



Deliverable D5.1

Haemocyte surface determinants

Point of Contact	Merel Braeckman
Institution	Ghent University
E-mail	Merel.Braeckman@UGent.be
Phone	+32(0)9 264 87 34

Project Acronym	Better-B
Project Title	Improving Bees' Resilience to Stressors by Restoring Harmony and Balance
Project No.	101081444
Topic	HORIZON-CL6-2022-BIODIV-02-03-two-stage
Project start date	1 June 2023
Nature	R
Dissemination level	Public
Due date	M15
Date of delivery	M15
Lead partner	USAMV
Contributing partners	UGent
Authors	Merel Braeckman (UGent), Lina De Smet (UGent), Dirk de Graaf (UGent)
Reviewer	Robert Paxton (MLU)



© Better-B 2023. This work is licensed under a [Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License](https://creativecommons.org/licenses/by-nc-nd/4.0/).



**UK Research
and Innovation**



Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

This work was supported by the Better-B project, which has received funding from the European Union, the Swiss State Secretariat for Education, Research and Innovation (SERI) and UK Research and Innovation (UKRI) under the UK government's Horizon Europe funding guarantee (grant number 10068544).

Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union, European Research Executive Agency (REA), SERI or UKRI. Neither the European Union nor the granting authorities can be held responsible for them.



Revision History

Version	Date	Author	Comment
0.1	22.05.2024	Merel Braeckman	First draft
	25.05.2024	Lina De Smet	Review
0.2	03.06.24	Merel Braeckman	Corrections
	06.06.24	Dirk de Graaf	Review
0.3	06.06.24	Merel Braeckman	Corrections
	14.07.24	Robert Paxton	Review
1.0	30.07.2024	Merel Braeckman	Final version, approved by the coordinator



Contents

Abbreviations, Participant short names	vi
Abbreviations	vi
Participant short names	vi
List of Figures	viii
List of Tables	viii
Summary	1
1. Background and objectives	2
1.1 Background	2
1.2 Objectives	2
2. Material and Methods	3
2.1 Haemolymph collection	3
2.2 Buffers	4
2.3 Haemocyte staining	4
2.4 Proteomics	5
2.4.1 Trypsin shaving	5
2.4.2 Separation of haemocytes based on adherence	6
2.4.3 LC-MS/MS	6
2.4.4 Data analysis	7
2.5 Transcriptomics	7
2.5.1 RNA extraction	7
2.5.2 Quality control with Fragment Analyser	7
2.5.3 Sequencing	7
2.5.4 Alignment & quantification of the reads	7
3. Results	8
3.1 Buffers	8
3.1.1 Trypsin digestion performance	8
3.1.2 Haemocyte viability	8
3.2 Haemocyte staining	10
3.3 Proteomics	10
3.3.1 Unfractionated haemocyte sample	10
3.3.2 Fractionated haemocyte sample	12
3.4 Transcriptomics	13
3.4.1 Fragment analyser	13
3.4.2 RNA sequencing	14
3.5 List of the surface determinants of <i>Apis mellifera</i> haemocytes	14
4. Discussion	14



5. Supplementary data	16
References	28



Abbreviations, Participant short names

Abbreviations

ABC	Ammonium bicarbonate
BSA	Bovine serum albumin
CA	Consortium Agreement
CD	Cluster of Differentiation
DoA	Description of Action
DPBS	Dulbecco's phosphate buffered saline
EDTA	Ethylenediaminetetraacetic acid
GA	Grant Agreement
GR	Granulocytes
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
LM	Lower marker
MES	2-(N-morpholino)ethanesulfonic acid
MGG	May-Grünwald Giemsa
PBS	Phosphate buffered saline
PL	Plasmotocytes
RFU	Relative fluorescence units
RQN	RNA Quality Number
TM9SF	Transmembrane 9 superfamily member
WP	Work Package

Participant short names

AU	Aarhus Universitet
COA	Co-Actions
IPB	Instituto Politécnico de Bragança
IRIAF	Instituto Regional de Investigación y Desarrollo Agroalimentario y Forestal de Castilla-La Mancha
IZSLT	Istituto Zooprofilattico Sperimentale delle Regioni Lazio e Toscana
KUL	Katholieke Universiteit Leuven
MLU	Martin-Luther-Universität Halle-Wittenberg
NB	Norges Biokterlag Forening [Non-governmental organisation]
SCIPROM	SCIPROM Sàrl
TNTU	The Nottingham Trent University
UCOI	Universidade de Coimbra



UGENT	Universiteit Gent
UJAG	Uniwersytet Jagiellonski
UM	Université de Montpellier
USAMV	Universitatea de Științe Agricole și Medicină Veterinară Cluj-Napoca
UU	Uppsala Universitet
VDSJ	Van Der Steen Joseph
WR	Stichting Wageningen Research



List of Figures

Figure 1. Haemolymph collection from a honeybee. Orange arrow indicates clear haemolymph droplet.	3
Figure 2. Incision site to extract haemolymph from fifth instar <i>A. mellifera</i> larva. Black arrow indicates ventral region of the head and the dashed line indicates the incision site. Photo by Butola (2020).	4
Figure 3. SDS-PAGE gel (10%) testing efficiency of trypsin in several buffers. Digest performed in (1) DPBS (2) ABC (3) Control (4) Tris (5) MES (6) HEPES buffer	8
Figure 4. Haemocytes stained with Trypan Blue in (A) EDTA-Tris buffer (B) EDTA-ABC buffer (C) EDTA-DPBS buffer. Orange circles show haemocytes.	9
Figure 5. Haemocytes stained with May-Grünwald Giemsa. GR = granulocyte, PL1 = plasmatocyte subtype 1, PL3 = plasmatocyte subtype 3, PL4 = plasmatocyte subtype 4	10
Figure 6. Venn diagramme of the number of surface determinants unique and shared across samples.	11
Figure 7. Unstained haemocytes viewed by phase contrast microscopy. (A) Non-adherent haemocytes (B) Adherent haemocytes.	12
Figure 8. Venn diagramme of the surface determinants for the adherent and non-adherent fractions.	12
Figure 9. Fragment Analyser results. (A) Adult RNA sample (B) Larval RNA sample (LM= Lower Size Marker, RFU = Relative Fluorescence Units). In insects, the 28S rRNA contains an endogenous hidden break, resulting in two fragments that migrate close to the 18S rRNA (pink peak) during denaturation. The RQN value is determined by the ratio of the 18S to 28S ribosomal RNA peaks and the separation between these peaks. However, this ratio is affected by the hidden break, making the RQN unreliable for most insects (Winnebeck et al., 2010).	13

List of Tables

Table 1. Overview of the different protocols tested for trypsin shaving	5
Table 2. Overview of the protein analysis for each sample	11
Table 3. Statistics of clean data (RNAseq)	14
Table 4. Summary of the haemocyte surface determinants	16



Summary

Work package 5 (WP5) focusses on the cellular immunity of honeybees. Until now, the haemocytes in the haemolymph have been characterized based on their morphological features, adherence, lectin binding properties, granularity and movement (De Graaf et al., 2002; Gábor et al., 2020). They are responsible for nodulation, encapsulation and phagocytosis (Strachecka et al., 2021). In Better-B we aim to differentiate between the functional haemocyte types by performing haemocyte transcriptome and surface proteome studies.

This deliverable (D5.1) is the first of three deliverables for WP5 'Immune Resilience'. It presents a set of results from task 5.1: 'Mapping of the haemocyte surface determinants'. D5.1 describes the materials and methods and presents a list of identified surface determinants of *Apis mellifera* haemocytes of different developmental stages. This list was the result of generated proteomic and transcriptomic data. Combining the identified surface determinants with a literature search allowed various potential markers for haemocyte subpopulations to be discovered. Based on these results, surface determinants will be selected as targets to generate monoclonal antibodies (task 5.2).

1. Background and objectives

1.1 Background

The insect's immune system, composed of highly developed cellular and humoral components, is activated to remove or neutralize harmful factors and return the individual to homeostasis (Strachecka et al., 2021). Our knowledge of honeybee immunity is mainly focussed on humoral immunity, manifested by the expression of antimicrobial peptides. Cellular responses, in contrast, are mediated by haemocytes in haemolymph, which are responsible for nodulation, encapsulation, and phagocytosis (Strachecka et al., 2021). Until now, the haemocytes of the honeybee have been characterized on the basis of their morphological features, adherence, lectin binding properties, granularity and movement, which has led to ambiguity in their classification (De Graaf et al., 2002; Gábor et al., 2020). Recently, functionally distinct haemocyte subpopulations (oenocytoids, aggregating cells and granulocytes) have been identified using monoclonal antibodies. In order to further differentiate between functional haemocyte types, we here explore the haemocyte transcriptome and surface proteome, establishing a detailed map of the surface determinants of honeybee haemocytes. In doing so, we set a benchmark that goes beyond the current state-of-the-art and opens the field for thorough screening of the innate immune competence of honeybees. Expanding our possibilities to study the immune processes of the honeybee is especially important as abiotic factors like heat stress, nutrition and pesticides affect immune competence of all castes of the honeybee.

1.2 Objectives

The general objectives of WP5 'Immune Resilience' are:

- To perform baseline studies on honeybee cellular immunity by mapping the haemocyte surface determinants and developing a novel haemocyte typology based on flow cytometric profiling
- To determine the impact of heat and nutritional stress on the cellular and humoral immune potential of honeybees

Deliverable 5.1, listing the surface determinants of haemocytes, is the first step in developing a novel haemocyte typology. The list with surface determinants will be used to select targets that could potentially reveal subpopulations of haemocytes. Against these targets, monoclonal antibodies will be raised. A new haemocyte typology will subsequently be developed using flow cytometric profiling (Deliverable 5.2). Finally, the effects of heat and nutritional stress on these novel haemocyte subpopulations will be investigated (Deliverable 5.3).



2. Material and Methods

2.1 Haemolymph collection

Haemolymph was collected from adult worker and larval honey bees.

To collect haemolymph from adults, closed brood frames from the UGENT apiary with ready-to-emerge worker honeybees were incubated in a constant temperature chamber at 34°C. Freshly emerged honeybees were collected and anesthetized by placing them on ice for 30 minutes. Following this, the bees were secured by their thoraces between the thumb and index finger of the researcher's non-dominant hand, with the head of the bee oriented towards the fingertips. The removal of one pair of wings was then executed using tweezers, with a quick pulling motion at the base of the wings. Following wing removal, the thorax was gently squeezed between the thumb and index finger of the non-dominant hand until a clear droplet of haemolymph emerged (Figure 1). This droplet was collected using a micropipette. The collected haemolymph was then transferred into a 1.5mL Eppendorf tube (low-protein binding) containing ice-cold Dulbecco's phosphate buffered saline (DPBS) pH 7.4 with 0.02% EDTA at a 1:1 ratio. EDTA is added to prevent melanisation and coagulation. This protocol was designed to reduce the risk of contamination from the digestive tract and fat body.



Figure 1. Haemolymph collection from a honeybee. Orange arrow indicates clear haemolymph droplet.

For the collection of larval haemolymph, the methodology described by Butola (2020) was followed. Brood frames with larvae from worker bees 8 days of age were removed from the hive and the larvae were immediately used for haemolymph collection. To do so, the larva was held between thumb and index finger of the non-dominant hand with the head of the larva oriented towards the fingertips. Gentle pressure was applied to extend the head region and an incision in the cuticle corresponding to the head area was made using ophthalmic scissors. A micropipette was utilized to collect the exuding haemolymph, which was then transferred into a 1.5 mL low-protein binding Eppendorf tube. The Eppendorf tube contained ice-cold Dulbecco's PBS with 0.02% EDTA at a ratio of 1:1 with the collected haemolymph.

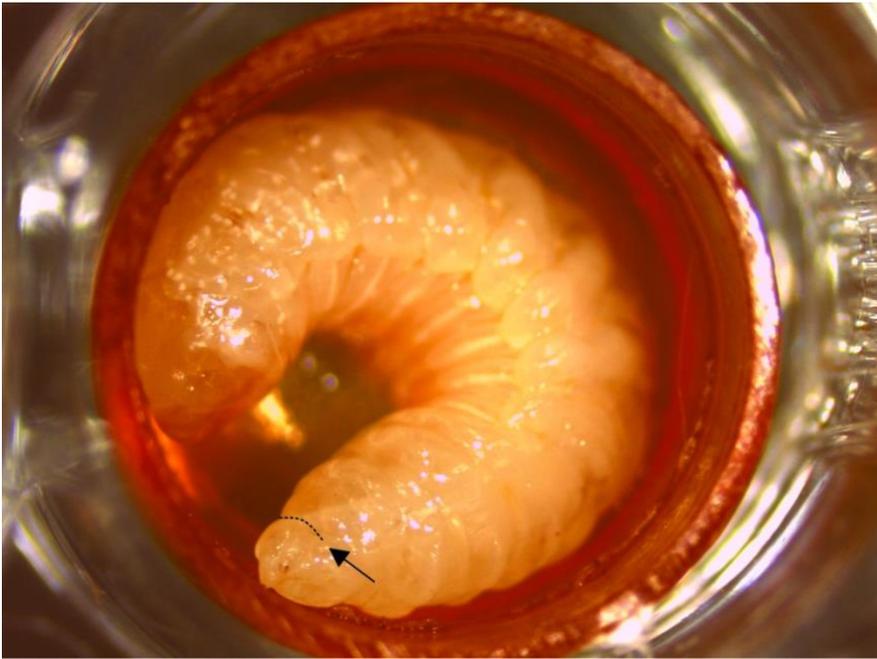


Figure 2. Incision site to extract haemolymph from fifth instar *A. mellifera* larva. Black arrow indicates ventral region of the head and the dashed line indicates the incision site. Photo by Butola (2020).

2.2 Buffers

2.2.1.1 Trypsin digestion performance

Five different buffers were tested to evaluate the efficiency of trypsin digestion of bovine serum albumin (BSA). The buffers used in the experiment were Dulbecco's phosphate-buffered saline pH 7.5 (DPBS), 100 mM ammonium bicarbonate buffer (ABC) pH 7.5, 50 mM Tris buffer pH 7.5, 50 mM MES buffer pH 7.0 and 50 mM HEPES buffer pH 7.5.

For each test, 10 µg of BSA was dissolved in one of the buffers, digested with 1 µg of trypsin and incubated for four hours at 37°C. The efficiency of trypsin digestion was determined by resolving products on a 10% SDS-PAGE gel. Additionally, a negative control consisting of BSA without trypsin was included.

2.2.1.2 Haemocyte viability

To evaluate haemocyte viability in the buffers, 12 µL of haemolymph was mixed with 12 µL of each buffer containing EDTA (0.02%) and 12 µL of Trypan Blue (0.4%). The integrity of the haemocytes was microscopically analysed in a Bürker counting chamber.

2.3 Haemocyte staining

For the staining of haemocytes, the protocol described by Van Steenkiste (1988) was followed. Twelve µL of haemolymph was collected in 10 µL of ice-cold EDTA-DPBS (0.02%) buffer. Two µL of methanol was added and the sample was then centrifuged for 5 minutes at 0.5G (4°C). After centrifugation, the plasma was removed and the pellet was resuspended in 10 µL of EDTA-DPBS (0.02%) buffer. A smear was prepared on a microscope slide and air dried briefly. The smear was covered with May-Grünwald stain for 2 minutes, followed by the addition of an equal amount of PBS for another 2 minutes. The smear was then washed with PBS. Next, the smear was covered with diluted Giemsa stain (0.2 mL Giemsa in 10 mL buffer) for 20 minutes. After staining, the smear was washed with distilled water. The stained smear was microscopically evaluated (40x10 magnification).



2.4 Proteomics

2.4.1 Trypsin shaving

To collect the proteins present on the haemocyte surface, trypsin (a protein-digesting enzyme) was used to shave the proteins off of the surface of the haemocytes. The following protocol was used: the haemolymph was centrifuged at 0.5G for 5 minutes at 4°C, after which the plasma was removed using a filter pipette tip. The pelleted haemocytes were washed once, twice or three times with EDTA-DPBS buffer. The haemocytes were collected using a 5-minute centrifugation step at 0.5G and 4°C. The wash buffer was removed using a filter pipette tip. The final haemocyte pellet was resuspended in EDTA-DPBS buffer. Subsequently the different samples were treated with trypsin (concentration ranging from 0.5 to 3 µg) or 1 µg proteinase K and incubated at 37°C for various durations (as specified in Table 1). Following incubation, the samples were centrifuged for 5 minutes at 0.5G and 4°C and the supernatant containing the surface determinants was collected and stored at -20°C till LC-MS/MS analysis. The shaved haemocytes were stored at -80°C for later RNA isolation which was to be used for RNA sequencing.

Table 1. Overview of the different protocols tested for trypsin shaving. Various amounts of different proteases were tested with shaving durations ranging from 5 to 60 minutes. In some cases, after an initial 20-minute shaving step and collecting the supernatant containing surface determinants, a second 20-minute shaving was performed on the same haemocyte pellet (20 min + 20 min).

Haemolymph	Protease	Shaving protocol
20 µl	1 µg trypsin	20 min
		20 min + 20 min
20 µl	1 µg proteinase K	20 min
		20 min + 20 min
40 µl	0.5 µg trypsin	30 min
40 µl	1 µg trypsin	30 min
80 µl	1 µg trypsin	60 min
		45 min
		30 min
		20 min
		15 min
		10 min
80 µl	2 µg trypsin	30 min
		20 min
		15 min
		10 min
		5 min
80 µl	3 µg trypsin	30 min
		10 min



2.4.2 Separation of haemocytes based on adherence

By separating haemocytes based on their adherence properties, the potential differences in surface proteins present can be evaluated, potentially identifying interesting markers. The following protocol was used to achieve this separation. A volume of 200 μL of haemolymph was added to 300 μL of ice-cold EDTA-DPBS buffer in a low-protein binding Eppendorf tube. The collected sample was then transferred to a well of a 24-well plate (SPL Life Sciences) containing a sterilized glass coverslip. The sample was incubated for 1 hour at 27°C to allow the haemocytes to adhere. The sample was pipetted up and down a few times to detach loosely attached haemocytes from the glass coverslip. The coverslip was then placed in a new well containing 500 μL of EDTA-DPBS buffer. The adherent haemocytes were removed using a sterilized, nonpyrogenic cell scraper (SPL Life Sciences) by scraping multiple times in different directions. Both the suspension and adherent cells were collected into a new 1.5 mL Eppendorf tube and centrifuged for 5 minutes at 0.5 G (4°C). The supernatant was discarded and the pellet was washed twice with 500 μL of EDTA-DPBS buffer. Finally, 1 μg of trypsin was added and the sample was incubated for 1 hour at 37°C. After incubation, the samples were centrifuged and the supernatants containing the surface determinants were stored at -20°C for further analysis.

2.4.3 LC-MS/MS

Twelve samples of surface determinants, each derived from 80 μL of haemolymph, were analysed using LC-MS/MS. Three samples from four different categories were analysed:

- Larval surface determinants in EDTA-DPBS buffer
- Adult surface determinants in EDTA-DPBS buffer
- Larval surface determinants in EDTA-ABC buffer
- Adult surface determinants in EDTA-ABC buffer

Within each category, samples were washed once, twice or three times. All samples were treated with 1 μg of trypsin and subjected to a 1-hour initial digestion (shaving step), followed by an additional 16-hour peptide digestion. One sample of surface determinants of the adherent fraction and one sample of surface determinants of the non-adherent fraction were analysed using LC-MS/MS.

Purified dried peptides were reconstituted in loading solvent A (0.1% TFA in water/ACN (98:2, v/v)). Approximately 2 μg of each sample was subjected to LC-MS/MS analysis using an Ultimate 3000 RSLC nano system with ProFlow Technology (Thermo scientific) linked to an Orbitrap Exploris mass spectrometer (Thermo Scientific) equipped with a Nanospray Flex Ion source (Thermo Scientific). The chromatographic separation utilized a trapping phase and a separating column in which peptides were trapped for 2 min at a flow rate of 20 $\mu\text{L}/\text{min}$ in solvent A on a 20 mm trapping column (20 mm trapping column (Thermo Scientific), 300 μm internal diameter (I.D.), 5 μm beads) while the sample was loaded on a 25 cm in needle packed column made in-house (75 μm internal diameter (I.D.), 3 μm beads, C18 Reprosil-HD, Dr. Maisch, Germany) mounted in the Ultimate 3000's column oven at 25°C. The elution gradient ranged from 0.5 to 26.4% MS solvent B (0.1% FA in water/ACN (2:8, v/v)) for 30 minutes, followed by a gradient from 26.4% to 44% solvent B in 8 min and from 44% solvent B to 56% solvent B in 2 minutes. This elution was first at a flow rate 250 nL/min, followed by a 15-minute wash reaching 99% MS solvent B and re-equilibration with MS solvent A (0.1% FA in water). The mass spectrometer operated in data-dependent mode - automatically switching between MS and MS/MS acquisition (using minimum intensity threshold = 8000), acquired full-scan MS spectra (350-1200 m/z) and subsequent MS/MS spectra (120-2000 m/z). The former MS spectra were acquired at a resolution of 60,000 in the orbitrap analyser, while the latter MS/MS spectra were acquired at a resolution of 15,000 in the orbitrap analyser. A dynamic exclusion window of 10 ppm was used. Fragmentation occurred at a normalized collision energy of 30% after filling the trap at a target value of 100,000 cps for maximum 100 ms. The S-lens RF level was set at 70 and precursor ions with single, unassigned and >7 charge states were excluded from fragmentation selection.



2.4.4 Data analysis

For data analysis, MaxQuant v2.4.4.0 was used together with the *Apis mellifera* database from Uniprot. Trypsin was designated as the digestion enzyme, allowing for a maximum of two missed cleavages and using default identification parameters. Carbamidomethylation of cysteine was set as a fixed modification, while acetylation of N-termini, oxidation of methionine and deamidation of asparagine and glutamine were considered as variable modifications. The false discovery rate (FDR) was set at 1%.

Because of the abundance of proteomic data, a script was developed to streamline analysis procedures. This script was designed to extract accession codes from the proteomic data and cross-reference them with the corresponding FASTA sequences on Uniprot. Based on these protein sequences, deep learning techniques were used, specifically DeepLoc 2.0, to predict the subcellular location of the protein.

2.5 Transcriptomics

2.5.1 RNA extraction

RNA was extracted from haemocytes from 320 μL of haemolymph, equivalent to approximately 5 million haemocytes. The haemolymph was centrifuged at 0.5G for 5 minutes at 4°C and the plasma was discarded. RNA extraction from the resulting haemocyte pellets was carried out using the Qiagen RNeasy Micro Kit. The pellet was resuspended in 350 μL of RLT lysis buffer supplemented with 1% β -mercapto-ethanol and vortexed. The lysate was then passed through a 20G needle ten times to ensure thorough disruption of the haemocytes. The RNA was purified using the Qiagen RNeasy spin column following the manufacturer's recommendations, including an on-column DNase step. Finally, the RNA was eluted in 14 μL of RNase-free water.

2.5.2 Quality control with Fragment Analyser

The quality and concentration of the RNA was measured using the Agilent 5200 Fragment Analyzer system, a capillary electrophoresis instrument. Separation was performed at 9.5 kV for 1 hour. The Fragment Analyzer generates an electropherogram showing both RNA fragment size and relative abundance. Additionally, an RNA Quality Number (RQN), a quantitative indication of RNA integrity, is given.

2.5.3 Sequencing

RNA sequencing was performed by BGI (China). mRNA enrichment was conducted using oligo dT beads to target mRNA molecules with poly A tails. In the next step, mRNA was fragmented into smaller pieces and then converted into first-strand cDNA using random primers. The second-strand cDNA synthesis utilized dUTP. The resulting cDNA underwent end-repair and 3' adenylation, followed by adapter ligation to the 3' adenylated cDNA fragments. PCR amplification was employed for library fragment enrichment, followed by quality control and library circularization steps. Subsequently, the library underwent amplification to generate DNA nanoballs, enabling sequencing on the DNBSEQ platform with paired-end reads of 150 base pairs. BGI also conducted basic bioinformatic analysis, involving the trimming of adaptors and removal of low-quality reads from the raw fastq (sequence) files.

2.5.4 Alignment & quantification of the reads

Utilizing Kallisto index in the command line, an index was generated based on the reference transcriptome of *Apis mellifera* (Zheng et al., 2023). Subsequently, the abundance of each transcript was estimated through a pseudo-alignment method using Kallisto quant.



3. Results

3.1 Buffers

3.1.1 Trypsin digestion performance

Trypsin was active in the following buffers DPBS- (lane 1), ABC- (lane 2) and Tris (lane 4) buffer (all in Figure 3). In MES (lane 5) and HEPES (lane 6) buffers, undigested BSA was still visible on the SDS-PAGE gel, suggesting these buffers to be unsuitable. Comparisons were made against the negative control (undigested BSA) (lane 3).

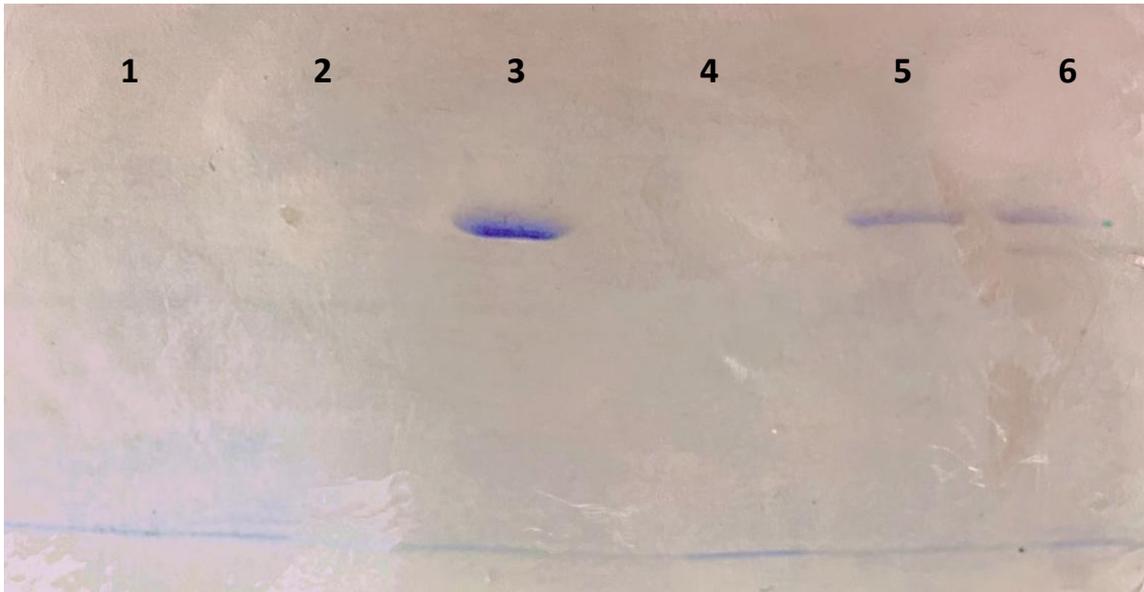


Figure 3. SDS-PAGE gel (10%) testing efficiency of trypsin in several buffers. Digest performed in (1) DPBS (2) ABC (3) Control (4) Tris (5) MES (6) HEPES buffer

3.1.2 Haemocyte viability

The three remaining buffers, EDTA-DBPS, EDTA-Tris and EDTA-ABC, were tested as sampling buffer for haemocytes. The haemocytes need to stay intact to achieve proper surface determinant shaving. Figure 4 demonstrates that EDTA-DBPS was the only buffer in which the haemocytes remained intact. The haemocytes in the EDTA-DBPS (Figure 4C) buffer did not take up the blue-purple stain (Trypan Blue), indicating their cell membranes were impermeable and, thus, the cells were still viable.

In contrast, Figures 4A and 4B reveal that haemocytes in EDTA-Tris and EDTA-ABC buffer exhibited significant shrinkage and darkening of their coloration. The dark staining is due to the uptake of the dye, a result of increased cell membrane permeability in dead haemocytes.

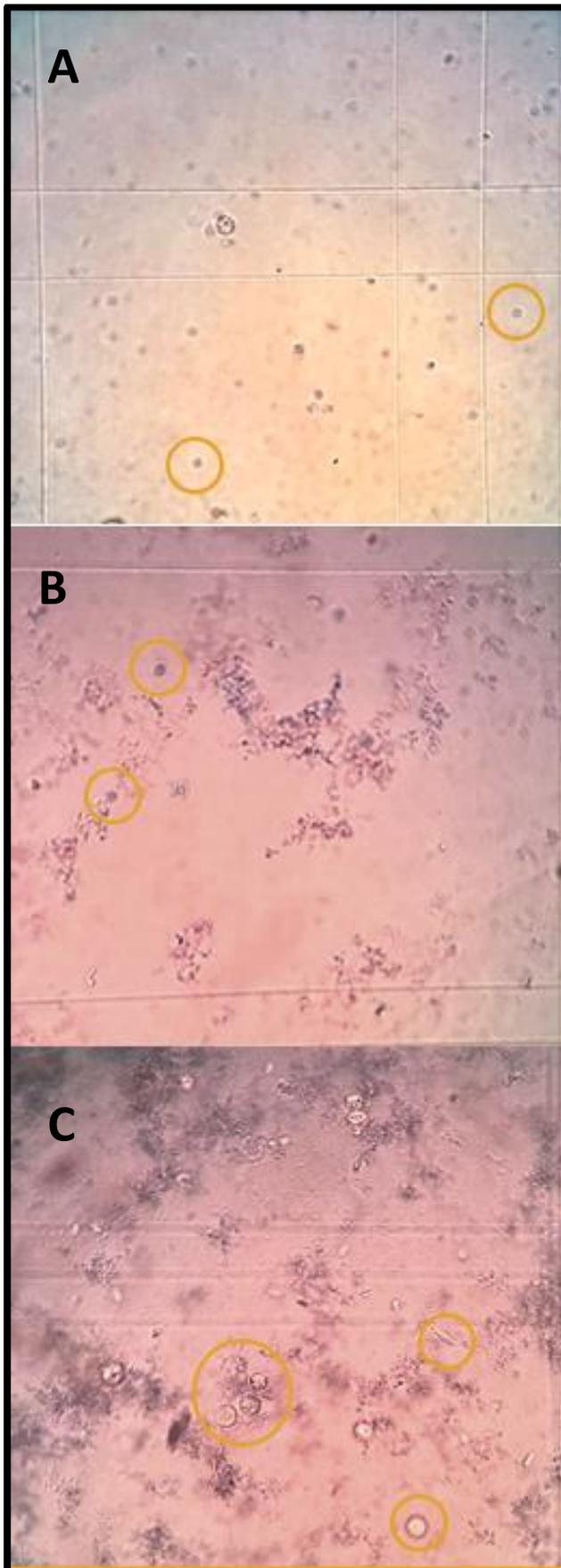


Figure 4. Haemocytes stained with Trypan Blue in (A) EDTA-Tris buffer (B) EDTA-ABC buffer (C) EDTA-DPBS buffer. Orange circles show haemocytes.

3.2 Haemocyte staining

Various morphological populations of haemocytes can be identified following staining with May-Grünwald Giemsa. Granulocytes (GR) are relatively large, round cells with cytoplasm that stains moderate to dark purple. Plasmatocytes (PL), on the other hand, are smaller cells with a high nucleus-to-cytoplasm ratio. They exhibit light or minimal staining, making them appear more neutrophilic (Richardson et al., 2018). Plasmatocytes are further classified into several subtypes. PL1 cells are small, round cells with a dense, central nucleus. PL2 cells represent a transitional stage between PL1 and PL3. PL3 cells are oval, discoid cells and PL4 cells are fusiform in shape. Figure 5 provides examples of these haemocyte subpopulations.

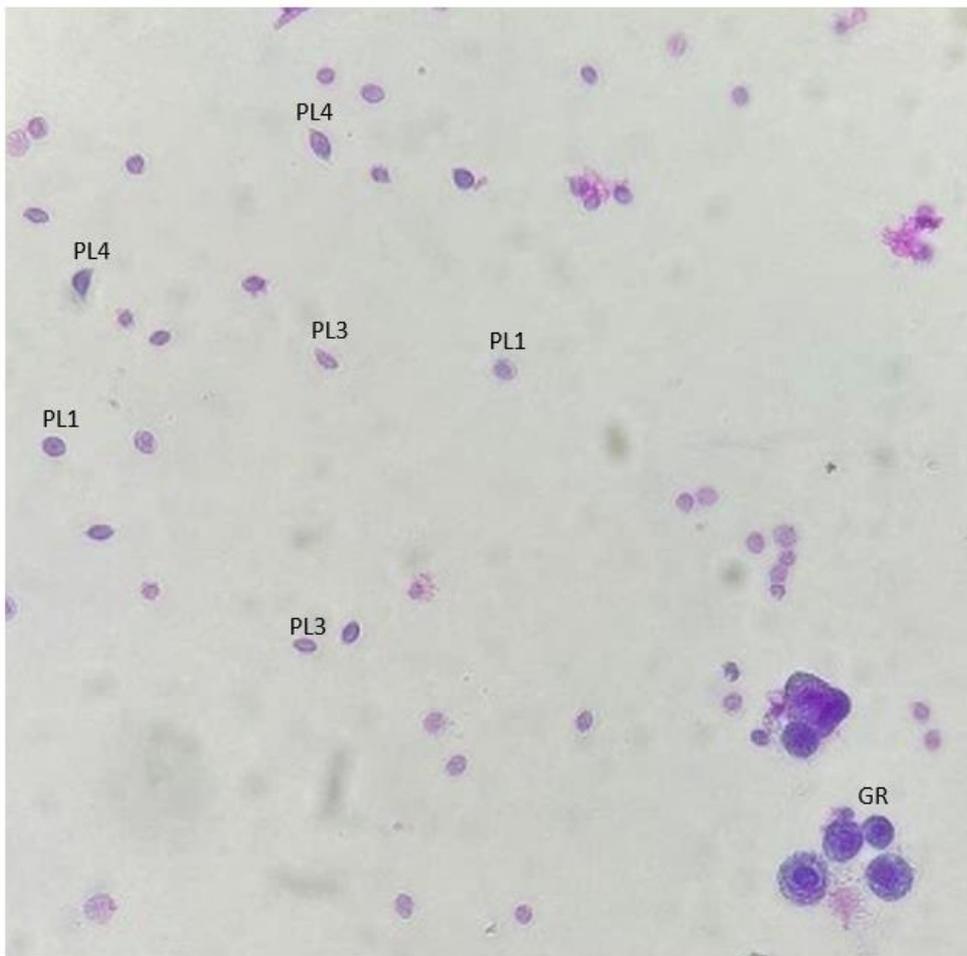


Figure 5. Haemocytes stained with May-Grünwald Giemsa. GR = granulocyte, PL1 = plasmatocyte subtype 1, PL3 = plasmatocyte subtype 3, PL4 = plasmatocyte subtype 4

3.3 Proteomics

3.3.1 Unfractionated haemocyte sample

Table 2 provides an overview of the protein analysis for each sample category; the first column lists the total number of proteins detected using LC-MS/MS. The second column gives the number of cell membrane proteins identified with DeepLoc 2.0. The final column shows the number of surface determinants remaining after manually reviewing the data based on existing literature and removing any identified, contaminating proteins.



Table 2. Overview of the protein analysis for each sample.

	Proteins identified with LC-MS/MS	After screening with DeepLoc 2.0	After manual filtering
Adult-DPBS	277	14	13
Larva-DPBS	428	27	23
Adult-ABC	4199	166	130
Larva-ABC	3855	138	105

As depicted in the Venn diagramme (Figure 6), three surface determinants were consistently present in every sample: Sodium/potassium-transporting ATPase subunit alpha, Talin-1 and Solute carrier family 12 member 7. Talin-1, previously identified in haemocytes of a mollusc (*Pomacea canaliculate*) (Boraldi et al., 2019), functions as a cytoskeletal protein (Carroll & Coyne, 2021). Sodium/potassium-transporting ATPase subunit alpha was found in the haemocytes of shrimp (*Litopenaeus vannamei*) and plays a role in regulating ion balance (Tang et al., 2018). In Tang's study, the gene was upregulated 12 hours after virus infection, leading to the suppression of haemocyte apoptosis. This mechanism potentially benefits the virus by promoting its proliferation. Similarly, the solute carrier family 12 functions as an ion electroneutral symporter. Although there is limited literature on solute carrier family 12 member 7, it is known that solute carrier family 12 member 9 is upregulated in virus-infected haemocytes of shrimp (Hernández-Pérez et al., 2019).

Samples in the ABC buffer revealed far more surface determinants than the samples in the DPBS buffer, respectively 147 versus 30. Minimal additional surface determinants (only five) were detected when using the DPBS buffer. Sodium/potassium-transporting ATPase subunit beta-2 and Collagen alpha chain were exclusively detected in adult haemocyte samples, while spectrin beta chain was exclusively detected in larval haemocyte samples. In total, there were 152 uniquely identified surface determinants for these 12 samples.

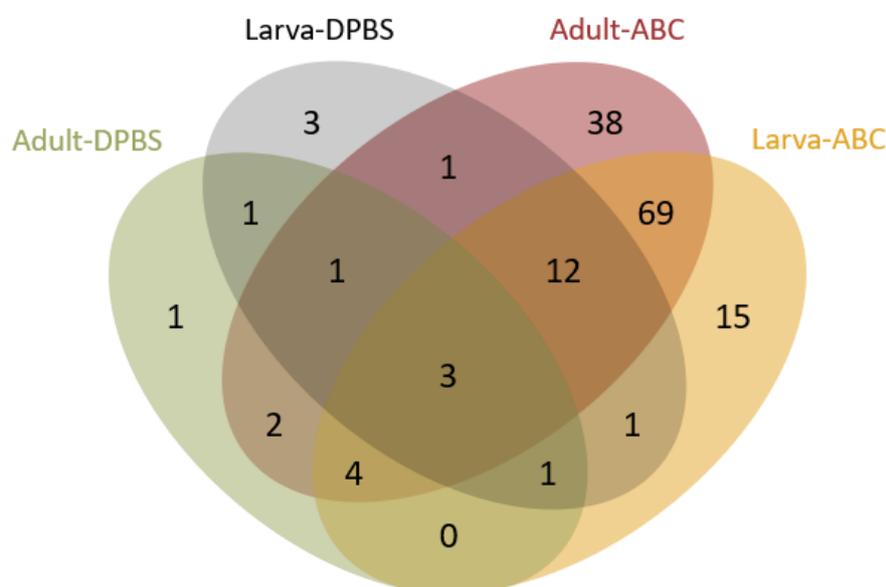


Figure 6. Venn diagramme of the number of surface determinants unique and shared across samples.

If the surface determinant's presence increases across samples, the surface determinant will localize more centrally within the Venn diagramme. Proteins exclusively identified in one sample are situated in the outer corners of the diagramme.



3.3.2 Fractionated haemocyte sample

To assess the separation of adherent and non-adherent fractions of haemocytes, the samples were examined by phase-contrast microscope (Figure 7). The observations revealed that granulocytes were exclusively present in the adherent fraction, while only a small portion of plasmatocytes adhered.

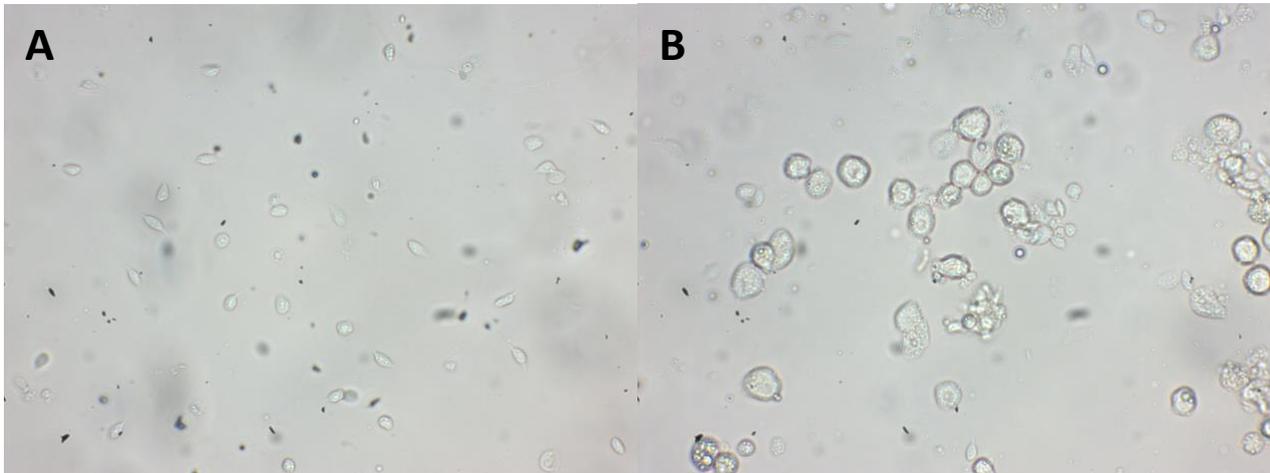


Figure 7. Unstained haemocytes viewed by phase contrast microscopy. (A) Non-adherent haemocytes (B) Adherent haemocytes.

In the samples collected from both adherent and non-adherent haemocytes, a total of 62 surface determinants were identified. Among these, 11 were found to be novel haemocyte surface determinants (not previously detected in the unfractionated haemocyte samples). Notably, only one protein, Transmembrane 9 superfamily member, was exclusively present in the adherent fraction (Figure 8) while 24 surface determinants were unique to the non-adherent fraction.

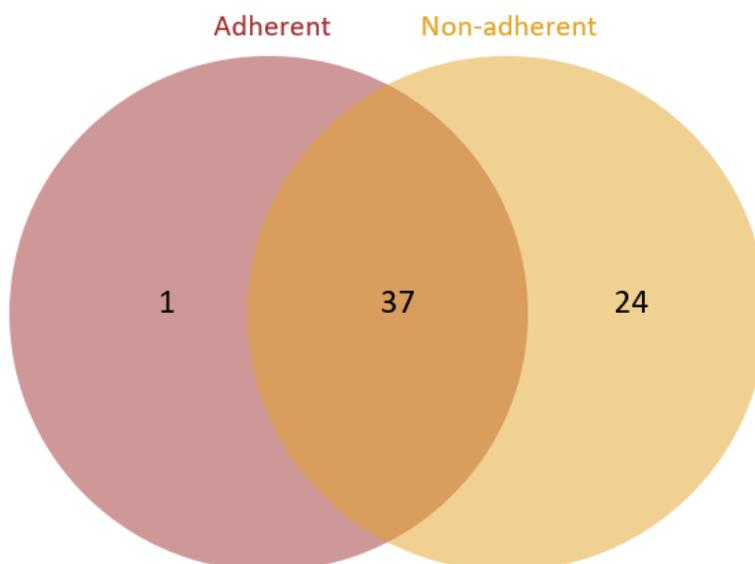


Figure 8. Venn diagramme of the surface determinants for the adherent and non-adherent fractions.

3.4 Transcriptomics

3.4.1 Fragment analyser

The RNA isolated from adults and larvae had a concentration of 101.1739 ng/ μ l and 44.6522 ng/ μ l respectively, which is sufficient for RNA sequencing, and with a quality score (RQN) of 4.0 and 4.3 respectively. The quality of the RNA was assessed with the Agilent 5200 Fragment Analyzer system and the size distribution of the RNA fragments is shown in Figure 9. The prominent pink peak corresponds to the 18S ribosomal RNA and the lower marker (LM) serves as an internal reference to ensure accurate sizing of the samples.

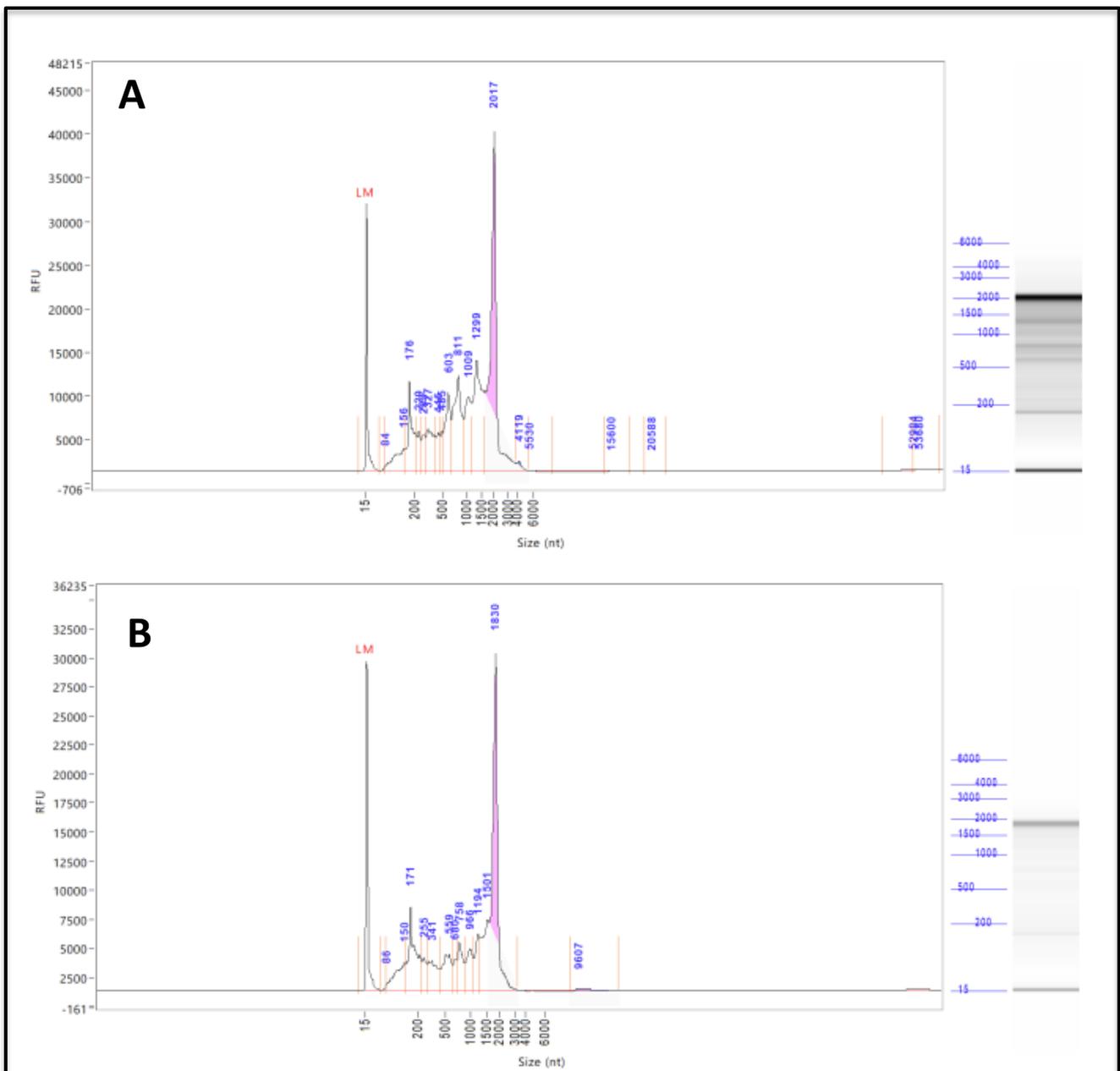


Figure 9. Fragment Analyser results. (A) Adult RNA sample (B) Larval RNA sample (LM= Lower Size Marker, RFU = Relative Fluorescence Units). In insects, the 28S rRNA contains an endogenous hidden break, resulting in two fragments that migrate close to the 18S rRNA (pink peak) during denaturation. The RQN value is determined by the ratio of the 18S to 28S ribosomal RNA peaks and the separation between these peaks. However, this ratio is affected by the hidden break, making the RQN unreliable for most insects (Winnebeck et al., 2010).



3.4.2 RNA sequencing

Table 4 shows the read counts and sequencing quality scores for each sample. Notably, 93.98% and 93.80% of the reads have a base call accuracy of at least 99.9%, while 98.34% and 98.32% of the reads have a base call accuracy of at least 99%. In addition, the GC content found in our samples aligns with the average genomic GC content of *Apis mellifera*, ranging from 33% to 39% (Jørgensen et al., 2007). The quality scores and the GC content indicate reliable, high-quality data.

Table 3. Statistics of clean data (RNAseq).

Sample	Clean Reads	Clean Base	Read Length	Q20 (%)	Q30 (%)	GC (%)
Adult	24,089,143	7,226,742,900	PE150	98.34	93.98	37.53
Larva	24,030,382	7,209,114,600	PE150	98.32	93.80	37.57

For all the surface determinants identified in the proteomic analysis, the corresponding transcript was found except for one. This surface determinant was not mentioned in the reference transcriptome, which explains why its corresponding transcript was not detected.

3.5 List of the surface determinants of *Apis mellifera* haemocytes

The supplementary data contains a comprehensive list of surface determinants of *A. mellifera* haemocytes, identified through proteomic analysis. A total of 163 surface determinants were identified. This list provides an overview of each surface determinant, specifying the developmental stage and buffer in which it was found, as well as whether it was detected in adhering or non-adhering haemocytes. Additionally, the determinants are ranked based on the reliability of their identification in LC-MS/MS, measured by the $-10\log P$ value, with the most reliable identifications listed at the top.

4. Discussion and conclusions

The goal of deliverable 5.1 was to list the surface determinants of haemocytes. In a first step, an appropriate buffer 1) preserving trypsin activity and 2) ensuring haemocyte viability was defined. The DBPS buffer was the only one in which trypsin was active and the viability of the haemocytes was conserved. EDTA (0.02%) was added to prevent coagulation and melanisation.

Until now, haemocytes have been mostly characterized based on their morphology, generally following the classification of Van Steenkiste (1988), who described the following subpopulations: granulocytes, plasmacytes, oenocytoids, coagulocytes and prohaemocytes. However, in this experiment, with the use of MGG staining, it was only possible to differentiate between granulocytes and plasmacytes. Additionally, prohaemocytes are small, round cells; oenocytoids are very similar to granulocytes; and coagulocytes are very unstable cells that mostly burst during sampling, leaving only a free nucleus. Morphological characterization can be challenging and fails to objectively classify subpopulations. For instance, the differentiation between plasmacytes and prohaemocytes, and between oenocytoids and granular cells is only possible by interference contrast or phase contrast microscopic examination (de Graaf et al., 2002).

This research aims to classify subpopulations of cells based on objective criteria. This has been achieved for various cell types in different organisms through the use of cell surface markers, which serve as targets for immunophenotyping. These targets are systematically named using the Cluster of Differentiation (CD) terminology. Initially applied to human leukocytes, the CD nomenclature has since been extended to



numerous other cell types and organisms, including mice (Iyer et al., 2013). The initial step in identifying valuable CD markers involves mapping the surface proteome using the LC-MS/MS approach.

The LC-MS/MS data collected was extensive, particularly for the samples in ABC buffer. The high number of intracellular proteins identified suggests cell lysis caused by the ABC buffer. Some of the identified surface determinants are associated with phagocytosis, such as Sortilin-related receptor (Zhu et al., 2022), Dyslexia-associated protein (Niu et al., 2022), Transmembrane 9 superfamily member (Bergeret et al., 2008), Basigin (Yang et al., 2023), Ninjurin-2 (Mishra et al., 2022), V-type proton ATPase (Mishra et al., 2022) and more. These phagocytosis-related proteins could potentially serve as markers for granulocytes considering that granulocytes are the major phagocytic cells in the haemolymph of *A. mellifera* (Richardson et al., 2018). Remarkably, the Transmembrane 9 superfamily member (TM9SF) was the only protein exclusively detected in the adherent fraction (containing all granulocytes). Sortilin-related receptor, along with other phagocytosis-related proteins, was identified in both fractions. Since plasmatocytes are not typically phagocytic cells in *Apis mellifera* (Gábor et al., 2020), this suggests that these proteins could have another function in plasmatocytes.

Additionally, several proteins that were previously associated with plasmatocytes were identified: Neuroglian (Nardi et al., 2006; Yokoi et al., 2018), Fasciclin-3 (Hirschhäuser et al., 2023), Integrin beta (Levin et al., 2005) and Prominin-like protein (Hu et al., 2021). These plasmatocyte-linked proteins were expected to be found in both fractions, as both contain plasmatocytes. However, Neuroglian, Fasciclin-3 and Integrin beta were identified only in the non-adherent fraction. This discrepancy is likely due to the presence of distinct plasmatocyte subpopulations, which differ in adhering ability and surface determinants. At last, proteins with limited or no literature regarding their relevance to haemocytes have been identified. Examples include CKLF-like MARVEL transmembrane domain-containing protein 8, Nephrylin-1 and protein FAM151B, which could potentially lead to novel discoveries in the context of haemocyte function.

Besides the valuable proteomics data, the transcriptomic data provided complementary insights. The RNA concentration achieved was 10 ng/μl, which was adequate for our needs. The RNA Quality Number (RQN) values were 4.0 and 4.3 on a scale of 10. These seemingly low values can be ignored due to the unique 28S ribosomal structure of insect RNA, which leads to unreliable RQN values (Winnebeck et al., 2010). The sequencing was successful, allowing assessment of transcript abundances. Each protein identified in the proteomic data was also present in the transcriptomic data, affirming the reliability of our proteomic data. The only exception was non-specific serine/threonine kinase, which was absent because it was not included in the reference transcriptome of *A. mellifera*.

The next step involves selecting targets for monoclonal antibody production, with the focus on being able to discover subpopulations of haemocytes. Potential differentiating targets will be selected based on existing literature about haemocytes. It is planned to develop monoclonal antibodies against nine targets, organized into three sets of three. For the first set, priority will be given to surface determinants present in multiple samples, thereby reducing the risk of coincidental identification. Additionally, results obtained using the DPBS buffer will minimize the risk of identifying intracellular proteins, as this buffer prevents cell lysis. Moreover, it will be verified whether the LC-MS/MS analysis specifically detected extracellular peptides or intracellular peptides of the protein targets, favouring the detection of extracellular peptides. Last but not least, the selected targets must be specific to avoid cross-reactivity of the monoclonal antibodies with other proteins. Hence, general proteins like ABC transporters and sodium/potassium channels will be avoided.

In conclusion, the experiments conducted for Deliverable 5.1 and the resulting list of surface determinants provide a solid foundation for selecting the initial targets for monoclonal antibody production



5. Supplementary data

Table 4. Summary of the haemocyte surface determinants. Two buffers were tested for each developmental stage (DPBS = Dulbecco's phosphate buffered saline, ABC = ammonium bicarbonate). Additionally, the haemocytes were fractionated into adhering and non-adhering haemocytes and the surface determinants were identified for each fraction. $-10\log P$ is a score that represents the reliability of an identification. A higher score indicates greater reliability.

Protein name	Accession code	Adult-DPBS	Larva-DPBS	Adult-ABC	Larva-ABC	$-10\log P$	Adherent	Non-adherent
Spectrin alpha chain	A0A7M7IE11			X	X	389.64	X	X
Spectrin beta chain	A0A7M7GUT9		X		X	371.21	X	X
Uncharacterized protein LOC410829	A0A7M7MWZ2			X	X	342.96	X	X
Sodium/potassium-transporting subunit alpha ATPase	A0A7M7IHQ2	X	X	X	X	337.52	X	X
Calcium-transporting ATPase	A0A7M7IJ13	X	X			276.16		
Neprilysin-1	A0A7M7GU97	X		X	X	271.62	X	X
Fibrous sheath CABYR-binding protein-like	A0A7M7L314		X	X	X	252.46		X
Sortilin-related receptor	A0A7M7H280		X	X	X	250.59	X	X
Calcium-transporting ATPase	A0A7M7MM31		X			242.63	X	X
Uncharacterized protein LOC411253	A0A7M7IRH3			X	X	228.64	X	X

Talin-1	A0A7M7SRL2	X	X	X	X	216.69		X
V-type proton ATPase subunit a	A0A7M7MNJ7		X			211.26		
V-type proton ATPase subunit a	A0A7M7MNN2	X	X		X	202.56		
Flotillin-2	A0A7M7GX08				X	199.55		X
Laminin subunit alpha	A0A7M7MR84			X	X	199.55	X	X
Neuroglian	A0A7M7R5T3		X	X	X	194.05		X
Spectrin beta chain, non-erythrocytic 5	A0A7M7GKC3			X	X	193.76		X
Tyrosine-protein kinase-like otk	A0A7M7GZA9		X	X	X	193.26		
Sodium/potassium-transporting subunit beta-2	A0A7M7R6N0	X		X		193.11	X	X
Motile sperm domain-containing protein 2	A0A7M7TEZ2			X	X	192.63	X	X
Secretory associated protein	A0A7M7R6Q3	X				189.7		
Multidrug resistance protein homolog 49	A0A7M7ILG0		X	X	X	188.51		
Protein FAM151B	A0A7M7L1B9	X		X	X	186.51	X	X
V-type proton ATPase proteolipid subunit	A0A8U0WQ65			X	X	179.89		



Basigin	A0A7M7IUP5		X	X	176.57		
Multidrug resistance-associated protein 1	A0A7M7GLJ8		X	X	174.49		X
Guanine nucleotide-binding protein G(Q) subunit alpha	A0A7M7ILG7			X	169.59	X	X
Extended synaptotagmin-1	A0A7M7GXP4		X	X	163.55	X	X
Monocarboxylate transporter 3	A0A7M7L296		X	X	161.35	X	X
Syntaxin-7	A0A7M7GXH1	X	X	X	160.79	X	X
ATP-binding cassette sub-family B member 6	A0A7M7G0A0	X			159.21		
ATP-binding cassette sub-family G member 1	A0A7M7LPJ7		X	X	155.03	X	X
Receptor protein-tyrosine kinase	A0A7M7SPC2			X	153.76	X	X
Non-specific serine/threonine protein kinase	A0A7M7MPF5			X	147.99		
Dystroglycan 1	A0A7M7L362		X		141.26		
Leucine-rich neuronal protein 3	A0A7M7GVU5		X	X	139.06	X	X
Syntaxin-12	A0A7M7RBR7		X	X	137.4		
Low-density lipoprotein receptor	A0A7M7L3H7	X	X	X	133.56		



Microtubule-associated protein futsch	A0A7M7L425		X	X	133.56		X
NPC intracellular cholesterol transporter 1	A0A7M7GZ33		X	X	129.41		
Dyslexia-associated protein KIAA0319	A0A7M7IEU7	X	X	X	127.86		
Piezo-type mechanosensitive ion channel component	A0A7M7SRJ6		X	X	127.56		
Uncharacterized protein LOC552843	A0A7M7RDU4		X	X	126.76	X	X
Phospholipid-transporting ATPase	A0A7M7GZR3		X	X	125.89		
Integrin beta	A0A7M7M1H3			X	125.34	X	X
Integrin alpha-8	A0A7M7MWT1		X	X	124.58		
Lysosome-associated membrane glycoprotein 1	A0A7M7R9E0	X	X	X	120.24	X	X
Zinc transporter 1	A0A7M7MVT4		X	X	119.38		
Sialin	A0A7M7L2G0		X	X	119.18		
Dipeptidase	A0A7M7MS33		X	X	118.68		X
Cell division control protein 42	A0A7M7MW83		X	X	114.4		
Aminopeptidase	A0A7M7L4N0		X		113.43		



Adenylyl cyclase-associated protein	A0A7M7R415			X	X	112.07	X	X
CD109 antigen	A0A7M7LT13			X	X	111.51		
Fibrillin-2	A0A7M7MVC3			X	X	109.59	X	X
Solute carrier family 12 member 7	A0A7M7R6R0	X	X	X	X	107.54		
Protein draper	A0A7M7SRS1		X	X	X	106.75		
Lachesin	A0A7M7R7G1		X	X	X	105.4	X	X
SPRY domain-containing protein 7	A0A7M7R3T0			X	X	104.4		
Multiple epidermal growth factor-like domains protein 11	A0A7M7GMV8			X		102.33	X	X
Integrin alpha-PS2	A0A7M7FXW2			X		101.29		
Low-density lipoprotein receptor-related protein 2	A0A7M7MMU4			X	X	101.27		
Calcium-transporting ATPase	A0A7M7LIA5			X	X	101.26		
Alpha/beta hydrolase domain-containing protein 17B	A0A7M7R934			X	X	100.39		
Uncharacterized protein LOC410600	A0A7M7IIX0			X	X	100.27		
Uncharacterized protein LOC409956	A0A7M7GKH6	X		X	X	99.87		X



Major facilitator superfamily (MFS) profile domain-containing protein	A0A7M7GZC1		X		99.8		
Synaptosomal-associated protein	A0A7M7GVB5		X	X	97.51		
Uncharacterized protein LOC408874	A0A7M7MKK8		X	X	97.28		
Plexin-A4	A0A7M7H236	X	X	X	97.09		X
DnaJ homolog subfamily C member 5	A0A7M7MUF5		X	X	95.71		
Flotillin-1	A0A7M7IGE5		X	X	93.59		X
DE-cadherin	A0A7M7GV76		X	X	92.89	X	X
Proton-coupled amino acid transporter-like protein pathetic	A0A7M7TFK1		X		92.19		X
Twinfilin	A0A7M7GQ01		X	X	90.6		
Fascin-3	A0A7M7MS04		X		90.13		X
CSC1-like protein 2	A0A7M7R3J0		X	X	87.72		
Ninjurin-2	A0A7M7FZF3		X		87.12	X	X
Protein spinster	A0A7M7GVT6		X		84.61		
Prominin-like protein	A0A7M7GR93		X	X	80.44		
Carboxypeptidase D	A0A7M7L4P0		X	X	77.27	X	X
Integrin alpha-8-like	A0A7M7L551		X		75.84		X



Guanine nucleotide-binding protein G(I) subunit alpha	A0A7M7TFD6		X		74.97	X	X
Plexin domain-containing protein 2	A0A7M7GPP4	X	X		74.57		X
Protein GPR107	A0A7M7LKL0	X	X		73.86		
Phosphatidylcholine:ceramide cholinephosphotransferase 2	A0A7M7GNV1	X	X		72.38		
Vinculin	A0A7M7GRL7	X	X		72.37		
Adenylate cyclase-stimulating G alpha	A0A7M7R568	X	X		72.28		
Multidrug resistance-associated prot. lethal 203659	A0A7M7GQY5	X			72.16		
Multidrug resistance-associated protein 4	A0A7M7GZ64	X	X		70.17		
Major facilitator superfamily domain-containing protein 1	A0A7M7R2X1	X	X		70		
Sodium/potassium-transporting ATPase subunit beta-2	A0A7M7H2X0	X			69.8		
Integrin beta	A0A7M7SR40	X			69.01		
Neurocalcin	A0A7M7R5H8	X	X		68.61		X



Guanine nucleotide-binding protein subunit alpha	A0A7M7R523	X	X	66.72		
Quiver family u-PAR/Ly-6-like domain protein	A0A7M7LNV0	X		66.65		
ATP-binding cassette sub-family G member 4	A0A7M7LSW0	X	X	65.29		
Probable chitinase 2	A0A7M7MTN1	X		62.86		
Amino acid transporter	A0A7M7L538	X	X	62.78		
Transmembrane superfamily member	⁹ A0A7M7H1B1		X	62.52		
Equilibrative nucleoside transporter 1	A0A7M7MLE9	X	X	61.77		
Multidrug resistance-associated protein 7	A0A7M7G2M8	X	X	61.47		
Transmembrane superfamily member	⁹ A0A7M7RAW8	X		60.85		
Metal transporter CNNM4	A0A7M7MUZ2	X	X	58.25		
V-type proton ATPase subunit a	A0A7M7ILA5	X		58.17		
T-cell immunomodulatory protein	A0A7M6UEA1	X	X	56.72		
Protein 4.1	A0A7M7L620	X		56.25	X	X
Chloride channel protein	A0A7M7H0V2	X	X	54.9		



Protein windpipe	A0A7M7MP55			X				53.94
Protein-tyrosine-phosphatase	A0A7M7MVA0						X	53.25
Monocarboxylate transporter 12	A0A7M7TF78		X	X		X		50.25
Sodium/potassium-transporting ATPase subunit beta-2	A0A7M7MK83			X		X		50.01
Neogenin	A0A7M7G330			X				48.8
Proton-coupled amino acid transporter-like protein pathetic	A0A7M7GS47						X	47.92
Frizzled-2	A0A7M7LM65						X	47.1
Neprilysin-2	A0A7M7R4N2			X		X		46.5
Facilitated trehalose transporter Tret1	A0A7M7LNR9			X		X		45.88
Uncharacterized protein LOC726407	A0A7M7FYB7						X	45.04
Probable serine incorporator	A0A7M7GVF6			X		X		44.18
Angiotensin-converting enzyme	A0A7M7GL03			X				41.43
CKLF-like MARVEL transmembrane domain-containing protein 8	A0A7M7R2R5	X	X	X				41.24
Aquaporin AQPAn.G	A0A7M7GZE5			X		X		41.23



Uncharacterized protein LOC725078	A0A7M7GZX8	X		40.62	
LMBR1 domain-containing protein 2 homolog	A0A7M6UPF4	X		39.85	
Fasciclin-2	A0A7M7IHD9	X	X	39.64	
Uncharacterized protein LOC100577325	A0A7M7GDK4	X	X	39.55	
Furin-like protease 2	A0A7M7MNF8	X		39.06	
Protein tyrosine phosphatase type IVA 1	A0A7M7M4Z1		X	38.38	
Tyrosine-protein kinase receptor	A0A7M7MLW6	X	X	37.92	
Fasciclin-1	A0A7M7MN23	X		36.8	X
Phospholipid-transporting ATPase	A0A7M7IER3	X		36.37	
ABC transporter A family member 1	A0A7M7R6V4	X		35.99	
Protein-tyrosine- phosphatase	A0A7M7MTK5	X		35.57	
Tetraspanin	A0A7M7LLV6	X		34.75	
Uncharacterized protein LOC727639	A0A7M7L7C0	X		34.02	
Solute carrier organic anion transporter family member	A0A7M7IHL0	X	X	33.85	



Actin interacting protein 3-like C-terminal domain-containing protein	A0A7M7GRN0		X		33.69		
Integrin alpha-PS1	A0A7M7RDN6		X		33.23		
Uncharacterized protein LOC100578472	A0A7M7GIN3			X	32.48		X
Ribonuclease kappa-B	A0A7M7FYX7		X		32.16		
Protein Malvolio	A0A7M7RC52				32.05	X	
Melanotransferrin	A0A7M7R699		X	X	30.3		
Uncharacterized protein LOC724680	A0A7M7GYR5		X		30.23		
Non-specific serine/threonine kinase	A0A7M7IM38		X		30.04		
G-protein coupled receptor Mth2	A0A7M7G1V7		X		29.88		
ABC transporter domain-containing protein	A0A7M7SS92		X		29.21		
Protein lap4	A0A7M7L0E6	X	X		28.56		
Methuselah-like protein 10	A0A8U0YCD9		X	X	27.12		
Leucine-rich repeat-containing protein 70	A0A7M7IRS5		X		26.68	X	X
Anion exchange protein	A0A7M7MRD7			X	26.37		



Collagen alpha chain CG42342	A0A7M7SR99	X	X	25.99		
Innexin	A0A7M7LN93		X	X	25.9	
Uncharacterized protein LOC551356	A0A7M7MXR6					X
Unc-112-related protein	A0A7M7TET9					X
Non-specific serine/threonine kinase	A0A7M7MMU6					X
Uncharacterized protein LOC100577517	A0A7M7L3P4					X
Leucine carboxyl methyltransferase 1	A0A7M7R3J8					X
Integrin beta	A0A7M7MS21					X
V-type proton ATPase subunit a	A0A7M7SQF7				X	X
Transmembrane superfamily member ⁹	A0A7M7RBZ4				X	
Aldehyde dehydrogenase	A0A7M7MNI1				X	X
Uncharacterized protein LOC411209	A0A7M7L3J9				X	X
Uncharacterized protein LOC726696	A0A7M7FZ72				X	X



References

- Bergeret, E., Perrin, J., Williams, M., Grunwald, D., Engel, E., Thevenon, D., Taillebourg, E., Bruckert, F., Cosson, P. & Fauvarque, M. O. (2008). TM9SF4 is required for *Drosophila* cellular immunity via cell adhesion and phagocytosis. *Journal of Cell Science*, 121(20), 3325–3334. <https://doi.org/10.1242/jcs.030163>
- Boraldi, F., Lofaro, F. D., Accorsi, A., Ross, E. & Malagoli, D. (2019). Toward the molecular deciphering of *Pomacea canaliculata* immunity: first proteomic analysis of circulating hemocytes. *Proteomics*, 19(4). <https://doi.org/10.1002/pmic.201800314>
- Butolo, N. P., Azevedo, P., de Alencar, L. D., Domingues, C. E. C., Miotelo, L., Malaspina, O. & Nocelli, R. C. F. (2020). A high quality method for hemolymph collection from honeybee larvae. *PLoS ONE*, 15(6). <https://doi.org/10.1371/journal.pone.0234637>
- Carroll, S. L., & Coyne, V. E. (2021). A proteomic analysis of the effect of ocean acidification on the haemocyte proteome of the South African abalone *Haliotis midae*. *Fish and Shellfish Immunology*, 117, 274–290. <https://doi.org/10.1016/j.fsi.2021.08.008>
- de Graaf, D. C., Dauwe, R., Walravens, K., & Jacobs, F. J. (2002). Flow cytometric analysis of lectin-stained haemocytes of the honeybee (*Apis mellifera*). *Apidologie*, 33(6), 571–579. <https://doi.org/10.1051/apido:2002041>
- Gábor, E., Cinege, G., Csordás, G., Rusvai, M., Honti, V., Kolics, B., Török, T., Williams, M. J., Kurucz, É., & Andó, I. (2020). Identification of reference markers for characterizing honey bee (*Apis mellifera*) hemocyte classes. *Developmental and Comparative Immunology*, 109. <https://doi.org/10.1016/j.dci.2020.103701>
- Hernández-Pérez, A., Zamora-Briseño, J. A., Ruiz-May, E., Pereira-Santana, A., Elizalde-Contreras, J. M., Pozos-González, S., Torres-Irineo, E., Hernández-López, J., Gaxiola-Cortés, M. G., & Rodríguez-Canul, R. (2019). Proteomic profiling of the white shrimp *Litopenaeus vannamei* (Boone, 1931) hemocytes infected with white spot syndrome virus reveals the induction of allergy-related proteins. *Developmental and Comparative Immunology*, 91, 37–49. <https://doi.org/10.1016/j.dci.2018.10.002>
- Hirschhäuser, A., Molitor, D., Salinas, G., Großhans, J., Rust, K., & Bogdan, S. (2023). Single-cell transcriptomics identifies new blood cell populations in *Drosophila* released at the onset of metamorphosis. *Development (Cambridge)*, 150(18). <https://doi.org/10.1242/dev.201767>
- Hu, J., Du, Y., Meng, M., Dong, Y., & Peng, J. (2021). Development of two continuous hemocyte cell sublines in the Asian corn borer *Ostrinia furnacalis* and the identification of molecular markers for hemocytes. *Insect Science*, 28(5), 1382–1398. <https://doi.org/10.1111/1744-7917.12854>
- Identification of novel markers for mouse CD4+ T follicular helper cells.* (n.d.).
- Jørgensen, F. G., Schierup, M. H., & Clark, A. G. (2007). Heterogeneity in regional GC content and differential usage of codons and amino acids in GC-poor and GC-rich regions of the genome of *Apis mellifera*. *Molecular Biology and Evolution*, 24(2), 611–619. <https://doi.org/10.1093/molbev/msl190>
- Levin, D. M., Breuer, L. N., Zhuang, S., Anderson, S. A., Nardi, J. B., & Kanost, M. R. (2005). A hemocyte-specific integrin required for hemocytic encapsulation in the tobacco hornworm, *Manduca sexta*. *Insect Biochemistry and Molecular Biology*, 35(5), 369–380. <https://doi.org/10.1016/j.ibmb.2005.01.003>
- Mishra, R., Hua, G., Bagal, U. R., Champagne, D. E., & Adang, M. J. (2022). *Anopheles gambiae* strain (Ag55) cultured cells originated from *Anopheles coluzzii* and are phagocytic with hemocyte-like gene expression. *Insect Science*, 29(5), 1346–1360. <https://doi.org/10.1111/1744-7917.13036>
- Nardi, J. B., Pilas, B., Bee, C. M., Zhuang, S., Garsha, K., & Kanost, M. R. (2006). Neuroglian-positive plasmatocytes of *Manduca sexta* and the initiation of hemocyte attachment to foreign surfaces. *Developmental and Comparative Immunology*, 30(5), 447–462. <https://doi.org/10.1016/j.dci.2005.06.026>
- Niu, G. J., Zhu, X. X., Lu, P. Y., Yang, M. C., Yuan, W. J., & Wang, J. X. (2022). The dyslexia-associated KIAA0319-like protein from *Marsupenaeus japonicus* recognizes white spot syndrome virus and restricts its replication. *Aquaculture*, 550. <https://doi.org/10.1016/j.aquaculture.2021.737827>
- Richardson, R. T., Ballinger, M. N., Qian, F., Christman, J. W., & Johnson, R. M. (2018). Morphological and functional characterization of honey bee, *Apis mellifera*, hemocyte cell communities. *Apidologie*, 49(3), 397–410. <https://doi.org/10.1007/s13592-018-0566-2>

- Strachecka, A., Migdał, P., Kuszewska, K., Skowronek, P., Grabowski, M., Paleolog, J., & Woyciechowski, M. (2021). Humoral and cellular defence mechanisms in rebel workers of *Apis mellifera*. *Biology*, *10*(11). <https://doi.org/10.3390/biology10111146>
- Tang, X., Li, X., Zhai, F., Xing, J., Sheng, X., & Zhan, W. (2018). Analysis and identification of tyrosine phosphorylated proteins in hemocytes of *Litopenaeus vannamei* infected with WSSV. *Fish and Shellfish Immunology*, *82*, 84–91. <https://doi.org/10.1016/j.fsi.2018.08.017>
- Winnebeck, E. C., Millar, C. D., & Warman, G. R. (2010). Why does insect RNA look degraded. In *Journal of Insect Science* | www.insectscience.org (Vol. 10). www.insectscience.org
- Yang, L., Wang, Z. ang, Gan, Y., Zuo, H., Deng, H., Weng, S., He, J., & Xu, X. (2023). Basigin binds bacteria and activates Dorsal signaling to promote antibacterial defense in *Penaeus vannamei*. *Fish and Shellfish Immunology*, *142*. <https://doi.org/10.1016/j.fsi.2023.109123>
- Zheng, S. Y., Pan, L. X., Cheng, F. P., Jin, M. J., & Wang, Z. L. (2023). A global survey of the full-length transcriptome of *Apis mellifera* by single-molecule long-read sequencing. *International Journal of Molecular Sciences*, *24*(6). <https://doi.org/10.3390/ijms24065827>
- Zhu, K., Yang, F., & Li, F. (2022). Molecular markers for hemocyte subpopulations in crayfish *Cherax quadricarinatus*. *Developmental and Comparative Immunology*, *132*. <https://doi.org/10.1016/j.dci.2022.104407>

