# Final report

**Dissecting equine respiratory disease through single-cell transcriptomics: a pilot study**

## Project number: H-19-47-475

**Project period:** 200530-210630


## Main applicant:
Amanda Raine, Uppsala University
Email: Amanda.Raine@medsci.uu.se

## Co-applicant(s):
Miia Riihimäki, Swedish University of Agricultural Sciences
Ida Waern, Swedish University of Agricultural Sciences
Jessica Nordlund, Uppsala University


## Part 1: Detailed summary


Inflammatoriska luftvägssjukdomar hos häst är, efter ortopediska problem, den vanligaste orsaken till nedsatt prestation hos hästar som lever i stallmiljö. Hästar kan lida av allvarliga astmaliknande problem eller lindrigare grad av inflammation i luftvägarna som enbart visar sig genom prestationsnedsättning. Hästens astma (ekvin astma) har varit föremål för forskning under de senaste decennierna men fortfarande kvarstår många frågetecken rörande de cellulära mekanismerna bakom de astmaliknande symptomen. Man kan dela in ekvin astma i flera undergrupper som troligen skiljer sig åt immunologiskt och svarar olika på behandling. Vi vet att immunceller i hästarnas lunga är involverade i sjukdomsförloppet och med hjälp av lungsköljprov kan man få ett prov av de immunceller som befinner sig i lungan.

Det finns ett stort behov av nya analysmetoder inom forskningsområdet för att förstå mekanismerna bakom inflammatoriska luftvägssjukdomar hos häst. Det här projektet var en pilotstudie där vi framgångsrikt har testat och optimerat en ny storskalig teknik för att studera genuttryck i tusentals enskilda immunceller i lungan. Principen är att ett mikrofluidiksystem används för att isolera enskilda celler i mycket små droppar i en emulsion. I dropparna frisätts cellens mRNA-molekyler som märks in med en cell-specifik ’barcode’. Cellernas genuttryck (transkriptom) sekvenseras sedan med sk massiv parallell sekvensering och sedan används bioinformatiska metoder för att rekonstruera varje cells genuttryck. Lungsköljprov från hästar med diagnosticerad astma samt från friska hästar har använts i studien. Genom att studera variation i genuttryck på individuell cellnivå har vi nu konstruerat ’cellulära kartor’ över sammansättningen av cellerna i luftvägarna på en mycket detaljerad nivå.

Med hjälp den här tekniken ska vi nu studera sammansättningen av celltyper i lungan vid astma, identifiera tidigare okända varianter av celltyper som är specifikt involverade i den inflammatoriska processen samt cell-varianter som är karaktäristiska för svårare respektive mildare former av sjukdomen. Vi kommer även framgent använda tekniken för att studera hur genuttryck förändras på enskilda cellnivå efter behandling. Genom ökad


The project has been financed by:
STIFTELSEN HÄSTFORSKNING

kunskap om sjukdomens orsak och fortskridande och hur behandlingssvar ser ut på enskilda cellnivå kan man utveckla bättre diagnostik och mer individanpassade behandlingsstrategier och förebyggande åtgärder.

# Part 2: Main report (max. 10 pages)

## Introduction

*Equine asthma*

Respiratory disorders in horses constitute a major horse-welfare problem in the equine industry and are second only to orthopedic disorders as a cause of poor performance in sporting horses (1-3). Clinical signs of equine asthma include airway obstruction, chronic intermittent cough, increased respiratory effort and mucus accumulation. Similar to its human equivalent, equine asthma is a heterogeneous and complex disease that is associated with both genetic and environmental factors (4-6). There is an established correlation between poor stable environment, feed quality and exacerbation of equine asthma, hence clinical symptoms are often relieved when affected horses are kept outdoors (7-10). In addition to environmental management, treatment includes corticosteroids in order to reverse inflammation and bronchodilating drugs, yet therapy outcome can be difficult to predict. Considering the heterogeneous nature of equine asthma, it is likely not a single disease but rather comprises several subtypes with distinct immunological characteristics. Severe equine asthma (formerly denoted Recurrent Airway Obstruction) is principally diagnosed in middle age and older horses whilst mild-to-moderate equine asthma (formerly Inflammatory Airway Disease) can be diagnosed in horses at any age (2-3). However, onset of mild-to-moderate equine asthma is more frequently observed in younger horses and it is the primary cause of exercise intolerance in sport and race horses. At present there is no clear-cut criteria or biomarkers identified for distinction between subtypes (2-3). Precise identification of asthma subtypes and their pathophysiological differences will be important to enable more precise treatment, prognostics and prevention.

Bronchoalveolar lavage (BAL) is an important clinical tool which reflects immunological and inflammatory processes at the alveolar level and is frequently used to support diagnosis of equine asthma. Traditional BAL cytology however, neither informs about the underlying processes in equine asthma nor has sufficient resolution for refined sub-classification of the disease. Targeted as well as genome-wide gene expression and biochemical measurements performed on bulk population of cells in BAL and bronchial biopsies have provided important insights into the pathophysiology of the disease. However, these studies suffer from the limitations of averaging measured values across different cell types or sub populations of cells and thus the underlying mechanisms of equine asthma remain incompletely defined.

*Single cell genomics in asthma research*

Our current understanding of cell “type” is mainly based on measurements from mixtures of cells in bulk samples, which average inter-cellular differences and prohibit detection of rare, or previously undetected cell populations. A number of next generation sequencing (NGS) based technologies have been developed which enable genome-wide analysis of several modalities such as DNA, chromatin, gene expression and to some extent also surface protein expression, at the level of individual cells, collectively referred to as ‘single cell genomics’. The ‘transcriptome’ refers to the complete set of gene transcripts (RNA species) in a given

The project has been financed by:
STIFTELSEN HÄSTFORSKNING

cell type and '*single cell transcriptomics'* or *'single cell RNA sequencing' (scRNA-seq)* has rapidly become an analysis method of choice in many areas of medical biology.

Using e.g droplet microfluidics, suspensions of viable cells or isolated nuclei of basically any origin (blood, tissue, tumour, etc) can be processed, and many thousands of cells analyzed per sample, without the requirement for prior enrichment of specific cell types (11-12). Cell identity can then be assigned after sequencing based on each single cell's transcriptome profile. Moreover, subtyping of known cell types and discovery of new/rare cell populations are enabled at substantially higher resolution compared to other methods.

*Single cell transcription profiling of horse samples*

To date, only two equine single cell RNA-seq studies have been published, in which phenotyping of equine PBMCs and mesenchymal stromal cells from healthy individual was performed (13-14). Utilizing unbiased single-cell RNA-sequencing approaches to study cell populations obtained from horses diagnosed with equine asthma presents an excellent opportunity to disentangle many of the cellular processes implicated in equine airway disease.

***Aim of the present study:***

In this 1-year pilot study the aim was to perform proof-of-principle experiments and establish single cell RNA-sequencing as a tool for investigating cell-to-cell variation of immune cells residing in the equine alveolar space. To this end we adapted the microfluidic droplet based workflow Drop-Seq (15) for use with equine BAL cells. This included optimizing sample preparation, single cell encapsulation, sequencing and to construct the first single cell transcriptome map of the immune cell landscape in the lung of healthy horses and horses diagnosed with equine asthma

***Long term aims:***

The long term objective of the project is to employ single cell RNA-sequencing to:

- ✓ Discover novel cell types, subtypes of cells and altered cell states implicated in equine asthma
- ✓ Investigate pathogenesis of equine asthma by understanding underlying inflammatory processes
- ✓ Take steps toward a refined disease classification of equine asthma
- ✓ Search for novel biomarkers with potential to predict prognosis and treatment outcome

## Material and methods

*Horses*

BAL cells were obtained from twelve horses (different breeds) diagnosed with equine asthma. The diagnosis was based on clinical- and endoscopic examination and cytology analysis of BAL. In addition, BAL samples were obtained from eight healthy standardbreed trotters with no history of airway disease, owned by the Department of Clinical Sciences (SLU).

*BAL cell sampling and preparation*

All horses airways were examined with endoscopy prior to BAL sampling. Healthy control horses showed no signs of airway inflammation as judged by clinical and endoscopy examination. BAL fluid samples were collected by bronchoscopy after intravenous premedication with sedative (detomidine and butorphanol). A local anaesthetic was instilled at the trachea before performing BAL ( Carbocain®). The BAL sampling was performed with

The project has been financed by:
STIFTELSEN HÄSTFORSKNING

a blind tube with 3 × 100 ml sterile isotonic saline solution (at 37 °C). Recovered BAL fluid was placed on ice, samples were pooled and a 30 ml aliquot was taken for cytology (performed by the Clinical Chemistry Laboratory, University Animal Hospital). 50 ml of the remaining BAL sample was filtered through a 70 µm strainer and BAL cells were spun down at 300g for 5 min. The cell pellet was resuspended in 2 ml of PBS containing 0.01 % ultrapure BSA. Cell viability and concentration were assessed by Acridine Orange/Propidium iodide staining using an automated fluorescence cell counter (Nexcelom K2). Cell samples with a viability score of > 85% were used for downstream single cell isolation. All aspects of the study have been approved by the Regional Ethical Review Board (Dnr5.8.18-20690/2020).

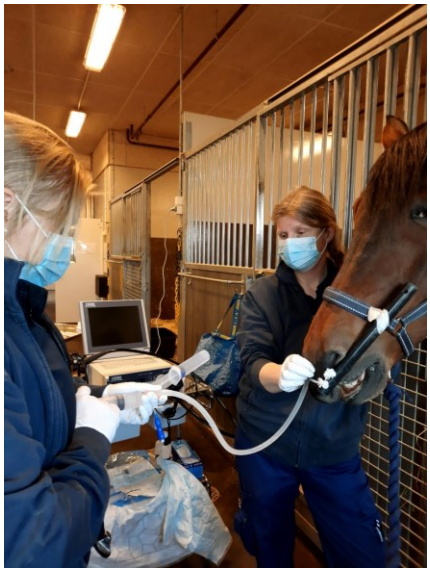

***Figure 1**. Miia Riihimäki DVM, PhD, performing airway endoscopy examination and BAL sampling on a healthy control horse at the University Animal Hospital, SLU*

*Single cell isolation and sequencing library preparation*

Single BAL cells were encapsulated together with microbeads in miniature droplets on a microfluidic device (Nadia, Dolomite Bio), according to a modified Drop-Seq protocol (15). The microbeads, which are coated with oligos comprising a PCR tag, a cell barcode, a unique molecular identifier (UMI) and a poly T sequence (ChemGenes) were resuspended in cell lysis buffer prior to droplet encapsulation. The microfluidic chip was loaded with 250 µl of cell suspension (300 cells/µl) and 250 µl of bead suspension (600 beads/µl). Two chips per sample were run in parallel and the resulting emulsions were pooled. In the miniature droplets the cells are lysed and mRNA is captured on the microbead (Figure 2). The emulsion was broken by filtration through an Uberstrainer filter (Pluriselect) and simultaneously washed in high salt 6X SSC buffer (ThermoFisher)). The beads were collected from the filter in 6X SSC buffer, spun down at 1000g and washed with 5x Maxima RT buffer (ThermoFisher) and finally resuspended in reverse transcription (RT) reaction mixture (Maxima H Minus Reverse Transcriptase, Thermo Fisher) containing a template switch oligo. The RT reaction was incubated for 30 min at 20 °C followed by 90 minutes at 42°C and gentle shaking to keep the beads in suspension. After RT, Exonuclease I was added in order to digest excess oligos. The beads were then washed, first in TE-SDS buffer, then in TE-Tween buffer and finally resuspended in H20. Beads were counted to determine the concentration and aliquoted in a 96 well plate at 5000 beads/well. cDNA was amplified in the 96 well plates for 14 cycles using Terra PCR polymerase (TakaraBio). Amplified cDNA was pooled and purified twice with 0.6x AMPure XP beads (BeckmanCoulter). The integrity of the purified cDNA was checked on a TapeStation HS5000 Screen Tape (Agilent) and concentration was measured by Qubit.

The project has been financed by:

STIFTELSEN HÄSTFORSKNING

Sequencing libraries were prepared from 3 x 1 ng of amplified cDNA using a modified Nextera XT protocol (Illumina Inc). Three Nextera XT reactions were prepared per sample, pooled, quantified by qPCR and sequenced on a NovaSeq6000 (Illumina Inc) with a custom recipe (Read 1: 25 bp, Read 2: 98 bp), at a read depth corresponding to at least 50,000 reads/cell.

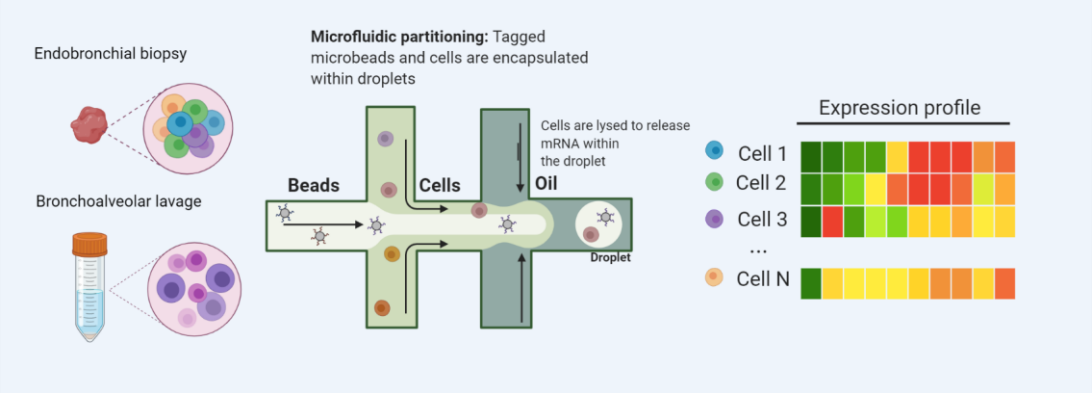


***Figure 2**. Schematic representation of the Drop-Seq single cell isolation and mRNA barcoding.*

*Bioinformatic analysis*

Read 2 (containing the gene mRNA sequence) were mapped to the horse reference genome (EquCab3.0, 16) with STAR v2.7.3. Read 1 (containing barcodes and UMIs) were used for cell demultiplexing, UMI deduplication and barcode quantification performed with either the DropSeqPipe tool (17) or with the corresponding workflows within the PartekFlow bioinformatic software. The EmptyDrops tool (18) were used to distinguish cells from empty droplets. The data were then further quality filtered to discard cells with > 5 % mitochondrial reads (indicative of dying cells), cells with > 5000 counts (putative doublets) and cells with less than 200 detected genes/cell. Furthermore, a noise filtering step was applied to remove non expressed genes and ribosomal protein genes. Downstream analysis was performed either using the R-package Seurat (19) or similar workflows in PartekFlow. Counts were normalized (log2(CPM+1)) and scaled before dimensionality reduction was performed by principal component analysis (PCA). Unsupervised graph-based clustering was performed on the first 25 principal components (based on the PCA scree-plot) and clusters were visualized by Uniform Manifold Approximation and Projection (UMAP). Harmony (20) was used for batch correction whenever data from several horse were clustered and visualized together.

## Results and discussion

During the 1-year project period we have optimized the Drop-Seq sample preparation for use with equine BAL cells. We then sequenced scRNA-seq libraries from twelve horses with equine asthma (collected from patients visiting the respiratory specialist clinic at the University Animal Hospital, SLU), at diagnosis) and eight healthy control horses, in total twenty samples. For each sample we have sequenced 3000-6000 cells, generating a large scRNA-seq data set of in total ~100,000 single cells.

The BAL cell Drop-Seq data was generally of high quality (Figure 3). The fraction of mitochondrial reads was low (less than 5%), which indicates low numbers of dead or dying cells (Figure 2, right panel). The number of genes detected per cell were between 200-2000 genes/cell and 400-5000 counts/cell (middle and left panels of Figure 3) which can

The project has been financed by:
STIFTELSEN HÄSTFORSKNING

be considered in the normal range for single cell RNA Drop-Seq data, especially for species with less well annotated reference genomes (especially annotation of 3'UTR regions as reads are enriched to 3'-end of genes in Drop-Seq).

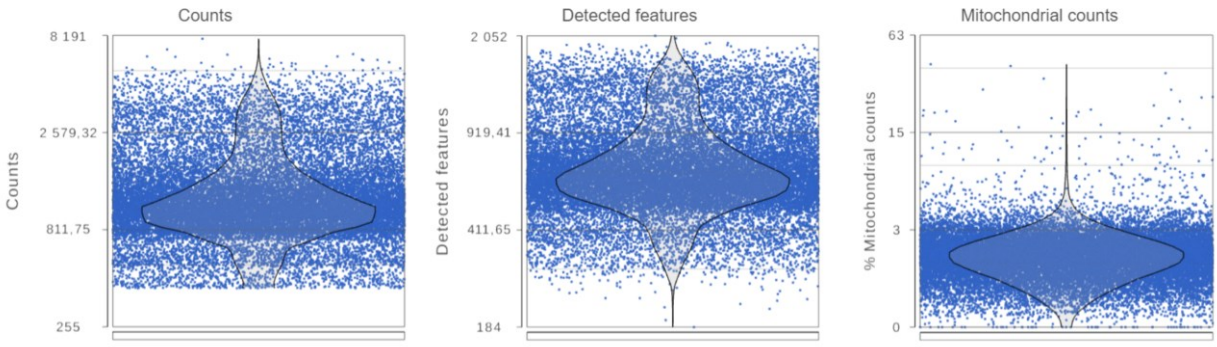


***Figure 3**. Example of quality control parameters for combined Drop-Seq data obtained from five equine asthma patients, in total 29000 cell. Each dot represents one cell. The violin plots show the number of counts/cell, number of genes (features)/cell and mitochondrial counts/cell.*

Unsupervised clustering and cluster annotation based on expressed marker genes identified all the expected cell types in BAL (Figure 4 and 5) i.e; T-cells, alveolar macrophages, neutrophils, mast cells and eosinophils. Examples of cluster specific gene expression markers are visualized in Figure 6. However, when comparing to cytology counts for the same samples, it became evident that the granulocyte populations (neutrophils, mast cells, eosinophils) were underrepresented in the Drop-Seq data (Figure 4 and 5). For instance, in the equine asthma patients the neutrophil count is expected to be > 8 % according to cytology analysis but only 2 % of cells could be annotated as neutrophils in the scRNA-seq data set. The corresponding number for mast cells were 7 % vs 1.2%.

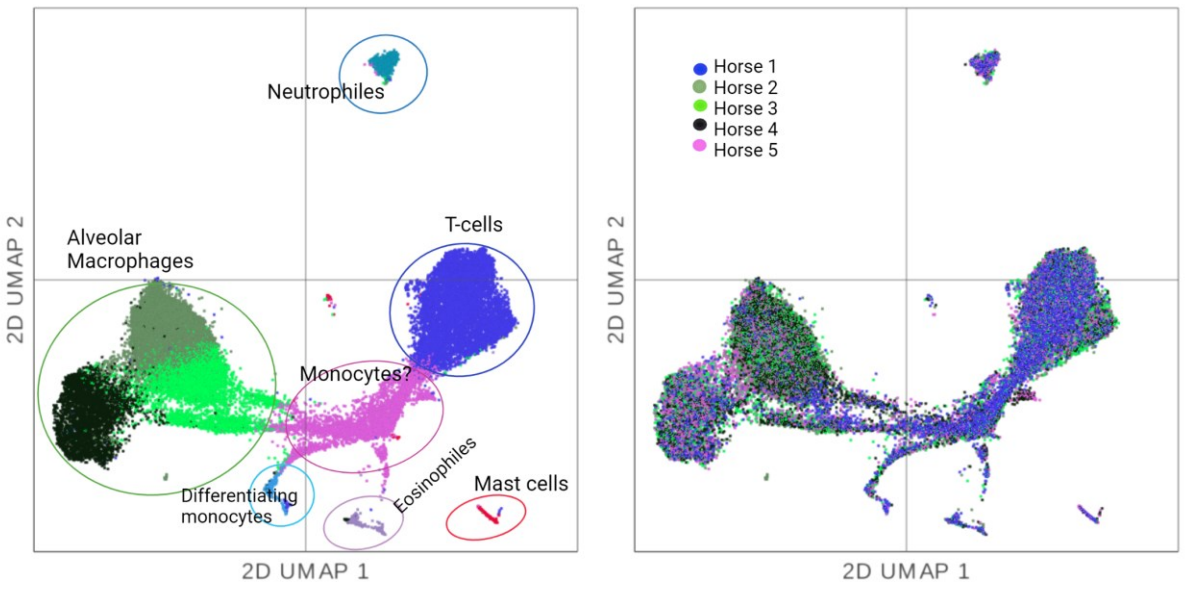


***Figure 4.** Example of UMAP visualization of ~29000 cells combined from five equine asthma patients. Left panel: cluster were annotated based on the expression of specific markers for each cell type. Right panel: the same UMAP plot but here cells are colored according to origin of sample/horse, showing successful correction of batch effects.*

The project has been financed by:
STIFTELSEN HÄSTFORSKNING

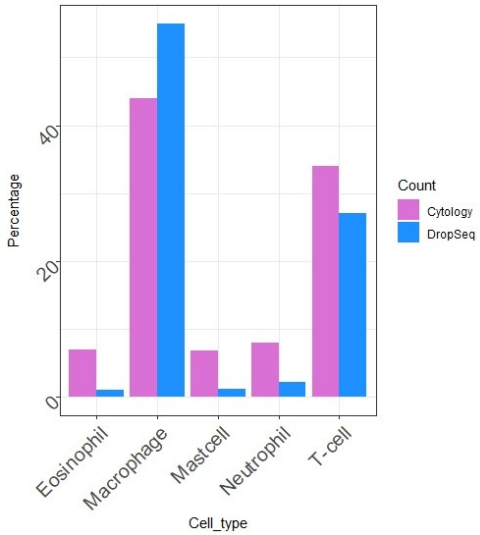


***Figure 5.*** *Barplot showing comparison of the percentage of cell types in equine asthma samples as measured by cytology (purple) and scRNA-seq (blue). The T-cell and Macrophage cell types are in good agreement. In contrast, the number of granulocytes detected in the scRNA-Seq data are only 10-15 % of those counted by cytology.*

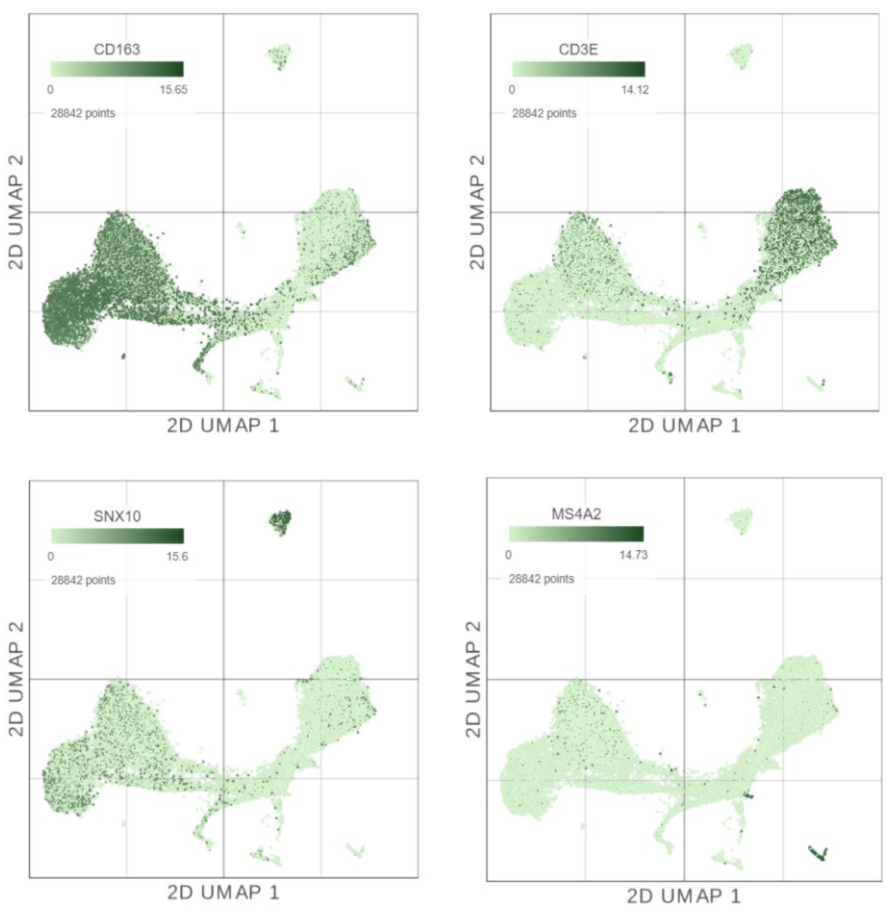


***Figure 6****. Examples of expression of cell cluster specific gene markers analyzed in 29000 equine asthma cells. CD163 in macrophages/monocytes. CD3 in T-cells, SNX10 in neutrophils, MS4A2 in mast cells*

The project has been financed by:

STIFTELSEN HÄSTFORSKNING

Infiltration of neutrophils and/or mast cells to the alveolar space is characteristic for equine asthma and these cell types are important players in the inflammatory process. However, it has become evident (in this study and by others) that fragile granulocytes are notoriously difficult to analyze with droplet based scRNA-seq methods. The lower-than-expected number of granulocytes detected in our data is most likely due to a combination of factors: i) fragile cells may break during prolonged sample handling and droplet encapsulation process ii) the low RNA-content of granulocytes makes them difficult to distinguish in the noisy single cell data, particularly when analyzing them together with large number of cells with higher levels of gene expression, such as e.g macrophages. Despite that, by analyzing combined samples obtained from several horses (n=20) we could in our data set detect sufficient numbers of neutrophils and mast cells for a preliminary view of these populations in equine asthma. However, for overall cell composition analysis and identification of rare neutrophil and mast cell subpopulations involved in equine asthma, the commonly employed droplet microfluidic based scRNA-seq methods are suboptimal. Therefore, we are currently in the process of evaluating a novel gravity-flow based device for single cell RNA-seq which is especially adapted for sensitive cell populations such as granulocytes (HIVE, Honeycomb Biotechnologies).

*Ongoing analysis and future plans*

The analysis of the full BAL scRNA-seq data set (~100 000) cells is now ongoing, where we are mapping out the BAL-cell landscape in more detail, identifying more fine-grained clustered subpopulations and performing differential gene expression analysis between equine asthma samples and healthy controls. These analyses are very computational heavy but a manuscript describing the full data set is planned for submission in Q3-Q4 of 2021.

In the original project plan, analysis of equine lung biopsy cells by scRNA-seq was also included. Due to Covid19 we were unfortunately not able to perform any biopsy sampling during the project period. Instead, we have only sequenced cells from BAL samples although increased the number of BAL samples to twenty (twelve equine asthma patients + eight healthy controls, while the original plan was; BAL+ biopsy from three patients and three controls). It should be mentioned however that the project has been granted additional funding to perform a large study (FORMAS) during 2021-2023, including biopsies and that sampling of biopsies for scRNA-seq are scheduled to be performed in the beginning of 2022. Future plans also include to investigate the single-cell transcriptional response to pharmacological treatment in horses that respond to treatment versus poor-responders.

## Conclusions

In this 1-year pilot project the primary aim was to test and optimize a single cell RNA-seq technique for detailed studies of immune cell populations in equine asthma. To this end we have optimized a Drop-Seq (15) workflow. Although the Drop-Seq method is technically more challenging than commercial scRNA-Seq solutions (e.g 10x Genomics) it has the advantage of being very cost efficient and therefore allowing for analysis of higher numbers of single cells and samples. Based on the results obtained in this pilot we conclude the following:

- A Drop-Seq workflow for analyzing equine single immune cells in bronchoalveolar lavage was successfully implemented and the generated data was of high quality. The evaluation of is based on sequencing of ~100 000 single cells.
- All expected BAL cell types were detected by the Drop-Seq method although granulocyte cell numbers were lower than expected.

The project has been financed by:

STIFTELSEN HÄSTFORSKNING

- Drop-Seq is a cost efficient option for single cell transcriptomics analysis of equine T-cells and macrophages in bronchoalveolar lavage.
- Alternative scRNA-seq methods should ideally be employed for a more comprehensive analysis of equine granulocytes in bronchoalveolar lavage.

## Relevance for the practical horse sector incl. recommendations

Respiratory conditions are a growing welfare issue in horses. Hence novel research strategies which may ultimately lead to enhanced diagnostic precision and better prediction of therapy outcome is of importance. In this project we have taken the first steps to study equine asthma using single cell RNA-sequencing techniques. A first single cell atlas of the immune cell landscape in the lung of healthy horses and horses diagnosed with asthma will generate new insights into the pathophysiology and inflammatory processes in equine asthma. Future directions include to apply the technique to an increased number of horses and take steps toward better disease stratification and to study transcriptional response to treatment at the single cell level. Moreover, for a more complete picture it will be important to also study pathophysiological changes in the cells of the airway walls.

## References

1. Riihimäki et al; A, Inflammatory Response In Equine Airways, Thesis, Swedish University of Agricultural Sciences, 2008
2. Couetil et al, Inflammatory Airway Disease of Horses-Revised Consensus Statement, J Vet Intern Med, 2016
3. Bond et al, Equine asthma: Integrative biologic relevance of a recently proposed nomenclature, J Vet Intern Med, 2018
4. Gerber et al, A, Mixed inheritance of equine airway obstruction, J Vet Intern Med, 2009
5. Gerber et al, B Genetics of upper and lower airway diseases in the horse, Equine Vet J, 2015
6. Mason, eQTL discovery and their association with severe equine asthma in European Warmblood horses, BMC Genomics, 2018
7. Robinson, Coughing, mucus accumulation, airway obstruction, and airway inflammation, Am J Vet Res, 2003
8. Holcombe, S.J. et al, Stabling is associated with airway inflammation in young Arabian horses, Equine Vet J, 2001
9. Laan, T.T et al, Evaluation of cytokine production by equine alveolar macrophages exposed to lipopolysaccharide, Aspergillus fumigatus, and a suspension of hay dust. Am J Vet Res, 2005
10. Gerber et al, Airway mucus in recurrent airway obstruction-short-term response to environmental challenge J Vet Intern Med, 2004
11. Stubbington M et al, Single-cell transcriptomics to explore the immune systemin health and disease, Science, 2017
12. Regev et al, The Human Cell Atlas, Elife, 2017
13. Patel et al, Single-cell resolution landscape of equine peripheral blood mononuclear cells reveal diverse cel types including Tbet+ B cells, BMC Biology, 2021
14. Harman et al, Single-cell RNA sequencing of equine mesenchymal stromal cells from primary donor-matched tissue sources revela functional heterogeneity in immune modulation and cell motility, Stem Cell Res Ther, 2020
15. Macosco et al, Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets, Cell, 2015
16. Kalbfleisch et al, Improved reference genome for the domestic horse increases assembly contiguity and composition, Commun Biol, 2018

The project has been financed by:
STIFTELSEN HÄSTFORSKNING

17. https://github.com/Hoohm/dropSeqPipe
18. Lun et al, EmptyDrops: Distinguishing cells from empty droplets in droplet-based single-cell RNA sequencing data. Genome Biology, 2019
19. Satija et al, Spatial reconstruction of single-cell gene expression data, Nat Biotechnol, 2015
20. Korsunsky et al, Fast, sensitive and accurate integration of single-cell data with Harmony, Nature Methods, 2019

# Part 3: Result dissemination

*State all result dissemination from the financed project into the appropriate section, including information as indicated. Additional rows can be added to the table.*

| | |
|---|---|
| **Scientific publications, *published*** | *Author(s), year, title, journal, Vol, No, pp. (doi/link if applicable)* |
| | |
| | |
| | |
| **Scientific publications, *submitted*** | *Author(s), title* |
| | |
| | |
| | |
| **Scientific publications, *manuscript*** | Riihimäki Miia*, Fegraeus Kim, Waern Ida, Nordlund Jessica and Raine Amanda* |
| | Working title: **Single cell landscape of bronchoalveolar immune cells in healthy horses and horses diagnosed with equine asthma** |
| | *corresponding authors |
| | Will be submitted to high ranked scientific journal in Q3-Q4 2021 |
| **Conference publications/ presentations** | No conferences were attended during 2020-21 due to Covid19 restrictions. We plan to present the results at several conferences during 2021 and 2022, depending on how the Covid19 situation develops |
| | |
| | |
| **Other publications, *media etc.*** | Publication in veterinary trade journal: |
| | **Fokus Luftvägar- Miia Riihimäki-Veterinär och forskare med hästens bästa i blicken**, Svensk Veterinär, Juni 2021 |
| | |
| | |
| **Oral communication, *to horse sector, students etc.*** | If circumstances allow, an equine asthma seminar will be organized at SLU in the fall/winter 2021, otherwise in spring 2022 |
| | |
| | |
| | |
| **Student theses** | *Author/Student, co-authors/supervisors, year, title, type of thesis (doi/link if applicable)* |
| | |
| | |

The project has been financed by:
STIFTELSEN HÄSTFORSKNING

| | |
|---|---|
| | |
| **Other** | Social media: A Twitter account: **Uppsala Equine Respiratory Research @equine asthma** has been launched as a channel for dissemination. |
| | Followers include; veterinarians, researchers, horse owners, racing stables, animal welfare organisations, research institutes etc |
| | |
| | |

The project has been financed by:

STIFTELSEN HÄSTFORSKNING