

Kerstin Bathon, Ines Kühnel, und Astrid Schönberger

4 Millionen mal um den Äquator, 400.000mal von der Erde zum Mond, 1000mal von der Erde zur Sonne oder 25mal von der Sonne zum Pluto. Diese Strecke, 150 Milliarden Kilometer, ergibt sich, wenn man alle DNA Moleküle des menschlichen Körpers ausgestreckt aneinanderlegen würde. Jede einzelne Zelle enthält etwa 2 Meter DNA auf denen 3 Milliarden Basenpaare liegen, die etwa 23.000 Gene enthalten. [1] In der Computersprache ausgedrückt, entspricht das zwar nur einer Speicherkapazität von 750 MB, aber in Anbetracht der Länge der DNA ist es nicht schwer vorstellbar, dass die Entschlüsselung des menschlichen Erbguts, des Genoms, sehr zeitaufwendig und teuer ist. Heutzutage kostet die Sequenzierung eines Genoms etwa 18.000 US-Dollar. [2] Um diesen Preis zu senken, forschen viele Arbeitskreise und Firmen an neuen Methoden. Das „1000 US-Dollar Genom Projekt“ wurde ins Leben gerufen, dessen erklärtes Ziel die Sequenzierung eines Genoms für unter 1000 US-Dollar ist. Seitdem herrscht ein regelrechtes Rennen darum, wer am schnellsten, günstigsten und einfachsten die Basenabfolge eines Genoms ermitteln kann. Dabei stehen sich meist Behauptungen über die Leistungsfähigkeit und tatsächlich wissenschaftlich publizierte Werte gegenüber. Gerade bei der Zahl der Basen, die angeblich am Stück sequenziert werden können, wird zunächst nach den Sternen gegriffen, dann jedoch wieder zu realistischeren Einschätzungen zurückgekehrt. Die bekannteste Methode zur DNA-Sequenzierung ist die 1975 von Sanger erfundene Kettenabbruchreaktion. [3] Ein wesentlicher Nachteil dieser Methode ist jedoch, dass nur ca. 1000 Basenpaare am Stück sequenziert werden können. Das bedeutet, die DNA muss in kleinere Fragmente zerteilt, sequenziert und anschließend die einzelnen Sequenzabschnitte wie ein Puzzle wieder zusammengesetzt werden. Ähnlich einem Puzzle mit mehreren gleichen Teilen gibt es jedoch verschiedene Möglichkeiten der Anordnung, weshalb Fehler entstehen können. Um sichere Aussagen zu erhalten, muss das Genom mehrfach in unterschiedliche Fragmente zerteilt und diese sequenziert werden. Mit dem Sanger-Verfahren ist dieses Vorgehen jedoch sehr zeitaufwändig. Daher wurde eine große Anzahl sogenannter „Next-Generation-Sequencing-Methoden“ auf den Markt gebracht. Wesentliches Ziel ist die Erfassung langer Sequenzen, wodurch das fehleranfällige Puzzle-Spiel unnötig wird. Eine Idee hierzu ist die von Branton und Deamer, Nanoporen - winzig kleine Kanäle, die durch eine Membran führen - zur DNA-Sequenzierung zu nutzen. Auf Grundlage dieser Idee wurde ein neues Gerät entwickelt, das angeblich bis zu 100 Kilobasen (kb) am Stück sequenzieren kann. Das 5,4 kb große Genom eines Virus soll es bereits in einem Lauf sequenziert haben. [4] Dem gegenüber stehen die bisher tatsächlich publizierten Werte mit nur etwa 1500 sequenzierten Basenpaaren. [5] Die dennoch sehr vielversprechende Methode der Sequenzierung mit Nanoporen wird im Folgenden dargestellt.

Nanoporen sind in diesem Fall kleine Kanalproteine, die biologische Membranen durchspannen. Durch solche Kanäle fließen Cl^- und K^+ Ionen und erzeugen dadurch einen Ionenstrom, der mit einem so genannten Patch-Clamp-Verstärker gemessen werden kann (Abbildung 1). Befindet sich jedoch einzelsträngige DNA (ssDNA) im Kanal, ist dieser enger, wodurch der Ionenstrom von etwa 50 pA auf etwa 30 pA verringert wird. Die Grundlage der Sequenzierung ist nun, dass die vier Basen, also die Bausteine der DNA,

unterschiedliche Größen und Formen aufweisen. Dadurch wird der Ionenstrom unterschiedlich stark verändert, was die Sequenzierung der DNA während ihrer Translokation durch die Nanopore ermöglicht. [6]

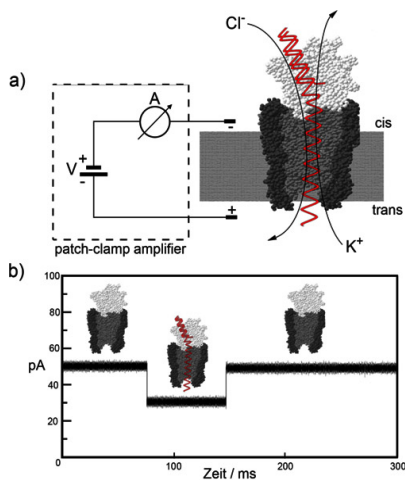


Abbildung 1: a) Aufbau des Patch-Clamp Verstärkers zur Messung des Ionenstroms. b) Schematische Darstellung einer Ionenstrommessung mit und ohne ssDNA. [6]

Bisher werden nur zwei verschiedene Poren zur Sequenzierung von DNA und RNA verwendet (Abbildung 2). Dabei handelt es sich zum einen um das Staphylokokken-Protein α -Hämolyisin, das als Kanal in die Lipidmembran von roten Blutzellen integriert wird. [6] Durch diesen Kanal können Ionen sowie größere Moleküle die Membran passieren. Die α -Hämolyisin-Pore hat eine pilzförmige Struktur und eine Engstelle mit einem Durchmesser von 1,4 nm, durch die nur ssDNA den Kanal passieren kann. Diese erzeugt ein spezifisches Signal, das hoch genug ist, um einzelne Basenverschiebungen zu beobachten. Allerdings beeinflussen alle Nukleotide, die sich im Kanal befinden, das Signal. Bei α -Hämolyisin befinden sich etwa zehn Basen gleichzeitig im Kanal, sodass die Signale nicht eindeutig zugeordnet werden können. [6]

Auch das Kanalprotein *Mycobacterium smegmatis* Porin A (MspA) kann zur Sequenzierung verwendet werden. Der Wildtyp von MspA trägt im Kanal negativ geladene Aminosäuren, die ersetzt werden müssen, da es sonst zu repulsiven Wechselwirkungen mit der ebenfalls negativ geladenen DNA kommt. Daher wird das Protein modifiziert, sodass sich positive Ladungen am Poreneingang befinden, die die DNA anziehen. [7] Der Kanal von MspA ist mit 1,3 nm schmäler und deutlich kürzer als der von α -Hämolyisin, sodass sich nur drei bis vier Basen gleichzeitig in der Pore befinden können. Damit ergibt sich eine zehnmal bessere Auflösung der Nukleotidsignale. Nicht nur die vier Nukleotide in der Pore beeinflussen den Ionenstrom, sondern auch die Sekundärstruktur der DNA, Wechselwirkungen der Nukleotide untereinander und mit der Pore sowie die exakte Positionierung der DNA im Kanal. Ein genereller Nachteil von Kanalproteinen ist die Auflösung von homopolymeren Sequenzen, wie z.B. einem Poly-A-Strang. [8] Ein passender Algorithmus, der dies differenzieren könnte, muss noch entwickelt werden. Ein weiteres Problem ist die Parallelisierung der Poresequenzierung, das heißt die gleichzeitige Sequenzierung vieler DNA-Stränge.

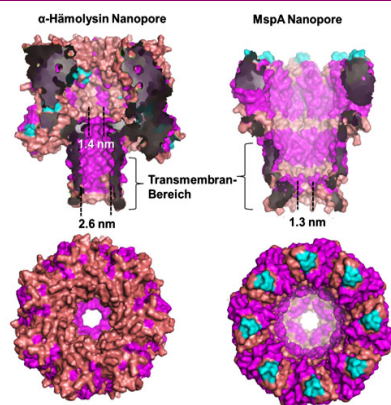


Abbildung 2: Die Nanoporen α -Hämolyisin (links, PDB-Nr.: 7AHL) und MspA (rechts, PDB-Nr.: 1UUN). Oben: Schnitt durch die Poren, unten: Aufsicht.

Zur Parallelisierung eignen sich synthetische Poren in Festkörpern besser als Proteinporen. Diese können basierend auf

einer Vielzahl verschiedener Materialien wie Siliziumwafer, Graphen und Carbon-Nanoröhren hergestellt werden. Grundlage der Sequenzierung mit Festphasen basierten Poren ist der Effekt des Elektronentunnels, der auftritt, wenn sich zwei Nanoelektroden sehr dicht gegenüber stehen. Befindet sich ein Nukleotid zwischen den Elektroden wird der durch das Elektronentunneln entstehende Strom nukleotidspezifisch verändert. Die Bewegung der DNA wird dabei durch das An- und Ausschalten eines elektrischen Feldes gesteuert. [2] IBM und Roche forschen zurzeit an einer Pore, die durch einen Chip aus alternierenden Metall- und Isolatorschichten führt. Dabei werden die Nanoelektroden in der Metallschicht mit organischen Molekülen modifiziert, die Wasserstoffbrückenbindungen zu den Nukleotiden ausbilden und so die DNA kurzzeitig in der Pore festhalten. Bisher ist Festphasen basiert lediglich die Sequenzierung von 20 Basenpaaren RNA (Toshiba)

möglich. Mit Proteinporen hingegen können bereits 1500 Basenpaare am Stück sequenziert werden. [5]

Ein wesentliches Problem bei der Sequenzierung von DNA mittels Nanoporen ist die zu rasche Durchflussrate mit durchschnittlich mehr als einem Nukleotid pro Mikrosekunde für MspA [8] und α Hämolyisin [6]. Dies ist ungefähr 1000mal zu schnell, um das basenspezifische Signal vom Hintergrundrauschen unterscheiden zu können. Zur Lösung dieses Problems gibt es verschiedene Ansätze. Naheliegende Methoden wie die Erniedrigung der Temperatur oder Erhöhung der Viskosität der Lösung verlangsamen zwar die Translokation, verbessern jedoch nicht die Signalauflösung, da sie zu einer Erniedrigung des Ionenstroms selbst führen. [7] Ansatzpunkt ist also ausschließlich, den Durchgang der DNA durch die Pore zu verlangsamen. Eine Möglichkeit besteht darin, DNA prozessierende Enzyme wie Polymerasen und Exonukleasen an die Pore zu koppeln. Polymerasen synthetisieren die DNA schrittweise entlang eines Templat-Strangs. Pro enzymatischen Zyklus wird ein Nukleotid eingebaut. Währenddessen verharrt eine bestimmte Sequenz der DNA in der Engstelle der Pore und sorgt dort einige Millisekunden lang für ein nukleotidspezifisches Signal. Die Geschwindigkeit der Translokation wird also von der Prozessierungsrate des Enzyms bestimmt. [6] Für den Einsatz bei der Nanoporen-Sequenzierung wurden die DNA-Polymerase I aus *E.coli* [6] und die aus den Bakteriophagen T7 und phi29 untersucht. [8] Vor allem die Polymerase aus phi29 (phi29 DNAP) erweist sich hierbei als gut geeignet, da sie in der Lage ist, sehr lange DNA-Fragmente (>70 kb) zu synthetisieren. [8]

Ein weiterer Ansatzpunkt zur Kontrolle der Translokation ist die Einführung doppelsträngiger DNA-Abschnitte, die im Vorhof der Pore stecken bleiben, während das einzelsträngige Ende in der Engstelle den Ionenstrom kontrolliert und für ein spezifisches Signal sorgt. Um so die Sequenz einer unbekanntes DNA zu ermitteln, müssten die einzelnen Nukleotide jeweils durch einen kurzen doppelsträngigen Bereich getrennt werden. Dies ist zwar theoretisch möglich, jedoch mit hohem Aufwand verbunden. [7] Einen ähnlichen Effekt haben Oligomere, die mit der DNA zum Doppelstrang hybridisiert werden. Bei einer Translokation der DNA wird dieses Oligomer dann zunächst mechanisch abgespalten, was den gesamten Prozess verlangsamt. [8] Meist werden Oligomere jedoch eingesetzt, um den Primer-Strang, der für die Synthese der DNA mit Polymerasen notwendig ist, zu blockieren. Dadurch wird eine Kontrolle der enzymatischen Prozessierung möglich. Auch die Variation der angelegten Spannung verursacht eine reproduzierbare Veränderung in der Durchflussrate und ist damit potentiell geeignet um die DNA-Translokation zu steuern. [6]

Die Kombination der MspA Nanopore mit phi29 DNAP, ist eine vielversprechende Methode. Hierbei wird das DNA-Templat mit einem Primer und einem blockierenden Oligomer zu einer Hairpin-Struktur verknüpft. Durch Anlegen einer Spannung wird das einzelsträngige Ende durch die Engstelle der Pore gezogen und erzeugt sequenzspezifische Signale, während das Oligomer unter Freilegung des Primers abgespalten wird. Dadurch kann auf der cis-Seite der Pore die DNA-Synthese durch phi29 DNAP beginnen. Durch den Einbau von Nukleotiden wird die DNA dabei wieder zurückgezogen und die Sequenz erneut abgelesen. Mit dieser Methode ist es möglich, einzelne Nukleotide anhand des verursachten Ionenstroms zu unterscheiden. [8]

Insgesamt ist die Methode der DNA-Sequenzierung an einer Pore ein vielversprechender Ansatz zur schnellen und günstigen Analyse. Mit Hilfe der Pore MspA und Polymerasen wie z.B. phi29 DNAP kann die Translokation der DNA kontrolliert und somit basenspezifische Signale erzeugt werden. Gegenüber anderen Methoden ist der wohl größte Vorteil, dass lange DNA-Abschnitte (möglicherweise bis zu 100 Kilobasen [4]) am

Stück sequenziert werden können, was die Fehlerrate enorm verringert. Außerdem ist die Nanoporen-Sequenzierung mit einer minimalen Probenvorbereitung und einem vergleichsweise geringen finanziellen Aufwand verbunden. Die Erkennung homopolymerer Regionen ist allerdings weiterhin ein Problem. In den letzten Jahren fand im Bereich der Nanoporen-Sequenzierung eine rasante Entwicklung statt. So will Oxford Nanopore Technologies Ende 2013 ein miniaturisiertes Gerät (MinION) basierend auf Nanoporen für nur etwa 900 US-Dollar auf den Markt bringen. [4] Damit würde diese Methode wohl das Rennen um das 1000-US-Dollar Genom gewinnen.

Take-Home-Messages:

- Genomsequenzierung mit Nanoporen hat das Potential für die schnelle und kostengünstige Sequenzierung sehr langer DNA-Abschnitte (möglicherweise bis zu 100 Kilobasen).
- Grundlage ist die Messung des Ionenstroms durch eine Pore, der sich beim Durchtritt einzelsträngiger DNA basenspezifisch verändert.
- Es können verschiedene Poren (MspA, α -Hämolysin) zur Sequenzierung verwendet werden.
- Die Geschwindigkeit der Translokation der DNA und damit die Ablesegenauigkeit kann durch Polymerasen, Veränderung der Spannung oder durch kurze eingefügte Stücke doppelsträngiger DNA kontrolliert werden.

Kontakt:	Schlauer Fuchs
<div data-bbox="172 215 411 517">  </div> <div data-bbox="172 517 411 819">  </div> <p data-bbox="432 215 963 338">Der Artikel wurde im Rahmen des Studienprojektes HighChem des Fachbereichs Chemie der Technischen Universität Darmstadt verfasst (s. Woche 2).</p> <p data-bbox="432 349 963 472">Die Autorinnen, Kerstin Bathon, Ines Kühnel, und Astrid Schönberger, sind Studierende des Masterstudiengangs Chemie an der TU Darmstadt</p> <p data-bbox="432 483 963 584">(E-Mail: kerstin.bathon@web.de, ines.kuehnel@stud.tu-darmstadt.de und astrid.schoenberger@stud.tu-darmstadt.de.)</p> <p data-bbox="432 618 963 719">Wissenschaftliche Betreuung: Prof. Dr. Katja Schmitz (E-Mail: schmitz@biochemie-tud.de).</p>	<p data-bbox="975 215 1369 271">Unsere Schlaue-Fuchs-Frage zu diesem Beitrag lautete:</p> <p data-bbox="975 327 1437 383">Von wem stammt die Idee, Nanoporen zur DNA-Sequenzierung</p>
Literatur:	
<p data-bbox="158 931 1437 992">[1] A. B. Daniel Lingenhöhl, http://www.spektrum.de/alias/welche-strecke-liesse-sich-mit-der-gesamten-dna-eines-menschlichen-koerpers-ueberbruecken-1000-mal/692182, 17.08.2012.</p>	
<p data-bbox="158 1010 676 1043">[2] E. Pennisi, Science 2012, 336, 534-537.</p>	
<p data-bbox="158 1066 1437 1126">[3] F. Sanger, S. Nicklen, A. R. Coulson, Proceedings of the National Academy of Sciences of the United States of America 1977, 74, 5463-5467.</p>	
<p data-bbox="158 1144 603 1178">[4] E. C. Hayden, Nature News 2012.</p>	
<p data-bbox="158 1200 611 1234">[5] H. Ledford, Nature 2011, 470, 155.</p>	
<p data-bbox="158 1256 874 1290">[6] D. Deamer, Annual review of biophysics 2010, 39, 79-90.</p>	
<p data-bbox="158 1312 1437 1373">[7] I. M. Derrington, T. Z. Butler, M. D. Collins et al., Proceedings of the National Academy of Sciences of the United States of America 2010, 107, 16060-16065.</p>	
<p data-bbox="158 1391 1278 1424">[8] E. A. Manrao, I. M. Derrington, A. H. Laszlo, et al., Nature biotechnology 2012, 30, 349-353.</p>	