

Systems Biology I: *iPACH22FFG*, A Genome-Scale Model For *Cutibacterium acnes* Within The Human Nose

Fehrenbach, Tobias; Frank, Simon; Ganske, Marcus

Submission on January 25, 2022
WS 21/22

Jun.-Prof. Dr. Andreas Draeger M.Sc. Nantia Leonidu

Dept. of Computer Science
Computational Systems
Biology of Infections

University Tuebingen

Abstract

Cutibacterium acnes is a commensal bacterium playing a major role in healthcare associated infections, but also in a healthy skin microbiome. There are indications that *C. acnes* also appears in the human nose environment. This project followed a pipeline to create an genome-scale model for *C. acnes*, *iPACH22FFG*, to enable further *in silico* research. In multiple steps the quality of the model was improved by increasing its consistency and annotate it properly. In a specialized medium for the human nose (SNM3), the model was able to grow. Sugar carbon sources increased its ability to grow. Thus, this project resulted in an working GEM for *C. acnes*.

Contents

Abstract	2
Contents	4
List of Figures	5
List of tables	5
1 Introduction	6
2 Results	8
2.1 Model generation, first analysis and evaluation	8
2.2 Model metabolites	9
2.2.1 Chemical formulae	9
2.2.2 Charges	9
2.3 Annotations	9
2.3.1 SBO-Terms	9
2.3.2 BiGG annotations	9
2.3.3 Annotations of other databases	10
2.4 Model improvements	10
2.4.1 Mass and Charge Imbalances	11
2.4.2 Growth on SMN3	11
2.4.3 Orphan and dead-end metabolites	11
2.4.4 KEGG Pathways	11
2.4.5 Groups plugin	11
2.5 Model visualization with Escher	11
2.5.1 Glycolysis and Pentose Phosphate Pathway	12
2.5.2 Citric acid (TCA) cycle	12
2.5.3 Pyruvate metabolism	13
2.5.4 Fructose and Mannose metabolism	14
2.5.5 Galactose metabolism	15
2.5.6 Inositol phosphate metabolism	16
2.6 Model analysis	17
2.6.1 Simulating growth for different carbon sources	17
2.6.2 Growth capabilities	17
3 Methods	19
3.1 Model generation	19
3.2 Model evaluation	19
3.3 Model metabolites	19
3.3.1 Chemical formulae	19
3.3.2 Charges	19
3.4 Annotations	19
3.4.1 SBO-Terms	19
3.4.2 BiGG annotations	20
3.4.3 Annotations of other databases	20
3.5 Model Improvements	20
3.5.1 Mass and charge imbalances	20
3.5.2 Growth on SMN3	20

3.5.3	Orphan and dead-end metabolites	20
3.5.4	KEGG Pathways	21
3.5.5	Groups plugin	21
3.6	Model visualization with Escher	21
3.6.1	Glycolysis and Pentose Phosphate Pathway	21
3.6.2	Citric acid (TCA) cycle	23
3.6.3	Pyruvate metabolism	23
3.6.4	Fructose and Mannose metabolism	23
3.6.5	Galactose metabolism	23
3.6.6	Inositol phosphate metabolism	23
3.7	Model analysis	23
3.7.1	Simulating growth for different carbon sources	23
3.7.2	Growth capabilities	23
4	Discussion	24
5	Appendix	25
5.1	Model generation, first analysis and evaluation	25
5.2	Model metabolites	25
5.3	Annotations	25
5.4	Model Improvements	25
5.5	Model visualization	26
5.6	Model analysis	26
6	References	27

List of Figures

1	Comparison of genes and proteins in the model vs the numbers stated by the NCBI database. The number of genes which are missing in the model added to the number of genes within the model equals the number of proteins without the hypothetical proteins.	8
2	Example of the added annotations to different databases for the metabolite 10fthf from the model SBML file.	10
3	Glycolysis and Pentose Phosphate pathway for HL096PA1 based on the KEGG pathways pach00010 [8] and pach00030 [10].	12
4	Map extension for HL096PA1 by the TCA cycle based on the KEGG pathway pach00020 [5].	13
5	Map extension for HL096PA1 by the Pyruvate metabolism based on the KEGG pathway pach00620 [11].	14
6	Map extension for HL096PA1 by the Fructose and Mannose metabolism based on the KEGG pathway pach00051 [6].	15
7	Map extension for HL096PA1 by the Galactose metabolism based on the KEGG pathway pach00052 [7].	16
8	Map extension for HL096PA1 by the Inositol phosphate metabolism based on the KEGG pathway pach00562 [9]	17
9	Binary heatmap to show growth on different carbon (sugars ending with -ose and Succinate, Acetate, Pyruvate) and nitrogen sources (Ammonium, Nitrite, Nitrate and Nitrogen). If the map is green for one metabolite, there is an increased increased growth rate after adding this metabolite to the medium	18
10	Glycolysis pathway for HL096PA1 as an example for a KEGG pathway (pach00010, [8]).	22

List of Tables

1	Reactions and metabolites added for the Inositol phosphate metabolism map.	16
---	--	----

1 Introduction

Nosocomial or healthcare-associated infections (HAI) are infections acquired during a medical intervention or an hospitalization. HAIs can be categorized in different types such as: (non-)CVC-associated bloodstream infections, Catheter-associated urinary tract infections, surgical site infections, (non-)ventilator-associated pneumonia and gastrointestinal infections. Due to limited data the consequences of HAI are not well proved but some studies indicate a mortality of 10% within 30-days in nosocomial infected patients. Other more commercial consequences are the increased length of hospitalization at around 12 days in Germany and increased healthcare cost associated with HAI of 7 billion Euro in the EU. However these numbers greatly vary on the location of infection and the causal pathogen [56].

While HAIs are also caused by fungi and viruses, this project aimed to create a genome-scale model for *Cutibacterium acnes*, a bacteria highly associated to HAIs and introduced in the following paragraphs.

The *Cutibacterium* genus refers to non-sporulating, gram-positive, slow-growing facultatively anaerobic bacteria which are usually seen as commensal skin bacteria and are therefore a healthy part of the skin microbiome [31][33]. Formerly, the *Cutibacterium*-genus belonged to the genus *Propionibacterium* but were reclassified in 2016 [55]. This was due to the fact that members of the genus *Cutibacterium* are mainly isolated from human skin, their G+C content ranges from 59-64 mol% and the diagnostic amino acid in peptidoglycan is *LL – A₂PM* [55]. The type species of the genus *Cutibacterium* is *C. acnes* which will discussed further on.

The most common appearance of *C. acnes*, despite other appearances in the human body (e.g. oral cavity, gastrointestinal and genitourinary tract), is the human skin [33]. While, at the skin surface *C. acnes* is only a small part (<2% of all bacteria) of the commensal microbiome in humans, more commonly *C. acnes* is found in the sebaceous follicles in contact with keratinocytes [31]. Within the sebaceous follicles it is mostly the only colonizing bacteria due to their metabolic ability to hydrolyse triglycerides. Those triglycerides are part of the sebum, the antimicrobial coating of the human skin [29]. The hydrolysis of triglycerides results in free fatty acids and therefore lower the pH of sebum, so that it can more efficiently inhibit pathogenic bacteria like *S. aureus* [33][29]. Therefore *C. acnes* is mostly seen as a commensal bacteria.

However upon entry to the body *C. acnes* can cause multiple diseases (e.g. endocarditis, septic arthritis, sarcoidosis) and is the causal agent of most post-surgical shoulder infections (51% of all infections) [24][33]. *C. acnes* is also associated to the name giving chronic skin disease Acne vulgaris, which is thought to be caused by an increase of sebum production going along with growth of the *C. acnes* population, which leads to inflammation [58]. More recent studies found out that some acne-associated phylotypes of *C. acnes* promote the disease by activating Th1 and Th17 responses [31]. Furthermore the ability to produce biofilms seems to be associated with Acne vulgaris, since the occurrence of *C. acnes* biofilms is higher in patients with Acne vulgaris (37%) than in patients without the disease (13%) [31]. Last but not least more and more stems of *C. acnes* are acquiring antibiotic resistances leading to difficulties in the treatment of all *C. acnes* associated infections [31].

The research on bacteria and their influence on the surrounding environment can often be difficult and expensive. A modern way to reduce costs are *in silico* experiments based on genome-scale models (GEMs). GEMs are computational models to predict phenotypes from genotypes and research metabolic networks, derived from systemically annotated

genome-scale networks based on knowledge [34]. In context of the above described antibiotic resistant microorganism GEMs can not only help to understand the antibiotic resistance but also could simulate the effectiveness of new antibiotics. Furthermore, since *C. acnes* is also present in the human nose microbiome, one could simulate their interactions and the impact on their virulence if parts of the microbiome are unstable due to genetic variations of the host or if the host is immunocompromised.

This project aims to create a GEM for *C. acnes* to set a base for further research.

2 Results

The model created and described in this project report is a genome-scale metabolic model (GEM) of *C. acnes* strain HL096PA1. Carey et al. proposed a naming convention for metabolic models [28]. Based on this convention the model is called *iPACH22FFG*, where PACH is the strain specific KEGG organism prefix [4], 22 the year of creation and FF the curators names. The final model has 90% as MEMOTE total score with 99.8% stoichiometric consistency, 97.3% mass balance, 98.3% charge balance, 100% metabolite connectivity and 90.9% unbounded flux in default medium, resulting in an overall consistency score of 97.9%. Metabolite annotations scored at 86.4%, reaction annotations at 84.4% and SBO annotations at 90.9%. The lowest score was the gene annotation score with 33.3%.

2.1 Model generation, first analysis and evaluation

The model was generated with CarveMe [47] as described in the Methods section 3.1. After creation some basic analysis on the model was done. At the point of creation the model contained 1759 reactions, 1230 metabolites and 769 genes with an optimized growth rate of $35.40 \frac{mmol}{gDW \cdot h}$. From the 769 genes were 36 essential for growth based on 1% of the optimized growth rate. The NCBI database stated that the accession HL096PA1 has a total of 2505 genes identified with 2332 proteins [1]. 373 of the proteins are stated as hypothetical, meaning there is no evidence that these regions in the DNA are really coding for a protein. Less the hypothetical proteins there should at least be 1959 proteins in a complete model. However if we can translate each gene in the model directly to a protein we are still missing 1190 proteins for a complete model. Figure 1 compares the genes and proteins within the model after creation against the genes and proteins stated by the NCBI database.

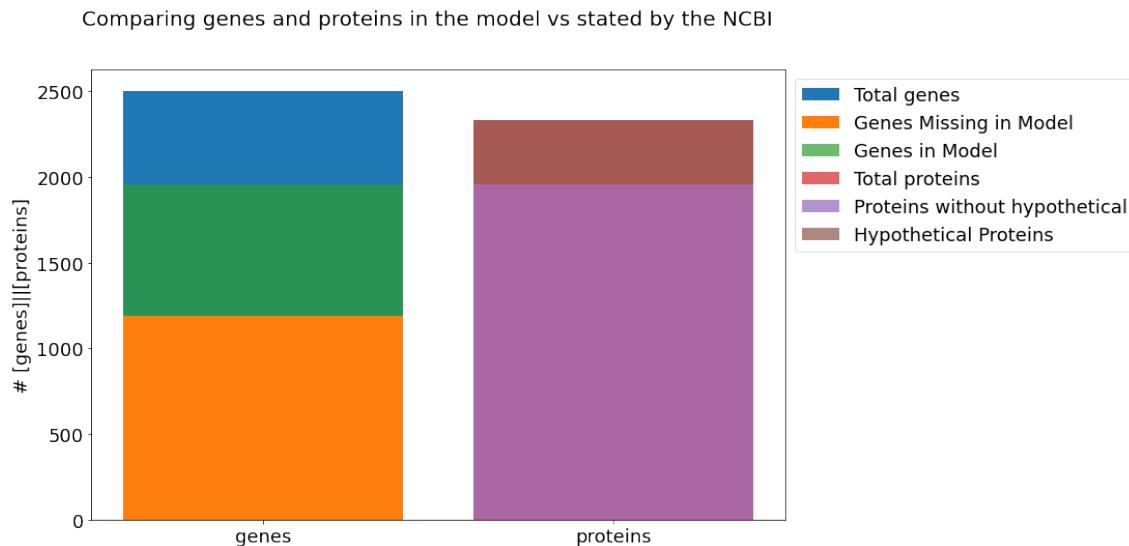


Figure 1: Comparison of genes and proteins in the model vs the numbers stated by the NCBI database. The number of genes which are missing in the model added to the number of genes within the model equals the number of proteins without the hypothetical proteins.

After the generation a benchmark of the model was created with Memote [46] as described in the Methods section 3.2. The first benchmark resulted in a total score of 38% with 87% consistency and no score for the annotations, except the uniform reaction and

metabolite identifier namespace, caused by the fact that all added metabolite and reaction names are based on the BiGG database used by CarveMe.

The appendix section 5.1 provides all files necessary for this part.

2.2 Model metabolites

The appendix section 5.2 provides all files necessary for this part.

2.2.1 Chemical formulae

Chemical formulae for the metabolites were added with CarveMe, which had no impact on the total score in the MEMOTE report.

2.2.2 Charges

The actual charge of a metabolite depends on its structural formulae and therefore its chemical property to absorb or emit ions (protons or electrons). However at this point there was no need to stretch the metabolite charge generation further, since a manual curation is a necessary step later. Thus, we have decided to take the first entry of the list of charges, whenever BiGG had multiple entries of charges for the corresponding metabolite. This led to some issues, for example “nad_c” should be charged negative (-1) but has no charge at all, in our result. The usual evaluation of the model with MEMOTE resulted in a slightly lower score of 35%, a decrease of 3% compared to the first benchmark (38%), due to the decreased consistency in the charge balance (from 100% to 48.9%). An explanation for this is that a neutral reaction can not be unbalanced and if there are no charges every reaction is neutral, therefore 100% of the reaction are charge balanced.

2.3 Annotations

The appendix section 5.3 provides all files necessary for this part.

2.3.1 SBO-Terms

Seven different SBO-Terms were added for seven groups of components. Genes were annotated with *SBO:0000243*, metabolites with *SBO:0000247*, metabolite reactions with *SBO:0000176*, exchange reactions with *SBO:0000627*, sink reactions with *SBO:0000632*, transport reactions with *SBO:0000185* and last but not least the growth reaction with *SBO:0000629*. Adding those SBO-terms, MEMOTE reports an Annotation - SBO Terms score of 91%. Demand reactions were skipped since there are none of this type in the model. The total score for the model is 77%, an increase by 37% compared to the 35% after adding the metabolite charges. This is a huge jump in the score, this indicates the importance of SBO-terms within systems biology.

2.3.2 BiGG annotations

An example for the added BiGG annotations is shown in figure 2, the specific BiGG entry is in line 51 within the annotation/rdf:Bag. For the BiGG annotations we did not encounter any difficulties.

2.3.3 Annotations of other databases

Annotations of other databases were done with ModelPolisher [54]. Here we encountered a slight problem for some annotations of the MetaNetX (MNX) [50] which probably is an error due to updated or outdated databases. For example, ModelPolisher adds a line `rdf:resource="MNXR98043"` for the reaction meta_R_3hoxpactex, which does not correspond to the usual "identifiers.org" format. When searching for the identifier in the MNX database, it states that the reaction ID "MNXR98043" is deprecated and replaced by MNXR94986" [23]. However the annotation is also added with the new reaction ID MNXR94986 in the usual format as:

```
rdf:resource="https://identifiers.org/resolve?query=metanetx.reaction:MNXR98043"
```

Since the problem raised some errors in the script for the MEMOTE report we removed the false identifiers with a python script from the model XML file.

After adding all annotations to the model, it scored a total of 85% in the MEMOTE report.

```
32      <species>
33          <metaid>meta_M_10fthf_c</metaid>
34          <sboTerm>SBO:0000247</sboTerm>
35          <id>M_10fthf_c</id>
36          <name>10-Formyltetrahydrofolate</name>
37          <compartment>c</compartment>
38          <hasOnlySubstanceUnits>false</hasOnlySubstanceUnits>
39          <boundaryCondition>false</boundaryCondition>
40          <constant>false</constant>
41          <fbc:charge>-2</fbc:charge>
42          <fbc:chemicalFormula>C20H21N7O7</fbc:chemicalFormula>
43
44      <notes>
45          <html xmlns="http://www.w3.org/1999/xhtml">
46              <p>FORMULA: C20H21N7O7</p>
47              <p>BioCyc: META:10-FORMYL-THF</p>
48              <p>SEED: Compound: cpd00201</p>
49              <p>UniPathway: Compound: UPC00234</p>
50              <p>KEGG: Compound: C00234</p>
51              <p>BioPath: Molecule: 10-Formyl-5,6,7,8-tetrahydrofolate</p>
52              <p>MetaNetX (MNX): Chemical: MNXM237</p>
53              <p>Reactome: 419151;5389850</p>
54              <p>Human Metabolome Database: HMDB00972</p>
55          </html>
56      </notes>
57      <annotation>
58          <rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#" xmlns:dcterms="http://purl.org/dc/terms/" xmlns:vCard="http://www.w3.org/2001/vcard-rdf/3.0#" xmlns:bqbiol="http://biomodels.net/biology-qualifiers/" xmlns:bqmodel="http://biomodels.net/model-qualifiers/">
59              <rdf:Description rdf:about="#meta_M_10fthf_c">
60                  <bqbiol:is>
61                      <rdf:Bag>
62                          <rdf:li rdf:resource="https://identifiers.org/bigg.metabolite/10fthf"/>
63                          <rdf:li rdf:resource="https://identifiers.org/biocyc/META:10-FORMYL-THF"/>
64                          <rdf:li rdf:resource="https://identifiers.org/chebi/CHEBI:11304"/>
65                          <rdf:li rdf:resource="https://identifiers.org/chebi/CHEBI:15637"/>
66                          <rdf:li rdf:resource="https://identifiers.org/chebi/CHEBI:19108"/>
67                          <rdf:li rdf:resource="https://identifiers.org/chebi/CHEBI:19109"/>
68                          <rdf:li rdf:resource="https://identifiers.org/chebi/CHEBI:57454"/>
69                          <rdf:li rdf:resource="https://identifiers.org/chebi/CHEBI:698"/>
70                          <rdf:li rdf:resource="https://identifiers.org/hmdb/HMDB00972"/>
71                          <rdf:li rdf:resource="https://identifiers.org/inchikey/AUFGTPPARQZWDO-YPMHNXCESA-L"/>
72                          <rdf:li rdf:resource="https://identifiers.org/kegg.compound/C00234"/>
73                          <rdf:li rdf:resource="https://identifiers.org/metanetx.chemical/MNXM237"/>
74                          <rdf:li rdf:resource="https://identifiers.org/reactome/R-ALL-419151"/>
75                          <rdf:li rdf:resource="https://identifiers.org/reactome/R-ALL-5389850"/>
76                          <rdf:li rdf:resource="https://identifiers.org/seed.compound/cpd00201"/>
77                  </rdf:Bag>
78              </bqbiol:is>
79          </rdf:Description>
80      </annotation>
81  </species>
```

Figure 2: Example of the added annotations to different databases for the metabolite 10fthf from the model SBML file.

2.4 Model improvements

Corresponding code, models and MEMOTE report files can be found in the appendix 5.4.

2.4.1 Mass and Charge Imbalances

The initial model had 647 mass, 35 charge and 21 mass and charge imbalanced reactions. After manual curation more than 89% of the unbalanced reactions were balanced. This increase is also confirmed by MEMOTE when looking at the mass and charge balance score. The mass balance score increased from 58.0% to 97.4%, and the charge balance score increased from 97.9% to 98.2%. The total score improved from 85% to 88%. Only 69 reactions are left unbalanced.

2.4.2 Growth on SMN3

C. acnes is known to grow in the human nose [33]. With the help of the chemically defined synthetic nasal medium (SNM3, [45]), which mimics the nasal habitat, the *in silico* growth rate of *Cutibacterium* was tested in this niche. In the initial medium there was no growth. Since Fe²⁺ could not be produced in a bacterium [36] it must be in the medium. However, there was still not growth. There are two amino acids (beta-alanine and l-Asparagine acid), which are not in the medium, but necessary for growth. Regarding BiGG there is one reaction in E.coli producing each metabolites [52]. Despite extensive literature research, including database searches on the KEGG database and BioCyc, no evidence was found that these reactions take place in the genus *Cutibacterium*. After adding Fe²⁺ to the medium and the two reactions to the medium, the model predicted a realistic growth rate of 0.44 $\frac{mmol}{gDW \cdot h}$ in the SNM3. Due to the change of the medium and lowering the amount of unbound fluxes, the MEMOTE score increased from 88% to 89%.

2.4.3 Orphan and dead-end metabolites

The current model had only one orphan metabolite pydam_c and no dead-end metabolites. Regarding KEGG there is a reaction in *C. acnes*, which produces this metabolites [21]. After this reaction was added there are no orphan and dead-end metabolites left in the model.

2.4.4 KEGG Pathways

The KEGG API was used to add all the pathways a reaction occurs in as a CV Term to the reaction. Reactions can be part of multiple pathways. For example the acetyl-CoA to Acetoacetyl-CoA reaction takes place in 15 pathways [12]. There was no change in the MEMOTE score.

2.4.5 Groups plugin

Since SBML Level 3 there is a groups plugin. Groups can be used as a collection of pathways. Each pathway was added as a group to the model. Altogether, 98 pathways were added as groups to model, with the plugin. All reactions, which are part of one pathway, were added to this group. There was no change in the MEMOTE score.

2.5 Model visualization with Escher

Corresponding code, models and map JSON files can be found in the appendix 5.5.

2.5.1 Glycolysis and Pentose Phosphate Pathway

For the Glycolysis and Pentose Phosphate Pathway all reactions marked by the KEGG pathways were already in the model. The visualization is shown in figure 3. In the upper right corner there are two reactions *DDGLK* and *EDA* which are not connected to the graph. This is probably caused by missing evidence for connecting reactions in the KEGG database. Therefore the connecting reactions are not colored green within the KEGG map and thus not added to the map for this model.

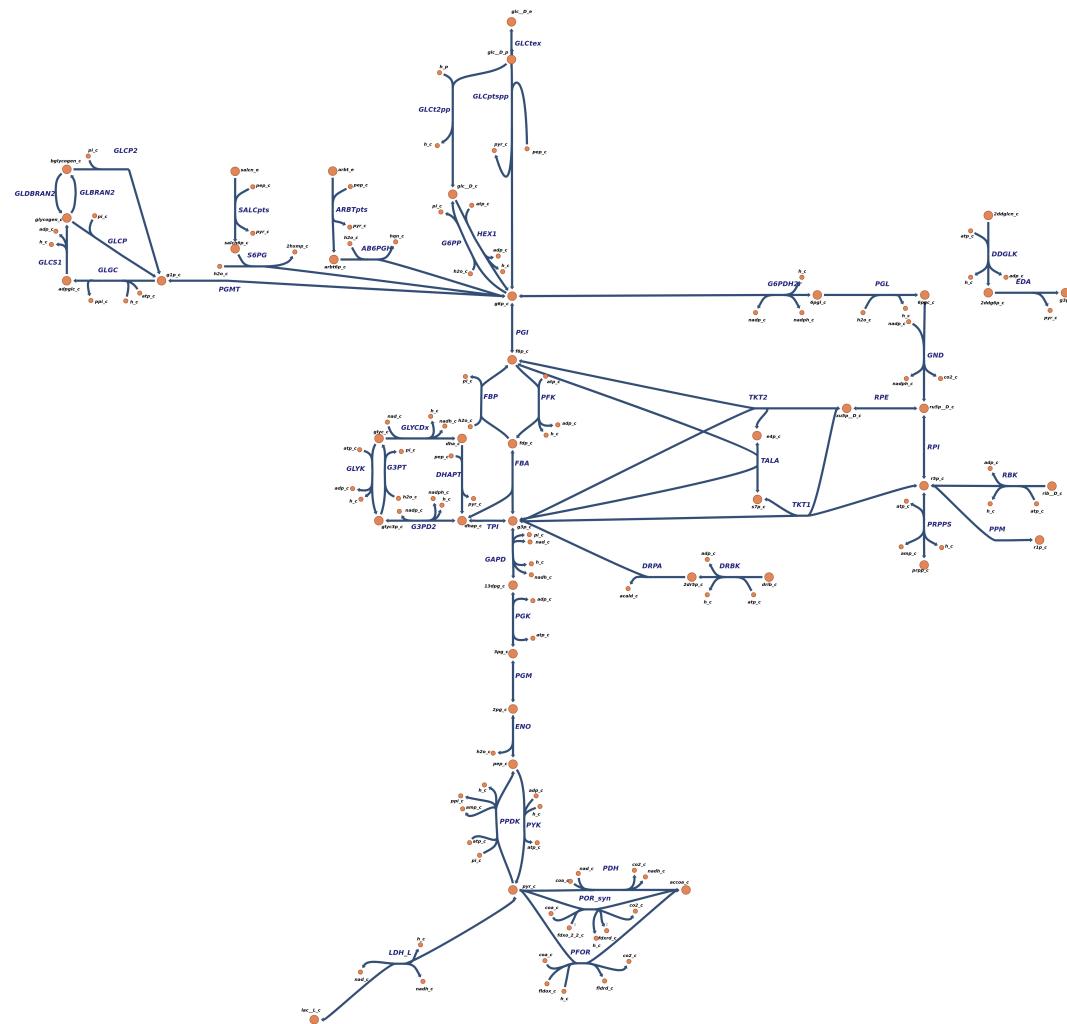


Figure 3: Glycolysis and Pentose Phosphate pathway for HL096PA1 based on the KEGG pathways pach00010 [8] and pach00030 [10].

2.5.2 Citric acid (TCA) cycle

The map from above was extended with the TCA cycle, again all corresponding reactions were already present in the model. The visualization is shown in figure 4.

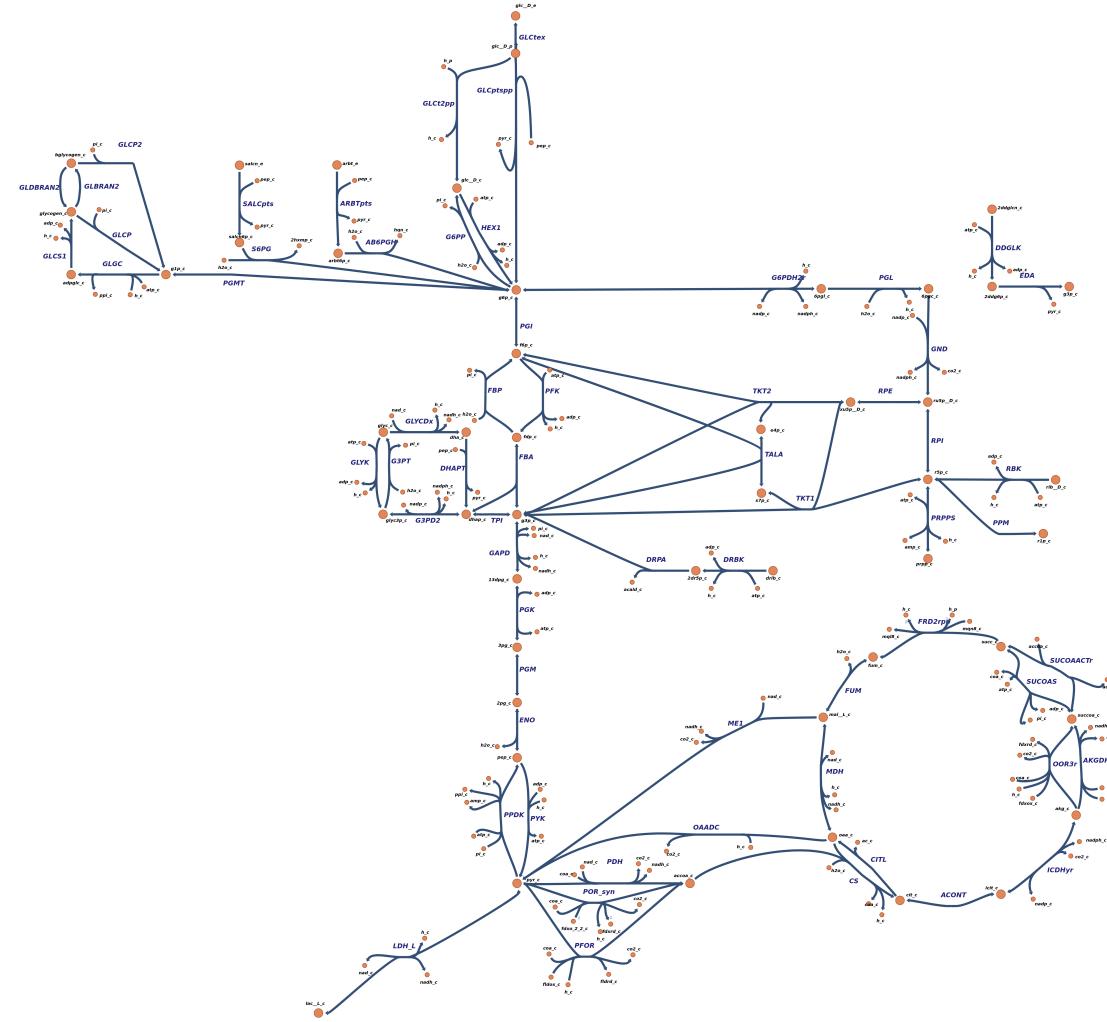


Figure 4: Map extension for HL096PA1 by the TCA cycle based on the KEGG pathway pach00020 [5].

2.5.3 Pyruvate metabolism

The model missed two reactions *DOACL* and *PDH_UBQ* to completely rebuild the map for the Pyruvate metabolism. These reactions were added to the model using a python script. The addition is based on the reference based genes *PAGK_2242* [22] for *DOACL* and *PAGK_0709* [20] for *PDH_UBQ* within the KEGG database. The map is shown in figure 5.

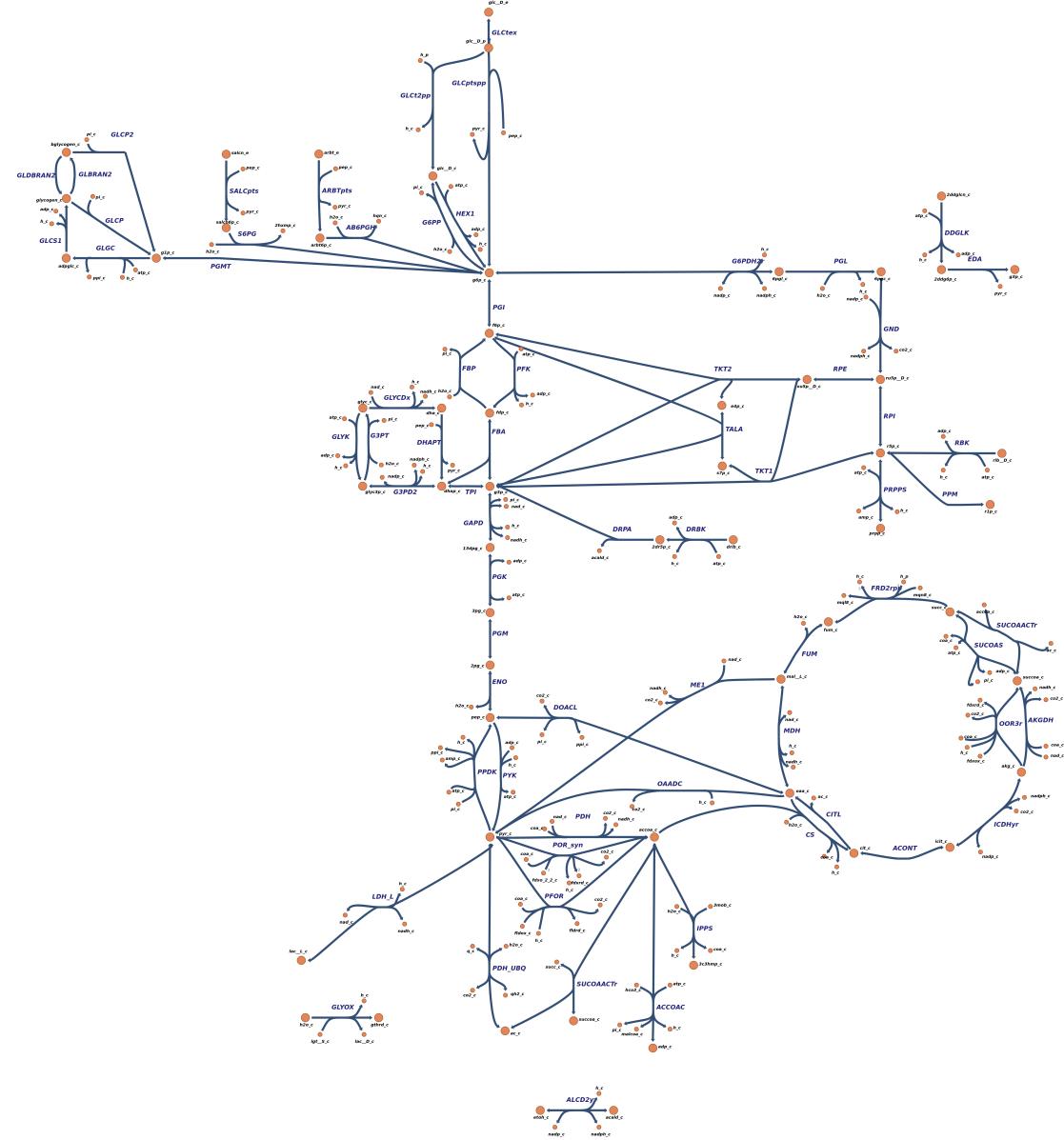


Figure 5: Map extension for HL096PA1 by the Pyruvate metabolism based on the KEGG pathway pach00620 [11].

2.5.4 Fructose and Mannose metabolism

For the Fructose and Mannose metabolism the model missed two reactions and were added again based on the KEGG database. The two reactions were:

MNL1PP based on the gene *PAGK_0027* [14] and

DFP based on the gene *PAGK_1063* [15].

Figure 6 shows the map.

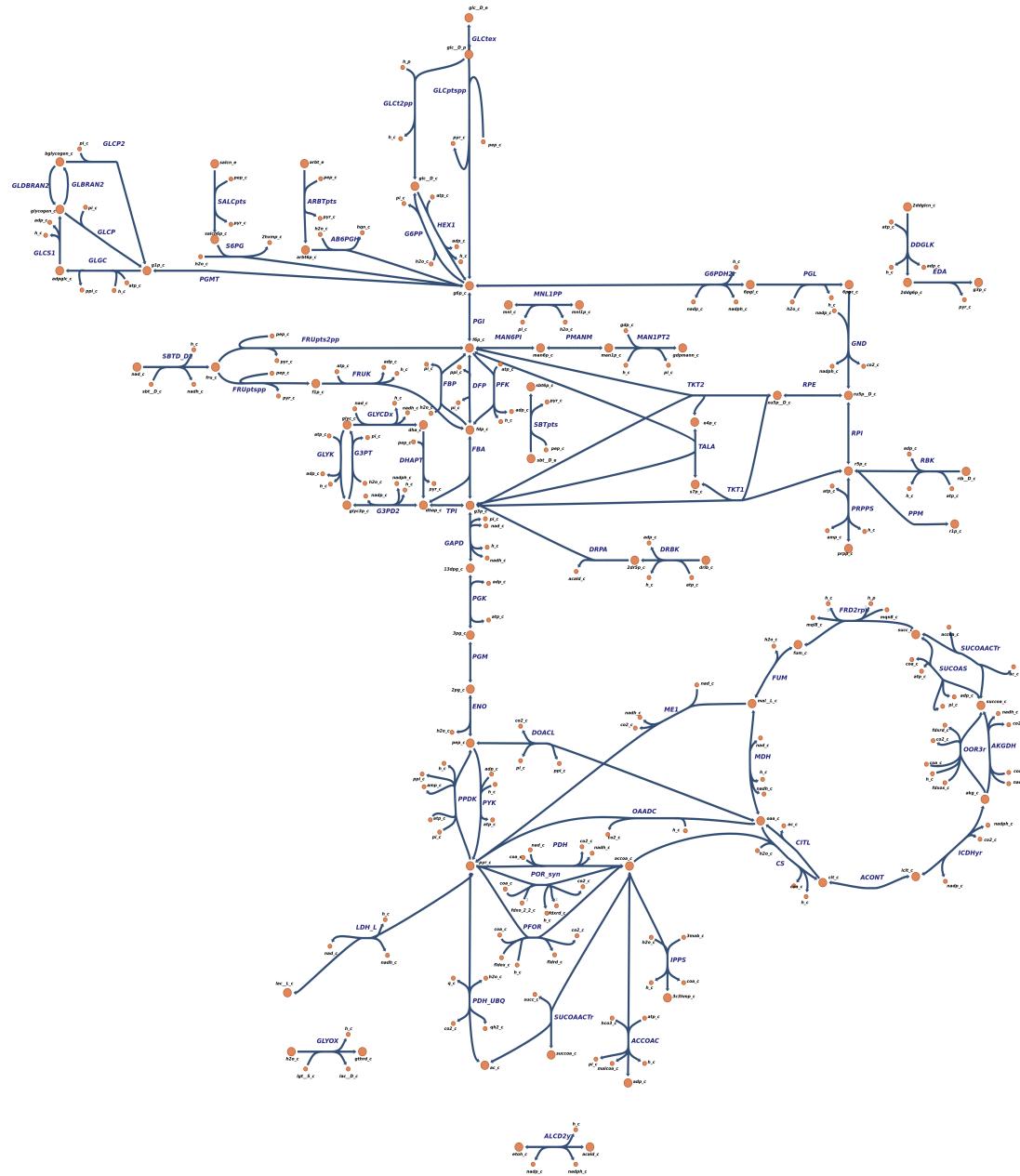


Figure 6: Map extension for HL096PA1 by the Fructose and Mannose metabolism based on the KEGG pathway pach00051 [6].

2.5.5 Galactose metabolism

For the Galactose metabolism no reactions were missing in the model. The map is displayed in figure 7.

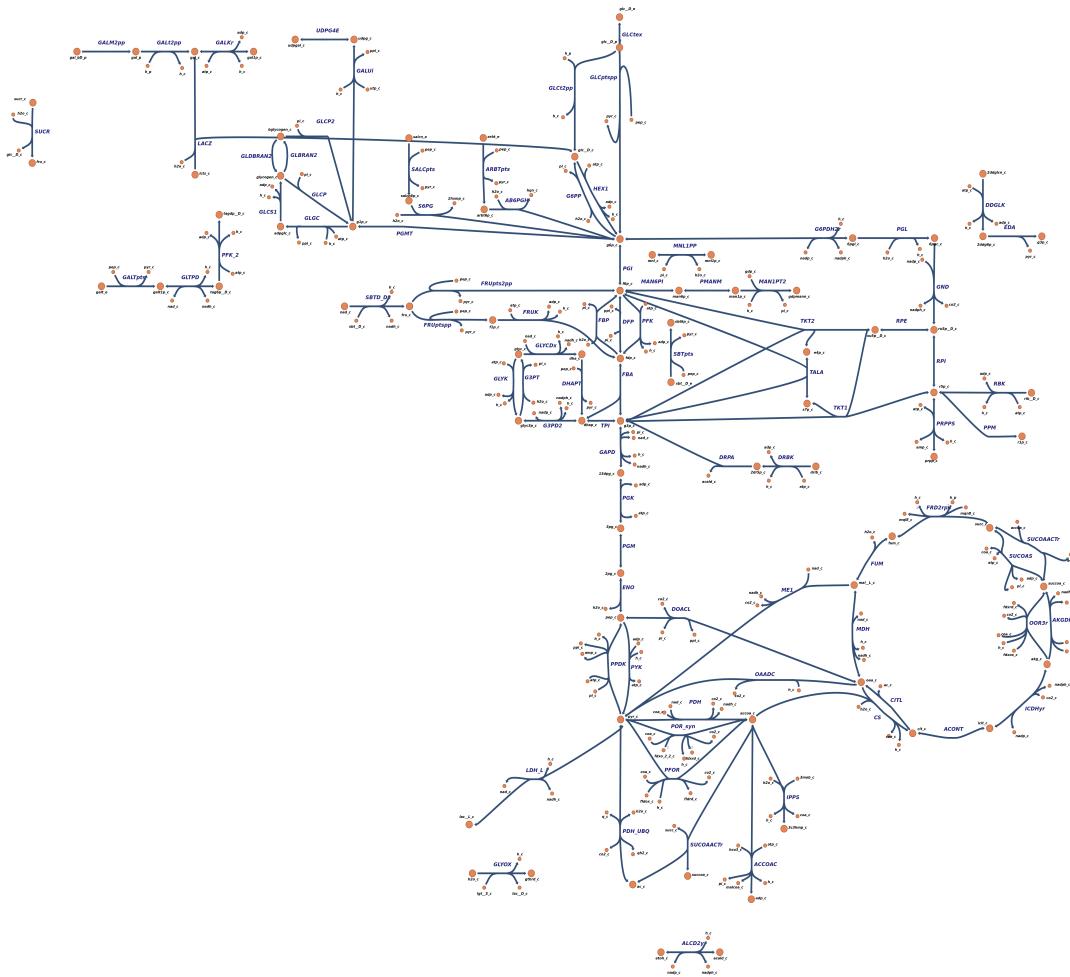


Figure 7: Map extension for HL096PA1 by the Galactose metabolism based on the KEGG pathway pach00052 [7].

2.5.6 Inositol phosphate metabolism

The Inositol Phosphate metabolism in this model lacked four reactions and corresponding metabolites. Reactions and metabolites added are based on the KEGG database (table 1). Figure 8 shows the final map containing all added pathways.

Table 1: Reactions and metabolites added for the Inositol phosphate metabolism map.

Reaction to add	Metabolites to add	Gene	KEGG Reference
<i>PHCHHL</i>	<i>thch_c</i>	<i>PAGK_0485</i>	[19]
<i>THCAH</i>	<i>ddgluc_c</i>	<i>PAGK_0476</i>	[18]
<i>5DG</i>	<i>dkh_c, dkhp_c</i>	<i>PAGK_0475</i>	[17]
<i>5DGP</i>		<i>PAGK_0473</i>	[16]

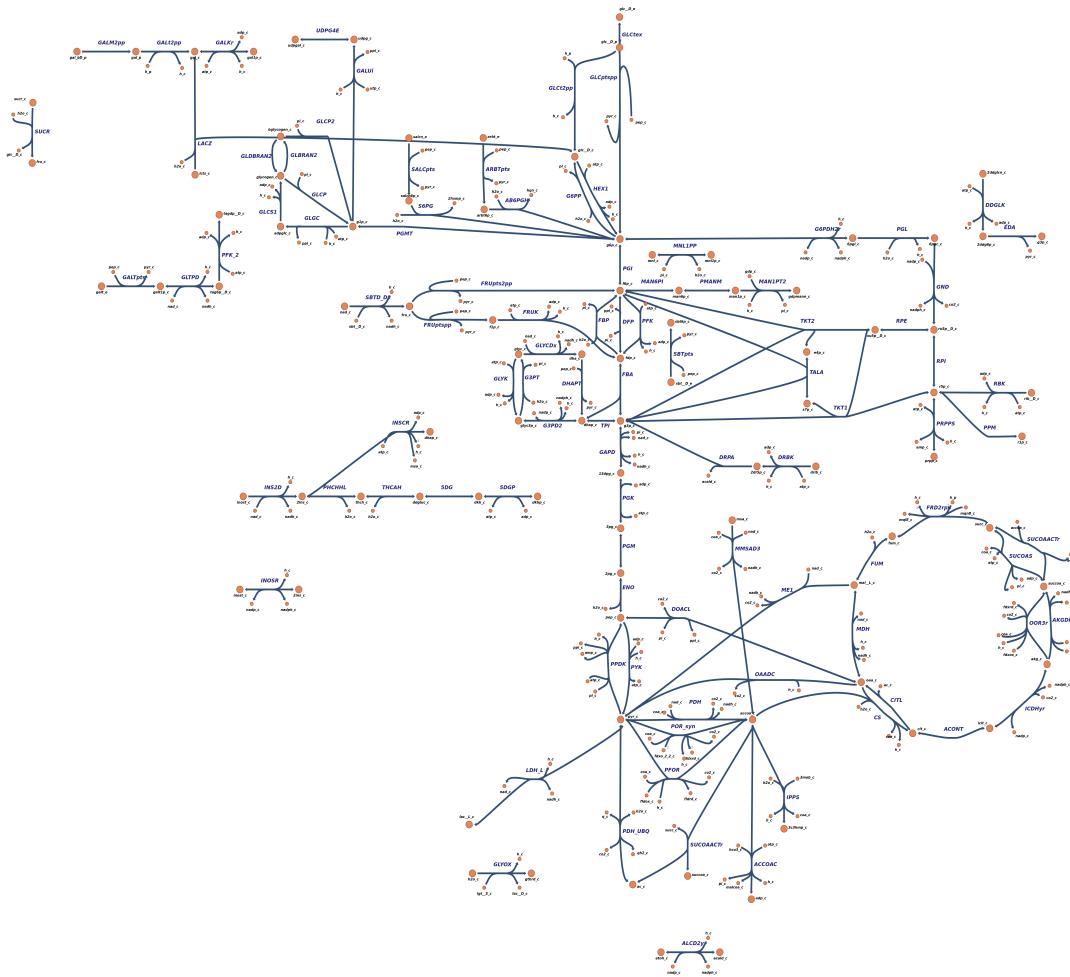


Figure 8: Map extension for HL096PA1 by the Inositol phosphate metabolism based on the KEGG pathway pach00562 [9]

2.6 Model analysis

Corresponding code, model and final MEMOTE report files can be found in the appendix 5.6.

2.6.1 Simulating growth for different carbon sources

To check if *C. acnes* growths in different carbon sources, the growth rate of *C. acnes* in the minimal medium was compared with the growth rate in the same medium with one additional carbon source. The growth rate in the minimal medium is $10.0 \frac{mmol}{gDW \cdot h}$. The growth rate increased to $12.6 \frac{mmol}{gDW \cdot h}$, if there was the carbon source glucose or fructose in the medium. There was no difference in the growth rate after adding succinate, acetate or pyruvate as carbon source to the medium.

2.6.2 Growth capabilities

Figure 9 shows the growth capabilities in different nutrient environments. The minimal medium was again used as comparison. For all sugars (ending with -ose) as carbon sources

there is an increased growth rate, if there were added to the minimal medium. Other carbon sources, such as Succinate, Acetate and Pyruvate, did not change the growth rate. No difference in the growth rate is observed if there are additional nitrogen sources (Ammonium, Nitrite, Nitrate or Nitrogen) in the environment (minimal medium).

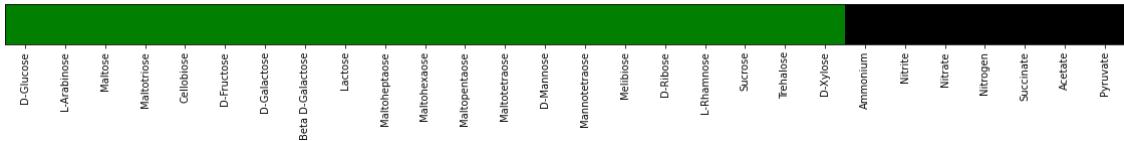


Figure 9: Binary heatmap to show growth on different carbon (sugars ending with -ose and Succinate, Acetate, Pyruvate) and nitrogen sources (Ammonium, Nitrite, Nitrate and Nitrogen). If the map is green for one metabolite, there is an increased increased growth rate after adding this metabolite to the medium

3 Methods

3.1 Model generation

CarveMe (V: 1.4.0) is a tool to reconstruct genome-scale metabolic models for microbes in a fast and automated manner [47]. CarveMe depends on two additional programs, Diamond (V: 0.9.14, [27]) and CPlex (V: 20.1.0, [2]), whereas the latter needs python (V: 3.8) to work properly.

The *C. acnes* model described here is based on the accession HL096PA1 and was created from the protein FastA file provided by the NCBI genome website [1]. With the console command:

```
$ carve GCF_000376705.1_ASM37670v1_protein.faa
```

Despite two other files, CarveMe creates a model XML file in an SBML format which is called:

```
GCF_000376705.1_ASM37670v1_protein.xml
```

3.2 Model evaluation

MEMOTE (V: 0.11.1, [46]) is a tool which produces an independent and comparable score for GEMs. The report produced by MEMOTE is an HTML file with two columns. The left column reports the score, while the right column reports an overview of the model. The score is calculated from the consistency (stoichiometric, mass balance, charge balance, metabolite connectivity and unbound fluxes) as well as the completeness of annotations for metabolites, reactions, genes and SBO-terms. For each improvement step the model was evaluated with a short python script calling MEMOTE.

3.3 Model metabolites

3.3.1 Chemical formulae

With the `-fbc2` flag set, CarveMe transfers all existing chemical formulae from the notes field to the species description. Therefore the model was regenerated from the protein FastA as stated above but with the `-fbc2` flag set.

3.3.2 Charges

In order to add charges for all metabolites, the metabolite ids had to be extracted from the model. Afterwards, the charges were downloaded via the BiGG api [3][43] and added to the model with COBRApy [32]. COBRApy is a python package for COnstraint Based Reconstruction and Analysis methods. It is used to easily manipulate the big model SBML files and for analyzing them for example in flux-balance analysis (FBA).

3.4 Annotations

3.4.1 SBO-Terms

To identify and interpret model components unambiguously the annotation with Systems Biology Ontology (SBO) terms is useful. SBO-terms allow a better understanding of the components [30]. Each MEMOTE report gives multiple lists of model components which are not annotated with an SBO-term. Additionally the report provides an SBO-term for each list. These lists and their corresponding SBO-terms were extracted from the

MEMOTE report and added to the model via libSBML [26]. LibSBML is a user-friendly tool to read, write, manipulate, translate, and validate SBML files.

3.4.2 BiGG annotations

The BiGG annotations for metabolites and reactions were added manually with the libSBML [30] package. This was done via extracting the IDs from the model and set a new CVTerm with the qualifier type "BIOLOGICAL_Qualifier" and "BQB_IS". Core of these annotations is the "<https://identifiers.org/database/component>" link added as a resource to the CVTerm. Identifiers.org is a web service which redirects the user to the corresponding database and reaction [25]. For example the link <https://identifiers.org/bigg.metabolite/10fthf> redirects the user to the BiGG database entry of *10fthf*.

3.4.3 Annotations of other databases

Annotations of other databases (e.g. BioCyc [41], CHEBI [35], MetaNetX (MNX) [50], KEGG [40]) were added with the ModelPolisher tool [54]. ModelPolisher needs a docker [48] container containing the corresponding databases. After establishing the container the tool can be called via the command line. The following command was used:

```
$ docker-compose run -u 1000 -v /path/to/working_directory polisher java -jar /ModelPolisher-2.1-beta.jar --input=/path/to/model.xml --output=/path/to/output_model.xml --annotate-with-bigg=true --add-adb-annotations=true --output-combine=true
```

Which returns a zip file containing the fully annotated model.

3.5 Model Improvements

After the automated creation of the model, manual refinement steps followed. They were done using CobraPy [32] and libSBML [26][37].

3.5.1 Mass and charge imbalances

The first model had 95 metabolites without a formula. The REST-API from the BiGG database [43][3] was used to add the formulae for this metabolites. For metabolites with more than one possible formula or charge regarding the BiGG Database always the tuple was chosen, which was causing the lowest mass and charge imbalances for the whole model.

3.5.2 Growth on SNM3

To validate the model further, the growth capabilities in the human nose was simulated. For this, there is a chemically defined synthetic nasal medium (SNM3) available [45]. To test the growth rate, the lowerbound of all exchange reactions for the metabolites, which are in the medium, was set to $-10 \frac{mmol}{gDW \cdot h}$. The uptake rate for oxygen was set to $20 \frac{mmol}{gDW \cdot h}$. To simulate the growth rate, the biomass objective function was maximized using COBRApy [32]. As no growth could be simulated with the model in the medium, the minimal medium calculated by COBRApy was compared with SNM3.

3.5.3 Orphan and dead-end metabolites

To resolve the orphan metabolite pyridoxamine, a reaction producing pyridoxamine was added to the model. The pyridoxamine kinase produces and consumes pyridoxamine [13]. Regarding KEGG [21] this reaction takes also place in *Cutibacterium*. The reaction and gene reaction rule was added to the model.

3.5.4 KEGG Pathways

To add the KEGG Pathways one reaction occurs in, the KEGG API [40] was used. The already annotated KEGG ID for every reaction was used to get pathway id from the KEGG API. The extracted pathway ids were added to the reaction as a CV term with the biological qualifier type BQB_OCCURS_IN using libSBML [26].

3.5.5 Groups plugin

The groups plugin [38] from libSBML[26] was enabled. A list of all the pathways from *Cutibacterium* was created with the KEGG API [40]. Each pathway was added to the model as a group. Every reaction occurring in the pathway was added to the group. Therefore the previously added pathways for each reaction was used.

3.6 Model visualization with Escher

The pathway maps were created with the Escher Python package (V: 1.7.3, [42]) within a Jupyter Notebook [44]. The maps are based upon the 'iJ01366.Central metabolism' map file, an example file natively provided by Escher. Each map was compared with and adapted to the corresponding KEGG pathway for the organism *C. acnes* accession HL096PA1 (KEGG ID: *PACH*, [4]). The KEGG pathway shows reactions belonging to the organism in green. If a reaction was missing, the reaction and corresponding metabolites were added to the model with a python script using COBRApy [32]. Those added reactions were named by the authors since these reactions had no corresponding BiGG entry. Reactions not colored green in the KEGG pathway were removed from the map. Starting with the Glycolysis and Pentose Phosphate pathways the maps were consecutively extended by the following pathways.

3.6.1 Glycolysis and Pentose Phosphate Pathway

The maps for Glycolysis and Pentose Phosphate Pathway are based on the KEGG pathway maps pach00010 [8] and pach00030 [10]. As an example the Glycolysis KEGG pathway pach00010 is shown in figure 10.

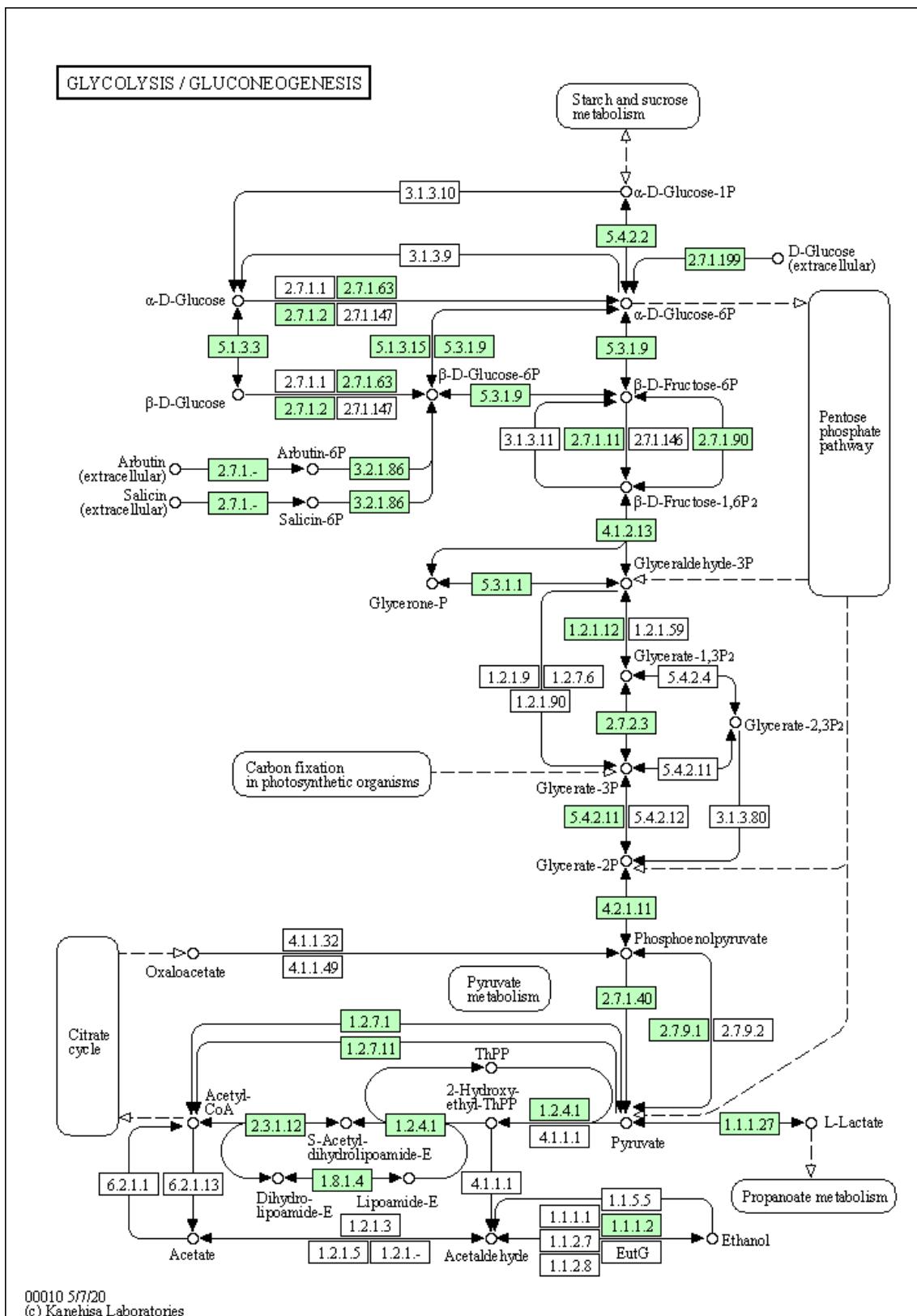


Figure 10: Glycolysis pathway for HL096PA1 as an example for a KEGG pathway (pach00010, [8]).

3.6.2 Citric acid (TCA) cycle

The map for the TCA pathway is based on the KEGG pathway map pach00020 [5].

3.6.3 Pyruvate metabolism

The map for the Pyruvate metabolism is based on the KEGG pathway map pach00620 [11].

3.6.4 Fructose and Mannose metabolism

The map for the Fructose and Mannose metabolism is based on the KEGG pathway map pach00051 [6].

3.6.5 Galactose metabolism

The map for the Galactose metabolism is based on the KEGG pathway map pach00052 [7].

3.6.6 Inositol phosphate metabolism

The map for the Inositol Phosphate metabolism is based on the KEGG pathway map pach00562 [9].

3.7 Model analysis

3.7.1 Simulating growth for different carbon sources

The minimal medium for the model returned by COBRApy [32] was used to examine the growth on different carbon sources. Glycogen was also part of the minimal medium. To test the growth on different carbon sources, there should not be any other carbon sources in the medium. Therefore Glycogen was removed from the minimal medium. The uptake rate (lowerbound) for each metabolite in the minimal medium was set to $10 \frac{mmol}{gDW \cdot h}$. Each carbon source was tested individually by only setting the tested carbon source's uptake rate to $10 \frac{mmol}{gDW \cdot h}$ and optimizing the model for growth.

3.7.2 Growth capabilities

With the previously used minimal medium the growth of the model on all possible carbon sources was tested. Therefore the uptake rate for all possible carbon sources was individually set to $10 \frac{mmol}{gDW \cdot h}$. To test the growth on different nitrogen sources ammonium, nitrate, nitrite, and nitrogen was individually added to the minimal medium. For all 28 different media the growth for the model was optimized using COBRApy [32]. If there was an increase in growth rate from the default minimal medium. The metabolite was marked green in the binary heatmap otherwise black. The heatmap was build with the python package MatPlotLib(V: 3.5.1, [39])

4 Discussion

In this work we described the creation of a genome-scale model (GEM) for *C. acnes* iPACH22FFG. Regarding the final MEMOTE report a total score of 90% was achieved, which is quite high and therefore is an indication of its quality. In comparison the GEM for *Dolosigranulum pigrum* iDPM21RW scored at 86% [53], while the first GEM for *Escherichia coli* iML1515 scored at 91% [49]. However there are still some things which can be improved. For example, the gene annotations scored only at 33.3%, which was the lowest score contribution to the total MEMOTE score. These could and should be annotated with different databases like the NCBI refseq [51] database, UniProt [57] or KeGG [40]. Also, other annotation categories could probably be improved but is a difficult task due to basic differences in different databases. Such differences include for example different abbreviations for metabolites, which makes a cross-reference very hard. A problem BiGG [43] tries to tackle and thus could be the big solution for the future.

Regarding the human nose environment, our *in silico* model was able to grow. However to achieve the growth, two reactions for the creation of beta-Alanine and l-Asparagine needed to be added to the model but this could only be done based on similar reactions in *E. coli* [52]. Now, a prove is needed that both reactions happen in *C. acnes*. This only can be done in the laboratory. Another possibility could be that the creation of beta-Alanine and l-Asparagine can be replaced by other reactions creating a necessary metabolite which was previously created from beta-Alanine and l-Asparagine. A third possibility would be that both amino acids are produced by other bacteria common within the human nose, so that *C. acnes* can only grow within a microbial community. To research this *in silico*, we could create a model containing multiple GEMs of different bacteria inherent to the human nose.

We conclude that the project model of *C. acnes* iPACH22FFG set a base for further researches.

5 Appendix

5.1 Model generation, first analysis and evaluation

Code, Protein FastA for model creation, model and Memote report are available in the *appendix*, sub-folder: *5-1_ModelGen_FirstAn_Eval*.

Code: A03_Ganske_Fehrenbach.ipynb
Protein FastA: GCF_000376705.1_ASM37670v1_protein.faa
Model: GCF_000376705.1_ASM37670v1_protein.xml
Memote report: Memote_first_aftercreation.html

5.2 Model metabolites

Code, Protein FastA for model creation, model and MEMOTE report are available in the *appendix*, sub-folder: *5-2_Model_Metabolites*.

Code: A04_Ganske_Fehrenbach.ipynb
Protein FastA: GCF_000376705.1_ASM37670v1_protein.faa
Model: GCF_000376705.1_ASM37670v1_protein.xml
Memote report: Memote_formulae_charges.html

5.3 Annotations

Code, model and MEMOTE report are available in the *appendix*, sub-folder: *5-3_Annotations* and the corresponding subsub-folder.

SBO
Code: A05_Ganske_Fehrenbach.ipynb
Model: cutie_cuti_a04_beforeSBO.xml, cutie_cuti_a05.xml
Memote report: MemoteResult_SBO.html

BiGG_otherDB
Code: A06_Problem12_Ganske_Frank_Fehrenbach.ipynb
Model: cutie_cuti_v6_BiGGterms.xml, cutie_cuti_v6_simon.xml,
cutie_cuti_v07.xml, cutie_cuti_v08_TESThand.xml
Memote report: memote_report_v7.html

5.4 Model Improvements

Code, model, MEMOTE reports and Medium files are available in the *appendix*, sub-folder: *5-4_Model_Improvements* and the corresponding subsub-folder.

Mass_Charge_Balance (Problem 1A)
Code: Problem1A_Ganske_Frank_Fehrenbach.ipynb
Model: cutie_cuti_v7.xml, cutie_cuti_v8_curated.xml
Memote report: MemoteResult_Curated.html

Problem_1B-E
Code: Problem1B-E_Ganske_Frank_Fehrenbach.ipynb
Model: cutie_cuti_v9_snm3.xml, cutie_cuti_v10.xml,
cutie_cuti_v11.xml, cutie_cuti_v12.xml
Memote report: MemoteResult_1B.html, MemoteResult_1C.html,
MemoteResult_1D.html, MemoteResult_1E.html

Medium: SNM3.csv

5.5 Model visualization

Code, model and map JSONs are available in the *appendix*, sub-folder: *5-5_Model_Visualization*.

Code: `DEV4_Ganske_Frank_Fehrenbach.ipynb`
Model: `cutie_cuti_v12.xml`, `cutie_cuti_v13.xml`
Map JSONs: `/maps/`

5.6 Model analysis

Code, model and MEMOTE report are available in the *appendix*, sub-folder: *5-6_Model_Analysis*.

Code: `Problem3.ipynb`
Model: `cutie_cuti_v13.xml`
Memote report: `final_Memote.html`

6 References

- [1] Cutibacterium acnes (ID 1140) - Genome - NCBI, Nov 2021. [Online; accessed 9. Nov. 2021] <https://www.ncbi.nlm.nih.gov/genome/?term=HL096PA1>.
- [2] Introducing IBM ILOG CPLEX Optimization Studio 20.1.0, Nov 2021. [Online; accessed 22. Jan. 2022] <https://www.ibm.com/docs/en/icos/20.1.0?topic=documentation-introducing-ilog-cplex-optimization-studio-2010>.
- [3] BiGG Web API, Jan 2022. [Online; accessed 22. Jan. 2022], http://bigg.ucsd.edu/data_access.
- [4] KEGG GENOME: Cutibacterium acnes HL096PA1, Jan 2022. [Online; accessed 22. Jan. 2022] https://www.genome.jp/dbget-bin/www_bget?gn:pach.
- [5] KEGG PATHWAY: Citrate cycle (TCA cycle) - Cutibacterium acnes HL096PA1, Jan 2022. [Online; accessed 23. Jan. 2022] <https://www.genome.jp/pathway/pach00020>.
- [6] KEGG PATHWAY: Fructose and mannose metabolism - Cutibacterium acnes HL096PA1, Jan 2022. [Online; accessed 23. Jan. 2022] <https://www.genome.jp/pathway/pach00051>.
- [7] KEGG PATHWAY: Galactose metabolism - Cutibacterium acnes HL096PA1, Jan 2022. [Online; accessed 23. Jan. 2022] <https://www.genome.jp/pathway/pach00052>.
- [8] KEGG PATHWAY: Glycolysis / Gluconeogenesis - Cutibacterium acnes HL096PA1, Jan 2022. [Online; accessed 23. Jan. 2022] <https://www.genome.jp/pathway/pach00010>.
- [9] KEGG PATHWAY: Inositol phosphate metabolism - Cutibacterium acnes HL096PA1, Jan 2022. [Online; accessed 23. Jan. 2022] <https://www.genome.jp/pathway/pach00562>.
- [10] KEGG PATHWAY: Pentose phosphate pathway - Cutibacterium acnes HL096PA1, Jan 2022. [Online; accessed 23. Jan. 2022] <https://www.genome.jp/pathway/pach00030>.
- [11] KEGG PATHWAY: Pyruvate metabolism - Cutibacterium acnes HL096PA1, Jan 2022. [Online; accessed 23. Jan. 2022] <https://www.genome.jp/pathway/pach00620>.
- [12] KEGG REACTION: R00238, Jan 2022. [Online; accessed 24. Jan. 2022] <https://www.kegg.jp/entry/R00238>.
- [13] KEGG REACTION: R02493, Jan 2022. [Online; accessed 24. Jan. 2022] <https://www.kegg.jp/entry/R02493>.
- [14] KEGG T02638: PAGK_0027, Jan 2022. [Online; accessed 23. Jan. 2022] https://www.genome.jp/entry/pach:PAGK_0027.
- [15] KEGG T02638: PAGK_0090, Jan 2022. [Online; accessed 23. Jan. 2022] https://www.genome.jp/entry/pach:PAGK_0090+pach:PAGK_1063.
- [16] KEGG T02638: PAGK_0473, Jan 2022. [Online; accessed 23. Jan. 2022] https://www.genome.jp/entry/pach:PAGK_0473.

- [17] KEGG T02638: PAGK_0475, Jan 2022. [Online; accessed 23. Jan. 2022] https://www.genome.jp/entry/pach:PAGK_0475.
- [18] KEGG T02638: PAGK_0476, Jan 2022. [Online; accessed 23. Jan. 2022] https://www.genome.jp/entry/pach:PAGK_0476.
- [19] KEGG T02638: PAGK_0485, Jan 2022. [Online; accessed 23. Jan. 2022] https://www.genome.jp/entry/pach:PAGK_0485.
- [20] KEGG T02638: PAGK_0709, Jan 2022. [Online; accessed 23. Jan. 2022] https://www.genome.jp/entry/pach:PAGK_0709.
- [21] KEGG T02638: PAGK_1642, Jan 2022. [Online; accessed 24. Jan. 2022] https://www.genome.jp/entry/pach:PAGK_1642.
- [22] KEGG T02638: PAGK_2242, Jan 2022. [Online; accessed 23. Jan. 2022] https://www.genome.jp/entry/pach:PAGK_2242.
- [23] MetaNetX: MNXR98043 - MNXR94986, Jan 2022. [Online; accessed 23. Jan. 2022] https://www.metanetx.org/equa_info/MNXR98043.
- [24] George S. Athwal, John W. Sperling, Damian M. Rispoli, and Robert H. Cofield. Deep infection after rotator cuff repair. *J. Shoulder Elbow Surg.*, 16(3):306–311, May 2007.
- [25] Manuel Bernal-Llinares, Javier Ferrer-Gómez, Nick Juty, Carole Goble, Sarala M. Wimalaratne, and Henning Hermjakob. Identifiers.org: Compact Identifier services in the cloud. *Bioinformatics*, 37(12):1781–1782, Jul 2021.
- [26] Benjamin J. Bornstein, Sarah M. Keating, Akiya Jouraku, and Michael Hucka. Lib-SBML: An API library for SBML. *Bioinformatics*, 24(6), 2008.
- [27] Benjamin Buchfink, Chao Xie, and Daniel H. Huson. Fast and sensitive protein alignment using DIAMOND - Nature Methods. *Nat. Methods*, 12:59–60, Jan 2015.
- [28] Maureen A. Carey, Andreas Dräger, Moritz E. Beber, Jason A. Papin, and James T. Yorkovich. Community standards to facilitate development and address challenges in metabolic modeling. *Mol. Syst. Biol.*, 16(8), Aug 2020.
- [29] G. J. M. Christensen and H. Brüggemann. Bacterial skin commensals and their role as host guardians. *Beneficial Microbes*, Dec 2013.
- [30] Mélanie Courtot, Nick Juty, Christian Knüpfer, Dagmar Waltemath, Anna Zhukova, Andreas Dräger, Michel Dumontier, Andrew Finney, Martin Golebiewski, Janna Hastings, et al. Controlled vocabularies and semantics in systems biology. *Mol. Syst. Biol.*, 7:543., Oct 2011.
- [31] B. Dréno, S. Pécastaings, S. Corvec, S. Veraldi, A. Khammari, and C. Roques. Cutibacterium acnes (Propionibacterium acnes) and acne vulgaris: a brief look at the latest updates. *J. Eur. Acad. Dermatol. Venereol.*, 32(S2):5–14, Jun 2018.
- [32] Ali Ebrahim, Joshua A. Lerman, Bernhard O. Palsson, and Daniel R. Hyduke. CO-BRApy: COnstraints-Based Reconstruction and Analysis for Python. *BMC Syst. Biol.*, 7(1):1–6, Dec 2013.

- [33] Marlee J. Elston, John P. Dupaix, Maria I. Opanova, and Robert E. Atkinson. *Cutibacterium acnes* (formerly *Propionibacterium acnes*) and Shoulder Surgery. *Hawaii. J. Health Soc. Welf.*, 78(11 Suppl 2):3, Nov 2019.
- [34] Xin Fang, Colton J. Lloyd, and Bernhard O. Palsson. Reconstructing organisms in silico: genome-scale models and their emerging applications - Nature Reviews Microbiology. *Nat. Rev. Microbiol.*, 18:731–743, Dec 2020.
- [35] J. Hastings, G. Owen, A. Dekker, M. Ennis, N. Kale, V. Muthukrishnan, S. Turner, N. Swainston, P. Mendes, and C. Steinbeck. ChEBI in 2016: Improved services and an expanding collection of metabolites. *Nucleic Acids Res.*, 44(D1):D1214–9, Oct 2015.
- [36] M Indriati Hood and Eric P Skaar. Nutritional immunity: transition metals at the pathogen-host interface. *Nature reviews. Microbiology*, 10:525–537, 07 2012.
- [37] Michael Hucka, Frank T. Bergmann, Claudine Chaouiya, Andreas Dräger, Stefan Hoops, Sarah M. Keating, Matthias König, Nicolas Le Novère, Chris J. Myers, Brett G. Olivier, Sven Sahle, James C. Schaff, Rahuman Sheriff, Lucian P. Smith, Dagmar Waltemath, Darren J. Wilkinson, and Fengkai Zhang. The Systems Biology Markup Language (SBML): Language specification for Level 3 Version 2 Core Release 2. *Journal of Integrative Bioinformatics*, 16(2):20190021, 2019.
- [38] Michael Hucka and Lucian P. Smith. Sbml level 3 package: Groups, version 1 release 1. *Journal of Integrative Bioinformatics*, 13(3):8–29, 2016.
- [39] J. D. Hunter. Matplotlib: A 2d graphics environment. *Computing in Science & Engineering*, 9(3):90–95, 2007.
- [40] Minoru Kanehisa, Miho Furumichi, Mao Tanabe, Yoko Sato, and Kanae Morishima. KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res.*, 45(D1):353–361, Jan 2017.
- [41] Peter D. Karp, Richard Billington, Ron Caspi, Carol A. Fulcher, Mario Latendresse, Anamika Kothari, Ingrid M. Keseler, Markus Krummenacker, Peter E. Midford, Quang Ong, et al. The BioCyc collection of microbial genomes and metabolic pathways. *Briefings Bioinf.*, 20(4):1085–1093, Jul 2019.
- [42] Zachary A. King, Andreas Dräger, Ali Ebrahim, Nikolaus Sonnenschein, Nathan E. Lewis, and Bernhard O. Palsson. Escher: A Web Application for Building, Sharing, and Embedding Data-Rich Visualizations of Biological Pathways. *PLoS Comput. Biol.*, 11(8):e1004321, Aug 2015.
- [43] Zachary A. King, Justin Lu, Andreas Dräger, Philip Miller, Stephen Federowicz, Joshua A. Lerman, Ali Ebrahim, Bernhard O. Palsson, and Nathan E. Lewis. BiGG Models: A platform for integrating, standardizing and sharing genome-scale models. *Nucleic Acids Research*, 44(D1):D515–D522, 10 2015.
- [44] Thomas Kluyver, Benjamin Ragan-Kelley, Fernando Pérez, Brian E Granger, Matthias Bussonnier, Jonathan Frederic, Kyle Kelley, Jessica B Hamrick, Jason Grout, Sylvain Corlay, et al. *Jupyter Notebooks-a publishing format for reproducible computational workflows.*, volume 2016. 2016.

- [45] Bernhard Krismer, Manuel Liebeke, Daniela Janek, Mulugeta Nega, Maren Rautenberg, Gabriele Hornig, Clemens Unger, Christopher Weidenmaier, Michael Lalk, and Andreas Peschel. Nutrient limitation governs staphylococcus aureus metabolism and niche adaptation in the human nose. *PLOS Pathogens*, 10(1):1–17, 01 2014.
- [46] Christian Lieven, Moritz E. Beber, Brett G. Olivier, Frank T. Bergmann, Meric Ataman, Parizad Babaei, Jennifer A. Bartell, Lars M. Blank, Siddharth Chauhan, Kevin Correia, et al. MEMOTE for standardized genome-scale metabolic model testing - Nature Biotechnology. *Nat. Biotechnol.*, 38:272–276, Mar 2020.
- [47] Daniel Machado, Sergej Andrejev, Melanie Tramontano, and Kiran Raosaheb Patil. Fast automated reconstruction of genome-scale metabolic models for microbial species and communities. *Nucleic Acids Res.*, 46(15):7542–7553, Sep 2018.
- [48] Dirk Merkel. Docker: lightweight linux containers for consistent development and deployment. *Linux journal*, 2014(239):2, 2014.
- [49] Jonathan M. Monk, Colton J. Lloyd, Elizabeth Brunk, Nathan Mih, Anand Sastry, Zachary King, Rikiya Takeuchi, Wataru Nomura, Zhen Zhang, Hirotada Mori, et al. iML1515, a knowledgebase that computes Escherichia coli traits. *Nat. Biotechnol.*, 35(10):904–908, Oct 2017.
- [50] Sébastien Moretti, Van Du T. Tran, Florence Mehl, Mark Ibberson, and Marco Pagni. MetaNetX/MNXref: unified namespace for metabolites and biochemical reactions in the context of metabolic models. *Nucleic Acids Res.*, 49(D1):D570–D574, Jan 2021.
- [51] Nuala A. O’Leary, Mathew W. Wright, J. Rodney Brister, Stacy Ciufo, Diana Had-dad, Rich McVeigh, Bhanu Rajput, Barbara Robbertse, Brian Smith-White, Danso Ako-Adjei, et al. Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res.*, 44(D1):733–745, Jan 2016.
- [52] Jeffrey D. Orth, Tom M. Conrad, Jessica Na, Joshua A. Lerman, Hojung Nam, Adam M. Feist, and Bernhard Ø. Palsson. A comprehensive genome-scale reconstruction of Escherichia coli metabolism–2011. *Mol. Syst. Biol.*, 7:535., Oct 2011.
- [53] Alina Renz, Lina Widerspick, and Andreas Dräger. First Genome-Scale Metabolic Model of *Dulosigranulum pigrum* Confirms Multiple Auxotrophies. *Metabolites*, 11(4):232, Apr 2021.
- [54] Michael Römer, Johannes Eichner, Andreas Dräger, Clemens Wrzodek, Finja Wrzodek, and Andreas Zell. ZBIT Bioinformatics Toolbox: A Web-Platform for Systems Biology and Expression Data Analysis. *PLoS One*, 11(2):e0149263, Feb 2016.
- [55] Christian F. P. Scholz and Mogens Kilian. The natural history of cutaneous propionibacteria, and reclassification of selected species within the genus *Propionibacterium* to the proposed novel genera *Acidipropionibacterium* gen. nov., *Cutibacterium* gen. nov. and *Pseudopropionibacterium* gen. nov. *Int. J. Syst. Evol. Microbiol.*, 66(11):4422–4432, Nov 2016.
- [56] Anna Sikora and Farah Zahra. Nosocomial Infections. In *StatPearls*. StatPearls Publishing, Aug 2021. Available at <https://www.ncbi.nlm.nih.gov/books/NBK559312>.

- [57] The UniProt Consortium. UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Res.*, 49(D1):D480–D489, Jan 2021.
- [58] Christos C. Zouboulis. Acne and sebaceous gland function. *Clin. Dermatol.*, 22(5):360–366, Sep 2004.