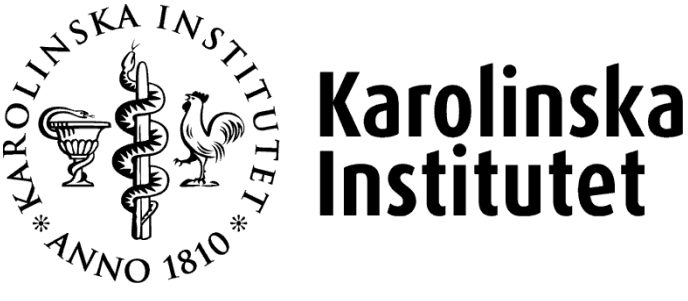
From Laboratory Medicine
Karolinska Institutet, Stockholm, Sweden

MULTI-DIMENSIONAL OMICS APPROACHES TO DISSECT NATURAL IMMUNE CONTROL MECHANISMS ASSOCIATED WITH RNA VIRUS INFECTIONS

Anoop Tholamplackil Ambikan



Stockholm 2023

Published by Karolinska Institutet.
Printed by Universitetsservice US-AB, 2023

ISBN 978-91-8016-991-2
Cover illustration: Drawn by Prasenjit Pal

Multi-dimensional omics approaches to dissect natural immune control mechanisms associated with RNA virus infections.

Thesis for Doctoral Degree (Ph.D.)

By

Anoop Tholamplackil Ambikan

The thesis will be defended in public at Karolinska Institutet, Alfred Nobels allé 8, Seminar Hall 4X, on Friday, April 21st, 2023, at 1:30 PM.

Principal Supervisor:
Associate Professor Ujjwal Neogi
Karolinska Institutet
Department of Laboratory Medicine
Division of Clinical Microbiology

Co-supervisor(s):
Dr. Rui Benfeitas
Karolinska Institutet
Department of Laboratory Medicine
Division of Clinical Microbiology

Docent Piotr Nowak
Karolinska Institutet
Department of Medicine
Division of Infectious Diseases

Opponent:
Professor Evangelia Petsalaki
European Molecular Biology Laboratory-EBI

Examination Board:
Professor Gerald McInerney
Karolinska Institutet
Department of Microbiology, Tumor and Cell Biology

Docent Marianne Jansson
Lund University
Division of Medical Microbiology

Docent Carsten Daub
Karolinska Institutett
Department of Biosciences and Nutrition

I dedicate the thesis to my parents.

Popular science summary of the thesis

Metabolism is a set of processes in our body for creating energy and other biochemical compounds from food. Viruses are parasites that survive on the human body for energy and other genetic materials to replicate themselves. There are many viruses affecting public health and causing mild to severe infections. The study presented in the thesis investigates the changes in the human metabolism that occur because of infections from three viruses, namely, acute respiratory syndrome coronavirus 2 (SARS-CoV2), human immunodeficiency viruses (HIV-1), and Crimean–Congo hemorrhagic fever virus (CCHFV). Many previous studies published in reputed science journals reported changes in human metabolism caused by infections from SARS-CoV2 and HIV-1. The human metabolism comprises many pathways that perform a unique metabolic function, with many reactions in each pathway. A reaction represents an activity in a cell that produces, consumes, or transports one or more metabolic compounds. However, very few (or none) previously published studies have examined metabolism changes in terms of their associated reactions. Also, research on CCHFV infections is still in the infant stage. The study found that the three viruses cause changes in similar metabolic pathways, but changes in their associated reactions are different. The results obtained from my studies can aid novel antiviral strategies targeting the host immunometabolic system. It can thus allow a fruitful route to the computational guiding of experimental antiviral drug discovery and drug repurposing in infectious diseases.

Abstract

In recent decades, global health has been challenged by emerging and re-emerging viruses such as severe acute respiratory syndrome coronavirus 2 (SARS-CoV2), human immunodeficiency viruses (HIV-1), and Crimean–Congo hemorrhagic fever virus (CCHFV). Studies have shown dysregulations in the host metabolic processes against SARS-CoV2 and HIV-1 infections, and the research on CCHFV infection is still in the infant stage. Hence, understanding the host metabolic re-programming on the reaction level in infectious disease has therapeutic importance. The thesis uses systems biology methods to investigate the host metabolic alterations in response to SARS-CoV2, HIV-1, and CCHFV infections.

The three distinct viruses induce distinct effects on human metabolism that, nevertheless, show some commonalities. We have identified alterations in various immune cell types in patients during the infections of the three viruses. Further, differential expression analysis identified that COVID-19 causes disruptions in pathways related to antiviral response and metabolism (fructose mannose metabolism, oxidative phosphorylation (OXPHOS), and pentose phosphate pathway). Up-regulation of OXPHOS and ROS pathways with most changes in OXPHOS complexes I, III, and IV were identified in people living with HIV on treatment ($PLWH_{ART}$). The acute phase of CCHFV infection is found to be linked with OXPHOS, glycolysis, N-glycan biosynthesis, and NOD-like receptor signaling pathways. The dynamic nature of the metabolic process and adaptive immune response in CCHFV-pathogenesis are also observed.

Further, we have identified different metabolic flux in reactions transporting TCA cycle intermediates from the cytosol to mitochondria in COVID-19 patients. Genes such as monocarboxylate transporter (SLC16A6) and nucleoside transporter (SLC29A1) and metabolites such as α-ketoglutarate, succinate, and malate were found to be linked with COVID-19 disease response. Metabolic reactions associated with amino acid, carbohydrate, and energy metabolism pathways and various transporter reactions were observed to be uniquely disrupted in $PLWH_{ART}$ along with increased production of α-ketoglutarate (αKG) and ATP molecules. Changes in essential (leucine and threonine) and non-essential (arginine, alanine, and glutamine) amino acid transport were found to be caused by acute CCHFV infection. The altered flux of reactions involving TCA cycle compounds such as pyruvate, isocitrate, and alpha-ketoglutarate was also observed in CCHFV infection.

The research described in the thesis displayed dysregulations in similar metabolic processes against the three viral Infections. But further downstream analysis unveiled unique alterations in several metabolic reactions specific to each virus in the same metabolic pathways showing the importance of increasing the resolution of knowledge about host metabolism in infectious diseases.

List of scientific papers

I. Ambikan, A. T., Yang, H., Krishnan, S., Svensson Akusjärvi, S., Gupta, S., Lourda, M., Sperk, M., Arif, M., Zhang, C., Nordqvist, H., Ponnan, S. M., Sönnerborg, A., Treutiger, C. J., O'Mahony, L., Mardinoglu, A., Benfeitas, R., & Neogi, U. (2022). Multi-omics personalized network analyses highlight progressive disruption of central metabolism associated with COVID-19 severity. Cell systems, 13(8), 665–681.e4. https://doi.org/10.1016/j.cels.2022.06.006

II. Ambikan, A. T.*, Svensson-Akusjärvi*, S., Krishnan, S., Sperk, M., Nowak, P., Vesterbacka, J., Sönnerborg, A., Benfeitas, R., & Neogi, U. (2022). Genome-scale metabolic models for natural and long-term drug-induced viral control in HIV infection. Life science alliance, 5(9), e202201405. https://doi.org/10.26508/lsa.202201405

III. Neogi, U., Elaldi, N., Appelberg, S., Ambikan, A., Kennedy, E., Dowall, S., Bagci, B. K., Gupta, S., Rodriguez, J. E., Svensson-Akusjärvi, S., Monteil, V., Vegvari, A., Benfeitas, R., Banerjea, A., Weber, F., Hewson, R., & Mirazimi, A. (2022). Multi-omics insights into host-viral response and pathogenesis in Crimean-Congo hemorrhagic fever viruses for novel therapeutic target. eLife, 11, e76071. https://doi.org/10.7554/eLife.76071

IV. Anoop T. Ambikan, Nazif Elaldi, Sara Svensson Akusjärvi, Adil Mardinoglu, Vikas Sood, Ákos Végvári, Rui Benfeitas, Soham Gupta, Ali Mirazimi, Ujjwal Neogi. Transcriptome-based systems-level temporal immune-metabolic profile in Crimean-Congo hemorrhagic fever virus infection (In Manuscript form)

* Equal contributions

List of scientific papers not included in the thesis

1. Svensson Akusjärvi, S., Krishnan, S., **Ambikan, A. T.**, Mikaeloff, F., Munusamy Ponnan, S., Vesterbacka, J., Lourda, M., Nowak, P., Sönnerborg, A., & Neogi, U. (2023). Role of myeloid cells in system-level immunometabolic dysregulation during prolonged successful HIV-1 treatment. AIDS (London, England), 10.1097/QAD.0000000000003512. Advance online publication. https://doi.org/10.1097/QAD.0000000000003512
2. O' Mahony, L., Buwalda, T., Blair, M., Forde, B., Lunjani, N., **Ambikan, A.**, Neogi, U., Barrett, P., Geary, E., O'Connor, N., Dineen, J., Clarke, G., Kelleher, E., Horgan, M., Jackson, A., & Sadlier, C. (2022). Impact of Long COVID on health and quality of life. HRB open research, 5, 31. https://doi.org/10.12688/hrbopenres.13516.1
3. Akusjärvi, S. S., **Ambikan, A. T.**, Krishnan, S., Gupta, S., Sperk, M., Végvári, Á., Mikaeloff, F., Healy, K., Vesterbacka, J., Nowak, P., Sönnerborg, A., & Neogi, U. (2021). Integrative proteo-transcriptomic and immunophenotyping signatures of HIV-1 elite control phenotype: A cross-talk between glycolysis and HIF signaling. iScience, 25(1), 103607. https://doi.org/10.1016/j.isci.2021.103607
4. Svensson Akusjärvi, S., Krishnan, S., Jütte, B. B., **Ambikan, A. T.**, Gupta, S., Rodriguez, J. E., Végvári, Á., Sperk, M., Nowak, P., Vesterbacka, J., Svensson, J. P., Sönnerborg, A., & Neogi, U. (2022). Peripheral blood CD4+CCR6+ compartment differentiates HIV-1 infected or seropositive elite controllers from long-term successfully treated individuals. Communications biology, 5(1), 357. https://doi.org/10.1038/s42003-022-03315-x
5. Munusamy Ponnan, S., Thiruvengadam, K., Tellapragada, C., **Ambikan, A. T.**, Narayanan, A., Kathirvel, S., Mathayan, M., Shankar, J., Rajaraman, A., Afshan Amanulla, M., Dinesha, T. R., Poongulali, S., Saravanan, S., Murugavel, K. G., Swaminathan, S., Velu, V., Shacklett, B., Neogi, U., & Hanna, L. E. (2021). Deciphering the Role of Mucosal Immune Responses and the Cervicovaginal Microbiome in Resistance to HIV Infection in HIV-Exposed Seronegative (HESN) Women. Microbiology spectrum, 9(2), e0047021. https://doi.org/10.1128/Spectrum.00470-21
6. Sperk, M., Mikaeloff, F., Svensson-Akusjärvi, S., Krishnan, S., Ponnan, S. M., **Ambikan, A. T.**, Nowak, P., Sönnerborg, A., & Neogi, U. (2021). Distinct lipid profile, low-level inflammation, and increased antioxidant defense signature in HIV-1 elite control status. iScience, 24(2), 102111. https://doi.org/10.1016/j.isci.2021.102111
7. Sperk, M., **Ambikan, A. T.**, Ray, S., Singh, K., Mikaeloff, F., Diez, R. C., Narayanan, A., Vesterbacka, J., Nowak, P., Sönnerborg, A., & Neogi, U. (2021). Fecal Metabolome Signature in the HIV-1 Elite Control Phenotype: Enrichment of Dipeptides Acts as an HIV-1 Antagonist but a Prevotella Agonist. Journal of virology, 95(18), e0047921. https://doi.org/10.1128/JVI.00479-21

8. Bai, X., Zhang, J., Hua, Y., Jernberg, C., Xiong, Y., French, N., Löfgren, S., Hedenström, I., **Ambikan, A.**, Mernelius, S., & Matussek, A. (2021). Genomic Insights Into Clinical Shiga Toxin-Producing Escherichia coli Strains: A 15-Year Period Survey in Jönköping, Sweden. Frontiers in microbiology, 12, 627861. https://doi.org/10.3389/fmicb.2021.627861
9. Andersson, E., **Ambikan, A.**, Brännström, J., Aralaguppe, S. G., Yilmaz, A., Albert, J., Neogi, U., & Sönnerborg, A. (2021). High-throughput sequencing reveals a high prevalence of pretreatment HIV-1 drug resistance in Sweden. AIDS (London, England), 35(2), 227–234. https://doi.org/10.1097/QAD.0000000000002740
10. Obasa, A. E., **Ambikan, A. T.**, Gupta, S., Neogi, U., & Jacobs, G. B. (2021). Increased acquired protease inhibitor drug resistance mutations in minor HIV-1 quasispecies from infected patients suspected of failing on national second-line therapy in South Africa. BMC infectious diseases, 21(1), 214. https://doi.org/10.1186/s12879-021-05905-2
11. Krishnan, S., Nordqvist, H., **Ambikan, A. T.**, Gupta, S., Sperk, M., Svensson-Akusjärvi, S., Mikaeloff, F., Benfeitas, R., Saccon, E., Ponnan, S. M., Rodriguez, J. E., Nikouyan, N., Odeh, A., Ahlén, G., Asghar, M., Sällberg, M., Vesterbacka, J., Nowak, P., Végvári, Á., Sönnerborg, A., ... Neogi, U. (2021). Metabolic Perturbation Associated With COVID-19 Disease Severity and SARS-CoV-2 Replication. Molecular & cellular proteomics : MCP, 20, 100159. https://doi.org/10.1016/j.mcpro.2021.100159
12. Appelberg, S., Gupta, S., Svensson Akusjärvi, S., **Ambikan, A. T.**, Mikaeloff, F., Saccon, E., Végvári, Á., Benfeitas, R., Sperk, M., Ståhlberg, M., Krishnan, S., Singh, K., Penninger, J. M., Mirazimi, A., & Neogi, U. (2020). Dysregulation in Akt/mTOR/HIF-1 signaling identified by proteo-transcriptomics of SARS-CoV-2 infected cells. Emerging microbes & infections, 9(1), 1748–1760. https://doi.org/10.1080/22221751.2020.1799723
13. Neogi, U., Hill, K. J., **Ambikan, A. T.**, Heng, X., Quinn, T. P., Byrareddy, S. N., Sönnerborg, A., Sarafianos, S. G., & Singh, K. (2020). Feasibility of Known RNA Polymerase Inhibitors as Anti-SARS-CoV-2 Drugs. Pathogens (Basel, Switzerland), 9(5), 320. https://doi.org/10.3390/pathogens9050320
14. Al-Farsi, H. M., Al-Adwani, S., Ahmed, S., Vogt, C., **Ambikan, A. T.**, Leber, A., Al-Jardani, A., Al-Azri, S., Al-Muharmi, Z., Toprak, M. S., Giske, C. G., & Bergman, P. (2019). Effects of the Antimicrobial Peptide LL-37 and Innate Effector Mechanisms in Colistin-Resistant Klebsiella pneumoniae With mgrB Insertions. Frontiers in microbiology, 10, 2632. https://doi.org/10.3389/fmicb.2019.02632
15. Aralaguppe, S. G., **Ambikan, A. T.**, Ashokkumar, M., Kumar, M. M., Hanna, L. E., Amogne, W., Sönnerborg, A., & Neogi, U. (2019). MiDRMpol: A High-Throughput Multiplexed Amplicon Sequencing Workflow to Quantify HIV-1 Drug Resistance Mutations against Protease, Reverse Transcriptase, and Integrase Inhibitors. Viruses, 11(9), 806. https://doi.org/10.3390/v11090806

16. Bai, X., Zhang, J., **Ambikan, A.**, Jernberg, C., Ehricht, R., Scheutz, F., Xiong, Y., & Matussek, A. (2019). Molecular Characterization and Comparative Genomics of Clinical Hybrid Shiga Toxin-Producing and Enterotoxigenic Escherichia coli (STEC/ETEC) Strains in Sweden. Scientific reports, 9(1), 5619. https://doi.org/10.1038/s41598-019-42122-z
17. Babu, H., Sperk, M., **Ambikan, A. T.**, Rachel, G., Viswanathan, V. K., Tripathy, S. P., Nowak, P., Hanna, L. E., & Neogi, U. (2019). Plasma Metabolic Signature and Abnormalities in HIV-Infected Individuals on Long-Term Successful Antiretroviral Therapy. Metabolites, 9(10), 210. https://doi.org/10.3390/metabo9100210
18. Babu, H., **Ambikan, A. T.**, Gabriel, E. E., Svensson Akusjärvi, S., Palaniappan, A. N., Sundaraj, V., Mupanni, N. R., Sperk, M., Cheedarla, N., Sridhar, R., Tripathy, S. P., Nowak, P., Hanna, L. E., & Neogi, U. (2019). Systemic Inflammation and the Increased Risk of Inflamm-Aging and Age-Associated Diseases in People Living With HIV on Long Term Suppressive Antiretroviral Therapy. Frontiers in immunology, 10, 1965. https://doi.org/10.3389/fimmu.2019.01965
19. Zhang, W., **Ambikan, A. T.**, Sperk, M., van Domselaar, R., Nowak, P., Noyan, K., Russom, A., Sönnerborg, A., & Neogi, U. (2018). Transcriptomics and Targeted Proteomics Analysis to Gain Insights Into the Immune-control Mechanisms of HIV-1 Infected Elite Controllers. EBioMedicine, 27, 40–50. https://doi.org/10.1016/j.ebiom.2017.11.031
20. Aralaguppe, S. G., Siddik, A. B., Manickam, A., **Ambikan, A. T.**, Kumar, M. M., Fernandes, S. J., Amogne, W., Bangaruswamy, D. K., Hanna, L. E., Sonnerborg, A., & Neogi, U. (2016). Multiplexed next-generation sequencing and de novo assembly to obtain near full-length HIV-1 genome from plasma virus. Journal of virological methods, 236, 98–104. https://doi.org/10.1016/j.jviromet.2016.07.010

Contents

List of abbreviations

CoV	Coronavirus
MERS-CoV	Middle East respiratory syndrome coronavirus
SARS-CoV	Severe acute respiratory syndrome coronavirus
COVID-19	Coronavirus disease 2019
ARDS	Acute respiratory distress syndrome
HIV-1	Human immunodeficiency virus type-1
AIDS	Acquired immune deficiency syndrome
NK cells	Natural killer cells
ART	Antiretroviral therapy
cART	Combination antiretroviral therapy
IRIS	Immune reconstitution inflammatory syndrome
EC	Elite controllers
CCHFV	Crimean-Congo hemorrhagic fever virus
LDH	lactate dehydrogenase
GEM	Genome-scale metabolic model
EHMN	Edinburgh Human Metabolic Network
Memote	Metabolic model testing
CORUM	Mammalian protein complexes database
INIT	Integrative Network Inference for Tissues
HPA	Human Protein Atlas
HMDB	Human Metabolome Database
GTPR	Gene-transcript-protein-reaction
tINIT	task-driven Integrative Network Inference for Tissues
FBA	Flux balance analysis
$PLWH_{ART}$	People living with HIV on antiretroviral therapy
$PLWH_{EC}$	People living with HIV (Elite Controllers)
$PLWH_{VP}$	People living with HIV (Viremic progressors)
EPIC	Estimating the Proportions of Immune and Cancer cells
DCQ	Digital cell quantification

PBMC	Peripheral blood mononuclear cells
TPM	Transcript per million
SNF	Similarity Network Fusion
CM	Classical monocytes
MAIT	Mucosa-associated invariant T cells
DC	Dendritic cells
OXPHOS	Oxidative phosphorylation
ROS	Reactive oxygen species

1 Introduction

Viruses are obligatory parasites that can cause mild to severe illness upon infection in animals and humans. Viruses require energy and other biochemical materials to replicate and survive. Viruses enter a host and interfere with their metabolic system for replication and reproduction. Since metabolism is the central mechanism for properly functioning all biological processes, such viral-mediated disruptions in the host system can cause severe health conditions. Proper metabolic processes are fundamental for efficient immune system function in health and disease. Significant interlink between immunity and metabolism is observed based on the findings that cellular metabolism dramatically affects the immune response to infectious diseases, inflammation, and cancer [1]. The study on the association between metabolism and immunity, known as immunometabolism, is getting greater attention. An in-depth understanding of cellular metabolic re-programming during viral infection and the associated immune regulation is critical to understand the host immune response and therapeutic interventions. Applications of system-level analytical methods and mathematical modeling approaches using multi-layered omics data can derive information about metabolic reprogramming in a more profound resolution. It can open the development of novel metabolism-based treatment strategies.

1.1 Emerging and re-emerging viruses

In recent decades, public health has been challenged persistently and intensively by newly identified pathogens, majorly zoonotic and vector-borne infectious agents which can cause fatal outbreaks. Human and veterinary health is constantly threatened by emerging and re-emerging virus diseases, with frequent new incidents originating mainly from the animal host. According to World Health Organization, out of the ~1400 pathogens affecting humans, over 50% originate from animal species through a zoonotic transmission naturally transmitted between humans and vertebrates through viral adaptation [2]. The term emerging disease represents the surfacing of a previously known infection or an as-yet-unknown infection, often with alterations in the pathogenicity [2]. They mainly denote frequently evolving infections in response to instantaneous changes in the association between pathogen and host. Most emerging and re-emerging pathogens are zoonotic, and viruses are over-represented in the group [3]. Over one-third of all emerging and re-emerging infections are caused by RNA viruses. Their host range typically spans several types of organisms in mammals, including humans.

Re-emergence has been observed for various viral diseases globally in recent years. Human mobility across national borders is helping the disease to spread very effectively and rapidly, causing any new viral disease to become a potential worldwide pandemic. The increase of emerging and re-emerging infectious diseases is correlated with several factors and is briefly categorized into three groups. The first factor is the change in human

and societal behavior concerning lifestyle. Secondly, environmental and ecosystem changes, and finally, the change in the biology of microorganisms through adaptation [4]. The risks for the emergence and transmission of pathogens and the opportunity for infection are considerably associated with the previously mentioned changes in nature. The most responsible factors for the emergence and spread of infectious diseases are the ones related to humans and society. The global population is overgrowing and is estimated to be over 7.5 billion, which is considered a population explosion. The population rise is more significant in southern regions of Asia and sub-Saharan Africa. Risk for human-to-human transmission of infectious diseases and infections due to poor hygienic conditions are significantly associated with high population and its density. Infectious agents spread worldwide effectively and efficiently because of the current development of airline networks. The previous outbreak of SARS (severe acute respiratory syndrome) in 2003 was linked with advancements in human mobility worldwide.

Socio-economic issues, comprising civil wars, increased number of refugees, and natural calamities, are also indirectly linked with the spread of emerging viruses. Changes in the global environment and ecosystem due to climate change related to global warming are also possible reasons for the increase in zoonosis and vector-borne diseases. Additionally, vectors are favored by natural calamities like floods, heat increases, and drought caused by climate variations and the reduction of natural rivals of the vectors. Globally the significant infectious disease can be grouped as follows with features such as (i) greater worldwide frequency (diarrhea, respiratory infections, etc.), (ii) concern for pandemic (Ebola viral disease, etc.), (iii) associated with eradication (e.g., poliomyelitis),

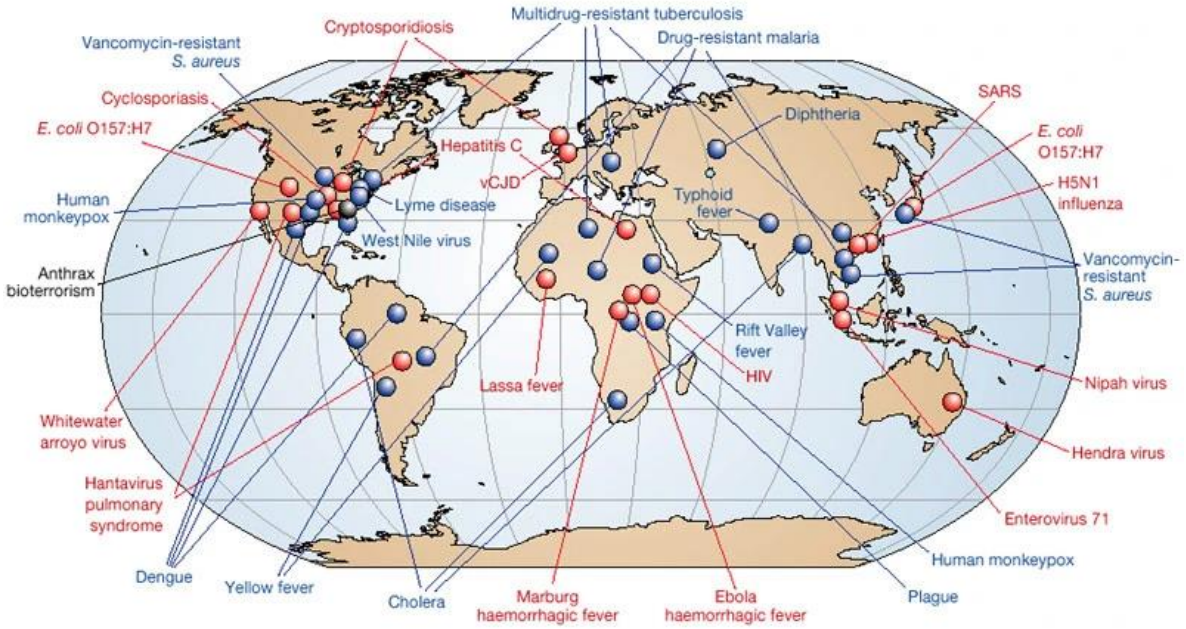


Figure 1: Pictorial illustration representing global occurrence of various emerging and re-emerging diseases. Red coloured bubbles show newly emerging diseases, blue coloured bubbles denote re-emerging/resurging diseases and black bubble represent a 'deliberately emerging' disease. Reproduced with permission from Springer Nature (Morens, David M et al, 2004).

and (iv) ignored tropical diseases (e.g., rabies) [4]. Figure 1 shows the worldwide incidents of emerging and re-emerging diseases caused by various pathogens [5]. The emergence of viral infections is practically challenging to predict, and it is near impossible to eradicate zoonosis because wild animals or vectors can also carry viral pathogens. But it can be efficiently controlled with the help of advanced therapeutic approaches. My thesis presented host immunometabolic research through systems biology studies on three emerging and re-emerging RNA viruses, namely, severe acute respiratory syndrome coronavirus 2 (SARS-CoV2), human immunodeficiency virus type-1 (HIV-1), and Crimean-Congo hemorrhagic fever virus (CCHFV).

1.1.1 Severe acute respiratory syndrome coronavirus 2

Severe acute respiratory syndrome coronavirus2 (SARS-CoV-2) is an emerging infectious virus with single-stranded and positive-sense RNA genetic material. It belongs to the β-coronavirus genus of the coronaviridae family, the seventh coronavirus (CoV) known to infect humans. Four of the seven coronaviruses are human CoV, NL63 (HCoV-NL63), HCoV-229E, HCoV-OC43, and HCoV-HKU1, known to be endemic in humans, and majorly the infection shows common-cold symptoms by affecting the upper respiratory track [6]. Infections in humans have been observed in recent decades by the rest of the three zoonotic coronaviruses, namely, Middle East respiratory syndrome coronavirus (MERS-CoV), severe acute respiratory syndrome coronavirus (SARS-CoV), and SARS-CoV-2, primarily transmitted from animals through adaptation [7, 8]. In 2003, SARS-CoV caused an epidemic in China, and MERS-CoV outbreaks occurred in the Middle East in recent years [6]. The coronavirus disease 2019 (COVID-19) pandemic was first identified in Wuhan city of China, in late 2019, and the causative agent was SARS-CoV-2 [9]. Infection in the lower respiratory tract, which causes fatal acute respiratory distress syndrome (ARDS), is the main characteristic of the three viruses.

Respiratory droplets and aerosols are the media of transmission of SARS-CoV-2, and it has a 4 to 5 days median incubation period [10, 11]. In most cases, asymptomatic infection is observed, but many patients develop mild to moderate respiratory disease showing cough, fever, headache, myalgia, and diarrhea [11]. Severe SARS-CoV-2 infection is commonly associated with dyspnoea (shortness of breath) as a result of the condition known as hypoxemia [12]. Progressive respiratory failure occurs immediately, followed by dyspnoea and hypoxemia in severe COVID-19. Often, severely infected patients develop the condition of ARDS [13, 14]. Systemic level hyperinflammation, including elevated concentrations of inflammatory markers, including D-dimer, ferritin, and C-reactive protein (CRP) and pro-inflammatory cytokines release such as interleukin-1 (IL-1), IL-6, IL-8, and tumor necrosis factor (TNF) are observed in COVID-19 patients having hypoxic respiratory failure [15]. Development of COVID-19 severity is found to be significantly correlated with age, obesity, and gender [16, 17], and conditions such as kidney failure, chronic pulmonary disease, hypertension, cardiac arrhythmia, heart failure, and diabetes

are observed as common co-morbidities [18]. Disruptions in metabolic flux in host cells are commonly observed in viral diseases. Experimental evidence suggests that the changes in host metabolism are associated with cell-specific viral replication and production [19]. Changes in glycolysis rate and ATP production in host cells are known to occur in viral diseases [20]. Since corroborative information shows significant disruptions in the host metabolic system in response to viral infection, a system-level biological understanding of the host can provide crucial knowledge to decide clinical intervention.

1.1.2 Human immunodeficiency virus type-1 (HIV-1)

HIV is another emerging virus of greater importance. Over the last four decades, the world has been suffering a terrible epidemic of the human immunodeficiency virus (HIV). Infection of HIV prompts progressive immunodeficiency in humans, and it can result in acquired immune deficiency syndrome (AIDS) [21]. If left untreated, AIDS can cause other opportunistic infections and, in the worst case, premature death. HIV is a single-stranded, enveloped RNA virus. HIV belongs to the genus Lentivirus in the family of Retroviridae (in the order of Ortervirales). Two strains of HIV have been identified so far, namely HIV type 1 (HIV-1) and HIV type 2 (HIV-2). HIV-1 strains cause most infections. In some countries, such as Guinea-Bissau and Portugal, HIV-2 is present as a significant minority. The most relevant differences between HIV-1 and HIV-2 are more homogeneity of HIV-2 than HIV-1 and more immune control over the HIV-2 infection [22]. There were several HIV-1 subtypes, and recombinant forms were identified in HIV-1 than in HIV-2 [23]. The HIV infection is generally characterized by (1) inefficient viral transmission, (2) an acute phase that involves active viral replication and spreading to lymphoid tissues, (3) a chronic phase which is usually an asymptomatic phase with viral replication and sustained immune activation (4) an advanced phase with the exhaustion of CD4+ T cells leading to acquired immune deficiency syndrome (AIDS) [24].

The most critical barriers to successfully treating HIV-1 infection are the quick establishment and perseverance of various HIV-1 reservoirs. HIV-1 reservoirs can be grouped into two main categories. The first category is lymphoid tissue, which facilitates the plenitude of target cells to HIV-1, efficient propagation among cells, and inefficient penetration of drugs. The second category is cellular reservoirs that include mainly CD4+ T cells which are more vulnerable to HIV replication when activated and it can also bear latent virus. HIV targets significant elements of the immune system. It demolishes and dysregulates CD4+ T cells. HIV also causes reduced immunologic functions of natural killer cells (NK cells), B cells, CD8+ T cells, and nonlymphoid cells through mechanisms that come with activation, high cell turnover, homeostatic responses, and differentiation [24].

HIV infection causes long-term consequences in infected individuals. In most untreated HIV-infected individuals, progressive destruction of CD4+ T cells can occur, also found in individuals taking antiretroviral therapy (ART). Treatment can lead to an increase in CD4+

T cells, but there is evidence of no increment in some individuals called immunological non-responders. Incomplete virologic suppression, as indicated by ongoing immune activation, can cause insufficient refurbishment of CD4+ T cell counts [25]. There is a high risk of developing a catastrophic complication named immune reconstitution inflammatory syndrome (IRIS) in individuals who start ART when CD4+ T cell counts are low and/or in the presence of other opportunistic infections [26]. Initialization of ART in heavily immunosuppressed individuals may lead to sudden activation of the immune response against the secondary pathogen, which triggers the specific inflammatory response, but that can cause significant damage to the tissue [27].

Studies have identified that a rare group of HIV-1 seropositive individuals called elite controllers (EC) can maintain very low virus levels in their bloodstream without ART [28]. Specifically, ECs are individuals with HIV-1 RNA levels below the detection limit (typically less than 50 copies/mL) on at least two separate occasions without receiving ART. It is estimated that only about 1% of people living with HIV-1 are EC. These individuals have a unique immune response that allows them to control the virus without developing AIDS. Research into EC is ongoing, as scientists are interested in understanding how these individuals can control the virus without medication. This knowledge could potentially lead to new treatments or a cure for HIV-1. Studies have shown that EC has higher levels of glucose uptake in their CD8+ T cells compared to individuals with progressive HIV-1 infection. This increased glucose metabolism appears necessary for CD8+ T cells to maintain their function in controlling the virus. In addition, studies from our group have shown the role of other metabolic pathways, such as lipid metabolism and mitochondrial function, in the immune response to HIV-1 [29]. These studies are helping to shed light on the complex interplay between metabolism and the immune system and how alterations in metabolic pathways can impact immune function.

1.1.3 Crimean-Congo hemorrhagic fever virus (CCHFV)

CCHFV is declared a critical emerging and re-emerging pathogen of interest by the World Health Organization for Animal Health and the World Health Organization [30]. The CCHFV belongs to the genus Nairovirus of the Bunyaviridae family. Viruses in the Nairovirus genus are tick-born in nature, and large L segments in their genome distinguish them from other bunyaviruses. They are primarily grouped into seven serogroups, and CCHFV and Hazara virus (HAZV) represent the CCHF serogroup. The Crimean–Congo hemorrhagic fever (CCHF) is a tick-born viral disease affecting humans, which is considered one of the most widespread infections globally. The primary geographic locations of the viral outbreaks include the Middle East to southeastern Europe, western China through southern Asia, and a majority of Africa. CCHFV is the causative agent of the disease, and it is transmitted through vertical and horizontal transmission through ticks belonging to the Ixodidae family. The ticks spread the virus to different kinds of wild and domestic mammals. Exposures to infected animals' blood, other body fluids, or direct tick bites are the reasons

for infection in humans. The primary source of human infection is the bite from ticks belonging to the genus Hyalomma. Nonspecific febrile syndrome through the vascular leak, multi-organ failure, shock, and hemorrhage are the main characteristics of CCHFV infection in humans [31]. Cases of CCHFV incidents have been recently reported since 2000 in Iran, Turkey, Greece, the Republic of Georgia, India, and some Balkan countries. Also, viral RNA was detected in Hyalomma ticks isolated from a deer in Spain. From the first case in Turkey reported in 2002, the total number of cases increased to over 6000 in ten years. Iran has also faced a large increase in cases since 1999 [32].

There are four phases in the CCHF course (1) incubation, (2) pre-hemorrhagic, (3) hemorrhagic, and (4) convalescent [33]. The mode of acquisition of the virus decides the incubation period length. The incubation usually spans 1-5 days after the tick bite, whereas in the case of exposure to infected blood or tissue, the period is mostly 5-7 days with a maximum of 13 days [34]. Lassitude, fever, and various nonspecific signs and symptoms are characteristics of the pre-hemorrhagic phase of CCHF, and it usually starts at 3-5 days of illness [34]. In the hemorrhage phase, which is fatal, death commonly happens on days 5-14 due to hemorrhage, multi-organ failure, and shock. Vaginal and abdominal bleeding and cerebral hemorrhage are common symptoms. Full recovery from the disease may take up to a year in surviving patients [34]. The recovery is often associated with health problems, including hair loss, poor appetite, weakness, hearing loss, polyneuritis, impaired vision and memory, and hepatorenal insufficiency [35].

We lack in-depth knowledge about the host-virus interactions and pathogenesis of the acute phase of CCHF disease and related health conditions after recovery because of inadequate availability of systematic studies, lack of infrastructure, and sporadic nature of CCHF outbreaks. Designing efficient therapeutic and containment strategies for CCHF is only possible with thorough knowledge of host response against CCHFV. The disease pathogenesis and host immune response mechanisms can be effectively derived from a system biology approach to omics data generated from patient material and infected cells [36].

1.2 Virus: A metabolic re-programmer

Viruses need the host's cellular metabolism to survive. A significant availability of host resources is required for the viral proliferation in the infected body, including nucleic acids, proteins, membranes, and energy required for viral synthesis. Many transcriptome-wide and metabolism studies identified a consequential shift in the metabolic system of infected cells. Cellular mechanisms such as replication and macromolecular synthesis are found to be inhibited by viral infection to use valuable resources for mass production of the virus [37]. Earlier studies on infection caused by various viruses have shown alterations in the glycolysis pathway. Increased level of glycolysis was identified in infections from viruses such as Rous sarcoma virus [38], feline leukemia virus [39], and poliomyelitis virus

[40]. In contradiction, inhibition of glycolysis was found in response to influenza infection by blocking the glucose-6-phosphate isomerase (GPI) reaction [41], but later studies found that early phases of influenza were associated with an increase in glycolysis [42]. Other aspects of metabolism are also identified to be dysregulated upon viral infection. Studies have shown that apart from increased glycolysis, redirection of cellular resources in favor of viral replication also occurs in the infection stage. Dysregulation in cellular metabolism by the increased rate of RNA breakdown to enlarge the nucleotide pool availability by the secretion of ribonuclease is observed in the poliomyelitis virus infection [43]. Similarly, a rapid decrease in available ATP was identified in the Rubella infection [44]. Invasion of mitochondria for the production of virus-specific compounds was identified in other viral infections as well [45]. My thesis is to identify various aspects of metabolic reprogramming of host cells in response to infections by SARS-CoV-2, HIV-1, and CCHF viruses.

Although, no systematic studies have fully characterized the effect of CCHFV infection in host metabolism, substantial characterizations have been employed in SARS-CoV2 and HIV-1. The SARS-CoV2 infection causes prominent host metabolic shifts (Figure 2) [46]. Regulation of immune responses and the assembly of progeny virus is mainly orchestrated by the amino acid metabolism [47]. Studies have identified significant roles of the kynurenine [48], arginine [49, 50], and glutamine [51] in COVID-19 disease response. The enhanced glycolysis pathway is a primary characteristic of many virus infections, which helps with viral replication by rapidly producing energy and other necessary

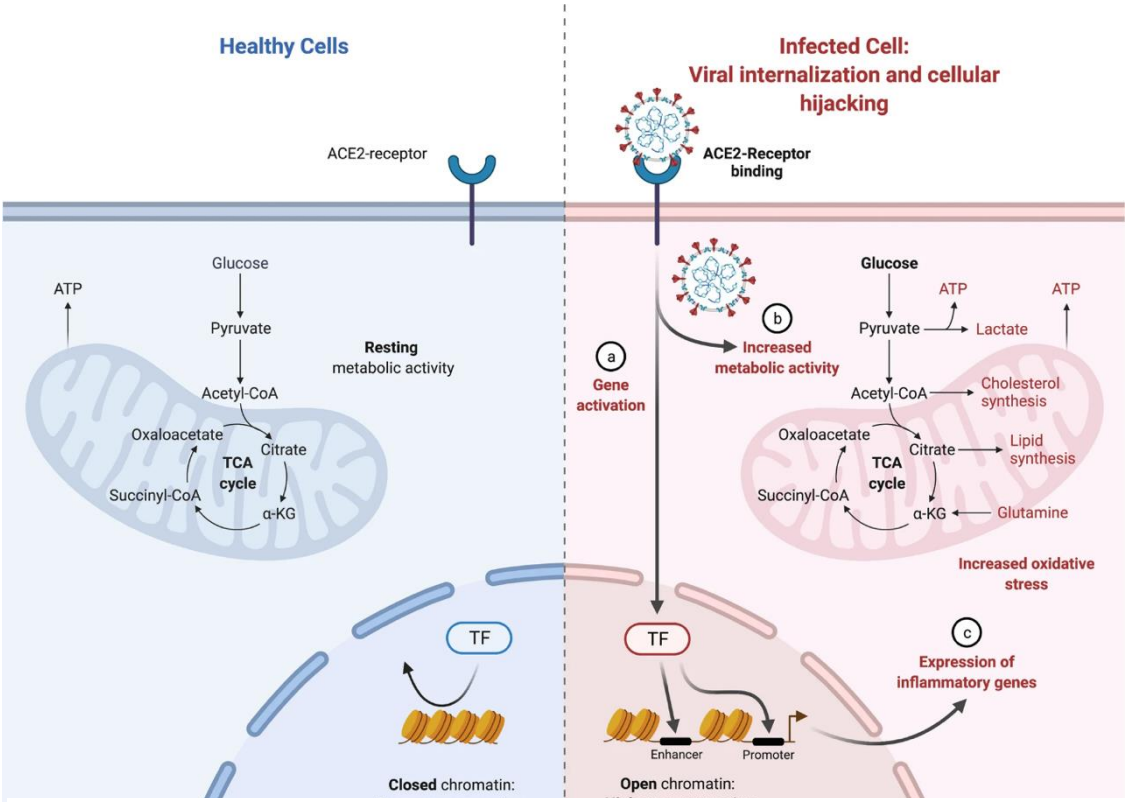


Figure 2: Illustration of host metabolic shift due to SARS-CoV2 viral replication and associated cellular dysregulation. Adapted from Jamison, David A Jr et al, 2022.

substrates [52]. Further, several studies have reported increased levels of metabolites or enzymes of glycolysis, such as pyruvate, pyruvate kinase, and lactate dehydrogenase (LDH) in COVID-19 [53]. Glycolysis is also crucial for SARS-CoV-2 to replicate [54]. Additionally, studies have also reported the importance of cholesterol [55] and fatty acid biosynthesis [56] [57] in the host response against COVID-19.

Several studies suggest that HIV-1 pathogenesis is driven by immunometabolism, where metabolic disruptions may cause an immune malfunction that leads to persistent inflammation, resulting in AIDS-associated comorbidities [58]. A significant association between HIV-1 pathogenesis and T-cell metabolic reprogramming is proposed since HIV-1 mainly infects CD4+ T-cells [59]. HIV-1 associated metabolic reprogramming in adaptive and innate immune cells is characterized by two major features: (1) upregulation of glycolysis pathway and downregulation of oxidative phosphorylation to favor more rapid ATP generation and biosynthesis for defense response and damage repair and (2) epigenetic reprogramming due to reduction in DNA methylation and increased histone acetylation [60]. The most abundant circulating amino acid, glutamine, which can also be converted to α-ketoglutarate, has a crucial role in the immune response. Glucose metabolism and glutamine are interlinked. Further, another study shows that glutaminolysis is the crucial pathway feeding the oxidative phosphorylation and TCA cycle in T-cell receptor-stimulated naïve and memory CD4 subsets essential for optimal HIV-1 infection [61]. It is evident from the literature that it holds therapeutic importance to accurately elucidate metabolic changes in the resolution of metabolic reactions associated with HIV-1 infection.

1.3 Systems biology to unveil the complexity

The scientific community has witnessed incredible technological advancement in the past decades that generate high-depth, multidimensional omics data and the evolution of advanced informatics approaches to analyze them. The multi-dimensional biological data include transcriptomics, proteomics, metabolomics, lipidomics, and microbiome. Integrative analysis of such multilayered data can generate a meaningful system-level understanding of the disease understudy. Network analysis and machine learning strategies are a few of the many techniques that can be applied to analyze complex and multi-layered data, but careful attention is required on the purpose of the method and the biological problem needed to be solved.

System biology is an applied science discipline that deals with the integrative analysis of multilayered data to understand complex biological mechanisms. Biological systems are often considered 'complex systems'. A complex system is usually represented by many simple and identical components interacting to give rise to complex behaviors. But the biological systems are rather different in reality. In the case of a biological system, many functionally divergent, often multifunctional sets of components interact among them

nonlinearly and selectively to generate coherent instead of complex behaviors. In a complex system formed by simple components, the functions are derived from the properties of the network and not from any specific components of the system. In contrast, biological systems depend on the combination of the network and its constituting components, which leads to the notion that biological systems are better considered symbiotic systems [62]. Another study performed on interactions data of yeast by Ideker, Trey et al. showed the possibility of formulating robust hypotheses for the fundamental mechanisms controlling the observed changes in gene expression by integrating the interactions data with mRNA expression data [63]. The availability of novel high-throughput data acquisition methods enables us to efficiently investigate the status of components of a cell and to identify the nature of interactions among these components. Understanding the topological and dynamical characteristics of the divergent networks that govern the cell's behavior is a major challenge in modern biology [64].

An explanation of biology in terms of its interacting components is the core philosophy of system biology [65]. For example, if an enzyme's action on a compound is considered a system, studying its interaction is crucial to understand the system. Now the advancement in technologies enables us to measure many distinct components of a system, making the system of study very complex; analyzing such a system requires great computational effort based on mathematics and statistics. Such a requirement paved the way for developing the discipline of systems biology. One of the important aspects of systems biology is the analysis of network forms by interacting with various biological components of the system under consideration. Apart from standard single-layer omics analysis, two different methodologies of systems biology, namely, genome-scale metabolic modeling (to study the metabolic network of a system based on the association between metabolites and genes) and co-expression network analysis (to analyze the coordinated functional response to stress in a system) are employed in the thesis to investigate the metabolic reprogramming aspects of three RNA viruses.

1.3.1 Genome-scale metabolic modeling

Metabolism in humans is an integral part of cellular function, and disruptions in metabolism are associated with adverse health conditions such as cancer, obesity, diabetes, etc. Studies have shown that viral diseases cause significant alterations in host metabolism by hijacking and reprograming host metabolic mechanisms to favor viral replications. An in-depth understanding of human metabolism in a holistic fashion in cells remains challenging despite the importance of metabolism and the availability of techniques to measure thousands of metabolites [66]. One of the possible reasons for this difficulty is that metabolism needs to be defined by metabolic fluxes through each reaction and not by the concentrations of biomolecules themselves (such as proteins, metabolites, or mRNAs). The concentrations of biomolecules should only be considered

as indirect proxies for biological activity [67]. Genome-scale metabolic models (GEMs) are mathematical representations of the metabolic system of a cell, tissue, or an entire organism. It represents the pool of total metabolic reactions as a series of stochiometric equations which can be used to study the phenotype of the biological entity in response to metabolic stress. Over the past decade, researchers have successfully developed and improved many GEMs for human metabolism. The development of GEMs has begun with the creation of Recon1 [68] and the Edinburgh Human Metabolic Network (EHMN) [69]. The Recon1 and EHMN became the commencement of two parallel model series: the Recon series (Recon1, 2, and 3D) [68, 70, 71] and the Human Metabolic Reaction series (HMR1 and 2) [72, 73]. These two models have been used to study diseases such as cancer, dysbiosis, fatty liver disease, and diabetes [74, 75], and they influenced each other during development and rectification. Many challenges persist in the development of a human GEM. Using non-standard identifiers for reactions, genes, and metabolites in GEM is one of them. Other problems include continuous use of errors from previous models, effort divided among multiple model lineages, delay in model rectification, non-transparent development, and difficulties in coordination among the scientific community [76].

1.3.2 Human 1 (Human-GEM)

Robinson et al. developed Human 1, the first version of an integrated model of human GEM lineages (Human-GEM) [76]. The human-GEM was created by extensively curating the Recon and HMR model lineages and unifying them. Integration of the information and components from HMR2, iHsa [77], and Recon3D yielded the Human-GEM containing 13,417 reactions, 10,138 metabolites (4,164 unique), and 3,625 genes. Version-controlled Git repository was efficiently used in the entire development process to make each of the modifications and changes publicly visible and thereby favor collaboration with the research community. While curating the model, the developers removed 8,185 duplicated reactions and 3,215 duplicated metabolites. Further, 2,016 metabolite formulas were revised, rebalanced 3,226 reaction equations based on stoichiometry, the reversibility of 83 reactions was corrected, and 576 inconsistent reactions that violated the conservation mass law were removed. A new, updated generic human biomass reaction was created based on different tissue and cell composition data to favor flux simulations. The metabolic model testing (Memote), a community-maintained platform for testing GEMs, was used to assess the quality of the Human-GEM [78]. While testing using Memote, Human 1 showed excellent performance with 100% stoichiometric consistency, 98.2% charge-balanced reactions, and 99.4% mass-balanced reactions.

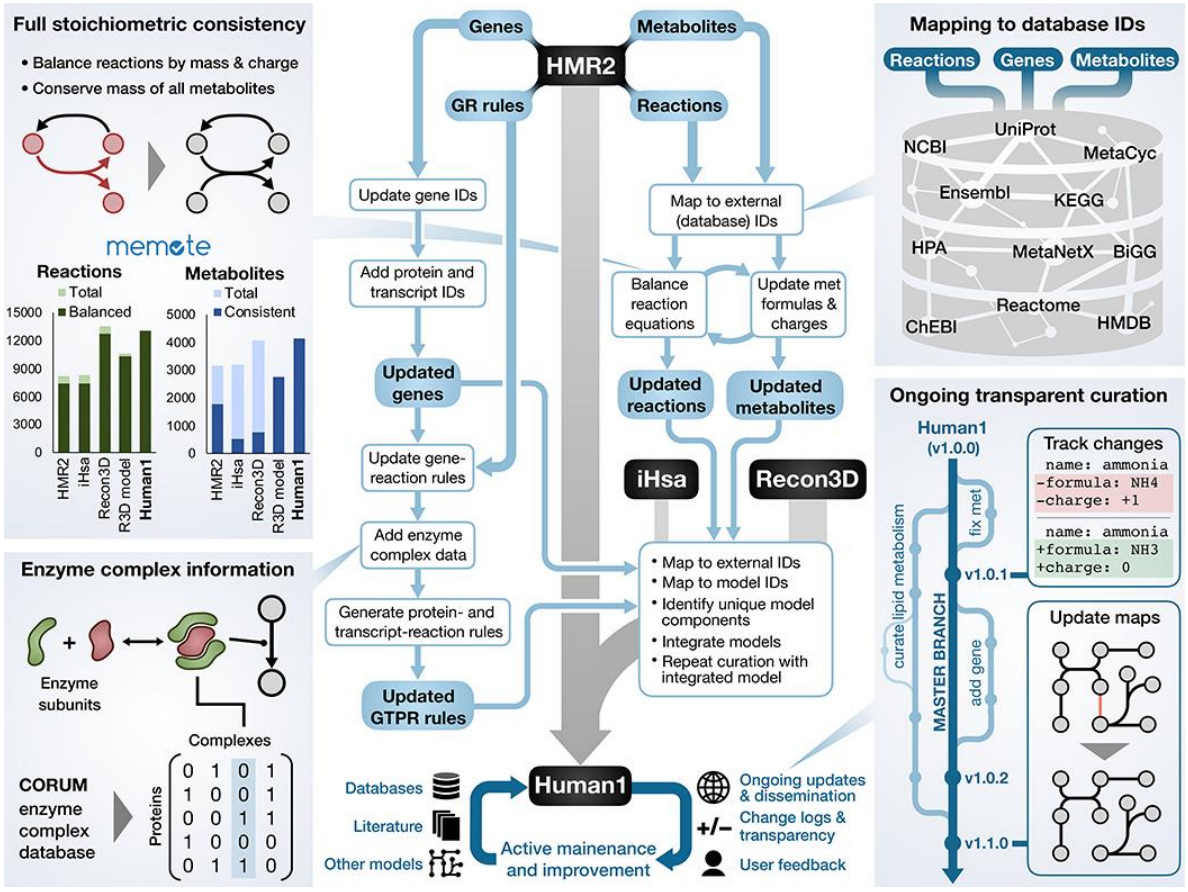


Figure 3: A pictorial illustration of the main steps involved in the creation of Human1 from HMR2, Recon3D, and iHsa. Reprinted with permission from AAAS (Robinson, Jonathan L et al, 2020).

Using Human 1, it is possible to integrate various molecular datatypes to facilitate better interpretation within the context of the metabolism [79]. The comprehensive resource of mammalian protein complexes database (CORUM [80]), iHsa, and enzyme complex information from Recon3D were combined and integrated with gene-reaction associations from HMR2, Recon3D, and iHsa to derive gene-reaction rules for Human 1. Direct integration of protein- or transcript-level data into the model is also possible due to the incorporation of transcript- and protein-reaction rules in the model [81]. The overview of the steps involved in the development process of Human 1 is illustrated in Figure 3 [76].

1.3.3 Contextualized/Context-specific GEM

The Human-GEM consists of metabolic reactions known to occur in any human cell, so it is a generic model of human metabolism. However, it does not represent a metabolic model of any one tissue or cell type in which only a subset of reactions is active. Therefore, the GEM extraction or contextualization process needs to be performed to create a GEM that represents any tissue or cell type of interest. The resultant model is called extracted or contextualized/context-specific GEM. The extraction process requires the integration of the corresponding dataset of interest, such as transcriptomics, proteomics, metabolomics, etc., on the metabolic reference model. The generic metabolic networks

such as Recon1, EHMN, and HumanCyc [82] are not tissue-specific, so they cannot be used to study particular human cell types or diseases. Earlier attempts have been made to generate models specific for ten different human tissues using tissue-specific transcription profiles [83], again subsets of Recon1. But these models can study the metabolic states of the tissues subject to different genetic and physiological conditions only on appropriate approximations [84]. Another approach was later published, which uses transcriptomic and proteomic data combinedly and Recon1 as a template to create a more flexible liver-specific metabolic model [84].

Integration of the data about the absence or presence of metabolic enzymes in a specific cell type while maintaining a well-connected network is the aim of the computational algorithm to construct a cell, type-specific metabolic model. A well-connected network assumes that metabolites consumed in one reaction must be generated in another reaction or consumed from the cell environment [85]. The mRNA expression differences are comparative to a reference state and not absolute; the transcriptome data are often noisy. Moreover, in most cases, enzyme levels do not correlate with the transcript expression levels [86]. Agren, Rasmus, et al. developed an automated computational pipeline for identifying expressed cell type-specific genome-scale metabolic networks [85]. The algorithm Integrative Network Inference for Tissues (INIT) forms a vital part of the pipeline, which uses information from the Human Protein Atlas (HPA) [87] about the presence or absence of metabolic enzymes in each human cell type. The INIT algorithm uses tissue-specific transcriptome expression [88] data as an additional source of information. INIT further constraint the model so that if a metabolite has been identified in a specific tissue, the resulting tissue-specific metabolic network should be able to produce this metabolite from the precursors using the metabolomics data derived from Human Metabolome Database (HMDB) [89].

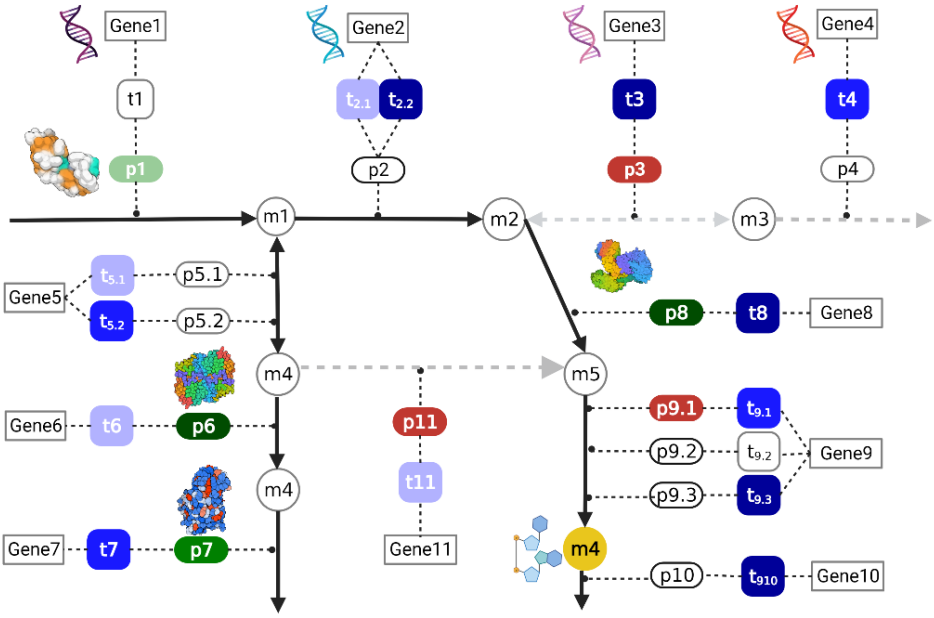


Figure 4 : The concepts and main steps of the INIT algorithm. Adapted from Agren, Rasmus et al., 2012.

Figure 4 here illustrates the main concepts of the INIT algorithm [85]. The gene-transcript-protein-reaction (GTPR) associations are the main

characteristic of the hierarchical structure of GEMs. One or more enzymes are associated with each metabolic reaction in the model, and they are also associated with transcripts and genes. The algorithm computes a score for each reaction catalyzed by an enzyme based on the knowledge of the presence or absence of the given enzyme or gene expression in a cell type. The HPA evidence scores are represented in figure 4. Components in red, light green, medium green, and dark green show negative, weak, moderate, and strong evidence scores. The gene expression evidence scores are represented as red, light blue, medium blue, and dark blue, respectively, representing zero, low, medium, and high expression. Metabolomics data are available for some metabolites to prove their presence in the cell types, and they are described as yellow circles. The algorithm then tries to identify the sub-network from the model, which consists of genes or proteins with strong evidence for their presence in the cell type under study. The algorithm does the extraction by maximizing the sum of evidence scores. Additionally, the algorithm ensures that all the reactions included in the extracted sub-network should be able to carry a flux, and all the metabolites identified experimentally should be generated from the precursors the cell is known to consume. The bold lines in figure 4 denote the resultant network.

Recently, a modification of the INIT algorithm named task-driven Integrative Network Inference for Tissues (tINIT) is proposed because the model generated by INIT is not a functional model which could be directly used for simulations; instead, they are just snapshots of functioning metabolism in a given tissue [90]. The algorithm tINIT facilitates the direct reconstruction of functional GEMs. Also, the algorithm provides the functionality to define metabolic tasks the extracted model must carry out. There are mainly two additional improvements in the tINIT over the INIT algorithm. Firstly, by constraining the solution, the algorithm does not let the reversible reactions have flux in both directions simultaneously. Secondly, it enables the user to decide whether net production of all metabolites should be allowed. A list of metabolic reactions known to occur in all cell types is used for applying constraints on the functionality of the reconstructed models.

1.3.4 Flux Balance Analysis (FBA)

Flux balance analysis can be defined as a mathematical methodology for analyzing the flow of metabolites through a metabolic network [91], and metabolic flux is defined as the rate of turnover of metabolites in a metabolic reaction. It is possible to predict an organism's growth rate or the production rate of biotechnologically important metabolites by using FBA, as it computes the flow of metabolites through a metabolic network. The mathematical representation of metabolic reactions is the first step in FBA. The representation is done in a tabulated structure containing a numerical matrix representing each reaction's stoichiometric coefficients. The stoichiometries of the reaction constrain the flow of each metabolite through the network.

Figure 5 shows each step involved in the flux balance analysis [91]. Each metabolic reaction in a network is denoted as stoichiometric matric (**S**) of dimension **m*n** where row (**m**) represents each metabolite and column (**n**) represents each reaction. Stoichiometric coefficients of the metabolites in each reaction are the values in the matrix. Metabolite consumption is expressed as a negative stoichiometric coefficient, whereas a positive coefficient means metabolite production by the corresponding reaction. The coefficient will be zero if any metabolite is not participating in a reaction. Most metabolic reactions consist of only a few different metabolites, due to which **S** will be a sparse matrix. A vector, **v** of length **n** is used to denote the flux through all the reactions in a network, whereas a vector, **x** of length **m** represents concentrations of all metabolites. Further, the system is assumed to be in a steady state (**dx/dt=0**) which gives **S*v = 0**. Any **v** that satisfies the equation is considered to be in the null space of **S**. In any large-scale model, there are always more reactions than compounds (**n>m**) which means more unknown variables than equations so that no unique solution is available to this system of equations. The constrain define a solution range, and it is possible to find and analyze single points within the solution space. FBA aims to compute such optimal points inside the constrained space. Conceptually, when there are no constraints applied, the flux distribution of the metabolic network falls at any point inside the solution space, and when the mass balance constraints by stoichiometric matrix (**S**) and capacity constraints by lower and upper bounds are applied, the network may get flux distribution within the constraint space. The FBA then optimizes an objective function and computes a single optimal flux distribution on the edge of the constraint solution space.

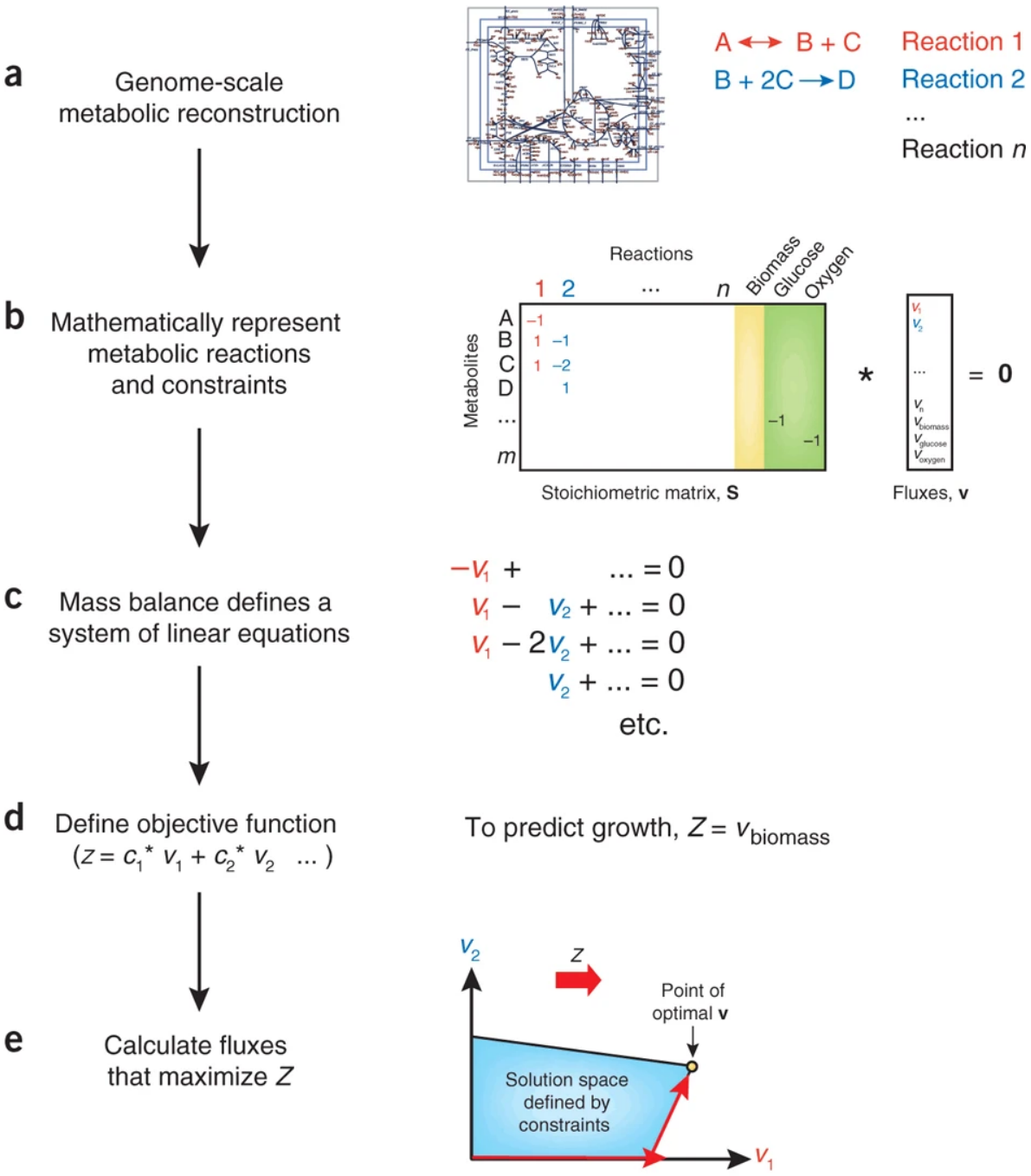


Figure 5 : Steps involved in flux balance analysis. Reproduced with permission from Springer Nature (Orth, Jeffrey D et al., 2010)

An objective function $\mathbf{Z=c^Tv}$ is getting maximized or minimized in FBA. The objective function is a linear combination of fluxes where **c** is a vector of weights that shows the contribution of each reaction to the objective function. If there is only one reaction is needed to maximize or minimize, then **c** is a vector of zeros with one at the position of the reaction of interest. During FBA, optimization of the system of equations is carried out by using linear programming, which is a mathematical approach to computing the best possible outcome from a mathematical model whose requirements are denoted by linear relationships. Given a set of upper and lower bounds on **v** and a linear combination of

fluxes as an objective function, the FBA tries to solve the equation **Sv=0** using linear programming. Defining an objective biological function relevant to the problem is the primary step in FBA. The objective function represents how much reach reactions contribute to the studied phenotype. Flux balance analysis uses various objective functions depending on the phenotype of the study to analyze. The objective of the study decides the choice of the objective function. Maximization of growth or biomass production is the most used objective function, which facilitates a variety of flux predictions consistent with study observations [92]. Examples of other objective functions include (1) minimization of ATP production, which allows predicting conditions of optimal metabolic energy efficiency (2) nutrient uptake minimization (3) Metabolite production maximization to predict the production abilities of the cell (4) biomass and metabolite production maximization (5) minimization of absolute norm of the flux vector to study optimal channeling of the metabolite [93]. Quality reconstruction of metabolic networks and selection of constraints are essential processes in FBA. A study on P.putida has shown that the metabolic network structure is the crucial factor in deciding the accuracy of FBA results, and the objective function has a lesser impact [94]. Incomplete protein annotation in the genome has a negative effect on FBA despite the fact that it can give hints to increase the current knowledge. A subset of the genome, which constitutes mainly the enzymes that catalyze different metabolic reactions in a cell, is the main focus of FBA. The incomplete nature of genome annotation has a severe impact on FBA. For example many reactions can be predicted to have zero fluxes only because of the non-characterization of the downstream or upstream reactions of them [93]. There are other types of flux balance analysis to study the reaction flux through a metabolic network. Flux variability analysis (FVA) and parsimonious flux balance analysis (pFBA) are two of them. The FVA can be applied to measure the minimum and maximum flux for each reaction in the metabolic network while conserving the state of the network [95]. A few applications of FVA include the investigation of alternative optima [95], exploring the distribution of flux under suboptimal growth [96], studying the flexibility and redundancy of network [97], process formulation optimization for antibiotic production [98], and optimal strain design process as a pre-processing step [99, 100]. Whereas pFBA seeks to minimize the metabolic flux of each reaction in the model while conserving the optimum flux through the objective function [101].

1.3.5 Weighted network analysis

In many systems-biology investigations, networks are used to model the experiment system computationally [102]. A network, in an organized manner, quantitatively represents information about complex biological systems. There are two primary components in any network, namely, nodes and edges. Figure 6 illustrates the visualization of essential components of a network [65]. In the network (Figure 6), the circles are nodes that denote the system's components (genes, proteins, metabolites, etc.). The lines connecting two nodes are called edges. Each edge represents the relationship between the components (nodes). Edges can also use to define the strength, type, and direction of the association between the nodes. The strength of an edge is called the edge weight in network analysis, and the edge weight can be derived from experimental data. For example, in a network representing a correlation between genes in response to external stress, the genes will be nodes, the correlation between genes will define the edge, and the correlation coefficient can be used to define the edge weight. Since in correlation, it is not possible the define the direction, the resultant network will be un-directed that means edge direction cannot be defined.

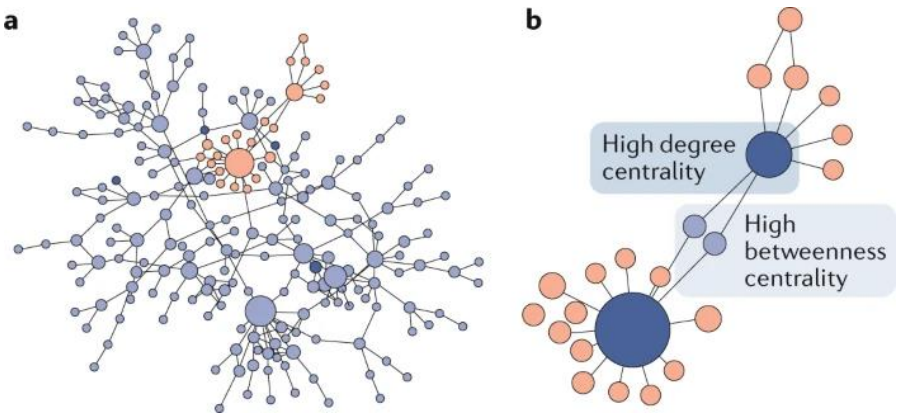


Figure 6: Representation of network showing its components and node centrality measures. Reproduced with permission from Springer Nature (Joshi, Abhishek et al, 2021)

Network science assigns properties for the nodes that define their importance for the existence of the corresponding network. One of the main properties of the node is centrality. Calculation of the centrality of nodes means the identification of which nodes are more "central" than others; in other words, central nodes are the nodes that sit "in the thick of things" or the focal points of the network [103]. The network shown in Figure 7 shows five nodes and associated connections (edges). In the network, the node in the middle (Node 1) has three advantages over the other nodes (1) the middle node has more connections, (2) it can get to other nodes more rapidly, and (3) it also governs

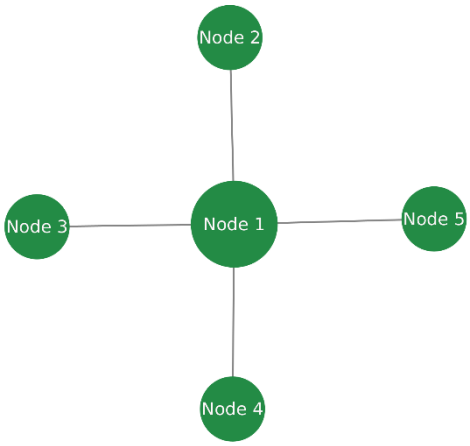


Figure 7 : A basic network of 5 nodes. Node size is relative to degree centrality. Created in Cytoscape

the flow among the other nodes. These three features formed the basic rules used by Freeman to define three various measures of node centrality [103]: degree, closeness, and betweenness. The number of nodes a node in focus is connected to defines the degree of centrality. The degree of a node implies how much the node is involved in the formation of the network. The advantage of degree is that it is simple to compute if the structure around the node is known. But the degree is not considering the global structure of the network, which is one of its limitations. If a node has a high degree, it is connected to many others, but it might not be able to access other nodes rapidly to gain resources such as knowledge or information [104]. Closeness centrality was introduced to overcome this limitation of degree. The closeness is defined as the inverse sum of the shortest distances to all other nodes from a node in focus. But the closeness centrality does not apply to networks with disconnected components, where components are a portion of the network that is disconnected from the main network. (there is no finite distance between two nodes belonging to two different components). Generally, the most significant component of the network is used to compute closeness. The betweenness centrality whereas defines as the degree to which a node falls on the closest path between two other nodes and can channel the flow through a network. The betweenness of a node also shows the control of information flows through the network. The global structure of the network is considered in the computation of betweenness, and it can also be used in networks with disconnected components [105].

The centrality measures defined by Freeman are explicitly defined for binary networks, and there were many attempts to generalize the three centrality measures in the context of weighted networks [106, 107]. Edge weight was the sole focus of all those attempts for generalization and not on the number of edges as in the original definition. The definition of degree was extended by Barrat et al., and they defined it as the sum of the weights attached to the edges connected to a node [108]. This means that a degree of 5 could mean that either the node has five edges with weight 1 or 1 edge with weight 5; a combination of both also could be possible in some cases. Next, the definition of closeness and betweenness were modified by Newman [109] and Brandes [107], respectively, by using Dijkstra's shortest path algorithm [110]. This Dijkstra algorithm defines the shortest path between two nodes as the least costly. The improvement of closeness and betweenness propose costs only depend on edge weights.

Table 1. Table shows definition of different centrality measurements

Centrality	Definition
Degree	The number of nodes a node in focus is connected to.
Closeness	The inverse sum of the shortest distances to all other nodes from a node in focus.
Betweenness	The degree to which a node lies on the closest path between two other nodes and can channel the flow through a network

1.3.6 Community detection

A cluster of nodes in many networks forms dense groups, which is often called communities [111]. The modular structure of the network which forms the community is often unknown beforehand. Therefore, the identification of network communities is a crucial problem, and it can help to derive valuable information about the network. Modularity is one of the best methods to detect community [112]. The modularity-based method maximizes the difference between the expected number of edges and the actual number of edges in a community. The expected number of edges can be denoted as $\mathbf{K_c^2/2m}$ where $\mathbf{K_c}$ is the sum of the degree of the nodes in a community **c,** and the total number of edges in the network is **m**. So the modularity can be given by, $H = \frac{1}{2m}\sum_c(e_c - \gamma\frac{K_c^2}{2m})$ where $\gamma > 0$ is a resolution parameter [113]. The lower resolution gives fewer communities, and the higher resolution gives rise to more communities. The most widely used and cited algorithm for modularity optimization was the Louvain algorithm [114]. Later, Traag, V A et al. showed that the Louvain algorithm is not efficient [115]. As per their observations, the issue known as resolution limit [116] may cause the algorithm to lead to badly connected communities. The identified communities even may be disconnected internally. In order to overcome these issues, Traag V A et al. proposed a new algorithm named Leiden which is faster and identifies better partitions and gives explicit guarantees and bounds [115].

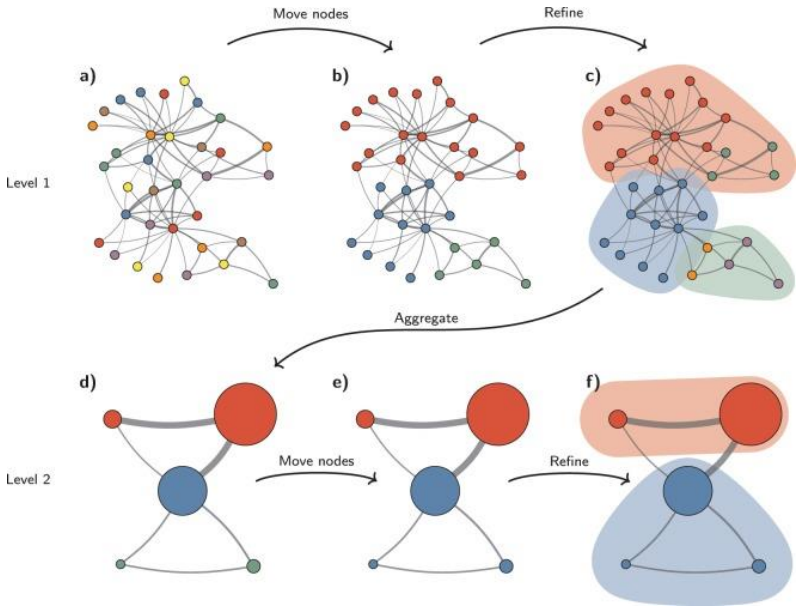


Figure 8 : Illustration of the steps involved in Leiden algorithm. The steps will be repeated until no further improvement is possible. Adapted from Traag, V A et al, 2019

The Leiden algorithm partly uses the smart local move algorithm [117]. The algorithm contains three steps, (1) local movement of nodes, (2) modification

of the partition, and (3) assembly of the network based on the modified partition, using the non-modified partition to create an initial partition for the assembled network. Figure 8 shows the steps involved in the algorithm pictorially. The weighted network analysis combined with centrality computation and community detection can provide valuable information about the biological network under study.

2 Research aims

The project's overall aim was to investigate the metabolic re-programming that occurs in humans in response to viral infections caused by three different viruses, namely, SARS-CoV2, HIV-1, and CCHFV using systems biology approaches.

The specific aims of each of the paper are written below,

Paper I: The clinical outcome and disease severity in COVID-19 are heterogeneous, and a single factor like age, ethnicity or co-morbidities cannot explain the progression or fatality of the disease. In this study, we aim to stratify patient groups solely based on multi-omics data signatures and to identify the Metabo-transcriptomics mechanisms associated with COVID-19 severity at personalized and group levels.

Paper II: During HIV-1 infection, cellular metabolic activity is significantly altered. After initiation of combination antiretroviral therapy (cART), virus-induced short-term metabolic changes do not restore the transient metabolic modulation caused by the infection. Here we sought to study the Metabo-transcriptomics changes in response to HIV-1 infection at the system level by comparing people living with HIV on prolonged cART ($PLWH_{ART}$) with the HIV-seropositive Elite Controllers ($PLWH_{EC}$) and HIV-negative controls using system biology methodologies.

Paper III: The pathogenesis and host-viral interactions of the CCHFV are complicated and poorly evaluated. The multi-omics system biology approaches, including biological network analysis, were used to derive knowledge about the complex host-viral response and viral pathogenesis. The study aimed to fingerprint the system-level alterations during acute CCHFV infection and the cellular immune responses during productive CCHFV replication in vitro.

Paper IV: The trajectory of system-level host response against infection provides a comprehensive host immune response, which is unknown in CCHFV infection. In the study, we aim to identify the impact of CCHF viral infection in the metabolic process and rearrangement of intracellular metabolic fluxes during progressive infection using system biology methods.

3 Materials and methods

3.1 Cohort descriptions and ethical considerations

All the papers presented in the thesis used data generated from patient samples. Descriptions of the cohort used and associated ethical details for each paper are written below.

3.1.1 Paper I

The patients were recruited from the South Hospital, Stockholm, and they were COVID-19 PCR positive (n=37) and were hospitalized in May 2020. The patients were grouped into hospitalized—mild (O2 consumption <4 l/min, n=26) and hospitalized—severe (O2 consumption ≥4 l/min, n=11) based on their oxygen requirement. COVID-19 PCR negative (n=31) patients were also recruited as the control group, of which ten patients were SARS-CoV-2 Ab positive. The regional ethics committees of Stockholm (DNR 2020-01865) approved the study and abided by the Declaration of Helsinki principles.

3.1.2 Paper II

The study was conducted on the Swedish InfCareHIV cohort established by Prof. Anders Sonnerborg in 2004, and by 2008 all HIV-1 clinics in Sweden jointly. There were four groups of patients in the study population (1) people living with HIV on cART ($PLWH_{ART}$)(n=19), (2) HIV Elite Controllers ($PLWH_{EC}$)(n=19), (3) people living with HIV with detectable viremia ($PLWH_{VP}$, n=19) (4) HIV negative controls (n=19). The study was authorized by the regional ethics committees of Stockholm (2013/1944-31/4 and 2009/1485-31) and amendment (2019-05585 and 2019-05584, respectively) and performed following the Declaration of Helsinki.

3.1.3 Paper III

The patients were diagnosed with CCHF and followed up by the Infectious Diseases and Clinical Microbiology clinical service of Sivas Cumhuriyet University Hospital, Sivas, Turkey. The samples were collected from 18 patients on the day they were admitted to the hospital (acute stage) and one year after their recovery. This study was authorized by the Local Research Ethics Committee of the Ankara Numune Education and Research Hospital, Turkey (Protocol # 17-1338) and the Regional Ethics Committee, Stockholm (Dnr. 2017-/1712-31/2).

3.1.4 Paper IV

There were 30 patients diagnosed and hospitalized with CCHFV infections in the study cohort. Data were generated from samples collected at the onset of symptoms (n=22), day of discharge (n=23), and post symptoms onset (n=24). The data was also generated

from CCHFV-seronegative patients as healthy controls (n=22). This study was approved by the Regional Ethics Committee, Stockholm (Dnr. 2017-/1712-31/2) and the Sivas Cumhuriyet University, Turkey. The study design and the sample collection are given in Figure 9.

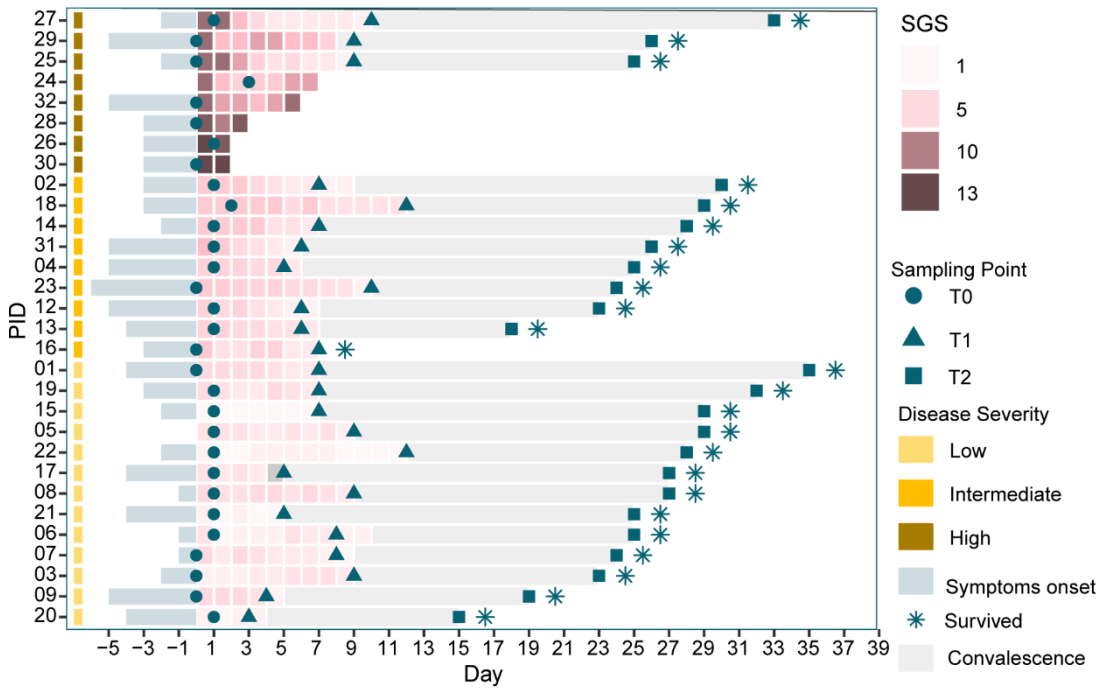


Figure 9: The study design and sample collection details of Paper IV. Adapted from Paper IV and created in R.

3.2 Transcriptomics data analysis

3.2.1 Data generation

In papers I and IV, total RNA was extracted from whole blood samples using TempusTM Spin RNA Isolation Kit (ThermoFisher, USA), followed by library preparation using SMARTer® Stranded Total RNA Sample Prep Kit. Illumina NovaSeq6000 platform in paired-end mode was used for the RNA sequencing. The Peripheral blood mononuclear cells (PBMC) were used as starting material to extract RNA in papers II and III, and the sequencing was performed using Illumina HiSeq2500 or NovaSeq6000.

3.2.2 Raw data processing

In Papers II and III, the quality of the raw sequence data was first assessed by the FastQC tool kit. Based on the FastQC results, low-quality bases and Illumina adapter sequences were removed using Trim Galore version 0.6.1. Low-quality bases were removed using a Phred score of 30 as the cut-off. The cleaned reads were aligned against the human reference genome (Ensembl build) using the aligner STAR version 2.7.3a [118]. Read counting was performed using the module featureCounts from the software subread version 2.0.0 [119]. A nextflow-based nf-core pipeline rnaseq v3.3 [120] was used to

process raw RNASeq reads in Paper I and IV. Gene-level estimated read counts and transcript per million (TPM) normalized count data were used for further downstream analysis.

3.2.3 Digital cell quantification (DCQ)

The transcriptomics data used in papers I, III, and IV were generated from whole blood, and in Paper II was from PBMC. In such cases, the tissue samples contain a mixture of cell types and possess different gene expression profiles unique to each cell type. Information about various cell type proportions in the sample is crucial as they are essential determinants of response to the disease condition. The aforementioned features make the proportion of cell types across samples a vital confounding factor in downstream gene expression data analyses such as differential expression and correlative analysis [121, 122]. We have used a deconvolution-based algorithm adapted from Estimating the Proportions of Immune and Cancer cells (EPIC) method [123] to computationally estimate the cell type proportions from gene expression data generated from bulk tissue samples. We denoted the method as digital cell quantification (DCQ) in the manuscript.

The algorithm of DCQ uses information from cell type-specific reference gene expression profiles for the different cell types to deconvolve. We have used reference cell type-specific gene expression profiles obtained from Human Protein Atlas (HPA) [124]. The reference data consists of all human genes' experimentally determined expression pattern in 18 different blood cell types. Online repositories such as CellMarker [125] and PanglaoDB [126] were used as a reference to derive signature genes corresponding to 18 blood cell types in the reference data. The method then uses normalized gene expression (Transcript per million, TPM) data of each sample as input to estimate the proportion of 18 different blood cell types based on a given reference gene expression profile and list of signature genes corresponding to the cell types. The estimated values were then adjusted while performing differential gene expression analysis.

3.2.4 Differential gene expression and pathway enrichment analysis

Differentially regulated genes were identified using R package DESeq2 v1.26.0 28 [127]. Confounding factors such as age, gender, cell type proportion, and other possible factors were adjusted while performing the analysis. R package RUVSeq v1.28.029 [128] was used to compute unwanted and hidden variation factors in the data. In Papers, I, III, and IV, the pathway enrichment analysis with directionality consideration were performed using R/Bioconductor package PIANO v2.2.030 [129]. KEGG pathway gene sets belonging to metabolism, environmental information processing, and organismal systems categories were chosen as the reference. Pathway enrichment without directionality was executed using the enrichr module from gseapy v0.10.5 [129, 130]. In Paper II, the GSEA v4.1.0 software [131] was used for pathway enrichment analysis, and MSigDB hallmark gene set v7.4 was used as the reference.

3.3 Metabolomics and Proteomics data analysis

Metabolomics data was generated using the Metabolon HD4 (Metabolon Inc, NC, US) platform. Olink Immuno-Oncology Panel (Olink, Sweden) generated the proteomics data. The metabolomics data were log2 scaled, and proteomics data were quantile normalized before analysis. The univariate statistical method Mann-Whitney U test and R package limma [132] were used to study the protein and metabolites enrichment differences wherever appropriate.

3.4 Network-based system biology analysis

3.4.1 Patient re-stratification using multi-omics data

Oxygen consumption rate was used to categorize COVID-19 patients in Paper I. However, we have observed from the previous study that oxygen requirement at hospitalization did not predict mortality [133]. Such single factorial phenotypic characterization may fail to capture molecular signatures, which leads to inaccurate cohort definition. We have used transcriptomics and metabolomics profiles of the patients to re-stratify them and thereby improve the cohort classification. Similarity Network Fusion (SNF) methodology was employed for the multi-omics patient classification.

Similarity network fusion (SNF) works by generating networks of samples based on each available data type and then systematically merging these networks into one that captures the full signatures of the inputted data [134]. The similarity network fusion method used metabolomics and transcriptomics generated from the samples. A preprocessing step was performed on the data before applying SNF. In the pre-processing, features with TPM < 5 (lowly expressed genes) and features having variance less than 0.01 from both transcriptomics and metabolomics data were removed. Further, similarity matrices were generated using the features followed by the application SNF method on the data with settings K=6, T=20, and alpha=0.7.

3.4.2 Weighted network analysis

A weighted gene-metabolite association network was generated in Paper I, and analyzed its structure to find node centralities and communities of interest. Genes and metabolites were the nodes in the network, and association computed using spearman's correlation metric was used to define the edges. A global network was created by associating the features among all the samples. Communities were further detected, and cohort-specific communities were identified by checking the expression of features in the corresponding cohort. TPM normalized transcriptomics data and quantitative metabolomics data were used as inputs. Transcriptomics and metabolomics features were first filtered based on their variance (row variance<0.1) and the expression level (row median TPM<1). Spearman's correlation coefficient was then computed, and the top 10% of the significant positive correlations (FDR<0.00005) were considered to make the network. The correlation

coefficient was used to define the edge weight. Further, the network was analyzed in igraph (https://igraph.org), and degree and betweenness centralities were computed. The modularity maximization by Leiden algorithm [115] was used to detect the communities. Finally, communities were analyzed based on size, centrality (average centralities of the features), and biological function of the features.

In Paper III, a weighted gene co-expression network was created and analyzed. Significant Spearman's correlation between genes was used to define the co-expression. Network was created using genes as nodes and correlation between genes as edges. FPKM normalized transcriptomics data was used as input. The correlation coefficient was used to apply weight to the edges. Firstly, lowly expressed (row median FPKM<1) and lowly varying (row variance<0.1) genes were filtered out from the data. Spearman's correlation was further calculated using the filtered data, and significant positive correlations (adjusted p<0.001) were used to generate the data. The network was then analyzed in igraph (https://igraph.org), and degree centrality was computed. The modularity maximization by the Leiden algorithm [115] was used to detect the communities. Finally, communities were analyzed based on size, centrality (average centralities of the features), and biological function of the features.

A metabolic network consisting of metabolites and associated genes were created and analyzed in both Paper I and Paper II. Metabolite – gene association was derived from the gene-protein-reaction rule depicted in Human-GEM [76]. Metabolic reactions found to have varying flux among the samples (or groups) were chosen for the network creation. Metabolites and genes were used as nodes, and edges were drawn between genes and metabolites in the same reaction. Non-expressed genes (TPM < 1) and currency metabolites were omitted. Metabolic flux value calculated using flux balance analysis was first scaled (-1 to 1) and used as edge weight. The network was then analyzed in igraph (https://igraph.org), and degree and betweenness centrality was computed. Metabolites and genes were further ranked based on their centrality to define their biological importance. The modularity maximization by the Leiden algorithm [115] was used to detect the communities. Finally, communities were analyzed based on size and centrality (average centralities of the features).

3.4.3 Genome-scale metabolic modeling and flux balance analysis

Context-specific genome-scale metabolic models were generated, and flux balance analysis was performed in papers I, II, and IV. TPM normalized data was used to generate the models. Personalized and (or) group-specific metabolic models were contextualized by incorporating transcriptomics data into a human generic metabolic model (Human 1) obtained from the Metabolic Atlas [76]. Individual samples' gene expression data was used to create personal models, and group-specific models were extracted using average gene-expression data. Context-specific metabolic model reconstruction was performed

by applying the Task-driven Integrative Network Inference for Tissues (tINIT) algorithm [85, 90]. An expression threshold of TPM=1 was used in the procedure. The model's biological feasibility was tested after each model's reconstruction. The feasibility was tested by checking the model's ability to perform 57 essential metabolic tasks [90] known to occur in any cell type. The entire analysis was performed using Matlab implementation of the tINIT algorithm (https://github.com/SysBioChalmers/Human-GEM). Since literature evidence suggests the presence of SARS-CoV2 in blood cells, in Paper 1, a viral biomass objective function (VBOF) (developed by Renz, Alina et al.) was added in the model created for SARS-CoV2 samples to emulate the energy demands required by the virus growth [135, 136]. For models created for SARS-CoV2 positive samples, a unidirectional pseudo-reaction consisting of the products of VBOF and ATP hydrolysis as reactants was used as the objective function for flux balance analysis. For models created for the rest of the samples (and for papers II and IV), ATP hydrolysis was used as the objective function. FBA was performed using solveLP function from the RAVEN toolbox v2.4.0 [137]. Constraining of the exchange reaction of the model was performed using plasma metabolomics data as reference. We assumed that the extracellular metabolite availability impacts the fluxes of exchange reactions, i.e., transport reactions are limited by metabolite abundance, effectively approaching first-order kinetics concerning extracellular metabolites. Log2 fold change values of metabolites in each group compared to the control group were calculated and were used proportionally to constrain the reaction bounds.

4 Results and Discussion

As viruses require energy and other biochemical materials to replicate and survive, they try to interfere with the metabolic system of the host to steal the metabolic compounds and energy molecules for their growth and reproduction. Therefore, it has great biological and therapeutic importance for studying host metabolic alterations in response to viral infections. In the thesis, we attempted to unveil the mechanisms associated with the metabolic re-programming in human cells resulting from infections from three different viruses, namely SARS-CoV2, HIV-1, and CCHFV. Advanced network-based system biology approaches were employed to investigate the metabolic system changes at the resolution of metabolic reactions. Studies on each virus were performed in each paper, and the resulting outcomes are explained and discussed here. Results are categorized into three sections as follows,

4.1 Infection-mediated changes in immune cell type abundance

Since the transcriptomics data was generated from bulk tissue samples, the gene expression differences we observe may be because of the differences in gene regulation in each cell type, whose abundances vary between patient groups. This can bias the differential gene expression analysis. Thus, using the digital cell quantification (DCQ) method, we have computationally estimated the abundances of 18 different immune cell types using their marker gene expression level in the input bulk transcriptomics data. The results were used as confounding factors while performing differential expression analysis. The DCQ results can also be used to study the distinct changes in immune cell types caused by infections.

Analysis in Paper-I identified, based on Krushkal-Wall's test, statistically significant (adj. $p<0.05$) differences among SARS-CoV2 positive patient cohorts and SARS-CoV2 negative patients in cell types such as classical monocytes (CM), gamma-delta T cells, neutrophils, and mucosa-associated invariant T (MAIT) cells (Figure 2A, Paper I). The neutrophils and classical monocytes were found to be displayed an increased abundance in SARS-CoV2+ hospitalized-mild and hospitalized severe patient cohorts compared with SARS-CoV2- healthy control cohort. Further, a decreased abundance (adj. $p < 0.05$) was identified in hospitalize-severe patients compared to healthy controls for cell types such as memory B cells, MAIT cells, and gamma-delta T-cells (Figure 2B, Paper I). Additionally, we have checked the pair-wise co-expression landscape of marker genes of specific cell types in all the study cohorts to check the communication disruptions between the cell types due to the viral infection. The analysis found that in the healthy control cohort, the cell types seemed to hold well the association among cell types. But in the case of convalescent, hospitalized-mild, and hospitalized severe patients, the communications were found to be greatly disrupted (Figure 2C, D, E, F, Paper I). Deeper analysis of the co-expression landscape revealed that, during SARS-CoV2 infection, the association

between marker genes of T cells and dendritic cells (DC) and between granulocytes (neutrophils/eosinophils) and DC or T cells are specifically disrupted. This suggests a system-level dysregulation in communication between these cell types due to COVID-19. This may cause failure in efficient adaptive immune response in severe COVID-19 patients by innate immune cells. Further, significant changes in the chemokine signaling pathway, NF-kB, MAPK, and PI3K-Akt signaling pathways were observed in the functional analysis of marker genes that lost associations in COVID-19.

Earlier studies have corroborated the DCQ-predicted increase in neutrophils and classical monocytes in COVID-19 [138]. An increased number of both mature and immature neutrophils was previously identified in the nasopharyngeal epithelium [139], lung [140], and blood [141, 142] in SARS-CoV2 infection. The neutrophil increase can potentially cause hyperinflammation in COVID-19 patients. It was also identified that classical proinflammatory monocytes and neutrophils could mediate myeloid-driven atypical cytokine storm, which can drive to the severity stage [143]. Additionally, the loss of communication between cell types identified in DCQ is indicative of disrupted interaction between T cells and DCs [144] and granulocytes with other immune cells [142], which could explain the improper control of the viral replication [145] causing severity in COVID-19.

In paper II, we checked the differences in cell type abundance associated with HIV-1 infections in people living with HIV with suppressed viremia ($PLWH_{EC}$ and $PLWH_{ART}$) and detectable viremia ($PLWH_{VP}$) in their bodies. We also checked the same in healthy controls for comparisons with negative infection states. Statistical analysis showed several cell types significantly differ in $PLWH_{VP}$ compared to other groups (Figure 1A & S1, Paper II). $PLWH_{EC}$ and HC showed similar levels of cell type proportions in them, while regulatory T-cells (Tregs) were found to be different between $PLWH_{EC}$ and $PLWH_{ART}$ ($p < 0.05$). In conclusion, DCQ results showed the expected results in Paper II. Significant alterations in cell type abundances are found in $PLWH_{VP,}$ and $PLWH_{EC}$ $PLWH_{ART}$ and HC showed almost similar profiles, which indicates the healthy-like status because of treatment efficiency.

Dynamic changes in immune cell populations in response to progressive CCHFV infection are studied in Paper IV. The analysis showed extreme changes in cell type abundances at the acute infection stage. The disruptions were then observed to be coming to normal (similar to HC) at the T2 time point (Figure 1D, Paper IV). The myeloid cells, neutrophils, NK cells, and to some extent, T-cells were most affected during the infection. The co-expression profile of marker genes of specific cell types (Figure 1F, Paper IV) showed disrupted correlations between the signature genes in the immune cells during the acute and early convalescent phase (T0 and T1) and normalized to the level of HC (Figure 1F, Paper IV) at the convalescent phase (T2). This data shows the disrupted immune cell dynamics and dysregulated interaction between these cell types during CCHFV acute infection but repaired back when the infection was cleared.

Such a longitudinal study on CCHFV is not reported earlier in humans. A previous study on Ebola virus infection showed a reduction in the monocytes and CD4 T-cells and an increase in memory CD8 T-cells and NK cells during the acute phase of infection [146]. In contrast, an increase in neutrophil and blood monocyte populations (CM and IM) was observed in SARS-CoV-2 (Paper I). While the early acute phase (T0) of CCHFV infection did not change in classical monocytes, T1 and T2 phases showed an increase. An increase in both pDCs and mDCs is observed at the acute phase (T0) and hospital discharge (T1) but normalized at HC-level at the T2 time point. This suggests the role of DCs during the early acute phase of CCHF infection. At the same time, classical monocytes played the role of effective control and clearance of the CCHFV infection during the recovery phase.

4.2 Virus-specific omics signatures

The main objective of the papers in the thesis was to investigate the system-wide host responses following infections from three different viruses. For this purpose, genome-wide bulk transcriptomics data was generated from patent samples and analyzed. Standard pipelines for transcriptomics data analysis were employed and identified differentially expressed genes and pathways linked with the three viruses. Additionally, weighted network analysis was also performed to identify virus-specific omics signatures.

In Paper I, firstly, the spatial distribution of the samples based on their gene expression profile was checked. UMAP-based distribution showed a clear separation of the COVID-19+ patients from HC and convalescent controls, but no clear separation was observed between hospitalized-mild and -severe patients (Figure 4 A, Paper I). Further, differentially expressed genes were identified. A threshold of adj. $p < 0.05$ and $|\log2\text{-fold change}| >= 1.5$ were applied to define significantly expressed genes. Analysis between HC and COVID-19 identified 581 genes significantly expressed, while analysis between hospitalized-mild and hospitalize-severe groups identified 154 genes expressed (Figure 4B, Paper I). Comparison between HC and hospitalized-mild and severe groups also identified 445 and 1,068 genes significantly expressed. Sample to samples similarities based on the gene expression pattern of significantly expressed genes was also accessed by computing Euclidean distance metric and performing hierarchical clustering. The procedure also showed a distinct pattern in the groups and heterogeneity among the COVID-19 patients (Figure 4D, Paper I). Functions of the significantly expressed genes were investigated by performing gene-set enrichment analysis. The analysis identified upregulations of pathways related to antiviral response (e.g., NOD-like receptor signaling, RIG-I-like signaling, and TNF signaling pathways), metabolism (e.g., fructose mannose metabolism, oxidative phosphorylation (OXPHOS), and pentose phosphate pathway), and pathways like C-type lectin receptor signaling pathway, complement, and coagulation cascades that are related to thromboembolism in COVID-19 patients compared to HC (Figure 4E,

Paper I). Interestingly, pathways such as PI3K-Akt, mTOR, AMPK, and HIF-1 signaling were found to be upregulated in severe patients compared to mild (Figure 4E, Paper I). These pathways are the key regulators of central carbon and energy metabolism (e.g., TCA cycle, OXPHOS, and pyruvate metabolism) and are regulated by the SARS-CoV-2 replication [57, 147, 148]. In the study, we have additionally re-stratified the patient groups exclusively based on metabolomics and transcriptomics signatures by applying similarity network fusion (SNF) methodologies. The process gave rise to four clusters and named them SNF 1,2,3,4 (Figure 5C and D, Paper I). we defined SNF-1 as COVID-19 mild/moderate, SNF-3 as COVID-19 severe, and SNF-4 as HC based on the number of samples using their original clinical definition. SNF-2 was found to be containing mixed types of samples, and thus omitted from the analysis. Weighted metabolite-gene association network analysis identified nine communities (c1–c9) (Figure 5G, Paper I). The community 1 (c1) was found to be specific to the severity group (SNF-3). Gene-set enrichment analysis of c1 found that it was associated with NOD-like receptor signaling, chemokine signaling, Fc gamma R-mediated phagocytosis, and platelet activation pathways, and other pathways such as NF-κ B, Notch, RIG-I, HIF-1, and FoxO signaling (Figure 5H, Paper I). Upregulation of pathways related to the antiviral response and metabolism in the COVID-19 patients and upregulation of pathways that are master regulators of central carbon and energy metabolism (e.g., TCA cycle, OXPHOS, pyruvate metabolism, etc.), the PI3K-Akt, mTOR, AMPK, and HIF-1 signaling in severe COVID-19 patients were shown in previous studies [57, 148].

Differential gene expression and pathway enrichment analysis were performed in Paper II to understand host response mechanisms associated with HIV-1 infection. The analysis was performed between each pair of the total four cohorts. The comparison between PLWHEC and HC did not show any genes significantly expressed (adj. $p<0.05$), whereas the comparison between $PLWH_{ART}$ and HC identified 949 genes significantly expressed. The study's main aim was to investigate the immune signatures during suppressive viremia that is naturally controlled or induced by cART. For this purpose, we performed differential expression analysis between $PLWH_{EC}$ and $PLWH_{ART}$. The analysis found 1,061 genes significantly expressed, out of which 400 genes were up-regulated and 661 genes were down-regulated in $PLWH_{ART}$ compared to $PLWH_{EC}$ (Figure 2B, Paper II). Additionally, hierarchical clustering based on the similarity between the samples showed distinct expression patterns in the two cohorts (Figure 2C, Paper II). Further, gene-set enrichment analysis showed that oxidative phosphorylation (OXPHOS) and reactive oxygen species (ROS) pathways are up-regulated in $PLWH_{ART}$ (Figure 2D, Paper II). The analysis did not find any pathways significantly downregulated. mTORC1 signaling and glycolysis pathways were also identified as upregulated in $PLWH_{ART}$ but without passing the significance threshold. The OXPHOS pathway was further looked at more deeply by checking the expression level of individual genes in the pathway. This identified that genes in complexes I III, and IV are mostly upregulated in $PLWH_{ART}$ compared with $PLWH_{EC}$. In

addition to the pair-wise differential expression analysis, we have also derived list of genes dysregulated uniquely in $PLWH_{ART}$ by applying basic set operation in the set of differential analysis results of all comparisons. The analysis provided 1037 genes as cART-specific, and gene-set enrichment analysis on them also identified OXPHOS and ROS pathways as top significantly enriched. A previous study has shown that OXPHOS is positively correlated with higher HIV-1 viral set point in untreated patients during acute HIV-1 infection, and viral replication was suppressed by in vitro pharmacological inhibition of complex I (by rotenone or metformin) and complex III (by antimycin A) [149]. Further, another study has observed that cell death in lymphocytic and pre-monocytic HIV-1 latent cell models was increased by blocking glycolysis with 2-deoxyglucose (2-DG) [149-151]. The studies also displayed a significant role of glycolysis and OXPHOS during HIV-1 pathogenesis. Several other studies have also indicated the crucial role of OXPHOS complex I, III, and IV in ATP generation for cellular energy requirement and the role of complex I and III n ROS production that leads to lymphocyte activation, proliferation, and differentiation of inflammatory macrophages and T helper 17 (Th17) cells [152].

Papers III and IV dealt with CCHFV infection, and transcriptomics data was generated from longitudinal samples and analyzed to understand the temporal dynamics of the CCHFV infection. In Paper III, differential expression analysis was performed between the acute stage of infection and 1-year post-recovery. The analysis found 2,891 genes upregulated and 2738 genes downregulated (adj. $p < 0.05$) at the acute phase compared to the recovered phase. Pathway enrichment analysis further identified the up-regulation of metabolic pathways such as one carbon pool by folate, oxidative phosphorylation (OXPHOS), glycolysis, N-glycan biosynthesis, and antiviral pathways like the NOD-like receptor signaling pathway (Figure 1B, Paper III) and downregulation of antiviral defense mechanism-associated pathways including innate immune responses like Th1, Th2, and Th17 cell differentiation, the NF-kB pathways, chemokine signaling pathway, etc. (Figure 1B, Paper III) in the acute phase. The study also performed reporter metabolite (The metabolites around which the majority of transcriptional changes are happening) analysis which identified the up-regulation of metabolites that are associated with OXPHOS, TCA-cycle, nucleotide metabolism, N-glycan metabolism, and amino acid-related pathways (Figure 1C, Paper III). To specifically study the association of disease severity with gene expression, analysis was performed using samples in severity group 1 (SG-1) and severity group 2 (SG-2) + severity group 3 (SG-3) in the acute phase. The analysis found 12 genes (ERG, PROM1, HP, HBD, AHSP, CTSG, PPARG, TIMP4, SMIM10, RNASE1, VSIG4, CMBL, MT1G) significantly upregulated in patients in the SG-2 and SG-3 combination group compared to SG-1. But no obvious association among these genes was observed. However, a distinct differential expression profile was found when analyzing the acute phase with the recovered phase in SG-1 and SG-2 separately. Functional analysis showed upregulation of the IFN-I signaling pathway (GO:0060337) and the regulation of viral genome replication (GO:0045069) (Figure 1F), which showed that the disease severity impacted

the gene expression of the interferon signaling pathway profiling during the acute phase, whereas it was comparable when they recovered. Additionally, weighted co-expression analysis in acute phase samples identified seven communities (c1-7). The most central community (c1) was found to be significantly associated with pyruvate metabolism, TCA-cycle, and to a smaller extent to, glycolysis. The analysis further showed negative correlations between community (c1) and those associated with Notch, mechanistic target of rapamycin (mTOR) and Forehead box protein O (FoxO) signaling (c5), and hypoxia inducing factor-1 (HIF-1) signaling (c7). The OXPHOS-associated community (c3) was also found to be negatively correlated with those involved in Notch/mTOR/FoxO signaling (c5) and HIF-1 signaling (c7) (Figure 3B and C, Paper III).

In Paper IV, the temporal dynamics of CCHFV infection were studied using longitudinal data generated from three-time points of infection, at acute phase (T0), at discharge (T1), and 30 days post symptom onset (T2). The transcriptomics time-series analysis identified 2,504 genes significantly expressed across the three-time points. To understand their temporal expression dynamics, the time series clustering of the 2,504 genes identified four expression trajectory modules (M1-M4) (Figure 2b, and Extended Data Table 2, Paper IV). The expression of the genes (n=879) in M1 initially decreased slowly between T0 and T1 time points and then decreased rapidly (Figure 2b, Paper IV). For genes (n=531) in M4, the expressions decreased drastically at the beginning and then slowly decreased until point T2. In contrast, the genes in M2 (n=634) gradually increased over time, while in M3 (n=460), the gene expression was first increased at the T1 while decreasing at the T2 (Figure 2b, Paper IV). Gene-set enrichment analysis of modules identified that M1 was associated mainly with metabolic pathways such as one carbon pool by folate and pyruvate metabolism, and M3 was found to be linked with metabolic pathways like OXPHOS, glucagon signaling pathways and adaptive immune response pathways like Th1 and Th2 receptor signaling processes indicate a dynamic nature of the metabolic process and adaptive immune response in CCHFV-pathogenesis. Further, M4 showed an association with innate antiviral response pathways like RIG-I-like receptor, NOD-like receptor, TNF signaling, and chemokine signaling, NF-kB signaling was highly exertive at the acute phase of the disease (T0). Their demand decreased rapidly as the adaptive immunity came into play (T1), and after that, a slight decrease to the normal level.

Our earlier study showed system-level host metabolic reprogramming towards central carbon metabolism with upregulation of oxidative phosphorylation (OXPHOS) during acute CCHFV infection [Paper I]. This is supported by the transcriptomics or proteomics study performed in-vitro infection models in cancer cell lines [153] and in vivo infection studies in non-human primates [154] and mice [155], displaying changes of the antiviral response, upregulation in the interferon (IFN) pathways, and disruptions of the metabolic process. While temporal quantitative proteomics analysis in the human hepatocellular carcinoma cell line, Huh7 infection model reported by us showed dynamic changes in the

interferon stimulating genes (ISGs) and metabolic process during the progressive replication [Paper I], to our knowledge to date, no study has investigated a longitudinal immune response against CCHFV pathogenesis in patients. The cumulative information obtained from both in vitro and patients can provide details on the viral pathogenesis and the cellular response during different phases of infection.

4.3 Metabolic dysregulations in infections

Transcriptomics analysis of all four papers identified severe disruption in response to the viral infection. Therefore, we further performed flux balance analysis on context (disease) specific genome-scale metabolic models that represent metabolic networks of cells in the blood (or PBMC) to find the specific metabolic reactions in the metabolic pathways which are altered upon virus entry. We also performed centrality-based prioritization using the metabolic networks created to find the top genes and metabolites which are contributing to the metabolic re-programming against the disease.

In Paper I, we have generated context-specific metabolic models for each SNF cluster (group-specific modeling) and for individual samples (personalized modeling). Flux balance analysis on SNF cluster-specific models identified 100 reactions with divergent flux among the SNF clusters. Also found 15 transporter reactions with distinct flux in SNF-1 (mild/moderate) and SNF-3 (severe) compared with SNF-4 (HC). The transporter reactions consisted of the transport of TCA-cycle intermediates such as cis-aconitate, α-ketoglutarate, succinate, malate, and fumarate between cytosol and mitochondria. Most of the transporter reactions had flux only in SNF-1 (mild/moderate) and SNF-3 (severe), displaying efficient transport of the metabolites in COVID-19 patients, which then possibly feed the TCA-cycle pathway (Figure 6 A, Paper I). We further checked the expression of transporter genes of TCA cycle intermediates in different cell types in COVID-19 using publicly available scRNASeq data. The analysis found that the transporter genes are mainly expressed in monocytes. These results were further corroborated by the results given by analyzing sctMetabolomics data targeting TCA-cycle intermediates in monocytes. FBA on personalized models found 274 metabolic reactions with unique fluxes across patients (Figure 6D, Paper I); most of them recapitulated the observations from group-specific FBA. Further, metabolic network topology-based analysis was performed to rank the genes and metabolites to measure their importance in the COVID-19 disease response. Betweenness centrality was used for the ranking. The analysis found genes and metabolites, namely monocarboxylate transporter SLC16A6 and nucleoside transporter SLC29A1 and metabolites such as α-ketoglutarate in the cytosol, succinate, and malate in mitochondria and cytosol, and butyrate in the cytosol and extracellular space as top-ranked features in COVID-19 disease response.

In addition to group-specific FBA, the study enabled us to identify and prioritize genes and metabolites at personal level as well. Here we performed an extensive

characterization of COVID-19 at a personalized level using transcriptomic data combining information from biologically meaningful metabolic flux predictions and topological network analysis. Our results showed disrupted flux in mitochondrial TCA cycle intermediates that display the altered central metabolic pathways in COVID-19 and could possibly be associated with the glycolytic modulation by mTOR/HIF-1 signaling and mitochondrial dysfunction [148]. We also showed that, viral replication and production are inhibited by blocking glycolysis and glutaminolysis pathways [148].

Genome-scale metabolic model contextualization and FBA were performed in Paper II to identify the metabolic reactions altered followed by HIV-1 infection. Also, metabolic network analysis was employed to prioritize the genes and metabolites for their importance in the existence of metabolic network coordination towards the HIV-1 infection. Here, we performed group-specific metabolic modeling and FBA, and the analysis found 80 reactions uniquely disrupted in $PLWH_{ART}$ compared with $PLWH_{EC}$ and HC cohorts. Those reactions were associated with amino acid, carbohydrate, and energy metabolism pathways and various transporter reactions. The energy metabolism pathways involving the tricarboxylic acid (TCA) cycle, glycolysis, glutaminolysis, and OXPHOS, were altered in $PLWH_{ART}$ (Figure 3C, Paper II). The OXPHOS reaction converting ADP to ATP showed positive flux in $PLWH_{ART,}$ whereas no flux was identified in $PLWH_{EC}$ and HC, which shows a higher energy requirement in $PLWH_{ART}$. Increased production of α-ketoglutarate (αKG) in the cytoplasm and glutamate production was also observed in $PLWH_{ART}$. Two other reactions were also found to be active in $PLWH_{ART,}$ which increased the production of αKG in the cytoplasm and that can potentially feed the TCA cycle in mitochondria. Feature ranking (based on centrality) in the metabolic network further identified fructose-6-phosphate, OAA, glutamate, and pyruvate as top features in $PLWH_{ART,}$ showing the role of the TCA cycle and glycolysis in differentiating $PLWH_{ART}$ from $PLWH_{EC}$ and HC.

It is observed from the FBA that there is a change in metabolic flux in pyruvate, glutamate, and αKG in the $PLWH_{ART}$ compared with $PLWH_{EC}$ and HC. A higher level of glutamate in $PLWH_{ART}$ in several cohorts compared with HC was found by us recently [150]. Another study also suggested an increased level of glutamate in PLWH with dementia [156]. The interlink between glutamate and pyruvate and its neuroprotective role in chronic HIV-1 infected patients on therapy needs further studies to understand neurological complications in HIV-1 infection after successful treatment.

Context-specific genome-scale modeling and FBA were employed in Paper IV to study the temporal metabolic re-programming against CCHFV infection at the resolution of individual reactions. The FBA found 546 metabolic reactions altered during infection and in HC (Figure 3A, Paper IV). The reactions were associated with central carbon metabolism (CCM) and the transport of the metabolites related to the CCM. High demand for the amino acids (AA) during the acute phase of infection was indicated by the altered flux of

essential (leucine and threonine) and non-essential amino acids (arginine, alanine, and glutamine) in T0 compared to T1, T2, and HC (Figure 3B, Paper IV). The transportation of TCA cycle compounds such as pyruvate, isocitrate, and alpha-ketoglutarate was also found to be altered (Figure 3C, Paper IV). Disrupted pentose phosphate pathway (PPP) and metabolic shift towards fatty acid oxidation were also identified during the T0 phase (Figure 3C, Paper IV). The mitochondrial TCA cycle was also changed, including the additional ATP creation using OXPHOS (Figure 3C, Paper IV), and the transports of amino acids were also significantly associated (adjusted $p<0.05$) with metabolic reactions (Figure 3d). The results altogether suggest a high transportation of essential and non-essential amino acids, a shift towards fatty acid oxidation, and extra ATP production through OXPHOS as the hallmark during the acute phase of infection that normalized to a certain extent during the recovery phase.

The study reports a robust metabolic change towards the central carbon metabolisms (e.g., pyruvate metabolism, OXPHOS) and its connection with the antiviral response. The severely affected transports of amino acids and the key TCA cycle intermediates during the acute phase of infection show a hypermetabolic state during the infection. Interestingly, the altered pentose phosphate pathway and metabolic shift towards fatty acid oxidation and the extra ATP production through OXPHOS during the T0 phase indicate dysregulated energy metabolism to give more energy to the cells to suppress the infection. This hypermetabolic state during the acute phase causes metabolic insufficiency with altered signaling cascades at the convalescent phase despite restoring the metabolic process. The observation corroborates that the essential metabolic processes (e.g., glycolysis, TCA cycle, and OXPHOS) and the regulatory signaling cascades (Akt/mTOR/HIF-1 signaling) were downregulated at T2 than HC.

4.4 Host metabolic reaction changes in acute SARS-CoV2 and CCHF infection

Additionally, the metabolic reactions altered specifically at acute stages of SARS-CoV2 and CCHF infections were compared to investigate the similarities and dissimilarities in the host metabolic reprogramming. Figure 10 shows the metabolic reactions part of central carbon metabolism pathways and transport reactions that are found to be changed. It is observed that acute COVID-19 activates several reactions that transport metabolites part of TCA cycle from the cytoplasm to mitochondria. Whereas in acute CCHF, few of those reactions are inactive; instead, it showed activation of oxalacetic acid

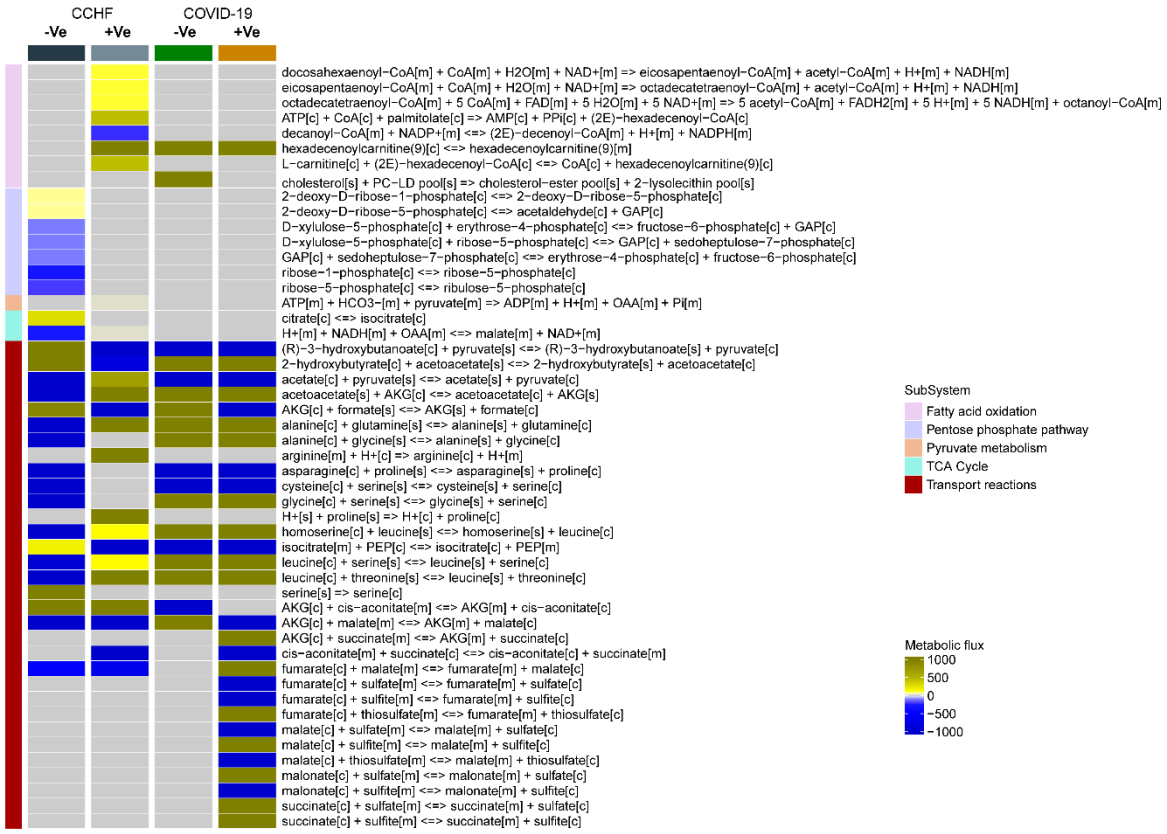


Figure 10: Metabolic reactions and associated flux computed from FBA at the acute stage of SARS-CoV2 and CCHFV infections and at corresponding negative controls. Created in R.

(OAA) to malate conversion in the TCA cycle and OAA production in pyruvate metabolism. Also, several of the reactions of fatty acid oxidation and pentose phosphate pathways were uniquely active in acute CCHF.

These results show the strength of network-based systems biology analysis in deducing the metabolic responses at the stage of individual reactions. Standard transcriptomics can provide only general information about the metabolic pathways being disrupted but cannot pinpoint the exact reaction of the affected pathways. This information is vital to therapeutically target them for treating the disease.

5 Conclusions and Future Directions

We have successfully performed advanced network-based system biology analysis and conventional single-layer omics analysis to identify biologically and therapeutically important information associated with viral infections. Paper, I utilized a combination of multi-modal systems-wide transcriptomics, DCQ, and immunophenotyping of MNPs. We observed the crucial role of the coordination of immune cells in COVID-19 severity. The novel data-driven patient stratification recapitulated many previously known clinical properties and enabled us to unveil the critical mechanical consequences of COVID-19 infection in immune cells. A progressive dysregulation of the central metabolic pathway (TCA cycle) concomitant with COVID-19 disease severity was identified by system biology methodologies such as network topology analysis and personalized GEMs, sctMetabolomics of monocytes, and gene expression mapping using published scRNA-seq data. COVID-19 disease severity at individual and group levels was associated with alterations in central carbon and energy metabolism, TCA-cycle intermediates like malate and α-ketoglutarate, and expression of metabolite transporters in monocytes. These observations suggest the metabolic reprogramming in monocytes in specific COVID-19 patient groups and potentiate personalized targets for treatment in severe patients.

A system-level up-regulation of OXPHOS and, to a certain extent, glycolysis in $PLWH_{ART}$ compared with the $PLWH_{EC}$ were observed in the study on HIV-1 infection. We further displayed how the dynamics of latent reservoir and immunosenescence in HIV-1–infected individuals with long-term successful therapy were impacted by the upregulation of OXPHOS. The study also showed the role of pharmacological inhibition of the OXPHOS complexes in latency reversal, apoptotic properties, and immunosenescence in latently infected cells. We concluded Paper II by warranting further studies investigating the molecular mechanisms underlying the observed shift in OXPHOS in $PLWH_{ART}$ and how its association with glutaminolysis can cause immune dysregulation during successful therapy.

Paper III concludes that the extensive transcriptomics analysis explains the host-immune response against CCHFV that can describe viral pathogenesis. Further, the study shows that efficient options for therapeutic intervention of CCHF could be provided by the interplay of the metabolic reprogramming toward the central carbon and energy metabolism and its negative association with biological signaling pathways like Notch/FoxO and PI3K/mTOR/HIF-1 and the IFN-mediated host antiviral mechanisms. More investigations on the role of mitochondrial biogenesis and dynamics in CCHFV infection, replication, and pathogenesis will increase our understanding of host-virus interactions, which can lead to the development of new antiviral strategies.

The analysis in Paper IV precisely identified the temporal antiviral response following the CCHFV infection and metabolic rearrangement. The study shows that the exhausted phenotype upon recovery can be promoted by dysregulation and hyperactivity of the critical metabolic process of the central carbon and energy metabolism and metabolic flux related to the amino acid metabolism during the acute progressive phases of CCHFV infection. Early recovery from the infection and quality of life improvement can be achieved by reprogramming the impaired metabolic pathways to improve the antiviral mechanism by harnessing immuno-metabolic regulations.

Though we have performed novel system biology methodologies and identified important host-viral entanglement and immunopathogenesis that can lead to therapeutic information, there is still room for improvement. All the studies presented here used bulk transcriptomics data representing the gene expression corresponding to a mixture of cells in the tissue. This could bias the analysis as they can significantly impact the expression change we see from the analysis. The application of single-cell transcriptomics technology can successfully overcome this issue. Moreover, the genome-scale metabolic model generated in the studies used bulk transcriptomics data, representing a metabolic model of a mixture of cells. This could be biased because metabolism is different in various cell types. Metabolic modeling using single-cell transcriptomics data is more advisable to generate biologically more meaningful models accurately. Another issue regarding flux balance analysis for models representing virus-infected tissues is the possible presence of viruses in the material. If the virus presence is detected in the material, incorporating viral biomass function in the model is necessary to adjust the metabolic requirement of the virus. In Paper I, we have used previously developed viral biomass function and development for other viruses recommended in future studies. Proper constraining of nutrient availability is also a crucial factor in flux balance analysis. Ideally, this should be done using temporal metabolomics data as a reference. Methodological advancement is required to accurately measure time-dependent metabolite concentration to be used for better analysis.

6 Acknowledgments

I want to express my sincere gratitude to my principal supervisor **Ujjwal Neogi**. I am thankful for his selecting me as his student and his supportive guidance during my Ph.D. I admire his aspirations towards learning new techniques outside of his background. Moreover, he has been a kind teacher in sharing his knowledge with everyone on the team. He has also let me participate in many collaborations that helped me meet other technology experts and expand my knowledge. I am grateful that his supervision and support helped me develop my independent research skill and successfully finish my studies.

My co-supervisor **Rui Benfeitas** has also been a great teacher to me. The techniques and knowledge I have learned under his guidance became the backbone of my thesis. I appreciate him being accessible and helpful even remotely. Furthermore, I am truly thankful that he helped me grow my technical skills, thereby increasing the quality of my thesis.

I want to thank my co-supervisor **Piotr Nowak** for his supervision. His expertise helped me to produce quality work, and he has been very supportive during my Ph.D.

I also want to thank **Soham Gupta**, who taught me several pieces of knowledge and helped interpret the results. I am thankful for his contribution, which significantly increases my papers' quality. I appreciate his kindness and support as a fellow researcher.

Next, I want to thank my colleagues **Sara Svensson Akusjärvi**, **Shuba Krishnan**, and **Maike Sperk**, who contributed significantly to the papers. I appreciate their wet laboratory skill and knowledge. They were also supportive and helped to learn and interpret the results. Also, I thank **Flora Mikaeloff**, **Xi Chen**, **Alejandra Escos Lopez,** and **Negin Nikouyan** for their support and being great co-workers. Further, I thank all the project collaborators and co-authors. I thank **Nazif Elaldi**, **Akos Vegvari** and **Jimmy E Rodriguez** for supporting the projects and teaching several techniques. Their contributions were vital for all the projects, and I appreciate this. I also thank the previous group members, **Shambhu Aralaguppe**, **Wang Zhang**, and **Duncan Njenda**, for being supportive and great friends.

I want to express my gratitude to the study subjects. Without their contribution and participation, the research would not be possible.

I am indebted to **Milner Kumar M** for teaching me NGS techniques, mentoring, and his recommendation that helped my career growth.

Above all, I am deeply grateful to my parents. Their prayer and love were the sole reason for my better life and successful career. Most importantly, I want to thank my lovely wife, who has been a supportive partner during the journey of the Ph.D.

7 References

1. Chapman, N.M. and H. Chi, *Metabolic adaptation of lymphocytes in immunity and disease*. Immunity, 2022. **55**(1): p. 14-30.
2. Howard, C.R. and N.F. Fletcher, *Emerging virus diseases: can we ever expect the unexpected?* Emerg Microbes Infect, 2012. **1**(12): p. e46.
3. Woolhouse, M.E. and S. Gowtage-Sequeria, *Host range and emerging and reemerging pathogens*. Emerg Infect Dis, 2005. **11**(12): p. 1842-7.
4. Kobayashi, N., *Impact of Emerging, Re-Emerging and Zoonotic Viral Infectious Diseases, in a Virologist's Perspective*. Open Virol J, 2018. **12**: p. 131-133.
5. Morens, D.M., G.K. Folkers, and A.S. Fauci, *The challenge of emerging and re-emerging infectious diseases*. Nature, 2004. **430**(6996): p. 242-9.
6. Lamers, M.M. and B.L. Haagmans, *SARS-CoV-2 pathogenesis*. Nat Rev Microbiol, 2022. **20**(5): p. 270-284.
7. Drosten, C., et al., *Identification of a novel coronavirus in patients with severe acute respiratory syndrome*. N Engl J Med, 2003. **348**(20): p. 1967-76.
8. Peiris, J.S., et al., *Coronavirus as a possible cause of severe acute respiratory syndrome*. Lancet, 2003. **361**(9366): p. 1319-25.
9. Zhu, N., et al., *A Novel Coronavirus from Patients with Pneumonia in China, 2019*. N Engl J Med, 2020. **382**(8): p. 727-733.
10. Lauer, S.A., et al., *The Incubation Period of Coronavirus Disease 2019 (COVID-19) From Publicly Reported Confirmed Cases: Estimation and Application*. Ann Intern Med, 2020. **172**(9): p. 577-582.
11. Guan, W.J., et al., *Clinical Characteristics of Coronavirus Disease 2019 in China*. N Engl J Med, 2020. **382**(18): p. 1708-1720.
12. Wang, D., et al., *Clinical Characteristics of 138 Hospitalized Patients With 2019 Novel Coronavirus-Infected Pneumonia in Wuhan, China*. JAMA, 2020. **323**(11): p. 1061-1069.
13. Huang, C., et al., *Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China*. Lancet, 2020. **395**(10223): p. 497-506.
14. Goh, K.J., et al., *Rapid Progression to Acute Respiratory Distress Syndrome: Review of Current Understanding of Critical Illness from Coronavirus Disease 2019 (COVID-19) Infection*. Ann Acad Med Singap, 2020. **49**(3): p. 108-118.
15. Chen, N., et al., *Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study*. Lancet, 2020. **395**(10223): p. 507-513.
16. O'Driscoll, M., et al., *Age-specific mortality and immunity patterns of SARS-CoV-2*. Nature, 2021. **590**(7844): p. 140-145.
17. Grasselli, G., et al., *Baseline Characteristics and Outcomes of 1591 Patients Infected With SARS-CoV-2 Admitted to ICUs of the Lombardy Region, Italy*. JAMA, 2020. **323**(16): p. 1574-1581.

18. Karagiannidis, C., et al., *Case characteristics, resource use, and outcomes of 10 021 patients with COVID-19 admitted to 920 German hospitals: an observational study*. Lancet Respir Med, 2020. **8**(9): p. 853-862.

19. Yu, Y., A.J. Clippinger, and J.C. Alwine, *Viral effects on metabolism: changes in glucose and glutamine utilization during human cytomegalovirus infection*. Trends Microbiol, 2011. **19**(7): p. 360-7.

20. Aller, S., et al., *Integrated human-virus metabolic stoichiometric modelling predicts host-based antiviral targets against Chikungunya, Dengue and Zika viruses*. J R Soc Interface, 2018. **15**(146).

21. Gilks, C., *Role of communities in AIDS response*. Indian J Med Res, 2019. **150**(6): p. 515-517.

22. Reeves, J.D. and R.W. Doms, *Human immunodeficiency virus type 2*. J Gen Virol, 2002. **83**(Pt 6): p. 1253-1265.

23. Sharp, P.M. and B.H. Hahn, *Origins of HIV and the AIDS pandemic*. Cold Spring Harb Perspect Med, 2011. **1**(1): p. a006841.

24. Moir, S., T.W. Chun, and A.S. Fauci, *Pathogenic mechanisms of HIV disease*. Annu Rev Pathol, 2011. **6**: p. 223-48.

25. Hunt, P.W., et al., *T cell activation is associated with lower CD4+ T cell gains in human immunodeficiency virus-infected patients with sustained viral suppression during antiretroviral therapy*. J Infect Dis, 2003. **187**(10): p. 1534-43.

26. Battegay, M., et al., *Antiretroviral therapy of late presenters with advanced HIV disease*. J Antimicrob Chemother, 2008. **62**(1): p. 41-4.

27. French, M.A., *HIV/AIDS: immune reconstitution inflammatory syndrome: a reappraisal*. Clin Infect Dis, 2009. **48**(1): p. 101-7.

28. Zhang, W., et al., *Transcriptomics and Targeted Proteomics Analysis to Gain Insights Into the Immune-control Mechanisms of HIV-1 Infected Elite Controllers*. EBioMedicine, 2018. **27**: p. 40-50.

29. Akusjarvi, S.S., et al., *Integrative proteo-transcriptomic and immunophenotyping signatures of HIV-1 elite control phenotype: A cross-talk between glycolysis and HIF signaling*. iScience, 2022. **25**(1): p. 103607.

30. Dzikwi-Emennaa, A.A., et al., *Detection of Crimean-Congo Hemorrhagic Fever Virus Antibodies in Cattle in Plateau State, Nigeria*. Viruses, 2022. **14**(12).

31. Bente, D.A., et al., *Crimean-Congo hemorrhagic fever: history, epidemiology, pathogenesis, clinical syndrome and genetic diversity*. Antiviral Res, 2013. **100**(1): p. 159-89.

32. Chinikar, S., et al., *Crimean-Congo hemorrhagic fever in Iran and neighboring countries*. J Clin Virol, 2010. **47**(2): p. 110-4.

33. Hoogstraal, H., *The epidemiology of tick-borne Crimean-Congo hemorrhagic fever in Asia, Europe, and Africa*. J Med Entomol, 1979. **15**(4): p. 307-417.

34. Ergonul, O., *Crimean-Congo haemorrhagic fever*. Lancet Infect Dis, 2006. **6**(4): p. 203-14.

35. Swanepoel, R., et al., *The clinical pathology of Crimean-Congo hemorrhagic fever*. Rev Infect Dis, 1989. **11 Suppl 4**: p. S794-800.

36. Neogi, U., et al., *Multi-omics insights into host-viral response and pathogenesis in Crimean-Congo hemorrhagic fever viruses for novel therapeutic target*. Elife, 2022. **11**.

37. Vaheri, A. and V.J. Cristofalo, *Metabolism of rubella virus-infected BHK 21 cells. Enhanced glycolysis and late cellular inhibition.* Arch Gesamte Virusforsch, 1967. **21**(3): p. 425-36.

38. Singh, V.N., et al., *Alterations in glucose metabolism in chick-embryo cells transformed by Rous sarcoma virus: intracellular levels of glycolytic intermediates*. Proc Natl Acad Sci U S A, 1974. **71**(10): p. 4129-32.

39. Bardell, D. and M. Essex, *Glycolysis during early infection of feline and human cells with feline leukemia virus*. Infect Immun, 1974. **9**(5): p. 824-7.

40. Levy, H.B. and S. Baron, *The effect of animal viruses on host cell metabolism. II. Effect of poliomyelitis virus on glycolysis and uptake of glycine by monkey kidney tissue cultures*. J Infect Dis, 1957. **100**(2): p. 109-18.

41. Fisher, T.N. and H.S. Ginsberg, *The reaction of influenza viruses with guinea pig polymorphonuclear leucocytes. II. The reduction of white blood cell glycolysis by influenza viruses and receptor-destroying enzyme (RDE).* Virology, 1956. **2**(5): p. 637-55.

42. Klemperer, H., *Glucose breakdown in chick embryo cells infected with influenza virus*. Virology, 1961. **13**: p. 68-77.

43. Salzman, N.P., R.Z. Lockart, Jr., and E.D. Sebring, *Alterations in HeLa cell metabolism resulting from poliovirus infection*. Virology, 1959. **9**: p. 244-59.

44. Bardeletti, G., *Respiration and ATP level in BHK21/13S cells during the earlist stages of rubella virus replication*. Intervirology, 1977. **8**(2): p. 100-9.

45. Gaitskhoki, V.S., et al., *[Reconstruction of autonomic genetic and protein-synthesizing systems from viral RNA and isolated mitochondria].* Dokl Akad Nauk SSSR, 1971. **201**(1): p. 220-3.

46. Jamison, D.A., Jr., et al., *A comprehensive SARS-CoV-2 and COVID-19 review, Part 1: Intracellular overdrive for SARS-CoV-2 infection*. Eur J Hum Genet, 2022. **30**(8): p. 889-898.

47. Grohmann, U. and V. Bronte, *Control of immune response by amino acid metabolism*. Immunol Rev, 2010. **236**: p. 243-64.

48. Lionetto, L., et al., *Increased kynurenine-to-tryptophan ratio in the serum of patients infected with SARS-CoV2: An observational cohort study*. Biochim Biophys Acta Mol Basis Dis, 2021. **1867**(3): p. 166042.

49. Rodriguez, P.C., D.G. Quiceno, and A.C. Ochoa, *L-arginine availability regulates T-lymphocyte cell-cycle progression*. Blood, 2007. **109**(4): p. 1568-73.

50. Reizine, F., et al., *SARS-CoV-2-Induced ARDS Associates with MDSC Expansion, Lymphocyte Dysfunction, and Arginine Shortage*. J Clin Immunol, 2021. **41**(3): p. 515-525.

51. Kimhofer, T., et al., *Integrative Modeling of Quantitative Plasma Lipoprotein, Metabolic, and Amino Acid Data Reveals a Multiorgan Pathological Signature of SARS-CoV-2 Infection*. J Proteome Res, 2020. **19**(11): p. 4442-4454.

52. Sanchez, E.L. and M. Lagunoff, *Viral activation of cellular metabolism*. Virology, 2015. **479-480**: p. 609-18.

53. Jia, H., et al., *Metabolomic analyses reveal new stage-specific features of COVID-19*. Eur Respir J, 2022. **59**(2).

54. Ajaz, S., et al., *Mitochondrial metabolic manipulation by SARS-CoV-2 in peripheral blood mononuclear cells of patients with COVID-19*. Am J Physiol Cell Physiol, 2021. **320**(1): p. C57-C65.

55. Wei, C., et al., *HDL-scavenger receptor B type 1 facilitates SARS-CoV-2 entry*. Nat Metab, 2020. **2**(12): p. 1391-1400.

56. Sanchez, E.L., et al., *Glycolysis, Glutaminolysis, and Fatty Acid Synthesis Are Required for Distinct Stages of Kaposi's Sarcoma-Associated Herpesvirus Lytic Replication*. J Virol, 2017. **91**(10).

57. Appelberg, S., et al., *Dysregulation in Akt/mTOR/HIF-1 signaling identified by proteo-transcriptomics of SARS-CoV-2 infected cells*. Emerg Microbes Infect, 2020. **9**(1): p. 1748-1760.

58. Palmer, C.S. and S.M. Crowe, *Immunometabolism may provide new insights into novel mechanisms of HIV reservoir persistence*. AIDS, 2016. **30**(18): p. 2895-2896.

59. Kang, S. and H. Tang, *HIV-1 Infection and Glucose Metabolism Reprogramming of T Cells: Another Approach Toward Functional Cure and Reservoir Eradication*. Front Immunol, 2020. **11**: p. 572677.

60. Sun, L., et al., *Metabolic Reprogramming in Immune Response and Tissue Inflammation*. Arterioscler Thromb Vasc Biol, 2020. **40**(9): p. 1990-2001.

61. Clerc, I., et al., *Entry of glucose- and glutamine-derived carbons into the citric acid cycle supports early steps of HIV-1 infection in CD4 T cells*. Nat Metab, 2019. **1**(7): p. 717-730.

62. Kitano, H., *Computational systems biology*. Nature, 2002. **420**(6912): p. 206-10.

63. Ideker, T., et al., *Discovering regulatory and signalling circuits in molecular interaction networks*. Bioinformatics, 2002. **18 Suppl 1**: p. S233-40.

64. Barabasi, A.L. and Z.N. Oltvai, *Network biology: understanding the cell's functional organization*. Nat Rev Genet, 2004. **5**(2): p. 101-13.

65. Joshi, A., et al., *Systems biology in cardiovascular disease: a multiomics approach*. Nat Rev Cardiol, 2021. **18**(5): p. 313-330.

66. Johnson, C.H., J. Ivanisevic, and G. Siuzdak, *Metabolomics: beyond biomarkers and towards mechanisms*. Nat Rev Mol Cell Biol, 2016. **17**(7): p. 451-9.

67. Sauer, U., *Metabolic networks in motion: 13C-based flux analysis*. Mol Syst Biol, 2006. **2**: p. 62.

68. Duarte, N.C., et al., *Global reconstruction of the human metabolic network based on genomic and bibliomic data*. Proc Natl Acad Sci U S A, 2007. **104**(6): p. 1777-82.

69. Ma, H., et al., *The Edinburgh human metabolic network reconstruction and its functional analysis*. Mol Syst Biol, 2007. **3**: p. 135.

70. Thiele, I., et al., *A community-driven global reconstruction of human metabolism*. Nat Biotechnol, 2013. **31**(5): p. 419-25.

71. Brunk, E., et al., *Recon3D enables a three-dimensional view of gene variation in human metabolism*. Nat Biotechnol, 2018. **36**(3): p. 272-281.

72. Mardinoglu, A., et al., *Integration of clinical data with a genome-scale metabolic model of the human adipocyte*. Mol Syst Biol, 2013. **9**: p. 649.

73. Mardinoglu, A., et al., *Genome-scale metabolic modelling of hepatocytes reveals serine deficiency in patients with non-alcoholic fatty liver disease*. Nat Commun, 2014. **5**: p. 3083.

74. Varemo, L., I. Nookaew, and J. Nielsen, *Novel insights into obesity and diabetes through genome-scale metabolic modeling*. Front Physiol, 2013. **4**: p. 92.

75. Lee, S., et al., *Network analyses identify liver-specific targets for treating liver diseases*. Mol Syst Biol, 2017. **13**(8): p. 938.

76. Robinson, J.L., et al., *An atlas of human metabolism*. Sci Signal, 2020. **13**(624).

77. Blais, E.M., et al., *Reconciled rat and human metabolic networks for comparative toxicogenomics and biomarker predictions*. Nat Commun, 2017. **8**: p. 14250.

78. Lieven, C., et al., *MEMOTE for standardized genome-scale metabolic model testing*. Nat Biotechnol, 2020. **38**(3): p. 272-276.

79. Bordbar, A., et al., *Constraint-based models predict metabolic and associated cellular functions*. Nat Rev Genet, 2014. **15**(2): p. 107-20.

80. Giurgiu, M., et al., *CORUM: the comprehensive resource of mammalian protein complexes-2019*. Nucleic Acids Res, 2019. **47**(D1): p. D559-D563.

81. Ryu, J.Y., H.U. Kim, and S.Y. Lee, *Framework and resource for more than 11,000 gene-transcript-protein-reaction associations in human metabolism*. Proc Natl Acad Sci U S A, 2017. **114**(45): p. E9740-E9749.

82. Romero, P., et al., *Computational prediction of human metabolic pathways from the complete human genome*. Genome Biol, 2005. **6**(1): p. R2.

83. Shlomi, T., et al., *Network-based prediction of human tissue-specific metabolism*. Nat Biotechnol, 2008. **26**(9): p. 1003-10.

84. Jerby, L., T. Shlomi, and E. Ruppin, *Computational reconstruction of tissue-specific metabolic models: application to human liver metabolism*. Mol Syst Biol, 2010. **6**: p. 401.

85. Agren, R., et al., *Reconstruction of genome-scale active metabolic networks for 69 human cell types and 16 cancer types using INIT*. PLoS Comput Biol, 2012. **8**(5): p. e1002518.

86. Olivares-Hernandez, R., S. Bordel, and J. Nielsen, *Codon usage variability determines the correlation between proteome and transcriptome fold changes*. BMC Syst Biol, 2011. **5**: p. 33.

87. Berglund, L., et al., *A genecentric Human Protein Atlas for expression profiles based on antibodies*. Mol Cell Proteomics, 2008. **7**(10): p. 2019-27.

88. Su, A.I., et al., *A gene atlas of the mouse and human protein-encoding transcriptomes*. Proc Natl Acad Sci U S A, 2004. **101**(16): p. 6062-7.

89. Wishart, D.S., et al., *HMDB: the Human Metabolome Database*. Nucleic Acids Res, 2007. **35**(Database issue): p. D521-6.

90. Agren, R., et al., *Identification of anticancer drugs for hepatocellular carcinoma through personalized genome-scale metabolic modeling*. Mol Syst Biol, 2014. **10**(3): p. 721.

91. Orth, J.D., I. Thiele, and B.O. Palsson, *What is flux balance analysis?* Nat Biotechnol, 2010. **28**(3): p. 245-8.

92. Edwards, J.S., R.U. Ibarra, and B.O. Palsson, *In silico predictions of Escherichia coli metabolic capabilities are consistent with experimental data*. Nat Biotechnol, 2001. **19**(2): p. 125-30.

93. Raman, K. and N. Chandra, *Flux balance analysis of biological systems: applications and challenges*. Brief Bioinform, 2009. **10**(4): p. 435-49.

94. Puchalka, J., et al., *Genome-scale reconstruction and analysis of the Pseudomonas putida KT2440 metabolic network facilitates applications in biotechnology*. PLoS Comput Biol, 2008. **4**(10): p. e1000210.

95. Mahadevan, R. and C.H. Schilling, *The effects of alternate optimal solutions in constraint-based genome-scale metabolic models*. Metab Eng, 2003. **5**(4): p. 264-76.

96. Reed, J.L. and B.O. Palsson, *Genome-scale in silico models of E. coli have multiple equivalent phenotypic states: assessment of correlated reaction subsets that comprise network states*. Genome Res, 2004. **14**(9): p. 1797-805.

97. Thiele, I., et al., *Functional characterization of alternate optimal solutions of Escherichia coli's transcriptional and translational machinery*. Biophysical journal, 2010. **98**(10): p. 2072-2081.

98. Bushell, M.E., et al., *The use of genome scale metabolic flux variability analysis for process feed formulation based on an investigation of the effects of the zwf mutation on antibiotic production in Streptomyces coelicolor*. Enzyme and Microbial Technology, 2006. **39**(6): p. 1347-1353.

99. Feist, A.M., et al., *Model-driven evaluation of the production potential for growth-coupled products of Escherichia coli*. Metab Eng, 2010. **12**(3): p. 173-86.

100. Pharkya, P. and C.D. Maranas, *An optimization framework for identifying reaction activation/inhibition or elimination candidates for overproduction in microbial systems*. Metab Eng, 2006. **8**(1): p. 1-13.

101. Lewis, N.E., et al., *Omic data from evolved E. coli are consistent with computed optimal growth from genome-scale models*. Mol Syst Biol, 2010. **6**: p. 390.

102. Chan, S.Y. and J. Loscalzo, *The emerging paradigm of network medicine in the study of human disease*. Circ Res, 2012. **111**(3): p. 359-74.

103. Freeman, L.C., *Centrality in social networks: Conceptual clarification*. Social network: critical concepts in sociology. Londres: Routledge, 2002. **1**: p. 238-263.

104. Borgatti, S.P., *Centrality and network flow*. Social networks, 2005. **27**(1): p. 55-71.

105. Opsahl, T., F. Agneessens, and J. Skvoretz, *Node centrality in weighted networks: Generalizing degree and shortest paths*. Social networks, 2010. **32**(3): p. 245-251.

106. Barrat, A., et al., *The architecture of complex weighted networks*. Proc Natl Acad Sci U S A, 2004. **101**(11): p. 3747-52.

107. Brandes, U., *A faster algorithm for betweenness centrality*. Journal of mathematical sociology, 2001. **25**(2): p. 163-177.

108. Barrat, A., et al., *The architecture of complex weighted networks*. Proceedings of the national academy of sciences, 2004. **101**(11): p. 3747-3752.

109. Newman, M.E., *Scientific collaboration networks. II. Shortest paths, weighted networks, and centrality*. Physical review E, 2001. **64**(1): p. 016132.

110. Dijkstra, E.W., *A note on two problems in connexion with graphs*, in *Edsger Wybe Dijkstra: His Life, Work, and Legacy*. 2022. p. 287-290.

111. Fortunato, S., *Community detection in graphs*. Physics reports, 2010. **486**(3-5): p. 75-174.

112. Newman, M.E. and M. Girvan, *Finding and evaluating community structure in networks*. Phys Rev E Stat Nonlin Soft Matter Phys, 2004. **69**(2 Pt 2): p. 026113.

113. Reichardt, J. and S. Bornholdt, *Statistical mechanics of community detection*. Phys Rev E Stat Nonlin Soft Matter Phys, 2006. **74**(1 Pt 2): p. 016110.

114. Blondel, V.D., et al., *Fast unfolding of communities in large networks*. Journal of statistical mechanics: theory and experiment, 2008. **2008**(10): p. P10008.

115. Traag, V.A., L. Waltman, and N.J. van Eck, *From Louvain to Leiden: guaranteeing well-connected communities*. Sci Rep, 2019. **9**(1): p. 5233.

116. Fortunato, S. and M. Barthelemy, *Resolution limit in community detection*. Proc Natl Acad Sci U S A, 2007. **104**(1): p. 36-41.

117. Waltman, L. and N.J. Van Eck, *A smart local moving algorithm for large-scale modularity-based community detection*. The European physical journal B, 2013. **86**: p. 1-14.

118. Dobin, A., et al., *STAR: ultrafast universal RNA-seq aligner*. Bioinformatics, 2013. **29**(1): p. 15-21.

119. Liao, Y., G.K. Smyth, and W. Shi, *featureCounts: an efficient general purpose program for assigning sequence reads to genomic features*. Bioinformatics, 2014. **30**(7): p. 923-30.

120. Ewels, P.A., et al., *The nf-core framework for community-curated bioinformatics pipelines*. Nat Biotechnol, 2020. **38**(3): p. 276-278.

121. Shen-Orr, S.S. and R. Gaujoux, *Computational deconvolution: extracting cell type-specific information from heterogeneous samples*. Curr Opin Immunol, 2013. **25**(5): p. 571-8.

122. Hagenauer, M.H., et al., *Inference of cell type content from human brain transcriptomic datasets illuminates the effects of age, manner of death, dissection, and psychiatric diagnosis*. PLoS One, 2018. **13**(7): p. e0200003.

123. Racle, J. and D. Gfeller, *EPIC: A Tool to Estimate the Proportions of Different Cell Types from Bulk Gene Expression Data*. Methods Mol Biol, 2020. **2120**: p. 233-248.

124. Uhlen, M., et al., *A pathology atlas of the human cancer transcriptome*. Science, 2017. **357**(6352).

125. Zhang, X., et al., *CellMarker: a manually curated resource of cell markers in human and mouse*. Nucleic Acids Res, 2019. **47**(D1): p. D721-D728.

126. Franzen, O., L.M. Gan, and J.L.M. Bjorkegren, *PanglaoDB: a web server for exploration of mouse and human single-cell RNA sequencing data*. Database (Oxford), 2019. **2019**.

127. Love, M.I., W. Huber, and S. Anders, *Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2*. Genome Biol, 2014. **15**(12): p. 550.

128. Risso, D., et al., *Normalization of RNA-seq data using factor analysis of control genes or samples*. Nat Biotechnol, 2014. **32**(9): p. 896-902.

129. Varemo, L., J. Nielsen, and I. Nookaew, *Enriching the gene set analysis of genome-wide data by incorporating directionality of gene expression and combining statistical hypotheses and methods*. Nucleic Acids Res, 2013. **41**(8): p. 4378-91.

130. Chen, E.Y., et al., *Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool*. BMC Bioinformatics, 2013. **14**: p. 128.

131. Subramanian, A., et al., *Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles*. Proc Natl Acad Sci U S A, 2005. **102**(43): p. 15545-50.

132. Ritchie, M.E., et al., *limma powers differential expression analyses for RNA-sequencing and microarray studies*. Nucleic Acids Res, 2015. **43**(7): p. e47.

133. Saccon, E., et al., *Distinct Metabolic Profile Associated with a Fatal Outcome in COVID-19 Patients during the Early Epidemic in Italy*. Microbiol Spectr, 2021. **9**(2): p. e0054921.

134. Wang, B., et al., *Similarity network fusion for aggregating data types on a genomic scale*. Nat Methods, 2014. **11**(3): p. 333-7.

135. Renz, A., L. Widerspick, and A. Drager, *FBA reveals guanylate kinase as a potential target for antiviral therapies against SARS-CoV-2*. Bioinformatics, 2020. **36**(Suppl 2): p. i813-i821.

136. Yaneske, E., et al., *Genome-scale metabolic modelling of SARS-CoV-2 in cancer cells reveals an increased shift to glycolytic energy production*. FEBS Lett, 2021. **595**(18): p. 2350-2365.

137. Wang, H., et al., *RAVEN 2.0: A versatile toolbox for metabolic network reconstruction and a case study on Streptomyces coelicolor*. PLoS Comput Biol, 2018. **14**(10): p. e1006541.

138. Reusch, N., et al., *Neutrophils in COVID-19*. Front Immunol, 2021. **12**: p. 652470.

139. Chua, R.L., et al., *COVID-19 severity correlates with airway epithelium-immune cell interactions identified by single-cell analysis*. Nat Biotechnol, 2020. **38**(8): p. 970-979.

140. Liao, M., et al., *Single-cell landscape of bronchoalveolar immune cells in patients with COVID-19*. Nat Med, 2020. **26**(6): p. 842-844.

141. Giamarellos-Bourboulis, E.J., et al., *Complex Immune Dysregulation in COVID-19 Patients with Severe Respiratory Failure*. Cell Host Microbe, 2020. **27**(6): p. 992-1000 e3.

142. Lourda, M., et al., *High-dimensional profiling reveals phenotypic heterogeneity and disease-specific alterations of granulocytes in COVID-19*. Proc Natl Acad Sci U S A, 2021. **118**(40).

143. Vanderbeke, L., et al., *Monocyte-driven atypical cytokine storm and aberrant neutrophil activation as key mediators of COVID-19 disease severity*. Nat Commun, 2021. **12**(1): p. 4117.

144. Hivroz, C., et al., *Crosstalk between T lymphocytes and dendritic cells*. Crit Rev Immunol, 2012. **32**(2): p. 139-55.

145. Saichi, M., et al., *Single-cell RNA sequencing of blood antigen-presenting cells in severe COVID-19 reveals multi-process defects in antiviral immunity*. Nat Cell Biol, 2021. **23**(5): p. 538-551.

146. Liu, X., et al., *Transcriptomic signatures differentiate survival from fatal outcomes in humans infected with Ebola virus*. Genome Biol, 2017. **18**(1): p. 4.

147. Codo, A.C., et al., *Elevated Glucose Levels Favor SARS-CoV-2 Infection and Monocyte Response through a HIF-1alpha/Glycolysis-Dependent Axis*. Cell Metab, 2020. **32**(3): p. 498-499.

148. Krishnan, S., et al., *Metabolic Perturbation Associated With COVID-19 Disease Severity and SARS-CoV-2 Replication*. Mol Cell Proteomics, 2021. **20**: p. 100159.

149. Guo, H., et al., *Multi-omics analyses reveal that HIV-1 alters CD4(+) T cell immunometabolism to fuel virus replication*. Nat Immunol, 2021. **22**(4): p. 423-433.

150. Gelpi, M., et al., *The central role of the glutamate metabolism in long-term antiretroviral treated HIV-infected individuals with metabolic syndrome*. Aging (Albany NY), 2021. **13**(19): p. 22732-22751.

151. Valle-Casuso, J.C., et al., *Cellular Metabolism Is a Major Determinant of HIV-1 Reservoir Seeding in CD4(+) T Cells and Offers an Opportunity to Tackle Infection*. Cell Metab, 2019. **29**(3): p. 611-626 e5.

152. Yin, M. and L.A.J. O'Neill, *The role of the electron transport chain in immunity*. FASEB J, 2021. **35**(12): p. e21974.

153. Mo, Q., et al., *Transcriptome profiling highlights regulated biological processes and type III interferon antiviral responses upon Crimean-Congo hemorrhagic fever virus infection*. Virol Sin, 2023. **38**(1): p. 34-46.

154. Arnold, C.E., et al., *Host response transcriptomic analysis of Crimean-Congo hemorrhagic fever pathogenesis in the cynomolgus macaque model*. Sci Rep, 2021. **11**(1): p. 19807.

155. Golden, J.W., et al., *The host inflammatory response contributes to disease severity in Crimean-Congo hemorrhagic fever virus infected mice*. PLoS Pathog, 2022. **18**(5): p. e1010485.

156. Ferrarese, C., et al., *Increased glutamate in CSF and plasma of patients with HIV dementia*. Neurology, 2001. **57**(4): p. 671-5.