

Das Projekt persönliches Genom

Spätestens in zehn Jahren dürfte es so weit sein: Jeder hätte die Möglichkeit, den Text seiner persönlichen DNA zu erfahren.

Von George M. Church

Nach seinem Start in den 1990er Jahren entfaltete sich das World Wide Web nur scheinbar fast wie über Nacht. Erfahrungsgemäß benötigen Neuerungen wenigstens ein Jahrzehnt, bis sie marktreif sind und sich durchsetzen. In diesem Fall existierte bereits eine Infrastruktur: Seit Jahren hatte sich das Internet aufgebaut, und plötzlich wurde klar, dass wesentliche benötigte Ressourcen, darunter PCs, in ihrer Zahl eine kritische Schwelle überschritten hatten.

Auch Visionen und Marktkräfte treiben die Entwicklung und Verbreitung neuer Technologien voran. So begann das amerikanische Raumfahrt-Programm als Vision der Regierung. Erst viel später forcierten die militärische und zivile Nutzung von Satelliten eine rentable Weltraumindustrie. Die nächste technologische Revolution könnte die Biotechnologie betreffen. Zunehmend lässt sich ausmalen, welche Märkte, Visionen, Entdeckungen und Erfindungen diese Umwälzung wohl vorantreiben werden

und welche Mindestanforderungen an Infrastruktur und Ressourcen dafür erfüllt sein müssen.

Etwa ein Dutzend Forscher, zu denen ich gehörte, brachte 1984/85 die Idee vor, das Genom des Menschen komplett zu entziffern. Ziel des Vorhabens war, eine menschliche DNA-Sequenz vollständig zu lesen – das menschliche Genom zu sequenzieren, wie wir dazu sagten. Die eigentliche Arbeit begann dann um 1990 und wurde unter dem Namen Humangenomprojekt bekannt. Für das Unterfangen, an dem Forscher vieler Länder teilnahmen, veranschlagten wir 15 Jahre. Im Jahr 2005 sollte ein menschliches Genom entziffert vorliegen. Bereits ein paar Jahre früher kam die Sequenzierung der einfacheren 93 Prozent dieses Erbguts zum Abschluss.

Das Humangenomprojekt sollte drei Milliarden Dollar kosten. In seinem Verlauf kamen so viele technische und methodische Neuerungen zu Stande, dass man heute die DNA eines Menschen für zwanzig Millionen Dollar sauber genug entziffern kann, um aus den Daten Nutzen zu ziehen. Das bedeutet aber: Der

Aufwand ist immer noch so hoch, dass fast nur ausgewiesene Sequenzierungszentren dies leisten können, zudem nur mit der finanziellen Ausstattung großer Forschungsprojekte.

Unter dem Schlagwort »1000-Dollar-Genom« wurde das Versprechen populär, eine Gesamtsequenzierung der DNA von Einzelpersonen erschwinglich zu machen. Der Preis müsste so niedrig sein, dass es vielen Menschen die Sache wert erscheint, einmal im Leben eine persönliche Sequenzierung ihres kompletten Genoms vornehmen zu lassen, sodass man die Daten auf einer Diskette gespeichert zum Arzt mitnehmen kann. Bis es so weit sein wird, muss die Entwicklung der Technologien weiter gewaltig vorangetrieben werden. Doch die Anstrengung lohnt. Durch die Möglichkeit, DNA kostengünstig zu lesen, wird sich nicht nur die Gemeinde der Genomforscher vervielfachen. Auch die Anzahl sequenzierter menschlicher Genome wird stark wachsen. Das wird helfen, genetische Unterschiede zwischen Individuen herauszuarbeiten – zum Nutzen des Einzelnen nicht nur im Krankheitsfall.

Die Vision reicht noch weiter. Eines Tages könnte man unter Humangenomik auch alles Genmaterial der gesamten äußeren und inneren Umwelt des Menschen fassen, widriges und günstiges. So wie wir heute Wetterkarten anschauen, werden wir zukünftig vielleicht täglich checken, welche Mikroben und potenziellen Allergieauslöser sich in und um uns befinden. Zugleich könnten Nanotechnologie und industrielle Biotechnologie mit ihrem großen Entwicklungspotenzial Ökosysteme nach »intelligenten« Materialien und Mikroorganismen absuchen, die sich zur industriellen Produktion oder

IN KÜRZE

- Um die Möglichkeiten der Biotechnologie voll auszuschöpfen, müssten die Verfahren so preiswert und einfach zu handhaben sein wie PCs. Das gilt auch für die **Genom-Sequenzierung für jeden**.
- Das Lesen von DNA wird durch **automatisierte und schnellere Verfahren** immer ökonomischer. Forscher versuchen beispielsweise, Einzelschritte einzusparen, mit immer weniger Material auszukommen und Prozesse millionenfach parallel ablaufen zu lassen.
- Wie sind umfassende persönliche genetische Daten zu verwalten? Wer darf und soll zu ihnen Zugang haben? Das **Projekt persönliches Genom** (*Personal Genome Project*) widmet sich auch diesen wichtigen Aspekten.

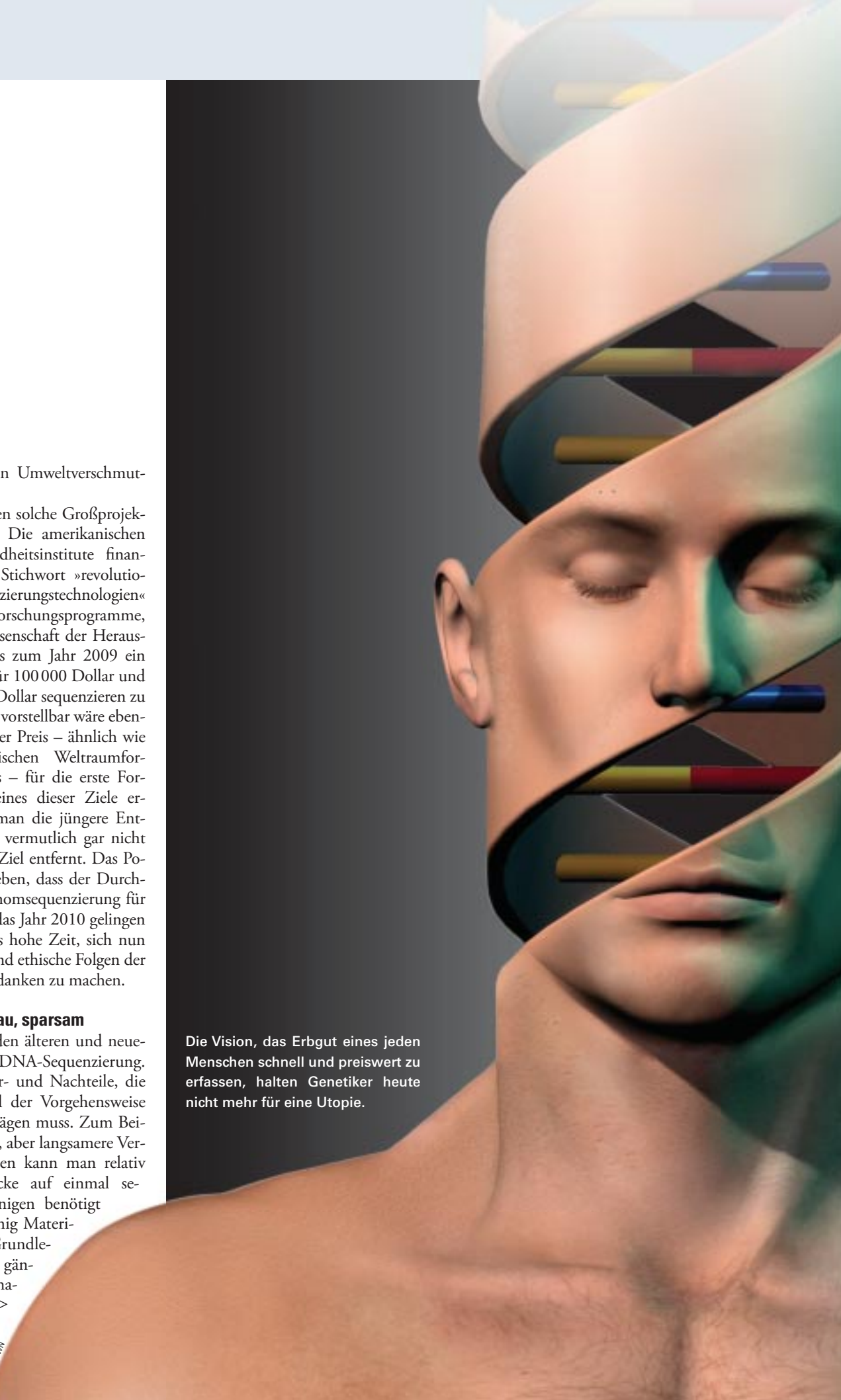
zur Beseitigung von Umweltverschmutzung nutzen lassen.

Allerdings kosten solche Großprojekte hohe Summen. Die amerikanischen Nationalen Gesundheitsinstitute finanzieren unter dem Stichwort »revolutionäre Genomsequenzierungstechnologien« gegenwärtig zwei Forschungsprogramme, wobei sich die Wissenschaft der Herausforderung stellt, bis zum Jahr 2009 ein Menschengenom für 100 000 Dollar und bis 2014 für 1000 Dollar sequenzieren zu können. Als Anreiz vorstellbar wäre ebenso ein hoch dotierter Preis – ähnlich wie in der amerikanischen Weltraumforschung der X-Preis – für die erste Forschergruppe, die eines dieser Ziele erreicht. Betrachtet man die jüngere Entwicklung, sind wir vermutlich gar nicht mehr so weit vom Ziel entfernt. Das Potenzial scheint gegeben, dass der Durchbruch zu einer Genomsequenzierung für 20 000 Dollar um das Jahr 2010 gelingen könnte. Da wird es hohe Zeit, sich nun auch über soziale und ethische Folgen der Errungenschaft Gedanken zu machen.

Ziel: schnell, genau, sparsam

Zunächst aber zu den älteren und neueren Verfahren der DNA-Sequenzierung. Jedes hat seine Vor- und Nachteile, die man bei der Wahl der Vorgehensweise gegeneinander abwägen muss. Zum Beispiel gibt es genaue, aber langsamere Verfahren; bei manchen kann man relativ längere DNA-Stücke auf einmal sequenzieren; bei einigen benötigt man besonders wenig Material und so weiter. Grundlegende Schritte der gängigen Verfahren haben die Forscher ▷

Die Vision, das Erbgut eines jeden Menschen schnell und preiswert zu erfassen, halten Genetiker heute nicht mehr für eine Utopie.



▷ gewissermaßen der Natur abgeschaut. Sie benutzen deren Werkzeugkasten, haben die Mechanismen nur für ihre Zwecke abgewandelt und verfeinert. Das menschliche Genom besteht aus 3 Milliarden Nucleotidpaaren, und jedes Nucleotid enthält eine der vier so genannten Kernbasen Adenin, Cytosin, Guanin oder Thymin, abgekürzt A, C, G und T. Diese vier Basen stellen sozusagen das Alphabet für die kodierte genetische Information dar. Ihre Abfolge galt es im Humangenomprojekt zu erfassen.

Die Genetiker machen sich zu Nutze, dass sich in einem DNA-Doppelstrang, in der DNA-Leiter, jeweils A nur mit T (und umgekehrt) und G nur mit C als Gegenüber paart. Diese beiden Basenpaare bilden quasi die Leitersprossen (siehe Kasten unten). Die strenge Paarungsregel bedeutet zugleich, dass die eine Sei-

te der Erbleiter die komplementäre Sequenz zur anderen Seite aufweist. Wenn man den einen Halbstrang liest, kennt man den anderen.

Die drei Milliarden Basenpaare verteilen sich auf 23 Chromosomen. Von denen besitzt ein Mensch normalerweise zwei Sätze, einen vom Vater, einen von der Mutter, also zusammen 46 Chromosomen. Die beiden Sätze unterscheiden sich allerdings zu 0,01 Prozent. Das bedeutet genau genommen, dass ein individuelles Genom des Menschen nicht drei, sondern sechs Milliarden Basenpaare aufweist.

Wie aber »liest« man eine Basensequenz? Im Prinzip könnte man dazu schlicht einen physikalischen Sensor einsetzen, der die Unterschiede der vier Moleküle noch unter der Nanometerebene erfasst. Mit einem Rastertunnelmikros-

kop zum Beispiel lassen sich die vier Basen zwar unterscheiden. Allerdings wäre das für unsere Zwecke viel zu langwierig. In einigen Schritten zumindest setzen die gebräuchlichen Sequenzierungsmethoden daher lieber chemische Verfahren ein.

Besonders wenige Ablesefehler

Zunächst gilt es bei den meisten üblichen Vorgehensweisen, die DNA-Stücke, die sequenziert werden sollen, beträchtlich zu vervielfachen. Erst die große Zahl nämlich liefert später von den einzelnen Basen ein ausreichend großes Messsignal (siehe Kasten rechts).

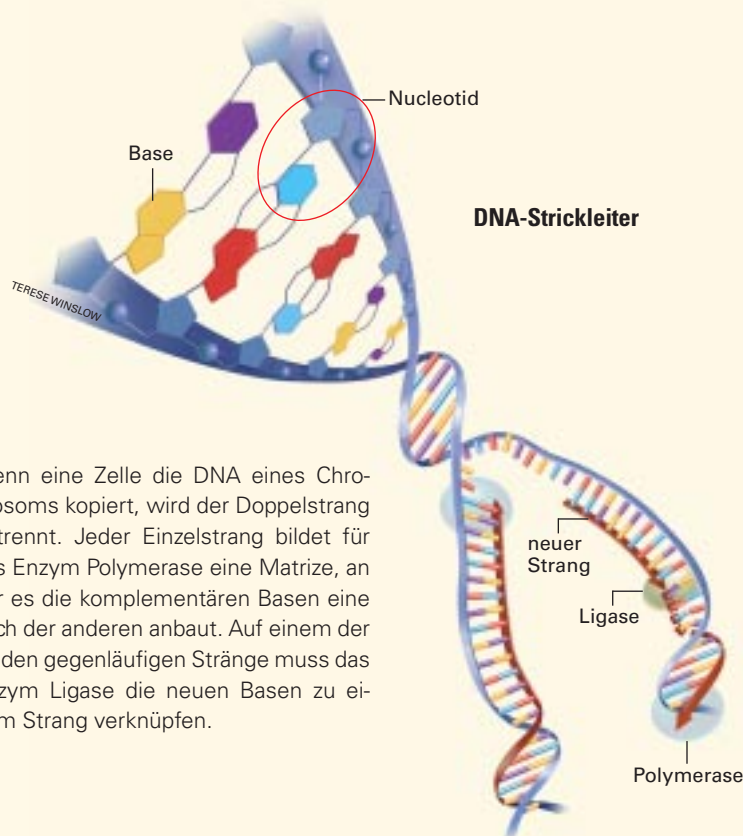
Bis heute eine der gebräuchlichsten Sequenzierungsmethoden, wenn auch mit den Jahren wesentlich weiterentwickelt, war und ist das Prinzip, das der englische Biochemiker Frederick Sanger in den 1970er Jahren erdachte. Er erhielt dafür 1980 den Nobelpreis für Chemie – seinen zweiten! (Den ersten bekam er 1958 für die Entschlüsselung des Insulins, die erste Entschlüsselung eines Proteins.) Das Sanger-Verfahren in seinen neueren Versionen bildete auch das Gerüst für das Humangenomprojekt. Im Labor sagen manche dazu »Sequenzierung durch Auftrennung«. Denn dabei werden viele unterschiedlich kurze Sequenzen erzeugt, die am Ende mit einem in Laserlicht fluoreszierenden Erkennungsmolekül bestückt sind. Diese Sequenzen werden dann nach ihrer Länge in einem Gel elektrophoretisch voneinander getrennt (Einzelheiten siehe Kasten rechts, unten). Am Ende liest ein Detektor die Fluoreszenzsignale in der Reihenfolge, in der sie ihn passieren. Weil diese Reihenfolge der Länge der erzeugten Sequenzen entspricht, gibt sie die Abfolge der Basen im ursprünglichen DNA-Strang wieder.

Das Sanger-Verfahren liefert zwar recht genaue Daten, das heißt besonders wenige Ablesefehler, doch auch die neueren Versionen bleiben immer noch recht zeitintensiv und teuer. Deswegen suchen Forscher nach alternativen Methoden, bei denen insbesondere die Zeitfordernde Prozedur der elektrophoretischen Auftrennung entfällt. Sie versuchen außerdem durch wesentlich verkleinerte Reaktionsansätze den Materialverbrauch zu minimieren – Stichwort Miniaturisierung. Und sie setzen auf Rationierung durch paralleles Erfassen von Millionen DNA-Fragmenten gleichzeitig im selben Durchgang.

Wie Zellen ihre DNA kopieren

Viele Verfahren nutzen die Regel, dass sich die vier Basen der Nucleotide – der Bausteine im Erbstrang – nur komplementär paaren: Adenin (A) nur mit Thymin (T), und Cytosin (C) nur mit Guanin (G). Diese Paare bilden gewissermaßen die Sprossen der DNA-Leiter.

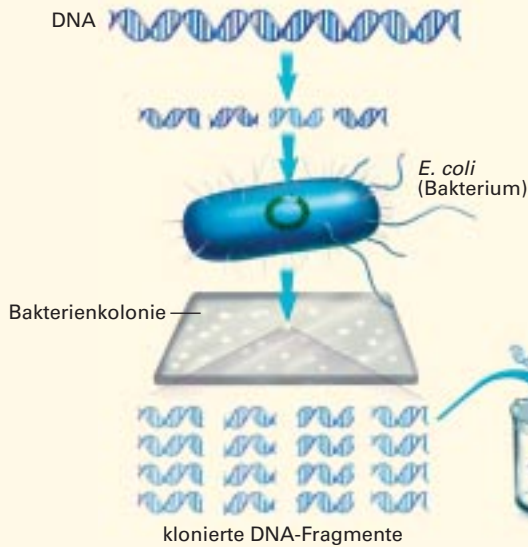
Dadurch ist die genetische Botschaft im Grunde auf beiden Einzelsträngen vorhanden, wenn auch »im Negativ« – in der komplementären Sequenz. Zellen nutzen das, um ihre DNA zu vervielfältigen und Ablesefehler zu beheben.



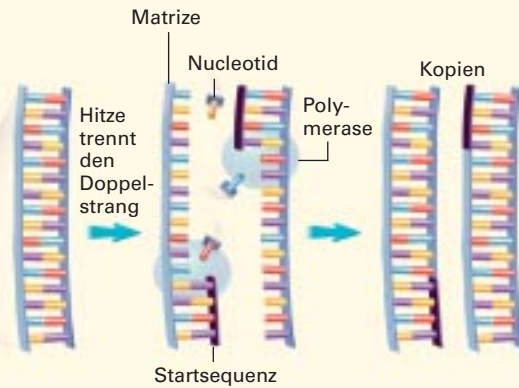
Wenn eine Zelle die DNA eines Chromosoms kopiert, wird der Doppelstrang getrennt. Jeder Einzelstrang bildet für das Enzym Polymerase eine Matrize, an der es die komplementären Basen eine nach der anderen anbaut. Auf einem der beiden gegenläufigen Stränge muss das Enzym Ligase die neuen Basen zu einem Strang verknüpfen.

Bewährte Methoden der DNA-Sequenzierung

1 Klonieren (Vervielfältigen) von DNA-Stücken in Bakterien



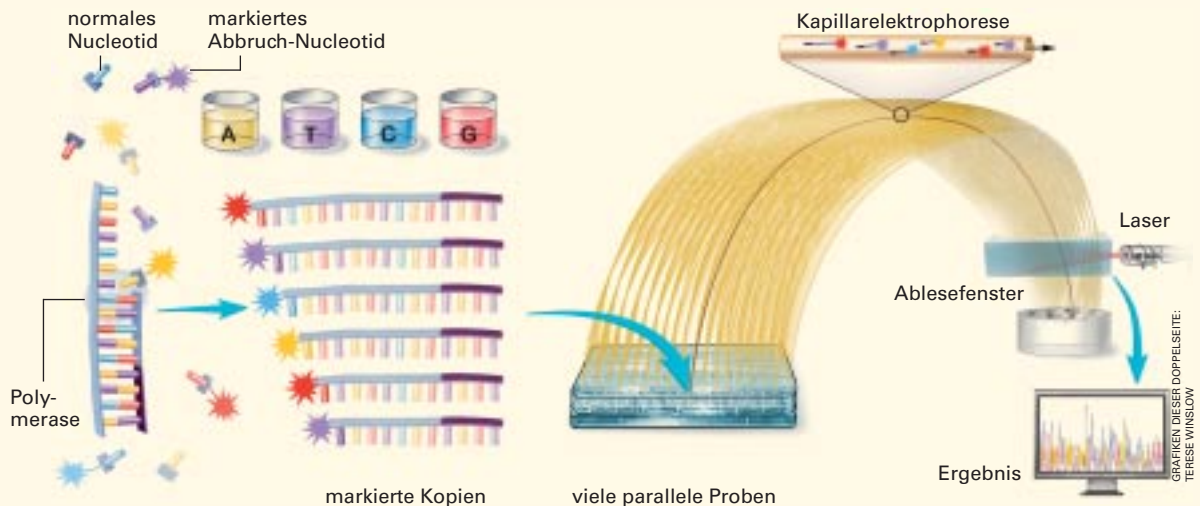
2 Die Polymerase-Kettenreaktion dient dazu, einzelne DNA-Fragmente in kurzer Zeit millionenfach zu kopieren. Immer wieder wird der Doppelstrang aufgetrennt und beide Einzelstränge werden komplementär ergänzt. Die Polymerase braucht dazu eine Startsequenz (einen Primer), ein kurzes, bereits komplementäres Stück, an dem sie ansetzt.



Modernes Sanger-Verfahren: Sequenzieren anhand der Länge und Markierung abgelesener Fragmente

3 Das Reaktionsgefäß enthält außer normalen Nucleotiden markierte Abbruchbausteine. Erwischt die Polymerase beim Kopieren des Strangs einen davon, kann sie nicht fortfahren. Je nach Länge sitzt am Ende jedes Strangs das zur Matrize an der Stelle komplementäre Nucleotid.

4 Auftrennen der markierten Fragmente nach der Länge durch Elektrophorese; Lesen der Markierung mit Laserlicht



Unter dem Schlagwort »Sequenzierung durch Synthese« werden oft jene Verfahren gefasst, bei denen direkt zum Lesen der Basen Mechanismen genutzt werden, die eine Zelle verwendet, wenn sie DNA kopiert oder repariert. Diese Vorgänge zeichnen sich in der Natur durch hohe Präzision aus. Zunächst einige Grundlagen: Zu einer Zellteilung

muss die Erbsubstanz verdoppelt werden. Die DNA-Leiter trennt sich dafür in zwei Einzelstränge. Nun tritt ein Enzym – die Polymerase – in Aktion. Es bewegt sich an den Einzelsträngen entlang und katalysiert Base für Base, genauer gesagt Nucleotid für Nucleotid, die komplementäre Sequenz. (Das Prinzip ist im Kasten links veranschaulicht.)

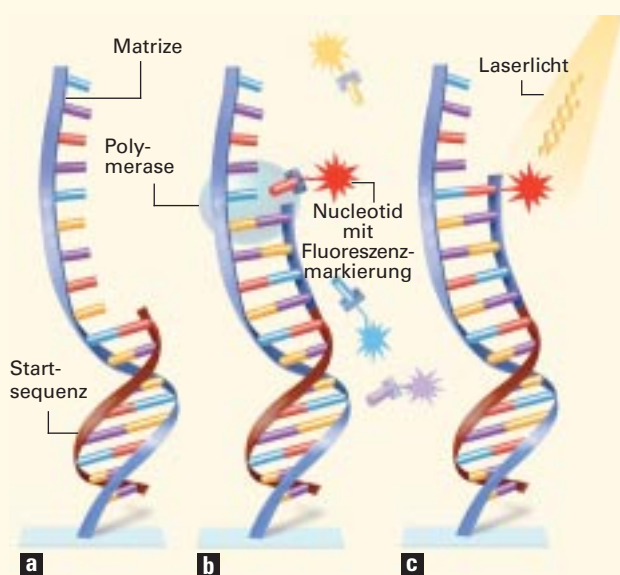
Die Polymerase nimmt also einen vorhandenen Strang als Vorlage, als Matrize, um auf die andere Seite die richtigen Basen zu setzen. Um diese miteinander zu verknüpfen, hilft ihr an einem der Stränge wegen der verschiedenen Leserichtung ein zweites Enzym – eine Ligase. Dieses kann auch sonst kurze Nucleotidsequenzen miteinander verbinden.

Sequenzieren bei der Synthese

Wie liest man DNA noch effizienter? Auch die meisten neuen Sequenziermethoden imitieren teilweise die Natur. Enzyme sorgen wiederum dafür, dass für jede Position auf der Matrize stets die richtige komplementäre Base gefunden wird. Welche das ist, wird aber sofort erfasst, wenn sich die Base anlagert. Zwei wichtige solche Verfahren sind hier dargestellt.

Jede neue Base wird gleich während der Synthese gelesen

Die Matrize ist an einer Oberfläche verankert. An einem Ende trägt sie eine Startsequenz für die Polymerase (a). Die neuen Nucleotide tragen, bis sie angelagert werden, einen Fluoreszenzmarker (b). Nach jedem einzelnen Schritt, Base für Base, werden die übrigen Reagenzien abgewaschen und die neu eingebaute Base mittels Laserlicht identifiziert (c). Vor dem nächsten Syntheseschritt wird ihre Markierung entfernt.

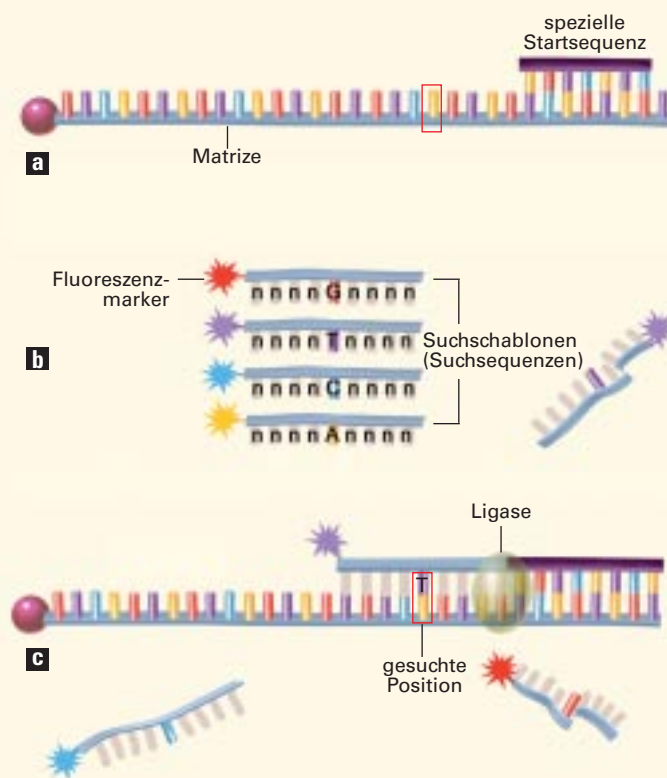


Variante Biolumineszenz:

Hierbei erzeugt ein Leuchtprotein beim Einbau eines jeden neuen Nucleotids ein basenspezifisches Lichtsignal.

Die Matrize wird mit Suchschablonen schrittweise abgegriffen

Die kurzen Suchschablonen tragen nur in der Mitte ein reguläres Nucleotid. Je nachdem, welche Base dieses Nucleotid aufweist, sind sie verschieden markiert (b). Die Startsequenzen (a) sind jeweils auf ein bestimmtes unbekanntes Nucleotid der Matrize ausgerichtet. Das Enzym Ligase bindet die passende Suchsequenz an (c). Zum Lesen eines anderen Nucleotids werden die vorige Startsequenz und die Suchschablone entfernt.



▷ Die Genomforscher imitieren jeweils Teile dieser Prozesse. Sie verfolgen so viele Ansätze dazu, dass hier nur wenige genannt werden können. Bei einem gern angewandten Verfahren, Basen-Verlängerung (*base extension*) genannt, registriert man schrittweise quasi in Echtzeit, welche Base eine Polymerase gerade angefügt hat (Kasten oben, links). Oder man benutzt kurze Suchsequenzen (Suchschablonen), die nur in der Mitte ein normales Nucleotid mit einer der vier Basen aufweisen, und erfragt damit gezielt eine bestimmte Basenposition. In dem Fall sorgt

eine Ligase dafür, dass an ein bestimmtes vorangehendes, zur Matrize komplementäres Sequenzstück jeweils die korrekte Suchsequenz angeknüpft wird – die mit dem passenden Nucleotid, also auch der passenden Base (Kasten oben, rechts). Die Technik heißt im Laborjargon Ligation, sozusagen Anbindung.

Die jeweils angefügten Basen kann man prinzipiell auf mehrere Arten nachweisen. Früher wurden oft radioaktive Markierungen benutzt. Heute verwendet man meist entweder Fluoreszenz oder Biolumineszenz. Die Fluoreszenzsignale,

für jede Basensorte eine andere Farbe, sind mit dem Lichtmikroskop nachweisbar. Damit arbeiten unter anderem Michael Metzker und seine Kollegen an der Baylor-Universität in Waco (Texas), Robi Mitra von der Washington-Universität in St. Louis (Missouri), mein eigenes Labor an der Harvard-Universität in Cambridge (Massachusetts) sowie unsere Projektmitarbeiter bei der Agencourt Bioscience Corporation.

Recht elegant, doch technisch zunächst aufwändig ist aber auch der Trick, Licht erzeugende Proteine einzusetzen,

die sich bei vielerlei Lebewesen finden. Glühwürmchen dürften das bekannteste Beispiel für solche Biolumineszenz sein. Bei dieser Art der Sequenzierung produzieren solche Proteine jeweils Licht, wenn an den Gegenstrang der DNA eine weitere Base angehängt wird (siehe Kasten links, Bild links unten). Die Lichtproteine werden infolge des Bindungsgeschehens durch dabei entstehendes Pyrophosphat angeregt. Dieses Verfahren entwickelte Mostafa Ronaghi, der heute an der Stanford-Universität (Kalifornien) arbeitet. Das System verwenden auch die Firmen Pyrosequencing/Biotage und 454 Life Sciences.

Der Trick mit den Polonien

Beide Sorten von Lichtsignalen sind natürlich winzigst und darum einzeln schwer zu erfassen. Eher gelingt das, wenn viele Moleküle gleichzeitig aufleuchten. Deswegen pflegen die Forscher zahlreiche Kopien einer DNA-Sequenz auf einmal durchzumessen – was natürlich mehr Kosten verursacht und im ganzen Ansatz auch länger dauert. Das ehrgeizige Ziel einiger Wissenschaftler ist nun, Fluoreszenz während der Sequenzierung eines einzelnen DNA-Strangs aufzufangen. An diesen Entwicklungen arbeiten Stephen Quake vom California Institute of Technology sowie Wissenschaftler von Helicos Biosciences und Nanofluidics. Bisher ist dieser Ansatz nicht ausgereift genug und darum zweischneidig, denn noch werden dabei etwa fünf Prozent der Signale übersehen. Das bedeutet, dass man jedes zu entziffernde DNA-Stück entsprechend öfter abarbeiten muss – so lange, bis seine Sequenz lückenlos vorliegt.

Ein Großteil der Arbeitsgruppen hält deswegen nach wie vor an Verfahren fest, bei denen viele Kopien der DNA-Fragmente zugleich untersucht werden. Früher ließ man die Genomstücke von Bakterien vervielfältigen (siehe Kasten S. 33, 1). Heute gibt es etliche so genannte zellfreie Methoden, um reichlich DNA-Kopien zu erhalten. Sie alle sind letztlich Varianten der so genannten Polymerase-Kettenreaktion (gern als PCR abgekürzt). Dabei verdoppelt das Enzym Polymerase zyklisch wiederholt die DNA-Menge eines Ansatzes, was schnell zu riesigen Mengen führt (siehe Kasten S. 33, 2). An neuen Methoden dieses Prinzips sind die Forscher recht erfindungsreich (siehe Kasten rechts, a und b).

Die Wortneuschöpfung »Polony« (oder Polonie), an sich im Plural verwendet, steht für ein winziges Reagenssystem, von denen sich eine Vielzahl nebeneinander unterhalten lassen. Der Begriff meint sozusagen eine Kolonie eines

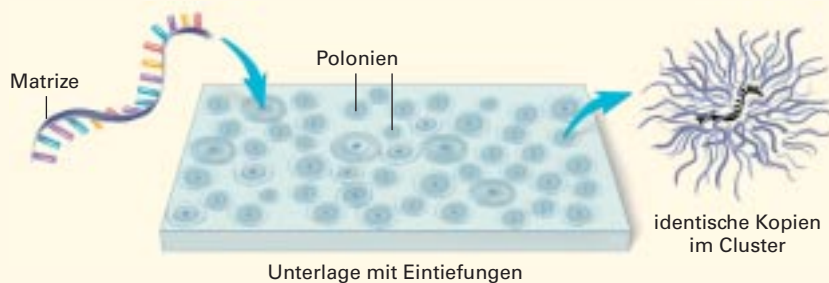
durch Polymerase vervielfältigten einzelnen DNA-Stücks – einen Pulk desselben Fragments.

Eric Kawashima vom Serono Pharmaceutical Research Institute in Genf, Alexander Chetverin von der Russischen

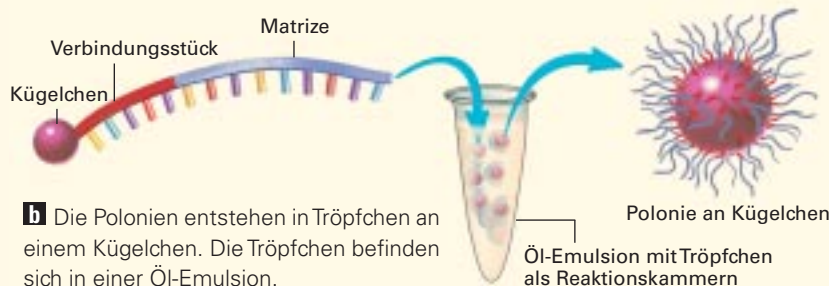
Miniaturisierte zellfreie Verfahren

Millionen Kopien durch »Polonien«

Damit die Fluoreszenzsignale stark genug sind, verwendet man gern Millionen identische DNA-Stücke gleichzeitig. In möglichst kleinen »Reaktionskammern« läuft eine Polymerase-Kettenreaktion für viele verschiedene Stränge parallel. So schafft man Polonien – regelrechte Kolonien der DNA-Stücke.



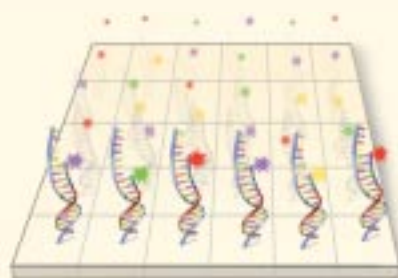
a Die Polonien entstehen in Vertiefungen auf einem Glaträger oder einem Gel.



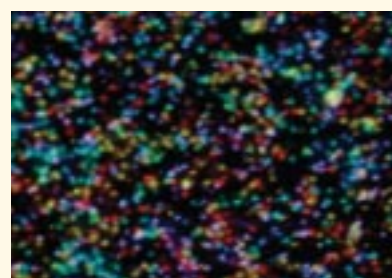
b Die Polonien entstehen in Tröpfchen an einem Kügelchen. Die Tröpfchen befinden sich in einer Öl-Emulsion.

Zigfach parallel gelesen

Beim »Multiplexing« werden Tausende oder Millionen verschiedene DNA-Fragmente gleichzeitig sequenziert. Das Prinzip lässt sich für die im Kasten auf der linken Seite links gezeigte Methode anwenden, wobei diesmal einzelne Stränge gelesen werden (unteres linkes Bild). Bei einer anderen Methode werden Polonien an Kügelchen (siehe oben, b) zu Millionen auf ein Gel aufgebracht, um sie dann mit Suchschablonen zu entziffern. Das Bild rechts unten zeigt nur 0,01 Prozent der Gesamtfläche eines solchen Ansatzes.



Arrangement mit Einzelmolekülen



Fluoreszenzsignale von Polonien

AUS: JAY SHENDURE ET AL., SCIENCE, 9. SEPT. 2005, BD. 309; GENEHMIGT VON AAAS

▷ Akademie der Wissenschaften und Mitra, der damals an der Harvard-Universität arbeitete, erzeugten viele solche Polonien nebeneinander direkt auf einer Gelschicht oder auf einem gläsernen Objektträger, wie er vielen vom Mikroskopieren her bekannt sein mag. Auf der Unterlage hatten sie dicht bei dicht, aber räumlich getrennt, einzelne DNA-Fragmente verteilt und jedes sich, wie oben beschrieben, millionenfach mittels Polymerase vervielfältigen lassen. Die Anordnung war so gewählt, dass jedes Fragment praktisch sein eigenes winziges Vervielfältigungslabor hatte und dass jedes Fragment schließlich einen getrennten Pulk, einen einzelnen Cluster bildete. Die Analogie mit einer Bakterienkolonie, die aus einer einzelnen Mikrobe entsteht, liegt dabei gar nicht so fern. Allerdings passen auf einen wenige Quadratcentimeter großen Objektträger Milliarden DNA-Polonien. Die einzelnen Cluster messen, bei einem Femtoliter Volumen, nur einen Mikrometer.

In Abwandlung dieser Methode lässt man Polonien auf winzigen Kügelchen wachsen, von denen sich jeweils eins in

einem Öltröpfchen befindet, wobei viele davon fein verteilt in einer Flüssigkeit schwimmen. In dem Fall sind die Öltröpfchen die Reaktionskammern. An jedem der Kügelchen hängt ein anderes DNA-Stück, von dem viele Kopien hergestellt werden. Zur eigentlichen Sequenzierung werden die Kügelchen auf Schablonen mit vielen feinen Vertiefungen übertragen oder auf einem Gel fixiert. Auch hierbei werden Millionen Proben nebeneinander aufgebracht und die darin enthaltene DNA gleichzeitig sequenziert.

Suche nach individuellen Genen

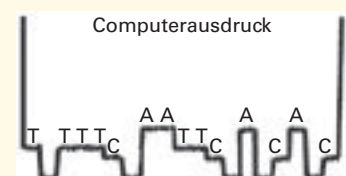
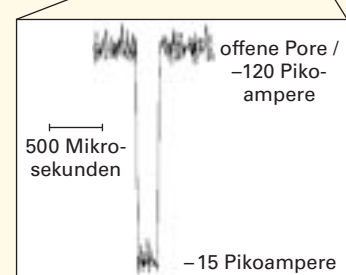
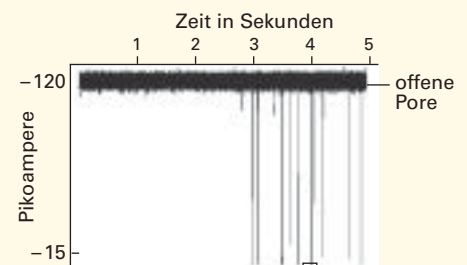
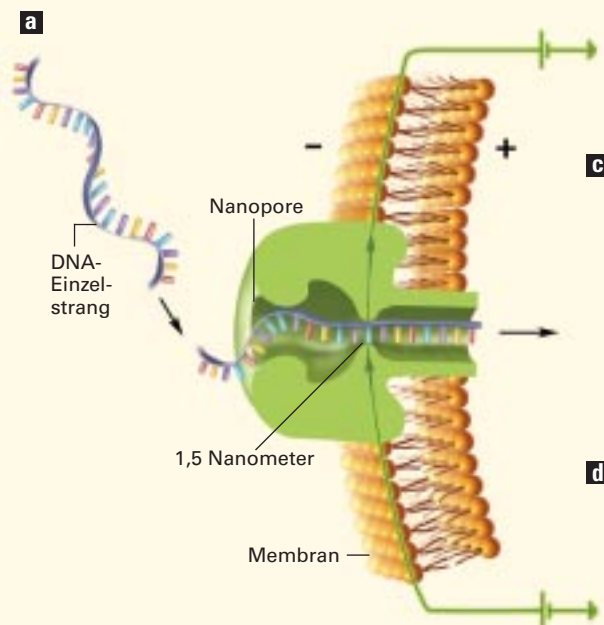
Ein anderes übergeordnetes Verfahren sollte erwähnt werden: das »Sequenzieren durch Hybridisierung« (ohne Abbildung). Diese Vorgehensweise wird gern angewandt, um bei einer schon bekannten Gensequenz nach Variationen zu suchen. Das System wird häufig kommerziell genutzt. Affymetrix, Perlegen Sciences und Illumina setzen es ein. Das Prinzip ist, zunächst viele kurze DNA-Einzelstränge in jeder denkbaren Basenkombination zu synthetisieren und diese geordnet auf einen großen Glasträger

aufzubringen. Darüber gibt man Kopien des DNA-Stücks, das untersucht werden soll. Sie lagern sich dort an, wo sie ihr komplementäres Gegenstück finden. Je genauer sich beide Teile aneinander fügen können, umso kräftiger wird das dabei auftretende Fluoreszenzsignal ausfallen. Illumina fügt anschließend noch eine Einzeldirektsequenzierung (wie im Kasten S. 34 links gezeigt) hinzu.

Ein völlig anderes Prinzip als die bisher genannten Methoden, das viel verspricht, ist die »Nanoporen-Sequenzierung« (siehe Kasten unten), eine Idee von Dan Branton von der Harvard-Universität, Dave Deamer von der Universität von Kalifornien in Santa Cruz und mir. Ausgenutzt werden Unterschiede in den physikalischen Eigenschaften der Basen im DNA-Strang. Langfristig könnte das eine Methode der Wahl werden, denn sie wäre schnell und kostengünstig. Man liest einzelne DNA-Einzelstränge, indem man sie durch feine Poren einer Membran schickt, über der eine elektrische Spannung liegt. Dabei verursacht jede Basensorte eine etwas andere Schwankung in der elektrischen

Sequenzieren mit Nanoporen

Nur noch Stunden würde es dauern, bis das komplette Genom eines Menschen mit dieser Methode erfasst wäre. Der einfache DNA-Strang wandert in einem elektrischen Feld durch eine Membran. Dazu muss er eine feine Pore queren, durch die er knapp hindurchpasst. Jedes Nucleotid blockiert kurz die Porenöffnung (a). Dadurch ändert sich vorübergehend die elektrische Leitfähigkeit der Membran, aber Größe und Dauer sind bei jeder der vier Basen etwas anders (b und c).



CGGATATTCGTC
ATGCTAAAGAT
MAIER IST DOOF
AGGGCTATTTCG
CAAGACTACTCA
TGGATATCCAGGT



Leitfähigkeit der Membran. Diese Methode entwickeln zurzeit Agilent Technologies und andere Unternehmen mit interessanten Varianten, zum Beispiel mit Fluoreszenzsignalen.

Welche Sequenzierungssysteme der nächsten Generation werden sich durchsetzen? Auf welche Eigenschaften wird es dabei ankommen? Zwei kürzlich veröffentlichte Studien bieten einen interessanten Vergleich. Beide lieferten Daten zur Effektivität, wenn man ganze Genome liest. Die eine führte meine eigene Forschungsgruppe an der Harvard-Universität durch, die andere machten Wissenschaftler der Firma 454 Life Sciences. In beiden Studien war der Sequenzumfang mit 30 Millionen Basenpaaren pro individuellem Durchgang gleich.

1700 Basenpaare pro Sekunde

Bei dem System, das mein Team beschrieb, verwendeten wir zur Sequenzierung Suchschablonen und Ligasen wie in Kasten S. 34, rechts, beschrieben. Um die untersuchte DNA zu vervielfältigen, nahmen wir Polonie-Kügelchen, zum Identifizieren der Fluoreszenzsignale der einzelnen Basen ein übliches digitales Mikroskop. Die Forscher von 454 Life Sciences benutzten einen ähnlichen Polonie-Vervielfältigungsansatz. Die Polonien kamen auf eine Schablone mit Eintiefungen und die Basensequenzen wurden per Biolumineszenz nach dem Einzelschrittmessprinzip erfasst wie in Kasten S. 34, links, beschrieben.

Der Vergleich ergab: Unser System las etwa 400 Basenpaare pro Sekunde, das der anderen Gruppe leistete 1700 Basenpaare pro Sekunde. Wir bekamen pro drei Millionen Basenpaaren weniger als einen Fehler – bei siebenfacher Abdeckung (sieben Durchläufen zur Fehlerabsicherung). Die anderen erhielten – mit einer 43fachen Abdeckung – einen Fehler

pro 2500 Basenpaare. Die Menge der erforderlichen Reagenzien hängt von der Größe der Trägerkügelchen ab. Unsere Kügelchen hatten einen Mikrometer Durchmesser, die im anderen Ansatz 28 Mikrometer in 75 Pikoliter fassenden Eintiefungen.

Und der Kostenvergleich? Die derzeit beste elektrophoresegestützte Sequenzierung kostet einen Dollar pro 150 Basenpaare für die fertige Erfassung. Unser System schaffte für einen Dollar 1400 Basenpaare, also etwa neunmal so viel wie üblich. Für die Studie von 454 Life Sciences liegen mir diesbezüglich keine Angaben vor.

Die Experten erwarten, dass die Sequenzierung der sechs Millionen Basenpaare eines individuellen menschlichen Genoms sehr bald für 100 000 Dollar zu machen sein wird. Die weitere Preisabsenkung mit Sequenziermethoden der nächsten Generation erfordert grundsätzliche Eingriffe an den Randbedingungen. Nachdem nun alle Systeme automatisiert laufen und darum menschliche Arbeitszeit immer weniger ins Gewicht fällt, stellen Reagenzien und Gerätschaften heute den größten Kostenfaktor dar. Die Miniaturisierung hat den Reagenzienverbrauch im Vergleich zu früher bereits um den Faktor eine Milliarde (10^9) reduziert. Beim konventionellen Sanger-Verfahren rechnete man noch in Mikrolitern (millionstel Liter) – heute sind es Femtoliter (10^{-15} Liter).

Beim Sammeln der Rohdaten erreichen die Apparaturen schon eine Rate von einem Gigabyte (einer Milliarde Bytes) pro Minute. Da Computer, die diese Informationen verarbeiten, mehrere Milliarden Operationen pro Sekunde leisten, sind das Entscheidende nun die Engpässe etwa durch eine Elektrophorese, eine Enzymreaktion oder andere Zeit raubende physikalische oder chemi- ▷

WICHTIGE ONLINE- ADRESSEN

» Das Universum – unendliche Weiten ...

Bücher, Lehrmaterial, DVDs, Puzzle, Globen

und mehr zum Thema Astronomie und Raumfahrt

www.spacebooks-etc.de

» Dipl.-Ing. Runald Meyer VDI

Entwicklung, Konstruktion,

Technische Berechnung

Strömungsmechanik

www.etastern.de

» DOK –

Düsseldorfer Optik-Kontor

Kontaktlinsen online bestellen

www.dok.de

» Kernmechanik –

Optimiertes Modell:

Kernspin + Dipolmomente

www.kernmechanik.de

Hier können Sie den Leserinnen und Lesern von Spektrum der Wissenschaft Ihre WWW-Adresse mitteilen. Für € 83,00 pro Monat (zzgl. MwSt.) erhalten Sie einen maximal fünfzeiligen Eintrag, der zusätzlich auf der Internetseite von Spektrum der Wissenschaft erscheint. Mehr Informationen dazu von

GWP media-marketing

Mareike Grigo

Telefon: 0211 6188-579

E-Mail: m.grigo@vhb.de

Auf dem Weg zur persönlichen Genomsequenz

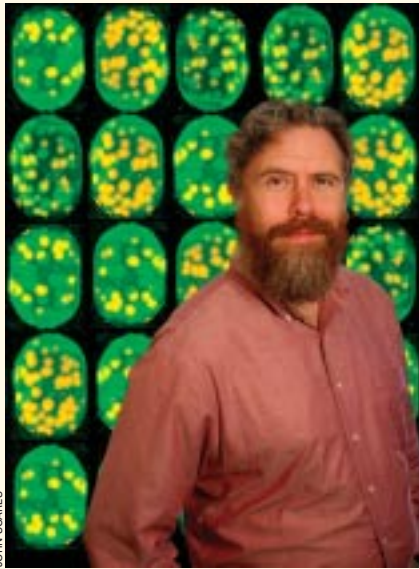
In vielen Ländern werden heute Blutproben Neugeborener auf die Erbkrankheit Phenylketonurie hin untersucht – eine Stoffwechselstörung, die unbehandelt schwere Entwicklungsdefekte erzeugt. Zunehmend wird es auch möglich, individuell genetisch zu testen, ob der Organismus eines Patienten auf ein bestimmtes Medikament

überhaupt ansprechen würde, etwa bei manchen Formen von Lungenkrebs. Auch sonst reagieren Menschen auf Wirkstoffe nicht gleich. Welche Dosierung im Einzelfall angebracht ist, können Ärzte inzwischen vielfach mit entsprechenden Labortests herausfinden (siehe Spektrum der Wissenschaft 5/2005, S. 96).

Doch die steigende Nachfrage nach persönlichen Genomdaten betrifft nicht nur den medizinischen Bereich. Wer wüsste nicht gern, ob er von Karl dem Großen abstammt oder einem der Krieger von Dschingis Khan? Und mit wem alles er näher verwandt ist? Mancher mag sich fragen, wie Gene im Wechselspiel miteinander und mit der Umwelt sein Aussehen und Wesen, seine Vorzüge und Schwächen gebildet haben. In der Zukunft dürfte die Wissenschaft manche Antwort hierauf aus riesigen Datensätzen herauslesen.

Schon wird Besorgnis laut, Betrüger und Geschäftemacher, Versicherungen, Arbeitgeber, Polizei, selbst Freunde oder Nachbarn könnten derlei Informationen missbrauchen. Niemand kann vorhersagen, wie es sich in der neuen Ära leben wird. Um für sie gerüstet zu sein, haben meine Kollegen und ich das »Projekt persönliches Genom« ins Leben gerufen. Im Grunde ist das nur der natürliche nächste Schritt nach dem Humangenomprojekt, der Entzifferung eines kompletten menschlichen Genoms.

Um mögliche Nutzen und Risiken zu erkennen, die auftreten können, wenn jedermanns Genomsequenz zugänglich ist, wollen wir Freiwillige finden, die der Öffentlichkeit die Daten ihres Genoms sowie aller körperlichen Merkmale und Eigenschaften zur Verfügung stellen. Das heißt, zu dieser Information werden die kompletten DNA-Sequenzen der 46 Chromosomen gehören, dazu Datensätze aller medizinischen Unterlagen, zudem aber auch umfangreiche Daten von Genaktivitäten und Proteinen, von Körper- und Gesichtsmaßen, von modernen medizinischen Aufnahmen des Körperinnern –



JOHN SOARES

George M. Church, Autor dieses Artikels, will wie andere Freiwillige seine Genomsequenz öffentlich machen. (Im Hintergrund Bilder von Polonien)

▷ sche Prozesse. Limitierend wirken auch insgesamt zu langsame oder zu raumgreifende, mithin zu locker gepackte Arrangements. Jedes Pixel zählt.

Zudem ist bei solchen Rationierungen abzuwägen, wozu man die Daten benötigt. Die neueren Methoden eignen sich meist nur für vergleichsweise kurze DNA-Stücke aus 400 bis 500 Basenpaaren. Das alte Verfahren las 800 Paare. Aus so kurzen Stücken das Genom eines neuen Organismus zu rekonstruieren, wäre wesentlich aufwändiger als früher mit den längeren Fragmenten. Darauf käme es jedoch in unserem Fall gar nicht an, wenn vor allem der medizinische Nutzen gefragt ist. Denn eine Gesamtversion des menschlichen Genoms, an der man sich orientieren kann, liegt ja schon vor. Beim Sequenzieren weiterer menschlicher Genome möchte man vor allem einzelne individuelle Varianten im Vergleich zu bekannten Sequenzen erkennen.

Auch die erforderliche Genauigkeit richtet sich nach dem Zweck der Daten. Benötigt man sie, wie gesagt, zur medizinischen Diagnostik, müsste man die Le-

sefehlerrate vielleicht noch unter den gegenwärtigen Standard von 0,01 Prozent senken (einen Fehler auf 10 000 Basen). Denn dies bedeutet immerhin 600 000 Fehler pro individuelles Genom. Gewebetypen klassifizieren und herausfinden, welche Gene darin aktiv sind, kann man schon anhand von Stichproben mit vier Prozent falschen Angaben. Auch in der ökologischen Forschung genügen oft grobe genetische Stichproben. Zwanzig Basenpaare können bei manchen Ansätzen ausreichen, um einen Organismus in einem Ökosystem zu identifizieren.

Bis die Genomsequenzierung für jedermann erschwinglich sein wird, bleibt noch viel zu tun. Angesichts des knappen gesetzten Zeitrahmens bedeutet das erhebliche Anstrengungen. Es kommt beileibe nicht nur darauf an, die Entzifferungsprozedur nochmals erheblich zu beschleunigen und zu verbilligen. Sondern wir müssen auf die Neuerung auch vorbereitet sein. So brauchen wir beispielsweise Datensysteme, mit denen ganz normale Ärzte etwas anfangen können. Von jedem einzelnen Patienten müssten

sie die – sagen wir zehn – speziell bei ihm vorhandenen wichtigsten genetischen Varianten kennen, die gerade für diesen Menschen besonders krankheits- oder behandlungsrelevant sein könnten.

Problem Privatsphäre

Ein zentraler Aspekt sind die möglichen Folgen für den Einzelnen, wenn solche Verfahren und Daten weiten Kreisen zugänglich sind. Zum Humangenomprojekt gehörte von Anfang an ein mit 10 Millionen Dollar pro Jahr finanziertes Programm, das den ethischen, juristischen und sozialen Auswirkungen des Vorhabens galt. Das führte zu einigen wichtigen Entscheidungen.

Die Mitwirkenden kamen überein, alle anfallenden Daten binnen einer Woche zu veröffentlichen, was es so noch nie gegeben hatte. Auch setzten wir uns dagegen ein, die Natur des Menschen zu kommerzialisieren. Spezielle Anstrengungen wurden unternommen, um die Anonymität der veröffentlichten Genomdaten zu wahren: Zur Sequenzierung wurden die Chromosomen mehrerer Per-



kurz alles, was zukünftig zu einem persönlichen Gesundheitsprofil gehören könnte. Zusätzlich sollen von jedem Teilnehmer Zelllinien zentral aufbewahrt werden. Der Zweck: Jedermann kann auf diese Daten zugreifen und daran Hypothesen und Modelle prüfen. Und die Hoffnung: Die Erkenntnisse führen zu neuen Erklärungsmodellen.

Hier ein einfaches Beispiel dafür, welche Vorteile die offene Handhabe von derlei Daten haben kann: Im Rahmen des Humangenomprojekts stehen einige persönliche medizinische Details im Netz – so auch von mir. Eines Tages meldete sich ein Hämatologe vom anderen Ende des Landes. Er meinte, eine Überprüfung meiner Cholesterinwerte sei überfällig. Ich kam dem nach, woraufhin meine Medikation geändert wurde und ich meine Ernährung umgestellt habe. Letztlich verdanke ich also der Veröffentlichung einiger meiner persönlichen Daten, dass zumindest ein Risikofaktor drastisch gesenkt wurde. In Zukunft würde eine genomische Software dieserart Warnungen geben.

Ein internes Gremium meiner Fakultät an der Harvard-Universität hat das Projekt persönliches Genom (englisch: *Personal Genome Project, PGP*) genehmigt. Wie immer bei Studien an Menschen müssen die Teilnehmer vorher über Risiken in Kenntnis gesetzt werden. Außerdem können sie sich über Erfahrungen früherer Teilnehmer informieren lassen, bevor sie der Datenfreigabe zustimmen. Letztlich ist die offene Handhabe für alle Beteiligten weniger riskant als die umgekehrte Situation, wenn geschützte Daten versehentlich oder absichtlich veröffentlicht werden.

Wie beim Humangenomprojekt auch soll der freie Datenzugang neue Entdeckungen erleichtern. Daneben lässt sich austesten, wie die Allgemeinheit darauf reagiert, zum Beispiel beim Versicherungsschutz.

Anfangs werden private Spenden noch helfen, die Versuchspersonen gegen genetische Diskriminierung abzusichern. In dieser ersten Phase muss Profit also keine große Rolle spielen. Dennoch dürften Versicherungen das Projekt interessiert beobachten.

sonen gemischt. Doch bleiben wichtige Fragen. Wie zum Beispiel lässt sich sicherstellen, dass niemand personenbezogene genetische Informationen missbraucht, ob nun Wissenschaftler, Versicherungen, Arbeitgeber, Gerichte, Schulen, Adoptionsstellen oder Regierungen? Wie kann man gewährleisten, dass die Privatsphäre bei Entscheidungen über Schwangerschaften oder medizinische Behandlungen gewahrt wird beziehungsweise dass Eingeweihte sich korrekt und fair verhalten?

All dem müssen wir genauso rigoros nachgehen wie im Labor den technologischen Neuerungen und rein biologischen Fakten. Ein Kreis von Wissenschaftlern hat darum das »Projekt persönliches Genom« ins Leben gerufen (siehe Kasten oben). Wir wollen Risiken und Nutzen untersuchen, die eine Ära mit sich bringen könnte, in der die Menschen ihr Genom sequenzieren lassen.

Die meisten wissen, dass es nirgends eine letzte Sicherheit gibt. Wer in Aktien investiert, akzeptiert das Risiko und das komplexe Verhalten der Märkte. Auch sonst im Leben wägen wir in vielem die

Wahrscheinlichkeiten für einen Gewinn gegen den Einsatz ab. Wer hätte wohl zu Zeiten der ersten Computer gedacht, dass sie einmal große ökonomische, soziale und wissenschaftliche Umwälzungen bringen würden? Für das Genomzeitalter sollten wir auf ähnlich revolutionäre Veränderungen gefasst sein. ◁

George M. Church ist Professor für Genetik an der Harvard-Universität in Cambridge (Massachusetts) und Direktor des dortigen Zentrums für Computergenetik. Er leitet mehrere andere amerikanische Forschungszentren, hält zehn US-Patente und berät mehr als 20 Firmen.

Das neue Genom. Spektrum der Wissenschaft, Dossier 1/2006

The race for the \$1000 Genome. Von R.F. Service in: Science, 17. März 2006, S. 1544

Advanced sequencing technologies: Methods and goals. Von J. Shendure et al. in: Nature Reviews Genetics, Bd. 5, Mai 2004, S. 335

Weblinks zum Thema finden Sie bei www.spektrum.de unter »Inhaltsverzeichnis«.

AUTOR UND LITERATURHINWEISE

spektrumdirekt
Die Wissenschaftszeitung im Internet

Die Redaktion von **spektrumdirekt** informiert Sie online schnell, fundiert und verständlich über den Stand der Forschung.



www.spektrumdirekt.de/info