

**Entwicklung der Array-CGH zur hochauflösenden,
genomweiten Untersuchung von DNA-
Veränderungen einzelner Tumorzellen**

Dissertation

der Fakultät für Biologie der Ludwig-Maximilians-Universität München
zur Erlangung des Doktorgrades der Naturwissenschaften (Dr. rer. nat.)



vorgelegt von
Christine Fuhrmann
aus Oldenburg

Dissertation eingereicht am: 13. September 2007

Erster Berichterstatter: Prof. Dr. T. Cremer

Zweiter Berichterstatter: Prof. Dr. H. Leonhardt

Tag der mündlichen Prüfung: 15. Juli 2008

Eidesstattliche Erklärung

Ich erkläre hiermit an Eides statt, dass ich die vorliegende Arbeit ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Die aus anderen Quellen direkt oder indirekt übernommenen Daten und Konzepte sind unter Angabe der entsprechenden Literaturzitate gekennzeichnet.

Die vorliegende Dissertation wurde weder in gleicher noch in ähnlicher Form einer anderen Prüfungsbehörde vorgelegt.

München, den 13. September 2007

Christine Fuhrmann

Ob eine Sache gelingt, erfährt man nicht indem man darüber nachdenkt, sondern wenn man es ausprobiert.

1	<u>EINLEITUNG</u>	4
1.1	DAS MAMMAKARZINOM	4
1.1.1	Histologie der Mamma	4
1.1.2	Brust-Stammzellen und ihre mögliche Bedeutung beim Brustkrebs	5
1.1.3	Krankhafte Veränderungen der Mamma	6
1.1.4	Klassifikation und Prognose des Mammakarzinoms	7
1.1.5	Therapie des Mammakarzinoms	9
1.2	IDENTIFIZIERUNG MOLEKULARER ZIELSTRUKTUREN ZUR THERAPIE DES MAMMAKARZINOMS	10
1.2.1	Fluoreszenz <i>in situ</i> Hybridisierung (FISH)	11
1.2.2	Allelverlust-Analyse (LOH)	11
1.2.3	Komparative Genom-Hybridisierung (CGH)	11
1.2.4	Genomische Untersuchung des Mammakarzinoms	14
1.3	DIE MINIMAL RESIDUALE RESTERKRANKUNG (MRD)	15
1.4	MOLEKULARE CHARAKTERISIERUNG EINZELNER DISSEMINIERTER TUMORZELLEN	16
1.5	ZIELSETZUNG	17
2	<u>MATERIAL UND METHODEN</u>	18
2.1	METHODEN	18
2.1.1	BAC-Plasmid Präparation	18
2.1.2	Qualitätskontrolle und Konzentrationsbestimmung der BAC-DNA	20
2.1.3	BAC-Plasmid Amplifikation und Fällung	20
2.1.4	Herstellung des 3K BAC-Arrays	21
2.1.5	Isolation von Einzelzellen und Amplifikation von Einzelzell-DNA	22
2.1.6	Genspezifische PCR	23
2.1.7	Reamplifikations- und Markierungs-PCR amplifizierter DNA	23
2.1.8	Hybridisierung markierter DNA auf BAC-Arrays	25
2.1.9	Hybridisierung markierter DNA auf Operon Oligonukleotid-Arrays	27
2.1.10	Hybridisierung markierter DNA auf Agilent Oligonukleotid-Arrays	28
2.1.11	Auswertung der Hybridisierung in Microsoft Excel	29
2.1.12	Auswertung der Hybridisierung in CAPweb	29
2.1.13	Statistische Auswertung von Array-CGH-Experimenten	31
2.1.14	Hybridisierung auf Nylonmembranen	33
2.1.15	Nachweis von Nukleotid Polymorphismen (SNPs) in Mse I-Fragmenten	34
2.1.16	Quantitative real time PCR	35
2.2	VERWENDETE STANDARDPROTOKOLLE	36
2.2.1	Zellkultur	36
2.2.2	Agarosegel Elektrophorese	36
2.2.3	Konzentrationsbestimmung von Nukleinsäuren	36
2.2.4	Herstellung elektrokompenter Bakterien und Elektroporation	36
2.2.5	Plasmid-Herstellung	36
2.2.6	Herstellung genomischer DNA aus Blut und Zelllinien	36
2.3	VERZEICHNIS DER VERWENDETEN ABKÜRZUNGEN	37
2.4	MATERIAL, GERÄTE, SOFTWARE, PUFFER UND LÖSUNGEN	38
2.5	OLIGONUKLEOTIDE UND PRIMER	41

2.6	ZELLINIEN, BAKTERIEN UND PLASMIDE	42
2.7	PATIENTENKOLLEKTIV	42
2.8	KOOPERATIONSPARTNER	42
3	<u>ERGEBNISSE</u>	43
3.1	ETABLIERUNG DER ARRAY-CGH-METHODE	43
3.1.1	Herkömmliche BAC-Arrays sind unbrauchbar für Einzelzellanalysen	43
3.1.2	Einzelzellanalysen auf BAC-Arrays mit reduzierter Kontamination	44
3.1.3	Optimierung der Spotting-Parameter	47
3.1.4	Präparation des 3K-Sets	50
3.1.5	Etablierung des Hybridisierungsprotokolls für Einzelzell-Array-CGH	52
3.1.6	Charakterisierung des Hybridisierungsverhaltens des 3K BAC-Arrays	58
3.1.7	Validierung der Hybridisierungsbedingungen durch Einzelzell-Analysen	64
3.1.8	Test von Oligonukleotid-Arrays für Array-CGH mit Einzelzellen	71
3.2	HYBRIDISIERUNG VON AMPLIFIKATEN DISSEMINIERTER TUMOREINZELZELLEN	78
3.2.1	Validierung von Alterationen in Einzelzell-Amplifikaten	78
3.2.2	Hochauflösende genomische Analyse von einzelnen disseminierten Tumorzellen aus dem Knochenmark von Mammakarzinom-Patientinnen	83
4	<u>DISKUSSION</u>	89
4.1	ENTWICKLUNG DES 3K BAC-ARRAYS ZUR EINZELZELL-DNA-ANALYSE	89
4.2	CHARAKTERISIERUNG DES 3K BAC-ARRAYS	92
4.3	DER 3K BAC-ARRAY IM VERGLEICH ZU OLIGONUKLEOTID-ARRAYS	97
4.4	ARRAY-CGH UNTERSUCHUNGEN VON DISSEMINIERTEN TUMORZELLEN	99
5	<u>ZUSAMMENFASSUNG</u>	106
6	<u>LITERATURVERZEICHNIS</u>	107
7	<u>DANKSAGUNG</u>	113
8	<u>LEBENS LAUF</u>	114
9	<u>VERÖFFENTLICHUNGEN</u>	115
10	<u>ANHANG</u>	116

1 Einleitung

1.1 Das Mammakarzinom

Die Brustkrebsinzidenz steigt seit 1970 in Deutschland stetig an, jährlich erkranken derzeit etwa 55.100 Frauen an Brustkrebs. Damit stellt diese Erkrankung 26,8%, bei Frauen unter 60 sogar 40% aller Krebsneuerkrankungen dar und ist somit häufigste Krebserkrankung bei Frauen¹. Das mittlere Erkrankungsalter liegt bei etwas über 62 Jahren. Ab einem Alter von 35 Jahren steigt die Inzidenz deutlich an, ab dem 55. Lebensjahr bleibt die Inzidenz gleich². Das mittlere Fünfjahresüberleben liegt derzeit bei 78%, also deutlich höher als beispielsweise beim Magen- oder Lungenkarzinom, konnte aber im Vergleich zu 73% Fünfjahresüberleben in den achtziger Jahren kaum gesteigert werden.

Seit 1997 ist die Brustkrebsmortalität leicht rückläufig, wobei allerdings unklar ist, welcher Anteil an dieser Entwicklung durch Früherkennung oder durch Therapie bedingt ist. Vieles spricht dafür, dass eine verbesserte Früherkennung das Überleben günstig beeinflusst, da kleine Tumore mit einer günstigen Prognose korrelieren³. Betrachtet man jedoch die erheblichen wissenschaftlichen und klinischen Bemühungen, so haben Patientinnen bislang vergleichsweise wenig von den Ergebnissen der Forschung profitiert.

1.1.1 Histologie der Mamma

Bereits während der Embryogenese werden Brust und Brustdrüse angelegt. Bei der Frau erfährt die Brust in der Pubertät eine weitere Entwicklung. Die Mamma ist im Wesentlichen aus Binde- und Fettgewebe sowie den epithelialen Drüsenstrukturen aufgebaut (siehe Abbildung 1). Eine Brust besteht aus bis zu 20 einzelnen Drüsenlappen, die sich zu einem ausgedehnten Gangsystem verzweigen. Die Milchgänge münden kurz vor der Brustwarze in den Milchsinus, der Ausführungsgang führt durch die Brustwarze und mündet an ihrer Spitze aus. Kleinere Milchgänge werden auch interlobuläre Gänge genannt, sie sind über die intralobulären Gänge mit den Endstücken der Drüsenläppchen verbunden, welche gemeinsam die terminale duktulobuläre Einheit (TDLE) bilden. Die Milchgänge sind aus einem zweischichtigen Epithel aufgebaut, dessen apikale Schicht aus luminalen Zellen und dessen basale Schicht aus myoepithelialen Zellen besteht. Die myoepithelialen Zellen stehen im Kontakt zur Basalmembran. In der nicht-laktierenden Mamma enden die Milchgänge in wenig verzweigten, tubulusähnlichen Endstücken, so genannten Tubuloalveoli, die aus Epithelzellen und myoepithelialen Basalzellen bestehen. Die Brust ist ein dynamisches

Organ, so kommt es während der Schwangerschaft zu einer starken Proliferation des Drüsengewebes. Außerdem wandelt sie sich in eine Milch produzierende Drüse um, die Laktation beginnt dann mit der Geburt des Kindes. Die engen Tubuloalveoli vergrößern sich zu weitlumigen Alveolen und die Milchdrüsenepithelzellen produzieren und sezernieren die Milch. Die myoepithelialen Zellen sind der glatten Muskulatur ähnlich und helfen, die Milch aus den Alveolen und Gängen zu pressen. Nach Abschluss der Laktation kommt es durch apoptotische Vorgänge zur Rückbildung (Involution) der Drüsenepithelien und der Basallamina. Nach der Menopause kommt es schließlich zur senilen Involution, bei der sowohl die Alveolen als auch das Gangsystem zurückgebildet werden⁴.

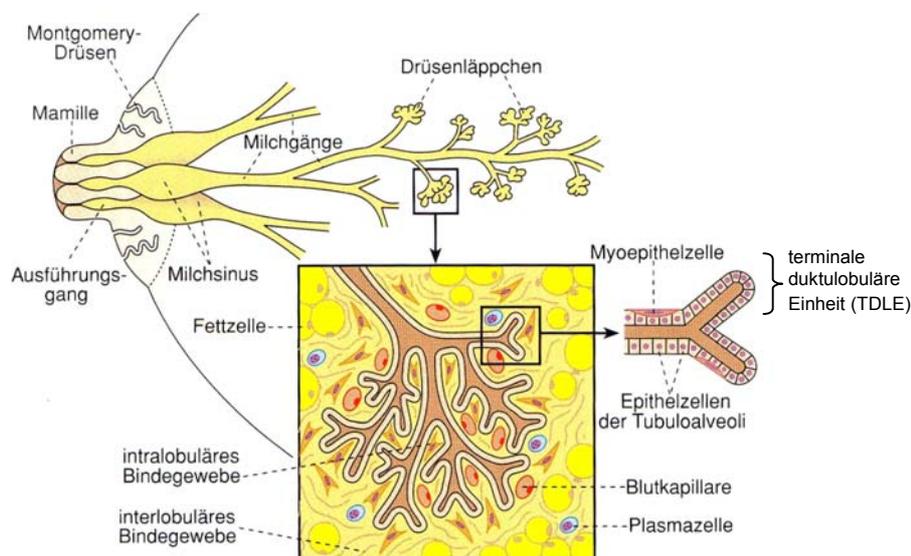


Abbildung 1 Aufbau der Brustdrüse

Gezeigt sind der allgemeine Aufbau der nicht-laktierenden Brustdrüse sowie eine Ausschnittsvergrößerung eines Drüsenläppchens und der detaillierte Aufbau einer terminalen duktulobulären Einheit (TDLE). (modifiziert nach⁴)

1.1.2 Brust-Stammzellen und ihre mögliche Bedeutung beim Brustkrebs

Während der Schwangerschaft kommt es zu einer starken Proliferation von Epithelzellen in der Brustdrüse und nach Abschluss der Laktation zum Abbau von Brustgewebe durch Apoptose (Involution). Eine Population von Zellen innerhalb der Mamma überlebt die Involution, weist ein starkes Proliferationspotential sowie die Fähigkeit zur Differenzierung auf und ist somit in der Lage, die Brustdrüse bei nachfolgenden Schwangerschaften erneut zu regenerieren. Die Vermutung, dass diese Population von Zellen Stammzellen beinhalten könnte, wurde dadurch untermauert, dass es in den vergangenen Jahren gelang, aus

Brustdrüsen von Mäusen^{5, 6} und Menschen⁷ Zellen zu isolieren, die Vorläuferzell- bzw. Stammzeleigenschaften besitzen und die in der Lage sind, sich selbst zu regenerieren und in die unterschiedlichen Zelltypen der Brustdrüse zu differenzieren. Außerdem können Zellen mit diesen Eigenschaften drüsige Strukturen ausbilden bzw. eine völlig neue Brustdrüse rekonstruieren⁵⁻⁷. Experimentelle und klinische Arbeiten belegen, dass Tumore aus mutierten Stammzellen entstehen können^{8, 9}. Die Beobachtung, dass Stammzellen und Tumorzellen eine Reihe von wichtigen phänotypischen Merkmalen teilen, unterstützt weiterhin die Hypothese, dass Stammzellen eine Rolle in der Tumorigenese besitzen. Dazu zählen unter anderem erhöhtes Proliferationspotential, Selbsterneuerungskapazität, die Möglichkeit in verschiedene Zelltypen zu differenzieren, die Aktivierung von Telomerase und anti-apoptotischen Mechanismen, erhöhte Membran-Transporter-Aktivität sowie die Fähigkeit, ohne Anhaftung überleben und migrieren zu können¹⁰. Die letztgenannten Fähigkeiten würden diese Zellen besonders zur Streuung befähigen. Diese Charakteristika könnten erklären, warum es mit konventionellen Chemotherapien zwar gelingt, den Primärtumor schrumpfen zu lassen, es aber trotzdem oft zu einem Rezidiv der Erkrankung kommt. Bislang entwickelte Therapieansätze, die auf Daten von Primärtumoren basieren und deren Erfolg am Schrumpfen des Tumors bemessen werden, zielen dementsprechend nicht auf die für die Progression der Krebserkrankung wichtigen Zellen⁸, die metastatischen Vorläuferzellen.

1.1.3 Krankhafte Veränderungen der Mamma

In der Mamma treten eine Vielzahl von gutartigen Veränderungen auf. Dazu zählt die fibrozystische Mastopathie, bei der entweder Zysten in den TDLEs entstehen, wobei es häufig zu einer Verkalkung der Strukturen kommt, oder Fibrosen des Stromas auftreten. Des Weiteren treten insbesondere bei jungen Patientinnen benigne Tumore auf, wie beispielsweise Papillom, Fibroadenom oder tubuläres Adenom.

Invasive Mammakarzinome gehen fast ausschließlich aus luminalen Zellen hervor, die sich über präinvasive Stadien, die Hyperplasie des Epithels und das Carcinoma *in situ* (CIS), entwickeln. Das Carcinoma *in situ* ist charakterisiert durch eine neoplastische Proliferation von malignen epithelialen Zellen innerhalb der Drüsengänge. Man unterscheidet dabei zwischen duktalem CIS (DCIS) (Abbildung 2b) und lobulärem CIS (LCIS). Das LCIS ist charakterisiert durch einen Verlust des Zellzusammenhangs und einer lobulären Differenzierung, wobei die Basalmembran intakt bleibt; eine Vorstufe des LCIS ist die atypische lobuläre Hyperplasie. Das DCIS zeigt eine duktale Differenzierung von

kohärierenden Tumorzellen und entsteht vorwiegend im Bereich der TDLEs aus dem Drüsenepithel. Das Anfangsstadium wird als atypische duktale Hyperplasie (ADH) bezeichnet, und auch hier ist die Basalmembran nicht durchbrochen (Abbildung 2a). Sowohl LCIS als auch DCIS erhöhen das Risiko einer Patientin etwa um das zehnfache, ein Mammakarzinom zu entwickeln.

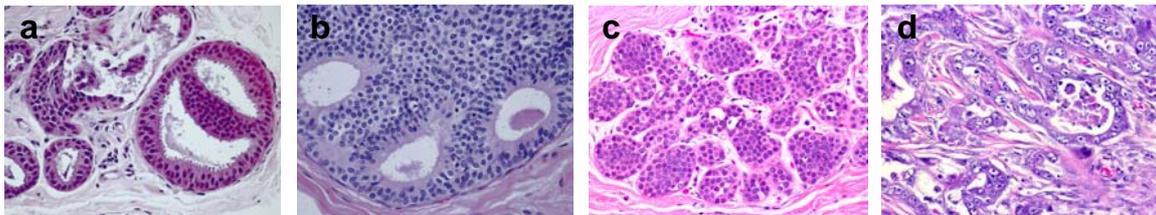


Abbildung 2 ADH und DCIS, duktales und lobuläres Mammakarzinom

(a) Atypische duktale Hyperplasie (ADH), (b) duktales Carcinoma *in situ* (DCIS), (c) lobuläres und (d) duktales Mammakarzinom. Die Bilder zeigen H/E-Färbungen von fixiertem Gewebe. ((a) und (b) nach Böcker et al.¹¹, (c) und (d) nach www.hopkinsbreastcenter.org/library/educational_information/mal.shtml)

Ein invasives Karzinom ist dadurch charakterisiert, dass die Basalmembran durchbrochen ist, der Tumor wächst infiltrativ. Wiederum unterscheidet man zwischen verschiedenen Typen, dem invasiven lobulären Karzinom und dem weitaus häufiger vorkommenden invasiven duktales Karzinom (Abbildung 2c und d), welches diverse Sonderformen wie das medulläre, tubuläre, papilläre, muzinöse und inflammatorische Mammakarzinom umfasst. Das lobuläre Karzinom macht etwa 10% aller Mammakarzinome aus und zeichnet sich durch dissoziiertes Wachstum aus. Dabei infiltrieren die Tumorzellen das umliegende Gewebe typischerweise im Gänsemarschmuster, also in Ketten hintereinander. Das duktales Karzinom weist histologisch eine unregelmäßige Struktur auf und ist durch unterschiedliche Differenzierung gekennzeichnet.

1.1.4 Klassifikation und Prognose des Mammakarzinoms

Um therapeutische und prognostische Aussagen in der Klinik zu erleichtern, wird das Mammakarzinom aufgrund verschiedener Kriterien in Gruppen mit unterschiedlicher klinischer Relevanz unterteilt. Das TNM-System¹² der UICC (Union Internationale Contre le Cancer) wird hierbei verwendet, um die Größe des Primärtumors (T), die Anzahl der befallenen Lymphknoten (N) und das Vorliegen von Fernmetastasen (M) zu klassifizieren. Das endgültige TNM-Stadium wird postoperativ erhoben, diese Daten werden mit dem Präfix „p“ gekennzeichnet (Tabelle 1). In Europa wird zusätzlich eine rein histopathologische Beurteilung durchgeführt¹³, bei der Zellkernmorphologie, Mitosen und Differenzierung der

Tumorzellen beurteilt werden. Bei diesem „Grading“ (G1-G3) entspricht ein Tumor der Stufe G1 einem gut differenzierten Tumor mit unauffälligen Zellkernen und einem relativen Zehnjahresüberleben von 87,9%, ein G2-Tumor hingegen ist mäßig differenziert und weist eine hohe Mitoserate und ein Zehnjahresüberleben von 69,7% auf. Ein G3-Tumor ist schlecht differenziert und weist mit 49,6% das schlechteste Zehnjahresüberleben auf¹⁴. Weiterhin sind biologische Kriterien wie Steroidhormonrezeptor-Expressionsmuster oder der Anteil proliferierender Zellen innerhalb des Tumors wichtig, um z.B. ein Ansprechen auf eine Therapie abschätzen zu können¹⁵.

Tabelle 1 TNM-Klassifikation des Mammakarzinoms

T: Ausdehnung des Primärtumors	
T ₀	Kein nachweisbarer Primärtumor
T _x	Primärtumor nicht bestimmbar
T _{is}	Carcinoma in situ; keine nachweisbare Tumormasse
T ₁	Tumor kleiner als 2 cm
T ₂	Tumor zwischen 2-5 cm
T ₃	Tumor größer als 5 cm
T ₄	Tumor mit direkter Ausdehnung auf die Brustwand oder Haut
N: Regionäre Lymphknotenmetastasen	
N ₀	Keine regionären Lymphknotenmetastasen
N _x	Lymphknotenstatus nicht definierbar
N _{mi}	Mikrometastasen (maximal 2 mm)
N ₁	Metastasen in 1-3 beweglichen axillären Lymphknoten derselben Seite
N ₂	an Strukturen fixierte axilläre Lymphknotenmetastasen; 4-9 befallene Lymphknoten derselben Seite
N ₃	Metastasen entlang der Arteria mammaria interna; 10 und mehr befallene Lymphknoten derselben Seite;
M: Fernmetastasen	
M ₀	Keine Fernmetastasen
M _x	Vorhandensein von Fernmetastasen kann nicht bestimmt werden
M _i	Isolierte Tumorzellen in Knochenmark oder parenchymatösen Geweben
M _{mi}	Mikrometastasen (maximal 2 mm)
M ₁	Fernmetastasen vorhanden

Der axilläre Lymphknotenstatus ist beim Mammakarzinom ein wichtiger Prognosefaktor für Rezidiv und Überleben. Während bei negativem Lymphknotenstatus das relative Zehnjahresüberleben noch 86,4% beträgt, sinkt es bei mehreren befallenen Lymphknoten auf nur 27,3%¹⁴. Es besteht außerdem ein enger Zusammenhang zwischen dem Lymphknotenstatus und der Größe des Primärtumors, die ihrerseits mit dem Zehnjahresüberleben der Patientinnen korreliert (74% Zehnjahresüberleben bei pT₁, 25% bei pT₄)¹⁶. Das Zehnjahresüberleben sinkt beim Vorliegen von Fernmetastasen, die häufigsten

Orte der Metastasierung sind Knochen (70%), Lunge (60%) und Leber (50%).

1.1.5 Therapie des Mammakarzinoms

Prinzipiell wird bei der chirurgischen Entfernung des Tumors zwischen der brusterhaltenden Therapie und der Mastektomie unterschieden. Bei kleineren Tumoren wird eine brusterhaltende Therapie durchgeführt, inzwischen werden 70% der Patientinnen auf diese Weise operiert. Während des Eingriffs werden gleichzeitig die axillären und/oder der drainierende Wächter-Lymphknoten zur Bestimmung des Nodalstatus entfernt. Zudem wird das Restparenchym der Brust postoperativ bestrahlt, um das Risiko eines Lokalrezidivs zu vermindern¹⁷. Bei lokal weit fortgeschrittenen oder inflammatorischen Mammakarzinomen wird vor dem chirurgischen Eingriff eine medikamentöse Tumorreduktion, eine so genannte neoadjuvante Therapie durchgeführt.

Im Anschluss an die Operation werden nahezu alle Patientinnen einer adjuvanten Therapie unterzogen. Zu den adjuvanten Therapien des Mammakarzinoms, die auf den Ergebnissen der Konsensus-Konferenz 2005 in St. Gallen beruhen, zählen die Hormon- und Chemotherapie. Zur Therapieentscheidung tragen unter anderem der TNM-Status, das Grading, der Menopausestatus und der Hormonrezeptorstatus des Tumors bei. Patientinnen mit Östrogen- oder Progesteronrezeptor-positiven Tumoren profitieren eher von einer endokrinen Therapie im Anschluss an eine Chemotherapie. Nach einer fünfjährigen endokrinen Therapie mit Tamoxifen wird bei Hormonrezeptor-positiven Tumoren das Risiko für Rückfälle um 14% signifikant reduziert¹⁸. Bei Hormonrezeptor-negativen Tumoren empfiehlt sich hingegen eine Polychemotherapie, bestehend aus einer Anthrazyklin-haltigen Dreifachkombination über drei bis sechs Monate, welche das Rückfallrisiko und die Sterblichkeit ebenfalls reduziert. Bei Frauen unter 50 Jahren bewirkt die adjuvante Chemotherapie eine absolute Verringerung der Sterblichkeit um ca. 15% nach fünf Jahren, bei Frauen zwischen 50 und 70 Jahren um ca. 6%. Eine Therapie mit dem monoklonalen, humanisierten HER2-Antikörper Trastuzumab hat sich bei Frauen mit einer nachweisbaren Überexpression von *HER2/neu*¹⁹, die mit einer schlechteren Prognose einhergeht²⁰, in diversen Studien bewährt. Allerdings ist *HER2/neu* nur in den Tumoren von 15-25% aller Mammakarzinompatientinnen amplifiziert, von denen wiederum nur maximal 50%, also insgesamt ca. 10% aller Brustkrebspatientinnen, auf eine Therapie mit Trastuzumab ansprechen²¹.

1.2 Identifizierung molekularer Zielstrukturen zur Therapie des Mammakarzinoms

Eine molekulare Charakterisierung der Tumore verspricht zum einen eine genauere Prognose, zum anderen eine maßgeschneiderte Therapie für die Patientinnen. Ein Beispiel für den Erfolg eines solchen Ansatzes ist die im vorherigen Abschnitt erwähnte Trastuzumab-Therapie bei *HER2/neu*-überexprimierenden Patientinnen. Ein weiteres Beispiel für die Ermittlung prognostischer Faktoren, kommt von Untersuchungen zum Mutationsstatus des Tumorsuppressorgens *p53*, wobei bestimmte Mutationen mit einem signifikant schlechteren Fünfjahresüberleben korrelierten²². Werden Moleküle wie *p53* als Zielstrukturen für Therapien verwendet, kann ihre Blockierung zu gravierenden Nebenwirkungen führen, da sie für den Organismus von entscheidender funktioneller Bedeutung sind. Im Gegensatz dazu weisen Proteine mit nicht-essentieller Funktion diese Limitierung nicht auf²³.

Um neue Ansatzmöglichkeiten für Therapien zu finden, die über die Hormon- und Chemotherapie hinausgehen, werden molekulare und genetische Analysen von Tumorgenomen durchgeführt, die zur Identifizierung von für die Tumorprogression wichtigen Genen oder relevanten Signalwegen führen sollen. Gleichzeitig ist ein Ziel solcher Untersuchungen, Muster von Veränderungen zu identifizieren, die Tumorsubtypen unterscheiden und Vorhersagen zum Therapieansprechen ermöglichen. Die Aktivierung von Onkogenen oder Inaktivierung von Tumorsuppressorgenen sind durch genetische Veränderungen bedingt, wobei es sich um chromosomale Gewinne oder Verluste, um Sequenz- oder Kopienanzahl-Polymorphismen oder Punktmutationen handeln kann. Single nucleotide Polymorphismen (SNPs) und Mikrosatelliten-Polymorphismen, die Individuen innerhalb der Bevölkerung voneinander unterscheiden, sind schon lange bekannt^{24, 25}. Vor kurzer Zeit wurden außerdem so genannte large copy number Polymorphismen (LCNPs) beschrieben²⁶, bei denen 100 kb bis 1 Mb große Bereiche der genomischen DNA in unterschiedlicher Kopienzahl vorliegen²⁷. Solche LCNPs wurden bislang mit neurologischen Erkrankungen in Verbindung gebracht, es ist allerdings nicht ausgeschlossen, dass diese Kopienzahl-Unterschiede auch zur Entstehung einer Tumorerkrankung beitragen können. Genomische Veränderungen können durch Methoden wie die komparative Genom-Hybridisierung (comparative genomic hybridization; CGH), die Fluoreszenz *in situ* Hybridisierung (FISH) oder Allelverlust-Analysen (loss of heterozygosity; LOH) untersucht werden.

1.2.1 Fluoreszenz *in situ* Hybridisierung (FISH)

Bei der Hybridisierung von DNA-Sonden auf humane Metaphasen-Chromosomen, der Fluoreszenz *in situ* Hybridisierungstechnik (FISH)²⁸, können gezielt sehr kleine Amplifikationen und Deletionen nachgewiesen werden. Durch eine entsprechende Wahl der Sonde können sogar Translokationen nachgewiesen werden. Obwohl die Mehrfarben-FISH²⁹,³⁰ eine simultane Untersuchung mehrerer Loci möglich macht, ist man bei dieser Technik immer auf Metaphasen-Präparationen angewiesen, was z.B. aus Tumorgewebe nicht immer möglich ist. Auch im Interphase-Nukleus können FISH-Untersuchungen durchgeführt werden³¹. Hierbei lässt sich zwar die chromosomale Lokalisation des Signals nicht nachweisen, aber anhand der Anzahl der Signale können Aussagen über Amplifikationen oder Deletionen bestimmter Loci getroffen werden. Der Vorteil dabei ist, dass keine aufwändigen Metaphasen-Präparationen hergestellt werden müssen.

1.2.2 Allelverlust-Analyse (LOH)

Neben der Untersuchung von Metaphasen besteht die Möglichkeit, genomische DNA aus Tumoren zu isolieren, um diese zu analysieren. Zur Darstellung von Deletionen in TumordNA kann der Nachweis des Verlustes von polymorphen Markern mittels spezifischer PCR verwendet werden (loss of heterozygosity, LOH). Hierbei können Mikrosatelliten-Polymorphismen oder Restriktionsfragment-Längenpolymorphismen (RFLPs)²⁴ untersucht werden. Letztere können ebenfalls für die Detektion von Punktmutationen dienen. Die Limitierung dieser Methoden besteht zum einen darin, dass bei der Untersuchung von Patientenmaterial auch immer Normalgewebe vorhanden sein muss, um den ursprünglichen Allel-Status eines Patienten feststellen zu können. Zum anderen können wie bei der FISH nur wenige Bereiche mit erheblichem Arbeitsaufwand untersucht werden.

1.2.3 Komparative Genom-Hybridisierung (CGH)

Die komparative Genom-Hybridisierung (CGH)^{32, 33} stellt ein Verfahren dar, mit dem Tumorgenome global auf Aberrationen getestet werden können. Bei der Metaphasen-CGH werden Tumor- und Referenz-DNA mit verschiedenen Fluoreszenzstoffen markiert und anschließend gemeinsam auf normale, humane Metaphasen hybridisiert. Die verschiedenen Farben werden getrennt voneinander fotografiert, und mit einer geeigneten Software werden die Signalintensitäten ermittelt. Die relative Hybridisierungsintensität des Test- und Referenz-Signals ist proportional zur relativen Kopienanzahl in den beiden Genomen

(schematische Darstellung in Abbildung 3). Numerische Aberrationen innerhalb des gesamten Genoms können so in einem Experiment detektiert werden. Mit der CGH können homozygote und heterozygote Verluste sowie Gewinne und Amplifikationen nachgewiesen werden. Man spricht von einem Gewinn, wenn ein genomischer Bereich verdoppelt wurde, und von einer Amplifikation, wenn die Kopienzahl für einen bestimmten Locus über der dreifachen Standardabweichung des Experiments liegt, in der Regel also sechs oder mehr Kopien vorliegen.

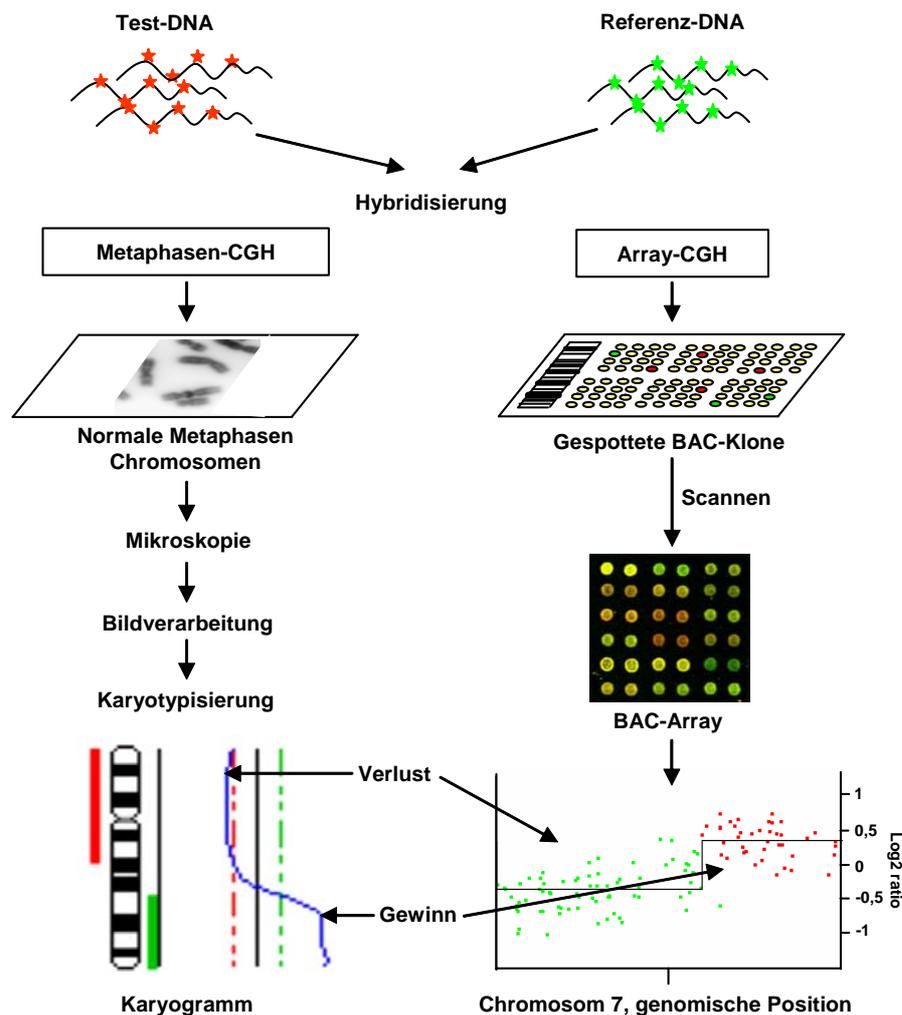


Abbildung 3 Schematische Übersicht der Metaphasen-CGH und Array-CGH

Test-DNA und Referenz-DNA werden mit unterschiedlichen fluoreszierenden Farbstoffen markiert und auf normale humane Metaphasen (Metaphasen-CGH, links) bzw. auf gespottete BAC-Klone (Array-CGH, rechts) hybridisiert. Die Fluoreszenzsignale der Test- und Referenz-DNA werden bei der Metaphasen-CGH mittels Mikroskopie erfasst, die Bildverarbeitungssoftware zeigt Gewinne und Verluste nach der Karyotypisierung an. Bei der Array-CGH werden die Signale mit einem Scanner gescannt, in der Auswertungssoftware wird das zur Basis 2 logarithmierte (Log_2) Verhältnis der Signalintensitäten für die einzelnen BAC-Klone berechnet, die BAC-Klone aufgrund ihrer genomischen Position geordnet, Gewinne und Verluste werden angezeigt. Es ist zu beachten, dass die Farbkodierung für Gewinne und Verluste bei Metaphasen- und Array-CGH gemeinhin entgegengesetzt ist.

Die optische Auflösung des Mikroskops und der stark kondensierte Zustand der Metaphasen limitieren die Auflösung auf ca. 10-30 Mb. Daher können Veränderungen von geringerer Größe mit der Metaphasen-CGH nicht entdeckt werden. Außerdem erschwert die nicht-lineare Kondensierung der Chromosomen, eine gegebene Aberration genau zu lokalisieren. Aus diesem Grund wurde die Entwicklung von Methoden mit immer höherer genomischer Auflösung vorangetrieben. 1997 wurde die so genannte Matrix-CGH³⁴ oder Array-CGH³⁵ entwickelt, bei der die humanen Metaphasen als herkömmliche Hybridisierungsplattform durch definierte, kurze Fragmente normaler menschlicher DNA ersetzt wurden (schematische Darstellung in Abbildung 3). Wichtige Voraussetzung für die Entwicklung dieser Methode war die Sequenzierung des humanen Genoms, bei der das gesamte Genom eines Spenders in 150-200 kb große Fragmente geschnitten und in bakterielle artifizielle Chromosomen – BAC-Vektoren – kloniert wurde. Die so entstandenen, charakterisierten DNA-Bibliotheken konnten für die Herstellung von so genannten BAC-Arrays verwendet werden. Zusätzlich konnte so die genaue Position eines DNA-Fragments innerhalb des Genoms festgestellt werden. Seitdem wurden BAC-Arrays mit 2000-3000 BAC-Klonen³⁵⁻³⁷ bis hin zu BAC-Arrays mit einer überlappenden Abdeckung des gesamten menschlichen Genoms produziert, die ca. 30.000 BAC-Klone umfassen³⁸. Inzwischen werden auch Arrays, die mit weniger komplexen Nukleinsäuren wie cDNAs³⁹, spezifischen PCR-Produkten⁴⁰ und Oligonukleotiden^{41, 42} hergestellt wurden, für Array-CGH-Analysen verwendet. BAC-Arrays werden seitdem immer häufiger durch Oligonukleotid-Arrays ersetzt, da die Herstellung von BAC-Arrays eine große technische Herausforderung darstellt. Die Hybridisierung auf repräsentierende Oligonukleotid-Arrays (ROMA)²⁷ ermöglicht die Untersuchung von genomischer DNA mit reduzierter genomischer Komplexität.

Die maximale technische Auflösung der Array-CGH entspricht momentan 140 kb, die überwiegende Mehrheit der konventionellen CGH-Arrays besitzt jedoch eine Auflösung von 1-2 Mb⁴³. Es ist außerdem möglich, ein Set von BAC-Klonen durch Fosmide, die wesentlich kürzer sind, zu ergänzen, um so Bruchpunkte genauer einzugrenzen und definieren zu können⁴⁴. Wird DNA aus Tumorgewebe hybridisiert, ist es möglich, Ein-Kopien-Gewinne bzw. -Verluste nachzuweisen, solange mindestens 60% der Zellen in der Probe Tumorzellen sind und diese die gleichen genomischen Veränderungen tragen. Beim Brustkrebs ist zu beachten, dass dieser besonders heterogen ist. Die einzelnen Zellen weisen zum einen nicht für alle chromosomalen Regionen die gleichen Kopienzahlen auf, zum anderen sind Brusttumore unterschiedlich stark mit gesunden Zellen durchsetzt, die das Stützgewebe und

das Gefäßsystem bilden. Genomische Veränderungen, die nur in Subpopulationen von bis zu 59% der Gesamtpopulation vorhanden sind, würden aus diesem Grund nicht entdeckt werden. Allgemein kann für die Untersuchung von genomischen Veränderungen in Tumorproben festgestellt werden, dass die Detektionsraten aufgrund folgender Parameter sinken: kleinere Aberrationen (in Basen), kleinere Aberrations-Amplituden und kleine Tumorsubklon-Repräsentierung.

1.2.4 Genomische Untersuchung des Mammakarzinoms

5-10% der Mammakarzinome werden durch erbliche Mutationen verursacht, im Vordergrund stehen dabei *BRCA1* (breast cancer associated protein 1)- und *BRCA2*-Mutationen⁴⁵. In den übrigen Fällen besteht zwar eine große Heterogenität bezüglich der Pathologie, des Krankheitsverlaufs und des Therapieansprechens, die zugrunde liegenden Mutationen konnten bislang aber noch nicht identifiziert werden. Wesentliche Ziele der genomischen Untersuchungen bestehen darin, 1.) Gene zu finden, die für die Progression der Erkrankung verantwortlich sind, 2.) neue Onkogene bzw. Tumorsuppressorgene zu identifizieren und 3.) Subtypen von Mammakarzinomen auf genomischer Ebene unterscheiden zu können.

Anhand der CGH-Profile können Tumore in *BRCA1*-Mutationsträger und -Nichtträger unterschieden werden⁴⁶, außerdem korrelieren bestimmte Aberrationsmuster mit unterschiedlichem Fünfjahresüberleben⁴⁷. Verschiedene Array-CGH-Studien konnten *HER2/neu* (human epidermal growth receptor 2) und *c-Myc* (myelocytomatosis cellular oncogene)⁴⁸ als bereits bekannte Onkogene verifizieren, neue Kandidaten-Onkogene wie *Rab25*⁴⁹, *FGFR1*⁵⁰ und *SPFH2*^{51, 52} konnten identifiziert werden.

Untersuchungen von gutartigen und prä-invasiven Brust-Läsionen wie der duktalem Hyperplasie (UDH: usual ductal hyperplasia), der atypischen duktalem Hyperplasie (ADH) und dem duktalem Carcinoma in situ (DCIS) sollen Aufschluss über die frühe Progression des Mammakarzinoms geben. Bereits in benignen Brust-Läsionen⁵³ und in normalen Brustdrüsen in Nachbarschaft zu Tumorgewebe kann Mikrosatelliten-Instabilität festgestellt werden⁵⁴. Auch in UDH-Proben können chromosomale Kopienanzahl-Variationen festgestellt werden, die denen gleichen, die in DCIS gefunden werden⁵⁵. LOH-Untersuchungen von ADH- und DCIS-Proben weisen darauf hin, dass ADH ein Vorläufer des DCIS ist, da eine starke Ähnlichkeit zwischen beiden festgestellt wurde⁵⁶. Es wurde lange Zeit angenommen, dass es sich beim DCIS um den Vorläufer des Karzinoms handelt. Allerdings konnten genetische Untersuchungen zeigen, dass DCIS eine sehr heterogene Erkrankung ist, wobei sich low-

grade DCIS und high-grade DCIS in der Anzahl und Art der Aberrationen deutlich unterscheiden. Auch beim invasiven Karzinom kann man gleichermaßen morphologisch wie genetisch low-grade und high-grade unterscheiden^{57, 58}. Des Weiteren ist der Mechanismus des 16q-Verlusts in low-grade und high-grade Karzinomen, wie durch LOH-Analysen gezeigt werden konnte, völlig verschieden⁵⁹. Das Paradigma der „schrittweisen“ Progression des Mammakarzinoms, welches analog zu dem durch Vogelstein für das Kolon-Karzinom vorgeschlagenen Modell ist, muss durch ein Progressions-Modell ersetzt werden, das multiple Evolutionsmöglichkeiten zulässt⁶⁰.

Größten Einfluss auf das Überleben der Patientin hat die Entwicklung manifester Metastasen. Es ist daher notwendig, die Evolution vom Primärtumor zur Metastase nachzuvollziehen, um zielgerichtete Therapien entwickeln zu können. Eine zwischen Primärtumor und asynchron entstandener Metastase vergleichende Studie konnte zeigen, dass in 31% der untersuchten Fälle Tumor und Metastase kaum übereinstimmende Aberrationen zeigten⁶¹. Zwar konnte in dieser Studie gezeigt werden, dass das X-Inaktivierungsmuster in Tumor und Metastase gleich war⁶¹, dennoch gibt es vermutlich keine lineare Progression vom Primärtumor zur Metastase, sondern eine parallele Evolution aus einer gemeinsamen Vorläuferzelle. Ähnlich wie im Primärtumor findet aber auch in der Metastase eine Selektion von Tumorzellen und somit eine Evolution statt, so dass die Untersuchung von klinisch manifesten Metastasen nicht die Frage beantworten kann, welche Aberrationen und Mutationen zur Disseminierung geführt haben. Um diese Antworten zu finden, müssen die Vorläuferzellen der Metastasen untersucht werden.

1.3 Die minimal residuale Resterkrankung (MRD)

Unter minimaler residueller Resterkrankung (MRD) versteht man den Nachweis von einzelnen gestreuten Tumorzellen in sekundären Organen, beispielsweise dem Knochenmark, nach der Entfernung des Primärtumors bei epithelialen Tumoren⁶². Durch immunzytochemische Färbung von epithelialen Antigenen wie dem Zytokeratin können diese Zellen im mesenchymalen Knochenmark nachgewiesen werden⁶³. Beim Mammakarzinom ist der Nachweis Zytokeratin-positiver (CK-positiver) Zellen unter der Bezeichnung $M_{(i)}$ Bestandteil der TNM-Klassifikation⁶⁴.

Bei Brustkrebspatientinnen können in 30-36% der Fälle CK-positive Zellen im Knochenmark nachgewiesen werden^{65, 66}, wohingegen in nur 1% der untersuchten gesunden Probanden solche Zellen detektiert wurden⁶⁵. Die Detektionsrate liegt zwischen einer und zehn Zellen

pro 10^6 Knochenmarkszellen. Die Frage nach der klinischen Relevanz dieser Zellen lässt sich mit einem signifikant schlechteren Gesamtüberleben von Patientinnen mit positiver Knochenmarkfärbung beantworten^{65, 66}, auch wenn nicht alle Patientinnen, bei denen CK-positive Zellen im Knochenmark detektiert wurden, tatsächlich Metastasen entwickeln. Auch in weiteren Tumorentitäten wie dem Kolonkarzinom und dem nicht-kleinzelligen Lungenkarzinom konnte eine signifikante Korrelation zwischen schlechterem Gesamtüberleben und dem Nachweis CK-positiver Zellen im Knochenmark der Patienten hergestellt werden^{67, 68}.

1.4 Molekulare Charakterisierung einzelner disseminierter Tumorzellen

Um den neoplastischen Ursprung CK-positiver Zellen zu belegen, wurden genetische Untersuchungen durchgeführt. Müller et al. konnten 1996 durch gleichzeitige CK-Färbung und FISH-Analyse zeigen, dass CK-positive Zellen im Knochenmark von Mammakarzinompatientinnen Amplifikationen des *HER-2/neu*-Lokus aufwiesen⁶⁹. Inzwischen wurden Methoden entwickelt, mit denen es möglich ist, sowohl das Transkriptom als auch das Genom einzelner Zellen global zu amplifizieren⁷⁰ und zu untersuchen⁷¹. Die Einzelzell-Metaphasen-CGH ermöglicht eine genomweite Detektion von Aberrationen⁷². In einer Studie wurden disseminierte Tumorzellen aus dem Knochenmark von Mammakarzinompatientinnen isoliert, ihre DNA amplifiziert und mittels CGH charakterisiert⁷³. Es konnte gezeigt werden, dass Tumorzellen von Patientinnen ohne Metastasen (M0-Zellen) völlig verschiedene und wesentlich weniger Aberrationen aufwiesen als Zellen von Patientinnen mit Metastasen (M1-Zellen)⁷⁴. Ein anschließender Vergleich mit den zugehörigen Primärtumoren zeigte, dass das Muster der Aberrationen verschieden von dem in disseminierten Zellen war, eine direkte Verwandtschaft zwischen Primärtumor und disseminierter Tumorzelle eines Patienten konnte in einer hierarchischen Ähnlichkeitsanalyse nicht nachgewiesen werden. Es liegt also nahe, dass die Disseminierung zu einem frühen Zeitpunkt in der Evolution des Primärtumors einsetzt, zu dem sich noch nicht die Aberrationen angehäuft haben, die zum Zeitpunkt der Operation detektiert werden können. Nach der Disseminierung scheinen Tumorzellen in ektopen Organen wie dem Knochenmark weitere Mutationen anzuhäufen und selektioniert zu werden. Die geläufige Hypothese, dass der Vorläufer der Metastase der am weitesten fortgeschrittene Klon des Primärtumors ist, wird hierdurch in Frage gestellt. Zudem kann aus der Korrelation von Tumorgröße und dem Zeitraum bis zur Entwicklung von Metastasen geschlossen werden, dass die Disseminierung

durchschnittlich bereits fünf Jahre vor der Diagnose stattfindet¹⁶.

Ein weiterer interessanter Befund der oben genannten Studie war, dass M0-Zellen zu mehr als 50% keine durch Metaphasen-CGH detektierbaren Aberrationen aufwiesen⁷⁴. Diese Zellen werden im Gegensatz zu Zellen, die genomische Veränderungen tragen, welche durch Metaphasen-CGH detektiert werden können (CGH-positiv), als CGH-negativ bezeichnet. Um den neoplastischen Ursprung dieser Zellen zu überprüfen, wurden umfassende LOH-Analysen für insgesamt 29 polymorphe Marker durchgeführt⁷⁵. CGH-negative M0-Zellen zeigten hierbei signifikant mehr LOH-Ereignisse als normale Kontrollzellen und fast ebenso viele LOH-Ereignisse wie CGH-positive M0- und M1-Zellen.

1.5 Zielsetzung

Um die genomweite, hochauflösende Untersuchung chromosomaler Veränderungen in einzelnen disseminierten Tumorzellen zu ermöglichen, war das Ziel der vorliegenden Arbeit die Entwicklung der Array-CGH für einzelne Zellen. Bei Beginn dieser Arbeit war lediglich bekannt, dass eine Reduzierung der bakteriellen Kontamination in BAC-Klon-Präparationen für die Herstellung eines BAC-Arrays notwendig war. Alle weiteren Schritte für die Präparation und die eingehende Validierung eines genomweiten BAC-Arrays für Einzelzellen sollten erarbeitet werden. Zunächst galt es geeignete Bedingungen für die Immobilisierung der BAC-DNA auf Objektträgern zu ermitteln. Im Anschluss an die Herstellung eines 3000 BAC-Klone umfassenden CGH-Arrays und die Etablierung eines Hybridisierungsprotokolls, sollte der Array charakterisiert werden. Dies sollte durch die Analyse von Einzelzellen mit bekannten Genomen sowie durch den Vergleich mit den aufkommenden Oligonukleotid-Arrays als Hybridisierungsplattformen erfolgen. Anschließend sollte eine erste Untersuchung von disseminierten Tumorzellen aus dem Knochenmark von Mammakarzinom-Patientinnen durchgeführt werden.

2 Material und Methoden

2.1 Methoden

2.1.1 BAC-Plasmid Präparation

Alle Gelelektrophoreseschritte, die für die BAC-Plasmid Präparation benötigt werden, finden in einem 0,5 x konzentrierten TAE-Puffersystem statt.

Es wurden Bakterienklone der FISHmappedClonesV3.1-Bibliothek verwendet. Innerhalb dieser Bibliothek wurden nur RP11-Klone verwendet, die 150-200 kb lange Fragmente humaner, genomischer DNA eines anonymen, männlichen Spenders enthalten, welche in den pBACe3.6-Vektor⁷⁶ kloniert worden waren⁷⁷ (Abbildung 4).

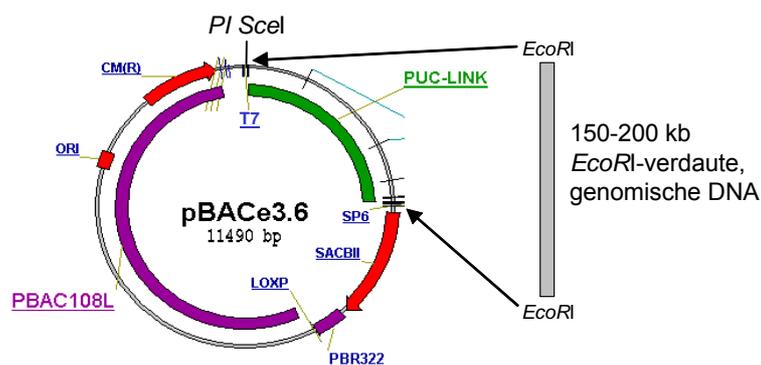


Abbildung 4 Schematische Darstellung eines RP11-BAC-Klons

Dargestellt ist der Aufbau des pBACe3.6-Vektors, in Violett ist das Rückgrat des Plasmids gekennzeichnet, mit dem origin of replication (ORI) und der Chloramphenicol-Resistenz - dem CM(R)-Gen (rot). Des Weiteren sind die Positionen der Sp6- und T7-Promotor/Primer-Bindungsstellen sowie die der *PI SclI*-Schnittstelle eingetragen. Die Positionen der *EcoRI*-Schnittstellen im PUC-Linker sind durch die schwarzen Pfeile angedeutet. 150-200 kb lange *EcoRI*-Fragmente von genomischer, humaner DNA wurden in den *EcoRI*-geschnittenen pBACe3.6-Vektor kloniert, um die RP11-BAC-Klone zu erhalten.

Für die Anzucht der Bakterien werden 2 ml LB-Chloramphenicol (LB-CA) mit -80°C-gelagerten Bakterienklonen angeimpft und 6-8 h bei 300 rpm und 37°C im Schüttel-Inkubator inkubiert. Mit dieser Vorkultur werden 50 ml LB-CA angeimpft und über Nacht (üN) unter identischen Bedingungen kultiviert. Nach 12-18 h erfolgt eine Zugabe von 10 µg/ml Tetrazyklin; nach 20 min ist die Zellteilung arretiert, aber die DNA-Synthese in allen Zellen abgeschlossen. Die Kulturen werden zunächst gekühlt und dann bei 4°C und 2600 g für 30 min abzentrifugiert. Die Bakterien werden mit 50 ml eisgekühltem PETT IV-Puffer gewaschen und erneut abzentrifugiert. Dann wird der Überstand verworfen und das Pellet im PETT IV-Rücklauf (ca. 400 µl) resuspendiert. Die Bakterien werden mit 2%-iger,

geschmolzener low-melt Agarose vorsichtig gemischt und in rechteckige Gießformen (Plugmolds) gegossen. Nach 30 min auf Eis sind die Bakterien-Agarose-Blöckchen geliert und werden in ein 50-ml-Gefäß überführt. Es werden 10 ml EC-Lyse-Puffer, 10 µl (10 U) RNase A und 5 mg Lysozym zugegeben, dann erfolgt ein Verdau der Zellwand und der RNA üN bei 37°C. Anschließend werden die Proteine durch eine Inkubation mit 300 U Proteinase K (PK) in 10 ml NDS-Puffer bei 50°C üN verdaut. Um die PK zu inaktivieren, werden die Blöckchen eine Woche in NDS bei Raumtemperatur (RT) gelagert.

Um Peptide und kleinere DNA-Fragmente aus den Agarose-Blöckchen zu entfernen, werden diese einer Reinigung durch Gelelektrophorese unterzogen. Hierfür findet im Blöckchen zunächst ein Pufferaustausch in TE-Puffer statt. Die Blöckchen werden dann in die Taschen eines 1%igen Agarose-Gels geladen, die Elektrophorese wird für 2 h bei 160 V durchgeführt. Danach werden die Blöckchen aus dem Gel isoliert. Die BAC-Plasmid-DNA wird anschließend mit 3 U PI-Sce I in einem 145 µl-Ansatz für 3 h laut Herstellerangaben linearisiert, nachdem der Enzym-Mix üN bei 4°C in das Blöckchen einwirken konnte. Für den folgenden PK-Verdau wird wie oben beschrieben verfahren.

Zur Trennung von bakterieller, chromosomaler DNA und den linearisierten BAC-Plasmiden wird eine pulsed-field-Gelelektrophorese (PFGE) angewendet. Zur Herstellung des Gels wird nach Herstellerangaben verfahren. Die Blöckchen werden in TE umgepuffert und anschließend in die Taschen des 1%igen low-melt Agarosegels geladen. Die Taschen werden mit low-melt Agarose versiegelt, dann folgt ein 20-stündiger Lauf bei 165 V. Das Puls-Schema verläuft hierbei wie folgt: Der erste Strom-Puls hat eine Länge von 3 sec und steigert sich innerhalb von 8 h extrapolierend bis auf 10 sec, in den verbleibenden 12 h steigert sich die Puls-Länge extrapolierend auf 80 sec. Die Spannung wird abwechselnd um 120° versetzt angelegt.

Zur Visualisierung der Banden wird das Gel im Anschluss für 30 min in ein Ethidium Bromid-Bad (20 µl Ethidium Bromid in 1 L 0,5x TAE-Puffer) gelegt. Die Banden werden auf dem UV-Schirm ausgeschnitten und die Gelstücke in 1,5-ml-Gefäße überführt. Die Agarose wird bei 70°C geschmolzen, auf 42°C temperiert und durch Zugabe von 1 U Agarase pro 100 mg Agarose bei 42°C üN verdaut. Die BAC-DNA wird aus der Lösung isoliert, indem diese zunächst mit 1/10 Volumenteil 3 M NaAc versetzt wird, 15 min bei 25000 g zentrifugiert wird, um unverdaute Oligo- und Polysaccharide abzutrennen, und dann mit einem Volumenteil Isopropanol versetzt, 30 min auf Eis inkubiert und anschließend 45 min bei 4°C und 25000 g zentrifugiert wird. Das DNA-Pellet wird mit 500 µl 70%igem

Ethanol gewaschen, getrocknet und in 10 µl TE üN bei 4°C gelöst. Die DNA wird in 96-well PCR-Platten überführt und bis zur weiteren Bearbeitung bei -20°C gelagert.

2.1.2 Qualitätskontrolle und Konzentrationsbestimmung der BAC-DNA

Um zu kontrollieren, ob die DNA vollständig rückgelöst ist, werden 2 µl einer 1:10-Verdünnung der rückgelösten BAC-DNA mittels Gelelektrophorese untersucht. Hierbei ist wichtig, dass die DNA nicht degradiert ist. Eine Degradation der DNA kann durch die Einwirkung starker Scherkräfte oder einen Verdau durch kontaminierende DNAsen bewirkt werden. Weiterhin wird die Konzentration der BAC-Plasmid DNA durch eine fluorimetrische Messung des PicoGreen-Einbaus in die Doppelstrang-DNA (dsDNA) bestimmt. Um alle 3000 Proben unter möglichst gleichen Bedingungen bearbeiten zu können, wurde dieser Prozess unter Verwendung des TECAN Freedom EVO-Pipettierroboters automatisiert. Es wurde ein Programm mit der Freedom EVOware-Software geschrieben, das die Abarbeitung der unten aufgeführten Schritte ermöglicht.

Hierbei werden 50 µl einer 1:100-Verdünnung der BAC-DNA mit 150 µl PicoGreen (1:1000 in TE verdünnt) in einer schwarzen 96-well-Photometer-Platte gemischt, anschließend wird die Fluoreszenz im GENios-Platten-Fluorimeter bei einer Anregungswellenlänge von 480 nm und Emissionswellenlänge von 520 nm gemessen. Aus diesen Werten wird mit Hilfe der Magellan-Software die DNA-Konzentration errechnet und gleichzeitig das Volumen bestimmt, das anschließend zugegeben wird, um eine Konzentration von 250 ng/µl in jedem Gefäß einzustellen. Die DNA wird bis zur weiteren Bearbeitung bei -20°C gelagert.

2.1.3 BAC-Plasmid Amplifikation und Fällung

Die aufgereinigte BAC-DNA wird im Anschluss an die Konzentrations-Einstellung mittels Adapter-Linker-PCR amplifiziert und anschließend durch Ethanol-Präzipitation aufgereinigt. Sowohl für die PCR als auch für die Fällung wurden automatische Abläufe programmiert. Mit Hilfe des Freedom EVO Pipettierautomaten werden jeweils 96 Proben parallel in einer 96-well-Platte bearbeitet, PCR-Platten werden mit hitzebeständiger, selbstklebender Folie verschlossen.

Jeweils 250 ng BAC-DNA in 8,8 µl TE-Puffer werden mit 0,2 µl (10 U) Mse I und 1 µl OnePhorAll-Puffer (OPA) bei 37°C für 3 h in der PCR-Maschine verdaut, dann wird das Enzym bei 65°C für 5 min inaktiviert. Während des Verdau wird der Ligationsansatz, bestehend aus 1 µl OPA-Puffer, jeweils 1 µl CAB21- und 1 µl ddCAB12-Oligonukleotid

(beide 100 μ M) und 3 μ l Wasser, in der PCR-Maschine einem Temperaturgradienten zur Basenpaarung unterworfen (65°C bis 15°C für je 1 min pro Temperaturschritt). Der Ansatz wird zusammen mit 5 U T4-DNA-Ligase und 0,01 mM ATP zum Produkt des Mse I-Verdau gegeben (Endvolumen 20 μ l) und über Nacht bei 15°C ligiert. Nach der Ligation wird der Ansatz zur Amplifikation mit 6 μ l Expand Long Template Puffer 1, 4 μ l dNTPs (10 mM), 70 μ l Wasser und 2 μ l Taq DNA-Polymerase aufgefüllt. Das PCR-Programm zur Amplifikation der BAC-DNA ist in Tabelle 2 aufgeführt.

Tabelle 2 PCR-Programm zur Amplifikation der BAC-DNA

1.	72°C	3:00 min
2.	95°C	1:00 min
3.	62°C	0:30 min
4.	72°C	3:00 min + 1 sec/Zyklus
5.	zurück zu 2. 34 Mal	
6.	4°C	∞

Das PCR-Produkt wird in eine 0,8-ml-96-well-deep-well-Platte umpipettiert und mit 1/10 Volumenteil 3 M NaAc und 1 μ l PAA vermischt. Die Fällung erfolgt durch Zugabe von 275 μ l 100%igem, unvergälltem Ethanol und Inkubation bei -20°C für 1 h. Die gefällte DNA wird 1h bei 4°C und 5000 g präzipitiert, der Überstand bis auf 30 μ l abgenommen, und es wird mit 500 μ l 70%igem unvergällten Ethanol gewaschen. Es erfolgt eine Zentrifugation für 30 min bei 5000 g und 4°C; der Waschschritt wird wiederholt und der Überstand vollständig abgenommen. Die DNA wird für 10 min bei 37°C im Ofen getrocknet und in 10 μ l Wasser über Nacht bei 4°C und 45 min bei 37°C im Schüttel-Inkubator rückgelöst.

Anschließend wird die DNA-Konzentration bestimmt, wobei ebenso verfahren wird, wie in 2.1.2 beschrieben. Es wird eine Konzentration von 800 ng/ μ l mit Wasser eingestellt, und die DNA wird bei -20°C gelagert.

2.1.4 Herstellung des 3K BAC-Arrays

Das Spotten der Test-Arrays erfolgte am DKFZ in Heidelberg mit Hilfe von Dr. Bernhard Radlwimmer. Die DNA wurde in verschiedenen Konzentrationen und in verschiedenen Spottingpuffern (siehe Ergebnisteil) mit einem Omnigrid-100-Spotter auf Corning GAPS II bzw. Quantifoil slides gespottet. Die DNA wurde anschließend durch UV-Quervernetzung und Backen bei 80°C für 30 min immobilisiert.

Das Spotten des 3K BAC-Arrays erfolgte bei der Firma SCIENION. Die DNA wurde auf eine Endkonzentration von 400 ng/ μ l in 2x sciSPOT-Puffer verdünnt und in Duplikaten auf

Corning GAPS II slides gespottet. Nach dem Trocknen wurden die Arrays einer Nachbehandlung mit *sciPROCESS* Puffer unterzogen. Die Qualität der Spots wurde durch optische Begutachtung nach einem Anfärben der gespotteten DNA mit Sybr Green bewertet. Die fertigen BAC-Arrays wurden luftdicht mit Trocknungspelen bei 4°C im Dunkeln bis zu ihrer Verwendung gelagert.

2.1.5 Isolation von Einzelzellen und Amplifikation von Einzelzell-DNA

Für die Erstellung dieser Arbeit werden Einzelzellen verschiedener Mammakarzinom-Zelllinien aus einer Suspension heraus isoliert. Hierbei werden die Zellen wie gewohnt trypsiniert, um sie zu lösen und zu vereinzeln und mit 1x PBS so verdünnt, dass unter dem Mikroskop einzelne Zellen mit Hilfe einer 10 µl-Pipettenspitze herausgesaugt werden können.

Tabelle 3 PCR-Mix (links) und PCR-Programm (rechts) für primäre Amplifikation

3 µl	Expand Long Template Puffer 1	1.	68°C	3:00 min
2 µl	10 mM dNTP-Mix	2.	94°C	0:40 min
1 µl	Expand Long Template PolMix	3.	57°C	0:30 min
35 µl	Wasser	4.	68°C	1:30 min + 1 sec/Zyklus
		5.	Zurück zu 2. 14 Mal	
		6.	94°C	0:40 min
		7.	57°C	0:30 min + 1°C/Zyklus
		8.	68°C	1:45 min + 1 sec/Zyklus
		9.	Zurück zu 6. 7 Mal	
		10.	94°C	0:40 min
		11.	65°C	0:30 min
		12.	68°C	1:53 min + 1 sec/Zyklus
		13.	Zurück zu 10. 21 Mal	
		14.	68°C	3:40 min
		15.	4°C	∞

Die isolierte Zelle wird zusammen mit 0,2 µl One Phor All Puffer (OPA), 0,13 µl 10% Igepal, 0,13 µl 10% Tween, 0,26 µl Proteinase K (1,6 U) und 1,28 µl Wasser für 15 h bei 42°C im Thermocycler inkubiert, wodurch das zelluläre Protein verdaut wird. Die Proteinase K wird anschließend für 10 min bei 80°C deaktiviert. Für die homogene Amplifikation des Genoms der Einzelzelle wird dieses zunächst mit dem Restriktionsenzym Mse I für 3 h bei 37°C verdaut. Zur verdauten Zelle werden hierfür 0,2 µl OPA, 0,2 µl Mse I (10 U) und 1,6 µl Wasser gegeben. Nach einem Inaktivierungsschritt für 5 min bei 65°C, wird ein Adapter an die entstandenen Überhänge ligiert, der aus den Oligonukleotiden LIB1 und ddMse11 besteht. Zuvor werden je 0,5 µl LIB1 und ddMse11 zusammen mit 0,5 µl OPA und 1,5 µl

Wasser auf 65°C erhitzt und mit einer Kühlungsrate von 1°C/min auf 15°C abgekühlt, wodurch eine Basenpaarung der beiden Oligonukleotide erreicht wird. Diese Mischung wird dann mit 1 µl 10 mM ATP und 1 µl T4 Ligase (5 U) versetzt, zur Mse I-verdauten DNA gegeben und über Nacht bei 15°C ligiert.

Für die abschließende Amplifikation wird der Ansatz mit dem PCR-Mix aufgefüllt und mit dem Primäramplifikationsprogramm im Thermocycler inkubiert (siehe Tabelle 3). Durch eine spezifische PCR (siehe 2.1.6) für zwei Mse I-Fragmente (CK19 und Ex2/3, Primer-Sequenzen siehe 2.5) wird kontrolliert, ob die Amplifikation erfolgreich war.

2.1.6 Genspezifische PCR

Ein 10-µl-Ansatz ist aus 1 µl PCR-Puffer + dNTPs, je 0,5 µl 8 µM 5'- und 3'-Primer, 0,25 µl BSA, 0,1 µl Taq-Polymerase (5 U) und 0,5 µl template zusammengesetzt. Die Amplifikation verläuft im Thermocycler, wie in Tabelle 4 aufgeführt.

Tabelle 4 PCR-Programm für spezifische PCR

1.	94°C	2:00 min
2.	T _m	0:30 min
3.	72°C	2:00 min
4.	94°C	0:15 min
5.	T _m	0:30 min
6.	72°C	0:20 min
7.	Zurück zu 4. 14 Mal	
8.	94°C	0:15 min
9.	T _m	0:30 min
10.	72°C	0:30 min
11.	Zurück zu 4. 24 Mal	
12.	72°C	2:00 min
13.	4°C	∞

Für die Qualitätskontrolle der primären Einzelzell-Amplifikate (siehe 3.1.7) werden diese 1:50 in Nuklease-freiem Wasser verdünnt. Von dieser Verdünnung werden 0,5 µl in die spezifische PCR eingesetzt.

2.1.7 Reamplifikations- und Markierungs-PCR amplifizierter DNA

Indirekte Markierung

Digoxigenin- und Biotin-gekoppelte Nukleotide werden während einer Reamplifikations-PCR in amplifizierte, genomische DNA eingebaut. Durch eine an die Hybridisierung

angeschlossene Detektion kann so die indirekt markierte DNA sichtbar gemacht werden. Referenz-DNA wird mit Biotin, Test-DNA mit Digoxigenin markiert. Die Zusammensetzung eines Markierungsansatzes und das PCR-Programm sind in Tabelle 5 aufgeführt. Der gleiche Ansatz (nur ohne gekoppelte Nukleotide und unter Verwendung einer 10-mM-dNTP-Mischung) und das gleiche PCR-Programm werden auch für die Reamplifikation von genomischer DNA ohne Markierung verwendet.

Tabelle 5 Ansatz für Dig/Bio-Markierung amplifizierter DNA (links) und PCR-Programm (rechts)

4 µl	Expand Long Template Puffer 1	1.	94°C	1:00 min
8 µl	Primer 10 µM (LIB1)	2.	60°C	0:30 min
1,5 µl	2/7 dNTP-Mix (siehe Tabelle 7)	3.	72°C	2:00 min
1,75 µl	Dig-/Bio-dUTP (1 mM)	4.	94°C	0:30 min
1,75 µl	Dig-dCTP (1 mM)	5.	60°C	0:30 min
4,375 µl	Bio-dCTP (0,4 mM)	6.	72°C	2:00 min + 20 sec/Zyklus
1 µl	Taq DNA-Polymerase (5 U/µl)	7.	zurück zu 4. 10 Mal	
21 µl	Wasser für Dig-dUTP	8.	4°C	∞
18,375 µl	Wasser für Bio-dUTP			
1 µl	Template			

Thermosequenz-Markierung

Fluoreszenz-gekoppelte Nukleotidderivate werden für eine direkte Markierung von amplifizierter, genomischer DNA in einer Reamplifikations-PCR verwendet, so dass diese nach einer Hybridisierung sofort detektiert werden kann. Referenz-DNA wird generell mit Cy3-gekoppelten, Test-DNA mit Cy5-gekoppelten Nukleotiden markiert. Die Zusammensetzung eines Ansatzes zur direkten Markierung mit Thermosequenz und das PCR-Programm sind in Tabelle 6 aufgeführt. Bei der Thermosequenz-Markierung wurden folgende Nukleotidderivate und dNTP-Mixe verwendet (für dNTP-Mixe siehe Tabelle 7):

- 9/10-Mix; je 1 µl Cy3/5-dCTP und -dUTP
- 15/20-Mix; 0,625 µl Cy3/5-dCTP und 0,475 µl Cy3/5-dUTP

Tabelle 6 Ansatz für Cy3/5 Thermosequenz-Markierung (links) und PCR-Programm (rechts)

2,5 µl	Thermosequenz Puffer	1.	94°C	1:00 min
2,5 µl	Primer 100 µM (LIB1)	2.	60°C	0:30 min
1 µl	dNTP-Mix (siehe Tabelle 7)	3.	0,2°C/sec bis 72°C	
X µl	Cy3/Cy5-dUTP (1 mM)	4.	72°C	5:00 min + 60 sec/Zyklus
X µl	Cy3/Cy5-dCTP (1 mM)	5.	zurück zu 1. 9 Mal	
0,5 µl	Thermosequenz (5 U/µl)	6.	4°C	∞
ad. 50 µl	Wasser			
0,8 µl	template			

Klenow-Markierung

Bei der Klenow-Markierung werden Fluoreszenz-markierte Nukleotidderivate in die DNA eingebaut. Zufalls-Oligonukletid-Decamere binden dabei an die denaturierte DNA, es folgt eine Strangverlängerung durch das Klenow-Fragment. Hierfür wurde der Invitrogen BioPrime-Labeling Kit verwendet. Hierbei wird die DNA-Konzentration zuvor photometrisch ermittelt. 250-1000 ng DNA in 21 µl Wasser und 20 µl 2,5x konzentrierte Random-Primer werden 5 min bei 95°C denaturiert und anschließend sofort für 5 min auf Eis gestellt. Es werden weiterhin 5 µl 10x dCTP-Nukleotid-Mix (siehe Tabelle 7), 2 µl Cy3/Cy5-dCTP (1 mM) und 1 µl Klenow-Fragment (5 U) zugegeben. Nach vorsichtigem Mischen wird der Ansatz für 5 sec zentrifugiert und 16 h bei 37°C in der PCR-Maschine inkubiert. Danach erfolgt die Zugabe von 5 µl Stop-Mix.

Tabelle 7 Zusammensetzung der dNTP-Mixe für die Markierungsreaktionen

2/7-Mix	dATP und dGTP 10 mM, dCTP und dTTP 8,4 mM
9/10-Mix	dATP und dGTP 10 mM, dCTP und dTTP 9 mM
15/20-Mix	dATP und dGTP 10 mM, dCTP 9,375 und dTTP 9,525 mM
10x dCTP-Nukleotid-Mix	dATP, dUTP und dGTP 1,2 mM, dCTP 0,6 mM
Reamplifikation ohne Markierung	dATP, dGTP, dCTP und dTTP alle 10 mM

2.1.8 Hybridisierung markierter DNA auf BAC-Arrays

Abtrennung nicht-eingebauter Nukleotide und Bestimmung der Einbaurrate

Jeweils ein Doppelansatz markierter Referenz- bzw. Test-DNA werden über eine Sephadex-G50-Säule von nicht-inkorporierten Nukleotiden gereinigt. Zur Herstellung der Säule wird eine 1-ml-Spritze mit silanisierter Glaswolle gestopft und bis zum Rand mit gequollenen Sephadex-G50-Perlen befüllt. Die Säule wird 2 min bei 300 g zentrifugiert, um überschüssigen Puffer zu entfernen, dann wird die Probe auf die Säule geladen und ebenso zentrifugiert. Anschließend wird die Einbaurrate im Eluat bestimmt, indem die optische Dichte (OD) einer 1:20 Verdünnung zunächst bei einer Wellenlänge von 260 nm gemessen wird, um die DNA-Konzentration zu ermitteln. Für Cy3 wird die OD₅₅₀, für Cy5 die OD₆₅₀ gemessen. Die Einbaurrate berechnet sich daraus wie folgt:

-Cy3-Einbau (Cy-Molekül pro bp): $OD_{260}/OD_{550} \times 23,15$

-Cy5-Einbau (Cy-Molekül pro bp): $OD_{260}/OD_{650} \times 38,58$

Eine Einbaurrate von einem Cy-Molekül alle 150-200 bp ist hierbei optimal. Bei einer

Einbaurate von weniger als einem Cy-Molekül pro 250 bp wurde die markierte Probe verworfen und ein neuer Markierungsversuch unternommen.

Vorbereitung der DNA-Probe für die Hybridisierung auf BAC-Arrays

Nach der Markierung mit dem jeweiligen Markierungsverfahren (s.o.) wurden je ein differentiell markierter Ansatz der Referenz- und der Test-DNA zusammen mit 80 µg Cot-1-DNA, 1/10 Volumenteil 3 M NaAc und zwei Volumenteilen 100%-igem Ethanol über Nacht bei -20°C gefällt. Ebenso werden für die Prähybridisierung sechs reamplifizierte Leerkontrollen zusammen mit 50 µg Heringsperma DNA gefällt. Die DNA wird durch Zentrifugation für 45 min bei 25000 g und 4°C präzipitiert, das Pellet wird mit 800 µl 70%-igem Ethanol gewaschen und luftgetrocknet. Die markierte DNA wird in 50 µl, die Prähybridisierungs-DNA in 100 µl Array-CGH-Mastermix gelöst. Zunächst wird 1 h bei 42°C im Schüttel-Heizblock inkubiert, dann 10 min bei 75°C denaturiert und zuletzt 2 h bei 42°C wieder im Schüttel-Heizblock renaturiert. Beim letzten Schritt kann die Cot-1-DNA mit repetitiven Sequenzen der Test- bzw. Referenz-DNA hybridisieren. Der Array-CGH-Mastermix ist hierbei wie folgt zusammengesetzt: 4% SDS, 50% Formamid, 8% Dextran Sulfat, pH 7,4 in 2x SSC.

Vorbehandlung und Hybridisierung des BAC-Arrays

Die BAC-Arrays werden 30 min bei 42°C im Wasserbad in Blocklösung (0,1 mg/ml BSA Fraktion V, 50% Formamid, 5x SSC, 0,1% SDS, pH 7,4) inkubiert, 1 min in Wasser und 1 min in Isopropanol gewaschen, bevor sie für 2 min bei 300 g trocken zentrifugiert werden. Die Hybridisierungskammern werden mit 500 µl Verdunstungspuffer befüllt (50% Formamid, 2x SSC, pH 7,4) und bei 42°C im Hybridisierungssofen vorgewärmt. Die Prähybridisierungs-DNA wird auf den Array aufgetragen und 1 h in der Hybridisierungskammer unter einem Deckglas bei 42°C im Schüttelinkubator bei ca. 5 rpm inkubiert.

Währenddessen wird der SlideBooster auf 42°C vorgeheizt, die Reservoirs der Hybridisierungskammern werden mit je 500 µl Verdunstungspuffer befüllt. Das Deckglas mit der Prähybridisierungslösung wird entfernt, die Probe mit einem Deckglas auf den Array aufgetragen, je 15 µl AdvaSon werden auf die Piezo-Elemente gegeben, dann wird der Array luftblasenfrei auf das AdvaSon gelegt und das Gerät verschlossen. Die Durchmischung der Probe erfolgt über die Vibrationen, die durch die Piezo-Elemente übertragen werden. Die

Hybridisierung erfolgt für 24 h bei 42°C in 10-sec-Zyklen mit einer Mischrate von 24 dBm für 3 sec gefolgt von 7 sec Pause.

Test-Arrays wurden im offenen Tropfen hybridisiert. Hierbei wurde nach der Vorbehandlung des Arrays die Hybridisierungsfläche mit einer dreifachen Schicht Fixogum umrahmt. In diese Begrenzung wurden die Prähybridisierungslösung und anschließend die Probe gegeben. Die Arrays wurden in den oben beschriebenen Hybridisierungskammern im Schüttelofen bei 42°C und ca. 5 rpm inkubiert. Die Probenvolumina waren die gleichen wie oben beschrieben. Nach Ende der Hybridisierung wird der Array aus der Hybridisierungskammer entnommen, gegebenenfalls wird der Fixogumring entfernt, und sofort in folgenden Puffern gewaschen: 10 min in 1x PBS + 0,05% Tween pH 7,4 bei Raumtemperatur, 30 min in 2x SSC, 50% Formamid pH 7,4 bei 42°C und 10 min in 1x PBS + 0,05% Tween pH 7,4 bei Raumtemperatur. Anschließend wird der Array durch Zentrifugation für 2 min bei 300 g getrocknet und sofort in einem Axon-Reader gescannt, wobei parallel eine Detektion des Cy3- und des Cy5-Signals erfolgt.

Mit Hilfe der GenePix-Software können die Scans ausgewertet werden. Dazu wird ein Raster über das Bild gelegt, das alle Spots mit den dazugehörigen Klonnenamen enthält. Das Raster wird mit den zugehörigen Spots in Deckung gebracht, fehlerhafte Spots und Artefakte werden entsprechend gekennzeichnet. Die Software erstellt anschließend eine Tabelle mit den Fluoreszenzintensitäten der Spots und des lokalen Hintergrunds sowie Angaben über ihre Spotqualität.

2.1.9 Hybridisierung markierter DNA auf Operon Oligonukleotid-Arrays

Es wurden humane Oligonukleotid-Arrays der Firma Operon (35K OpArray V4.0) verwendet. Hierbei sind 35.035 70 Basen-lange Oligonukleotide auf einem Corning GASP II slide immobilisiert, von denen 19592 Oligonukleotide für die CGH verwendet werden können. Die zu hybridisierenden Proben werden durch ein direktes Markierungsverfahren gelabelt. Nicht-amplifizierte genomische DNA wird durch Klenow-Labeling, amplifizierte DNA durch Thermosequenase-Labeling markiert. Wird nicht-amplifizierte DNA hybridisiert, wird ein doppelter Markierungsansatz verwendet, bei der Hybridisierung amplifizierter DNA wird, wie beim 3K BAC-Array, ein Markierungsansatz verwendet. Referenz und Probe werden wie in 2.1.8 beschrieben aufgereinigt und zusammen mit 50 µg Cot-1-DNA gefällt.

Die DNA wird wie in 2.1.8 beschrieben sedimentiert, gewaschen und getrocknet und anschließend in 30 µl Wasser aufgenommen. Die Probe wird 30 min bei 37°C rückgelöst und

dabei gelegentlich gemischt. Dann wird dasselbe Volumen 2x Probenpuffer (10x SSC, 50% Formamid, 0,4% SDS, pH 7,0) zugegeben. Es wird erneut bei 37°C für 1 h inkubiert, für 5 min bei 95°C denaturiert und die Probe anschließend bei 50°C für mindestens 1 h abgekühlt. Hierbei werden repetitive Sequenzen durch Anlagerung der Cot-1-DNA blockiert.

In der Zwischenzeit werden die Chips vorbehandelt, indem sie für 30 min bei 42°C im Oligo-Prähybridisierungspuffer (5x SSC, 1% BSA V, 0,1% SDS, pH 7,0) im Wasserbad inkubiert werden. Anschließend werden sie zweimal für eine Minute in Wasser und für eine Minute in Isopropanol gewaschen und bei 300 g für 2 min trocken zentrifugiert.

Der SlideBooster wird mit Oligo-Verdunstungspuffer (25% Formamid, 5x SSC) befüllt und auf 50°C vorgeheizt. Die Proben werden, wie in 2.1.8 beschrieben, auf die vorbehandelten slides aufgebracht. Die Hybridisierung erfolgt für 20 h bei 50°C in 10-sec-Zyklen mit einer Mischrate von 24 dBm für 5 sec gefolgt von 5 sec Pause.

Danach werden die Objektträger dem SlideBooster entnommen, das Deckglas mit der Probe wird entfernt. Es wird sofort mit folgenden Puffern gewaschen: 0.5x SSC /0,05%SDS bei 60°C für 2 min, den slide wenden und in einer frischen Küvette zwei weitere Minuten waschen, 0.5x SSC bei RT für 2 min, den slide wenden und in einer frischen Küvette zwei weitere Minuten waschen, 0.05x SSC bei RT für 2 min, den slide wenden und in einer frischen Küvette zwei weitere Minuten waschen, abschließend wird der slide in Wasser getaucht und für 3 min bei 250 g trocken zentrifugiert.

Die Detektion der Fluoreszenz-Signale und Auswertung mittels GenePix-Software erfolgt wie unter 2.1.8 beschrieben.

2.1.10 Hybridisierung markierter DNA auf Agilent Oligonukleotid-Arrays

Es wurde der 244K Oligonukleotid-Array der Firma Agilent verwendet. Alle 244.000 Oligonukleotide sind für eine Array-CGH-Anwendung optimiert und weisen eine Länge von 70 bp auf. Die amplifizierte DNA aus Einzelzellen und die nicht-amplifizierte DNA aus ca. 10^6 Zellen wurde wie in 2.1.7 beschrieben markiert. Nach der Markierung wurde die DNA über eine Microcon-YM 30-Säule aufgereinigt. Test- und Referenzprobe wurden in einem Gesamtvolumen von 30 µl Wasser aufgenommen und bei -20°C gelagert. Die Hybridisierung und Prozessierung der Arrays wurden, ebenso wie das Scannen und die Normalisierung der Rohdaten, von Dr. Andreas Polten, Agilent Technologies, durchgeführt, da hierfür spezielle Geräte und Programme benötigt werden. Die normalisierten Rohdaten wurden mit Hilfe der CGH Analytics Software visualisiert.

2.1.11 Auswertung der Hybridisierung in Microsoft Excel

Die von der GenePix-Software erstellte Datei wird in ein Excel-Format umgewandelt, dann erfolgt eine Normalisierung der Daten nach Quackenbush et al.⁷⁸, wobei die Summe aller balancierten Klone für Referenz- und Test-DNA gebildet wird. Teilt man die Summe der Referenz durch die Summe der Probe, erhält man einen Normalisierungsfaktor, mit dem man anschließend jeweils die Signalintensität der Referenz multipliziert; dieser Wert wird durch die Signalintensität der Probe geteilt. Aus diesem Quotienten bildet man für eine bessere graphische Darstellung den Log_2 . Als Signifikanzgrenze wird die dreifache Standardabweichung aller balancierten Klone definiert. Diese Art der Auswertung eignet sich für Arrays mit weniger als 200 Klonen, da eine Normalisierung, die räumliche Artefakte wie Hybridisierungsgradienten berücksichtigt, in diesen Fällen nicht möglich ist.

2.1.12 Auswertung der Hybridisierung in CAPweb

Zur Auswertung größerer Arrays kann die CAPweb-Software verwendet werden (CGH Array Analysis Platform on the WEB; <http://bioinfo-out.curie.fr/CAPweb/>). Eine Normalisierung der Daten erfolgt durch das MANOR-Paket (Micro-Array NORmalization of CGH-array data). Dabei werden lokale, räumliche Effekte wie z.B. Hybridisierungsgradienten durch einen dreistufigen Segmentierungs-Algorithmus beseitigt⁷⁹, der auf einem zweidimensionalen LOESS Verfahren beruht (Abbildung 5a). Zusätzlich werden diejenigen Spots aus der weiteren Auswertung ausgeschlossen, die eine schlechte Replikat-Übereinstimmung zeigen oder die ein zu niedriges Signal/Hintergrund-Verhältnis aufweisen (Abbildung 5b).

Um chromosomale Bruchpunkte zu ermitteln und Regionen zu identifizieren, die die gleiche DNA-Kopienzahl gemeinsam haben, wird anschließend das GLAD-Paket⁸⁰ (Gain and Loss Analysis of DNA) angewendet. Anhand der zur Basis 2 logarithmierten Verhältnisse (Log_2 -Verhältnis) der Proben- zu Referenzsignale werden Regionen definiert, deren Kopienzahl als entweder normal, amplifiziert oder deletiert gekennzeichnet wird. Die Parameter, die für die Ermittlung von Bruchpunkten, Gewinnen und Verlusten verwendet werden, entsprechen den Standardeinstellungen, die von den Entwicklern der Software bei der Untersuchung von insgesamt 399 BAC-Arrays⁷⁹ als optimal ermittelt wurden. In der Darstellung der Hybridisierungsprofile werden die Log_2 -Verhältnisse der einzelnen Datenpunkte dargestellt, zusätzlich zeigt eine so genannte Glättungslinie das durchschnittliche Log_2 -Verhältnis eines Abschnitts zwischen zwei Bruchpunkten an. Es muss dabei betont werden, dass keine Glättung der eigentlichen Daten durchgeführt wird. Außerdem werden in den einzelnen

Regionen „Ausreißer“ identifiziert und gekennzeichnet, die als nicht zuverlässig gelten. Für die Ausreißer-Detektion wird ein Wert von 0,001 für den Risiko-Faktor α festgelegt. In jeder Region werden die BAC-Klone, deren Log₂-Verhältnis im oberen oder unteren $\alpha/2$ -Bereich der Normalverteilung liegen, als Ausreißer definiert.



Abbildung 5 MANOR und VAMP Ausgabefenster der CAPweb-Software

(a) Der MANOR Algorithmus führt eine Normalisierung des Datensatzes durch. Im Ausgabefenster erscheint das Profil des Arrays vor und nach der Normalisierung, wobei weiß hinterlegte Quadrate Spots repräsentieren, die während der Normalisierung ausgeschlossen wurden. (b) Spot- und Klon-Informationen sowie Qualitätskriterien werden im selben Fenster dargestellt. (c) Die normalisierten Daten werden anschließend in VAMP visualisiert; hier dargestellt ist das Profil einer einzelnen Tumorzelle einer M0-Patientin. Schwarz eingefasste Klone sind „Ausreißer“, die schwarze, durchgezogene Linie ist die „Glättungslinie“. (Farbcodierung: grau = normal, rot = Gewinn, grün = Verlust, blau = Amplifikation)

In VAMP (Visualization and Analysis of CGH arrays, transcriptome and other Molecular Profiles; <http://bioinfo.curie.fr/vamp>) werden die Daten anschließend visualisiert. Die Daten

können mit dieser Software graphisch in einer klassischen Karyotyp-Ansicht oder in der Genom/Chromosom-Ansicht dargestellt werden, wie beispielhaft in Abbildung 5c gezeigt. Zusätzlich können Ähnlichkeitsanalysen mit verschiedenen Datensätzen durchgeführt werden, um Gemeinsamkeiten zwischen Profilen identifizieren zu können. Die Informationen zu Gewinnen und Verlusten werden durch den GLAD-Algorithmus erzeugt. Diese Daten werden nach der group-average-linkage Clustering-Methode mit Manhattan-Distanz gruppiert. Das Ergebnis dieser Gruppierung wird in VAMP graphisch dargestellt, wobei dem Cluster ein hierarchisches Dendrogramm zugeordnet wird. Hierbei wird die Ähnlichkeit zwischen Datensätzen in der Höhe des Bäumchens ausgedrückt.

2.1.13 Statistische Auswertung von Array-CGH-Experimenten

Student's *t*-Test

Mit dem Student's *t*-Test wurde untersucht, ob sich der GC-Gehalt und die Anzahl von Mse I-Fragmenten zwischen zwei Kollektiven von Sequenzen signifikant unterschieden. Das verwendete Programm wird im Internet unter: <http://faculty.vassar.edu/lowry/VassarStats.html> zur Verfügung gestellt.

Mann-Whitney-Test

Mit dem Mann-Whitney Rangsummen-Test wurden nicht-normalverteilte Datensätze auf signifikante Unterschiede überprüft. Zum einen wurde der Test angewendet, um zu überprüfen, ob sich eine Gruppe von einzelnen disseminierten Tumorzellen von einer Gruppe normaler Kontrollzellen hinsichtlich der Kopienanzahl für verschiedene durch qPCR untersuchte genomische Bereiche signifikant unterschieden. Der Test wurde außerdem angewendet, um den Anteil aberranter BAC-Klone in Hybridisierungen normaler Einzelzell-Amplifikate und Hybridisierungen von Amplifikaten aus M0- und M1-Zellen zu vergleichen. Die Berechnung wurde mit der „Mann-Whitney-Test“-Funktion des SPSS-Programms (V 13.0) durchgeführt.

Korrelationskoeffizienten

Korrelationskoeffizienten wurden berechnet, um den Grad des linearen Zusammenhangs (der Verwandtschaft) zwischen verschiedenen Array-CGH-Hybridisierungsprofilen zu ermitteln. Der Korrelationskoeffizient kann Werte zwischen 0 und 1 annehmen, bei einem Wert von 1

besteht ein vollständiger Zusammenhang zwischen den Proben, bei einem Wert von 0 besteht überhaupt kein Zusammenhang zwischen den Proben. Für die Berechnung wurde die „KORREL“-Funktion des Programms Excel verwendet.

Bestimmtheitsmaß

Das Bestimmtheitsmaß wurde berechnet, um zu ermitteln, in welchem Maß die Varianz für den prozentualen GC-Anteil in BAC-Klonen durch die Varianz für die Anzahl der Mse I-Fragmente bestimmt wird. Die Berechnung wurde mit dem Programm SigmaPlot durchgeführt. Das Bestimmtheitsmaß kann Werte zwischen 0 und 1 annehmen. Es gilt das gleiche, wie für den Korrelationskoeffizienten.

ROC-Kurven

In einer Receiver-Operating-Characteristic-Kurve (ROC-Kurve) werden die 1-Spezifität (falsch positive Ereignisse) und die Sensivität (richtig positive Ereignisse) gegeneinander aufgetragen, wobei diese einem Referenzwert zugrunde liegen. In dieser Arbeit wurden ROC-Kurven erstellt, um darzustellen, wie gut die Übereinstimmung zwischen einem Hybridisierungsprofil einer amplifizierten Einzelzell-DNA und einem Hybridisierungsprofil einer verwandten nicht-amplifizierten DNA (dem Referenzwert) ist. Die Werte der Referenz, nicht aber der Testprobe, werden in drei Kategorien eingeteilt: normal ($-0.2 < \log_2 \text{ratio} < 0.2$), Gewinn ($\log_2 \text{ratio} > 0.2$) und Verlust ($\log_2 \text{ratio} < -0.2$). Es wird unterschieden wie gut die Referenzprobe aus den Daten der Testprobe rekonstruiert werden kann. Dabei ist nicht entscheidend, dass der Wert der Testprobe sich in der selben Kategorie befindet, wie der der Referenz, vielmehr kommt es auf die Struktur der Daten an. Durch diese Darstellung können verschiedene Hybridisierungsplattformen miteinander verglichen werden. Die Fläche unter der Kurve (AUC) ist dabei ein Maß für die Übereinstimmung, je näher dieser Wert an 1 ist, desto besser ist die Übereinstimmung. Die ROC-Kurven wurden von Dr. Thomas Ragg erstellt. Eine ROC-Kurve zeigt nicht an, ob die Werte der Testprobe in die gleichen Kategorien eingruppiert werden, darüber gibt eine Verwechslungsmatrix Auskunft.

Um die Experimente der geglätteten Daten miteinander vergleichen zu können, wurden in diesen Fällen die geglätteten Werte der nicht-amplifizierten T47D-DNA als Referenz verwendet.

Daten-Glättung mit dem Savitzky-Golay-Filter

Die Glättung von Array-CGH Daten wurde einzeln für jedes Chromosom mit dem Savitzky-Golay-Filter⁸¹ durchgeführt. Der Savitzky-Golay-Filter verwendet n_L Datenpunkte links und n_R Datenpunkte rechts, die das Glättungsfenster definieren.

$$g_i = \sum_{n=-n_L}^{n=n_R} c_n f_{i+n}$$

Der Wert des Filters wird mit Hilfe des Koeffizienten c_n berechnet. Damit Auslenkungen erhalten bleiben, verwendet der Filter unterschiedliche Koeffizienten. Für den 3K BAC-Array wurde $n_L = n_R = 4$ verwendet, für den Operon-Array und den Agilent-Array wurde ein größeres Glättungsfenster verwendet ($n_L = n_R = 25$ und $n_L = n_R = 330$), das ihrer höheren Auflösung entsprach. Diese Parameter wurden so gewählt, dass die Auflösung der geglätteten Daten für alle drei Plattformen ähnlich war.

2.1.14 Hybridisierung auf Nylonmembranen

Herstellung von Nylonmembranen mit immobilisierter DNA

Zur Herstellung einer Nylonmembran mit immobilisierter DNA wird die DNA zunächst im Verhältnis 1:2 mit 1M NaOH versetzt, somit denaturiert und anschließend mit Wasser auf 50 μ l aufgefüllt. Zur Herstellung der Nylonmembranen wird eine Dot-Blot-Apparatur verwendet. Dabei wird eine positiv geladene Nylonmembran passend zugeschnitten und mit Wasser befeuchtet. Die Apparatur wird nach Hersteller-Angaben zusammengesetzt, die 96 Gefäß mit 200 μ l Wasser befüllt und anschließend ein Vakuum angelegt, um das Wasser abzusaugen. Die denaturierte DNA wird luftblasenfrei in die Gefäß geladen, nach dem Absaugen der DNA wird mit 200 μ l 0,4 M NaOH gespült. Die Membran wird anschließend auf Whatman-Papier getrocknet und mit UV-Strahlung quervernetzt. Es werden jeweils zwei Filter mit identischer Belegung hergestellt. Mit demselben Verfahren können Filter mit jeglichen PCR-Produkten und DNA hergestellt werden.

Hybridisierung markierter Sonden auf Nylonmembranen mit immobilisierter DNA

Der Filter wird in 2x SSC äquilibriert und in eine Hybridisierungsröhre mit 5 ml Dig Easy überführt. 50 μ g Cot-1-DNA werden zusammen mit 20 μ g Heringsperma-DNA in 1 ml Dig Easy 5 min bei 95°C denaturiert und anschließend in die Hybridisierungsröhre gegeben. Es

wird 30 min bei 42°C im Roller-Ofen prähybridisiert. Die Probe wird mit alkali-labilem Dig-dUTP nach Absatz 2.1.7 markiert, in 1 ml Dig Easy 5 min bei 95°C denaturiert, zum Prähybridisierungsansatz hinzugegeben und über Nacht bei 42°C im Roller-Ofen hybridisiert. Wird nicht-amplifizierte DNA als Sonde verwendet, wird für die Dig-Markierung das High Prime-Kit der Firma Roche nach Herstellerangaben verwendet.

Die Waschschritte nach der Hybridisierung erfolgen bei 68°C. Die Temperatur im Hybridisierungsofen wird auf 68°C erhöht, die Hybridisierungslösung verworfen und die Filter in 2x SSC + 0,1% SDS und anschließend 1x SSC + 0,1% SDS kurz geschwenkt. Danach wird für 15 min in 0,5x SSC + 0,1% SDS, dann zweimal für 30 min in 0,1x SSC + 0,1% SDS im Ofen rotierend gewaschen. Zur Visualisierung der Hybridisierungssignale wird die Membran zunächst mit Puffer 1 (0,1 M Maleinsäure, 0,1 M NaCl, pH 7,4) gespült und anschließend in einer Schale mit 25 ml 1%-iger Blocklösung für 30 min geschwenkt. Es werden 2,5 µl des anti-Dig-Fab-Alkalische-Phosphatase-Konjugats zugegeben. Dann wird für weitere 30 min inkubiert. Ungebundener Antikörper wird durch zweimaliges Waschen für 15 min in Puffer 1 + 0,003% Tween entfernt. Der Filter wird kurz in Puffer 3 (0,01 M Tris-HCl, pH 9,5, 0,01 M NaCl) umgepuffert, die überschüssige Flüssigkeit entfernt und der Filter zwischen zwei Folien gelegt. Das Entwicklungsreagenz CSPD wird 1:10 in Puffer 3 verdünnt, 750 µl der Verdünnung werden auf der Membran verteilt. Nach 5 min Inkubation wird die überschüssige Flüssigkeit abgetupft und der Filter 15 min bei 37°C im Ofen inkubiert.

Es werden Filme aufgelegt und nach jeweils 15 min, 30 min und einer Stunde entwickelt.

2.1.15 Nachweis von Nukleotid Polymorphismen (SNPs) in Mse I-Fragmenten

Nach einer spezifischen PCR-Amplifikation (2.1.5) des Mse I-Fragments werden 10 µl des PCR-Ansatzes mit dem für den einzelnen Nukleotid Polymorphismus (SNP) relevanten Restriktionsenzym verdaut. Der Verdau wird in einem Gesamtvolumen von 30 µl durchgeführt, wobei 2,5 µl des zehnfach konzentrierten Restriktionspuffers und 10 U Enzym hinzu gegeben werden. Falls erforderlich, werden 0,3 µl 100-fach konzentriertes BSA zugefügt. Gemäß den Angaben des Herstellers wird der Verdau bei der erforderlichen Temperatur (37, 60 oder 65°C) für 3 h inkubiert (siehe 2.5, Primer für LOH-Analysen).

Zur Auftrennung der Banden wird die Mischung auf einem 2,2%-igen Agarose-Gel (2.2.1) aufgetrennt, als Kontrollen dienen hierbei eine verdaute Negativkontrolle der PCR, ein unverdautes PCR-Amplifikat sowie eine Positivkontrolle mit bekannter Heterozygotie. Je

nachdem welche Allele des spezifischen Mse I-Fragments sich in welcher Zusammensetzung in einer Probe befinden, ergeben sich verschiedene Banden-Muster: Fehlt die Erkennungssequenz für das Restriktionsenzym in beiden Allelen, wird das PCR-Fragment nicht verdaut und man sieht im Gel eine Bande; weisen beide Allele die Erkennungssequenz auf, werden beide verdaut, so dass man im Gel zwei Banden sieht; wenn je ein Allel von jedem Typ vorhanden ist, bleibt eines unverdaut, das andere wird verdaut und das Muster besteht aus drei Banden. Um einen Allelverlust in einer Tumorzelle beurteilen zu können, musste zunächst bestimmt werden, ob der jeweilige Patient heterozygot für das bestimmte Allel ist.

2.1.16 Quantitative real time PCR

Für die quantitative real time PCR wurde das LightCycler-Faststart-DNA-master-SYBR-Green I-Kit verwendet. Ein Ansatz besteht aus 2 µl SYBR-Green/Polymerase-Mix, 3 mM MgCl₂, jeweils 1,6 µM 5'- und 3'-Primer, 1 µl template und Wasser. Der PCR-Ansatz wird in eine LightCycler-Kapillare gefüllt, die PCR mit dem Programm durchgeführt, das in Tabelle 8 aufgeführt ist.

Tabelle 8 PCR-Programm für die real time PCR im LightCycler

Schritt	Temperatur	Zeit	Heizrate	Fluoreszenzmessung
1.	50°C	2:00 min	20°C/min	Keine
2.	95°C	10:00 min	20°C/min	Keine
3.	95°C	0:10 min	20°C/min	Keine
4.	60°C	0:05 min	20°C/min	Keine
5.	72°C	0:10 min	20°C/min	Ein Mal am Ende
6.	Zurück zu 3. 39 Mal			
7.	Auf 95°C	0:00 min	20°C/min	Keine
8.	65°C	1:00 min	20°C/min	Keine
9.	Auf 95°C	0:00 min	0,1°C/min	*
10.	40°C	0:30 min	20°C/min	Keine

*während die Temperatur auf 95°C erhöht wird, wird die Fluoreszenz kontinuierlich gemessen (Schmelzkurve).

Zunächst wird für einen gegebenen Primer eine Standardkurve aus einer Verdünnungsreihe mit einer geeigneten Probe angelegt, um daraus den Korrekturkoeffizienten bestimmen zu können, mit dem verschiedene PCRs mit unterschiedlichen Effizienzen korrigiert werden können. In jede PCR wird eine der Standard-Verdünnungen integriert, die RelQuant-Software kann daraus die Effizienz der PCR berechnen. Für jede Probe wird eine PCR auf ein Referenzgen und eine PCR auf das Testgen durchgeführt. RelQuant ermittelt daraus die relative Menge an Testgen im Vergleich zum Referenzgen. In den letzten beiden Schritten

des oben aufgeführten PCR-Programms wird eine Schmelzkurve der entstandenen PCR-Produkte erstellt. Aus dieser Kurve ist ersichtlich, ob es sich um ein spezifisches Produkt handelt oder ob Kontaminationen vorhanden waren.

2.2 Verwendete Standardprotokolle

2.2.1 Zellkultur

Zelllinie	Medium	Passagierung
T47D (adhärent)	RPMI + 10% FCS	alle 2-3 Tage, 1:5 aufteilen

2.2.2 Agarosegel Elektrophorese

1,5% Agarose in 1xTBE nach Sambrook & Russell⁸². Zur Visualisierung der DNA wurde diese vor dem Auftrag auf das Gel mit 1µl SYBR Gold (Verdünnung 1:6000 in DMSO) versetzt.

2.2.3 Konzentrationsbestimmung von Nukleinsäuren

Eine Konzentrationsbestimmung von Nukleinsäuren erfolgt durch photometrische Messung nach Sambrook & Russell⁸², wenn nicht anders im jeweiligen Methodenabschnitt beschrieben.

2.2.4 Herstellung elektrokompetenter Bakterien und Elektroporation

Es wird nach Sambrook & Russell⁸² vorgegangen.

2.2.5 Plasmid-Herstellung

Es wird nach Sambrook & Russell⁸² vorgegangen, wenn nicht im Einzelfall anders beschrieben. Die Richtigkeit von Plasmiden wird durch entsprechenden Restriktionsverdau und/oder Sequenzierung überprüft. Für die Herstellung von Mini-Präps und Maxi-Präps werden die entsprechenden Qiagen-Kits verwendet, es wird nach Herstellerangaben verfahren. Werden Plasmide für eine weitere Herstellung von Lentiviren benötigt, wird die gefällte DNA 3 Mal mit 70%-igem Ethanol gewaschen.

2.2.6 Herstellung genomischer DNA aus Blut und Zelllinien

Periphere Blut-Lymphozyten (PBLs) werden aus heparinisiertem Blut durch eine Ficoll®-Gradienten-Zentrifugation gewonnen und mit 1xPBS gewaschen. Adhärenz Zellen werden

durch Trypsin-Behandlung abgelöst und ebenfalls mit 1xPBS gewaschen. Die genomische DNA wird anschließend mit dem Genomic-DNA-Extraction-Kit von Promega isoliert, es wird nach Herstellerangaben verfahren. Ein Aliquot der DNA wird auf ein Agarosegel aufgetragen, um zu überprüfen, dass sie nicht degradiert ist. Die Quantifikation verläuft photometrisch.

2.3 Verzeichnis der verwendeten Abkürzungen

Chemische Elemente wurden gemäß ihrer Bezeichnung im Periodensystem und Maßeinheiten mit den international gebräuchlichen SI-Einheiten abgekürzt. Wenn nicht anders angegeben, wurde bei der Abkürzung deutscher Wörter gemäß Duden verfahren.

A	Adenin
α	anti
AUC	area under the curve (ROC-Analyse)
BAC	bacterial artificial chromosome
bp	Basenpaar
C	Cytosin
cen	im Bereich des Centromers
CGH	comparative Genomhybridisierung
CGH-negativ	durch Metaphasen-CGH können keine Abberationen detektiert werden
CGH-positiv	durch Metaphasen-CGH können Abberationen detektiert werden
dBm	Dezibel pro Milliwatt (Leistungspegel)
dd	didesoxy
DNA	Desoxyribonukleinsäure
dNTP	2'-Desoxynukleosid-5'-triphosphat
ds	Doppelstrang
DTT	1,4-Dithiothreitol
FCS	fötales Kälberserum
g	Beschleunigung
h	Stunde
K	Tausend
kb	Kilobase
min	Minute
Mb	Megabase
mRNA	messenger RNA
M0	ohne klinisch manifeste Metastasen
M1	mit klinisch manifesten Metastasen
PCR	Polymerasekettenreaktion
qPCR	quantitative PCR
rfu	relative fluorescent units
rpm	Umdrehungen pro Minute
RT	Raumtemperatur
sec	Sekunde

ss.....	Einzelstrang
T	Thymin
tel.....	im Bereich des Telomers
U	unit (Restriktionsenzym Aktivitätseinheit)
üN	über Nacht

2.4 Material, Geräte, Software, Puffer und Lösungen

Enzyme und Konjugate

Die Schreibweise der Restriktionsenzyme richtet sich nach den Angaben des Herstellers.

Agarase (0,5 U/µl).....	MBI Fermentas, St. Leon-Rot
αDig-Cy5.....	Jackson Laboratories, Bar Harbor, ME
αDig-Fab-Alkalische Phosphatase-Konjugat	Roche, Mannheim
AlwN I (10 U/µl).....	New England Biolabs, Frankfurt
Avr II (4 U/µl).....	New England Biolabs, Frankfurt
Biotin-dCTP	Invitrogen, Karlsruhe
Biotin-dUTP	Roche, Mannheim
Bsr I (5 U/µl).....	New England Biolabs, Frankfurt
Cy3-dCTP, Cy3-dUTP, Cy5-dCTP, Cy5-dUTP	GE Healthcare, Chalfont St. Giles
Dde I (5 U/µl).....	New England Biolabs, Frankfurt
Dig-dUTP	Roche, Mannheim
Dig-dCTP	Roche, Mannheim
Lysozym (50.000 U/mg)	Sigma-Aldrich, Saint-Louis, MO
Mse I (50 U/µl)	New England Biolabs, Frankfurt
PI-Sce I (5 U/µl).....	New England Biolabs, Frankfurt
Proteinase K (0,6 U/µl)	Roche, Mannheim
RNAse A (DNase-frei) (50 U/µl).....	Roche, Mannheim
Streptavidin-Cy3	Jackson Laboratories, Bar Harbor, ME
Taq-Polymerase (50 U/µl).....	Roche, Mannheim
Thermosequenase (32 U/µl).....	GE Healthcare, Chalfont St. Giles
Tru I (50 U/µl).....	MBI Fermentas, St. Leon-Rot
T4-Ligase (5 U/µl)	Roche, Mannheim
Tsp45 I (5 U/µl)	New England Biolabs, Frankfurt

Chemikalien

AdvaSon	Advalytix, München
Agarose SeaKem (low-melt Agarose).....	Biozym, Hess. Oldendorf
Ampicillin.....	AppliChem, Darmstadt
ATP	Roche, Mannheim
BSA Fraktion V	Paesel + Lorei, Hanau
BSA für PCR	Roche, Mannheim
Chloramphenicol	Sigma-Aldrich, Saint-Louis, MO
Cot-1-DNA	Roche, Mannheim
CSPD.....	Roche, Mannheim
dATP, dCTP, dGTP, dTTP	GE Healthcare, Chalfont St. Giles
Dextran Sulfat	Sigma-Aldrich, Saint-Louis, MO

Dig Easy Hybridisierungspuffer.....	Roche, Mannheim
DMSO (Dimethylsulfoxid)	Sigma-Aldrich, Saint-Louis, MO
DNA-Größenmarker.....	Invitrogen, Karlsruhe
Ethanol unvergällt	J.T. Baker, Griesheim
Formamid	Sigma-Aldrich, Saint-Louis, MO
Heringsperma DNA.....	Invitrogen, Karlsruhe
Isopropanol.....	Roth, Karlsruhe
Montagekleber, Fixogum	Marabu, Tamm
Natriumazetat (NaAc)	Sigma-Aldrich, Saint-Louis, MO
OnePhorAll Puffer.....	GE Healthcare, Chalfont St. Giles
PCR-Öl.....	Sigma-Aldrich, Saint-Louis, MO
PicoGreen	Molecular Probes, Leiden, NL
SDS (Sodiumdodecyl-Sulfat).....	Sigma-Aldrich, Saint-Louis, MO
Sephadex	Sigma-Aldrich, Saint-Louis, MO
SYBR Green.....	Molecular Probes, Leiden, NL
Tetracyclin.....	Sigma-Aldrich, Saint-Louis, MO
Tween-20.....	Sigma-Aldrich, Saint-Louis, MO
Wasser (LiChrosolv für Chromatographie).....	Merk, Darmstadt
Yeast t-RNA.....	Invitrogen, Karlsruhe

Kits

BioPrime-Labeling-Kit.....	Invitrogen, Karlsruhe
Expand Long Template Kit	Roche, Mannheim
High Prime Labeling Kit	Roche, Mannheim
LightCycler FastStart DNA Master Sybr Green I	Roche, Mannheim
Maxi-Präp Kit.....	Qiagen, Hilden
Mini-Präp Kit	Qiagen, Hilden

Geräte

Elektroporator.....	Bio-Rad, Waltham, MA
Hybridisierungsofen Shake 'n Stak.....	Hybaid, Heidelberg
LightCycler	Roche, Mannheim
Microarray Scanner Genepix 4000A	Axon Instruments, Union City, CA
MJ Research Peltier Thermal Cycler PTC-200	Bio-Rad, Waltham, MA
MJ Research Peltier Thermal Cycler Tetrad	Bio-Rad, Waltham, MA
Pipetierautomat (Freedom EVO)	TECAN, Männedorf
Pulsed-field Gelelektrophorese-Aparatur.....	GE Healthcare, Chalfont St. Giles
SlideBooster Hybridisierungsstation	Implen, München
Spektrophotometer (GENios).....	TECAN, Männedorf

Zellkultur-Materialien

Fötale Kälber-Serum (FCS).....	PAA, Cölbe
L-Glutamin	PAA, Cölbe
Penicillin/Streptomycin (Pen/Strep).....	PAA, Cölbe
RPMI	PAA, Cölbe
Trypsin-EDTA	PAA, Cölbe

Verbrauchsmaterialien

GAPS II slides.....	Corning, New York, USA
Microcon YM-30 Säulen.....	Millipore, Billerica, MA
Microcon YM-30 Filterplatten.....	Millipore, Billerica, MA
Isopropylboxen.....	Nalgene, Rochester, NY
OpArray V4.0.....	Operon Biotechnologies, Köln
Plugmolds.....	Bio-Rad, Waltham, MA
Zellkulturschalen.....	NUNC, Roskilde
Quantifoil slides.....	Quantifoil, Jena
96-well Photometer-Platte (schwarz).....	Greiner, Kremsmünster, AU
96-well PCR-Platte.....	ABGene, Epsem, UK
96-well deep-well-Platte (0,8 ml).....	Sarstedt, Nümbrecht
96-well-Platten Verschlussfolie.....	ABGene, Epsem, UK

Software

BLAST NCBI.....	NCBI webpage
CAPweb.....	Institut Curie, Paris
CGH Analytics.....	Agilent Technologies Inc., Santa Clara, CA
Freedom EVOware Software.....	TECAN, Männedorf
GenePix Pro 4.1.....	Axon Instruments, Union City, CA
RelQuant.....	Roche, Mannheim
Sigma Plot 2001.....	Systat Software Inc., San Jose, CA
SPSS V13.0.....	SPSS Inc., Chicago, IL
TECAN Magellan.....	TECAN, Männedorf
VAMP.....	Institut Curie, Paris ⁷⁹
Vector NTI 6.0.....	InforMax Inc.

Puffer, Lösungen und Medien

EC-Lyse.....	6 mM Tris-HCl, 1 M NaCl, 100 mM EDTA, 0,5% Brij-58 0,2% Deoxycholat, 0,5% N-Lauroyl-Sarcosin; pH 7,5
PBS (Phosphate Buffered Saline) ..	8,5 mM Na ₂ HPO ₄ , 2 mM KH ₂ PO ₄ , 150 mM NaCl; pH 7,4
PCR-Puffer + dNTPs.....	10 mM MgCl ₂ , 100 mM Tris-HCl, 500 mM KCl; pH 8,5, 1 mM dNTPs
PAA.....	5% Acrylamid, 40 mM Tris, 20 mM NaAc, 1 mM EDTA, 1% Ammoniumpersulfat, 0,1% TEMED; pH 7,8
PETT IV.....	10 mM Tris-HCl, 1 M NaCl; pH 7,6
NDS.....	0,5 M EDTA, 1% N-Lauroyl-Sarcosin; pH 9
20x SSC (Saline-Sodiumcitrat).....	15 mM NaCitrat, 150 mM NaCl; pH 7,4
TAE (Tris-Acetat-Eisessig-Puffer) 2 M Tris, 1 M Eisessig, 5 mM EDTA; pH 8,0	
TBE (Tris-Borat-EDTA-Puffer).....	89 mM Tris, 89 mM Borat, 2 mM Borat; pH 8,3

TE 10 mM Tris-HCl, 1 mM EDTA; pH 7,4

LB (liquid broth) 1% NaCl, 1% Trypton, 5% Yeast Extract; pH 7,0

2.5 Oligonukleotide und Primer

Oligonukleotide und Primer für die primäre Amplifikation humaner genomischer DNA

LIB1 AGT GGG ATT CCT GCT GTC AGT
 ddMse1 TAA CTG ACA G-ddC
 LIG Mse21 AGT GGG ATT CCG CAT GCT AGT
 CAB21 CTG TGT CTG ACG ACT CAG TCT
 ddCAB12 TAA GAC TGA GT-ddC

Primer für spezifische PCRs

Gen/ Marker	5'-Primer	3'-Primer	Tm*
D3S1514	CAGCCAGCAGAATTATGAG	GGCAACAGAGCAAGATGC	60°C
D5S2117	CCAGGTGAGAACCTAGTCAG	ACTGAGTCCTCCAACCATGG	58°C
D6S1633	CTCATGGAGCTTATAGCCTG	TGTTCTTCTGGCTAGCATG	59°C
D17S1322	CTAGCCTGGGCAACAAACGA	GCAGGAAGCAGGAATGGAAC	55°C
BCR-TT1936	CGTGGACAACCTACGGAGTTG	TCAGCCTCAGGACTCTGTG	61°C
CK19	GAAGATCCGCGACTGGTAC	TTCATGCTCAGCTGTGACTG	58°C
Ex2/3	GAAGCGTCTCATGCTGGATC	CAGCCCAAGCCTTGTCTTA	58°C
PUC	ACTGGCTTCAGCAGAGCG	TGGCGGTTTCTCAATGCTC	58°C
pBACe3.6	CACGTACAACATTTTTTAGA	GGCTCTCAGTCCCCGTGGAT	58°C

*verwendete Annealing-Temperatur

Primer für LOH-Analysen

Marker	5'-Primer	3'-Primer	Tm	Enzym	T*
rs7935825	CCATGATTCATAGGTCAGCATGG	CTATTATCTTGGCACCATATTTGGC	63°C	<i>AvrII</i>	37°C
rs621456	CAGATAAGGACAATTTCTTTTG	GATAATGGCTCCAACCTCAAG	56°C	<i>BsrI</i>	65°C
rs2373547	GCTGGCACTTTACTATGCTG	GTTCAATCAAAAGAGCATGTTCAAC	61°C	<i>Tsp45I</i>	65°C
rs4945425	GGCCTATACAAAGCACAGGG	CTTCCATCACCTTTCAAAACCC	60°C	<i>DdeI</i>	37°C

*Temperatur, bei der der Restriktionsverdau durchgeführt wird

Primer für quantitative real time PCR

Gen	5'-Primer	3'-Primer
AKAP3 249	CAGAAGCAAAGCATAAGCAAAG	TAGCAGGCACTTACAGAAAGG
AKAP3 250	GGGATGTTCTTGCTCAGCTC	CCCTCCTCTTTGTCTGTTG
AKAP3 324	GGGTCATCACAAGGGCAGAT	AGCTATTCAGGGTGGAGTGG
AKAP6 259	TCTTCAACCCATGTCCTTCAG	GGCTCTCCAGGGTTACC
AKAP6 294	GGTGTGGACTGACGCAGTTA	ATTACCAGGGCTTTGTTCTCC
AKAP6 329	GGCTGAAACCAAACCTGTCTC	AACAGCCCTTAGAGCCCAATA
BC	GACTCTACCCAGGTTTCTAG	CTGTGCTTACCAAAGGCTTC
IL6	GACTCTACCCAGGTTTCTAG	CTGTGCTTACCAAAGGCTTC
SMS2	GCCAGGAGTTTGAGACCAGT	AGAGATTCACCTGCCTCAGC

2.6 Zelllinien, Bakterien und Plasmide

Zelllinie

Zelllinie	Beschreibung
T47D	Humane Mammakarzinom-Zelllinie; stabile Zelllinie mit komplexen, genomischen Veränderungen und tetraploidem Karyotyp; weist eine Amplifikation des INK4A-Lokus auf

Bakterienstämme und BAC-Library

DH10B ElectroMAX.....	Invitrogen, Karlsruhe
FISHmappedClonesV3.1.....	CHORI (Children's Hospital Oakland Research Institute)

Plasmide

Plasmid	Beschreibung
pBACe3.6	Vektor für bakterielle artifizielle Chromosomen, enthält die PUC-Klonierungskassette, den Replikationsursprung für das Plasmid und das Chloramphenicol-Resistenz-Gen

2.7 Patientenkollektiv

Einzelzell-Amplifikate von disseminierten Tumorzellen aus dem Knochenmark von Mammakarzinompatientinnen, die in dieser Arbeit verwendet wurden, wurden von Oleg Schmidt-Kittler im Rahmen seiner Dissertation hergestellt⁸³. Das Kollektiv dieser Einzelzellen ist in der benannten Arbeit vorgestellt und mittels Metaphasen-CGH charakterisiert worden. Die Patientenproben wurde vom Zentralklinikum Augsburg und der Universitätsklinik Frankfurt bereitgestellt.

2.8 Kooperationspartner

Für das Spotten der Test-BAC-Chips bestand eine Kooperation mit Dr. Bernhard Radlwimmer von der Arbeitsgruppe von Peter Lichter am DKFZ, Heidelberg. Der 3K-Array wurde bei der Firma SCIENION AG in Berlin gespottet. Alle BAC-Klone wurden in Kooperation mit Dr. Richard Reinhard am MPI für Molekulare Genetik in Berlin sequenziert. Dr. Andreas Polten führte die Hybridisierung des 244K Agilent Oligonukleotid-Arrays und die Erfassung und Normalisierung der Rohdaten durch. Frau Sybille Glöckl hat geholfen, die Sequenzierungsdaten zu BLASTen.

3 Ergebnisse

Der Ergebnisteil gliedert sich in zwei Teile: die Etablierung der Array-CGH-Methode (3.1) und die Untersuchung einzelner disseminierten Tumorzellen von Mammakarzinompatientinnen (3.2).

3.1 Etablierung der Array-CGH-Methode

3.1.1 Herkömmliche BAC-Arrays sind ungeeignet für Einzelzellanalysen

Zahlreiche BAC-CGH-Arrays wurden bereits beschrieben (siehe 1.2.3), deshalb sollte zuerst getestet werden, ob sich einer dieser Arrays für Einzelzellanalysen eignet. Von Peter Lichter (DKFZ, Heidelberg) wurden hauseigene BAC-Arrays³⁴ zur Verfügung gestellt und sowohl Einzelzell-Amplifikate als auch Leerkontrollen (Amplifikationsreaktion ohne initiale Zugabe humaner DNA; im weiteren Verlauf dieser Arbeit stets so benannt) wurden auf diesem Array hybridisiert (Daten von Christoph Klein, 2001). Die Leerkontrolle hybridisierte bei diesen Experimenten etwa genauso stark wie das Einzelzell-Amplifikat, zusätzlich konnte der Karyotyp der Einzelzelle nicht dargestellt werden.

Eine mögliche Erklärung hierfür ist, dass bei der Herstellung der amplifizierten Einzelzell-DNA, bzw. der Leerkontrolle, die Enzyme Proteinase K und Mse I verwendet werden (siehe 2.1.5), die beide rekombinant in *E. coli* hergestellt werden und deshalb Spuren von bakterieller DNA enthalten. Diese Kontamination wird somit in die Einzelzell-DNA-Präparation eingeschleust und amplifiziert. Je weniger spezifisches Ausgangsmaterial, also humane DNA, in die Amplifikationsreaktion eingebracht wird, desto mehr unspezifische DNA (Kontamination) wird aufgrund des Funktionsprinzips der PCR amplifiziert. Eine maximale Kontamination sollte dementsprechend in einer Leerkontrolle zu finden sein, wohingegen in nicht-amplifizierter DNA keine Kontamination zu finden sein sollte. Um diese Hypothese zu testen, wurden Amplifikationen von einer, 100 und 40.000 humanen Zellen und einer Leerkontrolle hergestellt. Diese verschiedenen Amplifikate wurden anschließend zusammen mit einer nicht-amplifizierten DNA als Kontrolle auf eine Nylonmembran gespottet (2.1.14). Um eine Kontamination nachzuweisen, wurde anschließend mit einer markierten Leerkontrolle hybridisiert (Abbildung 6). Erwartungsgemäß konnte in der Leerkontrolle die stärkste Kontamination nachgewiesen werden. An dem Hybridisierungssignal der nicht-amplifizierten DNA, bei der kein Signal erwartet wird, ist der Anteil unspezifischer Bindung abzulesen. Je weniger Zellen in den übrigen Amplifikaten als

Ausgangsmaterial eingesetzt wurden, desto mehr Kontamination konnte nachgewiesen werden, die stärkste Kontamination lag im Einzelzell-Amplifikat vor. Das Hybridisierungssignal der nicht-amplifizierten DNA ist stärker als das des Amplifikats von 40.000 Zellen. Eine Erklärung hierfür könnte sein, dass die nicht-amplifizierte DNA als hochmolekulare DNA – und nicht als Gemisch von 100-2000 bp langen Mse I-Fragmenten – vorliegt und deshalb möglicherweise eine stärkere unspezifische Bindung auftritt.

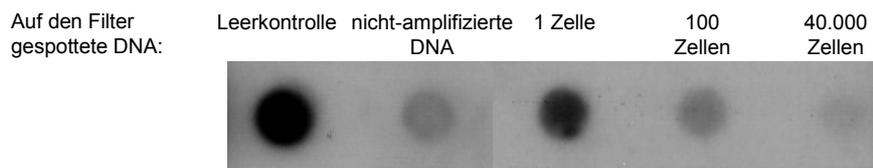


Abbildung 6 Hybridisierung markierter Leerkontrolle auf eine Nylonmembran mit Amplifikaten von Leerkontrolle, nicht-amplifizierter DNA, Einzelzelle, 100-Zell-Pool und 40.000-Zell-Pool.

Von links nach rechts wurde auf der Nylonmembran immobilisiert: Leerkontrolle, nicht-amplifizierte DNA, amplifizierte Einzelzell-, 100-Zell-Pool- und 40.000-Zell-Pool-DNA. Hybridisiert wurde mit markierter Leerkontroll-DNA. Der Leerkontroll-Spot dient als Positivkontrolle für die Hybridisierung, der Spot der nicht-amplifizierten DNA dient als Negativkontrolle für die Hybridisierung, da sie keine Kontamination mit bakteriellen Sequenzen enthält.

Es kann gefolgert werden, dass Einzelzell-Amplifikate aufgrund der geringen Menge an humanem Ausgangsmaterial stark mit der gleichen Kontamination verunreinigt sind, die auch in der Leerkontrolle amplifiziert wird. Da der DKFZ-BAC-Array aus konventionell präparierten und amplifizierten BAC-Plasmiden hergestellt wurde und die BAC-Sonden somit mit einem hohen Anteil an bakteriellen Sequenzen verunreinigt waren, erklärt das zum einen die starke Hybridisierung der Leerkontrolle. Zum anderen deutet dies darauf hin, dass die verhinderte Darstellung des Karyotyps in der amplifizierten Einzelzell-DNA durch die Kreuzhybridisierung der Kontamination bedingt ist.

3.1.2 Einzelzellanalysen auf BAC-Arrays mit reduzierter Kontamination

Die in diesem Abschnitt dargestellten Daten wurden von Oleg Schmidt-Kittler erhoben und werden hier zum besseren Verständnis der folgenden Experimente vorgestellt. Da eine Verminderung der Kontamination in den Einzelzell-Amplifikaten im Nachhinein nicht möglich ist, muss die Kontamination der Hybridisierungsplattform reduziert werden. Deshalb wurde eine Methode entwickelt, um BAC-Klone von kontaminierender, bakterieller DNA zu befreien. BAC-Plasmid-Präparationen werden hierbei zunächst mit der Endonuklease PI-Sce I verdaut. Die PI-Sce I-Schnittstelle ist nur in der pBACe3.6-Vektorsequenz enthalten und

Der Anteil der Kontamination wurde zwar insgesamt durch die Amplifikation verringert, die PFGE-gereinigten BAC-Klone waren aber wieder weniger kontaminiert als die konventionell gereinigten (0,2% im Gegensatz zu 0,002%; Faktor 109; Abbildung 8b). Das bedeutet, dass eine Aufreinigung von BAC-DNA mit der pulsed-field-Gelelektrophorese einer konventionellen Aufreinigung deutlich überlegen ist.

Um zu prüfen, ob PFGE-gereinigte BAC-Klone für die Einzelzell-Array-CGH geeignet sind, wurde ein Test-Array hergestellt, der 30 BAC-Klone von vier verschiedenen Chromosomen (Chromosomen 5, 15, 17 und 21) umfasst (0,03K-Array). Die aufgereinigte BAC-DNA wurde zunächst mittels Adapter-Linker-PCR amplifiziert (siehe 2.1.3), um ausreichend DNA zum Spotten vieler Arrays zur Verfügung zu haben. Zum direkten Vergleich wurden analog konventionell und PFGE-gereinigte BAC-Klone für die Herstellung des 0,03K-Arrays verwendet. Material eines Patienten mit Trisomie 21 eignet sich, um diesen Array zu testen, da nur die BAC-Klone des Chromosoms 21 einen Gewinn an genomischem Material anzeigen sollten, wohingegen die übrigen Chromosomen balanciert sein sollten. Da sowohl konventionell als auch PFGE-gereinigte BAC-Klone auf denselben Array gespottet sind, ist ein direkter Vergleich möglich.

Um zu zeigen, dass der Test-Array in seiner Funktionalität mit konventionellen Arrays vergleichbar ist, wurde mit nicht-amplifizierter DNA eines Patienten mit Down Syndrom hybridisiert (Abbildung 9a,b). Sowohl mit den konventionell als auch mit den PFGE-gereinigten BAC-Klonen konnte der Gewinn des Chromosoms 21 dargestellt werden. Die dreifache Standardabweichung (gestrichelte Linie), die als Signifikanzgrenze dient, ist gleichzeitig auch ein Maß für die Streuung innerhalb des Experiments. Da der verwendete Array zu wenige Spots für die Durchführung einer LOESS-Normalisierung umfasste, wurde die Auswertung nach Quackenbusch⁷⁸, wie in 2.1.11 beschrieben, durchgeführt. Anschließend wurde der 0,03K BAC-Array mit amplifizierter Einzelzell-DNA desselben Trisomie 21-Patienten hybridisiert (Abbildung 9c,d). Mit den konventionell gereinigten BAC-Klonen konnte der Gewinn des Chromosoms 21 nicht dargestellt werden (Abbildung 9c), was durch die starke Streuung (dreifache Standardabweichung = 1,78) der einzelnen Klone bedingt ist. Im Gegensatz dazu wurde die Trisomie 21 von den PFGE-gereinigten Klonen dargestellt (Abbildung 9d), außerdem wurden keine falsch-positiv hybridisierenden Klone detektiert. Dieses Ergebnis legt nahe, dass PFGE-gereinigte BAC-Klone durch eine verringerte Kontamination innerhalb der BAC-Sonden konventionell gereinigten BAC-Klonen überlegen sind. PFGE-gereinigte BAC-Klone stellen somit das Mittel der Wahl dar,

um hochspezifische und hochauflösende Array-CGH Analysen von Einzelzell-Amplifikaten durchführen zu können.

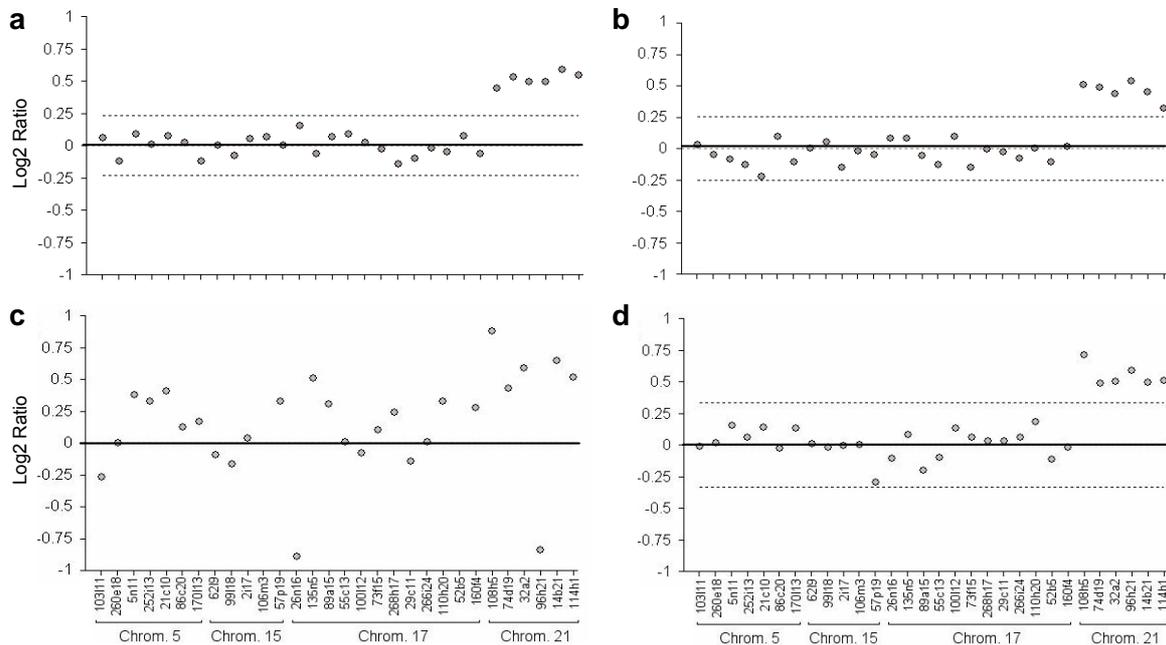


Abbildung 9 Hybridisierung von DNA eines Trisomie 21 Patienten auf den 0,03K BAC-Array aus konventionell- und PFGE-gereinigten BAC-Klonen

Hybridisiert wurde Cy5-markierte DNA eines Patienten mit Trisomie 21 gegen Cy3-markierte normale männliche DNA. Aufgetragen ist das logarithmierte Verhältnis des Fluoreszenz-Signals der Probe zu dem der Referenz gegen die Position der BAC-Klone im Genom. Die Signifikanzgrenze berechnet sich aus der dreifachen Standardabweichung aller balancierten Klone innerhalb eines Experiments und ist durch die gestrichelte schwarze Linie angezeigt. (a) und (c): Hybridisierung auf konventionell gereinigten BAC-Klonen, (b) und (d) auf PFGE-gereinigten BAC-Klonen. (a) und (b): Hybridisierung mit nicht-amplifizierter DNA aus 10^6 Zellen, (c) und (d): Hybridisierung amplifizierter Einzelzell-DNA ((c): Signifikanzgrenze=1,78, jenseits der dargestellten Skala)

3.1.3 Optimierung der Spotting-Parameter

Zusätzlich zur Aufreinigung der zu spottenden BAC-Klone mittels PFGE stellten die Spotting-Parameter und die Vorbereitung der Sonden für das Spotten weitere notwendige Optimierungsparameter dar, um auch mit größeren Arrays ähnliche Ergebnisse zu erzielen, wie mit dem 0,03K-Set. Dazu wurden weitere 120 BAC-Klone aus den Chromosomen 1, 4, 5, 6, 15, 17, 19, 20 und 21 ausgewählt, die isoliert, amplifiziert und gespottet wurden.

Um beurteilen zu können, ob ein Array gute Hybridisierungsergebnisse liefert, wurde stets mit derselben amplifizierten Einzelzell-DNA eines Patienten mit Trisomie 21 als Positivkontrolle hybridisiert. Anhand dieser Zelle mit bekanntem Karyotyp kann das Hybridisierungsverhalten der BAC-Klone beurteilt werden. BAC-Klone des Chromosoms 21 werden in diesem Fall in die Kategorie „richtig positiv“ eingruppiert, wenn sie einen Gewinn

anzeigen, und als „falsch negativ“ beurteilt, wenn sie keinen Gewinn anzeigen. Die übrigen Klone werden in die Kategorie „falsch positiv“ eingruppiert, wenn sie fälschlicherweise Gewinne oder Verluste anzeigen, und als „richtig negativ“ bezeichnet, wenn sie balanciert sind. Der Prozentsatz berechnet sich dabei aus dem Anteil der Klone des Chromosoms 21 bzw. der übrigen autosomalen Klone. Der Anteil „richtig positiver“ und „richtig negativer“ Klone sollte so hoch wie möglich sein, gleichzeitig sollte der Anteil der „falsch positiven“ und „falsch negativen“ Klone so gering wie möglich sein. Die Auswertung wurde wie in 2.1.11 beschrieben durchgeführt.

Die BAC-DNA muss nach der PCR-Amplifikation in einen geeigneten Spottingpuffer überführt werden, gleichzeitig muss die DNA-Konzentration erhöht werden. Dabei ist der Spottingpuffer ebenso entscheidend wie die Oberfläche, auf die gespottet wird; Größe und Homogenität der Spots werden durch den Spottingpuffer bestimmt⁸⁵, die Oberfläche des Objektträgers trägt zu unspezifischer Hintergrund-Fluoreszenz bei⁸⁶. Auf die Darstellung der Spottingversuche auf verschiedene Oberflächen wird an dieser Stelle verzichtet. Getestet wurden Epoxid-beschichtete Objektträger der Firma Quantifoil und verschiedene Gamma-Amino-Propyl-Silan-beschichtete Objektträger der Firma Corning (GAPSII und Ultra-GAPS). Die Epoxid-Beschichtung bewirkte, dass die Spots zu klein wurden (Durchmesser ca. 50 µm), um sie mit dem zur Verfügung stehenden Scanner auflösen zu können. Auf Epoxid-beschichteten Objektträgern und auf Ultra-GAPS Objektträgern wurde eine hohe unspezifische Hintergrund-Fluoreszenz beobachtet, nicht aber auf GAPSII Objektträgern, die deshalb für alle weiteren Versuche verwendet wurden.

Die Test-BAC-Arrays HD I bis HD V wurden in Kooperation mit Bernhard Radlwimmer am DKFZ in Heidelberg hergestellt. Als Spottingpuffer für diese Test-Arrays wurde 3 x SSC mit und ohne Zusatz von Betain verwendet, da 3 x SSC der in der Literatur verwendete Standardpuffer ist. Einen geringen Arbeitsaufwand für die Aufreinigung der PCR-Produkte erforderte die Verwendung von Microcon YM-30 Filterplatten (Abbildung 10, HD I), die im 96-well-Format mit Hilfe eines Vakuums verwendet werden. Mit dem Test-Array HD I konnte die Trisomie 21 aber nur mit 44,4% der Klone des Chromosoms 21 dargestellt werden (Tabelle, unten in Abbildung 10). Um auszuschließen, dass es sich dabei um ein Problem handelte, das durch die Behandlung im 96-well-Format verursacht worden war, wurde ein weiterer Probensatz zur Herstellung von Test-Array HD II mit einzelnen YM30-Säulen aufgereinigt (Abbildung 10, HD II). Das führte jedoch nicht zu einer Verbesserung des Ergebnisses, im Gegenteil, keiner der Klone des Chromosoms 21 konnte die Trisomie 21

darstellen. Da es möglich ist, dass kleine PCR-Fragmente bei der YM30-Säulen-Aufreinigung verloren gehen, wurde daraufhin eine Aufreinigung der PCR-Produkte durch eine Ethanol-Präzipitation durchgeführt (Abbildung 10, HD III bis Sci II). Die BAC-Klone für den Test-Array HD III wurden durch eine Platten-Fällung ohne Waschen mit 70%igem Ethanol aufgereinigt. Kleine, nicht-homogene Spots und schlechte Darstellung der Trisomie 21 wurden beobachtet. Die Proben für die Test-Arrays HD IV und V wurden manuell gefällt und mit 70%igem Ethanol gewaschen, um auszuschließen, dass Spot-Morphologie und Hybridisierungsverhalten von Test-Array HD III durch Salzrückstände verursacht wurden. 3 x SSC Spottingpuffer mit und ohne Zusatz von Betain wurden für die Test-Arrays HD IV bzw. HD V verwendet, um einen Einfluss des Betains auszuschließen, es konnte jedoch keine Verbesserung der Spot-Morphologie oder des Hybridisierungsverhalten festgestellt werden.

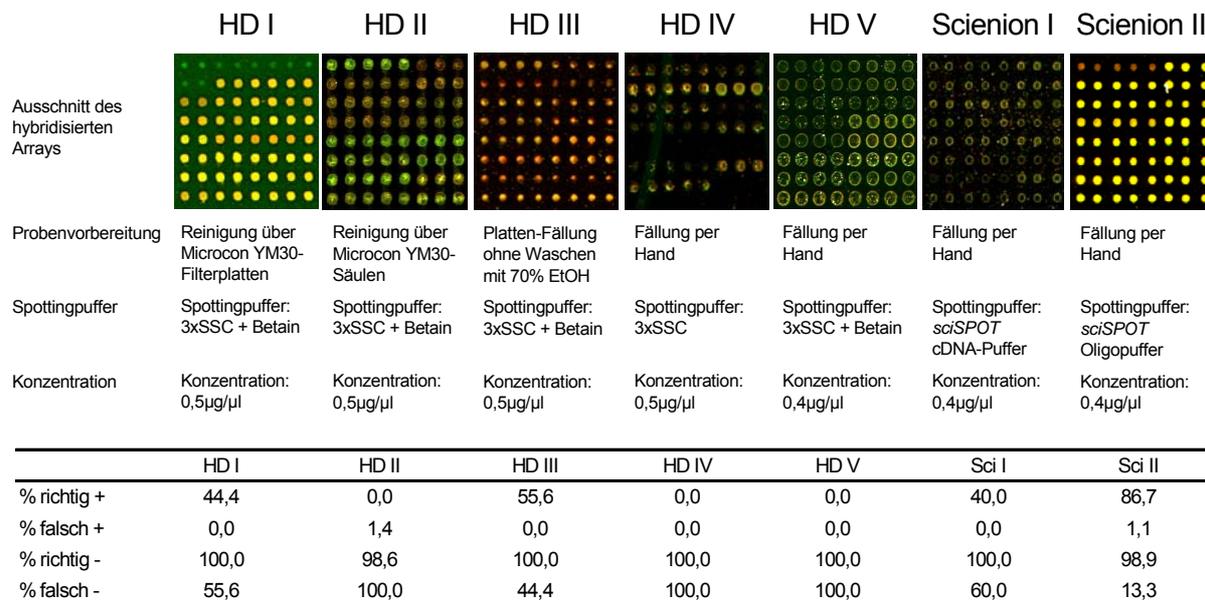


Abbildung 10 Optimierung von Spotting-Parametern und Probenvorbereitung

Es sind die Hybridisierungsergebnisse der einzelnen Test-Arrays HD I bis Scienion II dargestellt. Im oberen Teil der Abbildung ist ein exemplarischer Ausschnitt des hybridisierten Arrays zu sehen, die unterschiedliche Spot-Morphologie und Homogenität ist zu erkennen. In der Mitte sind die Parameter der Probenvorbereitung, der gewählte Spottingpuffer und die Konzentration der gespotteten DNA abzulesen. Die Tabelle im unteren Teil der Abbildung zeigt, wie hoch der prozentuale Anteil an richtig bzw. falsch positiv und negativ hybridisierenden BAC-Klonen auf jedem der Test-Arrays ist. Es wurde mit einer amplifizierten Einzelzell-DNA eines Patienten mit Trisomie 21 hybridisiert.

Die Test-Arrays Sci I und Sci II wurden von der Firma Scienion in Berlin hergestellt. Die BAC-Klone wurden mittels Ethanol-Präzipitation manuell aufgereinigt und zwei verschiedene Spottingpuffer getestet, die durch die Firma zur Verfügung gestellt wurden. Da die Herstellung der BAC-Klone teuer und zeitaufwändig ist, wurde die DNA-Konzentration auf 0,4 µg/µl verringert, so dass mehr Arrays aus den vorbereiteten BAC-Klonen gespottet

werden konnten. Die Verwendung des *sciSPOT*-Oligo-Spottingpuffers führte zu gleichmäßiger Spotmorphologie und Homogenität im Test-Array Sci II sowie zu einem guten Hybridisierungsverhalten. 86,7% der Klone, die das Chromosom 21 repräsentieren, hybridisierten „richtig positiv“ mit dem DNA-Amplifikat der Einzelzelle des Trisomie 21-Patienten. Mit den Spotting-Parametern für den Test-Array Sci II wurden die besten Hybridisierungsergebnisse erzielt, deshalb wurden diese Parameter – manuelle Ethanol Fällung, *sciSPOT*-Oligo-Spottingpuffer, eine DNA-Konzentration von 0,4 µg/µl und Herstellung des Arrays durch die Firma Scienion – für die Herstellung des 3K BAC-Arrays verwendet.

3.1.4 Präparation des 3K-BAC-Satzes

Automatisierung der Adapter-Linker-PCR-Amplifikation und Fällung der BAC-DNA

Die BAC-DNA des 3017 Klone umfassenden Satzes von FISH-kartierten BAC-Klonen (3K-Set) wurde mit dem in 2.1.1 beschriebenen Verfahren isoliert. Da eine manuelle Konzentrationsbestimmung und –angleichung, Amplifikation und Fällung von über 3000 BAC-Klonen zum einen sehr arbeitsaufwändig, vor allem aber die Verwechslungsgefahr hoch und die Reproduzierbarkeit gering ist, wurden die Prozesse der DNA-Konzentrationsbestimmung, der DNA-Konzentrationsangleichung, der Amplifikation mittels Adapter-Linker-PCR und der Ethanol-Präzipitation automatisiert. Folgende Geräte der Firma Tecan wurden für die Automatisierung verwendet: der Pipettierautomat FreedomEVO, das Plattenphotometer GENios und die EVOware- und Magellan-Software zum Programmieren der Skripte.

Um einen Prozess mit dem FreedomEVO-Pipettierautomaten zu automatisieren, muss ein Programm geschrieben werden, in dem sowohl die einzelnen Bewegungsabläufe des Automaten als auch die aufzunehmenden und abzugebenden Volumina genau definiert werden. In den zu automatisierenden Prozessen werden verschiedene Flüssigkeiten verwendet, die sich durch ihre Viskosität oder ihre Oberflächenspannung unterscheiden. Lösungen, die Enzyme enthalten, besitzen z.B. eine höhere Viskosität als wässrige Lösungen und wässrige Lösungen besitzen eine wesentlich höhere Oberflächenspannung als Alkohol.

In verschiedenen Testläufen wurden für jede der verwendeten Flüssigkeiten so genannte „Flüssigkeitsklassen“ definiert, die festlegen, mit welcher Geschwindigkeit die Flüssigkeit aufgenommen und abgegeben wird. Bei Flüssigkeiten mit niedriger Oberflächenspannung

stellte sich heraus, dass nach Aufnahme der Flüssigkeit Luft angesaugt werden musste, um ein Heraustropfen der Flüssigkeit zu verhindern. Die Einstellung der Flüssigkeitsklassen war vor allem für die Automatisierung der PCR wichtig, da es bei der PCR darauf ankommt, dass sehr kleine Volumina präzise pipettiert werden. So mussten z.B. Enzym-haltige Lösungen mit sehr geringer Geschwindigkeit angesaugt und abgegeben werden, damit korrekte Volumina pipettiert wurden. Sehr geringe Flüssigkeitsmengen wässriger Lösungen (wie die BAC-DNA) mussten mit hoher Geschwindigkeit abgegeben werden, um eine vollständige Entleerung der Spitzen zu erreichen.

Unter Verwendung der zuvor definierten Flüssigkeitsklassen wurde zunächst die Bestimmung der DNA-Konzentration der BAC-DNA automatisiert und die Konzentration auf 100 ng/µl angeglichen (2.1.2). Anschließend wurde getestet, ob eine Adapter-Linker-PCR zur Amplifikation der BAC-DNA, die durch den Roboter pipettiert wurde, ebenso effizient ist wie eine manuell pipettierte PCR. Produkte von parallel durchgeführten, maschinellen und manuellen PCRs von sieben BAC-DNAs wurden auf ein Gel aufgetragen (Abbildung 11a). Die Effizienz der PCR lässt sich durch die Intensität des Schmiers und die Bandbreite der verschiedenen langen Fragmente beurteilen. Nach erfolgreicher Einstellung aller kritischen Parameter (Definition der Flüssigkeitsklassen s.o.), konnte kein Unterschied in der Effizienz zwischen manuell und maschinell pipettierter PCR festgestellt werden, die Intensität des Schmiers und die Bandbreite der Fragmente war vergleichbar.

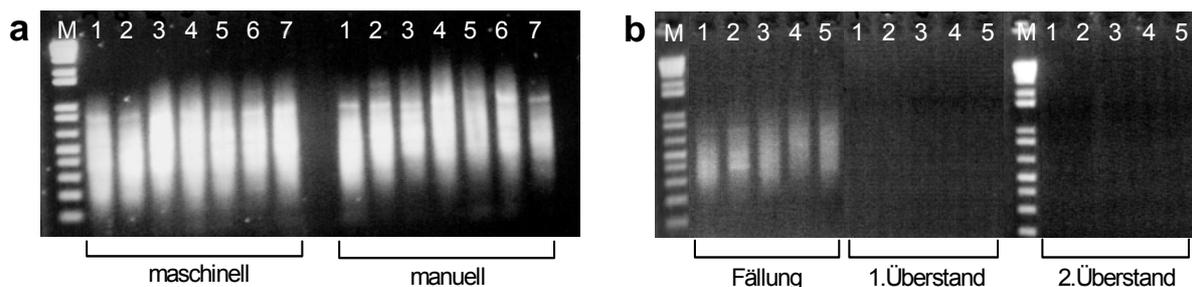


Abbildung 11 Manuell und maschinell pipettierte Adapter-Linker-Amplifikation und maschinelle Fällung von BAC-DNA

(a) Adapter-Linker amplifizierte BAC-DNA; links: mit Hilfe des Pipettierautomaten pipettiert, rechts: manuell durchgeführte PCR. Es wurden jeweils 3 µl des primären PCR-Produkts auf ein 1,5%iges Agarose-Gel aufgetragen. (b) Fällung der amplifizierten BAC-DNA mit Hilfe des Pipettierautomaten in 96-deep-well-Platten (Zentrifugation bei 5.000 g); links: Jeweils 1 µl der gefällten und rückgelösten DNA wurde auf ein 1,5%iges Agarose-Gel aufgetragen, Mitte: Der erste Überstand der Fällung wurde gesammelt und bei 25.000 g zentrifugiert, rückgelöst und das gesamte Volumen aufgetragen, ebenso wurde für den zweiten Überstand nach dem Waschen des Pellets (rechts) verfahren.

Anschließend wurde die Fällung der amplifizierten BAC-DNA automatisiert. Die DNA-Pellets nach der Präzipitation und Zentrifugation in 96-well-Platten sind sehr labil, da Platten

lediglich mit 5000 g, im Gegensatz zu 25.000 g bei einer manuellen Fällung, zentrifugiert werden können. Um trotzdem keine gefällte DNA zu verlieren, wurden die Pipettierparameter so angepasst, dass die Überstände nach der Präzipitation (Überstand 1) und nach dem Waschen (Überstand 2) abgesaugt werden können, ohne Turbulenzen im Gefäß zu erzeugen. Um zu überprüfen, dass die Überstände keine DNA enthielten, wurden die abgenommenen Überstände jedes Gefäßes in 1,5-ml-Reaktionsgefäße überführt, bei 25.000 g zentrifugiert und die Überstände verworfen. Nach dem Trocknen wurden 10 µl Wasser in das Reaktionsgefäß gegeben, und nach einer Inkubationszeit wurde das Gesamtvolumen auf ein Agarose-Gel aufgetragen. Die Pipettierparameter wurden so optimiert, dass die Überstände frei von DNA waren. Abbildung 11b zeigt, dass sowohl Überstand 1 als auch Überstand 2 keine DNA enthalten. Abschließend lässt sich festhalten, dass mit den speziell entwickelten Programmen zur Automatisierung der Amplifikation und Fällung der BAC-DNA 2940 BAC-Klone des 3K-Sets erfolgreich amplifiziert und gefällt werden konnten. Insgesamt 77 BAC-Klone (2,5%) konnten nicht erfolgreich isoliert oder amplifiziert werden.

Überprüfung der Annotation der BAC-Klone des 3K BAC-Satzes

Für die Auswertung von Array-CGH-Daten ist von entscheidender Bedeutung, dass die einzelnen BAC-Klone korrekt annotiert sind, damit ihre physikalische Lokalisation innerhalb des Genoms richtig dargestellt werden kann. Aus diesem Grund wurden alle BAC-Klone des 3K-Sets mit dem T7-Primer (siehe Abbildung 4) ansequenziert. Mit einer BLAST-Analyse wurden die Sequenzierungsdaten mit Sequenzen der NCBI-Datenbank verglichen. Das BLAST-Programm ermittelt homologe Sequenzen und gibt an, wie ähnlich sie sich sind. Die Sequenz, die die beste Übereinstimmung mit der BAC-Sequenz hatte, wurde verwendet, um die Identität der BAC-DNA zu überprüfen. Bei Abweichungen in Bezug auf die Angaben zur Lokalisationen eines BAC-Klons wurde das Ergebnis der BLAST-Analyse als neue Lokalisationsangabe verwendet. Insgesamt waren 329 der 2940 BAC-Klone (11,2%) falsch annotiert, 87 davon (3,0%) waren laut Sequenz auf anderen Chromosomen lokalisiert.

3.1.5 Etablierung des Hybridisierungsprotokolls für Einzelzell-Array-CGH

Bei der Etablierung des Hybridisierungsprotokolls zur Untersuchung von Einzelzell-DNA-Amplifikaten mit BAC-Arrays wurde untersucht, welchen Einfluss die Blockierung unspezifischer Bindungen auf das Hybridisierungsergebnis hat. Hierbei wird zwischen der

Blockierung unspezifischer Bindung der markierten DNA an den Array und der Blockierung repetitiver Sequenzen unterschieden. Außerdem wurde die Zusammensetzung der Hybridisierungslösung ebenso wie die Hybridisierungsparameter variiert, um die geeigneten Bedingungen zu identifizieren. Schließlich wurde getestet, welches Markierungsverfahren für die Markierung der Einzelzell-DNA die besten Ergebnisse erzielt.

Blockierung unspezifischer und unerwünschter Bindungen

Die Blockierung unspezifischer Bindungen der markierten DNA auf dem Glasobjektträger und unerwünschter Bindungen an die immobilisierten Sonden kann das Hybridisierungsergebnis verbessern. Verschiedene Blockierungsansätze wurden miteinander verglichen: 1.) Hering Sperm-DNA (Standardblockierung in der Literatur), 2.) Heringsperma-DNA zusammen mit 6 μg unmarkierter, amplifizierter Leerkontrolle und 3.) Heringsperma-DNA zusammen mit 3 bzw. 6 μg pBACe3.6-Leervektor. Nach der Blockierung wurde der Array mit differenziell markierter Positivkontrolle (Amplifikat einer normalen Einzelzell-DNA) und Leerkontrolle hybridisiert. Zur Beurteilung der Blockierung wurde die mittlere Signalintensität für Positiv- und Leerkontrolle aus fünf Replikaten bestimmt und die Differenz (Δ) zwischen beiden Werten berechnet (Abbildung 12).

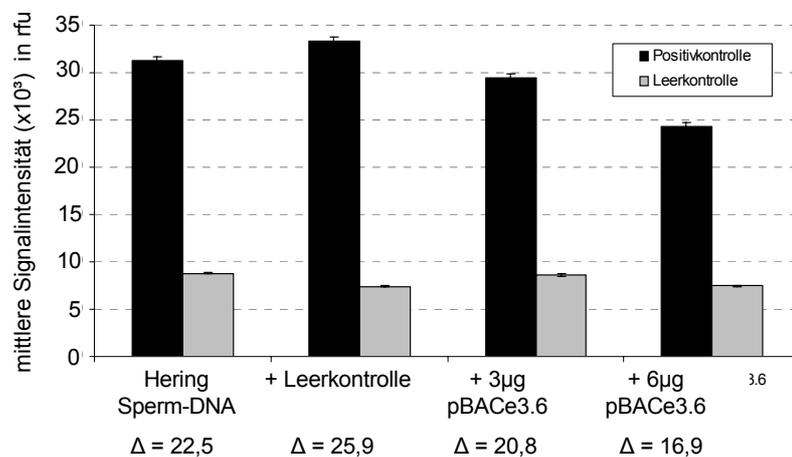


Abbildung 12 Einfluss verschiedener Blockierungsbedingungen auf die Signalintensität

Als Standardblockierung wurden 25 μg Hering Sperm-DNA (ganz links) verwendet. Weitere getestete Blockierungen enthielten zusätzlich 6 μg unmarkierter Leerkontrolle, 3 μg und 6 μg pBACe3.6-Leervektor (von links nach rechts). Anschließend wurde mit Cy3-markierter Leerkontrolle und Cy5-markierter Normalzell-DNA (Positivkontrolle) hybridisiert, und die mittleren Signalintensitäten für beide Kanäle wurden aufgetragen (Einheit: $\times 10^3$ rfu (relative Fluoreszenzeinheit)). Die Differenz Δ wurde zwischen beiden Werten berechnet. Je höher Δ , desto effektiver war die Blockierung. Wichtig ist außerdem, dass die Signalintensität für die Positivkontrolle nicht durch die Blockierung verringert wird. Die Standardabweichung aus fünf Replikaten ist durch die Fehlerbalken angezeigt.

Bei der Verwendung von Heringsperma-DNA³⁷ wurde eine Differenz von $22,5 \times 10^3$ rfu

(relative Fluoreszenzeinheit) bestimmt. Wurde der Heringsperma-DNA zusätzlich unmarkierte Leerkontrolle zugefügt, wurde eine Differenz von $25,9 \times 10^3$ rfu erzielt. Eine Blockierung mit Heringsperma-DNA und 3 bzw. 6 μg pBACe3.6-Leervektor führte zu einer Differenz von $20,8 \times 10^3$ rfu bzw. $16,9 \times 10^3$ rfu. Die Verwendung von pBACe3.6-Leervektor führte im Vergleich zum Standard zu einer Verschlechterung der Signalintensität der Positivkontrolle ohne die Bindung der Leerkontrolle zu blockieren und dadurch zu einer verringerten Differenz. Die Blockierung mit unmarkierter Leerkontrolle erhöhte die Signalintensität der Positivkontrolle und verringerte gleichzeitig die Signalintensität der Negativkontrolle und führte im Vergleich zum Standard zur besten Blockierung. Da bei einer Blockierung des BAC-Arrays mit unmarkierter Leerkontrolle das beste Verhältnis von Signalintensität einer Positivkontrolle zu der einer Negativkontrolle beobachtet wurde, wurde diese Blockierung für alle weiteren Experimente verwendet.

Blockierung repetitiver Sequenzen

Sowohl die zu hybridisierende DNA als auch die immobilisierte BAC-DNA weisen repetitive Sequenzen auf, die unspezifisch miteinander hybridisieren können und die durch humane Cot-1-DNA blockiert werden können (endogene repetitive Sequenzen). Eine weitere Quelle repetitiver Sequenzen wird durch die Primer bei der Markierungs-PCR eingeführt und könnte zu unspezifischen Bindungen beitragen (exogene repetitive Sequenzen). Um die optimale Blockierung endogener repetitiver Sequenzen zu identifizieren, wurden verschiedene Mengen Cot-1-DNA in Anlehnung an die Literatur^{37, 38, 87} getestet. Außerdem wurde untersucht, ob die Entfernung der Primer zur Eliminierung exogener repetitiver Sequenzen notwendig ist. Bei der Literatur-Recherche fiel auf, dass stets eine Hybridisierungstemperatur von 37°C verwendet wurde, wohingegen bei der Metaphasen-CGH von Einzelzellen eine Temperatur von 42°C ⁷² verwendet wird. Deshalb wurde untersucht, welche Hybridisierungstemperatur für die Array-CGH von Einzelzell-DNA geeigneter war.

Zur Beurteilung des Hybridisierungsergebnisses wurde mit einem Einzelzell-Amplifikat eines Patienten mit Trisomie 21 hybridisiert, die Auswertung erfolgte nach dem in 3.1.3 beschriebenen Vorgehen. Zunächst wurden 50 μg Cot-1-DNA und eine Hybridisierungstemperatur von 37°C verwendet (Tabelle 9, Spalte 1). Mit den Hybridisierungsbedingungen dieses Protokolls konnte die Trisomie 21 nicht dargestellt werden. Daraufhin wurde die Hybridisierungstemperatur auf 42°C erhöht, was zu einem verbesserten Ergebnis führte (Spalte 2). Zuletzt wurde die Menge Cot-1-DNA auf 80 μg

erhöht. Die Kombination von 80 µg Cot-1-DNA und einer Hybridisierungstemperatur von 42°C führte zum besten Ergebnis der erwarteten Darstellung des Trisomie 21-Karyotyps (Spalte 4). Diese Bedingungen wurden für alle weiteren Experimente verwendet. Bei diesen Experimenten wurden die zur Markierung des Einzelzell-Amplifikats verwendeten Primer nach der Markierungsreaktion, analog zum Protokoll der Einzelzell-Metaphasen-CGH^{72, 83}, durch einen Verdau mit dem Restriktionsenzym Tru I (Isoschizomer von Mse I) abgeschnitten. Da dieser Schritt zeit- und kostenaufwändig ist, wurde untersucht, ob der Tru I-Verdau notwendig ist. Es konnte gezeigt werden, dass das Hybridisierungsergebnis mit und ohne Tru I-Verdau gleich war (Spalten 2 und 3, 4 und 5). Bei allen weiteren Experimenten wurde deshalb auf den Tru I-Verdau verzichtet.

Tabelle 9 Wahl der Hybridisierungsbedingungen

	1	2	3	4	5
Tru I-Verdau	+	+	-	+	-
Cot-1-DNA	50µg	50µg	50µg	80µg	80µg
Hyb.-Temp.	37°C	42°C	42°C	42°C	42°C
% richtig +	0	83,3	83,3	100	100
% falsch +	0	0	0	0	0
% richtig -	100	100	100	100	100
% falsch -	100	16,6	16,6	0	0

Hybridisierung des 3K BAC-Arrays

Die bisherigen Experimente wurden mit dem Sci II-Test-Chip (siehe Abbildung 10) durchgeführt. Der Sci II-Test-Chip hat eine Hybridisierungsfläche von 2,2 cm² und kann im offenen Tropfen (siehe 2.1.8) hybridisiert werden, was den Vorteil einer optimalen Durchmischung des Hybridisierungsansatzes während der Hybridisierung bietet. Der 3K BAC-Array hat eine um den Faktor drei größere Hybridisierungsfläche von 6,5 cm². Zwar wäre eine Hybridisierung im offenen Tropfen auch für den 3K-Array möglich, das Hybridisierungsvolumen und damit das Probenvolumen müssten allerdings verdreifacht werden. Aufgrund der hohen Kosten für die Markierungsreaktion wurde deshalb ein alternatives Hybridisierungsverfahren angestrebt, bei dem ein geringeres Volumen verwendet werden kann. Wird beispielsweise eine Hybridisierung unter einem Deckglas durchgeführt, können sehr kleine Volumina eingesetzt werden. Eine Durchmischung der Hybridisierungsflüssigkeit unter dem Deckglas ist aber ohne entsprechende Hybridisierungsstation nicht möglich. Es wurde deshalb untersucht, ob ein besseres Hybridisierungsergebnis durch die Verwendung der Hybridisierungsstation SlideBooster der Firma Implen erzielt werden kann. Bei diesem Gerät werden Schallwellen in die

Hybridisierungslösung übertragen, die dadurch durchmischt wird. Zur Beurteilung der Hybridisierung wurde wieder mit einem Einzelzell-Amplifikat eines Patienten mit Trisomie 21 hybridisiert (s.o. und 3.1.3). Da nur solche Spots in der Auswertung berücksichtigt wurden, die eine Signalintensität von mindestens zweifacher Hintergrundfluoreszenz aufwiesen, führte eine hohe Gesamtsignalintensität zu einem höheren Anteil an auswertbaren Klonen (% Klonen). Der Anteil der auswertbaren Klone wurde berücksichtigt, um die Güte der Hybridisierung zu beurteilen. Bei einem Array mit einer großen Fläche spielt auch die Homogenität der Hybridisierung eine Rolle, die mit dem Ausmaß der Durchmischung des Hybridisierungsansatzes zusammenhängt. Es sollten möglichst keine Hybridisierungsgradienten entstehen, wie sie z.B. in der Hybridisierung 4 (Abbildung 13) zu sehen sind.

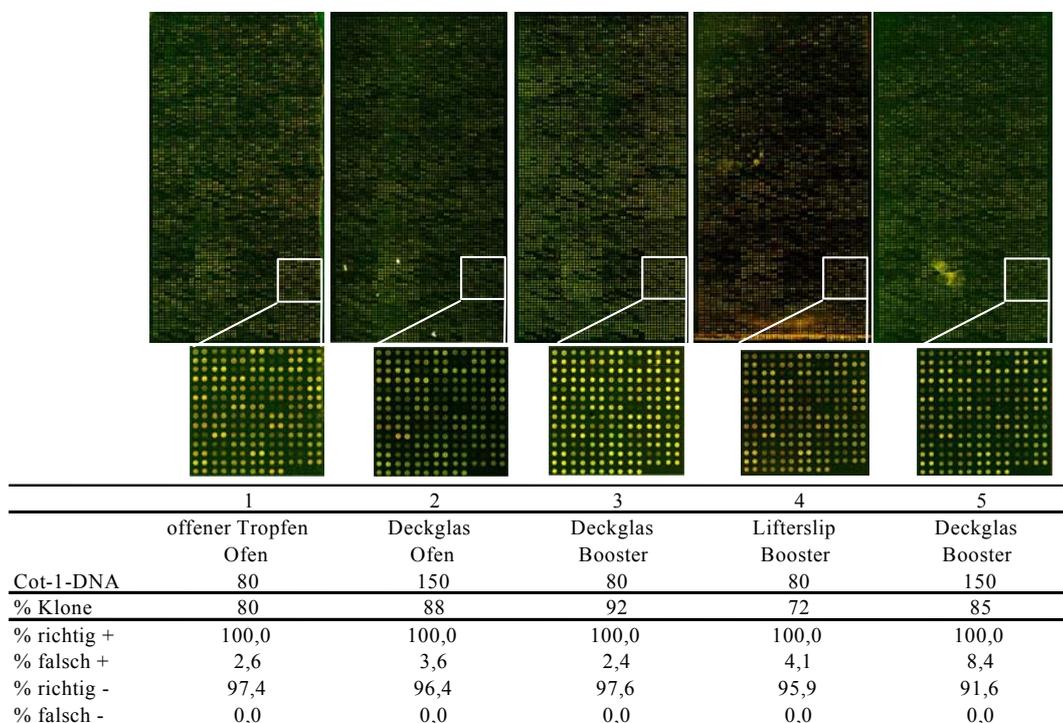


Abbildung 13 Test von Hybridisierungsbedingungen und Hybridisierungsstation für 3K-Array

Der 3K-Array wurde mit einer amplifizierten Einzelzell-DNA eines Patienten mit Trisomie 21 hybridisiert. Verglichen wurden die Hybridisierung im offenen Tropfen (1), unter dem Deckglas im Schüttelofen (2), unter dem Deckglas im SlideBooster (3) und unter dem Lifterslip im SlideBooster (4). Außerdem wurde untersucht, ob der Cot-1-DNA-Gehalt auf 150 µg erhöht werden muss (5). Das bereits beschriebene Bewertungssystem für Hybridisierungen (3.1.3) wurde erneut angewandt, zusätzlich wurde die Signalintensität der Hybridisierung bewertet, indem der Prozentsatz der auswertbaren Klone ermittelt wurde (nur Klone mit einem Signal zu Hintergrund-Verhältnis > 2 werden in der Auswertung berücksichtigt).

Eine Hybridisierung wurde im offenen Tropfen (Hybridisierung 1, Abbildung 13), weitere Hybridisierungen unter dem Deckglas im Hybridisierungssofen (Hybridisierung 2) und im SlideBooster (Hybridisierung 3) durchgeführt. Die Firma Implen bietet im Zusammenhang mit dem SlideBooster den „Lifterslip“ anstelle eines Deckglases an, der ebenfalls getestet

wurde (Hybridisierung 4). Der Lifterslip führte zu einer unerwünschten Gradientenbildung und wurde deshalb nicht verwendet. Die Hybridisierung unter dem Deckglas im SlideBooster führte zu den besten Hybridisierungsergebnissen und auch zu den höchsten Signalintensitäten und wurde deshalb für alle weiteren Experimente verwendet. Zusätzlich wurde untersucht, ob eine Erhöhung des Cot-1-DNA-Gehalts auf 150 µg aufgrund der Erhöhung der Klonzahl und der Hybridisierungsfläche notwendig war (Hybridisierung 5). Die Verwendung von 150 µg Cot-1-DNA führte zu einer Verschlechterung des Hybridisierungsergebnisses, es wurde deshalb weiterhin mit 80 µg Cot-1-DNA hybridisiert.

Untersuchung verschiedener Markierungsverfahren für Einzelzell-DNA

In den bisher gezeigten Experimenten waren Referenz- und Proben-DNA durch ein direktes Markierungsverfahren in einer PCR markiert worden. Das Enzym Thermosequenase wurde hierbei eingesetzt, da diese die großen Cy3- und Cy5-konjugierten Nukleotid-Analoga besser in die DNA inkorporieren kann als Taq-Polymerase. Da die direkte Markierung der DNA mit Thermosequenase durch Fluoreszenz-konjugierte Nukleotide teuer ist, sollte untersucht werden, ob andere Markierungsverfahren zu vergleichbaren oder verbesserten Hybridisierungsergebnissen führen.

In der Literatur beschriebene Markierungsverfahren sind das Klenow-Labeling und indirekte Markierungen. Beim Klenow-Labeling baut das Klenow-Fragment konjugierte Nukleotid-Analoga durch Strang-Verlängerung von Zufallsprimern in die DNA ein. Bei indirekten Markierungsverfahren werden, anstelle direkt-konjugierter Nukleotide, Dig- bzw. Bio-konjugierte Nukleotide in einer PCR in die DNA eingebaut. Diese werden anschließend durch eine anti-Dig-Cy5 und Streptavidin-Cy3 Färbung sichtbar gemacht. Letzteres Verfahren bietet den Vorteil einer höheren Einbaurrate, allerdings den Nachteil von zusätzlichen Färb- und Waschschrritten. Wie in Tabelle 10 gezeigt, waren das Klenow-Labeling und die indirekte Markierung dennoch schlechter, es wurde deshalb auch weiterhin das direkte Markierungsverfahren mittels Thermosequenase verwendet.

Tabelle 10 Vergleich verschiedener Labeling-Verfahren

	Thermosequenase-Labeling	Klenow-Labeling	indirektes Labeling
% richtig +	92,9	50	0
% falsch +	0	0	0
% richtig -	100	100	100
% falsch -	7,1	50	100

3.1.6 Charakterisierung des Hybridisierungsverhaltens des 3K BAC-Arrays

Nachdem die Parameter zur Herstellung des 3K BAC-Arrays sowie die Hybridisierungsbedingungen optimiert worden waren, wurde das Hybridisierungsverhalten des 3K BAC-Arrays charakterisiert. Die Charakterisierung des Hybridisierungsverhaltens beinhaltete die Untersuchung 1.) des Hybridisierungsverhaltens von nicht-amplifizierter DNA, 2.) der Hybridisierungsintensität der einzelnen Klone, 3.) des Auftretens von Ausreißern und 4.) des Hybridisierungsverhaltens von ganzen Chromosomen. Die Analyse der obigen Parameter erlaubt es, eine Aussage über die Zuverlässigkeit der Hybridisierungsdaten treffen zu können, und dient somit als Grundlage für die Beurteilung und Einordnung der mit dem 3K BAC-Array erzielten Ergebnisse.

Untersuchung des Hybridisierungsverhaltens nicht-amplifizierter DNA

Im nächsten Schritt sollte das Hybridisierungsverhalten des 3K BAC-Arrays mit den optimierten Hybridisierungsbedingungen getestet werden. Um zu zeigen, dass der 3K BAC-Array als Hybridisierungsplattform mit anderen publizierten BAC-Arrays vergleichbar ist, wurde zunächst mit normaler, männlicher und weiblicher nicht-amplifizierter DNA hybridisiert (Abbildung 14).

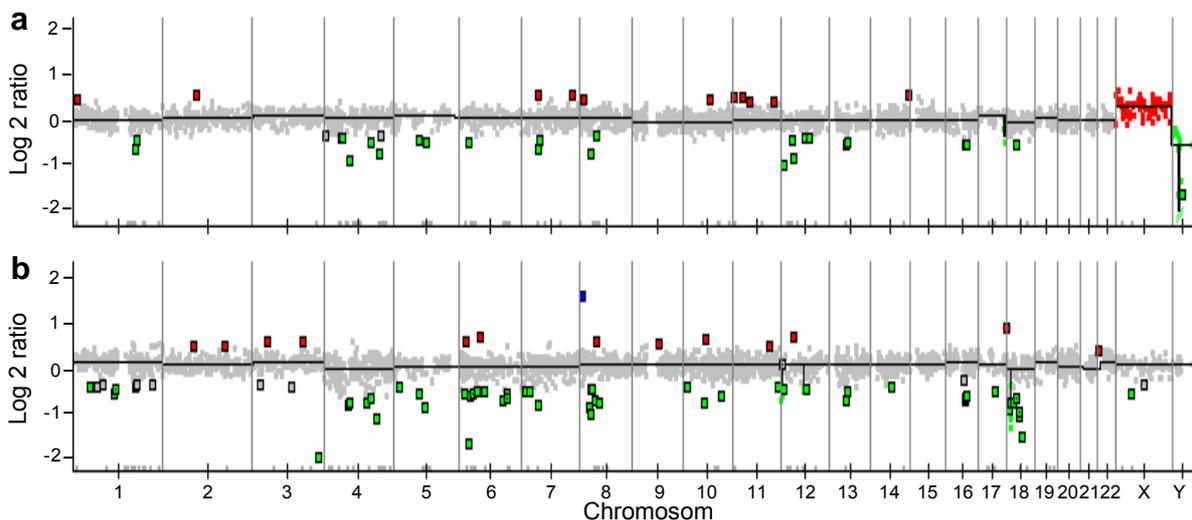


Abbildung 14 Hybridisierung von normaler, nicht-amplifizierter DNA auf den 3K BAC-Array

Der 3K BAC-Array wurde mit der nicht-amplifizierten DNA einer gesunden Probandin (a) und eines gesunden Probanden (b) hybridisiert. Dargestellt ist das Log₂-Verhältnis (Log₂Ratio) von Probe zu Referenz gegen die genomische Lokalisation. Graue senkrechte Linien trennen Chromosomen voneinander. Die schwarze durchgezogene Linie ist die Glättungslinie, die signifikante Aberrationen anzeigt. Die Farbe der Punkte zeigt an, ob ein Klon bei der Auswertung als normal (grau), deletiert (grün), gewonnen (rot) oder amplifiziert (blau) bestimmt wurde. Umrahmte Klone sind so als Ausreißer markiert und nicht vertrauenswürdig.

Als Referenz-DNA diente eine männliche, nicht-amplifizierte Normal-DNA, so dass das

zusätzliche X-Chromosom und fehlende Y-Chromosom als Gewinn bzw. Verlust detektiert wird, wenn weibliche Test-DNA verwendet wird. Alle Autosomen werden erwartungsgemäß als balanciert angezeigt. Beide Experimente wurden in Duplikaten durchgeführt, die Korrelationskoeffizienten betragen 0,83 bzw. 0,91 für die Hybridisierung der männlichen und weiblichen Test-DNA. Somit konnte gezeigt werden, dass sich der 3K BAC-Array grundsätzlich als Hybridisierungsplattform eignet.

Untersuchung der Hybridisierungsintensität der BAC-Klone

Die Untersuchung der Hybridisierungsintensität der BAC-Klone und die Untersuchung von Ausreißern auf den verschiedenen Chromosomen sollte Aufschluss darüber liefern, ob während der Präparation oder Amplifikation der BAC-DNA systematische Fehler unterlaufen waren. Die durchschnittliche Signalintensität lag mit 17.129 ± 906 rfu bei einem in der Literatur beschriebenen Wert (ca. 20.000 rfu)⁸⁶. Bei niedriger Gesamthybridisierungsintensität wurde der Anteil solcher BAC-Klone höher, die im Referenzkanal (männliche Pool-DNA) ein Signal-zu-Hintergrund-Verhältnis < 2 hatten, also von solchen Klonen, die nicht mit in die Auswertung einbezogen wurden (siehe 2.1.12). Die Signalintensität der Hintergrundfluoreszenz betrug in der Regel ca. 2.000 rfu (relative Fluoreszenzeinheit), Klone mit Signalintensitäten von < 4.000 rfu im Referenzkanal wurden deshalb nicht ausgewertet. Die Signal-zu-Hintergrund-Verhältnisse der übrigen auswertbaren Klone variierten zwischen 2 und 40. Abhängig von der Gesamthybridisierungsintensität und den Hybridisierungsartefakten waren im Mittel 86,5% (n=20 Hybridisierungen) der 2.940 BAC-Klone auswertbar, die Spannweite hierbei reichte von 76% bis 95%. Um die Verteilung von schwach hybridisierenden BAC-Klonen (Signalintensität < 4.000 rfu) auf den verschiedenen Chromosomen zu untersuchen, wurden vier Hybridisierungen ausgewertet, die eine normale Gesamthybridisierungsintensität ($89,5 \pm 2,5\%$ der Klone waren auswertbar) zeigten (Abbildung 15). Das Y-Chromosom zeigte mit 23,8% die höchste Rate von schwach hybridisierenden Klonen, das Chromosom 21 mit 0% die niedrigste Rate. Im Schnitt wiesen 9,7% der Klone eine zu schwache Hybridisierung auf, um in die Auswertung aufgenommen zu werden. Die übrigen 0,8% der Klone waren aufgrund von Hybridisierungsartefakten nicht auswertbar.

Es konnte festgestellt werden, dass oft die gleichen BAC-Klone eine niedrige Signalintensität, also eine schwache Hybridisierung aufwiesen. Es wurde deshalb überprüft, ob zwischen Klonen mit hoher und niedriger Signalintensität Unterschiede im Laufverhalten

bei der pulsed-field-Gelelektrophorese oder in der Effizienz der Amplifikations-PCR festzustellen waren. Da ganze Chromosomen zum Teil schlechter hybridisierten als andere und deren Präparation blockweise erfolgt war, wurden auch hier Zusammenhänge gesucht. In beiden Fällen konnten keine systematischen Zusammenhänge mit dem Hybridisierungsverhalten auf dem 3K BAC-Array gefunden werden.

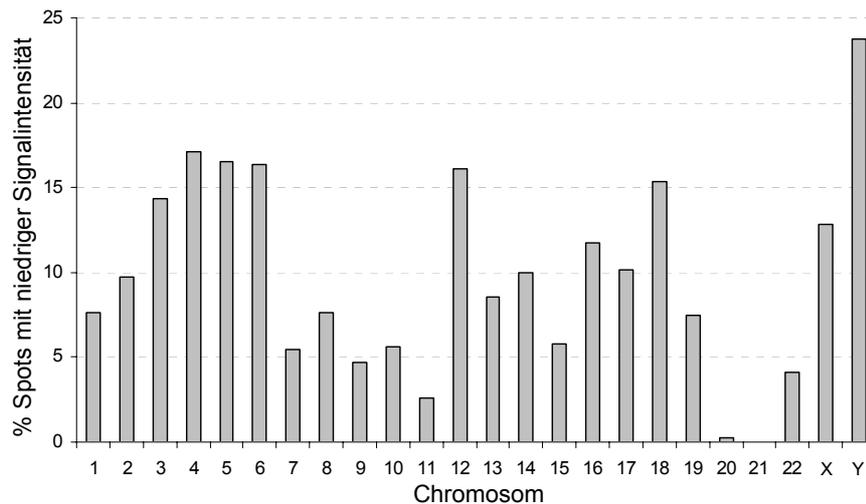


Abbildung 15 Verteilung der BAC-Klone mit schwacher Signalintensität

Die Verteilung von BAC-Klonen in den einzelnen Chromosomen mit Signal-zu-Hintergrund-Verhältnissen von < 2 im Referenzkanal (männliche Pool-DNA) wurde in drei Hybridisierungen von einzelnen Tumorzellen (weiblich) und einer weiblichen Normalzelle ausgewertet.

Untersuchung der Anzahl und Verteilung von Ausreißern

Zusätzlich wurden die Anzahl und Verteilung von Ausreißern auf den verschiedenen Chromosomen ermittelt. Für die Auswertung wurden insgesamt 12 Experimente untersucht. Ausreißer werden von der CAPweb-Auswertungssoftware ermittelt und stellen Klone dar, die aufgrund ihres Referenz-zu-Probe-Verhältnisses im Kontext der benachbarten Klone als unzuverlässig gelten (2.1.12). In Abbildung 16 ist die Verteilung von Ausreißern über die verschiedenen Chromosomen dargestellt. Chromosom 15 zeigt relativ zur Gesamtzahl der Klone eines Chromosoms die geringste Anzahl von Ausreißern (0,34%), Chromosom 22 und Y mit 2,9% die meisten Ausreißer. Die Anzahl und Verteilung von Ausreißern konnte in keinen Zusammenhang mit der Präparation oder anderen Arbeitsschritten gebracht werden (s.o.). Die Anzahl der Ausreißer in verschiedenen Experimenten mit amplifizierter Einzelzell-DNA variierte zwischen 12 (0,4%) und 82 (2,8%).

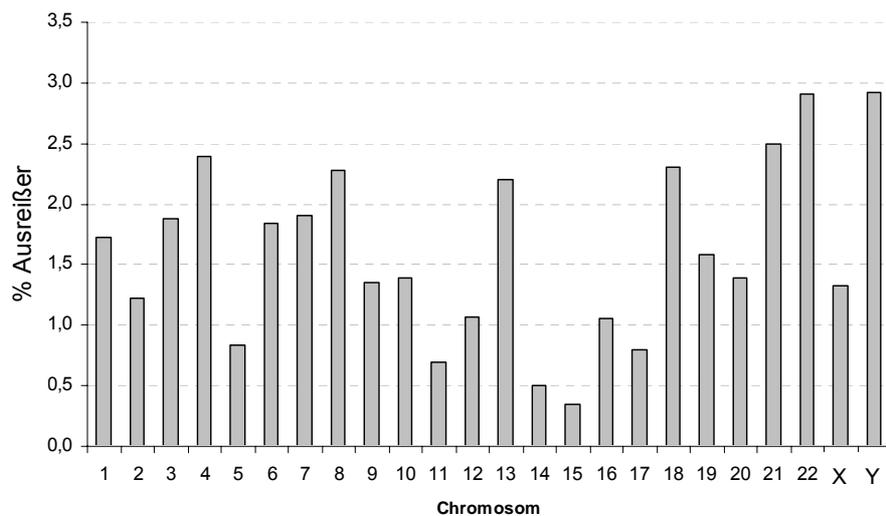


Abbildung 16 Verteilung der Ausreißer

Ausreißer werden durch die Auswertungssoftware ermittelt und stellen solche Klone dar, die als unzuverlässig eingestuft werden. Insgesamt zwölf Hybridisierungsexperimente wurden für diese Auswertung analysiert. Dargestellt ist der prozentuale Anteil von Ausreißern im Verhältnis zur Gesamtzahl an Klonen auf einem Chromosom (gemittelt für zwölf Experimente).

Zusätzlich zur Ermittlung der Verteilung von Ausreißern wurde untersucht, ob in Hybridisierungen von normalen Einzelzellen verschiedener Probanden immer wieder dieselben Klone als Ausreißer angezeigt wurden und es sich deshalb um reproduzierbar unzuverlässig hybridisierende Klone handelt, die aus den weiteren Auswertungen ausgeschlossen werden müssen. Insgesamt 14 Klone wurden in mindestens 40% der untersuchten Experimente als Ausreißer erkannt (Tabelle 11) und deshalb aus weiteren Untersuchungen ausgeschlossen.

Tabelle 11 Wiederkehrende Ausreißer in 12 unabhängigen Experimenten

BAC-Klon	Chromosom	% der Experimente als Ausreißer gewertet
449d11A	1	41.7
130b18	1	75.0
91l23	2	41.7
91m15	3	83.3
80d22	4	50.0
79m4	4	66.7
79j7	4	91.7
80l16	6	75.0
90c13	7	50.0
79o21	7	50.0
451o18A	8	50.0
91m2	9	75.0
90g13	18	41.7
15e10	21	41.7

GC-Gehalt und *Mse*I-Fragmente – das Chromosom 19-Artefakt

In wiederholten Hybridisierungsexperimenten mit normaler, amplifizierter Einzelzell-DNA wurde das Chromosom 19 als deletiert angezeigt. Auch in den meisten Hybridisierungen von Tumoreinzelzell-Amplifikaten (14 von 15) wurde dieses Artefakt beobachtet. Es wurde untersucht, ob für das auftretende Phänomen eine Erklärung und somit ein Lösungsansatz gefunden werden konnte. Eine mangelhafte Qualität der BAC-Präparationen für das Chromosom 19 könnte die Ursache des Artefakts sein. Dagegen spricht, dass das Deletions-Artefakt bei Hybridisierungen mit nicht-amplifizierter, genomischer DNA (siehe Abbildung 22a und X) nicht beobachtet wurde.

Aus der Sequenzierung des humanen Genoms ist bekannt, dass Chromosom 19 einen höheren GC-Gehalt als andere Chromosomen hat. Dies könnte eine geringere Frequenz von *Mse* I-Schnittstellen ($T^{\nabla}TAA$) und damit längere *Mse* I-Fragmente zur Folge haben. Deshalb wurde untersucht, ob sich die BAC-Klone von Chromosom 19 ($n=33$) hinsichtlich ihres GC-Gehalts und ihrer *Mse* I-Fragmentlänge von BAC-Klonen unterscheiden, die zuverlässig einen erwarteten Karyotyp darstellen (Abbildung 17). 31 BAC-Klone des Chromosoms 3 wurden hierfür ausgewählt.

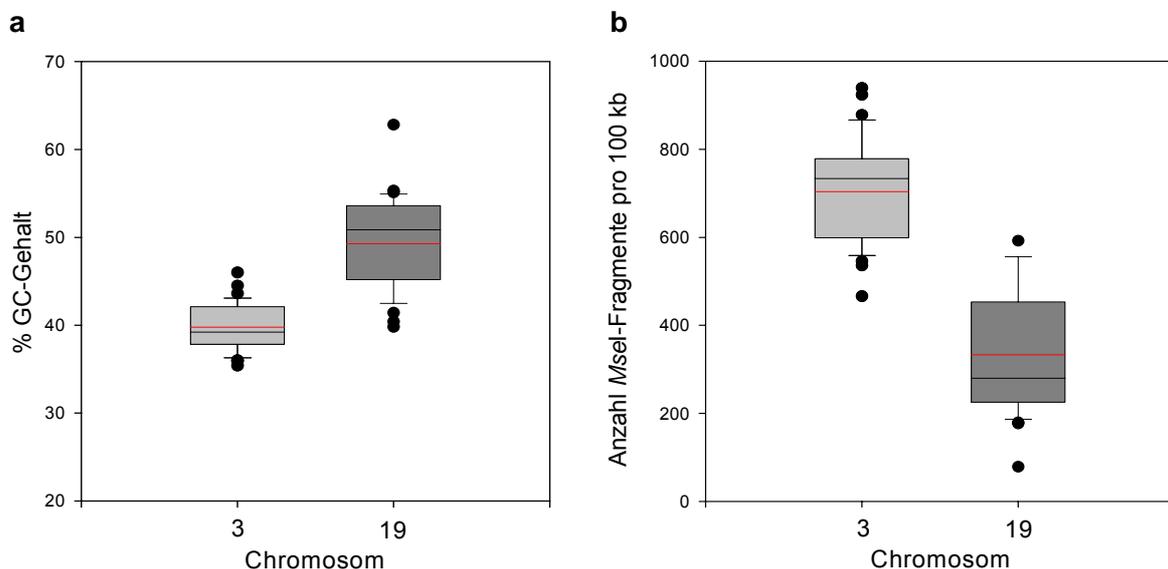


Abbildung 17 GC-Gehalt und Anzahl der *Mse*I-Fragmente von Chromosom 3- und 19-BAC-Klonen

(a) GC-Gehalt in % von 31 BAC-Klonen des Chromosom 3 und 33 BAC-Klonen des Chromosom 19. (b) Anzahl der *Mse* I-Fragmente pro 100 kb in denselben BAC-Klonen wie in (a). Die waagerechte, schwarze Linie zeigt den Mittelwert an, die rote Linie das arithmetische Mittel. $p=0,0001$ (Student's *t*-Test).

Der durchschnittliche GC-Gehalt der Klone von Chromosom 3 lag bei 39,8%, der der Klone von Chromosom 19 hingegen bei 49,3% (*t*-Test: $p \leq 0,0001$) (Abbildung 17a). Die

durchschnittliche Anzahl der Mse I-Fragmente lag bei 704 pro 100 kb für Chromosom 3 und bei 341 Schnittstellen pro 100 kb für Chromosom 19 (t -Test: $p \leq 0,0001$) (Abbildung 17b). Die Korrelation von GC-Gehalt und Anzahl der Mse I-Fragmente pro 100 kb wurde berechnet, um festzustellen, ob es einen Zusammenhang zwischen beiden gibt. Die Korrelation zwischen beiden Datensätzen ist in Abbildung 18 als Scatter-Plot dargestellt. Durch die Datenpunkte wurde eine lineare Regressionsgerade gelegt, aus der sich ein Korrelationskoeffizient von 0,97 (Bestimmtheitsmaß $R^2=0,95$) berechnen ließ. Daraus konnte gefolgert werden, dass beide Datensätze eng korreliert sind. Aus der angegebenen Formel ($y = -35,519x + 2.101,2$) kann aus dem GC-Gehalt in % (x) die Anzahl der Mse I-Fragmente pro 100 kb (y) annäherungsweise berechnet werden. Aus der Anzahl der Mse I-Fragmente lässt sich die durchschnittliche Mse I-Fragmentlänge berechnen: 142 bp für Chromosom 3 und 293 bp für Chromosom 19 (t -Test: $p \leq 0,0001$). BAC-Klone des Chromosoms 19 enthalten also im Schnitt ca. doppelt so lange Mse I-Fragmente wie BAC-Klone des Chromosoms 3.

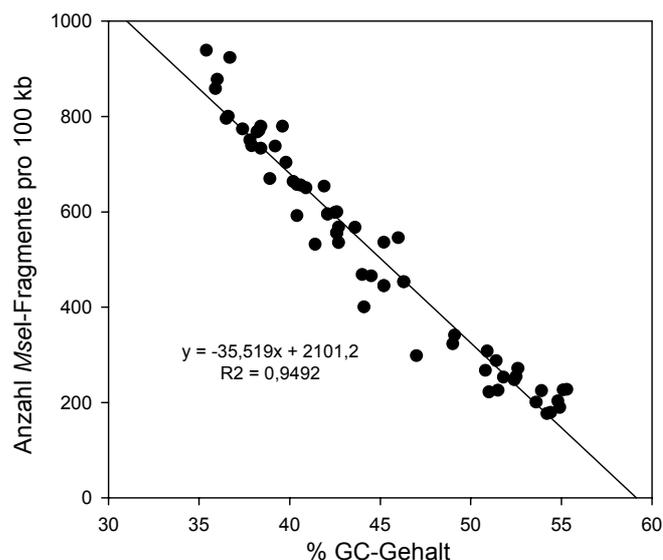


Abbildung 18 Korrelation der Anzahl der Mse I-Fragmente mit dem GC-Gehalt in BAC-Klonen von Chromosom 3 und 19

Die Korrelation zwischen der Anzahl der Mse I-Fragmente pro 100 kb und dem GC-Gehalt (in %) in BAC-Klonen von Chromosom 3 und 19 ist als Scatter-Plot dargestellt. Es wurde eine lineare Regressionsgerade durch die Datenpunkte gelegt. Das Bestimmtheitsmaß R^2 wurde auf 0,95 berechnet. Mit der angegebenen Formel ($y = -35,519x + 2101,2$) kann die Anzahl der Mse I-Fragmente pro 100 kb (y) aus dem % GC-Gehalt (x) berechnet werden.

Aus den BAC-Klonen des Chromosoms 19 wurden diejenigen ausgewählt, deren Mse I-Fragmentlängen denen von BAC-Klonen des Chromosoms 3 am nächsten kamen. Anhand dieser Klone sollte untersucht werden, ob derart selektionierte BAC-Klone ein besseres

Hybridisierungsverhalten zeigen. Es wurde eine Auswertung von Hybridisierungsdaten vorgenommen, bei denen für das Chromosom 19 nur solche BAC-Klone analysiert wurden, deren Anzahl von Mse I-Fragmenten mit denen von Chromosom 3 vergleichbar waren (> 500 pro 100 kb; Abbildung 17b). In Abbildung 19 ist exemplarisch für eine Hybridisierung einer Einzelzell-DNA eines Patienten mit Trisomie 21 eine vergleichende Auswertung gezeigt. In Abbildung 19a wurden alle Chromosom 19-BAC-Klone und in Abbildung 19b nur die Chromosom 19-BAC-Klone mit mehr als 500 Mse I-Schnittstellen pro 100 kb in die Auswertung einbezogen. Trotz der geringen Klonzahl ($n=6$), die in Abbildung 19b für das Chromosom 19 zur Verfügung steht, wird das Chromosom 19 als normal angezeigt. Daraus kann gefolgert werden, dass BAC-Klone, die durch vorwiegend kurze Mse I-Fragmente von unter 200 bp repräsentiert werden, gute Hybridisierungseigenschaften besitzen. In allen weiteren gezeigten Experimenten und Abbildungen wird das Chromosom 19 nicht berücksichtigt, da die Klonzahl zu gering ist, um eine Auswertung mit ausgewählten BAC-Klonen durchzuführen.

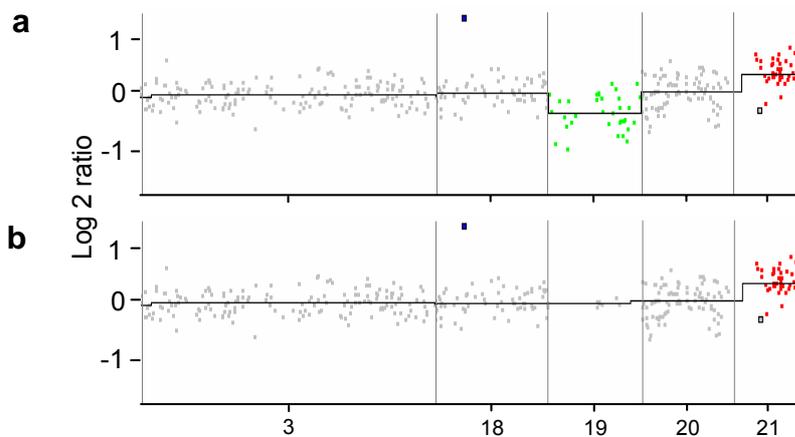


Abbildung 19 Auswertung der Hybridisierung einer Trisomie 21-Einzelzell-DNA

In (a) wurde die Auswertung der Hybridisierung einer Einzelzell-DNA eines Patienten mit Trisomie 21 unter Berücksichtigung aller BAC-Klone des Chromosoms 19 vorgenommen. In (b) wurden nur solche Chromosom 19-BAC-Klone verwendet, deren Anzahl von Mse I-Schnittstellen mit denen von Chromosom 3 vergleichbar war. Zur besseren Visualisierung sind ausschnittsweise nur die Chromosomen 3, 18, 19, 20 und 21 dargestellt. (Darstellung wie in Abbildung 14)

3.1.7 Validierung der Hybridisierungsbedingungen durch Einzelzell-Analysen

Als nächstes sollten Einzelzellen mit bekanntem Karyotyp hybridisiert werden. Eine weibliche und eine männliche, normale Einzelzell-DNA und die amplifizierte DNA aus einer Einzelzelle eines Patienten mit Trisomie 21 wurden gegen einen Referenz-DNA-Pool eines gesunden männlichen Probanden hybridisiert (Abbildung 20). Die beiden Profile der

normalen Einzelzellen sind balanciert (Abbildung 20a und b). Bei der weiblichen Einzelzelle (Abbildung 20a) sind das zusätzliche X-Chromosom und das fehlende Y-Chromosom als Gewinn bzw. Verlust zu erkennen, die sich durch die Hybridisierung gegen eine männliche Referenz-DNA ergeben. Bei der kleinen, als deletiert angezeigten Region auf dem Chromosom 4 des weiblichen Normalzell-Profiles (3,7 Mb) handelt es sich um ein falsch-positiv detektiertes Artefakt. Bei der Hybridisierung der Einzelzell-DNA des Patienten mit Trisomie 21 ist eindeutig die zusätzliche Kopie des Chromosoms 21 dargestellt, alle weiteren Autosomen sind, bis auf falsch-positive Artefakte auf den Chromosomen 8 (0,7 Mb Deletion) und 12 (0,8 Mb Gewinn), normal (Abbildung 20c). Das Chromosom 19 wird nicht in die Auswertungen einbezogen (siehe 3.1.6). Aus diesen Ergebnissen kann abgeleitet werden, dass Profile mit bekannten Karyotypen geringer Komplexität auch in Einzelzellen mit dem 3K BAC-Array korrekt dargestellt werden können.

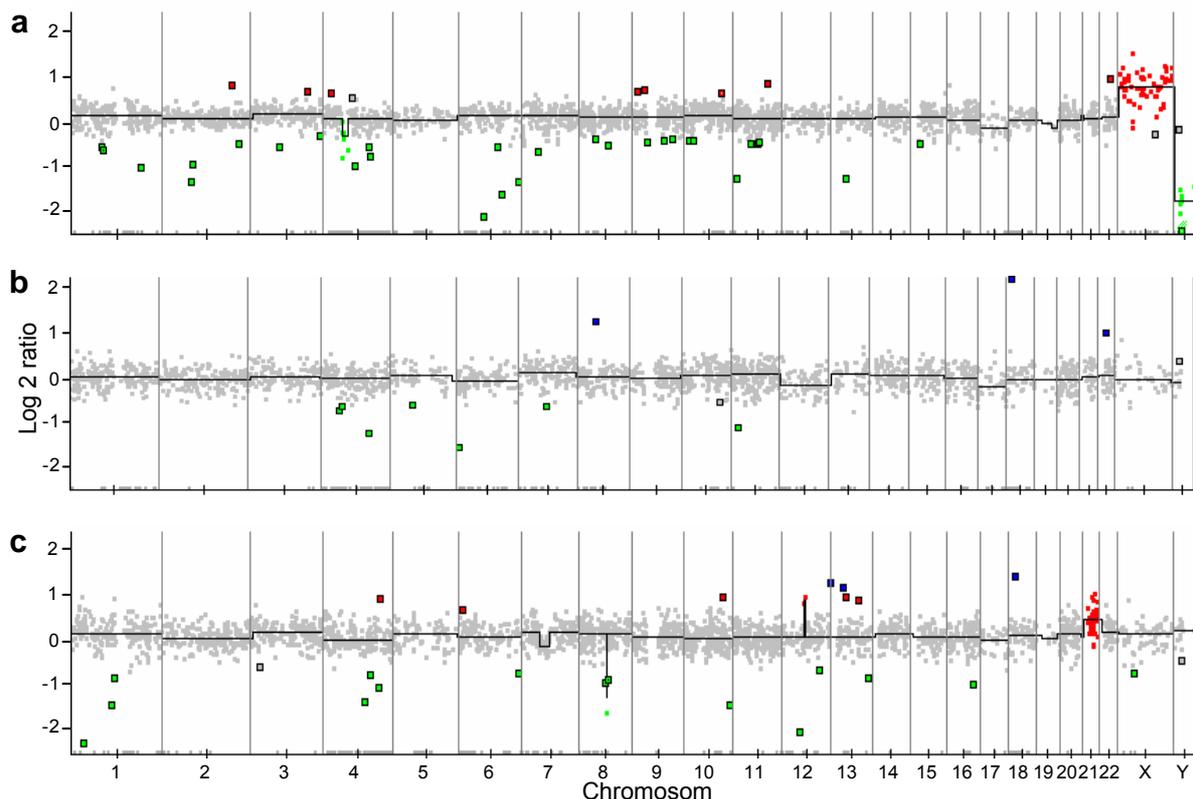


Abbildung 20 Hybridisierungsprofile normaler Einzelzellen und einer Trisomie 21 Einzelzelle

Hybridisierungsprofil einer normalen, weiblichen (a) und männlichen (b) Einzelzelle gegen eine männliche Referenz-DNA. (c) Hybridisierungsprofile einer Einzelzell-DNA eines männlichen Patienten mit Trisomie 21. (Darstellung wie in Abbildung 14; Chromosom 19 wird nicht berücksichtigt, siehe 3.1.6)

BAC-Klone mit nicht-erwartungsgemäßem Hybridisierungsverhalten und Streuung des 3K BAC-Arrays

Die Bestimmung der Rate an BAC-Klonen, die nicht den Erwartungen entsprechend hybridisieren, liefert ein Maß für die Zuverlässigkeit der Hybridisierungsdaten. In Hybridisierungen von Einzelzellen mit bekannten Karyotypen geringer Komplexität konnten solche BAC-Klone identifiziert werden. In fünf unabhängigen Experimenten betrug die Rate an nicht-erwartungsgemäß hybridisierenden BAC-Klonen 0,09% (ohne Berücksichtigung von BAC-Klonen des Chromosoms 19). Klone oder Regionen, die auch in Normalzellen wiederholt Alterationen anzeigen, werden als unzuverlässig bewertet und können in Experimenten mit Patientenproben nicht berücksichtigt werden.

Die Berechnung der Standardabweichung der Hybridisierungsdaten, also der experimentellen Streuung, liefert ein indirektes Maß für die Genauigkeit, mit der Aberrationen angezeigt werden können. Bei einer großen Streuung der Daten ist es z.B. unwahrscheinlich, dass kleine Aberrationen mit geringer Auslenkung detektiert werden können. In fünf Experimenten mit normaler Einzelzell-DNA und einem Experiment mit einer Einzelzelle eines Patienten mit Down Syndrom betrug die durchschnittliche Streuung (Standardabweichung) der balancierten BAC-Klone $-0,02 \pm 0,22$ (Log₂-Verhältnis Probe zu Referenz). In vier Hybridisierungsexperimenten mit nicht-amplifizierter Normal-DNA betrug die Standardabweichung der balancierten Klone $-0,01 \pm 0,15$. Die Reproduzierbarkeit von Hybridisierungsexperimenten auf dem 3K BAC-Array wurde bestimmt, indem Hybridisierungen von amplifizierter DNA aus Einzelzellen in Duplikaten durchgeführt wurden. Der durchschnittliche Korrelationskoeffizient betrug $0,84 \pm 0,1$ ($n=3$).

Assay zur Vorhersage einer erfolgreichen Array-CGH-Analyse

Im Verlauf der Validierungsexperimente mit normalen Einzelzellen fiel auf, dass einige dieser Zellen trotz unauffälliger Metaphasen-CGH-Profile (Abbildung 21a und b) stark streuende Array-CGH-Profile aufwiesen, die nicht den erwarteten Karyotyp darstellten (Abbildung 21c und d). Diese Profile waren nicht durch eine mangelhafte Hybridisierung bedingt, wie ein Ausschnitt des hybridisierten Arrays zeigt (Abbildung 21e und g). Deshalb wurde vermutet, dass der Grund für stark streuende Hybridisierungsprofile eine geringe Qualität der hybridisierten Einzelzell-DNA war. Bei Amplifikationen von mikrodissektiertem Paraffingewebe konnte bereits gezeigt werden, dass ein Zusammenhang zwischen geringer DNA-Qualität und dem Fehlen von langen Mse I-Fragmenten besteht⁸⁸. Des Weiteren wurde

zeigt, dass lange Fragmente in einem Gemisch verschieden langer Fragmente in einer PCR schlechter amplifiziert werden⁸⁹. Deshalb wurde überprüft, ob lange Mse I-Fragmente in den schlecht hybridisierenden Einzelzell-Amplifikaten vorlagen.

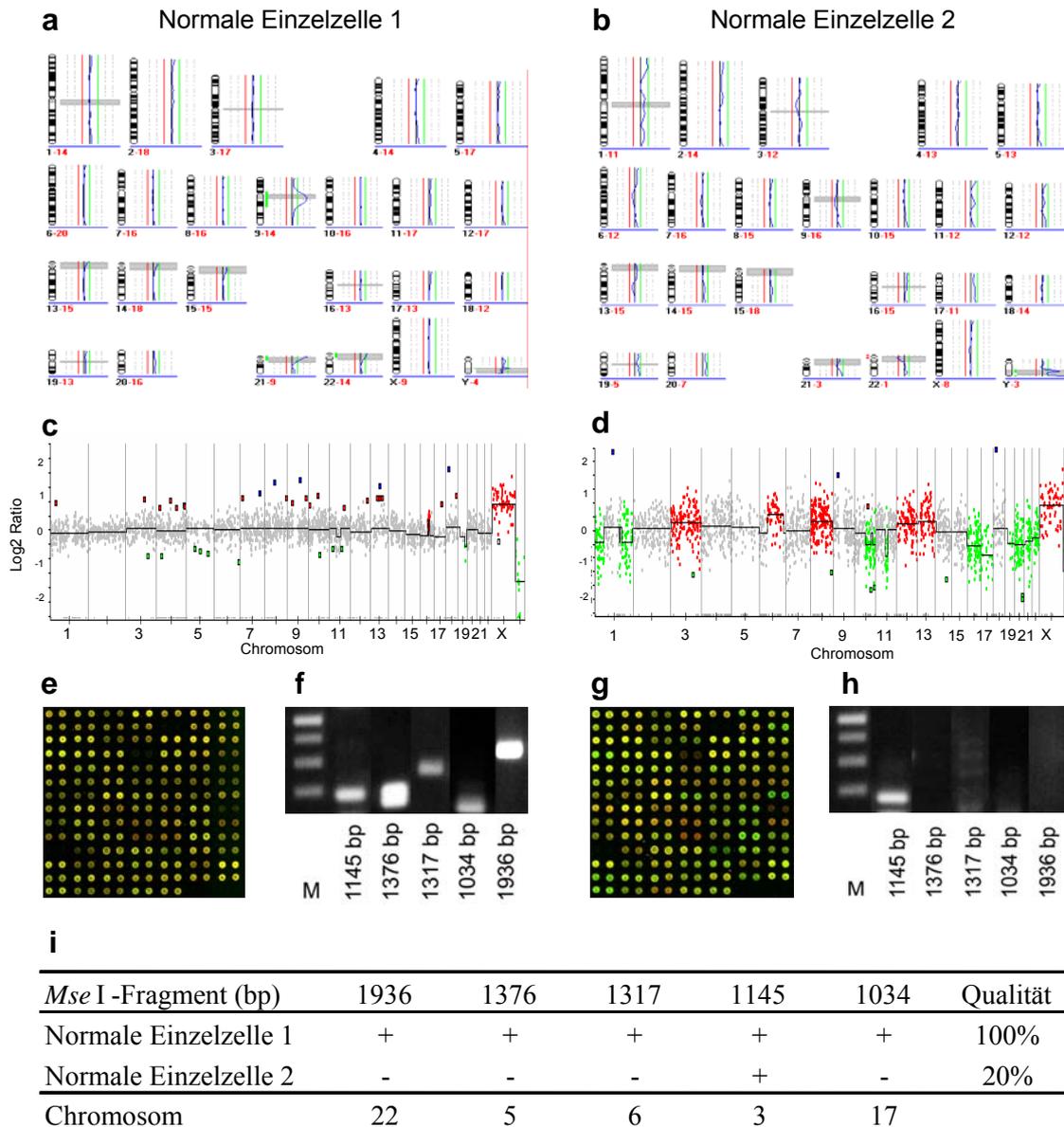


Abbildung 21 Assay zur Vorhersage einer erfolgreichen Array-CGH-Analyse

Beispielhaft dargestellt sind die Untersuchungen von zwei normalen Einzelzellen. (a) und (b) Metaphasen-CGH-Profile der beiden Einzelzell-Amplifikate, (c) und (d) Array-CGH-Profile derselben Proben, (e) und (g) Ausschnitte aus den Hybridisierungen des 3K BAC-Arrays zu den in (c) und (d) gezeigten Profilen, (f) und (h) das Ergebnis für die PCR von fünf verschiedenen Markern, mit denen das Vorliegen langer Mse I-Fragmente nachgewiesen wird. In (i) sind die Ergebnisse dieser PCR-Untersuchung zusammengefasst. (Darstellung der Array-CGH wie in Abbildung 14; Chromosom 19 wird nicht berücksichtigt, siehe 3.1.6)

Hierfür wurden fünf Primer verwendet, mit denen gezielt lange Mse I-Fragmente von mehr als 1000 bp Länge in einer PCR nachgewiesen werden konnten. Je ein Einzelzell-Amplifikat,

mit dem der erwartete Karyotyp einmal dargestellt und einmal nicht dargestellt werden konnte (Abbildung 21), wurde auf die Nachweisbarkeit dieser fünf Fragmente hin untersucht. In dem Einzelzell-Amplifikat, das ein stark streuendes Array-CGH-Profil zeigte, war nur eins der untersuchten fünf Fragmente (20%) nachweisbar (Abbildung 21h und i). Im Gegensatz dazu konnten in dem Einzelzell-Amplifikat, mit dem der erwartete Karyotyp dargestellt werden konnte, alle fünf Fragmente (100%) nachgewiesen werden (Abbildung 21f und i). Anschließend sollte überprüft werden, ob sich der Nachweis dieser fünf Mse I-Fragmente zur Vorhersage einer erfolgreichen Array-CGH-Analyse eignet. Hierfür wurden insgesamt 60 normale Einzelzellen und 59 disseminierte Tumorzellen, für die bereits Metaphasen-CGH-Profile vorlagen, auf die Nachweisbarkeit der fünf Marker hin untersucht. Bei einem Nachweis von $\geq 80\%$ der Fragmente, wurde die Probe als hoch-qualitativ beurteilt, bei 60% als mittel- und bei $\leq 40\%$ als nieder-qualitativ. 84% der untersuchten Zellen besitzen eine mittlere bis hohe Qualität (45% = hoch (n = 54), 39% = mittel (n = 46), 16% = niedrig (n = 19)). 20 Zellen aus den drei Qualitätsgruppen wurden anschließend mit Array-CGH untersucht, die Ergebnisse der Array-CGH wurden dann mit der Nachweisbarkeit der langen Mse I-Fragmente korreliert. Ein Array-CGH-Experiment wurde hierbei als erfolgreich bewertet, wenn das Profil eine geringe Streuung ($\leq 0,3$, siehe 3.1.6) aufwies und das Ergebnis mit dem der Metaphasen-CGH bzw. mit dem bekannten Karyotyp in Übereinstimmung war (Tabelle 12). Bei einer hohen Qualität der Proben war die Array-CGH zu 100% erfolgreich, bei einer schlechten Qualität nie. Der PCR-basierte Assay besitzt eine 90%ige Genauigkeit vorherzusagen, dass eine Probe von mittlerer bis hoher Qualität Erfolg in der Array-CGH haben wird.

Tabelle 12 PCR-basierter Assay zur Vorhersage einer erfolgreichen Array-CGH-Analyse

Mse I-Fragment	Erfolg (%)	erfolgreiche Array-CGH	erfolglose Array-CGH	n
hohe Qualität	100	11*	0**	11
mittlere Qualität	66	4*	2*	6
niedrige Qualität	0	0**	3**	3
gesamt		15	5	20

Genauigkeit = 90%, *Sensitivität = 88%, **Spezifität = 100%

Untersuchung komplexer Genome mit dem 3K BAC-Array

Als nächster Schritt sollte untersucht werden, ob auch komplexe Karyotypen mit zahlreichen genomischen Umlagerungen, wie sie in Tumorzellen vorkommen, in Einzelzellen korrekt durch den 3K BAC-Array dargestellt werden können. Als Modell für Zellen mit komplexem Karyotyp wurde die T47D-Zelllinie ausgewählt. In der Literatur wurde nicht-amplifizierte DNA aus T47D-Pool-DNA bereits mit Hilfe eines 3K-³⁷ und 32K-BAC-Arrays⁸⁷ charakterisiert. Diese Zelllinie wurde außerdem gewählt, da sie eine geringe Variabilität innerhalb der Population zeigte (C. H. Hartmann, unveröffentlichte Daten). Im Gegensatz zu anderen Mammakarzinom-Zelllinien wie MCF-7 wiesen verschiedene, durch Metaphasen-CGH untersuchte T47D-Einzelzellen den gleichen Karyotyp auf.

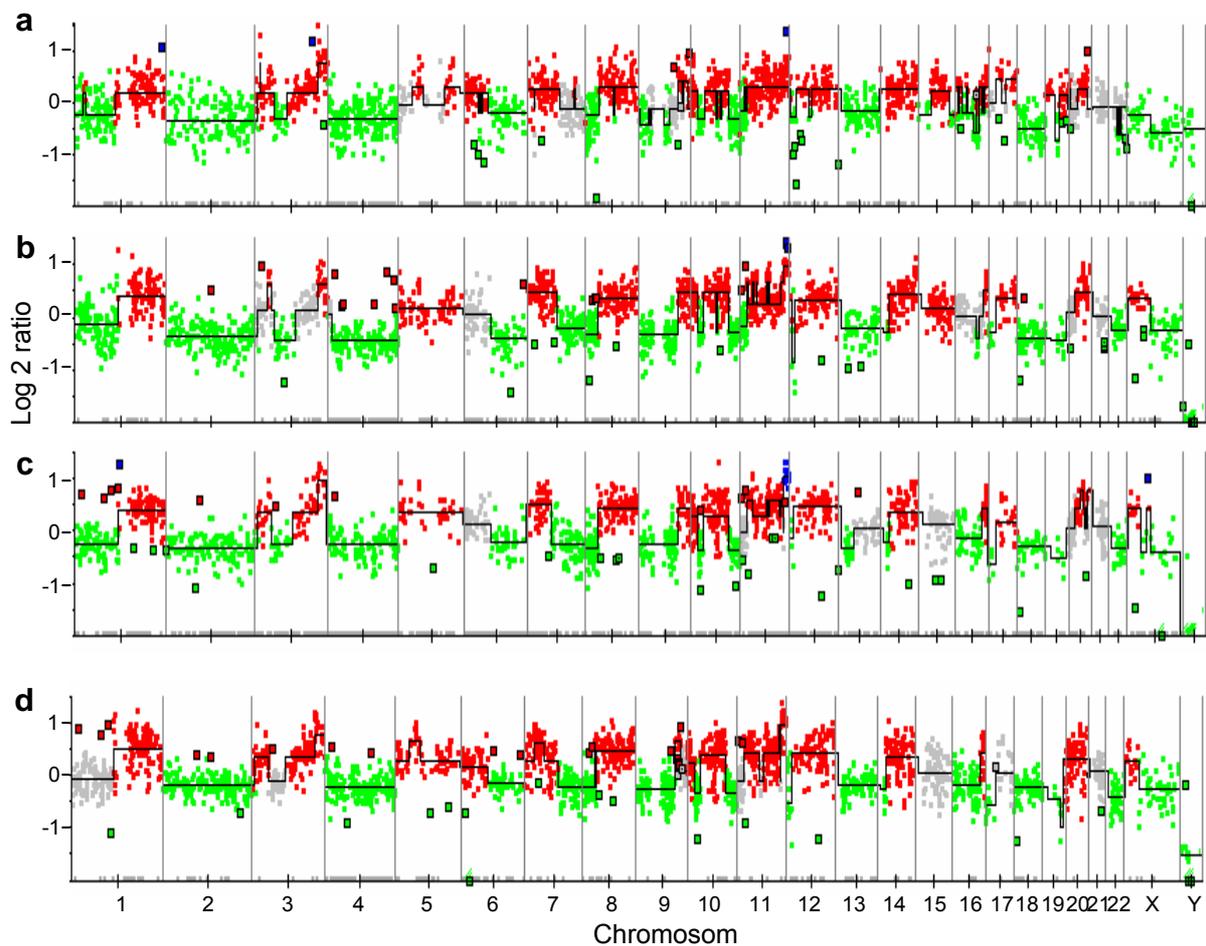


Abbildung 22 Hybridisierungsprofile auf dem 3K BAC-Array von Zellen der T47D-Zelllinie

(a) nicht-amplifizierte Pool-DNA aus ca. 10^6 T47D-Zellen, (b) und (c) amplifizierte Einzelzell-DNAs, (d) amplifizierte DNA aus einem 20-Zell-Pool. (Darstellung der Array-CGH wie in Abbildung 14; Chromosom 19 wird nicht berücksichtigt, siehe 3.1.6)

Um eine korrekte Darstellung des komplexen Karyotyps durch den 3K BAC-Array überprüfen zu können, sollten Analysen von T47D-Einzelzell-Amplifikaten mit Analysen nicht-amplifizierter Pool-DNA aus T47D-Zellen (ca. 10^6 Zellen) verglichen werden. Das Hybridisierungsprofil einer nicht-amplifizierten T47D-Pool-DNA ist in Abbildung 22a dargestellt. Im Vergleich dazu wurde amplifizierte DNA zweier T47D-Einzelzellen (Abbildung 22b und c) und eines Pools von 20 T47D-Zellen (Abbildung 22d) hybridisiert. Eine starke Ähnlichkeit zwischen den Profilen kann aufgrund einer visuellen Inspektion festgestellt werden. Um die Korrelationen quantitativ zu beurteilen, wurden die Korrelationskoeffizienten berechnet (Tabelle 13). Die beiden Hybridisierungsprofile der Einzelzellen und des 20-Zell-Pools ähneln stark dem Profil der genomischen DNA. Es konnte außerdem eine hohe Korrelation zwischen den Profilen der beiden Einzelzellen und dem Profil des 20-Zell-Pools festgestellt werden.

Tabelle 13 Korrelationskoeffizienten von T47D-Hybridisierungsprofilen auf dem 3K BAC-Array

T47D	Korrelationskoeffizient
gDNA vs.EZ1	0,98
gDNA vs.EZ2	0,83
gDNA vs.20-Zell-Pool	0,99
EZ2 vs.EZ1	0,97
EZ1 vs. 20-Zell-Pool	0,98
EZ2 vs.20-Zell-Pool	0,96

Da die T47D-Zelllinie in der Literatur bereits durch die Array-CGH charakterisiert worden war, konnten die mit dem 3K BAC-Array erhaltenen Profile außerdem mit den Literaturdaten verglichen werden. Der Vergleich der Literaturdaten (Hybridisierung einer nicht-amplifizierten, genomischen T47D-DNA auf einem 32K-BAC-Array) mit dem Hybridisierungsprofil einer nicht-amplifizierten, genomischen DNA auf dem 3K BAC-Array ist in Abbildung 23 (jeweils die linken zwei Profile) beispielhaft für die Chromosomen 3 und 10 gezeigt. Die 14,8 und 3,5 Mb großen Deletionen innerhalb des Gewinns auf Chromosom 10 werden auch auf dem 3K-Array dargestellt (Abbildung 23b). Das Aberrationsmuster auf Chromosom 3 ist ebenfalls vergleichbar (Abbildung 23a). Da hier ein 32K-Array mit einem 3K-Array verglichen wird, sind grundsätzlich Unterschiede durch die verschiedene Klonzahl möglich. Zum Vergleich ist das Profil der T47D-Einzelzelle 1 mit in die Abbildung 23 eingefügt, die Ähnlichkeit mit der genomischen DNA zum einen und dem Literatur-Profil zum anderen ist ersichtlich. Mit Hilfe des 3K BAC-Arrays kann demnach auch ein komplexer Karyotyp, wie der der T47D-Zelllinie, in Einzelzell-Amplifikaten korrekt dargestellt werden.

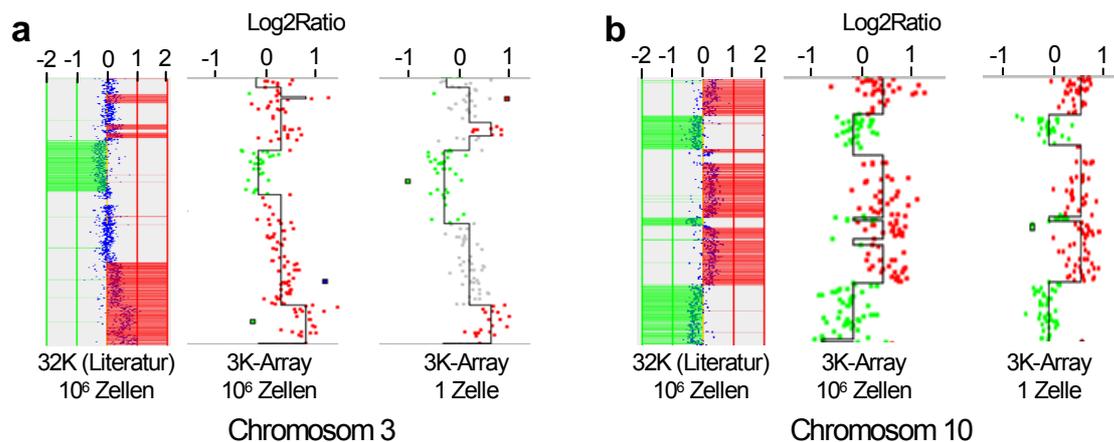


Abbildung 23 Vergleich von Literaturdaten einer Hybridisierung von T47D-DNA mit Hybridisierungen von T47D-DNA auf dem 3K BAC-Array

Dargestellt ist der Vergleich von Hybridisierungsprofilen der T47D-Zelllinie auf dem 3K BAC-Array und auf einem 32K-BAC-Array (Literaturdaten aus Shadeo et al.⁸⁷) ausschnittsweise für die Chromosomen 3 (a) und 10 (b). Auf dem 32K-BAC-Array wurde nicht-amplifizierte DNA aus ca. 10^6 Zellen der T47D-Zelllinie hybridisiert (jeweils auf der linken Seite). Auf dem 3K BAC-Array wurde zum einen nicht-amplifizierte DNA aus ca. 10^6 Zellen der T47D-Zelllinie (Mitte) und amplifizierte DNA aus einer T47D-Einzelle (rechts) hybridisiert. (Darstellung der Array-CGH wie in Abbildung 14) Bei den Hybridisierungsprofilen auf dem 32K-BAC-Array sind in Rot bzw. Grün signifikante Gewinne bzw. Verluste angezeigt.

3.1.8 Test von Oligonukleotid-Arrays für Array-CGH mit Einzelzellen

Im Jahr 2003 wurde das erste Mal beschrieben, dass Oligonukleotid-Arrays, die bis zu diesem Zeitpunkt lediglich für die Untersuchung der Genexpression verwendet wurden, als Hybridisierungsplattform für Array-CGH-Analysen verwendet werden können^{27, 41, 90}. Oligonukleotid-Arrays, die aus synthetischen Oligonukleotiden hergestellt werden, stellen völlig kontaminationslose Hybridisierungsplattformen dar. Deshalb sollte überprüft werden, ob sich solche Arrays für die Untersuchung von Einzelzell-Amplifikaten eignen.

Array-CGH von Einzelzell-Amplifikaten auf 35K Oligonukleotid-Arrays

Für die Experimente wurde der kommerzielle, humane 35K Oligonukleotid-Array der Firma Operon (35K OpArray) verwendet, der 35.035 Oligonukleotide umfasst, von denen 19.592 für eine Array-CGH Untersuchung verwendet werden können, da sich die übrigen Oligonukleotide nur für Expressionsstudien eignen. Zunächst wurden Hybridisierungs- und Waschbedingungen ermittelt, die zu guten Hybridisierungsergebnissen führten. Dabei wurde nicht-amplifizierte, genomische DNA aus der T47D-Zelllinie hybridisiert (Abbildung 24a), da nicht davon ausgegangen werden konnte, dass Hybridisierungen mit amplifizierter Einzelzell-DNA erfolgreich sein würden.

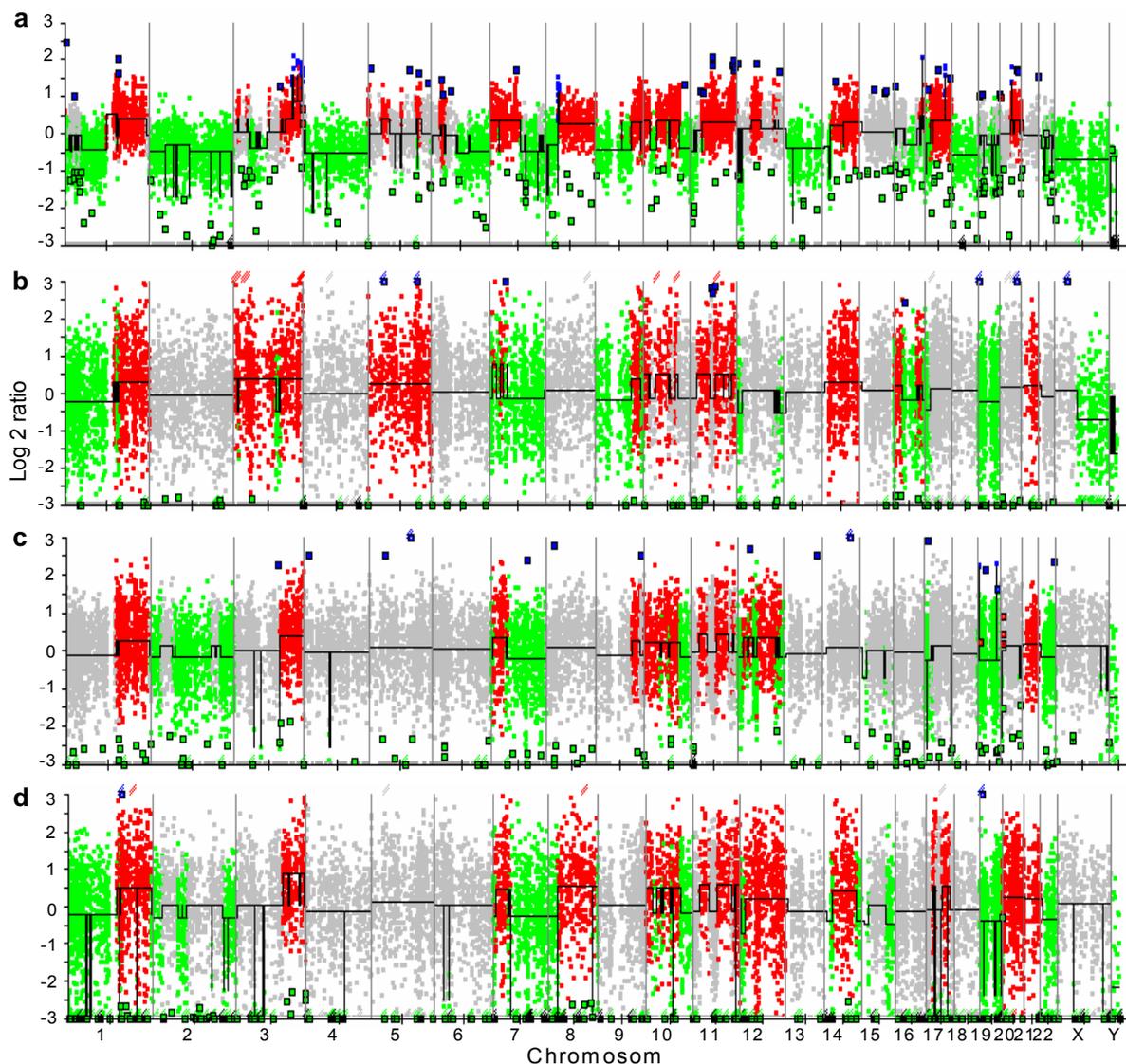


Abbildung 24 Hybridisierung von nicht-amplifizierter, genomischer DNA und von amplifizierter Einzelzell-DNA aus T47D-Zellen auf Operon Oligonukleotid-Arrays

Es wurde auf den 35K Oligonukleotid-Array der Firma Operon hybridisiert. In (a) ist das Hybridisierungsprofil von nicht-amplifizierter, genomischer DNA der T47D-Zelllinie zu sehen, (b) und (c) zeigen Hybridisierungsprofile zweier T47D-Einzelzell-DNAs, (d) zeigt das Hybridisierungsprofil der amplifizierten DNA eines T47D-20-Zell-Pools. (Darstellung der Array-CGH wie in Abbildung 14).

Die T47D-Zelllinie wurde in der Literatur bereits mit 3K-(Ref.³⁷) und 32K-BAC-Arrays⁸⁷ charakterisiert (siehe 3.1.7). Die erhaltenen Profile konnten so mit diesen Literaturdaten verglichen werden, und es konnte überprüft werden, ob Veränderungen unabhängig von der verwendeten Methode identifiziert werden konnten. Die Übereinstimmung mit Literaturdaten ist in Abbildung 25 exemplarisch für die Chromosomen 3 und 10 gezeigt. Sowohl die high-copy Amplifikation am Telomer von Chromosom 3q als auch die 14,8 Mb und 3,5 Mb großen Deletionen innerhalb der Gewinne auf Chromosom 10 konnten dargestellt werden.

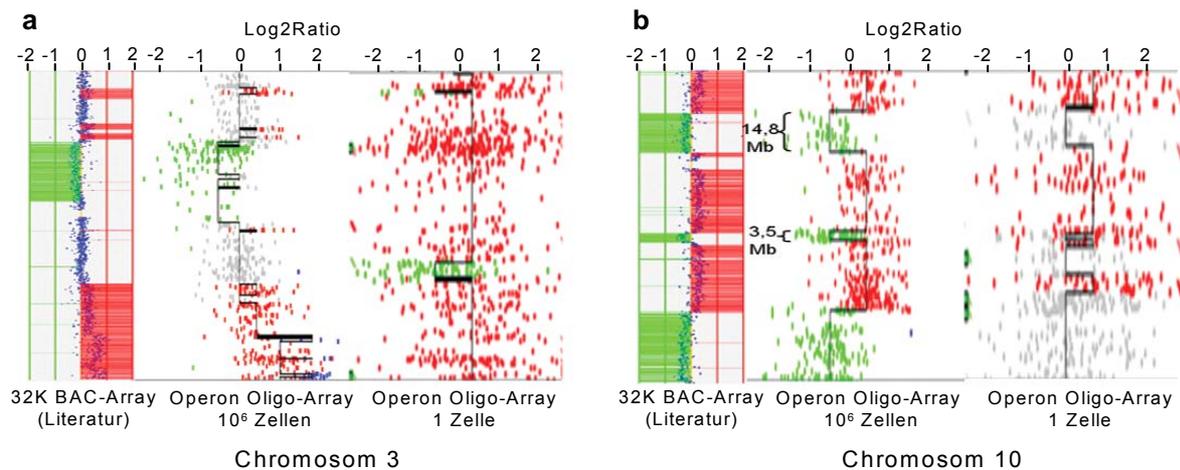


Abbildung 25 Vergleich von Literaturdaten einer Hybridisierung von T47D-DNA mit Hybridisierungen von T47D-DNA auf einem 35K Operon Oligonukleotid-Array

Dargestellt ist der Vergleich von Hybridisierungsprofilen der T47D-Zelllinie auf dem 35K Oligonukleotid-Array und auf einem 32K-BAC-Array (Literaturdaten aus Shadeo et al.⁸⁷), ausschnittsweise für die Chromosomen 3 (a) und 10 (b). Auf dem 32K-BAC-Array wurde nicht-amplifizierte DNA aus ca. 10^6 Zellen der T47D-Zelllinie hybridisiert (jeweils auf der linken Seite). Auf dem 35K Oligonukleotid-Array wurde zum einen nicht-amplifizierte DNA aus ca. 10^6 Zellen der T47D-Zelllinie (Mitte) und amplifizierte DNA aus einer T47D-Einzelzelle (rechts) hybridisiert (Darstellung der Array-CGH wie in Abbildung 14). Bei den Hybridisierungsprofilen auf dem 32K-BAC-Array sind in Rot bzw. Grün signifikante Gewinne bzw. Verluste angezeigt.

Dem Hybridisierungsprofil der nicht-amplifizierten T47D-Pool-DNA ist in Abbildung 24b und c die Hybridisierung zweier T47D-Einzelzell-Amplifikate auf dem 35K-Oligonukleotid-Array gegenüber gestellt. Außerdem wurde die amplifizierte DNA eines T47D-20-Zell-Pools (Abbildung 24d) hybridisiert. Zum einen ist eine starke Streuung des Profils (Log2-Ratios von 2 bis -2) zu erkennen. Zum anderen werden zwar deutliche Aberrationen, wie die Deletion von Chromosom 1p und der Gewinn von 1q dargestellt, kleinere Aberrationen, wie die in Abbildung 25 gezeigten, allerdings nicht. Es wurde außerdem berechnet, wie gut die Hybridisierungsprofile der Einzelzell- und 20-Zell-Pool-Amplifikate mit dem der nicht-amplifizierten DNA korrelieren (Tabelle 14). Die Einzelzell- und 20-Zell-Pool-Experimente korrelieren gut untereinander, die Korrelationskoeffizienten für den Vergleich von Einzelzell-Amplifikaten und nicht-amplifizierter DNA waren allerdings sehr niedrig (vergleiche Tabelle 13). Die Hybridisierungsergebnisse wurden nicht verbessert, wenn statt der DNA einer Zelle DNA von 20 Zellen eingesetzt wurde, was am niedrigen Korrelationskoeffizienten abzulesen war. Aus diesen Ergebnissen kann gefolgert werden, dass der Oligonukleotid-Array der Firma Operon zwar für Array-CGH-Analysen von nicht-amplifizierter genomischer DNA, nicht aber für die Untersuchung von genomischen Veränderungen in Adapter-Linker-PCR amplifizierter Einzelzell-DNA geeignet ist.

Tabelle 14 Korrelationskoeffizienten von T47D-Hybridisierungsprofilen auf 35K Oligonukleotid-Array

T47D	Korrelationskoeffizient
gDNA vs.EZ1	0.19
gDNA vs.EZ2	0.45
gDNA vs.20-Zell-Pool	0.29
EZ2 vs.EZ1	0.75
EZ1 vs. 20-Zell-Pool	0.79
EZ2 vs.20-Zell-Pool	0.77

Array-CGH von Einzelzell-Amplifikaten auf 244K Oligonukleotid-Arrays

Zusätzlich zu dem 35K Oligonukleotid-Array von Operon wurden Einzelzell-Amplifikate auf den 244K Oligonukleotid-Array von Agilent hybridisiert. Bei diesem Array sind alle 244.000 Oligonukleotide für eine Array-CGH-Anwendung optimiert worden. Es wurden die gleichen T47D-Proben hybridisiert, wie auch auf dem 3K BAC-Array und dem Operon 35K Oligonukleotid-Array (Abbildung 26).

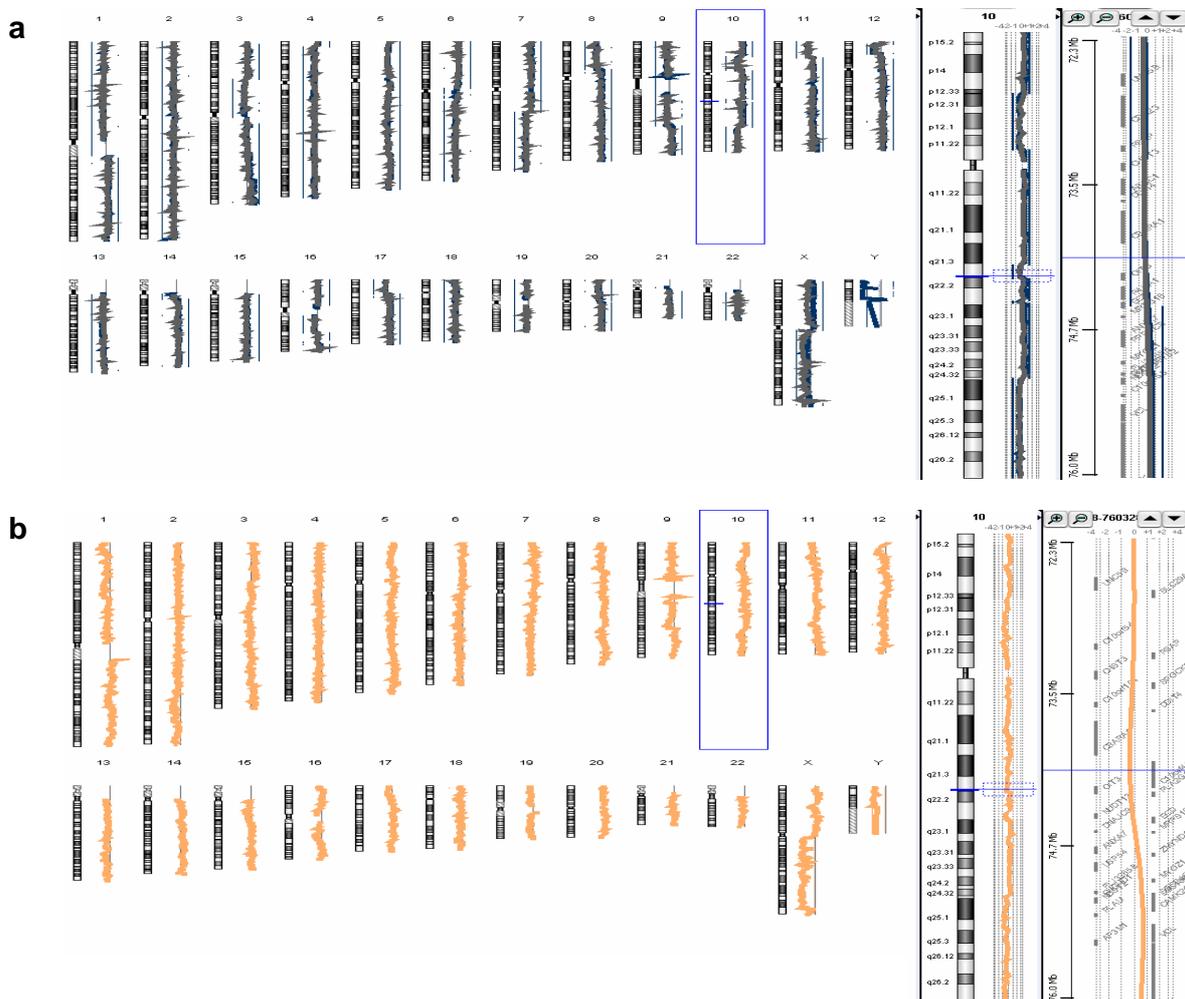


Abbildung 26 Hybridisierung von nicht-amplifizierter, genomischer DNA und von amplifizierter Einzelzell-DNA aus T47D-Zellen auf 244K Agilent Oligonukleotid-Arrays

Es wurde auf den 244K Oligonukleotid-Array der Firma Agilent hybridisiert. In (a) ist das Hybridisierungsprofil von nicht-amplifizierter, genomischer DNA der T47D-Zelllinie zu sehen, (b) zeigt das Hybridisierungsprofil einer amplifizierten T47D-Einzelzell-DNA. Dargestellt ist das Log10-Verhältnis von Probe zu Referenz gegen die genomische Lokalisation für jedes Chromosom. Rechts ist eine Ausschnittsvergrößerung des Chromosoms 10 dargestellt, ganz rechts ist der Ausschnitt eines Bruchpunkts auf 10q gezeigt. Die blaue Linie, die rechts bzw. links der Profile in (a) zu erkennen ist, zeigt signifikante Verluste bzw. Gewinne an. In (b) wurden weder signifikante Gewinne noch Verluste erkannt. Die Darstellung der Daten erfolgte mit der CGH Analytics Software von Agilent.

Bei der nicht-amplifizierten DNA sind die signifikanten Gewinne und Verluste durch Balken links bzw. rechts der Profile angezeigt. In der Ausschnittsvergrößerung der Chromosomen 3 und 10 ist zu erkennen, dass das erhaltene Profil dem Literaturprofil entspricht (Abbildung 27). Beim Hybridisierungsprofil der amplifizierten T47D-Einzelzell-DNA sieht man, dass weder signifikante Gewinne noch Verluste dargestellt werden (Abbildung 27). Trotzdem ist zu erkennen, dass das Hybridisierungsprofil des Einzelzell-Amplifikats eine Struktur aufweist, die der Struktur des Hybridisierungsprofils der nicht-amplifizierten DNA ähnelt. Die Agilent-Hybridisierungsdaten werden im Gegensatz zu den übrigen Darstellungen, die mit der CAPweb-Software erstellt wurden, als Log10-Verhältnisse dargestellt. Bei einer Darstellung als Log2-Verhältnis wäre deshalb die wesentliche größere Streuung der Daten ersichtlicher.

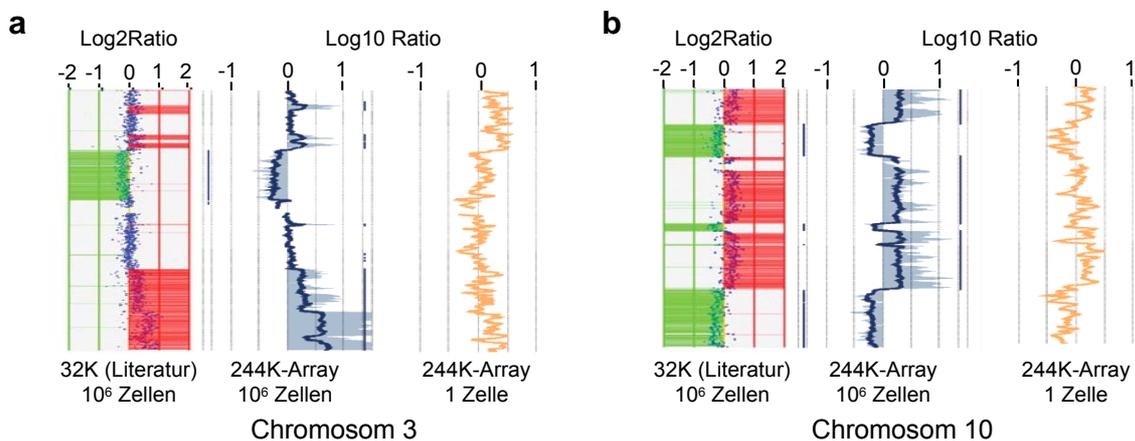


Abbildung 27 Vergleich von Literaturdaten einer Hybridisierung von T47D-DNA mit Hybridisierungen von T47D-DNA auf einem 244K Agilent Oligonukleotid-Array

Dargestellt ist der Vergleich von Hybridisierungsprofilen der T47D-Zelllinie auf dem 3K BAC-Array und auf einem 32K-BAC-Array (Literaturdaten aus Shadeo et al.⁸⁷), ausschnittsweise für die Chromosomen 3 (a) und 10 (b). Auf dem 32K-BAC-Array wurde nicht-amplifizierte DNA aus ca. 10⁶ Zellen der T47D-Zelllinie hybridisiert (jeweils auf der linken Seite). Auf dem 244K Oligonukleotid-Array wurde zum einen nicht-amplifizierte DNA aus ca. 10⁶ Zellen der T47D-Zelllinie (blau) und amplifizierte DNA aus einer T47D-Einzelzelle (orange) hybridisiert. Signifikante Auslenkungen sind durch Balken rechts und links des Profils dargestellt (Darstellung der Array-CGH wie in Abbildung 14). Bei den Hybridisierungsprofilen auf dem 32K BAC-Array sind in Rot bzw. Grün signifikante Gewinne bzw. Verluste angezeigt.

Vergleich von 3K BAC-, 35K und 244K Oligonukleotid-Array durch ROC-Analyse

Der 3K BAC-Array und die beiden untersuchten Oligonukleotid-Arrays unterscheiden sich beträchtlich in ihrer Auflösung, was ihren direkten Vergleich untereinander erschwert. Um dennoch einen Vergleich der Plattformen zu ermöglichen, wurde in einer ROC-Analyse (2.1.13) ermittelt, wie gut der Karyotyp der nicht-amplifizierten T47D-DNA in den Hybridisierungsprofilen der T47D-Einzelzell-Amplifikate durch jeden der drei Arrays dargestellt war (Abbildung 28). Die beiden Oligonukleotid-Arrays konnten den Karyotyp der T47D-DNA in den Einzelzell-Amplifikaten im Gegensatz zum 3K BAC-Array nicht darstellen.

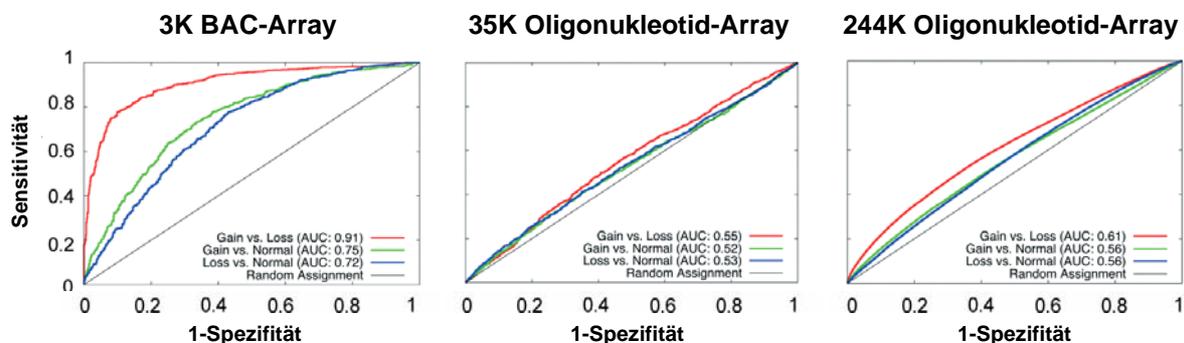


Abbildung 28 Übereinstimmung zwischen den Hybridisierungsprofilen einer T47D-Einzelzell-DNA und einer nicht-amplifizierten T47D-DNA auf verschiedenen Arrays, dargestellt durch ROC-Kurven

In den ROC-Kurven ist dargestellt, wie gut der Karyotyp der nicht-amplifizierten T47D-DNA in den Hybridisierungsprofilen der T47D-Einzelzell-Amplifikate durch den 3K BAC-Array (links), den 35K Operon Oligonukleotid-Array (Mitte) und den 244K Agilent Oligonukleotid-Array (rechts) dargestellt werden kann. Eine Zufallsverteilung ist durch die Winkelhalbierende dargestellt (graue Linie). In jedem der Diagramme sind ROC-Kurven dargestellt, die Gewinn gegen Verlust (grün), Gewinn gegen Normal (blau) und Verlust gegen Normal (rot) vergleichen. Die Fläche unter der Kurve (AUC) für jede der Kurven ist angegeben.

Einfluss von Glättung auf Hybridisierungsprofile von Einzelzell-Amplifikaten

Bei der Auswertung von Array-CGH-Hybridisierungsdaten wird in der Literatur häufig eine Glättung, also eine Zusammenfassung einzelner Datenpunkte vorgenommen⁹¹, um die Streuung des Datensatzes zu mindern. Dieses Verfahren wird angewandt, damit auch genomische Veränderungen mit geringer Amplitude detektiert werden können, die sonst in der Streuung nicht zu erkennen wären. Ein Glättungsverfahren bietet sich vor allem an, wenn viele Datenpunkte zur Verfügung stehen, so dass trotzdem auch kleine Veränderungen noch dargestellt werden. In der einfachsten Form der Glättung werden ausgehend von einem Datenpunkt unterschiedlich viele benachbarte Datenpunkte zu ihrem arithmetischen Mittel zusammengefasst. Das Mittel dieses Glättungsfensters ergibt so den geglätteten Datenpunkt.

Die Bioinformatik hat inzwischen Glättungsalgorithmen wie den Savitzky-Golay-Filter entwickelt⁸¹, die zum einen die Streuung der Datenpunkte mindern, zum anderen aber auch die Detektion von kleinen Aberrationen geringer Amplitude ermöglichen (2.1.13).

Es sollte untersucht werden, ob durch eine Glättung der Hybridisierungsdaten eine bessere Darstellung des Karyotyps der amplifizierten T47D-Einzelzell-DNA auf den drei getesteten Hybridisierungsplattformen möglich ist. Für die Glättung wurde der Savitzky-Golay-Filter wie in 2.1.13 beschrieben verwendet, die Glättungsfenster wurden so gewählt, dass alle drei Arrays eine ähnliche Auflösung hatten (3K BAC-Array: $n_L = n_R = 4$; 35K Operon Oligonukleotid-Array: $n_L = n_R = 25$ und 244K Agilent Oligonukleotid-Array: $n_L = n_R = 330$). Um alle drei Arrays untereinander vergleichen zu können, wurden in diesem Fall die geglätteten Datenpunkte der nicht-amplifizierten T47D-DNA auf dem 244K Agilent Oligonukleotid-Array als Referenz für alle drei Datensätze verwendet. Um den Referenz-Datensatz für die verschiedenen Arrays zu erhalten, wurde für jeden Datenpunkt innerhalb des Datensatzes der Einzelzell-Hybridisierung der Datenpunkt innerhalb des Referenz-Datensatzes ausgewählt, der dem Mittelpunkt des Glättungsfensters am nächsten war. Die Daten wurden wieder in Form von ROC-Kurven dargestellt (Abbildung 29, oben). Die Glättung hatte einen positiven Effekt auf die Darstellung des T47D-Karyotyps in den Einzelzell-Amplifikaten auf allen drei Plattformen, mit dem 3K BAC-Array konnte der T47D-Karyotyp aber am besten dargestellt werden.

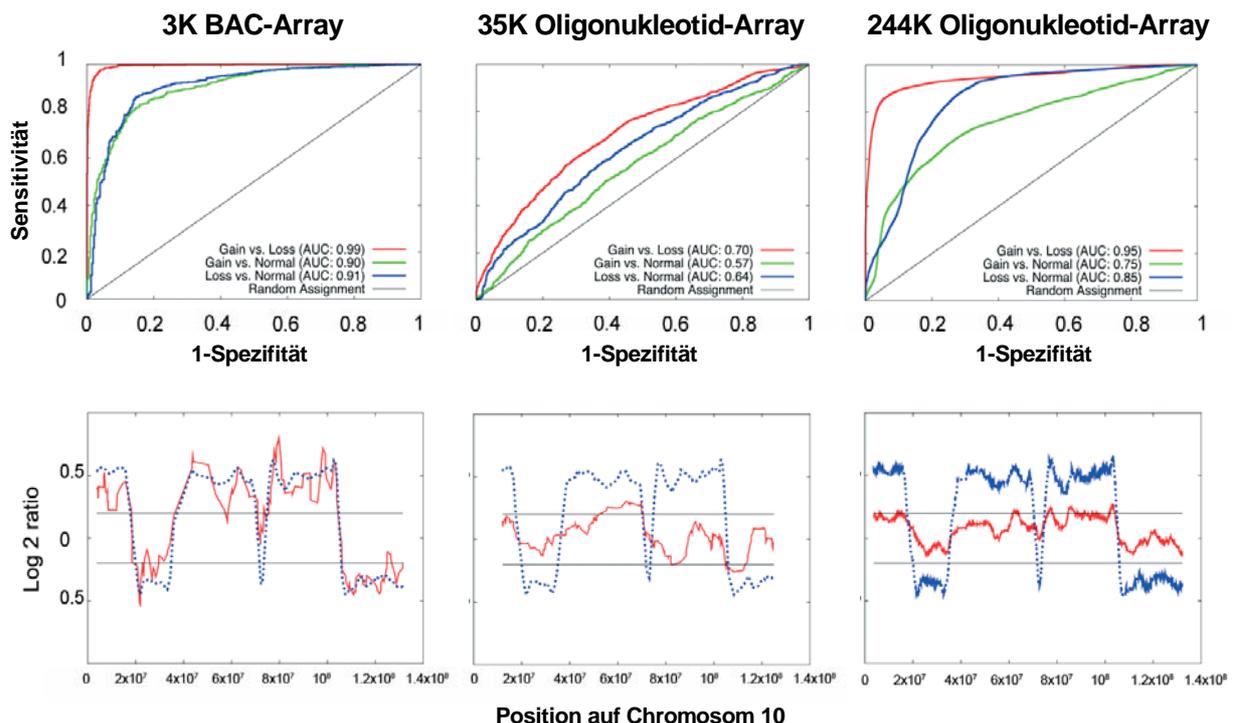


Abbildung 29 Einfluss einer Glättung durch den Savitzky-Golay-Filter auf die Darstellung des T47D-Karyotyps in Einzelzell-Amplifikaten auf verschiedenen Hybridisierungsplattformen

Oben: ROC-Kurven, wie in Abbildung 28. Die Hybridisierungsdaten wurden vorher durch die Anwendung eines Savitzky-Golay-Filters geglättet (2.1.13). Unten: Log₂-Verhältnisse von T47D-Einzelzell-Amplifikaten (rot) sind im Vergleich zu den Log₂-Verhältnissen der nicht-amplifizierten T47D-DNA (244K Oligonukleotid-Array) (blau) beispielhaft für das Chromosom 10 aufgetragen. 3K BAC-Array (links), 35K Oligonukleotid-Array (Mitte) und 244K Oligonukleotid-Array (rechts). Die schwarzen Linien zeigen die Signifikanzgrenzen für einen Gewinn (0,2) und einen Verlust (-0,2) an.

Aus der ROC-Analyse könnte geschlussfolgert werden, dass der 244K Oligonukleotid-Array nach einer Glättung für die Untersuchung von Einzelzell-Amplifikaten geeignet ist. Da die ROC-Analyse aber nur anzeigen kann, wie gut die Struktur der Daten übereinstimmt, sagt sie nichts darüber aus, ob der Karyotyp anhand der Log₂-Verhältnisse dargestellt werden kann. Deshalb wurden Hybridisierungsprofile der T47D-Einzelzell-Amplifikate aufgrund der geglätteten Datenpunkte erstellt (beispielhaft für Chromosom 10 in Abbildung 29 unten). Zum Vergleich wurden die Profile aus geglätteten Datenpunkten der nicht-amplifizierten T47D-DNA jeweils im selben Diagramm dargestellt. In dieser Darstellung der Daten ist zu sehen, dass die Datenstruktur des Einzelzell-Amplifikats und der nicht-amplifizierten DNA für den 244K Array zwar ähnlich ist, aber die Log₂-Verhältnisse in Gewinn- und Verlust-Regionen nicht die Signifikanzgrenzen von 0,2 und -0,2 erreichen. Das Einzelzell-Profil auf dem 35K Oligonukleotid-Array lässt hingegen nicht einmal die erwartete Datenstruktur erkennen. Mit dem 3K BAC-Array können sowohl die erwartete Datenstruktur als auch die erwarteten Log₂-Verhältnisse dargestellt werden. Es kann daher festgehalten werden, dass der 3K BAC-Array bei der Untersuchung von amplifizierter Einzelzell-DNA auch nach einer Datenglättung Oligonukleotid-Arrays überlegen ist.

3.2 Hybridisierung von Amplifikaten disseminierter Tumoreinzellen

3.2.1 Validierung von Alterationen in Einzelzell-Amplifikaten

Da mit dem 3K BAC-Array letztlich einzelne disseminierte Tumorzellen aus dem Knochenmark von Tumorpatienten untersucht werden sollen, wurden zur weiteren Charakterisierung des Arrays drei Tumorzellen aus dem Knochenmark einer Brustkrebspatientin ausgewählt. Zu diesen Zellen lagen bereits Metaphasen-CGH-Profile vor, die durch eine Untersuchung mit dem 3K BAC-Array verifiziert werden sollten. Die Zellen wurden ausgewählt, da die Metaphasen-CGH-Profile auf eine klonale Verwandtschaft zwischen den Zellen hinwiesen und sie zudem komplexe genomische Veränderungen aufwiesen. Array- und Metaphasen-CGH-Profile (Abbildung 30) dieser Zellen wurden

einander gegenübergestellt, der Übersichtlichkeit halber werden nur zwei der drei Zellen gezeigt.

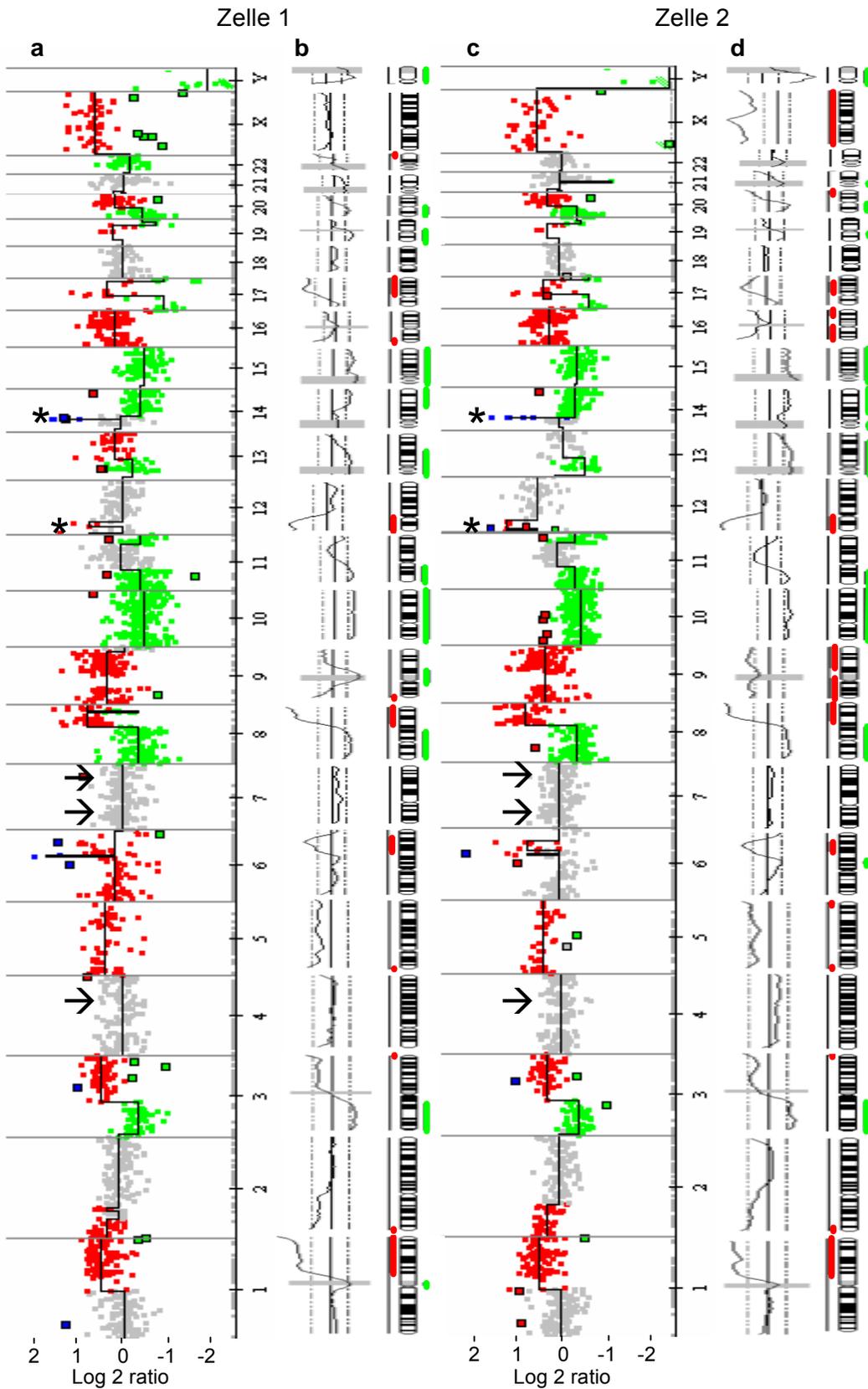


Abbildung 30 Array- und Metaphasen-CGH-Profile von zwei disseminierten Tumorzellen aus dem Knochenmark einer Mammakarzinompatientin

(a) und (c) zeigen die Array-CGH-Profile und (b) und (d) die Metaphasen-CGH-Profile von zwei einzelnen disseminierten Tumorzellen (Zelle 1 und 2; 113_1 und 113_2) aus dem Knochenmark einer Patientin mit Mammakarzinom (Darstellung der Array-CGH wie in Abbildung 14; Chromosom 19 wird nicht berücksichtigt, siehe 3.1.6). Bei der Metaphasen-CGH wird durch die grünen und roten Balken neben den stilisierten Chromosomen angezeigt, ob ein Bereich deletiert oder amplifiziert ist. Der * zeigt amplifizierte Bereiche an, die mittels quantitativer PCR (qPCR) validiert wurden, der → zeigt Regionen an, die bei der qPCR als Kontrollen dienten (s.u.).

Bei einem visuellen Vergleich von Metaphasen- und Array-CGH-Profilen der beiden Einzelzellen fallen genomische Bereiche auf, die gleich bewertet werden, z.B. der normale Bereich auf dem p-Arm von Chromosom 1 und der Gewinn von 1q oder der Verlust von 3p. Die Korrelation zwischen den Array-CGH- und den Metaphasen-CGH-Profilen wurde berechnet, wobei untersucht wurde, ob Aberrationen, die in der Metaphasen-CGH erkannt wurden, auch in der Array-CGH detektiert werden konnten. Korrelationskoeffizienten von 0,91 und 0,92 wurden für Zelle 1 und 2 berechnet. Trotzdem fallen Bereiche auf, in denen sich Metaphasen- und Array-CGH unterscheiden, so zeigt das Metaphasen-CGH-Profil für 3q zwar einen Ausschlag in Richtung Gewinn an, dieser wird jedoch nicht als signifikant beurteilt, wohingegen dieser Bereich in der Array-CGH als amplifiziert angezeigt wird. Ähnliche Beobachtungen wurden bei vergleichenden Analysen von drei weiteren disseminierten Tumorzellen mit bekannten Metaphasen-CGH-Profilen gemacht. Hierbei lagen die Korrelationskoeffizienten bei 0,94, 0,94 und 0,88 (Anhang 1 für die Darstellung der vergleichenden Profile).

Anhand von Ausschnittsvergrößerungen der Chromosomen 11 und 14 derselben zwei Tumorzellen ist exemplarisch dargestellt, wie sich Metaphasen- und Array-CGH-Profile unterscheiden (Abbildung 31). Auf dem Chromosom 11 zeigt die Metaphasen-CGH in beiden Zellen zwar einen Ausschlag in Richtung Gewinn, dieser ist allerdings nicht signifikant. Die Array-CGH zeigt in diesem Bereich einen Gewinn an, der aber von zwei 10,9 und 4,7 Mb großen Deletionen unterbrochen ist. Die Größe dieser Regionen ist zu klein, um von der Metaphasen-CGH erkannt werden zu können. Auf dem Chromosom 14 wird in der Array-CGH eine 4,0 Mb große Amplifikation angezeigt, die nicht durch die Metaphasen-CGH erkannt wird. Damit wird deutlich, dass die Array-CGH für die Untersuchung von Einzelzell-DNA sensitiver ist und eine höhere Auflösung besitzt als die Metaphasen-CGH.

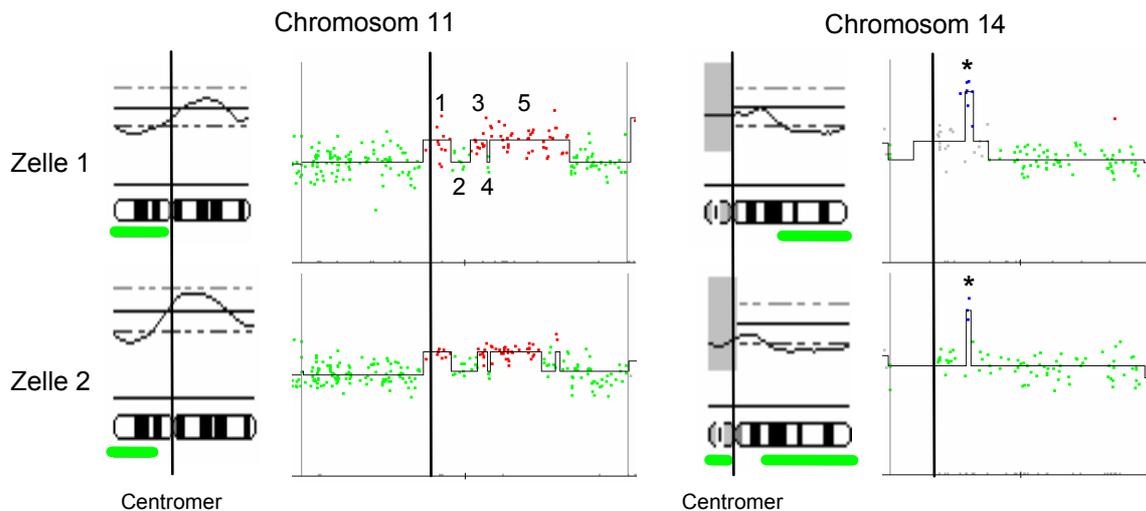


Abbildung 31 Ausschnittsvergrößerung der Chromosomen 11 und 14 der Array- und Metaphasen-CGH-Profile von zwei disseminierten Tumorzellen aus dem Knochenmark einer Mammakarzinompatientin

Ausschnittsvergrößerung der Chromosomen 11 und 14 aus Abbildung 30 für das Metaphasen-CGH und das Array-CGH-Profil (Darstellung der Array-CGH wie in Abbildung 14). Die senkrechte, schwarze Linie zeigt hier zusätzlich die Position des Centromers an. Beim Array-CGH-Profil von Chromosom 11 sind mit 1, 2, 3, 4 und 5 die abwechselnd amplifizierte und deletierte Regionen gekennzeichnet. Der * zeigt amplifizierte Bereiche an, die mittels quantitativer PCR (qPCR) validiert wurden (s.u.).

Um Array-CGH-Aberrationen zu verifizieren, wird in der Literatur meist eine FISH-Analyse durchgeführt. Um in Einzelzell-Amplifikaten zu verifizieren, ob eine detektierte Aberration korrekt ist, sollte eine Validierung mit Hilfe einer LOH-Analyse in Anlehnung an Cardoso et al.⁹² durchgeführt werden.

Der q-Arm von Chromosom 11 (Abbildung 31) weist in beiden Tumorzellen das gleiche Muster von Gewinn (Region 1), Verlust (2), Gewinn (3), Verlust (4) und Gewinn (5) auf. Durch eine LOH-Untersuchung der Regionen 11q13.4 (Region 3) bis 11q14.1 (Region 5) sollte dieses Muster verifiziert werden. Anhand eines Normalzellpools der Patientin wurden informative Marker (Marker, für die die Patientin heterozygot war) zunächst exemplarisch für die Regionen 3, 4 und 5 (s.o.) identifiziert. Daraufhin wurden die disseminierten Tumorzellen auf diese Marker getestet (Tabelle 15). Für die amplifizierte Region 5 stimmte das Ergebnis mit dem erwarteten Ergebnis überein – beide Allele waren vorhanden. In der ebenfalls amplifizierten Region 3 war allerdings nur eins von zwei Allelen nachweisbar, das Ergebnis stimmte nicht mit der Erwartung überein. In der deletierten Region 4 wurde erwartet, dass nur eins von zwei Allelen nachweisbar sein sollte, es konnten aber beide Allele detektiert werden. Eine Verifizierung dieser Aberrationen mittels Allelverlust-Analyse war daher nicht möglich.

Tabelle 15 LOH-Analyse von in der Array-CGH detektierten Aberrationen

Array-CGH Ergebnis	Gewinn	Verlust	Gewinn	Gewinn
Region	3	4	5	5
Marker	rs7935825	rs621456	rs2373547	rs4945425
Normalzell-Pool der Patientin	1			
Zelle 1	2			
Zelle 2				
Zelle 3				
Ergebnis Untersuchung				
erwartet				

¹ hellgrau: beide Allele vorhanden

² dunkelgrau: Allelverlust

Als zweite unabhängige Verifizierungsmethode zur Überprüfung von Amplifikationen, die mit der Array-CGH detektiert wurden, wurde daraufhin die quantitative PCR (qPCR) gewählt. Es sollten zwei amplifizierte Regionen mit Hilfe der qPCR untersucht werden. Da das Ausgangsmaterial für die qPCR Einzelzell-Amplifikate sind, wurden für jede zu untersuchende Region drei unterschiedliche PCRs durchgeführt, um statistisch auswertbare Ergebnisse erzielen zu können. Die Kopienzahl einer Region wurde berechnet, indem ein Verhältnis zwischen der gemessenen Kopienzahl der untersuchten Region und der gemessenen Kopienzahl einer Kontroll-Region gebildet wurde. Primer für Kontroll-Regionen wurden in balancierten Bereichen ausgewählt. Um zusätzlich auszuschließen, dass es sich bei den Ergebnissen um Artefakte handelte, die durch die primäre Amplifikation der Einzelzell-DNA entstanden waren, wurden qPCRs für die gleichen Marker mit zehn normalen Einzelzell-Amplifikaten durchgeführt. Durch einen Mann-Whitney-Test wurde untersucht, ob sich die Daten für die disseminierten Tumorzellen und die zehn Normalzellen signifikant voneinander unterscheiden.

Die 4,0 Mb Amplifikation auf Chromosom 14qcen (siehe Abbildung 30 und Abbildung 31), die nur durch die Array-CGH dargestellt wurde, und der 5,5 Mb Gewinn auf Chromosom 12ptel (siehe Abbildung 30), der sowohl in Metaphasen- als auch in der Array-CGH dargestellt wird, wurden mittels qPCR untersucht. Um die Region auf 12ptel zu untersuchen, wurden Primer im *AKAP3*-Lokus ausgewählt, für die Region auf 14qcen wurden Primer im *AKAP6*-Lokus gewählt. Die Primer für die Kontroll-Regionen wurden im *SMS2*-Lokus auf 4q, im *IL6*-Lokus auf 7p und im *BC*-Lokus auf 7q gewählt. In Abbildung 30 sind die zu untersuchenden Regionen durch * und die Kontroll-Regionen durch → gekennzeichnet. Zusätzlich zu den beiden Tumorzell-Amplifikaten, deren Array-CGH-Profile in Abbildung 30 dargestellt sind, wurde noch ein weiteres Amplifikat einer dritten disseminierten

Tumorzelle aus dem Knochenmark derselben Patientin untersucht (Zelle 3; 113_3). Die Kopienzahl für den *AKAP3*-Lokus auf 12ptel betrug in den drei untersuchten Zellen 4-5, die Kopienzahl für *AKAP6* auf 14qcen betrug 5-7 (Tabelle 16).

Tabelle 16 Kopienzahl für *AKAP3* und *AKAP6* und Signifikanz in drei disseminierten Tumorzellen aus dem Knochenmark einer Patientin mit Mammakarzinom

	Zelle 1	Zelle 2	Zelle 3
<i>AKAP 3</i>			
berechnete Kopienzahl	5	4	4
p (zweiseitig)*	0,001	0,005	0,001
<i>AKAP 6</i>			
berechnete Kopienzahl	5	7	6
p (zweiseitig)*	< 0,001	< 0,001	< 0,001

*Signifikanzunterschied der qPCR-Werte zwischen Tumorzellen und Normalzellen

Die berechnete Kopienzahl bezieht sich auf einen diploiden Chromosomensatz, der vermutlich nicht in den Zellen vorlag. Die Tumorzellen unterschieden sich im Ergebnis der qPCR signifikant von den normalen Einzelzellen (p-Werte (Mann-Whitney-Test) in Tabelle 16). Das zeigt, dass es sich bei den detektierten Amplifikationen um echte Alterationen und nicht um Artefakte handelte. Dieses Ergebnis bestätigt, dass die Untersuchung von Einzelzell-Amplifikaten mit Array-CGH mit erhöhter Sensitivität und Auflösung im Vergleich zur Metaphasen-CGH möglich ist.

3.2.2 Hochauflösende genomische Analyse von einzelnen disseminierten Tumorzellen aus dem Knochenmark von Mammakarzinom-Patientinnen

Nach erfolgreicher Etablierung der Hybridisierung und Charakterisierung des 3K BAC-Arrays wurden Pilotuntersuchungen von Patientenproben durchgeführt. Aus dem Kollektiv von einzelnen disseminierten Tumorzellen aus dem Knochenmark von Mammakarzinompatientinnen, die bereits durch Metaphasen-CGH charakterisiert worden waren^{73, 83}, wurden Zellen auf den 3K BAC-Array hybridisiert. Im vorangegangenen Kapitel wurde gezeigt, dass Veränderungen in disseminierten Tumorzellen von Patientinnen mit manifesten Metastasen (M1-Zellen), die durch die Metaphasen-CGH detektiert werden konnten, durch die Array-CGH reproduziert werden (3.2.1). Im Anschluss daran sollte untersucht werden, ob mit Hilfe des 3K BAC-Arrays Aberrationen in disseminierten Tumorzellen identifiziert werden können, die jenseits der Auflösungsgrenze der Metaphasen-CGH liegen. Hierfür wurden sieben einzelne Tumorzellen aus dem Knochenmark von Patientinnen ohne Metastasen (M0-Zellen) hybridisiert, die in der Metaphasen-CGH normale

Karyotypen zeigten⁸³, in der LOH-Untersuchung allerdings zahlreiche Allel-Verluste aufwiesen^{75,93} (Anhang 3).

Vergleich Metaphasen-CGH-negativer disseminierter Tumorzellen mit Normalzellen

Zunächst wurde untersucht, ob disseminierte Tumorzellen, die keine durch Metaphasen-CGH detektierbare Aberrationen aufwiesen, aufgrund ihrer Array-CGH-Profile von normalen Einzelzellen unterschieden werden können. Einzelzell-Amplifikate von vier normalen (000_1 bis 000_4; Anhang 2) und sieben disseminierten Tumorzellen ohne Metaphasen-CGH-Aberrationen (011_1, 012_1, 018_1, 022_1, 215_1, 221_1 und 230_1; Anhang 3) wurden auf dem 3K BAC-Array hybridisiert und die Array-CGH Profile anschließend einer Ähnlichkeitsanalyse unterworfen (2.1.12). Es zeigte sich, dass sich vier der sieben Tumorzellen (011_1, 018_1, 022_1 und 230_1) deutlich von den restlichen Zellen unterschieden (Abbildung 32). Davon abgetrennt waren die übrigen drei disseminierten Tumorzellen (012_1, 215_1, 221_1). Eine weitere Gruppe mit den vier Normalzell-Amplifikaten wurde ebenfalls klar abgetrennt (000_1 bis 000_4). Mit der Ähnlichkeitsanalyse wurde gezeigt, dass sich disseminierte Tumorzellen, die in der Metaphasen-CGH keine Veränderungen zeigten, aufgrund ihrer Array-CGH-Profile deutlich von normalen Einzelzellen unterscheiden.

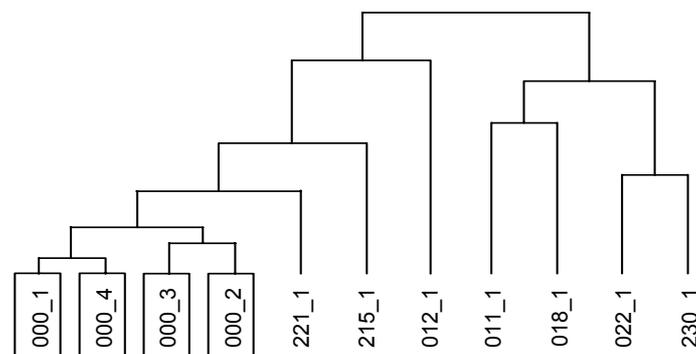


Abbildung 32 Ähnlichkeitsanalyse normaler Einzelzellen und disseminierter Tumorzellen von M0-Patientinnen, die in der Metaphasen-CGH keine Veränderungen zeigten

Array-CGH-Profile vier normaler Einzelzell-Amplifikate (eingerahmt) und sieben Metaphasen-CGH-negativer disseminierter Tumorzellen von M0-Patientinnen wurden einer Ähnlichkeitsanalyse unterworfen. Die Array-CGH-Profile der Normalzellen sind im Anhang 2, die der disseminierten Tumorzellen in Anhang 3 dargestellt. Zur Berechnung der Ähnlichkeit wurde die Manhattan-Distanz verwendet.

Durch die verbesserte Auflösung des 3K BAC-Arrays konnten in disseminierten M0-Zellen, die keine durch Metaphasen-CGH detektierbare Aberrationen aufwiesen, bislang unerkannte genomische Veränderungen nachgewiesen werden. Im Gegensatz dazu wurden in

Amplifikaten normaler Einzelzellen signifikant weniger veränderte BAC-Klone detektiert als in Amplifikaten disseminierter Tumorzellen von M0-Patientinnen (Tabelle 17).

Tabelle 17 Aberrante BAC-Klone in Array-CGH-Profilen von Einzelzell-Amplifikaten

	% aberrante BAC-Klone	p*	p*	p*
Normale Einzelzellen	0,4 ±0,6] 0,05] 0,05] 0,05
Disseminierte Tumorzellen aus M0-Patientinnen	19,1 ±8,3			
Disseminierte Tumorzellen aus M1-Patientinnen	67,3 ±12,6			

*zweiseitiger p-Wert (Mann-Whitney Rangsummen-Test)

Vergleich Metaphasen-CGH-negativer und –veränderter disseminierter Tumorzellen

Im nächsten Schritt wurden Array-CGH-Profile von Einzelzell-Amplifikaten aus disseminierten Tumorzellen von M0-Patientinnen (n=7; s.o.) mit solchen von M1-Patientinnen (n=6; 118_1, 113_1-3, 008_1-2) verglichen (Abbildung 33). Hierfür wurden die Profile wieder einer Ähnlichkeitsanalyse unterworfen. In dem resultierenden Cluster werden bei den Zellen von M1-Patientinnen zunächst die Zellen, die von einer Patientin stammen, zusammengefasst. Die Zellen der drei verschiedenen Patientinnen unterscheiden sich deutlich voneinander. Die disseminierten Tumorzellen von M0-Patientinnen werden zu einer Gruppe zusammengefasst und deutlich von den Zellen der M1-Patientinnen abgetrennt. Auch mit Hinblick auf die Anzahl der aberranten BAC-Klone unterscheiden sich sowohl M1-Zellen von normalen Einzelzellen als auch M1- von M0-Zellen signifikant (Tabelle 17). In disseminierten Tumorzellen von M1-Patientinnen konnten mehr als dreimal so viele veränderte BAC-Klone nachgewiesen werden, als in disseminierten Tumorzellen von M0-Patientinnen.

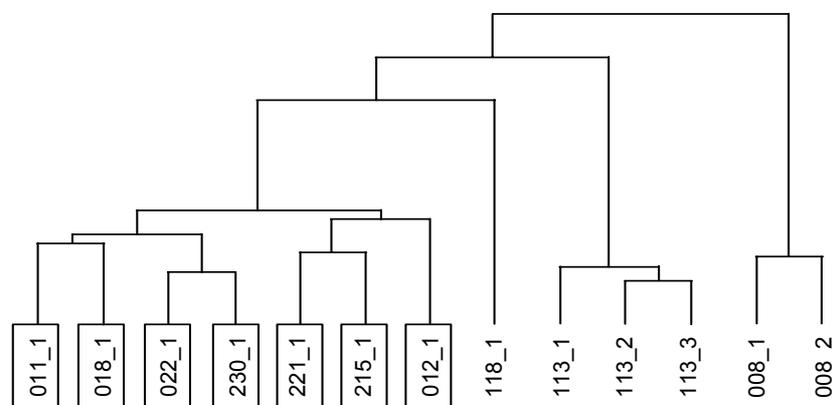


Abbildung 33 Ähnlichkeitsanalyse disseminierter Tumorzellen von M0- und M1-Patientinnen

Array-CGH-Profile sechs disseminierter Tumorzellen von Patientinnen mit manifesten Metastasen und sieben Metaphasen-CGH-negativer disseminierter Tumorzellen von M0-Patientinnen (eingerahmt) wurden einer

Ähnlichkeitsanalyse unterworfen. Die Array-CGH-Profile der Zellen 008_1-2 und 118_1 sind im Anhang 1, die der Zellen 113_1-2 in Abbildung 30 dargestellt. Zur Berechnung der Ähnlichkeit wurde die Manhattan-Distanz verwendet.

Gemeinsame Aberrationen in disseminierten Tumorzellen von M0-Patientinnen

Disseminierte Tumorzellen von Patientinnen ohne nachweisbare Metastasen (M0-Zellen) sind von besonderem Interesse, da diese Zellen vermutlich die genomischen Veränderungen tragen, die für eine frühe Disseminierung verantwortlich sind⁷⁵. Deshalb sollte in diesen Zellen nach gemeinsamen genomischen Veränderungen gesucht werden.

In den sieben Hybridisierungsprofilen von M0-Zellen (Anhang 3) wurde nach Veränderungen gesucht, die in mindestens 50% der Zellen detektiert werden konnten. Um auszuschließen, dass es sich bei den identifizierten Regionen um Artefakte handelte, wurden mit den gleichen Kriterien auch vier normale Einzelzellprofile (Anhang 2) untersucht. Solche Veränderungen, die in mindestens 50% der normalen Zellen gefunden wurden, wurden als Artefakt beurteilt und nicht berücksichtigt. Das Chromosom 17 wurde zwar in nur einer der vier Normalzellen als deletiert erkannt (004_1), in zwei weiteren Zellen (001_1 und 003_1) ist jedoch zu sehen, dass das durchschnittliche Log₂-Verhältnis für Chromosom 17 insgesamt geringer war als für die übrigen Chromosomen (Log₂-Verhältnis = -0,24 ±0,04 im Gegensatz zu -0,02 (3.1.7)). Die Deletion von Chromosom 17 wurde aus diesem Grund ausgeschlossen. Ebenso wurden das Chromosom 19 (3.1.6) und die Geschlechtschromosomen nicht berücksichtigt. In Tabelle 18 sind alle Veränderungen aufgelistet, die in mindestens 50% der M0-Zellen gefunden wurden. Es ist jeweils gekennzeichnet, ob es sich um einen Gewinn oder Verlust handelt. Des Weiteren ist die Anzahl der in den Regionen enthaltenen Gene aufgelistet.

Tabelle 18 Auflistung der minimalen Alterationen in 50% der M0-Zellen

Chr.	Name der Region	Größe (Mb)	Gewinn/Verlust	Gene (Ensembl Release 44, April 2007)
5	MinReg_L_5q14-5q21_1	18,7	Verlust	164 Gene
7	MinReg_L_7p21_2	1,8	Verlust	5 Gene
7	MinReg_G_7p14_1	5,0	Gewinn	18 Gene
10	MinReg_G_10q21_1	5,6	Gewinn	12 Gene
11	MinReg_L_11q12-11q13_4	12,7	Verlust	616 Gene
13	MinReg_G_13q14-13q31_2	41,4	Gewinn	181 Gene
15	MinReg_L_15q11.2-15q26_5	75,4	Verlust	1375 Gene
22	MinReg_L_22q11.1-22q13_9	31,9	Verlust	964 Gene

² ebenfalls in Normalzellen vorhanden; Artefakt, nicht berücksichtigt

Korrelation von Hybridisierungsprofilen mit klinischen Parametern in disseminierten Tumorzellen von M0-Patientinnen

Die Untersuchung von Zusammenhängen zwischen Aberrationsmustern in den untersuchten disseminierten Tumorzellen und den klinischen Eigenschaften der Tumorerkrankung kann dazu beitragen, genomische Veränderungen zu identifizieren, die für bestimmte Tumorsubgruppen relevant sind.

Die Hybridisierungsprofile der insgesamt sieben M0-Zellen aus sieben Patientinnen wurden hierfür einer hierarchischen Ähnlichkeitsanalyse unterworfen (Abbildung 34). Für einen besseren Vergleich der einzelnen Proben untereinander wurden die Hybridisierungsprofile als so genannte dot plots dargestellt, in denen jeder Balken einen BAC-Klon darstellt, dessen Farbe anzeigt, ob er balanciert (grau), amplifiziert (rot) oder deletiert (grün) ist; die einzelnen Klone werden ihrer genomischen Position entsprechend dargestellt. Das X- und das Y-Chromosom wurden nicht in die Ähnlichkeitsanalyse eingeschlossen. Klinische Informationen zu Typ und Rezeptoreigenschaften des Tumors wurden mit dem aus der Ähnlichkeitsanalyse resultierenden Dendrogramm verglichen, um Zusammenhänge zwischen Aberrationsmustern und klinischen Eigenschaften untersuchen zu können.

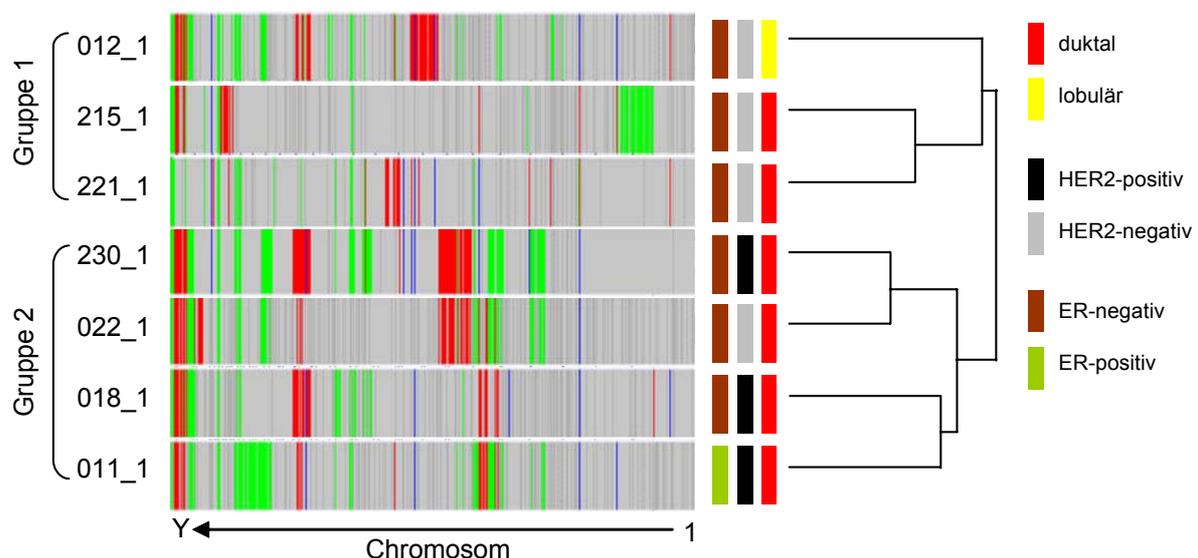


Abbildung 34 Ähnlichkeitsanalyse von disseminierten Tumorzellen aufgrund von Array-CGH-Daten und Korrelation mit klinischen Daten

Die Hybridisierungsdaten von einzelnen disseminierten Tumorzellen aus M0-Patientinnen wurden mittels hierarchischer Ähnlichkeitsanalyse (euklidische Distanz) auf ihre Ähnlichkeit zueinander untersucht. Aufgrund dieser Ähnlichkeit wurde ein Dendrogramm erstellt. Probenbezeichnungen, die mit „0“ beginnen, kennzeichnen Zellen von M0-Patientinnen, die sowohl Metaphasen-CGH-negative als auch -aberrante Zellen aufwiesen, Probenbezeichnungen, die mit „2“ beginnen, kennzeichnen Zellen von M0-Patientinnen, die nur Metaphasen-CGH-negative Zellen aufwiesen. Informationen zu klinischen Daten sind durch farbige Balken gekennzeichnet.

Die Zellen wurden in zwei Gruppen eingeteilt. Gruppe 1, die drei Zellen umfasste (012_1, 215_1 und 221_1), wies 10,1 ±4,0% veränderte BAC-Klone auf, Gruppe 2 (230_1, 022_1, 018_1 und 011_1) im Gegensatz dazu mehr als doppelt so viele veränderte BAC-Klone (24,9 ±3,7%; p(zweiseitig)=0,008). Die in Gruppe 1 enthaltenen Zellen hatten alle einen negativem HER2-Rezeptorstatus, wohingegen drei von vier Zellen in Gruppe 2 HER2-positiv waren. Zusätzlich zu den dargestellten Parametern wurde außerdem untersucht, ob sich die Gruppen hinsichtlich ihres TNM-Status und des Gradings unterschieden, es konnte jedoch kein Zusammenhang festgestellt werden. Auch in Bezug auf das Zehnjahresüberleben und die Entwicklung von Metastasen konnte kein Unterschied zwischen beiden Gruppen festgestellt werden.

4 Diskussion

4.1 Entwicklung des 3K BAC-Arrays zur Einzelzell-DNA-Analyse

Identifizierung der notwendigen Parameter für die Untersuchung von Einzelzellen mittels Array-CGH

Die Zielsetzung dieser Dissertation war die Entwicklung eines CGH-Arrays mit einer Auflösung von 1 Mb, mit dem amplifizierte DNA aus einzelnen Zellen auf genomische Aberrationen untersucht werden konnte. Bisher konnten genomische Veränderungen von Einzelzellen nur mit Hilfe der konventionellen Metaphasen-CGH untersucht werden, die eine Auflösung von ca. 10-30 Mb besitzt⁷². Sehr kleine Aberrationen zu entdecken sowie eine genaue Lokalisation der Bruchpunkte zu bestimmen, ist mit dieser Methode nicht möglich. Des Weiteren treten bei der Metaphasen-CGH „unzuverlässig“ hybridisierende Regionen auf, wie die Telomere, der terminale Bereich von 1p32-p, 16p, 19p und 22q, was vermutlich durch ihren hohen Anteil an repetitiven Sequenzen bedingt ist^{60, 94, 95}. Aufgrund dieser Limitierungen wurde 1997 die Array-CGH entwickelt^{34, 35}, bei der die humanen Metaphasen durch bakterielle artifizielle Chromosomen (BACs) ausgetauscht wurden. Da die genaue Position dieser BAC-Klone innerhalb des Genoms bekannt ist, konnten BACs selektiert werden, die in regelmäßigen Intervallen das Genom abdecken. Zu Beginn der Erstellung dieser Arbeit existierten bereits BAC-CGH-Arrays, die ca. 3000 BAC-Klone umfassten und somit eine Auflösung von 1 Mb erreichten³⁷, sowie Arrays mit geringerer Auflösung⁹⁶. Die BAC-Klone für diese sowie auch für weiterentwickelte Arrays wurden durch eine konventionelle Plasmid-Aufreinigung isoliert, bei der Reste der genomischen *E. coli* DNA in den Präparationen verbleiben und somit auch auf den Arrays vorkommen.

Bei der Hybridisierung von Einzelzell-Amplifikaten auf einem konventionell hergestellten Array (Peter Lichter, DKFZ, Heidelberg) konnten die erwarteten Karyotypen nicht dargestellt werden. Eine Erklärung hierfür war der hohe Anteil an kontaminierenden, bakteriellen Sequenzen in den Einzelzell-Amplifikaten (Abbildung 6), die mit der bakteriellen DNA auf dem Array kreuzhybridisierten und das spezifische Hybridisierungssignal überlagerten. Um die bakterielle Kontamination der BAC-Sonden zu eliminieren, wurde eine Aufreinigung der Plasmid-DNA durch eine Pulsed-field-Gelelektrophorese (PFGE) durchgeführt, was zu einer deutlichen Verringerung des bakteriellen Anteils führte (Abbildung 7). Für die Herstellung eines Test-Arrays mit 30 BAC-Klonen wurde die präparierte BAC-DNA durch Adapter-

Linker-PCR amplifiziert und anschließend auf Glasobjektträgern immobilisiert³⁷. Um einen direkten Vergleich zu konventionellen BAC-Präparationen zu ermöglichen, wurde auf dem Test-Array sowohl konventionell als auch durch PFGE gereinigte BAC-DNA immobilisiert und hybridisiert. Bei diesem Experiment zeigte sich die Überlegenheit der PFGE gegenüber der konventionell gereinigten BAC-DNA dadurch, dass in der hybridisierten Einzelzell-DNA eines Patienten mit Down Syndrom nur mit den PFGE gereinigten Klonen der Gewinn der Trisomie 21 dargestellt werden konnte (Abbildung 9). Die Herstellung von BAC-Arrays nach einer Aufreinigung der BAC-DNA durch PFGE und einer anschließenden Amplifikation mittels Adapter-Linker-PCR³⁷ ermöglichte es, genomische Veränderungen in Einzelzell-DNA nachzuweisen.

Etablierung der Spotting-Bedingungen zur Herstellung des 3K BAC-Arrays

Mit einem 120 BAC-Klone umfassenden Klon-Set wurden die Spotting-Parameter zur Herstellung eines zweiten Test-Arrays optimiert. Dabei wurde festgestellt, dass sowohl die Oberfläche auf die gespottet wurde als auch die Aufreinigung der amplifizierten BAC-DNA und der verwendete Spottingpuffer einen erheblichen Einfluss auf die Spotmorphologie hatten. Epoxid-beschichtete Glasobjektträger wurden aufgrund der kleinen Spotdurchmesser ausgeschlossen. Auf Amino-Poly-Silan-(APS)-beschichteten Objektträgern konnten hingegen Spotdurchmesser von 100-120 µm erreicht werden. Die Hintergrundfluoreszenz von Epoxid-beschichteten Objektträgern war zwar geringer als die von APS-beschichteten Objektträgern, dieser Effekt konnte aber, wie in der Literatur beschrieben, durch eine BSA-haltige Blockierung ausgeglichen werden⁸⁶. Bei der Aufreinigung der BAC-DNA nach der Amplifikation wurde festgestellt, dass nur eine Ethanol-Fällung mit anschließendem Waschen die PCR-Fragmente zuverlässig aufreinigen und gleichzeitig störende Salze eliminieren kann. Um sowohl eine homogene Spotmorphologie als auch gutes Hybridisierungsverhalten zu erhalten, stellte sich der *sciSPOT* Oligo-Spottingpuffer als geeignet heraus (Abbildung 10). Da die Zusammensetzung dieses Puffers nicht bekannt ist, können keine vergleichenden Aussagen zu den übrigen getesteten Puffern getroffen werden. Abschließend lässt sich sagen, dass neben den soeben aufgeführten Parametern auch die Expertise beim Spotten eines Arrays eine Rolle spielte. Die Kooperation mit dem auf Array-Herstellung ausgerichteten Unternehmen Scienion führte letztlich zur erfolgreichen Herstellung des BAC-Arrays.

Herstellung des 3K BAC-Arrays

Mit diesen Parametern wurde mit der Herstellung des 3K BAC-Arrays begonnen, der 2940 BAC-Klone umfasste, die in regelmäßigen Abständen von 1 Mb über das Genom verteilt waren, so dass eine durchschnittliche Auflösung von 1Mb erreicht wurde. Um sowohl die Amplifikation als auch die Ethanol-Fällung der BAC-DNA reproduzierbarer und schneller durchführen zu können, konnten diese Prozesse automatisiert werden, ohne dass eine Verringerung der Qualität im Vergleich zum manuellen Prozess festgestellt werden konnte. Diese Automatisierung stellt vor allem auch deshalb einen Vorteil dar, da so auch größere Mengen von BAC-Klonen, z.B. bei der Herstellung größerer Arrays, bearbeitet werden können. Mit der Präparation der BAC-Klone des 3K-Sets stellte sich die Frage nach der Überprüfung der Identität der einzelnen Klone. Es wird davon ausgegangen, dass bis zu 30% der Klone in BAC-Bibliotheken falsch annotiert sind (persönliche Mitteilung von Richard Reinhardt, MPI für Molekulare Genetik, Berlin). Da falsche Positionsangaben von BAC-Klonen die Ergebnisse von Array-CGH-Experimenten verfälschen, wurde das gesamte 3K BAC-Set einer End-Sequenzierung mit dem T7-Primer und anschließendem Sequenzabgleich unterworfen. Von den 2940 BAC-Klonen waren 11,2% falsch annotiert, was in guter Übereinstimmung mit dem erwarteten Wert war.

Optimierung der Hybridisierungsbedingungen für den 3K BAC-Array

Nach der Fertigstellung des Arrays wurden zunächst die Hybridisierungsbedingungen so optimiert, dass die unspezifische Bindung so gering wie möglich und das spezifische Signal so hoch wie möglich war. Dabei sollte die Rate von falsch-positiv und -negativ hybridisierenden Klonen idealerweise null betragen. Die bestmögliche Verringerung der unspezifischen Bindungen wurde durch eine Blockierung mit unmarkierter Leerkontrolle der Amplifikations-PCR erreicht. Dieses Ergebnis lässt sich dadurch erklären, dass bei der Adapter-Linker-Amplifikation der BAC-Klone vor dem Spotten geringe – wenn auch im Gegensatz zu konventionellen Präparationsmethoden kaum nachweisbare – Mengen bakterieller Kontamination eingeführt werden, die zu einer unspezifischen Bindung führen. Diese Bindung lässt sich durch die Prähybridisierung mit Leerkontrolle am besten blockieren (Abbildung 12). Für die Hybridisierung von BAC-Arrays lassen sich in der Literatur ebenso viele Protokolle wie Veröffentlichungen finden^{35, 37, 38, 50, 72, 85, 91, 97, 98}. Aus diesem Grund wurden verschiedene grundlegende Parameter getestet, um festzustellen, welches Protokoll für den 3K BAC-Array am besten geeignet war (Abbildung 12 und Tabelle 9). Obwohl die

Mehrheit der publizierten Protokolle eine Hybridisierungstemperatur von 37°C vorschlägt, wurden bei dieser Temperatur mit dem 3K BAC-Array die schlechtesten Ergebnisse erzielt. Eine Verbesserung ergab sich durch die Erhöhung der Temperatur auf 42°C, was in Anlehnung an die Metaphasen-CGH mit Einzelzellen⁷² geschah. Die Hybridisierungsergebnisse konnten durch die Erhöhung der Menge an humaner Cot-1-DNA zur Blockierung der repetitiven Sequenzen in den Proben auf 80 µg noch verbessert werden⁹⁸, eine weitere Erhöhung auf 150 µg³⁸ führte jedoch zu einem schlechteren Ergebnis. Diese Verschlechterung lässt sich durch eine Abnahme der Gesamtsignalintensität erklären, die durch eine übermäßige Blockierung zustande kommen könnte. Aufgrund der Größe der zu hybridisierenden Fläche wurde auf eine Hybridisierungsstation zurückgegriffen, die zu einer homogeneren Hybridisierung und höheren Gesamtsignalintensitäten führte. Die Signalintensitäten lagen mit durchschnittlich 17129 ±906 rfu innerhalb der in der Literatur beschriebenen Grenzen (ca. 20000 rfu)⁸⁶.

4.2 Charakterisierung des 3K BAC-Arrays

Hybridisierung des 3K BAC-Arrays mit nicht-amplifizierter DNA

In der anschließenden Charakterisierung des Hybridisierungsverhaltens wurde zunächst untersucht, ob der 3K BAC-Array als Hybridisierungsplattform mit bereits publizierten BAC-Arrays vergleichbar ist. Hierfür wurde nicht-amplifizierte DNA von gesunden Probanden nach Standardprotokollen hybridisiert (Abbildung 14). Die balancierten BAC-Klone zeigten ein durchschnittliches Log₂-Verhältnis von -0,01 ±0,15, die Streuung war damit im Vergleich zu publizierten BAC-Arrays höher (±0,15 im Vergleich zu ±0,08³⁷ und ±0,02⁹⁹). Hierbei muss allerdings angemerkt werden, dass die Streuung in den meisten BAC-Array-Publikationen nicht erwähnt wird und deshalb kein ausführlicher Vergleich mit Literaturdaten möglich ist. Obwohl die Streuung des 3K BAC-Arrays also höher als in einigen anderen Arbeiten war, konnte der erwartete normale Karyotyp in den untersuchten Proben dargestellt werden. Eine mögliche Ursache hierfür liegt evtl. darin, dass der gesamte 3K-Klonsatz in die Bewertung einbezogen wurde, wohingegen andere Arbeitsgruppen nach langjähriger Erfahrung eine Auswahl an Klonen identifizieren konnten, die die besten Hybridisierungseigenschaften besitzen.

Untersuchung von Artefakten bei der Hybridisierung von Einzelzell-Amplifikaten

Bei der weiteren Charakterisierung des 3K BAC-Arrays wurden bei der Hybridisierung von Einzelzell-Amplifikaten BAC-Klone mit systematischem Fehlverhalten identifiziert und von der weiteren Auswertung ausgeschlossen. Dabei wurden zuerst wiederholt (in 40% der Experimente) als Ausreißer identifizierte Klone eliminiert, da diese offensichtlich schlechte Hybridisierungseigenschaften besaßen, die, auch wenn sie nicht als Ausreißer identifiziert werden, trotzdem das Gesamtergebnis verfälschen können (Abbildung 16 und Tabelle 11).

Das Chromosom 19 zeigte in Kontrollexperimenten mit normaler DNA aus Einzelzellen stets eine Deletion an und wurde deshalb zunächst als unzuverlässige Region eingestuft. Dieses Hybridisierungsverhalten ist von der Metaphasen-CGH bekannt, bei der angenommen wird, dass die Begründung im hohen Anteil nicht-blockierbarer repetitiver Sequenzen liegt^{33, 95}. Im Fall des 3K BAC-Arrays wurden jedoch von Anfang an BAC-Klone mit hohem Anteil an repetitiven Sequenzen ausgeschlossen (<http://bacpac.chori.org/mapped-clones.htm>). Außerdem wurde Chromosom 19 als balanciert angezeigt, wenn mit nicht-amplifizierter DNA aus normalen Zellen hybridisiert wurde (Abbildung 14). Allerdings ist ein hoher GC-Gehalt des Chromosoms 19 bekannt, was eine Auswirkung auf den Restriktionsverdau mit dem Enzym Mse I (T[▼]TAA) vor der Adapter-Linker-Amplifikation haben könnte. In einer vergleichenden Untersuchung von GC-Gehalt und Mse I-Schnittstellen wurde zum einen gefunden, dass sich BAC-Klone des Chromosoms 19 sowohl im GC-Gehalt als auch in der Anzahl der Mse I-Schnittstellen deutlich von Chromosom 3-Klonen unterschieden (Abbildung 17). Zum anderen konnte eine enge Korrelation zwischen beiden Parametern identifiziert werden (Abbildung 18). Im Schnitt waren Mse I-Fragmente aus Chromosom 19-BAC-Klonen fast doppelt so lang (ca. 300 bp) wie die der Klone von Chromosom 3 (ca. 150 bp). Aus diesen Angaben ist allerdings die Größenverteilung der Mse I-Fragmente nicht ersichtlich. Zunächst konnte aber gezeigt werden, dass BAC-Klone von Chromosom 19, deren GC-Gehalt und Mse I-Fragmentlänge mit BAC-Klone des Chromosoms 3 vergleichbar waren, bessere Hybridisierungseigenschaften besitzen und bei der Hybridisierung normaler Einzelzellen als balanciert angezeigt wurden (Abbildung 19).

Eine Begründung für das beobachtete Hybridisierungsverhalten könnte sein, dass lange Fragmente in einer PCR schlechter amplifiziert werden als kurze Fragmente⁸⁹. Dagegen spricht jedoch, dass die BAC-Klone des Chromosoms 19 eine durchschnittliche Fragmentlänge von 300 bp aufwiesen, eine Länge, die in einer PCR problemlos amplifiziert wird. Um Aussagen über ein mögliches Amplifikationsartefakt treffen zu können, müsste die Größenverteilung der Mse I-Fragmente ermittelt werden. Weiterhin ist es denkbar, dass die BAC-Klone von Chromosom 19 aus bislang unbekanntem Gründen ein schlechteres Hybridisierungsverhalten zeigen oder eine schlechtere Bindung an die Oberfläche des Arrays haben und der GC-Gehalt lediglich einen Surrogat-Marker darstellt. Dennoch konnte ein Kriterium identifiziert werden, mit dem es möglich ist, BAC-Klone mit guten Hybridisierungseigenschaften vorherzusagen und zu selektionieren. Ein geplanter Schritt in der Weiterentwicklung des 3K BAC-Arrays ist die Untersuchung aller BAC-Klone anhand dieses Kriteriums. Besonders das Chromosom 17 könnte von dieser Untersuchung profitieren, da es bei Hybridisierungen von Amplifikaten aus normalen Einzelzellen zwar selten als deletiert angezeigt wird, aber ein durchschnittliches Log₂-Verhältnis von $-0,24 \pm 0,04$, im Gegensatz zu $-0,02$ bei den übrigen Chromosomen, aufweist (Anhang 2). Aus diesem Grund wurde neben dem Chromosom 19 auch Chromosom 17 aus den Auswertungen ausgeschlossen (Anhang 3). Diese ersten Befunde zeigen, dass eine systematische Analyse einzelner BAC-Klone dazu führen kann, Klone zu identifizieren, die gute oder schlechte Hybridisierungseigenschaften besitzen.

Charakterisierung der Hybridisierung von Einzelzell-Amplifikaten auf dem 3K Array

In den folgenden Hybridisierungen mit Adapter-Linker-amplifizierter DNA aus Einzelzellen gesunder Probanden und eines Patienten mit Down Syndrom konnte gezeigt werden, dass weitgehend balancierte Profile angezeigt wurden und die Trisomie 21 dargestellt werden konnte (Abbildung 20). Anhand der Hybridisierungsprofile, die mit amplifizierter Einzelzell-DNA erhalten wurden, sollten Angaben zum Hybridisierungsverhalten des 3K BAC-Arrays abgeleitet werden. Dazu gehörten unter anderem die Bestimmung der experimentellen Streuung und der Rate der BAC-Klone mit nicht-erwartetem Hybridisierungsverhalten. Die Grenzwerte für eine statistisch relevante Auslenkung hängen von der Streuung ab, so müssen bei einer größeren Streuung auch die Grenzwerte, die eine signifikante Aberration charakterisieren, höher liegen. Die experimentelle Streuung (Standardabweichung) wurde in fünf Hybridisierungsexperimenten mit amplifizierter DNA aus normalen Einzelzellen

ermittelt (Log2-Verhältnis: $-0,02 \pm 0,22$). Im Laufe der Anfertigung dieser Arbeit wurden verschiedene Arbeiten veröffentlicht, in denen Array-CGH Analysen mit sehr wenigen^{92, 100} bzw. einzelnen Zellen unternommen wurden^{91, 97, 101}. In diesen Untersuchungen betragen die Standardabweichungen $\pm 0,32$ ⁹¹, $\pm 0,37$ ¹⁰¹, $\pm 0,6$ ¹⁰⁰ und $\pm 0,84$ ⁹⁷, im Vergleich dazu war die Standardabweichung des 3K BAC-Arrays mit $\pm 0,22$ niedriger.

Aus den Log2-Verhältnissen von Hybridisierungen weiblicher und männlicher DNA lassen sich die durchschnittlichen Log2-Verhältnisse für die Autosomen, das X- und das Y-Chromosom bestimmen und mit den entsprechenden Literatur-Werten vergleichen (Literatur: Autosomen = $0,01$ ³⁷, X = $0,72$ ³⁷, $0,79$ ³⁶, Y = $-1,66$ ³⁷). Für den 3K BAC-Array betrug das durchschnittliche Log2-Ratio für die Autosomen $-0,02 \pm 0,22$, für das X-Chromosom $0,66 \pm 0,33$ und für das Y-Chromosom $-1,86 \pm 0,71$. Diese Werte sind vergleichbar mit den Literatur-Werten. Die Abweichung von den theoretischen Werten (X-Chromosom-Verhältnis von 2:1 und Y-Chromosom-Verhältnis von 1:2) lässt sich durch ähnliche repetitive Sequenzen auf den X- und Y-Chromosomen erklären³⁶ und wird allgemein bei Array-CGH-Untersuchungen beobachtet^{35, 37}.

Die Rate von BAC-Klonen, die nicht das erwartete Hybridisierungsverhalten zeigten, wurde bestimmt. In Hybridisierungen von Amplifikaten aus normalen Einzelzellen zeigten 1,1% der BAC-Klone Aberrationen an. Falsch-negativ hybridisierende BAC-Klone konnten in Hybridisierungen von Zellen mit bekanntem Karyotyp nicht detektiert werden. BAC-Klone, die in Hybridisierungen verschiedener normaler Einzelzellen als aberrant detektiert wurden, wurden als unzuverlässige Regionen gekennzeichnet und nicht in der Auswertung von Patientenproben berücksichtigt.

Bei Hybridisierungen von nicht-amplifizierter DNA tritt auf dem 3K BAC-Array bei konventionellen Array-CGH-Experimenten eine relativ starke experimentelle Streuung auf (Log2-Verhältnis $\pm 0,15$ im Gegensatz zu $\pm 0,02$ ⁹⁹, $\pm 0,08$ ³⁷ und $\pm 0,11$ ⁹¹). Interessanterweise nahm diese Streuung bei Hybridisierungen von amplifizierter Einzelzell-DNA auf den 3K BAC-Array lediglich um $\pm 0,07$ zu, wohingegen die Streuung in dem von Fiegler et al. publizierten Protokoll um $\pm 0,21$ ⁹¹ zunimmt. Hinzu kommt, dass die von Fiegler et al. beobachtete Streuung bei Hybridisierungen von Einzelzell-Amplifikaten nach Glättung des Datensatzes um den Faktor zehn bestimmt wurde, wohingegen sich die Streuung des 3K BAC-Arrays aus ungeglätteten Daten berechnete. Die stärkere Streuung bei nicht-amplifizierter DNA könnte also auf dem nicht-selektierten BAC-Satz beruhen. Durch die systematische Analyse der BAC-Klone des 3K BAC-Arrays ist eine Verminderung der

Streuung zu erwarten.

Entwicklung eines Assays zur Vorhersage einer erfolgreichen Array-CGH-Analyse

Im Verlauf der Hybridisierung des 3K BAC-Arrays mit amplifizierter DNA aus normalen Einzelzellen wurde festgestellt, dass mit einigen Amplifikaten bessere und mit anderen schlechtere Array-CGH-Profile erhalten wurden, obwohl mit allen Amplifikaten in der Metaphasen-CGH balancierte Profile gemessen worden waren. Dabei waren Kriterien für erfolgreiche Array-CGH-Hybridisierungsergebnisse, dass eine normale Streuung der Profile vorlag ($\text{Log}_2\text{-Verhältnis: } -0,02 \pm 0,22$) und ein balancierter Karyotyp angezeigt wurde. Da die Qualität der Hybridisierungen und die Gesamtsignalintensitäten vergleichbar waren, wurde untersucht, ob qualitative Unterschiede bei den hybridisierten Proben festgestellt werden konnten. Stoecklein et al. zeigten in LOH-Analysen von DNA aus fixiertem Paraffingewebe, dass Mse I-Fragmente, die länger als 400 bp sind, bei schlechter DNA-Qualität nicht nachgewiesen werden konnten⁸⁸. Da die Qualität der DNA aus Einzelzellen wesentlich besser ist als aus fixiertem Paraffingewebe, wurde die Nachweisbarkeit von fünf Mse I-Fragmenten, die länger als 1000 bp waren, in Einzelzell-Amplifikaten untersucht. Dabei konnte ein direkter Zusammenhang zwischen dem Nachweis der langen Mse I-Fragmente und einem Gelingen der Array-CGH festgestellt werden (Abbildung 21 und Tabelle 12). Der entwickelte PCR-basierte Test konnte mit 90%iger Genauigkeit vorhersagen, ob eine Array-CGH-Untersuchung erfolgreich verlaufen würde oder nicht. Bei der Untersuchung von 119 Einzelzell-Amplifikaten wurde gefunden, dass etwa 16% der Proben voraussichtlich nicht für eine Array-CGH-Analyse geeignet sind, für weitere 39% ist das Gelingen der Array-CGH-Analyse unsicher. Der PCR-basierte Test stellt insofern einen Vorteil dar, da Array-CGH-Experimente mit hohen Kosten verbunden sind.

Eine Begründung für den Einfluss der langen Mse I-Fragmente auf das Hybridisierungsergebnis könnte sein, dass während der PCR-basierten Markierung der DNA vor der Hybridisierung in der Summe zwar ähnlich viel Farbstoff in lange wie in kurze Fragmente eingebaut wird, die Einbaurate in kurze Fragmente aber sequenzabhängig größeren stochastischen Schwankungen unterliegt. Dafür spricht, dass weder konsequent geringere Einbauraten noch schwächere Gesamthybridisierungsintensitäten bei Proben gemessen wurden, in denen lange Mse I-Fragmente nicht nachgewiesen werden konnten. Der gleiche fehlende Zusammenhang zwischen Hybridisierungssignalintensität und erfolgreichem Ausgang eines Array-CGH-Experiments wurde auch von van Beers et al. in Experimenten

mit fixierten Parafingewebe beobachtet¹⁰². Bis jetzt stellt der fehlende Nachweis langer Mse I-Fragmente aber lediglich einen Surrogatmarker dar.

4.3 Der 3K BAC-Array im Vergleich zu Oligonukleotid-Arrays

Untersuchung komplexer Genome mit dem 3K BAC-Array

Um untersuchen zu können, ob auch komplexe genomische Veränderungen zuverlässig mit dem 3K BAC-Array detektiert werden können, wurde die Mammakarzinom-Zelllinie T47D als Modell ausgewählt. T47D war zum einen mit verschiedenen BAC-Arrays untersucht worden^{37, 87}, zum anderen zeigt diese Zelllinie eine relativ geringe Variabilität innerhalb der Population (C. H. Hartmann, unveröffentlichte Daten), so dass einzelne durch Metaphasen-CGH untersuchte T47D-Zellen den gleichen Karyotyp aufwiesen.

Hybridisierungsprofile von nicht-amplifizierter T47D-DNA auf dem 3K BAC-Array waren in Übereinstimmung mit den publizierten Profilen^{37, 87}, zudem konnten spezifische Aberrationen ebenfalls detektiert werden¹⁰³ (Abbildung 22a, Abbildung 23). Bei der Hybridisierung amplifizierter DNA aus einzelnen T47D-Zellen wurde eine gute Übereinstimmung mit den Profilen der nicht-amplifizierten T47D-DNA gefunden, wie durch hohe Korrelationskoeffizienten (0,98 und 0,83 für zwei verschiedene Einzelzellen) gezeigt wurde (Tabelle 13). Auch in den Einzelzell-Profilen konnte eine eindeutige Übereinstimmung von Gewinnen und Verlusten mit den Literaturdaten gezeigt werden (Abbildung 23). Dabei konnten unter anderem eine 14,8 Mb und eine 3,5 Mb große Deletion auf dem Chromosom 10 und ein vergleichbares Aberrationsmuster auf dem Chromosom 3 dargestellt werden, was belegt, dass auch relativ kleine Veränderungen in amplifizierter DNA aus Einzelzellen zuverlässig mit dem 3K BAC-Array detektiert werden.

Vergleich des 3K BAC-Arrays mit Oligonukleotid-Arrays

Der Arbeitsaufwand für die Herstellung eines BAC-Arrays mit reduzierter Kontamination ist hoch, und dementsprechend ist eine Erhöhung der Klonzahl, um eine höhere Auflösung zu erreichen, mühsam. Deshalb wurde untersucht, ob es möglich ist, Array-CGH-Untersuchungen mit Adapter-Linker amplifizierter Einzelzell-DNA auf kommerziellen

Oligonukleotid-Arrays durchzuführen, da auch sie das Kriterium einer minimalen Kontamination mit bakteriellen Sequenzen erfüllen. Array-CGH mit Oligonukleotid-Arrays wurde im Laufe der Anfertigung dieser Arbeit seit dem Jahr 2003 beschrieben^{27, 41, 42, 90}.

Um sicher zu stellen, dass mit den verwendeten Hybridisierungsprotokollen zuverlässige CGH-Untersuchungen vorgenommen werden können und um einen anschließenden Vergleich mit dem 3K BAC-Array zu ermöglichen, wurde zunächst nicht-amplifizierte DNA aus 10^6 T47D-Zellen auf zwei verschiedene 70-mer-Oligonukleotid-Arrays hybridisiert. Zum einen wurde der 35K Oligonukleotid-Array der Firma Operon, zum anderen der 244K Oligonukleotid-Array der Firma Agilent verwendet. Es konnte gezeigt werden, dass beide Hybridisierungsprofile mit den publizierten Profilen übereinstimmten (Abbildung 24a, Abbildung 25 und Abbildung 26a, Abbildung 27). Hybridisierungen von amplifizierter Einzelzell-DNA führten allerdings zu Profilen ohne Übereinstimmung mit publizierten Profilen und denen der nicht-amplifizierten DNA, wie durch niedrige Korrelationskoeffizienten und ROC-Analysen gezeigt werden konnte (Tabelle 14 und Abbildung 28 (3K BAC-Array: AUC=0,79, 35K Oligonukleotid-Array: AUC=0,53, 244K Array: AUC=0,54)). Da sowohl der 35K Oligonukleotid-Array als auch der 244K Oligonukleotid-Array aufgrund der Menge an Oligonukleotiden eine um den Faktor 6-80 höhere Auflösung ermöglichten als der 3K BAC-Array, wurde der Versuch unternommen, die Daten zu glätten, also viele Datenpunkte zu einem neuen Datenpunkt zusammenzufassen. Dieses Vorgehen wurde beispielsweise auch von Fiegler et al. angewendet⁹¹ und führte dort zu einer reduzierten Streuung der Daten und einer verbesserten Detektion genomischer Veränderungen.

Um sowohl eine Glättung der Daten zu erreichen als auch die Detektion von kleinen Aberrationen geringer Amplitude zu ermöglichen, wurde der Savitzky-Golay-Filter⁸¹ zur Glättung angewendet. Bei den Oligonukleotid-Arrays wurden jeweils so viele Datenpunkte zusammengefasst, dass die rechnerische Auflösung aller drei Plattformen vergleichbar war. Durch ROC-Analysen wurde überprüft, ob die Glättung in den T47D-Einzelzell-Amplifikaten zu einer verbesserten Darstellung des Karyotyps führte (Abbildung 29a). Auf allen drei Plattformen wurde eine verbesserte Darstellung beobachtet, den deutlichsten Effekt hatte die Glättung jedoch auf die Daten der 3K BAC- und 244K Oligonukleotid-Arrays (3K BAC-Array: AUC=0,93, 244K Oligonukleotid-Array: AUC=0,85).

Die ROC-Analyse kann allerdings nur zeigen, wie gut die Struktur des bekannten Referenz-Datensatzes aus der eines Test-Datensatzes rekonstruiert werden kann. Wird mit einem Array

eine gute Übereinstimmung der Datenstruktur in der vergleichenden ROC-Analyse gefunden, bedeutet das nicht, dass der Karyotyp einer unbekannt Probe anhand der gemessenen Log₂-Verhältnisse auch korrekt identifiziert wird. Bei der Erstellung vergleichender Profile zwischen amplifizierter Einzelzell-DNA und nicht-amplifizierter DNA auf den verschiedenen Arrays wurde deutlich, dass nur mit dem 3K BAC-Array der T47D-Karyotyp in Einzelzell-Amplifikaten dargestellt werden kann (Abbildung 29b).

Eine mögliche Erklärung hierfür könnte in der Natur der hybridisierten Einzelzell-Amplifikate liegen. Werden diese Amplifikate auf BAC-Klone hybridisiert, die durch ein analoges Protokoll amplifiziert wurden, können vermutlich jeweils Fragmente entsprechender Länge miteinander hybridisieren. Es könnte aber auch sein, dass ganz bestimmte Fragmente sowohl in den BAC-Klonen als auch in den Einzelzellproben hauptsächlich zum Gelingen der Hybridisierung beitragen. In diesem Fall wären diese Fragmente zwar nach wie vor in der Probe enthalten, nicht aber auf der Oligonukleotid-Plattform. Es ist allerdings auch anzunehmen, dass die Verwendung von Oligonukleotid-Arrays für die Einzelzell-CGH noch dadurch verbessert werden kann, dass die Oligonukleotide für eine optimale Hybridisierung von Mse I-Fragmenten angepasst werden. Diese Methode wurde auch bei der ROMA-Technologie²⁷ verwendet, bei der die Sequenzen der Oligonukleotide auf dem Array so gewählt wurden, dass sie an die Fragmente angepasst waren, die beim Restriktionsverdau der zu hybridisierenden DNA entstanden waren. Allerdings ermöglicht die ROMA-Technologie keine erfolgreiche Hybridisierung von Einzelzell-Amplifikaten (persönliche Mitteilung von Michael Wigler), so dass fraglich ist, ob dieser Ansatz für eine mögliche Optimierung der Einzelzell-Array-CGH mit Oligonukleotid-Arrays verfolgt werden sollte. Aus den gezeigten Daten lässt sich außerdem ableiten, dass eine weitere Erhöhung der Oligonukleotidzahl zu einer weiteren Verbesserung führen könnte, da mit dem 244K Array im Gegensatz zum 35K Array (ca. 12x weniger Oligonukleotide) zumindest die Struktur der Daten erkennbar war.

4.4 Array-CGH-Untersuchungen von disseminierten Tumorzellen

Validierung von Alterationen in Einzelzell-Amplifikaten

Um zu demonstrieren, dass Array-CGH-Untersuchungen auch von amplifizierter DNA aus einzelnen disseminierten Tumorzellen aus dem Knochenmark von Brustkrebspatientinnen möglich ist, wurden fünf einzelne Tumorzellen von drei Patientinnen auf dem 3K BAC-Array hybridisiert und mit den Ergebnissen der Metaphasen-CGH verglichen (Abbildung 30 und

Anhang 1). Aberrationen, die in der Metaphasen-CGH detektiert wurden, wurden auch durch die Array-CGH gefunden, wie die hohen Korrelationskoeffizienten (0,91, 0,92 und 0,94, 0,94 und 0,88) zeigten. Zusätzlich zu diesen wurden viele weitere Veränderungen entdeckt, die zu klein waren, um sie in der Metaphasen-CGH detektieren zu können, darunter eine Deletion von 4,7 Mb auf Chromosom 11q (Abbildung 31). Außerdem unterschieden sich die Profile in der Lokalisation der Bruchpunkte.

Um ausgewählte chromosomale Aberrationen mit einer unabhängigen Methode zu verifizieren, wurde zunächst mit einer Allelverlust-Analyse in Anlehnung an Cardoso et al.⁹² die Region 11q13.4 bis 11q14.1 (Abbildung 31) untersucht. Das Muster der in beiden disseminierten Tumorzellen detektierten Gewinne und des 4,7 Mb-Verlusts sollte so überprüft werden. Um diese Verifizierung erfolgreich durchführen zu können, wurde vorausgesetzt, dass der Prozess der genomischen Umlagerung auf eine bestimmte Weise stattgefunden hatte: Ursprünglich müsste ein diploider Chromosomensatz, also zwei Allele, vorgelegen haben. In den amplifizierten Regionen sollten beide Allele noch erhalten und dementsprechend nachweisbar sein. In den deletierten Regionen sollte nur noch eines von zwei Allelen erhalten und nachweisbar sein. Das Ergebnis der LOH-Untersuchung stimmte in zwei der drei untersuchten Regionen nicht mit dem erwarteten Ergebnis überein (Tabelle 15). In Region 4 wurden statt des erwarteten Allelverlusts zwei Allele nachgewiesen, in Region 3 wurde nur eines anstelle von zwei Allelen nachgewiesen. Dieser Befund legt nahe, dass die Zelle ihr Genom dupliziert hatte, bevor es zur Deletion kam, so dass in Region 4 nach wie vor die beiden verschiedenen Allele nachweisbar blieben. Wahrscheinlich liegt in den untersuchten Zellen ein polyploider Chromosomensatz vor, dessen genomische Umstrukturierung komplexer ist als erwartet war. Die Verifizierung mittels Allelverlust-Analyse war daher nicht geeignet.

Die unabhängige Bestätigung von zwei 4 Mb und 5,5 Mb großen Amplifikationen auf den Chromosomen 14 und 12 wurde durch eine qPCR-Analyse⁷⁵ erreicht (Tabelle 16). Die Kopienzahlen der beiden Loci wurden in Relation zu balancierten Kontroll-Loci auf 4-5 (*AKAP3*) und 5-7 Kopien (*AKAP6*) berechnet. Da jedoch aus der vorherigen LOH-Untersuchung abgeleitet werden kann, dass ein polyploider Chromosomensatz vorlag, ist die tatsächliche Kopienzahl der *AKAP3*- und *AKAP6*-Loci vermutlich höher. Aus diesem Ergebnis konnte abgeleitet werden, dass kleine Aberrationen, die unter dem Auflösungs-niveau der Metaphasen-CGH liegen, mit dem 3K BAC-Array zuverlässig in Einzelzell-Amplifikaten nachgewiesen werden können.

Mit Hilfe des 3K BAC-Arrays konnten also in Adapter-Linker amplifizierter DNA aus Einzelzellen 4 Mb große Gewinne bzw. Amplifikationen und 3,5-4,7 Mb große Deletionen detektiert werden. Im Gegensatz dazu besitzen die mittlerweile publizierten Array-CGH-Einzelzell-Protokolle eine schlechtere Auflösung. Hu et al. konnten mit DOP-PCR-amplifizierter Einzelzell-DNA lediglich Aneuploidien ganzer Chromosomen zeigen¹⁰¹. Le Caignec et al. hybridisierten multiple-strand-displacement-(MDA)-amplifizierte Einzelzell-DNA auf einem 4K BAC-Array und konnten eine Deletion von 34 Mb Größe dann nachweisen, wenn vorher bekannt war, dass diese Aberration vorhanden war⁹⁷. Unbekannte Aberrationen mussten eine Größe von mindestens 60 Mb aufweisen, damit diese detektiert werden konnten⁹⁷. Im Vergleich dazu die besten Ergebnisse erzielten Fiegler et al., die Genome-Plex®-amplifizierte Einzelzell-DNA auf einen 32K-BAC-Array hybridisierten und Aberrationen von 10,8 Mb aber nicht von 6,3 Mb Größe detektieren konnten⁹¹. Dieses Ergebnis wurde allerdings nur mit einer vorangegangenen Glättung der Hybridisierungsdaten erzielt. Obwohl aber in den beiden zuletzt beschriebenen Arbeiten BAC-Arrays mit teilweise wesentlich größerer Auflösung verwendet wurden, wurden mit Adapter-Linker-amplifizierter Einzelzell-DNA auf dem 3K BAC-Array um mindestens den Faktor 3 kleinere Aberrationen detektiert. Die bessere Leistung des 3K BAC-Arrays bei der Untersuchung von Einzelzell-DNA ist möglicherweise sowohl durch die homogenere Amplifikation der Einzelzell-DNA durch die Adapter-Linker-PCR als auch durch die reduzierte Kontamination des 3K BAC-Arrays bedingt. Hinzu kommt, dass die BAC-DNA für die Herstellung des 3K BAC-Arrays ebenfalls mittels Adapter-Linker-PCR amplifiziert wurde, so dass sowohl auf dem Array als auch in der hybridisierten Probe Mse I-Fragmente gleicher Länge miteinander hybridisieren können.

Bedeutung der Einzelzell-Array-CGH-Technologie

Die Entwicklung von der Metaphasen-CGH zur Array-CGH führte seit 1997³⁴ zu einer deutlichen Zunahme der Auflösung, mit der genomische Aberrationen gemessen werden können. Die technische Auflösung, die inzwischen mit BAC-Arrays erreicht wird, die das Genom überlappend abdecken, liegt bei 140 kb⁴³. In der Tumorbologie wird diese Technik in der Regel dazu verwendet, DNA aus primären Tumoren zu untersuchen, wobei DNA aus 50.000-500.000 Zellen isoliert wird⁴⁴. Zwar wurden seitdem verschiedene Array-CGH-Studien, in denen Brusttumore untersucht wurden, durchgeführt^{49, 51, 104-107}, der Informationsgewinn gegenüber Studien mit der konventionellen Metaphasen-CGH war

allerdings nur gering. Dies ist dadurch begründet, dass in fortgeschrittene Tumoren umfassende genomische Veränderungen selektiert werden, die oftmals ganze Chromosomen oder Chromosomen-Arme umfassen. Diese Art von Aberrationen kann auch durch die Metaphasen-CGH detektiert werden und benötigt nicht die hohe Auflösung der Array-CGH. Nur wenige neue Kandidatengene oder –Regionen wurden durch die erhöhte Auflösung der Array-CGH zusätzlich ermittelt, dazu zählten die potentiellen Onkogene *CCND3* und *p21/WAF1*¹⁰⁶, *Rab25*⁴⁹, *SPFH2* und *FLJI4299*^{51, 105} und *PROSC*⁵¹. Die hohe Auflösung der Array-CGH führte vielmehr in anderen Forschungsgebieten wie der Humangenetik dazu, dass beispielsweise so genannte 100 kb bis 3 Mb große Kopienanzahl-Polymorphismen (CNVs) entdeckt wurden^{27, 108, 109}, die mit verschiedenen Syndromen in Verbindung gebracht wurden¹¹⁰.

Untersuchungen von disseminierten Tumorzellen von Mammakarzinompatientinnen legen nahe, dass die Tumorzell-Disseminierung früh in der Entwicklung eines Mammakarzinoms erfolgt, wobei die Population der disseminierten Tumorzellen genomisch ausgesprochen heterogen ist^{73, 75}. Die genomische Entwicklung von Primärtumoren und disseminierten Tumorzellen wird frühzeitig entkoppelt, was durch eine fehlende klonale Verwandtschaft auf chromosomaler Ebene belegt wird⁷³. Disseminierte Tumorzellen als die Vorläufer der Metastasen⁶⁵ stellen die Zielzellen adjuvanter Therapien dar. Unklar ist deshalb, inwieweit die Untersuchung von Primärtumoren zur Identifizierung von Zielstrukturen für molekulare Therapien führen kann. Durch die genomische Charakterisierung von disseminierten Tumorzellen hingegen sollten Gene gefunden werden, die für die Metastasierung relevant und damit als Angriffspunkt für Therapien geeignet sind.

In einer vorangehenden Metaphasen-CGH-Studie konnte beim Mammakarzinom genomische Heterogenität der disseminierten Tumorzellen festgestellt werden. Tumorzellen von Patientinnen mit klinisch nachweisbaren Metastasen (M1) wiesen komplexe genomische Veränderungen auf, wohingegen 60% der disseminierten Zellen von Patientinnen ohne manifeste Metastasen (M0) wenige oder keine Aberrationen zeigten^{73, 83}. In diesen, als Metaphasen-CGH-negativ bezeichneten Zellen wurden allerdings durch hochauflösende LOH-Analysen Allelverluste nachgewiesen^{75, 93}, was sie als Tumorzellen kennzeichnet. Diese Zellen weisen vermutlich diejenigen Veränderungen auf, die ihre Disseminierung ermöglichen. Die Etablierung der Einzelzell-Array-CGH-Technologie soll die genomweite und hochauflösende Untersuchung dieses Zelltyps ermöglichen und erlauben, neue Erkenntnisse über den Prozess der Metastasierung zu gewinnen sowie relevante Zielgene zu

identifizieren.

Array-CGH-Veränderungen in disseminierten Tumorzellen

Bei der Untersuchung Metaphasen-CGH-veränderter disseminierter Tumorzellen, die aus dem Knochenmark von M1-Patientinnen isoliert worden waren, wurde gezeigt, dass die Ergebnisse der Metaphasen-CGH reproduziert und darüber hinaus noch zusätzliche Veränderungen detektiert werden konnten (Abbildung 30 und Anhang 1). Anhand eines kleinen Kollektivs von disseminierten Tumorzellen aus dem Knochenmark von M0-Patientinnen wurde gezeigt, dass in Metaphasen-CGH-negativen Zellen genomische Aberrationen detektiert werden können (Anhang 3). Aufgrund der detektierten Veränderungen wurden diese Zellen in Ähnlichkeitsanalysen sowohl von normalen Einzelzellen als auch von Zellen von M1-Patientinnen unterschieden (Abbildung 32 und Abbildung 33). Die Anzahl aberranter BAC-Klone war dabei 48-fach höher als in normalen Einzelzellen und 3,5-fach niedriger im Vergleich zu M1-Zellen (Tabelle 17). Metaphasen-CGH-negative Zellen von M0-Patientinnen weisen dementsprechend nicht nur Allel-Verluste im Submegabasen-Bereich⁷⁵, sondern auch größere Deletionen auf. In allen untersuchten Zellen wurden außerdem Gewinne bzw. Amplifikationen detektiert, die durch die LOH-Technik, die sich nur zur Detektion von Deletionen eignet, nicht nachgewiesen werden können.

Trotz der geringen Probenzahl wurde untersucht, ob in den analysierten disseminierten Tumorzellen aus M0-Patientinnen Regionen identifiziert werden konnten, die in der Mehrheit der Zellen verändert waren (Anhang 3 und Tabelle 18). Um auszuschließen, dass es sich bei den ermittelten Regionen um Artefakte handelte, die ebenfalls in normalen Zellen zu finden sind, wurden Regionen, die in den untersuchten Einzelzellen ebenfalls betroffen waren, aus den Analysen ausgeschlossen. Insgesamt wurden acht Regionen identifiziert, die in mindestens vier von sieben (>50%) untersuchten M0-Zellen verändert waren, bei fünf Regionen handelte es sich um Deletionen, bei drei Regionen um Gewinne bzw. Amplifikationen. Betroffen waren die Chromosomen 5, 7, 10, 11, 13, 15 und 22. Diese Untersuchung zeigt beispielhaft, wie die Array-CGH-Technologie dazu beitragen kann, Aberrationen in seltenen Tumorzellen zu identifizieren. Mit größeren Zell-Kollektiven könnten auf diese Weise relevante Regionen eingegrenzt werden. Eine Erhöhung der Auflösung des BAC-Arrays würde es ermöglichen Gene zu identifizieren, die an Mechanismen der Tumorprogression bzw. der Latenz beteiligt sein können.

Korrelation von Aberrationsmustern und klinischen Parametern

Inzwischen ist anerkannt, dass eine wirkungsvolle systemische Therapie des Mammakarzinoms nur möglich ist, wenn diese an die molekularen Eigenschaften des Tumors angepasst wird. So wird beispielsweise der HER2-Rezeptor-Status des Tumors erhoben, um zu entscheiden, ob eine Therapie mit Trastuzumab Vorteile für die Patientin bringt^{19, 20}. Auch kann aufgrund von Größe und Grad des Tumors und aufgrund des Lymphknoten-Status' der so genannte Nottingham prognostic Index (NPI) berechnet werden, der aussagt, ob eine Chemotherapie für eine Patientin möglicherweise vorteilhaft ist¹¹¹. Das geringe Ansprechen von Mammakarzinompatientinnen auf herkömmliche Therapien und die Heterogenität dieser Erkrankung machen deutlich, dass eine stärkere Subklassifikation der Patientinnen und die Entwicklung von maßgeschneiderten Therapeutika wichtig ist. Sorlie et al. gelang es z.B. aufgrund von Expressionsanalysen neue Subklassen von Mammakarzinomen zu identifizieren, die mit unterschiedlichen klinischen Eigenschaften und Überleben korrelierten¹¹². Ebenfalls durch die Untersuchung des Transkriptoms von Mammatumoren konnten van't Veer et al. eine Signatur von 70 Genen identifizieren, die mit der Prognose der Patientinnen korrelierte¹¹³. Dennoch ist es von diesen Ergebnissen ausgehend nicht möglich, einzelne Gene zu identifizieren, die sich als Zielgene für spezifische Therapieansätze eignen. Des Weiteren ist keine der identifizierten Subklassen prädiktiv. D.h., die Zugehörigkeit einer Patientin zu einer bestimmten Subklassen steht in keinem Zusammenhang damit, ob eine bestimmte Therapie erfolgreich verlaufen wird oder nicht, was für die Klinik von großem Nutzen wäre. Außerdem ist bei diesen Studien problematisch, dass aus frischem Tumorgewebe die äußerst labile mRNA isoliert werden muss, so dass durch unterschiedliche Handhabung der Proben eine große Variabilität eingeführt werden kann. Inzwischen konnte gezeigt werden, dass eine enge Korrelation zwischen der Expression und der Kopienzahl von Genen besteht^{39, 114, 115}, was die Bedeutung genomischer Veränderungen für den Phänotyp einer Tumorzelle belegt. Möglicherweise sind solche Gene, deren Expression aufgrund von Amplifikationen oder Deletionen verändert ist, für die Tumorprogression besonders wichtig. Vor kurzem konnte durch Chin et al. gezeigt werden, dass Muster von durch Array-CGH gemessenen Aberrationen einen unabhängigen prognostischen Marker für verschiedene Mammakarzinomsubklassen darstellen⁵⁰.

Im Rahmen der vorliegenden Dissertation wurde anhand weniger Patientinnen beispielhaft untersucht, ob es möglich ist, Gruppen ähnlicher disseminierter Tumorzellen aufgrund ihrer Array-CGH-Profile voneinander zu unterscheiden. Hierfür wurden die Array-CGH-Daten

einer hierarchischen Ähnlichkeitsanalyse unterworfen. Anschließend wurde überprüft, ob die gebildeten Gruppen mit klinischen Eigenschaften zusammenhängen (Abbildung 34). Innerhalb der sieben analysierten Zellen wurden zwei Gruppen gebildet. Die Gruppen unterschieden sich in der Anzahl der veränderten BAC-Klone, wobei in Gruppe 1 im Schnitt $10,4 \pm 4,0\%$ und in Gruppe 2 $24,9 \pm 3,7\%$ der BAC-Klone verändert waren. In Gruppe 1 waren nur Zellen von Patientinnen enthalten, die einen negativen HER2-Rezeptorstatus aufwiesen, in Gruppe 2 waren drei der vier Patientinnen HER2-positiv. Weitere gemeinsame Eigenschaften, die Hormon-Rezeptoren, den TNM-Status oder das Grading betreffend, konnten nicht gefunden werden. Um Aussagen darüber treffen zu können, ob spezifische Aberrationen mit klinischen Parametern korrelieren und ob bestimmte Aberrationen zu Latenz und andere zu einem aggressiven Krankheitsverlauf führen, ist es deshalb notwendig eine prospektive Studie durchzuführen.

5 Zusammenfassung

Trotz erfolgreicher Resektion des Primärtumors sterben bis heute ca. 40% aller Brustkrebspatientinnen an den Folgen der Metastasierung. Inzwischen ist zunehmend anerkannt, dass wenige selektierte Tumorzellen für die Progression der Erkrankung und eine Therapie-Resistenz verantwortlich sind. Es ist deshalb fraglich, ob genomische Veränderungen, die zur Disseminierung führen, entdeckt werden können, wenn die Masse der Zellen eines heterogenen Primärtumors untersucht wird. Einzelne Tumorzellen disseminieren in ektopy Organe, wie das Knochenmark, und gelten als die Vorläufer der metachronen Metastasen. Eine genetische Analyse dieser Zellen könnte zur Entdeckung neuer Zielstrukturen für adjuvante Therapien führen. Um genomweit und hochauflösend genomische Aberrationen identifizieren zu können, wurde im Rahmen der vorliegenden Dissertation ein Protokoll entwickelt, das Array-CGH-Analysen von einzelnen Zellen ermöglicht. Hierfür war es notwendig ein Verfahren zur Aufreinigung und Amplifikation von BAC-DNA zu entwickeln, die Bedingungen zur Array-Herstellung zu optimieren und ein Hybridisierungsprotokoll zu etablieren. Der daraufhin hergestellte, 3000 BAC-Klone umfassende 3K BAC-Array mit einer mittleren Auflösung von 1 Mb war bei der Analyse amplifizierter DNA aus Einzelzellen konventionellen BAC- und Oligonukleotid-Arrays überlegen. Bei Hybridisierungen von Einzelzell-DNA mit bekanntem Karyotyp wurde bei normalen Zellen ein balanciertes Profil, bei Einzelzell-DNA eines Patienten mit Trisomie 21 die zusätzliche Kopie des Chromosoms 21 und bei Einzelzell-DNA von T47D-Zellen der komplexe Karyotyp in Übereinstimmung mit der Literatur dargestellt. Zusätzlich wurden Kriterien zur Auswahl zuverlässig hybridisierender BAC-Klone identifiziert und ein Assay zur Vorhersage einer erfolgreichen Array-CGH Analyse von Einzelzell-Amplifikaten entwickelt. Bei der Untersuchung einzelner disseminierter Tumorzellen aus dem Knochenmark von Brustkrebspatientinnen wurden sowohl große Alterationen, die durch Metaphasen-CGH detektiert worden waren, als auch kleine zusätzliche Aberrationen zwischen 3,5 und 4,7 Mb Größe detektiert. Ausgewählte chromosomale Gewinne konnten durch eine qPCR unabhängig verifiziert werden. Der 3K BAC-Array erhöht die Auflösung der Metaphasen-CGH um den Faktor 10-30 und ist außerdem publizierten Verfahren zur Array-CGH mit Einzelzellen überlegen. Mit seiner Hilfe war eine hochauflösende Untersuchung gestreuter Metaphasen-CGH-negativer Tumorzellen möglich, was zur Identifizierung bislang unerkannter Veränderungen führte. Die Einzelzell-Array-CGH-Technologie ermöglicht nun die Charakterisierung mikrometastatischer Vorläuferzellen und damit neue Erkenntnisse über frühe Stadien der Krebserkrankung.

6 Literaturverzeichnis

1. Trends, K. i. D.-H. u. (Gesellschaft der epidemiologischen Krebsregister in Deutschland e.V.; in Zusammenarbeit mit dem Robert Koch Institut). (2006).
2. Brustkrebs. (Gesundheitsberichterstattung des Bundes; Robert Koch Institut und Statistisches Bundesamt). (2005).
3. Mammakarzinome. Manuale des Tumorzentrum Münchens; Empfehlungen zu Diagnostik, Therapie und Nachsorge. (2005).
4. Welsch, U. Lehrbuch Histologie. Urban & Fischer (2003).
5. Shackleton, M. et al. Generation of a functional mammary gland from a single stem cell. *Nature* 439, 84-8 (2006).
6. Stingl, J. et al. Purification and unique properties of mammary epithelial stem cells. *Nature* 439, 993-7 (2006).
7. Dontu, G. et al. In vitro propagation and transcriptional profiling of human mammary stem/progenitor cells. *Genes Dev* 17, 1253-70 (2003).
8. Reya, T., Morrison, S. J., Clarke, M. F. & Weissman, I. L. Stem cells, cancer, and cancer stem cells. *Nature* 414, 105-11 (2001).
9. Dontu, G., Liu, S. & Wicha, M. S. Stem cells in mammary development and carcinogenesis, implications for prevention and treatment. *Stem Cell Reviews* 1, 207-214 (2005).
10. Dontu, G., Al-Hajj, M., Abdallah, W. M., Clarke, M. F. & Wicha, M. S. Stem cells in normal breast development and breast cancer. *Cell Prolif* 36 Suppl 1, 59-72 (2003).
11. Bocker, W., Decker, T., Ruhnke, M. & Schneider, W. [Ductal hyperplasia and ductal carcinoma in situ. Definition--classification--differential diagnosis]. *Pathologe* 18, 3-18 (1997).
12. Sobin, L. H. & Wittekind, C. TNM classification of malignant tumors (John Wiley & Sons, Inc., New York, 1997).
13. Elston, C. W. & Ellis, I. O. Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up. *Histopathology* 19, 403-10 (1991).
14. Engel, J., Hölzel, D. & Schubert-Fritschle, G. Epidemiologie. Mammakarzinome; Manuale des Tumorzentrum Münchens; Empfehlungen zu Diagnostik, Therapie und Nachsorge (2005).
15. Funke, I. et al. Prognostische und prädiktive Faktoren beim primären Mammakarzinom. Manuale des Tumorzentrum Münchens; Empfehlungen zu Diagnostik, Therapie und Nachsorge (2005).
16. Engel, J. et al. The process of metastatisation for breast cancer. *Eur J Cancer* 39, 1794-806 (2003).
17. Gabka, C. J. et al. Operative Therapie des primären Mammakarzinoms. Manuale des Tumorzentrum Münchens; Empfehlungen zu Diagnostik, Therapie und Nachsorge (2005).
18. Kahlert, S., Anthuber, C. & Eiermann, W. Adjuvante Systemtherapie. Manuale des Tumorzentrum Münchens; Empfehlungen zu Diagnostik, Therapie und Nachsorge (2005).
19. Stebbing, J., Copson, E. & O'Reilly, S. Herceptin (trastuzumab) in advanced breast cancer. *Cancer Treat Rev* 26, 287-90 (2000).
20. Konecny, G. et al. Quantitative association between HER-2/neu and steroid hormone receptors in hormone receptor-positive primary breast cancer. *J Natl Cancer Inst* 95, 142-53 (2003).
21. Gonzalez-Angulo, A., Hortobagyi, G. & Esteva, F. Adjuvant therapy with trastuzumab for HER2/neu-positive breast cancer. *Oncologist* 11, 857-867 (2006).
22. Bergh, J., Norberg, T., Sjogren, S., Lindgren, A. & Holmberg, L. Complete sequencing of the

- p53 gene provides prognostic information in breast cancer patients, particularly in relation to adjuvant systemic therapy and radiotherapy. *Nat Med* 1, 1029-34 (1995).
23. Kamb, A., Wee, S. & Lengauer, C. Why is cancer drug discovery so difficult? *Nat Rev Drug Disc* 6, 115-120 (2007).
 24. Botstein, D., White, R. L., Skolnick, M. & Davis, R. W. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am J Hum Genet* 32, 314-31 (1980).
 25. Weber, J. L. & May, P. E. Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. *Am J Hum Genet* 44, 388-96 (1989).
 26. Feuk, L., Carson, A. R. & Scherer, S. W. Structural variation in the human genome. *Nat Rev Genet* 7, 85-97 (2006).
 27. Lucito, R. et al. Representational oligonucleotide microarray analysis: a high-resolution method to detect genome copy number variation. *Genome Res* 13, 2291-305 (2003).
 28. Pinkel, D., Straume, T. & Gray, J. W. Cytogenetic analysis using quantitative, high-sensitivity, fluorescence hybridization. *Proc Natl Acad Sci U S A* 83, 2934-8 (1986).
 29. Speicher, M. R., Gwyn Ballard, S. & Ward, D. C. Karyotyping human chromosomes by combinatorial multi-fluor FISH. *Nat Genet* 12, 368-75 (1996).
 30. Schrock, E. et al. Multicolor spectral karyotyping of human chromosomes. *Science* 273, 494-7 (1996).
 31. Manuelidis, L. Individual interphase chromosome domains revealed by in situ hybridization. *Hum Genet* 71, 288-93 (1985).
 32. Kallioniemi, A. et al. Comparative genomic hybridization for molecular cytogenetic analysis of solid tumors. *Science* 258, 818-21 (1992).
 33. Kallioniemi, O. P. et al. Optimizing comparative genomic hybridization for analysis of DNA sequence copy number changes in solid tumors. *Genes Chromosomes Cancer* 10, 231-43 (1994).
 34. Solinas-Toldo, S. et al. Matrix-based comparative genomic hybridization: biochips to screen for genomic imbalances. *Genes Chromosomes Cancer* 20, 399-407 (1997).
 35. Pinkel, D. et al. High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays. *Nat Genet* 20, 207-11 (1998).
 36. Fiegler, H. et al. DNA microarrays for comparative genomic hybridization based on DOP-PCR amplification of BAC and PAC clones. *Genes Chromosomes Cancer* 36, 361-74 (2003).
 37. Snijders, A. M. et al. Assembly of microarrays for genome-wide measurement of DNA copy number. *Nat Genet* 29, 263-4 (2001).
 38. Ishkanian, A. S. et al. A tiling resolution DNA microarray with complete coverage of the human genome. *Nat Genet* 36, 299-303 (2004).
 39. Pollack, J. R. et al. Genome-wide analysis of DNA copy-number changes using cDNA microarrays. *Nat Genet* 23, 41-6 (1999).
 40. Mantripragada, K. K. et al. DNA copy-number analysis of the 22q11 deletion-syndrome region using array-CGH with genomic and PCR-based targets. *Int J Mol Med* 13, 273-9 (2004).
 41. Brennan, C. et al. High-resolution global profiling of genomic alterations with long oligonucleotide microarray. *Cancer Res* 64, 4744-8 (2004).
 42. Carvalho, B., Ouwerkerk, E., Meijer, G. A. & Ylstra, B. High resolution microarray comparative genomic hybridisation analysis using spotted oligonucleotides. *J Clin Pathol* 57, 644-6 (2004).
 43. van Beers, E. H. & Nederlof, P. M. Array-CGH and breast cancer. *Breast Cancer Res* 8, 210 (2006).
 44. Pinkel, D. & Albertson, D. G. Array comparative genomic hybridization and its applications

- in cancer. *Nat Genet* 37 Suppl, S11-7 (2005).
45. Levy-Lahad, E. & Friedman, E. Cancer risks among BRCA1 and BRCA2 mutation carriers. *Br J Cancer* 96, 11-5 (2007).
 46. Wessels, L. F. et al. Molecular classification of breast carcinomas by comparative genomic hybridization: a specific somatic genetic profile for BRCA1 tumors. *Cancer Res* 62, 7110-7 (2002).
 47. Rennstam, K. et al. Patterns of chromosomal imbalances defines subgroups of breast cancer with distinct clinical features and prognosis. A study of 305 tumors by comparative genomic hybridization. *Cancer Res* 63, 8861-8 (2003).
 48. Reis-Filho, J. S. et al. Pleomorphic lobular carcinoma of the breast: role of comprehensive molecular pathology in characterization of an entity. *J Pathol* 207, 1-13 (2005).
 49. Cheng, K. W. et al. The RAB25 small GTPase determines aggressiveness of ovarian and breast cancers. *Nat Med* 10, 1251-6 (2004).
 50. Chin, S. F. et al. Using array-comparative genomic hybridization to define molecular portraits of primary breast cancers. *Oncogene* 26, 1959-70 (2007).
 51. Prentice, L. M. et al. NRG1 gene rearrangements in clinical breast cancer: identification of an adjacent novel amplicon associated with poor prognosis. *Oncogene* 24, 7281-9 (2005).
 52. Ray, M. E. et al. Genomic and expression analysis of the 8p11-12 amplicon in human breast cancer cell lines. *Cancer Res* 64, 40-7 (2004).
 53. Lakhani, S. R. et al. Genetic alterations in 'normal' luminal and myoepithelial cells of the breast. *J Pathol* 189, 496-503 (1999).
 54. Deng, G., Lu, Y., Zlotnikov, G., Thor, A. D. & Smith, H. S. Loss of heterozygosity in normal tissue adjacent to breast carcinomas. *Science* 274, 2057-9 (1996).
 55. Aubele, M. M. et al. Accumulation of chromosomal imbalances from intraductal proliferative lesions to adjacent in situ and invasive ductal breast cancer. *Diagn Mol Pathol* 9, 14-9 (2000).
 56. Amari, M. et al. LOH analyses of premalignant and malignant lesions of human breast: frequent LOH in 8p, 16q, and 17q in atypical ductal hyperplasia. *Oncol Rep* 6, 1277-80 (1999).
 57. Buerger, H. et al. Different genetic pathways in the evolution of invasive breast cancer are associated with distinct morphological subtypes. *J Pathol* 189, 521-6 (1999).
 58. Buerger, H. et al. Ductal invasive G2 and G3 carcinomas of the breast are the end stages of at least two different lines of genetic evolution. *J Pathol* 194, 165-70 (2001).
 59. Cleton-Jansen, A. M. et al. Different mechanisms of chromosome 16 loss of heterozygosity in well- versus poorly differentiated ductal breast cancer. *Genes Chromosomes Cancer* 41, 109-16 (2004).
 60. Reis-Filho, J. S., Simpson, P. T., Gale, T. & Lakhani, S. R. The molecular genetics of breast cancer: the contribution of comparative genomic hybridization. *Pathol Res Pract* 201, 713-25 (2005).
 61. Kuukasjarvi, T. et al. Genetic heterogeneity and clonal evolution underlying development of asynchronous metastasis in human breast cancer. *Cancer Res* 57, 1597-604 (1997).
 62. Schlimok, G. et al. Micrometastatic cancer cells in bone marrow: in vitro detection with anti-cytokeratin and in vivo labeling with anti-17-1A monoclonal antibodies. *Proc Natl Acad Sci U S A* 84, 8672-6 (1987).
 63. Pantel, K., Felber, E. & Schlimok, G. Detection and characterization of residual disease in breast cancer. *J Hematother* 3, 315-22 (1994).
 64. Hermanek, P. What's new in TNM? *Pathol Res Pract* 190, 97-102 (1994).
 65. Braun, S. et al. Cytokeratin-positive cells in the bone marrow and survival of patients with stage I, II, or III breast cancer. *N Engl J Med* 342, 525-33 (2000).
 66. Braun, S. et al. A pooled analysis of bone marrow micrometastasis in breast cancer. *N Engl J*

- Med 353, 793-802 (2005).
67. Pantel, K. et al. Frequency and prognostic significance of isolated tumour cells in bone marrow of patients with non-small-cell lung cancer without overt metastases. *Lancet* 347, 649-53 (1996).
 68. Lindemann, F., Schlimok, G., Dirschedl, P., Witte, J. & Riethmuller, G. Prognostic significance of micrometastatic tumour cells in bone marrow of colorectal cancer patients. *Lancet* 340, 685-9 (1992).
 69. Muller, P., Weckermann, D., Riethmuller, G. & Schlimok, G. Detection of genetic alterations in micrometastatic cells in bone marrow of cancer patients by fluorescence in situ hybridization. *Cancer Genet Cytogenet* 88, 8-16 (1996).
 70. Klein, C. A. et al. Combined transcriptome and genome analysis of single micrometastatic cells. *Nat Biotechnol* 20, 387-92 (2002).
 71. Hartmann, C. H. & Klein, C. A. Gene expression profiling of single cells on large-scale oligonucleotide arrays. *Nucleic Acids Res* 34, e143 (2006).
 72. Klein, C. A. et al. Comparative genomic hybridization, loss of heterozygosity, and DNA sequence analysis of single cells. *Proc Natl Acad Sci U S A* 96, 4494-9 (1999).
 73. Schmidt-Kittler, O. et al. From latent disseminated cells to overt metastasis: genetic analysis of systemic breast cancer progression. *Proc Natl Acad Sci U S A* 100, 7737-42 (2003).
 74. Schmidt-Kittler, O. Von einzelnen disseminierten Tumorzellen zur Metastase: Genomische Analyse der minimalen Resterkrankung des Mammakarzinoms. (2003).
 75. Schardt, J. A. et al. Genomic analysis of single cytokeratin-positive cells from bone marrow reveals early mutational events in breast cancer. *Cancer Cell* 8, 227-39 (2005).
 76. Frengen, E. et al. A modular, positive selection bacterial artificial chromosome vector with multiple cloning sites. *Genomics* 58, 250-3 (1999).
 77. Osoegawa, K. et al. A bacterial artificial chromosome library for sequencing the complete human genome. *Genome Res* 11, 483-96 (2001).
 78. Quackenbush, J. Microarray data normalization and transformation. *Nat Genet* 32 Suppl, 496-501 (2002).
 79. Neuvial, P. et al. Spatial normalization of array-CGH data. *BMC Bioinformatics* 7, 264 (2006).
 80. Hupe, P., Stransky, N., Thiery, J. P., Radvanyi, F. & Barillot, E. Analysis of array CGH data: from signal ratio to gain and loss of DNA regions. *Bioinformatics* 20, 3413-22 (2004).
 81. Press. *Numerical Recipes in C++*. Cambridge University Press (2002).
 82. Sambrook, J. & Russell, D. W. *Molecular Cloning* Cold Spring Harbor Laboratory Press (2001).
 83. Schmidt-Kittler, O. in *Institut für Immunologie (Ludwig-Maximilians-Universität, München, 2003)*.
 84. Foreman, P. K. & Davis, R. W. Real-time PCR-based method for assaying the purity of bacterial artificial chromosome preparations. *Biotechniques* 29, 410-2 (2000).
 85. Carter, N. P., Fiegler, H. & Piper, J. Comparative analysis of comparative genomic hybridization microarray technologies: report of a workshop sponsored by the Wellcome Trust. *Cytometry* 49, 43-8 (2002).
 86. Taylor, S., Smith, S., Windle, B. & Guiseppi-Elie, A. Impact of surface chemistry and blocking strategies on DNA microarrays. *Nucleic Acids Res* 31, e87 (2003).
 87. Shadeo, A. & Lam, W. L. Comprehensive copy number profiles of breast cancer cell model genomes. *Breast Cancer Res* 8, R9 (2006).
 88. Stoecklein, N. H. et al. SCOMP is superior to degenerated oligonucleotide primed-polymerase chain reaction for global amplification of minute amounts of DNA from microdissected archival tissue samples. *Am J Pathol* 161, 43-51 (2002).

89. Steinhaus, S. Eine Methode zur direkten Identifikation chromosomaler Translokationsbruchpunkte. (2007).
90. Barrett, M. T. et al. Comparative genomic hybridization using oligonucleotide microarrays and total genomic DNA. *Proc Natl Acad Sci U S A* 101, 17765-70 (2004).
91. Fiegler, H. et al. High resolution array-CGH analysis of single cells. *Nucleic Acids Res* 35, e15 (2007).
92. Cardoso, J. et al. Genomic profiling by DNA amplification of laser capture microdissected tissues and array CGH. *Nucleic Acids Res* 32, e146 (2004).
93. Eberhardt, A. Auf der Suche nach frühen Gendefekten beim sporadischen Mammakarzinom: Identifizierung zweier 4 und 6 Mb großer Deletionen auf Chromosom 6q24 bei einzelnen disseminierten Tumorzellen mit normalem CGH-Profil. Institut für Immunologie (2006).
94. Weiss, M. M. et al. Comparative genomic hybridisation. *Mol Pathol* 52, 243-51 (1999).
95. Kirchhoff, M. et al. Deletions below 10 megabasepairs are detected in comparative genomic hybridization by standard reference intervals. *Genes Chromosomes Cancer* 25, 410-3 (1999).
96. Wessendorf, S. et al. Hidden gene amplifications in aggressive B-cell non-Hodgkin lymphomas detected by microarray-based comparative genomic hybridization. *Oncogene* 22, 1425-9 (2003).
97. Le Caignec, C. et al. Single-cell chromosomal imbalances detection by array CGH. *Nucleic Acids Res* 34, e68 (2006).
98. Vermeesch, J. R. et al. Molecular karyotyping: array CGH quality criteria for constitutional genetic diagnosis. *J Histochem Cytochem* 53, 413-22 (2005).
99. Fiegler, H. et al. Accurate and reliable high-throughput detection of copy number variation in the human genome. *Genome Res* 16, 1566-74 (2006).
100. Guillaud-Bataille, M. et al. Detecting single DNA copy number variations in complex genomes using one nanogram of starting DNA and BAC-array CGH. *Nucleic Acids Res* 32, e112 (2004).
101. Hu, D. G., Webb, G. & Hussey, N. Aneuploidy detection in single cells using DNA array-based comparative genomic hybridization. *Mol Hum Reprod* 10, 283-9 (2004).
102. van Beers, E. H. et al. A multiplex PCR predictor for aCGH success of FFPE samples. *Br J Cancer* 94, 333-7 (2006).
103. Kim, S., Chin, K., Gray, J. W. & Bishop, J. M. A screen for genes that suppress loss of contact inhibition: identification of ING4 as a candidate tumor suppressor gene in human cancer. *Proc Natl Acad Sci U S A* 101, 16251-6 (2004).
104. Loo, L. W. et al. Array comparative genomic hybridization analysis of genomic alterations in breast cancer subtypes. *Cancer Res* 64, 8541-9 (2004).
105. Garcia, M. J. et al. A 1 Mb minimal amplicon at 8p11-12 in breast cancer identifies new candidate oncogenes. *Oncogene* 24, 5235-45 (2005).
106. Nessling, M. et al. Candidate genes in breast cancer revealed by microarray-based comparative genomic hybridization of archived tissue. *Cancer Res* 65, 439-47 (2005).
107. Devries, S. et al. Array-based comparative genomic hybridization from formalin-fixed, paraffin-embedded breast tumors. *J Mol Diagn* 7, 65-71 (2005).
108. Buckley, P. G., Mantripragada, K. K., Piotrowski, A., Diaz de Stahl, T. & Dumanski, J. P. Copy-number polymorphisms: mining the tip of an iceberg. *Trends Genet* 21, 315-7 (2005).
109. Freeman, J. L. et al. Copy number variation: new insights in genome diversity. *Genome Res* 16, 949-61 (2006).
110. Feuk, L., Marshall, C. R., Wintle, R. F. & Scherer, S. W. Structural variants: changing the landscape of chromosomes and design of disease studies. *Hum Mol Genet* 15 Spec No 1, R57-66 (2006).
111. Galea, M. H., Blamey, R. W., Elston, C. E. & Ellis, I. O. The Nottingham Prognostic Index in

- primary breast cancer. *Breast Cancer Res Treat* 22, 207-19 (1992).
112. Sorlie, T. et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci U S A* 98, 10869-74 (2001).
113. van 't Veer, L. J. et al. Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 415, 530-6 (2002).
114. Hyman, E. et al. Impact of DNA amplification on gene expression patterns in breast cancer. *Cancer Res* 62, 6240-5 (2002).
115. Chin, K. et al. Genomic and transcriptional aberrations linked to breast cancer pathophysiologies. *Cancer Cell* 10, 529-41 (2006).

7 Danksagung

Ganz besonders möchte ich Prof. Dr. Christoph Klein danken, für das ausgesprochen interessante Thema, die Betreuung, die Diskussionen, das Vertrauen und die Rückendeckung, die mir von Anfang an entgegengebracht wurden, sowie für die Möglichkeit, selbstständig und eigenverantwortlich arbeiten zu dürfen. Ich möchte mich auch bedanken für seinen Einsatz, mein Projekt und meine Stelle zu finanzieren, ebenso dafür, dass mir die Teilnahme an zahlreichen Konferenzen und Meetings und natürlich „Tiers“ ermöglicht wurde.

Vielen Dank an Herrn Prof. Dr. Thomas Cremer für sein Interesse an meinem Projekt und für die unkomplizierte Art der Betreuung.

Bei Oleg Schmidt-Kittler bedanke ich mich für die Einführung in die Welt der Blöckchen, bei Thomas Ragg für die bioinformatische Unterstützung und bei Bernhard Radlwimmer für seine Spotting-Bemühungen.

Danke auch an das wunderbare Team aus der Backstube: allen voran Tanja Siart, Karina Petat-Dutter, Christian Vay, Barbara Lindner und Kerstin Bockler. Danke „Schneewittchen“ für die viele Arbeit, die sie mir abgenommen hat, und danke Mani, der mir dabei geholfen hat, ihr Manieren beizubringen.

Ich danke außerdem meinen Kollegen für die schöne Zeit: besonders Claudia, Susanne und Christian für die vielen anregenden Gespräche, die besten Banknachbarn Marco und Claus und all den vielen anderen! Die tolle Arbeitsatmosphäre hat immer dazu beigetragen, dass ich mich auch während der Durststrecken im Verlauf der Arbeit gefreut habe zur Arbeit zu kommen.

Ich bedanke mich bei allen Korrekturlesern: Bernhard Polzer, Martin Fuhrmann, Claus Latrich, Claudia Hartmann und Andreas Fuhrmann, die sehr viel dazu beigetragen haben, die Arbeit lesbar zu machen, und die viele gute Anregungen hatten.

Meinen Eltern möchte ich besonders danken, dass sie mich immer unterstützt haben, für die vielen „Lektüren“, ihre Bereitschaft, sich auch in die technischen Details des Laboralltags hineinzusetzen, und vor allem für ihre moralische Rückendeckung.

Danke Martin, dafür, dass du dir immer die Zeit genommen hast, meine Arbeit aus einem anderen Blickwinkel kritisch zu hinterfragen und alles verstehen zu wollen. Vielen Dank auch für die vielen fruchtbaren Diskussionen und natürlich für deine Liebe.

8 Lebenslauf

Christine Fuhrmann geb. Nachtsheim
 geboren am 8. Juli 1979
 in Oldenburg, Oldenburg
 seit 10. Juli 2004, verheiratet mit Martin Fuhrmann

- 1985-1989 Grundschulen St. Cloud (Frankreich), Sigmaringen und Rheinbach
- 1989-1991 St. Joseph Gymnasium, Rheinbach
- 1991-1994 Gymnasium Schwarzenbek
- 1994-1996 Gymnasium der Deutschen Schule, SHAPE (Belgien)
- 1995-1997 American High School, Department of Defence Schools, SHAPE (Belgien)
- 1997 High School Diploma und Anerkennung der „Allgemeinen Hochschulreife“
 „gut“ (1,7)
- 1997-2002 Studium der „Technischen Biologie“ an der Universität Stuttgart
 1999 Diplom-Vorprüfung „sehr gut“ (1,3)
 2002 Diplom-Hauptprüfung „mit Auszeichnung“ (1,0)
- 2000 Anfertigung der Studienarbeit und Industriepraktikum bei der Novartis
 Pharma AG, Basel, in der Abteilung „Molekulare Onkologie“ mit dem
 Thema: „Expression und Aktivität der Cathepsine B, L und K in Tumoren
 verschiedener Tiermodelle und Etablierung eines *in vitro*
 Metastasierungsassays sowie dessen Validierung mit NCO700 “
- 2002 Diplomarbeit an der Universität Stuttgart unter der Leitung von Prof. Dr.
 Wolfgang Seufert mit dem Thema: „Characterization of Cdc123 and its
 interaction with Rbx8 and Rbx9 in *Saccharomyces cerevisiae*“
- 2003 Beginn der vorliegenden Dissertation an der Fakultät für Biologie der
 Ludwig-Maximilians-Universität München betreut von Prof. Dr. Thomas
 Cremer. Die praktischen Arbeiten wurden durchgeführt am Institut für
 Immunologie der Ludwig-Maximilians-Universität München, sowie am
 Institut für Pathologie, Abteilung Onkogenomik der Universität Regensburg
 (seit Oktober 2006), im Labor von Prof. Dr. Christoph Klein.

9 Veröffentlichungen

Eigene Veröffentlichungen

Julian A. Schardt, Manfred Meyer, Claudia H. Hartmann, Falk Schubert, Oleg Schmidt-Kittler, Christine Fuhrmann, Bernhard Polzer, Marco Petronio, Roland Eils, and Christoph A. Klein. (2005). Genomic analysis of single cytoke-
ratin-positive cells from bone marrow reveals early mutational events in breast cancer. *Cancer Cell* **8** (3) 227-39

Christine Fuhrmann¹, Oleg Schmidt-Kittler¹, Nikolas H. Stoecklein, Karina Petat-Dutter, Christian Vay, Kerstin Bockler, Thomas Ragg, Richard Reinhardt, Christoph A. Klein. (2008). High-resolution array comparative genomic hybridization of single micrometastatic tumor cells. *Nucleic Acids Research* **36** (7) e39¹: equal contributors

Posterpräsentation

Christine Fuhrmann, Oleg Schmidt-Kittler, Christian Vay, Karina Petat-Dutter, Nikolas H. Stoecklein, Kerstin Bockler, Christoph A. Klein. (2006). BAC array analysis of single disseminated tumor cells from bone marrow of breast cancer patients. *AACR 97th Annual Meeting 2006, Washington, DC*

Vorträge

Christine Fuhrmann. Array-based comparative genomic hybridization of single cells. (2005). *Array-CGH and Molecular Cytogenetics Meeting 2005, Bari, Italien*

Christine Fuhrmann. Identifizierung und funktionelle Charakterisierung von Tumor- und Metastasen-initiiierenden Mutationen beim Mammakarzinom als Ausgangspunkt für adjuvante Therapien. (2005). *SFB456-Retreat 2005, Schloss Hohenkammner*

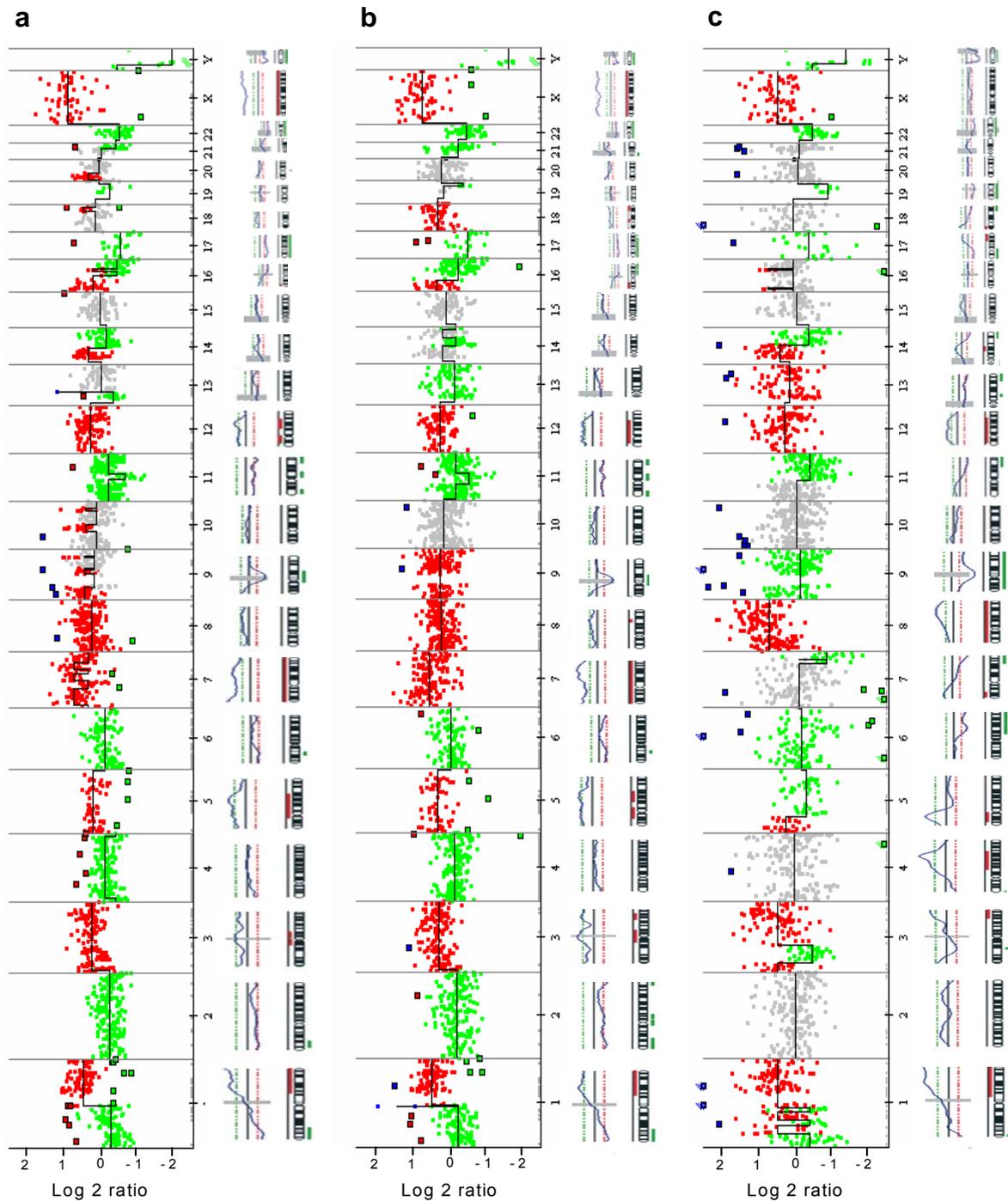
Christine Fuhrmann. Array-based comparative genomic hybridization of single cells. (2006). 3. Mitarbeitercolloquium des Bayerischen Genomforschungsnetzwerkes, **München**

Patent

Patent No.: 06020842.8-2403

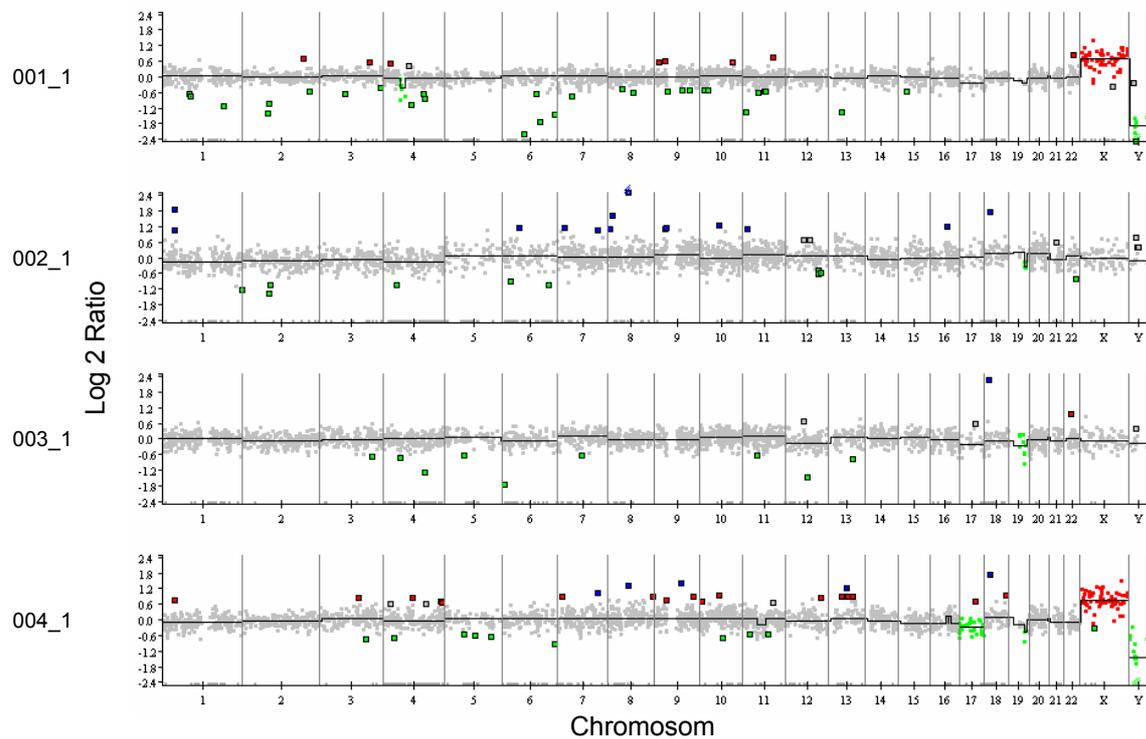
Titel: Method for preparing artificial chromosomes which are free of host DNA

10 Anhang



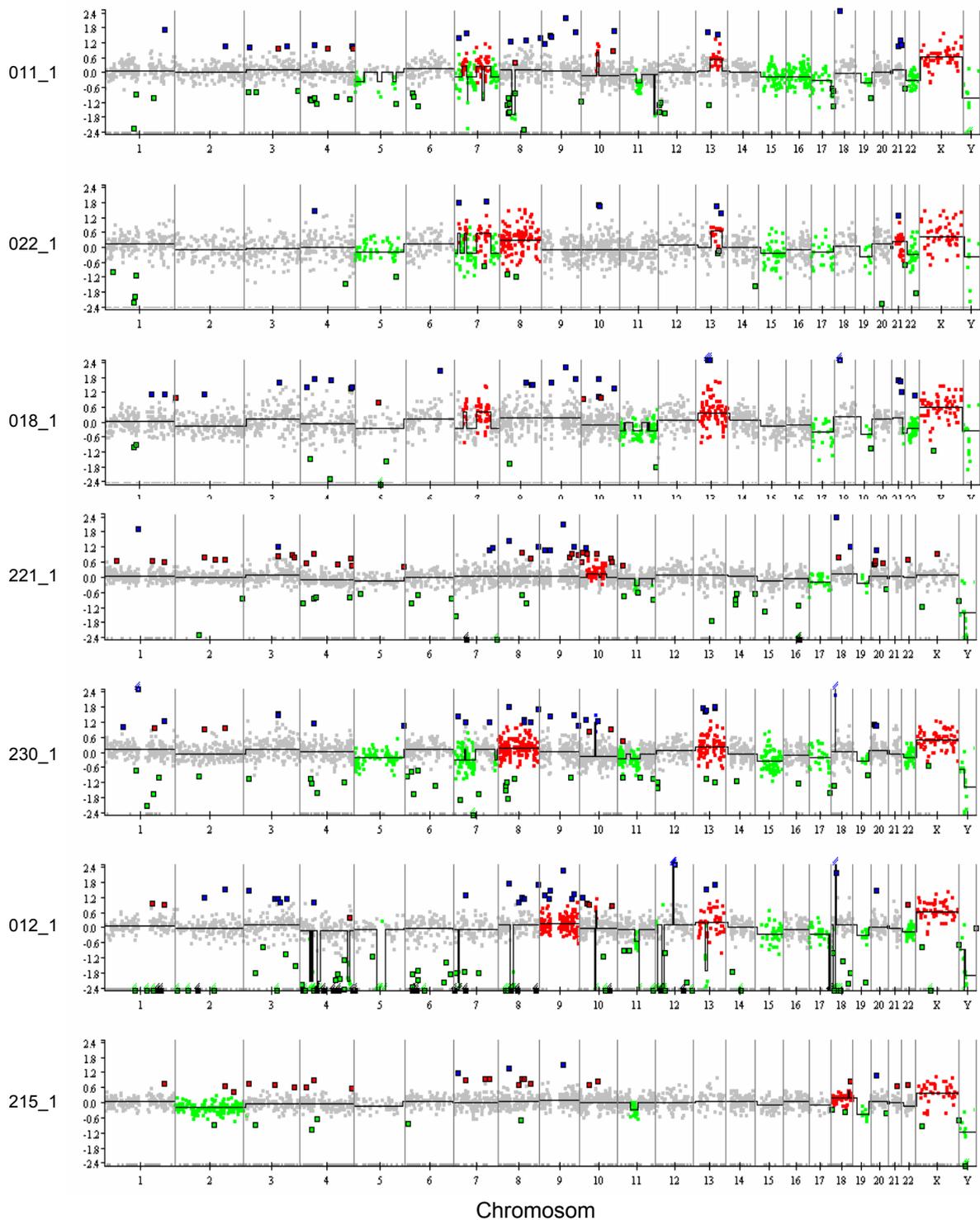
Anhang 1: Array- und Metaphasen-CGH-Profile von drei disseminierten Tumorzellen aus dem Knochenmark zweier Mammakarzinompatientinnen

Dargestellt sind jeweils nebeneinander die Metaphasen-CGH-Profile (rechts) und die Array-CGH-Profile (links) der selben Proben. **(a)** und **(b)** zeigen die Profile von zwei einzelnen disseminierten Tumorzellen aus dem Knochenmark einer Patientin mit Mammakarzinom (008_1 und 008_2). In **(c)** sind die Profile einer Tumorzelle aus dem Knochenmark einer weiteren Patientin dargestellt (118_1). (Darstellung der Array-CGH wie in Abbildung 14; Chromosom 19 wird nicht berücksichtigt, siehe 3.1.6)



Anhang 2: Array-CGH-Profil von vier normalen Einzelzellen

Dargestellt sind Array-CGH-Profile von Amplifikaten aus je zwei normalen Einzelzellen männlicher Probanden (002_1 und 003_1) und weiblicher Probanden (001_1 und 004_1). (Darstellung der Array-CGH wie in Abbildung 14; Chromosom 19 wird nicht berücksichtigt, siehe 3.1.6)



Anhang 3: Array-CGH-Profil von sieben disseminierten Tumorzellen von M0-Patientinnen

Dargestellt sind die Array-CGH-Profile von sieben disseminierten Tumorzellen von M0-Patientinnen. (Darstellung der Array-CGH wie in Abbildung 14; Chromosom 19 wird nicht berücksichtigt, siehe 3.1.6)