

Strukturvarianten im menschlichen Genom systematisch charakterisiert

Das menschliche Genom unterscheidet sich von Person zu Person. In der DNA-Sequenz sind zum Beispiel einzelne „Buchstaben“, sogenannte Nukleotide, verändert. Noch größere Unterschiede ergeben sich durch Strukturvarianten, die entstehen, wenn große DNA-Segmente eingefügt, gelöscht oder verschoben werden. Als Teil eines globalen Forscherteams haben Bioinformatiker der Universität des Saarlandes diese Unterschiede genauer untersucht und die Strukturvarianten in drei Familien umfassend charakterisiert. Die Wissenschaftler haben damit eine Basis geschaffen, um die Konsequenzen aus diesen genetischen Varianten systematisch zu erforschen. Die Ergebnisse wurden in Nature Communications publiziert.

Die Genomsequenzierung ist in den vergangenen zehn Jahren viel schneller, genauer und kostengünstiger geworden. Deswegen werden immer mehr menschliche Genome sequenziert. Das Wissen darüber, was diese Sequenzen tatsächlich für die Gesundheit und Krankheit von Menschen bedeuten, wächst rasant. Das menschliche Genom wird immer mehr als dynamische Einheit wahrgenommen, das erhebliche Unterschiede zwischen den Individuen aufweist. „Je besser wir genetische Varianten und ihre Folgen für jeden Einzelnen kennen, desto besser können wir dieses gesammelte Wissen in der Zukunft medizinisch nutzen“, erklärt Tobias Marschall vom Zentrum für Bioinformatik an der Universität des Saarlandes.

In der Forschung nennt man solche größeren Unterschiede „Strukturvarianten“. Sie beinhalten DNA-Segmente, die in das Genom eingefügt oder aus ihm entfernt werden, sowie Segmente, die dupliziert werden, und Segmente, deren Richtung vertauscht wird. Daher sind sie nicht nur schwieriger zu identifizieren als einzelne Nukleotidvarianten, bisher konnte die Wissenschaft auch nicht beantworten, wie viele Strukturvarianten tatsächlich in einem menschlichen Genom existieren.

Die Arbeitsgruppe des „Human Genome Structural Variation Consortium“ hat nun eine Palette von Analyse-Technologien genutzt, um die Genome von drei Familientrios (Eltern und Kind) umfassend zu erforschen. Zu den verwendeten Technologien gehören lang-, kurz- und strangspezifische Sequenzierungstechnologien, sogenanntes optisches Mapping und mehrere Computeralgorithmen zur Erkennung der Strukturvarianten.

Auf diese Weise stellten die Forscher den bisher umfassendsten Katalog von Strukturvarianten in individuellen Genomen zusammen, einschließlich der Informationen, von welchem Elternteil die einzelnen genetischen Varianten vererbt wurden. Das stellt einen großen Fortschritt dar, da dies bisher nicht in einem solchen Maßstab möglich war.

Zusammenfassend identifizierten die Forscher pro Genom durchschnittlich 818 054 kleine Einfügungen und Löschungen, also genomische Veränderungen, die jeweils weniger als 50 Basen DNA betrafen. Zudem erfassten sie 27 622 Strukturvarianten, die 50 Basen oder mehr DNA betrafen. „Bemerkenswert ist, dass wir durchschnittlich 156 Inversionen pro Genom, also um 180 Grad gedrehte Sequenzen, fanden. Von diesen lagen viele in Abschnitten des Genoms, die mit genetischen Erkrankungen in Verbindung stehen“, so Marschall.

Die Forscher fanden weiter heraus, dass mehr als 100 000 Varianten pro Individuum durch Routine-Sequenzierungstechnologien und häufig verwendete Computeralgorithmen übersehen werden. So werden mit Standardmethoden beispielsweise 83 Prozent der Einfügungen übersehen. Tatsächlich scheint die Anzahl der Strukturvarianten in menschlichen Genomen drei- bis siebenmal höher zu sein, als die meisten Studien typischerweise zeigen.

Diese Ergebnisse zeigen, dass Strukturvarianten einen großen Teil der genetischen Diversität ausmachen, die üblicherweise gerade nicht durch aktuelle Genomsequenzierungstechnologien und Analysemethoden erfasst wird. Dies, so die Forscher, bedeute, dass der Beitrag von Strukturvarianten zu menschlichen Krankheiten noch nicht gut quantifiziert sei. Dies sei deshalb wichtig, weil für viele Erkrankungen die genetischen Varianten, die zu einem erblichen Krankheitsrisiko beitragen, noch nicht identifiziert werden konnten. Die Wissenschaftler bezeichnen diese Tatsache als „missing heritability“. Strukturvarianten könnten daher eine wichtige Rolle beim Schließen dieser Lücke spielen. Während der flächendeckende Einsatz der von dem Forscherkonsortium verwendeten Methoden derzeit noch nicht praktikabel sei, könnten jedoch mehrere Technologien zur Erkennung der Strukturvarianten schon heute genetische Einflüsse auf Krankheiten offenlegen und bei der Aufklärung seltener genetischer Syndrome helfen.

Zu der aktuellen Publikation haben 97 Autorinnen und Autoren des Human Genome Structural Variation Consortium (HGSVC) beigetragen. Geleitet wurde die Studie von den Erstautoren Mark Chaisson, Ashley Sanders und Xuefang Zhao, den Senior-Autoren Paul Flicek, Ken Chen, Mark Gerstein, Pui-Yan Kwok, Peter Landsdorp, Gabor Marth, Jonathan Sebat, Xinghua Shi, Ali Bashir, Kai Ye, Scott Devine, Michael Talkowski, Ryan Mills, Tobias Marschall sowie den Leitern des HGSVC, Jan Korb, Evan Eichler und Charles Lee. Ihre Ergebnisse haben die Forscherinnen und Forscher in Nature Communications unter dem Titel „Multi-platform discovery of haplotype-resolved structural variation in human genomes“ veröffentlicht.

Über das Zentrum für Bioinformatik an der Universität des Saarlandes:

Das Zentrum für Bioinformatik ist das interdisziplinäre Gemeinschaftsprojekt der Fachrichtungen Biologie, Medizin und Mathematik/Informatik. Die Aufgabe des Zentrums ist es, Berechnungsmethoden zur Analyse, Modellierung und Simulation biologischer Prozesse auf molekularer Ebene zu entwickeln. Ziel ist es, Softwarewerkzeuge bereitzustellen, die die Entwicklung neuer Medikamente und neuer Diagnoseverfahren unterstützen und bei der Optimierung therapeutischer Strategien helfen.

Originalpublikation: <https://www.nature.com/articles/s41467-018-08148-z>

Weitere Informationen:

<https://zbi-www.bioinf.uni-sb.de>