detection, first observations in field and experimental infection of goats and sheep. *Veterinary Microbiology* **138**, 11–19.

Dallas, J. F., Cruickshank, R. H., Linton, Y. M., Nolan, D. V., Patakakis, M., Braverman, Y., Capela, R., Capela, M., Pena, I., Meiswinkel, R., Ortega, M. D., Baylis, M., Mellor, P. S. and

Mordue (Luntz), A. J. (2003). Phylogenetic status and matrilineal structure of the biting midge, *Culicoides imicola*, in Portugal, Rhodes and Israel. *Medical and Veterinary Entomology* **17**, 379–387.

Delécolle, J.-C. (1985). Nouvelle contribution à l'étude systématique et iconographique des espèces du genre *Culicoides* (Diptera: Ceratopogonidae) du Nord-Est de la France. Université Louis Pasteur, Strasbourg, France.

Feltens, R., Gorner, R., Kalkhof, S., Groger-Arndt, H. and von Bergen, M. (2010). Discrimination of different species from the genus *Drosophila* by intact protein profiling using matrix-assisted laser desorption ionization mass spectrometry. *BMC Evolutionary Biology* **10**, 95.

Goffredo, M. and Meiswinkel, R. (2004). Entomological surveillance of bluetongue in Italy: methods of capture, catch analysis and identification of *Culicoides* biting midges. *Veterinaria Italiana* **40**, 260–265.

Hellberg, W., Mellor, P. S., Torsteinsdottir, S. and Marti, E. (2009). Insect bite hypersensitivity in the horse: comparison of IgE-binding proteins in salivary gland extracts from *Simulium vittatum* and *Culicoides nubeculosus*. *Veterinary Immunology and Immunopathology* **132**, 62–67.

Kaufmann, C., Schaffner, F. and Mathis, A. (2009). Monitoring of biting midges (*Culicoides* spp.), the potential vectors of the bluetongue virus the 12 climatic regions of Switzerland. *Schweizer Archiv für Tierheilkunde* **5**, 205–513.

Kaufmann, C., Ziegler, D., Schaffner, F., Carpenter, S., Pflüger, V. and Mathis, A. (2011). Evaluation of matrix-assisted laser desorption/ionization time of flight mass spectrometry for characterization of *Culicoides nubeculosus* biting midges. *Medical & Veterinary Entomology* **25**, 32–38.

Mazzeo, M. F., Giulio, B. D., Guerriero, G., Ciarcia, G., Malorni, A., Russo, G. L. and Siciliano, R. A. (2008). Fish authentication by MALDI-TOF mass spectrometry. *Journal of Agricultural and Food Chemistry* **56**, 11071–11076.

Meiswinkel, R., Baldet, T., de Deken, R., Takken, W., Delécolle, J.-C. and Mellor, P. S. (2008). The 2006 outbreak of bluetongue in northern Europe – the entomological perspective. *Preventive Veterinary Medicine* **87**, 55–63.

Meiswinkel, R., van Rijn, P., Leijs, P. and Goffredo, M. (2007). Potential new *Culicoides* vector of bluetongue virus in northern Europe. *Veterinary Record* **161**, 564–565.

Mellmann, A., Bimet, F., Bizet, C., Borovskaya, A. D., Drake, R. R., Eigner, U., Fahr, A. M., He, Y., Ilina, E. N., Kostrzewa, M., Maier, T., Mancinelli, L., Moussaoui, W., Prevost, G., Putignani, L., Seachord, C. L., Tang, Y. W. and Harmsen, D. (2009). High inter-laboratory reproducibility of matrix-assisted laser desorption ionization-time of flight mass spectrometry-based species identification of nonfermenting bacteria. *Journal of Clinical Microbiology* **47**, 3732–3734.

Mellor, P. S., Boorman, J. and Baylis, M. (2000). *Culicoides* biting midges: their role as arbovirus vectors. *Annual Review of Entomology* **45**, 307–340.

Mellor, P. S. and Hamblin, C. (2004). African horse sickness. *Veterinary Research* **35**, 445–466.

Paweska, J. T., Venter, G. J. and Hamblin, C. (2005). A comparison of the susceptibility of *Culicoides imicola* and *C. bolitinos* to oral infection with eight serotypes of epizootic haemorrhagic disease virus. *Medical and Veterinary Entomology* **19**, 200–207.

Perera, M. R., Vargas, R. D. F. and Jones, M. G. K. (2005). Identification of aphid species using protein profiling and matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *Entomologia Experimentalis et Applicata* **117**, 243–247.

Santos, C., Paterson, R. R. M., Venancio, A. and Lima, N. (2010). Filamentous fungal characterizations by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *Journal of Applied Microbiology* **108**, 375–385.

Sauer, S. and Kliem, M. (2010). Mass spectrometry tools for the classification and identification of bacteria. *Nature Reviews Microbiology* **8**, 74–82.

Schwenkenbecher, J. M., Mordue Luntz, A. J. and Piertney, S. B. (2009). Phylogenetic analysis indicates that *Culicoides dewulfi* should not be considered part of the *Culicoides obsoletus* complex. *Bulletin of Entomological Research* **99**, 371–375.

Sloet van Oldruitenborgh-Oosterbaan, M. M., van Poppel, M., de Raat, I. J., van den Boom, R. and Savelkoul, H. F. (2009). Intradermal testing of horses with and without insect bite hypersensitivity in The Netherlands using an extract of native *Culicoides* species. *Veterinary Dermatology* **20**, 607–614.

Stephan, R., Ziegler, D., Pflüger, V., Vogel, G. and Lehner, A. (2010). Rapid genus- and species-specific identification of *Cronobacter* spp. by matrix-assisted laser desorption ionization-time of flight mass spectrometry. *Journal of Clinical Microbiology* **48**, 2846–2851.

Stevenson, L. G., Drake, S. K., Shea, Y. R., Zelazny, A. M. and Murray, P. R. (2010). Evaluation of matrix-assisted laser desorption ionization-time of flight mass spectrometry for identification of clinically important yeast species. *Journal of Clinical Microbiology* **48**, 3482–3486.

van Veen, S. Q., Claas, E. C. J. and Kuijper, E. J. (2010). High-throughput identification of bacteria and yeast by matrix-assisted laser desorption ionization-time of flight mass spectrometry in conventional medical microbiology laboratories. *Journal of Clinical Microbiology* **48**, 900–907.

Venter, G. J. and Meiswinkel, R. (1994). The virtual absence of *Culicoides imicola* (Diptera: Ceratopogonidae) in a light-trap survey of the colder, high-lying area of the eastern Orange Free State, South Africa, and implications for the transmission of arboviruses. *Onderstepoort Journal of Veterinary Research* **61**, 327–340.

Wenk, C. E., Kaufmann, C., Schaffner, F. and Mathis, A. (2011). Molecular characterisation of Swiss Ceratopogonidae (Diptera) and evaluation of real-time PCR assays for the identification of *Culicoides* biting midges. *Veterinary Parasitology* (in Press).