
RetroPath2.0: A retrosynthesis workflow for metabolic engineering and metabolome completion

Mathilde Koch^{*1}, Baudoin Delépine², Thomas Duigou¹, Pablo Carbonell³, and Jean-Loup Faulon^{1,4}

¹MICrobiologie de l'ALImentation au Service de la Santé humaine (MICALIS) – AgroParisTech, Institut national de la recherche agronomique (INRA) : UMR1319 – F-78350 JOUY-EN-JOSAS, France

²Institut de biologie systémique et synthétique (ISSB) – Génopole, Université d'Evry-Val d'Essonne : EA4527, CNRS : UPS3509 – 5 rue Henri Desbruères F-91030 ÉVRY CEDEX, France

³SYNBIOCHEM center, University of Manchester (UoM) – 131 Princess Street Manchester M1 7DN, UK, United Kingdom

⁴Institut de biologie systémique et synthétique (ISSB) – Génopole, Université d'Evry-Val d'Essonne : EA4527, CNRS : FRE3561 – 5 rue Henri Desbruères F-91030 ÉVRY CEDEX, France

Abstract

Synthetic biology applied to industrial biotechnology is transforming the way we produce chemicals. In order to expand the chemical repertoire for the production of next generation compounds, a major effort is required in the development of novel design tools that target chemical diversity through rapid and predictable protocols. Addressing that goal involves retrosynthesis approaches that explore the chemical biosynthetic space. Here, we present RetroPath2.0, an automated open source workflow for retrosynthesis based on generalized reaction rules that perform the retrosynthesis search from chassis to target through an efficient and well-controlled protocol. Its easiness of use and the versatility of its applications make of this tool a valuable addition into the biological engineer bench desk. We show through several examples the application of the workflow to biotechnological relevant problems, including the identification of alternative biosynthetic routes, the enumeration of aminoglycosides structures maximizing antibacterial activity and the development of biosensors. We also use our workflow on an *E. coli* model to complete *E. coli* metabolome, with novel molecules generated using promiscuous enzymatic reaction rules. These novel molecules are searched on the MS spectra of an *E. coli* cell lysate interfacing our workflow with OpenMS through the KNIME analytics platform. Because it is open source and freely available on MyExperiment.org, workflow community contributions should likely expand further the features of the tool, even beyond the use cases presented here.

Keywords: Retrosynthesis, metabolic engineering, biosensors, metabolome, CAD tools

^{*}Speaker