

# Genom-Entschlüsselung mit Hochgeschwindigkeitsverfahren: Bedeutung für Wissenschaft und Medizin

Shankar Balasubramanian\*

DNA · Genomik · Genomsequenzierung ·  
Personalisierte Medizin · Solexa-Verfahren

Die Aufklärung des genetischen Codes vor 50 Jahren<sup>[1]</sup> bedeutete einen Quantensprung für unser Verständnis der Nucleinsäuren.<sup>[1]</sup> Vorausgegangen war die Bestimmung der Regeln, nach denen sich komplementäre Nucleinsäuren durch Watson-Crick-Basenpaarung erkennen.<sup>[2]</sup> Der Code beruht auf einem genetischen Alphabet aus vier Nucleobasen, die als Sequenzen dreier Basen abgelesen werden, welche jeweils ein Codon bilden, das einer Aminosäure entspricht. Die codierenden Bereiche eines Genoms legen somit die Proteine fest, die in den DNA-Sequenzen verschlüsselt sind. Dieses Wissen zog unvermeidlich einen Bedarf an Methoden nach sich, um Nucleotidsequenzen der DNA im Laboratorium zu entschlüsseln. Es verdient Erwähnung, dass der Bedarf an einer Methode zur Bestimmung von Nucleotidsequenzen bereits vor der Aufklärung des genetischen Codes durch Brown und Todd erkannt wurde.<sup>[3]</sup> Zwei praktikable chemische Wege zur Entschlüsselung von DNA wurden in den 70er Jahren entwickelt: 1) die Methode der sequenzabhängigen chemischen Spaltung nach Maxam und Gilbert<sup>[4]</sup> und 2) die Methode der DNA-Synthese mit Kettenabbruch nach Sanger und Mitarbeitern.<sup>[5]</sup> Von beiden Methoden setzte sich der Sanger'sche Ansatz letztendlich durch, der in der Folge beträchtlich optimiert und automatisiert wurde, um die Routinesequenzierung von Nucleinsäuren und kleinen Genomen zu ermöglichen. Der ultimative Beweis für die Leistungsfähigkeit der Sanger-Methode war die Entschlüsselung der menschlichen Genomsequenz durch das internationale Humangenomprojekt.<sup>[6]</sup> Die Bestimmung von Nucleinsäuresequenzen ist ein integraler Bestandteil der modernen biologischen und medizinischen Forschung geworden.

In jüngster Vergangenheit sind neue Verfahren für die akkurate Entschlüsselung von DNA- und RNA-Sequenzen aufgetreten, welche die Art und Weise, wie wir Informationen über lebende Systeme herleiten und nutzen können, verändern dürften – vor allem im Hinblick auf Geschwindigkeit

Ausmaß und Kosten. Dieser Essay beschreibt die Erfindung der Methode der Solexa-Sequenzierung (nunmehr Solexa-Illumina-Sequenzierung)<sup>[7–9]</sup> und deren Bedeutung für die Naturwissenschaften. Außerdem wird ein Ausblick auf mögliche Auswirkungen auf die Medizin gegeben. Es gibt eine Anzahl anderer Sequenzierungskonzepte, die schon in der praktischen Anwendung sind oder sich noch in der Entwicklung befinden, und die an anderer Stelle beschrieben wurden.<sup>[10]</sup>

Im Jahr 1994 begann ich meine eigenständige akademische Laufbahn mit einem kleinen Team aus vier Doktoranden, die an völlig separaten Projekten arbeiteten. Bei einem der Projekte wurde die Methode des resonanten Förster-Energietransfers verwendet, um zu untersuchen, wie eine DNA-Polymerase im Detail ein DNA-Molekül verlängert. Im Verlauf der Arbeiten hielt ich es für sinnvoll, zeitlich aufgelöste Fluoreszenzmessungen durchführen zu lassen. Hierzu kontaktierte ich meinen Kollegen David Klennerman aus der physikalischen Chemie, der ungefähr zur selben Zeit wie ich seine eigene Arbeitsgruppe gegründet hatte. Neben dem Abschluss besagter Studie führte meine Zusammenarbeit mit David zu anregenden Diskussionen darüber, was wir noch durch die Untersuchung von Molekülen mit (damals neuen) optischen Methoden herausfinden konnten. Das sich daraus ergebende Projekt hatte die Einzelmolekülbildgebung einer DNA-Polymerase mit einem Matrizen/Primer-Substrat zum Ziel, um den Einbau von Nucleotiden im Verlauf der enzymatischen DNA-Synthese beobachten zu können. Wir entwarfen verschiedene Experimente, um durch unterschiedliche Platzierung von Fluorophoren an der Polymerase, der DNA und den Monomeren (den Desoxyribonucleotidtriphosphaten, dNTPs) jeweils verschiedene Aspekte der Struktur, der Funktion und des Mechanismus der DNA-Synthese visualisieren zu können. Eine bestimmte Versuchsanordnung bestand darin, das Konstrukt aus DNA-Matrize und Primer zu immobilisieren und den matrizen- und polymerasegelenkten Einbau von fluoreszenzmarkierten Nucleotiden abzubilden. Auf der Einzelmolekülebene bot uns dies die Aussicht, die DNA-Synthese in Echtzeit zu verfolgen. Es war offenkundig, dass ein solches Experiment das Potential hatte, die immobilisierte Matrize durch sequenzielle Detektion der eingebauten Fluorophore zu entschlüsseln. Zu jener Zeit (1997) waren wir uns der Tatsache bewusst, dass unsere Kollegen am nahegelegenen Wellcome Trust Sanger Institute gute Fortschritte bei der Entschlüsselung des menschlichen

[\*] Prof. S. Balasubramanian  
Department of Chemistry, University of Cambridge  
Lensfield Road, Cambridge, CB2 1EW (Großbritannien)  
und  
Cancer Research UK Cambridge Research Institute, Li Ka Shing  
Centre, Robinson Way, Cambridge, CB2 0RE (Großbritannien)  
und  
School of Clinical Medicine, University of Cambridge  
Cambridge, CB2 0SP (Großbritannien)  
E-Mail: sb10031@cam.ac.uk

Genoms im Rahmen des internationalen Humangenomprojekts machten. Unsere Studien waren außerdem von einem Konzept zur DNA-Entschlüsselung (oder Sequenzierung) inspiriert, demzufolge zuerst ein DNA-Strang synthetisiert wurde, in dem jede Base durch einen Fluorophor markiert war, sodass beim systematischen Abbau der DNA von einem Ende her jedes Nucleotid (und jede Markierung) nacheinander (sequenziell) detektiert werden konnte.<sup>[11]</sup> Im Verlauf des Jahres 1997 diskutierten David und ich mit einigen unserer Mitarbeiter sehr intensiv über das Thema. Im Gegensatz zum Abbauprozess stellten wir uns einen Ansatz vor, mit dem die akkurate Entschlüsselung von DNA durch eine DNA-Synthese gelingen sollte: eine kontrollierte schrittweise Festphasensynthese mittels DNA-Polymerase unter Verwendung markierter Bausteine, die mit 3'-O-Schutzgruppen und abspaltbaren Fluorophoren modifiziert sind, und wo der Einbauzyklus zum Einbau des korrekten Nucleotids führt, das dann anhand seiner Farbmarkierung identifiziert werden kann (Abbildung 1). Die Entfernung der Farbmarkierung und der 3'-O-Schutzgruppe setzt das System für den nächsten Decodierungszyklus an der nächsten Nucleotidposition zurück. Eine entscheidende Bedingung für das Funktionieren dieses Ansatzes war, dass man in der Lage sein musste, alle vier farbcodierten dNTPs zusammen einzusetzen, sodass die natürliche Konkurrenz der Nucleotide und damit deren korrekter Einbau durch die DNA-Polymerase gegeben wäre. Dies wiederum erforderte eine absolute chemische Kontrolle – und zwar mithilfe der Schutzgruppe –, um den Einbau auf ein einzelnes Monomer zu beschränken. Ein alternativer, potentiell einfacherer Ansatz hätte darin bestanden, auf die 3'-O-Schutzgruppe zu verzichten, nur jeweils ein markiertes Nucleotid einzubauen, das Fehlen oder Vorhandensein eines Signals zu registrieren und diese Schritte dann systematisch für jedes der vier Nucleotide zu wiederholen. Letzteren Ansatz haben wir verworfen, weil Genome oftmals Abfolgen gleicher Basen enthalten und wir daher zu der Überlegung kamen, dass ein genaues Abzählen der Guanine, Cytosine, Thymin oder Adenine in solchen Abschnitten am besten durch eine absolute Kontrolle über jeden Einbauschritt zu erreichen ist.

Das Prinzip, die Sequenzierung an einer festen Phase vorzunehmen, erschien deshalb attraktiv, da so die simultane Entschlüsselung einer potentiell riesigen Zahl von DNA-Proben oder -Fragmenten möglich wäre, woraus sich ein sehr hoher Durchsatz ergibt. Abbildung 2 zeigt eine frühe Schätzung, die David Klennerman und ich vorgenommen hatten.

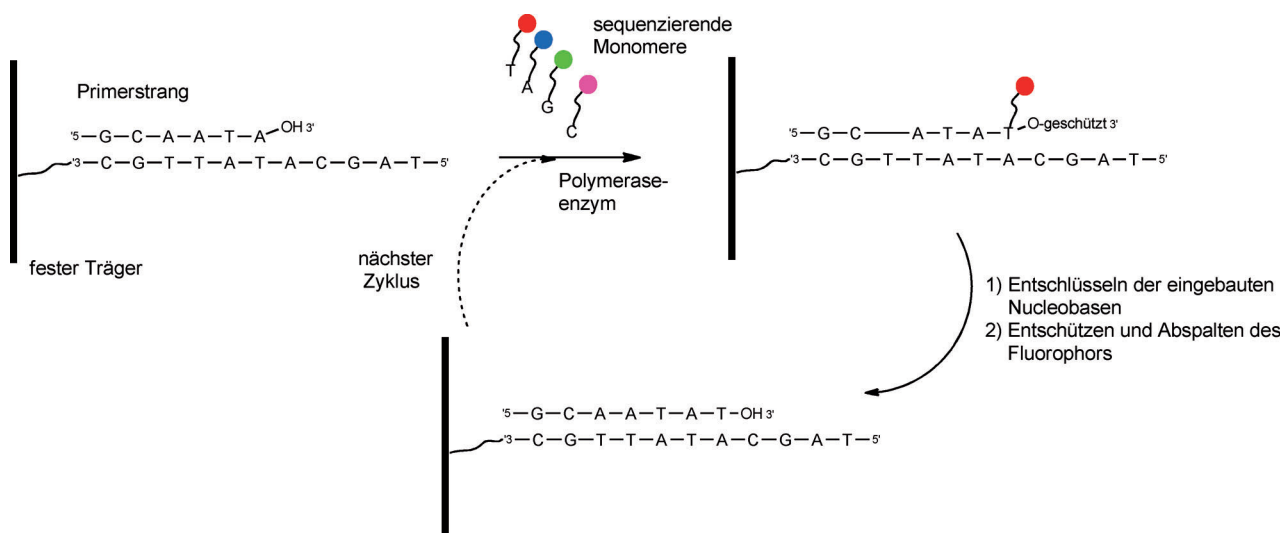


Shankar Balasubramanian ist Herchel Smith Professor für Pharmazeutische Chemie an der Cambridge University. Seine Forschungsinteressen gelten der chemischen Biologie von Nucleinsäuren und dem Genom sowie der Identifizierung, Aufklärung und Manipulation nichtcodierender genetischer Elemente, insbesondere von G-Quadruplexen. Er ist einer der Erfinder des Solexa-Sequenzierverfahrens.

Unsere Vorhersagen bezüglich Ableselänge und Grad an Parallelisierung sollten sich später als etwas „daneben“ herausstellen, dennoch ermöglichte es uns diese einfache Analyse, das Potential für die Sequenzierung des menschlichen Genoms vorauszusehen (im Bereich Milliarden Basen pro Experiment im Verhältnis zu den 3 Milliarden Basen des menschlichen Genoms). Eine praktische Vorgehensweise zum Aufbau eines sequenzierfertigen Festphasenarrays von DNA-Proben musste „in einem Topf“ durchzuführen sein, um die arbeitsintensive Präparation jedes einzelnen Probenfragments zu vermeiden. Unsere Vorstellung war, die genomische DNA zu fragmentieren und dann alle Fragmente in hoher Verdünnung auf der Oberfläche eines Chips zu immobilisieren, sodass jeder optisch auflösbare Spot nur von einem einzigen DNA-Fragment besetzt wäre. Mit diesem Einzelmolekül-Array würde es gelingen, eine Vielzahl von DNA-Anschnitten zu präparieren, die durch Festphasensequenzierung simultan decodierbar wären.

Trotz begrenzter Mittel gingen David und ich über unsere Grundlagenstudien hinaus und brachten eine Machbarkeitsstudie auf den Weg, die auf die Chemie und Enzymologie der Festphasensequenzierung von DNA in Verbindung mit einer Bildgebung der Decodierungsreaktionen auf der Oberfläche eines Glaschips abzielte. Wir waren überzeugt, den Ansatz zur Praxisreife führen zu können. Zur selben Zeit realisierten wir, dass die Umsetzung dieses Konzepts in einem Format, das auch von anderen Arbeitsgruppen verwendet werden könnte, erheblich mehr Mittel erforderte, als uns zur Verfügung standen, und dass wir außerdem mehr Expertise auf anderen Gebieten brauchten. Um unser Ziel zu erreichen, wendeten wir uns Ende 1997 an Investoren, und nachdem man unsere Konzepte eingehend geprüft hatte, gründeten wir im Sommer 1998 eine Firma. Wir nannten unseren Ansatz „Solexa“, zusammengesetzt aus „solo“ und „molekular“ (sowie, aus rein sprachlichen Gründen, einem „x“ in der Mitte), um damit unserem Denken in Einzelmolekülkategorien Rechnung zu tragen. Die zwei folgenden Jahre brachten wir unsere Machbarkeitsstudie an der Cambridge University voran, bevor wir die Forschungen dann im Jahr 2000 in ein eigenes Firmengebäude außerhalb von Cambridge verlagerten (Solexa Limited). Im Lauf der folgenden fünf bis sechs Jahre fügten wir ein herausragendes Team von Wissenschaftlern und Managern zusammen, um das Solexa-Konzept in ein robustes kommerzielles System zu entwickeln. Während dieser Zeit wurden weitere Mittel akquiriert und außerdem eine US-Firma (Lynx) übernommen, um eine Fertigungsstätte für Apparatebau und Produktion zur Verfügung zu haben.

Die Entwicklung eines Sequenziersystems erforderte eine Zusammenführung verschiedenster technischer Bereiche: Chemie, um die Nucleotide zu entwickeln und die Oberfläche des Sequenzierchips zu modifizieren; Protein-Engineering, um eine mit Hochdurchsatzverfahren kompatible DNA-Polymerase zu erzeugen; Molekularbiologie, um eine praktische Methode zur Probenvorbereitung zu entwickeln; Ingenieurwesen und Physik, um das Bildgebungssystem und andere Geräte zu entwickeln; sowie Computerkenntnisse, um die Software und Algorithmen für die Datenanalyse zu entwickeln. Eine detailliertere technische Beschreibung ist an-



**Abbildung 1.** Prinzip der Festphasen-DNA-Sequenzierung. Pro Zyklus wird eine einzelne Base pro DNA-Probe durch Detektion des codierenden Fluorophors entschlüsselt. Mit dem gleichen Prinzip werden heute simultan hunderte von Millionen bis Milliarden DNA-Proben in einem Mikrofluidiksystem analysiert.

menschliches Genom  $3 \times 10^9$  Basen

Arrayformat 300 X 300

10 s pro Zyklus

$$\therefore \frac{10^5}{10} \text{ Basen pro Sekunde} = 10 \text{ kb s}^{-1}$$

$$\therefore \text{eine Maschine } \frac{3 \times 10^9}{10^4} \text{ s menschliches Genom.}$$

$$\frac{3 \times 10^5}{60 \times 60} \text{ Std} \approx 100 \text{ Std} \approx 5 \text{ Tage}$$

**Abbildung 2.** Unsere erste Schätzung der möglichen Sequenzierkapazität des Solexa-Verfahrens aus dem Jahr 1997. Bei dieser Überschlagerrechnung gingen wir von nur ca.  $10^4$  DNA-Fragmenten in einem Probenarray aus. Es war aber klar, dass die Sequenzierung von ungefähr einer Milliarde Basen in einem Zeitraum von wenigen Tagen möglich sein würde. Der Genome Analyser erreichte dieses Ziel, allerdings mit einem weit höheren Parallelisierungsgrad (ca.  $10^7$ ). In der Folge übertraf dann das HiSeq2000-System die 1997 formulierten Erwartungen um das rund 1000-fache.

dernorts gegeben.<sup>[7,8]</sup> Als eine bedeutende Verfeinerung unseres ursprünglichen Konzept enthielt das fertige System einen Schritt zur Amplifizierung des Einzelmolekül-Probenarrays, um Kopien des ursprünglichen DNA-Fragments an einem spezifischen Ort auf dem Chip zu erzeugen. Erreicht wurde dies durch Adaptierung einer eleganten Brückenamplifikationsmethode, die andernorts erfunden<sup>[12]</sup> und unabhängig davon durch die Firma Mantea Predictive Medicine weiterentwickelt worden war. Die Amplifizierung des Einzelmolekül-DNA-Arrays führte zu stärkeren Signalen, wodurch ein kostengünstigeres Bildgebungssystem verwendet werden konnte und außerdem stochastische Fehler, wie sie beim Arbeiten mit Einzelmolekülen auftreten können, reduziert wurden.

Das allererste Genom, das mit der Solexa-Methode sequenziert wurde, war das des Bakteriophagen phiX174 im Jahr 2005. 2006 wurde dann das erste exportierbare Solexa-Sequenziersystem unter dem Namen „Genome Analyser“ auf den Markt gebracht, das in einem Versuchsdurchgang ca. 1 Milliarde Basen (1 Gigabase = 1 Gb) der DNA sequenzieren konnte. Damit hatten wir unsere 1 Gb Kapazität getroffen, die wir uns 1997 als verwegenes Ziel gesetzt hatten (Abbildung 2). Dieses System war in der Lage, mehrere 10 Millionen individuelle Fragmente von jeweils etwa 35 Basen Länge zu erzeugen. Mithilfe des Genome Analysers wurden im Jahr 2008 das Genom eines männlichen Yoruba,<sup>[7]</sup> das erste Genom eines Asiaten<sup>[13]</sup> und das Genom eines Patienten mit akuter myeloider Leukämie neben einem Vergleichs-genom einer normalen Hautzelle derselben Person<sup>[14]</sup> nach dem Solexa-Verfahren sequenziert. Mit der Veröffentlichung dieser Genome – in nur einer einzigen Ausgabe von *Nature*! – verdoppelte sich die Zahl der öffentlich zugänglichen Humangenomsequenzen, und jegliche Zweifel, ob tatsächlich ganze Genome durch parallele Entschlüsselung kurzer DNA-Stücke sequenzierbar seien, waren damit ausgeräumt. Die Kerntechnologie wurde in den folgenden Jahren durch die Firma Illumina (die Solexa und die Sequenzieretechnik zu Beginn des Jahres 2007 kaufte) auf vielen Ebenen verbessert und optimiert, hin zu einer um drei Größenordnungen gesteigerten Kapazität. Das heute bei Solexa-Illumina eingesetzte System „HiSeq 2000“ sequenziert im Routinebetrieb etwa 600 Gb pro Lauf; das entspricht einer um eine Million höheren Kapazität eines Sequenziersystems, als wir es 1997 vor Augen hatten. Die Kosten für die Sequenzierung eines menschlichen Genoms sanken in den letzten 10 Jahren von mehreren hundert Millionen US-Dollar auf unter 5000 US-Dollar – eine rund 100 000-fache Verbilligung.

Der Gang von den Kernkonzepten hin zu einem Hochgeschwindigkeits-Sequenziersystem war nicht unbegleitet von Zweifeln (und Zweiflern). Ein früher Einwand, der mir im Gedächtnis geblieben ist, war, ob jedwede Methode zur

routinemäßigen, kostengünstigen Sequenzierung menschlicher Genome von Interesse für die Fachwelt (und die Öffentlichkeit) wäre. Bei genauem Nachdenken erscheint dieser Einwand gar nicht unvernünftig, wenn man bedenkt, dass es zu dieser Zeit (1997) keine menschlichen Genomsequenzen gab (und auch keinen Markt für die Sequenzierung ganzer Genome). Wir hatten die Erwartung, dass die Entschlüsselung des ersten menschlichen Genoms (was 2003 geschehen sollte) Jahrzehnte weitergehender Forschungen entfesseln würde, in denen die Routinesequenzierung von Genomen entscheidend für die detaillierte Aufklärung von Genomabweichungen und deren Bedeutung z.B. für menschliche Krankheiten wäre. Unser Optimismus, der zumindest in Teilen naiv war, wurde nicht zuletzt gestärkt durch unseren Kontakt zu drei Hauptfiguren der Humangenomforschung,<sup>[15]</sup> die wir bei unserem Besuch des Wellcome Trust Sanger Institute 1998 treffen durften.

Heute wird die Sequenzierung ganzer Genome immer mehr Routine, was sich auch in einer Reihe prominenter Projekte niederschlägt, wie etwa dem 1000-Genome-Projekt<sup>[16]</sup> oder dem Internationalen Krebsgenomprojekt.<sup>[17]</sup> Wir sind zudem im Begriff, die Früchte solcher Studien zu sehen. Vergleichende Analysen der wachsenden Zahl akkurater, vollständiger Datensätze menschlicher Genome liefern tiefgehende Einsichten in die Evolutionsgeschichte menschlicher Populationen.<sup>[18]</sup> Die Totalsequenzierung von Tumoren liefert detaillierte Kenntnisse der somatischen Mutationen, die zur Entwicklung eines Tumorgenoms bis zu jenem Punkt, an dem Krebs Symptome in Erscheinung treten, beigetragen haben.<sup>[19]</sup> Es gibt auch erste Beispiele dafür, wie durch eine Sequenzierung die Entwicklung des Tumorgenoms eines Krebspatienten zurückverfolgt werden kann, und wie sich diese Informationen für effektivere Therapien zur Tumorstabilisierung nutzen lassen.<sup>[20]</sup> In einem jüngeren Beispiel wurde beschrieben, wie die Totalsequenzierung des Exons (d.h. der proteincodierenden Bereiche) eines Kleinkinds mit einer unbekannten, lebensbedrohenden Krankheit eine Mutation aufdeckte, die eine Diagnose erlaubte, die dann eine lebensrettende Knochenmarkstransplantation nach sich zog.<sup>[21]</sup> Diese und weitere Beispiele bieten uns einen kurzen Ausblick darauf, wie die akkurate Entschlüsselung menschlicher Genome eines Tages unsere Herangehensweise an die Klassifizierung, Diagnose und Behandlung von Krankheiten grundlegend verändern könnte. In der Zwischenzeit wird die fortwährende Erforschung von Anwendungsmöglichkeiten der Totalsequenzierung von Genomen die Vision einer personalisierten, genomischen Medizin immer näher bringen.

Das Aufkommen der routinemäßigen Sequenzierung menschlicher Genome hat klare ethische und soziologische Erwägungen von großer Tragweite nach sich gezogen; darunter sind Angelegenheiten wie Eigentum, Einverständnis, Privatsphäre, Datenschutz, Wahlfreiheit, um nur einige zu nennen. Ethische Grenzsetzungen finden sorgfältige Beachtung und sind in manchen Fällen in den Forschungsaktivitäten implementiert.<sup>[22,23]</sup> Dies wird auch in Zukunft ein zentraler Aspekt der gesellschaftlichen und politischen Debatte bleiben, die unverzichtbar ist, wenn die Sequenzierung menschlicher Genome einmal Teil unseres täglichen Lebens werden soll.

Die Fähigkeit zur schnellen Genomsequenzierung hat auch eine große Bedeutung für die Erforschung anderer Organismen als dem Menschen. Beispiele sind die Sequenzierung der vollständigen Genomsequenzen des Großen Panda (*Ailuropoda melanoleuca*)<sup>[24]</sup> sowie zahlreicher Krankheitserreger<sup>[25]</sup> und Pflanzen.<sup>[26]</sup> Das Humanmikrobiom-Projekt<sup>[27]</sup> hat zum Ziel, durch Tiefensequenzierung mikrobieller Genome die Wechselbeziehungen zwischen menschlicher Gesundheit und Veränderungen in der mikrobiellen Flora des Menschen zu charakterisieren und zu verstehen. Auch in strategisch wichtigen Forschungsbereichen wie der Agrarindustrie und Ernährungssicherung, den Umweltwissenschaften und dem aufkeimenden Gebiet der Bioenergie werden Methoden der Hochgeschwindigkeitssequenzierung zunehmend interessant.

Die Fähigkeit, „kurze“ Leselängen (35–100 Basen) im Hochdurchsatzverfahren zu sequenzieren (d.h. mehrere Milliarden DNA-Fragmente), hat auch der Grundlagenbiologie Perspektiven eröffnet, wie man sie vor einem Jahrzehnt nicht hätte erahnen können. Zu verdanken ist dies zum großen Teil der Kreativität und Vorstellungsgabe der Anwender dieser Technik. Anwendungen von Hochdurchsatzsequenzierungen nutzen die allgemeinen Prinzipien, dass 1) eine kurze Leselänge eines DNA-Fragment einen Sequenzkontext von z.B. 40–100 Basen zu identifizieren oder „markieren“ vermag, und dass 2) es über die Anzahl der Ablesungen, die diesen Sequenzkontext erfassen, gelingt, die Repräsentierung dieses Sequenzkontextes im Vergleich zu anderen digital zu zählen. Häufige Anwendungen sind die genomweite hochauflösende Kartierung von Proteinen (z.B. von Transkriptionsfaktoren oder Chromatinproteinen; ChIPseq),<sup>[28]</sup> die Ermittlung von Cytosinmethylierungsstellen zur Bestimmung des Epigenoms (genomweite Epigenetik)<sup>[29]</sup> und die hochauflösende dreidimensionale Kartierung der Chromatinstruktur in Zellen.<sup>[30]</sup> All diesen Unternehmungen ist gemeinsam, dass sie von der Kapazität der Hochdurchsatz-Sequenzierungsmethoden profitieren, indem sie ein Mittel zur Gewinnung von Ergebnissen zur Verfügung haben, ohne vorab Informationen über die betreffenden Sequenzen besitzen zu müssen. Derartige vorurteilsfreie Ansätze eignen sich auch gut für das Auffinden und die quantitative Profilierung von mRNAs und nichtcodierenden RNAs, wie z.B. Mikro-RNAs und anderen langen nichtcodierenden RNAs. Die Tatsache, dass ein Großteil der ca. 98 % des menschlichen Genoms, das nicht direkt für Proteine codiert (die nichtcodierenden Sequenzen), in RNA transkribiert wird, lässt uns klar erkennen, dass es in Bezug auf das Genom und die Bedeutung all dessen, was in der Gesamtheit seiner Sequenz verschlüsselt ist, noch beträchtlich viel mehr zu entdecken und funktionell aufzuklären gibt.

Die Veränderungen der letzten Jahre bedeuteten einen großen Sprung in der Methodik, der Geschwindigkeit, den Kosten und der Kapazität, mit der DNA sequenziert werden kann. Obgleich die hier beschriebene Methode die gegenwärtig am häufigsten eingesetzte Variante der Hochgeschwindigkeitssequenzierung ist, gibt es noch andere elegante Ansätze,<sup>[10]</sup> die Praxisreife erlangt haben, darunter die ligationsgestützte Sequenzierung an Mikrokügelchen<sup>[31]</sup> oder an DNA-Nanoarrays,<sup>[32]</sup> die Echtzeit-Einzelmolekülsequenzie-



rung<sup>[33]</sup> und ein kürzlich beschriebener protonendetektions-basierter Ansatz.<sup>[34]</sup> Weitere, kreative Sequenzierungskonzepte auf der Grundlage von Nanoporen befinden sich aktuell in der Entwicklung.<sup>[35]</sup> Die weitere Optimierung von Geschwindigkeit und Kosten der Genomsequenzierung werden zu einer wahren „Demokratisierung“ der Hochdurchsatzsequenzierung führen. Der nächste technologische Sprung wird die Entwicklung eines tragbaren Taschengerätes sein, das aus einem Tropfen Blut oder Speichel binnen Minuten das Genom, Transkriptom und Epigenom auslesen kann. Es ist nicht unvorstellbar, dass ein solches Gerät auf einem der bereits existierenden chemischen Sequenzierverfahren beruhen könnte. Ich sage voraus, dass in den nächsten zehn Jahren die Sequenzierung vollständiger Patientengenome Teil eines integrierten Therapiestandards werden wird. Die größten Hürden werden kultureller, ethischer und wirtschaftlicher Natur sein, nicht jedoch technologischer Natur.

Zum Schluss dieses Essays möchte ich hervorheben, dass die grundlegenden Konzepte und Experimente, die der Solexa-Illumina-Sequenzierung zugrundeliegen, neugiergetriebener Forschung entspringen sind,<sup>[36]</sup> die zu unerwarteten Ergebnissen führte, und also keineswegs einem strategischen Plan folgten. Auch ist anzumerken, dass die praktische Umsetzung der Ideen und Methoden in verschiedenste kommerziell verfügbare Systeme den vielen außergewöhnlichen und engagierten Mitarbeitern an der Cambridge University sowie bei Solexa und Illumina zu verdanken ist.

Eingegangen am 21. September 2011

Online veröffentlicht am 5. Dezember 2011

Übersetzt von Dr. Thomas Lazar, Paderborn

- [1] V. A. Erdmann, J. Barciszewski, *Angew. Chem.* **2011**, *123*, 9718–9724; *Angew. Chem. Int. Ed.* **2011**, *50*, 9546–9552.
- [2] J. D. Watson, F. H. Crick, *Nature* **1953**, *171*, 737.
- [3] D. M. Brown, A. R. Todd, *J. Chem. Soc.* **1952**, 52–58.
- [4] A. M. Maxam, W. Gilbert, *Proc. Natl. Acad. Sci. USA* **1977**, *74*, 560.
- [5] F. Sanger, F. Nicklen, A. R. Coulson, *Proc. Natl. Acad. Sci. USA* **1977**, *74*, 5463.
- [6] International Human Genome Project Consortium, *Nature* **2004**, *431*, 931.
- [7] D. Bentley, S. Balasubramanian, H. Swerdlow et al., *Nature* **2008**, *456*, 53.
- [8] S. Balasubramanian, *Chem. Commun.* **2011**, *47*, 7281–7286.
- [9] Es gibt aktuell über 1900 Publikationen, deren Ergebnisse mithilfe des Solexa-Illumina-Sequenzierverfahrens gewonnen wurden. Diese Arbeiten sind unter <http://www.illumina.com/> einsehbar.
- [10] M. L. Metzker, *Nat. Rev. Genet.* **2010**, *11*, 31–46.
- [11] J. H. Jett, R. A. Keller, J. C. Martin, B. L. Marrone, R. K. Moyzis, R. L. Ratliff, N. K. Seitzinger, E. B. Shera, C. C. Stewart, *J. Biomol. Struct. Dyn.* **1989**, *7*, 301–309.
- [12] C. Adessi, G. Matton, G. Ayala, G. Turcatti, J. J. Mermod, P. Mayer, E. Kawashima, *Nucleic Acids Res.* **2000**, *28*, 87e.
- [13] J. Wang, W. Wang, R. Li, *Nature* **2008**, *456*, 60–65.
- [14] T. J. Ley, E. R. Mardis, L. Ding, *Nature* **2008**, *456*, 66–72.
- [15] Dies waren: Dr. David Bentley, Dr. Richard Durbin und Dr. Jane Rogers.
- [16] The International 1000 Genomes Project Consortium, *Nature* **2010**, *467*, 1061.
- [17] The International Cancer Genome Consortium, *Nature* **2010**, *464*, 993.
- [18] H. Li, R. Durbin, *Nature* **2011**, *475*, 493–496.
- [19] E. D. Pleasance et al., *Nature* **2010**, *463*, 191–196.
- [20] S. M. Jones, J. Laskin, Y. Y. Li, *Genome Biol.* **2010**, *11*, R82.
- [21] E. A. Worthey, *Genet. Med.* **2011**, *13*, 255–262.
- [22] A. L. McGuire, T. Caulfield, M. K. Cho, *Nat. Rev. Genet.* **2008**, *9*, 152–156.
- [23] Zu sozialen, ethischen und rechtlichen Aspekten siehe: [http://www.ornl.gov/sci/techresources/Human\\_Genome/elsi/elsi.shtml](http://www.ornl.gov/sci/techresources/Human_Genome/elsi/elsi.shtml).
- [24] R. Li, W. Fan, G. Tian, *Nature* **2010**, *463*, 311–317.
- [25] C. Cunningham, D. Gatherer, B. Hilfrich, K. Baluchova, D. J. Dargan, M. Thomson, P. D. Griffiths, G. W. Wilkinson, T. F. Schulz, A. J. Davison, *J. Gen. Virol.* **2010**, *91*, 605–615.
- [26] V. Shulaev, D. J. Sargent, R. N. Crowhurst, *Nat. Genet.* **2011**, *43*, 109–116.
- [27] <https://commonfund.nih.gov/hmp/>.
- [28] A. Barski, *Cell* **2007**, *129*, 823; D. S. Johnson, A. Mortazavi, R. M. Myers, B. Wold, *Science* **2007**, *316*, 1497.
- [29] R. Lister, R. C. O'Malley, B. D. Gregory, C. C. Berry, A. H. Millar, J. R. Ecker, *Cell* **2008**, *133*, 523.
- [30] E. Lieberman-Aiden, *Science* **2009**, *326*, 289.
- [31] K. J. McKernan et al., *Genome Res.* **2009**, *19*, 1527–1541.
- [32] R. Drmanac et al., *Science* **2010**, *327*, 78–81.
- [33] J. Eid et al., *Science* **2009**, *323*, 133–138.
- [34] J. Rothberg et al., *Nature* **2011**, *475*, 348–352.
- [35] A. Meller, L. Nivon, E. Brandin, J. Golovchenko, D. Branton, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 1079–1084; J. Clarke, H.-C. Wu, L. Jayasinghe, A. Patel, S. Reid, H. Bayley, *Nat. Nanotechnol.* **2009**, *4*, 265–270.
- [36] Ich danke dem Biotechnology and Biological Sciences Research Council (Großbritannien) für die Förderung meiner Grundlagenforschungen, die zum Solexa-Illumina-Sequenzierverfahren führten.