

Open Source: Bioinformatik-Tools aus Jena erleichtern Phagen-Detektion und PCR-Design

In zwei wissenschaftlichen Publikationen haben InfectoGnostics-Forschende unter Beteiligung des Leibniz-IPHT kürzlich neue bioinformatische Tools vorgestellt, deren Open-Source-Code nun kostenlos genutzt und weiterentwickelt werden kann. Die Software *What the Phage* ermöglicht die Detektion von Phagen, während *ConsensusPrime* für die Entwicklung von verschiedenen molekularen Tests genutzt werden kann. Entwickelt wurden die Tools von Forschenden des Universitätsklinikums Jena (UKJ), des Leibniz-IPHT sowie des Startups nanozoo.

Bakteriophagen, kurz Phagen, sind Viren, die Bakterien befallen und deshalb in den letzten Jahren wieder stärker in den Fokus für therapeutische Anwendungen gelangt sind. Gerade bei antibiotikaresistenten Erregern könnten Phagentherapien eine vielversprechende Option darstellen. Bis heute sind allerdings Phagen und ihr Einfluss auf Mikrobiome in der Umwelt und im Menschen noch wenig erforscht.

Obwohl mehr und mehr Daten von sequenzierten Mikroorganismen vorliegen, wird bislang nicht systematisch untersucht, welche Phagen auftreten. Zur Lösung dieses Problems haben Wissenschaftlerinnen und Wissenschaftler gemeinsam mit nanozoo die Software *What the Phage* entwickelt. Die Software bietet einen Workflow, der mittels maschinellen Lernens und anderen Algorithmen Phagen detektieren und mögliche neue Varianten aus Sequenzdaten vorhersagen kann.

Bei *What the Phage* wurden dafür mehrere Programme zu einem Workflow vereint und für eine möglichst schnelle und einfache Auswertung von Nutzerinnen und Nutzern optimiert. Dabei haben die Forschenden bewusst auf ein modulares Open-Source-Prinzip gesetzt, um den Workflow kontinuierlich erweitern und verbessern zu können, wie zum Beispiel die Vorhersage für Prophagen.

***ConsensusPrime* erleichtert Primer-Design für neue PCR-Tests**

Die PCR, die Polymerase-Kettenreaktion, gilt als der methodische Goldstandard für exakte molekulare Labortests in der Infektiologie. Erbgut-Abschnitte, die charakteristisch für den gesuchten Erreger sind, können mit dem molekularbiologischen Verfahren vervielfältigt und so nachgewiesen werden.

Doch damit das namensgebende Enzym – die DNA-Polymerase – weiß, welche DNA-Abschnitte sie vervielfältigen muss, designen Wissenschaftlerinnen und Wissenschaftler sogenannte Primer – kurze DNA-Abschnitte, an denen die Polymerase bindet. Mit *ConsensusPrime* haben Forschende der Abteilung Optisch-Molekulare Diagnostik und Systemtechnologie am Leibniz-IPHT ein Bioinformatik-Tool entwickelt, das ein schnelleres und effizienteres Design von Primern ermöglicht.

Beim Primer-Design stehen Forschende oft vor dem Problem, dass sie nicht nur einen einzigen, sondern mehrere sehr nah verwandte Erregerstämme nachweisen oder unterscheiden möchten. Zugleich muss ein Primer aber auch möglichst spezifisch an das Ziel-Erbgut angepasst sein. Forschende suchen deshalb oft nach einem sogenannten Konsensus-Primer, das heißt einer

Sequenz, die die größte Ähnlichkeit zu mehreren Stämmen hat, sich aber dennoch optimal als Primer eignet.

„Ohne unsere Software muss man dafür meist händisch die Sequenzabschnitte herausuchen, die Ähnlichkeiten aufweisen. Das ist nicht nur unpräzise, sondern bei großen Datenmengen kaum noch überschaubar und auch sehr zeitaufwendig. ConsensusPrime filtert unpassende Primer direkt aus und berechnet exakt die Erbgutabschnitte mit den größten Ähnlichkeiten. Mit unserer Pipeline erhalten wir so eine Konsensussequenz für den optimalen Primer“, erklärt Dr. Maximilian Collatz vom Leibniz-IPHT.

Über den InfectoGnostics Forschungscampus Jena

Der InfectoGnostics Forschungscampus Jena beschreitet als öffentlich-private Partnerschaft neue Wege in der Vor-Ort-Diagnostik von Infektionen und Erregern, wie zum Beispiel Viren, Bakterien und Pilzen. InfectoGnostics wird durch das BMBF im Rahmen der Förderinitiative „Forschungscampus – öffentlich-private Partnerschaft für Innovationen“ mit zusätzlicher Unterstützung durch den Freistaat Thüringen gefördert. Etwa die Hälfte des benötigten Etats finanzieren die beteiligten Partner. www.infectognostics.de

Wissenschaftliche Veröffentlichungen

What the Phage:

Mike Marquet et. al, *GigaScience*, Volume 11, giac110, 2022,
<https://doi.org/10.1093/gigascience/giac110>

Weitere Informationen auf GitHub:

https://github.com/replikation/What_the_Phage

ConsensusPrime:

Maximilian Collatz et. al, *BioMedInformatics* 2, no. 4: 637-642, 2022,
<https://doi.org/10.3390/biomedinformatics2040041>

Weitere Informationen auf GitHub:

<https://github.com/mcollatz/ConsensusPrime>

Publikationen

What the Phage: A scalable workflow for the identification and analysis of phage sequences