

sbml_model

Collapse All

Readme

2019-06-03 10:00

Independent Section

Contains tests that are independent of the class of modeled organism, a model's complexity or types of identifiers that are used

Consistency

Stoichiometric Consistency

100.0%

x3

Stoichiometric inconsistency violates universal constraints:
1. Molecular masses are always positive, and 2. On each side of a reaction the mass is conserved. A single incorrectly defined reaction can lead to stoichiometric inconsistency in the model, and consequently to unconserved metabolites. Similar to insufficient constraints, this may give rise to cycles which either produce mass from nothing or consume mass from the model.

Implementation: This test first uses an implementation of the algorithm presented in section 3.1 by Gevorgyan, A., M. G Poolman, and D. A Fell. "Detection of Stoichiometric Inconsistencies in Biomolecular Models." Bioinformatics 24, no. 19 (2008): 2245. doi: 10.1093/bioinformatics/btn425
Should the model be inconsistent, then the list of unconserved metabolites is computed using the algorithm described in section 3.2 of the same publication.

This model contains 0 (0.00%) unconserved metabolites:

[]

Mass Balance

94.1%

Specific Section

Covers general statistics and specific aspects of a metabolic network that are not universally applicable. See readme for more

SBML

SBML Level and Version

Errored

This test reports if the model file is represented in the latest edition (level) of the Systems Biology Markup Language (SBML) which is Level 3, and at least version 1.

Implementation: The level and version are parsed directly from the SBML document.

null

FBC enabled

Errored

The Flux Balance Constraints (FBC) Package extends SBML with structured and semantic descriptions for domain-specific model components such as flux bounds, multiple linear objective functions, gene-protein-reaction associations, metabolite chemical formulas, charge and related annotations which are relevant for parameterized GEMs and FBA models. The SBML and constraint-based modeling communities collaboratively develop this package and update it based on user input. Implementation: Parse the state of the FBC plugin from the SBML document.

sbml_model

[Collapse All](#)[Readme](#)

2019-06-03 10:00

have a formula defined. In steady state, for each metabolite the sum of influx equals the sum of efflux. Hence the net masses of both sides of any model reaction have to be equal. Reactions where at least one metabolite does not have a formula are not considered to be balanced, even though the remaining metabolites participating in the reaction might be. Implementation: For each reaction that isn't a boundary or biomass reaction check if each metabolite has a non-zero elements attribute and if so calculate if the overall element balance of reactants and products is equal to zero.

A total of 25 (5.90%) reactions are mass unbalanced with at least one of the metabolites not having a formula or the overall mass not equal to 0: 3OAS80, DALTA_LLA, MCOATA, PTALEU, 3OAS60, ...

```
[ "3OAS80", "DALTA_LLA", "MCOATA", "PTALEU", "3OAS60", "EAR140x", "ACKVAL", "3OAS140", "ACKILE", "3OAS160", "3OAR140", "EAR160x", "2H3MBDH", "PTA
```

Charge Balance

32.5%

This will exclude biomass, exchange and demand reactions as they are unbalanced by definition. It will also fail all reactions where at least one metabolite does not have a charge defined. In steady state, for each metabolite the sum of influx equals the sum of efflux. Hence the net charges of both sides of any model reaction have to be equal. Reactions where at least one metabolite does not have a charge are not considered to be balanced, even though the remaining metabolites participating in the reaction might be. Implementation: For each reaction that isn't a boundary or biomass reaction check if each

Basic Information

Model Identifier

sbml_mo
del

The MIRIAM guidelines require a model to be identified via an ID. Further, the ID will be displayed on the memote snapshot report, which helps to distinguish the output clearly. Implementation: Check if the cobra.Model object has a non-empty "id" attribute, this value is parsed from the "id" attribute of the <model> tag in the SBML file e.g. <model fbc:strict="true" id="iJO1366">.

The model ID is sbml_model

"sbml_model"

Total Metabolites

510

To be useful a metabolic model should consist at least of a few metabolites that are converted by reactions. This test simply checks if there are more than zero metabolites. Implementation: Check if the cobra.Model object has non-empty "metabolites" attribute, this list is populated from the list of sbml:listOfSpecies which should contain at least one sbml:species.

510 metabolites are defined in the model.

sbml_model

[Collapse All](#)[Readme](#)

2019-06-03 10:00

A total of 286 (67.45%) reactions are charge unbalanced with at least one of the metabolites not having a charge or the overall charge not equal to 0: ALAt2r, DRPA, RBFSb, PHETA1, aratry2, ...

```
[ "ALAt2r", "DRPA", "RBFSb", "PHETA1", "aratry2",  
  "IGPS", "RNDR3", "DHORD1", "GPDDA5", "AB6PGH", "CYTGL",  
  "CSND", "GTHPi", "DURIK1", "ARGabc", "YPD
```

Metabolite Connectivity

100.0%



Disconnected metabolites are not part of any reaction in the model. They are most likely left-over from the reconstruction process, but may also point to network and knowledge gaps. Implementation: Check for any metabolites of the cobra.Model object with empty reaction attribute.

A total of 0 (0.00%) metabolites are not associated with any reaction of the model:

[]



Unbounded Flux In Default Medium

0.0%



A large fraction of model reactions able to carry unlimited flux under default conditions indicates problems with reaction directionality, missing cofactors, incorrectly defined transport reactions and more. Implementation: Without changing the default constraints run flux variability analysis.

```
  ", "13dpg_91_c_93", "23dhdp_91_c_93  
  ", "25aics_91_c_93", "25dhpp_91_c_93
```

Total Reactions

424



To be useful a metabolic model should consist at least of a few reactions. This test simply checks if there are more than zero reactions. Implementation: Check if the cobra.Model object has non-empty "reactions" attribute, this list is populated from the list of sbml:listOfReactions which should contain at least one sbml:reaction.

424 reactions are defined in the model.

```
[ "GLYCT", "PPIabc", "CTPS2", "OIVD3", "CTPS1", "G6PI", "GLCpts", "FOLR2", "ADSS", "PRPPS", "DTMPK",  
  "S6PG", "TRPTA", "UGT2_LLA", "ASNN", "UAAGLS1_1
```

Total Genes

364



A metabolic model can still be a useful tool without any genes, however there are certain methods which rely on the presence of genes and, more importantly, the corresponding gene-protein-reaction rules. This test requires that there is at least one gene defined. Implementation: Check if the cobra.Model object has non-empty "genes" attribute, this list is populated from the list of fbc:listOfGeneProducts which should contain at least one fbc:geneProduct.

364 genes are defined in the model.

```
[ "lp_2788", "lp_1670", "UNKNOWN", "lp_1674", "lp_1675", "lp_3054", "lp_2777", "lp_3662", "lp_033
```

[sbml_model](#)[Collapse All](#)[Readme](#)

2019-06-03 10:00

A fraction of 100.00% of the non-blocked reactions (in total 18 reactions) can carry unbounded flux in the default model condition. Unbounded reactions may be involved in thermodynamically infeasible cycles: TRPTA, SUCD1, G3PD, G3PD1ir, TYRTA, ...

```
[ "TRPTA", "SUCD1", "G3PD", "G3PD1ir", "TYRTA", "LETU", "aratry1", "aratry2", "PHETA1", "ILETA", "araphe3", "araphe1", "FRDx", "aratyr4", "aratyr3
```

Sub Total 75%

Annotation - Metabolites

Presence of Metabolite Annotation 0.0%

This test checks if any annotations at all are present in the SBML annotations field for each metabolite, irrespective of the type of annotation i.e. specific database cross-references, ontology terms, additional information. For this test to pass the model is expected to have metabolites and each of them should have some form of annotation.

Implementation: Check if the annotation attribute of each cobra.Metabolite object of the model is unset or empty.

A total of 510 metabolites (100.00%) lack any form of annotation: 10fthf_91_c_93_, 12dgr_LLA_91_c_93_, 13dpq_91_c_93_, 23dhdp_91_c_93_, 25aics_91_c_93_, ...

While simplified metabolic models may be perfectly viable, generally across the tree of life organisms contain at least one distinct compartment: the cytosol or cytoplasm. In the case of prokaryotes there is usually a periplasm, and eukaryotes are more complex. In addition to the internal compartment, a metabolic model also reflects the extracellular environment i.e. the medium/ metabolic context in which the modelled cells grow. Hence, in total, at least two compartments can be expected from a metabolic model. Implementation: Check if the cobra.Model object has a non-empty "compartments" attribute, this list is populated from the list of sbml:listOfCompartments which should contain at least two sbml:compartment elements.

A total of 2 compartments are defined in the model: c, e

```
[ "c", "e" ]
```

Metabolic Coverage 1.16

The degree of metabolic coverage indicates the modeling detail of a given reconstruction calculated by dividing the total amount of reactions by the amount of genes. Models with a 'high level of modeling detail have ratios >1, and models with a low level of detail have ratios <1. This difference arises as models with basic or intermediate levels of detail are assumed to include many reactions in which several gene products and their enzymatic transformations are 'lumped'. Implementation: Divides the amount reactions by the amount of genes. Raises an error if the model does not contain either reactions or genes.

sbml_model

Collapse All Readme

2019-06-03 10:00

Metabolite Annotations Per Database

Info



Specific database cross-references are paramount to mapping information. To provide references to as many databases as possible helps to make the metabolic model more accessible to other researchers. This does not only facilitate the use of a model in a broad array of computational pipelines, it also promotes the metabolic model itself to become an organism-specific knowledge base. For this test to pass, each metabolite annotation should contain cross-references to a number of databases. The currently selection is listed in `annotation.py`, but an ongoing discussion can be found at <https://github.com/opencobra/memote/issues/332>. For each database this test checks for the presence of its corresponding namespace ID to comply with the MIRIAM guidelines i.e. they have to match those defined on <https://identifiers.org/>. Since each database is quite different and some potentially incomplete, it may not be feasible to achieve 100% coverage for each of them. Generally it should be possible, however, to obtain cross-references to at least one of the databases for all metabolites consistently. Implementation: Check if the keys of the annotation attribute of each cobra.Metabolite of the model match with a selection of common biochemical databases. The annotation attribute of cobrapy components is a dictionary of key:value pairs.

[424, 364]



Metabolite Information

Unique Metabolites

510



Metabolites may be transported into different compartments, which means that in a compartmentalized model the number of metabolites may be much higher than in a model with no compartments. This test counts only one occurrence of each metabolite and returns this as the number of unique metabolites. The test expects that the model is compartmentalized, and thus, that the number of unique metabolites is generally lower than the total number of metabolites. Implementation: Reduce the list of metabolites to a unique set by removing the compartment tag. The cobrapy SBML parser adds compartment tags to each metabolite ID.

Not counting the same entities in other compartments, there is a total of 510 (100.00%) unique metabolites in the model: hqn_91_c_93_, arbt6p_91_c_93_, galt_91_e_93_, gdp_91_c_93_, fum_91_c_93_, ...

pubchem.compound

0.0%



The following 510 metabolites (100.00%) lack annotation for pubchem.compound: 10fthf_91_c_93_, 12dgr_LLA_91_c_93_, 13dpg_91_c_93_, 23dhdp_91_c_93_, 25aics_91_c_93_, ...

["hqn_91_c_93_", "arbt6p_91_c_93_", "g

[sbml_model](#) [Collapse All](#) [Readme](#)

2019-06-03 10:00

```
_","25aics_91_c_93_","25dhpp_91_c_93_
```

kegg.compound 0.0% [^](#)

The following 510 metabolites (100.00%) lack annotation for kegg.compound: 10fthf_91_c_93_, 12dgr_LLA_91_c_93_, 13dpq_91_c_93_, 23dhdp_91_c_93_, 25aics_91_c_93_, ...

```
["10fthf_91_c_93_","12dgr_LLA_91_c_93_"
_, "13dpq_91_c_93_","23dhdp_91_c_93_"
_, "25aics_91_c_93_","25dhpp_91_c_93_
```

seed.compound 0.0% [^](#)

The following 510 metabolites (100.00%) lack annotation for seed.compound: 10fthf_91_c_93_, 12dgr_LLA_91_c_93_, 13dpq_91_c_93_, 23dhdp_91_c_93_, 25aics_91_c_93_, ...

```
["10fthf_91_c_93_","12dgr_LLA_91_c_93_"
_, "13dpq_91_c_93_","23dhdp_91_c_93_"
_, "25aics_91_c_93_","25dhpp_91_c_93_
```

inchikey 0.0% [^](#)

The following 510 metabolites (100.00%) lack annotation for inchikey: 10fthf_91_c_93_, 12dgr_LLA_91_c_93_, 13dpq_91_c_93_, 23dhdp_91_c_93_, 25aics_91_c_93_, ...

```
c_93_","h2o_91_c_93_","h2s_91_c_93_
_","h_91_c_93_","h_91_e_93_","hco3_
91_c_93_","hcys_l_91_c_93_","hexACP_
91_c_93_","his_l_91_c_93_","hisp_9
1_c_93_","histd_91_c_93_","histp_9
```

Duplicate Metabolites in Identical Compartments 0 [^](#)

The main reason for having this test is to help cleaning up merged models or models from automated reconstruction pipelines as these are prone to having identical metabolites from different namespaces (hence different IDs). This test therefore expects that every metabolite in any particular compartment has unique inchikey values. Implementation: Identifies duplicate metabolites in each compartment by determining if any two metabolites have identical InChI-key annotations. For instance, this function would find compounds with IDs ATP1 and ATP2 in the cytosolic compartment, with both having the same InChI annotations.

There are a total of 0 metabolites in the model which have duplicates in the same compartment:

```
[]
```

Metabolites without Charge 174 [^](#)

To be able to ensure that reactions are charge-balanced, all model metabolites ought to be provided with a charge. Since it may be difficult to obtain charges for certain metabolites this test serves as a mere report. Models can still be stoichiometrically consistent even when charge information is not defined for each metabolite. Implementation: Check if each cobra.Metabolite has a non-empty "charge" attribute. This attribute is set by the parser if there is an fbc:charge attribute for the corresponding species in the SBML.

[sbml_model](#) [Collapse All](#) [Readme](#)

2019-06-03 10:00

12dgr_LLA_91_c_93_, 13dpq_91_c_93_,
23dhdp_91_c_93_, 25aics_91_c_93_, ...

["10fthf_91_c_93_", "12dgr_LLA_91_c_93_",
"13dpq_91_c_93_", "23dhdp_91_c_93_",
"25aics_91_c_93_", "25dhpp_91_c_93_

chebi

0.0%

The following 510 metabolites (100.00%) lack annotation
for chebi: 10fthf_91_c_93_,
12dgr_LLA_91_c_93_, 13dpq_91_c_93_,
23dhdp_91_c_93_, 25aics_91_c_93_, ...

["10fthf_91_c_93_", "12dgr_LLA_91_c_93_",
"13dpq_91_c_93_", "23dhdp_91_c_93_",
"25aics_91_c_93_", "25dhpp_91_c_93_

hmdb

0.0%

The following 510 metabolites (100.00%) lack annotation
for hmdb: 10fthf_91_c_93_,
12dgr_LLA_91_c_93_, 13dpq_91_c_93_,
23dhdp_91_c_93_, 25aics_91_c_93_, ...

["10fthf_91_c_93_", "12dgr_LLA_91_c_93_",
"13dpq_91_c_93_", "23dhdp_91_c_93_",
"25aics_91_c_93_", "25dhpp_91_c_93_

reactome

0.0%

The following 510 metabolites (100.00%) lack annotation
for reactome: 10fthf_91_c_93_,
12dgr_LLA_91_c_93_, 13dpq_91_c_93_,
23dhdp_91_c_93_, 25aics_91_c_93_, ...

2chdeacp_91_c_93_, 2cocdacp_91_c_93_, ...

["12dgr_LLA_91_c_93_", "26dap_M_91_c_93_",
"2ahhmp_91_c_93_", "2chdeacp_91_c_93_",
"2cocdacp_91_c_93_", "2ctdeacp_9

Metabolites without Formula

0

To be able to ensure that reactions are mass-balanced, all model metabolites ought to be provided with a chemical formula. Since it may be difficult to obtain formulas for certain metabolites this test serves as a mere report. Models can still be stoichiometrically consistent even when chemical formulas are not defined for each metabolite. Implementation: Check if each cobra.Metabolite has a non-empty "formula" attribute. This attribute is set by the parser if there is an fbc:chemicalFormula attribute for the corresponding species in the SBML.

There are a total of 0 metabolites (0.00%) without a formula:

[]

Medium Components

0

This test checks all boundary reactions in the model that permit flux towards creating a metabolite, and reports those metabolites. This test does not have any mandatory 'pass' criteria. Implementation: Identify the metabolite IDs of each reaction in the method cobra.Model.medium.

sbml_model

Collapse All

Readme

2019-06-03 10:00

","25aics_91_c_93_","25dhpp_91_c_93_

metanetx.chemical

0.0%



The following 510 metabolites (100.00%) lack annotation for metanetx.chemical: 10fthf_91_c_93_,
12dgr_LLA_91_c_93_, 13dpq_91_c_93_,
23dhdp_91_c_93_, 25aics_91_c_93_, ...

["10fthf_91_c_93_","12dgr_LLA_91_c_93_","13dpq_91_c_93_","23dhdp_91_c_93_","25aics_91_c_93_","25dhpp_91_c_93_"]

bigg.metabolite

0.0%



The following 510 metabolites (100.00%) lack annotation for bigg.metabolite: 10fthf_91_c_93_,
12dgr_LLA_91_c_93_, 13dpq_91_c_93_,
23dhdp_91_c_93_, 25aics_91_c_93_, ...

["10fthf_91_c_93_","12dgr_LLA_91_c_93_","13dpq_91_c_93_","23dhdp_91_c_93_","25aics_91_c_93_","25dhpp_91_c_93_"]

biocyc

0.0%



The following 510 metabolites (100.00%) lack annotation for biocyc: 10fthf_91_c_93_,
12dgr_LLA_91_c_93_, 13dpq_91_c_93_,
23dhdp_91_c_93_, 25aics_91_c_93_, ...

["10fthf_91_c_93_","12dgr_LLA_91_c_93_","13dpq_91_c_93_","23dhdp_91_c_93_","25aics_91_c_93_","25dhpp_91_c_93_"]

[]



Reaction Information

Purely Metabolic Reactions

374



If a reaction is neither a transport reaction, a biomass reaction nor a boundary reaction, it is counted as a purely metabolic reaction. This test requires the presence of metabolite formula to be able to identify transport reactions. This test is passed when the model contains at least one purely metabolic reaction i.e. a conversion of one metabolite into another. Implementation: From the list of all reactions, those that are boundary, transport and biomass reactions are removed and the remainder assumed to be pure metabolic reactions. Boundary reactions are identified using the attribute cobra.Model.boundary. Please read the description of "Transport Reactions" and "Biomass Reaction Identified" to learn how they are identified.

A total of 374 (88.21%) purely metabolic reactions are defined in the model, this excludes transporters, exchanges, or pseudo-reactions: NDPK3, ORPT, DRPA, NDPK1, RBFSb, ...

["NDPK3","ORPT","DRPA","NDPK1","RBFSb","PHET A1","aratry2","IGPS","RNDR3","DHORDi","GPDDA 5","AB6PGH","UAGPT1","CYSTGL","CSND","NADS1"]



sbml_model

[Collapse All](#)
[Readme](#)

2019-06-03 10:00

To identify databases and the identifiers belonging to them, computational tools rely on the presence of specific patterns. Only when these patterns can be identified consistently is an ID truly machine-readable. This test checks if the database cross-references in metabolite annotations conform to patterns defined according to the MIRIAM guidelines, i.e. matching those that are defined at <https://identifiers.org/>. The required formats, i.e., regex patterns are further outlined in `annotation.py`. This test does not carry out a web query for the composed URI, it merely controls that the regex patterns match the identifiers. Implementation: For those metabolites whose annotation keys match any of the tested databases, check if the corresponding values match the identifier pattern of each database.

pubchem.compound

0.0%



There are no metabolite annotations for the pubchem.compound database.

[]

kegg.compound

0.0%



There are no metabolite annotations for the kegg.compound database.

[]

seed.compound

0.0%



reaction nor a boundary reaction, it is counted as a purely metabolic reaction. This test requires the presence of metabolite formula to be able to identify transport reactions. This test simply reports the number of purely metabolic reactions that have fixed constraints and does not have any mandatory 'pass' criteria. Implementation: From the pool of pure metabolic reactions identify reactions which are constrained to values other than the model's minimal or maximal possible bounds.

A total of 11 (2.94%) purely metabolic reactions have fixed constraints in the model, this excludes transporters, exchanges, or pseudo-reactions: $0.0 \leq GTHPi \leq 0.0$, $0.0 \leq NADHPO \leq 0.0$, $0.0 \leq NOX2 \leq 0.0$, $0.0 \leq OIVD1r \leq 0.0$, $0.0 \leq GTHOr \leq 0.0$, ...

```
[ "0.0 <= GTHPi <= 0.0", "0.0 <= NADHPO <= 0.0", "0.0 <= NOX2 <= 0.0", "0.0 <= OIVD1r <= 0.0", "0.0 <= GTHOr <= 0.0", "0.0 <= PDH <= 0.0", "0.0 <= G3PD4 <= 0.0", "0.0 <= FDPK1 <= 0.0", "0.0 <= FDPK2 <= 0.0", "0.0 <= FDPK3 <= 0.0" ]
```

Transport Reactions

50



Cellular metabolism in any organism usually involves the transport of metabolites across a lipid bi-layer. This test reports how many of these reactions, which transports metabolites from one compartment to another, are present in the model, as at least one transport reaction must be present for cells to take up nutrients and/or excrete waste. Implementation: A transport reaction is defined as follows:

1. It contains metabolites from at least 2 compartments and
2. at least 1 metabolite undergoes no chemical reaction, i.e., the formula and/or annotation stays the same on both sides of the equation. A notable exception is transport via PTS, which also contains the following restriction:
3. The

sbml_model

Collapse All

Readme

2019-06-03 10:00

[]

inchikey

0.0%



There are no metabolite annotations for the inchikey database.

[]

inchi

0.0%



There are no metabolite annotations for the inchi database.

[]

chebi

0.0%



There are no metabolite annotations for the chebi database.

[]

hmdb

0.0%



There are no metabolite annotations for the hmdb database.

-> glucose-6-phosphate(c) + pyr(c) Reactions similar to transport via PTS (referred to as "modified transport reactions") follow a similar pattern: A(x) + B-R(y) -> A-R(y) + B(y) Such modified transport reactions can be detected, but only when the formula is defined for all metabolites in a particular reaction. If this is not the case, transport reactions are identified through annotations, which cannot detect modified transport reactions.

A total of 50 (11.79%) transport reactions are defined in the model, this excludes purely metabolic reactions, exchanges, or pseudo-reactions: DSERT2, MALTabc, GUAt2r, ALAt2r, MNabc, ...

["DSERT2", "MALTabc", "GUAt2r", "ALAt2r", "MNabc", "PIabc", "ACGAppts", "GLCpts", "GLUabc", "GALT2", "GLYBabc", "RIBabc", "CHLabc", "FE3abc", "CELB

Transport Reactions with Constraints

2



Cellular metabolism in any organism usually involves the transport of metabolites across a lipid bi-layer. Hence, this test reports how many of these reactions, which transports metabolites from one compartment to another, have fixed constraints. This test does not have any mandatory 'pass' criteria. Implementation: Please refer to "Transport Reactions" for details on how memote identifies transport reactions. From the pool of transport reactions identify reactions which are constrained to values other than the model's median lower and upper bounds.

A total of 2 (4.00%) transport reactions have fixed constraints in the model: $-1000.0 \leq GLYCt \leq 0.0$, $0.0 \leq CYTB_B2 \leq 0.0$

reactome 0.0%

There are no metabolite annotations for the reactome database.

[]

metanetx.chemical 0.0%

There are no metabolite annotations for the metanetx.chemical database.

[]

bigg.metabolite 0.0%

There are no metabolite annotations for the bigg.metabolite database.

[]

biocyc 0.0%

There are no metabolite annotations for the biocyc database.

Thermodynamic Reversibility of Purely Metabolic Reactions 1.00

If a reaction is neither a transport reaction, a biomass reaction nor a boundary reaction, it is counted as a purely metabolic reaction. This test checks if the reversibility attribute of each reaction agrees with a thermodynamics-based calculation of reversibility. Implementation: To determine reversibility we calculate the reversibility index \ln_{γ} (natural logarithm of gamma) of each reaction using the eQuilibrator API. We consider reactions, whose reactants' concentrations would need to change by more than three orders of magnitude for the reaction flux to reverse direction, to be likely candidates of irreversible reactions. This assume default concentrations around 100 μM ($\sim 3 \mu\text{M}$ —3 mM) at pH = 7, I = 0.1 M and T = 298 K. The corresponding reversibility index is approximately 7. For further information on the thermodynamic and implementation details please refer to <https://doi.org/10.1093/bioinformatics/bts317> and <https://pypi.org/project/equilibrium-api/>. Please note that currently eQuilibrator can only determine the reversibility index for chemically and redox balanced reactions whose metabolites can be mapped to KEGG compound identifiers (e.g. C00001). In addition to not being mappable to KEGG or the reaction not being balanced, there is a possibility that the metabolite cannot be broken down into chemical groups which is essential for the calculation of Gibbs energy using group contributions. This test collects each erroneous reaction and returns them as a tuple containing each list in the following order: 1. Reactions with reversibility index 2. Reactions with incomplete mapping to KEGG 3. Reactions with metabolites that are problematic

sbml_model

Collapse All

Readme

2019-06-03 10:00

Uniform Metabolite Identifier Namespace 100.0%

In well-annotated models it is no problem if the pool of main identifiers for metabolites consists of identifiers from several databases. However, in models that lack appropriate annotations, it may hamper the ability of other researchers to use it. Running the model through a computational pipeline may be difficult without first consolidating the namespace. Hence, this test checks if the main metabolite identifiers can be attributed to one single namespace based on the regex patterns defined at <https://identifiers.org/> Implementation: Generate a table with each column corresponding to one database from the selection and each row to a metabolite identifier. A Boolean entry indicates whether the identifier matches the regular expression of the corresponding database. Since the Biocyc pattern matches broadly, we assume that any instance of an identifier matching to Biocyc AND any other database pattern is a false positive match for Biocyc and thus set it to ``false``. Sum the positive matches for each database and assume that the largest set is the 'main' identifier namespace.

0 metabolite identifiers (0.00%) deviate from the largest found namespace (bigg.metabolite):

[]

to the reversibility index, are likely to be irreversible.

Out of 374 purely metabolic reactions, 0 have an absolute reversibility index greater or equal to 7 and are therefore likely candidates for being irreversible. 374 reactions could not be mapped to KEGG completely, 0 contained 'problematic' metabolites, and 0 are chemically or redox imbalanced.

```
[],  
["NDPK3", "ORPT", "DRPA", "NDPK1", "RBFSb", "PHET  
A1", "aratry2", "IGPS", "RNDR3", "DHORDi", "GPDDA  
5", "AB6PGH", "UAGPT1", "CYSTGL", "CSND", "NADS1"]
```

Reactions With Partially Identical Annotations 0.00

Identify reactions in a pairwise manner that are annotated with identical database references. This does not take into account a reaction's directionality or compartment. The main reason for having this test is to help cleaning up merged models or models from automated reconstruction pipelines as these are prone to having identical reactions with identifiers from different namespaces. It could also be useful to identify a 'type' of reaction that occurs in several compartments. Implementation: Identify duplicate reactions globally by checking if any two metabolic reactions have the same entries in their annotation attributes. The heuristic looks at annotations with the keys "metanetx.reaction", "kegg.reaction", "brenda", "rhea", "biocyc", "bigg.reaction" only.

Based on annotations there are 0 different groups of overlapping annotation which corresponds to a total of 0 duplicated reactions in the model.

Sub Total

25%

sbml_model

[Collapse All](#)[Readme](#)

2019-06-03 10:00

This test checks if any annotations at all are present in the SBML annotations field for each reaction, irrespective of the type of annotation i.e. specific database cross-references, ontology terms, additional information. For this test to pass the model is expected to have reactions and each of them should have some form of annotation.
Implementation: Check if the annotation attribute of each cobra.Reaction object of the model is unset or empty.

A total of 424 reactions (100.00%) lack any form of annotation: GLYCt, PPIabc, CTPS2, OIVD3, CTPS1, ...

```
[ "GLYCt", "PPIabc", "CTPS2", "OIVD3", "CTPS1", "G6PI", "GLCpts", "FOLR2", "ADSS", "PRPPS", "DTMPK", "S6PG", "TRPTA", "UGT2_LLA", "ASNN", "UAAGLS1_1
```

Reaction Annotations Per Database

[Info](#)

Specific database cross-references are paramount to mapping information. To provide references to as many databases as possible helps to make the metabolic model more accessible to other researchers. This does not only facilitate the use of a model in a broad array of computational pipelines, it also promotes the metabolic model itself to become an organism-specific knowledge base. For this test to pass, each reaction annotation should contain cross-references to a number of databases. The currently selection is listed in `annotation.py`, but an ongoing discussion can be found at <https://github.com/opencobra/memote/issues/332>. For each database this test checks for the presence of its corresponding namespace ID to comply with the MIRIAM

Duplicate Reactions

0.00

Identify reactions in a pairwise manner that use the same set of metabolites including potentially duplicate metabolites. Moreover, it will take a reaction's directionality and compartment into account. The main reason for having this test is to help cleaning up merged models or models from automated reconstruction pipelines as these are prone to having identical reactions with identifiers from different namespaces. Implementation: Compare reactions in a pairwise manner. For each reaction, the metabolite annotations are checked for a description of the structure (via InChI and InChIKey). If they exist, substrates and products as well as the stoichiometries of any reaction pair are compared. Only reactions where the substrates, products, stoichiometry and reversibility are identical are considered to be duplicates. This test will not be able to identify duplicate reactions if there are no structure annotations. Further, it will report reactions with differing bounds as equal if they otherwise match the above conditions.

Based on metabolites, directionality and compartment there are a total of 0 reactions in the model which have duplicates:

[]

Reactions With Identical Genes

0.45

feasible to achieve 100% coverage for each of them. Generally it should be possible, however, to obtain cross-references to at least one of the databases for all reactions consistently. Implementation: Check if the keys of the annotation attribute of each cobra.Reaction of the model match with a selection of common biochemical databases. The annotation attribute of cobrapy components is a dictionary of key:value pairs.

rhea 0.0%

The following 424 reactions (100.00%) lack annotation for rhea: GLYCt, PPIabc, CTPS2, OIVD3, CTPS1, ...

```
[ "GLYCt", "PPIabc", "CTPS2", "OIVD3", "CTPS1", "G6PI", "GLCpts", "FOLR2", "ADSS", "PRPPS", "DTMPK", "S6PG", "TRPTA", "UGT2_LLA", "ASNN", "UAAGLS1_1
```

kegg.reaction 0.0%

The following 424 reactions (100.00%) lack annotation for kegg.reaction: GLYCt, PPIabc, CTPS2, OIVD3, CTPS1, ...

```
[ "GLYCt", "PPIabc", "CTPS2", "OIVD3", "CTPS1", "G6PI", "GLCpts", "FOLR2", "ADSS", "PRPPS", "DTMPK", "S6PG", "TRPTA", "UGT2_LLA", "ASNN", "UAAGLS1_1
```

seed.reaction 0.0%

The following 424 reactions (100.00%) lack annotation for seed.reaction: GLYCt, PPIabc, CTPS2, OIVD3, CTPS1, ...

The main reason for having this test is to help cleaning up merged models or models from automated reconstruction pipelines as these are prone to having identical reactions with identifiers from different namespaces. Implementation: Compare reactions in a pairwise manner and group reactions whose genes are identical. Skip reactions with missing genes.

Based only on equal genes there are 63 different groups of identical reactions which corresponds to a total of 191 duplicated reactions in the model.

```
{"UNKNOWN,lp_0733,lp_0747,lp_0748,lp_0749,lp_0750":["PPIabc","PIabc"],"lp_0481":["CTPS1","CTPS2"],"lp_2151,lp_2152,lp_2153,lp_2154":["OIVD3"]}
```

Gene-Protein-Reaction (GPR) Associations

Reactions without GPR 0

Gene-Protein-Reaction rules express which gene has what function. The presence of this annotation is important to justify the existence of reactions in the model, and is required to conduct in silico gene deletion studies. However, reactions without GPR may also be valid: Spontaneous reactions, or known reactions with yet undiscovered genes likely lack GPR. Implementation: Check if each cobra.Reaction has a non-empty "gene_reaction_rule" attribute, which is set by the parser if there is an fbc:geneProductAssociation defined for the corresponding reaction in the SBML.

sbml_model

Collapse All

Readme

2019-06-03 10:00

metanetx.reaction

0.0%

The following 424 reactions (100.00%) lack annotation for metanetx.reaction: GLYCt, PPIabc, CTPS2, OIVD3, CTPS1, ...

```
[ "GLYCt", "PPIabc", "CTPS2", "OIVD3", "CTPS1", "G6PI", "GLCpts", "FOLR2", "ADSS", "PRPPS", "DTMPK", "S6PG", "TRPTA", "UGT2_LLA", "ASNN", "UAAGLS1_1
```

bigg.reaction

0.0%

The following 424 reactions (100.00%) lack annotation for bigg.reaction: GLYCt, PPIabc, CTPS2, OIVD3, CTPS1, ...

```
[ "GLYCt", "PPIabc", "CTPS2", "OIVD3", "CTPS1", "G6PI", "GLCpts", "FOLR2", "ADSS", "PRPPS", "DTMPK", "S6PG", "TRPTA", "UGT2_LLA", "ASNN", "UAAGLS1_1
```

reactome

0.0%

The following 424 reactions (100.00%) lack annotation for reactome: GLYCt, PPIabc, CTPS2, OIVD3, CTPS1, ...

```
[ "GLYCt", "PPIabc", "CTPS2", "OIVD3", "CTPS1", "G6PI", "GLCpts", "FOLR2", "ADSS", "PRPPS", "DTMPK", "S6PG", "TRPTA", "UGT2_LLA", "ASNN", "UAAGLS1_1
```

ec-code

0.0%

The following 424 reactions (100.00%) lack annotation for ec-code: GLYCt, PPIabc, CTPS2, OIVD3, CTPS1, ...

Fraction of Transport Reactions without GPR

0.00

As it is hard to identify the exact transport processes within a cell, transport reactions are often added purely for modeling purposes. Highlighting where assumptions have been made versus where there is proof may help direct the efforts to improve transport and transport energetics of the tested metabolic model. However, transport reactions without GPR may also be valid: Diffusion, or known reactions with yet undiscovered genes likely lack GPR. Implementation: Check which cobra.Reactions classified as transport reactions have a non-empty "gene_reaction_rule" attribute.

There are a total of 0 transport reactions (0.00% of all transport reactions) without GPR:

[]

Enzyme Complexes

63

Based on the gene-protein-reaction (GPR) rules, it is possible to infer whether a reaction is catalyzed by a single gene product, isozymes or by a heteromeric protein complex. This test checks that at least one such

sbml_model Collapse All Readme 2019-06-03 10:00

brenda	0.0%	^	(https://doi.org/10.1074%2Fmcp.M800490-MCP200). This might also be a relevant metric for other organisms. Implementation: Identify GPRs which contain at least one logical AND that combines two different gene products. A total of 63 reactions are catalyzed by complexes defined through GPR rules in the model.
biocyc	0.0%	^	["GLYct", "PPIabc", "CTPS2", "OIVD3", "CTPS1", "G6PI", "GLCpts", "FOLR2", "ADSS", "PRPPS", "DTMPK", "S6PG", "TRPTA", "UGT2_LLA", "ASNN", "UAAGLS1_1] ["PPIabc", "OIVD3", "GLCpts", "DALTA_LLA", "OIVD2", "GLCS1", "ACGALpts", "MNLpts", "ATPS3r", "SPMDabc", "DHAPT", "SERD_L", "PIabc", "PDH", "METab]
Reaction Annotation Conformity Per Database		Info	^
<p>To identify databases and the identifiers belonging to them, computational tools rely on the presence of specific patterns. Only when these patterns can be identified consistently is an ID truly machine-readable. This test checks if the database cross-references in reaction annotations conform to patterns defined according to the MIRIAM guidelines, i.e. matching those that are defined at https://identifiers.org/. The required formats, i.e., regex patterns are further outlined in `annotation.py`. This test does not carry out a web query for the composed URI, it merely controls that the regex patterns match the identifiers. Implementation: For those reaction whose annotation keys match any of the tested databases, check</p>			

Biomass

Biomass Reactions Identified 0 ^

The biomass composition aka biomass formulation aka biomass reaction is a common pseudo-reaction accounting for biomass synthesis in constraints-based modelling. It describes the stoichiometry of intracellular compounds that are required for cell growth. While this reaction may not be relevant to modeling the metabolism of higher organisms, it is essential for single-cell modeling. Implementation: Identifies possible biomass reactions using two principal steps: 1. Return reactions that include the SBO annotation "SBO:0000629" for biomass. If no reactions can be identified this way: 1. Look for the ``buzzwords`` "biomass", "growth" and "bof" in reaction IDs. 2. Look for metabolite IDs or names that contain the ``buzzword`` "biomass" and obtain the set of reactions they are involved in. 3. Remove boundary reactions from this set. 4. Return the union of reactions that match the buzzwords and of the reactions that metabolites are involved in that match the buzzword. This test checks if at least one biomass reaction is present.

In this model 0 of the following biomass reactions were identified.

sbml_model

[Collapse All](#)[Readme](#)

2019-06-03 10:00

rhea

0.0%

There are no reaction annotations for the rhea database.

[]

kegg.reaction

0.0%

There are no reaction annotations for the kegg.reaction database.

[]

seed.reaction

0.0%

There are no reaction annotations for the seed.reaction database.

[]

metanetx.reaction

0.0%

There are no reaction annotations for the metanetx.reaction database.

[]

Biomass Consistency

Skipped

This test only yields sensible results if all biomass precursor metabolites have chemical formulas assigned to them. The molecular weight of the biomass reaction in metabolic models is defined to be equal to 1 g/mmol. Conforming to this is essential in order to be able to reliably calculate growth yields, to cross-compare models, and to obtain valid predictions when simulating microbial consortia. A deviation from 1 - 1E-03 to 1 + 1E-06 is accepted. Implementation: Multiplies the coefficient of each metabolite of the biomass reaction with its molecular weight calculated from the formula, then divides the overall sum of all the products by 1000.

Biomass Production In Default Medium

Skipped

Using flux balance analysis this test optimizes the model for growth in the medium that is set by default. Any non-zero growth rate is accepted to pass this test. Implementation: Calculate the solution of FBA with the biomass reaction set as objective function and the model's default constraints.

sbml_model

[Collapse All](#)[Readme](#)

2019-06-03 10:00

[]

reactome 0.0%

There are no reaction annotations for the reactome database.

ec-code 0.0%

There are no reaction annotations for the ec-code database.

brenda 0.0%

There are no reaction annotations for the brenda database.

biocyc 0.0%

There are no reaction annotations for the biocyc database.

the fastest growing organism. This is based on a doubling time of *Vibrio natriegens* which was reported to be 14.8 minutes by: Henry H. Lee, Nili Ostrov, Brandon G. Wong, Michaela A. Gold, Ahmad S. Khalil, George M. Church in <https://www.biorxiv.org/content/biorxiv/early/2016/06/12/058487.full.pdf>. The calculation $\ln(2)/(14.8/60) \sim 2.81$ yields the corresponding growth rate. Implementation: Calculate the solution of FBA with the biomass reaction set as objective function and a model's default constraints. Then check if the objective value is higher than 2.81.

Biomass Production In Complete Medium

Skipped

Using flux balance analysis this test optimizes the model for growth using a complete medium i.e. unconstrained boundary reactions. Any non-zero growth rate is accepted to pass this test. Implementation: Calculate the solution of FBA with the biomass reaction set as objective function and after removing any constraints from all boundary reactions.

Blocked Biomass Precursors In Default Medium

Skipped

Using flux balance analysis this test optimizes for the production of each metabolite that is a substrate of the biomass reaction with the exception of atp and h2o. Optimizations are carried out using the

sbml_model

Collapse All

Readme

2019-06-03 10:00

Uniform Reaction Identifier Namespace 100.0%

In well-annotated models it is no problem if the pool of main identifiers for reactions consists of identifiers from several databases. However, in models that lack appropriate annotations, it may hamper the ability of other researchers to use it. Running the model through a computational pipeline may be difficult without first consolidating the namespace. Hence, this test checks if the main reaction identifiers can be attributed to one single namespace based on the regex patterns defined at <https://identifiers.org/>. Implementation: Generate a pandas.DataFrame with each column corresponding to one database from the selection and each row to the reaction ID. A boolean entry indicates whether the metabolite ID matches the regex pattern of the corresponding database. Since the Biocyc pattern matches quite, assume that any instance of an identifier matching to Biocyc AND any other DB pattern is a false positive match for Biocyc and then set the boolean to ``false``. Sum the positive matches for each database and assume that the largest set is the 'main' identifier namespace.

0 reaction identifiers (0.00%) deviate from the largest found namespace (bigg.reaction):

[]

Sub Total

25%

For each biomass precursor (except ATP and H₂O) add a temporary demand reaction, then carry out FBA with this reaction as the objective. Collect all metabolites for which this optimization is equal to zero or infeasible.

Blocked Biomass Precursors In Complete Medium

Skipped

Using flux balance analysis this test optimizes for the production of each metabolite that is a substrate of the biomass reaction with the exception of atp and h₂o. Optimizations are carried out using a complete medium i.e. unconstrained boundary reactions. This is useful when reconstructing the precursor biosynthesis pathways of a metabolic model. To pass this test, the model should be able to synthesis all the precursors. Implementation: First remove any constraints from all boundary reactions, then for each biomass precursor (except ATP and H₂O) add a temporary demand reaction, then carry out FBA with this reaction as the objective. Collect all metabolites for which this optimization is below or equal to zero or is infeasible.

Ratio of Direct Metabolites in Biomass Reaction

Skipped

Some biomass precursors are taken from the media and directly consumed by the biomass reaction. It might not be a problem for ions or metabolites for which the organism in question is auxotrophic.

sbml_model

Collapse All

Readme

2019-06-03 10:00

This test checks if any annotations at all are present in the SBML annotations field (extended by FBC package) for each gene product, irrespective of the type of annotation i.e. specific database, cross-references, ontology terms, additional information. For this test to pass the model is expected to have genes and each of them should have some form of annotation. Implementation: Check if the annotation attribute of each cobra.Gene object of the model is unset or empty.

A total of 364 genes (100.00%) lack any form of annotation: lp_2788, lp_1670, UNKNOWN, lp_1674, lp_1675, ...

```
["lp_2788", "lp_1670", "UNKNOWN", "lp_1674", "lp_1675", "lp_3054", "lp_2777", "lp_3662", "lp_0337", "lp_1680", "lp_1676", "lp_1679", "lp_1274", "
```

Gene Annotations Per Database

Info



Specific database cross-references are paramount to mapping information. To provide references to as many databases as possible helps to make the metabolic model more accessible to other researchers. This does not only facilitate the use of a model in a broad array of computational pipelines, it also promotes the metabolic model itself to become an organism-specific knowledge base. For this test to pass, each gene annotation should contain cross-references to a number of databases. The currently selection is listed in `annotation.py`, but an ongoing discussion can be found at <https://github.com/opencobra/memote/issues/332>. For

biomass precursors being "direct metabolites". This test reports the ratio of direct metabolites to the total amount of precursors to a given biomass reaction. It specifically looks for metabolites that are only in either exchange, transport or biomass reactions. Bear in mind that this may lead to false positives in heavily compartmentalized models. To pass this test, the ratio of direct metabolites should be less than 50% of all biomass precursors. This is an arbitrary threshold but it takes into account that while certain ions do not serve a relevant metabolic function, it may still be important to include them in the biomass reaction to account for the impact of their uptake energy costs. This threshold is subject to change in the future. Implementation: Identify biomass precursors (excluding ATP and H+), identify cytosol and extracellular compartment from an internal mapping table. Then, determine which precursors is only involved in transport, boundary and biomass reactions. Using FBA with the biomass function as the objective then determine whether the metabolite is taken up only to be consumed by the biomass reaction.

Number of Missing Essential Biomass Precursors

Skipped



There are universal components of life that make up the biomass of all known organisms. These include all proteinogenic amino acids, deoxy- and ribonucleotides, water and a range of metabolic cofactors. This test reports the amount of biomass precursors that have been reported to be essential constituents of the biomass equation. All of the following precursors need to be included in the biomass reaction to pass the test: Aminoacids: trp_L, cys_L, his_L, tyr_L, met_L, phe_L, ser_L, pro_L, asp_L, thr_L, gln_L, glu_L, ile_L, arg_L, lys_L, val_L, leu_L, ala_L, gly, asn_L DNA: dntp, dctp, dtcp, dgtp RNA: atp, ctp, utp, gtp Cofactors: nad, nadp, amet, fad,

[sbml_model](#) [Collapse All](#) [Readme](#)

2019-06-03 10:00

<https://identifiers.org/>. Since each database is quite different and some potentially incomplete, it may not be feasible to achieve 100% coverage for each of them. Generally it should be possible, however, to obtain cross-references to at least one of the databases for all gene products consistently. Implementation: Check if the keys of the annotation attribute of each cobra.Gene of the model match with a selection of common genome databases. The annotation attribute of cobrapy components is a dictionary of key:value pairs.

refseq 0.0%

The following 364 genes (100.00%) lack annotation for refseq: lp_2788, lp_1670, UNKNOWN, lp_1674, lp_1675,

...

```
["lp_2788", "lp_1670", "UNKNOWN", "lp_1674", "lp_1675", "lp_3054", "lp_2777", "lp_3662", "lp_0337", "lp_1680", "lp_1676", "lp_1679", "lp_1274", "
```

uniprot 0.0%

The following 364 genes (100.00%) lack annotation for uniprot: lp_2788, lp_1670, UNKNOWN, lp_1674, lp_1675,

...

```
["lp_2788", "lp_1670", "UNKNOWN", "lp_1674", "lp_1675", "lp_3054", "lp_2777", "lp_3662", "lp_0337", "lp_1680", "lp_1676", "lp_1679", "lp_1274", "
```

ecogene 0.0%

The following 364 genes (100.00%) lack annotation for ecogene: lp_2788, lp_1670, UNKNOWN, lp_1674,

(derivatives of tetrahydrofolate(B9) or tetrahydromethanopterin) as universal cofactors. We have omitted these from this check because there are many individual compounds that classify as C1 carriers, and it is not clear a priori which one should be preferred. In a future update, we may consider identifying these using a chemical ontology. Implementation: Determine whether the model employs a lumped or split biomass reaction. Then, using an internal mapping table, try to identify the above list of essential precursors in list of precursor metabolites of either type of biomass reaction. List IDs in the models namespace if the metabolite exists, else use the MetaNetX namespace if the metabolite does not exist in the model. Identifies the cytosol from an internal mapping table, and assumes that all precursors exist in that compartment.

Energy Metabolism

Non-Growth Associated Maintenance Reaction

Errored

The Non-Growth Associated Maintenance reaction (NGAM) is an ATP-hydrolysis reaction added to metabolic models to represent energy expenses that the cell invests in continuous processes independent of the growth rate. Memote tries to infer this reaction from a list of buzzwords, and the stoichiometry and components of a simple ATP-hydrolysis reaction. Implementation: From the list of all reactions that convert ATP to ADP select the reactions that match the irreversible reaction "ATP + H₂O → ADP +

sbml model

[Collapse All](#) [Readme](#)

Readme

2019-06-03 10:00

kegg.genes

0.0%

The following 364 genes (100.00%) lack annotation for kegg.genes: lp_2788, lp_1670, UNKNOWN, lp_1674, lp_1675, ...

ncbigi

0.0%

The following 364 genes (100.00%) lack annotation for
ncbigi: lp_2788, lp_1670, UNKNOWN, lp_1674, lp_1675

三

```
["lp_2788","lp_1670","UNKNOWN","lp_1674","lp_1675","lp_3054","lp_2777","lp_3662","lp_0337","lp_1680","lp_1676","lp_1679","lp_1274","lp_1671","lp_1672","lp_1673","lp_1675","lp_1676","lp_1677","lp_1678","lp_1679","lp_1680","lp_1681","lp_1682","lp_1683","lp_1684","lp_1685","lp_1686","lp_1687","lp_1688","lp_1689","lp_1690","lp_1691","lp_1692","lp_1693","lp_1694","lp_1695","lp_1696","lp_1697","lp_1698","lp_1699","lp_1671","lp_1672","lp_1673","lp_1675","lp_1676","lp_1677","lp_1678","lp_1679","lp_1680","lp_1681","lp_1682","lp_1683","lp_1684","lp_1685","lp_1686","lp_1687","lp_1688","lp_1689","lp_1690","lp_1691","lp_1692","lp_1693","lp_1694","lp_1695","lp_1696","lp_1697","lp_1698","lp_1699"]
```

ncbigene

0.0%

The following 364 genes (100.00%) lack annotation for
ncbigene: lp_2788, lp_1670, UNKNOWN, lp_1674,
lp_1675, ...

ncbiprotein

0.0%

The following 364 genes (100.00%) lack annotation for

It is assumed to be the compartment with the most metabolites. The resulting list of reactions is then filtered further by attempting to match the reaction name with any of the following buzzwords ('maintenance', 'atpm', 'requirement', 'ngam', 'non-growth', 'associated'). If this is possible only the filtered reactions are returned, if not the list is returned as is.

null

Growth-associated Maintenance in Biomass Reaction

The growth-associated maintenance (GAM) term accounts for the energy in the form of ATP that is required to synthesize macromolecules such as Proteins, DNA and RNA, and other processes during growth. A GAM term is therefore a requirement for any well-defined biomass reaction. There are different ways to implement this term depending on what kind of experimental data is available and the preferred way of implementing the biomass reaction: - Chemostat growth experiments yield a single GAM value representing the required energy per gram of biomass (Figure 6 of [1]). This can be implemented in a lumped biomass reaction or in the final term of a split biomass reaction. - Experimentally delineating or estimating the GAM requirements for each macromolecule separately is possible, yet requires either data from multi-omics experiments [2] or detailed resources [1], respectively. Individual energy requirements can either be implemented in a split biomass equation on the term for each macromolecule, or, on the basis of the biomass composition, they can be summed into a single GAM value.

sbml_model

Collapse All

Readme

2019-06-03 10:00

```
["lp_2788", "lp_1670", "UNKNOWN", "lp_1674", "lp_1675", "lp_3054", "lp_2777", "lp_3662", "lp_0337", "lp_1680", "lp_1676", "lp_1679", "lp_1274", "
```

ccds

0.0%

The following 364 genes (100.00%) lack annotation for ccds: lp_2788, lp_1670, UNKNOWN, lp_1674, lp_1675, ...

```
["lp_2788", "lp_1670", "UNKNOWN", "lp_1674", "lp_1675", "lp_3054", "lp_2777", "lp_3662", "lp_0337", "lp_1680", "lp_1676", "lp_1679", "lp_1274", "
```

hprd

0.0%

The following 364 genes (100.00%) lack annotation for hprd: lp_2788, lp_1670, UNKNOWN, lp_1674, lp_1675, ...

```
["lp_2788", "lp_1670", "UNKNOWN", "lp_1674", "lp_1675", "lp_3054", "lp_2777", "lp_3662", "lp_0337", "lp_1680", "lp_1676", "lp_1679", "lp_1274", "
```

asap

0.0%

The following 364 genes (100.00%) lack annotation for asap: lp_2788, lp_1670, UNKNOWN, lp_1674, lp_1675, ...

```
["lp_2788", "lp_1670", "UNKNOWN", "lp_1674", "lp_1675", "lp_3054", "lp_2777", "lp_3662", "lp_0337", "lp_1680", "lp_1676", "lp_1679", "lp_1274", "
```

Gene Annotation Conformity Per Database

Info

To identify databases and the identifiers belonging to them, computational tools rely on the presence of specific

it will only detect the use of a single GAM value as opposed to individual energy requirements of each macromolecule. Both approaches, however, have its merits. Implementation: Determines the metabolite identifiers of ATP, ADP, H₂O, HO₄P and H⁺ based on an internal mapping table. Checks if ATP and H₂O are a subset of the reactants and ADP, HO₄P and H⁺ a subset of the products of the biomass reaction. References: .. [1] Thiele, I., & Palsson, B. Ø. (2010, January). A protocol for generating a high-quality genome-scale metabolic reconstruction. *Nature protocols*. Nature Publishing Group. <http://doi.org/10.1038/nprot.2009.203> .. [2] Hackett, S. R., Zanotelli, V. R. T., Xu, W., Goya, J., Park, J. O., Perlman, D. H., Gibney, P. A., Botstein, D., Storey, J. D., Rabinowitz, J. D. (2010, January). Systems-level analysis of mechanisms regulating yeast metabolic flux *Science* <http://doi.org/10.1126/science.aaf2786>

Number of Reversible Oxygen-Containing Reactions

1

The directionality of oxygen-producing/-consuming reactions affects the model's ability to grow anaerobically i.e. create faux-anaerobic organisms. This test reports how many of these oxygen-containing reactions are reversible. This test does not have any mandatory 'pass' criteria. Implementation: First, find the metabolite representing atmospheric oxygen in the model on the basis of an internal mapping table or by specifically looking for the formula "O₂". Then, find all reactions that produce or consume oxygen and report those that are reversible.

sbml_model

[Collapse All](#)[Readme](#)

2019-06-03 10:00

annotations conform to patterns defined according to the MIRIAM guidelines, i.e. matching those that are defined at <https://identifiers.org/>. The required formats, i.e., regex patterns are further outlined in `annotation.py`. This test does not carry out a web query for the composed URI, it merely controls that the regex patterns match the identifiers. Implementation: For those genes whose annotation keys match any of the tested databases, check if the corresponding values match the identifier pattern of each database.

refseq 0.0% ^

There are no gene annotations for the refseq database.

[]

uniprot 0.0% ^

There are no gene annotations for the uniprot database.

[]

ecogene 0.0% ^

There are no gene annotations for the ecogene database.

[]

DHORDi

["DHORDi"]

Erroneous Energy-generating Cycles Info

When a model is not sufficiently constrained to account for the thermodynamics of reactions, flux cycles may form which provide reduced metabolites to the model without requiring nutrient uptake. These cycles are referred to as erroneous energy-generating cycles. Their effect on the predicted growth rate in FBA may account for an increase of up to 25%, which makes studies involving the growth rates predicted from such models unreliable.

Implementation: This test uses an implementation of the algorithm presented by: Fritzemeier, C. J., Hartleb, D., Szappanos, B., Papp, B., & Lercher, M. J. (2017).

Erroneous energy-generating cycles in published genome scale metabolic networks: Identification and removal. PLoS Computational Biology, 13(4), 1–14.

<http://doi.org/10.1371/journal.pcbi.1005494> First attempt to identify the main compartment (cytosol), then attempt to identify each metabolite of the referenced list of energy couples via an internal mapping table. Construct a dissipation reaction for each couple. Carry out FBA with each dissipation reaction as the objective and report those reactions that non-zero carry flux.

MNXM3

Skipped

[sbml_model](#)[Collapse All](#)[Readme](#)

2019-06-03 10:00

[]	MNXM63	Skipped
ncbigi [] There are no gene annotations for the ncbigi database.	MNXM51	Skipped
ncbigene [] There are no gene annotations for the ncbigene database.	MNXM121	Skipped
ncbiprotein [] There are no gene annotations for the ncbiprotein database.	MNXM423	Skipped
ccds [] There are no gene annotations for the ccds database.	MNXM6	Skipped

sbml_model

Collapse All Readme

2019-06-03 10:00

hprd

0.0%

There are no gene annotations for the hprd database.

[]



asap

0.0%

There are no gene annotations for the asap database.

[]



Sub Total

0%



Annotation - SBO Terms

Metabolite General SBO Presence

0.0%



The Systems Biology Ontology (SBO) allows researchers to annotate a model with terms which indicate the intended function of its individual components. The available terms are controlled and relational and can be viewed here <http://www.ebi.ac.uk/sbo/main/tree>. Implementation: Check if each cobra.Metabolite has a non-zero "annotation" attribute that contains the key "sbo".

A total of 510 metabolites (100.00%) lack annotation with any type of SBO term: 10ftfhf_91_c_93_,

MNXM38

Skipped



MNXM208

Skipped



MNXM191

Skipped



MNXM223

Skipped



MNXM7517

Skipped



sbml_model

Collapse All Readme

2019-06-03 10:00

	MNXM12233	Skipped
<pre>["10fthf_91_c_93_", "12dgr_LLA_91_c_93_", "13dpg_91_c_93_", "23dhdp_91_c_93_", "25aics_91_c_93_", "25dhpp_91_c_93_"]</pre> <p>Metabolite SBO:0000247 Presence 0.0% ^</p> <p>SBO:0000247 represents the term 'simple chemical'. Every metabolite should be annotated with this. Implementation: Check if each cobra.Metabolite has a non-zero "annotation" attribute that contains the key "sbo" with the associated value being one of the SBO terms above.</p> <p>A total of 510 transport reactions (100.00% of all metabolites) lack annotation with the SBO term "SBO:0000247" for 'simple chemical':</p> <ul style="list-style-type: none"> 10fthf_91_c_93_, 12dgr_LLA_91_c_93_, 13dpg_91_c_93_, 23dhdp_91_c_93_, 25aics_91_c_93_, ... 		
	MNXM558	Skipped
	MNXM21	Skipped
	MNXM89557	Skipped
	<h2>Network Topology</h2> <p>Universally Blocked Reactions 406</p>	

sbml_model

Collapse All

Readme

2019-06-03 10:00

```
[ "GLYct", "PPIabc", "CTPS2", "OIVD3", "CTPS1", "G6PI", "GLCpts", "FOLR2", "ADSS", "PRPPS", "DTMPK", "S6PG", "TRPTA", "UGT2_LLA", "ASNN", "UAAGLS1_1
```

Metabolic Reaction SBO:0000176 Presence 0.0%

SBO:0000176 represents the term 'biochemical reaction'. Every metabolic reaction that is not a transport or boundary reaction should be annotated with this. The results shown are relative to the total amount of pure metabolic reactions. Implementation: Check if each pure metabolic reaction has a non-zero "annotation" attribute that contains the key "sbo" with the associated value being the SBO term above.

A total of 374 metabolic reactions (100.00% of all purely metabolic reactions) lack annotation with the SBO term "SBO:0000176" for 'biochemical reaction': NDPK3, ORPT, DRPA, NDPK1, RBFSb, ...

```
[ "NDPK3", "ORPT", "DRPA", "NDPK1", "RBFSb", "PHET_A1", "aratry2", "IGPS", "RNDR3", "DHORD1", "GPDDA_5", "AB6PGH", "UAGPT1", "CYSTGL", "CSND", "NADS1"
```

Transport Reaction SBO:0000185 Presence 0.0%

'SBO:0000185', 'SBO:0000588', 'SBO:0000587', 'SBO:0000655', 'SBO:0000654', 'SBO:0000660', 'SBO:0000659', 'SBO:0000657', and 'SBO:0000658' represent the terms 'transport reaction' and 'translocation reaction', in addition to their children (more specific transport reaction labels). Every transport reaction that is not a pure metabolic or boundary reaction should be

which can be attributed to scope or knowledge gaps.

Implementation: Use flux variability analysis (FVA) implemented in cobra.flux_analysis.find_blocked_reactions with open_exchanges=True. Please refer to the cobrapy documentation for more information:

https://cobrapy.readthedocs.io/en/stable/autoapi/cobra/flux_analysis/variability/index.html#cobra.flux_analysis.variability.find_blocked_reactions

There are 406 (95.75%) blocked reactions in the model: GLYct, PPIabc, CTPS2, OIVD3, CTPS1, ...

```
[ "GLYct", "PPIabc", "CTPS2", "OIVD3", "CTPS1", "G6PI", "GLCpts", "FOLR2", "ADSS", "PRPPS", "DTMPK", "S6PG", "UGT2_LLA", "ASNN", "UAAGLS1_1", "3OAR140", "METGL", "ALAt2r", "IL
```

Orphan Metabolites 112

Orphans are metabolites that are only consumed but not produced by reactions in the model. They may indicate the presence of network and knowledge gaps. Implementation: Find orphan metabolites structurally by considering only reaction equations and reversibility. FBA is not carried out.

A total of 112 (21.96%) metabolites are not produced by any reaction of the model: 12dgr_LLA_91_c_93_, 25aics_91_c_93_, 26dap_M_91_c_93_, 2aeppn_91_e_93_, 2chdeacp_91_c_93_, ...

```
[ "12dgr_LLA_91_c_93_", "25aics_91_c_93_", "26dap_M_91_c_93_", "2aeppn_91_e_93_", "2chdeacp_91_c_93_", "2cocdacp_91_c_93_", "2cpr5p_91_c
```

sbml_model

[Collapse All](#)
[Readme](#)

2019-06-03 10:00

non-zero "annotation" attribute that contains the key "sbo" with the associated value being one of the SBO terms above.

A total of 50 metabolic reactions (100.00% of all transport reactions) lack annotation with one of the SBO terms: ['SBO:0000185', 'SBO:0000588', 'SBO:0000587', 'SBO:0000655', 'SBO:0000654', 'SBO:0000660', 'SBO:0000659', 'SBO:0000657', 'SBO:0000658] for 'biochemical reaction': DSERt2, MALTabc, GUAt2r, ALAt2r, MNabc, ...

["DSERt2", "MALTabc", "GUAt2r", "ALAt2r", "MNabc", "PIabc", "ACGAppts", "GLCpts", "GLUabc", "GALT2", "GLYBabc", "RIBabc", "CHLabc", "FE3abc", "CELB

Exchange Reaction SBO:0000627 Presence Skipped

SBO:0000627 represents the term 'exchange reaction'. The Systems Biology Ontology defines an exchange reaction as follows: 'A modeling process to provide matter influx or efflux to a model, for example to replenish a metabolic network with raw materials (eg carbon / energy sources). Such reactions are conceptual, created solely for modeling purposes, and do not have a physical correspondence. Exchange reactions, often represented as 'R_EX_', can operate in the negative (uptake) direction or positive (secretion) direction. By convention, a negative flux through an exchange reaction represents uptake of the corresponding metabolite, and a positive flux represent discharge.' Every exchange reaction should be annotated with this. Exchange reactions differ from demand reactions in that the metabolites are removed from or added to the extracellular environment only. Implementation: Check if

consumed by reactions in the model. They may indicate the presence of network and knowledge gaps. Implementation: Find dead-end metabolites structurally by considering only reaction equations and reversibility. FBA is not carried out.

A total of 95 (18.63%) metabolites are not consumed by any reaction of the model: 25dhpp_91_c_93_, 2aeppn_91_c_93_, 2ahbut_91_c_93_, 2hxmp_91_c_93_, 2mahmp_91_c_93_, ...

["25dhpp_91_c_93_", "2aeppn_91_c_93_", "2ahbut_91_c_93_", "2hxmp_91_c_93_", "2mahmp_91_c_93_", "2mba_91_c_93_", "2mpa_91_c_93_", "3mba_

Stoichiometrically Balanced Cycles

21

Stoichiometrically Balanced Cycles are artifacts of insufficiently constrained networks resulting in reactions that can carry flux when all the boundaries have been closed. Implementation: Close all model boundary reactions and then use flux variability analysis (FVA) to identify reactions that carry flux.

There are 21 (4.95%) reactions which participate in SBC in the model: TRPTA, NOX, SUCD1, G3PD, G3PD1ir, ...

["TRPTA", "NOX", "SUCD1", "G3PD", "G3PD1ir", "TYRTA", "LEUTA", "aratry1", "aratry2", "PHETA1", "ILETA", "DHORD6", "DHORD1", "araphe3", "araphe1", "FRDx", "aratyr4", "aratyr3

Metabolite Production In Complete Medium

510

In complete medium, a model should be able to divert flux to every metabolite. This test opens all the boundary reactions i.e. simulates

[sbml_model](#)[Collapse All](#)[Readme](#)

2019-06-03 10:00

The model has no exchange reactions.

[]

Demand Reaction SBO:0000628 Presence Skipped

SBO:0000628 represents the term 'demand reaction'. The Systems Biology Ontology defines a demand reaction as follows: 'A modeling process analogous to exchange reaction, but which operates upon "internal" metabolites. Metabolites that are consumed by these reactions are assumed to be used in intra-cellular processes that are not part of the model. Demand reactions, often represented 'R_DM', can also deliver metabolites (from intra-cellular processes that are not considered in the model).' Every demand reaction should be annotated with this. Demand reactions differ from exchange reactions in that the metabolites are not removed from the extracellular environment, but from any of the organism's compartments. Demand reactions differ from sink reactions in that they are designated as irreversible. Implementation: Check if each demand reaction has a non-zero "annotation" attribute that contains the key "sbo" with the associated value being one of the SBO terms above.

The model has no demand reactions.

[]

downstream of reactions with fixed constraints, or blocked by a cofactor imbalance. To pass this test all metabolites should be producible. Implementation: Open all model boundary reactions, then for each metabolite in the model add a boundary reaction and maximize it with FBA.

A total of 510 (100.00%) metabolites cannot be produced in complete medium: 10fthf_91_c_93_, 12dgr_LLA_91_c_93_, 13dpq_91_c_93_, 23dhdp_91_c_93_, 25aics_91_c_93_, ...

["10fthf_91_c_93_", "12dgr_LLA_91_c_93_", "13dpq_91_c_93_", "23dhdp_91_c_93_", "25aics_91_c_93_", "25dhpp_91_c_93_", "26dap_M_91_c_93_"]

Metabolite Consumption In Complete Medium

510

In complete medium, a model should be able to divert flux from every metabolite. This test opens all the boundary reactions i.e. simulates a complete medium and checks if any metabolite cannot be consumed individually using flux balance analysis. Metabolites that cannot be consumed this way are likely dead-end metabolites or upstream of reactions with fixed constraints. To pass this test all metabolites should be consumable. Implementation: Open all model boundary reactions, then for each metabolite in the model add a boundary reaction and minimize it with FBA.

A total of 510 (100.00%) metabolites cannot be consumed in complete medium: 10fthf_91_c_93_, 12dgr_LLA_91_c_93_, 13dpq_91_c_93_, 23dhdp_91_c_93_, 25aics_91_c_93_, ...

["10fthf_91_c_93_", "12dgr_LLA_91_c_93_", "13dpq_91_c_93_", "23dhdp_91_c_93_", "25aics_91_c_93_", "25dhpp_91_c_93_", "26dap_M_91_c_93_"]

sbml_model

Collapse All

Readme

2019-06-03 10:00

Sink Reactions SBO:0000632 Presence Skipped ^

SBO:0000632 represents the term 'sink reaction'. The Systems Biology Ontology defines a sink reaction as follows: 'A modeling process to provide matter influx or efflux to a model, for example to replenish a metabolic network with raw materials (eg carbon / energy sources). Such reactions are conceptual, created solely for modeling purposes, and do not have a physical correspondence. Unlike the analogous demand (SBO:....) reactions, which are usually designated as irreversible, sink reactions always represent a reversible uptake/secretion processes, and act as a metabolite source with no cost to the cell. Sink reactions, also referred to as R_SINK_, are generally used for compounds that are metabolized by the cell but are produced by non-metabolic, un-modeled cellular processes.' Every sink reaction should be annotated with this. Sink reactions differ from exchange reactions in that the metabolites are not removed from the extracellular environment, but from any of the organism's compartments. Implementation: Check if each sink reaction has a non-zero "annotation" attribute that contains the key "sbo" with the associated value being one of the SBO terms above.

No sink reactions found.

[]

Gene General SBO Presence 0.0% ^

Matrix Conditioning

Ratio Min/Max Non-Zero Coefficients 0.00 ^

This test will return the absolute largest and smallest, non-zero coefficients of the stoichiometric matrix. A large ratio of these values may point to potential numerical issues when trying to solve different mathematical optimization problems such as flux-balance analysis. To pass this test the ratio should not exceed 10^9 . This threshold has been selected based on experience, and is likely to be adapted when more data on solver performance becomes available. Implementation: Compose the stoichiometric matrix, then calculate absolute coefficients and lastly use the maximal value and minimal non-zero value to calculate the ratio.

The ratio of the absolute values of the largest coefficient 25.0 and the lowest, non-zero coefficient 0.005 is: 5E+03.

5000

Independent Conservation Relations 115 ^

This test will return the number of conservation relations, i.e., conservation pools through the left null space of the stoichiometric matrix. This test is not scored, as the dimension of the left null space is system-specific. Implementation: Calculate the left null space, i.e., the null space of the transposed stoichiometric matrix, using an algorithm based on the singular value decomposition adapted from <https://scipy.github.io/old-wiki/pages/Cookbook/RankNullspace.html> Then, return the estimated dimension of that null space

sbml_model

Collapse All

Readme

2019-06-03 10:00

are controlled and relational and can be viewed here http://www.ebi.ac.uk/sbo/main/tree . Check if each cobra.Gene has a non-zero "annotation" attribute that contains the key "sbo".	115
A total of 364 genes (100.00%) lack annotation with any type of SBO term: lp_2788, lp_1670, UNKNOWN, lp_1674, lp_1675, ...	
<pre>["lp_2788", "lp_1670", "UNKNOWN", "lp_1674", "lp_1675", "lp_3054", "lp_2777", "lp_3662", "lp_0337", "lp_1680", "lp_1676", "lp_1679", "lp_1274", "</pre>	
Gene SBO:0000243 Presence 0.0%	
SBO:0000243 represents the term 'gene'. Every gene should be annotated with this. Implementation: Check if each cobra.Gene has a non-zero "annotation" attribute that contains the key "sbo" with the associated value being one of the SBO terms above.	
A total of 364 genes (100.00% of all genes) lack annotation with the SBO term "SBO:0000243" for 'gene': lp_2788, lp_1670, UNKNOWN, lp_1674, lp_1675, ...	
<pre>["lp_2788", "lp_1670", "UNKNOWN", "lp_1674", "lp_1675", "lp_3054", "lp_2777", "lp_3662", "lp_0337", "lp_1680", "lp_1676", "lp_1679", "lp_1274", "</pre>	
Biomass Reactions SBO:0000629 Presence Skipped	
SBO:0000629 represents the term 'biomass production'. The Systems Biology Ontology defines an exchange	

Rank	395
The rank of the stoichiometric matrix is system specific. It is calculated using singular value decomposition (SVD). Implementation: Compose the stoichiometric matrix, then estimate the rank, i.e. the dimension of the column space, of a matrix. The algorithm used by this function is based on the singular value decomposition of the matrix.	
The rank of the S-Matrix is 395.	
Degrees Of Freedom	29
The degrees of freedom of the stoichiometric matrix, i.e., the number of 'free variables' is system specific and corresponds to the dimension of the (right) null space of the matrix. Implementation: Compose the stoichiometric matrix, then calculate the dimensionality of the null space using the rank-nullity theorem outlined by Alama, J. The Rank+Nullity Theorem. Formalized Mathematics 15, (2007).	
The degrees of freedom of the S-Matrix are 29.	

sbml_model

Collapse All

Readme

2019-06-03 10:00

reactants to produce multiple products. It is also assumed that parts of the reactants are also consumed in unrepresented processes and hence products do not have to reflect all the atom composition of the reactants.

Formulation of a biomass production process entails definition of the macromolecular content (eg. cellular protein fraction), metabolic constitution of each fraction (eg. amino acids), and subsequently the atomic composition (eg. nitrogen atoms). More complex biomass functions can additionally incorporate details of essential vitamins and cofactors required for growth.¹ Every reaction representing the biomass production should be annotated with this.

Implementation: Check if each biomass reaction has a non-zero "annotation" attribute that contains the key "sbo" with the associated value being one of the SBO terms above.

No biomass reactions found.

[]

Sub Total

0%

X2



Total Score

33%

▼

Total Score

33%

Experimental Data Comparison

Growth Prediction

Skipped

The in-silico growth prediction is compared with experimental data and the accuracy is expected to be better than 0.95. In principal, Matthews' correlation coefficient is a more comprehensive metric but is a little fragile to not having any false negatives or false positives in the output. Implementation: Read and validate experimental config file and data tables. Constrain the model with the parameters provided by a user's definition of the medium, then compute a confusion matrix based on the predicted true, expected true, predicted false and expected false growth. The individual values of the confusion matrix are calculated as described in https://en.wikipedia.org/wiki/Confusion_matrix

Gene Essentiality Prediction

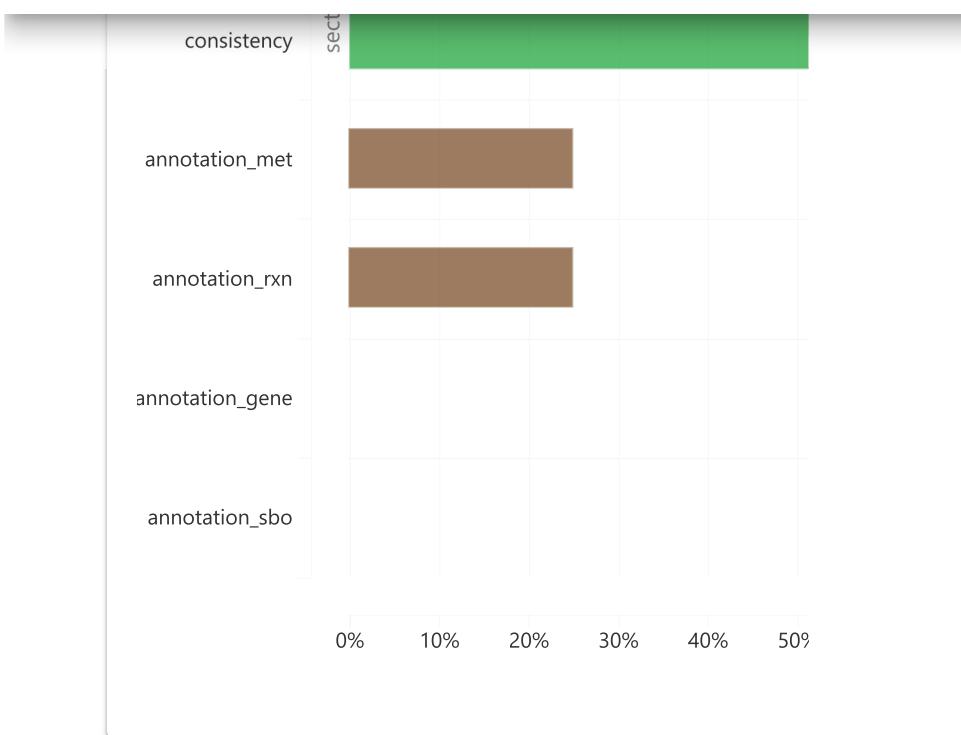
Skipped

The in-silico gene essentiality is compared with experimental data and the accuracy is expected to be better than 0.95. In principal, Matthews' correlation coefficient is a more comprehensive metric but is a little fragile to not having any false negatives or false positives in the output. Implementation: Read and validate experimental config file and data tables. Constrain the

sbml_model

[Collapse All](#)[Readme](#)

2019-06-03 10:00



Nonessential and expected nonessential genes. The individual values of the confusion matrix are calculated as described in https://en.wikipedia.org/wiki/Confusion_matrix

Misc. Tests

Environment

Python Version
Platform
Memote Version

3.6.8
Linux
0.9.11