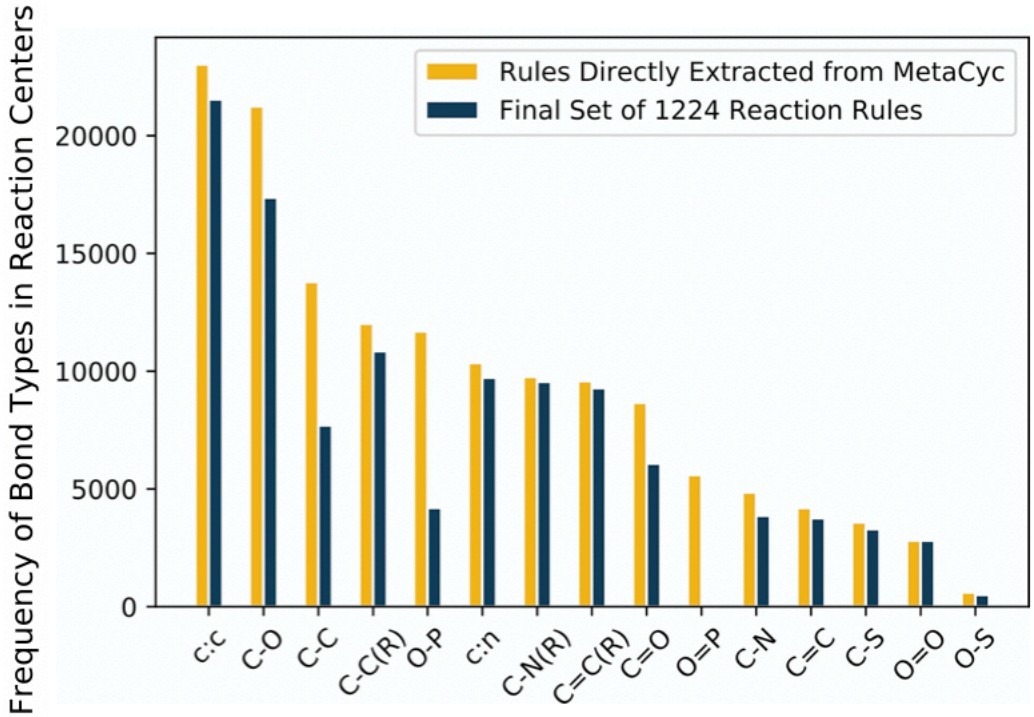
Supplementary Figures

Curating a Comprehensive Set of Enzymatic Reaction Rules for Efficient Novel Biosynthetic Pathway Design

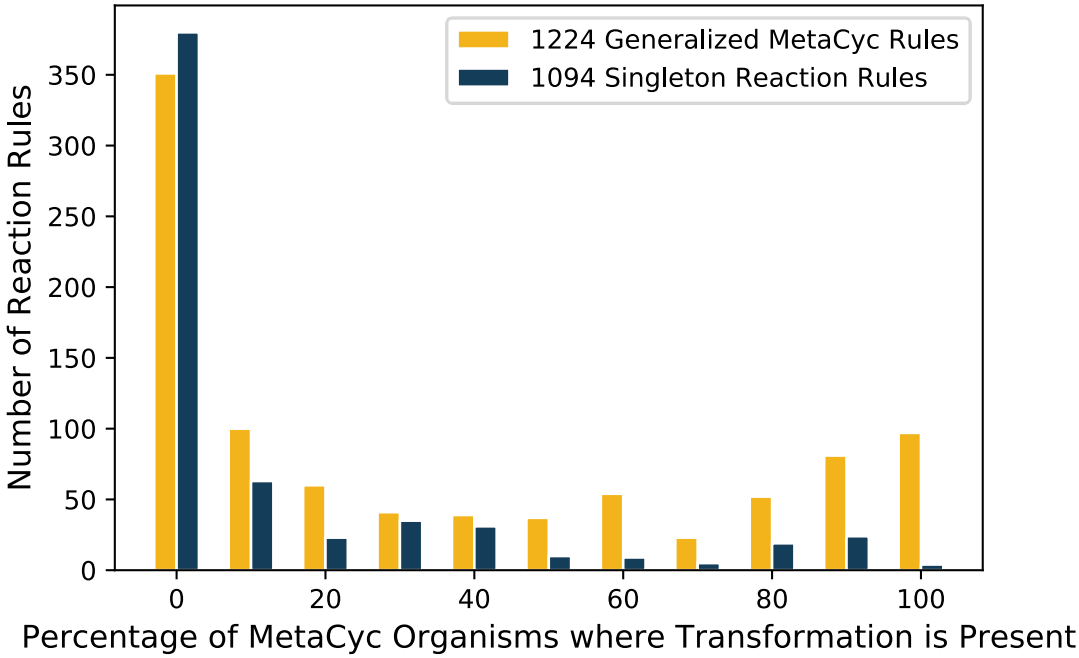
Zhuofu Ni, Andrew E. Stine, Keith E.J. Tyo, Linda J. Broadbelt

Department of Chemical and Biological Engineering, Northwestern University, Evanston, IL 60208, USA



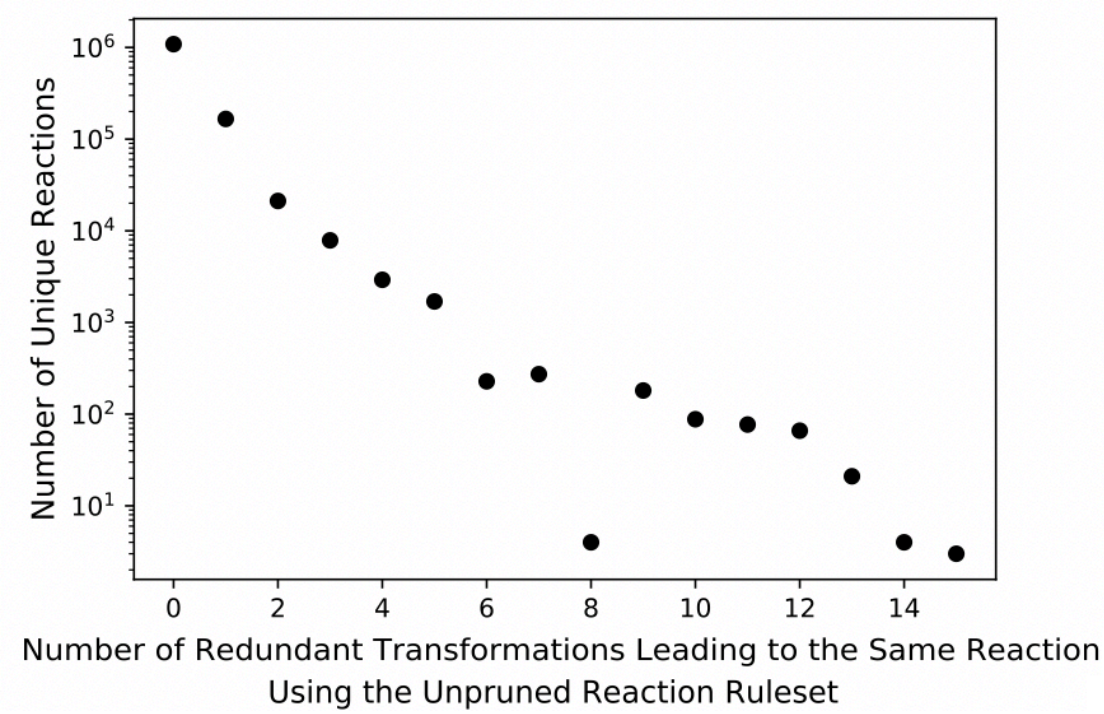
SI Fig 1. Reaction Rule Pruning Leads to More Concise Recognition of Reaction Centers for MetaCyc Reactions

A comparison of number of bonds in the reaction centers for all MetaCyc reactions (with common enzymatic transformations) indicates that fewer bond changes are now required to describe these reactions as a result of reaction rule pruning. The 15 most frequent bond changes are shown, categorized by atoms involved, bond types, and whether bonds are linear or cyclic.



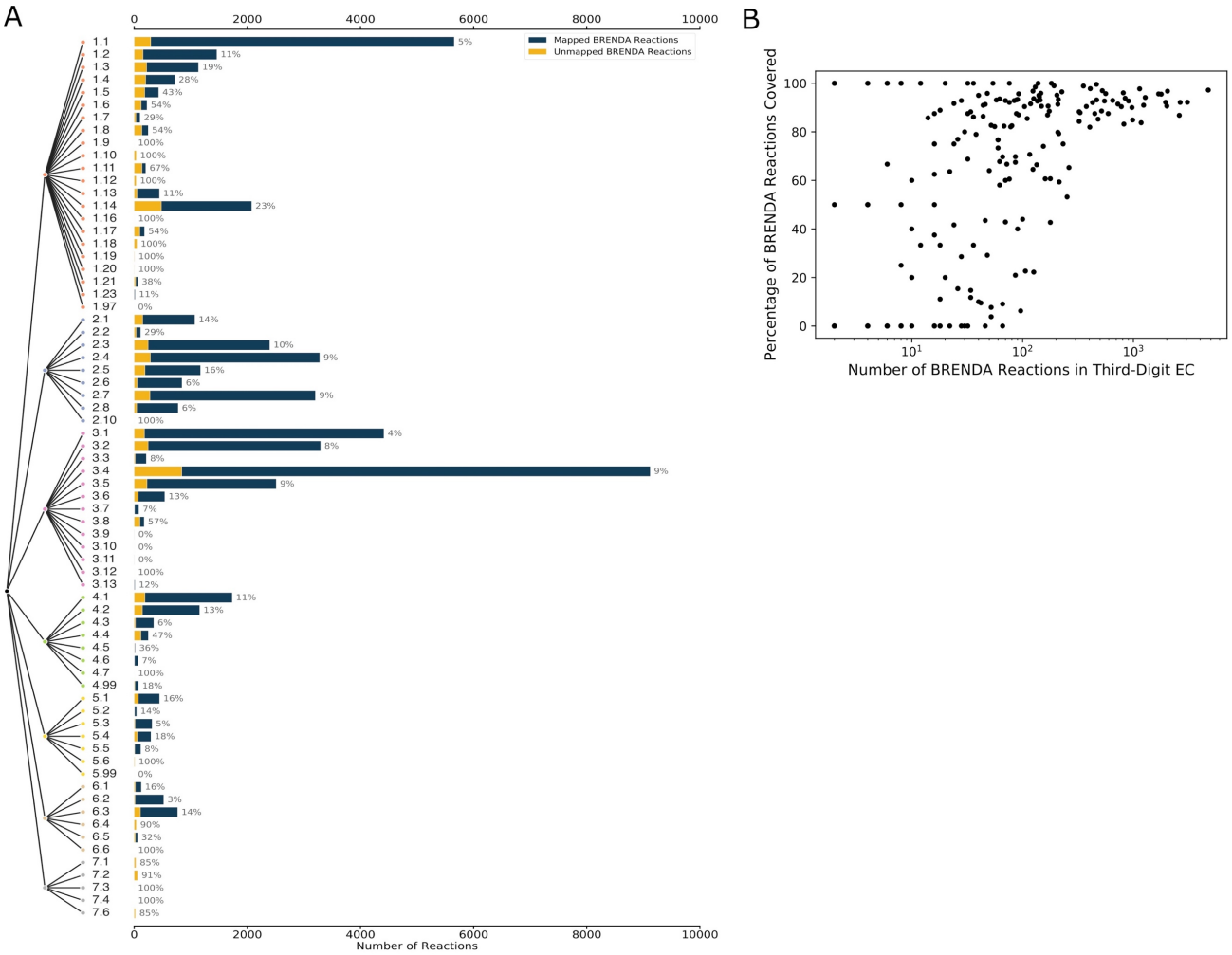
SI Fig 2. 1094 Filtered Out Singleton Reaction Rules Represent Enzymatic Transformations in Fewer MetaCyc Organisms than 1224 Generalized Ruleset

A histogram comparing distribution of reaction rules, in terms of percentage of MetaCyc organisms where transformations are present, indicates that the majority of 1094 filtered out singleton reaction rules represents peripheral metabolism present in less than 5% of MetaCyc organisms, whereas the finalized set of 1224 rules is more capable of describing enzymatic transformations common to all species.



SI Fig 3. Predicting Novel Reactions on iJO1366 Metabolites Using the Minimal Ruleset Reduces Number of Redundant Transformations Leading to the Same Reaction

Both the minimal set of 1224 enzymatic reaction rules and the set of 4722 unpruned rules directly extracted from MetaCyc were applied onto all iJO1366 metabolites for one iteration, in order to assess the reduction in level of redundancy due to the rule pruning process. In both rulesets, reaction rules involving only one non-cofactor reactant and one non-cofactor product were considered. 1.34 million unique reactions were enumerated with the minimal ruleset, while the unpruned ruleset predicted more than 0.26 million redundant transformations shown above. In some cases, there were up to 15 redundant transformations that predicted the same unique reaction. These redundancies were removed when applying the minimal ruleset.



SI Fig 4. BRENDA Reactions Covered vs Uncovered by Final Set of Reaction Rules

Categorizing covered vs uncovered BRENDA reactions by the EC nomenclature reveals that uncovered reactions predominantly originate from smaller groups of EC subclasses, while larger subclasses of EC tend to be well-covered. (A) shows the number of reactions and percentage of uncovered reactions, grouped up to the second-digit of EC. (B) shows the correlation between the number of reactions associated with each third-digit EC number and the coverage of these reactions by the final rule set.