

Eine kurze Einführung in die Pflanzengenetik

von Tim Blank

„Die DNA ist der Bauplan des Lebens“. Diesen Satz haben wir wohl alle schon mal gehört, doch was ist eigentlich die DNA?

DNA steht für ‚deoxyribonucleic acid‘ oder Desoxyribonukleinsäure (DNS) im deutschen. Der Aufbau der DNA ist zwar einfach erklärt, aber dennoch hoch komplex. Im Grunde genommen ist die DNA in zwei, miteinander verbundene Stränge aufgeteilt, den Leitstrang und den Folgestrang.

Jeder dieser zwei Stränge besteht aus einem Rückgrat, an welchem in einer bestimmten Reihenfolge immer eine von vier sogenannten Nukleinbasen angebracht sind, wie Perlen an einer Schnur. Die Namen dieser Nukleinbasen sind Adenin, Cytosin, Guanin und Thymin. Da diese Schreibweise aber auf Dauer zu umständlich ist, haben sich Wissenschaftler darauf geeinigt, diese einfach mit A, C, G und T abzukürzen. Jede Nukleinbase ist dabei immer auch mit einer Nukleinbase auf dem anderen Strang verbunden. Dabei gilt: Ein A ist immer mit einem T verbunden und ein C immer mit einem G. Die beiden verbundenen Nukleinbasen werden Basenpaar genannt (Abbildung 1).

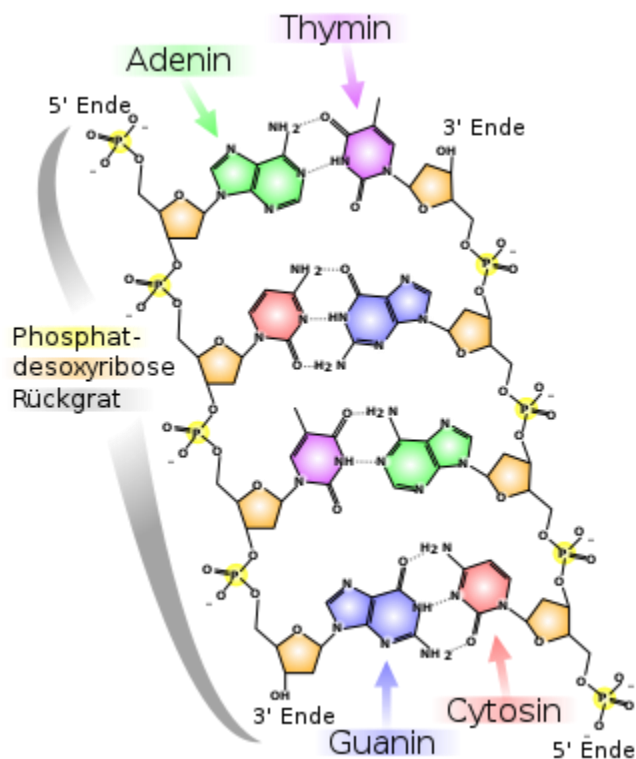


Abbildung 1: Aufbau der DNA. Am Rückgrat eines Strangs sind die Nukleinbasen Adenin (A), Cytosin (C), Guanin (G) und Thymin (T) angebracht. Jede Nukleinbase auf einem Strang ist auch immer mit einer bestimmten Nukleinbase auf dem Gegenstrang verbunden, und zwar immer A&T und C&G. Die Reihenfolge der Nukleinbasen auf dem Strang bestimmt die Information auf der DNA (Abbildung: Wikipedia, Desoxyribonukleinsäure).

Die Reihenfolge, in der diese Basenpaare an dem Rückgrat angebracht sind, bestimmt die Funktion dieses Abschnitts der DNA.

Am bekanntesten unter den funktionellen Abschnitten der DNA sind die Gene. Gene haben meistens eine Länge im Bereich von Ein- bis Zwei-Tausend Basenpaaren und liefern den Bauplan für Proteine, welche die eigentliche Arbeit in unserem Körper verrichten.

Neben den Genen gibt es aber noch viele weitere funktionelle Abschnitte. Ein Beispiel sind Promotoren, welche auf der DNA kurz vor den Genen liegen. Diese sind meistens nur wenige Basenpaare lang und teilen der Zelle mit, in welchen Situationen die Information von einem Gen gebraucht wird.

Die Zellen unseres Körpers können diesen Bauplan lesen und daraus alle Teile zusammenbauen, die sie zum Leben brauchen, und das alles auch noch fehlerfrei (unter normalen Umständen). Doch nur zu wissen, dass etwas funktioniert, reicht in der Wissenschaft nicht aus, und so hat sich das Feld der Genetik gebildet. Genetiker sind daran interessiert, die Funktion der DNA zu erforschen. Typische Fragestellungen sind dabei zum Beispiel „Welches Gen ist für ein bestimmtes Merkmal verantwortlich?“ oder „Unter welchen Bedingungen wird ein bestimmtes Gen aktiviert?“.

Wie viele andere Wissenschaften ist die Genetik aus der wissenschaftlichen Neugier entstanden, zu verstehen, wie die Welt funktioniert. So kennt quasi jeder die Namen Gregor Mendel oder Charles Darwin, welche die Vererbung und die Entwicklung von Arten (oder Spezies) untersuchten. Im Laufe der Zeit ist das Feld der Genetik enorm gewachsen und hat sich in viele Unterfelder aufgespalten. Da gibt es zum Beispiel die Humangenetik, welche sich auf den Menschen fokussiert und oft Themen im Bezug zu vererbaren Krankheiten bearbeitet. Oder die Pflanzengenetik, welche sich das Genom (also das gesamte Erbgut) von Pflanzen genauer ansieht, oft mit dem Ziel Eigenschaften zu finden, die Pflanzen resistenter gegen Krankheiten oder Umwelteinflüssen zu machen.

Einer der vielen Forschungsbereiche in der Pflanzengenetik ist der Sonderforschungsbereich 341 „Ökologische Genetik der Pflanzen“ (SFB341), welcher sich mit den Zusammenhängen zwischen der pflanzlichen Genetik und der Umwelt beschäftigt. Wir, die Forscher in diesem

Sonderforschungsbereich, wollen unter anderem herausfinden, wie sich Pflanzen an bestimmte Umweltbedingungen anpassen. Im Rahmen des „Pflanze Umwelt Klima Interaktion“ Bürgerwissenschaftsprojektes, kurz PUKI, wollen wir euch die Methoden, mit denen wir arbeiten, näherbringen und euch erklären, warum wir daran interessiert sind, möglichst viel über die Pflanzen an ihren natürlichen Standorten herauszufinden.

DNA Sequenzierung

Ein erster Schritt, um die DNA zu untersuchen ist, deren Sequenz herauszufinden, also die Anordnung der Nukleinbasen auf dem DNA-Strang. Je nachdem, wie viel und welcher Teil der DNA gelesen werden soll, gibt es unterschiedliche Herangehensweisen.

First Generation Sequencing: Die Sanger-Sequenzierung

Ist aus vorherigen Untersuchungen bereits bekannt, welcher Abschnitt der DNA von Interesse ist, reicht eine Sequenzierung dieses spezifischen Bereichs meistens aus. Dies ist zum Beispiel der Fall, wenn in einer Spezies ein Gen für eine bestimmte Funktion bekannt ist. Und nun möchte man feststellen, ob in unterschiedlichen Populationen die gleichen oder unterschiedliche Mutationen dieses Gens (also verschiedene Allelen oder Varianten desselben Genes) vorkommen.

(Populationen sind übrigens in der Biologie meistens so definiert: alle Individuen, die in einem bestimmten Gebiet vorkommen und auch miteinander in Beziehung stehen; z. B. alle Murmeltiere, die auf einem Berg vorkommen und sich untereinander fortpflanzen. Wenn Murmeltiere auf einem zweiten Berg sich nicht mit den Murmeltieren vom ersten Berg fortpflanzen, dann spricht man von zwei getrennten Populationen; sonst würde man von einer Population sprechen, die auf beiden Bergen vorkommt.)

Für so eine Fragestellung eignet sich die Sanger-Sequenzierung als Methode. Die Sanger-Sequenzierung ist eine der ältesten und die am häufigsten verwendete Sequenzierungsmethode. Mithilfe der Sanger-Sequenzierung können kurze Gen-Abschnitte von einer kurzen bekannten Sequenz aus bestimmt werden. Im Schnitt können knapp 500 Basenpaare mit einer hohen Genauigkeit sequenziert werden, wodurch wenige Sanger-Sequenzierungen meistens für die

Bestimmung der Sequenz eines Gens reichen (weil die meisten Gene um 1000-3000 Basenpaare haben).

Bei der Sanger-Methode wird die zu sequenzierende DNA als Vorlage genommen, um daraus neue DNA-Stränge zu synthetisieren. Der Startort der Synthese kann selbst bestimmt werden, indem ein ungefähr 20 Basenpaare langes DNA-Fragment mit passender Sequenz als „Primer“ genutzt wird. Der Primer wird an den vorhandenen DNA-Strang angelagert, und dann werden neue DNA-Stränge synthetisiert, indem ein Enzym namens Polymerase Schritt für Schritt die zum Vorlagestrang passende Nukleinbase in Form ihrer mobilen Chemikalie (dATP, dCTP, dGTP oder dTTP, als dNTPs zusammengefasst) anhängt. In vier Ansätzen werden jeweils die Polymerase und die dNTPs zusammengegeben, plus jeweils in einer geringen Menge eine von vier Nukleinbasen Analoga (ddATP, ddCTP, ddGTP und ddTTP, als ddNTPs zusammengefasst). Diese Analoga sehen den Nukleinbasen sehr ähnlich, aber wenn sie in den neusynthetisierten Strang eingebaut werden, ist eine Verlängerung dieses Strangs von dort aus nicht mehr möglich. So entstehen in den vier Ansätzen DNA-Stränge mit unterschiedlichen Längen, und für jeden Ansatz können die Stränge nur auf die Nukleinbase enden, dessen Analoga zum Ansatz gegeben wurde. Die Stränge werden anschließend je Ansatz nach ihrer Länge aufgetrennt und anhand der Reihenfolge der Längen in den Ansätzen kann die DNA-Sequenz bestimmt werden (Abbildung 2).

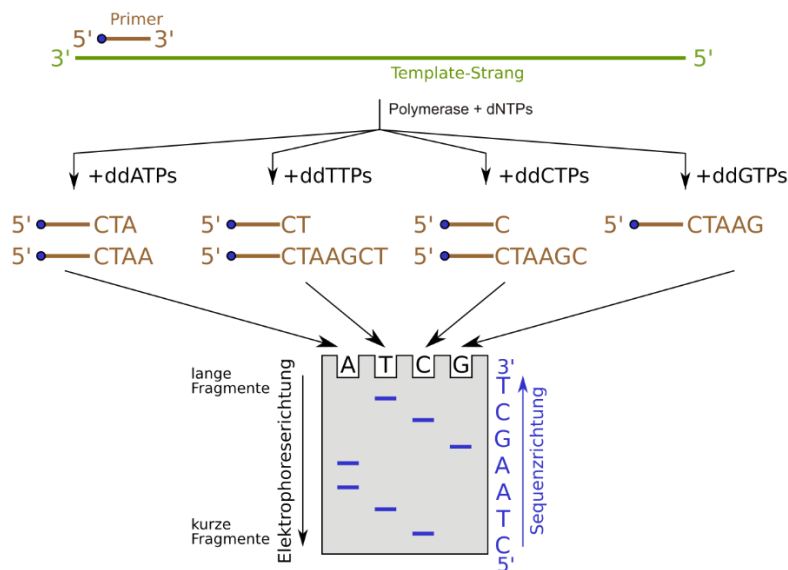


Abbildung 2: Prinzip der Sanger-Sequenzierung. Von einem Primer aus werden neue DNA-Stränge anhand einer Vorlage synthetisiert. Durch den Einbau von ddNTPs werden unterschiedlich lange Stränge mit einem bekannten Ende generiert. Die neuen Stränge werden der Länge nach aufgetrennt, und anhand der Reihenfolge der Längen kann die Sequenz des Vorlagestrangs bestimmt werden (Abbildung: Wikipedia, DNA-Sequenzierung).

Die Sanger-Sequenzierung kann dank weitreichender Automatisierungen und kommerziellen Anbietern heutzutage sehr schnell und kostengünstig angewandt werden. Der Vorlage DNA-Strang zusammen mit dem passenden Primer wird per Eillieferung an eine Firma geschickt und dort über Nacht sequenziert. Die fertige Sequenz wird einem bereits am nächsten Morgen per Mail zugeschickt. Eine Probe kostet dabei meistens zwischen 5 und 10 Euro.

Second Generation Sequencing: Sequenzierung durch Synthese

Ist hingegen noch nicht bekannt, welcher DNA-Abschnitt für eine bestimmte Fragestellung wichtig ist, hilft einem die Sanger-Sequenzierung nicht wirklich weiter. In so einem Fall kann z. B. versucht werden, DNA-Abschnitte mit systematischen Unterschieden in den Genomen verschiedener Populationen einer Spezies zu finden. Für solche großen und aufwendigen Vergleiche eignen sich Sequenzierungsmethoden der zweiten Generation hervorragend.

Eine dieser Methoden ist die Sequenzierung durch Synthese. Als Vorbereitung für eine solche Sequenzierung werden die Genome der verschiedenen Populationen in kleine Fragmente gespalten. Dies kann mittels molekularer Scheren gezielt an bestimmten Sequenzen oder rein zufällig mittels der „Schrotschussmethode“ geschehen. Diese Fragmente werden anschließend der Größe nach aufgetrennt, und Fragmente mit einer gewünschten Größe werden anschließend an beiden Enden mit einem Forward-Adapter und einem Reverse-Adapter sowie einem Index versehen. Die Adapter dienen dazu, den Anfang des Fragments vom Ende zu unterscheiden und die Fragmente an einer bestimmten Oberfläche zu befestigen. Der Index ist eine kurze Nukleinbasen-Sequenz und wird für jede Population unterschiedlich gewählt. Dank der Indizes kann während der Analyse jedes Fragment einer bestimmten Population zugewiesen werden.

Für die Sequenzierung werden die Fragmente aller Populationen gemischt und auf eine Sequenzierplatte (Flow Cell) gegeben. Auf dieser binden sich die Adapter der Fragmente mit den dazugehörigen Gegenstücken. Anschließend wird vom Forward-Adapter aus ein neuer DNA-Strang synthetisiert. Die für die Synthese verwendeten Nukleinbasen sind speziell markiert und senden nach einer Aktivierung ein Leuchtsignal aus, das spezifisch für die jeweilige Nukleinbase ist. Während der Synthese liest permanent ein Computer aus, wo auf der Flow Cell welche Leuchtsignale in welcher Reihenfolge aufblitzen. Der Computer kann so die Forward-Sequenz aller Fragmente auf der Flow Cell bestimmen. Das Ganze findet anschließend noch einmal vom Reverse-Adapter aus statt.

Das Ergebnis der Sequenzierung sind die Forward-Sequenzen und die Reverse-Sequenzen aller Fragmente auf der Flow Cell, inklusive Index zur Bestimmung, welcher Population die Fragmente angehören. In einem Computerprogramm werden die Sequenzen anschließend den

Populationen zugeordnet und entlang eines bekannten Genoms angeordnet. So können die Sequenzen unter den Populationen verglichen werden, um Zielregionen und molekulare Marker zu finden.

Die genomweite Assoziationsstudie

Das Genom eines Lebewesens kann gigantische Ausmaße annehmen: von „nur“ ungefähr 150 Tausend Basenpaaren bis zu über 150 Milliarden Basenpaaren! Da kann es schwierig werden, die paar tausend Basenpaare zu finden, die dafür sorgen, dass Pflanze A von einer Spezies gut auf nährstoffarmen Böden gedeiht, während Pflanze B von derselben Spezies unter gleichen Bedingungen nicht ganz so fit aussieht. Um den Bereich, der für ein Merkmal verantwortlich ist, einzugrenzen, wird gerne die **genomweite Assoziationsstudie**, kurz GWAS verwendet. Beim GWAS werden bestimmte genetische Marker innerhalb einer Spezies mit dem Vorhandensein oder nicht Vorhandensein eines Merkmals in Verbindung gebracht.

Als genetische Marker werden häufig sogenannte **Single Nucleotide Polymorphismen (SNPs)** genutzt. Diese SNPs sind Stellen im Genom, an denen sich die DNA zwischen Individuen um ein Basenpaar unterscheidet. So wurde zum Beispiel bei manchen Individuen ein Adenin-Thymin-Basenpaar gegen ein Guanin-Cytosin-Basenpaar getauscht. Kommt diese SNP bei genug Individuen vor und sind in einer Spezies genug solcher Stellen bekannt, können diese als Wegpunkte oder Marker auf der DNA genutzt werden.

Die Anwesenheit dieser Wegpunkte bei einem Individuum wird anschließend mit dessen gemessenen Merkmalen abgeglichen. Durch Vergleich der Merkmale eines Individuums mit dessen Markern auf der DNA können anschließend Zusammenhänge (oder in Fachsprache „Assoziationen“) zwischen der Ausprägung eines Merkmals und dem Auftreten der genetischen Marker gezogen werden.

Ein klassisches Ergebnis eines GWAS-Experiments ist die Beobachtung, dass ein bestimmtes Merkmal häufiger zusammen mit bestimmten Markern auftritt als mit anderen Markern. Dies bedeutet, dass der Bereich, der für das Merkmal verantwortlich ist, sehr wahrscheinlich in der Nähe dieser Marker liegen muss. Der biologische Hintergrund dazu ist, dass während der Vererbung nicht für jedes Basenpaar zufällig entschieden wird, von welchem Elternteil es stammt, sondern dass größere Fragmente der DNA beider Eltern zusammengebaut werden. Das

bedeutet, dass Bereiche der DNA, die näher aneinander liegen statistisch häufiger zusammen vererbt werden, als Bereiche, die weit voneinander entfernt liegen.

Um beim Beispiel der Pflanzen auf nährstoffarmen Böden zu bleiben, könnte sich zeigen, dass Pflanzen mit den genetischen Markern X, Y und Z auf nährstoffarmen Böden meistens große Blätter bilden, wohingegen die Blätter von Pflanzen ohne diese Marker meistens kleiner sind. Diese Beobachtung bedeutet, dass der Bereich der DNA, der für die Blattgröße bei Nährstoffstress verantwortlich ist, irgendwo in der Nähe der genetischen Marker X, Y und Z liegen muss.

Um diese Assoziationen zuverlässig zu erstellen, ist allerdings vor allem eins nötig: jede Menge Individuen! Denn es soll natürlich vermieden werden, dass die beobachteten Zusammenhänge zufällig gesehen wurden. Darum gilt: jedes Individuum, bei dem diese Zusammenhänge entweder gesehen oder nicht gesehen werden, erhöht die Zuverlässigkeit des Ergebnisses. Das ist letztendlich immer so bei der Statistik: je mehr, desto besser! Man würde ja auch nicht die Durchschnittsgröße der Bundesbürger dadurch bestimmen, dass man nur fünf Personen misst, sondern Hunderte oder Tausende.

Für GWAS-Experimente, die auf Merkmale mit kleinen Unterschieden abzielen oder eine hohe Präzision brauchen, müssen darum mehrere Tausende bis Millionen Individuen untersucht werden. Dies ist mit entsprechendem Arbeits- und Zeit-Aufwand verbunden. Darum können von der Idee eines GWAS-Experimentes bis hin zu einem Ergebnis viele Jahre vergehen.

Abbildung 3 zeigt das Ergebnis einer GWAS, die die Tendenz zur Bildung von Nierensteinen in Menschen untersucht hat. Der gezeigte Graph ist ein Manhattan Plot, und jeder Punkt repräsentiert einen SNP, also einen genetischen Marker. Auf der X-Achse ist die Position der SNPs auf der DNA angegeben und auf der Y-Achse wie wahrscheinlich es ist, dass diese SNPs mit Nierensteinen assoziiert sind. Die rot gepunktete Linie definiert im Plot die Wahrscheinlichkeitsgrenze, so dass davon auszugehen ist, dass nur SNPs, die darüber liegen, etwas mit der gesuchten Eigenschaft zu tun haben (also Nierensteinen). In diesem Fall sind es SNPs in den 20 genannten Genen.

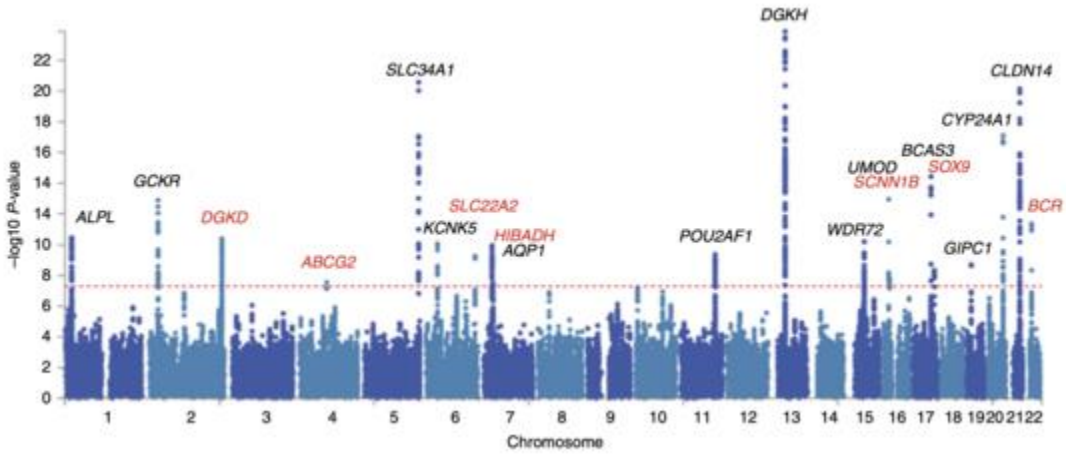


Abbildung 3: Ergebnis einer GWAS, welche genetische Marker für Nierensteine bestimmte (Abbildung: Wikipedia, genome wide association study).

Weiterführende Informationen

Vieles ist natürlich mittlerweile auf der deutschen und englischen Version von Wikipedia nachzulesen. Und weil Wikipedia einigermaßen gute Redakteure hat, sind die wissenschaftlichen Informationen dort normalerweise auch verlässlich, wenn manchmal auch unvollständig. Wo verlässliche wissenschaftliche Redakteure fehlen, so wie auf den meisten sozialen Medien, ist demzufolge die wissenschaftliche Information nicht zuverlässig und oft sogar völlig falsch und erfunden.

Eine der besten Quellen für Wissenschaftsinformationen sind weiterhin Textbücher, die das gesamte Wissen in einem Feld durch Expert:innen zusammenfassen. Obwohl Textbücher für Studierende geschrieben sind, sind sie natürlich nicht in allgemeinverständlicher Sprache geschrieben, und es bedarf etwas Mühe, sich durch die Informationen durchzuarbeiten.

Die Wissenschaftssprache auf Wikipedia und in Textbüchern ist also oft etwas schwer verständlich. Darum haben wir hier auch versucht, einige etwas umgangssprachlichere Beispiele aufzulisten (wobei diese Liste auf keinen Fall vollständig ist):

Vererbung erklärt bei Pflanzenforschung.de

<https://www.pflanzenforschung.de/de/pflanzenwissen/lexikon-a-z/vererbung-299>

Wissen und Fakten über Pflanzen bei Pflanzenforschung.de

<https://www.pflanzenforschung.de/de/pflanzenwissen/plantainments>

Grüne Gentechnik Video

<https://studyflix.de/biologie/grune-gentechnik-2659>

Auf dieser Seite sind noch viele andere Videos, die die Themen, die oben im Text erklärt werden, ebenfalls erklärt werden.

Gentechnik einfach erklärt: Methoden, Kritik und Gesetzeslage zu Grüner Gentechnik

<https://utopia.de/ratgeber/gentechnik-einfach-erklart-methoden-kritik-und-gesetzeslage-zu-gruener-gentechnik/>

Pflanzen genetisch verändern – die Verfahren im Überblick

<https://www.leopoldina.org/wissenschaft/gruene-gentechnik/gruene-gentechnik-verfahren/>

Die Werkzeug-Palette der Pflanzenzüchtung

<https://biooekonomie.de/themen/dossiers/die-werkzeug-palette-der-pflanzenzuechtung>

Genome Editing bei Kulturpflanzen: Nutzen und Gefahren der Genschere | Gut zu wissen | Bayerischer Rundfunk (BR)

https://www.youtube.com/watch?v=r0U_9Vwi32U

Gentechnik in der Landwirtschaft: Er wurde vom Fürsprecher zum Skeptiker | Unser Land | BR

<https://www.youtube.com/watch?v=DrAltBDtKjQ>