

# Urknall in der Biologie

Genetische Analysen von Krebszellen produzieren Millionen von Gigabytes DNA-Daten. Bioinformatiker Michael Baudis macht sie für die Forschung zugänglich und erforscht, wie sich Krebszellen verändern. Von Stefan Stöcklin

Michael Baudis ist ein gefragter Mann. Als Bioinformatiker und Molekularbiologe hat er ein Profil, das die Krebsmedizin dringend braucht, um aus den unübersichtlichen Gendaten Tausender von Patienten die massgebenden Muster herauslesen zu können. Baudis hat diese Entwicklung vor 25 Jahren antizipiert und entsprechende Weichen gestellt: Als Doktorand im Deutschen Krebsforschungszentrum der Universitätsklinik in Heidelberg analysierte er Mitte der 1990er-Jahre die genetischen Ursachen von bösartigem Lymphdrüsenkrebs. Gleichzeitig programmierte er einfache elektronische Systeme, um die Gendaten verschiedener Experimente zusammenzuführen. «Damals hat es mich gepackt», sagt der Bioinformatiker vom Institute of Molecular Life Sciences der UZH.

Heute ist er ein begehrter Experte für Tumordatenbanken. Wenn man ihn angesichts seines Werdegangs als Grenzgänger zwischen Medizin, Genetik und Informatik bezeichnet, versteht er das als Kompliment und nickt zustimmend: «Die Genomrevolution der letzten Jahre ist eine Informationsrevolution», sagt Baudis. Eine Art Urknall in der Biologie, der zu Millionen Gigabytes von Daten geführt hat, die von Spezialisten wie ihm analysiert werden müssen.

## Millionenfach billiger, tausendfach schneller

Um den atemberaubenden Umbruch in der Genomik zu verstehen, reicht ein Blick zurück auf die Geschichte der Analyse und Charakterisierung von DNA. Als Michael Baudis Medizin studierte, war die DNA-Sequenzierung noch weitgehend Handarbeit. Die Abfolge der Bausteine im Erbgut musste in einer Kaskade von Experimenten mühsam eruiert werden. In jene Zeit Anfang der 1990er-Jahre fällt auch der Startschuss des Human Genome Project. Das in den USA und

Grossbritannien durchgeführte Megavorhaben hatte die Analyse aller 23 menschlichen Chromosomen zum Ziel. Es dauerte 13 Jahre und kostete 2,7 Milliarden Dollar, um die Abfolge der rund drei Milliarden DNA-Bausteine eines menschlichen Genoms zu ermitteln. Heutige Sequenzier-

## Bioinformatik

### 3 Milliarden DNA-Bausteine

Tumoren entstehen meist durch Mutationen im Erbgut und führen dazu, dass sich Zellen unkontrolliert vermehren. Bioinformatiker analysieren diese Genomveränderungen und vergleichen dazu die drei Milliarden Bausteine der DNA von Krebspatienten mit denen von gesunden Personen.

geräte schaffen das in zwei, drei Tagen zu Kosten von rund 1500 Dollar. Das heisst, es ist heute millionenfach billiger und tausendfach schneller, das Genom eines Menschen zu sequenzieren. Diese Revolution hat einen Boom ausgelöst, der noch lange nicht zu Ende ist.

Die Mediziner interessieren sich für diese DNA-Daten, weil Krebs im Kern eine Genomkrankheit ist. Tumoren entstehen meistens durch Mutationen im Erbgut und führen dazu, dass sich die Zellen unkontrolliert vermehren. Bioinformatiker analysieren diese Genomveränderungen und vergleichen dazu die drei Milliarden Bausteine der DNA von Patienten mit denen von gesunden Personen. So einfach, wie man sich das ursprünglich vorgestellt hat, ist es allerdings nicht. Vor zwanzig Jahren dachte man, es gäbe vielleicht ein, zwei Dutzend Krebsgene. «Heute wissen wir, dass jeder Krebsfall ein individuelles

genetisches Muster zeigt», sagt Michael Baudis. Und aufgrund der instabilen DNA können sich sogar innerhalb ein und desselben Patienten Krebszellen genetisch voneinander unterscheiden.

Im Fall von Brustkrebs kennt man zum Beispiel die Gene BRCA1 und BRCA2, die in rund 50 Prozent der familiären Brustkrebsfälle verändert sind. Aber es gibt hunderte unterschiedlicher Mutationen in diesen Genen, die das Krebsgeschehen unterschiedlich beeinflussen. Bei anderen Krebstypen ist es ähnlich: Man findet zwar immer wieder die gleichen Krebsgene wie zum Beispiel das TP 53, ein Suppressor-Gen, oder RAS, das in die Signalverarbeitung der Zelle involviert ist. Aber daneben gibt es Tausende von Mutationen in vielen weiteren Genen, die zu einem höchst komplizierten Bild führen.

Kommt dazu, dass sich die Mutationen eines Krebses im Lauf der Zeit verändern und laufend neue Fehler akkumulieren. «Wir suchen nach den Mutationsmustern, um die Tumoren besser zu charakterisieren», sagt Baudis. Mit dem Ziel, die Behandlung anzupassen und zu individualisieren. Das Stichwort dazu lautet «Precision Medicine» (Präzisionsmedizin).

## Politisch unwillkommen

Michael Baudis darf sich als Pionier der Bioinformatik und des Precision Medicine rühmen, die heute als Schlüsseltechnologien der Zukunft gelten. Vorgezeichnet war dieser Weg nicht. Baudis ist in der ehemaligen DDR in Grossenhain (Sachsen) aufgewachsen und flog 1986 als «politisch unwillkommene» Person kurz vor dem Abitur gar von der Schule. Grund war ein Ausreiseantrag nach Westdeutschland, den er zuvor gestellt hatte. «Ich hatte Mühe mit diesem bizzaren System und vertrug die oppressive Mentalität nicht», sagt der Wissenschaftler. Wie er heute weiss, war es die für Bildung zuständige Margot Honecker und Ehefrau von Generalsekretär Erich Honecker höchstpersönlich, die den Rausschmiss veranlasst hatte. Baudis arbeitete nach dem Rauswurf noch während vier Jahren als Krankenpfleger in der DDR. Erst kurz vor der Wende Anfang Oktober 1989 konnte er das marode Land legal verlassen.

Danach holte er im westdeutschen Mannheim nach, was ihm der DDR-Staat verweigert hatte,

und beendete seine Schulkarriere als 24-Jähriger mit dem Abitur. «Zumindest einen positiven Aspekt hatte die Verzögerung», sagt Michael Baudis. Im Gymnasium lernte er seine Frau Laura kennen, die wie er aus politischen Gründen Rumänien verlassen hatte. Eine Verbindung, die halten sollte: Gemeinsam verfolgten die beiden ihre wissenschaftliche Karriere, sie als Teilchenphysikerin, er als Bioinformatiker und Tumorgenetiker. Parallel absolvierten sie ihre Lehr- und Wanderjahre in den USA und kehrten gemeinsam mit ihren Kindern zurück nach Europa. Nun sind sie zusammen an ein und derselben Fakultät der Universität Zürich als Professoren tätig.

Michael Baudis hat zusätzlich zum Medizinstudium eine vier Jahre dauernde Doktorarbeit in Tumorgenetik geschrieben. Das passt zum eigenständigen Profil des Wissenschaftlers. «Das ganze Feld der DNA-Analyse war noch in den Kinderschuhen, die neuen Möglichkeiten zeichneten sich erst unscharf ab», sagt Baudis. Klar war nur, dass man die Krankheit Krebs an der genetischen Wurzel packen konnte – «eine faszinierende Perspektive, an der ich mitarbeiten wollte», so Baudis. Erstmals war es zum Beispiel möglich, krebserzeugende Mutationen in Spezialmikroskopen abzubilden, bei der kleinere oder grössere Abschnitte der DNA multipliziert sind oder fehlen. Diese Duplikationen und Deletionen tragen häufig zur Entstehung von Krebs bei, indem sie andere Krebsgene aktivieren, Suppressor-Gene hemmen oder die DNA-Regulation stören.

#### «Wir sind Datenparasiten»

Da Michael Baudis schon immer ein Flair fürs Programmieren hatte, begann er mit dem Aufbau digitaler Archive, in denen diese DNA-Vervielfältigungen systematisch aufgelistet und dargestellt sind. 2001 veröffentlichte er die erste Version der Progenetix-Datenbank, die diese «Copy Number Aberrations» auflistet, wie dieser Mutationstyp im Jargon heisst. Sie enthält unterdessen die Daten von 364 verschiedenen Krebstypen aus über 32 000 Experimenten. 2012 erweiterte Baudis diese Datenbank mit der «arrayMap». Sie basiert auf einer Weiterentwicklung der ursprünglichen Technik und erlaubt eine höhere Auflösung in den DNA-Veränderungen. Das heisst, Mutationen kleinerer Sequenzabschnitte lassen sich vergleichen.

Die zugrundeliegenden DNA-Daten erhebt Baudis nicht selbst, stattdessen nutzen er und sein Team öffentlich zugängliche Daten aus grossen Repositorien, am wichtigsten sind die Ressourcen des amerikanischen NCBI (National Center for Biotechnology Information) und des European Bioinformatics Institute (EBI). «Wir sind Daten-

---

*Jeder Krebsfall zeigt ein individuelles genetisches Muster. Sogar bei einem einzelnen Patienten können sich die Krebszellen genetisch unterscheiden.*

---

parasiten», schmunzelt Baudis. Sein Gebiet der Tumordatenbanken profitiert von der wachsenden Bewegung für Open Access (siehe Kasten), die den Zugang und die Offenlegung von Forschungsdaten propagiert. Erst der freie Austausch ermöglicht den Aufbau der Datenbank, gleichzeitig ist sie selbst ein kostenloser Service für die Community.

Dank dieser Dienstleistung kann jeder Arzt oder Forscher in Progenetix nachschauen, wo bestimmte Mutationen mit welcher Häufigkeit in den archivierten Krebstypen auftreten. Die Fachkollegen aus der Onkologie und der Genetik nutzen die Dienste rege. «Besonders interessant sind die direkten Nachfragen zu speziellen Themen», sagt Baudis. Sie sind häufig Grundlage für Fachartikel: Eine kürzlich veröffentlichte Studie zur Metastasierung bei Brustkrebs im Fachblatt «Nature» stützte sich zum Beispiel auf Progenetix-Daten. Ebenso eine grosse Studie über die Genetik seltener und meist unheilbarer Tumoren im Hirnstamm von Kindern (Diffuse Intrinsic Pontine Glioma). Die Progenetix-Daten führten zu einem neuen Projekt zur Identifizierung von Subtypen des Tumors, was neue Ansätze zur Behandlung eröffnen dürfte.

Der Datenpionier sieht mit Genugtuung, dass die Bioinformatik in den letzten Jahren Fahrt aufgenommen hat. Angetrieben durch die Genomrevolution haben sich landauf, landab Initiativen entwickelt, um möglichst viele DNA-Daten von Krebspatienten zu generieren und zu analysieren. Laut einer aktuellen Untersuchung sind weltweit über 100 Konsortien daran, DNA-Krebsdaten zu erheben, hauptsächlich in den USA, Kanada und Westeuropa. In den USA läuft ein

entsprechendes Programm unter dem klingenden Titel «Cancer Moonshot». Angesichts dieses Booms sind übereinstimmende Formate gefragt, um die Daten der verschiedenen Institute und Kliniken vergleichen zu können.

Baudis engagiert sich deshalb in der Global Alliance for Genomics and Health (GA4GH), die genau dieses Ziel verfolgt: ein internationales Netzwerk, in dem genomische und klinische Daten ausgetauscht werden. Statt einer zentralen Datenbank verfolgt die 2013 lancierte Initiative eine dezentrale Lösung, in der lokale Datenbanken über gemeinsame Protokolle und Interface verbunden werden. Baudis arbeitet im sogenannten Beacon-Projekt mit, durch das DNA-Mutationen bis auf einzelne Basenbausteine der DNA abgefragt werden können.

#### Open Access

### Daten für alle

Was mit öffentlichen Geldern finanziert wird, soll auch frei zugänglich sein. Im Rahmen von Open Access verlangen Forschungsförderungsinstitutionen den kostenlosen und freien Zugang nicht nur zu den Publikationen, sondern auch zu den Daten. Auf nationaler Ebene ist es der Schweizerische Nationalfonds, der von den Forschenden «erwartet», dass sie ihre Daten anderen Wissenschaftlern und Wissenschaftlerinnen zugänglich machen und in öffentlichen Datenbanken speichern. Ab diesem Oktober müssen deshalb alle Gesuchsteller einen «Data Management Plan» als integralen Teil ihres Gesuchs einreichen. Auf internationaler Ebene ist die «European Science Cloud» im Entstehen begriffen. Das weltweit grösste Repositorium soll ab 2020 in Betrieb gehen und allen Forscherinnen und Forschern Datenzugang gewähren, darauf haben sich die wichtigsten Forschungsinstitutionen Europas im Juni 2017 geeinigt. 200 Millionen Euro will die Europäische Kommission dafür ausgeben. Wegweisend für den freien Zugang waren amerikanische und britische Förderungsinstitutionen, die im Rahmen des Human Genome Project bereits in den 1990er-Jahren genetische Primärdaten veröffentlicht hatten. Open Access umfasst aber alle Disziplinen.

Im Zusammenhang mit weltweiten, umfangreichen Netzwerken stellt sich ein neues Problem immer dringender: die Sicherstellung der Anonymität. Analysiert man DNA-Daten auf individuelle Mutationsmuster in Kombination mit klinischen Daten, steigen die Möglichkeiten der Identifizierung und machen aufwendige Vorkehrungen nötig. «Das Risiko einer Identifizierung der Patienten ist minimal», sagt Baudis, «aber ganz verhindern lässt sich dies nicht.» Nötig seien deshalb Rahmenbedingungen, die allfällige Missbräuche einschränken können.

Die Herausforderungen beim Datenschutz schmälern sein Engagement für Open Access jedenfalls nicht. Dass er von freien Daten überzeugt ist, entspringt letztlich der gleichen liberalen Grundhaltung, mit der er schon als Gymnasiast mit dem DDR-System in Konflikt geriet. Diese Konstante zeigte sich kürzlich auch in seiner Stellungnahme gegen den Einreisestopp für Bürger aus muslimischen Ländern, die Donald Trump Anfang Jahr verhängt hatte. «Gegen unbegründete Einschränkungen der persönlichen Freiheit bin ich aufgrund meiner Erfahrungen allergisch», sagte Baudis. Man könnte in diesem Zitat «persönliche Freiheit» auch mit «Forschungsdaten» ersetzen und wüsste, was den Forscher antreibt.



**Michael Baudis**

Der Professor für Bioinformatik erforscht genomische Veränderungen in Krebszellen. Sein besonderes Interesse gilt dem Zusammenhang von wiederkehrenden Mutationsmustern mit den klinischen und biologischen Erscheinungsformen der Krankheiten. Ein weiterer Schwerpunkt seiner Arbeit ist die Frage, wie man die dafür benötigten Daten insgesamt besser für die Forschung zugänglich machen kann.

Kontakt: michael.baudis@imls.uzh.ch

## «Globales Hirn»

Die Digitalisierung beschleunigt die Wissenschaft und verändert die Forschung. Abraham Bernstein und Michael Hengartner über die digitale Revolution und die Rolle der UZH. Von Roger Nickl und Thomas Gull

*Abraham Bernstein, Michael Hengartner, die Digitalisierung hat im Eiltempo in unserem Alltag Einzug gehalten. Wie revolutionieren Computer, Internet und soziale Medien unser Leben heute und in Zukunft?*

Abraham Bernstein: Die Digitalisierung ermöglicht, dass wir zu jeder Zeit an jedem Ort mit jedem und jeder kommunizieren können. Das führt zu einer räumlichen und zeitlichen Entgrenzung. Hinzu kommt, dass es immer billiger wird, Daten zu sammeln und zu verarbeiten. Wir können heute Informationen auf völlig neue Arten speichern und analysieren. So entstehen ganz neue Möglichkeiten: Systeme der künstlichen Intelligenz beispielsweise können uns helfen, Entscheidungen zu treffen oder Forscherteams können rund um die Uhr auf dem ganzen Erdball verteilt arbeiten. Wir werden ganz neue Arten entdecken, wie wir uns koordinieren, wie wir zusammenarbeiten und wie wir kommunizieren.

*Was heisst das konkret?*

Bernstein: Die Demokratie zum Beispiel wurde zu einer Zeit erfunden, als das schnellste Mittel zur Informationsübermittlung das Pferd war. Das ist ein Grund dafür, dass wir immer noch auf Papier abstimmen. Heute wäre es dagegen technisch möglich, dass eine Million Menschen in eine Nationalratsdebatte eingreifen und sich mittels sozialer Medien etwa zu einer Gesetzesvorlage äussern könnten. Das könnte den demokratischen Prozess völlig verändern.

Michael Hengartner: Auch das tägliche Leben verändert sich massiv. Das sehe ich bei meinen Töchtern. Sie tauschen sich viel intensiver mit ihren Freundinnen und Freunden aus, die Kommunikation findet aber ausschliesslich im virtuellen Raum statt. In meiner Jugend war ich oft bei Freunden zu Besuch oder sie waren bei mir. Das geschieht heute bei meinen Töchtern viel weniger.

*Was bedeutet die digitale Revolution für die Wissenschaft?*

Bernstein: Die Verfügbarkeit von immer mehr Daten wird zur Folge haben, dass Forschende immer häufiger Fragen stellen, die sie mit bereits existierenden Datensätzen zu beantworten versuchen. Wissenschaft wird auch zunehmend von Maschinen gemacht werden. In der Hefe-Forschung gibt es Geräte, die selber Hypothesen aufstellen und anschliessend ein Experiment dazu durchführen. Auch in der Statistik existieren Programme, die Daten auswerten und neue Theorien entwerfen, die sie dann gleich überprüfen. Weil die Kommunikation einfacher wird, werde auch die Bürgerwissenschaft, die Citizen Science, wichtiger: Wissenschaftler können das Wissen und die Daten von Laien aus der ganzen Welt in ihre Forschung einbeziehen. Um die Citizen Science zu fördern und weiterzuentwickeln, gründet die UZH gemeinsam mit der ETH Zürich ein Citizen Science Center.

*Was passiert mit den Wissenschaftlern, wenn die Maschine alles macht – werden sie bald überflüssig?*

Hengartner: Nein, aber sie müssen sich weiterbilden. Die datengetriebene Forschung muss auch in der Ausbildung und im Studium eine wichtigere Rolle spielen. Heute sollten Geisteswissenschaftlerinnen und -wissenschaftler wissen, wie man mit grossen Datensätzen umgeht. Auch in der Biologie wird zunehmend quantitativ gearbeitet. Damit entstehen neue Arten des Analysierens. Das ist natürlich total spannend und beschleunigt die Forschung.

*Sie, Herr Bernstein, haben die Vision eines globalen Hirns, eines «Global Brain», das die Digitalisierung ermöglicht. Was steckt hinter diesem Gedanken?*