Ralf Hofestädt
Universität Bielefeld

Ernst W. Mayr
TU München

Bioinformatik

Vor ca. 10 Jahren begann der systematische Aufbau und Ausbau der Bioinformatik in Deutschland. Dabei stellten die DFG-Initiative „Bioinformatik“ und die BMBF-Initiative „Bioinformatikzentren“ sicherlich die wesentlichen Bausteine für den erfolgreichen Verlauf der Aufbauphase dar. Diese finanziellen Mittel waren auch der Garant für den schnellen und weltweit herausragenden Aufbau der Bioinformatik in Deutschland. Der eigentliche Startschuss der Bioinformatik ist aber sicherlich mit dem humanen Genomprojekt verbunden. Dieses Projekt wurde von J. Watson – dem Entdecker der DNS (Bausteine des Lebens) – initiiert. Dabei wies J. Watson direkt darauf hin, dass dieses Projekt nur sinnvoll und erfolgreich geführt werden kann, wenn auch die entsprechende elektronische Infrastruktur zur Analyse dieser molekularen Massendaten dazu parallel aufgebaut wird. In dieser Phase reagierte die Wissenschaftspolitik in Deutschland hervorragend. Das BMBF leitete Anfang der 90er Jahre unter der Leitung von Prof. Krückeberg die Konzeption des ersten deutschen Förderkonzeptes ein. Unter diesem Dach fanden dann die weiteren Aktivitäten in Richtung Bioinformatik statt. Dabei standen zunächst die Forschungsschwerpunkte der Sequenzanalyse (der Gedanke Watsons) sowie des Protein Designs (Drug Designs) im Zentrum der wissenschaftlichen Aktivitäten. Beide Forschungsgebiete waren nicht neu – im Gegenteil. Unter dem Arbeitsfeld der Theoretischen und Mathematischen Biologie standen diese Forschungsgebiete schon viele Jahre zur Diskussion. Auch die Informatik als anwendungsorientierte Wissenschaft hatte sich in Deutschland bereits von Beginn an breit definiert, und somit erlaubten viele Informatik-Curricula das Nebenfach Biologie. An einigen Universitäten entwickelten sich bereits weit vor dem Jahr 1990 Arbeitsgruppen im Bereich der Bioinformatik (z. B. Bonn). Eine Sonderstellung nahm dabei sicherlich der Weg der Naturwissenschaftlichen Informatik in Bielefeld ein. Wenn man die Geschichte etwas genauer durchleuchtet, dann ist vielleicht sogar J. v. Neumann bereits als Vorreiter der Bioinformatik zu betrachten, weil er bei der Modellierung der Neuronalen Netze wissenschaftlich aktiv war und Werkzeuge wie den Zellularen Automaten dabei entwickelt hat. Aus der Sicht der Informatik sind noch ganz andere interessante Innovationen bekannt, die direkt aus der Analyse biologischer Systeme abgeleitet wurden. So sind neben den Neuronalen Netzen gerade die Genetischen Algorithmen herausragende Beispiele für die Symbiose von Biologie und Informatik. Leider hat gerade dieser Aspekt der direkten Innovation der Informatik keinen Einzug in die Bioinformatik-Curricula der deutschen Universitäten finden können. Vielmehr steht der Gedanke von J. Watson im Vordergrund – die Bioinformatik hat die elektronische Infrastruktur aufzubauen, um die Entwicklung der Molekularen Biologie massiv zu unterstützen. Die Bioinformatik hat sich in den vergangenen Jahren allerdings rasant weiter entwickelt. Natürlich steht auch heute die Sequenzanalyse noch immer im Blickpunkt der Aktivitäten. Jedoch hat auch der technologische Fortschritt, der die vollständige Sequenzierung von Genomen z. T. in wenigen Tagen ermöglicht, auch andere Ansprüche produziert. Die Analyse und Annotation dieser Genome stehen nun im Zentrum der Sequenzanalyse. Dazu nimmt die Entwicklung und Implementierung anspruchsvoller Algorithmen einen immer wichtigeren Teilbereich der Bioinformatik ein. Der Beitrag der Kollegen J. Stoye et al. entspringt diesem Arbeitsgebiet. Die Identifizierung der Gene stellt dabei einen Ausgangspunkt für die weitere Analyse dar. Gene können u. a. Proteine erzeugen, die wiederum als Enzyme biochemische Prozesse in der

Zelle steuern. Heute haben sich rund 1000 Datenbanken weltweit etabliert, die diese molekularen Datenbestände systematisch erfassen und verfügbar machen. Diese Datenbanken und Informationssysteme stellen das Rückgrat der geforderten elektronischen Infrastruktur dar. Die Entwicklung und Implementierung von Integrationswerkzeugen und komplexen Informationssystemen stellt einen weiteren neuen Arbeitsschwerpunkt dar, um gerade im Bereich der biologischen Netzwerkanalyse komplexe biologische Sachverhalte verstehen zu können. Der Beitrag des Kollegen F. Schreiber widmet sich dieser Thematik und zeigt, dass gerade die Konzepte der Visualisierung eine bedeutende Rolle bei der Analyse dieser biologischen Netzwerke einnehmen werden. Neben der Erfassung und Vorhersage von solchen Netzwerken spielt die Modellierung und Simulation gerade im Bereich der Systembiologie eine zentrale Rolle, um letztlich auch die Vision der virtuellen Zelle realisieren zu können. An dieser Stelle scheint gerade die deutsche Informatik/Bioinformatik eine tragende Rolle spielen zu können, weil z. B. Petrinetze einfach per Definition für diese Aufgabe geeignet sind. Seit nunmehr über 10 Jahren steht die Anwendung der Petrinetze in der Molekularen Biologie zur Diskussion und nimmt einen immer größeren und bedeutenderen Raum ein (2010 erscheint dazu ein Buch im Springer-Verlag in der Serie Computational Biology, herausgegeben von I. Koch, W. Reisig und F. Schreiber). Der Beitrag von R. Zimmer et al. wird dieser Situation gerecht und präsentiert aktuelle Ergebnisse dieser Arbeitsgruppe. Der größte Wandel der deutschen Bioinformatik ist aber ganz gewiss im Bereich der Anwendung sichtbar. In den 90er Jahren stand die Anwendung quasi nicht zur Diskussion – das hat sich geändert. Die Hauptanwendung liegt im Bereich der Molekularen Diagnostik und Therapie im humanen Bereich. Die Beiträge der Kollegen T. Lengauer et al. und H.-P. Lenhof et al. unterstreichen diese Phase der Bioinformatik. Dieser neue Schwerpunkt wird in den kommenden Jahren massiv ausgebaut. Dies wird auch durch die Strukturen der GI heute sichtbar. Im vergangenen Jahr wurde der neue Fachbereich „Informatik in den Lebenswissenschaften“ in direkter Kooperation mit der GMDS und den relevanten Untergliederungen der GMDS neu aufgebaut. Die Bioinformatik und die Medizinische Informatik bilden derzeit gemeinsam mit der Umweltinformatik und dem Fachausschuss ASIM den Schwerpunkt der Aktivitäten innerhalb dieses neuen Fachbereiches.

Die Bioinformatik wird im Wesentlichen von der Vision der virtuellen Zelle getrieben. Dahinter steckt die Hoffnung, die Zelle vollständig zu verstehen und virtuell im Rechner verfügbar zu machen. Von diesem Ziel sind wir weit entfernt. Bis heute ist selbst der fundamentale Mechanismus der Genregulation (das systematische An- und Abschalten der Gene) nicht vollständig geklärt. Internationale Projekte zur Erstellung virtueller Zellen laufen aber bereits mit mehr oder weniger sichtbaren Erfolgen. Hier ist das von Prof. Tomita initiierte E-Cell-Projekt (www.e-cell.org/ecell/) sicherlich wegweisend. Das vor einigen Jahren in Deutschland initiierte Projekt der virtuellen Leberzelle (BMBF Schwerpunktprojekt) wurde bereits vor Projektbeginn als nicht realistisch angesehen. Die Zeit für solche Projekte ist noch nicht reif, weil es u. a. an geeigneten Konzepten und Methoden zur Umsetzung der molekularen Phänomene mangelt. Die Bioinformatik wird sich diesem Thema in Teilbereichen sicherlich in den nächsten Jahren nähern. In diesem Themenheft sind wir dieser Vision der virtuellen Zelle indirekt gerecht geworden, indem das Cover des Themenheftes eine abstrakte virtuelle Zelle darstellt, die mit dem Zelleditor von B. Sommer erstellt wurde (www.Cellmicrocosmos.org). Zelleditoren können bereits heute sinnvoll eingesetzt werden und dann nicht nur im Bereich der Lehre Anerkennung finden.

Ralf Hofestädt und Ernst W. Mayr