

# Sequenzieren, bis der Rechner schlappmacht

Angesichts der Milliarden Informationen, die das Erbgut bereithält, benötigen Forschung und Industrie schnelle Computer und riesige Server. Über den Zusammenhang von Genen und Daten

Von Sabrina Gaisbauer

→ „Big Data“ in der Gentechnik bedeutet, enorme Gen-Datenmengen erfassen, vergleichen, analysieren und so die genetischen Ursachen für bestimmte Eigenschaften erkennen zu können. Da allein das menschliche Genom 3,27 Milliarden Basenpaare umfasst, braucht es leistungsfähige Rechner und treffsichere Algorithmen, um Zusammenhänge zwischen dem Erbgut verschiedener Menschen oder auch Tiere und Pflanzen verstehen und damit nutzbar machen zu können. Im genetischen Baukasten macht Big Data es möglich, Bausteine voneinander zu unterscheiden und zu bestimmen, wie sie zusammenpassen. Massenhafte Datensammlung und ihre Auswertung sind eine Voraussetzung für gentechnische Entwicklungen.

Im „1.000-Genome-Projekt“ etwa, das von 2008 bis 2015 lief, wurde bei 2.504 Menschen die Reihenfolge von Genbausteinen untersucht – dabei entdeckte man 84 Millionen genetische Variationen. Und das ist nur eines von vielen Genom-Projekten: Die USA haben letztes Jahr ein Projekt gestartet, bei dem das Erbgut von einer Million US-Amerikaner analysiert werden

soll, um individuelle Medikamente entwickeln zu können. Die Kosten für diese Genomuntersuchungen sind in den letzten Jahren dramatisch gesunken: Laut dem National Human Genome Research Institute hat die Sequenzierung eines einzigen Genoms vor rund zehn Jahren noch 14 Millionen Dollar gekostet, Ende letzten Jahres waren es nur noch 1.500 Dollar. 2015 gaben das Supercomputerunternehmen Cray und der Pharma-Riese GlaxoSmithKline an, für eine komplexe Wechselwirkungsanalyse von 36.000 Genen nur noch 20 Minuten zu benötigen.

Auch für die freie Forschung spielt Big Data eine große Rolle. Letztes Jahr hat ein Forscherteam von der University of Pittsburgh über zwei öffentlich zugängliche Datenbanken mit Gensätzen und Proteinen von Tausenden Brustkrebstumoren Hinweise dafür gefunden, wie Brustkrebs bei Frauen in den Wechseljahren verläuft.

Aber nicht nur für Menschen und Gentechnik ist die massenhafte Datenauswertung relevant, sondern auch für Züchtungen: Wo etwa früher jahrelange Versuche nötig waren, markieren heute Maschinen interessante Eigenschaften auf DNA-Strängen, und Computer rechnen für Millionen von Datensätzen aus, was bei Kreuzungen herauskommen würde. Was besonders vielversprechend aussieht, wird dann umgesetzt. Das Unternehmen Monsanto nutzt dieses „Smart Breeding“, auch „Präzisionszucht“ genannt, seit über zehn Jahren. Zu den ersten angekündigten Sorten gehörte eine Sojabohne, deren Öl u.a. nicht so schnell ranzig werden soll. Doch beim Züchten blieb es nicht: Für den US-amerikanischen Markt wurde die Sojabohne 2015 in einer gentechnisch veränderten Version zugelassen – und für die EU vor wenigen Wochen. ←



Reparatur an einem der größten Rechner der Welt. Noch vor zehn Jahren kostete das Auslesen eines Genoms Millionen Euro, mittlerweile nur noch ein paar Tausend