Universiteit Leiden
The Netherlands

Omics data integration with genome-scale modelling of dopaminergic neuronal metabolism

Preciat Gonzalez, G.A.

Citation

Preciat Gonzalez, G. A. (2022, December 21). *Omics data integration with genome-scale modelling of dopaminergic neuronal metabolism*. Retrieved from https://hdl.handle.net/1887/3503616

Version:	Publisher's Version
License:	Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden
Downloaded from:	https://hdl.handle.net/1887/3503616

Note: To cite this publication please use the final published version (if applicable).

Chapter 1

General introduction and scope

Living entities are part of the most complex systems we know. At the most basic level, a single cell comprises huge numbers of molecules and is structured in a very densely organised space. All of those molecules are involved in a variety of biochemical reactions that are driven by highly regulated enzymes and interfere with the cell in the form of hormones, drugs, or variations in the amount of nutrition available.

With the evolution of modern biosciences, multiple techniques have been developed to collect information about the elementary building blocks of an organism, also known as *omics* data[1]. Before the first genome was sequenced in 1995, the proteobacteria *Haemophilus influenza* (Fleischmann et al.), we believed we could comprehend the metabolic behaviour of a biological system and all its genetic elements by knowing its complete genomic sequence, starting with what it is known as genomics. This genotype-phenotype assumption might be correct for a simple biological system, such as *Mycoplasma genitalium*, consisting of 521 genes (Fraser et al.). Nevertheless, the metabolic behaviour of complex biological systems is difficult to fully understand due to the complexity of the interactions of their gene products, increasing the complexity of the genotype-phenotype relationship in a biological system. Velculescu et al. pioneered transcriptomics by characterising the entire set of RNA transcripts produced by the yeast *Saccharomyces cerevisiae*, where genes are expressed differently under different conditions. Additionally, the longer the genome is, the more likely it is to have non-coding DNA. For that reason, scientists aimed to find gene products in cells, isolate them and characterise them; the data acquired was called the bibliome (Grivell et al.). Later, it was discovered that a gene product does not do much by itself. The interaction with other gene products is necessary to manifest a biological function, e.g., in glycolysis, where all 12 chemical transformations are catalysed by their respective gene products need to be active (Berg et al.). Moreover, glycolysis is just a pathway within a larger biological system and will never function independently. Single-gene defects cause over 4,000 human diseases (Ropers). A single gene mutation can disrupt an entire biological system by affecting all low-molecular-weight molecules in the cellular environment, which later was described as the metabolome.

[1] In this context, the suffixes -ome and -omics are used to describe entire biological datasets, such as sets of biomolecules deriving from a single organism (e.g. proteome and lipidome), related to each other (e.g. genome, transcriptome and metabolome) or refer to the structure or function (e.g. chemome and glycome). *Omics* data can be obtained from big databases or technologies like next-generation sequencing, mass spectrometry or NMR spectroscopy.

From prokaryotes to eukaryotes, biological systems are highly organised in their structure and function. Additionally, they evolve by adapting to their environment. These modifications make their behaviour difficult to predict from the properties of individual parts, as it was believed before sequencing the genome of *Haemophilus influenza* (Fleischmann et al.).

Systems biology, an approach used in biomedical research, integrates the complexity of living entities, the evolution of modern biosciences, and the rise of *omics* to study biological systems. This approach provides new insights into the biochemical origins of diseases or allows us to predict novel targets for biomarkers or design new therapies (Heinken et al., Preciat, L. Moreno, Wegrzyn and others, Norsigian et al., Aurich et al., Brunk et al.). Systems biology combines *omics* data to obtain a mathematical description of a biological system. In principle, if we understand the set of instructions a biological system follows, we might be able to adjust it to comprehend the pathogenesis of complex diseases or develop novel therapies. This can be done through highly detailed computer models to precisely calculate the interactions of components to predict system behaviour. These models incorporate experimental conditions by incorporating physicochemical or experimental constraints. (Zierer et al., Das et al.)

Modelling biological systems

Mathematical models can be used to represent a biological system and recreate aspects of the organism's function. The models are useful for a variety of purposes, including developing intuition, filling in gaps in processes where we have made observations or testing all parameters at once (Figure 1).

However, a model doesn't behave exactly as an organism would do in the real world, but it can be used to approximate key aspects of it, allowing us to simulate experiments that would otherwise be impossible to run and allow us to test all of the parameters in the model at once. They can also be used to conduct experiments in order to test hypotheses and develop a theory.

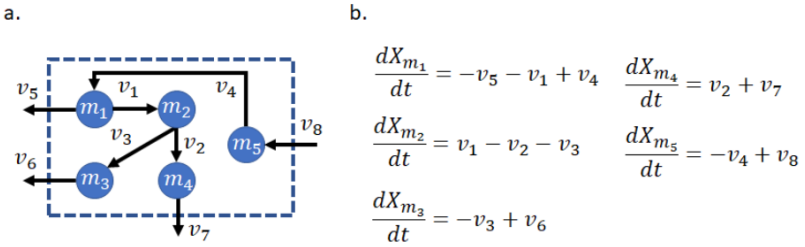

Figure 1: **An example of a model.**
A metabolic network with five metabolites (m_1, m_2, m_3, m_4 and m_5), four internal reactions (v_1, v_2, v_3 and v_4), and four exchange reactions (v_5, v_6, v_7 and v_8) within the boundary (a). Set of equations for estimating the concentration of each metabolite over time (b).

Mathematical models have different properties. Some of the properties can be summarised in Table 1.

With the evolution of modern biosciences, the information obtained from biological systems has grown and modelling approaches have become more complex, necessitating the implementation of computationally efficient algorithms. Various modelling approaches have been developed to study biological systems, where they can be used for basic exploration, disease biomarker discovery, and physiological phenotyping.

Machine learning. It is an explicit, statistical, and indicative model used to predict trends and patterns. These models learn from the experience provided in the form of data and eliminate irrelevant or redundant information to discover relevant patterns.

Kinetic models. It describes network dynamics using ordinary differential equations by modelling changes over time. This is accomplished by relating the rates of change for the elements in the model. They are typically dynamic, nonlinear, explicit, small, and deductive models.

Mechanistic models. They are based on fundamental natural science laws, such as physical and biochemical principles. They analyse the mechanism observed in an experimental scenario recreated through simulations based on a series of constructive principles. This approach includes different properties to represent the biological system, such as

steady-state, linearity, implicit, mechanistically, discrete, and deductive modelling.

Table 1: **Properties of mathematical models.**
The correct modelling approach is achieved by identifying the types of data available as well as the purpose of the study.

Property	Options	Description	Representation
Time	Steady state	It is time-invariant; it is used to calculate the elements of system in equilibrium.	$Sv = 0$
	Unsteady state	Time dependent.	$y_1 = x + y_0$
	Dynamic	Represent time-dependent changes in a system; involves the combination of algebraic and differential equations.	$y_k(t) = g_1(u_m(t), \ldots, x_n(t))$
Linearity	Linear	Represent systems that can be graphically represented as a straight line.	$y = x + 1$
	Nonlinear	The output difference is not proportional to the input difference.	$y = x^3$
Method	Explicit	Inputs are known.	$y = f(x)$
	Implicit	Outputs are known.	$R(x, y) = 0$
Approach	Mechanistic	A unique output is produced to a unique output for linear models and multiple outputs are possible for non-linear models.	$Sv = 0$
	Statistical	Sums up a collection of statistical assumptions from sample data.	$Y_i = \alpha + \beta X_i + \varepsilon_i$
	Probabilistic	Accounts uncertainties caused due to the varying behavioural characteristics; same inputs produce different outputs.	$x(t) = f(t) + h(t)\varepsilon(t)$
Frequency	Discrete	Represent states in a statistical model.	$y = \sum_{x=0}^{t}(x)$
	Continuous	The state change continuously over time	$y = \int_0^t (x)\,dx$
Scale	Small	Describe detailed processes.	$S \in \mathbb{R}^{3\times4}$
	Genome-scale	Describe complex/multidimensional processes.	$S \in \mathbb{R}^{1230\times3402}$
Nature	Deductive	Logical structures.	-
	Indicative	Based deductive models.	-
	Floating	Based on the invocation of the expected structure.	-

These approaches have been used to study complex diseases (Preciat, L. Moreno, Wegrzyn and others, Dinov et al.), which are the result of the convergence of several genomic variations, including diseases such as Parkinson's disease, diabetes, and cancer.

Parkinson's disease

After Alzheimer's, Parkinson's disease is the most common neurodegenerative disease (Balestrino and Schapira). The most noticeable

symptoms of this disease are slowness of movement, muscle rigidity, and resting tremors. Parkinson's disease affects many areas of the nervous system and different types of neurons and other cells, such as astrocytes and microglia. Still, much emphasis has been placed on dopaminergic neurons in brain regions associated with motor symptoms, particularly the substantia nigra pars compacta, an area of the midbrain (Oliveira et al.). This region is part of a major pathway in the brain that is important for movement facilitation. In Parkinson's disease, dopaminergic neurons in the substantia nigra die gradually, resulting in a malfunction of this pathway and the characteristic motor symptoms. Drugs that replace or mimic dopamine are frequently used to treat these types of deficits, but their efficacy diminishes over time. In addition, although deep brain stimulation may be used to treat symptoms, no current treatment slows neurodegeneration.

In most cases of Parkinson's disease, clumps of misfolded proteins within neurons are a distinct pathology. Lewy bodies are the most common and are formed with the aggregation of alpha-synuclein (Munoz et al., Lehtonen et al., Oliveira et al.). Parkinson's has also been linked to problems with mitochondria (Yao et al., Requejo-Aguilar et al., Hedrich et al., Ge et al.), which provides the neuron with the energy to perform vital functions and with neuroinflammation (Borsche et al., Jiang et al.). Another idea is that glial cells surrounding neurons may play a role in Parkinson's disease, causing neuroinflammation and therefore, damaging dopaminergic neurons (Belanger et al.).

Novel approaches to studying Parkinson's disease are being proposed. Lucumi Moreno et al. developed a three-dimensional cell culture bioreactor that employs microfluidics to induce neuroepithelial stem cell differentiation in dopaminergic neurons. Aggregation models, based on neuronal connectivity, predict the spread of misfolded proteins such as alpha-synuclein, which is likened to Parkinson's disease (Oliveira et al.). Dinov et al. used machine learning and big data on Parkinson's disease to analyse clinical, demographic, and genetic data to predict Parkinson's disease with an accuracy of 0.96. However, due to the multifactorial nature of the disease progression, understanding Parkinson's disease necessitates an interdisciplinary approach involving experimental and modelling studies in addition to clinical studies. In Borsche et al. is identified an association between Parkinson's disease genotypes, increased mtDNA release, and neuromodulation in Parkinson's disease patients after analysing different proteins and circulating cell-free mtDNA in serum of 245 participants in two cohorts from tertiary

movement disorder centres. Jiang et al. showed that KH176m inhibited the enzymatic activity of a metabolite up-regulated by inflammatory stimuli in both mouse macrophage-like cells and human fibroblasts, which may aid in treating patients with mitochondrial diseases and other diseases associated with inflammation.

Constraint-based modelling, the approach used in this thesis, is a mechanistic approach, which aims to represent the genotype-phenotype relationship of a biological system according to a multidimensional space representing the metabolism of a dopaminergic neuron, a neuron associated with the motor symptoms of Parkinson's disease.

Constraint-based modelling

Biological systems are subject to environmental constraints, such as temperature, osmolarity, or the availability of nutrients. Regulatory constraints enable cells to eliminate suboptimal conditions in order to confine themselves to fit behaviours; for example, a neuron will not use lactose for energy production if glucose is available.

Constraint-based modelling is a computational and mathematical approach that uses genome-scale metabolic models to simulate real-life biological activities *in silico*, providing an integrated view of a biological system. The goal is to predict the net flux of the metabolic reactions in the genome-scale metabolic model, given the constraints of the biological system.

The flux distribution is calculated mathematically as $Sv = dx/dt$, which represents a system with internal reactions occurring within the biological system and external reactions representing metabolites exchanged with the environment; $S \in R^{n \times m}$ is the stoichiometric matrix representing the biochemical network, m is the number of metabolites, and n is the number of reactions; v represents the flux vector, and dx/dt represents the metabolite concentrations over time.

A genome-scale metabolic model includes several physicochemical and context-specific constraints to predict the net flux, such as (1) the connectivity between metabolites and reactions, represented by the stoichiometric matrix S; (2) thermodynamic constraints indicate the reaction directionality since some reactions cannot be reversed, such as hexokinase, a glycolysis-related enzyme; (3) maximum flux rates, represent enzyme turnover rates, nutrient uptake or metabolite

secretions; and (4) the steady-state constraints, or null space, predicts the net flux with no metabolite accumulation, represented by $Sv = 0$.

Considering all the constraints, a multidimensional space is generated with all possible fluxes, allowing for the optimisation of specific features of a biological system e.g., Flux Balance Analysis (FBA). In FBA the feature of a genome-scale model to be optimised is mathematically represented by the objective function and may represent nutrient uptake minimisation, metabolite production maximisation, or Euclidean norm minimisation (Figure 2).

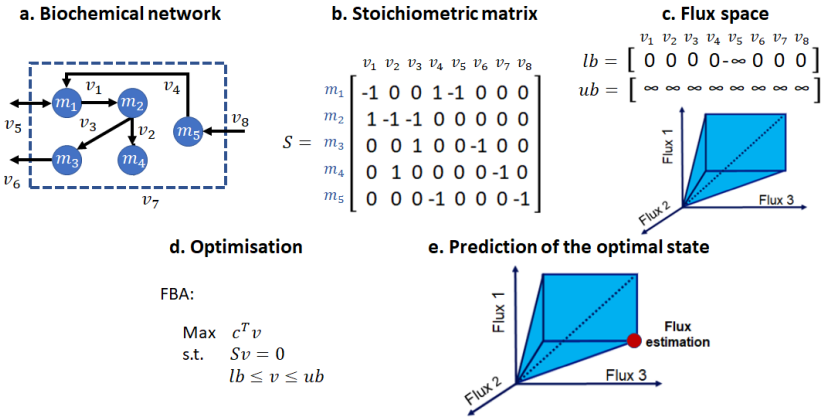

Figure 2: **An example of a constraint-based modelling workflow**
A biochemical network represents a constraint on the connectivity between metabolites and reactions (a) and its mathematical representation is the stoichiometric matrix S (b). Thermodynamic and flux rate constraints are indicated by the lower and upper bounds (lb and ub). For this example, the reactions v_5 and v_8 are the inputs. S, lb and ub form a unique convex flux space, representing all the possible fluxes (c). Optimisation methods such as FBA are used to identify an optimal solution for the given constraints in the flux space. c is the objective function, $Sv = 0$ is the steady-state constraint and the maximum and minimum reaction bounds were set between $\pm\infty$, therefore $-\infty \leq v \leq \infty$ (d). The optimal state is represented by a point in the flux space, where each coordinate represents the flux of a metabolic reaction (e).

Several protocols for generating genome-scale metabolic models have been developed (Thiele and Palsson, Norsigian et al., Preciat, Wegrzyn, Thiele and others), as well as algorithms for extracting them using context-specific information such as gene expression (Becker and Palsson, Zur et al., Agren et al.), a set of active reactions (Jerby et al.,

Preciat, Wegrzyn, Thiele and others, Vlassis et al.) or metabolite's concentrations (Aurich et al., Preciat, Wegrzyn, Thiele and others, Capela et al.). Different modelling approaches can also be used to integrate transcriptomics (Zur et al.) or metabolomics (Aurich et al., Preciat, Wegrzyn, Thiele and others) data.

The formulation of the objective function is critical in constraint-based modelling for accurately representing the phenotypic behaviour of a genome-scale model given by the flux vector v. An objective function represents a context-specific or a mathematical feature of a biological system. Context-specific objective functions may include reactions to ensure the system's growth or maintenance, such as biomass growth, cellular synthesis, and turnover requirements (c in Figure 1d). Among the mathematical objective functions is found the zero norm ($\|v\|_0$), which minimises the total number of non-zero elements in the flux vector; the one norm ($\|v\|_1$), which minimises the sum of the magnitudes in the flux vector; or the two norm ($\|v\|_2$), which minimises the distance between the elements in the flux vector. Furthermore, mathematical and context-specific objective functions could be combined by weighting the norms with biological or chemical data such as reaction expression based on transcriptomic data or the number of bonds broken and formed in a biochemical reaction ($q^T \|v\|_p$; where p is the norm and q a vector with context-specific data for each reaction). Alternatively, an unbiased description of a flux distribution can be provided by sampling the flux space uniformly to have the flux distribution of each metabolic reaction (Haraldsdottir et al.).

Finally, constraint-based modelling can be used in combination with experimental methods. Tracer-based metabolomics is an experimental approach for determining the distribution of stable isotope tracers across multiple metabolic stages (Wiechert). The distribution of the tracers enables a quantitative link between precursor and product, which aids in the identification of metabolic intermediates in biological systems. Data from tracer-based metabolomics can be analysed and interpreted using a constraint-based modelling approach. Modelling biological systems with efficient objective functions could be used to develop novel therapies or experiments in an attempt to better understand the neurodegeneration caused by Parkinson's disease (Figure 3).

Scope and outline

This thesis aims to generate a control model of a dopaminergic neuron derived from an induced pluripotent stem cell-derived dopaminergic neuronal culture, integrating *omics* data, and use it to design novel *in vitro*, tracer-based experiments.

It was assumed that the integration of *omics* information into a genome-scale model would increase the predictive capacity of the model. If the model holds up to reality, it could be used to interpret patient-specific metabolic data, combine these models with clinical data to develop diagnostic methods, identify biomarkers that reduce the selective death of dopaminergic neurons in the sustantia nigra, aid in the design of tracer-based metabolomics experiments or compare the flux predicted in the models to biomarkers found in clinical samples.

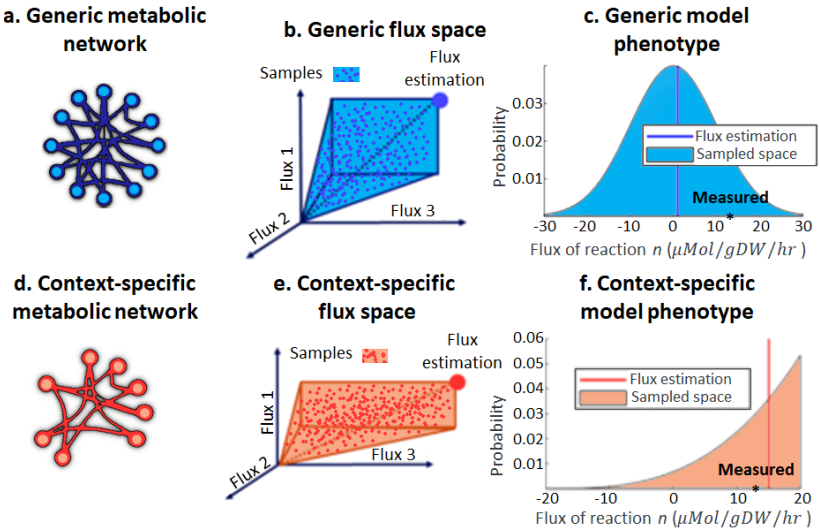

Figure 3: **Generic vs context-specific models**
A generic metabolic network with structural and thermodynamic constraints (a) has its own flux space (b) that can be sampled or used to predict the reaction's fluxes to analyse the phenotype of a biological system (c). Reactions, genes, and metabolites can be added or removed by incorporating context-specific data into a metabolic network, such as bibliomic, metabolomic, or transcriptomic data. Context-specific data could also be used to constrain reaction rates. (d). By adjusting the constraints of the generic model, the flux space can be reshaped (e), resulting in new predictions (f).

The specific aims were: (1) to integrate qualitative *omics* data into genome-scale models for the generation of context-specific models; (2) to generate a context-specific model using *omics* data from induced pluripotent stem cell-derived dopaminergic neuronal culture; (3) to compare existing atom mapping tools, a one-to-one correspondence between an atom in a substrate and an atom in a product using directed graphs, for the integration of chemomics data in genomic-scale models; and (4) to generate an atomically resolved, context-specific genome-scale model for neuronal dopaminergic metabolism (Figure 4).

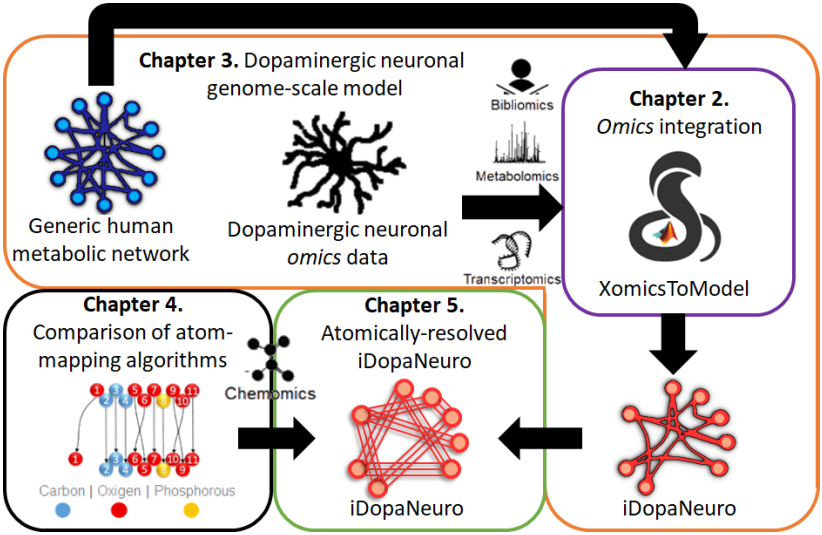

Figure 4: **Scope of the thesis**
Chapter 2, describes the development of the `XomicsToModel` pipeline, which allows integrating *omics* data into a generic genome-scale model for the generation of context-specific models. In **Chapter 3**, the `XomicsToModel` pipeline was implemented to generate a context-specific model, denoted as `iDopaNeuro`, using the generic human reconstruction Recon 3D, bibliomics, metabolomics, and transcriptomics data from induced pluripotent stem cell-derived dopaminergic neuronal culture. In **Chapter 4**, existing atom mapping algorithms were compared to integrate chemomics data into genomic-scale models. The most accurate atom mapping algorithm was used in **Chapter 5** to generate a pipeline that atomically resolves genome-scale models, including the `iDopaNeuro` model, generating a chemoinformatic database of metabolite structures and atom-mapped reactions.

In **Chapter 2**, is presented the `XomicsToModel` pipeline. It was integrated context-specific information from sources, such as proteo-

mics, metabolomics, bibliomics, and transcriptomics with a thermodynamic and flux-consistent, context-specific genome-scale metabolic model. The pipeline allowed flexible, modular integration of context-specific constraints into a genome-scale metabolic model. The pipeline also allowed several thermodynamic and flux-consistency checks, as well as quality checks on flux feasibility.

In **Chapter 3**, the `XomicsToModel` pipeline was used to generate a thermodynamically consistent and context-specific genome-scale model from a dopaminergic neuronal culture derived from induced pluripotent stem cells (`iDopaNeuro`) that allowed to analyse the phenotypes and genotypes seen in Parkinson's disease.

The `iDopaNeuro` model was generated using different context-specific information, such as (1) bibliomics, information about reactions, genes, and metabolites known to be active or inactive, known reaction rates and coupled reactions for neuronal maintenance, all based on a review of the literature; (2) transcriptomic, RNA-seq data from a dopaminergic neuronal culture; and (3) exometabolomics, maximum and minimum reaction rates for the uptake and secretion of metabolites based on their change in concentration in the spent media. The `iDopaNeuro` model was validated by performing *in silico* perturbations, inhibiting the mitochondrial complex I and mitochondrial complex V and comparing it to the exometabolomic phenotype of a dopaminergic neuronal culture treated with rotenone and oligomycin, which also inhibits the mitochondrial complex I and mitochondrial complex V respectively. Furthermore, a novel experimental design approach was developed to maximally shrink the set of external reaction fluxes by consistently identifying the top-ranked exometabolites whose corresponding external reactions would be the most important to constrain reducing model uncertainty. Finally, to go beyond stoichiometry, the `iDopaNeuro` model can be described at the atomic level. This description can be accomplished by obtaining the reaction mechanism described by the atom mappings in all of the reactions in the `iDopaNeuro` model.

In **Chapter 4,** the predictive capacity, features, and availability of various atom mapping algorithms were compared for the further integration of chemomics data in genome-scale metabolic models. The selected algorithm was chosen based on the accuracy of the tested algorithms by comparing them to >500 manually curated atom-mapped metabolic reactions. The decision was also influenced by technical and special features shared by the three tested algorithms, such as hydrogen atom mapping, reaction centre identification, and chemically equivalent

atom identification or availability. The algorithm chosen allows for the development of a cheminformatic database specific to a genome-scale metabolic model.

In **Chapter 5**, a pipeline was developed to generate a database of standardised metabolite structures and atom-mapped reactions (`generateChemicalDatabase`). The metabolite information in a genome-scale metabolic model was used to determine the most consistent metabolite structure for a genome-scale model based on an InChI-based comparison where features such as stereochemistry, charge, metabolite formula, and similarity with different sources were considered. The metabolite structures were represented in a variety of formats, including database identifiers or chemoinformatic formats, such as chemical tables, SMILES, and InChI. The standardised metabolite structures were used to generate a database of metabolic reactions with atoms mappings using the algorithm with the greatest predictive capacity tested in **Chapter 4**. The number of broken and formed bonds and the enthalpy change in a metabolic reaction were identified based on the reaction mechanism depicted with the predicted atom mappings. The pipeline was used to generate a chemoinformatic database of different genome-scale models, including the `iDopaNeuro` model and Recon3D, a generic human reconstruction. Additionally, the `iDopaNeuro` model was used to design tracer-based experiments.

Finally, **Chapter 6** provides a general conclusion to the thesis's genome-scale metabolic model generation approach and `iDopaNeuro` models. Perspectives and recommendations on how to improve and apply the proposed approach are also provided.

References

Agren, R., Bordel, S., Mardinoglu, A. and others: 2012, Reconstruction of Genome-Scale Active Metabolic Networks for 69 Human Cell Types and 16 Cancer Types Using INIT, *PLOS Computational Biology* **8**(5).

Aurich, M. K., Fleming, R. M. T. and Thiele, I.: 2016, MetaboTools: A Comprehensive Toolbox for Analysis of Genome-Scale Metabolic Models, *Frontiers in Physiology* **0**.

Balestrino, R. and Schapira, A. H. V.: 2020, Parkinson disease, *European Journal of Neurology* **27**(1), 27–42.

Becker, S. A. and Palsson, B. O.: 2008, Context-Specific Metabolic Networks Are Consistent with Experiments, *PLOS Computational Biology* **4**(5).

Belanger, M., Allaman, I. and Magistretti, P. J.: 2011, Brain Energy Metabolism: Focus on Astrocyte-Neuron Metabolic Cooperation, *Cell Metabolism* **14**(6), 724–738.

Berg, J. M., Tymoczko, J. L. and Stryer, L.: 2002, *Biochemistry*, W.H. Freeman.

Borsche, M., König, I. R., Delcambre, S. and others: 2020, Mitochondrial damage-associated inflammation highlights biomarkers in PRKN/PINK1 parkinsonism, *Brain* **143**(10), 3041–3051.

Brunk, E., Sahoo, S., Zielinski, D. C. et al.: 2018, Recon3d enables a three-dimensional view of gene variation in human metabolism, *Nature Biotechnology* **36**, 272.

Capela, J., Lagoa, D., Rodrigues, R. and others: 2021, merlin v4.0: an updated platform for the reconstruction of high-quality genome-scale metabolic models, *The FEBS Journal* **287**.

Das, T., Andrieux, G., Ahmed, M. and others: 2020, Integration of Online Omics-Data Resources for Cancer Research., *Frontiers in genetics* **11**.

Dinov, I. D., Heavner, B., Tang, M. and others: 2016, Predictive big data analytics: A study of parkinson's disease using large, complex, heterogeneous, incongruent, multi-source and incomplete observations, *PLoS One* **16**.

Fleischmann, R. D., Adams, M. D., White, O. et al.: 1995, Whole-genome random sequencing and assembly of haemophilus influenzae rd., *Science (New York, N.Y.)* **269**(5223), 496–512.

Fraser, C. M., Gocayne, J. D., White, O. et al.: 1995, The minimal gene complement of mycoplasma genitalium., *Science (New York, N.Y.)* **270**(5235), 397–403.

Ge, P., Dawson, V. L. and Dawson, T. M.: 2020, PINK1 and Parkin mitochondrial quality control: a source of regional vulnerability in Parkinson's disease, *Molecular Neurodegeneration* **15**.

Grivell, L., Allen, J., Ashburner, M. et al.: 2002, Mining the bibliome: searching for a needle in a haystack?: New computing tools are needed to effectively scan the growing amount of scientific literature for useful information, *EMBO Reports* **3**(3), 200–203.

Haraldsdottir, H. S., Cousins, B., Thiele, I. and others: 2017, CHRR: coordinate hit-and-run with rounding for uniform sampling of constraint-based models, *Bioinformatics* **33**(11), 1741–1743.

Hedrich, K., Hagenah, J., Djarmati, A. and others: 2006, Clinical Spectrum of Homozygous and Heterozygous PINK1 Mutations in a Large German Family With Parkinson Disease: Role of a Single Hit?, *Archives of Neurology* **63**(6), 833.

Heinken, A., Acharya, G., Ravcheev, D. and others: 2021, AGORA2: Large scale reconstruction of the microbiome highlights wide-spread drug-metabolising capacities | bioRxiv, *bioRxiv* .

Jerby, L., Shlomi, T. and Ruppin, E.: 2010, Computational reconstruction of tissue-specific metabolic models: application to human liver metabolism, *Molecular Systems Biology* **6**, 401.

Jiang, X., Renkema, H., Pennings, B. and others: 2021, Mechanism of action and potential applications of selective inhibition of microsomal prostaglandin E synthase-1-mediated PGE2 biosynthesis by sonlicromanol's metabolite KH176m, *Scientific Reports* **11**, 880.

Lehtonen, S., Sonninen, T.-M., Wojciechowski, S. and others: 2019, Dysfunction of Cellular Proteostasis in Parkinson's Disease, *Frontiers in Neuroscience* **13**.

Lucumi Moreno, E., Hachi, S., Hemmer, K. and others: 2013, Differentiation of neuroepithelial stem cells into functional dopaminergic neurons in 3D microfluidic cell culture, *Lab on a Chip* **5**.

Munoz, P., Huenchuguala, S., Paris, I. and others: 2012, Dopamine Oxidation and Autophagy, *Parkinson's Disease* **2012**.

Norsigian, C. J., Fang, X., Seif, Y. and others: 2021, A workflow for generating multi-strain genome-scale metabolic models of prokaryotes, *Nature Protocols* **15**(1), 1–14.

Oliveira, M. A. P., Arreckx, S., Monte, D. D., and others: 2019, The connectome is necessary but not sufficient for the spread of alpha-synuclein pathology in rats, *bioRxiv* .

Oliveira, M. A. P., Balling, R., Smidt, M. P. and others: 2017, Embryonic development of selectively vulnerable neurons in Parkinson's disease, *npj Parkinson's Disease* **3**(1).

Preciat, G., L. Moreno, E., Wegrzyn, A. and others: 2021, Mechanistic model-driven exometabolomic characterisation of human dopaminergic neuronal metabolism, *In preparation* .

Preciat, G., Wegrzyn, A., Thiele, I. and others: 2021, XomicsToModel: a COBRA Toolbox extension for generation of metabolic models from multi-omic data, *In preparation* .

Requejo-Aguilar, R., Lopez-Fabuel, I., Fernandez, E. and others: 2014, PINK1 deficiency sustains cell proliferation by reprogramming glucose metabolism through HIF1, *Nature Communications* **5**.

Ropers, H.-H.: 2010, Single gene disorders come into focus - again., *Dialogues in clinical neuroscience* **12**(1), 95–102.

Thiele, I. and Palsson, B. O.: 2010, A protocol for generating a high-quality genome-scale metabolic reconstruction, *Nature Protocols* **5**(1), 93–121.

Velculescu, V., Zhang, L., Zhou, W. and others: 1997, Characterization of the yeast transcriptome, *Cell* **2**(88).

Vlassis, N., Pacheco, M. P. and Sauter, T.: 2014, Fast Reconstruction of Compact Context-Specific Metabolic Network Models, *PLOS Computational Biology* **10**(1).

Wiechert, W.: 2001, 13c metabolic flux analysis, *Metabolic Engineering* **3**(3), 195–206.

Yao, Z., Gandhi, S., Burchell, V. S. and others: 2011, Cell metabolism affects selective vulnerability in PINK1-associated Parkinson's disease., *Journal of cell science* **124**(Pt 24), 4194–202.

Zierer, J., Menni, C., Kastenmüller, G. and others: 2015, Integration of 'omics' data in aging research: from biomarkers to systems biology., *Aging cell* **14**(6), 933–944.

Zur, H., Ruppin, E. and Shlomi, T.: 2010, iMAT: an integrative metabolic analysis tool, *Bioinformatics (Oxford, England)* **26**(24), 318–329.