



**Mayfly and fish species identification and sex determination in bleak (*Alburnus alburnus*)
by MALDI-TOF mass spectrometry**

^{1,2}Maasz, G. *, ³Takács, P., ⁴Boda, P., ⁴Varbiro, G. ⁺, ¹Pirger, Z. ⁺

¹MTA-ÖK BLI NAP_B Adaptive Neuroethology, Department of Experimental Zoology, Balaton Limnological Institute, MTA Centre for Ecological Research, 8237 Tihany, Hungary

²Department of Analytical Biochemistry, Institute of Biochemistry and Medical Chemistry, University of Pecs, 7624 Pecs, Hungary

³Department of Hydrozoology, Balaton Limnological Institute, MTA Centre for Ecological Research, 8237 Tihany, Hungary.

⁴Department of Tisza River Research, MTA Centre for Ecological Research, 4032 Debrecen, Hungary

*Address correspondence to Dr. Gabor Maasz, MTA-ÖK BLI NAP_B Adaptive Neuroethology, Department of Experimental Zoology, Balaton Limnological Institute, MTA Centre for Ecological Research, 8237, Tihany, Hungary

E-mail: maasz.gabor@okologia.mta.hu

+ These authors are equal

Abstract

Besides food quality control of fish or cephalopods, the novel mass spectrometry (MS) approaches could be effective and beneficial methods for the investigation of biodiversity in ecological research. Our aims were to verify the applicability of MALDI-TOF MS in the rapid identification of closely related species, and to further develop it for sex determination in phenotypically similar fish focusing on the low mass range. For MALDI-TOF MS spectra analysis, ClinProTools software was applied, but our observed classification was also confirmed by Self Organizing Map. For verifying the wide applicability of the method, brains from invertebrate and vertebrate species were used in order to detect the species related markers from two mayflies and eight fish as well as sex-related markers within bleak.

Seven *Ephemera* larvae and sixty-one fish species related markers were observed and nineteen sex-related markers were identified in bleak. Similar patterns were observed between the individuals within one species. In contrast, there were markedly diverse patterns between the different species and sexes visualized by SOMs. Two different *Ephemera* species and male or female fish were identified with 100% accuracy. The various fish species were classified into 8 species with a high level of accuracy (96.2%). Based on MS data, dendrogram was generated from different fish species by using ClinProTools software. This MS-based dendrogram shows relatively high correspondence with the phylogenetic relationships of both the studied species and orders.

In summary, MALDI-TOF MS provides a cheap, reliable, sensitive and fast identification tool for researchers in the case of closely related species using mass spectra acquired in a low mass range to define specific molecular profiles. Moreover, we presented evidence for the first time for determination of sex within one fish species by using this method. We conclude that it is a powerful tool that can revolutionize ecological and environmental research.

Keywords: Species identification; Mayfly; Sex identification; Fish; MALDI-TOF MS; Biotyping

1. Introduction

During the last decades, several studies focused on different levels of biodiversity ranging from cells to individuals. The identification of phylogenetically conserved markers enabling detection of different animal species and reconstruction of their evolutionary segregation is one of the current topics in biological diversity research. Recently, parallel with classical morphological (morphometric-, meristic-, descriptive characters) (Fischer, 2013) and DNA-based (barcodes) approaches (Radulovici et al., 2010; Zemlak et al., 2009) novel, mass spectrometry (MS) based methods give a useful, rapid, accurate, reproducible and relatively cheap tool for identification of species (Kim et al., 2015; Lee et al., 2015; Stephan et al., 2014; Volta et al., 2012).

Molecular tools can help to highlight the presence of potential cryptic species and identify where there are misidentified species due to the use of classical morphological approaches, and to provide new characteristics that can be used for correct identification. Traditional morphological approaches are still important and can be advantageous in some cases, but they are less accurate and time-consuming (Fischer, 2013; Takacs et al., 2016). Therefore, there has recently been developed a more accurate, reliable and relatively rapid sequence-based methods for quantification of biodiversity in the fields of molecular ecology and genomics. For molecular-based identification of a cryptic species, several polymerase chain reaction (PCR)-based tests have been developed including single and multiplexed assays both in conventional and real-time PCR formats (Balczun et al., 2009; Cetre-Sossah et al., 2004; Cetre-Sossah et al., 2008; Nolan et al., 2007; Pages and Monteys, 2005; Stephan et al., 2009). However, exact identification of an individuum is practically impossible using real-time PCR because of the limited availability of specific primer probes (Lee et al., 2015). The high demand for low-cost sequencing has driven the development of high-throughput sequencing, which also goes by the term Next-Generation Sequencing (NGS). These recent technologies allow us to sequence DNA and RNA 200 times quicker and significantly cheaper than the previously used Sanger sequencing method and revolutionised the study of genomics and molecular biology (Harismendy et al., 2009; von Bubnoff, 2008). Cardoso identified seven impediments to invertebrate conservation but similar problems may also affect non-invertebrate systematics (Cardoso et al., 2011). The NGS barcodes may help to overcome several of these impediments (Linnean, Wallacean shortfall) and barcodes will make it cheaper to find specimens for cryptic and unidentified species and provide information about distribution and abundance when samples from different sites are analysed and compared.

Beside the advantages provided by NGS barcodes, the matrix-assisted laser desorption ionization-time of flight mass spectrometry (MALDI-TOF MS) based method have also been established as a user-friendly species identification tool. The MALDI-TOF MS is useful for interpretation of the information encoded in genomes, complementing DNA-based approaches (Aebersold and Mann, 2003). MALDI-TOF MS has emerged as a promising tool for metabolite, peptide and protein identification and proteomic phenotyping of various invertebrate and vertebrate species due to its rapid turn-around time, high method sensitivity (0.5-1.5 µl amount of sample), high throughput and relatively low additional cost (Barbuddhe et al., 2008; Ilina et al., 2009; Zabbe et al., 2015). Using MALDI-TOF MS, phylogenetically conserved molecular markers (e.g.

metabolites, peptides, proteins) can also be identified as well as which is useful in the investigation of biological diversity and in the molecular taxonomy. It has already been used to identify bacteria (Barbuddhe et al., 2008; Ilina et al., 2009; Zabbe et al., 2015), microalgae (Emami et al., 2015; Lee et al., 2015) and even higher eukaryotes including nematodes (Perera et al., 2005a), molluscs (Lopez et al., 2005; Stephan et al., 2014), arthropods (Feltens et al., 2010; Kaufmann et al., 2011; Perera et al., 2005b; Salla and Murray., 2013), and fish (Mazzeo et al., 2008; Volta et al., 2012). For clinical and environmental microbiology of MALDI-TOF MS based bacterial and yeast identification there are two widely used system: Bruker BiotyperTM (Bruker Daltonics, Germany) and VITEK^R MS (bioMerieux, France) including databases (Nagy et al., 2009; Nomura, 2015; Patel, 2013; Seng et al., 2009). More recently, the MALDI-TOF MS based Biotyper system has also been applied for identification of animal pathogens from infected animals (Alispahic et al., 2012; Elbehiry et al., 2016; Goncalves et al., 2014; Kim et al., 2015). However, the applicability of these complex systems might be restricted and/or tricky due to the lack of comprehensive databases (Nomura, 2015). Based on recent data from food chemistry, it was also successfully applied for quality control of myofibrillar proteins in fish fillet (Feng et al., 2016; Feng et al., 2017a; Feng et al., 2017b).

Our aims were to verify the applicability of MALDI-TOF MS method in the fast and reliable identification of closely related species, and to further develop it for sex determination focusing on the low molecular range in both cases. For verification of the large applicability of the MS method in species identification and sex determination, both invertebrate and vertebrate species were applied. From invertebrates, two mayfly species were used from *Ephemera* genus, the common mayfly (*Ephemera vulgate*) and striped mayfly (*Ephemera lineata*). In general, morphological identification of *Ephemera* larvae is difficult but the authentication of the common and stripped mayfly is relatively easy. From vertebrates, eight fish species were collected representing phylogenetically distant and close species from Anguilliformes, Perciformes and Cypriniformes orders. Therefore, both vertebrate and invertebrate species were utilized for validation of MALDI-TOF MS results in our study. Brain samples were used in order to detect the species related molecular markers from mayflies and fish (1) as well as sex-related markers within bleak (*Alburnus alburnus*) (2). For MALDI-TOF MS spectra analysis a ClinProTools clustering software and a Self Organizing Map (SOM) were used. Furthermore, we investigated the usefulness of MALDI-TOF MS technique and described its advantages and disadvantages to promote research in the future.

2. Materials and Methods

2.1 Animals

2.1.1 Mayfly larvae - Larvae of two different but morphologically close mayfly species were used. Common mayfly (*Ephemera vulgate*) and striped mayfly (*Ephemera lineata*) can be identified by examining their abdomen under a stereo microscope. These two species were chosen because they usually occur on the same site and the morphological identification was able to be made in a living state and did not need preparation. *E. lineata* has three dark lines on the abdominal terga of 7, 8, and 9 while there is a distinct dark mark on segment 7, 8 and 9 in *E. vulgate*. (see Supplementary data, Figure S1) (Bauernfeind and Soldán, 2012; Elliott et al., 1988). To avoid site specificity, we selected a sample site where both species had stable populations. Both *E. vulgate* (n=14), and *E. lineata* (n=10) were collected from the River Ér (N47.279372 E21.805831) by using hand nets and transported to the lab in water tanks in March 2016. Individuals were identified based on their morphological motifs and after decapitation brains were removed and refrigerated at -80 °C until sample preparation.

2.1.2 Fish - The adult specimens of bleak (*Alburnus alburnus*, n=5-5 male and female), roach (*Rutilus rutilus*, n=10), topmouth gudgeon (*Pseudorasbora parva*, n=6), pumpkinseed (*Lepomis gibbosus*, n=6), European eel (*Anguilla Anguilla*, n=2), common bream (*Abramis brama*, n=6), gibel carp (*Carassius gibelio*, n=9) and European perch (*Perca fluviatilis*, n=6) were used in this study. Samples were collected (collection permit: VE-I-001/01890-9/2013) from Lake Balaton (N46.914315 E17.893004) by electrofishing in March 2016. Collected individuals were transported to the lab in a water tank. Animals were anaesthetized with clove oil and killed by decapitation. Their whole brains were removed while on ice and they were stored at -80 °C until sample preparation. Sex of bleaks was determined based on gonads (testis or ovary) using a stereo microscope (see Supplementary data, Figure S2).

2.2 Sample preparation for MALDI-TOF MS

We aimed to use a wide range of metabolites or peptides to provide evidence of the applicability of MALDI-TOF MS, therefore, the brain tissue was prepared from test animals. Brain samples from both mayflies and fish (0.4-0.9 mg tissue/mayfly and 20-50 mg tissue/fish) were homogenized in 200 µL 0.1% TFA by a glass homogenizer. Cells were dissociated for 6×10 sec with a high energy UIS250V ultrasonicator (Hielsher Ultrasound Technology, Germany) applying ice cooling between cycles. Samples were vortexed and centrifuged at 8,000 g for 10 min. The supernatant was transferred to centrifugal filter devices (Amicon Ultra - 0.5 mL, 10,000 NMWL, Merck Millipore Ltd., Germany) and the proteins were separated at 8,000 g for 40 min. One-hundred µl chloroform was used to remove the lipid section that provided a high grade of purity which led to a good degree of ionization. The mixture was gently shaken for several seconds and phase separation was performed by centrifugation at 3,000 g for 5 min. The organic phase was removed and the residual chloroform was evaporated using Speed Vac Concentrator (Concentrator Plus, Eppendorf, Austrian). The protein exclusion and delipidation steps resulted in a high purity of peptide and/or metabolite mixture. Samples were stored at -80 °C until further processing.

2.3 MALDI-TOF MS analysis

The samples were enriched on an Anchor chip target plate (MTPAnchorChip™384 T F, Bruker Daltonics, Germany) by using 1.5 µl sample solution. Thereafter, 1 µl freshly prepared 0.7 mg/ml matrix solution was added before each measurement. The matrix solution was made by dissolving α -cyano-4-hydroxycinnamic acid (CHCA) in acetonitrile/0.1 % TFA (1/2, v/v). Samples were analysed by using Autoflex Speed MALDI-TOF/TOF MS (Bruker Daltonics, Germany) operated in reflector mode for peptide mass fingerprinting (PMF). The FlexControl 3.4 software was used to control the instrument. The accelerating voltage was set to 20.00 kV. The instrument uses a 337 nm nitrogen laser (model MNL-205MC, Lasertechnik Berlin GmbH., Germany). Laser repetition rate was set at 50Hz and the global attenuator offset was 17%. The pulsed ion extraction delay was 120 ns. The reflector detector gain voltage was 1800V. External calibration was performed in each case using peptide calibration standards (#206195, Bruker Daltonics, Germany). In reflector mode, the mass resolutions were between 10,000 – 22,000 FWHM simultaneously across a wide m/z range. Reflector mass accuracy was greater than 20 ppm with external calibration. The masses were acquired in the range of m/z 500-5000. Each spectrum of individuals was produced by accumulating data from 1000 consecutive laser shots.

2.4 Data analysis

Potential species or sex-related markers of different groups were chosen manually from MALDI-TOF MS spectra, based on whether or not it presented in the spectra. The manually selected potential markers were statistically evaluated by Matlab peak by peak (see Supplementary data, Table S1). Individual mass spectra from mayflies and fish species were also verified statistically by ClinProTools 3.0 (Bruker Daltonics, Germany) clustering software (see Supplementary data, Table S2-4). Recalibration, spectral alignment, peak normalization, peak detection and peak area calculation of mass spectra were carried out automatically by ClinProTools. A logistic regression analysis was performed to identify the significant predictive peaks on the basis of the normalized peak areas. To demonstrate the predictive value of the identified species and sex-related marker, T-tests and one-way ANOVA tests were used for parametric statistical analysis of the different sample groups. For dendrogram generation, the ClinProTools software was used. Phylogenetic dendrogram presentation was constructed based on whole mass spectra from different fish species. The classification ration (cr) represents the number of correctly classified animals / total number of animals in the ClinProTools analysis. Generally, the ClinProTools software operates with a large number of samples within groups, therefore to confirm our results we used the SOM (Matlab SOM Toolbox), which is an independent mathematical analysis tool and followed the application suggested by Goodwin (Goodwin et al., 2015). The individual mass spectra data was used as an initial matrix for the SOM analysis. In the case of mayfly species identification, the range was between m/z 1200 to 1300, in the case of fish species identification the full MS range (m/z 500-5000) was used, while in fish sex identification the range was chosen between m/z 1300 to 2000.

3. Results

3.1 Species identification

3.1.1 Mayfly larvae

In this study, two morphologically identified *Ephemera* species were analysed by MALDI-TOF MS. Species related markers were observed by m/z 1252.0; 1789.0; 2697.8; 3166.0 to *E. vulgata* and m/z 1297.6; 1336.7; 3281.0 to *E. lineata* from MS spectra selected manually (Table 1). From among these, m/z 1252.0, 2697.8 and 3281.0 peaks were supported as significant differences between two larvae species by automated ClinProTools analysis (see Supplementary data, Table S2). Representative mass spectra from two species were present in Figure S3A including two different markers by m/z 1297.6 (*E. lineata*) and 1252.0 (*E. vulgata*). Figure S3B shows a graphical illustration of m/z 1252.0 in a smaller range (m/z 1232-1276), which was one of the differences between the two species. Figure S3C illustrates a representative dissociation based on m/z 1252.0 and 1297.6 markers using ClinProTools. Similar patterns were observed among the individuals within one species, in contrast, there were markedly diverse patterns between the two different species visualized by the component planes of the SOMs (see supplementary data, Figure S3D). The redder spots indicate the higher intensities of the markers in the range of m/z 1200 to 1300. Based on our MALDI-TOF MS analysis, two different *Ephemera* species were classified correctly as *E. vulgata* (cr: 14/14, 100%) and *E. lineata* (cr: 10/10, 100%).

3.1.2 Fish

Samples of eight freshwater fish species were analyzed by MALDI-TOF MS. Numerous species related markers were identified manually from MS spectra of fish brain homogenates for the species classification (Table 2). From among these, m/z 1020.6 (*P. parva*); 1507.0, 1578.1 (*A. anguilla*); 777.4, 949.6, 1112.6, and 1765.0 (*C. auratus gibelio*); 891.5, 936.6, and 979.6 (*P. fluviatilis*); 594.3, 1252.6 (*L. gibbosus*) peaks were supported as significant differences between fish species by automated ClinProTools analysis (see Supplementary data, Table S3). Based on species related markers, multivariate cluster analysis was made by ClinProTools among the studied fish species. Four cluster segregation was observed on the dendrogram (Fig. 1), where species were classified into three bigger groups according to their kind of relationship, as follows: Anguilliformes (*A. anguilla*), Perciformes (*L. gibbosus* and *P. fluviatilis*) and Cypriniformes (*A. alburnus*, *R. rutilus*, *P. parva*, *A. brama*, and *C. auratus gibelio*). The cluster formation was derived from typical order related markers, such as m/z 750.5, 759.5, 837.5, and 1651.9 to Perciformes (Cluster II) as well as m/z 1263.7, 1734.9, 2088.1, and 2127.3 to Cypriniformes (Cluster III and IV). Additionally, within Cypriniformes order, m/z 1901.1 was observed, as typical marker for species of Cluster III (*C. auratus gibelio*, *R. rutilus*, and *A. brama*). Species related patterns could be visualized by the component planes of the SOMs illustrating all species in the same range (Fig. 2). Similar patterns with a repetitive motif(s) were observed in the individuals within one species, however, the optimal distinction was observed between the species related pattern. Similarly to mayflies, in fish the redder spots indicate the higher intensities of the markers in the range of m/z 500 to 5000 in all cases. Using MALDI-TOF MS and two different analytical software, the various

fish species were properly classified as *A. anguilla* (cr: 2/2, 100%), *P. fluviatilis* (cr: 3/3 100%), *L. gibbosus* (cr: 6/6, 100%), *R. rutilus* (cr: 8/10, 80%), *A. brama* (cr: 6/6, 100%), *C. auratus gibelio* (cr: 9/9, 100%), *P. parva* (cr: 6/6 100%) and *A. alburnus* (cr: 10/10, 100%) species.

3.2 Sex identification

Morphologically identified male and female *A. alburnus* were applied for this investigation. Based on our observation, Table 3 shows ten sex-related markers representing male and nine specific masses indicating female *A. alburnus*. Sex-related markers were identified manually from MS spectra of male and female bleak brain homogenates. From among these, m/z 960.8, 1062.7, 1131.6, 1178.8, 1201.8, 1225.8, 1611.1, and 2088.5 (male); 945.8, 1749.1, 4821.1, 4893.2, and 4931.0 (female) peaks were supported as significant differences between two sexes by automated ClinProTools analysis (see Supplementary data, Table S4). Representative mass spectra from the two sexes was presented in Figure 3A including two different specific markers by m/z 1062.7 (male) and 808.6 (female). Figure 3B shows a simple graphical illustration of m/z 1062.7 in the lower range (m/z 1057.5-1072.5), which was one of the differences between the sexes. The distribution of m/z 808.6 female and 1062.7 male specific markers were illustrated in Figure 3C. Similar patterns were observed among the individuals within one sex but markedly diverse patterns could be visualized between the sexes by the component planes of the SOMs (Fig. 3D). The redder spots indicate the higher intensities of the markers in the range of m/z 1300.0 to 2000. Based on our test, *A. alburnus* sexes were correctly classified as male (cr: 5/5, 100%) and female (cr: 5/5, 100%) from a mixed sex population by MS. The MS results were confirmed by morphological identification performed on the intact animals before sample preparation (see Supplementary data, Figure S2).

4. Discussion

In previous studies the MALDI-TOF MS has already been used for identification of molecular markers both in invertebrates and vertebrates (Feltens et al., 2010; Kaufmann et al., 2011; Mazzeo et al., 2008; Perera et al., 2005b; Volta et al., 2012; Uhlmann et al., 2014). In this study, we confirmed further that this method is suitable for the rapid and reliable identification of species related molecular markers using mayfly and fish brain samples. Moreover, we presented evidence for the first time for the determination of sex by using MALDI-TOF MS. In contrast to other studies (Emami et al., 2015; Kaufmann et al., 2011; Kim et al., 2015; Lee et al., 2015; Mazzeo et al., 2008; Volta et al., 2012) where proteins (m/z 4000-20000) were analysed, our investigation focused on the range of metabolites or peptides (m/z 500-5000).

4.1 Identification of mayfly larvae and fish species by MALDI-TOF MS

Influenced by different environmental factors, the morphological identification of *Ephemera* larvae is generally difficult (e.g. due to adaptation) (Elliott et al., 1988). As our results proved, MALDI-TOF MS provides an alternative method, for the fast and reliable identification of *E. lineata* and *E. vulgate*, complementing conventional morphological analysis. In our experiments,

the two different *Ephemera* species were identified with 100% accuracy (24 out of 24 individuals). Similarly to earlier publications concerning the Aphididae family, *Drosophila* and *Culicoides* genus (Feltens et al., 2010; Kaufmann et al., 2011; Kaufmann et al., 2012; Perera et al., 2005b; Uhlmann et al., 2014), in our study 7 potential *Ephemera* species related markers were identified. Moreover, the applicability of these MS sex-related markers was also supported by an independent mathematical analysis (SOM) where raw MS data were evaluated. It is proposed that these MS sex-related markers can be used for MS database generation in the future.

Phylogenetically distant fish species (*Alosa agone*, *Coregonus macrophthalmus*, and *Rutilus rutilus*) have already been distinguished by a similar MS method based on mass spectra of muscle and liver tissues (Volta et al., 2012). Here, we successfully identified closely related species belonging to the Cypriniformes and Perciformes orders based on metabolite and/or neuropeptide markers. By using the ClinProTools analytical software, hierarchical clustering dendrogram was generated based on raw MS data obtained from different fish species. This MS-based dendrogram shows relatively high correspondence with the phylogenetic relationships (Broughton et al., 2013) of both the studied species and orders (Anguilliformes, Perciformes, Cypriniformes). The various fish species were identified with a high level of accuracy, 50 out of 52 individuals were correctly classified (96.2%) into 8 species. Within the same species (e.g. *P. fluviatilis*) individuals had similar patterns with the repetitive motif(s) but between species (e.g. *P. fluviatilis* and *L. gibbosus*) there was a distinction of the pattern. Additionally, some repetitive motifs were also observed within one cluster (e.g. cluster II.), which is important for MS based cluster generation. Closely related fish species belonging to the same order, family and genus were already analysed and discriminated based on structural characterization of species related parvalbumin isoforms by MALDI-TOF MS (Mazzeo et al., 2008). The exact characterisation of parvalbumin isoforms ($m/z \sim 11000$) required a time consuming and complicated protein analysis method (purification, digestion, peptide sequencing, and identification with database searching). Additionally, it is known that proteomic analysis provides an indirect measure of gene expression; therefore, it can serve as a supplementary tool for genomic approaches. On the contrary, we suggest that alterations in gene expression caused by changes in environmental factors could more accurately be detected by the analysis of neuropeptides and/or metabolites rather than measuring the profile of proteins. We believe that investigation of the low molecular range (m/z 500-5000) by the MS method is an easier way to identify the species or sex-related markers contributing to a routine species authentication independent from their phylogenetical relationship. By sampling both invertebrate and vertebrate species we also conclude that the MS based method can successfully be used regardless of the taxonomic category.

4.2 Sex identification in bleak by MALDI-TOF MS

In our earlier work, hypertrophied glandular tissue (gill gland) of adult male bloodfin tetra (*Aphyocharax anisitsi*) had already compared with gill tissues in females of the same species by LAESI MS method. Although, the aim of this study was different, a unique protein (m/z 11380) was found which only represented the male tissue. This unknown sex-related molecule as a

regulatory protein or enzyme is presumably involved in the production of the chemical species of fish (Shrestha et al., 2013). Here, the correct identification of sexes in all individuals within the same species proves the fast and high sensitivity of the MALDI-TOF MS method applied at the level of peptides and/or metabolites in the brain. We identified nineteen sex-related markers in bleak. Using these markers, all male and female individuals were correctly classified. The MS results were verified by morphological identification. The results showed 100% similarity to the determination of the sex of individual animals by morphological (testis or ovaries) observations before sample preparation. The data were also analysed by using two independent software evaluations. Using ClinProTools, the peak statistic of sex-related molecular markers provided significant differences between male and female *A. alburnus* (see Table 3 and Table S4). Additionally, the SOM analysis also showed differences between male and female fish, as well as similar patterns within male and female individuals. The sex-related molecular marker is a useful resource for studying sex-determining mechanisms and controlling fish sex. For example, intersex as an indicator for disruption of sexual differentiation, is a frequent phenomenon regarding different kinds of endocrine disrupting chemicals in fish. Earlier, the genetic sex was identified with male or female sex-linked PCR markers in various fish species (Chen et al., 2008; Hahlbeck et al., 2004; Patil et al., 2008; Wang et al., 2006). These PCR based methods are sensitive, accurate, and reliable techniques for genetic sexing of eggs, larvae, or adults. But in contrast to the MALDI-TOF MS method, the PCR based sex determination depends on sex-related primer pairs. Accordingly, we suggest that MALDI-TOF MS is better suited for subpopulation analysis, particularly sex determinate tools, in the ecological research.

4.3 Phylogenetically conserved markers in biological diversity research

Identification of phylogenetically conserved markers is a novel and instructive method in biological diversity research. Molecular biology and molecular and genetic markers are increasingly employed for both cryptic and new species identification and phylogenetic reconstruction. Several species related markers could also be described by classical morphological and DNA-based approaches (Stein et al., 2014). However, most of the described morphometric, meristic and descriptive markers are also influenced by various environmental factors, which complicates their interpretation. In contrast, molecular and genetic markers are transmitted across generations, independent from environmental factors (Fischer, 2013). DNA-based identification methods could increase the effect of species management plans to identify patterns of invasive species, especially for taxa that are difficult to identify morphologically. DNA-based identification needs only very low levels of biomass to detect the presence of a cryptic or new species. This method even allows the identification of species unavailable for morphological studies. By looking at a combination of morphological and DNA-based methods (metabarcoding), we suggest that the application of MALDI-TOF MS gives additional information that could help to discover markers and/or compounds related to new species. However, further studies are still needed in various species to fully recognise the usefulness of MALDI-TOF MS in species and sex identification.

4.4 Limits of MALDI-TOF MS utilization

Although, the utilized MS method is reliable, sensitive, fast and cheap, the chemical characteristic and exact type of observed molecular markers are unknown (metabolites or peptides), contrasted to slower and more expensive MS methods using structural characterization of protein by Mazzeo et al (2008). Generally, the routine application of MALDI-TOF MS methods mentioned above are dependent on comprehensive databases (Nomura, 2015). Additionally, the ClinProTools software is suitable for identification of specific markers in the case of *Ephemera* larvae and bleak sexes, in contrast to fish species. In mayflies and bleaks, only two different groups were compared so the majority (43% in mayfly, and 68% in bleak) of manually identified potential markers were supported also by automated ClinProTools evaluation (see Table 1 and S2; Table 3 and S4). However, in fish species, where eight different groups were compared, only 20% of manually identified species related markers were supported statistically by automated ClinProTools evaluation (see Table 2 and S3). We suppose that the relative number of evaluated ClinProTools peaks decrease (69/2 in mayfly, 37/2 in bleak, and 114/8 in fish) due to increasing the number of analysed groups and/or decreasing the number of individual within one group. We conclude that the applicability of ClinProTools alone in the mapping of specific molecular markers is questionable. Additionally, it is well suited for cluster classification (see Fig. 2) using complex marker pattern models, but only slightly well suited for the detection of all specific molecular markers. Therefore, research has needed to use a manual pre-searching method and an independent statistical analysis together for the validation of results or the development of data bases.

5. Conclusion

MALDI-TOF MS provides a fast, reliable and sensitive species and/or sex identification tool. Moreover, this approach has several other advantages such as high throughput, low sampling cost, and a quick and easy method. It is necessary to mention some disadvantages of the subject identification method: 1) the relatively high initial costs of MS, 2) the need to develop usable and accurate data bases for routine identification, 3) the ClinProTools software is only slightly suitable for identification in the case of small individual within one group or a large number of groups. Our results also demonstrate the usefulness of this method to investigate ecological questions. It is a powerful tool that can revolutionaries ecological and environmental research. In our future work, we plan to further investigate the application of this method to subpopulation analysis and to study habitat segregation due to ecological and environmental factors.

Authors contribution

The study was designed by VG and PZ, and the animal collection was performed by MG, VG, TP and PZ. The experimental work was performed by MG. The manuscript was written by MG and PZ with feedback from VG.

Declaration of interest

The authors declare that there is no conflict of interest that could be perceived as prejudicing the impartiality of the research reported.

Funding

This work was supported by PD-OTKA grant No. 115801 (TP), National Brain Project No. KTIA_NAP_13-2-2014-0006 (PZ), Bolyai Foundation No. BO/00952/16/8 (PZ).

Acknowledgement

The authors are indebted to Flextra-Lab Kft. (local distributor of Bruker Company in Hungary) for the help of the instrumental support.

References

- Aebersold R, Mann M. Mass spectrometry-based proteomics. *Nature* 2003; 422: 198-207.
- Alispahic M, Christensen H, Hess C, Razzazi-Fazeli E, Bisgaard M, Hess M. MALDI-TOF mass spectrometry confirms clonal lineages of *Gallibacterium anatis* between chicken flocks. *Veterinary Microbiology* 2012; 160: 269-273.
- Balczun C, Vorsprach B, Meiser CK, Schaub GA. Changes of the abundance of *Culicoides obsoletus* s.s. and *Culicoides scoticus* in Southwest Germany identified by a PCR-based differentiation. *Parasitology Research* 2009; 105: 345-349.
- Barbuddhe SB, Maier T, Schwarz G, Kostrzewa M, Hof H, Domann E, et al. Rapid identification and typing of *Listeria* species by matrix-assisted laser desorption ionization-time of flight mass spectrometry. *Applied and Environmental Microbiology* 2008; 74: 5402-5407.
- Bauernfeind E, and Soldán T. The mayflies of Europe (*Ephemeroptera*). Ollerup, Denmark, Apollo Books. 2012.
- Broughton RE, Betancur RR, Li C, Arratia G, Orti G. Multi-locus phylogenetic analysis reveals the pattern and tempo of bony fish evolution. *PLoS Currents* 2013; 5.
- Cardoso P, Erwin TL, Borges PAV, New TR. The seven impediments in invertebrate conservation and how to overcome them. *Biological Conservation* 2011; 144: 2647-2655.
- Cetre-Sossah C, Baldet T, Delecolle JC, Mathieu B, Perrin A, Grillet C, et al. 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* 2004; 35: 325-337.
- Cetre-Sossah C, Mathieu B, Setier-Rio ML, Grillet C, Baldet T, Delecolle JC, et al. Development and evaluation of a real-time quantitative PCR assay for *Culicoides imicola*, one of the main vectors of bluetongue (BT) and African horse sickness (AHS) in Africa and Europe. *Research in Veterinary Science* 2008; 85: 372-382.
- Chen SL, Li J, Deng SP, Tian YS, Wang QY, Zhuang ZM, et al. Isolation of female-specific AFLP markers and molecular identification of genetic sex in half-smooth tongue sole (*Cynoglossus semilaevis*). *Marine Biotechnology* 2007; 9: 273-280.

- Elbehiry A, Al-Dubaib M, Marzouk E, Osman S, Edrees H. Performance of MALDI biotyper compared with Vitek 2 compact system for fast identification and discrimination of *Staphylococcus* species isolated from bovine mastitis. *Microbiologyopen* 2016.
- Elliott, J. M., U. H. Humpesch, and T. T. Macan. Larvae of the British *Ephemeroptera*: a key with ecological notes. Ambleside, Cumbria: Freshwater Biological Association. 1988.
- Emami K, Hack E, Nelson A, Brain CM, Lyne FM, Mesbahi E, et al. Proteomic-based biotyping reveals hidden diversity within a microalgae culture collection: An example using *Dunaliella*. *Scientific Reports* 2015; 5.
- Feltens R, Gorner R, Kalkhof S, Groger-Arndt H, von Bergen M. Discrimination of different species from the genus *Drosophila* by intact protein profiling using matrix-assisted laser desorption ionization mass spectrometry. *Bmc Evolutionary Biology* 2010; 10.
- Feng X, Bansal N, Yang HS. Fish gelatin combined with chitosan coating inhibits myofibril degradation of golden pomfret (*Trachinotus blochii*) fillet during cold storage. *Food Chemistry* 2016; 200: 283-292.
- Feng X, Fu CL, Yang HS. Gelatin addition improves the nutrient retention, texture and mass transfer of fish balls without altering their nanostructure during boiling. *Lwt-Food Science and Technology* 2017a; 77: 142-151.
- Feng X, Ng VK, Miks-Krajnik M, Yang HS. Effects of fish gelatin and tea polyphenol coating on the spoilage and degradation of myofibril in fish fillet during cold storage. *Food and Bioprocess Technology* 2017b; 10: 89-102.
- Fischer J. Fish identification tools for biodiversity and fisheries assessments. Food and Agriculture Organization Of The United Nations, Rome, 2013.
- Goncalves JL, Tomazi T, Barreiro JR, Braga PAD, Ferreira CR, Araujo JP, et al. Identification of *Corynebacterium* spp. isolated from bovine intramammary infections by matrix-assisted laser desorption ionization-time of flight mass spectrometry. *Veterinary Microbiology* 2014; 173: 147-151.
- Goodwin CR, Covington BC, Derewacz DK, McNees CR, Wikswo JP, McLean JA, et al. Structuring microbial metabolic responses to multiplexed stimuli via self-organizing metabolomics maps. *Chemistry & Biology* 2015; 22: 661-670.
- Hahlbecka E, Griffiths R, Bengtsson BE. The juvenile three-spined stickleback (*Gasterosteus aculeatus* L.) as a model organism for endocrine disruption I. Sexual differentiation. *Aquatic Toxicology* 2004; 70: 287-310.
- Harismendy O, Ng PC, Strausberg RL, Wang XY, Stockwell TB, Beeson KY, et al. Evaluation of next generation sequencing platforms for population targeted sequencing studies. *Genome Biology* 2009; 10.
- Ilina EN, Borovskaya AD, Malakhova MM, Vereshchagin VA, Kubanova AA, Kruglov AN, et al. Direct bacterial profiling by matrix-assisted laser desorption-ionization time-of-flight mass spectrometry for identification of pathogenic *Neisseria*. *Journal of Molecular Diagnostics* 2009; 11: 75-86.
- Kaufmann C, Ziegler D, Schaffner F, Carpenter S, Pfluger V, Mathis A. Evaluation of matrix-assisted laser desorption/ionization time of flight mass spectrometry for characterization of *Culicoides nubeculosus* biting midges. *Medical and Veterinary Entomology* 2011; 25: 32-38.
- Kaufmann C, Schaffner F, Ziegler D, Pflüger V, Mathis A. Identification of field-caught *Culicoides* biting midges using matrix-assisted laser desorption/ionization time of flight mass spectrometry. *Parasitology* 2012; 139: 248-58.

- Kim SW, Jang HB, Lee JS, Im SP, Lazarte JMS, Seo JP, et al. Comparison of proteome typing and serotyping of *Streptococcus parauberis* isolates from olive flounder (*Paralichthys olivaceus*). Journal of Microbiological Methods 2015; 118: 168-172.
- Lee HW, Roh SW, Cho K, Kim KN, Cha IT, Yim KJ, et al. Phylogenetic analysis of microalgae based on highly abundant proteins using mass spectrometry. Talanta 2015; 132: 630-634.
- Lopez JL, Abalde SL, Fuentes J. Proteomic approach to probe for larval proteins of the mussel *Mytilus galloprovincialis*. Marine Biotechnology 2005; 7: 396-404.
- Mazzeo MF, De Giulio B, Guerriero G, Ciarcia G, Malorni A, Russo GL, et al. Fish authentication by maldi-tof mass spectrometry. Journal of Agricultural and Food Chemistry 2008; 56: 11071-11076.
- Nagy E, Maier T, Urban E, Terhes G, Kostrzewa M. Species identification of clinical isolates of bacteroides by matrix-assisted laser-desorption/ionization time-of-flight mass spectrometry. Clinical Microbiology and Infection 2009; 15: 796-802.
- Nolan DV, Carpenter S, Barber J, Mellor PS, Dallas JF, Mordue AJ, et al. Rapid diagnostic PCR assays for members of the *Culicoides obsoletus* and *Culicoides pulicaris* species complexes, implicated vectors of bluetongue virus in Europe. Veterinary Microbiology 2007; 124: 82-94.
- Nomura F. Proteome-based bacterial identification using matrix-assisted laser desorption ionization-time of flight mass spectrometry (MALDI-TOF MS): A revolutionary shift in clinical diagnostic microbiology. Biochimica Et Biophysica Acta-Proteins and Proteomics 2015; 1854: 528-537.
- Pages N, Monteys VSI. Differentiation of *Culicoides obsoletus* and *Culicoides scoticus* (Diptera: Ceratopogonidae) based on mitochondrial cytochrome oxidase subunit I. Journal of Medical Entomology 2005; 42: 1026-1034.
- Patel R. MALDI-TOF mass spectrometry: transformative proteomics for clinical microbiology. Clinical Chemistry 2013; 59: 340-342.
- Patil JG and Hinze HJ. Simplex PCR assay for positive identification of genetic sex in the Japanese medaka, *Oryzias latipes*. Marine Biotechnology 2008; 10: 641-644.
- Perera MR, Vanstone VA, Jones MGK. A novel approach to identify plant parasitic nematodes using matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. Rapid Communications in Mass Spectrometry 2005a; 19: 1454-1460.
- Perera MR, Vargas RDF, Jones MGK. Identification of aphid species using protein profiling and matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. Entomologia Experimentalis Et Applicata 2005b; 117: 243-247.
- Radulovici AE, Archambault P, Dufresne F. DNA barcodes for marine biodiversity: moving fast forward? Diversity 2010; 2:450-472.
- Salla V, Murray KK. Matrix-assisted laser desorption ionization mass spectrometry for identification of shrimp. Analytica Chimica Acta 2013; 794:55-59.
- Seng P, Drancourt M, Gourié F, La Scola B, Fournier PE, Rolain JM, et al. Ongoing revolution in bacteriology: routine identification of bacteria by matrix-assisted laser desorption ionization time-of-flight mass spectrometry. Clinical Infectious Diseases 2009; 49: 543-51.
- Shrestha B, Javonillo R, Burns JR, Pirger Z, Vertes A. Comparative local analysis of metabolites, lipids and proteins in intact fish tissues by LAESI mass spectrometry. Analyst 2013; 138: 3444-3449.

- Stein ED, Martinez MC, Stiles S, Miller PE, Zakharov EV. Is DNA barcoding actually cheaper and faster than traditional morphological methods: results from a survey of freshwater bioassessment efforts in the united states? Plos One 2014; 9.
- Stephan A, Clausen PH, Bauer B, Steuber S. PCR identification of *Culicoides dewulfi* midges (Diptera: Ceratopogonidae), potential vectors of bluetongue in Germany. Parasitology Research 2009; 105: 367-371.
- Stephan R, Johler S, Oesterle N, Naumann G, Vogel G, Pfluger V. Rapid and reliable species identification of scallops by MALDI-TOF mass spectrometry. Food Control 2014; 46: 6-9.
- Takacs P, Vital Z, Ferincz A, Staszny A. Repeatability, reproducibility, separative power and subjectivity of different fish morphometric analysis methods. Plos One 2016; 11.
- Uhlmann KR, Gibb S, Kalkhof S, Arroyo-Abad U, Schulz C, Hoffmann B, et al. Species determination of *Culicoides* biting midges via peptide profiling using matrix-assisted laser desorption ionization mass spectrometry. Parasites and Vectors 2014; 7:392.
- Volta P, Riccardi N, Lauceri R, Tonolla M. Discrimination of freshwater fish species by matrix-assisted laser desorption/ionization-time of flight mass spectrometry (MALDI-TOF MS): a pilot study. Journal of Limnology 2012; 71: 164-169.
- von Bubnoff A. Next-generation sequencing: The race is on. Cell 2008; 132: 721-723.
- Wang X, Williams E, Haasch ML, Dasmahapatra AK. Japanese medaka (*Oryzias latipes*): Developmental model for the study of alcohol teratology. Birth Defects Research (Part B) 2006; 77: 29-39.
- Zabbe JB, Zanardo L, Megraud F, Bessede E. MALDI-TOF mass spectrometry for early identification of bacteria grown in blood culture bottles. Journal of Microbiological Methods 2015; 115: 45-46.
- Zemlak TS, Ward RD, Connell AD, Holmes BH, Hebert PDN. DNA barcoding reveals overlooked marine fishes. Molecular Ecology Resources 2009; 9: 237-242.

Figure legends

Figure 1 – Cluster analysis of the overall signal pattern generated by MALDI-TOF MS analyses of various fish species. Dendrogram was made by ClinProTools (settings: Unit Variance 95 % sum explained Variance, Euclidean distance method, 1.5 Minkowski Exponent, Average linkage method). Four separated clusters were illustrated. Dendrogram also represents the phylogenetic relationships of the studied taxa (both species and families e.g. Anguilliformes, Perciformes, Cypriniformes).

Figure 2 – Identification of fish species based on SOMs. Species related pattern could be visualized by the component planes of the SOM illustrating in the range of m/z 500 to 5000. The redder colour indicates higher intensities of the different markers.

Figure 3 – Sex identification within *A. alburnus*. The A panel shows the intensity of m/z 1062.7 and 808.6, as identified sex-related differences between the two sexes. The B panel illustrates the intensities of m/z 1062.7 in a smaller range. Representative cluster analysis (C) from sample sets

of the male (green) and female fish (red) using the relative intensities of the markers with m/z 808.6 and 1062.7 by ClinProTools. Sex identification of *A. alburnus* based on SOMs (D) presenting 3-3 animals. The redder colour indicates higher intensities of the markers in the range of m/z 1300.0 to 2000.0.

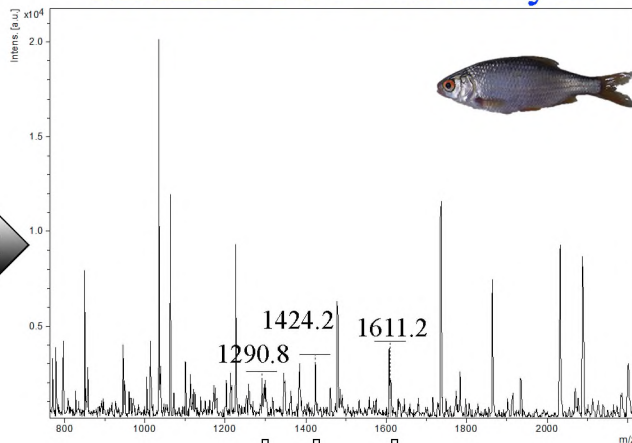
Table 1 – Different species related markers between the two mayfly larvae identified by MALDI-TOF MS. These markers allowed the identification of mayfly species. Use of different peak picking methods (monoisotopic mass - manually, and average mass - ClinProTools) caused the differences between m/z value of similar peaks from manually (e.g. m/z 1252.0, see Table 1) and ClinProTools data (e.g. m/z 1252.5, see Supplementary data Table S2). n.d.– not detected species related marker, * – indicates the significant differences ($p<0.05$) between the *Ephemera* species by ClinProTools (see Supplementary data, Table S2)

Table 2 – Species related markers for identification of fish by MALDI-TOF MS. Use of different peak picking methods (monoisotopic mass - manually, and average mass - ClinProTools) caused the differences between m/z value of similar peaks from manually (e.g. m/z 1020.6, see Table 2) and ClinProTools data (e.g. m/z 1020.8, see Supplementary data Table S3). * – indicates the significant differences ($p<0.05$) between the various fish species by ClinProTools (see Supplementary data, Table S3)

Table 3 – Different sex-related markers from *A. alburnus* based on MALDI-TOF MS analysis. These markers allowed the sex identification within one fish species. Use of different peak picking methods (monoisotopic mass - manually, and average mass - ClinProTools) caused the differences between m/z value of similar peaks from manually (e.g. m/z 1749.1, see Table 3) and ClinProTools data (e.g. m/z 1749.9, see Supplementary data Table S4). n.d.– not detected sex related marker, * – indicates the significant differences ($p<0.05$) between the male and female by ClinProTools (see Supplementary data, Table S4)

confirmation
by SOM

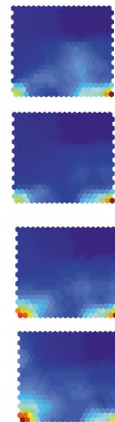
MALDI-TOF MS analysis



based on species related markers



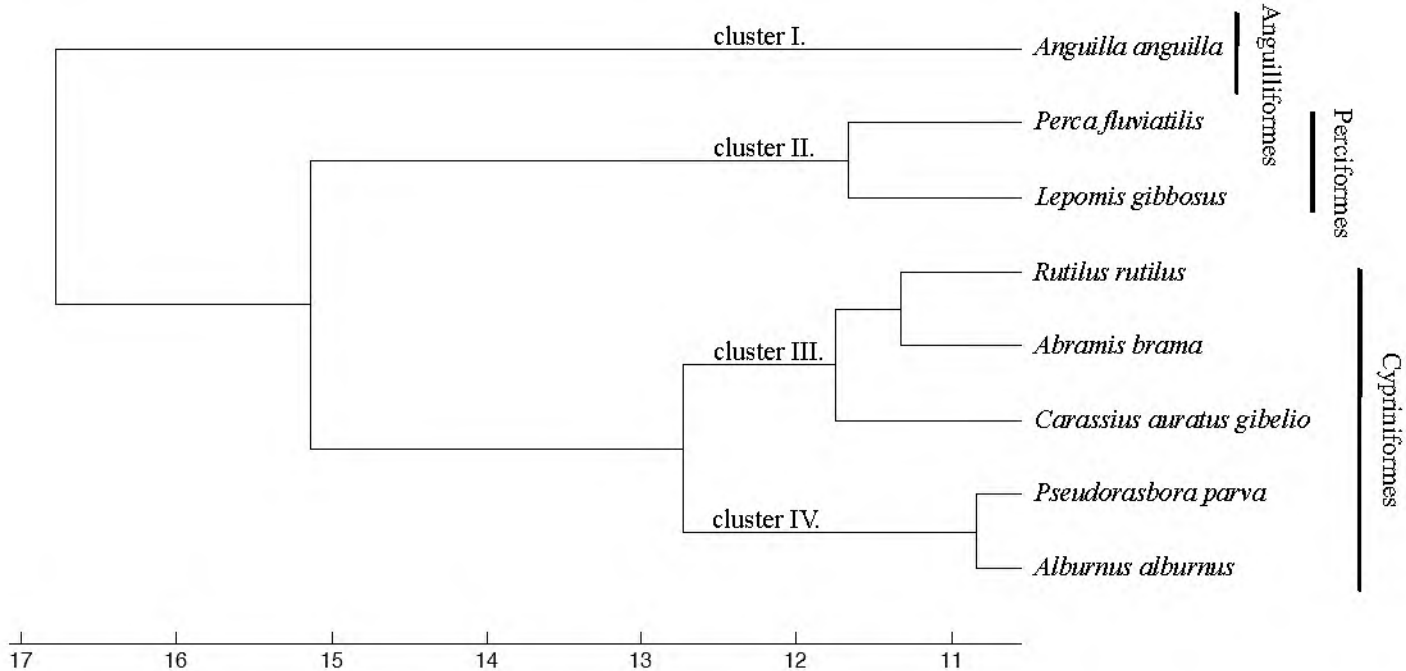
correct species ID (*Rutilus rutilus*)



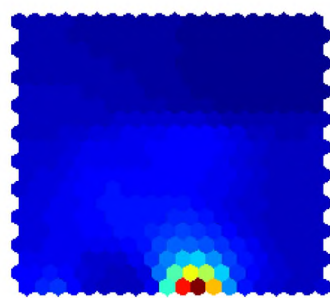
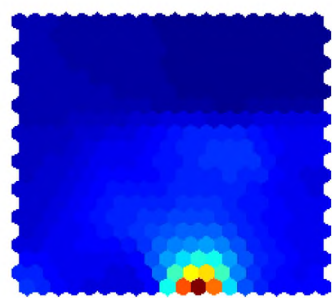
m/z	<i>Ephemera vulgata</i>	<i>Ephemera lineata</i>
1252.0*	+	n.d.
1297.6	n.d.	+
1336.7	n.d.	+
1789.0	+	n.d.
2697.8*	+	n.d.
3166.0	+	n.d.
3281.0*	n.d.	+

<i><u>Alburnus</u></i> <i><u>alburnus</u></i>	<i><u>Rutilus</u></i> <i><u>rutilus</u></i>	<i><u>Pseudorasbora</u></i> <i><u>parva</u></i>	<i><u>Anguilla</u></i> <i><u>anguilla</u></i>	<i><u>Abramis</u></i> <i><u>brama</u></i>	<i><u>Carassius</u></i> <i><u>auratus gibelio</u></i>	<i><u>Perca</u></i> <i><u>fluviatilis</u></i>	<i><u>Lepomis</u></i> <i><u>gibbosus</u></i>
m/z	m/z	m/z	m/z	m/z	m/z	m/z	m/z
745.5	1290.8	1020.6*	916.6	771.4	777.4*	794.5	594.3*
1169.8	1315.8	1030.6	987.7	1784.0	949.6*	891.5*	710.4
	1424.2	1109.6	1080.7	2221.7	1112.6*	936.6*	731.4
	1714.2	1522.8	1166.8	2256.2	1145.6	979.6*	1127.6
	1797.3	1768.9	1171.8	2308.6	1446.8	1098.6	1252.6*
	2031.2	3036.5	1207.7	2343.6	1559.8	1182.6	1405.8
			1378.0		1573.8	1414.8	1940.0
			1507.0*		1578.9	1448.9	4721.6
			1578.1*		1765.0*	1718.9	
			1805.1		1900.0	1940.3	
			1849.2		1909.0		
			2117.2				

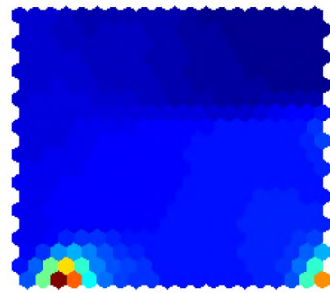
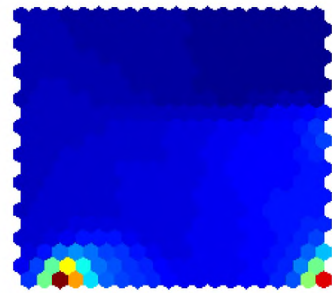
m/z	male <i>A. alburnus</i>	female <i>A. alburnus</i>
808.6	n.d.	+
945.8*	n.d.	+
960.8*	+	n.d.
1062.7*	+	n.d.
1131.6*	+	n.d.
1178.8*	+	n.d.
1201.8*	+	n.d.
1225.8*	+	n.d.
1315.9	+	n.d.
1388.1	n.d.	+
1460.9	+	n.d.
1611.1*	+	n.d.
1749.1*	n.d.	+
2088.5*	+	n.d.
3081.9	n.d.	+
4821.1*	n.d.	+
4858.2	n.d.	+
4893.2*	n.d.	+
4931.0*	n.d.	+



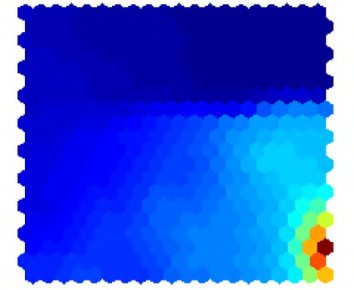
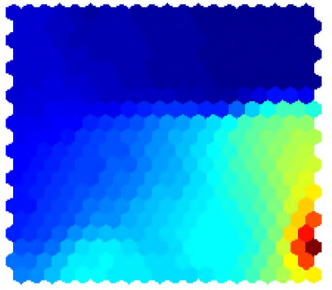
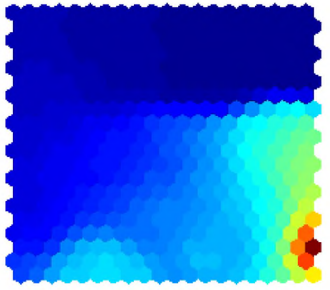
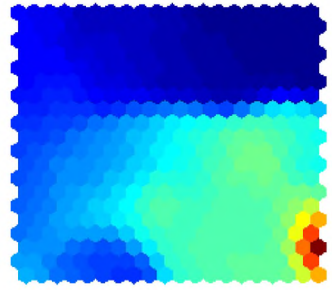
Anguilla anguilla



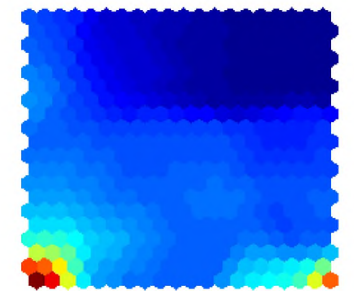
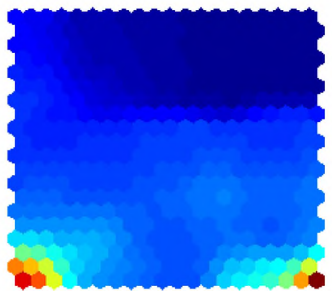
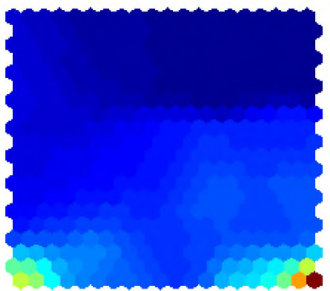
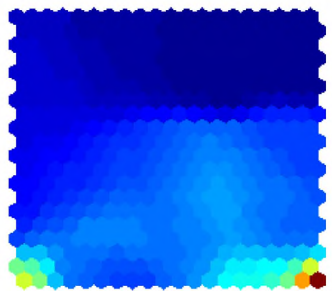
Perca fluviatilis



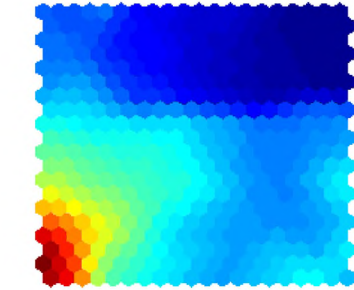
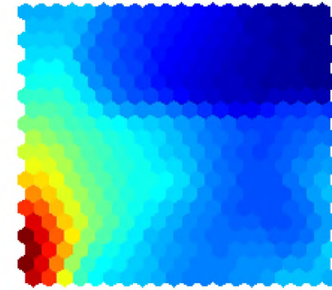
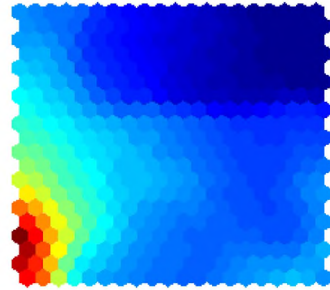
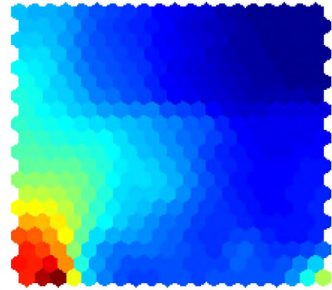
Lepomis gibbosus



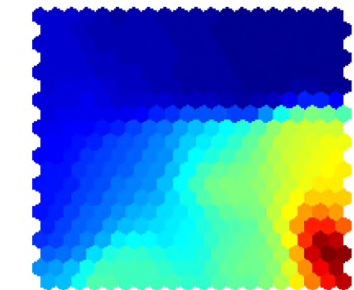
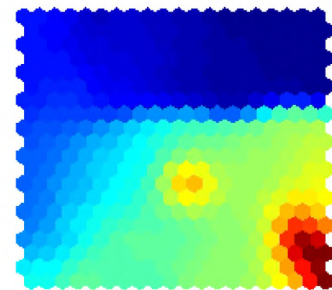
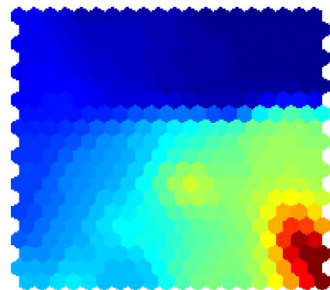
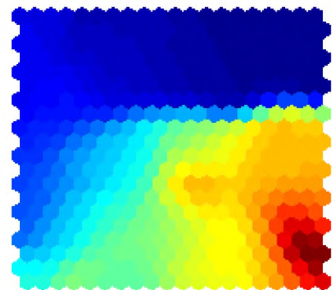
Rutilus rutilus



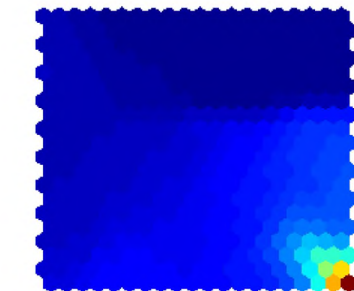
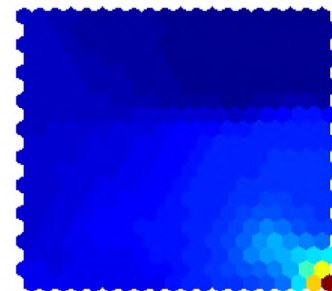
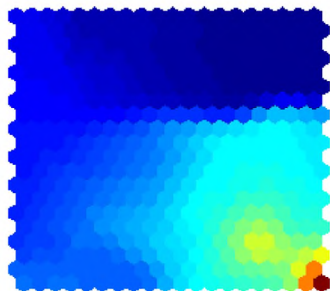
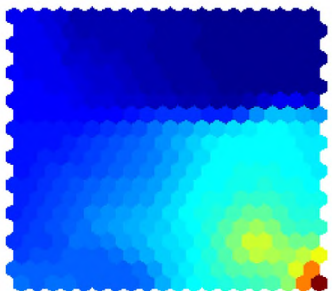
Abramis brama



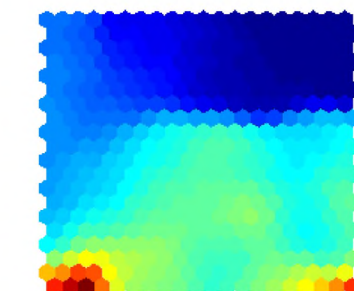
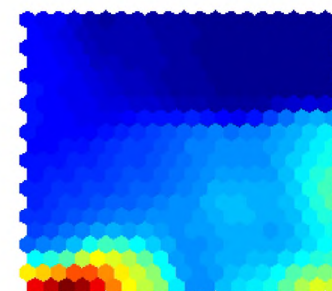
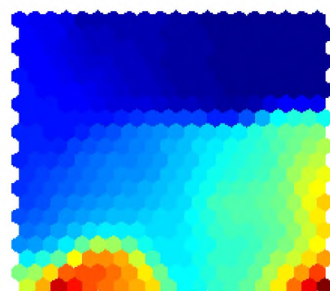
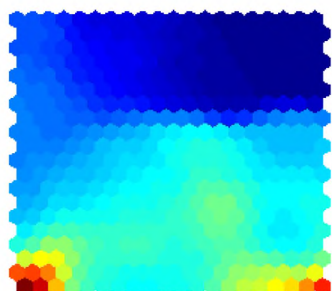
Carassius auratus gibelio

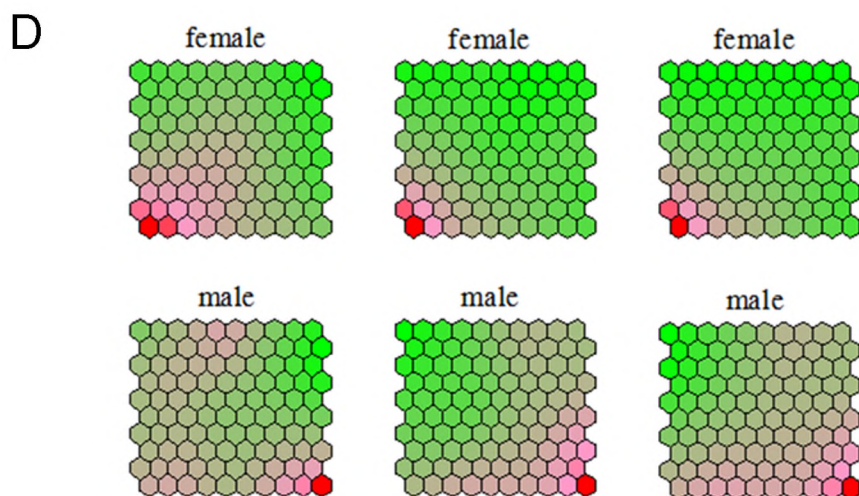
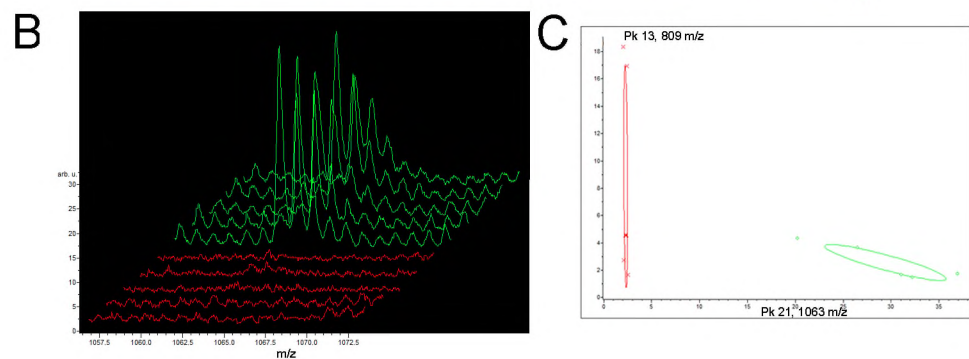
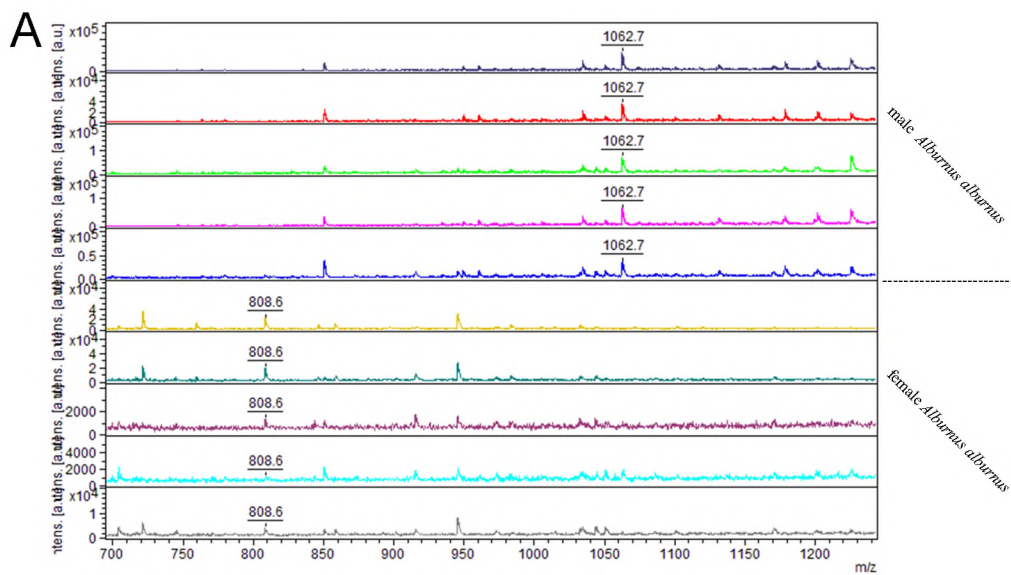


Pseudorasbora parva



Alburnus alburnus





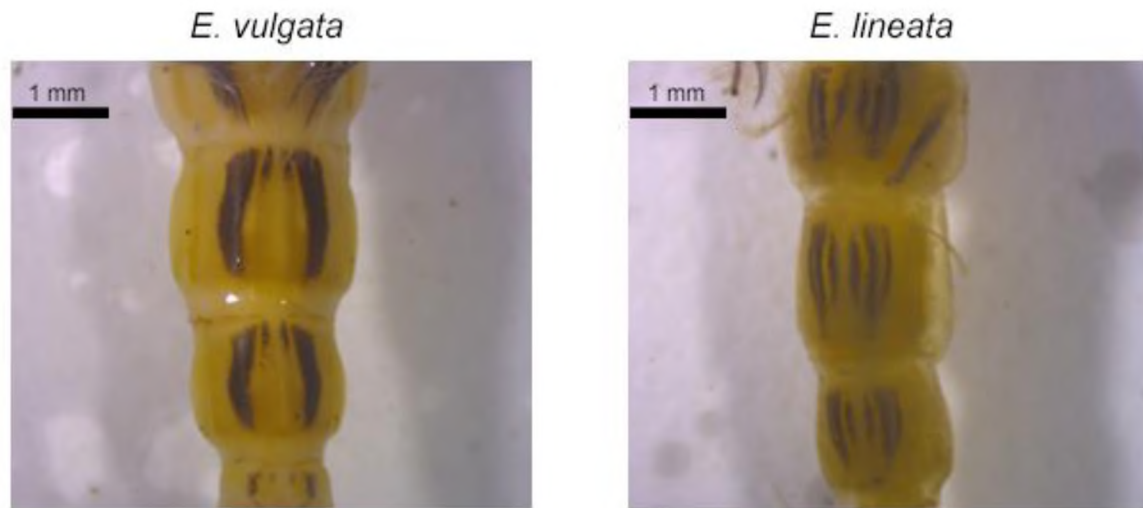


Figure S1 – Morphological sex identification of *Ephemera* species. The *E. vulgata* (A) and *E. lineata* (B) were identified based on the abdominal terga of 7, 8 and 9 (Elliott, et al., 1988; Bauernfeind and Soldán, 2012).

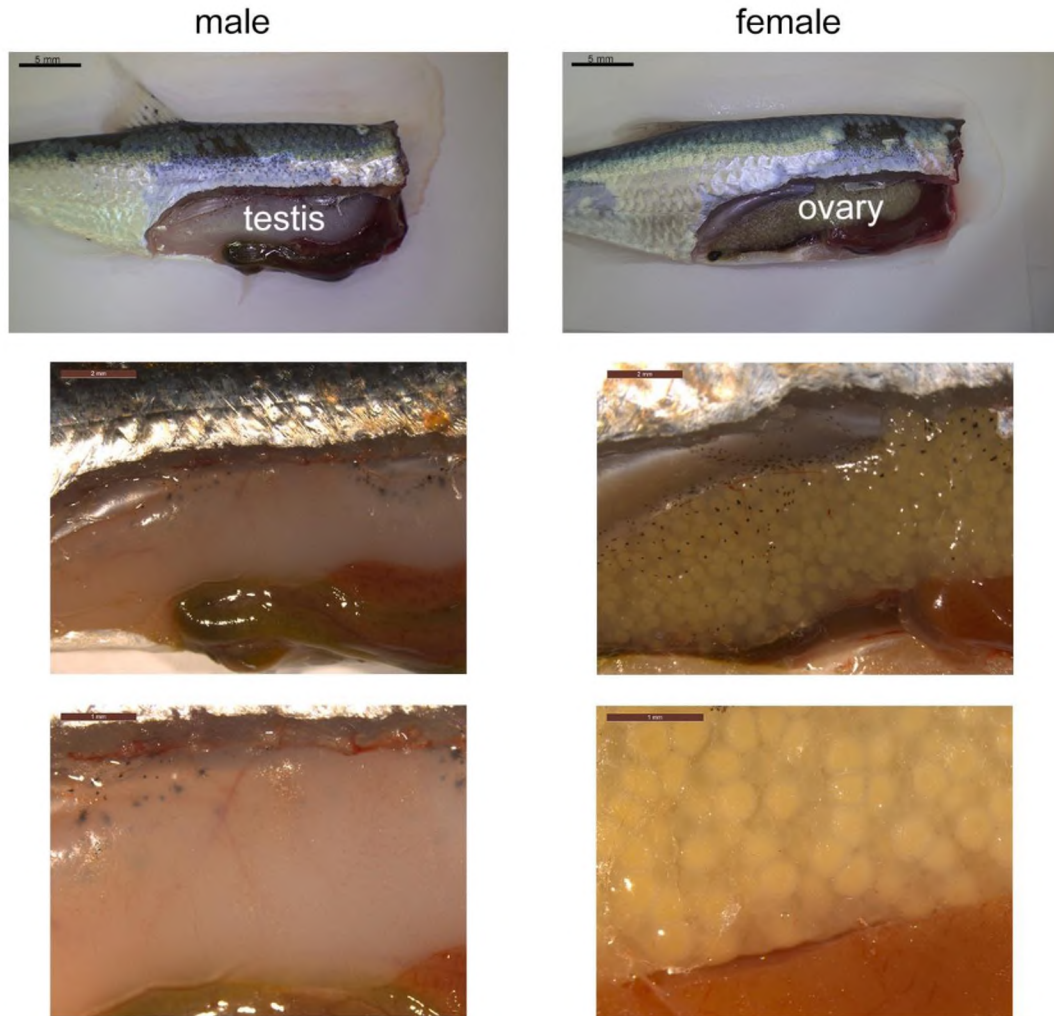


Figure S2 – Morphological sex identification of *A. alburnus*. The male (left side) and female (right side) were identified based on testis and ovary, respectively.

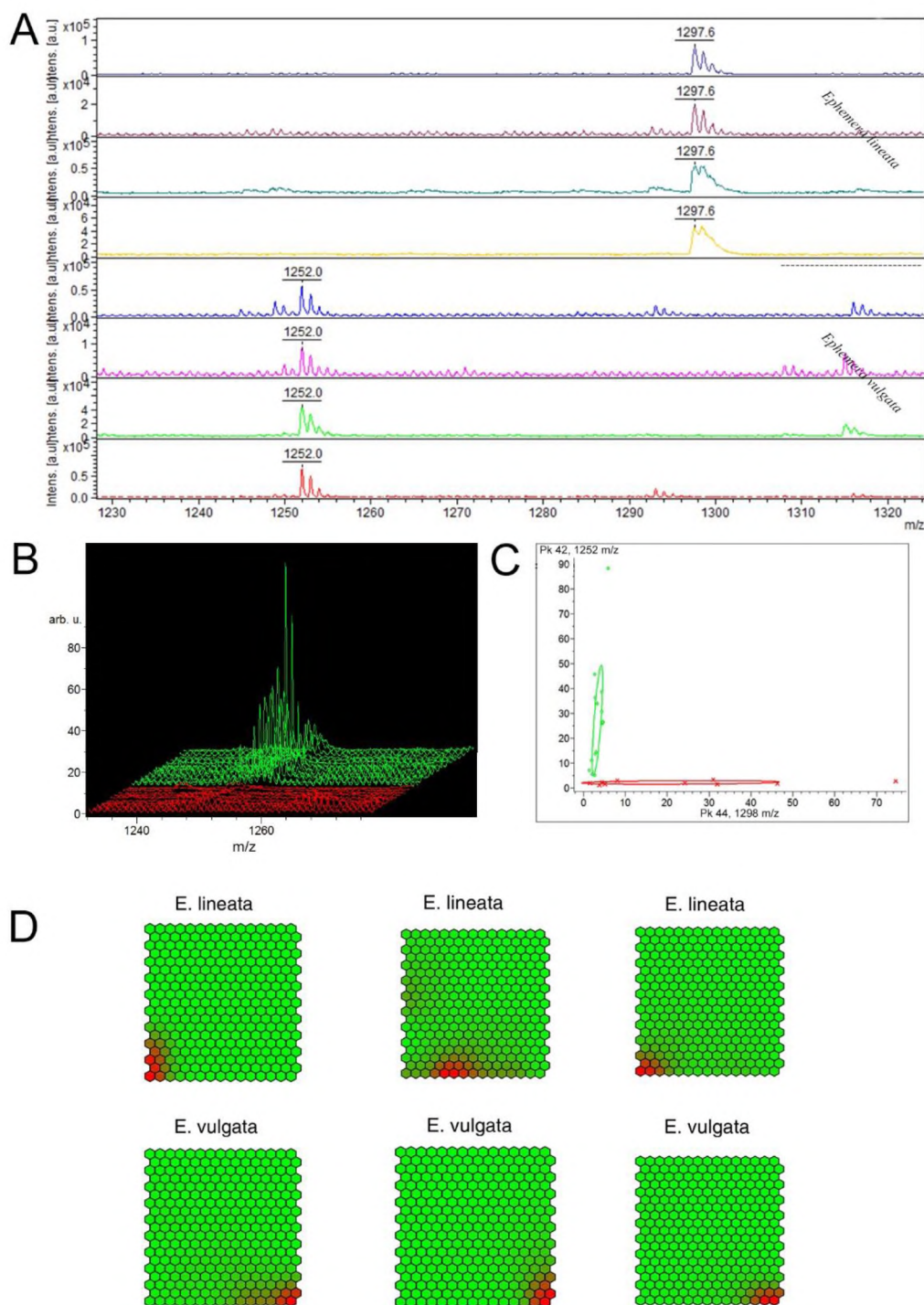


Figure S3 – Identification of mayfly species. The A panel shows the intensity of m/z 1252.0 and 1297.6, as an identified differences between the two species. The B panel graphically illustrates the intensities of m/z 1252.0 in a smaller range (m/z 1232-1276). Representative cluster analysis (C) from sample sets of the *E. vulgata* (green) and *E. lineata* (red) groups using the relative intensities of the markers with m/z 1252.0 and 1297.6 by ClinProTools. Identification of mayfly larvae based on SOMs (D) presenting 3-3 individuum. The redder color indicates higher intensities of the markers in the range of m/z 1200.0 to 1300.0.

Monoisotopic mass [m/z]	p value	Location	Specificity
1252.0	0.00330	Table 1	<i>E. vulgata</i>
1789.0	0.00771	Table 1	<i>E. vulgata</i>
2697.8	0.04325	Table 1	<i>E. vulgata</i>
3166.0	0.00956	Table 1	<i>E. vulgata</i>
1297.6	0.04197	Table 1	<i>E. lineata</i>
1336.7	0.00955	Table 1	<i>E. lineata</i>
3281.0	0.00278	Table 1	<i>E. lineata</i>
745.5	0.00121	Table 2	<i>A. alburnus</i>
1169.8	0.01164	Table 2	<i>A. alburnus</i>
1290.8	0.00217	Table 2	<i>R. rutilus</i>
1315.8	0.01624	Table 2	<i>R. rutilus</i>
1424.2	0.00573	Table 2	<i>R. rutilus</i>
1714.2	0.01687	Table 2	<i>R. rutilus</i>
1797.3	0.03812	Table 2	<i>R. rutilus</i>
2031.2	0.00463	Table 2	<i>R. rutilus</i>
1020.6	0.00286	Table 2	<i>P. parva</i>
1030.6	0.00130	Table 2	<i>P. parva</i>
1109.6	0.00183	Table 2	<i>P. parva</i>
1522.8	0.04811	Table 2	<i>P. parva</i>
1768.9	0.00722	Table 2	<i>P. parva</i>
3036.5	0.00674	Table 2	<i>P. parva</i>
916.6	0.00165	Table 2	<i>A. anguilla</i>
987.7	0.00092	Table 2	<i>A. anguilla</i>
1080.7	0.00126	Table 2	<i>A. anguilla</i>
1166.8	0.00114	Table 2	<i>A. anguilla</i>
1171.8	0.00226	Table 2	<i>A. anguilla</i>
1207.7	0.00367	Table 2	<i>A. anguilla</i>
1378.0	0.01632	Table 2	<i>A. anguilla</i>
1507.0	0.00819	Table 2	<i>A. anguilla</i>
1578.1	0.02191	Table 2	<i>A. anguilla</i>
1805.1	0.01452	Table 2	<i>A. anguilla</i>
1849.2	0.01824	Table 2	<i>A. anguilla</i>
2117.2	0.01136	Table 2	<i>A. anguilla</i>
771.4	0.01623	Table 2	<i>A. brama</i>
1784.0	0.01116	Table 2	<i>A. brama</i>
2221.7	0.01115	Table 2	<i>A. brama</i>
2256.2	0.00565	Table 2	<i>A. brama</i>

2308.6	0.00563	Table 2	<i>A. brama</i>
2343.6	0.00288	Table 2	<i>A. brama</i>
777.4	0.00116	Table 2	<i>C. auratus gibelio</i>
949.6	0.01134	Table 2	<i>C. auratus gibelio</i>
1112.6	0.00071	Table 2	<i>C. auratus gibelio</i>
1145.6	0.00273	Table 2	<i>C. auratus gibelio</i>
1446.8	0.01082	Table 2	<i>C. auratus gibelio</i>
1559.8	0.00935	Table 2	<i>C. auratus gibelio</i>
1573.8	0.02374	Table 2	<i>C. auratus gibelio</i>
1578.9	0.02769	Table 2	<i>C. auratus gibelio</i>
1765.0	0.02483	Table 2	<i>C. auratus gibelio</i>
1900.0	0.00782	Table 2	<i>C. auratus gibelio</i>
1909.0	0.04493	Table 2	<i>C. auratus gibelio</i>
794.5	0.00226	Table 2	<i>P. fluviatilis</i>
891.5	0.00214	Table 2	<i>P. fluviatilis</i>
936.6	0.00216	Table 2	<i>P. fluviatilis</i>
979.6	0.00084	Table 2	<i>P. fluviatilis</i>
1098.6	0.00087	Table 2	<i>P. fluviatilis</i>
1182.6	0.00084	Table 2	<i>P. fluviatilis</i>
1414.8	0.00088	Table 2	<i>P. fluviatilis</i>
1448.9	0.04999	Table 2	<i>P. fluviatilis</i>
1718.9	0.04824	Table 2	<i>P. fluviatilis</i>
1940.3	0.04826	Table 2	<i>P. fluviatilis</i>
594.3	0.00123	Table 2	<i>L. gibbosus</i>
710.4	0.00126	Table 2	<i>L. gibbosus</i>
731.4	0.00863	Table 2	<i>L. gibbosus</i>
1127.6	0.01662	Table 2	<i>L. gibbosus</i>
1252.6	0.01661	Table 2	<i>L. gibbosus</i>
1405.8	0.01157	Table 2	<i>L. gibbosus</i>
1940.0	0.04970	Table 2	<i>L. gibbosus</i>
4721.6	0.01724	Table 2	<i>L. gibbosus</i>
960.8	0.00495	Table 3	male <i>A. alburnus</i>
1062.7	0.00314	Table 3	male <i>A. alburnus</i>
1131.6	0.00499	Table 3	male <i>A. alburnus</i>
1178.8	0.00774	Table 3	male <i>A. alburnus</i>
1201.8	0.00652	Table 3	male <i>A. alburnus</i>
1225.8	0.00195	Table 3	male <i>A. alburnus</i>
1315.9	0.03325	Table 3	male <i>A. alburnus</i>
1460.9	0.04006	Table 3	male <i>A. alburnus</i>
1611.1	0.00322	Table 3	male <i>A. alburnus</i>
2088.5	0.00019	Table 3	male <i>A. alburnus</i>

808.6	0.00956	Table 3	female <i>A. alburnus</i>
945.8	0.00669	Table 3	female <i>A. alburnus</i>
1388.1	0.02567	Table 3	female <i>A. alburnus</i>
1749.1	0.00333	Table 3	female <i>A. alburnus</i>
3081.9	0.04420	Table 3	female <i>A. alburnus</i>
4821.1	0.00877	Table 3	female <i>A. alburnus</i>
4858.2	0.05000	Table 3	female <i>A. alburnus</i>
4893.2	0.00763	Table 3	female <i>A. alburnus</i>
4931.0	0.00761	Table 3	female <i>A. alburnus</i>

Table S1

The manually pre-selected potential markers were statistically confirmed by Matlab. Kruskal-Wallis ANOVA was used for parametric statistical analysis of the different sample groups ($p < 0.05$).

ClinProt Peak Statistic of *Ephemera* larvae

ClinProTools Version: 3.0 build 22

Number of peaks: 69

Sort Mode: p value tta

S	Index	Mass	DAve	PTTA	PWKW	PAD	Ave1	Ave2	StdDev1	StdDev2	CV1	CV2
X	50	1965.16	5.24	0.00379	0.00386	0.282	3.42	8.66	2.09	3.06	61.09	35.4
X	45	1769.86	6.63	0.00417	0.000352	0.0185	2.55	9.18	1.11	4.65	43.61	50.63
X	34	1199.38	10.06	0.00419	0.00237	0.00189	2.79	12.85	2.05	7.35	73.6	57.24
X	35	1252.45	27.81	0.0163	0.0000704	0.000102	2.77	30.59	1.01	24.48	36.56	80.04
X	40	1505.02	7.78	0.0252	0.00555	0.000952	11.96	4.17	5.87	1.75	49.08	41.86
X	58	2700.41	6.15	0.0252	0.00462	0.0000178	1.16	7.31	0.78	6.04	67.28	82.52
X	8	527.65	8.63	0.0275	0.000352	0.0000196	12	3.37	6.73	1.29	56.06	38.38
X	9	542.28	8.16	0.0305	0.000975	0.0000767	11.87	3.71	6.63	1.9	55.84	51.17
X	38	1316.33	7.03	0.0326	0.0436	0.000372	4.04	11.07	2.47	7.51	61.18	67.9
X	64	3284.67	5.85	0.0326	0.0000704	< 0.000001	6.54	0.69	4.97	0.25	75.96	36.9
X	49	1932.94	6.11	0.0376	0.0436	0.0113	4.23	10.34	2.33	6.83	55.1	66.03
X	53	2195.34	3.79	0.0376	0.0129	0.00969	7.2	3.41	3.36	1.77	46.72	52.02
X	17	664.6	22.27	0.105	0.000352	< 0.000001	25.9	3.63	25.28	1.6	97.59	44.13
X	18	665.57	10.69	0.105	0.000352	< 0.000001	13.09	2.4	12.15	0.92	92.83	38.4
X	37	1297.8	21.95	0.125	0.0236	< 0.000001	25.68	3.73	26.36	1.38	102.64	36.95
X	33	1093.07	17.49	0.125	0.076	< 0.000001	1.98	19.47	0.97	26.64	49.04	136.83
X	27	904.96	5.33	0.129	0.0596	< 0.000001	8.02	2.7	6.72	0.85	83.75	31.37
X	22	792.74	2.81	0.129	0.076	0.00502	5.47	2.66	3.24	2.26	59.28	84.89
X	65	3311.47	1.7	0.133	0.22	0.000402	1.28	2.98	1.54	2.19	119.95	73.67
X	39	1337.8	7.42	0.15	0.193	0.000305	12.44	5.02	10.23	3.09	82.23	61.52
X	10	543.29	3.02	0.15	0.0334	0.0000447	5.14	2.11	4.09	1.07	79.51	50.77
X	57	2656.86	2.42	0.15	0.234	0.524	3.46	5.88	1.59	3.89	46.09	66.1
X	63	3205.22	1.79	0.15	0.112	0.000967	3.28	1.49	2.46	0.9	74.89	60.36
X	56	2588.26	5.77	0.186	0.00663	< 0.000001	7.41	1.64	8.8	0.87	118.72	53.02
X	28	954.03	5.38	0.186	0.125	0.0489	7.35	12.73	6.31	7.77	85.84	61.01
X	25	848.65	4.79	0.186	0.548	< 0.000001	1.97	6.76	1.37	9.16	69.62	135.4
X	29	971.1	4.03	0.186	0.101	0.000199	7.19	3.16	5.96	3.19	82.93	100.96
X	30	976.95	3.55	0.186	0.154	0.000459	4.12	7.67	3.29	5.91	79.83	77.07
X	11	550.32	7.17	0.262	0.193	0.0000778	12.64	5.47	12.38	4.83	97.94	88.25
X	14	569.29	9.15	0.275	0.206	0.0000169	15.74	6.59	17.37	6.32	110.31	95.79

ClinProt Peak Statistic

X	5	524.34	8.77	0.275	0.188	0.0000196	15.37	6.6	15.98	5.82	104.02	88.22
X	1	506.34	7.43	0.275	0.188	0.00000605	12.02	4.59	13.4	4.1	111.51	89.31
X	6	525.3	3.82	0.275	0.22	0.00000937	6.89	3.07	7.19	2.52	104.37	82.26
X	12	551.3	2.89	0.275	0.234	0.000145	5.85	2.96	5.5	2.14	94.13	72.37
X	23	811.81	1.54	0.275	0.333	0.000983	3.1	4.64	1.62	3.15	52.46	67.96
X	62	3181.06	1.33	0.275	0.304	0.00956	3.18	1.84	2.44	1.24	76.72	67.48
X	13	568.35	16.26	0.283	0.193	0.000199	31.24	14.98	31.45	15.89	100.66	106.03
X	42	1636.62	8.33	0.283	0.139	0.0000998	29.17	20.83	15.39	11.47	52.77	55.06
X	7	526.36	3.64	0.283	0.256	0.00000366	6.78	3.14	7.48	2.56	110.27	81.47
X	15	570.32	3.27	0.283	0.206	0.00000974	6.22	2.95	6.53	2.39	104.88	80.82
X	4	522.35	3.09	0.283	0.28	0.00000366	5.86	2.77	6.3	2.21	107.49	79.75
X	46	1790.79	3.9	0.286	0.333	0.129	7.69	11.59	5.85	7.88	76.1	67.98
X	67	3739.2	0.26	0.286	0.333	0.546	0.83	1.09	0.37	0.55	44.59	50.69
X	2	507.31	3.53	0.306	0.333	0.0000352	7.19	3.66	7.68	3.01	106.84	82.29
X	69	4756.08	0.56	0.34	0.571	< 0.000001	1.02	0.46	1.31	0.4	128.93	86.1
X	20	688.3	2.06	0.342	0.571	< 0.000001	4.08	2.02	5	0.92	122.5	45.84
X	48	1920.83	1.91	0.384	0.548	0.627	10.56	8.66	4.37	3.29	41.38	38.02
X	19	682.34	4.69	0.398	0.674	< 0.000001	6.51	1.81	13.42	0.98	206.18	54
X	26	893.46	3.7	0.398	0.458	< 0.000001	5.9	2.21	10.56	1.11	178.8	50.24
X	41	1517.4	3.07	0.398	0.397	0.0000489	6.6	9.68	5.84	8.19	88.51	84.6
X	3	508.29	1.73	0.398	0.427	0.000521	5.1	3.37	4.53	2.71	88.71	80.41
X	68	4409.17	0.49	0.398	0.525	0.0000306	1.23	0.74	1.29	0.71	104.36	94.81
X	47	1853.84	2.06	0.436	0.333	0.00021	7.29	5.23	4.53	5.68	62.14	108.75
X	52	2164.65	2	0.442	0.584	0.00036	6.32	4.32	5.95	2.98	94.15	68.89
X	43	1719.66	4.17	0.476	0.525	0.0922	14.28	18.45	10.52	12.8	73.68	69.38
X	36	1293.24	2.29	0.476	0.502	0.0000248	6.07	8.36	5.17	7.8	85.21	93.19
X	60	2970.87	1.24	0.476	0.584	0.231	5.4	4.15	3.91	2.92	72.39	70.39
X	16	660.61	0.98	0.476	0.674	0.000467	2.8	3.78	2.15	3.47	76.6	91.86
X	51	2087.38	0.8	0.476	0.584	0.604	4.37	5.17	1.86	2.59	42.52	50.17
X	55	2491.13	0.56	0.517	0.502	0.663	3.05	3.61	1.64	1.91	53.69	52.87
X	61	3169.39	1.08	0.542	0.49	0.0602	5.81	6.89	2.74	4.56	47.1	66.15
X	66	3385.86	0.28	0.593	0.866	0.000932	1.35	1.06	1.29	0.67	95.53	63.18
X	32	1025.14	0.94	0.631	0.502	0.118	5.36	6.3	4.38	3.65	81.71	57.87
X	21	719.73	0.52	0.631	0.617	0.0443	3.76	4.29	2.2	2.27	58.56	52.87
X	24	829.87	0.55	0.688	0.945	0.000967	4.07	3.52	3.14	2.58	77.11	73.37
X	59	2856.88	0.31	0.688	0.832	0.000873	2.28	1.97	1.84	1.39	80.93	70.77
X	44	1730.65	0.54	0.761	0.651	0.324	6.2	6.74	4.09	3.53	65.94	52.29
X	31	1018.18	1.16	0.88	0.832	0.00113	18.05	19.21	17.64	14.96	97.77	77.91
X	54	2346.44	0.06	0.945	1	0.513	3.54	3.48	1.98	2.16	55.92	62.03

Table S2

Statistical analysis of *Ephemera* larvae. S – inclusion (X)/exclusion (-) state of the peak, Index – peak index, Mass – m/z value, DAve – difference between the maximal and the minimal average peak area/ intensity of all groups, PTTA – p-value of t-test and analysis of variance (ANOVA), PWKW – p-value of Wilcoxon test, PAD – p-value of Anderson-Darling test, Ave1 – average peak intensity of *Ephemera lineata* (peak by peak), Ave2 – average peak intensity of *Ephemera vulgate* (peak by peak), StdDev – the standard deviation of the peak area/ intensity average of group, CV – means coefficient of variation of group.

Use of different peak picking methods (monoisotopic mass - manually, and average mass - ClinProTools) caused the differences between m/z value of similar peaks from manually (e.g. m/z 1252.0, see Table 1) and ClinProTools data (e.g. m/z 1252.5, see Supplementary data Table S2).

ClinProt Peak Statistic of fish



ClinProtTools Version: 3.0 build 22
Number of peaks: 114
Sort Mode: p value ms

S	Index	Mass	DAve	PTTA	PWKW	PAD	Ave1	Ave2	Ave3	Ave4	Ave5	Ave6	Ave7	Ave8	StdDev1	StdDev2	StdDev3	StdDev4	StdDev5	StdDev6	StdDev7	StdDev8	CV1	CV2	CV3	CV4	CV5	CV6	CV7	CV8
X	87	1050.8	35.08	0.00000126	0.0718	0.00001102	2.97	3.27	1.89	36.97	2	22.86	10.93	31.91	1.51	0.95	0.65	0.15	0.06	3.8	11.58	5.03	51	29.06	34.35	0.4	2.78	16.63	105.99	15.76
X	43	667.34	5.2	0.00000351	0.0858	0.000647	1.37	6.58	2.59	2.49	4.6	1.53	6.43	2.05	0.01	0.83	0.58	NaN	0.42	0.28	4.36	0.65	0.68	12.6	22.55	NaN	9.13	18.11	67.79	31.59
X	98	1346.9	9.44	0.000185	0.198	0.00103	3.16	5.99	3.63	4.21	5.56	12.6	5.71	3.82	1.04	2.72	1.32	0.21	1.53	0.22	2.3	0.75	32.93	45.39	36.29	1.97	27.58	1.74	40.35	19.54
X	106	1578.32	72.82	0.000641	0.0718	<0.000001	74.95	3.47	2.12	12.14	2.76	5.05	6.57	9.08	11.7	0.27	0.16	0.27	0.96	0.47	0.81	2.17	15.61	7.7	7.74	2.26	34.82	9.26	12.3	23.93
X	77	936.7	20.35	0.000671	0.0718	<0.000001	1.52	1.97	1.38	2.2	2.42	2.73	2.85	21.73	0.67	0.3	0.32	0.18	0.41	0.91	0.49	0.85	44.14	15.21	23.46	8	17.11	33.25	17.29	3.9
X	83	1023.76	71.54	0.000887	0.0718	<0.000001	3.76	1.46	1.28	2.62	1.88	15.86	6.73	72.82	1.07	0.19	0.18	0.02	0.02	3.22	5.47	17.83	28.48	12.73	13.82	0.9	1.24	20.31	81.33	24.48
X	56	750.51	18.29	0.000887	0.0718	<0.000001	1.41	2.17	1.41	1.53	2.52	9.26	1.82	19.7	0.37	0.41	0.23	0.33	0.75	0.3	0.51	7.56	26.49	19.03	16.14	21.75	29.56	3.26	28.13	38.4
X	57	751.48	7.74	0.000887	0.0718	0.0000207	1.47	2.07	1.34	1.52	3	4.98	1.85	9.08	0.54	0.15	0.37	0.08	0.6	0.13	0.3	3.01	37.02	7.12	27.51	5.5	20.11	2.54	16.05	33.17
X	34	594.37	7.13	0.000887	0.0718	0.0404	1.35	4.27	1.85	1.55	4.96	8.48	1.89	6.49	0.52	0.6	0.59	0.52	1.94	0.14	0.82	1.75	38.25	14.07	31.78	33.25	39.07	1.61	43.4	27
X	81	1005.62	33.72	0.0016	0.0888	<0.000001	2.05	5.63	2.38	35.5	3.89	1.78	15.02	2.05	0.83	0.81	1.21	1.44	1.71	0.22	19.7	0.13	40.75	14.4	50.6	4.06	44.01	12.17	131.19	6.19
X	73	866.56	34.03	0.00439	0.0718	<0.000001	3.17	1.95	1.23	29.08	2.48	4.24	2.35	35.26	2.34	0.48	0.13	16.24	0.96	0.14	0.67	28.65	73.73	24.73	10.36	55.84	38.67	3.3	28.7	81.25
X	42	666.44	14.81	0.0048	0.0718	0.0000865	1.31	16.12	3.93	3.99	6.87	2.08	15.12	2.32	0.33	1.77	0.92	0.32	0.64	0.18	10.66	0.8	25.06	10.95	23.43	7.95	9.29	8.73	70.45	34.41
X	97	1323.01	48.12	0.00517	0.0718	<0.000001	49.93	2.66	1.81	3.74	2.58	5.87	4.33	7.73	6.1	0.65	0.27	0.33	0.38	0.25	0.36	3.02	12.23	24.38	14.7	8.92	14.84	4.26	8.33	39.04
X	75	892.55	26.12	0.00517	0.0718	<0.000001	1.26	3.29	1.46	2	3.27	1.54	4.24	27.38	0.34	0.76	0.32	0.29	0.12	0.08	3.23	12.59	27.11	23.23	22.03	14.37	3.71	5.2	76.29	45.99
X	46	681.47	17.9	0.00517	0.0718	<0.000001	1.56	2.19	1.61	1.97	2.71	19.33	1.43	5.78	0.27	0.36	0.64	0.13	0.87	6.48	0.38	0.32	17.02	16.42	39.79	6.73	31.71	33.55	26.31	5.53
X	64	778.4	14.99	0.00517	0.156	<0.000001	1.6	1.93	1.36	16.32	2.65	1.33	1.92	1.41	0.64	0.22	0.4	1.01	0.81	0.13	0.68	0.17	39.87	11.59	29.29	6.2	30.4	9.99	35.34	12.18
X	104	1507.24	65.09	0.00747	0.0718	<0.000001	67.24	3.04	2.15	3.3	2.55	10.71	5.6	8.5	7.87	0.66	0.3	0.07	0.47	0.87	0.65	3.21	11.7	21.69	13.83	2.26	18.28	8.08	11.69	37.79
X	96	1252.81	14.99	0.00924	0.0718	0.0000562	3.72	3.49	1.49	3.94	2.52	16.48	5.55	13.01	0.14	0.47	0.19	0.19	1.15	3.44	1.38	4.51	3.78	13.62	12.91	6.1	45.46	29.86	24.86	34.65
X	82	1020.79	108.5	0.0107	0.0849	<0.000001	1.75	3.07	2.95	1.87	2.06	2.24	110.25	3.27	0.72	0.38	1.6	0.07	0.4	0.65	101.44	0.12	40.99	12.33	54.23	3.7	19.62	28.89	92.01	3.75
X	66	787.54	4.57	0.0107	0.0718	0.00122	1.47	3.01	1.51	3.66	5.25	1.53	5.93	1.36	0.83	0.1	0.45	0.95	2.84	0.42	2.51	0.12	56.22	3.36	29.42	25.87	54.05	27.41	43.39	9.03
X	10	520.16	3.82	0.0107	0.166	0.156	1.92	3.95	2.06	5.24	5.74	5.49	3.38	3.94	0.28	0.11	0.15	1.29	2.11	3.78	1.73	1.35	14.39	2.68	7.4	24.67	36.73	68.74	51.24	34.35
X	74	891.56	47.39	0.0122	0.0718	<0.000001	1.32	2.21	1.15	2.73	2.21	1.79	2.4	48.54	0.43	0.22	0.09	0.29	0.48	0.31	0.11	25.72	32.28	9.79	7.53	10.63	21.69	17.6	4.42	52.98
X	90	1110.78	33.77	0.0122	0.0718	<0.000001	1.91	2.29	1.74	2.18	2.57	5.63	7.4	25.5	0.7	0.79	0.48	0.01	0.07	1.07	2.51	3.26	36.53	34.54	27.4	0.62	2.79	19	33.98	12.78
X	45	673.11	4.04	0.0122	0.257	0.524	2.05	3.09	3.39	5.41	5.68	5.18	5.01	6.09	0.28	0.75	1.83	0.19	2.51	0.62	2.31	2.89	13.62	24.32	54.03	3.46	44.23	11.92	46.12	47.39
X	65	779.52	6.02	0.0125	0.0849	0.132	1.75	7.49	3.66	5.23	6.26	1.47	4.17	1.67	0.75	1.4	1.17	0.28	1.62	0.56	3.08	0.34	42.88	18.7	32.1	5.32	25.82	37.86	73.97	20.36
X	108	1766	52.99	0.0152	0.0718	<0.000001	2.5	1.66	2.13	54.65	2	3.75	3.69	4.18	0.24	0.59	0.12	6.22	0.3	0.23	0.76	1.48	9.51	35.51	5.73	11.38	15.03	6.21	20.62	35.47
X	70	854.53	5.84	0.016	0.0718	0.0000571	1.38	7.22	1.44	1.47	4.15	3.53	2.22	2.15	0.27	4.53	0.13	0.34	1.78	0.19	0.92	0.16	19.47	62.76	8.82	16.48	43.01	5.3	41.25	7.41
X	9	519.16	3.94	0.0184	0.177	0.728	1.28	3.48	2.35	4.54	4.63	5.22	3.57	4.17	0.09	0.82	0.29	0.49	1.93	2.67	2.41	0.63	6.76	23.7	12.24	10.7	41.61	51.01	67.68	15.02
X	41	656.17	8.14	0.0197	0.0765	0.0365	1.72	2.79	2.3	8.47	4.85	9.85	4.55	7.94	0.39	0.42	0.94	1.04	1.71	0.98	2.38	4.18	22.76	15.2	40.7	12.32	35.32	9.91	52.42	52.62
X	112	4719.91	1.74	0.0199	0.0718	0.00287	0.43	1.62	2.08	0.95	0.57	1.36	0.4	0.34	0.02	0.2	0.42	0.7	0.37	0.22	0.11	0.08	5.03	12.35	20.08	73.15	65.38	16.6	27.96	24.02
X	85	1034.88	51.18	0.0219	0.0718	<0.000001	1.52	52.7	1.9	6.4	12.23	7.13	5.34	7.56	0.86	7.16	0.23	1.88	11.57	2.33	3.69	1.28	56.79	13.59	12.19	29.31	94.6	32.72	69.07	17.33
X	69	850.61	20.54	0.0226	0.0718	0.0000279	1.5	22.04	1.52	6	7.71	1.95	5.11	2.28	0.46	3.31	1.27	0.98	2.8	0.03	4.39	0.24	30.93	15.02	27.98	16.4	36.36	1.65	85.79	10.67

X	71	855.51	3.95	0.0226	0.0858	0.000314	5.62	5.14	1.67	1.69	3.42	3.49	2.43	2.08	5.11	1.89	0.38	0.58	0.52	0.13	0.42	0.13	90.92	36.74	22.81	34.12	15.22	3.85	17.22	6.09
X	32	587.25	3.83	0.0226	0.15	< 0.000001	6.11	3.69	2.28	2.81	6.09	2.32	2.53	2.82	4.84	0.11	0.23	0.08	3.41	1.01	0.94	0.17	79.23	2.86	9.99	2.86	56.03	43.24	37.3	5.95
X	80	979.75	26.98	0.0318	0.0718	< 0.000001	1.79	2.39	1.33	2	2.42	3.92	2.4	28.31	0.72	0.87	0.22	0.65	0.27	0.32	0.25	9.29	40.07	23.99	16.57	2.37	11.04	8.07	10.6	32.83
X	79	949.64	11.18	0.0318	0.0718	0.0000966	1.28	4.9	1.97	12.46	2.61	3.03	3.13	5.48	0.36	1.87	0.92	1.5	0.57	0.42	1.21	2.91	27.97	38.17	46.89	12	21.95	13.86	38.75	53.06
X	111	2089.79	17.16	0.0377	0.0718	0.000268	2	18.57	12.04	4.34	19.16	4.11	3.56	2.08	0.51	10.94	2.86	0.31	12.78	0.22	2.32	0.46	25.63	58.93	23.74	7.04	66.73	5.28	65.13	22.13
X	102	1478.86	14.7	0.0391	0.0718	0.011	3.91	13.88	3.21	2.74	7.02	11.62	4.95	17.44	0.5	3.95	0.75	0.09	2.88	1.96	1.28	4.13	12.79	28.49	23.52	3.43	41.09	16.89	25.83	23.68
X	100	1477.07	36.18	0.0397	0.0718	< 0.000001	2.91	2.46	1.7	2.86	2.66	24	4.05	37.88	0.5	0.66	0.44	0.03	0.22	4.2	0.66	9.84	17.21	26.96	25.55	1.18	8.17	17.49	16.23	25.98
X	91	1112.66	15.23	0.0397	0.0718	0.00063	1.95	5.62	1.65	16.88	3.03	8.09	6.07	13.91	1.12	0.56	0.37	15.85	0.52	3.23	1.78	5.76	57.98	9.92	22.52	82.05	17.19	39.98	29.29	41.32
X	89	1100.82	7.75	0.0397	0.102	0.00399	4.22	9.1	1.97	9.06	4.09	3.56	9.71	4.71	3.25	3.36	0.35	1.58	2.59	0.34	9.15	0.38	76.91	36.96	17.72	17.4	63.24	9.48	94.21	8.01
X	3	504.21	3.91	0.0397	0.181	0.00378	1.93	4.01	2.42	3.54	5.85	3.78	3.18	4.37	0.22	0.25	0.07	0.73	2.76	1.91	1.21	1.44	11.49	6.17	3.01	20.71	47.29	50.6	37.85	32.87
X	88	1062.9	49.22	0.0399	0.0858	0.0037	2.95	57.15	6.13	27.21	6.73	9.12	16.76	16.79	1.5	19.28	7.78	3.69	7.76	2.84	12.55	5.66	51.16	36.97	127.04	13.56	115.57	31.11	74.88	33.74
X	101	1477.85	28.91	0.0399	0.0718	< 0.000001	2.02	2.03	1.45	2.71	2.28	19.33	5.09	30.36	0.21	0.06	0.46	0.42	0.42	3.88	1.97	7.86	6.86	3.05	31.93	15.54	18.43	20.07	38.65	25.88
X	78	946.45	11.75	0.0399	0.0718	0.00637	1.75	5.83	3.27	2.48	8.5	4.38	2.09	13.5	0.88	2.92	1.71	0.46	2.15	0.25	0.61	3.14	50.35	50.04	53.09	18.69	28.86	5.74	29.28	23.29
X	53	706.34	8.93	0.0399	0.0858	0.0000164	1.4	3.1	2.26	10.33	3.74	1.72	2.22	2.42	0.01	0.38	0.63	5.04	0.95	0.24	0.95	0.36	0.36	12.2	28.1	48.8	25.38	14.16	42.8	14.85
X	114	4893.61	5.01	0.0407	0.0719	< 0.000001	0.46	2.17	5.38	0.69	0.63	0.44	0.41	0.37	0.14	1.49	0.78	0.17	0.49	0.08	0.09	0.01	31.29	68.99	14.55	25.17	77	17.13	22.76	4.05
X	44	672.12	13.2	0.0416	0.108	0.169	3.67	5.08	5.76	14.49	9.75	13.7	11.93	16.86	2.2	1.64	3.78	0.91	3.74	2.77	5.5	7.92	60.11	32.23	65.67	6.29	38.32	20.19	46.12	47
X	49	684.44	8.32	0.0461	0.0718	0.000159	1.29	8.14	2.75	2.55	9.62	1.72	1.72	3.24	0.14	3.79	1.42	0.97	1.66	0.36	0.54	1.35	10.58	46.48	51.84	37.85	17.21	21.09	31.61	41.53
X	103	1486.82	17.09	0.0529	0.0718	< 0.000001	2.72	3.98	2.11	8.69	3.86	4.24	19.2	3.5	0.51	1.37	0.28	3.79	0.7	0.35	5.42	0.66	18.62	34.45	13.43	43.59	18	8.24	28.23	18.95
X	113	4820.76	2.98	0.0529	0.0718	0.00000198	0.52	1.38	3.27	0.9	0.47	0.48	0.43	0.29	0.1	0.72	0.57	0.48	0.25	0.12	0.09	0.02	18.96	52.07	17.29	53.5	53.3	24.17	21.73	7.93
X	72	858.55	6.32	0.0538	0.11	0.000961	3.26	7.85	1.9	2.06	6.2	1.53	3.6	2.13	2.8	3.7	0.56	0.43	1.14	0.03	2.17	0.19	85.77	47.16	29.41	20.77	18.38	1.92	60.21	8.8
X	68	837.54	39.13	0.0559	0.0718	< 0.000001	2.7	2.3	1.32	3.35	2.69	22.67	5.04	40.45	1.07	0.19	0.24	1.65	0.47	7.31	2.1	10.63	39.69	8.45	18.29	40.33	17.66	32.24	41.78	26.28
X	22	550.17	17.7	0.0559	0.132	0.0866	4.63	7.99	5.69	12.96	8.22	13.88	14.24	22.34	4.08	3.96	4.9	11.03	3.43	5.97	7.08	1.76	87.96	49.63	86.17	85.11	41.78	43.02	49.73	7.88
X	92	1163.71	10.59	0.0644	0.0718	< 0.000001	1.85	2.77	1.51	2.43	2.67	4.98	12.1	3.68	0.34	0.24	0.35	0.71	0.98	1.06	11.4	0.33	18.44	8.6	22.93	29.38	36.85	21.27	94.17	8.94
X	26	559.36	12.47	0.0654	0.0718	0.00000164	1.88	13.65	4.5	2.5	14.35	2.67	2.34	3.28	0.18	4.54	1.78	0.46	7.81	1.11	0.84	0.27	9.49	33.29	39.5	18.55	54.43	41.38	35.79	8.36
X	99	1347.78	7.2	0.0654	0.0719	0.00677	3.26	5.84	2.63	2.96	5.36	9.84	8.89	4.06	1.04	1.55	0.49	0.23	2.94	1.26	4.83	1.24	31.94	26.5	18.66	7.63	54.84	12.86	54.35	30.46
X	94	1182.8	48.63	0.0699	0.0718	< 0.000001	1.77	2.79	1.48	3.6	1.65	4.31	4.22	50.12	0.77	0.6	0.11	1.28	0.29	0.81	0.74	27.21	43.42	21.37	7.67	35.59	17.32	18.91	17.6	54.3
X	29	570.21	1.59	0.0699	0.37	0.0698	3.38	5.19	3.96	6.16	6.71	6.53	7.07	7.97	0.68	1.62	2.2	4.36	2.93	3.6	3.62	0.4	20.13	31.25	55.68	70.78	43.72	55.04	51.24	5.02
X	35	595.37	13.62	0.0715	0.11	< 0.000001	1.21	5.36	3.22	2.89	14.83	3.43	1.78	4	0.47	2.3	1.54	1.71	10.73	0.13	0.52	0.76	38.99	42.87	47.69	59.11	72.32	3.69	29.4	18.97
X	105	1523.87	14.91	0.0746	0.0719	< 0.000001	2.36	2.67	4.5	3.56	2.93	3.68	17.27	5.74	0.06	0.43	3.24	0.27	0.45	0.73	12.96	1.32	2.33	16.1	71.98	7.57	15.41	19.96	75.03	22.97
X	11	522.16	4.96	0.0752	0.238	0.0608	2.78	7.43	3.79	6.18	6.88	7.06	5.84	7.74	0.8	1.55	0.74	2.77	2.43	3.01	3.08	0.73	28.77	20.81	19.54	44.8	35.36	42.56	52.69	9.46
X	17	535.15	9.5	0.0756	0.155	0.0000376	1.83	3.58	2.65	5.97	7.85	11.33	3.56	3.91	0.22	0.94	0.43	0.83	4.15	11.66	2.14	1.24	12.14	26.38	16.1	13.81	52.81	102.91	60.04	31.72
X	39	608.31	5.91	0.0756	0.0769	0.0214	1.75	6.09	3.09	2.77	7.67	2.32	2.57	3.08	0.64	0.99	1.12	0.21	1.85	0.37	1.07	0.74	36.53	16.23	36.18	7.54	24.13	15.79	41.38	23.92
X	8	516.23	5.41	0.0761	0.334	0.00457	7.8	4.6	2.39	4.37	5.98	3.56	3.81	3.79	7.2	0.32	0.32	1.21	2.44	1.27	1.8	1.22	92.37	7.05	13.29	27.62	40.91	35.6	47.25	32.13
X	54	710.39	7.71	0.0791	0.0718	0.0000353	2.48	2.32	1.63	2.17	3.71	9.33	1.86	3.61	1.2	0.31	0.44	0.37	0.56	1.84	0.2	1	48.51	13.39	27.04	16.99	15.05	19.7	10.94	27.85
X	109	1864.37	22.58	0.0793	0.0718	0.00000395	7.15	25.52	15.6	2.94	5.3	2.94	3.31	3.78	3.1	18.7	3.16	0.4	1.73	0.25	0.4	1.38	43.4	73.28	20.25	13.48	32.61	8.39	12.05	36.45
X	19	544.26	4.09	0.0793	0.264	0.000136	6.1	3.77	2.01	2.45	4.85	2.25	2.28	2.37	6.13	0.25	0.31	0.64	1.96	0.87	1.07	0.38	100.44	6.75	15.62	26.26	40.41	38.59	46.9	15.97
X	1	501.44	33.59	0.0824	0.269	0.006	7.08	9.89	17.68	36.4	13.77	13.97	10.36	2.81	6.54	1.07	12.56	24.63	6.11	16.53	8.35	1.15	92.38	10.8	71.05	67.67	44.35	118.38	80.6	40.8
X	52	704.08	7.6	0.0824	0.11	0.0641	2.71	4.69	4.43	3.08	9.47	1.87	5.7	4.3	1.11	1.85	2.34	0.73	2.13	0.28	1.95	2.2	40.81	39.51	52.85	23.59	22.51	14.77	34.16	48.83
X	58	753.51	6.69	0.0824	0.0718	0.0000175	3.44	8.11	1.42	1.81	3.94	1.5	2.02	2.01	0.33	4.77	0.4	0.16	1.52	0.39	0.79	0.47	9.72	58.83	28.25	9.05	38.58	26.19	39.06	23.47
X	63	777.44	32.94	0.0878	0.102	< 0.000001	1.62	1.82	1.06	33.99	3.01	1.3	2.16	1.68	0.3	0.46	0.21	7.41	1.12	0.31	0.72	0.39	18.51	25.16	20.05	21.79	37.32	23.65	33.58	23.3
X	48	683.42	10.44	0.088	0.0718	0.000104	1.18	7.68	1.27	7.25	3.35	3.04	1.82	11.63	0.24	5.46	0.38	5	0.52	0.96	0.79	7.55	20.38	71.04	29.92	68.94	15.42	31.59	43.25	64.94

X	15	526.2	5.97	0.088	0.339	0.11	2.25	6.22	3.89	6.01	6.5	6.85	5.78	8.22	1.09	2.1	2.01	2.82	3.01	3.47	3.18	0.4	48.66	33.74	52.42	46.98	46.38	50.67	54.95	4.85
X	93	1171.69	12.88	0.108	0.118	<0.000001	15.81	5.7	3.41	3.88	6.06	2.93	5.53	4.08	11.55	2.43	0.97	0.09	0.77	0.28	2.66	0.48	73.02	42.7	28.52	2.3	12.72	9.57	46.12	11.77
X	13	524.18	18.41	0.11	0.315	0.0314	5.68	15.27	8.29	15.36	13.28	18.4	17.54	24.09	3.97	11.01	7.31	14	7.28	10.34	8.82	1.31	69.85	72.07	88.15	91.16	54.8	56.18	50.3	3.44
X	36	600.41	5.05	0.11	0.104	0.00432	1.94	4.42	2.18	6.99	5.08	2.79	6.07	4.13	0.15	0.93	0.19	3.69	1.1	1.19	4.67	1.82	7.8	21.06	8.55	52.8	21.72	42.83	76.94	44.09
X	16	532.29	3.6	0.11	0.159	0.0583	2.7	4.65	2.7	3.51	6.3	3.2	4.05	3.72	0.06	0.57	0.4	0.9	2.3	1.62	1.44	0.68	2.35	12.18	14.87	25.62	36.46	50.59	35.64	18.18
X	6	508.2	3.29	0.11	0.387	0.176	2.29	4.47	3	4.51	5.34	4.73	4.17	5.58	0.6	1.39	0.93	1.59	2.2	2.27	2.01	0.45	26.16	31.03	30.92	35.29	41.25	47.93	48.13	8.07
X	62	771.48	7.06	0.111	0.0718	0.0156	1.58	6.38	4.74	2.99	8.38	1.31	4.43	1.58	0.42	2.73	2.76	1.08	2.11	0.42	2.06	0.21	26.52	42.75	58.23	36.09	25.24	31.58	46.38	13.43
X	38	604.24	4.35	0.114	0.297	0.0178	5.24	4.2	2.3	4.23	6.65	3.31	3.32	4.28	2.5	0.22	0.43	0.59	3.07	1.48	1.59	0.44	47.77	5.33	18.67	14	46.13	44.84	47.77	10.28
X	24	552.21	3.42	0.122	0.264	0.262	2.05	3.77	2.35	3.19	5.47	3.08	3.33	4.04	0.28	0.69	0.95	1.13	2.02	1.52	1.55	0.47	13.76	18.39	40.36	35.35	36.94	49.28	46.39	11.75
X	76	916.36	22.86	0.123	0.0888	<0.000001	21.46	2.41	1.6	3.58	2.87	2.08	3.48	2.45	22.9	0.97	0.25	0.67	0.55	0.16	1.25	0.1	93.61	40.37	15.46	18.82	19.31	7.74	35.88	4.13
X	61	770.55	7.17	0.123	0.0888	0.00109	1.54	8.56	2.12	3.41	6.47	1.38	7.72	1.86	0.62	5.17	0.88	3.01	2.21	0.03	4.83	0.19	40.36	60.45	41.52	88.04	34.15	1.83	62.51	10.04
X	55	731.43	8.02	0.13	0.164	<0.000001	1.46	2.27	1.59	1.72	2.85	9.49	1.82	1.59	0.53	0.28	0.51	0.2	1.09	2.03	0.8	0.22	36.15	12.24	32.1	11.43	38.25	21.42	44.25	13.66
X	40	617.34	4.5	0.13	0.0888	0.00177	1.7	3.88	1.89	3.05	3.88	2.76	5.85	6.2	0.89	0.97	0.34	0.12	1.24	0.42	3.51	3.16	52.23	24.95	18.11	3.87	32.04	15.36	60.05	50.99
X	95	1226.08	89.5	0.147	0.0718	0.0262	4.26	62.51	5.77	93.76	8.12	29.59	66.13	54.05	1.82	38.48	4.24	56.72	9.14	16.17	39.31	20.57	42.72	61.56	73.55	60.5	112.52	54.63	59.45	38.06
X	23	551.18	5.82	0.147	0.198	0.156	2.46	4.31	3.53	6.03	1.99	5.88	5.93	8.28	1.02	0.61	2.28	3.77	2.07	2.63	3.17	1.14	41.33	14.76	64.52	62.48	41.53	44.72	53.44	15.73
X	14	525.19	5.77	0.147	0.434	0.0235	3.34	6.27	4.39	7.31	6.35	6.48	6.11	9.11	1.54	3.73	2.54	4.17	2.75	2.95	2.98	0.06	46.22	59.42	57.96	56.95	43.37	45.45	48.73	0.61
X	86	1038.8	14.38	0.152	0.0718	<0.000001	1.47	15.86	2.02	2.65	10.23	2.41	3.22	2.54	0.33	14.89	0.36	0.04	11.94	0.39	1.68	0.12	22.19	93.9	17.89	1.49	116.64	16.23	52.16	4.81
X	47	682.4	31.15	0.152	0.0718	<0.000001	1.47	2.28	1.23	20.71	3.18	9.05	2.84	32.38	0.54	0.34	0.3	16.06	1.21	3.07	2.92	21.17	36.56	14.97	24.51	77.55	37.9	33.87	102.73	65.37
X	107	1736.21	21.12	0.152	0.0718	<0.000001	8.05	24.35	4.89	3.43	6.57	4.15	3.23	3.26	6.1	10.3	0.75	0.24	3.07	0.14	0.68	1.26	75.77	42.29	15.43	7.14	46.81	3.46	21.18	38.46
X	25	558.4	31.03	0.164	0.0718	0.0000575	2.41	33.07	10.04	2.15	19.79	2.04	2.27	2.05	1.19	11.48	3.76	0.18	12.23	0.88	0.84	0.04	49.3	34.71	37.44	8.33	61.83	43.35	37.04	1.86
X	4	506.18	15.58	0.164	0.154	0.0785	5.56	9.5	4.8	11.6	9.28	14.12	15.56	20.38	3.67	6.72	2.41	6.58	4.45	4.5	9.03	4.2	66.02	70.73	50.22	56.78	48	31.88	58.05	20.63
X	2	502.39	10.96	0.171	0.244	0.0386	3.12	5.29	7.97	13.52	6.31	5.33	4.72	2.56	2.37	0.41	4.76	7.11	2.56	5.27	1.79	0.72	75.9	7.66	59.75	52.59	40.49	98.81	38.01	28.27
X	67	835.58	8.27	0.177	0.11	0.0000249	2.11	3.31	1.61	2.87	2.96	2.97	6.24	9.88	1.69	0.41	0.44	1.27	0.58	0.22	4.41	6.68	79.97	12.26	27.1	44.37	19.51	7.25	70.69	67.66
X	50	688.08	9.88	0.178	0.11	0.103	3.87	6.96	7.1	7.93	13.75	5.86	12.92	12.42	1.34	1.6	4.35	0.05	4.55	1.55	5.32	3.53	34.55	22.92	61.22	0.67	33.11	26.5	41.16	28.38
X	27	568.18	65.42	0.179	0.214	0.064	15.75	29.36	21.83	49.89	30.99	58.34	56.99	81.17	16.23	18.07	18.92	48.43	15.97	32.94	27.69	14.58	103.11	61.55	86.67	97.08	51.55	56.46	48.95	17.93
X	28	569.18	21.19	0.18	0.24	0.0864	5.52	11.28	8.57	17.45	13.47	19.8	17.42	26.72	4.74	5.63	7.44	16.48	7.04	11.79	9.35	4.31	85.8	49.88	86.8	94.4	52.26	59.58	53.65	16.11
X	21	547.32	9.77	0.185	0.0718	0.0000353	1.88	11.66	4.53	2.79	10.51	2.02	2.49	2.34	0.58	7.42	2.06	0.22	3.21	0.82	0.42	0.57	30.58	63.61	45.56	7.83	30.58	40.35	16.9	24.34
X	31	586.26	9.2	0.188	0.144	<0.000001	11.27	3.88	2.07	2.19	5.89	2.48	2.22	2.37	12.09	0.57	0.61	0.15	2.46	1.53	0.81	0.43	107.27	14.74	29.47	6.9	41.75	61.81	36.57	18.25
X	18	537.25	7.2	0.236	0.159	0.00109	1.83	3.43	2.93	3.81	9.02	5.72	2.52	2.38	0.48	1.34	1.18	1.89	4.26	5.88	1.4	0.69	26.18	24.71	40.4	49.46	47.19	102.71	55.48	28.92
X	59	760.29	23.98	0.24	0.0718	<0.000001	25.58	2.27	1.61	3.19	3.18	7.02	2.83	8.02	27.9	0.66	0.36	1.16	1.45	1.98	1.35	4.43	109.06	29.21	22.59	36.4	45.7	28.14	47.77	55.19
X	37	603.23	7.16	0.241	0.11	0.064	3.48	4.69	3.11	9.39	10.27	6.9	6.94	9.91	2.23	0.39	0.81	3.29	6.88	2.44	2.39	3.33	64.19	82.2	25.98	35.06	66.95	35.37	34.47	33.64
X	110	1974.09	7.63	0.249	0.0718	0.000109	2.67	2.57	3.37	8.81	2.7	5.92	10.19	9.22	0.73	0.85	0.5	2.78	0.34	2.08	6.22	4.52	27.31	32.94	14.79	31.53	12.65	35.03	61.05	49.02
X	51	689.07	3.22	0.255	0.387	0.285	2.43	3.29	3.31	2.97	5.64	3.59	4.2	4.43	0.13	0.94	1.88	0.45	2.02	0.86	1.64	1.3	5.2	28.52	56.78	15	35.76	24.03	38.92	29.33
X	20	546.36	32.74	0.299	0.0718	0.0000584	2.41	34.84	11.16	2.71	24.28	2.38	2.09	2.53	1.07	27.68	6.86	0.16	10.75	1.34	0.84	0.77	44.36	79.44	61.52	5.82	44.3	56.48	40.3	30.46
X	30	584.35	8.77	0.316	0.0849	0.0000539	2.06	7.28	3.71	3.24	10.43	1.95	2.08	1.66	0.17	3.95	1.58	1.49	5.5	0.75	1.08	0.31	8.18	54.29	42.62	46.06	52.76	38.39	51.76	18.91
X	5	507.17	6.6	0.33	0.209	0.331	3.03	5.27	3.52	6.07	6.25	7.03	6.62	9.62	1.68	1.86	0.89	1.83	2.87	2.86	3.56	2.79	55.37	35.3	25.33	30.13	45.91	40.76	53.84	29.02
X	60	763.51	8.86	0.35	0.132	0.0000207	1.81	2.34	1.36	9.91	2.77	3.36	3.3	3.66	0.76	0.5	0.82	6.81	0.71	1.07	2.99	1.25	41.81	21.16	38.99	68.7	25.82	32	90.74	34.05
X	12	523.19	3.94	0.395	0.387	0.0235	3.02	4.53	3.15	3.47	6.96	3.57	3.44	4.81	0.72	0.54	0.67	1.4	3.34	1.87	1.33	0.82	23.81	11.86	21.28	40.47	48.02	52.41	38.6	17.08
X	33	588.31	1.68	0.447	0.167	0.000601	2.91	6.74	2.27	2.27	6.09	2.06	2.49	3.38	1.49	2.68	0.51	0.24	2.54	1	1.19	0.7	51	30.74	22.29	10.74	41.62	48.79	47.83	29.41
X	7	515.23	19.81	0.505	0.142	<0.000001	22.01	4.04	2.53	2.89	5.45	2.21	2.8	2.2	26.17	1.27	0.72	0.48	2.22	0.84	1.21	0.47	118.93	31.39	28.63	16.65	40.84	37.93	43.17	21.3
X	84	1030.7	47.52	0.763	0.177	<0.000001	2.3	2.09	1.6	2.31	2.16	2.08	49.11	1.79	0.7	0.45	0.57	0.29	0.44	0	55.81	0.42	30.61	21.55	35.56	12.38	20.19	0.16	113.63	23.74

Table S3

Statistical analysis of Fish. S – inclusion (X)/exclusion (-) state of the peak, Index – peak index, Mass – m/z value, DAve – difference between the maximal and the minimal average peak area/ intensity of all groups, PTTA – p-value of t-test and analysis of variance (ANOVA), PWKW – p-value of Wilcoxon test, PAD – p-value of Anderson-Darling test, Ave1-8 – average peak peak intensity of *Anguilla anguilla*, *Rutilus rutilus*, *Abramis brama*, *Carassius gibelio*, *Alburnus alburnus*, *Lepomis gibbosus*, *Pseudorasbora parva*, *Perca fluviatilis* (peak by peak), StdDev – the standard deviation of the peak area/ intensity average of group, CV – means coefficient of variation of group.

ClinProt Peak Statistic of male and female bleaks



ClinProTools 3.0 build 22
 Version:
 Number of peaks: 37
 Sort Mode: p value tta

S	Index	Mass	DAve	PTTA	PWKW	PAD	Ave1	Ave2	StdDev1	StdDev2	CV1	CV2
X	18	961.2	8.7	0.00459	0.0155	0.0584	10.79	2.08	1.97	0.68	18.22	32.74
X	12	779.85	2.98	0.00459	0.0155	0.388	5.09	2.11	0.78	0.61	15.29	28.96
X	21	1063.45	30.11	0.00626	0.0155	0.0302	32.66	2.55	7.04	0.23	21.57	9.01
X	14	850.99	17.05	0.00626	0.0155	0.0797	21.06	4.01	4.56	1.07	21.68	26.69
X	20	1035.26	15.06	0.00626	0.0155	0.21	19.19	4.13	4.8	1.51	25.01	36.54
X	26	1612.65	10.42	0.00626	0.0155	0.0318	13.51	3.09	3.15	0.31	23.35	10.05
X	17	950.1	8.38	0.00626	0.0155	0.108	10.61	2.23	2.51	0.84	23.69	37.67
X	27	1749.85	6.13	0.00626	0.0155	0.224	3.78	9.92	1.45	2.3	38.43	23.19
X	22	1171.29	3.98	0.00626	0.0155	0.388	7.57	3.59	1.12	1.47	14.8	40.86
X	6	565.57	3.12	0.00626	0.0155	0.517	5.98	2.86	0.94	1.1	15.63	38.44
X	23	1179.42	15.15	0.0091	0.0155	0.0476	17.19	2.04	5.15	0.33	29.94	15.93
X	10	666.72	7.45	0.0103	0.0155	0.324	10.76	3.31	2.95	0.96	27.44	28.91
X	25	1226.82	20.84	0.0114	0.0155	0.0429	23.97	3.13	7.86	0.48	32.79	15.43
X	32	2090.35	3.3	0.0143	0.0155	0.957	7	3.7	1.44	1.32	20.53	35.58
X	24	1202.53	14.01	0.019	0.0155	0.0476	17.04	3.03	6.48	0.61	38.04	20.09
X	28	1784.1	10.54	0.019	0.028	0.496	8.64	19.18	3.03	5.4	35.08	28.16
X	35	4822.46	1.53	0.0251	0.0155	0.0429	0.35	1.87	0.08	0.78	23.38	41.63
X	36	4895.74	1.84	0.0416	0.0155	0.0468	0.36	2.2	0.08	1.12	21.66	51
X	37	4933.6	1.04	0.0416	0.0155	0.0302	0.3	1.34	0.03	0.63	9.77	47.15
X	33	4664.2	1.05	0.0523	0.028	0.0398	0.61	1.66	0.08	0.7	12.54	42.29
X	29	1912.98	1.57	0.0823	0.305	0.476	6.16	4.6	1.16	0.91	18.78	19.89
X	34	4721.25	2.74	0.0931	0.0155	0.0188	1.08	3.82	0.13	2.29	12.28	59.91
X	11	721.72	11.53	0.207	0.141	0.00112	1.84	13.37	0.34	13.67	18.65	102.23
X	31	2077.53	3	0.207	0.141	0.352	5.69	8.69	2.33	3.22	40.94	37.11
X	3	546.49	8.35	0.21	0.141	0.00149	2.11	10.46	0.6	10.38	28.35	99.28
X	4	547.46	3.2	0.21	0.141	0.00151	1.96	5.17	0.43	4	21.84	77.35
X	13	808.85	6.97	0.218	0.305	0.00149	2.86	9.83	1.47	9.02	51.46	91.74
X	9	584.58	4.5	0.258	0.502	0.00151	2.68	7.18	0.84	6.47	31.44	90.13

ClinProt Peak Statistic

X	16	946.07	7.85	0.267	0.751	0.00928	6.35	14.2	2.79	11.64	43.88	81.98
X	30	1975.06	1.45	0.312	0.409	0.318	5.54	6.99	0.9	2.36	16.17	33.72
X	8	569.36	3.22	0.379	0.751	0.00174	2.63	5.84	0.91	6.27	34.76	107.42
X	19	984.11	1.53	0.422	0.502	0.34	6.17	4.63	1.51	3.15	24.43	67.9
X	1	506.41	3.65	0.437	0.889	0.0000651	3.12	6.76	0.55	8.69	17.53	128.49
X	5	550.43	3.55	0.437	1	0.0000651	2.36	5.91	0.82	8.25	34.95	139.65
X	7	568.41	3.84	0.494	1	0.0318	6.41	10.26	2.88	10.47	44.94	102.07
X	2	524.41	1.43	0.614	0.751	0.00149	4.22	5.65	1	5.52	23.64	97.58
X	15	916.01	0.74	0.651	0.502	0.0854	6.08	5.34	2.12	2.77	34.92	51.88

1/2

Table S4

Statistical analysis of male and female bleaks. S – inclusion (X)/exclusion (-) state of the peak, Index – peak index, Mass – m/z value, DAve – difference between the maximal and the minimal average peak area/ intensity of all groups, PTTA – p-value of t-test and analysis of variance (ANOVA), PWKW – p-value of Wilcoxon test, PAD – p-value of Anderson-Darling test, Ave1 – average peak intensity of male (peak by peak), Ave2 – average peak intensity of female (peak by peak), StdDev – the standard deviation of the peak area/ intensity average of group, CV – means coefficient of variation of group.

Use of different peak picking methods (monoisotopic mass - manually, and average mass - ClinProTools) caused the differences between m/z value of similar peaks from manually (e.g. m/z 1749.1, see Table 3) and ClinProTools data (e.g. m/z 1749.9, see Supplementary data Table S4).