Use of the human lung tissue explant model for functional biomarker studies in the early phase of infections

INAUGURALDISSERTATION
zur Erlangung des Grades eines Doktors
der Naturwissenschaften
-Doctor rerum naturalium-
(Dr. rer. nat.)

vorgelegt von

Aaqib Sohail
aus Sialkot, Pakistan
Hannover 2020
Angenommen durch den Senat: 19.06.2020

Präsident: Prof. Dr. med. Michael P. Manns
Wissenschaftliche Betreuung: PD Dr. med. Frank Pessler
Wissenschaftliche Zweitbetreuung: Prof. Dr. Armin Braun

1. Referent: PD Dr. med. Frank Peßler
2. Referent: Prof. Dr. Armin Braun
3. Referent: Prof. Dr. Dr. Burkhard Tümmler

Tag der mündlichen Prüfung: 19.06.2020

Prüfungsausschuss

Vorsitz: Prof. Dr. rer. nat. Jürgen Alves
1. Prüfer: PD Dr. med. Frank Peßler
2. Prüfer: Prof. Dr. Armin Braun
3. Prüfer: Prof. Dr. Dr. Burkhard Tümmler
# Contents

Abstract ......................................................................................................................................................... 5

Zusammenfassung ........................................................................................................................................ 7

1. Introduction .......................................................................................................................................... 9
   1.1. Lower respiratory tract infections ............................................................................................ 9
   1.2. Biomarkers of infection ............................................................................................................. 12
   1.3. Models for infection research .................................................................................................. 13
   1.4. Candidate biomarkers ............................................................................................................. 15
   1.5. Itaconic acid ............................................................................................................................. 16
   1.6. Metabolomics in biomarker research ......................................................................................... 17
   1.7. Importance of small noncoding RNAs as biomarkers ............................................................... 18
   1.8. Transcriptomics ........................................................................................................................ 19
   1.9. Single-cell RNA sequencing .................................................................................................... 20

2. Aims of the project ............................................................................................................................. 22

3. Materials and methods ...................................................................................................................... 23
   3.1. Materials ......................................................................................................................................... 23
         3.1.1. Lung tissue donors .............................................................................................................. 23
         3.1.2. Pathogens ............................................................................................................................ 24
         3.1.3. Cell lines .............................................................................................................................. 24
         3.1.4. Software/web tools ............................................................................................................. 24
         3.1.5. Buffers .................................................................................................................................... 25
         3.1.6. Reagents ............................................................................................................................. 26
         3.1.7. Kits ........................................................................................................................................ 27
         3.1.8. Primers ................................................................................................................................... 27
         3.1.9. Culture media ...................................................................................................................... 29
         3.1.10. Culture media .................................................................................................................... 29
   3.2. Methods .......................................................................................................................................... 30
         3.2.1. Preparation and maintenance of Human Lung Tissue Explants(HLTEs) ...................... 30
         3.2.2. IA treatments ....................................................................................................................... 30
         3.2.3. Infections ............................................................................................................................ 31
         3.2.4. Focus forming assay .......................................................................................................... 31
         3.2.5. LDH release assay ............................................................................................................. 31
3.2.6. ELISA .................................................................................................................. 32
3.2.7. RNA extraction ................................................................................................. 32
3.2.8. RNA quality control ......................................................................................... 32
3.2.9. Complementary DNA (cDNA) generation ......................................................... 32
3.2.10. Gene expression analysis using real-time PCR .................................................. 33
3.2.11. Metabolite extraction ....................................................................................... 33
3.2.12. Targeted metabolomics using Biocrates p180 kits ............................................ 33
3.2.13. Small non-coding RNA (sncRNA) sequencing .................................................. 33
3.2.14. sncRNA annotation ........................................................................................ 34
3.2.15. Bulk RNA sequencing ...................................................................................... 34
3.2.16. Transcript normalization and differential expression analysis ......................... 34
3.2.17. Preparation of single-cell suspension ............................................................... 35
3.2.18. Live/Dead and Apoptosis detection .................................................................. 35
3.2.19. Single-cell RNA-Seq ....................................................................................... 35
3.2.20. Single-cell RNA data analysis ......................................................................... 35
3.2.21. Gene Set Enrichment Analysis (GSEA) ............................................................. 36
3.2.22. Statistics .......................................................................................................... 36

4. Results ..................................................................................................................... 37

4.1. Influenza A virus infection establishment on HLTE model ...................................... 37
4.1.1. LDH Cytotoxicity Colorimetric Assay ................................................................. 37
4.1.2. Focus forming assay for IAV titer ..................................................................... 37
4.1.3. RNA quality assessment .................................................................................. 38
4.1.4. Quantification of viral and cytokine expression by RT-qPCR ............................... 38
4.2. Establishment of viral and bacterial infection ......................................................... 39
4.2.1. Tissue exclusion criteria .................................................................................... 39
4.2.2. Immune response to infections ....................................................................... 40
4.3. Evaluation of candidate biomarkers ..................................................................... 41
4.3.1. ACOD1 and A20 ............................................................................................... 41
4.3.2. Orosomucoid 2 ............................................................................................... 41
4.3.3. Prokineticin-2 ................................................................................................. 42
4.4. Metabolic changes in infected HLTE .................................................................... 43
4.4.1. Metabolite detection ....................................................................................... 43
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.4.2.</td>
<td>Metabolite-based relationships among the infections</td>
<td>44</td>
</tr>
<tr>
<td>4.4.3.</td>
<td>Metabolites expression profile</td>
<td>44</td>
</tr>
<tr>
<td>4.4.4.</td>
<td>Differentially regulated metabolites</td>
<td>46</td>
</tr>
<tr>
<td>4.5.</td>
<td>Differential reprogramming of sncRNA during infection</td>
<td>47</td>
</tr>
<tr>
<td>4.5.1.</td>
<td>sncRNA abundance and species distribution</td>
<td>47</td>
</tr>
<tr>
<td>4.5.2.</td>
<td>Principle component analysis</td>
<td>47</td>
</tr>
<tr>
<td>4.5.3.</td>
<td>Differentially expressed sncRNA</td>
<td>48</td>
</tr>
<tr>
<td>4.5.4.</td>
<td>Hierarchical clustering of sncRNA</td>
<td>49</td>
</tr>
<tr>
<td>4.5.5.</td>
<td>Global differences in expression of sncRNAs</td>
<td>51</td>
</tr>
<tr>
<td>4.5.6.</td>
<td>sncRNA expression changes in <em>P. aeruginosa</em> infection</td>
<td>52</td>
</tr>
<tr>
<td>4.5.7.</td>
<td>Functional annotation</td>
<td>54</td>
</tr>
<tr>
<td>4.5.8.</td>
<td>The transcriptomic response to infections</td>
<td>55</td>
</tr>
<tr>
<td>4.5.9.</td>
<td>Read alignment and quantification of expression levels</td>
<td>55</td>
</tr>
<tr>
<td>4.5.10.</td>
<td>Identification of differentially expressed genes (DEGs)</td>
<td>56</td>
</tr>
<tr>
<td>4.5.11.</td>
<td>Hierarchical clustering of samples based on DEGs</td>
<td>56</td>
</tr>
<tr>
<td>4.5.12.</td>
<td>Gene set enrichment analysis (GSEA)</td>
<td>58</td>
</tr>
<tr>
<td>4.5.13.</td>
<td>DEGs in IAV infection</td>
<td>59</td>
</tr>
<tr>
<td>4.5.14.</td>
<td>DEGs in BCG infection</td>
<td>60</td>
</tr>
<tr>
<td>4.5.15.</td>
<td>DEGs in <em>P. aeruginosa</em> infection</td>
<td>61</td>
</tr>
<tr>
<td>4.5.16.</td>
<td>Antiviral immune responses</td>
<td>62</td>
</tr>
<tr>
<td>4.5.17.</td>
<td>Regulatory cytokines</td>
<td>63</td>
</tr>
<tr>
<td>4.7.1.</td>
<td>Single cell transcriptomics</td>
<td>64</td>
</tr>
<tr>
<td>4.7.2.</td>
<td>Live/dead staining</td>
<td>64</td>
</tr>
<tr>
<td>4.7.3.</td>
<td>Quality control</td>
<td>65</td>
</tr>
<tr>
<td>4.7.4.</td>
<td>Clustering and visualization</td>
<td>66</td>
</tr>
<tr>
<td>4.7.5.</td>
<td>Differentially expressed genes</td>
<td>69</td>
</tr>
<tr>
<td>4.7.6.</td>
<td>Immune-related DEGs</td>
<td>70</td>
</tr>
<tr>
<td>4.7.7.</td>
<td>Gene set enrichment analysis</td>
<td>71</td>
</tr>
<tr>
<td>4.7.8.</td>
<td>Viral transcripts content in cells</td>
<td>72</td>
</tr>
<tr>
<td>4.8.</td>
<td>Itaconic acid (IA) as an anti-inflammatory drug</td>
<td>73</td>
</tr>
<tr>
<td>5.</td>
<td>Discussion</td>
<td>74</td>
</tr>
<tr>
<td>5.1.</td>
<td>Establishment of virus infection</td>
<td>74</td>
</tr>
<tr>
<td>Section</td>
<td>Title</td>
<td>Page</td>
</tr>
<tr>
<td>---------</td>
<td>------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>5.2</td>
<td>Establishment of viral and bacterial infection</td>
<td>74</td>
</tr>
<tr>
<td>5.3</td>
<td>Evaluation of candidate biomarkers</td>
<td>75</td>
</tr>
<tr>
<td>5.4</td>
<td>Differential metabolic reprogramming</td>
<td>76</td>
</tr>
<tr>
<td>5.5</td>
<td>Differential reprogramming of sacRNA</td>
<td>76</td>
</tr>
<tr>
<td>5.6</td>
<td>Transcriptomic response to infections</td>
<td>77</td>
</tr>
<tr>
<td>5.7</td>
<td>Transcriptomic profiling at single cell level</td>
<td>79</td>
</tr>
<tr>
<td>5.8</td>
<td>Itaconate as an anti-inflammatory drug</td>
<td>80</td>
</tr>
<tr>
<td>6</td>
<td>Conclusion</td>
<td>81</td>
</tr>
<tr>
<td></td>
<td>References</td>
<td>82</td>
</tr>
<tr>
<td></td>
<td>Abbreviation</td>
<td>93</td>
</tr>
<tr>
<td></td>
<td>List of figures</td>
<td>95</td>
</tr>
<tr>
<td></td>
<td>List of tables</td>
<td>96</td>
</tr>
</tbody>
</table>
Abstract

The lung is a frequent target of infections caused by various pathogens. The first 24 h of infection represent a critical time window in interactions between pathogen and host tissue as they include recognition of pathogens by cellular sensors and, in many cases, the first replication cycles of the pathogen. However, it is impossible to study these early events in lung from humans in vivo due to the lack of accessible tissue during early time points in infection. Even though there are well-established animal and cell culture models to study human lower respiratory tract infections, they may also not fully represent the events seen in intact human lung. To overcome these drawbacks, we have established human lung organotypic culture to evaluate the early transcriptional events following influenza A virus (IAV), Mycobacterium bovis Bacille Calmette-Guerin (BCG), or Pseudomonas aeruginosa infection.

We aimed to develop the human lung tissue explant (HLTE) system for functional infection studies, i.e. evaluation of candidate biomarkers and effects of immunomodulatory agents, and characterize global transcriptomic responses (comprising small noncoding RNA and mRNA) in this model during the first 24 h of infection. To pinpoint expression changes to specific cell types and to assess expression differences between IAV-infected and uninfected cells, we profiled IAV-infected lungs at the single cell.

We developed the model by culturing explanted lung tissue pieces (approx. 30 mg, in triplicates) from human donors (n=23) with a clinical indication for lung transplantation. HLTEs were infected overnight with IAV (H1N1pdm09/Giessen), BCG, or P. aeruginosa. We assessed cell viability, expression of selected mRNA, protein biomarkers and replication of pathogens to validate the model. Using this model, we evaluated the potential of ACOD1, PROK2 and ORM2 as candidate biomarkers, and itaconic acid (IA) as immunomodulatory agent. We also conducted targeted metabolomics to evaluate this model for metabolite biomarker discovery. To study the transcriptomic response to infections, we measured changes in global RNA expression by RNAseq. The 10x Genomics platform was used for single cell RNAseq.

Only a low degree of cell loss (10-15%) was seen during culture for 48 h. IAV infection led to viral RNA replication and increase in titers in supernatants. HLTEs infected with IAV, BCG, or P. aeruginosa induced the expression of expected interferon and cytokine responses as measured by RT-qPCR and ELISA, and cytokine responses differed significantly between viral and bacterial infections. ACOD1 levels were elevated in HLTEs infected with P. aeruginosa. ORM2 remained unchanged across all conditions, however PROK2 gene expression and protein in supernatant was significantly higher in both bacterial infections than in IAV or uninfected controls. Targeted metabolomics profiling of lung tissues showed a resemblance with human serum and plasma. Regulation of few metabolites (classes: biogenic amines and amino acids) was observed in response to the bacterial infections only. Treatment of IAV infected HLTEs with IA led to significant reduction of the inflammatory marker CXCL10.

Among the three major RNA classes, expression changes were greatest in mRNA, followed by long noncoding RNA, whereas changes in small noncoding RNA were modest. Differentially expressed genes in IAV infection were predominantly upregulated, with a strong presence of mRNA related to type I and II interferon signaling and virus restriction. In comparison to viral infection, both bacterial infections led to higher number of differentially expressed genes that are involved in immune responses, cell cycle
regulation, and tissue homeostasis. Gene set enrichment analysis furthermore revealed activation of cell signaling pathways in IAV infection, whereas in bacterial infections regulation of metabolic pathways was observed, e.g. downregulation of glycolysis, fatty acid metabolism, drug metabolism, and up-regulation of pantothenate and CoA biosynthesis. At single cell resolution, viral RNAs were detected predominantly in the known host cell types, macrophages and type II pneumocytes, but also inflammatory and stromal cells. Host transcriptional responses were most pronounced in type II pneumocytes and macrophages.

Conclusion: The results suggest that this model is a suitable alternative to replace animal models for the preclinical evaluation of (1) host responses to pulmonary infections, (2) evaluation and discovery of new biomarkers, and (3) novel therapeutic interventions. It provides an attractive model to study host transcriptomic responses during the early phase of viral and bacterial infections, both at the tissue and the single-cell level.
Zusammenfassung


Wir zielten darauf ab, das human lung tissue explant (HLTE) System für funktionelle Infektionsstudien zu entwickeln. Hierzu zählen beispielsweise die Bewertung potentieller Biomarker und Effekte immunmodulierender Substanzen, sowie die Charakterisierung globaler transkriptomischer Antworten (bestehend aus nichtkodierender RNA und mRNA) in diesem Modell während der ersten 24 Stunden einer Infektion. Um Expressionsänderungen bestimmten Zelltypen zuzuordnen und Expressionsunterschiede zwischen IAV-infizierten und nicht-infizierten Zellen zu bestimmen, haben wir IAV-infizierte Lungen auf der Ebene einzelner Zellen charakterisiert.

Wir haben dieses Modell entwickelt, indem wir explantiertes Lungengewebe (ca. 30 mg, als Triplikate) humaner Spender (n=23) mit klinischer Indikation für eine Lungentransplantation kultivierten. Die HLTEs wurden über Nacht mit IAV (H1N1pdm09/Giessen), BCG oder P. aeruginosa infiziert. Wir bewerteten hierbei die Zellviabilität, die Expression ausgewählter mRNA sowie Protein-Biomarker und die Vermehrung der Erreger, um unser Modell zu validieren. Mit diesem Modell beurteilten wir anschließend das Potential von ACOD1, PROK2 und ORM2 als mögliche Biomarker und Itaconsäure (IA) als immunmodulierende Substanz. Wir haben auch gezielte Metabolomik durchgeführt, um dieses Modell für die Entdeckung von Stoffwechsel-Biomarkern zu evaluieren. Um die transkriptomische Reaktion auf Infektionen zu untersuchen, haben wir mittels RNA-Sequenzierung Änderungen in der globalen RNA-Expression erfasst. Die 10x Genomics-Plattform wurde für Einzelzell-RNAseq verwendet.

Reaktion auf die bakteriellen Infektionen beobachtet. Die Behandlung von mit IAV infizierten HLTEs mit IA führte zu einer signifikanten Reduktion des Entzündungsmarkers **CXCL10**.


Zusammenfassend lassen die Ergebnisse darauf schließen, dass dieses Modell eine geeignete Alternative für vorklinische Untersuchungen über (1) die Reaktion des Wirts auf Atemwegsinfektionen, (2) Beurteilung und Enddeckung neuer Biomarker und (3) neuer Therapiemethoden ist und somit Tiermodelle ersetzen könnte. Es bietet ein attraktives Modell zur Untersuchung der transkriptomischen Reaktionen des Wirts während der frühen Phase viraler und bakterieller Infektionen, sowohl auf Gewebs- als auch auf Einzelzellenebene.
1. Introduction

1.1. Lower respiratory tract infections

Acute respiratory tract infections (ARIs) are the most common illnesses reported in primary health care worldwide [1, 2]. ARIs can be classified into upper and lower respiratory tract infections (URI, LTRI) [3]. LTRIs are common in the community around the world, and most of them are associated with the seasons. They include both acute bronchitis and pneumonia [4]. Pneumonia, by definition, is the acute inflammation of one or both lung’s parenchyma due to respiratory infections by bacteria, viruses, or rarely fungi [5, 6]. These pathogens invade the lung parenchyma and colonize there, leading to tissue damage and intra-alveolar exudate accumulation [6]. In the event that pathogens defeat the host immune system, they enter the lower respiratory tract and cause infection. Their systemic spread may lead to sepsis; indeed, pneumonia is the most known cause of sepsis in humans [7]. According to a World Health Organization (WHO) report, pneumonia is the most significant cause of death in children below the age of 5. Just in 2015, a total of 920,136 children under the age of 5 died due to Pneumonia, accounting for 16% of all deaths in this age group [8]. Preventive measures such as vaccination against pneumococci and influenza has helped in an overall reduction in the disease prevalence rate among children and adults [9].

A clinical diagnosis of pneumonia is based on characteristics of signs and symptoms such as dyspnea, cough, fever, and lung sounds. A few laboratory-based diagnostic biomarkers are available, such as procalcitonin and C-reactive protein, which can add diagnostic information to the clinical presentation [10]. Microbiology based isolation and identification of the causative agent is mostly done by infectious agent isolation by culture of sputum, lung (bronchoalveolar lavage) or blood samples [11]. New approaches for more rapid diagnosis such as lung imaging by ultrasound and microbiological detection by molecular methods are also promising [10]. The choice of treatment for patients depends on the identification of the specific causative pathogen and the clinical condition of the patient.

1.1.1. Early events in the lungs following respiratory infections

Due to the complex etiology of pneumonia, early events in the lung following infection by various pathogens are poorly understood. Major processes by the host in order to effectively prevent the onset of pneumonia are immune resistance against pathogens and tissue resilience [12]. Resistance against pathogens includes innate and adaptive immune responses to eradicate the invading pathogens, whereas tissue resilience refers to the prevention of tissue damage resulting from microbe or/and immune activities [12]. Transcriptome studies have been conducted in different setups to elucidate the early response to infections. The initial responses to infection are from resident lung cells, which include production and release of several cytokines, growth factors, anti-microbial substances, and many other factors, depending on the type of pathogen and host immunity [13-16]. Innate immunity, the first line of defense, comprises several different types of cells and chemokines that actively function to eradicate the microbes and maintain tissue integrity.

Alveolar macrophages (AMs) play the most vital role in immune resistance and tissue resilience (Figure 1-1) [17]. They ingest and phagocytize microbes to restrict and control the infection [18]. Activated AMs...
release cytokines and other mediators to alarm neighboring cells and recruit the other immune cells [19]. Under normal conditions, AMs suppress the inflammation through various channels.

**Airway Epithelial cells (AECs)** are comprised of different types of cells which collectively form the lung epithelium [20]. Following the interaction with a pathogen, AECs undergo dramatic transcriptional remodeling to play an immunomodulatory role [20]. AECs synthesize the surfactant proteins (SP) such as A (SPA) and D (SPD), which have potent antibacterial properties [21]. In the immediate response to infection, AECs, along with AMs, recruit the other immune cells, most commonly neutrophils, via released cytokines at the site of infection to kill invading pathogens [6].

Pathogens and an excess of immune activities at tissue level can damage lung tissue, and therefore, cells should act in order to limit injury. Failure to maintain tissue integrity develops pneumonia to severe forms of disease such as sepsis and acute respiratory distress syndrome (ARDS) [22, 23]. Tissue resilience of lungs following infection is maintained by controlling the immune response and number and activity of cells in the lung parenchyma. Cells control inflammation by releasing anti-inflammatory cytokines such as IL-10, IL-1 receptor antagonist (IL-1RA), and transforming growth factor (TGF)-β [24-27].

![Figure 1-1: Alveolar macrophages (AM) are located on the surface of the lower respiratory tract. Following bacterial or viral infection, AMs transit from their anti-inflammatory state to a center of immune activity. This protects host cells by a pro-inflammatory response against infection (i.e. immune resistance) and against tissue injury (i.e. tissue resilience). Figure modified from Quinton, L. et al. 2018.](image-url)
viability is maintained during pneumonia to protect tissue integrity and epithelial barrier. Signal transducer and activator of transcription 3 (STAT3) has been reported as the most prominent factor in controlling cytotoxicity and acute lung injury in response to infections caused by viruses, bacteria, lipopolysaccharides (LPS), and naphthalene [28, 29]. Also, macrophages support the growth of AECs [30].

1.1.1.1. Bacterial infection

Pulmonary bacterial pathogens are the most significant cause of pneumonia which have been originally identified almost entirely through microbial culture [11, 31]. *Streptococcus pneumoniae* and *Haemophilus influenzae* are the most common pathogens associated with pneumonia [11]. Although, different other bacterial pathogens are also associated with the epidemiology of pneumonia, such as *Staphylococcus aureus* with community-acquired pneumonia and *Pseudomonas aeruginosa* with ventilator-associated pneumonia [11, 32, 33]. *P. aeruginosa* infection of the lungs can result in two pathological conditions, 1) acute pneumonia in hospitalized patients, and 2) chronic inflammation in people suffering from the genetic disorder cystic fibrosis [34]. Another interesting factor in the pathogenesis of many *P. aeruginosa* infections is its ability to form biofilms, as this makes conditions more favorable for bacterial persistence by evading host defenses and antimicrobial therapy [35, 36]. The pathology of *P. aeruginosa* pneumonia is complex and depends on several virulence factors such as surface components, secretion systems, quorum sensing, iron scavenging, and many others as well as factors specific to the host [37]. Many virulence factors directly interfere with host cell activities. For instance, 3-oxo-C12-HSL modulates the activated form of myeloid cells, elastase, and alkaline protease degrades opsonins. ExoA and pyocyanin lead to cell apoptosis and impair macrophages for their ingestion activity of apoptotic cells [38-41]. Following bacterial infection, lung cells release pro-inflammatory cytokines to recruit immune cells to the site of infection to eliminate the pathogen [42]. CpG DNA induces the production of indoleamine-2,3-dioxygenase (IDO) which catalyzes the tryptophan (TRP) to kynurenine (KYN) reaction. Additionally, *P. aeruginosa* produces the enzyme kynurenine formidase (KF) which also induces the catabolism of TRP to KYN. Higher production of KYN stimulates the production of T regulatory lymphocytes [43]. So far, very little is known about tissue damage during bacterial pneumonia and how tissue initiates and sustains the regenerative process to promote tissue repair [44].

*Mycobacterium tuberculosis* (Mtb) remains the historically deadliest pulmonary bacterial pathogen [11]. Mtb is phagocytosed mainly by macrophages, which stimulates a local inflammatory response and infiltration of different immune cells at the site of infection. This large mass of different immune cells gather there and form granulomas, the characteristic marker of TB [45]. Mtb survives and replicates inside the host cells, but many aspects of the complex host cell-Mtb interactions remain unclear [45].

1.1.1.2. Viral infections

In viral pneumonia, causative pathogens vary greatly in different studies. This difference depends on the age group of the population and methods of disease diagnosis. Viral infection etiology of the respiratory tract illnesses also has unique seasonal distribution [46, 47]. Most of the viruses causing pneumonia in humans are RNA viruses, e.g. respiratory syncytial virus, rhinovirus, and influenza A virus (IAV) [48].
Influenza A or B cause seasonal influenza outbreaks in the winter season. Pneumonia caused by influenza viruses can be either a primary viral infection in the lungs or, more commonly, it is followed by a secondary bacterial infection [47]. IAV mainly targets epithelial cells of the respiratory tract through the binding of viral antigen hemagglutinin (HA) to sialic acid residues [49]. The presence of viral RNA within infected cells is recognized by various pattern recognition receptors (PRRs) of host cells. This stimulates the secretion of pro-inflammatory cytokines, activation of type I and II IFN cascade, and releases eicosanoids. These signaling molecules activate other non-infected immune cells which eventually stimulate the IFN gene cascade known as IFN-stimulated genes (ISGs). Alveolar macrophages are the first responders to the viral infection and AECs help to clear infection via phagocytosis and production of cytokines [50]. Pro-inflammatory cytokines and eicosanoids induce local and systemic inflammation, an excess of which can be lethal [51].

1.2. Biomarkers of infection

The United States National Institutes of Health defined [52] a biomarker as

“Biological marker (biomarker): A characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention [52].”

According to this definition, a biomarker could be any measurable indicator that can distinguish between a normal biological state and the presence or the stage of disease. Host biomarkers have the novelty to be used for clinical purposes because diagnosis and progression of a disease can be assessed by specific changes in the host induced by the pathological processes [53]. Host biomarkers can be clinical biometric data (such as physical or anatomical observation by a physician) or biomolecules such as metabolites, proteins or nucleic acids [54]. Recent developments of high-throughput technologies such as transcriptomics, metabolomics and proteomics have enabled us to detect and quantify a huge range of biomolecules from a single sample and identify new candidate biomarkers [53, 55]. In the recent decade, thousands of new molecular biomarker candidates have been identified to help us to diagnose a great variety of clinical conditions [56]. However, only very few novel biomarkers are currently used in clinics despite all these identifications [57]. From the discovery of a candidate biomarker to their introduction in clinical use requires additional verification of its diagnostic, prognostic, and predictive potential and revalidation in large cohort studies [56]. Some important features of an ideal biomarker are that it should be specific, sensitive, predictive, inexpensive, rapid to detect and, above all, reliable [58].

So far, it is often not possible to differentiate between bacterial and viral diseases on the basis of clinical presentation or even with currently available diagnostic tests [59]. A study by the Centers for Disease Control and Prevention shows that a pathogen (virus or bacteria) was detected in only 38% of patients with pneumonia [9]. Physicians therefore often prescribe antibiotics not knowing the exact causative agent of the disease (empiric therapy) [60]. This has led to over-prescription of antibiotics, almost twice higher than essential, in children with ARIs in the USA [59, 60]. Host-pathogen interaction and its outcomes need to be understood in order to validate new diagnostics and develop new treatments
against infectious diseases [61]. In infections, cells immediately resist invading pathogens via innate immune responses and thus maintain tissue resilience [20].

A useful biomarker adds new information to already available data from clinical assessment. In some cases, two or more biomarkers are combined to achieve the goal of “exact diagnosis” [62]. There are only a few biomolecules that fulfill the definition or criteria of a clinical biomarker [63]. An example of a biomarker that can help to distinguish between viral or bacterial infections is serum procalcitonin, and it has been approved by the United States Food and Drug Administration for use in clinics [64]. C-Reactive protein (CRP) is widely used in the diagnosis and follow-up of bacterial infections and antibiotic use in children. Furthermore, several studies have highlighted its usefulness in the prognosis of sepsis, community-acquired pneumonia or exacerbation of chronic obstructive pulmonary disease (COPD) [63]. Adrenomedullin (ADM) in combination with clinical data is being used to assess the severity of CAP and predict the risk of sepsis [65]. IL-6 has key roles in several immune and inflammatory responses through the induction of IL-2 [66]. IL-6 has advantages over PCT and CRP, for instance that it is an immediate response to infection and a more sensitive marker of localized infection [67].

### 1.3. Models for infection research

In infection research, classical cell culture or animal models are widely used to study infection pathology and the development of new therapeutics or diagnostics. Both models have been very useful in studying several aspects of respiratory infections, but still contain many drawbacks [68]. There have been many attempts to develop surrogate models based on in vitro cell culture techniques, but the majority of these models failed to recapitulate the structural and functional complexity of living tissues and organs [69]. Classical cell culture models consisting of one cell line are a very simple form of in vitro culture. They have been very productive in studying complexity and functional analysis of the interaction between host-pathogen in basic research [68]. Co-culture with another cell type constitutes an upgrade of this system, and can, for example, broaden the array of host factors [68]. However, these cell culture models lack cell diversity, the histological relationship among the cells and complete inflammatory mediator responses (Figure 1-2). A mediated immune response to respiratory pathogens and allergens at the organ level needs these complex cell to cell interactions between the lung epithelium and the underlying microvascular endothelium [70].

Animals are commonly used for infection studies, but their clinical relevance is questionable due to several reasons. Especially in molecular biomarker research, the structures of animal proteins and nucleic acids differ greatly from those in humans. In addition, lung anatomy, inflammatory responses, and immune systems to different stimuli are also different from humans [70]. Therefore, many findings achieved in animal models used in basic research could not be translated for human use in clinics [71]. Additionally, many pathogens involved in causing human lung infections does not cause infections in popular small animal models due to host specificity [72, 73]. Animal models for several infections have had to be established in species that are not the natural hosts [68]. In addition, there are ethical aspects of animal use in research. Good scientific practices and experimental design are very useful to replace or reduce the use of animals in research [74].
To circumvent these obstacles inherent to animal models, there are currently intense efforts to implement models based on human organs or organoids. Recent advances in microsystems engineering, stem cell research, and cell culture methods have enabled us to develop systems like lung-on-a-chip and lung organoids [75, 76]. These systems have helped us to overcome some challenges we face in cell culture or animal models, but require intense work to develop and maintain [77].

Lung tissue explant models have been established and evaluated for different infection studies [78-81]. Nicholas et al. used cultured lung biopsy tissue to study the efficacy of anti-influenza drugs in this system [78]. Previously, this tissue explant model was also used to study the host-pathogen interactions between *Legionella pneumophila* and lung cells [79]. Berg et al. identified the role of tyrosine kinase 2 in IFN receptor inhibition and restoration of interleukin (IL)-1β release in bacterial infection [81] using a lung tissue explant model. Different aspects of respiratory infections by bacteria e.g. *S. pneumoniae, M. tuberculosis, H. influenzae,* and *Coxiella burnetii,* and viruses e.g. coronaviruses, adenovirus 7 and influenza have been studied using human lung *ex vivo* infection models [68]. These systems have several advantages over animal or cell culture models: (1) lung tissue is the site of infection for respiratory tract infections; (2) it does not have certain complexities such as a functioning adaptive immune system, but it is capable of mounting early innate immune responses, which is ideal for biomarker research [78]; (3) it is often removed from humans for clinical reasons and therefore constitutes a good example of “upcycling” a human tissue that would otherwise be discarded. The precision-cut lung slices (PCLS) model is a more refined organotypic model also based on human or animal lung explants [82]. It is being used in drug testing and to study inflammation in different setups [82-84].

![Human lung consists of different cell types. Specialized cells for lung function and structure contain alveolar unit cells, including Type 1 cells, Type 2 cells, fibroblasts, endothelial cells, and mast cells. To protect the lung from extrinsic stimuli, lung harbors immune cells including alveolar macrophages, dendritic cells, lymphocytes, and inflammatory cells](image-url)
1.4. Candidate biomarkers

The translational process from the discovery of a biomarker to its clinical use involves an intensive process of evaluation and validation. In this development process, candidate biomarkers are tested in other model systems to qualify for validation [57]. Our research group has identified some candidate biomarkers for further qualification.

1.4.1. Aconitate decarboxylase 1 (ACOD1)

ACOD1 (previously known as immune responsive gene 1) was identified in 1995 from the cDNA library of murine macrophages following their treatment with lipopolysaccharides (LPS) [85]. ACOD1 codes for the enzyme cis-aconitate decarboxylase, which catalyzes the decarboxylation of cis-aconitate, a tricarboxylic acid cycle intermediate, to produce itaconic acid (IA) [86]. A higher expression of ACOD1 was observed in several infection studies [86-88]. Preuße et al. from our research group also identified ACOD1 as one of the most highly expressed genes in mouse lung following IAV infection [89]. Iqbal et al. from our research group evaluated the expression of ACOD1 and its downstream mediator A20 in different bacterial and viral infections in a cell line culture model [90]. It was found that expression of both ACOD1 and A20 was significantly increased by P. aeruginosa infection, but remained unchanged in influenza and S. aureus infections. Additionally, a synergetic effect in the upregulation of expression of these molecules was observed following influenza and P. aeruginosa co-infection. These findings were produced on differentiated THP1 cells.

1.4.2. Orosomucoid 2 (ORM2)

ORM2 (most commonly known as Alpha-1-acid glycoprotein 2) is an acute-phase protein that is involved in immunomodulatory functions, though its specific function is not known yet [91]. ORM1 is a paralog of ORM2 and, together, the two molecules modulate immune responses to stress. In different pathological conditions, such as acute infection, chronic inflammation and autoimmune disorders the levels of ORM2 increase two to five fold in human serum [92]. In one study, the role of ORMs against Gram-negative infections was identified and it was postulated that they contribute to nonspecific resistance to infection [93]. Paquette et al. identified elevated expression of the ORM2 gene in mice two days after infection with the influenza virus [94]. In a study conducted on mice by M. Preuße from our research group, ORM2 was one of the most highly upregulated genes following viral infection.

1.4.3. Prokineticin 2 (PROK2)

PROK2 is a cysteine-rich protein that possesses diverse biological activities including inflammatory modulation, angiogenesis, pain perception, tissue development and reproductive functions [95-98]. PROK2 acts as a cytokine and is highly expressed in inflamed tissue in order to modulate activity of infiltrating cells (here, neutrophils) [99]. It stimulates lipopolysaccharide-induced production of pro-inflammatory cytokines, i.e. IL-1 and IL-12, and reduces levels of the anti-inflammatory cytokine IL-10 [100]. Acute appendicitis patients’ blood mRNA profile showed elevated levels of RROK2 due to the presence of biofilm-forming bacteria [101]. In another study, higher expression of RROK2 gene was identified in mice treated with monosodium urate crystal in the murine air pouch model [102], which
shares important inflammatory features with sepsis in humans. Bacterial products stimulate PROK2 production in order to promote IL-1 signaling [103]. The importance of PROK2 as a biomarker for inflammation has been discussed previously [99, 103].

1.5. Itaconic acid

Itaconic acid (IA) is a dicarboxylic acid, synthesized by decarboxylation of cis-aconitate, an intermediate of the Krebs cycle, by the enzyme cis-aconitate decarboxylase 1 (ACOD1) mentioned above [86]. It is the most abundant metabolite in LPS stimulated human macrophages, highlighting its importance in innate immunity [104]. Its role as antibacterial and anti-inflammatory molecule has been studied previously [105, 106] [104]. This dual functionality shows its uniqueness and suggests possible use as treatment in infections. IA is responsible for inhibiting the glyoxylate cycle in several pathogenic bacteria. The glyoxylate cycle is important for the survival of bacteria particularly in eukaryotic host cells and for the infection pathogenesis. Some of these bacteria have the ability to degrade IA, thus enhancing their pathogenicity [107]. Among these bacteria, *P. aeruginosa* is an important example. It has a cluster of 6 genes, acting in the operon model which metabolize IA to pyruvate and acetyl-CoA [107].

IA has also been shown to directly inhibit succinate dehydrogenase (SDH) [108]. IA also alkylates cysteine residues of the KEAP1 protein, leading to activation of NRF2. Activated NRF2 then increases the expression of downstream genes with anti-oxidant and anti-inflammatory properties (Figure 1-3) [104]. Furthermore, it has been shown that NRF2 inhibits the stimulator of interferon genes (STING) by decreasing its mRNA stability. STING is a key component of a prominent pathway for inducing type I interferon production during viral infection, and IA repressed activation of STING and type I IFN production in immune cells [109].

Our research group has also observed that IA and its methylated derivatives (such as dimethyl-IA, abbreviated DI) exert major anti-inflammatory effects in influenza infection, notably by dampening IFN responses and other aspects of systemic inflammation. IA is a polar compound. It has therefore been chemically modified to make it less polar and facilitate entry into cells. There are two well-known forms: 1) a methylated derivative “dimethyl-IA” and 2) 4-octyl IA, which contains an 8-carbon chain at the C4-position. 4-octyl IA has an advantage over dimethyl-IA in that it is more stable and less toxic.
1.6. Metabolomics in biomarker research

Metabolites, by definition, are the molecules involved in chemical reactions in an organism. These chemical reactions, collectively known as metabolism, can be categorized as catabolism and anabolism. Metabolism drives several important cellular functions, mainly energy production and storage in body or cells, cell signaling, and apoptosis [110]. Foreign compounds, xenobiotics, such as drugs, environmental pollutants, cosmetics, food particles, and microorganisms also have a major impact on metabolomes [111]. The human metabolome consists of biomolecules in a cell and xenobiotics such as components of diet, drugs, toxins, and pollutants [112]. Immuno-metabolism has emerged as a key interest for several researchers. Several recent studies have described the role of metabolites particularly succinate and IA in innate immune cells such as macrophages [113, 114]. For a long time, it was believed that metabolites have no direct role in the regulation of the immune system, but, for instance, they act as danger signals for the immune cells to activate and enhance protection against pathogens [115]. Metabolomics is the large-scale study of metabolites in biofluids, cells, and tissues, which is now an extensively used tool for biomarker discovery and understanding the immune mechanisms [110].
Advances in informatics and analytical technologies, such as mass spectrometry-based techniques, have made metabolomics more effective in clinical research. Differential expression and modifications to these metabolites in pathological conditions are now intensively profiled, and this information is used to discover or evaluate diagnostic biomarkers [116]. There are two main mass spectrometry-based techniques applied for identification and quantification; (1) untargeted (global) and (2) targeted metabolomics [110]. One targeted quantitative and quality-controlled metabolomics method, the AbsoluteIDQ® p180 Kit (Biocrates Life Sciences AG, Austria) has been applied in several biomarker discoveries and evaluation studies [117-119]. The AbsoluteIDQ® p180 Kit can detect and quantify a broad range of metabolites involved in diverse physiological processes such as cell cycle control, mitochondrial function, and inflammation [120]. It can quantify up to 188 metabolites consisting of amino acids, biogenic amines, acylcarnitines, lysophosphatidylcholines, phosphatidylcholines, and sphingolipids in biological samples.

A recent study has identified prognostic metabolic signatures for the development of subclinical disease prior to active tuberculosis [121]. Several research studies have profiled metabolites in different biological samples such as plasma, cerebrospinal fluid (CSF), bronchoalveolar lavage fluid, saliva, or urine depending on disease etiology [118, 121-124]. Cell culture and animal models have also been also implied for the discovery and validation of novel biomarkers [125, 126]. M. Kuhn et al. from our research group has elaborated on the prognostic and diagnostic importance of disease-associated metabolite signatures in varicella zoster virus (VZV) reactivation in CNS [119]. The tryptophan (Trp)/kynurenine (Kyn) pathway is a major metabolic pathway involved in systemic inflammation and has been studied extensively [127]. Influenza infection induces the expression of IDO enzyme in the lung parenchyma and produces higher levels of Kyn, leading to immune dampening [128]. Biosynthesis of Kyn in macrophages increased in response to different viral and bacterial infections [129].

Besides biomarker research, metabolomics in combination with other “omics” are used to understand the role of metabolites in different biological functions [110].

### 1.7. Importance of small noncoding RNAs as biomarkers

A major part of the human genome transcribed into RNA is not translated into proteins. This “untranslated RNA” plays its role as housekeeping RNA (such as rRNA, tRNA) and regulatory RNA (such as microRNA (miRNA), Piwi-interacting RNA (piRNA), and long non-coding RNA (lncRNA)) [130]. Among regulatory RNA, miRNAs are a class of small (21-22 nt), non-coding RNAs that mainly regulate gene expression at the post-transcriptional level [131]. miRNAs bind to complementary sites at the 3’ untranslated region (UTR) of targeted mRNA leading to either mRNA degradation or protein translation suppression [132, 133]. A bioinformatics based study to predict gene regulation by miRNAs indicates that >60% of human mRNAs have been under selective pressure to maintain pairing with miRNAs [134]. miRNAs have structural properties that are required for diagnostic quantification: stability, resistance to high temperatures and freeze-thaw cycles, and are widely present in body fluids and tissues [135, 136].

Reprogramming of miRNAs in blood samples of sepsis patients has been reported in several studies [137-140]. Various intracellular miRNAs regulate the TLR/NF-κB mediated inflammatory response in sepsis
Several miRNAs have been identified as biomarkers in sepsis, although their association with the pathophysiology of sepsis is still controversial due to its complex etiology [141]. The biomarker potential of miRNA in infectious diseases has been reported in several studies [142, 143]. L. Araujo from our research group has shown the diagnostic importance of miRNAs in blood from patients with Mtb infection [144]. Recently, the potential of a panel of miRNAs has been evaluated to accurately distinguish between viral and bacterial respiratory tract infections [145]. In addition, the potential of miRNAs as therapeutic targets to treat immunological conditions, particularly mucosal inflammation has been evaluated [146]. In another study, the ability of microRNA-based therapy to promote lung cell regeneration and improve host recovery following bacterial pneumonia has been studied [44].

In recent years, other classes of small non-coding RNA have also been the focus of many studies to evaluate their diagnostic potential and functional roles [130]. piRNAs regulate gene expression via guiding PIWI proteins to cleave target RNA [147]. piRNAs have been reported to play a role in antiviral defenses [147]. Small nucleolar RNAs (snoRNAs) are another major class of sncRNA [148]. They are 60–300 nt in length and interact with their target proteins to form small nucleolar RNPs (snoRNPs). This complex guides RNAs during the post-transcriptional synthesis of 2-O-methylation or pseudouridylation [149]. Their potential as a clinical diagnostic marker has been proposed in several studies; e.g. two snoRNAs were recently identified as diagnostic biomarkers for lung cancer [150]. L. Araujo from our research group presented SNORD104 as an accurate biomarker for individuals with active TB and likely also those with latent infection at high risk of progressing to active disease [144].

1.8. Transcriptomics

Genome-wide transcriptional studies have become one of the most utilized –omics tools to study human diseases at the molecular level. They have emerged as a powerful tool to identify biomolecules for diagnostic, prognostic and therapeutic applicability in various pathological conditions [151]. The entire set of RNA molecules transcribed in a cell (or set of cells) at any one time can be defined as its transcriptome. RNA sequencing (RNA-seq) allows capturing the immediate reprogramming to a stimulus at the transcriptional level of the cell [152]. With the advancement in molecular biotechnology, RNA-based measurements have shown their potential for broad clinical application in disease diagnosis and prognosis, and selection of proper treatment. Quantitative reverse transcription PCR (RT-qPCR) assays have been established and are in use in the clinic, for instance, for viral detection and typing [152]. Rapid detection of multidrug resistance Mtb by bacterial DNA qPCR is commonly implemented in diagnostic labs [153]. Though these examples are from pathogen RNA detection, however, it shows the successful implementation of RNA molecules in diagnostics.

RNA-seq has advantages over other RNA transcripts quantification methods as it quantifies without pre-defining the RNA targets of interest and provides counts of rare and novel RNA transcripts. Besides the counts, RNA-seq can detect underlying genomic alterations even at single-nucleotide resolution in any of the expressed genes. In infectious diseases, host responses in the form of mRNA signatures can be a helpful tool to monitor specific infections [152]. Several studies have established gene set panels for infectious disease diagnosis, however, sometimes these gene sets are too large to translate into a useful clinical tool [154]. An 11 genes panel was evaluated to distinguish between noninfectious systemic
inflammation and sepsis. This panel has an area under the receiver operating characteristic (ROC) curve (AUC) above 0.9 [155]. In another study, Scicluna et al. used the \textit{FAIM3:PLAC8} mRNA ratio as a candidate biomarker to assist in the rapid diagnosis of CAP to no-CAP patients, giving an AUC of 0.84 [156]. qPCR based assays for gene set quantification to distinguish between bacterial and viral respiratory tract infections have been evaluated in several studies [157-159]. Andres-Terre et al. identified a common host gene expression signature between respiratory tract viral infections versus bacterial infection. In addition, they identified an influenza-specific host response signature at the mRNA level to diagnose and monitor pneumonia caused by influenza. This gene set panel was refined to study the response to the influenza vaccine [160]. These studies consist of large numbers of genes in their panel, used different models, lacked independent validation, or had low accuracy and precision, thus making them difficult to use in clinics [154]. The other challenges for the use of transcriptomics in clinics are difficulty in bioinformatic data analysis, assay development for clinical conditions, and reproducibility of the assay in different environments [152].

1.9. Single-cell RNA sequencing

Investigating cell heterogeneity in complex organisms and tissues has been historically difficult, though there has been progress in this field through antibody-based methods such as immunohistochemistry and fluorescence-activated cell sorting (FACS) [161]. However, these approaches require pre-selection of proteins of interest and are thus inherently biased and with limited output, although these are relatively cost-effective ways to assay hundreds of thousands of cells quickly. Single-cell transcriptome profiling (scRNA-seq) has enabled us to interrogate the transcriptome at the resolution of individual cells to study cellular heterogeneity, development, and activation states in healthy and diseased tissues [162]. It offers unprecedented options to study the interplay between different cell types in a tissue and response of cells to external stimuli such as infection by viruses or bacteria [163]. Additionally, single-cell approaches are being used to study the outcome of infection at the single cell level, drug or antibiotic resistance and targeted therapy aimed exclusively at infected cells [164-166]. This may help to identify new mechanisms involved in cellular processes to be later used in biotech or medical importance [167].

scRNA-seq has been implemented to investigate different cellular heterogeneity and mechanisms in different human tissues including lungs [168, 169]. AECs are the main target of IAV in the lungs [49]. There are also reports on IAV infecting other resident lung cells such as AMs, endothelial cells, natural killer (NK) cells, and dendritic cells (DCs) [170, 171]. The complexity of infection may also be related to (1) extremely wide cell-to-cell variation in the production of viral particles in infected cells, as well as (2) the heterogeneity of host-response states [172-174]. A few things are needed to be understood to resolve the complexity of IAV infection in the lung. So far, the degree and nature of intracellular IAV infection in different cell types in lungs has not been systematically determined. How do infected cells and bystander cell types in a tissue response to infection? Another complexity is derived from mediators or chemokines released from infected cells to bystanders [171].

Steuerman et al. have recently shown that murine lung cells (irrespective of the presence of viral transcripts) respond strongly to infection. In addition, they have identified novel markers differentially
expressed in influenza-infected cells compared with bystander cells (not infected but exposed to infection) [175].
2. **Aims of the project**

Lung inflammation is a complex pathophysiological entity with persistently high mortality and morbidity. Clinical diagnosis and prognosis are not sufficient for the accurate diagnosis of the etiology and severity of the disease. To date, there is a lack of sensitive and specific host biomarkers to reliably assess the severity of acute bacterial and viral infections at an early stage. The description of pneumonia as a pattern of genomic events in response to infection will enhance our understanding of sepsis and identify potential diagnostic and prognostic biomarkers, and therapeutic targets. In the context of this hypothesis, the overall aim of this work is to characterize the patterns of the early events in lung tissue following bacterial or viral infection.

The specific aims include:

- To establish the human lung tissue explant culture for functional infection studies.
- Evaluation of candidate biomarkers already under study in our research group in this model.
- Use of targeted metabolomics on lung tissue infection to elaborate its possible use in infection studies.
- To identify the genomic response in lung to infection by comparing samples derived from infection with IAV, *P. aeruginosa* and *M. bovis* (BCG strain).
- Understanding the complexity of IAV infection at the level of single lung cells.
3. Materials and methods

3.1. Materials

3.1.1. Lung tissue donors

Lungs explanted from people suffering from chronic lung diseases (cystic fibrosis, idiopathic pulmonary fibrosis, pulmonary arterial hypertension, emphysema) were examined by the Department of Pathology, Medical School Hannover. 10-20 grams lung parenchymal tissue was dissected and kept in AQIX® RS-I at 4°C while transferring to TWINCORE for research processing. A total of 23 samples were used in this study (Table 3.1).

**Table 3.1: List of lung tissue donors**

<table>
<thead>
<tr>
<th>Serial ID</th>
<th>Disease</th>
<th>Age</th>
<th>Gender</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>Emphysema</td>
<td>59</td>
<td>W</td>
</tr>
<tr>
<td>A2</td>
<td>Emphysema</td>
<td>50</td>
<td>W</td>
</tr>
<tr>
<td>A3</td>
<td>Fibrosis</td>
<td>46</td>
<td>M</td>
</tr>
<tr>
<td>A4</td>
<td>Fibrosis</td>
<td>61</td>
<td>W</td>
</tr>
<tr>
<td>A5</td>
<td>Cystic fibrosis</td>
<td>39</td>
<td>M</td>
</tr>
<tr>
<td>A6</td>
<td>Pulmonary arterial hypertension</td>
<td>26</td>
<td>M</td>
</tr>
<tr>
<td>A7</td>
<td>Lung retransplantation</td>
<td>59</td>
<td>M</td>
</tr>
<tr>
<td>A8</td>
<td>Pulmonary arterial hypertension</td>
<td>17</td>
<td>W</td>
</tr>
<tr>
<td>A9</td>
<td>Emphysema</td>
<td>59</td>
<td>W</td>
</tr>
<tr>
<td>A10</td>
<td>Cystic fibrosis</td>
<td>28</td>
<td>W</td>
</tr>
<tr>
<td>A11</td>
<td>Pulmonary arterial hypertension</td>
<td>32</td>
<td>W</td>
</tr>
<tr>
<td>A12</td>
<td>Emphysema</td>
<td>53</td>
<td>M</td>
</tr>
<tr>
<td>A13</td>
<td>Cystic fibrosis</td>
<td>25</td>
<td>M</td>
</tr>
<tr>
<td>A14</td>
<td>Emphysema</td>
<td>56</td>
<td>W</td>
</tr>
<tr>
<td>A15</td>
<td>Emphysema</td>
<td>54</td>
<td>M</td>
</tr>
<tr>
<td>A16</td>
<td>Emphysema</td>
<td>57</td>
<td>W</td>
</tr>
<tr>
<td>A17</td>
<td>Emphysema</td>
<td>64</td>
<td>M</td>
</tr>
<tr>
<td>A18</td>
<td>Emphysema</td>
<td>54</td>
<td>W</td>
</tr>
<tr>
<td>A19</td>
<td>Emphysema</td>
<td>53</td>
<td>W</td>
</tr>
<tr>
<td>A20</td>
<td>Emphysema</td>
<td>51</td>
<td>M</td>
</tr>
<tr>
<td>A21</td>
<td>Emphysema</td>
<td>48</td>
<td>W</td>
</tr>
<tr>
<td>A22</td>
<td>Emphysema</td>
<td>63</td>
<td>M</td>
</tr>
<tr>
<td>A23</td>
<td>Emphysema</td>
<td>46</td>
<td>M</td>
</tr>
</tbody>
</table>
3.1.2. Pathogens

In the current study, we used one virus pathogen and two bacterial pathogens to induce infection in lungs. Complete name of strains and source as Table 3.2.

Table 3.2: List of pathogens

<table>
<thead>
<tr>
<th>Name</th>
<th>Full name and strains</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>SOIV-WT (A/Giessen/6/2009 H1N1-WT)</td>
<td>Influenza A virus SOIV-WT (A/Giessen/6/2009 H1N1-WT)</td>
<td>Prof. Dr. Stephan Pleschka</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Institute for Medical Virology, Justus, Liebig University Gießen</td>
</tr>
<tr>
<td>PA14 WT strain</td>
<td>Pseudomonas aeruginosa (PA14 WT strain)</td>
<td>Prof. Dr. Susanne Häußler</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Institute for Molecular Bacteriology</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TWINCORE, Centre for Experimental and Clinical Infection Research</td>
</tr>
<tr>
<td>BCG</td>
<td>Mycobacterium bovis strain: Bacillus Calmette–Guérin (BCG) Mtb H37Rv</td>
<td>Prof. Dr. Ulrich Kalinke</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TWINCORE, Centre for Experimental and Clinical Infection Research</td>
</tr>
</tbody>
</table>

3.1.3. Cell lines

We used the Madin-Darby Canine Kidney (MDCK)-II cell line for virus propagation and titration. This cell line is susceptible to all strains of influenza A and B viruses and support replication [176]. MDCK-II cell line was used to titer the IAV by the focus forming assay.

3.1.4. Software/web tools

Several bioinformatic tools were used to design experiments and analyze the data (Table 3.3).

Table 3.3: Table of software and online tools

<table>
<thead>
<tr>
<th>Software</th>
<th>web link</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primer3 - PCR primer design tool</td>
<td><a href="http://bioinfo.ut.ee/primer3-0.4.0/primer3/">http://bioinfo.ut.ee/primer3-0.4.0/primer3/</a></td>
</tr>
<tr>
<td>GraphPad Prism8</td>
<td><a href="https://www.graphpad.com/scientific-software/prism/">https://www.graphpad.com/scientific-software/prism/</a></td>
</tr>
<tr>
<td>The R Project for Statistical Computing (R version)</td>
<td><a href="https://www.r-project.org/">https://www.r-project.org/</a></td>
</tr>
</tbody>
</table>
Buffers for enzyme-linked immunosorbent assay (ELISA), metabolite extraction, and Lactate dehydrogenase (LDH) release assay were prepared according to the manufacturer’s instructions. Tissue digestion buffer was developed from already published protocols (Table 3.4).

**Table 3.4: Buffers used**

<table>
<thead>
<tr>
<th>Use</th>
<th>Buffer</th>
<th>Ingredients</th>
</tr>
</thead>
<tbody>
<tr>
<td>ELISA</td>
<td>Washing buffer</td>
<td>0.05% Tween® 20 (Sigma-Aldrich, P1379) in PBS (Invitrogen, 00,3002)</td>
</tr>
<tr>
<td>ELISA</td>
<td>Coating buffer</td>
<td>3.03 g Na₂CO₃, 6.0 g NaHCO₃ 1000 ml distilled water pH 9.6</td>
</tr>
<tr>
<td>ELISA</td>
<td>Blocking buffer</td>
<td>3% BSA (Roth, 8076.5) in PBS (Invitrogen, 00,3002)</td>
</tr>
<tr>
<td>ELISA</td>
<td>Diluent</td>
<td>0.1% BSA (Roth, 8076.5) and 0.05% Tween® 20 (Sigma-Aldrich, P1379) in PBS (Invitrogen, 00,3002)</td>
</tr>
<tr>
<td>Single-cell suspension</td>
<td>Digestion</td>
<td>1 mg/ml collagenase D (Roche, 10269638001), 2.0 U/ml dispase II (Roche, 04942078001) and 0.1 mg/ml DNase I (DNase I, D4527) in Heps-buffered saline</td>
</tr>
<tr>
<td>Single-cell suspension</td>
<td>ACK lysing buffer</td>
<td>150 mM NH₄Cl, 10 mM KHCO₃ and 0.1 mM Na₂EDTA in dH₂O</td>
</tr>
<tr>
<td>LDH release assay</td>
<td>Tissue lysis buffer</td>
<td>1% Triton X-100 in dH₂O</td>
</tr>
<tr>
<td>Metabolite extraction</td>
<td>Tissue lysis buffer</td>
<td>Phosphate buffer (0.1 M NaH₄PO₄ + mL 0.1 M Na₂HPO₄) 15 ml in ethanol 85 ml</td>
</tr>
</tbody>
</table>
3.1.6. Reagents

The reagents used in all lab procedures are listed in Table 3.5.

**Table 3.5: List of reagents**

<table>
<thead>
<tr>
<th>Chemical</th>
<th>Company</th>
</tr>
</thead>
<tbody>
<tr>
<td>Biocoll</td>
<td>Biochrom AG, Berlin, Germany</td>
</tr>
<tr>
<td>Bromophenol blue</td>
<td>Merck Group, Darmstadt, Germany</td>
</tr>
<tr>
<td>Dimethylsulfoxide</td>
<td>Sigma-Aldrich, St. Louis, Missouri</td>
</tr>
<tr>
<td>Dimethyl-itaconate</td>
<td>Sigma-Aldrich, St. Louis, Missouri</td>
</tr>
<tr>
<td>DPBS (without Mg/Ca)</td>
<td>GIBCO Life Technologies, Darmstadt, Germany</td>
</tr>
<tr>
<td>EDTA</td>
<td>Merck Group, Darmstadt, Germany</td>
</tr>
<tr>
<td>Ethanol (70% and 99%)</td>
<td>Carl Roth Gmbh, Karlsruhe, Germany</td>
</tr>
<tr>
<td>Fetal calf serum (FCS)</td>
<td>Sigma-Aldrich, St. Louis, Missouri</td>
</tr>
<tr>
<td>GlutaMax</td>
<td>GIBCO Life Technologies, Darmstadt, Germany</td>
</tr>
<tr>
<td>Glycerol</td>
<td>Merck Group, Darmstadt, Germany</td>
</tr>
<tr>
<td>Hydroxyethylpiperazinyl-ethansulfonic acid (HEPES)</td>
<td>Carl Roth Gmbh, Karlsruhe, Germany</td>
</tr>
<tr>
<td>Itaconic acid</td>
<td>Sigma-Aldrich, St. Louis, Missouri</td>
</tr>
<tr>
<td>Lipofectamine-2000</td>
<td>Thermo Scientific, Massachusetts, USA</td>
</tr>
<tr>
<td>MACS BSA stock solution</td>
<td>Miltenyi Biotech Bergisch Gladbach, Germany</td>
</tr>
<tr>
<td>Non-essential amino acids</td>
<td>GIBCO Life Technologies, Darmstadt, Germany</td>
</tr>
<tr>
<td>Paraformaldehyde 37% (PFA)</td>
<td>Carl Roth Gmbh, Karlsruhe, Germany</td>
</tr>
<tr>
<td>Penicillin &amp; streptomycin</td>
<td>GIBCO Life Technologies, Darmstadt, Germany</td>
</tr>
</tbody>
</table>
3.1.7. Kits

Kits for ELISA, RNA extraction and RT-qPCR were purchased (Table 3.6).

Table 3.6: List of kits

<table>
<thead>
<tr>
<th>Kit name</th>
<th>Company</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human Prokineticin-2 (PROK2) ELISA</td>
<td>Cusabio technology llc, Houston, Texas.</td>
</tr>
<tr>
<td>Orosomucoid 2 (ORM2) ELISA</td>
<td>Cloud-clone corp. Katy, Texas</td>
</tr>
<tr>
<td>Human IFN-gamma ELISA</td>
<td>Biolegend, San Diego, California</td>
</tr>
<tr>
<td>Humn IL-10 ELISA</td>
<td>Biolegend, San Diego, California</td>
</tr>
<tr>
<td>Huamn IL-6 ELISA</td>
<td>Biolegend, San Diego, California</td>
</tr>
<tr>
<td>Human IL-1B ELISA</td>
<td>Biolegend, San Diego, California</td>
</tr>
<tr>
<td>Human IP-10 ELISA</td>
<td>Peprotech, Hamburg, Germany</td>
</tr>
<tr>
<td>Annexin V Staining</td>
<td>Invitrogen, Darmstadt, Germany</td>
</tr>
<tr>
<td>Propidium Iodide Staining Solution</td>
<td>Invitrogen, Darmstadt, Germany</td>
</tr>
<tr>
<td>Cytotoxicity Detection Kit (LDH)</td>
<td>Roche, Mannheim, Germany</td>
</tr>
<tr>
<td>miRNAeasy Qiagen Kit</td>
<td>Qiagen, Venlo, Niederland</td>
</tr>
<tr>
<td>Rnase-Free Dnase set</td>
<td>Qiagen, Venlo, Niederland</td>
</tr>
<tr>
<td>Eukaryote Total RNA Nano Series II chips</td>
<td>Agilent, Santa Clara, California</td>
</tr>
<tr>
<td>RT PrimeScript™ Master Mix</td>
<td>Takara, Göteborg, Sweden</td>
</tr>
<tr>
<td>SensiFast™ SYBR® No-ROX Kit</td>
<td>Bioline, London, UK</td>
</tr>
<tr>
<td>Absolute®p180 kit</td>
<td>Biocrates Life Sciences AG, Austria</td>
</tr>
<tr>
<td>Chromium™ Single Cell 3’ GEM, Library &amp; Gel</td>
<td>10x Genomics, Pleasanton, California</td>
</tr>
<tr>
<td>Bead Kit v3</td>
<td></td>
</tr>
</tbody>
</table>

3.1.8. Primers

Primers were designed using the primer3 tool and manufactured by Eurofins Genomics (Luxembourg). All primers were tested for specificity by product size by gel electrophoresis before using them in RT-qPCR.
<table>
<thead>
<tr>
<th>Gene name</th>
<th>Primer name</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>ISG15</td>
<td>ISG15-F</td>
<td>TGTCGGTGTCAGAGCTGAAG</td>
</tr>
<tr>
<td></td>
<td>ISG15-R</td>
<td>AGAGGTTCGTCGCATTTGTC</td>
</tr>
<tr>
<td>IL-6</td>
<td>IL6-F</td>
<td>CTACATTGCCGAAGAGCCC</td>
</tr>
<tr>
<td></td>
<td>IL6-R</td>
<td>CCCTGACCCAACCACAATATT</td>
</tr>
<tr>
<td>HPRT</td>
<td>HPRT-F</td>
<td>GAACGTCTTCTGAGATGTG</td>
</tr>
<tr>
<td></td>
<td>HPRT-R</td>
<td>CCAGCAGTCGCAAAGAATT</td>
</tr>
<tr>
<td>PROK2</td>
<td>PROK2-F</td>
<td>TGACAAGGACTCCAATGTG</td>
</tr>
<tr>
<td></td>
<td>PROK2-R</td>
<td>TACGAGTCAGTGAGGCGAG</td>
</tr>
<tr>
<td>ORM2</td>
<td>ORM2-F</td>
<td>GTTCTCTACCTGGACGATGA</td>
</tr>
<tr>
<td></td>
<td>ORM2-R</td>
<td>CTCTTCTCAGTGCTTCTT</td>
</tr>
<tr>
<td>IL-1β</td>
<td>IL-1β-F</td>
<td>TACCCAAAGAAGATGGAA</td>
</tr>
<tr>
<td></td>
<td>IL-1β-R</td>
<td>GAGGTGCTGATGACCAGTTG</td>
</tr>
<tr>
<td>CxCL10</td>
<td>CxCL10_F</td>
<td>CTGCTTTGGGGTTTATCGA</td>
</tr>
<tr>
<td></td>
<td>CxCL10_R</td>
<td>CACTGAAAGAATTTGGGC</td>
</tr>
<tr>
<td>A20</td>
<td>A20_F</td>
<td>ATGCACCGATACACACTGGA</td>
</tr>
<tr>
<td></td>
<td>A20_R</td>
<td>CACAAGCTTCCGGACCTCT</td>
</tr>
<tr>
<td>IL-10</td>
<td>IL10-F</td>
<td>TACCTGGGTTGCAAGCCT</td>
</tr>
<tr>
<td></td>
<td>IL10-R</td>
<td>AGAAATCGATGACGGCC</td>
</tr>
<tr>
<td>HA</td>
<td>HA-F</td>
<td>CTCGTGCTATGGGCGATTTCA</td>
</tr>
<tr>
<td></td>
<td>HA-R</td>
<td>TTGCAATCGTGACGGTGT</td>
</tr>
<tr>
<td>ACOD1</td>
<td>ACOD1_F</td>
<td>ATGCTGCTTTGTGAACGGTG</td>
</tr>
<tr>
<td></td>
<td>ACOD1-R</td>
<td>CTACCACGGAAGGGGATGGA</td>
</tr>
</tbody>
</table>
3.1.9. Culture media

Cell culture media were purchased and supplemented with other nutrients before use. Bacterial media were autoclaved to sterilize them.

**Table 3.8: List of media**

<table>
<thead>
<tr>
<th>Media</th>
<th>Company</th>
</tr>
</thead>
<tbody>
<tr>
<td>RPMI medium 1640</td>
<td>GIBCO Life Technologies, Darmstadt, Germany</td>
</tr>
<tr>
<td>MEM 10X</td>
<td>GIBCO Life Technologies, Darmstadt, Germany</td>
</tr>
<tr>
<td>Luria-Bertani (LB) medium</td>
<td>Invitrogen, Darmstadt, Germany</td>
</tr>
<tr>
<td>Difco™ Middlebrook 7H9 Broth</td>
<td>BD, New Jersey, USA</td>
</tr>
<tr>
<td>AQIX® RS-I 'Ready to Use' Solution</td>
<td>Aqix liquid life, Bedford, United Kingdom</td>
</tr>
</tbody>
</table>

3.1.10. Equipment

**Table 3.9: List of equipments**

<table>
<thead>
<tr>
<th>Name</th>
<th>Company</th>
</tr>
</thead>
<tbody>
<tr>
<td>ELISA plate reader Synergy 2</td>
<td>BioTek, Winooski, VT</td>
</tr>
<tr>
<td>Centrifuge 5415R</td>
<td>Eppendorf, Hamburg, Germany</td>
</tr>
<tr>
<td>Cell incubator</td>
<td>Thermo Scientific, Massachusetts, USA</td>
</tr>
<tr>
<td>Centrifuge table top falcon multifuge 1 S-R</td>
<td>Thermo Scientific, Massachusetts, USA</td>
</tr>
<tr>
<td>LightCycler480 machine</td>
<td>Roche, Mannheim, Germany</td>
</tr>
<tr>
<td>NanoDrop</td>
<td>Thermo Scientific, Massachusetts, USA</td>
</tr>
<tr>
<td>Neubauer counting chamber</td>
<td>Carl Roth Gmbh, Karlsruhe, Germany</td>
</tr>
<tr>
<td>Sterile workbench laminar flow</td>
<td>Thermo Scientific, Massachusetts, USA</td>
</tr>
<tr>
<td>Water bath</td>
<td>Julabo, Allentown, Pennsylvania</td>
</tr>
<tr>
<td>Cell Counter Scepter</td>
<td>Merck Millipore, Massachusetts, United States</td>
</tr>
<tr>
<td>T10 Homogenizer</td>
<td>IKA, Staufen, Germany</td>
</tr>
</tbody>
</table>
3.2. Methods

3.2.1. Preparation and maintenance of Human Lung Tissue Explants (HLTEs)

To develop the infection model, bronchial lung tissue (which was explanted from patients with terminal COPD for clinical reasons) was obtained from the Institute of Pathology, MHH. These tissues were further dissected into small pieces with an average size of approx. 27 mm$^3$ (3x3x3 mm) and an average weight of approx. 30 mg. These pieces (HLTEs) were cultured in RPMI medium without any supplement. In some experiments treatment IA compounds was applied, as mentioned in the Results section. Tissue pieces were cultured overnight in a humidified tissue culture incubator at 37°C, 5% CO$_2$ for overnight; this step was termed overnight washing. After 12-16 h of washing, tissues were infected with the respective pathogen. This protocol was developed by combining already published methods.

**Figure 3-1 Scheme of HLTEs preparation and infection.**

3.2.2. Treatments

Calculated amounts for DI 1 mM and IA 25 mM were added to RPMI medium and pH was adjusted to 7.2 by using 1 M KOH solution. Medium were sterilized by filtering through vacuum filters “Millipore Sterito Sterile Vacuum Bottle-Top Filters” with 0.22 μm pore size. HLTEs were washed/pre-treated overnight with IA or DI containing medium before start of infection.
3.2.3. Infections

3.2.3.1. Influenza A virus

Stock preparation: IAV was propagated in MDCK-II cells in infection medium (FBS-free DMEM medium) supplemented with 1 µg/mL TPCK-treated trypsin (Sigma Aldrich, St. Louis, MO, USA) at 35°C, 5% CO₂. Clarified infection supernatants were aliquoted and virus titer was calculated by focus forming assay.

Infection: For HLTE infection with IAV, infection medium containing 2.0 x 10⁵ ffu/ml in RPMI medium was added to a 24 well plate, placing one HLTE piece per well. Plates were incubated at 37°C for a specific duration, as shown in the Results section.

3.2.3.2. Mycobacterium bovis strain: Bacillus Calmette–Guérin (BCG)

M. bovis H37Rv were grown to mid-log phase in 50 ml of Middlebrook 7H9 (Difco) liquid culture medium supplemented with 0.5% glycerol, 0.15% Tween-80 and 10% oleic acid-albumin-dextrose-catalase (BD Biosciences). Bacterial cells were collected in a 50 ml Falcon tube and then washed twice with 45 ml PBS (Gibco) by centrifugation at 4000G. Bacterial optical density was measured with a spectrophotometer at an absorbance of 600 nm. Infection medium containing 5X10⁶ CFU/ml was prepared and HLTEs were subjected to infection in a 24 well plate.

3.2.3.3. Pseudomonas aeruginosa

P. aeruginosa was cultured in Luria-Bertani broth at 37°C in a shaker incubator and harvested at log phase. CFU were estimated according to the OD value measured as described by Kim et al [177], and infection medium was prepared containing 2x10⁵ CFU/ml in RPMI medium. HLTEs were incubated in 24 well plates for 2 h at 37°C. After infection, HLTEs were washed and incubated in RPMI medium containing 50 µg/ml gentamycin.

3.2.4. Focus forming assay

Cell culture supernatants containing IAV were serially diluted in PBS supplemented with 0.2% BSA and 1% Ca/Mg solution. Recently confluent monolayers of MDCK cells in 96-well plates were rinsed with PBS and each well infected with 50 µl of 10-fold serially diluted virus for 1 h at room temperature. After infection, the inoculum was aspirated and 150 µl Avicel-media 1% Avicel in MEM media supplemented with 0.5% BSA, 2 µg/ml Trypsin and 1% Dextran) was added to the cell monolayer. The plates were then incubated at 37°C, 5% CO₂ for 24 h, followed by removal of the Avicel media and fixation of cells by 4% PFA and 1% Triton X-100. Cells were stained by mouse-anti-NP-antibody followed by anti-mouse-HRP-antibody. AEC-staining-solution was used to visualize the IAV positive cells.

3.2.5. LDH release assay

Supernatants and tissue were collected from HLTE cultures at 0, 24, 48 and 72 h post incubation and stored at 4°C until collection of the 72 h time point. Tissue was lysed in 1% Triton X-100 lysis buffer, tissue lysate and supernatant were diluted and added to 96-well microplates. Catalyst and dye solution
were mixed and added to samples according to the protocol of the Roche LDH cytotoxicity detection kit. Color development was measured by OD at 492 nm (background control at 630 nm). LDH release in percentage was calculated as follows: (100 / LDH of lysate) x LDH of sample.

3.2.6. ELISA

Supernatants collected from HLTE cultures were immediately stored at -20°C. Samples were thawed at 4°C before measuring concentrations by commercial kits for the following proteins; IP10, IL-1β, IL-6, IL-10, IFN-γ, PROK2, and ORM2. In brief, ELISA plates were coated with specific capture antibodies overnight at 4°C, followed by washing and blocking with 3% BSA (Roth, 8076.5). Plates were washed and then incubated with standard proteins or sample supernatants (diluted in 1% BSA) for 2 h. After binding of protein antigens to coated antibodies, detection antibodies were added, followed by adding avidin conjugated with HRP. This was followed by a final quadruple wash step, after which the chromomeric substrate 3,3',5,5'-tetramethylbenzidine (TMB) was added and incubated in the dark. After approximately 15 minutes, the reaction was stopped by adding 0.3 M sulphuric acid (H₂SO₄) solution. Plates were read on an ELISA reader (BioTek, Synergy 2.0) at OD 450 nm and background OD that was earlier measured at 570 was subtracted. Protein concentrations were calculated by comparison with standard curves.

3.2.7. RNA extraction

HLTEs were snap-frozen in liquid nitrogen immediately after collection. RNA extraction of HLTEs was performed using miRNeasy Qiagen Kit (Qiagen, 74104) according to the manufacturer’s instructions with a few modifications. Briefly, Qiazol lysis reagent was added directly to frozen tissue, and disruption and homogenization of HLTE was performed using the Ultra-Turrax® (T10, IKA, 0003737000) for 20-30 sec. at the highest speed. Chloroform was added and the resulting mixture was centrifuged 12,000xg for 15 min at 4°C. The colorless upper layer was transferred to RNA binding columns with 100% ethanol. After two washings of RNA with RPE buffer, DNA was digested by treatment with DNAse I. RNA was eluted into 50 μl RNAse free water. The resulting sample contains total RNA including small RNA (<50 nt) and therefore can be used for both mRNA and snRNA profiling.

3.2.8. RNA quality control

The RNA quality and quantity were determined using the NanoDrop spectrophotometer (Manufacturer). A 260/280 ratio between 2.0 – 2.2 indicated RNA without contamination. The RNA Integrity Number (RIN) was determined with an Agilent Bioanalyzer using Eukaryote Total RNA Nano Series II chips (Agilent, 5067-1511).

3.2.9. Complementary DNA (cDNA) generation

Following RNA extraction, cDNA was prepared using the RT PrimeScript™ Master Mix (Takara) following the manufacturer's instructions. Briefly, 400 ng of RNA was reverse transcribed by adding 2 μl master mix (5x) and making a total reaction volume of 10 μl, followed by incubation at 37°C for 15 min. Afterward, all samples were heated to 85°C for 10 sec to inactivate the enzymes. The cDNA thus
synthesized was diluted further in RNase free water to get a final RNA concentration of 5 ng/µl of the cDNA.

### 3.2.10. Gene expression analysis using real-time qPCR

Primers ([Table 3.7](#)) were either designed using PrimerBLAST (National Center for Biotechnology Information, National Institutes of Health) or the Primer3 online tool [178]. Primers were designed to span exon-intron boundaries, have annealing temperatures around 60°C and generate amplicons of 100 bp - 300 bp. RT-qPCR reactions were set up in a final volume of 20 µl, using the Sensifast™ SYBR® No-ROX Kit (Bioline, Taunton, MA) and the primers listed. RT-qPCR was performed in a LightCycler® 2.0 instrument (Roche, Mannheim, Germany), using 45 cycles of the following program: 95°C for 15 sec., 60°C for 15 sec., and 72°C for 15 sec. To exclude artifacts resulting from primer dimer formation, melting curve analysis was performed using the sequence 95°C for 15 sec., 60°C for 15 sec., 95°C for 1 min. and 37°C for 30 sec. Relative expression of the mRNA targets was calculated using the \(2^{-\Delta\Delta CT}\) method [179]. Amplification of a single amplicon was confirmed by obtaining dissociation curve (melt curve) profiles as well as using gel electrophoresis to verify the size of the reaction product.

### 3.2.11. Metabolite extraction

Snap-frozen tissue samples were weighed and ice-cooled extraction solvent (ethanol/0.01 M phosphate buffer with 85/15, V/V) was added. Tissues were disrupted and homogenized using the Ultra-Turrax® (T10, IKA, 0003737000) for 40-60 sec. Lysates were centrifuged for 5 min, 4°C at 10,000xg and supernatants were collected in new tubes and stored at -80°C until further analysis.

### 3.2.12. Targeted metabolomics using Biocrates p180 kits

The absolute®p180 kit (Biocrates Life Science AG, Innsbruck, Austria) was used for metabolic profiling of all tissue extract samples on an HPLC-coupled triple-quadrupole mass spectrometer (Sciex, API 4000™, USA). The kit quantifies 188 metabolites belonging to six classes of analytes, comprising amino acids (n=21), biogenic amines (n=21) by liquid chromatography tandem mass spectrometry, and acylcarnitines (n=40), glycerophospholipids (n=90), sphingomyelins (n=15) and hexoses (n=1) by flow injection analysis (bypassing the HPLC step). All assays were performed according to the manufacturer’s recommendations (UM_p180_ABSciex_11 and Application Note 1003-1, Biocrates Life Science AG, Innsbruck, Austria).

### 3.2.13. Small non-coding RNA (sncRNA) sequencing

RNA sequencing was performed at Genome Analytics (GMAK) HZI - Helmholtz Centre for Infection Research, Braunschweig. The sequencing process is split into three parts: library preparation, cluster generation, and sequencing. Libraries were prepared using the NEBNext® Small RNA Library Prep Set for Illumina. 1 µg of input RNA was used and the synthesized cDNA was amplified with 12 PCR cycles. Libraries were prepared separately for each biological replicate. Samples were sequenced on Illumina HiSeq2500 (2x 50 bp paired-end), generating 50 bp single reads and approximately 16 million reads passing filter for each sample.
3.2.14. sncRNA annotation

FASTAQ files obtained after small RNA sequencing were used to annotate the transcripts as explained in the OASIS 2.0 web tool tutorial [180]. Briefly, FASTAQ files were compressed using the OASIS compressor. Compressed files were uploaded to the web tool for annotation. The annotation process comprises removing 3’ adapters, quality control, mapping to the reference genome (Homo sapiens – hg38), and the counting of reads in each sRNA for each sample. Obtained reads were subjected to detect new miRNAs by predicting their structures and genomic location. Contaminating viral or bacterial RNA sequences were removed.

3.2.15. Bulk RNA sequencing

Quality and integrity of bulk RNA, as well as RNA from different cell types, was checked on an Agilent Technologies 2100 Bioanalyzer (Agilent Technologies; Waldbronn, Germany). The RNA sequencing library was generated from 10 ng total RNA using NEBNext® Single Cell/Low Input RNA Library according to the manufacturer’s protocols. The libraries were sequenced on an Illumina NovaSeq 6000, using the NovaSeq 6000 S1 Reagent Kit (100 cycles, paired-end run 2x 50 bp) with an average of 5 x10⁷ reads per RNA sample. Library preparation and sequencing were performed by Genome Analytics Research Group (GMAK), HZI. Large sets of high-throughput sequencing reads were mapped against the Homo sapiens hg38 reference genome via STAR 2.7.3a tool.

3.2.16. Transcript normalization and differential expression analysis

All the following analyses were done using the R environment and programming code. DESeq2 was used to normalize the counts of each gene to account for differences in sequencing depth and low count variability. The DESeq2 normalization metric is based on the median of ratios method. Following normalization, DESeq2 was run in a pairwise manner, identifying significantly differentially expressed genes with higher levels of expression in the infection group in comparison to the uninfected group. This method provided functionality for a pair-wise comparison (identifying differentially expressed genes between two groups) in the form of a Wald test, and for multiple comparisons (classically used for time-series data) in the form of a likelihood ratio test. The output contained nominal p-values, False Discovery Rate (FDR) or P-adjusted values (correction for multiple tests computed with the Benjamini–Hochberg procedure) and fold change. DESeq2 used an empirical Bayes shrinkage to detect and correct for dispersion and log2-fold change estimates. The R software package Visualization of Differential Gene Expression using R (ViDGER) was used to display differential expression outputs uniformly.

ROC curve is a graphical approach to describe the ability of a biomarker to correctly classify “cases” and “non-cases”. It plots sensitivity and specificity (the sensitivity of our prediction against 1- the specificity), the outcome determines efficiency of the biomarker [181]. The area under the ROC curve (AUC) is taken as a measure for the discriminatory ability of a selected marker by combining two test outcomes, an AUC of 1.0 defining perfect discrimination. To determine the AUC, R package MetaboAnalystR was used.
3.2.17. Preparation of single-cell suspension

After 24 hpi, tissues were collected and processed to prepare a single-cell suspension for flow cytometry and cell sorting. Briefly, lung tissues were chopped in digestion buffer (1 mg/ml collagenase D (Roche, 10269638001), 2.0 U/ml dispase II (Roche, 04942078001) and 0.1 mg/ml DNase I (DNase I, D4527) in Hepes-buffered) with surgical scissors and homogenized. Tissue mixture was incubated for 30 minutes at 37°C with mild agitation. After 30 min, tissue was passed through a 30G syringe. The resulting single-cell suspension was filtered through a 40 µm nylon cell strainer and erythrocytes were lysed using ACK (Ammonium-Chloride-Potassium) lysing buffer. Cells were counted using an automated cell counter (Scepter, Millipore, PHCC00000).

3.2.18. Live/Dead and apoptosis detection

Single-cell suspension was washed 1x with PBS and 1x with binding buffer. Cells were incubated with fluorochrome-conjugated Annexin V (FITC) for 15 min. After incubation, cells were washed again and resuspended in propidium iodide staining solution. After staining, cells were sorted and analyzed by FACS (Cell Sorter Facility MHH). Dead and duplets cells were excluded from further analysis.

3.2.19. Single-cell RNA-Seq

A single-cell suspension obtained after cell sorting (see above) was processed for single-cell transcriptomics. Single-cell 3’ RNA-Seq libraries were prepared using Chromium Single Cell V3 Reagent Kit and Controller (10x Genomics) as described in the user guide. In brief, cells (n=4500) for a targeted recovery of 3500 along with gel beads, master mix and partitioning oil were loaded in designated wells on a chromium chip. The chip was placed in a Chromium Controller and run for Gel Bead-in-Emulsion (GEMs) preparation. After completion of the run, GEMs were run in PCR tubes for RT incubation in a thermal cycler. Post GEM-RT incubation, cDNA was cleaned up and amplified. cDNA was washed and quality was evaluated. Gene expression libraries were constructed and then libraries were assessed for quality (TapeStation 4200, Agilent). Libraries were sequenced by NextSeq550 using the NextSeq 500 High Output Kit v2.5 (1x75 cycles 400M cluster). We created a reference package for two species genomes (Human and IAV H1N1) and ran the sequences against this genome package. Initial data processing was performed using the Cell Ranger version 2.0 pipeline (10x Genomics).

3.2.20. Single-cell RNA data analysis

Cell Ranger output files “outs” contained raw counts for each sequenced genes (Human and H1N1 virus) of each cell. These files were used as input for further processing using the R package Seurat 3.0. Following the standard pre-processing workflow, cells were filtered that have unique feature counts over 2,500 or less than 200. Raw counts were normalized by the global-scaling normalization method “LogNormalize”. Principal component analysis score was used to determine the ‘dimensionality’ of all the cells. Cell clusters were identified as different cell populations based on the expression of their canonical markers recently published. Differential expression was determined for each cell type by comparing the IAV infected cells with uninfected cells. IAV infected cells were identified by the presence of IAV encoded transcripts.
3.2.21. Gene Set Enrichment Analysis (GSEA)

Hallmark, gene ontology and KEGG gene sets were downloaded from GSEA online portal. R-Bioconductor package “fgsea” was used to pre-rank the DEGs and annotate them against different pathways.

3.2.22. Statistics

Statistical tools for each analysis are described in detail with the results or detailed in the figure legends. Briefly, GraphPad Prism was used to analyze RT-qPCR, ELISA, metabolomics and LDH release assay results. Due to inter-and intra-donor variation, data is not “normally distributed” so a non-parametric test was applied. The Wilcoxon–Mann–Whitney test was performed to determine p-values for between group differences in medians. P values are summarized as values < 0.001 = ***, 0.001 to 0.01 = ** and 0.01 to 0.05 = *. Data are shown as mean ± SEM. Biological replicates are mentioned in the legend of all figures. For transcriptome data, DESeq2 test was applied which returned log2-fold change, p-adjusted values, and p-values.
4. Results

4.1. Influenza A virus infection establishment on HLTE model

4.1.1. LDH Cytotoxicity Colorimetric Assay

LDH is a cytosolic enzyme present in cells, and its release from cells is a good indicator of cell membrane damage, hence a measure of cell viability. LDH levels released into medium from HLTE cultured with or without IAV were quantified at 24, 48 and 72 hours post-infection (hpi). An increase \( p < 0.05 \) in LDH levels after 24 hpi was observed in control and infection groups, but cell death did not exceed 15% of the total (Figure 4-1 (A)). There was slightly higher release of LDH in supernatants from infected than from uninfected tissue. These changes were mainly observed in the first 48 hpi, and thereafter the levels of LDH did not change.

4.1.2. Focus forming assay for IAV titer

Titers of live virus particles in the medium of HLTE infected with IAV were determined by immune focus assay at 24, 48 and 72 hpi. Highest released virus titer was measured in the supernatant at 24 hpi, and then a decrease in live virus particles was observed at 48 and 72 hpi (Figure 4-1 (B)). A large difference in virus titers was seen among donors and among the different tissue pieces from the same donor at all tested time points.

Figure 4-1: LDH release and virus load in supernatant: Cell viability as measured by LDH release in HLTE culture medium at 0, 24, 48, and 72 hpi. Results are shown as percentage by comparing the LDH levels released in medium to LDH in tissue lysate as the positive control \((n=5)\). (B) Detection and quantification of IAV in supernatants of IAV-infected HLTEs was performed by counting foci formation on MDCK-II confluent monolayer \((n=5)\).
4.1.3. RNA quality assessment

Tissue samples were preserved initially in RNAlater or, after protocol modification, snap-frozen in liquid nitrogen. Total RNA was extracted from these tissues in different batches. RIN values of extracted total RNA from tissue preserved in RNAlater were highly variable, ranged from 1 to 10 (Figure 4-2). After protocol modification, RINs of RNA from snap-frozen tissue were stable across all samples and ranged from 7 to 10. This method was subsequently adapted for all further experiments and analyses. RNA quantities ranged from 5-15 µg, depending on tissue size and tissue homogenization method (beads or disperser).

Figure 4-2: Effect of tissue preservation method on RNA integrity: HLTE tissue was preserved in either RNAlater or snap-frozen in liquid nitrogen and stored at -80°C. Tissues were homogenized and total RNA was extracted (n < 35, mean± SEM, Wilcoxon–Mann–Whitney test).

4.1.4. Quantification of viral and cytokine expression by RT-qPCR

To assess the functionality of infection, viral mRNA and immune response genes were quantified after infection by RT-qPCR. HA, CXCL10 and ISG15 transcripts were quantified at 24, 48 and 72 h in tissues with or without IAV infection. The trend in viral transcript levels showed an increase until 48 hpi, and then a decrease at 72 hpi in tissue. Hypercytokinemia is a trademark of IAV infection [182]. Upregulation of CXCL10 and type I IFN signaling associated genes like ISG15 are key features of the early immune response to IAV infection [183]. A host response to infection was also observed at the gene expression level. An induction in the expression of both host genes tested, ISG15 and CXCL10, was seen during infection (Figure 4-3). The expression of ISG15 increased with time and was highest at 72 hpi, whereas, highest CXCL10 mRNA levels were detected at 48 hpi and declined by 72 hpi.
Figure 4-3: Quantification of viral and host mRNA transcripts by qPCR: total RNA from tissue infected with or without IAV was subjected to quantify the expression of viral HA mRNA (A) and host response genes ISG15 (B) and CXCL10 (C) at indicated time points. For each sample, relative expression of the selected host genes was calculated using the formula \( \Delta Ct = Ct_{\text{target gene}} - Ct_{\text{housekeeping}} \), where Ct values were the mean of two test replicates. (n = 7, mean± SEM, Wilcoxon–Mann–Whitney test).

4.2. Establishment of viral and bacterial infection

4.2.1. Tissue exclusion criteria

It was observed that stiff fibrotic lung tissues were less prone to IAV infection, so for further viral and bacterial infection study, these lung tissues were excluded. In addition, lung tissues from cystic fibrosis patient had already colonized bacterial infection, so these were also not used. Preexisting bacterial infection due to underlying diseases in lungs was screened by bacterial culture. After an overnight washing step, bacterial contamination was assessed by change in media color. Further confirmation was done by observing media under a microscope (Figure 4-4).

Figure 4-4: Bacterial contamination: HLTEs cultured overnight in RPMI medium show a major color change due to pH change. Bacterial biofilm formation was observed under the microscope in wells with color change.
4.2.2. Immune response to infections

HLTE tissues were cultured and infected separately with IAV, BCG or *P. aeruginosa*, as described in Methods. Supernatants were collected to quantify the released proteins using ELISA, and qPCR was performed to quantify the mRNA transcripts of an established cytokines panel. Released cytokines were quantified by specific ELISAs that showed a distinguished profile among different infections. IP10 and IFNγ levels in supernatant were higher in IAV infection as compared to both bacterial infections (Figure 4-5). IL-6, IL-10 and IL-1β levels produced in bacterial infections were significantly higher than the viral infection. Of the two bacterial infections, the response was more prominent in *P. aeruginosa* than compared to BCG. Quantification of these cytokines at the mRNA level showed a very similar profile of expression.

![Figure 4-5](image_url)

**Figure 4-5:** Immune response to infections: A panel of cytokines was measured to study the immune response to infection. HLTEs from different donors were separately infected with IAV, BCG and *P. aeruginosa*. Supernatants from infected HLTE of different donors were collected at 24 hpi. IP10, IL6, IL10 and IFNγ levels were analyzed by the ELISAs as described in Materials and Methods. RT-qPCR for some cytokine genes was carried out on mRNA extracted from HLTEs with the indicated infections at 24 hpi (F-I). Results are shown as n-fold change in gene expression (2^-ΔCt) as described in Methods. (n = 12, mean±SEM, Wilcoxon–Mann–Whitney test).
4.3. Evaluation of candidate biomarkers

4.3.1. ACOD1 and A20

Regulation of ACOD1 and A20 gene expression after P. aeruginosa infection in THP1 cells had been identified by our research group [90]. To validate this finding, relative expression of ACOD1 and A20 mRNA was quantified by RT-qPCR from HLTEs infected with IAV, BCG or P. aeruginosa at 24 hpi. Results showed significantly higher expression of ACOD1 and A20 genes in P. aeruginosa infection, but only slight changes in IAV and BCG infections (Figure 4-6). A synergetic effect was also observed in expression of both targets in response to IAV and P. aeruginosa co-infection.

![Figure 4-6: Expression of ACOD1 and A20 upon stimulation with bacterial and viral infections in the HLTE model. HLTEs were incubated and infected with IAV, BCG and P. aeruginosa separately and subjected to mRNA extraction. qPCR was performed to quantify the expression levels of ACOD1 and A20 mRNA transcripts. (n = 5, mean± SEM, Wilcoxon–Mann–Whitney test)](image)

4.3.2. Orosomucoid 2

Supernatants collected from HLTE infections were measured for ORM2 concentration by ELISA. To study the changes in ORM2 gene expression following the infection, ORM2 transcripts from HLTE tissue were quantified by RT-qPCR. ORM2 protein was detected in supernatants of both uninfected and infected cultured tissue, but there were no major differences among uninfected and infected groups (Figure 4-7). Also, there was no change in expression of the ORM2 gene in any of the studied infections.
ORM2 response to infection: HLTEs were subjected to viral and bacterial infections separately as indicated above. qPCR was performed to quantify the expression levels of ORM2 mRNA transcripts. qPCR results are shown as n-fold change in gene expression ($2^{-\Delta\Delta Ct}$) as described in methods. Supernatants were collected and ORM2 protein was quantified using specific commercial ELISA ($n = 7$, mean± SEM, Wilcoxon–Mann–Whitney test).

4.3.3. Prokineticin-2

To evaluate the potential of PROK2 as biomarker in infection, PROK2 protein and mRNA levels were quantified in this model. PROK2 proteins levels were significantly higher in supernatants of both bacterial infections, but it was unchanged for IAV infection (Figure 4-8). A similar trend was observed at the transcript levels. Results show the potential of PROK as biomarker in bacterial infections.

Figure 4-7: ORM2 response to infection: HLTEs were subjected to viral and bacterial infections separately as indicated above. qPCR was performed to quantify the expression levels of ORM2 mRNA transcripts. qPCR results are shown as n-fold change in gene expression ($2^{-\Delta\Delta Ct}$) as described in methods. Supernatants were collected and ORM2 protein was quantified using specific commercial ELISA ($n = 7$, mean± SEM, Wilcoxon–Mann–Whitney test).

Figure 4-8: PROK2 response to infection: HLTEs were subjected to viral and bacterial infections separately as indicated above. qPCR was performed to quantify the expression levels of PROK2 mRNA. qPCR results are shown as n-fold change in gene expression ($2^{-\Delta\Delta Ct}$) as described in Methods. Supernatants were collected and PROK2 protein was quantified using specific commercial ELISA ($n = 7$, mean± SEM, Wilcoxon–Mann–Whitney test).
4.4. Metabolic changes in infected HLTE

4.4.1. Metabolite detection

Out of 188 metabolites, 128 (68%) were detected above the limit of detection (LOD) in more than 75% of total samples and were thus included in subsequent analyses. These metabolites included 20 amino acid (AA), 11 biogenic amines (BA), 9 acylcarnitines, 75 glycerophospholipids, 13 sphingomyelins, and the sum of hexoses. Acylcarnitines and biogenic amines were poorly detected across all samples as compared to the glycerophospholipids and sphingomyelins (Figure 4-9). The observed efficiency of metabolite detection agreed well with information provided by the manufacturer and published data [184].

![Figure 4-9](image)

**Figure 4-9:** Class wise abundance of metabolites across all tissue samples (n=20): Each doughnut presents one class of metabolites. Abundance level (% of samples with concentration >LOD) of metabolites is categorized in 4 groups: 1) above 99%, 2) 75-99%, 3) below 75% and 4) without any detection. Size of pie slice and label legend presents the abundance levels of metabolites in each class. In total, 128 metabolites (dark blue and light blue in the pie chart) had concentrations >LOD in 75% of samples and were thus included in subsequent analyses. Of note, amino acids, sphingolipids and glycerophospholipids showed better detection than acylcarnitines.
4.4.2. Metabolite-based relationships among the infections

Principal component analysis (PCA) was used to visualize global differences in metabolite populations among different infections. Data shown in this study were obtained from only 3 donors with 1-2 technical replicates for each infection group. In general, there was no major separation of infection groups. It can be seen that uninfected and IAV infected groups overlap based on metabolite concentrations, whereas both bacterial infections overlap with each other, but are separate from uninfected and IAV infected groups (Figure 4-10 (A)).

4.4.3. Metabolites expression profile

Of the different classes of metabolites, BAs and AAs were the most strongly regulated classes among the different infections. Kyn, histamine, t4-OH-Pro and taurine among BAs and tryptophan, alanine and leucine among AAs were the most differentially abundant. Euclidean clustering based on differential expression of glycerophospholipids and biogenic amines showed uninfected control group distant to all infections, and IAV and BCG showed highest similarity among all infection groups (Figure 4-10 (B and C)). However, differential levels of amino acids show uninfected as the closest neighbor to IAV infection, whereas P. aeruginosa infection clusters farthest from the other groups (Figure 4-10 (D)). Heatmaps show only the top 10 regulated metabolites from each class of metabolites. Different infections lead to changes in the levels of metabolites, and each infection resulted differently for each class of metabolites.
Figure 4-10: Metabolite expression profile based on infection groups: (A) PCA plot was performed using the dataset comprising the 128 metabolites that were quantified above LOD in ≥75% of the samples. The distance among the groups shows the (dis)-similarity based on metabolite expression profile. Each dot represents one sample. The heatmaps shows the top ten most differentially abundant metabolites in class glycerophospholipids (B), biogenic amines (C), and amino acids (D). Euclidean clustering was performed to show (dis)-similarity among the infection groups. (Data presented is drawn from 3 donors with 1-2 technical replicated for each infection group)
4.4.4. Differentially regulated metabolites

Major changes observed in regulation are in levels of AA and BA in tissues to infections. The top 3 highly regulated metabolites (BAs and AAs) are presented in Figure 4-11. Kyn levels detected in IAV and *P. aeruginosa* infected tissues were higher than in uninfected and BCG-infected tissue. Tryptophan (Trp), precursor of Kyn, levels were decreased in tissue infected from *P. aeruginosa* and IAV. Kyn/Trp was significantly higher in tissue infected with *P. aeruginosa* as compared with uninfected tissue. Levels of putrescine were significantly higher in tissues infected with both types of bacteria (Figure 4-11). Increase in levels of putrescine can be due to production by bacteria. The Kyn/Trp ratio reflects activity of the enzyme IDO, which is highly increased in immune cells under inflammatory conditions [185].

![Graphs showing expression of biogenic amines](image)

**Figure 4-11:** Expression of biogenic amines: expressional changes of biogenic amines in HLTE tissues following bacterial and viral infections (n = 3, mean± SEM, Wilcoxon–Mann–Whitney test.).
4.5. Differential reprogramming of sncRNA during infection

4.5.1. sncRNA abundance and species distribution

FASTAQ files obtained from Illumina HiSeq2500 were subjected to trim the adapter sequence and length filter (15-40 nucleotides). An average count of 22 million (M) total reads was detected across all samples (Figure 4-12). Among these, an average of 14 M reads was uniquely mapped against the human genome hg38. Four different species of sncRNA were detected, a) microRNAs (miRNAs), b) small nucleolar RNAs (snoRNAs), c) PIWI-interacting RNA (piRNA) and d) small nuclear RNA (snRNA). miRNA was the most abundant group, making up roughly 92% (13 M) of total uniquely mapped reads, whereas piRNA (0.33 M), snoRNA (0.31 M) and snRNA (0.02 M) reads were very low. These reads were further subjected to a filter where sncRNA species with at least an average of 5 reads in any infection group were selected for further analysis. miRNA had the most species (n=890) passing this filter as compared to the other three groups, i.e. piRNA (394), snoRNAs (165) and snRNA (209). Among all these species, 51 sncRNA were significantly regulated in a multiple group comparison.

Figure 4-12: A universal presentation of sncRNA reads in all samples: (A) Total number of sequenced reads which passed the length filter (15-40 nucleotides) and sncRNAs which were uniquely mapped to the human genome hg38. (B) Distribution of different sncRNA species in all uniquely mapped reads. (C) sncRNA species were further filtered all by keeping only mapped sncRNA transcripts possessing with an average of 5 reads within each infection group. 1678 transcripts passed this filter and were used for subsequent analysis. (D) Number of sncRNA species regulated in a multiple group comparison using ANOVA. (n=5, each donor has uninfected and 3 infection groups with 2-3 technical replicates, RNA pooled)

4.5.2. Principle component analysis

PCA was conducted to assess whether there were global differences in sncRNA expression patterns among the different infection groups. This analysis showed that the uninfected and IAV infected groups clustered together, whereas the two bacterial infection groups overlapped (Figure 4-13). This indicates that the un- and IAV infected groups are more closely related to each other than to the BCG and P.
*P. aeruginosa* infection groups. Altogether, the *P. aeruginosa* infection differed the most from the other groups. This indicates that sncRNA expression changes are greater in *P. aeruginosa* infection than in the other 3 groups.

**Figure 4-13:** Principal component analysis (PCA) of sncRNA populations of (un)infected groups: Plot showing the principal components (PC1, PC4), as PC4 showed highest variance after PC1. The 4 groups are marked by different colors as indicated in the legend alongside the plot. One dot presents one individual of each infection treatment group.

### 4.5.3. Differentially expressed sncRNA

Differences in expression of sncRNAs in HLTEs after bacterial or viral infection were analyzed by using the DESeq2 tool. It provides its own method of normalization, which is based on median log deviation. Following normalization, DESeq2 was run in a pairwise manner, identifying significantly differentially expressed sncRNA in each infection group compared to the uninfected group. In addition to this, we report significantly differentially expressed sncRNA for pairwise comparisons between each infection group. The differential expression analyses were implemented with a q-value (p-adjusted value for False Discovery Rate [FDR]) threshold of 0.1 to enhance the scale of our downstream analysis.
We determined that a total of 90 sncRNAs were differentially expressed between the *P. aeruginosa* infected versus non-infected HLTEs (Figure 4-14 (B)). These DE sncRNAs were either upregulated or downregulated (ranging from 3 to -1.6 log2 fold change). In contrast, only 18 sncRNAs were differentially expressed between BCG-infected and uninfected HLTEs. Surprisingly, there were only 13 differentially expressed sncRNA when both bacterial infections are compared. There were no differentially expressed sncRNA between the IAV-infected and uninfected HLTEs.

**Figure 4-14:** Differentially expressed sncRNA: (A) number of differentially expressed transcripts (DETs) at an adjusted p-value (padj = 0.1) for each infection group in comparison to the others. Higher color intensity indicates a higher number of DETs. (B) Venn diagram showing the number of DETs unique for each infection type and shared among them.

### 4.5.4. Hierarchical clustering of sncRNA

A multigroup comparison (based on t-test) was performed to highlight the sncRNA species which were most differentially regulated between any two groups. The relationships among the infection groups seen in the PC analysis are further confirmed by visualization through heatmaps and by hierarchical clustering of all sncRNAs expression data. Heatmaps only show the top 50 sncRNAs (Highest variance based on ANOVA). There are two major clusters formed, one for bacterial infections and second for IAV infection and uninfected. This shows an expected separation in expression of both bacterial infections from viral and uninfected samples.

There are two major clusters of sncRNA observed based on their expression profile within each infection group. The bigger cluster presents upregulated sncRNA species in bacterial infection, whereas a smaller cluster shows a group of sncRNAs downregulated in bacterial infections.
Figure 4-15: Hierarchical clustering based on sncRNAs expression. Heatmap showing top 50 (based on ANOVA) differentially expressed sncRNAs in infection and uninfected control. The right side of the heatmap shows the names of these 50 transcripts. A histogram scale in the upper left corner demonstrates the variance-stabilized count value used for the heatmap. The hierarchical clustering was implemented with the Euclidean distance measure. The dendrogram above the heatmap shows similarity relationships among the infection groups. The dendrogram on the left side of the heatmap shows relationships in expression among the sncRNA species.
4.5.5. Global differences in expression of sncRNAs

A combination of differential expression analysis and binary ROC analysis were then used (1) to assess the degree of sncRNAs reprogramming in the different infections and (2) to obtain a first overview of differences in expression of sncRNA. There were no significant differences in sncRNA expression between IAV infection and uninfected samples. There were only 3 sncRNA in BCG infection which had p-adjusted below 0.05, and among these, has-miR-4773 and has-miR-212-5p had log2 fold change higher than 1. On the other hand, the calculated AUC values of both miRNA are 1 in comparison to uninfected, which is considered best for a biomarker.

![Graphs of differential expression analysis](image)

**Figure 4-16:** Visualization of differentially expressed sncRNA in the three types of infection: Pairwise comparison of IAV (A), BCG (B) and *P. aeruginosa* (C) was performed and plotted as p-adj on the X-axis; the Y-axis shows log2 fold change between infected and uninfected samples. An intercept line along the X-axis shows p-adj = 0.05; the Y-axis from -1 to 1 log2 fold change. The color scale presents the AUC of each transcript in comparison to uninfected tissue. Further, the class of each molecule is shown different shape as figure legend. (n=5)
Much greater differences in the expression of sncRNA were observed in *P. aeruginosa* infection as compared to uninfected. There are 66 differentially expressed transcripts which have adjusted p values < 0.05. These transcripts included all sncRNA species, i.e. miRNA (n=31), piRNA (n=21), snRNA (n=6) and snoRNA (n=6) **Figure 4-16.** Expression differences ranged from 3 to -1.6 log2 fold change, with an AUC value 0.8 to 1. Major fold change differences in expression are seen in piRNA, followed by fold change differences in miRNA.

**4.5.6. sncRNA expression changes in *P. aeruginosa* infection**

Several sncRNA species showed differential expression after *P. aeruginosa* infection (**Figure 4-17**). To further look into potential biological implications of the regulated miRNA, targets were predicted by applying a highly strict filter (score: 1) and comparing results obtained with 3 databases (TargetScan, miRDB and miRTarBase) (**Table 4.1**).

**Table 4.1:** List of miRNA differentially expressed in *P. aeruginosa* infection along with their predicted targets by miRWalk.

<table>
<thead>
<tr>
<th>miRNA name</th>
<th>Predicted targets</th>
<th><em>P. aeruginosa</em> infection outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>hsa-miR-194-3p</td>
<td>RAP1GDS1, MRI1, TPS3, PPP1R10</td>
<td>Upregulated</td>
</tr>
<tr>
<td>hsa-miR-221-3p</td>
<td>UBN2, RNF4, ERBB4, ZNF652, SOCS3, BCL2L11, FBX028, CTCF, TRPS1, ARID1A, NDFIP1, MIDN</td>
<td>Downregulated</td>
</tr>
<tr>
<td>hsa-miR-222-3p</td>
<td>PANK3, BCL2L11, TP53BP2, ZFYVE16, ESR1, PHACTR4, UBN2, TMCC1</td>
<td>Downregulated</td>
</tr>
<tr>
<td>hsa-miR-212-5p</td>
<td>POGZ, TMEM178B</td>
<td>Upregulated, also in BCG</td>
</tr>
<tr>
<td>hsa-miR-1247-3p</td>
<td></td>
<td>Downregulated</td>
</tr>
<tr>
<td>hsa-miR-148a-5p</td>
<td>AFF2</td>
<td>Downregulated</td>
</tr>
<tr>
<td>hsa-miR-148b-3p</td>
<td>USP48, ARL6IP1, RCC2, DDX6, TBL1XR1, TNRC6B, C1GALT1, GLRX5, BMP3, PHACTR2, ZDHHC17, LC7A11, OSBPL11</td>
<td>Downregulated</td>
</tr>
<tr>
<td>hsa-miR-665</td>
<td>PSMF1, SURF4, PPP2R2A, GNL3L, SMYD1, MEX3A, CACNG8, HEYL, PGPEP1, TMEM86A, KLFHC3, SRF</td>
<td>Upregulated</td>
</tr>
<tr>
<td>hsa-miR-4773</td>
<td>ZBTB39, HNRNPC, ARL8A</td>
<td>Upregulated</td>
</tr>
<tr>
<td>hsa-miR-181a-3p</td>
<td></td>
<td>Downregulated</td>
</tr>
</tbody>
</table>
**Figure 4-17:** Differentially expressed sncRNA profiles in response to the three types of infection: The column plots show the normalized number of transcripts (Y-axis) in each infection group (X-axis). The panel contains all studied species of sncRNAs (miRNA, piRNA, snRNA and snoRNA) (n = 5, mean ± SEM, Wilcoxon–Mann–Whitney test).
4.5.7. Functional annotation

To elucidate the possible roles of the 31 known differentially expressed miRNAs in response to *P. aeruginosa* infection, the potential targets of these miRNAs were predicted with the TargetScan, miRDB and miRTarBase programs using the miRWalk 3.0 tool. Only the overlapping targets identified by all three programs were considered as predicted targets for each miRNA. A total of 158 targets were identified for gene ontology (GO) analysis. GO enrichment analysis was implemented through PANTHER14.1. The results exhibited that these genes were categorized as follows: 49 as metabolic process, 47 as cellular process, and 33 as biological process. Most of the pathways involved were related to cell cycle (Figure 4-18).

![Biological Process](image)

![Molecular functions](image)

**Figure 4-18**: Gene Ontology enrichment terms for putative miRNA targets: Pie charts of the enriched biological processes (A), molecular functions generated from the putative targets of the 31 key differentially expressed miRNAs in *P. aeruginosa* infections (B).
4.6. The transcriptomic response to infections

4.6.1. Read alignment and quantification of expression levels

Libraries were prepared from total RNA using NEBNext® and were sequenced. FASTAQ files output from NovaSeq PE50 were subjected to adapter trimming and mapping against hg38 human genome via STAR (v2.4.2a) tool. RNA-sequencing produced a mean of ~50 million reads per sample. RNA Integrity Number (RIN) values were always ≥ 8.0, ensuring the RNA sequenced was of good quality. Due to the poly-A selection used in library preparation, rRNA contamination should be minimal in this dataset. Reads were distributed equally across all samples, irrespective of treatment or donor. After mapping, a mean of 29,000 genes expressed per sample were calculated via Ensembl gene models (Figure 4-19).

To account for the depth of sequencing impact, raw counts were normalized by 3 popular tools; edgeR, voom, and DESeq2. Comparison between these methods showed that reads were equally distributed along all samples by using edgeR and DESeq2 methods. After normalization via DESeq2, differential

![Figure 4-19: Read counts distribution and normalization.](image)

(A) Number of deferentially regulated genes among the infection groups using different methods. (B) A comparison of reads log10(FPKM) distributions for the experimental infection groups. (C) Number of genes expressed along all samples after normalization.
expression was calculated by use of negative binomial generalized linear models.

### 4.6.2. Identification of differentially expressed genes (DEGs)

DESeq2 was used to identify the differential expression among the groups in the normalized data. Differential expression analyses were performed in pairwise comparisons among the groups. The output file contains log2-fold change, p-adjusted values, and p-values.

The results showed that *P. aeruginosa* infection had the highest number of DEGs (n=749), followed by BCG infection (n=157). Interestingly, there were no DEGs when both bacterial infections were compared with each other (Figure 4-20). Additionally, there were 118 common DEGs in both bacterial infections when compared to uninfected tissue. IAV infection had only 79 DEGs when compared to the uninfected group.

![Figure 4-20](image)

**Figure 4-20**: Differentially expressed genes. (A) Heatmap of DEG counts by comparison among the groups; a DESeq2 data set with an adjusted p-value cutoff of 0.1 for classification as differentially expressed. (B) Venn diagram presenting the DEGs in all comparisons.

### 4.6.3. Hierarchical clustering of samples based on DEGs

The top 75 strongly DEGs are shown in (Figure 4-21). Based on their expression levels, hierarchical clustering was performed with the complete (maximum) linkage method and using Euclidean distance measures. The relationships among the samples clearly distinguished uninfected, IAV, BCG, and *P. aeruginosa* infections. There are two major clusters separating the uninfected and IAV infected HLTEs from the two bacterial infections. Further, samples of each group are clustered together, showing the similarity among them. Both up and down-regulation was observed, with greatest differences observed in *P. aeruginosa* infection. These DEGs consist of genes related to different cell functions such as immune response, cell adhesion, apoptosis, and cell proliferation.
Figure 4-21: Heatmap of DEGs with a conjoining hierarchical clustering dendrogram. Heatmap showing mean expression levels of top 75 DEGs across all transcriptomic samples. Dendrogram representing the (dis)-similarity among the samples; the hierarchical clustering was implemented with the Euclidean distance measure.
4.6.4. Gene set enrichment analysis (GSEA)

To further look into effects by infections, we performed GSEA using the Hallmark gene sets from MSigDB. All differentially regulated genes were used to annotate them against the Hallmark gene sets. Hallmark gene sets recapitulate and represent specific well-defined biological states or processes and display their coherent expression. It contains 50 gene sets, presenting a broader picture of cellular processes with relevant information. The functional profiles of both bacterial infections were similar, but differed greatly from IAV infection (Figure 4-22). In IAV infection, immune response pathways were upregulated, whereas in bacterial infections, in addition to immune response, several cell cycle-related pathways were downregulated.

**Figure 4-22:** Hallmark pathways: DEGs from each comparison (IAV, BCG and P. aeruginosa) were annotated against the Hallmark gene sets. The normalized enrichment score (NES) is calculated by ranking gene expression difference and normalizing for size of each gene set.
4.6.5. DEGs in IAV infection

Differential expression analyses via DESeq2 identified 65 DEGs with FDR < 0.05 (log2foldchange:1.2 - 6.7). All of these identified DEGs were upregulated in IAV infection. To determine the high confidence markers, AUCs of all these DEGs were calculated, which revealed high discriminatory ability (AUC: 0.92 to 1.0) (Figure 4-23).

Host immune related genes increased in IAV infected HLTE: A signature of immune response was clearly evident in these DEGs. Genes involved in the inflammatory response like CXCL9 and RTP4 were upregulated. Several genes involved in IFN-α and -γ signaling pathways were also highly upregulated (IFI44L, GBP4, MX1, EPST11, PARP9, RTP4, WARS, LAP3, USP18, ISG15, SAMD9L, UBE2L6, IFI35, PSMB9, GBP2, PSMB8, IFI44, HELZ2, PARP14, and CMPK2).

Figure 4-23: DEGs in IAV infection.: The X-axis presents log2 fold change, the Y-axis presents adjusted p-values (-log10) in differential expression of genes. Triangles indicate mRNA and round shapes present IncRNA. The color indicates AUC of each DEG (n=5).
**4.6.6. DEGs in BCG infection**

A pairwise comparison between uninfected and BCG-infected tissue identified 91 differentially regulated genes with FDR < 0.05. Expression of these DEGs was either upregulated or downregulated in the range of 6.7 to -3.2 log2 fold change (Figure 4-24). Among these DEGs, there were also lncRNA.

**Immune response to infection:** Several genes related to immune responses were induced following BCG infection. For instance, genes regulated by NF-kB in response to TNF-α (KYNU, DUSP4 and NR4A2) were upregulated. In addition, genes involved in the inflammatory response (EREG, PTAFR, IL10, LTA, CXCL9 and OSM) were also upregulated.

![Figure 4-24: DEGs in BCG infection. X-axis presents log2 fold change, Y-axis presents p adjusted values (-log10) in differential expression of genes. Triangle symbols indicate mRNA class and round symbols presents lncRNA class of RNA. The color scale shows AUC of each DEG (n=5).](image-url)
4.6.7. DEGs in *P. aeruginosa* infection

Differences in gene expression following infection with *P. aeruginosa* were analyzed by comparing *P. aeruginosa* infected samples with uninfected controls. 450 genes were identified as being differentially expressed with P-adjusted values < 0.05. 163 were upregulated and 287 downregulated (Figure 4-25).

**Immune response related genes:** Several genes involved in host immune response related pathways were upregulated. Inflammatory genes like EREG, PTAFR, CXCL6, CCR7, PDNP, GNA15, OSM, SLC1A2, ITGB8, SLC11A2, TNFAIP6, BEST1, NDP, and IL7R were upregulated.

![Figure 4-25: DEGs in *P. aeruginosa* infection. The plot presents the DEGs in *P. aeruginosa* infection. The X-axis presents log2 fold change, the Y-axis presents p adjusted values (-log10) in differential expression of genes. Triangle shape points indicate mRNA class and round shape presents lncRNA class of RNA. Color scale shows AUC of each DEG (n=5).](image-url)
4.6.8. **Antiviral immune responses**

In response to IAV infection, several viral restriction factors such as *MX1*, *MX2*, *SAMD9*, and *SAMHD1* as well as additional factors that are stimulated by type I and II IFNs were overexpressed. Top 25 DEGs in viral infection as compared to uninfected and bacterial infections are shown (Figure 4-26).

![Expression level graphs showing normalized number of transcripts (log) in Y-axis of each infection group (X-axis).](image)

*Figure 4-26:* Antiviral DEGs: The column plots show the levels of normalized number of transcripts (log) in Y-axis of each infection group (X-axis). (n=5, ± SEM, Wilcoxon–Mann–Whitney test).
4.6.9. Regulatory cytokines
Several pulmonary cell types work together to control pulmonary immune responses in order to avoid tissue injury and maintain tissue resilience. Anti-inflammatory cytokines restrict excessive inflammation, originally against pathogens, though some cytokines have dual function. Bacterial infections in this setup induced expression of several anti-inflammatory cytokines within 24 hpi (Figure 4-27).

![Figure 4-27: Regulatory cytokines: The column plots show the levels of normalized number of transcripts (log) in Y-axis of each infection group (X-axis) (n=5, ± SEM, Wilcoxon–Mann–Whitney test).](image-url)
4.7. Single cell transcriptomics

To study the complexity of IAV infection in lungs at the cellular level, we prepared single-cell suspensions of HLTEs for IAV infected and uninfected tissues, separately.

4.7.1. Live/dead staining

Apoptosis and cell death of single cell suspensions was determined using Annexin V (FITC) and PI. Duplets were excluded based on size. There were roughly 10% stained with PI, and 38% exhibited apoptosis (Figure 4-27). Dead cells were excluded from the cell suspension. Live cells (including cells undergoing apoptosis) were used for single cell transcriptomics.

Following the 10x genomics run, library preparation and sequencing, the retrieved cell number was 2964 for uninfected, and 2220 for infected samples. FASTAQ files obtained from NextSeq550 were subjected to Cell Ranger 3.1 analysis for counts and alignment. Aligned counts were used for downstream analysis using the Seurat v3.0.
4.7.2. Quality control
To learn about the quality of reads for pre-filtering and further analysis, we plotted unique features/cell, number of readouts/cell, and percentage of mitochondrial genomes/cell. Very lowly expressed unique genes can indicate empty droplets, while very high values can indicate droplets with two or more cells. So we filtered out cells that have unique feature counts over 2,500 or less than 200. Higher mitochondrial genome values can indicate low-quality or dying cells. We therefore filter out cells with >5% mitochondrial counts (Figure 4-29).

Figure 4-29: Data quality from single cell RNA-seq experiments. (A) Violin plots showing number of unique features per cells. (B) Number of genes identified in each cells. (C) Mitochondrial RNA content percentage in each cell.
4.7.3. Clustering and visualization

Seurat v3 determines the ‘dimensionality’ of the dataset using their PCA scores which is maintained while clustering the cells. It uses a graph-based clustering approach to identify transcriptionally distinct clusters among cells. Cell clusters based on the transcriptional profiles were visualized in 2D plots using the Uniform Manifold Approximation and Projection (UMAP) method. These algorithms place similar cells together on dimension reduction plots. Unsupervised clustering identified 13 clusters at a resolution of 0.5. These clusters were annotated to cell identities manually based on canonical marker genes and information from previously published scRNA articles on human lung cells (Figure 4-31 and Table 4.2). Among the different cell populations, lymphocytes and AET2 constituted the major portion of sequenced cells. In addition, other lung cell types identified were macrophages, AET1, basophils, fibroblasts, neutrophils, ciliated cells and endothelial cells (Figure 4-30).

![Figure 4-30: Integrated single-cell RNA-Seq analysis identifies diverse lung cell populations. Single-cell RNA-Seq was performed on single-cell suspensions generated from HLTEs (24 hpi uninfected and IAV infected) from a explant donor suffering with emphysema. Samples were analyzed using PCA score analysis within the Seurat R package. Cells were clustered using a graph-based shared nearest neighbor clustering approach and visualized using the Uniform Manifold Approximation and Projection (UMAP) method. The graph shows the identified cell populations.](image-url)
**Table 4.2: List of canonical cell markers used to label cell clusters**

<table>
<thead>
<tr>
<th>Cell types</th>
<th>Markers</th>
</tr>
</thead>
<tbody>
<tr>
<td>CD8 T cells</td>
<td>CD8A, CD3E, PTPRC</td>
</tr>
<tr>
<td>CD4 T cells</td>
<td>CD40LG, CD3E, PTPRC</td>
</tr>
<tr>
<td>AET2</td>
<td>SFTPD, SFTPC, EPCAM</td>
</tr>
<tr>
<td>NK cells</td>
<td>NKG7, NCR1</td>
</tr>
<tr>
<td>AET1</td>
<td>AGER</td>
</tr>
<tr>
<td>Fibroblast</td>
<td>COL1A1, ACTA2, DCN</td>
</tr>
<tr>
<td>Innate lymphoid cells</td>
<td>PTPRC, CD3E</td>
</tr>
<tr>
<td>Macrophages</td>
<td>MRC1, MARCO</td>
</tr>
<tr>
<td>Basophils/mast cells</td>
<td>TPSAB1, TPSB2</td>
</tr>
<tr>
<td>B cells</td>
<td>CD79A, CD24, MS4A1, CD19</td>
</tr>
<tr>
<td>Ciliated</td>
<td>CCDC78, TUBB1, TPPP3 FOXJ1</td>
</tr>
<tr>
<td>Endothelial cells</td>
<td>CDH5, CLDN5</td>
</tr>
<tr>
<td>Neutrophils</td>
<td>FCGR3B, S100A8, CSF3R</td>
</tr>
</tbody>
</table>
Figure 4-31: Canonical markers to identify the cell type: Cell populations were identified manually by canonical cell markers to label clusters. Cell types were classified as immune, endothelial, or epithelial as indicated in the legend. More than one marker was used for identification of cell type, as in Table 4.2.
4.7.4. Differentially expressed genes

To study the response of different cell types to IAV infection, we compared the transcription profiles of each cell type in a pairwise manner using the DESeq2 test. Macrophages and AET2 contained the highest number of DEGs (Figure 4-32). Interestingly, all DEGs in macrophages following the infection were upregulated, whereas a small number of downregulated DEGs were detected in AET2.

Figure 4-32: DEGs following IAV infection in different cell types. (A) Macrophages and (B) AET2. The X-axis presents log2 fold change and the Y-axis presents p adjusted values (-log10) in differential expression of genes.
4.7.5. Immune-related DEGs

Most of the highly DEGs were related to immune responses. In macrophages, CXCL10 and MX2 are among the top ten upregulated genes in IAV infection. AET2 also showed a strong response to IAV infection, and most DEGs were related to immune responses. A very mild lymphocyte response to IAV infection was observed. CXCL10 is the most significantly upregulated gene in macrophages following infection, but it remained unchanged in the other cell types. ISG15, IFI6 and MX1 were upregulated in most cell types following the infection (Figure 4-33).

![Figure 4-33: DEGs response to IAV infection. Cell types identified in lung tissue (X-axis) respond differently to IAV infection. The violin plots show the expression level of the top DEGs in IAV infection.](image_url)
4.7.6. Gene set enrichment analysis

The GSEA effectively summarizes and represents refined and concise specific biological states or processes and displays coherent expression. It contains 50 gene sets generated by a computational methodology based on original founder sets available on Molecular Signatures Database (MSigDB). Results across all cell types show regulation of different processes following IAV infection. Inflammatory immune response to infections related processes were mainly upregulated in macrophages, whereas cell proliferation related processes are downregulated in macrophages as well as in AET2 (Figure 4-34). Several metabolic processes such as oxidative phosphorylation and xenobiotic metabolism are also regulated in several cell types.

![MSigDB Hallmark GSEA Results](image)

**Figure 4-34:** GSEA in different cell types following infection. DEGs from each cell type were ranked based on their P-adjusted values and annotated against the GSEA Hallmark gene sets. Pathways were not identified in cell types with very low response to infection. The Y-axis shows different cell processes and the X-axis identifies the cell types.
4.7.7. Viral transcripts content in cells

All annotated cell types were checked for the presence of viral genes in the cells from infected tissue as well as uninfected controls. All cell types, except T cells, NK cells, basophils and neutrophils, returned with some number of cells containing the viral genes. Overall viral gene NS was detected in the highest number of cells across all cell types. A strong expression of all viral genes was detected in innate lymphoid cells and fibroblasts. The prevalence of infection among different cell types of the lung has been described previously [175].

Figure 4-35: Percentage of lung cell types containing viral transcripts: combined single cell analysis shows the presence of viral genes in several lung cell types. The percentage of infected in each cell type (Y-axis) is reported for cells derived from either the control (bottom cluster) or the IAV-treated tissue (upper cluster). The X-axis presents the viral genes.
4.8. Itaconic acid (IA) as an anti-inflammatory drug

Anti-inflammatory effects of IA in viral infection were previously reported in cell lines [90]. To test whether this would also apply at the tissue level, we used the HLTE model to study the anti-inflammatory effects of IA in IAV infections. Anti-inflammatory and immunosuppressive drugs may enhance the virus number, and we, therefore, quantified the virus particles in supernatants. HLTE tissues were untreated or treated with IA or di-methyl-IA (a chemically modified form of IA). After an overnight washing step, tissues were infected with IAV for 24 h. Virus titer in supernatants was quantified by focus forming unit assay (Figure 4-36 A). There were no significant differences in virus particles in supernatants. CXCL10 mRNA and the corresponding protein (IP10) were quantified to measure inflammation. Significant decreases in levels of both IP10 were observed after IA and DI treatments (Figure 4-36 B,C).

Figure 4-36: Anti-inflammatory effect of IA: HLTE tissue were pre-treated with IA or DI overnight and infected with IAV for 24 h. (A) Virus titer quantified by focus forming unit assay. (B) CXCL10 mRNA was quantified from RNA extracted from HLTE tissues using qPCR. (C) Released IP10 protein was quantified in supernatants of cultured HLTEs. (n=6, ± SEM, Wilcoxon–Mann–Whitney test).
5. Discussion

5.1. Establishment of virus infection

Influenza virus targets mainly AECs of the respiratory tract, though it has also been reported that other cell types, such as endothelial cells, NK cells, macrophages, and DCs, are also susceptible to this pathogen [175]. AECs are the main site for productive viral replication [186]. Whereas productive viral infection in macrophages is still controversial, only a subset of highly pathogenic avian H5N1 influenza virus strains productively replicate in macrophages [187]. Our results show a continuous decrease in released virus titer, lowest at last time point observed 72 hours. Considering the complexity of IAV infection pattern for each cell type, most cells were not susceptible to infection or have abortive replication of virus. The clearance of virus could be due to immune response by other cells. A previous study showed a decrease in virus due to apoptosis-dependent phagocytosis of virus-infected cells in a cell co-culture model with HeLa cells with macrophages [188]. In addition, technical replicates also responded differently in production of virus, which can be explained due to differences in diverse cell populations in each piece of tissue. Kinetics in levels of viral transcripts was different than viral load in supernatants, highest viral mRNA observed at 48 hpi. Most of IAV can enter into the nucleus of macrophages, but the majority fails to productively replicate, which results in late protein synthesis termed as defective replication [187, 189]. In another in vitro study, NK cells were shown infected with IAV, but viral mRNA load in cells decreased after 6 hpi, leading to incomplete replication of IAV [190].

In addition to viral load, the immune response by cells is another measure to study the efficiency of infection of any pathogen. We checked two cytokines related to innate immune response, namely CXCL10 and ISG15 mRNA, by qPCR. CXCL10 is a pro-inflammatory cytokine, whereas ISG15 mediates antiviral cytokine responses induced during IAV infection [191]. IAV infection induced the expression of both cytokines at all time points studied (8, 24, 48 and 72 hpi), showing the response of tissue to infection. This activity confirmed the usage of HLTe model for studying viral infection. During establishment of the protocol, initially we preserved the tissues in RNAlater for RNA extraction, but the quality of RNA was variable, from best to worst. To overcome this problem, we switched to another preservation method, flash freezing in liquid nitrogen. After adopting this method, RNA quality remain consistently good. Flash freezing immediately halts all biological activities, including RNase, whereas RNAlater might be unable to fix the enzymatic activities inside a bigger tissue.

Viability of tissue during in vitro culturing was assessed by quantifying released LDH. Tissues were cut with scissors, making it an open wound. We observed an increase in LDH levels in the first 24 h, and then a very mild increase until 72 h. Infection had also negative effects on viability of tissue. Most cells remained viable. Probably the cell death observed was due to incision, rather than culturing.

5.2. Establishment of viral and bacterial infection

Based on outcome of viral infection in HLTe model, we proposed to extend the number of pathogens by including two bacteria, *P. aeruginosa* and BCG, in the study. With an aim to study the early host response, we measured the cytokines in supernatants and their expression at the mRNA level at 24 hpi. A panel of five cytokines, IL-6, IL-10 and IL-1β as regulatory cytokines and IP-10 and IFNγ [192], reflecting
a pro-inflammatory host response, was established to study the differential response to viral and bacterial pathogens. Viral infection induced highest expression of inflammatory markers, IP10 and IFNγ, in comparison to the mild response seen in the bacterial infections. Induction of type I IFN-stimulated gene CXCL10 following IAV infection is a conserved and indispensable antiviral response already reported in mice and humans [193]. Type I IFNs released due to IAV infection inhibits IL-1β production [194]. We observed no change in IL1β in viral but an enormous production in bacterial infections, as IL-1β signaling is required for bacterial host defense [81, 194]. IL-1β activates the IL-17 pathway and helps the host in improved bacterial clearance. A bacterial co-infection following IAV infection is more lethal as virus attenuates the IL-1β production [81, 194]. IL-10 and IL-6 cytokines are powerful immune mediators with versatile functions [195, 196]. IL-10 has a major protective role against inflammation [195]. In a mouse study, it is shown that IL-10 during bacterial infection controls the expression of pro-inflammatory cytokines and the infiltration of neutrophils into the lungs [197]. IL-6 is known to have a differential role in lung inflammation, which varies with the stimulus used [198]. In a study conducted on the German CAPNETZ cohort, it is shown that people with severe CAP had high IL-6 and IL-10 levels [199]. IL-6 is a more sensitive biomarker of localized infection, thus having an advantage over procalcitonin in CAP diagnosis [200].

Taken together, we noted the expected differential profile of cytokine release in viral versus bacterial infections. Inflammatory cytokines are associated more with viral infection, whereas bacterial infections induced regulatory cytokines. In addition to released proteins, a similar signature was also seen at mRNA expression level of these cytokines. This shows the possible usage of the HLTE model for functional infection studies.

### 5.3. Evaluation of candidate biomarkers

Due to the complex etiology of pneumonia, the inflammatory response is complex and poorly controlled, making the exact diagnosis challenging. Several biomarkers for diagnosis and prognosis of pneumonia have been identified in animal models, cell culture models, or cohort studies. Our research had identified four candidate biomarkers (ACOD1, A20, ORM2 and PROK2) in different studies. With an aim to evaluate their diagnostic potential, we evaluated their differential expression in viral and both bacterial infections. The ACOD1 gene had been identified as one of the most highly upregulated genes following IAV infection in mice, but only a slight change was observed in a cell culture model. The significance of ACOD1 in other Gram-positive and -negative bacterial infections was studied by our research group. High induction in expression of ACOD1, and its downstream gene A20, was reported only in Gram-negative bacteria (P. aeruginosa). To confirm these findings from IAV and P. aeruginosa infections, we studied the induction in expression of ACOD1 and A20 following infections in HLTE model. IAV showed no significant changes in ACOD1 expression, contrary to finding in mouse model, but supporting the finding of cell culture model. In addition, P. aeruginosa infection (triggering TLR4 activation) highly induced the expression levels of ACOD1 and A20, supporting the findings in cell culture model. Induction of ACOD1 following infection might be due anti-inflammatory properties of to its downstream metabolite itaconic acid [201].

ORM2, an acute phase protein, was another feature identified by our research group as one of the most upregulated gene in the mice infected with IAV. To our knowledge, role of ORM2 in infections is not well explained yet, though its modulatory effect in inflammation has been shown [202]. Its protective role
against different viral and bacterial infections and potential as biomarker has been described in several animal models [93, 202-204]. In contrary to these reports and our mouse model finding, we did not observe any significant changes in ORM2 levels (protein and mRNA) in response to any of the infection studied.

PROK2, also known as BV8, was reported as one of highest responses to acute inflammation by monosodium urate crystals in the mouse air pouch model [102]. Bacterial products, especially LPS, induce PRKO2 overexpression to promote IL-1 signaling to modulate inflammation [70]. It was previously shown that PROK2 induces expansion of myeloid-derived suppressor cells in secondary pneumococcal pneumonia in mice [205]. Our results showed higher concentrations of PROK2 proteins in HLTE supernatants following both bacterial infections. Differently from bacterial infections, IAV infection did no change in levels of PROK2. These initial results in HLTE model support previous findings in other setups and require further confirmation. The HLTE model presents only local and immediate responses to stimuli. Confirmation of the findings in ongoing studies is mandatory before they can be applied in the development of novel diagnostic tools.

5.4. Differential metabolic reprogramming

Several metabolites play a role during inflammatory processes, their quantification in a variety of body fluids is now extensively used as a tool for biomarker discovery [110]. We applied a targeted MS-based assay to detect up to 188 metabolites per sample. These metabolites were from different classes, we were able to detect and quantify roughly 70% of metabolites in more than 75% of the samples. Interestingly, the abundance profile of different metabolites in lung tissue agreed with their profile in plasma [206]. In our study, currently we were interested to learn the possible use of the HLTE model to study metabolomics. The biological replicates were very low (n=3), but still we could see a clustering pattern among different infection groups based on their metabolic profiles. Uninfected or IAV-infected tissues were very similar, however, both bacterial infections were different from the uninfected controls. This indicates metabolic reprogramming in tissues during the bacterial infections. Among the different classes of metabolites, biogenic amines showed major changes in their levels following the infections.

A significant increase in the Kyn/Trp ratio was observed in P. aeruginosa infection, showing its discriminatory potential for distinguishing P. aeruginosa infection from Controls. Macrophages, dendritic cells and lung tissues produce the enzyme indoleamine 2,3-dioxygenase (IDO), which breaks down tryptophan into kynurenine [207]. Quorum sensing by P. aeruginosa also generates kynurenine formidase involved in the catabolism of tryptophan into kynurenine [43]. A study conducted by our research group has shown an increase Kyn/Trp in CAP patients [184]. Putrescine levels in P. aeruginosa were also strongly upregulated. Previous studies have highlighted the role of increased levels of TGF-β in enhancing decarboxylation of arginine, hence producing higher amounts of putrescine in macrophages [208].

5.5. Differential reprogramming of sncRNA

In this chapter, integrative small RNAseq analysis of HLTE samples was applied to identify a plethora of novel candidate biomarkers for lung infections. From all studied infection groups, we compared the
relative biomarker potential of the four major sncRNA classes, i.e. miRNA, piRNA, snoRNA and snRNA. Among these classes, the miRNA population was the most abundant, followed by piRNA and snoRNA. The differential expression profile of all infections shows 2 major clusters, uninfected and IAV-infected in one group and bacterial infections in second group. After applying statistics, no changes in sncRNA levels could be detected after IAV infection. Possible explanations for these negative findings are that the stimulus (infection dose) to induce regulatory sncRNA in the HLTE model was insufficient, that there was not enough incubation time to cause a measurable effect, or that the infection was too localized. miRNAs are not only produced at the site of infection, but they are also expressed in peripheral cells in response to infection. Regulatory miRNAs (e.g., miR-223) are expressed in neutrophils and transfer from neutrophils to pulmonary epithelial cells to dampen acute lung injury [209]. This indicates that miRNAs may perform this function in different cell than their origin cells.

*P. aeruginosa* infection induced the greatest expression changes of sncRNA among all infection groups. Among these DE sncRNA, several were piRNA and snoRNA, showing their possible use as disease biomarkers. miRNA targets mining using databases and available literature, predicts genes involved in various cellular processes. Clustered microRNAs miR-221-3p/miR-222-3p were significantly downregulated in *P. aeruginosa* infection. hsa-miR-221-3p targets the SOCS3 gene, which suppresses cytokine signaling in tuberculosis [210], thus helping to control the cytokines response. The miR-132/-212 family were highly upregulated in *P. aeruginosa* infection. miR-132/-212 function of miR-132/-212 in TLR2/TLR5 ligand–induced tolerance, which operates as a negative regulatory feedback mechanism to prevent uncontrolled inflammatory reaction potentially [211]. miRNA-362-5P, which enhances the production of IFN-γ, was downregulated in *P. aeruginosa* infection [212]. We observed a downregulation in miR-148 family in *P. aeruginosa* infection, in contrary to report by Chu et al. who showed a significant upregulation of miR-148 following LPS in marine fish miuy croaker. In addition, they showed that miR-148 overexpression suppresses the production of IL-1β and IL-6, whereas we observed higher production of these cytokines [213]. A possible explanation can be the observations in two different species and infection models. In short, microRNAs have emerged as important controllers involved in regulating the inflammatory response, though the mechanisms are not completely explained yet.

Three miRNAs (miR-194-3p, miR-181a-3p and miR-221-3p) and piRNA (piR-002410) were differentially regulated following the *P. aeruginosa* infection, with an AUC of 1.0 compared to uninfected controls. These molecules need further evaluation in different infection models and cohorts to study their potential as biomarker. To this end, we present HLTE an alternative model to study the function and biomarker potential of sncRNA in infection studies. A distinct expression profile of sncRNA for each infection group was shown in this model.

5.6. Transcriptomic response to infections

Host gene expression signatures derived from peripheral blood or site of infection have been reported for several diseases including bacterial and viral infections [214, 215]. Host bulk RNA transcriptome helps in understanding pathogenic processes and the immediate local response at the site of infection. We applied next-generation sequencing on bulk RNA from HLTE tissues following the infections, and we observed equal distribution of read counts across all the samples, irrespective of treatment or donor.
Distance clustering of differentially expressed genes returned with two major groups, 1) uninfected and IAV-infected and 2) both bacterial infections. A closer look at DEGs showed that IAV induced genes involved IFN type I and type II signaling. Both bacterial infections produced a similar expression pattern, as there was no significant DEGs when both infections were compared. In comparison to uninfected lung tissue, *P. aeruginosa* infection induced the highest number of DEGs; they were involved in several pathways such as immune response, cell cycle, and metabolic pathways.

The transcriptome profile of HLTE from IAV-infected lungs showed all the DEGs were upregulated. These responses were either by primary stimuli (IAV-infection) or by secondary (virus-induced IFN) responses to infection [216]. Among these upregulated genes, several play their role by restricting the replication of IAV. *MX1* and *MX2* restrict the virus replication post entry by interfering with functional viral complex assembly [217]. Several molecules involved in the IFN-induced ISG15 pathway were also highly regulated. In addition to mRNA, several long non-coding RNAs (lncRNAs) were also differentially upregulated. Our knowledge about role of host lncRNAs in viral infections is very limited, especially their function in host defense or virus replication. Some studies conducted on differentially induced lncRNAs in viral infections show their connection in general host immune response to viral infection. These identified lncRNA act by negative feedback to control innate immunity, which are regulated by the IFN pathway and would either benefit the host by blocking viral infection or assist the virus by subverting host immunity [218, 219]. In our data, AC005515.1 and AC093063.2 are significantly upregulated in IAV infection, and interestingly both are located on chromosome number 19 and share a sequence of 1004 bases. We observed LGALS17A and AC005515.1 among the top 5 upregulated transcripts in IAV infection. LGALS17A sequence is upstream of AC005515.1. LGALS17A (Galectin 14 Pseudogene) is a Pseudogene. One recent study suggests LGALS17A as a novel ISG [220]. Wang et al. identified a group of human lncRNAs that modulate IAV replication in a loss-of-function screen study. They evaluated particularly one IFN-independent lncRNA, called IPAN, which was specifically induced by IAV to assist IAV replication. It stabilizes the viral RNA dependent RNA polymerase PB1, enabling efficient viral RNA synthesis [221].

The greatest changes in transcriptome profile were observed in *P. aeruginosa* infection. To understand this massive response to *P. aeruginosa*, we must consider a large variety and complex virulence systems in *P. aeruginosa* [222]. Genes involved in the immune response to infection were upregulated, and cell cycle pathways were downregulated. Several genes involved in making the 26S proteasome, a tool for protein degradation, were also upregulated. We observed high levels of kynurenine in *P. aeruginosa* infected tissue, and at mRNA level we observed induction in expression level of KYNU, an enzyme that catalyzes the cleavage of L-kynurenine. The high expression of the IL10 family members at 24 hpi illustrates how host cells prepare themselves for immune suppression to control tissue damage. The complement system is also an important effector mechanism of the innate immune system for clearance of bacterial pathogens [223]. A cluster of DEGs which regulates the complement system activation were significantly downregulated. In addition to studying the various responses to infection, transcriptomic data can help us to discover novel molecules associated with a pathological condition. For example, the *VNN1* gene was the most highly induced gene in both bacterial infections, and slightly in IAV infection. Vascular non-inflammatory molecule-1 (abbreviated Vanin-1, VNN1) plays a major role in biosynthesis of CoenzymeA (CoA) [224]. In recent years, its importance in the context of oxidative stress and inflammation has been elaborated [224]. One study shows high induction of VNN1 gene in A549 cells in
response to inflammatory cytokines, and weak induction in response to IAV infection [225]. To our best of knowledge, no explanatory study has been yet conducted in relation to its role in bacterial infections. In another study, silencing of VNN1 by miR-203 miRNA in mice with septic shock enhanced the lung injury [226]. This indicates the possible role of VNN1 in tissue homeostasis.

Our work has research, clinical, and diagnostic value by making it possible to effectively associate observations in bulk transcriptomics data to specific immune subsets. One limitation we observed is the lower sensitivity of RNA seq, possibly due to limited sequencing depth (50k). One example is that CXCL10 upregulation in IAV-infected samples was observed by RT-qPCR, but not by RNAseq in the same samples. Nonetheless, we did see CXCL10 upregulation in the single cell RNAseq experiments, where sequencing depth was much higher (75k).

5.7. Transcriptomic profiling at single cell level

To investigate the host response of different lung cells to IAV infection, we applied the single cell transcriptomics technique in the HLTE model. To our knowledge, this is first time that human lung transcriptome is profiled at single-cell resolution following in vitro 48 h culturing and 24 h infection. The intensive process of tissue digestion to cell suspension (a process lasting 2 h) did not affect the viability of cells, which allowed us to proceed for RNAseq. We performed the clustering of cells based on variance in their transcriptomes to distinguish between different cell types. Recent studies to develop a human lung atlas suggest 40 discrete lung cell types, though many of them are not fully characterized [227]. Since we have cells from only one donor, diseased lung, and only 3000 cells, we clustered cells at low resolution to identify only major cell types. Identification of different cell types by canonical markers returned all well-known cell types including alveolar type II cells, ciliated and club airway epithelial cells, alveolar macrophages, and lymphocytes. Surprisingly, the lymphocyte population was quite large in proportion to the other cell types, which contradicts recent publication on lung single-cell characterization [228]. One possible explanation could be that it was a COPD; emphysema lung. COPD lung often suffers from chronic inflammation, characterized by an increased number of T cells and macrophages in lung tissue [229].

IAV contains eight segments of the negative-standard RNA genome, which is replicated in the host cell to mRNA and matured with poly(A) [230]. We adopted the strategy to map the sequences against a pool of human and IAV (H1N1) genomes, which helped us to identify intracellular viral genes even in the absence of productive infection [175]. We could identify viral genes in a particular number in immune and epithelial cells from tissues infected with IAV, but not in uninfected controls. This confirms the applicability of our strategy for viral infections and can be established for other viral pathogens. It is well documented that AET2 cells are the main targets of IAV and lead to productive infection [170]. We identified viral genes in different cell types, but cannot confirm productive infection. In several previous studies, IAV is shown to infect all cell types, but most usually only support an abortive infection [170, 171, 175, 231]. This method of identification of infected cells is simple and straightforward, but it has the drawback that it cannot distinguish between intracellular viral transcription and viral particles or RNA that are bound to the cell exterior or have been taken up. Interestingly, we observed that expression levels of different viral genes vary within cells. The viral segment NS was detected in the highest number of cells, whereas viral gene PB1, PB2, and P. aeruginosa were detected in lower numbers. This likely
reflects variation in virus infection and assembly in different host cells and host cell types. Besides, higher viral RNA content within a cell does not mean a “productive infection”. As the process of virus assembly is restricted by the host cell at different levels, a single infecting virion in a cell is mostly failed to produce at least one essential protein for its final assembly[232, 233].

Interestingly, innate lymphoid cells (ILCs) contained high levels of viral genes, but no induction of pro-inflammatory genes. ILCs play vital roles in tissue homeostasis during influenza infection [234, 235]. The absence of expression of antiviral genes may have provided the viruses a better environment for their genome replication. We also observed the same situation in fibroblasts, i.e. high viral gene content but no expression of pro-inflammatory and anti-viral genes. On the other hand, macrophages and AET2 responded greatly to IAV infection at the gene expression level. CXCL10, a pro-inflammatory cytokine mRNA, was highly induced in only macrophages, with no change in other cell types. This illustrates how lung macrophages immediately turn to an inflammatory state and activate neighboring cells to act against invading pathogens. The differential gene expression pattern observed in bulk RNA transcriptome (i.e. transcriptomes obtained from complete HLTE pieces) is very similar to the single-cell transcriptome data. The major feature observed is the upregulation of type I and II IFN signaling genes. A detailed comparative analysis between bulk RNA and single-cell RNA data can further help to understand the complexity of infection. In addition, one can explain the differential response of cells in an infected environment by differentiating the cells from infected tissue into cells expressing viral genes (infected cells) and those without viral genes (bystander cells).

5.8. Itaconate as an anti-inflammatory drug

As in previous sections we have observed that IAV infection in host cells induces a severe inflammatory response to restrict the replication of the virus. A major complication in influenza is a severe induction of immune cells and pro-inflammatory cytokines, a condition termed “cytokine storm” [51]. This hyperinflammation led to severe tissue damage is more lethal than the viral pathogenesis itself [236]. To overcome the rush of cytokines, adjunct immunomodulatory therapy is anticipated to be beneficial [51, 237]. However, these therapies modulate immune signaling pathways that restrict virus replication and may thus also lead to high viral load in host cells [238]. The role of itaconic acid as the anti-inflammatory compound has recently been extensively elaborated, with its possible use in clinics [239]. We have checked the potential of itaconic acid as potential as an adjunct therapy in influenza infection using the HLTE model. CXCL10 (IP10) is associated with inflammation, with multifunctional features in infectious diseases [240]. We observed a high induction of CXCL10 in lung tissue following IAV infection. Using CXCL10 as an inflammatory marker, we tested itaconic acid as an anti-inflammatory compound in HLTE infected with IAV. The results showed a major decrease in the expression of CXCL10 in infected tissue, reaching similar levels as in uninfected tissue. This supports our hypothesis and previous findings in a cell culture model. In addition, viral replication did not increase upon itaconate treatment, suggesting its possible use as immunomodulatory therapy.

Single-cell sequencing was performed only on one donor (pool of RNA from 8-10 HLTE pieces each for infection and control) due to the high cost of the experiment. Though recent studies support our findings [175], additional experiments are needed to confirm these preliminary results.
6. Conclusion
The HLTE model is a suitable alternative to animal models in lung infection research and is cost-effective. It has advantages over classic cell culture and may act as a bridge between basic research and clinical validation. Being the site of infection, it can help to study the immediate response to infectious and noninfectious stimuli. It can provide a platform to test candidate drugs and biomarkers, as well as screening for new candidate biomarkers. Nevertheless, this model contains some limitations such as variability due to differences in underlying disease, unavailability of tissue, and cell death due to aging.

We observed a differential host response of viral vs. bacterial infections, and between bacteria from two different phyla. The outcome of candidate biomarkers and the potential of IA as an immunomodulatory substance need further validations in larger cohorts before their use in clinics. Detection of numerous metabolites in tissue opens up opportunities to characterize their role in different host cellular functions during infections. Metabolic profiling in combination with transcriptomic information can enhance our knowledge about different biological functions. Since HLTEs contain all lung cells in their actual anatomical location and maintain cell-cell interactions, we can characterize cell responses to an infection and/or treatment at single-cell resolution.
82

References


Abbreviation
Aconitate decarboxylase 1 ACOD1
Acute respiratory distress syndrome ARDS
Acute respiratory tract infections ARIs
Airway Epithelial cells AECs
Alveolar macrophages AMs
Cerebrospinal fluid CSF
Chronic obstructive pulmonary disease COPD
C-Reactive protein CRP
Dendritic cells DCs
Differentially expressed genes DEGs
Dimethyl-itaconic acid DI
Enzyme-linked immunosorbent assay ELISA
Fluorescence-activated cell sorting FACS
Gene Set Enrichment Analysis GSEA
Hours post-infection hpi
Human Lung Tissue Explants HLTEs
Hypoxia-inducible factor 1α HIF-1α
IFN-stimulated genes ISGs
IL-1 receptor antagonist IL-1RA
*Indoleamine-2,3-dioxygenase* IDO
Itaconic acid IA
kelch-like ECH associated protein 1 KEAP1
Lactate dehydrogenase LDH
Limit of detection LOD
Lipopolysaccharides LPS
Long non-coding RNA IncRNA
Lower respiratory tract infections LRTIs
microRNA miRNA
Mycobacterium tuberculosis Mtb
Natural killer cells NKs
Orosomucoid 2 ORM2
<table>
<thead>
<tr>
<th>Term</th>
<th>Abbreviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Piwi-interacting RNA</td>
<td>piRNA</td>
</tr>
<tr>
<td>Precision-cut lung slices</td>
<td>PCLS</td>
</tr>
<tr>
<td>Principal component analysis</td>
<td>PCA</td>
</tr>
<tr>
<td>Prokineticin 2</td>
<td>PROK2</td>
</tr>
<tr>
<td>Quantitative reverse transcription PCR</td>
<td>RT-qPCR</td>
</tr>
<tr>
<td>Receiver operating characteristic</td>
<td>ROC</td>
</tr>
<tr>
<td>Receiver operating characteristic under curve</td>
<td>AUC</td>
</tr>
<tr>
<td>RNA Integrity Number</td>
<td>RIN</td>
</tr>
<tr>
<td>RNA sequencing</td>
<td>RNA-seq</td>
</tr>
<tr>
<td>Signal transducer and activator of transcription 3</td>
<td>STAT3</td>
</tr>
<tr>
<td>Small nuclear RNA</td>
<td>snRNA</td>
</tr>
<tr>
<td>Small nucleolar RNAs</td>
<td>snoRNAs</td>
</tr>
<tr>
<td>Standard error of the mean</td>
<td>SEM</td>
</tr>
<tr>
<td>Stimulator of interferon genes</td>
<td>STING</td>
</tr>
<tr>
<td>Succinate dehydrogenase</td>
<td>SDH</td>
</tr>
<tr>
<td>Surfactant proteins</td>
<td>SP</td>
</tr>
<tr>
<td>Uniform Manifold Approximation and Projection</td>
<td>UMAP</td>
</tr>
<tr>
<td>Upper respiratory tract infections</td>
<td>URIs</td>
</tr>
<tr>
<td>varicella zoster virus</td>
<td>VZV</td>
</tr>
<tr>
<td>World Health Organization</td>
<td>WHO</td>
</tr>
</tbody>
</table>
List of figures

Figure 1-1: Lung cells following viral and bacterial infections ........................................................... 10
Figure 1-2: Human lung cell types .................................................................................................... 14
Figure 1-3: Itaconic acid regulates immune response in cells .......................................................... 17
Figure 3-1 Scheme of HLTEs preparation and infection .................................................................. 30
Figure 4-1: LDH release and virus load in supernatant .................................................................... 37
Figure 4-2: Effect of tissue preservation method on RNA integrity ................................................. 38
Figure 4-3: Quantification of viral and host mRNA transcripts by qPCR ......................................... 39
Figure 4-4: Bacterial contamination ............................................................................................... 39
Figure 4-5: Immune response to infections ..................................................................................... 40
Figure 4-6: Expression of ACOD1 and A20 ................................................................................... 41
Figure 4-7: ORM2 response to infection ......................................................................................... 42
Figure 4-8: PROK2 response to infection ....................................................................................... 42
Figure 4-9: Class wise abundance of metabolites across all tissue samples ................................. 43
Figure 4-10: Metabolite expression profile based on infection groups ........................................... 45
Figure 4-11: Expression of biogenic amines .................................................................................... 46
Figure 4-12: A universal presentation of snRNA reads in all samples ............................................ 47
Figure 4-13: Principal component analysis (PCA) of snRNA populations of (un)infected groups .... 48
Figure 4-14: Differentially expressed snRNA ............................................................................... 49
Figure 4-15: Hierarchical clustering based on snRNA expression ............................................... 50
Figure 4-16: Visualization of differentially expressed snRNA in the three types of infection ......... 51
Figure 4-17: Differentially expressed snRNA profiles in response to the three types of infection ... 53
Figure 4-18: Gene Ontology enrichment terms for putative miRNA targets .................................. 54
Figure 4-19: Read counts distribution and normalization ............................................................... 55
Figure 4-20: Differentially expressed genes .................................................................................. 56
Figure 4-21: Heatmap of DEGs with a conjoining hierarchical clustering dendrogram .................... 57
Figure 4-22: Hallmark pathways .................................................................................................. 58
Figure 4-23: DEGs in IAV infection .............................................................................................. 59
Figure 4-24: DEGs in BCG infection ............................................................................................. 60
Figure 4-25: DEGs in P. aeruginosa infection .................................................................................. 61
Figure 4-26: Antiviral DEGs ......................................................................................................... 62
Figure 4-27: Regulatory cytokines ............................................................................................... 63
Figure 4-28: FACS gating strategies for human lung cells .............................................................. 64
Figure 4-29: Data quality from single cell RNA-seq experiments .................................................. 65
Figure 4-30: Integrated single-cell RNA-Seq analysis identifies diverse lung cell populations ....... 66
Figure 4-31: Canonical markers to identify the cell type ............................................................... 68
Figure 4-32: DEGs following IAV infection in different cell types ................................................ 69
Figure 4-33: DEGs response to IAV infection .............................................................................. 70
Figure 4-34: GSEA in different cell types following infection ....................................................... 71
Figure 4-35: Percentage of lung cell types containing viral transcripts .......................................... 72
Figure 4-36: Anti-inflammatory effect of IA ................................................................................... 73
List of tables

Table 3.1: List of lung tissue donors ............................................................................................................ 23
Table 3.2: List of pathogens ........................................................................................................................ 24
Table 3.3: Table of software and online tools ............................................................................................ 24
Table 3.4: Buffers used ................................................................................................................................ 25
Table 3.5: List of reagents ........................................................................................................................... 26
Table 3.6: List of kits .................................................................................................................................... 27
Table 3.7: List of primers ............................................................................................................................. 28
Table 3.8: List of media ............................................................................................................................... 29
Table 3.9: List of equipments ...................................................................................................................... 29
Table 4.1: List of miRNA differentially expressed in P. aeruginosa infection along with their predicted targets by miRWalk ............................................................................................................. 52
Table 4.2: List of canonical cell markers used to label cell clusters ........................................................................ 67
a. **Einverständniserklärung:**


04.02.2020, Hannover

Datum, Ort

Aaqib Sohail

Name und unterschrift
b. **Curriculum Vitae:**

**Personal Data**

**Name:** Aaqib Sohail

**Date of birth:** 30.11.1990

**Place of birth:** Sialkot

**Nationality:** Pakistan

**Education**

- **2008-2012** B.Sc Applied Microbiology, University of Veterinary and Animal Sciences, Lahore, Pakistan.
- **2012-2014** M.Sc Microbiology, Friedrich Schiller University Jena, Jena, Germany.
- **2015-** Ph.D. Dr. rer. Nat., 2020, Hannover Medical School, Hannover, Germany.

**Work Experience**

- Research work at TWINCORE, Centre for Clinical and Experimental Infection Research GmbH, Medizinische Hochschule Hannover, Hannover, Germany: October 2015 to date.
- M.Sc thesis at TWINCORE, Centre for Clinical and Experimental Infection Research Hannover. Research Project: “Pre-analytical and analytical evaluation of S100 proteins as biomarkers for acute bacterial infections”.

**Publications**


**Recent talks and at scientific meeting**

- 3D Tissue Infection Symposium University Würzburg (5-7 April 2019), Titled “Evaluation of a 3D human lung model for functional infection studies”
• 17th Workshop of the Study Group „Immunobiology of Viral Infections” of the Society for Virology (GfV) (26-28 September 2018), Titled “Evaluation of a 3D human lung model for functional infection studies”
• Poster presentations: 5 presentations in the last 2 years.

**Special Note**

In addition to this thesis project I was also actively involved in the following side project:

• Induction of mannitol metabolism in influenza A virus infected cells
• Development of liposomes for the delivery of itaconic acid

04.02.20, Hannover

Datum, Ort

Aaqib Sohail

Name und unterschrift
c. **Erklärungen:**


**TWINCORE, Centre for Clinical and Experimental Infection Research GmbH**

**Arbeitsgruppe:** Biomarker für Infektionskrankheiten  
**Leiter:** PD. Dr. Frank Peßler  
TWINCORE, Centre for Clinical and Experimental Infection Research  
Feodor-Lynen-Str. 07, 30625, Hannover  
**Leitung:** Prof. Dr. Armin Braun  
Research Group “Airway Immunology” Fraunhofer Institute for Toxicology and Experimental Medicine  
Nikolai-Fuchs-Straße 1, 30625, Hannover


04.02.2020, Hannover                                                                             Aaqib Sohail  
**Datum, Ort**                                                                                                                      **Name und unterschrift**
d.  **Erklärungen**

**Zur Verfügbarkeit der Originaldaten**


Datum       Aaqib Sohail
Acknowledgments

I would like to express my special appreciation and thanks to Dr. Frank Peßler for his patient guidance, encouragement, and advice throughout my time as a Ph.D. student. Likewise, past and present members of our research group have been invaluable for practical help, discussions, and constructive criticism.

In particular, I would like to thank Prof. Armin Braun for his valuable time, supervision, and support. I am truly grateful to Dr. Katherina Sewald and Dr. Sabine Wronski for their guidance on the project and very productive discussion on data.

I would like to express my appreciation to Dr. Peter Braubach for supporting this work by providing me with the human lung material, histological work, and guidance on the project. I want to thank other staff of the Lung Research group, especially Ms. Christina Petzold, who helped tirelessly to keep our lung pipeline running. Without their support, this work would not have been possible.

Furthermore, I would like to acknowledge all of our cooperation partners on this project. Prof. Michael Steinert for helping in developing the model. Prof. Stephan Pleschka for providing influenza strain and training on propagation and titration of virus. I am very thankful to Prof. Robert Geffers for conducting RNA sequencing and data analysis. Dr. Oliver Dittrich-Breiholz for helping in designing the single-cell transcriptomics experiments and data analysis. Simon Peter for translating the abstract to the German language. Prof. Mark Brönstrup, Dr. Raimo Franke, and Dr. Haroon Arshad for making the mass spectrometry measurements possible.

This work was supported by the Helmholtz Association Cross Programme Initiative on Individualized Medicine (iMed), and by the Innovative Medicines Initiative Joint Undertaking under grant agreement number [115523], resources of which are composed of financial contribution from the European Union’s Seventh Framework Programme (FP7/2007-2013) and EFPIA companies’ in kind contribution.

Finally, I extend my acknowledgment and heartfelt love to my parents, siblings and wife, for their love, patience, and support while pursuing these studies far away from home.