

Bachelorarbeit

Herleitung von Gen-Protein-Reaktionsregeln

1 Hintergrund

Mit der Verbreitung von Hochdurchsatzmessverfahren vollzieht sich ein kontinuierlicher Wandel in der biologischen Forschung. Durch die vielfältigen sogenannten Omik-Daten können Zellen ganzheitlich beobachtet werden. Dadurch konnte sich die Wissenschaft der Systembiologie als ein neuer Forschungszweig fest etablieren. Diese verfolgt das Ziel, biologische Systeme durch prädikative Computer-Modelle zu beschreiben. Langfristig sollen diese Modelle genutzt werden, um zahlreiche neue Hypothesen systematisch zu testen, wodurch beispielsweise Tierversuche eingespart werden können, indem unwirksame Versuche effektiv ausgeschlossen werden können.

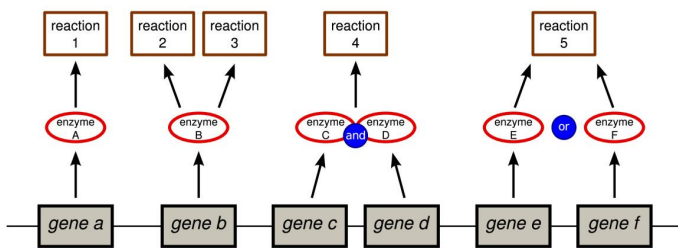


Abbildung 1: Gene-Protein-Reaktionsregeln (aus [6])

Aufgebaut sind die Modelle als Netzwerke biochemischer Reaktionen, die sowohl intrazelluläre als auch extrazelluläre Prozesse beschreiben. Enzyme katalysieren die überwiegende Mehrzahl der intrazellulären Reaktionen, welche wiederum durch genetische Elemente codiert sind. Kennt man den genauen Zusammenhang zwischen Genen und Enzymen, kann man effizient simulieren, wie sich bestimmte Mutationen oder die Unverfügbarkeit bestimmter Transkripte auf die metaboli-

schen Fähigkeiten der betrachteten Zelle auswirken.

Viele Enzyme bestehen aus mehreren Untereinheiten, die z. T. durch verschiedene Gene codiert sind. Dadurch kann das Zusammenspiel der Gene als logischer Ausdruck formuliert werden, welcher als Gen-Protein-Reaktionsregel bezeichnet wird (siehe schematische Abbildung neben dem Text).

2 Zielsetzung

In dieser Arbeit soll ein Verfahren etabliert werden, um automatisch von relevanten Datenbanken (wie z. B. KEGG [7] oder UniProt [1]) Gen-Protein-Reaktionsregeln (sog. GPR-Regeln) herzuleiten und diese mit systembiologischen Modellen zu verknüpfen, welche im SBML-Dateiformat [5, 10] vorliegen. Auch sollen mit diesem Verfahren bereits vorhandene GPR-Regeln überprüft und nötigenfalls korrigiert werden können. Relevante Vorarbeiten zum Thema wie z. B. der Algorithmus von Cardoso u. a. [3], TIGER [6], MERLIN [4], AutoKEGGRec [8], Pathway Tools [9], oder SuBliMinaL [12] sollen berücksichtigt werden. Als Anwendungsfall sollen zunächst Modelle wichtiger Humanpathogene wie *Staphylococcus aureus*, *Pseudomonas aeruginosa*, *Moraxiella catarrhalis* oder *Haemophilus influenzae* erweitert bzw. überprüft werden.

3 Voraussetzungen

Interesse an systembiologischer Modellierung sowie Kenntnisse im Programmieren mit einer geeigneten SBML-Bibliothek (z. B. libSBML [2] in Python oder Java™ mit JSBML [11]).

Literatur

- [1] BATEMAN, Alex: UniProt: A worldwide hub of protein knowledge. In: *Nucleic Acids Research* 47 (2019), Nr. D1, S. D506–D515. <http://dx.doi.org/10.1093/nar/gky1049>. – DOI 10.1093/nar/gky1049. – ISSN 13624962
- [2] BORNSTEIN, Benjamin J. ; KEATING, Sarah M. ; JOURAKU, Akiya ; HUCKA, Michael: LibSBML: An API library for SBML. In: *Bioinformatics* 24 (2008), Nr. 6, S. 880–881. <http://dx.doi.org/10.1093/bioinformatics/btn051>. – DOI 10.1093/bioinformatics/btn051. – ISSN 13674803
- [3] CARDOSO, João ; VILAÇA, Paulo ; SOARES, Simão ; ROCHA, Miguel: An algorithm to assemble gene-protein-reaction associations for genome-scale metabolic model reconstruction. In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 7632 LNBI (2012), S. 118–128. http://dx.doi.org/10.1007/978-3-642-34123-6_11. – DOI 10.1007/978-3-642-34123-6_11. – ISBN 9783642341229
- [4] DIAS, Oscar ; ROCHA, Miguel ; FERREIRA, Eugénio C. ; ROCHA, Isabel: Reconstructing genome-scale metabolic models with merlin. In: *Nucleic Acids Research* 43 (2015), Nr. 8, S. 3899–3910. <http://dx.doi.org/10.1093/nar/gkv294>. – DOI 10.1093/nar/gkv294. – ISSN 13624962
- [5] HUCKA, Michael ; BERGMANN, Frank T. ; HOOPS, Stefan ; KEATING, Sarah M. ; SAHLE, Sven ; SCHAFF, James C. ; SMITH, Lucian P. ; WILKINSON, Darren J.: The Systems Biology Markup Language (SBML): Language Specification for Level 3 Version 1 Core. In: *Journal of Integrative Bioinformatics (JIB)* 12 (2015), Nr. 2, S. 382–549. <http://dx.doi.org/10.2390/biecoll-jib-2015-266>. – DOI 10.2390/biecoll-jib-2015-266. – ISSN 1613-4516
- [6] JENSEN, Paul A. ; LUTZ, Kyla A. ; PAPIN, Jason A.: TIGER: Toolbox for integrating genome-scale metabolic models, expression data, and transcriptional regulatory networks. In: *BMC Systems Biology* 5 (2011). <http://dx.doi.org/10.1186/1752-0509-5-147>. – DOI 10.1186/1752-0509-5-147. – ISSN 17520509
- [7] KANEHISA, Minoru ; SATO, Yoko ; FURUMICHI, Miho ; MORISHIMA, Kanae ; TANABE, Mao: New approach for understanding genome variations in KEGG. In: *Nucleic Acids Research* 47 (2019), Nr. D1, S. D590–D595. <http://dx.doi.org/10.1093/nar/gky962>. – DOI 10.1093/nar/gky962. – ISSN 13624962
- [8] KARLSEN, Emil ; SCHULZ, Christian ; ALMAAS, Eivind: Automated generation of genome-scale metabolic draft reconstructions based on KEGG. In: *BMC Bioinformatics* 19 (2018), Nr. 1, S. 1–11. <http://dx.doi.org/10.1186/s12859-018-2472-z>. – DOI 10.1186/s12859-018-2472-z. – ISSN 14712105
- [9] KARP, Peter D. ; LATENDRESSE, Mario ; PALEY, Suzanne M. ; KRUMMENACKER, Markus ; ONG, Quang D. ; BILLINGTON, Richard ; KOTHARI, Anamika ; WEAVER, Daniel ; LEE, Thomas ; SUBHRAVETI, Pallavi ; SPAULDING, Aaron ; FULCHER, Carol ; KESELER, Ingrid M. ; CASPI, Ron: Pathway tools version 19.0 update: Software for pathway/genome informatics and systems biology. In: *Briefings in Bioinformatics* 17 (2016), Nr. 5, S. 877–890. <http://dx.doi.org/10.1093/bib/bbv079>. – DOI 10.1093/bib/bbv079. – ISSN 14774054
- [10] OLIVIER, Brett G. ; BERGMANN, Frank T.: *SBML Level 3 Package: Flux Balance Constraints version 2*. 2018. – 0–38 S. <http://dx.doi.org/10.1515/jib-2017-0082>. <http://dx.doi.org/10.1515/jib-2017-0082>. – ISBN 0000000252935
- [11] RODRIGUEZ, Nicolas ; THOMAS, Alex ; WATANABE, Leandro ; VAZIRABAD, Ibrahim Y. ; KOFIA, Victor ; GÓMEZ, Harold F. ; MITTAG, Florian ; MATTHES, Jakob ; RUDOLPH, Jan ; WRZODEK, Finja ; NETZ, Eugen ; DIAMANTIKOS, Alexander ; EICHNER, Johannes ; KELLER, Roland ; WRZODEK, Clemens ; FRÖHLICH, Sebastian ; LEWIS, Nathan E. ; MYERS, Chris J. ; LE NOVÈRE, Nicolas ; PALSSON, Bernhard O. ; HUCKA, Michael ; DRÄGER, Andreas: JSBML 1.0: Providing a smorgasbord of options to encode systems biology models. In: *Bioinformatics* 31 (2015), Nr. 20, S. 3383–3386. <http://dx.doi.org/10.1093/bioinformatics/btv341>. – DOI 10.1093/bioinformatics/btv341. – ISBN 13674811 (Electronic)
- [12] SWAINSTON, Neil ; SMALLBONE, Kieran ; MENDES, Pedro ; KELL, Douglas ; PATON, Norman: The SuBliMinaL Toolbox: automating steps in the reconstruction of metabolic networks. In: *Journal of integrative bioinformatics* 8 (2011), Nr. 2, S. 186. <http://dx.doi.org/10.2390/biecoll-jib-2011-186>. – DOI 10.2390/biecoll-jib-2011-186. – ISSN 16134516