

Einleitung

Modul 5 Fachübergreifendes Thema

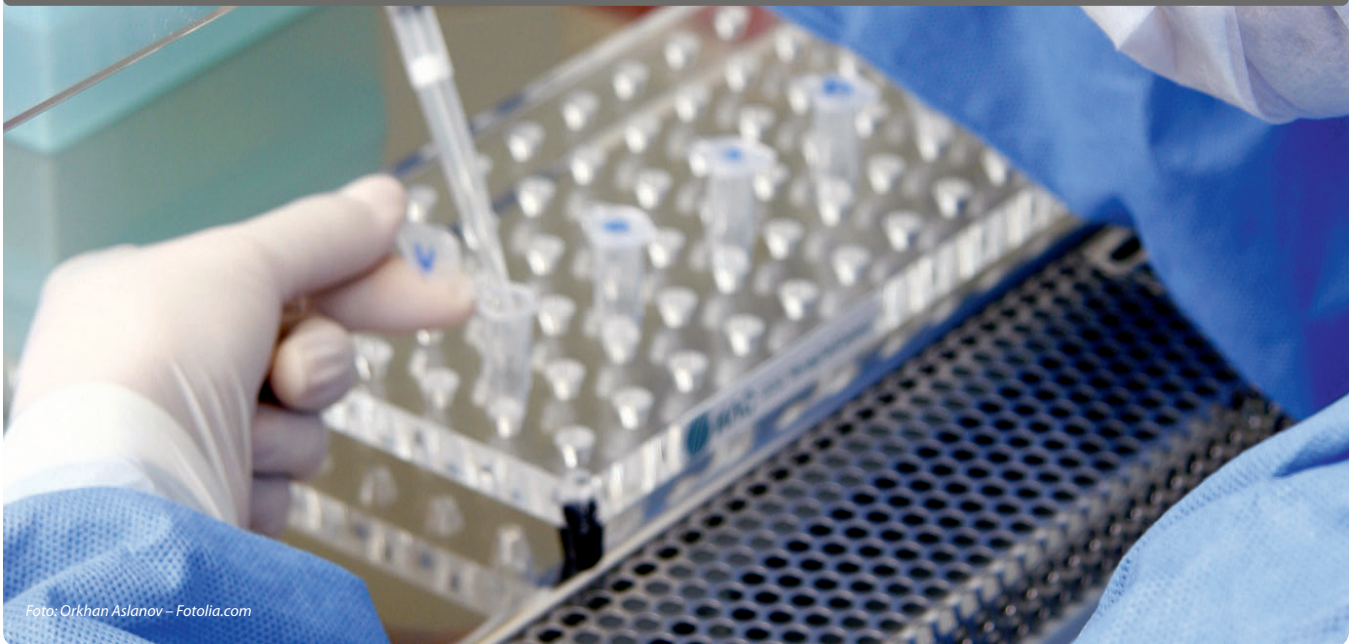


Foto: Orkhan Aslanov – Fotolia.com

Sequenzierung – eine bedeutende Methode der modernen Biowissenschaften

Als am 13. September 2011 im Berliner Max-Delbrück-Centrum für Molekulare Medizin (MDC) der Startknopf für eines der neuesten DNA-Sequenziergeräte gedrückt wurde, öffnete sich in Deutschland die Tür zur „dritten Generation“ der Sequenzierung („Third Generation Sequencing“). Mit Hilfe dieses Gerätes wird es den Forschern möglich sein, einzelne DNA-Moleküle in Echtzeit zu sequenzieren und damit schneller und kostengünstiger zu Genomanalysen zu kommen. Das Ziel, das Genom eines Menschen an einem Tag für weniger als 1000 US-Dollar sequenzieren zu können, ist in greifbare Nähe gerückt.

Davon konnten Frederick Sanger, Alan Coulson, Allan Maxam und Walter Gilbert nur träumen, als sie 1977 ihre Verfahren zur Sequenzierung von DNA-Molekülen vorstellten. Beide Methoden bauen darauf auf, die DNA zunächst zu vervielfältigen, und erlauben es, einzelsträngige DNA-Fragmente bis zu einer Länge von etwa 1000 Basen zu sequenzieren.

Die heute nicht mehr verwendete **Maxam-Gilbert-Methode** beruht auf der partiellen chemischen Spaltung doppelsträngiger DNA in vier getrennte Reaktionen. Die verwendeten Chemikalien (u. a. Piperidin und Dimethylsulfat) sind giftig und krebserregend. Auch aus diesem Grund setzte sich zuerst das **Kettenabbruch-Verfahren** (nach Sanger und Coulson) durch, das ohne schädliche Chemikalien auskommt und besser automatisierbar ist. Jahrzehnte lang wurde das Entziffern des genetischen Codes daher fast ausschließlich mit Hilfe dieser Sequenziermethode durchgeführt, einer sehr genauen, durch technische Neuerungen immer weiter verbesserten Methode. Das Prinzip basiert darauf, dass die mit Hilfe einer Polymerase erfolgte Vervielfältigung der Einzelstrang-DNA durch modifizierte Nukleotide (Didesoxinukleotide) basenspezifisch abgebrochen wird (siehe Abbildung 1A). Die entstehenden markierten DNA-Teilstränge werden mittels Gelelektrophorese nach Länge getrennt, so dass die Basensequenz des DNA-

Stückes anhand der Längenabfolge im Chromatogramm erkennbar ist (Abb. 1). In den Anfangsjahren der Sequenzierung wurde mit Radioaktivität markiert (^{32}P oder ^{35}S), heutzutage wird ausschließlich die Fluoreszenz-Markierung eingesetzt (Abb. Chromatogramm mit den Banden der Spuren G, C, A, T). Das Sanger-Verfahren war für fast dreißig Jahre das dominierende Sequenzierverfahren. Dieses zur Bestimmung der Basenabfolge bei kleineren DNA-Stücken sehr gut geeignete Verfahren stößt jedoch schnell an seine Grenzen, wenn es um das Auslesen ganzer Genome geht.

Sequenziertechnologien der zweiten und dritten Generation

Recht bald nach der Jahrtausendwende setzen sich daher neuartige, als „Second Generation Sequencing“ bzw. „Next-Generation Sequencing“-bezeichnete Technologien durch, die mittels verschiedener Methoden mit minimalem personellen Aufwand in kürzester Zeit enorme Datenmengen erzeugen können. Dank dieser Technologien, die gut automatisierbar sind, kommen die Forscher heute wesentlich schneller und kostengünstiger zu Genom-Analysen. Sie basieren nicht auf DNA-Strangabbruch wie die Sanger-Methode, sondern auf Sequenzierung durch Synthese oder Sequenzierung durch Ligation. Allerdings muss bei diesen Verfah-

Einleitung **Modul 5** Fachübergreifendes Thema

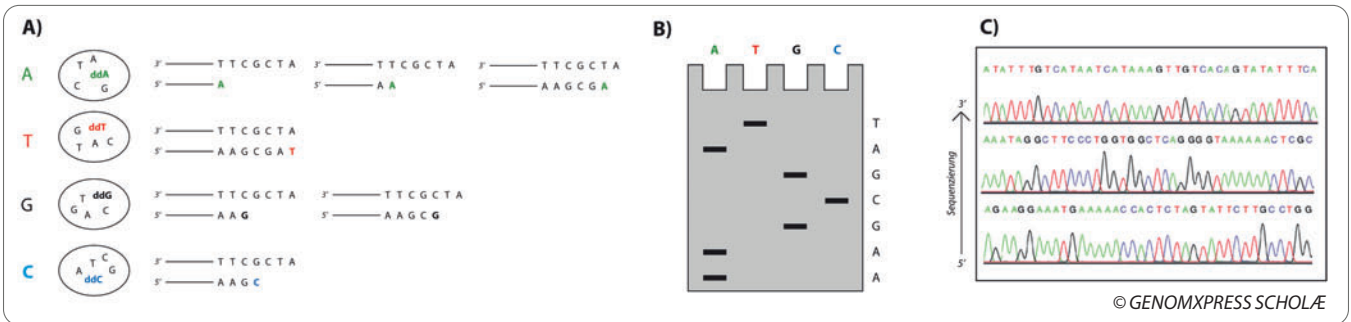


Abb. 1: Sequenzierung nach Sanger/Coulson: Von einem Vorlagestrang werden Kopien erstellt. Dabei werden zusätzlich Didesoxi-Nukleotide eingesetzt (in vier Reaktionen jeweils entweder ddA, ddG, ddC oder ddT), die zum Kettenabbruch führen (A). Es entsteht eine Mischung unterschiedlich langer Fragmente, die auf einem Gel aufgetrennt werden können (B). Anhand der Laufstrecke im Gel kann dann die Sequenz abgelesen werden. Heute werden die ddNukleotide farbig markiert, so dass ein einziger Reaktionsansatz ausreicht. Die Sequenz wird automatisiert ausgelesen, und in Form eines Elektroferrogramms ausgegeben (C).

ren wie auch bei der Sanger-Sequenzierung die zu sequenzierende DNA zunächst in kurze Abschnitte zerteilt und amplifiziert werden. Im Unterschied zur Sanger-Methode sind bei den Verfahren der zweiten Generation jedoch keine Elektrophorese-Verfahren zum Auslesen der Sequenz mehr notwendig

Die **Pyrosequenzierung** als Beispiel der Sequenzierung durch Synthese nutzt wie die Sanger-Sequenzierung eine DNA-Polymerase zur Synthese des DNA-Gegenstranges. Dabei kann man die DNA-Polymerase gewissermaßen „in Aktion“ beim Einbau der Nukleotide beobachten. Ein Einzelstrang dient der Polymerase als Matrize, an dem das Enzym den zweiten DNA-Strang neu synthetisiert. Nacheinander werden der Reaktion die vier verschiedenen Desoxyribonukleotidtriphosphate (dATP, dCTG, dGTP und dTTP, gemeinsam als Nukleotide bzw. dNTPs bezeichnet) kontrolliert zugegeben. Nach jeder Runde werden die nicht gebundenen dNTPs abgebaut. Ist ein passendes Desoxyribonukleotidtriphosphat gefunden, kommt es aufgrund einer enzymatischen Reaktion zu einem Lichtblitz, den ein Detektor erfasst (Abb. 2). So lässt sich die gesuchte DNA-Sequenz während der Synthesereaktion ermitteln.

Neben der Pyrosequenzierung gibt es noch weitere Sequenziermethoden der zweiten Generation. Beispiele sind die **Illumi-**

na- und die **SOLID-Technologie**. Diese Methoden führen zwar zu sehr kurzen Sequenz-Fragmenten von nur etwa 30-40 Basenpaaren und eine Neusequenzierung großer Genome ist daher nur schwer möglich. Allerdings sind diese Technologien im Vergleich zu Sanger sehr gut in riesigen Maßstäben anzuwenden und deutlich günstiger als die Pyrosequenzierung, deren Sequenz-Fragmente immerhin eine Größe von etwa 400 Basenpaaren erreichen. Durch den Einsatz großer Rechnerkapazitäten und bioinformatischer Methoden lassen sich die mit Illumina- oder SOLID-Technologie erzeugten Rohdaten jedoch gut für die erneute Sequenzierung bekannter Genome (Resequenzierung) oder die Diagnostik von Genveränderungen einsetzen. Die Pyrosequenzierung hingegen eignet sich aufgrund ihrer großen Leselängen für die Neusequenzierung von großen Genomen, wie am Beispiel der Gerste im Projekt BARLEX (siehe GXP 4.07) gezeigt werden konnte.

Das neue Gerät am MDC der dritten Generation, nutzt eine Technologie namens SMRT (Abkürzung für „**S**ingle **M**olecule **R**eal-**T**ime“), welche die direkte Sequenzierung einzelner DNA-Moleküle ohne vorherige Vermehrung der DNA erlaubt. Dadurch verringert sich die Fehleranfälligkeit bei gleichzeitiger Steigerung der Effizienz.

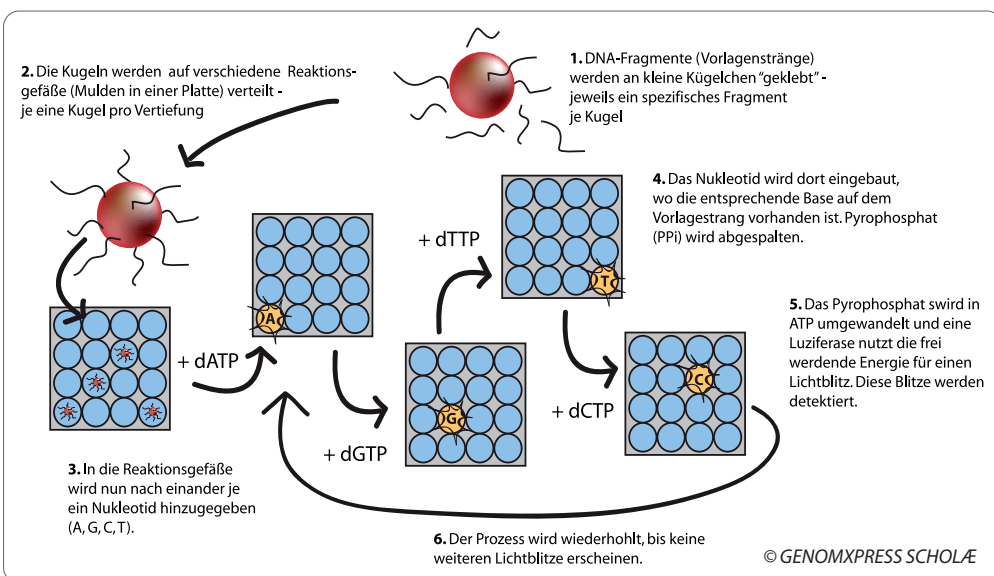


Abb. 2: Die zweite Generation: Bei der Pyrosequenzierung werden die Sequenzierreaktionen an kleinen Kügelchen durchgeführt. Nacheinander werden die Basen hinzugegeben, und bei Vorhandensein einer entsprechenden Template-Base eingebaut. Das frei werdende Pyrophosphat zu ATP umgewandelt und eine Luziferase nutzt diese Energiequelle, um einen Lichtblitz auszusenden, der von einer Kamera detektiert werden kann.