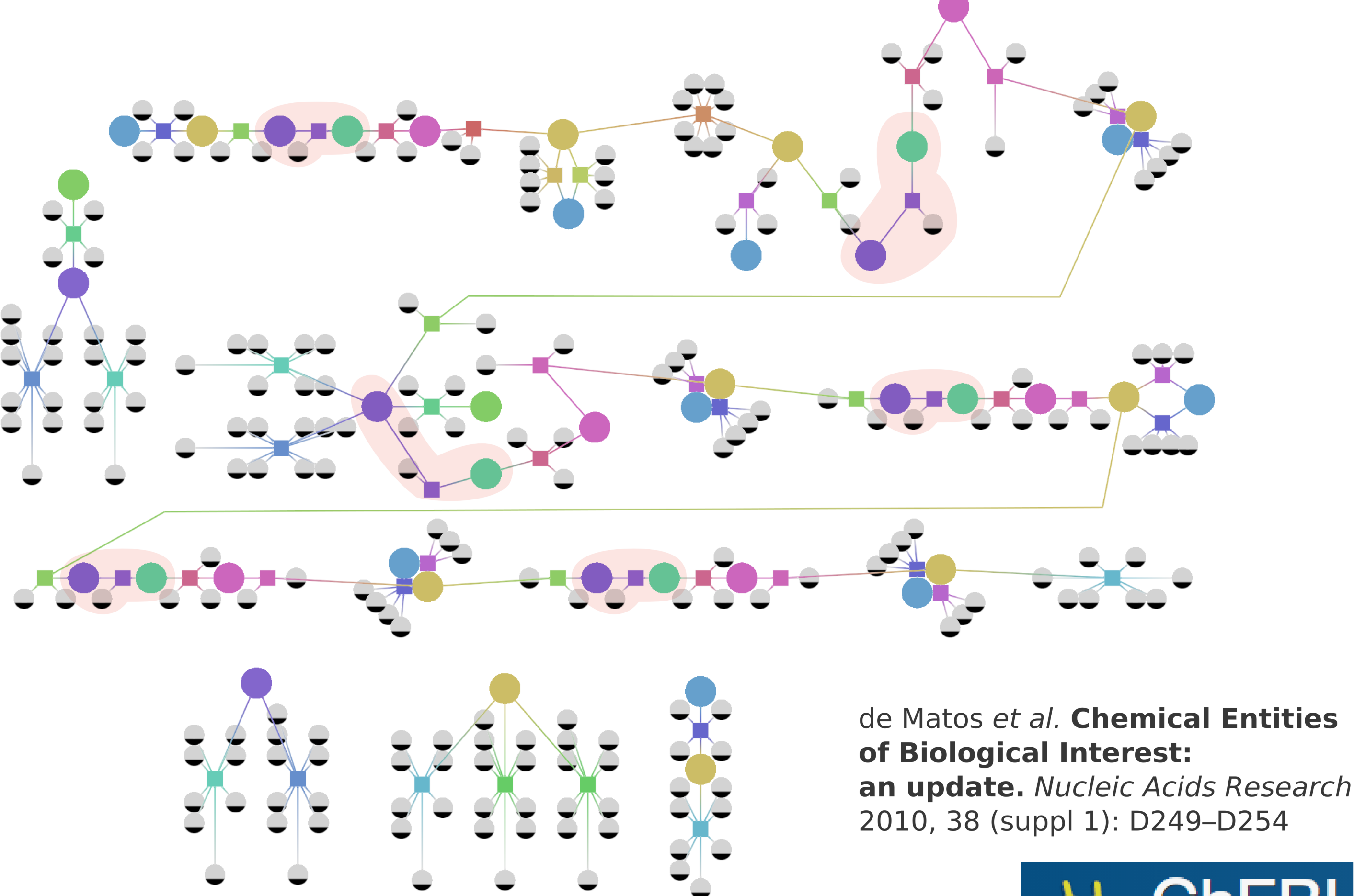


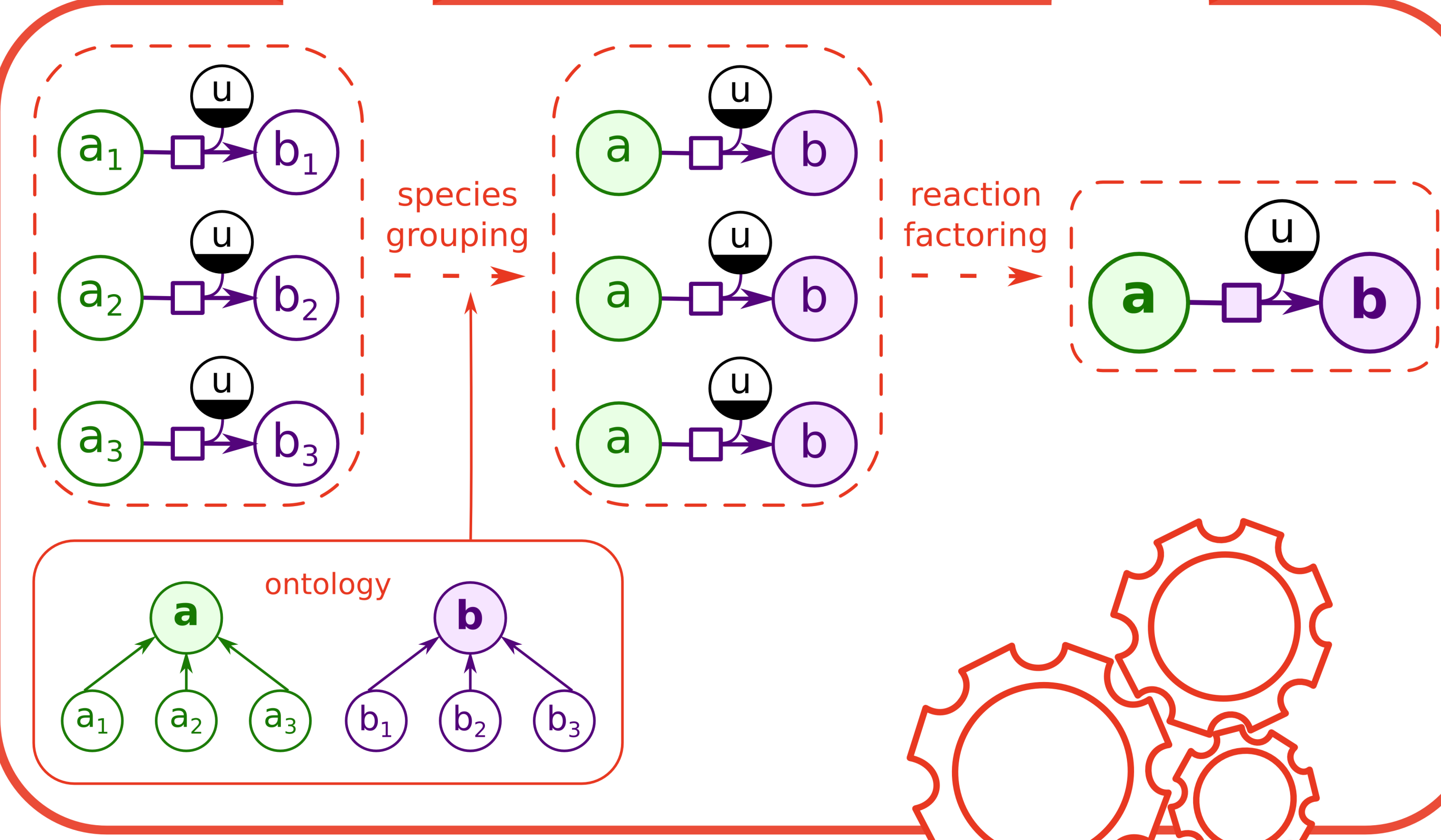
# Knowledge-Based Generalization of Metabolic Networks: An Application Study

Anna Zhukova, David James Sherman  
anna.zhukova@inria.fr, david.sherman@inria.fr

Loira et al. **A genome-scale metabolic model of the lipid-accumulating yeast *Yarrowia lipolytica***. *BMC Systems Biology* 2012, 6:35



de Matos et al. **Chemical Entities of Biological Interest: an update**. *Nucleic Acids Research* 2010, 38 (suppl 1): D249–D254



## Metabolic network generalization method

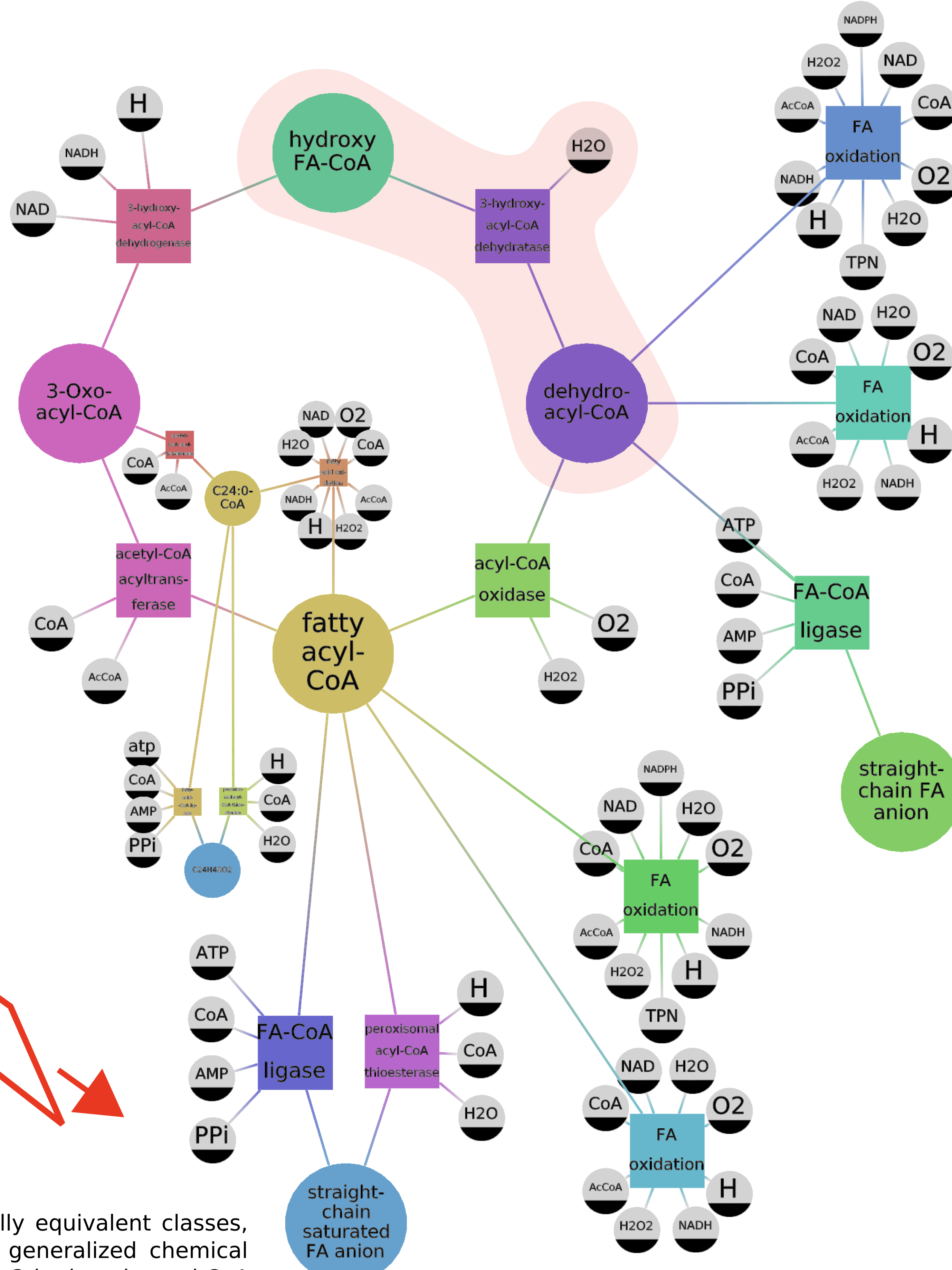
The generalization method groups chemical species present in the network into semantically equivalent classes, based on their hierarchical relationships in the ChEBI ontology, and merges them into a generalized chemical species. For instance, (S)-3-hydroxydecanoyl-CoA, (S)-3-hydroxyhexacosanoyl-CoA and (S)-3-hydroxylauroyl-CoA species can be generalized into hydroxy fatty acyl-CoA.

After the species generalization, reactions that share the same generalized reactants and the same generalized products, are factored together into a generalized reaction.

## Generalization of the *Yarrowia lipolytica* fatty acid oxidation network

The generalized network (in the center) describes the  $\beta$ -oxidation of fatty acids pathway, happening inside the peroxime compartment of the yeast *Yarrowia lipolytica*, in a more generic way: Fatty acyl-CoA (yellow) is transformed into dehydroacyl-CoA (violet), then into hydroxyacyl fatty acyl-CoA (green), 3-ketoacyl-CoA (magenta), and back to fatty acyl-CoA (with a shorter carbon chain). The initial network (top left) describes the same process in more details, specifying those reactions for each of the fatty acyl-CoA species present in the organisms' cell (e.g. decanoyl-CoA, dodecanoyl-CoA, etc.).

The more precise network is needed for simulation, while the more general one is clearer to a human, and reveals the main properties of organism's metabolism.



## Generalized network comparison

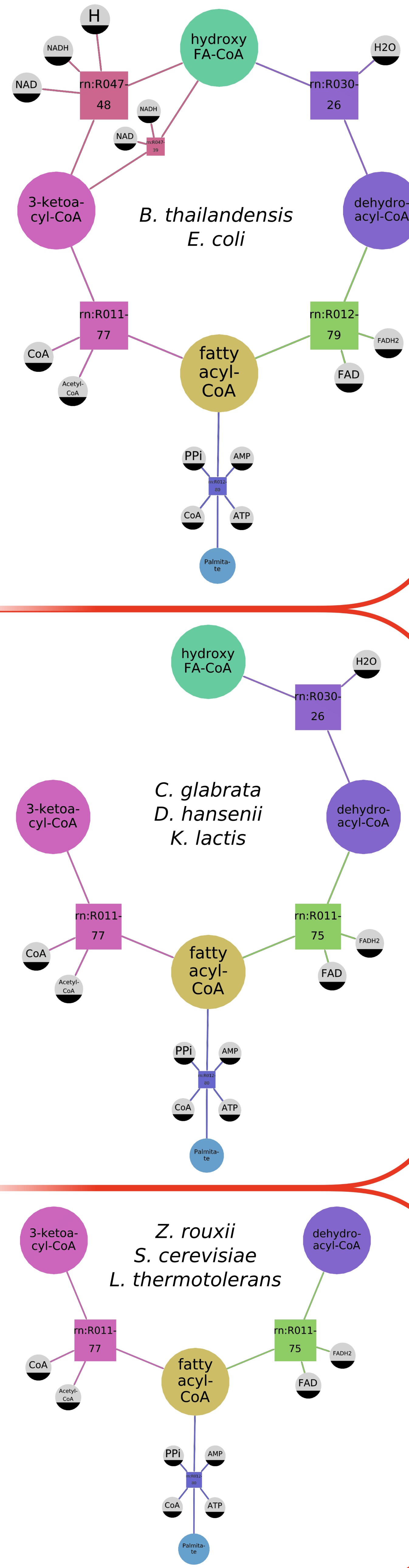
Along with  $\beta$ -oxidation of fatty acids of *Yarrowia lipolytica*, we have generalized (on theright) this pathway for several other yeasts: *Candida glabrata* (BMID45737), *Debaryomyces hansenii* (BMID61271), *Kluyveromyces lactis* (BMID109830), *Lachancea thermotolerans* (BMID43788), *Saccharomyces cerevisiae* (BMID89673), and *Zygosaccharomyces rouxii* (BMID110595); and bacteria: *Burkholderia thailandensis* (BMID103487) and *Escherichia coli* (strain 55989 / EAEC) (BMID83160). The networks were taken from the path2model project (<http://www.ebi.ac.uk/biomodels-main/path2models>).

In all the yeast generalized networks, but the one of *Y. lipolytica*,  $\beta$ -oxidation is not a cycle. This is caused by the missing reactions operating with hydroxy fatty acyls-CoA (green).

The generalized network of *Y. lipolytica* highlights the fact that there is a particularity concerning C24:0-CoA (tetracosanoyl-CoA) (yellow): There exists a "short-cut" reaction (orange), producing it directly from another fatty acyl-CoA (yellow), avoiding the usual four-reaction  $\beta$ -oxidation chain. This reaction is not present in the other networks.

The bacteria networks show a reaction alternative to the standart hydroxyacyl-CoA dehydrogenase, that does not produce a proton.

The generalization of a metabolic network highlights its particularities, and makes erroneous reactions and gaps easier to grasp for a human.



The figures were produced using Tulip:

Auber D. Tulip — A Huge Graph Visualization Framework. In *Graph Drawing Software, Mathematics and Visualization*, Springer Berlin Heidelberg 2004: 105–126

Zhukova A, Sherman D J. **Knowledge-based generalization of metabolic models**. (submitted)