Reconstruction and validation of a constraint-based metabolic network model for bone marrow-derived mesenchymal stem cells

H. Fouladiha*, S.-A. Marashi* and M. A. Shokrgozar†

*Department of Biotechnology, College of Science, University of Tehran, Tehran, 1417614411, Iran and †National Cell Bank of Iran, Pasteur Institute of Iran, Tehran, 1316943551, Iran

Received 26 January 2015; revision accepted 14 April 2015

Abstract

Objectives: Over recent years, constraint-based modelling of metabolic networks has become increasingly popular; the models are suitable for system-level modelling of cell physiology. The goal of the present work was to reconstruct a constraint-based metabolic network model of bone marrow-derived mesenchymal stem cells (BMMSCs).

Materials and methods: To reconstruct a BMMSC-specific metabolic model, transcriptomic data of BMMSCs, and additionally, the human generic metabolic network model (Recon1) were used. Then, using the mCADRE algorithm, a draft metabolic network was reconstructed. Literature and proteomic data were subsequently used to refine and improve the draft. From this, *i*MSC1255 was derived to be the metabolic network model of BMMSCs.

Results: *i*MSC1255 has 1255 genes, 1850 metabolites and 2288 reactions. After including additional constraints based on previously reported experimental results, our model successfully predicted BMMSC growth rate and metabolic phenotypes.

Conclusions: Here, *i*MSC1255 is introduced to be the metabolic network model of bone marrow-derived mesenchymal stem cells. Based on current knowledge, this is the first report on genome-scale reconstruction and validation of a stem cell metabolic network model.

Introduction

Mesenchymal stem cells

Mesenchymal stem cells (MSCs) are multipotent cells capable of differentiating into multiple cell types, including osteocytes, chondrocytes, adipocytes, myocytes and even non-mesenchymal cells such as hepatocytes (1). They also have immune-regulatory roles (2), for example, they do not greatly stimulate host immune systems in autologous or even allogenic cell transplantations (3). MSCs can be isolated from different tissues, including bone marrow and adipose tissue (4) and these features have made them good candidates for tissue engineering and regenerative medicine (5,6). Furthermore, MSCs do not express HLA type 2 and they cannot stimulate T cells (7); additionally, they are easy to manipulate genetically (8). Thus, they can be used as gene and protein carriers in gene therapy (9,10).

As a result of diverse applications of MSCs in (bio) medical and stem cell studies (11), understanding systems biology of these cells can be of great importance (12). Consequently, a reliable metabolic network model of MSCs can help in understanding physiological, metabolic and growth behaviour of these cells.

Metabolic network reconstruction and its applications

Reactions, metabolites and enzymes (which are controlled by metabolic genes), are responsible for managing the metabolism of cells (13). Constraint-based reconstruction and analysis of metabolic networks is a mathematical approach to provide *in silico* models of cell metabolism by considering gene-protein-reaction relationships (14–16).

Historically, classic small-scale metabolic network models have been reconstructed, by collecting the main metabolic reactions and pathways of a specific cell type [see (17–19)]. On the other hand, with the availability

Correspondence: S.-A. Marashi, Department of Biotechnology, College of Science, University of Tehran, Tehran, Iran. Tel.: +98-21-6649-1622; E-mail: marashi@ut.ac.ir

of 'omics' data over the past 15 years, hundreds of genome-scale metabolic network models have been reconstructed (20,21); these typically include hundreds to thousands of reactions, metabolites and genes (22).

In systems biology studies, metabolic network reconstruction and analysis has enabled constitution of a virtual '*in silico* laboratory' where computational analysis of cell metabolism can be performed faster and cheaper compared to 'wet lab' analyses. Various computational methods on hundreds of genome-scale metabolic models have provided a broad spectrum of applications in cell biology (23). For instance, metabolic models have been used in drug targeting (24) and in studying cancer (25,26). Assessing the consequences of regulation and/or mutation of a desired phenotype is a further application of metabolic network analysis (27).

Recon1 (28) is a generic genome-scale metabolic network of human cells that has been frequently used for modelling human metabolism. For example, by constraint-based analysis of fluxes in Recon1, some drug targets for hypercholesterolaemia and reactions involved in haemolytic anaemia have successfully been recognized (28). In a further study, biomarkers of human inborn errors of metabolism have been predicted using Recon1 (29), and the results have been shown to be in agreement with known mutations.

However, the generic metabolic network of human cells has limited applications, although many studies, use cell- and tissue-specific metabolic network models for accurate prediction of metabolism in different human tissues (30–32).

In the present work, using transcriptome and proteome data of bone marrow-derived mesenchymal stem cell (BMMSC), a constraint-based metabolic network model for these cells was reconstructed. The model was further validated using experimental data available in the literature to which it acquired a good level of consistency. Thus, this model is recommended for use in systems biology studies. In the light of current knowledge, this work is the first report on genome-scale reconstruction and validation of a stem cell metabolic network model.

Materials and methods

Data sets

To determine genes expressed in BMMSCs, transcriptome data were used from the Gene Expression Omnibus database (33). The selected five data series are as follows:

- GSE37470 data series (34): This includes microarray data of two normal early passage BMMSC samples,

GSM920586 and GSM920587, which are used for network reconstruction. Four other cell types (later passage BMMSCs and BMMSCs of large granular lymphocyte leukaemia patients) were NOT used in our work;

- GSE7637 data series (35): This includes microarray data of three early passage BMMSC samples, GSM184636, GSM184637 and GSM184638, while data of other cell types (later passages) were NOT used in our work;
- GSE7888 data series (35): Similar to the previous data series, this includes microarray data of three early passage BMMSC samples, GSM194075, GSM194076, GSM194077, GSM194078 and GSM194079. The data of other cells (later passages) were NOT used in our work;
- GSE30807 data series (36): This includes microarray data of a normal early passage BMMSC sample, GSM764199, which IS used in this work. The data of other cell types (osteosarcoma U2OS cells) were NOT used here;
- GSE32171 data series (37): This includes three early passage BMMSC samples, GSM797497, GSM797498 and GSM797499, which ARE used in the present study. The data of other type s (human MSCs in cardiomyocyte co-culture) were NOT used.

It should be noted that all these data series are based on Affymetrix Human Genome U133 Plus 2.0 Array platform. Using this platform, one obtains gene expression data of the sample, which was used directly by the network reconstruction algorithm (see below).

In the present study, a top-down strategy was used to reconstruct a cell-specific metabolic network; the generic model of human metabolism, Recon1 (28), was used in the first step. Then, those reactions of Recon1 (for which gene expression evidence exists), were chosen for inclusion in the initial draft of the BMMSC metabolic network (see below). For model refinement, a comprehensive proteome data of BMMSCs (38) was used, which includes 1676 proteins present both in BMMSCs and umbilical cord vein-derived MSCs.

Finally, cells of the A549 line, adipose tissue, blood and bone marrow, foetal cartilage, skeletal muscle and neutrophil metabolic network models [from (39)] were used for comparison with our model.

Reconstruction of the metabolic network model

Figure 1 represents a general overview of the framework used in this work. As mentioned above, to reconstruct a human cell-specific draft metabolic network, the top-down approach was used. This approach is based on

Figure 1. Schematic representation of the method used for BMMSC metabolic network reconstruction, refinement and validation. The first step, using mCADRE algorithm, the human generic metabolic network (Recon1) was pruned based on transcriptome data of BMMSC. Thus, unused reactions of Recon1, inactive in the mesenchymal stem cell, were removed. Consequently, an initial draft metabolic network was achieved. Then, the draft was refined and improved using the knowledge on the BMMSCs. The final metabolic network model of BMMSC was called *iMSC1255*. Fluxes in this network were then constrained using experimental data of BMMSC. Finally, the metabolic network, *iMSC1255*, was used for modelling metabolic phenotypes, which were validated by experimentally reported data. In summary, 14 transcriptome microarray data samples (from five data series) were used as the input of mCADRE algorithm. Then, the proteome data, together with some literature data, were used in the refinement step to achieve the final draft. The experimental data of consumption/production fluxes of the main metabolites were used for expanding the set of exchange reaction and for adding constraints to obtain the *iMSC1255* model. Finally, by comparing computational results with experimental data of changes in concentrations of amino acids, *iMSC1255* was validated.

pruning (reducing) a generic human metabolic network (like Recon1). Using this method, inactive reactions are recognized and 'removed' from the initial model, according to the omics data of the cell type or tissue. In the present work, the mCADRE algorithm was used for this purpose (39). Briefly, the first step of mCADRE algorithm is scoring genes of the generic model according to their presence in the transcriptome data. These scores were attributed to reactions, using gene-protein-reaction relations of the initial model. Furthermore, topology of the metabolic network is taken into account to update reaction scores. Finally, reactions with scores below a certain threshold were removed to complete reconstruction of the (initial draft) cell-specific metabolic network.

The initial draft of cell-specific metabolic network, that is, output of mCADRE algorithm, can be improved in a model refinement step; here, we used the literature and also BMMSC proteome data for this purpose. Subsequently, we obtained the final draft of the BMMSC metabolic network model. Using experimental data of

BMMSC, this final draft was further constrained, then validated (see below).

Mathematical concepts used for modelling

Over recent years, constraint-based modelling of metabolic networks has become increasingly popular for system-level modelling of cell physiology (40–42). In this modelling framework, metabolic fluxes are analysed subject to a number of constraints, including stoichiometric and reversibility constraints. Metabolic flux through reaction i is the rate of transforming reactants to products of i per gram of cell dry weight. In a real metabolic network, range of possible fluxes of each reaction is limited by physical and biochemical constraints. Some examples are:

Stoichiometric constraints. As changes in concentration of intracellular metabolites are much slower than of extracellular ones, the living cell can be considered to be under steady-state conditions (42). Considering a

metabolic network with m metabolites and n reactions, the stoichiometric matrix, $\mathbf{S}$, is an $m \times n$ matrix. In this matrix, S_{ij} represents the stoichiometric coefficient of metabolite i in reaction j . It can be shown that, assuming vector $\mathbf{v}$ as an n -dimensional vector of the reaction fluxes, the metabolic network is in a steady-state if:

$$\mathbf{S} \cdot \mathbf{v} = \mathbf{0} \quad (1)$$

This constraint, usually referred to as the stoichiometric constraint, limits flux space according to mass balance law.

Thermodynamic constraints. Some reactions are known to be irreversible. Thermodynamic constraints limit flux of a reaction, say reaction i , to be always non-negative:

$$v_i \geq 0 \quad (2)$$

Capacity constraints. Sometimes, for a certain reaction, say i , lower bound and upper bound of the flux (minimum and maximum value that can be taken by the flux) can be determined experimentally or theoretically. Such a constraint, in the general form, can be shown as:

$$\alpha_i \leq v_i \leq \beta_i \quad (3)$$

Equations 1–3 constrain fluxes that can pass through reactions of a metabolic network, but exact flux of the reactions cannot be determined. However, assuming that a cell is achieving an objective, one can find a flux distribution that satisfies the optimality criterion. Linear programming helps calculating flux distribution, $\mathbf{v}$, that maximizes an objective function, $\mathbf{c}^T \mathbf{v}$, where $\mathbf{c}$ is a vector of constant coefficients. Linear programming formulation of the problem, which is usually called flux balance analysis (FBA) (43), can be summarized as:

$$\begin{aligned} \text{Max} \quad & \mathbf{c}^T \mathbf{v} \\ \text{Subject to:} \quad & \mathbf{S} \cdot \mathbf{v} = \mathbf{0} \\ & 0 \leq v_i \quad \text{for any irreversible reaction } i \\ & \alpha_i \leq v_i \leq \beta_i \quad \text{for } i = 1, \dots, n \end{aligned}$$

Solution of the above linear programme is flux distribution, say $\mathbf{f}$, which satisfies the above-mentioned constraints. Maximum value of the objective function, Z_{obj} , is equal to $\mathbf{c}^T \mathbf{f}$.

Typically, it is assumed that biomass production rate is the objective function of FBA, which means that $\mathbf{c}^T \mathbf{v} = v_{\text{biomass}}$. In case of stem cells, proliferation occurs before differentiation. Thus, in this study, biomass production flux was a good choice for the objective function in FBA.

Computational analyses

COBRA toolbox (44) was used for constraint-based analysis of metabolic networks. This has many useful functions, such as *OptimizeCbModel* for performing FBA (see above). Another related COBRA toolbox function, *dynamicFBA*, incorporates time into FBA, that is, it sequentially performs FBA within a certain time point and updates metabolite concentrations accordingly (45). Thus, using this function one can simulate dynamics of cell metabolism during growth. This function uses initial values of biomass, concentrations of cell culture medium components and duration of modelling as input. Changes in amount of biomass and concentration of external metabolites as a function of time are computed as the output.

As mentioned above, output of *OptimizeCbModel* function comprises flux distribution of the metabolic network which optimizes the objective function. A further function in the COBRA toolbox, *enumerateOptimalSolutions*, finds all 'extreme' alternative flux distributions which optimize the objective function (46). In other words, this function helps in assessing state of all fluxes in a metabolic network, while the objective function has its optimal value.

Dead-end metabolites are those of a metabolic network which are either produced only or consumed only (47). Such metabolites highlight places of missing reactions. In the model refinement step, we found dead-end-related reactions of Recon1. If activity of each reaction was confirmed by transcriptome information, we manually added that reaction to the network model. For this purpose, the *addReaction* function of COBRA toolbox was used.

In silico culture medium

To model experimental cell culture conditions, an *in silico* culture medium was defined for the metabolic network model. More technically, only exchange fluxes of oxygen and components of the medium were allowed to be negative. Lower bounds of other exchange fluxes were set to zero. In this way, only oxygen and metabolites of the culture medium can be consumed by the cell model. Additionally, range of each negative exchange flux (uptake reaction) is defined according to concentration of each component in the culture medium. The *in silico* culture medium of BMMSC was defined using concentrations of α -MEM components (Table S1), similar to previous studies (25). That is, lower bounds of exchange fluxes were set to -0.05 for amino acids, -0.005 for vitamins, -5 for glucose, -0.7 for glutamine and -1000 for oxygen, H_2O , H^+ , sulphate,

pyruvate, bicarbonate, phosphate, chloride, sodium, calcium and potassium (in mmol/grDW/h units).

For *in vitro* experiments concerning stem cell growth, it is common to supplement cell culture medium by adding 10–15% foetal bovine serum (FBS) to it. Consequently, exchange fluxes of some of the main FBS components can become negative (Table S2). Therefore, lower bounds of exchange fluxes of urea, linoleic acid, cholesterol and triiodotyronine were set to -0.005 mmol/grDW/h.

Results

Reconstructing the metabolic network model

The goal of the present work was to model metabolism of BMMSC using a constraint-based approach. For this purpose, an ‘initial draft’ metabolic network model was reconstructed using mCADRE algorithm based on transcriptome data of the MSCs (Fig. 1).

The initial draft metabolic network had 1061 genes, 1817 reactions and 1344 metabolites (Fig. 2). This initial draft was improved over several consecutive steps to obtain the final model. First, according to the proteome data, we found the genes that need to be present in the BMMSC metabolic model (but not added to the model by mCADRE based on transcriptome data). Consequently, 38 new genes (encoding 98 new enzymatic

reactions) were added to the initial draft. In the latter step, 165 new metabolites had been added to the network. Then, to remove dead end metabolites, we found reactions of Recon1 involving at least one of the newly added metabolites in the previous step. From these 887 reactions, 324 were confirmed based on the transcriptome data, but were absent in the initial draft network. These reactions (together with their associated genes and metabolites) were added to the initial draft model. Therefore, the ‘final draft’ included 2239 (i.e. $1817 + 98 + 324$) reactions (Fig. 2).

Finally, the exchange reaction for each component of the culture medium (α -MEM) was checked to be present in the draft. In this step, uptake reactions of some medium components (such as ascorbic acid and lipoate), together with their consumption metabolic pathways, were manually added to the metabolic model. Then, an *in silico* culture medium was defined for the draft, according to α -MEM components (see Table S1).

Similar to other constraint-based models of metabolism (25,48,49), a previously-reported biomass production reaction (50) was added to our metabolic network. Flux through this reaction can be understood as a proxy for rate of growth and proliferation.

Following the above-mentioned steps, a metabolic network model of BMMSC was constructed, which will be hereafter called *i*MSC1255. The model has 1255 genes, 1850 metabolites and 2288 reactions (Fig. 2); this is available in SBML format in the File S3. The importance of this model is that it is the first reported genome-scale metabolic network model of stem cells.

To compare *i*MSC1255 with other metabolic network models, three tissue specific metabolic models (‘adipose’, ‘blood’ and ‘bone marrow’) were chosen. Adipose, bone marrow, and also peripheral blood, are the main sources for isolating MSCs (51). Thus, we needed to show that our BMMSC metabolic network model was not included in metabolic networks of these tissues. These networks had been reconstructed using mCADRE algorithm (39) (available from: <http://www.biomedcentral.com/content/supplementary/1752-0509-6-153-s2.zip>). Number of genes of the BMMSC metabolic network model was compared to numbers of genes in the above-mentioned models (Fig. 3). Based on this analysis, 89.5% (respectively 90.9%) of the genes of the bone marrow metabolic network were also present in the adipose (respectively blood) metabolic network. On the other hand, only 60.5% of the genes of *i*MSC1255 were also found in the adipose metabolic model, while this number drops to 58.9% and 55.9% for blood and bone marrow metabolic models respectively.

In a further analysis, we aimed to understand the relationship between each of the four (reduced)

Figure 2. Venn diagram representing evolution of metabolic model of BMMSC during the reconstruction procedure. The first step, mCADRE algorithm reconstructed an initial draft, which included 1061 genes, 1817 reactions and 1344 metabolites. After improving the initial draft by using proteomics data, a final draft was obtained, which included 1255 genes, 2239 reactions and 1837 metabolites. Finally, by manual curation and searching in the literature, we improved the model to obtain *i*MSC1255. This final metabolic network model includes 1255 genes, 2288 reactions and 1850 metabolites.

Figure 3. Number of metabolic genes shared by 'BMMSC', 'adipose', 'bone marrow' and 'blood' metabolic network models.

tissue-specific models and the (original) generic model, Recon1. For this analysis, reactions of Recon1 were categorized into seven different functional categories (Fig. 4), according to their metabolic subsystems. The complete list of reaction subsystems and the seven categories are included in Table S4. Figure 4 shows percentages of reactions of Recon1 used in the metabolic models of BMMSC, adipose tissue, bone marrow and blood. In all groups, especially in the case of carbohydrate metabolism, *i*MSC1255 included more reactions compared to other models.

We repeated the *i*MSC1255 comparison with four other metabolic network models (cells of the A549 line, foetal cartilage, skeletal muscle and neutrophil metabolic networks). Results were qualitatively the

same as those of three previously mentioned tissues (see File S7).

Constraining the metabolic network model

To determine culture medium constraints imposed on cells, experimental data from the literature were used. These data are as follows:

Amino acids exchange fluxes. Amino acid exchange fluxes of BMMSC have previously been reported in static and dynamic cultures (52) (see Table S6). In our study, using these observed fluxes for each amino acid exchange reaction, we determined a lower bound and an upper bound. From the biological data, we estimated standard deviations to be in the region of 13% of observed values. Therefore, lower and upper bounds of the fluxes were extended by 13% from each side. Lower and upper bounds, used in our simulations, are shown in Table 1.

Glucose, lactate, ammonia and oxygen exchange fluxes. Transport rates of glucose, oxygen, lactate and ammonium can greatly influence metabolism. Experimentally measured fluxes of these reactions in BMMSC cultures (Table 2) were found from the literature.

We decided to apply constraints on transport reactions in a way that all reported experimental data lie within the allowable range. Accordingly, bounds of exchange fluxes were set as shown in Table 3 (see Table S6 for conversion of measurement units).

Figure 4. The percentage of reactions shared by the generic human metabolic network, Recon1, and each of the four tissue-specific metabolic models (BMMSC, adipose, blood and bone marrow models).

Table 1. Lower and upper bounds of amino acid exchange fluxes used for simulations using *i*MSC1255 (see text for more detail)

Amino acid name	Lower bound of the exchange flux ($\mu\text{mol}/\text{grDW}/\text{h}$)	Upper bound of the exchange flux ($\mu\text{mol}/\text{grDW}/\text{h}$)
L-Alanine	-0.4017	0.0969
L-Arginine	-0.5731	0.0458
L-Asparagine	-0.1723	0.0170
L-Aspartate	-0.0291	-0.0179
L-Cysteine	-0.0174	0.0907
L-Glutamate	-0.0032	0.0962
L-Glutamine	-0.5880	-0.1295
Glycine	-0.0109	0.1680
L-Histidine	-0.0283	-0.0038
L-Isoleucine	-0.0462	0.0005
L-Leucine	-0.0451	-0.0089
L-Lysine	-0.0201	-0.0040
L-Methionine	-0.0070	-0.0010
L-Ornithine	-0.0163	0.7115
L-Phenylalanine	-0.0215	-0.0070
L-Proline	-0.0143	0.0975
L-Serine	-0.0450	-0.0100
L-Threonine	-0.0153	-0.0053
L-Tyrosine	-0.0254	-0.0076
L-Valine	-0.0472	-0.0211

Table 2. Experimental exchange flux measurements of D-glucose, oxygen, D-lactate and ammonium in BMMSC cultures, reported in the literature. These results were used to constrain exchange fluxes in *i*MSC1255. Negative fluxes (respectively, positive fluxes) denote consumption (respectively, production)

Metabolite name	Reported exchange fluxes	References
D-Glucose	-0.9 ± 0.7 , -0.4 ± 0.2 , -0.6 ± 0.02 ,	(52)
	-0.62 ± 0.03 pmol/cell/h	
	-8.04 , -10.73 , -11.47 , -8.63 ,	(53)
	-7.10 pmol/cell/day	
	13.5 , 12.5 , 15.5 pmol/cell/day	(54)
Oxygen	357.3 ± 45.2 , 327.6 ± 47.8 ,	(55)
	272.8 ± 43.3 ,	
	378.2 ± 127.7 femtomol/cell/h	
	225.4 ± 41.4 femtomol/cell/h	(56)
	105.7 ± 21.7 , 113 ± 22.3 , 101.7 ± 18.1 ,	(55)
D-Lactate	98.2 ± 24.2 femtomol/cell/h	
	119.2 ± 18.4 femtomol/cell/h	(56)
	1.2 ± 0.4 , 1.2 ± 0.04 , 1.1 ± 0.8 ,	(52)
	2.3 ± 0.7 , 1.2 ± 0.03 pmol/cell/h	
	14.23 , 22.85 , 20.99 , 17.00 ,	(53)
Ammonium	14.98 pmol/cell/day	
	37 , 37.5 , 29 , 31 , 26 pmol/cell/day	(54)
	960 ± 148 , 829.1 ± 92.5 ,	(55)
	637.5 ± 122.3 ,	
	741.9 ± 295.5 femtomol/cell/h	

Table 3. Lower and upper bounds of amino acid exchange fluxes used for simulations using *i*MSC1255 (see the text for more detail)

Metabolite name	Lower bound of exchange flux (mmol/grDW/h)	Upper bound of exchange flux (mmol/grDW/h)
D-Glucose	-3.0	-0.60
Oxygen	-0.50	-0.25
D-Lactate	+0.40	+8.0
Ammonium	+0.60	+6.0

Modelling BMMSC population growth rate

We used FBA for simulating cell population growth rate. By constraining transport fluxes of *i*MSC1255 (see above) and performing FBA, maximum biomass production rate was found to be 0.034 mmol/grDW/h.

Bone marrow-derived mesenchymal stem cell growth rate is largely dependent on cell culture conditions, including culture medium composition and its changes over culture time, seeding method, number of cells used for seeding and incubation situations. Moreover, cell population growth rate depends on individual features of the donor (age, gender and genetics). To the best of our knowledge, maximum reported *in vitro* BMMSC population growth rate is 0.027 h^{-1} (53).

Routinely, in the metabolic network analysis literature, each gram of cell dry weight (grDW) is assumed to be equivalent to one millimole of 'hypothetical biomass metabolite'. Therefore, optimal *in silico* specific growth rate is equal to: $0.034 \text{ mmol/grDW/h} \times 1 \text{ grDW/mmol} = 0.034 \text{ h}^{-1}$

Hence, maximum *in silico* growth rate (0.034 h^{-1}) is very close to maximum experimental growth rate (0.027 h^{-1}).

Amino acid consumption and production: *in vitro* versus *in silico* trends

To further assess the predictive power of *i*MSC1255, we examined dynamics of amino acid metabolism *in silico*; for this purpose, a dynamic FBA approach was used. For initial values in this analysis, we used cell culture composition and initial amount of biomass in static and dynamic cultures, based on the experimental values previously reported (52). Dynamic FBA modelling was run until the amount of *in silico* biomass became equal to the amount of experimentally observed end-point biomass (<50 h for dynamic culture and <100 h for static culture). Note that as the *in silico* growth rate was more than the *in vitro* growth rate, time required for production of a certain amount of biomass *in silico* is less than the required time *in vitro*.

Figure 5. Changes in the concentration of amino acids during mesenchymal stem cell culture. Experimental data ($\square$: static culture; $\circ$: dynamic culture) were obtained from (52). Computational results (dark grey curve: static culture; light grey curve: dynamic culture) were obtained by running *dynamicFBA* function of COBRA toolbox (44) on the *iMSC1255* model.

Figure 5 summarizes simulation results for production or consumption of 20 amino acids, together with the experimental data reported in (52).

Based on Figure 5, one can observe that in the case of most amino acids, good consistency exists between experimental data and computational results. This finding implies that, as an *in silico* modelling approach, dynamic FBA of *iMSC1255* successfully predicted trends of amino acid concentration changes.

For some amino acids such as arginine, cysteine, glutamine, glutamate and ornithine, very little concentration

changes occurred in the growth medium during simulation. Increasing duration of *dynamicFBA* simulation makes the trend of concentration changes and consistency with computational results clearer for arginine and glutamine, even though concentrations of cysteine, glutamate and ornithine remained almost constant. Using *enumerateOptimalSolutions* function of COBRA toolbox, 1312 series of flux distributions were attained, each leading to maximum biomass production flux. Averages of exchange fluxes of cysteine and glutamate in these series were small (0.013 and 0.002 mmol/grDW/h,

respectively). Thus, at least small concentration changes could be expected for these two amino acids.

Discussion

Relationship between iMSC1255 and automatically generated metabolic networks

In the present study, numbers of genes of the BMMSC network was compared to numbers of genes in three previously reported tissue-specific networks (Fig. 3). Interestingly, we found that a large number of genes were shared by adipose, blood and bone marrow metabolic networks. The BMMSC metabolic network, on the other hand, was found to be very different. We also compared reactions of these metabolic networks. Again, iMSC1255 was found to include more reactions compared to other models, in all functional categories. This difference was found to be more significant in the case of carbohydrate metabolism and energy, which presumably shows the importance of these reactions in stem cell proliferation. Thus, iMSC1255 includes more reactions compared to the other three metabolic models. There are two possible reasons for this observation. First, numbers of expressed genes in BMMSC was more than those of adipose tissue, blood and bone marrow (see File S5). This fact was reflected from the number of metabolic genes included in the initial draft (Fig. 2), which was greater than numbers of genes in adipose stem cell, bone marrow and blood models. Secondly, the other three networks were generated automatically, without manual addition of reactions. In contrast, a considerable number of reactions were inserted into the model (for example, in the proteome analysis), which in turn increased the size of the BMMSC metabolic network.

Model validation

For validation of the metabolic network of bacterial cells, often the maximal *in silico* growth rate is compared to experimentally determined optimal growth rate of the cells. This comparison is based on the assumption that the biological objective of bacterial cells is typically to grow, as fast as possible, in their highly competitive environments.

For bacterial cells, it has been shown that depending on culture conditions, cells may have different objectives. For example, in certain conditions, cells may aim to maximize their population growth, while in other conditions they may achieve other objectives (for example, minimization of glucose uptake, maximization of ATP production, and more) (57). The case must be the same

for MSCs. More precisely, these cells may maximize their population growth in certain conditions, while under other conditions they may have different objectives (which might not even be fully known). This is presumably the reason for having different reported values for experimental population growth rates. Consequently, we used the maximum reported value of growth rate, which is expected to be closest to maximum possible growth rate.

There is a wide range of BMMSC cell population growth rates reported in the literature. We found many by literature survey, for example, 0.01–0.02 h⁻¹ (52), 0.016–0.027 h⁻¹ (53), 0.009 h⁻¹ (54), 0.0036–0.0087 h⁻¹ (56), 0.0014–0.0042 h⁻¹ (58) and 0.003 h⁻¹ (59). Therefore, maximum reported *in vitro* BMMSC growth rate was 0.027 h⁻¹. This value is close to the value computed by FBA (0.034 h⁻¹), based on the assumption that BMMSCs grow optimally. Thus, the *in silico* model successfully predicted growth rate of these cells. Slight differences between these values (26%) might be related to cell-to-cell and culture-to-culture variations. Alternatively, differences may be related to missing regulatory constraints in our model.

To further assess the predictive power of iMSC1255, dynamics of amino acid metabolism were modelled using dynamic FBA (Fig. 5). Acceptable consistency between the model and experimental observations implies that, as an *in silico* modelling approach, dynamic FBA of iMSC1255 successfully predicted the trends of amino acid concentration changes.

For some of the amino acids, such as arginine, cysteine, glutamine, glutamate and ornithine, very little concentration changes occurred in the growth medium during simulation. Increasing duration of *dynamicFBA* simulation makes the trend of concentration changes and consistency with computational results clearer for arginine and glutamine, even though the concentrations of cysteine, glutamate and ornithine remain almost constant. Using *enumerateOptimalSolutions* function of COBRA toolbox, 1312 series of flux distributions were attained, each leading to maximum biomass production flux. Averages of exchange fluxes of cysteine and glutamate in these series are small (0.013 and 0.002 mmol/grDW/h respectively). Thus, at least small concentration changes could be expected for these two amino acids.

In conclusion, here, iMSC1255 has been presented as a metabolic network for BMMSCs. We have shown that our model is in agreement with experimental data concerning metabolism of BMMSCs; the intention is to use iMSC1255 as a platform for investigating their metabolism *in silico*. This model can potentially have direct biological implications, as BMMSCs are not very easily maintained in the laboratory conditions. As the

first step, our goal will be to increase BMMSC growth rate (increasing biomass production flux) using the model. Rapid proliferation of stem cells is essential for their use in medicine, as it reduces the time of stem cell therapy.

Exploring metabolic changes during BMMSC differentiation is also an interesting field of study. Metabolic network models could be applied in such research to better understand mechanisms of differentiation. Furthermore, metabolic networks can be integrated with other kinds of network, such as gene regulatory networks, to increase accuracy of computational results (60–63). Altogether, we recommend that *in silico* modelling of BMMSCs using iMSC1255 can constitute a basis for pre-testing creative ideas of cell researchers, including cell differentiation and genetic engineering, before performing time-consuming and expensive *in vitro* experiments.

Acknowledgements

S.-A. Marashi acknowledges the financial support of University of Tehran for this research under grant number 28791/1/2. We also thank M. Ruchel and Dr. J. White for editing the English text of this manuscript.

References

- Zaher W, Harkness L, Jafari A, Kassem M (2014) An update of human mesenchymal stem cell biology and their clinical uses. *Arch. Toxicol.* **88**, 1069–1082.
- Ma S, Xie N, Li W, Yuan B, Shi Y, Wang Y (2014) Immunobiology of mesenchymal stem cells. *Cell Death Differ.* **21**, 216–225.
- Ryan JM, Barry FP, Murphy JM, Mahon BP (2005) Mesenchymal stem cells avoid allogeneic rejection. *J. Inflamm.* **2**, 8.
- Bellavia M, Altomare R, Cacciabauda F, Santoro A, Allegra A, Concetta Gioviale M *et al.* (2014) Towards an ideal source of mesenchymal stem cell isolation for possible therapeutic application in regenerative medicine. *Biomed. Pap.* **158**, 356–360.
- Abdallah BM, Kassem M (2008) Human mesenchymal stem cells: from basic biology to clinical applications. *Gene Ther.* **15**, 109–116.
- Amorin B, Alegretti AP, Valim V, Pezzi A, Laureano AM, da Silva MA *et al.* (2014) Mesenchymal stem cell therapy and acute graft-versus-host disease: a review. *Hum. Cell* **27**, 137–150.
- Shi M, Liu ZW, Wang FS (2011) Immunomodulatory properties and therapeutic application of mesenchymal stem cells. *Clin. Exp. Immunol.* **164**, 1–8.
- Dzau VJ, Gnecci M, Pachori AS (2005) Enhancing stem cell therapy through genetic modification. *J. Am. Coll. Cardiol.* **46**, 1351–1353.
- Dwyer R, Khan S, Barry F, O'Brien T, Kerin M (2010) Advances in mesenchymal stem cell-mediated gene therapy for cancer. *Stem Cell Res. Ther.* **1**, 25.
- Ozawa K, Sato K, Oh I, Ozaki K, Uchibori R, Obara Y *et al.* (2008) Cell and gene therapy using mesenchymal stem cells (MSCs). *J. Autoimmun.* **30**, 121–127.
- Bianco P, Cao X, Frenette PS, Mao JJ, Robey PG, Simmons PJ *et al.* (2013) The meaning, the sense and the significance: translating the science of mesenchymal stem cells into medicine. *Nat. Med.* **19**, 35–42.
- MacArthur BD, Ma'ayan A, Lemischka IR (2009) Systems biology of stem cell fate and cellular reprogramming. *Nat. Rev. Mol. Cell Biol.* **10**, 672–681.
- Wegner A, Meiser J, Weindl D, Hiller K (2015) How metabolites modulate metabolic flux. *Curr. Opin. Biotechnol.* **34**, 16–22.
- Oberhardt MA, Palsson BØ, Papin JA (2009) Applications of genome-scale metabolic reconstructions. *Mol. Syst. Biol.* **5**, 320.
- Zomorodi AR, Suthers PF, Ranganathan S, Maranas CD (2012) Mathematical optimization applications in metabolic networks. *Metab. Eng.* **14**, 672–686.
- Rezola A, Pey J, Rubio Á, Planes FJ (2014) *In-silico* prediction of key metabolic differences between two non-small cell lung cancer subtypes. *PLoS One* **9**, e103998.
- Stelling J, Klamt S, Bettenbrock K, Schuster S, Gilles ED (2002) Metabolic network structure determines key aspects of functionality and regulation. *Nature* **420**, 190–193.
- Wiback SJ, Palsson BØ (2002) Extreme pathway analysis of human red blood cell metabolism. *Biophys. J.* **83**, 808–818.
- Fell DA, Small JR (1986) Fat synthesis in adipose tissue. An examination of stoichiometric constraints. *Biochem. J.* **238**, 781–786.
- Bordbar A, Monk JM, King ZA, Palsson BØ (2014) Constraint-based models predict metabolic and associated cellular functions. *Nat. Rev. Genet.* **15**, 107–120.
- Monk J, Nogales J, Palsson BØ (2014) Optimizing genome-scale network reconstructions. *Nat. Biotechnol.* **32**, 447–452.
- Ryu JY, Kim HU, Lee SY (2015) Reconstruction of genome-scale human metabolic models using omics data. *Integr. Biol.*, DOI: 10.1039/C5IB00002E.
- Lewis NE, Nagarajan H, Palsson BØ (2012) Constraining the metabolic genotype–phenotype relationship using a phylogeny of *in silico* methods. *Nat. Rev. Microbiol.* **10**, 291–305.
- Oberhardt MA, Yizhak K, Rupp E (2013) Metabolically re-modeling the drug pipeline. *Curr. Opin. Pharmacol.* **13**, 778–785.
- Folger O, Jerby L, Frezza C, Gottlieb E, Rupp E, Shlomi T (2011) Predicting selective drug targets in cancer through metabolic networks. *Mol. Syst. Biol.* **7**, 501.
- Hadi M, Marashi SA (2014) Reconstruction of a generic metabolic network model of cancer cells. *Mol. Biosyst.* **10**, 3014–3021.
- Martínez VS, Dietmair S, Quek L-E, Hodson MP, Gray P, Nielsen LK (2013) Flux balance analysis of CHO cells before and after a metabolic switch from lactate production to consumption. *Biotechnol. Bioeng.* **110**, 660–666.
- Duarte NC, Becker SA, Jamshidi N, Thiele I, Mo ML, Vo TD *et al.* (2007) Global reconstruction of the human metabolic network based on genomic and bibliomic data. *Proc. Natl. Acad. Sci. USA* **104**, 1777–1782.
- Shlomi T, Cabili MN, Rupp E (2009) Predicting metabolic biomarkers of human inborn errors of metabolism. *Mol. Syst. Biol.* **5**, 263.
- Shlomi T, Cabili MN, Herrgard MJ, Palsson BØ, Rupp E (2008) Network-based prediction of human tissue-specific metabolism. *Nat. Biotechnol.* **26**, 1003–1010.
- Mardinoglu A, Nielsen J (2015) New paradigms for metabolic modeling of human cells. *Curr. Opin. Biotechnol.* **34**, 91–97.
- Ghaffari P, Mardinoglu A, Asplund A, Shoaie S, Kampf C, Uhlen M *et al.* (2015) Identifying anti-growth factors for human cancer cell lines through genome-scale metabolic modeling. *Sci. Rep.* **5**, 8183.
- Barrett T, Troup DB, Wilhite SE, Ledoux P, Evangelista C, Kim IF *et al.* (2011) NCBI GEO: archive for functional genomics data sets – 10 years on. *Nucleic Acids Res.* **39**, 1005–1010.

- 34 Mailloux AW, Zhang L, Moscinski L, Bennett JM, Yang L, Yoder SJ *et al.* (2013) Fibrosis and subsequent cytopenias are associated with basic fibroblast growth factor-deficient pluripotent mesenchymal stromal cells in large granular lymphocyte leukemia. *J. Immunol.* **191**, 3578–3593.
- 35 Tanabe S, Sato Y, Suzuki T, Suzuki K, Nagao T, Yamaguchi T (2008) Gene expression profiling of human mesenchymal stem cells for identification of novel markers in early- and late-stage cell culture. *J. Biochem.* **144**, 399–408.
- 36 Ying M, Liu G, Shimada H, Ding W, May WA, He Q *et al.* (2013) Human osteosarcoma CD49⁺CD133⁺ cells: impaired in osteogenic fate while gain of tumorigenicity. *Oncogene* **32**, 4252–4263.
- 37 Muller-Borer B, Esch G, Aldina R, Woon W, Fox R, Bursac N *et al.* (2012) Calcium dependent CAMTA1 in adult stem cell commitment to a myocardial lineage. *PLoS One* **7**, e38454.
- 38 Miranda HC, Herai RH, Thomé CH, Gomes GG, Panepucci RA, Orellana MD *et al.* (2012) A quantitative proteomic and transcriptomic comparison of human mesenchymal stem cells from bone marrow and umbilical cord vein. *Proteomics* **12**, 2607–2617.
- 39 Wang Y, Eddy J, Price N (2012) Reconstruction of genome-scale metabolic models for 126 human tissues using mCADRE. *BMC Syst. Biol.* **6**, 153.
- 40 Bordbar A, Palsson BØ (2012) Using the reconstructed genome-scale human metabolic network to study physiology and pathology. *J. Intern. Med.* **271**, 131–141.
- 41 Maarleveld TR, Khandelwal RA, Olivier BG, Teusink B, Bruggeman FJ (2013) Basic concepts and principles of stoichiometric modeling of metabolic networks. *Biotechnol. J.* **8**, 997–1008.
- 42 Orman MA, Berthiaume F, Androulakis IP, Ierapetritou MG (2011) Advanced stoichiometric analysis of metabolic networks of mammalian systems. *Crit. Rev. Biomed. Eng.* **39**, 511–534.
- 43 Orth JD, Thiele I, Palsson BØ (2010) What is flux balance analysis? *Nat. Biotechnol.* **28**, 245–248.
- 44 Becker SA, Feist AM, Mo ML, Hannum G, Palsson BØ, Herrgard MJ (2007) Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox. *Nat. Protoc.* **2**, 727–738.
- 45 Varma A, Palsson BØ (1994) Stoichiometric flux balance models quantitatively predict growth and metabolic by-product secretion in wild-type *Escherichia coli* W3110. *Appl. Environ. Microbiol.* **60**, 3724–3731.
- 46 Reed JL, Palsson BØ (2004) Genome-scale in silico models of *E. coli* have multiple equivalent phenotypic states: assessment of correlated reaction subsets that comprise network states. *Genome Res.* **14**, 1797–1805.
- 47 Orth JD, Palsson BØ (2010) Systematizing the generation of missing metabolic knowledge. *Biotechnol. Bioeng.* **107**, 403–412.
- 48 Selvarasu S, Karimi IA, Ghim G-H, Lee D-Y (2009) Genome-scale modeling and *in silico* analysis of mouse cell metabolic network. *Mol. Biosyst.* **6**, 152–161.
- 49 Wagner A, Zarecki R, Reshev L, Gochev C, Sorek R, Gophna U *et al.* (2013) Computational evaluation of cellular metabolic costs successfully predicts genes whose expression is deleterious. *Proc. Natl. Acad. Sci. USA* **110**, 19166–19171.
- 50 Thiele I, Swainston N, Fleming RMT, Hoppe A, Sahoo S, Aurich MK *et al.* (2013) A community-driven global reconstruction of human metabolism. *Nat. Biotechnol.* **31**, 419–425.
- 51 Cheng H-Y, Ghetu N, Wallace CG, Wei F-C, Liao S-K (2014) The impact of mesenchymal stem cell source on proliferation, differentiation, immunomodulation and therapeutic efficacy. *J. Stem Cell Res. Ther.* **4**, 237.
- 52 Higuera GA, Schop D, Spitters TW, van Dijkhuizen-Radersma R, Bracke M, de Bruijn JD *et al.* (2012) Patterns of amino acid metabolism by proliferating human mesenchymal stem cells. *Tissue Eng. Part A* **18**, 654–664.
- 53 Schop D, Janssen FW, van Rijn LD, Fernandes H, Bloem RM, de Bruijn JD *et al.* (2009) Growth, metabolism, and growth inhibitors of mesenchymal stem cells. *Tissue Eng. Part A* **15**, 1877–1886.
- 54 Dos Santos F, Andrade PZ, Boura JS, Abecasis MM, da Silva CL, Cabral JM (2010) *Ex vivo* expansion of human mesenchymal stem cells: a more effective cell proliferation kinetics and metabolism under hypoxia. *J. Cell. Physiol.* **223**, 27–35.
- 55 Pattappa G, Heywood HK, de Bruijn JD, Lee DA (2011) The metabolism of human mesenchymal stem cells during proliferation and differentiation. *J. Cell. Physiol.* **226**, 2562–2570.
- 56 Pattappa G, Thorpe SD, Jegard NC, Heywood HK, de Bruijn JD, Lee DA (2013) Continuous and uninterrupted oxygen tension influences the colony formation and oxidative metabolism of human mesenchymal stem cells. *Tissue Eng. Part C* **19**, 68–79.
- 57 Schuetz R, Kuepfer L, Sauer U (2007) Systematic evaluation of objective functions for predicting intracellular fluxes in *Escherichia coli*. *Mol. Syst. Biol.* **3**, 119.
- 58 Zhao F, Pathi P, Grayson W, Xing Q, Locke BR, Ma T (2005) Effects of oxygen transport on 3-D human mesenchymal stem cell metabolic activity in perfusion and static cultures: experiments and mathematical model. *Biotechnol. Prog.* **21**, 1269–1280.
- 59 Grayson WL, Zhao F, Izadpanah R, Bunnell B, Ma T (2006) Effects of hypoxia on human mesenchymal stem cell expansion and plasticity in 3D constructs. *J. Cell. Physiol.* **207**, 331–339.
- 60 Chandrasekaran S, Price ND (2010) Probabilistic integrative modeling of genome-scale metabolic and regulatory networks in *Escherichia coli* and *Mycobacterium tuberculosis*. *Proc. Natl. Acad. Sci. USA* **107**, 17845–17850.
- 61 Covert MW, Xiao N, Chen TJ, Karr JR (2008) Integrating metabolic, transcriptional regulatory and signal transduction models in *Escherichia coli*. *Bioinformatics* **24**, 2044–2050.
- 62 Shlomi T, Eisenberg Y, Sharan R, Ruppin E (2007) A genome-scale computational study of the interplay between transcriptional regulation and metabolism. *Mol. Syst. Biol.* **3**, 101.
- 63 Karr JR, Sanghvi JC, Macklin DN, Gutschow MV, Jacobs JM, Bolival B Jr *et al.* (2012) A whole-cell computational model predicts phenotype from genotype. *Cell* **150**, 389–401.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Table S1. Concentrations of α -MEM components

Table S2. FBS components and their concentrations

File S3. Metabolic network model of bone marrow-derived mesenchymal stem cells (*i*MSC1255) in SBML format

Table S4. Categorization of the metabolic subsystems of Recon1 into seven functional categories

File S5. Comparison of expressed genes in BMMSC, adipose, blood and bone marrow tissues

Table S6. Experimental measurements of amino acid exchange rates

File S7. Comparison of *i*MSC1255 with cells of the A549 line, foetal cartilage, skeletal muscle and neutrophil metabolic network models