A sperm-specific proteome-scale metabolic network model identifies non-glycolytic genes for energy deficiency in asthenozoospermia

Arvand Asghari, Sayed-Amir Marashi, and Naser Ansari-Pour

*Department of Biotechnology, College of Science, University of Tehran, Tehran, Iran; †Faculty of New Sciences and Technology, University of Tehran, Tehran, Iran

ABSTRACT

About 15% of couples experience difficulty in conceiving a child, of which half of the cases are thought to be male-related. Asthenozoospermia, or low sperm motility, is one of the frequent types of male infertility. Although energy metabolism is suggested to be central to the etiology of asthenozoospermia, very few attempts have been made to identify its underlying metabolic pathways. Here, we reconstructed SpermNet, the first proteome-scale model of the sperm cell by using whole-proteome data and the mCADRE algorithm. The reconstructed model was then analyzed using the COBRA toolbox. Genes were knocked-out in the model to investigate their effect on ATP production. A total of 78 genes elevated ATP production rate considerably of which most encode components of oxidative phosphorylation, fatty acid oxidation, the Krebs cycle, and members of the solute carrier 25 family. Among them, we identified 11 novel genes which have previously not been associated with sperm cell energy metabolism and may thus be implicated in asthenozoospermia. We further examined the reconstructed model by in silico knock out of currently known asthenozoospermia implicated-genes that were not predicted by our model. The pathways affected by knocking out these genes were also related to energy metabolism, confirming previous findings. Therefore, our model not only predicts the known pathways, it also identifies several non-glycolytic genes for deficient energy metabolism in asthenozoospermia. Finally, this model supports the notion that metabolic pathways besides glycolysis such as oxidative phosphorylation and fatty acid oxidation are essential for sperm energy metabolism and if validated, may form a basis for fertility recovery.

Abbreviations: mCADRE: metabolic context-specificity assessed by deterministic reaction evaluation; ATP: adenosine triphosphate; RNA: ribonucleic acid; FBA: flux balance analysis; FVA: flux variability analysis; DAVID: database for annotation, visualization and integrated discovery; OXPHOS: oxidative phosphorylation; ETC: electron transfer chain; SLC: solute carrier; DLD: dihydrolipoamide dehydrogenase; SUCLG1: succinate-CoA ligase; SD: succinate dehydrogenase; OGDH: oxoglutarate dehydrogenase; FH: fumarate hydratase; IDH: isocitrate dehydrogenase; CS: citrate synthase; COBRA: constraint-based reconstruction and analysis toolbox; mCADRE: metabolic context-specificity assessed by deterministic reaction evaluation; FBA: flux balance analysis; FVA: flux variability analysis; SAF: synthetic accessibility factor; MI: metabolic interaction; GPB: gene product balance; mCADRE: metabolic context-specificity assessed by deterministic reaction evaluation; ATP: adenosine triphosphate; RNA: ribonucleic acid; KEGG: Kyoto Encyclopedia of Genes and Genomes; COBRA: constraint-based reconstruction and analysis toolbox; mCADRE: metabolic context-specificity assessed by deterministic reaction evaluation; SAF: synthetic accessibility factor; MI: metabolic interaction; GPB: gene product balance.

Introduction

Infertility is defined as failure to conceive after having unprotected intercourse for 12 months [Gurunath et al. 2011]. It is estimated that about 15% of couples suffer from infertility, with male factors thought to be responsible for about half of these cases [Hwang et al. 2011]. Sperm motility, morphology, and concentration are the main parameters affecting semen quality, and are thus routinely evaluated in clinical settings [Amaral et al. 2014]. Of these, low sperm motility (asthenozoospermia) is one of the most frequent types of male infertility explaining more than 40% of cases [Liu et al. 2015].

Although environmental factors such as unhealthy lifestyle, pollutants, and prolonged duration of sexual abstinence are thought to affect asthenozoospermia, genetic factors are thought to be mainly responsible for the underlying etiology [De Rosa et al. 2003; Sharpe 2010; Yu et al. 2014; Eslamian et al. 2015]. Asthenozoospermia has been extensively studied at the molecular level. For instance, comparative proteomic analyses of healthy and asthenozoospermic individuals were used to identify proteins linked to asthenozoospermia [Martínez-Heredia et al. 2008; Chan et al. 2009; Amaral et al. 2013a; Shen et al. 2013b].
2013; Amaral et al. 2014]. Other studies have focused on mouse knockout models and have identified novel genes such as Tekt4 and Smcp [Nayernia et al. 2002; Roy et al. 2007]. In addition, a number of studies have identified genetic polymorphisms associated with asthenozoospermia in multiple genes including Tektin-t, CATSPER1, and eNOS [Zuccarello et al. 2008; Visser et al. 2011; Song et al. 2015]. Despite all these, the underlying etiology of asthenozoospermia has yet to be explicated [Luconi et al. 2006; Martinez-Heredia et al. 2008; Du et al. 2015].

Several structural and functional proteins have been suggested to be responsible for sperm motility [Chan et al. 2009; Amaral et al. 2013a; Shen et al. 2013]. In addition to cytoskeletal and structural proteins in the sperm flagellum [Amaral et al. 2013c; Wang et al. 2013], metabolic enzymes have also been implicated [Ruiz-Pesini et al. 1998; Agarwal and Said 2004; Zhao et al. 2007]. Although the link between energy production and sperm motility has been explored in a number of studies [Ford 2006; Piomboni et al. 2012; Amaral et al. 2013b], a systematic study of metabolic genetic deficiencies pertaining to asthenozoospermia has not been undertaken [Amaral et al. 2014].

**Analysis of metabolic networks to understand disease etiology**

Constraint-based reconstruction and analyses of metabolic networks have shown a promising potential in modeling biochemical properties of cells [O’Brien et al. 2015]. To explore the complexity of cellular metabolism and to improve our understanding on the genotype-phenotype relationship, modeling and analysis of the omics-scale metabolic network of cells is of great importance [Mardinoglu and Nielsen 2012]. A genome-scale metabolic network model is a large system of chemical reactions and metabolites connected to each other by gene-protein-reaction relationships. In constraint-based modeling of cellular metabolism, based on an omics-scale metabolic model, metabolic fluxes which are required for biochemical functions are analyzed. Therefore, dysfunction of biochemical pathways in disease can be modeled by constraining the activity of certain reaction(s) in silico and determining the consequences [Ryu et al. 2015; Zhang and Hua 2015].

The first genome-scale human metabolic network, Recon1, was reconstructed in 2007 [Duarte et al. 2007], with its update, Recon2, released six years later [Thiele et al. 2013]. Cell- and tissue-specific metabolic network models have also been generated in a top-down approach by pruning these genome-scale models using proteome and transcriptome data obtained from that tissue [Jerby et al. 2010; Lewis et al. 2010; Fouladiha et al. 2015]. In addition, metabolic network models have been reconstructed to explore the underlying etiology of different tissue-specific disorders [Rezola et al. 2014; Salazar et al. 2016; Sohrabi-Jahromi et al. 2016], resulting in the identification of drug targets and biomarkers [Shlomi et al. 2009; Frezza et al. 2011; Mardinoglu et al. 2014]. Despite the significance of sperm cell metabolism in asthenozoospermia, to the best of our knowledge, no metabolic network model has been reconstructed for this cell. In this study, we reconstruct the first proteome-scale metabolic model of the sperm cell (Figure 1). The reconstructed model was then used to identify genes whose dysfunction can affect sperm motility through reduced ATP production. Our model successfully predicted several genes with known roles in asthenozoospermia (see Figure 2). Furthermore, the model identified 11 non-glycolytic genes that affect energy production efficiency. Mutations affecting the structure or expression of encoded enzymes may thus result in asthenozoospermia.

**Results and discussion**

**Proteome-scale reconstruction of the metabolic network in sperm cell**

Some studies in recent years provide evidence on the importance of the RNA pool in sperm cells [Johnson et al. 2011; Montjean et al. 2012], however, the mature spermatozoon is believed to be transcriptionally silent due to its chromatin undergoing condensation during spermatogenesis [Goodrich et al. 2013; Jodar et al. 2016]. Because of this, we employed the sperm proteomic data for model reconstruction. The model had been manually curated in order to have all essential reactions observed in sperm cell. The reconstructed proteome-scale metabolic network model of the sperm cell, namely SpermNet, is presented in SBML format in Supplementary file S1. This model can be read and analyzed by standard systems biology tools like the COBRA toolbox. SpermNet consists of 2,968 reactions, 2,034 metabolites and 1,242 genes, and is publicly available for those wishing to investigate different aspects of sperm cell metabolism and its related diseases.

**In silico gene knockout and model validation**

In order to validate the reconstructed model, we investigated its properties using flux balance analysis (FBA). The goal of this analysis was to predict those energy metabolism-related genes whose knockout changes the ATP production rate significantly (i.e., >10%). For this purpose, we knocked out all model genes one by one.
and explored their knockout effects on ATP production rate of the model. This analysis resulted in the identification of 78 genes listed in Supplementary file S2. Genes which encoded fully coupled reactions [Larhlimi et al. 2012] were grouped as a single gene set. Consequently, we obtained 29 independent gene sets in total. Of these, 18 had been previously reported to be linked to ATP production. The model therefore predicted 11 novel genes whose knockout decreased the ATP production rate. Furthermore, we independently analyzed each gene set to understand its role in ATP production in the sperm cell to better understand the underlying pathways affected in asthenozoospermia.

Prediction of down-regulated pathways

To evaluate the effects of each gene knockout on the pathways in our model, we in silico knocked out genes identified by FBA (N=78) one by one and explored their effect on model pathways using flux variability analysis (FVA). Especially, we were interested in high confidence decreased fluxes (Figure 3) because they are likely to be linked to the down-regulated pathways in ATP production, and hence, energy metabolism. We analyzed the high confidence decreased reactions with the database for annotation, visualization, and integrated discovery (DAVID) functional annotation tool [Huang et al. 2009]. Interestingly, down-regulated pathways predicted by the model were highly correlated to the experimentally known function of that certain gene. The main down-regulated pathways were ATP metabolism, oxidative phosphorylation, O-glycan metabolism, and fatty acid and lipid metabolism. The full list of down-regulated pathways after each gene knockout is available in Supplementary file S3. The role of each pathway in ATP production and asthenozoospermia is discussed below.

Oxidative phosphorylation

Oxidative phosphorylation (OXPHOS) is one of the most important pathways in all cells. There has been
a long debate about the role of OXPHOS in sperm cells. Originally, it was believed that glycolysis was the primary source of ATP production in sperm cells and OXPHOS plays a negligible role in providing energy for this cell type [Storey 2008]. However, this idea has been disputed by several studies suggesting an important role for OXPHOS in sperm motility [Marchetti et al. 2002; Amaral and Ramalho-Santos 2010; Sousa et al. 2011]. The emerging idea is that several non-glycolytic pathways are required to function in the sperm cell to produce sufficient ATP [Amaral et al. 2013; Amaral et al. 2013c]. In support of the OXPHOS role in sperm motility, it has been observed that inhibition of different compounds of the electron transfer chain (ETC), despite the availability of glucose, leads to a drop in motility [Ruiz-Pesini et al. 2000; John et al. 2005]. In support of previous observations, our model predicted genes related to OXPHOS whose knockout reduces the ATP production rate in sperm. These genes can be classified into three different groups of cytochrome \( c \) reductase (complex III), cytochrome \( c \) oxidase (complex IV), and ATP synthase (complex V) related genes.

The first group predicted by the model consisted of 10 cytochrome \( c \) reductase related genes. This protein complex, which is involved in the electron transport chain of mitochondria, catalyzes the oxidation of ubiquinol and the reduction of cytochrome \( c \). This oxidation/reduction reaction results in a net transfer of proton across the membrane which adds to the proton gradient. As a member of ETC, it is not surprising that problems in complex III leads to deficiency in ATP

<table>
<thead>
<tr>
<th>Novel genes</th>
<th>Previously reported genes</th>
</tr>
</thead>
<tbody>
<tr>
<td>GALT</td>
<td>Cytochrome ( c ) reductase family</td>
</tr>
<tr>
<td>NAGK</td>
<td>Cytochrome ( c ) oxidase family</td>
</tr>
<tr>
<td>PGM3</td>
<td>ATP synthase family</td>
</tr>
<tr>
<td>GCNT1</td>
<td>SLC25A1</td>
</tr>
<tr>
<td>CMPK</td>
<td>SLC25A3</td>
</tr>
<tr>
<td>ENTPD4</td>
<td>SLC25A20</td>
</tr>
<tr>
<td>GLB1</td>
<td>DLD</td>
</tr>
<tr>
<td>CTS1</td>
<td>DLST</td>
</tr>
<tr>
<td>NAGA</td>
<td>OGDH</td>
</tr>
<tr>
<td>SLC35A2</td>
<td>CPT2</td>
</tr>
<tr>
<td>SLC35D2</td>
<td>PPA2</td>
</tr>
</tbody>
</table>

Figure 2. From the total of 29 predicted essential gene sets, 18 gene sets (62%) were previously reported in the literature to be linked to ATP production, and thus, to asthenozoospermia. To the best of our knowledge, the other 11 gene sets (38%) have not previously reported to be linked to asthenozoospermia. The accompanying table shows a list of these genes.

Figure 3. The four reaction categories which show flux decrease or increase in the knockout model based on FVA [Hadi and Marashi 2014]. X-axis represents total reactions possible flux ranges. The solid color lines represent the flux of reaction in normal sperm cell model, (ranges from \( a \) to \( d \)). The dashed arrows indicate the flux of reaction in knockout model (ranges from \( a \) to \( b \)). If \( a < b \leq c < d \) this means a high confidence decrease, however if \( c \leq d \leq a < b \) it means a high confidence increase. If \( a < c \leq b < d \), it will result in a decrease, while if \( c < a \leq d < b \) it will result in an increase.
production and impaired sperm motility [Ruiz-Pesini et al. 1998]. Moreover, an inhibitor of complex III, antimycin A, is reported to decrease sperm motility [Ford and Harrison 1981; Ruiz-Pesini et al. 2000]. Interestingly, several studies have shown that treatment of asthenozoospermic patients with coenzyme Q10 (ubiquinol) results in improvement of sperm motility [Balercia et al. 2009; Safarinejad et al. 2012] which might be as a result of higher activity in complex III.

The second group is the cytochrome c oxidase or complex IV family. Cytochrome c oxidase is a large transmembrane protein complex that catalyzes the final step of the electron transfer chain in eukaryotic mitochondria [Li et al. 2006]. This complex is suggested to be a major regulatory site for oxidative phosphorylation [Kadenbach et al. 2000]. The main function of this enzyme is catalyzing the transfer of electrons to oxygen to produce water. Transfer of electrons to oxygen is coupled to the translocation of protons and is necessary for ATP synthesis. In line with the model predictions, several studies have shown that levels of different protein members of this family are lower in asthenozoospermic samples than those with normal motility [Ruiz-Pesini et al. 1998; Ruiz-Pesini et al. 2000; Martínez-Heredia et al. 2008]. In addition, it has been observed that inhibition of complex IV results in a dramatic decrease in sperm motility [Brown 1995; Pascual et al. 1996].

The third group includes ATP synthase related genes. ATP synthase is the last part of the OXPHOS pathway and is responsible for the most important part of this pathway, i.e., generation of ATP. This enzyme uses the energy stored in the proton gradient across the membrane to drive the synthesis of ATP. Different studies have explored the impact of deficiency in ATP synthase genes on sperm motility [Ruiz-Pesini et al. 2007; Piomboni et al. 2012]. Moreover, oligomycin, a drug that directly blocks ATP synthase by mitochondria, is shown to negatively influence sperm motility [Dreanno et al. 1999].

**Solute carrier family 25**

The solute carrier family includes a group of proteins responsible for transporting different metabolites across the mitochondrial membrane [Palmiéri 2013]. These proteins are reported to be essential for efficient synthesis of ATP in mitochondria [Piomboni et al. 2012]. Interestingly, three members of the solute carrier (SLC) 25 family, namely SLC25A1, SLC25A3, and SLC25A20 were among the genes predicted to be significant by the model. SLC25A1 is a citrate carrier, which promotes efflux of tricarboxylic citrate to the cytoplasm in exchange for dicarboxylic cytosolic malate [Palmiéri and Pierri 2010; Sun et al. 2010]. To the best of our knowledge, no direct evidence is reported in the literature to support the role of this protein in sperm motility. However, since it is a mitochondrial transporter, it is reasonable to suggest its inactivity may lead to asthenozoospermia, especially that inhibition of this protein is shown to reduce the fluxes of several pathways, including sperm hyperactivated motility related pathways [Cappello et al. 2012].

The second protein in this family, SLC25A3, is a phosphate transporter. This protein facilitates the transport of phosphate into mitochondria either by proton co-transport or in exchange for hydroxyl ions. This protein is known to play a fundamental role in ATP synthesis in mitochondria [Mayr et al. 2007]. A recent study has reported that the transcript of SLC25A3 is differentially regulated in asthenozoospermia in comparison with the normal control group, suggesting a possible role for this protein in asthenozoospermia [Bansal et al. 2015].

SLC25A20 is a carnitine carrier. This protein transports acyl-carnitines into the mitochondrial matrix for their oxidation by the mitochondrial fatty acid-oxidation pathway [Fukushima et al. 2013]. It works in conjunction with carnitine palmitoyl transferase II, which oxidizes long-chain fatty acids in mitochondria. Interestingly, our model predicted both genes to have a role in ATP production, and hence, in sperm motility. Carnitine is known to have an important role in sperm motility. Insufficient carnitine in sperm is shown to result in severe motility problems [Tanphaichitr 1976; Hinton et al. 1979; Agarwal and Said 2004]. Interestingly, it has been observed that in asthenozoospermic samples, the ratio of acyl-carnitine to carnitine is greatly reduced, suggesting an improper activity of carnitine transferase [Golan et al. 1984; Bartellini et al. 1986].

**Citric acid cycle**

Citric acid cycle (or Krebs cycle) is a chemical pathway used by all aerobic organisms to generate energy from oxidation of acetyl-CoA. In eukaryotic cells, this pathway occurs in mitochondria and is closely related to OXPHOS. The relationship between the Krebs cycle and asthenozoospermia has been investigated [Zhao et al. 2007; Martínez-Heredia et al. 2008], which showed that Krebs related gene deficiencies are linked to asthenozoospermia. In the present work, our model predicted a number of Krebs cycle genes involved in ATP production of sperm cells, and thus, in asthenozoospermia. Dihydrolypoamide dehydrogenase (DLD), dihydrolypoamide S-succinyl transferase (DLST), oxoglutarate dehydrogenase (OGDH), citrate synthase (CS), fumarate
hydratase (FH), isocitrate dehydrogenase (IDH), succinate-CoA ligase (SUCLG1), and succinate dehydrogenase (SD) are the genes predicted by SpermNet to influence ATP production in sperm cells.

DLD, DLST, and OGDH are three enzymatic subunits which work together to form the oxoglutarate dehydrogenase complex (α-ketoglutarate dehydrogenase). This complex converts oxoglutarate to succinyl-CoA. Increase in DLD precursor, which is a sign of decrease in mature DLD, has been observed in different asthenozoospermic patients [Martínez-Heredia et al. 2008; Ashrafzadeh et al. 2013]. Decrease in OGDH has also been linked with reduced sperm motility, consistent with the predictions of SpermNet [Agarwal et al. 2016]. Amaral et al. [2014] reported a lower DLST enzyme ratio in asthenozoospermic samples compared with control samples. SUCLG1, CS, FH, SD, and IDH are important enzymes of the Krebs cycle and deficiency in each of them have been shown to lead to impaired sperm motility [Ruiz-Pesini et al. 1998; Ruiz-Pesini et al. 2000; Zhao et al. 2007; Martínez-Heredia et al. 2008; Tomar et al. 2010; Tomar et al. 2012; Amaral et al. 2014]. It should be emphasized that the above-mentioned alterations are genetically heterogeneous and therefore, different patients may have different genetic deficiencies leading to asthenozoospermia.

Fatty acid oxidation

Fatty acid metabolism plays an important role in energy production of cells through the beta-oxidation of fatty acids [Costa et al. 1994]. In one study, inhibition of fatty acid oxidation using etomoxir was shown to result in a decrease in sperm motility, indicating the importance of beta-oxidation in energy production of the sperm cell [Amaral et al. 2013c]. SpermNet predicted two genes in the fatty acid beta-oxidation pathway to have significant roles in ATP production and thus related to asthenozoospermia. Hydroxyacyl-CoA dehydrogenase/3-ketoacyl-CoA thiolase/enoyl-CoA hydratase, subunit A (HADHA) and hydroxyacyl-CoA dehydrogenase/3-ketoacyl-CoA thiolase/enoyl-CoA hydratase, subunit B (HADHB) are two subunits of the mitochondrial trifunctional enzyme, which catalyzes three out of four steps of beta-oxidation [Kamijo et al. 1994]. Proteomic analysis of human sperm cells has shown that decreased levels of HADHA in asthenozoospermic patients, supporting the relevance of our model prediction [Amaral et al. 2014]. This prediction is also in line with the new emerging idea that pathways other than glycolysis and OXPHOS participate in energy production in the sperm cell [Amaral et al. 2013c].

Pyrophosphatase

Another protein predicted to be linked with ATP production by SpermNet is pyrophosphatase (inorganic) 2 (PPIA). This enzyme resides on the mitochondrial membrane and is involved in producing inorganic phosphate (PPI) through hydrolyzing pyrophosphate. PPI is a stable, easy-to-store, high energy compound, which is able to substitute for ATP under certain conditions in glycolysis-related reactions such as attenuated respiration [Chi and Kemp 2000]. Consistent with the prediction of SpermNet, Yi et al. [2012] showed that PPI plays an important role in sperm motility and impaired PPA2 leads to decreased motility.

V-type ATPase family

Another group of genes predicted by our model is the V-type ATPase family. They are mainly found in the apical membrane of cells that line the epididymis and vas deferens in the male reproductive system, controlling pH of the sperm environment [Nishi and Forgac 2002]. The sperm cell needs an acidified environment for efficient motility and viability [Forgac 2007]. Interestingly, the V-type ATPase subunit E1, predicted by SpermNet to affect ATP production significantly, is only found in sperm cells and in the acrosome part of spermatids and mature sperms, thus suggesting a dual role in energy metabolism and acidification of the acrosome [Sun-Wada et al. 2002]. Furthermore, ATPase subunit A2 has been studied in another work, which showed that lower amounts of this protein will lead to decreased sperm motility [Ota et al. 2013].

Double gene knockout in silico simulations

We also in silico knocked out genes in pairs to explore the double gene-knockout effects. For this purpose, first, we removed all genes that have been found in the single knockout simulation (N=78) from our gene list. We then knocked out all remaining genes in pairs and then selected those pairs which collectively reduced the ATP hydrolysis rate more than 10%. A total of 157 gene pairs were found (the complete list of gene pairs is available in Supplementary file S4). After consolidating the fully coupled genes together, 18 sets of gene pairs were identified. Main pairs were involved in OXPHOS and fatty acid beta oxidation complexes, which further confirm the role
of these pathways in sperm energy metabolism and motility. Simultaneous knockout of electron transfer flavoprotein, which participates in catalyzing the initial step of the mitochondrial fatty acid beta-oxidation, with the mitochondrial respiratory chain complex I and II proteins are also among the pairs of genes which reduces the ATP hydrolysis rate. Genes related to solute carrier family 25 and the Krebs cycle were also among the identified pairs, both of which were discussed above.

**Novel predicted genes as potential biomarkers**

SpermNet predicted 11 novel genes which significantly affect ATP production and may thus be associated with asthenozoospermia (see Table 1). To the best of our knowledge, none of these genes have been reported to be mutated or associated with asthenozoospermic patients. These genes may therefore become targets of future experiments aimed to identify unknown causes of asthenozoospermia and if validated, may potentially be used as biomarkers.

Interestingly, galactose-1-phosphate uridylyltransferase (GALT) which catalyzes the second step of the Leloir pathway of galactose metabolism was identified. It has previously been observed that this pathway can be active in the sperm cell and play a role in its energy metabolism [Amaral et al. 2013c], although GALT has not been specifically linked to asthenozoospermia. The other ten novel genes predicted by SpermNet are involved in nucleotide interconversion, glucosamine metabolism, and keratan sulfate degradation. Given the function of these genes, it is interesting to speculate that spermatozoa deposited in the female reproductive tract find substrates for these gene-products during their voyage to the oocyte. If validated, this provides support for an underlying gene-environment interaction responsible for efficient sperm motility. These genes therefore need to be examined further to establish their association with asthenozoospermia.

SpermNet, however, did not predict genes related to glycolysis. These results would have been surprising if it were not for those of Amaral et al. [2014] which, based on proteomic data generated from asthenozoospermic patients, showed that expression of glycolytic genes are not significantly altered. This finding contributes to the open discussion about the role of glycolysis and other pathways in sperm energy metabolism. It was initially believed that sperm cell energy is predominantly provided through glycolysis, however, this hypothesis was later challenged by several studies showing that OXPHOS and other pathways also play roles in sperm energy metabolism [Ford 2006]. Ford showed that in mouse glycolytic gene knockout sperms experienced a drop in sperm motility [Ford 2006]. However, incubation of the exact same cells in glucose free medium lead to proper sperm motility, indicating that glucose is not the only energy source for these cells [Ford 2006]. Moreover, independent studies have shown that incubation of human sperm with different ETC complex inhibitors in the presence of glucose will still lead to a rapid drop in sperm motility, suggesting that dysfunction in OXPHOS will also lead to asthenozoospermia even when glycolysis is fully functioning [Ruiz-Pesini et al. 2000; John et al. 2005].

Based on the findings of Amaral et al. [2014] and SpermNet predictions, we suggest that glycolysis is not the main energy source for sperm motility and correct functioning of pathways like OXPHOS and fatty acid oxidation play a significant role. This suggests that although glycolysis plays a significant role in energy metabolism, sperm cell genes that encode components of non-glycolytic pathways have important roles in asthenozoospermia.

**False negatives of SpermNet in predicting asthenozoospermia-implicated genes**

Although SpermNet predicted 78 genes with a significant role in ATP production and thus potentially implicated in asthenozoospermia, it failed to predict a number of asthenozoospermia-implicated metabolic genes (see Table 2). This failure is probably due to an incomplete reconstructed model as the generic human model and the proteomic dataset used does not cover all proteins encoded in the human genome. Furthermore, reconstruction algorithms are not error-free and some genes and pathways are likely not represented in the model.

In order to investigate the knockout effect of these genes on SpermNet, we undertook FBA and FVA analysis. After removing each gene, we detected all activated and inactivated reactions and compiled them separately. This procedure resulted in 307 inactivated and 287

---

Table 1. Novel genes predicted by the SpermNet model to play a role in asthenozoospermia.

<table>
<thead>
<tr>
<th>Entrez Gene ID</th>
<th>Abbreviated name</th>
<th>Gene full name</th>
</tr>
</thead>
<tbody>
<tr>
<td>2592</td>
<td>GALT</td>
<td>Galactose-1-phosphate uridylyltransferase</td>
</tr>
<tr>
<td>5377</td>
<td>NAGK</td>
<td>N-acetyl glucosamine kinase</td>
</tr>
<tr>
<td>5238</td>
<td>PGK3</td>
<td>Phosphoglucomutase 1</td>
</tr>
<tr>
<td>2650</td>
<td>GCNT1</td>
<td>Glucosaminyl (N-acetyl) transferase 1</td>
</tr>
<tr>
<td>51727</td>
<td>CMPK</td>
<td>Cytidine monophosphate kinase</td>
</tr>
<tr>
<td>9583</td>
<td>ENTPD4</td>
<td>Ectonucleoside triphosphatase diphosphohydrolase 4</td>
</tr>
<tr>
<td>2720</td>
<td>GLB1</td>
<td>Galactosidase beta 1</td>
</tr>
<tr>
<td>5476</td>
<td>CTSA</td>
<td>Cathepsin A</td>
</tr>
<tr>
<td>11046</td>
<td>SLC35D2</td>
<td>Solute carrier family 35, member D2</td>
</tr>
<tr>
<td>7355</td>
<td>SLC35A2</td>
<td>Solute carrier family 35, member A2 (UDP-galactose transporter), member A2</td>
</tr>
<tr>
<td>4668</td>
<td>NAGA</td>
<td>N-acetylgalactosaminidase</td>
</tr>
</tbody>
</table>
activated genes. The DAVID functional annotation tool identified multiple pathways silenced by deleting these genes based on their overall enrichment score, which is modified p-value. In other words, DAVID tries to rank the biological significance of gene groups based on their overall p-value scores of all enriched annotation terms. Figure 4A shows the top 10 enriched pathways based on the inactivated genes. Interestingly, most of the silenced pathways were related to energy metabolism, which suggests that SpermNet is capable of predicting their knockout effects correctly. Notably, among these enriched pathways, peroxisome-related pathways are also present, which is consistent with findings of a recent study indicating that these pathways are active in sperm cells and play a role in its metabolism [Amaral et al. 2013c].

The top 10 enriched pathways based on activated genes are shown in Figure 4B. The underlying basis for gene

Table 2. List of genes that are known to be linked to asthenozoospermia but have not been predicted to influence ATP production by the SpermNet model.

<table>
<thead>
<tr>
<th>Entrez ID</th>
<th>Gene symbol</th>
<th>Gene name</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>2821</td>
<td>GPI</td>
<td>Glucose 6-phosphate isomerase</td>
<td>Amaral et al. 2014</td>
</tr>
<tr>
<td>5162</td>
<td>PDHB</td>
<td>Pyruvate dehydrogenase beta</td>
<td>Amaral et al. 2014</td>
</tr>
<tr>
<td>1891</td>
<td>ECH1</td>
<td>Enoyl-CoA Hydratase 1</td>
<td>Martinez-Heredia et al. 2008</td>
</tr>
<tr>
<td>2597</td>
<td>GAPDH</td>
<td>Glyceroldehyde-3-Phosphate Dehydrogenase</td>
<td>Sheng et al. 2013</td>
</tr>
<tr>
<td>5322</td>
<td>PGK2</td>
<td>Phosphoglycerate kinase 2</td>
<td>Shen et al. 2013</td>
</tr>
<tr>
<td>4538</td>
<td>MT-ND4</td>
<td>Mitochondrially encoded NADH dehydrogenase 4</td>
<td>Pereira et al. 2008</td>
</tr>
<tr>
<td>4540</td>
<td>MT-ND5</td>
<td>Mitochondrially encoded NADH dehydrogenase 5</td>
<td>Dada et al. 2011</td>
</tr>
<tr>
<td>1080</td>
<td>CFTR</td>
<td>Cystic fibrosis transmembrane conductance regulator</td>
<td>Slezak et al. 2007</td>
</tr>
<tr>
<td>116369</td>
<td>SLC26A8</td>
<td>Solute carrier family 26 (anion exchanger), member 8</td>
<td>Dirami et al. 2013</td>
</tr>
</tbody>
</table>

Figure 4. Enriched pathways based on inactivated or activated genes in gene knockout simulations. (A) Enriched pathways based on 307 inactivated genes as a result of deleting asthenozoospermia-implicated metabolic genes not predicted by SpermNet. (B) Enriched pathways based on 287 activated genes as a result of deleting asthenozoospermia-implicated metabolic genes not predicted by SpermNet.
activation, after deleting asthenozoospermia-implicated metabolic genes from SpermNet, lies with the objective function of the model. Since its aim is to maximize the rate of ATP production, the model achieves this objective by activating other compensatory reactions, hence, up-regulating previously silent genes.

**Future study**

The first proteome-scale constraint-based metabolic network model of an asthenozoospermia sperm cell was constructed. The model was validated by comparing the predicted genes with asthenozoospermia-implicated genes in the literature, based mainly on mouse knockout models and inhibition assays. However, we believe that these findings have to be validated with in vitro experiments. For instance, knockout mouse models of their orthologs may be used to experimentally validate their functional effect. Mutational analysis of a large cohort of asthenozoospermic patients may also provide evidence for their association.

Additional confirmation can be obtained through drug targeting and then exploring their inhibitory effects on sperm energy metabolism and motility in vitro. This has been done for some genes predicted by SpermNet previously, such as antimycin A which inhibits complex III of ETC and oligomycin which blocks mitochondrial ATP synthase [Dreanno et al. 1999; Ruiz-Pesini et al. 2000]. For the novel predicted genes, we identified all possible drugs that may target these genes (for a complete list of drugs see Supplementary file S5). Notably, gemcitabine is a drug that inhibits the function of CMPK and can thus be used to analyze its effect on sperm motility [Vernejoul et al. 2006]. Further studies analyzing the interaction of potential drugs with these novel genes may have the potential to recover fertility in asthenozoospermic patients.

SpermNet is capable of predicting several related genes to asthenozoospermia and it confirms the emerging idea that several non-glycolytic pathways are implicated in asthenozoospermia and should be considered when analyzing asthenozoospermic patients. The findings presented above suggest that mitochondria and its related pathways (such as OXPHOS) play a significant role in energy metabolism of the sperm cell and any potential deficiency would likely lead to impaired sperm motility.

**Materials and methods**

**Sperm-specific metabolic network reconstruction**

In the present study, we used an updated version of Recon2 as the generic template model [Swainston et al. 2016]. This model includes 1,713 genes, 5,391 metabolites, and 7,852 reactions. To generate a sperm cell-specific model, proteomic data pertaining to the sperm cell were collected from two sources [Amaral et al. 2013a; Wang et al. 2013] and meta-analyzed. The consolidated protein list was then converted to a list of Entrez Gene IDs using DAVID [Huang et al. 2009]. For any gene being shared in both gene sets, a score of 1 was assigned in the expression array, denoting this gene as active in the sperm-specific metabolic network model. All other genes in the generic model absent in the converted gene set were assigned a score of 0. Among the different algorithms developed for tissue- or cell-specific model reconstruction, we used the metabolic context-specificity assessed by deterministic reaction evaluation (mCADRE) algorithm [Wang et al. 2012] for our model reconstruction due to its ability in considering network topology, reducing computational burden, and its known superiority in accurately inferring reactions. This algorithm uses a high-confidence set of core reactions from the human generic template model (e.g., Recon1, Recon2, etc.) based on tissue-specific expression evidence. Other non-core reactions are ranked according to their expression evidence and their link to other network reactions, and then removed in the inverse order of this ranking. Each reaction removal is checked by the consequent flux capacity of core reactions and a universal test of metabolic functionality to see whether that removal should be accepted or rejected [Wang et al. 2012]. The reconstructed model was then analyzed using multiple methods in the COBRA Toolbox [Schellenberger et al. 2011], such as FBA [Orth et al. 2010] and FVA [Mahadevan and Schilling 2003].

**Flux balance analysis (FBA)**

Metabolic network reactions can be represented as an $m \times n$ stoichiometric matrix $S$, in which each column corresponds to a reaction and each row represents a specific metabolite. Flux through all network reactions is represented by vector $v$. The main assumption in the network analysis is that the system is working on (quasi-)steady state conditions:

$$Sv = 0$$ (1)

Each reaction in the model is either reversible or irreversible. This classification is achieved by restricting each reaction flux with thermodynamic constraints ($v_i^{\text{min}} = 0$ for any irreversible reaction $i$). Furthermore, lower- and upper-bounds might be known for other reactions. All such constraints can be summarized in the “capacity” constraint

$$v^{\text{min}} < v < v^{\text{max}}$$ (2)
where $v^{\text{min}}$ and $v^{\text{max}}$ are the vectors of lower- and upper-bounds of reactions, respectively.

FBA is a mathematical optimization approach which relies on constraints (1) and (2) to gain insight about the reaction fluxes that maximize a biological objective in the model [Orth et al. 2010]. Here, we used FBA to explore the contribution of reactions to the maximization of the objective function, adenosine triphosphate (ATP) hydrolysis reaction (ATPasel). ATPasel works as an objective function for maximizing the ATP production, because in order to have the maximum ATP hydrolysis, the model first needs to produce maximum ATP. In this model, 994 reactions should carry flux in order to maximize the objective function.

**Knockout simulations**

Using constraint-based modeling, one can predict the effect of knocking out any gene in the model on the optimal value of the objective function [Bordbar et al. 2014]. Here, we simulated a gene knockout for every single gene in the model and studied these knockout effects on ATP production. This was implemented by the function "SingleGeneDeletion" in the COBRA toolbox [Schellenberger et al. 2011]. This function simulates the gene knockout by changing the gene associated reaction bounds to zero, thus inhibiting the flux of those reactions. Furthermore, in every knockout stage, FBA is used to evaluate the amount of ATP hydrolysis in the model by using ATPasel as the objective function. One of the important outputs of this function is $\text{grRatio}$, which is defined as follows:

$$\text{grRatio} = \frac{\text{max (ATP Hydrolysis rate in knockout model)}}{\text{max (ATP hydrolysis rate in normal model)}}$$

We decided to have a cutoff of 0.9 for $\text{grRatio}$, which means that we selected any gene whose knockout caused the model to have an ATP hydrolysis rate of less than 90% of that in the normal model (see Supplementary file S6).

A comprehensive list of genes thought to be responsible for asthenozoosperma through deficient energy metabolism was gathered from the literature. The metabolic consequences of gene knockout was simulated for each associated gene using our proteome-scale model. In specific, by using FBA and by setting the flux of ATPasel as the objective function, we identified those reactions that show flux changes in comparison with the fluxes of the normal sperm model. The reactions were then classified into two groups of those which show up-regulation or activation in comparison with the normal model fluxes and those which show the opposite. Reactions belonging to each category had been consolidated separately and then subjected to DAVID functional annotation clustering by DAVID [Huang et al. 2009], and ten of the most enriched clusters were selected for further analysis.

**Flux variability analysis (FVA)**

FVA is a tool for exploring alternative optimal solutions under certain thermodynamic and stoichiometric constraints [Bordbar et al. 2014]. It uses FBA-like linear programming problems to find the minimum and maximum flux of each reaction through the network while maintaining the objective function of the network at its optimal value [Mahadevan and Schilling 2003]. To evaluate the effect of knocking out each associated gene on the ATP hydrolysis rate, the gene was removed from the model and its reactions were silenced. The flux of ATPasel reaction was then fixed to its maximum value as determined by FBA. Finally, we performed FVA for the knockout model to determine the new maximal and minimal values of each reaction flux [Hadi and Marashi 2014]. Similar to the work of Hadi and Marashi [2014], we classified reactions into four groups by comparing the range of their flux in normal versus knockout states (see Figure 3). Furthermore, the DAVID functional annotation tool was used to find the positively influenced pathways in each knockout model [Huang et al. 2009].

**Acknowledgments**

The authors would like to thank Dr. K. Gilany (Avicenna Research Institute, ACECR, Iran) for his assistance in data collection and helpful discussions. The updated Recon2 model was kindly provided by Dr. Nathan Lewis (UCSD). S-AM acknowledges the financial support of the University of Tehran for this research under Grant Number 28791/1/2.

**Declaration of interest**

Financial support of the University of Tehran for this research under Grant Number 28791/1/2 (S-AM). The authors report no confliction of interest and none of the authors of the paper are employed by the Government of Iran.

**Notes on contributors**

Conceived and designed the experiments: S-AM, AA; Reconstructed the model, performed the computational experiments, and drafted the original report: AA; Involved in the interpretation of the findings and reviewing and revising the final manuscript: S-AM, AA, NAP.


Fukushima, T., Kaneoka, H., Yasuno, T., Sasaguri, Y., Tokuyasu, T., Tokoro, K., et al. (2013) Three novel references...


