2. Data sources for metabolic models

David Fell

OXFORD BROOKES UNIVERSITY

C1net gas fermentation

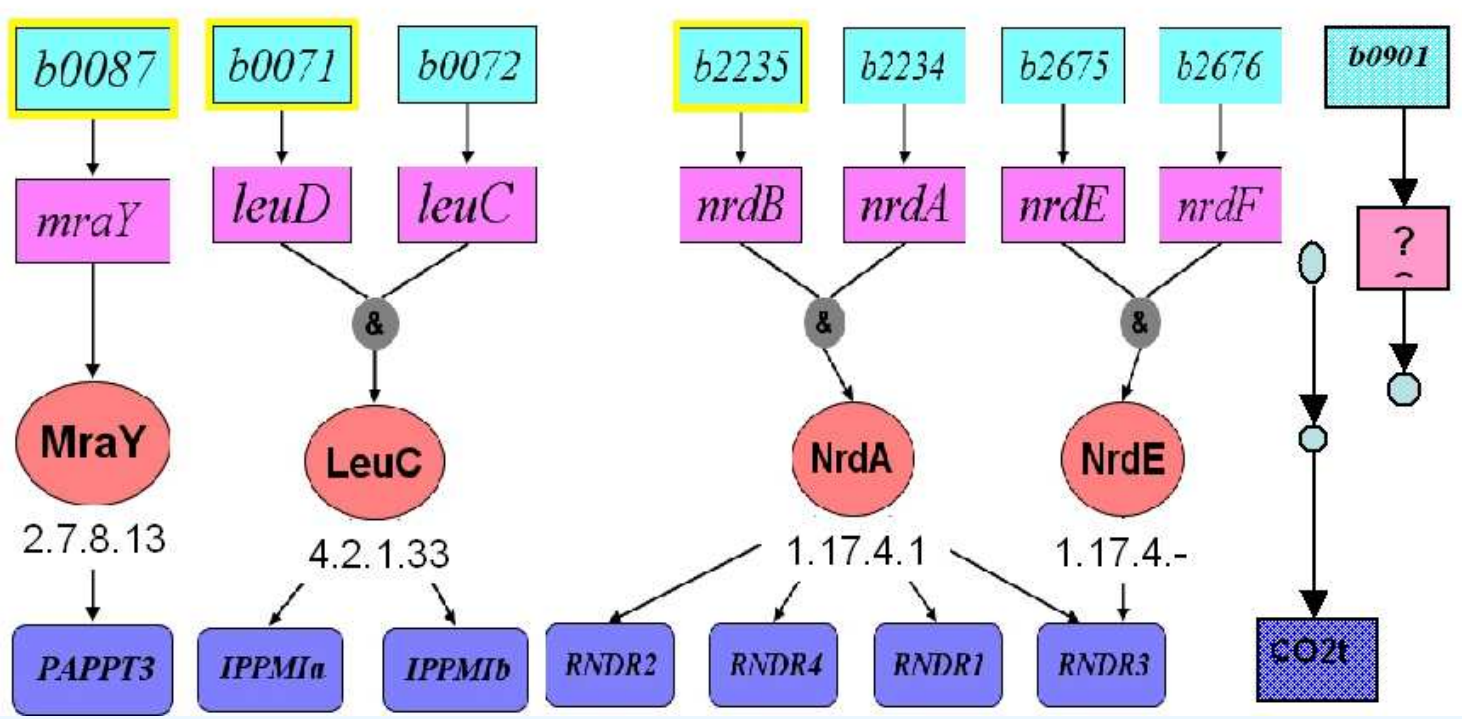
email: dfell@brookes.ac.uk

http://mudshark.brookes.ac.uk

January 16, 2018

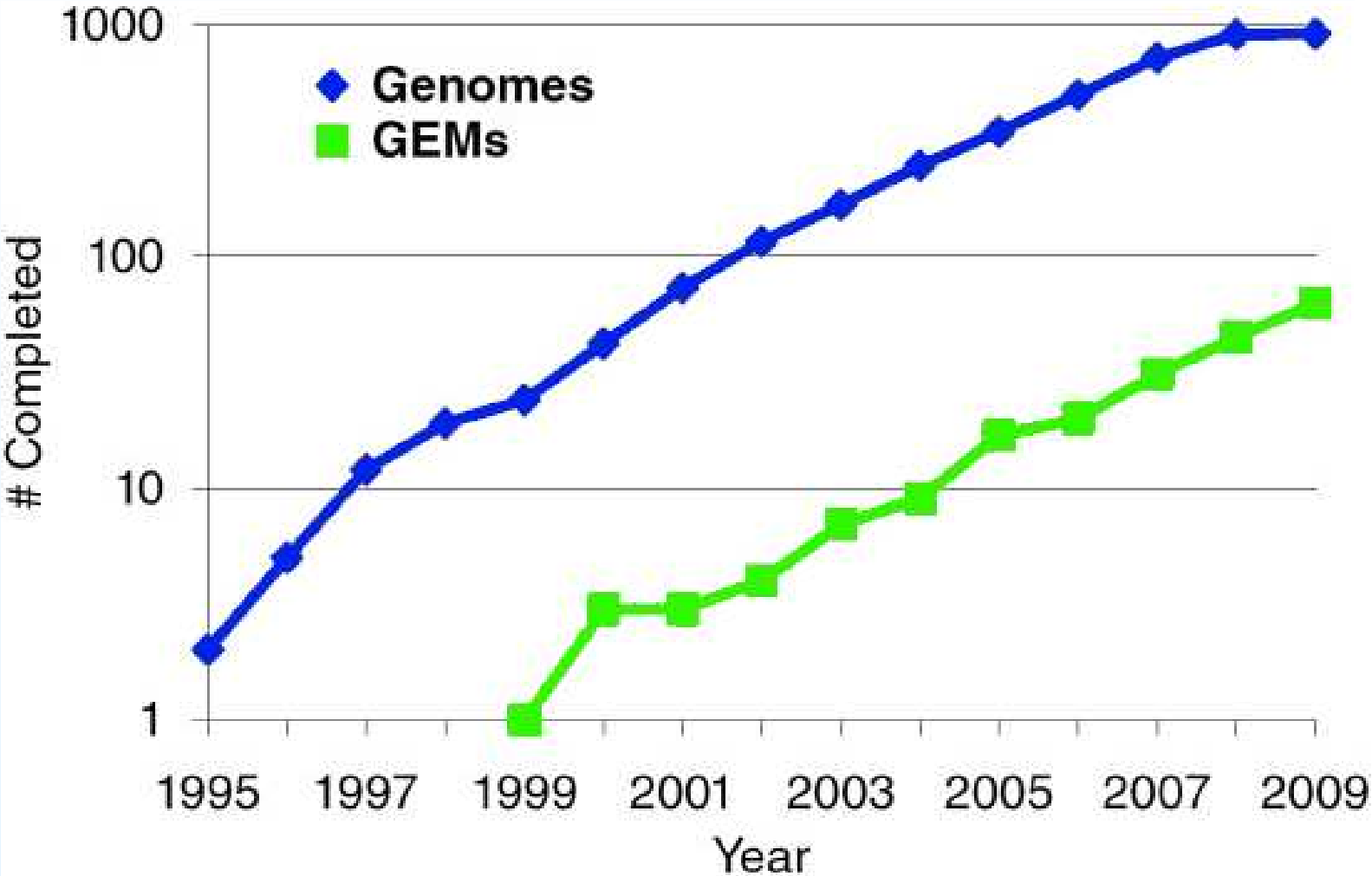
Genes to Reactions

Genome to reaction mapping
- Genes to Reactions
- The Interpretation Gap
- The Reaction List
- Data Sources
- BioCyc
- KEGG
- IntEnz: Enzymes at EBI
- Expasy Enzyme
- BRENDA
- Other Resources
- Requirements for a Structural Model



The Interpretation Gap

Genome to reaction mapping
- Genes to Reactions
- The Interpretation Gap
- The Reaction List
- Data Sources
- BioCyc
- KEGG
- IntEnz: Enzymes at EBI
- Expasy Enzyme
- BRENDA
- Other Resources
- Requirements for a Structural Model



Milne et al (2009) Biotech. J, **4**, 1653-1670

The Reaction List

Genome to reaction mapping
- Genes to Reactions
- The Interpretation Gap
- The Reaction List
- Data Sources
- BioCyc
- KEGG
- IntEnz: Enzymes at EBI
- Expasy Enzyme
- BRENDA
- Other Resources
- Requirements for a Structural Model

Where does the data for writing the reaction list for a metabolic model come from?

- Biochemical literature: books, reviews, journal articles.
- Genome databases plus annotation plus enzyme database.
  - BioCyc: http://biocyc.org/
  - KEGG: http://www.genome.jp/kegg/
  - IntEnz at http://www.ebi.ac.uk/intenz/ (now the primary source) or EXPASY Enzyme: https://enzyme.expasy.org/
  - Brenda: http://www.brenda-enzymes.info

The Reaction List

Genome to reaction mapping
- Genes to Reactions
- The Interpretation Gap
- The Reaction List
- Data Sources
- BioCyc
- KEGG
- IntEnz: Enzymes at EBI
- Expasy Enzyme
- BRENDA
- Other Resources
- Requirements for a Structural Model

Where does the data for writing the reaction list for a metabolic model come from?

- Biochemical literature: books, reviews, journal articles.

- Genome databases plus annotation plus enzyme database.

  - BioCyc: http://biocyc.org/
  - KEGG: http://www.genome.jp/kegg/
  - IntEnz at http://www.ebi.ac.uk/intenz/ (now the primary source) or EXPASY Enzyme: https://enzyme.expasy.org/
  - Brenda: http://www.brenda-enzymes.info

The Reaction List

Genome to reaction mapping
- Genes to Reactions
- The Interpretation Gap
- The Reaction List
- Data Sources
- BioCyc
- KEGG
- IntEnz: Enzymes at EBI
- Expasy Enzyme
- BRENDA
- Other Resources
- Requirements for a Structural Model

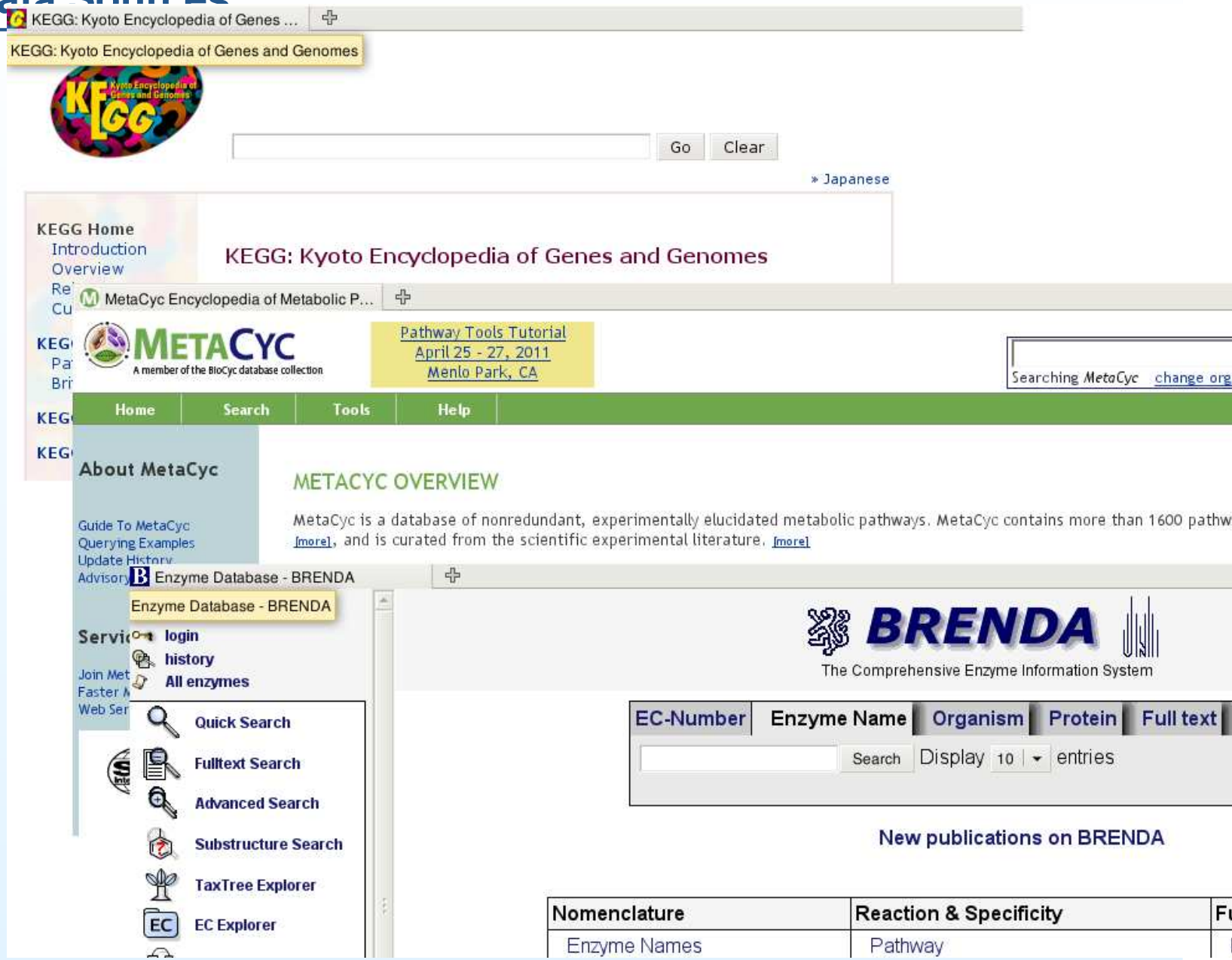
Where does the data for writing the reaction list for a metabolic model come from?

- Biochemical literature: books, reviews, journal articles.
- Genome databases plus annotation plus enzyme database.
  - BioCyc: http://biocyc.org/
  - KEGG: http://www.genome.jp/kegg/
  - IntEnz at http://www.ebi.ac.uk/intenz/ (now the primary source) or EXPASY Enzyme: https://enzyme.expasy.org/
  - Brenda: http://www.brenda-enzymes.info

Data Sources

Genome to reaction mapping

- Genes to Reactions
- The Interpretation Gap
- The Reaction List
- Data Sources
- BioCyc
- KEGG
- IntEnz: Enzymes at EBI
- Expasy Enzyme
- BRENDA
- Other Resources
- Requirements for a Structural Model



BioCyc

Genome to reaction mapping
- Genes to Reactions
- The Interpretation Gap
- The Reaction List
- Data Sources
- BioCyc
- KEGG
- IntEnz: Enzymes at EBI
- Expasy Enzyme
- BRENDA
- Other Resources
- Requirements for a Structural Model

http://biocyc.org/

Example entry: Pyruvate kinase

KEGG

Genome to reaction mapping
- Genes to Reactions
- The Interpretation Gap
- The Reaction List
- Data Sources
- BioCyc
- KEGG
- IntEnz: Enzymes at EBI
- Expasy Enzyme
- BRENDA
- Other Resources
- Requirements for a Structural Model

Kyoto Encyclopaedia of Genes and Genomes

http://www.genome.jp/kegg/

Example entry: TCA cycle

IntEnz: Enzymes at EBI

Genome to reaction mapping
- Genes to Reactions
- The Interpretation Gap
- The Reaction List
- Data Sources
- BioCyc
- KEGG
- IntEnz: Enzymes at EBI
- Expasy Enzyme
- BRENDA
- Other Resources
- Requirements for a Structural Model

IntEnz at http://www.ebi.ac.uk/intenz/

Example entry: Pyruvate kinase

Expasy Enzyme

Genome to reaction mapping
- Genes to Reactions
- The Interpretation Gap
- The Reaction List
- Data Sources
- BioCyc
- KEGG
- IntEnz: Enzymes at EBI
- Expasy Enzyme
- BRENDA
- Other Resources
- Requirements for a Structural Model

ENZYME database at https://enzyme.expasy.org/

Example entry: Pyruvate kinase; 2.7.1.40

BRENDA

Genome to reaction mapping
- Genes to Reactions
- The Interpretation Gap
- The Reaction List
- Data Sources
- BioCyc
- KEGG
- IntEnz: Enzymes at EBI
- Expasy Enzyme
- BRENDA
- Other Resources
- Requirements for a Structural Model

Brenda: http://www.brenda-enzymes.info/

Example entry: Pyruvate kinase

Other Resources

Genome to reaction mapping
- Genes to Reactions
- The Interpretation Gap
- The Reaction List
- Data Sources
- BioCyc
- KEGG
- IntEnz: Enzymes at EBI
- Expasy Enzyme
- BRENDA
- Other Resources
- Requirements for a Structural Model

Alternatives to reciprocal BLAST homology searching

(Deficiencies: Chinese whispers; failure to detect bi– or multifunctional– enzymes.)

Enzyme-finding tools based on sequence signatures:

- PRIAM
- theSEED (part of KBASE: `http://kbase.us`)
- EfiCAZ

Requirements for a Structural Model

Genome to reaction mapping
- Genes to Reactions
- The Interpretation Gap
- The Reaction List
- Data Sources
- BioCyc
- KEGG
- IntEnz: Enzymes at EBI
- Expasy Enzyme
- BRENDA
- Other Resources
- Requirements for a Structural Model

A valid structural model requires a reaction list in which:

- Every internal metabolite is both produced and consumed, either by an enzyme or a transporter.
- The metabolites that exchange with the environment or between compartments, usually via transporters, can be identified.
- Every enzyme/transporter in the list has a role in at least one pathway.
- Arguably, the metabolic network should form a single connected component.

Genome to reaction mapping
- Genes to Reactions
- The Interpretation Gap
- The Reaction List
- Data Sources
- BioCyc
- KEGG
- IntEnz: Enzymes at EBI
- Expasy Enzyme
- BRENDA
- Other Resources
- Requirements for a Structural Model

Requirements for a Structural Model

A valid structural model requires a reaction list in which:

- Every internal metabolite is both produced and consumed, either by an enzyme or a transporter.
- The metabolites that exchange with the environment or between compartments, usually via transporters, can be identified.
- Every enzyme/transporter in the list has a role in at least one pathway.
- Arguably, the metabolic network should form a single connected component.

Genome to reaction mapping
- Genes to Reactions
- The Interpretation Gap
- The Reaction List
- Data Sources
- BioCyc
- KEGG
- IntEnz: Enzymes at EBI
- Expasy Enzyme
- BRENDA
- Other Resources
- Requirements for a Structural Model

Requirements for a Structural Model

A valid structural model requires a reaction list in which:

- Every internal metabolite is both produced and consumed, either by an enzyme or a transporter.
- The metabolites that exchange with the environment or between compartments, usually via transporters, can be identified.
- Every enzyme/transporter in the list has a role in at least one pathway.
- Arguably, the metabolic network should form a single connected component.

Requirements for a Structural Model

Genome to reaction mapping
- Genes to Reactions
- The Interpretation Gap
- The Reaction List
- Data Sources
- BioCyc
- KEGG
- IntEnz: Enzymes at EBI
- Expasy Enzyme
- BRENDA
- Other Resources
- Requirements for a Structural Model

A valid structural model requires a reaction list in which:

- Every internal metabolite is both produced and consumed, either by an enzyme or a transporter.
- The metabolites that exchange with the environment or between compartments, usually via transporters, can be identified.
- Every enzyme/transporter in the list has a role in at least one pathway.
- Arguably, the metabolic network should form a single connected component.