

**Untersuchungen zur Struktur eines Amplikons im Intron 1  
des humanen epidermalen Wachstumsfaktor Rezeptors:  
Erstmutation in der Mammakarzinom-Entstehung**

**Dissertation**

zur Erlangung des Grades  
Doktor der Naturwissenschaften  
im Department Biologie,  
Fakultät für Mathematik, Informatik und Naturwissenschaften,  
Universität Hamburg

vorgelegt von  
**Sönke Meyer-Staeckling**  
aus Hamburg

Hamburg 2009

Genehmigt vom Department Biologie  
der Fakultät für Mathematik, Informatik und Naturwissenschaften  
an der Universität Hamburg  
auf Antrag von Herrn Professor Dr. B. BRANDT  
Weiterer Gutachter der Dissertation:  
Herr Professor Dr. U. WIENAND  
Tag der Disputation: 10. Juli 2009

Hamburg, den 25. Juni 2009



*Jörg Ganzhorn*

Professor Dr. Jörg Ganzhorn  
Leiter des Departments Biologie

Für meine geliebte Frau Verena,  
meine Familie,  
meine Freunde  
und alle, die mich auf meinem Weg begleiten.

Inhaltsverzeichnis.....	I
Abkürzungsverzeichnis .....	V
Zusammenfassung .....	VII
<b>1. Einleitung .....</b>	<b>1</b>
<b>1.1 Allgemeines Verständnis von Krebsentwicklung</b>	
nach der Entschlüsselung des menschlichen Genoms.....	1
<b>1.2 Krebsepidemiologie und Definition maligner Tumore.....</b>	<b>2</b>
<b>1.3 Tumorgenese und onkogene Mutationen in Rezeptortyrosinkinasen.....</b>	<b>2</b>
<b>1.4 Rezeptortyrosinkinasen und die ErbB-Familie.....</b>	<b>5</b>
<b>1.5 Der Epidermale Wachstumsfaktor-Rezeptor (EGFR) .....</b>	<b>9</b>
<b>1.6 Die Genstruktur und Transkriptionsregulation von <i>EGFR</i>.....</b>	<b>10</b>
<b>1.7 Modelle zur Entstehung von <i>EGFR</i>-Amplifikationen .....</b>	<b>13</b>
<b>1.8 Metastasierung und minimal residuale Erkrankungen.....</b>	<b>15</b>
<b>1.9 DNA-Quantifizierungsmethoden zur Untersuchung des <i>EGFR</i>-Amplikon .....</b>	<b>17</b>
1.9.1 Fine Tiling Array-CGH.....	17
1.9.2 Quantitative Real-Time PCR.....	19
<b>1.10 Fragestellung .....</b>	<b>20</b>
<b>2. Material und Methoden .....</b>	<b>22</b>
<b>2.1 Material .....</b>	<b>22</b>
2.1.1 Zelllinien.....	22
2.1.2 Verbrauchsmaterialien.....	22
2.1.2.1 Chemikalien .....	22
2.1.2.2 Antikörper.....	23
2.1.2.3 Oligonukleotide.....	23
2.1.2.4 Puffer und Lösungen.....	29
2.1.2.5 Zellkultur- und Nährmedien .....	29
2.1.2.6 Verwendete käufliche Systeme (Kits) .....	30
2.1.3 Geräte .....	30
<b>2.2 Methoden .....</b>	<b>31</b>
2.2.1 Aufarbeitung und Nachweis von disseminierten Tumorzellen in	
Knochenmarkaspiraten / Blut.....	31
2.2.1.1 Isolierung disseminierter Tumorzellen aus dem Knochenmark.....	31

2.2.1.2 Immunzytochemische Färbung zum Nachweis von Zytokeratin-positiven Zellen im Knochenmark / Blut.....	32
2.2.1.3 Immunomagnetische Anreicherung von DTCs aus Patientenblut.....	33
2.2.1.4 Fluoreszenzfärbung von Cytospins.....	33
2.2.2 Mikromanipulation einzelner DTZs / Zelllinienzellen.....	34
2.2.2.1 Transfer der Zellen in PCR-Reaktionsgefäße.....	35
2.2.2.2 Silanisierung der Glasstäbchen.....	35
2.2.2.3 Zellyse von transferierten Zellen.....	35
2.2.2.4 Ganz-genomische DNA Amplifikation (Whole Genome Amplification – WGA) durch Multiple Displacement Amplification (MDA).....	36
2.2.2.5 Aufreinigung des WGA-Produktes.....	37
2.2.3 Isolierung von Nukleinsäuren.....	38
2.2.3.1 Isolierung von genomischer DNA.....	38
2.2.3.2 Isolierung von Gesamt-RNA.....	38
2.2.3.3 Quantitäts- und Qualitätsbestimmung von Nukleinsäuren.....	38
2.2.3.4 Nanodrop Spektralphotometer.....	38
2.2.3.5 Agarose-Gelelektrophorese.....	39
2.2.3.6 Aufreinigung von Nukleinsäuren.....	39
2.2.3.7 Ethanolfällung.....	39
2.2.3.8 Bandenextraktion aus Agarosegelen.....	39
2.2.3.9 Erststrang cDNA-Synthese.....	39
2.2.3.10 Polymerase-Kettenreaktion.....	40
2.2.3.11 Aufarbeitung der Mikrosatelliten PCR-Produkte zur Fragmentanalyse.....	41
2.2.3.12 Auswertung der Fragmentanalysen der Mikrosatelliten-PCR.....	41
2.2.3.13 Quantitative Real Time PCR.....	42
2.2.3.14 Auswertungsmethoden der quantitativen Realtime PCR.....	44
2.2.3.15 Analyse der differentiellen PCR5 Fragmente mittels Restriktionsanalysen.....	45
2.2.3.16 DNA-Sequenzierung.....	45
2.2.3.17 Analysen zu Veränderungen der DNA-Kopienanzahl mittels Fine Tiling Array-CGH.....	46
2.2.3.18 Hybridisierung der Fine Tiling Array-CGH.....	47
2.2.3.19 Auswertung der Fine Tiling Array-CGH.....	48
<b>2.3 Zellkulturbedingungen.....</b>	<b>48</b>
<b>2.4 Verwendete Software, Onlinetools, Datenbanken.....</b>	<b>49</b>

3. Ergebnisse .....	50
3.1 Veränderung der <i>EGFR</i> -Genkopienzahl von MDA-MB-468 unter EGF-Druck .....	50
3.2 Quantitative Real-Time PCR: Etablierung eines TaqMan Assays zur Quantifizierung <i>EGFR</i> -Genkopienzahl .....	53
3.3 Fine Tiling Array-CGH: Untersuchung der <i>EGFR</i> -Amplikone von Tumorzelllinien.....	56
3.4 Validierung der errechneten Bruchpunkte über quantitative Real-Time PCR .....	65
3.5 PCR-basierte Detektion von Umlagerungssequenzen an der telomerwärtigen Startregion von MDA-MB-468.....	76
3.6 Analyse der differentiellen PCR-Fragmente.....	77
3.7 Expressionsanalyse von <i>EGFR</i> und <i>CHCHD2</i> in MDA-MB-468 Subklonen .....	80
3.8 <u>Whole Genome Amplification (WGA)</u> zur Bestimmung der <i>EGFR</i> -Gendosis in Einzelzelle .....	82
3.9 Validierung der WGA mittels eines Mikrosatelliten-PCR Panels .....	86
3.10 Multiplex PCR-Protokoll Optimierung für die Mikrosatelliten Primer .....	87
3.11 Fragmentlängenbestimmung in der DNA von Tumorzelllinien.....	88
3.12 Validierung der WGA mit 30pg Ausgangs-DNA durch anschließende Mikrosatelliten-PCR.....	89
3.13 Etablierung einer Methode zur genomischen Charakterisierung von Einzelzellen.....	92
3.14 Validierung der WGA amplifizierten DNA aus Einzelzellen durch Mikrosatelliten-PCR.....	98
3.15 Validierung der Replikationsgenauigkeit der WGA amplifizierten DNA aus Einzelzellen durch Sequenzierung.....	101
3.16 Validierung der WGA amplifizierten DNA aus Einzelzellen durch quantitative Real-Time PCR.....	103
3.17 Validierung des alternativen LINE-1 Referenz Assays für die qPCR .....	105
3.18 Etablierung weiterer <i>EGFR SYBR green Assays</i> für die qPCR .....	108
3.19 Quantifizierung von <i>EGFR</i> Amplifikationen in zirkulierenden Tumorzellen aus dem peripheren Blut und Knochenmark .....	109
3.20 Validierung der WGA amplifizierten DNA aus Einzelzellen durch <i>Fine Tiling Array CGH</i> .....	112

4. Diskussion .....	118
4.1 Methodische Aspekte.....	118
4.2 Analyse der Struktur des EGFR-Amplikon .....	123
4.3 Ausblick .....	132
5. Literaturverzeichnis.....	133
6. Appendix.....	144
Danksagung .....	153
Lebenslauf .....	154
Eidesstattliche Erklärung.....	155

## Abkürzungsverzeichnis

ACIS	<i>Automated Cellular Imaging System</i>
AGCN	durchschnittliche Genkopien Anzahl ( <i>Average Gene Copy Number</i> )
AI	<i>Allelic Imbalance</i>
APAAP	Alkalische Phosphatase Anti-Alkalische Phosphatase
Aqua dest.	Aqua destillata
Arg	Arginin
BAC	<i>Bacterial Artificial Chromosome</i>
BCIP/NBT	<i>5-Brom-4-Chlor-3-Indolyl-Phosphat / Nitro Tetrazolium Blue Chloride</i>
BFB	<i>Breakage Fusion Bridge</i>
Bps	Basenpaar ( <i>Basepairs</i> )
BSA	Rinderserumalbumin
CA	Carzinoma
cDNA	komplementäre DNA
CFS	<i>Common Fragile Site</i>
CGH	komparative genomische Hybridisierung ( <i>Comparative Genomic Hybridization</i> )
Chr	Chromosom
CK	Zytokeratin
CT	<i>Cycle Threshold</i> (Schwellenwert-Zyklus)
CV	Varianzkoeffizient ( <i>Coefficient of Variation</i> )
Cy3	<i>Cyanine Dye 3</i>
dATP	Desoxyadenosintriphosphat
DC	<i>Critical Difference</i>
dCTP	Desoxycytosintriphosphat
DEPC	Diethylenpyrocarbonat
dGTP	Desoxyguanosintriphosphat
DAPI	4,6-Diamidino-2-phenylindol Dihydrochlorid
DMSO	Dimethylsulfoxid
DNA	Desoxyribonukleinsäure
dNTP	Desoxynukleotid
DOP	<i>Degenerate Oligonucleotide Primed</i>
dTTP	Desoxythymidintriphosphat
DTC	disseminierte Tumorzelle ( <i>Disseminated Tumour Cell</i> )
dUTP	Desoxyuridintriphosphat
EDTA	Ethylendiamintetraessigsäure
EGFR	Epidermaler Wachstumsfaktor Rezeptor ( <i>Epidermal growth factor receptor</i> )
FCS	fötales Rinderserum
FISH	Fluoreszenz <i>in-situ</i> Hybridisierung
FRET	<i>Fluorescence Resonance Energy Transfer</i>
Gly	Glycin
h	Stunde
HE	Hämaltoxylin-Eosin
His	Histidin
HPLC	<i>High Performance Liquid Chromatography</i>
HPSF	<i>High Purity Salt Free</i>
kb	Kilobasenpaar
kDa	Kilodalton
KM	Knochenmark
LCR	<i>Low Copy Repeat</i>
LK	Lymphknoten
Log <sub>2</sub>	Logarithmus zur Basis 2
LOH	<i>Loss of Heterozygosity</i>
LTR	<i>Long Terminal Repeat</i>
mb	Megabasen

---

min	Minute
MRD	<i>Minimal Residual Disease</i>
mRNA	„messenger“ Ribonukleinsäure
NAHR	<i>Non Allelic Homologous Recombination</i>
NHEJ	<i>Non Homologous End Joining</i>
Nm	Nanometer
NTC	<i>Non Template Control</i>
OD	optische Dichte
PBS	Phosphat gepufferte Salzlösung
PCR	Polymerase-Kettenreaktion
Pe	Phytoerythrin
PEP	<i>Primer Extension Preamplification</i>
Pfa	Paraformaldehyd
Pg	Pikogramm
pH	negativer dekadischer Logarithmus der Wasserstoffionenkonzentration
qPCR	<i>quantitative Real-Time-PCR</i>
RFS	<i>Rare Fragile Site</i>
RNA	Ribonukleinsäure
RT	Raumtemperatur
RTK	Rezeptortyrosinkinase
RT-PCR	Reverse Transkriptase-PCR
rpm	Umdrehungen pro Minute
SD	<i>Standard Deviation</i>
sec	Sekunde
Ser	Serin
SNP	Einzelnukleotid-Polymorphismen ( <i>Single Nucleotid Polymorphism</i> )
SSC	Natriumchlorid und Natriumcitrat
TBE	Tris-Borat-EDTA
TE	<i>Tris-EDTA</i>
TMAC	Tetramethylammoniumchlorid
U	<i>Units</i>
UKE	Universitätsklinikum Hamburg-Eppendorf
UV	Ultraviolett
VK	Varianzkoeffizient
Vol.	Volumenanteil
WGA	<i>Whole Genome Amplifikation</i>
w/v	Masse pro Volumen
WT	Wildtyp

## Zusammenfassung

Der Epidermale Wachstumsfaktor Rezeptor (EGFR) gehört der Proteinklasse der transmembranen Rezeptortyrosinkinase (RTK) an (22). Die physiologische Aufgabe der RTKs besteht in der Regulation von Zellwachstum, Angiogenese, Adhäsion und Zellmotilität, sowie der Apoptose (24-26). Der Verlust bzw. die Störung von einem oder mehreren Autoregulations-Mechanismen löst seine Deregulation aus, die zur onkogenen Transformation der Zelle führen kann (27,28). Für EGFR wurden in zahlreichen Studien an unterschiedlichen Tumorentitäten hohe Überexpressionsraten ermittelt (29). Für die Karzinogenese der Brustdrüse ist ein frühes Auftreten der Überexpression von EGFR im Tumor und dem umgebenden Normalgewebe gezeigt worden (69), die mit einer schlechten Prognose sowie einer erhöhten Rate an Tumorrezidiven assoziiert war.

Der Fokus dieser Arbeit war die Untersuchung der Genamplifikation von *EGFR* im Bereich von Chromosom 7p11-14. Anhand des Modells der Mammakarzinom Zelllinie MDA-MB-468 sowie der Linien A431 und BT-20, die eine erhöhte durchschnittliche Genkopienzahl („Average Gene Copy Number“ – AGCN) für *EGFR* besitzen, sollten durch molekulargenetische Techniken Rückschlüsse auf mögliche Mechanismen der Entstehung der Genamplifikation gezogen werden. Die Analyse von Veränderungen der AGCN wurde basierend auf Fine Tiling Array-CGH („Competitive Genomic Hybridisation“) und quantitativer Real-Time PCR (qPCR) durchgeführt. Durch permanente EGF-Stimulation wurden MDA-MB-468 Sublinien mit absteigender AGCN für das *EGFR*-Gen generiert (MDA(+) 2-3; MDA0107 3-7; MDA0106 14-17). An ihnen konnte der Zusammenhang zwischen Genkopienzunahme und Zunahme der mRNA Expressionshöhe bestätigt werden (69).

Aus diesen Untersuchungen der *EGFR*-Amplikons von MDA-MB-468 (MDA-WT), drei seiner Sublinien mit unterschiedlicher AGCN für *EGFR* sowie A431 und BT-20 ergaben sich distinkte amplifizierte Bereiche der chromosomalen DNA. Diese schlossen, mit Ausnahme der MDA-MB-468 Sublinie mit nahezu diploider *EGFR*-Kopienzahl (MDA(+)), jeweils das *EGFR*-Gen in voller Länge ein. Das Amplikon der Sublinie MDA(+) zeigte nach den Exonen 4 bis 7 keine erhöhte AGCN für *EGFR*. Es wurden Regionen mit evolutionär-bedingter, erhöhter chromosomaler Instabilität an den Start- und Endregionen der Amplikons lokalisiert (177). Für die MDA-MB-468 Sublinie konnte ein mehrstufiges Amplifikationsmuster beschrieben werden, wobei die Länge des Amplikons mit steigender AGCN für *EGFR* zunahm. Dabei bildete das Amplikon der Sublinie MDA(+) das zentrale Element des Hauptamplikons aus. Dem Hauptamplikon wurden zwei weitere Regionen telomer- und

centromerwärts des zentralen Elementes zugeordnet. Diese Regionen zeigten ab dem Amplikon der Sublinie MDA0107 DNA-Zugewinne. Eine gering amplifizierte Region schloss sich centromerwärts dem Hauptamplikon von dem Subklon MDA0106 und MDA-WT an. In der telomerwärtigen Startregion des Amplikons wurden im Bereich eines retroviralen HERVK-Elements (179,180) durch qPCR Kopienzahlen bestimmt, die – bezogen auf die AGCN von *EGFR* – um Faktor 2-16 erhöht waren. In einem 4,8kb telomerwärts der HERVK-Sequenz lokalisierten retroviralen SVA\_A-Element konnten durch *Long-Range* PCR differentielle PCR-Fragmente amplifiziert werden, die auf chromosomale Umlagerungen in einem Sequenzbereich hindeuten, der eine Akkumulation repetitiver DNA Motive aufweist. Die Analyse der telomerwärtigen Startregion des Amplikons durch qPCR zeigte in der SVA\_A-Sequenz die ersten chromosomalen Bereiche mit erhöhter DNA-Kopienzahl. Ferner wurden durch vergleichende Softwareanalysen Homologien zwischen der repetitiven SVA\_A-Sequenz und dem HERVK-Element von >500 Basen nachgewiesen.

Zusammenfassend zeigen die Ergebnisse deutlich, dass alle analysierten *EGFR*-Amplikons innerhalb der rare fragile site (RFS) FRA7A lokalisiert sind. Es wäre demnach möglich, dass die Aktivierung der RFS durch Einzel- oder Doppelstrangbrüche erfolgen könnte (186, 187), die wiederum durch die Anwesenheit stark repetitiver Sequenzelemente begünstigt werden. Solche repetitiven Sequenzen führen häufig durch Leseungenauigkeiten zur Abnahme der Replikationsgeschwindigkeit an der Replikationsgabel. Aus einer möglichen Aktivierung der RFS FRA7A entstehende Fehler bei der NAHR (Nicht allelische homologe Rekombination) oder NHEJ („*Non homologous end joining*“) könnten als Auslöser des BFB-Mechanismus angesehen werden (89). Dieser wiederum könnte Ausgangspunkt für die beobachteten Amplifikationen mit *EGFR* im Zentrum sein.

Um die genomische Charakterisierung des *EGFR*-Status auf Einzelzell-Ebene zu ermöglichen, wurde eine Methode etabliert, die basierend auf Mikromanipulation mit anschließender „*Whole Genome Amplification*“ (WGA) das gesamte Genom der Zelle amplifiziert (155,156). An der durch WGA amplifizierten DNA wurden anschließend die Folgeapplikationen der Mikrosatelliten-PCR zur Detektion von allelischen Imbalancen (AIs), der Sequenzierung zur Überprüfung der Basensequenz, sowie Verfahren zur Analyse von Änderungen in der DNA-Kopienzahl, die qPCR und „*Fine Tiling Array CGH*“ validiert. Auf der Basis dieser Methode wurde die Heterogenität der MDA-MB-468 Zellen für den *EGFR*-Lokus in Einzelzellen quantitativ nachgewiesen.

Bei der Analyse von „*Disseminated Tumor Cells*“ (DTCs), die als Zytokeratin-positive Zellen im Knochenmark und Blut verschiedener metastasierender Mammakarzinom-Patientinnen identifiziert wurden, wurden durch qPCR in 58% der Zellen *EGFR*-Amplifikationen nachgewiesen, davon in 33% der Zellen mit einer AGCN für *EGFR*  $\geq 10$ . Im Vergleich dazu konnten in <5% der Primärtumore durch CGH-Analysen *EGFR*-Amplifikationen nachgewiesen werden (175). Ein biologischer Zusammenhang zwischen dem Metastasierungsprozess und der Genamplifikation des *EGFR* ist deshalb naheliegend.

## 1. Einleitung

### 1.1 Allgemeines Verständnis der Krebsentwicklung nach der Entschlüsselung des menschlichen Genoms

Noch bis vor wenigen Jahren beschrieb man in der Biologie, insbesondere der Zellbiologie, die Zellen in einem Tumor mit Attributen wie „entdifferenziert“ oder „retrograd differenziert“. Der Fokus lag dabei auf dem Verlust von organspezifischen Funktionen der Tumorzelle, wie z.B. dem Verlust der Polarität von Brustdrüsenzellen oder die geringere Adhärenz zu benachbarten Zellen und Basalmembranen. Diese Sichtweise basierte auf der Annahme, dass in den Tumorzellen eine „metabolisches Unordnung“ bestand. Die Pleiomorphie der Zellkerne von Tumorzellen mit ihrem deutlich höheren DNA-Gehalt oder die immunhistochemische Darstellung von Strukturproteinen, wie z.B. den Zytokeratinen des Zytoskeletts, die sich von außen betrachtet als unstrukturiert darstellen, verstärkten diesen Eindruck.

Diese Ansicht muss, basierend auf dem aktuellen Wissensstand der Molekularbiologie, in so weit revidiert werden, als dass Tumorzellen ganz im Gegenteil hochspezialisiert und strukturiert sind. Dadurch werden koordinierte Funktion wie z.B. gerichtete Zellbewegung bei der Invasion der Basalmembran, die Persistenz im Blutkreislauf und die Extravasation aus den Gefäßen in andere Organe als dem Ursprungsorgan, z.B. den Knochen, dem Gehirn oder der Leber überhaupt ermöglicht. Auf dieser Basis sind Tumorzellen, die die Fähigkeiten besitzen unkontrolliert zu proliferieren, invasiv über Basalmembranen zu wachsen und Metastasen auszubilden, als präzise koordinierte Systeme zu betrachten, die aufgrund dieser Fähigkeiten im menschlichen Organismus überleben und sich systemisch ausbreiten können.

Bereits vor dem Abschluss des Humanen Genomprojekts (HUGO) identifizierte Schlüsselmechanismen und –moleküle, die diese speziellen Funktionen der Tumorzellen vermitteln, könnten unter Einsatz neuer Technologien für die Analyse von DNA, RNA und Proteinen, als molekulare Grundlage dieser speziellen und komplexen Funktionen bestätigt werden und es gelang die Feinstruktur und Netzwerkstruktur weiter auszuleuchten. So wurden genomische Schlüssel-Aberrationen, wie z.B. chromosomale Deletionen oder Amplifikationen für einzelne Tumortypen identifiziert.

Es gelang mRNA Expressionsmuster für einzelne Tumortypen mit ihrer Progressionsgeschwindigkeit und dem Therapieansprechen zu verbinden. Auch spezifische Proteinmodifikationen z.B. für die Dominanz von Signalwegen der Zellmigration wurden gefunden. Dominante molekulare Signalwege der Zellen sind für die Diagnostik, sowie individualisierte Therapie als Indikatoren für eine schlechtere Prognose oder Therapieresistenzen von Bedeutung. Diese Molekularen Prognose- und Prädiktionsparameter werden derzeit zunehmend in die Klinik eingeführt.

## **1.2 Krebsepidemiologie und die Definition maligner Tumore**

Im Jahr 2004 starben in Deutschland ~208.000 Menschen an den Folgen von Krebs. Das macht Krebs zur zweithäufigsten Todesursache nach Herz-Kreislauf-Erkrankungen in unserem Land. Die meisten Todesfälle sind dabei auf bösartige Tumoren der Lunge, der weiblichen Brustdrüse, des Darms, der männlichen Prostata-drüse, des Magen und der Ovarien zurückzuführen (1). Karzinom, auch als „solide“ Tumore bezeichnet, entstehen aus den Epithelzellen der genannten Organe. Sie bleiben lange ohne Symptome und werden spät bzw. in einem fortgeschrittenen, kaum noch therapierbaren Stadium erkannt. Morphologisch wird ein maligner Tumor durch die Fähigkeit seiner Zellen definiert ohne Kontaktinhibition zu wachsen, invasiv über Basalmembranen zu wachsen und in anderen Organen Metastasen ausbilden zu können. Genetische Veränderungen innerhalb der Zellen, auf denen die Beobachtungen basieren, lassen sich in vielen Fällen als Chromosomenaberrationen bereits mikroskopisch feststellen. Die Entschlüsselung der DNA-Sequenz führte zur Identifikation von Genen, die unmittelbar mit der Transformation einer Zelle zu einer malignen Tumorzelle in Zusammenhang stehen.

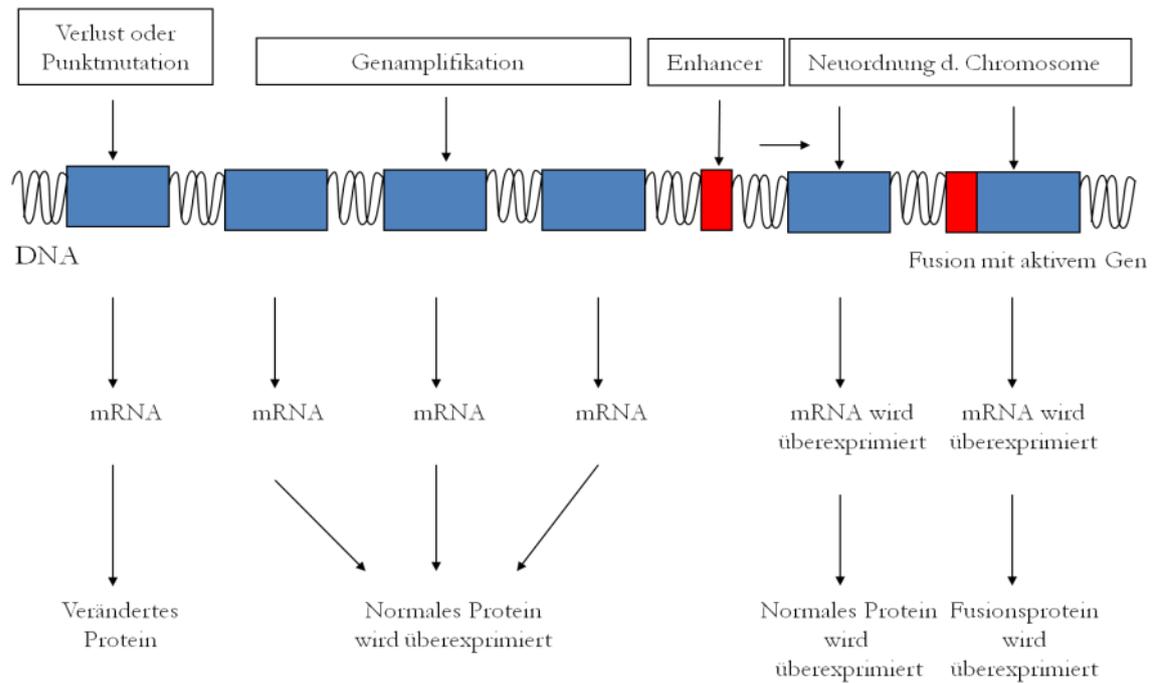
## **1.3 Tumorgenese und onkogene Mutationen in Rezeptortyrosinkinasen**

Der Tumorgenese liegt eine Störung bzw. Veränderung in der Signaltransduktion zu Grunde, die zu unreguliertem Wachstum führt. Diese kann in allen Signalwegen vorliegen, die Zellen modellieren oder verändern (2,3), sowie das empfindliche Gleichgewicht zwischen Zellteilung, und Wachstum gegenüber der Apoptose stören (4). Es wurden unterschiedliche Anfälligkeiten für onkogene Mutationen in den verschiedenen Klassen von Signal-Proteinen und ihren Signaltransduktionswegen beschrieben (5). Deren kodierende Gene werden als Protoonkogene bezeichnet.

Zu den wichtigsten Molekülen zählen hierbei Rezeptortyrosinkinasen (RTKs) und deren Liganden, sowie Transmittermoleküle, Transkriptionsfaktoren und proliferationshemmende Proteine (6-10).

Die Transformation einer benignen Zelle in eine Tumorzelle kann einerseits durch Genprodukte von Onkogenen, die aufgrund von Mutationen aus Protoonkogenen hervorgingen und andererseits durch Inaktivierung von Tumorsuppressorgenen bedingt sein. Ausgelöst werden können Mutationen durch Fehler in DNA-Reparaturmechanismen, der DNA-Rekombination (11), sowie der Chromosomenverteilung während des Zellzyklus (12).

Historisch bedingt unterscheidet man bei der Aktivierung von Onkogenen initiale Ereignisse, die während der frühen Phasen der neoplastischen Transformation stattfinden und Progression-assoziierten Ereignissen, die spätere transformierende Prozesse auslösen (13). Zu den initialen Ereignissen zählen die Mutationen, die eine konstitutive Aktivierung des veränderten Proteins bewirken. Als ein wichtiges Beispiel sind Deletionen im *EGFR*-Gen zu nennen, aus denen ein Verlust der Liganden-Bindedomäne resultiert, und die zu einer konstitutiven, Liganden-unabhängigen Aktivierung des Rezeptors führen (14). Ebenso kann ein Aminosäureaustausch durch eine Punktmutation zu einer konstitutiven Aktivierung der Kinase des Rezeptors führen (15,16). Die Bildung von Fusionsproteinen durch Translokationen kann ebenfalls zu einer Steigerung der Kinaseaktivität führen (17). Diese können durch die Fusion mit sehr aktiven Promotoren (18) oder einer Juxta-Position zu Enhancer Elementen bedingt sein (19,20). Bekannte Mechanismen, die zur Entstehung eines Onkogens führen sind in Abb.1.1 schematisch dargestellt.



**Abb. 1.1: Bekannte Mechanismen zur Transformation von Protoonkogenen zu Onkogenen, verändert nach Carlo, 2008 (13)**

Die Anzahl der betroffenen Onkogene, die Art der Mutation, das Zusammenspiel der Onkogene, sowie der daraus resultierenden Transformation der Zellen, ist sehr heterogen und während der Tumorgenese stetigen Veränderungen unterworfen. Dies lässt sich am Beispiel von Mammakarzinomen darstellen. Durch genomweite Genexpressionsanalysen von Mammatumoren mittels cDNA-Microarrays und anschließenden hierarchischen Clusteranalysen (*unsupervised hierarchical clustering*) konnte gezeigt werden, dass Brustkrebserkrankungen in verschiedene molekulare Subtypen eingeteilt werden können (Abb.1.2). Das kombinierte Expressionsmuster von mehr als 500 Genen wurde als intrinsisches Genset identifiziert und erlaubt die Zuordnung jedes einzelnen Mammatumors zu einer der molekularen Untergruppen, die sich auch durch ein immunhistochemisches Markersset charakterisieren lassen.

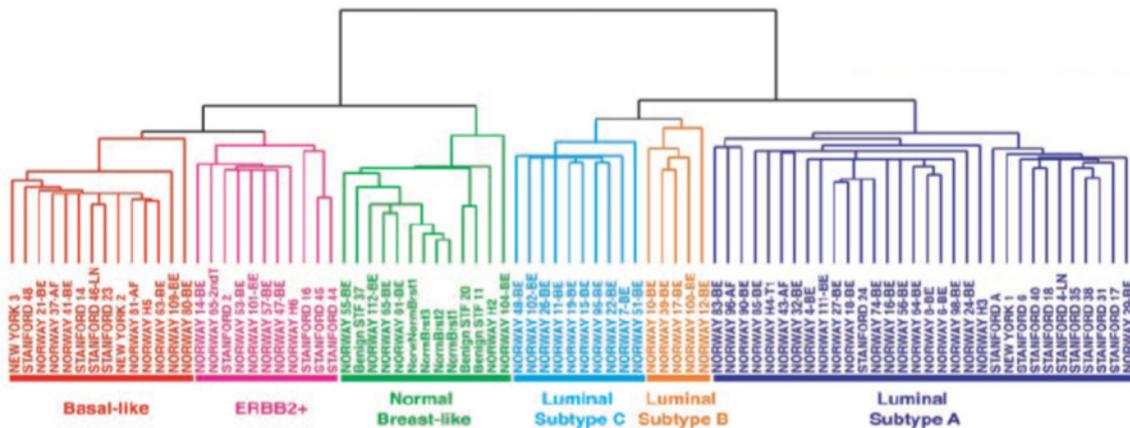


Abb.1.2: Molekulare Klassifizierung von Mammatumoren nach der Clusteranalyse von cDNA Microarrays, entnommen aus Dent et al., 2007 (21)

Luminale Tumoren sind häufig Hormonrezeptor-positiv und entwickeln sich möglicherweise aus den luminalen Zellen der Brustdrüse. Tumoren des ErbB2-Phänotyps weisen eine starke Amplifikation und Überexpression des HER2-Gens auf und sind in der Regel sensitiv für die Behandlung mit anti-HER2-gerichteter Therapie. Als dritte wichtige Gruppe sind die „*basal-like*“-Tumoren zu unterscheiden. Es handelt sich hierbei um Hormonrezeptor- und HER2-negative, häufig jedoch Cytokeratin 5/6- oder EGFR-positive Karzinome mit Eigenschaften der regenerativen Basalzellen. Die molekularen Subtypen des Mammakarzinoms zeigen aufgrund ihrer spezifischen biologischen Besonderheiten starke Unterschiede im Krankheitsverlauf und im Ansprechen auf (neo-)adjuvanten Behandlung. Vor allem die „*basal-like*“-Karzinome sind schnell proliferierende Tumoren mit schlechter Prognose, deren weitere molekulare Charakterisierung von großer klinischer Bedeutung ist.

#### 1.4 Rezeptortyrosinkinasen und die ErbB-Familie

Durch das Human Genom Projekt weiß man, dass bis zu 20% der etwa 32.000 menschlichen Gene für Proteine kodieren, die an der Signaltransduktion mitwirken. Zu dieser Gruppe zählen unter anderem die über 520 Proteinkinasen, welche sich nach Tyrosin- und Serin/Threonin-Spezifität einteilen lassen. Es sind bis jetzt über 90 Proteinkinase-Gene bekannt, von denen 58 für transmembrane Rezeptortyrosinkinasen kodieren, die man wiederum in 20 Klassen unterteilt (22). Da die RTKs eine wichtige Rolle im Wachstum, Differenzierung, Migration und dem Tod einer Zelle spielen, sind sie auch ein Schlüssel zum Verständnis der onkogenen Aktivierung.

Die wichtigsten Ereignisse der Entdeckung der RTK-vermittelten Signaltransduktion und der daraus resultierenden RTK-basierten Krebstherapie sind in Abb.1.3 chronologisch dargestellt.

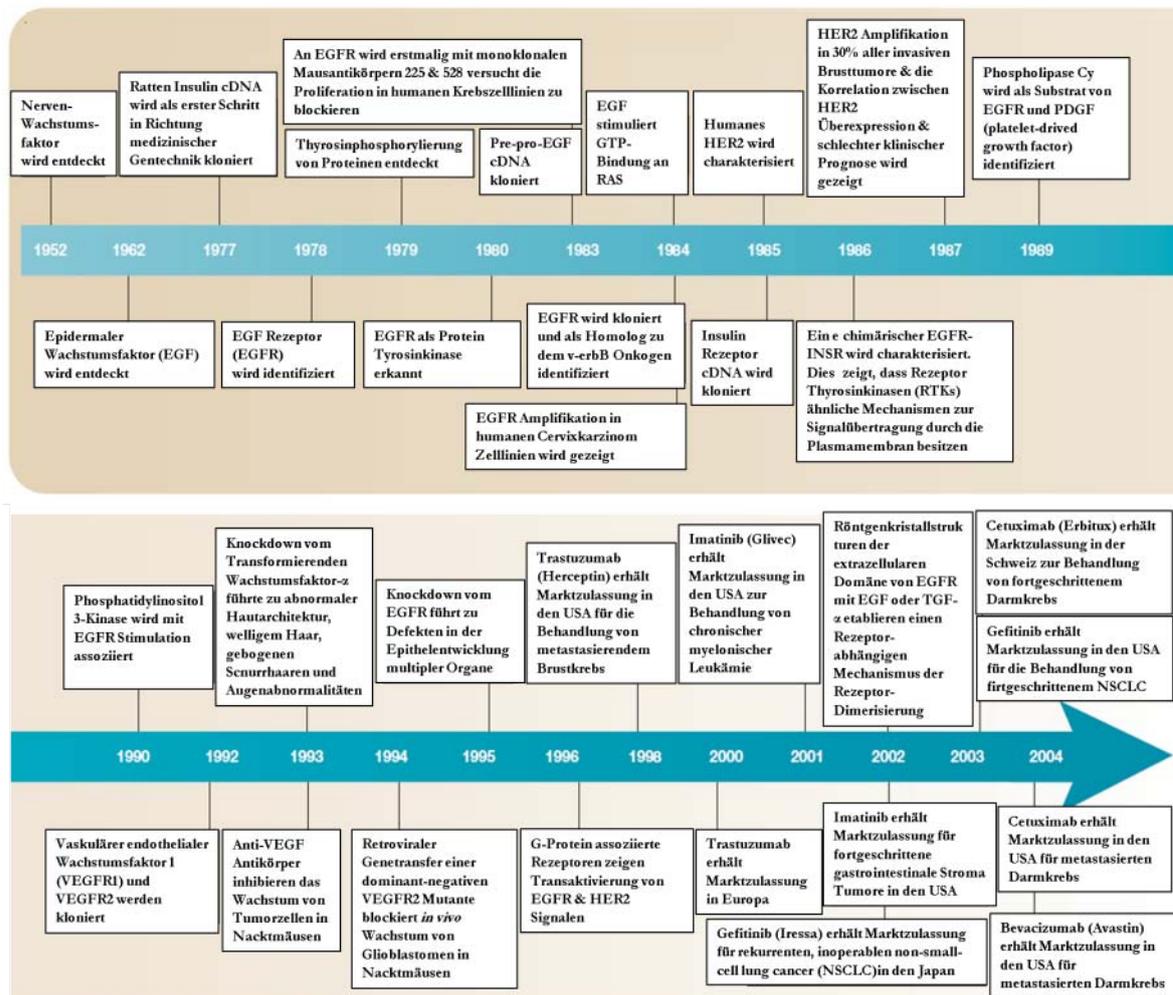
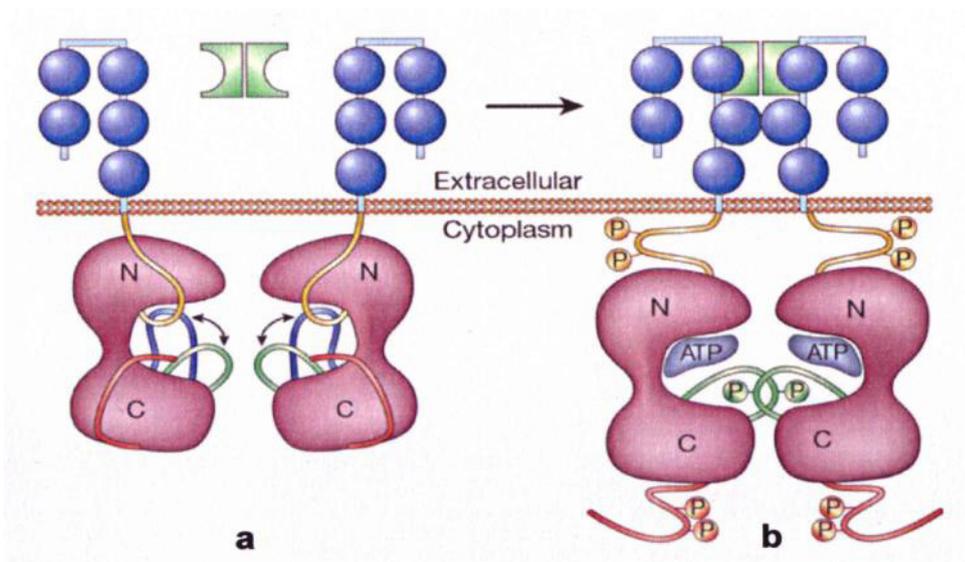


Abb.1.3: Zeitliche Abfolge der Entdeckung der RTK-vermittelten Signaltransduktion und RTK-basierter Krebstherapie, nach Gschwind et al., 2004 (23).

RTKs sind Zelloberflächenproteine die durch eine einzelne Transmembrandomäne, eine intrazelluläre Domäne mit Kinaseaktivität und eine extrazelluläre ligandenbindende Domäne charakterisiert sind. Ligandenbindung führt zu Homo- bzw. Heterodimerisierung. Durch die sterisch günstigere Lage der beiden intrazellulären Domänen zueinander erfolgt eine Aktivierung der Tyrosinkinase-Domänen mit der daraus resultierenden Aktivierung der Auto- bzw. Transphosphorylierung der carboxyterminalen Region (siehe Abb. 1.4)(24-26). Die von den RTKs gebundenen Signalproteine induzieren nun ihrerseits Zellwachstum und Differenzierung. Sie besitzen als Erkennungssequenz ein Modul von 100 Aminosäuren, welches als SH2-Domäne (src-homologe Domäne 2) bezeichnet wird.

Diese erkennen phosphorylierte Tyrosinpeptide und durch Wechselwirkungen von zwei Argininresten und einem Lysinrest mit der Phosphatgruppe und dem aromatischen Ring des Phosphotyrosins erfolgt eine Phosphorylierung des Signalproteins und initiiert eine Signalkaskade aus hintereinander geschalteten biochemischen Reaktionen. Diesen Kaskaden, die auch als „downstream signaling“ bezeichnet werden, unterliegen die Kontrolle über Zellwachstum und Vermehrung, der Apoptose, der Angiogenese, der Adhäsion und der Beweglichkeit von Zellen.



**Abb.1.4: Aktivierung der RTK, entnommen aus Blume-Jensen & Hunter, 2001 (27)**

**a:** Die Tyrosinkinaseaktivität ist im unstimulierten Zustand stark reprimiert. Die Aktivierungsschleife und die katalytische Schlinge liegen im Gleichgewicht zwischen substratabweisender (blau) und substratzugänglicher (grün) Konformation vor. Die juxtamembranäre Region (orange) und die C-terminale Region (rot) könnten die Konformation des N-terminalen Kinaselappens („N<sup>C</sup>“) oder den Substratzugang stören.

**b:** Die Hemmung durch die Aktivierungsschleife und die juxtamembranären und C-terminalen Regionen wird durch Liganden-induzierte Rezeptor-Dimerisierung und Tyrosin-Autophosphorylierung aufgehoben.

Allgemein gilt für alle an malignen Erkrankungen beteiligten RTKs, dass durch Verlust oder Störung einer oder mehrerer Regulations-Mechanismen onkogene Transformation ausgelöst wird. Dieses kann durch folgende Prozesse ausgelöst werden:

- Retrovirale Insertionsmutagenese. Die Insertionen von C-Typ Retroviren in der Nähe eines Onkogens kann zu einer Deregulation der Expression führen. Insertionen von Promotoren mit erhöhter Aktivität sind nur stromaufwärts eines Gens in gleicher transkriptioneller Orientierung aktiv. Enhancer Insertionen können auch stromabwärts zu einer Deregulation führen. C-Typ Retroviren können ebenfalls *v-onc* Gene enthalten. Ein klassisches Beispiel hierfür ist das *v-src* Onkogen des Rous-Sarkom-Virus. Das *v-src* Gen kodiert für eine Tyrosinkinase mit verstärkter Aktivität (27).

- Genomische Umlagerung, z.B. chromosomale Translokation, aus der Fusionsproteine resultieren können, die mindestens eine katalytische Domäne und ein Protein mit Dimerisierungsfunktion beinhalten. Als Beispiel hierfür ist das Philadelphia Chromosom zu nennen, das in über 95% aller Fälle von chronischer myeloischer Leukämie zur Entstehung einer chimäre RTK führt (28).
- „Gain-of-Function“ Mutationen oder kleine Deletionen. Als Beispiel ist das EGFRvIII Transkript zu nennen, bei dem es durch eine Deletion zum Verlust der Ligandenbindedomäne kommt, die eine konstitutive Aktivierung der RTK auslöst (14).
- Überexpression durch Genamplifikation. Diese bedingt eine ansteigende Konzentration aktiven Proteins. Als Beispiel hierfür ist die *HER2-neu* Genamplifikation zu nennen, die häufig in Mammakarzinomen auftritt (58)

Die Proteine der ErbB-Familie stellen von allen im Zusammenhang mit der Entstehung und Entwicklung des Mammakarzinoms untersuchten Wachstumsfaktoren die am besten beschriebenen dar. Sie spielen in der Differenzierung des Brustdrüsengewebes eine entscheidende Rolle und es gibt Anzeichen dafür, dass bestimmte genetische Veränderungen innerhalb der ErbB-Familie mit verschiedenen morphologischen Typen des Mammakarzinoms assoziiert werden können (29,30).

Die Bezeichnung “erbB” resultiert aus einer Homologie der kodierenden Gene zu dem retroviralen *v-erbB*-Onkogens, das im allgemeinen Krebskrankungen in Zellen der erythropoetischen Systems hervorruft (12,32). Die ErbB-Familie, für die als Synonym die Bezeichnung HER verwendet wird, setzt sich aus erbB1 (EGFR), erbB2 (HER2), erbB3 und erbB4 zusammen (31), die alle den in Abb.1.6 beschriebenen Aufbau besitzen (33). Die Gruppe der als Liganden für die ErbB-Familie bekannten Wachstumsfaktoren setzt sich aus EGF-ähnlichen Liganden und Neuregulinen zusammen (35), die über ihre bevorzugte Bindung an bestimmte Rezeptoren zu unterschiedlichen Rezeptor-Dimeren führen und somit verschiedene Signaltransduktionswege aktivieren (31,34). Der Rezeptor HER2 besitzt keinen eigenen Liganden und dient in erster Linie als Dimerisierungspartner der drei übrigen Rezeptoren. Neben der Überexpression von HER2, die zur Transaktivierung durch Rezeptor-Dimere führt, wurden für die ErbB-Familie aktivierende Mutationen, sowie der Ausfall von inaktivierenden Mechanismen beschrieben (36).

### 1.5 Der Epidermale Wachstumsfaktor-Rezeptor (EGFR)

Für EGFR wurden in zahlreichen Studien an unterschiedlichen Tumorentitäten hohe Überexpressionsraten erhoben (Tab.1.1). Da die Überexpression von EGFR zu einer Steigerung der Zellproliferation, Angiogenese und Metastasierungsrate mit gleichzeitiger Verringerung der Apoptose führt, ist EGFR ein interessantes Ziel für die „*Targeted Therapy*“ (Gezielte Therapie).

Tumorart	EGFR-Überexpressionsrate
Mamma-CA	14%-91%
Colon-CA	25%-77%
Ovarial-CA	35%-70%
Pankreas-CA	30%-50%

**Tab.1.1: EGFR-Überexpression in menschlichen Tumoren, verändert nach Buerger et al. 1999 (29)**

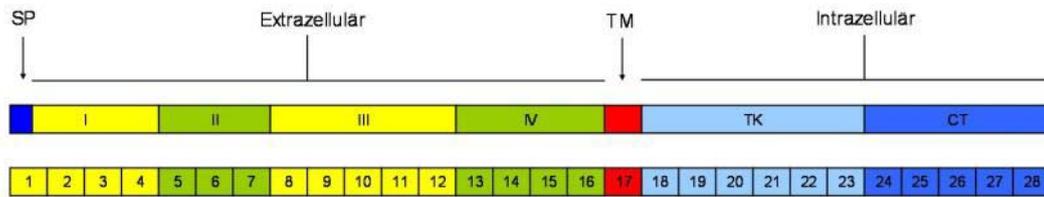
Die erhobenen Daten sind sehr stark abhängig von dem verwendeten Antikörper, des Protokolls des immunhistochemischen Nachweises, der Auswahl des Scoring Systems und der Festlegung des Cutoffs (37). Eine Vereinheitlichung dieser Parameter konnte bis jetzt nicht erzielt werden.

Der EGF-Rezeptor, ein transmembranes Glykoprotein, ist als Prototyp der erbB-Rezeptor-Familie zu nennen, das über eine intrinsische Tyrosinkinaseaktivität verfügt. Das Protein setzt sich aus 1186 Aminosäuren (38) zusammen und besitzt ein Molekulargewicht von 170 kDa (39). Die extrazelluläre, ligandenbindende Domäne besteht aus 612 Aminosäuren (40), besitzt ein Molekulargewicht von 100 kDa (41) und wird in vier Bereiche eingeteilt (42). In den N-Terminus (43), zwei cysteinreiche Regionen (44) und den Bereich, der die beiden cysteinreichen Regionen voneinander trennt. Am N-Terminus findet man an etwa zwölf Asparaginresten die Möglichkeit zur Glycosylierung (38), welche die Bindungsaffinität von Ligand und Rezeptor beeinflusst (45). Die cysteinreichen Regionen sind zwischen den Aminosäuren 134-313 und 446-612 lokalisiert und beinhalten insgesamt 51 Cysteinreste (38). Durch die Ausbildung von Disulfidbrücken untereinander führt zur Tertiärstruktur der extrazellulären Domäne (45). Ein weiterer wichtiger Bestandteil der extrazellulären Domäne sind Kohlenwasserstoffe, die in Form von 10-11 N-verknüpften Oligosaccharid-Ketten vorliegen. Die 23 Aminosäuren lange transmembrane Domäne befindet sich zwischen den Aminosäuren 622 und 644. Dieser hydrophobe Teil des Rezeptors bildet eine  $\alpha$ -Helix-Struktur aus (44,42,46).

Die intrazelluläre Domäne besteht aus 542 Aminosäuren (47), ist 60 kDa (40) schwer und setzt sich aus der Tyrosinregion, sowie der juxtamembranären und der Carboxylregion (42) zusammen. Innerhalb der juxtamembranären Region (645-689) (48) an der Position 654 befindet sich ein Threoninrest, der ein Phosphorylierungsziel der Proteinkinase C bildet (49). Dies führt zu einer Senkung der Tyrosinkinaseaktivität des Rezeptors und fördert dessen Abbau (48). Das Molekulargewicht der 250 Aminosäuren-langen Tyrosinkinase-region (690-940)(Ref.38) beträgt 40 kDa (50). Ihre zentrale Einheit setzt sich aus zwei Peptidschlaufen und einer Peptidasche zusammen. Das an der größeren Schlaufe gebundene ATP, ebenso wie ein Teil des an der kleineren Schlaufe gebundenen Substrates reichen in die Tasche hinein, in der dann die Phosphorylierung stattfindet (51). Die Autophosphorylierungsstellen des Rezeptors befinden sich in der 246 Aminosäuren-langen und 20 kDa schweren (49) carboxyterminalen Region. Die im C-Terminus befindlichen Tyrosinreste 1148, 1068 und vor allem 1173 werden am häufigsten phosphoryliert (32).

## 1.6 Die Genstruktur und Transkriptionsregulation von *EGFR*

Das kodierende Gen für den humanen epidermalen Wachstumsfaktor-Rezeptor ist auf dem Chromosom 7p11.2 lokalisiert und umfasst 193 kbp (52). Durch Alignements der *EGFR*-cDNA mit der genomischen Sequenz konnte gezeigt werden, dass der EGF-Rezeptor von insgesamt 28 Exonen kodiert wird (43,53,54). Alle Exone werden flankiert von GT-AG Konsensus-Splice Stellen (39,55). Wie von Callghan und Mitarbeitern bereits 1993 (56) beschrieben, umfasst das humane Exon 1 die 5'-UTR (Untranslated Region) und Sequenzen, die für das Signalpeptid und die ersten fünf Aminosäuren des reifen EGF-Rezeptors kodieren. Die extrazelluläre Region wird durch die Exone 2-16 kodiert, wobei die beiden cysteinreichen Regionen von den Exonen 5-7 und 13-16 kodiert werden. Exon 17 kodiert für die transmembrane Region, die Exone 18-24 für die intrazelluläre Tyrosinkinasedomäne und die Exone 25-28 für die intrazelluläre carboxyterminale Region des Rezeptors (siehe Abb.1.5). Besonders auffällig an der Genstruktur des EGF-Rezeptors ist das mit 123 kbp extrem große Intron 1 (55).

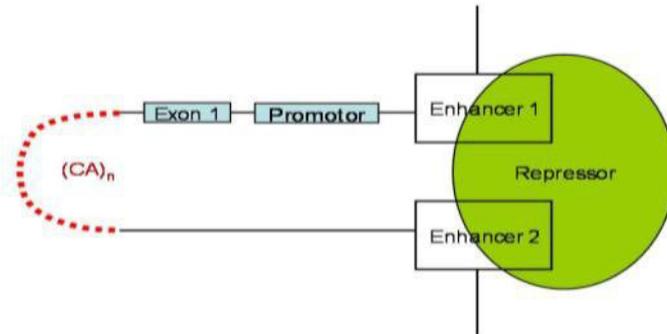


**Abb.1.5: Schematische Darstellung der Subdomänen des EGF-Rezeptors und die dafür kodierenden Exone, verändert nach Reiter et al., 2001 (55): SP: Signalpeptid; TM: Transmembrane Domäne; II und IV: zwei cysteinreichen Regionen; TK: Tyrosin-Kinase-Domäne; CT: Carboxyterminale Region**

Die Richtung der *EGFR*-Transkription von Telomer nach Zentromer wurde durch den Abgleich von 17 STS (Sequence Tagged Site) Markern mit der genomischen Sequenz ermittelt (55).

Anhand der Zelllinie NIH-3T3 konnte gezeigt werden, dass die Überexpression von EGFR eine zelluläre Transformation bedingen kann (57). Eine Überexpression durch Amplifikationen von *EGFR* wurde für verschiedene Tumorentitäten mit unterschiedlicher Häufigkeit beschrieben. In Glioblastomen tritt diese mit einem Anteil von 60% häufig auf. Für Mammakarzinome hingegen war eine Amplifikation, wie sie z. Bsp. für das *HER2/neu*-Gen beschrieben wurden (58), in weniger als 5% aller Fälle nachweisbar. Es wird daher davon ausgegangen, dass beim Brustkrebs eine Überexpression von EGFR normalerweise nicht durch eine erhöhte Gen-Dosis ausgelöst wird, sondern die Expression auf der Transkriptionsebene reguliert wird. Das *EGFR*-Gen besitzt verschiedene Initiationsstellen oberhalb des GC-reichen Promotors, an denen die Transkription beginnt. Der Promotor besitzt weder eine TATA- noch eine CAAT-Box (59). Wie auch von anderen Genen bekannt, spielt das Intron 1 eine wichtige regulatorische Rolle. Es finden sich sowohl upstream als auch downstream des Promotors im Intron 1 zwei Enhancer (60). Nahe am Downstream-Enhancer-Element findet man einen polymorphen SSR (simple sequence repeat), der aus 14-21 CA-Dinukleotiden (CA-SSR I) besteht und in einer kaukasischen Referenzbevölkerung eine Heterozygotie von 72% aufweist (61). Die Länge dieses CA-Repeats korreliert negativ mit der Transkriptionsrate (62). Man versucht, sich dieses Phänomen wie folgt zu erklären (63): Es konnte gezeigt werden, dass der Bereich des Poly-CA-Stückes auffallend biegsam ist und der Hang zu intrinsischer DNA-Krümmung wenig ausgeprägt. Je länger der CA-Repeat, desto biegsamer der Abschnitt. Dies könnte bedeuten, dass ein Repressorprotein eventuell beide Enhancer binden und damit Sekundärstrukturen der DNA, z. B. Schleifenbildung, erzwingen könnte.

In dieser würde der CA-Repeat als eine Art Gelenk fungieren und es würde eine Erklärung dafür geben, dass die Repression der Translation bei längerem CA-Repeat, also biegsamerer DNA, begünstigt wird (Abb.1.6).



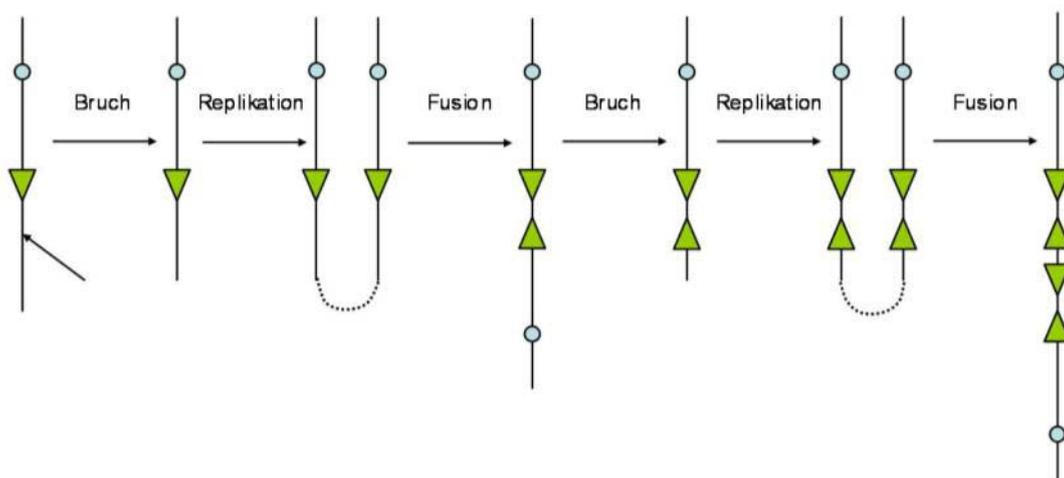
**Abb.1.6: Modell zur transkriptionellen Regulation der EGFR-Expression, verändert nach Gebhardt et al., 1999 (62):** Die Biegsamkeit der 5'-Region des *EGFR*-Gens wird durch den CA-SSR I im Intron 1 bestimmt. Ein möglicher Repressor kann durch den längeren CA-SSR I begünstigt binden und die DNA in Sekundärstrukturen bringen.

Ein „Loss of heterozygosity“ (LOH) bezeichnet die Unterrepräsentation eines Allels in einem Gewebe, wie sie mittels Mikrosatelliten-PCR ermittelt werden kann. Diesem Phänomen können entweder der Verlust oder die Amplifikation eines Allels zugrunde liegen. Die Allele liegen also im Ungleichgewicht vor, man spricht daher von Allelic Imbalance (AI). Tumore, in denen sich die Allelverteilung zugunsten des kürzeren CA-Repeats verschoben hat, zeigten eine höhere intratumorale EGFR-Konzentration (64). CA-SSR I ist innerhalb des *EGFR*-Gens im Mammakarzinom ständiges Ziel von Mutationen. Hierbei handelt es sich um Amplifikationen, welche wahrscheinlich die ersten Schritte in der Mammakarzinogenese repräsentieren und zu einem Brustkrebstyp führen, der mit schlechten klinischen Prognosen, wie einer erhöhten Rate an Tumorrezidiven, in Verbindung gebracht wird. In Gen-Dosis-Studien konnten jedoch nur minimale Anzeichen für die Neuordnung oder Amplifikation des *EGFR*-Gens nachgewiesen werden (66). Die Überexpression von EGFR wurde deshalb hauptsächlich als Resultat gestörter Transkriptionsmechanismen betrachtet. In 20% der Patientinnen mit einer Amplifikation wurden von 6 oder mehr Genkopien gemessen. Weitere 20% der Patientinnen mit erhöhter EGFR-Expression zeigten hingegen eine gesenkte Kopienzahl (67). Die verwendeten Primer befanden sich in der bekannten regulatorischen Sequenz des *EGFR*-Gens in der 5'-Region an der Grenze zwischen Exon 1 und Intron 1 des *EGFR*-Gens. Da in beiden Fällen eine erhöhte intratumorale EGFR-Konzentration festgestellt werden konnte, werden beide Patientengruppen der Hochrisikogruppe für den frühen Beginn der Tumorprogression zugeordnet (68).

Die genetischen Abweichungen finden sich in den ersten 1200 Basen des Intron 1 und Hinweise deuten auf ausgeprägte Amplifikationen des CA-SSR-I Bereiches im Intron 1 des *EGFR*-Gens (69). Da der Mechanismus, welcher diese Amplifikationen bedingt, noch ungeklärt ist, muss zunächst auf Modellvorstellungen zurückgegriffen werden.

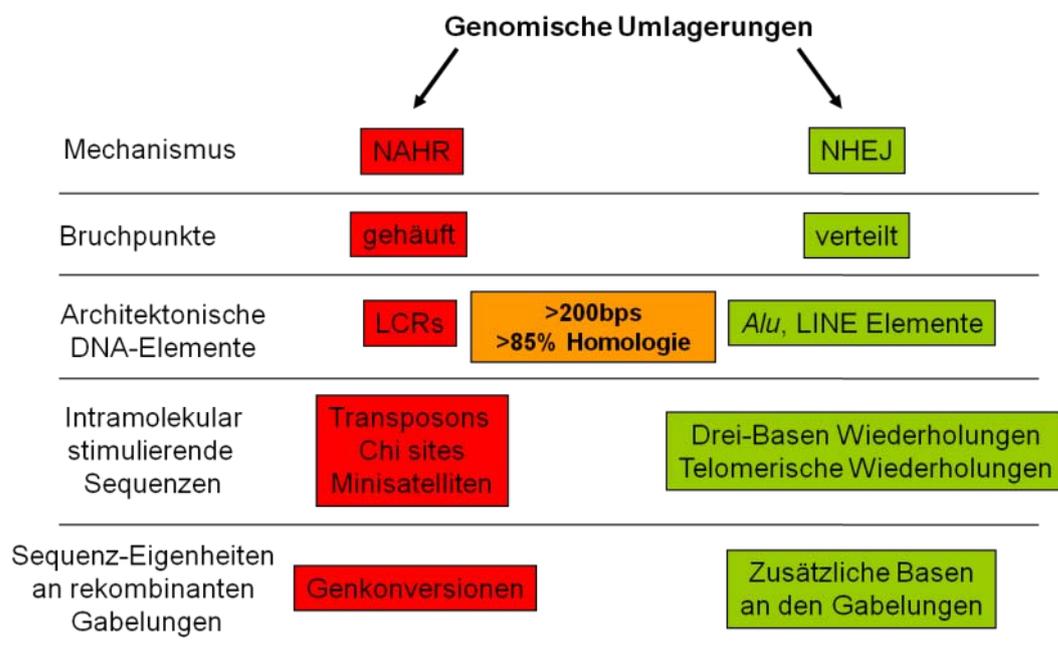
### 1.7 Mögliche Ursachen für die Entstehung von *EGFR* Amplifikationen

Ein möglicher Ansatzpunkt für den Entstehungsprozess von *EGFR* Amplifikationen bieten „Fragile Sites“, bei denen es sich um wenig verdichtete Regionen in Metaphase-Chromosomen handelt. An diesen kommt es während der DNA Replikation vermehrt zu Strangbrüchen, wobei zwischen Common Fragile Sites (CFSs) und Rare Fragile Sites (RFSs) differenziert wird, die nur in <5% aller Individuen einer Population auftreten (70). Mutagene Agenzien erhöhen die Wahrscheinlichkeit eines Strangbruches (71-72). Der Zusammenhang zwischen genetischer Prädisposition und der Aktivierung von CFSs, konnte für eine CFS auf Chromosom 7p13-14, nahe dem *EGFR* Lokus (74), sowie andere CFSs (75), bestätigt werden. Dabei waren häufig auch chromosomale Regionen mit anderen Genen betroffen, zu denen eine Assoziation mit humanen malignen Erkrankungen bestand (73). Durch die Aktivierung einer Fragile Site kann der von McClintock beschriebene BFB (Breakage Fusion Bridge) Zyklus (76) initiiert werden (Abb.1.9). Bei diesem kann es während der Anaphase, nach dem Verlust des schützenden Telomers (78), zur Fusion und anschließenden Ungleichverteilung von zwei Schwesterchromatiden auf die Tochterzellen kommen und somit zur Amplifikation definierter genomischer Bereiche (62,72) führen.



**Abb.1.7: Schematische Darstellung des strangbruchinduzierten BFB-Mechanismus, verändert nach Shuster et al. 2000 (77):** Dieser führt zu Inverted Repeats bestehend aus den amplifizierten DNA-Segmenten. Die blauen Kreise symbolisieren Zentromere, die grünen Dreiecke Gene.

Ein weiterer Ansatzpunkt zur Erklärung des Amplifikationsmechanismus bietet der DNA Reparaturmechanismus des „non-homologous end joining“ (NHEJ), bei dem es regelmäßig zum Verlust von Basen kommt. Als Folge eines fehleranfälligen NHEJ zwischen zwei Chromosomen kann es zu somatischen, tumorrelevanten Chromosomenveränderungen, wie unbalancierten Translokationen kommen (80). Als Initiatoren für ein NHEJ können, unter anderem, die im Intron 1 lokalisierten *Alu I* Elemente dienen. Die Existenz einer RFS auf Chromosom 7p11.2, deren Induktion in einem Einzelstrangbruch nahe des *EGFR* Lokus resultierte(73,79), könnte gemeinsam mit dem Prozess der nicht-allelichen homologen Rekombination (NAHR) einen weiteren Erklärungsansatz darstellen. NAHR kann durch ungleiches „cross-over“ zur Entstehung von Mikrodeletionen und Mikroduplikationen führen. Diese molekularzytogenetisch darstellbaren Imbalancen (81) können zum Verlust bzw. Gewinn von mehreren Genen führen (82). Der Prozess der NAHR findet zwischen repetitiven Elementen, den sogenannten „low copy repeats“ (LCR) statt (83-84). Die LCRs, zu denen auch die im Intron 1 lokalisierten MER22 gehören, tragen hierbei Signalstrukturen, die bei der Rekombination eine essentielle Rolle spielen (85-88). Eine Zusammenfassung der für NHEJ und NAHR beschriebenen Charakteristika ist in Abb.1.8 dargestellt.

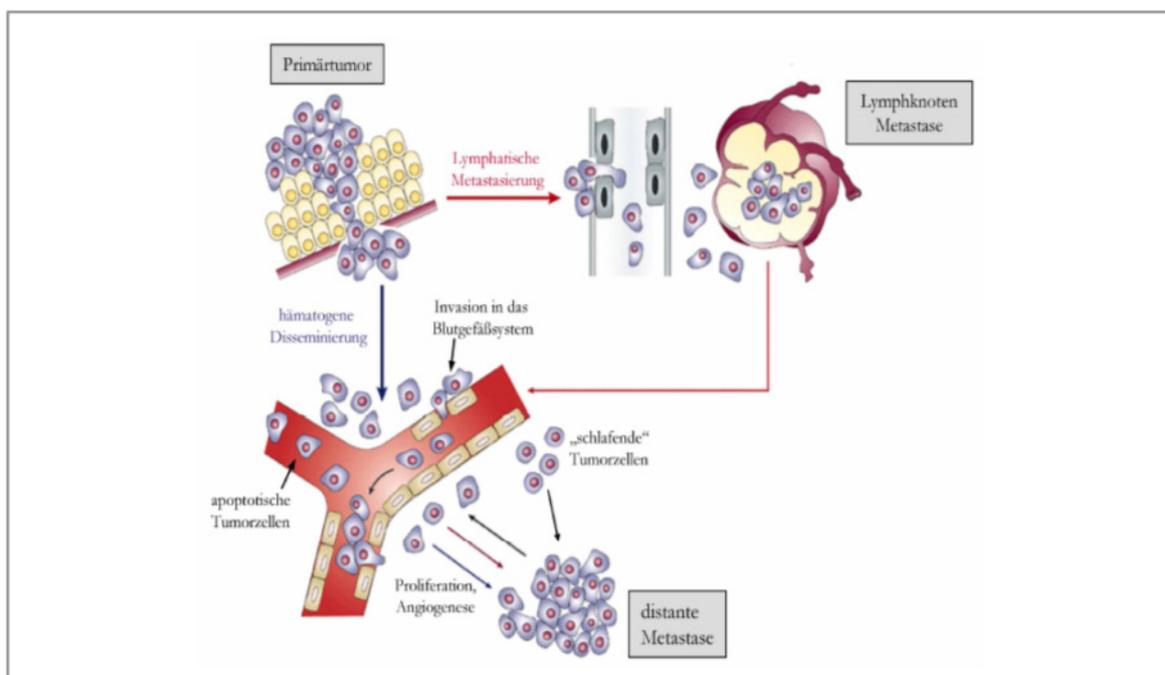


**Abb.1.8: Charakteristika der NAHR und NHEJ Mechanismen, verändert nach Shaw & Lupski, 2004 (89)** Dargestellt sind DNA architektonische Elemente und Charakteristika, welche für die NAHR und NHEJ Mechanismen jeweils verschieden sind, sowohl als aktivierende Elemente, als auch Eigenschaften, die nach Rekombinationsprozessen Hinweise auf einen möglichen Mechanismus geben.

Eine Aufklärung des Amplifikationsmechanismus von *EGFR* könnte zu Tage fördern, dass hier eine genetische Prädisposition vorliegt und z.B. durch die Entwicklung eines PCR-basierten Tests zu einer Risikoabschätzung für Patienten führen.

### 1.8 Metastasierung und minimal residuale Erkrankungen

Die Ursache der hohen Mortalitätsrate bei invasiven Karzinomen liegt aus klinischer Sicht nicht im Primärtumor. Mammakarzinome sind z. Bsp. in den meisten Fällen chirurgisch zu entfernen und die Rate der In-Brust-Rezidive liegt in Kombination mit einer Strahlentherapie der betroffenen Brust bei ~4-9%. Maligne Tumore besitzen jedoch die Eigenschaft, im Körper zu migrieren und in anderen Teilen des Körpers Metastasen auszubilden. Maßgebliche Risikofaktoren hierfür stellen zum einen die Anzahl karzinomatös infiltrierter, axillärer Lymphknoten (LKs) dar, die direkt mit der Rückfallrate korrelieren (90), zum anderen die Größe des Primärtumors, der direkt mit der Überlebensrate korreliert (91). Die Verbreitung von Tumorzellen kann über den lymphatischen oder über den hämatogenen Weg stattfinden (s. Abb.1.9).



**Abb.1.9: Schematische Darstellung der Metastasierungskaskade, verändert nach Pantel & Brakenhoff, 2004 (92):** Die Metastasierung der Tumorzellen ist sowohl auf lymphatischem (rote Pfeile), als auch auf hämatogenem Weg (blaue Pfeile) möglich. Lymphatisch disseminierende Tumorzellen wandern hierbei in die lokalen LKs ein und proliferieren zu soliden Metastasen. Eine spätere Disseminierung in distante Organe durch die LK Metastasen und die Formation von sekundären Metastasen ist ebenfalls möglich. Die hämatogene Disseminierung kann sowohl von Primärtumor, LK Metastasen, als auch sekundär Metastase ausgehen.

Maßgeblich am Metastasierungsprozess beteiligt sind die Bildung von proteolytischen Enzymen, der Verlust von Zelladhäsionsmolekülen, sowie die Angiogenese. Durch sie wird das Ablösen von Tumorzellen aus dem Zellverband des Primärtumors und das Eintreten in Lymph- bzw. Blutbahnen nach Invasion und Penetration des Stromagewebes ermöglicht (93-96). Die zirkulierenden Tumorzellen müssen den Scherkräfte und Angriffen des Immunsystems widerstehen, sowie die Apoptose, die durch Verlust von Zell-Zell-Kontakt induziert wird, umgehen können (97).

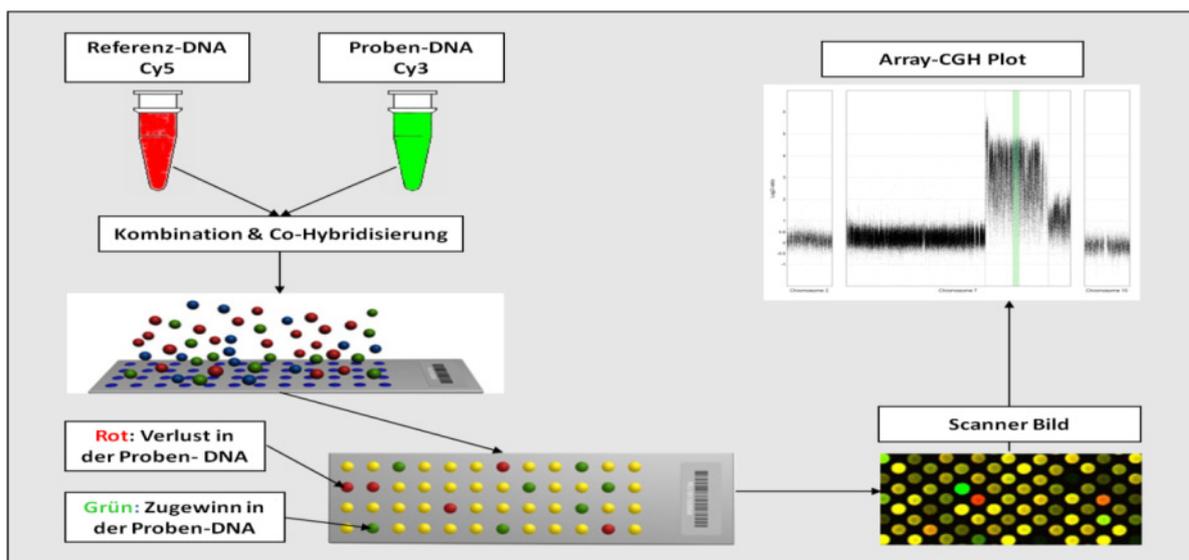
Adhäsionsproteine können anschließend den Kontakt zum Kapillarbett anderer Organe herstellen. Nach Durchdringen des Endothels, der endothelialen Basalmembran (98), Invasion des Stromas und Aktivierung der Proliferation kann sich eine solide Metastase ausbilden. Aus unbekanntem Gründen kommt es vor, dass die Proliferation erst nach einer teilweise jahrelangen Latenzphase (Dormancy) einsetzt (99,100). Die Metastasierung einzelner Tumorzellen scheint entgegen klassischen Modellen (101) ein frühes Ereignis in der Tumorgenese darzustellen (92,102). Nur ein geschätzter Anteil von <1% der nachweislich großen Anzahl an vom Tumor freigesetzten Tumorzellen (103) besitzt die Fähigkeit, Metastasen zu bilden (104,105). Das Einwandern einzelner Tumorzellen, den „disseminierten Tumorzellen“ (DTC), ohne die Ausbildung von soliden Metastasen, bezeichnet man als minimale residuale Krebserkrankung (*Minimal Residual Disease, MRD*) (106). Der Nachweis dieser Zellen wurde 1981 erstmals erfolgreich in Knochenmarks-Aspiraten durch epitheliale Marker geführt (107). Der immunzytochemische Nachweis von DTC über epithelspezifische Proteine in den mesenchymalen Kompartimenten Blut, Lymphknoten und Knochenmark zeigte eine Korrelation zwischen dem Vorhandensein von DTC und der Entstehung von Metastasen (108,109). Als spezifischer Marker für DTC in zytologischen Präparaten des KM etablierte sich der Nachweis von Zytokeratin-positiven Zellen (CK<sup>+</sup>-Zellen) durch einen Pan-Antikörper, der sich gegen ein gemeinsames Epitop verschiedener Zytokeratine richtet (110,111). Diese werden, als integraler Bestandteil des Zytoskeletts, von epithelialen Zellen stabil exprimiert. Da in ca. 1% der gesunden Probanden ebenfalls CK<sup>+</sup>-Zellen im KM nachgewiesen wurden, kann eine ektopische Expression in mesenchymalen Zellen nicht vollständig ausgeschlossen werden (112). In acht von zwölf Studien wurde gezeigt, dass das Vorhandensein von DTC im KM ein unabhängiger Prädiktor für eine schlechte Prognose darstellt (108,112-119). Aufgrund der prognostischen Relevanz des Auftretens der DTC soll eine genomische Charakterisierung der Zellen, neben einem verbesserten biologischen Verständnis der Disseminierung, ebenfalls Aufschluss über den möglichen Nutzen des Einsatzes von „Targeted Therapies“ geben.

Neben der bereits beschriebenen Therapie mit Herceptin für HER2-überexprimierende Tumore stehen ebenfalls Tyrosinkinaseblocker für EGFR-überexprimierende Tumore zur Verfügung. Als zugelassene Substanzen sind unter anderem Gefitinib (Iressa), Erlotinib (Tarceva), sowie die monoklonalen Antikörper Cetuximab (Erbix) und Panitumumab (Vectibix) zu nennen.

## 1.9 DNA-Quantifizierungsmethoden zur Untersuchung des *EGFR* Amplikon

### 1.9.1 *Fine Tiling Array-CGH*

Die Methode der CGH (Comparative Genomic Hybridization) dient zum spezifischen Nachweis von Veränderungen in der DNA Kopienzahl. Die Methode beruht auf der gleichzeitigen Hybridisierung gleicher Anteile von unterschiedlicher Fluoreszenzmarkierter Test- (Tumor-DNA) und Referenz-DNA (normale DNA) auf Metaphasechromosomen (CGH) oder auf Mikroarrays mit cDNA-, BAC-Klone oder Oligonukleotiden (siehe Abb. 1.10). Die Mikroarrays werden anschließend gescannt und die Stärke der Fluoreszenz gemessen. Das Verhältnis der Fluoreszenzintensitäten von Test- und Referenz-DNA gibt Aufschluss über das Vorliegen von chromosomalen Imbalanzen. Ein Zugewinn der Test-DNA wird dabei durch eine relative Erhöhung von der Test-Fluoreszenz zu der Referenz-Fluoreszenz detektiert. Bei einem Verlust in der Test-DNA wird dementsprechend eine relative Erhöhung Referenz-Fluoreszenz detektiert.



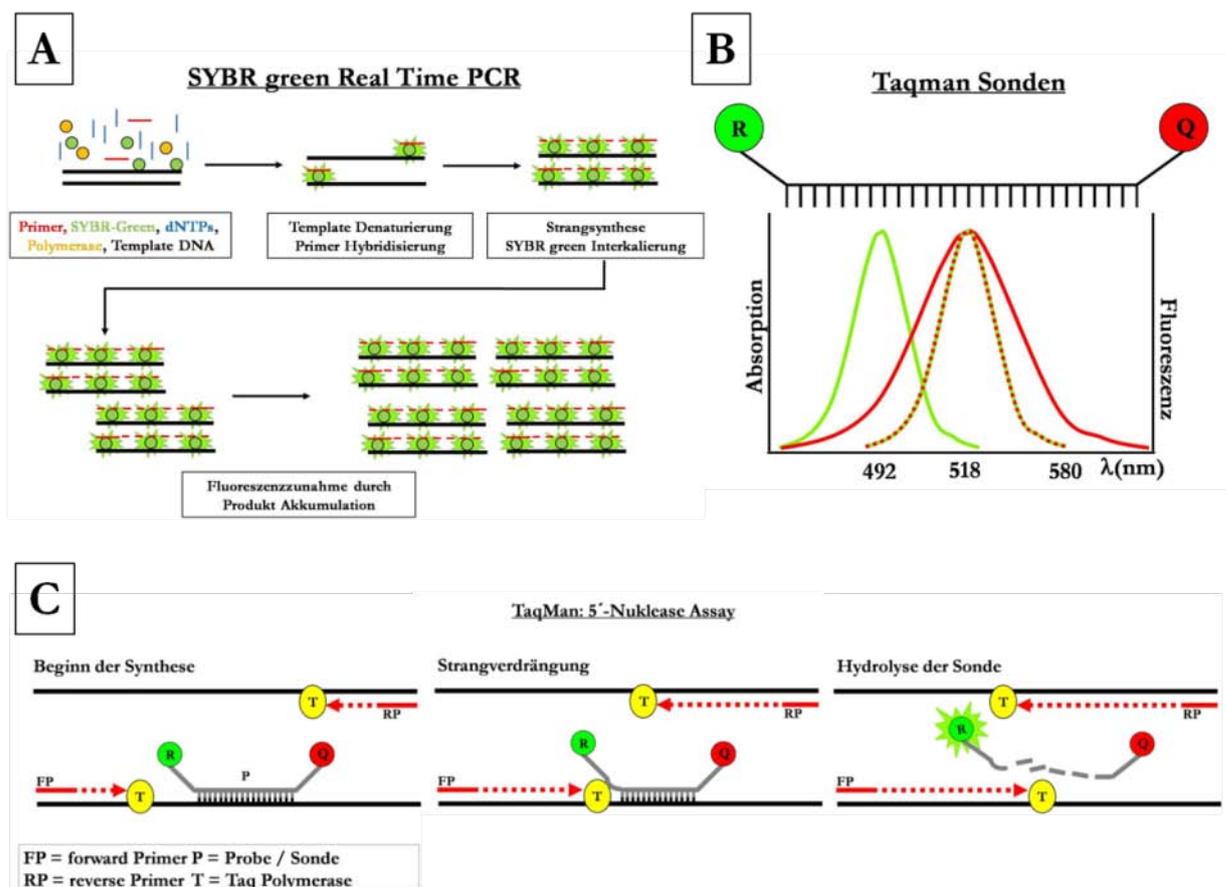
**Abb.1.10: Übersicht zur Durchführung einer Array-CGH.** Referenz- und Proben-DNA werden mit unterschiedlichen Fluoreszenzfarbstoffen markiert (Cy5 bzw. Cy3) und in gleichen Anteilen zusammengegeben. Bei der kompetitiven Hybridisierung konkurrieren die Tumor- und die Referenz-DNA um die homologen Bindungsstellen des Arrays. Anhand des Verhältnisses aus den Fluoreszenzintensitäten der Tumor- und der Referenz-DNA können DNA-Überrepräsentationen bzw. DNA-Verluste der Tumor-DNA bestimmt werden.

Die klassische CGH besitzt mit 2mb für DNA-Gewinne und 10mb für DNA-Verluste, was in etwa einer chromosomalen Bande entspricht, ein verhältnismäßig geringes Auflösungsvermögen. Durch die Verwendung von Mikroarray-CGHs wurde das Auflösungsvermögen um mehr als eine Zehnerpotenz erhöht. Der Einsatz von cDNA-Klonen ermöglichte, die Untersuchung von Veränderungen in der DNA-Kopienanzahl und der Genexpression auf einer gemeinsamen Plattform. Mit einem wesentlich höheren Auflösungsvermögen von 267kb zeigten cDNA-Mikroarrays aufgrund ihrer begrenzten Anzahl an Klonen und ihrer schwachen Signalintensität allerdings starke Schwächen. Kurze Zeit später erschien der erste Mikroarray mit DNA aus 2.600 BAC Klonen und einer Auflösung von 1,2mb auf dem Markt. Die Verwendung eines 30.000 BAC-Arrays erhöht das Auflösungsvermögen um eine weitere Zehnerpotenz, ihre Herstellung ist allerdings zeitaufwendig und kostspielig. Kostengünstiger und flexibler in der Zusammenstellung sind Mikroarray-CGHs, die mit mehr als 30.000 Oligonukleotiden das vollständige humane Genom abdecken und mit ~100kb ein hohes Auflösungsvermögen besitzen. Da die Auswertungen auf Gen-Ebene stattfinden, erlaubt die Verwendung des gleichen Arrays für CGH und Expressionsanalysen einen direkten Vergleich von mRNA-Expression und DNA-Kopienanzahl. Das *Fine Tiling Array* ist dafür konzipiert einen DNA-Sequenzbereich von mit einer Auflösung von ~0,5kb darzustellen. Die als Sonden verwendeten Oligonukleotide werden dabei so angeordnet, dass der maximale Abstand zwischen den Sonden kleiner sein muss, als deren jeweilige Länge. Dies führt zu einer Überlappung (Tiling) der Sonden mit singulärer Sequenz, wodurch die Konsistenz der Daten erhöht wird. Hoch-homologe DNA Regionen, die keine singulären 50-75 Basen langen Sequenzen aufweisen, können mit *Fine Tiling Arrays* nicht erfasst werden. Das in dieser Arbeit verwendete Array bestand aus 395000 Oligonukleotiden einer Länge von 50-75 Basen und einem Tiling von 15 Basen. Dadurch konnte eine DNA-Sequenzlänge von ~4,7mb erfasst werden.

Eine weitere Mikroarray-Art zum Nachweis von Veränderungen in der DNA-Kopienanzahl stellen SNP-Arrays dar. Einzelnukleotid-Polymorphismen (Single Nucleotide Polymorphism, SNP) stellen die häufigste Quelle genetischer Variation beim Menschen dar. SNPs eignen sich aufgrund ihrer hohen Dichte als Marker zur Lokalisation von krebsassoziierten Genen, können aber auch selbst zur Tumorentstehung und dessen Progression beitragen. SNP-Arrays erlauben die gleichzeitige Bestimmung von LOH und DNA-Kopienanzahl. Mit bis zu 100.000 SNPs wird eine höhere Auflösung als bei Oligonukleotid Mikroarray-CGHs oder gängige Mikrosatellitenstudien erreicht, wobei die Sensitivität im Vergleich zu Mikroarrays basierend auf Oligonukleotiden verringert ist.

### 1.9.2 Quantitative Real-Time PCR

Die Quantitative Real-Time PCR (qPCR) ist eine Methode zur Vervielfältigung von DNA, basierend auf dem Prinzip der herkömmlichen Polymerase-Kettenreaktion (PCR), die zusätzlich die Quantifizierung der gebildeten DNA in Echtzeit erlaubt. Die Quantifizierung erfolgt jeweils am Ende eines jeden PCR-Zyklus über die Fluoreszenz eines Reporter-moleküls, die proportional zur gebildeten DNA-Menge zunimmt. Der Methode liegt die Annahme zugrunde, dass die erste signifikant messbare Zunahme an Fluoreszenz proportional zu der Menge des eingesetzten Startmaterials ist. Dies setzt eine maximale Effizienz der Reaktion voraus, die einer Verdopplung pro PCR-Zyklus entspricht. Die korrekte Quantifizierung kann deshalb nur im exponentiell verlaufenden Teil der PCR-Reaktion durchgeführt werden, der oft nur wenige Zyklen andauert. Als Reportermoleküle werden hauptsächlich der DNA-interkalierende Farbstoff *SYBR green* sowie doppelt gelabelte Hydrolyse-Sonden (TaqMan-Sonden) verwendet (Abb.1.11).



**Abb.1.11: Quantitative Real-Time PCR: A.** Prinzip des *SYBR green Assays*, Fluoreszenz über Interkalierung; **B.** TaqMan Sonde mit sich überlagernden Extinktions- und Absorptionsspektren bei 518nm von Reporter (grün) und Quencher (rot); **C.** TaqMan 5'-Nuklease Assay, Fluoreszenz-Zunahme durch Hydrolyse der Sonde

Dem *SYBR green Assay* liegt eine Fluoreszenzzunahme durch Interkalierung des Reporter-moleküls in das sich akkumulierende Produkt zugrunde. Bei dem TaqMan Prinzip wird ein doppelt gelabeltes Oligonukleotid, das zwischen den beiden PCR-Primern bindet, während der Strangsynthese hydrolysiert. Solange die Sonde intakt ist, befinden sich das Reporter-molekül und der sogenannte Quencher in unmittelbarer räumlicher Nähe. Wird das Reporter-molekül angeregt, wird die abgegebene Energie durch FRET (fluorescence resonance energy transfer) auf das Quencher-molekül übertragen. Aufgrund der sich überlagernden Emissions- und Absorptionsspektren der Moleküle ist keine Fluoreszenz messbar. Wird die Sonde hydrolysiert ist FRET nicht mehr möglich und die Fluoreszenz nimmt zu.

### 1.10 Fragestellung

In zahlreichen Studien an unterschiedlichen Tumorentitäten wurden hohe Überexpressionsraten für EGFR erhoben (29). Für Mammakarzinome konnte eine Überexpression von EGFR in 30-40% der Tumore nachgewiesen werden. Diese scheint auf Ereignissen zu beruhen, die in einem frühen Stadium der Tumorgenese auftreten. Die Überexpression von EGFR konnte ebenfalls in prä-malignen Vorstufen und dem umgebenden morphologisch normalen Gewebe nachgewiesen werden. Es konnte gezeigt werden, dass die Expression des *EGFR* Gens über die Länge eines „CA dinucleotide repeats“ im Intron 1 beeinflusst wird (64). Für 33% der Mammakarzinome und 75% des morphologisch normalen Drüsengewebes wurde dabei eine Ungleichverteilung der Allele (allelic Imbalance) in CA-SSR (simple sequence repeat) nachgewiesen. Durch quantitative PCR 5'-Nuklease-Assays (TaqMan) zeigte sich, dass die AIs auf Amplifikationen zurückzuführen waren, die zu einer Überexpression des Rezeptors führten (69). Die gleiche Amplifikation des CASSR im *EGFR*- Gen konnte in 30% des umgebenden morphologisch normalen Brustgewebes und 100% der Lymphknotenmetastasen nachgewiesen werden. Als Basismodell dieser Arbeit dient die Zelllinie MDA-MB-468, die den Phänotyp eines invasiven Mammakarzinoms besitzt. Die Zelllinie trägt im Bereich des Chromosoms 7p11-14 eine distinkte Amplifikation, die das *EGFR*-Gen einschließt.

Während dieser Arbeit sollten folgende Ziele verfolgt werden:

- Es sollten, basierend auf der Heterogenität der *EGFR* Genkopienzahl innerhalb der Population der Zelllinie MDA-MB-468, Sublinien mit unterschiedlichen durchschnittlichen Genkopienzahlen („Average Gene Copy Number“ – AGCN) erzeugt werden.
- Für eine stabile und reproduzierbare Quantifizierung der AGCN im diploiden Bereich sollte ein quantitatives Real-Time PCR (qPCR) Assay etabliert werden.
- Anhand von quantitativer „Reverse-Transkriptase-PCR“ (q-RT-PCR) der Sublinien mit unterschiedlicher AGCN für *EGFR* sollte der kausale Zusammenhang zur Expressionshöhe von *EGFR* untersucht werden.
- Ferner sollten Analysen zur Struktur des *EGFR* Amplikons auf genomischer Ebene durchgeführt werden. Diese sollten mit Bezug auf *EGFR*-Gendosis, Amplikonlänge und genomischer Charakteristika der Grenzregionen Aufschluss über mögliche Amplifikationsmechanismen geben.
- Dies sollte ebenfalls auf der Einzelzell-Ebene durchgeführt werden, um die Heterogenität von MDA-MB-468 Zellen für den *EGFR* Lokus zu untersuchen. Hierfür sollte eine Methode etabliert werden, mit der Einzelzellen isoliert und das ganze Genom linear amplifiziert werden konnten.

Die Untersuchungen sollten unter Verwendung der folgenden molekulargenetischen Methoden durchgeführt werden:

- Für eine gesicherte Quantifizierung der AGCN von *EGFR* sollten TaqMan Assays für das *EGFR*-Gen und stabile Referenzregionen etabliert werden.
- Die Expressionsanalysen sollten mit TaqMan RT-PCR Assays durchgeführt werden.
- Die genomische Charakterisierung der Struktur des Amplikons sollte durch *Fine Tiling Array-CGH* („Competitive Genomic Hybridisation“) durchgeführt werden.
- Die Resultate der *Fine Tiling Array-CGH* sollten anschließend durch qPCR, „Long-Range“ PCR und Sequenzierung überprüft werden.
- Zur genomischen Charakterisierung von Einzelzellen sollte eine Methode basierend auf Mikromanipulation und anschließender MDA („Multi Displacement Amplification)-WGA („Whole Genome Amplification“) etabliert werden.

- Für die WGA amplifizierte DNA sollten anschließend die molekulargenetischen Applikationen der Mikrosatelliten DNA, Sequenzierung, qPCR und *Fine Tiling Arrays* validiert werden.
- In den abschließenden experimentellen Arbeiten sollten die etablierten Assays an Patientenmaterial in Form von „Circulating Tumor Cells“ (CTCs) aus Blutproben und „Disseminated Tumor Cells“ (DTCs) aus dem Knochenmark von Patienten mit primärem, metastasierendem Mammakarzinom validiert werden.

## 2. Material und Methoden

### 2.1 Material

#### 2.1.1 Zelllinien

Tab.2.2: Herkunft der verwendeten humanen Zelllinien.

Zelllinie	Typ	Ursprungsgewebe	Literatur
MDA-MB-468	Adenokarzinom der Brust	Pleuraerguss	(124)
A431	Epidermoidkarzinom	Haut	(125)
HBL100	Adenokarzinom der Brust	Brustdrüse	(126)
MCF-7	Adenokarzinom der Brust	Pleuraerguss	(127)
SK-OV-3	Ovariales Adenokarzinom	Aszites	(128)
SKBR-3	Adenokarzinom der Brust	Pleuraerguss	(129)
BT-20	Adenokarzinom der Brust	Brustdrüse	(130)
G22	Gliom	NA	(131)

#### 2.1.2 Verbrauchsmaterialien

##### 2.1.2.1 Chemikalien

Die Lösungen wurden, wenn nicht anders erwähnt, mit demineralisiertem Wasser angesetzt. Sollten Lösungen für Arbeiten mit RNA verwendet werden, wurden sie mit Diethylenpyrocarbonat (DEPC, Sigma) behandeltem Wasser angesetzt.

AB-Serum	Biotest AG, Dreieich
Aceton	J. T. Baker, Deventer, Niederlande
Borsäure (H <sub>3</sub> BO <sub>3</sub> )	Sigma-Aldrich, Deisenhofen
DAKO Blocking Solution	DAKO, Hamburg
DAPI Fluorescent Nuclear Stain	Thermo Scientific, Bonn
Desoxynukleosidtriphosphat-Set	Roche, Mannheim
Diethylprocarbonate (DEPC)	Sigma-Aldrich, Deisenhofen
Dimethylsulfoxid (DMSO)	Serva, Heidelberg
D-MEM-Medium	Gibco BRL, Life Technologies, Eggenstein
EDTA (Diethylen-diamin-tetraessigsäure)	Sigma-Aldrich, Deisenhofen
Ethanol, Isopropanol	Merck Eurolab, Darmstadt
Fötale Kälberserum (FCS)	PAA Laboratories GmbH, Pasching, Österreich
Ficoll-Plaque	Amersham Pharmacia Biotech, Freiburg
Formaldehydlösung, mindestens 37%	Merck Eurolab, Darmstadt
Formamid, deionisiert	Merck Eurolab, Darmstadt
Gel Star Nucleic Acid Gel Stain	Lonza, Rockland, USA
L-Glutamin 200 mM	Gibco BRL, Life Technologies, Eggenstein
Hämalaun	Merck Eurolab, Darmstadt
HIDI Formamid	Applied Biosystems, Darmstadt
Kalimchlorid (KCl)	Merck Eurolab, Darmstadt
Kaliumhydrogenphosphat (K <sub>2</sub> HPO <sub>4</sub> )	Merck Eurolab, Darmstadt
Kaliumdihydrogenphosphat (K <sub>2</sub> H <sub>2</sub> PO <sub>4</sub> )	Merck Eurolab, Darmstadt
Methylenblau	Sigma-Aldrich, Deisenhofen
Natriumchlorid (NaCl)	Sigma-Aldrich, Deisenhofen
Natriumdihydrogenphosphatdihydrat	Merck Eurolab, Darmstadt

Natriumacetat	Merck Eurolab, Darmstadt
Neufuchsin	Merck Eurolab, Darmstadt
Nuklease-freies Wasser	Promega GmbH, Mannheim
Oktadecyltrichlorsilan	Fluka, Erlangen
Oktanfraktion	Fluka, Erlangen
Paraformaldehyd	Merck Eurolab, Darmstadt
Penicillin-Streptomycin	Gibco BRL, Life Technologies, Eggenstein
PBS, Dulbecco's	Gibco BRL, Life Technologies, Eggenstein
Protease Lösung	Qiagen, Hilden
Proteinase K-Lösung	Qiagen, Hilden
Seakem® LE Agarose	FMC-Bioproducts, Rockland, USA
Sodiumdodecylsulfat (SDS)	Amersham Pharmacia Biotech GmbH, Freiburg
Natriumcitrat-Dihydrat	Merck Eurolab, Darmstadt
Neodisher	Dr. Weigert, Hamburg
Tetramethylammoniumchlorid	MWG Biotech, Hamburg
Triton X 200	Sigma-Aldrich, Deisenhofen
Toluidinblau	Sigma-Aldrich, Deisenhofen
Trypsin-EDTA 10 x konzentriert	Gibco BRL, Life Technologies, Eggenstein
Trypsinlösung	ICN Biomedicals

### 2.1.2.2 Antikörper

**Tab.2.3: Antikörper zum Nachweis von Zytokeratin-positiven Zellen im KM oder Blut.**

Antikörper	Verdünnung	Hersteller
Anti-Pan CK A45-B/B3	1:100	Micromet, Martinsried
DAKO Z0259	1:20	Dako, Hamburg
A45/CY3	1:300	Micromet, Martinsried
CD45/Alexa488	1:150	Denovo Biolabels, Münster
CK 5/8 oyster red	1:500	Progen Scientific GmbH, Heidelberg
APAAP-Komplex	0,1 mg/ml	Dako, Hamburg

### 2.1.2.3 Oligonukleotide

Die aufgelisteten Oligonukleotide wurden von den Firmen Operon und TibMolBiol bezogen und wurden mit dem Reinheitsgrad HPSF (*High Purity Salt Free*) im 0,02 mmol-Maßstab synthetisiert. Einfach und doppelt Fluoreszenz-gelabelte Oligonukleotide wurden HPLC gereinigt und die zweifach markierten Oligonukleotide zusätzlich mit der Option 3'-*minor groove binder* synthetisiert.

**Tab.2.4: Verwendete Oligonukleotide für AI-Analysen.**

Bezeichnung *	Sequenz (5' → 3')	Annealing-Temperatur [°C]	Amplikon [Bp]	Literatur
D7S522_F	[6-FAM]GCAGGACATGAGATGACTGA	56	121	eigene
D7S522_R	GTTATGCCACTCCCTCACAC			
D8S258_F	[6-FAM]AGCTGCCAGGAATCAACTGAGAG	56	229	eigene
D8S258_R	GATGCTCACATAAAGGAGGGAGG			
D16S400_F	[6-FAM]GGTTCACAATTGGACAGTAT	56	179	eigene
D16S400_R	GAACCCTCCATGCTGACATT			
NEFL_F	[6-FAM]CCAATACCTGCAGTAGTGCC	56	109	eigene
NEFL_R	GAGCTGCTTAACACATAGGG			
D13S153_F	[6-FAM]AGGGTTATGTATAACCGACTCC	56	194	eigene
D13S153_R	GTCTAAGCCCTCGAGTTGTGG			

Bezeichnung *	Sequenz (5' → 3')	Annealing-Temperatur [°C]	Amplikon [Bp]	Literatur
D17S855_F	[6-FAM]GGATGGCCTTTTAGAAAAGTGG	56	151	eigene
D17S855_R	ACACAGACTTGTCTACTGCC			
D10S541_F	[6-FAM]CACCACAGACATCTCACAACC	56	161	eigene
D10S541_R	CCAGTGAATAGTTCAGGGATGG			
D16S402_F	[6-FAM]GTACCCATGTACCCCAATA	56	112	eigene
D16S402_R	CAAAGCACCACATAGACTAA			
D16S422_F	[JOEE]GAGAGGAAGGTGGAAAATACA	56	122	eigene
D16S422_R	GTTTAGCAGAATGAGAATAT			
P53CA_F	[6-FAM]AAGAAATCCCCTGCCACTC	56	171	eigene
P53CA_R	GCTGTAGAGTGAAGCTCAGGCT			
CASSR1_F	[6-FAM]GTTTGAAGAATTTGAGCCAACC	56	121	eigene
CASSR1_R	TTCTGTCTGCACACTTGGCAC			
CASSR3_F	[6-FAM]AGGCCACAGAGGAGATAACAGA	62	126	eigene
CASSR3_R	CAGGTGTGGTAGATGCCAAAGAA			
CASSR4_F	[6-FAM]GCAACTTATCCAAACCTGACC	62	197	eigene
CASSR4_R	AGAGTGGACTAGGAAATGCTAGGAG			
D7S2429_F	[6-FAM]CAACTGCCACACACATCTTTC	62	218	eigene
D7S2429_R	TGGGAGCTGGGAGTCAAGTG			
D7S494_F	[6-FAM]AGCTATGACCACCACTGAACTCAAG	62	116	eigene
D7S494_R	TGAGTCTTTGCAAAACATGCCTG			
D7S499_F	[6-FAM]GCAGGCTCAGTAAGTGGTTGC	62	216	eigene
D7S499_R	CATGAGTGTTCCTGCTGTTTCCCT			
D7S2467_F	[6-FAM]AAGGAGGAATTACTTGGCTGTGC	62	178	eigene
D7S2467_R	GTGGTAACGGTCATCTGTGTTCCG			
CASSR2_F	[6-FAM]CTCGAGGTCTCATCCTCTTCC	62	169	eigene
CASSR2_R	GCAGAGGTGCACAAAGGAGTAATG			
D7S2550_F	[6-FAM]TTCATTCAGCTCTCCTCGTCTCAC	62	248	eigene
D7S2550_R	CTAAGTTCATTTGTCTCGGTTCCA			
CASSR6_F	[6-FAM]AGTTCCTGACTGGGAATTCGAT	62	158	eigene
CASSR6_R	TTGGCCAAATTACACACCTTTG			

\* F: *forward* Primer (*sense*); R: *reverse* Primer (*antisense*); [6-FAM] Fluoreszenzfarbstoff 6-FAM-phosphoramidit

Tab.2.5: Verwendete Oligonukleotide zur DNA-Sequenzierung der Exone 5-9 von TP53.

Bezeichnung *	Sequenz (5' → 3')	Annealing-Temperatur [°C]	Amplikon [Bp]	Literatur
p53_5/6_F	CACTTGTGCCCTGACTTTC	58	477	eigene
p53_5/6_R	ACTGACAACCACCCTTAACC			
p53_7_F	CTGCTTGCCACAGGTCTC	58	267	eigene
p53_7_R	TGGAAGAAATCGGTAAGAGG			
p53_8/9_2F	TTAGGCTCCAGAAAGGACAAG	58	560	eigene
p53_8/9_2R	CAGGAGCCATTGTCTTTGAG			

\* F: *forward* Primer (*sense*); R: *reverse* Primer (*antisense*); 2F/2R zweites Primer Set; 5/6, 7, 8/9 bezeichnen die zu amplifizierenden Exone

**Tab.2.6: Verwendete Oligonukleotide zur quantitativen Analyse von DNA Genkopienzahlen:  
TaqMan Assay basierte Quantifizierung von *CHR2*, *CHR10* & *EGFR***

Bezeichnung *	Sequenz (5' → 3')**	Annealing-Temperatur [°C]	Amplikon [Bp]	Literatur
QEGFR_F	TCTGCATTCTGCCGAGTTC	58	94	eigene
QEGFR_R	GCAGTCTCCACTCCATGCTCA			
QEGFR_CA+1.5k	P-CAGCCCTCTGTTGGGTCACCTTCC-P	58	115	eigene
QCHR2_F	TGCTGCAGGAAGCTTGAACAC			
QCHR2_R	CACCATCCAGGATCTGTCTGC			
QCHR2FAM	P-CACACTCTTCACACTCACCTTCCATT-P	58	126	eigene
QCHR10_F	CAGAGATGCGGTGTTTCAGGAC			
QCHR10_R	CTGGCCTCATCCTTGGTATCG			
QCHR10FAM	P-AGGCACCTTTCGAAGGAGAACCACCA-P			

\* F: *forward* Primer (*sense*); R: *reverse* Primer (*antisense*) \*\* P: Doppelt gelabeltes Oligonukleotid, 5'-6-FAM-phosphoramidit und 3'-Black-Hole-Quencher-1a gegen 6-FAM-phosphoramidit

**Tab.2.7: Verwendete Oligonukleotide zur quantitativen Analyse von DNA Genkopienzahlen:  
*SYBR green* Assay basierte Quantifizierung von *Line1* & *EGFR***

Bezeichnung *	Sequenz (5' → 3')	Annealing-Temperatur [°C]	Amplikon [Bp]	Literatur
Line-1_F	AAAGCCGCTCAACTACATGG	58	149	(155)
Line-1_R	TGCTTTGAATGCGTCCCAGAG			
qEGFR3-F	TACCTGGACCTTGAGGGATTG	58	115	eigene
qEGFR3-R	ACTGCGTGAGCTTGTACTCG			
qEGFR4-F	CCATGACTGCAATCGTCTACC	58	62	eigene
qEGFR4-R	TTGAACCCTAATGCACACGAG			
qEGFR4b-F	GGTCAAAGGCTAACGTGCAG	58	119	eigene
qEGFR4-R	TTGAACCCTAATGCACACGAG			
qEGFR5-F	TACATCATGCGACCATTCCAG	58	147	eigene
qEGFR5-R	TGGGCTTTGAGCAATGATTAG			
qEGFR7-F	CACCACTCACTGAGACCTTGG	58	52	eigene
qEGFR7-R	GAAATGGAGGCATGGTAGTCC			
qEGFR9-F	CACCGTCATCACCTTCCTTTC	58	107	eigene
qEGFR9-R	TCCTCCATCTCATAGCTGTCTG			
qEGFR14-F	CCAAGGGAGTTTGTGGAGAAC	58	122	eigene
qEGFR14-R	TGAAATGGACGTGGATAGCAG			
qEGFR15-F	CCTACGGGTGAGTGGAAAGTG	58	107	eigene
qEGFR15-R	ACAAACCTCGGCAATTTGTTG			
qEGFR16-F	TGTCCAACGAATGGGTAAGTG	58	73	eigene
qEGFR16-R	CCACAGCAGTGTGGTCATTTC			
qEGFR17-F	TACATCCATCCGAGGAAATGG	58	61	eigene
qEGFR17-R	CAGCTTTGTGGACATTGAAGG			
qEGFR19-F	TCTAGACCCTGCTCATCTCCAC	58	104	eigene
qEGFR19-R	AGGCCAGTGCTGTCTCTAAGG			
qEGFR20-F	TCGCAAAGGTAATCAGGGAAG	58	57	eigene
qEGFR20-R	GGATCCTGGCTCCTTATCTCC			
qEGFR21-F	AGCCATAAGTCCTCGACGTG	58	124	eigene
qEGFR21-R	GAGAAGACCCTGCTGTGAGG			

\* F: *forward* Primer (*sense*); R: *reverse* Primer (*antisense*)

**Tab.2.8: Verwendete Oligonukleotide zur quantitativen Analyse von DNA Genkopienzahlen:  
SYBR green Assay basierte Validierung der Tiling Array Daten für MDA-MB-468 WT**

Bezeichnung *	Sequenz (5' → 3')	Annealing-Temperatur [°C]	Amplikon [Bp]	Literatur
WT-F01-F	CAGTCAGCACATTGACACTGG	60	108	eigene
WT-F01-R	AAGTGTGGACCAGCTCTGTCTC			
WT-F02-F	AGCCAGCTCTTTGATTCTTGG	60	109	eigene
WT-F02-R	CAAAGGAAACCTATTGCAAGC			
WT-F03-F	TCTGTAGCTCATGGATTACTTGG	60	209	eigene
WT-F-03-R	TACCCGTGGGCTGTGTACC			
WT-F04-F	GAATCCAGGTCAAACCATGTG	60	105	eigene
WT-F04-R	TACACAGATGCCTGCTCAAAG			
HERV-K1-F	CCAAAGAAGAGACTCTGAGAGGAC	60	59	eigene
HERV-K1-R	AGTCTTGACATGGCTGTTATTGAG			
HERV-K2-F	GAAATCCAAGCTGTATGTTCAATG	60	268	eigene
HERV-K2-R	ATTAGAAGTCAGCCTAATGCCATC			
WT-F05-F	GGCTCCCATGATTCTCCTG	60	219	eigene
WT-F05-R	GGGCAACACTGAGCATTAAAG			
N-WT-F01-F	AACCAGCCATATTCATTTCATC	60	212	eigene
N-WT-F01-R	GCCCAAAGTCAGGTCATATAACAG			
N-WT-F02-F	TCACCCAAGTCGTGTATGTCAG	60	251	eigene
N-WT-F02-R	CCACTCTCGAAATTTGTTTCAGC			
N-WT-F03-F	AAGAAAGGTGATGAGCATGGAC	60	50	eigene
N-WT-F-03-R	CCCTTAAATTTGAGAACAAGTGG			
N-WT-F04-F	GAAAGCATTTTCAGCCTTTTCAC	60	174	eigene
N-WT-F04-R	CACATGTTCAAATGAATTGATGC			
N-WT-F05-F	TTCTTTGCACATTAATTTGTTTG	60	111	eigene
N-WT-F05-R	AGAGGAATAGGCAGATCAATGG			
N-WT-F06-F	ACTGTTCTTGGGTAGATTGTTCTG	60	257	eigene
N-WT-F06-R	TGTTGGTCGTTGTATTTGTTTC			
N-WT-F07-F	TTTAGAAGATGGATTGGTGAGGAG	60	50	eigene
N-WT-F07-R	CCTGTTCTTGTTCCTGACCATTG			
WT-E01-F	CTCCAGAAGCCCTGACATTG	60	91	eigene
WT-E01-R	CTTCCTCCTATCCCTCACTGC			
WT-E02-F	CTGGGAGGACAAGGTAAGAGG	60	208	eigene
WT-E02-R	AATGGTGGAGTCACAGCTCAC			
N-WT-E01-F	GGGACCACATCTTACTCTCCAG	60	149	eigene
N-WT-E01-R	ACAGCACCTTCTCCTTACATC			
N-WT-E02-F	AGGAGAGTGCAGTGAGGGATAG	60	71	eigene
N-WT-E02-R	ACAGCACCTTCTCCTTACATC			
N-WT-E03-F	TACCAACAGTTCTCCCTCATCC	60	144	eigene
N-WT-E-03-R	AGTCTGCGGTGTAGAGGAAAATC			
N-WT-E04-F	ATATAGATTTCGGCAGCCTTCAG	60	120	eigene
N-WT-E04-R	TCCAATATCCAAGAGGCAGAAC			
N-WT-E04-R	TCCAATATCCAAGAGGCAGAAC			
N-WT-E05-F	TCTCAAAGATTGGTTTGTTTG	60	288	eigene
N-WT-E04-R	TCCAATATCCAAGAGGCAGAAC			

\* N-: Nachdesign; WT: MDA-MB-468 WT; F0x: telomerseitige Amplikongrenze; E0x centromerseitige Amplikongrenze; F: *forward* Primer (*sense*); R: *reverse* Primer (*antisense*)

**Tab.2.9: Verwendete Oligonukleotide zur quantitativen Analyse von DNA Genkopienzahlen:  
SYBR green Assay basierte Validierung der Tiling Array Daten für MDA-MB-468 0106**

Bezeichnung *	Sequenz (5' → 3')	Annealing-Temperatur [°C]	Amplikon [Bp]	Literatur
0106-F01-F	CGGGTCATTGCTTTACAGG	60	300	eigene
0106-F01-R	TGGGTTCTGAAGGATGGTTG			
0106-F02-F	ACATCCCTAGCACAGGCATC	60	60	eigene
0106-F02-R	CAGAAAGCAGCACCTCACAC			
0106-F03-F	ATTTGGGAAGAGCCACGAG	60	65	eigene
0106-F03-R	AATCCAGATTTGGGTAGGG			
WT-F01-F	CAGTCAGCACATTGACACTGG	60	108	eigene
WT-F01-R	AAGTGTGGACCAGCTCTGTCTC			
WT-F02-F	AGCCAGCTCTTTGATTCTTGG	60	109	eigene
WT-F02-R	CAAAGGAAACCTATTGCAAGC			
0106-E01-F	CTACGGGCAGAGGGTTTCAG	60	221	eigene
0106-E01-R	CACAGACAGTTCGTGCAAATC			
0106-E02-F	GCCACATCAGAGCCATAAGC	60	124	eigene
0106-E02-R	TTCCAAGAAGGTCTGTGAGGAC			
N0106-E1-F	CCTCCTATCCAATGGTCTCTCTC	60	55	eigene
N0106-E1-R	TTCTCCTCTGCAAGTAAAGTCAG			
N0106-E2-F	GTCTCAAAGTAAATGTGGGATGTG	60	107	eigene
N0106-E2-R	TCACCAAACCATATCATATTCAGC			
N0106-E3-F	TTTCCCACCTAAAGATAGGATTG	60	265	eigene
N0106-E3-R	TATTGAGTCATTGCTGTACCCATC			
N0106-E4-F	TGGCTGTCTGTGTATGTCTGTATG	60	224	eigene
N0106-E4-R	AGGGAGAGTATTCACACATGCTTC			

\* N-: Nachdesign; 0106: MDA-MB-468 0106; F0x: telomerseitige Amplikongrenze; E0x centromerseitige Amplikongrenze; F: *forward* Primer (*sense*); R: *reverse* Primer (*antisense*)

**Tab.2.10: Verwendete Oligonukleotide zur quantitativen Analyse von DNA Genkopienzahlen:  
SYBR green Assay basierte Validierung der Tiling Array Daten für MDA-MB-468 0107**

Bezeichnung *	Sequenz (5' → 3')	Annealing-Temperatur [°C]	Amplikon [Bp]	Literatur
0107-F01-F	CCATGATTTCGGGATATGGTC	60	132	eigene
0107-F01-R	GGACTCAAACATCTTTAGCCAAAC			
0107-F02-F	ATTGGTGAGGAGCAGACAATG	60	73	eigene
0107-F02-R	TGGTTTGACCTGGATTCACCTC			
0107-F03-F	CTGCATCCCAGCTACATCAG	60	78	eigene
0107-F03-R	ACACCCTGTAAAGCAATGACC			
0107-F04-F	GCTCACCAGCACTACCTTCTG	60	121	eigene
0107-F04-R	AGAAAGCAGCACCTCACACTC			
N-0107-F01-F	GCACTTTGTTTGACTTTGTAGGG	60	143	eigene
N-0107-F01-R	TTGAACTAAGAAACCAAGGAAGC			
0107-E01-F	CATGGGTATTGAGGCTCTTTG	60	160	eigene
0107-E01-R	ATTGTCCACCAGACAGCTGAG			
0107-E02-F	TACCTCCTCACAACGTCCATC	60	124	eigene
0107-E02-R	TGAAGCCAGAATACCACCAAG			
0107-E03-F	GTTTGAATGTGGTTTCGTTGG	60	115	eigene
0107-E03-R	CCTGTGAGCTGAAGAGTGAGG			

\* N-: Nachdesign; 0107: MDA-MB-468 0107; F0x: telomerseitige Amplikongrenze; E0x centromerseitige Amplikongrenze; F: *forward* Primer (*sense*); R: *reverse* Primer (*antisense*)

**Tab.2.11: Verwendete Oligonukleotide zur quantitativen Analyse von DNA Genkopienzahlen:  
SYBR green Assay basierte Validierung der Tiling Array Daten für MDA-MB-468 +**

Bezeichnung *	Sequenz (5' → 3')	Annealing-Temperatur [°C]	Amplikon [Bp]	Literatur
MDA+F01-F	TGGTTCATCAACTTGGTAGGG	60	160	eigene
MDA+F01-R	TTCGCAATAAGCATGTGAG			
MDA+F02-F	ACCAGTGTCAACCACAAATGC	60	58	eigene
MDA+F02-R	CTTAGCCTTCCCAGTGTTC			
MDA+F03-F	AAATTGGCAGCAACAGCAC	60	166	eigene
MDA+F03-R	TTCACTGCCTGAAACTGTGG			
MDA+F04-F	CCTGCCTGACTTACCCACTC	60	250	eigene
MDA+F04-R	TTATGGAAAGGAGCGTGGAG			
MDA+E01-F	CTGCGTAGGACCTTGCTTTC	60	65	eigene
MDA+E01-R	GACGTTGTGAGGAGGTAGGC			
MDA+E02-F	TCGTTGGAAGCAAATGTGTC	60	245	eigene
MDA+E02-R	CTTACCAGGCAGTCGCTCTC			
MDA+E03-F	CGTGTGGGTCAGGTACTTTG	60	284	eigene
MDA+E03-R	CTCATAACGAGGCTGCTTCC			
MDA+E04-F	TGCTCTTTGCTTCCATGTTG	60	209	eigene
MDA+E04-R	CATCTATCCACCACCCAACC			
N-MDA+E01-F	AATGTGGTTCGTTGGAAGC	60	122	eigene
N-MDA+E01-R	GAGCAAAGGTTCCCTGTGAG			
N-MDA+E02-F	GAATGGATAGAGGGACAGGTTG	60	59	eigene
N-MDA+E02-R	ACTTCAAAGGTGTGGGTTTCATC			
N-MDA+E03-F	GTTGTGTCACCACCAGATCAAG	60	124	eigene
N-MDA+E03-R	CAACCTGTCCCTCTATCCATTC			
N-MDA+E04-F	AAGGTGGTCTGGAGAAACAAAG	60	139	eigene
N-MDA+E04-R	GCGTAATCCCAAGGATGTTATG			
N-MDA+E05-F	CACATCTTATCACAGGGACCAG	60	99	eigene
N-MDA+E05-R	GTGAGCACCTGCATGTCTATTC			
N-MDA+E06-F	TACCATGCCTCCATTTCCCTTAC	60	100	eigene
N-MDA+E06-R	GTCCTAATGACGGACATCACAG			

\* N-: Nachdesign; +: MDA-MB-468 +; F0x: telomerseitige Amplikongrenze; E0x centromerseitige Amplikongrenze; F: forward Primer (*sense*); R: reverse Primer (*antisense*)

**Tab.2.12: Verwendete Oligonukleotide zur Charakterisierung der telomerwärts gerichteten Amplikongrenze von MDA-MB-468 WT**

Bezeichnung *	Primerkombination Nr.	Annealing-Temperatur [°C]	Amplikon [Bp]	Literatur
N-WT-F7-F	1	60	8071	eigene
HERV-K1-R				
N0107-F1-F	2	60	6426	eigene
HERV-K1-R				
N-WT-F3-F	3	60	4192	eigene
N-WT-F7-R				
WT-F3-F	4	60	2669	eigene
N-WT-F7-R				
WT-F5-F	5	60	2021	eigene
N-WT-F7-R				

Bezeichnung *	Primerkombination Nr.	Annealing-Temperatur [°C]	Amplikon [Bp]	Literatur
N-WT-F1-F	6	60	1119	eigene
WT-F3-R				
N-WT-F1-F	7	60	1397	eigene
N-WT-F6-R				
N-WT-F1-F	8	60	1777	eigene
WT-F5-R				
N-WT-F1-F	9	60	3579	eigene
N-WT-F7-R				
N-WT-F6-F	10	60	2439	eigene
N-WT-F7-R				
HERV-K1-F	11	60	2611	eigene
HERV-K2-R				
HERV-K2-F	12	60	3873	eigene
0106-F2-R				

\* N-: Nachdesign; +: MDA-MB-468 +; F0x: telomerseitige Amplikongrenze; E0x centromerseitige Amplikongrenze; F: *forward* Primer (*sense*); R: *reverse* Primer (*antisense*)

**Tab.2.13: Verwendete TaqMan® Gene Expression Assays für die Genexpressionsmessungen von *EGFR*, *CHCHD2* und *B2M* in MDA-MB-468 und deren Subklonen**

Assay ID	Zielgen*	Hersteller
Hs00193306_m1	<i>EGFR</i>	Applied Biosystems, Darmstadt
Hs00853326_g1	<i>CHCHD2</i>	Applied Biosystems, Darmstadt
Hs99999907_m1	<i>B2M</i>	Applied Biosystems, Darmstadt

\* *EGFR*: Epidermaler Wachstumsfaktor Rezeptor; *CHCHD2* coiled-coil-helix-coiled-coil-helix domain containing 2; *B2M*: Beta-2-Mikroglobulin

#### 2.1.2.4 Puffer und Lösungen

PBS, pH 7,4 (10x)	1,54 M NaCl, 0,08 M Na <sub>2</sub> HPO <sub>4</sub> x 2 H <sub>2</sub> O, 0,02 M KH <sub>2</sub> PO <sub>4</sub>
TBE (10x, pH 8)	0,89 M Tris, 0,89 M Borsäure, 0,02 M Na <sub>2</sub> -EDTA x 2 H <sub>2</sub> O
TE-Puffer (10 mM, pH 7,4)	10 mM Tris, 1 mM EDTA
EGF-Puffer	10 mM Essigsäure, 0,1% BSA

#### 2.1.2.5 Zellkultur- und Nährmedien

DMEM-Medium

90-95% (v/v)	DMEM Medium (Glukose 4,5g/l; 0,11g/l Natriumpyruvat)
X% (v/v)	FCS (Hitze inaktiviert),
0,5% (v/v)	Penicillin-Streptomycin Mix (Endkonzentration 50 U/ml)
2 mM	L-Glutamin
10%	FCS für die Linien: MCF-7, HBL100, SK-OV-3, MDA-MB-468
5%	FCS für die Linie: A431

Dem Suppressionsmedium für die Linie MDA-MB-468 wurde zusätzlich EGF (Endkonzentration 5 nm/ml) bei jeder Passage frisch zugesetzt. Das EGF (MW: ~6 kD ~ 6000 g/mol) wurde in EGF-Puffer mit einer Endkonzentration von 5 µM/ml aufgenommen, aliquotiert und bis zur Verwendung bei -20°C gelagert. Nach dem Auftauen ist die EGF-Lösung eine Woche bei 4°C haltbar und sollte nicht wieder eingefroren werden.

### 2.1.2.6 Verwendete käufliche Systeme (*Kits*)

Tab.2.14: Übersicht der verwendeten käuflichen Systeme.

Bezeichnung	Verwendung	Hersteller
BCIP/NBT, AP Konjugat Substrat Kit	Immunfärbung	BioRad Laboratories
BigDye® Terminator v. 1.1 Cycle Sequencing Ready Reaction Kit	Sequenzierung	Applied Biosystems, Darmstadt
Failsafe PCR Premix Selection Kit	PCR	Epicentre Technologies / Biozym Scientific GmbH, Oldendorf
Genomiphi kit V1 / V2	Ganz-genomische DNA Amplifikation	GE Healthcare, München
GeneScan-500-Rox size standard	LOH-PCR Fragmentanalyse	Applied Biosystems, Darmstadt
NucleoSEQ kit	Sequenzierungs-, WGA-Aufreinigung	Macherey –Nagel, Düren
QIAmp® DNA Micro Kit	DNA Isolation	Qiagen, Hilden
QIAquick® Gel Extraction Kit	DNA Isolation	Qiagen, Hilden
qPCR MasterMix Plus, No ROX	Real-Time quantitative PCR	Eurogentec, Köln
<i>QuantiTect SYBR green PCR Kit</i>	Real-Time quantitative PCR	Qiagen, Hilden
REPLI-g kit	Ganz-genomische DNA Amplifikation	Qiagen, Hilden
RNeasy® Micro Kit	RNA Isolation	Qiagen, Hilden
SuperScript First-Strand Synthesis	cDNA Konvertierung	Invitrogen, Karlsruhe

### 2.1.3 Geräte

Applied Biosystems 3130 Genetic Analyzer	Applied Biosystems, Darmstadt
Analysenwaage, Satorius BP6100	MS Laborgeräte, Heidelberg
Analysenwaage, Satorius CP2245	MS Laborgeräte, Heidelberg
Brutschrank, Hera150	Heraeus Kendro, Langselbold
CellTram vario / Kapillare Kappa MFKII	Eppendorf AG, Hamburg
Durchlichtmikroskop	Wilovert S Helmut Hund GmbH, Wetzlar

Durchlichtmikroskop mit Fluoreszenzkanal und Kamera	Leica Mikroskopie und Systeme GmbH, Wetzlar
Elektrophoreseapparaturen	Bio-Rad Laboratories GmbH, München
GeneAmp PCR System 9700	Applied Biosystems, Darmstadt
Geldokumentations-System, GeneGenius 2	Syngene, Cambridge, UK
Gelkammer, ComPor L Mini	Bioplastics RV, Landgraaf, Niederlande
Heizblock	Eppendorf AG, Hamburg
Kühlzentrifuge, 5415R	Eppendorf AG, Hamburg
Mastecycler Gradient	Eppendorf AG, Hamburg
Mastecycler EP-Gradient S – Realplex <sup>4</sup>	Eppendorf AG, Hamburg
Nanodrop ND-1000 Spectrometer	PeqLab, Erlangen
pH-Meter	inoLab WTW, Heidelberg
Pipettierhilfe	Hirschmann Laborgeräte, Eberstadt
Rollmischer, Stuart SRT1	Bibby Sterilin, Staffordshire, UK
Scanner, Epson 1680	LaserSoft Imaging AG, Kiel
Schüttler	Heidolph Instruments GmbH, Schwabach
Sterilwerkbank Herasafe HS12	Heraeus, Frankfurt
Thermocycler	Flexigene Techne, Staffordshire, UK
Tischzentrifuge, Heraeus B12	Kendro, Langselbold
TransferMan NK2, Mikromanipulator	Eppendorf AG, Hamburg
Wasserbad, GFL-1003	GmbH für Labortechnik, Burgwedel
Vortex-Genie 2	Scientific Industries
Zentrifuge, Heraeus 3S-R	Kendro, Langselbold
Zentrifuge, Rotofix 32	Hettich, Villingen-Schwenningen
Zytozentrifuge	Hettich, Villingen-Schwenningen

## 2.2 Methoden

### 2.2.1 Aufarbeitung und Nachweis von disseminierten Tumorzellen in Knochenmarkspiraten / Blut

#### 2.2.1.1 Isolierung disseminierter Tumorzellen aus dem Knochenmark

Die Trennung von mononuklearen Blutzellen und epithelialen Zellen von Erythrozyten und Granulozyten wurde mit Hilfe der Ficoll-Dichtegradientenzentrifugation nach Braun et al. durchgeführt (131, 132). Das durch Punktion aus dem oberen Beckenkamm bzw. den Rippen entnommene Knochenmark wurde in Heparin enthaltenen Monovetten aspiriert. Die mononuklearen Zellen wurden durch eine Ficoll-Dichtegradientenzentrifugation (Dichte 1,077 g/mol) bei 800 x g für 30 min bei RT separiert. Die Interphase mit den mononuklearen Zellen wurde abgenommen und mit 50 ml PBS für 10 min bei 1.400 rpm gewaschen. Abhängig von dem Erythrozytenanteil wurde ggf. eine zusätzliche Erythrozyten-Lyse nach Angaben des Herstellers durchgeführt (Whole Blood Erythrocyte Lysing Kit, R&D Systems, Minneapolis). Von den mononuklearen Zellen wurden jeweils  $7 \cdot 10^5$  Zellen für 3 min bei 1.000 rpm auf einen Objektträger (Superfrost Plus, Fischer) zentrifugiert (Zytospins). Die Zytospins wurden über Nacht getrocknet und bis zur weiteren Verwendung bei  $-80^\circ\text{C}$  eingefroren.

### 2.2.1.2 Immunzytochemische Färbung zum Nachweis von Zytokeratin-positiven Zellen im Knochenmark / Blut

Das Vorhandensein von DTZ im KM wurde anhand der Auswertung von drei mit A45-B/B3 und drei mit MOPC21 gefärbten Zytospins mit je  $7 \cdot 10^5$  Zellen bestimmt. Patienten, die mindestens eine Zytokeratin-positive Zelle in ihrem KM-Aspirat aufwiesen, wurden als KM-positiv deklariert.

DTZ in der mononuklearen Fraktion der Knochenmarkaspirate wurden mit Hilfe einer immunzytochemischen Untersuchung nach der APAAP-Methode (Alkalische Phosphatase Anti-Alkalische Phosphatase) nachgewiesen (133). Als monoklonaler Primärantikörper wurde A45-B/B3 (IgG1, Micromet, München, Germany) eingesetzt, der sich spezifisch gegen ein gemeinsames Epitop verschiedener Zytokeratinkomponenten (u.a. Zytokeratin 8, 18 und 19 sowie Zytokeratin 8/19 Heterodimere) richtet (132, 134). Zytokeratine stellen als intrazelluläre Polypeptide einen essentiellen Bestandteil des Zytoskeletts aller epithelialer Zellen dar, werden jedoch nicht in mesenchymalen Zellen exprimiert. Daher sind sie als spezifische Marker für die Detektion von epithelialen Tumorzellen im mesenchymalen KM geeignet. Zur Kontrolle jeder immunzytochemischen Färbung wurden als Positivkontrollen Zytospins mit Zellen der Zytokeratin-positiven Mammakarzinomzelllinien MCF7 oder MDA-MB-468 verwendet. Als Isotyp-Kontrolle diente der MOPC21 Antikörper (Sigma).

Die mononuklearen Zellen auf den Zytospins wurden 5 min in Aceton fixiert. Unspezifische Bindungsstellen wurden durch Inkubation mit 10% AB-Serum in PBS für 20 min bei RT blockiert. Anschließend wurden die Zellen 45 min mit dem monoklonalen Primärantikörper A45-B/B3 (2  $\mu\text{g}/\text{ml}$ , Micromet) inkubiert. Nach Waschen in PBS erfolgte eine 30-minütige Inkubation mit dem sekundären Kaninchen-Anti-Maus Brückenantikörper Z 0259 (125  $\mu\text{g}/\text{ml}$ , Dako). Die in PBS gewaschenen Zytospins wurden nun für 30 min mit einem APAAP-Konjugat (1  $\mu\text{g}/\text{ml}$ , Dako) inkubiert, dem eine 20-minütige Färbung mit der Substratlösung folgte. Das APAAP-Konjugat baut sich aus zwei Molekülen alkalischer Phosphatase und einem gegen den Sekundärantikörper gerichteten Maus-Antikörper auf. Der Sekundärantikörper kann dadurch den Primärantikörper mit dem APAAP-Konjugat verbinden. Die alkalische Phosphatase hydrolysiert Naphtholphosphatester (Substrat) zu Phenolkomponenten und Phosphaten. Die Phenole reagieren mit Diazoniumsalzen (hier Neufuchsin) und bilden unlösliche Azofarbstoffe. Endogene alkalische Phosphatasen des Knochengewebes und einiger Leukozyten wurden durch den Zusatz von Levamisol gehemmt. Nach dem Waschen der Zytospins in PBS folgte eine Hämatoxylin-Gegenfärbung in Mayer's Hämalaun Lösung (1:5) und das Eindeckeln der Objektträger mit Glyzergel Mounting Medium. Das automatische Scannen und die Auswertung erfolgte mit dem *Automated Cellular Imaging System* (ACIS; Clariant, Kalifornien, USA) (135).

Substratlösung (frisch angesetzt)

- |          |  |
|----------|--|
| Lösung 1 | 78,8 ml Aqua dest., 4,2 ml 1 M Tris (pH 9,5), 200 µl 0,625 M Levamisol |
| Lösung 2 | 415 µl 0,56 M NaNO <sub>2</sub> , 166 µl 5% Neufuchsin                 |
| Lösung 3 | 42 mg Naphthol-AS-Bi-Phosphat, 500 µl Dimethylformamid                 |

Lösung 2 wird mit Lösung 1 vermischt, Lösung 3 wird hinzugeben und filtriert.

Alternativ wurde ebenfalls eine Färbung über BCIP/NBT durchgeführt. Das Protokoll wurde wie oben beschrieben durchgeführt. Es wurden jedoch unfixierte mononukleare Zellen verwendet und BCIP/NBT als chromogenes Substrat. Hierfür wurden die BCIP und NBT Lösungen nach Herstellerangaben 1:100 mit dem Farbwicklerpuffer gemischt. Anschließend wurde Levamisol mit einer Endkonzentration von 0,2 mM, sowie ein Tropfen Inhibitor pro ml Substrat hinzugegeben. Nach einer Filtration durch einen 0,2 µm Filter wurden 200 – 500 µl der Lösung auf die Objektträger gegeben. Nach 10 min. Inkubation bei RT wurden die Zellen dreimal für drei min in PBS gewaschen und bis zur Verwendung in PBS aufbewahrt.

### 2.2.1.3 Immunomagnetische Anreicherung von DTCs aus Patientenblut

Für die Anreicherung von DTCs aus Patientenblut wurde das automatisierte *CellsSearch System* (Veridex, Warren, NJ) verwendet. Das System besteht aus drei Komponenten: dem Veridex System, dem *MagNest cell presentation device* und dem *CellSpotter Analyzer*. Das Veridex System dient der Anreicherung der DTCs im Blut. Das System reichert über Ferrofluid Partikel, die an anti-EpCAM Antikörper konjugiert sind (137) immunomagnetisch epitheliale Zellen aus dem Blut an. Diese werden zusätzlich mit zwei Fluoreszenz-markierten Antikörpern, einem Phycoeritrin-konjugierten anti-Cytokeratin Antikörper (CK 8, 18 & 19), sowie einem Allophycocyanin-konjugierten anti-CD45 Antikörper für weiße Blutzellen gefärbt. Die Zellen werden zudem DAPI gefärbt. Nach der Anreicherung der Zellen aus 7,5ml Blut bleibt ein Probenvolumen von ~300µl, das die angereicherten DTCs enthält. Diese werden dann bei RT in einer *MagNest Cartridge* im Dunkeln durch ein starkes Magnetfeld nochmals aufgetrennt und können anschließend mit dem *CellSpotter Analyzer* durch die Fluoreszenz-Markierung identifiziert werden (138).

### 2.2.1.4 Fluoreszenzfärbung von Zytospins

Zur Detektion der DTCs in Zytospins aus aufgearbeitetem KM oder Blut wurden parallel weitere Fluoreszenz-markierte Antikörper verwendet. Hierfür wurden die mononuklearen Zellen für 15 min bei RT mit einer 4 %igen PFA- bzw. 3,5 %igen Formalin-Lösung blockiert und anschließend einmal mit PBS gewaschen. Nach 5 min Permeabilisierung mit 0,2 % Triton X (DAKO) wurden die Zellen dreimal für 2 min in PBS gewaschen und 20 min mit DAKO Blocking Solution (DAKO) blockiert. Für die bei RT durchgeführte Inkubation mit den primär markierten

Antikörpern wurden jeweils einer der gegen Zytokeratine gerichteten Antikörper A45/CY3 (1:300) bzw. CK5/8-oyster-red (1:500) zusammen mit CD45/Alexa488 (1:150) verwendet, welcher gegen das Leukozyten spezifische Antigen CD45 gerichtet ist. Nach einmaligem Waschen in PBS wurde eine Kernfärbung mit DAPI für 2 min bei RT durchgeführt. Nach zweimaligem waschen mit PBS wurden die Zellen in PBS eingedeckelt und bis zur Verwendung dunkel bei RT aufbewahrt.

### 2.2.2 Mikromanipulation einzelner DTZs / Zelllinienzellen

Zur genaueren genomischen Charakterisierung einzelner DTZs, CTCs oder Zelllinien-Zellen wurde ein Protokoll erarbeitet, welches durch eine Kombination aus Mikromanipulation und gesamt-genomischer Amplifikation (Whole Genome Amplifikation - WGA) ausreichend qualitativ hochwertige DNA erzeugt, um daran weitere Applikationen wie Mikrosatelliten-Analysen, Sequenzierungen, quantitative Real-Time PCR und *Fine Tiling CGH Arrays* durchzuführen. Der Transfer der einzelnen Zellen wurde aufgrund der ganzheitlichen visuellen Kontrolle durch Mikromanipulation durchgeführt. Es wurde das inverse Fluoreszenzmikroskop Axiovert 200 (Zeiss) in Kombination mit dem Mikromanipulator CellTram Vario (Eppendorf) und dessen Steuereinheit TranferMan NK2 (Eppendorf) verwendet. Ausgestattet wurde der Mikromanipulator mit der speziell entwickelten Kapillare Kappa MFKII (siehe Abb. 2.1). Die Kapillare verfügt über eine speziell angeschliffene Spitze, die sowohl das Transferieren von Zellen in Lösung, wie auch das Aufnehmen adhärenter Zellen von Oberflächen ermöglicht. Es stehen zwei Innendurchmesser von 30 und 60µm zur Verfügung um eine bessere Adaption an die Zellgröße zu erlangen.

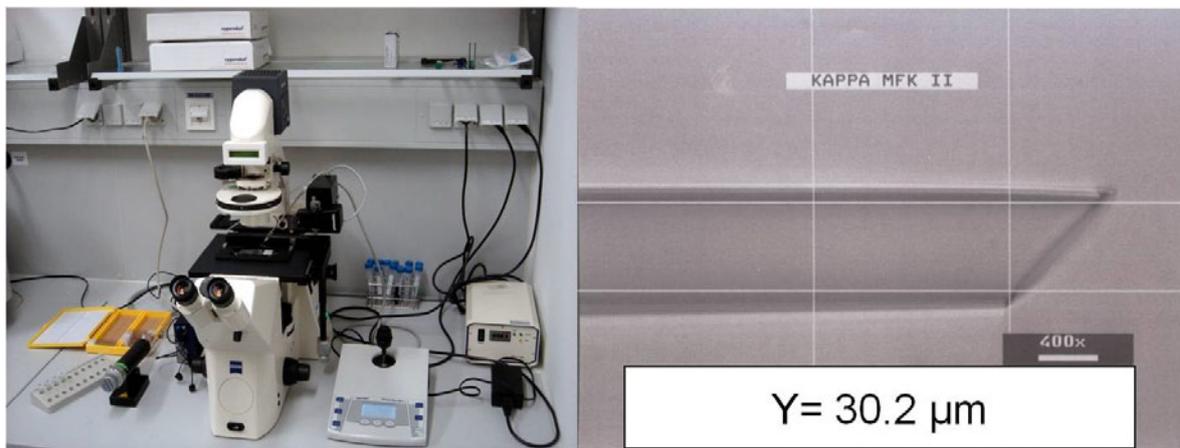


Abb. 2.1: Fluoreszenzmikroskop mit Mikromanipulator, Kapillare Kappa MFKII (400x Vergrößerung)

### 2.2.2.1 Transfer der Zellen in PCR-Reaktionsgefäße

Um einen Verlust der zu transferierenden Zelle oder Teile des Kerns auszuschließen, wurden zum Überführen spezielle, aus 0,19-0,23 µm starkem Glas geschnittene Stäbchen (Marienfeld, auf Anfrage) als Zellträger verwendet. Diese Stäbchen wurden vor ihrer Verwendung mit einem hydrophoben Silan beschichtet, um eine irreversible Bindung der bei der Zellyse freigesetzten DNA an die Glasoberfläche zu verhindern.

### 2.2.2.2 Silanisierung der Glas-Stäbchen

Die Stäbchen wurden in einer 1:1 Mischung aus NeoDisher (Dr.Weigert) und HPLC sauberen H<sub>2</sub>O in stark alkalischen Niveau 30 min bei 80-90°C geschwenkt, getrocknet und für 1 h in konzentrierter H<sub>2</sub>SO<sub>4</sub> geschwenkt, um Verunreinigungen und Fettrückstände von der Oberfläche zu entfernen. Nach rückstandslosem Entfernen der H<sub>2</sub>SO<sub>4</sub> durch HPLC sauberes H<sub>2</sub>O wurden die Stäbchen 15 min bei 115°C im Heisschrank getrocknet. Die Silanisierung erfolgte durch schwenken in einer 0,5% Lösung Octadecyltrichlorsilan in Oktanfraktion (Fluka) für 2 h bei RT. Während dieser Phase band das Silan kovalent an die Glasoberfläche. Die nach außen gerichtete C<sub>18</sub>-Kette erzeugte eine hydrophobe Oberfläche und verhinderte die DNA-Bindung. Ungebundenes Silan wurde durch schwenken in EtOH >99,9% für eine Stunde bei RT entfernt. Abschließend wurde das EtOH mit einem scharfen Strahl HPLC sauberen H<sub>2</sub>O aus einer Spritzflasche von jedem Stäbchen einzeln entfernt. Nach Trocknung bei RT konnten die Stäbchen trocken, dunkel für mindestens 8 Wochen gelagert werden.

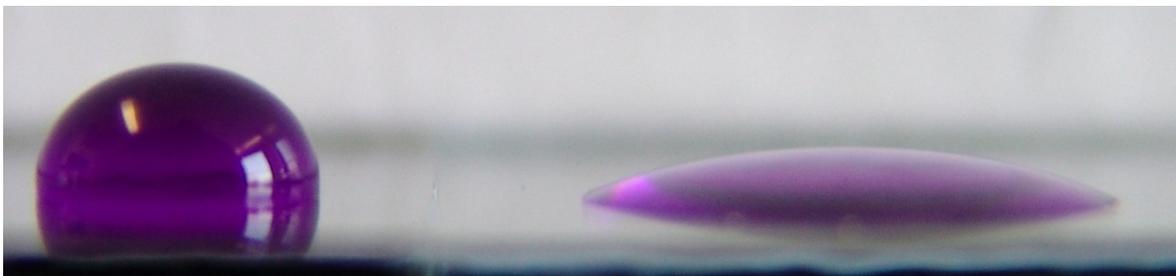
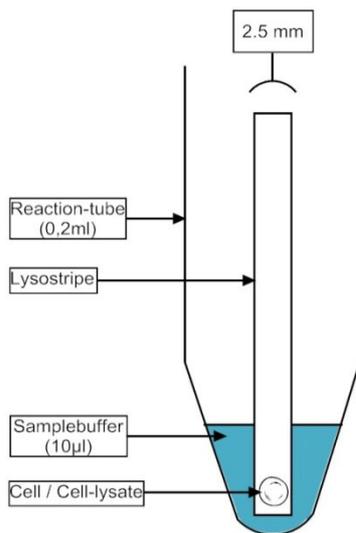


Abb. 2.2: 50 µl eingefärbtes H<sub>2</sub>O auf silanisiertem (links) und unsilanisiertem (rechts) Objektträger zeigen deutlich den Effekt der Oberflächenversiegelung

### 2.2.2.3 Zellyse von transferierten Zellen

Der Lysepuffer (LP) bestehend aus 0.25 % Sarcosyl, 2 M GITC (Guanidinthiocyanat), 0.01 M Natriumzitat (pH 7.0), 1 % DMSO (Dimethylsulfoxid) und Nuklease freiem H<sub>2</sub>O enthielt Detergenzien und Schwefelbrücken spaltende Chemikalien. Der Puffer wurde aliquotiert bei -20°C aufbewahrt. Jedem Aliquot wurde bei Verwendung frisch 60 mM DTT (Dithiotreitol) zugesetzt, da dieses unter O<sub>2</sub> sehr schnell oxidiert (abgewandelt nach 139). Um eine effiziente Zellyse ohne Inhibition der anschließenden WGA zu erreichen, wurden 20 nl Lysepuffer auf ein Ende des Stäbchens gegeben (0,2 µl 1:10 verdünnter LP) und bei RT eingetrocknet. Die Silanbeschichtung bewirkte ein Ankonzentrieren der Salze des Lysepuffers auf eng begrenzten Raum (Lysospot).

Auf dem Lysospot wurde die zu transferierende Zelle in einem möglichst kleinen Volumen Flüssigkeit abgesetzt. Nach Ablegen der Zelle wurde das Stäbchen in ein PCR-Reaktionsgefäß überführt, in das zuvor 9  $\mu\text{l}$  *Sample Buffer* des Genomiphi WGA Kits vorgelegt wurde. Der herausragende Teil des Stäbchens wurde durch das Schließen des Deckels abgebrochen und der Rest verblieb bis zum Ende der WGA im Reaktionsgefäß (siehe Abb. 2.3). Das Reaktionsgefäß wurde zur Effizienzsteigerung der Lyse >20 min bei  $-80\text{ }^{\circ}\text{C}$  eingefroren (140). Anschließend wurde dem Lysat 1  $\mu\text{l}$  1:10 verdünnte Protease (Qiagen) zugegeben und das Reaktionsgefäß 15 min bei  $50\text{ }^{\circ}\text{C}$  (Verdau) und 15 min bei  $70\text{ }^{\circ}\text{C}$  (Inaktivierung der Protease) inkubiert.



**Abb.2.3: Glasstäbchen mit Zellysate im PCR-Reaktionsgefäß**

#### 2.2.2.4 Whole Genome Amplifikation durch Multiple Displacement Amplification (MDA)

Die WGA des Genomiphi Kit basiert auf der isothermalen Amplifikation der DNA durch die strangverdrängende Eigenschaft der phi29-DNA Polymerase (141). Die fast ringförmig geschlossene Polymerase wurde mit 3'-Exonuklease geschützten Random Hexamer Primern kombiniert, die zwei Thiophosphat Modifikationen am Zuckerrückrad des 3'-Endes tragen. Das WGA Produkt ist doppelsträngig und hochmolekular (>10 kb). Die durchschnittlichen Produktmengen lagen bei Einzelzellen mit einem genomischen Gewicht von 6-7 pg bei 2,5-3  $\mu\text{g}$ . Aus 1 ng genomischer DNA wurden 5-7  $\mu\text{g}$  Produkt gebildet. Nach dem Verdau wurde die Probe, nach Herstellervorschrift, zwei min auf  $95\text{ }^{\circ}\text{C}$  erhitzt und sofort auf Eiswasser gestellt, um einzelsträngige DNA zu bilden. Anschließend wurden 9  $\mu\text{l}$  Genomiphi *Reaction Buffer* und 1  $\mu\text{l}$  *Enzyme Mix* hinzugegeben, vorsichtig gemischt, abzentrifugiert und anschließend 16 h (V1) oder 2,5 h (V2) bei  $30\text{ }^{\circ}\text{C}$  im Thermocycler inkubiert. Die Inaktivierung der Polymerase erfolgte für 15 min bei  $65\text{ }^{\circ}\text{C}$ .

### 2.2.2.5 Aufreinigung des WGA-Produktes

Zur Aufreinigung wurden NucleoSEQ-Säulen (Macherey – Nagel) verwendet. Diese auf dem Prinzip einer Gelfiltration basierenden Säulen separierten kleine DNA-Moleküle wie Random Hexamer Primer und nicht verbrauchte dNTPs von dem WGA-Produkt. Die Säulen wurden 2 h mit 600 µl Nuklease freiem H<sub>2</sub>O äquilibriert. Nach einer 1 min. Zentrifugation bei 3500 rpm wurden die Säulchen nochmals mit 300 µl Nuklease freiem H<sub>2</sub>O beladen und wiederum 1 min zentrifugiert. Anschließend wurden die Säulchen auf ein frisches Eppendorf Reaktionsgefäß gesetzt, mit den 20 µl WGA-Produkt beladen und 5 min. bei 3500 rpm zentrifugiert. Das gereinigte Produkt wurde anschließend mit 20 µl Nuklease freiem H<sub>2</sub>O verdünnt. Die Qualität und Quantität des WGA-Produktes wurden anschließend am Nanodrop Spektralphotometer bestimmt. Der schematische Arbeitsablauf der Einzelzell Analyse wurde in Abb. 2.4 zusammengefasst.

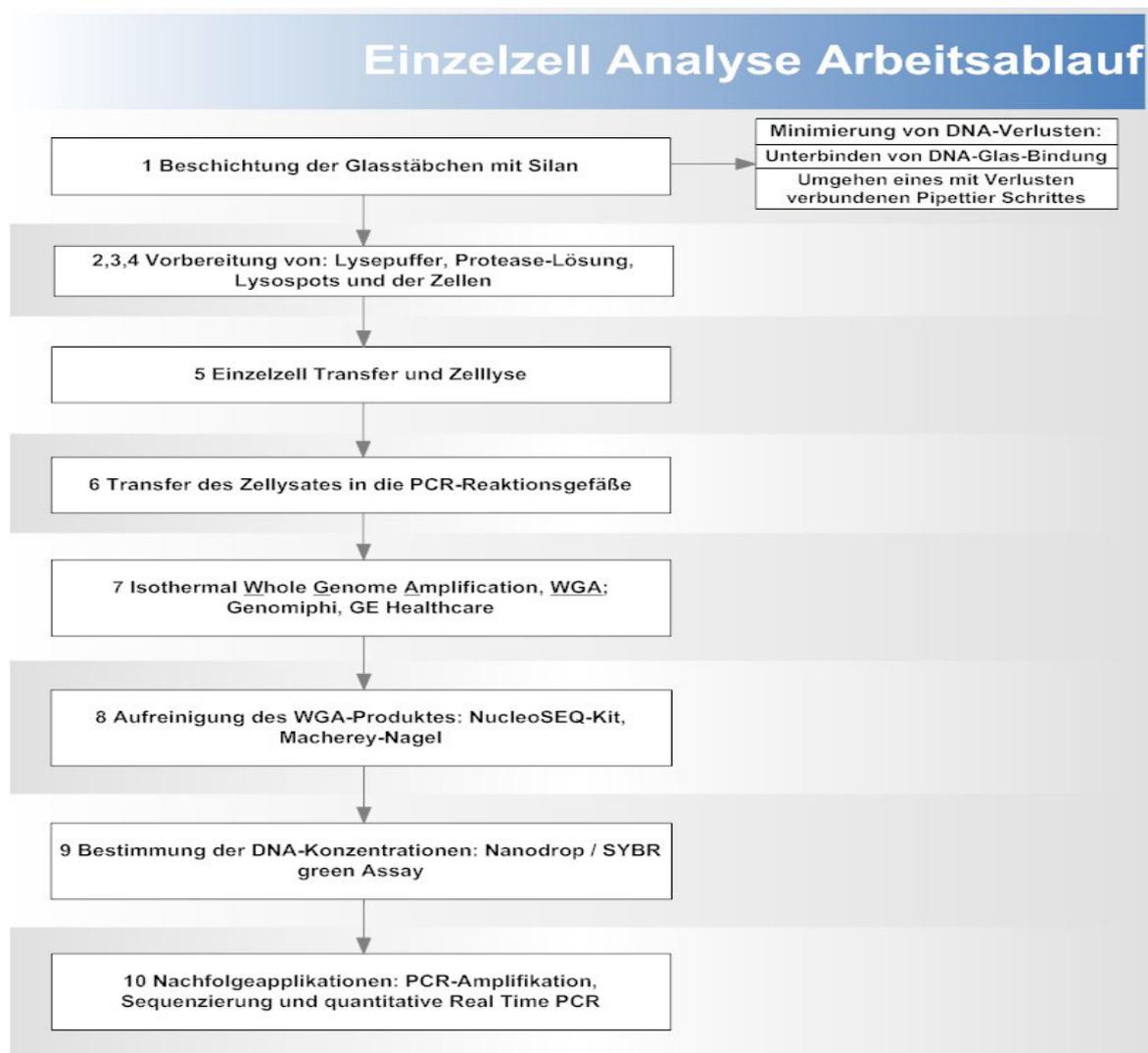


Abb. 2.4: Schematischer Arbeitsablauf der genomischen Analyse einer Einzelzelle

### 2.2.3 Isolierung von Nukleinsäuren

#### 2.2.3.1 Isolierung von genomischer DNA

Die DNA-Isolierung wurde unter Verwendung des QIAmp® DNA Micro (< 1 µg) oder QIAmp® DNA Mini Kits (Qiagen) nach Angaben des Herstellers ohne Verwendung von *carrier*-DNA oder Glycerol durchgeführt. Die Zellen wurden in 180 µl Lyse-Puffer und 20 µl Proteinase K (20 mg/ml) für 6-12 Stunden bei 56°C im Wasserbad inkubiert. Die weitere DNA-Isolation erfolgte nach dem Protokoll des Herstellers. Zur Elution der DNA, wurde zunächst 25 µl bzw. 60 µl DNase-freies Wasser auf die Säule gegeben, diese bei RT für 3 min inkubiert und anschließend bei maximaler Geschwindigkeit zentrifugiert.

Die Qualität und Quantität wurde durch Messungen am Nanodrop Spektralphotometer (PeqLab, Erlangen). Die Ausbeute an genomischer DNA betrug im Allgemeinen 10 µg für Zellen. Bis zur weiteren Verwendung wurde die genomische DNA bei -20°C gelagert.

#### 2.2.3.2 Isolierung von Gesamt-RNA

Für alle Lösungen wurde mit DEPC behandeltes Wasser verwendet und unter RNase-freien Bedingungen gearbeitet.

Die Isolierung von Gesamt-RNA erfolgte mit Hilfe des RNeasy® Micro bzw. des RNeasy® Mini Kits (Qiagen). Das Tumorgewebe bzw. die Zellen wurden in 250 µl Lyse-Puffer und 2,5 µl Mercaptoethanol homogenisiert. Zur Isolation der RNA aus Zellkulturen wurden die Zellen zunächst mit Hilfe des QIAshredders (Qiagen) aufgeschlossen. Die weitere Aufreinigung erfolgte nach Angaben des Herstellers. Da für die spätere Verwendung hochreine RNA ohne Kontaminationen mit DNA benötigt wurde, wurde ein DNA-Verdau für 15 min bei RT mit 30 U DNase (Qiagen) durchgeführt. Die Gesamt-RNA wurde mit 14 µl bzw. 60 µl RNase-freiem Wasser von der Säule eluiert. Die Qualität und Quantität der erhaltenen Gesamt-RNA wurde zum einen mit dem Nanodrop Spektralphotometer. 750ng Gesamt-RNA aus Zelllinien wurde zur Qualitätskontrolle auf ein 1%iges Agarosegel aufgetragen. Bis zur weiteren Verwendung wurde die Gesamt-RNA bei -80°C gelagert.

#### 2.2.3.3 Quantitäts- und Qualitätsbestimmung von Nukleinsäuren

##### 2.2.3.4 Nanodrop Spektralphotometer

Die Bestimmung der Konzentration wurde mit dem Nanodrop ND-1000-Spektrophotometer (PeqLab) durchgeführt. Für jede Messung wurde 1 µl der DNA- bzw. RNA-Lösung eingesetzt. Bei einer reinen DNA-Lösung sollte der Absorptionsquotient A<sub>260 nm</sub>/A<sub>280 nm</sub> einen Wert von 1,8 und für eine reine RNA-Lösung einen von 2,0 aufweisen. In dem Wellenlängenbereich von 220 nm bis 350 nm konnten zusätzlich die Absorptionsmessung beeinflussende Substanzen wie Alkohole und Verunreinigungen durch Proteine detektiert werden.

### **2.2.3.5 Agarose-Gelelektrophorese**

Agarose-Gelelektrophorese von Nukleinsäuren für analytische und präparative Zwecke erfolgte in horizontalen Elektrophoreseapparaturen. Je nach erwarteter Fragmentlänge wurden 0,8 – 2%ige (w/v) TBE-Agarosegele verwendet, denen 0,1 µg/ml Ethidiumbromid bzw. Gel-Star zugesetzt wurde. Die DNA-Proben wurden mit 0,2 Vol. DNA-Probenpuffer (50% Glycerol, 1 mM EDTA, 0,4% Bromphenolblau, 0,4% Xylencyanol) versetzt, auf die Gele aufgetragen und bei einer Spannung von 90 V aufgetrennt. Zur Abschätzung der Molekulargröße wurde ein Molekulargewichts-Marker (Roche bzw. Invitrogen) verwendet. Im Anschluss an die elektrophoretische Auftrennung wurde mit Hilfe eines Transilluminators (Anregungswellenlänge 234 nm) in die DNA eingelagertes Ethidiumbromid bzw. Gel-Star zur Fluoreszenz angeregt und die Gele zur Dokumentation fotografiert (Geldokumentationssystem Gene Genius 2 der Firma Syngene, Cambridge, UK). Bei Fragmentgrößen über 1000 Basen wurde grundsätzlich Gel-Star verwendet, da Ethidiumbromid durch seine Ladung das Laufverhalten der großen Fragmente im Gel verändert.

### **2.2.3.6 Aufreinigung von Nukleinsäuren**

#### **2.2.3.7 Ethanolfällung**

Zur Aufreinigung und ggf. Aufkonzentrierung der DNA wurde eine Alkoholfällung durchgeführt. Die DNA wurde mit 2,5 Vol. absolutem Ethanol und 0,1 Vol. 3 M Natriumacetat (pH 5,4) versetzt, so dass sie während einer Inkubation bei -20°C ausfiel. Nach Pelletierung durch 30 min Zentrifugation bei 14.000 x g wurde die DNA durch Waschen mit 70%igem Ethanol von Salz- und wasserlöslichen Substanzen gereinigt. Die getrocknete DNA wurde in TE (pH 7,4) bzw. Nuklease-freiem Wasser gelöst.

#### **2.2.3.8 Bandenextraktion aus Agarosegelen**

Um PCR-Produkte aufzureinigen und zu isolieren, wurden diese durch Gelelektrophorese in einem 0,8-2%igem (w/v) TBE-Agarosegele mit 0,1 µg/ml Ethidiumbromid aufgetrennt. Die gewünschten Banden wurden unter Benutzung von UV-Licht herausgeschnitten und in ein Reaktionsgefäß überführt. Die DNA wurde entsprechend der Herstellervorschrift mit dem QIAquick® Gel Extraction Kit (Qiagen) aus dem Gel isoliert, in TE eluiert und bei -20°C gelagert.

#### **2.2.3.9 Erststrang cDNA-Synthese**

Die Konvertierung von Gesamt-RNA in komplementäre DNA (cDNA) erfolgte mit dem SuperScript First-Strand Synthesis System (Invitrogen) nach Anleitung des Herstellers. Als Primer wurden Oligo(dT)<sub>12-18</sub>-Primer verwendet, welche mit den poly(A)-Schwänzen der mRNA-Moleküle hybridisieren können. Ein µg Gesamt-RNA wurde mit 1 µl dNTP-Mix (10 mM) und 1 µl Oligo(dT)<sub>12-18</sub> (0,5 µg/µl) in 8 µl H<sub>2</sub>O für 5 min bei 65°C denaturiert und anschließend für 3

min auf Eis gekühlt. Anschließend wurden 2 µl 10x Reaktionspuffer, 4 µl MgCl<sub>2</sub>-Lösung (25 mM), 2 µl DTT-Lösung (0,1 M) und 1 µl RNase Inhibitor zugegeben. Nach Inkubation für 2 min bei 42°C wurden 50 U SuperScript II Reverse Transkriptase hinzugefügt. Das Umschreiben der mRNA in cDNA erfolgte bei 42°C für 50 min. Die Reaktion wurde durch Erhitzen (70°C, 15 min) gestoppt, und die RNA durch Zugabe von 1 µl RNase H (37°C, 20 min) abgebaut. Für Expressionsanalysen mittels PCR wurden Gebrauchslösungen mit einer Verdünnung von 1:10 in Aqua dest. angesetzt. Die cDNA wurde bei -80°C bis zur weiteren Verwendung aufbewahrt.

### 2.2.3.10 Polymerase-Kettenreaktion

Die Polymerase-Kettenreaktion (PCR) dient zur Amplifizierung spezifischer Nukleinsäure-Fragmente. Zu analytischen Zwecken wurden die Amplifikationen unter Verwendung der AmpliTaq Gold DNA-Polymerase (Applied Biosystems) durchgeführt. Zur Bestimmung der optimalen Annealing-Temperatur für eigene Primer wurden zunächst Gradienten-PCRs durchgeführt. Das PCR-Additiv Tetramethylammoniumchlorid (TMAC) wurde eingesetzt um unspezifische Primerbindung zu minimieren.

#### Standard-Ansatz:

1	µl	10x PCR-Buffer	
1	µl	MgCl <sub>2</sub>	(3,5 mM)
1	µl	dNTP	(8 mM, Roche Diagnostics)
0,1	µl	Primer <i>forward</i>	(100 pmol/µl)
0,1	µl	Primer <i>reverse</i>	(100 pmol/µl)
0,1	µl	DNA-Polymerase	(5 U/ µl)
1	µl	TMAC	(200mM)
4	µl	DNA (5ng/µl)	(1 µl cDNA 1:20)
		add Aqua dest.	10 µl

Zyklen	Schritt	Temperatur [°C]	Zeit [min]
1x	Initiale Denaturierung	95	10:00
	Annealing	Variation A	0:30
	Extension	72	Variation B
35-40x	Denaturierung	95	0:30
	Annealing	Variation A	0:30
	Extension	72	Variation B
1x	Finale Amplifikation	72	7:00
	Kühlung	4	∞

Variation A: Annealing-Temperatur lag zwischen 56°C und 62°C und wurde abhängig vom AT-Gehalt der jeweiligen Primer variiert (s. Tabelle 2.4 und 2.5).

Variation B: Die Extensionszeit lag zwischen 0:30 min und 8:00 min und wurde je nach erwarteter Länge des PCR-Produktes variiert (s. Tabelle 2.5 und 2.12).

### 2.2.3.11 Aufarbeitung der Mikrosatelliten-PCR Produkte zur Fragmentanalyse

Nach Vorverdünnung mit 80 µl HPLC-H<sub>2</sub>O wurden jeweils 0,5 µl der PCR-Produkte mit jeweils 20 µl HIDI Formamid und 0,15 µl GeneScan-500-Rox size standard ( Applied Biosystems) in ein PCR-Reaktionsgefäß pipettiert. Nach der Denaturierung der DNA wurden die Produkte bei -20°C aufbewahrt bis zur Fragmentanalyse auf dem Applied Biosystems 3130 Genetic Analyzer (siehe Kapitel 2.2.2.3).

### 2.2.3.12 Auswertung der Fragmentanalysen der Mikrosatelliten-PCR

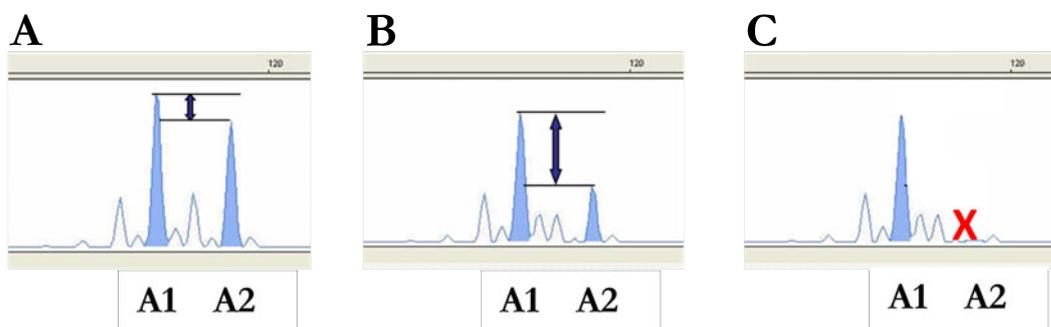
Die Genescans wurden mit der Software Genemapper V2.03 (Applied Biosystems) ausgewertet. Die Fläche oder Höhe eines jeden Allels wurde dabei markiert und zur Berechnung der Ratios eingesetzt wie im Folgenden beschrieben (64):

1. Standardabweichung (SD) = Peak-Höhen A1/A2; wird für jede Dreifachbestimmung einzeln berechnet. A1 entspricht dabei Allel1, A2 dementsprechend Allel2 (siehe Abb. 2.5)
2. Berechnung der kritischen Differenz:  $d_c = 2 * \sqrt{2} * SD$  ( $d_c = \text{critical difference}$ )
3. Auf der Basis der berechneten  $d_c$ -Werte wurden für jede Dreifachbestimmung bzw. jedes *Assay* die *Cutoff*-Werte zur Detektion eines AI wie folgt berechnet:

Verlust langes Allel:  $1 - d_c$

Verlust kurzes Allel:  $1 / (1 - d_c)$

Werte für Allel-Ratios, die kleiner als der berechnete Cutoff  $1 - d_c$  oder größer als  $1 / (1 - d_c)$  waren, konnten zweifelsfrei als AI eingestuft werden, da sie außerhalb des Assay-spezifischen Schwankungsbereiches lagen. Der Varianzkoeffizient, der als Maß der Präzision der Messung berechnet wird lautet:  $VK = (SD / \text{Mittelwert}) * 100$  Der Varianzkoeffizient wird in Prozent angegeben.



**Abb. 2.5: Fragmentanalysen mit unterschiedlichen Ratios für A1/A2.** Die blau hinterlegten Peaks wurden durch Fläche oder Höhe zur Berechnung verwendet. Zwischen A und B ist eine klare Verschiebung der Ratios zu erkennen, was einem Ungleichgewicht der Allele zueinander entspricht. Bei C fehlt A2, was einen vollkommenen Verlust des längeren Allels darstellt.

### 2.2.3.13 Quantitative Real-Time PCR

Die Real-Time-quantitative-PCR (RTQ-PCR) wurde im Mastercycler S Realplex (Eppendorf) entweder unter Verwendung des *QuantiTect SYBR green Master Mixes* (Qiagen) oder des qPCR MasterMix Plus, No ROX (Eurogentec) für TaqMan Probe Assays durchgeführt. In einer RTQ-PCR wird – anders als in einer konventionellen PCR - die Anzahl der gebildeten DNA Moleküle mit Hilfe von Fluoreszenzfarbstoffen am Ende von jedem Zyklus gemessen. Der im ersten Master Mix verwendete Farbstoff *SYBR green* fluoresziert nach dem Interkalieren in doppelsträngige DNA, so dass die Fluoreszenz proportional mit der Menge an PCR-Produkten zunimmt. Das Reportersystem für den zweiten Master Mix basiert auf einem zweifach markiertem Oligonukleotid (TaqMan Probe), welches am 5'-Ende das Fluorochrom FAM und am 3'-Ende das Quencher Molekül BHQ1 trägt. Das Oligonukleotid wurde so hergestellt, dass es zwischen den beiden dazugehörigen Primern bindet. Bleiben Reporter und Quencher in unmittelbarer räumlicher Nähe zueinander wurde über einen Fluoreszenz-Resonanz-Energie-Transfer (FRET) bei Anregung des FAM Moleküls (Reporter) die Energie auf den BHQ1 (Quencher) übertragen. Es war keine Fluoreszenz messbar. Dies war dadurch bedingt, dass die Extinktion des Reporter-Moleküls gleich der Absorption des Quencher-Moleküls war. Wurde ein PCR-Produkt gebildet, hydrolysierte die Taq Polymerase die TaqMan Probe vom 5'-Ende her, was zur Folge hatte, dass sich Reporter und Quencher nicht mehr in räumlicher Nähe zueinander befanden. Die Fluoreszenz des Reporter-Moleküls wurde dadurch messbar. Bevor ein Signal vor dem Hintergrundsignal detektiert werden konnte, wurde eine Mindestmenge von ungefähr  $10^{10}$ - $10^{11}$  Kopien an DNA benötigt. Dieser Schwellenwert wird als CT-Wert (*Threshold Cycle*) bezeichnet und beschreibt den Zyklus, an dem die Fluoreszenz erstmalig signifikant über die Hintergrund-Fluoreszenz ansteigt. Dieser Wert war gleichzeitig ein direkter Hinweis auf die vorliegende Konzentration der Ausgangs-DNA in der Probe (je geringer die Konzentration, desto später wurde der Schwellenwert erreicht und dementsprechend höher war der CT-Wert). Nach einem exponentiellen Anstieg des Signals näherte sich die Fluoreszenz einem Maximum an. Anschließend konnte für *SYBR green Assays* eine Schmelzkurvenanalyse durchgeführt werden, anhand derer sich indirekt über deren Schmelztemperatur die Fragmentlänge(n) und die Spezifität bestimmen ließ. Indem die Temperatur kontinuierlich von 60°C auf 90°C erhöht wurde, wurde die DNA aufgeschmolzen, so dass einzelsträngige DNA entstand. Der Farbstoff *SYBR green* wurde freigesetzt, welches als Fluoreszenzabnahme registriert wurde. Da unspezifische Primerdimere einen niedrigeren Schmelzpunkt besaßen als die doppelsträngige DNA von spezifischen PCR-Produkten, wurde eine Unterscheidung möglich. Die Höhe des Peaks der Schmelzkurve gab ebenfalls Auskunft über die Menge des gebildeten Fragments und somit über die Anfangskonzentration der DNA oder cDNA (142). Aufgrund der Sensitivität der Methode wurden alle Proben in Triplikaten gemessen, um abweichende Messwerte identifizieren zu können.

Ansatz unter Verwendung *QuantiTect SYBR green Master Mixes*

7,5 µl	2x Master Mix	
0,4 µl	Primer <i>forward</i>	(100 pmol/µl)
0,4 µl	Primer <i>reverse</i>	(100 pmol/µl)
2 µl	DNA (5 ng/µl)	
	add Aqua dest. 15 µl	

Zyklen	Schritt	Temperatur [°C]	Zeit [min]
1x	Initiale Denaturierung	95	15:00
45x	Denaturierung	95	0:15
	Annealing	58	0:30
	Extension & Fluoreszenzmessung	68	0:30
1x	Schmelzung	95	0:15
		60	0:15
		95	0:15
1x	Kühlung	RT	∞

Ansatz unter Verwendung des qPCR Master Mixes, No Rox zur Quantifizierung von DNA

7,5 µl	2x Master Mix	
3 µl	Primer <i>forward</i>	(100 pmol/µl)
3 µl	Primer <i>reverse</i>	(100 pmol/µl)
1,25 µl	TaqMan Sonde	(20 pmol/µl)
2 µl	DNA (5 ng/µl)	
	add Aqua dest. 15 µl	

Zyklen	Schritt	Temperatur [°C]	Zeit [min]
1x	Denaturierung	95	10:00
45x	Denaturierung	95	0:15
	Annealing	58	0:30
	Extension	68	0:30
1x	Kühlung	RT	∞

Ansatz unter Verwendung des qPCR Master Mixes, No Rox zur Quantifizierung von cDNA

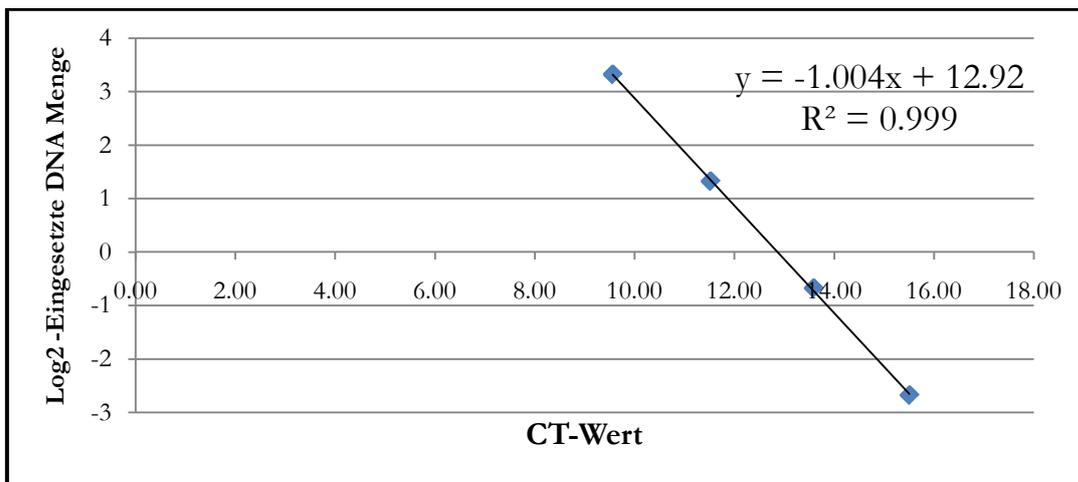
10 µl	2x Master Mix	
1 µl	TaqMan Gene Expression Assay (20x)	
2 µl	cDNA	
	add Aqua dest. 20 µl	

Zyklen	Schritt	Temperatur [°C]	Zeit [min]
1x	Denaturierung	95	10:00
40x	Denaturierung	95	0:15
	Annealing + Extension	60	1:00
1x	Kühlung	RT	∞

### 2.2.3.14 Auswertungsmethoden der quantitativen Real-Time PCR

Zur Auswertung der Daten wurden je nach Versuchsansatz zwei unterschiedliche Methoden für die absolute oder relative Quantifizierung verwendet. Grundsätzlich unterschieden sich die beiden Methoden durch den Einsatz einer Kalibrationsgerade (143) für die absolute Quantifizierung, die bei relativer Quantifizierung durch eine einzelne Referenzprobe ersetzt wurde, zu der die Proben jeweils relativ in Bezug gesetzt wurden.

Die Relative Quantifizierung (144,145) setzte voraus, dass die Effizienzen der zu vergleichenden Assays identisch waren. Die Effizienz der Assays wurde durch eine Kalibrationsreihe über die Steigung (*slope*) bestimmt (siehe Abb. 2.6):



**Abb. 2.6: Kalibrationsgerade für die Effizienzberechnung.** Aufgetragen wurden die Log2-Werte gegen die gemessenen CT-Mittelwerte der Dreifachbestimmungen. Der  $R^2$ -Wert ist ein unmittelbares Maß für die Qualität der Eichgeraden. Je näher der Wert an eins lag, umso höher war die Qualität der Messung und somit die Präzision der errechneten Werte.

Die Effizienz berechnete sich nach:  $Effizienz = 2^{\frac{-1}{slope}} - 1$  und ergab für die Steigung von -1.004 eine Effizienz von 0,994 bzw. 99,4%. Die Relative Quantifizierung des Zielgens ( $X_{test}$ ) gegen die endogene Referenz ( $X_{control}$ ) wurde wie folgt berechnet:

$$\frac{X_{test}}{X_{control}} = 2^{\Delta\Delta C_T} = 2^{(C_{T,X} - C_{T,R})_{control} - (C_{T,X} - C_{T,R})_{test}}$$

Der Wert  $C_{T,X}$  entsprach dabei dem CT Wert des Zielgens und  $C_{T,R}$  dem CT Wert der endogenen Kontrolle. Als Kontrolle für die Quantifizierung von DNA wurde jeweils Leukozyten DNA eingesetzt, für das WGA-Produkt äquivalentes Leukozyten-WGA-Produkt und für die Expressionsmessungen cDNA von *Universal Human Reference* mRNA. Da bei dieser Art der Berechnung die Kontrolle als arbiträre Konstante betrachtet wurde, floss diese nicht in die Fehlerberechnung ein. Die Fehlerberechnung wurde wie folgt durchgeführt:

$CV = \sqrt{CV_{test,X}^2 + CV_{test,R}^2}$  Der Wert CV (Koeffizient der Variation) entsprach dem Mittelwert der Standardabweichung der Dreifachbestimmungen.

Für die absolute Quantifizierung (144,145) wurden in jedem PCR-Lauf Kalibrationsgeraden für das Zielgen sowie die Referenzen gemessen. Für die Quantifizierung von DNA wurde für die Messung der Kalibrationsgerade 10 ng bis 15 pg DNA in einer seriellen 1:4 Verdünnung eingesetzt. Als Template für die Quantifizierung von DNA wurde jeweils Leukozyten-DNA, für WGA-Produkte entsprechendes Leukozyten-WGA-Produkt verwendet. Es war notwendig sicher zustellen, dass die gemessenen DNA Mengen innerhalb der Kalibrationsgerade lagen, da nur so ein linearer Verlauf der PCR-Reaktion voraus gesetzt werden konnte. Für die Kalibrationsgeraden wurde jeweils wie unter Kapitel 2.2.3.14 beschrieben, die lineare Geradengleichung ermittelt. Zur Überprüfung des PCR-Laufes wurden jeweils die Effizienzen und der R<sup>2</sup>-Wert überprüft. Die DNA Mengen der Proben wurden unter Verwendung der Geradengleichung wie folgt ermittelt:

$$\text{Proben DNA (ng)} = 2^{\left(\frac{C_{T,Probe} - \text{intercept}}{\text{slope}}\right)}$$

Hierbei entsprach der *intercept* dem y-Achsenabschnitt der Geradengleichung. Die Berechnung der relativen Verhältnisse (Normalisierung) zwischen Zielgen (DNA) und Referenzgen (DNA) erfolgte anschließend durch Quotientenbildung nach folgender Gleichung:

$$\text{Zielgen (Referenzgen}^{-1}) = \frac{\text{Ziel DNA (ng)}}{\text{Referenz DNA (ng)}}$$

Die Fehlerberechnung über die CV-Werte der Proben- und Referenzmesswerte der Dreifachbestimmungen wurde nach folgender Gleichung durchgeführt:

$$CV = \sqrt{CV_{\text{Ziel DNA}}^2 + CV_{\text{Referenz DNA}}^2}$$

### 2.2.3.15 Analyse der differentiellen PCR5-Fragmente mittels Restriktionsanalysen

Die beiden differentiell in MDA-MB-468 auftretende verkürzten PCR-Produkt wurden durch einen Restriktionsverdau mit den Enzymen *PvuII*, *NcoI* und *PstI* überprüft. Dazu wurden 250 ng PCR-Produkt mit je 10 U Enzym (Alle von New England BioLabs, Frankfurt am Main), 1 µl des entsprechenden Puffers und 1 µl Aqua dest. vermischt und 2 h bei 37°C inkubiert. Anschließend wurde das Enzym 20 min bei 65°C inaktiviert. Das Ergebnis des Verdau wurde mit einer Agarose-Gelelektrophorese kontrolliert (s. Kapitel 2.2.5).

### 2.2.3.16 DNA-Sequenzierung

Die hier durchgeführte Sequenzierung erfolgte nach der klassischen Didesoxy-Methode (117) unter Verwendung des BigDye® Terminator v. 1.1 Cycle Sequencing Ready Reaction Kits (Applied Biosystem, Darmstadt). Der vom Hersteller gelieferte BigDye Mix enthielt AmpliTaq-Polymerase, dNTPs und die vier unterschiedlich fluoreszenz-markierten Didesoxy-Analoga.

Für einen 10 µl-Ansatz wurden 200 ng gereinigte DNA in maximal 6,5 µl H<sub>2</sub>O, 2 µl BigDye Mix und 1,5 pmol Primer eingesetzt. Anschließend wurde folgende Sequenzierungsreaktion im Thermocycler durchgeführt:

Zyklen	Schritt	Temperatur [°C]	Zeit [min]
1x	Initiale Denaturierung	96	2:00
25x	Denaturierung	96	0:30
	Annealing	50	0:15
	Extension	60	1:00
1x	Kühlung	4	∞

Die Entfernung von Primern, Primerdimeren, nichtpolymerisierten dNTP's und ddNTP's erfolgte mittels Gelfiltration durch NucleoSeq Säulen (Macherey-Nagel), siehe Kapitel 2.2.2.4. Die Kapillarelektrophoretische Auftrennung der DNA-Fragmente erfolgte mit dem Applied Biosystems 3130 Genetic Analyzer unter Verwendung des Polymers POP7 und 36 cm langen Kapillaren (Applied Biosystems). Die Auswertung der Sequenzanalysen erfolgte mittels der Sequence Analysis Software V3.07 (Applied Biosystems), FinchTV V1.4 und dem Online Alignment Tool ClustalW. Die Referenzsequenzen wurden entweder aus der NCBI oder aus der UCSC Datenbank bezogen.

### 2.2.3.17 Analysen zu Veränderungen der DNA-Kopienanzahl mittels *Fine Tiling Array-CGH*

Die CGH ist eine molekularzytogenetische Methode, mit der sich chromosomale Aberrationen nachweisen lassen. Tumor- und Referenz-DNA werden dabei gleichzeitig auf normale Metaphasechromosomen (klassische CGH), BAC-Klone, cDNA oder Oligonukleotiden (Array-CGH) hybridisiert. Sowohl die Proben DNA, als auch Referenz DNA (bezogen von Nimblegen) wurden vor der Hybridisierung mit Fluoreszenzfarbstoffen markiert, um nach der Hybridisierung durch einlesen an einem Microarray Scanner quantifiziert werden zu können. Die Markierung der Proben DNA erfolgte mit CY3 (Kanal 532nm), die der Referenz DNA mit CY5 (Kanal 633nm). Bei der hier verwendeten *Fine Tiling Array-CGH* wurden nach unseren Spezifikationen erstellte Arrays verwendet. Pro Array standen ohne die Firmeneigenen Qualitätskontrollen 395.000 Oligonukleotide zur Verfügung, welche bei einer Länge von 50-75 Basen und einem *Tiling* (Überlappung) von 15 Basen (alle 15 Basen beginnt ein neues Oligonukleotid) einen maximalen abzubildenden Bereich von ~5,7 Megabasen ergeben. Das Design des Arrays wurde von der Firma Nimblegen durchgeführt. Die Oligonukleotide wurden auf eine einheitliche Hybridisierungstemperatur von 76°C eingestellt. Dabei sollten sie einzigartig für das menschliche Genom sein. Dies führte jedoch dazu, dass in einigen Bereichen mit hochkonservierten oder

repetitiven Sequenzmotiven kein uniques Probedesign möglich war, so das Lücken in der abzubildenden Sequenz auftraten. Basierend auf Daten von älteren Arrays mit BAC-Klonen umfasste unser Array Design einen Bereich von ~4,7 Megabasen auf Chromosom 7 im Bereich des *EGFR* Gens, sowie 177 Kilobasen auf Chromosom 2 und 196 Kilobasen auf Chromosom 10, den beiden Bereichen, in denen die ReferenzAssay unserer quantitativen Real-Time PCR liegen (Auswahl basierend auf 146). Dieses Array Layout erlaubt eine hochauflösende Abbildung der Sequenzbereiche, welche zuvor mit quantitativer Real-Time PCR untersucht wurden.

Als Proben wurden Leukozyten-DNA als Normalreferenz, drei Tumorzelllinien mit einer Amplifikation des *EGFR* Gens (A431, BT20, MDA-MB-468), die hemizygote Tumorzelllinie MCF-7 (trägt nur noch eine Kopie des *EGFR*-Gens), drei Subklone der Zelllinie MDA-MB-468 mit unterschiedlich hoher *EGFR*-Genkopienzahl, sowie WGA-Produkte dieser Linie aus 500 Zellen und sechs Einzelzell-WGAs verwendet. Von diesen Proben wurden jeweils 2,5 µg hochmolekulare DNA mit einer Konzentration von ~250 ng/µl benötigt. Die  $A_{260}/A_{280}$  Ratio sollte dabei >1,7 und die  $A_{260}/A_{230}$  Ratio >1,5 entsprechen.

#### **2.2.3.18 Hybridisierung der *Fine Tiling Array-CGH***

Die Probenvorbereitung, die Markierung, die Hybridisierung, sowie die Datenerfassung wurden von der Firma Nimblegen durchgeführt. Detaillierte Protokolle sind im Appendix hinterlegt.

#### **2.2.3.19 Auswertung der *Fine Tiling Array-CGH***

Die erhaltenen Datensätze wurden mit dem Algorithmus Quantsmooth unter Verwendung der Regionen von Chr. 2 und Chr. 10 normalisiert und die Signale geglättet. Die Normalisierung und Anpassung der Daten wurden am Institut für Bioinformatik der Universität Hamburg durchgeführt (Dr. H. Pospisil). Quantsmooth ist Teil eines R-basierten Softwarepackets (147). Anschließend wurde die Glättungsfunktion des Algorithmus verwendet um die maximalen Änderungen in den Anstiegen und Gefällen der Arrays zu detektieren. Diese wurden dann als Start- bzw. Endpunkte der jeweiligen Amplikone definiert.

### **2.3 Zellkulturbedingungen**

Alle Zelllinien wurden in sterilen Kulturflaschen der Firma Nunc (Wiesbaden) kultiviert. Die Kultivierung erfolgte in Hera150-Brutschränken (Kendro) bei 37°C in wassergesättigter Atmosphäre mit 5% CO<sub>2</sub>. Die Zelllinien wurden abhängig der Zellteilungsrate ein- bis zweimal wöchentlich unter sterilen Bedingungen bei einer Konfluenz von ~80% passagiert. Hierzu wurden die adhären wachsenden Zellen mit PBS bei 37°C gewaschen und anschließend mit Trypsin/EDTA-Lösung (0,05%/0,02%) von der Kulturschale abgelöst. Der Vorgang wurde durch Zugabe von frischem vorgewärmtem Vollmedium gestoppt und die resuspendierten Zellen in einer Dichte von 20-30% ausgesät.

Zur längeren Lagerung von eukaryontischen Zellen wurden die Zellen in Flüssigstickstoff bei -196°C eingefroren. Um die Bildung von Eiskristallen in den Zellen zu verhindern, wurden den Zellen durch Zugabe des stark hygroskopischen Dimethylsulfoxids (DMSO) langsam das Wasser entzogen.  $2 \cdot 10^6$  Zellen wurden in 1 ml Einfriermedium (10% (v/v) DMSO, 90% (v/v) Vollmedium) resuspendiert, in Kryotubes (Nunc, Wiesbaden) überführt und auf -80°C gekühlt. Nach 24 h wurden die Kryoröhrchen in flüssigem Stickstoff gelagert. Um die kryokonservierten Zellen wieder in Kultur zu nehmen, wurden die Zellen nach dem Auftauen (37°C im Wasserbad) mit Vollmedium gewaschen, um das toxische DMSO zu entfernen. Die im Vollmedium resuspendierten Zellen wurden in T75-Zellkulturflaschen ausgesät.

## 2.4 Verwendete Software, Onlinetools, Datenbanken

### Software / Onlinesoftware

- Realplex Software V2.02 (Eppendorf): Quantitative PCR Steuerungs- und Auswertungssoftware
- GeneMapper (Applied Biosystems) Auswertungssoftware für PCR-Fragmentanalysen
- Sequence Analysis Tool (Applied Biosystems) Auswertungssoftware für Sequenzierungen
- FinchTV ([www.geospiza.com/finchtv/](http://www.geospiza.com/finchtv/)) Freie Software zur Visualisierung von Sequenzierungsdateien (.asf).
- Perlprimer ([perlprimer.sourceforge.net](http://perlprimer.sourceforge.net)) Freie Software zum Primerdesign, enthält Blastfunktion, Sekundärstrukturvoraussage, Schmelzpunktberechnung uvm.
- Signalmap V1.9 (Nimblegen) Software zur Visualisierung der *Fine-Tiling-Array* Plots
- ClustalW V2.0 ([www.ebi.ac.uk/clustalw/](http://www.ebi.ac.uk/clustalw/)) Onlinetool zur Erstellung von Alignmentsequenzen zwischen einer oder mehrerer Proben und einer Referenzsequenz.
- Primer3plus ([www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi](http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi)) Onlinesoftware zum Design von Primern/Probes für Detektion, Sequenzierung, Quantitative Assays uvm.
- NEBcutter V2.0 ([tools.neb.com/NEBcutter2/index.php](http://tools.neb.com/NEBcutter2/index.php)) Onlinetool zur Suche von Restriktionssites in DNA Sequenzen
- REPFIND (<http://zlab.bu.edu/mfrith/cgi-bin/repfind.cgi>) Onlinesoftware zur Detektion repetitiver DNA Sequenzen

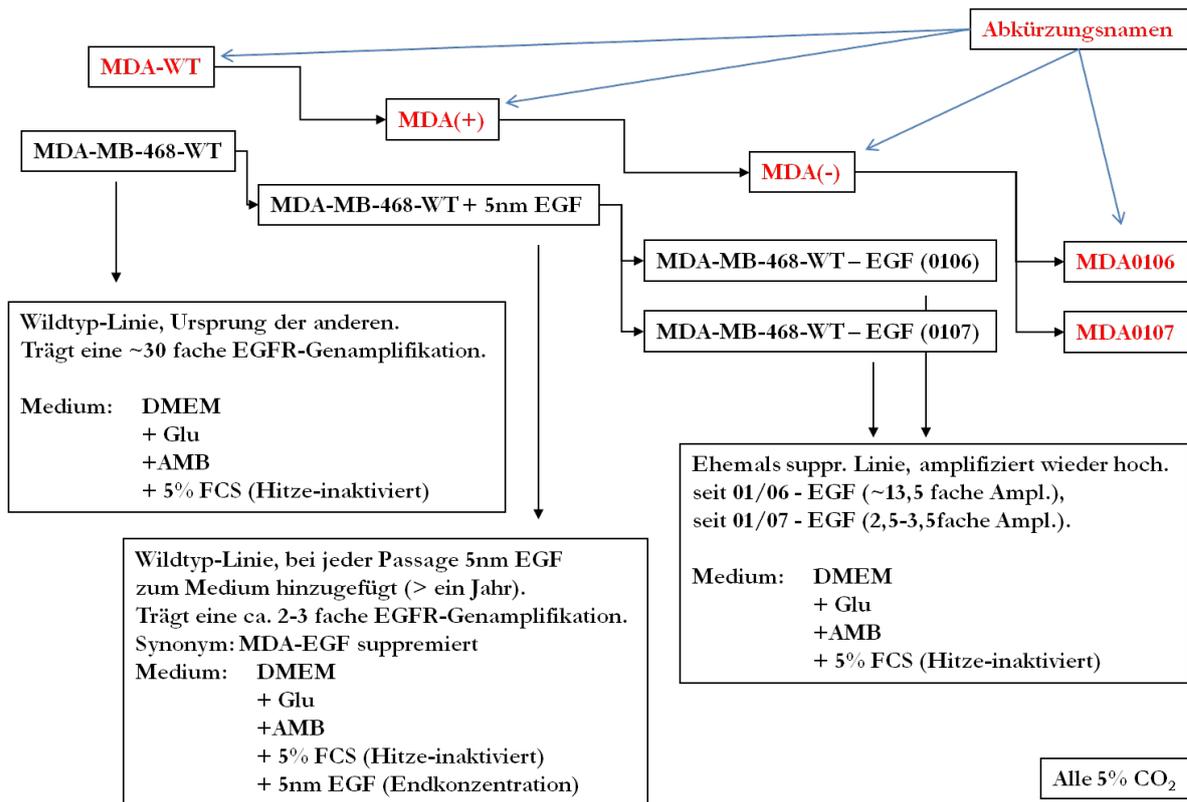
### Datenbanken:

- NCBI (<http://www.ncbi.nlm.nih.gov>) Literatur, DNA, RNA, Proteindatenbanken, BLAST Tools zum Sequenzabgleich uvm.
- UCSC (<http://genome.ucsc.edu/>) DNA, RNA, Proteindatenbank, PCR-Primer Test tool uvm.
- Repeatmasker ([www.repeatmasker.org](http://www.repeatmasker.org)) Datenbank für Repetitive DNA Elemente, Retrovirale Elemente, LTRs uvm.
- IARC TP53 Datenbank ([www-p53.iarc.fr](http://www-p53.iarc.fr)) P53-Datenbank mit Mutationslisten, Fallregistern uvm.

### 3. Ergebnisse

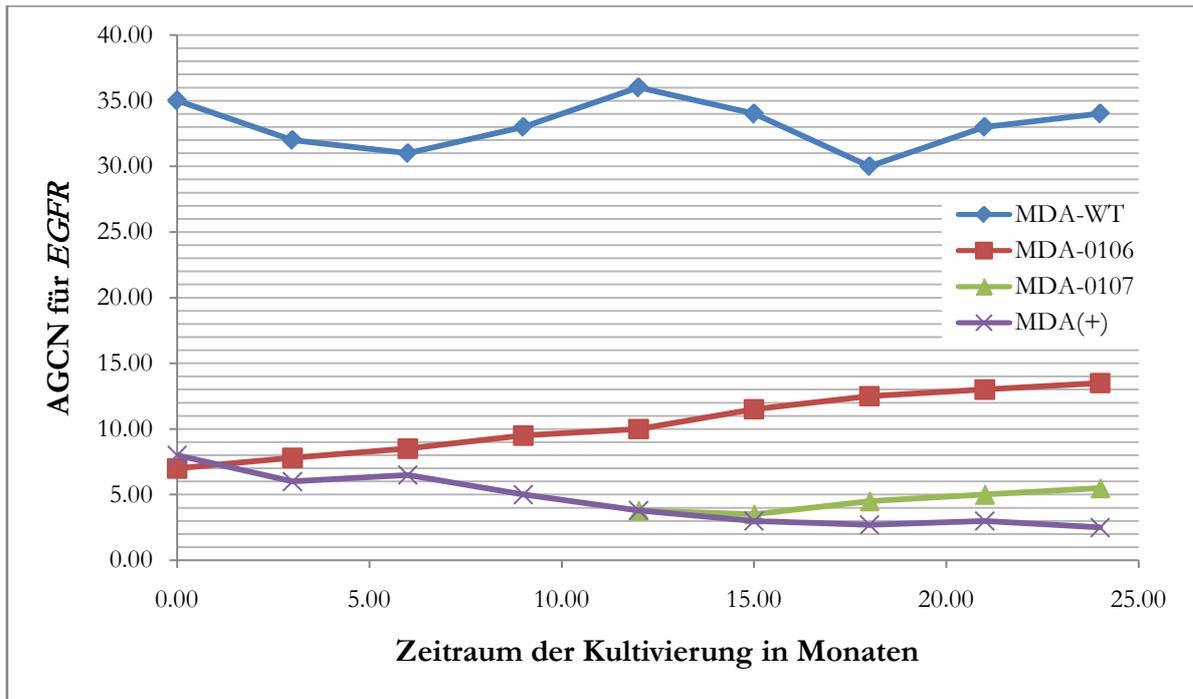
#### 3.1 Veränderung der *EGFR*-Genkopienzahl von MDA-MB-468 unter EGF-Druck

Die Untersuchungen des *EGFR*-Amplikons wurden an dem Model der Mammakarzinom Zelllinie MDA-MB-468 durchgeführt. Diese besaß den Phänotyp eines invasiven Mammakarzinoms und eine distinkte Amplifikation unter Einschluss des *EGFR*-Gens im Bereich des Chromosoms 7p11-14. Die Zelllinie MDA-MB-468 trug eine durchschnittliche Genkopienzahl (*Average Gene Copy Number – AGCN*) für *EGFR* von  $\sim 30$ , wobei der Wert 1 als diploider Satz definiert wurde. Mittels Durchflusszytometrie und Anti-*EGFR*-Antikörper wurden in einer Population von MDA-MB-468 Zellen Subpopulationen isoliert, die eine erhöhte bzw. verminderte *EGFR*-Expression zeigten. Die anschließende Untersuchung durch CGH, FISH und quantitative Real-Time PCR (qPCR) zeigten, dass MDA-MB-468 heterogen für den *EGFR*-Lokus und die AGCN war. Es bestand eine positive Korrelation zwischen der AGCN für *EGFR* und dessen Expression. Die Auswertung der CGH Analysen zeigte keine signifikanten genomischen Unterschiede zwischen den beiden Sublinien, außer dem Zugewinn auf Chromosom 7p11-14. Dieser konnte in beiden Sublinien gezeigt werden, jedoch in unterschiedlich starker Ausprägung. Ferner wurde gezeigt, dass sich die AGCN für *EGFR* der Sublinien unter normalen Zellkulturbedingungen wieder der AGCN für *EGFR* von  $\sim 30$  annäherten (K. Agelopoulos Dissertation). Durch Veränderung der Zellkulturbedingungen wurden Sublinien von MDA-WT (Abb. 3.1) generiert, die verschiedene AGCN für *EGFR* zeigten.



**Abb. 3.1: MDA-MB-468 Sublinien mit unterschiedlichen AGCN für *EGFR*** Dargestellt ist der Stammbaum der Sublinien, die verwendeten Abkürzungsnamen und Kultivierungsmedien, sowie die AGCN für *EGFR* der letzten untersuchten Passagen der Sublinien.

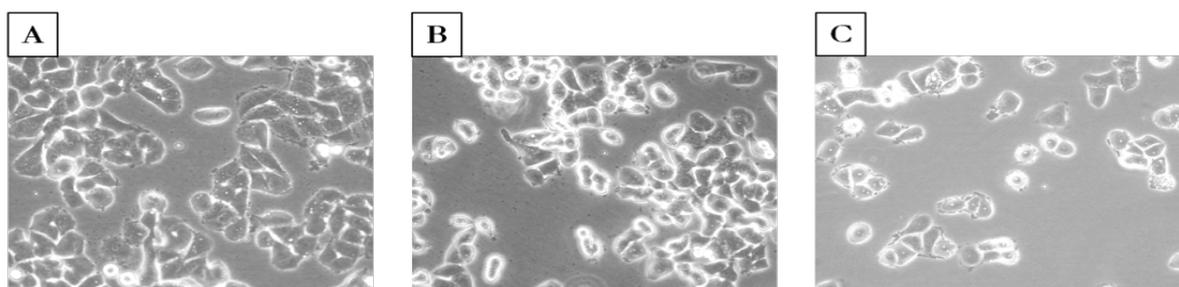
Durch Supplementierung des Standard-Mediums mit 5nM EGF bei jeder Passage wurde ein künstlicher Selektionsdruck geschaffen, der zur Abnahme der AGCN für *EGFR* in MDA(+) führte. Der Subklon MDA(+) zeigte bei andauernder Supplementierung des Mediums eine stabile AGCN für *EGFR* von 2-3. Nach jeweils 12 Monaten wurde ein Teil der Zellen von MDA(+) wieder unter Standardbedingungen kultiviert, was zu einer Zunahme der AGCN für *EGFR* in den Zellen der Sublinien MDA0106 und MDA0107 führte. Im Abstand von 3 Monaten wurde die AGCN für *EGFR* in den 4 Sublinien durch qPCR überprüft. Die Ergebnisse der Messungen der AGCN für *EGFR* der 4 Sublinien wurden in Abb. 3.2 dargestellt. Die Quantifizierung der AGCN für *EGFR* der verwendeten Zelllinien durch qPCR wird im folgenden Kapitel 3.2 erläutert.



**Abb. 3.2: Änderungen der AGCN für *EGFR* der MDA-MB-468 Linie und ihrer Sublinien** Das Diagramm zeigt die Änderungen der AGCN für *EGFR* (vertikale Achse) der 4 Sublinien über einen Zeitraum von 24 Monaten (horizontale Achse).

MDA-WT zeigte über den Zeitraum von 24 Monaten eine stabile AGCN für *EGFR* von 30-35. Bei MDA(+) konnte über die Zeit ein nahezu lineares Absinken auf eine AGCN für *EGFR* von 2-3 beobachtet werden, die bei diesem Wert stagnierte. Die Linien MDA0106 und -0107 zeigten beide eine stetige Zunahme der AGCN, welche in beiden Fällen nahezu linear und mit derselben Steigung verlief. Beim Zeitpunkt 24 Monate wurde für MDA0106 eine AGCN für *EGFR* von 13-14 und MDA0107 eine AGCN für *EGFR* von 5-6 gemessen.

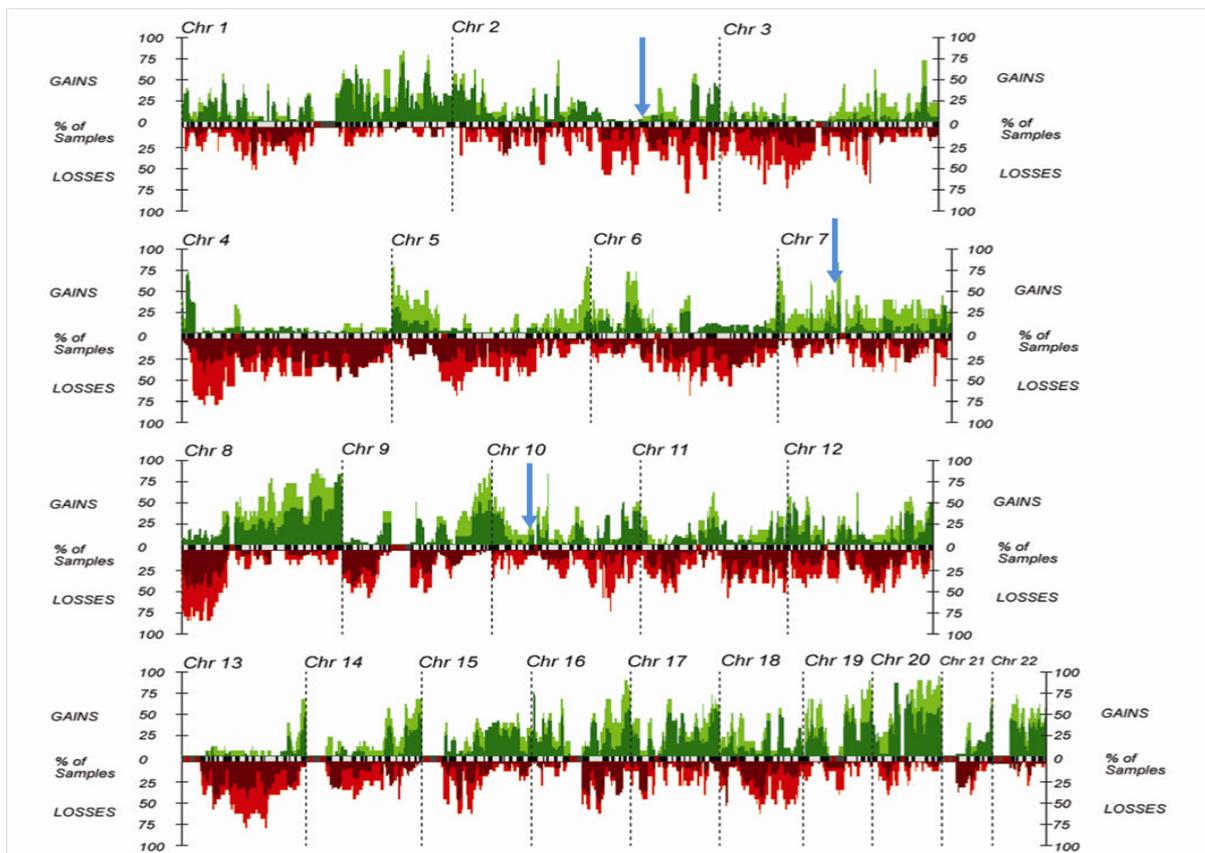
Über den Zeitraum von 24 Monaten konnten zwischen den drei Sublinien und der Mutterlinie keine Unterschiede in Morphologie (siehe Abb.3.3) oder Proliferationsrate festgestellt werden. Die Sublinien zeigten einen identischen epithelialen Phänotyp, eine identische Zellgröße und ein Wachstum als Monolayer. Alle vier Zelllinien erreichten nach dem Splitten (eins zu vier) binnen 3-4 Tagen wieder eine Konfluenz von 80-85%.



**Abb. 3.3: Morphologie der MDA-MB-468 Sublinien.** Lichtmikroskopische Aufnahme (400-fache Vergrößerung) von MDA-WT (A) bei 85% Konfluenz, MDA0106 (B) bei 80% Konfluenz und MDA(+) (C) bei 20% Konfluenz nach dem Passagieren.

### 3.2 Quantitative Real-Time PCR: Etablierung eines TaqMan Assays zur Quantifizierung *EGFR*-Genkopienzahl

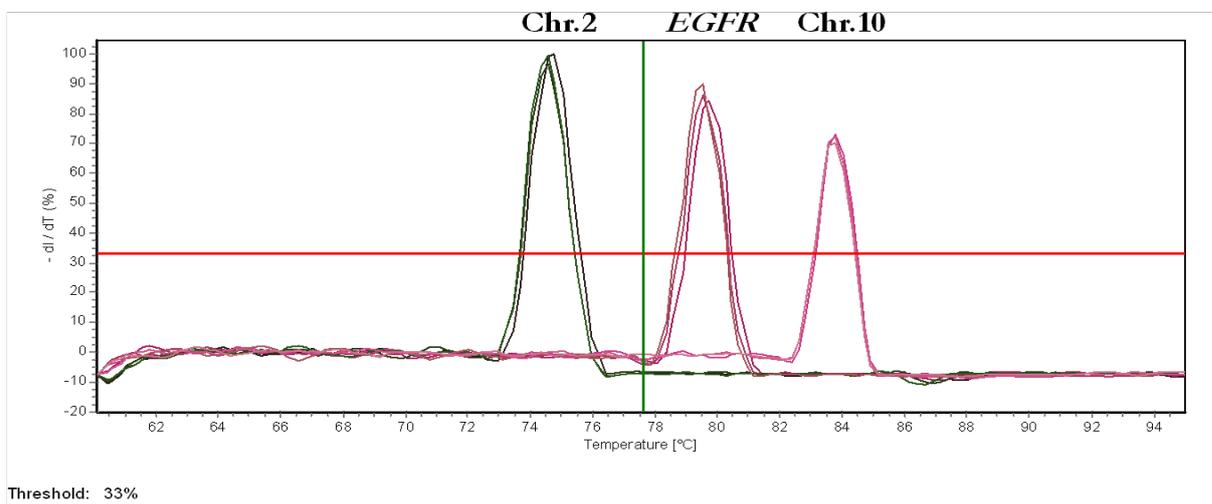
Um eine reproduzierbare Messung der AGCN für *EGFR* zu erreichen, wurde das Assay auf Basis von TaqMan Sonden und unter Verwendung von zwei Referenz Regionen entwickelt. Da Tumore mit fortschreitender Entwicklung immer mehr genomische Aberrationen akkumulieren, zu denen Punktmutationen, Deletionen, Amplifikationen sowie Zugewinne oder Verluste von Chromosomenarmen oder ganzen Chromosomen bis hin zu polyploiden Chromosomensätzen gehören, war die Auswahl der Referenzregionen von entscheidender Bedeutung. Um die Wahrscheinlichkeit abweichender Messungen zwischen den Referenzen zu minimieren, die eine gesicherte Quantifizierung ausgeschlossen hätten, wurden diese basierend auf der Auswertung von 65 High Resolution Array-CGHs ermittelt (162). Das Panel der analysierten Proben umfasste 47 primäre Mammakarzinome sowie 18 Mammakarzinom Zelllinien. Das Array System besaß eine Auflösung von durchschnittlich 0,9Mb. Die Auswahl der Referenz-Loci wurde in Abb. 3.4 dargestellt.



**Abb.3.4: Locus Auswahl der Referenzregionen basierend auf 65 High Resolution Array-CGHs (verändert Naylor et al. 2005 (163))** Die Daten der Array-CGHs wurden nach Chromosomen unterteilt dargestellt, wobei DNA Zugewinne (grün) und DNA Verluste (rot) in Prozent aller Proben aufgetragen wurden. Die blauen Pfeile markieren die Loci der Referenzassays auf Chr.2 und Chr.10 sowie des *EGFR*-Assays auf Chr.7.

Die Auswahl der Referenzregionen wurde basierend auf den Array Daten und einem anschließenden Literatur Screening getroffen. Die Loci auf Chr.2 und Chr.10 zeigten für Mammakarzinome geringe Aberrationsraten von <10%. Der Locus des *EGFR*-Gens, wies in >55% aller analysierter Array-CGHs DNA-Zugewinne auf.

Nach dem Primer- und TaqMan Sonden Design wurden die Primer anhand einer Kalibrationsgerade in einem SYBR green Assay auf ihre Effizienzen überprüft (siehe 2.2.3.13). Es ergaben sich Effizienzen von 98,2% (Chr.2), 98,4% (Chr.10) und 98,7% (*EGFR*). Durch die Schmelzkurvenanalyse (Abb. 3.5) sowie die Überprüfung der PCR Produkte auf einem Agarosegel wurden Primer-Mismatching und Dimerbildung ausgeschlossen.



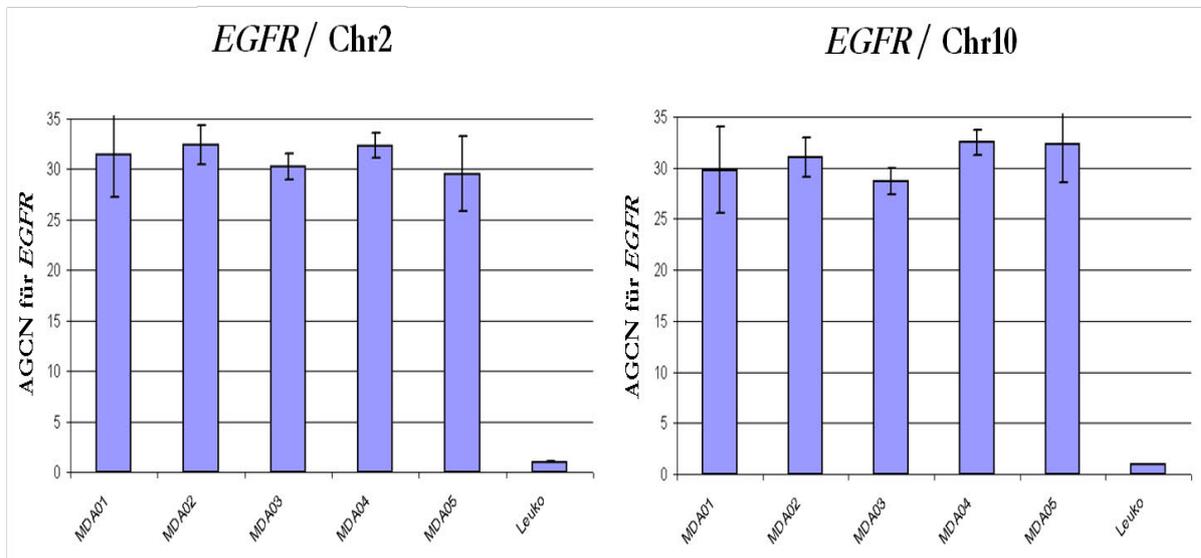
**Abb. 3.5: Schmelzkurven der SYBR green qPCR Assays für Chr.2, Chr.10 und *EGFR***

Auf der horizontalen Achse ist der Temperaturgradient in °C, auf der vertikalen Achse die negative 1. Ableitung der Fluoreszenz gegen die Temperatur dargestellt.

Über den Verlauf des Temperaturgradienten wurden singuläre Schmelzkurven für die drei Assays mit Schmelztemperaturen von ~75°C (Chr.2), ~80°C (*EGFR*) und ~84°C (Chr.10) ermittelt. Die singulären Peaks ergaben sich aus den jeweils spezifischen PCR Produkten. Die Auftrennung der PCR Produkte über einen Agarosegel verifiziert die Resultate der Schmelzkurvenanalyse.

Die Präzision der Messungen der TaqMan Assays wurde anhand der Variationskoeffizienten (VK) ermittelt. Der VK beschreibt die relative Standardabweichung in Prozent. Die Messungen wurden in fünf unabhängigen Wiederholungen als absolute Quantifizierungen durchgeführt.

Als Template wurde in allen Messungen die identische Präparation von genomischer MDA-WT DNA verwendet. Die Ergebnisse wurden in Abb. 3.6 als relative AGCN für *EGFR* gegen die jeweilige Referenz dargestellt.



**Abb. 3.6: Relative AGCN für *EGFR* gegen die Referenzregionen Chr.2 und Chr.10** Dargestellt sind die relativen AGCN für *EGFR* gegen Chr.2 oder Chr.10 für MDA-WT und eine Leukozyten Diploid-Kontrolle.

Der berechnete VK für die Quantifizierung gegen die Chr.2 Referenz lag bei maximal 8%, gegen die Chr.10 Referenz bei maximal 11%. Der VK bei der Quantifizierung der beiden Referenzregionen gegeneinander lag bei maximal 9%. Die Quantifizierung der Leukozyten Diploid-Kontrolle ergab in allen Wiederholungen einen maximalen VK von <5%. Für MDA-WT wurden keine Aberrationen der Referenzregionen nachgewiesen. Bedingt durch die beiden Referenzassays war der Probendurchsatz auf einer 96well Platte auf 6 Proben für ein Zielgen limitiert.

Die an MDA-MB-468 Sublinien, sowie an den Zelllinien A431 und BT-20 und MCF-7 mit unterschiedlichen AGCN für *EGFR*, durchgeführten Fine Tiling Array-CGHs zur Untersuchung der *EGFR*-Amplikons werden im folgenden Kapitel beschrieben.

### 3.3 Fine Tiling Array-CGH: Untersuchung der *EGFR*-Amplifikons von Tumorzelllinien

Die CGH ist eine Methode, mit der sich chromosomale Aberrationen nachweisen lassen. Tumor- und Referenz-DNA wurden dabei bei der Array-CGH gleichzeitig auf Oligonukleotiden hybridisiert. Sowohl die Proben DNA als auch Referenz wurden vor der Hybridisierung mit Fluoreszenzfarbstoffen markiert, um nach der Hybridisierung durch Einlesen an einem Microarray Scanner quantifiziert werden zu können. Die Markierung der Proben DNA erfolgte mit CY3 (532nm), die der Referenz DNA mit CY5 (633nm).

Die verwendeten Fine Tiling Arrays wurden nach unseren Spezifikationen erstellt. Jedes Array bestand aus 395.000 Oligonukleotiden. Bei einer Oligonukleotidlänge von 50-75 Basen und einem Tiling von 15 Basen ergab sich ein maximaler abzubildender Bereich von ~5,7Mb. Die Oligonukleotide wurden mit einer einheitlichen Hybridisierungstemperatur von 76°C und einer singulären Sequenz designed. Die schematische Verteilung der Oligonukleotide eines Fine Tiling Array wurde in Abb. 3.7 dargestellt.

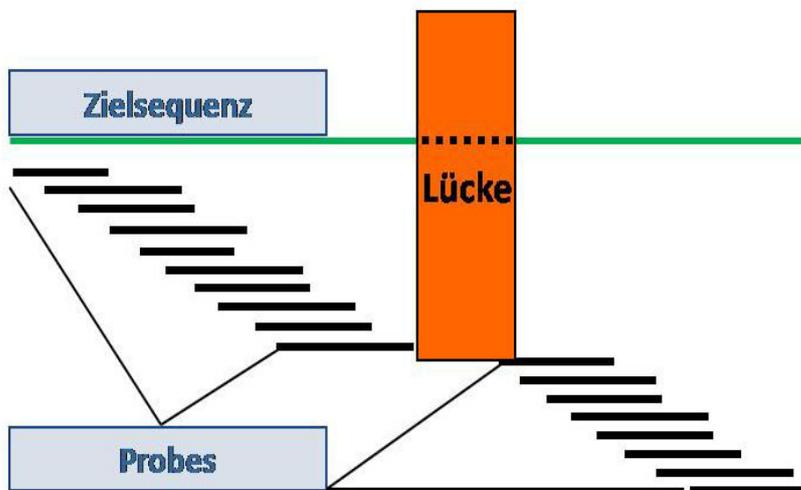


Abb. 3.7: Verteilung der Oligonukleotide auf einem Fine Tiling Array CGH.

Die Zielsequenz wurde durch sich überlappende Oligonukleotide abgedeckt. Statistisch begann um jeweils 15 Basen versetzt ein neues Oligonukleotid (Probe). Bedingt durch die Länge von 50-75 Basen konnte bei hochrepetitiven Zielsequenzen keine singuläre Oligonukleotid Sequenz ermöglicht werden. An diesen Positionen entstanden Lücken, die vom Array nicht darstellen werden konnten.

Die Auswahl der durch die Arrays abgedeckten Bereiche wurde auf der Basis von zwei Datensätzen definiert. Der erste Datensatz war ein an genomischer MDA-MB-468 DNA durchgeführtes BAC-Array (Prof. B. Brandt). Dabei wurden BAC-Klone als Sonden zur Detektion von Kopienzahl Veränderungen verwendet, was eine Auflösung im Bereich von  $\sim 100\text{kb}$  ermöglichte (Tab. 3.1). Die Angaben der Basenposition auf Chr.7 für Klon Start, Ende und Mittelpunkt beziehen sich auf die Annotation des Human Genome Built 18 (März 2006).

Klon Name	Lokalisation	Klon Start	Klon Ende	Klon Mittelpunkt	log2-Ratio
RP4-718N17	Chr. 7p11.2	51.370.001	<b>51.505.527</b>	51.437.764	0.394
<b>RP11-449G3</b>	<b>Chr. 7p11.2</b>	<b>54.381.610</b>	<b>54.558.811</b>	<b>54.470.211</b>	<b>4.414</b>
<b>RP5-1091E12</b>	<b>Chr. 7p11.2</b>	<b>54.784.158</b>	<b>54.784.461</b>	<b>54.784.310</b>	<b>4.756</b>
<b>RP11-339F13</b>	<b>Chr. 7p11.2</b>	<b>54.997.134</b>	<b>55.122.186</b>	<b>55.059.660</b>	<b>4.453</b>
<b>RP4-725G10</b>	<b>Chr. 7p11.2</b>	<b>55.903.906</b>	<b>56.039.803</b>	<b>55.971.855</b>	<b>1.456</b>
<b>RP11-760D2</b>	<b>Chr. 7p11.2</b>	<b>56.239.132</b>	<b>56.413.678</b>	<b>56.326.405</b>	<b>1.368</b>
RP11-10F11	Chr. 7p11.2	<b>56,379,679</b>	56,543,344	56,461,512	0.146

**Tab. 3.1: BAC-Array Daten von MDA-MB-468 DNA** Dargestellt wurden die jeweiligen Klone mit Start-, Mittel- und Endpunkten auf Chr.7 und die jeweiligen ermittelten log2-Ratios gegen Normal-DNA. Die außerhalb des Amplikon liegenden Grenzen wurden dunkelgrau hinterlegt. Die hellgrau hinterlegten Klone werden von dem Bereich vollständig abgedeckt.

Aufgrund der log2-Ratios  $<1$  wurden die Grenzen, die das Amplikon einschlossen, für Position 51.505.527 und 56.379.679 festgelegt. Der zweite Datensatz, der in das Design der Arrays mit einbezogen wurde, waren die Referenzregionen des TaqMan Assays. Diese wurden mit dem Fokus auf möglichst Aberrations-freie Bereiche als Referenzen für die quantitative Real-Time PCR verwendet und sollten auch für die Arrays als interne Kontrolle zur Normalisierung herangezogen werden. Die Messungen der *EGFR*-Amplifikationen der Proben, die mittels der Fine Tiling Arrays charakterisiert werden sollten, wurden auf der Basis des TaqMan Assays mit Chr.2 und Chr.10 als Referenz erhoben. Die für die Arrays verwendeten Bereiche wurden mittels der UCSC-Datenbank (<http://genome.ucsc.edu>) auf Basis der Annotation des Human Genome Built 18 (März 2006) ermittelt. Die ausgewählten Bereiche basierten auf Assemblys, die durch Alignments von sequenzierten BAC Klonen erstellt wurden. Für die Referenzregionen wurde jeweils ein einzelnes Assembly, für den Bereich auf Chr.7 insgesamt 8 sich überlappende Assemblys verwendet (siehe Tab. 3.2)

Chromosom	Assembly Accession Nr.	Startpunkt	Endpunkt	Länge	Überlappung
2	AC018890	175,272,764	175,446,728	180kb	
7	AACC02000111	51,551,170	53,166,847	1,61mb	
	AC073318	53,085,403	53,247,862	162kb	81kb
	CH236957	53,179,720	55,702,834	2,52mb	68kb
	AC091812	55,692,954	55,815,917	122kb	10kb
	AC092647	55,813,918	55,977,808	164kb	2kb
	AC092579	55,975,809	56,094,766	119kb	2kb
	AC092101	56,092,767	56,131,641	39kb	2kb
	AC006970	56,129,667	56,265,748	136kb	2kb
10	AC079269	23,123,769	23,319,773	196kb	

**Tab. 3.2 Für das Array Design verwendete Assemblys.** Angegeben wurden jeweils Chromosom, Accession Nr., Start- und Endpunkte auf dem jeweiligen Chromosom, die Länge des Assemblys und die Überlappungen mit dem jeweils vorangegangenen Assembly.

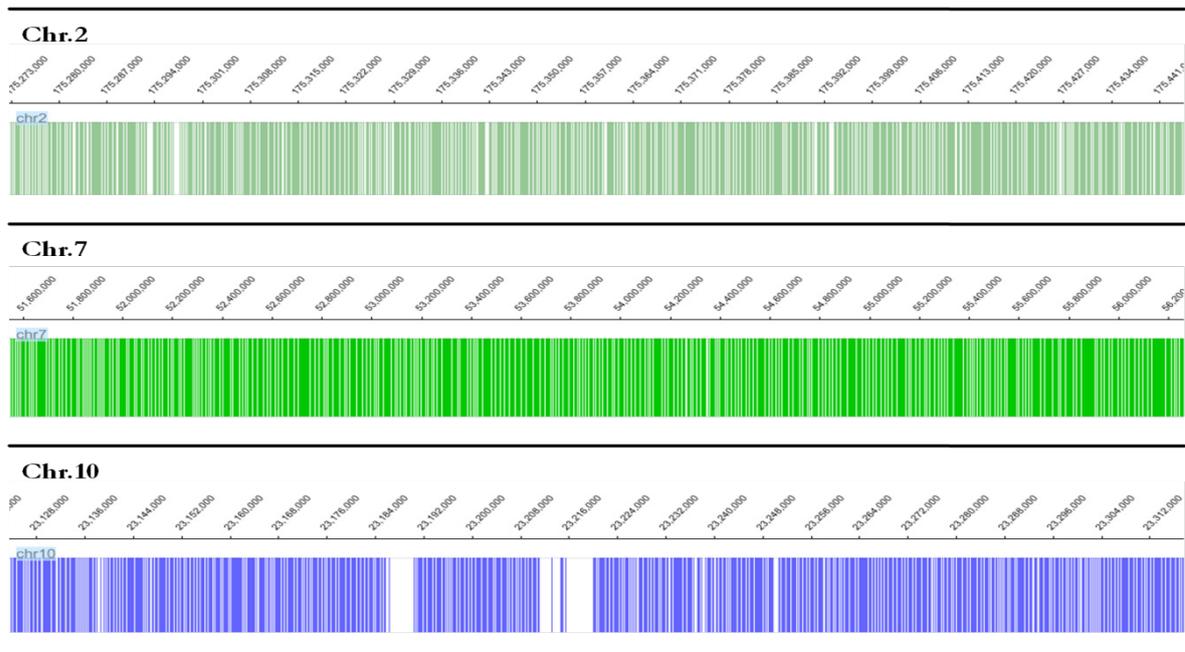
Zusammenfassend deckte das Array Design, das für alle durchgeführten Array-CGHs beibehalten wurde, folgende chromosomale Bereiche ab:

Chr7: 51,551,170 - 56,265,748 ~ 4.7mb

Chr2: 175,272,764 - 175,446,728 ~ 180kb

Chr10: 23,123,769 - 23,319,773 ~ 196kb

Die Abdeckung der Sequenzen durch die Oligonukleotide wurde in Abb. 3.8 anhand des Designfiles dargestellt.



**Abb. 3.8 Designfile des Fine Tiling Array** Dargestellt sind die chromosomalen Positionen der Bereiche auf Chr.2, Chr.7 und Chr.10. Die farbigen vertikalen Linien geben jeweils die Positionen der Oligonukleotide wieder. Weiße Bereiche konnten aufgrund repetitiver DNA Sequenzen nicht vom Array abgebildet werden.

Die acht verwendeten Proben beinhalteten Leukozyten DNA als Normalreferenz, MCF-7 DNA aufgrund der Hemizygotie für *EGFR*, MDA-MB-468 und deren Sublinien, sowie die Zelllinien A431 und BT-20. Das Probenet mit den durch qPCR bestimmten AGCN für *EGFR* wurde in Tab. 3.3 zusammengefasst.

Probe	Typ	AGCN für <i>EGFR</i>
Leukozyten	DNA	keine
MCF-7	DNA	0,5
BT-20	DNA	12,5
A431	DNA	40
MDA-WT	DNA	27
MDA(+)	DNA	2,5
MDA0107	DNA	7,5
MDA0106	DNA	17,5

Tab. 3.3: AGCN für *EGFR* der Fine Tiling Array Proben

Die erhaltenen Datensätze wurden mit der Software Quantsmooth unter Verwendung der Regionen von Chr.2 und Chr.10 normalisiert und die Signale geglättet. Anschließend wurde die Glättungsfunktion des Algorithmus von Quantsmooth verwendet, um die maximalen Änderungen in den Anstiegen und Gefällen der Messpunkte der Arrays zu detektieren. Diese wurden dann als Start- bzw. Endpunkte der jeweiligen Amplikons definiert. Die berechneten Punkte wurden anschließend visuell anhand der Plots verifiziert und mittels qPCR validiert (Kap. 3.4). Als erstes wurden die MDA-MB-468 Sublinien verglichen und anhand der Leukozyten DNA Referenz die Qualität des Array Design überprüft (Abb. 3.9). Dargestellt wurden die jeweiligen Plots der einzelnen Proben für die Regionen von Chr.2, Chr.7 und Chr.10 (horizontale Achse) unter Angabe der jeweiligen Start und Endpunkte, sowie die log<sub>2</sub>-Ratios der normalisierten und geglätteten Werte der hybridisierten Proben-DNA gegen die Referenz-DNA (vertikale Achse). Ebenfalls angegeben wurden die über qPCR bestimmten AGCN für *EGFR* der einzelnen Proben. Der Bereich des *EGFR*-Gens wurde durch die breite grüne Linie markiert. Die dünnen grünen Linien beschrieben die errechneten Amplikongrenzen für MDA-WT und dienten der Erleichterung der Orientierung zum Vergleich mit den anderen Proben. Diese Darstellungsweise wurde für Abb. 3.11 übernommen.

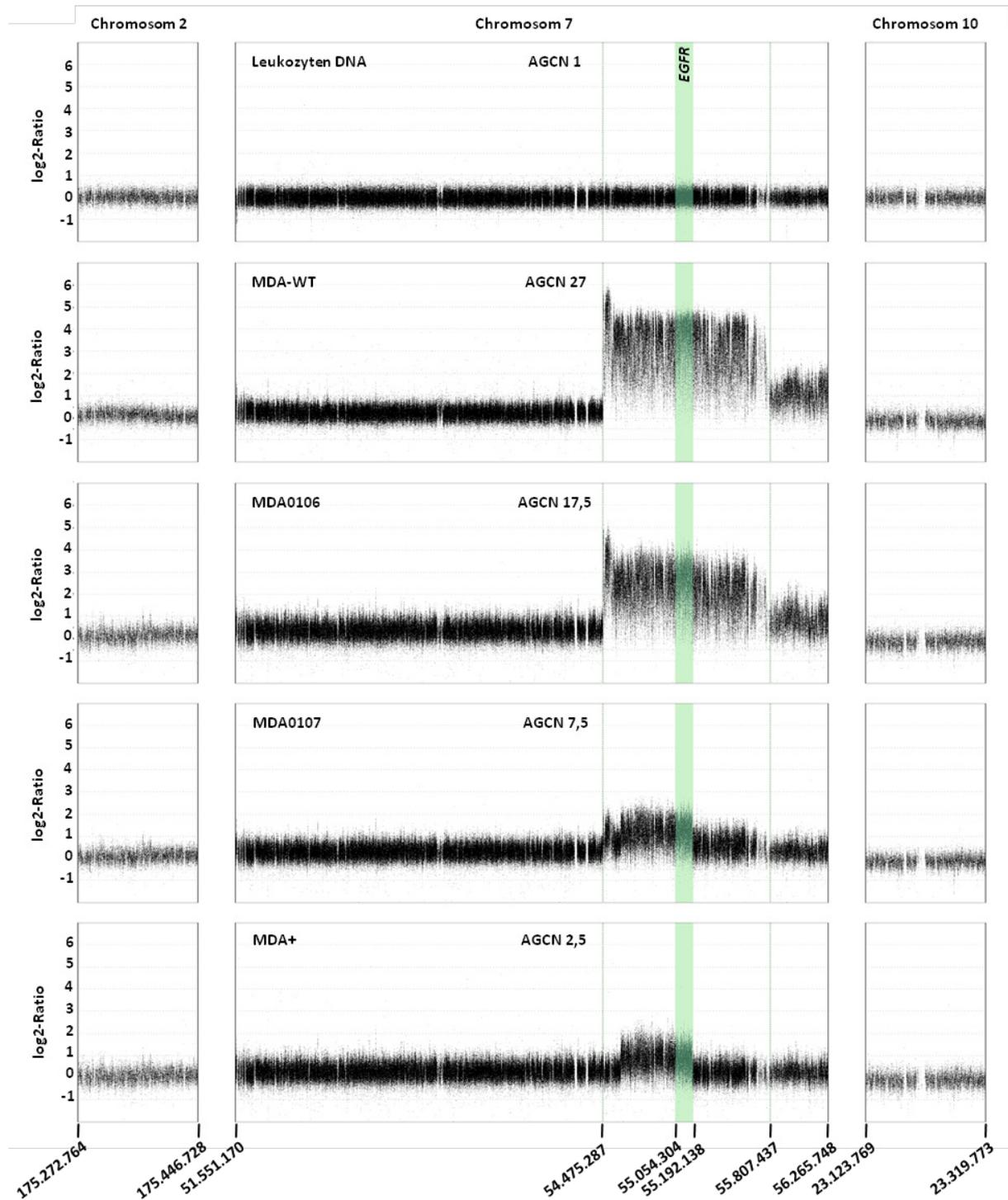


Abb. 3.9: Array CGH Plots von Leukozyten und MDA-MB-468 und deren Sublinien

Die Leukozyten DNA zeigte keine Veränderungen auf DNA-Kopienebene. Die Plots für MDA-WT und deren Sublinien zeigten keine Aberrationen in den Bereichen von Chr.2 und Chr.10. Insgesamt korrelierten die semiquantitativen log<sub>2</sub>-Ratios der Arrays mit den durch qPCR gemessenen AGCN für *EGFR*. Die Plots von MDA-WT und MDA0106 verliefen für das Amplikon identisch in Struktur und Länge.

Die log<sub>2</sub>-Ratios des Amplikon für MDA-WT lagen im Mittel bei 4-5, die von MDA0106 im Mittel bei 3-4. Centromerwärts des Hauptamplikons lag bei MDA-WT und MDA0106 ein amplifizierter Bereich mit log<sub>2</sub>-Ratios von ~2, der sich bis zum Endbereich der Chr.7 Region des Arrays ausdehnte. Der Plot von MDA0107 zeigte ein Amplikon, das in der Länge mit dem MDA-WT Amplikon übereinstimmte. Das Amplikon zeigte am Ende des *EGFR*-Gens eine Stufe, bei der die mittleren log<sub>2</sub>-Ratios von ~2 auf ~1 abfielen. Diese Stufe bildete, sowohl für MDA0107 als auch für MDA(+), das errechnete Ende des Hauptamplikons. Der Plot von MDA(+) zeigte, verglichen mit dem MDA0107 Amplikon, telomerwärts eine Verkürzung des Amplikon und im Mittel log<sub>2</sub>-Ratios von ~2. Telomerwärts entfiel bei MDA(+) die Peak-artig aufragende Struktur, die bei MDA-WT und MDA0106 das Amplikon um 1 log<sub>2</sub>-Ratio überragte. Bei MDA0107 lag diese Struktur mit log<sub>2</sub>-Ratios von 2 auf dem Niveau des Amplikons.

Mit ansteigender Amplifikation vergrößerte sich das Amplikon der MDA-MB-468 Sublinien. Ab dem Amplifikationsgrad von MDA0106 schien neben dem für den MDA-WT definierten Hauptamplikon ein weiterer Bereich mit zu amplifizieren, der centromerwärts verlief. Die Zugewinne der Amplikons sowie *EGFR*-Amplifikationen der MDA-MB-468 Sublinien wurden in Abb. 3.10 zusammenfassend dargestellt.

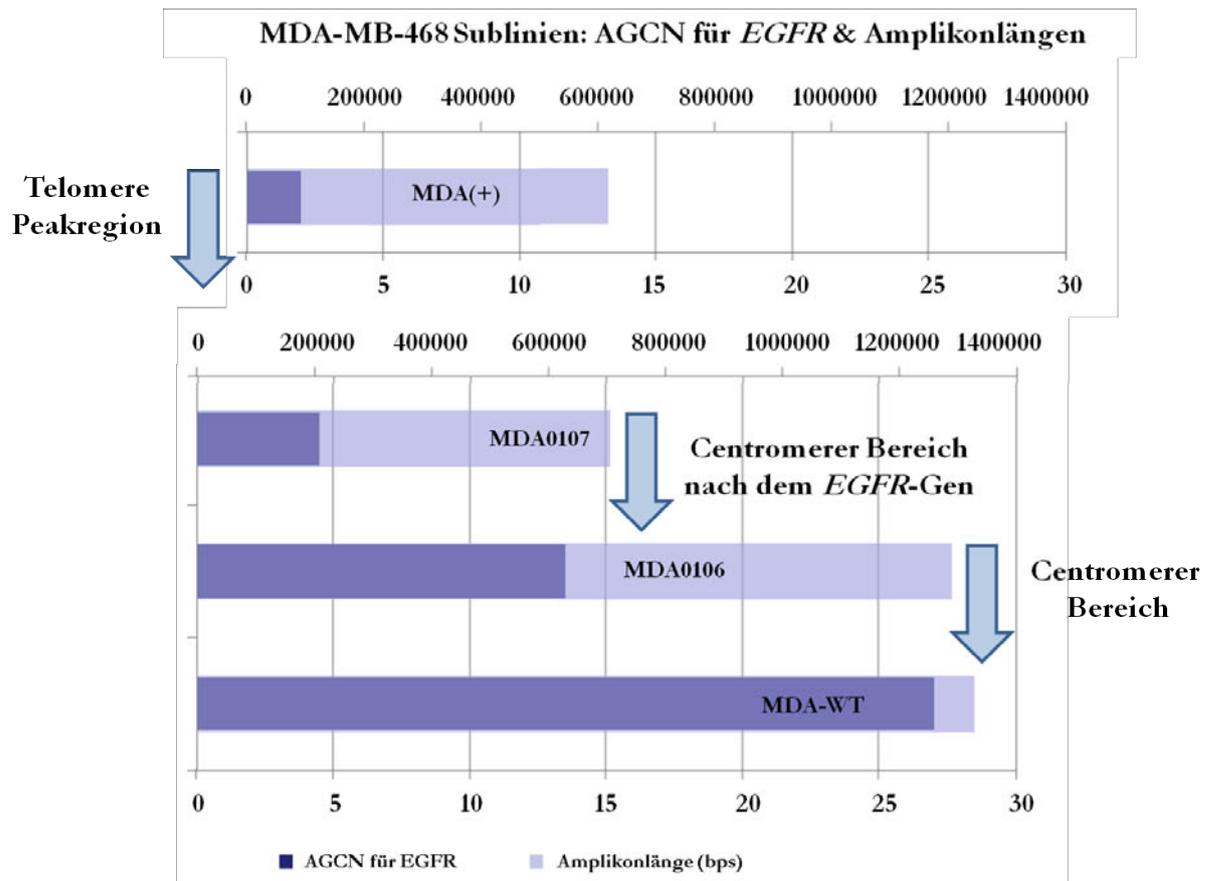


Abb. 3.10: Gegenüberstellung der Amplikonlängen und *EGFR*-Amplifikationen der MDA-MB-468 Sublinien

Eine Zunahme der AGCN für *EGFR* führte zu einer nicht linearen Zunahme der Amplikonlänge. Der Zugewinn zwischen MDA(+) und MDA0107 bedingte sich durch die telomerwärtige Verlängerung des Amplikons bei MDA0107 um ~87kb bei einer um Faktor 3 erhöhten AGCN für *EGFR*. Zwischen den Sublinien MDA0107 und MDA0106 kam es zu einem Ansteigen der Amplikonlänge um ~511kb bei einer um Faktor 2,3 erhöhten AGCN für *EGFR*. Zwischen MDA0106 und MDA-WT kam es bei einer um Faktor 1,5 erhöhten AGCN für *EGFR* lediglich zu einem Längenzugewinn von ~34kb. Die Zugewinne zwischen MDA0107, MDA0106 und MDA-WT bedingten sich durch centromerwärtige Verlängerungen des Amplikons.

Die Array Plots der Zelllinien MCF-7, A431 und BT-20 wurden vergleichend zu den Strukturen und Grenzen der Amplifikons der MDA-MB-468 Sublinien untersucht (Abb. 3.11).

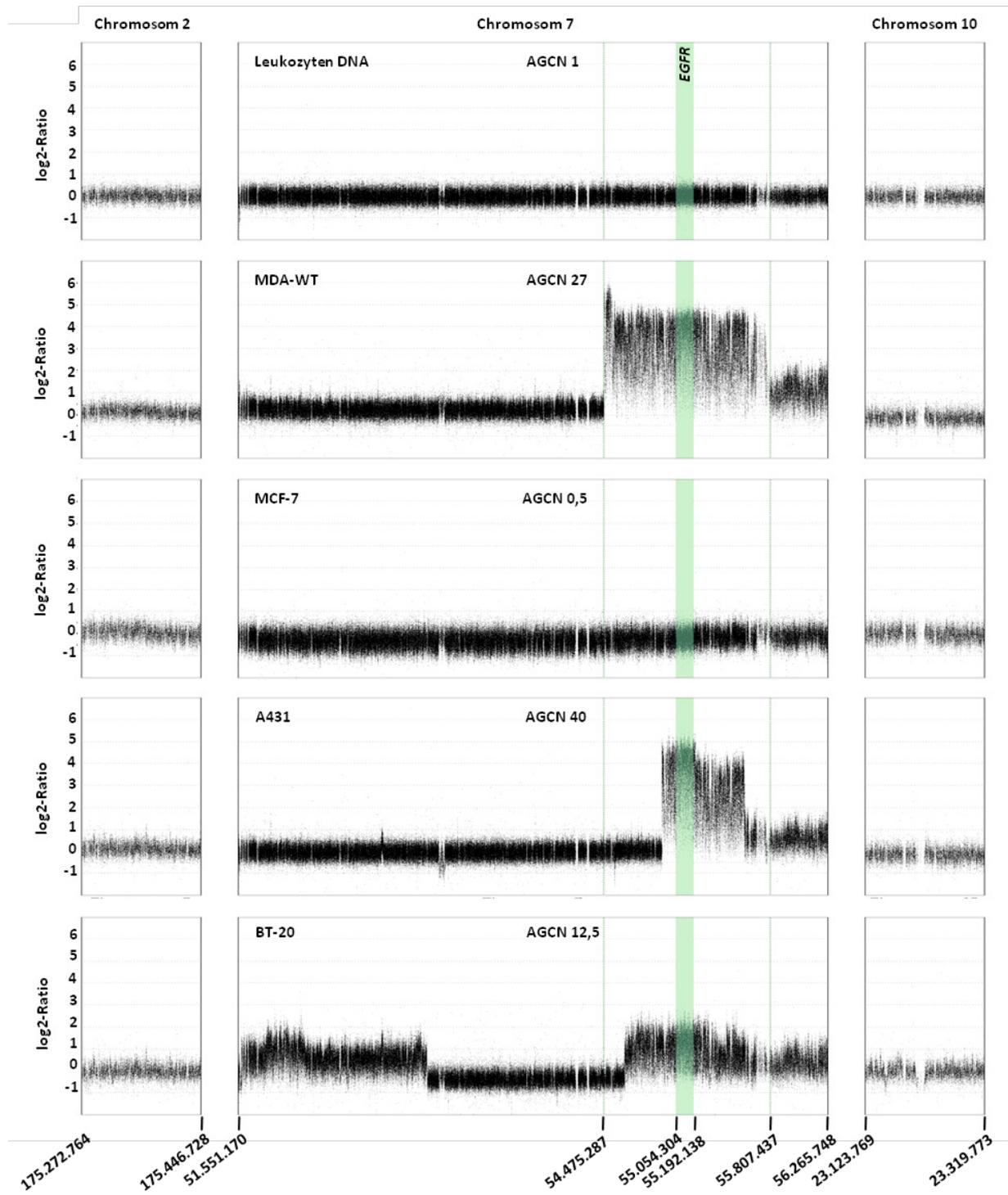
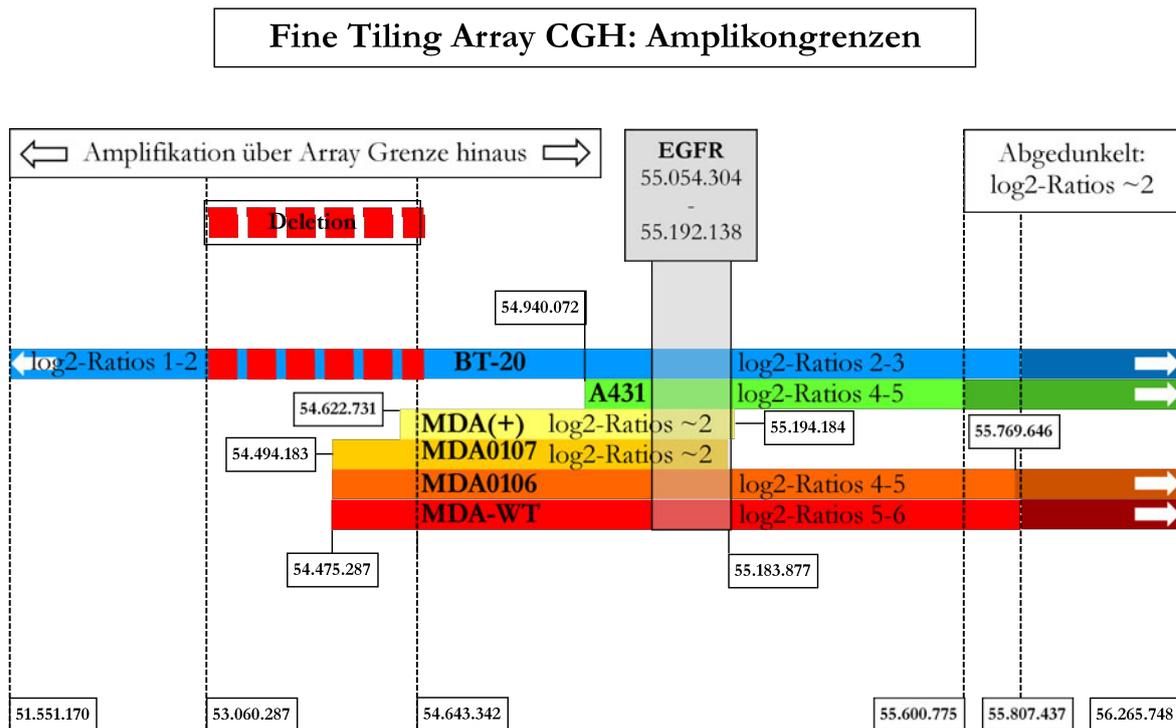


Abb. 3.11: Array CGH Plots von Leukozyten, MDA-MB-468, MCF-7, A431 und BT-20

Die Zelllinie MCF-7 zeigte über den gesamten Bereich von Chr.7 eine auf den Verlust eines der Allele zurückzuführende log<sub>2</sub>-Ratio von 0,5. Die Bereiche auf Chr.2 und 10 zeigten keine Aberrationen, wie auch für die Linien A431 und BT-20. Der Plot für die Linie A431 stellte im Vergleich zu MDA-WT ein verkürztes Amplikon mit scharf definierten Kanten dar. Das Amplikon besaß am Ende des *EGFR*-Gens eine Stufe, bei der die log<sub>2</sub>-Ratio im Mittel von 5 auf 4 abfiel. Centromerwärts folgte wie bei MDA-WT und MDA0106 ein amplifizierter Bereich mit log<sub>2</sub>-Ratios von ~2. Dieser Bereich wurde ebenso für die Linie BT-20 gefunden. Die telomerwärtigen Grenzen der Hauptamplikons von BT-20 und MDA(+) lagen ~21kb voneinander entfernt. Das centromerwärtige Ende des Amplikons verlief parallel zu dem von MDA-WT. Neben dem Hauptamplikon mit log<sub>2</sub>-Ratios von ~2 konnte für BT-20 noch ein weiterer amplifizierter Bereich von ~1,5mb Länge im Anfangsbereich der Chr.7 Region des Arrays gezeigt werden. Die log<sub>2</sub>-Ratios in diesem Bereich lagen bei ~1. Zwischen den Amplikons war ein Bereich einer ~1,6mb langen Deletion eingeschlossen. Das für die MDA-WT Sublinien beschriebene Verhältnis zwischen der AGCN für *EGFR* und der Amplikonlänge konnte für die Linien A431 und BT-20 nicht beschrieben werden. Das Amplikon von A431 war bei einer AGCN für *EGFR* von 40 im Vergleich zum MDA-WT Amplikon um 665kb verkürzt. Das Amplikon für BT20 besaß einen abweichenden strukturellen Aufbau. Der Teil des Amplikon, in dem der *EGFR*-Lokus lag, war für BT-20 (log<sub>2</sub>-Ratios ~2-3) im Vergleich zu dem Amplikon von MDA0107 (log<sub>2</sub>-Ratios ~2) um 464kb verlängert. Die für die Zelllinien festgesetzten Bruchpunkte, sowie strukturelle Gemeinsamkeiten der verschiedenen Amplikons wurden abschließend in der schematischen Abb. 3.12 zusammengefasst.



**Abb. 3.12: Schematische Darstellung der Amplikongrenzen für alle Linien und Sublinien** Proportional zueinander werden Position und Länge der amplifizierten Bereiche innerhalb der Zelllinien dargestellt. Die Position von *EGFR* werde durch den grauen Kasten markiert. Weiße Pfeile deuten das Fortlaufen von amplifizierten Bereichen über die Grenzen des Arrays hinaus an. Die roten Kästen in BT-20 decken den deletierten Bereich ab. Grau abgedunkelte Bereiche in der centromerwärtigen Region symbolisieren die amplifizierten Regionen mit log2-Ratios ~2.

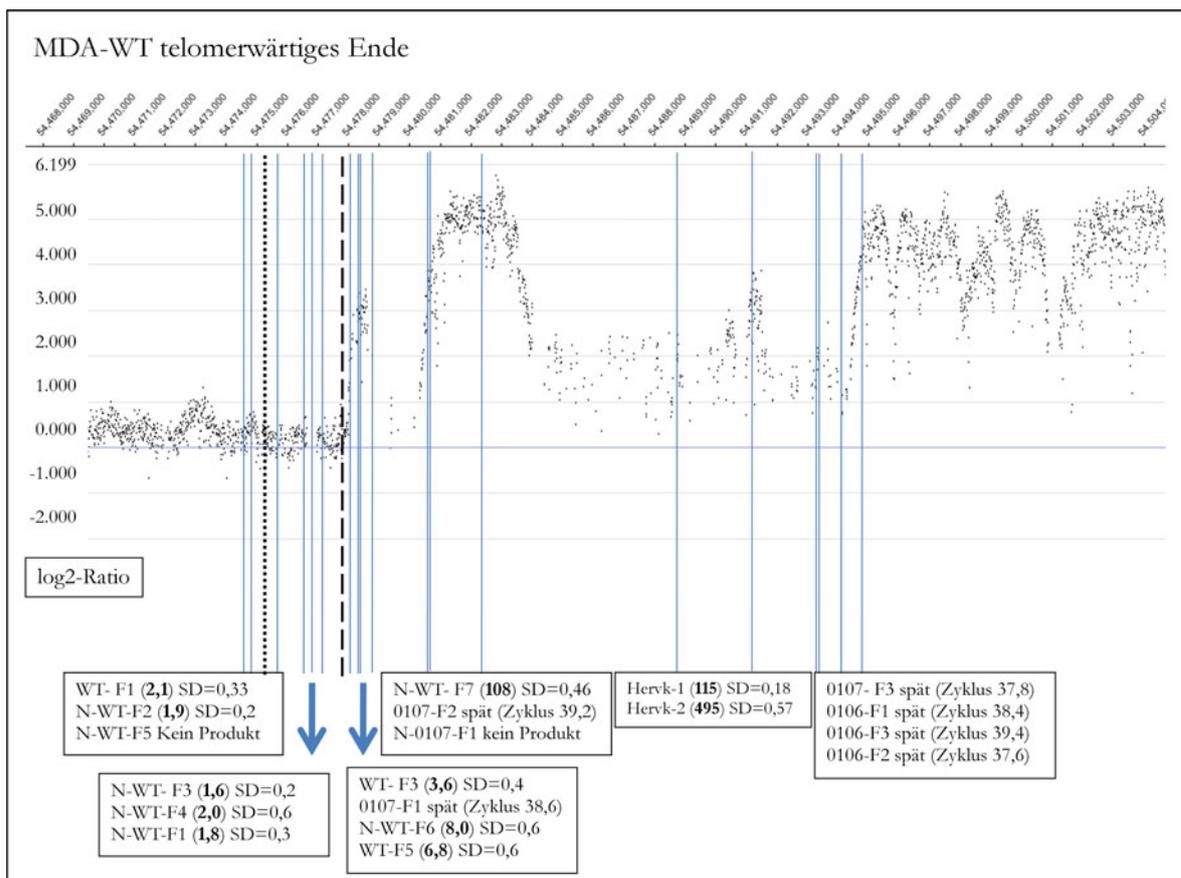
Die berechneten Start- und Endpunkte der amplifizierten Bereiche der Zelllinie MDA-MB-468 und ihrer Sublinien wurden anschließend durch qPCR validiert. Die errechneten Sequenzbereiche wurden in 3'- und 5'-Richtung mit jeweils mindestens 2 quantitativen Assays validiert.

### 3.4 Validierung der errechneten Bruchpunkte über quantitative Real-Time PCR

Die Quantifizierung der Assays wurde an genomischer DNA des jeweiligen MDA-MB-468 Klons und Leukozyten DNA, als diploider Referenz zur Normalisierung, durchgeführt. Die Proben wurden relativ zu einer definierten Menge Leukozyten DNA quantifiziert. Die relative Quantifizierung erhöhte die SD der Messungen, steigert aber den Probendurchsatz pro 96well Platte. Da relative Verhältnisse der Messpunkte zueinander erhoben werden sollten, war der Versuchsansatz der Fragestellung angepasst.

In den folgenden Abbildungen wurden die hochaufgelösten Plots der Start- bzw. Endregionen der Amplikons der Linie MDA-MB-468 und ihrer Sublinien dargestellt. Die Position der quantitativen Assays wurde durch blaue Linien markiert.

Diesen wurden Kästen zugeordnet, in denen Assay-Namen, die AGCN (Fettschrift in Klammern) und die SDs der Messungen aufgelistet wurden. Die Auflistung von oben nach unten entsprach den Assays von links nach rechts. Assays mit CT-Werten >35 wurden nicht quantitativ ausgewertet. Die errechneten Start- bzw. Endsequenzen des Amplikons wurden durch eine gepunktete Linie markiert. Resultierte aus der kombinierten Auswertung der log<sub>2</sub>-Ratios des Plots und der AGCN der qPCR Assays eine abweichender Start- bzw. Endsequenz, wurde diese durch eine gestrichelte Linie markiert (Abb. 3.13-3.20).



**Abb. 3.13: Validierung der telomerwärtigen Startsequenz für MDA-WT**

Durch die log<sub>2</sub>-Ratios des Plots und die AGCN der qPCR wurde eine Startsequenz bei ~54.478.000 festgelegt. Auffallend waren die AGCN der qPCRs für den Bereich N-WT-F7 bis HervK-2 mit Werten von 108-495. Diese wurden mehrfach validiert und als konsistent betrachtet. Ein Endbereich der Hochamplifikation konnte nicht festgelegt werden, da die centromerwärts gelegenen Assays nicht quantitativ auswertbar waren.

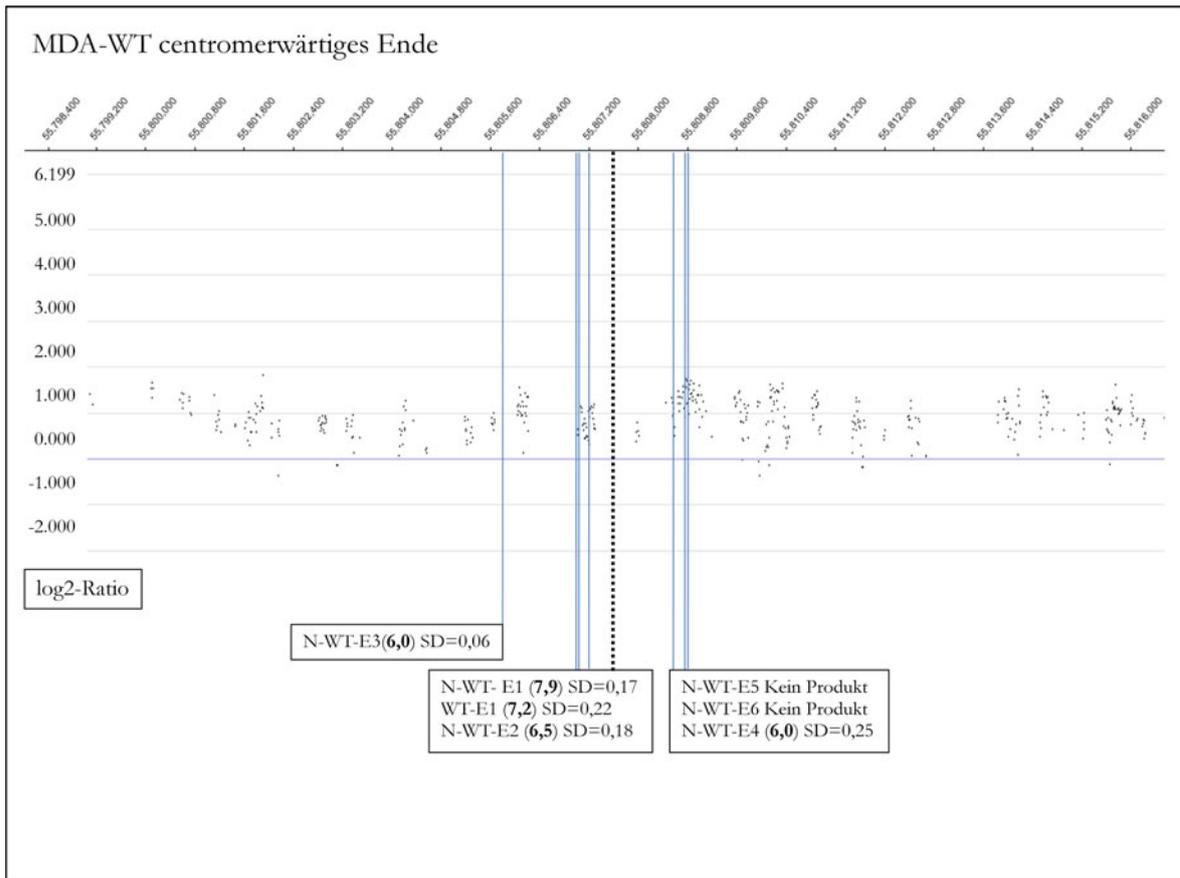


Abb. 3.14: Validierung der centromerwärtigen Endsequenz für MDA-WT

Die AGCN der qPCRs bestätigten in Kombination mit den log<sub>2</sub>-Ratios des Plots die berechnete Endsequenz des Amplikons.

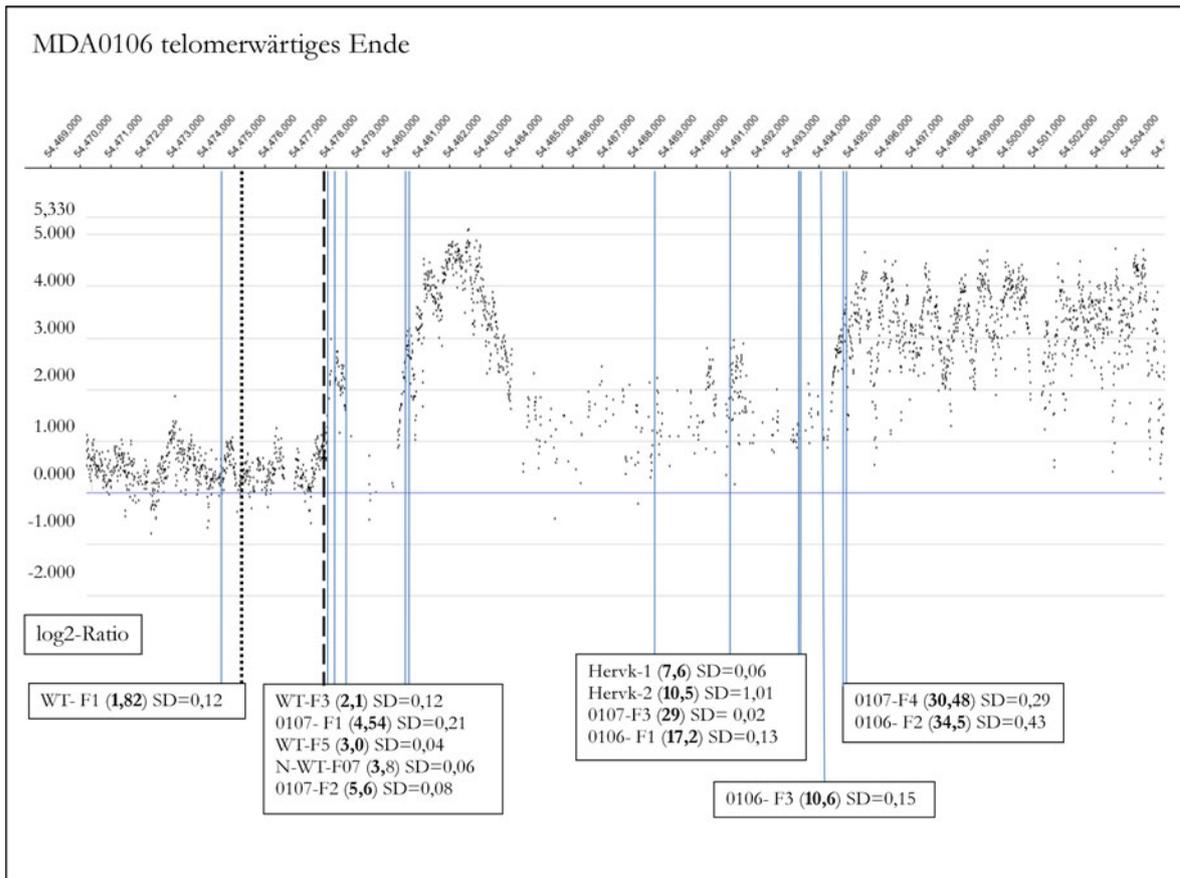


Abb. 3.15: Validierung der telomerwärtigen Startsequenz für MDA0106

Basierend auf den log<sub>2</sub>-Ratios des Plots und den AGCN der qPCRs wurde für MDA0106 die Startsequenz des Amplikons bei ~54.478.000 festgelegt. Die erhöhten AGCN der qPCRs von N-WT-F07 bis Hervk1+2 konnten für MDA0106 mit Werten von 3,8-29 bestätigt werden. Die centromerwärts gelegenen qPCR Assays waren für MDA0106 quantitativ auswertbar, wobei die AGCN der qPCRs den Verlauf der log<sub>2</sub>-Ratios des Plots bestätigten.

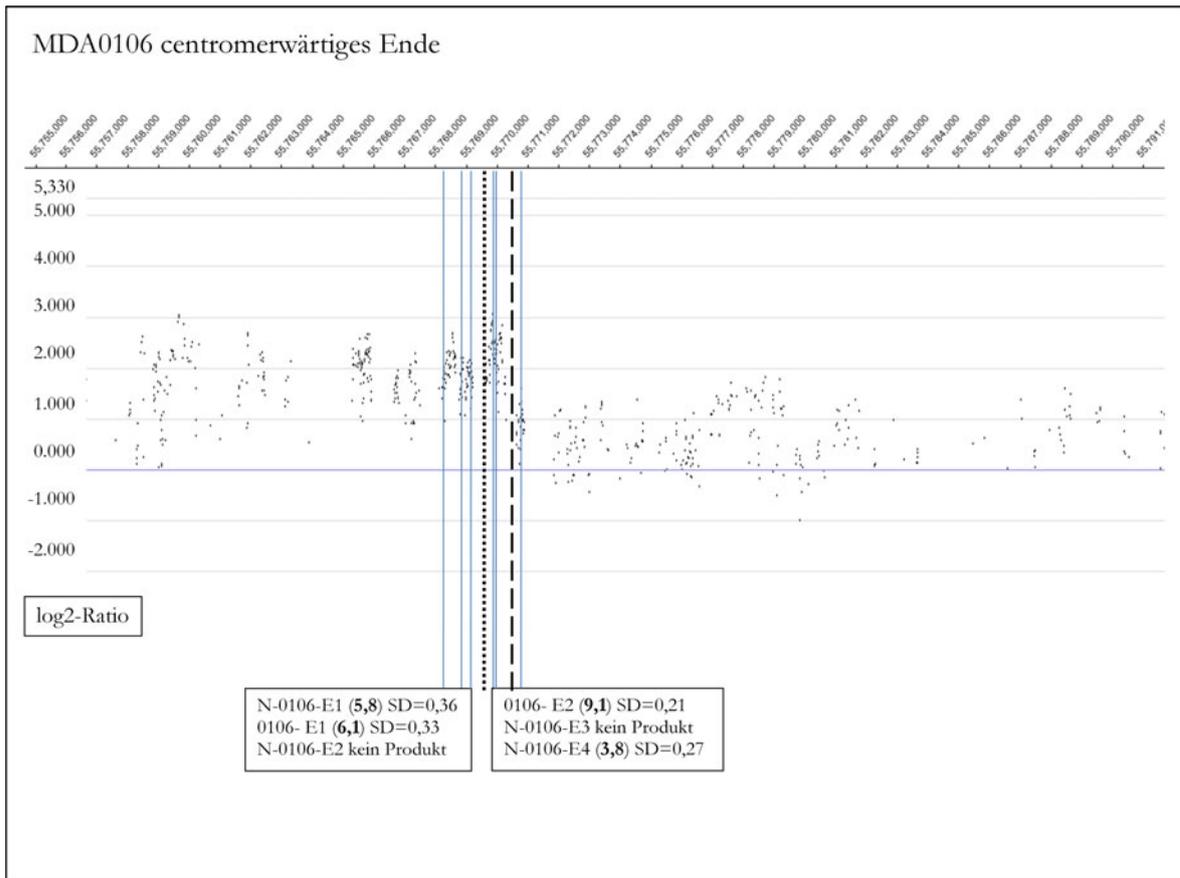


Abb. 3.16: Validierung der centromerwärtigen Endsequenz für MDA0106

Durch die AGCN der qPCRs in Kombination mit den log<sub>2</sub>-Ratios des Plots wurde die Endsequenz des Amplikons bei ~55.770.550 festgelegt.

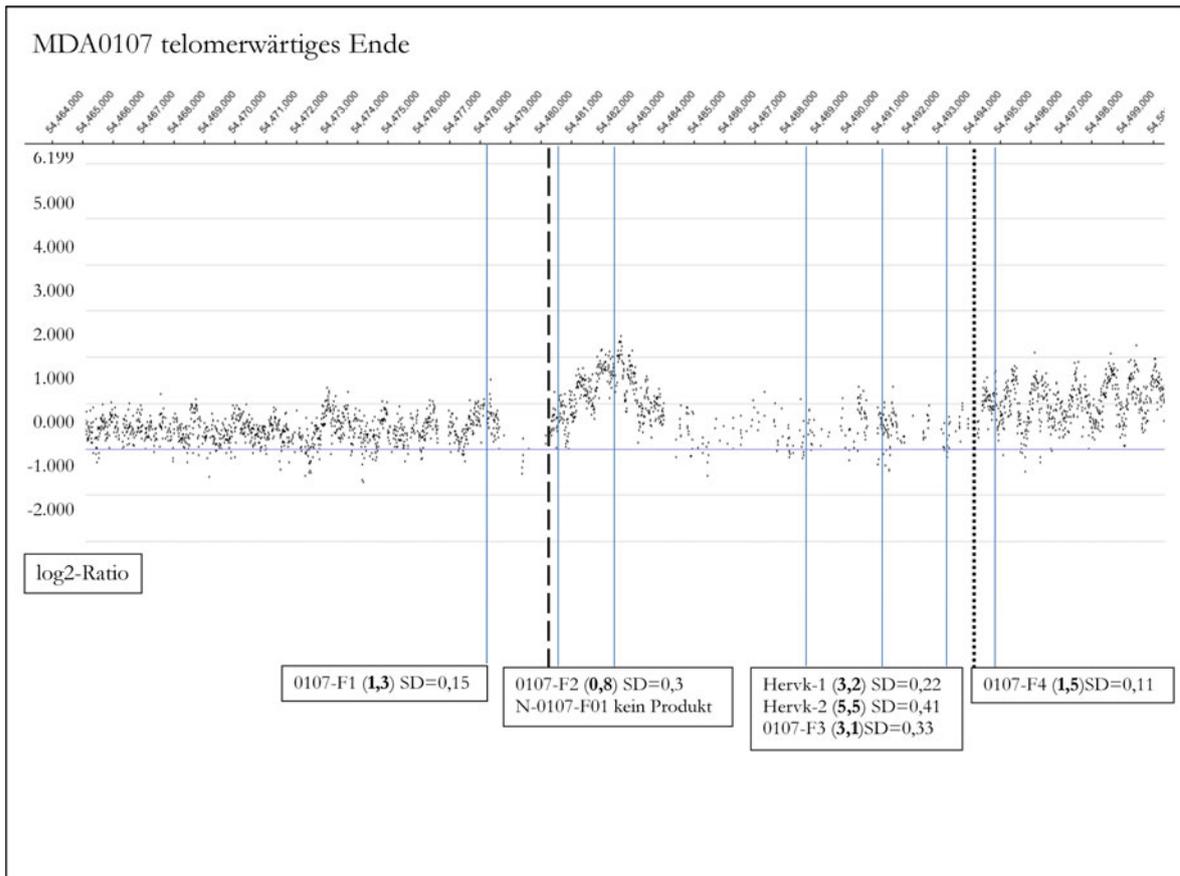


Abb. 3.17: Validierung der telomerwärtigen Startsequenz für MDA0107

Aufgrund des Verlaufs der log<sub>2</sub>-Ratios des Plots und der AGCN der qPCRs wurde die identische Startsequenz bei 54.480.500 festgelegt, wie für MDA-WT und MDA0106. Die AGCN der qPCR Assays von Hervk-1 bis 0107-F3 zeigten relativ zu den Assays 0107-F1 und 0107-F4 erhöhte Werte im Bereich von 3,2-5,5.

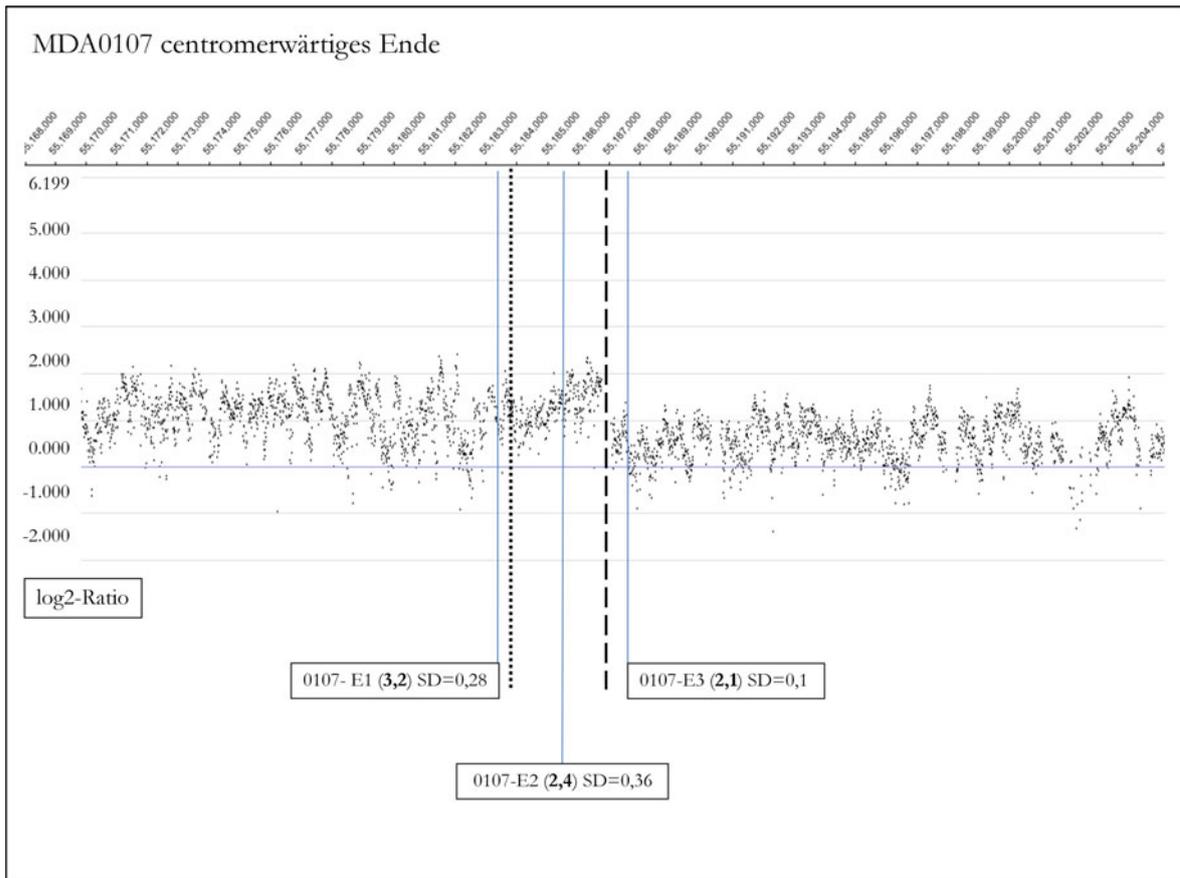


Abb. 3.18: Validierung der centromerwärtigen Endsequenz für MDA0107

Durch die AGCN der qPCRs in Kombination mit den log2-Ratios des Plots wurde die Endsequenz des Amplikons bei  $\sim 55.187.000$  festgelegt.

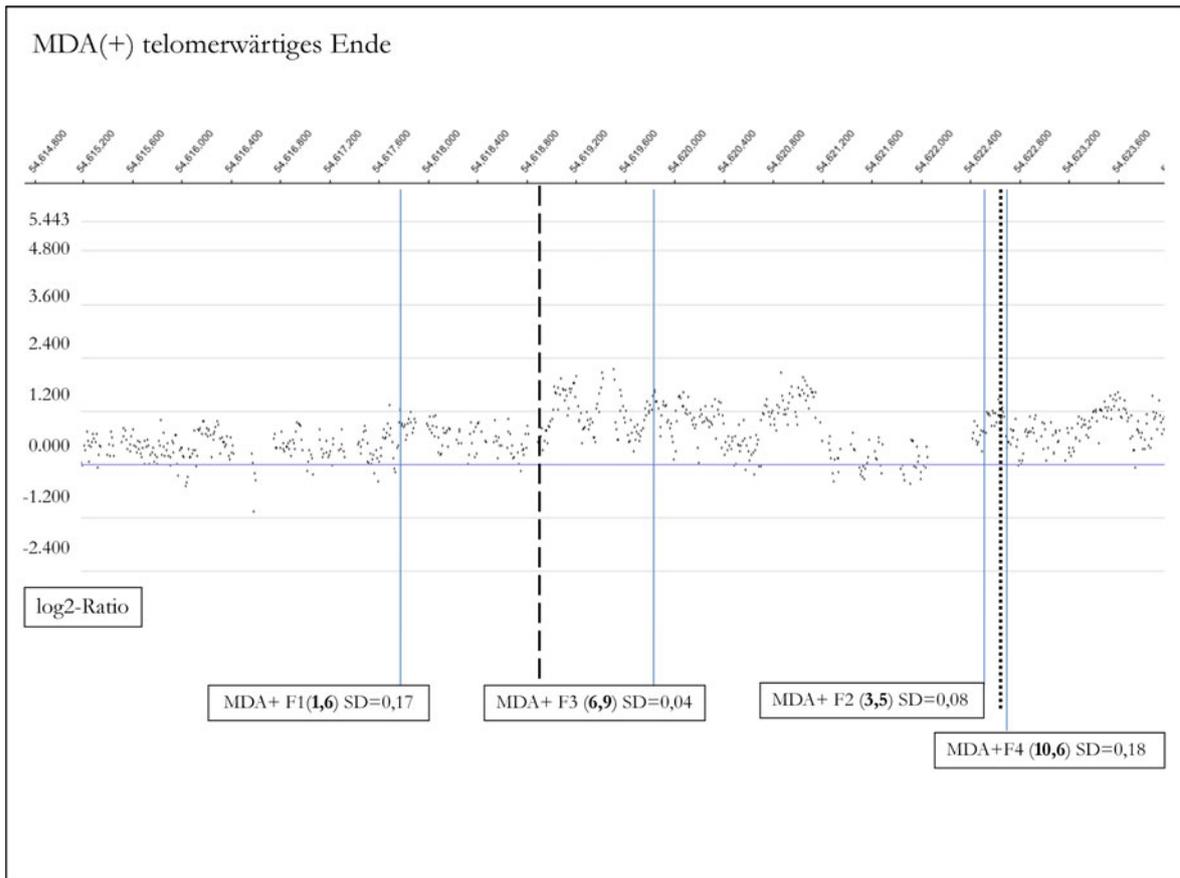


Abb. 3.19: Validierung der telomerwärtigen Startsequenz für MDA(+)

Die AGCN der qPCRs und der Verlauf der log2-Ratios des Plots zeigten einen mit den Startregionen der anderen Sublinien vergleichbaren Verlauf, aufgrund dessen die Startregion bei ~56.618.850 festgelegt wurde.

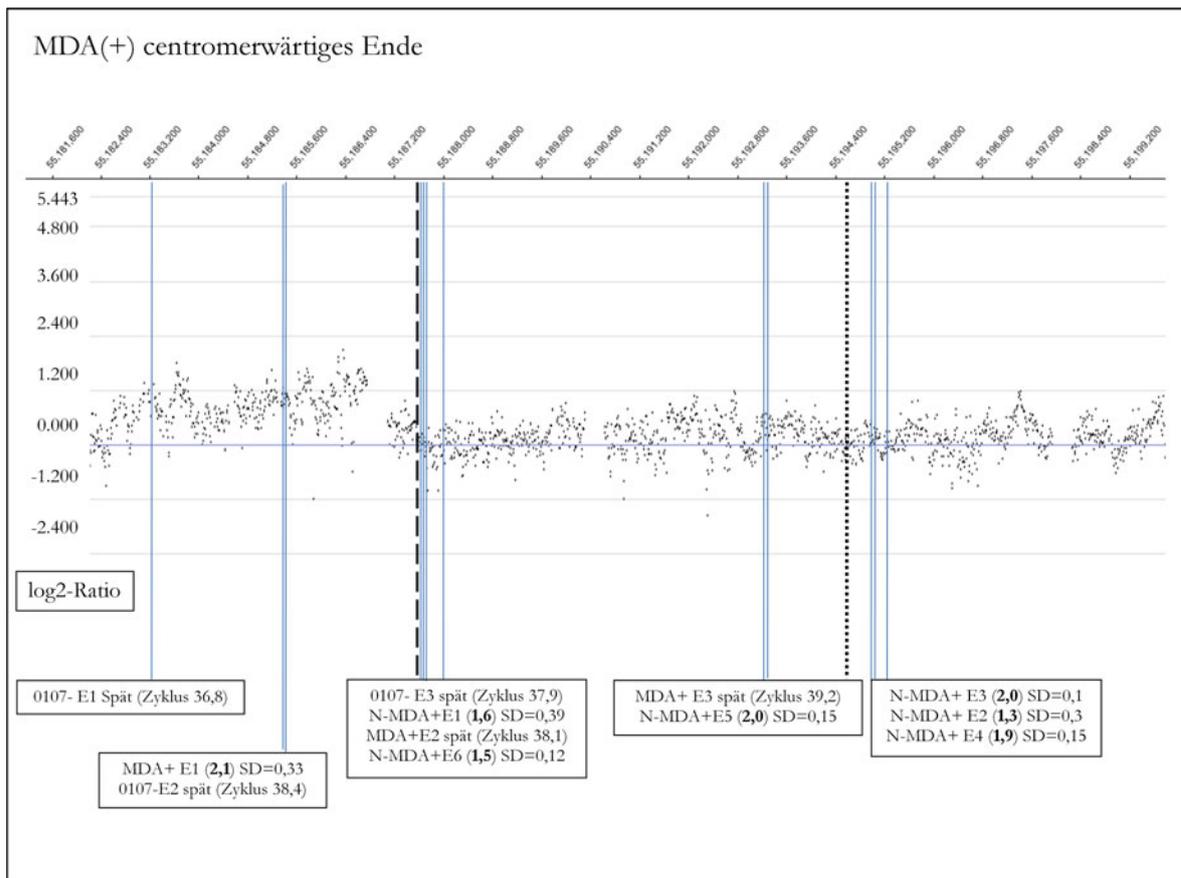
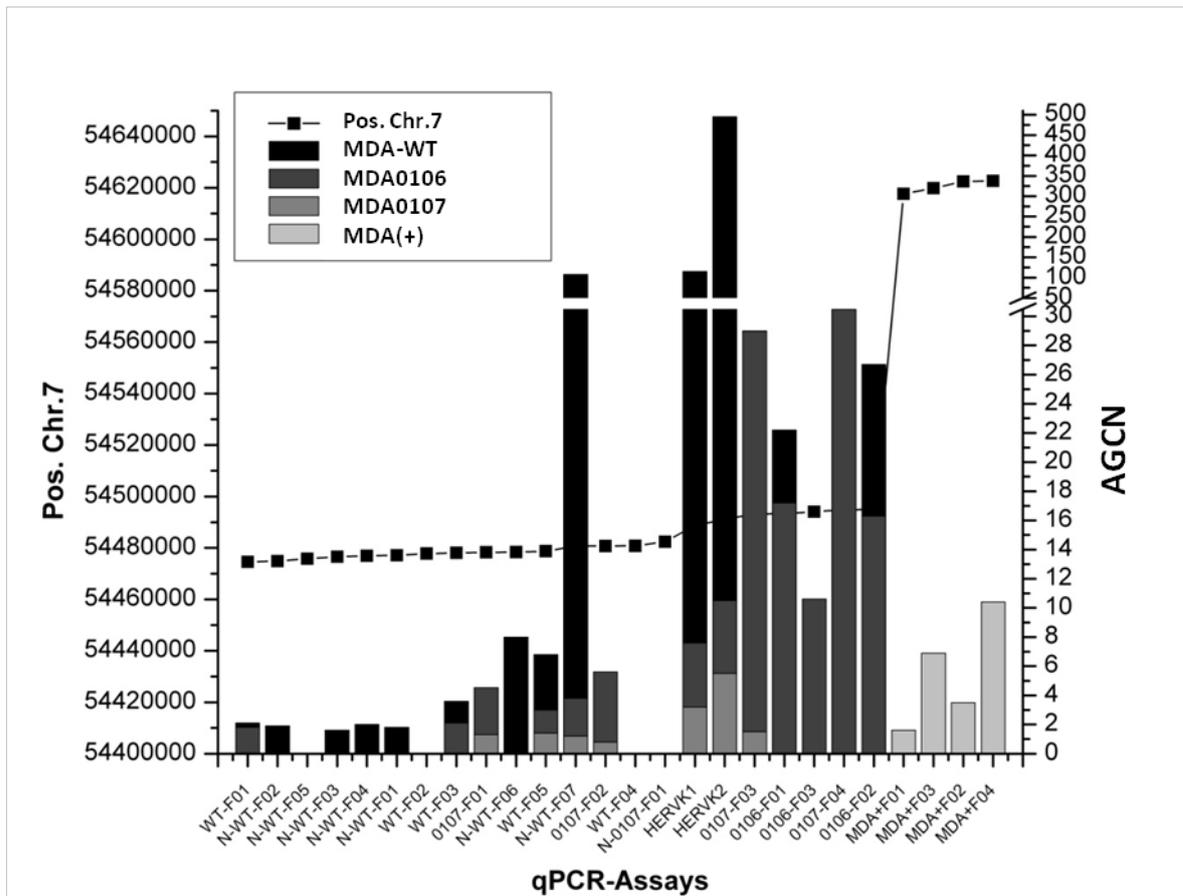


Abb. 3.20: Validierung der centromerwärtigen Endsequenz für MDA(+)

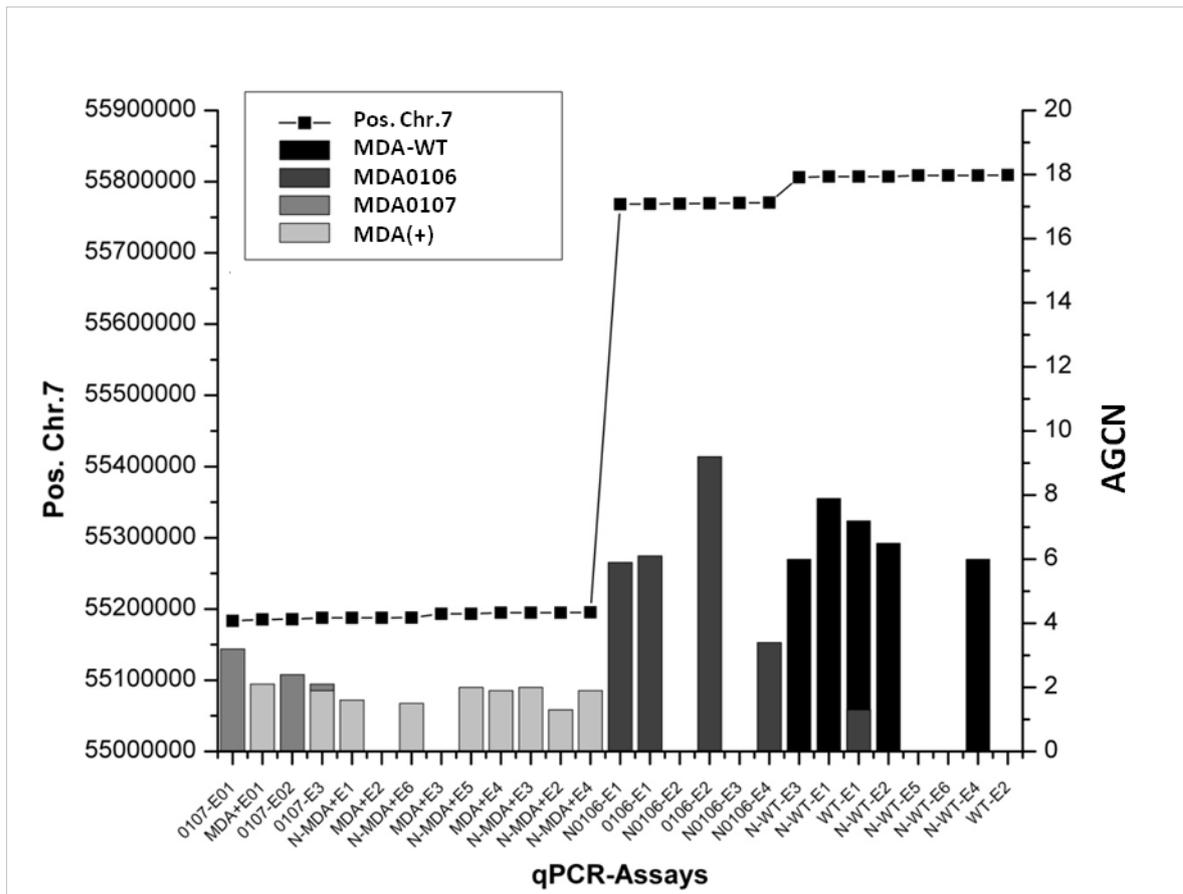
Aufgrund des Verlaufs der log<sub>2</sub>-Ratios des Plots mit einer klar definierte Kante sowie der AGCN der qPCRs wurde die Endsequenz bei ~55.187.500 festgelegt.

Die Berechnung der Start- und Endsequenzen der Amplikons ergab im Vergleich zur Validierung anhand der log<sub>2</sub>-Ratios der Plotverläufe und der AGCN der qPCRs, für die klarer definierten telomerwärtigen Regionen, Abweichung von ~2000-4000 Basen. Die Präzision der Berechnungen nahm mit kleineren log<sub>2</sub>-Ratios ab. Für die nicht klar definierten Endsequenzen wurden Abweichungen von ~2000–14000 Basen zwischen den berechneten Endsequenzen und den auf den log<sub>2</sub>-Ratios der Plotverläufe sowie AGCN der qPCRs basierenden Endsequenzen erreicht. Auch hier bedingten kleinere log<sub>2</sub>-Ratios eine Abnahme in der Präzision der Berechnung. Die Verläufe der AGCN der qPCR Assays wurden jeweils für die Start- bzw. Endregion der MDA-MB-468 Sublinien zusammengefasst und in Kombination mit der chromosomalen Position dargestellt (Abb. 3.21-3.22).



**Abb. 3.21: Zusammenfassung der AGCN der qPCRs der telomerwärtigen Startregion der Amplikons**  
Auf der horizontalen Achse wurden die Assaynamen, auf der linken vertikalen Achse die Position auf Chr.7 und auf der rechten horizontalen Achse die AGCN der qPCRs dargestellt.

Für MDA-MB-468 und die Sublinien konnte ein ähnlicher Verlauf der AGCN der qPCRs innerhalb der Startsequenzen festgestellt werden, der in der Höhe der Messwerte sowie der chromosomalen Position variierte. Dem eigentlichen langgestreckten Amplikon mit dem *EGFR*-Gen war jeweils ein Sequenzbereich mit erhöhten AGCN der qPCRs vorangestellt. Verglichen mit den AGCN für *EGFR* waren die AGCN der qPCRs in der vorgeschobenen, hochamplifizierten Region für die Sublinien um Faktor 2-4 und für MDA-MB-468 um Faktor 5-16 erhöht.



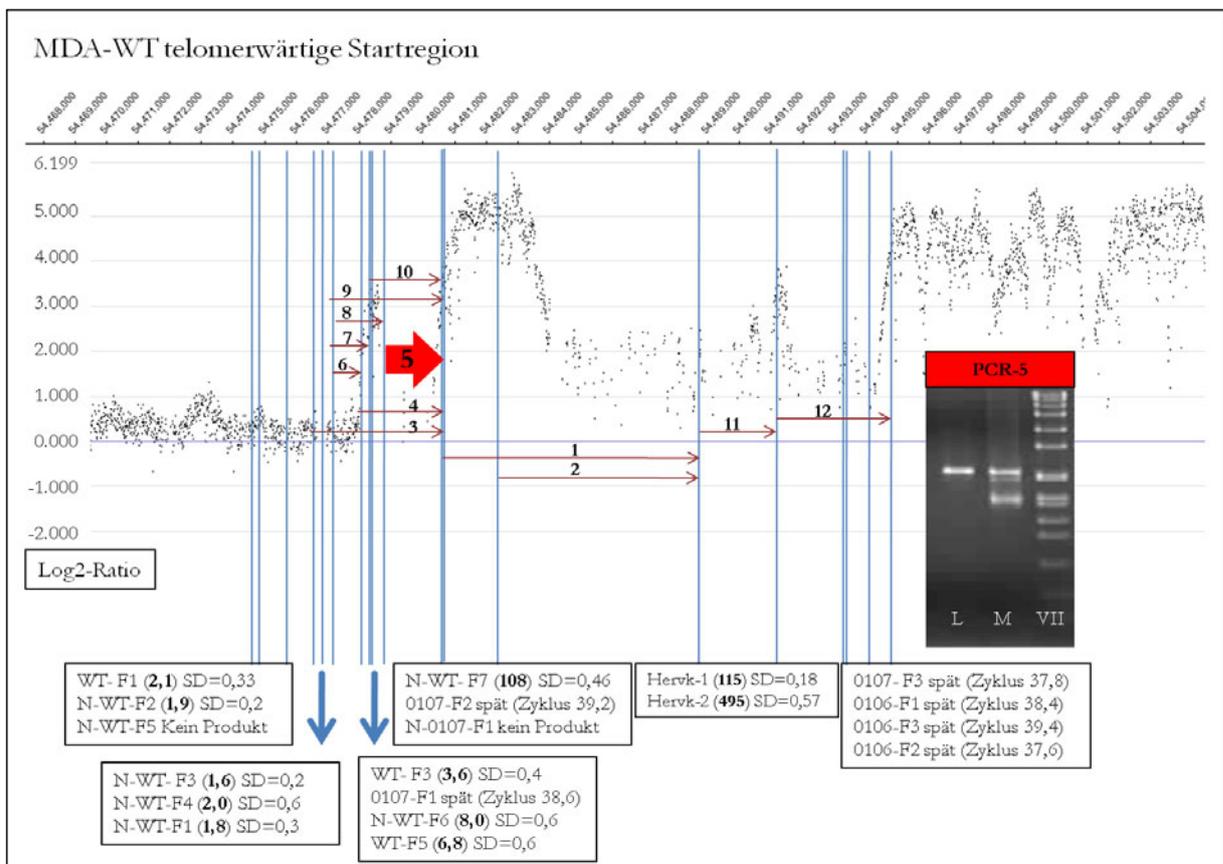
**Abb. 3.22: Zusammenfassung der AGCN der qPCRs der centromerwärtigen Endregion der Amplikons**  
 Auf der horizontalen Achse wurden die Assaynamen, auf der linken vertikalen Achse die Position auf Chr.7 und auf der rechten horizontalen Achse die AGCN der qPCRs dargestellt.

Die Verläufe der AGCN der qPCRs für die Endregionen der MDA Sublinien zeigten ein homogenes Bild. Alle Linien zeigten ein von der Höhe der Amplifikation abhängiges lineares Auslaufen der Endbereiche. Die Position der Endsequenzen musste mit stärkerer Gewichtung auf den Verlauf der log<sub>2</sub>-Ratios der Array Plots festgelegt werden.

Die telomerwärtige, durch besser definierte Grenzen charakterisierte Startregion der Amplikons wurde anschließend PCR-basiert auf Sequenzen untersucht, die von der Normalsequenz abweichend Insertionen oder Deletionen zeigten. Dies sollte Rückschlüsse auf mögliche chromosomale Umlagerungsprozesse ermöglichen.

### 3.5 PCR-basierte Detektion von Umlagerungssequenzen an der telomerwärtigen Startregion von MDA-MB-468

Um potentielle Bruchpunkte in der Startregion des Amplikons aufzufinden und zu charakterisieren, wurde der Bereich mit Long PCRs abgedeckt. Hierzu wurden Kombinationen von Forward und Reverse Primern (siehe Tab.2.12) der bestehenden qPCR Assays verwendet. Die PCRs sollten in Regionen, in denen Strangbrüche und anschließende Fusionen in gleicher Orientierung stattgefunden hatten, PCR-Fragmente mit abweichender Länge erzeugen. Es wurden für die Versuche die Enzyme Taq Polymerase (Qiagen), AmpliTaq Gold (ABI) und Phusion Polymerase (Finnzymes) verwendet. In Kombination mit der Phusion Polymerase wurde außerdem das Failsafe PCR Kit (Finnzymes) getestet, das mit 18 verschiedenen Pufferbedingungen zur PCR Optimierung genutzt wurde. Es wurden für jede Primerkombination sowohl genomische MDA-MB-468 DNA als auch Leukozyten DNA als Kontrolle verwendet. Eine schematische Darstellung der langen PCR Assays über die Startregion wurde in Abb. 3.23 dargestellt.

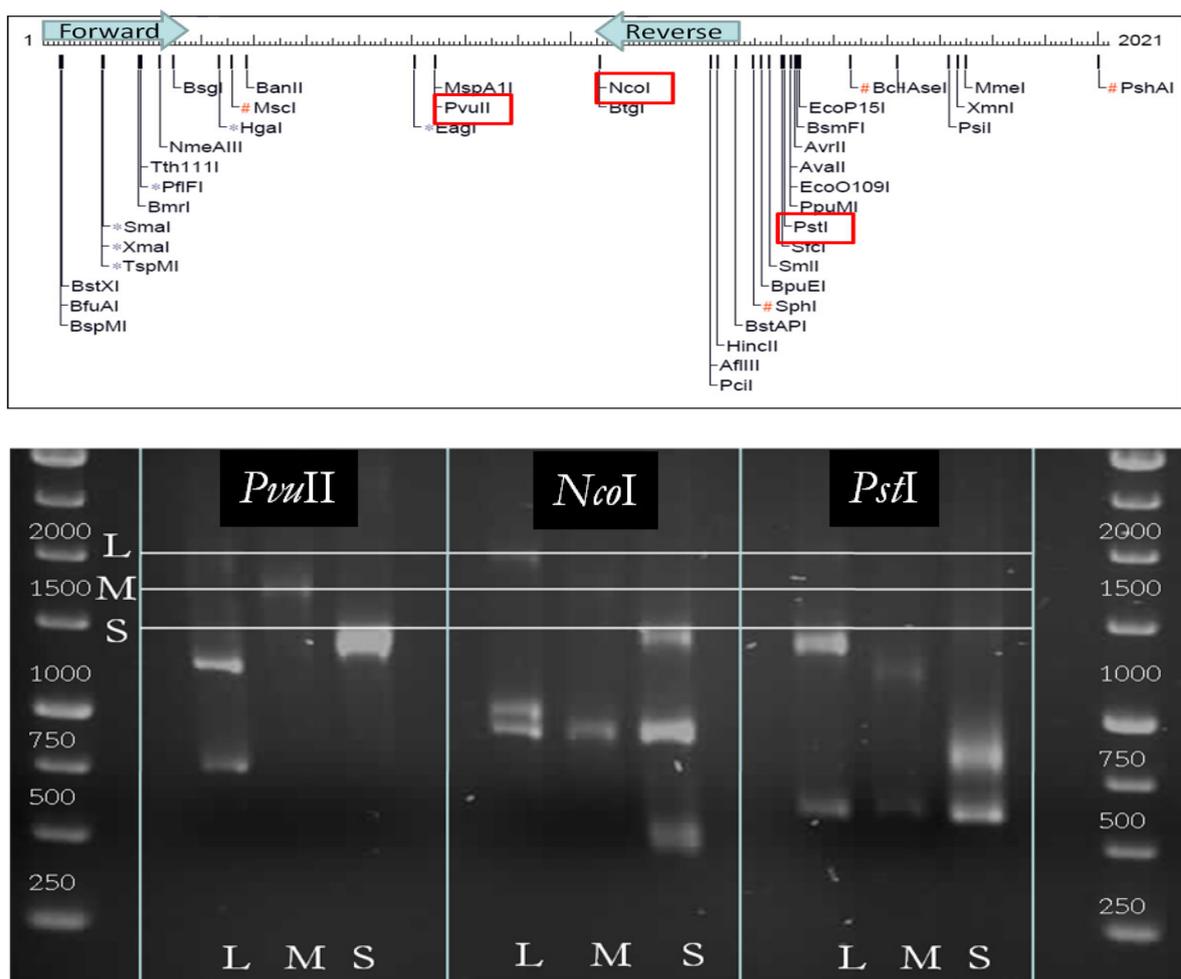


**Abb. 3.23: Ergebnisse der Long PCRs in der telomerwärtigen Startregion von MDA-MB-468** In der Abbildung werden die verschiedenen Primerkombinationen der Long PCRs durch braune Pfeile dargestellt. Das Agarosegel zeigt für MDA-WT DNA (M) im Vergleich zur Leukozyten DNA (L) differenzielle PCR Produkte für die Primerkombination 5. Marker VII (VII) wird zum Abschätzen der Fragmentlängen verwendet.

Die Long PCR mit der Primerkombination 5 erzeugte für MDA-WT DNA neben dem berechneten Fragment von 2021bps Länge noch 2 verkürzte Fragmente, deren Längen ~1700bps und ~1400bps betragen. Diese PCR überspannte genau den Bereich der Startregion, für die die Startsequenz des Amplikons festgelegt wurde (siehe 3.4). Alle anderen Primerkombinationen erzeugten für MDA-WT DNA und Leukozyten DNA identische Fragmente. Für die Primerkombinationen 1, 2, 11 und 12 konnte mit keiner Kombination der Enzyme und Pufferbedingungen ein PCR Produkt der errechneten Größe amplifiziert werden.

### 3.6 Analyse der differentiellen PCR Fragmente

Die verkürzten PCR Fragmente wurden durch Restriktionsanalysen näher charakterisiert. In Abb. 3.24 wurden die Restriktionskarte der Sequenz des PCR Fragments sowie ein Agarosegel mit den Resultaten der Restriktionsanalyse dargestellt.



**Abb.3. 24: Restriktionsanalyse der PCR-5 Fragmente aus MDA-MB-468 DNA** Für die Analyse wurden die in der Restriktionskarte des Fragments rot umrahmten Enzyme *PvuII*, *NcoI* und *PstI* verwendet. Die Positionen auf dem Agarosegel wurden für das 2021bps lange Normalfragment mit L, das mittlere ~1700bps Fragment mit M und das kurze ~1400bps Fragment mit S (short) markiert.

Durch die Restriktionsenzyme *NcoI* und *PstI* wurde das PCR Fragment jeweils vollständig geschnitten. Für die Restriktion mit *PvuII* war nur ein Verdau des Normalfragmentes zu erkennen. Das Fehlen der *PvuII* Schnittstelle in den beiden differentiellen PCR Fragmenten, schränkte den Bereich, in dem die Verkürzung lokalisiert sein musste, auf ~1060bps ein. Dieser Bereich wurde durch Sequenzierung charakterisiert. Als Sequenzierungsprimer wurden der Forward Primer der PCR5 Kombination, sowie ein neu designter Primer mit der Sequenz CATGGTCGCTGTCTCTTCG direkt 3'-wärts von der *NcoI*-Schnittstelle verwendet. In Abb. 3.25 wurde der Sequenzabschnitt mit den Sequenzierungsergebnissen dargestellt.

```

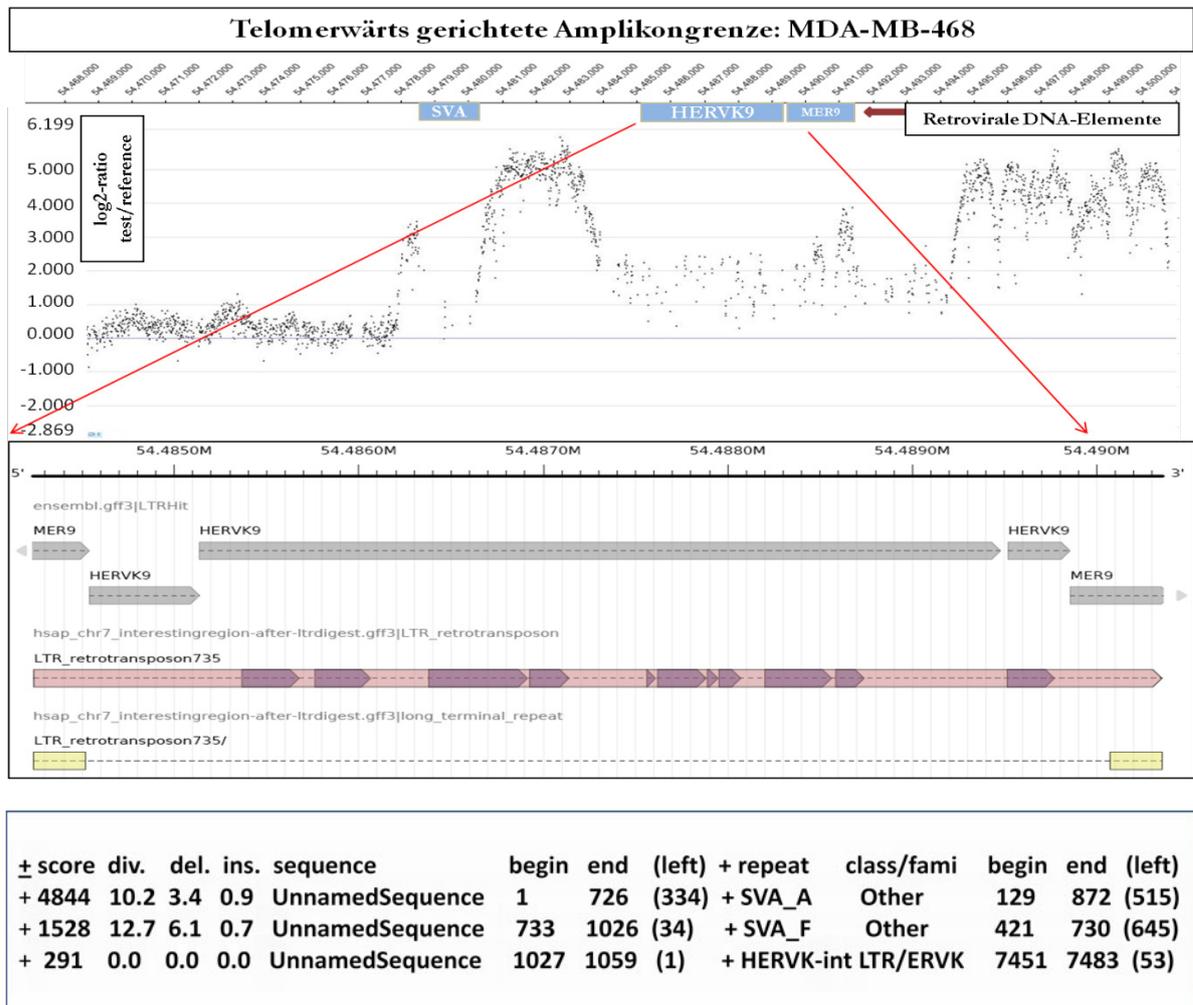
GGCTCCCATGATTCTCCTGCTTCGACCTGCCAAGTGCCTGGGATTCCAGGCACACGCCGCCACGCCTGAC
TGGTTTTTTGTATTTTTGGTGGAGACGGGGTTTTGCGGTGTTGCCCGGGCTGGTCTCCAGCTCCTGGCCTC
GGGTGGTCTGCCTGCCTCGACCTCCCGAGGTACTGGGATTGCAGACGGAGTCTCGTTCACTTAATGCTCA
GTGTTGCCCAGGCTGGAGTGCAGTGGCGTGATCTCGGCTGGCTACAACCTCCACCTCCAGCCGCCTGCC
TTGGCCTCCCAAAGTGCTAAGATTACAGCCTCTGCCCGCCCGCCACCCCGTCTAGGAAGTGAGCAGCGTC
TCTGCCTGGCCACCCATCGTCTGGGATGTGAGGAGCCCTCTGCCAGCCGCCCATCTGGGAAGTGAGG
AGTGCCTCTGCCAGCCACCACCCATCTAGGAAGTGAGGAGCATCTCTGCCTGGCTGCCCATCGTCTGG
GTTGTGAGGAGCGCCTCTGCCAGCCGCCCGTCTGGGAGGAAGTGAGGAGCACCTCCGCCCGGCTGCC
CGAATGGGAAGTGAGGAGCGCCTCTGCCTGGCCGCCCATCTGGGAAGTGAGGAGCGCCTCTGCCCGGTC
TCCCCCCCCTAGGAAGTAAGGAGTGTCTCTGCCCGACCGCCACTGTCTGGGATGTGAGGAGCACCTC
TGCCCGGCCGCCCATCTGGGAAGTGAGGAGCGCCTCTGCCAGCTGCCCGTCTGGGATGTGAGGAGCG
CCTCTGCCCGGCCGCCCATCGTCTGGGAAGTGAGGAGCGCCTCTGCCCGGCCACCCGTCTGGGAAGTGA
GGAGTGCCTGTGCCTGGCCGCCCATCTGGGTAGTGAGGAGTGTCTCTGCCTGTCCGCCATCGTCTGGG
AGGTGAGGAGCGCCTCTTCCAGCCGCCCATCTCTGGGATGTGAGGAGCGCCTCTGCCAGCCACCCTG
TCTGGGAAGTGAGGAGCACCTCTGCATGGCCGCCCATCTGGGAGGTGTACCCAACAGCTCCGAAGAGAC
AGCGACCATG

```

Abb. 3.25: Ergebnisse der Sequenzierung der M- und S-Fragmente von PCR-5 Dargestellt sind die **flankierenden Primersequenzen**, die ersten *nicht lesbaren Basen* der einzelnen Sequenzierungen, die **lesbaren Bereiche**, und der Bereich, in dem die Sequenzierungen aus 3'- und 5'-Richtung jeweils in jeder der 3fachen Wiederholungen abbrachen. Teile dieser Region wurden unterstrichen dargestellt, um eine bessere Darstellung der repetitiven DNA Bereiche zu ermöglichen.

Die in 3fachen Wiederholungen durchgeführten Sequenzierungen der beiden verkürzten Fragmente erlaubten nur eingeschränkte Leseweiten von 361bps mit dem Forward- und 117bps mit dem Reverse Primer. Der gesamte in Abb.3.25 grau dargestellte Bereich konnte nicht erfasst werden. Dies könnte mit der repetitiven Struktur der DNA zusammenhängen, die für die Polymerase des verwendeten BigDye nicht amplifizierbar war. An den Grenzen dieser Bereiche fielen die Fluoreszenzsignale der Sequenzanalysen abrupt auf das Niveau des Hintergrundrauschens ab.

Dieser Sequenzbereich in der Startregion von MDA-WT wurde anschließend durch die am Institut für Bioinformatik Hamburg entwickelte Software LTR-Digest untersucht. LTR-Digest analysierte DNA Sequenzen auf der Basis von Datenbanksequenzen (RepBase, NCBI) auf unbekannte retrovirale DNA Elemente (Abb. 3.26).



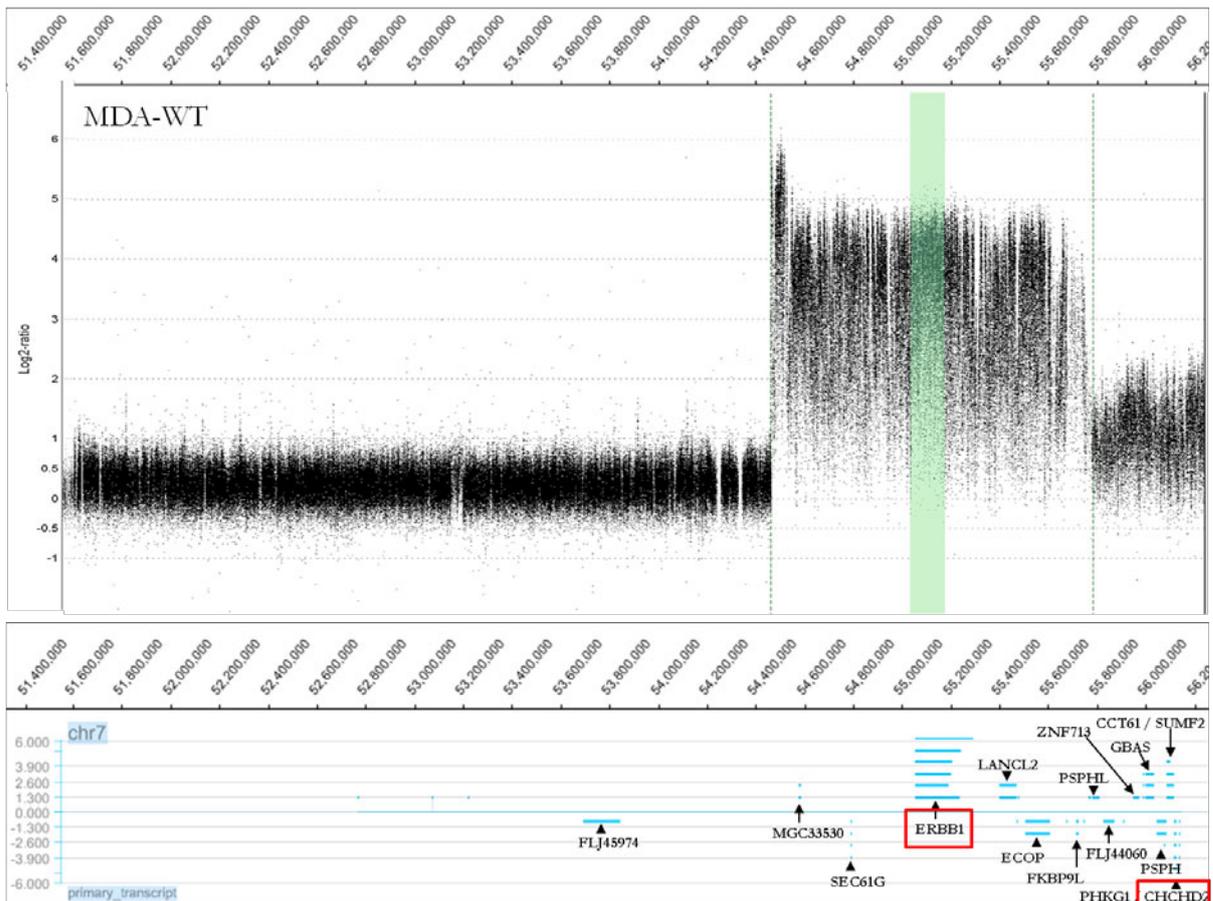
**Abb. 3.26: Ergebnisse der LTR-Digest Analyse der verkürzten PCR-5 Fragmente** Dargestellt ist die telomerwärtige Startsequenz von MDA-WT mit den durch LTR-Digest detektierten retroviralen Elementen.

Die Analyse der telomerwärtigen Startregion von MDA-WT ergab ein HERV-K Element, das als einziges HERV-K Element auf Chr.7 eine vollständige, aber mosaikartig unterbrochene Übereinstimmung mit der HERV-K Ursequenz besaß. Die flankierenden MER9 Elemente bildeten die langen terminalen Repeats (LTRs) des retroviralen HERV-K Elements. Die Suche nach verwandten retroviralen DNA Elementen ergaben für die HERV-K Sequenz hohe Homologien mit dem ~4800bps entfernt liegenden SVA\_A Element von >500bps. Innerhalb des retroviralen SVA\_A Elements waren die differentiellen PCR Fragmente der Primerkombination 5 lokalisiert.

Abschließend wurde der Einfluss der Genamplifikation von *EGFR* auf die Transkription untersucht. Dies sollte Aufschluss darüber geben, inwieweit eine direkte Abhängigkeit der Expression von *EGFR* zu der Höhe der genomischen Amplifikation bestand und die Ergebnisse der Anti-EGFR gerichteten FACS basierten Analysen mit anschließender qPCR (siehe Kapitel 3.1 K. Agelopoulos) bestätigen.

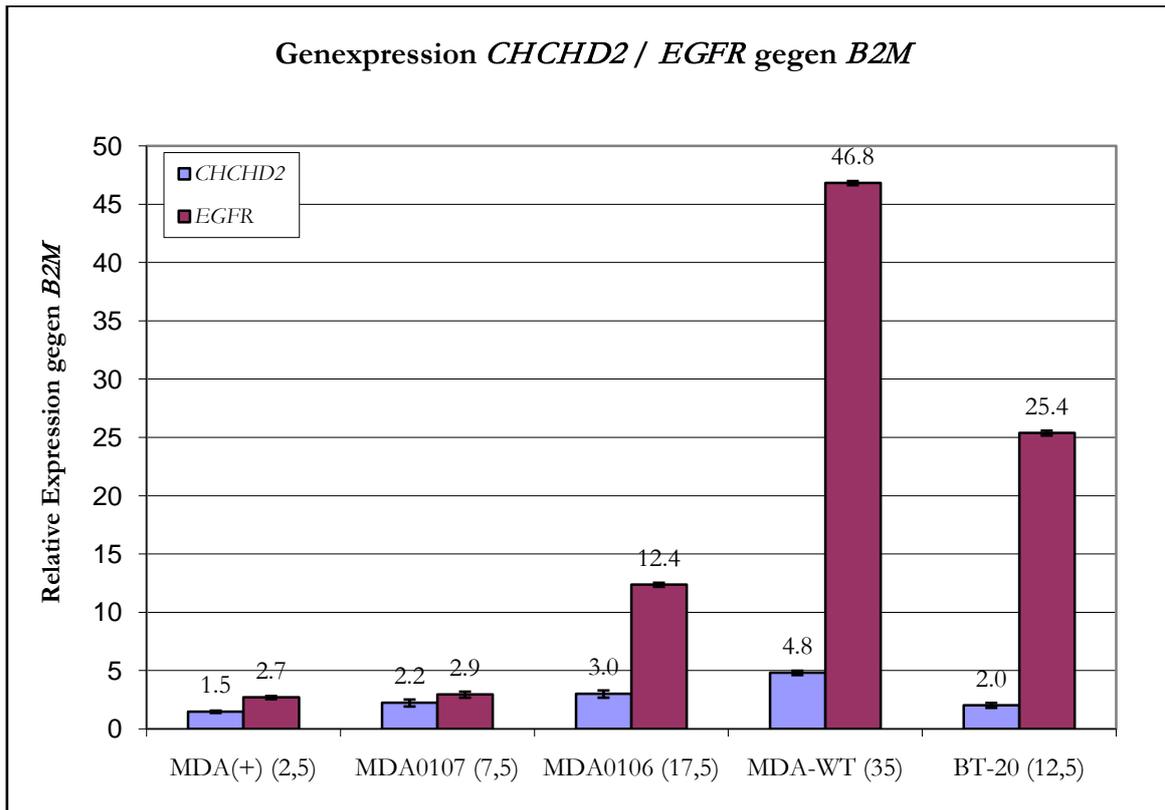
### 3.7 Expressionsanalyse von *EGFR* und *CHCHD2* in MDA-MB-468 Sublinien

Für die Analyse wurden käufliche TaqMan Assays (Applied Biosystems) für *EGFR* und das Gen *CHCHD2* (Coiled Helix Coiled Helix like Domain 2), welches im centromerwärtigen Bereich des Hauptamplikons mit log<sub>2</sub>-Ratios von ~2 lokalisiert war, verwendet (siehe Abb. 3.27).



**Abb. 3.27: Lokalisierung bekannter Gene auf der Chr.7 Region der Array-CGH** Dargestellt wurde der Amplifikationsplot für MDA-WT mit der log<sub>2</sub>-Ratio auf der vertikalen Achse und der chromosomalen Position auf der horizontalen Achse. Im unteren Bereich wurden, durch blaue Linien markiert, bekannte Gentranskripte bzw. deren Isoformen eingezeichnet. *ERBB1* (*EGFR*) und *CHCHD2* wurden durch rote Kästen hervorgehoben.

Als Referenzgen wurde *B2M* (Beta2-Mikroglobulin) gewählt, welches sich als stabiles Housekeeping-Gen für die Quantifizierung von Expressionen erwiesen hatte. Für die Analyse wurden aus den Zelllinien MDA-MB-468 und deren Sublinien und BT-20 Gesamt-mRNA isoliert und parallel zusammen mit universeller humaner Referenz-RNA, die zur Normalisierung verwendet wurde, in cDNA umgeschrieben. Die Ergebnisse der relativen Quantifizierung wurden in Abb. 3.28 dargestellt.



**Abb. 3.28: Relative Expressionen von *EGFR* und *CHCHD2* gegen *B2M* der MDA Sublinien** Auf der horizontalen Achse wurden die jeweiligen MDA Sublinien mit AGCN für *EGFR* (Zahl in Klammern) und auf der vertikalen Achse die relativen Expressionen von *EGFR* bzw. *CHCHD2* gegen *B2M* dargestellt.

Für die MDA Sublinien konnte für *EGFR* und *CHCHD2* eine steigende Expression mit steigender AGCN für *EGFR* bzw. *CHCHD2* gemessen werden. Die *EGFR*-Expression schien, im Vergleich zu der eher linear ansteigenden AGCN für *EGFR* der verwendeten Sublinien, exponentiell anzusteigen. Die log<sub>2</sub>-Ratios der Array-CGH im Hauptamplikon von 2-6 für den *EGFR*-Lokus und die log<sub>2</sub>-Ratios von 1-2 für den *CHCHD2* Locus spiegelten die Höhe der relativen Expression von *EGFR* mit Werten von 2,7-46,8 und *CHCHD2* mit Werten von 1,5-4,8 wieder. In der Zelllinie BT-20 war die Expression von *EGFR* bezogen auf die AGCN für *EGFR* im Vergleich zu MDA0107 und MDA0106 um Faktor ~2-3 erhöht. Die Expression von *CHCHD2* in BT-20 lag etwa auf dem Niveau von MDA0107.

Um den Amplifikationsmechanismus von *EGFR* in Modelzelllinie MDA-MB-468 zu untersuchen, wurde ein Verfahren entwickelt, das die genetische Charakterisierung einzelner Zellen ermöglichte. Die Zelllinie MDA-MB-468 war heterogen für den *EGFR*-Lokus. Das Verfahren sollte Aufschluss über den *EGFR*-Status einzelner Zellen der Modelzelllinie geben. Außerdem sollte das Verfahren zur Untersuchung des *EGFR*-Status von DTC in Blut oder KM von Patienten verwendet werden, um einen möglichen Zusammenhang zwischen Aberrationen im *EGFR*-Lokus und dem Prozess der Metastasierung aufzuklären. Das zentrale Element dieses Verfahrens bildete die ***Whole-Genome-Amplification*** (WGA), die für die Amplifikation der gesamten genomischen DNA verwendet wurde.

### **3.8 *Whole Genome Amplification* (WGA) zur Bestimmung der *EGFR*-Gendosis in Einzelzellen**

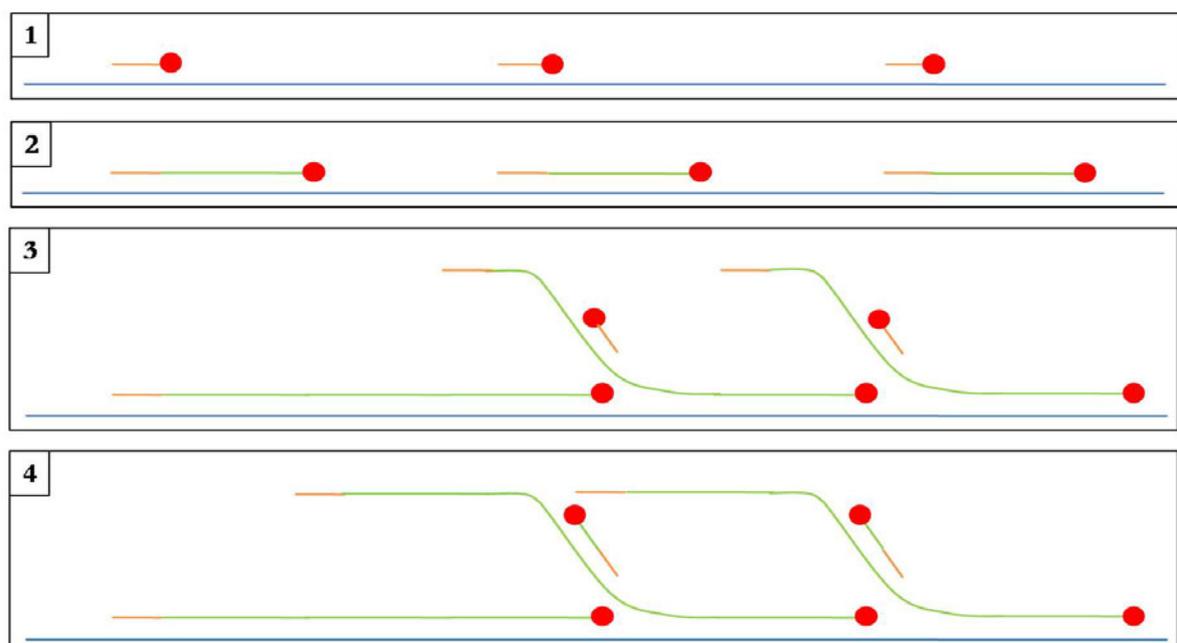
Die verschiedenen Strategien, die bei der WGA verfolgt, wurden ließen sich in zwei unterschiedliche Gruppen einteilen: die PCR-basierten Verfahren und die „Multi-Displacement-Amplification“ (MDA). Die erste Gruppe der Verfahren spaltete sich in drei Untergruppen auf, wobei jedes der Verfahren auf einer Taq Polymerase basierte:

1. Random PEP (Primer Extension Preamplification): Die WGA basierte auf der Amplifikation durch zufällige 15 Basen lange Oligonukleotide. Die Oligonukleotide sollten aufgrund der zufälligen Sequenzen mit einer ähnlichen statistischen Wahrscheinlichkeit binden und so zu einer Amplifikation des gesamten Genoms führen. Mit dieser Methode wurde eine maximale Amplifikation des Ausgangsmaterials um Faktor 30 erreicht (148).
2. DOP-PCR (Degenerate Oligonucleotide Primed): Dieses Verfahren verwendete ein Primersystem, bei dem der Primer am 3'-Ende eine definierte Sequenz trug, die in ca.  $10^6$  Kopien auf dem menschlichen Genom vorhanden war, eine weitere definierte Sequenz am 5'-Ende, welche eine XhoI Schnittstelle trug und im mittleren Teil eine zufällige 6 Basen lange Sequenz. In den ersten 6-8 Zyklen der PCR, die mit niedriger Annealing Temperatur abliefen, lagerten die definierten 3'-Enden an die DNA an. Die zufälligen Hexamer Sequenzen banden dann ebenfalls, jedoch unvollständig an die DNA. In den finalen 25 Zyklen wurde die Annealing Temperatur erhöht und somit die Spezifität der PCR auf die bereits vom dem Primer markierten Bereiche erhöht (149).

3. Die Verfahren der dritten Gruppe basierten auf jeweils drei Schritten. Im ersten Schritt wurde die DNA durch Restriktionsverdau (150-152), zufälliges Scheren der DNA (153) oder „nick translation“ (154) fragmentiert. Im zweiten Schritt wurde durch Ligation von Linkeradaptoren an die fragmentierte DNA eine sogenannte DNA-Bibliothek erstellt. Die Amplifikation der DNA aus der Bibliothek in Schritt drei wurde durch universelle Primer mit der Komplementärsequenz der Linkeradapter erreicht.

Die Verfahren erreichten im Mittel Fragmentlängen  $<3\text{kb}$  und zeigten eine Fehlerrate von  $3 \times 10^{-5}$ . Desweiteren wurde keine vollständige Abdeckung der genomischen DNA erreicht und die Amplifikation verlief nicht mit gleichen Effizienzen für alle genomischen Loci. Für PEP wurden  $10^2$ - $10^4$ -fache, für die DOP-PCR  $10^3$ - $10^6$ -fache Abweichungen durch ungleiche Effizienzen der WGA in unterschiedlichen Loci beschrieben.

Die MDA verwendete Exonuklease-resistente, zufällige Hexamer Primer. Diese trugen jeweils eine Thiophosphat Modifikation zwischen den drei terminalen Nukleotiden. Die Reaktion verlief über einen Zeitraum von 16h isothermal bei einer Temperatur von  $30^\circ\text{C}$ . Durch die Strang-verdrängenden Eigenschaften der Polymerase in Kombination mit den Hexamer Primern wurde eine kontinuierlich fortlaufende Strangsynthese erreicht. Das Prinzip der MDA wurde in Abb. 3.29 schematisch dargestellt.



**Abb. 3.29: Prinzip der Multi Displacement Amplifikation:** (1) Random Hexamere (orange Linie) binden an die Einzelstrang DNA (blaue Linie); (2) Die phi29 DNA Polymerase (Roter Kreis) verlängert die Primer, bis sie neusynthetisierte Doppelstrang DNA erreicht (grüne Linie); (3) Das Enzym verdrängt den neusynthetisierten Strang während es mit der eigenen Synthese fortfährt. Random Hexamer Primer binden an neusynthetisierte Stränge; (4) Die Neusynthese beginnt an den neuen Strängen, es bildet sich eine hyperverzweigte Struktur aus.

Die phi29 Polymerase besaß eine extrem hohe Effizienz und generierte im Mittel eine 300.000-fache Amplifikation des Startmaterials mit einer durchschnittlichen Produktgröße >10kb. Die Fehlerrate der „*Proofreading*“ Polymerase lag bei  $10^{-6}$ - $10^{-7}$ . Es wurde eine maximale 3-fache Amplifikationsabweichung beschrieben (155) und somit eine homogenere Amplifikation der genomischen Loci erreicht. Es wurde ferner gezeigt, dass MDA-amplifizierte genomische DNA als Ausgangsmaterial für gängige genetische Analysen wie „SNP“ (Single Nucleotide Polymorphisms), Genotypisierungen, „chromosomal painting“, Southern Blotting, subcloning und Sequenzierungen (156) verwendet werden konnte.

Um aus Einzelzellen mit einem durchschnittliches Genomgewicht von 6-7pg ausreichend Material für eine Vielzahl von genetischen Analysen zu erzeugen, das sowohl auf den Ebenen der Sequenz, der genomweiten Abdeckung und gleichmäßigen Amplifikation aller Loci möglichst nah am Ausgangsmaterial war, wurde die MDA verwendet.

Für das MDA-basierte Genomphi-Kit wurde vom Hersteller eine minimale Menge Ausgangsmaterial von 1ng DNA empfohlen. Deshalb wurde eine serielle Verdünnungsreihe mit genomischer DNA als Ausgangsmaterial durchgeführt, um die generelle Funktionalität der WGA im pg-Bereich zu überprüfen. Die durch die WGA amplifizierte DNA wurde mit einer Multiplex PCR überprüft. In einer Multiplex PCR wurden mehrere genomische Loci in einer PCR Reaktion amplifiziert. Dadurch wurde zum einen die Amplifikation der drei Loci in der WGA überprüft und zum anderen zwischen spezifischem und unspezifischem WGA Produkt unterschieden. Dies war erforderlich, da in der Negativkontrolle der WGA immer eine vergleichbar hohe Menge hochmolekularer DNA gebildet wurde, wie bei der Amplifikation der eigentlichen Proben. Die in der Negativkontrolle gebildete DNA wurde von der Polymerase durch Verlängerung der Primer erzeugt und war nicht PCR-amplifizierbar. Die Ergebnisse der seriellen Verdünnungsreihe wurden in Tab. 3.4 dargestellt.

Template DNA (ng)	WGA amplifizierte DNA ( $\mu\text{g}$ )	Multiplex PCR 1
10	3,8	3 / 3
1	3,6	3 / 3
0,5	3,7	3 / 3
0,25	3,5	3 / 3
0,125	3,5	3 / 3
0,063	3,6	3 / 3
0,032	3,4	3 / 3
0,016	3,4	1 / 3
0,008	3,2	0 / 3
NTC	2,8	0 / 3

**Tab. 3.4: Serielle Template-Verdünnungsreihe für das Genomphi-Kit** In der Tabelle sind die Mengen der eingesetzten Template DNA, der WGA amplifizierten DNA sowie die Anzahl der genomischen Loci, die durch die Multiplex PCR dargestellt werden konnten, aufgelistet.

Im Bereich von 1ng bis 32pg ( $\sim 5$  Zellen) gab es geringe Differenzen in der gebildeten Produktmenge. Es konnten in diesem Bereich alle genomischen Loci in der WGA amplifizierten DNA durch die Multiplex PCR nachgewiesen werden. Bei 16 pg Template wurde nur noch ein, bei 8pg Template keine Multiplex Loci in der WGA amplifizierten DNA nachgewiesen. Die in der NTC gebildete Produktmenge lag bei 2,8 $\mu\text{g}$ .

Die folgende Validierung der WGA an Mikrosatelliten wurde, basierend auf der Verdünnungsreihe, mit 30pg Template durchgeführt. Es war davon auszugehen, dass ab einer Konzentration von  $<30\text{pg}/\mu\text{l}$  die statistische Wahrscheinlichkeit, DNA durch Bindung an die Pipettenspitze, das Reaktionsgefäß oder die inhomogene Verteilung des Makromoleküls in Lösung zu verlieren, zu groß wurde, um reproduzierbare Ergebnisse zu erzielen.

Vor der eigentlichen Validierung wurden die Mikrosatelliten Primer zu Multiplex PCRs zusammengefasst und deren Protokolle optimiert. Anschließend wurden die PCR-Fragmentlängen der Zelllinien A431, HBL100, MCF-7, MDA-MB-468 und SK-OV-3 für alle Multiplex PCRs bestimmt, um für die Validierung ein DNA-Set zusammenstellen zu können, das für jeden Locus heterozygot war.

### 3.9 Die Validierung der WGA mittels eines Mikrosatelliten PCR Panels

Als Mikrosatelliten oder SSR (Simple Sequence Repeats) werden kurze, repetitive nicht kodierende DNA-Sequenzen bezeichnet, die sich in dem Genom eines Organismus häufig wiederholen. Bei den SSRs handelt es sich um einfache 2-64 Nukleotide lange Elemente, die zwischen 10 und 100mal wiederholt sein können. Den häufigsten Mikrosatelliten stellt der Typ  $(CA)_n$  dar, welcher  $\sim 0,5\%$  des humanen Genoms ausmacht. Aufgrund ihrer Häufigkeit und der hohen Wahrscheinlichkeit auf polymorphe Längen der Mikrosatelliten, auch zwischen den beiden Allelen eines Organismus, werden diese häufig zur Genanalyse verwendet. Eine weitere Anwendung, neben dem klassischen Vaterschaftstest, stellte der Vergleich von Normalgewebe und Primärtumor, bzw. Primärtumor und Metastase dar. Es wurden hierzu bevorzugt Mikrosatelliten verwendet, die in der Nähe von bekannten potentiellen Onkogenen lokalisiert waren. Durch Mikrosatelliten Analysen können allelische Imbalancen (AIs) detektiert werden, welche Hinweise auf DNA Zugewinne bzw. Verluste einzelner Allele gaben. Die Lokalisierungen bzw. Genassoziationen des in dieser Arbeit verwendeten Mikrosatelliten Panels wurden in Tab. 3.5 aufgelistet.

Name	Zielgen / Region	Genomische Lokalisation
D7S522	<i>Caveolin 1+2</i>	CHR7-q31.2
D8S258	8p22	CHR8-p21.3
D16S400	16q22.2-23.1	CHR16-q21
NEFL	8p21	CHR8-p21.2
D13S153	<i>Rb1</i>	CHR13-q14.2
D17S855	<i>BRCA1</i>	CHR17-q21.31
D10S541	<i>PTEN</i>	CHR10-q23.31
D16S402	16q24.2	CHR16-q23.3
D16S422	16q24.2	CHR16-q23.3
P53CA	<i>TP53</i>	CHR17-p13.1
CASSR1	<i>EGFR</i>	CHR7-p11.2
CASSR3	<i>EGFR</i>	CHR7-p11.2
CASSR4	<i>EGFR</i>	CHR7-p11.2
D7S2429	<i>EGFR</i>	CHR7-q11.21
D7S494	<i>EGFR</i>	CHR7-p11.1
D7S499	<i>EGFR</i>	CHR7-p11.2
D7S2467	<i>EGFR</i>	CHR7-p12.1
CASSR2	<i>EGFR</i>	CHR7-p11.2
D7S2550	<i>EGFR</i>	CHR7-p11.2
CASSR6	<i>EGFR</i>	CHR7-p11.2

Tab. 3.5: Lokalisierung der verwendeten Mikrosatelliten

Das Panel umfasste 10 Primerpaare, die Chr.7p11-12 lokalisiert waren und zur Charakterisierung der Region um *EGFR* dienten. Zehn weitere Primerpaare verteilten sich auf weitere Chromosomen und beinhalten Regionen nahe der Gene *BRCA1*, *TP53*, *Rb1*, *Caveolin1+2* und *PTEN*.

### 3.10 Multiplex PCR-Protokoll Optimierung für die Mikrosatelliten Primer

Mit dem Ziel von Zeit-, Kosten- und Proben DNA Ersparnis wurden die PCR Protokolle darauf optimiert, als Multiplex PCR zu laufen. Als Erstes wurde eine Gradienten-PCR (56-62°C) mit jedem Primerpaar durchgeführt, um die optimale Anlagerungstemperatur zu bestimmen. Anschließend wurden PCR Additive hinsichtlich ihres Einflusses auf die Bildung von Primerdimeren, unspezifischer Primerbindung und Produktausbeute getestet. Die Primerpaare wurden dann, gemäß ihrer errechneten Produktgrößen und optimalen Anlagerungstemperatur, in 2er bis 3er Gruppen zusammengefasst. Der Mindestunterschied der Produktgrößen innerhalb jeder Multiplex PCR betrug 10 Basen. Abschließend wurden die Effizienzen der Mikrosatelliten PCRs innerhalb einer Multiplex PCR durch das Anpassen der Primerkonzentrationen angeglichen. Die optimierten Multiplex-PCR Bedingungen wurden in Tab. 3.6 dargestellt.

Gruppe	Primerpaar	Produktgröße	Primerkonz.	Anlagerungstemp.	TMAC	Q-Solution
Multiplex 1	D7S522	121	200µM	56°C	200mM	Kein Effekt
	D8S258	229	100µM			
	D16S400	179	100µM			
Multiplex 2	NEFL	109	100µM			
	D13S153	194	100µM			
	D17S855	151	100µM			
Multiplex 3	D10S541	161	100µM			
	D16S402	112	150µM			
	D16S422	122	250µM			
Multiplex 4	P53CA	171	50µM			
	CASSR1	125	100µM			
Multiplex 5	CASSR3	126	100µM	62°C	200mM	Kein Effekt
	CASSR4	197	100µM			
	D7S2429	218	100µM			
Multiplex 6	D7S494	117	100µM			
	D7S499	216	100µM			
	D7S2467	177	100µM			
Multiplex 7	CASSR2	169	100µM			
	D7S2550	255	100µM			
	CASSR6	158	100µM			

Tab. 3.6: Optimierte Multiplex-PCR Bedingungen

Die 20 Primerpaare konnten zu 7 Multiplexen gruppiert werden, die auf 2 Anlagerungstemperaturen eingestellt wurden. Das PCR Additiv TMAC zeigte eine Reduktion von unspezifischer Primerbindung, wohingegen Q-Solution (Qiagen) keinen sichtbaren Effekt hatte. CASSR6 konnte aufgrund von starker Dimerbildung nicht in Multiplex 7 integriert werden.

### 3.11 Fragmentlängenbestimmung in der DNA von Tumorzelllinien

Es wurden die Fragmentlängen aller Mikrosatelliten aus dem Panel für fünf Tumorzelllinien bestimmt, mit dem Ziel, für jeden Mikrosatelliten einen heterozygoten Locus zu finden, anhand dessen die Validierung der WGA gegen genomische DNA durchgeführt werden konnte. Die Ergebnisse der Fragmentlängenanalyse wurden in Tab. 3.7 aufgelistet.

Gruppe	Primerpaar	Produktgröße	A431	HBL100	MCF-7	MDA-MB-468	SK-OV-3
Multiplex 1	D7S522	121	Deletiert	121	121 / 123	117	123
	D8S258	229	225	223 / 225	225 / 227	209	225 / 229
	D16S400	179	167	167 / 177	175 / 177	173	175 / 177
Multiplex 2	NEFL	109	102	94	90	Deletiert	90
	D13S153	194	176	182	146 / 176	Deletiert	172
	D17S855	151	137 / 139	141 / 149	147	151	143
Multiplex 3	D10S541	161	155	159	165	155	165
	D16S402	112	108 / 116	116 / 118	116	108 / 110	114
	D16S422	122	121	119 / 121	123	119 / 121	103 / 123
Multiplex 4	P53CA	171	167	165 / 167	183	171	167
	CASSR1	125	114	118	124	116	112
Multiplex 5	CASSR3	126	121	121	119	119	119
	CASSR4	197	190 / 192	184	188	196	194 / 196
	D7S2429	218	221	219	219	215	219
Multiplex 6	D7S494	117	107	107	125	107	121
	D7S499	216	212 / 214	210 / 212	214 / 216	208 / 210	210 / 212
	D7S2467	177	170 / 172	170 / 172	172 / 174	168 / 170	168 / 170
Multiplex 7	CASSR2	169	162 / 166	166 / 168	166 / 168	166 / 168	166 / 168
	D7S2550	255	267 / 269	277 / 279	273 / 275	281 / 283	277 / 279
	CASSR6	158	155	153 / 155	153	155	153

**Tab. 3.7: Fragmentlängenbestimmung der Mikrosatelliten für 5 Tumorzelllinien** In der Tabelle wurden die Fragmentlängen der einzelnen Primerpaare in Nukleotiden für die 5 Zelllinien aufgelistet. Die auf der Basis von UCSC berechnete mittlere Fragmentlänge der Mikrosatelliten wurde ebenfalls aufgelistet. Die für die Validierung der WGA verwendeten Zelllinien bzw. Zellliniengemische wurden grau hinterlegt.

Unter den getesteten Linien befanden sich mit der Multiplex-1 für MCF-7 und Multiplex-7 für HBL100 nur zwei Ansätze, die für alle Primer der jeweiligen Multiplex heterozygot waren. Es mussten deshalb Gemische der Zelllinien A431/HBL100 und HBL100/MCF-7 im Verhältnis 1:1 hergestellt werden, um heterozygoten Ausgangsmaterial für die anderen Multiplex PCRs zu schaffen. Von den 100 bestimmten Mikrosatelliten waren 40 % heterozygot, 57% homozygot und 3% nicht durch PCR amplifizierbar.

### 3.12 Validierung der WGA mit 30pg Ausgangs-DNA durch anschließende Mikrosatelliten PCR

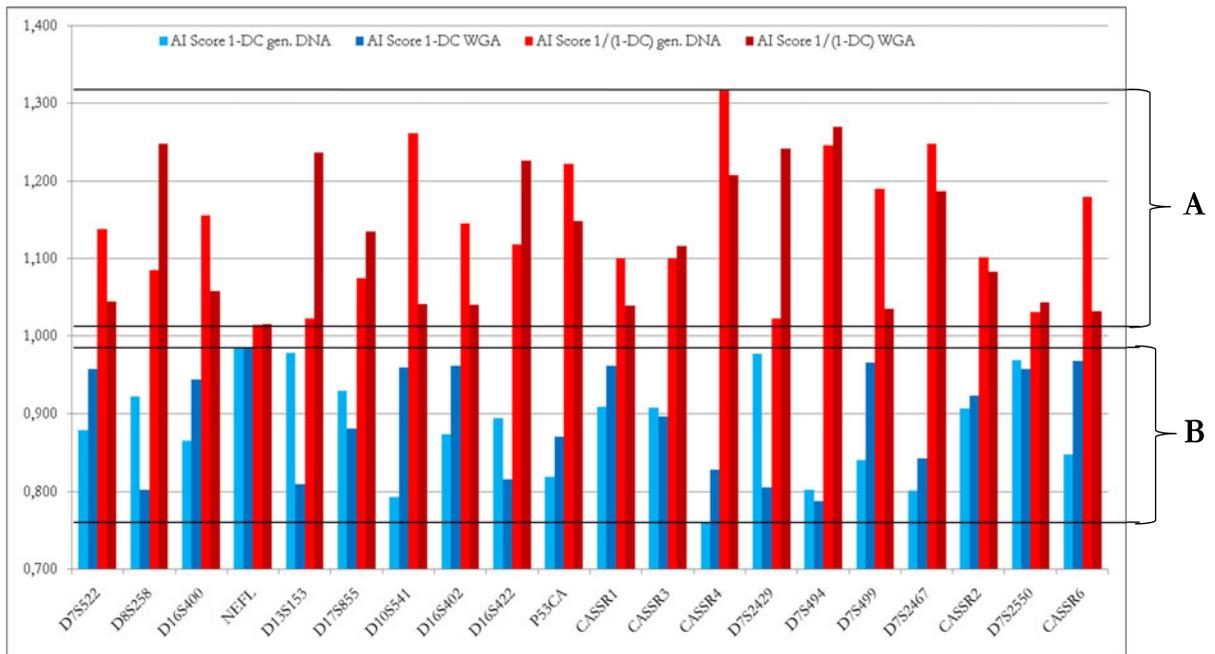
Um die PCR-bedingten Schwankungen der einzelnen Mikrosatelliten Assays erfassen zu können, wurden die PCRs in unabhängigen 3-fach Bestimmungen durchgeführt und anschließend aus den Fragmentanalysen der PCR Produkte die Grenzen für die Bestimmung eines Verlustes im kurzen Allel (AI Score 1-DC) und eines Verlustes im längeren Allel (AI Score 1/(1-DC)) bestimmt. Die DC-Werte (*Critical Difference*) wurden über die Standardabweichung der jeweiligen Dreifachbestimmungen berechnet. Die verwendete Funktion lautete  $DC = 2 \cdot \sqrt{2} \cdot SD$ . Dieselben Bestimmungen wurden parallel für genomische DNA und WGA amplifizierte DNA durchgeführt. Als Ausgangsmaterial für die WGA wurde 30pg genomische DNA eingesetzt. Die Ergebnisse wurden in Tab. 3.8 für genomische DNA und in Tab. 3.9 für WGA amplifizierte DNA aufgelistet. Eine zusammenfassende Auswertung der Daten wurde als Diagramm in Abb. 3.30 dargestellt.

Primer	Fläche A1	Fläche A2	Ratio A1/A2	SD	DC	AI Score 1-DC	AI Score 1/(1-DC)
D7S522	7385	11624	0,635	0,043	0,121	<b>0,879</b>	<b>1,138</b>
	6463	10903	0,593				
	7349	10541	0,697				
D8S258	8040	7284	1,104	0,027	0,078	<b>0,922</b>	<b>1,084</b>
	6119	5647	1,084				
	8043	6998	1,149				
D16S400	8483	17988	0,472	0,048	0,135	<b>0,865</b>	<b>1,156</b>
	12925	22584	0,572				
	6193	13162	0,471				
NEFL	7360	7358	1,000	0,005	0,014	<b>0,986</b>	<b>1,014</b>
	7387	7361	1,004				
	7298	7356	0,992				
D13S153	5472	5429	1,008	0,008	0,022	<b>0,978</b>	<b>1,022</b>
	11614	11418	1,017				
	6454	6287	1,027				
D17S855	6038	5306	1,138	0,025	0,070	<b>0,930</b>	<b>1,075</b>
	9383	8221	1,141				
	6771	5681	1,192				
D10S541	6576	3885	1,693	0,073	0,207	<b>0,793</b>	<b>1,261</b>
	13165	7472	1,762				
	15653	8368	1,871				
D16S402	11544	3158	3,655	0,045	0,127	<b>0,873</b>	<b>1,145</b>
	12076	3292	3,668				
	13710	3650	3,756				
D16S422	10307	5807	1,775	0,037	0,106	<b>0,894</b>	<b>1,119</b>
	9494	5576	1,703				
	8679	5136	1,690				
P53CA	21449	10705	2,004	0,064	0,182	<b>0,818</b>	<b>1,222</b>
	25931	12000	2,161				
	17855	8575	2,082				
CASSR1	7304	3713	1,967	0,032	0,091	<b>0,909</b>	<b>1,100</b>
	12270	6388	1,921				
	6022	3012	1,999				
CASSR3	16511	7706	2,143	0,054	0,152	<b>0,848</b>	<b>1,180</b>
	18998	9311	2,040				
	18126	8977	2,019				
CASSR4	13542	7044	1,922	0,085	0,240	<b>0,760</b>	<b>1,316</b>
	19988	9596	2,083				
	10668	5038	2,118				
D7S2429	5597	5011	1,117	0,008	0,022	<b>0,978</b>	<b>1,023</b>
	6021	5442	1,106				
	4287	3905	1,098				
D7S494	2848	1975	1,442	0,070	0,198	<b>0,802</b>	<b>1,246</b>
	8684	6422	1,352				
	4770	3753	1,271				
D7S499	7530	15814	0,476	0,056	0,159	<b>0,841</b>	<b>1,190</b>
	12043	24827	0,485				
	5209	14417	0,361				
D7S2467	8669	9710	0,893	0,070	0,199	<b>0,801</b>	<b>1,248</b>
	21983	22131	0,993				
	9330	8767	1,064				
CASSR2	20871	13418	1,555	0,033	0,093	<b>0,907</b>	<b>1,102</b>
	38621	25517	1,514				
	20937	13138	1,594				
D7S2550	5200	10955	0,475	0,011	0,030	<b>0,970</b>	<b>1,031</b>
	6841	14832	0,461				
	5138	10544	0,487				
CASSR6	7316	7193	1,017	0,032	0,092	<b>0,908</b>	<b>1,101</b>
	7386	7327	1,008				
	7148	6613	1,081				

Tab. 3.8: AI Scores für genomische DNA

Primer	Fläche A1	Fläche A2	Ratio A1/A2	SD	DC	AI Score 1-DC	AI Score 1/(1-DC)
<b>D7S522</b>	10234	13699	0,747	0,015	0,042	<b>0,958</b>	<b>1,044</b>
	6980	8919	0,783				
	6140	8109	0,757				
<b>D8S258</b>	7999	7603	1,052	0,070	0,199	<b>0,801</b>	<b>1,248</b>
	5791	6542	0,885				
	3853	4131	0,933				
<b>D16S400</b>	10016	18793	0,533	0,019	0,055	<b>0,945</b>	<b>1,058</b>
	6599	13034	0,506				
	6815	14040	0,485				
<b>NEFL</b>	7392	7208	1,026	0,005	0,015	<b>0,985</b>	<b>1,015</b>
	7218	7054	1,023				
	7134	7041	1,013				
<b>D13S153</b>	7773	6684	1,163	0,068	0,191	<b>0,809</b>	<b>1,237</b>
	5456	4366	1,250				
	3688	2776	1,329				
<b>D17S855</b>	11230	8168	1,375	0,042	0,119	<b>0,881</b>	<b>1,135</b>
	7035	4768	1,475				
	7482	5317	1,407				
<b>D10S541</b>	12348	6179	1,998	0,014	0,040	<b>0,960</b>	<b>1,042</b>
	7951	3918	2,029				
	9413	4705	2,001				
<b>D16S402</b>	18403	6192	2,972	0,014	0,038	<b>0,962</b>	<b>1,040</b>
	9761	3248	3,005				
	9406	3149	2,987				
<b>D16S422</b>	14896	9009	1,653	0,065	0,185	<b>0,815</b>	<b>1,226</b>
	6929	4024	1,722				
	6867	3788	1,813				
<b>P53CA</b>	13370	9894	1,351	0,046	0,129	<b>0,871</b>	<b>1,148</b>
	10111	7884	1,282				
	10033	8086	1,241				
<b>CASSR1</b>	9392	5418	1,733	0,013	0,038	<b>0,962</b>	<b>1,039</b>
	8127	4774	1,702				
	6342	3709	1,710				
<b>CASSR3</b>	14255	7155	1,992	0,011	0,031	<b>0,969</b>	<b>1,032</b>
	14085	6978	2,018				
	14055	6988	2,011				
<b>CASSR4</b>	19596	9236	2,122	0,061	0,172	<b>0,828</b>	<b>1,207</b>
	14779	6713	2,202				
	14430	6356	2,270				
<b>D7S2429</b>	4140	3996	1,036	0,069	0,195	<b>0,805</b>	<b>1,242</b>
	3415	3928	0,869				
	4676	4799	0,974				
<b>D7S494</b>	4376	2677	1,635	0,075	0,212	<b>0,788</b>	<b>1,270</b>
	7371	5074	1,453				
	8088	5319	1,521				
<b>D7S499</b>	10210	19100	0,535	0,012	0,034	<b>0,966</b>	<b>1,035</b>
	10776	21080	0,511				
	9758	18134	0,538				
<b>D7S2467</b>	2454	2657	0,924	0,056	0,158	<b>0,842</b>	<b>1,187</b>
	14964	14277	1,048				
	9940	9612	1,034				
<b>CASSR2</b>	26815	16946	1,582	0,027	0,077	<b>0,923</b>	<b>1,083</b>
	29380	18719	1,570				
	26153	17212	1,519				
<b>D7S2550</b>	5125	11158	0,459	0,015	0,042	<b>0,958</b>	<b>1,043</b>
	8154	17377	0,469				
	6312	12770	0,494				
<b>CASSR6</b>	7401	7201	1,028	0,037	0,104	<b>0,896</b>	<b>1,116</b>
	7333	7212	1,017				
	7233	6578	1,100				

Tab. 3.9: AI Scores für WGA Produkte



**Abb. 3.30: AI Scores der Assays für genomische DNA und WGA amplifizierter DNA** Auf der horizontalen Achse wurden die jeweiligen Namen der Assay und auf der vertikalen Achse die AI Scores aufgetragen. Die Bereiche der minimalen und maximalen Schwankungen der AI Scores für genomische DNA wurden für 1/(1-DC) mit A und für 1-DC mit B markiert.

Die Grenzen des Schwankungsbereichs A, für die AI Scores 1/(1-DC), wurden von den AI Scores der Assays NEFL (1,014) und CASSR4 (1,316) für genomische DNA gebildet. Die Grenzen des Schwankungsbereichs B, für die AI Scores 1-DC, wurden von den AI Scores der Assays NEFL (0,986) und CASSR4 (0,76) für genomische DNA gebildet. Alle AI Scores die für WGA amplifizierte DNA lagen innerhalb der Schwankungsbereiche A bzw. B.

Der gemittelte Varianzkoeffizient (VK) der Assays mit genomischer DNA betrug 3,7%, wobei der höchste VK für D7S499 bei 12,8% lag. Der gemittelte VK der Assays mit WGA amplifizierter DNA lag bei 3,1%, wobei der höchste VK für D7S2429 bei 7,1% lag.

### 3.13 Etablierung einer Methode zur genomischen Charakterisierung von Einzelzellen

Nach den ersten Validierungen der WGA wurde ein Protokoll zur Amplifikation der DNA von Einzelzellen etabliert. Das Verfahren wurde unter der Berücksichtigung folgender Aspekte optimiert:

1. Minimierung von: Möglichen DNA Verlusten, Kontaminationsrisiko durch den Anwender oder andere äußere Einflüsse
2. Maximierung von: Der visuellen Kontrolle während des gesamten Verfahrens, dem Spektrum verwendbarer Färbungen und Fixierungen für das Probenmaterial

Der Arbeitsablauf der Einzelzellanalyse wurde in Abb. 3.31 schematisch dargestellt.

## Einzelzell Analyse Arbeitsablauf

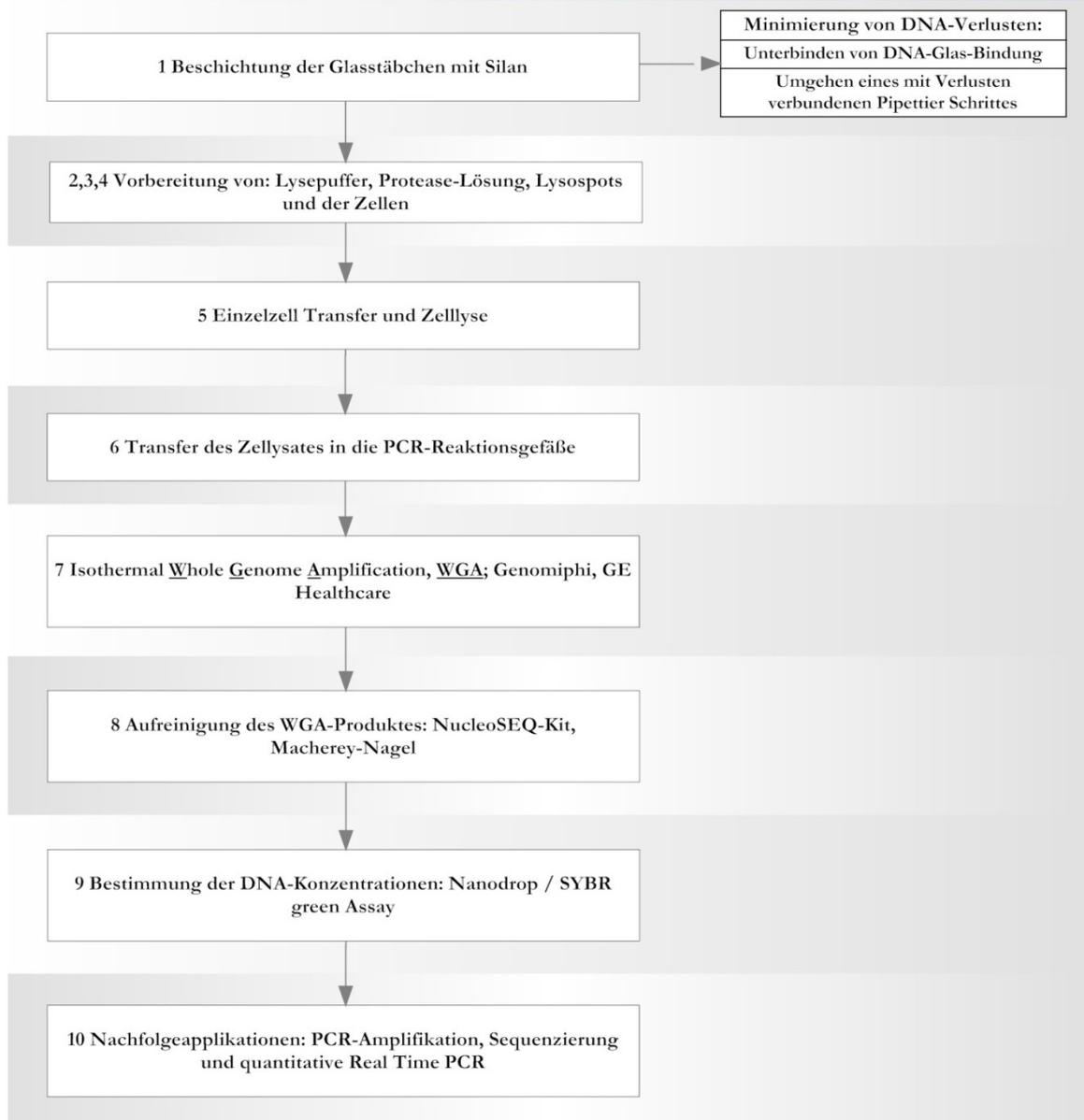


Abb. 3.31: Schematische Darstellung des Ablaufs der genomischen Charakterisierung von Einzelzellen. Die einzelnen Punkte werden im Folgenden weiter behandelt.

**Zu 1:** Zur Minimierung potentieller Verluste beim Transfer von Einzelzellen wurde der Einsatz von Glasstäbchen als Träger gewählt (Details siehe 2.2.2.2). Die Oberfläche der Stäbchen wurde mit einem hydrophoben Silan versiegelt, um das irreversible Binden von DNA nach der Zellyse zu unterbinden. Die optisch klaren Glasstäbchen boten zudem die Möglichkeit uneingeschränkter visueller Kontrolle während der Mikromanipulation.

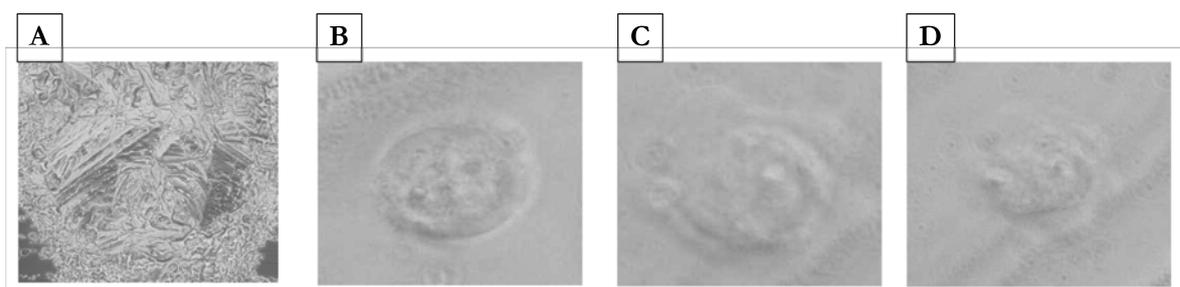
**Zu 2:** Der Lysepuffer sollte eine effiziente Zellyse gewährleisten und ebenfalls das Arbeiten mit RNA ermöglichen. Um die WGA nicht zu inhibieren, wurde ein geringes Volumen von 20nl des Lysepuffers auf die Glasstäbchen aufgebracht (0,2µl einer 1:10 Verdünnung). Dieses wurde anschließend bei RT eingetrocknet. Die Salze des Puffers konzentrierten sich durch die hydrophobe Beschichtung auf einem sehr kleinen Bereich. Auf dem so entstandenen Lysospot wurden die durch Mikromanipulation isolierten Zellen in einem möglichst kleinen Volumen Flüssigkeit abgesetzt.

**Zu 3:** Es wurde ein Protease Verdau zur Steigerung der Effizienz der Zellyse gewählt. Um inhibitorische Effekte während der WGA zu verhindern, wurde die Protease der Proteinase K vorgezogen (beides Qiagen), da diese bei 70°C vollständig Hitze-inaktivierbar war. Die Protease wurde in einer 1:10 Verdünnung verwendet.

**Zu 4:** Es wurden verschiedene Fixierungen und Färbeprotokolle (Fluoreszenz und Substratbasierend) getestet. Auch der Einfluss von längerfristiger Lagerung bei -80°C wurde untersucht (siehe **Zu 5**).

**Zu 5:** Die Isolierung und der Transfer der Cytokeratin positiven Zellen wurde durch Mikromanipulation durchgeführt. Die verwendeten Kapillaren mit unterschiedlich großem Innendurchmesser besaßen einen speziellen Schliff, der sowohl das Manipulieren von fixierten Zellen als auch die Verwendung von zelladhäsiven Objektträgern ermöglichte.

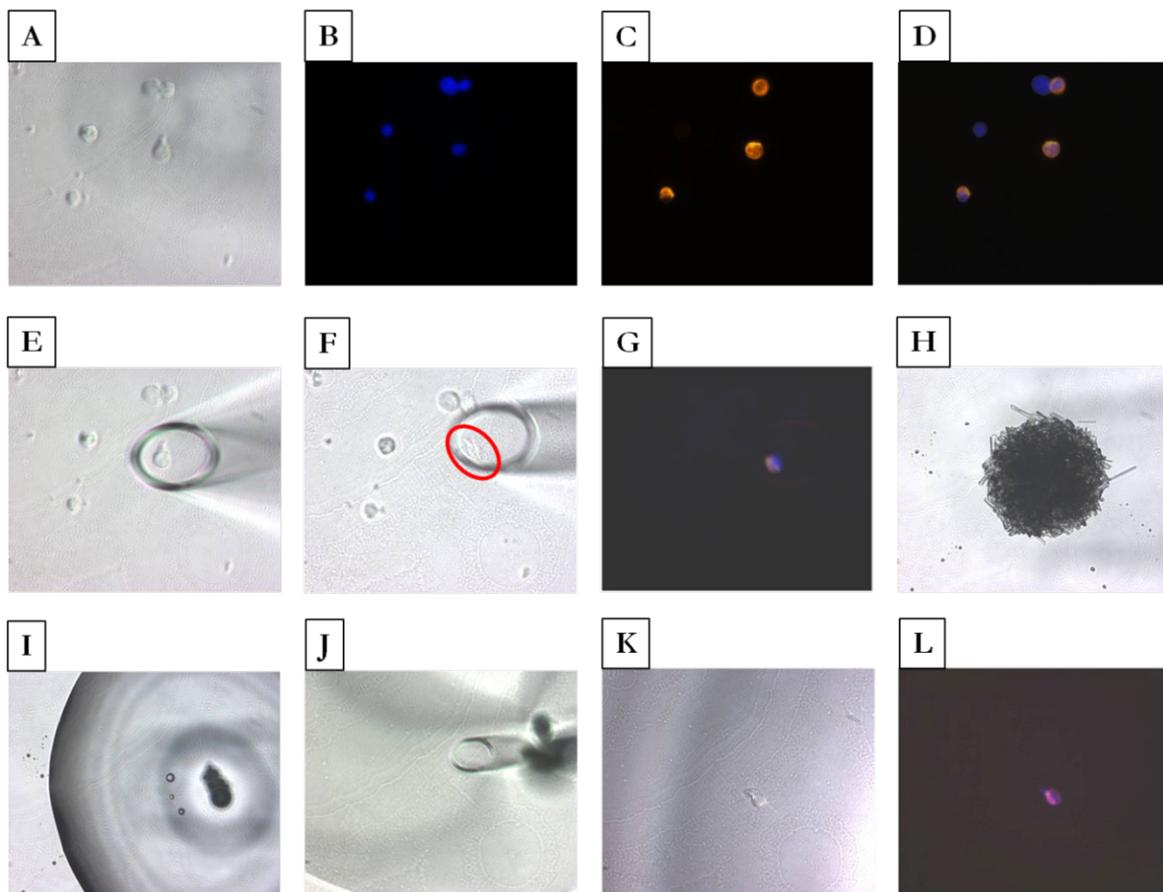
Die Wirksamkeit des Lysepuffers wurde anhand von unfixierten Zellen getestet (Abb. 3.32).



**Abb. 3.32: Lyse einer unfixierten MDA-MB-468 Zelle** Lichtmikroskopische Aufnahmen bei 400-facher Vergrößerung: **A** Lysospot; **B** abgesetzte Zelle nach 30sek; **C** nach 60sek; **D** nach 120sek

Das Verschwimmen der Zellkonturen über die Zeit spiegelte den fortschreitenden Lyseprozess wieder. Nach 150sek waren keine Konturen mehr erkennbar.

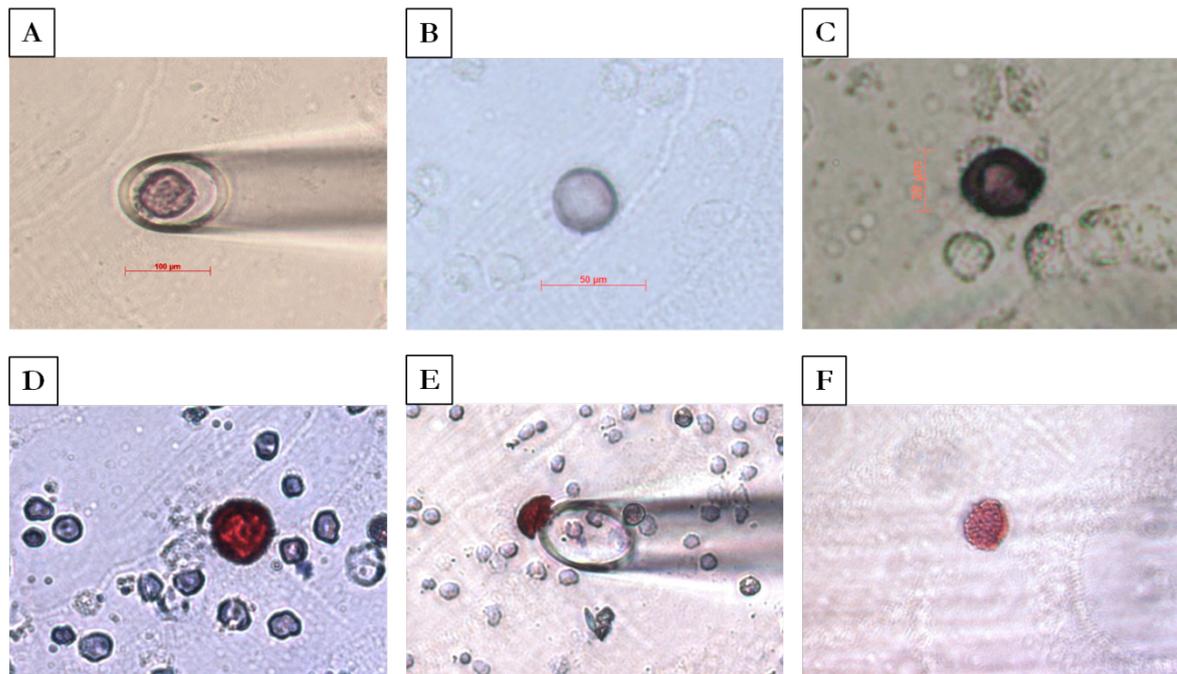
Durch die kombinierte DAPI Kern- und A45-CY3 Fluoreszenzfärbung gegen Zytokeratine ließen sich einerseits Zellen mit intaktem Zellkern identifizieren und zum anderen MDA-MB-468 Zellen eindeutig von den Cytokeratin negativen Fibroblasten unterscheiden. Die Fluoreszenzfärbung ermöglichte eine effiziente Kontrolle über die Position der Zelle während der gesamten Manipulation (Abb. 3.33).



**Abb.3. 33: Sequenzaufnahmen der Mikromanipulation von A45-CY3 Fluoreszenz gefärbten MDA-MB-468 Zellen** **A** lichtmikroskopische Aufnahme (400fach) von humanen Fibroblasten mit eingespikten MDA-MB-468 Zellen; **B** identische Aufnahme im DAPI Kanal zeigte blau fluoreszierende Zellkerne; **C** identische Aufnahme im CY3 Kanal zeigte orange fluoreszierende Cytokeratin positive MDA-MB-468 Zellen; **D** Überlagerung von **B+C**; **E** Die Kapillare wurde über einer MDA-MB-468 Zelle platziert; **F** Die Zelle wurde vorsichtig vom Träger gelöst; **G** Überlagerungsaufnahme der Zelle innerhalb der Kapillare; **H** Lysospot auf Glasstäbchen (200fach); **I** resuspendierender Lysospot; **J** Kapillare in Lysepuffer; **K** Abgelegte MDA-MB-468 Zelle auf Glasstäbchen; **L** Überlagerungsaufnahme beider Fluoreszenzkanäle

Die CK-positiven MDA-MB-468 Zellen wurden durch die Fluoreszenzfärbung mit dem Antikörper A45-CY3 von den CK-negativen Fibroblasten unterschieden (C+D). Durch die DAPI Färbung konnte überprüft werden, ob der Zellkern intakt war (B). Die MDA Zelle wurde mit der Kapillare des Mikromanipulators von der Oberfläche gelöst und aufgenommen (E-G). Bei dem Absetzen der Zelle resuspendierte der Lysospot (H-J). Die Position der Zelle konnte während der Manipulation durch die Fluoreszenzfärbung nachverfolgt werden (G+L).

Neben Fluoreszenz basierten Färbungen wurden auch die Substratumsatz-basierten Färbungen APAAP/Neufuchsin und BCIP/NBT getestet (Abb. 3.34).



**Abb. 3.34: BCIP/NBT und APAAP/Neufuchsin gefärbte CK-positive Zellen** **A-C** bläulich violette BCIP/NBT-gefärbte Cytokeratin positive Zellen; **D-F** rote APAAP/Neufuchsin gefärbte Cytokeratin positive Zellen; **A** SK-BR-3 Zelle (~80µm Durchmesser) mit umgebenden Monozyten-Zellen; **B** MDA-MB-468 Zelle (~40µm Durchmesser) mit umgebenden Knochenmarkzellen; **C** Cytokeratin positive Patientenzelle (~20µm Durchmesser) mit umgebenden Knochenmarkzellen; **D** MDA-MB-468 Zelle mit umgebenden Monozyten Zellen; **E** MDA-MB-468 Zelle während der Mikromanipulation; **F** MDA-MB-468 Zelle im Lysepuffer auf einem Glasstäbchen nach der Mikromanipulation

Die beiden Substratumsatz-basierten, alkalische Phosphatase-abhängigen BCIP/NBT- und APAAP/Neufuchsin-Färbungen ermöglichten eine eindeutige Differenzierung zwischen den CK-positiven Tumorzellen und den umgebenden CK-negativen Blut- oder Knochenmarkzellen. Die visuelle Kontrolle über die manipulierte Zelle war ebenfalls gegeben. Der Nachteil der Färbemethoden bestand darin, dass keine kombinierte Kernfärbung möglich war. Auf die Intaktheit des Kerns konnte ausschließlich über die Intaktheit der ganzen Zelle rückgeschlossen werden.

In der folgenden Tabelle 3.10 wurden alle Kombinationen aus Vorbehandlungen, Fixierungen und Färbungen aufgelistet, die im Rahmen dieser Arbeit auf ihre Lyseeffizienz bzw. WGA Amplifizierbarkeit getestet wurden. Die WGA amplifizierte DNA wurde durch die CASSR1 Mikrosatelliten PCR überprüft.

Probe	Färbung	Fixierung	Lagerung -80°C	DAPI Färbung	CASSR1 PCR
MDA- MB-468	ohne	ohne	ohne	ja	10 / 10
	ohne	15 min. Aceton	ohne	ja	1 / 5
	A45-CY3	15 min. Aceton	ohne	ja	1 / 5
	APAAP/ Neufuchsin	15 min. Aceton	ohne	nein	0 / 5
	ohne	20 min. PFA	ohne	ja	3 / 5
	Fluoreszenz	20 min. PFA	ohne	ja	3 / 5
	ohne	10 min. PFA	ohne	ja	4 / 5
	A45-CY3	10 min. PFA	ohne	ja	4 / 5
Apaap Neufuchsin	10 min. PFA	ohne	nein	1 / 5	
ohne	ohne	> 4 Wochen	ja	5 / 5	
Fluoreszenz	10 min. PFA	> 4 Wochen	ja	4 / 5	
BCIP-NBT	ohne	> 4 Wochen	nein	13 / 15	
A45-PE	Veridex Profile Kit	ohne	nein	3 / 4	
Monozyt	ohne	ohne	ohne	ja	6 / 15
	ohne	10 min. PFA	ohne	ja	2 / 5
	A45-CY3	10 min. PFA	ohne	ja	2 / 5
BCIP-NBT	ohne	> 4 Wochen	nein	1 / 4	

**Tab. 3.10: PCR-basierte Einzelzell WGA Effizienz Validierung an unterschiedlich vorbehandeltem Probenmaterial** Für diese Testserie wurden in Monozyten eingespikete MDA-MB-468 Zellen verwendet. Als Parameter wurde der Einfluss verschiedener Färbeprotokolle in Kombination mit verschiedenen Fixierungsmethoden getestet. Als Material wurden frisch präparierte oder länger als vier Wochen bei -80°C gelagerte Cytospins verwendet. Die WGA wurde anhand der CASSR1-Mikrosatelliten-PCR als erfolgreich oder fehlgeschlagen eingestuft.

Im Hinblick auf die Färbungen ließ sich feststellen, dass sowohl BCIP-NBT als auch die Fluoreszenzfärbungen (A45-CY3, -PE und DAPI) einen geringeren Einfluss auf die Amplifizierbarkeit des CASSR1 Locus hatten als die APAAP/Neufuchsin Färbung. Für unfixierte BCIP-NBT gefärbte Zellen konnten 87% der CASSR1 Loci amplifiziert werden. Für Fluoreszenz-gefärbte Zellen in Kombination mit einer 10min. PFA-Fixierung wurden noch 80% der CASSR1 Loci amplifiziert. Für die stark alkalische APAAP/Neufuchsin Färbung konnten in Kombination mit einer 10min. PFA-Fixierung 20%, in Kombination mit der Aceton-Fixierung keiner der CASSR1 Loci amplifiziert werden. Für ungefärbte oder Fluoreszenz-gefärbte und Aceton fixierte Zellen konnten jeweils 20% der CASSR1 Loci amplifiziert werden. Für die Protein-quervernetzende PFA-Fixierung wurden bei einer Fixierdauer von 10min. 80%, bei einer Fixierdauer von 20min. 60% der CASSR1 Loci amplifiziert. Eine Lagerung der Zellen bei -80°C hatte keinen erkennbaren Einfluss. Für ungefärbte und unfixierte Monozyten konnten <50% der CASSR1 Loci amplifiziert werden. Wurde WGA Produkt von Monozyten als Referenzmaterial verwendet, wurden jeweils ~5 gepoolte Monozyten amplifiziert. Dadurch wurde die Amplifizierbarkeit der CASSR1 Loci auf 100% erhöht.

**Zu 6:** Für den verlustfreien Transfer der manipulierten Zelle wurde das Glasstäbchen in ein PCR-Reaktionsgefäß überführt, in dem bereits der Probenpuffer des WGA Kits vorgelegt war. Zur Verbesserung der Zellyse wurden die manipulierten Zellen für mindestens 20min bei  $-80^{\circ}\text{C}$  eingefroren. Anschließend wurde die Protease hinzugefügt und für min. 15min ein enzymatischer Verdau durchgeführt. Die Protease wurde durch Hitze inaktiviert. Das Glasstäbchen verblieb bis zur abschließenden Aufreinigung des WGA Produktes im Reaktionsgefäß.

**Zu 7:** Nach Hitze basiertem Aufschmelzen der DNA zu Einzelsträngen wurde der mit Enzym-Mix versetzte Reaktionspuffer zu dem Zellysat gegeben und die WGA nach Herstellerangaben durchgeführt.

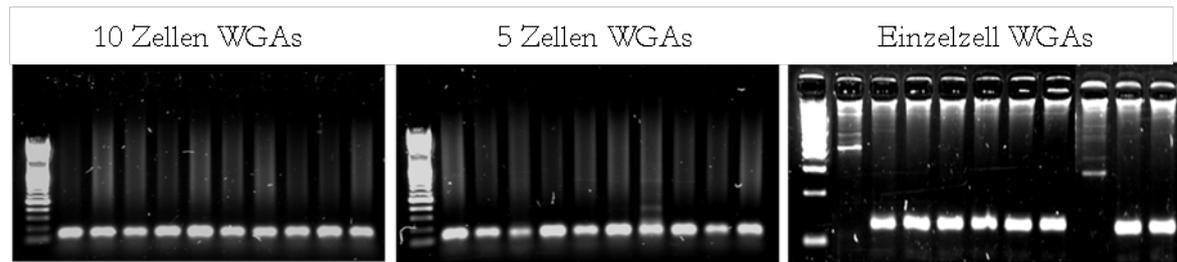
**Zu 8:** Durch die Aufreinigung der WGA Reaktion über eine Gelfiltration wurden freie dNTPs und Hexamer Primer entfernt. Diese hätten in Folgeapplikationen störende Einflüsse zur Folge gehabt.

**Zu 9:** Zur Quantifizierung wurden sowohl das Spektralphotometer Nanodrop als auch SYBR green basierte Quantifizierung über Fluoreszenzmessungen getestet. Die Vergleiche beider Messverfahren zeigten keine Unterschiede in den gemessenen DNA-Konzentrationen. Aus Gründen der Kostensenkung und Zeitersparnis wurden standardmäßig Nanodrop Messungen verwendet.

**Zu 10:** Die Folgeapplikationen wurden in den folgenden Kapiteln 3.14-3.19 ausführlich behandelt.

### **3.14 Validierung der WGA amplifizierten DNA aus Einzelzellen durch Mikrosatelliten PCR**

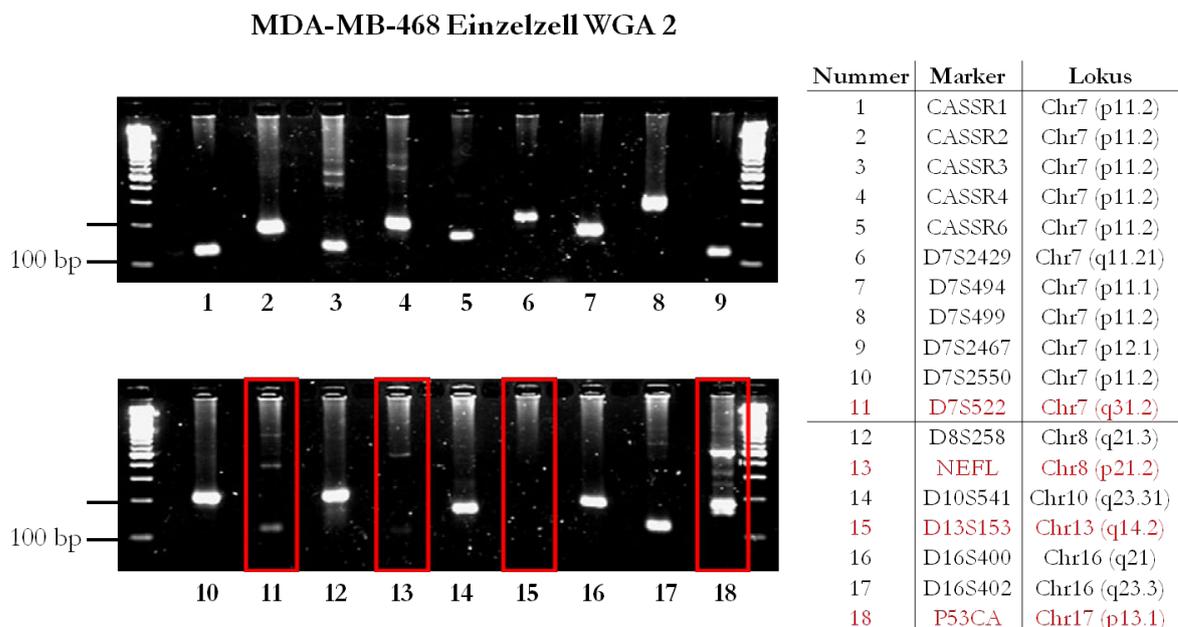
Die Validierung der WGA an Einzelzellen wurde an der Zelllinie MDA-MB-468 durchgeführt, die in Monozyten eingespiked, 10min PFA fixiert und A45-CY3 und DAPI gefärbt wurden. Die Zellen wurden nach dem beschriebenen Protokoll manipuliert und amplifiziert. Initial wurde eine Testreihe durchgeführt, in der jeweils 10, 5 oder eine MDA-MB-468 Zelle als Ausgangsmaterial für die WGA eingesetzt wurden. Die Ansätze wurden jeweils 10-mal wiederholt und anschließend mit der CASSR1 PCR überprüft (Abb. 3.35).



**Abb. 3.35: CASSR1 PCR auf MDA-MB-468 WGA Produkten aus 10, 5 und Einzelzellen** Die PCR Produkte wurden auf einem 2% Agarose-Gel aufgetrennt und mit einer 100bps Leiter aufgetragen.

Für die WGA amplifizierte DNA aus 10 und 5 Zellen konnten 20/20 CASSR1 Loci amplifiziert werden (100%). Für die WGA amplifizierte DNA aus Einzelzellen konnten noch 8/10 CASSR1 Loci amplifiziert werden (80%). Diese Ergebnisse deckten sich mit den Resultaten von der Testreihe aus Kapitel 3.13. Das leiterähnliche Produkt in den fehlgeschlagenen WGAs wurde aus dem Gel eluiert und als Template für eine weitere CASSR1 PCR eingesetzt, ließ sich jedoch nicht reamplifizieren. Die Banden enthielten also keine Template DNA. Die Herkunft der Banden ist weiterhin unbekannt.

Um einen Überblick über die genomweite Abdeckung der WGA zu erhalten, wurde ein Panel von 18 Mikrosatelliten anhand der Einzelzell WGA2 überprüft. Die Ergebnisse sind in Abb. 3.36 in Form eines Agarosegels dargestellt.



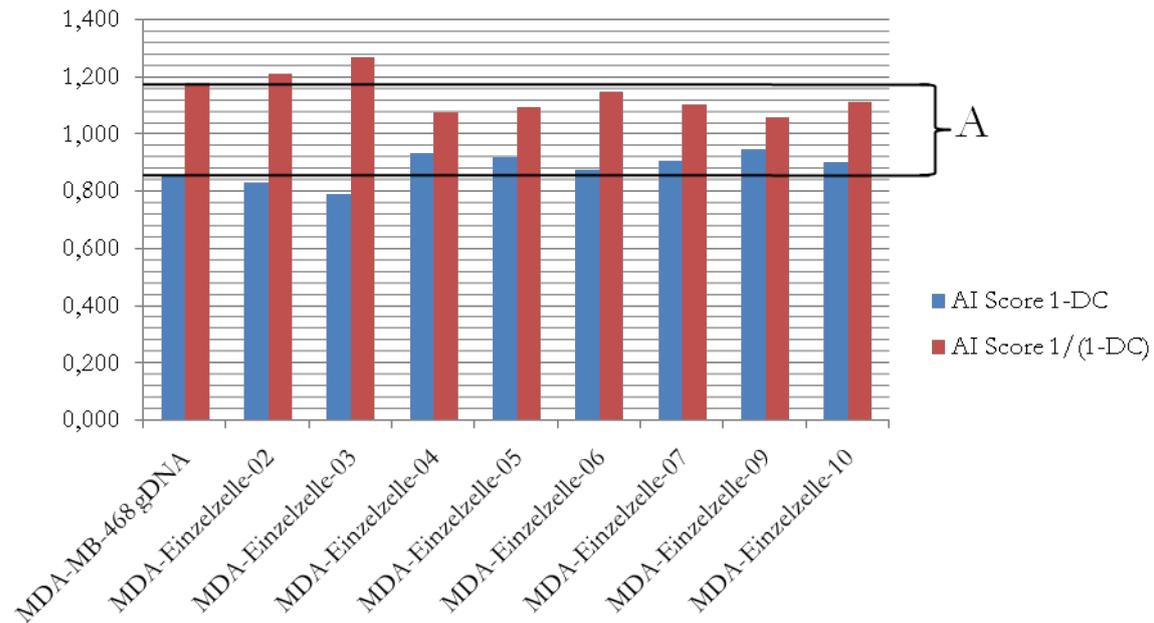
**Abb. 3.36: Mikrosatelliten-PCR Panel auf Einzelzell WGA2** Die PCR Produkte wurden wie unter Abb. 3.31 beschrieben aufgetragen.

Für das Panel von 18 Mikrosatelliten wurden 14 (78%) PCR Produkte der erwarteten Größe amplifiziert (siehe Tab. 3.7). Die Mikrosatelliten P53CA und D7S522 (11%) zeigten neben dem erwarteten PCR Produkt noch hochmolekulare Nebenprodukte. Für die Mikrosatelliten NEFL und D13S153 (11%) konnte gar kein PCR Produkt amplifiziert werden. In der Mikrosatelliten Analyse von der MDA-MB-468 Einzelzell WGA 2 wurde in 16/18 PCR Reaktionen (89%) ausreichend Produkt für eine Fragmentanalyse generiert. Diese Produkte zeigten in der Fragmentanalyse die erwarteten Längen, wohingegen für NEFL und D13S153 kein Fluoreszenzsignal detektiert werden konnte.

Zur abschließenden Validierung der Mikrosatelliten PCR auf Einzelzell Ebene wurde der für MDA-MB-468 heterozygote Marker D7S499 für die acht erfolgreichen Einzelzell WGAs wie unter 3.6 beschrieben gegen genomische DNA getestet. Die Ergebnisse der Validierung wurden in Tab. 3.11, sowie Abb. 3.37 dargestellt.

Primer	Fläche A1	Fläche A2	Ratio A1/A2	SD	DC	AI Score 1-DC	AI Score 1/(1-DC)
MDA-MB-468 genomisch	521	1030	0,506	0,013	0,036	0,964	1,037
	547	1151	0,475				
	1141	2351	0,485				
MDA-1Cell-02	665	1392	0,478	0,015	0,042	0,958	1,044
	1056	2057	0,513				
	2213	4535	0,488				
MDA-1Cell-03	908	1689	0,538	0,024	0,067	0,933	1,072
	421	707	0,595				
	366	652	0,561				
MDA-1Cell-04	779	1483	0,525	0,007	0,020	0,980	1,020
	493	937	0,526				
	419	775	0,541				
MDA-1Cell-05	710	1375	0,516	0,008	0,023	0,977	1,024
	499	932	0,535				
	518	995	0,521				
MDA-1Cell-06	852	1694	0,503	0,011	0,032	0,968	1,033
	324	664	0,488				
	455	883	0,515				
MDA-1Cell-08	1537	3132	0,491	0,008	0,023	0,977	1,023
	452	936	0,483				
	414	824	0,502				
MDA-1Cell-09	837	1459	0,574	0,006	0,018	0,982	1,018
	507	894	0,567				
	588	1009	0,583				
MDA-1Cell-10	3569	7191	0,496	0,009	0,025	0,975	1,026
	651	1270	0,513				
	732	1488	0,492				

Tab. 3.11: AI Scores von D7S499 für genomische MDA-MB-468 DNA und WGA amplifizierte DNA aus Einzelzellen



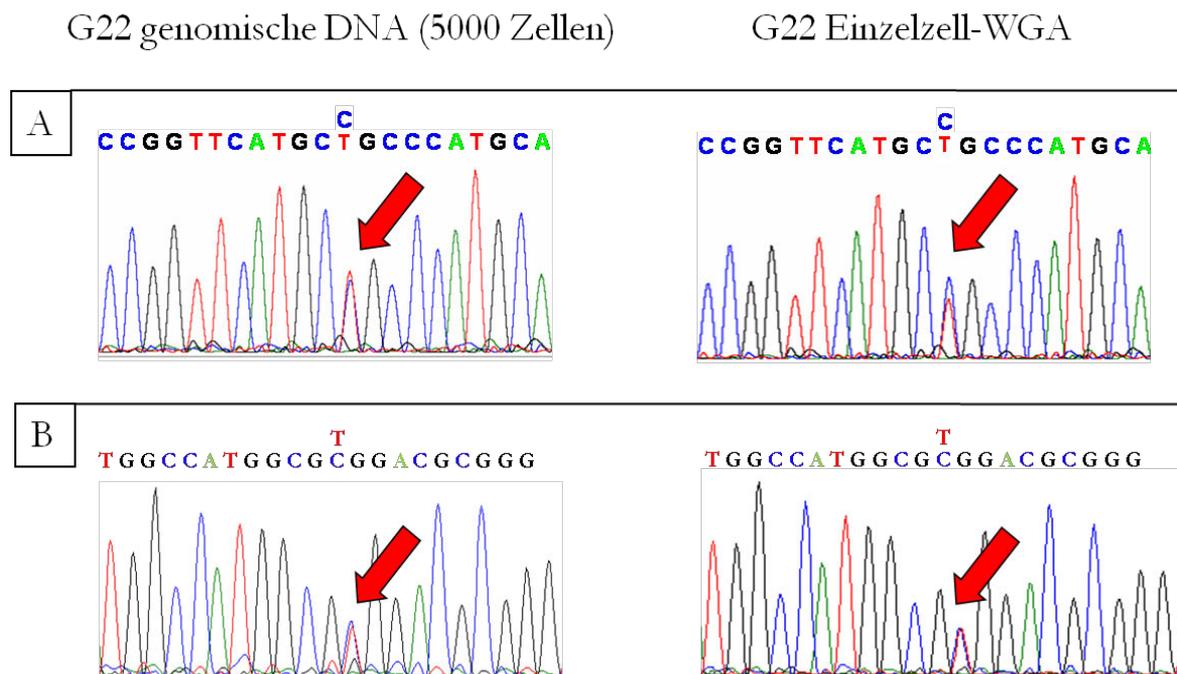
**Abb. 3.37: AI Scores von D7S499 für genomische MDA-MB-468 DNA und WGA amplifizierte DNA aus Einzelzellen** A markiert die für genomische MDA-MB-468 DNA erhobenen AI Scores für 1-DC und 1/(1-DC)

Die AI Scores für 1-DC und 1/(1-DC) der Einzelzell WGAs 4-10 lagen innerhalb des durch die genomische DNA vorgegebenen Grenzen (A). Die AI Scores der Einzelzell WGAs 2+3 lagen außerhalb dieser Grenzen. Mit einer maximalen Abweichung für die Einzelzell WGA 3 von 2,9% des AI Score 1-DC und 3,5% für den AI Score 1/(1-DC) lagen alle Messwerte unter 5% Abweichung. Der VK des Assays mit genomischer DNA lag bei 2,8%. Alle Assays mit WGA amplifizierter DNA hatten VK <2,3% außer MDA-Einzelzelle-03 mit einem VK von 4,1%.

### 3.15 Validierung der Replikationsgenauigkeit der WGA amplifizierten DNA aus Einzelzellen durch Sequenzierung

Die Sequenzierung WGA amplifizierter DNA aus Einzelzellen wurde zur Validierung der Replikationsgenauigkeit durch die phi29 Polymerase durchgeführt. Zusätzlich wurde die Kontaminationswahrscheinlichkeit bei dem bestehenden Manipulations- und Amplifikationssystem durch die Anwender oder andere äußere Einflüsse abgeschätzt. Der sequenzierte DNA Bereich der Exone 5-9 des *TP53* Gens wurde durch drei Primerpaare abgedeckt und umfasste eine Sequenzlänge von 1300 Basen.

Für den Validierungsversuch wurde die Gliomzelllinie G22 verwendet, die zwei bekannte Punktmutationen innerhalb der DNA-Bindedomäne trug. Es wurde hochmolekulare genomische DNA von G22 als Template für die Sequenzierungen mit drei unabhängig voneinander amplifizierten G22 Einzelzell WGA's verglichen. Die Ergebnisse der Sequenzierungen wurden exemplarisch an vier Sequenzausschnitten, die jeweils eine der Mutationen tragen, in Abb. 3.38 dargestellt.



**Abb. 3.38: Punktmutationen der TP53 DNA-Bindedomäne in G22** Gegenübergestellt wurden jeweils identische genomische Sequenzen generiert aus genomischer DNA und Einzelzell WGA Produkt. **A** Punktmutation an Aminosäure-Position 245 in Exon 7 von TP53 (C/T) **B** Punktmutation an Aminosäure-Position 158 in Exon 5 von TP53 (T/C)

Die Punktmutation in Exon 7 hatte einen Aminosäure Austausch von Glyzin nach Serin zur Folge. Die Mutation in Exon 5 führte zu einem Austausch von Arginin nach Histidin. Der durch die Sequenzierung abgedeckte Bereich ist in Abb. 3.39 noch einmal mit allen kodierten Aminosäuren dargestellt.

101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120
AAA	ACC	TAC	CAG	GGC	AGC	TAC	GGT	TTC	CGT	CTG	GGC	TTC	TTG	CAT	TCT	GGG	ACA	GCC	AAG
K	T	Y	Q	G	S	Y	G	F	R	L	G	F	L	H	S	G	T	A	K
121	122	123	124	125	E5126	127	128	129	130	131	132	133	134	135	136	137	138	139	140
TCT	GTG	ACT	TGC	ACG	TAC	TCC	CCT	GCC	CTC	AAC	AAG	ATG	TTT	TGC	CAA	CTG	GCC	AAG	ACC
S	V	T	C	T	Y	S	P	A	L	N	K	M	F	C	Q	L	A	K	T
141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160
TGC	CCT	GTG	CAG	CTG	TGG	GTT	GAT	TCC	ACA	CCC	CGG	CCC	GGC	ACC	CGC	GTC	CGC	GCC	ATG
C	P	V	Q	L	W	V	D	S	T	P	P	P	G	T	R	V	R	A	M
161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180
GCC	ATC	TAC	AAG	CAG	TCA	CAG	CAC	ATG	ACG	GAG	GTT	GTG	AGG	CGC	TGC	CCC	CAC	CAT	GAG
A	I	Y	K	Q	S	Q	H	M	T	E	V	V	R	R	C	P	H	H	E
181	182	183	184	185	186	E187	188	189	190	191	192	193	194	195	196	197	198	199	200
CGC	TGC	TCA	GAT	AGC	GAT	G/GT	CTG	GCC	CCT	CCT	CAG	CAT	CTT	ATC	CGA	GTG	GAA	GGA	AAT
R	C	S	D	S	D	G	L	A	P	P	Q	H	L	I	R	V	E	G	N
201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220
TTG	CGT	GTG	GAG	TAT	TTG	GAT	GAC	AGA	AAC	ACT	TTT	CGA	CAT	AGT	GTG	GTG	GTG	CCC	TAT
L	R	V	E	Y	L	D	D	R	N	T	F	R	H	S	V	V	V	P	Y
221	222	223	224	E7225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240
GAG	CGG	CCT	GAG	GTT	GGC	TCT	GAC	TGT	ACC	ACC	ATC	CAC	TAC	AAC	TAC	ATG	TGT	AAC	AGT
E	P	P	E	V	G	S	D	C	T	T	I	H	Y	N	Y	M	C	N	S
241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260
TCC	TGC	ATG	GGC	GGC	ATG	AAC	CGG	AGG	CCC	ATC	CTC	ACC	ATC	ATC	ACA	CTG	GAA	GAC	TCC
S	C	M	G	G	M	N	R	R	P	I	L	T	I	I	T	L	E	D	S
E8281	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280
AGT	GGT	AAT	CTA	CTG	GGA	CGG	AAC	AGC	TTT	GAG	GTG	CGT	GTT	TGT	GCC	TGT	CCT	GGG	AGA
S	G	N	L	L	G	R	N	S	F	E	V	R	V	C	A	C	P	G	R
281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300
GAC	CGG	CGC	ACA	GAG	GAA	GAG	AAT	CTC	CGC	AAG	AAA	GGG	GAG	CCT	CAC	CAC	GAG	CTG	CCC
D	R	R	T	E	E	E	N	L	R	K	K	G	E	P	H	H	E	L	P
301	302	303	304	305	306	E9307	308	309	310	311	312	313	314	315	316	317	318	319	320
CCA	GGG	AGC	ACT	AAG	CGA	G/GCA	CTG	CCC	AAC	AAC	ACC	AGC	TCC	TCT	CCC	CAG	CCA	AAG	AAG
P	G	S	T	K	R	A	L	P	N	N	T	S	S	S	P	Q	P	K	K
321	322	323	324	325	326	327	328	329	330	331	E10332	333	334	335	336	337	338	339	340
AAA	CCA	CTG	GAT	GGA	GAA	TAT	TTC	ACC	CTT	CAG	ATC	CGT	GGG	CGT	GAG	CGC	TTC	GAG	ATG
K	P	L	D	G	E	Y	F	T	L	Q	I	R	G	R	E	R	F	E	M
341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360
TTT	CGA	GAG	CTG	AAT	GAG	GCC	TTG	GAA	CTC	AAG	GAT	GCC	CAG	GCT	GGG	AAG	GAG	CCA	GGG
F	R	E	L	N	E	A	L	E	L	K	D	A	Q	A	G	K	E	P	G

CGC – CAC  
Arg - His

GGC – AGC  
Gly - Ser

Abb. 3.39: Durch Sequenzierungen abgedeckter Bereich des TP53 Gens Rot umrandet dargestellt wurden die Exone 5-9. Jeweils im Leseraster dargestellt wurden die Aminosäure-Position, darunter die kodierenden Basen und der Einbuchstabencode der daraus resultierenden Aminosäure (Entnommen: TP53 Datenbank).

Durch die dreifache Wiederholung der Sequenzierung sowohl aus der 3', als auch aus der 5'-Richtung wurde eine Gesamtlänge von 7200 Basen sequenziert. Gemessen an der hohen Proofreading-Aktivität der Polymerase stimmte das Ergebnis, keine durch die phi29 Polymerase erzeugte Mutation zu finden, mit den Angaben des Herstellers überein. Die beiden Punktmutationen zeigten in allen Sequenzierungen eine identische Allelfrequenz und eine Kontamination mit Fremd-DNA oder Zellen konnte daraufhin ausgeschlossen werden.

### 3.16 Validierung der WGA amplifizierte DNA aus Einzelzellen durch quantitative Real-Time PCR

Die Validierung der WGA amplifizierte DNA aus Einzelzellen durch quantitative Real-Time PCR wurde unter Verwendung des Genomphi V2 Kit durchgeführt. Das V2 Kit stellt die Weiterentwicklung von V1 dar. Die Versionen unterscheiden sich in zwei maßgeblichen Punkten:

1. Die Reaktionszeit wurde von 16h bei V1 auf 2,5h in V2 verkürzt.
2. Version 1 generiert im Gegensatz zu Version 2 immer auch hochmolekulare DNA in der Negativkontrolle (NTC). Laut Hersteller handelt es sich hierbei um verlängerte Hexamer Primer.

Es wurde eine Testreihe zur Maximierung der Reaktionszeit von V2 durchgeführt, um den Zeitpunkt zu bestimmen, an dem noch kein unspezifisches Produkt in der Negativkontrolle generiert wurde. Dies sollte die Produktausbeute der WGA amplifizierten DNA aus Einzelzellen maximieren. Dieser Versuchsansatz wurde bereits zu einem früheren Zeitpunkt für V1 durchgeführt. Die Ergebnisse der Testreihen sind in Abb. 3.40 dargestellt.

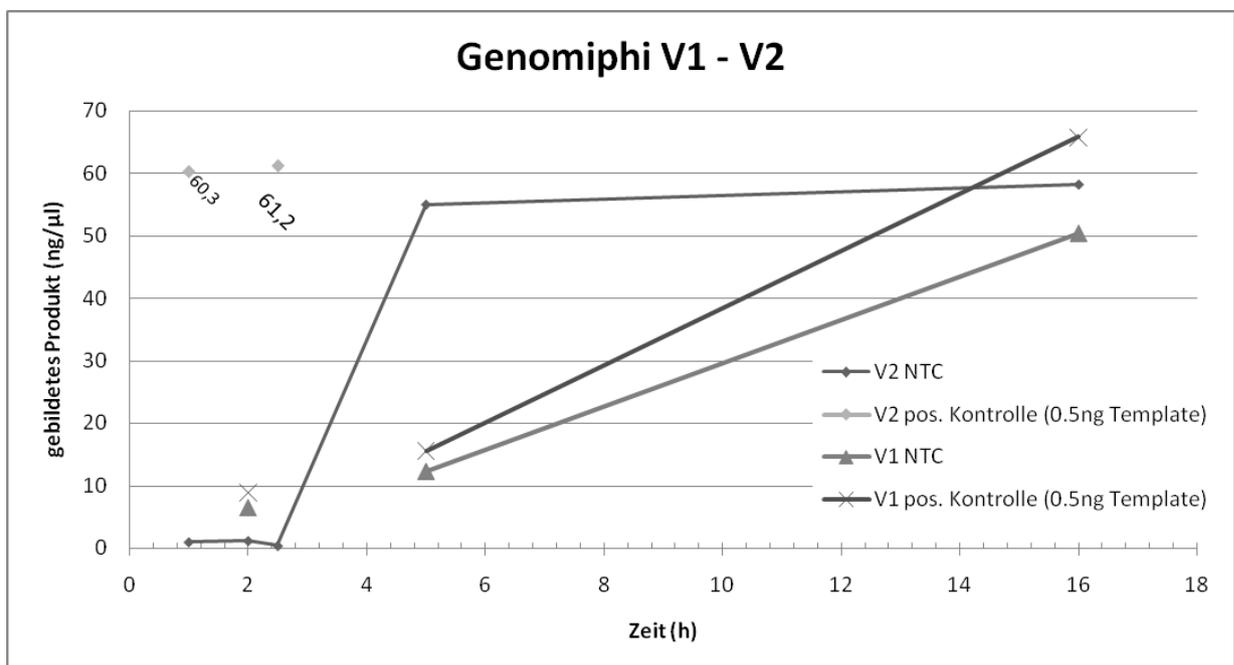


Abb. 3.40: Genomphi V1 & V2 Testreihen zum Vergleich der Reaktionskinetiken in Negativ- und Positiv-Kontrollen

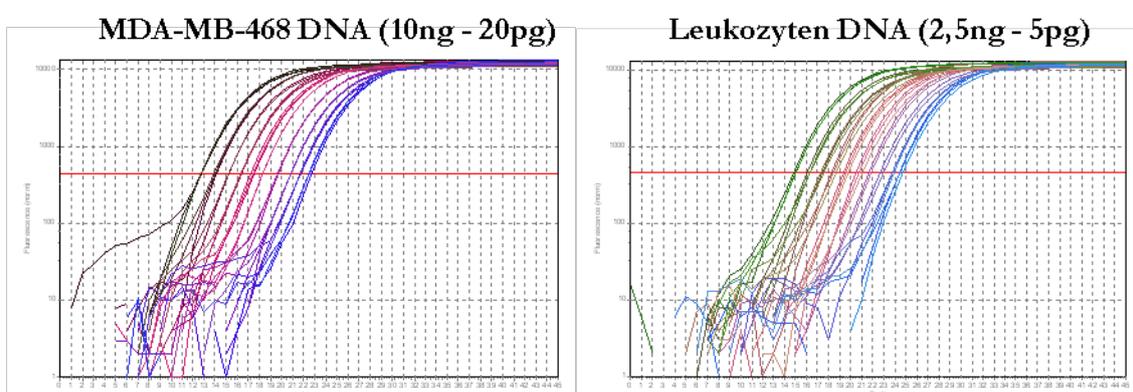
Die vom V1 Kit gebildete Produktmenge stieg über die Zeit linear an. Die gebildete Produktmenge in der Negativkontrolle lag bei jedem Messpunkt bei etwa ~70% der gebildeten Produktmenge der Positivkontrolle. Die vom V2 Kit nach 5h gebildete Produktmenge in der Negativkontrolle entsprach ~80% der gebildeten Produktmenge der Positivkontrolle nach 2,5h. Die Produktmenge der Negativkontrolle stieg nach 16h Reaktionszeit um ~5% an. Bis zu einer Reaktionszeit von 2,5h war kein gebildetes Produkt in der Negativkontrolle nachweisbar. Die gebildete Produktmenge beider Versionen war nach der vom Hersteller empfohlenen Amplifikationszeit identisch.

Basierend auf den Ergebnissen der Testreihe wurde die vom Hersteller empfohlene Amplifikationszeit von 2,5h beibehalten, um die Bildung von unspezifischen Produkten zu unterbinden.

Um bei der Quantifizierung von Einzelzell WGAs Impräzisionen identifizieren zu können, die durch die WGA entstanden, wurde ein weiteres Assay Konzept erstellt. Dieses basierte auf einem einzelnen Referenz Assay, welches eine LINE1 Sequenz amplifizierte, die Genom-weit in  $\sim 900$  Kopien vorlag und dadurch wesentlich unanfälliger für WGA bedingte Schwankungen war als normale diploide Regionen. Bedingt durch die hohe Kopienanzahl wurde eine SYBR green basierte Methode gewählt. Durch die Einsparung eines zweiten Referenzassays wurde Raum für ein zweites parallel durchführbares Assay im *EGFR*-Gen geschaffen. Dies ermöglichte die Verifizierung der Stabilität der Einzelzell WGAs am *EGFR*-Lokus innerhalb desselben qPCR Laufs. Die LINE1 Primer wurden ebenfalls über die Schmelzkurve und ein Agarosegel überprüft.

### 3.17 Validierung des alternativen LINE-1 Referenz Assays für die qPCR

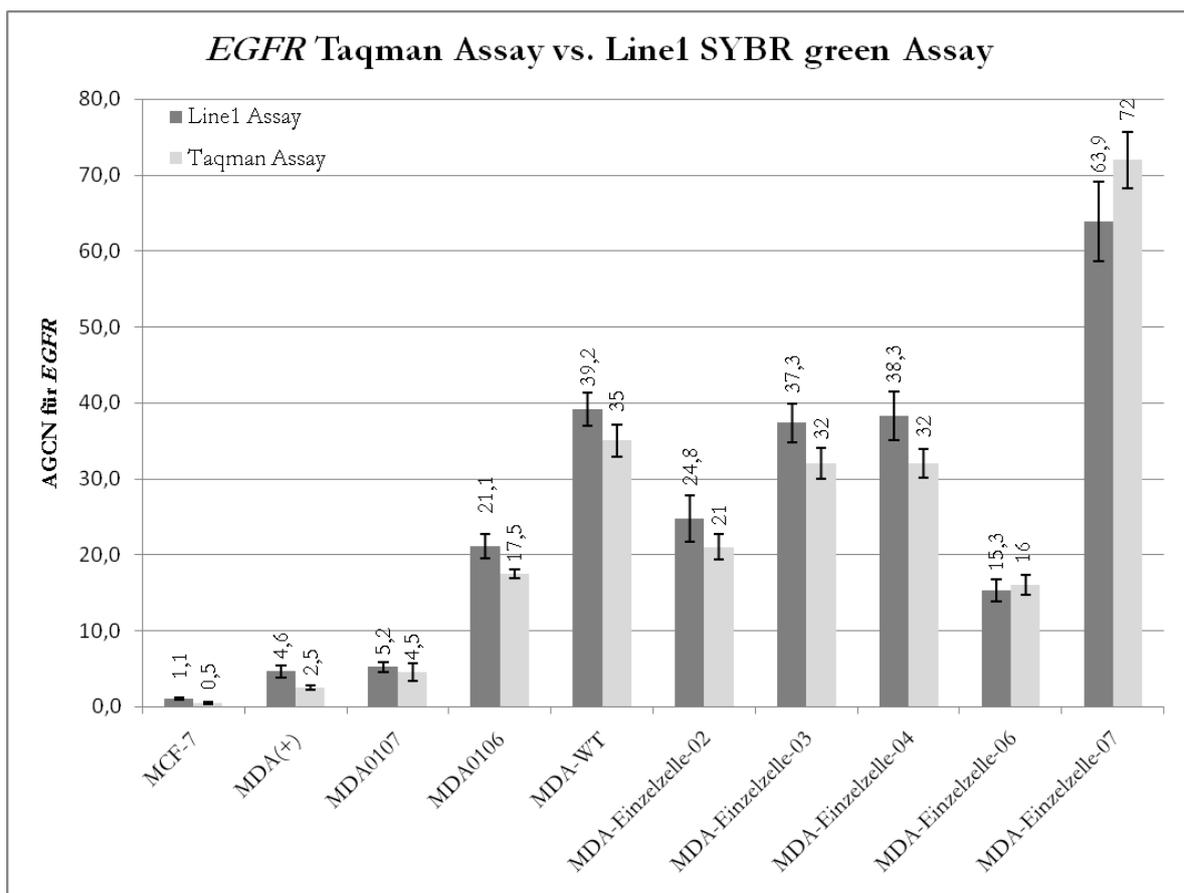
Als erstes musste ausgeschlossen werden, dass es aufgrund der  $\sim 900$ -fach erhöhten Anzahl an LINE1 Sequenzen im Vergleich zu diploiden Regionen zu Effizienzänderungen zwischen den einzelnen Assays oder innerhalb des LINE1 Assays selbst kam. Der Faktor 900 entsprach in der Real-Time PCR einer Differenz von  $\sim 10$  CT-Werten. Zu diesem Zweck wurde die DNA von MDA-MB-468 und Leukozyten in einer geometrischen Reihe mit dem Faktor 2 verdünnt. Die Ergebnisse der geometrischen Verdünnungsreihen wurden in als Amplifikationskurven dargestellt (Abb. 3.41).



**Abb. 3.41: Amplifikationskurven des LINE1 Assays der geometrischen Verdünnungsreihen von MDA-MB-468 und Leukozyten DNA** Der horizontalen Achse wurden die PCR Zyklen und der vertikalen Achse die relativen Fluoreszenzintensitäten zugewiesen.

Durch das Abdecken eines Template Mengen Bereichs von 10ng-20pg für MDA-MB-468 und einen 2CT-Werte größeren Mengen Bereich von 2,5ng-5pg für Leukozyten DNA konnten keine inhibitorischen Effekte festgestellt werden. Wie erwartet zeigten sich für beide Tests jeweils ein CT-Wert Differenz zwischen den einzelnen Verdünnungsstufen. Die Steigung der Amplifikationskurven verlief jeweils parallel zueinander. Die Effizienz der PCR lag für das Assay mit MDA-MB 468 DNA bei 98,8% und mit Leukozyten DNA bei 98,6%.

Da keine inhibitorischen Effekte innerhalb der PCR Assays erkennbar waren, wurde als nächstes ein direkter Vergleich zwischen TaqMan basiertem System und dem LINE1 SYBR green Assay durchgeführt. Zu diesem Zweck wurden die Zelllinien MCF-7, MDA-MB-468 und seine Sublinien, sowie 5 verschiedene Einzelzell WGA's aus MDA-MB-468 Zellen parallel mit beiden Systemen, unter Nutzung identischer *EGFR*-Primer absolut quantifiziert. Die Gegenüberstellung der ermittelten relativen AGCN für *EGFR* wurde in Abb. 3.42 dargestellt.



**Abb. 3.42: Vergleich der durch das TaqMan Assay und SYBR green LINE-1 Assay berechneten AGCN für *EGFR*** Die höchste Abweichung zwischen den beiden Systemen wurde für die hemizygoten Linie MCF-7 und den EGF-supprimierten Subklon MDA(+) gemessen. Die Abweichung für diese beiden Proben lag bei ~1 CT-Wert. Der durch das TaqMan Assay für MCF-7 ermittelte Wert entspricht der erwarteten halben Genkopie. Für alle anderen Quantifizierungen betrug die maximale Abweichung <20,7% (<0,2 CT-Werte).

Die höchsten berechneten Differenzen der AGCN für *EGFR* ergaben sich für die Zelllinien MCF-7 und MDA(+). Diese Differenzen entsprachen Abweichungen der Messwerte der qPCR von  $\sim 1$  CT-Wert. Für die anderen Zelllinien lagen die Abweichungen der Messwerte der qPCR in dem Bereich  $< 0,2$  CT-Werte. Der gemittelte VK für die TaqMan Assays lag bei 4,7%, wobei der höchste VK mit 6,1% für die Messung der Linie MDA(+) berechnet wurde. Für die LINE-1 Assays lag der gemittelte VK bei 9,9%. Der höchste VK wurde für die Messung der Linie MDA(+) mit 16,2% berechnet.

Um die Grenzen des LINE1 Assays im diploiden Bereich validieren zu können, wurde die AGCN für *EGFR* von WGA amplifizierter DNA aus Einzelzellen von Leukozyten und Fibroblasten bestimmt. Diese sollten ein Aberrations-freies Genom besitzen. Die Ergebnisse der absoluten Quantifizierung wurden in Abb. 3.43 dargestellt.

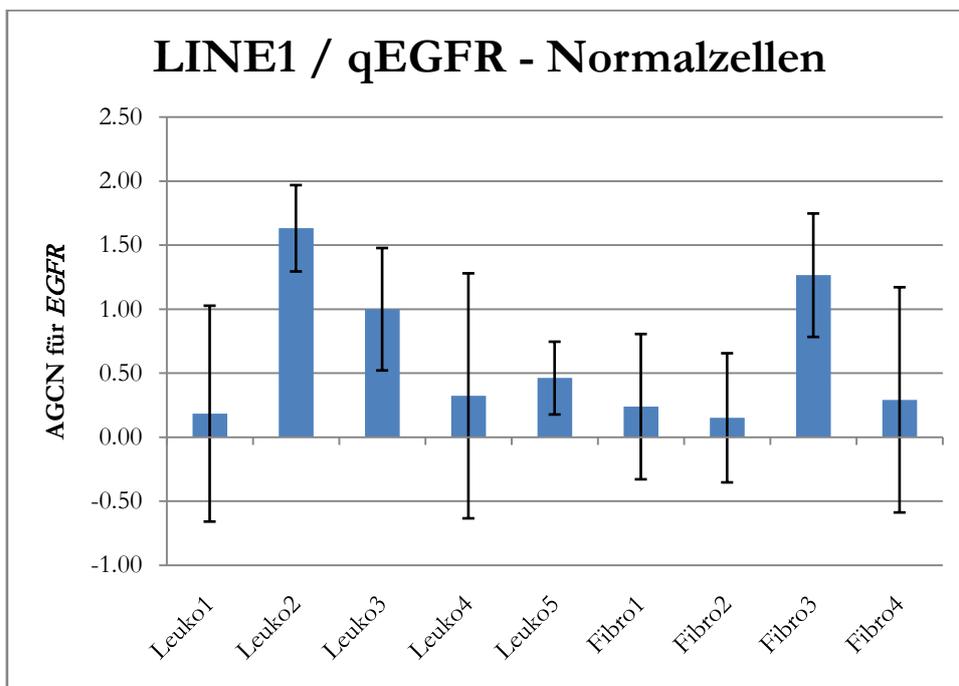


Abb. 3.43: AGCN für *EGFR* aus WGA amplifizierter DNA aus Einzelzellen von Leukozyten und Fibroblasten

Die AGCN für *EGFR* der Einzelzell WGAs der diploiden Zellen lagen im Bereich von 0,2-1,7. Bei einem erwarteten Wert von 1 entsprachen der höchste gemessene Wert einer Abweichung von 70% (0,7 CTs) und der niedrigste Messwert einer Abweichung von 80% (2,25 CTs). Die VK der Messung lag im Mittel bei 1,5%, wobei der höchste VK von 3,6% für die Messung der Einzelzell WGA von Fibroblast 4 berechnet wurde. Es war davon auszugehen, dass die Impräzisionen auf der Amplifikation der WGA beruhten.

### 3.18 Validierung weiterer *EGFR* SYBR green Assays für die qPCR

Zusätzlich zu dem für das TaqMan Assay designten Primerpaares, welches im Intron 1 des *EGFR*-Gens hybridisierte, wurden fünf weitere Primerpaare designed und wie oben beschrieben getestet, die in Exon 4, 7, 9, 15 und 21 hybridisierten. Für alle Primerpaare wurden Effizienzen >98% ermittelt. Validiert wurden die Assays gegen die TaqMan basierten *EGFR*-Amplifikationswerte von MDA-MB-468 und deren Sublinien auf der Grundlage von genomischer DNA. Die Ergebnisse der absoluten Quantifizierung wurden vergleichend in Abb. 3.44 dargestellt.

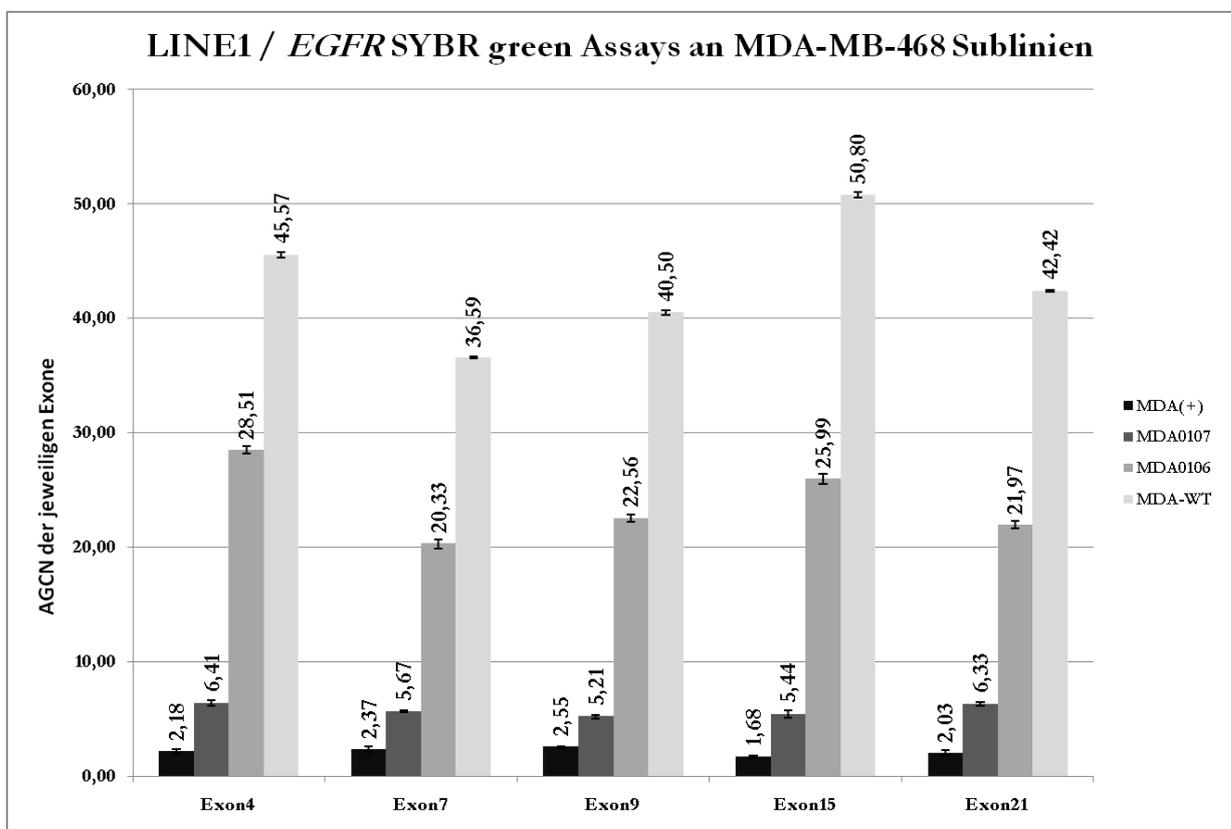


Abb. 3.44 Validierung der SYBR green *EGFR*-Assays an MDA-MB-468 und deren Sublinien

Die berechneten AGCN für die einzelnen Exone wiesen im Vergleich zu den AGCN des TaqMan Assays Abweichungen in einem Bereich von <0,5 CT-Werten auf. Der mittlere VK aller Messungen lag bei 1,7%.

Die maximalen errechneten Abweichungen der AGCN zum TaqMan Assay wurden in Tab. 3.12 als Prozentangaben dargestellt.

Name	Exon 4	Exon 7	Exon 9	Exon 15	Exon 21
MDA+	14,5	5,7	2,0	48,7	23,3
MDA0107	29,8	20,6	13,6	17,3	28,9
MDA0106	38,6	13,9	22,4	32,7	20,3
MDA-WT	23,2	4,3	13,6	31,1	17,5

Tab. 3.12: Prozentuale Abweichungen zwischen SYBR green und TaqMan Assays

Basierend auf den ermittelten Messwerten wurden die Assays von Exon 7 und 9 als Standard Assays in Kombination mit der LINE1 Referenz gewählt. Diese zeigten mit 20,6% für Exon 7 auf MDA0107 DNA und 22,4% für Exon 9 auf MDA0106 DNA die geringsten maximalen Abweichungen und wurden deshalb als Assays mit der größten Präzision angesehen.

### 3.19 Quantifizierung von *EGFR*-Amplifikationen in zirkulierenden Tumorzellen aus dem peripheren Blut und Knochenmark

Es wurden Zellen von drei Patienten mit metastasierendem Mammakarzinom untersucht. Dabei handelte es sich um zwei Knochenmarksproben, eine davon aus der Rechtsmedizin, sowie eine Blutprobe. Die Anwesenheit der DTCs / CTCs innerhalb der Patientenproben wurden durch vorherige immunomagnetische Anreicherung im Veridex System und anschließende immunhistochemische Färbung als Cytokeratin positive Zellen am CellSearch System nachgewiesen. Jeweils ein zelladhäsiver Objektträger mit unfixierten, ungefärbten und vorher bei -80°C gelagerten Zellen wurde BCIP-NBT gefärbt. Anschließend wurden die CK-positiven Zellen nach dem oben beschriebenen Verfahren manipuliert und über WGA amplifiziert. Für die Knochenmarkprobe 1 konnten 4/4 Zellen, für die Knochenmarkprobe aus der Rechtsmedizin 4/5 und für die Blutprobe ebenfalls 4/5 Zellen erfolgreich amplifiziert werden. Die absolute Quantifizierung wurde, wie oben beschrieben, für Exon 7 und Exon 9 des *EGFR*-Gens mit LINE-1 als Referenz durchgeführt. Die Ergebnisse der Quantifizierung wurden in Tab. 3.13 aufgelistet.

	Probe	Exon 7		Exon 9	
		<i>EGFR</i> -Amplifikation	SD	<i>EGFR</i> -Amplifikation	SD
<b>Knochenmark 1</b>	Zelle 1	2,02	0,08	2,18	0,11
	Zelle 2	9,78	0,66	11,43	0,69
	Zelle 3	2,39	0,06	2,18	0,07
	Zelle 4	0,02	0,01	0,40	0,04
<b>Knochenmark 2</b>	Zelle 1	20,95	0,96	15,25	1,23
	Zelle 2	1,42	0,19	1,98	0,27
	Zelle 3	0,21	0,02	0,37	0,04
	Zelle 4	0,10	0,01	0,02	0,01
<b>Blutprobe</b>	Zelle 1	0,11	2,97E-03	0,38	0,03
	Zelle 2	5,05	0,17	8,17	0,81
	Zelle 3	0,03	0,00	0,00	0,00
	Zelle 4	0,58	0,15	15,52	1,50
<b>Monozyten</b>	5 Zellen	0,7	0,16	0,9	0,14

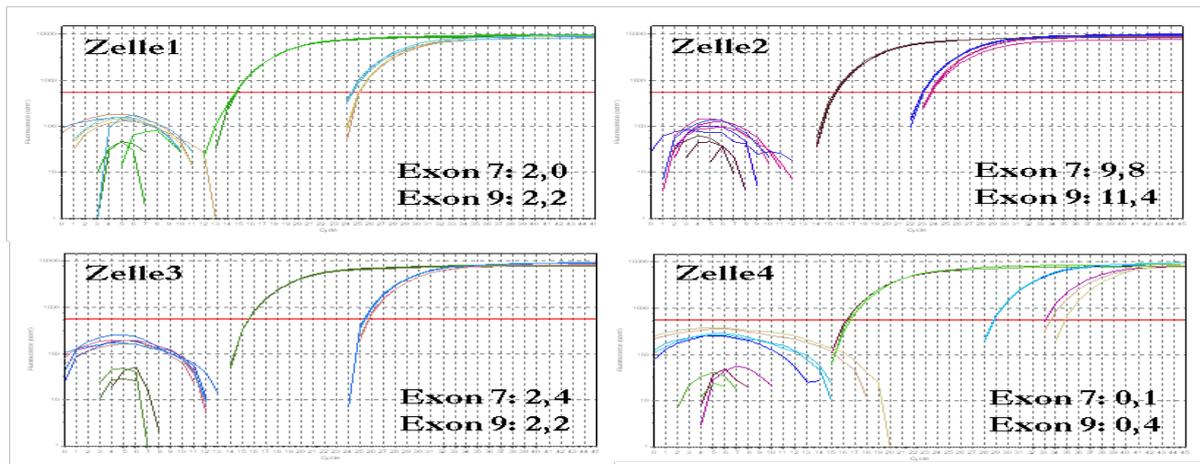
**Tab.3. 13: Quantifizierung der AGCN für Exon 7 und Exon 9 des *EGFR*-Gens in DTZs** In jeder der Proben wurde mindestens eine Zelle mit einer hohen AGCN für *EGFR* nachgewiesen (dunkles Grau). In den Knochenmarksproben wurden drei Zellen nachgewiesen, deren AGCN für *EGFR* sich im Grenzbereich einer sicher detektierbaren Amplifikation befanden (helles Grau). Aufgrund der geringen Standardabweichungen könnten die Werte als Amplifikationen interpretiert werden.

In allen drei Proben konnten *EGFR*-amplifizierte Zellen nachgewiesen werden. Die Amplifikationen zeigten Werte, die sich in drei Gruppen einteilen ließen:

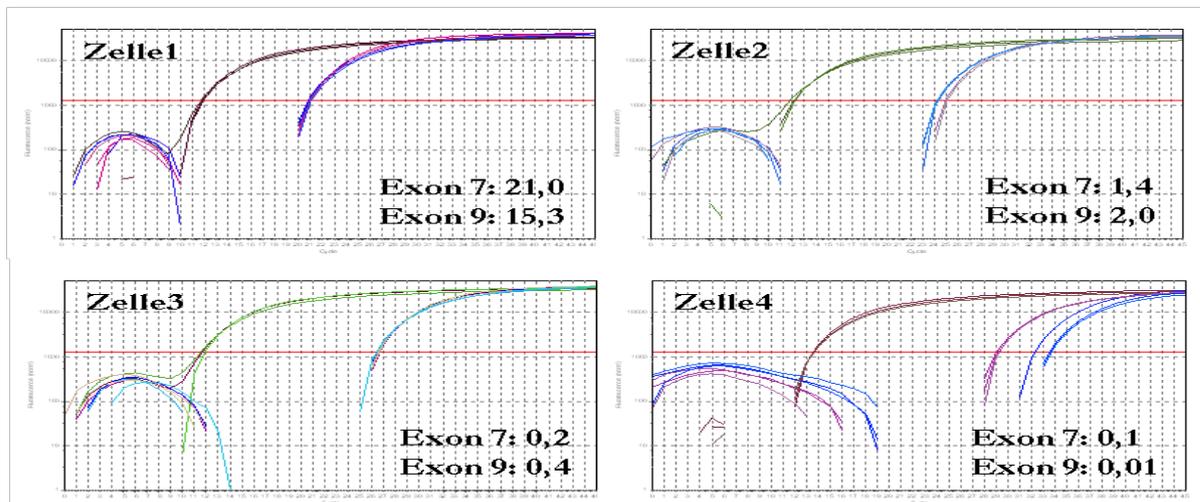
1. 5 - 21fache Amplifikation: stark amplifiziert (4/12 Zellen)
2. 1,4 - 2,4fache Amplifikation: leicht amplifiziert, aufgrund niedriger SDs (3/12 Zellen)
3. 0,0 - 0,58fache Amplifikation: aufgrund von Abb.6.18 als Normal einzustufen (5/12 Zellen)

Die *EGFR*-Amplifikationen in den Patientenzellen wurden aufgrund der folgenden Amplifikationskurven bestimmt (Abb. 3.45).

## Knochenmark 1



## Knochenmark 2 (Rechtsmedizin)



## Blutprobe 1

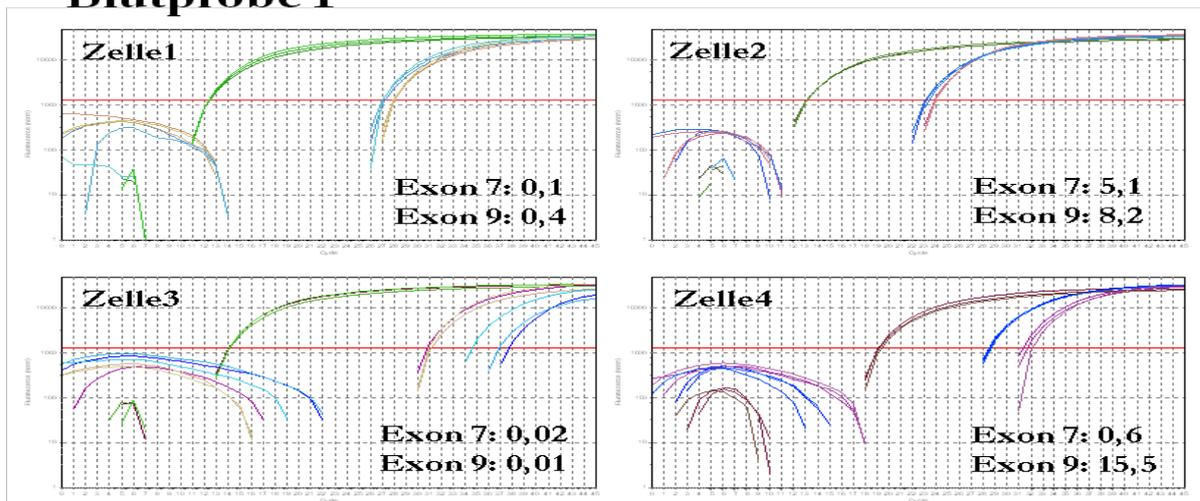


Abb. 3.45 Amplifikationskurven der quantitativen PCR Assays an Patientenmaterial Dargestellt wurden die Amplifikationskurven der jeweils vier quantifizierten Einzelzell WGA's pro Patient. Die Kurven des LINE1 Assays kreuzten jeweils als erstes den Threshold (rote Linie). Die Amplifikationskurven danach stellten die Assays Exon 7 und Exon 9 dar. Die Amplifikationskurven des Assays mit der höheren EGFR-Amplifikation zeigten jeweils die kleineren CT-Werte.

Die CT-Werte des LINE1 Assays verliefen in allen Assays konstant im Bereich von 12-14, mit Ausnahme der WGA von Zelle 4 der Blutprobe. Für diese war der CT-Wert für LINE1  $\sim 5$ CTs größer, was etwa einem Faktor 32 weniger eingesetzter DNA entsprach. Aufgrund der großen Diskrepanz zwischen den Exon 7 und Exon 9 Assays für die WGA der Zelle 4 der Blutprobe wurden zusätzlich Quantifizierungen von Exon 4 und Exon 15 durchgeführt. Diese ergaben beide eine 0,8-fache Amplifikation (SDs 0,2). Ab CT-Werten von  $\sim 35$ -36 begannen die Dreifachbestimmungen aller Quantifizierungen stark zu schwanken, so dass eine quantitative Aussage nicht mehr möglich war. Diese Schwankungen beruhten darauf, dass die Zielsequenzen in der WGA mit verminderter Effizienz amplifiziert wurden. Darüber, ob dies auf Impräzisionen der Amplifikation der WGA oder eine Deletion der Zielsequenz der manipulierten Zelle zurückzuführen war, konnte durch die Methode der qPCR keine Aussage getroffen werden.

### 3.20 Validierung der WGA amplifizierten DNA aus Einzelzellen durch Fine Tiling Array CGH

Als nächstes wurden die Proben aus WGA amplifizierter DNA durch Fine Tiling Array-CGH hinsichtlich der in Kapitel 3.3 genannten Parameter untersucht. Die AGCN für *EGFR* der Proben der WGA amplifizierten DNA wurden in Tab. 3.14 aufgelistet.

Probe	Typ	AGCN für <i>EGFR</i>
MDA-WT 500 Zellen	WGA	27
MDA-Einzelzelle-02	WGA	21
MDA-Einzelzelle-03	WGA	32
MDA-Einzelzelle-04	WGA	32
MDA-Einzelzelle-05	WGA	13
MDA-Einzelzelle-06	WGA	16
MDA-Einzelzelle-07	WGA	72

**Tab. 3.14: TaqMan basierte AGCN für *EGFR* der Fine Tiling Array Proben** Die WGA aus 500 MDA-WT Zellen WT zeigte die identische AGCN für *EGFR* von 27 wie genomische MDA-WT DNA. Die Einzelzell WGA's zeigten AGCN für *EGFR* im Bereich von 13-72.

Von Interesse waren ebenfalls die Konsistenz der Ergebnisse sowie die Qualität bezogen auf das Grundlinienrauschen und die ganzheitliche Abdeckung der Sequenz. Des Weiteren wurde ein Abgleich der für die Einzelzell WGA's berechneten Start- und Endpunkte mit den Grenzen von MDA-WT durchgeführt (Abb. 3.46). Die Art der Darstellung wurde aus Kapitel 3.3 übernommen.

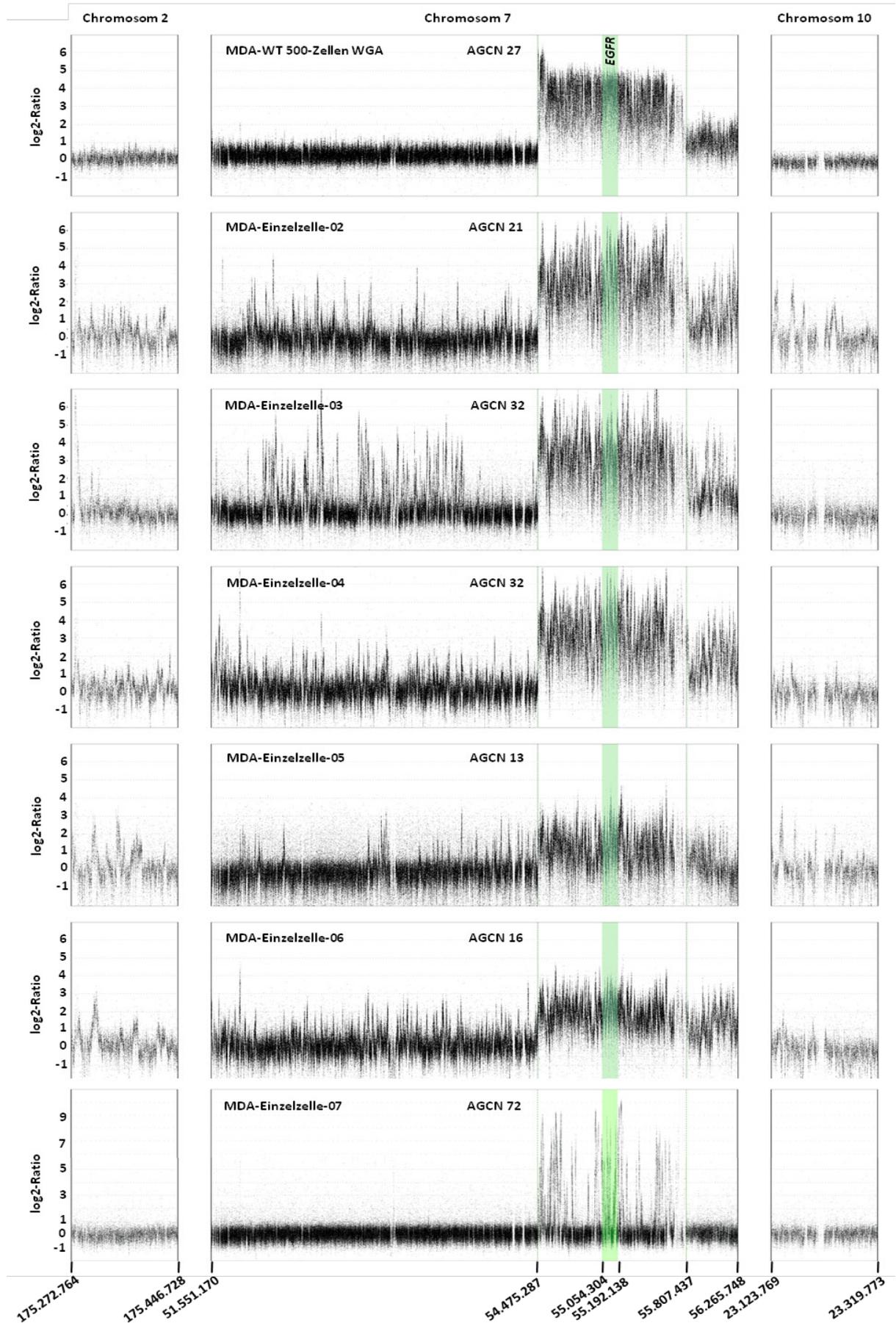


Abb. 3.46: Array-CGH Plots aus WGA amplifizierter DNA von MDA-MB-468 Zellen

Die WGA aus 500 Zellen zeigte keine signifikanten Unterschiede zu dem Plot auf der Basis von genomischer DNA. Für die Einzelzell WGAs konnte ein verstärktes Grundlinienrauschen bedingt durch unregelmäßige Amplifikation einzelner DNA Bereiche festgestellt werden. Sequenzbereiche mit redundanten Signalverlusten oder Zugewinnen traten nicht auf. Klar definierbar blieben die Bereiche des Hauptamplikons und des sich centromerwärts anschließenden leicht amplifizierten Bereichs im Vergleich zu genomischer MDA-WT DNA. In allen Einzelzell WGAs war zudem eine Korrelation zwischen den AGCN für *EGFR* der qPCR und den semiquantitativen log<sub>2</sub>-Ratios der Arrays erkennbar. Die Ausnahme bildete MDA-1Cell-07. Der Plot zeigte kein Grundlinienrauschen und nur vereinzelte Messpunkte im Bereich des *EGFR*-Amplikon. Bedingt durch einen Fehler bei der Hybridisierung des Arrays hatte wahrscheinlich fast ausschließlich die Referenz-DNA hybridisiert.

Die Plots der Einzelzell WGAs zeigten keine Lücken in der Abdeckung der Messpunkte des Arrays, was darauf schließen ließ, dass während des Manipulationsprozesses und der anschließenden Amplifikation keine Verluste der Ausgangs-DNA eintraten. Das Grundlinienrauschen für die Plots der Einzelzell WGAs ließ die Berechnung der Start- und Endregionen des Hauptamplikons in den meisten Fällen zu. Die maximalen Abweichungen der Start- bzw. Endregionen der Einzelzell WGAs von den Regionen der genomischen MDA-WT DNA wurden in Tab. 3.15 dargestellt. Die Abweichungen wurden in Basen und Prozent, bezogen auf die Gesamtlänge des MDA-WT Amplikon angegeben.

Probe	Amplikon Start	Abweichung (bps)	Abweichungen (%)
MDA-WT	54475287		
MDA-5000-Zell-WGA	54475287	0	0
MDA-1Cell-02	54485699	10412	0,8
MDA-1Cell-03	54479868	4581	0,3
MDA-1Cell-04	54485699	10412	0,8
MDA-1Cell-05	54558689	83402	6,3
MDA-1Cell-06	54479868	4581	0,3
MDA-1Cell-07	54503631	28344	2,1
Probe	Amplikon Ende	Abweichung (bps)	Abweichungen (%)
MDA-WT	55807437		
MDA-5000-Zell-WGA	55769646	37791	2,8
MDA-1Cell-02	55859238	51801	3,9
MDA-1Cell-03	nicht möglich	nicht möglich	nicht möglich
MDA-1Cell-04	55832696	25259	1,9
MDA-1Cell-05	55255411	552026	41,4
MDA-1Cell-06	55813259	5822	0,4
MDA-1Cell-07	55797989	9448	0,7

**Tab. 3.15: Abweichungen der errechneten Amplikongrenzen für WGA amplifizierte DNA aus MDA-WT Zellen bezogen auf genomische MDA-WT DNA**

Mit Ausnahme der Endregion des Amplikon für MDA-1Cell-03 konnten für alle Arrays Grenzen berechnet werden. Die maximalen Abweichungen lagen für die Startregion des Amplikon bei 6,3%. Für den Endbereich des Amplikon wurde bei MDA-1-Cell-05 die größte Abweichung mit 41,4% berechnet. Alle anderen Abweichungen der Grenzberechnung im Endbereich lagen unter 4%. Der verwendete quantsmooth Algorithmus schien eine zuverlässige Methode zu sein, um Bruchpunkte in Fine Tiling Array CGHs zu detektieren. Den höchsten prozentualen Abweichungen lagen zum einen das erhöhte Grundlinienrauschen des Plots und zum anderen die AGCN für *EGFR* von 13 der MDA-Einzelle-03 zugrunde. Alle anderen Werte lagen stabil unter 4% Abweichung. Aufgrund der extrem hohen Dichte von 390.000 Messpunkten wurde exemplarisch der Bereich des Amplikonstart (54.475.287) für einen Bereich von 30.000bps (54.470.000-54.500.000) hochaufgelöst dargestellt, um die einzelnen Messpunkte und Feinstrukturen des Amplikons erkennbar zu machen (Abb. 3.47).

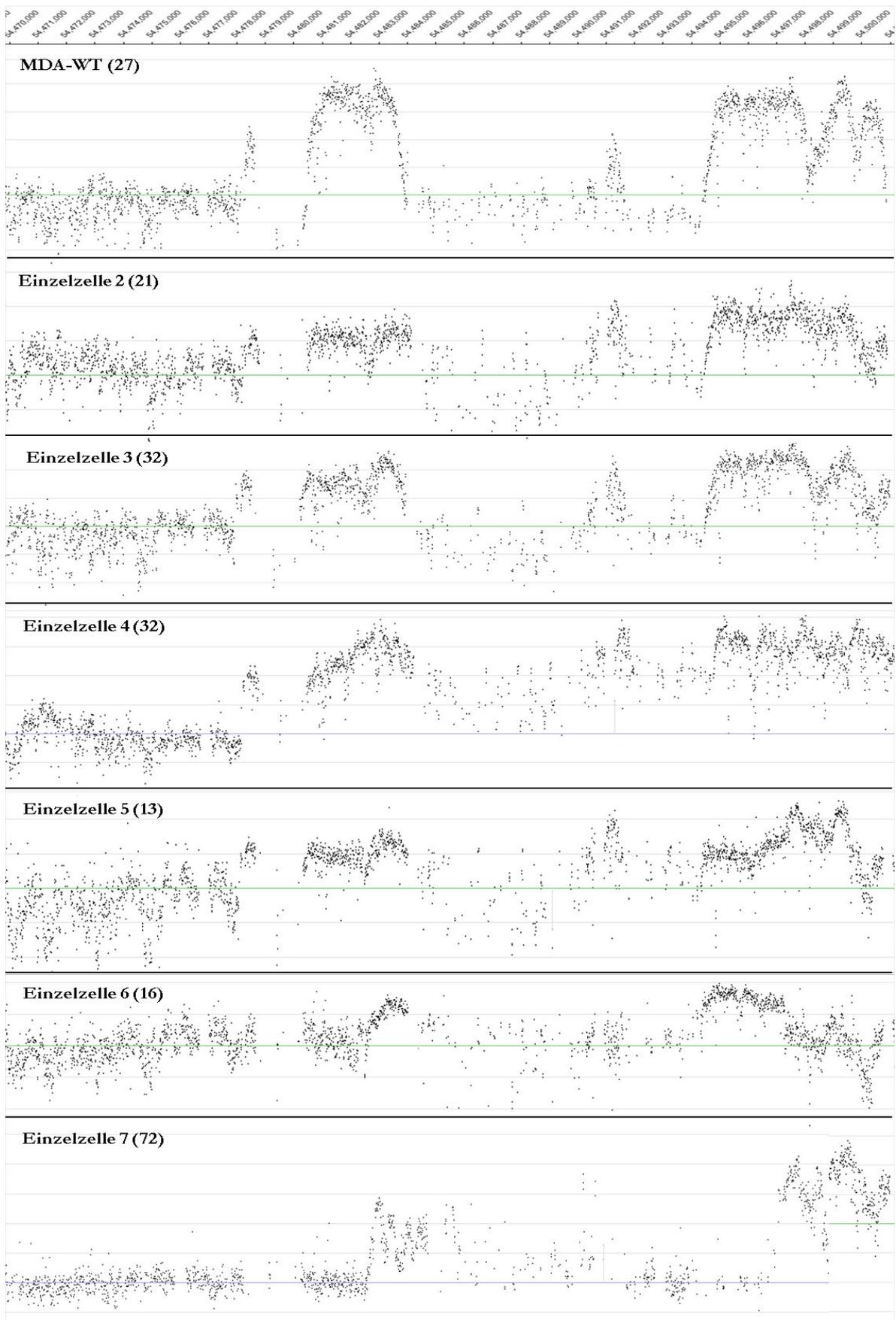


Abb. 3.47: Hochaufgelöste Darstellung des telomerwärtigen Startbereiches des MDA-MB-468 Amplikons von genomischer und WGA amplifizierter DNA aus Einzelzellen Die AGCN für *EGFR* wurde in Klammern angegeben.

---

Für MDA-WT DNA zeichneten sich folgende Strukturen ab:

- 54.478.000-54.479.000 unamplifizierter Normalbereich, am Ende ansteigend
- 54.479.000-54.481.000 Bereich geringer Sondendichte
- 54.481.000-54.484.000 amplifizierte Region (log<sub>2</sub>-Ratios ~5-6)
- 54.484.000-54.495.000 Bereich geringer Sondendichte mit leicht amplifiziertem Bereich bei 54.491.000 (log<sub>2</sub>-Ratios ~3-4)
- 54.495.000-54.500.000 langgestreckte amplifizierte Region (log<sub>2</sub>-Ratios ~5-6)

Aufgrund der hohen Auflösung der Arrays konnten einzelne Feinstrukturen der Amplikons erst durch Verkleinern des dargestellten Bereichs sichtbar gemacht werden. Der strukturelle Feinaufbau des MDA-WT Amplikons konnte für alle Einzelzell WGA in dem analysierten Bereich, mit Ausnahme von WGA 07, reproduzierbar dargestellt werden.

## 4. Diskussion

Das Ziel dieser Arbeit war die Untersuchung einer genomischen Amplifikation des auf dem Chromosom 7p11-14, die das *EGFR*-Gen einschließt. Anhand des Modells der Mammakarzinom Zelllinie MDA-MB-468, einiger unterschiedlich stark *EGFR* amplifizierter Sublinien, sowie Einzelzellen dieser Linie wurden Untersuchungen zur Aufklärung eines möglichen Amplifikationsmechanismus durchgeführt. Die hierbei verwendeten Methoden der *Fine Tiling Array-CGH* und qPCR zum Nachweis von Veränderungen in der DNA Kopienzahl, sowie der Mikrosatellitenanalyse zur Detektion von AIs und der Sequenzierung, wurden für WGA amplifizierte DNA aus Einzelzellen validiert. Ein möglicher biologischer Zusammenhang, zwischen einer *EGFR* Genamplifikation und dem Selektionsprozess der zur Ausbildung von DTCs führt, wurde durch qPCR Analysen des *EGFR*-Gens von CK-positiven Zellen aus dem Blut und Knochenmark von metastasierenden Mammakarzinom Patientinnen untersucht.

### 4.1 Methodische Aspekte

Die genomische Charakterisierung von Einzelzellen ist von mehreren kritischen Faktoren abhängig. Der erste wird von dem Startmaterial selbst gebildet, also der zu analysierenden Zelle. Vorbehandlungen wie Fixierung, Färbung, Alter der Probe, Art und Dauer der Lagerung, sowie der Zelltyp selbst haben einen direkten Einfluss auf die Applizierbarkeit der genomischen DNA und der Reproduzierbarkeit der quantitativen Ergebnisse. Die durch die WGA amplifizierte DNA aus Einzelzellen wurde durch die CASSR-1 Mikrosatelliten-PCR überprüft. Wenn ein singuläres PCR-Fragment der erwarteten Größe amplifiziert wurde, konnte die WGA als erfolgreich eingestuft werden. Die Kombinationen aus PFA- bzw. Aceton-Fixierung (158-160) und Fluoreszenz- bzw. Substratumsatz-basierten Färbungen (20,34) beeinflussten die Rate mit der der CASSR-1 Locus amplifiziert werden konnte unterschiedlich stark. Eine PFA-Fixierung für 10min. senkte die Rate um 20% und für 20min. um 40%. Die Proteinquervernetzung der Fixierung erschwerte wahrscheinlich die Zellyse und somit die Zugänglichkeit der DNA für die WGA. Durch eine Aceton Fixierung wurde die Rate um 80% gesenkt. Die Fluoreszenzfärbungen (A45-CY3, -PE, DAPI) und auch die Substratumsatz-basierte BCIP-NBT Färbung senkten die Rate um 0-13%. Die Substratumsatz-basierte Färbung APAAP-Neufuchsin senkte die Rate um 60%. Das stark alkalische Milieu der Färbung führte wahrscheinlich zu einer Schädigung der DNA. Bei einer Kombination aus Fixierungen und Färbungen addierten sich die negativen Effekte. Bei einer

optimalen Probenvorbereitung konnten mit dem bestehenden Protokoll 87-100% der CASSR-1 Loci für Tumorzelllinien- und CK<sup>+</sup>-Patientenzellen erfolgreich amplifiziert werden. Eine Lagerung der Zellen bei -80°C hatte keine Auswirkungen auf die Rate der erfolgreich amplifizierten CASSR-1 Loci. Die Rate der erfolgreich amplifizierten CASSR-1 Loci lag für Monozyten bei maximal 38% (161). Diese Diskrepanz könnte einerseits auf dem höheren DNA-Gehalt in Tumorzellen, andererseits durch Apoptose-bedingte Degradierung der DNA (162) oder die differierende Nukleus-Struktur in Monozyten erklärt werden. Die Effizienz der WGA war für die Proben mit erfolgreich amplifizierbarem CASSR-1 Locus relativ konstant, wobei jeweils 2-2,5µg WGA amplifizierte DNA gebildet wurden. Ein weiterer essentieller Punkt ist der Prozess der Manipulation der Zelle selbst, deren Lyse und der Transfer in das Reaktionsgefäß der WGA. Der Manipulationsprozess muss mit einer Präzision durchführbar sein, der es erlaubt intakte Einzelzellen aus einem Zellverbund zu lösen, möglichst ohne mesenchymale umgebende Zellen mitzuführen, die die Ergebnisse der Folgeapplikationen verfälschen würden. Für Fluoreszenzfärbungen, welche den Vorteil der DAPI Kern- und CD45 Gegenfärbung für Monozyten hatten, wurde eine 10min PFA Fixierung gewählt. Für unfixierte, BCIP-NBT gefärbte Zellen wurde mit Zell-adhäsiven OTs gearbeitet. Eine effiziente Zellyse, die in der WGA keinen Einfluss auf die Quantität und Qualität des gebildeten Produkts haben sollte, wurde durch den Lysospot erreicht, der das benötigte Volumen Lysepuffer auf 20nl senkte. Ein verlustfreier Transfer in das dafür vorgesehene Reaktionsgefäß wurde durch die oberflächenversiegelten Glasstäbchen erreicht. Der Einfrierschritt mit nachfolgendem Proteaseverdau wurden für eine effiziente Zellyse von fixierten Zellen benötigt (140). Der kritischste Faktor bleibt jedoch die WGA selbst (155,159,163). Diese muss mit möglichst hoher Effizienz, Sequenzstabilität, Homogenität der Amplifikation verschiedener Loci und einer möglichst ganzheitlichen Genomrepräsentation arbeiten, um eine interpretierbare Auswertung einer breiten Basis an Folgeapplikationen zu gewährleisten.

Die Validierung der WGA für die STR Analyse wurde auf Basis von 30pg genomischer DNA aus Zelllinien oder Zellmischungen, mit bekannten Fragmentlängen für 20 STR Marker, als Ausgangsmaterial mit dem WGA-Kit V1 durchgeführt. Die AI Scores für WGA amplifizierte DNA lagen alle innerhalb des durch genomische DNA vorgegebenen Schwankungsbereichs. Der gemittelte VK der Assays mit WGA amplifizierter DNA von 3,1% mit einem maximalen VK von 7,1% war geringer als der gemittelte VK für die Assays mit genomische DNA von 3,7% mit einem maximalen VK von 12,8%. Damit ist die Präzision der Assays mit WGA amplifizierter DNA höher und rechtfertigt den Einsatz zur STR Analyse.

Diese Werte korrelieren gut mit den Literaturwerten von 95-99,9% Übereinstimmung zwischen genomischer und WGA amplifizierter DNA (165,166). Auf der Basis von Einzelzellen wurden für STR Analysen 80-89% der analysierten Marker erfolgreich amplifiziert und Abweichungen für die AI Scores zur Detektion eines Verlusts im kurzen Allel lagen bei nur 2,9%, für den Verlust im langen Allel bei nur 3,5%. Der maximale VK der Assays von WGA amplifizierter DNA aus Einzelzellen lag bei 4,12%. Die Fragmentlängenanalyse zeigte zudem ausschließlich die erwarteten Größen, sodass eine Kontamination durch Fremd- Zellen oder DNA ausgeschlossen wurde. Die relativ hohen in der Literatur beschriebenen Dropout Raten von 28-40% bei Lymphozyten (167) und 7-24% bei humanen Spermazellen (168) begründen sich wahrscheinlich durch die oben beschriebenen Effekte und konnten für die Einzelzell STR Analyse der amplifizierten Tumorzellen nicht bestätigt werden. Die Dropout Raten für Monozyten konnten durch die simultane Amplifikation von 5 Zellen vollkommen kompensiert werden (167,169).

Die Validierung der Replikationsgenauigkeit der phi29 Polymerase in der Einzelzell WGA wurde durchgeführt, um die Sequenzstabilität zu überprüfen und die Wahrscheinlichkeit einer Kontamination abzuschätzen. Hierbei wurden in 3 Einzelzell-WGAs jeweils 2 bekannte Punktmutationen im *TP53* Gen der Zelllinie G22 nachgewiesen. Auf einer sequenzierten Gesamtlänge von 7200 Basen konnten keine Basenfehleinsbauten bedingt durch die phi29 Polymerase festgestellt werden. In einer größeren Studie wurden bei einer Gesamtlänge von 500.000 sequenzierten Basen 5 Sequenzvarianten in WGA amplifizierter DNA im Vergleich zu genomischer DNA detektiert (170). Da sowohl Allelfrequenz, Position und Basenaustausch bei allen Punktmutationen übereinstimmte konnten Kontaminationen ausgeschlossen werden.

Zur Validierung der Einzelzell-WGA für die qPCR, sowie die Charakterisierung der Amplifikation von *EGFR* in Zelllinien und Patientenproben wurden zwei unterschiedliche Assaysysteme etabliert. Das erste System, basierend auf TaqMan Sonden, wurde mit einem Assay für *EGFR* im Intron 1 des Gens, sowie 2 unabhängigen Referenzassays auf Chr.2 und Chr.10 konzipiert. Die Referenzregionen wurden vorher durch Array-CGHs von Mamma-CAs auf ihre genomische Stabilität hin ausgewählt. Die maximalen Assay bedingten Schwankungen wurden mit VK von 8-11% im Bereich >30 Genkopien und einem VK von 5% im diploiden Bereich bestimmt. Damit bot die Methode eine ausreichende Sensitivität, um heterozygote Deletionen zu detektieren. Die Referenzregionen erwiesen sich in 7 getesteten Mamma-CA Linien als nicht aberriert. Die Quantifizierung wurde jeweils absolut, unter der Verwendung von Kalibrationsreihen durchgeführt (171-173).

Um WGA bedingte Impräzisionen innerhalb der Referenzregionen bei der Quantifizierung von Einzelzellen zu umgehen, wurde ein weiteres Assaysystem auf *SYBR green* Basis etabliert. Dieses wurde so konzipiert, dass als Referenz eine LINE-1 Region gewählt wurde, die mit 912 Kopien pro Genom unanfälliger gegenüber WGA bedingten Impräzisionen sein sollte (155). Simultan wurden jeweils 2 von 5 etablierten *EGFR* Assays zur Quantifizierung eingesetzt. Die beiden präferierten Assays zeigten im direkten Vergleich mit dem TaqMan Assay SD von 11,1% für Exon 7 und 12,9% für Exon 9. Bei Bedarf konnte zusätzlich auf Assays in Exon 4, 15 und 21 zurückgegriffen werden. Bedingt durch *SYBR green* basierte Impräzision und die relativ hohe CT-Wert Differenz zwischen Referenzassay und den *EGFR* Assays von ~10 CT-Werten, durch die sich minimale Abweichungen rechnerisch stärker auswirken, wurden im diploiden Bereich Abweichungen von bis zu 100% (1 CT-Wert) erreicht. Dies macht das Assay ungeeignet zur Detektion von heterozygoten Deletionen und setzt die Assay bedingte Grenze für den gesicherten Nachweis einer Amplifikation, je nach Qualität der Messwerte, bei einer AGCN für *EGFR* >2-3 fest. Eine generelle Unterrepräsentierung von repetitiven LINE-1 Sequenzen in WGA-Produkten, wie sie für die Bestimmung mittels Array-CGHs beschrieben wurde (155,169) konnte für die qPCR nicht nachgewiesen werden. Diese hätte zu einer generellen Erhöhung der AGCN für *EGFR* in WGA amplifizierter DNA führen müssen, was nicht der Fall war. Bei der Quantifizierung von 12 CK<sup>+</sup>-Zellen, die über BCIP-NBT Färbung in 2 KM-Präparaten und einer Blutprobe von 3 unterschiedlichen metastasierten Mamma-CA Patienten nachgewiesen wurden, konnten in 7/12 Zellen *EGFR* Amplifikationen gezeigt werden. Hohe Amplifikationen (AGCN 5-21) wurden in 4/12 Zellen und leichte Amplifikationen (AGCN 1,4-2,4) in 3/12 Zellen gemessen. Die niedrigen SDs von 0,06-0,27 der leicht amplifizierten Zellen ließen eine Einstufung als solche zu. Die als Normalreferenz mitgeführten gepoolten 5 mesenchymalen, ungefärbten Zellen zeigten eine AGCN für *EGFR* von 0,7-0,9 und SD von 0,14-0,16. Die Einzelzell-WGA 4 der Blutprobe zeigte nur in Exon 9 eine starke Amplifikation und wurde deshalb noch für Exon 4 und Exon 15 quantifiziert. Da keine Amplifikation in den Exons nachweisbar war, ist davon auszugehen, dass die Amplifikation sich entweder maximal von Exon 8 bis 14 erstreckt, oder die Amplifikation in Exon 9 auf einem WGA generierten Artefakt beruht. Betrachtet man die LINE-1 Sequenz als indirektes Maß für die Effizienz der jeweiligen WGA, so war diese im Vergleich zu den übrigen WGAs mit ~5 CT-Werten um einen Faktor 32 verringert. Eine Aussage über den Auslöser der verminderten Effizienz und deren Auswirkung auf die Amplifikation anderer Loci zu treffen ist nicht möglich. Der Anteil an *EGFR* Amplifikationen in den DTCs von 33% mit einer AGCN  $\geq 10$  liegt weit über dem für Mamma-CA beschriebenen *EGFR* Amplifikationsraten von <5% in Primärtumoren (175).

Ein Zusammenhang zwischen dem Metastasierungsprozess und der Amplifikation von *EGFR* ist deshalb naheliegend. Es ist anzunehmen, dass sich die Genamplifikation von *EGFR* und die daraus resultierende Überexpression des Rezeptors positiv auf den Selektionsprozess der Differenzierung zu einer DTC auswirken. Dabei werden Faktoren, wie Zellwachstum und Vermehrung, Angiogenese, Adhäsion und Beweglichkeit, sowie die Apoptose, an deren physiologischer Regulation *EGFR* beteiligt ist, maßgeblichen Einfluss ausüben.

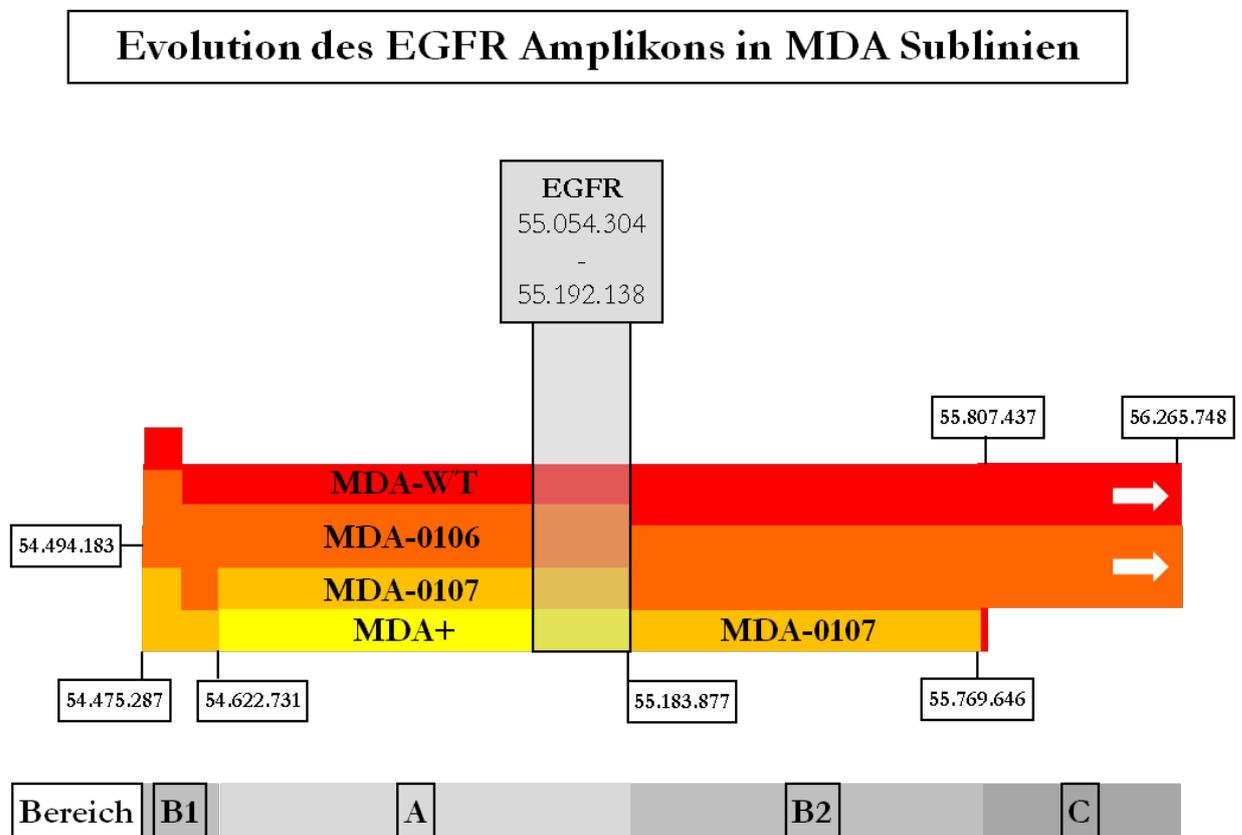
Die zur genaueren Charakterisierung der Länge, Lokalisierung und Struktur der *EGFR* Amplikons durchgeführten *Fine Tiling Array*-CGHs wurden ebenfalls an WGA-Produkten von 500 Zellen und 6 Einzelzellen von MDA-MB-468 durchgeführt. Die Plots der CGHs aus genomischer DNA der Zelllinien A431, BT-20, MDA-MB-468 und deren Sublinien, der für *EGFR* hemizygoten Linie MCF-7 und Leukozyten-DNA, als diploide Referenz, zeigten im Vergleich zu den Plots der CGHs aus WGA amplifizierter DNA der MDA-MB-468 Einzelzellen kaum Grundlinienrauschen. Die Leukozyten Referenz war frei von Aberrationen und die Hemizygotie für Chr.7 in MCF-7 war anhand der log<sub>2</sub>-Ratios von 0,5 erkennbar. Die Referenzregionen auf Chr.2 und Chr.10 zeigten in allen Plots keine Aberrationen, übereinstimmend mit den Befunden der TaqMan qPCR. Für alle *EGFR* amplifizierten Linien konnten klare Strukturen mit Start- und Endpunkten der Amplikons, sowohl Algorithmus basiert, wie auch visuell anhand des Plots, festgesetzt werden. Die 500-Zell-WGA bildete das genaue Bild der genomischen MDA-WT DNA ab. Bei den Einzelzellen war ein erhöhtes Maß an Grundlinienrauschen zu verzeichnen, ebenso wie kleine über- oder unterrepräsentierte DNA Bereiche, die diffus verteilt waren und keinem Muster zwischen den einzelnen Arrays folgten. Trotz dieser Effekte war keine Glättung der Daten über größere DNA Bereiche erforderlich, was die Auflösung des Array verringert hätte, wie für andere Array-CGHs an Einzelzell-WGAs beschrieben (176). Die Algorithmus basierte Detektion der Amplikongrenzen war aufgrund der beschriebenen Effekte in einigen Fällen ungenau oder nicht möglich. Die visuelle Auswertung der Plots zeigte jedoch in allen Fällen eine klare Abbildung des MDA-WT Amplikons. Die durch qPCR erhobenen AGCN für *EGFR* stimmten sowohl für die Arrays der Einzelzell-WGAs, wie auch der Arrays auf Basis genomischer DNA gut mit den semiquantitativen log<sub>2</sub>-Ratios der Arrays in ihren Tendenzen überein. Die einzige Ausnahme bildete Einzelzell-WGA 7, die nur vereinzelte Messwerte im Bereich des MDA-WT Amplikons zeigte, sowie kaum merkliches Grundlinienrauschen. Aufgrund der zuvor für MDA-MB-468 als korrekt validierten STR Marker ist von einem Hybridisierungsfehler bei der CGH auszugehen und eine Kontamination mit anderer DNA auszuschließen.

Zusammenfassend lässt sich festhalten, dass das auf der MDA WGA basierende Protokoll zur genomischen Charakterisierung von Einzelzellen, für die validierten Folgeapplikationen, eine solide Grundlage bildet. Die Auswertung der nun ermöglichten Charakterisierung verschiedenster genetischer Marker in DTCs sollte jedoch immer unter dem kritischen Aspekt durchgeführt werden, dass man die Analysen an der momentan, technisch gesehenen bestmöglichen Kopie der Original-DNA durchführt. Die Ergebnisse der Untersuchung des *EGFR* Amplikons durch *Fine Tiling Array-CGH* und qPCR zum Nachweis von Veränderungen der DNA Kopienanzahl, sowie der näheren Charakterisierung bestimmter Sequenzen durch lange PCRs und Sequenzierungen werden im Folgenden ausführlich behandelt.

## 4.2 Analyse der Struktur des *EGR* Amplikon

Die Zelllinie MDA-MB-468 besteht aus einem heterogenen Mix von Klonen mit unterschiedlichen Genotypen für den *EGFR* Locus. Eine weite Spanne an *EGFR* Genkopienzahlen, die zu unterschiedlichen *EGFR* Expressionen führen, existiert dabei simultan unter normalen Zellkulturbedingungen. In vorangegangenen Untersuchungen konnte, durch FACS-Zellsortierung nach *EGFR* Expression und angeschlossene Analysen (FISH, CGH und qPCR), eine positive Korrelation von *EGFR* Amplifikation und *EGFR* Expression nachgewiesen werden (K. Agelopoulos). Diese direkte Auswirkung der *EGFR* Genkopienzahl wurde ebenfalls für Glioblastome (181), sowie Karzinome des Kopf- und Halsbereichs (182) beschrieben. Vergleichbare Gendosiseffekte wurden für zahlreiche Gene bei Karzinomen der Pankreas (183) und Mamma (184,185) beschrieben. Unter Stimulation des *EGFR* Pathways mit EGF konnte ein Wachstums-inhibierender Effekt beobachtet werden, der zur positiven Selektion der niedrig amplifizierten Sublinien führte (183-185). Dieser Effekt war reversibel und es konnte eine stetige Reamplifikation ohne den EGF Stimulus beobachtet werden. Betrachtete man das Verhalten der MDA Sublinien unter Zellkulturbedingungen, so konnte man feststellen, dass der Verlust an *EGFR* Kopien in der EGF-supprimierten Sublinie MDA(+) linear und mit einer vergleichbaren Geschwindigkeit verläuft, wie der Reamplifikationsprozess in den beiden parallel zueinander ansteigenden Sublinien MDA0106 und MDA0107. Es konnten keine sprunghaften Veränderungen in den Genkopienzahlen nachgewiesen werden. Dies in Kombination mit dem linearen Verlauf der drei Kurven lässt auf einen oder mehrere ablaufende Prozesse rückschließen, die mit einer vergleichbar seltenen Wahrscheinlichkeit auftreten. Die MDA-WT Linie war über die Zeit im Mittel konstant bei  $\sim 32$  Genkopien und schwankte maximal um  $\sim 3$  Kopien.

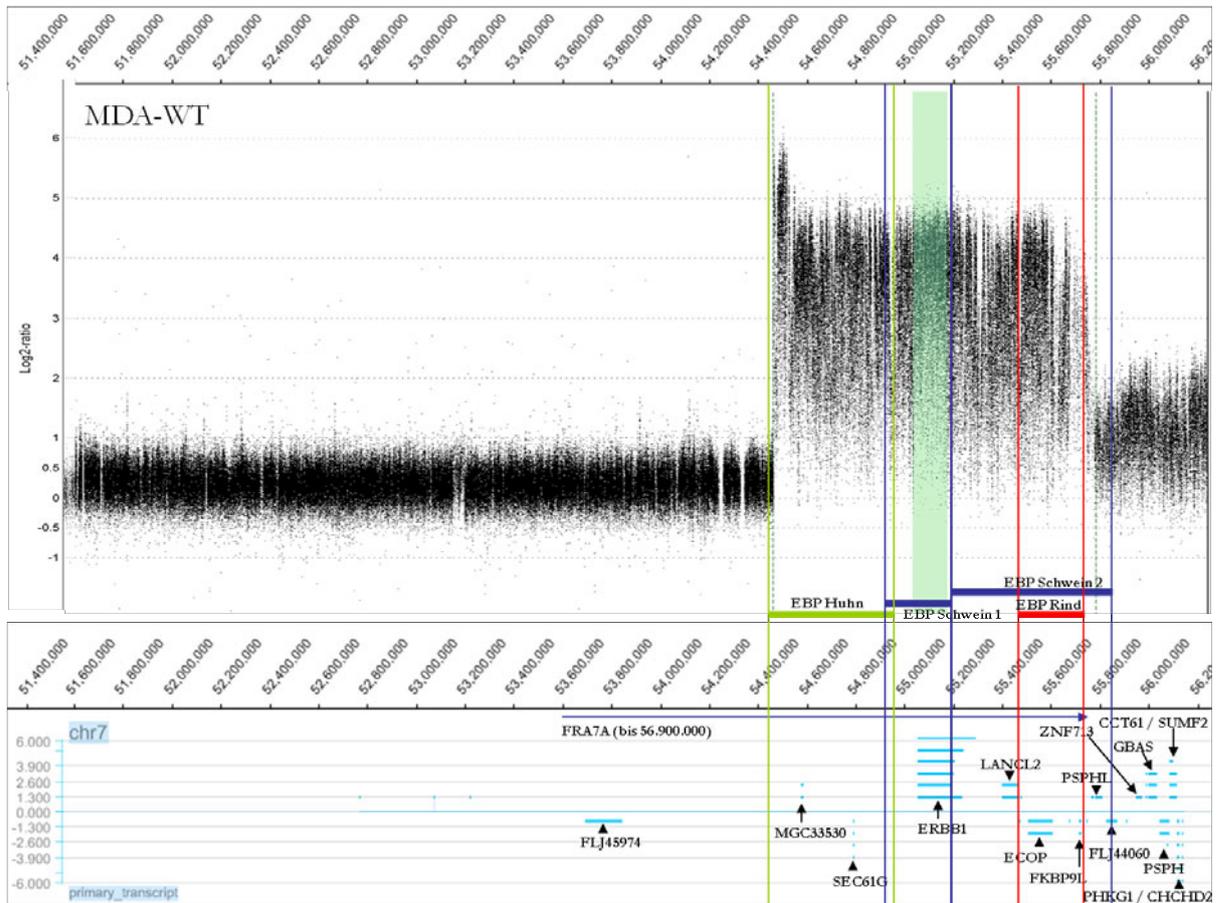
Sowohl die Heterogenität, der leiterartige Aufbau des Amplikons, der immer auf ein Chromosom pro Zelle beschränkt war, als auch der Verlust der Telomerregion von Chr.7p legten einen BFB-basierten Amplifikationsmechanismus nahe, der durch die Aktivierung der Fragile Site FRA7A ausgelöst werden könnte (125,177). Die *Fine Tiling Array CGHs* ermöglichten eine genaue Analyse von Struktur und der Länge der Amplikons. Betrachtet man die Sublinie MDA(+) als Ausgangsform der Genamplifikation, so konnte zwischen den Sublinien ein mehrstufiger Verlauf des Amplifikationsprozesses abgeleitet werden. Die, für die Sublinien anhand der Array-CGHs erhobenen, strukturellen Grenzen und Regionen der Zugewinne wurden durch qPCR validiert. Die Zusammenfassung der so für die MDA Sublinien erhobenen Daten wurde in Abb.4.1 zusammengefasst.



**Abb.4.1: Evolution des *EGFR* Amplikons in MDA Sublinien.** Proportional zueinander dargestellt wurden die Längen der amplifizierten Bereiche in den MDA Sublinien. Einzelne Strukturen sind in identischer Farbe für die jeweilige Sublinie unterlegt. Leicht amplifizierte Bereiche wurden abgedunkelt dargestellt. Des Weiteren wurden die Positionen der Grenzen auf Chr.7 eingetragen, sowie die Lokalisation des *EGFR*-Gens innerhalb der Amplikons. Es wurde zusätzlich eine Einteilung in die Bereiche A, B1+2, sowie C vorgenommen, die den regionsbezogenen Zugewinnen des Amplifikationsprozesses entsprechen.

Das Amplikon der MDA Sublinien zeigte mit ansteigender Genkopienzahl eine Zunahme an Länge, welche nicht linear zu verlaufen schien und die Unterteilung des Amplikons in mehrere Abschnitte (A-C) implizierte. Die telomerwärtige Ausbreitung des Amplikons beschränkt sich auf den Bereich B1. Dieser zeigt bereits in MDA0107 seine volle Länge und eine Ausbildung einer Peak-ähnlichen Struktur direkt am Startpunkt des Amplikons. Ab dem Amplifikationsstatus von MDA0106 erreicht der Bereich das Niveau des Bereichs A. Centromerwärtig lassen sich parallel dazu Zugewinne des Bereichs B2 bereits für MDA0107 erkennen, die für im Amplikon von MDA-WT das Niveau von Bereich A erreichen. Ab dem Amplifikationsstatus von MDA0106 ist ein weiterer, amplifizierter Bereich C zu erkennen, der über die Grenzen des Arrays hinausreicht. Zusammenfassend konnte man sagen, dass es mit steigendem Amplifikationsgrad schnell zu der Ausbildung einer scharfen telomerwärtigen Amplikongrenze mit einer, das Hauptamplikon überragenden, Peak-artigen Struktur kam. Der zentrale Bereich des Amplikons wurde von Region A und B2 gebildet, die in zwei Stufen amplifizierte, wobei B2 der Region A folgte. Der centromerwärtige Bereich C wurde ab dem Amplifikationsgrad von MDA0106 ebenfalls erfasst. Es kam also zu einer schrittweisen, centromerwärtigen Ausdehnung des amplifizierten Bereichs. Alle Einzelzell-WGAs zeigten einen, mit dem MDA0106 oder MDA-WT assoziierten Amplikon-Status, was mit den qPCR-Ergebnissen korrelierte.

Beim Abgleich der ermittelten Amplikongrenzen und der Bereiche A-C mit der UCSC Datenbank fanden sich in den Amplikon und Bereichsgrenzen Anhäufungen von Homologien zu anderen chromosomalen Bereichen, die von repetitiven DNA Elementen wie LCRs, *Alu*-Sequenzen oder LINE-Elementen flankiert wurden. Pro Grenzbereich fanden sich zum Teil >50 homologe Bereiche zu anderen Chromosomen, die in ihrer Länge von ~200bps bis >10kb reichten und Homologien von >85% aufwiesen (RepBase). Desweiteren wurden durch Sequenzvergleiche verschiedener Spezies, bei denen evolutionär konservierte Umlagerungsregionen berechnet wurden (177), mehrere Bruchregionen gefunden, die mit den Amplikongrenzen der MDA-WT Sublinien korrelierten. Die im Bereich des MDA-WT Amplikons gelegenen, evolutionär bedingten Bruchpunkte (EBPs) sowie die Position der Fragile Site FRA7A und der in diesem Bereich lokalisierten Gene wurden in Abb.4.2 dargestellt.



**Abb.4.2: Übersicht über das MDA-WT Amplikon: Amplikon lokalisierte Gene und evolutionär bedingte Bruchpunkte (EBPs)** Dargestellt ist der Amplifikationsplot von MDA-WT auf Chr.7. Horizontal abzulesen ist die chromosomale Position (schwarze Zahlen), in der Region lokalisierte Gene (blaue Balken) und die Fragile-Site FRA7A (blauer Pfeil). Vertikal aufgetragen sind die log<sub>2</sub>-Ratios. Die vertikalen Linien beschreiben die zwischen Mensch und anderen Spezies berechneten evolutionär konservierten Umlagerungspositionen.

Die für MDA-WT als Hauptamplikon definierte Region beinhaltete neben *EGFR* noch 5 weitere Gene, die jedoch soweit beschrieben, keinen erkennbaren Zusammenhang mit der Entstehung von Karzinomen zeigten. Dies galt auch für weitere 9 Gene, die in dem centromerwärts liegenden leicht amplifizierten Bereich lokalisiert waren. Die durchgeführte Expressionsanalyse für *EGFR* (Bereich A) und *CHCHD2* (Bereich C) zeigte in den MDA-WT Sublinien im Bezug auf die Genkopienzahl für *CHCHD2* eine etwa linear, für *EGFR* eine eher exponentiell zunehmende Expression. Die Zelllinie BT-20 bewegte sich in ihrem *CHCHD2* Expressionsniveau etwa bei MDA0107, was gut mit den Array Daten korreliert. Die *EGFR* Expression war jedoch im Vergleich zu MDA0106 trotz geringerer Genkopien 2-fach erhöht. Die erhöhte *EGFR* Expression in BT-20 konnte unter Umständen durch einen, im Vergleich zu MDA-MB-468 verkürzten CA-Repeat im *CASSR1* bedingt sein (62,63). Die Zelllinie BT-20 zeigte ein homozygoten 113 Basen Fragment, das im Vergleich zu MDA-MB-468 mit einem homozygoten 117 Basen Fragment einer Verkürzung von 2 CAs entsprach.

Die Expressionsanalysen der MDA-MB-468 Sublinien bestätigten die positive Korrelation zwischen Expressionshöhenänderung und Genkopienzunahme, die neben *EGFR* (69) auch für das leichter amplifizierte *CHCHD2* Gen nachgewiesen werden konnte.

Sehr auffällig waren die präzisen Übereinstimmungen zwischen den Amplikongrenzen der MDA-WT Sublinien und den EBPs. So gab es exakte Überlappungen des Bereichs B1 mit einem EBP für Huhn, der Grenze zwischen Bereich A und B2 mit einem EBP für Schwein und eine weitere für den Grenzbereich zwischen B2 und C mit einem EBP für Rind. Für die ermittelten Grenzregionen der beiden *EGFR* amplifizierten Zelllinien A431 und BT-20, die auf den ersten Blick wenig Übereinstimmungen in Struktur, Position und Länge mit den Amplikons der MDA-WT Sublinien zeigten, konnten ebenfalls Übereinstimmungen mit EBPs nachgewiesen werden, die in Abb.4.3 zusammenfassend dargestellt wurden (177).

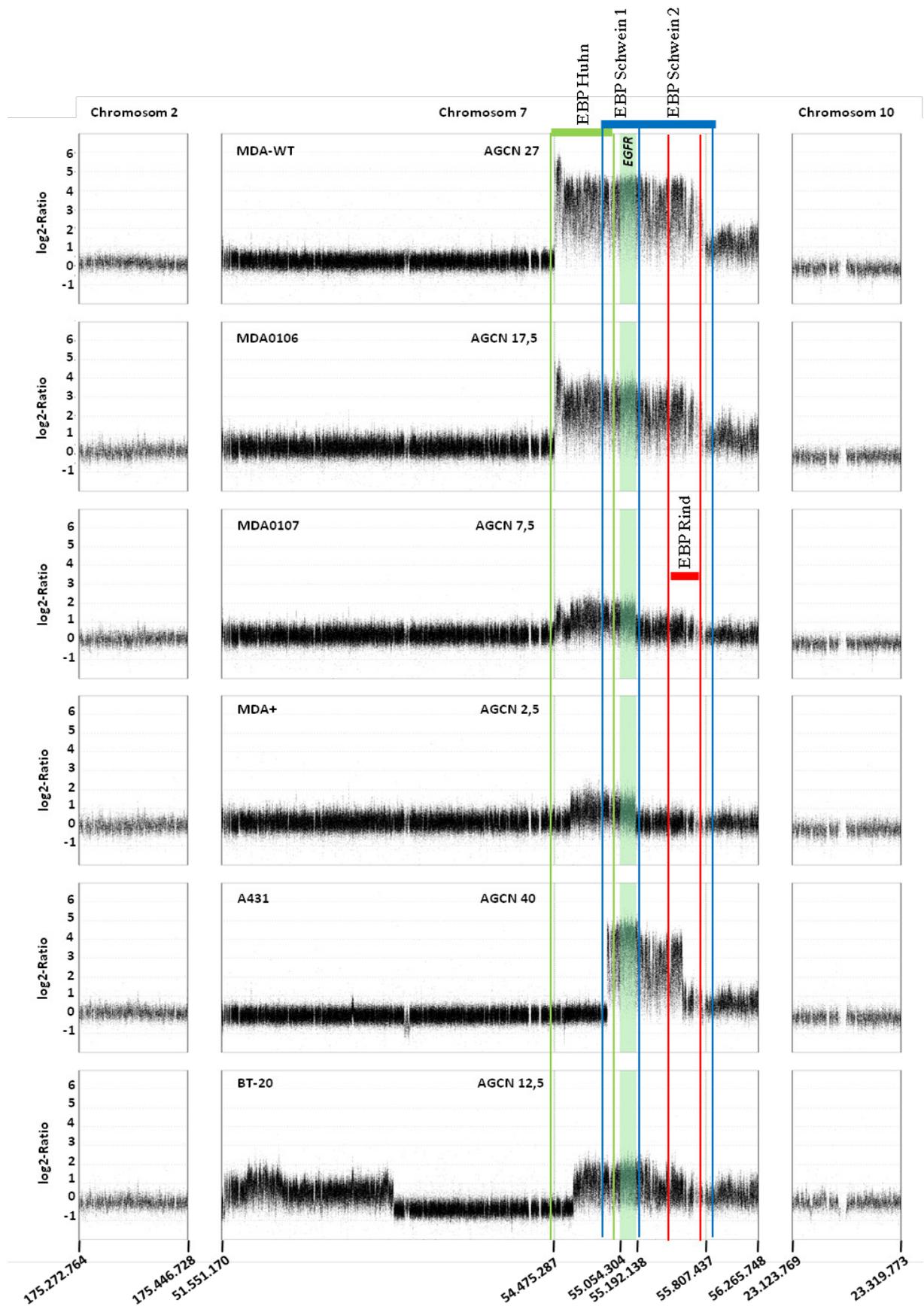
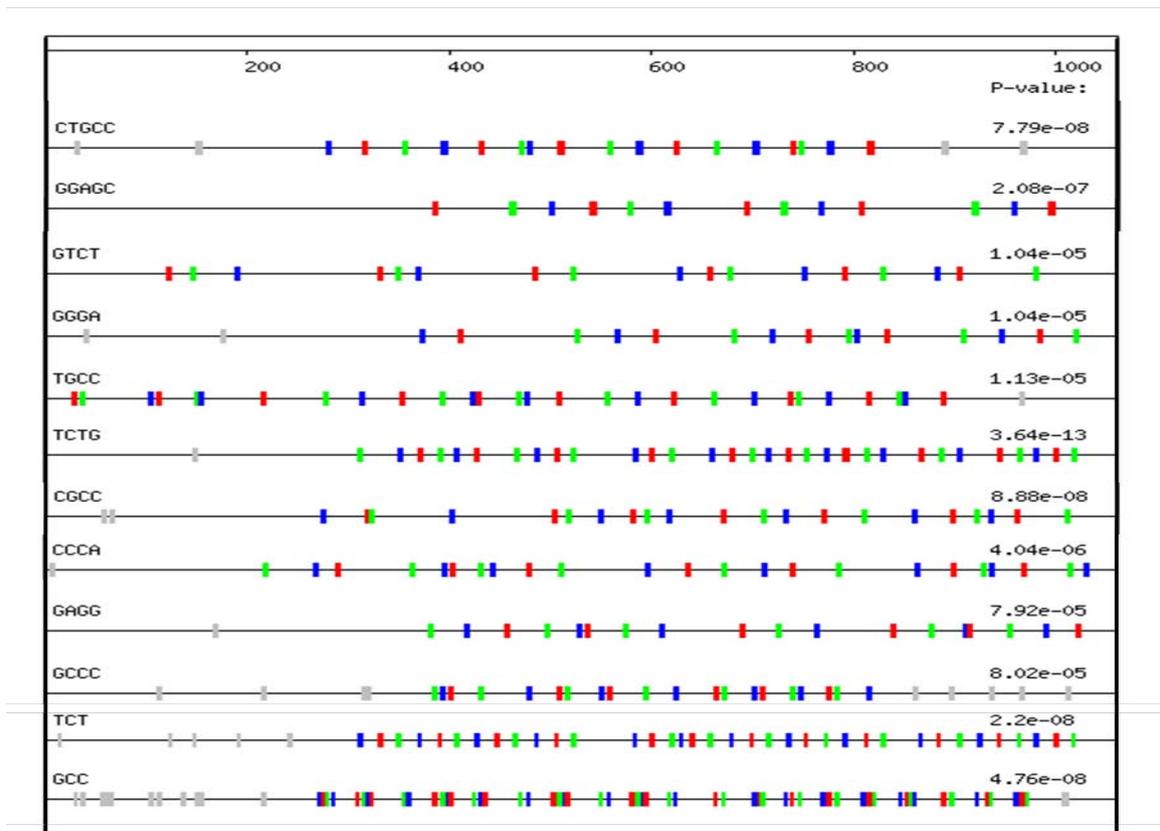


Abb.4.3: EBPs – Amplikongrenzen Korrelation der *EGFR* amplifizierten Linien Um den oben erläuterten Sachverhalt besser visuell erfassen zu können, wurden hier nochmal alle Plots der *EGFR* amplifizierten Linien in Kombination mit den EBPs dargestellt. Verändert nach Ruiz-Herrera et al. 2006 (177).

Zusätzlich zu den oben beschriebenen Überlappungen mit EBPs bei den MDA Sublinien, die mit Ausnahme der Grenze von Bereich B1 ebenso auf BT-20 zutreffen, konnten für A431 in dessen telomerwärtigen Amplikonstart eine Übereinstimmung mit einem EBP für Schwein / Huhn, sowie im centromerwärtigen Amplikonende Übereinstimmungen mit einem EBP für Rind beschrieben werden. Durch die Arrays konnten für die Zelllinien Hotspots für Start- bzw. Endpunkte der Amplifikationen gezeigt werden, die teilweise miteinander übereinstimmen. Eine Vielzahl an Überlappungen mit EBPs konnte ebenfalls demonstriert werden, die als Hinweis für evolutionär bedingte chromosomale Instabilität ausgelegt werden können. Für alle untersuchten Linien konnte gezeigt werden, dass *EGFR* in voller Länge im Amplikon eingeschlossen war. Desweiteren war allen Linien mit einer bestimmten Minimalgröße des Amplikon, die jenseits der Größe von MDA0107 lag, eine Amplifikation des, centromerwärtig vom Hauptamplikon, gelegenen Bereichs C gemein. Der Bereich A des Amplikons wurde bei der Untersuchung von metaplastischen Mamma-Cas mit *EGFR* Amplifikation, als minimal amplifizierter Bereich beschrieben (186). Einen gesondert zu betrachtenden Fall bildete BT-20, welche zwei Amplikons zeigte, die eine Deletion umschlossen. Der Aufbau der Amplikonstrukturen in BT-20 wurde ebenfalls als typisch für den BFB Mechanismus beschrieben, da es bei diesem häufig zu Deletionen zwischen amplifizierten DNA Bereichen kommen kann (177).

Die durch „*Long-Range*“-PCRs durchgeführte Suche nach Strangbruch-bedingten chromosomalen Umlagerungen im distinkt definierten telomerwärtigen Ende des MDA-WT Amplikon zeigte zwei verkürzte Segmente in einer repetitiven SVA\_A Sequenz, die der Familie der retroviralen SVA Elemente zuzuordnen war. Die um 300 bzw. 600 Basen verkürzten PCR-Fragmente der PCR-Primerkombination 5 traten nur in MDA0106 und MDA-WT auf. Eine auf der Software LTR-Digest basierte Analyse der SVA\_A Sequenz zeigte Homologien von >500 Basen zu der 4,8kb entfernt liegenden HERVK-Sequenz innerhalb der telomerwärtigen Startregion des MDA-WT Amplikon. Bei der Analyse der MDA Sublinien durch qPCR wurden in dem Bereich des HERVK-Elements für alle Sublinien die jeweils höchsten Amplifikationsgrade bestimmt. Eine Umlagerung durch aktive Transposition der retroviralen Sequenz ist jedoch auszuschließen, da zwar in Brusttumoren eine erhöhte Expression von HERVK nachgewiesen wurde, jedoch auf dem gesamten Genom keine intakte HERVK-Sequenz mehr erhalten ist (179,180). Die Nähe der telomerwärtigen Amplikongrenze zu der evolutionär instabilen HERVK-Sequenz ließ eine direkte Beteiligung am Amplifikationsprozess vermuten.

Diese Annahme wurde durch die differentiellen PCR-Fragmente innerhalb der nahegelegenen SVA\_A Sequenz unterstützt, die Homologien zu der HERVK-Sequenz aufweist. Die HERVK-Sequenz könnte, bei einem NAHR basierten Mechanismus als Rekombinations-site für andere Sequenzen dienen, wie z.B. die in unmittelbarer Nähe befindliche SVA\_A Sequenz. Der alternative Prozess des NHEJ unter Beteiligung der HERVK-Sequenz ist ebenso denkbar. Ausgelöst werden könnten beide Prozesse durch Regressionen in der Replikationsgeschwindigkeit an der Replikationsgabel bedingt durch repetitive Sequenzen und daraus resultierenden Einzel- oder Doppelstrangbrüchen. Ein Zusammenhang zwischen dem Auftreten bestimmter repetitiver DNA Motive und zahlreichen humanen hereditären Erkrankungen wurden impliziert (186,187). Es konnten zwar keine direkten Sequenzrepeats in der beteiligten HERVK-Sequenz nachgewiesen werden, jedoch eine hohe Dichte sich überlappender Repeats, die vermutlich ebenfalls zu Regressionen der Replikationsgeschwindigkeit an der Replikationsgabel führen können. Diese sind in Abb. 4.4 zusammenfassend dargestellt.



**Abb.4.4: Akkumulation der überlappenden repetitiven Sequenzmotiven des PCR5-Fragments** Dargestellt wurden die sich überlappenden repetitiven Sequenzmotive. Die Art des jeweiligen Repeats ist auf der linken Seite angegeben. Die farbigen Balken symbolisieren jeweils eine Kopie des Repeats. Die Farben dienen ausschließlich dazu die Repeats visuell trennen zu können. Den Farben selbst kommt keine weitere Bedeutung zu. Die p-Werte auf der rechten Seite sind ein Maß der statistischen Wahrscheinlichkeit des zufälligen Auftretens der jeweiligen Repeatfolge. Ein p-Wert von  $1e-5$  bedeutet, ein zufälliges auftreten in 10.000 Basen.

Die Akkumulation der unterschiedlichen repetitiven Sequenzmotive und die Höhe ihrer p-Werten verstärkt die Annahme, dass die Sequenz nicht zufällig entstanden ist, sondern häufig durch evolutionär bedingte Umlagerungen betroffen war.

Der Umstand, dass die Sequenzierung der differentiellen PCR-Fragmente im repetitiven Sequenzbereich nicht möglich war, unterstützt die Annahme der erschwerten Amplifizierbarkeit. Dieselben durch die DNA Sequenz bedingten Effekte, die der Polymerase innerhalb der Sequenzierungsreaktion eine Strangsynthese erschweren müssten sich auch auf die an der DNA Replikation beteiligten Polymerasen auswirken. Ein Mehrstufiger Amplifikationsprozess, wie ihn die erhobenen Daten nahe legen, könnte basierend auf der Lokalisierung der EBPs, chromosomalen Homologien und repetitiven DNA Elemente ebenfalls auf NAHR oder NHEJ Mechanismen zurückzuführen sein (Abb.1.10, 133). Ob die Amplifikation durch sequenzspezifische Mutation, sporadisch oder bedingt durch einen Defekt im DNA Reparaturapparat ausgelöst wurde konnte nicht ermittelt werden.

Zusammenfassend ließ sich festhalten, dass das Model des BFB basierten Amplifikationsprozesses durch Aktivierung der FRA7A erhärtet werden konnte. Die gehäuft an den Grenzen der Bereiche A-C auftretenden EBPs, Homologen chromosomalen Bereiche, sowie flankierende repetitive DNA Elemente sind alle innerhalb der Rare Fragile Site FRA7A lokalisiert.

Die Ergebnisse legen für MDA-MB-468 einen mehrstufigen Amplifikationsmechanismus nahe, an dem eine direkte Beteiligung einer HERVK- und einer SVA-Sequenz angenommen werden könnten. Für die SVA Sequenz konnten verkürzte, wahrscheinlich durch Umlagerungen der DNA entstandene verkürzte Sequenzabschnitte durch PCR nachgewiesen werden. Diese Sequenzabschnitte besaßen Homologien zu der HERVK-Sequenz, die unmittelbar im telomerwärtigen Startbereich des Amplikon lag und bei der Analyse durch qPCR die höchsten Amplifikationen des Bereichs zeigte. Die Umlagerungsprozesse werden mit hoher Wahrscheinlichkeit durch fehlgeschlagene NAHR oder NHEJ Mechanismen nach Replikation-bedingten Einzel- oder Doppelstrangbrüchen induziert.

### 4.3 Ausblick

In den nächsten Schritten könnte damit begonnen werden eine Bibliothek von DTC Einzelzell-WGAs aufzubauen, die zur Charakterisierung verschiedener genetischer Marker, in statistisch signifikanten Zellzahlen, für unterschiedliche Tumorentitäten genutzt wird. Ein vergleichendes Screening von Primärtumoren, Metastasen und der DTCs könnte neue relevante Informationen zum Disseminierungsprozess, Minimal Residual Diseases und neuen Zielen für Targeted Therapys erbringen. Bezogen auf die *EGFR* Amplifikation wäre es von großem Interesse neben dem Amplifikationsstatus ebenfalls die Grenzbereiche der Amplifikationen über Array-CGHs oder qPCR in DTCs, Primärtumoren und Metastasen zu charakterisieren, um eventuell, wie an dem differentiellen PCR5-Fragment in MDA-WT demonstriert, einfach per PCR nachweisbare Marker zu generieren. Nach der Analyse einer großen Anzahl *EGFR* amplifizierter DTCs wäre es z.B. denkbar, Marker zu entwickeln, die basierend auf bestimmten genetischen Prädispositionen, zur Vorhersage eines möglichen *EGFR* Amplifikationsrisikos genutzt werden könnten. Um den Amplifikationsprozess anhand der Modelzelllinie MDA-MB-468 weiter aufzuklären, wäre es naheliegend in den Start- und Endbereichen der Sublinien, die mit EBPs assoziiert sind, durch *Long-Range* PCR nach weiteren Bruchpunkten zu suchen und die betroffenen Bereiche auf Sequenzebene zu charakterisieren. Da das Amplikon von MDA(+) wahrscheinlich dem Ausgangszustand am ähnlichsten ist, sollte dabei mit den Grenzbereichen des MDA(+) Amplikons begonnen werden. Durch den Einsatz von speziellen Polymerasen, die für lange PCR-Fragmente optimiert wurden, sollte versucht werden Produkte für die bislang fehlgeschlagenen *Long-Range* PCRs in der telomerwärtigen Startregion des MDA-WT Amplikon zu amplifizieren. Diese könnten weitere wichtige Informationen zur Beteiligung der HERVK-Sequenz am Prozess der Genamplifikation beitragen.

**5. Literaturverzeichnis**

1. Arbeitsgemeinschaft Bevölkerungsbezogener Krebsregister in Deutschland in Zusammenarbeit mit dem Robert-Koch-Institut (2008)
2. Hunter T. Tyrosine phosphorylation in cell signaling and disease. *Keio J Med.* 2002 Jun;51(2):61-71. PMID: 12125907
3. Pawson T, Nash P. Protein-protein interactions define specificity in signal transduction. *Genes Dev.* 2000 May 1;14(9):1027-47. PMID: 10809663
4. Reed JC. Dysregulation of apoptosis in cancer. *J Clin Oncol.* 1999 Sep;17(9):2941-53. Review. PMID: 10561374
5. Hanahan D, Weinberg RA. The hallmarks of cancer. *Cell.* 2000 Jan 7;100(1):57-70. PMID: 10647931
6. Spittler A, Oehler R, Goetzinger P, Holzer S, Reissner CM, Leutmezer F, Rath V, Wrba F, Fuegger R, Boltz-Nitulescu G, Roth E. Low glutamine concentrations induce phenotypical and functional differentiation of U937 myelomonocytic cells. *J Nutr.* 1997 Nov;127(11):2151-7. PMID: 9349841
7. Klein G. Oncogenes and tumor suppressor genes. *Acta Oncol.* 1988;27(4):427-37. Review. PMID: 2849463
8. Yamamoto T, Toyoshima K. Neoplasma induced by retrovirus and cellular oncogene *Nippon Rinsho.* 1982;40(8):1880-7. PMID: 6292551
9. Kinzler KW, Vogelstein B. Lessons from hereditary colorectal cancer. *Cell.* 1996 Oct 18;87(2):159-70. PMID: 8861899
10. Vogt PK, Bos TJ. jun: oncogene and transcription factor. *Adv Cancer Res.* 1990;55:1-35. PMID: 2166997
11. Miyagawa K. Genetic instability and cancer. *Int J Hematol.* 1998 Jan;67(1):3-14. Review. PMID: 9594439
12. Zeng X, Saunders WS. The *Saccharomyces cerevisiae* centromere protein Slk19p is required for two successive divisions during meiosis. *Genetics.* 2000 Jun;155(2):577-87. PMID: 10835382
13. Croce CM. Oncogenes and cancer. *N Engl J Med.* 2008 Jan 31;358(5):502-11. PMID: 18234754
14. Arteaga CL. Epidermal growth factor receptor dependence in human tumors: more than just expression? *Oncologist* 2002;7:Suppl 4:31-39. PMID: 12202786
15. Davies H, Bignell GR, Cox C, et al. Mutations of the BRAF gene in human cancer. *Nature* 2002;417:949-954. PMID: 12068308
16. Frattini M, Ferrario C, Bressan P, et al. Alternative mutations of BRAF, RET and NTRK1 are associated with similar but distinct gene expression patterns in papillary thyroid cancer. *Oncogene* 2004;23:7436-7440. PMID: 15273715
17. Konopka JB, Watanabe SM, Singer JW, Collins SJ, Witte ON. Cell lines and clinical isolates derived from Ph1-positive chronic myelogenous leukemia patients express c-abl proteins with a common structural alteration. *Proc Natl Acad Sci U S A* 1985;82:1810-1814. PMID: 3856862
18. Tomlins SA, Rhodes DR, Perner S, et al. Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. *Science* 2005;310:644-648. PMID: 16254181
19. Tsujimoto Y, Yunis J, Onorato-Showe L, Nowell PC, Croce CM. Molecular cloning of the chromosomal breakpoint of B cell lymphomas and leukemias with the t(11;14) chromosome translocation. *Science* 1984;224:1403-1406. PMID: 6610211
20. Tsujimoto Y, Cossman J, Jaffe E, Croce CM. Involvement of the bcl-2 gene in human follicular lymphoma. *Science* 1985;228:1440-1443. PMID: 3874430

21. Dent R, Trudeau M, Pritchard KI, Hanna WM, Kahn HK, Sawka CA, Lickley LA, Rawlinson E, Sun P, Narod SA. Triple-negative breast cancer: clinical features and patterns of recurrence. *Clin Cancer Res*. 2007 Aug 1;13(15 Pt 1):4429-34. PMID: 17671126
22. Robinson DR, Wu YM, Lin SF. The protein tyrosine kinase family of the human genome. *Oncogene*. 2000 Nov 20;19(49):5548-57. PMID: 11114734
23. Gschwind A, Fischer O.M, Ullrich A. The discovery of receptor tyrosine kinases: targets for cancer therapy. *Nature Reviews Cancer* 4, 361-370 (May 2004).
24. Carraway KL 3rd, Cantley LC. A new acquaintance for erbB3 and erbB4: a role for receptor heterodimerization in growth signaling. *Cell*. 1994 Jul 15;78(1):5-8. Review. PMID: 8033211
25. van der Geer P, Hunter T, Lindberg RA. Receptor protein-tyrosine kinases and their signal transduction pathways. *Annu Rev Cell Biol*. 1994;10:251-337. Review. PMID: 7888178
26. Heldin CH. Dimerization of cell surface receptors in signal transduction. *Cell*. 1995 Jan 27;80(2):213-23. Review. PMID: 7834741
27. Blume-Jensen P, Hunter T. Oncogenic kinase signalling. *Nature*. 2001 May 17;411(6835):355-65. Review. PMID: 11357143
28. Walker LC, Ganesan TS, Dhut S, Gibbons B, Lister TA, Rothbard J, Young BD. Novel chimaeric protein expressed in Philadelphia positive acute lymphoblastic leukaemia. *Nature*. 1987 Oct 29-Nov 4;329(6142):851-3. PMID: 3478595
29. Buerger H, Otterbach F, Simon R, Schäfer KL, Poremba C, Diallo R, Brinkschmidt C, Dockhorn-Dwornczak B, Boecker W. Different genetic pathways in the evolution of invasive breast cancer are associated with distinct morphological subtypes. *J Pathol*. 1999 Dec;189(4):521-6. PMID: 10629552
30. Ignar-Trowbridge DM, Nelson KG, Bidwell MC, Curtis SW, Washburn TF, McLachlan JA, Korach KS. Coupling of dual signaling pathways: epidermal growth factor action involves the estrogen receptor. *Proc Natl Acad Sci U S A*. 1992 May 15;89(10):4658-62. PMID: 158480
31. Pérez-Soler R. HER1/EGFR targeting: refining the strategy. *Oncologist*. 2004;9(1):58-67. Review. PMID: 14755015
32. Downward J, Yarden Y, Mayes E, Scrace G, Totty N, Stockwell P, Ullrich A, Schlessinger J, Waterfield MD. Close similarity of epidermal growth factor receptor and v-erb-B oncogene protein sequences. *Nature*. 1984 Feb 9-15;307(5951):521-7. PMID: 6320011
33. Jimenez C, Jones DR, Rodríguez-Viciana P, Gonzalez-García A, Leonardo E, Wennström S, von Kobbe C, Toran JL, R-Borlodo L, Calvo V, Copin SG, Albar JP, Gaspar ML, Diez E, Marcos MA, Downward J, Martínez-A C, Mérida I, Carrera AC. Identification and characterization of a new oncogene derived from the regulatory subunit of phosphoinositide 3-kinase. *EMBO J*. 1998 Feb 2;17(3):743-53. PMID: 9450999
34. Aloy I, Yarden Y. The ErbB signaling network in embryogenesis and oncogenesis: signal diversification through combinatorial ligand-receptor interactions. *FEBS Lett*. 1997 Jun 23;410(1):83-6. Review. PMID: 9247128
35. Mendelsohn J, Baselga J. The EGF receptor family as targets for cancer therapy. *Oncogene*. 2000 Dec 27;19(56):6550-65. Review. PMID: 11426640
36. Vlahovic G, Crawford J. Activation of tyrosine kinases in cancer. *Oncologist*. 2003;8(6):531-8. PMID: 14657531
37. Dacic S. EGFR assays in lung cancer. *Adv Anat Pathol*. 2008 Jul;15(4):241-7. PMID: 18580100
38. Ullrich A, Coussens L, Hayflick JS, Dull TJ, Gray A, Tam AW, Lee J, Yarden Y, Libermann TA, Schlessinger J, et al. Human epidermal growth factor receptor cDNA sequence and aberrant expression of the amplified gene in A431 epidermoid carcinoma cells. *Nature*. 1984 May 31-Jun 6;309(5967):418-25. PMID: 6328312

39. Pratt RM, Kim CS, Grove RI. Role of glucocorticoids and epidermal growth factor in normal and abnormal palatal development. *Curr Top Dev Biol.* 1984;19:81-101. Review. PMID: 6094115
40. Carpenter G. Receptors for epidermal growth factor and other polypeptide mitogens. *Annu Rev Biochem.* 1987;56:881-914. Review. PMID: 3039909
41. Gill GN, Bertics PJ, Santon JB. Epidermal growth factor and its receptor. *Mol Cell Endocrinol.* 1987 Jun;51(3):169-86. Review. PMID: 3109978
42. Khazaie K, Schirmacher V, Lichtner RB. EGF receptor in neoplasia and metastasis. *Cancer Metastasis Rev.* 1993 Sep;12(3-4):255-74. Review. PMID: 8281612
43. Browne CA. Epidermal growth factor and transforming growth factor alpha. *Baillieres Clin Endocrinol Metab.* 1991 Dec;5(4):553-69. PMID: 1755808
44. Wells A. The epidermal growth factor receptor and its ligands. *Cancer Treat Res.* 1989;47:143-68. PMID: 2576995
45. Earp HS, Dawson TL, Li X, Yu H. Heterodimerization and functional interaction between EGF receptor family members: a new signaling paradigm with implications for breast cancer research. *Breast Cancer Res Treat.* 1995 Jul;35(1):115-32. Review. PMID: 7612898
46. Stoscheck CM, King LE Jr. Role of epidermal growth factor in carcinogenesis. *Cancer Res.* 1986 Mar;46(3):1030-7. Review. PMID: 3002608
47. Gullick WJ, Downward J, Waterfield MD. Antibodies to the autophosphorylation sites of the epidermal growth factor receptor protein-tyrosine kinase as probes of structure and function. *EMBO J.* 1985 Nov;4(11):2869-77. PMID: 2415353
48. Martín-Nieto J, Villalobo A. The human epidermal growth factor receptor contains a juxtamembrane calmodulin-binding site. *Biochemistry.* 1998 Jan 6;37(1):227-36. PMID: 9425043
49. Hunter T. The epidermal growth factor receptor gene and its product. *Nature.* 1984 Oct 4-10;311(5985):414-6. PMID: 6090939
50. Gullick WJ, Downward J, Foulkes JG, Waterfield MD. Antibodies to the ATP-binding site of the human epidermal growth factor (EGF) receptor as specific inhibitors of EGF-stimulated protein-tyrosine kinase activity. *Eur J Biochem.* 1986 Jul 15;158(2):245-53. PMID: 3015611
51. Cadena DL, Gill GN. Receptor tyrosine kinases. *FASEB J.* 1992 Mar;6(6):2332-7. PMID: 1312047
52. Eley GD, Reiter JL, Pandita A, Park S, Jenkins RB, Mähle NJ, James CD. A chromosomal region 7p11.2 transcript map: its development and application to the study of EGFR amplicons in glioblastoma. *Neuro Oncol.* 2002 Apr;4(2):86-94. PMID: 11916499
53. Zarnegar R, Michalopoulos G. Purification and biological characterization of human hepatopoietin A, a polypeptide growth factor for hepatocytes. *Cancer Res.* 1989 Jun 15;49(12):3314-20. PMID: 2524251
54. Ekstrand AJ, Sugawa N, James CD, Collins VP. Amplified and rearranged epidermal growth factor receptor genes in human glioblastomas reveal deletions of sequences encoding portions of the N- and/or C-terminal tails. *Proc Natl Acad Sci U S A.* 1992 May 15;89(10):4309-13. PMID: 1584765
55. Reiter JL, Threadgill DW, Eley GD, Strunk KE, Danielsen AJ, Sinclair CS, Pearsall RS, Green PJ, Yee D, Lampland AL, Balasubramaniam S, Crossley TD, Magnuson TR, James CD, Mähle NJ. Comparative genomic sequence analysis and isolation of human and mouse alternative EGFR transcripts encoding truncated receptor isoforms. *Genomics.* 2001 Jan 1;71(1):1-20. PMID: 11161793
56. Callaghan T, Antczak M, Flickinger T, Raines M, Myers M, Kung HJ. A complete description of the EGF-receptor exon structure: implication in oncogenic activation and domain evolution. *Oncogene.* 1993 Nov;8(11):2939-48. PMID: 8414496

57. Velu TJ, Beguinot L, Vass WC, Willingham MC, Merlino GT, Pastan I, Lowy DR. Epidermal-growth-factor-dependent transformation by a human EGF receptor proto-oncogene. *Science*. 1987 Dec 4;238(4832):1408-10. PMID: 3500513
58. Slamon DJ, Godolphin W, Jones LA, Holt JA, Wong SG, Keith DE, Levin WJ, Stuart SG, Udove J, Ullrich A, et al. Studies of the HER-2/neu proto-oncogene in human breast and ovarian cancer. *Science*. 1989 May 12;244(4905):707-12. PMID: 2470152
59. Ishii S, Xu YH, Stratton RH, Roe BA, Merlino GT, Pastan I. Characterization and sequence of the promoter region of the human epidermal growth factor receptor gene. *Proc Natl Acad Sci U S A*. 1985 Aug;82(15):4920-4. PMID: 2991899
60. Maekawa T, Imamoto F, Merlino GT, Pastan I, Ishii S. Cooperative function of two separate enhancers of the human epidermal growth factor receptor proto-oncogene. *J Biol Chem*. 1989 Apr 5;264(10):5488-94. PMID: 2925616
61. Chi DD, Hing AV, Helms C, Steinbrueck T, Mishra SK, Donis-Keller H. Two chromosome 7 dinucleotide repeat polymorphisms at gene loci epidermal growth factor receptor (EGFR) and pro alpha 2 (I) collagen (COL1A2). *Hum Mol Genet*. 1992 May;1(2):135. PMID: 1301150
62. Gebhardt F, Zänker KS, Brandt B. Modulation of epidermal growth factor receptor gene transcription by a polymorphic dinucleotide repeat in intron 1. *J Biol Chem*. 1999 May 7;274(19):13176-80. PMID: 10224073
63. Gebhardt F, Bürger H, Brandt B. Modulation of EGFR gene transcription by a polymorphic repetitive sequence--a link between genetics and epigenetics. *Int J Biol Markers*. 2000 Jan-Mar;15(1):105-10. Review. PMID: 10763151
64. Gebhardt F, Bürger H, Brandt B. Modulation of EGFR gene transcription by secondary structures, a polymorphic repetitive sequence and mutations--a link between genetics and epigenetics. *Histol Histopathol*. 2000 Jul;15(3):929-36. Review. PMID: 10963135
65. Buerger H, Gebhardt F, Schmidt H, Beckmann A, Hutmacher K, Simon R, Lelle R, Boecker W, Brandt B. Length and loss of heterozygosity of an intron 1 polymorphic sequence of egfr is related to cytogenetic alterations and epithelial growth factor receptor expression. *Cancer Res*. 2000 Feb 15;60(4):854-7. PMID: 10706093
66. Ro J, North SM, Gallick GE, Hortobagyi GN, Gutterman JU, Blick M. Amplified and overexpressed epidermal growth factor receptor gene in uncultured primary human breast carcinoma. *Cancer Res*. 1988 Jan 1;48(1):161-4. PMID: 3334990
67. Brandt B, Vogt U, Harms F, Bosse U, Zänker KS, Assmann G. Double-differential PCR for gene dosage estimation of erbB oncogenes in benign and cancer tissues and comparison to cellular DNA content. *Gene*. 1995 Jun 14;159(1):29-34. PMID: 7607569
68. Werkmeister R, Brandt B, Joos U. The erbB oncogenes as prognostic markers in oral squamous cell carcinomas. *Am J Surg*. 1996 Dec;172(6):681-3. PMID: 8988677
69. Tidow N, Boecker A, Schmidt H, Agelopoulos K, Boecker W, Buerger H, Brandt B. Distinct amplification of an untranslated regulatory sequence in the egfr gene contributes to early steps in breast cancer development. *Cancer Res*. 2003 Mar 15;63(6):1172-8. PMID: 12649171
70. Richards RI. Fragile and unstable chromosomes in cancer: causes and consequences. *Trends Genet*. 2001 Jun;17(6):339-45. Review. PMID: 11377796
71. Chen AT, Reidy JA, Annett JL, Welty TK, Zhou HG. Increased chromosome fragility as a consequence of blood folate levels, smoking status, and coffee consumption. *Environ Mol Mutagen*. 1989;13(4):319-24. PMID: 2737183
72. Coquelle A, Pipiras E, Toledo F, Buttin G, Debatisse M. Expression of fragile sites triggers intrachromosomal mammalian gene amplification and sets boundaries to early amplicons. *Cell*. 1997 Apr 18;89(2):215-25. PMID: 9108477

73. Yunis JJ, Soreng AL. Constitutive fragile sites and cancer. *Science*. 1984 Dec 7;226(4679):1199-204. PMID: 6239375
74. Ceçener G, Egeli U, Taşdelen I, Tunca B, Duman H, Kizil A. Common fragile site expression and genetic predisposition to breast cancer. *Teratog Carcinog Mutagen*. 1998;18(6):279-91. PMID: 10052563
75. Richards RI. Dynamic mutations: a decade of unstable expanded repeats in human genetic disease. *Hum Mol Genet*. 2001 Oct 1;10(20):2187-94. Review. PMID: 11673400
76. McCLINTOCK B. Chromosome organization and genic expression. *Cold Spring Harb Symp Quant Biol*. 1951;16:13-47. PMID: 14942727
77. Shuster MI, Han L, Le Beau MM, Davis E, Sawicki M, Lese CM, Park NH, Colicelli J, Gollin SM. A consistent pattern of RIN1 rearrangements in oral squamous cell carcinoma cell lines supports a breakage-fusion-bridge cycle model for 11q13 amplification. *Genes Chromosomes Cancer*. 2000 Jun;28(2):153-63. PMID: 10825000
78. Hellman A, Zlotorynski E, Scherer SW, Cheung J, Vincent JB, Smith DI, Trakhtenbrot L, Kerem B. A role for common fragile site induction in amplification of human oncogenes. *Cancer Cell*. 2002 Feb;1(1):89-97. PMID: 12086891
79. Agelopoulos K, Tidow N, Korsching E, Voss R, Hinrichs B, Brandt B, Boecker W, Buerger H. Molecular cytogenetic investigations of synchronous bilateral breast cancer. *J Clin Pathol*. 2003 Sep;56(9):660-5. PMID: 12944548
80. Aten JA, Stap J, Krawczyk PM, van Oven CH, Hoebe RA, Essers J, Kanaar R. Dynamics of DNA double-strand breaks revealed by clustering of damaged chromosome domains. *Science*. 2004 Jan 2;303(5654):92-5. PMID: 14704429
81. Tönnies H. Modern molecular cytogenetic techniques in genetic diagnostics. *Trends Mol Med*. 2002 Jun;8(6):246-50. PMID: 12067603
82. Emanuel BS, Shaikh TH. Segmental duplications: an 'expanding' role in genomic instability and disease. *Nat Rev Genet*. 2001 Oct;2(10):791-800. Review. PMID: 11584295
83. Lupski JR. Genomic disorders: structural features of the genome can lead to DNA rearrangements and human disease traits. *Trends Genet*. 1998 Oct;14(10):417-22. Review. PMID: 9820031
84. Lupski JR, Stankiewicz P. Genomic disorders: molecular mechanisms for rearrangements and conveyed phenotypes. *PLoS Genet*. 2005 Dec;1(6):e49. Review. PMID: 16444292
85. Landers ES et al Initial sequencing and analysis of the human genome. *Nature*. 2001 Feb 15;409(6822):860-921. Erratum in: *Nature* 2001 Aug 2;412(6846):565. *Nature* 2001 Jun 7;411(6838):720. Szustakowski, J [corrected to Szustakowski, J]. PMID: 11237011
86. Bailey JA, Gu Z, Clark RA, Reinert K, Samonte RV, Schwartz S, Adams MD, Myers EW, Li PW, Eichler EE. Recent segmental duplications in the human genome. *Science*. 2002 Aug 9;297(5583):1003-7. PMID: 12169732
87. Thomas NS, Durkie M, Potts G, Sandford R, Van Zyl B, Youings S, Dennis NR, Jacobs PA. Parental and chromosomal origins of microdeletion and duplication syndromes involving 7q11.23, 15q11-q13 and 22q11. *Eur J Hum Genet*. 2006 Jul;14(7):831-7. Epub 2006 Apr 12. PMID: 16617304
88. Naume B, Zhao X, Synnestvedt M, et al. Presence of bone marrow micrometastasis is associated with different recurrence risk within molecular subtypes of breast cancer. *Molecular oncology* 2007:1-12. PMID: 19383292
89. Shaw CJ, Lupski JR. Implications of human genome architecture for rearrangement-based disorders: the genomic basis of disease. *Hum Mol Genet*. 2004 Apr 1;13 Spec No 1:R57-64. Epub 2004 Feb 5. PMID: 14764619

90. Nemoto T, Vana J, Bedwani RN, Baker HW, McGregor FH, Murphy GP. Management and survival of female breast cancer: results of a national survey by the American College of Surgeons. *Cancer*. 1980 Jun 15;45(12):2917-24. PMID: 7388735
91. Carter CL, Allen C, Henson DE. Relation of tumor size, lymph node status, and survival in 24,740 breast cancer cases. *Cancer*. 1989 Jan 1;63(1):181-7. PMID: 2910416
92. Pantel K, Brakenhoff RH. Dissecting the metastatic cascade. *Nat Rev Cancer*. 2004 Jun;4(6):448-56. Review. PMID: 15170447
93. Hart IR, Saini A. Biology of tumour metastasis. *Lancet* 1992;339:1453-7. PMID: 1376386
94. Nabeshima K, Inoue T, Shimao Y, Sameshima T. Matrix metalloproteinases in tumor invasion: role for cell migration. *Pathol Int* 2002;52:255-64. PMID: 12031080
95. Salon C, Lantuejoul S, Eymin B, Gazzeri S, Brambilla C, Brambilla E. The E-cadherin/beta-catenin complex and its implication in lung cancer progression and prognosis. *Future Oncol* 2005;1:649-60. PMID: 16556042
96. Egeblad M, Werb Z. New functions for the matrix metalloproteinases in cancer progression. *Nat Rev Cancer* 2002;2:161-74. PMID: 11990853
97. Frisch SM, Francis H. Disruption of epithelial cell-matrix interactions induces apoptosis. *J Cell Biol* 1994;124:619-26. PMID: 8106557
98. Orr FW, Wang HH. Tumor cell interactions with the microvasculature: a rate-limiting step in metastasis. *Surg Oncol Clin N Am* 2001;10:357-81, ix-x. PMID: 11382592
99. Chambers AF, Groom AC, MacDonald IC. Dissemination and growth of cancer cells in metastatic sites. *Nat Rev Cancer* 2002;2:563-72. PMID: 12154349
100. Uhr JW, Scheuermann RH, Street NE, Vitetta ES. Cancer dormancy: opportunities for new therapeutic approaches. *Nat Med* 1997;3:505-9. PMID: 9142117
101. Fidler IJ, Kripke ML. Metastasis results from preexisting variant cells within a malignant tumor. *Science* 1977;197:893-5. PMID: 887927
102. Bernards R, Weinberg RA. A progression puzzle. *Nature* 2002;418:823. PMID: 12192390
103. Eberhard A, Kahlert S, Goede V, Hemmerlein B, Plate KH, Augustin HG. Heterogeneity of angiogenesis and blood vessel maturation in human tumors: implications for antiangiogenic tumor therapies. *Cancer Res* 2000;60:1388-93. PMID: 10728704
104. Fidler IJ. Metastasis: quantitative analysis of distribution and fate of tumor emboli labeled with 125 I-5-iodo-2'-deoxyuridine. *J Natl Cancer Inst* 1970;45:773-82. PMID: 5513503
105. Liotta LA, Kohn E. Cancer invasion and metastases. *Jama* 1990;263:1123-6. PMID: 8491804
106. Pantel K, Cote RJ, Fodstad O. Detection and clinical importance of micrometastatic disease. *J Natl Cancer Inst* 1999;91:1113-24. PMID: 10393719
107. Dearnaley DP, Sloane JP, Ormerod MG, et al. Increased detection of mammary carcinoma cells in marrow smears using antisera to epithelial membrane antigen. *Br J Cancer* 1981;44:85-90. PMID: 7020739
108. Pantel K, Izbicki JR, Angstwurm M, et al. Immunocytological detection of bone marrow micrometastasis in operable non-small cell lung cancer. *Cancer Res* 1993;53:1027-31. PMID: 7679945
109. Coombes RC, Dearnaley DP, Buckman R, et al. Detection of bone metastases in patients with breast cancer. *Invasion Metastasis* 1982;2:177-84. PMID: 6234254
110. Pantel K, Schlimok G, Angstwurm M, et al. Methodological analysis of immunocytochemical screening for disseminated epithelial tumor cells in bone marrow. *J Hematother* 1994;3:165-73. PMID: 7530132

111. Schlimok G, Funke I, Holzmann B, 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* 1987;84:8672-6. PMID: 2446326
112. Braun S, Pantel K, Muller P, 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* 2000;342:525-33. PMID: 10684910
113. Passlick B, Kubuschok B, Izbicki JR, Thetter O, Pantel K. Isolated tumor cells in bone marrow predict reduced survival in node-negative non-small cell lung cancer. *Ann Thorac Surg* 1999;68:2053-8. PMID: 10616976
114. Cote RJ, Beattie EJ, Chaiwun B, et al. Detection of occult bone marrow micrometastases in patients with operable lung carcinoma. *Ann Surg* 1995;222:415-23; discussion 23-5. PMID: 7574923
115. Ohgami A, Mitsudomi T, Sugio K, et al. Micrometastatic tumor cells in the bone marrow of patients with non-small cell lung cancer. *Ann Thorac Surg* 1997;64:363-7. PMID: 9262576
116. Sugio K, Kase S, Sakada T, et al. Micrometastasis in the bone marrow of patients with lung cancer associated with a reduced expression of E-cadherin and beta-catenin: risk assessment by immunohistochemistry. *Surgery* 2002;131:S226-31. PMID: 11821816
117. Kasimir-Bauer S, Schleucher N, Weber R, Neumann R, Seeber S. Evaluation of different markers in non-small cell lung cancer: prognostic value of clinical staging, tumour cell detection and tumour marker analysis for tumour progression and overall survival. *Oncol Rep* 2003;10:475-82. PMID: 12579292
118. Yasumoto K, Osaki T, Watanabe Y, Kato H, Yoshimura T. Prognostic value of cytokeratin-positive cells in the bone marrow and lymph nodes of patients with resected nonsmall cell lung cancer: a multicenter prospective study. *Ann Thorac Surg* 2003;76:194-201; discussion 02. PMID: 12842539
119. Braun S, Vogl FD, Naume B, Janni W, Osborne MP, Coombes RC, Schlimok G, Diel IJ, Gerber B, Gebauer G, Pierga JY, Marth C, Oruzio D, Wiedswang G, Solomayer EF, Kundt G, Strobl B, Fehm T, Wong GY, Bliss J, Vincent-Salomon A, Pantel K. A pooled analysis of bone marrow micrometastasis in breast cancer. *N Engl J Med*. 2005 Aug 25;353(8):793-802. PMID: 16120859
120. Woelfle U, Cloos J, Sauter G, et al. Molecular signature associated with bone marrow micrometastasis in human breast cancer. *Cancer Res* 2003;63:5679-84. PMID: 14522883
121. Bubendorf L, Muller P, Joos L, et al. Multitarget FISH analysis in the diagnosis of lung cancer. *Am J Clin Pathol* 2005;123:516-23. PMID: 15743749
122. Halling KC, Rickman OB, Kipp BR, Harwood AR, Doerr CH, Jett JR. A comparison of cytology and fluorescence in situ hybridization for the detection of lung cancer in bronchoscopic specimens. *Chest* 2006;130:694-701. PMID: 16963665
123. Brinkley, B.R., Beall, P.T., Wible, L.J., Mace, M.L., Turner, D.S., Cailleau, R.M. Variations in cell form and cytoskeleton in human breast carcinoma cells in vitro. *Cancer Res*, 40: 3118-3129, 1980. PubMed: 7000337
124. Giard DJ, et al. In vitro cultivation of human tumors: establishment of cell lines derived from a series of solid tumors. *J. Natl. Cancer Inst.* 51: 1417-1423, 1973. PubMed: 4357758
125. Vanhamme L, Szpirer C. Transforming activity of the human mammary line HBL100 DNA is associated with SV40 large T antigen genetic information integrated in its genome. *Carcinogenesis*. 1988 Apr;9(4):653-5. PubMed: 3281770
126. Dickson, R.B., Bates, S.E., Mc Manaway, M.E., Lippman, M.E. Characterization of estrogen responsive transforming activity in human breast cancer cell lines. *Cancer Res*, 46: 1707-1713, 1986. PMID: 2418952
127. J, editor. *Human tumor cells in vitro*. 43: New York: Plenum Press; 1975, pp. 115-159.
128. Fogh J, editor. *Human tumor cells in vitro*. 43: New York: Plenum Press; 1975, pp. 115-159.

129. Sugarman BJ , et al. Recombinant human tumor necrosis factor-alpha: effects on proliferation of normal and transformed cells in vitro. *Science* 230: 943-945, 1985. PubMed: 3933111
130. Moroni M, Veronese S, Benvenuti S, Marrapese G, Sartore-Bianchi A, Di Nicolantonio F, Gambacorta M, Siena S, Bardelli A. Gene copy number for epidermal growth factor receptor (EGFR) and clinical response to antiEGFR treatment in colorectal cancer: a cohort study. *Lancet Oncol.* 2005 May;6(5):279-86. PubMed: 15863375
131. Westphal M, Hänsel M, Hamel W, Kunzmann R, Hölzel F. Karyotype analyses of 20 human glioma cell lines. *Acta Neurochir (Wien).* 1994;126(1):17-26. PMID: 8154317
132. Pantel K, Schlimok G, Angstwurm M, et al. Methodological analysis of immunocytochemical screening for disseminated epithelial tumor cells in bone marrow. *J Hematother* 1994;3:165-73. PMID: 7530132
133. Braun S, Pantel K, Muller P, 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* 2000;342:525-33. PMID: 10684910
134. Cordell JL, Falini B, Erber WN, et al. Immunoenzymatic labeling of monoclonal antibodies using immune complexes of alkaline phosphatase and monoclonal anti-alkaline phosphatase (APAAP complexes). *J Histochem Cytochem* 1984;32:219-29. PMID: 6198355
135. Stigbrand T, Andres C, Bellanger L, et al. Epitope specificity of 30 monoclonal antibodies against cytokeratin antigens: the ISOBM TD5-1 Workshop. *Tumour Biol* 1998;19:132-52. PMID: 9486565
136. Bauer KD, de la Torre-Bueno J, Diel IJ, et al. Reliable and sensitive analysis of occult bone marrow metastases using automated cellular imaging. *Clin Cancer Res* 2000;6:3552-9. PMID: 10999743
137. Momburg F, Moldenhauer G, Hämmerling GJ, Möller P. Immunohistochemical study of the expression of a Mr 34,000 human epithelium-specific surface glycoprotein in normal and malignant tissues. *Cancer Res.* 1987 Jun 1;47(11):2883-91. PMID: 3552208
138. Allard WJ, Matera J, Miller MC, Repollet M, Connelly MC, Rao C, Tibbe AG, Uhr JW, Terstappen LW. Tumor cells circulate in the peripheral blood of all major carcinomas but not in healthy subjects or patients with nonmalignant diseases. *Clin Cancer Res.* 2004 Oct 15;10(20):6897-904. PMID: 15501967
139. Hartshorn C, Anshelevich A, Wangh LJ. Rapid, single-tube method for quantitative preparation and analysis of RNA and DNA in samples as small as one cell. *BMC Biotechnol.* 2005 Jan 13;5:2. PubMed: 15649321
140. Spits C, Le Caignec C, De Rycke M, Van Haute L, Van Steirteghem A, Liebaers I, Sermon K. Whole-genome multiple displacement amplification from single cells. *Nat Protoc.* 2006;1(4):1965-70. Epub 2006 Nov 30. PubMed: 17487184
141. Frank B. Dean, John R. Nelson, Theresa L. Giesler, and Roger S. Lasken Rapid Amplification of Plasmid and Phage DNA Using Phi29 DNA Polymerase and Multiply-Primed Rolling Circle Amplification *Genome Res.* 2001 June; 11(6): 1095–1099. doi: 10.1101/gr.180501. PMCID: PMC311129
142. Ferré F. Quantitative or semi-quantitative PCR: reality versus myth. *PCR Methods Applications* 1992;2:1–9. PMID: 1490169
143. R. G. Rutledge\* and C. Cote Mathematics of quantitative kinetic PCR and the application of standard curves *Nucleic Acids Research*, 2003, Vol. 31, No. 16 e93 PMID: 12907745
144. Michael Walter Pfaffl Real-time RT-PCR: Neue Ansätze zur exakten mRNA Quantifizierung. *BIOspektrum* · 1/04 · 10. Jahrgang
145. Bianca Holzapfel & Lucia Wickert *Die quantitative Real-Time-PCR (qRT-PCR)*. In: *Biologie in unserer Zeit*. 2007 Band 37, Nr. 2, S. 120-126.

146. Naylor TL, Greshock J, Wang Y, Colligon T, Yu QC, Clemmer V, Zaks TZ, Weber BL. High resolution genomic analysis of sporadic breast cancer using array-based comparative genomic hybridization. *Breast Cancer Res.* 2005;7(6):R1186-98. Epub 2005 Nov 24. PMID: 16457699
147. Eilers PH, de Menezes RX. Quantile smoothing of array CGH data. *Bioinformatics.* 2005 Apr 1;21(7):1146-53. Epub 2004 Nov 30. PMID: 15572474
148. Zhang L, Cui X, Schmitt K, Hubert R, Navidi W, Arnheim N. Whole genome amplification from a single cell: implications for genetic analysis. *Proc Natl Acad Sci U S A.* 1992 Jul 1;89(13):5847-51. PMID: 1631067
149. Telenius H, Carter NP, Bebb CE, Nordenskjöld M, Ponder BA, Tunnacliffe A. Degenerate oligonucleotide-primed PCR: general amplification of target DNA by a single degenerate primer. *Genomics.* 1992 Jul;13(3):718-25. PMID: 1639399
150. Klein CA, Schmidt-Kittler O, Schardt JA, Pantel K, Speicher MR, Riethmüller G. Comparative genomic hybridization, loss of heterozygosity, and DNA sequence analysis of single cells. *Proc Natl Acad Sci U S A.* 1999 Apr 13;96(8):4494-9. PMID: 10200290
151. Lucito R, Nakimura M, West JA, Han Y, Chin K, Jensen K, McCombie R, Gray JW, Wigler M. Genetic analysis using genomic representations. *Proc Natl Acad Sci U S A.* 1998 Apr 14;95(8):4487-92. PMID: 9539764
152. Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, et al. AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Res.* 1995 Nov 11;23(21):4407-14. PMID: 7501463
153. Tanabe C, Aoyagi K, Sakiyama T, Kohno T, Yanagitani N, Akimoto S, Sakamoto M, Sakamoto H, Yokota J, Ohki M, Terada M, Yoshida T, Sasaki H. Evaluation of a whole-genome amplification method based on adaptor-ligation PCR of randomly sheared genomic DNA. *Genes Chromosomes Cancer.* 2003 Oct;38(2):168-76. PMID: 12939744
154. Langmore JP. Rubicon Genomics, Inc. *Pharmacogenomics.* 2002 Jul;3(4):557-60. PMID: 12164778
155. Dean FB, Hosono S, Fang L, Wu X, Faruqi AF, Bray-Ward P, Sun Z, Zong Q, Du Y, Du J, Driscoll M, Song W, Kingsmore SF, Egholm M, Lasken RS. Comprehensive human genome amplification using multiple displacement amplification. *Proc Natl Acad Sci U S A.* 2002 Apr 16;99(8):5261-6. PMID: 11959976
156. Lovmar L, Syvänen AC. Multiple displacement amplification to create a long-lasting source of DNA for genetic studies. *Hum Mutat.* 2006 Jul;27(7):603-14. Review. PMID: 16156098
157. Moroni M, Veronese S, Benvenuti S, Marrapese G, Sartore-Bianchi A, Di Nicolantonio F, Gambacorta M, Siena S, Bardelli A. Gene copy number for epidermal growth factor receptor (EGFR) and clinical response to antiEGFR treatment in colorectal cancer: a cohort study. *Lancet Oncol.* 2005 May;6(5):279-86. PMID: 15863375
158. Lage, J. M., Leamon, J. H., Pejovic, T., et al. 2003, Whole genome analysis of genetic alterations in small DNA samples using hyperbranched strand displacement amplification and array-CGH, *Genome Res.*, 13, 294–307. PMID: 12566408
159. Wang G, Maher E, Brennan C, Chin L, Leo C, Kaur M, Zhu P, Rook M, Wolfe JL, Makrigiorgos GM. DNA amplification method tolerant to sample degradation. *Genome Res.* 2004 Nov;14(11):2357-66. PMID: 15520297
160. Lee CI, Leong SH, Png AE, Choo KW, Syn C, Lim DT, Law HY, Kon OL. An isothermal method for whole genome amplification of fresh and degraded DNA for comparative genomic hybridization, genotyping and mutation detection. *DNA Res.* 2006 Apr 30;13(2):77-88. Epub 2006 Mar 27. PMID: 16766515
161. Ballantyne KN, van Oorschot RA, Mitchell RJ. Comparison of two whole genome amplification methods for STR genotyping of LCN and degraded DNA samples. *Forensic Sci Int.* 2007 Feb 14;166(1):35-41. Epub 2006 May 9. PMID: 16687226

162. Wang G, Maher E, Brennan C, Chin L, Leo C, Kaur M, Zhu P, Rook M, Wolfe JL, Makrigiorgos GM. DNA amplification method tolerant to sample degradation. *Genome Res.* 2004 Nov;14(11):2357-66. PMID: 15520297
163. Cardoso J, Molenaar L, de Menezes RX, Rosenberg C, Morreau H, Möslein G, Fodde R, Boer JM. Genomic profiling by DNA amplification of laser capture microdissected tissues and array CGH. *Nucleic Acids Res.* 2004 Oct 28;32(19):e146. PMID: 15514107
164. Hellani A, Coskun S, Benkhalifa M, Tbakhi A, Sakati N, Al-Odaib A, Ozand P. Multiple displacement amplification on single cell and possible PGD applications. *Mol Hum Reprod.* 2004 Nov;10(11):847-52. Epub 2004 Oct 1. PMID: 15465849
165. Dickson PA, Montgomery GW, Henders A, Campbell MJ, Martin NG, James MR. Evaluation of multiple displacement amplification in a 5 cM STR genome-wide scan. *Nucleic Acids Res.* 2005 Jul 29;33(13):e119. PMID: 16055919
166. Lovmar L, Fredriksson M, Liljedahl U, Sigurdsson S, Syvänen AC. Quantitative evaluation by minisequencing and microarrays reveals accurate multiplexed SNP genotyping of whole genome amplified DNA. *Nucleic Acids Res.* 2003 Nov 1;31(21):e129. PMID: 14576329
167. Handyside AH, Robinson MD, Simpson RJ, Omar MB, Shaw MA, Grudzinskas JG, Rutherford A. Isothermal whole genome amplification from single and small numbers of cells: a new era for preimplantation genetic diagnosis of inherited disease. *Mol Hum Reprod.* 2004 Oct;10(10):767-72. Epub 2004 Aug 20. PMID: 15322224
168. Jiang Z, Zhang X, Deka R, Jin L. Genome amplification of single sperm using multiple displacement amplification. *Nucleic Acids Res.* 2005 Jun 7;33(10):e91. PMID: 15942023
169. S, Faruqi AF, Dean FB, Du Y, Sun Z, Wu X, Du J, Kingsmore SF, Egholm M, Lasken RS. Unbiased whole-genome amplification directly from clinical samples. *Genome Res.* 2003 May;13(5):954-64. Epub 2003 Apr 14. PMID: 12695328
170. Paez JG, Lin M, Beroukhim R, Lee JC, Zhao X, Richter DJ, Gabriel S, Herman P, Sasaki H, Altshuler D, Li C, Meyerson M, Sellers WR. Genome coverage and sequence fidelity of phi29 polymerase-based multiple strand displacement whole genome amplification. *Nucleic Acids Res.* 2004 May 18;32(9):e71. PMID: 15150323
171. Lai KK, Cook L, Krantz EM, Corey L, Jerome KR. Calibration curves for real-time PCR. *Clin Chem.* 2005 Jul;51(7):1132-6. PMID: 15976098
172. Burns MJ, Nixon GJ, Foy CA, Harris N. Standardisation of data from real-time quantitative PCR methods - evaluation of outliers and comparison of calibration curves. *BMC Biotechnol.* 2005 Dec 7;5:31. PMID: 16336641
173. Michael Walter Pfaffl, Real-time RT-PCR: Neue Ansätze zur exakten mRNA Quantifizierung *BIOspektrum Spezial*, 1/04 - 10. Jahrgang
174. Bhargava R, Gerald WL, Li AR, Pan Q, Lal P, Ladanyi M, Chen B. EGFR gene amplification in breast cancer: correlation with epidermal growth factor receptor mRNA and protein expression and HER-2 status and absence of EGFR-activating mutations. *Mod Pathol.* 2005 Aug;18(8):1027-33. PMID: 15920544
175. Fuhrmann C, Schmidt-Kittler O, Stoecklein NH, Petat-Dutter K, Vay C, Bockler K, Reinhardt R, Ragg T, Klein CA. High-resolution array comparative genomic hybridization of single micrometastatic tumor cells. *Nucleic Acids Res.* 2008 Apr;36(7):e39. Epub 2008 Mar 15. PMID: 18344524
176. Gisselsson D, Pettersson L, Höglund M, Heidenblad M, Gorunova L, Wiegant J, Mertens F, Dal Cin P, Mitelman F, Mandahl N. Chromosomal breakage-fusion-bridge events cause genetic intratumor heterogeneity. *Proc Natl Acad Sci U S A.* 2000 May 9;97(10):5357-62. PMID: 10805796
177. Ruiz-Herrera A, Castresana J, Robinson TJ. Is mammalian chromosomal evolution driven by regions of genome fragility? *Genome Biol.* 2006;7(12):R115. PMID: 17156441

178. Frank O, Verbeke C, Schwarz N, Mayer J, Fabarius A, Hehlmann R, Leib-Mösch C, Seifarth W. Variable transcriptional activity of endogenous retroviruses in human breast cancer. *J Virol*. 2008 Feb;82(4):1808-18. Epub 2007 Dec 12. PMID: 18077721
179. Mayer J, Meese EU. The human endogenous retrovirus family HERV-K(HML-3). *Genomics*. 2002 Sep;80(3):331-43. PMID: 12213204
180. Sauter G, Maeda T, Waldman FM, Davis RL, Feuerstein BG. Patterns of epidermal growth factor receptor amplification in malignant gliomas. *Am J Pathol*. 1996 Apr;148(4):1047-53. PMID: 8644846
181. Ishitoya J, Toriyama M, Oguchi N, Kitamura K, Ohshima M, Asano K, Yamamoto T. Gene amplification and overexpression of EGF receptor in squamous cell carcinomas of the head and neck. *Br J Cancer*. 1989 Apr;59(4):559-62. PMID: 2713242
182. Kawamoto T, Mendelsohn J, Le A, Sato GH, Lazar CS, Gill GN. Relation of epidermal growth factor receptor concentration to growth of human epidermoid carcinoma A431 cells. *J Biol Chem*. 1984 Jun 25;259(12):7761-6. PMID: 6330080
183. Filmus J, Pollak MN, Cailleau R, Buick RN. MDA-468, a human breast cancer cell line with a high number of epidermal growth factor (EGF) receptors, has an amplified EGF receptor gene and is growth inhibited by EGF. *Biochem Biophys Res Commun*. 1985 Apr 30;128(2):898-905. PMID: 2986629
184. Filmus J, Trent JM, Pollak MN, Buick RN. Epidermal growth factor receptor gene-amplified MDA-468 breast cancer cell line and its nonamplified variants. *Mol Cell Biol*. 1987 Jan;7(1):251-7. PMID: 3494191
185. Reis-Filho JS, Pinheiro C, Lambros MB, Milanezi F, Carvalho S, Savage K, Simpson PT, Jones C, Swift S, Mackay A, Reis RM, Hornick JL, Pereira EM, Baltazar F, Fletcher CD, Ashworth A, Lakhani SR, Schmitt FC. EGFR amplification and lack of activating mutations in metaplastic breast carcinomas. *J Pathol*. 2006 Aug;209(4):445-53. PMID: 16739104
186. Fouché N, Ozgür S, Roy D, Griffith JD. Replication fork regression in repetitive DNAs. *Nucleic Acids Res*. 2006;34(20):6044-50. Epub 2006 Oct 28. PMID: 17071963
187. Arnaudeau C, Lundin C, Helleday T. DNA double-strand breaks associated with replication forks are predominantly repaired by homologous recombination involving an exchange mechanism in mammalian cells. *J Mol Biol*. 2001 Apr 13;307(5):1235-45. PMID: 11292338

Sample Submission Instructions for

## Comparative Genomic Hybridization (CGH) Services

To facilitate order processing, review these instructions and prepare your samples accordingly. NimbleGen will label and hybridize your samples. If you have questions, please contact customer service at [customer-services@nimblegen.com](mailto:customer-services@nimblegen.com) or (608)218-7600ext.3.

### Sample Requirements

- Purified, unamplified genomic DNA is required to obtain optimal labeling of the sample for the assay hybridization.
  - If your sample is derived from a class I organism, submit 2.5µg of genomic DNA for each sample. (If requesting hybridization of technical replicates, submit an additional 1mg for each replicate hybe.)
  - If your sample is derived from a class II or III pathogen, submit 5mg of sample. (If requesting hybridization of technical replicates, submit an additional 1µg for each replicate hybe.)

The requested amount of sample enables performance of the incoming sample QC ( $A_{260}$  quantification and an agarose gel analysis) and of 2 labeling reactions (a repeat experiment is performed if the first experiment fails NimbleGen's hybridization QC specifications).

Note: Heating sample at 45°C for 10 minutes is recommended prior to quantification.

- OPTIONAL: NimbleGen can provide human male and/or female reference sample. Your preorder form must indicate this intention prior to your receipt of barcodes. NimbleGen's stock reference samples Human Male Genomic DNA, Promega catalog #G1471, or Human Female Genomic DNA, Promega catalog #G1521, are a pool of 6 individuals that can be used with human test samples. Note: The Promega DNA pool of 6 individuals changes with new lot numbers. If you anticipate using NimbleGen Male or NimbleGen Female for a large study, we recommend directly obtaining the DNA from Promega to ensure sufficient DNA for completion of the study.
- Samples should be analyzed for quality and purity by gel electrophoresis using a 1% agarose gel. They must be free of RNA contamination and show no signs of degradation. Submit proof of quality with sample submission.
- Samples must have a concentration of ~250ng/ml in 1X TE buffer or water.
- Samples have an  $A_{260}/A_{280} \geq 1.7$  and  $A_{260}/A_{230} \geq 1.5$ . Have these ratios ready when filling out the submission form.
- Samples must be submitted in sterile, RNase- and DNase-free, 1.5ml, micro-centrifuge tubes.
  - If you are requesting hybridization technical replicates, submit the sample in one tube.
  - If you are requesting dye-swap experiment (the same sample labeled with Cy3 and Cy5 dyes), submit the sample in 2 separate tubes with the individual required amounts.

## CGH Services

### Labeling Your Samples

1. NimbleGen mails barcodes to you after we receive your purchase order. After preparing samples, attach a barcode onto each sample tube.  
Note: If you have multiple projects and/or multiple sample submissions, each project and/or submission receives its own barcodes.
2. You will receive an email with access to an online form that requests sample information. You need to complete this form before sending your samples to NimbleGen. Follow the onscreen instructions to complete the form, and print a copy of your sample information for your records.

### Shipping Your Samples

The online form you use to identify your samples (see previous section) includes shipping instructions.

### After You Submit Your Samples

Upon receipt of your samples, NimbleGen does the following:

1. Compares the samples to the information you entered into the online form.
2. Sends you an email notifying you of the receipt and acceptance of your samples and the sample information submitted in the online form.
3. Analyzes your samples for purity, concentration, and quality using either the Agilent Bioanalyzer or gel electrophoresis.
4. Labels your samples with Cy dye for hybridization.
5. Performs the hybridizations.
6. Sends you a data CD or DVD with the results of your experiment.

### If There Are Problems

NimbleGen will contact you and discuss the options available. Additional fees may be applied to your order based on the option you select.

### What Happens to Unused Samples

If you provide more sample than required and want to use it in future experiments, contact your NimbleGen Account Manager. Otherwise, any materials not used in NimbleGen's performance of the services will be returned to you only upon your written request, provided you give such notice within 6 months from completion of the services and pay for the shipment of materials back to you.

---

## Labeling Your Samples

1. NimbleGen mails barcodes to you after we receive your purchase order. After preparing samples, attach a barcode onto each sample tube.  
**Note:** *If you have multiple projects and/or multiple sample submissions, each project and/or submission receives its own barcodes.*
2. You will receive an email with access to an online form that requests sample information. You need to complete this form before sending your samples to NimbleGen. Follow the onscreen instructions to complete the form, and print a copy of your sample information for your records.

---

## Shipping Your Samples

The online form you use to identify your samples (see previous section) includes shipping instructions.

---

## After You Submit Your Samples

Upon receipt of your samples, NimbleGen does the following:

1. Compares the samples to the information you entered into the online form.
2. Sends you an email notifying you of the receipt and acceptance of your samples and the sample information submitted in the online form.
3. Analyzes your samples for purity, concentration, and quality using either the Agilent Bioanalyzer or gel electrophoresis.
4. Labels your samples with Cy dye for hybridization.
5. Performs the hybridizations.
6. Sends you a data CD or DVD with the results of your experiment.

## If There Are Problems

NimbleGen will contact you and discuss the options available. Additional fees may be applied to your order based on the option you select.

## What Happens to Unused Samples

If you provide more sample than required and want to use it in future experiments, contact your NimbleGen Account Manager. Otherwise, any materials not used in NimbleGen's performance of the services will be returned to you only upon your written request, provided you give such notice within 6 months from completion of the services and pay for the shipment of materials back to you.

## Protocols for

# Comparative Genomic Hybridization

## Outline

This protocol describes the process for comparative genomic hybridization (CGH) to NimbleGen arrays. The process requires both test and reference DNA samples. These samples are independently sheared by sonication and labeled with dye-modified random primers. Labeled test and reference DNA are combined and hybridized to a NimbleGen CGH array using the MAUI Hybridization System. The array is stringently washed, dried, and scanned.

## Protocol Information & Safety

- Wear gloves and take precautions to avoid sample contamination.
- Cy dyes are light sensitive. Be sure to minimize light exposure of the dyes during use and store in the dark immediately after use.
- Cy dyes are ozone sensitive. See <http://genomics.princeton.edu/dunham/ozone.html> for information on monitoring and remediation for atmospheric ozone levels exceeding 5ppb.
- We have found that using VWR water and DTT for all post-hybridization washes is useful for obtaining higher signal from Cy dyes.

## Required Apparatus & Labware\*

\* See last page for a reagent supplier list

- Thermocycler
- Microcentrifuge with 12,000 x g capability
- Heat blocks at 37°C, 45°C, and 95°C
- Vortexer
- Spectrophotometer (A NanoDrop Spectrophotometer is recommended).
- Digital Sonifier (Branson Model #450)
- Electrophoresis box and gel documentation system
- Speed-Vac
- MAUI Hybridization System (BioMicro Systems, Model # 02-A002-03)
- MAUI Mixer SL Low Temp. Hybridization Chambers (BioMicro Systems, Catalog # 02-A008-03)
- Method of drying arrays after washing (e.g., custom NimbleGen Array-Go-Round unit, ArrayIt Microarray Highspeed Centrifuge, Cat# MHC110V)
- Dark Desiccator

## Step 1. Sonication and QC

1. Measure the  $A_{260}$  of the test and reference samples and calculate the DNA concentration.
2. Reserve 0.25 $\mu$ g each of the test and reference samples for gel analysis.
3. Dilute 2 $\mu$ g each of test and reference DNA to 80 $\mu$ l with VWR water in a 1.5ml microcentrifuge tube.
4. Clean the sonicator tip with 70% ethanol and wipe dry with a Kim-Wipe.
5. Set the sonicator to the following:
  - Time = 10 sec
  - Amplitude = 10%
  - Pulse On = 0.5 sec
  - Pulse Off = 0.5 sec

### Important!

Wear hearing protection when operating the sonicator.

6. Lower the sonicator tip to the bottom of the tube and push the start button. Make sure the sonicator is at the bottom or near the bottom of the tube or it may splash and your full sample may not get sonicated.
7. To QC the sonicated and unsonicated genomic samples, prepare a 1% agarose gel with ethidium bromide.
8. Load 250ng of sonicated or unsonicated genomic sample with loading buffer into wells. Include a molecular weight marker ladder.
9. Run the gel until the loading dye is halfway down gel.
10. Photograph or digitally document the gel.

### Important!

Unsonicated genomic samples should have one strong band at the top of the gel. If there is more than one band or smearing in the unsonicated genomic sample, it may be degraded or have a contaminant that could affect the labeling procedure.

11. Analyze gel results to verify the sonication step. Sonicated sample should be a smear from ~500–2000bp.

Genomic DNA exhibiting significant degradation is unsuitable for CGH analysis.

## Step 2. Sample Labeling

1. Prepare the following three solutions:

50X dNTP Mix			Aliquot 50X dNTP mix into single-use amounts. Avoid freeze-thaw cycles, which can result in diminished yields.
VWR deionized water	250	μl	
10X TE	50	μl	
100 mM dATP	50	μl	
100 mM dGTP	50	μl	
100 mM dTTP	50	μl	
100 mM dCTP	50	μl	
<b>Total</b>	<b>500</b>	<b>μl</b>	
Random 9mer Buffer			Make fresh each time primers are diluted.
VWR deionized water	8.6	ml	
1M Tris-HCl	1.25	ml	
1M MgCl <sub>2</sub>	125	μl	
β-Mercaptoethanol	17.5	μl	
<b>Total</b>	<b>10</b>	<b>ml</b>	
10X TE			Mix and store at room temperature.
1M Tris-HCl	1.5	ml	
0.5M EDTA	300	μl	
VWR deionized water	13.2	ml	
<b>Total</b>	<b>15</b>	<b>ml</b>	

2. Dilute Cy3 and Cy5 dye-labeled 9mers to 1 O.D./42μl Random 9mer Buffer. Aliquot to 40 μl individual reaction volumes in 0.2 ml thin-walled PCR tubes and store at -20°C.

### Important!

NimbleGen recommends Cy3 for the test sample and Cy5 for the reference sample. These can be reversed for dye swap experiments.

3. Assemble the following components in separate 0.2ml thin-walled PCR tubes:

	1. Test Sample	2. Reference Sample
Fragmented DNA (1μg)	1μg	1μg
Cy3-9mer Primers	40μl	
Cy5-9mer Primers		40μl
VWR Water (to 80 μl)	To volume	To volume
<b>Total</b>	<b>80μl</b>	<b>80μl</b>

### Important!

Snap-chilling step after denaturation is critical for high efficiency labeling.

4. Heat denature samples in thermocycler at 98°C for 10 min. Quick chill in ice water bath. Add the following (for larger quantities of samples, you may use a master mix for this step):

	1. Test Sample	2. Reference Sample
Reaction volume from Step 2.3	80μl	80μl
50X dNTP mix	10μl	10μl
VWR deionized water	8μl	8μl
Klenow (50 U/μl)	2μl	2μl
<b>Cumulative Total</b>	<b>100μl</b>	<b>100μl</b>

**Important!**

Do not vortex after addition of Klenow.

5. Mix well by pipetting 10 times.
6. Spin down at low RPM to force contents to bottom of tube.
7. Incubate at 37°C for 2 hours in a thermocycler protected from light.
8. Stop the reaction by addition of EDTA:

	1. Test Sample	2. Reference Sample
Reaction volume from step 2.4	100µl	100µl
Stop Solution (0.5 M EDTA)	10µl	10µl
<b>Cumulative Total</b>	<b>110µl</b>	<b>110µl</b>

**Important!**

Vortex briefly after addition of 5M NaCl prior to addition of isopropanol.

9. Transfer the reaction to a 1.5 ml tube. Precipitate the labeled samples by adding NaCl and isopropanol to each tube as indicated.

	1. Test Sample	2. Reference Sample
Reaction volume from step 2.7	110µl	110µl
5M NaCl	11.5µl	11.5µl
Isopropanol	110µl	110µl
<b>Cumulative Total</b>	<b>231.5µl</b>	<b>231.5µl</b>

10. Vortex. Incubate 10 minutes at room temperature in the dark.
11. Centrifuge at maximum speed for 10 minutes. Remove supernatant with pipette.
12. Rinse pellet with 500µl 80% ice-cold ethanol. Dislodge pellet from tube wall.
13. Centrifuge at maximum speed for 2 minutes. Remove supernatant with pipette.
14. Speed-Vac on low heat for 5 minutes until dry, protected from light.
15. Store labeled samples at -20° C, protected from light.
16. Rehydrate dried pellets in 25µl VWR deionized water.
17. Vortex for 30 seconds and quick spin to collect contents at bottom of tube. Continue to vortex or let sit in the dark at room temp until the pellet is completely rehydrated, then vortex again and quickly spin.
18. Measure the  $A_{260}$  in each sample.

**Important!**

Spin tubes down briefly prior to opening.

It is important to consume as little sample as possible when performing this measurement. NimbleGen recommends the use of a NanoDrop spectrophotometer ([www.nanodrop.com](http://www.nanodrop.com)) that allows the measurement of a 1µl sample without dilution. Typical yields range from 25µg to 50µg per reaction.

### Step 3. Hybridization of Cy-labeled DNA

1. Set MAUI Hybridization System to 42°C and allow time for the temperature to stabilize.
2. Based on the  $A_{260}$  measurement performed in step 2.17, combine 13µg each of the Test and Reference samples into a single 1.5ml microcentrifuge tube.
3. Dry the combined contents in a Speed-Vac on low heat.
4. Resuspend the sample in 10.9µl VWR water and vortex to completely dissolve the sample. Spin the tube down briefly to collect the contents in the bottom.

#### Important!

Protect tube from light during handling to prevent photobleaching of the light-sensitive Cy dyes.

#### Important!

CPK6 48mer Oligos are included in the hybridization as controls that hybridize to alignment features on NimbleGen arrays. They are required for proper extraction of array data from the scanned image.

Note: All volumes in this step have changed due to a change in the volume of the MAUI Hybridization Chamber.

5. Using the NimbleGen Array Reuse Kit, add the following to the resuspended sample:

Hybridization Solution	
Cy labeled Test & Reference samples	10.9µl
2X Hybridization Buffer	19.5 µl
Hybridization Component A	7.8µl
Cy3 CPK6 50mer Oligo, 50nM	0.4µl
Cy5 CPK6 50mer Oligo, 100nM	0.4µl
<b>Total</b>	<b>39µl</b>

6. Mix the tube briefly then spin down to collect the contents in the bottom and place at 95°C for 5 minutes.
7. Immediately transfer the tube to the MAUI 42°C sample block and hold at this temperature until you are ready for sample loading.
8. Place the MAUI Mixer SL Hybridization Chamber on the array using the provided assembly/disassembly jig and carefully follow MAUI setup instructions. Use the braying tool to remove all air bubbles from the adhesive gasket around the outside of the hybridization chamber.
9. Put the array and Hybridization Chamber on the MAUI and allow 30 seconds for the chips to come up to temperature.
10. Load the sample using the pipet supplied with the MAUI Station and following manufacturer instructions. During loading, a small amount (3-7µl) of the sample may flow out of the outlet port.
11. Confirm that there are no bubbles in the chamber.  
  
If there are, very gently massage any bubbles to either of the ends, away from the center of the array. Avoid applying too much pressure since this will force liquid out of the ports.
12. Place the loaded array into one of the four MAUI bays and let equilibrate for 30 seconds. Wipe off any sample leakage at the ports with a Kim-Wipe, and adhere MAUI stickers to both ports.
13. Close the bay clamp and select mix mode B. Hold down the mix button to start mixing. Confirm that the mixing is in progress before closing the cover.
14. Hybridize the sample overnight (16-20 hours).

#### Important!

Hybridization chambers have a finite shelf life. Check the expiration date. Use the chambers within 30 minutes of opening the package.

## Step 4. Washing of Arrays

1. Prior to removing the array from the MAUI Hybridization Station, prepare the following solutions. You will need **two** 250 ml dishes of Wash I, and one each for Wash II and Wash III. One dish for Wash I should be shallow and be wide enough to accommodate the array and mixer loaded in the MAUI assembly/disassembly jig. The lid from a 1000  $\mu$ l pipet tip box works well. Also, the buffer in the first Wash I dish should be heated to 42°C to help soften the adhesive on the hybridization chamber; this will help prevent braking the array. Place the remaining three wash solutions in 300ml Tissue-Tek slide staining dishes.

	Wash I x 2	Wash II	Wash III
VWR Water	225ml	225ml	225ml
10X Wash Buffer I, II, or III	25ml	25ml	25ml
1M DTT	25 $\mu$ l	25 $\mu$ l	25 $\mu$ l
<b>Total</b>	<b>250ml</b>	<b>250ml</b>	<b>250ml</b>

### Important!

Peel the hybridization chamber off very slowly to prevent the slide from cracking. Do not let the surface of the slide dry out at any point during washing.

The first volume of Wash I should be heated to 42°C to enable the removal of the mixer.

2. Remove chip from MAUI Hybridization Station, load it back into the MAUI assembly/disassembly jig, and immerse in the shallow 250ml Wash I at 42°C. While the chip is submerged, carefully peel off the mixer. Gently agitate the chip in Wash I for 10-15sec.
3. Transfer the slide into a slide rack in the second dish of Wash I and incubate 2 min with agitation.
4. Transfer to Wash II and incubate 1 min with agitation. Rock the dish to move the wash over the tops of the arrays.
5. Transfer to Wash III and incubate for 15 sec with agitation.
6. Remove array and spin dry in an array-drying unit (e.g. the NimbleGen Array-Go-Round) for 1 min.
7. Store the dried array in a dark desiccator and proceed immediately to the protocol for Two-Color Scanning of NimbleGen Arrays.

## Component and Supplier List

Component	Supplier	Package Size	Item Number
0.5M EDTA	Sigma Aldrich	100ml	E-7889
Absolute Ethanol	Sigma Aldrich	500ml	E702-3
Water, Reagent Grade, ACS, Nonsterile, Type 1	VWR	2.5 gallon	RC915025
100mM dNTP's	Invitrogen	4 X 25umole	10297-018
1M Tris HCL pH 7.4	Sigma Aldrich	1L	T-2863
1M MgCl <sub>2</sub>	Sigma Aldrich	100ml	M-1028
β-Mercaptoethanol	Sigma Aldrich	25ml	M3148
Klenow Fragment 3' ->5' exo-	NEB	50,000u/ml	M0212M
5M NaCl	Sigma Aldrich	250ml	71386
Isopropanol	Sigma Aldrich	500ml	I-9516
Cy3 9mer Wobble	TriLink Biotechnologies	50 O.D. 200 O.D.	N46-0001-50 N46-0001-200
Cy5 9mer Wobble	TriLink Biotechnologies	50 O.D. 200 O.D.	N46-0002-50 N46-0002-200
Cy3 CPK6 50mer (see below)	IDT	250nmole	Custom oligo synthesis
Cy5 CPK6 50mer (see below)	IDT	250nmole	Custom oligo synthesis
MAUI Mixer SL	BioMicro	12 Lids	02-A008-03
NimbleGen Array Reuse Kit 40	NimbleGen	40 Rxn	KIT001-2
NimbleGen Hybridization Kit 40 – CGH & ChIP	NimbleGen	40 Rxn	KIT002-2
NimbleGen Hybridization Kit 40 Refill – CGH & ChIP	NimbleGen	40 Rxn	KIT005-2

## Oligo Ordering Information

Source: Integrated DNA Technologies (800-328-2661)

Scale: 250 nmole, HPLC purified

3' Cy3 CPK6-48mer:

5' TTC CTC TCG CTG TAA TGA CCT CTA TGA ATA ATC CTA TCA AAC AAC  
TCA 3'

3' Cy5 CPK6-48mer:

5' TTC CTC TCG CTG TAA TGA CCT CTA TGA ATA ATC CTA TCA AAC AAC  
TCA 3'

## Danksagung

Abschließend möchte ich mich bei all den Menschen bedanken, die mich bei der Erstellung meiner Dissertation nach Kräften unterstützt haben. Es muss hierbei deutlich gesagt werden, dass ich mich immer auf ein starkes, einzigartiges und motivierendes Team und Umfeld verlassen konnte. Ihnen und Euch allen möchte ich schon vorweg ein herzliches Dankeschön aussprechen! Dies richtet sich an alle gleichermaßen, diejenigen die schon lange dabei, neu hinzugekommen, oder leider weitergezogen sind.

Mein ganz besonderer Dank gilt Herrn Prof. Dr. Burkhard Brandt für das interessante und aktuelle Thema sowie die Unterstützung und die netten und hilfreichen Diskussionsrunden. Darüber hinaus bedanke ich mich bei ihm für das Vertrauen, dass er mir schon vor Antritt meiner Arbeit entgegengebracht hat und die umfangreiche Erfahrungen, die ich im Rahmen meiner Arbeit sammeln durfte.

Ich möchte mich bei Herrn Dirk Kemming für die intensive, gute und kompetente Hilfe bei meiner Arbeit bedanken. Sein Wissen und die Bereitschaft dieses zu teilen hat mich auf sehr vielen Ebenen vorangebracht.

Herr Prof. Dr. Udo Wienand danke ich für seine Bereitschaft diese Dissertation als Vertreter des Fachbereiches Biologie zu begutachten.

Frau Dr. Heike Pospisil und Herrn Simon Joosse möchte ich für die Unterstützung bei der Aufbereitung und statistischen Auswertung der Array-CGH Daten danken.

Mein größter Dank gilt dem gesamten ITB-Team für die gute Zusammenarbeit und das tolle Arbeitsklima – ganz besonders Antje Andreas, Iris Alpers und Andreas Bauche. Dank geht an Roswitha und Michaela, die meine Arbeit Korrektur gelesen haben. Ich bedanke mich auch im Besonderen bei Frau Dr. Juliane Hannemann, die meine Arbeit Korrektur gelesen hat, mir viele Anregungen gab und sich sehr für mich eingesetzt hat. Danke.

Ein ganz besonderer Dank gilt meiner Frau Verena: ohne Deine Zuneigung, Geduld – vor allem während der Schreib- und Lernphase – und Unterstützung hätte ich das alles nicht geschafft! Ich danke meiner Familie und meinen Freunden für die liebevolle Begleitung während all dieser Jahre.