

Die Rolle der Mathematik in der Biologie des genomischen Zeitalters

Ich habe mir zu der Frage Gedanken gemacht, ob Biologie mathematisierbar ist. Rein empirisch denke ich sofort: Nein! Vor allem, wenn mir einfällt, daß ein Drittel der Abiturienten in die Biologie oder Medizin geht, weil sie für den Rest ihres Lebens genug von Mathematik haben und hoffen, damit nicht weiter behelligt zu werden. Wenn man sich die Grundlagen dieser beiden Fächer näher ansieht, sind doch Begriffswelt, Denkweise und Gesetzmäßigkeit fundamental verschieden. Die Axiomatik und Deduktivität der Mathematik, ihr reduktionistischer, ja minimalistischer Ansatz, die völlig verschiedenen Charaktere von Beweisbegriff und Beweisstruktur zwischen diesen beiden Fächern – nein, eigentlich sind sie inkompatibel. Das ist meine These. Aber andererseits habe ich nun auch schon viele Jahre, sogar Jahrzehnte Spaß gehabt, als Biomediziner immer wieder einmal auf Mathematik zu stoßen. Hierzu mußte ich eine Reihe von Gebieten der Mathematik erlernen, die strategische Bedeutung in unserem Fach, in unserer Forschung wie in unserer Praxis hatten. Also vor allem selbstverständlich die mathematische Statistik, ihre Basis: die Probabilistik und die Stochastik – das sind Gebiete der Mathematik, deren Entwicklung zum Teil von der Biologie direkt bestellt worden ist. Denken Sie nur an die Varianzanalyse als mathematische Disziplin, die ausdrücklich für biologische Sachverhalte erfunden worden ist. Oder denken Sie an die Koaleszenztheorie, mit der man heute versucht, den gegenwärtigen Zustand von genetischen Populationen zurückzuverfolgen, um unter gewissen, wiederum minimalistischen Grundannahmen den wahrscheinlichen Verlauf der Evolution zu rekonstruieren. Oder an die Theorie dynamischer Systeme, die räumlichen und zeitlichen hierarchischen Systeme, die man mit ihr erzeugen und analysieren kann, einschließlich Multistabilität und Chaotik. Hinzu kommt alles, was mit Systemtheorie von Wechselwirkungen zu tun hat, Graphentheorie, Netzwerktheorie, Steuerungstheorien und dann schließlich die ganzen Entwürfe, die Systemtheorie des Zellstoffwechsels oder die Kinetik und Thermodynamik biochemischer Reaktionen. Das alles muß man lernen, wenn man theoretisch in der Biologie und Medizin arbeitet. Es ist ganz erstaunlich, in welchem Maße die Rolle der Mathematik sich in den Jahrzehnten, die ich überblicke, von einem mehr oder weniger freundlich akzeptierten Status als Hilfsdienst, als Magd der biologi-

schen oder medizinischen Forschungspraxis zur Inhaberin einer strategischen Position gemausert hat. Die moderne Genomforschung etwa ist nicht nur vom Einsatz von Computern, sondern auch vom Einsatz mathematischer Begriffe und Sätze und von Methoden abhängig. Nur ein Beispiel: Alle Welt, die in der Genomforschung arbeitet, kennt das Verb „Blasten“. Das ist ein Suchverfahren, mit dem man in einer unübersehbaren Menge von DNS oder Eiweißstrukturen homologe, das heißt ihrer Herkunft nach verwandte Strukturen herausfindet. Die Methode, mit der das möglich ist, ist im Grunde genommen eine erweiterte Form eines ganz primitiven, wahrscheinlichkeitstheoretischen Problems, nämlich: Wenn ich sehr oft eine Münze hochwerfe und Kopf und Zahl registriere – wie häufig ist es da, oder wann ist es zu erwarten, oder wie überraschend ist es, wenn eine ununterbrochene Folge von Köpfen herauskommt. Zehn oder fünfzehn Mal ist wohl das äußerste, wenn die Münze nicht einseitig beklebt ist oder irgendwie sonst unfair geworfen wird. Und die ganze Theorie, die dazu gehört, ist von Erdős und Renyi in dieser einfachen Form ausgearbeitet worden und von Samuel Karlin erweitert worden für Gensequenzen, von denen wir ja Millionen und bald auch Milliarden haben und immer mehr bekommen. Dann hat der US-amerikanische Mathematiker Karlin eine Theorie zufälliger Sequenzähnlichkeit entwickelt, schließlich noch ein Tool, mit dem heute Sequenzanalyse auf dem Computer durchgeführt wird, und das für die Zuverlässigkeit und Genauigkeit und den heuristische Wert solcher Analysen unabdingbar, von strategischer Bedeutung ist. Dabei ist interessant, daß der Ansatz nicht aus einer Reduktion des Lebendigen folgt, aus einer physikochemischen Theorie der Evolution, sondern ein rein heuristischer Ansatz ist, bei dem die Probabilistik direkt ein Modell liefert, wie man einen in zahllosen Daten versteckten Sachverhalt aufklären kann.

Die These ist also, daß Mathematik und Biologie gar nicht vereinbar sind, obwohl sie praktisch unabdingbar miteinander strategisch verbunden sind. Ein zweites Paradoxon: Die meisten mathematischen Anwendungen, die ich in unserem Fach sehe, sind eigentlich physikalisch-chemische Modelle – zum Beispiel eine Nervenzelle mit ihren Ionen-elektrischen Zuständen, oder in einem chemischem System etwa die Belousov-Zhabotinsky-Reaktion. Und dort trifft man dann natürlich soviel Mathematik, wie in Physik und Chemie bereits als Grundlage vorhanden sind. Ähnlich ist es mit der Kristallographie von Proteinen. Aber erheblich zugenommen hat ein anderer Zugang, nämlich die direkte Heuristik, die über Simulation großer Systeme erfolgt. Simulationsmodelle können ganz phänomenologisch aufgebaut sein, was nicht unbedingt auf irgendwelche fundamentalen Gleichungen zurückführbar sein muß. Auch das sogenannte *data mining* in großen Datenbanken gehört hierher. Das sind mathematische Methoden, um aus einer riesigen Zahl

von Literaturstellen die passenden herauszusortieren, ohne daß die wichtige Information in einem Müllberg von irrelevanten Treffern versinkt. Auch die diskrete Mathematik und speziell Kombinatorik gehen ohne Umweg über Physikochemie in die Analyse genetischer Systeme von Populationen ein. Ich habe den Eindruck, daß diese direkte Ehe zwischen angewandten mathematischen Disziplinen und der Biologie stark zunimmt.