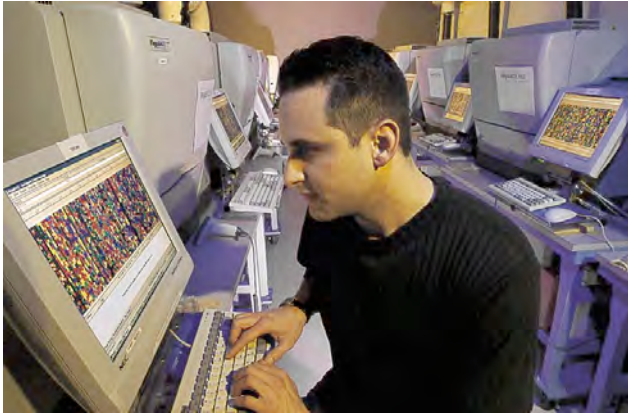


## Gentechnik und Genomics



Das Foto auf zeigt ein Labor; in dem Automaten die menschliche DNA mit einer Rate von 350000 Nucleotiden pro Maschine und Tag sequenzieren.

### Einführung

Die Kartierung und Sequenzierung des menschlichen Genoms, eine der grössten Leistungen der modernen Naturwissenschaften, wurde durch Fortschritte in der Gentechnik möglich, beginnend mit der Entwicklung von Methoden zur Herstellung rekombinanter DNA. Das ist DNA, bei der Gene von zwei verschiedenen Quellen - oft sind dies unterschiedliche Spezies - *in vitro* in einem Molekül vereinigt werden. Heute werden Hunderte nützlicher Produkte gentechnisch hergestellt. Gentechnik oder Gentechnologie ist die direkte Manipulation von Genen zur praktischen Anwendung. Hierbei sind Methoden zur Herstellung rekombinanter DNA von zentraler Bedeutung. Diese kann dann in kultivierte Zellen eingeschleust werden, welche die DNA replizieren und das Gen exprimieren, wodurch das gewünschte Protein entsteht. Das Bakterium *E. coli* wird oft als Wirtsorganismus für rekombinante DNA eingesetzt, weil es einfach zu kultivieren ist und man seine Biochemie weitgehend kennt.

Die Gentechnik hat zu einer industriellen Revolution der Biotechnologie geführt. Biotechnologie kann demnach definiert werden als die Veränderung lebender Organismen oder deren Bestandteile mit dem Ziel, bestimmte Verfahren durchzuführen beziehungsweise besondere Produkte herzustellen. Viele biotechnologische Verfahren sind Jahrhunderte alt, beispielsweise die Verwendung von Mikroorganismen zur Herstellung von Wein und Käse. Natürliche genetische Prozesse, wie zum Beispiel Mutation und genetische Rekombination, waren schon immer an solchen biotechnologischen Verfahren beteiligt. Biotechnologie, die auf der Veränderung von DNA *in vitro* basiert, unterscheidet sich von früheren Praktiken, indem sie es möglich macht, spezifische Gene

zu modifizieren und sie zwischen ganz verschiedenen Organismen wie beispielsweise Bakterien, Pflanzen und Tieren zu übertragen.

Gentechnische Methoden werden heute in zahlreichen praktischen Bereichen eingesetzt, von der Landwirtschaft bis hin zur Gerichtsmedizin. Die wichtigsten - und von Laien oft übersehenen - Fortschritte durch diese Methoden liegen jedoch in der Grundlagenforschung. Die Gentechnik hat auf fast allen Gebieten der Biologie zu einem enormen Erkenntnisgewinn geführt, indem sie den Wissenschaftlern ein ganz neues Instrumentarium liefert, mit denen sie an die Natur gestellte Fragen beantworten können. Die Gentechnik versorgt uns in atemberaubender Geschwindigkeit mit detaillierten Kenntnissen über das Genom des Menschen und anderer Organismen, und über die Struktur der von den Genen codierten Proteine. Viele dieser Kenntnisse galten noch vor wenigen Jahrzehnten als unerreichbar.

In diesem Kapitel befassen wir uns mit den wichtigsten gentechnischen Methoden, besprechen, wie Genome auf dem Niveau der DNA analysiert werden und verschaffen uns einen Überblick über praktische Anwendungen der Gentechnik.

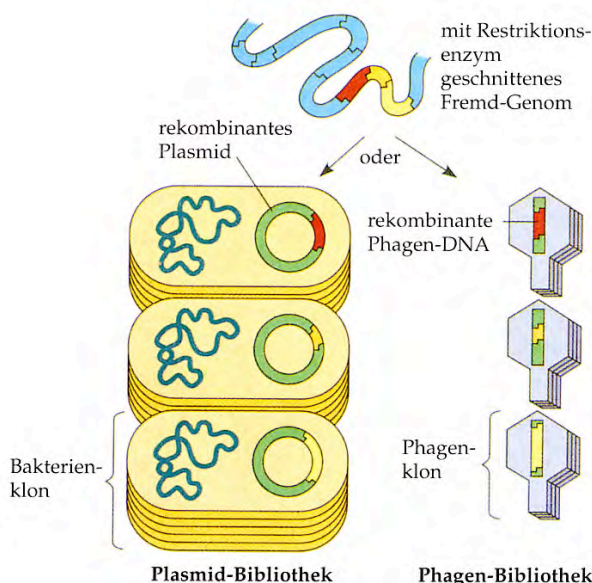
### DNA-Klonierung

Will ein Molekularbiologe ein bestimmtes Gen untersuchen, so hat er eigentlich ein enormes Problem. Natürlich vorkommende DNA-Moleküle sind sehr lang, und ein einzelnes Molekül trägt gewöhnlich viele Gene. Zudem beanspruchen Gene nur einen kleinen Teil der chromosomalen DNA, während der Rest aus nichtcodierenden Sequenzen besteht. Ein menschliches Gen umfasst nur rund 1/100000 eines chromosomalen DNA-Moleküls. Als weitere Schwierigkeit ist die Unterscheidung zwischen einem Gen und der umgebenden DNA sehr problematisch, da sie nur auf Unterschieden in der Nucleotidsequenz beruht. Um direkt mit einem spezifischen Gen zu arbeiten, mussten die Forscher also Methoden zur Herstellung wohldefinierter DNA-Fragmente in Gen-Grösse in vielen identischen Kopien entwickeln. Mit anderen Worten: Sie benötigten Techniken zur Genklonierung.

## Klonierte Gene werden in DNA-Bibliotheken gespeichert

Weil Genklonierungen mit einem Gemisch von Fragmenten aus dem ganzen Genom eines Organismus beginnen, wird dies als "Schrotschuss-Ansatz" (shotgun) bezeichnet - es wird also nicht ein einzelnes Gen gezielt kloniert. Tausende verschiedener rekombinanter Plasmide werden produziert, und jeweils ein Klon bringt eine Kolonie hervor. Der vollständige Satz von Tausenden rekombinanter Plasmid-Klone, von denen jeder Kopien eines bestimmten Segments des ursprünglichen Genoms enthält, wird als genomische Bibliothek (genomic library) bezeichnet. Man kann solch eine Bibliothek konservieren und sie als Quelle anderer Gene von Interesse oder zur Genomkartierung verwenden. Neben den Plasmiden sind auch bestimmte Bakteriophagen als Klonierungsvektoren zum Erstellen genomischer Bibliotheken geeignet.

Fragmente von Fremd-DNA können wie in ein Plasmid auch in das Phagengenom eingebaut werden. Dazu benutzt man ebenfalls Restriktionsenzyme (Endonucleasen) und Ligase. Die rekombinante Phagen-DNA wird nun in Capside in vitro verpackt und in einem normalen Infektionsprozess in die bakterielle Zelle eingebracht. In der Zelle angelangt, repliziert sich die Phagen-DNA, und neue Phagen werden produziert, von denen jeder die Fremd-DNA einhält. Eine genomische Bibliothek aus Phagen wird als Kollektion von Phagen-Klonen aufbewahrt. Unabhängig davon, welcher Klonierungsvektor benutzt wird, nehmen die Restriktions-Endonucleasen beim Zerschneiden der genomischen DNA in keiner Weise Rücksicht auf die Gengrenzen. Daher können manche Gene in einer genomischen Bibliothek auf zwei oder mehr Klone verteilt sein.

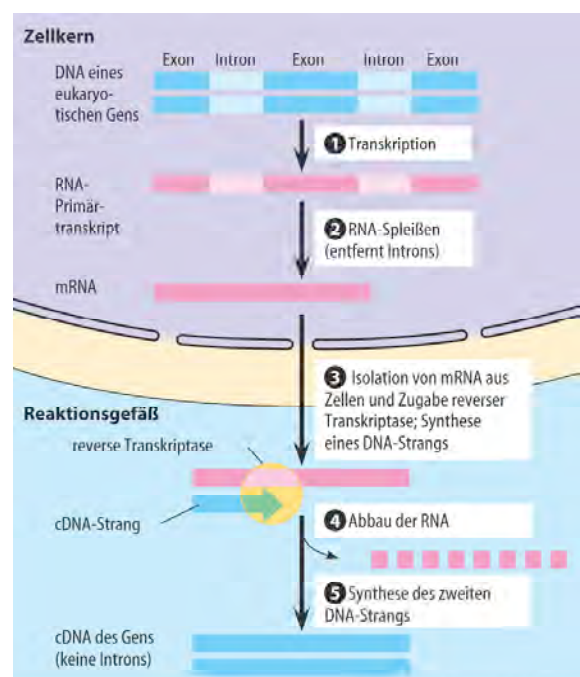


## Mithilfe der Reversen Transkriptase lässt sich von mRNA eine DNA-Kopie erzeugen

Aus komplementärer DNA oder cDNA lässt sich eine viel kleinere DNA-Bibliothek erzeugen, die nur die Gene enthält, die in einem bestimmten Gewebe transkribiert werden.

Der erste Schritt zur Erzeugung von cDNA ist die Extraktion von mRNA aus einem Gewebe. Die mRNA wird dann mit einem Primer hybridisiert, der die Poly(A)-Schwänze der mRNA erkennt. Primer und die mRNA dienen als Matrize für die Reverse Transkriptase, die von RNA ausgehend DNA synthetisiert. Auf diese Weise entstehen cDNA-Stränge, die zur mRNA komplementär sind.

Eine Sammlung von cDNAs aus einem bestimmten Gewebe, die zu einem bestimmten Zeitpunkt im Lebenszyklus eines Organismus isoliert wurde, bezeichnet man als cDNA-Bibliothek. mRNAs überdauern im Cytoplasma nicht lange und sind häufig nur in geringen Mengen vorhanden, sodass eine cDNA-Bibliothek eine Art Momentaufnahme des Transkriptionsmusters einer Zelle darstellt. cDNA-Bibliotheken leisten bei Vergleichen der Genexpression in verschiedenen Geweben zu verschiedenen Entwicklungsphasen unschätzbare Dienste. Durch ihre Verwendung ließ sich beispielsweise zeigen, dass bis zu einem Drittel aller Gene, die ein Tier besitzt, nur während der vorgeburtlichen Entwicklung exprimiert wird. Komplementäre DNA ist auch Ausgangspunkt für die Klonierung von eukaryotischen Genen. Sie ist besonders hilfreich bei der Klonierung von Genen, die nur von einigen wenigen Zelltypen und in geringen Mengen exprimiert werden.



## DNA-Analyse und Genomics

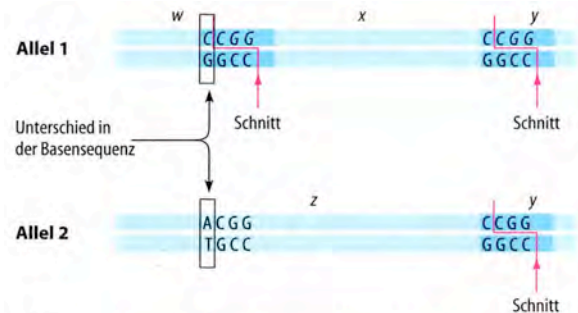
Ausgerüstet mit Methoden zur Präparation homogener DNA-Proben, von denen jede eine grosse Zahl identischer Segmente des Genoms enthält, können wir beginnen, weiterreichende Fragen zu stellen. Stellen Sie sich vor, Sie hätten ein DNA-Segment mit einem Sie interessierenden menschlichen Gen kloniert. Ist das Gen in verschiedenen Menschen unterschiedlich, und gibt es bestimmte Allele davon, die mit einer Erbkrankheit verbunden sind? Wo im Körper und wann wird das Gen exprimiert? Wo innerhalb des Genoms ist dieses Gen lokalisiert? Sie können auch Fragen zur Evolution angehen, wenn Sie untersuchen, wie sich das Gen bei verschiedenen Organismenarten unterscheidet.

Um solche Fragestellungen vollständig zu klären, brauchen Sie möglicherweise die komplette Nucleotidsequenz des Gens und seiner Verwandten in anderen Individuen und Arten. Und schliesslich möchten die Forscher die Sequenzen ganzer Genome zur Verfügung haben, um sie als Basis zum Studium ganzer Sätze von Genen und ihren Interaktionen zu verwenden - ein Ansatz, der Genomics genannt wird. Aber wir können auch beginnen, die klonierte DNA durch mehr indirekte Methoden zu analysieren, die sehr schnell wertvolle vergleichende Informationen liefern. Die meisten dieser Methoden - ebenso wie übrigens auch die DNA-Sequenzierung - basieren auf einer Technik, die Gelelektrophorese genannt wird. Diese Technik, trennt Makromoleküle - Nucleinsäuren oder Proteine - aufgrund ihrer Grösse, ihrer elektrischen Ladung und anderer physikalischer Eigenschaften. Bei linearen DNA-Molekülen hängt die Trennung nur von der Grösse der Fragmente ab. Die Gelelektrophorese trennt eine Mischung von DNA-Molekülen in Banden auf, wobei jede Bande DNA-Moleküle derselben Länge enthält.

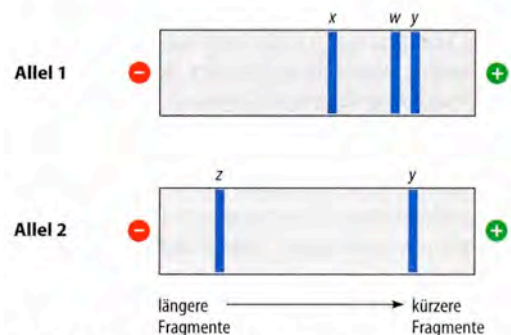
### Restriktionsfragment-Analyse zeigt jedes DNA-Molekül aufgrund seiner Restriktions-Schnittstellen ein typisches Muster

Die Restriktionsfragment-Analyse erkennt indirekt bestimmte Unterschiede in der Nucleotidsequenz von DNA-Molekülen. Bei dieser Methode benutzt man die Gelelektrophorese, um jene DNA-Fragmente der Grösse nach zu ordnen, die aus der Inkubation langer DNA-Moleküle mit einem Restriktionsenzym entstehen. Man braucht sich nur das Muster der gefärbten Banden in einem Gel anzusehen, um wissenschaftlich nützliche Informationen zu erhalten. Trennt man eine Mischung von Restriktionsfragmenten eines bestimmten DNA-Moleküls gelelektrophoretisch auf, so erhält man ein Bandenmuster, das

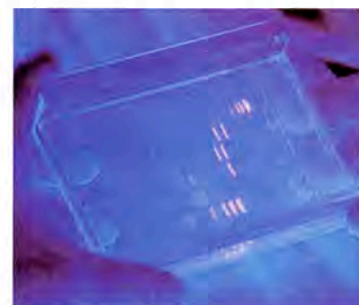
charakteristisch für das Ausgangsmolekül und die eingesetzten Restriktionsenzyme ist. In der Tat können die relativ kleinen DNA-Moleküle von Viren und Plasmiden einfach durch ihre Restriktionsfragment-Muster identifiziert werden. (Längere DNA-Moleküle, wie beispielsweise die DNA eukaryotischer Chromosomen, liefern viel zu viele Fragmente, als dass sie in Form diskreter Banden erscheinen könnten.) Weil man DNA unbeschadet aus Gelen isolieren kann, eröffnet diese Prozedur eine Möglichkeit, reine Proben individueller DNA-Fragmente zu erhalten.



(a) **DNA zweier Allele.** Hier sind zwei homologe DNA-Abschnitte abgebildet, die unterschiedliche Allele eines Gens tragen, nur die relevanten Basen sind gezeigt. Ein Unterschied in einem einzigen Basenpaar führt zum Allel 2, das eine Erkennungssequenz (Schnittstelle) weniger für ein bestimmtes Restriktionsenzym besitzt. Dieses Enzym schneidet die DNA des Allels 1 in drei Stücke (w, x und y), die DNA des Allels 2 dagegen nur in zwei Stücke (z und y).



(b) **Gelelektrophorese von Restriktionsfragmenten.** Die Elektrophorese trennt die Restriktionsfragmente, die aus jedem Allel gewonnen wurden. Die Bandenmuster im Gel zeigen klare Unterschiede zwischen den beiden Allelen. Allel 1 liefert drei Banden, welche den Fragmenten w, x und y entsprechen; Allel 2 liefert zwei Banden, die z und y entsprechen.



(c) **Fertiges Gel.** Nach Zugabe eines DNA-bindenden Farbstoffs (Ethidiumbromid) fluoreszieren die Banden in ultraviolettem Licht in rosa Farbe. In dem hier gezeigten Gel sind sechs Proben aufgetragen. Jede Probe enthält eine Mischung von Fragmenten aus einer DNA-Präparation, die mit einem Restriktionsenzym geschnitten wurde. Die rosa Banden entsprechen DNA-Restriktionsfragmenten unterschiedlicher Grösse.

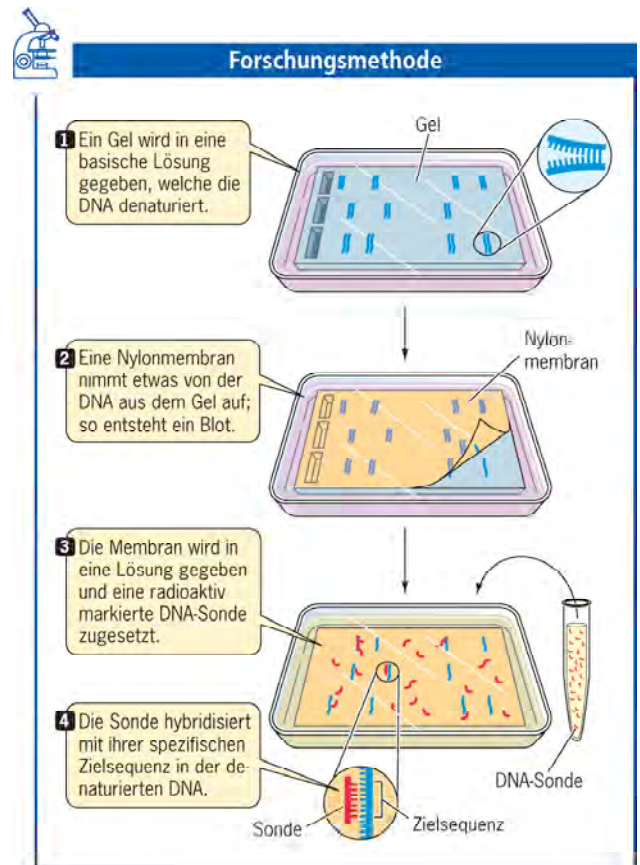
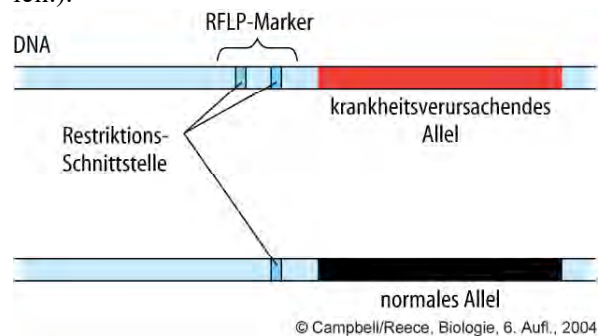
Verwenden wir nun die Restriktionsfragment-Analyse, um zwei unterschiedliche DNA-Moleküle zu vergleichen, die beispielsweise verschiedene Allele eines Gens tragen. Wir beginnen, indem wir die beiden DNA-Proben mit demselben Restriktionsenzym verdauen. Weil sich die beiden Allele in ihrer DNA-Sequenz zwangsläufig ein wenig unterscheiden, könnten sie in einer oder mehreren Schnittstellen voneinander abweichen. Ist dies der Fall, sollten die beiden Restriktionsfragment-Gemische in der Gelelektrophorese unterschiedliche Bandenmuster zeigen. Die Abbildung (c) zeigt, wie die Restriktionsfragment-Analyse durch Gelelektrophorese zwischen zwei Allelen eines Gens unterscheiden kann, die in nur einem Basenpaar voneinander abweichen - falls eine Restriktions-Schnittstelle davon betroffen ist.

Das Ausgangsmaterial sind Proben klonierter und gereinigter Gene. Durch Kombination von Gelelektrophorese mit Nucleinsäure-Hybridisierung kann man denselben Vergleich auch ohne vorhergehende Klonierung durchführen. Wir beginnen unsere Analyse mit der DNA des gesamten Genoms. Zwar ergeben sich zu viele Fragmente, als dass man sie elektrophoretisch sauber voneinander trennen könnte, aber durch Nucleinsäure-Hybridisierung mit einer spezifischen Sonde können wir diskrete Banden markieren, die sich von unserem untersuchten Gen ableiten.

Diese Methode wird Nucleinsäure-Hybridisierung genannt: Eine radioaktiv markierte einzelsträngige DNA bildet spezifische Wasserstoffbrücken mit der komplementären DNA-Sequenz; dies kann dann durch Autoradiographie detektiert werden. Nur verwenden wir eine Sonde in Kombination mit einer zusätzlichen Technik, die man als „Southern-Hybridisierung“ oder Southern-Blotting bezeichnet. Das Southern-Blotting hat sich als ausserordentlich wichtige Technik erwiesen, als Biologen ihr Interesse auch auf nichtcodierende DNA richteten, aus welcher die meiste DNA von Tier- und Pflanzengenomen besteht. Gibt es auch bei den nichtcodierenden Sequenzen Unterschiede, wie sie bei den verschiedenen Allelen eines Gens gefunden wurden? Die Untersuchung nichtcodierender DNA verursachte bei den Forschern einige Aufregung, weil sie viele individuelle Unterschiede in den Bandenmustern entdeckten. Unterschiede in der DNA-Sequenz homologer Chromosomen, die sich in unterschiedlichen Restriktionsfragment-Mustern äussern, sind beim Menschen und anderen Organismen über das ganze Genom verstreut. Diese Unterschiede bezeichnet man als Restriktionsfragment-Längenpolymorphismen (RFLPs; ausgesprochen „Riflips“). Solche Variationen im Genom entsprechen vom Konzept her den Variationen bei codierenden Sequenzen; auch sie können als genetische Marker für bestimmte Bereiche des Genoms dienen. Ein bestimmter RFLP-Marker kommt

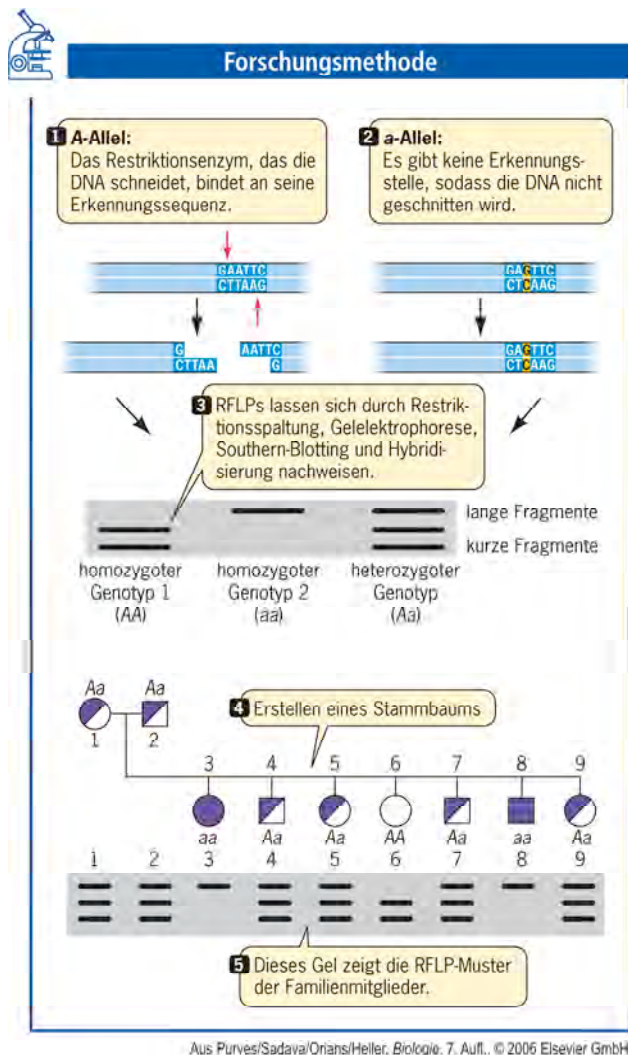
in einer Population oft in einer Vielzahl von Varianten vor. (Polymorphismus stammt aus dem Griechischen, wo es etwa "viele Formen" bedeutet.)

RFLPs werden durch Southern-Blotting detektiert und analysiert. Die Detektion eines RFLP kann ebenso in einer nichtcodierenden Sequenz als auch einen Unterschied in der codierenden Sequenz zweier Allele darstellen. Wegen der Sensitivität der DNA-Hybridisierung kann das ganze Genom als Ausgangsmaterial benutzt werden (Proben menschlicher DNA erhält man typischerweise aus weissen Blutzellen.).



Weil RFLP-Marker in mendelscher Weise vererbt werden, kann man sie als genetische Marker zur Erstellung von Kopplungskarten verwenden. Die Häufigkeit, mit der zwei RFLP-Marker - oder ein RFLP-Marker und ein bestimmtes Allel eines Gens - zusammen vererbt werden, gilt als ein Hinweis auf die relative Entfernung der beiden Loci auf dem

Chromosom. Die Entdeckung der RFLPs erhöhte die Anzahl der Marker für die Kartierung des menschlichen Genoms deutlich. Dadurch sind die Genetiker nicht mehr beschränkt auf genetische Veränderungen, die zu einem erkennbaren Phänotyp führen (wie beispielsweise Erbkrankheiten) oder auf Unterschiede in den Protein-Produkten.



## Genomprojekte eröffnen den Zugang zu grundlegenden biologischen Fragen.

Genomforschung (Genomics), das Studium ganzer Genome auf der Basis ihrer DNA-Sequenzen, brachte neue Antworten auf fundamentale Fragen zur Genomorganisation, Expressionskontrolle, Wachstumsregulation, Keimesentwicklung und Evolution. Die Eleganz der Gentechnik besteht darin, dass sie Molekularbiologen ermöglicht, Gene direkt zu studieren, ohne Umweg über den Phänotyp. Aber dieser neue Ansatz beschert uns heute das umgekehrte Problem, nämlich den Phänotyp vom Genotyp ableiten zu müssen. Wenn man am Anfang lediglich eine lange DNA-Sequenz hat, wie erkennt man Gene, und wie kann man ihre Funktion bestimmen?

## Die Auswertung von DNA-Sequenzen

DNA-Sequenzen werden in Computer-Datenbanken gespeichert, die Forschern aus aller Welt über das Internet zur Verfügung stehen. Sie benutzen eine spezielle Software, um eine solche Sequenz auf Signatursequenzen proteincodierender Gene zu durchsuchen, wie zum Beispiel Start- und Stoppsignale für Transkription und Translation (beispielsweise Promotoren) und Sequenzen, die mit dem RNA-Spleissen zusammenhängen. Die Software sucht auch Sequenzen, die in ähnlicher Weise in anderen bekannten Genen vorhanden sind. Tausende solcher Sequenzmarker sind in den Computer-Datenbanken katalogisiert. Auf diese Weise erhalten die Wissenschaftler eine Liste von Genkandidaten.

## Sequenzierung des menschlichen Genoms

Jedes menschliche Chromosom besteht aus einem doppelsträngigen DNA-Molekül. Die 46 menschlichen Chromosomen lassen sich aufgrund ihrer unterschiedlichen Größe voneinander trennen (siehe Abbildung 9.13). So ist es möglich, die DNA von jedem Chromosom für eine Sequenzierung zu isolieren. Bei einer ganz direkten Vorgehensweise würde man an einem Ende eines Chromosoms beginnen und nacheinander die 50 Millionen Basenpaare sequenzieren. Dieser Ansatz ist jedoch nicht durchführbar. Die DNA eines Moleküls, das 50 Millionen Basenpaare lang ist, lässt sich nicht auf einmal sequenzieren; pro Schritt sind lediglich etwa 700 Basenpaare möglich.

Um ein vollständiges Genom zu sequenzieren, wird die chromosomale DNA zuerst in Fragmente von etwa 500 Basenpaaren geschnitten, und dann wird jedes Fragment sequenziert. Für das menschliche Genom mit etwa 3,2 Milliarden Basenpaaren sind also über sechs Millionen solcher Fragmente erforderlich. Das nächste Problem besteht darin, diese Millionen Fragmente wieder zusammzusetzen, wie die Stücke eines Puzzlespiels. Dies lässt sich realisieren, indem man die DNA in grössere Abschnitte zerlegt, die sich überlappen, und diese überlappenden Fragmente aneinander ausrichtet. Dafür gibt es zwei Strategien:

Das öffentlich finanzierte, internationale Human-genom-Sequenzierungskonsortium IHGSC bediente sich einer Methode, die man als hierarchische Sequenzierung bezeichnet. Zuerst bestimmte man systematisch kurze Markersequenzen entlang der Chromosomen, wobei jedes DNA-Fragment, das sequenziert werden sollte, einen Marker enthielt. Dieses Verfahren lässt sich mit der Erstellung einer Strassenkarte vergleichen, auf der Städte und die Entfernungen zwischen ihnen verzeichnet sind. Den

Städten entsprechen die Markersequenzen, und die Entfernungen dazwischen werden in Basenpaaren angegeben. Die einfachsten Marker sind die Erkennungssequenzen von Restriktionsenzymen.

Einige Restriktionsenzyme erkennen DNA-Sequenzen von acht bis zwölf Basenpaaren und nicht die üblichen vier bis sechs Basenpaare. Ein DNA-Molekül mit mehreren Millionen Basenpaaren enthält relativ wenige dieser grösseren Erkennungsstellen, sodass das Enzym eine geringe Zahl von verhältnismässig grossen Fragmenten erzeugt. Diese grossen Fragmente können in einen Vektor (Bakterium oder Virus) eingefügt werden. Das Ziel ist die Erzeugung einer Genbibliothek.

Die einzelnen Bücher (Fragmente) in dieser Bibliothek lassen sich entlang des Chromosoms in der korrekten Reihenfolge anordnen. Um die DNA-Fragmente auf der Genomkarte richtig anzuordnen, vergleicht man Bibliotheken miteinander, die mit unterschiedlichen Restriktionsenzymen erzeugt wurden. Wenn zwei grosse DNA-Fragmente, die mit verschiedenen Enzymen geschnitten wurden, denselben Marker enthalten, müssen sie überlappen. Dieses Verfahren funktioniert, ist aber langsam.

Anstatt Marker zu definieren, die DNA stufenweise zu fragmentieren und dann erst zu sequenzieren, schneidet man die DNA beim Schrotschuss-

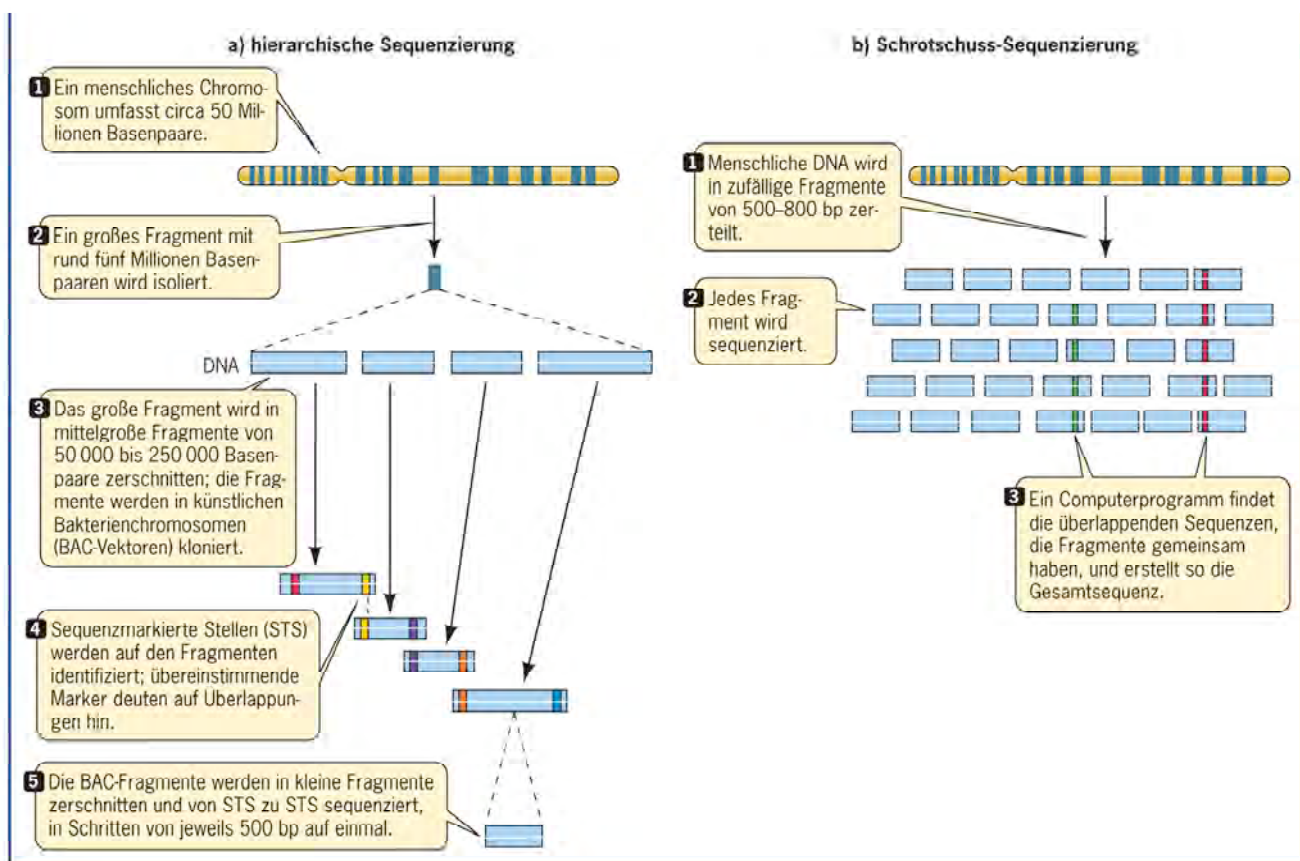
Verfahren sofort in kleine sequenzierbare Fragmente und lässt leistungsstarke Computersysteme diejenigen Sequenzstellen finden, die sich überlappen. Dann können die Sequenzen der Fragmente zusammengesetzt werden.

Die von der Privatindustrie (Firma Celera Genomics) angewandte Schrotschuss-Sequenzierung ist viel schneller als der hierarchische Ansatz, da man keine Karte des Genoms erstellen muss. Zuerst stand man der neuen Strategie mit grosser Skepsis gegenüber. Es gab Bedenken, dass die Computer ohne eine vorher auf den Chromosomen konsequent durchgeführte Markerkartierung die repetitiven Sequenzen auswählen würden, die in vielen DNA-Fragmenten vorhanden sind, und dann die Fragmente falsch zuordnen würden. Aber die schnelle Entwicklung von Hochleistungscomputern und ausgeklügelten Softwareprogrammen hat es ermöglicht, die Schrotschuss-Methode zu verbessern, sodass eine falsche Zuordnung kein wesentliches Problem mehr darstellt. Das gesamte Genom der Taufliege mit 180 Millionen Basenpaaren wurde innerhalb von etwas über einem Jahr mithilfe der Schrotschuss-Methode sequenziert. Nach diesem Erfolg hielt man die Schrotschuss-Methode auch für die Sequenzierung des viel grösseren menschlichen Genoms für geeignet.

Die beiden Forschungsgruppen kündigten die Er-



### Forschungsmethode



Aus Purves/Gadava/Orians/Heller, *Biologie*, 7. Aufl., © 2006 Elsevier GmbH

stellung einer ersten Rohsequenz im Jahr 2000 an und veröffentlichten ihre Daten gleichzeitig im Februar 2001. Anfang 2003 war die endgültige Sequenz fertiggestellt, zwei Jahre vor dem ursprünglichen Zeitplan, der auf über ein Jahrzehnt angelegt worden war, ausserdem wurde das angesetzte Budget deutlich unterschritten. Die Sequenzierung offenbarte einige interessante Eigenschaften des menschlichen Genoms:

- Von den 3,2 Millionen Basenpaaren sind weniger als zwei Prozent codierende Regionen, die insgesamt 21 000 Gene enthalten. Noch während der Sequenzierung beliefen sich die Schätzungen auf 80 000 bis 100 000 Gene, entsprechend 100 000 Proteinen.
- Ein durchschnittliches Gen enthält 27 000 Basenpaare. Bei der Grösse der Gene gibt es erhebliche Unterschiede, sie reicht von 1 000 bis 2,4 Millionen Basenpaaren. Das war zu erwarten, da die menschlichen Proteine (wie auch die RNAs) unterschiedliche Grössen aufweisen, von 100 bis etwa 5000 Aminosäuren in einer Polypeptidkette. Praktisch alle menschlichen Gene enthalten Introns .
- Über 50 Prozent des Genoms bestehen aus hochrepetitive, Sequenzen. Wiederholungssequenzen in der Nähe von Genen sind GC-reich, während weiter entfernt liegend Sequenzen AT-reich sind.
- Fast das gesamte Genom (99,9 Prozent) stimmt bei allen Menschen überein. Aber selbst diese erstaunliche Homogenität bedeutet, dass es viele unterschiedliche Sequenzen gibt. Man hat über zwei Millionen Einzelnucleotid

### Erstaunlich wenige Gene im menschlichen Genom?

Das überraschendste - und etwas demütigende - Ergebnis des Humangenom-Projekts war die zunächst berechnete Zahl an möglichen Genen. Diese war mit 30 000 bis 40 000 viel niedriger als die erwarteten 50000 bis 100000 und damit nur zwei- oder dreimal so gross wie bei Fruchtfliegen und Nematoden. Allerdings hatte man dabei offenbar zahlreiche Gene übersehen, und neuere Berechnungen gehen wieder von bis zu 80000 Genen aus. Allerdings ist diese Debatte im Moment (Sommer 2002) noch voll im Fluss. Im Vergleich zu anderen bisher untersuchten Organismen besteht bei der menschlichen DNA eine viel kleinere Fraktion aus Genen. Viel von der enormen Menge an nichtcodierender DNA im menschlichen Genom ist repetitive DNA, aber auch ungewöhnlich lange Introns tragen signifikant zur nichtcodierenden DNA bei. Menschliche Introns sind typischerweise zehnmal länger als die der Fruchtfliege *Drosophila* oder des Wurmes *Caenorhabditis*.

Tabelle 20.1: Genomgrößen und Anzahl der Gene

Organismus	Genomgröße	geschätzte Zahl der Gene	Gene pro Mb*
<i>H. influenzae</i> (Bakterium)	1,8 Mb *	1 700	950
<i>S. cerevisiae</i> (Hefe)	12 Mb	6 000	500
<i>C. elegans</i> (Nematode)	97 Mb	19 000	200
<i>A. thaliana</i> (Pflanze)	100 Mb	25 000	200
<i>D. melanogaster</i> (Fliege)	180 Mb	13 000	100
<i>H. sapiens</i> (Mensch)	3 200 Mb	30 000–80 000 (je nach Schätzung)	10–20

\* Mb = (Megabasenpaare = Millionen Basenpaare)

Was macht nun den Menschen - und Wirbeltiere im Allgemeinen - so viel komplexer als Fliegen und Würmer? Als ein Grund bekommt unser Genom, verglichen mit dem anderer Organismen, "mehr für sein Geld", also für seine codierenden Sequenzen, indem die RNA-Transkripte menschlicher Gene mehr alternatives Spleissen erfahren. Viele menschliche Gene codieren durch unterschiedliche Kombinationen von Exons zwei oder mehrere verschiedene Proteine. Dies führt zur Bildung von über 100000 verschiedenen menschlichen Polypeptiden.

Des Weiteren weist die menschliche DNA-Sequenz (ebenso wie andere Daten) darauf hin, dass unsere Proteine tendenziell komplizierter gebaut sind als die der Wirbellosen. Wie Sie sich erinnern, bestehen viele Proteine aus mehr als einer "Domäne". Domänen sind diskrete Abschnitte mit einer besonderen Struktur und Funktion. Zwar besitzt der Mensch im Vergleich zu Wirbellosen nicht mehr Typen von Proteindomänen, aber diese werden viel mannigfaltiger kombiniert.

Wie sich zusammenfassend sagen lässt, verwendet das menschliche Genom die verfügbaren modularen Elemente - Exons und Proteindomänen - durch Mischen und Anpassen erheblich plastischer als einfacher gebaute Organismen, und dies ist zweifellos ein Charakteristikum der Wirbeltiere im Allgemeinen.

### Analyse und Vergleich von Genen

Kaum die Hälfte der menschlichen Gene war bekannt, bevor das Humangenom-Projekt begann; welche Aufgaben haben also die neu entdeckten Gene? Wissenschaftler vergleichen die Sequenzen neuer Genkandidaten mit bereits sequenzierten Genen aus verschiedenen Organismen. In einigen Fällen passt die gerade bestimmte Gensequenz, zumindest teilweise, auf ein Gen, dessen Funktion bekannt ist. Beispielsweise könnte ein Abschnitt des neuen Gens mit einem bekannten Gen übereinstimmen, das eine Enzym codiert. Dies legt nahe, dass das neue Gen ebenfalls eine Enzym codiert. In anderen Fällen ist die neue Gensequenz zwar einer früher gefundenen

Sequenz ähnlich, aber deren Funktion ist nicht bekannt. In weiteren Fällen ähnelt die neue Sequenz keiner anderen Sequenz in den Datenbanken. Bei den bislang sequenzierten Organismen erwiesen sich viele Genkandidaten als vollständig neu für die Wissenschaft. Beispielsweise ist bei *E. coli*, dem bestuntersuchten Organismus in der biologischen Forschung, ein Drittel der Gene für uns völlig neu!

Trotz der hohen Komplexität des menschlichen Genoms und der von ihm codierten Proteine bestätigt der Vergleich von Genomsequenzen über alle Zweifel die evolutionären Verbindungen zwischen ganz unterschiedlichen Organismen und unterstreicht damit die Bedeutung der Forschung an einfach gebauten Lebewesen für das Verständnis auch der

menschlichen Biologie. Die Ähnlichkeit zwischen Genen ganz unterschiedlicher Organismen kann erstaunlich sein, was einen Forscher zu der Bemerkung veranlasste, er betrachte jetzt die Fruchtfliegen als "kleine Leute mit Flügeln". Hefe zum Beispiel besitzt eine Anzahl von Genen, die so weitgehende Übereinstimmungen mit den entsprechenden menschlichen Genen zeigen, dass sie diese in menschlichen Zellen funktionell ersetzen zu können. In der Tat können Forscher manchmal die Wirkung eines menschlichen krankheitsbedingenden Gens durch das Studium des homologen Gens in Hefe untersuchen.

## Methoden

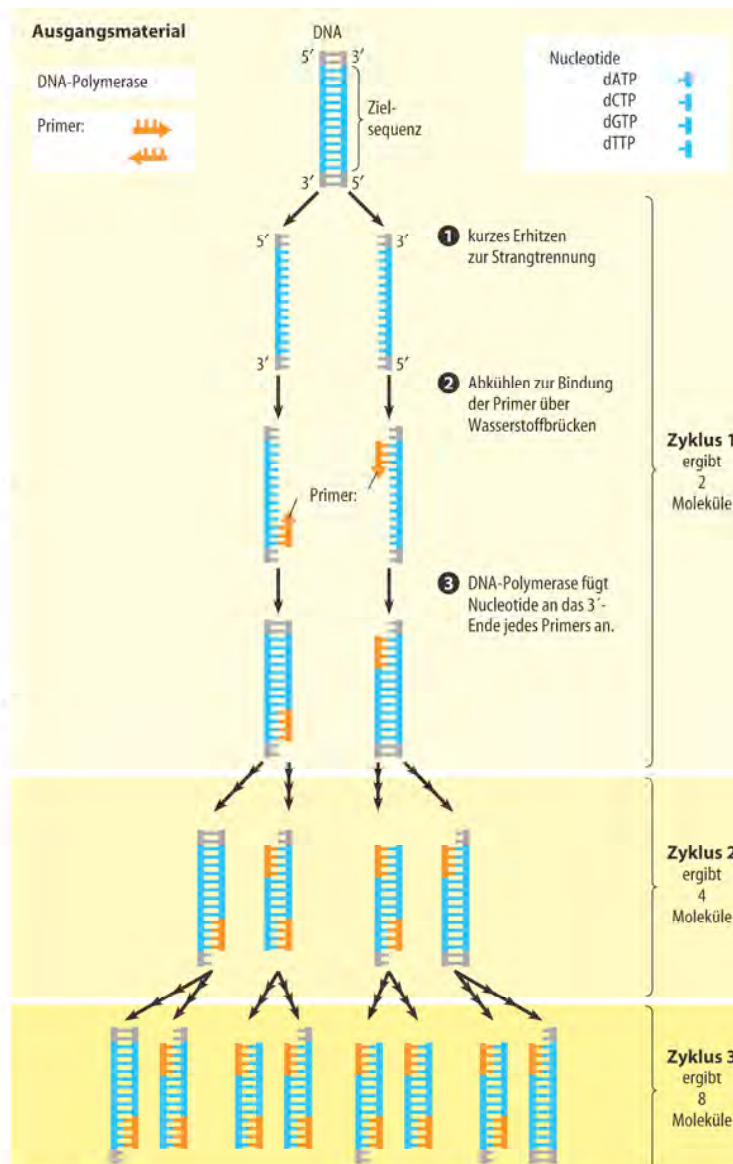
### PCR -Polymerase-Kettenreaktion

Das Ausgangsmaterial für die PCR ist eine Lösung doppelsträngiger DNA, welche die Sequenz enthält, die zur Vermehrung bestimmt ist. Es werden eine hitzeresistente DNA-Polymerase, alle vier Nucleotide sowie Primer zugegeben.

Die Primer, die für die Initiation der DNA-Synthese in der PCR gebraucht werden, sind kurze synthetische Moleküle einzelsträngiger DNA, die komplementär zu den Enden der Ziel-DNA sind. Die Primer bestimmen also das zu vermehrende DNA-Segment.

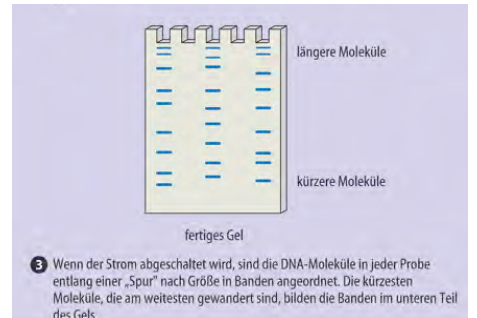
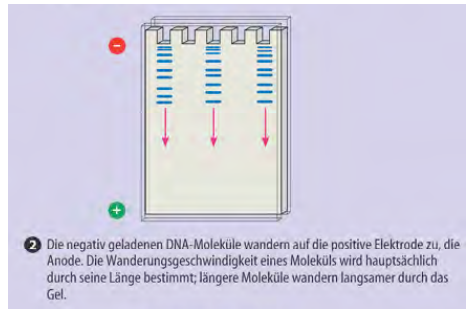
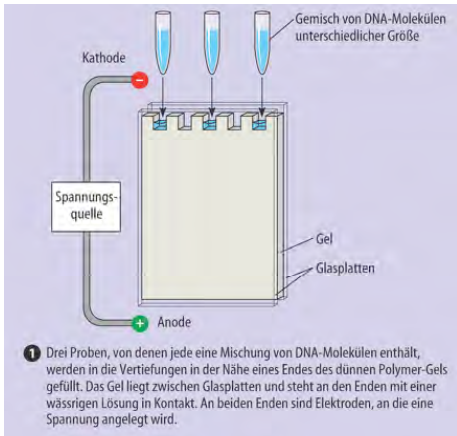
Jeder Zyklus der PCR benötigt nur etwa 5 Minuten. Am Ende eines Zyklus ist die zu vermehrende DNA-Sequenz verdoppelt – auch wenn sie hunderte von Basenpaaren lang ist.

Die Lösung wird wieder erhitzt, und der nächste Zyklus der Strangtrennung, Primerbindung und DNA-Synthese beginnt.



© Campbell/Reece, Biologie, 6. Aufl., 2004

# Gelelektrophorese



# Southern-Blotting

