

2 Einleitung

2.1 Untersuchungen von Aneusomien im humanen Genom

Seit 1944 von Avery, MacLeod und McCarthy die DNS als genetisches Material identifiziert wurde, ermöglichten vor allem technische Weiterentwicklungen neue wissenschaftliche Entdeckungen in der Genetik. So führte die Anwendung der Röntgenstrukturanalyse von DNS- Molekülen (Franklin, Watson und Crick, 1953) zur Aufklärung der räumlichen Struktur. Mit dem abgeleiteten Modell bestehend aus den DNS- Doppelsträngen und der Aufklärung der Basenpaarung wurden auch die Grundlagen für die DNS- Hybridisierung (Southern, 1975) gelegt. Diese Methode führte dann zur ersten Lokalisierung eines Gens (α 1- Hämoglobin, HBA1) auf molekularer Ebene (Kan *et al.*, 1976). Seitdem wurden für verschiedene Fragestellungen Modifikationen eines ursprünglich entwickelten Verfahrens mit Erfolg verwendet. So wie bei der genannten Identifizierung des für die α - Thalassämie verantwortlichen Gens konnte immer wieder der Verlust oder auch die Amplifikation von genetischem Material (numerische oder unbalancierte Chromosomenaberrationen) als Ursache für eine vererbte Krankheit nachgewiesen werden. Auch bei der Entstehung von Tumoren wurde diese Art von Mutationen gefunden. Der letztere Zusammenhang ist besonders interessant, da im Verlauf der neoplastischen Transformation genetische Veränderungen akkumuliert werden. Bestimmten Malignitätsgraden sind spezifische Aberrationen zuzuordnen. Sie sind u.a. in Form von Amplifikationen, Deletionen und Rearrangements einzelner Gene bis zu vollständiger Chromosomen (McEvoy *et al.*, 2003) zu finden. Diese Veränderungen erlangen diagnostische und prognostische Bedeutung.

Die klassische Methode, solche Veränderungen zu untersuchen, ist die vergleichende Genom- Hybridisierung (*Comparative Genomic Hybridization*, Kallioniemi, A. *et al.*, 1992). Wie aus der Bezeichnung abzuleiten ist, werden dabei ganze Genome vergleichend hybridisiert. DNS- Hybridisierungstechniken (siehe oben) gehören zu den klassischen molekularbiologischen Methoden. Grundlage hierfür ist die Ausbildung von DNS- Doppelsträngen. Diese können aufgetrennt werden, das heißt, die DNS liegt dann

als Einzelstrang vor. Unter geeigneten, renaturierenden, Bedingungen bilden sich dann wieder Doppelstränge aus, die nicht notwendigerweise aus den beiden Ausgangsmolekülen bestehen müssen, sondern es hybridisieren alle DNS- Moleküle miteinander, deren Sequenz komplementär ist (Hybride). Für eine klassische Hybridisierung werden die zu testenden Nukleinsäuren auf ein Trägermaterial transferiert, immobilisiert und eine markierte DNS- Sonde bekannter Zusammensetzung detektiert dann die passenden Sequenzen des zu untersuchenden Nukleinsäuregemischs. Für das weitere Verständnis ist es wichtig, die Begriffe Sonde und Probe zu definieren: als Sonde wird die bekannte DNS bezeichnet, als Probe das unbekannte, zu testende DNS- Gemisch. Für die CGH benutzt man immobilisierte Metaphase- Chromosomen als Sonde und markierte, gesamtgenomische DNS- Fragmente als Probe. Das Verfahren wird komparativ durchgeführt. Dabei werden zwei Proben gleichzeitig eingesetzt, die um die vorhandenen Sondenmoleküle konkurrieren. Das Verhältnis der an die Sonden gebundenen markierten Probenmoleküle spiegelt so das DNS- Mengenverhältnis an einem bestimmten Locus wieder. Die Proben- Genome werden dabei so gewählt, daß immer ein Referenz- oder Kontrollgenom mit dem zu untersuchenden, unbekanntem Genom verglichen wird. Durch die Auswertung der so entstandenen Bilder erhält man dann ein Profil, das Abweichungen der Kopienzahl von der Kontroll- DNS locus-spezifisch dargestellt. Dem Vorteil, ein ganzes Genom in einem Experiment untersuchen zu können, stehen aber auch Nachteile gegenüber. Zunächst können nur Veränderungen untersucht werden, die zu einem Verlust oder zu einer Zunahme von genomischen Abschnitten führen, d.h. unbalancierte Veränderungen. Sogenannte

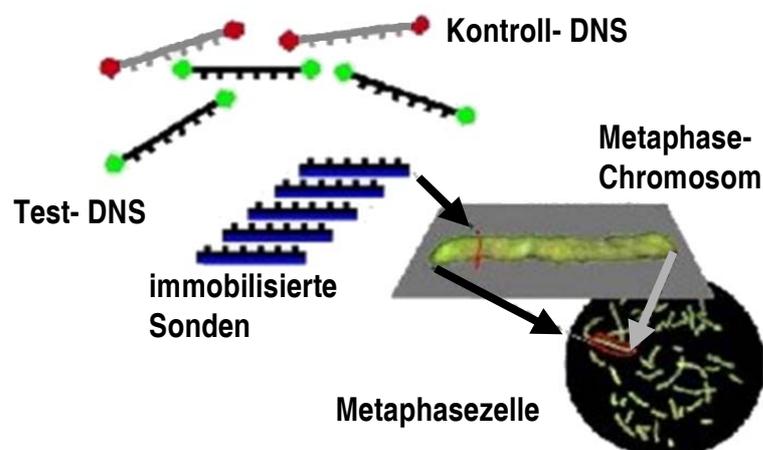


Abbildung 2-1: Prinzip der CGH: verschieden markierte Test- und Kontroll- DNS- Fragmente werden auf immobilisierte Metaphasechromosomen hybridisiert, die Methode ist schematisch im molekularen und mikroskopischen Maßstab dargestellt

balancierte Translokationen können dagegen nicht detektiert werden. Limitationen ergeben sich auch hinsichtlich der erreichbaren Auflösung. Da das primäre Ergebnis ein digitales Bild eines Chromosoms ist, kann die Auflösung nicht größer sein als ein Pixel. Die technischen Grenzen der CGH wurden untersucht und dabei gezeigt, daß die maximale Auflösung von der numerischen Änderung der Kopienzahl abhängig ist. Je höher die Anzahl der zusätzlichen Kopien, desto genauer kann die Größe der amplifizierten Region eingegrenzt werden. Bestenfalls kann das Produkt aus Kopienzahländerung und Größe der amplifizierten Region mit zwei Mbp angegeben werden (Joos *et al.*, 1993). Diesem Grenzwert entsprechen zum Beispiel eine zusätzliche Kopie von zwei Megabasenpaaren Größe oder 40 Kopien von 50 Kilobasenpaaren. Das Erkennen von heterozygoten (ein Allel deletiert) Deletionen ist besonders schwer, da die Änderung nur eine Kopie umfaßt. Tatsächlich wurde das Limit für die Detektion von heterozygoten Deletionen experimentell mit 10-12 Mbp bestimmt (Bentz *et al.*, 1998). Die Anzahl zusätzlicher Kopien läßt sich nur ungenau abschätzen. (Kallioniemi, A. *et al.*, 1992). Die Lokalisierung der gefundenen Aberrationen erfolgt über ihre Zuordnung zum cytogenetischen Bandenmuster. Ist eine nähere Charakterisierung der mutierten Region erwünscht, um zum Beispiel Krankheitsgene zu identifizieren, sind weitere Hybridisierungen bis zur Klonierung unerlässlich. Die

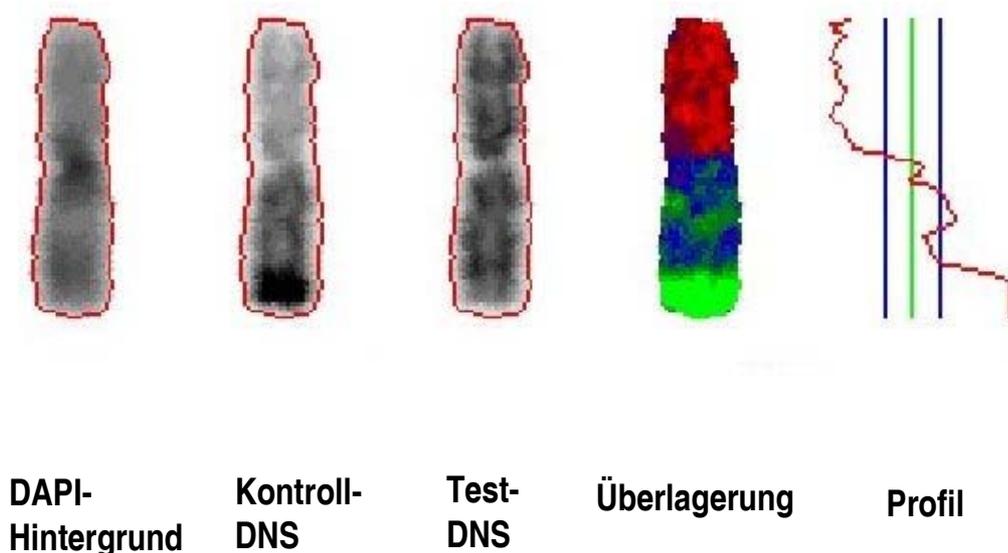


Abbildung 2-2: CGH- Auswertung: die DAPI- Hintergrundfärbung ermöglicht die Zuordnung der Signale zum cytogenetischen Bandenmuster, Kontroll- und Test- DNS werden unterschiedlich markiert und die Überlagerung beider Signale wird rechnerisch zu einem Profil zusammengesetzt, das die Kopienverhältnisse am jeweiligen chromosomalen Locus wiedergibt.

Technologie der klassischen CGH darf als etabliert angesehen werden. Weitere Optimierungen der Methode sind nicht zu erwarten, so daß für höhere Auflösungen und eine exaktere Kartierung der gefundenen Mutationen neue Ansätze nötig werden. Da das Grundprinzip der vergleichenden Hybridisierung auch bei der Expressionsanalyse auf Mikroarrays Anwendung findet, liegt es nahe, diese Technologie auch auf die CGH zu übertragen. Das Prinzip der CGH kombiniert mit der Technologie der Expressionsanalyse auf Mikroarrays wird als ‚Matrix- CGH‘ (engl.: ‚Array CGH‘) bezeichnet.

2.2 Technische Grundlagen der Mikroarrays

Die Grundlagen für die Hochdurchsatzvariante zur Hybridisierung von Nukleinsäuren im Mikroformat wurden Mitte der 90-er Jahre gelegt (Fodor *et al.*, 1993, Schena *et al.*, 1995). Mit dieser Methode eröffnen sich neue Möglichkeiten für die Analyse der Genexpression, für die Detektion von genetischen Variationen (Mutationen und Polymorphismen), für die Entschlüsselung bisher unbekannter Gene und für die Entwicklung neuer diagnostischer Verfahren. Zwei verschiedene technische Ansätze haben sich herausgebildet, die sich vor allem in der Art der verwendeten Sonden und deren Immobilisierung unterscheiden. Da zur Herstellung von Oligonukleotid- Rastern Methoden aus der Halbleitertechnologie verwendet werden, wird diese Form der Microarrays auch DNS- Chips genannt. Die Oligonukleotide werden dabei Base für Base auf der Oberfläche synthetisiert. Die aufwendige Fertigung beschränkt die Herstellung aber zumeist auf kommerzielle Anbieter.

Hingegen ermöglicht das Aufbringen von vorgefertigten DNS- Klonen (PCR- Produkte, Plasmide, PACs, BACs) auf Glasoberflächen eine breitere Benutzerbasis. Microarray- Experimente lassen sich am ehesten mit *dotblot*- Hybridisierungen (Kafatos *et al.*, 1979) vergleichen. Der auffälligste Unterschied liegt dabei in dem verwendeten Trägermaterial: beschichtete Glasobjektträger statt Nylonmembranen. Die Verwendung von Glas ermöglicht den Einsatz von Fluoreszenzfarbstoffen zur Markierung der Proben, da Glas im Gegensatz zu diversen Kunststoffen eine sehr niedrige Eigenfluoreszenz besitzt. Die mit Fluoreszenzfarbstoffen erreichbaren Intensitäten haben einen dynamischen Meßbereich von bis zu 1.000 : 1. Dagegen lassen sich

Autoradiogramme nur sehr grob quantifizieren¹, Werte von 10 : 1 gelten als obere Grenze. Da mehr als nur ein Farbstoff eingesetzt werden kann, können mehrere Hybridisierungen in einem Experiment durchgeführt werden. Erst dadurch werden die komparativen Ansätze für Expressionsanalyse und Matrix- CGH möglich. Das als Trägermaterial verwendete Glas ist chemisch und thermisch sehr beständig, so daß hohe Temperaturen und Ionenstärken oder Lösungsmittelkonzentrationen während Hybridisierung und Waschschritten toleriert werden können. Glasoberflächen sind plan und nicht porös, somit kann das Reaktionsvolumen klein und die Hybridisierungszeit kurz gehalten werden. Aufgrund der genannten Vorteile hat sich der Einsatz von Glas in Form von 75x 25mm Objektträgern durchgesetzt.

Glasobjektträger lassen sich auch leicht durch Roboter bearbeiten. Die Notwendigkeit des Einsatzes von Robotern wird bei möglichen Sondenzahlen von über 10.000 pro Objektträger leicht ersichtlich. Dazu haben sich verschiedene Verfahren für den Transfer der Sonden- DNS auf die Glasoberfläche herausgebildet. Zu nennen sind hier einfache *Split Pin- Arrayer*, *Pin and Ring-* oder *InkJet-* Systeme. Ein Gerät aus der ersten Gruppe fand für die vorliegende Arbeit Verwendung. Dabei übernimmt ein in allen drei Achsen beweglicher, computergesteuerter Arm den Transfer der DNS- Lösung vom Vorratsgefäß (Mikrotiterplatte) auf den Objektträger. Die DNS wird von einer Metallnadel transportiert, die in der Mitte geschlitzt ist. Der Schlitz dient hierbei als Vorratsreservoir für die DNS- Lösung. Die Spitzen sind meist gehärtet, da die Geometrie der Kontaktfläche maßgeblich die Qualität der erzeugten Tröpfchen bestimmt, die nach deren Eintrocknen als ‚Spots‘ bezeichnet werden. Um eine gleichmäßige Form und Größe der Spots zu erreichen, muß die Nadel langsam und gleichmäßig bewegt werden. Der Transport der Flüssigkeit beruht auf der Ausnutzung von Kapillarkräften. Die vor jeder neuen Sonde gereinigte und getrocknete Spitze wird in der Mikrotiterplatte in die DNS- Lösung getaucht. Die Flüssigkeit steigt in dem engen Schlitz empor und verbleibt dort bis die Spitze auf die Glasoberfläche aufsetzt. Die Oberflächenspannung der wäßrigen Lösung bewirkt dann, daß sich eine Flüssigkeitssäule zwischen Glasoberfläche und Spitze ausbildet, die bei weiterer Entfernung der Spitze schließlich reißt. Die hydrophobe Oberfläche des Objektträgers

¹ Mengenunterschiede von 3: 2 gelten als radioaktiv nicht unterscheidbar (Geschwind *et al.*, 1998)

hat zur Folge, daß das gebildete Tröpfchen sehr klein bleibt, bis es dann eintrocknet. Ein typischer Spot (bezogen auf die in der vorliegenden Arbeit verwendeten Geräte) entsteht aus einem Tröpfchen von 300 µm Durchmesser mit einem Volumen von 5nl, das 2,5 ng DNS enthält. Die Bindung der DNS an die Oberfläche erfolgt primär durch elektrostatische Wechselwirkung zwischen den negativ geladenen Phosphatgruppen des DNS- Rückgrats mit den NH₂- Gruppen von Aminosilan oder Poly-L-Lysin. Durch 'UV- *Crosslinking*' oder 'Backen' können zusätzlich kovalente Bindungen zwischen den Thymidingruppen und den genannten NH₂- Gruppen erzeugt werden. Die eigentliche Hybridisierung findet unter einem Deckgläschen in einem geschlossenen Behältnis bei gesättigter Luftfeuchtigkeit (Feuchtekammer) und definierter Temperatur statt. Da das Hybridisierungsvolumen nur einige µl beträgt, reichen wenige Stunden Reaktionszeit aus. Im allgemeinen werden Hybridisierungen aber über Nacht durchgeführt. Die verwendeten Puffersysteme sind grundsätzlich die gleichen, wie sie auch für Hybridisierungen auf Nylonmembranen benutzt werden. Die Puffer enthalten Substanzen wie Denhardt's Reagenz, tRNS oder C₀t1- DNS, um unspezifische Bindungen durch die markierte Probe zu unterdrücken. Nach der Hybridisierung werden die Objektträger gewaschen, um nicht gebundene Proben zu entfernen..

Die Detektion gebundener, fluoreszenzmarkierter Sonden erfolgt mit einem konfokalen Laser- Scanner. Diese Scanner können verschiedene Fluoreszenzfarbstoffe nebeneinander detektieren. Am weitesten verbreitet ist die Kombination Cyanin3 (Cy3) mit Cyanin5 (Cy5) (Mujumdar *et al.*, 1993). Die Objektträger werden nacheinander in beiden Kanälen ausgelesen, d.h. beide Farbstoffe werden mit einem Laser spezifisch angeregt und über einen Photomultiplier ausgewertet. Je höher die gemessene Farbintensität, desto höher ist die Menge an gebundener, markierter DNS- Probe. Über einen A/D- Wandler werden die gemessenen Werte standardmäßig in ein TIF- Bild mit einer Auflösung von 16 bit pro Pixel ausgegeben. Die erhaltenen Bilder sind Grundlage für die Auswertung. Zum einen werden die schwarz- weiß- Bilder beider Kanäle übereinandergelegt und die Unterschiede der Intensitäten pro Pixel in einer Falschfarbendarstellung wiedergegeben, die Abweichungen farblich hervorhebt. Zur genaueren Analyse werden die Daten noch statistisch ausgewertet, um die Signifikanz der erhaltenen Ergebnisse beurteilen zu können.

Tabelle 2-1: Vergleich wichtiger Parameter von Hybridisierungen auf verschiedenen Trägermaterialien

	<i>Nylon- Membranen</i>	<i>Glas- Microarrays</i>	<i>Oligonukleotid-Chips</i>
Probenmenge (z.B. mRNS)	0,5µg	2µg	10µg
Markierung	³² P, ³³ P	Fluoreszenz	Fluoreszenz
Hybridisierungsvolumen	40ml	2-10µl	200µl
Detektion	Phosphorimager	Konfokaler Scanner	Konfokaler Scanner
Detektionslimit [unter wie vielen unspezifischen kann ein spezifisches Molekül erfaßt werden]	1 : 20.000	1 : 100.000	1 : 300.000
Minimale Probenmenge (pro Sonde) [wie viele markierte Moleküle werden zur Detektion eines Spots benötigt]	25x10 ⁶ Moleküle	20x10 ⁶ Moleküle	30x10 ⁶ Moleküle

[nach: Technologies Avancées pour le Génome et la Clinique –TAGC, siehe Internetquellen]

2.3 Matrix- CGH - Stand der Technik

Matrix- CGH beruht auf denselben technischen Grundlagen wie die Expressionsanalyse mit Microarrays (Schena *et al.*, 1995). Während letztere Methode als molekular-biologisches Werkzeug bereits etabliert ist (Shalon *et al.*, 1996, DeRisi *et al.*, 1996, Iyer *et al.*, 1999), war die Matrix- CGH zu Beginn dieser Arbeit für eine breite Anwendung in der allgemeinen Forschung und Diagnostik technisch noch nicht ausgereift genug. Fortschritte werden vor allem durch die Lösung der folgenden beiden Probleme erwartet:

- a.) die Häufigkeit von repetitiven Sequenzen (siehe 1.4.1): repetitive Sequenzen führen in Hybridisierungen zu unspezifischen Signalen und müssen deshalb unterdrückt werden. Während repetitive Elemente in mRNA nur sehr selten zu finden sind (5%, Makalowski *et al.*, 1996), ist ihr Vorhandensein in genomischen Klonen die Regel (50%, siehe 1.4.1). Die Auswahl der Sonden und Maßnahmen zum Blockieren unspezifisch bindender Sequenzen sind für DNS- Hybridisierungen besonders kritisch.
- b.) die Komplexität der Probe: die Komplexität der Probe findet ihre direkte Entsprechung in der Signalintensität der einzelnen Experimente. Zu schwache Signale führen zu schwer interpretierbaren Aussagen. Die Signalintensität ist abhängig von der Menge und der Länge der gebildeten DNS- Hybride. Die komplexe Probe enthält unterschiedlich viele verschiedene Spezies von Nukleinsäuren. Je höher die Anzahl verschiedener Spezies, desto höher ist der effektive Verdünnungsfaktor für die eingesetzte Menge an markierter Probe.

Das humane Genom (= Probe bei der Matrix- CGH) ist sehr viel komplexer als das Transkriptom (= Probe bei der Expressionsanalyse), welches die Gesamtheit aller transkribierten Gene einer Zelle umfaßt. Die Angaben für die Anzahl der Gene im menschlichen Genom schwanken zwischen 30-35.000. Das Humane Genom Projekt rechnet mit 30-40.000 Genen, wobei 31.000 als momentan wahrscheinlichste Anzahl angesehen wird (THE GENOME SEQUENCING CONSORTIUM, 2001). Die Daten der Firma Celera weisen 26.383 validierte Gene aus dazu kommen weitere 12.731 bisher noch nicht bestätigte Genvorhersagen (Venter *et al.*, 2001). Diese Zahlen sind aber als ein Maximalwert anzusehen, da eine große Anzahl von Genen nur vorübergehend exprimiert werden und in der mRNA ausdifferenzierter, spezialisierter

Zelltypen (z.B. Lymphozyten) gar nicht oder kaum vertreten sind. Daraus ergibt sich ein Schätzwert von 15.000 mRNS- Spezies pro Zelle (Alberts *et al.*, 1994: 15-20.000 , Bertoli *et al.*, 1995: 11-14.000). Rechnet man mit einem Wert von 15.000 mRNS- Spezies mit einer durchschnittlichen Länge von rund 1,5 kbp (1.340 bp: THE GENOME SEQUENCING CONSORTIUM, 2001) so erhält man eine Gesamtlänge von rund 22,5 Mbp. Im Vergleich zum humanen Genom mit einer Größe von rund 3.000 Mbp ist das Transkriptom einer einzelnen Zelle somit 133 mal weniger komplex. Theoretisch benötigt man also eine mehr als hundertmal höhere Empfindlichkeit für Systeme zur Auswertung von Matrix- CGH- Experimenten, was technisch nur sehr schwer realisierbar ist. Es müssen also andere Wege gefunden werden, diese Sensitivitätsbarriere zu umgehen. Zwei Ansätze stehen sich dabei zur Zeit gegenüber: zum einen benutzt man längere genomische Klone wie zum Beispiel BACs als Sonden (Solinas-Toldo *et al.*, 1997; Pinkel *et al.*, 1998). Dabei können mehr, verschiedene Probenmoleküle an eine Sonde binden und man erhält so eine höhere Signalintensität auf Kosten der Auflösung, da die minimale Auflösung durch die Größe der verwendeten Klone vorgegeben ist. Eine kurze Kalkulation verdeutlicht diesen Ansatz: Ein PCR-Produkt als Sonde für die Expressionsanalyse habe eine Länge von 500 bp, dieser Wert wird mit dem Faktor 133 multipliziert und man erhält eine Länge von 66,5 kbp. Da aber rund 50% der genomischen Sequenz repetitiv sind, und somit effektiv nicht zur Verfügung stehen, ist der erhaltene Wert von 66,5kbp noch zu verdoppeln. Man erhält somit eine Mindestlänge von 133kbp, das entspricht in etwa der Größe der verwendeten BAC- Klone.

Eine andere Methode arbeitet unverändert mit PCR- Produkten als Sonden und reduziert dafür die Komplexität der Sonde mindestens um den Faktor 133 (Geschwind *et al.*, 1998, Lucito *et al.*, 2000). Die Verringerung der Komplexität der Sonde geschieht durch selektive PCR- Amplifikation einer repräsentativen Auswahl von genomischen Sequenzen. Diese Auswahl bestimmt dann die Komplexität der erzeugten Probe. Es ist zum Beispiel möglich, genomische DNS mit Restriktionsenzymen zu zerschneiden und man erhält auf diese Art eine Verteilung unterschiedlich langer Fragmente. An diese Fragmente können PCR- Adaptoren ligiert werden und mit diesen Adaptoren anschließend PCR- Produkte mit einem bestimmten Längenspektrum erhalten werden. Die Verringerung der Komplexität wird dadurch bestimmt, daß nur Restriktions-

fragmente geeigneter Länge (meist kleiner als 2,5 kbp) amplifiziert werden und der Großteil der längeren Fragmente nicht vermehrt wird. Als Sonden werden die auf diese Weise erzeugten PCR-Produkte verwendet (Lucito *et al.*, 2000). Alternativ kann man auch die im Genom enthaltenen repetitiven Elemente für eine entsprechende PCR benutzen (Geschwind *et al.*, 1998). Entsprechende Methoden, wie zum Beispiel die Inter-Alu-PCR (Nelson *et al.*, 1989), werden schon seit längerem für andere Anwendungen eingesetzt (Kass und Batzer, 1995) und im weiteren detailliert dargestellt (siehe 2.4.3).

Die ersten Ergebnisse zur Matrix-CGH wurden 1997 (Solinas-Toldo *et al.*, 1997) veröffentlicht. Als immobilisierte Sonden wurden Cosmide und PACs verwendet. DNS von Tumorzelllinien wurde mit Referenz-DNS verglichen. Es konnte gezeigt werden, daß stark amplifizierte genomische Bereiche mit Hilfe von Cosmiden (40kb) als Sonden nachweisbar sind. Für das Auffinden von genomischen Abschnitten mit einer oder zwei zusätzliche Kopien waren dagegen Sonden in der Größenordnung von PACs (75-130 kbp) notwendig. Die Anzahl der zusätzlichen Kopien wurde nicht ermittelt. Um das Problem der schwachen Signalintensität zu lösen, wurden somit PCR-Produkte durch größere genomische Klone ersetzt. In diesen Experimenten wurde nachgewiesen, daß ein direkter Zusammenhang zwischen erreichbarer Auflösung und Sensitivität besteht. Eine ausreichende Sensitivität ist gegeben, wenn sich das Signal im dynamischen Bereich befindet, in dem Änderungen der Kopienzahl auch quantitativ erfaßt werden können. Diese Ergebnisse wurden von Pinkel *et al.*, 1998 bestätigt, der anhand von Zelllinien, die ein bis fünf X-Chromosomen pro Kerngenom enthielten, einen linearen Zusammenhang zwischen gemessenen Signalintensitäten und der Anzahl der vorhandenen DNS-Kopien nachweisen konnte. Ein experimentell anderer Ansatz (Geschwind *et al.*, 1998) zur Überwindung der limitierenden Sensitivitätsbarriere (siehe oben) benutzt nicht die genomische DNS als Probe, sondern nur eine repräsentative Auswahl von kurzen genomischen Fragmenten der Ausgangs-DNS. Diese Auswahl mit verringerter Komplexität ermöglicht Hybridisierungen mit Fragmentlängen im Bereich von PCR-Produkten. Da der mittlere Abstand der dabei durch Inter-Alu-PCR (siehe 2.4.4) erzeugten genomischen Klone im Bereich von etwa einem pro 50-100 kbp liegt, ist die Auflösung in etwa die gleiche wie in den zuvor genannten Veröffentlichungen.

Allerdings hat dieses Verfahren offensichtlich eine größere Streuung der Meßwerte zur Folge, wie man aus den dargestellten Ergebnissen ableiten kann.

2.4 Das humane Genom - repetitive Elemente

2.4.1 Übersicht

Der Begriff "humanes Genom" steht für die Gesamtheit der genetischen Information einer menschlichen Zelle. Dieses umfaßt das komplexe Kerngenom und das mitochondriale Genom. Letzteres enthält 37 Gene in 16,6 kb Gesamtlänge. Das Kerngenom ist dagegen 3,2 Gbp groß und beinhaltet wie erwähnt schätzungsweise 31.000 Gene. Sequenzen, die auch translatiert werden, stellen aber nur 3% des humanen Genoms dar. Der Rest verteilt sich auf Abschnitte, deren Zuordnung zu spezifischen Funktionen meist nicht möglich ist. Bei diesen Sequenzen fällt auf, daß einzelne Bereiche mehrfach ($n > 2$) kopiert vorliegen. Das für die Erkennung solcher klassifizierter repetitiver Elemente konzipierte Computer- Programm ‚RepeatMasker‘ sagt für das humane Genom einen Repeat- Anteil von 50% voraus, einen Wert, den man als untere Grenze annehmen sollte (THE GENOME SEQUENCING CONSORTIUM, 2001). Diese Abschnitte teilen sich entsprechend ihrer Kopienzahl ($n = 10^5 - 10^6$) in mittel- und hochrepetitive DNS- Sequenzen auf. Eine weitere Einteilung erfolgt nach der Lage der repetitiven Elemente zueinander. Lange Folgen desselben Elements werden als „Tandem- Repeats“ bezeichnet. Sie werden nach der Größe der gebildeten Blöcke in Satelliten- DNS, Mini- und Mikrosatelliten unterteilt. Andere repetierte Sequenzen liegen unabhängig voneinander, verteilt über das gesamte Genom vor. Sie werden als verstreute repetitive (‚interspersed‘) Elemente bezeichnet.

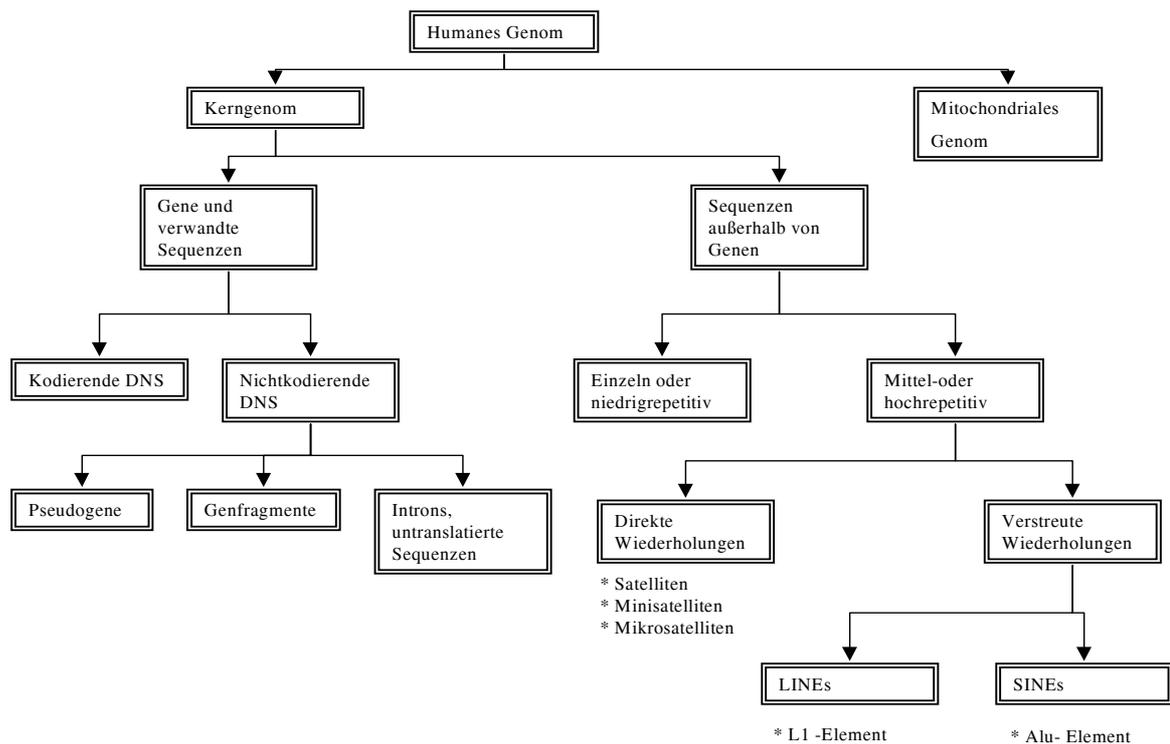


Abbildung 2-3: Schematische Darstellung des Aufbaus des humanen Genoms mit unterschiedlichen Klassen repetitiver Elemente (nach Strachan, T. und Read, A.P., 1996, S.173)

2.4.2 Repetitive Elemente

Hochrepetitive verstreute DNS- Familien in Säugetieren werden anhand ihrer Länge in zwei Gruppen eingeteilt: SINEs und LINEs (*Short bzw. Long interspersed nuclear elements*) (Singer, 1982). Demnach sind SINEs kürzer als 1000 nt. Sie kommen in Kopienzahlen von mehr als 10.000 im Genom vor. Nukleotidsequenzen von SINEs lassen durch ihre Ähnlichkeit zu Genen für tRNS oder die 7SL- RNS auf ihren Ursprung als prozessierte Pseudogene schließen. Sequenzen viraler Herkunft werden nicht zu den SINEs gezählt. Durch einen Retrotransposition (Rogers, 1983) genannten Prozeß werden die Gene über eine RNS- und eine cDNS- Zwischenstufe stark vervielfältigt („amplifiziert“) ins Genom integriert. Dieser Mechanismus erfordert das Vorhandensein einiger Strukturelemente, die man bei allen SINEs findet, wie den internen RNS- Polymerase III- Promoter und das Adenosin- reiche 3'-Ende. Die RNS- Zwischenstufe wird dabei nicht translatiert. Einige repetitive Elemente fusionierten zu Homo- oder Heterodimeren. Der Prozeß der Integration erfordert keine spezifischen

Erkennungsmotive in der Zielsequenz. Dennoch ist die Verteilung einzelner SINEs im Genom nicht ganz gleichmäßig. Die meisten SINEs sind inaktiv. Im Verlauf der Evolution ist bei ihnen durch Ansammlung von Mutationen ihre Fähigkeit zur Vermehrung durch Retrotransposition verloren gegangen.

2.4.3 Die Alu- Wiederholung

Die häufigste Form von SINEs sind die sogenannten ‚Alu- Repeats‘. Die Bezeichnung Alu ist von der Restriktionsendonuklease AluI abgeleitet, mit der erste Versuche zur Charakterisierung dieser Sequenzen durchgeführt wurden (Houck *et al.*, 1979). Vorläufige Auswertungen der veröffentlichten Sequenz des humanen Genoms haben ergeben, daß das Genom zu 10% aus Alu- Elementen besteht. Dies deutet auf eine Gesamtkopienzahl von einer Million hin (THE GENOME SEQUENCING CONSORTIUM, 2001, Venter *et al.*, 2001). Man findet sie in der DNS aller Primaten

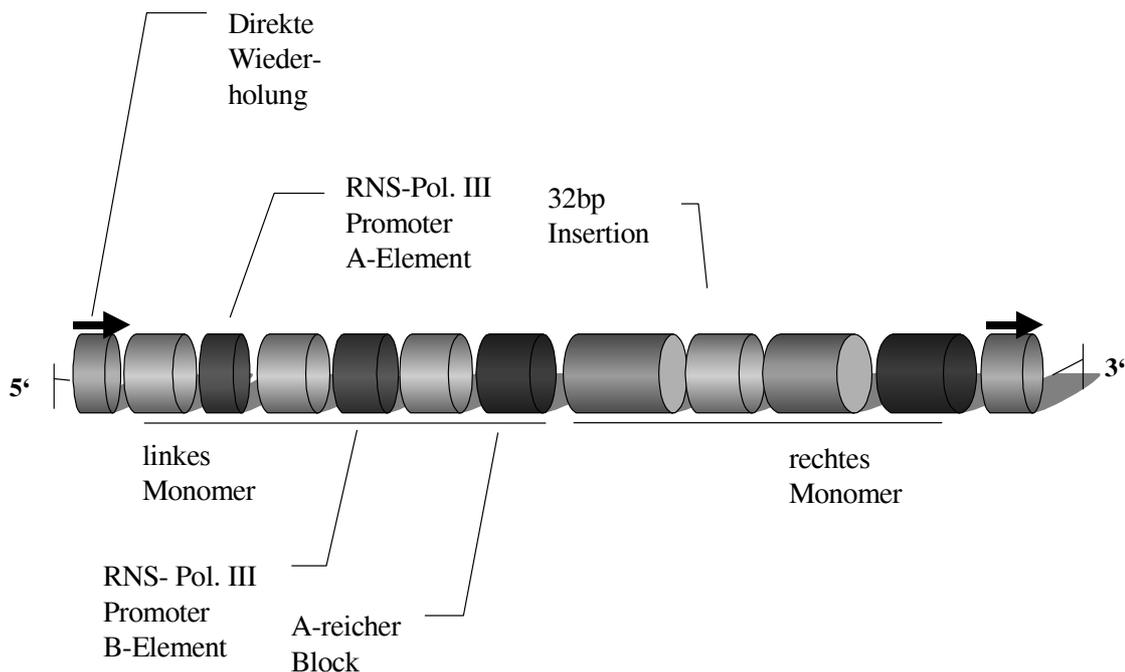


Abbildung 2-4: Schematische Darstellung der Bestandteile der Alu- Wiederholung: der interne RNS-Polymerase III- Promoter, die direkten Wiederholungen und der degenerierte poly- A- Bereich weisen auf eine Herkunft als Retrotransposon hin.

und schätzt, daß sie seit 65 Mio. Jahren existieren (Shen *et al.*, 1991). Die Retrotranspositionsaktivität ist mittlerweile fast vollständig zum Erliegen gekommen, so daß die exakte Position einzelner Alu- Elemente konserviert ist. In der Tat findet man 99,8% einander entsprechende Alu- Sequenzen an der gleichen Position im Genom aller Primaten (Deininger und Batzer, 1999). Schätzungen zu Folge hat unter 200 Neugeborenen eines ein Alu- Element mehr als die Eltern (Deininger und Batzer, 1995). Dieses ist zwar eine sehr geringe Zunahme, sie kann im Einzelfall aber durchaus zu phänotypisch relevanten Mutationen führen. Die Länge einzelner Alu- Elemente beträgt ungefähr 280 bp. Sie werden von kurzen, direkten, das heißt gleich orientierten, Wiederholungen flankiert. Die typische Alu- Sequenz ist ein Dimer. Eine Wiederholung besteht aus rund 120 bp. An beide Monomere schließt sich ein kurzes, Adenosinreiches Ende an. Der zweite Teilbereich beinhaltet weiterhin eine "Insertion" von 32 bp (tatsächlich handelt es sich um eine Deletion im ersten Fragment), die im ersten Monomer nicht vorkommt. Die Sequenzen der beiden repetitiven Teilbereiche zeigen eine auffallende Ähnlichkeit zur *elongation- arresting- domain* der 7SL- RNS (Siegel und Walter, 1986), einem Bestandteil des Signalerkennungspartikels, welches den Transport durch die Membran des endoplasmatischen Reticulums ermöglicht. Im Genom existieren auch Einzelkopien der Monomere. Sie werden als ‚Freies Linkes Alu- Monomer‘ (FLAM) oder ‚Freies Rechtes Alu- Monomer‘ (FRAM) bezeichnet und treten sehr viel seltener auf als die Dimere. Die Mehrzahl aller Alu- Sequenzen differieren untereinander um bis zu 20-30% (Deininger, 1981). Diese Differenzen ermöglichen die Einteilung in verschiedene Familien, welche durch ein unterschiedliches evolutionäres Alter erklärt werden. Die älteste Gruppe ist AluJ, die mittlere wird als AluS bezeichnet (Untergruppen: AluSx, AluSq, AluSp und AluSc); AluY beinhaltet die jüngsten Sequenzen (Untergruppen z.B.: AluYa5, AluYa8 oder AluYb8). Alle noch aktiv transponierenden Alu- Sequenzen gehören zur AluY- Familie (Batzer *et al.*, 1996). Obgleich die Anzahl der aktiv transponierenden Alu- Elemente gering ist, so ist doch ein bedeutender Anteil von phänotypisch relevanten Mutationen auf Insertionen von Alu- Sequenzen zurückzuführen. Der einfachste Fall ist eine Einfügung direkt in kodierende Sequenzen. Es wurden Fälle beschrieben, die zu pathologischen Veränderungen führten (Wallace *et al.*, 1991, Vidaud, *et al.*, 1993) oder auch ohne Auswirkungen auf den Phänotyp blieben (Brosius, 1991). Neben direkten

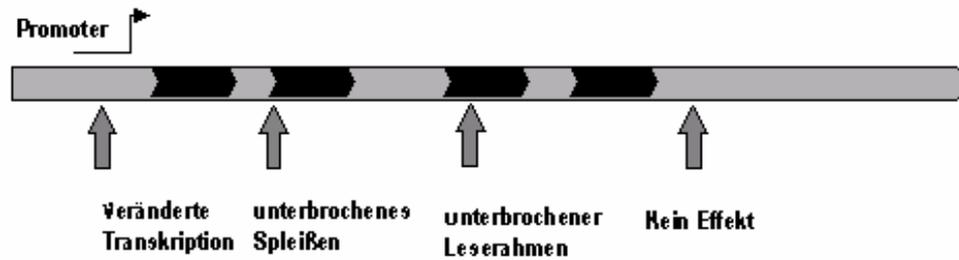


Abbildung 2-5: Auswirkungen von Alu-Insertionen in Abhängigkeit vom Ort der Insertion (dunkle Bereiche: Exons des Gens).

Auswirkungen (Exonverlängerung, Leserasterverschiebungen, vorzeitiger Abbruch) sind auch kryptische Spleißstellen innerhalb der Alu-Sequenz als Ursache für mutierte Proteine beschrieben worden (Makalowski, 1994). Durch Insertion von Alu-Elementen können außerdem Promotersequenzen und Polyadenylierungssignale in die Nähe von Genen geraten; auf analoge Weise kann es zur Reaktivierung von Pseudogenen kommen. Der umgekehrte Effekt wurde beobachtet, wenn nukleäre Faktoren an den internen Alu-Promoter binden und so die RNS-Polymerase blockieren (Wu *et al.*, 1990). Weiterhin wurde gezeigt, daß Alu-Sequenzen die Chromatinstruktur beeinflussen, und so die Stärke der Transkription beeinflussen können (Englander *et al.*, 1993). Weitaus häufiger beruhen Mutationen aber auf ungleichem *Crossing Over*, hervorgerufen durch Alu-Alu-Interaktionen. Das ist verständlich, da alle Alu-Elemente dazu beitragen können, nicht nur die der aktiven AluY-Familie. Alu-Elementen kommt statistisch eine dominierende Rolle bei ungleichem *Crossing Over* zu (Rüdiger *et al.*, 1995). So wird zum Beispiel die am häufigsten vorkommende Nicht-Robertson-Translokation auf den Einfluß von Alu-Elementen zurückgeführt. Die t(11;22)(q23;q11) Translokation konnte in 160 unabhängigen Familien nachgewiesen werden und führt bei den Betroffenen zu einem Syndrom, das mentale Retardierung, cranofaciale und Herzmißbildungen umfaßt (Hill *et al.*, 2000). Es sei aber auch erwähnt, daß dieser Typ von Mutationen im Verlauf der Evolution des humanen Genoms zur Erweiterung der Proteinvielfalt durch Bildung von Genfamilien führte. Beispiele sind die Glycophorin-Familie (Kudo *et al.*, 1989), α -Globine (Hess *et al.*, 1983), α 1-saure Glycoproteine (Merrit *et al.*, 1990) und die Haptoglobinfamilie (Erickson *et al.*, 1992).

Alu- Elemente sind über das gesamte Genom verteilt. Wenn man mit einer Million Kopien rechnet, ergibt sich daraus ein mittlerer Abstand von rund 3 kbp. In den R-Banden von Metaphase- Chromosomen ist die Dichte von Alu- Wiederholungen bis zu doppelt so hoch wie in den G- Banden (Korenberg und Rykowski, 1988). Das heißt, Alu- Elemente finden sich vor allen in genreichen (Korenberg und Engels, 1978) und rekombinationsaktiven (Kuhn und Therman, 1986) Regionen des humanen Genoms. Verbunden mit der Tatsache, daß Alu- Sequenzen relativ kurz sind, ist zu erwarten, daß sie sich häufig in unmittelbarer Nähe von Genen befinden. Dies wurde zum Beispiel ausgenutzt, um Promoterbereiche zu klonieren (Jendraschek und Kaminski, 1998).

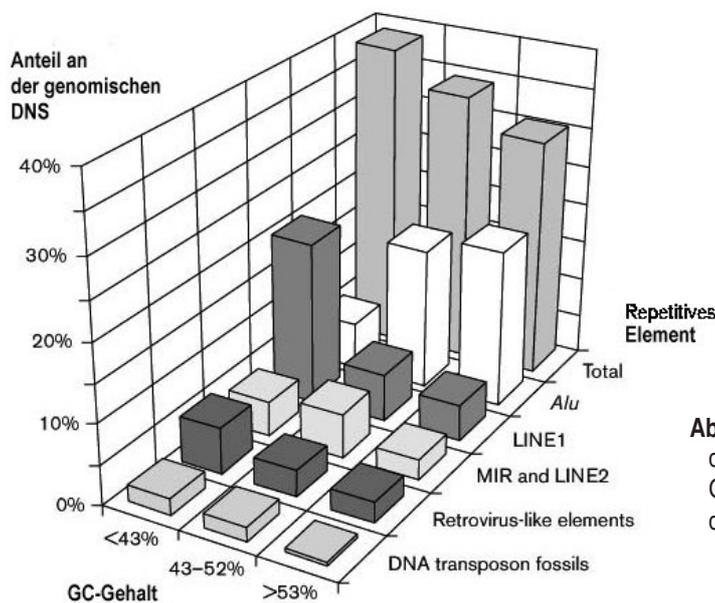


Abbildung 2-6: Ungleiche Verteilung der repetitiven Elemente im humanen Genom in Abhängigkeit vom GC- Gehalt der umgebenden Sequenz

Die ungleiche Verteilung der Alu- Elemente und die Konservierung ihrer Positionen im Genom deuten daraufhin, daß Alu- Sequenzen im Verlauf der Evolution eine für das Wirtsgenom positive Wirkung erworben haben, so daß ein Evolutionsdruck auf die Alu- Elemente angenommen werden kann. Die vorläufige Auswertung der Sequenzdaten des humanen Genoms unterstützt diese Hypothese (THE GENOME SEQUENCING CONSORTIUM, 2001). Sequenzhomologien zu konservierten Motiven in den Genomen von Bakterien lassen eine Funktion als Bruchpunkt bei Rekombinationsereignissen möglich erscheinen (Gebow *et al*, 2000). Eine gemeinsame Kernsequenz von 26bp wurde in allen untersuchten, an Rekombinationen beteiligten Alu- Elementen nachgewiesen. Diese Kernsequenz enthält das Pentamer 5'-CCAGC-3', das komplementär zum Rekombinations- *hotspot Chi* des *recBC*- abhängigen

Rekombinationssystems in *E.coli* ist (Rüdiger *et al.*, 1995). Die Annahme, daß diese Sequenz auch beim Menschen *Crossing Over* induzieren kann, wird dadurch unterstützt, daß für das Element 5'-CCAG-3' eine Beteiligung unabhängig von Alu-Elementen gezeigt werden konnte (Chou und Morrison, 1993)

2.4.4 Die Inter- Alu- Polymerase- Kettenreaktion (IA- PCR)

Die günstige Verteilung von Alu- Sequenzen im Genom wurde mit Hilfe der Polymerase- Ketten- Reaktion (PCR, Saiki *et al.*, 1985) ausgenutzt, um chromosomen-spezifische Sequenzen zu generieren (Ledbetter *et al.*, 1990), polymorphe Marker zu erschließen (Charlieu *et al.*, 1992), unbekannte Sequenzen ausgehend von Alu-Elementen zu klonieren (Minami *et al.*, 1995) oder neue Promotersequenzen zu finden (Jendraschek und Kaminski, 1998). Alle genannten Ansätze basieren auf der Inter- Alu-PCR (Nelson *et al.*, 1989). Dabei werden Primer eingesetzt, die spezifisch an die Alu-Sequenzen binden. Wenn zwei Alu- Elemente in der passenden Orientierung zueinander sowie dicht beieinander liegen, so kann die von den beiden Primern eingeschlossene genomische Sequenz durch PCR amplifiziert werden. Es gibt Protokolle, die mit einem Primer oder mit je zwei Primern für das 5'- und das 3'- Ende der Alu- Elemente arbeiten. Eine Fülle von Modifikationen des ursprünglichen Protokolls ist möglich (Kass und Batzer, 1995). Je nach Anwendung kann die Zusammensetzung des komplexen Amplikons optimiert werden. Eine wichtige Rolle für die Zusammensetzung des Inter- Alu- Amplikons spielen die verwendeten Primer, welche die folgenden Anforderungen erfüllen sollen (Tagle und Collins, 1992):

- Primer- Erkennungssequenzen müssen ganz am Rande der Alu-Konsensussequenz liegen, um den Anteil der eingeschlossenen repetitiven Sequenzen zu minimieren
- sie sollten hoch konserviert sein und so die Amplifikation möglichst vieler Inter-Alu- PCR- Produkte erlauben
- da das Alu- Element ein Dimer ist, muß darauf geachtet werden, daß der Primer nur eine Bindungsstelle in der Alu- Sequenz besitzt
- weiterhin gelten die allgemeinen Grundsätze zur Auswahl von Primersequenzen, wie die Berücksichtigung einer angemessenen *annealing*- Temperatur, das Vermeiden von *self- annealing*, etc.

Da die Inter- Alu- PCR oftmals verwendet wird, um spezifisch humane DNS-Fragmente neben Hamster- oder Maus- DNS zu amplifizieren (Nelson *et al.*, 1989), dürfen die verwendeten Primer möglichst wenig Sequenzähnlichkeit zu den entsprechenden Nager- DNS- Sequenzen aufweisen. Eine Auswahl von Primern, die für die Inter- Alu- PCR verwendet werden, ist in der folgenden Abbildung dargestellt. In

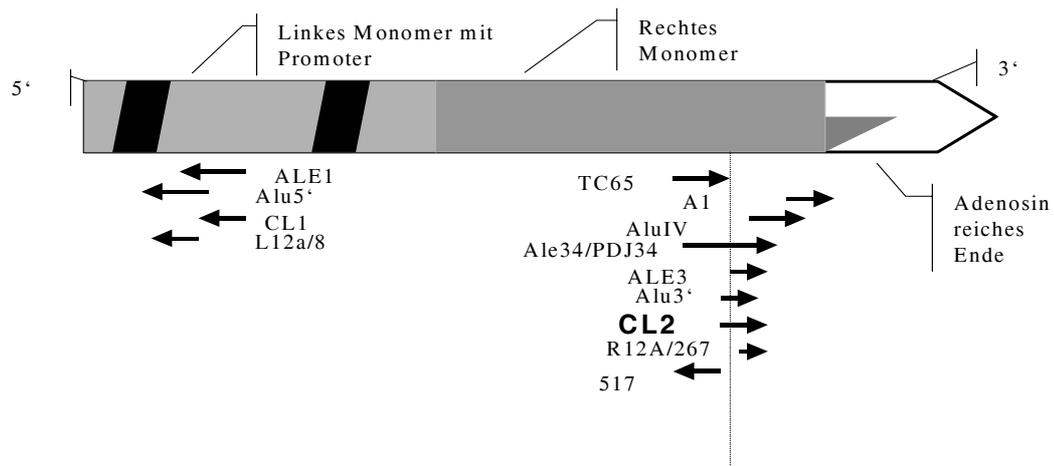


Abbildung 2-7: Schematische Darstellung der Alu- Sequenz und Lokalisation verschiedener Alu- Primer

der vorliegenden Arbeit wurde mit dem CL-2- Primer gearbeitet (Lengauer *et al.*, 1992), da mit diesem Primer in Pilotexperimenten besonders viele verschiedene Inter- Alu- PCR- Produkte erzeugt werden konnten.

Inter- Alu- PCR- Produkte erfüllen aufgrund ihrer Verteilung im Genom wichtige Voraussetzungen, um als Sonden in Matrix- CGH- Experimenten zu fungieren. Wie in der klassischen CGH wird mit diesen Sonden nach Deletionen oder Amplifikationen in der zu testenden DNS gesucht. Dazu müssen sie im amplifizierten bzw. deletierten Bereich liegen. Es ergeben sich aber auch Möglichkeiten für die Detektion von balancierten Translokationen. Eine Translokation wird als balanciert eingeordnet, wenn kein Verlust oder keine Zunahme an DNS festgestellt werden kann. Viele Translokationen sind aber nur cytogenetisch charakterisiert, d.h. relativ grob (4 Mbp, Strachan und Read, 1996). In der Tat werden bei einer anschließenden genaueren Charakterisierung häufig zusätzlich deletierte Bereiche an den Bruchpunkten gefunden

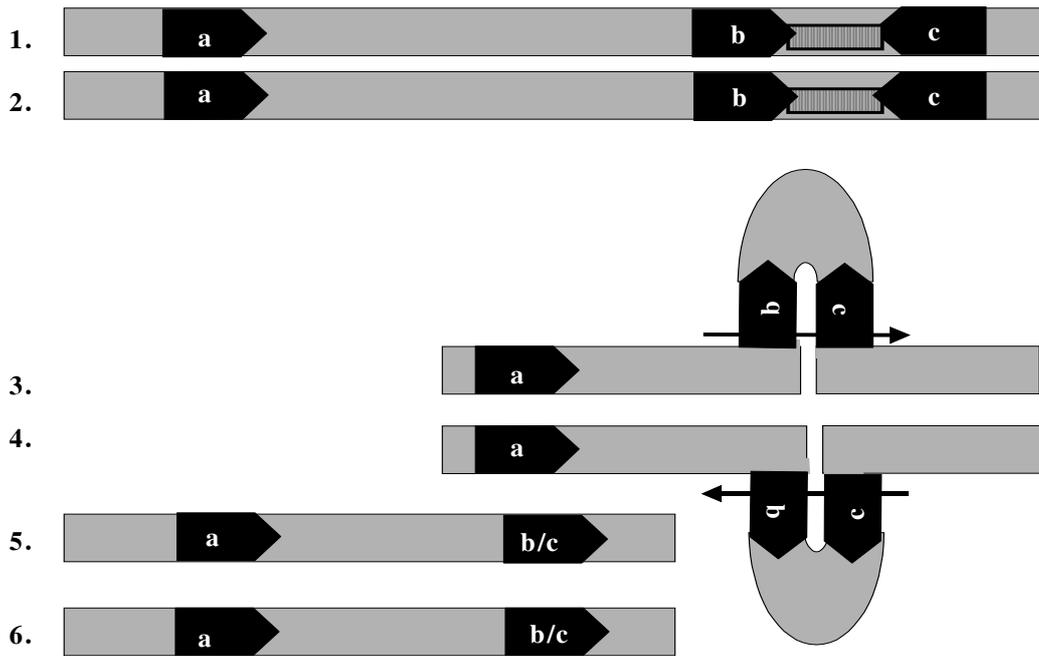


Abbildung 2-8: Excision von invertierten repetitiven Elementen (z.B. Alu), a, b, c stellen Alu- Elemente mit ihrer Orientierung dar, zwischen b und c gibt es in der Referenz- DNS ein putatives Inter- Alu- Produkt (1. und 2.). Invertierte repetitive Elemente können Kreuzformen ausbilden (3. und 4.) die von speziellen zellulären DNS-Reparatursystemen herausgeschnitten werden (Lobachew *et al.*, 2000) (5. und 6.), als Ergebnis ergibt sich in der mutierten DNS eine Deletion, die genau ein IA- PCR Produkt umfasst.

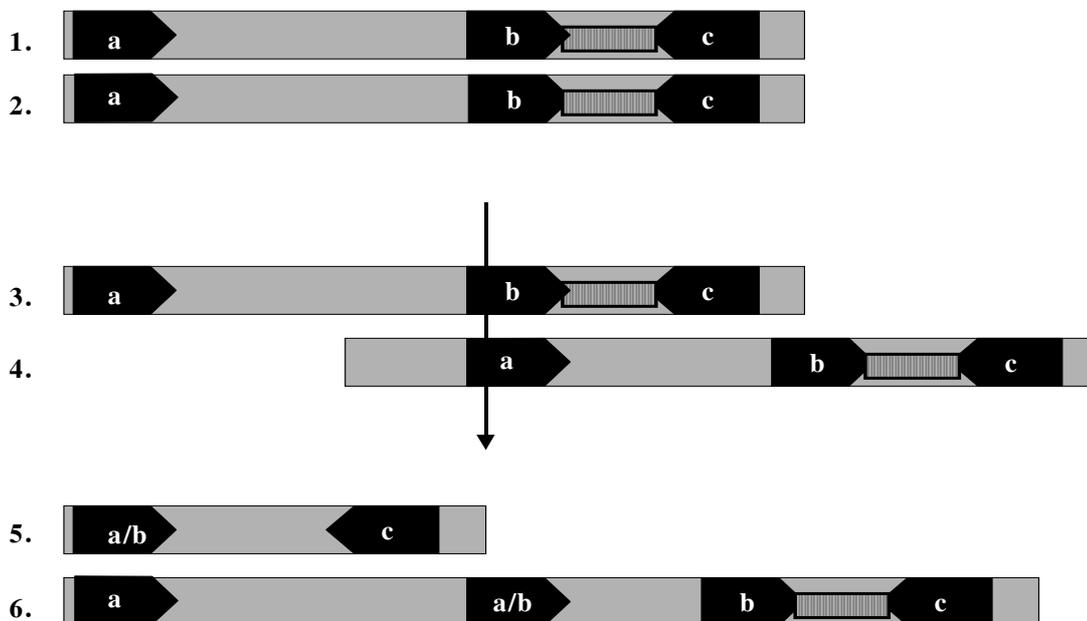


Abbildung 2-9: Ungleiches *Crossing Over* induziert durch Alu- Elemente. In der Referenz- DNS (1. und 2.) befinden sich mehrere Alu- Elemente, zwischen den Elementen b und c existiert ein putatives IA- PCR- Produkt, in 3. und 4. sind die beiden Schwesterchromatiden zueinander verschoben, das *Crossing Over* erfolgt durch die beiden Alu- Elemente b und c in 5. und 6., sind die beiden mutierten Abschnitte dargestellt, im Allel mit der Deletion kann kein IA- PCR mehr gebildet werden, in dem duplizierten Allel bleibt das PCR- Produkt erhalten.

(Kumar *et al.*, 1998). Als eine Methode für die genauere Charakterisierung würde sich die Matrix- CGH anbieten. Wie bereits erwähnt, sind die zugrunde liegenden nichthomologen Rekombinationen oftmals durch die repetitiven Alu- Elemente selbst induziert und man erwartet dann an den Bruchpunkten ein solches Element (siehe oben), dessen genomischer Kontext sich auf einer Seite verändert hat. Wenn auf dieser Seite in der Referenz- DNS ein IA- PCR- Produkt mit diesem Alu- Element vorhanden ist, so erwartet man in der veränderten DNS das Fehlen dieses spezifischen Produktes. Eine solche Konstellation wird in Bereichen mit hoher lokaler Konzentration von Alu- Elementen wahrscheinlicher, da sowohl die Möglichkeit von IA-PCR- Produkten als auch die Wahrscheinlichkeit von Rekombinationen in diesem Bereich erhöht sein sollte. Mögliche Mechanismen der durch Alu- Elemente hervorgerufene nicht homologe Rekombinationen und ihre Auswirkungen auf eventuell vorhandene Inter- Alu- PCR- Produkte sind in den beiden Abbildungen der vorherigen Seite dargestellt.

2.4.5 Das humane X-Chromosom

Der menschliche Chromosomensatz umfaßt 46 Chromosomen (46,XY oder 46,XX). Das X- Chromosom nimmt eine Sonderstellung ein, da es bei Mann (XY) und Frau (XX) in unterschiedlicher Kopienzahl vorliegt. Die Zelle muß deshalb in der Lage sein, diesen Dosisunterschied zu kompensieren. Der Mechanismus der X- Inaktivierung (Lyon, 1999) führt dazu, daß in somatischen Zellen sich immer nur ein X- Chromosom im aktiven Zustand befindet. Dies gilt nicht nur für Zellen des normalen weiblichen Organismus, sondern auch für Patienten mit überzähligen X- Chromosomen, was die verhältnismäßig ‚milden‘ Phänotypen dieser Krankheitsbilder erklärt (Klinefelter- Syndrom, OMIM 278850). Bei weiblichen Patienten konnten bis zu fünf X- Chromosomen pro Zelle nachgewiesen werden (Penta- X- Syndrom).

Das Vorhandensein zusätzlicher Autosomen ist dagegen meist letal. Einzig die kleinen und genarmen Chromosomen 13, 18 (beide selten) und 21 können in drei Kopien auftreten, führen dann aber zu schweren Phänotypen (z.B. Down- Syndrom, OMIM 190685). Zelllinien mit multiplen X- Chromosomen sind ein wichtiges Modellsystem zur Etablierung neuer Techniken für die Bestimmung von Aneuploidien oder Aneusomien (Pinkel *et al.*, 1998).

Mit einer Länge von 164 Mbp ist das X- Chromosom das achtgrößte humane Chromosom. Wichtige Erbkrankheiten konnten bereits früh aufgrund ihres spezifischen

Erbgangs mutierten Genen auf dem X- Chromosom zugeordnet werden (Rot- Grün-Farbblindheit (OMIM 303800), Hämophilie A (OMIM 306700) und B (OMIM 306900)). Die durchschnittlichen Gendichte ist aber vergleichsweise geringer (etwa 50% der autosomalen Gendichte, NCBI- *gene distribution map*), so daß eine vergleichsweise hohe Anzahl an nichtletalen Aberrationen erwartet werden kann. Zum anderen ist die Sequenzierung des X- Chromosoms weiter vorangeschritten, als die der meisten anderen großen Chromosomen. Obwohl X und Y zwei verschiedene Chromosomen sind, so ist ihr gemeinsamer Ursprung doch an ausgedehnten homologen Bereichen wie den pseudoautosomalen Regionen (0,35 bzw. 2,5 Mbp) zu erkennen. Es existiert ein weiterer homologer Bereich bei Xq21.3 (und Yp11.1), der ungefähr 4 Mbp groß ist (Mumm *et al.*, 1997). Diese Bereiche erfordern besondere Beachtung bei der Auswertung von komparativen Hybridisierungen. Aufgrund der aufgeführten Eigenschaften wurde das X- Chromosom als Modell für das gesamte Genom ausgewählt und untersucht.

2.5 Aufgabenstellung

Für die genomweite Suche nach chromosomalen Aberrationen stand zu Beginn der vorliegenden Arbeit nur die konventionelle CGH zur Verfügung. Diese Methode erlaubt aufgrund technischer Limitationen nur eine verhältnismäßig grobe Charakterisierung der gefundenen Veränderungen. Die geringe Auflösung und die fehlende Möglichkeit zur exakten Kartierung der gefundenen Mutationen machten eine Weiterentwicklung nötig, die diese technischen Grenzen aufhebt. Kurz nach Beginn der vorliegenden Arbeit (Oktober 1997) wurde eine Methode veröffentlicht, die die Grundlagen einer CGH auf Microarrays (Matrix- CGH) beschrieb und somit versprach, die genannten technischen Grenzen zu überwinden. Weitere Veröffentlichungen zeigten auf, daß dieses Ziel erreichbar ist.

Die Matrix- CGH wird in naher Zukunft die klassische CGH bei der Untersuchung von Amplifikationen und Deletionen im menschlichen Genom ersetzen. Zur Zeit existieren mehrere methodische Ansätze, dieses Ziel zu erreichen. Eine dieser Methoden beinhaltet die Erstellung einer repräsentativen Auswahl von Sequenzen der Ausgangs-DNS mit verringerter Komplexität und deren Hybridisierung auf korrespondierende Sonden. Die in dieser Arbeit verfolgte Aufgabenstellung umfaßte zunächst die

Etablierung der technischen Grundlagen der Matrix- CGH. Im folgenden sollte untersucht werden, ob mittels Inter- Alu- PCR geeignete, genomische Repräsentationen von verringerter Komplexität erhalten werden können. Nach erfolgter Charakterisierung des Inter- Alu- Amplikons sollte dann stellvertretend für das humane Genom das X- Chromosom untersucht werden. Es standen dabei Zelllinien zur Verfügung, die mögliche Aberrationen wie Amplifikationen und Deletionen des X- Chromosoms enthielten. Anhand dieser Zelllinien, die ein bis fünf X- Chromosomen pro Genom aufwiesen, sollte untersucht werden, welche Kopienzahlunterschiede detektiert werden können und wo die Grenzen der angewendeten Methode liegen. Weitere Zelllinien von Patienten mit verzögerter geistiger Entwicklung enthielten X- Chromosomen, denen jeweils eine deletierte Region bei Xq21 fehlte. Die Lage und die Größe dieser Deletionen waren bereits durch konventionelle Methoden bestimmt worden. Durch Vergleich mit den zu ermittelnden Daten der Matrix- CGH sollte bestimmt werden, wie genau sich solche Mutationen eingrenzen lassen. In Abhängigkeit von den erreichten Zwischenzielen sollten dann in Verbindung mit den zur Verfügung stehenden Daten des Humanen Genomprojekts neue Kandidatengene für Mentale Retardierung bestimmt werden.