

Timo Bäroth, Sonja Wendenburg

Im Jahre 2141...

Erschöpft und mit pochendem Schmerz im Unterleib sitzt Ute Klopfer vor Dr. med. Hennings. Die Diagnose: Krebs. Eine behandelbare Krankheit, wie der Arzt Klopfer erklärt: „Nachdem wir Ihre DNA sequenziert hatten, konnten wir mit Hilfe des „*Cancer Genome Atlas*“ * feststellen, dass Sie unter Eierstockkrebs leiden. In der Datenbank des „*Personal Genome Project*“ * konnten wir die optimal auf Ihr Genom abgestimmte Behandlung ermitteln.“ Das heißt für Klopfer 6 Wochen lang regelmäßig eine speziell auf sie zugeschnittene Tablette einzunehmen.

Krebs problemlos behandeln können: Eine schöne Vorstellung, die keine Zukunftsmusik bleiben muss. Denn Krebs kann durch verschiedene Gendefekte ausgelöst werden. Können diese schnell ermittelt werden, ist eine effektive Behandlung der Krankheit möglich. Die wichtigste Voraussetzung zur Realisierung einer effektiven Krebstherapie ist folglich die Entwicklung von neuen Methoden für die DNA-Sequenzierung, sogenannte „*Next Generation Sequencing Techniken*“. Das heißt, die Bestimmung der Basenpaarabfolge der DNA (Sequenzierung) mit möglichst geringen Kosten, möglichst schnell und mit einer möglichst geringen Fehlerrate durchzuführen.

Vom Ursprung der DNA-Sequenzierung bis hin zu Zukunftsprojekten

Vater der DNA-Sequenzierung ist der Nobelpreisträger Frederick Sanger, der 1977 das Prinzip der Sequenzierungsmethode publizierte, welche die darauffolgenden 30 Jahre das Feld beherrschte. Forscher bezeichnen sie häufig auch als „First Generation“ der DNA-Sequenzierung.

In den frühen 90er Jahren ersetzen fluoreszenzmarkierte Nukleotide - die Bausteine eines DNA-Strangs - die ursprünglich von Sanger verwendeten radioaktiv markierten Nukleotide. Nebenbei entwickelte sich die Automatisierung des Verfahrens (zu den Details der historischen Entwicklung siehe [1]). Doch die Sequenzierungsmethode nach Sanger konnte mit Beginn des 21. Jahrhunderts immer weniger den stetig steigenden Ansprüchen der Forschung genügen. Mit Kosten von etwa 70 Millionen US-Dollar [2] zur Bestimmung der Basenpaarabfolge eines menschlichen Genoms - unter anderem begründet durch eine Sequenzierungsgeschwindigkeit von nur 24 Basenpaaren pro Sekunde [3] - wären groß angelegte medizinische Projekte nicht realisierbar. Dazu gehören der *Cancer Atlas**, der sich mit dem Zusammenhang zwischen Krebserkrankungen und dem menschlichen Genom befasst, sowie das *Personal Genome Project**, das eine Optimierung der Behandlungsmethoden basierend auf genetischen Informationen zum Ziel hat.

Aufgrund der Kostenfrage begannen verschiedene Firmen, neue Ansätze zur Sequenzierung von DNA zu entwickeln: die „*Next Generation Sequencing Techniken*“ (NGS). Bekannte Vertreter sind die 454-Sequenzierung von Roche oder die Solexa-Sequenzierung von Illumina.

Bereits heute sind die meisten NGS-Techniken mit Kosten von unter 1 Millionen US-Dollar [2] für ein menschliches Genom deutlich günstiger als die herkömmliche Methode nach Sanger. Das angestrebte Ziel der Forscher ist allerdings, die Basenabfolge eines Genoms für nur 1.000 US\$ bestimmen zu können. [2]

DNA-Sequenzierung der nächsten Generation

Alle Methoden der NGS-Techniken lassen sich in drei Arbeitsschritte unterteilen: Probenvorbereitung, Sequenzierung/Datenauswertung und Genomanpassung beziehungsweise -aufbau. Bei der Probenvorbereitung werden die zu untersuchenden DNA-Stränge zunächst zufällig zu kürzeren Abschnitten zerkleinert, da die neuen Techniken noch nicht in der Lage sind, die DNA-Sequenz eines gesamten Genoms am Stück zu ermitteln. Die Länge der sequenzierbaren Fragmente ist abhängig von der verwendeten Methode. Da die meisten Methoden darauf beruhen, dass die Synthese des komplementären DNA-Strangs beobachtet wird, müssen die DNA-Fragmente als Einzelstränge vorliegen (für Details siehe Metzker, 2010 [2]). Anschließend folgen die Sequenzierung und Datenauswertung, deren Vorgehensweise je nach Sequenzierungsmethode variiert. Bei den meisten Methoden wird der Einbau einzelner Nukleotide an den zu sequenzierenden DNA-Strang mittels eines Fluoreszenz- oder Chemolumineszenzsignals beobachtet. Hierzu werden normalerweise Nukleotide verwendet, die nach ihrem Einbau durch ihre chemische Struktur den Einbau weiterer Nukleotide verhindern, bis das Messsignal detektiert und die den Einbau weiterer Nukleotide blockierende Gruppe abgespalten wurde.

Allen Methoden gemeinsam ist, dass die ermittelten Sequenzen der einzelnen DNA-Fragmente wieder zu einem vollständigen Genom zusammengesetzt werden müssen. Da die DNA-Fragmente während der Probenvorbereitung zufällig entstanden sind, gestaltet sich dies schwierig. Die schnellere und einfachere Methode ist die Anpassung an eine bereits bekannte Basenpaarabfolge, sofern eine Referenzsequenz existiert. Die zweite Methode ist deutlich zeitaufwändiger und anfälliger für Fehler, aber immer dann nötig, wenn keine Vergleichssequenz existiert. Hier wird versucht, die einzelnen Fragmente basierend auf überlappenden Sequenzen durch statistische Berechnungen spezieller Computerprogramme in die richtige Reihenfolge zu bringen. [2] (Zu weiterführenden Informationen zu NGS-Techniken siehe Metzker [2] und Mardis [4].)

Fleißiger Helfer: Die DNA-Polymerase

Die *Zero-Mode Waveguide* Sequenzierungstechnik von *Pacific Biosciences* ist eine NGS-Technik, die in Echtzeit mit bis zu 100 Basenpaaren pro Sekunde [5] die Basenabfolge eines DNA-Strangs ermittelt. Während der komplementäre DNA-Strang synthetisiert wird, wird dessen Sequenz zeitgleich detektiert. Dazu bindet der zu sequenzierende DNA-Einzelstrang an ein DNA-Polymerase-Molekül, das sich am Boden einer Kavität eines *Zero-Mode Waveguides* befindet (Abbildung 1 oben links). Diese Nanostruktur besteht aus einer Anordnung von Kavitäten mit etwa 100 nm Durchmesser auf einem Metallfilm, der auf einem transparenten Trägermaterial aufgebracht ist. Unterhalb des transparenten Trägers befindet sich eine Lichtquelle. Das emittierte Licht weist eine zu große Wellenlänge auf, um in die Kavitäten eindringen zu können, sodass die Lichtstärke ausgehend vom Boden ins Innere der Kavität exponentiell abnimmt. So kann das beleuchtete Volumen auf wenige Zeptoliter (10^{-21} l) am Boden der Kavität eingeschränkt werden. [5]

In diesem Bereich befindet sich ein einzelnes DNA-Polymerase-Molekül, das den komplementären DNA-Strang zum Einzelstrangfragment synthetisiert, und die für die DNA-Synthese benötigten Desoxyribonukleosidtriphosphate (dNTPs), die über einen Linker am dritten Phosphatrest fluoreszenzmarkiert sind (Abbildung 2). Jeder der vier verschiedenen dNTP-Typen ist dabei mit einem anderen Fluoreszenzfarbstoff markiert. Beim Einbau eines dNTPs in den DNA-Strang – und nur in diesem Fall – gelangt ein markiertes Molekül für einen längeren Zeitraum in den beleuchteten Bereich der Kavität

und wird zur Fluoreszenz angeregt. Welches markierte dNTP eingebaut wurde, lässt sich anhand der Wellenlänge des detektierten Signals feststellen (Abbildung 1 unten). Während des Einbaus wird das terminale Pyrophosphat einschließlich des konjugierten Farbstoffes abgespalten und diffundiert aus dem beleuchteten Bereich der Kavität heraus. Die Fluoreszenz erlischt. Sobald das nächste dNTP eingebaut wird, kann erneut ein Fluoreszenzsignal detektiert werden. Die Sequenz ergibt sich folglich aus der zeitlichen Abfolge der detektierten Fluoreszenzfarben (Abbildungen 1 oben rechts und Mitte). [5], [6]

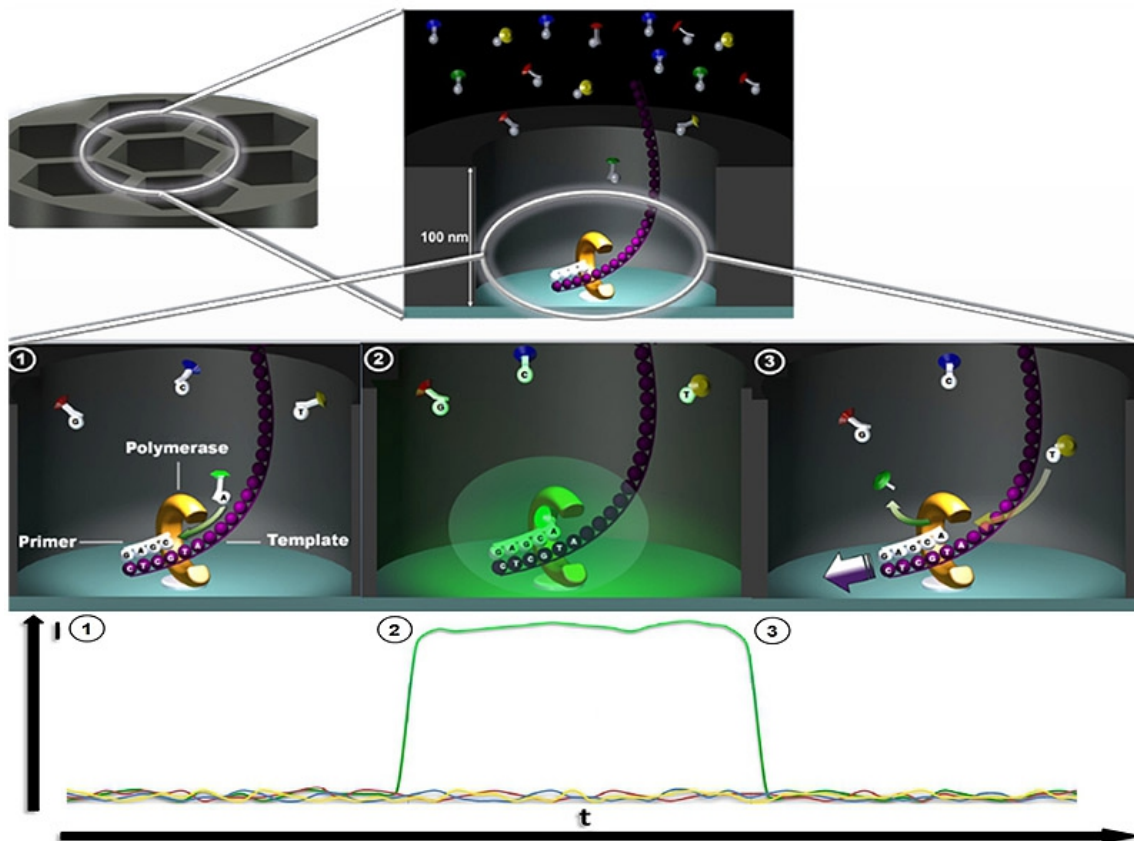


Abbildung 1: Funktionsprinzip der Zero-Mode-Waveguide-Sequenzierung. Oben: Schema der 100 nm-Kavitäten auf dem Zero-Mode-Waveguide und Kavität mit Polymerase und Nucleotiden. Mitte: Synthese eines DNA-Strangs durch immobilisierte Polymerase unter Verwendung fluoreszenzmarkierter Nucleotide. Unten: Resultierendes Fluoreszenzsignal in Echtzeit.

Korlach et al. [5] beschreiben diverse Methoden zur Produktion eines *Zero-Mode Waveguides*. Allen gemeinsam ist, dass durch eine lithographische Methode eine Anordnung von Kavitäten von ca. 100 nm Durchmesser in einem etwa 100 nm dicken Metallfilm (Aluminium oder Gold) entsteht. Dieser befindet sich auf einem transparenten Träger aus Siliziumdioxid. Jede Kavität muss anschließend selektiv mit einer einzelnen Polymerase versehen werden, während die übrige Oberfläche vor unspezifischer Adsorption – beispielsweise der dNTPs – geschützt wird. Um die Wahrscheinlichkeit einer Adsorption zu vermindern, wird die Oberfläche der Kavität mit Polyphosphonaten passiviert. Eine Streptavidin-Biotin-Bindung ermöglicht die gezielte Fixierung der mit Streptavidin konjugierten Polymerase an der selektiv mit Biotin-Polyethylenglykol-Silan funktionalisierten Glasoberfläche (für Details der Methode siehe *Korlach*, 2010 [5]). Die verwendete DNA-Polymerase muss die fluoreszenzmarkierten dNTPs effizient und fehlerfrei einbauen können. Empirisch wurde ermittelt, dass sich $\phi 29$ DNA Polymerase aufgrund ihrer hohen Stabilität, schnellen Einbaugeschwindigkeit, Genauigkeit und hohen Zahl an Katalysezyklen gut eignet. Durch einige Mutationen und die Verwendung eines Linkers konnte erreicht werden, dass diese Polymerase auch in oberflächengebundener

Form die fluoreszenzmarkierten dNTPs genauso gut einbaut, wie eine unmodifizierte Polymerase herkömmliche dNTPs. [5]

Die Messung des Fluoreszenzsignals beruht auf dem Prinzip eines gewöhnlichen Fluoreszenzspektrometers mit kleinen Abwandlungen. Zunächst werden zur Anregung der Fluoreszenz zwei Laser verwendet, um die verschiedenen Anregungswellenlängen der vier Fluoreszenzfarbstoffe abzudecken. Des Weiteren wird das Fluoreszenzsignal vor der Detektion durch ein Prisma in die verschiedenen Spektralbereiche aufgetrennt, sodass alle Emissionsmaxima separat detektiert werden können. Dies ermöglicht während der Auswertung die Zuordnung, welcher dNTP-Typ das Signal verursacht hat. Schließlich wurde ein Aufbau entwickelt, mit dem gleichzeitig jede „einzelne“ Polymerase in den verschiedenen Kavitäten der Anordnung eines *Zero-Mode Waveguides* untersucht werden kann. Diese Möglichkeit der Parallelisierung der Messung zeigt, dass das Verfahren für den Hochdurchsatz geeignet ist (für Details zum Aufbau siehe [5]).

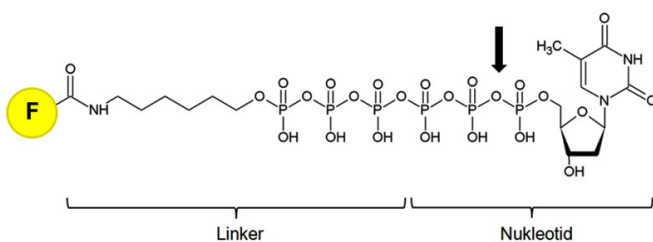


Abbildung 2: Struktur eines fluoreszenzmarkierten Nucleotids. Der Fluorophor (gelb) ist über einen Trophosphatlinker an das Nucleotid (hier Thymidin) gebunden. Der Pfeil markiert die Stelle, an der die Polymerase die Oligophosphatkette spaltet.

Die *Zero-Mode Waveguide* Sequenzierungstechnik weist gegenüber der herkömmlichen Sanger-Methode, aber vor allem auch gegenüber den anderen NGS-Techniken einige entscheidende Vorteile auf: Es können einzelne DNA-Fragmente in Echtzeit - also mit einer Umsatzgeschwindigkeit von bis zu 100 Basen pro Sekunde - sequenziert werden; die Polymerase ermöglicht die Sequenzierung relativ langer Fragmente; und es entsteht dabei ein natürliches DNA-Molekül, das für weitere Analysen zur Verfügung steht. Diese Vorteile

ermöglichen zuvor undenkbare Anwendungen, wie beispielsweise die Detektion methylierter DNA. [7]

Im Jahre 2141... 6 Wochen später....

Wieder ist Ute Klopfer bei Dr. med. Hennings. Keine zwei Monate ist es her, dass sie hier vor Schmerzen kaum aufrecht sitzen konnte. Doch nun fühlt sie sich gut und gesund, was der Arzt ihr auch bestätigt: „Sie sind geheilt!“ Durch die Entwicklung der *Next Generation DNA-Sequencing* Techniken – besonders der *Zero-Mode Waveguide* Sequenzierungstechnik – ist bereits ein erster Schritt zur Entwicklung neuer medizinischer Diagnosemethoden, die Geschichten, wie die von Ute Klopfer, wahr werden lassen könnten, gemacht worden. Doch dazu muss den Forschern zunächst die Erhöhung der Sequenzierungsgeschwindigkeit bei gleichzeitiger Reduzierung der Kosten von derzeit 1 Millionen auf maximal 1000US\$ pro Genom gelingen.

Auf einen Blick:

- Wichtige Projekte zur Verbesserung der medizinischen Diagnostik erfordern eine weitere Reduzierung der Kosten und des Zeitaufwands von DNA-Sequenzierung im Vergleich zu aktuell angewendeten Methoden.
- Verschiedene „*Next Generation Sequencing Techniken*“ wie die 454-Sequenzierung von Roche oder die Solexa-Sequenzierung von Illumina werden zurzeit entwickelt.
- Bei der *Zero-Mode Waveguide* Sequenzierungsmethode erfolgt die Sequenzierung in Echtzeit. Die Sequenz wird durch die Beobachtung der Synthese eines

komplementären DNA-Strangs aus fluoreszenzmarkierten Nukleotiden durch eine DNA-Polymerase mittels Fluoreszenzspektroskopie ermittelt.

- Besonders vorteilhaft an der ZMW Sequenzierungstechnik im Vergleich zu anderen Techniken ist die Sequenzierung verhältnismäßig langer DNA-Fragmente in kurzer Zeit.

Kontakt:	Schlauer Fuchs
 <p>Der Artikel wurde im Rahmen des Studienprojektes HighChem des Fachbereichs Biochemie der Ernst-Moritz-Arndt-Universität Greifswald verfasst (s. Woche 2).</p> <p>Die Autoren, Tima Bäröth und Sonja Wendenburg, waren zum Zeitpunkt des Verfassens Studierende des Masterstudiengangstudiengang Chemie an der TU Darmstadt.</p> <p>Wissenschaftliche Betreuung: Prof. Dr. Katja Schmitz (E-Mail: schmitz@biochemie-tud.de).</p> 	<p>Unsere Schlaue-Fuchs-Frage zu diesem Beitrag lautete:</p> <p>Welche drei Arbeitsschritte müssen bei allen Methoden der NGS-Techniken durchgeführt?</p>
Literatur:	
<p>[1] C. A. Hutchison, 3rd, <i>Nucleic acids research</i> 2007, 35, 6227-6237.</p>	
<p>[2] M. L. Metzker, <i>Nature reviews. Genetics</i> 2010, 11, 31-46.</p>	
<p>[3] E. Hitt, <i>Genomics and Proteomics</i> 2005, 5, 17-21.</p>	
<p>[4] E. R. Mardis, <i>Annual review of genomics and human genetics</i> 2008, 9, 387-402.</p>	
<p>[5] J. Korlach, K. P. Bjornson, B. P. Chaudhuri, R. L. Cicero, B. A. Flusberg, J. J. Gray, D. Holden, R. Saxena, J. Wegener, S. W. Turner, <i>Methods in enzymology</i> 2010, 472, 431-455.</p>	