Developing metabolomics for a systems biology approach to understand Parkinson's disease

Willacey, C.C.W.

Addendum

Summary

Neurodegenerative diseases, including Parkinson's disease (PD), are increasing in prevalence due to the aging population. Despite extensive study, these diseases are still not fully understood and the lack of personalised treatment options that can target the cause of the diseases, rather than the symptoms, has led to a greater demand for improved disease understanding, therapies and diagnostic procedures. In this thesis, we use systems biology approaches to construct disease-specific models intended for biomarker discovery, therapeutic treatment strategy identification and drug repurposing in PD. Systems biology is a mathematical field of research that analyses biological systems via construction of a computational model using experimental data. This is achieved by integration of omics data, including genomics, proteomics, transcriptomics and metabolomics. A specific approach used to identify the physico- and biochemical bounds within a biological system is constraint-based modelling, which requires the input of absolute quantitative metabolomics data. To improve our absolute quantitative coverage of the metabolome, we developed and improved new quantitative metabolomics methods using a targeted MS workflow to obtain data intended to be integrated into constraint-based metabolic models for the study of PD. A subset of PD is associated with mitochondrial dysfunction for a range of genetic mutations, consistent with the strong link between mitochondrial function and PD, with many of the associated metabolites involved in the TCA cycle and energy metabolism. Therefore, metabolomics methods were used to absolutely quantify metabolites of the central carbon and energy metabolism, and specific neurochemicals. Due to the vastness of the metabolome and the difficulty in achieving a high coverage of metabolites within a biological matrix, we employed chemical derivatization to enhance the detection and quantitation of a larger proportion of metabolites from human urine, SUIT-2 cells, HepG2 cells, rat brain and induced pluripotent stem cell (iPSC)-derived dopaminergic neurons. To evaluate the realistic biological environment, the use of 3D cell culture microfluidic devices are becoming more popular. However, this leads to a reduced sample size. Material-limited samples are providing challenges for the metabolomics community due to the required sensitivity. Chemical derivatization can aid detection by increasing sensitivity and offering other benefits, such as greater selectivity and separation. There is also a lack of existing absolute quantitative metabolite reference values (with specific focus on the mammalian brain) which is

essential for integration into metabolic models. In addition, metabolite profiles change in relation to the cell type, brain region and function. For these reasons, we also pursue the absolute quantitation of neurochemicals found in rat brain, i.e., the mammalian brain. Finally, we integrate multi-omics data from the iPSC-derived dopaminergic neurons into a genome scale constraint-based reconstruction and analysis model that can be used to understand the complexities of PD.

To begin the journey to produce a comprehensive PD model, **chapter 2** focused on the issues met with achieving broad coverage and absolute quantitation of the metabolome by creating a high-throughput, reliable, single, pre-column derivatization RPLC-MS/MS analysis with a 10-minute acquisition time using only positive ionization mode. Using the chemical derivatization reagent dimethylaminophenacyl bromide (DmPABr), we were able to simultaneously label carboxylic acids, thiols and amines of which only few published methods have successfully targeted together. These groups are abundant in the metabolites of the central carbon and energy metabolism, thus enabling us to quantify the concentrations of these metabolites. To further enhance quantitation, we also applied isotope-coded derivatization (ICD) by using internal standards with an isotopically labelled reagent (DmPABr-D_6). From human urine and SUIT-2 cells, we detected and quantified 64 central carbon and energy-related metabolites, including amino acids, N-acetylated amino acids, metabolites from the TCA cycle and pyruvate metabolism, acylcarnitines and medium-/long-chain fatty acids. We demonstrated the ability of the method to identify mitochondrial dysfunction by exposure of SUIT-2 cells to 100 nM rotenone. Following treatment, 50% of the metabolites detected showed significant alterations.

Another common problem faced by scientists in metabolite detection and quantitation is sensitivity, particularly with low volumes or concentrations of samples. **Chapter 3** applies the derivatization method and ICD described in **chapter 2** to material-limited cell samples, in turn improving chromatographic separation and enhancing MS ionization. After fast and accessible derivatization with DmPABr, we applied our novel RPLC-MS/MS method to HepG2 cells, ranging from 250 cells to 1×10^5 cells. In sub-10,000 cells, we were able to detect and quantify 37 metabolites, and a further 11 metabolites were detected below LLOQ.

In **chapter 4**, we address the current lack of available absolute quantitative mammalian brain metabolite reference values by quantifying neurochemicals across 25 regions isolated from rat brain to produce a comprehensive metabolic atlas of the mammalian brain. The brain regions associated with PD and other neurological disorders are focused on, including the orbitofrontal cortex, cerebral cortex, frontal lobe, ventromedial prefrontal cortex, subcortical structure and brain stem. We optimized LLE extraction before applying the derivatization LC-MS/MS technique. In this case, we used the benzoyl chloride derivatization reagent because, unlike DmPABr, it has the ability to stabilise vulnerable catecholamines and capture neuroactive metabolites. With our method, we were able to create a concentration profile of 43 metabolites including important neurotransmitters, such as dopamine, epinephrine, norepinephrine, GABA and serotonin, and key metabolites involved in specific pathways associated with PD, such as the urea cycle, and polyamine and tyrosine metabolism.

The integration of omics data into a model can vastly improve the understanding of the mechanism and function behind complex neurological conditions. **Chapter 5** introduces *iNESC2DN*, a validated constraint-based metabolic model in human dopaminergic neurons designed to aid the understanding of PD. We conducted multi-omics analysis on iPSC-derived, human neuroepithelial stem cells (hNESC) differentiated into midbrain-specific dopaminergic neurons and integrated the obtained data into a genome scale constraint-based reconstruction and analysis model that focused on the generic human metabolome, *Recon3D*, as a basis to generate stoichiometrically and flux consistent constraint-based model of dopaminergic neuron metabolism. AccQ-Tag derivatization (RPLC-MS) and GC-MS were applied to assess the central carbon and energy metabolism, and capture the neurochemical profile. In addition, GC-MS quantified sugars. Manual literature curation, transcriptomics and the metabolomics input were used to constrain the metabolism.

The final piece in the jigsaw of understanding disease may be in the metabolome. We developed and validated a new derivatization technique which has potential to absolutely quantify over 90% of the human metabolome. Within this thesis, we utilize targeted LC-MS/MS approaches and this limits the metabolic coverage required to truly piece together the human metabolism. From the data presented in

this thesis, with slight modifications, the methodology can potentially be expanded to untargeted derivatization and coupled to sensitive analytical equipment (i.e., nanospray ESI-MS) which would bring single-cell metabolomics within its reach.

With this thesis, we have demonstrated the applicability of systems biology to enhance the understanding of PD. We must advance from single compartment models onto multiple matrices/organs in order to form the whole picture of an organism. Successful execution of organ or multi-organ metabolic models will enhance the understanding of PD; this aligns with the disease's association with the metabolic changes that are exhibited not only in the substantia nigra, but also other brain regions. This is strongly dependent on access to accurate and reproducible omics data, and the ability to remove bias caused from data obtained from experiments that are do not realistically represent the human metabolome, which include issues such as inaccuracy in brain region cell line association and cell life cycle stage, variability in co-culture cell line expression and culturing cells in an artificial environment. In our study, we focused on dopaminergic neurons in the substantia nigra. Further future prospects of these computational models include the ability to explore the causal factors of PD and other complex neurological diseases, and pinpoint potential therapeutic targets. The combination of cell models, high quality measurements, a set of reference values in animal models and computational modelling could bridge the gap in reaching personalised medicines.