# BioModels Curation

Introduction

BioModels is an online repository for mathematical equations that model biological processes in peer-reviewed literature. Each paper has a record in the repository identified by a unique alphanumeric code (BIOMD##....), and each record is split up into six sections (denoted by the tabs on the top ribbon). An individual record should look something like this:
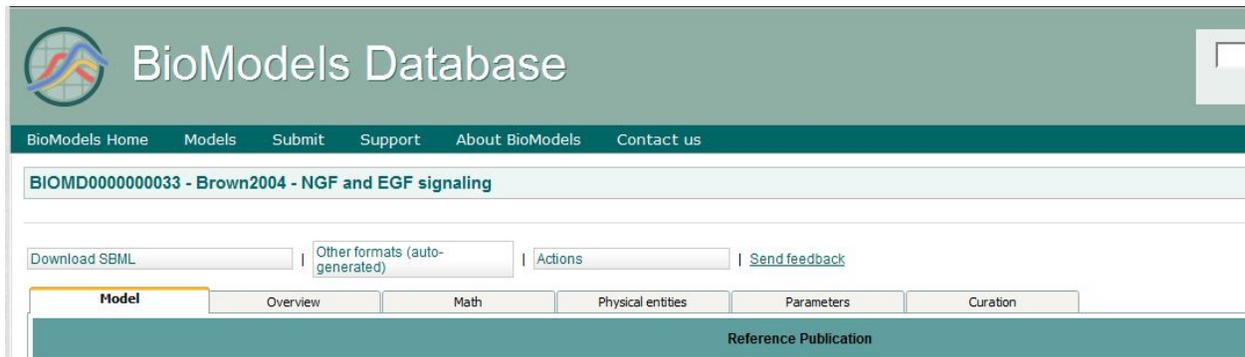
There will be several entities that you will tag in the BioModel after you have validated it in COPASI, *the most important are the molecular species and the reactions that they are involved in*. The qualifiers you will use to describe these entities will derive from a metadata language called SBML (Systems Biology Markup Language). The following are the most common:

Bqmodel: is → directly links to a record. (Example: Bqmodel: is BIOMD0000000033).

bqmodel : isDescribedBy → where the entity is described in. (Example: bqmodel: isDescribedBy PMID##)

Bqbiol: is → directly identifies the entity. Could be a species or a reaction. (Example: bqbiol: is Epidermal Growth Factor Binding.)

Bqbiol: isVersionOf → describes the 'parent' in the relationship. (Example: if bqbiol:is Epidermal Growth Factor Binding, then bqbiol: isVersionOf positive regulation of epidermal growth factor-activated receptor activity).

As you can see, there is a formula to creating these tags (qualifier followed by a term). The terms are taken from the following databases:

UniProt
Gene Ontology
ChEBI
KEGG
NCBI Gene
Interpro

Resources on these databases are available in other documents. For now, we will go over the tabs described above.

Model

This tab describes the information about the paper (metadata). You will find the citation, keywords used to describe the model, and notes on the simulation of the model.

A breakdown of the qualifiers used to describe the model is follows:

Bqbiol: hasPart → Used to describe the subjects that involve the model. In this case, the model is describing the epidermal growth factor receptor signaling pathway (taken from the Gene Ontology database) and the neurotrophin TRK receptor signaling pathway (also taken from the Gene Ontology database).

Bqbiol: hasProperty → In this instance, the qualifier is used to describe the type of mathematical model. The Mathematical Modelling Ontology (MAMO) term is 'Ordinary Differential Equation Model', which is the record that is linked to the 'MAMO_0000046' hyperlink.

Bqbiol: hasTaxon → This qualifier denotes the Taxonomy that the mathematical model works under. Since the model talks about cellular pathways, the term is 'cellular organisms' (which can be found in UniProt Taxonomy).

Bqbiol: isVersionOf → As stated above, this term with this qualifier describes the 'parent' of the relationship. The epidermal growth factor receptor signaling pathway and the neurotrophin TRK receptor signaling pathway are both part of the ERK1 and ERK2 cascade.

Overview

This section lists all of the reactions, species involved in the reactions, and the compartment the reaction or process is happening in (usually 'Cell'). Each name is linked to its respective area in the "Math" tab.

Math

This is where the reactions are described and annotated. You may use the following qualifiers to annotate the reaction: bqbiol: is (insert term that best fits reaction), bqbiol: isVersionOf (insert term that fits parent reaction, if applicable).

The reactions can usually be found in KEGG and Reactome. If there is a complex reaction that involves multiple parts, then it is appropriate to tag all relevant reactions to describe the complex reaction.

It should also be noted that Gene Ontology can also be used to tag reactions. "If the reaction is a generic one, sometimes a Gene Ontology term is sufficient." (EMBL-EBI, 2016).

Physical Entities

The species that are part of the reactions are listed in this tab. You may use the following qualifiers to annotate the species:

Bqbiol: is → used when the term exactly defines the species.
Bqbiol: hasPart → used when the term describes part of the species.
Bqbiol: isVersionOf → used when the term is the larger entity the species is part of.

Terms can be taken from the following databases:

UniProtKB
ChEBI
KEGG Compound
InterPro
Gene Ontology

Parameters

The Parameters tab is dedicated to the association of appropriate numeric values to the given species. In the notes section, you should indicate whether the value is based on: direct measurements, a number taken from another source, or some other method.

Curation

This is where you can input your notes on the model, specifically evidence of model validation from COPASI.

References
http://www.ebi.ac.uk/biomodels-main/BIOMD0000000033#metaid_0000054

http://co.mbine.org/standards/qualifiers
http://www.ebi.ac.uk/biomodels-main/annotationtips