



# Die neue Zweisamkeit im Erbgut



Sequenziert ja – aber entschlüsselt? Wirklich verstanden ist das Erbgut des Menschen noch nicht. Die Lösung vieler Rätsel liegt in der diploiden Natur des Genoms, das Chromosomensätze von Vater und Mutter vereint. **Margret Hoehe** vom **Max-Planck-Institut für molekulare Genetik** in Berlin hat erstmals beide Versionen des Erbguts eines Menschen getrennt sequenziert und festgestellt: Das Individuum ist individueller als gedacht.

TEXT **CATARINA PIETSCHMANN**

**W**enn Margret Hoehe von ihrer Arbeit erzählt, kommt einem bald der Gedanke, wie logisch ihr Vorgehen ist, wie offensichtlich. Und warum das eigentlich so nicht längst schon jemand gemacht hat. Vielleicht weil, wer zu dicht dran ist, leicht den Blick für das Gesamtbild verliert: Er schaut durch eine Lupe, blickt auf eine zerklüftete graue Landschaft und weiß nicht, was da vor ihm liegt. Tritt ein paar Schritte zurück, will man ihm zurufen, leg die Lupe beiseite! Dann siehst du, dass vor dir ein ausgewachsener Elefant steht.

Margret Hoehe ist eine von denen, die ein paar Schritte zurückgegangen sind, um besser zu sehen. Sie hat sich der genetischen Grundlagen besonnen, ohne die Gregor Mendels bereits 1866 aufgestellte Regeln zur Vererbung von Merkmalen nicht zu verstehen sind und die in keinem Biologieunterricht fehlen. Er hatte reinerbige Erbsenpflanzen mit roten und weißen Blüten miteinander gekreuzt. Die Pflanzen der Tochtergeneration hatten alle gleichermaßen rote Blüten, denn das Gen für die Blütenfarbe Rot war dominant. In der Enkelgeneration traten dann wie-

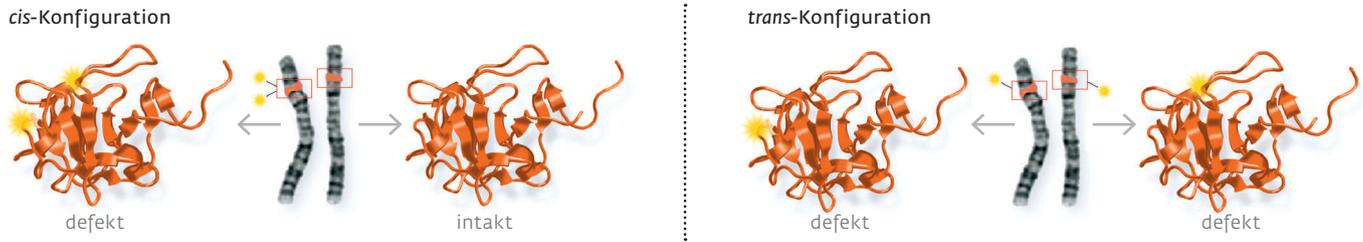
der zu bestimmten Anteilen rote und weiße Blüten auf – ein Phänomen, das nur durch die Existenz zweier Chromosomensätze erklärt werden kann.

#### GENE VON VATER UND MUTTER

Auch wenn die Kreuzungen beim Menschen nicht immer so klar berechenbar sind, eines haben wir Menschen mit vielen anderen Organismen gemein: Jedes Individuum besitzt ein Genom, das aus zwei verschiedenen Chromosomensätzen besteht. Daher kommen auch alle Gene doppelt vor: einmal auf dem entsprechenden Chromosom der Mutter und einmal auf dem des Vaters. Die unterschiedlichen elterlichen Sequenzversionen eines jeden Chromosoms werden von Fachleuten auch als Haplotypen bezeichnet.

Weniger bekannt ist, dass der amerikanische Biophysiker Seymour Benzer bereits 1957 in einem simplen Experiment gezeigt hat, dass es einen Unterschied macht, ob zwei unterschiedliche Mutationen eines Gens auf demselben Chromosom sitzen oder ob sie auf beide Chromosomen verteilt sind. Die Entdeckung Benzers geriet wieder in Vergessenheit. >

Chromosomen im Moment der Zellteilung: Das menschliche Erbgut besteht aus zwei Chromosomensätzen, einem vom Vater und einem von der Mutter. Von jedem Chromosom gibt es also zwei Varianten. Eine spezielle Färbemethode erzeugt für jedes Chromosom ein charakteristisches Bandenmuster, sodass die Paare an ihrer vergleichbaren Größe, Form und Farbe identifiziert werden können.



Die Verteilung von Mutationen in einem Chromosomenpaar kann über Gesundheit und Krankheit entscheiden: In der *cis*-Konfiguration treten zwei Mutationen in ein und derselben Genkopie auf, das zugehörige Protein wird dadurch funktionsuntüchtig. Die zweite Kopie und das davon abgelesene Protein bleiben unbehelligt. In der *trans*-Konfiguration sind dagegen beide Genkopien mutiert und produzieren zwei geschädigte Proteine.

Wäre es da nicht sinnvoll, sich die beiden Versionen des Genoms getrennt anzusehen? Bislang wurde ein Genom ausgelesen, quasi ein Mischprodukt aus den mütterlichen und väterlichen Sequenzen. Dies hat historische Gründe: Zu den Zeiten des Humangenomprojekts in den 1990er-Jahren war es technisch kaum möglich, die rund drei Milliarden Basenpaare des menschlichen Genoms in doppelter Ausführung zu analysieren. Dies wäre einfach zu aufwendig gewesen.

### ERBGUT IN ZWEIFACHER AUSFÜHRUNG

Gewohnheiten, die sich eingeschliffen haben, lassen sich nur schwer ablegen. „Selbst hochrangige Experten haben bis vor Kurzem beim Genom *einfach*, nicht *zweifach* gedacht“, sagt Margret Hoehe. Sie leitet am Berliner Max-Planck-Institut die Arbeitsgruppe „Genetische Variation, Haplotypen und Genetik komplexer Erkrankungen“ und hat als Erste das Erbgut eines Menschen – eines Deutschen – getrennt nach Haplotypen sequenziert. Sie hat dabei die beiden Chromosomensätze vollständig und in bisher unerreichter Genauigkeit analysiert. Die Vorgehensweise, die sie dafür zusammen mit ihren Mitarbeitern entwickelt hat, ist vom Fachmagazin *NATURE* zu einer der vielversprechendsten Methoden des Jahres 2011 gekürt worden.

„Moment, gleich haben wir's. Es muss doch möglich sein...!“ Auf ihrem großen Schreibtisch sucht Hoehe nach

Platz für zwei Teetassen und die Notizen ihrer Besucherin. Was nicht so leicht ist, denn Literaturquellen und Datensätze stapeln sich nahezu auf dem ganzen Schreibtisch zu Türmen auf. Seit August ist die Forscherin fast nur noch am Schreiben, um alle Analysen zu Papier zu bringen.

Warum ist es wichtig zu wissen, wie bestimmte Mutationen zwischen beiden Erbanteilen verteilt sind? „Weil es zum Beispiel den Unterschied machen kann zwischen Krebs oder nicht Krebs“, sagt sie. „Finden sich zwei Mutationen – etwa die mit Brustkrebs assoziierten Mutationen des Risikogens *BRCA1* –, so müssen diese nicht zwangsläufig krank machen.“ Denn es gibt zwei Möglichkeiten: Die Mutationen betreffen beide Formen des Gens (*trans*), dann ist die Wahrscheinlichkeit, an Brustkrebs zu erkranken, sehr hoch. Oder aber sie finden sich nur auf einem Chromosom (*cis*), dann hat der Mensch auch eine völlig „gesunde“ Genkopie.

*Cis* oder *trans*? Die Antwort auf diese Frage kann über gesund oder krank, ja sogar Leben oder Tod entscheiden. Und Hoffnung machen: Es kann schließlich auch eine „gesunde“ Form des Gens an die Nachkommen weitergegeben werden.

Gut zehn Jahre ist es nun her, dass Craig Venter und Francis Collins – erst Kollegen, dann Konkurrenten – Seite an Seite im Oval Office bei Bill Clinton die Entschlüsselung des menschlichen Genoms verkündeten. Das Ende der Unwissenheit, wie Venter sagte. Was für ein Triumph! Kennen wir den genetischen

Code, werden wir auch bald die Funktion aller Gene erkennen. Wir werden die Entstehung von Krankheiten verstehen – und sie heilen können. So weit die Hoffnung.

In den Jahren danach machte die Medizin zweifellos Fortschritte, der große Durchbruch blieb bisher allerdings aus. Das könnte sich jetzt mit Margret Hoehes Haplotypen-Projekt ändern.

Das Alphabet der DNA besteht nur aus den vier Buchstaben ACGT. Sie stehen für die Basenmoleküle Adenin, Cytosin, Guanin und Thymin. Rund 90 Prozent der genetischen Unterschiede zwischen Menschen sind Stellen, an denen ein Buchstabe des genetischen Codes – also eine Base – durch einen anderen ersetzt wurde. Die Forscher nennen diese Basenaustausche auch *Single Nucleotide Polymorphisms* (SNPs). „Jeder Mensch besitzt im Schnitt einen Basenaustausch auf 1700 Buchstaben“, erzählt Hoehe.

### FEHLER IM GENETISCHEN CODE

In der Reihenfolge der Basen liegt jedoch die Information für die Bildung von Proteinen aus einzelnen Aminosäuren. Deshalb kann beispielsweise der Austausch eines Cytosins durch ein Thymin zum Einbau einer anderen Aminosäure führen. Die mögliche Folge: Das resultierende Protein ist funktionsunfähig. Werden Basen in Kontrollregionen der DNA ausgetauscht, können biochemische Vorgänge im Körper blockiert oder beschleunigt sein.

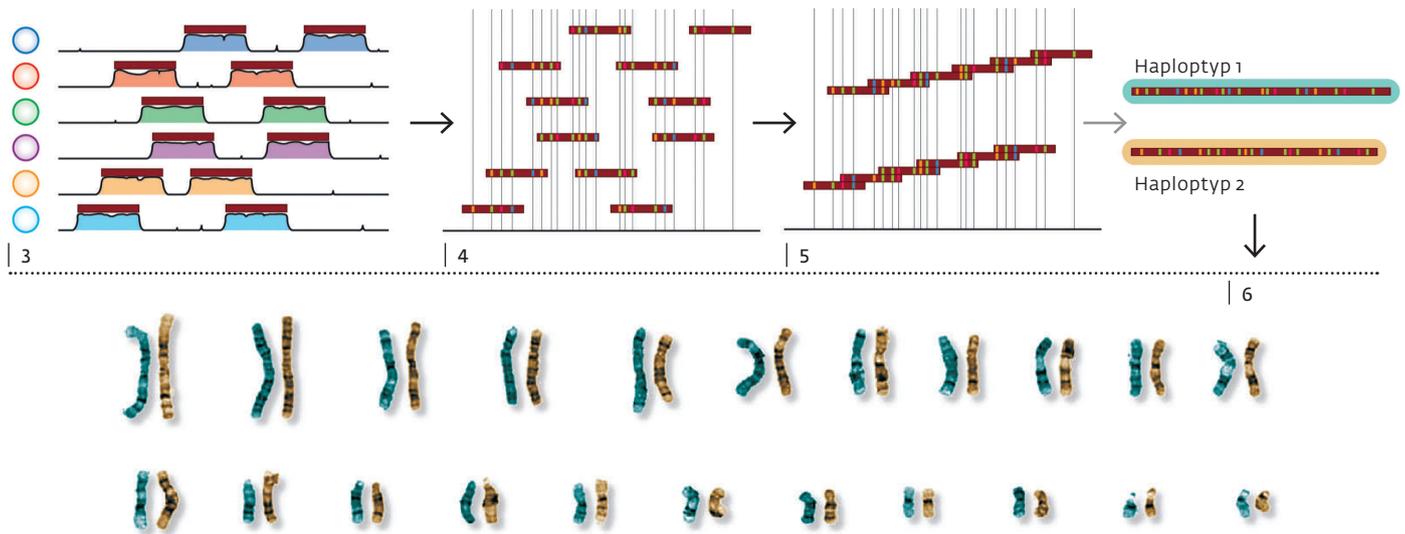
Bibliothek im Gefrierschrank: Die in bakterielle DNA-Moleküle eingefügten menschlichen Erbgutfragmente werden bei minus 80 Grad aufbewahrt. Jede der in Aluminiumfolie verpackten Genom-Bibliotheken stammt von einer Person. Das Erbgut von einhundert Menschen füllt so einen ganzen Gefrierschrank.

Für die Sequenzierung isolieren Forscher DNA aus weißen Blutkörperchen, zerschneiden sie mit Ultraschall und sortieren die Bruchstücke nach Größe. Dann werden die DNA-Schnipsel mithilfe kleiner Ankerküle auf einer Oberfläche fixiert und vervielfältigt. Das Auslesen der Basensequenz erfolgt zeitgleich für Millionen von DNA-Fragmenten. Heute braucht man nur noch wenige Tage, um mit den Sequenziermaschinen der zweiten Generation ein komplettes Genom auszulesen. „Die modernsten Geräte schaffen in einem Durchlauf bis zu 300 Milliarden Basen“, sagt Hoehe. „Angefangen hat es mal mit 900 Basen.“ Ein Hochleistungsrechner speichert Billionen dieser Sequenzabschnitte. Dann beginnt das große Puzzeln: Er vergleicht die Basenabfolge der einzelnen Schnipsel und legt sie überlappend zu größeren Stücken aneinander, bis die DNA komplett ist.

Woher weiß die Software, was an welche Stelle gehört? „Sie vergleicht die Basenfolge der Schnipsel permanent mit jener der Referenzsequenz“, erklärt Hoehe. „Aber genau das ist der kritische Punkt.“ Denn diese Referenz – das Humangenom von 2001, auf das sich bis heute alle beziehen – ist ein Kunstprodukt. Ein Cocktail aus dem Erbgut von mindestens sechs verschiedenen Menschen. Dabei ist jeder Mensch einzigartig. *Die* Referenzsequenz existiert von Natur aus nicht. Sie gibt also nicht den genetischen Code einer bestimmten lebenden Person wieder. Das war Absicht, aber viele Experten halten es inzwischen für einen Fehler. Denn das macht







Basen über Generationen gemeinsam vererbt werden, ist bei ihnen am kürzesten, im Schnitt etwa 10000 Basen“, erzählt Margret Hoehe. Bei Europäern sind es 20 000 bis 120 000 und mehr Basen. „Offenbar verließ nur eine kleine Gruppe ihren Ursprungskontinent Afrika, um den Rest der Welt zu bevölkern, und nur wenige zogen nach. Die genetische Vielfalt war also deutlich reduziert – ein genetischer Flaschenhals sozusagen.“

**AFRIKANISCHES GENOM ALS REFERENZ**

Bislang beziehen sich alle humangenetischen Forschungen auf die Referenzsequenz, also auf ein Genom mit europäischen Wurzeln. Denn am Humangenomprojekt 2001 waren nun einmal die Forschernationen USA, Frankreich, Deutschland, England und Japan beteiligt. „Biomarker, Rezeptorvarianten, Krankheitsregionen, Ziele für Arzneimittel – all das lässt sich nicht eins zu eins auf das afrikanische Genom übertragen.“ Ein afrikanisches Referenzgenom ist also überfällig, um wirksamere Medikamente für Menschen mit afrikanischen Wurzeln entwickeln zu können.

Um die Genomhälften getrennt sequenzieren zu können, musste Margret Hoehe gemeinsam mit ihrem Team eine neue molekulargenetische Methode sowie die komplette Bioinformatik dazu entwickeln. Ein wesentlicher Un-

terschied zur herkömmlichen Technik besteht darin, dass die DNA-Abschnitte nicht wie üblich 25 bis 40, sondern etwa 40 000 Basen lang sind. Da sie charakteristische Basenaustausch-Muster aufweisen und nicht exakt gleich geschnitten sind – beispielsweise an den Basen 128 und 40200 oder an 14000 und 55030 –, erkennt der Computer beim Überlappen leicht, ob ein Schnipsel zu Teil A oder B des Genoms gehört. Ob A jedoch von Vater oder Mutter stammt, lässt sich erst durch zusätzlichen Vergleich mit mindestens einem Elternteil feststellen.

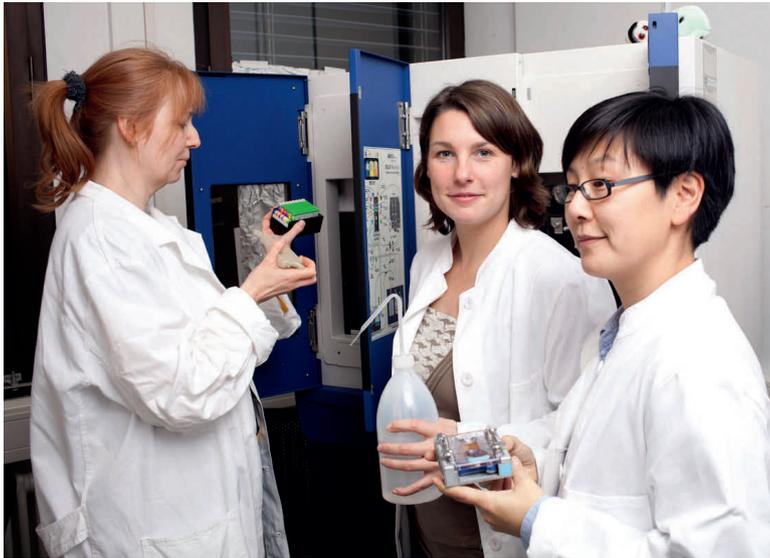
Fast alle der 17861 für Proteine codierenden Gene des Deutschen konnten so in ihrer doppelten Ausführung entschlüsselt werden. 90 Prozent davon kommen in zwei unterschiedlichen molekularen Formen vor. Die beiden Chromosomensätze unterscheiden sich dabei an rund zwei Millionen Stellen. Der Vergleich mit der Referenzsequenz zeigt zudem, dass 60 bis 75 Prozent der molekularen Genformen so nur bei dieser einen Person vorkommen. Wir sind also wesentlich individueller als gedacht. Darüber hinaus identifizierten die Wissenschaftler 159 Gene mit zwei und mehr potenziell krankheitsfördernden Mutationen, 86 davon nur in einer Genkopie.

Wie verhalten sich die beiden molekularen Genformen zueinander? Kooperieren sie, oder arbeiten sie gegeneinander? Oder wechseln sie sich ab?

Welche ist dominant – und warum? Denn nur wenn ein defektes Gen sein intaktes Pendant übertrumpft, entsteht Krankheit. Hoehe ist sicher, es wird eine neue, tiefere Sichtweise der Biologie geben. Sie könnte tatsächlich der fehlende Schlüssel zum Verständnis des Erbguts sein und der Medizin die ersehnte individualisierte Therapie ermöglichen. Damit öffnet sich die Tür zu einem ganz neuen Forschungsfeld.

**UNBEKANNTE REGULATION DER GENE**

Gute Gene mitbekommen zu haben – diese Redewendung bekommt eine neue, eine molekulare Bedeutung. „Manchmal kommt es mir so vor, als ob sich bei mir die Gene von Mutter und Vater gerade bekämpfen“, sagt Margret Hoehe lachend. Nein, im Ernst, da sei was dran. „Man muss es sich nur mal wieder ins Bewusstsein rufen. Vielleicht werden in verschiedenen Lebensphasen die Gene unterschiedlich reguliert? Vielleicht wird sogar eine Version des Chromosoms komplett abgeschaltet?“ Darüber weiß man praktisch nichts. Die Pufferkapazität des Systems scheint hoch. Zwischen gesund und krank gibt es viele kleine Übergangsstufen. „Funktioniert das eine nicht, dann eben das andere. Doch wenn zu viel zusammenkommt, entgleist es irgendwann.“ Letztlich ist wohl vieles eine Frage aller individuellen Anlagen zusammen. >



links: Stefanie Palczewski (links), Sabrina Schulz (Mitte) und Eun-Kyung Suk (rechts) vor einem Hochdurchsatz-Sequenziersystem zur DNA-Analyse. rechts: Jede der beiden Fließzellen des Geräts kann mehr als 600 Millionen DNA-Fragmente aufnehmen.

Um den genetischen Variantenreichtum besser zu erfassen, folgte auf das Humangenomprojekt 2008 das 1000-Genome-Projekt, das sich die Entschlüsselung von 2500 Genomen zum Ziel gesetzt hat. Während dabei viele Forscher weltweit zusammenarbeiten und Teilsequenzen beisteuern, hat Hoehe mit ihrem Team aus gerade mal einer Handvoll Mitarbeitern ein Genom innerhalb weniger Monate komplett sequenziert und analysiert. Bis Jahresende will sie mit zwölf weiteren individuellen haploid analysierten Genomen fertig sein.

Individualität – dieses Thema zieht sich durch Margret Hoehes gesamtes Forscherleben. Sie studierte Medizin und Psychologie in München, promovierte 1986 über „die Wirkung von Opiaten auf neuroendokrine und psychische Parameter“. Bei einer Studie mit depressiven und schizophrenen Patienten merkte sie bald, dass manche auf die Medikamente gar nicht ansprachen. Aber ein Teil der gesunden Probanden ebenfalls nicht. „Die klassische Mittelwertmethode ebnete schließlich alle in-

dividuellen Unterschiede ein.“ Stattdessen müsste man untersuchen, wie das Erbgut jedes Einzelnen die Wirkung von Medikamenten beeinflusst.

### GENOM VERRÄT WIRKSAMKEIT VON MEDIKAMENTEN

Also nahm sie ein Kapitel über die interindividuelle Variabilität der Wirksamkeit von Medikamenten in ihre Dissertation auf. „Es gab einen Aufstand“, erinnert sie sich. Es macht sie noch immer wütend, dass sie das Kapitel streichen musste. Aber der Gedanke an sich ließ sich nicht wegwischen. Und so wird heute den Unterschieden zwischen Respondern und Non-Respondern in medizinischen Studien besondere Beachtung geschenkt.

Der nächste Schritt waren Familienstudien. Hoehe wollte die individuellen Marker für Krankheitsanfälligkeiten und die Wirksamkeit von Medikamenten, die sie entdeckt hatte, auf ihre möglichen genetischen Wurzeln hin untersuchen. Dazu ging sie 1987 in die USA an das National Institute of Mental Health.

„Am Horizont erschien schon das Humangenomprojekt.“ Als ihr Chef fragte, ob sie weiter Neuroendokrinologie betreiben oder mit DNA arbeiten wolle, entschied sie sich kurzerhand für die Erbsubstanz. „Dabei hatte ich keine Ahnung davon und musste mir das Handwerkszeug erst beibringen“, sagt sie und schmunzelt. Bald wurden die Gene der Opiat- und adrenergen Rezeptoren entdeckt, die sie zuvor pharmakologisch untersucht hatte. „Mit heute steinzeitlichen Methoden haben wir die ersten Variationen gefunden und dann das Genom nach Regionen für Depression und Schizophrenie abgesehen.“

Damals glaubte man noch, dass im Erbgut nur wenige Gene mutiert seien, dass diese Mutationen aber schwere Auswirkungen hätten – ein Irrtum, wie man heute weiß. Tatsächlich sind nur sehr wenige Erkrankungen auf ein einzelnes Gen zurückzuführen. Und was die Schizophrenie betrifft: 2009 veröffentlichte NATURE eine Studie, nach der Tausende Veränderungen auf sehr vielen Genen zum Ausbruch der Krankheit

führen können, wobei das Risiko mit einer einzigen dieser Mutationen meist nur sehr wenig steigt. „Ein Pearl Harbour der Schizophrenieforschung“ nannte es die *NEW YORK TIMES*.

## NEUER ANSATZ FÜR GENOMANALYSE

Doch zurück zum Anfang der 1990er-Jahre. Man müsste das Ganze umdrehen, dachte Hoehe bald: nicht von der komplexen Krankheit ausgehen und die verantwortlichen Gene suchen, sondern das Genom nach interindividuellen DNA-Sequenzunterschieden absuchen und *in vitro* nachsehen, welchen Effekt sie haben. Sie schrieb einen entsprechenden Projektantrag und schickte ihn George Church, einem der führenden Genforscher. Ihm gefiel die Idee auf den ersten Blick, und kurze Zeit später war sie als Postdoktorandin an der Harvard Medical School akzeptiert. Dort entwickelte sie mit seinem Team die sogenannte Multiplex-PCR-Sequenzierungstechnologie. 20, später 50 DNA-Schnipsel ließen sich damit simultan auslesen. Dass Kary B. Mullis, Erfinder der Polymerase-Kettenreaktion und Nobelpreisträger, damals sagte, das könne nicht funktionieren, „niemals!“ – das erfuhr sie erst später.

1995 kam sie zurück, baute am Max-Delbrück-Centrum in Berlin eine Arbeitsgruppe auf – noch so ein Abenteuer. Die Maschinen dafür ließ sie vom Machine Shop der Harvard University einfliegen. Sie re-sequenzierte das  $\mu$ -Opiatrezeptor-Gen bei 250 Patienten und Gesunden und fand erste Haplotypen, die mit Suchterkrankungen assoziiert sind. Es war eine der zwei ersten Arbeiten überhaupt, die zeigten, dass es wichtig ist, beide Haplotypen eines Gens zu betrachten und nicht nur eine zufällig herausgegriffene Einzelmutation. 2002 wechselte sie zur Max-Planck-Gesellschaft. Ein Hauptprojekt war die Sequenzierung der Haplotypen der komplexesten Region des menschlichen Genoms, des Histokompatibilitätskomplexes: Bauplan für An-

tikörper des Immunsystems, aber auch reich an Krankheitsgenen. Daraus entstand schließlich die Sequenzierung ganzer haploider Genome.

All dies hat Kraft gekostet. „Es war nicht leicht, über so lange Zeit nicht im Mainstream zu schwimmen.“ Aber das ändert sich gerade. Und so treibt Margret Hoehe bereits erneut die Individualität um. Diesmal geht es um ein personalisiertes Genomprojekt. Krankheit lässt sich am besten dort untersuchen, wo sie stattfindet – am Patienten. Im Fokus wird zuerst eine Brustkrebspatientin stehen, deren haploide Genome Hoehe im Blut und im Tumorgewebe untersuchen wird. Beide Frauen kennen sich bereits geraume Zeit, führen gemeinsam Tagebuch über das Projekt. Vielleicht wird einmal ein Buch daraus.

Ungewöhnlich ist daran, dass die Patientin, eine Journalistin mit medizinischer Vorbildung, an der Studie beteiligt ist und Zugang zu ihren genomischen Informationen erhält. Denn noch ist anonymisierte Forschung in der Medizin die Regel. „Menschen stellen ihre Proben der Wissenschaft freiwillig zur Verfügung. Dafür sollten wir den Menschen auch etwas zurückgeben“, meint Margret Hoehe. „In den USA wird das Thema bereits heiß diskutiert.“ Patienten, die es wissen wollen und ein entsprechendes Vorwissen besitzen, sollten die Informationen zu ihrem Genom auch erhalten, nach dem Motto: Mein Genom gehört mir!

Craig Venter soll, so hört man, prophylaktisch Medikamente nehmen, seit er sein Genom kennt. Der Deutsche, dessen Erbgut sequenziert wurde, kann das nicht. Da seine Blutprobe aus einer Biobank stammt, war sie aus ethischen Gründen anonymisiert. Über ihn ist nur bekannt, dass er zum Zeitpunkt der Blutspende 51 Jahre alt und gesund war. Die Wissenschaftler hätten eine schlechte und eine gute Nachricht für ihn: Er besitzt auch die mutierten Formen des Brustkrebs-Gens *BRCA1*. Aber zum Glück befinden sie sich auf derselben Genkopie. Er hat also auch eine völlig gesunde Version. ◀



Mit ihrer Methode, ein Genom getrennt nach den beiden Haplotypen zu sequenzieren, hat Margret Hoehe die Grundlagen für eine personalisierte Medizin gelegt.

## GLOSSAR

### Chromosom

Das Erbgut von Lebewesen mit Zellkern ist auf unterschiedlich viele Chromosomen verteilt. Ein Chromosom besteht aus dem fadenförmigen DNA-Molekül und verschiedenen Proteinen. Die Chromosomen einer Zelle haben unterschiedliche Größe und tragen unterschiedlich viele Gene. Beim Menschen ist das größte Chromosom rund 250 Millionen Basenpaare lang, das kleinste enthält nur rund 50 Millionen Basenpaare. Das genreichste Chromosom trägt etwa 3000 Gene, das männliche Y-Chromosom dagegen nur rund 100 Gene.

### Diploidie

Diploide Organismen haben einen doppelten Chromosomensatz. Der Mensch beispielsweise besitzt 46 Chromosomen, die in Form von 23 Chromosomenpaaren vorkommen. Jedes homologe Chromosomenpaar enthält je ein Chromosom von der Mutter und vom Vater. Das 23. Chromosomenpaar besteht entweder aus zwei XX-Geschlechtschromosomen (Frauen) oder je einem XX- und einem XY-Geschlechtschromosom (Männer).

### Haplotyp

Die beiden elterlichen Sequenzversionen eines jeden Chromosoms. Da der Mensch jedes Chromosom in doppelter Ausführung besitzt, liegt jedes Chromosom in zwei Haplotypen vor.