



## BIOINFORMATIK

# Ordnung in die Datenflut

Es war ein Meilenstein in der Geschichte der Wissenschaften: die Entzifferung des Humangenoms. Doch die erfolgreiche Sequenzierung dieses Genoms hatte eine solche Datenflut zur Folge, dass die riesigen Mengen neuer Informationen nur computergestützt zu bewältigen waren. Nun ist es erforderlich, alle Daten, die sich zum Verständnis der Funktion einzelner Gene als notwendig erwiesen haben, für möglichst alle Gene eines Organismus zu generieren, zu sammeln und zu analysieren. Ohne die Bioinformatik wäre das unmöglich ...

**Hamburg/Lübeck** > Die Bioinformatik, also die Anwendung von Methoden aus der Informatik auf wissenschaftliche Probleme aus der Biotechnologie, entwickelt sich zu einer Schlüsseldisziplin für die Life Sciences. In der modernen molekularbiologischen und pharmazeutischen Forschung ist der Computer zu einem integralen Bestandteil geworden. Die moderne Forschung in den Lebenswissenschaften ist gekennzeichnet durch eine Vielzahl von Hochdurchsatzexperimenten, zum Beispiel die Sequenzierung kompletter Genome, die Strukturaufklärung von Proteinen, die Bestimmung von Expressionsprofilen oder die Messung von intermolekularen Wechselwirkungen zwischen Proteinen oder Proteinen und niedermolekularen Verbindungen. Die Aufgabe der Bioinformatik hierbei ist es, die entsprechenden Software-Werkzeuge zu liefern, die zur Vorbereitung, Auswertung und Analyse dieser Daten notwendig sind. Nachdem die Bioinformatik bisher vor allem für die Interpretation und Verarbeitung der komplexen biologischen Genom-, Proteom- und Strukturdaten wichtige Unterstützung leisten konnte, muss ihre künftige Rolle im Hinblick auf die mittelfristige Entwicklung der Biowissenschaften weit darüber hinausgehen.

Nach der Sequenzierung des Humangenoms und der Genome der wichtigsten Modellorganismen wie Reis, Ratte, Maus, Zebra- und Kugelfisch richtet sich der Forschungsfokus zunehmend auf die Funktions-

&gt;

> analyse, zum Beispiel die Struktur, die Funktion und das Zusammenspiel von Proteinen. Aufgrund der Komplexität von Regelsystemen in Zellen und Organismen fällt hier immer wieder eine immense Flut von Daten an, die es zu analysieren und zu nutzen gilt. Daher wird die bereits gegenwärtig bestehende erhebliche Bedeutung der Bioinformatik weiter wachsen. Mittlerweile ist es gerechtfertigt, bei der Bioinformatik von einem eigenständigen Wissenschaftszweig zu sprechen. Synergien bestehen nicht mehr nur darin, dass der Biologe moderne Methoden der Datenverarbeitung einsetzt und der Informatiker in der Biologie interessante Anwendungsgebiete für seine Arbeit findet. Vielmehr sind zwischen beiden Disziplinen Grenzgebiete entstanden, bei denen es wenig sinnvoll wäre, die Inhalte und Ziele der wissenschaftlichen Forschung alleine der einen oder anderen Disziplin zuzurechnen.

Mit der Aufklärung und weit reichenden Funktionsanalyse verschiedener vollständiger Genome (zum Beispiel des Fadenwurms *Caenorhabditis elegans*) verlagert sich der Schwerpunkt bioinformatischer Arbeit auf Fragestellungen der Proteomik, wie zum Beispiel das Problem der Proteinfaltung und Strukturvorhersage, also die Frage nach der Sekundär- oder Tertiärstruktur bei gegebener Aminosäuresequenz. Auch die Frage nach der Interaktion von Proteinen mit verschiedenen Nukleinsäuren, anderen Proteinen oder auch kleineren Molekülen wird untersucht, da sich daraus neben Erkenntnissen für die Grundlagenforschung auch wichtige Informationen für Medizin und Pharmazie ableiten lassen, beispielsweise darüber, wie ein durch eine Mutation verändertes Protein Körperfunktionen beeinflusst oder welche Medikamente wie an verschiedenen Proteinen wirken.

### Schnelles Auffinden von Mustern

Die ersten reinen Bioinformatikanwendungen wurden für die DNA-Sequenzanalyse entwickelt. Dabei ging es primär um das schnelle Auffinden von Mustern in langen DNA-Sequenzen und die Lösung des Problems, wie man zwei oder mehr ähnliche Sequenzen so übereinander legt und gegeneinander ausrichtet, dass man eine möglichst optimale Übereinstimmung erzielt. Zur Anwendung kommen dabei Algorithmen der dynamischen und heuristischen Programmierung. Seltener benötigt man bei biologischen Fragestellungen die Suche nach exakten Übereinstimmungen von kurzen Sequenzabschnitten, typischerweise für hochkonservierte Signale wie Start- oder Stoppcodons.

Des Weiteren wurden Methoden zum Auffinden von Genen in unbekanntem DNA-Sequenzen entwickelt, die so genannten Genvorhersagen. Dieses Problem wird mit verschiedenen Rechenmethoden und Algorithmen angegangen, darunter statistische Sequenzanalyse, Markow-Ketten oder künstliche neuronale Netze zur Mustererkennung.

Die Weiterentwicklung der Biotechnologie mit ihren Auswirkungen auf Medizin, Pharmazie, Landwirtschaft und Umwelt, aber auch auf die Nanotechnologie, die Informationstechnologie sowie die Materialforschung hängt damit unmittelbar von der Weiterentwicklung der Bioinformatik ab. Dabei herrscht – wie in den übrigen Industrienationen – auch in Deutschland ein deutlicher Mangel an ausgebildeten Bioinformatikern.

Doch Abhilfe ist in Sicht. Vor allem in Norddeutschland, und hier speziell in Hamburg, hat die Bioinformatik nicht zuletzt durch das Zentrum für Bioinformatik (ZBH) der Universität Hamburg ein tragfähiges Fundament für die Weiterentwicklung von Forschung und Ausbildung. Im Bereich der Forschung deckt das ZBH die Hauptfelder der methodischen Bioinformatik ab: die Speicherung und Analyse genomischer Sequenzen, die Vorhersage von Proteinstrukturen und Protein-Design und das computergestützte molekulare Design.



Fotos: Sher

Die ersten Bioinformatikanwendungen wurden für die DNA-Sequenzanalyse entwickelt. Dabei ging es primär um das Auffinden von Mustern in langen DNA-Sequenzen

Im Bereich der Ausbildung bietet das ZBH einen Diplom-Hauptstudiengang Bioinformatik an. Aufbauend auf einem Grundstudium in Informatik oder einem naturwissenschaftlichen Fach werden in einem fünfsemestrigen modularisierten Studium neben einer fundierten Bioinformatik-Ausbildung notwendige Kenntnisse aus der Informatik und den Life Sciences vermittelt. Das ZBH-Modell erlaubt eine individuelle Schwerpunktbildung hin zu einer methodisch- oder anwendungsorientierten Bioinformatik und spiegelt damit die Bedürfnisse des Arbeitsmarktes wieder. »Bioinformatiker können die Daten am besten auswerten, weil sie einerseits die Informatik beherrschen und andererseits die Sprache der Biologen und Chemiker verstehen«, so Prof. Matthias Rarey vom ZBH.

Welchen Stellenwert das ZBH auch im überregionalen Umfeld hat, macht die Tatsache deutlich, dass die Bioinformatik-Jahrestagung »German Conference on Bioinformatics« im Oktober 2005 erstmals in Hamburg stattfand. Führende Wissenschaftler aus Deutschland und

dem Ausland präsentierten innovative Forschungsansätze aus dem Gebiet der Bioinformatik. Unter anderem wurde die strukturelle Bioinformatik aus Sicht der Anwendung und der Methodenentwicklung beleuchtet, und neue computerbasierte Methoden zur Entwicklung von Diagnostika für pathogene Viren und Bakterien wurden vorgestellt. Aufgrund der rasanten Weiterentwicklung molekularbiologischer Methoden ergab sich eine Vielzahl neuer Fragestellungen. Als Beispiel sei das Themenfeld »Metabolomics« genannt, die Analyse aller an Zellabläufen beteiligten niedermolekularen Verbindungen.

## Hoher Bedarf an ausgebildeten Bioinformatikern

Auch die TU Harburg hat erkannt, dass eine schnelle und sichere Informationsverarbeitung für die Medizintechnik ebenso unerlässlich ist wie für die Biochemie. In diesem besonders zukunftsorientierten Schwerpunkt sind in Harburg Gebiete zusammengefasst, in denen Techniken aus anderen Disziplinen als Alternativen zur oder in Kombination mit der elektronischen Informationsverarbeitung erforderlich sind. Aus der Biochemie stammen die Techniken des DNA-Computing, die erhebliche Grundlagenfragen aufwerfen. Beispiele sind neuere bioinformatische Ansätze zur molekularen Modellierung und Biokatalyse sowie Biomathematik und systemtheoretische Arbeiten zur Systembiologie.

Nicht nur wegen des Studienfachs »Computational Life Sciences/ Neuroinformatik« ist auch die Universität Lübeck zu einem wichtigen Player im Bereich der Bioinformatik geworden: Die Bündelung von Medizintechnik und Informatik am Standort Lübeck steht im Mittelpunkt einer neuen Fraunhofer-Arbeitsgruppe, des FIT Fraunhofer Institutes für Angewandte Informationstechnik.

Die FIT-Arbeitsgruppe soll unter der Bezeichnung VIBE für »Virtuelle Biomedizintechnik Entwicklung« am Multifunktionscenter im neuen Hochschulstadteil etabliert werden. Gegenstand der Forschungs- und Entwicklungstätigkeit ist die dynamische Simulation und 3-D-Animation von biomedizinischen und klinischen Abläufen. Damit soll vor allem die räumliche und bewegte Darstellung komplizierter medizinischer, biologischer und technischer Wirkprozesse ermöglicht werden.

Ziel sind neuartige intelligente Medizintechnikgeräte und chirurgische Assistenzsysteme, aber auch virtuell animierte Trainings- und Ausbildungszentren in der Biomedizin- und Sicherheitstechnik. Derzeit werden Projektplanungen für einen Anästhesie-Navigator und einen Chirurgie-Eingriffssimulator erarbeitet sowie ein Wissensportal für Entzündungsprozesse geprüft.

## Informationsprovider für molekulare Informationen

Nicht nur im akademischen, sondern auch im industriellen Umfeld ist die Bioinformatik auf dem Vormarsch. So unterstützen die Werkzeuge der Bioinformatik (vor allem Softwareprogramme) mittlerweile nicht nur die Forschung, sondern finden auch Anwendung in Unternehmen wie Evotec AG oder Lilly Research Laboratories. Und hier gibt es gerade einen enormen Fortschritt zu vermelden: den ersten vorinstallierten, sofort einsatzfähigen Bioinformatik-Supercomputer. Das neue Superhirn für die Biotechnologie, das von der Hamburger emplies AG in Kooperation mit der Norderstedter c.a.r.u.s. Information Technology AG entwickelt wurde, soll dank der bereits eingerichteten Bioinformatik-Basisanwendungen laut Unternehmensangaben sofort in Forschungsabteilungen einsetzbar sein.

Der Spezialist für Linux-Clustereysteme und der Anbieter für High Performance Hard- und Software bündelten hierbei ihr Know-how, um den weltweit steigenden Bedarf an Hochleistungsrechnern in der Bioinformatik zu bedienen. Diese Systeme werden bei der Entwicklung von Medikamenten eingesetzt, um die Forschungszeiten deutlich zu verkürzen – etwa bei der Simulation von Wirkstoffen gegen Krankheiten wie Alzheimer oder Aids. Das Cluster-Prinzip, aus herkömmlichen Prozessoren zusammengestellte Rechnersysteme, ermöglicht extrem hohe Leistung bei deutlich geringeren Investitionskosten als mit klassischen Supercomputern.

Es ist viel Bewegung zu verzeichnen in dieser neuen Schlüsseldisziplin, die das Wissen von Biologen, Chemikern und Informatikern zum Nutzen der Life Sciences bündelt. Der Service-Aspekt für die Forschung tritt dabei immer stärker in den Vordergrund. Immer mehr Bioinformatiker verstehen sich als Informationsprovider für molekulare biologische Informationen. Bioinformatik wird nicht nur zu enormer Zeitersparnis in der Forschung führen, sondern in Zukunft auch mehr Möglichkeiten bieten, Behandlungen zum Wohle des Patienten stärker zu individualisieren.

### Weitere Informationen:

[www.zbh.uni-hamburg.de](http://www.zbh.uni-hamburg.de) · [www.carus-it.com](http://www.carus-it.com) · [www.inb.uni-luebeck.de](http://www.inb.uni-luebeck.de)

