

Das Wissenschaftsmagazin im Falter

Mathematik (2/09)

Lotsen in der Informationsflut

Um die Unmengen neuer genetischer Daten zu interpretieren, braucht es Bioinformatiker und Biomathematiker. Zehn Fragen und zehn Antworten zu einem boomenden Forschungsfeld.
Von Oliver Hochadel

Wieso wird jetzt in der Biologie gerechnet?

Weiße Mäuse und Fruchtfliegen, Mikroskop und Petrischale – das ist doch die Grundausstattung eines modernen Biologielabors. Oder? Ja, aber nicht mehr nur das. In manchen Abteilungen stehen nämlich nur mehr Rechner und Server.

Das Bild von der Biologie als „weicher, organischer“ Naturwissenschaft im Gegensatz zur „harten“, mathematischen Physik stimmt schon lange nicht mehr. In der Ökologie oder der Populationsgenetik nutzen die Biologen schon seit vielen Jahrzehnten mathematische Methoden.

Längst wird auch in der Molekularbiologie modelliert und statistisch ausgewertet. Zu den traditionellen Bereichen *in vivo* (Organismus) und *in vitro* (Reagenzglas) ist in den letzten 25 Jahren ein drittes Standbein hinzugekommen: *in silico* (Silizium, also der Computerchip), kurz: die Bioinformatik. Der Computer, die Datenbank und das Netzwerk sind zu mächtigen Werkzeugen der biologischen Forschung geworden.

Lässt sich die Information so einfach digitalisieren?

In den 90er-Jahren sorgte das Human Genome Project für weltweite Aufmerksamkeit. Elektronische Datenverarbeitung spielte bei der Erfassung der mehr als drei Milliarden Basenpaare des menschlichen Erbguts eine zentrale Rolle. Das Buch des Lebens, wie es damals vollmundig hieß, ist damit freilich noch lange nicht entschlüsselt, bestenfalls aufgeschlagen. Die Sequenzierung der DNA mit den immer wiederkehrenden Basenpaaren AT und CG ist ihrer Natur nach digital. Wie aber etwa Gene miteinander interagieren oder wie Proteine gefaltet sind, diese Fragen stellen Biologen und Informatiker vor ungleich schwierigere Probleme.

Wo lagern die Daten?

Das erfasste biologische Material (DNA, Gene, Proteine etc.) des Menschen und vieler anderer Organismen ist auf öffentlichen Datenbanken allen zugänglich. „Wir haben mehr Nukleotide in unserer Datenbank, als es Sterne in der Milchstraße gibt“, rühmte sich die US-Datenbank GenBank schon 2005. Es gibt zwei weitere Datenbanken in Europa und Japan; die drei gleichen ihre Bestände täglich untereinander ab. Nach eigenen Angaben verdoppelt sich ihr Datenbestand alle 18 Monate, nicht zuletzt deshalb, weil viele wissenschaftliche Zeitschriften von ihren Autoren verlangen, ihre Daten dort einzustellen.

„Man muss natürlich genau wissen, wonach man sucht“, sagt [Jacques Colinge](#), Bioinformatiker am CeMM, dem Zentrum für Molekulare Medizin der ÖAW in Wien. „Aber allein durch die Nutzung öffentlich zugänglicher Daten kann ein cleverer Biologe wahre Schätze heben.“

Was tun Bioinformatiker?

Die digitalisierten Datenmengen sind häufig nicht nur enorm groß, sondern auch sehr heterogen. „Wir entwickeln Algorithmen und statistische Modelle, um Muster aufzuspüren“, beschreibt Colinge seine Arbeit. Auch [Joachim Hermisson](#), Stiftungsprofessor für Mathematik und Biowissenschaften an der Universität Wien, betont: „Die jeweiligen mathematischen Werkzeuge sind nicht einfach gegeben, sondern müssen für die gegebene Fragestellung maßgeschneidert werden.“

Welche Methoden kommen zur Anwendung?

Immer schneller, immer präziser: Die Entwicklung neuer Methoden hält die Biologen schwer auf Trab. Was noch vor wenigen Jahren in war, sogenannte Genchips oder Microarrays, gelten heute schon wieder als veraltet, nicht zuletzt weil sie schwer zu interpretieren sind. Der letzte Schrei ist die neue Generation an Sequenzierern bzw. das sogenannte „High-throughput Sequencing“. Durch die Mehrfachsequenzierung des genetischen Materials wird das Ergebnis nicht nur genauer, auch die molekulare Variation zeigt sich. „Die Daten sind von einer nie dagewesenen Menge und einer Genauigkeit“, schwärmt der österreichische Biomathematiker [Günter Wagner](#) von der Yale University in den USA. Man erhalte ein völlig neues Bild von den Vorgängen auf molekularer Ebene und könne etwa die Interaktion zwischen Zelle und Antikörper quasi „live“ mitverfolgen. Dank der schnellen Lesetechniken gehen Datenerwerb und Analyse Hand in Hand, und man muss nicht mehr Monate oder Jahre warten, bis man Rückmeldungen erhält, so Wagner.

Wo liegen die Probleme?

Wie immer bei großen Datenmengen müssen Standards für die Aufbereitung und Interfaces für den Austausch der Daten geschaffen werden. Die Schwierigkeiten inhaltlicher Art, also der Datenanalyse, klingen ein klein wenig nach Luxusproblem. „Wir laufen immer hinterher“, berichtet [Zlatko Trajanoski](#) vom Institut für Genomik und Bioinformatik der TU Graz. „Haben wir eine Methode für ein bestimmtes Sequenzierungsverfahren ausgearbeitet, gibt es schon wieder ein schnelleres. Und gerade bei den neuesten Verfahren werden die Daten aufgrund der gigantischen Menge gar nicht mehr gespeichert. Die bleiben auf einer Festplatte, die im Keller gelagert wird, und sind kaum mehr transferierbar.“ „Wir werden bald in Daten ertrinken“, stimmt Joachim Hermisson zu, was ihn als Theoretiker freilich freut. Die Zeit des Flaschenhalses, der knappen Daten, ist jedenfalls vorbei.

Ist Bioinformatik ein eigenes Fach?

Ursprünglich war die Bioinformatik eine Art Hilfswissenschaft. Mittlerweile ist es aber längst ein eigenes Fach geworden, das man studieren kann und das über eigene Zeitschriften und Fachgesellschaften verfügt. Nach wie vor besteht ein großer Teil der Arbeit im „Service“. Universitäten wie auch außeruniversitäre Forschungseinrichtungen haben eigene Abteilungen für Bioinformatik, um die experimentell gewonnenen Datenmengen aufbereiten und auswerten zu können. Aber selbst im Supportbereich brauche es eigene Forschung, sonst bleibe die Disziplin steril, so Zlatko Trajanoski. Mittlerweile geht der Austausch aber in beide Richtungen, d.h. die Bioinformatiker generieren selbst Hypothesen, die die experimentellen Biologen überprüfen.

An was wird konkret geforscht?

Trajanoski nennt als Beispiel das GOLD-Projekt des österreichischen Genomforschungsprogramms GEN-AU, in dem es um die Ursachen von Fettleibigkeit geht. Mittlerweile schlagen die Bioinformatiker den Laborbiologen „Kandidaten“ für Gene vor, die mit Adipositas korrelieren könnten.

Biomathematiker Joachim Hermisson modelliert evolutionäre Prozesse, wobei neben den „klassischen“ Größen wie Mutationsrate und Selektionsstärke auch die Geschwindigkeit der Umweltveränderung „mitberechnet“ wird. Dadurch soll simuliert werden, wann sich Spezies schnell und wann langsam anpassen.

[Christian Schmeiser](#), Mathematiker an der Universität Wien, möchte mithilfe von Modellierungen und Simulationen die Mechanik in Zellen beschreiben, also wie etwa weiße Blutkörperchen durch die Zellen „kriechen“. Ziel ist ein besseres Verständnis der Immunabwehr.

Wie funktioniert die Zusammenarbeit?

„Die Mathematiker müssen auf die Biologen zukommen, wir müssen die Distanz überbrücken“, sagt Christian Schmeiser, der mit dem Institut für Molekulare Biotechnologie (IMBA) der ÖAW kooperiert. „Wir zeigen den Biologen keine Formeln, sondern kommunizieren über Bilder. Denn die sind grafische Darstellungen gewohnt.“ Ja, der Mathematiker habe die schwerere zu verstehende Information, stimmt Günter

Wagner zu. Er weist aber auch darauf hin, dass sich die disziplinären Hintergründe der Forscher immer stärker vermischen. Sein Team in Yale erledigt sowohl die experimentellen als auch die mathematisch-modellierenden Aufgaben selbst.

Wie gut ist die Bioinformatik in Österreich?

Was die Biomathematik angeht, also die Anwendung mathematischer Methoden auf biologische Fragen, gab (und gibt) es in Österreich schon seit Jahrzehnten auch international beachtete Forschung. Zu den Vorreitern gehören der Wiener Chemiker und frühere ÖAW-Präsident Peter Schuster und der Wiener Mathematiker Karl Sigmund.

In der Bioinformatik hingegen war Österreich Nachzügler. In den letzten fünf Jahren hat sich die Situation allerdings wesentlich verbessert, etwa durch die Einrichtung von Stiftungsprofessuren durch den WWTF (Wiener Wissenschafts- und Technologiefonds). Mittlerweile gibt es Bioinformatik-Professuren in Wien (2), Graz, Linz, Salzburg und Innsbruck und etwa zehn weitere Arbeitsgruppen. „Gute Forscher wollen vor allem ein gutes Netzwerk haben, sagt Jacques Colinge, und das entsteht hier mehr und mehr.“ Zur Weltspitze der Bioinformatik reiche es freilich noch nicht.