

## Bioinformatik...

► ...ein Wort, das immer häufiger im Labor Einzug hält und ein rasch expandierendes Gebiet beschreibt, welches noch Mitte der achtziger Jahre von Experimentatoren als exotisch angesehen wurde. Mit dem Siegeszug des PC entstand neben der klassischen theoretischen Biologie das „Biocomputing“. Getrieben durch das Entstehen und Wachsen internationaler biologischer Datenmassive aller Art, insbesondere durch Sequenzdatenbanken, bedurfte es immer mehr des Vergleiches von Laborergebnissen mit vorhandenem Wissen durch Computeranalyse. Die Ausrichtung der Pharma-, Nahrungsmittel- und Agrarindustrie seit Anfang der neunziger Jahre auf die Genomforschung ging einher mit stark steigender staatlicher Unterstützung der „high throughput“-Biologie (z.B. genomweiten Expressionsanalysen, Proteininteraktionsstudien und Strukturbestimmungen). Dadurch steht nun eine Datenflut bevor, die man ohne Bioinformatik (Bio-computing als Name kam aus der Mode) nicht mehr bewältigen kann – die Nachfrage an Bioinformatik als Service zur Datenauswertung ist derzeit so extrem hoch, daß die wenigen qualifizierten Fachleute mit konkurrierenden und äußerst attraktiven Angeboten aus der Industrie überschwemmt werden. Nur zaghaft angedacht sind jedoch die Möglichkeiten und Notwendigkeiten, die sich aus einer Bioinformatik als Wissenschaftszweig ergeben.

Viele Molekularbiologen bedienen sich schon heute bioinformatischer Werkzeuge (das Schreiben von Manuskripten mit einem Textverarbeitungsprogramm gehört allerdings nicht dazu). Neben den Computern werden auch das Internet, Datenbanken und Software, die heterogene Daten verschiedenster Herkunft miteinander verbindet, zur täglichen Laborarbeit gehören. Je größer das Wissen um die Möglichkeiten der *in silico*-Biologie, desto besser kann die Planung von Experimenten erfolgen und umso effektiver wird auch deren Auswertung sein. Schon heute sind wertvolle Daten jedem Labor zugänglich, von denen viele Forscher noch keinen Gebrauch machen. Nehmen wir als Beispiel eine neue Mutation und die Charakterisierung des entsprechenden Gens in *Drosophila*: Es ist üblich geworden, Sequenzdatenbanken nach Homologien abzusuchen. Wenn als einziges ein hypothetisches Gen im Menschen gefunden wird, reicht es oftmals nur zu einer mageren Abbildung in der

entsprechenden Veröffentlichung. Kaum ein *Drosophila*-Forscher überprüft heutzutage, ob das menschliche Gen einem bestimmten Chromosomenabschnitt zugeordnet ist und ob in diesem Abschnitt schon eine Krankheit kartiert ist, die phänotypische Ähnlichkeiten zur *Drosophila*-Mutation aufweist. Dann wäre die Mutation ein Volltreffer in Richtung Identifizierung eines Krankheitsgens. Man kann dies beliebig weiterführen: Jeder menschliche Chromosomenabschnitt ist syntänisch zu entsprechenden Regionen im Mausgenom, von denen Homologe oder vielleicht auch nur Fragmente (expressed sequence tags; ESTs) bekannt sein können, die über „drei Ecken“ Zugang zu phänotypischer Information liefern, die noch nicht in Sequenzdatenbanken erfaßt ist. Solche Analysen sind nur der Anfang. Mit der Sequenzierung von rund 90% des menschlichen Genoms (als „working draft“, also noch mit Lücken und Fehlern behaftet) bis zum Frühjahr nächsten



Jahres werden auch Methoden reifen, die Kartierungsinformation integrieren. Neue Herausforderungen durch die Ausrichtung der molekularen Biologie und Medizin auf andere „high throughput“-Technologien sind jedoch schon an der Tagesordnung und führen zu globalen und vergleichenden Untersuchungen, orthogonal zu klassischen Vorgehensweisen. Die meisten Bioinformatiker werden dabei mit Datenmanagement und Serviceaufgaben zugedeckt. Eine vorherschaubare Entwicklung ist jedoch die Anwendung von Computersimulationen für Vorhersagen und das Design von Experimenten. „Computational Biology“ schwärmen die Amerikaner – Theoretische Biologie wird es oftmals noch skeptisch bis abwertend in Deutschland genannt. Die Modellierung ganzer Zellen am Computer rückt (zumindest bei Optimisten und Visionären) in greifbare Nähe – die „neuen“ Daten bringen mit Sicherheit die Möglichkeit, bestimmte zelluläre Prozesse zu formalisieren. In den USA werden deshalb Physiker und Biologen stimuliert, zusammen in gemeinsamen Instituten solche Visionen Wirklichkeit werden zu lassen.

Doch wie sieht es in Deutschland aus? Das Potential der Bioinformatik ist durchaus erkannt (wobei der Begriff nicht klar definiert ist); die Förderung geht derzeit jedoch hauptsächlich in Drittmittel. Dies führt zur notwendigen Expansion existierender Arbeitsgruppen, läßt jedoch die Anzahl von Bioinformatikgruppen nicht gerade explodieren. Bio-

informatikkurse an den Universitäten laufen an, und erste Bioinformatikprofessuren werden installiert (wobei Kandidaten, die in internationalen Schwerpunktrichtungen arbeiten, schwer zu finden sind und die Ausschreibungen auch nicht unbedingt auf solche Schwerpunktrichtungen zielen) –, es wird also noch eine Weile dauern, bis dringend benötigter Nachwuchs heranreift. Erforderlich ist aber auch eine lokale Bündelung von Aktivitäten, möglichst zusammen mit experimentellen Gruppen, um eine kritische Masse für die Bearbeitung wichtiger biologischer oder medizinischer Fragestellungen zu finden. Eine einzelne Arbeitsgruppe reicht dafür häufig nicht aus. Die Anforderungen an die Bioinformatik steigen und werden mit der Einführung von DNA-Chips eine neue Qualität annehmen. Die Bioinformatik von morgen wird sich viel komplexer darstellen, da neben der Expansion auch ein Trend zur Vernetzung mit anderen Gebieten wie der Chemie und Medizin absehbar ist. Schnittstellen zu „Cheminformatics“ und „Pharmainformatics“ sind schon heute gegeben (zum Beispiel bei der Verbindung von kombinatorischer Chemie mit Proteinstruktur- und Ligandenmodellierung). Bioinformatik wird direkt an den neuen Technologien und Methoden benötigt (z.B. bei der Digitalisierung von Hybridisierungsmustern oder mikroskopischen Aufnahmen) und ist auch essentiell für die Integration von Daten in allgemein verfügbaren Datenbanken. Bioinformatik wird schon bald zur Allgemeinbildung jedes Biologen gehören und im Labor allgegenwärtig sein und wird ein gleichberechtigtes Wissenschaftsgebiet darstellen, das biologische Fragestellungen bearbeiten und Lösungen anbieten kann.

Der Computer führt unausweichlich seinen Siegeszug in der Biologie fort, und es bleibt zu hoffen, daß auch in Deutschland das Potential der Bioinformatik nicht nur vollständig erkannt, sondern auch rechtzeitig (besser gestern als heute) und langfristig genutzt wird. Dazu gehören Ausbildung, strukturbildende Maßnahmen, Kompetenzvernetzung aktiver Bioinformatikgruppen und Bündelung von Aktivitäten sowie die Forcierung interdisziplinärer Teams mit Technologieentwicklungs- und Anwendungsgruppen. Nicht zuletzt ist die Akzeptanz der Bioinformatik nicht nur als Service, sondern auch als eigenständige Forschungsrichtung gefordert.

Peer Bork  
EMBL Heidelberg, und  
Max-Delbrück-Centrum für Molekulare Medizin, Berlin

UB/TIB Hannover



ZL 246

1999 : 5,3

# BIO

# spektrum

ZEITSCHRIFT DER GESELLSCHAFT FÜR BIOCHEMIE UND MOLEKULARBIOLOGIE (GBM),  
DER VEREINIGUNG FÜR ALLGEMEINE UND ANGEWANDTE MIKROBIOLOGIE (VAAM),  
DER GESELLSCHAFT FÜR GENETIK (GfE)  
UND DER GESELLSCHAFT FÜR ENTWICKLUNGSBIOLOGIE (GFE)

NR. 3 • 5. JAHRGANG



- C<sub>α</sub>-FORMYLGLYCIN — EINE NEUARTIGE PROTEINMODIFIKATION IM AKTIVEN ZENTRUM VON SULFATASEN
- BIOSYNTHESE VON CYTOCHROM C IN *E. COLI*
- DAS GENOM VON *C. GLUTAMICUM*
- REGULATION VON PROTEINKINASEN DER MLCK-FAMILIE
- SPECIAL: FLUORESCENZTECHNIKEN
- AKTUELLES AUS WISSENSCHAFT UND TECHNOLOGIE

30167 Hannover

Welfengarten 1 b

Universitätsbibliothek und TIB

400318 68 20 St.

Postvertriebsstück • D 13808 F • Entgelt bezahlt  
polycom GmbH • Potsdamer Str. 77-87 • 10785 Berlin

**ktrum**  
HEMISCHER VERLAG