DynCubeProd: Dynamic Solution Space Division-based Methods for Calculating Reaction Deletion Strategies for Constraint-based Metabolic Networks for Substance Production

数理生物情報

Metabolic engineering is a DNA recombination-based technology proposed in 1991 to improve the designated substance production and the cell properties by manipulating and introducing specific biochemical reactions. In many cases, current metabolic engineering technology focuses on the utilization of microorganisms. In metabolic engineering analysis, metabolic pathways in organisms are often represented by metabolic networks, in which nodes represent metabolite molecules and biochemical reactions. Any two metabolites (biochemical reactions) cannot be directly connected, and a metabolite must be connected to at least two biochemical reactions. The biochemical reactions can be irreversible or reversible. Nodes of external metabolites form the input and output nodes of the entire network. Constraint-based modeling is a mathematical method to identify the best solution within a set of possible choices subject to pre-specified constraints. In the constraint-based models of metabolic networks, the cell growth reaction and the target metabolite production reactions are of particular interest. The cell growth reaction has been virtually designed to simulate the efficient conversion of uptake resources into cellular energy and chemical components, which support cell growth in response to selection pressure to construct the system in the most plausible physiological state. The target metabolite production reaction produces a chemical of interest. Growth coupling is a fundamental design principle in metabolic engineering and computational strain design. The purpose of growth coupling is to make the target metabolite a mandatory by-product of the cell growth reaction. Constraint-based modeling methods, such as linear programming (LP) and mixed integer linear programming, are widely used effective optimization techniques.

Flux balance analysis (FBA) is a crucial method to analyze large-scale constraint-based metabolic networks and computing design strategies for strain production in metabolic engineering by leveraging LP and MILP techniques. However, as it is often non-straightforward to obtain such design strategies to produce valuable metabolites, many tools have been proposed based on FBA. Among them, GridProd, which divides the solution space into small squares by focusing on the cell growth rate and the target metabolite production rate to efficiently find the reaction deletion strategies, was extended to CubeProd, which divides the solution space into small cubes. However, as GridProd and CubeProd naively divide the solution space into equal sizes, even places where solutions are unlikely to exist are examined. To address this issue, we introduce dynamic solution space division methods based on CubeProd for faster computing by avoiding searching in places where the solutions do not exist. We applied the proposed method DynCubeProd to iJO1366, which is a genome-scale constraint-based model of Escherichia coli. Compared with CubeProd, DynCubeProd significantly accelerated the calculation of the reaction deletion strategy for each target metabolite production. In addition, under the anaerobic condition of iJO1366, DynCubeProd could obtain the reaction deletion strategies for almost 40% of the target metabolites that the elementary flux vector-based method, which is one of the most effective methods in existence, could not.