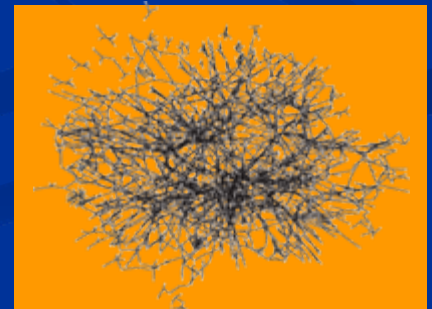
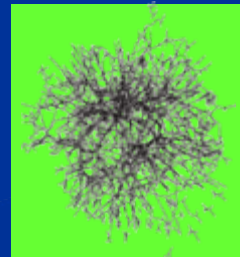
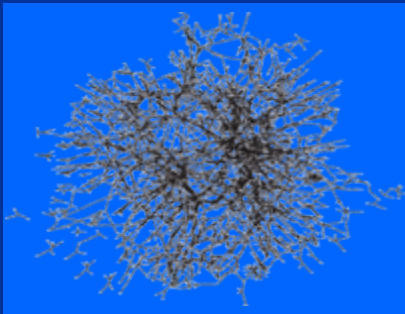


Progress report: SBML Level 3 extension “FBA”

Brett G. Olivier¹ & Frank T. Bergmann² et al

¹VU University Amsterdam, Netherlands

²University of Washington, USA



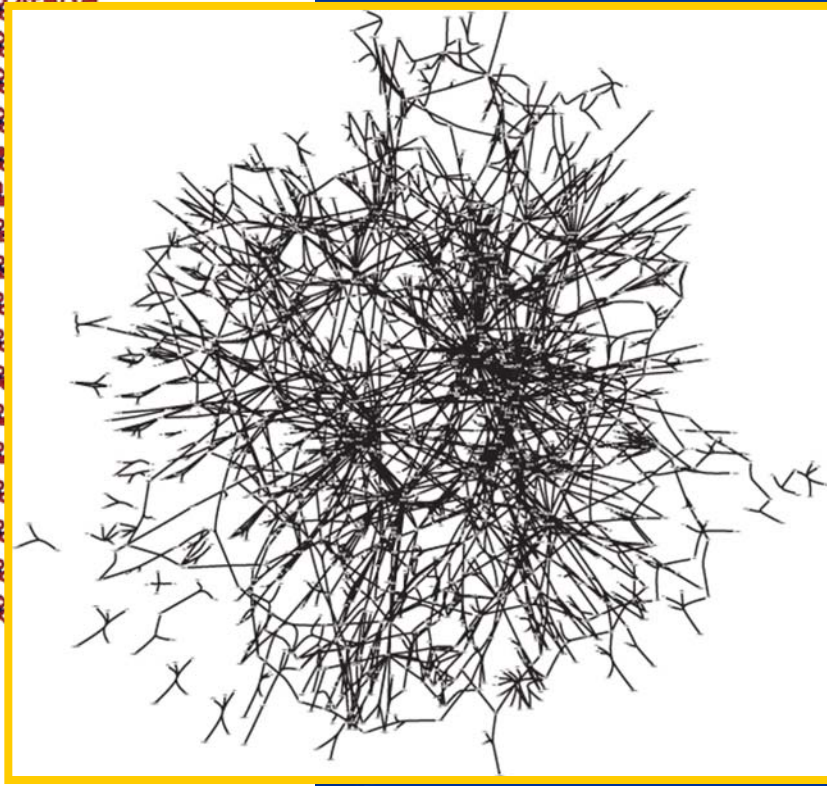
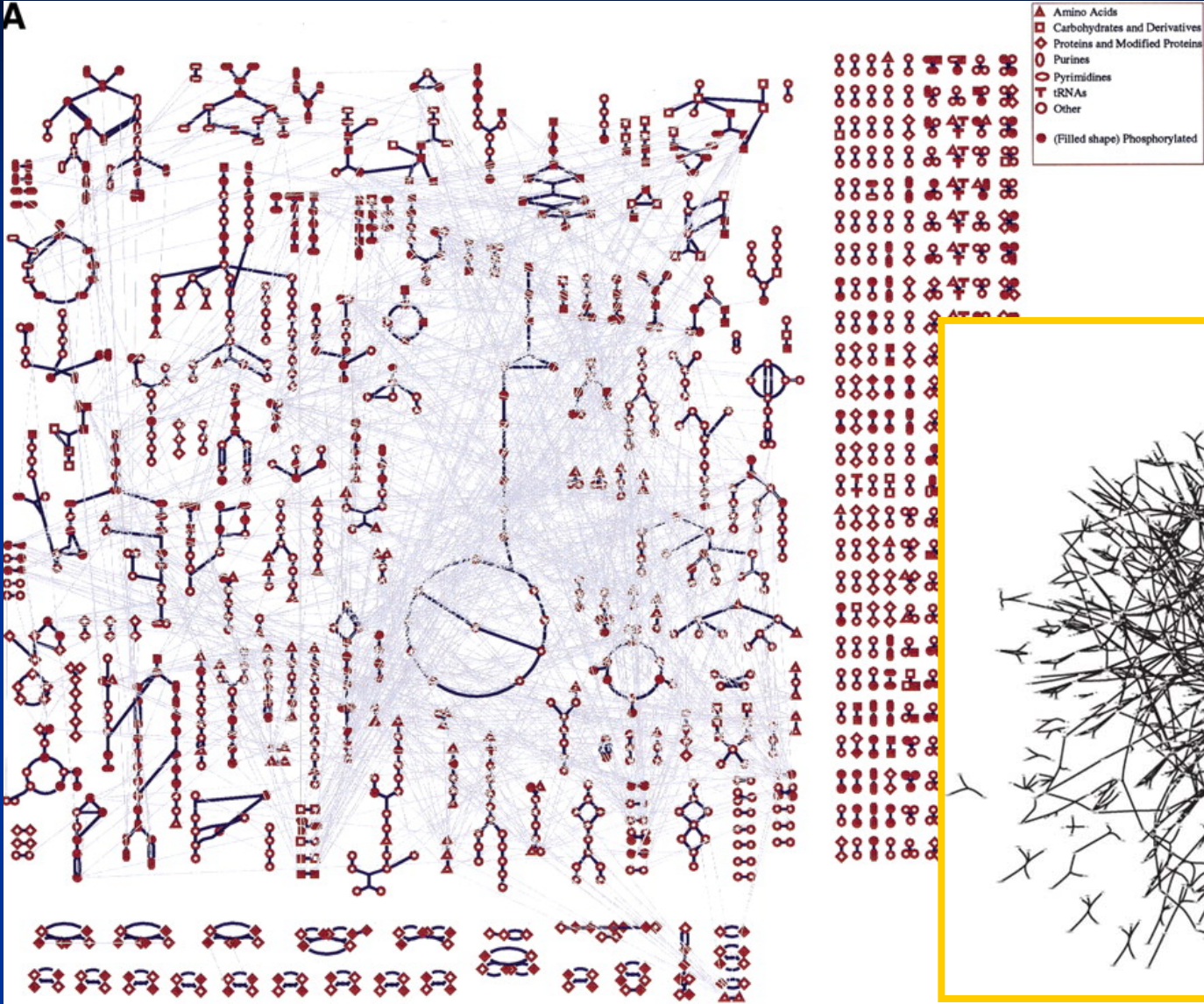
Background / history

Initial draft 2010

First revision and beyond

Escherichia coli metabolic network

A



Ouzounis et al. *Genome Res.* 2000. 10: 568-576

Vallabhajosyula et al. *Bioinformatics* 2006 22:346-353

Flux balance analysis

- **Optimise** a specific property: typically biomass production
- Constrained by the **stoichiometry** (N)

Maximize
biomass

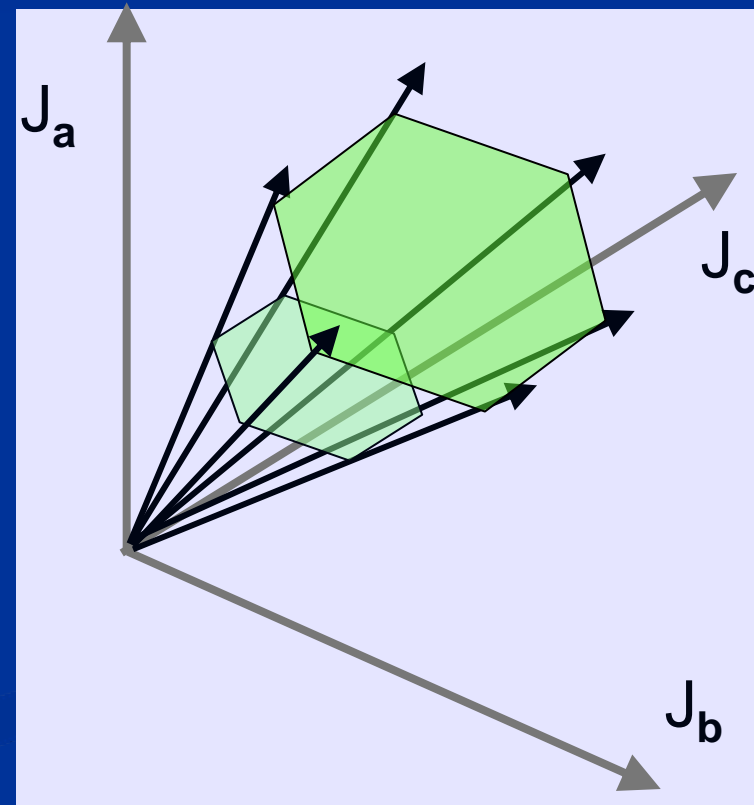
Subject to
 $NJ = 0$

With bounds

$$0 \leq J_{\text{irrev}} \leq \text{inf}$$

$$-\text{inf} \leq J_{\text{rev}} \leq \text{inf}$$

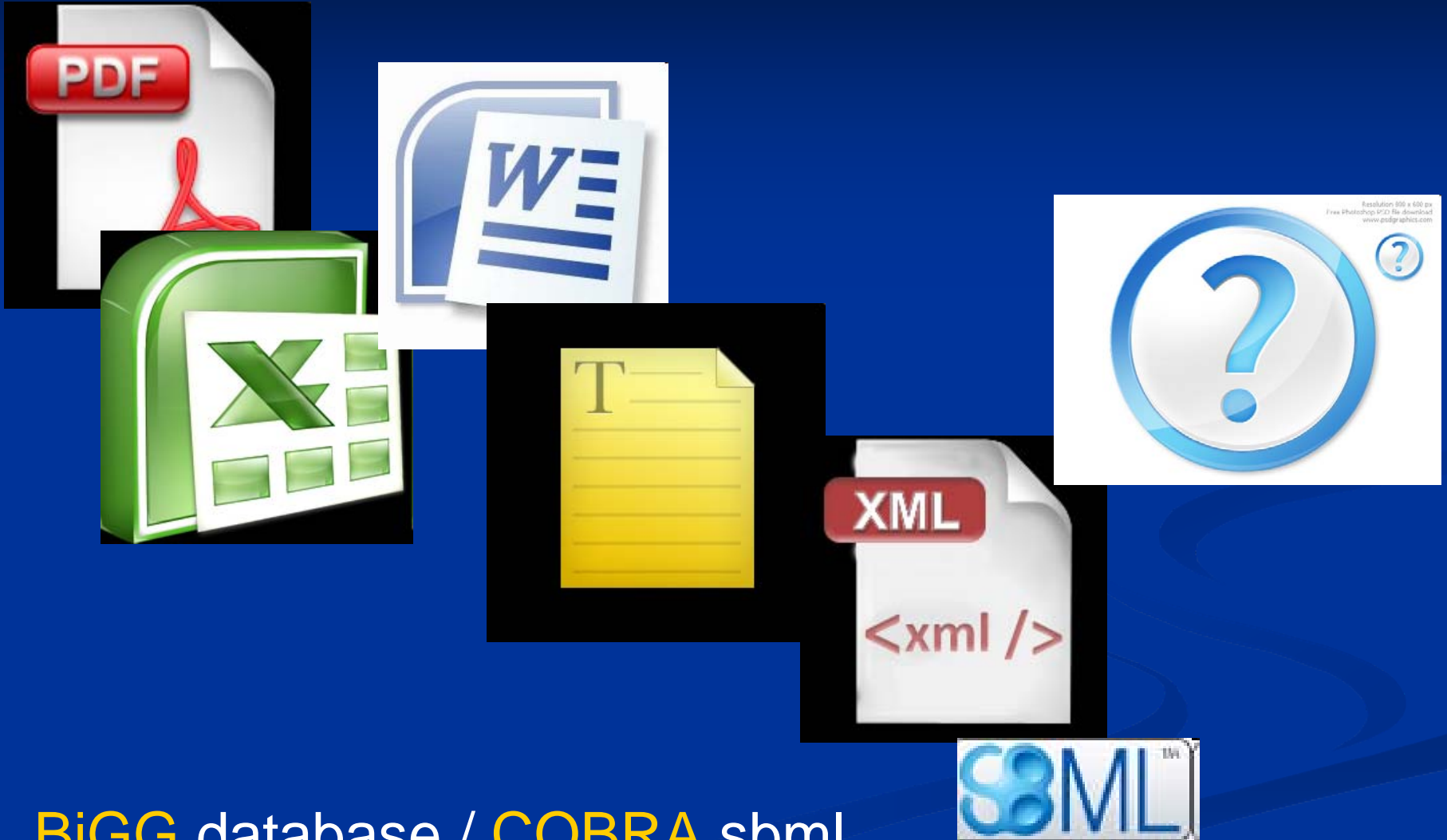
$$\text{l.b} \leq J_n \leq \text{u.b}$$



Critical points

- FBA is an *operation* performed on a model, essentially specified by the objective function, unlike a kinetic model *definition*
 - Flux bounds may be biochemical constraints on the model (or based on the simulation)
- **Fundamental question:** should SBML also allow
 - exchange of simulation results (flux values), esp. for FBA?
 - Specification of a model *instance*, in terms of a particular objective function

Genome scale models



- **BiGG** database / **COBRA** sbml
 - currently *the way* to exchange GSR's

Genome reconstructions 2009

Available reconstructions (all formats): 55+ **

BiGG database (<http://bigg.ucsd.edu/>) 10 GSR's (COBRA)

nature
biotechnology

Nature Biotechnology 28 (9) , pp 977–982, September 2010

2010

COBRA

High-throughput generation, optimization and analysis of genome-scale metabolic models

Christopher S Henry¹, Matthew DeJongh², Aaron A Best³, Paul M Frybarger^{2,3}, Ben Linsay⁴ & Rick L Stevens^{4,5}

Genome-scale metabolic models have proven to be valuable for predicting organism phenotypes from genotypes. Yet efforts to develop new models are failing to keep pace with genome sequencing. To address this problem, we introduce the Model SEED, a web-based resource for high-throughput generation, optimization and analysis of genome-scale metabolic models. The Model SEED integrates existing methods and introduces techniques to automate nearly every step of this process, taking ~48 h to reconstruct a metabolic model from an assembled genome sequence. We apply this resource to generate 130 genome-scale metabolic models representing a taxonomically diverse set of bacteria. Twenty-two of the models were validated against available gene essentiality and Biolog data, with the average model accuracy determined to be 66% before optimization and 87% after optimization.

Background / history

Initial draft 2010

First revision and beyond

SBML³ FBA

- Extends SBML by adding classes for **flux bounds** and **objective functions**
- Implemented as an **SBML L2V4** annotation

```
<?xml version="1.0" encoding="utf-8"?>
<sbml xmlns="http://www.sbml.org/sbml/level2" level="2" version="1" >
  <model id="BranchMultipleCycle" name="BranchMultipleCycle">
    <annotation>
      <fba:fluxBalance xmlns:fba="http://www.sbml.org/sbml/level3/version1/fba/version1" >
        <fba:listOfConstraints>
          <fba:constraint fba:reaction="J0" fba:operation="equal" fba:value="10" />
        </fba:listOfConstraints>
        <fba:listOfObjectives fba:activeObjective="obj1">
          <fba:objective id="obj1" fba:type="maximize">
            <fba:listOfFluxes>
              <fba:fluxObjective fba:reaction="J8" fba:coefficient="1" />
            </fba:listOfFluxes>
          </fba:objective>
        </fba:listOfObjectives>
      </fba:fluxBalance>
    </annotation>
  </model>
</sbml>
```

Bergmann, Frank and Olivier, Brett. (2010) *SBML Level 3 Package Proposal: Flux*.
Nature Precedings <<http://dx.doi.org/10.1038/npre.2010.4236.1>>

SBML³ FBA

- **Implemented** as SBML L2 extension :
 - **Systems Biology WorkBench** (FBA module)



- **PyscesCBM** (MultiState FBA module)



- **ConverterTool** (Converts BiGG/COBRA into L2 + FBA extension)

Flux Bounds

Bounds a model flux (SBML reaction) with a value.
Multiple bounds on single reaction possible.

```
<listOfFluxBounds>  
  
<fluxBound id="fb1" reaction="Glc_in"  
  operation="lessEqual" value="10"/>  
  
<fluxBound id="fb2" reaction="R_HEX1"  
  operation="greaterEqual" value="0"/>  
  
</listOfFluxBounds>
```

Objective function

Defines one or more optimization targets.

```
<listOfObjectives activeObjective="ObjFun1">  
  <objective id="ObjFun1" type="maximize">  
    <listOfFluxes>  
      <fluxObjective reaction="PFK" coefficient="1"/>  
    </listOfFluxes>  
  </objective>  
</listOfObjectives>
```

Extended species

```
<species id="glc"  
  name="D-Glucose"  
  compartment="Cytosol"  
  fba:chemicalEquation="C6H12O6"  
  fba:charge="0">
```

- First draft captured the basic **model structure** but did not deal with annotation.
- Simply left **COBRA** annotation as is: a reaction <note>

```
<notes>  
<html:p>Confidence Level: 0</html:p>  
<html:p>SUBSYSTEM: Glycolysis/Gluconeogenesis</html:p>  
<html:p>GENE ASSOCIATION: (b3916) or (b1723)</html:p>  
...  
</notes>
```

- Most important is **GENE ASSOCIATION**

Background / history

Initial draft 2010

First revision and beyond

Species annotation

chemicalEquation replaced with a CHEBI annotation

```
<species metaid="atp" id="atp" name="ATP_C10H12N5O13P3"
  compartment="Cytosol" initialConcentration="1"
  boundaryCondition="true" fba:charge="-4">
  <annotation>
    <rdf:RDF>
      <rdf:Description rdf:about="#atp">
        <bqbiol:is>
          <rdf:Bag>
            <rdf:li
  rdf:resource="urn:miriam:obo:chebi:CHEBI%3A30616" />
          </rdf:Bag>
        </bqbiol:is>
      </rdf:Description>
    </rdf:RDF>
  </annotation>
</species>
```

Gene association

- Add **genes** to model as boundary **species** with proper MIRIAM annotation
- Add a **listOfGeneAssociations** to FBA

```
<listOfGeneAssociations>  
  <geneAssociation id="ga3" reaction="R_PFK">  
    <association>  
      <and>  
        <gene>b3916</gene>  
        <gene>b1723</gene>  
      </and>  
    </association>  
  </geneAssociation>  
</listOfGeneAssociations>
```

Open questions / issues

- Proceed to rewrite proposal into official document?
- **fluxBounds** or something more general?
- **Objective function** could **SED-ML** be used (L3FBA)?
- The **gene association** fix?
- Anything else ... ?

Thanks

- Herbert Sauro
- Neil Swainston
- Kieran Smallbone
- sbml-discuss
- Bas Teusink
- Frank Bruggeman
- Steven Kelk
- ... and others