



Foto: Montage fotolia.com

Der Mensch und sein Genom

I. Welche Merkmale sind spezifisch für Menschen?

Moderne Sequenzierungsmethoden eröffnen den Zugang zu Informationen über das komplette Erbgut (Genom) von Organismen. Vergleiche zwischen dem Genom des Menschen und dem Genom anderer Organismen, vor allem von Menschenaffen werden häufig herangezogen, um bereits etablierte Vorstellungen bezüglich deren Evolution zu untermauern. Ein Überblick über die bisher bekannten Daten aus der Erforschung des Humangenoms offenbart viele Merkmale, die spezifisch für Menschen sind.

Harald Binder

Mit diesem Artikel wird eine Serie von Beiträgen eröffnet, die den aktuellen Stand der Erforschung des menschlichen Genoms[★] in unterschiedlichen Aspekten darstellen. Wir starten mit einer Zusammenfassung der charakteristischen Besonderheiten des menschlichen Genoms. Dabei wird auch die Abgrenzung zum Genom anderer Primaten angesprochen; das Thema Genomvergleich zwischen Menschen und den großen Menschenaffen wird jedoch in einem

gesonderten Beitrag eigens vorgestellt. Weitere Beiträge werden sich mit genetischen Phänomenen befassen, die üblicherweise mit evolutionären Vorgängen in Verbindung gebracht werden. Eines der vielen Projekte zur Entschlüsselung des gesamten menschlichen Erbguts ist die Suche nach Hinweisen für Funktionen in der gesamten DNA. Dazu wurden, bei aller Vorläufigkeit, erstaunliche Resultate veröffentlicht; auch darüber wird in folgenden Beiträgen berichtet.

Vom ersten DNA-Modell zur Genomforschung

WATSON und CRICK vereinigten 1953 die damals bekannten physikalisch-chemischen Daten zur stofflichen Grundlage der Erbinformation im berühmt gewordenen Modell der DNA-Doppelhelix (Abb. 1). Die Vorstellung der spezifischen komplementären Basenpaarung löste unmittelbar Ideen zu molekularen Mechanismen der Vererbungsvorgänge aus.

Die Einsicht, dass Nukleinsäuren als materielle Grundlage der genetischen Information in allen Lebewesen vorkommen, motivierte intensive Bemühungen, deren Strukturen aufzuklären und die Vorgänge zu verstehen, durch die die abgespeicherte Information genutzt wird. Das Erbgut des Menschen stand schon bald aus naheliegenden Gründen im Fokus des Interesses.

Das Projekt zur systematischen Erforschung des *gesamten* menschlichen Erbguts (= Genom) wurde 1990 unter maßgeblicher Beteiligung von James WATSON¹ in internationaler Kooperation begonnen. Als Strategie war bereits zu Beginn festgelegt worden, dass zunächst der Schwerpunkt auf die Entwicklung schnellerer und effektiverer Methoden zur Sequenzierung der DNA gelegt wird. Im Jahr 2001 wurde die DNA-Sequenz des Menschen in einer ersten Rohfassung durch das internationale *Human Genome Project* und durch die dazu in Konkurrenz arbeitende Gruppe um Craig VENTER (2001) veröffentlicht. Das Projektziel wurde mit der 2004 erschienenen Veröffentlichung – gemäß den für das Projekt formulierten Maßstäben – erreicht und seit dieser Zeit liegt offiziell das vollständig „entschlüsselte“ Erbgut des Menschen vor.²

Zur Untersuchung umfangreicher DNA-Sequenzen stehen heute sehr leistungsfähige Technologien zur Verfügung. Diese ermöglichen die Entschlüsselung der Genome von verschiedenen Lebewesen und deren Unterschiede zwischen Individuen einer Gruppe. In der vergleichenden Genomik versucht man durch die Erforschung der Unterschiede im Erbgut verschiedener Arten Erkenntnisse auch über deren vermutete Evolution zu gewinnen.

Nachdem das Genom des Menschen als Rohentwurf 2001 veröffentlicht worden war, mündeten die weiteren Erkenntnisse in die annähernd vollständige Beschreibung des Euchromatins* des Menschen (2004). Unter dem Euchromatin versteht man den größeren Teil des Erbguts, der sich in vergleichsweise aufgelockerter Form im Zellkern befindet und dort für zelluläre Vorgänge zur Verfügung steht. Das bedeutet, dass sich andere Moleküle an die DNA anlagern können, die als Translationsfaktoren bezeichnet werden. Diese Anlagerung stellt

Kompakt

Komplexe Organismen sind aus Zellen aufgebaut, die in unterschiedlichster Weise differenziert sind. Nach wenigen Zellteilungen in der frühen Embryogenese unterscheiden sich die einzelnen Zellen und entwickeln sich im weiteren Verlauf zu unterschiedlichen Gewebetypen. Bei allen Unterschieden weisen die Zellen aber auch für alle Zellen typische Strukturen und biochemische Prozesse auf. Daher benötigen Zellen auch hinsichtlich ihrer genetischen Ausstattung Erbinformationen, die man praktisch in allen biologischen Zellen erwarten würde, damit sie funktionieren können. Es ist daher nahe liegend, dass es auch bei ganz unterschiedlichen Lebewesen einen Anteil an genetischer Information gibt, der dafür notwendig ist, dass in den Zellen die grundlegenden biochemischen Prozesse ablaufen können und daher sehr große Ähnlichkeit aufweist. Worin unterscheidet sich aber das Erbgut unterschiedlicher Lebewesen und gibt es im Genom* des Menschen Bereiche, die spezifisch für ihn sind? In diesem Beitrag wird dargestellt, was in der Genomforschung bisher dazu gefunden worden ist und es werden auch die Herausforderungen dieser Untersuchungen aufgezeigt. Der enormen Menge an Genomdaten steht bisher ein sehr lückenhaftes Verständnis der Bedeutung einzelner Abschnitte des Erbguts für den jeweiligen Organismus gegenüber. Dieser Mangel begrenzt derzeit die Aussagekraft von Genomvergleichen. Die Forschung hat vor allem unter humanmedizinischen Gesichtspunkten bisher eine Menge genetischer Besonderheiten zutage gefördert, die spezifisch für Menschen sind. Ein spannender Aspekt der vergleichenden Genomforschung ist, ob ein tieferes Verständnis von Genomen verschiedener Organismen deren jeweilige Besonderheit vergrößert oder eher die Vorstellung von deren gemeinsamer Abstammung stützt. Bisherige Erfahrungen enthalten jedenfalls auch Hinweise darauf, dass bestimmte Erwartungshaltungen die Interpretation von Genomdaten, wie z. B. im Zusammenhang mit „junk-DNA“, erschweren kann.

einen der ersten Schritte dar, der dazu führt, dass auf DNA codierte Information in RNA umgeschrieben und damit in der Zelle genutzt werden kann (z. B. im Stoffwechsel). Dagegen findet man das Heterochromatin dicht gepackt vor, mit nur schwer zugänglicher DNA, hauptsächlich an der Membrannenseite des Zellkerns konzentriert. Über diesen Teil des Genoms ist bisher vergleichsweise wenig bekannt.

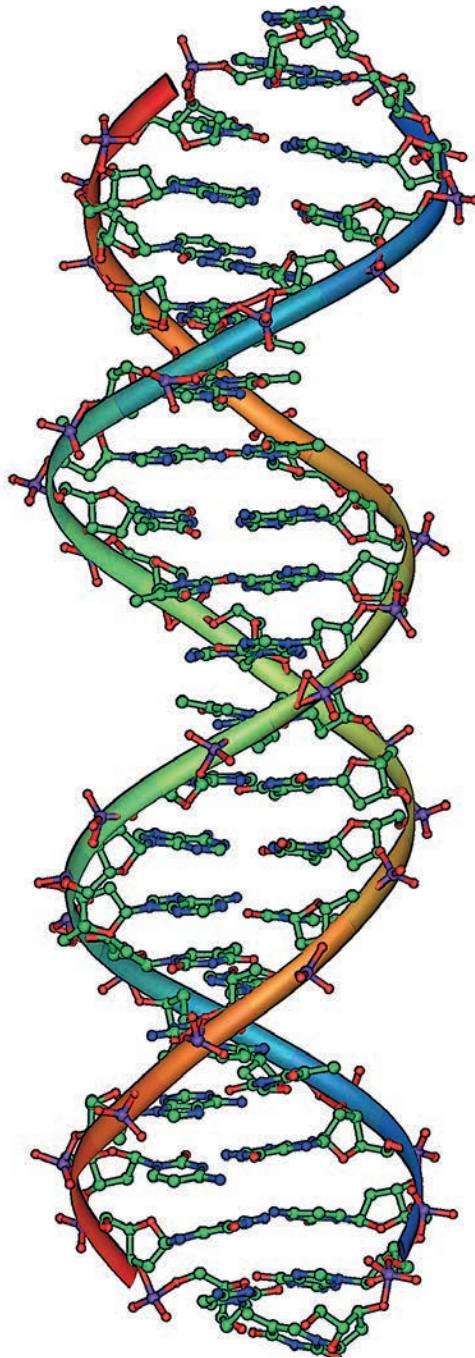
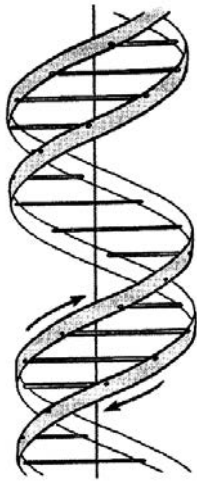
Die Sequenzierung der menschlichen DNA eröffnete für die medizinische Forschung die Möglichkeit, genetische Faktoren zu finden, die im Zusammenhang mit Krankheiten stehen. Weiter erhoffen sich viele Forscher aus den genetischen Daten von Menschen und anderen Lebewesen Hinweise darüber, welche genetischen Bereiche ursächlich für typisch menschliche Eigenschaften sind.

Bedeutende Fragen stellen sich auch nach dem Zusammenhang zwischen der Struktur des Erbguts, der DNA-Sequenz (also dem, was man auch als Genotyp bezeichnet) und dem Phänotyp, dem Erscheinungsbild bzw. der Gestalt eines Organismus.

Spezifische Merkmale des menschlichen Genoms

O'BLENESS et al. (2012) präsentieren in einem Überblicksartikel besondere Merkmale im menschlichen Genom, die spezifisch für Menschen sind, d. h. solche, die nur beim Menschen

Abb. 1 Rechts: Ein Modell der DNA-Doppelhelix (Wikimedia Commons); unten: die historische Darstellung aus der Veröffentlichung von WATSON & CRICK (1953). Die Helix ist 2 nm breit und eine Umdrehung 3,4 nm lang. (1 nm = 1 Millionstel Millimeter)



vorkommen und nicht bei anderen heute lebenden Primaten zu finden sind. Eine der Herausforderungen für derartige Studien ist, dass praktisch alle Genome nach wie vor nur unvollständig bekannt sind und die vorliegenden Ergebnisse Fehler aufweisen. Fehler entstehen z. B. beim Zusammenfügen von Teilsequenzen zum Genom oder beim Sequenzieren. Es ist also möglich, dass es sich bei ermittelten Unterschieden um tatsächliche Abweichungen handelt oder aber um besagte Fehler.

Zunächst führen die Autoren die großräumigen Veränderungen in den Chromosomen an, die bereits in den 1970er Jahren beschrieben worden waren. Die Zahl der Chromosomen unterscheidet sich bei Mensch (23) und Primaten (24). Bei Affen treten zwei kleine Chromosomen auf, die als 2 a und b bezeichnet werden und die, so

vermutet man, zum menschlichen Chromosom 2 fusioniert sein sollen.³ Außerdem konnte man mit verschiedenen Färbetechniken Hinweise auf Unterschiede in der Struktur einzelner Chromosomen finden, denn je nach Packungsdichte und Unterschieden in der Sequenz lassen sich bestimmte Abschnitte der Chromosomen verschieden anfärben. Auch im Bereich der Telomere*, den Enden der Chromosomen, die für deren Stabilität von Bedeutung sind und bei der Replikation* (Genverdopplung bei der Zellteilung) eine wichtige Rolle spielen, sind für Menschen typische Änderungen bekannt. Mehr als 60 duplizierte Segmente, die eine Größe von mehr als 65 000 Basen (> 65 kb) aufweisen, sind im menschlichen Genom abweichend von Affen beschrieben.

Neben den genannten Besonderheiten des menschlichen Genoms gibt es weitere, die im Vergleich zu den vorgenannten als „kleinräumig“ aufgefasst werden. Darunter versteht man den Austausch einzelner Basenpaare (single nucleotide polymorphism, SNP), Einschübe (Insertionen) und Verluste (Deletionen, beides zusammengefasst als Indels bezeichnet) von unterschiedlich großen DNA-Stücken, sowie Unterschiede in der Anzahl von Kopien einzelner Gene im Genom.

Die Angabe zum 1,2 %-Unterschied lässt diejenigen Bereiche des Genoms außer Betracht, die gar nicht miteinander verglichen werden können.

Bezogen auf den Austausch einzelner Nukleotide im Genom wurde der Unterschied zwischen Mensch und Schimpanse mit 1,2 % angegeben (The Chimpanzee Sequencing and Analysis Consortium 2005). Diese Angabe lässt aber gerade diejenigen Bereiche des Genoms außer Betracht, die aufgrund von Indels und unterschiedlichen Kopienzahlen gar nicht miteinander verglichen werden können (das Ausrichten der Sequenzen für den Vergleich – das „alignment“ – ist aufgrund der genannten Besonderheiten problematisch). Bezieht man diese Unterschiede mit ein, so unterscheiden sich die Genome von Mensch und Schimpanse um mehr als 5 % (BRITTEN 2002; vgl. zur ca. 1 %-Angabe auch COHEN 2007)⁴.

In einer ersten Studie, in der das gesamte Genom des Menschen und das aller vier großen Menschenaffen einbezogen worden war, dokumentierten FORTMA et al. (2004) 140 Gene, die spezifisch für den Menschen sind. Diese fanden sich interessanterweise an besonderen Positionen im Genom konzentriert, in der Nähe der Centromere, der Telomere und in Bereichen mit vielen Genduplikationen.

DNA-Abschnitte, die sich selbstständig aus einem Bereich des Genoms ausschneiden und an anderen Stellen wieder einfügen können (so genannte springende Gene oder Transposons), stellen einen weiteren bedeutenden Teil der spezifisch menschlichen Merkmale im Genom dar (vgl. Abb. 2). Bei Transposons unterscheidet man zum einen solche aus DNA-Sequenzen und zum anderen sogenannte Retrotransposons*, bei denen RNA-Sequenzen in DNA rückübersetzt und dann ins Genom integriert wurden. Retrotransposons können weiter unterteilt werden und zwar in solche, die lange Endstücke mit vielen Wiederholungselementen (long terminal repeat; LTR) enthalten und diejenigen, die diese nicht aufweisen. In Genomen finden sich auch Retrotransposons mit LTR, die gleichzeitig typische Merkmale von Viren aufweisen. Diese als endogene Retroviren* bezeichneten DNA-Sequenzen finden sich in spezifischer Verteilung im Genom von Menschen und bei den großen Menschenaffen. Allerdings sind die Unterschiede zwischen den Arten im Bereich der sich wiederholenden DNA-Sequenzabschnitte aufgrund der Schwierigkeiten, gerade diese Abschnitte zuverlässig zu sequenzieren, schwer zu bestimmen und entsprechend mit Unsicherheit behaftet.

Unter den für Menschen typischen Retrotransposons, die bereits sequenziert werden konnten, findet sich eine Untergruppe, die als LINE 1-Hs bezeichnet wird. Sie zeichnet sich dadurch aus, dass sie die einzigen Transposons darstellen, für die bisher im menschlichen Genom Aktivität gezeigt werden konnte. Ca. 80-100 Kopien kommen im Genom vor. Es gibt Hinweise darauf, dass diese Gene eine bedeutende Rolle bei der Flexibilität von Nervensystemen (neuronale Plastizität) spielen. Diese Gene sind während der Bildung von Nerven hochaktiv und tragen zur genetischen Vielfalt bei Nervenzellen, und damit zu der für jeden Menschen einzigartigen „Verdrahtung“ des Gehirns bei.

Nach O'BLENESS et al. belegen bereits diese Befunde, dass die Genome von Mensch und großen Menschenaffen weit unterschiedlicher sind als bisher angenommen wurde. Weiter schreiben sie, dass es schwierig ist, den genauen Umfang des für Menschen spezifischen Anteils des Genoms derzeit zu bestimmen. Dafür sind, bei aller Verlässlichkeit der bisherigen Kenntnisse, noch zu viele Lücken in komplexen Bereichen des Genoms vorhanden (die Lückenhaftigkeit der Kenntnis der Genome anderer Primaten ist noch größer).

Die Anzahl der für Menschen spezifischen Unterschiede im Genom ist in den vergangenen Jahren stark angestiegen und die Autoren mutmaßen, dass in Zukunft noch viele weitere entdeckt werden.



Zusammenhang zwischen Genen und Merkmalen

Die Bemühungen, phänotypische Unterschiede, d. h. Unterschiede im körperlichen Erscheinungsbild eines jeden Menschen, mit den genetischen Unterschieden zwischen den Menschen zu verknüpfen, stecken noch in den Anfängen. Derzeit bezieht sich etwa die Hälfte der bekannten Zusammenhänge auf die Gehirnstruktur und unsere Wahrnehmung. Die Autoren äußern allerdings Zweifel daran, dass dieser Wert repräsentativ ist. Sie vermuten dahinter eine entsprechende Vorliebe in der Forschung. Weitere Zusammenhänge sind aus den Bereichen Krankheitsabwehr und Immunsystem bekannt, beim Stoffwechsel, bei physiologischen und anatomischen Unterschieden sowie bei der Reproduktion und Entbindung beim Menschen.

Einzelne Beispiele. Im Erbgut können auch DNA-Sequenzen, die nur für einen Teilbereich eines Proteins (sog. Domäne) codieren, unabhängig vom restlichen Gen vervielfältigt werden. Ein bemerkenswertes Beispiel ist das Gen für die Domäne DUF1220. Die Domäne zeigt die größte Zunahme der Kopienzahl im Vergleich zum Schimpansen-genom im Bereich der für Proteine codierenden DNA*. Im menschlichen Genom finden sich 272 Kopien des DUF1220-Gens, während sich im Schimpansen-genom nur 126 Kopien, also weniger als die Hälfte davon finden (bei Nagern, wie z. B. Maus und Ratte liegt nur eine einzige Kopie vor). Aus Untersuchungen von krankhaft veränderten Gehirnen (Mikro- und Makrozephalie) liegen Hinweise auf einen

Abb. 2 Springende Gene bzw. Transposons beschrieb Barbara McClintock 1948 (Nobelpreis 1983) erstmals an bunten Maiskolben. Den Farbvariationen der Maiskörner liegen transponierbare genetische Elemente zugrunde. Solche mobilen DNA-Sequenzen findet man auch im Erbgut anderer Lebewesen einschließlich dem des Menschen. (Bild: <http://de.academic.ru/pictures/dewiki/99/corn-cobs.jpg>)

Die Anzahl der für Menschen spezifischen Unterschiede im Genom ist in den vergangenen Jahren stark angestiegen.

Glossar

Codierende DNA (cDNA): DNA-Abschnitte, die in RNA umgeschrieben und anschließend in Polypeptide übersetzt werden; also DNA, die für Proteine codiert. Weniger als 5 % der DNA im Genom des Menschen werden derzeit als cDNA angesehen.

Endogene Retroviren: DNA-Sequenzen, die sich wie springende Gene bzw. Transposons selbst aus dem Genom heraus schneiden und an anderer Stelle wieder integrieren können. Endogene Retroviren zeigen in ihrer DNA-Sequenz charakteristische Merkmale von Viren und werden vererbt.

Euchromatin: Im Euchromatin liegt die DNA gut zugänglich für verschiedene Proteine vor, die sich für die Nutzung und Umsetzung der darauf abgespeicherten Information daran anlagern müssen.

Genfamilie: Eine Gruppe von Genen, die auffällige Ähnlichkeiten in ihrer Sequenz aufweisen. Unter evolutionärem Gesichtspunkt werden sie auf einen ursprünglichen DNA-Abschnitt zurückgeführt, aus dem die Genfamilie dann durch Duplikationen und Mutationen entstanden ist.

Genom: Das gesamte Erbgut eines Organismus, das dieser an seine Nachkommen vererbt.

Heterochromatin: Ein Teil der DNA in den Chromosomen ist im Zellkern so kompakt, dass er z.B. für die Wechselwirkung mit Proteinen unzugänglich ist. Dieser Teil des Erbguts findet sich typischerweise innen an der Membran des Zellkerns.

Lektine: Kohlenhydratstrukturen können durch (Glyco-)Proteine erkannt werden, die man als Lektine bezeichnet. Lektine weisen keine enzymatische Aktivität auf.

Nichtcodierende DNA: Der weit überwiegende Teil der DNA codiert nicht für Proteine. Seine Funktion ist oft nicht einfach zu bestimmen. Viele Bereiche der nichtcodierenden DNA werden in RNA umgeschrieben und haben regulierende Funktion.

Pseudogen: Ein DNA-Abschnitt, der sehr große Ähnlichkeit mit einem funktionsfähigen Gen zeigt, bei dem aber durch Variation die bekannte und erwartete Funktion nicht mehr erfüllt wird. Für einzelne Pseudogene liegen Hinweise auf regulatorische Funktionen vor (z.T. gerade an den Sequenzabschnitten, denen sie ähnlich sind).

Replikation: Bei der Replikation wird der DNA-Doppelstrang vergleichbar einem Reißverschluss unter dem Einfluss von Enzymen geöffnet und entsprechend der komplementären Basenpaarung zu zwei DNA-Molekülen komplettiert.

Retrotransposon: Retrotransposons gehören zu den transponierbaren Elementen, also zu Genombereichen, die sich selbst aus dem Genom heraus schneiden und an anderen Stellen wieder integrieren können. Retrotransposons gehen dabei über RNA als Zwischenform und müssen für die Rückführung ins Genom wieder in DNA (mit dem Enzym Reverse Transkriptase) zurückübersetzt werden.

Telomer: DNA-Sequenz am Ende der Chromosomen, die charakterisiert ist durch sich wiederholende Elemente (TTAGGG). Telomere sind für die Stabilität der Chromosomen von Bedeutung und sind bei der Verdoppelung der Chromosomen beteiligt.

Zusammenhang zwischen der Kopienzahl von DUF1220 und der Gehirngröße vor: Je öfter die Domäne kopiert und ins Genom re-integriert wurde, umso stärker wächst das Gehirn (DUMAS et al. 2012).

Andere Beispiele betreffen genetische Veränderungen, die Aminosäuren im Protein austauschen. Hier nennen O'BLENESS et al. zwei Beispiele: Erstens COX5A, eine Untereinheit der Cytochrom C-Oxidase, ein Enzym, das in Mitochondrien am Ende der Elektronentransportkette zur Energiegewinnung beteiligt ist. Die Gene aller anderen Untereinheiten sind typischerweise hoch konserviert, d.h. sie zeigen keinerlei Variationen. Das wird als Hinweis dafür gesehen, dass die Funktion auch bei kleinsten Änderungen verloren geht. Beim Menschen sind aber in COX5A im Vergleich zu anderen Lebewesen an zwei Stellen Aminosäuren verändert. Leider ist die Funktion von COX5A

noch nicht vollständig verstanden, doch es gibt Hinweise darauf, dass diese Änderungen mit der Regulation des Fettstoffwechsels zu tun haben.

Im zweiten Fall handelt es sich um einen bedeutsamen Aminosäureaustausch in FOXP2, einem Protein, bei dem ein Einfluss auf die Entwicklung der Sprachfähigkeit des Menschen vermutet wird.

Pseudogene. Als Pseudogene* werden DNA-Abschnitte bezeichnet, die auffallende Ähnlichkeiten zu funktionierenden Genen haben, aber aufgrund irgendwelcher Veränderungen ihre Funktion nicht mehr erfüllen. Pseudogene werden oft als Relikte aus der evolutionären Vergangenheit dargestellt (vgl. Abb. 3). Aufgrund veränderter Bedingungen sei ihre Funktion nicht mehr notwendig. An dieser Stelle sei aber erwähnt, dass inzwischen für eine Reihe von Pseudogenen doch Hinweise auf eine Funktion gefunden wurden (PINK et al. 2011).

KIM et al. (2010) beschreiben 38 für Menschen spezifische Pseudogene, davon liegen 9 als Einzelkopien vor. Eines davon ist Apolipoprotein C1 (APOC1), das am Lipoprotein-Stoffwechsel beteiligt ist. Im Genom der großen Menschenaffen finden sich zwei leicht unterschiedliche Gene für zwei APOC1-Proteinvarianten. Eine Variante wird im menschlichen Genom aufgrund eines Stopp-Codons nicht hergestellt, so dass Menschen nur eine Form von APOC1 besitzen. Die Konsequenzen dieser für den Menschen spezifischen Ausprägung im Genom sind derzeit noch nicht verstanden; die Autoren diskutieren gesundheitliche Zusammenhänge.

Besonders auffällig sind Pseudogene im Zusammenhang mit dem Geruchssinn. In der Familie der Geruchsrezeptoren werden beim Menschen im Vergleich zu den Großaffen mehr als 60 % der Gene als nicht funktional betrachtet, d.h. als Pseudogene. Die funktionierenden Gene für Geruchsrezeptoren beim Menschen zeigen oft eine erhöhte Kopienzahl, d.h. sie liegen mehrfach im Genom vor.

Während der Replikation können Fehlarrangements von Genen auftreten. Dabei werden z.B. Teile eines Gens an ein anderes Gen angefügt. Solche Vorgänge führen erwartungsgemäß zu Pseudogenen. Es wurden aber Beispiele beschrieben, in denen solche Änderungen zu einem funktionsfähigen Gen führen. Ein Beispiel ist das Protein, welches spezifisch an Sialinsäure bindet (Sialinsäure bindendes Ig Superfamilie Lektin 11). Im Vergleich zu anderen Lebewesen liegt es beim Menschen in veränderter Form vor. Wie es scheint wurde am vorderen Ende des Proteins ein ehemals angrenzendes Pseudogen angefügt, so dass ein neuer Promotor und eine neue Proteinsequenz am vorderen Teil des Proteins eingeführt

wurden. Dadurch ist die spezifische Bindung modifiziert und das Protein bindet jetzt an andere Moleküle. Durch den neuen Promotor wird die Bildung dieses neuen Lektins* nach bisherigen Studien ausschließlich in den Mikrogliazellen initiiert, das sind Zellen, die für die Immunabwehr im Zentralnervensystem verantwortlich sind. Die modifizierten Mikrogliazellen spielen beim Überleben von Nervenzellen eine Rolle.

Änderungen in Genfamilien. Eine Genfamilie* umfasst einander Sequenz-ähnliche Gene, die – so die Vermutung – im Laufe der Evolution dadurch entstanden, dass ein Gen kopiert wurde, die Kopie sich an anderer Stelle ins Genom einfügte und dann beide Gene (Original und Kopie) sich durch Mutationen unterschiedlich entwickelten. So kann eine Genfamilie durch zusätzliche Genkopien bereits vorhandener Gene erweitert werden. Die oben angeführte Domäne DUF1220, die bereits an sich eine erhöhte Kopienzahl aufweist, trug vermutlich auch zu vier weiteren Sequenz-veränderten Genkopien bei.

Eine Genfamilie umfasst die Homöobox-Gene. Sie codieren für Proteindomänen, die an die DNA binden und dort Schalterfunktionen ausführen. Typischerweise schalten sie ganze Kaskaden von nachgeordneten Genen ein oder aus. Ein berühmtes Beispiel ist das Gen Pax6; wenn dieses aktiviert wird, beginnt die Entwicklung des Auges. Homöobox-Gene einer Unterfamilie der „double homeobox“ oder DUX-Familie sind an der Regulation der Transkription und bei der Entwicklung des Embryos beteiligt. Drei Kopien (DUXY2, DUXY3, DUXY4) liegen auf dem menschlichen Y-Chromosom und ein sonst typischerweise vorhandenes Stopp-Codon fehlt, so dass am Ende des eigentlichen DUX-Gens die Aminosäurekette bis zum nächsten Stopp-Codon erweitert wurde. Diese drei Gene haben beim Menschen eine neue, bisher noch nicht bekannte Funktion.

In seltenen Fällen sind beim Menschen einzelne Vertreter von Genfamilien vollständig inaktiviert. Dies konnte am Beispiel eines Sialinsäure erkennenden Lektins (SIGLEC13) gezeigt werden.

Neben der bereits erwähnten extrem gestiegenen Kopienzahl für DUF1220 im menschlichen Genom machen solche Veränderungen in den Kopienzahlen von Genen den Hauptunterschied der Genome von Menschen und der großen Menschenaffen aus. So liegen z. B. vom Gen für Aquaporin7 mehrere zusätzliche Kopien im menschlichen Genom vor. Bei Aquaporin handelt es sich um in die Zellmembran integrierte Proteine, die dort den Transport von Wassermolekülen und weiteren kleinen Molekülen (wie z. B. Glycerin) durch die Membran

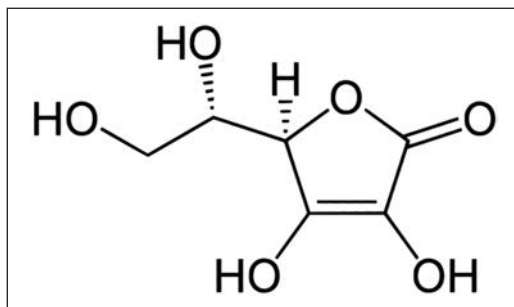


Abb. 3 Der Mensch muss seinen Bedarf an Vitamin C (Ascorbinsäure; Abb. links) durch Nahrungsaufnahme (z. B. Zitrusfrüchte) decken, da er kein funktionsfähiges Gen für das Enzym L-Gulonolacton-Oxidase besitzt, das für die Synthese von Ascorbinsäure erforderlich ist. Es gibt Lebewesen, die dieses Gen besitzen (z. B. die Maus) und Ascorbinsäure biosynthetisch herstellen. Ein DNA-Abschnitt im Genom des Menschen, der dem Gen bei der Maus ähnlich ist, wird als Pseudogen bezeichnet. Neben dem Menschen sind eine ganze Reihe von Lebewesen auf Vitamin C-Quellen angewiesen. Foto: Alterfalter, fotolia.com

erleichtern. Sie stehen im Zusammenhang mit der Möglichkeit von Zellen, Fett für die Deckung des Energiebedarfs zu nutzen. Darüber hinaus wird über die Rolle von Aquaporinen bei der Anpassung des menschlichen Stoffwechsels an ausdauerndes Laufen und der Thermoregulation beim Schwitzen diskutiert. Beides wird von manchen Autoren als bedeutsam für die Entwicklung des Menschen zum ausdauernden Jäger vermutet.

Neue Gene. Im menschlichen Genom finden sich wie in jedem Lebewesen Gene, die nur beim Menschen vorkommen und die nicht aufgrund von Sequenzähnlichkeiten auf Gene anderer Lebewesen zurückgeführt werden können. Damit müssen diese Gene, geht man von Evolution aus, neu entstanden sein (*de novo* origin). Dabei wird das Kriterium, nach welchem ein Gen als neu bezeichnet werden soll, durchaus kontrovers diskutiert. Für diese Gruppe „neuer“ Gene wurden bisher erst in einem Fall Hinweise auf eine Funktion beschrieben. Weitere Untersuchungen müssen zeigen, ob sich solche „neuen“ Gene bestätigen lassen und in welchem Umfang sie die Besonderheit des menschlichen Genoms prägen.

Gene und Krankheiten. Viele der hier angeführten Besonderheiten, die spezifisch für das menschliche Genom sind, wurden im Zusammenhang mit Krankheiten erforscht. Der Zusammenhang zwischen „Gen“ und „Merkmal“, in diesem Fall Krankheit, ist aber äußerst schwierig zu bestimmen. Nur in ganz wenigen Fällen liegt der Zusammenhang klar auf der Hand: Beispielsweise haben Menschen in Malaria-gebieten oft ein verändertes Hämoglobin, welches Immunität gegen Malaria verleiht. Welche Genvarianten aber z. B. mit einem erhöhten Alzheimer-Risiko einhergehen, ist weit weniger klar, denn oft werden genetisch bedingte Krankheiten durch ein Zusammenspiel vieler daran beteiligter Gene ausgelöst. Ist die Korrelation zwischen Krankheitsbild beim Menschen und dessen genetischer Struktur wirklich eine ursächliche Wirkung des Genoms auf den Gesundheitszustand? Das lässt sich heutzutage in den allermeisten Fällen kaum beantworten.

Hoffnung auf weitere Forschung

O'BLENESS et al. versprechen sich durch die fortlaufend gesteigerte Leistungsfähigkeit der DNA-Sequenzierungstechnologien eine exponentiell wachsende Zahl von Genomdaten und damit verbunden eine weitere Klärung darüber, welche Bereiche des Genoms spezifisch für Menschen sind. Dazu werden aber vor allem auch Methoden notwendig sein, die gerade diejenigen Bereiche des Genoms erschließen, die sich bisher nicht oder nur mit großem Aufwand sequenzieren lassen, aufgrund der variablen Bereiche, in denen kurze Nukleotidfolgen sich häufig wiederholen. Diese Genom-Abschnitte sind aus Sicht der Review-Autoren beklagenswert unbekannt. Durch den immer stärkeren Einsatz von sogenannten „next generation sequencing“ Technologien, bei denen kleine Genomschnipsel kostengünstig sequenziert werden können, sind zuletzt auch die Genome vom Zwergschimpanse (Bonobo) und Gorilla zugänglich geworden. Computer setzen die Genomschnipsel wie ein Puzzle wieder zusammen. Durch die vergleichsweise kurzen Fragmente, die sequenziert und dann zusammengesetzt werden, sind aber gerade die genannten Problembereiche der Genome (immer und immer wiederholte Kopien ein und derselben Sequenz) nicht bestimmbar. Das bedeutet, dass die veränderte Anzahl solcher sich wiederholenden Sequenzabschnitte, die ja einen großen Anteil der Unterschiede von Genomen zwischen Mensch und andere Primaten ausmachen, nicht genau bekannt ist.

Auf ein weiteres Problem weisen O'BLENESS et al. im Bereich der sogenannten *de novo* Gene

Der Zusammenhang zwischen „Gen“ und Krankheit ist äußerst schwierig zu bestimmen.

hin. Die Schwierigkeiten beim Versuch, die Resultate erster Studien in Kontrollen zu reproduzieren, unterstreichen die Notwendigkeit von klaren Anleitungen für zuverlässige und vergleichbare Charakterisierungen von Genen in den Genomen der jeweiligen Organismen. Da das Genom des Menschen am detailliertesten untersucht ist, wird es typischerweise als Orientierungshilfe genutzt, wenn die Zusammensetzung eines anderen Genoms aus dessen sequenzierten Fragmenten ermittelt wird. Die nicht immer einheitlichen Beschreibungen von Genen im Genom des Menschen führen zu Schwierigkeiten und Widersprüchen, wenn bei der Rekonstruktion von Genomen anderer Organismen das des Menschen zugrundegelegt wird. Zu guter Letzt erwarten O'BLENESS et al. durch die Nutzung von Genomdaten ein besseres Verständnis dessen, was die Menschen ausmacht und wie sie geworden sind.

Vorläufige Schlussfolgerungen

Die hier angeführten Befunde belegen, dass eine ganze Reihe von Merkmalen im Genom des Menschen nach bisherigen Erkenntnissen für diesen spezifisch und einzigartig ist. Der Mensch kann also anhand genetischer Merkmale klar von anderen Lebewesen abgegrenzt werden. In den Fachpublikationen werden diese unterschiedlichen Merkmale typischerweise vor dem Hintergrund einer gemeinsamen Entwicklungsgeschichte mit anderen Primaten interpretiert und so dargestellt, als ob sie diese belegen. Die Festlegung auf diese Perspektive ist jedoch lange vor den Untersuchungen von Genomen erfolgt und hat die Interpretation der Resultate in der Genomforschung nachhaltig geprägt.

Der Mensch kann anhand genetischer Merkmale klar von anderen Lebewesen abgegrenzt werden.

Der größte Teil der Erbinformation ist derzeit nur als formale Sequenzdaten verfügbar. Nur für einen kleinen Teil davon ist experimentell untersucht und verstanden, ob der entsprechende Bereich eine Funktion hat und wenn ja welche. Aus der Erforschung verschiedener Genome liegen bereits Erfahrungen vor die zeigen, dass scheinbar naheliegende Erklärungen, die von der genannten Evolutionsvorstellung geleitet sind, zu falschen Annahmen führten

und so die Bearbeitung wichtiger Fragen (z. B. im Zusammenhang mit den sogenannten „Pseudogenen“) verzögerten. So haben PINK et al. (2011) in einem Review die bisher bekannten Befunde zusammengestellt, die zeigen, dass zumindest einige Pseudogene entgegen der zunächst unterstellten Funktionslosigkeit sehr wohl in RNA transkribiert werden und die ähnlichen, für Proteine codierenden Bereiche regulieren können

Zunächst jedoch bedarf es großer Anstrengungen, „das Genom“ – sowohl des Menschen als auch das anderer Lebewesen – noch besser zu verstehen.

Anmerkungen

- ¹ WATSON leitete das Human Genome Projekt vom Beginn 1990 bis 1992; 1993 übernahm Francis COLINS die Leitung des Projekts bis zu dessen Abschluss.
- ² Im Internet bieten verschiedene Institutionen umfangreiche Plattformen zur Erforschung des menschlichen Genoms an, die ständig weiter entwickelt und gepflegt werden: <http://genome.ucsc.edu/>; <http://www.ensembl.org/>; <http://www.ncbi.nlm.nih.gov/genome/guide/human/>
- ³ Dabei finden sich in der DNA-Sequenz in dem entsprechenden Bereich vom menschlichen Chromosom 2 allerdings nur wenige Merkmale, die man erwarten würde, wenn zwei funktionsfähige Chromosomen an ihren beiden Kopfen miteinander fusionieren.
- ⁴ EBERSBERGER et al. (2007) schreiben in einer vergleichenden Studie von Mensch, Schimpanse, Gorilla, Orang Utan und Rhesusaffen von ca. 23 % des menschlichen Genoms, in denen der Mensch keine direkte genetische Abstammung mit Schimpansen aufweist. Die Autoren folgern aus ihren Ergebnissen, dass etwa 1/3 unsere Gene begonnen haben, sich in einer für Menschen spezifischen Weise zu entwickeln, bevor sich die Linien von Mensch, Schimpanse und Gorilla getrennt haben.

Literatur

- BRITTEN RJ (2002) Divergence between samples of chimpanzee and human DNA sequences is 5%, counting indels. *Proc. Natl. Acad. Sci USA* 99, 13633–13635.
- COHEN J (2007) Relative differences: the myth of 1%. *Science* 316, 1836.
- EBERSBERG I, GALGOCZY P, TAUDIEN S, TAENZER S, PLATZER M & VON HAESELER A (2007) Mapping human genetic ancestry. *Mol. Biol. Evol.* 24, 2266–2276.
- DUMAS LJ, O'BLENESS MS, DAVIS JM, DICKENS CM, ANDERSON N, KEENEY JG, JACKSON J, SIKLELA M, RAZNAHAN A, GIEDD J, RAPPORT J, NAGAMANI SSC, EREZ A, BRUNETTI-PIERI N, SUGALSKI R, LUPSKI JR, FINGERLIN T, CHEUNG SW & SIKELA JM (2012) DUF1220-Domain copy number implicated in human brain-size pathology and evolution. *Am. J. Hum. Gen.* 91, 44–454.
- FORTNA A, KIM Y, MACLAREN E, MARSHALL K, HAHN G, MELTESEN L, BRENTON M, HINK R, BURGER S, HERNANDEZ-BOUSSARD T, KARIMPOUR-FARD A, GLUECK D, MCGAVRAN L, BERRY R, POLLACK J & SIKELA JM (2004) Lineage-specific gene duplication and loss in human and great ape evolution. *PLOS Biol.* 2, 937–954.
- KIM HL, IGAWA T, KAWASHIMA A, SATTI Y & TAKAHATA N (2010) Divergence, demography and gene loss along the human lineage. *Phil. Trans. R. Soc. B* 365, 2451–2457.
- O'BLENESS M, SEARLES VB, VARKI A, GAGNEUX P & SIKELA JM (2012) Evolution of genetic and genomic features unique to the human lineage. *Nature Rev. Gen.* 13, 853–866.
- PINK RC, WICKS K, CALEY DP, PUNCH EK, JACOBS L & CARTER DRF (2011) Pseudogenes: Pseudo-functional or key regulators in health and disease? *RNA*; doi: 10.1261/rna.2658311.
- ROMERO IG, RUVINSKY I & GILAD Y (2012) Comparative studies of gene expression and the evolution of gene regulation. *Nature Rev. Genet.* 13, 505–516.
- The Chimpanzee Sequencing and Analysis Consortium (2005) Initial sequence of the chimpanzee genome and comparison with the human genome. *Nature* 437, 69–87.
- UDDIN M (2008) Molecular evolution of the cytochrome c oxidase subunit 5A gene in primates. *BMC Evol. Biol.* 8, 8; doi:10.1186/1471-2148-8-8
- WU DD, IRWIN DM & ZHANG YP (2011) De novo origin of human protein-coding genes. *PLoS Genet.* 7, e1002379.

Anschrift des Verfassers:

*Dr. Harald Binder, Taborweg 8, 78467 Konstanz
email: h-e.binder@t-online.de*