

Computerjagd auf das **AIDSVIRUS**

Resistenzen sind das gravierendste Problem bei so wandlungsfähigen Viren wie dem Aids-erreger HIV. Wie kann man den Infizierten helfen? Die Antwort erfordert umfassende Informationen über Varianten von HIV – und Computerhilfe.

Von Thomas Lengauer und Rolf Kaiser

Als Virus ist der Aids-erreger streng genommen kein Lebewesen, sondern lediglich ein Stück verpacktes Erbgut – knapp 10 000 genomische Buchstaben, umgeben von einer Proteinhülle. Diese 10 000 aneinandergereihten Lettern definieren den vielleicht gefährlichsten Erreger überhaupt. Wie alle Viren nutzt er die von ihm befallene Zelle, um sich zu vermehren. Doch das Humane Immunschwäche-Virus HIV ist besonders tückisch, denn es versklavt dafür just Zellen des menschlichen Immunsystems, die der Krankheitsabwehr dienen.

Derzeit sind über zwei Dutzend Medikamente gegen HIV auf dem Markt, und fast alle zielen darauf ab, eines der viruseigenen Proteine – das Zielprotein des jeweiligen Arzneistoffs – unschädlich zu machen. Es gibt mehrere Möglichkeiten, das anvisierte Eiweißmolekül zu blockieren. Eine beliebte Methode ist, den Wirkstoff so zu konzipieren, dass er sich an die aktive Stelle am Zielprotein heftet – an diejenige, welche die eigentliche Arbeit verrichtet. Das ist etwa so, als würde man den Kopf einer Zange mit Knete verkleben.

Doch warum braucht man so viele verschiedene Wirkstoffe? Warum reicht nicht ein einzelner aus, der ein Protein ausschaltet und so den Lebenszyklus von HIV unterbricht? Der Grund ist, dass sich das Virus rasant – praktisch bei jedem Kopiervorgang – wandelt. Sein Genom, das die Bauinformation der viralen Eiweißstoffe trägt, verändert sich ständig durch Mutationen – vor allem durch Austausch einzelner Basen, quasi der Buchstaben im genomischen Text. Damit ändern sich auch oft Aufbau und Gestalt der von ihm kodierten viralen Proteine, und plötzlich passt der Wirkstoff nicht mehr auf das Zielprotein: Das Virus ist resistent geworden.

Da dies sehr schnell geschieht, gibt man am besten gleich mehrere Wirkstoffe, die sich an verschiedene Proteine heften – in der Hoffnung, dass wenigstens eine der Zielstrukturen sich nicht allzu rasch verändert und der darauf zugeschnittene Arzneistoff eine Weile wirksam bleibt. Darum sind heute Kombinationen von drei bis sechs Wirkstoffen verbreitet, die der Patient gleichzeitig erhält.

Aber am Ende gewinnt das Virus immer. Es werden so viele virale Varianten erzeugt, dass über kurz oder lang, meistens im Verlauf von einigen Monaten, eine neue Variante entsteht und dann dominiert, die gegen alle bisher verabreichten Mittel resistent ist. Damit sind wir an dem Punkt, der uns Bioinformatiker interessiert: Welche Wirkstoffkombination kann nun weiterhelfen?

Tabellen für die Therapie

Die Antwort hängt natürlich von der im Körper des Patienten herausgebildeten Virusvariante ab – genauer: vom Aufbau der in ihrem Genom vorprogrammierten Proteine. Darum wird dem Patienten Blut entnommen, das Virusgenom daraus extrahiert und die genaue Abfolge der genetischen Buchstaben bestimmt. Diese Sequenz – aus der sich wiederum Art und Reihenfolge der Aminosäuren in den entsprechenden Proteinen herauslesen lässt – ist die Grundlage für die Auswahl einer neuen Therapie.

Das in der klinischen Praxis erworbene vielfältige Wissen über die HIV-Resistenz hat man bisher in so genannten Mutationstabellen gesammelt; sie geben an, welche Mutationen im Virusgenom nach klinischen Erfahrungen von Experten mit Resistenzen gegen einzelne Wirkstoffe einhergehen. Die gebräuchlichste Tabelle wird von der International Aids Society fortgeschrieben. Die herkömmliche Therapieauswahl geschieht folgendermaßen. Das aus

In Kürze

- ▶ Das **Aidsvirus mutiert fortwährend** und wird daher über kurz oder lang gegen eine Wirkstoffkombination, die ein Patient erhält, resistent.
- ▶ Mit Hilfe von Datenbanken und mit **bioinformatischen Methoden wie statistischen Lernverfahren** ist es nun möglich, aus der Genomsequenz der jeweiligen Virusvariante die wahrscheinlich effizientesten alternativen Wirkstoffe zu ermitteln.
- ▶ Das Verfahren lässt sich im Prinzip beispielsweise auch **auf Krebszellen anwenden**, um den Zustand eines Tumors besser zu beurteilen.

Eine einmal gewählte Kombination von Wirkstoffen unterdrückt die individuelle Erregervariante eines HIV-Infizierten immer nur begrenzte Zeit. Wie lässt sich möglichst rasch und treffsicher eine andere wirksame Zusammenstellung ermitteln? Bei solchen und ähnlichen Problemen kann die klinische Bioinformatik Lösungen bieten.



ISTOCKPHOTO / STEFAN KLEIN

dem Patientenblut isolierte Virusgenom wird auf Resistenzmutationen analysiert, die in der Mutationstabelle vorkommen. Der behandelnde Arzt verordnet dann dem Patienten eine Kombination solcher Wirkstoffe, gegen die das Virus keine der gelisteten Mutationen aufweist, wobei die genaue Zusammenstellung viele weitere Einflussfaktoren und Nebenbedingungen berücksichtigt.

Eine Auswahl auf dieser Basis hat jedoch entscheidende Nachteile:

- Es gibt Hunderte von sinnvollen Wirkstoffkombinationen, und diese Zahl wächst sprunghaft mit jedem Aidsmedikament, das neu auf den Markt kommt. Dem Arzt fällt es schwer, bei dieser Vielfalt die Übersicht zu behalten.

- Das Resistenzverhalten des Virus lässt sich im Allgemeinen nicht auf einzelne Mutationen zurückführen. Mutationen können interagieren, dabei eine Resistenz verstärken oder umgekehrt abmildern oder gar aufheben. Zudem wird das Virus bei Verabreichung einer neuen Therapie bildlich gesprochen einen evolutionären Fluchtweg in die Resistenz suchen – durch Anhäufung neuer Mutationen. Dieser Weg ist ohne komplexe Analyse praktisch nicht vorherzusagen.

Seit zehn Jahren werden im klinischen Feld systematisch Daten über den Zusammenhang zwischen Virusgenom und Resistenz gesammelt. Dabei geht es um zweierlei.

Zum einen werden in der klinischen Praxis der entwickelten Länder nahezu flächendeckend so genannte genotypische Daten katalogisiert. Sie umfassen wesentliche Sequenzabschnitte des Genoms der im Patientenblut gefundenen HIV-Variante, ergänzt um klinische Parameter, etwa die Zahl der Viruspartikel und der kritischen Immunzellen im Blut des Patienten. Aus den Genabschnitten geht zugleich der biochemische Aufbau der entsprechenden Proteinabschnitte hervor.

Zum anderen sammeln spezielle Forschungsprojekte in geringerem Umfang Daten über die Resistenz des jeweiligen Virus gegen individuelle Wirkstoffe, also über den so genannten Resistenz-Phänotyp. Dabei testet man im Prinzip die im Patientenblut gefundenen Erregervarianten in Zellkultur: Man fügt ihre medikamentenrelevante Genomregion in eine standardisierte Variante von HIV ein, setzt dann dieses Virus unterschiedlichen Konzentrationen jedes einzelnen Aidsmedikaments aus und misst seine Vervielfältigungsrate. Ein beträchtlicher Aufwand.

Resistente Viren zeichnen sich im Test dadurch aus, dass sich ihre Vermehrungsrate erst bei höherer Wirkstoffdosis als sonst reduziert. Das Verhältnis zwischen den Wirkstoffdosen, welche die Rate der untersuchten HIV-Variante beziehungsweise diejenige eines nicht resistenten Referenzstamms zu halbieren vermögen, bildet ein quantitatives Maß für die Resistenz der HIV-Variante, den so genannten Resistenzfaktor. Ist er hoch, so widersteht das Virus, ist er niedrig, so wirkt das Medikament – zumindest im Labor. Ein mittelhoher Faktor bedeutet schwache Resistenz.

Das Ziel unserer Forschung ist nun, den in der Mutationstabelle gesammelten Expertenmeinungen eine Resistenzbestimmung entgegenzusetzen, die sich systematisch und direkt aus klinischen Daten ableitet. Wir tun dies mit Computern. So hoffen wir, eine Systematik und Objektivität zu erzielen, die erlaubt, komplexe Zusammenhänge bei der Resistenzentwicklung zu berücksichtigen.

Grundlage für unsere Analysen bilden die oben beschriebenen genotypischen und phänotypischen Daten. Gesammelt wurden derartige Informationen für mittlerweile über 1000 HIV-Varianten in der deutschen Datenbank Arevir (für: Analyse von Resistenzmutationen bei Viren) – unter Mitarbeit des Nati-

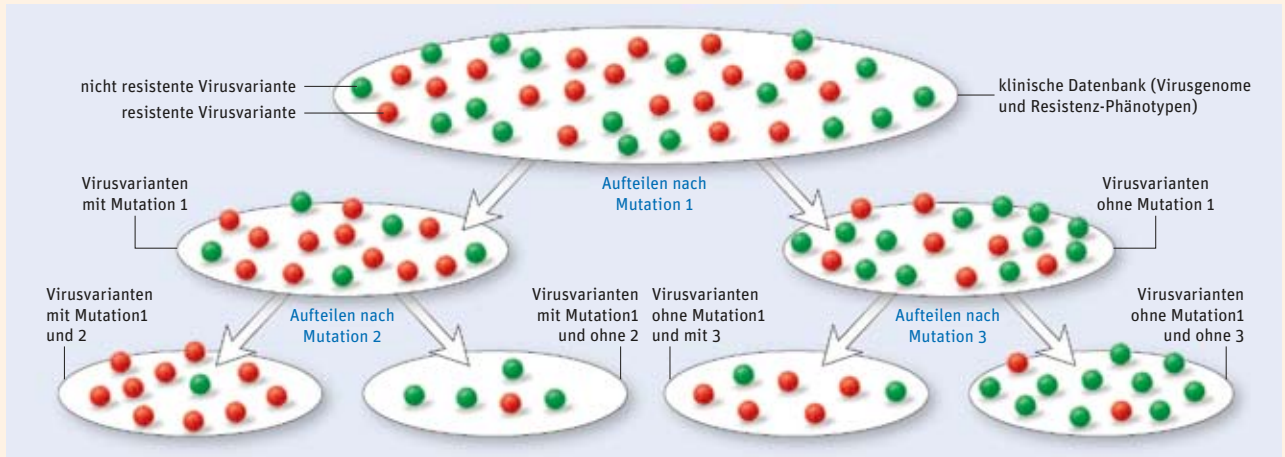
KOMPLEXES ZUSAMMENSPIEL

Im Allgemeinen gibt es nicht nur eine einzige Mutation, die ein Virus gegen ein bestimmtes Medikament resistent macht, sondern viele verschiedene, die sich zudem wechselseitig beeinflussen, verstärken oder auch abschwächen können. Eine so genannte **sekundäre Resistenzmutation** beispielsweise kann den Vermehrungserfolg eines Virus steigern, aber nur im Zusammenhang mit einer bestehenden Resistenz.

WIE KANN DER COMPUTER DIE RESISTENZ DES PATIENTENVIRUS VORHERSAGEN?

Bislang konnten Ärzte sich bei der Auswahl der Medikamente gegen das jeweilige Patientenvirus nur von gewissen typischen Mutationen und ihrer klinischen Erfahrung leiten lassen. Mittels Datenbanken und bioinformatischer Methoden versuchen wir schneller und genauer ein Ergebnis zu erreichen. Dazu wird eine

Abfragefolge an das virale Genom abgeleitet, an deren Ende sich entscheiden lässt, ob eine individuelle Erregervariante gegen ein Medikament resistent ist oder nicht. Die resultierenden Entscheidungsbäume sind hierarchische Strukturen und erlauben es, schwer verständliche Datenbestände zu interpretieren.



Ableiten der Abfragefolge

Ausgangspunkt ist eine Datenbank vieler verschiedener Virusgenomvarianten, die auf Grund von Labortests als resistent (rot) oder nicht resistent (grün) gegen ein Medikament klassifiziert sind. Die Menge wird nun als Erstes so aufgespalten, dass der eine Teil alle Virusvarianten mit einer bestimmten Mutation enthält (gleichgültig, welche Mutationen sie sonst noch tragen), der andere die Varianten ohne diese. Als erste Mutation wählt man eine, welche die resistenten Varianten möglichst stark in einem der beiden »Töpfe« anreichert. Die Aufteilung wird auf der Basis wei-

terer Mutationen fortgesetzt – und zwar so, dass die Trennung von resistenten und nicht resistenten Varianten am Ende möglichst deutlich ausfällt. Im dargestellten, stark verkürzten Schema hieße das: Alle Virusvarianten, welche die Kombination aus Mutation 1 wie auch 2 tragen, sind überwiegend resistent; Ähnliches gilt für Varianten mit Mutation 3, ohne 1. Diese Abfragefolge, dann auf eine neue Erregervariante eines Patientenvirus angewandt, erlaubt dem Computer beispielsweise die Aussage: mit hoher Wahrscheinlichkeit resistent gegen Wirkstoff X (siehe rechte Hälfte des Kastens).

MASCHINELLES LERNEN

Mit diesem Begriff bezeichnen wir eine computerbasierte Methode, die statistische Verfahren auf einen schwer interpretierbaren Datenbestand anwendet – um darin manuell nur schwer zu findende Muster zu erkennen. Im hier geschilderten Fall sind dies die komplexen Zusammenhänge zwischen dem Aufbau des jeweiligen Virusgenoms und seinem im Labor ermittelten so genannten Resistenz-Phänotyp.

onalen Referenzzentrums für Retroviren an der Universität Erlangen-Nürnberg und anderer im Verein Genafor (Gesellschaft für nachhaltige Forschung) organisierten Labors, Praxen und Institute. Diesen Fundus, zusammen mit der Sammlung klinischer Daten, haben wir jüngst in Euresist, eine europäische Datenbank mit weiteren 5400 Datensätzen, eingebracht.

Mit Hilfe des Datenbankmaterials und mit bioinformatischen Methoden aus dem Bereich des so genannten maschinellen oder statistischen Lernens (siehe linke Randspalte) erstellen wir statistische Modelle, welche die erwähnte Mutationstabelle ersetzen sollen. Im Gegensatz zu ihr sind unsere Modelle in der Lage, Abhängigkeiten und Wechselwirkungen zwischen verschiedenen Mutationen darzustellen. Ein bekanntes Prinzip zur Mustererkennung in schwer verständlichen Datenbestän-

den ist der Entscheidungsbaum (siehe Kasten oben). Aus den phänotypischen Daten in Arevir entwickeln wir algorithmisch möglichst genaue Entscheidungsbäume. Ein so genannter Kreuzvalidierungstest zeigte, dass unsere Bäume in etwa 85 Prozent der Fälle korrekte Aussagen über die Resistenz eines Virus liefern – allerdings über die in einer Zellkultur im Labor gemessene Resistenz.

Flucht in die Resistenz

Doch in der klinischen Praxis werden nicht einzelne Medikamente verabreicht, sondern Wirkstoffkombinationen. Außerdem agiert das Virus nicht in einer Zellkultur, sondern im Körper des Patienten. Erst wenn die Resistenzanalyse diese beiden Aspekte berücksichtigt, kann sie als klinisch relevant gelten. Es hieß also, die Bioinformatik-Methoden in dieser Hinsicht weiterzuentwickeln.

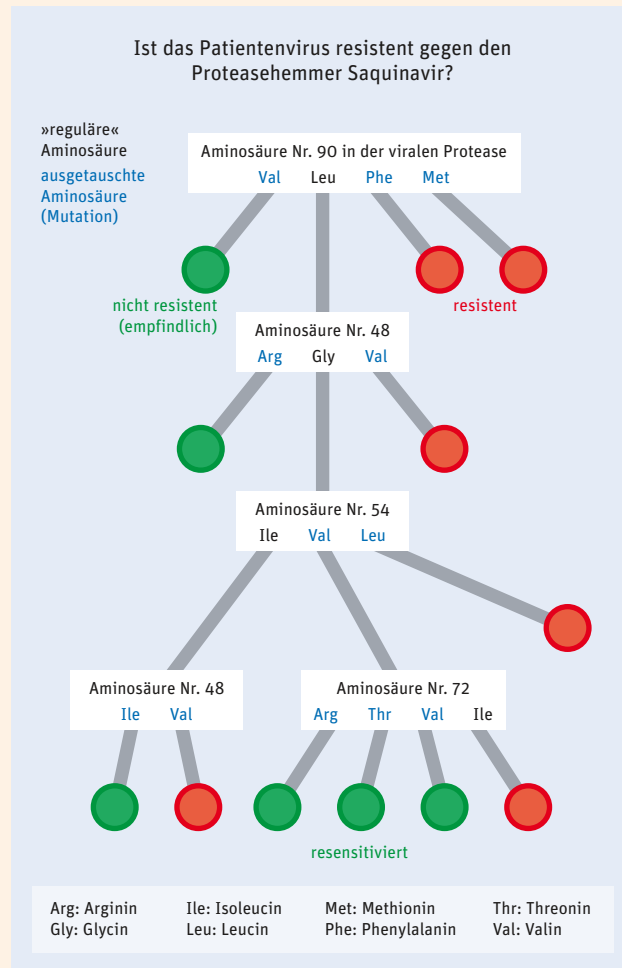
Resistent gegen Saquinavir?

Wie komplex das reale Zusammenspiel von Mutationen sein kann, zeigt dieser Entscheidungsbaum, nach dem bestimmt wird, ob das untersuchte Patientenvirus gegen den Proteasehemmer Saquinavir resistent ist. Die Zahlen entsprechen Positionen in der Aminosäuresequenz des Zielproteins, hier der viralen Protease. Der Name der »normalen« Aminosäure an der jeweiligen Stelle im Protein des Referenzvirus ist schwarz dargestellt, blau der einer durch Mutation ausgetauschten Aminosäure. Die Farbkreise an den einzelnen Enden der Entscheidungshierarchie repräsentieren die endgültige Entscheidung über die Resistenz einer Virusvariante.

Ist an Sequenzposition 90 der Protease statt des ursprünglichen Leucins die Aminosäure Valin zu finden, dann spricht das Patientenvirus dennoch auf Saquinavir an. Steht aber dort durch Mutation ein Phenylalanin oder Methionin, so ist es resistent. Bleibt die Stelle genauso wie beim Referenzvirus, müssen wir weiter an Position 48 schauen und so fort.

Wie statistische Modelle die Abhängigkeiten zwischen Mutationen berücksichtigen, zeigt sich beispielhaft bei den Entscheidungswegen, die über Position 54 und 72 führen. Steht an der 54 mutationsbedingt ein Valin, dann ist das Patientenvirus nur resistent, wenn die Position 72 bei ihm die normale Aminosäure enthält. Ist diese letzte Position beispielsweise ebenfalls zu Valin mutiert, ist es dagegen empfindlich. Dieser Austausch bewirkt also eine Resensitivierung: Das Medikament wird wieder wirksam. Insgesamt können drei verschiedene Mutationen an Position 72 die Resistenz eines Virus aufheben, die sonst durch die Mutation bei 48 hervorgerufen würde.

Neben ihrer klinischen Bedeutung für die Resistenzvorhersage geben die Entscheidungsbäume den Forschern im Labor Hinweise über mögliche Mechanismen der Resistenzbildung und können somit Grundlage weiterer Experimente zur Aufklärung dieser Mechanismen darstellen.

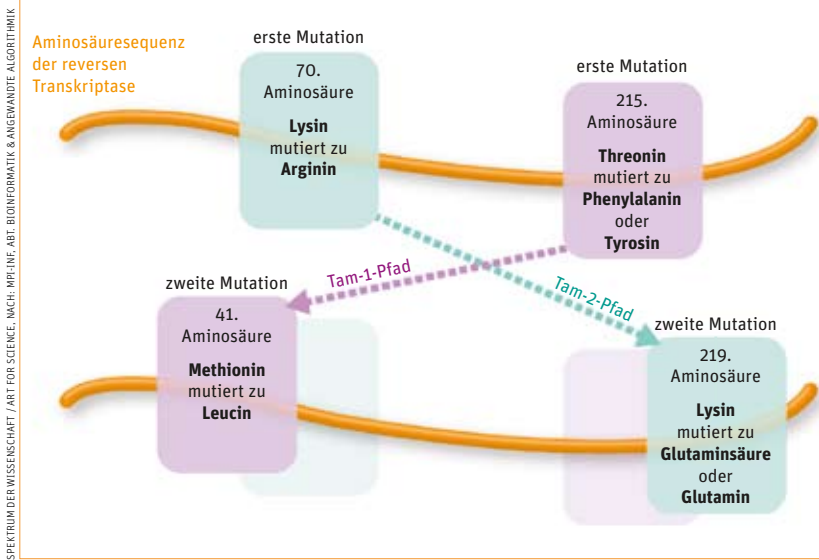


In Kombinationstherapien wirken Medikamente auf komplexe Weise zusammen. Um das zu modellieren, benutzen wir wiederum statistische Lernverfahren, die subtile nichtlineare Abhängigkeiten zwischen den verschiedenen Wirkstoffen erkennen. Dabei betrachten wir nicht nur die Resistenz des gegenwärtig beobachteten Virusstamms, sondern wir extrapolieren auch in die Zukunft. Denn für den Erfolg einer Therapie ist nicht nur wichtig, wie resistent der gegenwärtige Stamm gegen die verabreichten Mittel ist, sondern auch, nach welcher Zeit ein kritisches Resistenzniveau gegen eine neue Kombination erreicht sein dürfte.

Mit jeder neuen Therapie mutiert das Virus schließlich zur Resistenz. Es tut dies aber nicht beliebig, sondern folgt erfahrungsgemäß vorzugsweise so genannten Mutationspfaden. Gäbe unsere Datenbank Aufschluss über die

langfristigen Verläufe vieler Patienten, so könnten wir die Flucht in die Resistenz am einzelnen Patienten beobachten. Durch statistische Mittelung wären dann die am häufigsten verfolgten Fluchtpfade zu ermitteln. Leider besitzen wir solche Daten nicht in ausreichendem Maß, sondern nur einzelne oder wenige Messungen für jeden Patienten. Trotzdem können wir die Wegverläufe statistisch ermitteln.

Ein Beispiel bietet sich im Zusammenhang mit dem am längsten verfügbaren HIV-Medikament Azidothymidin (AZT), das sich gegen ein virales Enzym namens reverse Transkriptase richtet. Im Verlauf einer Therapie damit kann das Virus in unterschiedlicher Weise Resistenz dagegen ausbilden. Im Wesentlichen beschreitet es zwei verschiedene Wege, die zielgerichtet erscheinen, weil sich die dafür nötigen Mutationen gewöhnlich nicht in zu-



HIV nutzt bevorzugte »Fluchtwege« in die Resistenz, bei denen bestimmte Mutationen aus ungeklärten Gründen fast immer in bestimmter Reihenfolge auftreten. Bei der Resistenzentwicklung gegen den Wirkstoff Azidothymidin, der die reverse Transkriptase des Virus hemmen soll, sind das der so genannte TAM-1- und der TAM-2-Pfad. Pfade wie diese lassen sich mit geeigneten statistischen Methoden aus der Datenbank extrahieren.

fälliger, sondern in fester Reihenfolge ansammeln (über die Ursache kann man bisher nur spekulieren). Fachlich werden diese beiden Wege als TAM-1- und TAM-2-Pfad bezeichnet (das Kürzel steht für Thymidinanalogmutation). Der erste führt über eine Mutation an Position 215 der reversen Transkriptase zu einer Mutation an Position 41 (siehe Grafik oben). Dieser Pfad zeichnet sich nun folgendermaßen in unserer Datenbank ab: Es kommen Viren ohne Mutationen vor, solche mit Mutationen an Position 215 im Protein, aber ohne den Austausch an Position 41, sowie solche mit Mutationen an beiden Positionen. Hingegen sehen wir keine oder wenige Viren mit der Veränderung an der Position 41 ohne gleichzeitige Mutationen an der Position 215.

Gelungene Bewährungsprobe

Die Information über bevorzugte Fluchtpfade ist in unserer Datenbank somit implizit vorhanden, und mit entsprechenden statistischen Methoden lassen sich die Wege extrahieren. Damit haben wir ein statistisches Maß für die so genannte genetische Barriere entwickelt – das heißt für die Wahrscheinlichkeit, dass das Virus nach einer gewissen Zeit diese Hürde überspringen, also gegen das betrachtete Medikament resistent wird. Hat die Erregervariante beispielsweise den ersten Schritt auf einem erkannten Vorzugsweg gemacht, wird das Mittel wahrscheinlich nur kurzfristig greifen. Die durch unsere Modelle vorhergesagten Resistenzen gehen in eine statistische Lernmethode ein, die letztlich über die geschätzte Effektivität einer jeden möglichen Kombinationstherapie Auskunft gibt. Das Berechnen der Reihung dauert bloß einige Sekunden.

Natürlich kann die computergestützte Analyse nur Vorschläge machen. Viele wich-

tige Elemente der HIV-Infektion bleiben unberücksichtigt. Der wohl wichtigste Aspekt, der dabei fehlt, ist das Genom des Patienten, also die individuelle Variante der Wirtszelle.

Bei unvollständigen Modellen, wie wir sie hier vorlegen, ist die Überprüfung ihrer Aussagekraft, die so genannte Validierung, ein wichtiger Teil der Forschung. Wir wenden dafür eine Reihe von statistischen Tests an, die nachweisen, dass unsere Aussagen statistisch signifikant sind. So liegt etwa die Fehlerrate bei der Bewertung des Erfolgs von Kombinationstherapien bei 15 Prozent. Das ist immerhin wesentlich besser als die 25 Prozent, die man in der klinischen Praxis sonst erzielt. Damit hat unsere Methode ihre Relevanz unzweifelhaft unter Beweis gestellt.

Noch eindrucksvoller ist vielleicht folgende Geschichte aus der Praxis: Einer der mit uns zusammenarbeitenden Ärzte hat seit 1987 einen Aidspatienten in Behandlung, nennen wir ihn Georg. Aber es war keine Therapie zu finden, die Georg wirklich half, das heißt die Virusvermehrung erfolgreich unterdrückte. 15 Jahre lang wurden bei allen verabreichten Therapien hohe Zahlen von HI-Viren im Blut gemessen. Dabei durchlief Georg mehrmals den Zyklus von Resistenzentwicklung und Therapiewechsel. Im Oktober 2003 boten die Mutationstabellen schließlich keine Alternative mehr. Der im Blut von Georg isolierte Virenstamm enthielt Resistenzmutationen gegen alle Wirkstoffe.

Nach der klassischen Vorgehensweise war Georg »austherapiert«. In dieser Situation wurde unser Server zu Rate gezogen. Der machte einen Vorschlag, den der Arzt auf Grund seiner medizinischen Expertise noch etwas modifizierte. Seit diesem Therapiewechsel konnte bei Georg zum ersten Mal in seiner Patientenlaufbahn die Erregervermehrung so erfolgreich unterdrückt werden, dass im Blut keine Viruspartikel (freies Virus) mehr nachzuweisen waren. (HIV lässt sich freilich nicht ganz aus dem Körper entfernen, da es in seiner DNA-Form im Erbgut befallener Zellen zu ruhen und sich so in verschiedenen Organen längere Zeit zu verstecken vermag.) Dieser Zustand hält jetzt schon über fünf Jahre an, und Georg kann seinen Beruf ausüben.

Aber auch bei bester Verabreichung von Anschlusstherapien ist eine Anreicherung von Viren mit vielfältigen Resistenzmutationen gegen die heute gängigen Medikamente zu erwarten – sowohl im einzelnen Patienten als auch langfristig in der Bevölkerung. Daher müssen fortwährend neue Medikamente entwickelt werden, die den Lebenszyklus des Virus an immer anderen Stellen unterbrechen. Besonders interessant sind Wirkstoffe, die

den Eintritt des Virus in die Wirtszelle blockieren. Er ist ein komplexer Vorgang, bei dem virale Hüllproteine und Proteine an der Oberfläche der Wirtszelle zusammenwirken. Daher versucht man mit Wirkstoffen auf diese zelleigenen, also auf menschliche Proteine zu zielen.

Zum Eintritt in die Wirtszelle benötigt HIV das Zellprotein CD4 sowie einen weiteren zellulären Rezeptor, der in diesem Zusammenhang Korezeptor genannt wird. Je nach Virusstamm, also genetisch vorgegeben, nutzt HIV als Korezeptor entweder nur das Protein CCR5 oder nur das Protein CXCR4. Dabei gibt es auch Virusvarianten, die jedes der beiden Moleküle missbrauchen können (die beiden vermitteln der Zelle üblicherweise Signale von speziellen Botenstoffen, den Chemokinen).

Anwendung bei Krebs

Neuerdings sind Medikamente entwickelt worden, die CCR5 blockieren. Einer dieser so genannten Korezeptorblocker ist seit dem Herbst 2007 auf dem Markt. Doch auch hier steht dem Virus ein Weg in die Resistenz offen: Sein Hüllprotein, das an CD4 und den Korezeptor andockt, verändert sich in einer Weise, dass es entweder sich in etwas anderer Form an CCR5 heftet oder stattdessen CXCR4 als Korezeptor vorzieht. Vor allem bei fortgeschrittenem Krankheitsverlauf treten derartige X4-Viren im Patienten auf. In diesen Fällen bliebe eine Therapie mit den CCR5-Korezeptorblockern wirkungslos. Daher ist vor ihrer geplanten Verabreichung ein Test auf X4-Viren vorgeschrieben. Wir haben eine bioinformatische Variante dieses Tests entwickelt, die auf der Sequenz des Virusgenoms basiert und inzwischen vielfach verwendet wird. All unsere Verfahren sind im Internet unter www.geno2pheno.de frei zugänglich und werden bei der Krankenversorgung bereits genutzt.

Unsere statistischen Modelle beruhen allerdings auf einer wesentlichen Annahme, die das Problem unzulässig vereinfacht. Wir sind bisher davon ausgegangen, dass im Patienten eine einzige virale Variante vorherrscht. Dem ist aber nicht so. Im Patienten werden außerordentlich viele Viruspartikel gebildet und wieder vernichtet. Zu Beginn der Infektion sind es Millionen bis Milliarden täglich – und die sind wegen der häufigen »Tippfehler« beim Umkopieren des viralen RNA-Genoms in DNA nicht alle gleich. Vielmehr bilden sie eine heterogene Population, eine so genannte Quasispezies.

Im Patienten verbirgt sich daher ein Gemisch aus nicht resistenten und resistenten Viren; manche benutzen CCR5, andere CXCR4.

Traditionelle Methoden erfassen immer nur das Erbgut der dominierenden – häufig vorkommenden – Varianten. Manchmal spielen aber auch selten auftretende Formen eine Rolle. Falls diese Minoritäten hoch resistent sind, können gerade sie sich mit der Zeit durchsetzen und eine Therapie wirkungslos machen.

Seit kurzer Zeit gibt es neue Messmethoden, die praktisch die gesamte Quasispezies – also alle im Patienten vorhandenen genetischen Varianten – zu erfassen vermögen. So haben wir in einer Untersuchung zum Korezeptorgebrauch bei einem Patienten den relevanten Genomabschnitt von über 12 500 Viren erfasst und dabei über 1000 Varianten vorgefunden. Wir entwickeln nun unsere statistischen Methoden so weiter, dass sie die gesamte Quasispezies berücksichtigen und nicht nur den dominanten Virenstamm. Davon erwarten wir uns deutlich genauere Vorhersagen. Obwohl die Verfahren wichtige Faktoren noch nicht berücksichtigen – zum Beispiel das Immunsystem des jeweiligen Patienten –, verbessern sie bereits jetzt die klinische Versorgung der Betroffenen.

Die hier beschriebenen Methoden lassen sich im Prinzip überall dort einsetzen, wo es klinisch relevante evolutionäre Prozesse gibt sowie Daten darüber, wie diese Prozesse ablaufen. Konkret möchten wir in den nächsten Jahren unsere Methoden auf die Krankheiten Hepatitis C und Hepatitis B anwenden. Aber auch bei Krebs gibt es Einsatzmöglichkeiten.

Der Verlauf von Krebs ist oftmals mit einem fortschreitenden Zerfall des Genoms der Tumorzelle verbunden: Teile von Chromosomen gehen verloren oder werden vervielfacht, ganze Chromosomen verschwinden oder erscheinen in Überzahl. Dadurch verwandelt sich eine gesunde Zelle, die gemeinsam mit anderen im jeweiligen Organ ihre vielfältigen Aufgaben erfüllt, in eine Krebszelle, die allein auf Vermehrung und Metastasierung spezialisiert ist.

Diese Veränderung des Zellgenoms während des Tumorwachstums entspricht im Prinzip der Flucht des HI-Virus in die Resistenz. Mit unseren Methoden können wir anhand des vorliegenden degenerierten Zellgenoms den Zustand des Tumors genauer erfassen, als es bisher möglich war. Das bietet zwar noch keinen Ansatz zu einer neuen Therapie, aber der Arzt kann nun besser entscheiden, ob aggressive Therapieansätze noch fruchten oder ob man lieber palliativ – schmerzlindernd – vorgehen sollte. Da Krebs vor allem in späteren Stadien vielfach übertherapiert wird, ist dies ein nützlicher Beitrag zum besseren Ressourceneinsatz und zum Schutz des Patienten vor unnötigem Leiden.



Thomas Lengauer (links) promovierte 1976 an der Freien Universität Berlin in Mathematik sowie 1979 an der Stanford University (US-Bundesstaat Kalifornien) in Informatik. Von 1979 bis 1981 war er bei den Bell Laboratories in Murray Hill (New Jersey) beschäftigt. Er ist seit 2001 Direktor der Arbeitsgruppe »Computational Biology and Applied Algorithmics« des Max-Planck-Instituts für Informatik in Saarbrücken.
Rolf Kaiser promovierte 1990 in Bonn in Biologie und gehört seit 1999 dem Institut für Virologie der Universität Köln sowie dem Vorstand der Gesellschaft für nachhaltige Forschung (Genafor) an.

Beerenwinkel, N. et al.: Computational Methods for the Design of Effective Therapies against Drug Resistant HIV Strains. In: *Bioinformatics* 21(21), S. 3943–3950, 2005.

Lengauer, T., Sing, T.: Bioinformatics-Assisted Anti-HIV Therapy. In: *Nature Reviews Microbiology* 4(10), S. 790–797, 2006.

Lengauer, T. et al.: Bioinformatics Prediction of HIV Coreceptor Usage. In: *Nature Biotechnology* 25(12), S. 1407–1410, 2007.

Rahmenführer, J. et al.: Estimating Cancer Survival and Clinical Outcome Based on Genetic Tumor Progression Scores. In: *Bioinformatics* 21(10), S. 2438–2446, 2005.

Weblinks zu diesem Thema finden Sie unter www.spektrum.de/artikel/999553.