

Masterarbeit in Bioinformatik

Prof. Dr. Dominik Heider
dominik.heider@uni-marburg.de

Thema: Identifikation von bakteriellen Replikationsursprüngen auf Sequenzfragmenten (Software Development)

Der Lebenszyklus bakterieller Zellen ist bestimmt von der Vervielfältigung ihrer genetischen Information, die meist in Chromosomen organisiert ist. Diese zirkulären DNA-Makromoleküle enthalten genau einen fest definierten Replikationsursprung (origin of replication), der als zentraler Kontrollpunkt der DNA-Replikation dient. An diesem DNA-Fragment interagieren mehrere Proteine miteinander, um die DNA-Doppelhelix zu öffnen, die DNA-Replikations-Maschinerie versammeln und um so mit der DNA-Synthese zu beginnen.

Auf diese Art wirkt der Replikationsursprung wie ein Signalprozessor, der aus mehreren Inputsignalen (Nährstoff- und Giftstoffsituation, Replikationszustand, Umgebungstemperatur etc.) ein einzelnes Initiationssignal berechnet. Obwohl einige der hierbei beteiligten Proteine und Signalwege bekannt und in großem Detail erforscht sind, ist es bisher unmöglich, einen funktionellen Replikationsursprung nachzubauen.

Ein zentrales Hindernis der bisherigen Forschung ist die karge Datenlage: In der größten Datenbank für Replikationsursprüngen sind etwa 2800 bakterielle Sequenzen hinterlegt. Diese Zahl steht im Kontrast zu den >40.000 assemblierten bakteriellen Genomen und unzähligen Genomfragmenten, die in öffentlich zugänglichen Datenbanken hinterlegt sind.

In diesem Projekt sollen bewährte Replikationsursprung-Vorhersagealgorithmen (wie z.B. Ref. 3) adaptiert und mit Hilfe von Machine Learning-Ansätzen weiterentwickelt werden, sodass auch Genomfragmente mit ihnen untersucht werden können. Das so entwickelte Programm soll als Online-Tool der wissenschaftlichen Community zugänglich gemacht werden.

Literatur

[1] Messer, W. (2002). The bacterial replication initiator DnaA. DnaA and oriC: The bacterial mode to initiate DNA replication. *FEMS Microbiology Reviews*, 26(4):355-374.

[2] Leonard, A. C. and Mechali, M. (2013). DNA replication origins. *Cold Spring Harbor Perspectives in Biology*, 5(10):a010116-a010116.

[3] Gao, F. and Zhang, C.-T. (2007). DoriC: a database of oriC regions in bacterial genomes. *Bioinformatics*, 23(14):1866-1867.

Bitte melden Sie sich bei Interesse per Email bei:
Prof. Dr. Dominik Heider - dominik.heider@uni-marburg.de

20.04.17