

PRESSEMITTEILUNG

Die Ableitung des Humangenoms zu einem Bruchteil der aktuellen Kosten wird die biomedizinische Forschung beschleunigen

Lausanne, 13. Januar 2021 – Bisher wurden bereits tausende genetische Marker robust mit komplexen menschlichen Zügen wie beispielsweise Alzheimer, Krebs, Übergewicht oder Körpergrösse assoziiert. Um diese Assoziationen zu entdecken, müssen Forscher die Genome vieler Einzelpersonen an Millionen von genetischen Abschnitten, sogenannten Markern, vergleichen, was kostengünstige Genotypisierungstechnologien erfordert. Eine neue statistische Methode, die von der Gruppe um Olivier Delaneau vom SIB Swiss Institute of Bioinformatics und der Universität Lausanne entwickelt wurde, bietet nun revolutionäre Möglichkeiten. Für Berechnungskosten von weniger als 1 \$ kann GLIMPSE ein komplettes Humangenom aus einer sehr geringen Datenmenge statistisch ableiten. Die Methode ist eine erste realistische Alternative zu aktuellen Ansätzen, die sich auf einen vordefinierten Satz genetischer Marker stützen, und dadurch eine grössere Inklusion von unterrepräsentierten Populationen ermöglicht. Die Studie, die einen Paradigmenwechsel für die Datenerzeugung in der biomedizinischen Forschung einläuten könnte, wurde in *Nature Genetics* veröffentlicht.

Ein preiswert und effizienter Ansatz für die Untersuchung von genetischen Markern

Bei der Sequenzierung des Gesamtgenoms mit niedriger Abdeckung (Low Coverage Whole Genome Sequencing, LC-WGS), handelt es sich um eine Methode, mit der ein Gesamtgenom statistisch abgeleitet werden kann. Neben einem sehr geringen Sequenzierungsaufwand stellt sie zudem eine unverfälschtere und leistungsstärkere Alternative zu aktuellen Technologien dar, die auf vordefinierten Sätzen von genetischen Markern beruhen (siehe Kasten 'Genetische Assoziationsstudien'). Die hohen Rechenkosten dieses Ansatzes haben seine breite Anwendung bislang jedoch verhindert. Jetzt stellt das [Wissenschaftlerteam um Olivier Delaneau](#), Gruppenleiter am SIB und der Universität Lausanne, eine Open-Source-Software namens GLIMPSE bereit, die dieses grosse Berechnungsproblem endlich behebt. «GLIMPSE bietet einen Rahmen, der 10-1.000 Mal schneller und somit günstiger ist als andere LC-WGS Methoden. Gleichzeitig ist die Genauigkeit für seltene genetische Marker sehr viel grösser», erklärt Olivier Delaneau. «GLIMPSE kann die Qualität eines Genoms, das durch eine Sequenzierungstechnik mit geringer Abdeckung generiert wurde, für weniger als 1 \$ Rechenkosten signifikant verbessern. Die Software ist somit die erste realistische Alternative zu aktuellen Ansätzen.»

Von unvoreingenommenen Daten zu einer inklusiveren Gesundheitsversorgung

Genomweite Assoziationsstudien konzentrierten sich bisher hauptsächlich auf Europäer: 80 % aller Studienteilnehmer waren europäischer Abstammung, obwohl sie nur 16 % der Weltbevölkerung darstellen. Dies ist ein grosses ethisches Problem im Hinblick auf die Inklusivität der Gesundheitsversorgung und den fairen Zugang zu den Vorteilen der biomedizinischen Forschung, da der Beitrag von genetischen Markern zur Krankheitsanfälligkeit in den menschlichen Populationen unterschiedlich ist. LC-WGC umgeht diese Verzerrung, die durch einen vorerstellten Satz von genetischen Markern entsteht. Daher kann die Methode erfolgreich für unterrepräsentierte Populationen angewendet werden, wie in dieser Machbarkeitsstudie für eine afrikanisch-amerikanische Population gezeigt ist. «Neben der Überwindung der wirtschaftlichen Hürde, um

Genomweite Assoziationsstudien basierend auf LC-WGS zu ermöglichen, ist dieser Ansatz besonders spannend, da er es Forschern ermöglicht, Assoziationen in unterrepräsentierten Populationen effizient zu entdecken», kommentiert Simone Rubinacci, Postdoktorand in Delaneaus Gruppe und Erstautor der Arbeit.

Bereits sequenzierte Genome nutzen

«Anfangs fragten wir uns, ob es möglich sei, die Vielfalt von bereits sequenzierten Genomen zu nutzen, um neu sequenzierte Genome zu verbessern. Anders gesagt, mehr für weniger zu bekommen: Genau das ermöglicht GLIMPSE», erklärt Diogo Ribeiro, Postdoktorand in Delaneaus Gruppe und Co-Autor der Arbeit. Wie? Indem die Idee zugrunde gelegt wird, dass wir alle relativ junge Vorfahren teilen, von denen wir kleine Teile unserer DNA geerbt haben. Kurz gesagt, GLIMPSE nutzt grosse Sammlungen von Humangenomenen, die sehr genau sequenziert wurden (High-Coverage WGS), um mit neu sequenzierten Genomen geteilte DNA-Abschnitte zu identifizieren. So kann GLIMPSE die Lücken in den Low-Coverage-Daten füllen.

Ein neues Paradigma für zukünftige Genomstudien mit weitreichenden Anwendungen

GLIMPSE wird als Teil einer Open-Source-Tool-Suite zur Verfügung gestellt und ebnet den Weg für eine breite Einführung von Low-Coverage WGS. Die Software läutet einen Paradigmenwechsel bei der Datenerzeugung für künftige Genomstudien ein. Seit der ersten Veröffentlichung als Preprint im April 2020 wurde das Tool bereits für die laufende Forschung verwendet. Zum Beispiel, um anhand von alter DNA die Genome von Menschen zu rekonstruieren, die vor tausenden von Jahren lebten, oder die von COVID-19-Patienten, anhand von Nasen-Rachen-Abstrichen für SARS-CoV-2 als Teil einer Genomweite Assoziationsstudie.

Genetische Assoziationsstudien

Um Assoziationen zwischen genetischen Markern und komplexen Merkmalen zu finden, untersuchen Forscher heute die miteinander verknüpften Daten aus den medizinischen Aufzeichnungen und genetischen Daten von Tausenden von Teilnehmern (Genomweite Assoziationsstudien). Um dies zu erreichen, stützt sich eine häufig verwendete Technologie auf vordefinierte Sets genetischer Marker, die SNPs (für Single Nucleotide Polymorphisms) genannt werden. SNPs sind einzelne Bausteine der DNA im Genom, die von Mensch zu Mensch unterschiedlich sind. Allerdings hat diese Technik, obwohl sie relativ schnell und kostengünstig ist, grosse Nachteile, da neue oder seltene Varianten, wie z. B. solche, die in wenig untersuchten Populationen vorkommen (siehe oben), unbemerkt bleiben können.

REFERENZ

Rubinacci *et al.*, Efficient phasing and imputation of low-coverage sequencing data using large reference panels, Nat. Genet. 2021. DOI : [10.1038/s41588-020-00756-0](https://doi.org/10.1038/s41588-020-00756-0)

ÜBER DAS SIB SWISS INSTITUTE OF BIOINFORMATICS

Das [SIB](#) ist eine international anerkannte gemeinnützige Organisation, die den biologischen und biomedizinischen Datenwissenschaften gewidmet ist. Die Data-Science-Experten des SIB teilen die Leidenschaft, Wissen zu schaffen und komplexe Probleme auf verschiedenen Gebieten zu lösen, von der Biodiversität über die Evolution bis zur Medizin. Sie stellen essentielle Datenbanken und Software-Plattformen sowie Expertise im Bereich der Bioinformatik und Dienstleistungen für akademische, klinische und industrielle Gruppen bereit. Das SIB vereint die schweizerische Bioinformatik-Community und besteht aus 800 Wissenschaftlern, die die Zusammenarbeit und den Wissenstransfer



fördern. Das Institut trägt dazu bei, dass die Schweiz an der Spitze der Innovation bleibt, indem es die biologische Forschung voranbringt und das Gesundheitswesen verbessert.

MEDIENKONTAKTE

Olivier Delaneau – Gruppenleiter – olivier.delaneau@unil.ch – 021 692 40 55

Maia Berman – Kommunikationsmanagerin – maia.berman@sib.swiss – 021 692 40 54

MEDIENMATERIAL

[Bilder](#) der Autoren, von links nach rechts: Robin Hofmeister, Diogo Ribeiro, Simone Rubinacci und Olivier Delaneau.