Supplementary Material for “Saint: A Lightweight Integration Environment for Model Annotation” by Lister et al.

1 Introduction

This is the supplementary material for the Lister et al. publication “Saint: A Lightweight Integration Environment for Model Annotation”. It provides a short comparison, using three SBML models, of these models before and after annotation using Saint.

Saint makes use of two main features as starting points for adding annotation to models: the name/id of a species, and the existing resource URIs. Depending on the type of model loaded, one method may be more suitable than the other. The following examples illustrate when each method performs best. Saint can provide new MIRIAM\(^1\)-compliant annotation with even a minimal amount of data, thus allowing the addition of new information both for early-stage and already-annotated models.

The capabilities of Saint are being extended to include more data sources and web services, improve the user interface, and increase the coverage and quality of the annotation. Therefore the examples given in this supplementary material have a limited lifetime. To run these examples at any time, download the following models from the BioModels repository\(^2\) and run them within Saint\(^3\):

- BIOMD0000000087\(^4\), Modelling the checkpoint response to telomere uncapping in budding yeast, Proctor CJ, Lydall DA, Boys RJ, Gillespie CS, Shanley DP, Wilkinson DJ, Kirkwood TB.
- BIOMD0000000189\(^5\), Explaining oscillations and variability in the p53-Mdm2 system, Proctor Carole J, Gray Douglas A.
- BIOMD0000000032\(^6\), Modelling the dynamics of the yeast pheromone pathway, Kofahl B, Klipp E.

For each of the above models, the original as well as re-annotated versions are available from the Saint Wiki pages\(^7\). Further information on Saint can be found on the wiki and developer pages\(^8\). If you have questions or comments, please contact the CISBAN\(^9\) helpdesk (helpdesk@cisban.ac.uk).

2 Notes on These Examples

In all examples, we have begun by downloading the SBML Level 2 Version 4 files from BioModels. Then, “stripped” models are created from these original models and used as a starting point for Saint annotation. Then the re-annotated model as downloaded from Saint is compared to the original, intact model. Additionally, in the final example, Saint is loaded with the original model to determine what extra annotation is possible when Saint is provided with all of the available annotation.

\(^{1}\)http://www.ebi.ac.uk/miriam/
\(^{2}\)http://www.ebi.ac.uk/biomodels-main/
\(^{3}\)http://www.cisban.ac.uk/saint
\(^{4}\)http://www.ebi.ac.uk/biomodels-main/BIOMD0000000087
\(^{5}\)http://www.ebi.ac.uk/biomodels-main/BIOMD0000000189
\(^{6}\)http://www.ebi.ac.uk/biomodels-main/BIOMD0000000032
\(^{7}\)http://sourceforge.net/apps/mediawiki/saint-annotate/
\(^{8}\)http://saint-annotate.sourceforge.net
\(^{9}\)http://www.cisban.ac.uk
Saint is not only able to add names, SBO Terms and MIRIAM annotations, but can also add new reactions based on pathway data. However, such reactions are necessarily brand new and cannot be usefully compared against the original models. This functionality of Saint is therefore not included in this comparison.

In some cases, species names or ids are created as a hybrid of two protein names, e.g. “proteinA_proteinB”. In such cases, Saint suggests annotation based around both proteins. However, as a composite or complex formed from multiple proteins will not necessarily behave in the same way as its constituents, all annotation suggested by Saint except for the SBO terms and UniProt URIs will be removed by the user for the purposes of these examples. An expert modeller would be able to correctly choose the appropriate name and GO terms, but for the purposes of this example they were simply removed from consideration.

In this document, those species with informative IDs or names are termed “useful” species. Useful species have identifiers which closely match actual protein or gene names, and are therefore suitable for querying external resources. They do not have to be exact matches. There is a limit to how much information Saint can pull from non-informative IDs and names. In these cases, it is important to use any information already stored in the Annotation elements such as UniProt URIs. This style of annotation is covered Section 4.

2.1 Creating “Stripped” Models

After each model is downloaded from BioModels, the sboTerm and the Annotation element of each species is removed. Though all URIs within the Annotation elements are stripped from the model, only a proportion of those URIs are currently supported by Saint and therefore only those supported types may be added back by Saint. These stripped models are then loaded into Saint, and new annotation is retrieved. After the new annotation is presented to the user, basic editing of the annotation is performed, including the choosing of sensible species names from the list and removal of obvious false positive annotation. Then the newly-annotated model is downloaded and compared to the original, unmodified model.

Even though sboTerm and name attributes, together with RDF Annotation elements were added by Saint in these examples, only the sboTerm attribute and Annotation elements were removed during the stripping operation. This is because names cannot be stripped without losing the ability to match the query against the remote data sources. Examination of a number of BioModels entries shows a tendency for modellers to use either descriptive id attributes without names, or for modellers to use descriptive name attributes together with an arbitrary identifier scheme.

3 Comparison Against Stripped Models

The first example of using Saint illustrates the benefits of searching based on the id or name attributes of an SBML species element. Two models (BIOMD0000000087 and BIOMD0000000189) are annotated in this section. These models have a number of descriptive IDs for their species. However, each model is developed differently and without a common naming convention. Therefore, even within a single model, some names and IDs may be useful to Saint while others are not.

3.1 BIOMD0000000087

Basic information on the annotation of both the original and re-annotated BIOMD0000000087 model is presented in Table 1, together with a summary of the comparison between the two models. A detailed listing of the modifications made to the model is available in Table 2.

The total number of Saint-supported URIs (those from GO and UniProt) have increased from 14 to 163, with further editing of URIs by the modeller possible if required. In addition, Saint recovered 100% of the URIs from the useful species (14 URIs), as shown in Table 1. The original model contained no SBO terms or species names, and Saint has added 15 SBO Terms and 14 species names, which is almost 100% coverage of the useful species.

The species “ssDNA” was incorrectly annotated by Saint. “ssDNA” is difficult for Saint to filter: it is not a protein, but can be used as a phrase within protein names. However, Saint can keep track of terms...
| Species | 55 |
|---------|----|
| Species with useful identifiers | 15 |
| Species with Names (total) | 0 |
| **New species names (total after re-annotation)** | 14 |
| Species with annotation / name / SBO term | 51 (92% of 55) |
| Useful species with annotation / name / SBO term supported by Saint | 14 (93%) |
| **Useful species with annotation / name / SBO term supported by Saint after re-annotation** | 15 (100%) |
| SBO Terms (total) | 0 |
| **New SBO Terms (total after re-annotation)** | 15 |
| URIs / Resources (total) | 59 |
| **URIs / Resources after re-annotation** | 163 |
| URIs / Resources in useful species supported by Saint | 14 |
| **URIs / Resources recovered from useful species** | 14 (100%) |

Table 1: Basic information on the annotation present in the original, unmodified version of BIOMD0000000087 and in bold for the re-annotated version. This model had no species name attributes, and therefore Saint relies on the id attribute. There were also no SBO terms in the original model. Useful identifiers are those which match actual protein or gene names. The one useful species without Saint-supported annotation is “RPA”. This species has a PIRSF resource URI. A full explanation of this species and how it is annotated by Saint is in the main text.

that are both commonly used by modellers and not proteins. Further, the Saint user can (and has in this example) remove the suggested annotation using the Saint interface prior to downloading the model.

The one useful species which was not completely annotated with SBO term, species name and URIs was “RPA”. A large amount of annotation was suggested by Saint, but all except the SBO term was rejected by the user. In the original annotation, the “RPA” annotation is just a single URI which points to the PIR family (PIRSF) “replication protein A, RPA70 subunit” (http://pir.georgetown.edu/cgi-bin/ipcSF?id=PIRSF002091). This URI is marked as having an “isVersionOf” relationship the RPA species, rather than as an exact “is” match, as the PIRSF URI is linked to mouse, and not yeast. A search for an equivalent for yeast in UniProt shows a number of yeast subunits. In theory, an appropriate choice would be made by the modeller. In PIRSF002091, the equivalent yeast UniProt accession is listed as P22336 (http://www.uniprot.org/uniprot/P22336), which is “Replication factor A protein 1”, a.k.a. “RFA1”. This leads to the conclusion that the species in the model was named after the mouse protein and not the yeast protein. As “RFA” is not a name that matched the original ID (“RPA”), there is no way to retrieve the correct information from the data sources in the otherwise-annotationless species in the stripped-out SBML model used in this example.

In conclusion, in comparing the annotation on the useful species for this model, all URIs supported by Saint were recovered, and new URIs, SBO terms and species names were added. A detailed listing of all new and recovered annotation present in the useful species are shown in Table 2.

### 3.2 BIOMD0000000189

Basic information on the annotation of both the original and re-annotated BIOMD0000000189 model is presented in Table 3, together with a summary of the comparison between the two models. A detailed listing of the modifications made to the model is available in Table 4.

The total number of Saint-supported URIs (those from GO and UniProt) have increased from 7 to 89, with further editing of URIs by the modeller possible if required. Saint recovered 71% of the URIs from the useful species (5 URIs), as shown in Table 3. Saint also added 84 GO URIs. The original model contained 2 SBO terms, 100% of which were recovered with a further two being added. The original model did not include species names, and Saint has added two, which is one-third of the useful species.

In contrast to BIOMD0000000087, this model contained many species whose id attribute was created through a hybrid of two protein names, e.g. “Mdm2.p53”. Of the six useful species, four have IDs formed in this way. For each of these four, Saint suggested annotation based around both of their associated proteins. However, as a composite of two proteins will not necessarily behave in the same
Table 2: A detailed breakdown of recovered (in the “MATCH” column) and new or modified (in the “NEW” column) annotation by Saint for BIOMD0000000087. Only those annotations recognised by Saint are shown in this table. Additionally, new GO terms are omitted for clarity (there are a large number).

| SPECIES ID | MATCH | NEW |
|------------|-------|-----|
| Cdc13      |       | CDC13 SBO:0000245 |
| Cdc13      | urn:miriam:uniprot:P32797 |    |
| Rad17      |       | RAD17 SBO:0000245 |
| Rad17      | urn:miriam:uniprot:P48581 |    |
| Rad24      | urn:miriam:uniprot:P32641 | RAD24 SBO:0000245 |
| RPA        |       | SBO:0000245 |
| Mec1       | urn:miriam:uniprot:P38111 | MEC1 SBO:0000245 |
| Exo1I      | urn:miriam:uniprot:P39875 | EXO1 SBO:0000245 |
| Exo1A      | urn:miriam:uniprot:P39875 | EXO1 SBO:0000245 |
| Rad9I      | urn:miriam:uniprot:P14737 | RAD9 SBO:0000245 |
| Rad9A      | urn:miriam:uniprot:P14737 | RAD9 SBO:0000245 |
| Rad53I     | urn:miriam:uniprot:P22216 | RAD53 SBO:0000245 |
| Rad53A     | urn:miriam:uniprot:P22216 | RAD53 SBO:0000245 |
| Chk1I      | urn:miriam:uniprot:P38147 | CHK1 SBO:0000245 |
| Chk1A      | urn:miriam:uniprot:P38147 | CHK1 SBO:0000245 |
| Dun1I      | urn:miriam:uniprot:P39009 | DUN1 SBO:0000245 |
| Dun1A      | urn:miriam:uniprot:P39009 | DUN1 SBO:0000245 |

way as its constituents, all annotation except for SBO terms and UniProt URIs was removed by the user: specifically, the GO annotations and species names were removed. An expert modeller would be able to correctly choose the appropriate name and GO terms, but for the purposes of this example they were simply removed. The UniProt URIs were kept, but with a relationship of “hasPart”.

4
One species, “ARF”, was incorrectly identified. In the original model, the “ARF” species has an exact “is” match with UniProt Q8N726, whose names and synonyms are “Cyclin-dependent kinase inhibitor 2A, isoform 4”, “p14ARF”, and “p19ARF”. As such, the correct entry could not be returned due to the mismatch of the query term. However, a number of alternative suggestions were provided by Saint.

| Species | Useful identifiers | Species with Names (total) |
|---------|-------------------|---------------------------|
|         | 18                | 6                         |

**Species with Names (total after re-annotation)**

| Species with annotation / name / SBO term | Useful species with annotation / name / SBO term supported by Saint after re-annotation |
|------------------------------------------|------------------------------------------------------------------------------------------------|
| 13 (72%)                                 | 6 (100%)                                                                                      |

**SBO Terms (total after re-annotation)**

| SBO Terms (total) | SBO terms recovered | New SBO terms |
|-------------------|---------------------|---------------|
| 2                 | 6                   | 4             |

**URIs / Resources (total)**

| URIs / Resources (total) | URIs / Resources after re-annotation | URIs / Resources in useful species supported by Saint | URIs / Resources recovered |
|--------------------------|--------------------------------------|-----------------------------------------------------|---------------------------|
| 20                       | 89                                   | 7                                                   | 5 (71%)                   |

Table 3: Basic information on the annotation present in the original, unmodified version of BIOMD0000000189 and **in bold for the re-annotated version**. This model had no species name attributes, and therefore Saint relies on the id attribute. There were also no SBO terms the original model.

Useful identifiers are those which match actual protein or gene names. The two unrecovered URIs are for the “ARF” species and are described in the main text.

**SPECIES ID**

| MATCH Old | CHANGED New |
|------------|-------------|
| Mdm2       | E3 ubiquitin-protein ligase Mdm2 |
| p53        | Cellular tumor antigen p53        |
| Mdm2,p53   | SBO:0000245               |
| Mdm2,p53   | urn:miriam:uniprot:P04637                 |
| Mdm2,p53   | urn:miriam:uniprot:Q00987               |
| ARF_Mdm2   | SBO:0000245               |
| ARF_Mdm2   | urn:miriam:uniprot:Q8N726                 |
| ARF_Mdm2   | SBO:0000245               |

Table 4: A detailed breakdown of recovered (in the MATCH column) and new or modified annotation (in the CHANGED columns) by Saint for BIOMD0000000189. Only those annotations recognised by Saint are shown in this table. Additionally, new GO terms are omitted for clarity (there are a large number).
4 Comparison Against An Intact Model

This example shows an additional feature of Saint, which illustrates the benefits of searching based on the resource URI of an SBML `species` element. BIOMD0000000032 is annotated in this section, first starting from a stripped model and second starting from the original, intact model. Comparison using URIs is suitable primarily for those models whose IDs and names are uninformative with respect to their biologically-relevant names or descriptions.

Please note that, as in the previous examples, species composed of more than one protein had all of their suggested GO URIs removed. The authors of the model would themselves be able to choose the correct GO terms, but for simplicity GO URIs were only kept when there was just a single UniProt URI per species to ensure no false positives were retained.

4.1 Annotating a Stripped Version of BIOMD0000000032

In this example, we have begun by downloading the SBML Level 2 Version 4 file for the model. If this model is then stripped as described in the earlier examples, a reasonable amount of suitable annotation is retrieved by Saint. This section summarises the annotation received by the stripped version of this model. The next section summarises the improvements in annotation gained by retaining the original model and its MIRIAM URIs.

Basic information on the annotation of both the original and re-annotated BIOMD0000000032 model (when starting from a stripped version of the original model), as well as a summary of the comparison between them, is presented in Table 5. A detailed listing of the modifications made to the model is available in Table 6.

The total number of Saint-supported URIs (those from GO and UniProt) have increased from 7 to 89, with further editing of URIs by the modeller possible if required.

Saint recovered 71% of the URIs from the useful species (5 URIs), as shown in Table 5. Saint also added 84 GO URIs. The original model contained 2 SBO terms, 100% of which were recovered with a further two being added. The original model did not include species names, and Saint has added two, which is one-third of the useful species.

With some species such as “α factor” it was unclear which α factor was intended from the original model, as no further annotation was provided. Therefore one of the most likely candidates was chosen, and only those GO URIs which had the most sources of evidence were kept. The complex names, such as “complexB” were completely uninformative and were therefore not useful.

| Species | 37 |
|---------|----|
| Usefull identifiers | 15 |
| Species with Names (total) | 10 |
| Species with Names (total after re-annotation) | 21 |
| Species names changed | 4 |
| New species names | 11 |
| Species with annotation / name / SBO term | 36 (97%) |
| Useful species with annotation / name / SBO term supported by Saint | 14 (93%) |
| Useful species with annotation / name / SBO term supported by Saint after re-annotation | 15 (100%) |
| SBO Terms (total) | 0 |
| New SBO Terms (total after re-annotation) | 15 |
| URIs / Resources (total) | 104 |
| URIs / Resources supported by Saint | 101 |
| URIs / Resources supported by Saint after re-annotation | 174 |
| URIs / Resources in useful species supported by Saint | 14 |
| URIs / Resources recovered | 14 (100%) |

Table 5: Basic information on the annotation present in the original, unmodified version of BIOMD0000000032, and in bold for the re-annotated version. Useful identifiers are those which match actual protein or gene names.
| SPECIES ID | MATCH            | Old               | CHANGED             |
|------------|------------------|-------------------|---------------------|
| alpha      |                  | α-factor          | Mating factor alpha-1 |
| alpha      |                  |                   | SBO:0000245         |
| alpha      |                  |                   | urn:miriam:uniprot:P01149 |
| Ste2       | SBO:0000245      |                   | Pheromone alpha factor receptor |
| Ste2       | urn:miriam:uniprot:P06842 | Ste2active          | Pheromone alpha factor receptor |
| Ste2       | urn:miriam:uniprot:P06842 |                   | SBO:0000245         |
| Ste3       | SBO:0000245      |                   | Protein STE5        |
| Ste3       | urn:miriam:uniprot:P32917 |                   | SBO:0000245         |
| Ste11      | SBO:0000245      |                   | Serine/threonine-protein kinase STE11 |
| Ste11      | urn:miriam:uniprot:P23561 |                   | SBO:0000245         |
| Ste7       | SBO:0000245      |                   | Serine/threonine-protein kinase STE7 |
| Ste7       | urn:miriam:uniprot:P06784 |                   | SBO:0000245         |
| Fus3       | SBO:0000245      |                   | Mitogen-activated protein kinase FUS3 |
| Fus3       | urn:miriam:uniprot:P16892 |                   | SBO:0000245         |
| Ste20      | SBO:0000245      |                   | Serine/threonine-protein kinase STE20 |
| Ste20      | urn:miriam:uniprot:Q03497 |                   | SBO:0000245         |
| Ste12      | SBO:0000245      |                   | Protein STE12       |
| Ste12      | urn:miriam:uniprot:P13574 | Ste12active          | SBO:0000245         |
| Ste12a     | SBO:0000245      |                   | Protein STE12       |
| Ste12a     | urn:miriam:uniprot:P13574 | Ste12active          | SBO:0000245         |
| Bar1       | SBO:0000245      |                   | Barrierpepsin       |
| Bar1       | urn:miriam:uniprot:P12630 |                   | SBO:0000245         |
| Bar1a      | SBO:0000245      |                   | Barrierpepsin       |
| Bar1a      | urn:miriam:uniprot:P12630 | Bar1active           | SBO:0000245         |
| Far1       | SBO:0000245      |                   | Cyclin-dependent kinase inhibitor FARI |
| Far1       | urn:miriam:uniprot:P21268 |                   | SBO:0000245         |
| Cdc28      | SBO:0000245      |                   | Cell division control protein 28 |
| Cdc28      | urn:miriam:uniprot:P00546 |                   | SBO:0000245         |
| Sst2       | SBO:0000245      |                   | Protein SST2        |
| Sst2       | urn:miriam:uniprot:P11972 |                   | SBO:0000245         |
| Gabc       | SBO:0000245      |                   |                       |
| Gabc       | urn:miriam:uniprot:P08539 |                   |                       |
| Gabc       | urn:miriam:uniprot:P18851 |                   |                       |
| Gabc       | urn:miriam:uniprot:P18852 |                   |                       |
| Gbc        | SBO:0000245      |                   |                       |
| Gbc        | urn:miriam:uniprot:P08539 |                   |                       |
| Gbc        | urn:miriam:uniprot:P18851 |                   |                       |
| Gbc        | urn:miriam:uniprot:P18852 |                   |                       |
| Gbc        | urn:miriam:bind:50058 |                   |                       |
| GbcTP      | urn:miriam:uniprot:P08539 |                   |                       |
| GbcTP      | urn:miriam:uniprot:P18851 |                   |                       |
| GbcTP      | urn:miriam:uniprot:P18852 |                   |                       |
| GbcTP      | urn:miriam:bind:50058 |                   |                       |
| Fus3PP     | urn:miriam:uniprot:P16892 |                   |                       |
| Bar1ax     | urn:miriam:uniprot:P12630 |                   |                       |
| Far1PP     | urn:miriam:uniprot:P21268 |                   |                       |
| Far1U      | urn:miriam:uniprot:P21268 |                   |                       |

Table 6: A detailed breakdown of recovered (in the MATCH column) and new or modified annotation (in the CHANGED columns) by Saint for BIOMD0000000032, starting from a stripped model. Only those annotations recognised by Saint are shown in this table. Additionally, new GO terms are omitted for clarity (there are a large number). Please note that whether the original or new species name is used is entirely at the discretion of the user, and that these particular examples are not necessarily the only choice a user could make.
4.2 Annotating an Intact Version of BIOMD0000000032

When annotating the stripped model, there were a number of instances (e.g., “Fus3PP”) which had no annotation retrieved as the ID and name were unsuitable. In these cases, if we re-run Saint using the intact, original model, we increase the coverage. This increase in coverage is due to Saint’s ability to use existing URLs to retrieve annotation if the IDs and names of the species are uninformative. For instance, with the UniProt URLs available for species such as “GaGTP” and “complexC”, Saint can suggest appropriate species names, GO URLs and SBO terms as it now knows the URLs used to identify the species.

Basic information on the annotation of the intact (rather than stripped) BIOMD0000000032 model is presented in Table 5, together with a summary of how the new annotation compares both with the original model and the re-annotation of the stripped model in Section 4.1. A detailed listing of the modifications made to the model is available in Table 8.

The total number of Saint-supported URLs (those from GO and UniProt) have increased from 101 to 325, with further editing of URLs by the modeller possible if required. The original model contained 0 SBO terms and the initial annotation from Section 4.1 added 15. However, a further 47% were added with the annotation based on the intact model, bringing the total to 21. Further, while the Section 4.1 model was able to add 11 more species names, a further two names were added when starting with the intact model.

| Species with Names (total) | 23 | 10 (130%) | 21 (10%) |
|----------------------------|----|-----------|----------|
| SBO Terms (total)          | 21 | 0         | 15 (47%) |
| URLs / Resources supported by Saint | 325 | 101 (222%) | 174 (87%) |

Table 7: Basic information on the annotation present in the re-annotated BIOMD0000000032 created from the intact version of the model. Useful identifiers are those which match actual protein or gene names. Also shown are the percentage increase in annotations compared to the original model and to the initial re-annotation performed on the stripped model from Section 4.1. The total amount of annotation has increased compared to the re-annotated model based on this stripped starting model. This is due to the ability of Saint to pull information based on URLs even if there are no annotations returned based on id or name attributes.

5 Conclusions

By examining how Saint behaves with both stripped and intact models, we have isolated the two main query types used within the application: id and name querying, and querying based on pre-existing MIRIAM resources (URIs). The use of URIs to retrieve information is one way to get around the variable quality of IDs and names. A number of new techniques are in development to increase the coverage of annotation provided by Saint.

These examples highlight the importance of both a common naming scheme and the use of MIRIAM resources (URIs) for systems biology models. It is difficult for automated procedures such as Saint to disambiguate the IDs and names used to identify species. While a number of checks and automated modifications of ids and names can and are performed, there are many modellers with many different naming conventions. Saint performs well with a minimal amount of data, thus allowing annotation of models that are still at an early stage of development. However, the use of a common naming scheme would improve the coverage of new annotations. Additionally, suitable resource URLs would be even more useful. These MIRIAM-compliant URIs can point not only to entries in databases, but also terms in ontologies, thus providing the means to tag elements in a model with precise information about those elements. If elements such as species were annotated, then their names would conceivably become merely a helpful aid for humans, which is their main purpose in the first place.
| SPECIES ID | MATCH | Old                           | New                           |
|------------|-------|-------------------------------|-------------------------------|
| GaGTP      |       | GaGTP                         | Guanine nucleotide-binding protein alpha-1 subunit SBO:0000245 |
| GaGTP      | urn:miriam:uniprot:P08539 |                               |                               |
| GaGTP      |       |                               |                               |
| GaGDP      |       | GaGDP                         | Guanine nucleotide-binding protein alpha-1 subunit SBO:0000245 |
| GaGDP      | urn:miriam:uniprot:P08539 |                               |                               |
| GaGTP      |       |                               |                               |
| Fus3PP     |       | Fus3PP                        | Mitogen-activated protein kinase FUS3 SBO:0000245 |
| Fus3PP     | urn:miriam:uniprot:P16892 |                               |                               |
| Fus3PP     |       |                               |                               |
| Bar1aex    |       | Bar1activeEx                  | Barrierpepsin SBO:0000245     |
| Bar1aex    | urn:miriam:uniprot:P12630 |                               |                               |
| Bar1aex    |       |                               |                               |
| Far1PP     |       | Far1PP                        | Cyclin-dependent kinase inhibitor FARI SBO:0000245 |
| Far1PP     | urn:miriam:uniprot:P21268 |                               |                               |
| Far1PP     |       |                               |                               |
| Far1U      |       | Far1ubiquitin                 | Cyclin-dependent kinase inhibitor FARI SBO:0000245 |
| Far1U      | urn:miriam:uniprot:P21268 |                               |                               |
| Far1U      |       |                               |                               |

Table 8: A detailed breakdown of recovered (in the MATCH column) and new or modified annotation (in the CHANGED columns) by Saint for BIOMD000000032, starting from an intact model. Saint discovers all annotation found in Table 6, plus the annotation shown in this table. Only those annotations recognised by Saint are shown in this table. Additionally, new GO terms are omitted for clarity (there are a large number). Please note that whether the original or new species name is used is entirely at the discretion of the user, and that these particular examples are not necessarily the only choice a user could make.