

Memorandum zur Entwicklung der Bioinformatik in Deutschland

Stand Juni 2000

Erarbeitet von der Präsidiumskommission der GMDS (P. Dirschedl, L. Gierl, K.-H. Jöckel (Vorsitz), R. Klar, H. Schäfer, M. Schumacher)

1. Zielsetzung des Memorandums, Adressatenkreis

Daten aus Molekularbiologie und Genomforschung haben nicht zuletzt durch das Human Genome Project und die damit zusammenhängende Intensivierung der Forschungstätigkeit auf diesem Gebiet einen Umfang angenommen, der einen Bedarf für die informatische und statistische Unterstützung dieser Forschungstätigkeit unabweisbar macht. Gleichzeitig ist ein Defizit an ausgebildeten Wissenschaftlern, die zur Bearbeitung dieser Aufgabenstellungen qualifiziert sind, zu konstatieren. Dieser Bedarf für Statistik und Informatik auf der einen Seite und der Mangel an entsprechendem Personal auf der anderen Seite hat vor mehr als 40 Jahren zur Gründung einer Fachgesellschaft geführt, die heute den Namen Gesellschaft für Medizinische Informatik, Biometrie und Epidemiologie (GMDS) e.V. führt. Auch damals wurde in der Medizin, weniger allerdings in der grundlagenorientierten, mehr in der anwendungsnahen Forschung sowie in der Krankenversorgung ein entsprechender Bedarf festgestellt. Eine Zielstellung des Memorandums ist es daher, Lehren aus der vergangenen Situation und deren Bewältigung für die Problemstellung der Gegenwart zu ziehen.

Aber nicht nur in einer historischen, sondern auch in einer fachlichen Perspektive sollen mit diesem Memorandum Hinweise für die weitere Entwicklung der Bioinformatik gegeben werden. Wenngleich die Genomforschung und Molekularbiologie biologische

Artengrenzen überwinden und zu Recht die fachbereichsübergreifende Kooperation gefordert wird, so will dieses Memorandum doch auch spezifische Hinweise für die Entwicklung der Bioinformatik **in der Medizin** geben.

Adressaten des Memorandums sind demzufolge

- Mitglieder aus der Fachgesellschaft, die sich wissenschaftlich und/oder organisatorisch zur Bioinformatik positionieren wollen,
- Medizinische Fakultäten sowie kollegiale Organe, die Modelle für die Organisation bioinformatischer Forschung überlegen,
- Hochschulplaner und Wissenschaftsministerien sowie
- Forschungsförderer.

Ihnen sollen aus Sicht der GMDS Hinweise zur Schaffung einer bioinformatischen Infrastruktur für die Grundlagenforschung (perspektivisch auch anwendungsnah) in der Medizin gegeben werden.

2. Begriffsbestimmung, nationale und internationale Ausgangssituation

Nach der DFG-Stellungnahme "Perspektiven der Genomforschung" beschäftigt sich die Bioinformatik mit der "Analyse und informatischen Unterstützung von Daten aus Molekularbiologie und Genomforschung". Ob sie damit gleich, wie in der zitierten Stellungnahme geschehen, als "neue Wissenschaft" zu bezeichnen ist oder eher als neues interdisziplinäres Arbeitsgebiet etablierter Wissenschaften, sei zunächst einmal dahingestellt. Unstrittig ist, dass es sich um ein stark expandierendes Gebiet mit einer ausgeprägten biologischen Komponente handelt. Es existieren engste Berührungen zur Molekularbiologie, aber auch zu anderen medizinischen Grundlagenfächern wie etwa der Biochemie oder der Genetik. In vorher nicht gekannter Weise werden Datenmengen generiert, die die in der Grundlagenforschung üblichen Dimensionen an Zahl und Komplexität sprengen. International hat dies zur Etablierung einer wissenschaftlichen Infrastruktur für Bioinformatik geführt. Zu nennen sind hier zum Beispiel:

- wissenschaftliche Publikationsmedien wie Computational Biology, Bioinformatics, J. Biopharm. Stat., J. Comp. Chem., Comput. Chem., Microb. Comp. Genomics, J. Comput. Biol., Genetic Testing

- aktive Forschergruppen, Bioformatik-Netzwerk der HGF (GSF, DKFZ, MDC, GBF, GMD), Universität Stanford, Vereinigung für Allg. und Angewandte Mikrobiologie (VAAM) Fachgruppe: funktionelle Genomanalyse, The Bioperl Project <http://bio.perl.org>, The Sanger Centre <http://www.sanger.ac.uk/>, St. Louis Genome Sequencing Center <http://genome.wustl.edu/gsc/index.shtml>., The EMBL – European Bioinformatics Institute (EBI) <http://www.ebi.ac.uk/>
- zentrale Strukturen wie das NCBI (National Center for Biotechnology Information) in den USA

In Deutschland dagegen gibt es nur vereinzelte Gruppen hochmotivierter Forscher, und im universitären Bereich sieht die DFG die Bioinformatik sogar noch als Sache einzelner Doktoranden, in der Ausbildung als "beinahe nicht existent".

3. Aufgabenstellung der Bioinformatik in Forschung und Entwicklung

Molekularbiologie und Genomforschung werden in Zukunft eine Schlüsselstellung der medizinischen Grundlagenforschung erhalten. Ein wichtiges Ziel der Humangenomforschung ist dabei die Aufklärung der molekularen Ursachen genetisch bedingter Erkrankungen.

Hierzu zählen sowohl die klassischen monogenen Krankheiten als auch die multifaktoriell bedingten Volkskrankheiten. Nach der Gensequenzierung mit ihren erheblichen Datenvolumina steht die Frage der Aufklärung der Funktion, Interaktion und Regulation von Genen im Vordergrund. Aus deren Kenntnis heraus lassen sich wichtige Schlussfolgerungen über ätiologische Faktoren, diagnostische Verfahren und therapeutische Ansätze ableiten.

Allen genannten Gebieten ist gemein, dass beträchtliche Datenmengen anfallen, die einer Strukturierung und anschließenden Aufbereitung zugeführt werden müssen. Dabei bezieht sich das Adjektiv "groß" auf das, was bislang im Bereich der Biologie und medizinischen Grundlagenforschung üblich war.

Es werden dabei Anforderungen gestellt, die sich in drei Gruppen unterteilen lassen:

1. die Anwendung von vorhandenem, technologischem Wissen und die Übertragung aus anderen Anwendungsgebieten der Informatik und Statistik in den Bereich der Molekularbiologie. Zu nennen ist hierbei etwa die Modellierung, Erstellung, Integration und das Transaktionsmanagement von Datenbanken, sowie das Retrieval in heterogenen Datenbanken.
2. die Bewertung vorhandener Verfahren der Informatik und Statistik für ihre Eignung im Einsatz zur Analyse von Daten aus der Genomforschung und Molekularbiologie. Hierzu zählen z. B. statistische Methoden der Klassifikation wie die Clusteranalyse, die bspw. für einen Einsatz in der Analyse von DNA-Chipdaten zu evaluieren sind.
3. das Arbeiten mit den molekularbiologischen Daten zur Hypothesengenerierung und dem Auffinden von Gesetzmäßigkeiten. Hierin ist bspw. das sogenannte Data mining eingeschlossen, wobei die Bioinformatik im Anschluss an die Bildung von Hypothesen aus der Datenanalyse auch Hinweise für die Bestätigung derselben in adäquaten Experimenten geben muss. Als Fernziel ließe sich hier sicherlich auch die Mathematisierung der Biowissenschaften formulieren.

Gemeinsam ist den o.g. Desiderata, dass sich methodische mit inhaltlichen Anforderungen vermengen, wobei die Bioinformatik perspektivisch zentrale Kompetenzen in beiden Bereichen aufweisen muss. Zu den Forschungsaufgaben der Bioinformatik gehört somit sowohl die Schaffung neuer inhaltlicher Erkenntnisse (z.B. Untersuchung genomweiter Expressionsmuster) als auch die Bewertung bekannter und die Entwicklung neuer Methoden zur Erreichung dieser Ziele (bspw. der Clusteranalyse). Medizinische Informatik, Biometrie und Epidemiologie können hierzu substantielle Beiträge liefern, da sie viele dieser Methoden bereits aus anderen Anwendungsgebieten kennen.

Darüber hinaus sind klinische Anwendungen bereits vorherzusehen (z. B. Pharmakogenetik, Charakterisierung von Tumoren anhand von Expressionsmustern, etc.), die den Einsatz und die Modifikation etablierter Verfahren der klinischen Forschung (Therapiestudien, Prognosemodelle, etc.), aber auch methodische Neuentwicklungen erforderlich machen werden.

4. Grundlagen der Bioinformatik

Damit ergibt sich die Bioinformatik als eine interdisziplinäre Aufgabenstellung der Life Sciences. Der/die Bioinformatiker/in muss über Kernkompetenzen sowohl in den methodischen Disziplinen als auch im Inhalt verfügen. Die inhaltlichen Kompetenzen erstrecken sich dabei auf die biochemischen, biologischen und medizinischen Grundlagen und ein gründliches Verständnis der Genetik. Auf der methodischen Seite werden Grundlagen der Mathematik und Informatik, insbesondere der Statistik erwartet. Essentiell für eine effiziente Arbeit sind ein interdisziplinäres Selbstverständnis und Erfahrungen in der empirischen Forschung.

5. Strukturelle Entwicklung (Ist und Soll-Situation)

5.1 Lehre und Forschung an den Hochschulen

Im derzeitigen Medizinstudium findet eine Ausbildung in Bioinformatik praktisch nicht statt. Allerdings lernen die Medizinstudenten bereits jetzt in der Veranstaltung Biomathematik, sowie im Kursteil Medizinische Statistik und Medizinische Informatik im Rahmen des Kursus des ökologischen Stoffgebietes Basiskomponenten aus Statistik und Informatik kennen, die auch für die Bioinformatik relevant sind. Der von den Fachvertretern für Medizinische Informatik, Biometrie und Epidemiologie derzeit in Vorbereitung befindliche neue Gegenstandskatalog für die drei Fächer sieht unter anderem auch die Berücksichtigung der Bioinformatik vor. Dennoch ist das Fachgebiet in der Medizin hoffnungslos unterentwickelt, vergleichbares gilt aber auch für die übrigen Life Sciences an den deutschen Universitäten.

Zur Soll-Situation:

1. In der Medizin: Im Rahmen der grundständigen Medizinausbildung sollte eine Basisinformation zum Thema Bioinformatik erfolgen. Die Studenten sollten erkennen, welche grundsätzlichen Methoden mit welchen erkenntnistheoretischen Grenzen für bioinformatische Fragestellungen zum Einsatz kommen. Im Rahmen von Spezialkursen für Doktoranden sollten spezifische Kenntnisse in Abhängigkeit von konkreten Projekten der medizinischen und klinischen Grundlagenforschung vermittelt werden. Die Vermittlung dieser spezifischen Kenntnisse für Doktoranden und die Betreuung molekularbiologischer Projekte in der Medizin sollte durch eine

Arbeitsgruppe erfolgen, die über die erforderlichen informatischen und biostatistischen Kompetenzen verfügt. Minimal sollte eine drittmittelunabhängige, grundfinanzierte Arbeitsgruppe bestehend aus einem Wissenschaftler und einer Medizinischen Dokumentarin / einem Medizinischen Dokumentar für die Aufgabe vorgesehen werden. Wünschenswert wäre, wenn an jeder Universität mittelfristig für diese Arbeitsgruppe eine Professur mit enger Anbindung an (optimal Integration in) ein Institut für Medizinische Informatik, Biometrie und Epidemiologie vorgesehen würde.

2. Außerhalb der Medizin: An Standorten mit einer starken molekularbiologischen Forschung außerhalb der Medizin erscheint die Einrichtung eines Lehrstuhls für Bioinformatik in der naturwissenschaftlichen Fakultät sinnvoll und notwendig. Verbunden hiermit sollte die Einrichtung eines interdisziplinären Studienganges (Diplom, Bachelor oder Master) sein, der die vorhandenen Kompetenzen der Universität zur Bioinformatik bündelt. Als minimale Ausstattung einer solchen Professur muss neben der Stelle des Lehrstuhlinhabers (C4) die Einrichtung von drei wissenschaftlichen Mitarbeitern / Mitarbeiterinnen, zwei technischen Mitarbeitern / Mitarbeiterinnen und einer Sekretariatsstelle angesehen werden. Als Minimalvoraussetzung gilt darüber hinaus das Vorhandensein einer biologischen und informatischen Grundausbildung an der Universität, wobei eine Beteiligung des Lehrstuhls / Instituts für Medizinische Informatik, Biometrie und Epidemiologie zwingend erscheint.

5.2 Außeruniversitäre und industrielle Forschung

Unbeschadet bedeutsamer Ansätze in der außeruniversitären Forschung (Bioinformatiknetzwerk und zahlreiche Einzelaktivitäten, Großforschungseinrichtungen des Bundes und der Länder) und wichtigen Gruppen in der Industrie muss ein genereller Mangel kompetitiver Forschung in der Bioinformatik in der Bundesrepublik Deutschland konstatiert werden. Es erscheint zwingend, dass in den einschlägigen Forschungseinrichtungen (z. B. in Instituten der Helmholtz-Gemeinschaft, aber auch der Industrie) eigenständige Forschergruppen etabliert werden, die zumindest aus einem Senior Scientist, ein bis zwei Junior Scientists und entsprechendem technischem Personal zusammengesetzt sind.

6. Studium, Aus- und Weiterbildung in Bioinformatik

Das generell, insbesondere in Deutschland, zu diagnostizierende Defizit an qualifizierten Bioinformatikern sollte Anlass sein, **auf allen Ebenen** die Ausbildung von qualifiziertem bioinformatisch tätigen Personal zu fördern. Hierzu ist die Einrichtung akademischer Studiengänge (vgl. 5.1), aber auch eine Weiterbildung naturwissenschaftlicher und / oder medizinisch vorqualifizierter Personen und eine Ausbildung bioinformatisch ausgebildeten technischen Personals dringlich erforderlich. Die GMDS rät dringend, Studiengänge für Bioinformatik einzurichten und gleichzeitig sinnvolle Weiterbildungsperspektiven für ärztliches und naturwissenschaftliches Personal zu entwickeln. Durch geeignete Kombination vorhandener Ressourcen im Bereich der Informatik, Statistik und Biologie könnten und sollten grundständige Studiengänge für Bioinformatik etabliert werden. Darüber hinaus sollten Weiterbildungsstudiengänge eingerichtet werden, die motivierten und qualifizierten Naturwissenschaftlern und Medizinern die Möglichkeit eröffnen, einen entsprechenden berufsqualifizierenden Abschluss zu erreichen. Insbesondere sollte über die Möglichkeit einer ärztlichen Zusatzbezeichnung für "Medizinische Bioinformatik" nachgedacht werden.

Zusätzlich zum akademischen Abschluss und zur akademischen / ärztlichen Weiterbildung sollte der Bedarf an bioinformatisch qualifizierten Mitarbeitern durch die Einrichtung entsprechender Fachschulen / Weiterbildungsangeboten Rechnung getragen werden.

Sind entsprechend Lehrangebote außerhalb der Medizin angesiedelt, sollte eine gebührende Vermittlung medizinisch orientierter Inhalte durch Beteiligung der Medizinischen Fakultäten sichergestellt werden.

7. Projektförderung, Förderprogramme nationale und internationale Kooperationen

Die von den Ländern (z. B. NRW), aber auch der DFG initiierten Förderprogramme können relevant zur Etablierung einer bioinformatischen Forschungslandschaft in Deutschland beitragen. Darüber hinaus erscheint es dringend erforderlich, dass der Bund mit geeigneten Fördermaßnahmen (Projekt- und Verbundförderung) die nötige

Anschubfinanzierung zur Etablierung einer Forschungsinfrastruktur leistet. Besonderes Gewicht sollte dabei auf die Etablierung nationaler und internationaler Kooperationen gelegt werden.

8. Ausblick

Ziel der gemeinsamen Bemühungen interessierter Gruppen aus Forschung, Lehre und industriellen Anwendungen muss es sein, möglichst schnell eine bioinformatische Infrastruktur zu schaffen, die die deutsche Forschung international konkurrenzfähig macht, bzw. bereits erreichte Forschungsvorsprünge sichert. Hierzu sollte sowohl im Gebiet der akademischen Ausbildung als auch der Schaffung kompetitiver Forschung / Infrastrukturen ein entsprechendes Commitment des Bundes und der Länder, aber auch anderer Forschungsförderungsinstitutionen erfolgen.

9. Zusammenfassung

Um die für die weitere bioinformatische Forschung, insbesondere im Bereich der Medizin, erforderliche Infrastruktur zu schaffen, erscheint es unerlässlich, folgende Strukturen kurz- bis mittelfristig aufzubauen:

- Die Schaffung einer personell belastbaren Infrastruktur (mindestens eine Wissenschaftlerstelle, möglichst eine Professur mit wissenschaftlichem und technischem Personal) in enger Anbindung an die oder durch Erweiterung der Institute für Medizinische Informatik, Biometrie und Epidemiologie der Medizinischen Fakultäten.
- Der Etablierung grundständiger bioinformatischer Studiengänge an mindestens 5 Universitäten/Hochschulen der Bundesrepublik Deutschland
- Der Einrichtung von Fachschulen für das erforderliche bioinformatisch ausgerichtete, technische Personal in mindestens drei bis vier Bundesländern
- Der Etablierung von Forschungsförderungsprogrammen des Bundes und der Länder, die einen engen Bezug zur Einrichtung entsprechender Aus- und Weiterbildungsprogramme der jeweiligen Universitäten aufweisen.

Für alle diese Aktivitäten steht die GMDS mit den bei ihren Mitgliedern vorhandenen Kompetenzen beratend zur Verfügung.