

Die Lösung des Sherman-Paradoxons oder die erfolgreiche Suche nach dem Huntington-Gen

Von Ernst-L. Winnacker*

Es war ein Chemiker, Linus Pauling, der das heute so aktuelle Konzept von der „molekularen Krankheit“ formulierte. Ende der vierziger Jahre, lange bevor überhaupt die erste Sequenz eines Proteins bestimmt war, zeigte er, daß sich Hämoglobin von gesunden Personen von dem von Patienten mit Sichelzellanämie in seiner Beweglichkeit im elektrischen Feld unterscheidet^[1]. Erst sieben Jahre später gelang es Vernon Ingram, diesen Unterschied im elektrophoretischen Verhalten auf einen einzigen Aminosäureaustausch in der β -Kette des Hämoglobins von Glutaminsäure nach Valin zurückzuführen^[2].

Seither sind dank der großen Fortschritte in der Molekularbiologie die Ursachen vieler weiterer molekularer Krankheiten aufgeklärt worden. Bei den zugrundeliegenden Mutationen kann es sich um einzelne Basenaustausche, also um sogenannte Punktmutationen, aber auch um Deletionen und Insertionen von DNA-Abschnitten handeln. In *einem* Gen kommen oft mehrere Mutationstypen vor, die gelegentlich auch zu unterschiedlich schweren Krankheitsbildern Anlaß geben. Gemeinsam ist aber (fast) allen bekannten Erbkrankheiten, daß die ihnen zugrunde liegenden Mutationen, einmal entstanden, stabil weiter vererbt werden, so daß sich in einer Familie die Krankheitsverläufe bei Eltern und ihren Nachkommen nicht wesentlich unterscheiden. Ausnahmen sollten auch hier die Regel bestätigen. Oder vielleicht auch einmal nicht?

Nun, schon lange sind Humangenetikern einige wenige Erbkrankheiten bekannt, bei denen das oben gesagte nicht zutifft, bei denen sich also das jeweilige Krankheitsbild von Generation zu Generation verändert, meist verschlimmert, als böte eine einmal ausgelöste Mutation die Grundlage zu immer neuen Veränderungen im Erbmaterial. Da man sich hierauf jedoch keinen Reim machen konnte, sprach man in diesen Fällen, in denen also das Krankheitsrisiko für einen Patienten von der Stellung in seinem Familien-Stammbaum abhängig ist, vom Sherman-Paradoxon^[3].

Was geschieht hier? Liegt diesem Phänomen tatsächlich eine (lokale) Instabilität des Erbmaterials zugrunde? In der Tat ist dies der Fall. Am Prototyp dieser Krankheiten, der Huntingtonschen Chorea (oft auch als HD für *Huntington's Disease* abgekürzt), sei die Natur dieses neuartigen Mutationstyps beschrieben.

Die Huntington-Chorea, auch Veitstanz genannt, wurde erstmals 1872 vom amerikanischen Land- und Nervenarzt George Huntington beschrieben und auch bereits als erblich erkannt^[4]. Die seltene Krankheit – sie tritt mit einer Häufigkeit von 7–10 auf 100 000 Personen auf – bricht meist zwischen dem 35. und 50. Lebensjahr aus. Sie manifestiert sich in motorischen und neurologischen Störungen, wobei sie den ruckartigen ungewollten Bewegungen der Patienten den Namen „Chorea“ verdankt. Mit der Zeit verschlimmern sich diese Symptome immer mehr; begleitet von einem dramati-

schen Abfall der intellektuellen Fähigkeiten führt die Krankheit nach wenigen Jahren unweigerlich zum Tode.

Es ist schon lange bekannt, daß die Symptome mit einem Verlust an Neuronen im Gehirn einhergehen; die genauen biochemischen Grundlagen dieses Phänomens blieben allerdings bislang ein Rätsel. Der klassische Weg zur Identifizierung von Genen, das Aufspüren und Charakterisieren eines modifizierten Proteins, dem dann, wie im Fall der Sichelzellanämie, auch ein defektes Gen zugeordnet werden kann, dieser Weg konnte also hier nicht gegangen werden. Es blieb nur der umgekehrte, dementsprechend als „reverse Genetik“ bezeichnete Ansatz: Die direkte und unmittelbare Suche nach dem defekten Gen, von dessen DNA-Sequenz anschließend das zugehörige Protein abgeleitet und in seinen biochemischen Eigenschaften untersucht werden kann. Ein solcher Weg ist nicht einfach, enthält das menschliche Genom doch mindestens 100 000 Gene. Nur eines davon konnte das fragliche Huntington-Gen sein, aber welches?

Zweier Voraussetzungen bedarf die Lösung einer solchen Aufgabe. Einmal benötigt man Großfamilien mit einer möglichst großen Anzahl von Personen aus mehreren Generationen, um den Erbgang der Krankheit verfolgen zu können. Ferner bedarf es der Existenz charakteristischer Wegweiser oder Markierungspunkte auf dem Erbmaterial. Je enger sie auf dem Chromosom angeordnet sind, desto eindeutiger läßt sich später ein „verdächtiges“ DNA-Fragment einem bestimmten Abschnitt auf dem Erbmaterial zuordnen. Derartige Wegweiser gibt es in großer Zahl entlang der Chromosomen in Form von DNA-Sequenzvarianten – man spricht auch von DNA-Polymorphismen –, die meist außerhalb der codierenden Bereiche liegen. Wenn sie zufällig die Erkennungsstellen für Restriktionsenzyme verändern, dann führt dies zu unterschiedlich langen DNA-Fragmenten, die für ein bestimmtes Genom und damit für ein Individuum charakteristisch sind. Man spricht von „Restriktionsfragment-Längenpolymorphismen“ (RFLPs oder RFLIPs). Solche Veränderungen vererben sich, wie richtige Mutationen, von den Eltern auf ihre Nachkommen und ermöglichen den Nachweis von Verwandtschaftsverhältnissen. Um auf unsere Frage zurückzukommen: Für den Fall, daß sich ein solcher Wegweiser in der Nachbarschaft des defekten Huntington-Gens befindet, wird er auch mit diesem gemeinsam vererbt. Man muß also in „Huntington-Familien“ nach einem DNA-Fragment suchen, das immer gemeinsam mit der Krankheit vererbt wird, aber eben nicht in Individuen oder Familien auftritt, in denen die Krankheit nicht vorkommt. Zu diesem Zweck zerlegt man das Genom von Personen aus Huntington-Stammbäumen^[5] mit einem Restriktionsenzym, das polymorphe DNA-Fragmente erkennt, in die entsprechenden Stücke und hybridisiert die nach ihrer Größe aufgetrennten Fragmente mit radioaktiv markierten menschlichen DNA-Fragmenten. Zunächst ist die Auswahl dieser markierten Fragmente, die man aus Genbanken gewinnt, völlig wahllos; es ist ja nicht bekannt, welcher DNA-Abschnitt in der Nähe des Huntington-Gens liegt. Im Prinzip kann es daher passieren, daß Tausende und Abertausende von DNA-Fragmenten darauf getestet werden müssen, ob sie ein Fragmentmuster

[*] Prof. Dr. E.-L. Winnacker
Institut für Biochemie der Universität München
D-85152 Martinsried
Telefax: Int. + 89/8578-3795

entstehen lassen, das für Huntington-Patienten charakteristisch ist und sich von dem gesunder Personen aus dem gleichen Stammbaum unterscheidet. James Gusella und seine Mitarbeiter^[6] am Massachusetts General Hospital (MGH) in Boston hatten daher 1983 maßloses Glück, als sie schon mit dem achten Fragment, das sie wahllos, wie bei einer Lotterie, aus einer Sammlung von Millionen von klonierten DNA-Fragmenten herausgriffen, fündig wurden. Das Fragment stammte vom Chromosom 4, so daß damit bereits feststand, daß sich das Huntington-Gen auf diesem Chromosom befinden mußte.

Ausgehend von diesem Fragment wurde im Laufe der vergangenen zehn Jahre dann der Genort immer stärker eingengt. Gegen Ende 1992 war man sicher, daß sich das Gen auf einem 500 000 Basenpaare großen Fragment am Ende des kurzen Arms von Chromosom 4 befinden mußte. Selbst ein solches Stück ist aber immer noch gut für zehn bis zwanzig Gene normaler Länge. Um deren codierende Bereiche im Meer nicht-codierender Abschnitte zu finden – wobei immer die zusätzliche Schwierigkeit im Auge zu behalten ist, daß diese Bereiche im Zweifelsfall nicht an einem Stück, sondern aufgeteilt in zahlreiche Stückchen, den sogenannten Exons, vorliegen – stehen inzwischen zahlreiche Verfahren zur Verfügung. Darunter kommt folgenden besondere Bedeutung zu: a) der Identifizierung evolutionär verwandter Sequenzen, die eventuell bereits vorher im Genom eines anderen Organismus identifiziert wurden und durch eine Suche in Sequenzdatenbanken gefunden werden können; b) dem Nachweis hypomethylierter CpG-Inseln, von denen bekannt ist, daß sie sich oft im Promotorbereich von Genen befinden; c) der Identifizierung langer, offener Leserahmen durch DNA-Sequenzierung. Im vorliegenden Fall wurde ein neues Verfahren, die sogenannte Exon-Amplifikation, eingesetzt^[7]. Man nimmt dabei einen Expressionsvektor für Säugerzellen, der ein funktionelles, leicht erkennbares Gen enthält. In dessen Intron wird ein Stück DNA einkloniert, von dem man gerne wissen möchte, ob es ein Exon enthält. Wenn nicht, dann wird nach Transfektion des Vektors in eine Säugerzelle einfach die vom Vektor-Gen abgeleitete mRNA gebildet; wenn ja, wenn also ein funktionelles Exon vorhanden ist, dann wird die neu gebildete mRNA um die Länge dieses Exons größer sein. Über PCR-Verfahren (PCR = Polymerase-Kettenreaktion) kann dieses amplifiziert und charakterisiert werden.

Angewandt auf das 500 kBp lange Segment aus dem kurzen Arm des Chromosoms 4 ergab diese Suche unter anderem einen Leserahmen, genannt IT15 (IT steht für „interesting transcript“), der für ein bislang unbekanntes Protein mit einer Molekülmasse von insgesamt 3448 kD codierte^[8]. Der Leserahmen fiel den Autoren auf, weil er in der Nähe seines 5'-Endes zahlreiche Kopien des Triplets CAG aufwies, das für die Aminosäure Glutamin (Q) codiert (Abb. 1). Der Vergleich dieses Abschnitts mit entsprechenden Abschnitten aus anderen, aber ebenfalls gesunden Personen ergab sehr schnell, daß er in seiner Länge bezüglich dieses einen Codons variabel ist, wobei dessen Anzahl zwischen 11 und 34 Repetitionen schwankt. Offensichtlich ist das Gen in diesem Bereich instabil.

Ist vielleicht dieser Abschnitt die Basis für die Huntington-Mutationen? In der Tat fanden sich bei Huntington-Patienten noch längere Repetitionen, wobei die bislang längste mit 86 Repetitionen identifiziert wurde. Und noch mehr: Die

```
-----AGACCCGCCATGGGGACCTGGAAAAGCTGATGAAGGCCTTCGAGTCCCCTCAAG
M A T L E K L M K A F E S L K
TCCTTCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG
S F Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q Q P P P P P P P P P P P P Q L P G
CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG
Q Q Q Q Q P P P P P P P P P P P P P Q L P G
CCG----P
```

Abb. 1. Sequenzausschnitt aus dem Amino-Terminus des Huntington-Proteins (Huntingtins) und der zugehörigen DNA-Sequenz (große bzw. kleine Buchstaben). Deutlich sind die 23 Glutaminreste aus dem Protein eines gesunden Individuums zu erkennen. In gesunden Individuen schwankt diese Zahl zwischen 11 und 34; in Patienten ist sie stark erhöht.

Anzahl der Repetitionen korreliert mit dem Zeitpunkt des Ausbruchs der Krankheit. Je größer ihre Zahl, desto jünger die Patienten! Damit war nicht nur das Sherman-Paradoxon gelöst, es bestand und besteht nun auch kaum mehr Zweifel daran, daß das gefundene Gen in der Tat das Huntington-Gen ist. Das zugehörige Protein hat 3144 Aminosäuren und wurde Huntingtin genannt. Es weist keinerlei Homologien mit anderen Proteinen auf, so daß weiterhin keine Anhaltspunkte für seine biochemische Funktion existieren. Dennoch ist dieser Befund in vieler Hinsicht bedeutungsvoll. Einmal zeigt er, daß die Technik der reversen Genetik funktioniert, daß es also möglich ist, einzelne Gene ohne Kenntnis ihrer Funktion in großen Genomen zu finden. Darüber hinaus weist er auf einen neuen, bislang unbekannten Mutationstyp, eine Art dynamischer Mutation, hin, die durch die Zunahme von Codon- bzw. Trinucleotidrepetitionen charakterisiert ist. Inzwischen wurde gefunden, daß vier weitere Erbkrankheiten auf diesem Prinzip der instabilen Codonrepetitionen beruhen (siehe Abb. 2). Die zugehörigen Krankheitsbilder, so verschieden sie alle sind, weisen gewisse Gemeinsamkeiten auf: sie gehen generell mit dem Verfall von motorischen und intellektuellen Fähigkeiten einher, und die Symptome verschlimmern sich von Generation zu Generation; die Zunahme der Repetitionen korreliert mit dem Alter, in dem die Krankheit ausbricht.

Wie immer in solchen Fällen, wirft der Fortschritt mehr Fragen auf, als er beantwortet. Gibt es bestimmte Codons, die für diese Art der generationsabhängigen Instabilitäten bevorzugt sind? In diesem Zusammenhang fällt es auf, daß

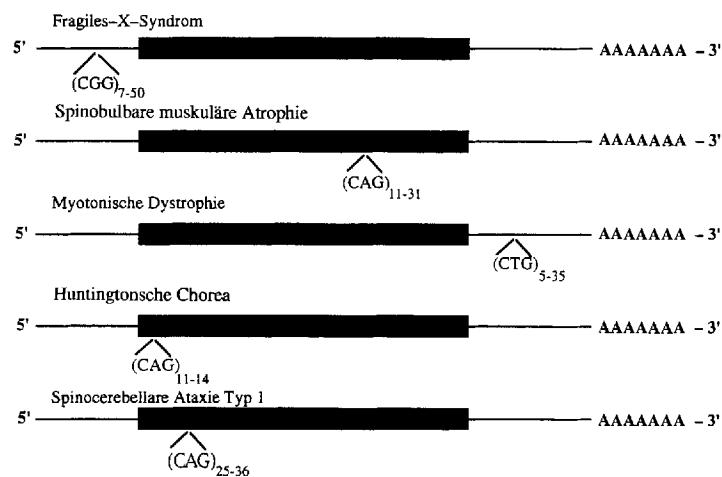


Abb. 2. Anordnung der Trinukleotid-Repetitionen in den verschiedenen Genen, bei denen bislang derartige Repetitionen als variabel und für das Auftreten entsprechender Krankheitsbilder verantwortlich erkannt wurden (nach *Science* 1993, 260, 1422–1423). Die erwähnten Syndrome sind ausnahmslos seltene Erbkrankheiten und gehen meist mit neurodegenerativen Symptomen einher.

die drei bislang beobachteten Codons CAG, CTG und CGG (Abb. 2) an der ersten und dritten Position jeweils die komplementären Basen C und G tragen. Längere Abschnitte aus diesen Codons können daher auf sich selbst zurückfalten und stabile Sekundärstrukturen, sogenannte Stem-Loop-Strukturen, bilden. Die hierfür postulierten A-A, T-T- und G-G-Homo-Purin bzw. -Homo-Pyrimidin-Basenpaare sind dabei keineswegs verboten, sondern tragen vielmehr signifikant zur freien Energie der Bildung solcher Sekundärstrukturrelemente bei^[9].

Warum sind es ausgerechnet Trinucleotide, die diese Instabilitäten aufweisen und nicht Di- oder Tetranucleotide, denn auch solche Repetitionen sind im menschlichen Genom bekannt? Welche Auswirkungen haben die Codonrepetitionen auf die Genexpression? Einige von ihnen, wie etwa im Fall des Fragilen-X-Syndroms befinden sich außerhalb der codierenden Region, betreffen also „nur“ die mRNA und nicht das Protein^[10]. Selbstverständlich kann man sich vorstellen, daß und wie eine solche Repetition die Sekundärstrukturbildung, die Stabilität, die Modifikation, die Lokalisation und was sonst noch alles die Translatierbarkeit einer mRNA steuert, beeinflussen könnte. Vielleicht bilden sich auf diese Weise auch Bindungsstellen für mRNA-bindende Proteine, die – ähnlich wie im Fall des Ferritin/Transferrin-Rezeptors^[11] – die Effizienz des Translationsprozesses kontrollieren. Auch für den Fall, daß sich die Repetitionen in codierenden Bereichen befinden, lassen sich die Auswirkungen der Anwesenheit solcher Repetitionen gut vorstellen. Ein Polyglutamin-Abschnitt am Amino-Terminus eines Proteins muß dessen Faltung, Löslichkeit, Lebensdauer und auch seine Wechselwirkungen mit anderen Proteinen beeinflussen, wobei all diese Erklärungen zu berücksichtigen haben, daß die Huntington-Mutation dominant ist, daß also das „defekte“ Protein nicht nur für sich allein krankheitserregend ist, sondern auch schon in Gegenwart des „normalen“ Proteins. Dieses muß also in seiner Wirkung durch das mutierte Huntingtin irgendwie negativ beeinflußt werden.

Schließlich werfen diese Beobachtungen die Frage auf, was denn die molekularen Grundlagen dieser Instabilitäten sind? Offensichtlich entstehen sie in den Keimzellen. Diese wiederum bilden sich durch eine charakteristische Folge von Zell- und sogenannten Reduktionsteilungen (Meiose) aus entsprechenden Vorläuferzellen. Dabei kommt es auch zu DNA-Synthese und Reparaturprozessen, die mit dem Austausch von Erbmaterial, der sogenannten homologen Rekombination, in diesen Differenzierungsphasen einherge-

hen. Es ist denkbar, daß die hier beteiligten Enzyme andere Spezifitäten aufweisen als die, die an den üblichen DNA-Syntheseprozessen beteiligt sind. Man wird diese und viele andere Fragen jetzt untersuchen können. Biochemie und Chemie sind hier in ganz unerwarteter Weise gefordert.

Am Ende sei erwähnt, daß und wie die hier diskutierten Ergebnisse auch vielfältige ethische Fragen aufwerfen. Wenn das Huntington-Gen nun identifiziert ist, wird man die Krankheit auch diagnostizieren können und über die Länge der aufgefundenen Repetitionen sogar Prognosen über ihren Verlauf machen können. Wie kann und darf man dies gesunden Patienten vermitteln, die auf diese Weise vielleicht ein erst nach Jahrzehnten wirksames, aber sicheres Todesurteil erfahren? Für die Humangenetik sind derartige Fragen nicht neu; die modernen Entwicklungen der Molekularbiologie können sie in der einen oder anderen Weise jedoch verschärfen, insbesondere, wenn – wie so oft – die Therapieentwicklung hinterherhinkt. Es ist hier nicht der Raum, dies weiter auszuführen. Dennoch darf ein solcher Triumph der modernen Genetik, der die Identifizierung des Huntington-Gens sicherlich ist, uns allen den Blick auf diese Fragen, die derartige Projekte eben auch aufwerfen, nicht verwehren^[12].

- [1] L. Pauling, *Science* **1949**, *110*, 543–548.
- [2] V. M. Ingram, *Nature* **1957**, *180*, 326–328.
- [3] Y. H. Fu, D. P. Kuhl, A. Pizutti M. Pieretti, J. S. Sutcliffe, S. Richards, A. J. M. H. Verkerk, J. A. Holden, R. G. Fenwick, S. Warren, B. A. Oostra, D. L. Nilson, C. T. Caskey, *Cell* **1991**, *67*, 1047–1058.
- [4] J. F. Gusella, *Adv. Hum. Genet.* **1991**, *125*–151; J. B. Martin, J. F. Gusella, *New Engl. J. Med.* **1986**, *315*, 1267–1276.
- [5] Solche Stammbäume sind in den USA, aber auch und gerade aus Venezuela bekannt, wo um den Maracaibo-See herum ca. 9000 Personen leben, die in ihrer Abstammung auf eine einzige Frau zurückgeführt werden können, die im frühen 19. Jahrhundert an der Huntington-Chorea erkrankte und starb.
- [6] J. F. Gusella, N. S. Wexler, P. M. Conneally, S. L. Naylor, M. A. Anderson, R. E. Tanzi, P. C. Watkins, K. Ottina, M. R. Wallace, A. Y. Sakaguchi, A. B. Young, I. Shoulson, E. Bonilla, J. B. Martin, *Nature* **1983**, *306*, 234–238.
- [7] A. J. Buckler, D. D. Chang, S. L. Graw, J. D. Brook, D. A. Haber, P. A. Sharp, D. E. Hausman, *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 4005–4009.
- [8] The Huntington's Disease Collaborative Research Group, *Cell* **1993**, *72*, 971–983.
- [9] I. Tinoco, P. N. Borer, B. Dengler, M. D. Levine, O. C. Uhlenbeck, D. M. Crothers, J. Gralla, *Nature London New Biol.* **1973**, *346*, 40–41.
- [10] C. T. Ashley, J. S. Sutcliffe, C. B. Kunst, H. A. Leiner, E. E. Eichler, D. L. Nelson, S. T. Warren, *Nat. Genet.* **1993**, *4*, 244–251.
- [11] E. C. Theil, *J. Biol. Chem.* **1990**, *265*, 4771–4774.
- [12] Siehe z.B.: H. Markl, *Wissenschaft im Widerstreit*, VCH, Weinheim, **1990**, S. 145–177 oder E.-L. Winnacker, *Am Faden des Lebens*, Piper, München, **1993**, S. 227–256.