

**DNA-Microarrays  
zur therapiebegleitenden Prognose  
bei Brustkrebs**

Von der Fakultät für Geo- und Biowissenschaften  
der Universität Stuttgart  
zur Erlangung der Würde eines  
Doktors der Naturwissenschaften (Dr. rer. nat.)  
genehmigte Abhandlung

vorgelegt von  
**Andreas Horschinek**  
aus Leonberg

Hauptberichter: Prof. Dr. techn. Herwig Brunner  
Mitberichter: Prof. Dr. rer. nat. Peter Scheurich

Tag der mündlichen Prüfung: 2. August 2006

Hiermit versichere ich, daß ich die vorliegende Arbeit selbst und nur unter Verwendung der angegebenen Quellen und Hilfsmittel verfaßt habe.

Stuttgart, April 2006

Andreas Horschinek

---

<b>Inhaltsverzeichnis .....</b>	<b>I</b>
<b>Abkürzungen.....</b>	<b>IV</b>
<b>Zusammenfassung.....</b>	<b>V</b>
<b><i>Abstract</i>.....</b>	<b>VI</b>
<b>1. Einleitung</b>	
<b>1.1 Brustkrebsentstehung.....</b>	<b>1</b>
<b>1.2 Brustkrebsformen und Statistik.....</b>	<b>1</b>
<b>1.3 Diagnose.....</b>	<b>3</b>
<b>1.4 Therapieformen und Behandlungsmöglichkeiten.....</b>	<b>3</b>
<b>1.5 Prognose.....</b>	<b>4</b>
<b>1.5.1 Gegenwärtige Marker.....</b>	<b>4</b>
<b>1.5.2 Weitere prognostische Marker in klinischer Praxis.....</b>	<b>6</b>
<b>1.6 Signalwege.....</b>	<b>7</b>
<b>1.7 DNA-Microarrays in der Tumordiagnostik.....</b>	<b>8</b>
<b>1.8 Ziel der Arbeit und Projektbeschreibung.....</b>	<b>10</b>
<b>2. Material und Methoden.....</b>	<b>11</b>
<b>2.1 Material.....</b>	<b>11</b>
<b>2.1.1 Chemikalien.....</b>	<b>11</b>
<b>2.1.2 Enzyme.....</b>	<b>12</b>
<b>2.1.3 Antikörper.....</b>	<b>12</b>
<b>2.1.4 Gebrauchsfertige Reagenziensätze (Kits).....</b>	<b>12</b>
<b>2.1.5 Geräte.....</b>	<b>12</b>
<b>2.1.6 Verbrauchsmaterialien.....</b>	<b>13</b>
<b>2.1.7 Verwendete Zelllinien.....</b>	<b>13</b>
<b>2.1.8 Gesamt-RNA Isolate und Referenzprobe.....</b>	<b>13</b>
<b>2.2 Methoden.....</b>	<b>13</b>
<b>2.2.1 Kultur von <i>E. coli</i>.....</b>	<b>13</b>
2.2.1.1 Kulturmedium.....	13
2.2.1.2 Kultur-/Gefriermedium.....	13

<b>2.2.2</b>	<b>Übernacht-Kultur von <i>E. coli</i>.....</b>	<b>14</b>
2.2.3	Zellkultur.....	14
2.2.4	Herstellung von DNA-Matrizen zur PCR-Amplifikation.....	14
2.2.5	PCR-Amplifikation und Aufreinigung der cDNA-Fragmente.....	15
2.2.6	Isolation von Gesamt RNA aus Zellkultur.....	17
2.2.7	Umschreiben der RNA in cDNA.....	18
2.2.7.1	Direktmarkierung der cDNA aus Gesamt RNA.....	18
2.2.7.2	Direktmarkierung der cDNA aus amplifizierter RNA.....	19
2.2.7.3	Indirektmarkierung der cDNA.....	19
2.2.8	Herstellung fluoreszenz-gekoppelter Antikörper.....	19
2.2.9	Dot Blots und Nachweis der modifizierten cDNAs durch..... Antikörper-gekoppelte Farbstoffreaktion	20
2.2.10	RNA-Amplifikation.....	20
2.2.11	Bearbeitung der Objektträger.....	21
2.2.12	Bedrucken und Nachbehandlung der Objektträger.....	21
2.2.13	Hybridisierung.....	22
2.2.14	Auslesen der DNA-Microarrays und Rohdatenextraktion.....	23
2.2.15	Auswertung der DNA-Microarrays und Datenanalyse.....	24
<b>3.</b>	<b>Ergebnisse.....</b>	<b>29</b>
3.1	Genauswahl anhand von Studien.....	29
3.2	Überprüfung der Reproduzierbarkeit.....	31
3.2.1	Nutzung von Transkriptionsprofilen zur Klassifikation von Zelllinien...33	33
3.3	Lineare RNA-Amplifikation.....	35
3.4	Markierungsstrategien für Microarray Experimente.....	44
3.4.1	Biotin-dUTP Monocolor-Markierung.....	44
3.4.2	Etablierung und Validierung des DNA-Microarrays.....	45
3.4.3	Biotin- /Digoxigenin-dUTP Dualcolor-Markierung.....	51
3.5	Tumorscreening.....	56
3.5.1	Differentiell exprimierte Gene in ER <sup>+</sup> - und ER <sup>-</sup> -Gewebe.....	60
3.5.2	Differentiell exprimierte Gene in Normal- und Tumorgewebe.....	61
3.6	Datenauswertung.....	58
3.6.1	Vergleich der beiden Chip-Systeme.....	62
3.7	Bewertung des Genkollektivs.....	65
3.7.1	Differentielle Gene zwischen erbB2 <sup>+</sup> - und erbB2 <sup>-</sup> -Tumoren.....	65
3.7.2	Differentielle Gene zwischen ER <sup>+</sup> - und ER <sup>-</sup> -Tumoren.....	65
3.7.3	Differentielle Gene zwischen Normal- und Tumorgewebe.....	66

## 4. Diskussion

4.1 Literaturrecherche nach relevanten Markergenen.....	67
4.2 Methodenetablierung.....	68
4.3 Bewertung der differentiellen Gene nach dem <i>Tumorscreen</i> .....	69
4.3.1 Signifikante Gene innerhalb ER <sup>+</sup> - und ER <sup>-</sup> -Gewebe.....	69
4.3.2 Signifikante Gene innerhalb Normal- und Tumorgewebe.....	72
4.4 Ausblick.....	75

5. Literaturverzeichnis.....	77
------------------------------	----

## 6. Anhang

Veröffentlichungen.....	94
Lebenslauf.....	95
Abbildungsverzeichnis.....	96
Tabellenverzeichnis.....	97
Aktuelle Genliste .....	98
Danksagungen	

**Abkürzungen**

Die Abkürzungen der Gennamen sind gesondert im Anhang ab S. 98 aufgeführt.

A	Absorption
Abb	Abbildung
adj. p-Wert (BH)	Adjustierter p-Wert nach Benjamini Hochberg
amp	Ampicillin
ATCC	American type culture collection
BLAST	Basic local alignment search tool
BSA	Bovines Serum Albumin
BCIP	5-bromo,4-chloro,3-indolylphosphat
CCD	Charge-coupled device
CDS	Coding sequence
Cam	Chloramphenicol
DEPC	Diethylpyrocarbonat
DTT	1,4 Dithiothreitol
EB	Elution-Buffer
EDTA	Ethylendiamintetraessigsäure
EtOH	Ethanol
EST	Expressed sequence tag
g	Erdbeschleunigung
FCS	Fötales Kälberserum
H.M.F.M	Hogness modified freezing medium
kb	Kilobase
LB	Luria broth
LS-Puffer	Labelingstop-Puffer
limma	Linear Models for Microarray Data
M	Molar
MCS	Multiple cloning site
mJ	Milli Joule
MOPS	3-(N-morpholino)propansulfonic acid
NBT	Nitroblue tetrazolium
NCBI	National Center of Biotechnology Information
ÖR	Östrogenrezeptor
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PgR	Progesteronrezeptor
P/S	Penicillin/Streptomycin
RPA	RNase protection assay
RT	Raumtemperatur
RZPD	Resourcenzentrum / Primärdatenbank
s	Sekunde
SDS	Sodium dodecylsulfate
SSC	Sodium chloride / sodium citrate
TIFF	Tagged Information File Format
TNM	Tumor Nodus Metastase
Tris	Tris-(hydroxymethyl)Aminomethan
ÜN	Übernacht

## Zusammenfassung

Der Kern der Arbeit befasst sich mit der Entwicklung und Validierung eines DNA-Microarrays, der zur therapiebegleitenden Prognose bei Brustkrebserkrankungen eingesetzt werden soll. Das Fernziel ist es, diesen DNA-Microarray in der klinischen Routine einzusetzen.

Das Vorhaben gründet sich auf der Tatsache, dass ca. 70 - 80 % aller erkannten Mammakarzinome keine Chemotherapie zur Ausheilung benötigen. Derzeit gibt es jedoch keine Methoden diese Brustkrebstypen zu erkennen und entsprechend andere Therapieformen anzuwenden. Trotz therapeutischer Behandlungsmaßnahmen liegt die Überlebensrate bei den restlichen 30 - 20 % der Mammakarzinom-Patientinnen, über einen Zeitraum von 5 Jahren hinaus betrachtet, bei lediglich 10 %.

In diesem Ansatz wird angestrebt neue Marker auf Transkriptionsebene zu finden, um hier eine verbesserte Prognose zu ermöglichen, da es bisher nur 2-3 etablierte Marker (ER-Status, HER2), sowie eine Beurteilung nach dem Lymphknotenstatus gibt. Das hier verwendete Genkollektiv gründet sich auf ausführlicher Literaturrecherche und deckt sowohl therapeutisch bedeutsame Gene, als auch Sequenzen, die Tumortyp, den Östrogenrezeptorstatus und die Signaltransduktionswege repräsentieren, ab.

Ein erstes Ergebnis dieser Arbeit war, dass anhand basaler Transkriptionsmuster zweier humaner Brustkrebszelllinien, ein sog. Set an Klassifikatorgenen bestimmt wurde, mit denen die beiden Zelllinien eindeutig voneinander unterschieden werden konnten.

In ersten Versuchen konnte mit der linearen RNA-Amplifikation gezeigt werden, dass mit dieser Technik lediglich 60 % der ursprünglich detektierbaren Gene erhalten blieb, diese gemeinsamen Gene jedoch das gleiche Expressionsverhalten zeigten.

Die Validierung des DNA-Microarrays an einer mit Antiöstrogen stimulierten humanen Brustkrebszelllinie zeigte im Vergleich zum unstimulierten Zustand 33 differentiell regulierte Gene. Unter diesen Genen befanden sich CDKN1A und TFF1, welche direkt von Östrogen und Antiöstrogen reguliert werden können. Diese Ergebnisse wurden durch unabhängige Methoden überprüft und zeigten sehr ähnliche Expressionsverhältnisse.

Nach Entwicklung eines Zweifarben-Systems zur Detektion von indirekt-markierten Proben erfolgte das Tumorscreening an zunächst 24 Gewebeproben (21 Tumor-, 3 Normalgewebe). Die statistische Auswertung ergab zwei Genkollektive. Das erste Genset zur Unterscheidung von ER<sup>+</sup>- und ER<sup>-</sup>-Gewebe erstreckte sich über 33 Gene. Das zweite Genkollektiv, mit dem Normal- von Tumorgewebe unterschieden werden kann, beinhaltete 23 Gene. Die Überprüfung dieser Ergebnisse erfolgte parallel mit einem kommerziell erhältlichen DNA-Microarray.

Neben einer Übereinstimmung von gemeinsam exprimierten Genen, zeigten die beiden Chipsysteme vor allem aber auch ähnliche Ergebnisse bei nicht detektierbaren Genen, d.h. von solchen Sequenzen deren Expression vom Hintergrundsignal nicht zu unterscheiden war.

Die vorliegende Arbeit ist Teil eines Verbundforschungsprojektes zwischen dem Fraunhofer Institut für Grenzflächen- und Bioverfahrenstechnik (IGB) Stuttgart, dem Robert-Bosch-Krankenhaus (RBK) Stuttgart, dem Institut für Zellbiologie und Immunologie (IZI), dem Mathematischen Institut A (MIA) (beide Universität Stuttgart) sowie dem Interfakultären Institut für Zellbiologie (IZT) der Universität Tübingen und wurde vom Ministerium für Wissenschaft, Forschung und Kunst Baden Württemberg aus Mitteln der Landesstiftung finanziell unterstützt (BW PTJ-BIO/24-720.431-1-1/D31).

**Abstract**

The main part of this work dealt with development and validation of a DNA-microarray, which should be applied to accompany measures in therapy of breast cancer. Long-term objective of this project is using this DNA-microarray in clinical practice.

The intention is based on the fact, that about 70 - 80 % of all diagnosed breast carcinomas need any chemotherapy for curing. Actually there are no methods to indentify these types of breast tumors in order to choose other, maybe better suited therapies. The remaining 30 - 20 % of mamma carcinoma patients will only have an overall survival rate within 5 years by 10 %, despite of a therapeutic treatment.

This approach is aimed to identify new markers on transcriptional level improving prognosis, because there are only 2 to 3 established markers (i.e. status of ER or HER2/neu), like also nodal status. The applied gene set is based on extensive literature research and covers genes significant in therapy, tumor type specific and ER representing sequences, like also many components of signal transduction pathways.

A first result of this work was identifying a gene set, which distinguishes the two human breast cancer cellines, MCF 7 and MDA-MB 231, based on their expression profiles. Initial experiments of linear RNA-amplification have shown, that only 60 % of primarily detectable genes remained. In comparison to experiments without RNA-amplification these set of genes indicated similar expression tendencies.

Array validation was performed by a model system containing the antiestrogen- and unstimulated human breast cancer celline MCF 7. The experiment provided 33 differential genes ( $p \leq 0.05$ ). This set of genes contains also, CDKN1A and TFF1, both regulated by estrogen and antiestrogen directly. This results were checked and confirmed by independent methods.

After developing a two color detection system for indirect-labeled probes, a initial tumor screen was performed with 24 specimens (21 tumors and 3 normal tissues).

Statistical data analysis points out two sets of genes. The first set contains 34 genes which distinguishes ER<sup>+</sup> from ER<sup>-</sup> tumors. The second set including 23 genes could be used to differentiate normal tissues from tumors.

Examination of these results was carried out with a commercial available DNA-microarray. The other system showed an accordance in the case of collectively expressed genes, as well by genes which couldn't be detected by both chip systems, because there are no differences to background.

This work is part of a network research project between Fraunhofer Institut für Grenzflächen- und Bioverfahrenstechnik (IGB) Stuttgart, Robert-Bosch-Krankenhaus (RBK) Stuttgart, Institut für Zellbiologie und Immunologie (IZI), Mathematischen Institut A (MIA) (both University of Stuttgart) and Interfakultären Institut für Zellbiologie (IZT), University of Tübingen and was financially supported of means from Landesstiftung by ministry of science, research and arts Baden Württemberg (BW PTJ-BIO/24-720.431-1-1/D31).



## 1. Einleitung

### 1.1 Brustkrebsentstehung

Brustkrebs ist eine komplexe genetische Erkrankung der weiblichen und auch sehr selten der männlichen Brustdrüse, welche durch Akkumulation mehrerer molekularer Veränderungen charakterisiert ist, aber auch durch verschiedene Umweltfaktoren entstehen kann (Bertucci et al., 2002; Zugmaier et al., 2003).

Die Summe dieser Veränderungen, wie beispielsweise Mutationen, chromosomale Instabilität oder epigenetische Ereignisse (Bsp. DNA-Methylierung) sind allen Neoplasien gemeinsam und beeinträchtigen die Funktion der Zellen hinsichtlich Wachstum und Tod erheblich (Børresen-Dale, 2003). Insbesondere Mutationen beeinflussen die Krebsentstehung, wobei mindestens vier bis fünf Mutationen für eine maligne Transformation notwendig sind. Diese genetischen Veränderungen betreffen hauptsächlich zwei verschiedene Genklassen: Onkogene und Tumorsuppressoren. Dabei werden die Onkogene aktiviert und die Tumorsuppressor-Gene inaktiviert. Üblicherweise aktivieren Onkogene die Zellproliferation, wohingegen die Tumorsuppressor-Gene als Inhibitoren zellulärer Proliferation wirken (Cho and Vogelstein, 1992).

Dabei entstehen Brusttumore über einen Mehrstufen Prozess, welcher allgemein als Transformation normaler Zellen in Krebszellen aufgefasst werden kann. Ein Modellsystem dafür haben Vogelstein und Mitarbeiter anhand des Kolonkarzinoms beschrieben (Vogelstein et al., 1989, Fearon and Vogelstein, 1990). Die maligne Transformation erfolgt über eine Hyperplasie, ein prämaligues Stadium und resultiert meist in einem *in situ* Karzinom (Beckmann et al., 1997). Aus dem *in situ* Karzinom entstehen dann häufig die invasiven und damit auch metastasierenden Formen.

Die Komplexität dieser Erkrankung bindet sowohl unterschiedliche Gewebetypen, als auch spezifische Zellen verschiedener Funktionen mit ein, wodurch viele unterschiedliche Subtypen dieser Tumorart resultieren (Jeffery et al., 2002).

Trotz der Vielzahl genetischer Faktoren entsteht Brustkrebs zum überwiegenden Teil durch Umweltfaktoren, wobei die eigentlichen Entstehungsursachen noch ungeklärt sind. Dazu gehören neben Lebensalter, Zeitpunkt der Menarche und Erstgeburt, auch Herkunftsland, familiäre Disposition, Ernährung (Bsp. erhöhter Alkoholkonsum) sowie Östrogenspiegel (McPherson et al., 2000; Boyd 2002, Zugmeier et al., 2003).

### 1.2 Brustkrebsformen und Statistik

Das Mammakarzinom ist bei Frauen in den westlichen Industrienationen die häufigste Tumorart und die zweithäufigste Todesursache. Brustkrebs hat somit einen Anteil von ca. 20 % an der Gesamtkrebsmortalität<sup>1</sup>. So erkrankt etwa jede siebte Frau während ihres Lebens an Brustkrebs. Für Deutschland ergeben sich daraus jährlich 45000 neue Fälle, wovon ca. 15000 tödlichen Ausgang haben (Zahlen 2003, Quelle Robert Bosch Krankenhaus).

<sup>1</sup> [www.dkfz-heidelberg.de](http://www.dkfz-heidelberg.de)

Der Begriff Mammakarzinom umfasst all diejenigen malignen Tumore, die entweder aus den Epithelien der Milchgänge oder von den lobulären Drüsenendstücken hervorgehen.

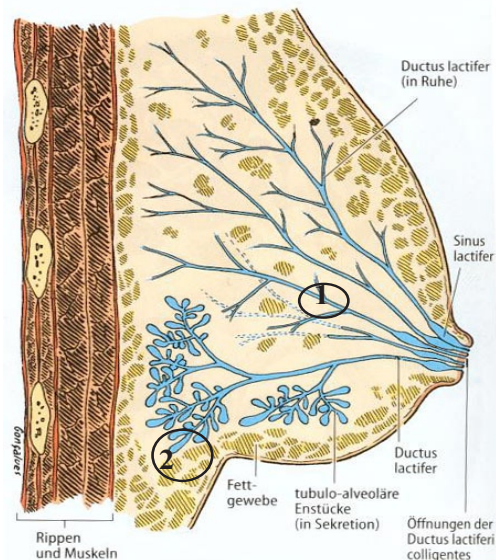


Abb. 1: Drüsen und Gangsystem der weiblichen Brust 1= Ductus lactifer = Milchgang, 2 = Drüsenendstücke (Junqueira et al, 1996)

Die häufigsten Brustkrebsformen sind duktales und lobuläres Ursprungs, wobei das invasiv duktales Karzinom einen Anteil von 80 % besitzt (Riede and Schaefer 1999). Seltener auftretende Mammakarzinome mit geringerem Malignitätsgrad sind das tubuläre Karzinom, das papilläre Karzinom, das muzinöse Karzinom und das medulläre Karzinom.

Histologisch lassen sich die *in situ* Form von der invasiven bzw. infiltrierenden Form gut auseinanderhalten (Abb.2).

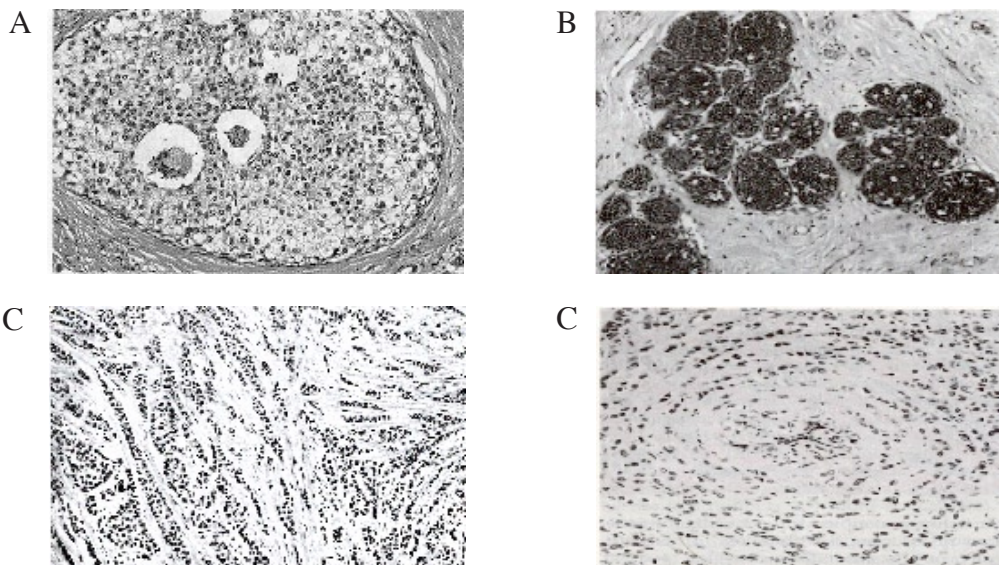


Abb.2: HE-Färbungen verschiedener Mammakarzinomformen. A : Duktales Karzinom *in situ*, klassischer Typ 1:200. B: Lobuläres Karzinom *in situ*, klassischer Typ 1:100; C: Infiltrierendes duktales Karzinom 1:200; D: Infiltrierendes lobuläres Karzinom 1: 200 (Quelle Robert-Bosch Krankenhaus Stuttgart)

### 1.3 Diagnose

Die gegenwärtige Diagnostik beim Mammakarzinom beginnt üblicherweise mit dem Er tasten von Knoten in der Brust. Diese Knoten können durch visualisierte Verfahren, wie beispielsweise die Mammographie, Sonographie oder Kernspintomographie weiter untersucht werden. Erhärtet sich der Tumorverdacht, wird das verdächtige Tumorgewebe entfernt. Vor dem chirurgischen Eingriff kann eine Feinnadelbiopsie erfolgen. In anschließenden pathologischen Untersuchungen wird das Gewebe auf Integrität sowie auf gewebe- und zelltypische Merkmale (s. Tab. 1) untersucht, um weiter mögliche Behandlungsmaßnahmen, ggf. auch bereits vor der Operation, zu treffen.

Aber auch die Früherkennung von Brustkrebs spielt für die Diagnose eine bedeutende Rolle. Obgleich hier die Mammographie als diagnostischer Test die Methode der Wahl ist bleiben ca. 20 % aller Brusttumore unentdeckt (Vlahou et al., 2003). So haben neuere Studien gezeigt, dass sich serologische Biomarker durchaus zur Früherkennung von Brustkrebs eignen (Li et al., 2002, Vlahou et al., 2003, Li et al., 2005).

### 1.4 Therapieformen und Behandlungsmöglichkeiten

Zur Behandlung des Mammakarzinoms gibt es prinzipiell zwei unterschiedliche Therapieansätze. Hier unterscheidet man zwischen invasivem Eingriff, also der chirurgischen Entfernung des Tumors, sowie einer medikamentösen Therapie. Nichtinvasive Therapieformen können beispielsweise Strahlentherapie, Chemotherapie, Hormontherapie sowie adjuvant systemische Therapie sein. Die adjuvant systemische Therapie, findet hauptsächlich bei nodalpositiven Tumoren und der Mehrzahl nodalnegativer Tumore Anwendung (Zugmaier et al. 2003). Hier werden neben dem chirurgischen Eingriff, weitere medikamentöse Maßnahmen durchgeführt, um eventuell vorhandene, jedoch noch nicht diagnostizierbare Mikrometastasen, zu bekämpfen.

Die Hormontherapie wird vorwiegend bei prä- und postmenopausalen Patientinnen mit Östrogenrezeptor-positiven Tumoren (ca. 60-75% aller Brusttumore) eingesetzt, da hier oftmals der klinische Verlauf des Mammakarzinoms vom Östrogenspiegel beeinflusst wird (Knabbe, 1997). Somit ist der Östrogenrezeptorgehalt ein wichtiger prädiktiver Marker, hinsichtlich einer endokrinen Therapie. Der Östrogenrezeptor, wie auch der Progesteronrezeptor gehören zu den hormoninduzierbaren, intrazellulären Transkriptionsfaktoren die u. a. Zellproliferation und Zelldifferenzierung beeinflussen. Ihre Funktion wird durch Steroide reguliert. Bei hohem Östrogenrezeptorgehalt ist die Zelle also entsprechend sensitiv gegenüber Östrogen und besitzt damit eine erhöhte Proliferationsrate. Bei weitergehender Tumorprogression sinkt die Östrogenrezeptoranzahl und der Tumor wächst hormonunabhängig. Mit der Anti-hormontherapie wird mit Östrogenanaloga versucht, den Östrogenrezeptor zu blockieren und damit die verstärkte Zellproliferation einzudämmen.

Derzeit lassen sich Antiöstrogene in zwei Hauptgruppen einteilen. Tamoxifen bzw. seine Metaboliten besitzen sowohl östrogene, als auch antiöstrogene Wirkung. ICI 187.780 zeigt ausschließlich antiöstrogene Wirkung (MacGregor and Jordan, 1998).

Die Validierung des in der Arbeit verwendeten DNA-Microarrays wurde an unbehandelten, wie auch mit ICI 187.780 stimulierten MCF 7 Zellen, aufgrund bereits vorhandener Daten, durchgeführt.

## 1.5 Prognose

Aufgrund der Heterogenität, wie auch dem Auftreten vieler Subtypen des Mammakarzinoms ist es oftmals schwierig eine geeignete Behandlung auszuwählen. Selbst bei histopathologisch identischen Tumoren können diese bei gleicher Behandlung unterschiedliche klinische Verläufe zeigen. Aufgrund dieses Sachverhalts ist hier eine weiter differenzierbare Diagnostik dringend erforderlich (s. Kap 1.8), um bestehende Therapien zu optimieren. Dennoch liefern histopathologische Daten, vor allem das *Grading*, wichtige prognostische Informationen. Einen Ansatz das histologische *Grading* auf molekularer Ebene zu charakterisieren wurde von Sortiriou et al. durchgeführt (Sortiriou et al., 2006). Anhand ihres Genexpressionsprofils konnten Brusttumore mit histologischem *Grading* 1 und 3 deutlich voneinander unterschieden werden. Die Brusttumoren mit histologischem *Grading* 2 zeigten dagegen ein variables Genexpressionsprofil, dessen Werte sich zwischen den beiden anderen Genexpressionsprofilen bewegten. Allerdings konnte innerhalb dieser Gruppe eine Einteilung zwischen Tumoren mit hohem und niederem Rückfallrisiko gemacht werden. Üblicherweise sind *Grading* 2 klassifizierte Tumoren mit mittlerem Rückfallrisiko assoziiert und bieten damit keine verwertbaren Informationen für klinische Entscheidungen. Dies ist insofern von Bedeutung, da zwischen 30 und 60 % der Brusttumoren mit histologischem *Grading* 2 klassifiziert werden. Die Aufteilung der *Grading* 2 klassifizierten Tumoren könnte die Genauigkeit des *Gradings* und damit seinen prognostischen Wert verbessern.

Analog der endokrinen Therapie (s. Kap 1.4) soll dabei auch möglichst eine Vorhersage über das Ansprechen einer adjuvanten systemischen Behandlung gemacht werden (Mitra, 2002). Zur Unterstützung einer Therapie sollen prognostische und prädiktive Faktoren herangezogen werden. Ein prognostischer Faktor gibt hinsichtlich Überlebenszeit ohne Therapie bzw. unter Anwendung einer Standardtherapie Auskunft. Ein prädiktiver Faktor kann dazu genutzt werden, eine Vorhersage über das Ansprechen für eine bestimmte Therapie machen zu können (Zugmaier et al., 2003, Di Leo et al., 2002).

### 1.5.1 Gegenwärtige Marker

Neben Umweltfaktoren werden den histopathologischen Faktoren (Bsp. Östrogenrezeptorstatus) derzeit die größte Bedeutung für eine Prognose zugesprochen. Obwohl es neben den histopathologischen Befunden eine ganze Reihe von weiteren Markern, wie Onko- oder Tumorsuppressorgene sowie Proliferationsindices (Bsp. Ki67) gibt, stützt sich die Prognose gegenwärtig auf nur 2 - 3 etablierte Marker. Dazu gehören der Östrogenrezeptor-, HER2- sowie der Lymphknotenstatus.

Der Östrogenrezeptorstatus (ÖR) begünstigt dabei eine endokrine Therapie. Ist gleichzeitig auch der Progesteronrezeptor (PgR) vorhanden, so lässt sich neben dem Ansprechen auf eine endokrine Therapie auch eine Aussage über eine verlängerte beschwerdefreie Überlebenszeit treffen (el-Ashry and Lippmann, 1994). Auf der anderen Seite zeigt die endokrine Therapie bei einem Drittel der Östrogen- und oder Progesteronrezeptor-positiven Tumoren keine Wirkung (Di Leo et al., 2002). Eine mögliche Ursache dafür kann das Mengenverhältnis der beiden Formen des Progesteronrezeptors (PR-A und PR-B) zum Östrogenrezeptor sein. Während PR-B ein starker Transkriptionsaktivator ist, wirkt PR-A als dominanter Repressor auf PR-B und den Östrogenrezeptor. Ein verändertes Gleichgewicht dieser Steroidrezeptoren kann die Wirkung von Tamoxifen beeinflussen (Hopp et al., 2004).

Aber der Erfolg einer endokrinen Therapie ist nicht alleine von den Steroidrezeptoren abhängig. So konnten Oh et al. zeigen, dass bei Östrogenrezeptor-positiven Brusttumoren vor allem Östrogen-regulierte Gene eine Vorhersage über den klinischen Verlauf einer Tamoxifenbehandlung ermöglichen (Oh et al., 2006). Innerhalb des untersuchten Tumorkollektivs ergab die Studie zwei Gruppen mit unterschiedlicher Vorhersage hinsichtlich dem klinischen Verlauf bei einer Tamoxifenbehandlung. Ferner ist mit diesen Daten auch eine Aussage zur Überlebensrate der Patienten möglich.

Das Proto-Onkogen HER2/neu gehört neben ÖR und PgR zu den am meisten evaluierten Markern. HER2, auch erbB-2 genannt, ist das humane Homolog des in der Ratte charakterisierten Proto-Onkogens c-neu, welches zuerst in Ratten-Neuroblastomen identifiziert wurde (Shih et al., 1981). Das Gen erbB-2 liegt in ca. 25 % aller Brusttumoren in einer erhöhten Kopienzahl vor (Amplifikation) bzw. ist überexprimiert. Bei zusätzlich Lymphknoten-positiven Brusttumoren korreliert dies mit einer schlechten Prognose sowie erhöhter Rückfallrate (Übersicht: Perren 1991, Gullick et al., 1991). Darüberhinaus korreliert eine HER2/neu Überexpression negativ mit ÖR und PgR (Heintz et al., 1990) und mit der Wirksamkeit einer adjuvanten Hormontherapie (Carlomagno et al., 1996).

Bei HER2-positiven Mammakarzinomen wird gegenwärtig eine sog. *First-line*-Chemotherapie in Verbindung mit dem monoklonalen HER2-Rezeptor-Antikörper Trastuzumab oder mit Trastuzumab alleine angewendet (Vogel et al., 2001, Slamon et al., 2001).

Der Lymphknotenstatus ist wohl einer der am stärksten gewichteten Faktoren hinsichtlich eines prognostischen und prädiktiven Markers bei Brustkrebs (Zugmaier et al., 2003, Nemoto et al., 1980, Carter et al., 1989).

Die Methode selbst gilt als universell reproduzierbar und ist quantitativ. Die Überlebenszeit korreliert negativ mit der Anzahl befallener Lymphknoten.

Bei Patientinnen mit Brusttumoren ohne Lymphknotenbefall liegt die 5-Jahres-Überlebenszeit zwischen 70 - 80 %, wohingegen bei Mammakarzinomen mit mehr als 10 Lymphknotenmetastasen die 5-Jahres-Überlebenszeit lediglich bei 20 - 30 % liegt (Mittra, 2002).

Die Einbeziehung der Lymphknoten erfolgte meist im Zusammenhang mit einer Brustamputation und wurde durch großflächige Entfernung des Lymphsystems im axiliären Bereich realisiert. Damit sollte eine Ausbreitung des Tumors verhindert bzw. eventuelle Metastasen beseitigt werden. In jüngerer Zeit kann jedoch bei ca. 2/3 aller Mammakarzinome auf eine Brustamputation verzichtet werden, da mit Hilfe der sog. Sentinel-Node Biopsie lediglich der Wächterlymphknoten (*sentinel node*), also der erste im Abstromgebiet des Tumors liegende Lymphknoten, entfernt wird (McMasters et al., 1998, Giuliano, 2001). Besitzt dieser Lymphknoten keine Metastasen, so wird der Brusttumor als metastasenfrei eingestuft und es erfolgt keine Brustamputation oder Entfernung weiterer Lymphknoten. Neben diesen etablierten Markern wurden eine Vielzahl weiter potentieller Marker untersucht, die in Tabelle 1 (s. Kap 1.5.2) aufgeführt sind.

### 1.5.2 Weitere prognostische Marker in klinischer Praxis

Neben histopathologischen Faktoren wie den Proliferationsindices werden in der nachfolgenden Tabelle zum überwiegenden Teil genetische Marker aufgeführt. Ein bereits häufig beschriebener Faktor ist der Epidermale Wachstumsfaktor Rezeptor 1 (EGFR). EGFR ist ein wichtiger prognostischer Faktor in Brusttumoren mit variabler Expression, der negativ mit der ÖR-Expression korreliert (Fitzpatrick et al., 1984, Sainsbury et al., 1985, Klijn et al., 1992, Beckmann et al., 1996, Bucci et al., 1997, Walker, Dearing, 1999).

Mit TP53 (p53) erweitert sich die Liste von potentiellen Markern. Wildtyp p53 spielt eine zentrale Rolle bei der Zellzyklusarretierung, der DNA-Reparatur sowie bei apoptotischen Vorgängen. Bei Brustkrebs korreliert die Überexpression von p53 mit erhöhtem nukleärem Differenzierungsgrad, verstärkter Proliferation, Aneuploidie, HER2/neu Überexpression sowie negativem Hormonrezeptorstatus (Davidoff et al., 1991). So sprechen einige Hinweise aus retrospektiven Studien für eine günstige Prognose bei Expression von p53 hinsichtlich einer Therapieantwort (Beenken and Bland, 2002).

Unabhängig von ER-Status, der Tumorgröße und Immunhistochemie gilt die Expression von MDM2 als ein negativ prognostischer Marker bei Brustkrebs hinsichtlich der Überlebenszeit (Turbin et al., 2006). MDM2, ein Onkogen, inhibiert u.a. die Funktion von TP53 und damit die apoptotische Kontrolle. Dagegen werden genetische Instabilität und Resistenz bei Chemotherapie toleriert.

Oftmals reicht jedoch ein Marker für eine sichere Prognose nicht aus. Aufgrund dessen wurden in der Vergangenheit unselektive Ansätze, häufig mit erheblichen Nebenwirkungen, angewandt. Deshalb zielen neuere Konzepte auf die Entwicklung molekularer Therapieansätze, wobei die Wirkung neuer Substanzen auf molekulare Mechanismen (Signaltransduktionswege) ausgedehnt wird (Sauer et al., 2002).

Diese subzellulären Signaltransduktionswege umfassen neben Apoptose, Angiogenese, Invasion und Metastasierung vor allem auch die Wachstumsfaktoren sowie deren Rezeptoren.

Eine anderer Ansatz ist die Kombination von mehreren Genen als sog. Multimarker. Dabei könnte beispielsweise die Prognose bei der Behandlung von Brustkrebs verbessert werden. So wurden einer aktuellen Studie TP 53, BCL2, erbB2 und MYC als sog. Multimarker-Modell genutzt, um den klinischen Verlauf bei Tamoxifenbehandlung von Brustkrebspatientinnen vorherzusagen (Linke et al., 2006). Dabei zeigte dieses Multimarker-Modell eine höhere Zuverlässigkeit in der Vorhersage des klinischen Verlaufs bei Tamoxifenbehandlung, als die gegenwärtigen Standardrichtlinien.

<b>Prognostischer Faktor</b>	<b>Positiv</b>	<b>Negativ</b>
Alter	40 - 49	andere Altersgruppen
Axilläre Lymphknoten	kein Befall	Befall vorhanden
Tumorgröße	< 2 cm	> 2 cm
Histologisches Grading	niedrig	hoch
Histologischer Typ	Tubulär, kolloid, medullär oder papillär	nicht weiter spezifiziert
Lymphatisch und Blutgefäß Invasion	nicht vorhanden	vorhanden
Östrogen- und Progesteronrezeptor	vorhanden	nicht vorhanden
Proliferationsindices(Ki67, Mitose-und Markierungsindex)	niedrig	hoch
S-Phase Fraktion	niedrig	hoch
Ploidiegrad	diploid	aneuploid
c-erbB2 Onkogen	exprimiert	nicht exprimiert
bcl-2 Protein	überexprimiert	nicht exprimiert
p53 Tumorsuppressorgen	nicht mutiert	mutiert
Nm-23 Gen	exprimiert	nicht exprimiert
Cathepsin-D	niedrig exprimiert	hoch exprimiert
Urokinase-Plasminogen Aktivator	niedrig exprimiert	hoch exprimiert
Angiogenese (mikrovaskuläre Dichte)	niedrig	hoch
Knochenmark-Mikrometastasen	nicht vorhanden	vorhanden
BRCA1 und BRCA2	*	*
pS2	exprimiert	nicht exprimiert
Hitzeschockproteine	*	*

Tab 1. Zusammenfassung prognostischer Faktoren bei Brusttumoren. Positiv begünstigt die Prognose, negativ wirkt sich nachteilig auf die Prognose aus. Bei den mit \* gekennzeichnet Spalten sind die Untersuchungen noch nicht abgeschlossen (Mittra, 2002)

## 1.6 Signalwege

Einige der Wachstumsfaktorfamilien, deren Signalübertragung durch Rezeptor-Tyrosinkinasen erfolgt, sind stark in die Tumorgenese von Brustkrebs eingebunden. Im einzelnen sind dies das EGF- (*Epidermal growth factor*) und das IGF- (*Insulin-like growth factor*) System (Ausnahme: IGF2R ist ein Mannose-6-Phosphatrezeptor).

Zwei Komponenten des EGF-Systems wurden bereits näher beschrieben, dennoch sind die EGFs selbst mögliche Mitogene und damit sog. „*downstream*“ Effektoren im Signalweg, in den auch c-myc und Cyclin D eingebunden sind.

Die Familie der EGF-Rezeptoren besteht aus vier Mitgliedern (EGFR/HER/erbB1, erbB2/HER2/neu, erbB3 und erbB4).

Nach Bindung von Liganden wie EGF, Amphiregulin, Heregulin oder *Transforming growth factor alpha* kommt es zur Rezeptordimerisierung und anschließender Autophosphorylierung an den Tyrosinresten (Ullrich and Schlessinger, 1990).

Nachfolgend kommt es zur Initiation sog. Signaltransduktionskaskaden, die das äußere Signal ins Zellinnere bis zum Zellkern weiterleiten.

Neben der HER2-Expression wird auch der *Transforming growth factor alpha* in 30 - 50 % aller primären und metastasierenden Brusttumoren festgestellt. Interessanterweise zeigen TGFA exprimierende Mammakarzinome eine Koexpression von EGF-Rezeptoren, was auf einen autokrinen Mechanismus hindeutet<sup>2</sup>.

IGF-I und -II sind pleiotrophe Wachstumsfaktoren die über Komplexbildung mit hochaffinen Bindeproteinen reguliert werden. IGFs haben sowohl mitogene wie auch antiapoptotische Eigenschaften (Zhang and Yee, 2000). Ferner wurden bereits in früheren Untersuchungen bei IGF-I und -II parakrine bzw. autokrine Wirkungen nachgewiesen (Yee et al. 1988, Yee et al. 1989).

Dabei sind Funktionsstörungen von para- und autokrinen Regelkreisen ein wesentlicher Bestandteil der malignen Transformation und Progression (Sporn and Roberts, 1985, Lui and Grandis, 2002).

Das IGF-System zeigt außerdem aktive Interaktionen mit dem ÖR-System. So stimulieren IGF-I und -II die Proliferation ÖR-positiver Brustkrebszellen und umgekehrt induziert Östrogen einige Mitglieder des IGF-Systems (Zhang and Yee 2000). Neben Interaktionen mit dem ÖR-System, dokumentieren einige Studien auch Wechselwirkungen mit dem Integrinrezeptor-System (Vuori and Ruoslahti, 1994, Baron et al. 1998). Vuori und Mitarbeiter konnten dabei erstmals eine Verbindung zwischen dem Insulinrezeptor Signalweg und Zelladhäsionsmolekülen, wie den Integrinen herstellen. Bei Baron et al. wurde gezeigt, dass p125Fak durch Insulin und IGF-I induziert werden kann. Dabei ist p125 *focal adhesion kinase* ein wichtiges Molekül bei Integrin und anderen Wachstumsfaktor-Signalwegen. Auf Chemotaxisbasierend induzieren IGFs Migration in Brustkrebszellen, wobei die IGFs hier möglicherweise direkten Einfluß auf Integrine haben (Doerr and Jones, 1996). Auf der anderen Seite konnten Perks et al zeigen, dass die Aktivierung von Integrinen den mitogenen Effekt von IGF-I in Brustkrebszelllinien inhibiert (Perks et al., 1999).

Anhand der beiden Wachstumsfaktorsysteme soll im Ansatz veranschaulicht werden, welche Möglichkeiten Signaltransduktionswege bei der Suche neuer Marker zur verbesserten Brustkrebsprognose bieten können.

## 1.7 DNA-Microarrays in der Tumordiagnostik

Aufgrund der Heterogenität und Komplexität von Krebs hat sich die DNA-Microarray Technologie in den letzten Jahren zu einem wichtigen Werkzeug der Diagnostik entwickelt. Sie ermöglicht die gleichzeitige und quantitative Analyse von tausenden Genen in einem Experiment.

So haben Studien aus neuerer Zeit gezeigt, daß die Genexpressionsanalyse mit DNA-Microarrays neben einer molekularen Phänotypisierung auch eine Subklassifizierung von Tumoren und Tumorzelllinien ermöglicht (Golub et al., 1999, Alizadeh et al., 2000, Sørli et al., 2001, Perou et al., 2000, Ahr et al., 2001, Khan et al., 2001, Hampton et al., 2003, Perou et al., 1999, Zajchowski et al., 2001), die mit den konventionellen, histopathologischen Methoden, in dieser Weise nicht möglich gewesen wäre (Gruvberger et al., 2002, Bertucci et al., 2003).

<sup>2</sup> [http://emice.nci.nih.gov/emice/mouse\\_models/organ\\_models/mammary\\_models/human\\_breast\\_cancer/growth\\_factors\\_receptors](http://emice.nci.nih.gov/emice/mouse_models/organ_models/mammary_models/human_breast_cancer/growth_factors_receptors)



Zusätzlich konnten anhand tumorspezifischer Genkollektive teilweise auch Aussagen über Krankheitsverlauf oder hinsichtlich der Überlebenszeit gemacht werden (van de Vijver et al., 2002, van't Veer et al., 2002). Mit diesen Studien erhofft man sich außerdem, neue diagnostische, prognostische wie auch prädiktive Biomarker zu finden, um zukünftig verbesserte Therapiekonzepte bzw. Behandlungsmöglichkeiten Patientinnen anbieten zu können. Dabei geht die Entwicklung von Microarray Studien verstärkt in Richtung Prognose hinsichtlich dem Ansprechverhalten verschiedener Therapieansätze. So zeigten jüngere Microarray Studien, dass mit dem Genexpressionsprofil von Brustumoren sowohl eine Aussage über das Ansprechen von prä- als auch postoperativen Behandlungsschemata möglich ist (Nagahata et al., 2004, Folgueira et al., 2005, Gyorffy et al., 2005, Rouzier et al., 2005).

Allerdings basierten all diese Studien häufig auf gesamtgenomischen Transkriptionsanalysen, wo anschließend durch statistisch aufwendige Methoden ein zur Fragestellung relevantes Genset ermittelt wurde.

In dieser Arbeit soll mit der Entwicklung und Validierung eines wissensbasierten DNA-Microarrays eine Weiterentwicklung aus diesen gesamtgenomischen Studien angestrebt werden. Dazu enthält der in dieser Arbeit untersuchte DNA-Microarray nur wenige, wissensbasiert ausgewählte Gene. Die damit identifizierten Markerkombinationen sollen dann in Folgestudien genutzt werden, um Vorhersagen hinsichtlich des Ansprechens auf eine bestimmte Therapie oder dem Überleben zu ermöglichen. Zusätzlich sollen die Reduktion der Genanzahl sowie die Erhöhung der Genreplikate robustere Daten liefern und damit die statistische Datenanalyse vereinfachen. Fernziel ist es, damit ein kostengünstiges Hilfsmittel zur Therapieauswahl im klinischen Alltag zu besitzen.

Ein Beispiel solch eines spezialisierten, wissensbasierten DNA-Microarrays wurde in der Arbeit von Sawiris et al. gezeigt. Dabei konnte sehr präzise eine Unterscheidung zwischen Ovarial- und Kolonkarzinomen (analog dem Mammakarzinom auch als Adenokarzinome bezeichnet), sowie aus Ovarialkarzinomen abgeleiteten Zelllinien gezeigt werden (Sawiris et al., 2002).

Dies ist vor allem insofern von Bedeutung, da Ovarialkarzinome erst sehr häufig in fortgeschrittenem Stadium diagnostiziert werden und aufgrund ihrer unmittelbaren Nachbarschaft zum Kolon nicht immer einwandfrei als Ovarialkarzinom erkennbar sind.

In einer anderen Studie wurde ein brustkrebspezifischer DNA-Microarray (145 Gene) zur molekularen Charakterisierung von Brustkrebszelllinien benutzt (Longueville et al., 2005). Die Vorteile dieses „Low density“ Microarrays liegen hier in der kostengünstigen Anschaffung sowie in der schnellen Durchführbarkeit von Phänotypisierungsstudien bei Populationen von Brustkrebszellen oder Tumorproben.

Die vorliegende Arbeit ist Teil eines Verbundforschungsprojektes zwischen dem Fraunhofer Institut für Grenzflächen- und Bioverfahrenstechnik (IGB) Stuttgart, dem Robert-Bosch-Krankenhaus (RBK) Stuttgart, dem Institut für Zellbiologie und Immunologie (IZI), dem Mathematischen Institut A (MIA) (beide Universität Stuttgart) sowie dem Interfakultären Institut für Zellbiologie (IZT) der Universität Tübingen.

## 1.8 Ziel der Arbeit und Projektbeschreibung

Mit dieser Arbeit wird die Entwicklung und Validierung eines DNA-Microarrays angestrebt, der zur therapiebegleitenden Prognose bei Brustkrebskrankungen eingesetzt werden soll. Das Fernziel des Projektes besteht darin, diesen DNA-Microarray in der klinischen Routine anzuwenden.

Die Arbeit besteht im wesentlichen aus drei Teilen. Im ersten Teil erfolgte die Literaturrecherche nach relevanten Markergenen (s.u.). Der zweite Teil umfaßt die Methodenetablierung sowie den Vergleich zwischen den einzeln angewandten Methoden. Der *Tumorscreen* sowie die zugehörige Bewertung der als differentiell exprimierten Gene wird in Teil 3 dargestellt.

Im Gegensatz zu den bisher durchgeführten Studien (s. S. 8,9), die meist auf gesamtgenomischen Transkriptionsanalysen basieren, soll hier ein wissenschaftlicher Ansatz zur Anwendung kommen. Diese Vorgehensweise soll eine kostengünstige und zielgerichtete Herstellung des DNA-Microarrays ermöglichen, das „Grundrauschen“ irrelevanter Gene reduzieren und die damit verbundene statistische Auswertung erleichtern. Das für diesen DNA-Microarray ausgewählte Genkollektiv gründet sich auf ausführlicher Literaturrecherche und deckt sowohl therapeutisch bedeutsame Gene als auch Sequenzen, die Tumortyp, den Östrogenrezeptorstatus und die Signaltransduktionswege repräsentieren, ab.

Ferner soll in diesem Ansatz eine größere Anzahl von Markern als bei der bisher etablierten Diagnostik (ER-, PgR- und Lymphknotenstatus, HER2/neu) untersucht werden. Dabei wird angestrebt, neue Marker auf Transkriptionsebene zu finden, um hier eine verbesserte Prognose zu ermöglichen. Die Validierung des DNA-Microarrays soll am Modell einer vom Mammakarzinom abgeleiteten Zelllinie erfolgen und durch unabhängige Methoden bestätigt werden.

Im nächsten Schritt sollen Expressionsprofile von zunächst 24 Tumorproben erstellt und anschließend aufgrund gemeinsamer Genexpressionsmuster bzw. gemeinsamen Tumortyp gruppiert werden. Dabei liegt ein Schwerpunkt auf der differentiellen Genexpression, zwischen den Gewebegruppen.

Die hier erzielten Ergebnisse sollen mit einem erweiteren Tumorkollektiv (Erhöhung des Stichprobenumfangs) überprüft werden. Ferner sollen über diese Arbeit hinaus die gewonnenen Daten mit klinischen Parametern wie der Tumorgroße, dem histologischen *Grading* und dem Lymphknotenstatus korreliert werden.

## 2. Material und Methoden

### 2.1 Material

#### 2.1.1 Chemikalien

Soweit nicht anders genannt wurden die verwendeten Chemikalien von den Firmen Roth oder Sigma bezogen.

2log DNA-Ladder, Invitrogen

Ammoniumsulfat

Ampicillin

$\beta$ -Mercaptoethanol

Biotin -11- dUTP, R0081, MBI Fermentas,

BSA

Chloramphenicol

Cy3-, Cy5-dUTP, PA53022, PA55022, Amersham Biosciences

Cy3-, Cy5 Mono-Reactive Dye Pack, PA23001, PA25001, Amersham Biosciences

Diethylpyrocarbonat

Digoxigenin-11-dUTP, Roche

Dinatriumhydrogenphosphat

Dikaliumhydrogenphosphat

EDTA

Ethanol

Ethidiumbromid

Formaldehyd

Glycerin

Hefeextrakt, Difco

Isopropanol

Lithiumchlorid

Kaliumdihydrogenphosphat

Magnesiumsulfat

MOPS

Natriumacetat

Natriumchlorid

Natriumbicarbonat

Natriumdihydrogenphosphat

Natriumhydrogencarbonat

Natriumhydroxid

Nukleotide, MBI Fermentas

polyLysinlösung (0,1%)

Seakem LE Agarose, BMA

Sybr Gold, Molecular Probes

Tris

Trypton Pepton, Difco

### 2.1.2 Enzyme

Taq DNA-Polymerase, ~ 8 U/ $\mu$ l aus *E.coli* pT4 isoliert  
Trypsin , Invitrogen

### 2.1.3 Antikörper

Antinbiotin Antikörper, Maus, monoklonal, Cy3-Konjugat, Sigma  
Antidioxigenin Antikörper, Schaf, polyklonal, Acris Antibodies  
Antibiotin Antikörper, Ziege, polyclonal, Acris Antibodies

### 2.1.4 Gebrauchsfertige Reagenziensätze (Kits)

E.Z.N.A. Total RNA-Kit, TRK PEQLAB  
Nucleo Fast™ 96, Machery Nagel  
Nucleobond AX Machery Nagel  
LabelStar™ Array Kit, Qiagen  
Riboamp™ RNA-Amplification, Arcturus  
Reverse-iT™ 1<sup>st</sup> Strand Synthesis Kit, Abgene

### 2.1.5 Geräte

Heraeus Megafuge 1.0 R, Kühlzentrifuge  
Heraeus Megafuge 1.0  
Heraeus Biofuge pico, Tischzentrifuge  
Heraeus Biofuge fresco, Tisch-Kühlzentrifuge  
Sterilbank Herasafe HSP 12, Heraeus  
Mikroskop CK40, Olympus  
Arrayer: Microgrid II, Biorobotics  
Pins: SMP3, Telechem Array IT  
Scanner: Arrayworx™ Biochip Reader, Applied Precision  
Arrayersoftware TAS Application Suit, Version 2.2.2.5  
Scannersoftware: arrayWoRx 2.0  
Auswertungssoftware AIDA Image Analyzer Version 3.45  
Wasserbad W22, Prüfgeräte-Werk Medingen GmbH  
Waagen Satorius BP 410, Satorius BP1215  
pH-Meter Greisinger Electronics Meß- und Regeltechnik  
PCR-Cycler DNA-Engine, PTC-200 Peltier Thermal Cycler, MJ Research  
PCR-Cycler T3 Thermocycler, Biometra  
Thermoblock Digiblock, Laboratory Devices Inc., USA  
UV-Quelle Stratalinker, Stratagene  
Gelelektrophorese Owl Separation Systems (v. PEQLAB) ,  
Power Supply, Amersham Pharmacia  
Geldokumentation Vilber Lourmat, Fröbel Labortechnik  
Orbitalschüttler Belly Dancer, Stovall Life Sciences Inc.  
Wippschüttler Ltf Labortechnik GmbH, Rocky RT-1S

### 2.1.6 Verbrauchsmaterialien

1,5 ml Zentrifugenröhrchen, Eppendorf  
15 ml und 50 ml PP-Röhrchen, Greiner bio-one  
250 ml Zellkulturflächen, steril, DNase-/RNasefrei, Greiner bio-one  
96 Well Platten, steril, Greiner bio-one  
96 Well PCR-Platten, Steinbrenner Laborsysteme  
Microcon, YM30 Filtereinheiten, Millipore  
Dialyse Membran,  
Nylonmembran Hybond N<sup>+</sup>, Amersham

### 2.1.7 Verwendete Zelllinien

Die in der Arbeit verwendeten Zelllinien MCF-7 und MDA-MB231, (beides humane Brustkrebszelllinien) wurden von ATCC ([www.atcc.org](http://www.atcc.org)) bezogen.

### 2.1.8 Gesamt-RNA Isolate und Referenzprobe

Das hier verwendete Biopsiematerial, bzw. die daraus isolierte Gesamt-RNA wurde im Rahmen des Kooperationsprojektes vom Robert Bosch Krankenhaus (Prof. C. Knabbe, Dr. Fritz) zur Verfügung gestellt. Ferner erfolgte die Bereitstellung einer Referenzprobe, bestehend zu gleichen Anteilen Gesamt-RNA von humanen Brustkrebszelllinien MCF 7, MDA-MB 231 und MDA-MB 435, ebenfalls durch das Robert Bosch Krankenhaus (Dr. Buck).

## 2.2 Methoden

### 2.2.1 Kultur von *E. coli*

#### 2.2.1.1 Kulturmedium

1 x LB: 10 g Trypton Pepton, 5 g Hefeextrakt, 5 g NaCl, ad 1 l H<sub>2</sub>O

#### 2.2.1.2 Kultur-/Gefriermedium

1x LB amp (100µg/ml) in 1x H.M.F.M:

10 x H.M.F.M (*Hogness modified freezing medium*) Gefriermedium:

0,76 g MgSO<sub>4</sub>·7H<sub>2</sub>O, 4,50 g C<sub>6</sub>H<sub>5</sub>Na<sub>3</sub>O<sub>7</sub>·2H<sub>2</sub>O (Natriumcitratdihydrat), 9,00 g (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>,  
440,0 g Glycerin, mit H<sub>2</sub>O auf ein Volumen von 800 ml ergänzen und autoklavieren.  
18 g KH<sub>2</sub>PO<sub>4</sub>, 47 g K<sub>2</sub>HPO<sub>4</sub> mit H<sub>2</sub>O auf ein Volumen von 200 ml ergänzen und autoklavieren.  
Beide Lösungen mischen und bei 4 ° C aufbewahren.

### 2.2.2 Übernacht-Kultur von *E.coli*

Der zur Plasmidisolierung notwendige *E.coli*-Stamm (DH10B), wurde aus der Schrägagarkultur in 1 ml LB + 500 mg/ml Ampicillin überimpft und bei 37° C, 180 rpm ÜN inkubiert. Für die Herstellung von Dauerkulturen wurden 0,5 ml der ÜN-Kultur mit 680 µl 87 %-igem Glycerin gemischt und bei – 80° C gelagert. Weitere ESTs wurden bereits in 96 Well-Platten mit glycerinhaltigem Medium (LB amp, 1x H.M.F.M) angeliefert und durch Überimpfen mit einem 96 Nadel Impfstempel in 1x LB amp (100µg/ml) oder 1 x LB Cam (30µg/ml, nur bei Vektoren pOTB7 und pDNR LIB) mit 1x H.M.F.M., ÜN bei 37° C im Brutschrank repliziert.

### 2.2.3 Zellkultur

Sofern nicht anders angegeben wurden die Zellkulturmedien, sowie die Supplemente von der Firma Invitrogen bezogen. Die verwendeten Brustkrebszelllinien wurden in folgenden Medien bei 37 ° C und 5 % CO<sub>2</sub> kultiviert:

MCF7 in 1640 RPMI mit 10 % FCS und 1 % P/S. MDA-MB 231 in DMEM mit 10 % FCS, 1% P/S, 2 mM L-Glutamax™ und 1 mM Natrium-Pyruvat. Bei konfluenten Zellen erfolgte die Ablösung mit 1 x Trypsin EDTA.

### 2.2.4 Herstellung von DNA-Matrizen zur PCR-Amplifikation

Ein Teil der hier verwendeten ESTs wurde aus einer früheren Lieferung des RZPD, nach Sequenzbestätigung, zur Bestückung des Microarrays ausgewählt.

Zur Generierung dieser PCR-Fragmente wurden entsprechende genspezifische ESTs mit Hilfe des *Unigene-clusters* ausgewählt und vom Ressourcenzentrum des Humangenomprojekts (RZPD, Berlin) bezogen. Die Gensequenzen stammen aus unterschiedlichen Genbibliotheken, die in der Regel in dem Vektor pT7T3 kloniert wurden. Diese Klone wurden als Bakterienkulturen im Schrägagar geliefert.

Zur PCR-Amplifikation wurden mit Vektorprimern aus dem 5'- und 3'- Bereich der *multiple cloning site* gen-spezifische Fragmente amplifiziert. Hierzu wurden Bakterien aus der Schrägagarkultur direkt in das PCR-Reaktionsgemisch gegeben und somit deren Plasmid-DNA durch die Temperaturvariationen während des Denaturierungszyklus freigesetzt und amplifiziert.

Konnte mit dieser Technik keine Amplifikation erzielt werden, erfolgte aus ÜN Bakterienkulturen eine Plasmidisolierung (Puffersystem aus Nucleobond AX). Hierzu wurden 1,5 ml Kulturmedium in Microtiterplatten mit 24 Kavitäten bebrütet.

Zur Zellernte wurde die Microtiterplatte 4 min bei 1550 x g zentrifugiert. Das Pellet wurde in 150 µl S1-Puffer durch vorsichtiges Schwenken der Platte resuspendiert.

Bei der Zellyse wurden 150 µl S2-Puffer zugegeben und durch Schwenken gemischt. Das Gemisch wurde 5 min bei RT inkubiert. Anschließend wurden 150 µl S3-Puffer zugegeben, gemischt und für 15 min bei 0 ° C inkubiert.

Danach wurde die Kulturplatte für 20 min bei 2756 x g (4 ° C) zentrifugiert. Der Überstand wurde durch Invertieren in eine neue Microtiterplatte überführt. Zum Überstand wurden 360 µl Isopropanol pipettiert, gemischt und bei 2359 x g (4 ° C) für 20 min zentrifugiert.

Der Überstand wurde verworfen und das DNA Pellet mit 200 µl 70 %-igem Ethanol gewaschen. Der Platteninhalt wurde erneut gemischt und bei 2359 x g (4 °C) für 5 min zentrifugiert. Der Überstand wurde verworfen und das Pellet bei RT getrocknet (ca. 10 – 15 min).

Das Pellet wurde in 50 µl ddH<sub>2</sub>O und für 5 min bei 37 ° C in Lösung gebracht. Die Effizienz der Plasmid-Isolation wurde mittels Gelelektrophorese, in einem 1 %-igen Agarosegel überprüft. Die Proben wurden bei – 20 ° C gelagert.

Die neuen Gensequenzen liegen in den Vektoren Lafmid BA, p7T3d-Pac, pOTB7 und pCMV-Sport vor.

### 2.2.5 PCR-Amplifikation und Aufreinigung der cDNA-Fragmente

Zur Durchführung einer PCR wurde entweder direkt gereinigte *template*-DNA (s. Kap. 2.2.4) benutzt oder ein Aliquot aus den ÜN-Kulturen mittels des 96 er Impfstempels, in den PCR-Reaktionsansatz überführt.

In der PCR-Reaktion mit Endvolumen von 50 µl wurden folgende Komponenten zugesetzt:

- 5 µl 10 x Reaktionspuffer (100 mM Tris pH 8.8, 600 mM KCl, 15 mM MgCl<sub>2</sub>)
- 1 µl Nukleotid-Mix (je 10 mM)
- 10 µl 5 M Betain
- 0,5 µl 30 mM Cresolrot
- 25,5 bzw 27,5 µl ddH<sub>2</sub>O
- 1 µl Taq Polymerase (8 U/µl)

Separat zu diesem Mastermix wurden in PCR-Reaktionsgefäße je 2 µl *template*-DNA und 5 µl *Primermix* bzw. mit dem 96er Impfblock Bakteriensuspension direkt in den PCR-Mastermix überführt (T7- und T3-Primer, RZPD M13 forward/reverse, M13 (-21) forward/M13 reverse und RZPD sp6/T7 je 10 pmol) vorgelegt.

Für die PCR mit dem Vektor pT7T3 wurde folgendes PCR-Programm gewählt:

1. 95 ° C 3 min 1 Zyklus
2. 94 ° C 1 min
3. 50 ° C 1 min
4. 72 ° C 1 min
5. 72 ° C 3 min 1 Zyklus
6. 4 ° C 15 min

Die Schritte 2. bis 4. wurden mit 30 Zyklen wiederholt.

Für einen Großteil der in den Vektoren Lafmid BA und pT7T3-Pac klonierten Sequenzen konnte o.g. Programm, mit Änderung der *Annealing* Temperatur von 50 auf 61 ° C, übernommen werden.

Bei einer kleinen Fraktion dieser ESTs wurde die *Annealing* Temperatur von 61 auf 58° C abgesenkt.

Durch nochmalige Erniedrigung der Annealing Temperatur, von 58 auf 53° C, gelang die Amplifikation weiterer Sequenzen. Einige Sequenzen konnten jedoch überhaupt nicht amplifiziert werden.

Die in den Vektoren pOTB7 und pDNR LIB klonierten Sequenzen konnten bei einer *Annealing* Temperatur von 48 °C amplifiziert werden.

Jedoch 50 % dieser Sequenzen konnte anfangs, mit den vom RZPD empfohlenen Primern, zunächst überhaupt nicht amplifiziert werden. Nach einer Plasmidisolation (s. Kap. 2.4) gelang die Amplifikation der betroffenen Fragmente.

Für die Amplifikation der im Vektor pCMV-Sport klonierten Sequenzen wurde ebenfalls o.g. PCR-Programm verwendet, wobei die *Annealing* Temperatur auf 42° C festgelegt wurde.

Die PCR-Fragmente wurden im 1%-igen Agarosegel visuell überprüft und verifiziert. Da nicht alle Fragmente gleichwertig amplifizierbar waren, d.h. betr. Produkte zeigten unterschiedlich starke Signale, erfolgte die Konzentrationsbestimmung an ausgewählten Proben photometrisch. Die Messungen ergaben Produktkonzentrationen zwischen 0,1 und 1 µg/µl. Die Fragmentgrößen variierten zwischen 600 und 2000 bp.

Im Folgenden werden die benutzten Primer und ihre Sequenzen aufgeführt.

Vektor	Primername	Primersequenz
pT7T3D-Pac	T7 T3	5'-AAT ACG ACT CAC TAT AGG G-3' 5'-ATT ACC CCT CAC TAA AGG G-3'
Lafmid BA, pT7T3D-PacI	RZPD M13 for RZPD M13 rev	5'-GCT ATT ACG CCA GCT GGG GAA AGG GGG ATG TG-3' 5'-CCC CAG GCT TTA CAC TTT ATG CTT CCG GCT CG-3'
pOTB7, pDNR LIB	M13 for (-21) M13 rev	5'-TGT AAA ACG ACG GCC AGT-3' 5'-CAG GAA ACA GTC ATG AC-3'
pCMV-Sport	RZPD Sp6 for RZPD T7	5'-ATT TAG GTG ACA CTA TAG-3' 5'-TAA TAC GAC TCA CTA TAG GG-3'

Tabelle 2: Primerübersicht zur Amplifikation aus den verschiedenen Vektoren

Die Aufreinigung der PCR-Produkte erfolgte mit Nucleo Fast 96™ Aufreinigungsplatten (Machery Nagel). Die Methode basiert auf Durchflusszentrifugation, wobei Salze, Primer sowie andere Verunreinigungen abgetrennt, die PCR-Produkte jedoch auf der Membran angereichert werden.

Insgesamt 700 Proben wurden aus den 96er PCR-Platten in die 96er Aufreinigungsplatten übertragen.

Anschließend wurde 25 min bei 3750 x g und RT zentrifugiert. Das Waschen der Platten erfolgte mit jeweils 100 µl Reinstwasser pro Kavität, 25 min bei 3750 x g und RT. Die gereinigten PCR-Produkte wurden mit je 40 µl Reinstwasser durch 5 minütige Inkubation auf einem Wippschüttler resuspendiert und mit einer Multikanalpipette in eine 384 Kavitätenplatte transferiert.



Die Proben werden anschließend unter der Sterilbank ÜN eingengt und danach in jeweils 12 µl Druckpuffer (3 x SSC, 1,5 M Betain, 150 mM Na-Phosphat Puffer, pH 8,5) resuspendiert. Die Lagerung der Druckplatten erfolgt bei – 20 ° C. Nach jedem Druckvorgang wurden die Druckplatten unter der Sterilbank erneut eingengt und vor Druckbeginn mit je 12 µl ddH<sub>2</sub>O wieder gelöst. Dies garantiert minimale Probenverluste und konstante Pufferzusammensetzung.

### 2.2.6 Isolation von Gesamt RNA aus Zellkultur

Zur Isolation hochreiner RNA wurden unterschiedliche Methoden angewandt. Es wurden die Reagenziensätze der Fa. Qiagen (RNAeasy™) und PEQLAB (E.Z.N.A. Total RNA Kit™) verwendet. Mit der Methode von PEQLAB wurden die besten Ergebnisse erzielt. Die Vorgehensweise ist im Folgenden beschrieben:

Das Prinzip dieser Methode ist die Kombination der selektiven und reversiblen Bindungseigenschaften der Silikasäulenmembranen, mit den schnell und einfach durchführbaren Zentrifugationssäulenmethoden. So können mit dieser Methode bis zu 100 µg hochreine RNA isoliert werden.

Vor der RNA-Isolation erfolgte die Homogenisierung und Lyse von Zellen bzw. Gewebe unter denaturierenden Bedingungen. Durch ein optimiertes Puffersystem wurden alle vorhandenen RNAsen und sonstigen störenden Enzyme wirkungsvoll inhibiert. Das eingesetzte Puffervolumen richtete sich nach der Zellmenge. 5 x 10<sup>6</sup> Zellen wurden mit 800 µl TRK-Lysepuffer lysiert.

Die Lyse der Zellen wurden durch 20 maliges Hochziehen mit einer 1 ml Spritze und Kanüle (0,8 x 120 mm) unterstützt. Zu dem Lysat wurde eine äquivalente Menge von 70 %-igem Ethanol (mit DEPC-H<sub>2</sub>O) gegeben und vermischt.

Die Säule wurde in das Sammelröhrchen gesteckt, das Lysat auf die Säule gegeben und für 15 s mit 10000 x g zentrifugiert. Der Säulendurchfluss wurde verworfen. Im Anschluss wurde die Säule mit 750 µl Waschpuffer I gewaschen (15 s, 10000 x g). Die Säule wurde in ein neues Sammelröhrchen transferiert und mit 500 µl Waschpuffer II gewaschen (15 s, 10000 x g). Dieser Waschschritt erfolgte zweimal. Die entstandenen Säulendurchflüsse wurden jeweils verworfen. Durch einminütiges Zentrifugieren bei maximaler Geschwindigkeit (16000 x g) erfolgte die vollständige Trocknung der Säule. Die gebundene RNA wurde mit 75 µl DEPC-H<sub>2</sub>O (auf 70 °C vorgewärmt) von der Säule eluiert (1 min, 16000 x g).

Durch diese Technik wurden ca. 65 – 70 µl RNA-Lösung (mit 0,8 – 1,3 µg Gesamt-RNA/µl) isoliert. Die RNA wurde bei – 80 ° C gelagert. Zur Quantifizierung der isolierten RNA wurde die Absorption bei 260 und 280 nm gemessen (OD 260 = 1 entspricht der RNA-Konzentration von 40 mg/ml). Die RNA-Konzentration wurde folgendermaßen berechnet:

$$\text{RNA-Konzentration } (\mu\text{g/ml}) = \text{Absorption}_{260} \times 40 \times \text{Verdünnungsfaktor}$$

Der Quotient A<sub>260</sub> / A<sub>280</sub> zeigte die Reinheit der isolierten RNA an. Diese Werte schwankten zwischen 1,7 und 1,9 (Werte von 1,8 und 2,0 entsprechen einer Reinheit von 90 – 100 %).

In einem Formaldehyd- bzw. RNA-Gel erfolgte die Überprüfung auf mögliche Degradation. Zur Volumenreduktion sowie zur Entfernung von genomischer DNA wurde die RNA mit 4M LiCl auf 2 M eingestellt und ÜN bei – 20 °C eingefroren.

Für die Aufkonzentrierung wurde die RNA 30 Minuten lang bei 16000 x g und 4 ° C zentrifugiert, zweimal mit 70 %-igem Ethanol gewaschen (je 10 min bei 16000 x g, 4° C) und anschließend, nach vollständiger Entfernung des Überstandes für 5 Minuten auf Eis an Luft getrocknet. Anschließend wurde die RNA mit DEPC-H<sub>2</sub>O auf Konzentrationen zwischen 4 und 6 µg/µl eingestellt.

### **2.2.7 Umschreiben der RNA in cDNA**

Für Transkriptionsanalysen der beiden humanen Brustkrebszelllinien MFC 7 und MDA-MB 231 erfolgte die reverse Transkription unter zwei verschiedenen Markierungsmethoden. Ferner wurden auch Transkriptionsprofile mit linear amplifizierter RNA aus o.g. Zelllinien, erstellt. Die Markierung der amplifizierten RNA erfolgte zunächst direkt.

#### **2.2.7.1 Direktmarkierung der cDNA aus Gesamt RNA**

Für die Direktmarkierung mit Cyanine 3- Cyanine-5-dUTP wurden je 50 µg Gesamt RNA eingesetzt. Die Reaktion wurde mit dem Label Star Array Kit<sup>TM</sup> von Qiagen durchgeführt. Das vom Hersteller angegebene Protokoll ist für cDNA-Synthesen ausgehend von 0,2 – 50 µg Gesamt-RNA optimiert und wird wie folgt für Direktmarkierung mit Cyanine 3- oder 5-dUTP beschrieben:

Nach dem Auftauen der Reagenzien (Reaktionskomponenten siehe Herstellerhandbuch) wurde die RNA-Lösung mit RNasefreiem Wasser auf ein Volumen von 18 µl eingestellt und zusätzlich mit 2 µl Denaturierungslösung versetzt. Die Probe wurde kurz gemixt, zentrifugiert und für 5 Minuten bei 65° C inkubiert. Danach wurde die Probe sofort auf Eis gestellt, erneut zentrifugiert und bis zur weiteren Verwendung auf Eis inkubiert. Für jede reverse Transkription wurde separat ein Mastermix folgender Zusammensetzung erstellt:

5 µl 10 x RT-Puffer, 5 µl dNTP-Mix (dATP, dCTP, dGTP, je 5 mM, dTTP 0,8 mM), 1 µl Cyanine 5- bzw. Cyanine 3-dUTP (1 mM), 5 µl Oligo-dT Primer, 0,5 µl RNase Inhibitor, 11 µl RNasefreies Wasser und 2,5 µl Reverse Transkriptase.

Der Mastermix wurde vorsichtig gemischt und zentrifugiert. Anschließend wurde der Mastermix mit der denaturierten RNA komplementiert und die Probe bei 37 °C zwei Stunden inkubiert. Nach zwei Stunden wurde die Reaktion mit je 2 µl Stopsolution terminiert und die Proben mittels Durchflusszentrifugation über Säulen aufgereinigt. Sofern keine weiteren Angaben zur Zentrifugation gemacht sind, wurden alle Zentrifugationsschritte bei 16000 x g, 1 Minute bei RT durchgeführt.

Zunächst wurden die Proben mit je 260 µl PB-Puffer gemischt, auf die Reinigungssäulchen übertragen, zentrifugiert und der ÜS verworfen.

Nach Durchflussentfernung wurden die Säulchen mit 750 µl LS Puffer beladen, erneut zentrifugiert und der ÜS verworfen. Im nächsten Schritt wurden die Säulchen mit 750 µl PE-Puffer beladen, zentrifugiert und der ÜS verworfen. Danach wurden die Säulchen zusätzlich zur vollständigen Trocknung zentrifugiert und in ein neues 1,5 ml Reaktionsgefäß überführt.

Die auf den Säulchen gebundene cDNA wurde mit 10 µl EB-Puffer eluiert. Zur verbesserten Ausbeute und bei cDNA-Synthesen aus mehr als 25 µg Gesamt RNA war eine weitere Elutionsrunde erforderlich.

Die cDNA kann direkt für die Hybridisierung eingesetzt oder bis zur weiteren Verwendung bei – 20 °C gelagert werden.

Zur Hybridisierung erfolgte die Aufnahme in einem Hybridisierungspuffer (s. Kap. 2.2.13).

### 2.2.7.2 Direktmarkierung der cDNA aus amplifizierter RNA

Für die Direktmarkierung mit Cyanine 3- Cyanine-5-dUTP wurden je 1 µg amplifizierter RNA eingesetzt. Die Reaktion wurde mit dem Label Star Array Kit™ von Qiagen durchgeführt. Denaturierung der RNA und die Aufreinigungsschritte blieben unverändert. Die einzige Änderung ergab sich in der Zusammensetzung des Mastermixes hinsichtlich des verwendeten Primers. Da in der RNA-Amplifikation sog. *antisense*-RNA generiert wird und diese nicht über ein polyA-Ende verfügt, wurde hier statt des Oligo-dT Primers ein 9mer Random-Primer verwendet:

5 µl 10 x RT-Puffer, 5 µl dNTP-Mix (dATP, dCTP, dGTP, je 5 mM, dTTP 0,8 mM), 1 µl Cyanine 5- bzw. Cyanine 3-dUTP (1 mM), 5 µl 9mer Random-Primer (100 mM), 0,5 µl RNase Inhibitor, 11 µl RNasefreies Wasser und 2,5 µl Reverse Transkriptase.

### 2.2.7.3 Indirektmarkierung der cDNA

Die indirekte Markierung mit Biotin-11- und Digoxigenin-11-dUTP wurde ebenfalls unter Verwendung des Label Star Array Kit™ von Qiagen, mit 5 µg RNA-Ausgangsmenge, durchgeführt. Der Denaturierungsschritt und das Aufreinigungsprozedere blieben unverändert. Die einzige Änderung ergab sich in der Zusammensetzung des Mastermixes:

5 µl 10 x RT-Puffer, 5 µl dNTP-Mix (dATP, dCTP, dGTP, je 5 mM, dTTP 0,9 mM) 0,5 µl Biotin-11- und Digoxigenin-11-dUTP (je 1mM), 5 µl Oligo-dT Primer, 0,5 µl RNase Inhibitor, 11,5 µl RNasefreies Wasser und 2,5 µl Reverse Transkriptase.

### 2.2.8 Herstellung fluoreszenz-gekoppelter Antikörper

Die Detektion von Expressionsprofilen bei Hybridisierungen mit biotin-/digoxigenin markierter cDNA erfolgte über eine Antikörperfärbung des Arrays. Neben einem kommerziell erhältlichen Cyanin 3 konjugierten Antikörper wurde zur Detektion von digoxigeninmarkierter cDNA ein Cy5 konjugierter Anti-Digoxigenin Antikörper benötigt. Da dieser Antikörper mit Cyanin 5 Kopplung kommerziell so nicht erhältlich war, wurde ein unkonjugierter Antikörper mit Cyanin 5 nach folgender Vorschrift gekoppelt: Der Antikörper wurde mit 1 ml 0,1 mM Natriumcarbonat-Natriumbicarbonat Puffer, pH 9,3, auf eine Konzentration von 1 mg/ml eingestellt. Diese Antikörperlösung wurde komplett in das Röhrchen mit dem Cyanin 5 Farbstoff (*Cy5 Mono-Reaktive Dye Pack*) überführt und gründlich gemixt.

Die Reaktion wurde für 30 min bei Raumtemperatur inkubiert und in Intervallen von 10 Minuten zusätzlich gemixt. Nach der Markierung erfolgte die Trennung von ungebundenem Farbstoff durch Dialysieren in 1 x PBS bei 4 °C für 2 Stunden. Dieser Dialyseschritt wurde insgesamt 3 x wiederholt. Bei der kompetitiven Hybridisierung und anschließender Antikörperfärbung zeigte sich, dass der kommerziell erhältliche Cy3-Antikörper im Gegensatz zum selbst hergestellten Cy5-Antikörper wesentlich schlechtere Signalintensitäten zeigte. Daraufhin wurde ein zweiter Antikörper mit Cyanin 3 gem. obiger Anleitung markiert.

### **2.2.9 Dot Blots und Nachweis der modifizierten cDNAs durch antikörper-gekoppelte Farbstoffreaktion**

Zur Überprüfung der in Kap. 2.2.7.3 synthetisierten cDNAs wurden diese auf eine zuvor in ddH<sub>2</sub>O equilibrierte Nylonmembran aufgebracht und UV-fixiert (0,12 J/cm<sup>2</sup>). Nach der Fixierung wurden die Blots 5 min in Waschpuffer (0,1 M Maleinsäure, 3 M NaCl pH 8,0 anschließend auf 1000 ml auffüllen, 0,3% Tween 20) inkubiert. Im nächsten Schritt wurde die Membran 60 min. im sog. Blockierungspuffer (9,5 ml Waschpuffer 0,5 ml 10 % Blockierungsreagenz, Roche) inkubiert. Anschließend erfolgte eine 30-minütige Inkubation im sog. Konjugat-Puffer (Blockierungspuffer mit a) 1: 15000 Antidigoxigenin AP; b) 1:2000 Antistreptavidin AP). Zur Entfernung ungebundener Antikörper wurde die Membran 4 x 5 min in Waschpuffer gewaschen. Ein weiterer Waschschrift erfolgte in 1 M Tris-Puffer, pH 7,5 für 2 x 5 min. Der Farbnachweis wurde mit einer BCIP/NBT-haltigen Lösung (45 µl BCIP und 35 µl NBT auf 10 ml TBS) durchgeführt und nach der 30-minütigen Inkubation der Blots durch Spülen unter Leitungswasser beendet.

### **2.2.10 RNA-Amplifikation**

In der Praxis steht zur Transkriptionsanalyse von Tumoren/Normalgewebe häufig wenig Ausgangsmaterial zur Verfügung, da der Hauptteil des Tumors zur histochemischen Analyse verwendet wird.

Da jedoch für Microarray Experimente mindestens 10 µg Gesamt-RNA zur Verfügung stehen und diese Experimente mindestens im Duplikat ausgeführt werden sollten, wird dieser Materialmangel durch RNA-Amplifikation kompensiert.

Gegenwärtig stehen zwei RNA Amplifikationsmethoden zur Verfügung, auf der einen Seite die lineare, T7 RNA Polymerase basierte sowie auf der anderen Seite die PCR-basierte RNA-Amplifikation. Des Weiteren stehen für diese Methode sowohl Kits wie auch von Arbeitsgruppen entwickelte bzw. modifizierte Protokolle zur Verfügung. Die Wahl der Methode richtet sich in erster Linie nach der RNA-Ausgangsmenge und dem Budget der Experimentatoren. So werden bei der exponentiellen RNA Amplifikation Mengen von weniger als 100 ng, bei der linearen RNA Amplifikation Mengen über 100 ng eingesetzt. Allerdings können bei der linearen RNA-Amplifikation auch Mengen < 100 ng eingesetzt werden, wenn die Amplifikation wiederholt wird. In der vorliegenden Arbeit wurde die lineare RNA-Amplifikation mit dem RiboAmp™ RNA Amplification Kit von Arcturus durchgeführt. Die Amplifikation beinhaltet fünf Teilschritte (cDNA-Erststrangsynthese, cDNA-Zweitstrangsynthese, cDNA-Aufreinigung, *in vitro* Transkription und RNA-Aufreinigung; Prinzip s. Kap 3.3).

Alle temperaturabhängigen Schritte erfolgen in einem Thermocycler mit beheizbarem Deckel.

Zur cDNA-Erststrang Synthese wurden jeweils 3,5 und 1,75 µg Gesamt RNA aus den Brustkrebszelllinien MDA-MB 231 und 1,0 und 0,5 µg Gesamt RNA aus MCF 7, in jeweils 10 µl RNasefreiem Wasser aufgenommen. Die weiteren Arbeitsschritte erfolgten nach Herstellerangaben.

### 2.2.11 Bearbeitung der Objektträger

Zur Herstellung von DNA-Microarrays wurden Mikroskopieobjektträger aus Glas, im weiteren *slides* genannt, verwendet. Die *slides* wurden in eine mit Waschlösung (2,5 N NaOH mit 60 % Alkoholanteil) gefüllte Glaskammer (350 ml) gestellt, für 2 Stunden auf einem Orbitalschüttler vorsichtig bewegt und damit gereinigt. Es ist absolut notwendig keinen 100 %-igen Alkohol zu verwenden, da dieser einen sehr starken Hintergrund in der Fluoreszenzanalyse erzeugt<sup>3</sup> (Eisen and Brown, 1999). Nach dem zweistündigen Waschschrift wurden die *slides* viermal mit ddH<sub>2</sub>O gewaschen und anschließend für 1 Stunde in einer polyLysinlösung (0,01 % polyLysin in 0,1 x PBS) durch vorsichtiges Bewegen auf einem Orbitalschüttler beschichtet.

Im Anschluß wurden die *slides* viermal mit ddH<sub>2</sub>O gewaschen und getrocknet (5 min, 43 x g, RT).

Die *slides* müssen vor der Weiterbearbeitung mindestens zwei Wochen bei RT lagern (Eisen and Brown, 1999). Diese Zeit ist notwendig, damit sich auf der *slide*-Oberfläche eine ausreichende Hydrophobizität ausbilden kann, die das Druckverhalten der *slides* beeinflusst. Hierfür wurden die *slides* in einer staubdichten Box aufbewahrt und bei RT gelagert.

### 2.2.12 Bedrucken und Nachbehandlung der Objektträger

Das Bedrucken der *slides*, im folgenden *printing* bzw. *spotting* genannt, erfolgte automatisiert mit einem sogenannten *Arrayer*. Der *Arrayer* ist ein computergesteuerter, dreidimensional beweglicher Roboterarm und kann mit seiner Druckvorrichtung (*pins*) aufgrund von Kapillarkräften ein definiertes Volumen an Flüssigkeit aufnehmen (lt. Herstellerangabe 250 nl), um diese wieder in definierter Menge (etwa 0,6 nl) auf der *slide*-Oberfläche abzugeben. Aufgrund des Arrayformats wurden vier Nadeln genutzt.

Die Anordnung der DNA-*dots* (*arraydesign*) wurde in einem vierfachen Feld, je als Replikat angelegt, vorgegeben.

Nach dem *printing* wurden die *slides* im *Arrayer* gelagert (25 °C, Luftfeuchtigkeit 40 %) und anschließend einer Nachbehandlung, im folgenden *postprocessing* genannt, unterzogen. Das *postprocessing* gliederte sich in drei Arbeitsschritte: „Backen“ der *slides*, UV-*cross-linking*, Denaturierung (s. unten). Zusätzlich wurden die *slides* unmittelbar vor Gebrauch prähybridisiert (s. unten).

Die *spots* weisen nach dem Drucken häufig unterschiedliche Größen auf. Durch zwei-stündiges „Backen“ bei 80 °C im Trockenschrank erfolgt eine Trocknung des betainhaltigen Druckpuffers (Zusammensetzung s. oben) und damit eine Homogenisierung der *spot*-Größe.

<sup>3</sup> [http://cmgm.stanford.edu/pbrown/protocols/1\\_slides.html](http://cmgm.stanford.edu/pbrown/protocols/1_slides.html)

Das *crosslinking* erfolgte mit UV-Strahlung im Stratalinker. Hierfür wurden die *slides* mit der bedruckten Seite nach oben mit einer Energiedosis von 250 mJ bestrahlt. Beim *crosslinking* wird die DNA kovalent an die *slide*-Oberfläche gebunden. Anschließend wurde durch Waschen der *slides* in 0,1% SDS und ddH<sub>2</sub>O, für jeweils 15 Sekunden, ungebundene DNA entfernt. Im nächsten Arbeitsschritt erfolgt die Denaturierung der Sonden.

Zur Denaturierung wurden 1,2 – 1,6 l H<sub>2</sub>O zum Kochen gebracht und die *slides* in das heiße, jedoch nicht mehr kochende Wasser für 3 Minuten eingetaucht. Im Anschluß erfolgte die Trocknung durch Zentrifugation der *slides* (5 min, 43 x g). Die *slides* sind nun zum sofortigen Gebrauch fertig bzw. können ca. 2 Monate bei RT lagert werden.

Vor unmittelbarem Gebrauch der *slides* erfolgt eine Prähybridisierung (4 x SSC, 10 % BSA, 0,1 % SDS) für 45 Minuten bei 42 °C. Nach der Prähybridisierung erfolgt ein Waschschritt für 10 s in ddH<sub>2</sub>O. Die Prähybridisierung dient zur Vermeidung von unspezifischer Bindung der markierten cDNA sowie zur Reduktion von störenden Hintergrundsignalen.

Neben selbst hergestellten polyLysin-*slides* wurden für den Hauptteil der Arbeit auch kommerziell erhältliche Aminosilan-*slides* der Fa. PEQLAB verwendet. Das *postprocessing* dieser *slides* kann entweder nach den Herstellerangaben oder nach obigem Protokoll durchgeführt werden. Die Lagerung dieser *slides* erfolgt nach Anbruch der Verpackung sowohl vor, als auch nach dem *postprocessing* unter einer Argonschutzatmosphäre in einem Exikator. Die Verwendungszeit nach dem *postprocessing* beträgt etwa 3 Monate.

### 2.2.13 Hybridisierung

Das Hybridisierungsvolumen bei dem hier verwendeten DNA-Microarray beträgt für eine kompetitive Hybridisierung 15 µl. Die Aufkonzentration der cDNA erfolgt nach Paarung einer Cy5-markierten mit einer Cy3-markierten Probe über Amicon Filtereinheiten der Fa. Millipore (YM 30). Das hier angewandte Prinzip basiert auf Durchflusszentrifugation, wobei die cDNA nicht auf der Membran gebunden, sondern aufgrund der Membranausschlussgröße, angereichert wird. Die Filtereinheit wurde mit der Probe beladen und zweimal für zwei Minuten bei 10000 x g zentrifugiert. Anschließend wurde in 30 s Intervallen bei 16000 x g zentrifugiert bis neben der Probe auch die Membranoberfläche wieder sichtbar wurde.

Zu Elution der cDNA wurde die Filtereinheit invertiert und 1 Minute bei 2000 x g zentrifugiert. Das resultierende Volumen betrug zwischen 5 und 7 µl. Zur Komplementierung der Hybridisierungslösung wurde die cDNA mit 3 µl 20 x SSC, 0,2 µl 1 % SDS vermischt und mit EB-Puffer auf 15 µl ergänzt.

Vor der Hybridisierung erfolgt eine Denaturierung der cDNA für 3 Minuten bei 100 °C, mit anschließender Zentrifugation (20 s bei 16000 x g, RT).

Die Probe bzw. das Proben-Gemisch wurde nun blasenfrei auf den gekennzeichneten Bereich des *slides* pipettiert und mit einem Deckgläschen (20 x 20 mm) belegt. Das *slide* wurde in einer verschließbaren Hybridisierungskammer im Wasserbad bei 65 °C ÜN (Zeitraum 16 – 20 h) inkubiert.

Die mit Biotin-, Digoxigenin-dUTP-markierten Proben wurden nach der Elution mit 9,2  $\mu$ l 20 x SSC, 4,6  $\mu$ l 1 % SDS und 23,2  $\mu$ l ddH<sub>2</sub>O auf ein Endvolumen von 46  $\mu$ l eingestellt. Die Denaturierung der Proben erfolgte analog den direktmarkierten Proben. Auf den gekennzeichneten Bereich des *slides* wurden sog. Lifterslips positioniert und zusammen mit der Hybridisierungskammer auf 65 °C erwärmt. Anschließend wurde die Probe zügig und blasenfrei unter das Lifterslip pipettiert. Die Hybridisierung erfolgte wie oben beschrieben. Vor der Auswertung wurden die *slides* folgendem Waschprozedere unterzogen:

Erster Waschschrift der *slides* erfolgte in 2 x SSC, 0,1 % SDS-haltigem, auf 42 °C vorgewärmtem Waschpuffer bis sich das Deckgläschen gelöst hat. Anschließend wurden die *slides* im Waschpuffer gleicher Zusammensetzung (ebenfalls auf 42 °C vorgewärmt) fünf Minuten lang auf dem Wippschüttler, bei langsamer Bewegung, inkubiert. Im Waschpuffer2 (0,1 x SSC, 0,1 % SDS) wurden die *slides* für 10 Minuten auf dem Wippschüttler inkubiert und anschließend 15 Sekunden lang im Waschpuffer3 (0,1 x SSC) gewaschen. Im nächsten Waschschrift wurden die *slides* für zwei Minuten in 0,1 x SSC auf dem Wippschüttler, inkubiert. Dieser Waschschrift wurde wiederholt und anschließend nochmals in frischem 0,1 x SSC für eine weitere Minute fortgesetzt. Der letzte Waschschrift erfolgte für 15 Sekunden im Waschpuffer4 (0,01 x SSC). Danach wurden die *slides* fünf Minuten bei 43 x g getrocknet.

Die *slides* mit den fluoreszenzmarkierten Proben konnten nun direkt im Microarray Scanner ausgelesen werden.

Bei den *slides* mit biotin-, digoxigenin-markierten Proben erfolgte zuvor die Antikörperfärbung. Dazu wurden *slides* erneut mit einem Lifterslip belegt, ein Cy3 konjugierter Antibiotin-Antikörper 1:100, ein Cy5 konjugierter Antidigoxigenin Antikörper 1:40 (je 10  $\mu$ g/ml) in TBST-Puffer (25mM Tris HCl, pH 7.5, 150 mM NaCl, 0,05 % Tween20) verdünnt, unter das Lifterslip pipettiert und in einer Hybridisierungskammer (s. o.) zwei Stunden bei RT, im Dunkeln inkubiert. Das Waschen der *slides* erfolgte zunächst bis zum Ablösen der Lifterslips in TBST-Puffer. In einem weiteren Waschschrift wurden die *slides* für 30 min in TBST-Puffer, in einem abgedunkelten Gefäß, auf dem Orbitalschüttler bei langsamen Bewegungen inkubiert. Anschließend wurden die Arrays für 15 s in ddH<sub>2</sub>O gewaschen und 5 min bei 43 x g trocken zentrifugiert.

#### 2.2.14 Auslesen der DNA-Microarrays und Rohdaten Extraktion

Vor der Signalanalyse wurden die *slides* beidseitig mit N<sub>2</sub> abgeblasen und die unbedruckte Seite zusätzlich mit 100 % EtOH gereinigt. Die Fluoreszenzanalyse erfolgte mit einem computergesteuerten Laserscanner (ArrayWorx, Applied Precision). Das optische System des Scanners besteht aus einer CCD-Kamera und einer Weißlichtquelle. Von dieser Weißlichtquelle aus kann über einen Anregungsfilter nur eine bestimmte Bandbreite des Lichtes passieren, welches weiter über eine Fiberoptik auf die zu untersuchende Probe trifft. Die Emissionsfilter und weitere Optik stellen sicher, dass nur vom Fluorophor emittierte Wellenlängen wieder die CCD Kamera erreichen. In der CCD Kamera wird das von den Emissionsfiltern gesammelte Licht digitalisiert. Anschließend generiert die Scanner Software die zur Fluoreszenzanalyse notwendigen Expressionsbilder (TIFF-files). Die *slides* wurden mit unterschiedlichen Belichtungszeiten gescannt.

Beim Scannen mit unterschiedlichen Belichtungszeiten soll mit der Erweiterung des dynamischen Messbereichs die Sensitivität der Microarrays verbessert werden.

Mit den beiden Belichtungszeiten sollen zum einen diejenigen Gene mit starker Expression erfasst werden, deren Signalintensitäten bereits an der oberen Grenze des linearen Messbereichs liegen. Mit einer zweiten, meist verlängerten Belichtungszeit werden hauptsächlich die schwach exprimierten Gene detektiert. Grundsätzlich gilt dabei, daß die Anzahl der gemessenen Sonden, wie auch die Größe des Fluoreszenzsignals steigt, je länger die Belichtungszeit ist. Allerdings ist es für eine korrekte Auswertung notwendig, daß die Signalintensitäten im linearen Meßbereich liegen. Beim hier verwendeten Scanner umfasst der Bereich, durch die 16bit-Skalierung, 1- 65535 Graustufen ( $2^{16}$ ).

Wird die Belichtungszeit zu lang gewählt, fallen einige Signale in die Sättigung und liegen damit außerhalb des linearen Bereichs, so dass der Scanner die Signale nicht mehr unterscheiden kann.

Die Rohdatenerstellung erfolgte der Auswertungssoftware AIDA Image Analyser™ Vers.3.45 (Raytest).

### 2.2.15 Auswertung der DNA-Microarrays und Datenanalyse

Ausgehend von den Rohdaten wurde die weitere Auswertung vom Mathematischen Institut A der Universität Stuttgart, mit den Software Applikationen R<sup>4</sup> und dem *Bioconductor-Package* limma<sup>5</sup> durchgeführt.

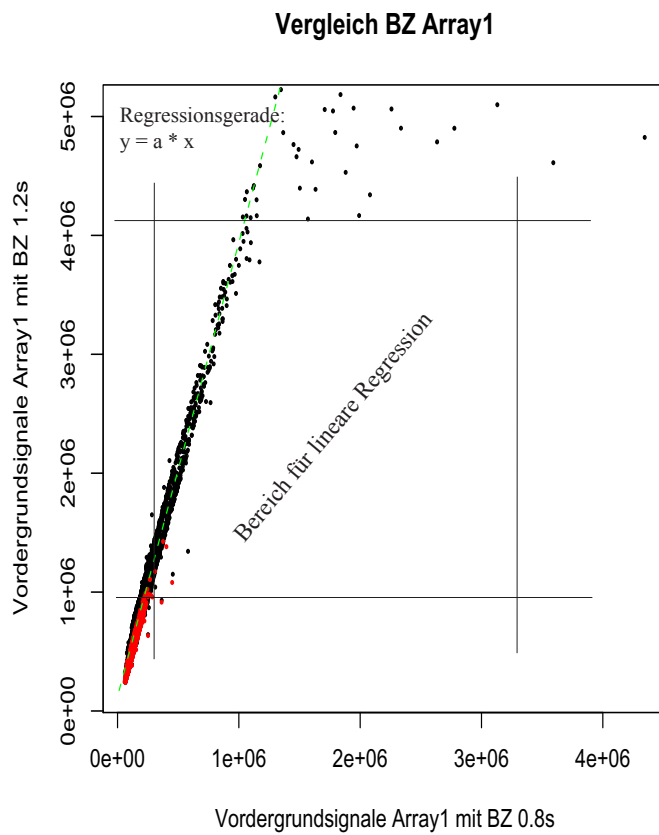
Dabei wurde mit nachfolgendem Schema gearbeitet (s. Abb. 3): Zunächst wurden für die Rohdaten Qualitäts- bzw. Filterkriterien aufgrund der Intensitäten (Signal-/Hintergrundverhältnis, Höhe der Signalintensität, Artefakte etc.) festgelegt. Nach diesen Kriterien erfolgte im Anschluß die Gewichtung der einzelnen Gene (*Spots*). Neben der Gewichtung wurde auch für jede Arrayanalyse eine sog. Hintergrundkorrektur durchgeführt. Damit möchte man Inhomogenitäten im Hintergrundsignal des Arrays, wie sie beispielsweise durch unspezifische Hybridisierungen bzw. Autofluoreszenz hervorgerufen werden, korrigieren um Datenverzerrungen zu vermeiden. Bei der Hintergrundkorrektur gibt es verschiedene Verfahren (globale, regionale und lokale Hintergrundkorrektur), die je nach Bedarf anzuwenden sind (Simon et al., 2003) Die bei dieser Arbeit vorgenommenen Hintergrundkorrekturen basierten auf globalen und lokalen Hintergrundkorrekturen. Dabei wird ein konstanter Wert für den Hintergrund festgelegt und dieser von jedem Gen subtrahiert. Anschließend wurden die unterschiedlichen Belichtungszeiten der Arrays unter Berücksichtigung der Gewichtung zusammengefasst.

Die weitere Auswertung der Daten erfolgte über die Berechnung einer linearen Regression für geeignete Intensitätsbereiche. Dabei werden, falls zwei oder mehrere Belichtungszeiten (BZ) vorhanden sind, für jede BZ die Intensitätsbereiche (der Vordergrundsignale) ermittelt, in denen *Spots* (Gene) weder gesättigt sind noch in den Niedrigsignalbereich fallen (d.h. Vordergrundsignal < Vordergrundsignal der Negativkontrollen). Außerdem werden beim Zusammenfassen der einzelnen BZ die bereits genannten Qualitäts-/Filterkriterien berücksichtigt. Für je zwei Belichtungszeiten wird in der Schnittmenge der beiden ermittelten Intensitätsbereiche eine Regressionsgerade an die Nettosignale der beiden BZs angelegt. Dabei wird für alle Gene, deren Nettosignale in den jeweils ermittelten Intensitätsbereiche liegen, ein Faktor  $a$  bestimmt, so dass die Gerade  $y = a \cdot x$ , die Daten möglichst gut beschreibt. In der nachfolgenden Abbildung ist solch eine lineare Regression, am Beispiel eines Arrays aus den Validierungsexperimenten (s. Kap. 3.5) veranschaulicht.

<sup>4</sup> [www.r-project.de](http://www.r-project.de)

<sup>5</sup> <http://bioinf.wehi.edu.au/limma/usersguide.pdf>





Dabei ist die grüngestrichelte Linie eine Regressionsgerade mit  $y = a \cdot x$ , wobei  $a$  die Steigung der Regressionsgeraden ist. Die waagerechte obere und untere graue Linie kennzeichnet den Bereich für die lineare Regression, die zerstreuten Punkte oberhalb der waagerechten Linie sind die Gene/*Spots*, die in der Signalsättigung liegen und somit nicht für die lineare Regression berücksichtigt werden. Analog verhält es sich mit den rot gekennzeichneten Punkten, die wiederum im Niedrigsignalbereich liegen und somit ebenfalls nicht bei der linearen Regression berücksichtigt werden. Die X- und Y-Achsen des Graphen repräsentieren die beiden Intensitätsbereiche.

Je nach Hybridisierungsart (kompetitiv / nicht-kompetitiv) wurden die *log-ratios* und *log-adds* bzw. die *log-2-Nettosignale* gebildet.

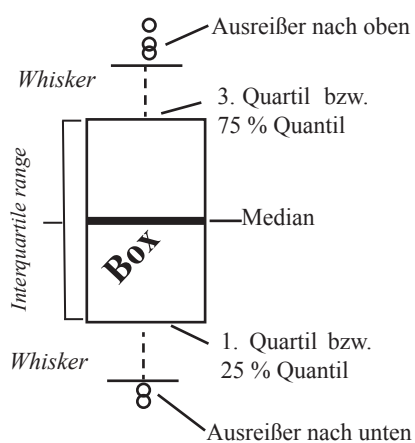
Mit der Normalisierung der Microarrays wurde die Auswertung fortgesetzt. Je nach Hybridisierungsart stehen verschiedene Normalisierungsverfahren zur Verfügung.

Für Zwei-Farben Experimente (kompetitive Hybridisierung) stehen die lineare oder die globale Normalisierung, die intensitäts-abhängige Normalisierung, sowie lokal-abhängige Normalisierung zur Auswahl (Simon et al., 2003). Bei den kompetitiven Hybridisierungen erfolgte die Normalisierung zunächst innerhalb eines Arrays über eine sog. *location normalization* mit *loess*-Regression für jede *print tip group*. Der Vorteil dieser *print tip group* Normalisierung ist, dass eventuell vorhandene, systematische Verzerrungen nur innerhalb der *print tip groups* beseitigt werden, was mit einer globalen Normalisierung über den gesamten Array, nicht erreicht wird. Im nächsten Schritt erfolgte die Normalisierung zwischen den Arrays über eine sog. *scale normalization*. *Log ratios* verschiedener Arrays können meist unterschiedliche Streuungen aufweisen, die eine Skalenanpassung erfordern. Mit der *scale normalization* wird die Skalenanpassung durchgeführt.

Bei der nicht-kompetitiven Hybridisierung, auch *Single*-Hybridisierung genannt, erfolgt lediglich eine Normalisierung zwischen den Arrays. Auch hier bieten sich je nach Fragestellung verschiedene Verfahren an. Für die hier durchgeführten *Single*-Hybridisierungen wurde eine sog. *quantile normalization* benutzt. Dabei geht man von der Annahme aus, dass alle zu untersuchenden Arrays dieselbe Intensitätsverteilung haben und passt die Daten der einzelnen Arrays entsprechend an. Im nachfolgenden Schritt wurden die Spot-Replikate zusammengefasst und mittels statistischer Analyse ausgewertet.

Neben der Reproduzierbarkeit der Experimente, wurde das Augenmerk vor allem auf die Suche nach differentiellen Genen gelegt. Desweiteren erfolgte ein Vergleich der Ergebnisse mit unabhängigen Methoden (quantitativer *real-time* PCR, kommerzieller DNA-Microarray).

Zur Beurteilung der Reproduzierbarkeit z.B. von technischen Replikaten (Kap. 3.6) kann ein sog. Boxplot (exakt: *Box and Whisker-Plot*) verwendet werden. Bei dem Boxplot handelt es sich um eine graphische Darstellung des Medians und der Variabilität der Daten in einer Fünf-Punkte-Zusammenfassung:



Die dicke waagerechte Linie in der Mitte der Box stellt den Median dar, über und unter dem 50 % der Daten liegen. Die obere bzw. untere Begrenzung der Box ist das 75 % bzw. 25 % -Quantil, unterhalb dem mindestens 75 % bzw. 25 % der Daten liegen. Der *Interquartile range* (IQR) ist die Differenz zwischen 75 %- und 25 %-Quantile und beschreibt die Größe der Box. Die sog. *Whisker*, als waagerechte Linien mit senkrechtgestrichelten Linien eingezeichnet, grenzen den Bereich ein, unter bzw. über dem die „Ausreißer“ der Daten liegen. Die Endpunkte der *Whisker* sind hier für den kleinsten Wert  $\geq 25\%$  Quantil  $- 1,5 \times \text{IQR}$  und für den größten Wert mit  $\leq 75\%$  Quantil  $+ 1,5 \times \text{IQR}$  festgelegt (in der Zeichnung nicht maßstabgetreu dargestellt).

Des weiteren können mittels Boxplot auch Signal-/ Hintergrundverhältnisse bewertet werden (s. Kap. 3.4.3). So ist der Boxplot je nach Fragestellung vielseitig verwendbar.

Nach dem sog. *preprocessing* der Daten stehen gem. Abb. 3 verschiedene statistische Analysemethoden (z. B. Multiples Testen, Klassifikation und *Clustering*) zur Verfügung. Mit dem multiplen Testen lassen sich die Genexpressionsprofile auf differentielle Genregulation jedes einzelnen Genes hin untersuchen. Dabei erfolgt die statistische Prüfung zwischen der Nullhypothese, dass keine Unterschiede zwischen zwei untersuchten Zuständen (hier beispielsweise der Vergleich zwischen zwei Tumorkollektiven) auftreten. Demgegenüber steht die Alternativhypothese, die besagt, dass Unterschiede zwischen den Zuständen bestehen. Die Alternativhypothese wird dann angenommen, wenn ein bestimmter Wert (*p*-Wert) eine zuvor festgelegte Grenze (Irrtumswahrscheinlichkeit  $\alpha$ ) unterschreitet. Mit dem statistischen Test, beispielsweise *t*-Test, wird über die Annahme oder Ablehnung der Nullhypothese entschieden. Dieser Test macht jedoch keine Aussage über das Expressionsmaß. Mit dem berechneten Wert aus der Teststatistik ist es dann möglich, aufgrund zweier Stichproben zu entscheiden, ob differentielle Expression zwischen den betrachteten Zuständen besteht.

Die Klassifikation gehört zu den überwachten Lernverfahren. Hier ist die Klassenzugehörigkeit bereits erstellter Genexpressionsprofile bekannt. Dabei wird über eine Trainingsmenge (Bsp. Tumor / Nichttumor) eine Funktion, ein sog. Klassifikator bestimmt, welcher die Klassen aufgrund der Genexpressionsprofile möglichst genau vorhersagen soll. Aufgrund dieser Information können neue Genexpressionsprofile von Testmengen in eine der bestehenden Klassen zugeordnet werden. An einem Beispiel wird dieses Verfahren in Kap. 3.2.1 noch etwas näher beschrieben, wo aufgrund von Genexpressionsprofilen ein Klassifikator zur Unterscheidung zweier humaner Brustkrebszelllinien bestimmt wurde.

Neben der Klasseneinteilung können mit der Klassifikation aber auch Vorhersagen hinsichtlich des Therapieansprechens oder Überlebens gemacht werden. Des Weiteren kann die Klassifikation aber auch eine Aussage über den Typ einer Erkrankung machen.

Das sog. *Clustering* gehört zu den unüberwachten Lernverfahren, da hier zwar aus einer Experimentenserie eine gewisse, meist jedoch nicht exakte Anzahl von Klassen bestimmt wird bzw. u. U. man nicht einmal weiß, ob überhaupt Klassen vorhanden sind und zudem die tatsächliche Klassenzugehörigkeit jedoch nicht zum Lernen zur Verfügung steht.

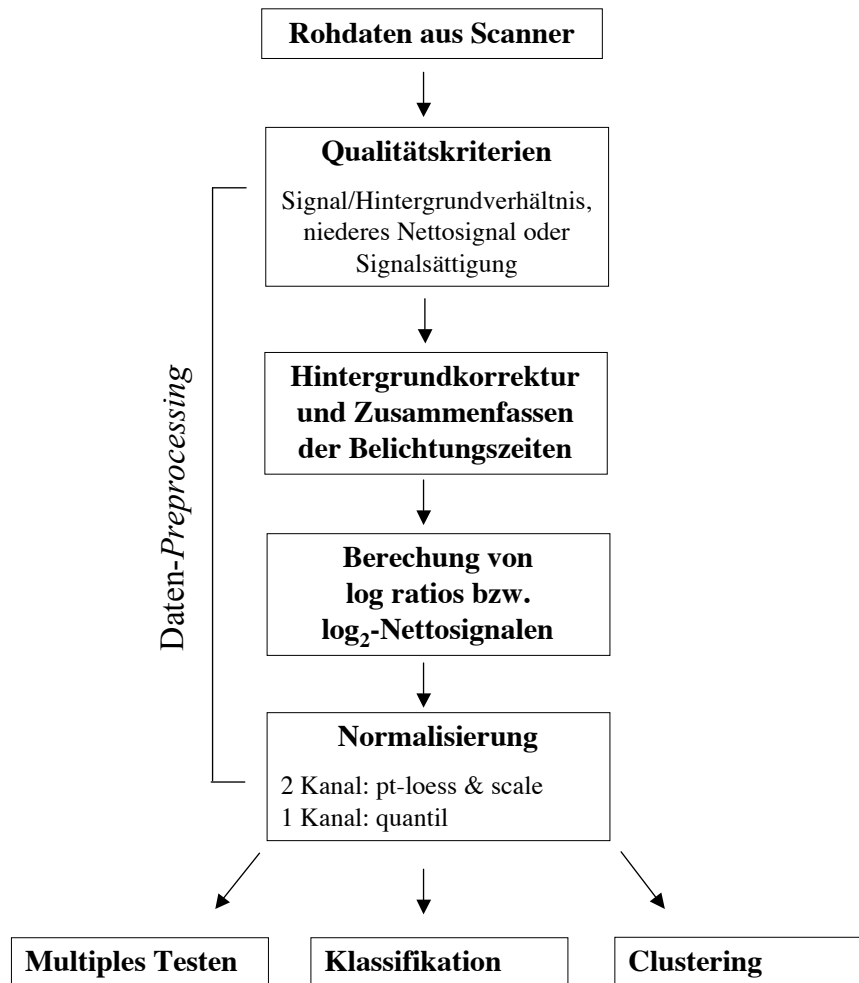


Abb. 3: Auswertungsschema der Rohdaten mit unterschiedlichen Möglichkeiten der Datenanalyse. Nach dem Daten- *Preprocessing* kann die weitere Datenanalyse, je nach Fragestellung durch multiples Testen, der Klassifikation, dem *Clustering*, oder allen drei Schritten zusammen, erfolgen.

### 3. Ergebnisse

#### 3.1 Genauswahl anhand von Studien

Grundlage für diese Studie war ein in meiner Diplomarbeit aufgebauter, EST-basierter DNA-Microarray mit 280 Sequenzen, die aus verschiedenen Publikationen ausgewählt wurden (Iyer et al., 1999, Kononen et al., 1999, Schramel et al., 1999, Golub et al., 1999, Khan et al., 1999, Alizadeh et al., 2000). Auf diesem DNA-Microarray befanden sich bereits cDNAs von bekannten Schlüsselmolekülen, welche bei der Kontrolle von Zellzyklus, Proliferation, Differenzierung und Apoptose eine Rolle spielen (z. B. Cycline, Wachstumsfaktoren wie EGF, FGF, TGF beta sowie deren Rezeptoren, Transkriptionsfaktoren, Onko- bzw. Tumorsuppressorgene oder verschiedene Cytokeratine).

Bei einer Sequenzüberprüfung von 45 Genen wurden 40 Gene in ihrer Sequenz bestätigt (im Anhang mit \* gekennzeichnet). Aufgrund seines späteren Einsatzgebietes, nämlich zur therapiebegleiteten Prognose bei Brustkrebserkrankungen, wurde bei der Neukonzeption des DNA-Microarrays darauf geachtet, die vorhandenen ESTs möglichst durch bereits vom RZPD sequenzbestätigte Klone zu ersetzen. Dieses Genset wurde auf zunächst 700 Gene erweitert. Das erweiterte Genkollektiv gründet sich auf ausführlicher Literaturrecherche unabhängiger gesamtgenomischer Mikroarrayanalysen und deckt sowohl therapeutisch bedeutsame Gene, als auch Sequenzen, die Tumortyp, den Östrogenrezeptorstatus und die Signaltransduktionswege repräsentieren, ab. Unter Einbeziehung der Studie von West et al. wurde das Genkollektiv um 100 Sequenzen, die den Östrogenrezeptorstatus repräsentieren, erweitert (West et al., 2001). Da die gesamtgenomischen DNA-Microarrays teuer in der Anschaffung und komplex in der Datenanalyse sind, schlägt Lacroix in seiner literaturbasierten Übersichtsarbeit einen sog. *Low density* DNA-Microarray mit ca. 250 Sequenzen für spezifische Fragestellungen vor. (Übersicht: Lacroix et al., 2002). Ein Teil dieser Sequenzen (ca. 40) war bereits Bestandteil des o.g. DNA-Microarrays. Das verbliebene Kollektiv wurde Bestandteil des neuen DNA-Microarrays. Diese Genauswahl erstreckt sich über die gesamte Bandbreite zellulärer Funktionen, wie beispielsweise Zellzyklusregulation, Angiogenese, Proteolyse, Chemoresistenz und Hormonsensitivität.

Weitere 70 Sequenzen wurden aus einer Studie von van de Vijver et al. aufgenommen. Diese Studie beschreibt, dass ein Expressionsprofil von 70 Genen ausreicht, eine Aussage bei Patienten mit guter und auch schlechter Prognose-Signatur hinsichtlich einer 10 jährigen Überlebensrate ( $94,5\% \pm 2,6$  und  $54,6\% \pm 4,4$ ) bzw. einer Bildung von Fernmetastasen innerhalb von 10 Jahren ( $85,2\% \pm 4,3$  und  $50,6\% \pm 4,5$ ) zu treffen und zwar deutlich besser als mit herkömmlichen Standardmethoden der klinischen Praxis, oder nach histologischen Kriterien (van de Vijver et al., 2002). Im weiteren ist diese Studie als eine Fortsetzung aus der Arbeit von van't Veer anzusehen (van't Veer L et al., 2002). Die Unterschiede der beiden Studien liegen in der Vorhersage des Krankheitsverlauf innerhalb 5 Jahren, einem vergrößerten Patientenkollektiv (295 : 98) sowie in der zusätzlichen Untersuchung von lymphknotenpositiven Patienten bei van de Vijver et al.. In der Arbeit von Sørli et al wurde eine Klassifikation von Tumoren hinsichtlich ihrer variierenden Genexpression durchgeführt mit der Absicht, die Tumorcharakteristika mit dem klinischen Verlauf zu korrelieren (Sørli et al., 2001). Eines der Ergebnisse war ein Gencluster, das hoch mit dem klinischen Bild der Patienten korrelierte. Ein Teil dieses Genkollektivs wurde ebenfalls auf dem DNA-Microarray implementiert.

Die Studie von Sørliie ist als Folgearbeit von Perou et al. zu betrachten, der durch ein DNA-Microarray basiertes Tumorscreening das untersuchte Tumorkollektiv subklassifizieren konnte (Perou et al., 2000).

Neben Genen aus den o.g. Studien wurden weitere Gene auch von den beiden Kooperationspartnern (RBK, IZI) vorgeschlagen. Dabei sind für die Klinik besonders die Gene mit therapeutischer Bedeutung (Bsp. Hormonsensivität oder Chemoresistenz) interessant (Kudoh et al., 2000, Sortiriou et al., 2002, Inoue et al., 2002, Faneyte et al., 2003, Cunliffe et al., 2003). Für die Kooperationspartner IZI sind vor allem Gene für die Regulation apoptotischer Prozesse relevant. Aus diesen drei Genkollektiven (IGB, RBK und IZI) wurde nach Abgleich hinsichtlich mehrfach vorhandener Gene ein DNA-Microarray mit zunächst 743 Sequenzen konzipiert (vollständige Genliste s. Anhang). In dieser alphabetisch geordneten Genliste sind neben der Abkürzung des Gennamens, die Angabe der mRNA-Referenzsequenz Nummer, der vollständige Name sowie die RZPD-Bestell-Nr. des Klonen angegeben.

Die Auswahl der ESTs von o.g. DNA-Microarray wurde ebenfalls bereits in der Diplomarbeit Horschinek von 2001 näher beschrieben.

Die ESTs für den IGB-Chip wurden wie folgt herausgesucht. Über den Gennamen wurde unter <http://www.dkfz-heidelberg.de/GeneCards/> das betr. *Unigene Cluster*, also die *Hs. No* herausgesucht. Die *Hs. No* wurde bei der RZPD-Homepage unter <http://www.rzpd.de/cgi-bin/products/getGene.pl.cgi> bei „search for cluster ID“ eingegeben. Zusätzlich wurden über die Buttons „Library“ die Genbibliothek „Human Unigene Set-RZPD 3“ und über „specified sequence“ nur die verifizierten Sequenzen ausgewählt. So ergab sich beispielsweise für das Gen FGF2 (*fibroblast growth factor 2*) die *Hs. No* 284244, die wiederum über die ausgewählten Kriterien zu der Klon-Nummer: HU3\_p972D1296D führt, unter welcher dann diese Sequenz beim RZPD bezogen werden kann. Der andere Teil der bestätigten Sequenzen stammt aus der RZPD-Library No 998. Auch hier handelt es sich um cDNA-Klone, allerdings aus einer bereits seit 2000 bestehenden Genbibliothek. Diese Klonsammlung wurde vom *Lawrence Livermore National Laboratory* übernommen, wobei die Klonannotation nicht durch Sequenzierung bestätigt ist. So sollen lt. RZPD bis zu 20 % dieser Klone nicht mit ihrer Annotation übereinstimmen. Eine stichprobenartige Sequenzierung dieser Klone ergab jedoch, dass 40 von 45 Sequenzen in ihrer Annotation bestätigt werden konnten.

Allerdings konnten nicht alle Gene aus o.g. Bibliothek (*Human Unigene Set-RZPD 3*) herausgesucht werden. Ersatzweise wurden die verbliebenen Sequenzen (etwa 30 %) aus der RZPD-Bibliothek No. 958 bezogen. Ein wesentlicher Unterschied zu den beiden anderen Genbibliotheken besteht darin, dass die cDNA-Klone aus der *Library No 958* eine Chloramphenicol- statt einer Ampicillinresistenz tragen, lt. RZPD auch nicht sequenzbestätigt sind und als Konsequenz etwa 20 % dieser Klone nicht mit der betr. Genannotation übereinstimmen (analog *Library No 998*). Das RZPD empfiehlt in beiden Fällen eine Überprüfung der Genannotation, um fehlerhafte Sequenzen bzw. Artefakte zu identifizieren. Weitere Angaben über die Genbibliotheken des RZPD sind unter [http://www.rzpd.de/cgi-bin/products/rzpd\\_products.pl.cgi](http://www.rzpd.de/cgi-bin/products/rzpd_products.pl.cgi) zu finden.

### 3.2 Überprüfung der Reproduzierbarkeit

Ausgehend von einem Mini-Array (192 Sequenzen), sollte die prinzipielle Machbarkeit anhand eines zellbasierten Modellsystems untersucht werden. Der hier maßgebliche Parameter ist vor allem die Reproduzierbarkeit unter Einfluss selbstbeschichteter Oberflächen.

Zunächst wurden die basalen Transkriptionsprofile der beiden humanen Brustkrebszelllinien, MCF 7 und MDA-MB 231 erstellt. MCF7 ist eine ÖR-positive Zelllinie mit niedrigerem Invasionspotential und wird seit Jahrzehnten als Modellsystem für ÖR-positive Brusttumore benutzt. Ihr „Gegenspieler“ MDA-MB 231 ist eine ÖR-negative Brustkrebszelllinie mit hohem Invasionspotential. In dieser Experimentenserie erfolgten 6 kompetitive Hybridisierungen mit je 50 µg Gesamt-RNA, zwischen Cy5-markierter cDNA von MCF 7 und Cy3-markierter cDNA von MDA-MB 231. Mit anschließenden *dye-swap*-Experimenten sollten die Einbau-raten der markierten Nukleotide untersucht werden.

In den beiden Abbildungen sind die Effekte des *dye-swap*-Experiments deutlich sichtbar. Gene die in 4a) als grün erscheinen werden in 4b) rot dargestellt (siehe markierter Bereich). Zur Überprüfung der Qualität bzw. der Reproduzierbarkeit wurden die jeweiligen Korrelationen zwischen den Arrays berechnet. Die Berechnung erfolgte auf der Basis normalisierter, logarithmierter Signalverhältnisse (*log ratios*), wobei hier nur diejenigen Gene mit einbezogen wurden, die auf mindestens 2 von 3 Arrays unter Berücksichtigung der Qualitätskriterien (s. Kap. 2.2.15), detektierbar waren.

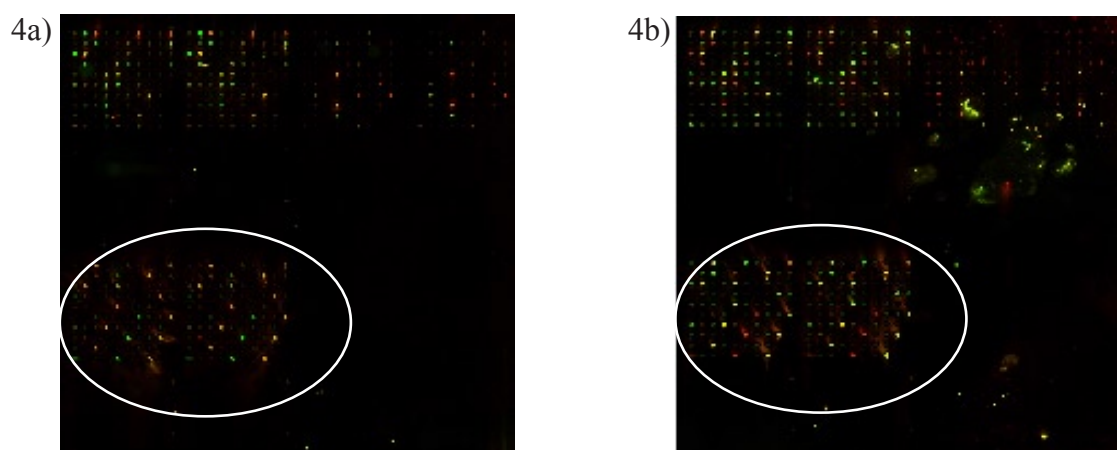


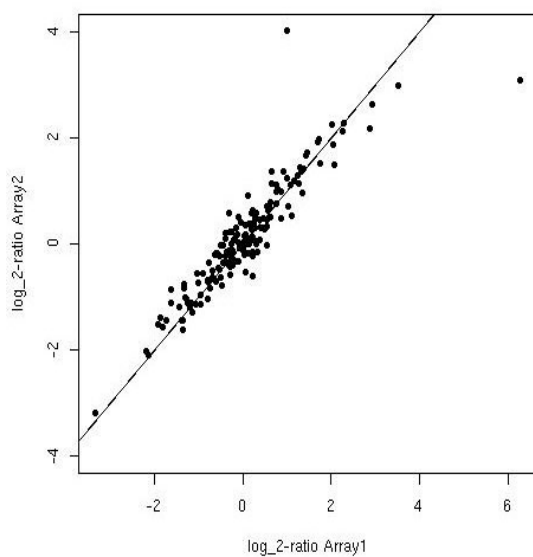
Abb. 4: Microarray Hybridisierungsbilder 4a) Transkriptionsprofil von MCF 7 Cy5- und MDA-MB 231 Cy3-markierter cDNA; 4b) Transkriptionsprofil von MCF 7 Cy3- und MDA-MB 231 Cy5- markierter cDNA. Die Expression einzelner Gene wird in der Analyse als Mischfarbe entsprechend ihrer Anteile der jeweiligen cDNA dargestellt.

cDNA Markierung	Array Vergleich	Korrelation
MCF 7 Cy5, MDA-MB 231Cy3	<b>1 mit 2</b>	<b>0,96</b>
	1 mit 3	0,94
	2 mit 3	0,95
MCF 7 Cy3, MDA-MB 231Cy5	<b>1 mit 2</b>	<b>0,88</b>
	1 mit 3	0,77
	2 mit 3	0,93

Tab. 3: Qualitätsüberprüfung der durchgeführten Experimente mit Gesamt-RNA anhand der berechneten Korrelationen zwischen den Arrays. Diese und alle folgenden Berechnungen sowie die zugehörigen Abbildungen wurden von Herrn S. Winter, Mathematisches Institut A der Universität Stuttgart, durchgeführt bzw. erstellt.

Neben der Korrelation zwischen den Arrays sind aber auch die Signalverhältnisse (*ratios*) von Bedeutung. In einem sog. *Scatter Plot* lassen sich außer über die Reproduzierbarkeit (Korrelations Koeffizient  $\geq 0,90$ ) auch Aussagen über die Höhe von Signalverhältnissen ( $> 0$ ) sowie deren Verteilung machen. Aus den Nettosignalen der beiden Farbkanäle erfolgt die Bildung der sog. *log ratios* (logarithmierte Signalverhältnisse), die anschließend normalisiert und über die Spotreplikate gemittelt werden. Diese Werte werden dann in *Scatter Plots* für je zwei Arrays gegeneinander aufgetragen (Abb.5 a,b). Exemplarisch sind hier die Ergebnisse aus den Arrayexperimenten 1 und 2 (s. in Tab. 3 hervorgehoben) dargestellt.

5a)



5b)

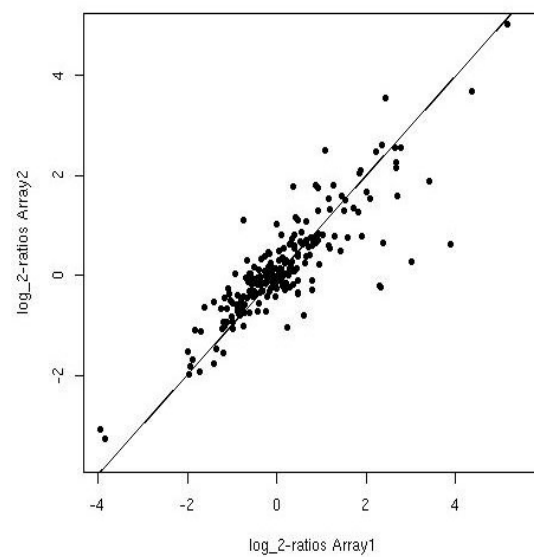


Abb.5: *Scatter Plots* aus den Hybridisierungsexperimenten mit Proben aus Gesamt RNA (50  $\mu\text{g}$ ). 5a) Darstellung der *log ratios* von den *Arrays* 1 und 2 mit MCF 7 Cy5, MDA-MB 231 Cy3. 5b) Darstellung der *log ratios* von den *Arrays* 1 und 2 des entsprechenden *dye-swap* Experimentes.

Gemäß Tabelle 3 zeigen die *dye-swap* Experimente deutlich schlechtere Korrelationen (Ausnahme: Array 2 mit 3). Neben den bekanntermaßen unterschiedlichen Einbauraten der Fluoreszenznukleotide könnte dieser Effekt zusätzlich auch durch Inhomogenitäten der selbstbeschichteten Oberflächen einzelner Arrays oder aufgrund unterschiedlicher Hybridisierungsqualität, hervorgerufen bzw. weiter verstärkt worden sein.

Ogleich die Korrelationen zwischen den beiden Experimenten erheblich schwanken, konnte mit diesem Modellsystem die prinzipielle Machbarkeit gezeigt werden.

Um allerdings künftig störende Einflüsse zu reduzieren, sollen für weitere Arrayexperimente nur kommerziell beschichtete Oberflächen verwendet werden. Ferner soll durch die Erweiterung des Genkollektivs sowie durch Erhöhung der Spotreplikate die statische Auswertung vereinfacht und damit robustere Daten generiert werden. In Kapitel 3.2.1 wurden diese Änderungen bereits mit einbezogen.



### 3.2.1 Nutzung von Transkriptionsprofilen zur Klassifikation von Zelllinien

Mit dieser Versuchsreihe wurden erneut basale Transkriptionsprofile der beiden humanen Brustkrebszelllinien MCF7 und MDA-MB 231 als Modellsystem erstellt. Die Hybridisierungen erfolgten kompetitiv (s. Kap. 3.2). Statistische Auswertung der Daten mittels eines t-Tests ergab ca. 170 signifikant differentiell exprimierte Gene zwischen den beiden Zelllinien.

Um zu untersuchen, inwieweit es möglich ist, anhand des entwickelten Microarrays Vorhersagen über den Subtyp eines untersuchten Tumorgewebes (und damit auch über das Therapieansprechen, die Überlebenszeit u. ä.) zu treffen, erfolgte eine Klassifikation der Proben innerhalb dieser Versuchsreihe. Ziel war hierbei, eine Funktion (sog. Klassifikator) zu ermitteln, die auf der Basis weniger Gene (sog. Klassifikatorgene) den Zelllinien-Typ, also MCF7 oder MDA-MB 231, mit möglichst kleinem Vorhersagefehler richtig bestimmt.

Insgesamt wurden die Daten von je 4 MCF7- und 4 MDA-MB 231-Proben verwendet. Als Kandidaten für die Klassifikatorgene kamen dabei die ca. 170 differentiell exprimierten Gene in Frage, anhand denen nun die Berechnung des Klassifikators erfolgte. Dazu wurde zunächst für jede Kombination von je 2 MCF7- und 2 MDA-MB 231-Proben (sog. Trainingsmenge) sowie für jede Anzahl von Genen (also von einem Gen bis zu den ca. 170 Genen) eine Funktion ermittelt, die die beiden Klassen (MCF7 bzw. MDA-MB 231) aufgrund der Genexpressionswerte voraussagt. Zur Bestimmung, welche Gene konkret für eine betrachtete Anzahl von Genen dabei berücksichtigt werden, wurde ausgehend von allen Kandidatengen in jedem Schritt das Gen ausgeschlossen, welches die geringste Vorhersagekraft zur Unterscheidung der beiden Zelllinien hat (d.h. so dass der Vorhersagefehler für die betrachtete Anzahl von Genen bei Ausschluss gerade dieses und nicht eines anderen Genes am kleinsten ist).

Für die beiden jeweils nicht in den Trainingsmengen verwendeten MCF7- bzw. MDA-MB 231-Proben (sog. Testmenge) wurde nun der Vorhersagefehler bestimmt und diejenige Funktion als Klassifikator ausgewählt, die den kleinsten Vorhersagefehler besitzt. Der hierdurch ermittelte Klassifikator ist optimal (d.h. ohne Vorhersagefehler) und berechnet sich auf der Basis von nur 9 Klassifikatorgenen (Tab. 4).

Das obige Vorgehen mit einer Einteilung der Proben in eine Trainings- und eine Testmenge ist wichtig, um zu gewährleisten, dass der Klassifikator auch auf neuen Daten, die nicht seine Berechnung einfließen, einen geringen Vorhersagefehler besitzt.

Klassifikator Gene	Mittelwert		Standard-Abw.		BH-adj. p-Werte	Statistik (t-test)
	MCF 7	MDA	MCF 7	MDA		
PLAUR	15,82	19,32	0,28	0,15	0,0081	-17,70
PTN	21,27	17,78	0,19	0,25	0,0045	45,02
KRT19	22,66	18,92	0,37	0,18	0,0051	32,36
TPD52L1	20,43	15,64	0,74	0,10	0,0103	13,72
EMP2	19,90	16,48	0,37	0,20	0,0116	12,60
VEGF	21,81	18,65	0,13	0,32	0,0052	31,35
FLT4	18,04	14,63	0,11	0,43	0,0115	12,82
C17orf28	18,91	15,43	0,40	0,27	0,0088	15,14
CRABP2	21,06	16,83	0,71	0,32	0,0086	15,36

Tab. 4: Klassifikator-Gene zur Untersuchung der beiden Zelllinien.

Bei den Mittelwerten in Tabelle 4 handelt es sich um die jeweiligen gemittelten, logarithmierten Signalintensitäten, die den Unterschied zwischen den beiden Zelllinien zeigen. Des weiteren steht in Tabelle 4 eine negative Teststatistik für eine stärkere Expression in MDA-MB 231 als in MCF7. Bei einer positiven Teststatistik verhält es sich umgekehrt. Die Bestimmung der 170 differentiellen Gene erfolgte über einen t-Test und die anschließende Auswahl der Gene, die einen (nach dem Verfahren von Benjamini-Hochberg (BH)) p-Wert  $< 0,05$  aufwiesen.

Die gewöhnliche t-Teststatistik für 2 unverbundene Stichproben wird wie folgt berechnet:

$$(1) \quad t = \frac{\text{Mittelwert}(MCF7) - \text{Mittelwert}(MDA)}{\sqrt{\left(\frac{1}{n} + \frac{1}{m}\right) * s^2}}$$

$$(2) \quad s^2 = \frac{(n-1) * SD\_MCF7 + (m-1) * SD\_MDA}{n + m - 2} \quad \begin{array}{l} n = \text{Gruppengröße MCF 7} \\ m = \text{Gruppengröße MDA-MB 231} \end{array}$$

Hier wurde die gewöhnliche t-Teststatistik gegen eine sog. *moderated* t-Teststatistik ersetzt.

In einem Folgeexperiment mit den beiden Zelllinien MCF 7 vs MDA-MB 231 über 4 Arrays konnte der zuvor berechnete Klassifikator ohne Vorhersagefehler auf die neuen Daten angewendet werden. Somit konnte anhand der Expressionsprofile deutlich eine präzise Trennung der beiden Zelllinien gezeigt werden.

### 3.3 Lineare RNA-Amplifikation

In den Experimentenfolgen aus Kap. 3.2 und 3.2.1 war der Einsatz von Gesamt-RNA mit je 50 µg pro cDNA-Synthese recht hoch. Bei Verwendung von Zelllinien als Quelle stellt dies jedoch kein Limit dar. Anders verhält es sich bei der Nutzung von Tumorproben. Lt. Angaben des klinischen Kooperationspartners können je nach Größe des Tumors, oft nur mit Mühe zweistellige µg-Mengen an Gesamt-RNA isoliert werden. Damit steht oftmals kaum genug Material für eine Hybridisierung zur Verfügung. Gegenwärtig liegt der Bedarf für Transkriptionsanalysen mit Gesamt-RNA zwischen 10 µg bei Direktmarkierung und 5 µg bei indirekter Markierung. Um jedoch auch Biopsiematerial analysieren zu können, wird von den Klinikpartnern 1 µg Ausgangsmenge an Gesamt-RNA angestrebt.

Im Rahmen dieser Arbeit soll die lineare RNA-Amplifikation als Teilaspekt einer partiellen Sensitivitätssteigerung untersucht werden. Deshalb soll mit der RNA-Amplifikation der „Materialmangel“ kompensiert werden, um damit genügend Ausgangsmaterial für Transkriptionsanalysen bereitzustellen. So kann je nach Amplifikationsmethode eine bis zu tausendfache Erhöhung der mRNA-Ausgangsmenge (unter Annahme von 1% mRNA) erzielt werden.

Zur Ermittlung der Amplifikationsrate nach einer Amplifikationsrunde (der hier verwendete Kit bietet grundsätzlich die Möglichkeit für zwei Amplifikationsrunden) wurde die RNA-Amplifikation mit Gesamt-RNA aus den beiden humanen Brustkrebszelllinien MCF 7 und MDA-MB 231, in je zwei Ansätzen unterschiedlicher Ausgangsmengen, parallel durchgeführt. Die notwendigen Arbeitsschritte erfolgten nach Hersteller Protokoll (Arcturus) und sind in Abb. 7 schematisch dargestellt.

Probe	Ausgangsmenge [µg]	Erzielte Menge [µg]	Amplifikationsrate
Kontroll RNA	2	40	2000*
MCF 7	0,5	18	3600*
MCF 7	1,0	23,3	2330*
MDA-MB 231	1,75	50	2857*
MDA-MB 231	3,5	37,7	1077*

Tab. 5: Ergebnisse der linearen RNA-Amplifikation unterschiedlicher Gesamt-RNA-Ausgangsmengen. \* Amplifikationsrate ausgehend von 1 %-igen mRNA-Anteil der Proben.

Nach der Amplifikation wurden die Proben in einem RNA-Gel überprüft. Von jeder Probe wurden 2 µg aufgetragen.

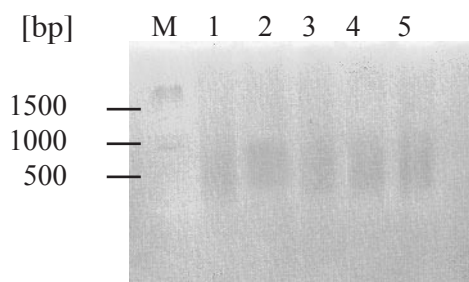


Abb. 6. RNA-Gel, 1 %, sybrgold-gefärbt: M = 1 kb Marker, 1 = 2 µg Kontroll RNA, 2 = MDA-MB 231 aus 3,5 µg-Ansatz, 3 = MDA-MB 231 aus 1,75 µg-Ansatz, 4 = MCF 7 aus 1 µg-Ansatz, 5 = MCF 7 aus 0,5 µg-Ansatz

Gemäß Abbildung 6 ist bei allen 5 Proben deutlich ein charakteristischer Schmier zu sehen. Dementsprechend hat die Amplifikation bei allen 5 Proben stattgefunden. Nach dem Amplifikationsprotokoll soll der „RNA-Schmier“ zwischen 250 - 1800 bp liegen, wobei der Schwerpunkt der RNAs nach der ersten Amplifikationsrunde laut dem Protokoll bei etwa 600 bp lokalisiert ist. Im RNA-Gel erscheinen die RNAs gehäuft zwischen 1000 und 500 bp.

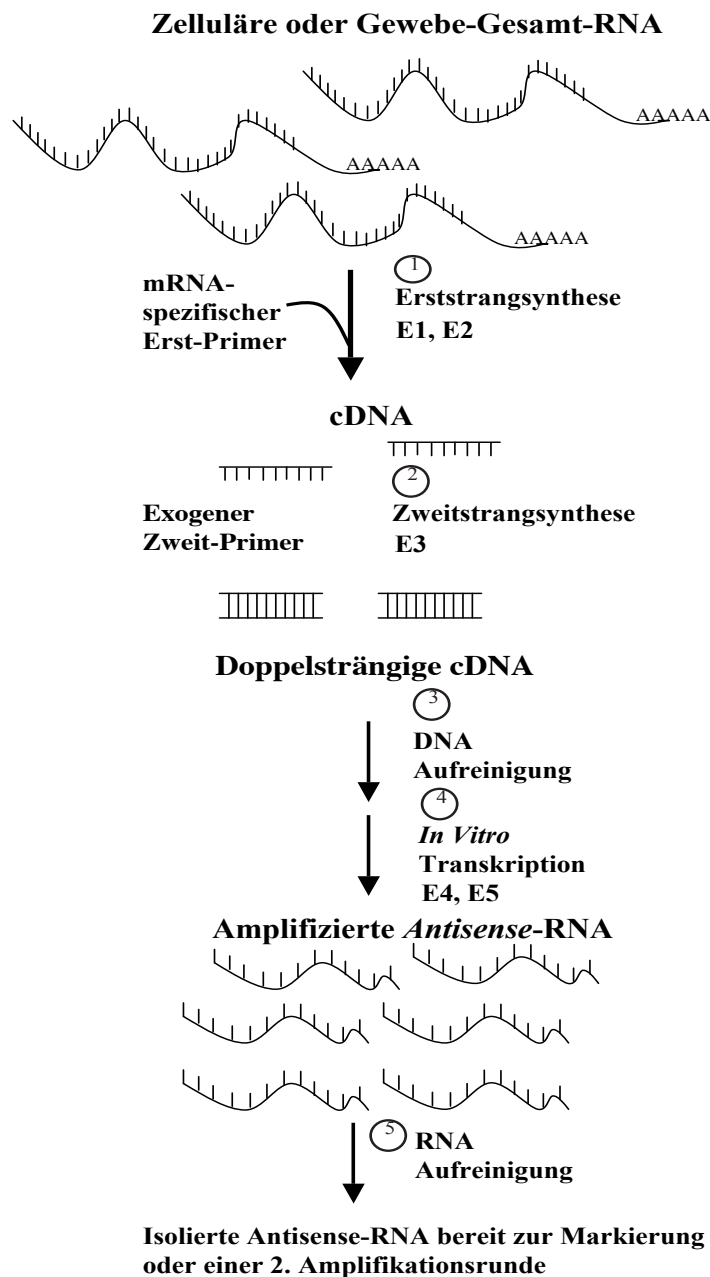


Abb. 7: Prinzip der linearen RNA-Amplifikation in 5 Teilschritten: Erststrangsynthese u.a. mit *Reverser Transkriptase* (E1 = Enzymmix 1) anschließend Verdau der RNA mit Nukleasemix (E2), Zweitstrangsynthese u.a. mit *DNA Polymerase* (E3 = Enzymmix 3), Aufreinigung doppelsträngiger cDNA, *in vitro* Transkription mit u.a. mit *T7 RNA Polymerase* (E4 = Enzymmix 4), anschließender Verdau der DNA mit DNase Mix (E5) und RNA-Aufreinigung. Genaue Zusammensetzung der Enzymgemische ist Firmengeheimnis der Arcturus GmbH.

Für die Hybridisierungsexperimente wurden die jeweiligen *Antisense*-RNAs gepoolt. Analog den Experimenten mit Gesamt-RNA wurden auch hier 6 kompetitive Hybridisierungen mit jeweils 1  $\mu\text{g}$  Antisense RNA, zwischen Cy5-markierter cDNA von MCF 7 und Cy3 markierter cDNA von MDA-MB 231 durchgeführt. Mit den *dye swap*-Experimenten sollten die Einbauraten der markierten Nukleotide untersucht werden.

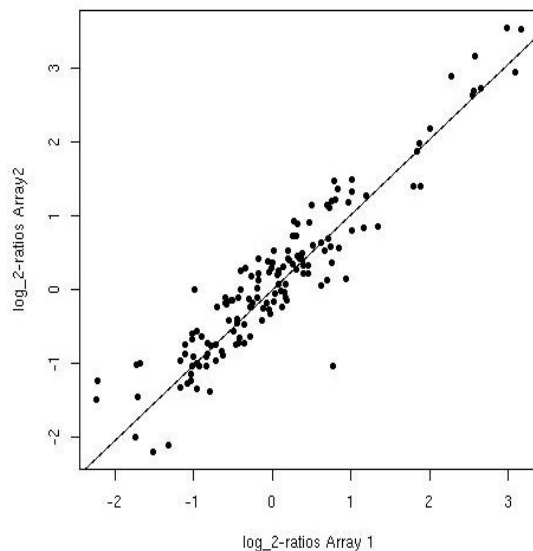
Zur Überprüfung der Qualität wurden die jeweiligen Korrelationen zwischen den Arrays, analog denen in Kapitel 3.2, berechnet.

cDNA-Markierung	Array Vergleich	Korrelation
MCF 7 Cy5, MDA-MB 231 Cy3	<b>1 mit 2</b>	<b>0,96</b>
	1 mit 3	0,96
	2 mit 3	0,95
MCF 7 Cy3, MDA-MB 231 Cy5	<b>1 mit 2</b>	<b>0,81</b>
	1 mit 3	0,84
	2 mit 3	0,93

Tab. 6: Qualitätsüberprüfung der durchgeführten Experimente mit amplifizierter RNA anhand der berechneten Korrelationen zwischen den Arrays.

Die *dye-swap* Experimente mit cDNA aus amplifizierter RNA zeigen hier ebenfalls unterschiedliche Korrelationen analog den *dye-swap* Experimenten mit Gesamt-RNA (vgl. Tab 3, 6). Abb. 8 a,b zeigt exemplarisch zwei ausgesuchte Arrayexperimente (s. in Tab. 6 hervorgehoben).

8a)



8b)

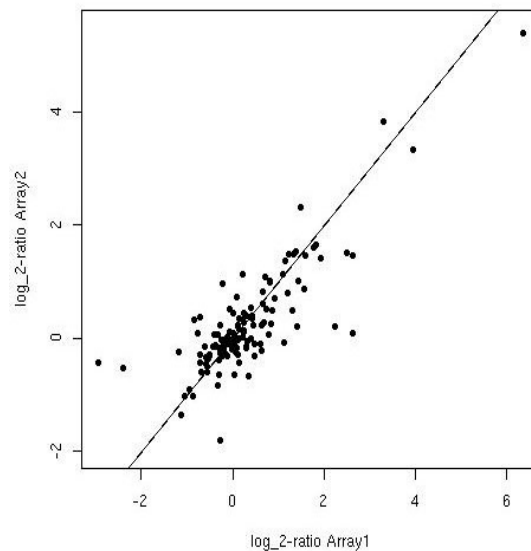


Abb.8: *Scatter Plots* aus den Hybridisierungsexperimenten mit Proben aus amplifizierter RNA (1 $\mu\text{g}$ ). 8a) Darstellung der *log ratios* von *Arrays* 1 und 2 mit MCF 7 Cy5, MDA-MB 231 Cy3. 8b) Darstellung der *log ratios* von *Arrays* 1 und 2 des entsprechenden *dye-swap* Experimentes.

In weiteren Versuchen erfolgte neben der Qualitätsprüfung von den dye-swap Experimenten auch die Berechnung der Korrelationen zwischen Hybridisierungsexperimenten mit cDNA aus Gesamt- bzw. amplifizierter RNA (s. Tab. 7). Für die Reproduzierbarkeit der linearen RNA-Amplifikation selbst wurden zwei identische Experimente mit Proben aus parallel durchgeführten Amplifikationen überprüft (s. Tab.7, 3. Zeile).

Markierung	Korrelation
Gesamt: MCF 7 Cy5, MDA-MB 231 Cy3 vs MCF 7 Cy3, MDA-MB 231 Cy5	0,79
Amp: MCF 7 Cy5, MDA-MB 231 Cy3 vs MCF 7 Cy3, MDA-MB 231 Cy5	0,74
Amp: MCF 7 Cy5, MDA-MB 231 Cy3 vs MCF 7 Cy5, MDA-MB 231 Cy3	0,96
Gesamt: MCF 7 Cy5, MDA-MB 231 Cy3 vs Amp. MCF 7 Cy5, MDA-MB 231 Cy3	0,85
Gesamt: MCF 7 Cy3, MDA-MB 231 Cy5 vs Amp. MCF 7 Cy3, MDA-MB 231 Cy3	0,81

Tab. 7: Korrelationen zwischen den *dye-swap* Experimenten, sowie zwischen zwei parallelen Amplifikationen und aus dem Vergleich zwischen cDNA aus Gesamt und amplifizierter RNA.

Interessanterweise zeigen die *dye-swap* Experimente (MCF 7 Cy3 , MDA-MB 231 Cy5) in fast allen Fällen etwas schlechtere Korrelationen. (s. Tab. 3,6,7). Ein Grund dafür könnten die verschiedenen Einbauraten der unterschiedlichen Cyanine-dUTPs sein.

Bei der RNA-Amplifikation ist neben der Reproduzierbarkeit der Methode selbst vor allem auch der Vergleich zwischen Array-Experimenten mit Gesamt- und amplifizierter RNA wichtig. Der kritische Reaktionsschritt beginnt bereits bei der reversen Transkription (s. 1. Erststrangsynthese in Abb. 7), wo eine Verfälschung der ursprünglichen mRNA-Population durch mangelhafte Lesefehleraktivität der reversen Transkriptase (Williams and Loeb, 1992) oder der Ausbildung höherer RNA-Strukturen auftreten kann (Strobel and Doudna, 1997). Diese beiden Ursachen führen häufig zu Verzerrungen und damit fehlerbehafteten Ergebnissen.

Exemplarisch ist in Abb. 11 der Vergleich gemeinsam exprimierter Gene zwischen den Hybridisierungsexperimenten von MCF 7 Cy5-markierter und MDA-MB 231 Cy3-markierter cDNA sowie den analog markierten Proben aus der Amplifikation dargestellt. In Abb. 9 sind die Gene anteilig aus den beiden Experimentenserien, in Abb. 10 die Korrelation der Experimente, dargestellt. So sind bei den Hybridisierungen mit Gesamt-RNA von 192 Genen insgesamt 71 Gene detektierbar. Aus den Hybridisierungen mit amplifizierter RNA ergeben sich insgesamt 73 von 192 Genen. Die Gene, welche sich nur bei Gesamt- bzw. amplifizierter RNA detektieren lassen sind nicht identisch. Dafür lassen sich zwei Ursachen finden: a) durch die Amplifikation wird die ursprüngliche Transkriptlänge verkürzt, so dass sich die Detektion von 5' repräsentierenden Sequenzen auf dem Array reduziert (Kenzelmann et al., 2004) und /oder b) dass durch die Amplifikation die Anzahl von niedrig vorkommender Sequenzen in der Zelle erhöht wird und diese damit erst detektierbar werden.

In den untersuchten Experimenten bestätigen sich o.g. Sachverhalte. So kommt es durch die Amplifikation auf der einen Seite zur Reduktion der 5' repräsentierenden Sequenzen, da ca. 80 % dieser Sequenzen, die sich nur bei Gesamt RNA detektieren lassen, näher am 5' Ende der betr. mRNA liegen und somit die verkürzten Transkripte dort nicht mehr hybridisieren können. Auf der anderen Seite wurde der Anteil sog. *low copy templates*, durch die Amplifikation erhöht, so dass diese Sequenzen überhaupt erst detektierbar wurden. Dieses Ergebnis zeigt außerdem, dass durch die Amplifikation lediglich 60 % von den ursprünglich detektierbaren Sequenzen (71 bzw. 73 Gene) erhalten bleibt.

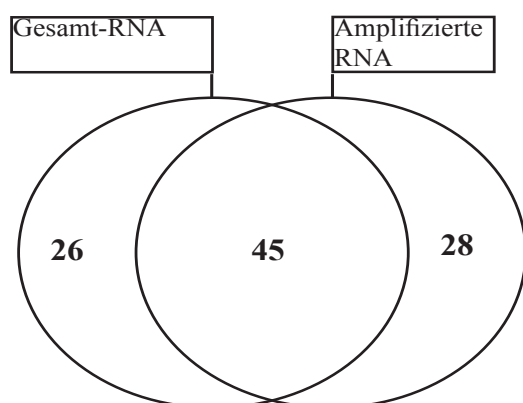


Abb. 9: Darstellung der jeweiligen Anzahl detektierbarer Gene zwischen Gesamt und amplifizierter RNA. Experimentenserie: MCF 7 Cy5, MDA-MB 231 Cy3 .

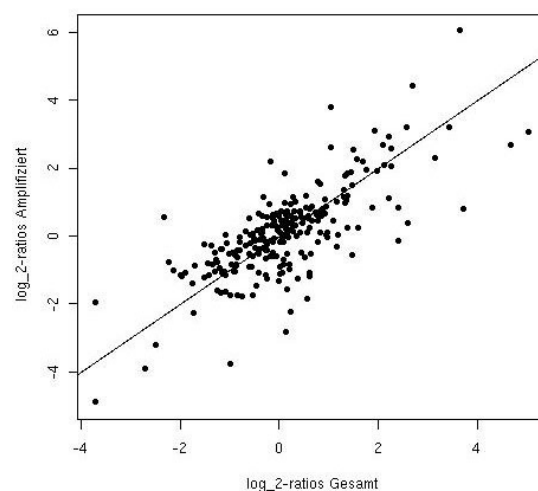


Abb. 10: *Scatter Plot* zwischen den Experimenten mit Gesamt- und amplifizierter RNA, bzgl. der Schnittmenge in Abb. 9.

Für den Vergleich von Expressionsdaten zwischen den Experimenten mit Gesamt- und amplifizierter RNA sind insbesondere die 45 Gene aus der Schnittmenge interessant. Dabei wurden sowohl die Signalintensitäten, wie auch die *log ratios* dieser Gene untersucht (s. Abb. 11).

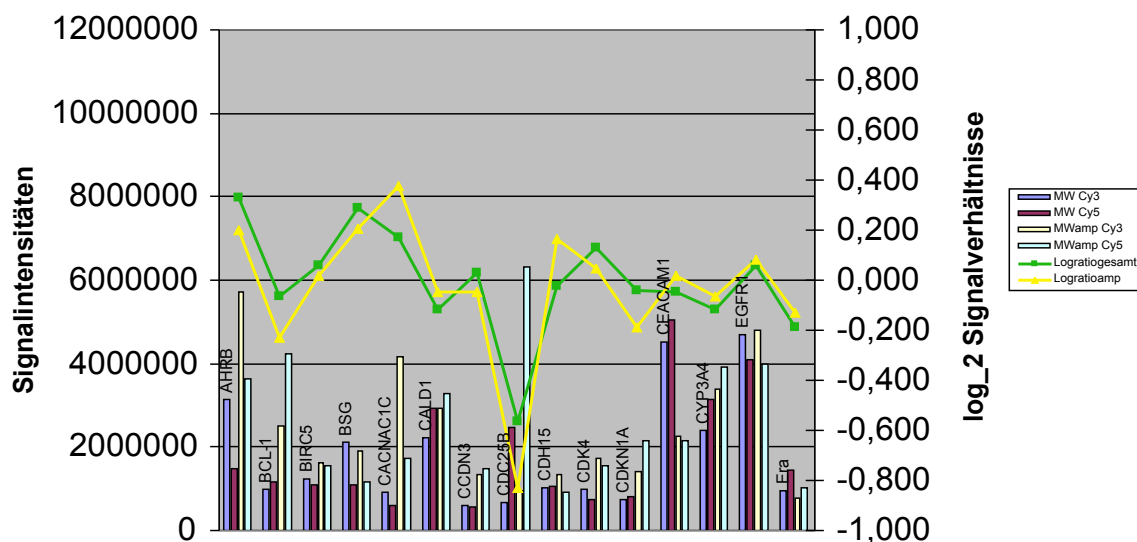


Abb. 11: Auszug (15 von 45 Genen) der Schnittmenge aus Abb. 9 mit Darstellung der Signalintensitäten sowie der logarithmierten Signalverhältnisse (*log<sub>2</sub>-ratios*).

Bei Betrachtung von Abb. 11 fällt auf, dass etwa 50 % der hier erfassten Gene durch die Amplifikation eine Erhöhung der Signalintensitäten aufzeigen. Bei den *log ratios* dagegen zeigen diese Gene nahezu ein gleiches Verhalten. Von den insgesamt 45 gemeinsamen Genen zeigen 3 gegensätzliche Expression, 12 Sequenzen werden differentiell exprimiert (Signalverhältnisse  $< 1/3$  bzw.  $> 3$ ), wohingegen die restlichen 30 Gene keine differentielle Expression aufweisen (Signalverhältnisse  $> 1/3$  bzw.  $< 3$ ).

Für eine erste Bewertung sind jedoch auch die Gene interessant, welche lediglich in einem der beiden Typen von Hybridisierungsexperimenten detektierbar sind. Dabei stellt sich in erster Linie die Frage, inwieweit das Verfahren der linearen RNA-Amplifikation damit einen Vorteil für Microarray Experimente aufzeigt.

Dies ist vor allem unter dem Aspekt zu sehen, dass nach der Amplifikation noch lediglich 60 % der Gene in beiden Experimentenserien (also in der Schnittmenge) detektierbar waren und diese Methode zunächst mit einem „Verlust“ einhergeht.

Allerdings müssen auf der anderen Seite auch die Gene außerhalb der Schnittmenge (linker bzw. rechter Kreis, s. Abb. 9) betrachtet werden.

So sind im Nachfolgenden diejenigen Gene aufgelistet die entweder nur bei Hybridisierungsexperimenten mit amplifizierter oder Gesamt-RNA detektiert werden konnten (s. Tab. 8,9).



Im Vergleich der biologischen Funktion scheinen mit der Amplifikation nahezu die doppelte Anzahl von Genen aus den Kategorien Zellzyklus, Zellproliferation und Differenzierung detektierbar zu sein. Gene aus den übrigen Kategorien, wie Zelladhäsion, Zellbewegung, Tumorsuppressorgene und Sonstige (als Zusammenfassung von Immunantwort, Signaltransduktion etc.) zeigen dagegen ähnliche Zahlen (vgl. Tab. 8 und 9).

Name	GeneID	Beschreibung	Biologische Funktion*
ABCG2 (ABCP)	NM_004827	ATP-binding cassette, sub-family G (WHITE), member 2	Medikamentenresistenz
ACTA1	NM_001100	Actin, alpha 1, skeletal muscle	Zytoskelettbestandteil
ATF3	NM_001674	Activating transcription factor 3	Negativer Zellzyklusregulator
CCNA1 (1169)	NM_003914	Cyclin A1, spez PCR Produkt**	Zellzykluskontrolle
CCNA2	NM_001237	Cyclin A2	Zellzykluskontrolle G1/S u. G2/M
CCNB1	NM_031966	Cyclin B1	Zellzykluskontrolle G2/M
CD9	NM_001769	CD9 antigen (p24)	Zellbeweglichkeit, -adhäsion
CDC25A	NM_001789	Cell division cycle 25A	Induktor mitotischer Kontrolle
CDKN1A (1965)	NM_001389	Cyclin-dependent kinase inhibitor 1A, spez. PCR Produkt**	Zellzykluskontrolle
CDKN2C	NM_001262	Cyclin-dependent kinase inhibitor 2C	Tumorsuppressor
CTSD	NM_001789	Cathepsin D	unab. Prognosemarker bei Brustkrebs
ERA 3199	Bastain	Östrogenrezeptor alpha, spez. PCR Produkt**	Zelluläre Proliferation und Differenzierung
FCRL4 (IGFP2)	AF343661	Fc receptor-like 4	B-Zellentwicklung
FLNA	NM_001456	Filamin A, alpha	Reorganisation Aktinzytoskelett u Zellmigration
FN1 5113	Bastain	Fibronectin 1, spez. PCR Produkt**	Zelladhäsion und -beweglichkeit
FOS	NM_005252	v-fos FBJ murine osteosarcoma viral oncogene homolog	Einfluß a. Zellproliferation u. -differenzierung
ITGA2	NM_002203	Integrin, alpha 2	Zell-Matrix-Adhäsion
KRT8	NM_002273	Keratin 8	Zytoskelettorganisation
MAP2K1	NM_002755	Mitogen-activated protein kinase kinase 1	Signaltransduktion
MKI67	NM_002417	Antigen identified by monoclonal antibody Ki-67	Proliferationsmarker
MYB	NM_005375	v-myb myeloblastosis viral oncogene homolog (avian)	Kontrolle von Proliferation u. Differenzierung
PABPC1	NM_002568	Poly(A) binding protein, cytoplasmic 1	mRNA-Metabolismus
PDGFA	NM_006206	Platelet-derived growth factor receptor, alpha polypeptide	Wundheilung und Zellzykluskontrolle

Fortsetzung Tabelle 8:

<b>Name</b>	<b>GeneID</b>	<b>Beschreibung</b>	<b>Biologische Funktion*</b>
PRG2 (BMPG)	NM_024888	Proteoglycan 2, bone marrow (natural killer cell activator, eosinophil granule major basic protein)	Immunantwort
PTPRF	NM_002840	Protein tyrosine phosphatase, receptor type, F	Zelladhäsion
SOX4	NM_003107	SRY (sex determining region Y)-box 4	Transkriptionsaktivator für Tumorgenese /Zelltod
TFRC (4740)	NM_003274	Transferrinrezeptor, spez. PCR Produkt**	Eisentransport und -gleichgewicht
TIMP1	NM_003254	Tissue inhibitor of metalloproteinase 1 (erythroid potentiating activity, collagenase inhibitor)	Positiver Zellzyklusregulator

Tab. 8: Auflistung der durch Amplifikation detektierbaren Gene in alphabetischer Reihenfolge. Die Tabelle enthält neben dem abgekürzten Gennamen, die Referenz-mRNA Nr., den vollen Gennamen, sowie die biologische Funktion. \* Bei der biologischen Funktion wurde lediglich eine, von oftmals mehreren Funktionen genannt. Unter <http://genome-www5.stanford.edu/cgi-bin/source/sourceResult> sind neben der biologischen Funktion, auch die molekularen und zellulären Funktionen aufgeführt. Die mit \*\* gekennzeichneten spezifischen PCR-Produkte wurden in der Diplomarbeit von S. Pfeiffer hergestellt und verwendet, um die Datenqualität des dort eingesetzten DNA-Microarrays mit anderen Methoden vergleichen zu können.

Name	GeneID	Beschreibung	Biologische Funktion*
ACTB	NM_001101	Actin, beta	Zellbeweglichkeit
ACTB	50 Oligo	Actin, beta	Zellbeweglichkeit
BARD1	NM_000465	BRCA1 associated RING domain 1	Tumorsuppressor
CEACAM1	NM_001712	Carcinoembryonic antigen-related cell adhesion molecule 1 (biliary glycoprotein)	Immunantwort
DUSP1	NM_004417	Dual specificity phosphatase 1	Zellzyklus
ELF3	NM_004433	E74-like factor 3	Transkriptionsregulator
ERA	70 Oligo	Östrogenrezeptor alpha	Zelluläre Proliferation und Differenzierung
GAPDH	NM_002046	Diglycerinaldehyde-3-phosphate dehydrogenase	Glukosemetabolismus
JUNB	NM_002229	jun B proto-oncogene	Transkription
KLF10 (TIEG)	NM_005655	Kruppel-like factor 10	Zellwachstum
KRT18	NM_000224	Keratin 18	Morphogenese
LTBP3	NM_021070	Latent transforming growth factor beta binding protein 3	
LTBP4	NM_003573	Latent transforming growth factor beta binding protein 4	Regulation v. Zellwachstum u. -differenzierung
MDM2	NM_002392	Mdm2, transformed 3T3 cell double minute 2, p53 binding protein (mouse)	Zellzyklus
MIF	NM_002415	Macrophage migration inhibitory factor (glycosylation-inhibiting factor)	Immunantwort
MTA1	NM_004689	Metastasis associated 1	Signaltransduktion
NME1	NM_000269	Non-metastatic cells 1, protein (NM23A) expressed in	Negativer Zellzyklusregulator
PSMC5	NM_002805	Proteasome (prosome, macropain) 26S subunit, ATPase, 5	Proteinkatabolismus
RB1	NM_000321	Retinoblastoma 1	Tumorsuppressor
RPS19	NM_001022	Ribosomal protein S19	Proteinbiosynthese
SMAD3	NM_005902	SMAD, mothers against DPP homolog 3 (Drosophila)	Transkription
TGFR1	NM_004612	Transforming growth factor, beta receptor I	Signaltransduktion
TGM2	NM_004613	Transglutaminase 2	positive Regulation der Zelladhäsion
TUBB	NM_178014	Tubulin, beta polypeptide	
TXN	NM_003329	Thioredoxin	Zellbeweglichkeit und -proliferation
VCAM1	NM_001078	Vascular cell adhesion molecule 1	Zelladhäsion

Tab. 9: Auflistung der detektierbaren Gene ohne vorausgegangene Amplifikation in alphabetischer Reihenfolge. Die Tabelle enthält neben dem abgekürzten Gennamen, die Referenz-mRNA Nr., den vollen Gennamen, sowie die biologische Funktion. \* Bei der biologischen Funktion wurde lediglich eine, von oftmals mehreren Funktionen genannt. Unter <http://genome-www5.stanford.edu/cgi-bin/source/sourceResult> sind neben der biologischen Funktion, auch die molekularen und zellulären Funktionen aufgeführt. Die mit \*\* gekennzeichneten Oligomere wurden in der Diplomarbeit von S. Taifel verwendet, um die Datenqualität des dort eingesetzten DNA-Microarrays mit anderen Methoden vergleichen zu können.

Somit ermöglicht die RNA-Amplifikation hier eine weitere Detektion von Genen aus Zellzyklus, Zellproliferation und Differenzierung, deren Regulation häufig bei malignen Erkrankungen gestört ist und deren Transkripte offensichtlich in geringerer Kopienzahl vorkommen. Diese Ergebnisse liefern erste Hinweise dafür, dass die lineare RNA-Amplifikation eine Möglichkeit zur partiellen Sensitivitätssteigerung bei geringen Ausgangsmengen von Gesamt-RNA bietet. Allerdings sind zur Überprüfung und Validierung dieser Methode eine größere Anzahl Experimente sowie eine weitere RNA-Amplifikationsmethode notwendig. Eine andere Möglichkeit die Sensitivität zu verbessern, liegt in der Auswahl anderer Markierungsmethoden, die nachfolgend beschrieben werden.

### 3.4 Markierungs-Strategien für Microarray Experimente

Grundsätzlich bieten sich bei Microarray Experimenten verschiedene Markierungsmöglichkeiten an. So gibt es neben der direkten Markierung mit Cyanine 5-/3-dUTP bzw. Biotin-/Digoxigenin-dUTP, die radioaktive Markierung mit  $^3\text{H}$ -dXTP bzw.  $^{32}\text{P}$ -/ $^{33}\text{P}$ -dCTP, oder die indirekte Markierung mit Aminoallyl-dUTP. Alle markierten cDNAs, mit Ausnahme der radioaktiv markierten Proben, können sowohl einzeln oder kompetitiv hybridisiert werden. Im weiteren Teil dieser Arbeit wurden mit dem neu konzipierten DNA-Microarray Hybridisierungsexperimente mit Biotin-/Digoxigenin-dUTP-markierten Proben durchgeführt.

#### 3.4.1 Biotin-dUTP Monocolor-Markierung

Für die bisher durchgeführten Microarray Experimente war der Bedarf an Gesamt-RNA mit jeweils 50  $\mu\text{g}$  / Hybridisierung recht hoch. Mit der linearen RNA-Amplifikation wurde eine Möglichkeit aufgezeigt, auch mit geringen RNA-Ausgangsmengen (1  $\mu\text{g}$  amplifizierte RNA) Hybridisierungsexperimente durchzuführen.

Eine andere Möglichkeit zur Sensitivitätssteigerung liegt in abgewandelten Markierungsverfahren, beispielsweise mit Biotin-dUTP (s. Kap. 2.2.7.3) und nachfolgender Detektion mit einem Cyanin-3 Anti-Biotin konjugierten Antikörper (s. Kap. 2.2.13). Dabei wurde lediglich 1/10 der ursprünglichen Gesamt-RNA-Ausgangsmenge, also 5  $\mu\text{g}$ , zur cDNA-Synthese eingesetzt.

Diese Versuchsreihe umfasste insgesamt 8 Einzelhybridisierungen mit den o.g. Zelllinien (4 x MCF 7, 4 x MDA-MB 231). Im Gegensatz zu den direkt mit Cyanine-dUTP markierten Proben (s. Kap. 3.2.1) ergab die statistische Datenanalyse keine signifikant differentiell exprimierten Gene. Dies begründet sich durch die größeren Varianzen der Expressionsdaten beim indirekten Probenvergleich im Gegensatz zum direkten Probenvergleich. Als Konsequenz entsteht eine Verzerrung, auch als Interchipvarianz bezeichnet, die zu weniger deutlich ausgeprägten Expressionsunterschieden führt, welche wiederum die Qualitätskriterien verfehlen (s. Kap. 2.2.15) und damit nicht mehr signifikant sind. Hinzu kommt ein kleiner Stichprobenumfang, der ebenfalls eine gewisse Verzerrung erzeugt.

Durch Erhöhung des Stichprobenumfangs bei den Einzelhybridisierungen könnte zumindest teilweise die Signifikanz der Expressionsdaten verbessert werden, woraus sich wiederum differentielle Genexpression ergibt (Persönliche Kommunikation m. Herrn Winter, MIA Universität Stuttgart). Diese Vorgehensweise wurde zur Validierung des Microarrays in Kap. 3.4.2 erfolgreich angewandt.

### 3.4.2 Etablierung und Validierung des DNA-Microarrays

Von Seiten des klinischen Kooperationspartners (RBK) erfolgte mit einer anderen Microarray Plattform (Oligo-Array der Fa. Eppendorf) eine parallele Datenerhebung an einem bereits etablierten Zellmodellsystem. Als Modellsystem wurde hier die humane Brustkrebszelllinie MCF 7 verwendet und Expressionsprofile unter Antiöstrogenstimulation sowie im unstimulierten Zustand, erstellt. Diese Arraydaten wurden mit unabhängigen Methoden, wie der quantitativen RT-PCR (ebenfalls RBK) und dem *RNAse protection assays* (IZI) überprüft. Aufgrund dieser Daten und zur besseren Vergleichbarkeit zwischen den beiden Arrayplattformen wurden im Anschluß mit dem IGB-Chip auch Einzelhybridisierungsexperimente mit biotinylierter cDNA aus dem o.g. Modellsystem durchgeführt. Die Stimulation mit dem Antiöstrogen ICI 182780 (10 nM, 120 h; Kontrolle 0,1% EtOH) erfolgte zu drei voneinander unabhängigen Zeitpunkten. Nach der Stimulation wurde die RNA aus den so behandelten Zellen von den klinischen Kooperationspartnern bereitgestellt. Die Array-Experimente mit dem IGB-Chip wurden mit 3 voneinander unabhängigen RNA-Probenpaaren im Triplikat durchgeführt ( $n = 9$ ). Die statistische Auswertung ergab 33 differentiell exprimierte Gene ( $p \leq 0,05$ , s. Tab. 10 und Abb. 12).

Der hier vom RBK verwendete DNA-Microarray der Fa. Eppendorf (im weiteren Eppendorf-Chip genannt) ist eine oligo-basierte Arrayplattform mit 222 brustkrebspezifischen Genen, gehört damit zu den *low density* Arrays und ist lediglich ein Beispiel aus der themenspezifischen Microarrayproduktpalette. Dieser Array deckt die wichtigsten Signalwege wie Angiogenese (KDR), Apoptose (BCL2, BAK1, BAG1), Zellzyklus (CDKN1, p53, CCNE1) und -differenzierung (KRT 5, 8, 17-19), DNA-Reparatur (TOPO2A,) Onkogenese (MDM2, c-myc), Transkription (ESR, RB1) etc. mit ihren charakteristischen Vertretern ab (Annotation der Gennamen siehe Anhang).

Der Array wird auch als Dualchip<sup>TM</sup> bezeichnet, weil dieser zwei voneinander getrennte Hybridisierungskammern mit jeweils identischen Arrays auf einem Chip besitzt und damit eine leichte Handhabung ermöglicht. Lt. Eppendorf ist dieser Array sowohl für 1-Farb- wie auch für 2-Farbexperimente geeignet. Die Arrayplattform wird als sog. „*one hand solution*“ mit allen zugehörigen Komponenten (z. B. Reagenzien, Hybridisierungskammer und Auswertungssoftware; weitere Informationen unter <http://www.eppendorf.com/microarrays/>) angeboten.

Name	ID	adj.p-Werte (BH)	t	MW-ICI	MW-unstim	Varianz-ICI	Varianz-unstim
TRA1	NM_003299	0.0044	-3.5762	19.8338	20.7726	0.1560	0.0731
EIF2S1	NM_004094	0.0044	-3.5762	18.2531	19.0287	0.0798	0.0575
NUSAP1	NM_016359	0.0044	-3.5762	17.2975	18.8159	0.7178	0.1501
BCAR3	NM_003567	0.0044	-3.5762	22.5044	22.8347	0.1258	0.0027
CDKN1A	NM_000389	0.0044	3.5762	19.2696	18.2231	0.0454	0.1011
CLDN1	NM_021101	0.0044	-3.5762	21.5590	22.6190	0.1212	0.0229
PTTG1	NM_004219	0.0044	-3.5762	16.9603	18.5509	0.5375	0.0808
DIAPH3	NM_030932	0.0069	3.4879	19.1422	18.5630	0.0475	0.0563
MIF	NM_002415	0.0069	-3.4879	19.6962	20.7508	0.1653	0.1461
NSEP1	NM_004559	0.0088	-3.3996	19.6468	21.0256	0.2691	0.2111
CSE1L	NM_001316	0.0088	-3.3996	18.2154	19.3430	0.1634	0.0965
K-ALPHA-1	NM_006082	0.0088	-3.3996	20.5826	21.9661	0.2468	0.2526
MMP14	NM_004995	0.0088	3.3996	19.0064	18.0897	0.0570	0.2779
DUSP1	NM_004417	0.0088	3.3996	18.9796	17.6675	0.0868	0.4004
ANXA2	NM_004039	0.0114	3.3113	19.0745	17.8175	0.2842	0.2024
JUP	NM_002230	0.0114	3.3113	19.6172	18.9652	0.1057	0.0619
MKI67	NM_002416	0.0114	-3.3113	17.1199	18.2966	0.3786	0.1801
FOS	NM_005252	0.0114	3.3113	18.8465	17.5515	0.6368	0.0786
PRC1	NM_003981	0.0114	-3.3113	17.4180	18.9302	0.4614	0.4547
SLC7A5	NM_003486	0.0169	-3.2230	18.0534	19.2987	0.4571	0.1541
TFF1	NM_003225	0.0169	-3.2230	19.4377	20.0849	0.0746	0.0572
RPS6KB1	NM_003161	0.0169	-3.2230	18.9892	19.6552	0.0957	0.1134
CTGF	NM_001901	0.0226	3.1347	18.4167	17.7148	0.1411	0.0965
E2-EPF	NM_014501	0.0226	-3.1347	18.4760	19.5054	0.2700	0.2096
DHPS	NM_001930	0.0226	-3.1347	18.2874	18.9292	0.1249	0.0643
COL4A1	NM_001845	0.0226	-3.1347	19.5647	21.0156	0.3872	0.3045
ITGAV	NM_002210	0.0343	3.0464	17.8877	17.2828	0.1956	0.0856
LTBP1	NM_000627	0.0421	2.9581	18.7052	18.0036	0.1813	0.0796
ITGB3	NM_000212	0.0421	-2.9581	17.9816	18.5012	0.0785	0.0653
MCM6	NM_005915	0.0421	-2.9581	18.3250	18.9686	0.1160	0.1433
MAGEA1	NM_004988	0.0421	2.9581	18.3741	17.4767	0.5138	0.1680
DC13	NM_020188	0.0421	-2.9581	18.2163	18.8836	0.1491	0.0492
BUB1	NM_004336	0.0421	-2.9581	16.8180	17.6431	0.2826	0.1658

Tab. 10: Differentielle Gene zwischen mit ICI187780 und unstimulierten MCF 7 Zellen. adj. p-Wert (BH) = adjustierter p-Wert nach Benjamini-Hochberg t = t-Teststatistik. Erklärung der Spalten siehe Kap.3.2.1

Bei der Betrachtung der adjustierten p-Werte sowie der zugehörigen Werte der Teststatistiken fällt auf, dass die Werte bei mehreren Genen, ungeachtet der verschiedenen Mittelwertdifferenzen, identisch sind. Im Gegensatz zu den Auswertungen hinsichtlich differentiell exprimierter Gene zwischen ÖR<sup>+</sup>- und ÖR<sup>-</sup>-Tumoren, sowie zwischen Tumor- und Normalgewebe wurde hier ein sog. Wilcoxon-Vorzeichen-Rangsummentest durchgeführt. Dazu wurden für jedes Gen die Differenzen der Werte der 3 RNA-Probenpaare gebildet (jeweils ICI-stimuliert - nicht stimuliert), die Beträge dieser Differenzen ihrer Größe nach sortiert und jedem dieser Differenzen der Wert (Rang) zugeordnet, der beschreibt an welcher Stelle sie in dieser sortierten Reihenfolge auftreten (also 1 für die kleinste, 2 für die mittlere und 3 für die größte Differenz).

Die Wilcoxon-Teststatistik wird nun allein aufgrund dieser Werte (Ränge) und der Vorzeichen der Differenzen berechnet. Aufgrund der Tatsache, dass hier nur drei unabhängige Probenpaare untersucht wurden, kommt es häufiger vor, dass für mehrere Gene die Reihenfolge der Differenzen und deren Vorzeichen übereinstimmen. Für diese Gene ergeben sich dadurch die selben Werte für die Teststatistiken und damit auch identische p-Werte (wie dies z.B. für die ersten sieben Gene in Tab. 10 der Fall ist).

Zur besseren Veranschaulichung der Ergebnisse aus Tabelle 10 wurden die logarithmierten Mittelwerte der Probenpaare in die Mittelwerte relativer Signalintensitäten zurückgerechnet und in Abb. 12 veranschaulicht.

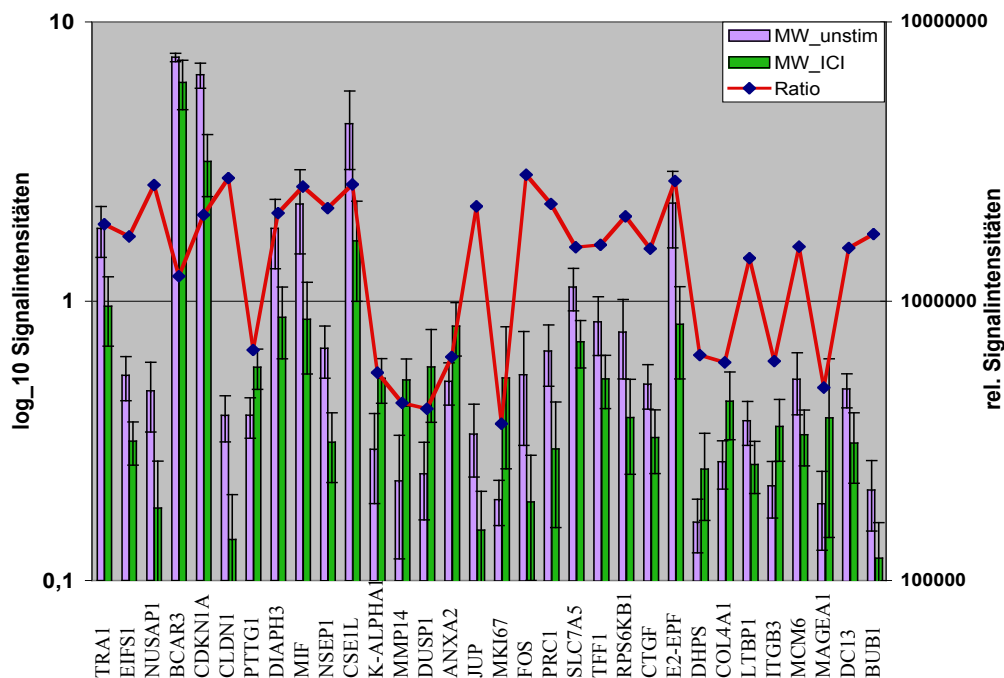


Abb. 12: Differenziell exprimierte Gene zwischen unstimulierten und mit ICI187780 stimulierten MCF 7 Zellen; MW = unlogarithmierter Mittelwert relativer Signalintensitäten sowie logarithmierter Signalverhältnisse ( $\log_{10}$  ratios)

In obigem Genset befinden sich auch 2 Gene, die ebenfalls im Genkollektiv des Eppendorf-Chips vertreten sind. CDKN1A ( $p21^{CIP}$ ), ein cyclin-abhängiger Kinase Inhibitor, welcher auch bei der Onkogenese von Brustkrebs eine Rolle spielt (Musgrove et al., 1995), wird unter Antiöstrogen-Stimulation reprimiert. Desweiteren steht CDKN1A unter direktem Einfluss von p53 und besitzt damit Funktionen hinsichtlich der Kontrolle von Zellzyklus und Apoptose (Wagner, 1999). TFF1 (pS2) das andere Gen, ist östrogen-induzierbar und begünstigt Patientinnen hinsichtlich der Anwendung einer Antihormontherapie (Gillesby and Zacharewski, 1999). Unter Antiöstrogenstimulation wurde dieses Gen ebenfalls reprimiert. Das Expressionsverhalten dieser beiden Gene folgt tendenziell dem des Eppendorf-Chips. Abb. 13 zeigt die  $\log$  ratios weiterer, gemeinsam exprimierter Gene der beiden Chip-Systeme. Obwohl diese  $\log$  ratios in ihrem Wert teilweise sehr große Unterschiede weisen, zeigen IGB- und Eppendorf-Chip in die gleiche Richtung.

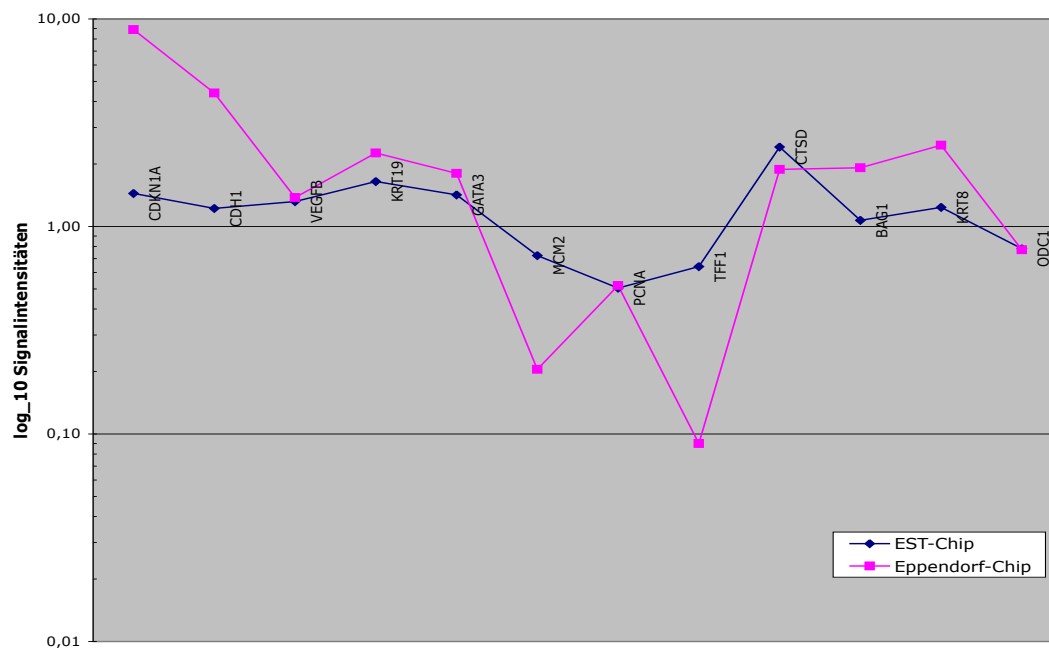


Abb. 13: Vergleich der *log-ratios* zwischen IGB- und Eppendorf-Array. Annotation der Gennamen siehe Genliste im Anhang; logarithmierte Signalverhältnisse (*log<sub>10</sub> ratios*)

Obwohl hier nur ein sehr kleines Genkollektiv verglichen wurde, beträgt die Übereinstimmung bzgl. der Expressionstendenz zwischen den beiden Chipsystemen ca. 40 % (5 von 12 Genen). Anhand eines ausgesuchten Genkollektivs erfolgte zwischen dem IGB-Array und dem *RNase protection assay* (RPA) ein weiterer Methodenvergleich. Dabei sind die Ergebnisse des RPA nicht quantitativ dargestellt, sondern zeigen eine Expressionstendenz an (s. Tab. 11).

Genname	Expression EST-Chip	Expression RPA
CDKN1A	1,06	=
CASP8	1,44	+
JUN	1,81	+
FOS	2,74	+
TGFB3	2,21	=
TGFB2	1,75	+

Tab. 11: Expressionsvergleich zwischen IGB-Array und RPA. = unverändert, + = nach oben reguliert, - = nach unten reguliert. Annotation der Gennamen siehe Genliste im Anhang.

Der IGB-Array folgt der Tendenz des RPA mit einer Ausnahme. Im Falle von TGFB3 zeigt der IGB-Array eine Regulation nach oben, wohingegen der RPA keine Veränderung zeigt. In einem weiteren Methodenvergleich sollte nun der IGB-Array gegen die quantitative *real time*-PCR getestet werden. Anhand bereits vorhandener Daten sowie der Expertise des klinischen Kooperationspartners mit auto- und parakrinen Signalprozessen wurde an ausgesuchten Komponenten des TGF-Beta-Systems die durch Antiöstrogen-Stimulation hervorgerufenen Expressionsänderungen untersucht (Abb. 14, 15).



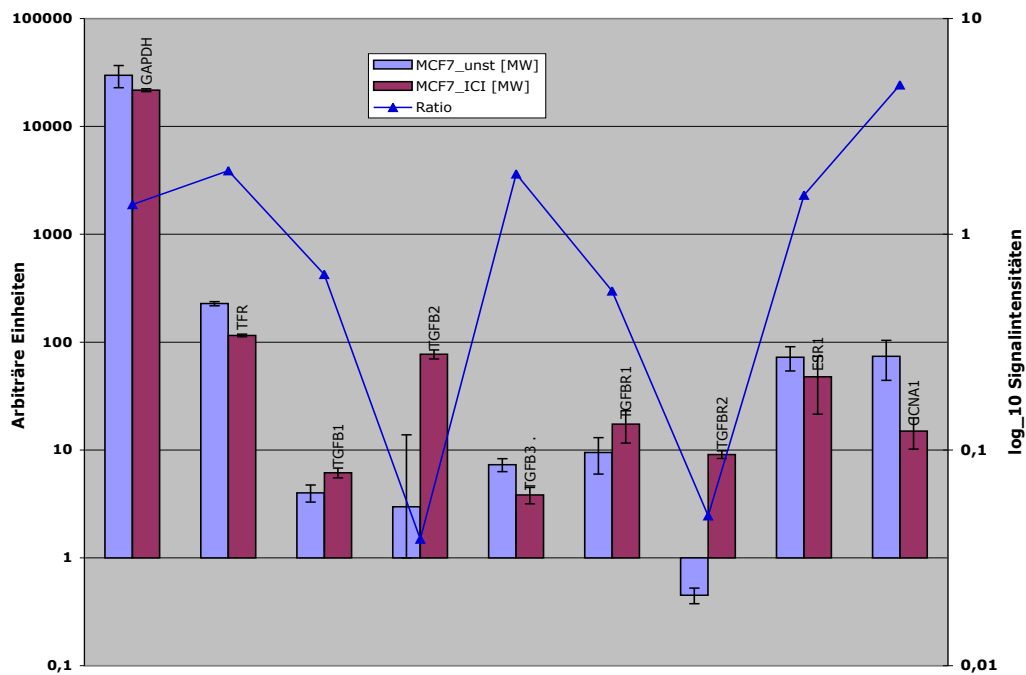


Abb. 14: Signalintensitäten und logarithmierte Signalverhältnisse ( $\log_{10}$  ratios) von *real-time* PCR nach Antiöstrogen-Stimulation (n = 3). Annotation der Gennamen siehe Genliste im Anhang.

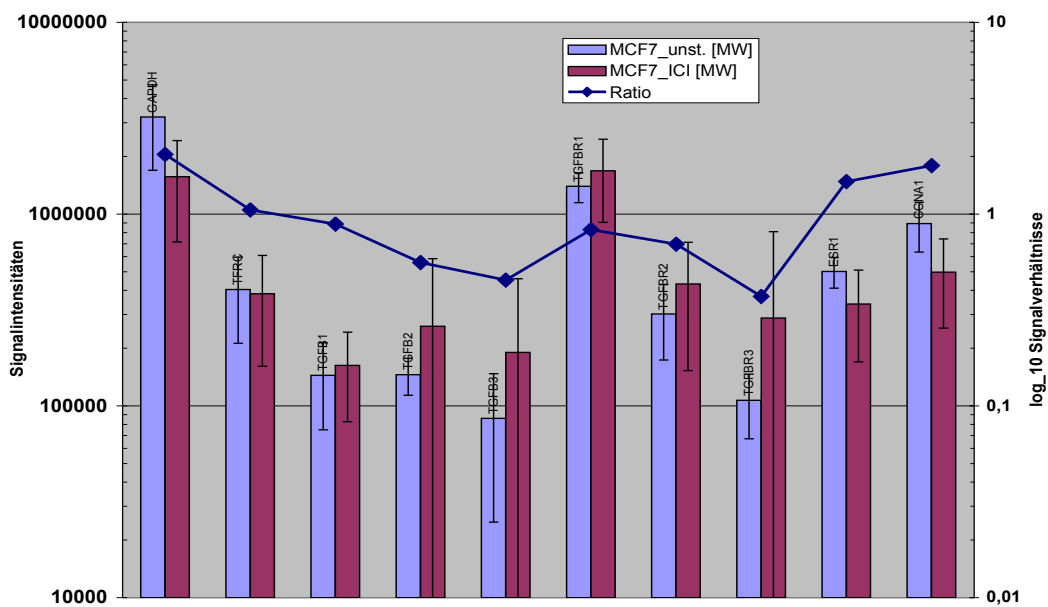


Abb. 15: Signalintensitäten und logarithmierte Signalverhältnisse ( $\log_{10}$  ratios) des IGB-Chip nach Antiöstrogen-Stimulation (n = 9). Annotation der Gennamen siehe Genliste im Anhang.

Analog den vorausgegangenen Methodenvergleichen ist beim Betrachten der beiden Diagramme zu erkennen, dass der IGB-Array hinsichtlich der Expressionsverhältnisse tendenziell denen der quantitativen *real-time* PCR folgt. Allerdings sind, wie im Vergleich zum Eppendorf-Array, auch hier die Unterschiede nicht so deutlich ausgeprägt.

Zusammenfassend kann gesagt werden, dass sich die Arrayergebnisse mittels dieser dreier unabhängigen Methoden recht gut überprüfen lassen, da die Übereinstimmung der Expressionstendenz bei den einzelnen Methoden zwischen 40 und 80 % liegt.

Die Korrelationen der Experimente beim IGB-Array liegen für die unstimulierten Proben ( $n = 9$ ) zwischen 0,74 und 0,94. Im Fall der Antiöstrogen-Stimulation liegt die Korrelation im Bereich von 0,7 - 0,9 ( $n = 5$ ). Damit ist die biologische Reproduzierbarkeit Proben unabhängig. Allerdings treten bei der Antiöstrogen-Stimulation Schwankungen auf, die wohl auf technische Gründe zurückzuführen sind.

Durch diese veränderte cDNA-Markierungsstrategie konnte hier eine verbesserte Sensitivität bzgl. einer deutlich reduzierten Ausgangsmenge an Gesamt-RNA (5 statt 50  $\mu\text{g}$ ) erzielt werden.

### 3.4.3 Biotin-/Digoxigenin-dUTP Dualcolor-Markierung

Neben den in Kap. 3.4.2 durchgeführten 1-Farbexperimenten lassen sich mit der kompetitiven Hybridisierung (2-Farbexperimente) bei cDNA-Chips meist stabilere Daten generieren. Einer der Vorteile bei den 2-Farbexperimenten ist dabei der direkte Vergleich zweier Proben auf einem Chip, wodurch sich Verzerrungen, wie die Interchipvarianz, besser kontrollieren lassen und somit kleinere Varianzen entstehen. Beim Design von 2-Farbexperimenten sollte auch stets ein Kontrollsystem mit einbezogen werden, wodurch die Vergleichbarkeit von Expressionsdaten zwischen mehreren Chips erleichtert wird. Ein Kontrollsystem kann entweder mit sog. Haushaltsgenen als interne oder externe Kontrollen wie *spiked*-RNAs oder mittels einer Referenzprobe (in dieser Arbeit RNA-Pool verschiedener Zelllinien) realisiert werden.

Um die indirekte Markierung auch als 2-Farbexperiment durchführen zu können, wurde zunächst nach einem weiteren modifizierten Nukleotid gesucht. Aufgrund verwandter Molekülstruktur zu Biotin-dUTP fiel die Wahl dabei auf Digoxigenin-dUTP.

Im nächsten Schritt wurde mittels RT die cDNA-Synthese überprüft, ob der Einbau von Digoxigenin-dUTP überhaupt erfolgt ist. Nach photometrischer Bestimmung der Proben 1, 2 und 3 wurden äquivalente Mengen cDNA in einem Sybrgold-gefärbten, 1 % Agarosegel elektrophoretisch aufgetrennt. Gemäß Abb. 16 haben die cDNA-Synthesen in allen drei Fällen stattgefunden. Im *Dot Blot* sollte der Einbau der modifizierten Nukleotide aus dem Reaktionsansatz von Abb. 16 überprüft werden. Der *Dot Blot* zeigt, dass die modifizierten cDNAs detektierbar sind und damit der Einbau von Biotin- sowie Digoxigenin-dUTP stattgefunden hat.

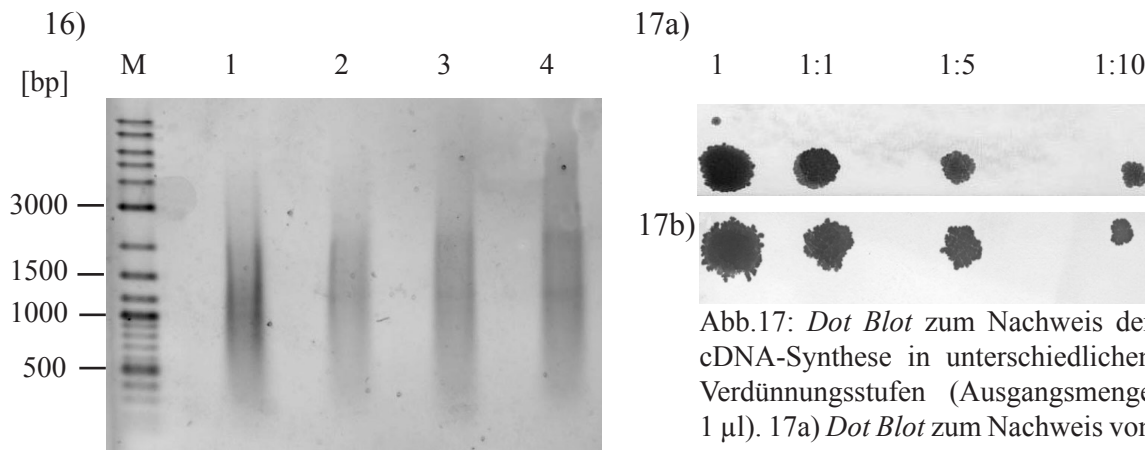


Abb. 16: cDNA-Synthese Überprüfung im 1 % Agarosegel; M. 2\_log Ladder, 1. cDNA ohne modifizierte Nukleotide. 2. biotin-modifizierte cDNA. 3. digoxigenin-modifizierte cDNA. 4. 2. und 3. im Verhältnis 1:1

Abb.17: *Dot Blot* zum Nachweis der cDNA-Synthese in unterschiedlichen Verdünnungsstufen (Ausgangsmenge 1  $\mu$ l). 17a) *Dot Blot* zum Nachweis von biotin-modifizierter cDNA ( $\sim 0,15 \mu\text{g}/\mu\text{l}$ ) detektiert mit Streptavidin im alkalischen Phosphatase Assay; 17b) *Dot blot* zum Nachweis von digoxigenin-modifizierter cDNA ( $\sim 0,12 \mu\text{g}/\mu\text{l}$ ), detektiert mit Anti-Dig-oxigenin im alkalischen Phosphatase Assay.

Die cDNA-Synthese mit unmodifizierten Nukleotiden erscheint auf dem Gel trotz Verwendung gleicher Ausgangsmenge etwas stärker (vgl. 1. mit 2. und 3.). Dies ist möglicherweise auf die unterschiedlich modifizierten Nukleotide in den jeweiligen cDNAs zurückzuführen.

Zur Detektion der modifizierten cDNAs wurden unterschiedliche Fluorophore verwendet. Da gegenwärtig kein kommerziell erhältlicher Cyanine 5 konjugierter Antikörper für Digoxigenin verfügbar war, wurde folgende Alternative gewählt: Die Detektion von biotin-modifizierter cDNA erfolgte nun mit Alexa Fluor 647 (Abs. 650 nm, Em. 668 nm als Analog für Cyanine 5), digoxigenin-modifizierte-cDNA wurde über Anti-Digoxigenin-Rhodamin Fab-Fragmente (Abs. 558 nm, Em. 580 nm als Analog für Cyanine 3) detektiert.

Das hier benutzte Detektionssystem konnte allerdings nur eingeschränkt verwendet werden. Einzelhybridisierungen mit den modifizierten cDNAs und die Detektion mit den o.g. Fluorophoren zeigten geringe Signal-/Hintergrundverhältnisse, d.h. Signale waren nicht eindeutig vom Hintergrund zu unterscheiden. Des Weiteren zeigte das Detektionssystem auch insgesamt schwächere Signalintensitäten als die in Kap. 3.4.2 beschriebene Monocolor-Markierung.

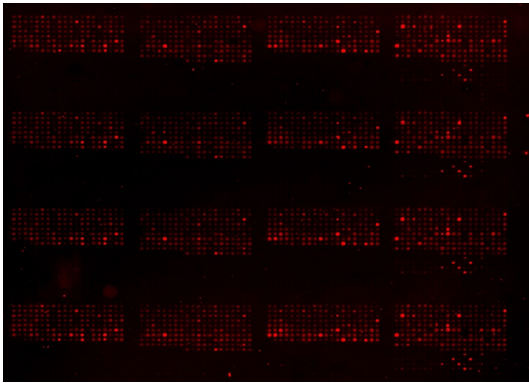
Ferner zeigten beide Fluorophore auch unspezifische Bindungen über den gesamten Array hinweg (Abb. 18 a,b). Dieser Sachverhalt zeigte sich auch bei der kompetitiven Hybridisierung, wobei hier lediglich die Anti-Digoxigenin-Rhodamin Fab-Fragmente zur unspezifischen Bindung neigten. Aufgrund dieser Nachteile wurde das ausgesuchte System nicht weiter verwendet.

Bei den Einzelhybridisierungen (s. Kap. 3.4.2) hat sich der verwendete Cyanine 3 konjugierte Biotin-Antikörper als zuverlässig erwiesen. Als weiteres Detektionsmolekül wurde deshalb ein mit Cyanine 5 gekoppelter Digoxigenin-Antikörper eingesetzt (s. Kap. 2.2.8). Mit einem ausgewählten Tumorkollektiv von 4 Brusttumoren sowie einer Referenzprobe, bestehend aus einem Zellpool von 3 humanen Brustkrebszelllinien (MCF7, MDA-MB 231 und MDA-MB 435 im Verhältnis 1:1:1), erfolgten mit diesem neuen Detektionssystem sowohl Einzel- wie auch kompetitive Hybridisierungen (Tab. 12,13). Des Weiteren wurden diese Experimente auch im *dye-swap* ausgeführt. Das Genkollektiv dieser Arrayserie wurde im Duplikat (2 x 778 Sonden) gedruckt. Die Arrays wurden mit zwei Belichtungszeiten ausgelesen (Mono-Cy5: 1,0 s u. 2,3 s; Mono-Cy3: 1,4 s u. 2,3 s).

Hybridisierung	Fluorophor	Gewebe* /Probe	Spotanteil (%)**
Mono	Cyanin 5	2227/02	70,19
Mono	Cyanin 5	2611/02	54,26
Mono	Cyanin 5	4160/02	62,64
Mono	Cyanin 5	11341/02	65,11
Mono	Cyanin 3	2227/02	36,13
Mono	Cyanin 3	2611/02	22,53
Mono	Cyanin 3	4160/02	59,75
Mono	Cyanin 3	11341/02	31,59

Tab. 12: Monohybridisierungen an ausgesuchtem Gewebekollektiv. \* Journal Nr. aus Gewebekollektiv des RBK. \*\* Alle Spots nach Zusammenfassen von Belichtungszeiten und Duplikaten nach Erfüllung der Qualitätskriterien wie Nettosignal  $\geq 1,5 \cdot$  Hintergrundsignal; Nettosignale  $> 95 \%$  Quantil der Negativkontrollen, keine Signalsättigung.

18a)



18b)

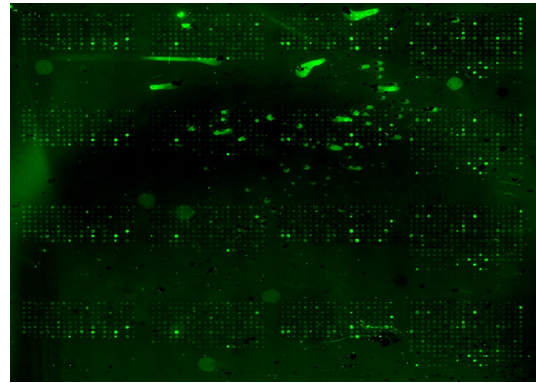


Abb. 18: Einzelhybridisierungsbilder in Falschfarbendarstellung. 18a) Alexa Fluor 647 und 18b) Antidigoxigenin-Rhodamin Fab-Fragmente. Bilder zeigen in beiden Fällen unspezifische Bindungen über den ganzen Array.

Es ist auffällig, dass bei den Einzelhybridisierungen mit Cyanine 5 detektiert der Spot-Anteil rund doppelt so hoch ist wie bei der Detektion mit Cyanine 3. Eine Ausnahme ist das *dye-swap* Experiment mit der Probe 4160/02.

Hybridisierung	Fluorophor	Gewebe*/Probe	Spot-Anteil (%)**
Kompetitiv	Cyanin 5	2227/02	90,25
Kompetitiv	Cyanin 3	Referenz	45,88
Kompetitiv	Cyanin 5	2611/02	79,53
Kompetitiv	Cyanin 3	Referenz	53,30
Kompetitiv	Cyanin 5	4160/02	77,20
Kompetitiv	Cyanin 3	Referenz	46,43
Kompetitiv	Cyanin 5	11341/02	59,34
Kompetitiv	Cyanin 3	Referenz	45,19
Kompetitiv	Cyanin 5	Referenz	54,40
Kompetitiv	Cyanin 3	2227/02	28,16
Kompetitiv	Cyanin 5	Referenz	78,57
Kompetitiv	Cyanin 3	2611/02	9,75
Kompetitiv	Cyanin 5	Referenz	71,57
Kompetitiv	Cyanin 3	4160/02	48,49
Kompetitiv	Cyanin 5	Referenz	71,43
Kompetitiv	Cyanin 3	11341/02	41,48

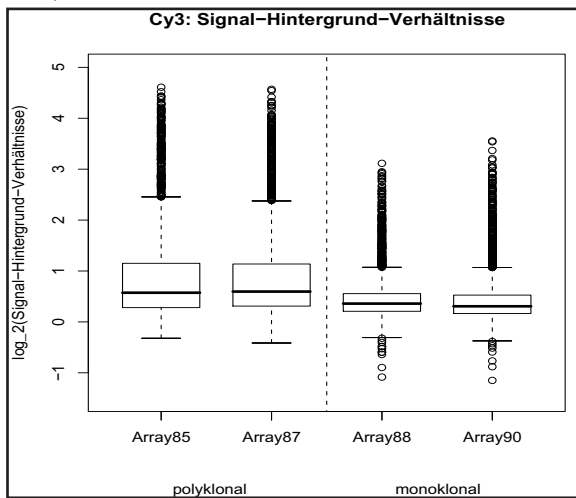
Tab. 13: Kompetitive Hybridisierungen mit ausgesuchtem Gewebekollektiv und Zellpool als Referenz. \* Journal Nr. aus Gewebekollektiv des RBK. Beschreibung analog Tab. 12.

Auch bei den kompetitiven Hybridisierungen ist der Spot-Anteil im Fall der Detektion mit Cyanine 5 stets ein Drittel höher als bei der Detektion mit Cyanine 3. In einigen Ausnahmefällen ist der Spot-Anteil fast doppelt so hoch, obwohl für Cyanine 3 die Belichtungszeiten wesentlich höher gewählt wurden (Duo-Cy5: 1,5 s u. 2,3 s; Duo-Cy3: 2,3 s u. 3,5 s), ohne dabei zumindest für einen Teil der Spots die Signalsättigung erreicht zu haben. Dies legt die Vermutung nahe, dass der aktuelle Cyanine 3 konjugierte Biotin-Antikörper sich sowohl für Einzel- wie auch kompetitive Hybridisierungen nicht eignet.

Da dieser Antikörper aus einer anderen Charge stammt, als der in Kap. 3.4.2 verwendete Antikörper, könnte dies ein Grund für seine wesentlich schlechtere *Performance* sein.

Zum Vergleich dieses Antikörpers sollte nun ein selbst hergestellter Cyanine 3 Biotin-Antikörper getestet werden. Im Gegensatz zum seither verwendeten Antikörper ist dieser polyklonal. Die Fluoreszenzgekoppelung mit Cyanine 3 erfolgt analog dem Cyanine 5 konjugierten Antikörper für Digoxigenin. In einer Testserie wurden je zwei Doppelhybridisierungsexperimente mit biotin- und digoxigenin-modifizierten cDNAs aus Gesamt-RNA mit der humanen Brustkrebszelllinie MDA-MB 231 durchgeführt. Dabei wurde das jeweilige Signal/Hintergrundverhältnis sowie die Streuung zwischen den beiden Biotin-Antikörpern untersucht (s. Abb. 19a,b).

19a)



19b)

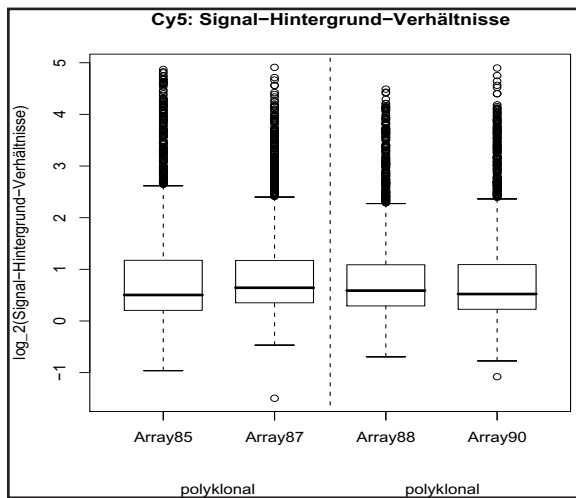


Abb. 19: Boxplots zur Darstellung der Signal-/Hintergrundverhältnisse der unterschiedlichen Antikörper, aufgetrennt in die jeweiligen Farbkanäle Cyanine 3 und 5. 19a) Vergleich zwischen den beiden Biotin-Antikörpern. Der polyklonale Antikörper wurde selbst mit Cyanine 3 fluoreszenz-gekoppelt, wohingegen der monoklonale Antikörper bereits in Cyanine 3 konjugierter Form kommerziell erhältlich ist. 19b) Vergleich der selbst Cyanin 5 fluoreszenz-gekoppelten Antikörper zwischen den beiden Doppelhybridisierungsexperimenten Array 85\* u. 87\* im Vergleich mit Array 88\* u. 90\* (\* interne Bezeichnungen).

In Abb. 19a) ist der Unterschied der beiden Antikörper im Signal-/Hintergrundverhältnis eindeutig zu erkennen (Zur Erklärung der einzelnen Bestandteile des Boxplots s. S.24). So ist hier der jeweilige Median bei Verwendung des monoklonalen Antikörpers erheblich kleiner als bei den polyklonalen Antikörpern (Darstellung der Signal-/Hintergrundverhältnisse erfolgt auf log-2-Basis). Des Weiteren ist beim monoklonalen Antikörper der *Interquartile range* klein und damit sind die Boxen wesentlich schmäler, was auf eine geringe Streuung der Signal-/Hintergrundverhältnisse schließen lässt, so dass in diesem Fall viele Signale nicht eindeutig vom Hintergrundrauschen zu unterscheiden sind. Zur eindeutigen Unterscheidung zwischen Signal und Hintergrund ist jedoch eine möglichst große Streuung zwischen den Signal/Hintergrundverhältnissen erwünscht. Exakt diesen Sachverhalt hat der Cyanine 5 Digoxigenin Antikörper bereits gezeigt (Vgl. die Größe der Boxen in Abb. 19 sowie den Anteil der detektierbaren Spots in Tab. 12 und 13). Ferner treten beim monoklonalen Antikörper eine größere Anzahl von sog. „Ausreißern nach unten“ auf. Als Konsequenz verschwinden viele Signale mit geringer Intensität im Hintergrund.

Dadurch wird auch ein wesentlich kleinerer Anteil der Signale als solche erkannt und gewertet. Ein weiterer Punkt ist, dass das 75 % Quantil bereits kleiner als der Median bei den polyklonalen Antikörpern ist, was allgemein auf geringere Signalintensitäten, im Gegensatz zu den polyklonalen Antikörpern, schließen lässt. Anders verhält es sich bei der Verwendung der polyklonalen Antikörper, bei denen viele der Datenpunkte deutlich oberhalb des Medians liegen (siehe Größe der Box, sowie Abstand zum *Whisker*). Somit ist auch das Ergebnis aus den Tabellen 12 und 13 zu erklären, in denen dargelegt wird, dass mit dem Cyanine 5 markierten Antikörper stets erheblich mehr prozentuale Spot-Anteile detektiert werden konnten.

### 3.5 Tumorscreening

Im Rahmen dieser Studie sollte neben der Etablierung und Validierung des EST-Chips auch ein erstes *Tumorscreening* an insgesamt 24 Gewebeproben (21 Tumor- und 3 Normalgewebe) durchgeführt werden. Die Proben stammen aus der Gewebebank vom RBK. Die Ausführung der Experimente erfolgte im Duplikat. In Tabelle 14 sind die untersuchten Gewebe histopathologisch näher beschrieben.

Jounalnr.	Int. Bez.	pT	pN	M	ER	PR	erbB2	Eppendorf	Fraunhofer
2227/02	T1								+
2611/02	T2	2	1bi	0	0	0		+	+
7254/03	T3	4b	3	0	12	2	1	+	+
4160/02	T4	1	1	0	0	0	1		+
4742/03	T5								+
3409/03	T6	3	3a	1	8	2	3	+	+
11341/02	T7	4d	2	k.A.	0	0		+	+
14157/02	T8	3	2a	0	9	9	1	+	+
13244/03	T9	3	1mic	0	12	12	0	+	+
7168/02	T10	1	1biii	0	12	9	1	+	+
9787/02	T11	2	1a	0	3	2	1	+	+
3562/03	T12	2	0	0	12	4	1	+	+
13194/03	T13	2	0	0	0	0	0	+	+
7185/03	T14	2	0	0	4	9	1		+
8133/03	T15	2	0	0	6	9	0		+
185/03	T16	2	0	0	12	12	0		+
10616/02	T17	2	0	0	4	4	1		+
11415/02	T18	3	0		9	9	1	+	+
12603/02	T19	2	0	0	6	0	2	+	+
8631/03	T20	2	0	0	0	0	3	+	+
2708/03	T21	3	0	0	12	9	2		+
9561/02*	N1							+	+
11754/02*	N2							+	+
11809/02*	N3							+	+

Tab. 14: Histopathologische Details zu den untersuchten Geweben. TNM-Klassifikation + Hormonrezeptoren- und erbB2-Status; Int. Bez. = interne Bezeichnung; pT = Tumorstadium nach TNM: T0 - T4a -d; pN = Lymphknotenstatus nach TNM: N0 - N3, N1; M = Fernmetastasen nach TNM: M0-M1; ER = Östrogenrezeptor: (0-12: 0-2 eindeutig negativ, 3+4 grenzwertig, 5-8 positiv, 9-12 stark positiv); PR = Progesteronrezeptor (anaolg ÖR); erbB2 = Wachstumsfaktor-Rezeptor-Status (0-3; 0-1 negativ, 2 = grenzwertig, 3 = eindeutig positiv); (TNM-Klassifikation nach Wittekind et al. und Ergänzungen aus Therapieleitlinie Mammakarzinom, OSP Stuttgart); +Eppendorfchip = durchgeführte Experimente Eppendorfchip; +Fraunhofer = durchgeführte Experimente IGB-Chip.



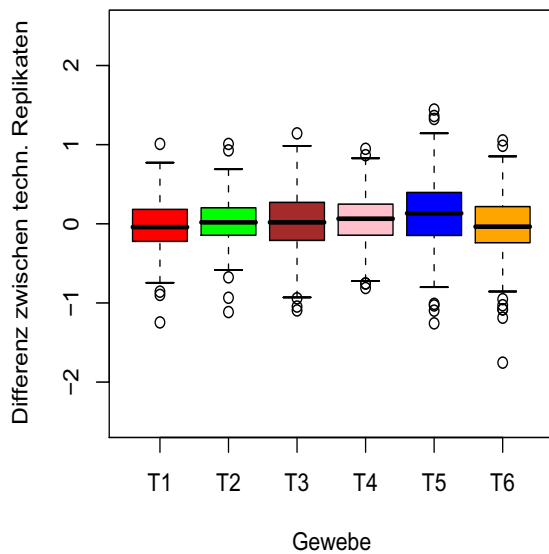
Für die Tumore 2227/02 und 4742/03 wurde keine histopathologische Analyse durchgeführt. Bei den mit \* gekennzeichneten Proben handelt es sich um Normalgewebe, die ebenfalls histopathologisch nicht näher untersucht wurden, wobei diese lt. Klinik als schwach östrogenrezeptor-positiv anzusehen sind. Nicht alle Proben wurden mit beiden DNA-Microarray-Systemen untersucht. So erfolgte die Auswertung aller o. g. Proben zunächst mit dem IGB-Chip.

Dabei fokussierte sich das Interesse hauptsächlich auf die differentielle Genexpression zwischen den Östrogenrezeptor-positiven (ER<sup>+</sup>) und -negativen Proben (ER<sup>-</sup>) sowie auf den Vergleich zwischen Tumor- und Normalgewebe.

Neben der Unterscheidung zwischen Östrogenrezeptor-positiven und Östrogenrezeptor-negativen Tumoren war für die klinischen Kooperationspartner auch die Unterscheidung zwischen erbB2<sup>+</sup>- und erbB2<sup>-</sup>-Tumore von Interesse. Der erbB2-Rezeptor (auch HER2) stellt neben dem Östrogenrezeptor die zweite, wichtige Tumorgruppe dar. Aufgrund des Genkollektivs von der Eppendorf-Arrayplattform, konnte mit den für erbB2<sup>+</sup>/erbB<sup>-</sup>-Tumore charakteristischen Genen erbB2, GRD7 und STARD3 auch diese Tumorgruppe unterschieden werden. Deshalb wurden mit dem Eppendorf-Array nicht alle Tumorproben aus Tabelle 14 untersucht. Bevor die Daten der statistischen Analyse unterzogen wurden, erfolgte zunächst die Qualitätskontrolle hinsichtlich der Reproduzierbarkeit bei den technischen Replikaten. Die Hybridisierungen der 24 Gewebeproben wurden im Duplikat ausgeführt. Mit 5 der Gewebe wurde ein drittes technisches Replikat erstellt, da bei diesen Proben teilweise eine große Diskrepanz aufgetreten war. Die Reproduzierbarkeit wird in diesem Fall mittels Boxplots aus den Differenzen der *log ratios* von den Duplikaten und in sog. CAT-Plots dargestellt. Für diese Plots wurden im Fall von 3 durchgeführten technischen Replikaten jeweils die beiden mit der geringsten Diskrepanz der *log ratios* ausgewählt. Die nachfolgenden Boxplots deuten auf eine gute Reproduzierbarkeit hin, wenn die Mediane nahe bei Null liegen und zusätzlich eine geringe Streuung aufweisen. Diese Kriterien werden beispielsweise von den Tumoren T1, T2 und T6 erfüllt. Deutlich größere Streuungen zeigen hingegen T3 und T5.

Daneben ist es auch interessant die Übereinstimmung des *Rankings* der Gene hinsichtlich des Expressionsmaßes zu vergleichen. Bei der Verwendung eines von Irizarry et al. entwickelten, deskriptiven Plots (CAT-Plots, *correspondence at the top*) lässt sich diese Übereinstimmung darstellen (Irizarry et al., 2005). Für die Erstellung eines CAT-Plots erfolgte zunächst eine Anordnung der *log ratios* in absteigender Reihenfolge (nach Absolutwerten) für je zwei technische Replikate. Die beiden daraus resultierenden Genlisten werden unter Berücksichtigung unterschiedlicher Listengrößen miteinander verglichen. Bezogen auf die jeweilig betrachtete Listengröße erfolgt die Ermittlung des Anteils gemeinsamer Gene der beiden Listen. In den nachfolgenden CAT-Plots ist ein Ausschnitt von 1 bis 10 % Listengröße dargestellt. Im Fall der hier erfassten Tumore (T1-T6) liegt z.B. der Anteil gemeinsamer Gene in den Listen bei einer Listengröße von 28 Genen (ca. 3% der maximalen Listengröße bei 20 - 50 %). Ein Vergleich mit den Werten in der von Irizarry ausgewerteten Daten zeigt, daß die in dieser Arbeit ermittelten Werte dort im mittleren Bereich liegen würden.

20a)



20b)

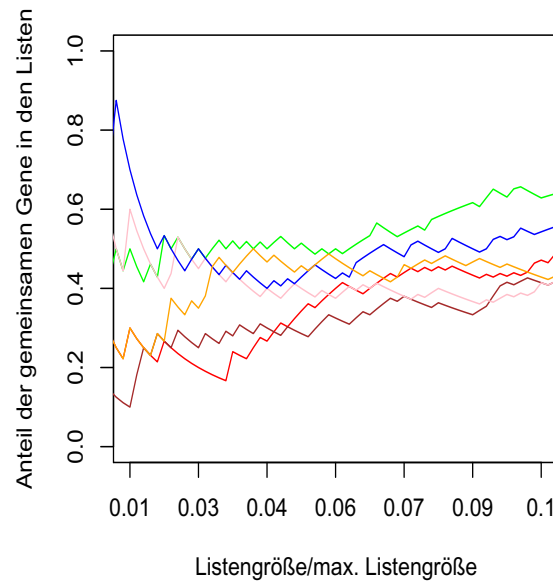
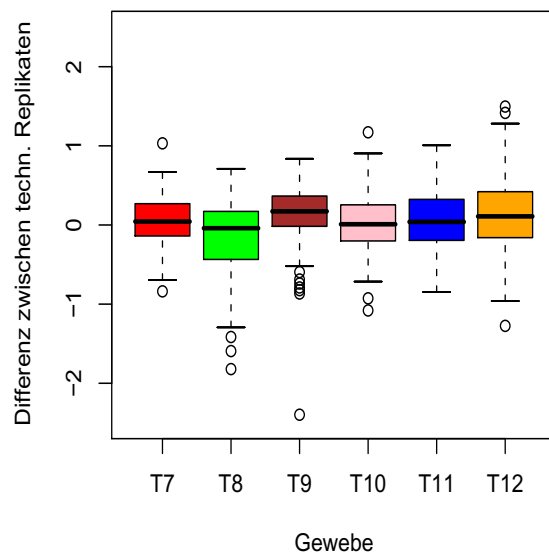


Abb. 20) Qualitätsüberprüfung der Hybridisierungen für Tumore T1-T6. 20a) Boxplots aus den Differenzen der *log ratios* der beiden technischen Replikate für die Tumore T1-T6. 20b) CAT-Plots der Vergleiche für die beiden technischen Replikate der Listen bezogen auf die jeweilige Listengröße. Die Farbgebung innerhalb der Boxplots wurde für die Kurven des CAT-Plots beibehalten.

21a)



21b)

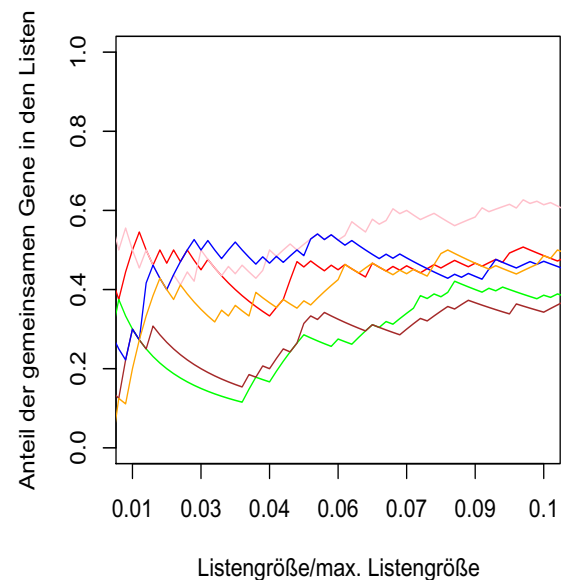
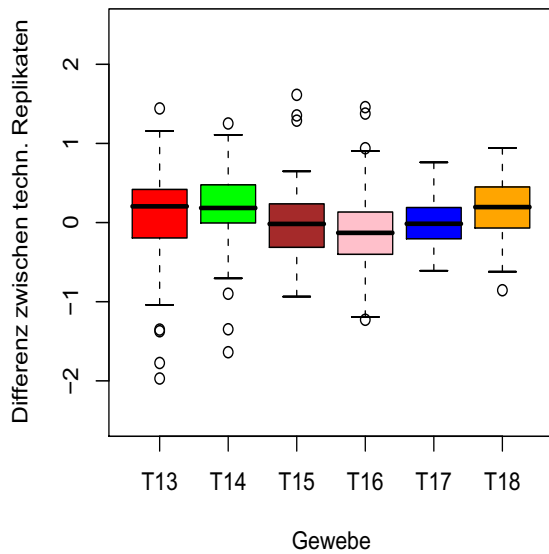


Abb. 21) Qualitätsüberprüfung der Hybridisierungen für Tumore T7-T12. 21a) Boxplots aus den Differenzen der *log ratios* der beiden technischen Replikate für die Tumore T7-T12. 21b) CAT-Plots der Vergleiche für die beiden technischen Replikate der Listen bezogen auf die jeweilige Listengröße. Die Farbgebung innerhalb der Boxplots wurde für die Kurven der CAT-Plots beibehalten.

Bei den Tumoren T7, T10 und T11 sind die Mediane der technischen Replikate nahe bei Null und die Streuung eher gering. Größere Streuungen dagegen zeigen die Tumore T8 und T12. Die Übereinstimmung der gemeinsamen Gene mit den betragsmäßig größten *log ratios* der technischen Replikate untersuchten Proben liegt hier bei etwa 40 - 50 %.

22a)



22b)

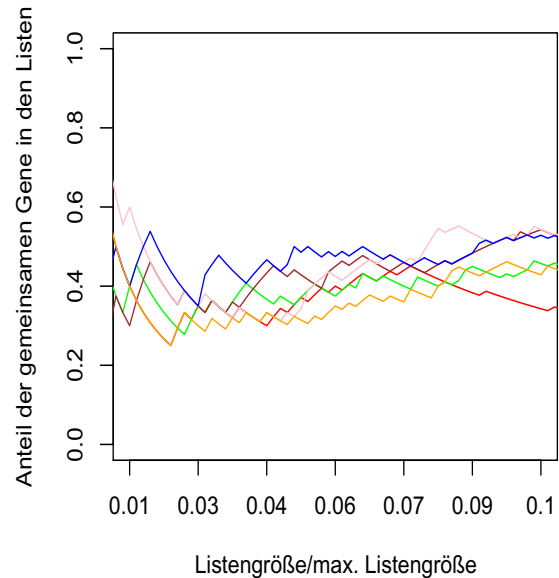
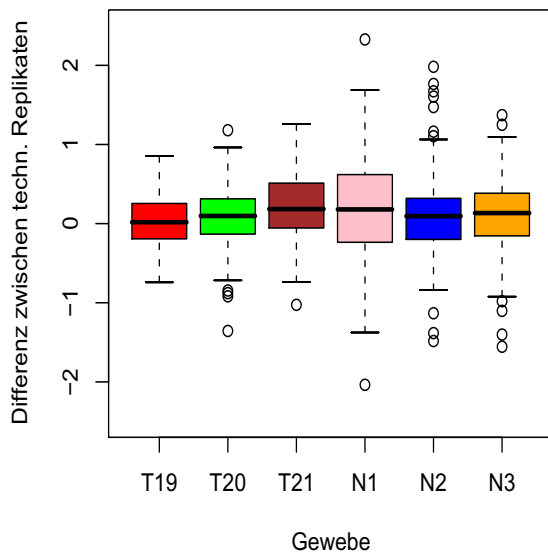


Abb. 22) Qualitätsüberprüfung der Hybridisierungen für Tumore T13-T18. 22a) Boxplots aus den Differenzen der *log ratios* der beiden technischen Replikate für die Tumore T13-T18. 22b) CAT-Plots der Vergleiche für die beiden technischen Replikate der Listen bezogen auf die jeweilige Listengröße. Die Farbgebung innerhalb der Boxplots wurde für die Kurven der CAT-Plots beibehalten.

23a)



23b)

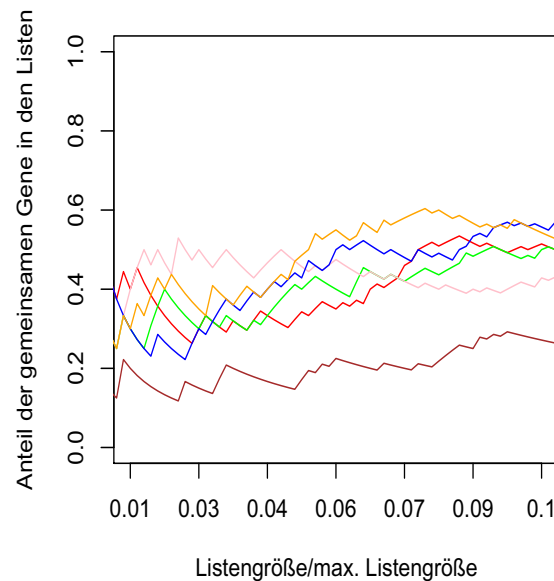


Abb. 23) Qualitätsüberprüfung der Hybridisierungen für Tumore T19-T21 und Normalgewebe N1-N3. 23a) Boxplots aus den Differenzen der *log ratios* der beiden technischen Replikate für die Tumore T19-T21 und für die Normalgewebe N1-N3. 23b) CAT-Plots der Vergleiche für die beiden technischen Replikate der Listen bezogen auf die jeweilige Listengröße. Die Farbgebung innerhalb der Boxplots wurde für die Kurven der CAT-Plots beibehalten.

Diese Plots haben von allen (T18-T21) die große Streuung (vor allem T21 und N1) und die niedrigsten Werte im CAT-Plot.

Anhand der Boxplots zeigen die technischen Replikate eine mittlere bis gute Reproduzierbarkeit.

### 3.5.1 Differentiell exprimierte Gene in ER<sup>+</sup>- und ER<sup>-</sup>-Gewebe

Die statistische Auswertung der Arrays des *Tumorscreens* ergab 35 differentiell regulierte Gene, die hier nach aufsteigendem p Wert sortiert wurden ( $p < 0,05$ ). Eine positive Teststatistik weist dabei auf stärkere Expression im ER<sup>+</sup>-Gewebe als in ER<sup>-</sup>-Gewebe. Bei negativer Teststatistik verhält es sich umgekehrt. Des Weiteren sind hier die gemittelten logarithmierten Signalintensitäten der ER<sup>+</sup> und ER<sup>-</sup>-Gewebe sowie die zugehörigen Werte der Referenz gezeigt.

Name	ID	adj. p-Wert (BH)	t	MW Gewebe ER-	MW Ref. Er-	MW Gewebe ER+	MW Ref. Er+
BTF3	NM_001207	4.23E-05	6.44	19.79	20.98	20.37	20.39
CSTB	NM_000100	1.18E-04	-5.99	19.37	18.16	18.84	18.69
TNC	NM_002160	3.86E-04	-5.60	18.99	18.38	18.53	18.84
PLAT	NM_000930	1.35E-03	5.13	18.34	18.40	18.75	17.99
RPL23A	NM_000984	2.66E-03	4.73	20.82	21.26	21.16	20.91
EMP1	NM_001423	2.66E-03	-4.78	18.29	17.78	17.94	18.13
ABCC1	NM_004996	5.45E-03	-4.46	18.41	17.95	18.19	18.17
JDP1	NM_021800	7.39E-03	4.40	17.88	18.21	18.21	17.89
NFIB	NM_005596	8.05E-03	-4.26	19.02	18.24	18.64	18.62
KRT7	NM_005556	9.92E-03	-4.11	19.40	18.27	18.88	18.78
CIRBP	NM_001280	9.92E-03	4.11	18.64	19.11	18.87	18.88
TOP2A	NM_001067	1.06E-02	4.06	19.81	20.54	20.05	20.30
LTBP1	NM_000627	1.24E-02	-3.99	19.12	18.78	18.91	18.99
CCNB1	NM_031966	1.45E-02	3.91	17.76	18.33	18.01	18.08
MMP7	NM_002423	1.45E-02	-3.92	19.73	17.77	19.24	18.26
PABPC1	NM_002568	1.60E-02	-3.82	21.06	20.26	20.68	20.64
ESR1	NM_000125	1.69E-02	3.85	18.30	18.45	18.89	17.86
ITGB1	NM_002211	1.70E-02	-3.77	19.14	19.29	18.92	19.51
BZRP	NM_000714	1.79E-02	-3.78	18.46	18.14	18.22	18.38
NUMA1	NM_006185	1.83E-02	5.55	16.51	18.38	17.60	17.28
EIF4E	NM_001968	2.35E-02	3.63	17.93	18.29	18.13	18.09
CDH3	NM_001793	2.35E-02	3.60	20.96	21.05	21.20	20.81
CSE1L	NM_001316	2.35E-02	3.60	17.59	19.49	17.96	19.12
MGC15419	NM_024735	2.78E-02	3.53	19.17	18.72	19.41	18.48
CENPA	NM_001809	3.18E-02	3.62	17.84	18.28	18.08	18.03
LCN2	NM_005564	3.18E-02	-3.55	18.49	17.31	18.03	17.77
ATF4	NM_001675	3.18E-02	-3.41	19.18	19.86	18.90	20.13
XTP1	NM_018369	3.41E-02	3.39	18.87	18.88	19.16	18.60
PQLC2	NM_017765	3.61E-02	3.36	19.16	19.11	19.34	18.93
B3GNT4	NM_030765	3.67E-02	3.35	19.64	18.83	19.85	18.63
ACVRL1	NM_000020	3.67E-02	3.44	18.63	18.11	18.95	17.79
SLC39A6	NM_012319	3.75E-02	3.32	19.11	19.03	19.83	18.31
FAT	NM_005245	3.97E-02	-3.33	18.06	17.90	17.80	18.15
CBFB	NM_001755	4.22E-02	3.24	19.41	19.02	19.65	18.79
S100A6	NM_014624	4.61E-02	-3.20	20.66	19.70	20.03	20.33

Tab. 15: Differentiell exprimierte Gene zwischen ER<sup>-</sup> und ER<sup>+</sup> Gewebe unter jeweiliger Einbeziehung der Referenzproben. adj. p-Wert (BH) = adjustierter p-Wert nach Benjamini-Hochberg, t = t-Teststatistik, MW-Gewebe = gemittelte logarithmierte (log<sub>2</sub>) Signalintensitäten Gewebe, MW-Ref. = gemittelte logarithmierte (log<sub>2</sub>) Signalintensitäten Referenz.

Die unterschiedlichen Werte der jeweiligen Referenzen sind auf experimentielle Gründe, wie beispielsweise unterschiedliche Arrays oder unterschiedliche Gewebe, gegen die gemessen wurde, zurückzuführen.

### 3.5.2 Differentiell exprimierte Gene in Normal- und Tumorgewebe

Die statistische Auswertung der Arrays des *Tumorscreens* ergab 23 differentiell regulierte Gene, die hier nach aufsteigendem p Wert sortiert wurden ( $p < 0,05$ ). Eine positive Teststatistik weist dabei auf eine stärkere Expression im Tumorgewebe als im Normalgewebe, bei negativer Teststatistik verhält es sich umgekehrt. Alle anderen Werte sind analog denen in Tabelle 15 dargestellt.

Name	ID	adj. p-Wert (BH)	t	MW Normal-gewebe	MW Ref. Normal	MW Tumor-gewebe	MW Ref. Tumor
VDR	NM_000376	5.31E-06	-7.03	19.32	18.21	18.63	18.91
SPP1	NM_000582	2.12E-03	5.01	18.41	18.94	18.96	18.39
COL4A1	NM_001845	2.12E-03	4.94	19.07	21.94	19.68	21.33
KRT19	NM_002276	2.45E-03	4.80	19.69	20.19	20.31	19.57
KRT8	NM_002273	4.71E-03	4.50	19.18	20.81	19.72	20.26
CDH3	NM_001793	4.71E-03	-4.48	21.44	20.57	21.01	20.91
CSTB	NM_000100	4.71E-03	4.47	18.50	19.03	19.04	18.49
CLDN1	NM_021101	7.53E-03	4.27	20.06	22.90	20.80	22.16
BBC3	NM_014417	9.65E-03	-4.19	18.68	17.70	18.09	18.29
CX3CL1	NM_002996	9.65E-03	-4.21	18.23	17.50	17.92	17.82
CASP2	NM_032982	1.15E-02	-4.02	19.88	18.59	19.62	18.86
IGFBP4	NM_001552	1.21E-02	-3.98	19.20	18.50	18.83	18.87
A2M	NM_000014	1.39E-02	-4.08	19.13	17.22	18.67	17.68
MIF	NM_002415	1.39E-02	3.90	17.97	19.33	18.46	18.84
VEGF	NM_003376	2.00E-02	3.76	18.27	18.71	18.74	18.24
IGF2	NM_000612	2.14E-02	-3.70	19.11	19.08	18.84	19.35
KIT	NM_000222	2.15E-02	-3.74	18.92	17.97	18.57	18.32
ARHB	NM_004040	3.01E-02	-3.53	20.74	19.57	20.57	19.75
JUN	NM_002228	3.01E-02	-3.71	18.32	16.88	17.69	17.51
IFI30	NM_006332	3.01E-02	3.50	18.57	19.59	18.97	19.19
BSG	NM_001728	3.13E-02	3.48	17.90	19.32	18.21	19.01
KRT18	NM_000224	3.55E-02	3.42	18.31	19.44	18.70	19.06
LAMA3	NM_198129	3.64E-02	-3.52	18.17	17.60	17.84	17.93

Tab. 16: Differentielle Genexpression zwischen Tumor und Normalgewebe unter jeweiliger Einbeziehung der Referenzproben; adj. p-Wert (BH) = adjustierter p-Wert nach Benjamini-Hochberg, t = t-Teststatistik, MW-Tumorgewebe = gemittelte, logarithmierte ( $\log_2$ ) Signalintensitäten Tumorgewebe; MW-Normalgewebe. = gemittelte logarithmierte ( $\log_2$ ) Signalintensitäten Normalgewebe.

### 3.6 Datenauswertung

#### 3.6.1 Vergleich der beiden Chip-Systeme

Von den 778 Sequenzen auf dem IGB-Chip und den 222 Sequenzen des Eppendorf-Chips gibt es eine Überschneidung von 125 Genen. Diese 125 Gene können in zwei Kategorien A und B sowie in weitere 4 Gruppen unterteilt werden (s. Abb. 24, von links nach rechts).

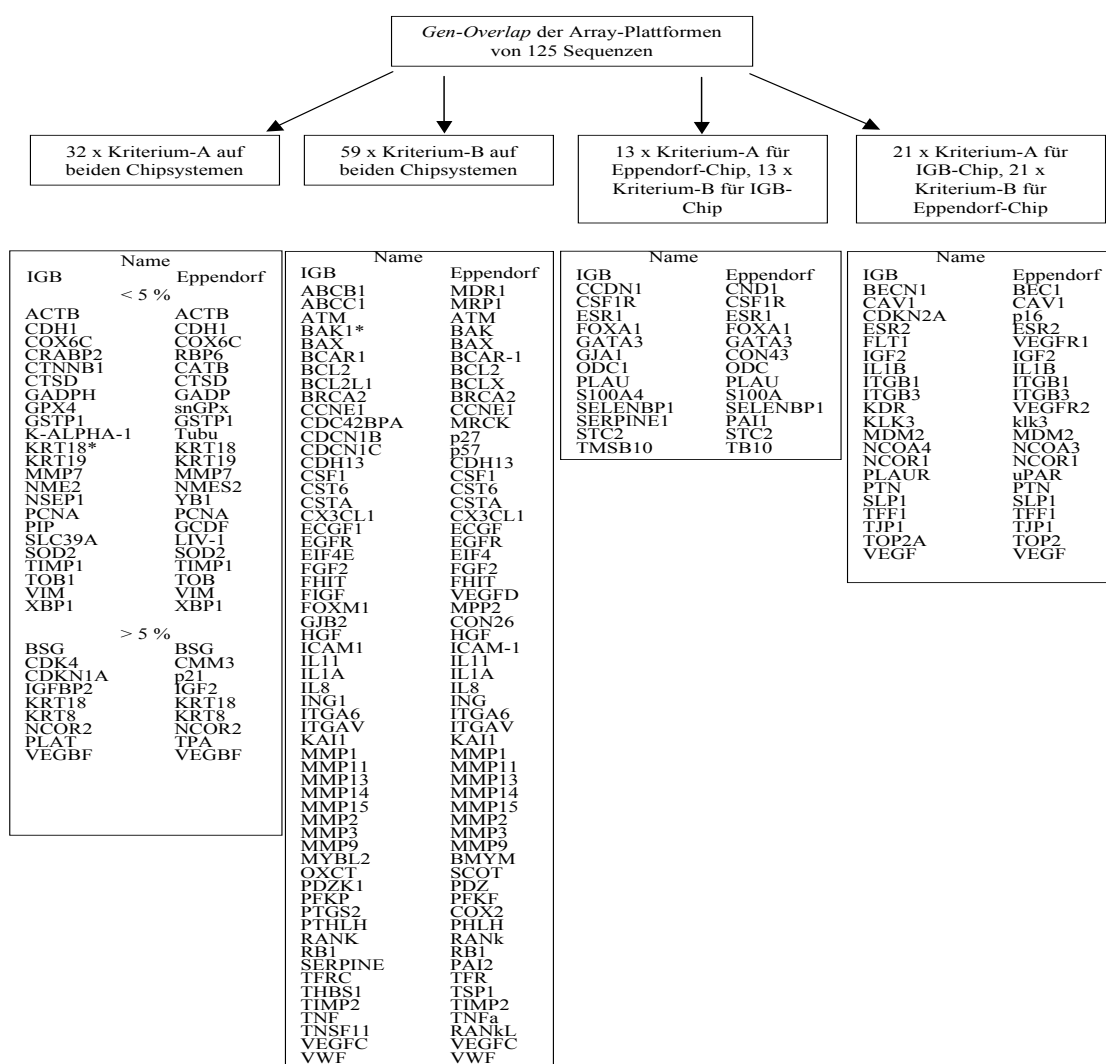


Abb. 24: Darstellung des *Gen-Overlap* der Array-Plattformen mit weiterer Aufteilung (v. links nach rechts). Die 1. Gruppe unter Kriterium-A umfasst diejenigen Gene, die auf beiden Chip-Systemen, bei > 50 % der untersuchten Gewebe vorkommen und deren log-2 ratios bzw. log-2 Nettosignale hier einen positiv, signifikanten Zusammenhang (adj. p-Wert BH < 5 % bei Test auf Assoziation, d.h. Korrelation > 0) zeigen und zudem alle in Kap. 2.2.15 genannten Qualitätskriterien erfüllen. Ein weiteres Set an Genen der 1. Gruppe unter Kriterium-A ist zwar ebenfalls auf beiden Chipsystemen bei >50 % aller untersuchten Gewebe vorhanden, allerdings zeigen hier die log-2 ratios bzw. log-2 Nettosignale beim Test auf Assoziation einen adj. p-Wert BH > 5 %. Gene der 2. Gruppe sind zwar auf beiden Chipsystemen detektierbar, erfüllen jedoch nicht die Qualitätskriterien, haben zudem einen adj. p-Wert > 5% und fallen damit unter das Kriterium B. Bei der Gruppe 3 ist ein Genset vorhanden, das die Qualitätskriterien zwar für den Eppendorf- jedoch nicht für IGB-Array erfüllt. Umgekehrt verhält es sich beim Genset der 4. Gruppe. Gelegentlich haben die gleichen Gene zwischen den beiden Chipsystemen unterschiedliche Genabkürzungen. Weitere Informationen zu den hier aufgelisteten Genen finden sich im Anhang.

Die Kategorie A umfasst dabei diejenigen Gene die auf beiden Arraysystemen detektierbar sind, deren  $\log_2$  ratios bzw.  $\log_2$  Nettosignale beim Test auf Assoziation einen adj. p-Wert  $< 5\%$  haben sowie alle in Kap. 2.2.15 genannten Qualitätskriterien erfüllen. In Kategorie B sind alle Gene erfasst, welche zwar ebenfalls auf beiden Chipssystemen detektierbar sind, deren  $\log_2$  ratios bzw.  $\log_2$  Nettosignale beim Test auf Assoziation jedoch einen adj. p-Wert  $> 5\%$  und gleichzeitig nicht die Qualitätskriterien erfüllen.

Innerhalb o.g. Schnittmenge (Kategorie A, adj. p-Wert BH  $< 5\%$ ) gibt es insgesamt 23 Sequenzen deren Korrelationskoeffizienten, berechnet aus den  $\log$  ratios bzw. den  $\log$ -Signalintensitäten der beiden Chip-Systeme,  $\geq 0,78$  ist (ein Bsp. s. Abb. 25). Die übrigen Gene besitzen beim Test auf Assoziation einen Korrelationskoeffizienten  $> 0$ , aber  $< 0,78$ , haben einen adj. p-Wert BH  $> 5\%$  und sind damit nicht mehr signifikant.

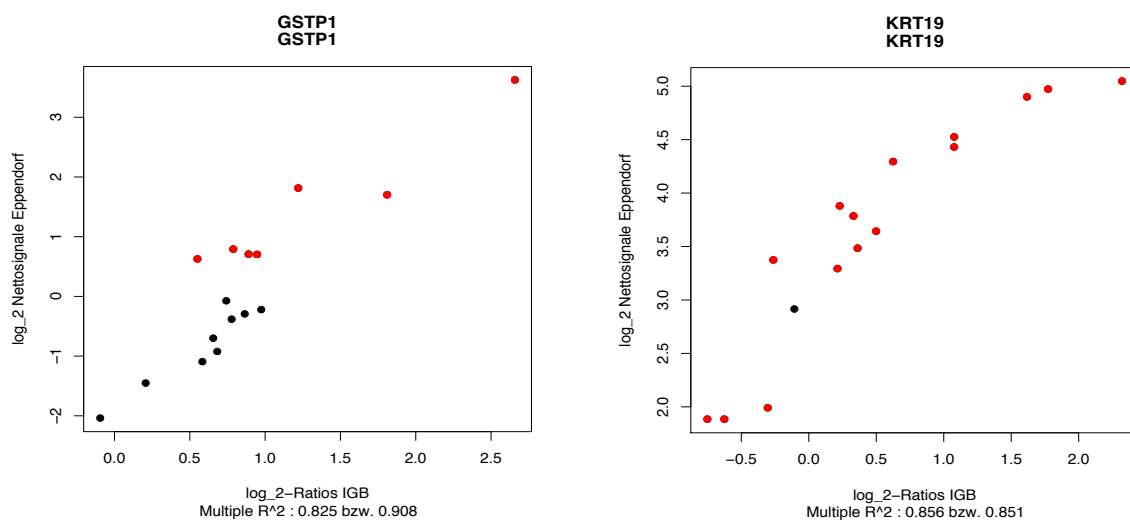


Abb. 25: Beispiele des Vergleichs gemeinsamer Sequenzen mit hoher Korrelation. Rote Punkte repräsentieren Spots, die auf beiden Chip-Systemen durch keines der in Kap. 2.2.15 genannten Qualitätskriterien ausgeschlossen wurden und damit volles Gewicht haben.

In den beiden Plots wurden die  $\log_2$  Nettosignale des Eppendorf-Chips über den  $\log_2$  Ratios des IGB-Chips aufgetragen. Dabei repräsentiert jeder Punkt ein Gewebe. Aus den multiplen Regressionskoeffizienten ( $R^2$  entspricht den Quadraten des Korrelationskoeffizienten) der beiden Beispiele ergeben sich Korrelationen von 0,91 bzw. 0,93.

Etwa 85 % der Sequenzen aus der o.g. Schnittmenge zeigen jedoch deutlich schlechtere Korrelationen u.a. auch Gene aus Kategorie B. Die häufigste Ursache dafür ist, dass diese Sequenzen auf beiden Chipssystemen teilweise ein sehr geringes Expressionsmaß haben, dazu noch große Streuungen aufzeigen, nicht detektierbar sind und damit letztlich nicht die in Kap. 2.2.15 genannten Qualitätskriterien erfüllen (Abb. 24, aus 2. Gruppe).

Abb. 26 zeigt zwei gegenteilige Beispiele von Abb. 25. Neben ausschließlich negativen  $\log_2$  Werten bzw.  $\log_2$  Ratios  $< 0,3$  erfüllt auch kein Gewebe für diese Gene die Qualitätskriterien aus Kap. 2.2.15. Somit fallen diese Gene in das Hintergrundrauschen.

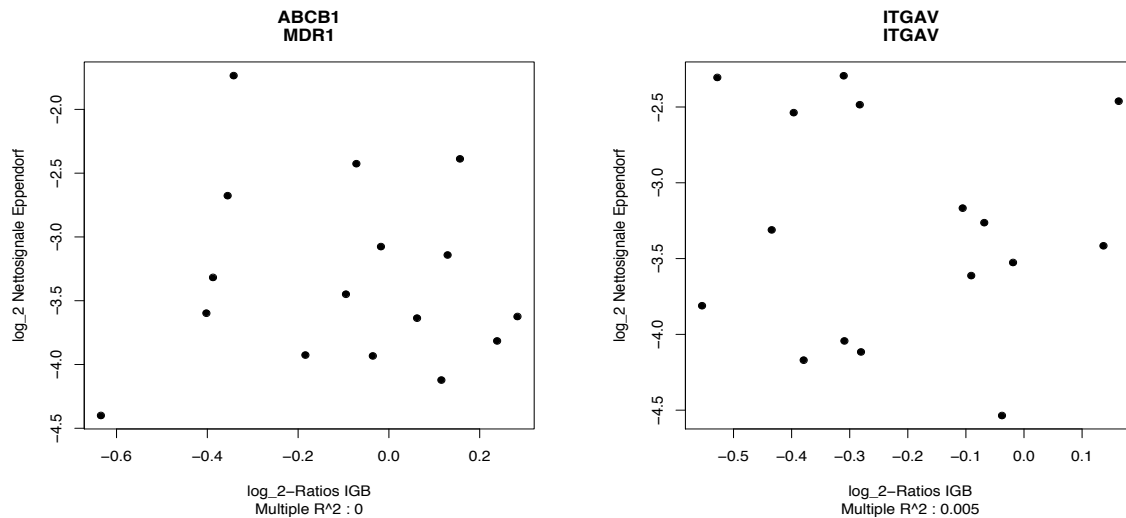


Abb. 26: Beispiele des Vergleichs gemeinsamer Sequenzen mit niedriger Korrelation ( $\geq 0$ ). Schwarze Punkte repräsentieren Spots die auf beiden Chip-Systemen keines der in Kap. 2.2.15 genannten Qualitätskriterien erfüllen und damit nicht weiter Gewicht haben.

Unter Einbeziehung dieser Gene (Korrelation  $> 0$ , aber  $< 0,78$ ) wird der Vergleich beider Chip-Systeme hinsichtlich gemeinsam differentiell exprimierter Gene erschwert. Ein weiterer Punkt bei der Analyse beider Chipsysteme sind die Gene aus den Gruppen 3 und 4 (34 Gene), wo Gene auf dem einen Chipsystem, die Qualitätskriterien erfüllen und deren adj. p-Werte BH beim Test auf Assoziation  $< 5\%$  sind, auf dem anderen wiederum nicht. Damit sind diese Gene für den Vergleich der Chipsysteme nicht geeignet.

Hinzu kommt, dass etwa die Hälfte der Gene aus der Schnittmenge eine außerordentlich geringe Expression nahe dem Hintergrund zeigten (59 Gene von 125, s. Gruppe 2). Allerdings besitzen diese Gene zumindest eine ähnliche Expressionstendenz auf beiden Chipsystemen. Somit bleibt lediglich eine kleine Anzahl Gene, die in beiden Systemen starke Signale und einen signifikant positiven Zusammenhang zeigen (23 Gene mit adj. p-Wert  $< 5\%$ , anteilig ca. 20%).

Unter Zusammenfassung der Gene aus den Gruppen 1 und 2 lässt sich jedoch auch sagen, dass der IGB-Chip mit über 70% (genau 72,8%) der Expressionstendenz des Eppendorf-Chips folgt. Dies ist dennoch ein durchaus akzeptables Ergebnis, vor allem wenn man noch den sog. „Labor-Effekt“, mitberücksichtigt. So berichtet die Studie von Irizarry et al., dass bei Arrayexperimenten mit identischen RNA-Proben in verschiedenen Laboratorien sowohl erhebliche Unterschiede hinsichtlich der Daten innerhalb derselben Arrayplattform, als auch zwischen unterschiedlichen Arraysystemen auftreten können und dies letztlich zu einem nicht unerheblichen Anteil auf den sog. Laboreffekt zurückzuführen ist (Irizarry et al, 2004).

Die weiteren Auswertungen beziehen sich auf die Unterteilung der Gewebe in verschiedene Gruppen anhand der histopathologischen Daten wie erbB2<sup>+</sup> und erbB2<sup>-</sup>, ER<sup>+</sup> und ER<sup>-</sup>-Tumoren sowie zwischen Normal- und Tumorgewebe (s. Tab. 14).



### 3.7 Bewertung des Genkollektivs

Die Bewertung des Genkollektivs erfolgt von zwei Seiten. Einerseits werden hier die Ergebnisse der Tabellen 15 und 16 im Hinblick aller mit dem IGB-Chip untersuchten Gewebepopulationen (21 Tumor- und 3 Normalgewebe) bewertet. Dabei werden einzelne Gene exemplarisch mit Erkenntnissen bzw. Ergebnissen aus der Literatur verglichen. Im weiteren erfolgt der Vergleich zwischen den beiden Arrayplattformen anhand der gemeinsam untersuchten Proben (vgl. Tab. 14, insgesamt 16 Gewebe).

#### 3.7.1 Differentielle Geneexpression zwischen erbB2<sup>+</sup>- und erbB2<sup>-</sup>-Tumoren

Für die Untersuchung zwischen den erbB2<sup>+</sup> und erbB2<sup>-</sup>-Tumoren wurden für den IGB-Array zwar Daten erhoben, wobei sich daraus allerdings keine differentiell exprimierte Gene ergaben. Einer der Gründe dafür war, dass die zu untersuchenden Proben einerseits in zu geringem Stichprobenumfang vorlagen und andererseits das Fehlen der für erbB2<sup>+</sup>/erbB2<sup>-</sup>-Tumoren charakteristischen Gene wie erbB2, GRB7 und STARD3. So liegen für diese Analyse ausschließlich Daten hinsichtlich differentieller Gene vor, die mit dem Eppendorf-Chip erhoben wurden. Aus diesen Daten wurden insgesamt 7 differentiell exprimierte Gene zum 5 %- Signifikanzniveau bestimmt. Darunter die drei bereits genannten Sequenzen wie erbB2, GRB7 und STARD3.

Für die restlichen 4 Gene, namentlich TOP2A (*Topoisomerase 2 alpha*), ITGB1 (*Integrin beta 1*), TIMP1 (*Tissue inhibitor of metalloproteinase 1*) und CTSD (*Cathepsin D*) konnte mit dem IGB-Chip keine differentielle Expression (p-Wert > 5%) festgestellt werden.

#### 3.7.2 Differentielle Geneexpression zwischen ER<sup>+</sup>- und ER<sup>-</sup>-Tumoren

Aus den 35 differentiell exprimierten Genen konnten 12 Gene detektiert werden, die in Verbindung mit dem Östrogenrezeptor stehen (West et al., 2001, Grubberger et al., 2001). Neben dem Östrogenrezeptor (ER) selbst, sind dies das östrogenregulierte Gen LIV-1 oder auch SLC39A6 (*solute carrier family 39 member 6*, Dressmann et al., 2001) sowie das in überwiegend ER<sup>-</sup>-Brusttumoren vorkommende Gen LCN2 (*lipocalin 2*, Stoesz et al., 1998). Aus einer anderen Studie (Übersicht: Lacroix et al., 2002) konnten 12 weitere Gene detektiert werden, die verstärkt in ER<sup>-</sup>-Brusttumoren vorkommen. Eine nähere Beschreibung dieser und weiterer Gene erfolgt im Diskussionsteil dieser Arbeit. So lassen sich bereits anhand dieser wenigen Gene ER<sup>+</sup>-von ER<sup>-</sup>-Tumoren eindeutig unterscheiden.

Der Eppendorf-Chip zeigte bei der Untersuchung dieses Tumorkollektivs (s. Tab. 14) insgesamt 20 differentiell exprimierte Gene. Bei diesen 20 Genen ergibt sich eine Schnittmenge zwischen den beiden Arrayplattformen von lediglich 12 Genen. Innerhalb dieser 12 Gene ist nur das Gen MMP7 auch beim IGB-Chip differentiell exprimiert. Die Funktion für MMP7 (*Matrix metalloproteinase 7* oder *matrilysin*) wird im Diskussionsteil unter Kapitel 4.3 weiter erklärt. Interessanterweise war mit dem Eppendorf-Chip der für ER<sup>+</sup>-Gewebe charakteristische Östrogenrezeptor (ÖR) nicht als differentiell exprimiertes Gen detektierbar, wohl aber mit dem IGB-Chip (s. Tab. 15).

### 3.7.3 Differentielle Geneexpression zwischen Normal- und Tumorgewebe

Aus dieser Untersuchung ergaben sich 23 differentiell exprimierte Gene. Neun dieser Gene stammen wiederum aus der Übersichtsarbeit von Lacroix (Lacroix et al., 2002). Mit den Zytokeratinen 8,9 und 19 können Tumor- und Normalgewebe gut unterschieden werden, da diese Gene ausschließlich in Tumorzellen vorkommen. (Trask et al, 1990). Des weiteren konnte mit VEGF (*vascular endothelial growth factor*) ein Schlüsselmolekül der Angiogenese detektiert werden, Die Angiogenese ist für das Tumorwachstum essentiell (Heer et al., 2003). Mit hohen VEGF-Pegeln im Gewebe ergibt sich eine schlechte Prognose bei Brustkrebs. (93. Lee JE., 2004). Auch anhand von diesem Genset (s. Tab 15) lassen sich hier Normal- und Tumorgewebe eindeutig unterscheiden. Die weitere Beschreibung und Bewertung dieser Gene wird im Diskussionsteil fortgeführt.

Beim Eppendorf-Chip ergaben sich ebenfalls 23 differentiell regulierte Gene, wobei hier die Schnittmenge der beiden Arrayplattformen bei 15 Genen liegt. Innerhalb dieser Menge findet sich keines der Gene aus Tabelle 15 wieder.

Wie bereits in Kapitel 3.6 erwähnt, beträgt die Schnittmenge an gemeinsamen Genen der beiden Arrayplattformen 125, oder etwa 16 % (von 778 Sequenzen ausgehend). Damit besteht ein relativ kleiner *Overlap* zwischen den beiden Arraysystemen. Hinzu kommt, dass sich viele der Gene beim Eppendorf-Chip kaum vom Hintergrund abheben bzw. von diesem sogar überlagert werden. So lassen sich möglicherweise die teilweise erheblichen Unterschiede bei den differentiell exprimierten Genen zwischen den beiden Arrayplattformen erklären.

Auch zeigt der IGB-Chip beim Vergleich der beiden Zellpopulationen (s. Kap. 3.4.2) wesentlich mehr differentiell exprimierte Gene (33 Gene vs. 20 Gene) als der Eppendorf-Chip.

An diesen beiden Beispielen konnte gezeigt werden, dass mit beiden Arrayplattformen, durchaus eine Unterscheidung der untersuchten Gewebe- und Zellpopulationen möglich ist. Allerdings ist dabei auch zu prüfen, inwieweit die beiden Genkollektive ergänzt werden müssen.

## 4. Diskussion

Das Ziel dieser Arbeit war die Herstellung und Validierung eines wissensbasierten DNA-Microarrays zur verbesserten Prognose bei Brustkrebserkrankungen. Die Arbeit umfasst im wesentlichen drei Teile:

- Literaturrecherche nach relevanten Markergenen
- Methodenetablierung
- Bewertung der differentiellen Gene nach *Tumorscreen*

Bei der Vorstellung der Ergebnisse erfolgte bereits eine Diskussion, um die einzelnen zur Fortführung der Experimente erforderlichen Entscheidungen logisch nachvollziehbar zu machen. Nachfolgend sollen weiterführende Gedanken und Informationen zu den einzelnen Punkten dargestellt werden.

### 4.1 Literaturrecherche nach relevanten Markergenen

Die Zusammenstellung des Arrays erfolgte unter Berücksichtigung verschiedener Aspekte bzgl. der Abdeckung möglichst vieler Genfunktionen (z. B. therapeutisch bedeutsame Gene, tumortyp-spezifische Sequenzen oder Gene wichtiger Signaltransduktionswege) mit dem Ziel hier einen wissensbasierten Array zu entwickeln. Ein maßgeblicher Grund dafür war, dass die bisher durchgeführten gesamtgenomischen Transkriptionsanalysen eine sehr kleine Schnittmenge bzgl. der verwendeten Genkollektive zeigten (West et al., 2001, van de Vijver et al., 2002, van't Veer et al., 2002, Sorlie et al., 2001, Perou et al., 2000) und als Folge davon auch eine sehr geringe Übereinstimmung gemeinsam exprimierter Gene. Des Weiteren waren diese Transkriptionsanalysen von erheblich finanziellem Aufwand und komplexer Datenanalyse begleitet.

Ausgehend von dem Mini-Array (s. Kap. 3.2) war das Genkollektiv zunächst nicht ausschließlich auf brustkrebsrelevante Gene beschränkt, sondern eher auf Sequenzen, die als Schlüssel-moleküle neben der Kontrolle des Zellzyklus in Proliferation, Differenzierung und Apoptose eine Rolle spielten.

Bei der Neukonzeption des DNA-Microarrays wurden zudem auch Gene, die mit dem Östrogenrezeptorstatus assoziiert sind und mit denen sich eindeutig ER<sup>+</sup>- von ER<sup>-</sup>-Tumoren unterscheiden lassen, berücksichtigt (West et al., 2001, Gruvberger et al., 2001). Des Weiteren wurden vor allem Gene, die möglichst die gesamte Bandbreite zellulärer Funktionen, wie Zellzyklusregulation (z. B. CDKN1A, *cyclin depend kinase inhibitor 1 A*; CDKN1B *cyclin depend kinase inhibitor 1 B*), Angiogenese (z. B. EDN1, *endothelin 1*; FLT1, *Fms-related tyrosine kinase 1*) Proteolyse (z. B. MMP7 *matrix metalloprotease 7* oder *martrilysin*; PLAU, *plasminogen activator, urokinase*), Adhäsion (z. B. GJA1, *gap junction protein alpha 1*; ITGA6, *intergrin alpha 6*), Chemoresistenz (z. B. TOP2A, *topoisomerase II alpha*; ABCB1, *ATP-binding cassette subfamily B, member 1* oder *P-glycoprotein*) und Hormonsensitivität (z. B. TFF1 *trefoil factor 1*; VDR *vitamine D receptor*) abdecken, ausgewählt (Übersicht: Lacroix et al., 2002). Das Genkollektiv wurde um weitere 70 Sequenzen ergänzt, mit denen eine Aussage bzgl. einer 10 jährigen Überlebensrate unter Bildung von Fernmetastasen, ermöglicht werden soll (van't Veer et al., 2002).

Etwa ein Drittel dieser Sequenzen besaß zum damaligen Zeitpunkt noch keine korrekte Annotation. Bei Recherchen in der *Genecards*-Datenbank<sup>6</sup> wurden diese Gene häufig als sog. *hypothetical proteins* bezeichnet, von denen lediglich ihre Chromosomenlokalisierung bekannt war.

Mit Genen, die eine Subklassifizierung von Tumoren u.a. in den basalähnlichen, den normalbrust-ähnlichen sowie den luminalen-epithelialen Typ ermöglichten, wurde das Genkollektiv weiter ergänzt (Sorlie et al., 2001, Perou et al., 2000).

Als Sonden wurden ESTs gewählt. Diese ESTs wurden maßgeblich aus zwei Genbibliotheken des RZPDs, dem *Human Unigene Set*-RZPD 3 oder der Bibliothek Nr. 972 und der Bibliothek Nr. 958 gewählt. Bei beiden Bibliotheken handelt es sich um cDNA-Klonsammlungen, wobei die ESTs aus Nr. 972 als sequenzbestätigt ausgewiesen sind. Im Gegensatz dazu sind die ESTs aus der Bibliothek 958 nicht durch das RZPD sequenzbestätigt. So sind lt. RZPD ungefähr 20 % dieser Klone mit Fehlern bzgl. ihrer Annotation behaftet. Weitere Informationen zu den o.g. Genbibliotheken finden sich unter [http://www.rzpd.de/cgi-bin/products/rzpd\\_products.pl.cgi](http://www.rzpd.de/cgi-bin/products/rzpd_products.pl.cgi).

#### 4.2. Methodenetablierung

Zur Überprüfung hinsichtlich der Zuverlässigkeit des neuen DNA-Microarrays (IGB-Chip) wurde von Seiten der klinischen Kooperationspartner ein bereits etabliertes System verwendet. Als Modellsystem diente dabei die humane Brustkrebszelllinie MCF 7, deren Genexpression zwischen unstimuliertem Zustand und unter Antiöstrogenstimulation (ICI 182780) untersucht wurde. Die Ergebnisverifikation erfolgte mittels verschiedener, von einander unabhängiger Methoden (Eppendorf-Chip, *RNase protection assay* und quantitative *real-time* PCR).

Die statistische Datenanalyse ergab insgesamt 33 differentiell exprimierte Gene. Darunter auch CDKN1A und TFF1, beide durch Östrogen bzw. Antiöstrogene (Musgrove et al., 1995, Gillesby and Zacharewski, 1999) beeinflussbar.

Im Vergleich zwischen dem IGB- und dem Eppendorf-Chip zeigte der IGB-Chip anhand eines Sets von 12 Genen tendenziell in die Richtung vom Eppendorf-Chip (s. Kap. 3.4.2, Abb. 15). Dabei ist jedoch das Expressionsverhalten mancher Gene (z. B. CDH1, MDM2 oder Keratin 8) nicht so deutlich ausgeprägt wie beim Eppendorf-Chip. Mögliche Unterschiede könnten durch die verschiedenen Sondentypen (IGB-Chip: EST-basiert; Eppendorf-Chip: oligo-basiert), hervorgerufen worden sein. Allerdings herrscht andererseits innerhalb dieser 12 Sequenzen bei 5 Genen (VEGFB, Keratin 19, PCNA, CSTD und ODC1) recht gute Übereinstimmung zwischen den beiden Chipsystemen, was etwa 40 % entspricht.

Mit dem RPA (*RNase protection assay*) als zweiter, unabhängiger Methode erfolgte anhand eines kleinen Sets von 6 Genen der Vergleich des IGB-Chips. Die Expressionsdaten des IGB-Chips zeigten auch hier tendenziell in dieselbe Richtung des RPA (s. Tab. 13). Allerdings wurden die Daten des RPA rein qualitativ erfasst.

Die quantitative *real-time* PCR wurde als dritte unabhängige Methode zum Vergleich herangezogen. Die Untersuchung ergab, dass die Expressionsdaten des IGB-Chips in der Tendenz denen der quantitativen *real-time* PCR entsprachen (Vgl. Abb. 14 und 15). Analog zu den beiden anderen Methoden (Eppendorf-Chip und RPA) sind auch hier im Vergleich die Expressionsverhältnisse des IGB-Chips nicht so deutlich ausgeprägt. Hinzu kommt noch, dass sich die quantitative *real-time* PCR durch eine sehr hohe Sensitivität auszeichnet.

<sup>6</sup> <http://www.dkfz.de/GeneCards>

Aufgrund dieser ausgeprägten Sensitivität eignet sich die quantitative *real-time* PCR hervorragend zur Überprüfung einzelner Genexpressionsdaten. So konnten beim Vergleich dieser beiden Methoden immerhin bei 8 von 10 Genen eine übereinstimmende Expressionstendenz gezeigt werden.

Zusammenfassend kann für diese unabhängigen Methoden gesagt werden, dass sie die Ergebnisse des IGB-Chips zu einem großen Teil in ihrer Tendenz bestätigen und sich hier gut zur Ergebnisverifikation eignen. Dieses Ergebnis setzt sich auch beim Vergleich der beiden Chipsysteme anhand der gemeinsam untersuchten Tumorproben fort (s. Kap. 3.6). Dort konnte immerhin für über 70 % der Gene (91 von 125) die gleiche Expressionsrichtung festgestellt werden. Allerdings können unterschiedliche Messsysteme nicht immer hohe Übereinstimmung zeigen. Ein gegenteiliges Beispiel zeigt die Arbeit von Tan et al.. So wurden in dieser Studie mit 3 kommerziell erhältlichen Arrayplattformen unterschiedlicher Hersteller Hybridisierungsexperimente mit identischen RNAs sowie unter gleichen Bedingungen durchgeführt und miteinander verglichen. Bei der Expression differentieller Gene zeigten sich jedoch erhebliche Unterschiede (Tan et al., 2003, Marshall, 2004). So ergaben sich von den unterschiedlichen Arrayplattformen insgesamt 185 differentiell exprimierte Gene, wobei die Schnittmenge hinsichtlich gleichem Expressionsverhalten bei lediglich 4 Genen lag, was etwa 2,2 % Übereinstimmung entspricht.

### 4.3 Bewertung der differentiellen Gene nach dem *Tumorscreen*

Anhand des ausgewählten Genkollektivs konnten mit dem IGB-Array innerhalb der 4 untersuchten Gewebegruppen (ER<sup>+</sup>- ER<sup>-</sup>-Tumore sowie zwischen Normal- und Tumorgewebe) differentiell exprimierte Gene bestimmt werden.

Beim Vergleich zwischen ER<sup>+</sup>-, ER<sup>-</sup>-Tumoren wurden insgesamt 35 differentiell exprimierte Gene gefunden. Innerhalb der beiden anderen Gruppen (Tumor- vs. Normalgewebe) ergaben sich differentiell exprimierte 23 Gene.

#### 4.3.1 Signifikante Gene innerhalb ER<sup>+</sup> -und ER<sup>-</sup>-Gewebe

Wie bereits in Kap. 3.7.1 genannt, beinhaltet dieses Genset zum überwiegenden Teil Gene, die mit dem Östrogenrezeptor in Verbindung stehen bzw. mit diesem koexprimiert werden. Innerhalb dieses Genkollektivs zeigt das Gen BTF3 (*Basic transcription factor 3*) die größte Signifikanz ( $p = 4,23 \times 10^{-5}$ ) für eine verstärkte, differentielle Expression in ER-positiven Geweben. In der Literatur wird ein Zusammenhang zwischen BTF3 und der AF-1-Funktion (*Activation function-1*) des Östrogenrezeptors beschrieben, wodurch dieser aktiviert wird (el Tanni and Green, 1998). AF-1 und AF-2 sind nukleäre Rezeptoren, die als liganden-induzierbare Transkriptionsfaktoren wirken (Kato, 1999). Durch die Vollaktivierung von AF-1 kommt es wiederum zu Proliferation in Brustkrebszellen (Fujita et al., 2003). Inwieweit sich dieses Gen dadurch jedoch als ein relevanter Marker eignet ist unklar, da es in der gegenwärtigen Literatur dazu keine weiteren Abhandlungen gibt. Es bleibt noch zu erwähnen, dass bei sporadisch auftretendem Darmkrebs BTF3 in einem Tumorgenese-Signalweg involviert ist (Dunican et al., 2002). Dieser Tumorgenesetyp erfolgt über chromosomale Instabilität und kommt in 85 % aller Fälle von Darmkrebs vor.

Das Gen CSTB oder *Cystatin B* (*Stefin B*) ist ein Cystein-Proteaseinhibitor, der in diesem Genset mit hoher Signifikanz ( $p = 1,8 \times 10^{-4}$ ) in ER<sup>-</sup>-Gewebe exprimiert wird. Dieses Gen wurde auch in der Studie von Sorlie als Marker bei basalähnlichen ER<sup>-</sup>-Tumoren gefunden (Sorlie et al., 2001). In der Studie von Levicar et al. werden die Cysteinproteasen Cathepsin B, L, sowie deren Inhibitoren Stefin A, B mit der Serin Protease u-PA und deren Inhibitor (PAI-1) bzgl. klinikopathologischer Daten verglichen (Levicar et al., 2002). Ein Ergebnis dieser Arbeit war, dass PAI-1, unabhängig von dem Cysteinprotease/-inhibitor-System, mit den meisten klinikopathologischen Daten korreliert. Auf der anderen Seite bleibt u-PA der biologische Marker mit größter Signifikanz, obwohl hohe Expression der Cathepsine B,L ebenfalls mit einem kürzeren Zeitraum zwischen Therapie und Tod bei lymphknoten-negativen Patientinnen verknüpft sind. Aufgrund dieser Fakten wird CSTB vermutlich nicht weiter in der Literatur beschrieben und ist somit als einzelnes Markergen nicht verwendbar. Beim hier untersuchten Gewebekollektiv konnte uPA bzw. PLAU (*plasminogen activator; urokinase*) nicht als differentiell exprimiertes Gen (adj. p-Wert >5%) detektiert werden.

*Tenascin-C* (TNC) ist ein ebenfalls mit hoher Signifikanz, in überwiegend ER<sup>-</sup>-negativen Tumoren exprimiertes Gen. Bei diesem Gen wird eine Beteiligung in der Angiogenese durch Regulation von VEGF (*vascular epithelial growth factor*) während der Tumorgenese, vermutet (Tanaka et al., 2004). Die Studie hat gezeigt, daß die Expression beider Gene durch das umgebende Mesenchym beeinflusst wird und damit diese Mikroumgebung allgemein sehr wichtig für die Angiogenese in der Tumorentwicklung ist. Weshalb im *Tumorscreen* nun *Tenascin-C* überwiegend in ER<sup>-</sup>-Gewebe stärker exprimiert wird und VEGF im Vergleich der beiden Gewebe überhaupt nicht als differentiell reguliert erscheint, konnte nicht geklärt werden. Im Vergleich zwischen Normal- und Tumorgewebe wird VEGF jedoch mit erhöhter Signifikanz stärker im Tumorgewebe exprimiert ( $p = 0,02$ , s. Tab. 16).

Ob damit *Tenascin-C* alleine als ein weiteres Markergen in der Tumorgenese zu betrachten ist, kann danach schwer beurteilt werden, zudem es auch in der Literatur keine weiteren Hinweise gibt.

Im Falle von VEGF erscheint es jedoch interessanter, den VEGF-Signaltransduktionsweg selbst als potentiell Target zu untersuchen, da es hier bereits einige Strategien gibt (Sauer et al., 2002, Belletti et al., 1999, Oku et al., 1998, Asano et al., 1999, Wang et al., 1998).

PLAT (*plasminogen activator tissue* oder auch tPA) erscheint als Gen mit zweitgrößter Signifikanz in ER<sup>+</sup>-Gewebe ( $p = 1,35 \times 10^{-3}$ ). PLAT ist eine Serin-Protease und in erster Linie bei der intravasulären Auflösung von Blutgerinnsel eingebunden. Neben der Tatsache, dass intratumorale PLAT-Pegel mit der Prognose einiger Tumore assoziiert sind, handelt es sich bei PLAT um ein östrogen-induzierbares Protein in humanen Brustkrebszelllinien (Corte et al., 2005). In der Studie von Corte et al. wurde PLAT hinsichtlich seines klinischen Nutzens untersucht. Dabei ergab sich, dass niedere intratumorale PLAT-Pegel mit Tumoraggressivität und einer schlechten Prognose für Brustkrebspatienten verknüpft sind, dies jedoch nicht ausreicht, um eine Vorhersage bzgl. dem Ansprechen einer systemisch, adjuvanten Therapie treffen zu können. Bereits zuvor haben Chappuis et al. in einem zymografischen *Assay* nachgewiesen, dass eine reduzierte PLAT-vermittelte Plasminproduktion, sich als ein unabhängiger, prognostischer Faktor ungünstig bei Brustkrebs auswirkt (Chappuis et al., 2001). Trotz dieser Tatsachen wird PLAT bisher noch nicht als etablierter Marker verwendet.

Ein weiteres Gen, welches im Vergleich zu ER<sup>+</sup>-Gewebe stärker in ER<sup>-</sup>-Gewebe exprimiert wird, ist ABCC1 (*ATP-binding cassette, sub-family C* oder auch MRP). ABCC1 ist ein Membranprotein und gehört zur Familie der Transporter. Im humanen System besteht diese Familie aus 49 Mitgliedern. Überexpression dieser Transporter, wie ABCB1 oder ABCC1, verursacht eine sog. *multidrug resistance* in der Krebsbehandlung (Liu et al., 2005). Inwieweit die Überexpression dieses einzelnen Transporters jedoch eine Aussage über das Ansprechen ein bestimmten Therapie ermöglicht, lässt sich schwer abschätzen, denn oftmals wird eine *multidrug resistance* durch die Expression zweier weiterer Gene, MDR 1 (*multidrug resistance gene 1*) und BCRP (*breast cancer resistance protein*) hervorgerufen (Wijnholds, 2002). Obgleich diese beiden Gene auch Bestandteil des Genkollektiv sind (s. Anhang), weisen ihre Expressionsdaten, bei dem hier untersuchten Gewebekollektiv, auf keine differentielle Expression (adj. p-Wert > 5%) hin.

MMP7 (*matrix metallo protease 7* oder *matrilysin*) ist wie PLAT und ABCC1 ein weiteres Gen aus der Übersichtsarbeit von Lacroix (Lacroix et al., 2002). Obwohl es gemeinsam mit MMP11 (*Stromelysin-3*) und MMP2 (*Gelatinase A*) verstärkt in kanzerösem Gewebe exprimiert wird, konnte MMP7 weder allein, noch im Verbindung mit den beiden anderen Metalloproteasen in Korrelation zu den klassisch prognostischen Markern (Alter, Tumorgröße, Steroidrezeptoren, Lymphknotenstatus etc.) gebracht werden (Pacheco et al., 1998). Interessanterweise tritt hier in der durchgeführten Transkriptionsanalyse lediglich MMP7 verstärkt in ER<sup>-</sup>-Gewebe auf ( $p = 1,45 \times 10^{-2}$ ). Es bleibt zu erwähnen, daß MMP7 bei Magenkrebs bzgl. lokalbeschränkter Invasivität als glaubwürdiger prognostischer Marker angesehen werden kann (Liu et al., 2002).

BZRP (*benzodiazapine receptor peripheral* oder auch PBR *peripheral-type benzodiazepin receptor*) erscheint im Mittelfeld der Transkriptionsanalyse und ist ebenfalls verstärkt im ER<sup>-</sup>-Gewebe exprimiert ( $p = 1,72 \times 10^{-2}$ ). Erhöhte BZRP-Pegel treten vor allem in Tumorgewebe auf und sind damit als ein neuer prognostischer Indikator hinsichtlich des aggressiven Phänotyps einzustufen (Han et al., 2003). Allerdings sind erhöhte BZRP-Werte nicht automatisch mit dem Aggressionspotential von Tumoren assoziiert. Die verstärkte Expression wurde zum überwiegenden Teil in Brust-, Kolorectal- und Prostatatumoren festgestellt und beschränkt sich damit lediglich auf wenige Krebsarten.

EIF4E (*eukaryotic translation initiation factor 4E*) wird gemäß der Transkriptionsanalyse verstärkt in ER<sup>+</sup>-Gewebe exprimiert ( $p = 2,35 \times 10^{-2}$ ). Dieses Gen gilt als Schlüsselregulator der malignen Progression und wird über MAP-Kinase- und PI3-Kinase-Signaltransduktionswege induziert (Zimmer et al., 2000). In einer neueren Studie wurde gezeigt, dass die Überexpression von EIF4E und TLKB1 (*tousled-like kinase 1B*) u.a. unabhängig vom Steroidrezeptor-Status ist (Norton et al., 2004). In diesem Zusammenhang bleibt unklar, weshalb bei den hier untersuchten Tumorproben EIF4E nur in ER<sup>+</sup>-Gewebe exprimiert wird.

SLC39A6 oder LIV-1 (*solute carrier family 39 member 6*) zur Unterfamilie der ZIP (*Zrt-, Irt-like proteins*) Zinktransporter gehörend, wurde in den hier untersuchten Geweben mit erhöhter Signifikanz (adj. p-Wert < 5%) verstärkt in ER<sup>+</sup>-Gewebe exprimiert. Neben der Assoziation mit ER<sup>+</sup>-Brusttumoren ist LIV-1 auch in die Metastasenverbreitung in der Nähe liegender Lymphknoten involviert (Taylor et al., 2003).

Aufgrund dieser Tatsachen wurde LIV-1 in dieser Studie als potentieller Marker hinsichtlich der Progression von Brustkrebs untersucht. Bekannt war zu dem Zeitpunkt, dass Zink essenziell für Wachstum von Bedeutung ist. In diesem Zusammenhang besitzen Zinktransporter eine wichtige Funktion in der Erhaltung des intrazellulären Zinkgleichgewichts. Aberrationen des Gleichgewichts können zu Erkrankungen wie Krebs führen. In einer Folgestudie wurde LIV-1 als eines der Gene identifiziert, welches zusammen mit den Transkriptionsfaktoren STAT3 (*signal transducer and activator of transcription 3*) und *Snail* in der EMT (*epithelial-mesenchymal transition*) bei der Embryogenese des Zebrafisches mitwirkt (Taylor et al., 2004). Generell ist die ETM ein essentielles Element in der Embryonalentwicklung, der Gewebeumgestaltung sowie in der Wundheilung. Inwieweit diese Erkenntnisse LIV-1 als einen weiteren potentiellen Marker bei Brustkrebs auszeichnen, lässt sich zum gegenwärtigen Zeitpunkt nicht abschätzen.

Zwei weitere Gene, die hier verstärkt in ER<sup>+</sup>-Gewebe vorkommen, sind TOP2A (*topoisomerase II alpha*) und CDH3 (*cadherin3* oder P-cadherin), obgleich in der Literatur beide Gene stets mit ER-Tumoren in Verbindung gebracht werden (Gruvberger et al., 2001, Kovacs et al., 2003, Paredes et al., 2004, Bofin et al., 2003). TOP2A ist ein molekulares Target von Doxorubicin, welches seit Jahren in der Therapie zur Behandlung von Brustkrebs eingesetzt wird (Lynch et al., 1997).

E- und P-Cadherine spielen eine wichtige Rolle in der Architektur und Funktion der normalen Brustdrüse, können jedoch durch Über- bzw. Unterexpression unterschiedliche Brustkrebstypen hervorbringen (Knudsen and Wheelock, 2005). Die genaue Rolle der hier genannten Cadherine sowie ihre abwärtsliegenden Signalereignisse bzgl. Entwicklung und malignem Verhalten von Brustkrebs erfordern jedoch weitergehende Forschung.

Mit dem hier ermittelten Genexpressionsprofil können beide Gewebegruppen durchaus auch unabhängig vom histochemisch bestimmten Hormonrezeptor-Status, eindeutig unterschieden werden. Die Expressionsdaten einzelner Gene zeigen ca. 80 % Übereinstimmung mit Literaturangaben, wo diese Gene verstärkt in dem einen oder anderen Gewebetyp exprimiert werden.

Mit den Genen, wie beispielsweise LIV-1, TOP2A oder LCN2, könnte die histologische Bestimmung vom Östrogenrezeptor-Status untermauert werden und in Kombination mit histopathologischen Daten eventuell veränderte Therapieschemata ermöglichen.

#### 4.3.2 Signifikante Gene innerhalb Normal - und Tumorewebe

Bereits in Kap. 3.5.2 wurden Gene kurz vorgestellt, die mit Signifikanz verstärkt in Tumorgewebe (z. B. Zytokeratine 8,18 und 19) exprimiert sind. Zwei Gene (IGF2, *insulin-like growth factor 2*, A2M *alpha-2macroglobulin*), die in der Studie von Gruvberger positiv mit dem Östrogenrezeptor korrelieren, bekommen bei der Auswertung hier eine negative t-Teststatistik, womit sie verstärkt in Normalgewebe exprimiert sind. Nach Angaben der Klinik wurden die hier untersuchten Normalgewebe als schwach Östrogenrezeptor-positiv eingestuft.



Gemäß den hier durchgeführten Transkriptionsanalysen ist das Gen VDR (*vitamin D receptor*) mit größter Signifikanz stärker im Normalgewebe exprimiert ( $p = 5,3 \times 10^{-6}$ ).

Der Vitamin D (3) Rezeptor ist ein ligand-abhängiger Transkriptionsfaktor, der in die Regulation des Zellzyklus, in Differenzierung und Apoptose von normalen sowie transformierten Brustzellen mit eingebunden ist (Zinser and Welsch., 2004).

Aus der Literatur ist weiter bekannt, dass VDR in Brust- und Eierstockkrebs stärker exprimiert ist als im entsprechenden Normalgewebe (Friedrich et al., 2003).

Im Normalgewebe dient die VDR-Expression als Gegenspieler der östrogen-gesteuerten Proliferation und ist damit für die Erhaltung der Differenzierung wichtig (Welsh et al., 2003). Obwohl VDR bei Brustkrebs als mögliches Target für therapeutische Vitamin D Analoga gilt, wird er nicht als prognostischer Faktor eingesetzt (Friedrich et al., 2002).

SPP1 (*secreted phosphoprotein* oder *osteopontin*) ist im hier untersuchten Tumorgewebe mit hoher Signifikanz ( $p = 2,12 \times 10^{-3}$ ) stärker als im entsprechenden Normalgewebe exprimiert. *Osteopontin*, ein integrin-bindendes Protein, ist u.a. auch in die Krebsentstehung verwickelt und spielt sowohl bei Progression der Brustkrebserkrankung, wie auch in der Entwicklung von normalen Brustdrüsen eine Rolle (Tuck and Chmabers., 2001). Diese Effekte werden teilweise durch spezifische Oberflächenintegrine vermittelt, die wiederum verstärkte Zellmigration, aber auch Aktivierung von Wachstumsfaktor/-rezeptor Signalwegen verursachen (Tuck et al., 2003). Bereits im Einleitungsteil der Arbeit wurde darauf verwiesen, dass o.g. Signaltransduktionswege neue molekulare Ziele darstellen können (Sauer et al., 2002).

Die Zytokeratine 8,18 und 19 treten als einzige Gruppe von Genen auf, die verstärkt im Tumorgewebe exprimiert wird. Die Keratine gehören zu den Intermediärfilamenten, mit denen Zellen epithelialen Ursprungs identifiziert werden können (Trask et al., 1990). Untersuchungen der Keratine 8, 18 und 19 zeigten übereinstimmende Expression hinsichtlich Tumorgrad, wobei Keratin 19 allein signifikant mit zunehmender Tumorgröße korreliert (Brotherick et al., 1998). Insbesondere die Quantifizierung von Keratin 19 mRNA-Pegeln haben sich nützlich bei der Stadienbestimmung von Krebs, der Krankheitsüberwachung und dem prognostischen Vergleich unter Brustkrebspatienten erwiesen (Wong et al., 2001). Ferner gilt Keratin 19 als ein Detektionsmarker mit höchster Sensivität für verborgene Tumorzellen, insbesondere bei operablem und metastasierendem Brustkrebs (Stathopoulou et al., 2003).

BBC3 (*BCL2 binding component 3*) oder auch PUMA ist verstärkt in Normalgewebe exprimiert ( $p = 9,65 \times 10^{-3}$ ). BBC3 ist ein proapoptotisches Gen und wird durch p53 induziert (Nakano and Vousden., 2001). Neben BBC3 wird auch CASP2 (*caspase-2*) stärker in Normalgewebe exprimiert ( $p = 1,15 \times 10^{-2}$ ). Beide Gene stehen nicht unmittelbar im Zusammenhang mit Brustkrebs, sondern haben größere Bedeutung in der Apoptose.

IGFBP4 (*Insulin-like growth factor-binding protein 4*) wird überraschenderweise in Normalgewebe stärker als im Tumorgewebe exprimiert. Dies steht vermutlich im Zusammenhang mit dem schwach Östrogenrezeptor-positiven Normalgewebe. Normalerweise ist IGFBP4 zusammen mit HDAC6 (histone deacetylase 6) ein Östrogen-reguliertes Gen, die bei den ER<sup>+</sup>-Patientinnen, diejenigen mit schlechter Prognose kennzeichnen (Yoshida et al., 2004).

MIF (*macrophage migration inhibitory factor*) gehört zu den Zytokinen und ist in die zellvermittelte Immunität, die Immunregulation sowie in Entzündungsprozessen eingebunden. Im Zusammenhang mit Brustkrebs scheint MIF eine Rolle bei den Interaktionen von Tumor und Stroma zu spielen. Sowohl in intratumoralen Konzentrationen wie auch als Serummarker korreliert MIF invers mit befallenen Lymphknoten (Bando et al., 2002).

Mit dieser Tatsache könnte MIF zusätzlich zur Ermittlung des Lymphknotenstatus bei primärem Brustkrebs herangezogen werden.

KIT (*v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homologue*) tritt stärker im Normal- als im Tumorgewebe auf. Dieses Protoonkogen zeigt sich beständig in Normalgewebe, jedoch mit variabler Expression innerhalb von Brustkrebsgruppen. Dies legt die Vermutung nahe, dass KIT eine Rolle in der Tumorprogression spielt und damit diagnostische, prognostische und therapeutische Auswirkungen hat (Ko et al., 2003, Yared et al., 2004).

Des weiteren dienen KIT- und EGFR-Überexpression zur Unterscheidung hochgradiger Brusttumore mit mesenchymaler und/oder myoepithelialer Differenzierung von denjenigen mit erbB2-Überexpression (Tsuda et al., 2005).

BSG (*basigin* oder auch *EMMPRIN*) ist ein multifunktionelles Transmembran Glykoprotein und spielt u.a. bei Entzündungsprozessen und Tumorinvasion eine Rolle (Muramatsu and Miyauchi, 2003). Im Zusammenhang mit Tumor-Stroma-Interaktionen induziert BSG die Produktion von Metalloproteasen, die eine Schlüsselrolle in der Tumorprogression spielen (Tang et al., 2004).

Anhand der hier vorgestellten, differentiell exprimierten Gene ist es durchaus möglich, Normalgewebe von Tumorgewebe zu unterscheiden u.a. auch, da sich die Genexpression mit Daten aus der Literatur deckt. Allerdings ist es bei einem solchen Vergleich wichtig, die Zahl der Normalgewebe, soweit verfügbar, deutlich zu erhöhen. Bei 21 Tumor- zu 3 Normalgeweben erscheinen die Normalgewebe eher unterrepräsentiert. Hinzu kommt noch, dass eines der Normalgewebe (N1) lediglich mittlere Reproduzierbarkeit zeigte (vgl. Boxplot S. 59) und die Normalgewebe zudem noch als schwach Östrogenrezeptor-positiv einzustufen sind.

#### 4.4 Ausblick

Mit dem hier entwickelten DNA-Microarray war die Unterscheidung von 4 Gewebegruppen (ER<sup>+</sup>-, ER<sup>-</sup>-, Normal- und Tumorgewebe) prinzipiell möglich. Von dieser Plattform ausgehend, ergeben sich einige Aspekte, die es weiter zu untersuchen gilt.

Das Wichtigste erscheint hier die Bestätigung der als differentiell ausgewiesenen Gene durch Sequenzierung sowie deren Überprüfung mittels unabhängiger Methoden, wie es bereits in Kap. 3.4.2 an ausgesuchten Genen mit der quantitativen *real-time* PCR erfolgt ist.

Anschließend wäre es von Seiten der Kliniker interessant, ob sich eines oder eine Gruppe dieser Gene zusätzlich zu den klinikopathologischen Daten in die Tumoranalyse miteinbeziehen lässt, um dadurch verbesserte Aussagen hinsichtlich Diagnose oder Therapierbarkeit zu erreichen.

Zur Untersuchung von erbB2<sup>+</sup>- und erbB2<sup>-</sup>-Tumoren sollten auf jeden Fall auch die entsprechenden Statusgene (erbB2, GRB7 und STARD3) sowie damit verbundene Sequenzen in das Genkollektiv mit aufgenommen werden. In diesem Zusammenhang erscheint es auch notwendig, die Genanzahl zu überprüfen. Lt. Aussagen der Statistiker lagen bei der Untersuchung dieses Tumorkollektivs etwa 30 % aller Gene ständig im Hintergrund. Daraus ergibt sich die Überlegung, ob diese Gene überhaupt weiter benötigt werden.

Ein weiterer Aspekt ist die Etablierung der linearen RNA-Amplifikation im Hinblick auf Transkriptionsanalysen aus Biopsiematerial. In der vorliegenden Arbeit wurde lediglich mit einem Kit gearbeitet. Allerdings sollte hier auch ein Methodenvergleich zwischen einem Kit und einem etablierten Protokoll durchgeführt werden, um möglichst geringe Verzerrungen im Vergleich zur Originalprobe zu erreichen.

Obgleich die DNA-Microarray-Technologie das Methodenrepertoire vielerorts bereits erweitert hat, sollte sich der Anwender stets bewußt sein, hier in erster Linie ein sog. *Screening Tool* in der Hand zu haben, um sich damit einen allerersten Überblick zu verschaffen.

Derzeit werden themenspezifische DNA-Microarrays für unterschiedliche Tumortypen von mehreren Herstellern (u.a. Affymetrix, Eppendorf, Biomol) angeboten. Allerdings sind diese Arrays weder klinisch validiert, noch finden sie in der klinischen Routine Anwendung.

Nachfolgend sollen weitere Techniken aufgezeigt werden, die ebenfalls in der Tumordiagnostik genutzt werden können.

Eine parallele Entwicklung zur DNA-Microarray-Technologie sind PCR-basierende Methoden. Die quantitative *real time*-PCR ist hier die Methode der Wahl. Die Vorteile dieser Methode gegenüber der DNA-Microarray Technologie liegen bei einem wesentlich geringeren Bedarf an Probenmaterial, ihrer hohen Sensitivität sowie ihrer sehr guten Reproduzierbarkeit. Allerdings ist die Methode nicht für *Screening*-Experimente geeignet, da mit ihr nur eine begrenzte Anzahl von Genen untersucht werden kann. Um dieses Hindernis zu umgehen bieten einige Firmen (z. B. Roche Diagnostics) die quantitative *real time*-PCR auch im sog. Mikrotiterplattenformat an. Als Anwendung für so einen „plattenbasierten“ Assay bietet z. Zt. die Firma Genomichealth mit Oncotype DX<sup>TM</sup> ein Testsystem für Brustkrebs an, das bereits klinisch validiert (Paik et al., 2004) und von der FDA (*Food and Drug Administration*) zugelassen ist.

Oncotype DX<sup>TM</sup> ist ein auf RT-PCR basierendes Nachweissystem, mit dem anhand von individuellen Tumorprofilen Aussagen bzgl. verbesserter Behandlungsmöglichkeiten sowie über krankheitsbedingte Rückfälle bei Brustkrebserkrankungen, gemacht werden können (weitere Informationen unter <http://www.genomichealth.com/oncotype/>).

Auch Techniken wie beispielsweise *Comperative genomic hybridisation* (CGH) und *single nucleotide polymorphism* (SNP) eignen sich zur Analyse von Brustkrebs (Espinosa et al., 2006). Die CGH erlaubt eine simultane Erfassung unterschiedlicher genetischer Veränderungen, wie beispielsweise Imbalancen der Kopienanzahl von Chromosomen oder von Amplifikationen bzw. Deletionen. Üblicherweise werden in einem CGH Experiment die relativen Hybridisierungsintensitäten zweier gesamtgenomischer DNA Populationen (z. B. Normal vs. Tumor) verglichen. Dabei wird angenommen, dass die gemessenen Unterschiede proportional zur veränderten Kopienzahl der Chromosomen sind. Mit der array-basierten CHG ist es möglich das gesamte Genom hochauflösend, d.h. in Abschnitten von 40 -150 kb, zu untersuchen. Das Genom wird dabei über Sonden, wie Oligonukleotide, BAC- (*bacterial artificial chromosome*) oder cDNA-Klone repräsentiert (Pinkel and Albertson, 2005).

SNPs sind Variationen von einzelnen Basenpaaren im DNA-Strang und kommen etwa bei 1% der Bevölkerung vor (Syvänen, 2005). Ferner stellen SNPs die häufigste genetische Variation im humanen Genom dar. SNP-Analysen können sowohl array-basiert (z. B. Gene Chips<sup>TM</sup>, Affymetrix) oder im Mikrotiterplattenformat (z. B. Genestream<sup>TM</sup>, Beckman Coulter) durchgeführt werden. Ihre Anwendung findet die SNP-Analyse sowohl in Genotypisierungstudien von Krankheiten (Hirschhorn and Daly, 2005) oder zur Identifikation von krankheits-bezogenen Genen (Benusiglio et al., 2005, Hoyal et al., 2005).

Neben diesen genombasierten Ansätzen kommen auch immer mehr Methoden der Proteomik zum Einsatz. Nachfolgend wird stellvertretend eine Technik genannt und kurz beschrieben. Mit der Massenspektroskopie ist es möglich tausende von Proteinen aus komplexen Gemischen zu identifizieren und auch zu quantifizieren (Aebersold and Mann, 2003). Ein Massenspektrometer trennt Proteine nach ihrem Masse-Ladungszahl Verhältnis (*m/z ratio*). Prinzipiell besteht ein Massenspektrometer aus drei Einheiten: Einer Ionisierungseinheit, einem Massenanalysator sowie einem Detektor. Die Ionisierung kann über verschiedene Techniken erfolgen. Gewöhnlich finden in der Biologie zwei Ionisierungstechniken Anwendung: Die MALDI-TOF-MS (*matrix-assisted laser-desorption/ionization time-of-flight mass-spectrometry*) sowie die SELDI-MS (*surface enhanced laser desorption/ionization mass spectroscopy*), eine Abwandlung der MALDI-TOF-MS. Der Unterschied zwischen den beiden Verfahren besteht in der chemischen Modifikation der Probeneller beim SELDI-MS. Damit eignet sich das SELDI-MS auch zur Detektion von Serumproteinen. Die Massenspektroskopie zeichnet sich besonders durch kurze Analysezeiten, hohe Sensitivität und der Möglichkeit zur Automatisierung aus. In einer kürzlich veröffentlichten Studie wurde die SELDI-MS zur Diagnose von Brustkrebs eingesetzt (Li et al., 2005).

Welche der hier vorgestellten Techniken sich als Instrument für maßgeschneiderte Therapiekonzepte behaupten werden, bleibt offen, da zukünftig die Kostenexplosion im Gesundheitswesen eine immer größere Bedeutung für die zu ergreifenden Behandlungsmaßnahmen gewinnt. Denkbar wäre jedoch, dass die Kombination dieser Techniken neue Zusammenhänge bzw. Erkenntnisse in der Krebsforschung ergeben könnten.

**Literaturverzeichnis**

Aebersold R, Mann M

Mass spectrometry-based proteomics  
Nature. 2003 Mar 13;422(6928):198-207

Ahr A, Holtrich U, Solbach C, Scharl A, Strebhardt K, Karn T, Kaufmann M.  
Molecular classification of breast cancer patients by gene expression profiling.  
J Pathol. 2001;195:312-20

Alizadeh AA, Eisen MB, Davis RE, Ma C, Lossos IS, Rosenwald A, Boldrick JC,  
Sabet H, Tran T, Yu X, Powell JI, Yang L, Marti GE, Moore T, Hudson J Jr, Lu L,  
Lewis DB, Tibshirani R, Sherlock G, Chan WC, Greiner TC, Weisenburger DD,  
Armitage JO, Warnke R, Levy R, Wilson W, Grever MR, Byrd JC, Botstein D,  
Brown PO, Staudt LM  
Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling  
Nature 2000;403:503-11

Asano M, Yukita A, Suzuki H

Wide spectrum of antitumor activity of a neutralizing monoclonal antibody to human  
vascular endothelial growth factor.  
Jpn J Cancer Res 1999;90:93-100

Bando H, Matsumoto G, Bando M, Muta M, Ogawa T, Funata N, Nishihira J,  
Koike M, Toi M

Expression of macrophage migration inhibitory factor in human breast cancer:  
association with nodal spread  
Jpn J Cancer Res 2002;93:389-96

Baron V, Calleja V, Ferrari P, Alengrin F, Van Obberghen E

p125Fak focal adhesion kinase is a substrate for the insulin and insulin-like growth  
factor-I tyrosine kinase receptors  
J Biol Chem 1998;273:7162-8

Beckmann MW, Niederacher D, Massenkeil G, Tutschek B, Beckmann A, Schenko G,  
Schnurch HG, Bender HG

Expression analyses of epidermal growth factor receptor and HER-2/neu: no advantage of  
prediction of recurrence or survival in breast cancer patients  
Oncology 1996;53:441-7

Beckmann MW, Niederacher D, Schnurch HG, Gusterson BA, Bender HG

Multistep carcinogenesis of breast cancer and tumour heterogeneity  
J Mol Med 1997;75:429-39

Beenken SW, Bland KI

Biomarkers for breast cancer  
Minerva Chir 2002;57:437-48

Belletti B, Ferraro P, Arra C, Baldassarre G, Bruni P, Staibano S, De Rosa G, Salvatore G, Fusco A, Persico MG, Viglietto G

Modulation of in vivo growth of thyroid tumor-derived cell lines by sense and antisense vascular endothelial growth factor gene

Oncogene 1999;18:4860-9

Benusiglio PR, Lesueur F, Luccarini C, Conroy DM, Shah M, Easton DF, Day NE, Dunning AM, Pharoah PD, Ponder BA

Common ERBB2 polymorphisms and risk of breast cancer in a white British population: a case-control study

Breast Cancer Res 2005;7:R204-8

Bertucci F, Houlgatte R, Granjeaud S, Nasser V, Loriod B, Beaudoin E, Hingamp P, Jacquemier J, Viens P, Birnbaum D, Nguyen C

Prognosis of breast cancer and gene expression profiling using DNA arrays

Ann N Y Acad Sci 2002;975:217-31

Bertucci F, Viens P, Hingamp P, Nasser V, Houlgatte R, Birnbaum D

Breast cancer revisited using DNA array-based gene expression profiling

Int J Cancer 2003;103:565-71

Bofin AM, Ytterhus B, Hagmar BM

TOP2A and HER-2 gene amplification in fine needle aspirates from breast carcinomas

Cytopathology 2003;14:314-9

Borresen-Dale AL

Genetic profiling of breast cancer: from molecular portraits to clinical utility

Int J Biol Markers 2003;18:54-6

Boyd, N.F. (Einl.)

The epidemiology of breast cancer

in Oxford Textbook of Oncology Oxford University Press 2. Aufl. 2002

Brotherick I, Robson CN, Browell DA, Shenfine J, White MD, Cunliffe WJ, Shenton BK, Egan M, Webb LA, Lunt LG, Young JR, Higgs MJ

Cytokeratin expression in breast cancer: phenotypic changes associated with disease progression

Cytometry 1998;32:301-8

Bucci B, D'Agnano I, Botti C, Mottolese M, Carico E, Zupi G, Vecchione A

EGF-R expression in ductal breast cancer: proliferation and prognostic implications

Anticancer Res 1997;17:769-74

Carlomagno C, Perrone F, Gallo C, De Laurentiis M, Lauria R, Morabito A, Pettinato G, Panico L, D'Antonio A, Bianco AR, De Placido S

c-erb B2 overexpression decreases the benefit of adjuvant tamoxifen in early-stage breast cancer without axillary lymph node metastases

J Clin Oncol. 1996;14:2702-8

Carter CL, Allen C, Henson DE

Relation of tumor size, lymph node status, and survival in 24,740 breast cancer cases

Cancer 1989;63:181-7

Chappuis PO, Dieterich B, Sciretta V, Lohse C, Bonnefoi H, Remadi S, Sappino AP

Functional evaluation of plasmin formation in primary breast cancer

J Clin Oncol 2001;19:2731-8

Cho KR, Vogelstein B

Genetic alterations in the adenoma-carcinoma sequence

Cancer 1992;70:1727-31

Corte MD, Verez P, Rodriguez JC, Roibas A, Dominguez ML, Lamelas ML,

Vazquez J, Garcia Muniz JL, Allende MT, Gonzalez LO, Fueyo A, Vizoso F

Tissue-type plasminogen activator (tPA) in breast cancer: relationship with clinicopathological parameters and prognostic significance

Breast Cancer Res Treat 2005;90:33-40

Cunliffe HE, Ringner M, Bilke S, Walker RL, Cheung JM, Chen Y, Meltzer PS

The gene expression response of breast cancer to growth regulators: patterns and correlation with tumor expression profiles

Cancer Res 2003;63:7158-66

Davidoff AM, Kerns BJ, Pence JC, Marks JR, Iglehart JD

p53 alterations in all stages of breast cancer

J Surg Oncol 1991;48:260-7

Di Leo A, Cardoso F, Durbecq V, Giuliani R, Mano M, Atalay G, Larsimont D,

Sotiriou C, Biganzoli L, Piccart MJ

Predictive molecular markers in the adjuvant therapy of breast cancer: state of the art in the year 2002

Int J Clin Oncol 2002;7:245-53

Doerr ME, Jones JI

The roles of integrins and extracellular matrix proteins in the insulin-like growth factor

I-stimulated chemotaxis of human breast cancer cells

J Biol Chem 1996;271:2443-7

- Dressman MA, Walz TM, Lavedan C, Barnes L, Buchholtz S, Kwon I, Ellis MJ, Polymeropoulos MH  
Genes that co-cluster with estrogen receptor alpha in microarray analysis of breast biopsies  
Pharmacogenomics J 2001;1:135-41
- Dunican DS, McWilliam P, Tighe O, Parle-McDermott A, Croke DT  
Gene expression differences between the microsatellite instability (MIN) and chromosomal instability (CIN) phenotypes in colorectal cancer revealed by highdensity cDNA array hybridization  
Oncogene 2002;21:3253
- Eisen MB., Brown PO  
DNA-Arrays for Analysis of Gene Expression  
Methods in Enzymology 1999; 303:179-217
- el-Ashry D, Lippman ME  
Molecular biology of breast carcinoma  
World J Surg 1994;18:12-20
- el-Tanani MK, Green CD  
Transcription factor, BTF3, and the AF-1 function of the estrogen receptor  
Biochem Soc Trans 1998;26:S252
- Espinosa E, Redondo A, Vara JA, Zamora P, Casado E, Cejas P, Baron MG  
High-throughput techniques in breast cancer: A clinical perspective  
Eur J Cancer 2006;42:598-60
- Faneyte IF, Schrama JG, Peterse JL, Remijnse PL, Rodenhuis S, van de Vijver MJ  
Breast cancer response to neoadjuvant chemotherapy: predictive markers and relation with outcome  
Br J Cancer 2003;88:406-12
- Fearon ER, Vogelstein B  
A genetic model for colorectal tumorigenesis  
Cell. 1990;61:759-67
- Fitzpatrick SL, Brightwell J, Wittliff JL, Barrows GH, Schultz GS  
Epidermal growth factor binding by breast tumor biopsies and relationship to estrogen receptor and progestin receptor levels  
Cancer Res 1984;44:3448-53
- Folgueira MA, Carraro DM, Brentani H, Patrao DF, Barbosa EM, Netto MM, Caldeira JR, Katayama ML, Soares FA, Oliveira CT, Reis LF, Kaiano JH, Camargo LP, Vencio RZ, Snitcovsky IM, Makdissi FB, e Silva PJ, Goes JC, Brentani MM  
Gene expression profile associated with response to doxorubicin-based therapy in breast cancer  
Clin Cancer Res 2005;11:7434-43



- Friedrich M, Villena-Heinsen C, Tilgen W, Schmidt W, Reichrat J, Axt-Flidner R  
Vitamin D receptor (VDR) expression is not a prognostic factor in breast cancer  
Anticancer Res 2002;22:1919-24
- Friedrich M, Rafi L, Mitschele T, Tilgen W, Schmidt W, Reichrath J  
Analysis of the vitamin D system in cervical carcinomas, breast cancer and ovarian cancer  
Recent Results Cancer Res 2003;164:239-46
- Fujita T, Kobayashi Y, Wada O, Tateishi Y, Kitada L, Yamamoto Y, Takashima H,  
Murayama A, Yano T, Baba T, Kato S, Kawabe Y, Yanagisawa J  
Full activation of estrogen receptor alpha activation function-1 induces proliferation  
of breast cancer cells  
J Biol Chem 2003;278:26704-14
- Gillesby BE, Zacharewski TR  
pS2 (TFF1) levels in human breast cancer tumor samples: correlation with clinical  
and histological prognostic markers  
Breast Cancer Res Treat 1999;56:253-65
- Giuliano AE.  
Current status of sentinel lymphadenectomy in breast cancer  
Ann Surg Oncol 2001;8(9 Suppl):52S-55S
- Golub TR, Slonim DK, Tamayo P, Huard C, Gaasenbeek M, Mesirov JP, Coller H,  
Loh ML, Downing JR, Caligiuri MA, Bloomfield CD, Lander ES  
Molecular classification of cancer: class discovery and class prediction by gene  
expression monitoring  
Science 1999;286:531-7
- Gruvberger S, Ringner M, Chen Y, Panavally S, Saal LH, Borg A, Ferno M,  
Peterson C, Meltzer PS  
Estrogen receptor status in breast cancer is associated with remarkably distinct gene  
expression patterns  
Cancer Res 2001;61:5979-84
- Gullick WJ, Love SB, Wright C, Barnes DM, Gusterson B, Harris AL, Altman DG  
erbB-2 protein overexpression in breast cancer is a risk factor in patients with  
involved and uninvolved lymph nodes  
Br J Cancer 1991;63:434-8
- Gyorffy B, Serra V, Jurchott K, Abdul-Ghani R, Garber M, Stein U, Petersen I, Lage H,  
Dietel M, Schafer R  
Prediction of doxorubicin sensitivity in breast tumors based on gene expression profiles of  
drug-resistant cell lines correlates with patient survival  
Oncogene 2005;24:7542-51

- Hampton GM, Frierson HF  
Classifying human cancer by analysis of gene expression  
Trends Mol Med 2003;9:5-10
- Han Z, Slack RS, Li W, Papadopoulos V  
Expression of peripheral benzodiazepine receptor (PBR) in human tumors:  
relationship to breast, colorectal, and prostate tumor progression  
J Recept Signal Transduct Res 2003;23:225-38
- Heer K, Kumar H, Read JR, Fox JN, Monson JR, Kerin MJ  
Serum vascular endothelial growth factor in breast cancer: its relation with cancer type and  
estrogen receptor status  
Clin Cancer Res 2001;7:3491-4
- Heintz NH, Leslie KO, Rogers LA, Howard PL  
Amplification of the c-erb B-2 oncogene and prognosis of breast adenocarcinoma  
Arch Pathol Lab Med 1990;114:160-3
- Hermanek P, Wittekind C, Seib H  
Current TNM aspects 5. TNM classification and stage grouping  
Zentralbl Chir 1997;122:934-7
- Hirschhorn JN, Daly MJ  
Genome-wide association studies for common diseases and complex traits  
Nat Rev Genet 2005;6:95-108
- Hopp TA, Weiss HL, Hilsenbeck SG, Cui Y, Allred DC, Horwitz KB, Fuqua SA  
Breast cancer patients with progesterone receptor PR-A-rich tumors have poorer disease-  
free survival rates  
Clin Cancer Res 2004;10:2751-60
- Hoyal CR, Kammerer S, Roth RB, Reneland R, Marnellos G, Kiechle M, Schwarz-Boeger  
U, Griffiths LR, Ebner F, Rehbock J, Nelson MR, Braun A  
Genetic polymorphisms in DPF3 associated with risk of breast cancer and lymph node  
metastases  
J Carcinog 2005;4:13
- Inoue A, Yoshida N, Omoto Y, Oguchi S, Yamori T, Kiyama R, Hayashi S  
Development of cDNA microarray for expression profiling of estrogen-responsive genes  
J Mol Endocrinol. 2002 Oct;29(2):175-92
- Irizarry RA, Warren D, Spencer F, Kim IF, Biswal S, Frank BC, Gabrielson E, Garcia JG,  
Geoghegan J, Germino G, Griffin C, Hilmer SC, Hoffman E, Jedlicka AE, Kawasaki E,  
Martinez-Murillo F, Morsberger L, Lee H, Petersen D, Quackenbush J, Scott A, Wilson M,  
Yang Y, Ye SQ, Yu W  
Multiple-laboratory comparison of microarray platforms  
Nat Methods 2005;2:345-50

Iyer VR, Eisen MB, Ross DT, Schuler G, Moore T, Lee JC, Trent JM, Staudt LM, Hudson J Jr, Boguski MS, Lashkari D, Shalon D, Botstein D, Brown PO  
The transcriptional program in the response of human fibroblasts to serum  
Science 1999;283:83-7

Jeffrey SS, Fero MJ, Borresen-Dale AL, Botstein D  
Expression array technology in the diagnosis and treatment of breast cancer  
Mol Interv 2002;2:101-9

Junqueira L., Carneiro J., Kelley R.  
Histologie  
Springer Verlag Berlin Heidelberg New York 4. Aufl. 1996

Kato S  
Function of estrogen receptor (ER) in gene expression  
Jpn J Clin Oncol 1999;29:321-2

Kenzelmann M, Klaren R, Hergenbahn M, Bonrouhi M, Grone HJ, Schmid W, Schutz G  
High-accuracy amplification of nanogram total RNA amounts for gene profiling  
Genomics 2004;83:550-8

Khan J, Bittner ML, Chen Y, Meltzer PS, Trent JM  
DNA microarray technology: the anticipated impact on the study of human disease  
Biochim Biophys Acta 1999;1423:M17-28

Khan J, Wei JS, Ringner M, Saal LH, Ladanyi M, Westermann F, Berthold F,  
Schwab M, Antonescu CR, Peterson C, Meltzer PS  
Classification and diagnostic prediction of cancers using gene expression profiling and  
artificial neural networks  
Nat Med 200;7:673-9

Klijn JG, Berns PM, Schmitz PI, Foekens JA  
The clinical significance of epidermal growth factor receptor (EGF-R) in human breast  
cancer: a review on 5232 patients  
Endocr Rev 1992;13:3-17

Knabbe, C  
Steroide und Wachstumsfaktoren in der Proliferationskontrolle von Mammakarzinomen  
Akt. Onkol. Zuckschwerdt Verlag München 1997; 100: 1-4

Knudsen KA, Wheelock MJ  
Cadherins and the mammary gland  
J Cell Biochem. 2005;95:488-96

Ko CD, Kim JS, Ko BG, Son BH, Kang HJ, Yoon HS, Cho EY, Gong G, Ahn SH  
The meaning of the c-kit proto-oncogene product in malignant transformation in human mammary epithelium  
Clin Exp Metastasis 2003;20:593-7

Kononen J, Bubendorf L, Kallioniemi A, Barlund M, Schraml P, Leighton S, Torhorst J, Mihatsch MJ, Sauter G, Kallioniemi OP  
Tissue microarrays for high-throughput molecular profiling of tumor specimens  
Nat Med 1998;4:844-7

Kovacs A, Dhillon J, Walker RA  
Expression of P-cadherin, but not E-cadherin or N-cadherin, relates to pathological and functional differentiation of breast carcinomas  
Mol Pathol 2003;56:318-22

Kudoh K, Ramanna M, Ravatn R, Elkahloun AG, Bittner ML, Meltzer PS  
Trent JM, Dalton WS, Chin KV  
Monitoring the expression profiles of doxorubicin-induced and doxorubicin-resistant cancer cells by cDNA microarray  
Cancer Res 2000;60:4161-6

Lacroix M, Zammateo N, Remacle J, Leclercq G  
A low-density DNA microarray for analysis of markers in breast cancer  
Int J Biol Markers 2002;17:5-23

Lee JE, Chung KW, Han W, Kim SW, Kim SW, Shin HJ, Bae JY, Noh DY  
Effect of estrogen, tamoxifen and epidermal growth factor on the transcriptional regulation of vascular endothelial growth factor in breast cancer cells  
Anticancer Res 2004;24:3961-4

Levicar N, Kos J, Blejec A, Golouh R, Vrhovec I, Frkovic-Grazio S, Lah TT  
Comparison of potential biological markers cathepsin B, cathepsin L, stefin A and stefin B with urokinase and plasminogen activator inhibitor-1 and clinicopathological data of breast carcinoma patients  
Cancer Detect Prev 2002;26:42-9

Li J, Zhang Z, Rosenzweig J, Wang YY, Chan DW  
Proteomics and bioinformatics approaches for identification of serum biomarkers to detect breast cancer  
Clin Chem 2002;48:1296-304

Li J, Orlandi R, White CN, Rosenzweig J, Zhao J, Seregini E, Morelli D, Yu Y, Meng XY, Zhang Z, Davidson NE, Fung ET, Chan DW  
Independent validation of candidate breast cancer serum biomarkers identified by mass spectrometry  
Clin Chem 2005;51:2229-35

- Linke SP, Bremer TM, Herold CD, Sauter G, Diamond C  
A multimarker model to predict outcome in tamoxifen-treated breast cancer patients  
Clin Cancer Res 2006;12(4):1175-83
- Liu XP, Kawauchi S, Oga A, Tsushimi K, Tsushimi M, Furuya T, Sasaki K  
Prognostic significance of matrix metalloproteinase-7 (MMP-7) expression at the  
invasive front in gastric carcinoma  
Jpn J Cancer Res 2002;93:291-5
- Liu Y, Peng H, Zhang JT.  
Expression profiling of ABC transporters in a drug resistant breast cancer cell line  
using AmpArray  
Mol Pharmacol 2005 May 18
- de Longueville F, Lacroix M, Barbuto AM, Bertholet V, Gallo D, Larsimont D, Marcq L,  
Zammatteo N, Boffe S, Leclercq G, Remacle J  
Molecular characterization of breast cancer cell lines by a low-density microarray  
Int J Oncol. 2005;27:881-92
- Lui VW, Grandis JR  
EGFR-mediated cell cycle regulation  
Anticancer Res 2002;22:1-11
- Lynch BJ, Guinee DG Jr, Holden JA  
Human DNA topoisomerase II-alpha: a new marker of cell proliferation in invasive  
breast cancer  
Hum Pathol 1997;28:1180-8
- MacGregor JI, Jordan VC  
Basic guide to the mechanisms of antiestrogen action  
Pharmacol Rev 1998;50:151-96
- Marshall E  
Getting the noise out of gene arrays  
Science 2004;306:630-1
- McMasters KM, Giuliano AE, Ross MI, Reintgen DS, Hunt KK, Byrd DR,  
Klimberg VS, Whitworth PW, Tafra LC, Edwards MJ  
Sentinel-lymph-node biopsy for breast cancer--not yet the standard of care  
N Engl J Med 1998;339:990-5
- McPherson K, Steel CM, Dixon JM  
ABC of breast diseases. Breast cancer-epidemiology, risk factors, and genetics  
BMJ 2000;321:624-8

Mitra I

Prognostic Factors and Staging

in Oxford Textbook of Oncology Oxford University Press 2. Aufl. 2002

Muramatsu T, Miyauchi T

Basigin (CD147): a multifunctional transmembrane protein involved in reproduction, neural function, inflammation and tumor invasion

Histol Histopathol 2003;18:981-7

Musgrove EA, Lilischkis R, Cornish AL, Lee CS, Setlur V, Seshadri R, Sutherland RL

Expression of the cyclin-dependent kinase inhibitors p16INK4, p15INK4B and p21WAF1/CIP1 in human breast cancer

Int J Cancer 1995;63:584-91

Nagahata T, Onda M, Emi M, Nagai H, Tsumagari K, Fujimoto T, Hirano A, Sato T,

Nishikawa K, Akiyama F, Sakamoto G, Kasumi F, Miki Y, Tanaka T, Tsunoda T

Expression profiling to predict postoperative prognosis for estrogen receptor-negative breast cancers by analysis of 25,344 genes on a cDNA microarray

Cancer Sci 2004;95:218-25

Nakano K, Vousden KH

PUMA, a novel proapoptotic gene, is induced by p53

Mol Cell 2001;7:683-94

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;45:2917-24

Norton KS, McClusky D, Sen S, Yu H, Meschonat C, Debenedetti A, Li BD

TLK1B is elevated with eIF4E overexpression in breast cancer

J Surg Res 2004;116:98-103

Oh DS, Troester MA, Usary J, Hu Z, He X, Fan C, Wu J, Carey LA, Perou CM

Estrogen-Regulated Genes Predict Survival in Hormone Receptor-Positive Breast Cancers

J Clin Oncol 2006 Feb 27

Oku T, Tjuvajev JG, Miyagawa T, Sasajima T, Joshi A, Joshi R, Finn R, Claffey KP, Blasberg RG

Tumor growth modulation by sense and antisense vascular endothelial growth factor gene expression: effects on angiogenesis, vascular permeability, blood volume, blood flow, fluorodeoxyglucose uptake, and proliferation of human melanoma intracerebral xenografts

Cancer Res 1998;58:4185-92

- Pacheco MM, Mourao M, Mantovani EB, Nishimoto IN, Brentani MM  
Expression of gelatinases A and B, stromelysin-3 and matrilysin genes in breast carcinomas: clinico-pathological correlations  
Clin Exp Metastasis. 1998;16:577-85
- Paik S, Shak S, Tang G, Kim C, Baker J, Cronin M, Baehner FL, Walker MG, Watson D, Park T, Hiller W, Fisher ER, Wickerham DL, Bryant J, Wolmark N  
A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer  
N Engl J Med 2004;351:2817-26
- Paredes J, Stove C, Stove V, Milanezi F, Van Marck V, Derycke L, Mareel M, Bracke M, Schmitt F  
P-cadherin is up-regulated by the antiestrogen ICI 182,780 and promotes invasion of human breast cancer cells  
Cancer Res 2004;64:8309-17
- Perks CM, Gill ZP, Newcomb PV, Holly JM  
Activation of integrin and ceramide signalling pathways can inhibit the mitogenic effect of insulin-like growth factor I (IGF-I) in human breast cancer cell lines  
Br J Cancer 1999;79:701-6
- Perou CM, Jeffrey SS, van de Rijn M, Rees CA, Eisen MB, Ross DT, Pergamenschikov A, Williams CF, Zhu SX, Lee JC, Lashkari D, Shalon D, Brown PO, Botstein D  
Distinctive gene expression patterns in human mammary epithelial cells and breast cancers  
Proc Natl Acad Sci U S A. 1999 Aug 3;96(16):9212-7
- Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, Pollack JR, Ross DT, Johnsen H, Akslen LA, Fluge O, Pergamenschikov A, Williams C, Zhu SX, Lonning PE, Borresen-Dale AL, Brown PO, Botstein D  
Molecular portraits of human breast tumours  
Nature 2000;406:747-52
- Perren TJ  
c-erbB-2 oncogene as a prognostic marker in breast cancer  
Br J Cancer 1991;63:328-32
- Riede U.-N., Schaefer H.-E.  
Allgemeine und spezielle Pathologie  
Thieme Verlag Stuttgart New York 4. Aufl 1999 aktualisierter Nachdruck
- Rouzier R, Perou CM, Symmans WF, Ibrahim N, Cristofanilli M, Anderson K, Hess KR, Stec J, Ayers M, Wagner P, Morandi P, Fan C, Rabiul I, Ross JS, Hortobagyi GN, Pusztai L  
Breast cancer molecular subtypes respond differently to preoperative chemotherapy  
Clin Cancer Res 2005;11:5678-85

- Sainsbury JR, Farndon JR, Harris AL, Sherbet GV.  
Epidermal growth factor receptors on human breast cancers  
Br J Surg 1985;72(3):186-8
- Sauer G, Deissler H, Kurzeder C, Kreienberg R  
New molecular targets of breast cancer therapy  
Strahlenther Onkol. 2002 Mar;178(3):123-33
- Sawiris GP, Sherman-Baust CA, Becker KG, Cheadle C, Teichberg D, Morin PJ  
Development of a highly specialized cDNA array for the study and diagnosis of  
epithelial ovarian cancer  
Cancer Res 2002;62:2923-8
- Schraml P, Kononen J, Bubendorf L, Moch H, Bissig H, Nocito A, Mihatsch MJ,  
Kallioniemi OP, Sauter G  
Tissue microarrays for gene amplification surveys in many different tumor types  
Clin Cancer Res 1999 ;5:1966-75
- Shih C, Padhy LC, Murray M, Weinberg RA  
Transforming genes of carcinomas and neuroblastomas introduced into mouse mouse  
fibroblasts  
Nature 1981;290:261-4
- Simon RM, Korn EL, McShane L, Radmacher MD, Wright GW, Zhao Y,  
Design and analysis of DNA-Microarray investigations  
Springer Verlag Berlin Heidelberg New York 2003
- Slamon DJ, Leyland-Jones B, Shak S, Fuchs H, Paton V, Bajamonde A, Fleming T,  
Eiermann W, Wolter J, Pegram M, Baselga J, Norton L  
Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast  
cancer that overexpresses HER2  
N Engl J Med 2001;344(11):783-92
- Sorlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, Hastie T, Eisen MB,  
van de Rijn M, Jeffrey SS, Thorsen T, Quist H, Matese JC, Brown PO, Botstein D,  
Eystein Lonning P, Borresen-Dale AL  
Gene expression patterns of breast carcinomas distinguish tumor subclasses with  
clinical implications  
Proc Natl Acad Sci U S A 2001;98:10869-74
- Sotiriou C, Powles TJ, Dowsett M, Jazaeri AA, Feldman AL, Assersohn L,  
Gadisetti C, Libutti SK, Liu ET  
Gene expression profiles derived from fine needle aspiration correlate with response  
to systemic chemotherapy in breast cancer  
Breast Cancer Res 2002;4:R3



Sotiriou C, Wirapati P, Loi S, Harris A, Fox S, Smeds J, Nordgren H, Farmer P, Praz V, Haibe-Kains B, Desmedt C, Larsimont D, Cardoso F, Peterse H, Nuyten D, Buyse M, Van de Vijver MJ, Bergh J, Piccart M, Delorenzi M

Gene expression profiling in breast cancer: understanding the molecular basis of histologic grade to improve prognosis

J Natl Cancer Inst. 2006;98:262-72

Sporn MB, Roberts AB

Autocrine growth factors and cancer

Nature 1985 Feb 28-Mar 6;313(6005):745-7

Stathopoulou A, Mavroudis D, Perraki M, Apostolaki S, Vlachonikolis I, Lianidou E, Georgoulas V.

Molecular detection of cancer cells in the peripheral blood of patients with breast cancer: comparison of CK-19, CEA and maspin as detection markers

Anticancer Res 2003;23:1883-90

Stoesz SP, Friedl A, Haag JD, Lindstrom MJ, Clark GM, Gould MN

Heterogeneous expression of the lipocalin NGAL in primary breast cancers

Int J Cancer. 1998;79:565-72

Strobel SA, Doudna JA

RNA seeing double: close-packing of helices in RNA tertiary structure

Trends Biochem Sci 1997;22:262-6

Syvanen AC

Toward genome-wide SNP genotyping

Nat Genet 2005;37 Suppl:S5-10

Tan PK, Downey TJ, Spitznagel EL Jr, Xu P, Fu D, Dimitrov DS, Lempicki RA, Raaka BM, Cam MC

Evaluation of gene expression measurements from commercial microarray platforms

Nucleic Acids Res 2003;31:5676-84

Tanaka K, Hiraiwa N, Hashimoto H, Yamazaki Y, Kusakabe M

Tenascin-C regulates angiogenesis in tumor through the regulation of vascular endothelial growth factor expression

Int J Cancer 2004;108:31-40

Tang Y, Kesavan P, Nakada MT, Yan L.

Tumor-stroma interaction: positive feedback regulation of extracellular matrix metalloproteinase inducer (EMMPRIN) expression and matrix metalloproteinase-dependent generation of soluble EMMPRIN

Mol Cancer Res 2004;2:73-80

- Taylor KM, Morgan HE, Johnson A, Hadley LJ, Nicholson RI  
Structure-function analysis of LIV-1, the breast cancer-associated protein that belongs to a new subfamily of zinc transporters  
Biochem J 2003;375:51-9
- Taylor KM, Hiscox S, Nicholson RI  
Zinc transporter LIV-1: a link between cellular development and cancer progression  
Trends Endocrinol Metab 2004;15:461-3
- Therapieleitlinie Mammakarzinom des onkologischen Schwerpunkts (OSP) Stuttgart e. V.  
Arbeitsgruppe Mammakarzinom erstellt März 2002
- Trask DK, Band V, Zajchowski DA, Yaswen P, Suh T, Sager R  
Keratins as markers that distinguish normal and tumor-derived mammary epithelial cells  
Proc Natl Acad Sci U S A 1990;87:2319-2
- Tsuda H, Morita D, Kimura M, Shinto E, Ohtsuka Y, Matsubara O, Inazawa J, Tamaki K, Mochizuki H, Tamai S, Hiraide H.  
Correlation of KIT and EGFR overexpression with invasive ductal breast carcinoma of the solid-tubular subtype, nuclear grade 3, and mesenchymal or myoepithelial differentiation  
Cancer Sci. 2005;96:48-53
- Tuck AB, Chambers AF  
The role of osteopontin in breast cancer: clinical and experimental studies  
J Mammary Gland Biol Neoplasia 2001;6:419-29
- Tuck AB, Hota C, Wilson SM, Chambers AF  
Osteopontin-induced migration of human mammary epithelial cells involves activation of EGF receptor and multiple signal transduction pathways  
Oncogene 2003;22(8):1198-205
- Turbin DA, Cheang MC, Bajdik CD, Gelmon KA, Yorida E, De Luca A, Nielsen TO, Huntsman DG, Gilks CB  
MDM2 protein expression is a negative prognostic marker in breast carcinoma  
Mod Pathol 2006;19:69-74
- Ullrich A, Schlessinger J.  
Signal transduction by receptors with tyrosine kinase activity  
Cell 1990 Apr 20;61:203-12
- van de Vijver MJ, He YD, van't Veer LJ, Dai H, Hart AA, Voskuil DW, Schreiber GJ, Peterse JL, Roberts C, Marton MJ, Parrish M, Atsma D, Witteveen A, Glas A, Delahaye L, van der Velde T, Bartelink H, Rodenhuis S, Rutgers ET, Friend SH, Bernards R.  
A gene-expression signature as a predictor of survival in breast cancer  
N Engl J Med. 2002;347:1999-2009

- van 't Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AA, Bernards R, Friend SH  
Expression profiling predicts outcome in breast cancer  
Breast Cancer Res 2003;5:57-8
- Vlahou A, Laronga C, Wilson L, Gregory B, Fournier K, McGaughey D, Perry RR, Wright GL Jr, Semmes OJ  
A novel approach toward development of a rapid blood test for breast cancer  
Clin Breast Cancer 2003 Aug;4:203-9
- Vogel CL, Cobleigh MA, Tripathy D, Gutheil JC, Harris LN, Fehrenbacher L, Slamon DJ, Murphy M, Novotny WF, Burchmore M, Shak S, Stewart SJ  
First-line Herceptin monotherapy in metastatic breast cancer  
Oncology 2001;61 Suppl 2:37-42
- Vogelstein B, Fearon ER, Kern SE, Hamilton SR, Preisinger AC, Nakamura Y, White R  
Allelotype of colorectal carcinomas  
Science 1989;244:207-11
- Vuori K, Ruoslahti E  
Association of insulin receptor substrate-1 with integrins  
Science 1994;266:1576-8
- Wagner C., 1999  
Molekulare Onkologie  
Thieme Verlag Stuttgart New York 2. Aufl. 1999
- Walker RA, Dearing SJ  
Expression of epidermal growth factor receptor mRNA and protein in primary breast carcinomas  
Breast Cancer Res Treat 1999;53:167-76
- Wang G, Dong Z, Xu G, Yang Z, Shou C, Wang N, Liu T  
The effect of antibody against vascular endothelial growth factor on tumor growth and metastasis  
J Cancer Res Clin Oncol 1998;124:615-20
- Welsh J, Wietzke JA, Zinser GM, Byrne B, Smith K, Narvaez CJ.  
Vitamin D-3 receptor as a target for breast cancer prevention  
J Nutr 2003;133(7 Suppl):2425S-2433S
- West M, Blanchette C, Dressman H, Huang E, Ishida S, Spang R, Zuzan H, Olson JA, Marks JR, Nevins JR  
Predicting the clinical status of human breast cancer by using gene expression profiles  
Proc Natl Acad Sci U S A 2001;98:11462-7

Wijnholds J

Drug resistance caused by multidrug resistance-associated proteins

Novartis Found Symp 2002;243:69-79

Williams KJ, Loeb LA

Retroviral reverse transcriptases: error frequencies and mutagenesis

Curr Top Microbiol Immunol 1992;176:165-80

Wong IH, Yeo W, Chan AT, Johnson PJ

Quantitative correlation of cytokeratin 19 mRNA level in peripheral blood with disease stage and metastasis in breast cancer patients: potential prognostic implications

Int J Oncol 2001;18:633-8

Yared MA, Middleton LP, Meric F, Cristofanilli M, Sahin AA

Expression of c-kit proto-oncogene product in breast tissue

Breast J. 2004;10(4):323-7

Yee D, Cullen KJ, Paik S, Perdue JF, Hampton B, Schwartz A, Lippman ME, Rosen N

Insulin-like growth factor II mRNA expression in human breast cancer

Cancer Res 1988;48:6691-6

Yee D, Paik S, Lebovic GS, Marcus RR, Favoni RE, Cullen KJ, Lippman ME, Rosen N

Analysis of insulin-like growth factor I gene expression in malignancy: evidence for a paracrine role in human breast cancer

Mol Endocrinol 1989;3:509-17

Yoshida N, Omoto Y, Inoue A, Eguchi H, Kobayashi Y, Kurosumi M, Saji S, Suemasu K, Okazaki T, Nakachi K, Fujita T, Hayashi S

Prediction of prognosis of estrogen receptor-positive breast cancer with combination of selected estrogen-regulated genes

Cancer Sci 2004;95:496-502

Zajchowski DA, Bartholdi MF, Gong Y, Webster L, Liu HL, Munishkin A, Beauheim C, Harvey S, Ethier SP, Johnson PH

Identification of gene expression profiles that predict the aggressive behavior of breast cancer cells

Cancer Res 2001;61:5168-78

Zhang X, Yee D

Tyrosine kinase signalling in breast cancer: insulin-like growth factors and their receptors in breast cancer

Breast Cancer Res 2000;2:170-5

Zimmer SG, DeBenedetti A, Graff JR.

Translational control of malignancy: the mRNA cap-binding protein, eIF-4E, as a central regulator of tumor formation, growth, invasion and metastasis.

Anticancer Res 2000;20:1343-5

Zinser GM, Welsh J.  
Vitamin D receptor status alters mammary gland morphology and tumorigenesis in  
MMTV-neu mice  
Carcinogenesis 2004;25:2361-72

Zugmaier, G., Knabbe, C., Aigner, A  
Das Mammakarzinom – Risikofaktoren, Prognose und medikamentöse Therapie  
Med Welt 2003; 54: 96-101

**Veröffentlichungen**

I. Brändlin, S. Hübner, T. Eiseler, M. Martinez-Moya , A. Horschinek , A. Hausser, G. Link, S. Rupp, P. Storz, K. Pfizenmeier, F.-J. Johannes.  
Protein kinase C (PKC)  $\eta$ -mediated PKC $\mu$  activation modulates ERK and JNK signal pathways

J. Biol. Chem. **8**, 6490 – 6497 (2002)

## Lebenslauf

### Persönliche Daten

Name: Andreas Horschinek

Geburtsdatum/-ort: 07.04.72, Leonberg

Familienstand: ledig

### Ausbildungsdaten

Schulbildung: 09/80 - 06/84 Grundschule in Korntal  
09/84 - 06/89 Realschule in Korntal  
09/89 - 05/92 Technisches Gymnasium Max-Eyth-Schule Stuttgart;  
Abschluß: Allgemeine Hochschulreife

Wehrdienst: 07/92 – 06/93 Grundwehrdienst bei JaboG 32 in Landsberg/Lech

Studium: 10/93 - 09/95 Studium der Verfahrenstechnik Universität Stuttgart;  
10/95 - 01/01 Studium der technischen Biologie, Universität Stuttgart  
Abschluß: Diplom-Biologe (t.o.)  
Diplomarbeit am Fraunhofer-Institut für Grenzflächen-  
und Bioverfahrenstechnik (IGB), Stuttgart  
Betreuer: PD Dr. Franz-Josef Johannes

Promotion 02/01 - 08/06 Dissertation am Fraunhofer-Institut für Grenzflächen-  
und Bioverfahrenstechnik (IGB), Stuttgart  
Betreuer: PD Dr. Franz-Josef Johannes †  
PD Dr. Steffen Rupp  
Prof. Dr. Herwig Brunner

Berufspraxis ab 08/05 Anstellung am Institut für Pathologie,  
Universität Lübeck, Prof. Dr. med A.C. Feller,  
Bereich molekulare Diagnostik

## Abbildungsverzeichnis

Abb. 1: Drüsen und Gangsystem der weiblichen Brust	S. 2
Abb. 2: HE-Färbungen verschiedener Mammakarzinomformen	S. 2
Abb. 3: Auswertungsschema der Rohdaten mit unterschiedlichen Möglichkeiten der Datenanalyse	S. 28
Abb. 4: Microarray Hybridisierungsbilder	S. 31
Abb. 5: Scatterplot aus den Hybridisierungsexperimenten mit Proben aus Gesamt RNA (50µg)	S. 32
Abb. 6: RNA-Gel, 1%, sybrgold-gefärbt	S. 35
Abb. 7: Prinzip der linearen RNA-Amplifikation in 5 Teilschritten	S. 36
Abb. 8: Scatter Plots aus den Hybridisierungsexperimenten mit Proben aus amplifizierter RNA (1µg)	S. 38
Abb. 9: Darstellung der Anzahl detektierbarer Gene zwischen Gesamt und amplifizierter RNA	S. 39
Abb. 10: Scatter Plot zwischen den Experimenten mit Gesamt und amplifizierter RNA bzgl. auf die Schnittmenge in Abb. 9	S. 39
Abb. 11: Auszug der Schnittmenge (15 von 45 Genen) aus Abb. 9 mit Darstellung der Signalintensitäten und log-ratios	S. 40
Abb. 12: Differentiell exprimierte Gene zwischen unstimulierten und mit ICI182780 stimulierten MCF7 Zellen	S. 47
Abb. 13: Vergleich der <i>log-ratios</i> zwischen IGB- und Eppendorf-Array	S. 48
Abb. 14: Signalintensitäten und Expressionsverhältnisse von <i>real-time</i> PCR nach Antiöstrogen-Stimulation (n=3)	S. 49
Abb. 15: Signalintensitäten des IGB-Chip nach Antiöstrogen-Stimulation (n=9)	S. 49
Abb. 16: cDNA-Synthese Überprüfung im 1 % Agarosegel	S. 51
Abb. 17: Dot Blot zum Nachweis der cDNA-Synthese in unterschiedlichen Verdünnungsstufen	S. 51
Abb. 18: Einzelhybridisierungsbilder in Falschfarbendarstellung	S. 53
Abb. 19: Boxplots zur Darstellung der Signalhintergrundverhältnisse der unterschiedlichen Antikörper aufgetrennt in die jeweiligen Farbkanäle Cyanine 5 und 3	S. 54
Abb. 20: Qualitätsüberprüfung der Hybridisierungen für Tumore T1 - T6	S. 58
Abb. 21: Qualitätsüberprüfung der Hybridisierungen für Tumore T7 - T12	S. 58
Abb. 22: Qualitätsüberprüfung der Hybridisierungen für Tumore T13 - T18	S. 59
Abb. 23: Qualitätsüberprüfung der Hybridisierungen für Tumore T19 - T21 und Normalgewebe N1 - N3	S. 59
Abb. 24: Darstellung des Gen-Overlap der Array-Plattformen mit weiterer Aufteilung	S. 60
Abb. 25: Beispiele des Vergleichs gemeinsamer Sequenzen mit hoher Korrelation	S. 63
Abb. 26: Beispiele des Vergleichs gemeinsamer Sequenzen mit niedriger Korrelation ( $\geq 0$ )	S. 64



## Tabellenverzeichnis

Tab. 1: Zusammenfassung prognostischer Faktoren bei Brusttumoren	S. 7
Tab. 2: Primerübersicht zur Amplifikation aus den verschiedenen Vektoren	S. 16
Tab. 3: Qualitätsüberprüfung der durchgeführten Experimente mit Gesamt-RNA anhand der Korrelationen zwischen den Arrays	S. 31
Tab. 4: Klassifikator-Gene der 170 differentiell exprimierten Gene aus den Transkriptionsprofilen der beiden Zelllinien	S. 33
Tab. 5: Ergebnisse der linearen RNA-Amplifikation unterschiedlicher Gesamt-RNA Ausgangsmengen	S. 35
Tab. 6: Qualitätsüberprüfung der durchgeführten Experimente mit amplifizierter RNA anhand der Korrelationen zwischen den Arrays	S. 37
Tab. 7: Korrelationen zwischen den dye-swap Experimenten, sowie zwischen zwei parallelen Amplifikationen und aus dem Vergleich zwischen cDNA aus Gesamt und amplifizierter RNA	S. 38
Tab. 8: Auflistung der durch Amplifikation detektierbaren Gene	S. 41
Tab. 9: Auflistung der detektierbaren Gene ohne vorausgegangene Amplifikation	S. 43
Tab. 10: Differentielle Gene zwischen un- und mit ICI187780 stimulierten MCF7 Zellen	S. 46
Tab. 11: Expressionsvergleich zwischen IGB-Array und RPA	S. 48
Tab. 12: Monohybridisierungen an ausgesuchtem Gewebekollektiv	S. 52
Tab. 13: Kompetitive Hybridisierungen mit ausgesuchtem Gewebekollektiv und Zellpool als Referenz	S. 53
Tab. 14: Histopathologische Details zu den untersuchten Geweben	S. 56
Tab. 15: Differentiell exprimierte Gene zwischen ER- und Er+ Gewebe unter jeweiliger Einbeziehung der Referenzproben	S. 60
Tab. 16: Differentielle Genexpression zwischen Tumor und Normalgewebe unter jeweiliger Einbeziehung der Referenzproben	S. 61

## Aktuelle Genliste

Genname	Gene ID	Beschreibung	RZPD No
A2M	NM_000014	Alpha-2-macroglobulin	HU3_p972C06273D8
ABCB1	NM_000927	P-glycoprotein, Multidrugresistance protein 1 (MDR1)	HU3_p972A03120D8
ABCC1	NM_004996	Multidrug resistance-associated protein (MRP, MRP-1)	IMAGp958B041158Q2
ABCG2 (ABCP)	NM_002748	Breast cancer resistance protein (BCRP), Placenta-specific ATP-binding cassette transporter (ABCP)	HU3_p972D0882D8
ACADSB	NM_001609	Acyl-Coenzyme A dehydrogenase, short/branched chain	HU3_p972A0760D
ACTA	NM_001100	Actin alpha	HU3_p972A0988D
ACTB	NM_001101	Actin beta	IMAGp998D09782
ACTG1*	NM_001614	Actin gamma 1	IMAGp998I041725
ACVRL1*	NM_000020	Activin A receptor II-like	IMAGp950E1054
ADFP	NM_001122	Adipose differentiation-related protein	HU3_p972G1097D
ADRM1	NM_007002	Adhesion regulating molecule 1	HU3_p972H11273D
ADSS	NM_001126	Adenylosuccinate synthase	HU3_p972A09101D8
AKT1	NM_005163	Akt-1, PKB, RAC, PRKBA, RAC-ALPHA	IMAGp958A011667Q2
ALDH1A1	NM_000689	aldehyd dehydrogenase 1A1	HU3_p972A0388D8
ALDH4A1	NM_003748	Aldehyde dehydrogenase 4 family, member A1	HU3_p972A106D
ANXA2	NM_004039	Annexin A2	DKFZp686D04105Q2
AOX1	NM_001159	Aldehyde oxidase 1	IMAGp958A151583Q2
AP1S2	NM_003916	Adaptor-related protein complex 1, sigma 2 subunit	HU3_p972A0378D
AP2B1	NM_001282	Adaptor-related protein complex2, beta 1 subunit	HU3_p972F0739D
AP3S2	NM_005829	Adaptor-related protein complex 3, sigma 2 subunit	HU3_p972A0717D8
APC	NM_000038	Adenomatosis polyposis coli	HU3_p972H11268D
APOD	NM_001647	Apolipoprotein D	HU3_p972E0620D8
APPBP2	NM_006380	Amyloid beta precursor protein (cytoplasmic tail) binding protein 2	HU3_p972F10141D
AR	NM_000044	Antrogen receptor	HU3_p972H12234D
ARHB	NM_004040	Ras homolog gene family, member B	HU3_p972H05162D
ARHH	NM_004310	Ras homolog gene family, member H	HU3_p972A0690D8
ASS	NM_054012	Arginiosuccinate synthetase	HU3_p972D0562D
ATF3*	NM_001674	Activating transcription factor 3	IMAGp958A041577Q2
ATF3	NM_004024	Activating transcription factor 3	IMAGp998I095718Q2
ATF4	NM_001675	Activating transcription factor 4	HU3_p972H05289D
ATM	NM_000051	Ataxia telangiectasia mutated	DKFZp686F17258Q2
B3GNT4	NM_030765	UDP-GlcNa:betaGal beta-1,3-N-acetylglucosaminyltransferase4	HU3_p972D07137D8
B4GALT1	NM_001497	UDP-Gal:betaGlcNAc beta 1,4- galactosyltransferase, polypeptide 1	IMAGp958A092529Q2
BACE2	NM_012105	Beta-site APP-cleaving enzyme 2	HU3_p972E01147D
BAG1	NM_004323	BCL-2-associated athanogene	IMAGp958A15821Q2
BAK1*	NM_001188	BCL2-antagonist/killer 1	IMAGp998I23465Q2
BAK1	NM_001188	BCL-2-antagonis/killer 1	HU3_p972F0961D8
BARD1*	NM_000465	BRCA1 associated RING domain 1	IMAGp998H05531
BAX	NM_004324	BCL2-associated X protein	HU3_p972D04189D8
BAZ2A	NM_013449	Bromodomain adjacent to zinc finger domain, 2A	DKFZp686A2292Q2

Tab. 17: Die Tabelle enthält neben dem abgekürzten Gennamen, die Referenz-mRNA Nr., den vollen Gennamen sowie die Klon-Bestellnummer (RZPD No). Mit \* gekennzeichnete Klone sind sequenzbestätigt.

Genname	GeneID	Beschreibung	RZPD No
BBC3	NM_014417	Bcl-2 binding component 3	IMAGp998E242692Q3
BCAR1	NM_014567	Breast cancer anti-estrogen resistance 1, p130Cas adaptor protein	HU3_p972F0795D8
BCAR3	NM_003567	Breast cancer anti-estrogen resistance 3	IMAGp958B02761Q2
BCAS1	NM_003657	Breast carcinoma amplified sequence 1	HU3_p972A06276D
BCAS2	NM_005872	Breast carcinoma amplified sequence 2	HU3_p972F0391D
BCKDHA	NM_000709	Branched chain keto acid dehydrogenase E1, alpha polypeptide (maple syrup urine disease)	HU3_p972D0449D
BCL2	NM_000633	B-cell CLL/lymphoma 2	HU3_p972B0227D8
BCL2A1	NM_004049	BCL2-related protein A1	IMAGp958A171067Q2
BCL2L1	NM_001191	Bcl-2-like 1	HU3_p972B07121D8
BCL2L10	NM_020396	BCL2-like 10	HU3_p972G01280D8
BCL2L2	NM_004050	BCL2-like 2	IMAGp958D05987Q2
BECN1	NM_003766	Bcl-2-interacting protein beclin-1	HU3_p972B067D8
BID	NM_004049	BH3 interacting domain death agonist	IMAGp998F10286Q3
BIN1	NM_009668	Bridging integrator 1	IMAGp958A181767Q2
BIRC1	NM_004536	Baculoviral IAP repeat-containing 1	HU3_p972E11294D8
BIRC2	NM_001166	Baculoviral IAP repeat-containing 2	HU3_p972F04125D8
BIRC3	NM_001165	Baculoviral IAP repeat-containing 3	HU3_p972D0349D
BIRC4	NM_001167	Baculoviral IAP repeat-containing 4	HU3_p972F0976D8
BIRC5	NM_001168	Baculoviral IAP repeat-containing 5 (survivin)	HU3_p972F07150D
BPGM	NM_001724	2,3-bisphosphoglycerate mutase	HU3_p972G1131D
BRCA1	NM_007274	Breast cancer 1, early onset	HU3_p972B0264D8
BRCA2	NM_000059	Breast cancer 2, early onset	HU3_p972G0443D8
BRMS1L	NM_032352	Breast cancer metastasis-suppressor 1-like	HU3_p972H06146D
BRRN1	NM_015341	Barren homolog (Drosophila) // KIAA0074 protein	HU3_p972C0846D
BSG	NM_001728	Basigin (OK blood group)	IMAGp958A012512Q2
BTF3	NM_001207	Basic transcription factor 3	IMAGp958A031490Q2
BTG2	NM_006763	BTG family, member 2	HU3_p972A05278D
BTG3	NM_006806	BTG family, member 3	HU3_p972E0563D
BUB1	NM_004336	Budding uninhibited by benzimidazoles 1 homolog	HU3_p972E05140D
BZRP	NM_000714	Benzodiazapine receptor (peripheral)	HU3_p972H12147D
C17orf28	NM_030630	Chromosome 17 open reading frame 28	HU3_p972H01130D
CA1	NM_001738	Carbonic Anhydrase I	IMAGp998B074743Q3
CA2	NM_000067	Carbonic Anhydrase II	IMAGp958C1481Q2
CA9	NM_001216	Carbonic anhydrase-9	HU3_p972G08322D8
CACNA1C	NM_000719	Calcium channel, voltage-depend, L-type, alpha 1C subunit	IMAGp998D061933Q3
CALD1	NM_033138	Caldesmon 1	HU3_p972C1226D
CALML3	NM_005185	Calmodulin like 3	IMAGp998A085279Q3
CARD10	NM_014550	Caspase recruitment domain family, member 10	HU3_p972G0928D8
CARD12	NM_021209	Caspase recruitment domain family, member 12	HU3_p972B11218D8
CARD15	NM_022166	Caspase recruitment domain family, member 15	HU3_p972E12107D8
CARD8	NM_014959	Caspase recruitment domain family, member 8	HU3_p972G104D8
CARD9	NM_022352	Caspase recruitment domain family, member 9	HU3_p972A05324D8
CASP1	NM_032292	Caspase-1	HU3_p972C077D8
CASP10	NM_001230	Caspase-10	IMAGp998H18113Q3

Genname	GeneID	Beschreibung	RZPD No
CASP2	NM_032982	Caspase-2	IMAGp998F242013Q3
CASP3	NM_004346	Caspase-3	HU3_p972C04110D8
CASP4	NM_001225	Caspase-4	IMAGp958A202462Q2
CASP5	NM_004347	Caspase-5	IMAGp998F04780Q3
CASP6	NM_032992	Caspase-6	HU3_p972C0194D8
CASP7	NM_001227	Caspase-7	IMAGp958G102623Q2
CASP8	NM_032992	Caspase-8	HU3_p972C0194D8
CASP8	NM_001228	Caspase-8	IMAGp998P21595Q3
CASP9	NM_001229	Caspase-9	HU3_p972C0810D8
CAV1	NM_001753	Caveolin 1	HU3_p972B0268D
CBFB	NM_001755	Core-binding factor, $\beta$ subunit	HU3_p972B0592D
CCNA1	NM_003914	Cyclin A1	IMAGp958N0692Q2
CCNA2*	NM_001237	Cyclin A2	IMAGp998P222037Q2
CCNA2	NM_001237	Cyclin A2	HU3_p972B01108D8
CCNB1*	NM_031966	Cyclin B1	IMAGp998P171890Q2
CCND1	NM_053056	Cyclin D1 (PRAD1: parathyroid adenomatosis 1)	IMAGp958A111279Q2
CCND3	NM_001760	Cyclin D3	HU3_p972A11130D
CCNE1	NM_001238	Cyclin E1	HU3_p972A08101D
CCNE2	NM_004702	Cyclin E2	HU3_p972F1063D
CCNH	NM_001239	Cyclin H	HU3_p972A1227D
CCT4	NM_006430	Chaperonin containing TCP1, subunit 4 (delta)	HU3_p972C01275D
CD34	NM_001773	CD34 antigen	HU3_p972G10116D
CD9	NM_001769	CD9 antigen (p24)	IMAGp958A081775Q2
CDC20	NM_001255	CDC20, cell division cycle 20 homolog (S. cerevisiae)	HU3_p972B0619D
CDC25A	NM_001789	Cell division cycle 25A	HU3_p972G043D8
CDC25B*	NM_021874	Cell division cycle 25B	IMAGp998E13114
CDC25B	NM_021874	Cell division cycle 25B	HU3_p972A08297D
CDC25C	NM_001790	Cell division cycle 25C	IMAGp958N061713Q2
CDC27	NM_001256	Cell division cycle 27	HU3_p972D052D8
CDC42BPA	NM_003607	CDC42 binding protein kinase alpha DMPK-like	HU3_p972F12275D
CDC42EP1	NM_007061	CDC42 effector protein 1	HU3_p972D0434D8
CDCA7	NM_031942	Cell division associated 7	HU3_p972A0849D
CDH1	NM_004360	Cadherin 1, type 1, E-cadherin (epithelial)	HU3_p972G0844D8
CDH11	NM_033664	Cadherin 11, type 2, OB-cadherin (osteoblas)	HU3_p972B0214D
CDH13	NM_001257	Cadherin 13, H-cadherin (heart)	HU3_p972G10294D
CDH15	NM_004933	Cadherin 15, M-cadherin (myotubule)	HU3_p972G0896D
CDH2*	NM_001792	Cadherin 2	IMAGp998C07737Q2
CDH3	NM_001793	Cadherin 3, type 1, P-cadherin (placental)	IMAGp958A11812Q2
CDH5	NM_001795	Cadherin 5, type 2, VE-cadherin (vascular epithelium)	HU3_p972H12292D
CDK2	NM_001798	Cyclin-dependent kinase 2	IMAGp958A112053Q2
CDK4	NM_000075	Cyclin-dependent kinase 4	HU3_p972A03270D
CDK9	NM_001261	Cyclin-dependent kinase 9 (CDC2-related kinase)	HU3_p972H045D8
CDKN1A	NM_000389	Cyclin-dependent kinase inhibitor 1A, p21/waf1/cip1	IMAGp958A05179Q2
CDKN1B*	NM_004064	Cyclin-dependent kinase inhibitor 1B, p27/kip1	IMAGp998O201484Q2
CDKN1B	NM_004064	Cyclin-dependent kinase inhibitor 1B, p27/kip1	HU3_p972A04106D8

Gennamen	GeneID	Beschreibung	RZPD No
CDKN1C	NM_000076	Cyclin-dependent kinase inhibitor 1C, p57/waf2	IMAGp958F16163Q2
CDKN1C*	NM_000076	Cyclin-dependent kinase inhibitor 1C, p57/waf2	IMAGp998F13194Q2
CDKN2A	NM_000077	Cyclin-dependent kinase inhibitor 2A, p16/ink4/mts1	HU3_p972B09317D8
CDKN2C	NM_001262	Cyclin-dependent kinase inhibitor 2C	HU3_p972C1081D
CENPA	NM_001809	Centromere protein A (17kD)	HU3_p972G10130D
CENPE	NM_001813	Centromere protein E, 312kDa	HU3_p972D0161D
CFLAR (FLIP)	NM_003879	CASP8 and FADD-like apoptosis regulator	HU3_p972D01265D8
CH3L1	NM_001276	Chitinase 3-like 1 (cartilage glycoprotein-39)	IMAGp958F142633Q2
CIRBP	NM_001280	Cold inducible RNA-binding protein	IMAGp958A047Q2
CITED2	NM_006079	Cbp/p300-interacting transactivator, with Glu/Asp-rich carboxy terminal	IMAGp998A084417Q3
CLDN1	NM_021101	Claudin 1	IMAGp958A1848Q2
CNAP1	NM_014865	Chromosome condensation-related SMC-associated protein1	HU3_p972D05280D
CNN1	NM_001299	Calponin 1, basic smooth muscle	DKFZp686N1615Q2
COL4A1	NM_001845	Collagen, type IV, alpha 1	DKFZp434M1215Q2
COL4A2	NM_001846	Collagen, type IV, alpha 2, chain	HU3_p972E0865D
COL7A1	NM_000094	Collagen, type VII, alpha 1	HU3_p972D0898D
CORO1C	NM_014325	Coronin, actin binding protein, 1c	HU3_p972H1068D
COTL1	NM_021149	Coactosin-like 1 (Dictyostelium) // gelsolin amyloidosis, Finish type	HU3_p972E1181D
COX6C	NM_004374	Cytochrome c oxidase subunit VIc	HU3_p972H011D8
CPB1	NM_001871	Carboxypeptidase B1	IMAGp958B07816Q2
CRABP2	NM_001878	Cellular retinoic acid binding protein	HU3_p972F0932D8
CRMP1	NM_001313	Collapsin response mediator 1	HU3_p972H0140D8
CRYAB	NM_001885	Crystallin, aB	HU3_p972F03124D
CSDE1	NM_007158	Cold shock domain containing E1, RNA-binding	HU3_p972E0947D8
CSE1L	NM_001316	Chromosome segregation 1-like	HU3_p972B0710D
CSF1	NM_000757	Colony stimulating factor 1 (macrophage)	HU3_p972F1010D
CSF1R	NM_005211	Colony stimulating factor 1 receptor, formerly McDonough feline sarcoma viral (v-fms) oncogene homolog	HU3_p972G0752D
CSPG2	NM_004385	Chondroitin sulfate proteoglycan 2 (versican)	HU3_p972C0648D
CST3	NM_000099	Cystatin C	HU3_p972E10262D
CST6	NM_001323	Cystatin E/M	HU3_p972A0587D
CSTA	NM_005213	Cystatin A (stefin A)	HU3_p972D0177D
CSTB	NM_000100	Cystatin B (stefin)	HU3_p972A105D
CTGF	NM_001901	Connective tissue growth factor	HU3_p972C1271D
CTNNA1	NM_001903	Catenin (cadherin-associated protein), alpha 1	HU3_p972F0488D
CTNNB1	NM_001904	Catenin beta 1	IMAGp958B031495Q2
CTSC	NM_001814	Cathepsin C	HU3_p972B1027D
CTSD*	NM_001909	Cathepsin D	IMAGp998G02229Q3
CTSS	NM_004079	Cathepsin S	HU3_p972G1118D
CUTL1	NM_181552	Cut-like 1, CCAAT displacement protein (Drosophila)	HU3_p972A11273D
CX3CL1	NM_002996	Chemokine (C-X3-C motif) ligand 1	HU3_p972H09296D8
CYP11A1	NM_000781	Cytochrom P450, family 11, subfamily A, polypeptide 1	HU3_p972B07120D

Genname	GeneID	Beschreibung	RZPD No
CYP19A1	NM_031226	Cytochrome P450, family 19, subfamily A, polypeptide 1	HU3_p972E1221D
CYP1A1	NM_000499	Cytochrome P450, family 1, subfamily A), polypeptide 1	IMAGp958L021543Q2
CYP1B1	NM_000104	Cytochrom P450, family 1, subfamily B, polypeptide 1	HU3_p972E0468D8
CYP2B6	NM_000767	Cytochrom P450, subfamily I, (dioxin-inducible)	HU3_p972B1142D
CYP2C9*	NM_000771	Cytochrome P450, family 2, subfamily B, polypeptide 6	IMAGp950E1329
CYR61	NM_001554	Cysteine-rich, angiogenic inducer, 61	IMAGp958A231259Q2
DC13	NM_020188	DC13 protein	HU3_p972D01139D
DCC	NM_005215	Deleted in colorectal carcinoma	HU3_p972G10197D
DCK	NM_000788	Deoxycytidine kinase	HU3_p972G04144D
DCN	NM_001920	Decorin	HU3_p972H0374D
DCTD	NM_001921	dCMP deaminase	HU3_p972F0953D8
DDIT3	NM_004083	DNA-damage-inducible transcript 3	HU3_p972D11144D
DEDD	NM_032998	Death effector domain-containing	HU3_p972D06282D8
DEPDC1B (XTP1)	NM_018369	HBxAG transactivated protein 1 (DEP domain containing 1B)	HU3_p972B0931D
DFNA5	NM_004403	Deafness, autosomal dominant 5	HU3_p972C09119D
DHPS	NM_001930	Deoxyhypusine synthase	HU3_p972C02274D
DIAPH3	NM_030932	Diaphanous homolog 3 (Drosophila)	HU3_p972D06178D
DKFZp4341-L0117	NM_022778	Hypothetical protein DKFZp434L0117	HU3_p972E011D
DLC1	NM_006094	Deoxyhypusine synthase	HU3_p972F0515D8
DLG7	NM_014750	Deleted in liver cancer 1	HU3_p972D06103D
DNAL11	NM_003462	Dynein, axonemal, light intermediate polypeptide	HU3_p972A07277D8
DPYD	NM_000110	Dihydropyrimidine dehydrogenase	IMAGp958C07920Q2
DRIM	NM_014503	Down-regulated in metastasis	HU3_p972D0732D
DSP	NM_004415	Desmoplakin (DPI, DPII)	HU3_p972H0260D
DUSP1	NM_004417	Dual specificity protein phosphatase 1, dephosphorylates map kinase erk2	HU3_p972D1150D
E2-EPF	NM_014501	Ubiquitin-conjugating enzyme E2S	IMAGp958A062462Q2
E2F1	NM_005225	E2F transcription factor 1	IMAGp958A101833Q2
E2F3	NM_001949	E2F transcription factor 3	HU3_p972G0757D
EBP	NM_006579	Emopamil binding protein (sterol isomerase)	IMAGp958A17318Q2
ECGF1	NM_001953	Endothelial cell growth factor 1 (platelet-derived),	IMAGp958A201720Q2
ECT2	NM_018098	Epithelial cell transforming sequence 2 oncogene	HU3_p972H09176D
EDN1	NM_001955	Endothelin1	IMAGp998A086498Q3
EDNRB	NM_000115	Endothelin receptor type B	HU3_p972F01291D
EGF	NM_001963	Epidermal growth factor (beta-urogastrone)	IMAGp958A121058Q2
EGFR*	NM_005228	Epidermal growth factor receptor (avian erythroblastic leukemia viral (v-erb-b) oncogene homolog)	IMAGp998L13400Q2
EGFR	NM_005228	Epidermal growth factor receptor (avian erythroblastic leukemia viral (v-erb-b) oncogene homolog)	HU3_p972F0749D8
EGLN1	NM_022051	EGL nine homolog 1	HU3_p972E11276D
EIF2S1	NM_004094	Eukaryotic translation initiation factor 2, subunit 1 alpha, 35kDa	HU3_p972A0359D8
EIF3S4	NM_003755	Eukaryotic translation initiation factor 3, subunit 4 (44 kD)	HU3_p972A01332D8
EIF4E	NM_001968	Eukaryotic translation initiation factor 4E	HU3_p972F0579D

Gennamen	GeneID	Beschreibung	RZPD No
ELF3	NM_004433	E74-like factor 3 (ets domain transcription factor, epithelial-specific)	HU3_p972E01142D
EMP1	NM_001423	Epithelial membrane protein 1	HU3_p972E0149D
EMP2	NM_001424	Epithelial membrane prtotein 2	HU3_p972B1097D
ENG	NM_000118	Endogolin; CD 105	HU3_p972A03288D
ENO1	NM_001428	Enolase 1 (alpha)	HU3_p972A02307D
EP300	NM_001429	E1A- binding protein p300	HU3_p972E0164D
EPHB4	NM_004444	Eph receptor B4	IMAGp958A10197Q2
EPHX1	NM_000120	Epoxide hydrolase 1, microsomal	HU3_p972H11291D8
ERBB2IP	NM_018695	ERBB2 interacting protein	HU3_p972B1281D8
ERBB3*	NM_001982	erbB-3 receptor protein-tyrosine kinase	IMAGp998B03782
ERG	NM_004449	v-ets erythroblastosis virus E26 oncogene like (avian)	IMAGp998J12119Q2
ESR1	NM_000125	Estrogen receptor alpha	HU3_p972B08122D8
ESR2	NM_001437	Estrogen receptor beta	IMAGp958H211521Q2
ETS1	NM_005238	v-ets erythroblastosis virus E26 oncogene homolog 1 (avian)	HU3_p972D127D
ETS2	NM_005239	v-ets erythroblastosis virus E26 oncogene homolog 2	HU3_p972A09229D
ETV4	NM_001986	Ets variant gene 4 (E1A enhancer binding protein, E1AF)	HU3_p972H01110D
EXT1	NM_000127	Exostoses (multiple) 1	HU3_p972A096D
EZH2	NM_004456	Enhancer of zeste homolog 2 (Drosophila)	HU3_p972A0565D
F2R	NM_001992	Coagulation factor II (thrombin) receptor	HU3_p972B08146D
FABP3	NM_004102	Fatty acid binding protein 3, mammary-derived growth inhibitor (MDGI)	IMAGp958C081223Q2
FABP7	NM_001446	Fatty acid binding protein 7, brain	HU3_p972F10115D8
FADD	NM_003824	Fas-(TNFRSF6)-associated via death domain	IMAGp958A191708
FAF1	NM_007051	Fas-associated factor 1	HU3_p972G07290D8
FAP	NM_080685	Fibroblast activation protein	HU3_p972E1197D8
FAT	NM_005245	FAT tumor suppressor	HU3_p972C0298D
FBP1	NM_000507	Fructose-1,6-bisphosphatase 1	IMAGp958A102617Q2
FGF18	NM_003862	Fibroblast growth factor 18	HU3_p972B1285D
FGF2	NM_002006	Fibroblast growth factor 2 (basic)	HU3_p972D1296D
FGFR1	NM_023109	Fibroblast growth factor receptor 1	HU3_p972C04140D
FGFR2	NM_023028	Fibroblast growth factor receptor 2 (bacteria-expressed kinase, keratinocyte growth factor receptor, craniofacial dysostosis 1)	HU3_p972A01273D
FHIT	NM_002012	Fragile histidine triad gene	HU3_p972E1216D
FIGF	NM_004469	c-fos induced growth factor (vascular endothelial growth factor D)	HU3_p972C0233D
FLJ20489	NM_017842	Hypothetical protein: FLJ20489	IMAGp958A14787Q28
FLJ20605	NM_017898	Hypothetical protein: FLJ20605	IMAGp958C061730Q2
FLJ21062	NM_024788	Hypothetical protein: FLJ21062	IMAGp998A054765Q3
FLJ22719	AK026372	CDNA:FLJ22719 fis, clone HSI14307	HU3_p972B07183D
FLJ42602 (MT1E)	AK124593	Homo sapiens cDNA FLJ42602 fis, clone BRACE3011271	HU3_p972A07162D
FLNA	NM_001456	Filamin A, alpha (actin binding protein 280)	DKFZp434A0129
FLT1	NM_002019	fms-related tyrosine kinase 1 (vascular endothelial growth factor/vascular permeability factor receptor)	DKFZp686I18126Q2

Genname	GeneID	Beschreibung	RZPD No
FLT4	NM_182925	fms-related tyrosine kinase 4	IMAGp958F072042Q2
FOS	NM_005252	v-fos FBJ murine osteosarcoma viral oncogene homolog	HU3_p972D05272D8
FOXA1	NM_004496	Forkhead box A1	HU3_p972A01153D
FOXM1	NM_021953	Forkhead box M1	IMAGp958A021820Q2
FRZB	NM_001463	Frizzled-related protein	HU3_p972C0922D
FSCN1	NM_003088	Fascin homolog 1, actin-bundling protein	HU3_p972E0197D
FUCA1	NM_000147	Fucosidase, a-L-1	HU3_p972C0937D
FXYD3	NM_005971	FXYD domain containing ion transport regulator 3	IMAGp958A092327Q2
GADD45A	NM_001924	Growth arrest and DNA-damage-inducible, alpha	HU3_p972H0693D
GADP	NM_002046	Glyceraldehyde-3-phosphate dehydrogenase	HU3_p972C01134D8
GALNT3	NM_004482	UDP-N-acetyl-D-galactosamine:polypeptide N-acetylglactosaminyltransferase 3 (GalNAc-T3)	HU3_p972C0729D8
GAPDH*	NM_002046	Glyceraldehyde-3-phosphate dehydrogenase	IMAGp998N1236
GARP	NM_005512	Glycoprotein A repetitions predominant	HU3_p972F1053D
GATA3	NM_002051	GATA-binding protein 3	HU3_p972H11141D8
GCLM	NM_002061	Glutamate-cysteine ligase, modifier subunit	IMAGp958A01955Q2
GGH	NM_003878	Gamma-glutamyl hydrolase (conjugase, folylpolygammaglutamyl hydrolase)	HU3_p972A02287D
GJA1	NM_000165	Gap junction protein, alpha 1, 43kDa (connexin 43)	HU3_p972F1124D
GJB1	NM_000166	Gap junction protein, beta 1, 32 kD, connexin 32 (Cx32)	IMAGp958B141201Q2
GJB2	NM_004004	Gap junction protein, beta 2, 26kDa (connexin 26)	IMAGp998B237495Q3
GLUL	NM_002065	Glutamate-ammonia ligase	HU3_p972H08134D
GMPS	NM_003875	Guanine monphosphate synthetase	HU3_p972G0664D
GNAZ	NM_002073	Guanine nucleotide binding protein (G protein), alpha z polypeptide	HU3_p972G03283D
GNB2L1*	BC010119	Guanine nucleotide binding protein, beta polypeptide 2 like-1	IMAGp998I16819Q2
GPR160	NM_014373	G-protein coupled receptor 160	HU3_p972E0292D
GPX1	NM_000581	Glutathione peroxidase 1	HU3_p972F0664D
GPX4	NM_002085	Glutathione peroxidase 4	IMAGp958A012632Q2
GRK6	NM_002082	G-protein coupled receptor kinase 6	HU3_p972C10289D
GRN	NM_002087	Granulin	HU3_p972D129D8
GSN	X04412	Gesolin	HU3_p972B053D
GSTM3	NM_000849	Glutathione S-transferase M3 (brain)	HU3_p972C1122D
GSTP1	NM_000852	Glutathione-S-transferase P(i)1	IMAGp958A072458Q2
GUCA2B	NM_007102	Guanylate cyclase activator 2B	HU3_p972H03243D
HDAC1	NM_004964	Histone deacetylase 1	IMAGp958A191686Q2
HGF	NM_000601	Hepatocyte growth factor (hepapoietin A; scatter factor)	HU3_p972C02270D
HIF1A	NM_001530	Hypoxia-inducible factor 1, alpha subunit (basic helix-loop-helix transcription factor)	HU3_p972B1297D
HMGA1	NM_145904	High mobility group AT-hook 1 // high mobility group nonhistone chromosomal I and V	HU3_p972G09135D
HMGB1	NM_002128	High-motility group box 1	HU3_p972D0711D
HMMR	NM_012484	Hyaluronan-mediated motility receptor (RHAMM)	HU3_p972A12147D
HNRPA1	NM_031157	Heterogeneous nuclear ribonucleoprotein A1	HU3_p972C106D
HNRPH1	NM_005520	Heterogenous nuclear ribonucleoprotein H1	HU3_p972H0283D
HPN	NM_182983	Hepsin (transmembrane protease, serine1)	IMAGp958B191687Q2



Genname	GeneID	Beschreibung	RZPD No
HPSE	NM_006665	Heparanase	HU3_p972B0869D
HRASLS	NM_020386	HRAS like supressor 1	HU3_p972C09130D
HSPA1B	NM_005346	Heat shock 70 kDa protein 1B	DKFZp686M1914Q2
HSPA5	NM_005347	Heat shock 70kDa protein 5 (glucose-regulated protein, 78kDa)	HU3_p972A02258D
HSPCA*	NM_005348	Heat shock protein hsp 90-alpha	IMAGp998E091219Q2
HSPCA	AK056446	Heat shock 90kDa protein 1, alpha	DKFZp434A122Q2
HSPE1	NM_002157	Heat shock 10kDa protein 1 (chaperonin 10)	IMAGp998A201925Q3
HYOU1	NM_006389	Hypoxia up-regulated 1	HU3_p972C09290D
ICA1	NM_022307	Islet cell autoantigen 1, 69kDa	HU3_p972B03290D
ICAM1	NM_000201	Intercellular adhesion molecule 1 (CD54), human rhinovirus receptor	DKFZp686A2315Q2
ID1	NM_002165	Inhibitor of DNA binding 1, dominant negative helix-loop-helix protein	HU3_p972H0352D
IDUA	NM_000203	Iduronidase, a-L	HU3_p972C04125D
IFI30	NM_006332	Interferon, gamma-inducible protein 30	HU3_p972F01119D8
IFITM2	NM_006435	Interferon induced transmembran protein 2, Human 1-8D gene from interferon-inducible gene family	HU3_p972A0923D
IGF1R	NM_000875	Insulin-like growth factor 1 receptor	HU3_p972F1028D8
IGF1R*	NM_000875	Insulin-like growth factor 1 receptor	IMAGp950J0720
IGF2	NM_000612	Insulin-like growth factor 2 (somatomedin A)	IMAGp958A02191Q2
IGF2*	NM_000612	Insulin-like growth factor 2 (somatomedin A)	IMAGp998N04349Q2
IGFBP2	NM_000597	Insulin-like growth factor binding protein 2 (36 kD)	DKFZp686F2261Q2
IGFBP4	NM_001552	Insulin-like growth factor-binding protein 4	HU3_p972F0332D
IGHG1	AK090464	Immunoglobulin heavy constant gamma 1	HU3_p972B06135D8
IL11	NM_000641	Interleukin 11	HU3_p972D0797D
IL15	NM_000585	Interleukin 15	HU3_p972D0981D
IL1A	NM_000575	Interleukin 1, alpha	IMAGp998C157882Q3
IL1B	NM_000576	Interleukin 1, beta	HU3_p972G0594D
IL6R	NM_000565	Interleukin 6 receptor	HU3_p972G08137D
IL6ST	NM_002184	Interleukin 6 signal transducer (gp130 oncostatin M-receptor)	HU3_p972G0831D
IL8	NM_000584	Interleukin 8	IMAGp958A15957Q2
IMP-1	NM_006546	IGF-II mRNA-binding protein 1	HU3_p972G0323D
IMPA2	NM_014214	Inositol(myo)-1(or4)-monophosphatase 2	IMAGp958A12809Q2
ING1	NM_005537	Inhibitor of growth family, member 1	HU3_p972G1224D
INPP4B	NM_003866	Inositol polyphosphat-4-phosphatase, type II, 105 kDa	HU3_p972B05150D
IRS1	NM_005544	Insulin receptor substrate-1	HU3_p972B09110D
ITGA2	NM_002203	Integrin alpha 2	HU3_p972B06145D
ITGA5	NM_002205	Integrin, alpha 5 (fibronectin receptor, alpha polypeptide)	HU3_p972D0911D8
ITGA6	NM_000210	Integrin, alpha 6	HU3_p972G03232D
ITGAV	NM_002210	Integrin, alpha V (vitronectin receptor, alpha polypeptide, antigen CD51)	HU3_p972E0242D
ITGB1	NM_002211	Integrin, beta 1 (fibronectin receptor, beta polypeptide, antigen CD29 includes MDF2, MSK12)	HU3_p972C0489D
ITGB3	NM_000212	Integrin, beta 3 (platelet glycoprotein IIIa, antigen CD61)	IMAGp958B081319Q2

Genname	GeneID	Beschreibung	RZPD No
ITGB4	NM_000213	Integrin, beta 4	HU3_p972F0935D8
ITGB5	NM_002213	Integrin, beta 5	HU3_p972C0457D
ITR	NM_180989	Intimal thickness-related receptor	HU3_p972C09291D
JAG1	NM_000214	Jagged1 (NOTCH1 ligand)	HU3_p972D08109D
JARID1B	NM_006618	Jumonji, AT rich interactive domain, 1B	HU3_p972D0619D
JDP1	NM_021800	J domain containing protein 1	HU3_p972A0266D
JMJD1C (TRIP8)	NM_004241	Jumonji containing domain 1C	HU3_p972A06289D
JUN*	NM_002228	v-jun sarcoma virus 17 oncogene homolog (avian)	IMAGp998E10129Q2
JUNB	NM_002229	jun-B proto-oncogen	HU3_p972B019D
JUND	NM_005354	jun-D proto oncogen	IMAGp958A231824Q2
JUP	NM_002230	Junction plakoglobin	HU3_p972D02280D
KAI1	NM_002231	Kangai 1 (suppression of tumorigenicity 6, prostate; CD82 antigen)	HU3_p972D0333D
K-ALPHA-1	NM_006082	Tubulin, alpha,	IMAGp958A011161Q2
KCNN4	NM_002250	Potassium intermediate/small conductance caldium- activated channel, subfamily N, member 4	IMAGp958B172431Q2
KDR *	NM_002253	Kinase insert domain receptor (a type III receptor tyrosine kinase)	IMAGp998L12441
KIAA0430	NM_014647	KIAA00430 gene product	HU3_p972B0552D
KIAA1442	XM_044921	KIAA1442 protein	HU3_p972B04176D
KIAA1750 (TSPYL5)	NM_033512	KIA Protein 1750	HU3_p972B10288D
KIF23	NM_138555	Kinesin family member 23	HU3_p972A09301D
KIFC1	BC000712	Kinesin family member C1	HU3_p972A07258D
KIP2	NM_006383	DNA-dependent protein kinase catalytic subunit- interacting protein 2	HU3_p972A0891D8
KIT	NM_000222	v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog	HU3_p972F0173D8
KLK3	NM_001648	Kallikrein 3, (prostate specific antigen)	HU3_p972F06241D
KPNA2	NM_002246	Karyopherin alpha 2 (RAG cohort 1, importin alpha 1)	HU3_p972H122D
KRT14	NM_000526	Keratin 14 (epidermolysis bullosa simplex, Dowling- Meara, Koebner)	HU3_p972G0334D
KRT18*	NM_000224	Keratin 18	IMAGp998K231365Q2
KRT18	NM_000224	Keratin 18	IMAGp958A012282Q2
KRT19	NM_002276	Cytokeratin 19	HU3_p972H10122D
KRT20	NM_019010	Cytokeratin 20	HU3_p972C09151D
KRT7	NM_005556	Keratin 7	IMAGp958A012614Q2
KRT8*	NM_002273	Keratin 8	IMAGp998P16404Q2
LAD1	NM_005558	Ladinin 1	HU3_p972F086D8
LAF4	NM_002285	Lymphoid nuclear protein related to AF4	HU3_p972H09201D
LAMA3	NM_198129	Laminin, alpha 3	HU3_p972B0920D
LAMB1	NM_002291	Laminin, beta 1	DKFZp686A21147Q2
LAMB3	NM_000228	Laminin, beta 3	HU3_p972A09322D
LAMC1	NM_002293	Laminin, gamma 1 (formerly LAMB2)	HU3_p972A08175D
LAPTM4B	NM_018407	Lysosomal associated protein transmembrane 4 beta	HU3_p972H02291D
LASP1	NM_006148	LIM and SH3 protein 1	HU3_p972F03271D

Genname	GeneID	Beschreibung	RZPD No
LCN2	NM_005564	Lipocalin 2 (oncogene 24p3)	HU3_p972E09130D8
LDHB	NM_002300	Lactat dehydrogenase B	DKFZp686110144Q2
LGP2	NM_024119	Likely ortholog of mouse D11pg2	IMAGp998A012574Q3
LMO4	NM_006769	LIM domain only 4	HU3_p972C05293D8
LOX	NM_002317	Lysyl oxidase	HU3_p972E1252D
LOXL1	NM_005576	Lysyl oxidase-like 1	HU3_p972G11133D
LRBA	NM_006726	LPS-responsive vesicle trafficking, beach and anchor containing	HU3_p972C0774D
LRDD (PIDD)	NM_0184934	Leucin-rich and death domain containing	IMAGp958B11987Q2
LRRC17	NM_005824	Leucine rich repeat containing 17	HU3_p972C04290D
LRRC8	NM_019594	Leucine rich containing repeat 8	HU3_p972C10141D
LTBP1	NM_000627	Latent transforming growth factor beta binding protein 1	HU3_p972E0789D
LTBP3	NM_021070	Latent transforming growth factor beta binding protein 3	HU3_p972H01147D
LTBP3	NM_021070	Latent transforming growth factor beta binding protein 3	HU3_p972H01147D
LTBP4	NM_003573	Latent transforming growth factor beta binding protein 4	HU3_p972C0435D
LU	NM_005581	Lutheran blood group	HU3_p972G0633D
LUM	NM_002345	Lumican	HU3_p972H0324D
LYZ	NM_000239	Lysozyme (renal amyloidosis)	IMAGp958A011308Q2
MAD2L1	NM_002358	MAD2 (mitotic arrest deficient, yeast, homolog)-like 1	HU3_p972D1116D8
MADH1	NM_005900	MAD, mothers against decapentaplegic homolog 1	HU3_p972H0115D
MADH2	NM_005901	Smad2, mothers against decapentaplegic homolog 2	IMAGp958A23979Q2
MADH3	NM_005902	MAD, mothers against decapentaplegic homolog 3	HU3_p972H1298D8
MADH4	NM_005359	MAD, mothers against decapentaplegic homolog 4	HU3_p972E1193D8
MADH6	NM_005585	MAD, mothers against decapentaplegic homolog 6	HU3_p972A11112D
MADH7	NM_005904	Smad7, mothers against decapentaplegic homolog 7	IMAGp958G221347Q2
MAGEA1	NM_004988	Melanoma antigen, family A,1	IMAGp998C236475Q3
MAP2K1	NM_002755	Mitogen-activated protein kinase kinase 1	HU3_p972H04189D
MAP2K3	NM_002756	Mitogen-activated protein kinase kinase 3	IMAGp958A031774Q2
MAP2K6	NM_002758	Mitogen-activated protein kinase kinase 6	IMAGp958D031589Q2
MAP3K7	NM_003188	Mitogen-activated kinase kinase kinase 7	HU3_p972E09120D
MAP3K8	NM_005204	Mitogen-activated protein kinase kinase kinase 8	HU3_p972H0393D
MAPK13	NM_002754	Mitogen-activated protein kinase 13	HU3_p972G0295D
MAPK14	NM_001315	Mitogen activated protein kinase 14	HU3_p972E0445D8
MAPK8	NM_002750	Mitogen-activated protein kinase 8	IMAGp998A023415Q3
MAPT	NM_005910	Microtubule-associated protein tau	HU3_p972D02271D8
MAWBP	NM_022129	MAWD binding protein	IMAGp958A071351Q2
MCM2	NM_004526	Minichromosome maintenance deficient 2, mitotin (S. cerevisiae)	HU3_p972C0548D8
MCM5	NM_006739	Minichromosome maintenance deficient 5, cell division cycle 46 (S. cerevisiae)	HU3_p972A12107D
MCM6	NM_005915	Minichromosome maintenance deficient (mis5 S. pombe) 6	HU3_p972D113D
MDM2	NM_002392	Mouse double minute 2, human homolog of p53 binding protein	IMAGp958A091751Q2
MDM2 *	NM_002392	Mouse double minute 2, human homolog of p53 binding protein	IMAGp998H181668Q2

Genname	GeneID	Beschreibung	RZPD No
MEIS4	AL359938	Meis (mouse) homolog 4	IMAGp998D117174Q3
MELK	NM_014791	Maternal embryonic leucin zipper kinase	HU3_p972E11147D
MGC15419	NM_024735	MGC15419 protein	HU3_p972A0740D
MGC17337	NM_080655	Transmembrane protein with EGF-like and two follistatin-like domains 1 similar to RIKEN cDNA 5730528L13 gene	HU3_p972E09109D
MGC22960	NM_001441	MGC22960 gene	HU3_p972D1015D
MGLL	NM_007283	Monoglyceride lipase	HU3_p972F03290D
MIF	NM_002415	Macrophage migration inhibitory factor	IMAGp958A012053Q2
MKI67	NM_002416	Antigen identified by monoclonal antibody Ki-67, Ki-67 antigen, Mib-1 antigen	HU3_p972C112D8
MLH1	NM_000249	MutL (E. coli) homolog 1 (colon cancer, nonpolyposis type 2)	HU3_p972B0872D
MMP1	NM_002421	Matrix metalloproteinase 1	HU3_p972G0697D
MMP11	NM_005940	Matrix metalloproteinase 11 (stromelysin 3)	HU3_p972C10116D
MMP13	NM_002427	Matrix metalloproteinase 13	HU3_p972D02144D
MMP14	NM_004995	Matrix metalloproteinase 14 (membrane-inserted)	HU3_p972F0274D
MMP15	NM_002428	Matrix metalloproteinase 15 (membrane-inserted)	HU3_p972F04143D
MMP17	NM_016155	Matrix metalloproteinase 17	HU3_p972D12275D
MMP2	NM_004530	Matrix metalloproteinase 2 (gelatinase A, 72kDa gelatinase, 72kDa type IV collagenase)	HU3_p972E0897D
MMP3	NM_002422	Matrix metalloproteinase 3	HU3_p972H0297D
MMP7	NM_002423	Matrix metalloproteinase 7 (matrilysin)	HU3_p972A08114D
MMP9	NM_004994	Matrix metalloproteinase 9	HU3_p972H02268D
MN1	NM_002430	Meningioma (disrupted in balanced translocation) 1	HU3_p972D09296D
MRPL40	NM_003776	Mitochondrial ribosomal protein L40	HU3_p972H1015D
MRPS22	NM_020191	Mitochondrial ribosomal protein S22 //coatomer proteincomplex, subunit beta 2 prime	IMAGp958A082426Q2
MS4A7	NM_021201	Membrane-spanning, 4-domains, subfamily A, member 7	HU3_p972G1025D
MSH2	NM_000251	mutS (E. coli) homolog 2 (colon cancer, nonpolyposis type 1)	HU3_p972F05278D
MST1R	NM_002447	Macrophage stimulating 1 receptor (c-met-related tyrosine kinase)	IMAGp958B142570Q2
MTA1	NM_004689	Metastasis-associated gene	HU3_p972G0786D8
MVP	NM_005115	Major vault protein, lung resistance protein, LRP, VAULT1	IMAGp958A031941Q2
MYB	NM_005375	v-myb avian myeloblastosis viral oncogene homolog	IMAGp958A191449Q2
MYBL2	NM_002466	V-myb avian myeloblastosis viral oncogene homolog-like 2	HU3_p972C10107D8
MYO1C	NM_033375	Myosin IC	DKFZp686C1316Q2
NAT1	NM_000662	N-acetyltransferase 1	HU3_p972F06313D
NCOA4	NM_006534	Nuclear receptor coactivator 4	HU3_p972H08284D
NCOR1	NM_006311	Nuclear receptor corepressor 1, KIAA1047 protein	IMAGp998C1118Q3
NCOR2	NM_006312	Nuclear receptor corepressor 2, silencing mediator of retinoid and thyroid hormone action (SMRT)	HU3_p972B0333D8
NF2	NM_181826	Neurofibromin 2 (bilateral acoustic neuroma)	IMAGp958A24757Q2
NFIB	NM_005596	Nuclear factor I/B	HU3_p972C0118D

Genname	GeneID	Beschreibung	RZPD No
NFKB1	NM_003998	Nuclear factor of kappa light chain gene enhancer in B-cells	HU3_p972D0766D8
NFYB	NM_006166	Nuclear transcription factor Y, beta	HU3_p972C0991D8
NK4	NM_004221	Natural killer cell transcript 4	HU3_p972G0992D
NME1	NM_000269	Protein expressed in non-metastatic cells 1, NM23-h1, nm23A, nucleoside diphosphatase kinase A (NDKA)	HU3_p972F0987D8
NME1	NM_000269	Non-metastatic cells 1, protein expressed in	IMAGp998H2356
NME2	NM_002512	Non-metastatic cells 2, protein expressed in	HU3_p972D0936D8
NME3	NM_002513	Non-metastatic cells3, protein expressed in	HU3_p972D10118D
NMT1	NM_021079	N-myristoyltransferase 1	IMAGp958A061294Q2
NMU	NM_006681	Neuromedin U	IMAGp958D01515Q2
NPY1R	Hs.519057	Neuropeptide Y receptor Y1	HU3_p972A0330D
NRG1	NM_013957	Neuregulin1	HU3_p972E0233D
NRIP1	NM_003489	Nuclear receptor interacting protein 1	HU3_p972H0777D8
NSEP1	NM_004559	Nuclease sensitive element binding protein 1, YB-1	HU3_p972H02275D8
NUDT1	NM_002452	Nudix (nucleoside diphosphate linked moiety X)-type motif 1	HU3_p972C0718D
NUMA1	NM_006185	Nuclear mitotic apparatus protein 1	HU3_p972B094D
NUSAP1	NM_016359	nucleolar and spindle associated protein 1	HU3_p972B0986D
ODC1	NM_002539	Ornithine decarboxylase 1	IMAGp958A012000Q2
ORC6L	NM_014321	Origin recognition complex, subunit 6 (yeast homolog)-like	IMAGp998H222413Q3
OXCT	NM_000436	3-oxoacid CoA transferase (OXCT)	HU3_p972H08168D
P2RX4	NM_032554	Purinergic receptor P2X, ligand-gated ion channel 4	HU3_p972G10167D
P8	NM_012385	P8 protein (candidate of metastasis 1)	IMAGp958C09138Q2
PABPC1	NM_002568	Poly(a)-binding protein 1	HU3_p972A11285D
PAK1IP1	NM_017906	PAK1 interacting protein 1	IMAGp998B16150Q3
PALM2	NM_007203	Paralemmin 2	HU3_p972G0877D
PAPPA	NM_002581	Pregnancy-associated plasma protein A	IMAGp958A131573Q2
PBX3	NM_006195	Pre-B-cell leukemia transcription factor 3	HU3_p972F05144D
PCNA*	NM_002592	Proliferatin cell nuclear antigen	IMAGp998C24845Q2
PCNA	NM_002592	Proliferatin cell nuclear antigen	HU3_p972B023D8
PCSK7	NM_004716	Proprotein convertase subtilisin/kexin type 7	DKFZp686A01205Q2
PDGFA*	NM_002607	Platelet-derived growth factor alpha	IMAGp998B064598
PDZK1	NM_002614	PDZ domain containing 1	HU3_p972G0346D8
PECAM1	NM_000442	Platelet/endothelial cell adhesion molecule CD31 antigen	HU3_p972E0320D
PECI	NM_006117	Peroxisomal D3, D2-enoyl-CoA isomerase	IMAGp958A051426Q2
PEPD	NM_000285	Pepidase D	IMAGp998L1082
PES1	NM_014303	Pescadillo	IMAGp958A02124Q2
PFKP	NM_002627	Phosphofructokinase, platelet	DKFZp686A22223Q2
PHB	NM_002634	Prohibitin	HU3_p972H1143D
PIB5PA	NM_014422	Phosphatidylinositol (4,5) biphosphate 5-phosphatase, A	HU3_p972A02277D8
PIP	NM_002652	Prolactin-induced protein	HU3_p972E05154D
PLAT	NM_000930	Plasminogen activator, tissue	IMAGp958A01336Q2
PLAU	NM_002658	Plasminogen activator, urokinase	HU3_p972D09234D
PLAUR	NM_002659	Plasminogen activator, urokinase receptor	HU3_p972H01267D

Genname	GeneID	Beschreibung	RZPD No
PLD1	NM_002662	Phospholipase D1, phosphatidylcholine-specific	HU3_p972C01232D
PLK	NM_005030	Polo-like kinase (Drosophila)	IMAGp958A02121Q2
POLD4	NM_021173	Polymerase (DNA-directed), d4	IMAGp958A012086Q2
POLR2J	NM_006234	RNA polymerase II	IMAGp958A011431Q2
POPDC2	NM_022135	Popeye domain containing 2	HU3_p972E047D8
POR	NM_000941	P450 (cytochrome) oxidoreductase	IMAGp958A032427Q2
PPARBP	NM_004774	Peroxisome proliferator-activated receptor binding protein, Vitamin D receptor-interacting protein complex component	HU3_p972B03278D8
PPP1R15A	NM_014330	Protein phosphatase 1, regulatory (inhibitor) subunit 15A	HU3_p972G1117D8
PQLC2	NM_017765	PQ-loop repeat containing 2	HU3_p972A01197D
PRC1	NM_003981	Protein regulator of cytokinesis 1	HU3_p972G10143D
PRKCA	NM_002737	Protein kinase C alpha	IMAGp958A101900Q2
PSCD1	NM_004762	Pleckstrin homology, Sec 7 and coiled-coil domains 1	HU3_p972E07150D
PSMA6	NM_002791	Proteasome subunit alpha type 6	HU3_p972D1092D
PSMC5	NM_002805	Proteasome (prosome, macropain) 26S subunit	HU3_p972A01283D
PTEN	NM_000314	Phosphatase and tensin homolog (mutated in multiple advanced cancers 1	HU3_p972G12136D
PTGS2	NM_000963	Prostaglandin-endoperoxid synthase 2	HU3_p972F0827D
PTH1H	NM_002820	Parathyroid hormone-like hormone	HU3_p972H05144D
PTK2	NM_005607	PTK2 protein tyrosine kinase 2	HU3_p972A06122D8
PTN*	NM_002825	Pleiotrophin	IMAGp998I193889
PTP4A2	NM_003479	Protein tyrosine phosphatase type IVA, member 2	HU3_p972A02293D
PTPRF	NM_002840	Protein tyrosine phosphatase, receptor type, F	HU3_p972F10286D
PTTG1	NM_004219	Pituitary tumor-transforming 1	IMAGp998A161825Q3
PTX3	NM_002852	Pentaxin-related gene, rapidly induced by IL-1 beta	IMAGp998A18737Q3
PURA	NM_005859	Purine-rich element binding protein A	HU3_p972E04147D
PVALB	NM_002854	Parvalbumin	HU3_p972E0141D8
PXN*	NM_002859	Paxillin	IMAGp998K04671
QSCN6L1	NM_181701	Quiescin 6-like 1	HU3_p972H09287D
RAB2L	NM_004761	RAB2, member of RAS oncogene family-like	HU3_p972G07148D
RAB6B	NM_016577	RAB6B, member RAS oncogene family	IMAGp958I0196Q2
RAD51C	NM_058216	RAD51 homolog C (S. cerevisiae)	HU3_p972D11150D
RAF1	NM_002880	v-raf-1 murine leukemia viral oncogene 1	HU3_p972F0462D8
RAMP	NM_016448	RA-regulated matrix associated protein	HU3_p972E0114D
RANK	NM_003839	Tumor necrosis factor receptor superfamily, member 11a, activator fo NfKB	HU3_p972E1098D8
RARA	NM_000964	Retinoic acid receptor, alpha	HU3_p972G1238D
RARB	NM_016152	Retinoic acid receptor beta-2	IMAGp998H15416
RARG	NM_000966	Retinoic acid receptor, gamma	IMAGp998P211965
RARRES1	NM_002888	Retinoic acid receptor responder 1	HU3_p972E02258D
RB1	NM_000321	Retinoblastoma-1	IMAGp958B051650Q2
RBBP8	NM_002894	Homo sapiens retinoblastoma binding protein 8 (RBBP8)	HU3_p972C06294D8
RBP1	NM_002899	Retinol-binding protein, cellular	IMAGp958A07313Q2
RCF4	NM_002916	Replication factor C (activator 1) 4	HU3_p972A02135D

Genname	GeneID	Beschreibung	RZPD No
RCV1	NM_002903	Recoverin	HU3_p972G1253D8
REA	NM_007273	Repressor of estrogen receptor activity, B-cell associated protein (BAP), D-prohibitin	HU3_p972C08134D8
REL	NM_002908	v-rel reticuloendotheliosis viral oncogen homolog (avian)	IMAGp958C211362Q2
RELB	NM_006509	v-rel reticuloendotheliosis viral oncogen homolog B, nuclear factor of NfKB	HU3_p972G032D8
RERG	NM_032918	RAS-like, estrogen-regulated, growth-inhibitor	HU3_p972G01267D
RPL23A	NM_000984	Ribosomal protein L23A	HU3_p972B0846D
RPL37	NM_000997	Ribosomal Protein L37	IMAGp958A031951Q2
RPLP1	NM_001003	Ribosomal protein, large, P1	HU3_p972D0123D
RPS19	NM_001022	Ribosomal protein S19	HU3_p972D0758D
RPS23	NM_001025	Ribosomal protein S23	HU3_p972D05296D
RPS6KB1	NM_003161	Ribosomal protein S6 kinase polypeptide 1	HU3_p972C0932D
RRM2	NM_001034	Ribonucleotide reductase M2 polypeptide	HU3_p972F08109D
RTN4RL1	NM_178568	Reticulon 4 receptor-like 1	HU3_p972D05148D
RXRA	NM_002957	Retinoid X receptor, alpha	HU3_p972A0360D
S100A4	NM_002961	S100 calcium binding protein A4 (metastasin)	IMAGp958A062469Q2
S100A6	NM_014624	S100 calcium binding protein A6 (calcyclin)	IMAGp958A051318Q2
S100A7	NM_002963	S100 calcium binding protein A7 (psoriasin 1)	IMAGp998B117653Q3
S100A8	NM_002964	S100 calcium-binding protein A8(calgranulin A)	IMAGp958A142587Q2
S100P	NM_005980	S100 calcium-binding protein P	HU3_p972H0523D8
SAFB	NM_002967	Scaffold attachment factor B, Hsp27 ERE-TATA-binding protein (HET), HnRNP A1 associated protein (HAP)	HU3_p972C12104D8
SAT	NM_002970	Spermidine/spermine N1-acetyltransferase	HU3_p972F1060D8
SCA2	NM_002973	Spinocerebellar ataxia 2 (olivopontocerebellar ataxia 2, autosomal dominant, ataxin 2)	IMAGp958A021250Q2
SCGB2A1	NM_002407	Secretoglobin family 2A, member 1	HU3_p972B01128D
SCUBE2	NM_020974	Signal peptide, CUB domain, EGF-like 2	HU3_p972B01103D
SELENBP1	NM_003944	Selenium binding protein 1	IMAGp958A022385Q2
SERPINA3	NM_000624	Serine protease inhibitor, clade A, member 3	IMAGp958A062495Q2
SERPINB2	NM_002575	Serine (or cysteine) proteinase inhibitor clade B (ovalbumin), member 2	HU3_p972G0467D
SERPINB6	NM_004568	Serine proteinase inhibitor, clade B, member 6	HU3_p972D02128D
SERPINE1*	NM_000602	Serine (or cysteine) proteinase inhibitor, clade E, member 1	IMAGp998F161320
SERPINI1	NM_005025	Serine (or cysteine) proteinase inhibitor, clade I (neuroserpin), member 1	IMAGp958A11112Q2
SFRS5	NM_006925	Splicing factor, arginine/serine-rich 5	HU3_p972G1020D8
SIAT4C	NM_006278	Sialyltransferase 4C	IMAGp958A121278Q2
Sim. DIAPH3 (lLoc341692)	BC041395	Similar to diaphanous homolog 3 (Drosophila)	HU3_p972D02141D
SKIL*	NM_005414	SKI-like	IMAGp956E0896
SLC12A2	NM_001046	Solute carrier family 12 (sodium/potassium/chloride transporters), member 2	HU3_p972G12204D8
SLC39A6	NM_012319	Solute carrier family 39 (metall ion transporter) member 6	HU3_p972A1136D

Genname	GeneID	Beschreibung	RZPD No
SLC43A1	NM_003627	Solute carrier family 43 member 1	HU3_p972A0520D
SLC7A5	NM_003486	Solute carrier family 7 (cationic amino acid transporter, y+ system), member 5	HU3_p972G1061D8
SLC9A3R1	NM_004252	Solute carrier family 9 (sodium/hydrogen exchanger), isoform 3 regulatory factor 1	HU3_p972G0834D8
SLPI	NM_003064	Secretory leukocyte protease inhibitor	HU3_p972H06106D
SMURF1	NM_020429	Smad specific E3 ubiquitin protein ligase 1	HU3_p972F1230D8
SNCG	NM_003087	Synuclein, gamma (breast cancer-specific protein 1)	HU3_p972H03124D
SOD2	NM_000636	Superoxide dismutase 2, mitochondrial	HU3_p972D0332D
SOD3	NM_003102	Superoxide dismutase 3, extracellular	HU3_p972B09322D8
SOX4	NM_003107	SRY (sex determining region Y)-box 4	HU3_p972A1288D
SP1	NM_138473	Sp1 transcription factor	HU3_p972C0574D
SPARC	NM_003118	Secreted protein, acidic, cysteine-rich	HU3_p972B02274D
SPP1	NM_000582	Secreted phosphoprotein 1 (osteopontin, bone sialoprotein I, early T-lymphocyte activation 1)	HU3_p972G09285D
SRA1	AF092038	Steroid receptor RNA activator (1)	HU3_p972F0978D8
SRD5A1	NM_001047	Steroid-5-alpha reductase, alpha polypeptide 1	IMAGp998A025084Q3
SRPK1	NM_003137	SFRS protein kinase 1	HU3_p972G0563D
STAR	NM_003491	Steroidogenic acute regulatory protein	IMAGp998M141785Q3
STC2	NM_003714	Stanniocalcin 2	HU3_p972D0318D8
STK11	NM_000455	Serine/threonine kinase 11	IMAGp958A041998Q2
STK32B (HSA250839)	NM_018401	Gene for serine/threonine protein kinase 32 B	HU3_p972D0399D
SYT1	NM_005639	Synaptotagmin I	HU3_p972F08279D8
TACSTD1	NM_002354	Tumor-associated calcium signal transducer 1	IMAGp958B231234Q2
TBX2	NM_005994	T-box 2	HU3_p972A11117D
TCEAL1	NM_004780	Transcription elongation factor A (SII)-like 1, Homo sapiens (pp21)	HU3_p972H07117D
TEGT	NM_003217	Testis enhanced gene transcript	HU3_p972E0833D
TEK	NM_000459	TEK tyrosine kinase, endothelial (venous malformations, multiple cutaneous and mucosal)	HU3_p972D0635D
TERT	NM_003219	Telomerase reverse transcriptase	IMAGp998F093325Q3
TFF1	NM_003225	Trefoil factor 1, pS2, estrogen-inducible sequence expressed in breast cancer (BCEI)	IMAGp958A162622Q2
TFRC	NM_003234	Transferrin receptor	HU3_p972A0818D8
TGFA	NM_003236	Transforming growth factor alpha	HU3_p972C0336D8
TGFB1	NM_000660	Transforming growth factor beta 1	HU3_p972A0623D8
TGFB2	NM_003238	Transforming growth factor beta 2	HU3_p972C1277D8
TGFB3*	NM_003239	Transforming growth factor beta 3	IMAGp998E145734
TGFB3	NM_003239	Transforming growth factor beta 3	HU3_p972C04102D8
TGFBR1	NM_004612	Transforming growth factor beta receptor type I	DKFZp686I10211Q2
TGFBR2	NM_003242	Transforming growth factor beta receptor type II	DKFZp686C10189Q2
TGFBR3	NM_003243	Transforming growth factor beta receptor type III	HU3_p972D04286D8
THBS1	NM_003246	Thrombospondin 1	HU3_p972F0169D8
TIE	NM_005424	Tyrosine kinase with immunoglobulin and epidermal growth factor homology domains	HU3_p972E01280D
TIEG	NM_005655	TGFb inducible early growth response	HU3_p972E1281D8



Genname	GeneID	Beschreibung	RZPD No
TIMP1	NM_003254	Tissue inhibitor of metalloproteinase 1	IMAGp958A101240Q2
TIMP2	NM_003255	Tissue inhibitor of metalloproteinase 2	HU3_p972E0269D
TIMP3	NM_000362	Tissue inhibitor of metalloproteinase 3 (Sorsby fundus dystrophy, pseudoinflammatory)	HU3_p972H0751D8
TJP1	NM_003257	Tight junction protein 1 (zona occludens 1)	HU3_p972G06285D
TJP2	NM_004817	Tight junction protein 2 (zona occludens 2)	HU3_p972A0350D
TK1	NM_003258	Thymidine kinase	HU3_p972F11244D8
TLE3	NM_005078	Transducin-like enhancer of split 3 (E(sp1) homolog, Drosophila)	HU3_p972E09135D
TM9SF1	NM_006405	Transmembrane 9 superfamily member 1, multispinning membrane protein	HU3_p972C03138D
TMSB10	NM_021103	Thymosin beta-10	HU3_p972G0123D8
TNC	NM_002160	Tenascin C (hexabrachion)	HU3_p972H10123D
TNF	NM_000594	Tumor necrosis factor	HU3_p972F02195D8
TNFAIP3 (A20)	NM_006290	Tumor necrosis factor, alpha induced protein 3	IMAGp958C1853Q2
TNFRSF1A	NM_001065	Tumor necrosis factor receptor superfamily, member 1A	IMAGp958A061001Q2
TNFRSF1B (TNF-R2)	NM_001066	TNF-receptor superfamily, member 1b	HU3_p972B11295D8
TNFRSF6 (FAS)	NM_000043	Fas	HU3_p972D07134D8
TNFRSF6B	NM_032957	Tumor necrosis factor receptor superfamily, member 6b, decoy	IMAGp958A06132Q2
TNFSF10 (TRAIL)	NM_009425	Tumor necrosis factor receptor superfamily, member 10	IMAGp958A031378Q2
TNFSF11 (RANK-L)	NM_003701	Tumor necrosis factor (ligand) superfamily, member 11	HU3_p972E05151D8
TNFSF6	NM_000639	FasLigand	IMAGp958J192014Q2
TOB1	NM_005749	Transducer of ERBB2,1	HU3_p972D12144D
TOP2A*	NM_001067	Topoisomerase (DNA) II alpha (170 kD)	IMAGp998D17536Q2
TOP2A	NM_001067	Topoisomerase (DNA) II alpha (170 kD)	HU3_p972E0266D8
TPBG (ST4)	NM_006670	Trophoblast glycoprotein	IMAGp958A102576Q2
TPD52L1	NM_003287	Human D53 (hD53)	IMAGp958A06515Q2
TRA1	NM_003299	Tumor rejection antigen (gp96) 1	HU3_p972E1293D8
TRADD	NM_153425	TNFRSF1A-associated via death domain	HU3_p972F09326D8
TRAF1	NM_005658	TNF receptor-associated factor 1	IMAGp958B121623Q2
TRAF2	NM_021138	TNF receptor-associated factor 2	HU3_p972B11283D8
TRAF3	NM_145725	TNF receptor-associated factor 3	IMAGp958A162016Q2
TRAF4	NM_004295	TNF receptor-associated factor 4	HU3_p972B08273D8
TRAF5	NM_004619	TNF receptor-associated factor 5	HU3_p972G0997D8
TRAF6	NM_145803	TNF receptor-associated factor 6	IMAGp958B131721Q2
TRAIL-R1	NM_003844	Tumor necrosis factor receptor superfamily, member 10a	IMAGp958C04757Q2
TRAIL-R3	NM_003841	Tumor necrosis factor receptor superfamily, member 10c	IMAGp958B241535Q2
TRAIL-R4	NM_003840	Tumor necrosis factor receptor superfamily, member 10d, decoy with truncated death domain	HU3_p972E062D8
TRIP13	NM_004237	Thyroid hormone receptor interactor 13	HU3_p972H0845D

Genname	GeneID	Beschreibung	RZPD No
TRPV6	NM_018646	Transient receptor potential cation channel, subfamily V, number 6	IMAGp958G161074Q2
TSN	NM_004622	Translin	HU3_p972B11128D
TSP50	NM_013270	Testes-specific protease 50	HU3_p972C12313D
TUBB	NM_001069	Tubulin beta polypeptide	IMAGp958A022324Q2
TUBB5*	BC021909	Tubulin, beta 5	IMAGp998P09738Q2
TXN	NM_003329	Thioredoxin	HU3_p972F1141D
TXNIP	NM_006472	Thioredoxin interacting protein	HU3_p972B03172D
TYMS	NM_001071	Thymidylate synthetase	HU3_p972D08100D8
U2AF1	NM_006758	U2(RNU2) small nuclear RNA auxilliary factor 1, small subunit	HU3_p972F0861D
UGCG	NM_003358	UDP-glucose ceramide glucosyltransferase	HU3_p972A0786D
UMPS	NM_000373	Uridine monophosphate synthetase	IMAGp998N031749Q3
VCAM1	NM_001078	Vascular cell adhesion molecule 1	HU3_p972C11287D
VDR	NM_000376	Vitamin D receptor	IMAGp958B10257Q2
VEGF*	NM_003376	Vascular endothelial growth factor	IMAGp998P101162Q2
VEGFB	NM_003377	Vascular endothelial growth factor B	HU3_p972F12142D
VEGFC	NM_005429	Vascular endothelial growth factor C	HU3_p972E10286D
VIM	NM_003380	Vimentin	HU3_p972A1286D
VWF	NM_000552	von Willebrand Factor	HU3_p972F09275D
WEE1*	NM_003390	Wee1-like protein kinase	IMAGp998M03222Q2
WISP1	NM_003882	WNT1 inducible signaling pathway protein 2	HU3_p972C05120D
WISP1	NM_003882	WNT1 inducible signaling pathway protein 1	HU3_p972C05120D
WISP3	NM_198239	WNT1 inducible signaling pathway protein 3	IMAGp998K205294Q3
WNT10B	NM_003394	Wing-less type MMTV integration site family member 10B	HU3_p972F09143D
WNT5A	NM_003392	Wing-less type MMTV integration site family, member 5A	HU3_p972D0796D
WT1	NM_024426	Wilms tumor 1	HU3_p972C05115D
XBP1	NM_005080	X-box-binding protein 1	HU3_p972C1230D
XIAP	NM_001167	XIAP-associated factor 1	IMAGp958B061914Q2
XRCC1	NM_006297	X-ray repair complementing defective repair in Chinese hamster cells 1	HU3_p972H11269D8
ZAP128	NM_006821	Peroxisomal long-chain acyl-CoA thioesterase	IMAGp958A041154Q2
ZFP36	NM_003407	Zinc finger protein 36, C3H type homolog (mouse)	HU3_p972E04122D
ZNF217	NM_006526	Zinc finger protein 217	HU3_p972E0985D
ZNF533	NM_152520	Zinc finger protein 533	HU3_p972G12145D

## Danksagungen

Herrn Professor Brunner möchte ich für die Möglichkeit danken, die Doktorarbeit bei der Fraunhofer Gesellschaft durchzuführen und insbesondere nach dem unerwarteten und plötzlichen Tod von Herrn Dr. Johannes diese auch zu beenden.

Herrn Professor Scheurich danke ich für die Übernahme des Mitberichters sowie für den Zugang zum Institut für Zellbiologie und Immunologie.

Herrn PD Dr. Johannes möchte ich für die ersten beiden Betreuungsjahre der Doktorarbeit danken. Sein unerwarteter und plötzlicher Tod hat mich sehr getroffen.

Besonderer Dank gilt Herrn PD Dr. Steffen Rupp, der die Betreuung der Doktorarbeit übernommen hat und stets ein offenes Ohr für mich hatte.

Ebenfalls besonderer Dank gilt Frau Dr. Nicole Hauser, die mir in experimentiellen Fragen stets zur Seite stand.

Herrn Prof. Dr. C. Knabbe, Herrn Dr. P. Fritz, sowie Fr. Dr. M. Buck (Robert Bosch Krankenhaus) danke ich für die Bereitsstellung von Probenmaterial und Durchführung von Validierungsexperimenten.

Besonderer Dank gilt auch Herrn Dr. Jürgen Dippon und Herrn Stefan Winter, beide Mathematisches Institut A der Universität Stuttgart, bei der statistischen Datenanalyse.

Für die stets angenehme Arbeitsatmosphäre möchte mich bei allen Mitarbeiterinnen und Mitarbeitern der Abteilung *Genomics-Proteomics-Screening* ganz herzlich bedanken.

Meiner Familie möchte ich auch danken, da sie stets für mich da ist.