

# Die Decodierung des Lebens

Biologie und Informationstechnologie besaßen vor einigen Jahren eine geringe Schnittmenge. Doch die Bioinformatik führt beide zusammen. Spätestens seit dem erfolgreichen Abschluss des Human Genome Project ist der Computer nicht mehr aus der Wissenschaft des Lebens wegzudenken. Und neue Aufgaben fordern mehr denn je seinen Einsatz.



► Selten positionierten sich die politischen Führer der USA und Großbritanniens so prominent und einmütig zu einer naturwissenschaftlichen Fragestellung. Als Bill Clinton und Tony Blair Anfang des Jahres 2000 vor die Presse traten, ging es um nichts Geringeres als das Human Genome Project (HGP). Die Forderung nach freiem Zugang zu den Daten des Projekts, das sich die Entschlüsselung der kompletten DNA-Sequenz des menschlichen Genoms zur Aufgabe gesetzt hatte, traf den Kern einer jungen interdisziplinären Wissenschaft, der Bioinformatik. Denn ähnlich wie im kommerziell genutzten Internet bestimmt seit einigen Jahren auch in den Biowissenschaften die Datenhoheit über die Möglichkeiten von Forschungslaboren und Unternehmern.

In den 80er Jahren des vergangenen Jahrhunderts war der Computer noch kein zentrales Arbeitsmittel für die Kernfragen der Biologie. Erste Ansätze in den USA stellten die Ausnahme dar. Doch eine Vision brachte, ähnlich wie für die Raumfahrt der Flug zum

Mond zwanzig Jahre zuvor, gewaltigen Schub in die technologische Entwicklung. Das HGP wurde 1989 gegründet. Unter der Führung des prominenten Fürsprechers James Watson, einem der beiden Entdecker der Doppelhelixstruktur des Erbguts und Nobelpreisträger für Medizin, startete das Projekt unter der Federführung des Human Genome Research Office. Das Ziel: Die komplette Entschlüsselung der menschlichen DNA – immerhin eine Sequenz von drei Milliarden Basenpaaren (s. Kasten 1) – innerhalb weniger Jahre. „Das war damals ein enormes, visionäres Ziel, das nur durch die Entwicklung neuer Techniken auch im Bereich der Informatik gelingen konnte. Die Hard- und Software-Möglichkeiten waren ja lange nicht so groß wie heute. Das HGP holte die Informatik zur Bewältigung und Auswertung der enormen Datenmengen mit ins Boot“, sagt Dr. Ute Willhöft vom Zentrum für Bioinformatik der Universität Hamburg in der Rückschau.

## Ist die Karte des Lebens nur ein Telefonbuch?

Wenige Jahre später, Anfang des 21. Jahrhunderts, war das zuvor Unvorstellbare erreicht: Das menschliche Genom lag – publiziert in den wissenschaftlichen Weltblättern „Nature“ und „Science“ – vollständig kartographiert vor der wissenschaftlichen Gemeinde. Die Euphorie war groß, die Ernüchterung folgte wie so oft auf dem

► Proteine bestehen aus Aminosäuren. Ihre Sequenz wird durch die DNA kodiert. Innerhalb der Ribosomen – den Proteinfabriken der Zelle – werden in der Zelle die Informationen aus der DNA verwendet, um aus einzelnen Aminosäurebausteinen ein Protein in einer durch die DNA vorgegebenen Reihenfolge zusammenzufügen.

Fuße. Denn die neuen Fragen überstiegen die gerade gegebenen Antworten. „Man muss sich das Ergebnis vorstellen wie ein Telefonbuch von München“, erklärt Dr. Ralf Blossey, Direktor am Interdisziplinären Forschungsinstitut des Centre National de la Recherche Scientifique, der in Lille eine Arbeitsgruppe über biologische Nanosysteme leitet. Die bloße Aneinanderreihung der Zahlen ist nutzlos, es geht vielmehr um die damit verbundenen Informationen und Beziehungen. Eine der neuen Fragestellungen an die Informatik war und ist das Auffinden der ca. 25.000 menschlichen Gene innerhalb der gesamten Sequenz von Basenpaaren nach ausgefeilten Suchalgorithmen. „Nur zwei bis drei Prozent der DNA bestehen überhaupt aus Genen, die an einer charakteristischen Sequenz von Basenpaaren erkennbar sind“, erklärt Ute Willhöft. Und Ralf Blossey fügt hinzu: „Deshalb hat die Bioinformatik sich in den letzten Jahren vor allem mit Mustererkennung beschäftigt. Welche bereits bekannten Gene kann man innerhalb der riesigen Datenmengen entdecken, wie kann man sie in die Gesamtheit der langen DNA-Stränge einordnen?“ Auch bei dieser Frage ist die Biologie heute mit Hilfe der Bioinformatik sehr weit vorgedrungen.

## Informatik als Hirnerweiterung für Biologen

Doch die komplexen biologischen Fragestellungen stellen weiterhin hohe Anforderungen an die interdisziplinären Kompetenzen. So sieht Professor Dr. Michael Berthold, Inhaber des Lehrstuhls für Bioinformatik und Information Mining an der Universität Konstanz, den momentanen Stand der Forschung eher nüchtern: „Das Lebewesen an sich ist auch heute noch weitgehend unverstanden. Vieles läuft in der Bioinformatik noch immer über Versuch und Irrtum.“ Vor allem bei der Analyse von Daten, so Berthold, sei der informatische Teil der Bioinformatik noch weitgehend auf die Intuition und Erfahrung der Naturwissenschaftler angewiesen: „Wir sehen uns im Moment noch sehr stark als eine Art Hirnerweiterung für Biologen.“

### *„Vieles läuft in der Bioinformatik noch immer über Versuch und Irrtum.“*

Ralf Blossey, von Hause aus Physiker, sieht allerdings neue Herausforderungen für die Bioinformatik am Horizont: „Die Mustererkennung ist mit Sicherheit noch immer ein heißes Thema für uns. Aber in Zukunft wird es eher um das Verständnis der Dynamik innerhalb von Systemen gehen.“ Denn die eigentlich interessanten Fragen, etwa für die Medizin, stellen die Wissenschaftler heute nicht an die DNA oder die in ihr liegenden Gene, sondern vor allem an die Proteine im Zellinneren, die durch die Gene kodiert werden (s. Kasten 2). „Um die Interaktionen innerhalb einer Zelle zu verstehen, setzen wir seit einiger Zeit Modelle ein, die aus dem Mobile Computing kommen.“ Auch weitere Strukturen und deren Funktionen sowie die Betrachtung des individuellen Genoms rücken ins Zentrum des bioinformatischen Interesses.

## Genvergleich eröffnet neue Möglichkeiten

Beim Verständnis der komplexen Wechselwirkungen zwischen Genen und der dadurch kodierten Proteine kann die Informatik bisher nur erste Hilfestellung leisten. Dafür macht sie in einem anderen Bereich rasante Fortschritte: Die Geschwindigkeit, mit der Genome entschlüsselt und verglichen werden können, wächst derzeit schneller als es bei der Entwicklung der Mikroprozessoren der Fall war, bei gleichzeitig rapide fallenden Kosten (derzeit Faktor 3000 in 6 Jahren).

Der Vergleich verschiedener Genome wird eine der Aufgaben von morgen sein. Denn bislang ist das menschliche Genom nur allgemein erfasst, individuelle Unterschiede, die dafür sorgen, dass sich die sechs Milliarden Menschen auf der Erde nicht wie ein Ei dem anderen gleichen, werden zunehmend Bedeutung erlangen. Ute Willhöft: „Bestimmte Medikamente schlagen bei manchen Menschen an, andere müssen vor allem die Nebenwirkungen erdulden und profitieren nicht vom positiven Effekt.“ Grund ist unter anderem die je nach Individuum unterschiedliche genetische Information. „Wenn diese individuellen Unterschiede und deren Umsetzung innerhalb der Zellen und Organe besser verstanden sind, dann können wir künftig eine individuelle Medizin entwickeln.“

## Schutz der eigenen Erbinformation als zentrale Frage der Zukunft

Angesichts der Entwicklung der Bioinformatik ist es technisch durchaus vorstellbar, dass in absehbarer Zukunft genetische Analysen in größerem Umfang als Dienstleistung für Privatpersonen angeboten werden können. Die Diskussion um Vaterschaftstests anhand von DNS, die ohne Wissen der betreffenden Person im Alltag relativ leicht beschafft werden kann, ist vermutlich nur die Spitze des Eisbergs, der hier auf uns zukommt. Wie lässt sich dann das eigene Erbgut und das Recht, über die damit verbundenen Informationen selbst bestimmen zu können, noch effektiv schützen? Noch haben wir Zeit, aber um Antworten auf diese Fragen werden wir wohl nicht herumkommen. ◀

► Die Entdecker der chemischen Struktur der DNA, Watson und Crick, stellten fest, dass die so genannten Basen Adenin (A) und Thymin (T) immer im Verhältnis eins zu eins innerhalb der doppelsträngigen Helixstruktur der DNA vorkommen. Auch die Basen Guanin (G) und Cytosin (C) liegen in dieser Form vor. Daraus schlossen die beiden Wissenschaftler, dass A-T und G-C jeweils komplementäre Basenpaare bilden, die über so genannte Wasserstoffbrückenbindungen miteinander verbunden sind.