

5 Diskussion

5.1 Grundlagen

Chromosomale Aberrationen sind bereits oft als Ursache von Erbkrankheiten identifiziert worden. Auch bei der Entstehung und der Progression von Tumoren sind Verlust und Amplifikation von genomischen Bereichen entscheidend für den Verlauf der Erkrankung. Die Untersuchung solcher Aneusomien stellt deshalb ein wichtiges Hilfsmittel für die Erstellung der Diagnose oder einer möglichen Prognose dar. Die Karyotypanalyse ermöglicht zum Beispiel die Diagnose des Down- Syndromes, dem das Auftreten eines zusätzlichen Chromosom 21 zugrunde liegt. Zumeist umfassen genomische Aberrationen aber kleinere Bereiche, die auch mikroskopisch nicht mehr erkannt werden können. Die Entwicklung der vergleichenden Genom- Hybridisierung (CGH) eröffnete neue Möglichkeiten im Bereich dieser submikroskopischen Veränderungen. Insbesondere in der Krebsforschung stellt die CGH ein Routineinstrument zur Aufklärung der für die Tumorprogression typischen genomischen Aberrationen dar. Die Auflösung dieser Methode reicht jedoch nicht aus, den ermittelten Veränderungen einzelne Gene zuzuordnen. Die Bestimmung der molekularen Ursachen dieser Erkrankungen erfordert aber die Identifizierung der mutierten Gene. Die Matrix- CGH ist eine Weiterentwicklung der vergleichenden Genom- Hybridisierung, die durch eine höhere Auflösung eine Eingrenzung der Amplikons oder Deletionen auf wenige betroffene Gene ermöglichen soll. Die technische Entwicklung der genannten Methode ist bis zum jetzigen Zeitpunkt noch nicht abgeschlossen. Verschiedene Ansätze (siehe 2.3) werden verfolgt, um eine möglichst hohe Auflösung erreichen zu können.

Die in dieser Arbeit vorgestellte Variante der Matrix- CGH ist durch die Verwendung einer gegenüber der Ausgangs- DNS in ihrer Komplexität verringerten repräsentativen Auswahl von genomischen Sequenzen gekennzeichnet. Für die Reduktion der Komplexität fand die Inter- Alu- PCR Anwendung. Die Proben wurden durch PCR- Amplifikation mit dem Alu- spezifischen CL-2- Primer generiert. Die

korrespondierenden Sonden wurden *In Silico* aus bekannten Datenbanksequenzen ausgewählt und durch PCR mit spezifischen, internen Primerpaaren erzeugt.

5.2 Charakterisierung des Inter- Alu- Amplikons

Wie bereits erwähnt, verwendet die in dieser Arbeit vorgestellte Methode anstelle der genomischen DNS des Patienten ein IA- PCR- Produkt dieser DNS für die Hybridisierung. Das durch selektive Amplifikation der Inter- Alu- Sequenzen erzeugte PCR- Produkt muß dabei bei verringerter Komplexität repräsentativ für die Ausgangs- DNS sein. Das heißt zum einen, daß die einzelnen IA- Klone möglichst gleichmäßig über das Genom verteilt sind und zum anderen, daß die Mengenverhältnisse einzelner Klone untereinander erhalten bleiben. Die grundsätzliche Eignung der Inter- Alu- PCR für die Erzeugung von repräsentativen Klonen einer Ausgangs- DNS ist bekannt (Nelson *et al.*, 1989) und wurde u.a. durch FISH- Analyse der erzeugten Amplikons nachgewiesen (Ledbetter, S. *et al.*, 1990). In der vorliegenden Arbeit wurde untersucht, ob die gewählte Variante der IA- PCR die in der Einleitung geforderten Werte für Auflösung und Verringerung der Komplexität erfüllen kann.

Für die Bestimmung dieses Faktors wurde die Anzahl verschiedener IA- PCR- Klone pro Mbp genomischer Sequenz experimentell ermittelt. Dazu wurden IA- Amplikons von Megabasen- YACs erzeugt und kloniert. Durch die gelektrophoretische Auswertung der einzelnen Kolonie- PCR- Produkte wurde ermittelt, wie viele voneinander verschiedene Klone pro YAC erhalten wurden und wie groß die durchschnittliche Länge der einzelnen IA- PCR- Klone war. Die erhaltenen Werte von ungefähr fünf Klonen pro Mega- YAC mit einer mittleren Länge von etwa 500bp ergaben eine Anreicherung um den Faktor 400 ($1\text{Mbp} / (5 \cdot 500\text{bp})$). Theoretisch mußte mindestens ein Faktor von 133 erreicht werden (siehe 2.3), diese Bedingung wurde somit erfüllt. Die Zahl der IA- PCR- Klone pro Megabase bestimmte auch die maximale Auflösung der Methode. Die erreichten fünf Klone pro Mbp entsprechen einer Auflösung von 200 kbp.

Nachdem sich die direkte Hybridisierung von IA- PCR- Amplikons auf IA- PCR- Klone aufgrund der noch in ihnen enthaltenen hochrepetitiven Alu- Endsequenzen als nicht erfolgreich erwies, wurde ein anderer Weg eingeschlagen. Dazu wurden die klonierten IA- PCR- Produkte sequenziert und die so erhaltenen Daten zur Auswahl von

klonspezifischen Primerpaaren benutzt. Dabei konnte durch Maskieren der repetitiven Sequenzen sichergestellt werden, daß in der anschließenden Hybridisierung keine durch *Repeats* induzierten, falschpositiven Signale zu erwarten waren. Dieser Ansatz machte jedoch einen hohen technischen Aufwand für jede einzelne Sonde notwendig. Mit der schnell wachsenden Zahl an in Datenbanken verfügbaren genomischen Sequenzen konnte dazu übergegangen werden, nicht die durch Sequenzierung erhaltenen Daten sondern die Datenbanksequenzen selbst für die Auswahl von spezifischen Primerpaaren zu benutzen. Für diese Methode wurde ein Programm geschrieben, daß die Extraktion von putativen IA-PCR- Sequenzen aus genomischen Sequenzdaten übernahm. Es wurde zunächst überprüft, welche Programmparameter die experimentell erhaltenen Daten zur Anzahl und Verteilung der gefundenen IA- PCR- Klone am besten wiedergeben. Die Auswertung der experimentell bestimmten Primererkennungssequenzen (siehe 4.1.1.4) ergab, daß im Durchschnitt etwa eine Fehlpaarung pro Primer unter den eingesetzten PCR- Bedingungen toleriert wird. Mit diesem Wert und einer Begrenzung der Produktlänge auf 500- 2000bp kann die Anzahl der theoretisch vorhergesagten IA-PCR- Klone am besten mit der Anzahl der experimentell erhaltenen Klone korreliert werden. Aus den für verschiedene Programm- Parameter ermittelten Werten läßt sich ableiten, daß eine Veränderung der Stringenz der Reaktionsbedingungen, zum Beispiel durch Erhöhung der Magnesiumkonzentration in der PCR, die Anzahl der gefundenen IA-PCR- Klone sehr stark erhöhen würde, was zu einer Erhöhung der Komplexität des Amplikons führen würde. Eine Erhöhung der Komplexität bedeutet schwächere Signale bei einer höheren Auflösung. Somit könnte die Methode schnell besseren technischen Möglichkeiten bei der Signaldetektion angepaßt werden. Alle *In Silico* generierten Sonden konnten experimentell detektiert werden.

5.3 Ermittlung der experimentellen Grenzen der vorgestellten Methode

Ein wichtiger Parameter zur Beurteilung einer Meßmethode ist das erreichbare Verhältnis von spezifischem Signal zu unspezifischen Hintergrund. Im Abschnitt 4.1.4.1 wurde beschrieben, daß unspezifische Signale maximal 2% des Wertes für spezifische Signale erreichen. Unabhängig vom Signal/ Rauschverhältnis schwanken die Werte für die gemessenen DNS- Mengenverhältnisse um den erwarteten Mittelwert. Die in den Abbildungen 4-16 bis 4-21 dargestellten Ergebnisse lassen erkennen, daß die

Werte dabei sehr stark differieren. Zur Abschätzung dieser Schwankungen wurde für jedes Experiment die Standardabweichung der Signalintensitätsverhältnisse der autosomalen Kontrollklone bestimmt. Dazu wurden die Signalverhältnisse der vom Chromosom 22 abgeleiteten Klone auf 1,00 normiert und von diesen 96 Werten die Standardabweichung ermittelt. Die Auswertung aller, nicht nur der hier dargestellten, Versuche ergab, daß die Standardabweichung zwischen den Werten 0,1 und 0,3, entsprechend 10-30% der gemessenen Signalintensität schwankt. Ein Wert von 0,2 oder 20% kann für ein durchschnittliches Experiment als typisch angesehen werden und stellt so die Auflösungsgrenze für die gesamte Methode dar. Generell bedeuten kleinere Werte für die Standardabweichung bessere Möglichkeiten, kleine DNS-Mengenunterschiede erfassen zu können. Außerdem können über diesen Parameter verschiedene methodische Ansätze aus der Literatur miteinander verglichen werden:

Tabelle 5-1: Vergleich der ermittelten durchschnittlichen Standardabweichungen verschiedener Matrix- CGH-Veröffentlichungen.

Quelle	Ermittelte Standardabweichung	Sonden-Typ
Albertson <i>et al.</i> , 2000	14,5%	BAC
Pinkel <i>et al.</i> , 1998	7%	BAC
Bruder <i>et al.</i> , 2001	4-10%	BAC
Pollack <i>et al.</i> , 1999	9-40%	cDNS
Diese Arbeit	20% (10-30%)	IA-PCR

Offensichtlich können für diesen Parameter bei Experimenten mit als Sonden verwendeten BAC- oder PAC- Klonen bessere (niedrigere) Werte erreicht werden. Die längeren genomischen Klone entsprechen praktisch im Durchschnitt mehr als hundert IA- PCR- Produkten, was die individuellen klonspezifischen Abweichungen minimiert. In einer methodisch ähnlichen Arbeit, in der ebenfalls Inter- Alu- PCR verwendet wurde (Geschwind *et al.*, 1998), sind keine Vergleichsdaten angegeben. Die graphische Darstellung der ermittelten Werte läßt jedoch vermuten, daß die Standardabweichung etwa 1,5-2x so hoch sind wie in der vorliegenden Arbeit.

In hier nicht aufgeführten Experimenten wurde ohne Erfolg versucht, durch Optimierung der experimentellen Bedingungen diese Schwankungen zu minimieren. Auch die Auswertung mehrerer paralleler Experimente führt nicht zu ‚besseren‘ Werten, da die verwendeten IA- PCR- Klone dazu tendierten, bei verschiedenen

Experimenten entweder immer zu hohe oder immer zu niedrige Werte zu liefern. Dazu sei erwähnt, daß verschiedene Klone sehr stark in ihrer Signalstärke differieren. Intensitätsunterschiede bis zu einem Faktor von 100 sind beobachtet worden. Bei geringen Signalstärken ist die korrekte Subtraktion des Hintergrundsignals Voraussetzung für verlässliche Ergebnisse.

Bei der Auswertung von Matrix- CGH- Experimenten fiel auf, daß die experimentell ermittelten Signalstärken die eingesetzten DNS- Mengen nicht im Verhältnis 1:1 wiedergaben. Es wurden stets niedrigere Werte ermittelt als theoretisch erwartet. Für dieses auch als Kompression bezeichnet Phänomen, werden u.a. unspezifisches Hybridisieren einiger Sequenzen, Probleme bei der korrekten Subtraktion des Hintergrundsignals oder die gegenseitige Beeinflussung der verwendeten Farbstoffmoleküle („Crosstalk“) verantwortlich gemacht. Es ist gegenwärtig nicht möglich, diese Faktoren rechnerisch zu erfassen oder sie sogar herauszurechnen (Mohapatra *et al.*, 1997 (für CGH)). Der Zusammenhang zwischen den eingesetzten DNS- Mengen und den ermittelten Signalintensitätsverhältnissen wird für kleine Werte zumeist als lineare Gleichung der Form:

$$(5.1) \quad y = mx + n \quad \left| \begin{array}{l} \text{mit } x = \text{theoretisch erwarteter Wert,} \\ y = \text{experimentell ermittelter Wert} \end{array} \right.$$

angegeben. Die lineare Regression der in den Abb. 4-16 bis 4-21 dargestellten Meßwerte ergibt den folgenden Zusammenhang zwischen eingesetzten DNS- Mengen (x) und gemessenen Intensitätsverhältnissen:

$$(5.2) \quad y = 0,33x + 0,63 \quad \left| \begin{array}{l} y = \text{Meßwert, } x = \text{tatsächlicher Wert} \end{array} \right.$$

oder

$$(5.3) \quad x = (y - 0,63) / 0,33$$

Da in einigen Publikationen ebenfalls komparative Hybridisierungen mit einem bis fünf X-Chromosomen pro Genom (oder entsprechende Experimente) erfolgten, können die so ermittelten Werte für den Anstieg (m) und für das absolute Glied (n) mit den in der vorliegenden Arbeit ermittelten Daten verglichen werden:

Tabelle 5-2: Vergleich der Parameter der Ausgleichsfunktion für Matrix- CGH- Ergebnisse verschiedener Autoren

m	n	Quelle
0,28 ¹	0,4	Pollack <i>et al.</i> , 1999
0,37	0,27	Pinkel <i>et al.</i> , 1998
0,25 ²	0,8	Solinas-Toldo <i>et al.</i> , 1997
0,33	0,63	diese Arbeit

Generell bedeuten steilere Anstiege (höhere Werte für m) bessere Chancen, geringere Kopienzahlunterschiede bei gleicher Standardabweichung voneinander unterscheiden zu können.

Dem Wert für n kann keine Bedeutung zugeordnet werden. Da der Anstieg m für große genomische Klone, die repetitive Sequenzen enthalten (Pinkel *et al.*, 1998), größer ist als für kleine, fast repetitionsfreie PCR- Produkte (Pollack *et al.*, 1999, diese Arbeit) kann man davon ausgehen, daß der Einfluß von unzureichend unterdrückten repetitiven Sequenzen auf den Faktor m sehr gering sein muß. Der in der vorliegenden Arbeit ermittelte Wert für m ordnet sich zwischen die vorhandenen Literaturwerte ein.

Auf die Bedeutung der vorgestellten Parameter von Standardabweichung und Anstieg der Ausgleichsgeraden für die Beurteilung der Meßergebnisse soll im folgenden noch einmal gesondert eingegangen werden. Ausgehend von den experimentell erreichten Werten für $m=0,33$ und $\sigma=0,2$ sind in der Abbildung 5-1 die für jedes eingesetzte DNS-Mengenverhältnis die zu erwartenden Meßwerte mit den Intervallgrenzen einer Standardabweichung dargestellt. Die Intervalle geben somit den Bereich an, in dem mit einer Wahrscheinlichkeit von 68% (entspricht 1σ) die gemessenen Daten erwartet werden können. Es ist erkennbar, daß benachbarte Intervalle stark überlappen. Diese können daher nicht signifikant voneinander unterschieden werden. Im vorliegenden Beispiel können deshalb Klone erst als amplifiziert erkannt werden, wenn sie gegenüber der Referenz in mindestens drei Kopien vorliegen. Kleinere Standardabweichungen und höhere Werte für den Anstieg der Ausgleichsgeraden führen zu geringeren

¹ aus Diagramm durch Extrapolation abgeschätzt

² lineare Regression für angegebene Werte für 1, 2 und 3 Kopien des EBV- Genoms

Überlappungen und somit zu höherer Trennschärfe. Die bisher gemachten

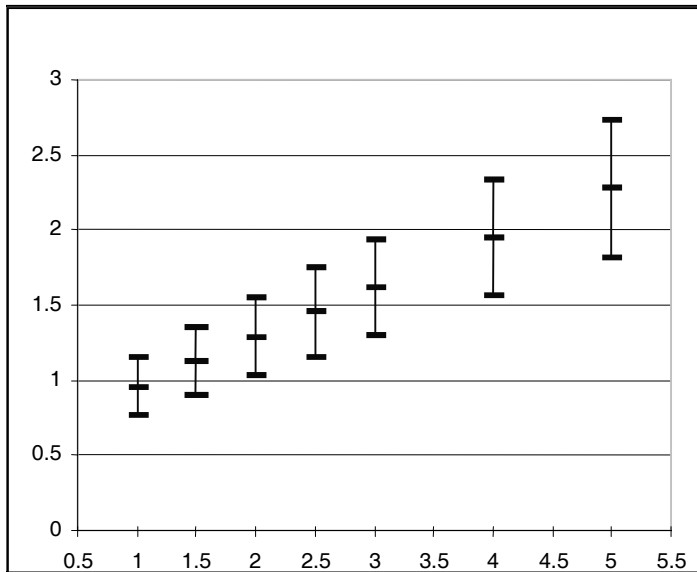


Abbildung 5-1: Zusammenhang zwischen den eingesetzten Kopienzahlverhältnissen (Abzisse) und den erwarteten Werten mit den zugehörigen Konfidenzintervallen von einer Standardabweichung (Ordinate)

Ausführungen beschreiben die Detektion eines einzelnen amplifizierten Klones. Wenn eine genomische Aberration mehrere Klone umfaßt, so verkleinern sich die Konfidenzintervalle und somit die Überlappungen mit steigender Anzahl der Klone (entspricht den Stichproben). Dadurch ist zu erwarten, daß auch der Verlust einer Kopie detektiert werden kann, sofern sich diese Deletion über genügend auswertbare IA- PCR Klone erstreckt.

Bei der Betrachtung der Meßwerte für die Hybridisierungen der verschiedenen multiplen X- Zelllinien fällt auf, daß nicht nur die Meßwerte untereinander differieren, sondern, daß einzelne Klone unabhängig vom eingesetzten Kopienzahlverhältnis immer ein Verhältnis von eins zu eins ergeben. Eine Zuordnung zum X- Chromosom wurde für alle aufgeführten IA- PCR- Klone durch BLAST- Analyse sichergestellt, der Einfluß von bekannten repetitiven Sequenzen wurde ebenfalls ausgeschlossen, so daß für dieses Verhalten keine Erklärung gefunden werden konnte. Aufgrund dieser Beobachtung ergibt sich die Frage, wie groß der Anteil der IA- PCR- Klone ist, die diese Eigenschaft nicht aufweisen. Dazu wurde im weiteren unterschieden, ob die Klone zumindest überhaupt als amplifiziert erkannt wurden oder ob sie auch die exakte Kopienzahl

wiedergeben. Wie der Abbildung zu entnehmen ist, nimmt der Anteil der Klone,

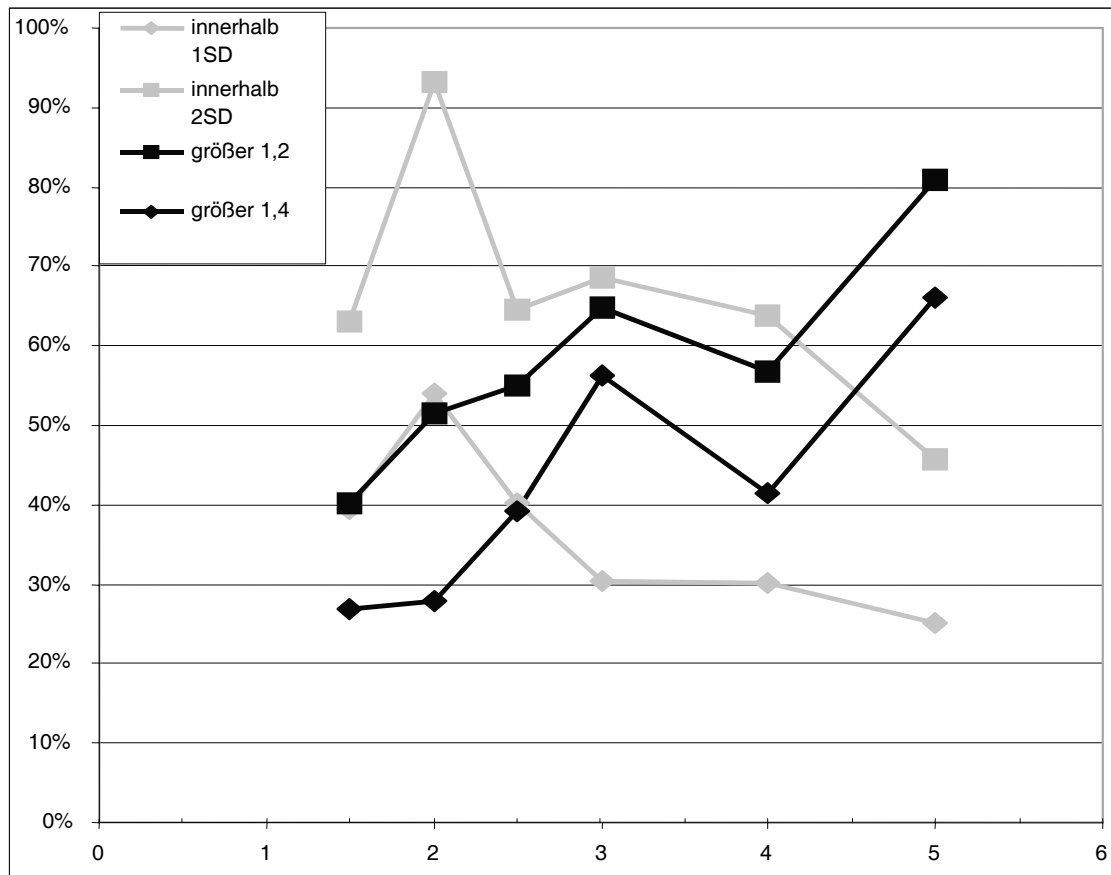


Abbildung 5-2: Anteil der IA- PCR- Klone, die zu den eingesetzten Kopienzahlverhältnissen korrespondierende Meßwerte ergeben, Ordinate: Anteile der Klone in Prozent, Abzisse: eingesetzte Kopienzahlverhältnisse, die grau markierten Kurven stellen den Anteil der Klone dar die im Konfidenzintervall, begrenzt durch eine (1SD) oder zwei (2SD) Standardabweichungen vom theoretisch erwarteten Mittelwert liegen, die schwarz markierten Kurven geben den Anteil von Klonen wieder, die signifikant (1,2 entspricht 1σ , 1,4 entspricht 2σ) als amplifiziert klassifiziert werden können

die signifikant als amplifiziert eingeordnet werden können, mit der Höhe des eingesetzten Kopienzahlverhältnisses zu und erreicht für das 1σ - Konfidenzintervall 80,95% und das 2σ - Intervall 66,07% bei einem eingesetzten Verhältnis von fünf zu eins. Der Anteil der Klone, die innerhalb eines Konfidenzintervalls auch um den theoretisch erwarteten Mittelwert liegen, beträgt bei dem eingesetzten Verhältnis von fünf zu eins 25,15 % für das 1σ - Konfidenzintervall und 45,61% für das 2σ -Konfidenzintervall. Die Werte für die numerisch kleineren Verhältnisse erreichen höhere Anteile, da die Konfidenzintervalle für diese stärker mit dem Normierungswert eins überlappen und so auch Anteil falsch zugeordneter Klone mitgezählt wird. Verhältnisse mit einem Faktor größer als zwei ergeben offensichtlich keine großen

Änderungen des Anteils der im Konfidenzintervall liegenden Klone mehr, so daß davon ausgegangen werden kann, daß mehr als 25,15% (für das 2σ - Konfidenzintervall) der Klone die eingesetzten Kopienzahlverhältnisse korrekt wiedergeben (45,65% für das 1σ Konfidenzintervall).

Nachdem bisher die Gesamtheit der Meßwerte einer Hybridisierung betrachtet wurde, sind im folgenden die ermittelten Signalintensitätsverhältnisse einzelner IA- PCR- Klone in Abhängigkeit von den eingesetzten Kopienzahlverhältnissen dargestellt. Dabei kann die Gültigkeit der aus der Gleichung 5.1 ermittelten Formel überprüft werden. Die

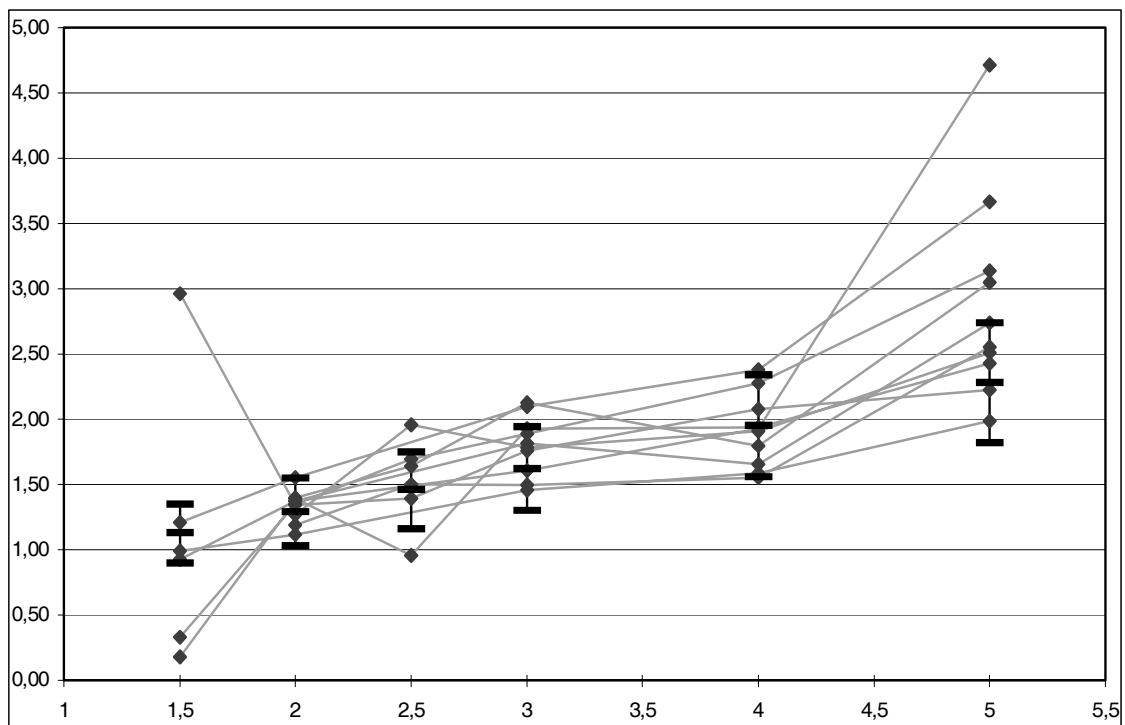


Abbildung 5-3: Zusammenfassende graphische Darstellung aller gemessenen Signalintensitätsverhältnisse von zehn einzelnen IA- PCR- Klonen, Ordinate: ermittelte Signalverhältnisse, Abzisse eingesetztes Kopienzahlverhältnis, die 1σ - Konfidenzintervalle wurden aus der Abbildung 5-1 übertragen

vorgestellten Daten ermöglichen zusammenfassend die Aussage, daß für eine Matrix- CGH mit aus Datenbank- Sequenzen generierten IA- Sonden Ergebnisse erwartet werden können, wie sie mit vergleichbaren, ähnlichen (konventionellen) Ansätzen erzielt wurden.

5.4 Kartierung einer hemizygoten Deletion

Für die Kartierung der Deletionen bei Xq21 wurde ein möglichst dichtes Raster von IA-PCR- Klonen der Region erzeugt. Die korrespondierende, umgebende Region von 70-90 Mbp konnte mit 28 IA- PCR- Klonen abgedeckt werden. Davon waren 26 durch *In Silico*- Vorhersage und zwei durch Sequenzierung von klonierten YAC- Inter- IA- PCR- Produkten bestimmt worden. Damit konnte die Region mit einer Auflösung von durchschnittlich 700 kbp untersucht werden. Die erreichte Auflösung spiegelte die damalige (Stand Juli 2000) Verfügbarkeit von genomischen Sequenzen sowie die vergleichsweise geringe lokale Konzentration von Alu- Elementen wieder. Aufgrund der in 4.1.1.2 gemachten Abschätzung von 15.000 IA- PCR- Klonen im Genom hätte man dagegen maximal 200kbp Auflösung erwarten können.

Um die vorhandenen IA- PCR- Sonden maximal ausnutzen zu können, wurde die Auswertung so gestaltet, daß jeder einzelne Spot untersucht wurde. Um keine Signale wegen zu geringer Intensität zu verlieren wurde die Verstärkung beim Einlesen (,Scannen‘) erhöht, wobei in Kauf genommen wurde, daß einzelne Meßwerte den Maximalwert (Signalsättigung) erreichten. Durch diese Maßnahme verschlechterte sich auch die Standardabweichung der gemessenen Signalintensitätsverhältnisse. Da vollständige Deletionen aber sehr große Unterschiede in der Signalintensität erwarten lassen, konnte diese Abweichung toleriert werden. Zur Lokalisierung der Deletionen wurde der gleitende Durchschnitt über fünf benachbarte IA- PCR- Klone betrachtet. Durch dieses Glätten der Meßwerte lassen sich die Aberrationen leichter detektieren. Um die so erkannte Deletion einzugrenzen, wurden dann aber wieder die eigentlichen Werte Signal für Signal betrachtet. Daten, die um eine Standardabweichung kleiner als der normierte Mittelwert der Kontrollspots waren, wurden der Deletion zugeordnet. Dadurch konnten die Grenzen des deletierten Bereichs auf ein Intervall eingengt werden, das durch den letzten IA- PCR- Klon, der als ,vorhanden‘ und den ersten, der als ,deletiert‘ klassifiziert wurde, bestimmt wird. Im Durchschnitt waren diese sechs bestimmten Intervalle 596 kbp groß, was ungefähr der erwarteten Auflösung entspricht. Die erhaltenen Ergebnisse wurden durch PCR mit den Sondenprimern auf der Patienten- DNS überprüft. Dabei wurden die erhaltenen Daten zu mehr als 90% bestätigt. Ein weiterer Beweis für die Plausibilität der ermittelten Daten ist die Position

des Gens TBX22. Je nach Größe der Deletion besitzen die Patienten noch eine Kopie des Gens oder es wurde verloren. In der Zelllinie 303/99 ist dieses Gen (siehe Abb. 4-19) deletiert, in den beiden anderen (sehr wahrscheinlich) nicht. Der entsprechende Patient hat als einziger der drei Patienten eine Lippen-, Kiefer- Gaumenspalte. Das TBX22-Gen wurde erst vor kurzem entdeckt (Laugier-Anfossi und Villard, 2000) und gehört zu einer Familie evolutionär konservierter Transkriptionsfaktoren (T-Box- Gene, Bollag *et al.*, 1994). Für einige Mitglieder der T-Box- Genfamilie konnte eine Funktion in der frühen Phase der Embryogenese, bei Morpho- und Organogenese nachgewiesen werden. Ein spezifischer Zusammenhang zwischen mutierten T- Box- Genen und Mittelliniendefekten (die sich zum Beispiel als Lippen-, Kiefer- Gaumenspalten manifestieren) konnte im Modellsystem Zebrafisch nachgewiesen werden (Dheen *et al.*, 1999). Die Position dieses Gens stellt also einen Marker dar, mit dem die Ausdehnung der Deletionen auf der dem Centromer zugewandten Seite bestimmt werden kann. Nur der Patient, bei dem die Deletion wahrscheinlich über diese Position hinwegreicht, hat den Phänotyp des mutierten TBX- Gens. Distal konnte die Begrenzung des deletierten Bereichs für die drei Patienten einheitlich festgelegt werden. Insbesondere konnte gezeigt werden, daß der deletierte Bereich nicht in die XY- Homologie- Region hinein reicht, in der die Auswertung der Meßwerte erschwert wäre. Die Lokalisierung von genomischen Markern aus der homologen Region (Schwartz *et al.*, 1998³) durch BLAST- Analyse legt den Beginn dieses Bereichs auf den Index 85,1 Mbp fest. Das Intervall, in dem Bruchpunkt der Deletion gefunden wurde, liegt zwischen 82,9 und 83,2 Mbp.

Der Vergleich der in dieser Arbeit durch Matrix- CGH ermittelten Daten mit den ursprünglich durch *Southern Blotting* bestimmten Grenzen des deletierten Bereichs (van der Maarel *et al.*, 1995) ergibt ähnliche Ergebnisse. Die für alle drei Patienten ungefähre Größe von weniger als zehn Mbp und die Lage hinsichtlich schon kartierter Gene stimmen für beide Untersuchungen überein. Bei der Betrachtung der gefundenen Unterschiede muß beachtet werden, daß die Darstellung der Ergebnisse abhängig von der Qualität der verwendeten genomischen Karten ist. Der lokal relativ hohe Anteil von repetitiven Sequenzen erschwert eine exakte Kartierung von genomischen Klonen in

³ die in dieser Arbeit ebenfalls aufgeführte Darstellung der kartierten Deletionen für die Patienten 303/99 (=NP) und 526/00 (=XL-62) sind nicht plausibel

dieser Region. Mit der Veröffentlichung der vorläufigen Version der Sequenz des menschlichen Genoms (THE GENOME SEQUENCING CONSORTIUM, 2001) stehen nahezu komplette Daten nicht nur für die Xq21- Region zur Verfügung. Mit diesen Daten wurde mittels Matrix- CGH beim Patient 303/99 die größte Deletion gefunden, bei van der Maarel *et al.*, 1995 wird dagegen für Patient 525/00 die größte Deletion detektiert. Die Position des Gens TBX22 unterstützt aber die erste Version.

Die Tatsache, daß trotz Deletion der entsprechenden Gene Patient 303/99 keine Hördefekte und Patient 525/00 keine Choroideremie aufweist, war schon bekannt und kann mit unvollständiger Penetranz der beteiligten Gene erklärt werden.

5.5 Ausblick

Die in dieser Arbeit vorgestellte Variante der Matrix- CGH beruht auf der Auswertung einer repräsentativen Auswahl von genomischen Sequenzen die gegenüber der Ausgangs- DNS in ihrer Komplexität verringert ist. Für diesen Schritt fand die Inter-Alu- PCR Anwendung. Die Proben wurden durch PCR- Amplifikation mit dem CL-2-Primer generiert. Die korrespondierenden Sonden wurden in *In Silico* aus bekannten Datenbanksequenzen bestimmt und durch PCR mit spezifischen, internen Primerpaaren erzeugt. Eine ähnliche Variante der Matrix- CGH wurde von Geschwind *et al.*, 1998 beschrieben. Sie unterscheidet sich von der in dieser Arbeit vorgestellten Methode nur durch die Beschaffenheit der verwendeten Sonden, die aus bereits charakterisierten genomischen Klonen durch IA- PCR gewonnen wurden.

Durch den Datenbank- gestützten Ansatz ergeben sich einige Vorteile gegenüber dem konventionellen Ansatz:

- es werden keine charakterisierten genomischen Klone zur Erzeugung der Sonden benötigt
- Schritte zur Klonierung der Sonden entfallen
- die aus den Datenbanksequenzen generierten IA- PCR- Klone können eindeutig lokalisiert werden, die Gesamtheit an bereits vorhandenen Positionsdaten von Genen, genomischen Markern etc. steht somit ohne Umwege direkt zur Verfügung
- bei der Auswahl der internen Primerpaare können repetitive Sequenzen in den Sonden vermieden werden

- PCR- Produkte sind die technisch am einfachsten zugänglichen genomischen Klone

Nachteilig ist die Generierung der Sonden durch je zwei spezifische Primer. Dadurch wird für ein Matrix- CGH- Experiment eine große Anzahl an Oligonukleotiden benötigt, ein Aufwand der durch eine große Anzahl an herzustellenden Mikrorastern gerechtfertigt werden kann.

Ein methodisch ähnlicher Ansatz wurde von Lucito *et al.*, 2000 beschrieben. Dabei wurde die Repräsentation mit verringerter Komplexität durch Schneiden der DNS durch Restriktionsendonukleasen und anschließende Klonierung der Fragmente erreicht, die Limitierung des Größenspektrums war ebenfalls per PCR möglich. Die Ergebnisse zeigten, daß die Signifikanz der Meßwerte durch eine verringerte Komplexität erhöht wird.

Ein anderer Ansatz, bei dem die Komplexität der Probe nicht reduziert wird, sondern große genomische Klone als Sonden und nicht amplifizierte DNS als Probe Verwendung finden, wurde in den Publikationen von Solinas-Toldo *et al.*, 1997, Pinkel *et al.*, 1998, Albertson *et al.*, 2000 und Bruder *et al.*, 2001 beschrieben. Gegenüber dem in dieser Arbeit dargestellten Ansatz bleiben alle zuvor genannten Vorteile bestehen, zusätzlich kann angeführt werden, daß bei den oben genannten Methoden die zu untersuchende DNS nicht amplifiziert wird und somit mehr Ausgangsmaterial benötigt wird. Die Auflösung ist durch die Größe der als Sonden verwendeten genomischen Klone begrenzt. Der Wegfall der PCR und die größeren genomischen Klone ermöglichen dagegen Meßwerte, die weniger Schwankungen aufweisen als die der zuvor beschriebenen PCR- basierten Ansätze.

Weitere Versuche wurden unternommen, die cDNS- Microarrays aus der Expressionsanalyse auch für die Matrix- CGH zu verwenden. Für diesen Ansatz kann die zu untersuchende DNS als Probe nicht amplifiziert werden. Es wird erwartet, daß bei dieser Hybridisierung von gesamtgenomischer DNS auf PCR- Produkte die Signale zu schwach für die Detektion sind. Pollack *et al.*, 1999 beschreiben, wie sie mit der auch in dieser Arbeit verwendeten Markierungsmethode trotzdem auswertbare Signale erhalten haben. Obwohl dieses Ergebnis ein grundlegendes Problem der Matrix- CGH löst, wurde diese Methode nach dem derzeitigen Stand der Veröffentlichungen noch nicht wiederholt. Die Resultate stehen im Widerspruch zu den Arbeiten von Lucito *et al.*,

2000 und Heiskanen *et al.*, 2000. In der letzteren Arbeit wurde ebenfalls versucht, markierte genomische DNS auf cDNS- Mikroarrays zu hybridisieren. Um detektierbare Signale zu erhalten wurde ein Signalamplifikationssystem eingesetzt (Tyramidsystem – Bobrow *et al.*, 1991). Mit dieser Methode konnten allerdings nur Unterschiede von mehr als fünf Kopien erkannt werden.

Im Vergleich zu Karyotypisierung und CGH liefert die Matrix- CGH eine vielfach höhere Auflösung. Gegenüber der CGH reduziert sich der technische Aufwand und vereinfacht sich die Auswertung. Es ist zu erwarten, daß die Matrix- CGH die CGH in naher Zukunft als Standardmethode zur Untersuchung nichtbalancierter chromosomaler Aberrationen ablösen wird. Zur Zeit ist jedoch die Bereitstellung eines dichten Rasters charakterisierter Klone zur Erzeugung von Sonden das größte technische Hindernis. Neben gemeinsamen Anstrengungen, ein solches genomweites Raster (Cheung *et al.*, 2001) zu etablieren, bietet auch der in dieser Arbeit vorgestellte Ansatz der Generierung von Sonden durch Auswahl von spezifischen Primerpaaren direkt aus Datenbanksequenzen eine Möglichkeit, dieses Hindernis zu überwinden. Er zeigt auch auf, wie die Daten aus der Entschlüsselung des humanen Genoms genutzt werden können, um schneller zu genetisch relevanten Informationen zu gelangen (Auffinden eines Kandidatengens für die Lippen-, Kiefer- Gaumenspalte – TBX22). Für einen möglichen Einsatz in Klinik oder Forschung muß das Auflösungsvermögen für kleine Unterschiede der Kopienzahlen noch verbessert werden. Dazu müssen Werte für die beiden wichtigen Parameter – den Anstieg in der Ausgleichsgerade (siehe 4.1.4.3) und die Standardabweichung der gemessenen Signalintensitätsverhältnisse weiter verbessert werden. Dazu wäre es nötig, durch technische Maßnahmen das Signal-Hintergrundverhältnis zu erhöhen. Bei der Auswertung könnte eine Konzentration auf ungefähr 30% der Sonden, für die die ermittelten Meßwerte sehr gut mit den eingesetzten Kopienzahlverhältnissen korrelieren, eine Verbesserung der Detektion von kleinen Kopienzahlunterschieden ergeben. Auch eine individuelle Ausgleichsgerade für jede einzelne Sonde kann zu optimalen Ergebnissen führen. Dann kann die vorgestellte Methode nicht nur starke Amplifikationen, sondern auch den Verlust (LOH) oder den Erwerb eines Allels erkennen, eine Konstellation, die nicht nur für die Tumorforschung interessant ist.

Die dargestellten Ergebnisse der Detektion von vollständigen Deletionen lassen dagegen an einen direkten Einsatz in der Diagnostik denken. Insbesondere Krankheiten, deren bekanntes Mutationsspektrum durch große hemizygote Deletionen geprägt sind (Muskuläre Dystrophie, Typ Duchenne (DMD – OMIM 310200), Hämophilie A (OMIM 306700) und Hämophilie B (OMIM 306900), lassen den Einsatz der Methode sinnvoll erscheinen.