A peer-reviewed version of this preprint was published in PeerJ on 7 July 2017.

Stopkova R, Klempt P, Kuntova B, Stopka P. 2017. On the tear proteome of the house mouse (*Mus musculus musculus*) in relation to chemical signalling. PeerJ 5:e3541 https://doi.org/10.7717/peerj.3541

Dual roles of tear lipocalins as 'chemical signalling' and 'toxic waste disposal' systems of the house mouse

Pavel Stopka Corresp., [1] , **Petr Klempt** [1] , **Barbora Kuntova** [1] , **Romana Stopkova** [1]

[1] Department of Zoology, Faculty of Science, Charles University, Prague, -- Please Select --, Czech Republic

Corresponding Author: Pavel Stopka
Email address: pstopka@natur.cuni.cz

Mammalian tears are produced by lacrimal glands to protect eyes and to function in chemical communication and immunity. However, excess tears flow through nasolacrimal ductsto nasal tissues, and via the nasopharyngeal duct to the oral cavity where digestion starts. Tears contain soluble proteins that attack pathogens, as well as proteins from the lipocalin family that – with their capacity to transport volatile organic compounds (VOCs) in their eight-stranded beta barrel – are involved in sexual signalling and may also transport toxic VOCs towards digestion. Therefore, we generated the tear proteome of the wild-living house mouse (*Mus musculus musculus*) and detected a total of 719 proteins in tears with 20% being sexually dimorphic. Those proteins that showed the most elevated sexual dimorphisms are VOC transporters from the recently discovered odorant binding protein (OBP), and major urinary protein (MUP) families, thus demonstrating that tears have the potential to elicit sex-specific signals in combination with different lipocalins. Moreover, some tear lipocalins are non-dimorphic – with MUP20/Darcin, LCN11, and LCN13 being good examples – thus suggesting that they are involved in other biological processes besides sexual signalling.

Dual roles of tear lipocalins as 'chemical signalling' and 'toxic waste disposal' systems of the house mouse

Pavel Stopka*, Petr Klempt, Barbora Kuntová, and Romana Stopková
BIOCEV group, Department of Zoology, Faculty of Science, Charles University, Prague, Czech Republic

*corresponding author
E-mail: pstopka@natur.cuni.cz

Abstract
Mammalian tears are produced by lacrimal glands to protect eyes and to function in chemical communication and immunity. However, excess tears flow through nasolacrimal ducts to nasal tissues, and via the nasopharyngeal duct to the oral cavity where digestion starts. Tears contain soluble proteins that attack pathogens, as well as proteins from the lipocalin family that – with their capacity to transport volatile organic compounds (VOCs) in their eight-stranded beta barrel – are involved in sexual signalling and may also transport toxic VOCs towards digestion. Therefore, we generated the tear proteome of the wild-living house mouse (*Mus musculus musculus*) and detected a total of 719 proteins in tears with 20% being sexually dimorphic. Those proteins that showed the most elevated sexual dimorphisms are VOC transporters from the recently discovered odorant binding protein (OBP), and major urinary protein (MUP) families, thus demonstrating that tears have the potential to elicit sex-specific signals in combination with different lipocalins. Moreover, some tear lipocalins are non-dimorphic – with MUP20/Darcin, LCN11, and LCN13 being good examples – thus suggesting that they are involved in other biological processes besides sexual signalling.

Key words
lipocalins, tears, pheromone, sex dimorphism, immunity, toxic waste hypothesis

Introduction

The genome of the mouse contains at least 55 genes for lipocalins (Stopkova et al. 2014; Stopková et al. 2009) that are – due to heir beta barrel structure – able to transport VOCs (Zidek et al. 1999). Almost half of lipocalins belongs to major urinary proteins that are almost exclusively cited within the context of mouse chemical communication (Stopková et al. 2009). Investments in chemical communication are expected to be costly and it is evident that scent-marking signals have strong effects on the reproductive success of the signaller (Thonhauser et al. 2013). In mice, these signals are manifested via expression of large quantities of sexually dimorphic (Stopková et al. 2007) major urinary proteins in the liver. They bind volatile organic compounds (VOCs) in their eight-stranded beta barrel and transport them to the urine (Kwak et al. 2013; Sharrow et al. 2002; Timm et al. 2001), where they act as an honest, cheat-proof display of an individual's health and condition (Zala et al. 2004). VOCs are slowly released from different urinary MUPs, and have been proposed to function in a variety of social signals, including identity, territorial marking, mate choice etc. (Hurst & Beynon 2004; Hurst et al. 2001; Mucignat-Caretta & Caretta 1999; Nelson et al. 2015). Some studies, however, suggest that MUPs may be used as carriers of various degradation products and of potentially toxic waste (Kwak et al. 2011; Kwak et al. 2016), which can be seen as their parallel – and presumably

ancestral – function within the 'Toxic waste hypothesis' (Stopková et al. 2009; Stopkova et al. 2016). Although, the urinary profiles of the wild male house mice *M. m. musculus* are relatively homogenous - i.e. not individually unique (Enk et al. 2016; Thoss et al. 2016; Thoß et al. 2015), their expression is dynamic over time with significant changes after puberty and during adulthood (Thoß et al. 2015). The signals that are transported by MUPs or MUPs themselves have been shown to regulate reproductive behaviour of the receiver (Janotova & Stopka 2011; Ma et al. 1999; Novotny et al. 1986; Roberts et al. 2010; Stopka et al. 2007), MUPs have a predictive value for the onset of aggressive behaviour and dispersal tendency in male wild house mice (Rusu et al. 2008), and one particular MUP - MUP20 or 'Darcin', which in *M. m. domesticus* is prevailingly expressed by males, has been reported to predict the outcome of male-male territorial competition (Nelson et al. 2015), stimulates inherent female attraction for particular males (Roberts et al. 2010), and its level decreases in immune-challenged male mice (Lopes & Konig 2016) which presumably shows that the production of MUP20 / MUPs is costly. In *M. m. musculus,* however, MUP20 was also detected in the saliva of males and females (Stopka et al. 2016). Moreover, MUPs and other lipocalins (e.g. OBPs, LCNs) are also expressed by various oro-facial tissues and glands, including sensory but also lymphoid tissues (Stopka et al. 2016; Stopkova et al. 2016), which further extends their interesting roles besides chemical signalling.

For example, in our recent study, we have determined the mRNA expression sites for a newly described family of odorant binding proteins (OBP) and provided evidence that the extraorbital lacrimal glands produce high quantities of mRNAs coding OBP5, OBP6, OBP7 and also MUP4, MUP5, and LCN11 (Stopkova et al. 2016). Interestingly, most of those lipocalins that are produced by olfactory, vomeronasal, and nasal-associated lymphoid tissues, are finally transported to the oral cavity where digestion starts (Stopka et al. 2016). Furthermore, lacrimal glands contain large quantities of transporters of chemical signals essential for sexual signalling during lacrimation, and then during selfgrooming the signals are spread onto the fur with saliva (i.e. containing also lacrimal signals) or move further to the digestive tract. Interestingly, when lacrimal glands are removed it impairs sexual behaviour (Cavaliere et al. 2014). In tears, MUPs are particularly important for their affinity to several biologically active compounds where MUP4 revealed strong affinity to the male-derived pheromone 2-*sec*-butyl-4,5-dihydrothiazole – SBT (Sharrow et al. 2002) which causes inter-male aggression and estrus synchrony (Jemiolo et al. 1986; Novotny et al. 1985). Lacrimal expression of *Mup4* and the presence of MUP4 with its ligands in the mouse tears (i.e. along with other signals) and saliva (Stopka et al. 2016) may explain the observation of Luo et al. (Luo et al. 2003; Luo & Katz 2004), who reported that mouth and facial areas are the first and the most frequently investigated areas during mouse social contacts. Moreover, these areas are investigated longer and more frequently in comparison with investigation of the anogenital region. Furthermore, the facial areas elicit strong neuronal activity responses in accessory olfactory bulbs, whilst the investigation of the anogenital region does not(Luo et al. 2003).

Along with the above-described roles of MUP4 in estrus synchronization, the role of exocrine gland-secreted peptides (ESPs) was shown to be in parallel with MUP4 roles. The ESP1, peptide pheromone from male tears also activates the vomeronasal neurons of female mice and enhances sexual receptive behaviour through a specific vomeronasal receptor (Kimoto et al. 2005). A completely opposite role, however, is exhibited by another peptide - ESP22, which when present

in tears of juveniles protects them from adult male mating behaviour (Ferrero et al. 2013). The *Esp* family includes members in which expression is both sexually-dimorphic and strain-specific (Kimoto et al. 2007). Because of the strain-specificity we have realized that given findings should be further investigated in wild-living house mouse subspecies, thus avoiding experiments with laboratory mice. This is due to the differential contribution of blocks of genes from the two subspecies *M. m. domesticus* and *M. m. musculus* to current laboratory strains (Abril et al. 2002) that may mask natural intra- and inter-specific differences. Furthermore, tears have also important protective and antimicrobial roles.

It has been well documented that tears keep exposed and non-vascularized parts of the eyeball healthy and hostile to pathogens (Walcott 1998; Zoukhri 2006). Therefore, numerous are proteins secreted in the aqueous layer of tears and each of them has specific functions. For example, secretory IgA inhibits pathogen adhesion, phospholipase A2 hydrolyses phospholipids in bacterial membranes and various growth factors maintain cornea proliferation and regeneration, reviewed in (Fluckinger et al. 2004). Specific antimicrobial activity has been demonstrated for the mouse lipocalin LCN2, which is up-regulated as a response to inflammation in mucosal tissues (Flo et al. 2004; Goetz et al. 2002). Thus, a strategy called "nutritional immunity" prevents pathogens from acquiring host iron (Porcheron et al. 2013), which is an essential nutrient, but only small amounts of free iron are accessible. Therefore, bacteria acquire iron by secretion of high-affinity iron sequestrating siderophores. The mammalian host, however, limits this process by the production of LCN2 (Goetz et al. 2002) which efficiently scavenges for catecholate-type siderophores (Flo et al. 2004), and is equally expressed by individuals of both sex in mouse saliva (Stopka et al. 2016). Other mechanisms of defence involve bactericidal proteins from the PLUNC (palate, lung, and nasal epithelium clone) protein family, defending the mucosal layers of the body against pathogenic microbiota. These include for example the bactericidal/permeability-increasing proteins - BPI (Leclair 2003a; LeClair 2003b). In mouse saliva, BPI abundances are male biased and include BPIA1, BPIB1, BPIB2, BPIB3, BPIFA2, BPIFB5, BPIFB9B (Stopka et al. 2016).

The aim of this paper was to characterize the tear proteome from wild individuals of the house mouse (*M. m. musculus*) to detect abundant and sexually-dimorphic proteins potentially involved in sexual signalling, as well as those that are monomorphic and, thus, may have other interesting biological roles.

Materials and Methods

Ethical Standards

All animal procedures were carried out in strict accordance with the law of the Czech Republic paragraph 17 no. 246/1992 and the local ethics committee of the Faculty of Science, Charles University in Prague chaired by Dr. Stanislav Vybíral specifically approved this study in accordance with accreditation no. 27335/2013-17214 valid until 2019. Animals were sacrificed by cervical dislocation.

Animals

Fourteen individuals of the House mouse (the eastern form, *M. m. musculus*) used in this study were captured in the Czech Republic near Bruntál - 49.9884447N, 17.4647019E (1male; 1 female), in Velké Bílovice - 48.8492886N, 16.8922736E (3 males; 3 females), Prague-Bohnice -

50.1341539N, 14.4142189E (3 males; 3 females). All animals were trapped in human houses and garden shelters. On the day of capture or the next day, all animals were transferred to our animal facility. Each animal was caged individually with *ad libitum* access to water and food.

Sample collection

Eye lavage was used as a non-invasive method of tear collection. Eyes were carefully rinsed with 10 µl of the saline physiology solution by a gentle pipetting. The process was repeated three times with at least a two hour interval between every rinsing, and each sample was analysed twice with MS to produce mean values from the methodology duplicates. This was done in the 'in-house' Mass Spectrometry and Proteomics Service Laboratory, Faculty of Science, Charles University in Prague.

Individual mice were sacrificed next day after the tear sampling. The exorbital lacrimal glands were dissected and immediately placed into RLT buffer (Qiagen) and homogenised in MagNALyser (Roche) for 30s at 6000rpm. RNA was isolated using the RNeasy Mini Kit (Qiagen) according to the manufactures protocol with on-column DNase I treatment. The purity and concentration of eluted RNA was measured with a NanoDrop ND1000. The quality of RNA was checked on agarose gel electrophoresis (AGE). RNA was stored at -70°C pending further use.

Protein Digestion

Protein samples were precipitated with the ice-cold acetone and followed by a re-suspension of dried pellets in the digestion buffer (1% SDC, 100mM TEAB – pH=8.5). Protein concentration of each lysate was determined using the BCA assay kit (Fisher Scientific). Cysteines in 20µg of proteins were reduced with a final concentration of 5mM TCEP (60° C for 60 min) and blocked with10mM MMTS (i.e. S-methyl methanethiosulfonate, 10 min Room Temperature). Samples were cleaved with trypsin (i.e. 1/50, trypsin/protein) in 37°C overnight. Peptides were desalted on a Michrom C18 column.

nLC-MS^2 Analysis

Nano Reversed phase columns were used (EASY-Spray column, 50 cm x 75 µm ID, PepMap C18, 2 µm particles, 100 Å pore size**)**. Mobile phase buffer A was composed of water, 2% acetonitrile and 0.1% formic acid. Mobile phase B contained 80% acetonitrile, and 0.1% formic acid. Samples were loaded onto a trap column (Acclaim PepMap300, C18, 5 µm, 300 Å Wide Pore, 300 µm x 5 mm, 5 Cartridges) for 4 min at 15 µl/min loading buffer was composed of water, 2% acetonitrile and 0.1% trifluoroacetic acid. After 4 minutes ventile was switched and Mobile phase B increased from 2% to 40% B at 60 min, 90% B at 61 min, hold for 8 minutes, and 2% B at 70 min, hold for 15 minutes until the end of run.

Eluting peptide cations were converted to gas-phase ions by electrospray ionization and analysed on a Thermo Orbitrap Fusion (Q-OT-qIT, Thermo). Survey scans of peptide precursors from 400 to 1600 m/z were performed at 120K resolution (at 200m/z) with a 5×10^5 ion count target. Tandem MS was performed by isolation at 1.5 Th with the quadrupole, HCD fragmentation with normalized collision energy of 30, and rapid scan MS analysis in the ion trap. The MS^2 ion count target was set to 10^4 and the max injection time was 35ms. Only those precursors with charge state 2–6 were sampled for MS^2. The dynamic exclusion duration was set to 45s with a 10ppm

tolerance around the selected precursor and its isotopes. Monoisotopic precursor selection was turned on. The instrument was run in top speed mode with 2s cycles.

Protein analysis

All data were analysed and quantified with MaxQuant software (version 1.5.3.8) (Cox et al. 2014). The false discovery rate (FDR) was set to 1% for both proteins and peptides and we specified a minimum peptide length of seven amino acids. The Andromeda search engine was used for the MS/MS spectra search against the Uniprot *Mus musculus* database (downloaded on June, 2015), containing 44,900 entries. Enzyme specificity was set as C-terminal to Arg and Lys, also allowing cleavage at proline bonds(Rodriguez et al. 2008) and a maximum of two missed cleavages. Dithiomethylation of cysteine was selected as fixed modification and N-terminal protein acetylation and methionine oxidation as variable modifications. The "match between runs" feature of MaxQuant was used to transfer identifications to other LC-MS/MS runs based on their masses and retention time (maximum deviation 0.7 min) and this was also used in all quantification experiments. Quantifications were performed with the label-free algorithms described recently (Cox et al. 2014) using a combination of unique and razor peptides. To detect differentially expressed / abundant proteins, we used the Power Law Global Error Model (PLGEM) (Pavelka et al. 2004) within the *Bioconductor package* in R software (Gentleman et al. 2004).

Expression profile of the mouse exo-orbital lacrimal gland

We conducted 454 RNA-sequencing with a desktop pyro-sequencer GS Junior from Roche using the long reads mode. The sequencing was conducted on whole lacrimal gland of four wild-caught adult female and four adult male biological replicates. To increase the precision of transcript mapping we excised from a gel and sequenced only transcripts between ~400 and 1300 bp. Transcripts of this length include those of genes, described for their involvement in chemical communication. This method is amenable to further analyses because the nebulization step is skipped and, therefore, whole transcripts instead of their fragments are further pyro-sequenced and mapped. We estimated particular expression levels from the number of uniquely mapped transcripts assigned to each annotated gene.

Size-selected transcriptome preparation

cDNA was prepared using the SMARTer PCR cDNA Synthesis Kit (Clontech) and amplified with Advantage 2 PCR Kit (Clontech). Both procedures were handled according to protocol for Trimmer-2 Normalization Kit (Evrogen). The products of optimalized cDNA amplification were then loaded on AGE. For each sample, only the area of product in range from ~400bp to ~1300bp (well visible area full of bands) was excized from the gel and the DNA products were extracted using the Gel/PCR DNA Fragments Extraction Kit (Geneaid). Appropriate amounts of size-selected products were then secondarily amplified according to the recommended protocol from Evrogen. Products of secondary amplification were purified using MiniElute PCR Purification Kit (Qiagen). Purified products (and the range where they emerge) were checked on AGE. Purity was analysed with NanoDrop ND1000. Concentration was measured/determined using Quant-it Pico Green dsDNA Assay Kit (Invitrogen) and fluorimeter (Hoefer DQ 300).

Rapid Library Preparation and GS Junior Transcriptome Sequencing

Rapid Library (RL) was prepared for each transcriptome (4 males and 4 females) according to Rapid Library Preparation Manual (my454.com). Equal amounts from each of 8 Rapid Libraries (10^7 molecules per µl dilution) were mixed and then used for emPCR. Further steps followed the provider´s instructions for sequencing with GS Junior (Roche; emPCR Amplification Method Manual Lib-L and Sequencing Method Manual, my454.com). We obtained >165000 high quality (HQ) reads. HQ 454 Reads were multiplexed, trimmed (i.e. using a trimming database that contains primers used for library preparations), filtered and aligned into contigs against *Mus musculus* cDNA database ("the super-set of all known, novel and pseudo gene predictions"; ensembl.org, 17-FEB-2015 version) and using GS Reference Mapper (Roche). Differential expression was analysed in R software using the *DEseq* routine within the *Bioconductor package* (Gentleman et al. 2004).

RNA-seq data availability

The transcriptome data is provided as bam files in 'Sequencing Read Archive' (www.ncbi.nlm.nih.gov/sra) under the accession number – SRP063762, BioProject: PRJNA295909.

Protein surface modelling

The surface electrostatics modelling involved several steps. First, we downloaded the structures from the RSCB Protein Data Bank (http://www.rcsb.org/) under accession IDs: 3S26, 1I04 and 2L9C, respectively. Because the mouse OBP1 structure has no record in the database we had to predict it with i-TASSER (Iterative Threading ASSEmbly Refinement) program (http://zhanglab.ccmb.med.umich.edu/I-TASSER/) with the rat ortholog OBP1F (PDB ID: 3FIQ, 76% similarity) as template for the homologous modelling. Next, we used PyMOL - Molecular Graphic System (version 1.7.0.0) with APBS (Adaptive Poisson-Boltzman Solver) plugin to model the electrostatics with the default software settings.

Results

The tear proteome and the level of sexual dimorphism

We have generated the tear proteome of the house mouse, *M. m. musculus* and detected a total of 719 proteins at 0.01 FDR (i.e. False Discovery Rate for all peptides and proteins). Successful identifications of these proteins resulted from a relatively high number of peptides per identification (6.5±8.1, mean±sd), sequence coverage (24.1±19.2%), and unique sequence coverage (20.1±17.4%).

Next, we searched for differentially abundant proteins between males and females using the Power Law Global Error Model (PLGEM) (Pavelka et al. 2004). This model was first developed to quantify microarray data (Pavelka et al. 2004), however, due to similar statistical properties – namely the distribution of signal values deviating from normality – it has proved to be an amenable model for the quantification of label-free MS-based proteomics data (Pavelka et al. 2008). First of all, we reduced our data such that only the proteins that were detected in three or more individuals were used (i.e. 457 proteins). Next, we calculated the signal-to-noise ratio – STN (equation provided in (Pavelka et al. 2008)), because it explicitly takes unequal variances into account and because it penalizes proteins that have higher variance in each class more than those proteins that have a high variance in one class and a low variance in another (Pavelka et al.

2004). Because PLGEM can only be fitted on a set of replicates from the same experimental condition we have done this for female data. Correlation between the mean values and standard deviations was high (r^2= 0.96, Pearson=0.94) so we continued with the resampled STNs and calculated differences with corresponding p-values between males and females.

PLGEM analysis of the level of sexual dimorphism revealed that 68 (14.9%) out of 457 proteins identified at 1% FDR and p<0.05 were sexually dimorphic, Fig. 1. Male biased proteins included 36 (7.8%) and female biased proteins included 32 (7%) successful identifications. Thus, male-biased proteins were not more common than female-biased proteins in the tear proteome of the house mouse subspecies *M. m. musculus*. The most dimorphic proteins included the female-biased OBP5, and OBP7, the male-biased MUP4, the male-only ESP1, male-biased ESP36, and several male-biased secretoglobins (SCGB1B19, SCGB1B20/25, SCGB1B24, SCGB1B3, SCGB2A2 – Mammaglobin, SCGB2B3, SCGB2B7). Kallikrein 1-related peptidases were also significantly sexual dimorphic (i.e. female-biased), however, this pattern (though significant) was not consistent across all the females tested, Fig. 2. Interestingly, we have also detected sexually dimorphic BPI proteins. Bactericidal/permeability-increasing proteins (BPI) are ~50kDa proteins that are a part of the innate immune system, and have an antibacterial activity against the gram-negative bacteria (LeClair 2003b). We have detected three BPIs, of which BPIFA2 was male biased, BPIFA6 was female biased, whilst males and females equally expressed BPIFB9B.

The most abundant tear proteins

Based on the median value we sorted our data to detect the most abundant proteins in the tear proteome, Fig. 2. We have filtered out potential contaminants such as keratins and also trypsins which are the enzymes that cleave all peptides before LC-MS in this study. The top 5% of the most abundant proteins that characterize the soluble tear proteome of the mouse are depicted in Fig. 2, and include for example the female-biased lipocalins OBP5, OBP7, the unbiased lipocalins OBP1 and LCN11, and the male-biased lipocalin MUP4. Other proteins dominating the soluble tear-proteome included three male-biased secretoglobins (SCGB1B3, SCGB1B20, SCGB2B20/ SCGB2B27), two unbiased secretoglobins (SCGB1B2, SCGB2B2), male-biased carbonic anhydrase 6 (CAH6), (unbiased) exocrine secreted peptide ESP6, Lacrein, and female-biased prolactin inducible protein (PIP). Interestingly, out of the top 5% most abundant proteins, 50% of them (i.e. 11) were significantly sexual dimorphic. Thus, even though the level of sexual dimorphism is rather low within the complete tear proteome (i.e. 15%), those few proteins that were most abundant were often the most sexually dimorphic.

Sex-unique proteins

It is always difficult to presume that some peptides/proteins are sex unique, because some of them might have been lost during data filtering at the level of a particular FDR or they may be below the limit of the equipment detection. Thus, we provide visual representation of all proteins using MA plot, also including potentially sex-unique proteins (Fig. 1), where significant points are colored from green (p<0.05) to blue (p<0.01). Female-unique proteins included S10A8/S10A9 which are calcium- and zinc-binding proteins and which play important roles in the regulation of inflammatory processes and immune responses, and can induce neutrophil chemotaxis and adhesion (Vogl et al. 2007). We have also detected the secretoglobin SCGB2B20 which is female-unique in tears under this study but male-biased in the saliva proteome(Stopka et al. 2016), thus suggesting that there are multiple sources of expression of

this protein. We have also detected the female-unique kallikreins KLK1B22, KLK1B1, and KLK1B3. They are, however, female-biased in the mouse saliva (Stopka et al. 2016) and not unique. Other female-unique proteins invlolved RENI2, LIPR1 and one keratin (KT33A).

Male-unique proteins included the Secretoglobin SCGB1B19, the exocrine gland-secreted peptide ESP1, ZA2G (i.e. Zn-Alpha2-Glycoprotein), CUZD1 (zona pellucida-like domain-containing protein 1), and products of two predicted genes Gm12887 and Gm1330. Previously we provided evidence that ESP1 is male-biased in mouse saliva and that it presumably cannot function as a male pheromone if females produce this peptide too. ESP1 has been described as a 7kDA male-specific signalling protein in the laboratory mouse and was named as the exocrine gland-secreted peptide-1 or ESP1 (Kimoto et al. 2005; Kimoto et al. 2007). ESP1 is produced by the mouse lacrimal glands, secreted with tears and when experimentally transferred to the female vomeronasal organ, it stimulates V2R-expressing vomeronasal chemosensory neurons, and thus elicits an electrical response (Kimoto et al. 2005). In mouse tears, ESP1 was also male-unique and co-expressed with other ESPs (ESP3, ESP4, ESP6, ESP15, ESP16, ESP18, ESP34, ESP38). Previously we have suggested that ESP1 (and potentially also other ESPs) may simply be involved in the defence system against bacteria because their structures have a strong electrostatics antipathy (Stopka et al. 2016). This theory, however, needs to be further tested *in vitro* with different cultures of pathogens.

Transcriptome: mRNAseq based analysis of exo-orbital lacrimal glands

We used transcriptomic analysis to detect the most likely site for tear protein expression. We also searched for a sexually dimorphic expression pattern with the *DESeq* routine within the *Bioconductor package* (Gentleman et al. 2004) to detect protein-coding transcripts that may account for sex-specific differences. We have filtered for further analysis only the data where the sum of counts per row ≥ 10. Then, we normalised the data with a size factor vector to make the libraries comparable. Because *DESeq* calculates sexual dimorphisms from the original non-transformed number of counts we first looked at the level of variation between replicates within sex. When dispersion values are plotted against the means of the normalised counts (Fig. 3b) it is evident from the slope of the red fitting curve that data with a low mean of normalized counts have higher levels of dispersion than high expression data.

Having estimated the dispersion for each gene we next performed the analysis of differentially expressed genes by calling the *nbinomTest* in *DESeq*. The resulting pattern is plotted using MA plot (Fig. 3c) with red colouring of those genes that are significant at FDR=0.1 (i.e. False discovery rate). Significantly female-biased genes with a p-value < 0.05 at FDR=0.1 include *Obp5*, *Obp7, Obp8, Spt1, Hba,* and *Scgb2b1*. Similarly, male biased genes with a the p-value < 0.05 at FDR=0.1 included for example *Mup4, Esp1, Esp16, Esp18* and several secretoglobins graphically demonstrated with the heat-map in Fig. 3d.
Next we asked which of the above sex-biased genes are most differentially expressed. Using p-adjusted values ($p<0.05$) these genes included a total of 13 genes with female-biase *Obp5, Obp7,* and *Spt1*, whilst male-biased genes included the male-biased *Mup4,* six male-biased secretoglobins, and two ESPs (ESP16, ESP18). Potentially interesting data though marginally significant or with potential trends and those that were not sexually dimorphic but still highly expressed are provided as a Supplementary Dataset.

Anti-microbial peptides

BPI proteins have an antibacterial activity against gram-negative bacteria (LeClair 2003b). The saliva proteome contains seven members of the bactericidal/permeability-increasing proteins (i.e. BPI (Leclair 2003a; LeClair 2003b)) which are male biased (Stopka et al. 2016) and include BPIA1, BPIB1, BPIB2, BPIB3, BPIFA2, BPIFB5, BPIFB9B (Stopka et al. 2016). However, tears only contain BPIFA2/*Bpifa2*, BPIFA6, BPIFB9B. Thus, we searched for other proteins/peptides which may have similar roles due to their amphipathic structural properties or proteolytic activities. Recently, WFDC proteins (i.e. 'Whey acidic proteins four disulphide core') were shown to have anti-microbial properties (Scott et al. 2011) and the two members WFDC12 and WFDC18 are present in mouse saliva as proteins encoded by submandibular gland transcripts (i.e. *Wfdc12*, *Wfdc18*) (Stopka et al. 2016). In this study, we have detected WFDC12 and WFDC18 as transcripts of the extraorbital lacrimal glands (i.e. *Wfdc12*, *Wfdc18*), but only WFDC18 was detected in tears on the proteomic level and just in two males. Our results, however, provide evidence that the major antimicrobial protein in tears is TRFL (Lactotransferrin). Lactotransferrin also known as lactoferrin (LF) has antimicrobial properties (bactericidal, fungicidal) and is a part of the innate immune system, mainly at mucoses (Sanchez et al. 1992). In the tear proteome, we detected TRFL as one of the most abundant proteins and similar amounts were previously also detected in mouse saliva (Stopka et al. 2016).

Supporting evidence for the theory entitled 'The origin of chemical communication by means of toxic-waste perception'

John Maynard Smith and David Harper defined a signal as '…any act or structure which alters the behaviour of other organisms, which evolved because of that effect, and which is effective because the receiver's response has also evolved' (Maynard Smith & Harper 2003). Thus, evolution of chemical communication seems to require two steps. However, the 'toxic waste hypothesis' (Stopkova et al. 2014; Stopková et al. 2009) or as we rephrased it here as the theory entitled 'The origin of chemical communication by means of toxic waste perception' requires only one step because it presupposes that only the receiver's response has evolved as an adaptation to already existing sources of individual VOCs/odours which resulted from metabolic degradation. Moreover, this theory expects that the level of degradation correlates with energy intake and immune system efficiency, and thus reflects an inherent quality of the signaller.

The tear proteome of the house mouse provides an essential support for this hypothesis on several levels. First, the house mouse tears – as we show in this study – contain a wide spectrum of anti-microbial peptides/proteins that yield various products of bacterial degradation and it is known that female mice are able to recognize infected males (Zala et al. 2015; Zala et al. 2004). Mouse tears also contain various products of ocular lipid peroxidation such as 4-Hydroxynon-2-enal (HNE). HNE excess amounts cause chronic inflammation, however, it has an affinity to a binding pocket of OBPs (OBP5, OBP7)(Grolli et al. 2006) which – in tears – presumably diminish ocular damage by transporting HNE to the oral cavity. Second, we have detected group-A and group-B MUPs in the mouse tears – namely MUP4, MUP5, MUP10 and yet another unspecified group-B MUP member (most likely MUP17, identification provided in Supplementary Dataset). MUPs (via their ligands) are known to elicit behavioural responses in the receiver and at the same time they are known to transport toxic substances out of the body (Kwak et al. 2016). Third, we show that several proteins from the lipocalin family are produced by lacrimal glands (OBPs, MUPs) but move to the oral cavity where they were detected as

proteins in the saliva but not as submandibular-gland transcripts (Stopka et al. 2016). Thus, there is a continuous flush of liquid containing proteins from several orofacial tissues to the oral cavity. This is also evidenced by several lipocalins that are produced exclusively by VNO (e.g. LCN3, LCN4) but are also detectable in high quantities in the oral cavity where digestion starts (Stopka et al. 2016). All together, it is evident that some (if not all) lipocalins have dual functions – in that they are preferentially used for sexual signalling (e.g. group-B MUPs) and thus they are sexually dimorphic in some tissues (e.g. in the liver/urine, saliva), or they are not sexually dimorphic or are less dimorphic and may aid to removing those toxic VOCs that are not recognized as signals (e.g. lacrimal glands/tears).

Discussion

Tears are a source of chemical signals involved in sexual signalling, however, most studies to date concentrated on particular genes or group of genes for putative chemical signals and ignored or underestimated the roles of others. Thus, we focused on the detection of differentially abundant proteins in mouse tears with label-free LC-MS/MS techniques to obtain more complex and hopefully also more detailed comparative view. Furthermore, we assumed that sex-specific differences that we detected may have roles in sexual signalling - i.e. a process which is driven by sexual selection, whilst those that are not sexually dimorphic may have other – presumably ancestral – roles.

The most interesting result of this study is evidence that males differ from females by a cocktail-like composition of significant sexually dimorphic genes. Previously, we have demonstrated on the level of mRNA, that lacrimal glands produce high quantities of *Mup4, Lcn11, Obp5, Obp6,* and *Obp7* transcripts in both subspecies of the house mouse *M. m. domesticus* and *M. m. musculus* (Stopkova et al. 2016). This led us to an idea that sex-specific and sex-biased expression of several different lipocalins is combinatorial, thus differentially contributing to individual scents. The combinatorial and context dependent effect of signalling (i.e. different composition of urinary MUPs yields different behavioural responses) has recently been described for the urinary MUPs in mice (Kaur et al. 2014). In tears, similar effects may potentially be achieved by differential expression of eleven lipocalins detected in this study with abundances being unique for each sex and with a notable variation between individuals. Even a greater potential for a combinatorial mode of lipocalin functioning was recently documented in mouse saliva (Stopka et al. 2016), where we detected 20 (out of 55) mouse lipocalins belonging to the groups of LCNs (LCN2, LCN3, LCN4, LCN11, LCN12, LCN13, LCN14), OBPs (OBP1, OBP2, OBP5, OBP6, OBP7 (Stopkova et al. 2014; Stopkova et al. 2016)), and MUPs (MUP4, MUP5, MUP6, MUP8, MUP14, MUP17, MUP20, MUP21). A total of 10 salivary lipocalins (50%) was significantly sexually dimorphic (OBP1, OBP2, LCN3, LCN4, LCN13, LCN14, MUP4, MUP8, MUP14, and MUP20). Only MUP8 was female biased, while all other sexually dimorphic lipocalins were male biased (Stopka et al. 2016). Moreover, the assumption that different pheromone transporters may have complementary roles is supported by our recent study showing that MUPs and OBPs have different biochemical properties with OBPs being less hydrophobic and having higher iso-electric points than MUPs that are more acidic and hydrophobic (Stopkova et al. 2016). In Fig. 4, we provide representative structures from homology modelling. It is notable that different lipocalins have similar structures but different electrostatics properties. These differences correspond to the previously detected biochemical

differences (Stopka et al. 2016), however, the electrostatics modelling in Fig. 4 demonstrates that the distribution of negative and positive residua is not random in OBP1 and even less so LCN2. Their structures are amphipathic and may fit the description of antimicrobial peptides (i.e. similar to CRAMP (Gallo et al. 1997)). LCN2 (already) is antimicrobial, as it efficiently scavenges for catecholate-type siderophores (i.e. such as enterochelin, mycobactin) which bacteria produce to scavenge for free iron (Flo et al. 2004). However, such amphipathic structure may aid to a direct attack upon bacterial membranes by its oppositely charged protein surface. On the other hand, MUP structures are not amphipathic but rather homogeneously negative, which corresponds to their low pI. Further details about MUP structures are provided or reviewed elsewhere (Phelan et al. 2014).

To follow the main aim of this study, we are providing evidence that female tears have a unique protein content which is relatively different from that of males. The number of female-biased and male-biased proteins was almost equal. Females in this study, however, were characteristic in producing higher quantities of OBP5, OBP7, OBP8 (OBP8 was detected only on the level of *Obp8* transcript) and SPT1, whilst males produced more secretoglobins (i.e. as in the laboratory mouse (Karn & Laukaitis 2015)), exocrine-secreted peptides (ESPs) and MUPs (i.e. MUP4, MUP5). Male biased expression of mRNA coding MUP4 has already been reported by Shaw et al. (Shaw et al. 1983) for its affinity to the male-derived pheromone 2-*sec*-butyl-4,5-dihydrothiazole (SBT). It is also known that MUP4's ligands cause intermale aggression and estrus synchrony (Sharrow et al. 2002). Along with the sexually dimorphic MUPs we have also detected the major urinary protein MUP20 or 'Darcin' which was previously detected only in males of *M. m. domesticus* and which was demonstrated to stimulate female attraction for particular males (Roberts et al. 2010). In *M. m. musculus* in this study MUP20 was found in male and female tears and because the tear content is continuously moving via naso-lacrimal ducts to the nasal, vomeronasal, and oral cavities, it is difficult to imagine that this protein may function as a pheromone in this sub-species.

Mammalian OBPs were thought to include only a few genes per species (Cavaggioni & Mucignat-Caretta 2000; Nagnan-Le Meillour et al. 2014; Pes et al. 1992; Pes & Pelosi 1995). However, Stopkova et al. (Stopkova et al. 2014) demonstrated that there are more genes and/or predicted transcripts for odorant binding proteins in the mouse genome (Stopková et al. 2009; Stopkova et al. 2010) and, therefore, provided alternative names based on their position on chromosome X as *Obp1*, *Obp2*, *Obp5* (synonym in C57Bl – *Obp1a* (Pes et al. 1998)), *Obp6*, *Obp7* (synonym in C57Bl – *Obp1b* (Pes et al. 1998)), and *Obp8*, where *Obp3* and *Obp4* are pseudogenized. In this study we have detected the expression of *Obp1/*OBP1, *Obp2*, *Obp6/*OBP6, and sexually dimorphic *Obp5/*OBP5, *Obp7/*OBP7, and *Obp8* in lacrimal glands/tears. OBP8, which is 99% similar to OBP7, was not detected on the level of protein (most likely) due to a low incidence of unique peptides for OBP8.

Potential secretory roles suggested for OBPs in mice became clearer when Stopkova et al. (Stopkova et al. 2014; Stopková et al. 2009; Stopkova et al. 2010) provided bioinformatics evidence that in other rodent taxa there exist true orthologs of mouse X-linked OBPs with a CxxxC motif which are involved in chemical communication (Stopkova et al. 2010). These include hamster Aphrodisin a protein pheromone transporter present in female vaginal secretion that elicits copulatory behaviour in males (Abril et al. 2002; Singer et al. 1986) via ligands that

Aphrodisin-OBPs transport. In pig some OBPs undergo extracellular protein modifications (e.g. glycosilation) to finely modulate their specificities to odours and pheromones (Nagnan-Le Meillour et al. 2014). The C57Bl mouse OBP member OBP1a (synonym in *M. m. musculus* – OBP5) was implicated in playing a major role in rapid internalization of OBP-odorant complexes into lysosomes and to scavange for toxic products of free radical exposure (Grolli et al. 2006; Strotmann & Breer 2011). However, due to the sexual dimorphism detected in this study, the presence of OBP5, OBP7, and OBP8 in tears implies roles that are parallel to ligand internalization. It is likely that some OBPs (e.g. non-dimorphic OBP1, OBP2, OBP6) are required for the internalization of degradation products or for transport of these harmful substances to the oral cavity where digestion starts (Stopka et al. 2016), whilst those that were detected as sexually dimorphic (i.e. the female-biased OBP5, OBP7, OBP8) could be essential for female sexual signalling with products of metabolic degradation that correlate with inherent quality of the receiver. This hypothesis, however, requires further testing.

To conclude, mammalian tears contain various proteins with protective roles to keep eyes healthy. In mice, however, several tear proteins function as pheromone transporters. Because most studies focused on male pheromones and how they influence female behaviour, we were interested in whether female tears also contain proteins important for signalling. Thus we have generated the tear proteome and identified that female-biased proteins are as frequent as male-biased proteins. Furthermore, we are providing evidence that female mice produce the recently identified odorant binding proteins rather than major urinary proteins, exocrine gland-secreted peptides, and secretoglobins, which are produced by males.

Acknowledgement

We are very grateful to Helena Uhlířová for her careful and patient lab assistance and to Karel Harant and Pavel Talacko from the Mass Spectrometry and Proteomics Service Laboratory, Faculty of Science, Charles University in Prague for performing the LC-MS/MS run.

References

Abril JF, Agarwal P, Alexandersson M, Antonarakis SE, Baertsch R, Berry E, Birney E, Bork P, Bray N, Brent MR, Brown DG, Butler J, Bult C, Chiaromonte F, Chinwalla AT, Church DM, Clamp M, Collins FS, Copley RR, Couronne O, Cawley S, Cuff J, Curwen V, Cutts T, Daly M, Dermitzakis ET, Dewey C, Dickens NJ, Diekhans M, Dubchak I, Eddy SR, Elnitski L, Emes RD, Eswara P, Eyras E, Felsenfeld A, Flicek P, Frankel WN, Fulton LA, Furey TS, Gnerre S, Glusman G, Goldman N, Goodstadt L, Green ED, Gregory S, Guigo R, Hardison RC, Haussler D, Hillier LW, Hinrichs A, Hlavina W, Hsu F, Hubbard T, Jaffe DB, Kamal M, Karolchik D, Karlsson EK, Kasprzyk A, Keibler E, Kent WJ, Kirby A, Kolbe DL, Korf I, Kulbokas EJ, III, Kulp D, Lander ES, Letunic I, Li M, Lindblad-Toh K, Ma B, Maglott DR, Mauceli E, Mesirov JP, Miller W, Mott R, Mullikin JC, Ning Z, Pachter L, Parra G, Pevzner P, Poliakov A, Ponting CP, Potter S, Reymond A, Roskin KM, Sapojnikov V, Schultz J, Schwartz MS, Schwartz S, Searle S, Singer JB, Slater G, Smit A, Stabenau A, Sugnet C, Suyama M, Tesler G, Torrents D, Tromp J, Ucla C, Vinson JP, Wade CM, Weber RJ, Wheeler R, Winter E, Yang S-P, Zdobnov EM, Waterston RH, Whelan S, Worley KC, and Zody MC. 2002. Initial sequencing and comparative analysis of the mouse genome. *Nature* 420:520-562.

552 Cavaggioni A, and Mucignat-Caretta C. 2000. Major urinary proteins, a_{2U}-globulins and
553 aphrodisin. *Biochimica et Biophysica Acta* 1482:218-228.
554 Cavaliere RM, Ghirardi F, and Tirindelli R. 2014. Lacrimal gland removal impairs sexual
555 behavior in mice. *Front Neuroanat* 8:101.
556 Cox J, Hein MY, Luber CA, Paron I, Nagaraj N, and Mann M. 2014. Accurate proteome-wide
557 label-free quantification by delayed normalization and maximal peptide ratio extraction,
558 termed MaxLFQ. *Molecular & Cellular Proteomics* 13:2513-2526.
559 Enk VM, Baumann C, Thoss M, Luzynski KC, Razzazi-Fazeli E, and Penn DJ. 2016. Regulation
560 of highly homologous major urinary proteins in house mice quantified with label-free
561 proteomic methods. *Mol Biosyst* 12:3005-3016.
562 Ferrero DM, Moeller LM, Osakada T, Horio N, Li Q, Roy DS, Cichy A, Spehr M, Touhara K,
563 and Liberles SD. 2013. A juvenile mouse pheromone inhibits sexual behaviour through
564 the vomeronasal system. *Nature* 502:368-371.
565 Flo TH, Smith KD, Sato S, Rodriguez DJ, Holmes MA, Strong RK, Akira S, and Aderem A.
566 2004. Lipocalin 2 mediates an innate immune response to bacterial infection by
567 sequestrating iron. *Nature* 432:917-921.
568 Fluckinger M, Haas H, Merschak P, Glasgow BJ, and Redl B. 2004. Human tear lipocalin
569 exhibits antimicrobial activity by scavenging microbial siderophores. *Antimicrob Agents*
570 *Chemother* 48:3367-3372.
571 Gallo RL, Kim KJ, Bernfield M, Kozak CA, Zanetti M, Merluzzi L, and Gennaro R. 1997.
572 Identification of CRAMP, a cathelin-related antimicrobial peptide expressed in the
573 embryonic and adult mouse. *J Biol Chem* 272:13088-13093.
574 Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, Ellis B, Gautier L, Ge Y,
575 Gentry J, Hornik K, Hothorn T, Huber W, Iacus S, Irizarry R, Leisch F, Li C, Maechler
576 M, Rossini AJ, Sawitzki G, Smith C, Smyth G, Tierney L, Yang JY, and Zhang J. 2004.
577 Bioconductor: open software development for computational biology and bioinformatics.
578 *Genome Biol* 5:R80.
579 Goetz DH, Holmes MA, Borregaard N, Bluhm ME, Raymond KN, and Strong RK. 2002. The
580 Neutrophil Lipocalin NGAL Is a Bacteriostatic Agent that Interferes with Siderophore-
581 Mediated Iron Acquisition. *Molecular Cell* 10:1033-1043.
582 Grolli S, Merli E, Conti v, Scaltriti E, and Ramoni R. 2006. Odorant binding protein has the
583 biochemical properties of a scavenger for 4-hydroxy-2-nonenal in mammalian nasal
584 mucosa. *FEBS Journal* 273:5131-5514.
585 Hurst JL, and Beynon RJ. 2004. Scent wars: the chemobiology of competitive signalling in mice.
586 *BioEssays* 26:1288-1298.
587 Hurst JL, Payne CE, Nc., Marie AD, Humphries RE, Robertson DHL, Cavaggioni A, and
588 Beynon RJ. 2001. Individual recognition in mice mediated by major urinary proteins.
589 *Nature* 414:631-634.
590 Janotova K, and Stopka P. 2011. The level of major urinary proteins is socially regulated in wild
591 Mus musculus musculus. *J Chem Ecol* 37:647-656.
592 Jemiolo B, Harvey S, and Novotny M. 1986. Promotion of the Whitten effect in female mice by
593 synthetic analogs of male urinary constituents. *PNAS* 83:4576-4579.
594 Karn R, and Laukaitis C. 2015. Comparative Proteomics of Mouse Tears and Saliva: Evidence
595 from Large Protein Families for Functional Adaptation. *Proteomes* 3:283.

596 Kaur AW, Ackels T, Kuo TH, Cichy A, Dey S, Hays C, Kateri M, Logan DW, Marton TF, Spehr
597 M, and Stowers L. 2014. Murine pheromone proteins constitute a context-dependent
598 combinatorial code governing multiple social behaviors. *Cell* 157:676-688.
599 Kimoto H, Haga S, Sato K, and Touhara K. 2005. Sex-specific peptides from exocrine glands
600 stimulate mouse vomeronasal sensory neurons. *Nature* 437:898-901.
601 Kimoto H, Sato K, Nodari F, Haga S, Holy TE, and Touhara K. 2007. Sex- and strain-specific
602 expression and vomeronasal activity of mouse ESP family peptides. *Curr Biol* 17:1879-
603 1884.
604 Kwak J, Grigsby CC, Preti G, Rizki MM, Yamazaki K, and Beauchamp GK. 2013. Changes in
605 volatile compounds of mouse urine as it ages: their interactions with water and urinary
606 proteins. *Physiol Behav* 120:211-219.
607 Kwak J, Josue J, Faranda A, Opiekun MC, Preti G, Osada K, Yamazaki K, and Beauchamp GK.
608 2011. Butylated hydroxytoluene is a ligand of urinary proteins derived from female mice.
609 *Chem Senses* 36:443-452.
610 Kwak J, Strasser E, Luzynski K, Thoss M, and Penn DJ. 2016. Are MUPs a Toxic Waste
611 Disposal System? *PLoS One* 11:e0151474.
612 Leclair EE. 2003a. Four BPI (bactericidal/permeability-increasing protein)-like genes expressed
613 in the mouse nasal, oral, airway and digestive epithelia. *Biochem Soc Trans* 31:801-805.
614 LeClair EE. 2003b. Four reasons to consider a novel class of innate immune molecules in the
615 oral epithelium. *J Dent Res* 82:944-950.
616 Lopes PC, and Konig B. 2016. Choosing a healthy mate: sexually attractive traits as reliable
617 indicators of current disease status in house mice. *Animal Behaviour* 111:119-126.
618 Luo M, Fee MS, and Katz LC. 2003. Encoding pheromonal signals in the accessory olfactory
619 bulb of behaving mice. *Science* 299:1196-1201.
620 Luo M, and Katz LC. 2004. Encoding pheromonal signals in the mammalian vomeronasal
621 system. *Current Opinion in Neurobiology* 2004:428-434.
622 Ma W, Miao Z, and Novotny MV. 1999. Induction of Estrus in Grouped Female Mice (Mus
623 domesticus) by Synthetic Analogues of Preputial Gland Constituents
624 10.1093/chemse/24.3.289. *Chem Senses* 24:289-293.
625 Maynard Smith J, and Harper DD. 2003. *Animal signals*. Oxford: Oxford University Press.
626 Mucignat-Caretta C, and Caretta A. 1999. Protein-bound odorants as flags of male mouse
627 presence. In: Johnston RE, Muller-Schwarze D, and Sorensen P, eds. *Advances in*
628 *Chemical Communication in Vertebrates*. New York: Plenum Press, 359-364.
629 Nagnan-Le Meillour P, Vercoutter-Edouart AS, Hilliou F, Le Danvic C, and Levy F. 2014.
630 Proteomic Analysis of Pig (Sus scrofa) Olfactory Soluble Proteome Reveals O-Linked-N-
631 Acetylglucosaminylation of Secreted Odorant-Binding Proteins. *Front Endocrinol*
632 *(Lausanne)* 5:202.
633 Nelson AC, Cunningham CB, Ruff JS, and Potts WK. 2015. Protein pheromone expression
634 levels predict and respond to the formation of social dominance networks. *Journal of*
635 *Evolutionary Biology* 28:1213-1224.
636 Novotny MV, Harvey S, Jemiolo B, and Alberts J. 1985. Synthetic pheromones that promote
637 inter-male aggression in mice. *Proc Natl Acad Sci USA* 82:2059-2061.
638 Novotny MV, Jemiolo B, Harvey S, Wiesler D, and Marchlewska-Koj A. 1986. Adrenal-
639 mediated endogeneous metabolites inhibit puberty in female mice. *Science* 231:722-725.

640 Pavelka N, Fournier ML, Swanson SK, Pelizzola M, Ricciardi-Castagnoli P, Florens L, and
641 Washburn MP. 2008. Statistical similarities between transcriptomics and quantitative
642 shotgun proteomics data. *Molecular & Cellular Proteomics* 7:631-644.
643 Pavelka N, Pelizzola M, Vizzardelli C, Capozzoli M, Splendiani A, Granucci F, and Ricciardi-
644 Castagnoli P. 2004. A power law global error model for the identification of differentially
645 expressed genes in microarray data. *BMC Bioinformatics* 5:203.
646 Pes D, Dal Monte M, Ganni M, and Pelosi P. 1992. Isolation of two odorant-binding proteins
647 from mouse nasal tissue. *Comp Biochem Physiol* 103B:1011-1017.
648 Pes D, Mameli M, Andreini I, Krieger J, Weber M, Breer H, and Pelosi P. 1998. Cloning and
649 expression of odorant-binding proteins Ia and Ib from mouse nasal tissue. *Gene* 212:49-
650 55.
651 Pes D, and Pelosi P. 1995. Odorant-binding proteins of the mouse. *Comp Biochem Physiol*
652 112B:471-479.
653 Phelan MM, McLean L, Hurst JL, Beynon RJ, and Lian LY. 2014. Comparative study of the
654 molecular variation between 'central' and 'peripheral' MUPs and significance for
655 behavioural signalling. *Biochem Soc Trans* 42:866-872.
656 Porcheron G, Garenaux A, Proulx J, Sabri M, and Dozois CM. 2013. Iron, copper, zinc, and
657 manganese transport and regulation in pathogenic Enterobacteria: correlations between
658 strains, site of infection and the relative importance of the different metal transport
659 systems for virulence. *Front Cell Infect Microbiol* 3:90.
660 Roberts SA, Simpson DM, Armstrong SD, Davidson AJ, Robertson DH, McLean L, Beynon RJ,
661 and Hurst JL. 2010. Darcin: a male pheromone that stimulates female memory and sexual
662 attraction to an individual male's odour. *BMC Biol* 8: doi:10. 1186/1741-7007-1188-
663 1175.
664 Rodriguez J, Gupta N, Smith RD, and Pevzner PA. 2008. Does trypsin cut before proline? *J*
665 *Proteome Res* 7:300-305.
666 Rusu AS, Krackow S, Jedelsky PL, Stopka P, and Konig B. 2008. A qualitative investigation of
667 major urinary proteins in relation to the onset of aggressive behavior and dispersive
668 motivation in male wild house mice (Mus musculus domesticus). *Journal of Ethology*
669 26:127-135.
670 Sanchez L, Calvo M, and Brock JH. 1992. Biological role of lactoferrin. *Arch Dis Child* 67:657-
671 661.
672 Scott A, Weldon S, and Taggart CC. 2011. SLPI and elafin: multifunctional antiproteases of the
673 WFDC family. *Biochem Soc Trans* 39:1437-1440.
674 Sharrow SD, Vaughn JL, Žídek L, Novotny MV, and Stone MJ. 2002. Pheromone binding by
675 polymorphic mouse major urinary proteins. *Protein Science* 11:2247-2256.
676 Shaw PH, Held WA, and Hastie ND. 1983. The gene family for major urinary proteins:
677 expression in several secretory tissues of the mouse. *Cell* 32:755-761.
678 Singer A, Macrides F, clancy AN, and Agosta WC. 1986. Purification and Analysis of a
679 Proteinaceous Aphrodisiac Pheromone from Hamster Vaginal Discharge. *Journal of*
680 *Biological Chemistry* 261:13323-13326.
681 Stopka P, Janotova K, and Heyrovsky D. 2007. The advertisement role of major urinary proteins
682 in mice. *Physiology & Behavior* 91:667-670.
683 Stopka P, Kuntova B, Klempt P, Havrdova L, Cerna M, and Stopkova R. 2016. On the saliva
684 proteome of the Eastern European house mouse (Mus musculus musculus) focusing on
685 sexual signalling and immunity. *Sci Rep* 6:32481.

686 Stopkova R, Dudkova B, Hajkova P, and Stopka P. 2014. Complementary roles of mouse
687 lipocalins in chemical communication and immunity. *Biochemical Society Transactions*
688 42:893-898.
689 Stopková R, Hladovcová D, J. K, Vyoral D, and Stopka P. 2009. Multiple roles of secretory
690 lipocalins (MUP, OBP) in mice. *Folia Zool* 58:29-40.
691 Stopková R, Stopka P, Janotová K, and Jedelsky PL. 2007. Species-specific expression of major
692 urinary proteins in the house mice (Mus musculus musculus and Mus musculus
693 domesticus). *J Chem Ecol* 33:861-869.
694 Stopkova R, Vinkler D, Kuntová B, Sedo O, Albrecht T, Suchan J, Dvorakova-Hortova K,
695 Zdrahal Z, and Stopka P. 2016. Mouse lipocalins (MUP, OBP, LCN) are co-expressed in
696 tissues involved in chemical communication. *Frontiers in Ecology and Evolution* 4.
697 Stopkova R, Zdrahal Z, Ryba S, Sedo O, Sandera M, and Stopka P. 2010. Novel OBP genes
698 similar to hamster Aphrodisin in the bank vole, Myodes glareolus. *BMC Genomics* 11:45.
699 Strotmann J, and Breer H. 2011. Internalization of odorant-binding proteins into the mouse
700 olfactory epithelium. *Histochem Cell Biol* 136:357-369.
701 Thonhauser KE, Raveh S, Hettyey A, Beissmann H, and Penn DJ. 2013. Scent marking increases
702 male reproductive success in wild house mice. *Anim Behav* 86:1013-1021.
703 Thoss M, Enk V, Yu H, Miller I, Luzynski KC, Balint B, Smith S, Razzazi-Fazeli E, and Penn
704 DJ. 2016. Diversity of major urinary proteins (MUPs) in wild house mice. *Sci Rep*
705 6:38378.
706 Thoß M, Luzynski K, Ante M, Miller I, and Penn DJ. 2015. Major urinary protein (MUP)
707 profiles show dynamic changes rather than individual 'barcode' signatures. *Frontiers in*
708 *Ecology and Evolution* 3.
709 Timm DE, Baker LJ, Mueller H, Zidek L, and Novotny MV. 2001. Structural basis of
710 pheromone binding to mouse major urinary protein (MUP-I). *Protein Science* 10:997-
711 1004.
712 Vogl T, Tenbrock K, Ludwig S, Leukert N, Ehrhardt C, van Zoelen MA, Nacken W, Foell D,
713 van der Poll T, Sorg C, and Roth J. 2007. Mrp8 and Mrp14 are endogenous activators of
714 Toll-like receptor 4, promoting lethal, endotoxin-induced shock. *Nat Med* 13:1042-1049.
715 Walcott B. 1998. The Lacrimal Gland and Its Veil of Tears. *News Physiol Sci* 13:97-103.
716 Zala SM, Bilak A, Perkins M, Potts WK, and Penn DJ. 2015. Female house mice initially shun
717 infected males, but do not avoid mating with them. *Behavioral Ecology and Sociobiology*
718 69:715-722.
719 Zala SM, Potts WK, and Penn DJ. 2004. Scent-marking displays provide honest signals of health
720 and infection. *Behavioral Ecology* 15:338-344.
721 Zidek L, Stone MJ, Lato SM, Pagel MD, Miao Z, Ellington AD, and Novotny MV. 1999. NMR
722 Mapping of the Recombinant Mouse Major Urinary Protein I Binding site Occupied by
723 the Pheromone 2-*sec*-Butyl-4,5-dihydrothiazole. *Biochemistry* 38:9850-9861.
724 Zoukhri D. 2006. Effect of inflammation on lacrimal gland function. *Exp Eye Res* 82:885-898.
725
726
727

Figure 1(on next page)

Figure 1

Graphical representation of signal intensities (x axis) and particular fold differences between males and females. Significant differences between males and females are continuously scaled from green ($p<0.05$) to blue ($p<0.01$).

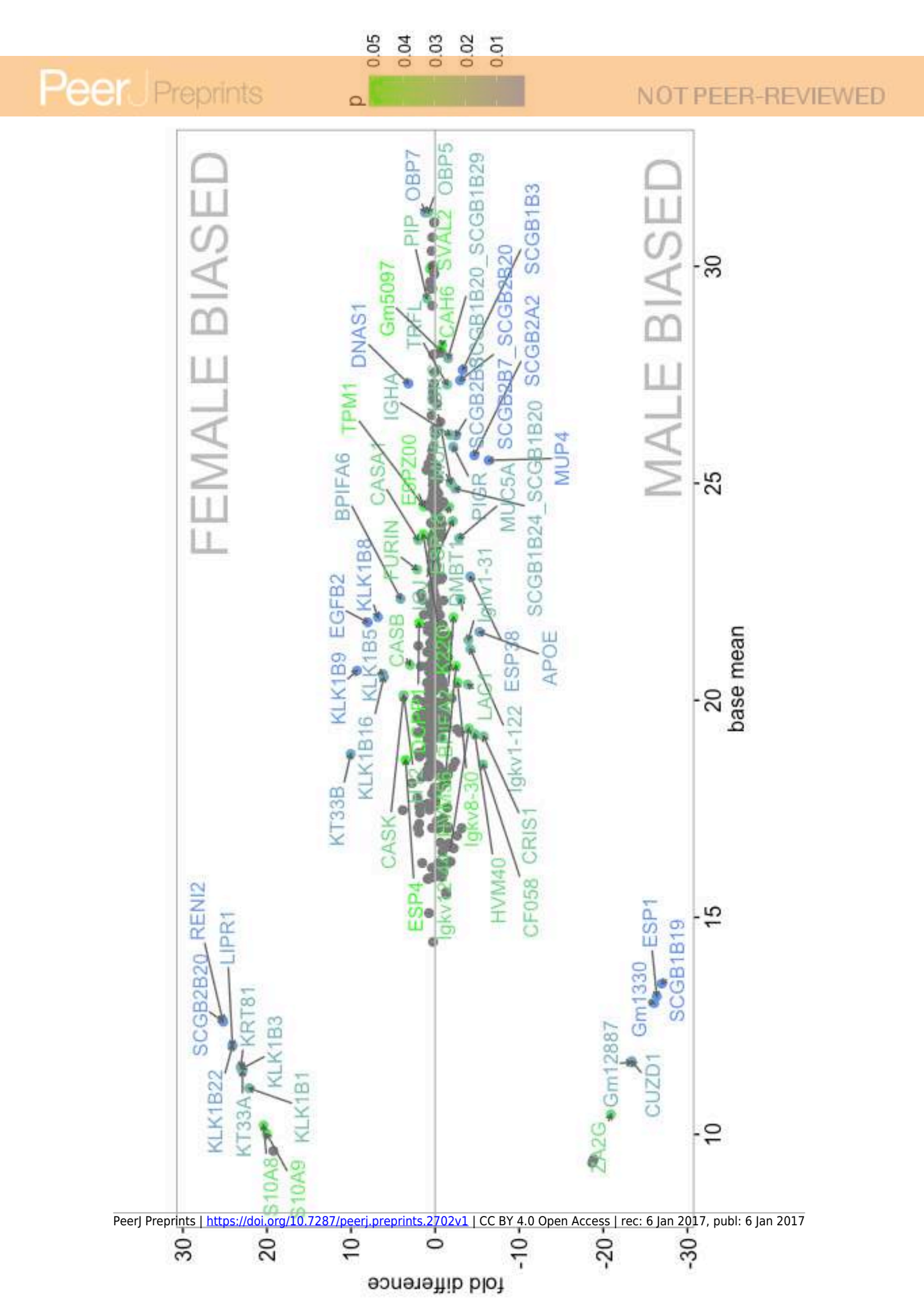


Figure 2(on next page)

Figure 2

Graphical representation of individual variation with heat maps produced by a hierarchical clustering method shows (a) the top 5% of highly expressed proteins, (b) significant sexually dimorphic proteins (i.e. labelled with stars - * $P \leq 0.05$, ** $P \leq 0.01$, *** $P \leq 0.001$) with a notable variation between individuals. Note that the abundance of MUP20 is invariant over individuals whilst exocrine-secreted peptides show a variation between individuals.

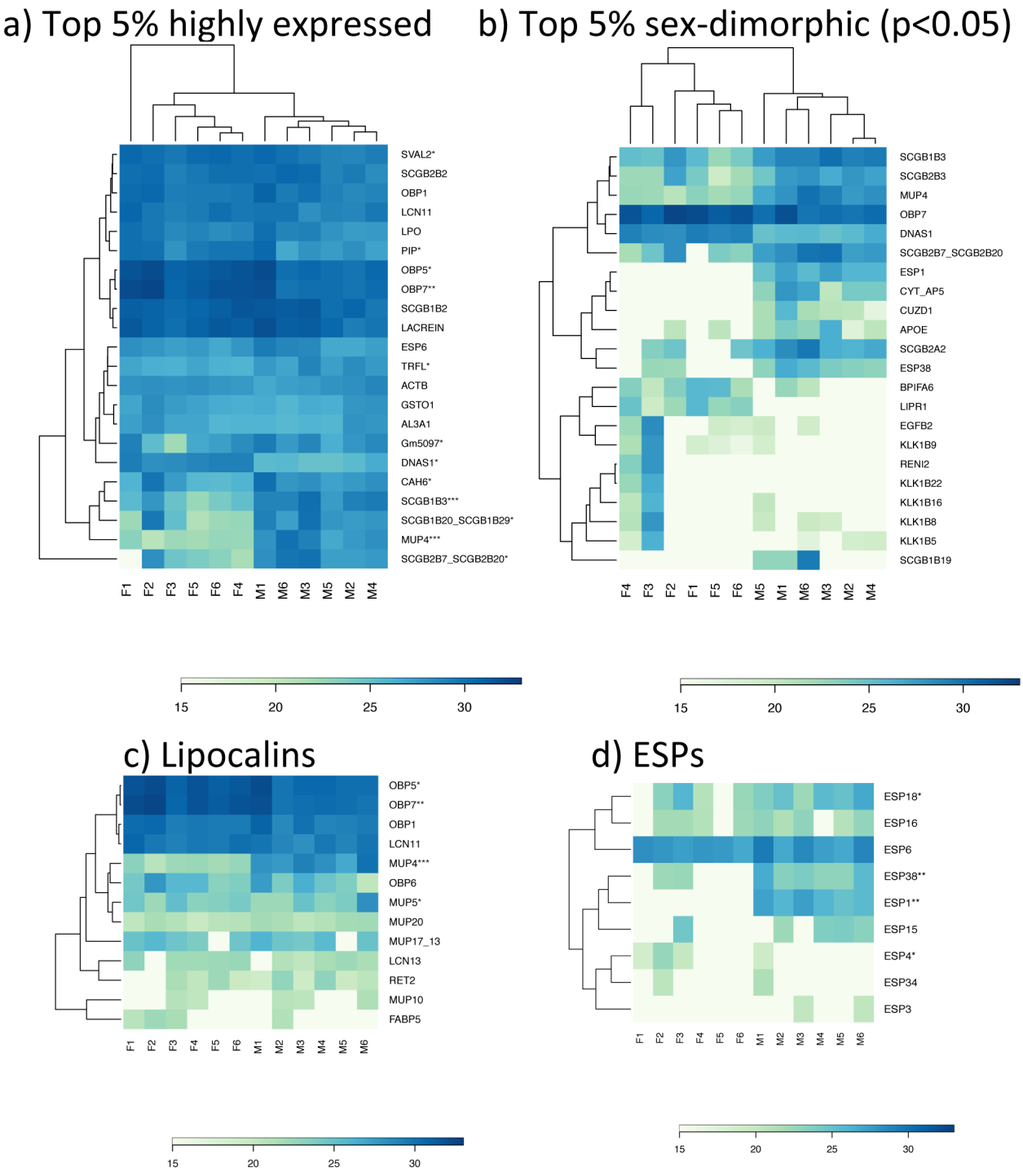

a) Top 5% highly expressed
SVAL2*
SCGB2B2
OBP1
LCN11
LPO
PIP*
OBP5*
OBP7**
SCGB1B2
LACREIN
ESP6
TRFL*
ACTB
GSTO1
AL3A1
Gm5097*
DNAS1*
CAH6*
SCGB1B3***
SCGB1B20_SCGB1B29*
MUP4***
SCGB2B7_SCGB2B20*
F1 F2 F3 F5 F6 F4 M1 M6 M3 M5 M2 M4
15 20 25 30
b) Top 5% sex-dimorphic (p<0.05)
SCGB1B3
SCGB2B3
MUP4
OBP7
DNAS1
SCGB2B7_SCGB2B20
ESP1
CYT_AP5
CUZD1
APOE
SCGB2A2
ESP38
BPIFA6
LIPR1
EGFB2
KLK1B9
RENI2
KLK1B22
KLK1B16
KLK1B8
KLK1B5
SCGB1B19
F4 F3 F2 F1 F5 F6 M5 M1 M6 M3 M2 M4
15 20 25 30
c) Lipocalins
OBP5*
OBP7**
OBP1
LCN11
MUP4***
OBP6
MUP5*
MUP20
MUP17_13
LCN13
RET2
MUP10
FABP5
F1 F2 F3 F4 F5 F6 M1 M2 M3 M4 M5 M6
15 20 25 30
d) ESPs
ESP18*
ESP16
ESP6
ESP38**
ESP1**
ESP15
ESP4*
ESP34
ESP3
F1 F2 F3 F4 F5 F6 M1 M2 M3 M4 M5 M6
15 20 25 30

Figure 3(on next page)

Figure 3

Histograms of mRNA contig length (a) are consistent over individuals and show that more than 50% of contigs is longer than 400bps. Dispersion (b) and MA (c) plots are demonstrating that the transcripts with lower number of reads have a higher dispersion. Most sexually dimorphic and significant transcript abundances are demonstrated with the heat map (d).

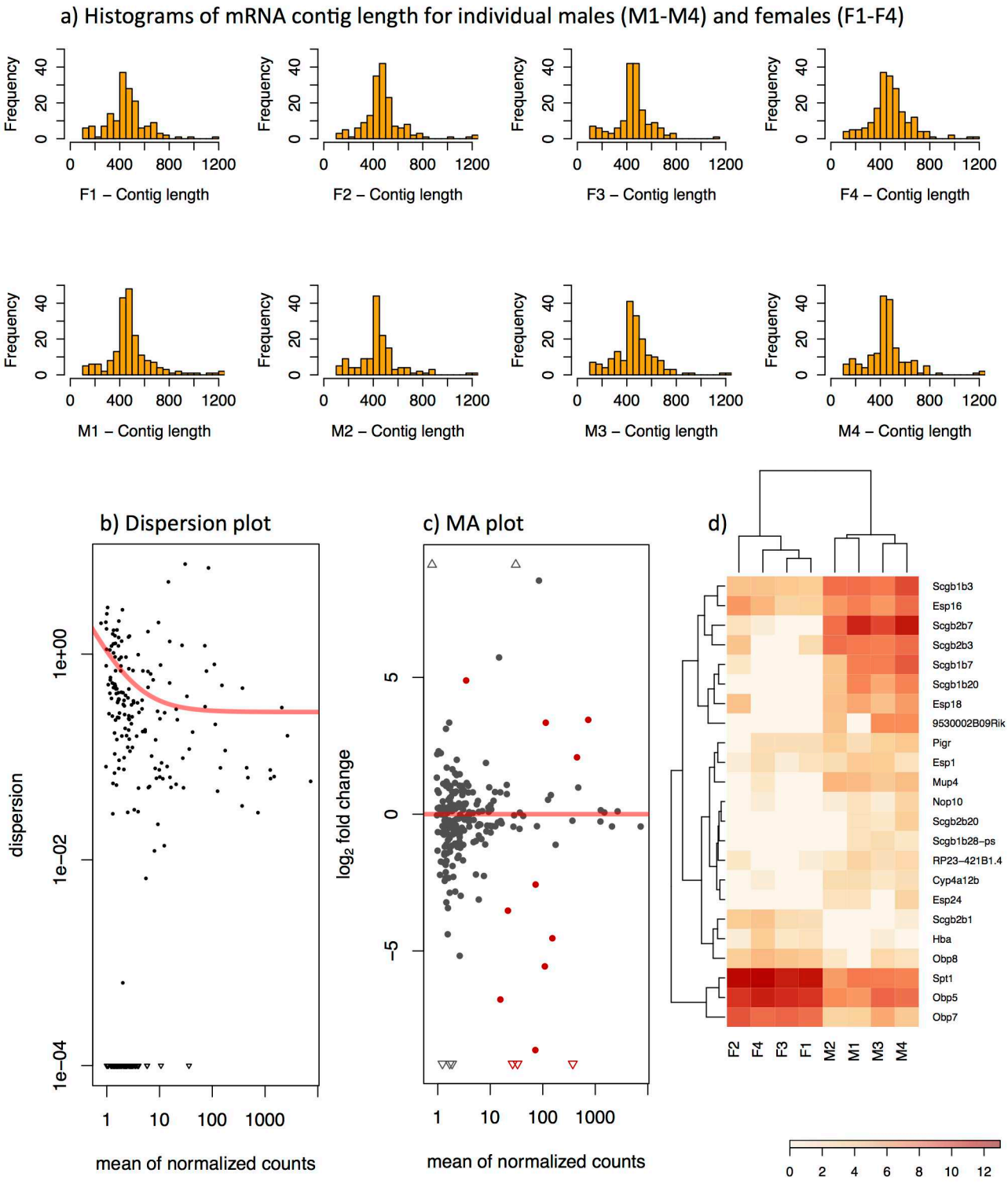
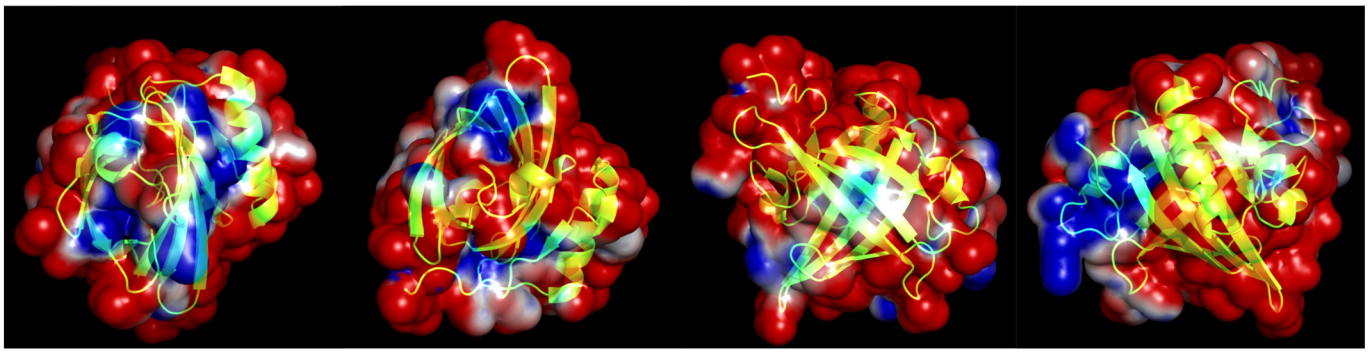

a) Histograms of mRNA contig length for individual males (M1-M4) and females (F1-F4)
Frequency
F1 – Contig length
F2 – Contig length
F3 – Contig length
F4 – Contig length
M1 – Contig length
M2 – Contig length
M3 – Contig length
M4 – Contig length
b) Dispersion plot
dispersion
mean of normalized counts
c) MA plot
log2 fold change
mean of normalized counts
d)
Scgb1b3
Esp16
Scgb2b7
Scgb2b3
Scgb1b7
Scgb1b20
Esp18
9530002B09Rik
Pigr
Esp1
Mup4
Nop10
Scgb2b20
Scgb1b28–ps
RP23–421B1.4
Cyp4a12b
Esp24
Scgb2b1
Hba
Obp8
Spt1
Obp5
Obp7
F2 F4 F3 F1 M2 M1 M3 M4

Figure 4(on next page)

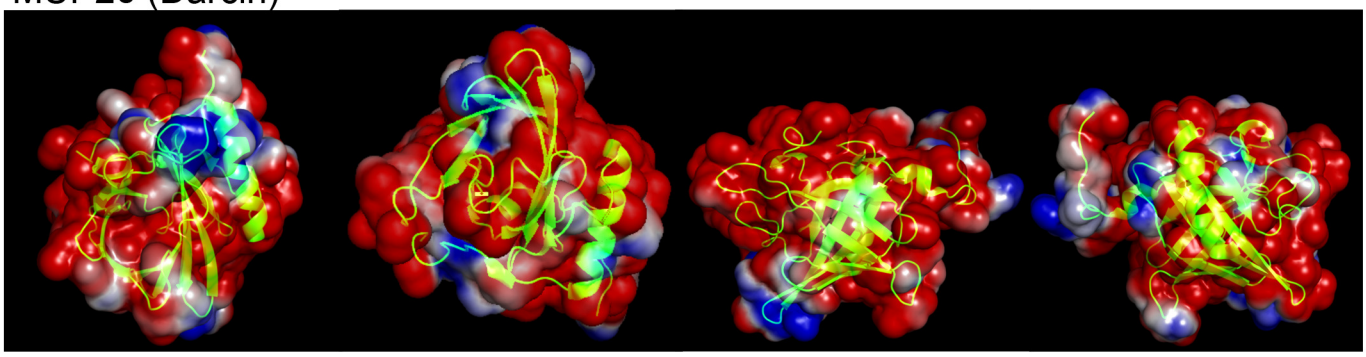
Figure 4

Graphical representation of the tertiary structure of MUP20, MUP1, OBP1, and LCN2 with electrostatics modelling and scaled from -1kTe (red, negative) to +1kTe (blue, positive). Although, their structures are highly similar due to their beta-barrel pocket, the distribution of positive and negative charges are non-random with OBP1 and LCN2 showing a great level of amphipathy.

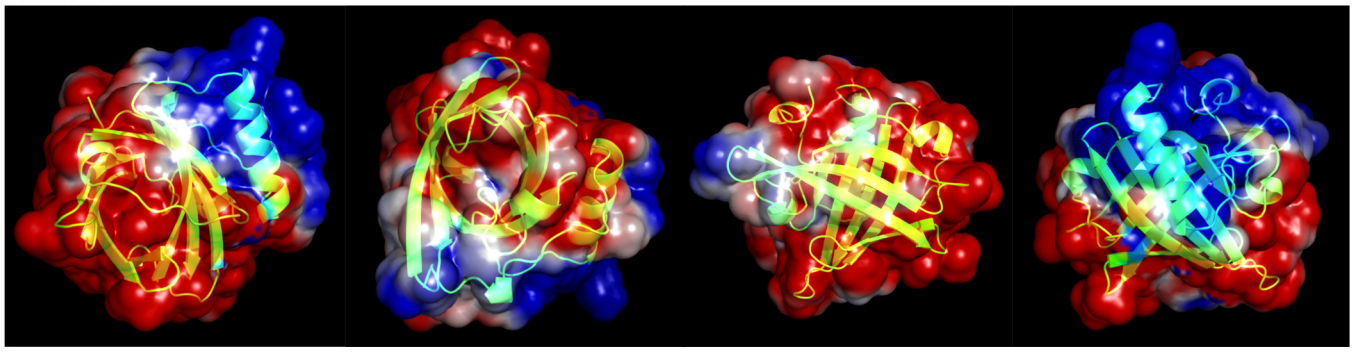
MUP1



MUP20 (Darcin)



OBP1



LCN2