

# 1. Einleitung

Bis Anfang 1970 war die Desoxyribonukleinsäure (DNA), das am schwersten zu analysierende zelluläre Molekül. Durch ihre enorme Länge und ihren monotonen chemischen Aufbau konnte man sich ihrer Primärstruktur (Sequenz) nur auf indirektem Wege, wie zum Beispiel über genetische Analyse, RNA- und Protein-Sequenzanalyse nähern. Heute hat sich die Situation vollständig geändert. Die DNA wurde zum am leichtesten zu analysierenden Makromolekül der Zelle. Momentan befindet sich die Biologie und Biomedizin an einem Wendepunkt, der durch das massive Wachstum an DNA-Sequenzinformation und die Entwicklung von Techniken, die diese verarbeiten, bestimmt wird.

*Die zeitliche Entwicklung der DNA-Analyse lässt sich in den wesentlichen Schritten wie folgt zusammenfassen:*

- 1944 DNA Träger der Erbsubstanz (Avery *et al.*, 1944)
- 1953 Entschlüsselung der Struktur der DNA:  
Die Doppelhelix (Watson, J.D., and Crick, F.H.C, 1953)
- 1975 Erste direkte DNA-Sequenzierung (Sanger and Coulson, 1975)
- 1998 Bäcker-Hefe (*Saccaromyces cerevisiae*) vollständig sequenziert
- 2000 Erstes pflanzliches Genom vollständig sequenziert: *Arabidopsis thaliana*
- 2001 Menschliches Genom „vollständig“ sequenziert

## 1.1 DNA-Sequenzierverfahren

Das erste direkte DNA-Sequenzier- oder DNA-Sequenzanalyse-Verfahren, die enzymatische Sequenzierung, wurde 1975 von Fred Sanger vorgestellt (Sanger and Coulson, 1975) und 1977 durch das enzymatische Kettenabbruchsverfahren verfeinert (Sanger, 1977). Das heute am meisten verwendete Verfahren der DNA-Sequenzierung, das "Cycle-Sequencing", ist eine Kombination aus dem enzymatischen Kettenabbruchsverfahren nach Sanger und der PCR Technik, die hitzbeständige Enzyme für die DNA-Amplifikation verwendet (siehe Abschnitt 2, Grundlagen). Sanger verwendete noch radioaktiv markierte DNA und belichtete nach der Elektrophorese mit dem Polyacrylamidgel einen Film. Dieser Film wurde dann manuell

ausgelesen. Die Elektrophorese ermöglicht die Trennung der DNA-Moleküle und damit den Erhalt der Sequenz-Information. Als Elektrophorese bezeichnet man eine Technik zur Trennung geladener, kolloidaler Teilchen auf der Grundlage ihrer unterschiedlichen Beweglichkeit in einem elektrischen Feld. Die Elektrophorese wurde 1809 von dem russischen Physiker Reuss (Reuss *et al.*, 1809) entdeckt. Diese heute sehr verbreitete Methode erlaubt unter anderem die Separation makromolekularer Ionen, Viren, biologischer Zellen, submolekularer Organellen, Aminosäuren und Nukleinsäuren.

Die Trennung von DNA-Molekülen muss in einem unterstützenden Medium (Agarose, Papier oder Polyacrylamid) geschehen, da DNA-Moleküle unterschiedlicher Größe bei der Elektrophorese in freier Lösung identische Beweglichkeiten aufweisen.

## 1.2 Fluoreszente DNA-Sequenzierung

Die heute am meisten verbreiteten DNA-Sequenziersysteme basieren auf laserinduzierter Emission fluoreszent markierter DNA nach einer vorgegebenen Trennstrecke in einem Polyacrylamidgel. Die Sequenzinformation ergibt sich im Gegensatz zur ursprünglichen Methode mit radioaktiv markierter DNA nicht aus der räumlichen Trennung, sondern der zeitlichen Differenz der Laufzeiten, der unterschiedlich langen DNA-Fragmente bis zum Laserstrahl.

Man unterscheidet hier nochmals zwischen Flachbettgel- und Kapillar-Systemen zur elektrophoretischen Trennung. Bei den Flachbettgel-Systemen befindet sich das Trenngel zwischen zwei durch Platzhalter (200-300 µm) getrennten Glasplatten, bei den Kapillar-Systemen in Glasröhrchen mit geringem Innendurchmesser (100-150µm). Die Anfänge der fluoreszenten DNA-Sequenzierung gehen auf Ansorge (Ansorge *et al.*, 1986) und Smith (Smith *et al.*, 1986) zurück, ein Jahr nach der ersten erfolgreichen Synthese eines fluoreszent markierten Primers durch Smith (Smith *et al.* 1985).

Die großen Vorteile der fluoreszenten gegenüber der radioaktiven DNA-Sequenzierung sind:

- a) Der Verzicht auf die gesundheitsschädlichen und eine besondere Entsorgung erfordernden Radionuklide (z.B.  $^{32}\text{P}$ ).
- b) Die Detektion durch den Laser gibt auch die Trennstrecke im Gel wieder. Der Großteil aller Moleküle migriert eine längere Strecke und ist deshalb besser getrennt.

- c) Die Daten werden „online“ erfasst und können deshalb durch geeignete Algorithmen automatisch in Sequenzinformationen umgewandelt werden. Dies vereinfacht die sich anschließende Sequenz-Analyse.

Das ursprüngliche System von Smith ist ein Flachbettgel-Verfahren und bedient sich 4 unterschiedlicher fluoreszenter Farbstoffe für die 4 Basen einer Reaktion. Durch die Verwendung von 4 unterschiedlichen Farbstoffen (Fluoreszein, NBD, Tetramethyl-Rhodamin und Texas-Red) für eine Reaktion können alle 4 Basen einer Reaktion in einer Spur laufen. Allerdings ist die automatische Bestimmung der Sequenz auf Grund der unterschiedlichen Mobilitäten der Farbstoffe im Gel relativ fehleranfällig. Zur Beobachtung des Geles über seine ganze Breite wird ein Scanning-Mechanismus verwendet. Die unterschiedlichen Absorptionsspektren der Farbstoffe erfordern die Anregung mit zwei unterschiedlichen Wellenlängen. Dies wurde durch einen Argon Laser im Multi-Line Modus realisiert. Zur Vermeidung der spektralen Überlappung des Emissionslichts wurde ein Photomultiplier hinter ein rotierendes Filtrerrad gesetzt.

Um die Nachteile der unterschiedlichen Mobilitäten der Farbstoffe und des die Sensitivität verringernden Filtrerrades zu vermeiden wurden folgende Modifikationen vorgenommen:

- 1) Anstatt des Filtrerrades wurde eine CCD-Kamera eingesetzt, die das vierfache Abfahren eines jeden Peaks vermeidet.
- 2) Durch das Einführen modifizierter Farbstoffe konnten die Mobilitätsunterschiede der einzelnen Basen minimiert werden (Energy transfer Primer).

Aufbauend auf diesem Verfahren wird das System „ABI PRISM“ 377 mit einer Vielzahl von Modifikationen von der Firma Applied Biosystems vertrieben.

Das ursprüngliche System von Ansorge ist ein Flachbettgel-System, bei dem alle 4 Basen der DNA mit dem gleichen Farbstoff markiert sind und in unterschiedlichen Spuren im Gel laufen. Der Laserstrahl eines Argon Lasers wird seitlich in das Trenngel eingekoppelt. Nach einer bestimmten Strecke im Trenngel wird der Farbstoff Fluoreszein der markierten DNA durch den Laserstrahl angeregt. Das Emissionslicht wird durch Photomultiplier in elektrische Signale umgewandelt, die dann über einen

analog-digital Konverter auf einem Computer abgespeichert und ausgewertet werden. Den einzelnen Multipliern sind die einzelnen Basen der unterschiedlichen Klone fest zugeordnet.

Auf diesem Prinzip der Ein-Farbstoff-4-Spur Analyse pro Probe basieren die kommerziellen Systeme der Firmen Amersham-Pharmacia (ALF, ALF-Express, 10 Proben pro Lauf) und Li-Cor (Li-Cor, 32 Proben pro Lauf). Der Li-Cor DNA Sequenzer besitzt im Unterschied zum ursprünglichen Ansorge System ein über die Gelbreite abtastendes Anregungs- und Detektionssystem.

Die grundlegenden Weiterentwicklungen des Ansorge Systems sind:

- 1) Der Übergang von einem Einfarb- zu einem Fünffarbsystem. Das heißt, bei gleichbleibender Gelddimension eine Verfünfachung der Kapazität.
- 2) Die Möglichkeit von simultanen Sequenzierungsreaktionen für alle fünf Farbstoffe. Das heißt, eine Verminderung der Sequenzierungskosten.
- 3) Die separaten Photomultiplier oder später Photodioden wurden durch ein durchgehendes Photodiodenarray ersetzt. Der Einsatz eines Photodiodenarrays, mit 0,8mm Pixelweite, erlaubt theoretisch das Auftragen von 48 Proben pro Farbstoff und 305 mm Gelbreite (Ventzki, 2000); in der Praxis haben sich 36 Proben durchgesetzt. Das heißt, bei 5 unterschiedlichen Farbstoffen theoretisch 240 Proben (praktisch 180 Proben) pro Lauf. Der Nachteil ist, dass den einzelnen Spuren keine festen Photodioden des Arrays zugeordnet werden können. Deshalb ist eine Nachbearbeitung des Laufs notwendig. Diese geschieht zwar zum Großteil über eine spezielle Software (Tracking software); eine manuelle Kontrolle ist aber unerlässlich.

Die Sequenziergeschwindigkeiten hängen von den Vorgaben an die Qualität und Leselänge der Resultate ab. Leselängen von über 1000 Basen pro Probe werden in ungefähr 20 Stunden erreicht. Solche von knapp 600 Basen können mit sehr dünnen Gelen (0,1 mm) schon in einer Stunde Laufzeit erreicht werden (Ansorge, W., and de Maeyer, L. 1980, Stegemann *et al*, 1991).

Der Einsatz der letzten Photodiodenarray-Generation mit 0,2 mm Pixelbreite sollte theoretisch eine Vervierfachung der Kapazität bringen und wird in dieser Arbeit beschrieben.

Kapillar-Sequenzierereinheiten:

Ein Verfahren, das anstelle der Flachgele, Kapillaren zur elektrophoretischen Trennung verwendet. Die zwei meistverarbeiteten werden im Folgenden vorgestellt.

Die „ABI-PRISM 3700“ Kapillar-Sequenzierereinheit der Firma „Applied Biosystems“ verfügt über 96 Kapillaren. Die Umlaufzeit beträgt 2,5 Stunden, bei einer durchschnittlichen Leselänge von 550 Basen. Die Detektion der fluoreszent markierten DNA-Moleküle findet nach dem Austritt aus der Kapillare statt. Durch dieses Detektionsverfahren, das den fluoreszenten Untergrund des Gels und des Glases umgeht konnte das Signal zu Rauschen-Verhältnis um den Faktor 10 verbessert werden. Das System verwendet 2-dimensionale CCD-Kameras und gestattet den Einsatz von bis zu 5 unterschiedlichen Farbstoffen. Der Einsatz von linearen Polymeren erlaubt den automatischen Austausch der Trennmatrix. Das Beladen der Kapillaren mit DNA erfolgt durch elektrokinetische Injektion.

Der „MegaBACE“ der Firma „Amersham Pharmacia“ ist eine Sequenzierereinheit mit 96 Kapillaren. Die Umlaufzeit beträgt 2 Stunden mit maximal 800 lesbaren Basen, bei einer durchschnittlichen Leselänge von mehr als 550 Basen. Der Austausch der Trennmatrix (lineares Polymer) und die Beladung der Kapillaren mit DNA (elektrokinetische Injektion) erfolgt automatisch. Die Lebensdauer der Kapillaren beträgt ungefähr 130 Läufe.

Vorteile der Kapillarsequenzierereinheiten gegenüber den Flachbettgel-Sequenzierereinheiten sind ihre hohe Automatisierbarkeit, die hohe Elektrophorese-Geschwindigkeit und der Wegfall des Trackings. Nachteile sind geringere Leseweiten und höhere Sequenzierkosten.

### 1.3 Beladung von Trenngelen mit DNA

Um die DNA-Moleküle einer Lösung in einem Gel der Größe nach zu trennen, müssen sie vor dem Start der Elektrophorese entweder in das Gel eingebracht oder mit dem Gel in Kontakt gebracht werden. Diesen Vorgang bezeichnet man als Beladung von Trenngelen mit DNA.

Kapillar-Sequenziersysteme werden mit der Methode der elektrokinetischen Injektion beladen. Die DNA wird direkt aus Mikrotiter-Platten durch Anlegen eines elektrischen Feldes in die Kapillaren transferiert. Das kleinste Probenvolumen aus dem die gelöste DNA aufgenommen werden kann ist 5 µl, was die Miniaturisierung und somit Kostensenkung der Probenvorbereitung limitiert.

Bei Flachbett-Sequenziersystemen wird die Probe traditionell manuell durch Unterdruck mit einer Pipette aus der Mikrotiter-Platte aufgenommen und durch Überdruck in Aussparungen im Gel, sogenannte Probenaschen, abgegeben. Um den Durchsatz und die Reproduzierbarkeit zu erhöhen wurde eine Reihe unterschiedlicher Wege eingeschlagen.

- Ein System, das peristaltische Pumpen einsetzt wurde von Smith und Watson (Smith und Watson 1993) vorgestellt.
- Ein anderer Ansatz ist es, die Probenlösung durch ein Kapillarsystem aus einer Druckkammer - in der die Probenplatte plaziert ist - in die Probenaschen zu transferieren (Panussis *et al.*, 1998). Dieses System erlaubt die parallele Beladung von 96 Proben. Die Kapillaren besitzen einen inneren Durchmesser von 0,1 mm (äußerer 0,15 mm). Die Länge beträgt 200 mm. Diese Methode kann ab Probenvolumina von etwa 3µl eingesetzt werden.
- An der Universität Washington wurde eine Methode entwickelt, die Agarose Gele direkt mit Glaskapillaren lädt (Evensen *et al.* 1999).
- In Zusammenarbeit mit der elektronischen und mechanischen Werkstatt des EMBL wurde von mir ein automatischer Beladungsroboter für eine A.L.F-Sequenziereinheit entwickelt (EMBL Research Reports 1996). Der Aufbau besteht aus einem Spritzensystem der Firma Hamilton und einem Dreiachsen-Schrittmotorsystem der Firma "Isel".

Diese Einheit wurde über Tastaturprogrammierung so eingerichtet, dass nach dem Beladen der Geltaschen, die Elektrophorese und Datenaufnahme gestartet werden. Hierbei wurde ein Vorteil der „On-Line“ Sequenzier-Geräte ausgenutzt, der bei kurzen Läufen auftritt; die Wiederbeladbarkeit der Gele.

Nach der Elektrophorese wird der Lauf automatisch über die Tastatursteuerung angehalten und die Gele werden erneut automatisch beladen. Dieser Vorgang wird wiederholt, bis die Proben der beiden Mikrotiter-Platten vollständig analysiert sind. Dieses System wurde zur automatischen Fragmentanalyse eingesetzt.

#### **1.4 Zielsetzung**

Um die neuentwickelten und schon existierenden hochauflösenden Detektionssysteme für Flachbettssysteme und fluoreszent markierte DNA-Moleküle effektiv zu nutzen, wird das theoretische Maximum von Proben pro Trenngel angestrebt.

Durch die in dieser Arbeit vorgestellte Entwicklung eines neuartigen Systems zum Beladen von Gelen mit DNA soll nicht nur die Effizienz der Flachbett-Systeme erhöht werden, sondern auch die Genauigkeit der Daten und der Grad der Automatisierung. Mit dieser Methode, die auf porösen Membranen beruht, konnten in der ersten Testphase 200 Proben (Erfler et al. 1997) nach der Trennung im Gel durch Photodiodenarrays mit 0,8 mm Pixelbreite analysiert werden.

Eine weitere Zielsetzung war, die mit dieser Methode erreichbare Anzahl analysierbaren Proben pro Trenngel unter Anwendung von Photodiodenarrays mit 0,2 mm Pixelabstand nochmals zu erhöhen. Um dies zu erreichen, musste das Verständnis von Membran-DNA Komplexen in äußeren elektrischen Feldern vertieft werden.