THE CONSTRUCTION AND ANALYSIS OF A GENOME-SCALE METABOLIC MODEL OF MESENCHYMAL STEM CELLS CULTURED ON SURFACE TOPOGRAPHIES

Master thesis of	:	Manouk Groels, 0904057
Department	:	Biomedical Engineering – Computational Biology
Course code	:	8ZM104
Organization	:	Eindhoven University of Technology
Graduation committee	:	Prof. dr. Natal van Riel (chairman) Prof. dr. Peter Hilbers Prof. dr. Jan de Boer MSc. Andrea Cabbia (advisor)
Period	:	11/2019 – 10/2020

TU/e EINDHOVEN UNIVERSITY OF TECHNOLOGY

Abstract

An increasing interest in human mesenchymal stem cells (hMSCs) in regenerative medicine is growing forth from their amazing capabilities, such as their multipotency, immunomodulatory capabilities, and their ability to maintain proliferating. To use these cells in stem cell therapies, it is important to multiply them in vitro. However, during this in vitro cultivation, a loss of stem cell-like characteristics occurs, while it is precisely these features that make hMSCs effective in stem cell therapy. In order to maintain stem cell-like characteristics, topographically enhanced surfaces were used to mimic the mechanical microenvironment of the hMSCs in the bone marrow. This study focuses on the metabolic adjustment of hMSCs to culturing on topographies in comparison with culturing on flat surfaces. A genome-scale metabolic reconstruction (GSMM) is used as a computational model, to predict main metabolic ATP production and hMSC growth rates. We hypothesize that cells growing on topographies are better able to maintain their stem cell-like characteristics and so, reduce their metabolism and proliferation.

A tissue-specific base model for hMSCs was constructed with the use of the CORDA algorithm, a switch-based method. The culture medium composition functions as input for the hMSC-base model to which a biomass objective function was added to be able to simulate cell growth. The hMSC-base model was further constrained via a valve-based method. This method used gene expression data of hMSCs cultured in vitro on a flat surface and seven different topographically enhanced surfaces and resulted in 24 different topography-specific GSMMs.

The topography-specific models predicted that the hMSCs are mainly dependent on glycolysis (95.5% - 99.3%) for their ATP production, indicating undifferentiated cells. The in silico growth rates were predicted to be between 0.0082 h^{-1} and 0.0107 h^{-1}.

The combination of a switch- and a valve-based method to construct condition-specific GSMMs resulted in accurate predictions for the energy production and growth rates of cells cultured in vitro and so, this combination of methods is proven to be effective. However, the acquired topography-specific models were not yet able to distinguish different metabolic phenotypical functions, such as ATP production or growth, induced by topographies. To see metabolic adjustment to the topographies, additional steps are required, for example, by adding the effects of signaling pathways to the models. The gene expression of genes involved in metabolic regulation (HIF1A and polyamine biosynthesis pathway) was promising, as those genes were upregulated in hMSCs cultured on a flat surface compared to hMSCs cultured on topographically enhanced surfaces.

Keywords: human mesenchymal stem cells, metabolism, genome-scale metabolic network, topographically enhanced surfaces

Contents

List of Figures

List of Tables

List of Abbreviations

AMEM alpha minimal essential medium.

CHO carbohydrate.

CORDA cost optimization reaction dependency assessment.

ECS extracellular space.

ESC embryonic stem cell.

FBA flux balance analysis.

FBS fetal bovine serum.

FtB flux through biomass reaction.

gDW gram dryweight.

GPR gene-protein-reaction.

GSMM genome-scale metabolic model.

HC high confidence.

HIF1A hypoxia-inducible factor 1 alpha.

hMSC human mesenchymal stem cell.

LA linoleic acid.

MC medium confidence.

NC negative confidence.

OT other reaction.

OXPHOS oxidative phosphorylation.

PCA principal component analysis.

TCA tricarboxylic acid.

1 Introduction

In medicine, there is an increasing interest in human mesenchymal stem cells (hMSC) [1]. This type of stem cells are versatile as they are multipotent; they can, for example, undergo adipogenic (fat tissue), chondrogenic (cartilage tissue), and osteogenic (bone tissue) differentiation. hMSCs also have the ability for immunomodulation, self-renewal, and proliferation i.e. population expansion [2], properties that make the use of hMSCs interesting in gene therapy, regenerative medicine, and tissue engineering [3, 4]. These cell-based therapies are proven to be effective in the treatment of different diseases such as multiple sclerosis [5], diabetes [6], and Crohn's disease [7]. One of the main advantages, compared to embryonic stem cells (ESCs), is that the cells are easily accessible as they can be isolated from various tissues including adipose tissue, bone marrow, umbilical cord, and blood [8].

To better understand the biology of hMSCs and to prepare them for therapy, hMSCs are cultured in vitro. Unfortunately, after prolonged in vitro expansion senescence will occur, with a loss of stem cell-like characteristics: cells will stop proliferating and lose their immunomodulatory and differentiating abilities [9]. This attenuation of stemness already occurs after the second passage. Nowadays, stem cells are recommended to be used in therapy within the fifth passage, as they lose their effectiveness after that [10]. At the same time a sufficient amount of cells is needed for therapy, so finding ways to maintain their stemness while expanding them would substantially benefit the development of stem cell therapies. Inside the bone marrow, the hMSCs have been observed to be in a quiescent state. In this state, they are not proliferating, but due to self-renewal, the cells can maintain their stem cell properties [11]. This quiescence is regulated by multiple factors in the bone marrow niche, such as protein 53 and retinoblastoma regulation, inhibition of cyclin-dependent kinases, Notch signaling, and the post-transcriptional regulation by miRNAs [12]. Different approaches have been used to recreate the features of the bone marrow niche in a petri dish. These approaches concerned gene manipulation, targeting signaling pathways, oxygen regulation, high-density seeding, or alteration of the culture serum among others [13, 14, 15]. In a more recent approach, the in vivo environment of the bone marrow was mimicked by culturing the cells not on a flat surface, but on a topographically enhanced surface. These enhanced surfaces impair the cells' proliferation due to a changed morphology, shrinkage of cytoplasm and nucleus, and a decreased overall cellular metabolism [16, 17].

The metabolism of the cell is a life-sustaining complex network of biochemical reactions and enzymes. It is found to be an important factor in hMSC proliferation, differentiation, and immunomodulatory function [18]. For example, undifferentiated hMSCs have a glycolytic phenotype while differentiated cells shift to a more oxidative phosphorylation (OXPHOS)-dependent metabolism [19]. Genome-scale metabolic models (GSMM) can be a useful tool to understand the effect of metabolism on the proliferation and differentiation of hMSCs cultured on topographically enhanced surfaces.

These computational models function as predictive metabolic network reconstructions. For example, they are used to understand drug targeting, production of chemicals and materials, interactions between organisms, enzyme functions, and human diseases [20]. They are constructed by integrating multiple metabolic information sources such as chemical reactions, gene-protein-reaction (GPR) associations, metabolite interactions, genome annotations, and metabolic behavior under certain circumstances. These models contain a stoichiometric matrix, which indicates the reactants and products of every reaction and their stoichiometric values. This matrix includes the laws for mass- and energy-balance. The metabolic objective of the cell is modeled by an objective function, which is a biologically interesting feature, such as ATP or biomass production. The in silico model enables the computation of an allowable flux distribution to optimize this feature.

In every tissue, different metabolic reactions can occur. The generic metabolic reconstruction needs to be adapted to the specific metabolism of the cell line or tissue of interest. There are multiple methods available to integrate transcriptional regulatory information, for example, iMAT [21], INIT [22], mCADRE [23], GIMME [24] and CORDA [25]. In this study, we use the Cost Optimization Reaction Dependency Assessment (CORDA) algorithm. This method creates a tissue-specific reconstruction retaining only those reactions with experimental evidence showing they take place in the specific tissue. This evidence can be in the form of literature or a combination of omics data and computational methods.

A GSMM has been used to study the metabolism of hMSC, but not while cultured on different surfaces [26]. This tissue-specific model, iMSC1255, was constructed with the use of publicly available transcriptomics data of hMSCs, the global Recon1 model, and the mCADRE algorithm. Later, it was used to find the influence of medium substrate on the differentiation and proliferation [27]. They found that supplementation of the medium with phosphatidylserine and phosphatidylethanolamine leads to increased proliferation. This was also verified by in vitro experiments.

The goal of this study is to find out how hMSCs metabolically adjust to culturing on topographies compared to flat surfaces. To accomplish this goal, condition-specific GSMMs of hMSCs on different surface topographies are constructed. This is done with the use of gene expression data from Vasilevich and coworkers [17]. To acquire this data, hMSCs were cultured in vitro on 8 different surfaces: a flat polystyrene surface and seven different topographically enhanced surfaces. The model will allow to identify the metabolic adjustments caused by each topography and their effect on growth rate and differentiation. We hypothesize that cells growing on topographies are better able to maintain their stem cell-like characteristics and so, reduce their metabolism and proliferation.

In the first chapter more background information is given about GSMMs, the gene expression data, and the characteristics of hMSCs. The following chapter presents the methods used to construct and analyze the hMSC- and topography-specifc models. In the third chapter, these methods will be applied to construct the tissue-specific models. When using the culture medium as input and the maximization of biomass or ATP production as cellular objective, it is possible to simulate the growth rate and energy maintenance requirements of hMSCs on different topographies.

2 Background

2.1 Characteristics of hHMCs

2.1.1 Quiescence

In the bone marrow, stem cells are in a quiescent state. This means that they are not proliferating as the cells remain in the so-called G_0 phase due to a cell cycle arrest. This phase is a reversible state as the cell cycle arrest can be lifted and stem cells can perform mitosis again. A lack of growth-promoting signals or the abundance of growth-inhibiting signals keeps the cells in this state. The reversibility of the quiescent state is important for the self-renewal of cells [28]. Also via cell-extrinsic environmental signaling the cell can go back into the cell cycle. This is necessary in case of physical tissue damage. Migration of hMSCs to the damaged tissue site will occur where they can assist in wound repair [29].

2.1.2 Metabolic regulation

The function of hMSCs is linked to their cellular metabolism, since it influences their differentiation, proliferation, and immunomodulation abilities. This makes the metabolism an interesting topic in stem cell research. During proliferation in an undifferentiated state the metabolism of hMSCs mainly relies on the supply of energy from glycolysis with a distribution of 97% glycolysis versus 3% OXPHOS [30]. A phenomenon that is called the Warburg effect and is common in rapidly proliferating cells such as cancer cells. The Warburg effect ensures the production of energy while minimizing the accumulation of reactive oxygen species [31]. During osteogenic and adipogenic differentiation a shift towards OXPHOS and an increase of oxygen consumption is observed [18, 32, 33].
This shift is regulated by the availability of oxygen, among other factors. A lack of oxygen will activate hypoxia-inducible factor 1 (HIF1A) which mediates the response of cells to hypoxia [34]. This transcription factor is capable of promoting self-renewal and proliferation and regulating differentiation in hMSCs [35, 36]. HIF1A is downregulated during osteogenic differentiation, while this factor is upregulated in undifferentiated cells [18]. HIF1A also reduces the activity of pyruvate dehydrogenase, which stimulates the conversion of pyruvate to acetyl coenzyme A, this molecule enters the tricarboxylic acid (TCA) cycle [37]. Pyruvate dehydrogenase is thus the reaction that links the glycolysis to the TCA cycle, so this reaction would be interesting to regulate, as it has a big effect on the overall metabolism. Stem cells are sensitive to mitochondrial reactive oxygen species (mROS). This is a by-product of the OXPHOS, caused by the incomplete reduction of oxygen [38]. mROS function as signaling molecules, they can regulate cell death, cell renewal, and differentiation [39]. A recent study shows that excessive mROS production could stimulate adipogenic differentiation through inhibition of the Hedgehog signaling pathway. On the other hand, the excessive production of mROS inhibits osteogenic differentiation [40].

2.1.3 Impact of mechanical microenvironment

Multiple research groups are investigating the effect of the mechanical microenvironment on hMSCs. The obtained information can benefit the design of implants and help with a better understanding of in vivo cell fate. During research regarding biomaterials used for orthopedic implants, adhesion turned out to be an important aspect of the reaction of cells to a nano-patterned substrate. A large adhesion site leads to increased intracellular tension and changes in the cytoskeleton, these changes in turn result downstream in osteogenic differentiation [41].

Another study found a connection between the width of micro-scale ridges on polymide culture surfaces and differentiation capabilities of hMSCs. A 15 µm ridge enhanced adipogenic differentiation, whereas a 2 µm ridge increased osteogenic differentiation. However, this differentiation was only observed when the cells were cultured in differentiation media [42].
Research concerning the relation between hMSC cell shape (induced by surface topographies) and the cells' phenotypical functions, showed a positive correlation between cell area and proliferation. Changes were observed in cell-cycle signaling and in the signaling pathways JAK, IKK, and HIF in cells cultured on topographies [17].

2.2 Recon3D

Recon3D is, together with HMR 2.0, a widely used and extensive GSMM [43]. These generic models include all reactions proved to take place in a human body. The Recon3D model consists of 10600 reactions and 5835 metabolites spread over eight different cellular compartments: cytosol, nucleus, lysosome, endoplasmatic reticulum, mitochondria, peroxisome/glyoxysome, Golgi apparatus, and the inner mitochondrial compartment and the extracellular space (ECS). The network is defined by a stoichiometric matrix (**S**), with m (number of metabolites) rows and n (number of reactions) columns. This matrix is filled with stoichiometric coefficients that indicate the contribution of a metabolite to a reaction. A negative coefficient indicates a consumed metabolite, a positive coefficient indicates a produced metabolite and a zero means that the metabolite does not participate in that reaction. Extracellular metabolites can enter or leave the model through exchange reactions. Via these reactions, it is possible to indicate the metabolites in the cells surrounding which are available for uptake, simulating the culture media. An example of a a small metabolic network and its corresponding stoichiometric matrix are given in Figure 1. To regulate the amount of flux a reaction is able to carry, reactions are restricted with a lower and upper boundary. These boundaries can be adjusted to fit biological characteristics and thermodynamic constraints.

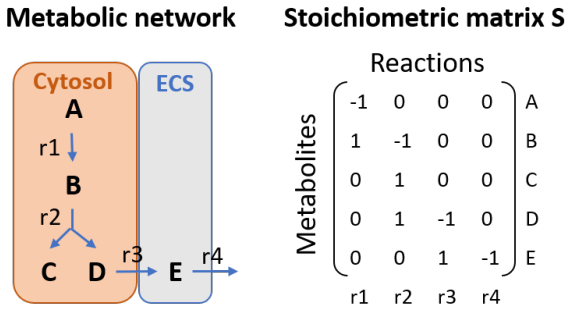

Figure 1: On the left a toy model of a small metabolic network and on the right its corresponding stoichiometric matrix **S**. The metabolic network consists of five metabolites (A - E) and four reactions (r1 - r4) divided over two compartments, the cytosol and the extracellular space. Reaction r3 is a transport reaction between the two compartments. Reaction r4 is an exchange reaction, it excretes metabolite E from the ECS. This is indicated in the stoichiometrix matrix with a -1 in column r4, row E.

The model also includes 2248 metabolic genes. 5938 of the 10600 reactions are associated with these genes through gene-reaction-rules. Furthermore, Recon3D contains gene and metabolite annotations, protein and metabolite 3D structural information, and metabolite formulas. The biomass reaction is an

artificial reaction, that makes it possible to simulate the growth rate of a cell or organism. This reaction models the synthesis of all known precursors needed to build a cell [44, 45]. A positive flux through the biomass reaction (FtB) simulates the generation of cellular biomass products from available inputs. A FtB of 1 mmol gDW^{-1} h^{-1} is in correspondence with a growth rate of $1\ h^{-1}$ [46].

2.3 Data

hMSCs were obtained from human bone marrow of a 29 years old female (donor d016). The cells are harvested from the acetabulum part of the hip bone [47]. In all experiments, cells at passage 4 are used. This means that the cells have been subcultured four times; they have been harvested and reseeded into multiple daughter cell cultures. The stem cells were cultured on polystyrene surfaces from TopoChip with randomly generated patterns [48]. These patterns contain elements with a height of 5 µm and they are build up by combining three primitive shapes: circles, triangles and rectangles. Alpha minimal essential medium (AMEM, Gibco 22-571-038) supplemented with 10% fetal bovine serum (FBS, Sigma-Aldrich batch number 013M3396) was used as culture medium. The seeding density was 15,000 cells/cm^2. After 24 hours of culturing, total RNA was isolated from the cells. Microarray analysis was performed using Illumina Beadchips to measure gene expression. The gene expression data was further processed with the online portal arrayanalysis.org and quantile normalization was applied. The principal component analysis (PCA) of the gene expression data is given in Appendix A.
Besides gene expression, information about proliferation rate was acquired. With the use of an EdU assay, cells that have undergone DNA synthesis were stained after 24 hours of culturing on the topographies. Images were acquired and the number of stained cells was counted with image analysis. With the following formula, the growth rate is calculated:

$$N_{\text{total}} = N_0(1+r)^t \quad \text{with} \quad N_0 = N_{\text{total}} - 0.5 * N_{\text{EdU}} \tag{1}$$

N_{total} is the total number of cells, N_0 the number of cells at time t = 0, r the growth rate, t the time, which is in this case 24 hours and N_{EdU} the number of cells that are EdU stained after 24 hours. For more information about the culturing process, see the article of Vasilevic et al. [17]. All the experiments are carried out in triplicate on one flat surface and on seven different topographically enhanced surfaces. This gives a total of 24 samples.

3 Methods

3.1 Simulation details

For the modeling, the COnstrained-Based Reconstruction and Analysis for Python (COBRApy) package is used [49]. This software package supports the representation and analysis of complex metabolic networks. It includes model building functions and functions to read and write models from/to files. To visualize the Recon3D model, ReconMap is utilized with the web interface of Virtual Metabolic Human [50]. In this study, the Systems Biology Markup Language (SBML) format is used to describe the model. This is an XML-based standard format for communicating and storing computational models of biological processes [51]. The package contains multiple functions to simulate the metabolism in such a model. Most of those functions are based on the mathematical approach of Flux Balance Analysis (FBA). Furthermore, functions are included in the package concerning gap-filling, growth media, and gene knockouts.

3.2 Flux balance analysis

FBA is a constraint-based modeling approach which simulates the metabolism of an organism using GSMMs [46]. With the use of linear programming, it calculates the flow of metabolites through the network. So, besides the stoichiometric matrix **S** of *m* rows and *n* columns to describe the network, the vector **v** represents the fluxes through all the *n* reactions. The lower and upper boundaries for each reaction are given in vectors **lb** and **ub** respectively. This gives the following system at steady state:

$$\mathbf{Sv} = 0 \quad \text{with} \quad lb_i \leq v_i \leq ub_i \tag{2}$$

This system of equations can be solved using linear programming. As with all GSMMs, $n > m$, in other words, there are more unknown variables than equations, which means there is no unique solution. A large solution space is the result with multiple feasible flux distributions **v**. To narrow down the solution space an objective function can be assigned. In vector **c** of length *n* it is indicated how much each reaction contributes to the objective function and whether its flux should be minimized ($c_i < 0$) or maximized ($c_i > 0$). FBA now finds a flux distribution **v** that satisfies the optimization of:

$$\max(\mathbf{c}^\mathbf{T}\mathbf{v}) \tag{3}$$

This could also lead to multiple solutions, but the solution space is much smaller. Parsimonious FBA (pFBA) is a variation of FBA. This method optimizes the flux through the objective function, while also minimizing the sum of absolute fluxes through the model.

3.3 CORDA

The CORDA algorithm is used to build the hMSC tissue-specific model [25]. This algorithm selects reactions from a general reconstruction, in this case Recon3D, to build a tissue-specific reconstruction. As input, the algorithm requires all the reactions of the general reconstruction divided into four categories. Those groups are specified as follows:

- High confidence (HC) reactions: Reactions that should definitely be in the tissue-specific model.
- Medium confidence MC) reactions: These reactions could be in the model if they are not depen-

dent on negative confidence reactions that are associated with only a few HC or MC reactions.

- Negative confidence (NC) reactions: Reactions that should not be in the model, unless they are associated with a HC reaction or multiple MC reactions.
- Other reaction (OT): These are the remaining reactions of the general reconstruction which are not included in the HC, MC, or NC group.

The assignment of reactions to a specific group is based on gene expression, proteomics data, or other biochemical information available for that cell type. The algorithm starts by constructing a tissue-specific model that only consists of the HC reactions. With the use of dependency assessment, MC, NC, and OT reactions are added to the tissue-specific reconstruction. In this study, gene expression data of the hMSCs grown on flat surfaces is used to construct the tissue-specific model. The expression of the metabolic genes is mapped to the reactions using GPR rules. This gave a confidence per reaction. For more information about reaction confidence, see the next section. The 4662 reactions that are not associated with any genes are classified as OT reactions. Of the other 5938 reactions, the one-third of the reactions with the highest confidence are classified as HC, the one third with the lowest confidence are labeled as NC and the middle 33% as MC. Furthermore, the biomass reaction and several reactions related to essential energy pathways are also classified as HC, regardless of their gene expression. These reactions are part of the glycolysis, tricaboxylic acid (TCA) cycle, urea cycle, OXPHOS, pentose phosphate pathway, or nucleotide interconversion. An overview of these reactions can be found in Appendix B.

3.4 Reaction confidence

Gene expression can be mapped to reactions via the GPR associations that are included in the model. These associations are based on enzymes encoded by the genes and the reactions that these enzymes catalyze. Boolean logic is used to describe the relationship between a gene and a reaction. In Figure 2 four examples are given of such a GPR association. Reaction 1 is able to carry flux if gene A is expressed. Enzyme B can catalyze both reactions 2 and 3. The third example shows that only one of the genes has to be expressed for reaction 4 to be able to carry flux. The AND statement indicates that the reaction is only able to carry flux if both genes E and F are expressed. This information is used to calculate a confidence number for every reaction.

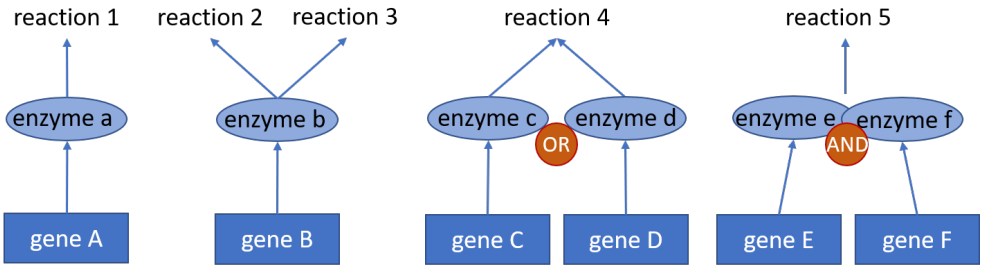

Figure 2: Examples of gene-protein-reaction (GPR) associations. Reaction 1 is catalyzed by enzyme a which is the product of gene A. Gene B encodes for enzyme b. This enzyme can catalyze reaction 2 and reaction 3. Enzyme b is a promiscuous enzyme. Gene C and D encode for isozymes c and d, both enzymes can catalyse reaction 4. Enzymes e and f form an enzyme complex, both enzymes are required to catalyze reaction 5. The figure is adapted from Jensen [52].

To calculate a confidence score for a reaction, the expression of a gene is not a boolean but the gene expression value. The confidence number is the result of the GPR association. So, if gene C has a gene expression value of 10 and gene D a value of 20, then the confidence number of reaction 4 is 20. Reaction 5 will have a confidence score of 10 if gene E has an expression value of 10 and gene F a value of 20.
The dataset contained gene expression information in the form of Illumina data. Recon3D uses entrez ID as gene identifier. Not all genes in the model are directly connected to the Illumina genes, for some genes the expression is retrieved from the data consulting its BiGG ID, synonyms, name, or with the help of the KEGG database. For gene 10050785AT1 there is no expression data available. When multiple Illumina probes are provided for a single gene, the average gene expression of those probes is taken.

4 Model construction

The construction of the hMSC- and topography-specific models will be explained in this chapter. This process is illustrated in Figure 3. At first, a tissue-specific model for hMSCs is constructed with the use of the CORDA algorithm. The general reconstruction Recon3D, the assignment of important energy pathways, and the genome data of the hMSCs cultured on flat surfaces are the input for this algorithm. The composition of the culture medium is set as input and a biomass objective function is added to finalize the hMSC-specific model. The final model is referred to as the hMSC-base model. By constraining the fluxes through the reactions of the hMSC-base model based on gene expression data, the topography-specific models are constructed. These steps are further elucidated in this chapter.

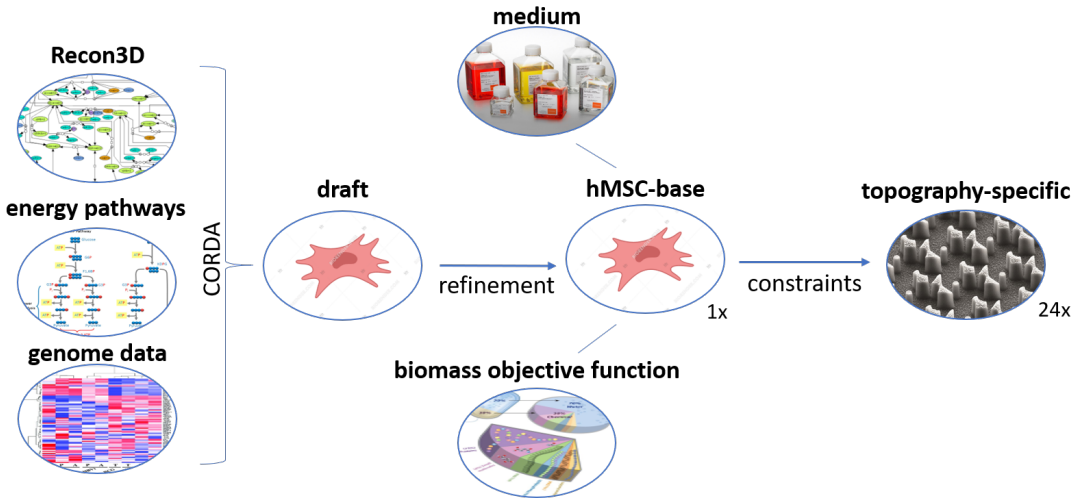

Figure 3: A schematic representation of the process of constructing a hMSC-base and 24 topography-specific metabolic reconstructions. The first step is applying the CORDA algorithm to the Recon3D reconstruction to omit reactions, based on genome data, that are not likely to be active in hMSCs. The draft model returned by CORDA is improved during a refinement step. The culture medium is set as input and the biomass objective function is added to the model. Constraining reactions based on gene expression data of hMSCs cultured on topographies leads to the construction of 24 topography-specific models.

4.1 hMSC-base model

The CORDA algorithm, as explained in the methods section, is used to build the hMSC-base model. The model returned by the CORDA algorithm is not yet complete, as the model is not able to produce all components of the biomass reaction. This is due to a lack of gene-associated reactions. To fix this problem, three refinement steps are performed. The first one is a dead end analysis in which dead end metabolites are detected with the detectDeadEnds function in the COBRA toolbox in MATLAB [53]. Those metabolites are only connected with one reaction to the reconstruction graph. Because of this, no flux can flow through that reaction or a series of reactions connected to it. Two operations can be performed to repair this. The first is to add a reaction from the general reconstruction (Recon3D) to make the metabolite fully connected (with two reactions) to the hMSC-specific reconstruction. The second operation is to remove the reaction in which the metabolite occurs. This last step can only be done if the reaction does not contribute to the model.
Secondly, exchange reactions are added to make sure the uptake, and possible secretion, of metabolites

that are present in the medium is fully supported. In some cases, it was not enough to only add exchange reactions, because the transport reaction from the ECS to the cytosol was still missing. In that case, the transport reaction is also added to the model. It is important to notice that only reactions that are present in the general reconstruction, are added.
The last step in the refinement process is to fill some gaps that retained the model form being able to optimize for biomass production with the given medium composition as input. A pruning method is used to do this. All reactions of the Recon3D reconstruction are again added to the hMSC-specific model. The model is then optimized to produce biomass. One-by-one the reactions are removed from the model and if the model is still able to optimize for biomass after the removal, that reaction is definitely removed from the model. Else, the reaction is added to the hMSC-specific model.
The two steps mentioned above are executed three times with gene expression data collected from cells cultured on three equal flat surfaces. To make the model as robust as possible, the three acquired models are merged, to create a model with the reactions from all three models. The three tissue-specific models of stem cells on a flat surface are very similar. The flat 1 model consists of 4545 reactions, flat 2 of 4535 reactions, and flat 3 of 4542 reactions. A total of 4457 reactions occur in all three of the models as could be seen in Figure 4. Merging those three models gives the hMSC-base model. This model contains 4625 reactions.

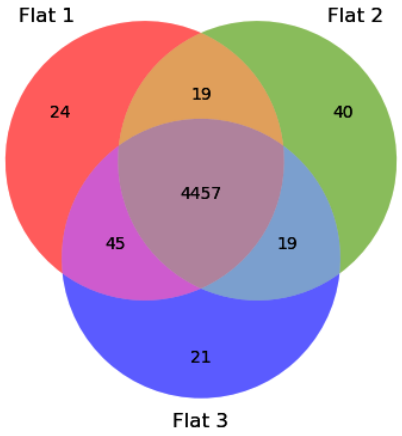

Figure 4: Venn diagram representing the number of common reactions of the three hMSC-specific models on a flat surface. These models are constructed by applying the CORDA algorithm and a refinement step. Gene expression of hMSCs cultured on flat polystyrene surfaces is used as input for the CORDA algorithm.

4.1.1 In silico culture medium

A set of input metabolites is required as input for the model and needs to be defined. The culture medium is a combination of alpha-minimal essential medium (a-MEM) and fetal bovine serum (FBS). The first is a fixed medium, while the composition of FBS can vary per batch. The composition of the input used in this study is based on the average concentration of metabolites in FBS [54]. The in silico culture medium is introduced to the model via the so-called exchange reactions. A negative flux through this reaction indicates the influx of a metabolite to the ECS of the model, a positive flux indicates excretion of the metabolite. The size of this flux can be restricted with an upper and a lower boundary. Excretion of all metabolites is unconstrained, therefore the upperbound of all exchange reactions is

set to 1000 mmol gDW^{-1} h^{-1}. Only metabolites that are available to the cells, i.e. metabolites that are in the culture medium, are allowed to enter the model. The lowerbound of exchange reactions of metabolites that are not allowed to enter the model is set to zero. There are two ways to allow medium metabolites to enter the model.
In the first case, the lowerbound is set equal to the concentration of that metabolite in the medium. This is referred to as *true medium*. Exact values of concentrations can be found in Appendix C. In this case it is assumed that the concentration of a metabolite in the medium is equal to the uptake-rate of that metabolite. The second method does not rely on this assumption but simulates the infinite availability of the metabolites. In this case, the lowerbound of the exchange reaction of medium metabolites is set to a large value such as -1000 mmol gDW^{-1} h^{-1}. This is referred to as *infinite medium*. This approach is used by Yousefi and coworkers [55]. The uptake of water, oxygen, hydrogen, and carbon dioxide is unconstrained in both cases.
Linoleic acid (LA) and alpha-linolenic acid are two essential fatty acids. In Recon3D only one of those lipids is essential for the production of biomass [56]. As the average concentration of alpha-linolenic acid in FBS is zero, only LA is present in the in silico medium with a concentration of 9.52e-8 mM. This value turned out to be too low as LA is required in the formation of different phospholipids, important building blocks of the cell membrane [57]. A value of 0.005 mM is chosen, similar to the work of Fouladiha and coworkers to ensure enough LA is available to the model [26].

4.1.2 Biomass objective function

The biomass reaction contains the synthesis of all precursors necessary to build a new cell. Those precursors can be divided into five macromolecular classes: proteins, lipids, DNA, RNA, and carbohydrates (CHO). Within these classes, it is specified how much each metabolite contributes to that class. For example, the proteins class contains amino acids such as alanine and arginine and the lipid class covers metabolites such as cardiolipin, sphingomyelin, and phosphatidylethanolamine. Besides the macromolecular classes, energy is required in the form of ATP to account for biosynthetic processes that are not included in the model. RNA error checking in transcription and the polymerization of amino acids into proteins are two of those processes [58]. This results in the following reaction:

$$x_i\text{Protein} + x_i\text{Lipid} + x_i\text{DNA} + x_i\text{RNA} + x_i\text{CHO} + x_i\text{ATP} + x_i\text{h2o} \rightarrow x_i\text{ADP} + x_i\text{h} + x_i\text{pi}$$

In which the values of x stand for the fractional contribution of each group/metabolite to the cellular biomass. Of course, not all cells have the same macromolecular composition, and so the biomass reaction varies between species, cell types, and physiological conditions [59]. As the exact macromolecular composition of hMSCs is unknown, constructing a realistic biomass reaction is a challenge. That's why multiple biomass reactions are considered and tested. Table 1 gives an overview of these biomass reactions.

Table 1: Overview of tested biomass reactions of human cells.

Reaction	Model	Cell type	Contribution in weight% Protein|Lipid|RNA|DNA|CHO	ATP yield	Source
BIOMASS	Recon2	Generic	66.6 | 17.8 | 7.7 | 1.8 | 6.1	20.65	[44]
Lipids	Recon2	Generic/hMSc	66.6 | 17.8 | 7.7 | 1.8 | 6.1	20.65	[44, 60]
Folger	Recon1	Liver/muscle	66.6 | 14.2 | 3.7 | 0.9 | 15.2	x	[61]
Opdam	Recon1	Generic/cancer	72.1 | 8.0 | 5.4 | 1.4 | 5.5	30	[62]
Bordbar	Recon1	Macrophage	75.3 | 7.6 | 6.1 | 2.4 | 8.6	25.17	[63]

The BIOMASS reaction is the primary reaction present in the Recon2 model. This reaction is not specific for a cell type. The Lipids reaction is similar to the BIOMASS reaction, only the metabolite contribution to the lipid class is adjusted based on the information of Levental and coworkers who researched the lipid composition of the plasma membrane of hMSCs. The other biomass reactions describe the composition of liver/muscle cells, cancer cells, and macrophages. The effect of the biomass reactions is tested by analyzing the maximal flux through the reaction in the hMSC-base model with a true and infinite medium. The results are given in Table 2.

Table 2: Optimization of the hMSC-base model for the production of the different biomass reactions with a true and infinite medium.

Reaction	FTB [mmol gDW^{-1} h^{-1}]	
	True medium	Infinite medium
BIOMASS	0.009	116
Lipids	0.015	123
Folger	0.035	1000
Folger + ATP yield	0.035	96
Opdam	0.015	90
Bordbar	0.032	109

As the Folger biomass reaction did not have an ATP yield, the FtB with the infinite medium is 1000 mmol gDW^{-1} h^{-1}, which is the maximum flux of the model. Adding the same ATP yield as Opdam to the Folger reaction (30 mmol gDW^{-1} h^{-1}) the models are better comparable. Furthermore, the sensitivity of each coefficient of the biomass metabolites is tested. The coefficient is increased and decreased by 30% and the effect is measured on the FtB. With the true medium as input, the coefficient of phosphatidylcholine is the most sensitive: changing this coefficient by 30% resulted in a change of FtB between -23.0% and +22.0%. The sensitivity of the other metabolites is a lot lower. The sensitivity of phosphatidylcholine could be explained by the structure of the model. Since this metabolite is interconvertible into multiple other metabolites of the lipid class, such as phosphatidylglycerol, phosphatidylserine, and sphingomyelin. With the infinite medium as input, the FtB is the most sensitive to the coefficient of ATP in all biomass reaction. A lower coefficient leads to a lower FtB. Changing this coefficient by 30% resulted in a change of FtB between -23.0% and +42.6%.
According to the data of Vasilevich et al. the average growth rate of hMSCs on a flat surface is equal to 0.015 h^{-1}. With the true medium, the Lipids and Opdam reactions result in the same FtB. The Lipids reaction is then in favor because it is altered towards hMSCs. The exact composition of the Lipids biomass reaction can be found in Appendix D.

4.1.3 Model characteristics

Table 3 shows the main characteristic of the hMSC-base model. The resulting model is less than half the size of the Recon3D model, considering the number of reactions. The FtB with infinite medium is higher in the Recon3D model. In this general model, all possible pathways can be used to produce the components of the biomass reaction. The FtB with the true medium is identical. This suggests that in this case, the FtB does not rely on the number of reactions in the model, but on the supply of available medium.

Table 3: Comparison between the generalized Recon3D model and the hMSC-base model.

	Number of reactions	Number of metabolites	Number of genes	FtB [mmol gDW^{-1} h^{-1}] True medium	FtB [mmol gDW^{-1} h^{-1}] Infinite medium
Recon3D	10601	5835	2248	249	0.015
hMSC	4626	3193	1715	139	0.015

A representation of the distribution of reactions over the subsystems can be found in Figure 5. The fatty acid oxidation is the system with the most reactions. This could be explained by the extensiveness of the lipid metabolism in the general Recon3D reconstruction. Reactions concerning the transport of metabolites between compartments is also highly presented. Extracellular transport is the shift of metabolites between the cytosol and the ECS. Intracellular transport contains the remaining transport reactions in the cell. The category "Others" contains the subsystems: pentose phosphate pathway, TCA cycle, OXPHOS, mROS detoxification, and more.

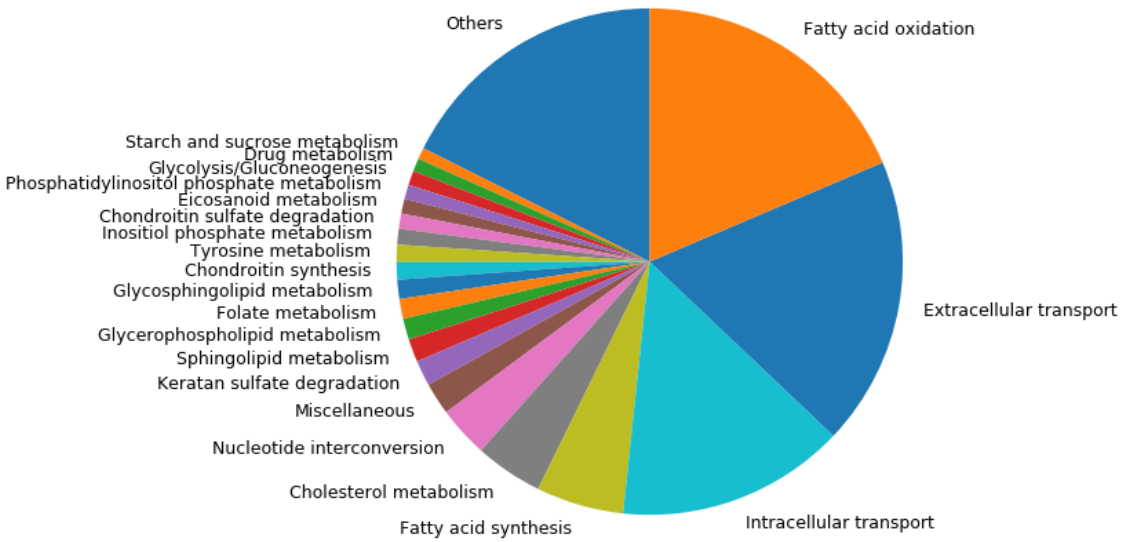

Figure 5: Distribution of the reactions in the hMSC-base model per subsystem.

4.2 Topography-specific models

To adjust the hMSC-base model specifically for the cells cultured on topographies, constraints are added to the model. This is the last step of the process depicted in Figure 3. As no reactions are added or removed, the topography-specific models contain the same number of reactions as the hMSC-base model.

Gene expression data is available of 24 cell cultures, so 24 different topography-specific models are created. Three models represent the cells grown on a flat surface and the other 21 represent cells grown on 7 different topographically enhanced surfaces. At first, the reaction confidence is determined for every reaction by mapping the gene expression to the hMSC-base model. Then, the reactions are constrained by adjusting the maximal flux the reaction is able to carry. This is a valve-based method, the maximum capacity of a reaction is limited by the gene expression of its associated enzymes. The reduction of capacity is imposed by altering the boundaries of that reaction. Per default, the boundaries of reactions are set to 1000 mmol gDW^{-1} h^{-1}. With this method, all boundaries are set linearly proportional to their reaction confidence, via the following formula:

$$\text{boundary} = lf + \frac{hf}{hc}\, conf \quad \text{with} \quad lf \leq \text{boundary} \leq hf \tag{4}$$

In which $conf$ is the confidence number, hc the highest confidence number, hf the highest boundary value, which is set to 3 and lf the lowest boundary value which is 0. Reactions with no confidence number, in other words, with no associated genes, are left with their default boundary values. This approach is similar to the E-Flux method, although different parameters are used [64]. Instead of a linear approach, a logarithmic approach is also evaluated. However, this did not improve the resulting models.

5 Results

5.1 In silico metabolism

In all topography-specific models, the cells mainly have a glycolytic metabolism. Only a small part (0.8%-4.9%) of the ATP is produced in the mitochondria, by the TCA cycle and the OXPHOS. This came forward by optimizing the models for ATP production and using parsimonious FBA (pFBA) to solve for the flux distribution. During glycolysis, ATP is produced by Phosphoglycerate kinase and pyruvate kinase reactions. In the TCA cycle ATP is produced by the succinyl coenzyme A synthetase reaction and in OXPHOS by ATP synthetase. In Figure 6, the ATP production and consumption for every topography-specific model is depicted. The difference between production and consumption gives the quantity of ATP that could be used for processes such as growth or protein biosynthesis. ATP is mainly consumed in the first part of the glycolysis by hexokinase and phosphofructokinase. Other consuming reactions are part of the fructose and mannose metabolism pathway, the pentose phosphate pathway, and the nucleotide interconversion.

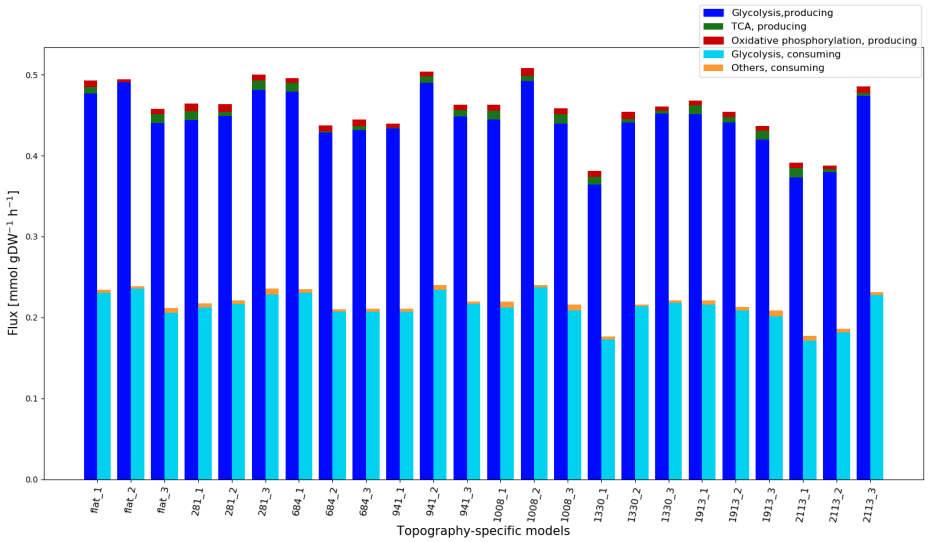

Figure 6: ATP production (dark colors) and consumption (light colors) per subsystem for every topography-specific model. On the x-axis the topography-specific models are represented. The y-axis gives the production or consumption of ATP in mmol gDW^{-1} h^{-1}. The ATP production and consumption is given per subsystem.

There is no significant difference in mitochondrial ($P = 0.67$), glycolytic ($P = 0.16$), or total ($P = 0.11$) ATP production between the models depicting flat and topographically enhanced surfaces.
The uptake of glucose in the topography-specific models is between 0.088 and 0.119 mmol gDW^{-1} h^{-1} and the excretion rate of lactate is between 0.173 and 0.239 mmol gDW^{-1} h^{-1}, leading to a ratio $Y_{lac/gluc}$ of 1.90 - 2.02. The oxygen uptake lies within the range of 0.005 - 0.022 mmol gDW^{-1} h^{-1}. Cutting off the oxygen supply in the models results in solely glycolytic (anaerobic) energy production. The production of reactive oxygen species is very limited.

5.2 In silico growth

The optimization of a model for the production of biomass is a good way to simulate growth of the cells as it shows the cells' ability to synthesize the precursors that are needed for proliferation. In this section, the flux distributions of the topography-specific models are optimized for the production of biomass.

5.2.1 Experimental growth rate

As the FtB is equal to the growth rate per hour, the results of the model can be compared to the experimentally measured growth rates. In Figure 7, the experimentally measured growth rates and the in silico growth rates are given. For the samples 941_1 and 941_2 there is no experimental growth rate information available. The experimental growth rates range from 0.0023 - 0.0187 h^{-1}. The in silico measured growth rates fall within that range. The differences in in silico growth rates are very small compared to the differences measured in the experimental setting. There is not a good agreement between the in silico predicted and the experimentally measured growth rates.

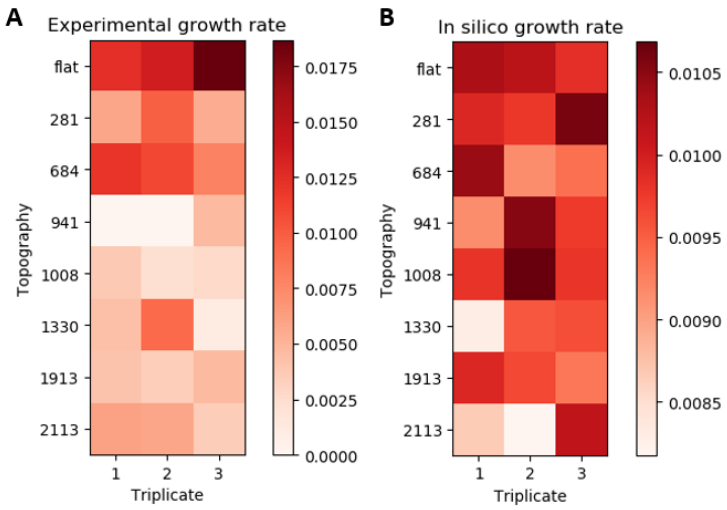

Figure 7: A. Experimentally measured growth rates of hMSCs cultured on different topographies. B. In silico growth rates predicted by the topography-specific models. On the y-axis the topographies are depicted. The x-axis represents the triplicates. The colorbars map the growth rate values in h^{-1} to their color.

5.2.2 Essential nutrients

The correctness of the model is tested by the essentiality of different nutrients for cell maintenance. A nutrient is considered essential when the model is not able to produce all biomass components if the metabolites' influx is restrained. This is indeed the case for all essential amino acids (His, Ile, Leu, Lys, Met, Phe, Thr, Trp, and Val). For the nonessential amino acids (Ala, Asn, Asp, Glu, Ser) however, the influx of serine turned out to be indispensable. This was due to the missing phosphoserine phosphatase (PSP_L) reaction. When this reaction is added to the model, it is able to synthesize all biomass components without the influx of serine. Regarding the essential fatty acids (LA and alpha-linolenic acid), only LA is essential in the Recon3D reconstruction. In all topography-specific models,

the withholding of LA leads to an FtB smaller then $1.3{\cdot}10^{-7}$ mmol gDW^{-1} h^{-1}. The pathways of vitamins are not fully integrated in the Recon3D structure, this makes it impossible to test their essentiality.

5.2.3 Reactions and genes limiting in silico growth

To find out which reactions are restraining the in silico growth a deeper analysis is performed. During this analysis, it becomes clear that the energy production to meet the ATP yield, which is enclosed in the biomass reaction, is restraining in silico growth. This limitation of ATP production is not due to the supply of glucose or any other medium metabolite, but it's due to the constraints that are put on reactions based on enzymatic gene expressions associated with them. Reactions are considered limiting as their flux is equal to their boundary and setting their boundary wider immediately results in a bigger FtB. Table 4 shows the limiting reactions and the genes that are responsible for the bounds on these reactions. As can be seen in Figure 8, all the reactions are directly or indirectly connected to the glycolysis, TCA cycle, or OXPHOS.

Table 4: Limiting reactions and genes that restrict in silico growth in topography-specific models. These reactions are indicated in Figure 8.

Reaction name	**Limiting gene expression**	**System**	**Indication**
PGK	PGK1	Glycolysis	A
PGM	PGAM1	Glycolysis	B
DGSNt, DADNt4	SLC29A2	Transport $\rightarrow$ Glycolysis	C
CYOOm2i, CYOOm3i	COX4I2, COX6A2, COX6B2, COX7B2, COX8C	OxPhos	D
ATPS4mi	ATP5S	OxPhos	E

The fact that ATPS4mi and PGK reactions are limiting could easily be explained as they are ATP producing reactions. A higher flux through these reactions would right away lead to a higher ATP production. The PGK reaction is followed by the PGM reaction. A lower flux through the PGK reaction results in a lower flux through the PGM reaction, vice versa, this is also the case. The DGSNt and DADNt4 reactions supply glyceraldehyde 3-phosphate (g3p) to the glycolysis. DGSNt and DADNt4 transport deoxyguanosine and deoxyadenosine respectively to the cytosol, where it is converted to g3p during three successive reactions: PUNP2/4, PPM2, DRPA which are part of the purine catabolism (PUNP4) and the pentose phosphate pathway (PPM2 and DRPA). This way, less flux will need to flow through the first part of the glycolysis (conversion of glucose to g3p), where the reactions HEX1 and PFK consume ATP.
The reactions CYOOm2i and CYOOm3i represent complex IV of the OXPHOS. The only difference between the two reactions is the fact that, next to water, reactive oxygen species are released in the CYOOm3i reaction. A wider constraint of these reactions leads to a bigger flux through part of the TCA cycle. This leads to more ATP production by the SUCOASm reaction, during the conversion of succinyl coenzyme A to succinate.

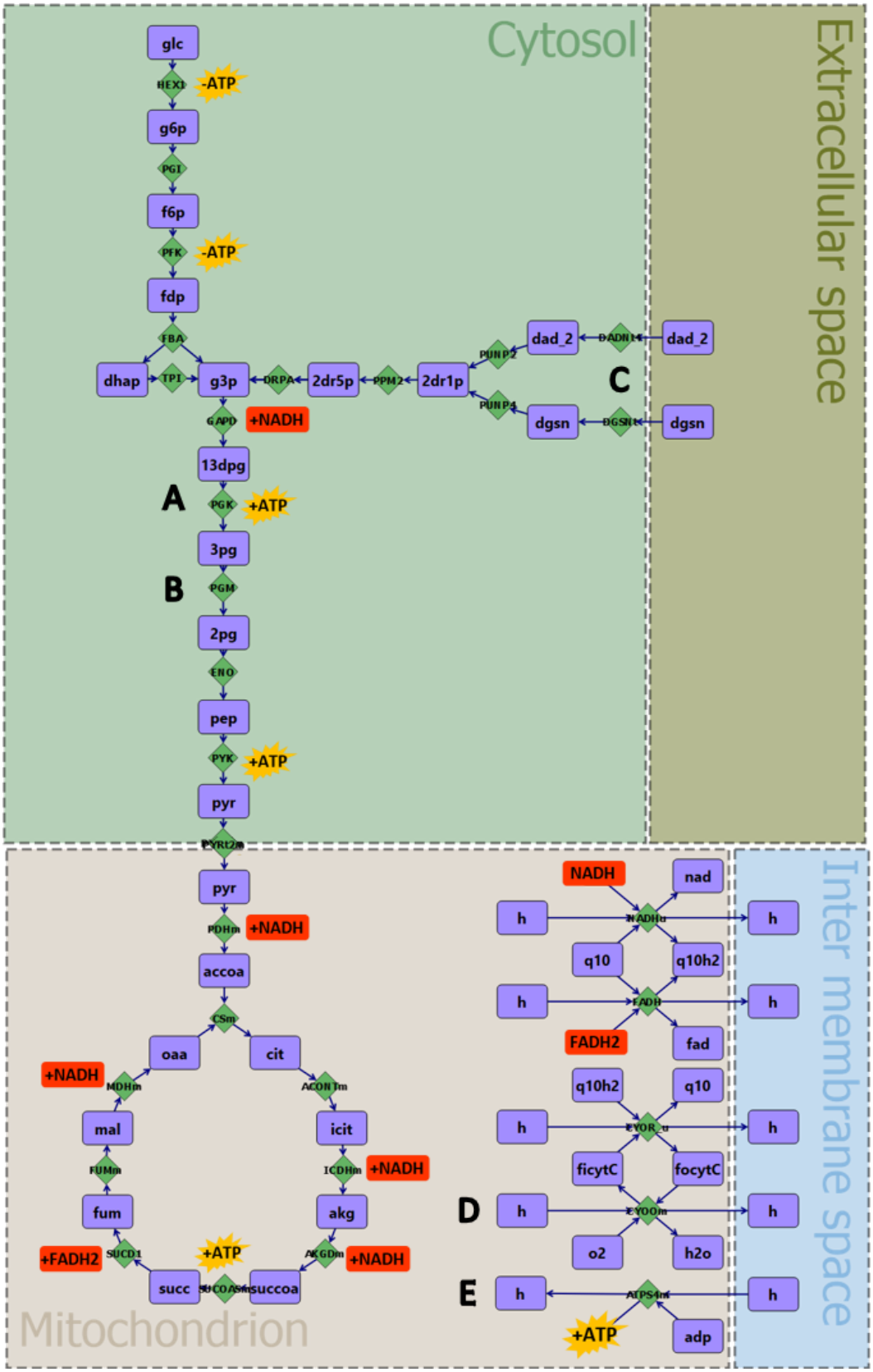
Cytosol
Extracellular space
glc
HEX1
-ATP
g6p
PGI
f6p
PFK
-ATP
fdp
FBA
dhap
TPI
g3p
DRPA
2dr5p
PPM2
2dr1p
PUNP2
PUNP4
dad_2
DADN1
dad_2
dgsn
DGSN1
dgsn
C
GAPD
+NADH
13dpg
A
PGK
+ATP
3pg
B
PGM
2pg
ENO
pep
PYK
+ATP
pyr
PYRt2m
pyr
PDHm
+NADH
accoa
CSm
oaa
cit
+NADH
MDHm
ACONTm
mal
icit
FUMm
ICDHm
+NADH
fum
akg
+FADH2
SUCD1
+ATP
AKGDm
+NADH
succ
SUCOASm
succoa
Mitochondrion
NADH
nad
h
NADHm
h
q10
q10h2
h
FADH
h
FADH2
fad
q10h2
q10
h
CYOR_u
h
ficytC
focytC
D
h
CYOOm
h
o2
h2o
E
h
ATPS4m
h
+ATP
adp
Inter membrane space

Figure 8: Representation of the main energy-producing pathways, glycolysis, TCA cycle, OXPHOS, and a small part of the pentose phosphate pathway. Letters A to E indicate the reactions that are limiting the in silico growth. This image is created with the use of the Omix Visualization toolbox [65]. Metabolite abbreviations: glc, glucose; g6p, glucose-6-phosphate; f6p, fructose-6-phosphate; fdp, D-fructose 1,6-bisphosphate; dhap, dihydroxyacetone phosphate; g3p, glyceraldehyde 3-phosphate; 2dr5p,2-deoxy-D-ribose 5-phosphate; 2dr1p, 2-deoxy-D-ribose 1-phosphate; dad_2, deoxyadenosine; dgsn, deoxyguanosine; 13dpg, 3-phospho-D-glyceroyl phosphate; 3pg, 3-phospho-D-glycerate; 2pg, D-glycerate 2-phosphate; pep, phosphoenol pyruvate; pyr, pyruvate; accoa, acetyl coenzyme A; oaa, oxalacetic acid; cit, citrate; icit, isocitrate; akg, 2-oxoglutarate; succoa, succinyl coenzyme A; suc, succinate; fum, fumarate; mal, malate; q10, ubiquinone-10; q10h2, ubiquinol-10; ficytC, ferricytochrome C; focytC, ferrocytochrome C.

5.2.4 Medium composition

The composition of the medium was expected to have an influence on the FtB. This influence was tested by altering the influx of every metabolites in the medium by 30%. It turns out to be true that some medium metabolites influence the FtB, namely cholesterol and LA. Increasing the influx of cholesterol by 30% leads to an increase of FtB between 1.29% and 1.79%. The same effect is seen if the influx is reduced by 30%. This resulted in a decrease of the FtB between 1.53% and 2.15%. By adding more cholesterol to the model, the consumption of ATP decreases as less ATP is needed for the Acetyl-Coa carboxylation in the mitochondrion. This is a step in the process of fatty acid oxidation. As stated before, ATP is the limiting factor in the optimization of biomass production.
Increasing the influx of LA does not affect the FtB, a decrease of 30% results in a decrease of FtB only in the models 281_3 and 1008_1. In the other models it has no effect. The topography-specific models in which the decrease is observed are the models with the highest FtB. This means that these models also have the highest yield for building blocks for membrane lipids such as (glycero)phospholipids. LA is essential for the production of the phospholipids phosphatidic acid, phosphatidylcholine, and phosphatidylserine among others. When the influx of LA is below 0.33 times the FtB, the shortage of LA limits the production of phospholipids and so the production of biomass. So, when the influx of LA drops below 0.33 times the FtB the production of phospholipids becomes the limiting factor instead of the production of ATP.

5.2.5 Gene knockout

A single gene knockout is performed to identify lethal genes. A gene is considered lethal if a knockout of this gene results in the model not being able to produce all the components of the biomass reaction. The following genes were found to be lethal, sorted by their subsystem:

- **Sphingolipid metabolism:** SPTLC1, SPTLC2, SPTLC3, UGCG, KDSR, SGMS1
- **Glycerophospholipid metabolism:** GPD1, PCYT2, PGS1, LCAT
- **Transport:** SLC7A5, COL4A3BP, SLC27A5
- **Nucleotide interconversion:** GUK1, CMPK1
- **Pentose phosphate pathway:** RPIA
- **Cholesterol metabolism:** LIPA

The majority of these genes are associated with the lipid metabolism. An explanation could be that there is only one pathway to build the specific lipid metabolites of the biomass reaction. For the production of the other macromolecular classes, there are more alternative routes. In case of the amino acids, these metabolites are provided in the medium, so the synthesis of these metabolites is not necessary. Only a transport reaction from the ECS to the cytoplasm is required.

5.3 Metabolic regulation by HIF1A

HIF1A is an important factor in the regulation of the glycolysis. Therefore, the gene expression of genes associated with HIF1A and its signaling pathways are given in Figure 9.

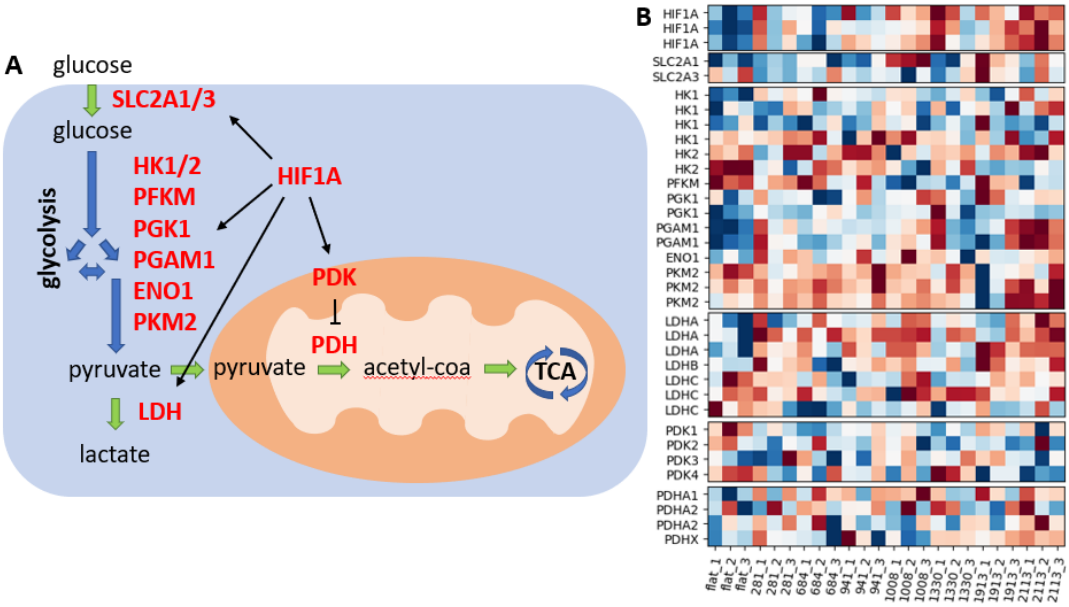

Figure 9: A. Influence of hypoxia-inducible factor 1 alpha (HIF1A) on genes associated with the glycolysis and the transfer to the TCA cycle. During the glycolysis glucose is converted to pyruvate. In anaerobic conditions, pyruvate is shifted to lactate by lactate dehydrogenase (LDH). In oxidative conditions, there is a shift from pyruvate to Acetyl-coenzyme A (acetyl-coA) which enters the TCA cycle. The scheme is adapted from Jeoung [66] and Semenza [67]. B. Gene expression of cells grown on flat surfaces and various topographically enhanced surfaces. The expression of each gene (rows) in each sample (columns) is represented with a color. Blue represents the highest expression value and red the lowest.
Remaining gene abbreviations: SLC2A1/3, solute carrier family 2 member 1/3; HK, hexokinase; PFKM, phosphofructokinase; PGK1, phosphoglycerate kinase 1; PGAM1, phosphate glycerate mutase 1; ENO1, enolase 1; PKM2, pyruvate kinase M2; PDK, pyruvate dehydrogenase kinase; PDH, pyruvate dehyrdogenase.

Especially the expression of HIF1A is higher in the cells grown on flat surfaces in comparison with cells grown on topographies. This can also be seen, but to a lesser extent, for the genes PGK, PGAM1, ENO1, and LDHA. PDK1 and PDK4 which inhibit the pyruvate dehydrogenase, are slightly underexpressed in cells cultured on a flat surface. For the genes, SLC2A, HK, PFKM, PKM, LDHB/C and PDH no clear difference in expression is seen between cells on a flat and on a topographically enhanced surface.

5.4 Polyamine biosynthesis

Another interesting pathway is the polyamine biosynthesis, which produces putrescine, spermidine, and spermine. A part of the polyamine biosynthesis pathway is depicted in Figure 10. The gene expression of genes involved in this pathway are upregulated in cells cultured on a flat surface. A downregulation can be seen in the cells cultured in topography 281. In the first place, this was not visible in the in silico models because of a shortage of S-adenosylmethionine (amet). But after adding enough amet to the system via a sink reaction, the flat topography-specific models predict a higher maximal production of spermine. The maximal production of spermine is the highest in flat_3, which is in line with the gene expression data. The spermine production in the models representing topography 281 are not significantly lower.

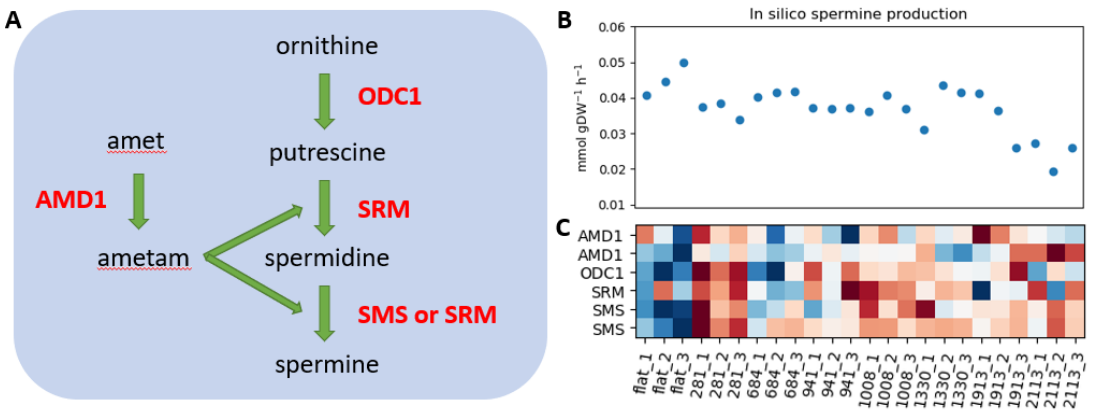

Figure 10: A. Part of the polyamine biosynthesis pathway. This network shows the production of the polyamines putrescine, spermidine, and spermine from ornithine and S-adenosylmethionine. Involved genes are depicted in red. B. Maximal in silico production of spermine for the topography-specific models. C. Expression of genes involved in the biosynthesis of the polyamines. Blue represents the highest expressionvalue and red the lowest.
abbreviations: amet, S-adenosylmethionine; ametam, decarboxylated S-adenosylmethionine; AMD1, adenosylmethionine decarboxylase 1; ODC1, ornithine decarboxylase 1; SRM, spermidine synthase; SMS, spermine synthase

6 Discussion

In this study, two methods are combined to construct situation-specific models. The first method, CORDA, is a switch-based method and ensures the exclusion of reactions that are not typically active in hMSCs. Its application to Recon3D results in the hMSC-base model. The second method is a valve-based method in which gene expression of cells on topographies is mapped to the hMSC-base model to constrain reactions. This method results in 24 different topography-specific models. Optimization of the ATP production in those models leads to a predicted glycolytic ATP production between 95.5% and 99.3%. These numbers are in agreement with data of hMSCs in an undifferentiated state, specifying that 97.5% of ATP is produced by the glycolysis [30]. So, the topography-specific models are capable to describe ATP production distributions for hMSCs. Also, the ratio between lactate production and glucose consumption is in correspondence with the ratio found during the culturing of proliferating hMSCs [68, 69]. Optimization of the ATP production in the hMSC-base model (the model build with the CORDA algorithm) leads to a prediction of 87% mitochondrial and 13% glycolytic ATP production. This is significantly different from the predictions of the 24 topography-specific models and not in agreement with literature. This shows that the last step of the model construction, in which reactions are constraint according to their associated gene expression, is necessary to construct a model with realistic predictions for the ATP production. There is a downside to using the switch-base method (CORDA) with just the expression of the cells cultured on a flat surface. As it could have lead to the exclusion of reactions that only take place on topographies. In a subsequent study, it would be valuable to also include gene expression data of the cells cultured on topographies to construct the hMSC-base model. All in all, the combination of a switch-based and a valve-based method ensures the construction of a smaller model (fewer reactions) with accurate predictions for energy production. As far as we know, a combination of those two methods is not used before to build situation-specific GSMMs.

GSMMs have already been used to analyze diseases and tissues, the effect of different culturing media, medicine targeting, and microbial interactions. This is the first time genome-scale metabolic modeling is used to analyze the effect of culturing cells on topographies. A GSMM has been made of hMSCs on a flat surface before, the iMSC1255 [26]. Reactions were selected from Recon1 by the mCADRE algorithm (a switch based method) and later literature and proteome data were used to refine the model. So all reactions in this model are unconstrained, the boundaries have their default value. The biggest difference between the generic Recon1 and Recon3D is the extent of the model [43, 70]. Recon3D contains more than three times as many reactions and more than two times as many metabolites. The lipid metabolism is considerably expanded causing the incorporation of LA to be essential. Furthermore, the inner mitochondrial compartment is added in Recon3D, which makes the representation of the OXPHOS more realistic. iMSC1255 predicted a maximum growth rate of 0.034 h^{-1} which is higher than the predictions from our topography-specific models and higher than growth rates mentioned in literature. Optimizing this model for ATP production resulted in 64.2% ATP produced by the mitochondrion and 31.9 % from glycolysis. This is not a realistic energy production distribution for undifferentiated hMSCs.

Next to the already performed refinement steps, it would be recommended to perform the essentiality test during the model construction phase. This way it is possible to add or remove reactions from the model to handle essential and non-essential nutrients correctly.

The topography-specific models were used to predict the growth rate of the hMSCs. These growth rates were compared to experimentally measured growth rates. The spread of in silico growth rates is

much smaller, they lie between 0.0082 h^{-1} and 0.0107 h^{-1}, whilst experimental growth rates range from 0.0010 to 0.0187 h^{-1}. In literature, there are numerous hMSC population growth rates reported. These growth rates are in the range 0.0014-0.027 h^{-1} [68, 69, 71, 72, 73, 74]. Both the topography-specific model predictions and experimentally measured growth rates (except for hMSCs grown on topography 1330 sample 3) fall within that range. So, it is clear to state that the topography-specific models are able to predict realistic growth rate values.

In the experimental setting, hMSCs on a flat surface showed a higher proliferation rate than the cells cultured on a topographically enhance surface (Figure 7.A). This phenomenon did not come forward in the in silico predictions. Several aspects could explain this poor agreement between in silico predictions and experimental measurements. First of all, it must be taken into account that the expression of only 1714 genes is used during the construction of the topography-specific model. However, from the PCA in Appendix A can be seen that those genes are also able to separate the cells cultured on a flat surface from those cultured on topographies. This means that these metabolic genes should contain enough information to create different results in the flat models compared to the topography models. Besides those 1714 metabolic genes, other genes could influence proliferation. Growth factors, for example, are not included in the model. Also, the composition and stoichiometric coefficients of the metabolites enclosed in the biomass objective function influence the FtB. In the topography-specific models, the FtB is most sensitive to the ATP yield. This yield has been taken from a generic biomass objective function, but it is not verified whether this energy requirement is also applicable to hMSCs. Altering the biomass reaction more towards this specific type of cell might help predict a more precise in silico growth rate.

During this study, multiple targets came to light to stimulate the growth of hMSCs in vitro. Increasing the proliferation rate for hMSCs could be accomplished by manipulating the expression of the limiting genes mentioned in Table 4. Increasing the gene expression of limiting genes by 10% leads to an average increase of the in silico growth rate of 12.7%. Nonetheless, it must be taken into account that increasing the expression of genes related to the mitochondrion can cause the cells to differentiate. As research has shown that mitochondrial respiration is necessary for the differentiation of stem cells [33, 75, 76].

In addition, medium supplementation could help increase the growth rate. Increasing the concentration of cholesterol turned out to elevate the in silico growth rate. Proliferation also benefits from an increased concentration of LA in some topography-specific models. An alternative for LA is the addition of phosphatidic acid to the medium. This way, this phospholipid can directly be consumed from the ECS, which is more efficient than its biosynthesis from LA. LA is desaturated and elongated to arachidonic acid which is incorporated in glycerophospholipids of the cells membrane [57]. In two studies, however, the medium supplementation with LA turned out to decrease hMSC proliferation [77, 78]. This could be due to the fact that besides being a building block for membrane lipids, LA also has a downstream effect on multiple signaling pathways. Moreover, gene expression and secretion of angiogenic mediators VEGF and IL-8 are elevated through exposure to LA [77]. On the other hand, is ESC proliferation stimulated in the presence of LA via Ca^{2+}/PKC, PI3K/Akt, and MAPKs [79]. Cholesterol acts, besides his role as a phospholipid membrane, also as a retardant in the senescence of hMSCs by regulating ROS/p53/$p21^{Cip1/Waf1}$ signaling pathways and the process of autophagy [80]. In GSMM, LA and cholesterol are just used as a building block for membrane lipids, the effects of signaling pathways are very limited. This stresses the importance of conducting in vitro experiments to further investigate the role of LA and cholesterol supplementation in hMSC culturing.

The expression of HIF1A in hMSCs has been examined in different studies [18, 34, 35, 36]. We found a

difference in gene expression of HIF1A between hMSCs cultured on a flat surface and on topographically enhanced surfaces. This upregulation, however, did not lead to an increased glycolytic ATP production in the models representing the flat surface as was expected. The differential expression of HIF1A due to surface topographies is a novel finding. In literature, no connection has been found between culturing cells on surface topographies and the expression of HIF1A. Normally HIF1A is activated by a lack of oxygen (hypoxia). During the experiments, the oxygen supply was the same in every situation. The ability of the cells to take up oxygen could be the reason for the difference in HIF1A expression. On the flat surface, cells are proliferating faster. This leads to a higher cell density which could have led to a decreased oxygen consumption rate [81].

In gene expression of genes involved in polyamine biosynthesis, a difference between cells cultured on flat and topographically enhanced surfaces is also seen. Research with embryonic stem cells (ESCs) showed that the genes AMD1 and ODC1 may promote self-renewal in ESCs [82, 83]. Other studies suggest a role for the polyamine pathway and ODC1 in the osteogenic differentiation of hMSCs [84, 85]. From the gene expression data and the in silico production of spermine, culturing on topographically enhanced surfaces seems to have a negative influence on the polyamine biosynthesis pathways. How this influences the differentiation or self-renewal of hMSCs should be further investigated.

7 Conclusion and future work

7.1 Conclusion

In this study, models are constructed to represent the metabolic behavior of hMSCs cultured on flat and topographically enhanced surfaces. The combination of a switch-based and a valve-based method is proven to be effective for the construction of models with accurate predictions for the energy production and realistic growth rates.
We hypothesized that cells grown on topographies are better able to maintain their stem cell-like characteristics and so, reduce their metabolism and proliferation. Our acquired topography-specific models do not provide a decisive answer. While they represent metabolic behavior of hMSCs culturing very well, they are not yet able to distinguish different metabolic phenotypes induced by topographies. In order to see metabolic adjustment to the topographies, additional steps are required, for example, by adding the effects of signaling pathways to the models. The gene expression of genes involved in metabolic regulation (HIF1A and polyamine biosynthesis pathway) was promising, as those genes were upregulated in hMSCs cultured on a flat surface compared to hMSCs cultured on topographically enhanced surfaces.

7.2 Future work

To generate a more accurate model, the integration of proteomics or metabolomics data would be useful. Measuring in vitro fluxes, also called Metabolic Flux Analysis (MFA), is hard to achieve as it requires mass spectrometry approaches and isotope labeling techniques [86]. An alternative could be time-coarse measurements of metabolites in the medium during culturing. Ranges of metabolite consumption and excretion could then be used as additional constraints in the model. Besides using this data to further constrain the model, it could also be used as a tool to validate the model: the production of for example, lactate or amino acid can be a reflection of the metabolism.
The current method of mapping gene expression to GSMM is based on the expression of enzymes catalyzing reactions. GSMMs might broaden their usefulness if the expression of signaling pathways could be included. This way, more gene expression information could be used and the models would be more representative for the specific situation. For example, genes of inhibitors could be implemented in the model via an inverse method, the higher the expression, the lower the flux that reaction is able to carry.
Another improvement would be to take into account that enzymes can only be used once to catalyze a reaction. So if an enzyme catalyzes multiple reactions, the gene expression should be divided over those reactions while calculating the reaction confidence.
The influence of HIF1A turned out to be an interesting factor. It is not sure however, how this gene is affected by the culturing on different surfaces. What is the connection between the downregulated expression of HIF1A and the mechanical stimulation of topographically enhanced surfaces? To answer this question, more research is needed.
The use of this model can be expanded to see how the metabolism reacts to changes in the metabolic gene expression. Also, the adjustment of the metabolism to changes in the medium can be further examined. This way you can first test the effect in silico before experimenting in vitro. It is also possible to optimize the models for the production of every metabolite in the model to see if a metabolites' production capacity is higher in cells on a distinct topography. Furthermore, it would be interesting to build tissue-specific models of differentiated hMSCs towards different lineages and see how the

metabolism differs upon differentiation. These models could be constructed with genomics datasets of differentiated cells which are publicly available.
In this study, the emphasis was on the difference between the cells cultured on flat and on topographically enhanced surfaces. But the different topographies also impose different phenotypes on the cells. The expression of PTGS2 for example, is only upregulated in topographies 281 and 1330. Normally, its expression is only elevated during inflammation. So it would be interesting to investigate the differences between the topographies further.

References

[1] Sigmarsdóttir, S. McGarrity, Ó. Rolfsson, J. T. Yurkovich, and Ó. E. Sigurjónsson, "Current status and future prospects of genome-scale metabolic modeling to optimize the use of mesenchymal stem cells in regenerative medicine," *Frontiers in Bioengineering and Biotechnology*, vol. 8, p. 239, Mar 2020.

[2] A. J. Rosenbaum, D. A. Grande, and J. S. Dines, "The use of mesenchymal stem cells in tissue engineering: A global assessment," *Organogenesis*, vol. 4, pp. 23–27, Jan 2008.

[3] A. I. Caplan, "Mesenchymal stem cells and gene therapy," *Clin. Orthop. Relat. Res.*, pp. 67–70, Oct 2000.

[4] A. I. Caplan, "Adult mesenchymal stem cells for tissue engineering versus regenerative medicine," *J. Cell. Physiol.*, vol. 213, pp. 341–347, Nov 2007.

[5] M. Mohyeddin Bonab, S. Yazdanbakhsh, J. Lotfi, K. Alimoghaddom, F. Talebian, F. Hooshmand, A. Ghavamzadeh, and B. Nikbin, "Does mesenchymal stem cell therapy help multiple sclerosis patients? report of a pilot study," *Iranian Journal of Immunology*, vol. 4, no. 1, pp. 50–57, 2007.

[6] R. E. B. Fitzsimmons, M. S. Mazurek, A. Soos, and C. A. Simmons, "Mesenchymal stromal/stem cells in regenerative medicine and tissue engineering," *Stem cells international*, vol. 2018, pp. 8031718–8031718, Aug 2018.

[7] M. Duijvestein, A. C. W. Vos, H. Roelofs, M. E. Wildenberg, B. B. Wendrich, H. W. Verspaget, E. M. C. Kooy-Winkelaar, F. Koning, J. J. Zwaginga, H. H. Fidder, A. P. Verhaar, W. E. Fibbe, G. R. van den Brink, and D. W. Hommes, "Autologous bone marrow-derived mesenchymal stromal cell treatment for refractory luminal crohn's disease: results of a phase i study," *Gut*, vol. 59, no. 12, pp. 1662–1669, 2010.

[8] R. Secunda, R. Vennila, A. M. Mohanashankar, M. Rajasundari, S. Jeswanth, and R. Surendran, "Isolation, expansion and characterisation of mesenchymal stem cells from human bone marrow, adipose tissue, umbilical cord blood and matrix: a comparative study," *Cytotechnology*, vol. 67, pp. 793–807, Oct 2015.

[9] J. Liu, Y. Ding, Z. Liu, and X. Liang, "Senescence in mesenchymal stem cells: Functional alterations, molecular mechanisms, and rejuvenation strategies," *Frontiers in Cell and Developmental Biology*, vol. 8, p. 258, 2020.

[10] T. Jiang, G. Xu, Q. Wang, L. Yang, L. Zheng, J. Zhao, and X. Zhang, "In vitro expansion impaired the stemness of early passage mesenchymal stem cells for treatment of cartilage defects," *Cell Death & Disease*, vol. 8, pp. e2851–e2851, Jun 2017.

[11] C. Nombela-Arrieta, J. Ritz, and L. E. Silberstein, "The elusive nature and function of mesenchymal stem cells," *Nature reviews. Molecular cell biology*, vol. 12, pp. 126–131, Feb 2011.

[12] T. H. Cheung and T. A. Rando, "Molecular regulation of stem cell quiescence," *Nature Reviews Molecular Cell Biology*, vol. 14, pp. 329–340, Jun 2013.

[13] Y. Zhou, T.-L. Tsai, and W.-J. Li, "Strategies to retain properties of bone marrow-derived mesenchymal stem cells ex vivo," *Annals of the New York Academy of Sciences*, vol. 1409, pp. 3–17, Dec 2017.

[14] N. Chosa and A. Ishisaki, "Two novel mechanisms for maintenance of stemness in mesenchymal stem cells: Scrg1/bst1 axis and cell-cell adhesion through n-cadherin," *The Japanese Dental Science Review*, vol. 54, p. 37–44, Feb 2018.

[15] H. Obradovic, J. Krstic, D. Trivanovic, S. Mojsilovic, I. Okic, T. Kukolj, V. Ilic, A. Jaukovic, M. Terzic, and D. Bugarski, "Improving stemness and functional features of mesenchymal stem cells from wharton's jelly of a human umbilical cord by mimicking the native, low oxygen stem cell niche," *Placenta*, vol. 82, p. 25–34, 2019.

[16] N. Beijer, Z. Nauryzgaliyeva, E. Arteaga, L. Pieuchot, K. Anselme, J. van de Peppel, A. Vasilevich, N. Groen, N. Roumans, D. Hebels, and J. Boer, "Dynamic adaptation of mesenchymal stem cell physiology upon exposure to surface micropatterns," *Scientific Reports*, vol. 9, 12 2019.

[17] A. S. Vasilevich, S. Vermeulen, M. Kamphuis, N. Roumans, S. Eroumé, D. G. Hebels, J. van de Peppel, R. Reihs, N. R. Beijer, A. Carlier, A. E. Carpenter, S. Singh, and J. de Boer, "On the correlation between material-induced cell shape and phenotypical response of human mesenchymal stem cells," *bioRxiv*, 2020.

[18] L. C. Shum, N. S. White, B. N. Mills, K. L. d. M. Bentley, and R. A. Eliseev, "Energy metabolism in mesenchymal stem cells during osteogenic differentiation," *Stem cells and development*, vol. 25, pp. 114–122, Jan 2016.

[19] G. Pattappa, H. K. Heywood, J. D. de Bruijn, and D. A. Lee, "The metabolism of human mesenchymal stem cells during proliferation and differentiation," *Journal of Cellular Physiology*, vol. 226, p. 2562–2570, Oct 2011.

[20] C. Gu, G. B. Kim, W. J. Kim, H. U. Kim, and S. Y. Lee, "Current status and applications of genome-scale metabolic models," *Genome Biology*, vol. 20, no. 1, p. 121, 2019.

[21] H. Zur, E. Ruppin, and T. Shlomi, "iMAT: an integrative metabolic analysis tool," *Bioinformatics*, vol. 26, pp. 3140–3142, 11 2010.

[22] R. Agren, S. Bordel, A. Mardinoglu, N. Pornputtapong, I. Nookaew, and J. Nielsen, "Reconstruction of genome-scale active metabolic networks for 69 human cell types and 16 cancer types using init," *PLOS Computational Biology*, vol. 8, pp. 1–9, 05 2012.

[23] Y. Wang, J. A. Eddy, and N. D. Price, "Reconstruction of genome-scale metabolic models for 126 human tissues using mcadre," *BMC Systems Biology*, vol. 6, p. 153, Dec 2012.

[24] S. A. Becker and B. O. Palsson, "Context-specific metabolic networks are consistent with experiments," *PLOS Computational Biology*, vol. 4, pp. 1–10, 05 2008.

[25] A. Schultz and A. A. Qutub, "Reconstruction of tissue-specific metabolic networks using corda," *PLOS Computational Biology*, vol. 12, pp. 1–33, 03 2016.

[26] H. Fouladiha, S.-A. Marashi, and M. A. Shokrgozar, "Reconstruction and validation of a constraint-based metabolic network model for bone marrow-derived mesenchymal stem cells," *Cell proliferation*, vol. 48, pp. 475–485, Aug 2015.

[27] H. Fouladiha, S.-A. Marashi, M. A. Shokrgozar, M. Farokhi, and A. Atashi, "Applications of a metabolic network model of mesenchymal stem cells for controlling cell proliferation and differentiation," *Cytotechnology*, vol. 70, pp. 331–338, Feb 2018.

[28] K. Mohammad, P. Dakik, Y. Medkour, D. Mitrofanova, and V. I. Titorenko, "Quiescence entry, maintenance, and exit in adult stem cells," *International Journal of Molecular Sciences*, vol. 20, p. 2158, May 2019.

[29] I. Ullah, R. B. Subbarao, and G. J. Rho, "Human mesenchymal stem cells - current trends and future prospective," *Bioscience Reports*, vol. 35, Apr. 2015.

[30] N. Fillmore, A. Huqi, J. S. Jaswal, J. Mori, R. Paulin, A. Haromy, A. Onay-Besikci, L. Ionescu, B. Thébaud, E. Michelakis, and G. D. Lopaschuk, "Effect of fatty acids on human bone marrow mesenchymal stem cell energy metabolism and survival," *PLOS ONE*, vol. 10, p. e0120257, Mar. 2015.

[31] N. Shyh-Chang and H.-H. Ng, "The metabolic programming of stem cells," *Genes & Development*, vol. 31, pp. 336–346, Feb. 2017.

[32] A. V. Meleshina, V. V. Dudenkova, M. V. Shirmanova, V. I. Shcheslavskiy, W. Becker, A. S. Bystrova, E. I. Cherkasova, and E. V. Zagaynova, "Probing metabolic states of differentiating stem cells using two-photon FLIM," *Scientific Reports*, vol. 6, Feb. 2016.

[33] Y. Zhang, G. Marsboom, P. T. Toth, and J. Rehman, "Mitochondrial respiration regulates adipogenic differentiation of human mesenchymal stem cells," *PloS one*, vol. 8, pp. e77077–e77077, Oct 2013.

[34] A. J. Majmundar, W. J. Wong, and M. C. Simon, "Hypoxia-inducible factors and the response to hypoxic stress," *Molecular Cell*, vol. 40, pp. 294–309, Oct. 2010.

[35] I.-H. Park, K.-H. Kim, H.-K. Choi, J.-S. Shim, S.-Y. Whang, S. J. Hahn, O.-J. Kwon, and I.-H. Oh, "Constitutive stabilization of hypoxia-inducible factor alpha selectively promotes the self-renewal of mesenchymal progenitors and maintains mesenchymal stromal cells in an undifferentiated state," *Experimental & Molecular Medicine*, vol. 45, pp. e44–e44, Sept. 2013.

[36] M. Kanichai, D. Ferguson, P. J. Prendergast, and V. A. Campbell, "Hypoxia promotes chondrogenesis in rat mesenchymal stem cells: A role for AKT and hypoxia-inducible factor (HIF)-1," *Journal of Cellular Physiology*, vol. 216, pp. 708–715, Sept. 2008.

[37] J. C. Estrada, C. Albo, A. Benguría, A. Dopazo, P. López-Romero, L. Carrera-Quintanar, E. Roche, E. P. Clemente, J. A. Enríquez, A. Bernad, and E. Samper, "Culture of human mesenchymal stem cells at low oxygen tension improves growth and genetic stability by activating glycolysis," *Cell Death & Differentiation*, vol. 19, pp. 743–755, Dec. 2011.

[38] R. J. Burgess, M. Agathocleous, and S. J. Morrison, "Metabolic regulation of stem cell function," *Journal of Internal Medicine*, vol. 276, pp. 12–24, June 2014.

[39] N. S. Chandel, "Mitochondria as signaling organelles," *BMC Biology*, vol. 12, May 2014.

[40] Q. Li, Z. Gao, Y. Chen, and M.-X. Guan, "The role of mitochondria in osteogenic, adipogenic and chondrogenic differentiation of mesenchymal stem cells," *Protein & cell*, vol. 8, pp. 439–445, Jun 2017.

[41] M. J. Biggs, R. G. Richards, N. Gadegaard, C. D. Wilkinson, R. O. Oreffo, and M. J. Dalby, "The use of nanoscale topography to modulate the dynamics of adhesion formation in primary osteoblasts and ERK/MAPK signalling in STRO-1 enriched skeletal stem cells," *Biomaterials*, vol. 30, pp. 5094–5103, Oct. 2009.

[42] G. Abagnale, M. Steger, V. H. Nguyen, N. Hersch, A. Sechi, S. Joussen, B. Denecke, R. Merkel, B. Hoffmann, A. Dreser, U. Schnakenberg, A. Gillner, and W. Wagner, "Surface topography enhances differentiation of mesenchymal stem cells towards osteogenic and adipogenic lineages," *Biomaterials*, vol. 61, pp. 316–326, Aug. 2015.

[43] E. Brunk, S. Sahoo, I. Thiele, and B. O. Palsson, "Recon3d enables a three-dimensional view of gene variation in human metabolism," *Nature Biotechnology*, vol. 36, pp. 272 EP –, Feb 2018.

[44] I. Thiele, N. Swainston, R. M. T. Fleming, and B. Ø. Palsson, "A community-driven global reconstruction of human metabolism," *Nature biotechnology*, vol. 31, pp. 419–425, May 2013.

[45] A. M. Feist and B. O. Palsson, "The biomass objective function," *Current opinion in microbiology*, vol. 13, pp. 344–349, Jun 2010.

[46] J. D. Orth, I. Thiele, and B. Ø. Palsson, "What is flux balance analysis?," *Nature biotechnology*, vol. 28, pp. 245–248, Mar 2010.

[47] A. Mentink, M. Hulsman, N. Groen, R. Licht, K. J. Dechering, J. van der Stok, H. A. Alves, W. J. Dhert, E. P. van Someren, M. J. T. Reinders, C. A. van Blitterswijk, and J. de Boer, "Predicting the therapeutic efficacy of msc in bone tissue engineering using the molecular marker cadm1," *Biomaterials*, vol. 34, p. 4592—4601, June 2013.

[48] H. V. Unadkat, M. Hulsman, K. Cornelissen, B. J. Papenburg, R. K. Truckenmüller, A. E. Carpenter, M. Wessling, G. F. Post, M. Uetz, M. J. T. Reinders, D. Stamatialis, C. A. van Blitterswijk, and J. de Boer, "An algorithm-based topographical biomaterials library to instruct cell fate," *Proceedings of the National Academy of Sciences*, vol. 108, no. 40, pp. 16565–16570, 2011.

[49] A. Ebrahim, J. A. Lerman, B. O. Palsson, and D. R. Hyduke, "Cobrapy: Constraints-based reconstruction and analysis for python," *BMC Systems Biology*, vol. 7, p. 74, Aug 2013.

[50] A. Noronha, A. D. Daníelsdóttir, P. Gawron, F. Jóhannsson, S. Jónsdóttir, S. Jarlsson, J. P. Gunnarsson, S. Brynjólfsson, R. Schneider, I. Thiele, and R. M. T. Fleming, "ReconMap: an interactive visualization of human metabolism," *Bioinformatics*, p. btw667, Dec. 2016.

[51] M. Hucka, A. Finney, H. M. Sauro, H. Bolouri, J. C. Doyle, H. Kitano, A. P. Arkin, B. J. Bornstein, D. Bray, A. Cornish-Bowden, A. A. Cuellar, S. Dronov, E. D. Gilles, M. Ginkel, V. Gor, I. I. Goryanin, W. J. Hedley, T. C. Hodgman, J.-H. Hofmeyr, P. J. Hunter, N. S. Juty, J. L. Kasberger, A. Kremling, U. Kummer, N. Le Novère, L. M. Loew, D. Lucio, P. Mendes, E. Minch, E. D. Mjolsness, Y. Nakayama, M. R. Nelson, P. F. Nielsen, T. Sakurada, J. C. Schaff, B. E. Shapiro, T. S. Shimizu, H. D. Spence, J. Stelling, K. Takahashi, M. Tomita, J. Wagner, J. Wang, and the rest of the SBML Forum:, "The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models," *Bioinformatics*, vol. 19, pp. 524–531, 03 2003.

[52] P. A. Jensen, K. A. Lutz, and J. A. Papin, "Tiger: Toolbox for integrating genome-scale metabolic models, expression data, and transcriptional regulatory networks," *BMC Systems Biology*, vol. 5, p. 147, Sep 2011.

[53] L. Heirendt, S. Arreckx, T. Pfau, S. N. Mendoza, A. Richelle, A. Heinken, H. S. Haraldsdóttir, J. Wachowiak, S. M. Keating, V. Vlasov, and et al., "Creation and analysis of biochemical constraint-based models using the cobra toolbox v.3.0," *Nature Protocols*, vol. 14, p. 639–702, Mar 2019.

[54] G. Gstraunthaler and T. Lindl, *Zell- und Gewebekultur*. Springer Berlin Heidelberg, 2013.

[55] M. Yousefi, S.-A. Marashi, A. Sharifi-Zarchi, and S. Taleahmad, "The metabolic network model of primed/naive human embryonic stem cells underlines the importance of oxidation-reduction potential and tryptophan metabolism in primed pluripotency," *Cell & bioscience*, vol. 9, pp. 71–71, Aug 2019.

[56] M. Masid, M. Ataman, and V. Hatzimanikatis, "Analysis of human metabolism by reducing the complexity of the genome-scale models using redhuman," *Nature Communications*, vol. 11, p. 2821, Jun 2020.

[57] M.- Ouellette, J.-C. Bérubé, J.-M. Bourget, M. Vallée, Y. Bossé, and J. Fradette, "Linoleic acid supplementation of cell culture media influences the phospholipid and lipid profiles of human reconstructed adipose tissue," *PLOS ONE*, vol. 14, pp. 1–22, oct 2019.

[58] I. Thiele and B. Ø. Palsson, "A protocol for generating a high-quality genome-scale metabolic reconstruction," *Nature protocols*, vol. 5, pp. 93–121, Jan 2010.

[59] H. Yuan, C. Y. M. Cheung, P. A. J. Hilbers, and N. A. W. van Riel, "Flux balance analysis of plant metabolism: The effect of biomass composition and model structure on model predictions," *Frontiers in Plant Science*, vol. 7, p. 537, Apr. 2016.

[60] K. R. Levental, M. A. Surma, A. D. Skinkle, J. H. Lorent, Y. Zhou, C. Klose, J. T. Chang, J. F. Hancock, and I. Levental, "ω-3 polyunsaturated fatty acids direct differentiation of the membrane phenotype in mesenchymal stem cells to potentiate osteogenesis," *Science advances*, vol. 3, Nov 2017.

[61] O. Folger, L. Jerby, C. Frezza, E. Gottlieb, E. Ruppin, and T. Shlomi, "Predicting selective drug targets in cancer through metabolic networks," *Molecular systems biology*, vol. 7, pp. 501–501, Jun 2011.

[62] S. Opdam, A. Richelle, B. Kellman, S. Li, D. C. Zielinski, and N. E. Lewis, "A systematic evaluation of methods for tailoring genome-scale metabolic models," *Cell systems*, vol. 4, pp. 318–329.e6, Mar 2017.

[63] A. Bordbar, N. E. Lewis, J. Schellenberger, B. Ø. Palsson, and N. Jamshidi, "Insight into human alveolar macrophage and m. tuberculosis interactions via metabolic reconstructions," *Molecular systems biology*, vol. 6, pp. 422–422, Oct 2010.

[64] S. Chandrasekaran and N. D. Price, "Probabilistic integrative modeling of genome-scale metabolic and regulatory networks inEscherichia coliandMycobacterium tuberculosis," *Proceedings of the National Academy of Sciences*, vol. 107, pp. 17845–17850, Sept. 2010.

[65] P. Droste, K. Nöh, and W. Wiechert, "Omix - a visualization tool for metabolic networks with highest usability and customizability in focus," *Chemie Ingenieur Technik*, vol. 85, pp. 849–862, Apr. 2013.

[66] N. H. Jeoung, "Pyruvate dehydrogenase kinases: Therapeutic targets for diabetes and cancers," *Diabetes & Metabolism Journal*, vol. 39, no. 3, p. 188, 2015.

[67] G. L. Semenza, "Hypoxia-inducible factors: mediators of cancer progression and targets for cancer therapy," *Trends in Pharmacological Sciences*, vol. 33, pp. 207–214, Apr. 2012.

[68] G. A. Higuera, D. Schop, T. W. Spitters, R. van Dijkhuizen-Radersma, M. Bracke, J. D. de Bruijn, D. Martens, M. Karperien, A. van Boxtel, and C. A. van Blitterswijk, "Patterns of amino acid metabolism by proliferating human mesenchymal stem cells," *Tissue Engineering Part A*, vol. 18, pp. 654–664, Mar. 2012.

[69] D. Schop, F. W. Janssen, L. D. van Rijn, H. Fernandes, R. M. Bloem, J. D. de Bruijn, and R. van Dijkhuizen-Radersma, "Growth, metabolism, and growth inhibitors of mesenchymal stem cells," *Tissue Engineering Part A*, vol. 15, pp. 1877–1886, Aug. 2009.

[70] N. C. Duarte, S. A. Becker, N. Jamshidi, I. Thiele, M. L. Mo, T. D. Vo, R. Srivas, and B. O. Palsson, "Global reconstruction of the human metabolic network based on genomic and bibliomic data," *Proceedings of the National Academy of Sciences*, vol. 104, pp. 1777–1782, Jan. 2007.

[71] F. dos Santos, P. Z. Andrade, J. S. Boura, M. M. Abecasis, C. L. da Silva, and J. M. Cabral, "Ex vivo expansion of human mesenchymal stem cells: A more effective cell proliferation kinetics and metabolism under hypoxia," *Journal of Cellular Physiology*, pp. 27–35, Dec. 2009.

[72] G. Pattappa, S. D. Thorpe, N. C. Jegard, H. K. Heywood, J. D. de Bruijn, and D. A. Lee, "Continuous and uninterrupted oxygen tension influences the colony formation and oxidative metabolism of human mesenchymal stem cells," *Tissue Engineering Part C: Methods*, vol. 19, pp. 68–79, Jan. 2013.

[73] F. Zhao, P. Pathi, W. Grayson, Q. Xing, B. R. Locke, and T. Ma, "Effects of oxygen transport on 3-d human mesenchymal stem cell metabolic activity in perfusion and static cultures: Experiments and mathematical model," *Biotechnology Progress*, vol. 21, pp. 1269–1280, Sept. 2008.

[74] W. L. Grayson, F. Zhao, R. Izadpanah, B. Bunnell, and T. Ma, "Effects of hypoxia on human mesenchymal stem cell expansion and plasticity in 3d constructs," *Journal of Cellular Physiology*, vol. 207, pp. 331–339, May 2006.

[75] Y. Tang, B. Luo, Z. Deng, B. Wang, F. Liu, J. Li, W. Shi, H. Xie, X. Hu, and J. Li, "Mitochondrial aerobic respiration is activated during hair follicle stem cell differentiation, and its dysfunction retards hair regeneration," *PeerJ*, vol. 4, p. e1821, May 2016.

[76] K. V. Tormos, E. Anso, R. B. Hamanaka, J. Eisenbart, J. Joseph, B. Kalyanaraman, and N. S. Chandel, "Mitochondrial complex III ROS regulate adipocyte differentiation," *Cell Metabolism*, vol. 14, pp. 537–544, Oct. 2011.

[77] A. N. Smith, L. A. Muffley, A. N. Bell, S. Numhom, and A. M. Hocking, "Unsaturated fatty acids induce mesenchymal stem cells to increase secretion of angiogenic mediators," *Journal of Cellular Physiology*, vol. 227, pp. 3225–3233, May 2012.

[78] F. Tigistu-Sahle, M. Lampinen, L. Kilpinen, M. Holopainen, P. Lehenkari, S. Laitinen, and R. Käkelä, "Metabolism and phospholipid assembly of polyunsaturated fatty acids in human bone marrow mesenchymal stromal cells," *Journal of Lipid Research*, vol. 58, pp. 92–110, Nov. 2016.

[79] M. H. Kim, M. O. Kim, Y. H. Kim, J. S. Kim, and H. J. Han, "Linoleic acid induces mouse embryonic stem cell proliferation via ca^{2+}/PKC, PI3K/Akt, and MAPKs," *Cellular Physiology and Biochemistry*, vol. 23, no. 1-3, pp. 053–064, 2009.

[80] M. Zhang, Y. Du, R. Lu, Y. Shu, W. Zhao, Z. Li, Y. Zhang, R. Liu, T. Yang, S. Luo, M. Gao, Y. Zhang, G. Zhang, J. Liu, and Y. Lu, "Cholesterol retards senescence in bone marrow mesenchymal stem cells by modulating autophagy and ROS/p53/p21cip1/waf1pathway," *Oxidative Medicine and Cellular Longevity*, vol. 2016, pp. 1–10, 2016.

[81] C. Magliaro, G. Mattei, F. Iacoangeli, A. Corti, V. Piemonte, and A. Ahluwalia, "Oxygen consumption characteristics in 3d constructs depend on cell density," *Frontiers in Bioengineering and Biotechnology*, vol. 7, Oct. 2019.

[82] D. Zhang, T. Zhao, H. S. Ang, P. Chong, R. Saiki, K. Igarashi, H. Yang, and L. A. Vardy, "AMD1 is essential for ESC self-renewal and is translationally down-regulated on differentiation to neural precursor cells," *Genes & Development*, vol. 26, pp. 461–473, Mar. 2012.

[83] T. Zhao, K. J. Goh, H. H. Ng, and L. A. Vardy, "A role for polyamine regulators in ESC self-renewal," *Cell Cycle*, vol. 11, pp. 4517–4523, Dec. 2012.

[84] R. M. Borzì, S. Guidotti, M. Minguzzi, A. Facchini, D. Platano, G. Trisolino, G. Filardo, S. Cetrullo, S. D'Adamo, C. Stefanelli, A. Facchini, and F. Flamigni, "Polyamine delivery as a tool to modulate stem cell differentiation in skeletal tissue engineering," *Amino Acids*, vol. 46, pp. 717–728, Nov. 2013.

[85] Y.-H. Tsai, K.-L. Lin, Y.-P. Huang, Y.-C. Hsu, C.-H. Chen, Y. Chen, M.-H. Sie, G.-J. Wang, and M.-J. Lee, "Suppression of ornithine decarboxylase promotes osteogenic differentiation of human bone marrow-derived mesenchymal stem cells," *FEBS Letters*, vol. 589, pp. 2058–2065, June 2015.

[86] M. R. Antoniewicz, "Methods and advances in metabolic flux analysis: a mini-review," *Journal of Industrial Microbiology & Biotechnology*, vol. 42, pp. 317–325, Jan. 2015.

[87] T. F. S. US, "32571 - mem alpha, glutamax(tm), nucleosides."

[88] M. K. Gregory, H. W. King, P. A. Bain, R. A. Gibson, D. R. Tocher, and K. A. Schuller, "Development of a fish cell culture model to investigate the impact of fish oil replacement on lipid peroxidation," *Lipids*, vol. 46, pp. 753–764, Aug 2011.

Appendices

Appendix A Principal component analysis

Figure 11 shows the PCA of the gene expression data. In the first plot, a clear distinction is visible between the cells grown on flat surfaces (red circles) and on topographically enhanced surfaces. This separation is mainly due to the first component. The other topographies form two main clusters. The first cluster contains topographies 281, 1008, 1913 and 2113 the second consists of 684, 941 and 1330.

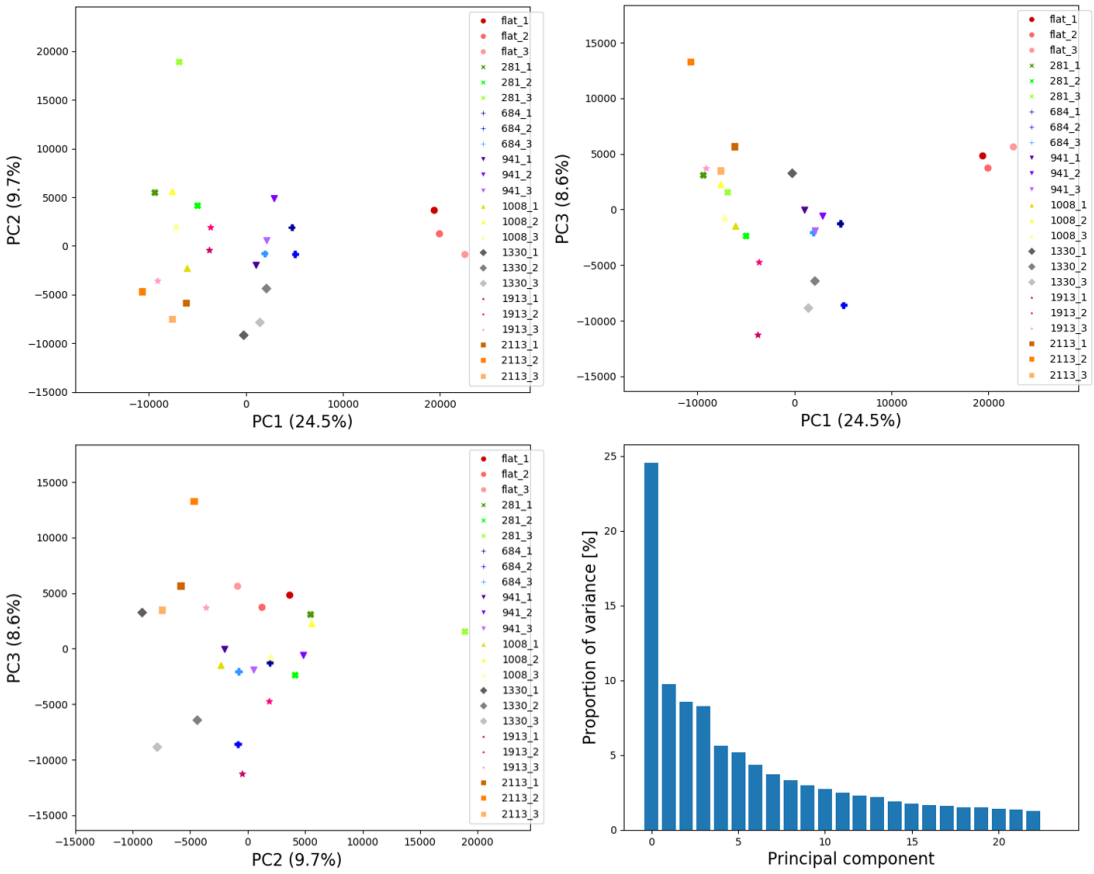

Figure 11: Principal component analysis of gene expression data of cells cultured on different topographies. The data contains the expression of 47273 genes. Different topographies are indicated both by color and shape. The bottom right graph shows the proportion of explained variance for every principal component.

In Figure 12 a PCA is performed with only metabolic genes. These are the genes that are present in the hMSC-base model. From these PCA plots we can still conclude that some genes are differentially expressed in the cells cultured on a flat surface compared to the topographically enhanced surface. Those genes are captured in PC1.

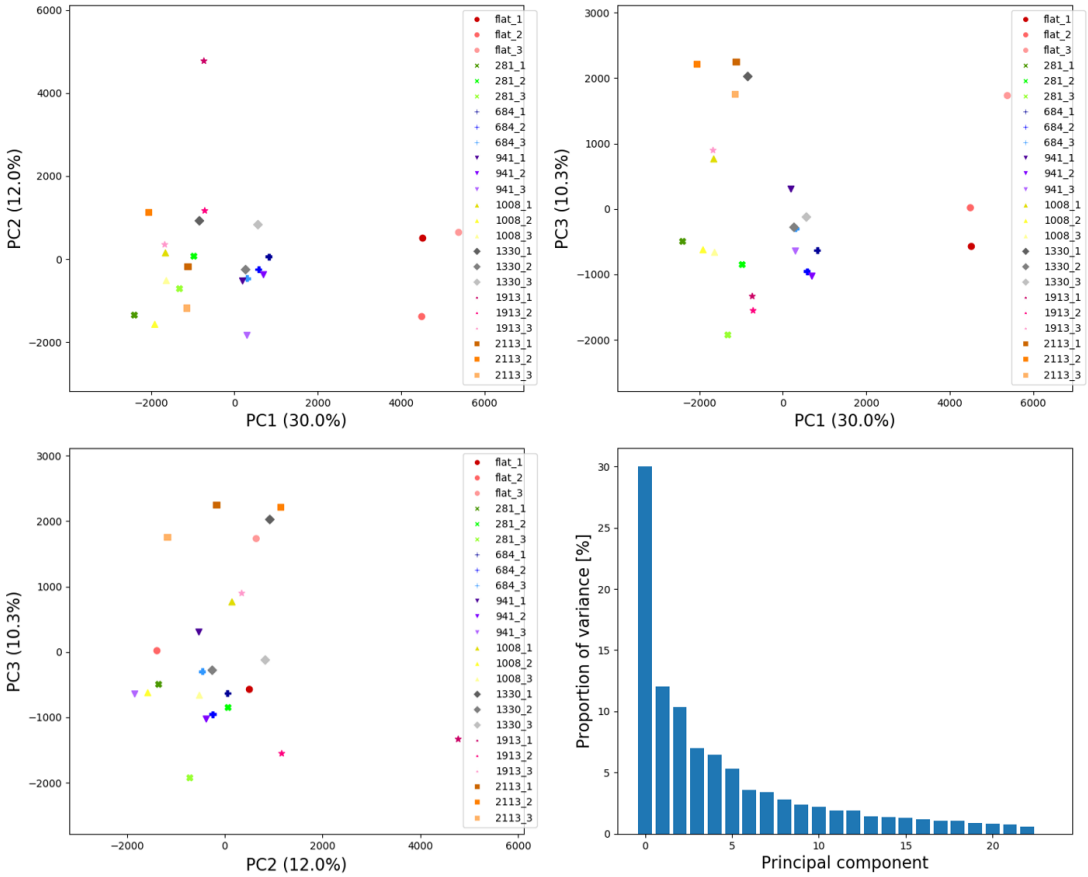

Figure 12: Principal component analysis of metabolic genes of cells cultured on different topographies. Metabolic genes are the 1714 genes that are comprised in the hMSC-base GSMM. Different topographies are indicated both by color and shape. The bottom right graph shows the proportion of explained variance for every principal component.

Appendix B Important metabolic reactions

Table 5: Important metabolic reactions that are given the status of HC reactions regardless of their gene expression.

HEX1	FBP	ALATA_Lm	ORNt4m	RPI	ADK1
PGI	PFK	THD1m	ORNt3m	TKT1	ADK3
r0191	PGK	NADH2_u10mi	OCBTm	TALA	NDP1
FBA	r0354	CYOR_u10mi	CBPSam	TKT2	IMPD
TPI	r0355	FADH2ETC	CPS_m	PGMT	GMPS
GAPD	PDHm	ETFQO	r2412	GALUi	GK1
ACYP	LDH_D	r0205	GDHm	UDPGP	CYTK4
PGM	ME2m	CYOOm2i	LEUTAm	GLGNS1	NDPK1
ENO	ME1m	CYOOm3i	ILETAm	ORPT	CYTK3
PYK	r1109	NADK	PHETA1m	OMPDC	DADK
PYK2	CSm	ATPasel_1	TYRTAim	UMPK	CYTK1
PYK3	MDHm	ATPM	GLUDym	UMPK2	CYTK2
PYK4	FUMm	CYOR_u10mi	ASPTAm	UMPK3	NDPK3
PCm	SUCD1m	FADH2ETC	VALTAim	UMPK4	NDPK9
r0280	SUCOAS1m	ETFQO	ARGNm	UMPK5	NDPK10
AGPOP	SUCOASm	r0205	GNK	UMPK6	DGK1
PYRt2m	AKGDm	CYOOm2i	PGLc	UMPK7	URIDK2r
PEPCK_re	ICDHym	CYOOm3i	PGL	CYTK14	RNDR1
OAADC	ICDHxm	NADK	GLCNte	CYTK7	RNDR2
LDH_L	ACONTm	ATPasel_1	G6PDH2r	NDPK2	RNDR3
MALOAAtm	r0509	ATPM	G6PDH2c	CTPS1	RNDR4
L_LACDcm	r0317	ARGSS	GND	CTPS2	CYTK8
AKGMALtm	MDH	ARGSL	GNDc	PYK6	CYTK10
THD1m	NADH2_u10mi	ARGN	RPEc	HXPRT	CYTK11
AMPDA	NDPK6	NDPK8	NDPK4	CPK1	CYTK13
CYTK2_1	NDPK7	NDPK5	RE0453C		

Appendix C Medium composition

Table 6: Composition of α-minimal essential medium [87].

Begin of Table 6		
Metabolite	**ID**	**Influx**
Glycine	gly	0.666667
L-Alanine	ala__L	0.280899
L-Arginine	arg__L	0.49763
L-Asparagine	asn__L	0.378788

Continuation of Table 6		
Metabolite	**ID**	**Influx**
L-Aspartic acid	asp__L	0.225564
L-Cysteine	cys__L	0.667223
L-Glutamic Acid	glu__L	0.510204
L-Glutamine	gln__L	1.870968
L-Histidine	his__L	0.2
L-Isoleucine	ile__L	0.4
L-Leucine	leu__L	0.4
L-Lysine	lys__L	0.39726
L-Methionine	met__L	0.100671
L-Phenylalanine	phe__L	0.193939
L-Proline	pro__L	0.347826
L-Serine	ser__L	0.238095
L-Threonine	thr__L	0.403361
L-Tryptophan	trp__L	0.04902
L-Tyrosine	tyr__L	0.198895
L-Valine	val__L	0.393162
Ascorbic Acid	ascb__L	0.284091
Choline	chol	0.007143
Pantothenate	pnto__R	0.002096
Folic Acid	fol	0.002268
Niacinamide	ncam	0.008197
Pyridoxal hydrochloride	pydx	0.004902
Riboflavin	ribflv	2.66E-04
i-Inositol	inost	0.011111
Adenosine	adn	0.037453
Cytidine	cytd	0.041152
Guanosine	gsn	0.035336
Uridine	uri	0.040984
2'Deoxyadenosine	dad_2	0.039841
2'Deoxycytidine	dcyt	0.037879
2'Deoxyguanosine	dgsn	0.037453
Thymidine	thymd	0.041322
D-Glucose (Dextrose)	glc__D	6
Lipoic acid	lipoate	9.71E-04
Pyruvate	pyr	1

Table 7: Composition of fetal bovine serum [54, 88].

Metabolite	**ID**	**Influx**
Tetradecanoate (n-C14:0)	ttdca	1.23E-08
Hexadecanoate (n-C16:0)	hdca	3.15E-07
Octadecanoate (n-C18:0)	ocdca	2.32E-07
Hexadecenoate (n-C16:1)	hdcea	1.8E-08
Vaccenic acid	vacc	2.49E-08
Linoleic acid	lnlc	0.005
Gamma-linolenic acid	lnlncg	3.78E-09
Dihomo-gamma-linolenic acid (n-C6)	dlnlcg	3.9E-08
Arachidic acid	arach	1.11E-07
Timnodonic acid	tmndnc	3.48E-09
Clupanodonic acid	clpnd	3.93E-08
Cervonic Acid (n-C22:6)	crvnc	4.28E-08
D-Glucose	glc__D	6.9
Bilirubine	bilirub	0.0103
Urea	urea	2.664
Urate	urate	0.1725
Cortisol	crtsl	1.38E-09
Testosterone	tststerone	1.39E-10
Progesteron	prgstrn	2.55E-10
Prostaglandin E	prostge2	1.66E-08
Prostaglandin E	HC02213	1.68E-08
Prostaglandin F	prostgf2	3.74E-08
Vitamine A	retinol	3.14E-07
Vitamine A	retinal	3.16E-07
Cholesterol	chsterol	0.0008

Appendix D Biomass reaction

Table 8: Metabolites contributing to the Lipids biomass reaction, their sbml ID and their stoichiometric coefficient.

Begin of Table 8		
Metabolite	**ID**	**Stoichiometric coefficient**
L-alanine	ala__L_c	-0.505626
L-arginine	arg__L_c	-0.359260
L-asparagine	asn__L_c	-0.279425
L-aspartate	asp__L_c	-0.352607
L-cysteine	cys__L_c	-0.046571
L-glutamine	gln__L_c	-0.325996
L-glutamate	glu__L_c	-0.385872

Continuation of Table 8		
Metabolite	**ID**	**Stoichiometric coefficient**
Glycine	gly_c	-0.538891
L-histidine	his__L_c	-0.126406
L-isolycine	ile__L_c	-0.286078
L-leucine	leu__L_c	-0.545544
L-lysine	lys__L_c	-0.592114
L-methionine	met__L_c	-0.153018
L-phenylalanine	phe__L_c	-0.259466
L-proline	pro__L_c	-0.412484
L-phenylalanine	phe__L_c	-0.259466
L-proline	pro__L_c	-0.412484
L-serine	ser__L_c	-0.392525
L-threonine	thr__L_c	-0.312690
L-tryptophan	trp__L_c	-0.013306
L-tyrosine	tyr__L_c	-0.159671
L-valine	val__L_c	-0.352607
Ceramide	crm_hs_c	-0.000851
Cholesterol	chsterol_c	-0.115354
Diaglycerol	dag_hs_c	-0.008037
Lysophosphatidic acid	alpa_hs_c	-0.000359
Lysophosphatidylcholine	lpchol_hs_c	-0.003136
Phosphatidylcholine	pchol_hs_c	-0.099066
Phosphatidylethanolamine	pe_hs_c	-0.033549
Phosphatidylglycerol	pglyc_hs_c	-0.000079
Phosphatidylinositol	pail_hs_c	-0.004397
Phosphatidylserine	ps_hs_c	-0.007749
Sphingomyelin	sphmyln_hs_c	-0.018731
CTP	ctp_c	-0.039036
GTP	gtp_c	-0.036117
UTP	utp_c	-0.053446
dATP	datp_n	-0.013183
dCTP	dctp_n	-0.009442
dGTP	dgtp_n	-0.009898
dTTP	dttp_n	-0.013091
D-glucose-6-phosphate	g6p_c	-0.275194
ATP	atp_c	-20.704451
Water	h2o_c	-20.650823
ADP	adp_c	20.650823
Hydrogen	h_c	20.650823
Phosphate	pi_c	20.650823