Genome-scale modelling of human adipocyte metabolism

Impact paragraph

Humans can be considered as a network of organs (e.g., brain, heart, liver) where each organ contains dynamic networks of cellular and molecular components (e.g., genes, enzymes, reactions). Many diseases are multifactorial in nature and several physiological entities are affected within tissues and organs. Interaction of these entities can profoundly modify the progression and extent of the disease. Although these interactions are complex to understand, detailed knowledge of (dys)functions at the molecular level is necessary for identifying new targets, developing therapeutic strategies and determining the efficacy of a given therapeutic approach. The past two decades have seen tremendous developments in experimental techniques (-omics technologies) that measure several cellular molecules such as DNA (genomics), RNA (transcriptomics), proteins (proteomics) and metabolites (metabolomics). Such detailed molecular information is crucial to investigating cellular complexity in biological systems. These technological advancements have resulted in a huge amount of data becoming available which will facilitate identifying molecular processes that underlie disease development, progression and response to treatment. Each omics technology obtains data from a specific layer of cellular functioning and the integration of two or more omics can connect these different layers. Such a combination allows extracting information that would otherwise remain hidden if each dataset was considered alone. Current research is focussing on data integration and building mathematical models of biological systems for understanding complex disease mechanisms.

Towards this goal, the work presented here address analytical challenges for integrating multiple omics data with models of human metabolism. In this thesis, we developed two computational methods to analyse large-scale data and showed how these methods can be applied for studying specific aspects of adipocyte metabolism

in the context of obesity.

Visualisation approaches are an indispensable part of data analysis workflows. We developed a versatile and scalable workflow, EFMviz [99](Chapter 2), for visualising omics data with metabolic pathways. In addition, we presented several ways to combine network visualisation with existing methods for GEM-based analysis and study different biological scenarios (as shown in Chapters 2 [99], 3 [152] and 4 [206]). Through this demonstration, we showed that our versatile visualisation approach can be easily incorporated into multi-omics integration workflows. By enabling graphical visualisation and data mapping in (subsets of) metabolic models, EFMviz has made a significant contribution to interpreting molecular changes in metabolic pathways.

We also developed a model-based method, ComMet (Chapter 4, [206]), for comparing any metabolic phenotypes of interest. Moreover, we presented how ComMet can be used for comparing simulated (Chapter 4, [206]) and data-driven metabolic states of adipocytes (Chapter 5, [231]). These demonstrations highlight the potential of ComMet as a hypothesis generation tool. The predictions from ComMet provided insightful information and generated several hypotheses on metabolic adaptations in adipocytes. Future research can build upon these hypotheses to design experiments that test and validate ComMet predictions. The methodological advancement of ComMet enables investigating a wide range of metabolic conditions even in large human models which are highly relevant in a biomedical context. We comment in Chapter 4 that some potential applications include comparing metabolic shifts between healthy and diseased cells (ex. Cancer), (b) comparing metabolic capabilities of different members of gut microbiota and (c) identifying

metabolic differences between developing cell types. Therefore, our work will be interesting to a broad community of researchers, including but not limited to Systems Biologists and experimental biologists studying uni- or multi-cellular organisms.

We applied the methods developed during the course of this thesis to obtain insights into adipocyte metabolism. We identified sex-specific differences in adipocyte metabolism and in its role in regulating plasma BCAA levels (Chapter 3, [152]). We also characterised the metabolic differences in adipocytes that occur during weight loss (Chapter 5, [231]). The results of these biological applications are particularly interesting for researchers investigating metabolic dysfunctions in obesity. It is worth noting that our results focussed on adipocyte metabolism which is one part of the puzzle as many molecular aspects are dysregulated in different tissues in obesity. Thus, studies in future can follow our computational approaches and extend the study of obesity-specific metabolic dysfunctions into other organs (such as liver and muscle). This will help in building a more complete picture on a whole-body level. The molecular players in adipocytes that we identified to be affected can be starting point for future research to identify potential therapeutic targets. Therefore, these outcomes not only create avenues for further research but also can have societal implications in the long term.

The work done during this PhD been presented in several international conferences on systems biology and metabolic pathway analysis in the form of scientific talks, posters and workshops. Methods will be useful for systems biologists and have been published in academic peer-reviewed journals. Our work has been well received by the scientific community. Particularly, the utility of our method, ComMet, was

commended at International Conference on Intelligent Systems for Molecular Biology (ISMB 2020, a leading conference in bioinformatics) where our poster was awarded as the best poster. Both EFMviz and ComMet describe important methodological developments that are useful tools in constraint-based modelling. They are available under an open source licence to improve their usability. EFMviz has also been integrated into the COBRA toolbox platform [88, 99] which is widely used in the field of constraint-based modelling. Moreover, the scripts and workflows are available as educational scripts (e.g. MATLAB LiveScript, R Markdown). This was done with the intention of supporting open knowledge transfer which, in turn, will promote collaborative efforts within the scientific community. Unrestricted access to the methods will also facilitate easy adaptation by other researchers.

In summary, the work presented in this thesis can be seen as building blocks for applications in human health. Our strategies for multi-omics integration and visualisation can be applied to investigate human diseases and therefore help further biomedical research. For instance, they are also being applied in ageing and cardiovascular research at MaCSBio. Ultimately, the principles of data integration, such as the strategies presented here, have great potential in using omics data from individuals to generate personalised models of diseased functions and thus, contribute to developing personalised medical treatments.