

Aus dem Institut für Immunologie der Ludwig-Maximilians-Universität zu München

Vorstand: Prof. Dr.rer.nat. Thomas Brocker

**Analyse einzelner disseminierter
Mammakarzinomzellen durch
Suppression Subtractive Hybridization:
Identifizierung eines differentiell exprimierten
ERV9-LTR Transkriptes**

Dissertation

zum Erwerb des Doktorgrades der Humanbiologie

an der Medizinischen Fakultät der

Ludwig-Maximilians-Universität zu München

vorgelegt von

Karina Petat-Dutter

aus Saint Martin d'Hères (Frankreich)

2006

Mit Genehmigung der Medizinischen Fakultät
der Universität München

Berichterstatter: PD. Dr. med. Christoph A. Klein

2. Berichterstatter Prof. Dr. R. Bartl

.....

Mitberichterstatter: Prof. Dr. Dr. R. Kürzl

PD. Dr. A. Brehm

Prof. Dr. J. Diebold

Mitbetreuung durch den
promovierten Mitarbeiter:

Dekan Prof. Dr. med. D. Reinhardt

Tag der mündlichen Prüfung: 23 Oktober 2006

Inhaltsverzeichnis

<u>1</u>	<u>Einleitung</u>	7
1.1	<u>Das Mammakarzinom</u>	7
1.1.1	<u>Epidemiologie</u>	7
1.1.2	<u>Die Normale Brustdrüse</u>	7
1.1.3	<u>Pathogenese des Mammakarzinoms</u>	9
1.1.4	<u>Die Metastasierung</u>	10
1.2	<u>Nachweis von disseminierten Tumorzellen</u>	12
1.2.1	<u>Nachweis von epithelialen Zellen</u>	12
1.2.2	<u>Nachweis von nicht malignen Zellen</u>	13
1.2.3	<u>Notwendigkeit der Entwicklung eines sensitiven und spezifischen Markers</u>	14
1.3	<u>Analyse der Genexpression</u>	15
1.3.1	<u>Methoden zur Analyse der differentiellen Genexpression</u>	16
1.3.2	<u>Suppression Subtractive Hybridization (SSH)</u>	17
1.4	<u>Zielsetzung der Arbeit</u>	20
<u>2</u>	<u>Material und Methoden</u>	21
2.1	<u>Material</u>	21
2.1.1	<u>Chemikalien</u>	21
2.1.2	<u>Enzyme und Antikörper</u>	22
2.1.3	<u>Kit und Systeme</u>	23
2.1.4	<u>Puffer, Lösungen und Medium</u>	23
2.1.5	<u>Verbrauchsmaterial</u>	25
2.1.6	<u>Geräte</u>	26
2.1.7	<u>Oligonukleotide</u>	26
2.2	<u>Methoden</u>	29
2.2.1	<u>Isolierung einzelner Zellen</u>	29
2.2.2	<u>Globale cDNA-Amplifikation</u>	30
2.2.3	<u>Subtraktion</u>	34
2.2.4	<u>Klonierung und Screening</u>	39
2.2.5	<u>Klonierung der vollständigen Sequenz #288</u>	47
<u>3</u>	<u>Ergebnisse</u>	57
3.1	<u>Subtraktion mit Einzelzellen</u>	57
3.1.1	<u>Notwendige Modifikationen und Optimierung der SSH</u>	57
3.1.2	<u>Subtraktion mit Mamma-Ca Zellen</u>	64
3.2	<u>Analyse der subtrahierten Produkte</u>	69
3.2.1	<u>Klonierung und Screening</u>	69
3.2.2	<u>Validierung differentiell exprimierter Sequenzen</u>	69
3.3	<u>Die Sequenz #288</u>	71
3.3.1	<u>Zuordnung der Sequenz #288</u>	72
3.3.2	<u>Identifizierung der vollständigen cDNA vom Fragment #288</u>	73
3.3.3	<u>Suche nach der Sequenz #288 in einer cDNA Bank</u>	76
3.3.4	<u>Expressionsmuster der Sequenz #288 in verschiedenen Geweben</u>	79
3.3.5	<u>Elektronische Analyse</u>	80
3.3.6	<u>Die ERV9-LTR-Hypothese</u>	85
3.4	<u>Zusammenfassende Übersicht der Ergebnisse</u>	91
<u>4</u>	<u>Diskussion</u>	93
4.1	<u>Auswahl der Methode und des Materials</u>	94
4.2	<u>Entwicklung einer SSH für Einzelzellen</u>	95
4.3	<u>SSH von Tumorzellen und Knochenmark</u>	96
4.3.1	<u>Anzahl der humanen Sequenzen</u>	96

4.3.2	Bakterielle Sequenzen	97
4.3.3	Identität und Expression der Sequenzen	98
4.4	Die Sequenz #288	100
4.4.1	Analyse des chromosomalen Bereichs 3q25.1	100
4.4.2	Die ERV9-LTR-Hypothese	100
4.4.3	Die Promotoraktivität des ERV9 LTR	101
4.4.4	Schlußfolgerungen	102
5	Zusammenfassung	107
6	Literaturverzeichnis	109
7	Danksagung	119
8	Erklärung	121
9	Lebenslauf	123

Abkürzungsverzeichnis

bp	Basenpaare
cDNA	komplementäre Desoxyribonukleinsäure
DNA	Desoxyribonukleinsäure
dNTP	Desoxyribonukleotidtriphosphat
EST	<i>Expressed Sequence Tag</i>
g	Erdbeschleunigung (<i>gravity</i> , $g = 9,81 \text{ m/s}^2$)
h	Stunde (<i>hour</i>)
HERV	humaner endogener Retrovirus
KM	Knochenmark
min.	Minute(n)
mRNA	messenger Ribonukleinsäure
nt	Nukleotide
LTR	lange terminale Sequenzwiederholung (<i>Long Terminal Repeat</i>)
ORF	offener Leserahmen (<i>open reading frame</i>)
PCR	Polymerasekettenreaktion (<i>polymerase chain reaction</i>)
RDA	<i>Representational Difference Analyse</i>
RNA	Ribonukleinsäure (<i>ribonucleic acid</i>)
rRNA	ribosomale Ribonukleinsäure
rpm	Umdrehung pro Minute (<i>round per minute</i>)
RT	Raumtemperatur
RT	Reverse Transkriptase
RT-PCR	Reverse Transkription-Polymerasekettenreaktion
SAGE	<i>Serial Analysis of Gene Expression</i>
sek.	Sekunde(n)
SSH	Subtraktive Hybridisierung (<i>suppression subtractive hybridization</i>)
T _m	Schmelztemperatur (<i>Melting Temperature</i>)
U	Einheit (<i>unit</i>)
w/v	Gewicht/Volumen (<i>weight/volume</i>)

1 Einleitung

1.1 *Das Mammakarzinom*

1.1.1 Epidemiologie

In Deutschland erkrankten 1997 45 800 von 42 Millionen Frauen an Brustkrebs, was 26 % der jährlichen Krebsneuerkrankungen entspricht. In Deutschland erkrankt somit jede zehnte Frau im Laufe ihres Lebens an einem Mammakarzinom. Trotz eines Rückganges der Mortalität seit Anfang der 90er Jahre in Westeuropa, der die Fortschritte bei der Früherkennung und der Therapie widerspiegelt (Lacey *et al.*, 2002; Levi *et al.*, 2001), starben noch im Jahr 1997 und 1998 in Deutschland über 18 000 Frauen an den Folgen dieser Krankheit. Die mittlere Überlebenszeit nach der Diagnose eines Mammakarzinoms beträgt statistisch 6,4 Jahre (Engel *et al.*, 2001). Für 90 % der Sterbefälle an einem Karzinom ist die Metastasierung verantwortlich (Sporn, 1996). Deswegen ist es entscheidend, diesen Prozess zu verstehen. Trotz verbesserter Kenntnisse der Biologie dieser Erkrankung während der letzten zehn Jahre, wird die Metastasierung noch zu wenig verstanden. Aber gerade dieses Wissen ist ein wichtiger Ansatzpunkt, die Krankheit erfolgreich zu heilen, bzw. das Leben der Patientinnen zu verlängern. Die Fünfjahresüberlebensrate ist in der Tat abhängig vom Diagnosezeitpunkt: sie liegt bei 89 %, wenn der Tumor noch lokal begrenzt ist, sinkt aber auf 22 %, wenn bereits eine Metastasierung vorliegt (Lacey *et al.*, 2002).

1.1.2 Die Normale Brustdrüse

Das Wissen über die Entstehung der Brustdrüse ist noch unvollständig. Die Entwicklung beginnt *in utero* durch epithelial-mesenchymale Interaktionen und führt zu einem deutlich erkennbaren *solid epithelial bud* am Ende des zweiten Monats der Schwangerschaft (Übersicht in (Howard & Gusterson, 2000; Pitelka, 1988). In der 13. Woche beginnen Verzweigungen auszuwachsen. Erst um die 18. Woche treten die eigentlichen Mammazellen auf. Sie sind von der Basalmembran des angrenzenden Periderms durch ihre negative Färbung für Keratin 14 und *alpha smooth Muscle Actin* unterscheidbar. Im zweiten Trimenon der Schwangerschaft entstehen durch ständige Vermehrung und Verzweigung 10 bis 20 rudimentäre Ductus, die Lumina entwickeln. Die epithelialen Zellen der Gänge sind kubisch, polarisiert und tragen apicale Mikrovilli. Während des dritten Trimenons der Schwangerschaft werden die Ductus länger und bilden immer mehr Verzweigungen bis zum subkutanen Bindegewebe aus. Je nach Entwicklungsmaß sind die Ductus röhrenförmig oder verzweigt mit Ausbildung von Acini (beerenförmige Drüsen).

Weibliche und männliche Brustdrüsen durchlaufen *in utero* die gleiche Entwicklung und bis zur Pubertät sind keine morphologischen Unterschiede zu erkennen. Brustgewebe, das histologisch eine starke Verzweigung der Drüsen mit Acinus-Entwicklung zeigt, unterscheidet sich im interlobulären und intralobulären Stroma, dessen Fibroblasten teilweise Kollagen IV sezernieren. Die intralobulären Fibroblasten können zum brustspezifischen Stromalen Tumor führen. Nach der Geburt, wenn es keine mütterlichen hormonellen Einflüsse mehr gibt, bildet sich die Brust zurück. Im Alter von zwei Jahren sind nur kleine duktales Strukturen im fibroblastenreichen Stroma zu sehen.

Die Entwicklung der Brust bei Frauen findet während der Pubertät durch hormonelle Einflüsse statt. Die ersten Änderungen treten im Brustgewebe auf. Eine normale Brust besteht aus mindestens 80 % Bindegewebe. Die zwölf bis 20 Hauptductus, die die Brustwarze erreichen, werden länger und verzweigen sich (Dichotomie), um segmentale und kleinere subsegmentale Ductus zu bilden. Die subsegmentalen Ductus führen zu terminalen Ductuli, die als kleine Acini enden. Alle Acini, die aus einem einzigen terminalen Ductulus stammen, bilden die funktionale Einheit der Brust, den Lobulus (Terminal Duct Lobular Unit, TDLU) (**Abbildung 1**). Der Lobulus ist durch eine einzige myoepitheliale Zellmembran vom Stroma getrennt. Während der Stillzeit sind die Acini in der Brust mit Muttermilch gefüllt. Nach dem Abstillen bilden sich die Acini zurück, die Ductus aber sind wenig von der Rückbildung betroffen. Im Gegensatz dazu sind nach der Menopause Lobuli und Ductus quantitativ stark reduziert. Das intralobuläre Stroma wird durch Kollagen, das Epithelium und das interlobuläre Bindegewebe werden durch Fettgewebe ersetzt.

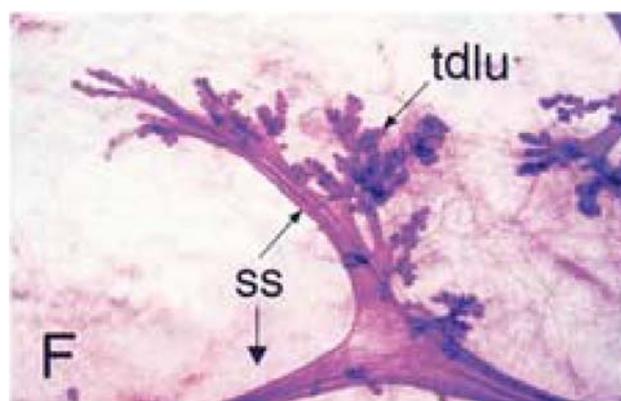


Abbildung 1. Vorbereitung des Brustgewebes einer nullipara 18-jährigen Frau.
Ein segmentaler Ductus teilt sich in zwei subsegmentalen Ductus (ss), der dann zu den terminalen Ductuli (tdlu) führt (Howard & Gusterson, 2000).

Das Brustdrüsengewebe ist das ganze Leben lang hormonellen Einflüssen wie zum Beispiel Menstruationszyklen oder Schwangerschaft unterworfen. Jeder Zyklus führt zur Vermehrung und Apoptose der epithelialen Zellen. Das Stroma steht ebenfalls unter hormoneller

Kontrolle. Möglicherweise hat die ständige zyklische Stimulation und Involution einen Einfluß auf die Krankheitsentwicklung.

1.1.3 Pathogenese des Mammakarzinoms

Es gibt eine Vielzahl von Erkrankungen, die die Brust betreffen: Mikrobiologische Infektionen, endokrine Funktionsstörungen, angeborene Abnormalitäten (Größe, Form, Drüsenzahl), sowie abnormale Zellvermehrungen (Hyperplasie). Gutartige Vermehrungen entsprechen Adenomen, sklerosierenden Läsionen, fibrozystischen Veränderungen und Papillomen. Die Tumorentwicklung kann sowohl Binde- und Epithelgewebe betreffen als auch die Umwandlung in eine bösartige Erkrankung. Bindegewebskrebse (Sarkome) sind seltene Krankheiten. Die meisten Krebsarten der Brust haben einen epithelialen Ursprung und werden „Karzinom“ genannt. Die Karzinome werden in nicht-invasive und invasive eingeteilt.

Die nicht-invasiven Karzinome werden auch als „In-situ-Karzinome“ bezeichnet. Das gemeinsame histopathologische Merkmal der In-situ-Karzinome ist der fehlende Nachweis der Stromainvasion (intakte Basalmembran; sobald diese durchbrochen ist, wird das Karzinom „invasiv“). Aufgrund der Morphologie wird zwischen Intraduktalem Karzinom (DCIS) und Lobulärem Karzinom (LCIS) unterschieden, obwohl es keine anatomische Rechtfertigung für diese Nomenklatur gibt. Man nimmt an, dass alle bösartigen Veränderungen aus dem TDLU stammen, unabhängig davon ob sie einen duktalem oder lobulären Phänotyp zeigen. Außerdem werden viele Läsionen als „duktalem“ bezeichnet, obwohl sie deutlich einen lobulären Ursprung haben (Mallon *et al.*, 2000).

Über 17 verschiedene Karzinome der Mamma sind bislang klassifiziert. Darunter ist das invasiv duktalem Karzinom mit 65-80 % der bei weitem häufigste Tumortyp. Er ist definiert als ein Tumor, der keiner Kategorie zugeordnet werden kann und deshalb als „*not otherwise specified*“ (NOS) bezeichnet wird. Alle andere Karzinome haben typische morphologische Charakteristika, die eine reproduzierbare Klassifizierung ermöglichen. In der Häufigkeit folgen dann das invasiv lobuläre (6-15 %), das medulläre (ca. 3 %), das tubuläre (1-15 %), das muzinöse (1-2 %) und das papilläre Karzinom (1-7 %). Die weiteren Karzinomformen treten in einer Häufigkeit von weniger als 1 % auf (Lebeau *et al.*, 2001).

Sobald die Basalmembran durchbrochen ist, wird das Karzinom als „invasiv“ bezeichnet. Der Tumor kann sich im Stroma ausbreiten und Lymph- und Blutgefäße erreichen. Maligne Zellen können sich dann im ganzen Körper ausbreiten und Metastasen bilden, besonders in Lunge, Leber, Gehirn und Knochen.

1.1.4 Die Metastasierung

Für die Entstehung eines Karzinoms ist kein solitäres Ereignis ausschlaggebend. Es werden zur Zeit fünf Mechanismen diskutiert, die zu Änderungen der Physiologie von Epithelzellen führen und in einer bösartigen Vermehrung resultieren: Stimulation durch Wachstumsfaktoren, Unempfindlichkeit gegen antiproliferative Signale oder programmierten Zelltod (Apoptose), Erhöhung des replikativen Potentials und Angiogenese (Ausbildung neuer Blutgefäße). Um sich im ganzen Körper ausbreiten zu können, benötigen die Zellen eine weitere Eigenschaft: die Fähigkeit zur Gewebeinvasion und Metastasierung (Hanahan & Weinberg, 2000). Weil die Metastasierung für 90 % der Sterbefälle an einem Karzinom verantwortlich ist (Sporn, 1996), ist es entscheidend diesen Prozess zu verstehen. In letzter Zeit wurden viele Gene identifiziert, die an der Metastasierung beteiligt sind (Debies & Welch, 2001). Beispielsweise geht die Zelladhäsion, die normalerweise für das Überleben epithelialer Zellen nötig ist, verloren. Selektine, Integrine, Lektine und Cadherine sind Moleküle, deren veränderte Expression mit der Metastasierung korreliert. Um zu wandern, müssen die Zellen durch die Basalmembran gelangen. Hierbei spielen die Proteolyse und die Migration eine entscheidende Rolle. An der Proteolyse sind zum Beispiel Matrix-Metalloproteinasen, Serin-Proteinasen und Cathepsine beteiligt, bei der Migration Autotaxin oder HGF (*Hepatocyte Growth Factor*).

Andere Prozesse wie Organotropismus oder Anoikis fördern ebenfalls die Metastasierung. Jacob *et al.* haben gezeigt, dass Osteonectin Brust- und Prostatakarzinomzellen in den Knochen leiten kann (Jacob *et al.*, 1999). Brustkarzinomzellen, die Metastasen im Lymphknoten und Knochenmark bilden können, weisen die Chemokin Rezeptoren CXCR4 und CCR7 auf (Müller *et al.*, 2001). Um Gewebe eines anderen histogenetischen Ursprungs besiedeln zu können, müssen die metastatischen Zellen der „Anoikis“ widerstehen. Diese spezielle Art der Apoptose betrifft Zellen, die ihre normalen Gewebeinteraktionen verloren haben. Warum Anoikis bei Tumorzellen nicht ausgelöst wird, ist noch weitgehend unverstanden.

Seit 1987 gibt es beim Mammakarzinom einen Rückgang der Mortalität um 25%, was auf die verbesserte Detektion durch die Mammographie und systemische Therapien, besonderes durch Tamoxifen, zurückgeführt wird (Lacey *et al.*, 2002). Die zwei stärksten prognostischen Marker bei Brustkrebs sind die Grösse des invasiven Tumors und die Beteiligung von axillären Lymphknoten. Da die Fünfjahresüberlebensrate auf 22 % sinkt (statt 89 % für einen lokalen Tumor) wenn eine Metastasierung vorliegt, ist eine frühzeitige Diagnose von überragender Bedeutung.

Das „Tumor-Node-Metastasis“ (TNM) System beschreibt die anatomische Ausdehnung der Erkrankung:

- T- Ausdehnung des Primärtumors,
- N- Fehlen oder Vorhandensein und Ausdehnung von regionären Lymphknotenmetastasen,
- M- Fehlen oder Vorhandensein von Fernmetastasen.

Jahrelang wurde in dieser Klassifikation die Anwesenheit von einzelnen disseminierten Tumorzellen im Blut oder Knochenmark (KM) nicht berücksichtigt. Aus diesem Grund wurden von verschiedenen Autoren unterschiedliche Bezeichnungen für diese Zellen verwendet, wie zum Beispiel ITC (*isolated-tumor-cells*, (Hermanek *et al.*, 1999), CTC (*circulating-tumor-cells*, (Marth *et al.*, 2002) oder OTC (*occult-tumor-cells*, (Gerber *et al.*, 2001). Erst in den neunziger Jahren wurde vorgeschlagen, sie unter der Bezeichnung pM_(i) in die TNM-Klassifikation aufzunehmen (Hermanek, 1994; Wittekind & Wagner, 1997). Die disseminierten Tumorzellen stellen einen wichtigen prognostischen Faktor dar, obwohl die unterschiedlichen immunzytochemischen Nachweissysteme Unterschiede in den prognostischen Aussagen zur Folge haben. Tatsächlich ändert sich je nach benutztem Marker der disseminierten Tumorzellen (u.a. epithelial Membranantigen (EMA) und Zytokeratin) der Zusammenhang zwischen der „minimal residualen Tumorerkrankung“ und dem rezidivfreien Überleben (*Disease-free survival*) bzw. dem Gesamtüberleben (*overall survival*) oder anderen prognostischen Faktoren (Übersicht in (Klein, 2003). 1998 haben Funke und Schraut in einer Meta-Analyse von insgesamt 2494 Patientinnen gezeigt, dass nur wenige Studien (5/11) einen positiven Nachweis von disseminierten Tumorzellen als unabhängig prognostischen Faktor für das rezidivfreie Überleben bestätigen (Funke & Schraut, 1998). Eine neue Meta-Analyse von 4703 Patientinnen zeigt aber, dass in neun Studien disseminierte Tumorzellen als unabhängiger prognostischen Faktor identifiziert wurden. Ein positiver Nachweis von disseminierten Tumorzellen korreliert mit einer geringeren Gesamtüberlebensrate im Vergleich zu Patienten bei denen keine disseminierten Tumorzellen detektiert wurden (Braun *et al.*, 2005). Die klinische Anwendung der Detektion der disseminierten Tumorzellen im Knochenmark ist aber immer noch aus Mangel an standardisierten Methoden eingeschränkt. Biologisch betrachtet könnten die disseminierten Tumorzellen der erste Schritt zur Mikrometastase (<2 mm Durchmesser oder <50 Zellen) bzw. zur Metastase (Klein, 2000) sein. Da einzelne dissiminierte Tumorzellen frühzeitig nachweisbar sind, könnten sie für die Planung einer adjuvanten Therapie von Bedeutung sein. Deswegen beschäftigen sich immer

mehr Arbeitsgruppen mit einzelnen disseminierten Tumorzellen im Blut oder Knochenmark, um ihre Biologie und Bedeutung besser zu verstehen. Die Möglichkeit der Unterscheidung zwischen Zellen, die lediglich disseminieren und denjenigen, die Metastasen bilden, ist eine wichtige Anforderung für die Zukunft. Mehrere Untersuchungen berichten, dass ca. die Hälfte der Patienten (160/315 and 46/89), die bei Diagnosestellung Tumorzellen im KM hatten, nach 6,5 Jahren keinen Rückfall (*relapse-free*) zeigten (Mansi *et al.*, 1991; Solomayer *et al.*, 2001). Nach 10 Jahren waren es noch 44 % (Mansi *et al.*, 1999). Je nach Tumortyp treten bei 30 % bis 70 % der Patienten mit disseminierten Tumorzellen in KM oder Lymphknoten keine Metastasen auf (Klein, 2003). Dieses stimmt mit experimentellen Beobachtungen überein, dass nur wenige disseminierte Tumorzellen (0,05 %) überleben und Metastasen bilden (Abati & Liotta, 1996).

Was soeben dargelegt wurde, kann in drei Punkten zusammengefaßt werden:

1. Die disseminierten Tumorzellen in Knochenmark oder Lymphknoten spiegeln vielleicht zwei unterschiedliche Metastasierungsprozesse wider (Gerber *et al.*, 2001).
2. Die disseminierten Tumorzellen sind eventuell die Vorläuferzellen der späteren Metastasen (Klein, 2000).
3. Der Nachweis der disseminierten Tumorzellen muss weiter verbessert werden.

1.2 **Nachweis von disseminierten Tumorzellen**

Mit der Computertomographie (CT), Röntgen, Ultraschall und Knochenszintigraphie kann eine minimale Streuung einzelner Tumorzellen nicht nachgewiesen werden. Eine frühe Detektion von disseminierten Tumorzellen würde die Entwicklung neuer Techniken benötigen.

1.2.1 **Nachweis von epithelialen Zellen**

Beim Nachweis von disseminierten Tumorzellen macht man sich ihren epithelialen Ursprung zu nutze. Epitheliale Zellen exprimieren Zytokeratine, die in mesenchymalen Geweben wie Knochenmark, Blut und Lymphknoten nicht vorhanden sind. Verschiedene Epithelien können anhand ihrer spezifischen Zytokeratin-Expression charakterisiert werden (Moll *et al.*, 1982). Die Entwicklung von Antikörpern gegen Zytokeratin 18 (Debus *et al.*, 1982) erlaubt eine immunozytochemische Färbung für den Nachweis von Mikrometastasen in Lymphknoten (Wells *et al.*, 1984) und von disseminierten Tumorzellen im Knochenmark (Schlimok *et al.*, 1987). Mittlerweile ist ein kommerzieller Kit für die zytochemische Routinefärbung erhältlich (Epimet Kit), der die intrazellulären Zytokeratine 8, 18 und 19 erkennt. Seltener wird auch eine direkte Anreicherung von Tumorzellen mittels Antikörpern, die an paramagnetische

Partikel gekoppelt sind, durchgeführt (Übersicht in (Krag *et al.*, 1999). Beide Methoden erlauben eine Isolierung einzelner Zellen, sind aber sehr zeitraubend. Deshalb versuchen mehrere Arbeitsgruppen die Tumorzellen mit Hilfe der RT-PCR direkt nachzuweisen (Ghossein & Rosai, 1996).

Für eine frühe Detektion von Tumorzellen wäre eine einfache Routinemethode, wie zum Beispiel die periphere Blutabnahme, wünschenswert. Die meisten Analysen werden allerdings bisher mit Knochenmarkproben durchgeführt, da die Tumorzellen im Knochenmark zahlreicher als im Blut (Braun *et al.*, 2001; Marth *et al.*, 2002) oder Lymphknoten zu finden sind (Gerber *et al.*, 2001). Die Frequenz der disseminierten Tumorzellen im Knochenmark nach Anreicherung über einen Dichtegradienten liegt bei 2 % für Patienten ohne Tumor. Bei einem Drittel der Karzinom-Patienten können 10^{-5} - 10^{-6} disseminierte Tumorzellen detektiert werden (Pantel *et al.*, 1999). Verschiedene Arbeiten zeigen, dass die Detektion einzelner disseminierter Tumorzellen im Knochenmark mit einer ungünstigeren Prognose für die Patienten korreliert. Die biologische Charakterisierung dieser Zellen kann von großer Bedeutung sein, um adaptierte Behandlungen der Patienten zu entwickeln. Die biologische Charakterisierung positiver Zellen kann sich zum Beispiel auf genomische Untersuchungen stützen (Klein *et al.*, 1999). Dadurch kann man nützliche Erkenntnisse über chromosomale Aberrationen gewinnen, eine Gen- und Proteinexpressionsanalyse, die einen weiteren Einblick in die Biologie der disseminierten Zellen geben könnte, bleibt aber unmöglich. Die für eine intrazelluläre Färbung nötige Fixierung der Zellen und Permeabilisierung der Zellmembran schaden der mRNA Qualität. Um eine möglichst gute mRNA Isolierung zu erreichen, ist eine Oberflächenfärbung besser geeignet.

Eine solche Färbung kann mittels des hochaffinen monoklonalen Antikörpers (mAb 3B10-C9) gegen das epitheliale Zelladhäsionsmolekül EpCAM erfolgen. Dieses Glycoprotein, das auf der Zelloberfläche lokalisiert ist, wird auf normalem epitheliales Gewebe schwach und bei vielen Karzinomformen stark exprimiert (Balzar *et al.*, 1999; Gottlinger *et al.*, 1986) und erlaubt eine *in vivo* Detektion von Tumorzellen im mesenchymalen Gewebe (Schlimok *et al.*, 1987).

1.2.2 Nachweis von nicht malignen Zellen

Beim Nachweis von Tumorzellen im Knochenmark müssen verschiedene Faktoren beachtet werden. Der Zeitpunkt der Knochenmarkpunktion kann die Ergebnisse beeinflussen, zu einer falschen Diagnose und folglich zur falschen Therapie führen. Es wurde zum Beispiel beobachtet, dass nach einer Biopsie der Brust die Zahl von epithelialen Zellen im

Lymphknoten zunimmt. Diese Zellen können entweder von einem Tumor stammen oder es sind gutartige epitheliale Zellen (Carter *et al.*, 2000). Sie spiegeln aber vermutlich keine Metastasierung wider.

Zweitens werden mit der Immunfluoreszenzfärbung schwach 3B10-C9-positive Zellen in den mononukleären Zellen des Knochenmarks (nach der Trennung durch Dichtegradientenzentrifugation) von gesunden Patienten detektiert. Die Frequenz liegt bei 10^{-3} - 10^{-4} . Die EpCAM Expression dieser Zellen konnte mit einer RT-PCR bestätigt werden (Klein *et al.*, 2002). Schließlich zeigt eine neue Arbeit, dass das EpCAM Glykoprotein nicht nur von epithelialen Zellen exprimiert wird, sondern auch von unreifen erythropoetischen Zellen im KM (Lammers *et al.*, 2002). Damit ist EpCAM ein weiteres Beispiel für einen Tumormarker, der zunächst als spezifisch für epitheliale Zellen betrachtet wurde, wie bereits E-cadherin and Muzin (MUC-1). Auch diese Oberflächenproteine werden von unreifen erythropoetischen Zellen exprimiert (Brugger *et al.*, 1999; Buhring *et al.*, 1996).

1.2.3 Notwendigkeit der Entwicklung eines sensitiven und spezifischen Markers

Die Entwicklung eines zuverlässigen Nachweises lebender disseminierter Tumorzellen ist von großer Bedeutung. Die Detektion nicht maligner Zellen könnte in der Klinik zu fehlerhaften Diagnosen und ungeeigneten Therapien führen. Die Isolierung nicht maligner Zellen im Labor ist kostspielig und zeitaufwendig.

Die beste Nachweismethode in der Klinik bleibt noch immer die intrazelluläre Zytokeratinfärbung. Für den Nachweis lebender Tumorzellen ist EpCAM einer der besten zur Verfügung stehenden Marker. Aber neben der EpCAM-Expression unreifer erythropoetischer Zellen (s.o.) gibt es weitere Schwierigkeiten, die bei der Anwendung von EpCAM zur Detektion maligner Zellen beim Mammakarzinom auftreten. Nicht alle Zellen dieses Tumors exprimieren EpCAM, obgleich dieses zum Beispiel bei Kolorektal- oder Magenkarzinom beobachtet wurde (Gottlinger *et al.*, 1986; Spizzo *et al.*, 2002). Es ist demnach möglich, dass es disseminierte Tumorzellen gibt, die keine EpCAM Expression zeigen und die folglich unentdeckt bleiben. Außerdem wird in der normalen Brust EpCAM von den Luminal- aber nicht von den Basalzellen exprimiert (MacDougall & Matrisian, 2000). Wenn disseminierte Tumorzellen einen basalen Ursprung haben, was nach der Arbeit von Perou *et al.* vorkommen kann (Perou *et al.*, 2000), bleiben sie auch undetektierbar.

Zusätzliche unspezifische Reaktionen können bei Antikörperfärbungen auftreten, zum Beispiel wenn die Konzentration des benutzten Antikörpers falsch gewählt wird. Apoptotische Zellen werden gelegentlich unspezifisch gefärbt. Die Auswertung der

Färbungen hängt von der Beurteilung und von der Erfahrung jedes Prüfers ab. Damit ist immer ein gewisses Maß an Subjektivität verbunden und Beobachtungsfehler können nicht ausgeschlossen werden.

Bislang ist eine spezifische Identifizierung der Tumorzellen mit Antikörpern gegen Oberflächenproteine unmöglich. Eine Analyse der Genexpression von disseminierten Tumorzellen könnte einen weiteren Einblick in ihre Biologie geben und zur Entdeckung neuer Zielmoleküle für ihre Detektion führen.

1.3 **Analyse der Genexpression**

Die Anzahl der Gene im humanen Genom, die Proteine kodieren, ist umstritten. Die unterschiedlichen Einschätzungen können an den verschiedenen Methoden liegen, aber auch mit ähnlichen methodischen Grundlagen können die Abweichungen voneinander sehr ausgeprägt sein. Zwei Arbeiten, die auf **Expressed Sequence Tag (EST)** Analysen beruhen, sprechen von 142634 Genen (Dickson, 1999), bzw. von maximal 40 000 Genen (Venter *et al.*, 2001). Die jüngste Einschätzung liegt zwischen 20 000 und 25 000 Genen, die Proteine kodieren (Consortium, 2004).

In einer Zelle gibt es verschiedene Genexpressionslevel. So zeigt eine Arbeit über die Quantität der RNA, dass ~20 % der mRNA zahlreiche Transkripte (1000-12 000 Kopien/Zelle) sind (Hastie & Bishop, 1976). Etwa 25 % der mRNA kommt in einer Zahl von 100-1000 Kopien/Zelle vor, die restlichen 50 % sind seltene Transkripte (<13 Kopie/Zelle) (Lewin, 1975). Eine neuere Untersuchung, die mit **Serial Analysis of Gene Expression (SAGE)** durchgeführt wurde, geht mit diesem Ergebnis konform: 75 % der mRNA Gesamtmenge wird von Transkripten mit einer hohen Kopienzahl (>5 Kopie/Zelle) und 25 % von Transkripten mit einer niedrigen Kopienzahl gebildet (<5 Kopie/Zelle). Die Menge der seltenen Transkripte enthält dabei 86 % der verschiedenen mRNA-Typen (Zhang *et al.*, 1997).

Jeder Zelltyp exprimiert unterschiedliche Gene. Meistens zeigen Zellen des gleichen Ursprungs ähnliche Genexpressionsmuster. Differentiell exprimierte Gene sind Gene, die in einem besonderen Zelltyp bei verschiedenen Differenzierungs- und Entwicklungsstadien transkribiert werden, bzw. deren Expression sich im Laufe des Zellzyklus, der Tumorigenese oder einer pharmakologischen Behandlung ändert. Um solche Gene isolieren zu können, stehen mehrere Methoden zu Verfügung, wie *Random Sampling* und Sequenzierung von cDNA Klonen, differentielles Screening von cDNA Library oder Differential Display.

1.3.1 Methoden zur Analyse der differentiellen Genexpression

Mit der Automatisierung des Sequenzierungsprozesses wurde es möglich, die Genexpression durch zufällige direkte Sequenzierung zu erreichen (Frigerio *et al.*, 1995). Die Methode wird auch benutzt, um EST-Banken zu etablieren, welche viele nützliche Informationen zur Identifizierung der Transkripte enthalten (Hillier *et al.*, 1996).

Das humane Genom ist komplex (im Sinne hoher Sequenzdiversität) und die Redundanz mancher Transkripte sehr hoch. Deswegen ist eine zufällige direkte Sequenzierung von cDNA Klonen (*Random Screening*) kein guter Ansatz zur Identifizierung neuer seltener Transkripte. Eine Untersuchung des ganzen Transkriptoms (Gesamtheit aller mRNA Transkripte) ist schwierig, so dass eine quantitative Analyse der Expression unzugänglich ist. 1995 wurde die **Serial Analysis of Gene Expression (SAGE)** Methode veröffentlicht, die eine schnellere Expressionsanalyse erlaubt (Velculescu *et al.*, 1995). Trotz dieser wesentlichen Verbesserung dauert die Bearbeitung sehr lang und die Redundanz bleibt ein Problem. Um dies zu reduzieren und die ständige Sequenzierung von häufig vorkommenden Transkripten zu vermeiden, sind differentielles Screening und/oder Normalisierung/Subtraktions Strategien hilfreiche Ansätze.

1.3.1.1 Differentielles Screening

Eine jüngere Methode der Genexpressionsanalyse ist die **cDNA Array Technologie** (*Nylon* oder *Glass Arrays*). Obwohl die Methode aufwendig und kostenintensiv ist, wird sie immer öfter im Labor verwendet, um verschiedene Zellpopulationen zu vergleichen (Brown & Botstein, 1999). Gute Ergebnisse können sogar mit Einzelzellen erreicht werden (Klein *et al.*, 2002). Leider erlaubt die Methode keine Identifizierung neuer Gene, da sie auf bekannten Sequenzen beruht.

1.3.1.2 Subtraktionsmethoden

Subtraktionsmethoden verwenden Proben, in denen die Konzentration der hoch exprimierten cDNA reduziert wird, ohne niedrig exprimierte cDNA zu beeinflussen. Ein Überschuss von cDNA einer Population (Driver) wird mit der cDNA der zu untersuchenden Population (Tester) hybridisiert. Die nicht-hybridisierte seltene cDNA (Zielsequenzen) wird dann von den hybridisierten häufigen Sequenzen getrennt. Zuerst wurden die Einzelstränge von der Doppelstrang cDNA beim letzten Schritt getrennt, mittels Hydroxylapatit Chromatographie oder Avidin-Biotin Bindung, so dass viel Ausgangsmaterial ($> 20 \mu\text{g}$ von Poly(A)⁺) benötigt wurde. 1993 wurde eine PCR-basierte Methode veröffentlicht, **Representational Difference Analyse (RDA)**, die beim letzten Schritt keine physikalische Trennung der Einzel- und

Doppelstrang cDNA erfordert. Wenig Ausgangsmaterial wird hierbei benötigt (>120 pg), leider löst RDA das Problem der mRNA Redundanz nicht, so dass mehrere Subtraktionsrunden dafür gemacht werden müssen und seltene Transkripte oft verloren gehen (Hubank & Schatz, 1994). 1996 wurde eine Methode veröffentlicht, die das Problem durch einen Normalisierungsschritt der Sequenzfülle überwindet: die **Suppression Subtractive Hybridization (SSH)** (Diatchenko et al., 1996; Diatchenko et al., 1999).

1.3.2 Suppression Subtractive Hybridization (SSH)

Die Methode beruht auf dem 1995 von Siebert and al. etablierten *Suppression PCR Effect* (Siebert *et al.*, 1995). Dabei werden lange Adapter an DNA-Fragmente ligiert. Sie sind so gebaut, dass eine unerwünschte PCR Amplifikation nicht möglich ist. Der Suppressioneffekt kommt zustande, wenn beide Enden eines Einzelstrangs komplementäre Sequenzen tragen. Während jedes Bindungsschrittes ist die Bildung so genannter Haarnadelstrukturen (*Hairpin Like*) begünstigt, so dass die Bindung von Primer-Oligonukleotiden erschwert ist. Wenn eine Oligonukleotidbindung ab und zu erfolgt und zu einer Extension führt, trägt der neue Strang auch die zu einer Haarnadelstruktur führende Sequenz. Dadurch werden die unspezifischen Amplifikationen vermieden und die spezifische Amplifikation von cDNA-Molekülen mit verschiedenen Adaptoren gefördert (**Abbildung 2**).

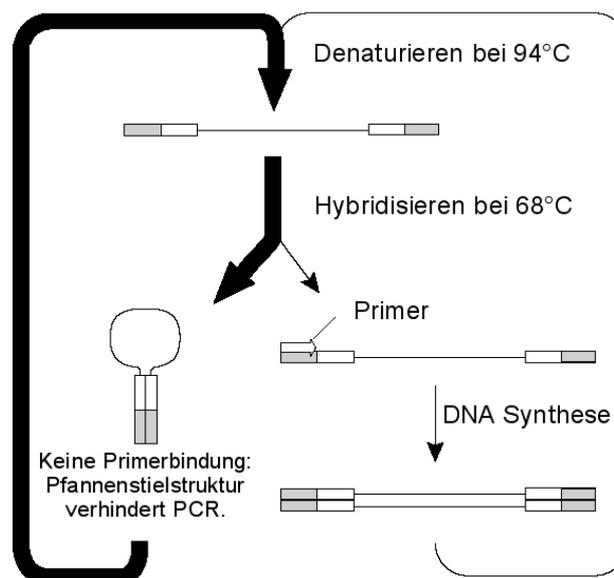


Abbildung 2. Mechanismus der Suppressions PCR.

Die cDNA Moleküle, die eine Adapter-Sequenz an beiden Enden trägt, können nicht mit dem korrespondierenden Primer effizient amplifiziert werden: Die Adapter-Sequenzen sind so konzipiert, dass sie nach jedem Denaturierungs/Hybridisierungs-Zyklus eine Haarnadelstruktur (oder Pfeifenstielstruktur) bilden.

Die SSH besteht hauptsächlich aus zwei subtraktiven Hybridisierungen und einer einzigen PCR Amplifikation (**Abbildung 3**). Für die erste Hybridisierung wird die Tester-cDNA, die mit Adapter 1 oder 2 ligiert ist, mit Überschuss von Driver cDNA gemischt. Die zwei

entstehenden Mischungen werden denaturiert und bei hoher Temperatur getrennt hybridisiert, so dass nur ein eingeschränktes Annealing stattfinden kann. Doppelstrang cDNA, die bei dieser Hybridisierung entsteht, kann bei der darauffolgenden PCR nicht amplifiziert werden, weil entweder eine Adaptersequenz fehlt oder Haarnadelstrukturen gebildet werden (s.o.). Nur Einzelstrang cDNA, die einen Adapter trägt, spielt bei der zweiten Subtraktion eine Rolle. Die zwei Subtraktionen werden dann zusammen mit einem Überschuss an denaturiertem Driver gemischt. Die zweite Hybridisierung dauert so lange an, bis jede Einzelstrang cDNA aus dem ersten Hybridisierungsschritt doppelsträngig wird. Nur aus doppelsträngiger cDNA mit zwei verschiedenen Adaptoren 1 und 2 können PCR-Produkte amplifiziert werden. Die entstehenden PCR-Produkte werden kloniert und weiteruntersucht.

Die SSH zeigt neben der Normalisierung weitere Vorteile, die sie für die Analyse weniger Zellen besonders geeignet macht. So wird zum Einen weniger Ausgangsmaterial als bei vorherbeschriebenen Subtraktionsmethode benötigt und zum Anderen gehen vor allem seltene Transkripte nicht verloren, da nur eine Subtraktionsrunde nötig ist. Die differentiell exprimierten Gene können theoretisch über 1000 Mal durch eine einzige Subtraktionsrunde angereichert werden. Obwohl die Methode kompliziert ist, ist sie ziemlich populär geworden, seitdem ein kommerzieller SSH Kit verfügbar ist (CLONTECH Laboratories, USA).

Wie andere offene Analyse Systeme ("*open systems*," z.B. SAGE) hat die Suppression Subtractive Hybridization (SSH) die Flexibilität neue, unbekannt Sequenzen zu identifizieren. Mit dieser Methode können auch Gene mit einer geringen Expression detektiert werden. Diese niedrigexprimierten Gene spielen eine Rolle für die Bildung unterschiedlicher Phänotypen, weshalb ihre Identifizierung von größter Wichtigkeit für das vollständige mechanistische Verständnis von zellulären Veränderungen ist. Die SSH wurde für die Untersuchung der differentiellen Expression in verschiedenen Versuchssystemen verwendet, bei Bakterien, Tieren sowie beim Menschen. Die Methode wird besonders oft in der Krebsforschung benutzt, unter anderem beim Kolonkarzinom (Hufton *et al.*, 1999; Li *et al.*, 2003), Lungenkarzinom (Difilippantonio *et al.*, 2003) oder Brustkarzinom (Kuang *et al.*, 1998; Xie *et al.*, 2002). Bei all diesen Untersuchungen wurden differentiell exprimierte Sequenzen isoliert, die mit keinen bekannten Sequenzen der öffentlichen Datenbanken identisch waren.

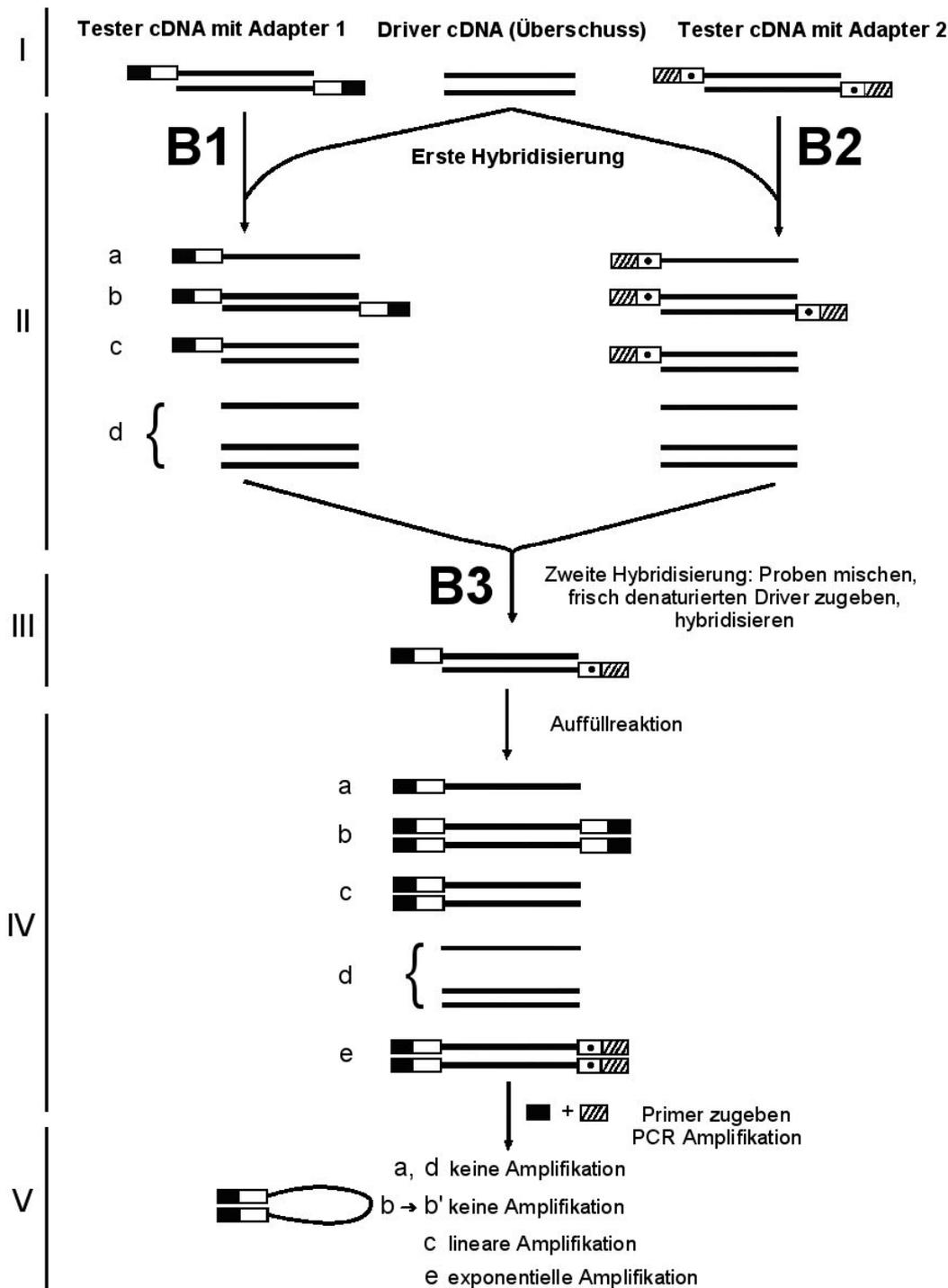


Abbildung 3. Algorithmus der SSH Methode.

(I) Adapter 1 und 2 werden in zwei getrennten Ansätzen an die Tester cDNA ligiert. (II) Jeder Ansatz (Reaktionsgefäß B1 bzw. B2) wird mit einem Überschuss an Driver cDNA gemischt. Bei der ersten Hybridisierung werden die differenziell exprimierten Sequenzen der einzelsträngigen Tester Moleküle normalisiert und angereichert. (III) Die beiden Ansätze werden anschließend zusammen mit einem Überschuss an denaturiertens Driver cDNA gemischt (Reaktionsgefäß B3) und eine zweite Hybridisierung wird durchgeführt. Dabei entstehen verschiedene Hybridtypen (a-d). (IV) Die Enden dieser Hybridisierungsprodukte werden aufgefüllt, so dass die differenziell exprimierten Gene durch die nächsten PCR-Schritte mit geeigneten Primern bevorzugt amplifiziert werden. (V) Zwei weitere PCRs werden durchgeführt, um differenziell exprimierte Gene bevorzugt zu amplifizieren.

1.4 **Zielsetzung der Arbeit**

Das Thema dieser Arbeit ist die Suche nach neuen spezifischen Markern für den Nachweis von Zellen epithelialen Ursprungs im mesenchymalen Gewebe, zum Beispiel für den Nachweis von disseminierten Tumorzellen im Knochenmark, da diese Zellen einer der Schlüssel zum Verständnis der Metastasierung sind. Zu diesem Zweck sollte die Genexpression einzelner disseminierter Tumorzellen analysiert werden. Da einerseits die Zellen, die für die Analyse zur Verfügung stehen, sehr selten sind, und andererseits die Suche nach unbekanntem differentiell exprimierten Genen eine empfindliche Technik erfordert, wurde hierfür die *Suppression Subtractive Hybridization* (SSH) Methode gewählt.

Daraus leiten sich zwei Aufgabenstellungen ab: Erstens die Anpassung der SSH Methode an die zur Verfügung stehende Zell-cDNA von einzelnen disseminierten Tumorzellen. Diese Tester cDNA wurde im Labor mit der Globalen cDNA-Amplifikations Methode isoliert und mit Sicherheit als Material aus Tumorzellen identifiziert. Zweitens die Isolierung und Analyse der differentiell exprimierten Gene, die aus der SSH mit metastatischen Brustzellen als Tester und normalen Knochenmarkzellen als Driver identifiziert wurden. Es sollte versucht werden, ob sich spezifische Gene von Zellen epithelialen Ursprungs, Gene, die nur oder hauptsächlich in Brustzellen exprimiert werden, oder Gene, die in Krebszellen überexprimiert werden, identifizieren lassen.

2 Material und Methoden

2.1 *Material*

2.1.1 Chemikalien

Produkt	Bestell Nr.	Firma
AB-Serum	805 135	Biotest
Agar-Agar	-	Nierle GmbH
Agarose GTG Seakem	50004	Cambrex Bio Sc. Inc
Ampicillin Na-Salt Cryst.	01503	Biomol
Aqua-Roti-Phenol	A980.1	Roth
Blocking Reagent	1096176	Roche
Bovine Serum Albumin (BSA)	B-8667	Sigma
Bromphenolblau	18030	Fluka (Biochemika)
BSA für die Molekularbiologie (20 mg/ml)	711454	Roche
Chloroform	C-2432	Sigma
CSPD [®] : Disodium 3-(4-Methoxy Spiro {1,2-dioxethane-3,2'-(5'-Chloro) Tricyclo [3.3.1.1 ^{3,7}] Decan}-4-yl) Phenylphosphat	1655884	Roche
Cetyltrimethyl Amonium Bromid (CTAB)	195029	ICN Biomedicals Inc.
Human Cot-1 DNA [®] (1mg/ml)	15279-011	Invitrogene
dATP, dTTP, dCTP, dGTP (Sequencing Grade)	27-2035-01	Amersham Biosciences
Dig-Easy Hybridisierungslösung (Dig Luminescent Detection Kit)	1796895	Roche
Dig-UTP (Digoxigenin-11-2'Uridin-5'-Triphosphat, Alkali-labil)	1573152	Roche
Dulbecco's Modified Eagle Medium (DMEM)	31885-023	Gibco
DNA Molecular Weight Marker X (250 µg/ml)	1498037	Roche
Dodecyl sulfat natriumsalz (SDS)	20760	Serva
Dithiothreitol (DTT) (0,1 M)	Y00174	Invitrogen
Epicurian Coli [®] XL1-Blue Electroporation Competent cells	200228	Stratagene
Ethanol 100 %	8006	J.T.Baker
Ethidiumbromid	46067	Fluka (BioChemika)
Ethylendiamintetraessigsäure (EDTA)	ED2SS	Sigma
ExpressHyb Hybridisierungslösung	8015-1	Clontech
Ficoll-400	F-4375	Sigma
Formamid	F-5786	Sigma
Fötale Kälberserum "GOLD" (FCS)	A15-649	PAA Laboratories
Glycerin	G-2289	Sigma
Glycogen für die Molekularbiologie (20 mg/ml)	901393	Roche
Guanidinium Thiocyanat (GT)	50990	Fluka (BioChemika)
Hank's Salzlösung (10x)	L 2045	Biochrom KG
Hefeextrakt (Select Yeast Extract)	30393-029	Gibco
4-(2-Hydroxyethyl)piperazine-1-ethanesulfonic acid (HEPES)	H-3375	Sigma
Herring Sperm DNA (10 mg/ml)	15634.017	Gibco
Hexanukleotid-Mix (10x)	1277081	Roche

Human Placenta DNA	D-4642	Sigma
Igepal CA-630	I-3021	Sigma
Isopropyl-beta-D-thiogalactopyranoside (IPTG)	05684	Biomol
Isopropanol	6752.2	Roth
Kaliumchlorid (KCl)	104936	Merck
Kaliumdihydrogenphosphat (KH ₂ PO ₄)	104873	Merck
L-Glutamin	25030-024	Gibco
LIDS Wasch-Puffer (mRNA Direct Kit)	610.11	Dynal
Lysispuffer (mRNA Direct Kit)	610.11	Dynal
Magnesiumchlorid (MgCl ₂ , 1M)	M-1028	Sigma
Magnesiumsulfat-Heptahydrat (MgSO ₄ ·7H ₂ O)	105882	Merck
Maleinsäure	M-0375	Sigma
NaCl-Lösung 0,9 % (isotonische Spüllösung)	-	B/Braun
Natrium Acetat-Trihydrat (CH ₃ COONa·3H ₂ O)	106267	Merck
Natrium Hydroxid (NaOH)	6771.1	Roth
Natriumchlorid (NaCl)	0278	J.T. Baker
Natriumcitrat-Dihydrat (tri-) (C ₆ H ₅ Na ₃ O ₇ ·2H ₂ O)	106448	Merck
NEBuffer 1 (10x)	B7001S	New England Biolabs
NEBuffer 3 (10x)	B7003S	New England Biolabs
N,N'-dimethyl-Formamid (C ₃ H ₇ NO)	A529.1	Roth
Oligo-dT ₍₂₅₎ magnetischer Beads (mRNA Direct Kit)	610.02	Dynal
PCR Öl	M-5904	Sigma
Penicillin-Streptomycin (100x)	P11-010	PAA Laboratories
Pepton (Select Pepton 140)	30392-021	Gibco
Percoll	17-0891-01	Amersham Biosciences
Phenol-Chloroform-Isoamylalkohol (25:24:1, v/v)	A156.2	Roth
Puffer 1 (10x, mit 17,5 mM MgCl ₂ ELT System)	1742647	Roche
Puffer H (SuRE/Cut Buffer H for restriction enz, 10x)	1417991	Roche
Rapid Ligation Puffer (2x) für T4 DNA Ligase	C6711	Promega
Roswell Park Memorial Institute Medium (RPMI 1640)	31870-025	Gibco
N-laurylsarcosin Natrium Salz (Sarkosyl)	61743	Fluka BioChemica
Streptavidin magnetische Beads (10 mg/ml)	112.05	Dynal
Sybr-Gold	S-11494	Molecular Probes
Tris (C ₄ H ₁₁ NO ₃)	50005	Biomol
Tris/HCl (1 M)	T-3038	Sigma
Trypton/Pepton aus Casein	8952.2	Roth
TWEEN [®] 20	P-9416	Sigma
Wasser (LiChrosolv für Chromatographie, RNase frei)	1.15333.2500	Merk
X-Gal (5-bromo-4-chloro-3-indolyl-β-D-galactoside)	0651745	Roche

2.1.2 Enzyme und Antikörper

Produkt	Bestell Nr.	Firma
Antikörper: 3B10 (anti EpCAM, 1,5 mg/ml)	-	Micromet
Antikörper: Ziege Anti-Maus IgG, Cy3 markiert (1,4 mg/ml)	115-166-071	Jackson Inc.
Anti-Digoxigenin-AP Fab Fragment	1093274	Roche
Bgl I (+ NEBuffer 3)	143 S	New England Biolabs

DNA Polymerase Mix (5U/μl, ELT System)	1759060	Roche
EcoR I (40 U/μl)(+ Puffer H)	0606189	Roche
Mung-Bean Nuklease I (50U/μl) (+ MB Puffer)	1134485	Roche
Rsa I (+ NEBuffer 1)	R0167 S	New England Biolabs
Taq DNA Polymerase (5U/μl)	1435094	Roche
T4 DNA Ligase	0716359	Roche
Terminale Desoxynukleotidyl-Transferase (TdT, 14 U/μl)	E2230Y	Amersham Biosciences
Reverse Transcriptase (RT-Superscript II)	18064014	Invitrogen

2.1.3 Kit und Systeme

Produkt	Bestell Nr.	Firma
QIAprep Miniprep Kit	27106	Qiagen
QIAquick Gel Extraction Kit	28706	Qiagen
Dig Luminescent Detection Kit for Nucleic Acids	1363514	Roche
Dynabeads mRNA Direct Kit	610-11	Dynal
Cloning Matrix System for Human Universal cDNA Library	937840	Stratagene
Expand Long Template PCR System (ELT System)	1759060	Roche
Human Multiple Tissue Expression (MTE™) Array 2	636901	BD Biosciences
Nuc Trap Probe Purification Columns	400701	Stratagene
Push Column Beta Shield Device	400700	Stratagene
pGEM®-T Easy Vector System I	A1360	Promega

2.1.4 Puffer, Lösungen und Medium

Puffer, Lösung oder Medium	Zusammensetzung
Hybridisierungspuffer (4x)	4 M NaCl, 200 mM HEPES (pH:8,3), 4 mM cethyltrimethyl ammonium bromid (CTAB)
Verdünnungspuffer (1x)	20 mM HEPES-HCl (pH:8,3), 50 mM NaCl, 0,2 mM EDTA
PBS (Phosphatbuffer-Saline)	8,5 mM Na ₂ HPO ₄ , 2 mM KH ₂ PO ₄ , 150 mM NaCl, pH:7,4 einstellen
PCR-Puffer (Standard;10x)	500 mM KCl, 100 mM Tris HCl (pH:8,5), 10 mM MgCl ₂ , 1 mM jedes dNTPs
Gel-loading Buffer (10x)	25 % Ficoll-400 + 0,4 % Bromphenolblau in TBE 1x
Percoll 100 %	400 ml Percoll, 36 ml Hank's Salzlösung (10x), pH:7,4 mit HCl 3 M einstellen, Densitometerwert kontrollieren (1,3520), 0,2 μm filtrieren.
Percoll 60 %	300 ml Percoll 100 % + 200 ml NaCl-Lösung 0,9 %, Densitometerwert kontrollieren (1,3450.)

cDNA-Wasch-Puffer+Igepal (1x)	50 mM Tris/HCl (pH:8,0), 75 mM KCl, 10 mM DTT, 0,25 % Igepal.
cDNA-Wasch-Puffer+Tween (1x)	50 mM Tris/HCl (pH:8,0), 75 mM KCl, 10 mM DTT, 0,50 % TWEEN 20.
Tailing-Wasch-Puffer (1x).	50 mM KH ₂ PO ₄ , 1 mM DTT, 0,25 % Igepal.
10 % SDS	100 g SDS (Dodecyl-Sulfat Natrium-Salz) mit 900 ml destilliertem H ₂ O bei 68°C lösen, pH 7,2 einstellen, auf 1 l H ₂ O auffüllen.
20x SSC	175,3 g NaCl, 88,2 g Natriumcitrat, mit 800 ml H ₂ O auffüllen, pH 7 einstellen, auf 1l H ₂ O auffüllen
2x SSC + 0,1 % SDS	50 ml 20x SSC + 5 ml 10 % SDS + 445 ml H ₂ O
1x SSC + 0,1 % SDS	25 ml 20x SSC + 5 ml 10 % SDS + 470 ml H ₂ O
0,5x SSC + 0,1 % SDS	12,5 ml 20x SSC + 5 ml 10 % SDS + 482,5 ml H ₂ O
0,1x SSC + 0,1 % SDS	2,5 ml 20x SSC + 5 ml 10 % SDS + 492,5 ml H ₂ O
Puffer 1 (1x)	4,64 g Maleinsäure und 3,50 g NaCl in 400 ml H ₂ O lösen. Mit NaOH-Plätzchen auf pH 7,5 einstellen
10 % Block	10 g Blocking Reagent mit 100 ml Puffer 1 mischen und zum Lösen in Intervallen von je 30 sek in der Mikrowelle erhitzen (max 10x, ungefähr 3 min).
Blocking-Lösung	5 ml 10 % Block mit 45 ml Puffer 1 mischen
Puffer 1-Tween-Lösung	500 ml Puffer 1 mit 1,5 ml Tween-20 mischen.
Puffer 3 (1x):	5 ml Tris 1M, pH 9,5(Tris 100 mM), 5 ml NaCl 1M (NaCl 100 mM) in 40 ml H ₂ O
TBE (Tris-Borat-EDTA, 10x)	539 g Tris, 275 g Borsäure, 37 g EDTA, auf 5 l H ₂ O auffüllen, pH:8,3.
TE (Tris-EDTA)	100 mM Tris-HCl (pH:8,0), 10 mM EDTA (pH:8,0).
RPMI / DMEM Medium	RPMI 1640 oder DMEM, Foetales Kälberserum (10 %), Penicillin-Streptomycin(1x), L-Glutamin (2 mM)
SOB Medium	2 % Trypton, 0,5 % Hefeextrakt, 10 mM NaCl, 25 mM KCl, 10 mM MgCl ₂ [*] , 10 mM MgSO ₄ [*] . [*] erst kurz vor Gebrauch zugeben.
SOC Medium	SOB Medium + 20 mM Glucose [*]

GT Mix	4 M Guanidinium-Thiocyanat, 25 mM Natriumcitrat, 0,5 % (w/v) Sarkosyl
Bindungspuffer (2x)	10 mM Tris-HCl (pH.7,5), 1 mM EDTA, 2,0 M NaCl.
Bindungspuffer + Igepal	1x Bindungspuffer, 7 % Igepal.
Pepton/PBS (2 % w/v in 1x PBS)	2 g Pepton in 100 ml PBS 1x, pH 7,4 einstellen. Autoklavieren
LB Platte / Ampicillin	500 ml LB Medium mit 6 g Agar Agar mischen und zum Lösen in der Mikrowelle erhitzen. Auf 60°C abkühlen, dann mit 100 µg/ml Ampicillin ergänzen.
LB Platte / Ampicillin / X-Gal/IPTG	Mit zusätzlichen 0,5 mM IPTG und 80 µg/ml X-Gal ergänzen. Platten sofort gießen.
X-Gal (50 mg/ml)	100 mg X-Gal in 2 ml N,N'-dimethyl-Formamid lösen. In Dunkelheit bei -20°C gelagert, ist die Lösung für 2-4 Monate stabil.
IPTG (0,1 M)	1,2 g IPTG in 50 ml Wasser lösen, 0,2 µm filtrieren. Bei 4°C lagern.
LB Medium	50 g Select Pepton, 25 g Hefeextrakt, 50 g NaCl mit 5 l H ₂ O auffüllen, pH 7 einstellen, autoklavieren.

2.1.5 Verbrauchsmaterial

Produkt	Bestell Nummer	Firma
8-Feld Objekt Träger (Chamber Slides)	177402	Nalge Nunc
Elektroporationsküvette (2 mm)	ECU-102	Equibio
Filter (0,2 µm)	FP30/0,2 CA-S	Schleicher & Schuell
Glasskapillar (Micro haematocrit tubes)	7 493 21	Brand
Magneten (Dynal MPC-9600)	120.06	Dynal
Microcon 10, 30 und 100 Säulen	42407, 42410, 42413	Millipore
Nylonmembran	1209299	Roche
Petrischale	632161	Greiner Bio-One
Pipetten (P2, P20, P200, P1000)	-	Gilson
Reaktionsgefäß (0,2 ml-Thermo-Tubes)	AB-0337	Abgene
Reaktionsgefäß (1,5 ml)	7805 00	Brand
Röhrchen (15 und 50 ml)	227 270 / 227 261	Greiner Bio-One
Röntgenfilmen (Biomax Light Film)	153 8966	Kodak
Schlauchfolie (Filter Entwicklung)	129 2126	VWR International

2.1.6 Geräte

Gerät	Bestell Nr.	Firma
Bio-Dot	170-6545	Bio-Rad
Brutschrank für die Bakterien	2737	Köttermann
Brutschrank für die Zellkultur BB 6220	51007413	Heraeus Instruments
Elektroporator	Gene Pulser [®] II System	Biorad
Elektrophorese Geräte	Horizon 11.14	Gibco
Fluoreszenz Mikroskop	DM IRB	Leica
Hybridisierung-Ofen Shake "n" Stack	HBSNSRS220	Thermo Electron Corp.
Mikroinjektor	5242	Eppendorf
Mikromanipulator	5171	Eppendorf
Phosphoimager "Storm 860"	63-0035-64	Amersham Biosciences
Roller Mixer SRT1	SRT1	Stuart Scientific
Stratalinker [®] 2400 UV Crosslinker	400076	Stratagene
Thermocycler "PE GeneAmp 2400"	-	Perkin Elmer
Thermocycler "PTC 200"	PTC-0200 + ALD-1244	MJ Research
Thermocycler "Robocycler [®] 96 Gradient "	400886	Stratagene
Tischzentrifuge 5417C	5417 000.315	Eppendorf
Tischzentrifuge 5402	5402	Eppendorf
Zentrifuge Hettich	Rotix/RP	Hettich

2.1.7 Oligonukleotide

Alle Oligonukleotide wurden von Metabion, Martinsried, synthetisiert.

(N = A, C, G oder T; V = A, G, oder C)

T_m: Errechnete Schmelztemperatur [$T_m = 69.3 + GC\% \cdot 0.41 - 650 / (nA + nC + nG + nT)$]

PCR: Verwendete Temperatur (v = variable experimentelle Bedingungen)

Name	Sequenz	T _m	PCR
<u>PCR Primer</u>			
α-Tubulin 1 (Tub-1)	CAC TCT GAT TGT GCC TTC ATG	55,9	58
α-Tubulin 2 (Tub-2)	GTC AAC ATT CAG GGC TCC AT	57,3	58
α-Tubulin 3 (Tub-3)	TGC TAC AGA AAG CTG TTC ATG	57,9	58
β-Aktin-3'	CTC CTT AAT GTC ACG CAC GAT TTC	61,0	58
β-Aktin-5'	GTG GGG CGC CCC AGG CAC CA	69,6	58
089-3'	CAT GCC CAT TGG CCT TGA	56,0	55
089-5'	AGG CTT CAG CAG GCA CTT	56,0	55
090-3'	ATC TGG AGC AGA GTG GGA	56,0	60

090-5'	TCC CAC TCT TGC CAA AGC		56,0	60
092-3'	CGC GGT CCT ACA ACC CCA		60,5	60
092-5'	GCC TCT AGG CCT GAC TGC		60,5	60
121-3'	CTT TAT TTC AGT ATG AGG		46,8	45
121-5'	ACT TTG TGG AAG GAA AAG		49,1	45
121bis-3'	GGC TCC TTG TGT TTA TAC TC		55,3	55
121bis-5'	TGA GAC TAG ATT GGT ACC AG		55,3	55
131-3'	TAA CAT TTT GCC ACA TTT GCT		52,0	52
131-5'	TTG GAA GTG AGG TAA ACA G		52,4	52
132-3'	TAC AAC ATC CCG CAC GGG		58,2	49
132-5'	AGT CAG AGA AGC TAT AAG AGG		58,2	49
155-3'	GGT CGA AGT TGA CGA ACT		53,7	60
155-5'	GGA TTC TGC TGT CAG TTA		51,4	60
155bis-3'	GTG ATG ACT TTT TGA CCG TG		55,3	55
155bis-5'	GAA TTT GAG AAT GTC GAC GC		55,3	55
157-3'	GGA AAT TCC TGG AAA GCA C		54,5	58
157-5'	GTC TAG AGC CTT CGT AG		52,8	58
195-3'	ACC TGA CCT GGG CCA CTC		58,2	55
195-5'	GGA GAA AAC CTA TGG TGG C		56,7	55
208-3'	AAT GAG GTC CAT TTG CTA GG		55,3	55
208-5'	GGT CTC ATT CTC TAC CTT CT		55,3	55
217-3'	AAA GTG ATG ATG AGG AGC		51,4	52
217-5'	AGC TCG GTG AGG ATT CA		52,8	52
218-3'	GTG ATG ATG AGG AGC AAG G		56,7	58
218-5'	GTC TAG CTC GGT GAG GAT T		56,7	58
219-3'	GCT CAA GTT CTC GGC CTT C		58,8	60
219-5'	GTA ACC GGG CGC TGC CTT		60,5	60
267-3'	GCT GCT CCT TCT TGG ATC		56,0	58
267-5'	GTC TAG ACA GGA CGA ACG		56,0	58
288#1	AAG AAG GTA TCC TTT AGT TCC	(AL045287-5-1)	54,0	v
288#2	TGG TGA AAT CTC GGT TCA CT	(AL045287-5-2)	55,3	v
288#3	CTT GGC GAT AGG TGA TTG T	(Chr3-161130970)	54,5	v
288#4	GAC AAG CAG ATG AAT GGT GA	(Chr3-161133441)	55,3	v
288#5	ACA ATC ACC TAT CGC CAA	(Chr3-161130952)	54,5	v
288#6	GAC AGC CAT CTT ACA GAG G	(AL045287-3-1)	56,7	v
288#7	GGT GAG GTG AAA TGT TGC A	(AL045287-3-2)	54,5	v
288#8	TCC AAA TGC TTT CTT CAA TGC	(AL045287-3-3)	54,0	v
288#9	GTG AGA CCA TCG CCT ATA G		56,7	v
288#10	CTA GAT GGG ACA ATC GCC		56,0	v
288#11	ACC CAC CAA TTC CAG ACA C		56,7	v
288#12	GAT GAA TGG TGA AAC GAT AC		53,2	v
288#13	CGG GAA GAT AAT CTT CCT GC		57,3	v
288-15C	(CCC) ₅ AGT TCA GCC GAG GGT GGT		81,9	v
288-3'	TCT AGA GCG GCG AGA CAA T		56,7	58
288-5'	GGA TGG CAC GGG AAG ATA A		56,7	58
314-3'	GCC TGT ATG TGT CCT GAA		53,7	55
314-5'	CAC CTC TAT TGG GCA CAT		53,7	55
337-3'	TAG AAG GTG ATA TCA ATC CG		53,2	56
337-5'	TCA GGT TTA AAG TTT CAG GTT		52,0	56
346-3'	TCT AGA TGC ACC TAG AAC A		52,4	55
346-5'	ACC TGA TCA TAA TTC CTG G		52,4	55

356-3'	CCG CCC AGC TCT GCC TCA	62,8	60
356-5'	AAG TGC CGA GAA GTT CGC TC	59,4	60
376-3'	TCC TCA TGC GAT CCT CCT	56,0	54
376-5'	TCT AGA TGG GCA GTG AAG	53,7	54
383-3'	GCT TAG AAG TGG GAT TCC TG	60,5	60
383-5'	AGG ATG GCT TGA GCC CGG	52,4	60
Cytokeratin-7 3'	CGG TAG GTG GCG ATC TCG A	50,0	52
Cytokeratin-7 5'	AGA GCT GCA GTC CCA GAT C	48,0	52
EF1- α 3'	CTG TGT CGG GGT TGT AGC CA	61,4	58
EF1- α 5'	TGC CCC AGG ACA CAG AGA CT	61,4	58
EGP 3'	CAC ATC AGC TAT GTC CAC ATC	57,9	58
EGP 5'	GCG AGT GAG AAC CTA CTG	55,6	58
Lib1	AGT GGG ATT CCT GCT GTC AGT	59,8	59
M13 Forward	TTT TCC CAG TCA CGA C	49,2	51
M13 Reverse	CAG GAA ACA GCT ATG AC	50,4	51
Mage4-3'	CAC AGA ACT GTG TGA CTG CT	57,3	58
Mage4-5'	GCG TGA AGC TTT GAG AG	52,8	58
NT1	ACA GGA CTG GCA GGC AGC T	61,0	v
NT10	GCC TTAAGC AAG GCT TGT TT	55,3	v
NT11	TCC TTC TGA TGT TCG GAT GT	55,3	v
NT12	CCA TTG CAG ATT GGA CAG GG	59,4	v
NT13	TCA CCA GGT TCA GAA CCA AT	55,3	v
NT2	GTG AAT GCA CCA ATC GAC AC	57,3	v
NT3	TGG TCT GTT TTG ACA GGG C	56,7	v
NT4	CCC ACC AGA AGG AAG AAT CT	57,3	v
NT5	ATT GGT GGG TTC TTC GTC TC	57,3	v
NT6	TTA GTC CTA GAA AGG CAG ACT G	58,4	v
NT7	CAG TCT GCC TTT CTA GGA CTA A	58,4	v
NT8	CAT CAG TGG AAC TTT TCT TC	53,2	v
NT9	CCA GTT TGA ACT GAG CTG AA	55,3	v
PBP-3'	TCT GCT TGG CCT CGC CAT	58,2	60
PBP-5'	CTG TAT TTG GGA TCC TTC CT	55,3	60
Pseudo-3'	CTG AGG AAA AGA AAT CAA CC	53,2	58
Pseudo-5'	TTC GCG TAC AGG GCT ATC	56,0	58
Tapasin-3'	CTG CAA CTC CTG GGC TGA	58,2	60
Tapasin-5'	AGC AAG CTC CAG GGT GAC	58,2	60

Primer für die cDNA Synthese

288-Fish-Bio	Biotin-AGT TCA GCC GAG GGT GGT	58,2	44
288-Fish-Nest	CGG AGA GGT TTG GCC ACC CAG CAG GAA G	72,4	44
CFL5c8	(CCC) ₅ GTC TAG ANN NNN NNN	77,7	44
CFL5cT	(CCC) ₅ GTC TAG ATT (TTT) ₄ TVN	72,8	44
CP2	TCA GAA TTC ATG (CCC) ₅	74,1	68
CPBgl	TCA GAA TTC ATG CCG CCC CCC CGG CCC	74,1	63
TLib	GTC TAG AAG TGG GAT TCC TGC TGT CAG (TTT) ₅ TVN	68,7	44

SSH Adaptoren und Primer

Eco 12 I	AAT TCC TGC CCG	39,0	-
Eco 12 II	AAT TCG CCC TCC	39,0	-
Eco 43 II	TGT AGC GTG AAG ACG ACA GAA AGG TCG CGT GGT GCG GAG GGC G	79,9	-

Eco 44 I	GTA ATA CGA CTC ACT ATA GGG CTC GAG CGG CTC GCC CGG GCA GG	80,6	-
P1(-22)	GTA ATA CGA CTC ACT ATA GGG C	58,4	66
P1-30	GTA ATA CGA CTC ACT ATA GGG CTC GAG CGG	69,5	66
P1-33	GTA ATA CGA CTC ACT ATA GGG CTC GAG CGG CTC	71,9	68
P2(-22)	TGT AGC GTG AAG ACG ACA GAA	57,9	66
P2-30	TGT AGC GTG AAG ACG ACA GAA AGG TCG CGT	69,5	66
P2-33	TGT AGC GTG AAG ACG ACA GAA AGG TCG CGT GGT	71,9	68
PN1	TCG AGC GGC TCG CCC GGG CAG G	73,3	68
PN1-30	CGA CTC ACT ATA GGG CTC GAG CGG CTC GCC	75,0	68
PN2	AGG TCG CGT GGT GCG GAG GGC G	71,4	68
PN2-30	GTG AAG ACG ACA GAA AGG TCG CGT GGT GCG	72,2	68

2.2 Methoden

2.2.1 Isolierung einzelner Zellen

2.2.1.1 Dichtegradienten-Zentrifugation

Die Knochenmarkaspirate werden in Hank's Salzlösung gewaschen und 10 min. bei 170 g zentrifugiert. Die Fettphase und die Thrombozyten werden abpipettiert, das Restvolumen (ca. 10 ml) resuspendiert und dann vorsichtig in ein 15 ml Röhrchen auf 5 ml Percoll 60 % pipettiert. Die benötigte Percoll Konzentration wird mit steriler, isotonischer NaCl-Lösung eingestellt. Durch eine 20 min. Zentrifugation (RT, 1000 g) sedimentieren die Erythrozyten und Granulozyten, während sich die mononukleären Zellen in der Interphase anreichern. Die Phase mit den *Peripheral blood mononuclear cells (PBMC)* wird abgenommen, in ein frisches 50 ml Röhrchen überführt und in 1x PBS gewaschen (10 min., RT, 540 g). Anschließend wird das Pellet in 1 ml 1x PBS resuspendiert und die Zellzahl in einer Neubauerzählkammer bestimmt.

2.2.1.2 EpCAM Färbung

Eine Million angereicherte Zellen wird in ein steriles 1,5 ml Reaktionsgefäß überführt und pelletiert (4°C, 5 min, 540 g). Das Pellet wird sofort mit 189 µl Pepton/PBS (2 % w/v in 1x PBS), 10 µl AB-Serum und mit 1 µl des Anti-EpCAM Antikörpers mAK 3B10 (Endkonzentration: 7,5 µg/ml) resuspendiert. Die Zellen werden für 10 min. bei 4°C unter Schütteln inkubiert, dann wird erneut mit Pepton/PBS gewaschen, um ungebundene Antikörper zu entfernen. Als Sekundärantikörper dient ein Cy3-markierter Ziege-anti-Maus Antikörper. Von diesem werden 4 µl mit 186 µl Pepton/PBS und 10 µl AB-Serum zusammen gemischt und zu den Zellen gegeben (Endkonzentration: 7 µg/ml). Die Behandlung erfolgt wie für den

Primärantikörper. Anschließend wird das Zellpellet in 200 µl PBS aufgenommen und bei 4°C und Lichtschutz bis zum nächsten Schritt aufbewahrt.

Die Zellen werden auf drei Felder von einem 8-Feld Objektträger verteilt. Wichtig ist hierbei, dass die Felder zunächst mit BSA bestrichen werden, um ein Anhaften der Zellen auf der Glasfläche zu vermeiden, was das folgende Picken wesentlich erleichtert. Nach der Sedimentation werden die Zellen mittels Fluoreszenzmikroskopie untersucht. Wenn eine Zelle als positiv erkannt wird, wird sie mit einer Glaskapillare eingesaugt und in ein neues Zell-freies Feld überführt, das mit PBS bedeckt ist. Die Zelle wird mit einer P2 Pipette in ein PCR-Gefäß transferiert, welches 10 µl magnetische Beads in Lysispuffer enthält. Mit Hilfe Oligo-dT₍₂₅₎-Beads wird im Folgenden die mRNA isoliert und diese anschließend in cDNA umgeschrieben (s.u.).

2.2.2 Globale cDNA-Amplifikation

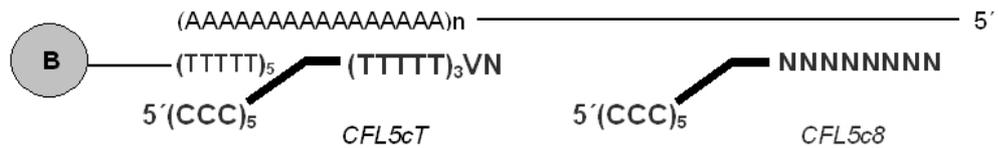
Der Zellaufschluss erfolgt in einem LiDS-haltigen Lyse-Puffer. Bei der mRNA-Isolierung macht man sich den Poly-A-Schwanz der mRNA zu Nutze, über den diese an Oligo-dT₍₂₅₎-Nukleotide der magnetischen Beads gebunden werden (**Abbildung 4**, Schritt 1). Mittels eines Magneten kann dann die mRNA von der DNA und dem Zelldebris getrennt werden. Nach mehreren Waschschritten, bei denen die genomische DNA, die sich in den Überständen befindet, gesammelt werden kann, werden die Reagenzien für die cDNA-Synthese zugegeben. Zunächst lässt man bei Raumtemperatur die cDNA-Synthese-Primer binden. Bei den Primern handelt es sich um die Oligonukleotide CFL5cT und CFL5c8. Das CFL5cT-Oligonukleotid kann durch die VN-Nucleotide an das 5'-Ende des Poly-A-Schwanzes der mRNA binden. Die Anlagerung des CFL5c8-Primers an die mRNA erfolgt auf Grund seiner acht degenerierten Nukleotide zufällig (Schritt 2). Von diesen Primern ausgehend startet dann die Reverse Transkriptase die cDNA-Synthese (Schritt 3).

Auf die Denaturierung der RNA/cDNA Hybridmoleküle bei 94°C folgt das G-Tailing (Schritt 4). Die Terminale Desoxynukleotidyl-Transferase (TdT) besitzt die Eigenschaft am 3'-Ende von einzelsträngiger DNA unspezifisch Desoxynukleotide mit sehr hoher Effizienz anzuhängen (Effizienz bei doppelsträngiger DNA deutlich niedriger). Auf diese Weise erhält man durch Anhängen nur eines der vier Nukleotide einen homopolymeren Schwanz variabler Länge. In dem hier beschriebenen Protokoll werden der TdT nur G-Nukleotide zur Verfügung gestellt, die diese dann am 3'-Ende der cDNA anhängt. Die TdT muss anschließend bei 70°C inaktiviert werden. Der Einbau der G-Nukleotide endet aufgrund sterischer Behinderung der TdT nach etwa 15-30 Basen (Schritt 4).

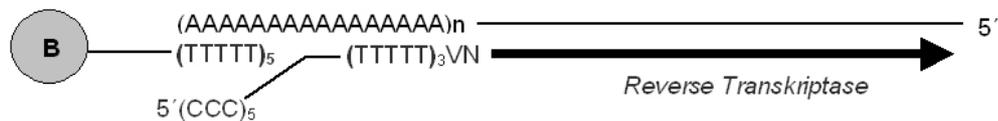
1. mRNA Isolation:



2. PRIMER-Anlagerung:



3. cDNA-SYNTHESE:



4. RNA LYSE + G-TAILING:



5. GLOBALE cDNA AMPLIFIKATION:

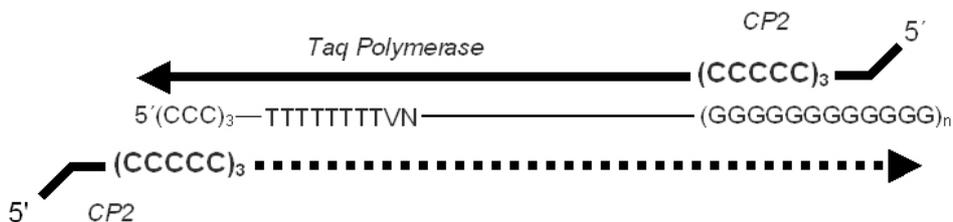


Abbildung 4. Schema der reversen Transkription und der globalen cDNA-Amplifikation. Die mRNA wird durch magnetische Beads (B) isoliert (1). Zunächst lässt man bei Raumtemperatur die cDNA-Synthese-Primer CFL5c8 und CFL5cT binden (2). Die reverse Transkription erfolgt von beiden Primern aus (3; CFL5c8 wird in Abbildung ab dieser Schritt weggelassen). Nach der Lyse der mRNA wird ein Poly-G Schwanz mit der Terminalen Desoxynukleotidyl-Transferase (TdT) eingefügt (4). Mit dem CP2 Primer können alle Sequenzen amplifiziert werden (5).

Die Amplifikation der cDNA erfolgt anschließend mit Hilfe einer PCR. Die verwendeten Primer CP2 binden dabei im ersten Zyklus an das von der Terminalen Transferase angefügte G-Ende und im zweiten Zyklus an die neu entstandene G₍₁₅₎-Sequenz, die zur Poly-C-Sequenz der cDNA-Synthese Primer komplementär ist. In den folgenden Zyklen findet eine exponentielle Amplifikation statt (Schritt 5).

Um Kontaminationen der Proben und dem Abbau der RNA vorzubeugen, wird die Herstellung der primären Amplifikate in einem separaten und RNase-freien Raum durchgeführt. Alle Schritte müssen mit Filter-Pipettenspitzen pipettiert werden.

2.2.2.1 Vorbereitung der Dynal-Beads

Für eine isolierte Zelle werden 10 µl Oligo-dT₍₂₅₎ magnetische Beads 1:1 mit LiDS-Waschpuffer gemischt und an den Magneten gegeben. Der Überstand wird abgenommen und die Beads erneut mit 20 µl LiDS-Waschpuffer gewaschen, wieder an den Magneten gegeben und der Überstand erneut abgenommen. Die Beads werden in 10 µl Lysispuffer aufgenommen und anschließend auf Eis gestellt. Werden die vorbereiteten Beads nicht sofort verwendet, können diese bei -20°C gelagert werden.

2.2.2.2 RNA-Isolierung

Die isolierten Zellen werden in 10 µl Beadslösung gegeben. Die Bindung der mRNA an die Beads erfolgt unter Rotation bei Raumtemperatur für 30 min. Danach werden 10 µl cDNA-Igepal-Waschpuffer dazu pipettiert und jedes Gefäß wird in einen speziellen Ständer mit einem Magneten gestellt. Nach ungefähr 1 min. sind die mit der mRNA gekoppelten Beads an den Magneten gebunden. Die genomische DNA befindet sich im Überstand. Dieser wird gesammelt, wobei darauf geachtet werden muss, dass keine Beads gewegewaschen werden. Anschließend werden in 20 µl cDNA-Tween-Waschpuffer die mit mRNA gekoppelten Beads aufgenommen, in ein frisches Reaktionsgefäß pipettiert und an den Magneten gegeben. Erneut werden sie mit 20 µl cDNA-Igepal-Waschpuffer resuspendiert und an den Magneten gegeben. Es wird mit Zugabe von Mix 1 sofort fortgefahren.

2.2.2.3 cDNA-Synthese

Die cDNA-Synthese wird unter Rotation in einem Hybridisierungssofen durchgeführt, damit die Beads, an die die mRNA gebunden ist, nicht sedimentieren. Mix 1 und Mix 2 werden angesetzt während die RNA auf dem Roller an die Beads hybridisiert. Der Hybridisierungssofen wird auf 44°C vorgewärmt.

10 µl Mix 1 enthalten 1x RT-Puffer, 10 mM DTT, 1% Igepal, 60 µM CFL5c8 und 30 µM CFL5cT. Mix 2 wird zunächst ohne Reverse Transkriptase (RT-Superscript II) angesetzt und diese wird erst vor Zugabe von Mix 2 dazu pipettiert. 10 µl Mix 2 enthalten 1x RT-Puffer, 10 mM DTT, 1 mM dNTPs und 100 U RT-Superscript II.

Jeweils 10 µl von Mix 1 zu den Beads geben und anschließend die Primer für die cDNA-Synthese binden lassen. Nach Zugabe der Reversen Transkriptase zu Mix 2, ebenfalls 10 µl

zu den Proben geben und auf Eis stellen. Die Reaktionsgefäße rotieren im Hybridisierungssofen langsam 45 min. bei 44°C.

2.2.2.4 G-Tailing

Während der cDNA-Synthese wird der Tailing-Mix vorbereitet und eine PCR-Maschine auf 94°C vorgeheizt. 10 µl Tailing-Mix enthalten 4 mM MgCl₂, 0,1 mM DTT, 0,2 mM dGTP und 10 mM KH₂PO₄.

Nach der cDNA-Synthese muss man darauf achten, dass keine Beads mehr an der Gefäßwand kleben, eventuell müssen sie abzentrifugiert werden. Die Beads werden am Magneten einmal mit 20 µl Tailing-Waschpuffer gewaschen, dann in 10 µl Tailing-Mix resuspendiert und mit 40 µl PCR-Öl überschichtet. Die Reaktion wird 4 min. bei 94°C inkubiert und anschließend von 94°C direkt auf Eis gestellt. Pro Reaktionsgefäß werden 0,8 µl TdT (14 U/µl) zugegeben, gemischt und 1 Stunde bei 37°C und dann mindestens 1 Stunde bei 22°C inkubiert (spätestens am nächsten Morgen entnehmen). Die Inaktivierung der TdT erfolgt bei 70°C für 5 min. im Thermocycler.

2.2.2.5 PCR

Für die PCR werden zwei PCR-Stammlösungen vorbereitet: 35 µl PCR-Stammlösung 1 enthalten 4 µl Puffer 1, 7,5 µl Formamid 20 % und 24 µl Wasser. 5 µl PCR-Stammlösung 2 enthalten 2,5 µl CP₂-Oligonukleotid (24 µM), 1,75 µl dNTP (10 mM) und 1,5 µl Pol.Mix (3,5U/µl, Expand Long Template System). Jeweils 35 µl vom PCR-Stammlösung 1 werden pro Tube unter das Öl pipettiert und in die PCR-Maschine gestellt. Bei 78°C muss man die PCR unterbrechen und jeweils 5,5 µl vom PCR-Stammlösung 2 dazugeben, mischen und nach letzter Probe Pause lösen („Hot-start“).

Dieser Thermocycler wird wie folgt programmiert:

	1 Zyklus (Hot Start)	19 Zyklen	10 Zyklen	1 Zyklus
78°C	30 sek.			
94°C		15 sek.	15 sek.	
65°C		30 sek.		
68°C		2 min.	3,5 min.+ 5 sek./Zyklus	7 min.

Als erste Kontrolle der Globalen cDNA-Amplifikation werden 5 µl des Reaktionsgemisches auf ein 1,5 % Agarosegel aufgetragen. Die Elektrophorese wird nach Standardprotokollen durchgeführt (Sambrook & Russel, 2001). Bei erfolgreicher Amplifikation sollte man einen Schmier von 300 bp bis zu 1000 bp beobachten können. Es ist zu bemerken, dass die Negativkontrolle ebenfalls einen Schmier zeigt, der jedoch schwächer ist und kürzere Fragmente aufweist.

2.2.2.6 Kontroll-PCR (β -Aktin, EF1- α , EGP) / Standard Protokoll

Die Qualität der synthetisierten cDNA wird mit drei PCR Reaktionen für das zytoplasmatische β -Aktin, den Elongation Factor EF-1-alpha und das epitheliale Glycoprotein EGP/EpCAM getestet. β -Aktin (Accession # M10277) ist eine Komponente des Zytoskeletts und die alpha Untereinheit des Elongation Factor-1 (Accession # J04617) ist an der Bindung von tRNA an 80S Ribosomen beteiligt. Die mRNA von diesen zwei Genen sollte in jeder humanen Zelle zu isolieren sein. Positive Ergebnisse bei der PCR sind ein Hinweis darauf, dass eine Zelle erfolgreich gepickt, isoliert und anschließend das Transkriptom erfolgreich amplifiziert wurde. Ein positives Ergebnis für EpCAM (Accession # M32306) zeigt, dass die aus den Knochenmark isolierte Zelle das Oberflächenmolekül EpCAM exprimiert hat und dass die Antikörperbindung spezifisch war (siehe auch 1.2.1 und 1.2.2).

Die PCR wird nach einem **Standard Protokoll** durchgeführt. Als PCR-Stammlösungen werden für jeden Ansatz 1 μ l PCR-Puffer (10x), 0,25 μ l BSA (20 mg/ml), 7,25 μ l Wasser und 0,1 μ l Taq Polymerase (5 U/ μ l) mit 0,5 μ l von jedem entsprechenden 3'- und 5'-Oligonukleotid (8 mM) gemischt. Zu diesen 9,5 μ l Ansatz wird 1 μ l von der Probe (d.h. von der cDNA Globale Amplifikation Reaktion aus 2.2.2.5, 1/10 im Wasser verdünnt) zugemischt.

Der Thermocycler wird folgendermaßen programmiert:

	1 Zyklus	15 Zyklen	45 Zyklen	1 Zyklus
94°C	2 min.	40 sek.	40 sek.	
58°C		30 sek.	30 sek.	
72°C		20 sek.	30 sek.	2 min.

Die PCR wird im Robocycler[®] durchgeführt und anschließend auf ein 1,5 % Agarosegel aufgetragen.

2.2.3 Subtraktion

2.2.3.1 Driver Vorbereitung

2.2.3.1.1 Reamplifikation

Der Driver wird mit 1 μ l unverdünnter Probe in Puffer 1 (Roche), 0,4 mM dNTPs, 3% Formamid, 2,4 μ M CPBgl Oligonukleotid und 1,75 U DNA Polymerase Mix in einem Gesamtvolumen von 50 μ l amplifiziert. Das CPBgl Oligonukleotid beinhaltet eine Restriktionsschnittstelle für Bgl I.

Die Bedingungen für dieses PCR Programm sind:

	1 Zyklus	15 Zyklen	10 Zyklen	1 Zyklus
94°C	2 min.	15 sek.	15 sek.	
63°C		30 sek.	30 sek.	
68°C		3 min.	3,5 min.+ 5 sek./Zyklus	7 min.

Die Konzentration des PCR Produkts wird auf einer Agarose-Ethidiumbromid-Platte gegen eine Standard-Verdünnungsreihe bestimmt.

2.2.3.1.2 Rsa I und Bgl I Verdau

Vom Driver werden 5 µg mit 50 U Rsa I Restriktionsenzym, Neb Puffer 1 und Wasser in 50 µl 5 Stunden bei 37°C verdaut. Das Enzym wird bei 65°C 20 min. inaktiviert. Danach schließt sich ein Bgl I-Verdau 3 Stunden bei 37°C an, mit 100 U Bgl I, 7,5 µl Neb Puffer 3, Wasser ad 100 µl.

Der verdaute Driver wird auf einer Microcon 10 Säule gereinigt, die Konzentration bestimmt, auf 500 ng/µl in Wasser verdünnt und bei -20°C gelagert.

2.2.3.2 Tester Vorbereitung

2.2.3.2.1 Tester Amplifikation

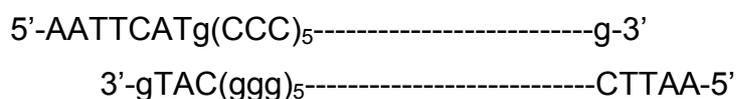
Der Tester wird mit dem CP2 Oligonukleotid, dessen Sequenz eine Schnittstelle für das EcoR I Restriktionsenzym beinhaltet, amplifiziert. Die PCR wird mit 1 µl unverdünnter Probe in Puffer 1, 0,4 mM dNTPs, 3% Formamid, 2,4 µM CP2 Oligonukleotid und mit 1,75 U DNA Polymerase Mix in einem Gesamtvolumen von 50 µl durchgeführt.

Die Bedingungen für das PCR Programm sind:

	1 Zyklus	12 Zyklen	3 Zyklen	1 Zyklus
94°C	2 min.	15 sek.	15 sek.	
68°C	4 min.	4 min.	4 min.+ 10 sek./Zyklus	7 min.

2.2.3.2.2 EcoR I Verdau

5 µg vom Tester werden mindestens 90 min. bei 37°C in einen Volumen von 50 µl mit 50 U EcoR I in Puffer H verdaut, dann 20 min. bei 65°C inaktiviert. Aus dieser Reaktion ergeben sich die folgenden Produkte, die später mit den Adaptoren ligiert werden :



Zugabe von 200 µl Verdünnungspuffer (1x) und Erhitzen auf 68°C für 7 min. erfolgt die Inaktivierung.

2.2.3.3 Kontrolle für die Subtraktion

Als Kontrolle für die Subtraktion dient die folgende Reaktion: 100 ng vom Tester werden gleichzeitig mit 2,5 µl Adapter A1 (50 µM) und Adapter A2 (50 µM) ligiert und wie oben beschrieben inaktiviert. Die Reaktion läuft wie für die Subtraktionsreaktion beschrieben ab. Sie wird als "nicht subtrahiert" (nicht sub.) bezeichnet.

2.2.3.4 Suppressive PCR

Um differentiell exprimierte Sequenzen zu amplifizieren, wird eine PCR auf alle Subtraktionsprodukte durchgeführt. Weil die Hybridisierungsprodukte **a** und **d** keine Bindungsstelle für die Primer haben, können sie nicht amplifiziert werden (**Abbildung 3**). Dank der speziellen Sequenz der Adaptoren bilden die meisten Moleküle vom Typ **b** eine Haarnadelstruktur aus, die ihre exponentielle Amplifikation verhindert (**Abbildung 2**, Seite 17). Die Moleküle vom Typ **c** haben nur eine einzige Primer-Bindungsstelle und können daher nur linear amplifiziert werden. Nur bei Hybridisierungsprodukten vom Typ **e** erlauben die zwei verschiedenen Adaptoren eine exponentielle Amplifikation.

Anschließend an die Amplifikation wird noch eine Nested-PCR durchgeführt: die Produkte der suppressiven PCR werden als Matrize für eine zweite Amplifikation mit anderen Primern verwendet. Das zweite Primerpaar liegt dabei zwischen dem ersten. Auf diese Weise werden die unspezifischen Amplifikationsprodukte der ersten Reaktion nicht amplifiziert. So wird der Hintergrund reduziert und die differentiell exprimierten Sequenzen angereichert. Nach diesem Schritt kann man die amplifizierten cDNAs direkt in einen *T/A-Cloning* Vektor ligieren.

2.2.3.4.1 Suppressive PCR (PCR 1)

Zwei PCR Reaktionsansätze werden vorbereitet: 19 µl PCR-Ansatz 1 enthalten 2 µl Puffer 1 (10x), je 1 µl von P1-30 und P2-30 Oligonukleotid (8 µM), 1 µl dNTPs (10 mM) und 14 µl Wasser. 5 µl PCR-Ansatz 2 enthalten 0,5 µl Puffer 1 (10x, Roche), 0,5 µl Taq Polymerase (5 U/µl) und 4 µl Wasser. Jeweils 19 µl vom PCR-Ansatz 1 werden pro Tube pipettiert und 1 µl der Subtraktion in Verdünnungspuffer dazu gegeben. Die Reaktionsansätze werden anschließend mit Öl überschichtet und in einer PCR-Maschine bis 75°C erhitzt. Dann wird die Reaktion angehalten, um 5 µl vom PCR-Ansatz 2 dazu zu geben. Die Bedingungen für das PCR Programm sind:

	Hot Start + Adapter Auffüllung	27 Zyklen	1 Zyklus
75°C	7 min.		
94°C		30 sek.	
66°C		30 sek.	
72°C		1,5 min.	7 min.

Die PCR-Produkte werden durch eine Gelelektrophorese kontrolliert.

2.2.3.4.2 Nested-PCR (PCR 2)

Die Nested-PCR wird wie oben beschrieben angesetzt, wobei hier die Oligonukleotide PN1-30 und PN2-30 (8 µM) verwendet werden. Als Template dient 1 µl einer 1:10 in Wasser Verdünnung des PCR-Produkts von der vorigen PCR1. Die 5 µl von Mix II werden erst zugegeben wenn die Temperatur im Reaktionsgefäß 72°C erreicht. Die Bedingungen für das PCR2 Programm sind:

	Hot Start	12 Zyklen	1 Zyklus
94°C		30 sek.	
68°C		30 sek.	
72°C	1 min.	1,5 min.	7 min.

Die Produkte werden durch eine Gelelektrophorese kontrolliert.

2.2.3.4.3 Test für die Effizienz der Subtraktion.

2.2.3.4.3.1 PCR

Standard-PCR-Reaktionen (s. 2.2.2.6) werden mit 1 µl den Produkten von der PCR2 durchgeführt. Getestet wird die relative Quantität des β-Aktin Gens oder einer früher isolierten kontaminierenden Sequenz, die mit verschiedenen 23S rRNA der Pseudomonasgattung korreliert. Für jede Reaktion wird durch Austesten von PCR-Programmen mit unterschiedlichen Zykluszahlen sichergestellt, dass die Reaktion das Plateau nicht erreicht.

Die Bedingungen für das "Pseudo-PCR" Programm sind:

	1 Zyklus	2 Zyklen	15,20,25,30 Zyklen	1 Zyklus
94°C	4 min.	30 sek.	20 sek.	
58°C		30 sek.	30 sek.	
72°C		3 min.	30 sek.	4 min.

2.2.3.4.3.2 Spot Test.

Um die Anreicherung von Testersequenzen bzw. Abnahme von Driversequenzen in der Subtraktion zu quantifizieren, kann ein Spot Test durchgeführt werden. Dazu werden 2 ng, 200 pg, 20 pg und 2 pg jeder PCR2-Reaktion in 0,5 M NaOH denaturiert und auf eine positive geladene Nylonmembran punktförmig aufgebracht. Nach UV-Licht-Bestrahlung führt

man eine Hybridisierung mit dem Driver und dann mit dem Tester als Sonde durch. Es wird dabei die Hybridisierungsintensität der Tester- bzw. Driver-Sonden auf den Substraktionsprodukten verglichen. Das genaue Hybridisierungsprotokoll für Nylonmembranen ist unter 2.2.4.5 beschrieben.

2.2.4 Klonierung und Screening

Da die Sequenzen der Subtraktionsprodukte unbekannt sind, müssen diese in einen Vektor kloniert werden, um weiter analysiert werden zu können. In dieser Arbeit wurde das pGEM[®]-T Easy Plasmid von der Firma Promega benutzt.

Das pGEM[®]-T Easy Plasmid ist für die Klonierung von PCR-Produkten besonders geeignet. Die Taq-Polymerase produziert keine Fragmente mit glatten Enden, sondern meist einen unspezifischen Überhang von einer Base (Clark, 1988). Diese Base ist in der Mehrzahl der Fälle ein Adenosin. Durch diesen Basenüberhang lassen sich die PCR-Produkte leicht in Vektoren mit Thymidin-Überhang, wie das pGEM[®]-T Easy Plasmid klonieren. Diese Methode wird als TA-Klonierung bezeichnet (Holton & Graham, 1991; Marchuk *et al.*, 1991).

Die Klonierungsstelle liegt in dem lacZ-Gen, so dass es inaktiviert wird, wenn ein DNA-Fragment erfolgreich in den Vektor kloniert wurde. Dies erlaubt eine Blau-Weiß-Selektion der Kolonien: Bakterienkolonien, in denen das LacZ-Gen durch die Insertion eines DNA-Fragments zerstört ist, bleiben nach Inkubation mit IPTG und X-Gal weiß, während Klone ohne Insertion sich blau färben. Durch ein Resistenzgen des Vektors für Ampicillin werden nur Bakterien, die eine Plasmid tragen, selektioniert.

2.2.4.1 TA-Klonierung

10 µl der PCR2-Reaktion werden 10 min. bei 95°C denaturiert, danach wird dazu eine Mischung von 3 µl dATP (2 mM) und 0,2 µl Taq Polymerase (5 U/µl) gegeben, für 15 min. bei 70°C inkubiert, mit 100 µl Wasser gemischt und mit einer Microcon 100 Säule gereinigt, um die Fragmente unter 125 bp abzutrennen.

Das entstandene Produkt wird in Wasser auf 10 ng/µl verdünnt und 2 µl davon werden mit 4 µl pGEM[®]-T Easy Vector (50 ng/µl, Promega), 27 µl Rapid Ligation Puffer (2x), 5 µl T4 DNA Ligase (3 U/µl) und 16 µl Wasser über Nacht bei 15°C inkubiert.

Da man für die folgende Elektroporation entsalzte Produkte braucht, werden die Ligationsprodukte durch eine Fällung gereinigt. Zu den 50 µl Ligationsansatz werden 250 µl TE gegeben; die Lösung wird nacheinander mit einem Volumen (300 µl) Phenol-Chloroform-

Isoamylalkohol (25:24:1, v/v) ausgeschüttelt, zentrifugiert und die wässrige Phase, unter Vermeidung der Interphase in ein neues Gefäß überführt. Dieser Vorgang wird unter Verwendung von Chloroform anstatt von Phenol-Chloroform-Isoamylalkohol wiederholt. Der Überstand wird mit 1 µl Glycogen, 30 µl Natrium Acetat (3M, pH 5,2), 660 µl Ethanol 100 % gemischt, über Nacht bei -20°C gefällt und anschließend zentrifugiert (45 min., 4°C , 14000 rpm). Das entstandene Pellet wird dreimal mit 1 ml EtOH 70% gewaschen (10 min., 4°C , 14000 rpm) und in 50 µl Wasser gelöst. Die üblichen erhaltenen Konzentrationen lagen zwischen 3 und 12 ng/µl.

2.2.4.2 Elektroporation der Subtraktionsprodukte in *E. coli*

Bakterien besitzen die Fähigkeit freie DNA, wie z.B. Plasmide, aus dem umgebenden Medium aufzunehmen. Diese Eigenschaft wird als Kompetenz bezeichnet. Einige Organismen besitzen eine natürliche Kompetenz, aber im Labor werden die Bakterien erst durch besondere Wachstumsbedingungen zur Aufnahme von DNA befähigt. Um eine hohe Effizienz bei dem Verfahren zu erreichen, wurde hier die Elektroporation der Transformation mit Calciumchlorid vorgezogen. Die Elektroporation ist ein physikalischer Prozess, bei dem die Zellmembran durch kurze ($\sim 10\ \mu\text{s}$) elektrische Pulse hoher Feldstärke durchlässig gemacht wird. Die Methode gilt für hochmolekulare Moleküle besonders zuverlässig und allgemein als sehr effizient.

Für einen Ansatz werden 50 µl elektrokompetente *E. coli* XL1-Blue (Stratagene) auf Eis aufgetaut, zu den vorgekühlten 4 µl (0,1 bis 5 ng) DNA-Lösung aus dem Ligationsansatz pipettiert, vermischt und in eine vorgekühlte Elektroporationsküvette (2 mm) überführt. Nach 1 min. Inkubationszeit auf Eis wird der Ansatz in einem Genepulser (Biorad) bei 2,5 kV, 25 µF und 200 Ω mit vorgekühltem Schlitten durchgeführt. Übliche Zeitkonstanten lagen im Bereich von 3 bis 4 Millisekunden. Sofort nach der Elektroporation werden die Bakterien mit 950 µl 37°C warmen SOC Medium erwärmt, aus der Küvette in ein 1,5 ml Reaktionsgefäß überführt und für weitere 30 min. bei 37°C inkubiert.

2.2.4.3 Blau-Weiß-Selektion

Zwei Verdünnungen der Bakterien (1/100 und 1/1000 mit LB Medium) werden auf Ampicillin (0,1 mg/ml), X-gal (80 µg/ml) und IPTG (0,5 mM) enthaltende LB-Platten aufgetragen (der Rest der Bakterien wird bei -80°C in Glycerin-Stocks weggefroren). Nach Inkubation der Platten über Nacht bei 37°C wird die Koloniezahl bestimmt und die Zahl plasmidhaltiger (weißer) Klone errechnet. Weiße Kolonien werden gepickt, in 50 µl LB mit Ampicillin (0,1 mg/ml) Medium überführt und bei 4°C gelagert.

2.2.4.4 Größenbestimmung der Inserts mit PCR

Für jeden Klon muss eine Kontroll-PCR durchgeführt werden, um die Größe des Inserts zu bestimmen. Eine Standard PCR wird mit den PN1 und PN2 Oligonukleotiden mit 1 µl des in LB Medium gelagerten Klons durchgeführt. Die Reaktion wird in einem Gesamtvolumen von 30 µl statt der üblichen 10 µl durchgeführt und mit längerer Anfangsdenaturierung.

Die Bedingungen für das PCR-Programm im Robocycler sind:

	1 Zyklus	16 Zyklen	11 Zyklen	1 Zyklus
94°C	6 min.	40 sek.	40 sek.	
63°C	30 sek.	30 sek.	30 sek.	
72°C	2 min.	1 min.	1,5 min.	4 min.

Es werden 7 µl auf ein 1,5% Agarose Gel aufgetragen.

2.2.4.5 Filter-Herstellung / Hybridisierung

Bei der Filter-Hybridisierung versucht man die auf einer Membran immobilisierte DNA mit Hilfe einer markierten Sonde nachzuweisen. Es wurden hier positiv geladene Nylonmembranen verwendet. Mit Hilfe eines spezifischen enzymgekoppelten Antikörpers kann die markierte Sonde nachgewiesen werden. In unserem Fall handelt es sich um einen nicht-radioaktiven Nachweis, bei dem Digoxigenin (DIG)-markierte Sonden verwendet werden. Die Markierung wird von einem spezifischen Antikörper gegen DIG, an den das Enzym Alkalische Phosphatase (AP) gekoppelt ist, erkannt. Der Nachweis der alkalischen Phosphatase erfolgt über 5'-Brom-4-chlor-3-indolylphosphat (CSPD, Chemiluminescence Substrate). Das chemilumineszente Substrat wird enzymatisch dephosphoriliert und emittiert Licht bei 477 nm, das sich durch Exposition eines Röntgenfilms detektieren lässt.

2.2.4.5.1 Filter Herstellung

Die PCR-Produkte werden mit Hilfe eines Dot-Blot Apparates (Bio-Dot, Biorad) auf einer Nylonmembran (Filter) aufgebracht. Durch Hybridisierungsversuche mit Tester-DNA als auch mit Driver-DNA sollten sich Klone differentiell exprimierter cDNA identifizieren lassen. Zwei Filter werden nach dem folgenden Protokoll hergestellt.

Die Membranen werden auf die passende Größe zugeschnitten, mit Wasser befeuchtet und in den Dot-Blot Apparat eingespannt. Die Proben werden folgenderweise vorbereitet: 10 µl von jedem PCR Produkt werden mit 15 µl EDTA (100 mM) und 5 µl Wasser gemischt und 20 µl frisch hergestellte NaOH (1 M) werden dazu gegeben, um die cDNA zu denaturieren. Nachdem jede Vertiefung des Dot-Blot Apparates durch Absaugen von 200 µl Wasser mit Hilfe eines Vakuums gewaschen wurde, werden je 50 µl Probe aufgetragen und alle Löcher ohne Probe mit 50 µl Wasser gefüllt. Dabei sollten Luftblasen in den Proben vermieden

werden. Die Flüssigkeit wird abgesaugt und alle Löcher noch zweimal mit 200 µl NaOH (400 mM) gewaschen. Der Apparat wird dann auseinandergebaut und die Membranen auf ein Filterpapier gelegt. Anschließend werden die DNA-Proben durch UV-Licht Bestrahlung auf der Membran vernetzt (*Crosslinking*) und diese getrocknet. Die Membranen können bei 4°C bis zur weiteren Verwendung gelagert werden.

2.2.4.5.2 Herstellung der Hybridisierungssonden

Damit die gebundene DNA-Sonde auf den Filtern sichtbar gemacht werden kann, muss sie vorher markiert werden. In der vorliegenden Arbeit wurde eine nichtradioaktive Markierung mit Digoxigenin-11-dUTP (Dig-UTP) gewählt. Dig-UTP ist ein Substrat für die Taq-DNA-Polymerase und wird anstelle von dTTP in der labeling-PCR in einem Verhältnis von 35% Dig-UTP und 65% dTTP eingesetzt. Die Digoxigenin-markierte DNA kann anschließend mit einem spezifischen Antikörper, an den das Enzym Alkalische Phosphatase (AP) gekoppelt ist, erkannt werden.

49 µl Markierungsmix enthalten 5 µl Puffer 1 (10x), 5 µl vom CP2 Oligonukleotid (24 µM), 7,5 µl Formamid 20%, 1,75 µl 1/7-dNTP-Mix (s. unten), 2,5 µl Digoxigenin-11-UTP (1 mM), 27,5 µl Wasser und 0,5 µl DNA Polymerase Mix (5 U/µl, ELT PCR System, Roche). Zu 49 µl Markierungsmix wird jeweils 1 µl unverdünnte Probe pipettiert.

Der 1/7-dNTP-Mix wird aus der Mischung 5 µl dCTP, 5 µl dGTP, 5 µl dATP, 4,2 µl dTTP und 30,8 µl Wasser hergestellt.

Die Bedingungen für das PCR-Programm im PE 2400 (Perkin Elmer) sind:

	1 Zyklus	10 Zyklen	2 Zyklen	1 Zyklus
94°C	2 min.	15 sek.	15 sek.	
68°C	4 min.	4 min.	4 min.+ 10 sek. / Zyklus	7 min.

Die Reaktion wird bei -20°C gelagert.

2.2.4.5.3 Prähybridisierung der Filter

Um einen Hintergrund bei der Hybridisierung zu vermeiden, müssen zuerst alle verbleibenden freien Bindungsstellen blockiert werden. Dazu wird der Filter mit 2 x SSC befeuchtet. Sechs Milliliter der Dig Easy Hyb™ Lösung werden auf 45°C erwärmt; davon werden 5 ml in ein Hybridisierungsgefäß gefüllt, der Filter dazugegeben und bei langsamer Rotation im Hybridisierungssofen bei 45°C inkubiert. Die restlichen 1 ml Dig Easy Hyb™ werden mit verdautem pBS-E.coli-DNA (100 µg/µl) und 100 µl Herringssperma-DNA (1 mg/ml)

gemischt, bei 95°C für 5 min. denaturiert und zum Filter gegeben. Die Prähybridisierung läuft über Nacht bei 45°C unter Rotation.

2.2.4.5.4 Hybridisierung der Filter

Ein Milliliter Prähybridisierungsmix wird abgenommen und 6 µg von der Sonde (Endkonzentration: 1 µg/ml) werden hinzupipettiert. Anschließend denaturiert man 4 min. bei 94°C und gibt das Gemisch sofort wieder zum Filter. Die Hybridisierung läuft bei 45°C über Nacht.

2.2.4.5.5 Entwicklung

Zuerst werden die Filter gewaschen. Die Waschschrte dienen dazu, überschüssige, nicht gebundene Sonden zu entfernen und unspezifische Bindungen zu lösen. Durch Erhöhung der Waschtemperatur werden unspezifische Bindungen zwischen Sonde und Ziel-DNA gelöst.

Der Hybridisierungsofen und die Waschlösungen werden auf 68°C erwärmt. Alle Waschrte werden in dem Hybridisierungsgefäß unter Rotation im Hybridisierungsofen bei 68°C durchgeführt. Im ersten Waschrte wird die Hybridisierungslösung durch 2 x SSC + 0,1 % SDS ersetzt (die Hybridisierungslösung kann bei -20°C aufbewahrt werden und wieder verwendet werden). Die Filter werden kurz gedreht, dann kurz in 1 x SSC + 0,1 % SDS gewaschen. Man lässt die Filter in 0,5 x SSC + 0,1 % SDS für 15 min. drehen, dann noch einmal für 30 min. in 0,1 x SSC + 0,1 % SDS. Dieser letzte Schrt wird wiederholt und schließlich werden die Filter kurz in einer Schale mit Puffer 1 gewaschen.

Nach den Waschrten der Hybridisierung werden für den Sondennachweis mit Anti-Digoxigenin-Antikörper zunächst unspezifische Bindungsstellen blockiert. Dafür werden die Filter 30 min. bei Raumtemperatur auf dem Schüttler in 25 ml Blockierungslösung langsam geschwenkt.

Die Antikörperbindung erfolgt nach der Zugabe von 2,5 µl eines mit alkalischer Phosphatase (AP) konjugiertem Anti-Digoxigenin-Antikörper (150 U/200 µl) zu der Blockierungslösung und einer Inkubationszeit von 30 min. bei Raumtemperatur. Danach werden die Filter zweimal für jeweils 15 min. bei Raumtemperatur in Puffer 1-Tween-Lösung inkubiert, um den ungebundenen Antikörper weg zu waschen und um die Membran zu äquilibrieren.

Die Filter werden abschließend zweimal kurz in Puffer 3 gewaschen und nach Abtropfen der Flüssigkeit in saubere Entwicklungsbeutel überführt. 990 µl Puffer 3 werden mit 10 µl CSPD

gemischt und auf die Filter gegeben. Das Substrat wird sofort mit der Deckfolie des Beutels gleichmäßig und ohne Luftblasen über die gesamte Membran verteilt. Nach 5 min. Inkubationszeit bei RT wird die überschüssige Flüssigkeit vorsichtig entfernt und der Beutel vollständig geschlossen, um ein Austrocknen der Membran zu vermeiden. Eine 10 minütige Inkubation der feuchten Membran bei 37°C verstärkt die Lumineszenz-Reaktion.

Zur Filmentwicklung werden die im Beutel feucht gelagerten Filter in einer Entwicklungsbox fixiert und ein Film aufgelegt. Die Standard-Entwicklungszeiten liegen zwischen 15 min. und 2 Stunden. Mehrere Exponierungen mit mehreren Filmen können erfolgen, um die gewünschte Signalintensität zu erreichen.

2.2.4.6 Plasmid Extraktion / Sequenzierung

Zur Isolierung von Plasmiden wird der QIAprep Miniprep Kit benutzt. Die Birnboim-Doly-Methode (Birnboim & Doly, 1979) nutzt das unterschiedliche Verhalten chromosomaler DNA und Plasmid-DNA im pH-Bereich von 12-12,5 aus. Bei diesen pH-Werten wird bakterielle chromosomale DNA selektiv denaturiert, während kovalent geschlossene, zirkuläre Plasmid-Ringe intakt bleiben. Nach der Neutralisierung renaturiert die chromosomale DNA und bildet einen unlöslichen Klumpen, der durch Zentrifugation entfernt werden kann. Die Plasmid-DNA verbleibt im Überstand.

Im QIAprep Miniprep Kit ist diese Technik leicht modifiziert worden. Die Neutralisierungslösung hat eine sehr hohe Salzkonzentration, die optimale Bindungsbedingungen für Plasmid-DNA an Silica-Gel-Membranen bietet. Die gebundene Plasmid-DNA kann unter niedriger Salzkonzentration (sogar mit Wasser) und in einem kleinen Volumen eluiert werden, ohne dass weitere Präzipitierung, Konzentration oder Entsalzung nötig ist.

Die selektierten Klone werden in 5 ml LB Medium in Kultur genommen und über Nacht bei 37°C inkubiert. Die Extraktion erfolgt am nächsten Tag wie im QIAprep Miniprep Kit beschrieben. Die Sequenzierung der gewonnenen Plasmid-DNA wurde entweder von der Firma Sequiserve (85591, Vaterstetten) oder von Doktor Klaus Witter (Institut für Anthropologie und Humangenetik der LMU, 80336 München) durchgeführt.

2.2.4.7 Screening

2.2.4.7.1 Elektronische Sequenzanalysen

Ein BLAST (<http://www.ncbi.nlm.nih.gov/BLAST/>) Suchlauf wird mit allen gefundenen DNA-Sequenzen durchgeführt. Sequenzen, die mit diesem Programm nicht identifiziert werden konnten, werden mit dem BLAT Programm (<http://genome.ucsc.edu/cgi-bin/hgBlat>) weiteranalysiert. Für alle menschlichen Sequenzen wird ein Primerpaar bestellt (s. 2.1.7).

2.2.4.7.2 PCR - Bestätigung der isolierten Sequenzen

Die PCR wird nach dem Standard Protokoll (s.2.2.2.6) in 10 µl Volumen durchgeführt. Die optimale Bindungstemperatur wird für jedes Primerpaar zuerst mittels einer Standard-PCR mit einem Temperaturgradienten ermittelt.

Ein Thermocycler wird wie folgt programmiert (die Bindungstemperatur „X“ richtet sich nach dem jeweiligen Primerpaar):

	1 Zyklus	20 Zyklen	20 Zyklen	1 Zyklus
94°C	5 min.	30 sek.	30 sek.	
X°C		30 sek.	30 sek.	
72°C		45 sek.	1 min.	4 min.

2.2.4.7.3 Expression in Driver und Tester

Um die Expression der Sequenzen im Driver und Tester analysieren zu können, werden Verdünnungen (50 ng/µl) von jeder Patientenprobe und von normalen Knochenmarksproben, die im Tester bzw. Driver benutzt wurden, hergestellt. Die Expression jeder Sequenz wird durch die PCR mit dem entsprechenden Primerpaar geprüft.

2.2.4.7.4 Expression in Zelllinien und Knochenmark

Die Sequenzen werden auch für eine differentielle Expression in Tumorzelllinien und Knochenmark von gesunden Patienten geprüft. Als Template für die PCR dient cDNA.

Mononukleäre Zellen aus dem Knochenmark werden wie unter 2.2.1.1 beschrieben isoliert. Als Tumorzelllinien werden verwendet: SkBr3, BT20, MCF7 (Mamma), LNCAP, PC3 (Prostata), HT29, SW480, Lovo (Kolon), H69, A549 (Lunge) und A431 (Epidermoid). Als Wachstumsmedium für die Zelllinien wird ein auf RPMI basierendes Zellkulturmedium verwendet. Eine Ausnahme bildet die Zelllinie A549, die ein Medium mit DMEM benötigt (s. 2.1.4).

2.2.4.7.4.1 RNA Isolierung

Um die Gesamt-RNA zu isolieren, wird eine modifizierte Guanidium Thiocyanat Methode verwendet. Eine Million der zu untersuchenden Zellen (Mononukleäre Zellen aus dem Knochenmark oder Zellen von Tumorzelllinien) werden pelletiert und sofort im 500 µl GT-

Mix mit einem Vortex homogenisiert. Der Reihe nach werden 50 µl Natrium-Acetat (2 M, pH 4), 500 µl Aqua-Roti-Phenol und 100 µl Chloroform zugegeben und homogenisiert. Nach einer Inkubation auf Eis für 15 min. wird die Reaktionsmischung 20 min. bei 4°C und 10 000 rpm zentrifugiert. Die wässrige Phase wird in ein frisches Tube mit 500 µl Isopropanol überführt und mindestens eine Stunde oder über Nacht bei -20°C präzipitiert. Das Tube wird 20 min. bei 4°C und 10 000 rpm zentrifugiert, der Überstand abpipettiert und das Pellet mit 300 µl GT-Mix und 300 µl Isopropanol erneut bei -20°C über Nacht gefällt. Nach 30 minütiger Zentrifugation bei 4°C und 10 000 rpm wird das RNA Pellet mit 70 % Ethanol gewaschen und anschließend in 10 µl RNase-freies Wasser aufgenommen. Diese RNA-Stocklösung wird bei -80°C gelagert.

2.2.4.7.4.2 cDNA-Synthese und Spezifische PCR

Bei der cDNA-Synthese binden zunächst Hexanukleotide ("random primer", Hexanukleotid-Mix, Roche) mit zufälligen Basensequenzen an die RNA. Ein komplementärer DNA-Strang (cDNA) wird anschließend von den 3'-Enden der „random primer“ mit Hilfe des Reverse-Transkriptase-Enzyms synthetisiert. Bei dem als Primer dienenden Hexanukleotid-Mix sind praktisch alle Sequenzkombinationen enthalten, so dass sie an die eingesetzte RNA in zufälliger Verteilung binden. So wird weitgehend eine Reverse-Transkription der gesamten RNA erreicht.

Auf Eis werden pro Probe in einem Reaktionsgefäß 2 µl First Strand RT Puffer (5x), 1 µl DTT (0,1 M), 0,5 µl Igepal 10 %, 2 µl Hexanukleotid-Mix, 3,5 µl Wasser und 1 µl der RNA zu einem Endvolumen von 10 µl gemischt. Die Primerbindung auf die RNA erfolgt nach einer Annealing-Zeit von 5 min. bei Raumtemperatur. Während dieses Schrittes werden 2 µl First Strand RT Puffer (5x), 1 µl DTT (0,1 M), 1 µl dNTP (je 10 mM), 5 µl Wasser und 1 µl Reverse-Transkriptase (200 U/µl) zusammengemischt und der Mix wird dann zu der Annealingreaktion gegeben. Die cDNA-Synthese ist nach einer Inkubation von 45-60 min. bei 44°C beendet. Anschließend werden die erhaltenen Produkte 1/20 mit Wasser verdünnt und ihre Qualität mittels einer Standard-PCR auf β-Aktin geprüft.

Die PCR-Bedingungen sind unter 2.2.4.7.2 beschrieben. Als Template dient 1 µl von der 1/20 verdünnten cDNA Lösungen.

2.2.5 Klonierung der vollständigen Sequenz #288

2.2.5.1 Suche nach der Sequenz #288 in der Gesamt-RNA

2.2.5.1.1 Suche mit dem Primer 288-Fish-Bio

Es wird versucht die gesuchte cDNA mit einem spezifischen biotinylierten Primer zu synthetisieren (**Abbildung 5**, (1)).und sie über Streptavidin-Beads selektiv anzureichern (2). Das Anfügen eines dG-Schwanzes (G-Tailing) an das 3'-Ende der Fragmente (3) sollte die darauffolgende Amplifikation mit dem spezifischen Primer 288-Fish-Nest und dem CP2-Primer (4) ermöglichen.

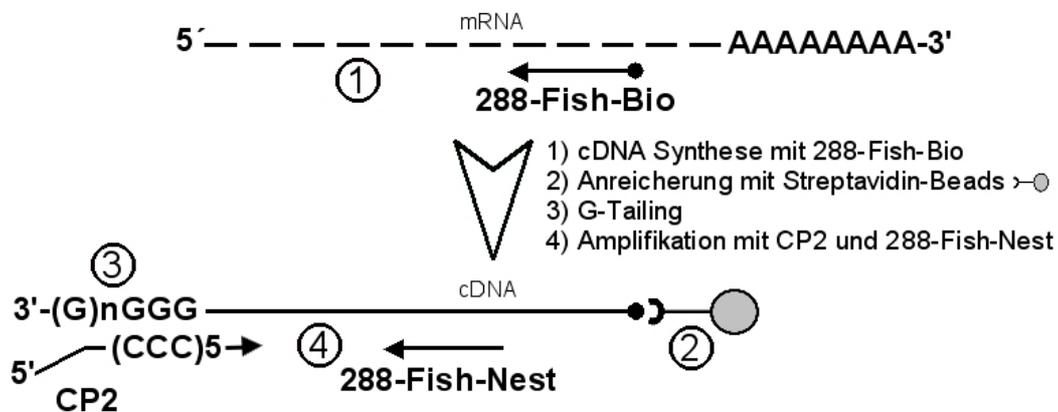


Abbildung 5. Methode zur Isolierung der Sequenz #288 mit dem Primer 288-Fish-Bio.

Um eine cDNA zu isolieren, wird sie aus der Gesamt-RNA mit einem spezifischen, biotinylierten Primer synthetisiert (1) und die cDNA über Streptavidin-Beads selektiv angereichert (2). Nach dem Anfügen eines dG-Schwanzes (G-Tailing) an das 3'-Ende der Fragmente, eine Reaktion, die durch das Enzym Terminale Transferase katalysiert wird (3), wird das Amplifikat mit dem spezifischen Primer 288-Fish-Nest und dem CP2-Primer erneut amplifiziert (4).

2.2.5.1.1.1 cDNA-Synthese

Pro Isolierung werden zwei Reaktionsansätze vorbereitet. Der Reaktionsansatz I enthält 2 µl First Strand RT Puffer (5x), 1 µl DTT (0,1 M), 1 µl 288-Fish-Bio Primer (0,1 M) und 5 µl Wasser. Nach der Zugabe von 1 µl Gesamt-RNA (s.2.2.4.7.4.1) wird die Reaktionsmischung 3 min. bei 70°C denaturiert und eine Hybridisierung des Primers durch eine Abkühlung auf 45°C erlaubt, wobei die Temperatur alle 20 sek. um 1°C reduziert wird. Während der Abkühlung werden 10 µl Reaktionsansatz II angesetzt: 2 µl First Strand RT Puffer (5x), 1 µl DTT (0,1 M), 1 µl dNTPs (10 mM), 5 µl Wasser und 1 µl Reverse Transkriptase. Die cDNA-Synthese erfolgt nach der Zugabe von Reaktionsansatz II 1 Stunde bei 44°C unter Rotation in einem Hybridisierungshofen.

2.2.5.1.1.2 Anreicherung

Bevor man die selektive Anreicherung durchführt, müssen die benötigten Beads vorbereitet werden. Pro Ansatz werden 15 µl Streptavidin-Beads, zweimal mit 15 µl Bindungspuffer (2x) am Magnet gewaschen und anschließend in 15 µl Bindungspuffer (2x) resuspendiert. Zu den 15 µl Beads werden 5 µl Wasser und 10 µl des cDNA-Syntheseansatzes gegeben. Die Bindung der cDNA an die Beads erfolgt bei 4°C während einer 30 minütigen Rotation.

Die Beads werden dann viermal mit jeweils 30 µl Puffer gewaschen. Zuerst mit Lids-Wasch-Puffer (1x), dann mit cDNA-Igepal-Wasch-Puffer (1x), mit cDNA-Tween-Wasch-Puffer (1x) und zuletzt mit Tail-Wasch-Puffer (1x).

2.2.5.1.1.3 G-tailing

Pro Ansatz werden 30 µl Tailing-Mix angesetzt: 3 µl MgCl₂ (40 mM), 3 µl DTT (0,1 M), 3 µl dGTP (0,2 mM), 1,5 µl KH₂PO₄ (200 mM) und 19,5 µl Wasser. Darin werden die Beads nach dem vorigen Waschschritt resuspendiert. Die Reaktion wird 4 min. bei 94°C inkubiert und anschließend direkt auf Eis gestellt. Pro Reaktionsgefäß wird 1 µl TdT (14U/µl) zugegeben. Das G-Tailing erfolgt während einer Inkubation von einer Stunde bei 37°C und mindestens einer Stunde bei 22°C. Die Inaktivierung der TdT erfolgt durch 5 min. bei 70°C.

2.2.5.1.1.4 PCR

Um Fehlhybridisierung oder Primerdimerbildung bei der PCR mit dem Primer CP2 zu vermeiden, wird die PCR mit einem „Hot start“ Schritt gestartet. In einem ersten Reaktionsansatz werden 5 µl PCR Puffer (10x), 7,5 µl Formamid 20 %, 1 µl CP2 Primer (24 µM), 1 µl 288-Fish-Nest Primer (24 µM) und 20,5 µl Wasser pipettiert. Zu diesen 35 µl werden 10 µl der G-Tailing Reaktion zugegeben und in einem Thermocycler auf 78°C erhitzt. In 3 µl Wasser werden 1 µl dNTPs und 0,5 µl Taq Polymerase (5 U/µl) gemischt und zu dem auf 78°C erhitzten Ansatz zugegeben.

Der Thermocycler wird folgenderweise programmiert:

	1 Zyklus	40 Zyklen	1 Zyklus
78°C (Hot Start)	Mix II geben		
94°C	2 min.	15 sek.	
68°C		30 sek.	
72°C		2 min.	5 min.

Jeder PCR-Ansatz wird auf einem präparativen Agarosegel aufgetrennt und die amplifizierten DNA-Banden ausgeschnitten, wobei auch mehrere Banden pro Ansatz auftreten können. Die DNA-Banden werden mittels Qiaquick Gel Extraktion-Kit (Qiagen) aufgearbeitet.

2.2.5.1.2 Suche mit dem Primer 288-15C

Eine andere Isolierungsstrategie wird getestet. Der Primer 288-15C trägt einen Poly-C-Schwanz an 5'-Ende, dessen fünfzehn Cytosin-Basen keine magnetische Selektion (wie für die vorige Methode beschrieben), sondern eine Amplifikation der Fragmente mit einem einzigen CP2 Oligonukleotid (**Abbildung 6**, (1)) ermöglichen.

Zu diesem Zweck wird das Protokoll für die globale cDNA-Amplifikation (2.2.2), allerdings mit folgenden Modifizierungen verwendet:

- Zur Isolierung der mRNA werden 0,6 µl der Gesamt-RNA (s.2.2.4.7.4.1) an 10 µl magnetische Beads gebunden und dann gewaschen.
- Die cDNA-Synthese erfolgt mit dem 288-15C Primer anstelle der CFL5cT- und CFL5c8-Oligonukleotide. Auf diese Weise kann spezifisch die gesuchte Sequenz transkribiert werden. Nach dem G-Tailing (2) erfolgt eine Amplifikation mit dem CP2-Primer (3).

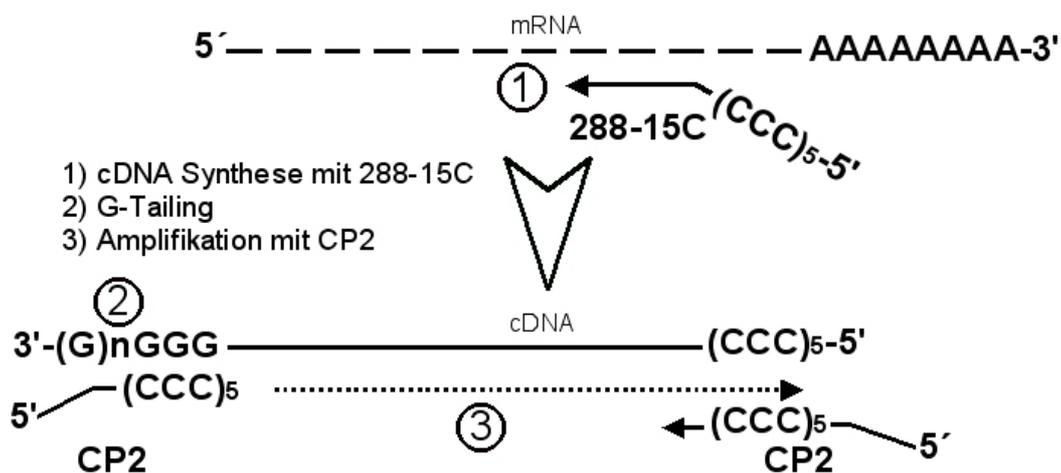


Abbildung 6. Methode zur Isolierung der Sequenz #288 mit dem Primer 288-15C.

Im ersten Schritt (1) wird eine cDNA Synthese mit dem spezifischen Primer 288-15C durchgeführt. Dann wird am 3' Ende der cDNA ein poly-G-Schwanz von einer Terminalen Transferase angefügt (2). Bei der Amplifikation der cDNA (3) bindet der Primer CP2 im ersten Zyklus an das poly-G-Ende und im zweiten Zyklus an die neu entstandene G₍₁₅₎-Sequenz, die zur Poly-C-Sequenz des Primers 288-15C komplementär ist. Die Produkte dieser sekundären PCR werden auf ein Gel aufgetragen.

2.2.5.1.3 Suche mit dem Primer CFL5cTLib

Da die Benutzung des CP2 Primers bei der globalen cDNA-Amplifikation zur Bildung unspezifischer Produkte führen kann, wird die cDNA-Synthese mit dem Primer CFL5cTLib (anstelle der CFL5cT- und CFL5c8-Oligonukleotide, s. 2.2.2.3) durchgeführt. Die gezielte cDNA-Amplifikation der Sequenz #288 erfolgt mit der Kombination von Primern Lib1 und 288-5' (**Abbildung 7**).

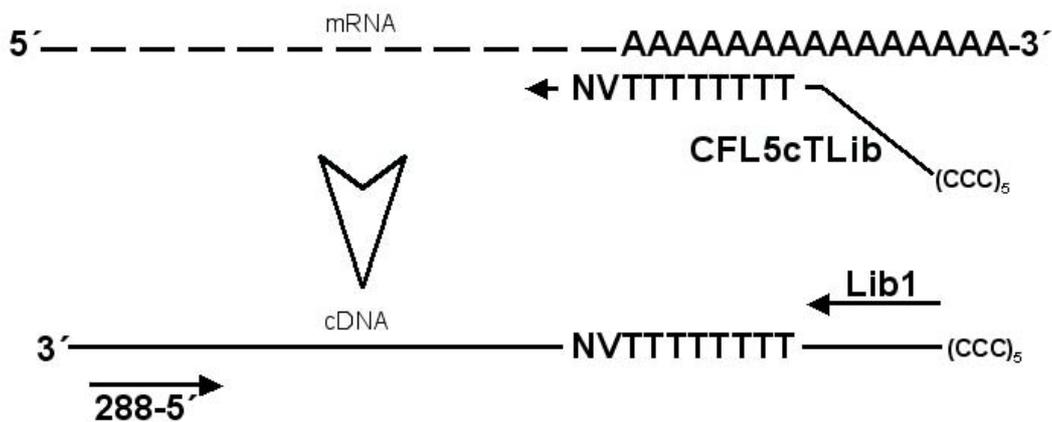


Abbildung 7. Prinzip der Isolierung der Sequenz #288 mit dem CFL5cTLib Primer.

Dank seines 3'-poly-T Endes kann der CFL5cTlib Primer an den Poly-A Schwanz der mRNA binden und die cDNA-Synthese initiieren. Mit den Primern 288-5' und Lib1 kann in der folgenden PCR die Sequenz #288 spezifisch amplifiziert werden. Es wurde mit dieser Methode versucht, die Sequenz aus den Zelllinien BT20, Lovo und PC3 zu isolieren; Die cDNA-Synthesen der drei Zelllinien wurden vereinigt (cDNA pool). Als Positivkontrolle wurde mit entsprechenden Primern die Beta-Aktin Sequenz isoliert.

(Die Isolierung der Sequenz #288 mit dem TLib Primer (s. unten) läuft nach dem gleichen Prinzip, der Tlib Primer trägt aber keine Cytosin Basen an seinem 5'-Ende.)

Diese letzte Amplifikation wird nach dem Standard-PCR-Protokoll (s. 2.2.2.6) in 20 µl durchgeführt; als Probe dienen 2 µl der cDNA-Synthese Produkte.

Die Bedingungen für das PCR Programm sind:

	1 Zyklus	40 Zyklen	1 Zyklus
94°C	2 min.	30 sek.	
59°C		30 sek.	
72°C		1,5 min.	4 min.

2.2.5.1.4 Suche mit dem Primer TLib

Die fünfzehn Cytosin Basen können am 5'-Ende des CFL5cTLib möglicherweise starke Sekundärstrukturen bilden, was die Effektivität der cDNA-Synthese einschränken kann. Deswegen wird der Primer TLib, der diese fünfzehn Cytosin Basen nicht trägt, anstatt des CFL5cTLib Primers benutzt. Das verwendete Protokoll ist in 2.2.5.1.3 beschrieben.

2.2.5.2 Spezifische magnetische Anreicherung nach der globale cDNA-Amplifikation

Im Vergleich zu der im Abschnitt 2.2.5.1.1 beschriebenen Methode, findet die Bindung des biotinylierten Primers an Streptavidin-Beads **vor** der Zugabe von cDNA (**Abbildung 8, [1]**) statt (in diesem Fall wird cDNA benutzt, die direkt aus frischen Zellen global amplifiziert wurde

(s. 2.2.2)). Dieses soll die Bindung des Primers fördern und eine effizientere magnetische Anreicherung ermöglichen

Für vier Ansätze werden 20 µl Streptavidin-Beads zweimal in Bindungspuffer (BP, 2x) gewaschen, dann in 20 µl BP (1,6x = 16 µl BP (2x) + 4 µl Wasser) aufgenommen, und mit 12 µl vom 288-Fish-Bio Primer (100 µM) gemischt. Die Bindung des Primers an die Beads erfolgt unter Rollen, 30 min. bei 4°C. Die ungebundenen Primer werden abgewaschen (Igepal-BP (1x), BP (1x)), dann wird der Ansatz in 40 µl Wasser und 8 µl PCR-Puffer (10x) aufgenommen.

Als Proben dienen drei globale cDNA-Amplifikationen von 3 Zelllinien (BT20, LNCAP und PC3). Von jeder Probe werden 2 µl mit 5 µl Wasser gemischt, 5 min. bei 95°C denaturiert und anschließend auf Eis gestellt. Dazu werden 11 µl von den Beads in PCR Puffer gegeben und das Ganze (18 µl) wird 15 min. bei 60°C in einem Hybridisierungssofen unter Rotation annealt.

Zwei Amplifikationsmethoden entweder mit den CP2 und 288-Fish-Nest Primern (**Abbildung 8, 2a**) oder mit dem Primer CP2 allein (**2b**) werden für nachfolgende Bearbeitung getestet.

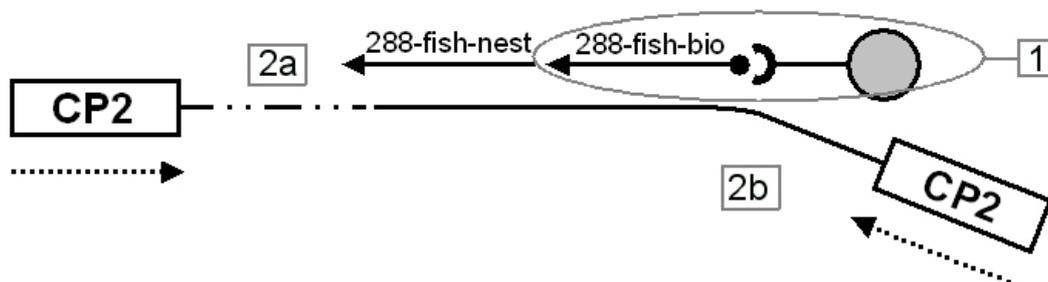


Abbildung 8. Methode zur Isolierung der Sequenz #288 mit dem 288-fish-bio Primer.

In Schritt 1 wird der Primer 288-Fish-Bio an die Streptavidin-Beads gebunden. Nach der Hybridisierung dieses Primerkonstruktes an die cDNA wird es gründlich gewaschen um die ungebundenen und unspezifischen cDNA Fragmente zu entfernen. Dann wird versucht, die Sequenz entweder mit den 288-Fish-Nest und CP2 Primern (2a) oder mit dem Primer CP2 allein (2b) zu amplifizieren. Diese beiden PCR-Strategien wurden auf Proben der drei Zelllinien BT20, LNCAP und PC3 angewendet.

2.2.5.2.1 PCR mit CP2 und 288-Fish-Nest

Für vier Ansätze werden 2 µl dNTPs (10 mM), 4 µl BSA (10 mg/ml), 1 µl Taq Polymerase (5 U/µl) und 4 µl Wasser zusammengemischt. 2 µl davon werden zu dem im oberen Schritt vorbereiteten 18 µl Ansatz zugegeben. Eine 5 minütige Inkubation bei 72°C erlaubt die Synthese des spezifischen Strangs ab dem 3'-Ende des 288-Fish-Bio Primers.

Die neu synthetisierten Stränge, die magnetische Streptavidin-Beads an ihrem 5'-Ende tragen, werden am Magnet isoliert. Dafür werden die Stränge in 20 µl NaOH (0,1 M) resuspendiert

und 15 min. bei 37°C inkubiert, um Einzelstränge zu generieren. Nach einem Waschschrift (Igepal-BP (1x), BP (1x)), werden die Beads in 30 µl Wasser aufgenommen.

Für die PCR mit den Primern CP2 und 288-Fish-Nest werden 20 µl von den Beads benutzt, die restlichen 10 µl werden für eine Kontroll-PCR verwendet. Der Master PCR Ansatz enthält 28 µl PCR-Puffer (10x), 5,6 µl dNTPs (10 mM), 42 µl Formamid 20 %, 1,4 µl Taq Polymerase und 81,2 µl Wasser.

- 112 µl werden mit 4 µl CP2 (24 µM) und 4 µl 288-Fish-Nest (24 µM) Primern gemischt, dann auf vier Aliquots à 30 µl geteilt und zu den 20 µl Beads gegeben.
- 22,5 µl werden mit 2 µl CP2 (24 µM) gemischt, auf vier Ansätze mit je 5,6 µl geteilt und zu jedem Ansatz werden 4 µl Beads dazu gegeben.
- 22,5 µl werden mit 2 µl 288-Fish-Nest (24 µM) gemischt, auf je vier Ansätze mit 5,6 µl geteilt und zu jedem Ansatz werden 4 µl Beads dazu gegeben.

Die Bedingungen für das PCR-Programm sind:

	1 Zyklus	40 Zyklen	1 Zyklus
94°C	2 min.	15 sek.	
68°C		30 sek.	
72°C		2 min.	5 min.

2.2.5.2.2 PCR mit CP2

Bevor man die CP2-Amplifikation durchführt, werden die Produkte noch einmal am Magnet gewaschen, um möglichst alle cDNA-Stränge, die an den biotinylierten 288-Fish-Bio Primer unspezifisch gebunden haben, zu entfernen. Bleiben unspezifische cDNA-Stränge hängen, werden sie mit dem Primer CP2 amplifiziert und verhindern die Spezifität bei der Isolierung der Zielsequenzen. Jeder Ansatz wird einmal mit Igepal-BP (1x) und einmal mit BP (1x) gewaschen, und in 20 µl Wasser aufgenommen.

Für vier Ansätze werden 20 µl Puffer 1 (10x), 8 µl dNTPs (10 mM), 30 µl Formamid 20 %, 20 µl CP2-Primer (24 µM), 2 µl Expand Long Template Enzym Mix (3,5 U/µl) und 40 µl Wasser zusammengemischt. Für jede cDNA-Probe (20 µl) werden 30 µl vom Reaktionsansatz verwendet.

2.2.5.3 Suche nach der Sequenz #288 in einer kommerziellen cDNA Bank

Die Human Universal cDNA Library (HUCL) der Firma Stratagene wird auf das Vorhandensein der Sequenz #288 geprüft. Die Library ist entweder als Filter zu hybridisieren oder als Mikrotiterplatte mit einer PCR zu untersuchen. Die Länge der Sequenz #288 beträgt 400 bp, was für eine gute spezifische Hybridisierung zu kurz ist. Deshalb wird die PCR als sensitivere Methode verwendet.

Die Library ist in 768 Mikrotiterplatten zu je 384 Einzelklonen archiviert, jedoch so organisiert, dass man die gesamte Library mit nur zwei PCR-Analysen untersuchen kann. Die cDNA von jeder Platte werden gemischt und dann in drei virtuelle Dimensionen (Spalte, Zeile und Fläche) eingeteilt. Mit 88 Reaktionen erlaubt die erste PCR die Identifizierung der Platte, die den zu suchenden Klon enthält. Die identifizierte Platte wird gekauft und mit der zweiten PCR werden alle der 384 Klonen getrennt analysiert. Wenn ein positiver Klon identifiziert wird, kann er einzeln für weitere Analysen erhalten werden.

2.2.5.3.1 "Screening" PCR (Primäre PCR)

Die PCRs werden mit den Primern 288-5' und 288#4 und dem Expand Long Template System durchgeführt. Pro 10 µl Ansatz werden 1 µl Puffer 1, 0,2 µl dNTPs (10 mM), 0,5 µl Primer 288-5' (8 µM) und 0,5 µl Primer 288#4 (8 µM), 6,7 µl Wasser, 0,1 µl Expand Long Template Enzym Mix (3,5 U/µl) und 1 µl Probe.

Die von Stratagene empfohlenen Bedingungen für das PCR Programm im RoboCycler werden folgenderweise angepasst:

	1 Zyklus	60 Zyklen	1 Zyklus
95°C	5 min.	30 sek.	
53°C		30 sek.	
72°C		30 sek.	10 min.

Für jede weitere durchgeführte primäre PCR mit anderen Primerpaaren werden die gleichen Bedingungen benutzt, mit Ausnahme der Hybridisierungstemperatur: für das Primerpaar 288-5' und 288#10 liegt sie bei 53°C, für das Primerpaar 288-5' und 288#11 bei 56°C.

Die Produkte werden auf einem 3 %-igen Agarosegel (170 V, 30 min.) getrennt. Positive Banden werden isoliert, mit dem Qiaquick Gel Extraktion Kit gereinigt und sequenziert, um die Spezifität der PCR zu bestätigen. Um falsch positive Signale auszuschließen, werden zwei zusätzliche Kontrollen durchgeführt: eine Nested-PCR und ein Verdau.

2.2.5.3.2 Nested-PCR

Die Bedingungen für die Nested-PCR sind die gleichen wie unter 2.2.5.3.1 beschrieben, mit Ausnahme der verwendeten Primer und der Zyklenzahl. Für das Primerpaar 288-5' und 288#12 wird die Hybridisierungstemperatur bei 53°C eingestellt, für das Primerpaar 288#12 und 288#13 bei 58°C. Die Zyklenzahl wird für beide Fälle auf 25 Zyklen beschränkt. Als Probe dient 1 µl der unverdünnten primären PCR.

2.2.5.3.3 Ear I Verdau

Die Fragmente, die mit den Primern 288#12 und 288#13 entstehen, werden aus dem Agarosegel ausgeschnitten, mit dem Qiaquick Gel Extraktion Kit in 50 µl Wasser eluiert und anschließend mit der Restriktionsendonuklease Ear I behandelt. Die Erkennungsstelle dieses Enzyms ist:



Damit werden die 119 bp langen Fragmente in zwei Stücke (86 und 33 bp) gespalten.

In einem Reaktionsgefäß werden 10 µl von der eluierten DNA mit 1,5 µl NEB Puffer 1 (10x) und 0,5 µl Ear I (10 u/µl) gemischt, Wasser ad 10 µl. Die Reaktion wird mindestens eine Stunde bei 37°C inkubiert und dann auf einem 3 %-igen Agarosegel getrennt.

2.2.5.4 Expression in menschlichen Gewebe

Die Expression der Sequenz #288 in verschiedenen Geweben wird mit Hilfe einer Hybridisierung auf einen kommerziellen Nylon Array überprüft. Der hierfür verwendete Multiple Tissue Expression (MTE) Array (BD Biosciences) enthält mRNA von 57 Gewebearten, acht Tumorzelllinien, sieben fetale Gewebearten und acht unterschiedliche Negativkontrollen. Die Konzentration der fixierten Poly(A)⁺mRNA auf der Nylonmembran ist mit Hilfe von acht Haushaltsgenen normalisiert worden und schwankte zwischen 40 und 300 ng. Diese Konzentration ist ausreichend, um damit Transkripte nachweisen zu können, die zu 0,1-0,01% (i.e. 1000-100 Transkripte) in der gesamten Poly(A)⁺mRNA vertreten sind (BD Biosciences, Multiple Tissue Expression (MTE) Array Manual).

2.2.5.4.1 Vorbereitung der Radioaktiven Probe

Da für die Radioaktivmarkierung einer Probe eine sehr hohe Konzentration an DNA benötigt wird, wird ein 500 µl PCR Ansatz vorbereitet: 50 µl PCR-Puffer (10x), 25 µl BSA, 360 µl Wasser, 25 µl 288-5' Primer und 25 µl 288#4 Primer, 5 µl Taq Polymerase und 10 µl Plasmid (s. 2.2.4.6) werden gemischt. Dieser Ansatz wird in 15 PCR Ansätze à 30 µl verteilt. Die anschließende PCR wird in RoboCycler mit folgendem Programm durchgeführt:

	1 Zyklus	25 Zyklen	14 Zyklen	1 Zyklus
94°C	6 min.	40 sek.	30 sek.	4 min.
58°C	30 sek.	30 sek.	30 sek.	30 sek.
72°C	2 min.	45 sek.	1 min.	4 min.

Alle 15 PCR Ansätze werden zusammen gemischt, auf einem präparativen Agarosegel aufgetrennt und die amplifizierte DNA-Bande ausgeschnitten. Die DNA-Bande wird mit dem

Qiaquick Gel Extraktion-Kit (Qiagen) extrahiert und 3 µl des 50 µl Eluats werden als Kontrolle auf einem Agarosegel aufgetragen.

Die Radioaktivmarkierung wird mit [³²P]dATP und 3 µl des PCR-Eluats nach dem Protokoll von Feinberg und Vogelstein durchgeführt ((Feinberg & Vogelstein, 1983). Die radioaktiv markierte Sonde wird anschließend auf einer Säule (NucTrap® Probe Purification Columns, Stratagene) von nicht eingebauten Nukleotiden getrennt (und somit aufgereinigt).

2.2.5.4.2 Prähybridisierung und Hybridisierung des Filters

Der MTE Filter wird während der Vorbereitung der Radioaktivprobe bei 65°C mit 10 ml ExpressHyb™ Hybridization Solution (BD Bioscience) angefeuchtet. Dieser Prähybridisierungslösung werden anschließend 1,5 mg Herringssperma-DNA und 30 µg Human Cot-1 DNA (beide bei 95°C für 5 min. vorher denaturiert) zugegeben.

Die radioaktiv markierte Probe wird sofort nach ihrer Reinigung über die Säule zu der Prähybridisierungslösung gegeben. Die Hybridisierung findet über Nacht bei 65°C statt.

2.2.5.4.3 Entwicklung des Filters

Der Filter wird wie in Paragraph 2.2.4.5.5 beschrieben mit SSC/SDS Lösungen gewaschen. Zur Filmentwicklung wird der Filter in einem Beutel feucht gelagert und in einer Entwicklungsbox auf ein "Storage Phosphor Screen" aufgelegt. Die Exponierungszeit entspricht sechs Tagen. Die Ergebnisse der Hybridisierung werden mit einem Phosphoimager (Storm 860, Amersham Biosciences) detektiert.

3 Ergebnisse

Die folgende Übersicht soll die komplexen Zusammenhänge dieser Arbeit zum besseren Verständnis schematisch darstellen.

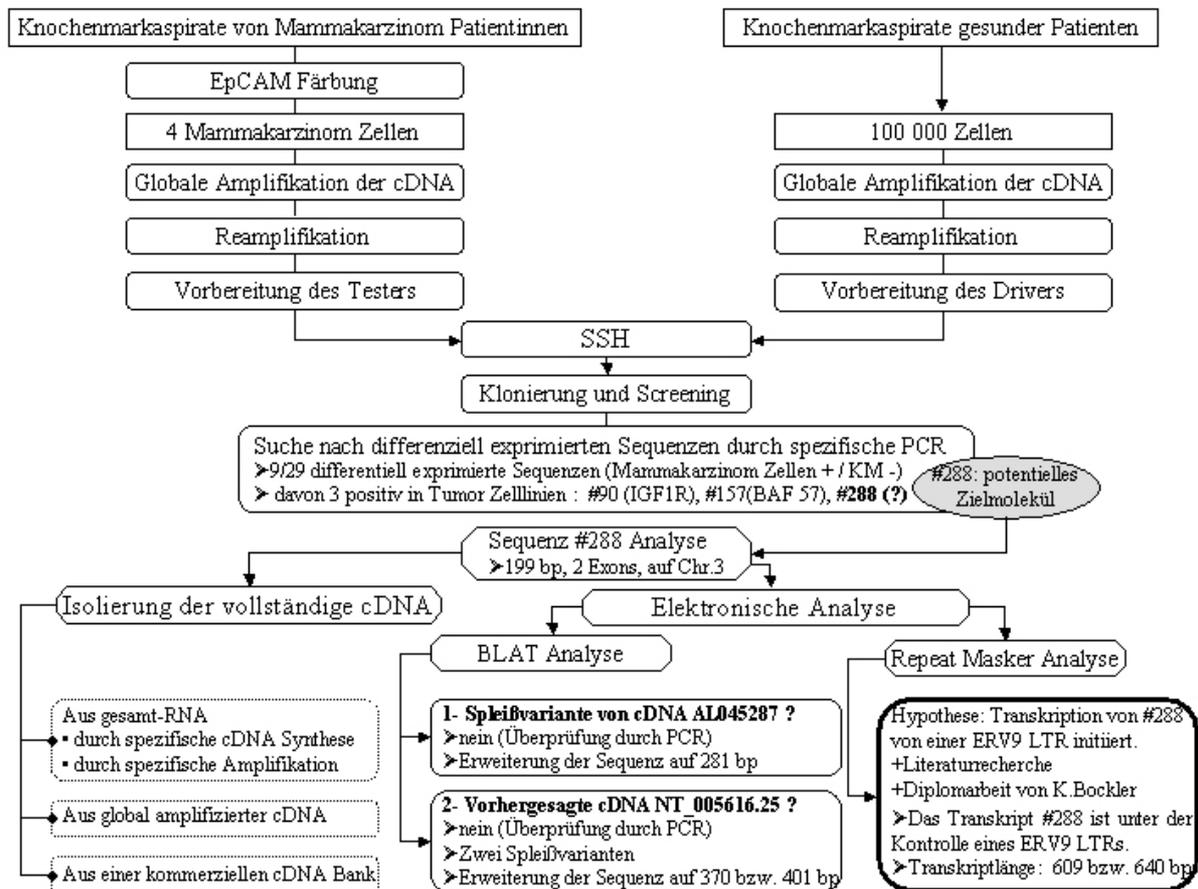


Abbildung 9. Übersicht der Vorgehensweise und Ergebnisse der Arbeit.

Das verwendete Material ist in Rechtecken wiedergegeben, die verwendete Methoden in abgerundeten Rechtecken (gepunktete Linien zeigen Methoden, die keine Ergebnisse gebracht haben). Die erhaltenen Ergebnisse sind mit Pfeilen (>) angeführt.

3.1 Subtraktion mit Einzelzellen

3.1.1 Notwendige Modifikationen und Optimierung der SSH

3.1.1.1 Anpassung der Adaptoren an Einzelzellamplifikate

Im Gegensatz zu dem ursprünglichen Protokoll der SSH (Diatchenko *et al.*, 1996; Diatchenko *et al.*, 1999), bei dem mRNA als Ausgangsmaterial verwendet wird, werden in dieser Arbeit PCR-Amplifikate einzelner disseminierter Tumorzellen in der Subtraktion eingesetzt. Um die nach globaler Amplifikation erhaltenen PCR-Produkte dafür verwenden zu können, mussten Modifikationen an der SSH-Methode (**Abbildung 10**) vorgenommen werden.

Die cDNA-Fragmente dürfen nicht größer als 2 kb sein. Aufgrund des Random-Primings mit Cfl5c8 bei der globalen Amplifikation liegt die Größe der Produkte zwischen 200-1000 bp, so dass eine Restriktion mit Rsa I (wie im originalen Protokoll empfohlen) nicht unbedingt erforderlich ist (**Abbildung 12**, "Ausgangsmaterial"). Durch die PCR-Amplifikation mit dem CP2-Primer wurden an den Fragmentenden EcoR I-Schnittstellen eingeführt. Mit den leicht modifizierten Adaptoren A1 und A2 konnte nun eine „sticky end“ Ligation durchgeführt werden, was die Ligationseffizienz erhöhte (2.2.3.2). Die Sequenz der P- und PN-Primer wurde entsprechend angepasst (**Abbildung 12**, "Tester").

3.1.1.2 Neue Primer für die PCR1 und für die Driver-Amplifikation

3.1.1.2.1 Amplifikation des Drivers mit dem CP2-Primer (Subtraktion #1)

Die erste Subtraktion wurde mit den genannten Modifizierungen durchgeführt. Als Tester diente ein Amplifikat einer einzelnen K562 Zelle, als Driver ein Amplifikat eines cDNA-Pools aus 100 000 normalen KM Zellen von drei Patienten.

Zur Kontrolle der Subtraktion wurden nach der PCR1 und PCR2 (s. 2.2.3.4, Seite 37) β -Aktin und Mage 4 in einer spezifischen PCR nachgewiesen und zusammen mit jeweils 10 μ l der Ausgangsamplifikate auf ein Agarosegel aufgetragen (Im gesunden KM ist Mage-4, im Gegensatz zu K562, normalerweise nicht nachzuweisen). Auf dem Gel war, im Gegensatz zur PCR1, nach der PCR2 kein Schmier mehr sichtbar. Darüber hinaus war weder eine Reduktion des β -Aktin Fragmentes, noch eine Anreicherung von Mage-4 zu erkennen. Eine Optimierung der Anzahl der PCR-Zyklen, wie im Originalprotokoll vorgeschlagen, verbesserte das Ergebnis nicht.

Die wahrscheinliche Ursache der geringen PCR-Ausbeute liegt in den (Cp)₁₅ und (Gp)₁₅-Sequenzen im 5'- bzw. 3'-Bereich der Fragmente, die während cDNA-Synthese eingeführt werden. Dadurch können die Fragmente eine relativ stabile Haarnadelstruktur ausbilden, die eine effiziente PCR-Amplifikation verhindert. Auch die gewünschten differentiellen Fragmente, die endständig zwei unterschiedliche Primer tragen, sind möglicherweise von diesem Problem betroffen. Hinzu kommt, dass die in der PCR1 verwendeten Primer P1 und P2 mit 22 Basen Länge in Relation mit dieser CG-Haarnadelstruktur eine zu geringe Bindungsenergie besitzen, so dass letztere thermodynamisch bevorzugt ist. Aufgrund dessen wurden 30 und 33 Basen lange Primer in der PCR1 getestet (P1-30 und P1-33), deren entsprechend höhere Annealingtemperatur die Bildung der Haarnadelstruktur vermeidet (s. Material und Methoden).

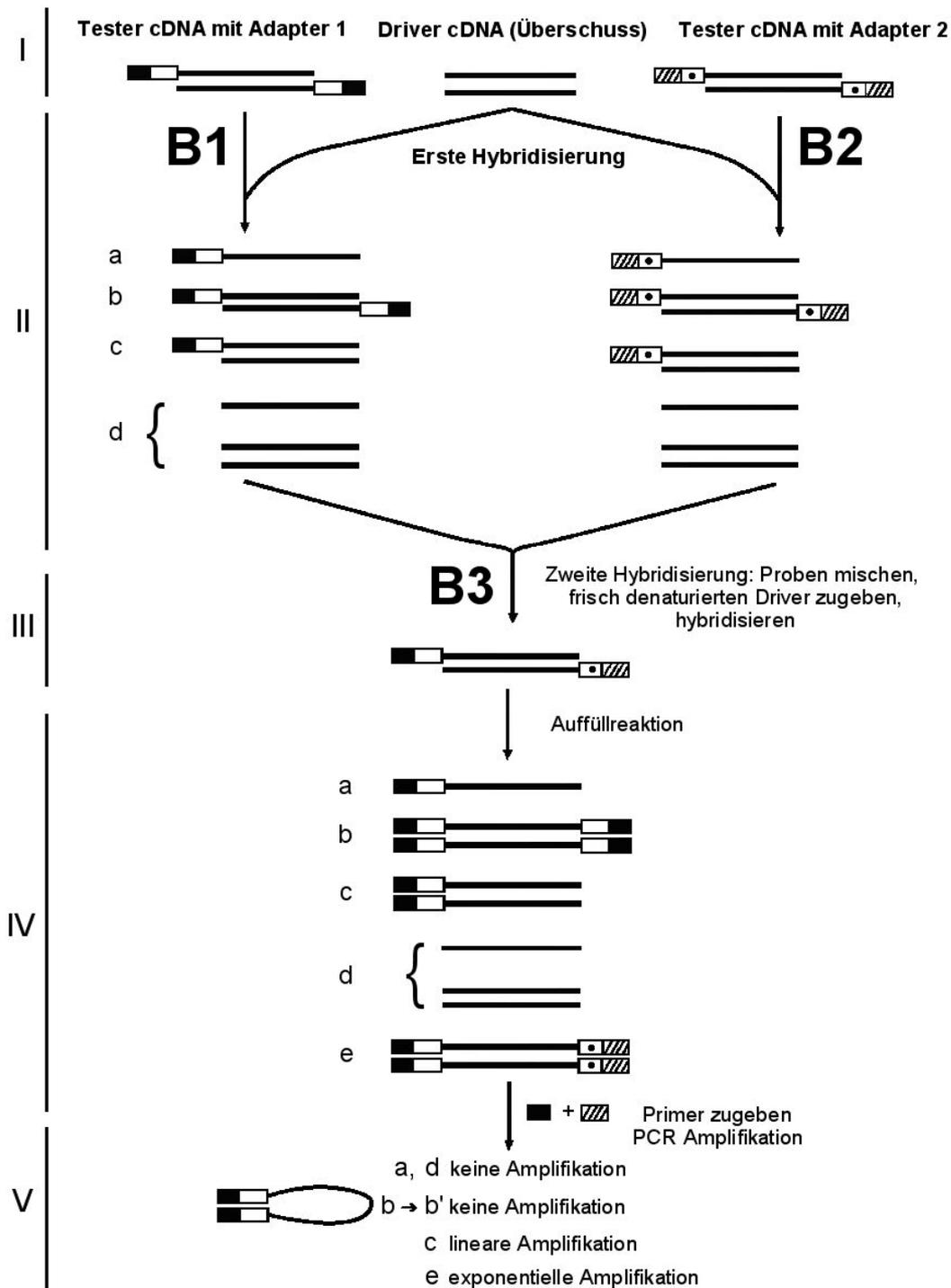


Abbildung 10. Algorithmus der SSH Methode.

(I) Adapter 1 und 2 werden in zwei getrennten Ansätzen an die Tester cDNA ligiert. (II) Jeder Ansatz (Reaktionsgefäß B1 bzw. B2) wird mit einem Überschuss an Driver cDNA gemischt. Bei der ersten Hybridisierung werden die differenziell exprimierten Sequenzen der einzelsträngigen Tester Moleküle normalisiert und angereichert. (III) Die beiden Ansätze werden anschließend zusammen mit einem Überschuss an denaturierten Driver cDNA gemischt (Reaktionsgefäß B3) und eine zweite Hybridisierung wird durchgeführt. Dabei entstehen verschiedene Hybridtypen (a-d). (IV) Die Enden dieser Hybridisierungsprodukte werden aufgefüllt, so dass die differenziell exprimierten Gene durch die nächsten PCR-Schritte mit geeigneten Primern bevorzugt amplifiziert werden. (V) Zwei weitere PCRs (PCR1 und PCR2) werden durchgeführt, um differenziell exprimierte Gene bevorzugt zu amplifizieren.

Ausgehend von der Subtraktion #1 wurden dazu mit den Primern P22, P30 und P33 drei neue suppressive Amplifikate (PCR1) synthetisiert. Es zeigte sich jedoch nach der Nested-PCR (PCR2), dass auch hier mit keinem der neuen Primersets ein PCR-Produkt nachzuweisen war.

3.1.1.2 Amplifikation des Drivers mit dem Primer CPBgl (Subtraktion #2)

Ein modifizierter Driver wurde hergestellt, indem das gepoolte Ausgangs-PCR-Produkt, das die C₁₅-Sequenz trägt, mit einem im Vergleich zu CP2 an drei Basen veränderten CPBgl-Primer reamplifiziert wurde. Dadurch konnte eine Bgl I-Schnittstelle (GCCNNNN⁰NGGC) sechs Basen vor dem 3'-Ende des Primers eingefügt werden. Der Verdau mit Bgl I erlaubte es nach der Amplifikation, 21 Basen des Primers abzuschneiden, so dass der Großteil der C₁₅-Sequenz entfernt werden konnte (**Abbildung 12**, "Driver"). Eine neue Subtraktion wurde mit diesem Driver durchgeführt. Die suppressive PCR1 ergab mit dem Bgl I verdauten Driver und den Primern P30 bzw. P33 auf einem Gel sichtbare PCR-Produkte. Weiterhin negativ waren im Gegensatz dazu die PCR-Ergebnisse unter Verwendung des P22-Primers sowie der nicht subtrahierten Kontrolle (Ligation von Tester mit beiden Adaptoren A1 und A2, 2.2.3.3.3, Seite 37) (**Abbildung 11**).

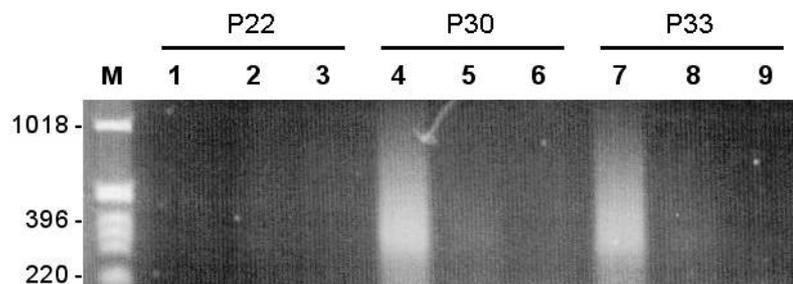


Abbildung 11. Einfluss der Primerlänge auf die PCR1

Die PCR1 in Anschluß an die Subtraktion #2 wurde mit drei Primerpaaren von unterschiedlicher Länge durchgeführt. Das ursprünglich empfohlene Primerpaar, mit 22 Basen Länge, ergab keinen Schmier. Der erwartete Schmier wurde mit den längeren Primern von 30 bzw. 33 Basen erhalten. Die nicht subtrahierte Kontrolle sowie die Negativkontrolle zeigten keine Amplifikation. Primerlänge: 22 bp (Spuren 1-3), 30 bp (Spuren 4-6) und 33 bp (Spuren 7-9). Spuren 1, 4, 7: Subtraktion #2. Spuren 2, 5, 8: nicht subtrahierte Kontrolle. Spuren 3, 6, 9: Negativkontrolle. M: Marker. 1,5 % Agarosegel, Sybr Gold.

Daher kann man annehmen, dass sich erstens die C₁₅-Sequenz aufgrund der verstärkten Haarnadelstruktur negativ auf die Subtraktion auswirkt und zweitens eine Primerlänge von 22 Basen für eine effiziente Amplifikation des Testers nicht ausreicht.

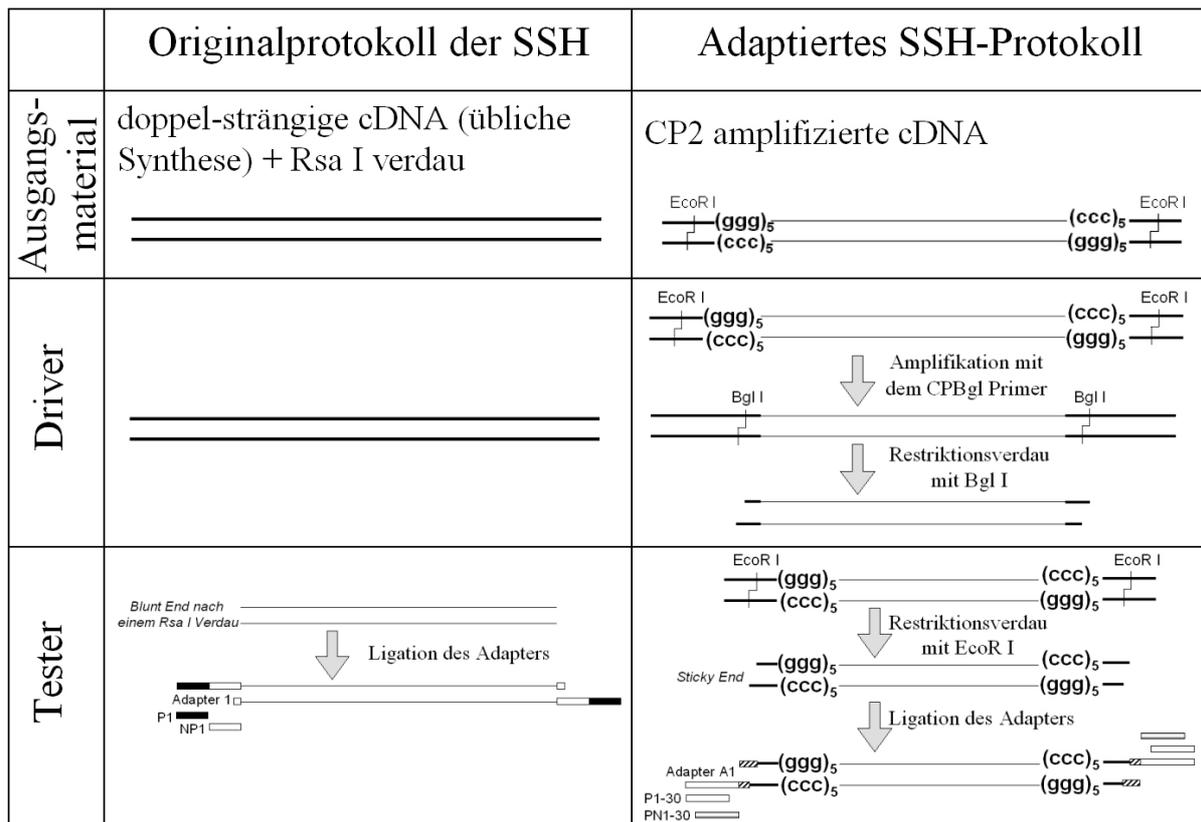


Abbildung 12. Modifikationen der SSH.

Nach der Amplifikation mit dem Primer CP2 trägt das Ausgangsmaterial $(Cp)_{15}$ und $(Gp)_{15}$ -Sequenzen im 5'-bzw.-3' Bereich. Diese Sequenzen müssen im **Driver** entfernt werden, da sie die Bildung störender Haarnadelstrukturen fördern. Die Entfernung erfolgt durch eine Amplifikation mit dem Primer CPBgl und eine Restriktionsverdau mit Bgl I. Bei dem **Tester** stören diese $(Cp)_{15}$ und $(Gp)_{15}$ -Sequenzen ebenfalls, so dass alternative Primer P1-30 und PN1-30 für die PCR1 und PCR2 neu gestaltet wurden. Der Primer CP2 trägt eine Restriktionsschnittstelle für EcoR I, die für die Herstellung des **Testers** benutzt wird: Eine *sticky end* Ligation des Adapters ist effizienter als eine *blunt end* Ligation. Dafür muss die Sequenz der Adapter A1 et A2 angepasst werden.

3.1.1.2.3 PCR-Kontrolle

Um die Effizienz der Subtraktion #2 zu evaluieren, wurde in den Produkten der Nested-PCR mit Hilfe einer PCR die Transkripte der Gene β -Aktin und Mage-4 nachgewiesen. Die Quantifizierung im Vergleich zur nicht subtrahierten Kontrolle erfolgte durch eine PCR-Serie mit ansteigender Zyklenzahl (β -Aktin: 18,23,28,33, Mage-4: 25,30,35,40).

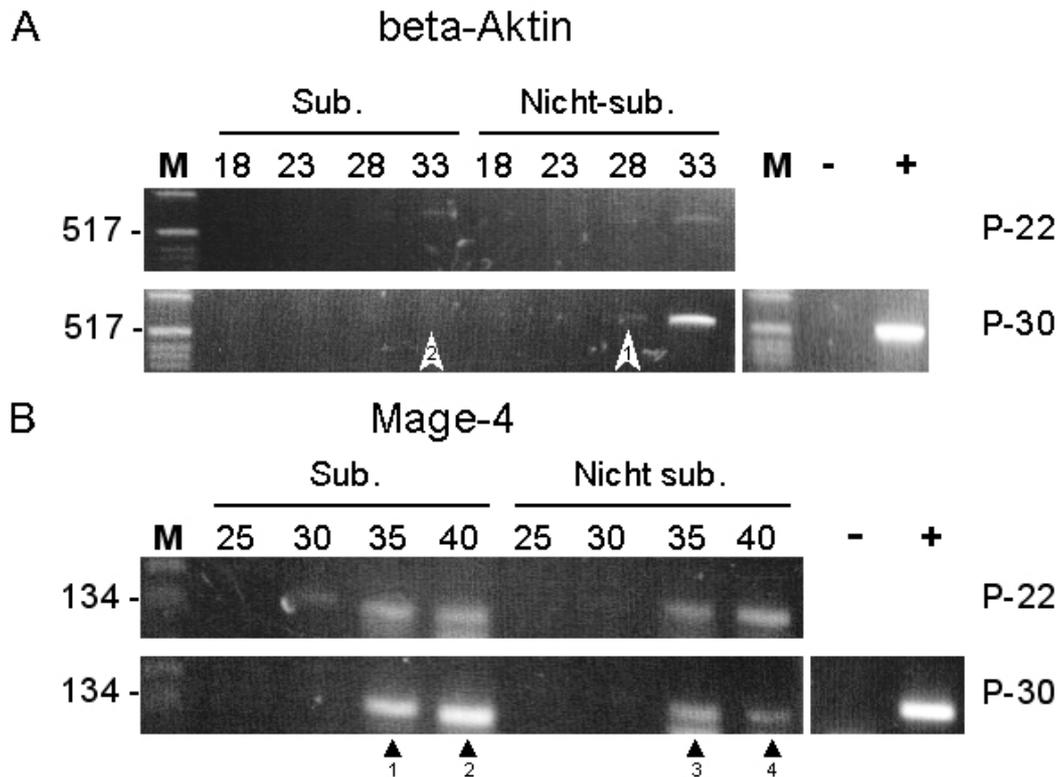


Abbildung 13. Verminderung an β -Aktin (A) und Anreicherung an Mage-4 (B) in Subtraktion #2.

Als Tester-cDNA diente das Amplifikat einer einzelnen K562-Zelle, als Driver-cDNA die eines Amplifikats von 100 000 normalen KM-Zellen, das mit dem CPBgl Primer amplifiziert und anschließend mit dem Bgl I Restriktionsenzym verdaut worden war. Für die PCR1 wurden drei verschiedene Primerlängen getestet, 22, 30 und 33 Basen (Zeilen P-22 und P-30; mit dem P-33 Primer wurde keine Verbesserung festgestellt, die Ergebnisse werden nicht gezeigt). Als Template der β -Aktin und Mage-4 PCR wurde die PCR2 der Subtraktion (Sub.) und die nicht subtrahierte Kontrolle (Nicht sub.) eingesetzt. Die Zyklenzahl bei der β -Aktin PCR betrug: 18, 23, 28 und 33 Zyklen. Die Zyklenzahl der Mage-4 PCR betrug: 25, 30, 35 und 40 Zyklen. M: Marker. Positiv Kontrolle: HT29 cDNA. 2,5 % Agarosegel, Sybr Gold. Angaben zu den Pfeilen siehe Text.

Wie die **Abbildung 13-A** zeigt, wurden die höchsten relativen PCR-Produktkonzentrationen in der PCR1 mit den P1-30 und P2-30 Primern erzielt. β -Aktin ist nach 28 PCR-Zyklen in der nicht subtrahierten Kontrolle als Bande zu detektieren (weißer Pfeil 1) aber noch nicht nach 33 Zyklen bei erfolgter Subtraktion (weißer Pfeil 2). Dass nach der Subtraktion 5 bis 15 zusätzliche PCR-Zyklen benötigt werden um ein Fragment nachzuweisen, wird von Diatchenko et. al. beschrieben (Diatchenko *et al.*, 1999). Im Gegensatz dazu ist Mage-4 in der Subtraktion nach 35 und 40 Zyklen besser zu detektieren als bei der Kontrolle (**Abbildung 13-B**, schwarze Pfeile 1-2 und 3-4 vergleichen). Diese PCR-Ergebnisse weisen darauf hin, dass die Subtraktion #2 erfolgreich war. Mit einem anderen Tester wurde die Subtraktion wiederholt, um die Reproduzierbarkeit der Methode zu bestätigen.

3.1.1.2.4 Spot-Test Kontrolle (Subtraktion #3)

Als Tester der erneuten Subtraktion diente ein Amplifikat einer einzelnen Bronchialkarzinomzelle (Loh3), der Driver entsprach demjenigen der Subtraktion #2. Außerdem wurde eine Gegen-Subtraktion ("Driver-Tester", was hier als "KM-Loh3 geschrieben wird) durchgeführt. Diese Gegen-Subtraktion wird eingesetzt, um die differentiellen Transkripte anzureichern, die im Driver aber nicht im Tester vorhanden sind.

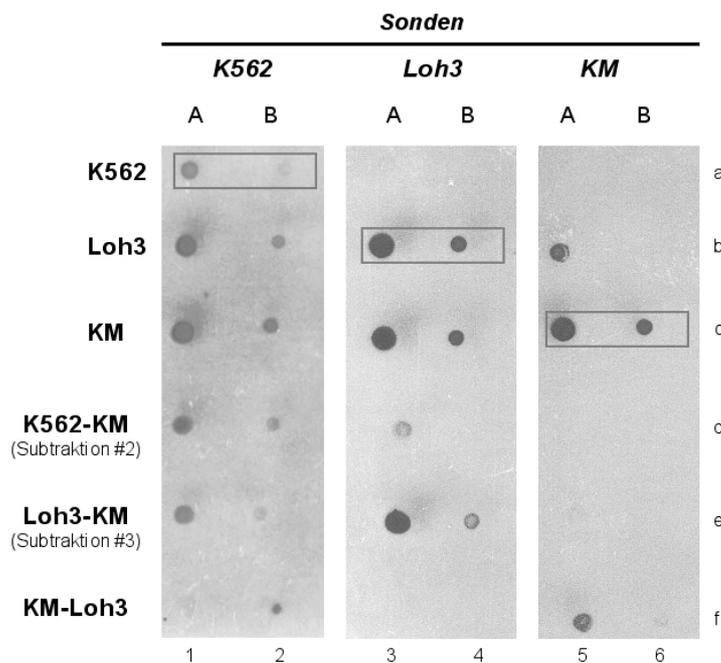


Abbildung 14. Spot Test für die Effizienz von Subtraktionen #2 und #3.

Die Effizienz der Subtraktion #2 und #3 wurde durch einen Dot-Blot überprüft. Zu vergleichen ist jeweils die Hybridisierung der Sonde mit sich selbst (grauen Kästchen) bzw. mit den entsprechenden Subtraktions-Produkten. Für den Test wurden 63 ng (A) und 6,3 ng (B) des Testers, des Drivers und der Subtraktion auf einer Nylonmembran immobilisiert und mit Digoxigenin markiertem Tester und Driver hybridisiert (Sonden). Die Entwicklung erfolgte mittels eines anti-Digoxigenin Alkalische-Phosphatase gekoppelten Antikörpers. Exponierungszeit: 1 min.

Die PCR1 wurde mit den P1-30 und P2-30 Primern durchgeführt, welche die besten Ergebnisse in der obigen Etablierung erbracht hatten. Wie erwartet, war in der Subtraktion #3 der erhaltene PCR-Schmier sowohl nach der PCR1 als auch nach der PCR2 stärker als derjenige der nicht subtrahierten Kontrolle.

Um die Subtraktionen zu quantifizieren wurde ein Dot-Blot durchgeführt. Dazu wurden Verdünnungen von Driver, Tester, Subtraktion und Kontrolle hergestellt und diese auf Nylon-Filtern immobilisiert. Die Filter wurden mit Tester- und Driver-Sonden hybridisiert (**Abbildung 14**). Bei einer erfolgreichen Subtraktion sollte man eine Verminderung der Driversequenzen und eine Anreicherung der Testersequenzen erhalten. Bei der Loh3-KM

Subtraktion (#3) ist eine Abnahme der Driversequenzen (d.h. der nicht differentiellen Sequenzen) deutlich zu sehen (b5 mit e5 vergleichen). Eine Abnahme ist auch bei der Gegen-Subtraktion "KM-Loh3" zu sehen (c-3,4 mit f-3,4 vergleichen). Der Spot Test kann aber keine Anreicherung nachweisen.

3.1.2 Subtraktion mit Mamma-Ca Zellen.

Die Subtraktion wurde nun mit der Ziel-cDNA durchgeführt.

3.1.2.1 Kontamination durch bakterielle Sequenzen (Subtraktion #4)

Als Tester für die Subtraktion diente eine cDNA Mischung von vier einzelnen Mammakarzinomzellen aus dem KM von vier Patientinnen. Die Subtraktion #4 wurde parallel mit 500 ng (#4.1), wie im publizierten Protokoll vorgeschlagen, bzw. 2,5 µg Driver (#4.2) durchgeführt und das Ergebnis durch eine PCR auf β -Aktin und mit Hilfe eines Spot-Tests überprüft. Durch eine spezifische PCR wurde die Reduktion der relativen Häufigkeit des β -Aktin Fragments nachgewiesen. Die Verminderung des β -Aktin Fragments ist zum Beispiel in der Subtraktion #4.1 nach 35 und 40 Zyklen zu sehen, wie in **Abbildung 15** dargestellt (Spuren 3-4 mit Spuren 7-8 vergleichen). Der Spot-Test der Subtraktion #4.1 zeigt eine Anreicherung der Testersequenzen (**Abbildung 16**, b6 mit d6 vergleichen); Bei der Gegen-Subtraktion "KM-Tumorzellen" ist eine Verminderung der Driversequenzen (Tumorzellen) wie erwartet deutlich zu erkennen (a4-6 mit c4-6 vergleichen).

Beide Subtraktionen #4.1 und #4.2 ergaben ähnliche Ergebnisse: Der Überschuss an Driver scheint die Subtraktion weder verbessert noch verschlechtert zu haben, so dass im Folgenden mit dem "normalen" Subtraktionsprotokoll #4.1 weiter gearbeitet wurde.

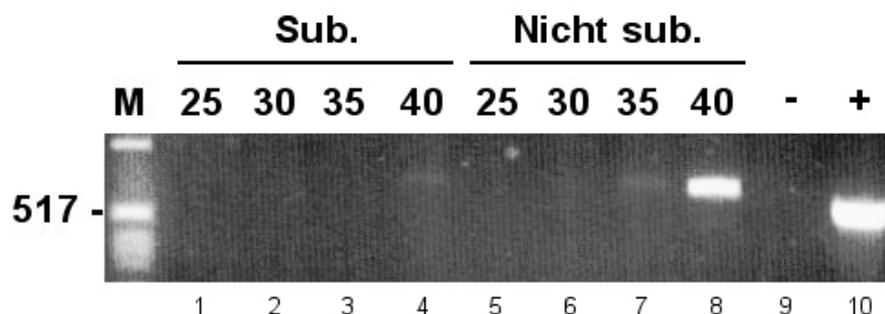


Abbildung 15. Verminderung der β -Aktin Häufigkeit in Subtraktion #4.1.

Als Tester-cDNA der Subtraktion diente das Amplifikat einer Mischung vier einzelner Mammakarzinom-Zellen, als Driver-cDNA ein Amplifikat von 100 000 gesunden KM Zellen, das mit dem CPBgl Primer amplifiziert und anschließend mit dem Bgl I Restriktionsenzym verdaut wurde. Die PCR1 wurde mit den P-30 Primern durchgeführt. Die β -Aktin Kontroll-PCR wurde auf der PCR2 der Subtraktion (Sub.) und der nicht subtrahierten Kontrolle (Nicht sub.) durchgeführt mit aufsteigenden Zyklenzahlen (25, 30, 35 und 40). M: Marker. Positivkontrolle: HT29 cDNA; 2 % Agarosegel, Sybr Gold.

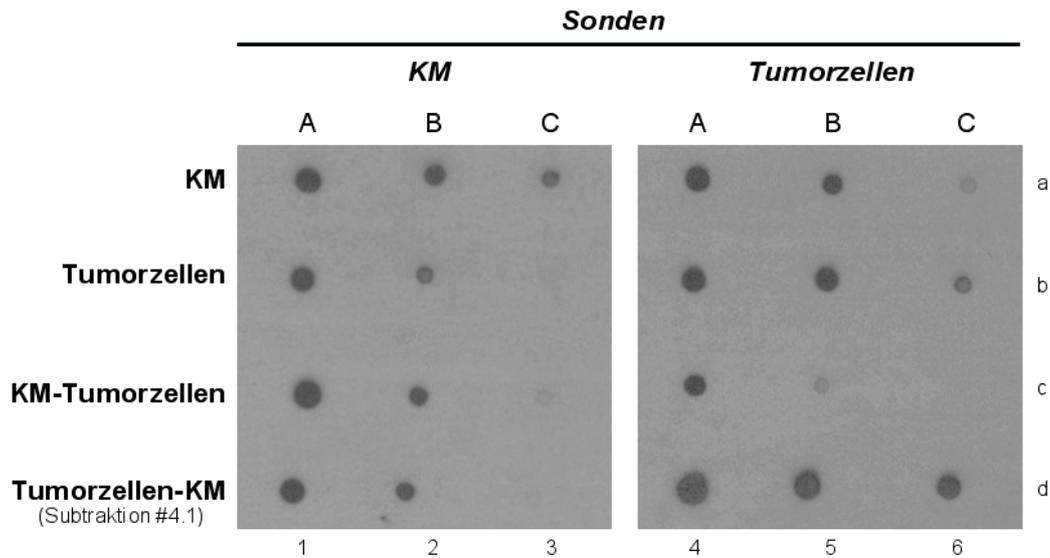


Abbildung 16. Spot Test für die Effizienz der Subtraktion #4.1.

Für diesen Test wurden 100 ng (A), 10 ng (B) und 1 ng (C) des Knochenmarks (KM; cDNA), der Tumorzellen (cDNA) und der beiden Subtraktionen auf zwei Nylonmembranen aufgebracht. Die Knochenmark-Tumorzellen "Gegen-Subtraktion" wurde als Kontrolle durchgeführt. Als Sonden wurden Dig-markierten Knochenmark und Tumorzellen benutzt. Die Entwicklung erfolgte mittels eines anti-Digoxigenin Alkalische-Phosphatase gekoppelten Antikörpers. Exponierungszeit: 1 min.

Die Subtraktions-Produkte wurden für die weiteren Analysen in ein pGEM-T Easy Plasmid ligiert, in kompetente *E. coli* XL1-Blue Zellen elektroporiert und auf Ampicillin enthaltende LB Platten ausplattiert.

Bei 16 Klonen wurde die Größe des Inserts mittels einer PCR unter Verwendung der Primer M13 analysiert. Weil die Hälfte (8/16) der Plasmide kein Insert aufwiesen, wurde für die zukünftigen Untersuchungen auf die Blau-weiß-Selektion der Klone mit Hilfe von IPTG-X-Gal-Medium zurückgegriffen. Von den positiven Klonen trugen viele sehr kleine Fragmente (<30 bp, 6/8). Nur zwei Klone enthielten relativ große Fragmente (>250 bp) (**Abbildung 17-A**). Zwei Southern-Blots wurden hergestellt und mit Digoxigenin-markierten Driver und Tester cDNA-Sonden hybridisiert (**Abbildung 17-B** und **-C**). Kein Klon hybridisierte mit dem Driver, im Gegensatz dazu hybridisierten jedoch fast alle klonierten Fragmente mit dem Tester, obgleich nur Klon #2 und #16 eine zufriedenstellende Signalintensität erreichten. Die anderen Signale wurden als unspezifisch betrachtet.

Die Klone #2 und #16 wurden sequenziert, allerdings konnte nur für Klon #2 eine Sequenz erhalten werden: Es handelte sich um eine 23S ribosomale RNA-Sequenz aus *Pseudomonas*, die auch eine hohe Homologie mit 23S ribosomalen RNA-Sequenzen von anderen Bakterien zeigte.

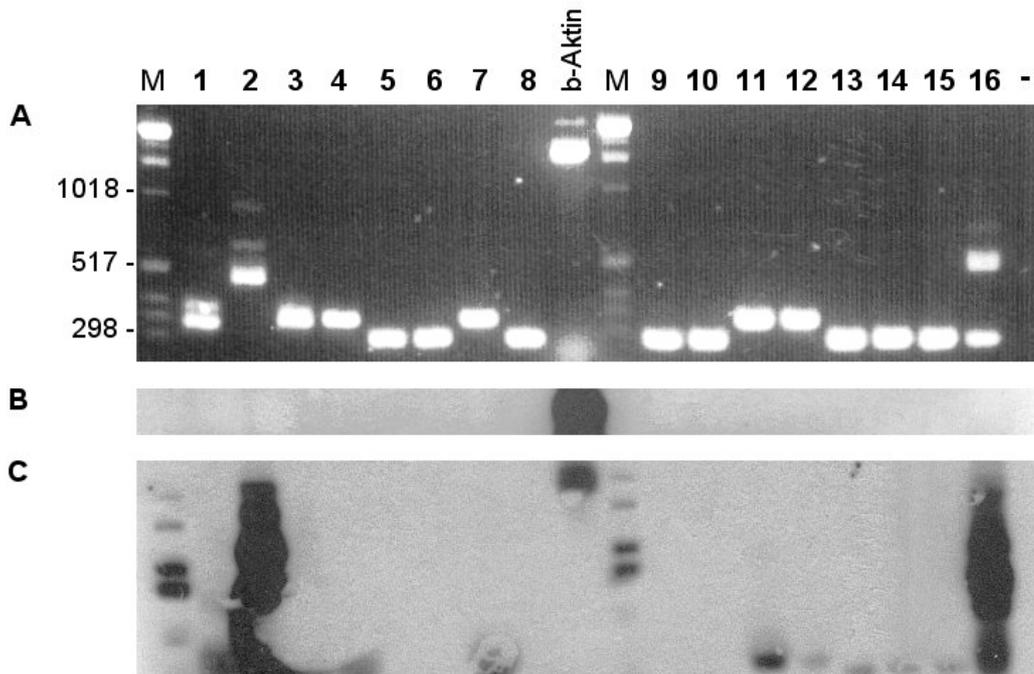


Abbildung 17. Screening von 16 Klonen.

(A)-Screening PCR: 16 Klone wurden mittels einer PCR mit den Primern M13 überprüft; die Produkte wurden auf einem 2 %igen Gel aufgetrennt und anschließend auf eine Nylonmembran transferiert. Die Klone 5,6,8,9,10,13,14 und 15 tragen kein Insert. β -Aktin: 300 ng von einem β -Aktin Insert wurden als Kontrolle für die weiteren Hybridisierungen aufgetragen. M: Marker. (B)-Southern Blot mit dem Driver als Sonde. Das einzige Signal zeigte das β -Aktin Insert. (C) Southern Blot Analyse mit dem Tester als Sonde. Die Klone 2 und 16 reagierten stark positiv.

Um unspezifische Hybridisierungssignale zu vermeiden, wurde die Screening-PCR mit den Primern PN-30 statt mit den Primern M13 durchgeführt und zum anderen wurden die Filter zusätzlich zur Herring-Sperma-DNA auch mit *E. coli*- und Plasmid-DNA blockiert. Die PCR wies ebenso wie die Hybridisierung weniger unspezifische Banden auf.

Um das Screening zu beschleunigen, wurde versucht die PCR-Produkte direkt auf eine Nylonmembran zu spotten und mit mRNA des Testers und Drivers zu hybridisieren, so dass auf eine Überprüfung der Fragmentgröße und einen Southern-Blot verzichtet werden konnte.

96 Klone wurden mit den Primern PN-30 amplifiziert, die PCR-Produkte gespottet und zum einen mit dem Tester und zum anderen mit einer Mischung aus Negativkontrolle (Negativ-Mix) von drei verschiedenen globalen cDNA-Amplifikationen (siehe 2.2.2) hybridisiert. Als Kontrolle wurden zusätzlich acht PCR-Produkte auf ein Gel aufgetragen (**Abbildung 18**). Wie aus der Abbildung hervorgeht, enthielten nicht alle PCR-Ansätze ein Produkt (Spuren 1c und 1h), so dass negative Hybridisierungssignale des Dot-Blots sehr vorsichtig beurteilt werden müssen, wenn auf die Gelelektrophorese verzichtet wird.

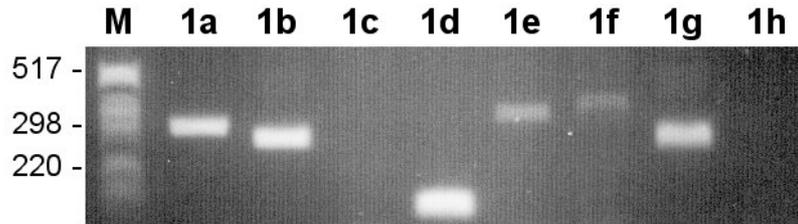


Abbildung 18. Screening PCR von 8 Klonen.

96 Klone wurden in 50 μ l LB-Medium aufgenommen und jeweils 1 μ l in eine Screening PCR mit den PN-30 Primern eingesetzt. 5 μ l von acht PCR-Ansätzen wurden auf einem 1,5 %igen Agarosegel aufgetrennt, um den Erfolg der PCR vor der Filterherstellung zu bestätigen. Die Bezeichnung der Proben ergibt sich aus der Spaltennummerierung und Zeilenziffer von **Abbildung 19**. Es ist anzumerken, dass die Stärke der PCR-Banden mit der Stärke des Signals der Hybridisierung korreliert.

Mit dem Negativ-Mix hybridisierten ca. 70 % der PCR-amplifizierten Klone (**Abbildung 19-B**). Die Signale der Tester-Sonde zeigten eine vollständige Übereinstimmung mit den positiven Klonen der Negativkontrolle (**Abbildung 19-A**).

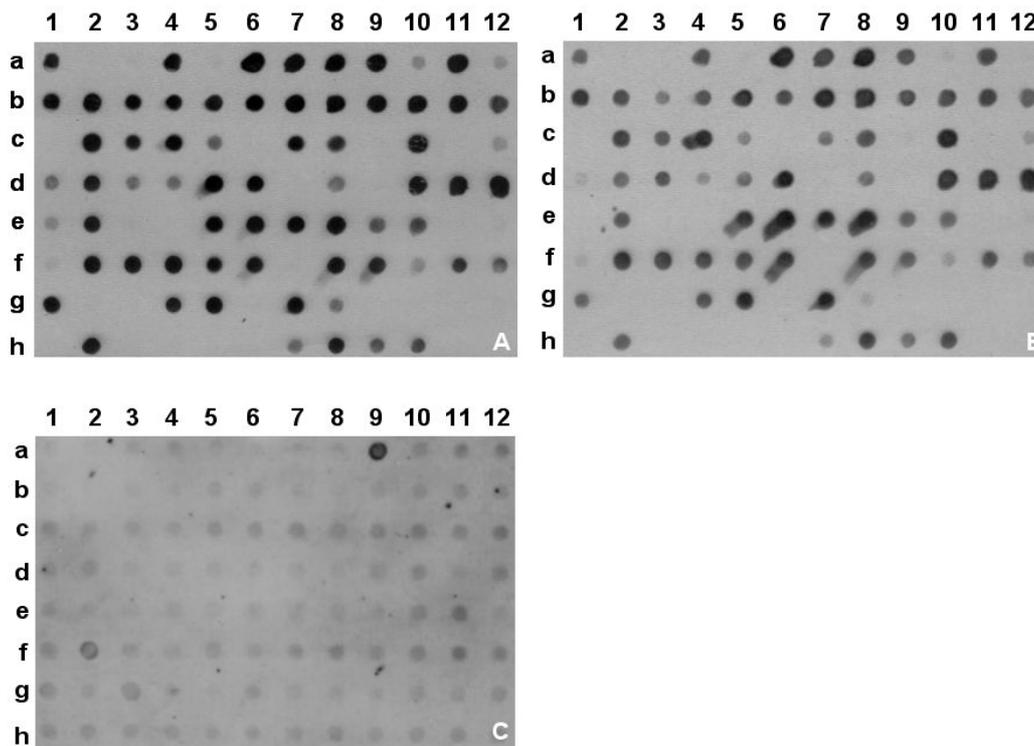


Abbildung 19. Hybridisierungen mit dem Tester (A) oder mit dem Negativ-Mix (B) als Sonde bestätigen die bakterielle Kontamination. Die Hybridisierung mit dem Driver (C) bestätigt den Erfolg der Subtraktion.

Pro Spot wurden 10 μ l von den Screening-PCR-Produkten mit NaOH denaturiert und auf einer Nylonmembran fixiert. Exponierungszeit: (A): 5 min., (B): 10 min., (C): 40 min.

Eine Hybridisierung markierter ribosomaler 23S RNA ergab 80-90 % positive Klone. Sie diente als Hinweis, dass die Kontamination schon im Ausgangsmaterial vorhanden war und daher nicht zu vermeiden ist. Wahrscheinlich sind die Enzyme, die für die globale

Amplifikation verwendet werden, mit bakterieller DNA kontaminiert. Diese bakteriellen DNA Sequenzen werden durch die globale Amplifikation ebenfalls amplifiziert. Sie sind als Schmier bei der Negativkontrolle zu sehen. Da sowohl der Tester als auch der Driver nach diesem Verfahren vorbereitet werden, ist die bakterielle Kontamination in beide vorhanden (s.4.3.2).

Die Stärke der Kontamination des Drivers und Testers wurde mit Hilfe einer PCR auf die ribosomale Pseudomonas 23S-RNA abgeschätzt (**Abbildung 20**). Da die Kontamination im Tester die des Driver überwog, wird sie wahrscheinlich durch die Subtraktion nicht reduziert, sondern im Gegenteil wie differentielle Produkte angereichert. Deshalb wurde beschlossen, dem Driver diese kontaminierenden Sequenzen zuzugeben und die Subtraktion zu wiederzuholen.



Abbildung 20. Abschätzung der Kontamination in Tester und Driver.

Eine PCR zum Nachweis der ribosomalen Pseudomonas 23S-RNA wurde auf 0,1 µg cDNA-Amplifikat der einzelnen Mammakarzinom-Zellen (1-6; im Tester sind 1, 3, 5 und 6 vorhanden) und dem cDNA-Pool von 100 000 normalen KM-Zellen dreier Patienten (7-9) durchgeführt. Primer: Pseudo-5'/-3', 58°C, 25 Zyklen; 2,5 % Agarosegel, EtBr

3.1.2.2 Neuer Driver (Subtraktion #5)

Die Konzentrationsverringerng einer bestimmten Sequenz im Endprodukt durch die Zugabe dieser spezifischen Sequenz zum Driver wurde bereits bei einem anderen Subtraktionsansatz erfolgreich getestet (Hubank & Schatz, 1994). Das gleiche Vorgehen wurde hier durchgeführt. Durch die Reamplifikation einer Negativkontrolle eines globalen cDNA-Amplifikationsexperiments wurden kontaminierende Sequenzen gewonnen und diese wurden zu dem Knochenmark gemischt. Die Mischung von Knochenmark und kontaminierenden Sequenzen wurde als Driver für eine neue Subtraktion (#5) verwendet.

Die Reduktion der Kontamination wurde nach der Subtraktion durch eine PCR auf die ribosomale 23S RNA (verwendete Primer: Pseudo) und einen Spot Test überprüft. Diese zwei Kontrollen zeigten, dass die Zugabe störender Sequenzen zum Knochenmark keine deutliche Reduktion der Kontamination brachte. Andere Kontrollen zeigten, dass die Subtraktion #5 eine ähnliche Qualität wie die Subtraktion #4.1 hatte, so dass beschlossen wurde, die Subtraktion #5 weiter zu analysieren.

3.2 Analyse der subtrahierten Produkte

3.2.1 Klonierung und Screening

Die Produkte der Subtraktion #5 wurden in das pGEM-T Easy Plasmid ligiert, in kompetente *E. coli* XL1-Blue Zellen elektroporiert und einer Blau-Weiß-Selektion untergezogen. Isoliert wurden 386 weiße Klone und die Größe der Inserts mittels PCR bestimmt. Obwohl durch die Unterbrechung des LacZ-Reportergens des pGEMT-Easy-Vektors Klone mit Insert weiß erscheinen, zeigten trotzdem viele dieser Klone kein Signal. Viele trugen nur sehr kleine Inserts. Mit 110 der 386 Klone wurde nach der PCR-Analyse weiter gearbeitet. Nach dem vorhergegangenen Dot-Blot-Experiment der Subtraktion #4 wurde eine Selektion von differentiellen Transkripten mittels Hybridisierung von Tester und Driver als nicht erfolgsversprechend angesehen. Daher wurden alle 110 Klone ohne zusätzliches Screening sequenziert.

Von den 110 Klonen ergaben 35 keine Ergebnisse und 13 zeigten Mischsequenzen. Dreiunddreißig Klone enthielten bakterielle ribosomale Sequenzen, und nur 29 Inserts waren humanen Ursprungs. Neun dieser Sequenzen konnten Tapasin (TAPBP) zugeordnet werden, die übrigen 20 waren einmalig auftretende humane Sequenzen (**Abbildung 22**).

3.2.2 Validierung differentiell exprimierter Sequenzen

Die mögliche differentielle Expression der 29 humanen Sequenzen wurde dann mit einer PCR auf Driver und Tester überprüft. Elf Sequenzen waren auch im Driver exprimiert und wurden aufgrund dessen nicht weiter verfolgt (in **Abbildung 22**, "+" im Driver). Neun Sequenzen zeigen eine deutlich differentielle Expression, das heißt die PCR-Reaktionen waren positiv im Tester und negativ im Driver. Bei der Sequenz #288 war die Natur ihrer Expression nach den ersten Experimentens noch unklar.

Das Ziel dieser Arbeit ist es einen Marker zu finden, mit welchem Tumorzellen in Knochenmarkproben spezifisch nachgewiesen werden können. Daher wurde die Expression dieser 9 Sequenzen in Karzinomzelllinien (ZL) und Knochenmarkproben (KM) gesunder Patienten analysiert. Sechs Sequenzen waren sowohl in den Karzinomzelllinien, als auch in den Knochenmarkproben der Normalspender nachzuweisen. Drei Sequenzen (#90, #157 und #288) erschienen differentiell exprimiert (in **Abbildung 22** als "-" in der KM Spalte zu identifizieren und **Abbildung 21**). Die Sequenzen #90 und #157 konnten als Gen-Fragment des Insulin-Like Growth Faktor I Receptors (IGF1R) bzw. der Untereinheit BAF57 des menschlichen SWI/SNF Komplexes identifiziert werden.

Der IGF1R spielt eine Rolle bei der Entwicklung vieler Tumortypen (Mauro & Surmacz, 2004), sodass es interessant sein könnte, seine Expression weiter zu untersuchen. Das hier gefundene Fragment ist allerdings eine Intron-Sequenz. Sie kann entweder eine mögliche unbekannte Spleißvariante darstellen oder aus einer DNA-Kontamination stammen. Außerdem ist der IGF1R ein ubiquitär vorkommendes Transmembranprotein, was ihn für diese Arbeit zu einem ungeeigneten Zielmolekül macht. Aus diesem Grund wurde die Expression der Sequenz #90 nicht weiter untersucht.

Die Sequenz #288 wies keine Homologien zu bekannten Sequenzen auf und wurde daher weiterverfolgt.

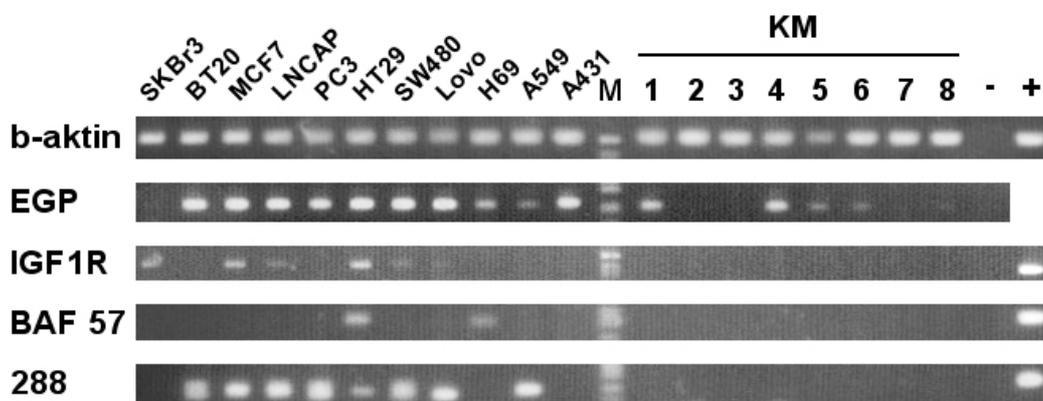


Abbildung 21. Differentielle Expression IGF1R, BAF 57 und der Sequenz #288 in Tumorzelllinien und im Knochenmark.

Die Gesamt-RNA wurde aus elf Tumorzelllinien und acht Knochenmarkproben gesunder Patienten isoliert, dann durch eine Reverse-Transkription in cDNA umgewandelt und 1/20 in Wasser verdünnt. Vor den Expressionsanalysen wurde die Qualität der cDNA mittels zwei PCR's auf β -Aktin und EGP überprüft. M: Marker. Primer: β -Aktin-5'/3' bzw. EGP-5'/3' bzw. 288-5'/3', 157-5'/3' und 90-5'/3'. 40 Zyklen; 1,5 % Agarosegel, Sybr Gold.

Klon	Homologie zu	Accession N ^o	Expression			
			Tester Tumorzellen	Driver 3 KM	ZL	KM
16-30-50- 256-284- 307-363- 369-384	TAP binding protein (Tapasin) (TAPBP)	XM 004114	+	+	n. d.	n. d.
40	phosphatidylethanolamine binding Protein (PBP)	X75252	+	+	n. d.	n. d.
89	12p13.3 PAC RPCI5-1063M23	AC005865	+	-	+	+
90	Insuline-Like Growth Factor I Receptor (IGFIR)	NM_000875	+	-	+	-
121	KIAA0152 gene product	NM_014730	+	-	+	+
131	clone RP11-197P3 on chromosome 9	AL158825	+	+	n. d.	n. d.
132	Emerin	X82434	+	+	n. d.	n. d.
157	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily e, member 1 (SMARCE1/BAF 57)	NM_003079	+	-	+	-
195	Transcription Elongation factor B (SIII), polypeptide 1 (15kD, elongin C) (TCEB1)	NM_005648	+	+	n. d.	n. d.
208	Xp22 PAC RPCI1-108M6	AC003036	+	-	+	+
217	KIAA0978	AB023195	+	+	n. d.	n. d.
218	KIAA0978	AB023195	+	+	n. d.	n. d.
219	Brain Glycogen Phosphorylase (PYGB)	NM_002862	+	+	n. d.	n. d.
267	12 BAC RP11-571M6	AC025165	+	+	n. d.	n. d.
288	Chromosome 3 BAC RP11-206M11	AC093001	+	(-)?	+	-
292	Homo sapiens mitochondrion		+	+	n. d.	n. d.
314	Procollagen-proline, 2-oxoglutarate 4-dioxygenase, alpha polypeptide II (P4HA2)	NM_004199	+	+	n. d.	n. d.
346	Lipoma preferred partner (LPP) gene, exon 5	U49962	+	-	+	+
356	FLJ00068 protein	AK024475	+	+	n. d.	n. d.
376	Chromosome 19 clone LLNLR-231D4	AC092306	+	-	+	+
383	genomic DNA of 9q32 anti-oncogen of flat epitherium cancer, segment 9/10	AB020877	+	-	-	-

Abbildung 22. Isolierte humane Sequenzen und ihre Expression in Tumor- und Knochenmarkszellen.

Alle Sequenzen wurden durch das BLAST Programm analysiert und anschließend ihre Expression in Tester (disseminierten Tumorzellen) und Driver (3 Knochenmarkproben), 11 Tumorzelllinien (ZL) und 8 Knochenmarkproben gesunder Patienten (KM, unabhängig von den Driverproben) und durch eine PCR untersucht (+ = positive PCR; - = negative PCR; n.d. = nicht durchgeführt). Sequenzen, die auch im Driver exprimiert werden, wurden nicht weiter verfolgt und ihre Expression in ZL und KM nicht getestet (n.d).

3.3 Die Sequenz #288

Die Sequenz #288 besitzt drei Merkmale, die sie zu einem interessanten Markerkandidaten machen. Zum einen zeigt die Sequenz eine differentielle Expression zwischen Tumorzelllinien und Knochenmark (3.2.2). Außerdem entspricht sie keinem bereits bekannten Gen (3.2.2). Und zuletzt besteht die klonierte Sequenz aus zwei Exons (3.3.1), weshalb man annehmen kann, dass es sich um ein mRNA Bruchstück handelt. Aus diesen Gründen wurde versucht die Sequenz #288 durch verschiedene Methoden vollständig zu isolieren und die daraus resultierende Ergebnisse zu analysieren.

3.3.1 Zuordnung der Sequenz #288

Mit Hilfe der University of California Santa Cruz (UCSC) Genome Browser Database (<http://genome.ucsc.edu>) (Karolchik *et al.*, 2003) wurde die Sequenz #288 dem 3q25.1 genomischen Bereich zugeordnet. Die im Plasmid klonierte Sequenz von 199 Basen besteht aus zwei Exons von 170 bp und 29 bp (Exon 1 bzw. Exon 2 genannt) Länge, die durch einen Intronbereich von 2438 bp getrennt sind.

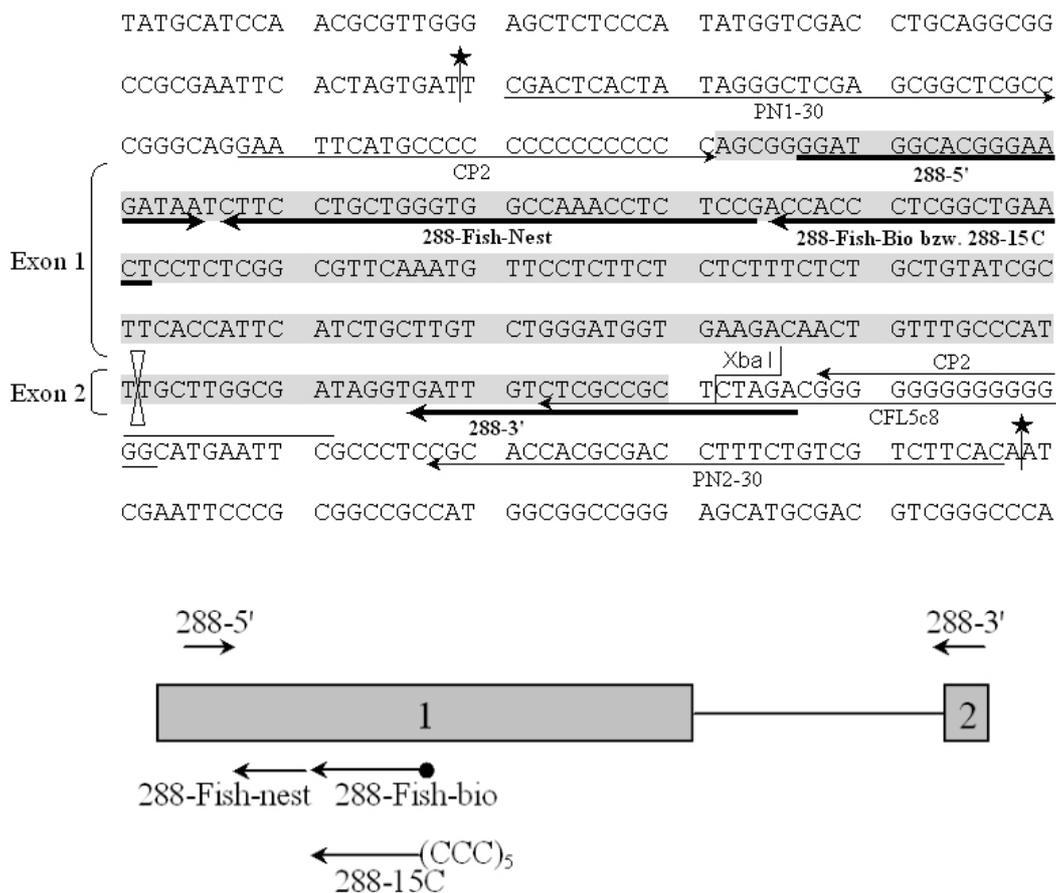


Abbildung 23. In Plasmid klonierte Sequenz #288.

Die Sequenz #288 (grau markiert) ist von den Primern, die zu ihrer Isolierung geführt haben, umrahmt (schmale Pfeile; CFL5c8 für die cDNA-Synthese, CP2 für die globale cDNA-Amplifikation, PN1-30 und PN2-30 für die Nested-PCR (PCR 2) der Subtraktion). Die Sequenzspezifischen Primer, die für die Isolierung aus Gesamt-RNA benutzt wurden, sind auch gezeigt (fett markierte Pfeile, "Fish" Primer). Sterne: Insertionspunkt im Plasmid.

Der Cfl5c8 Primer (der für die globale Amplifikation der cDNA verwendet wurde (Abbildung 4, Seite 31)) am 5'- und 3'-Ende des Fragments ermöglicht die Orientierung der mRNA (Abbildung 24). Die Transkription eines Gens, das für ein Protein kodiert, ist ein asymmetrischer Prozess, bei dem nur einer der zwei DNA-Stränge (positiv oder negativ, immer derselbe), in mRNA umgeschrieben wird. Deshalb ist die Orientierung wichtig für das korrekte Design von Oligonukleotiden zur Bestimmung des genauen 5'- und 3'-Endes der mRNA.

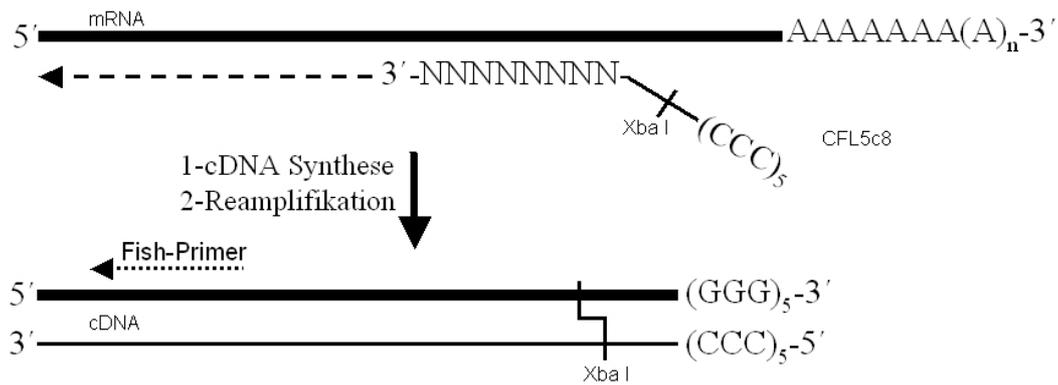


Abbildung 24. Bestimmung der Orientierung einer Sequenz nach der globalen cDNA-Amplifikation

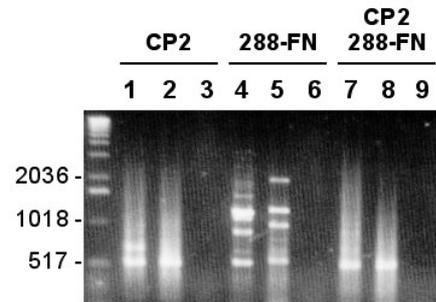
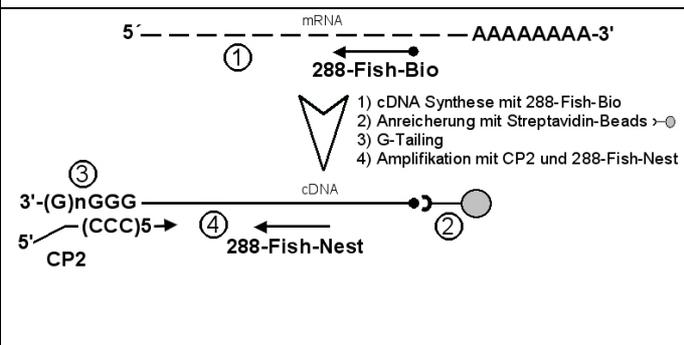
Die Orientierung der mRNA ist von der Sequenz des cDNA Stranges, der die ergänzende Sequenz des Cfl5c8 Primers trägt, zu erkennen (fett markierte Linien).

3.3.2 Identifizierung der vollständigen cDNA vom Fragment #288

Da die isolierte Sequenz #288 aus zwei Exons besteht, kann man annehmen, dass es sich um mRNA und nicht um genomische DNA handelt. Deshalb wurde versucht, die vollständige cDNA zu isolieren.

Um die gesuchte cDNA zu synthetisieren und zu amplifizieren, wurden verschiedene Strategien verwendet. Da die unterschiedlichen Versuche keine Ergebnisse gebracht haben, sind sie hier nur kurz tabellarisch angeführt.

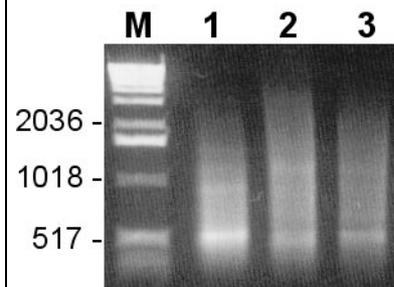
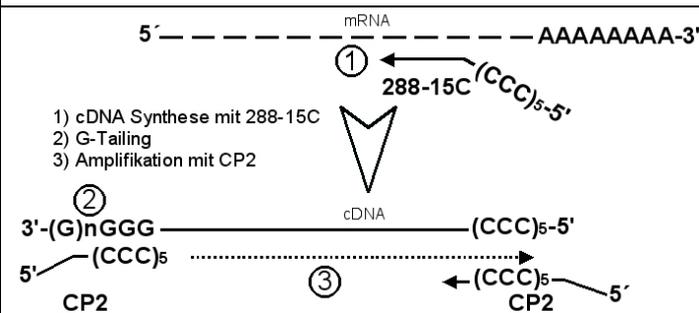
cDNA-Synthese mit dem Primer 288-Fish-Bio und magnetische Anreicherung



Methode zur Isolierung der Sequenz #288 mit dem biotinylierten Primer 288-Fish-Bio. Aus den Zelllinien Lovo und LNCAP wurde die Gesamt-RNA extrahiert, um daraus die #288-cDNA mit dem biotinylierten Primer zu isolieren. Die Produkte der sekundären PCR (4) mit den Primern 288-Fish-Nest (288-FN) und CP2 wurden auf ein Gel aufgetragen.

Die PCR mit beiden Primern (7-9) amplifizierte die gleichen Banden wie der CP2-Primer alleine (1-3), so dass es sich offensichtlich um unspezifische PCR-Produkte handelte. Die Kontroll-PCR des 288-Fish-Nest Primers ergab zahlreiche Banden (4-6), was darauf hindeutet, dass dieser Primer viele unspezifische Bindungsstellen hat. Lovo: 1, 4, 7; LNCAP: 2, 5, 8. Negativkontrolle: 3, 6, 9. 2 % Agarosegel, Sybr Gold.

cDNA-Synthese mit dem Primer 288-15C und CP2 Amplifikation



Um das Problem der Unspezifität des 288-Fish-Nest Primers zu umgehen und eine effektivere Amplifikation der #288 Produkte zu erreichen, wurde auf einen biotinylierten Primer verzichtet, an dessen Stelle trat ein #288 spezifischer Primer dessen 5'-Ende aus einer poly-C Sequenz bestand (288-15C). Eine Anreicherung mit Streptavidin-Beads war daher nicht mehr möglich, jedoch konnten nach dem G-Tailing (2) die #288 Fragmente mit einem einzigen Primer (CP2) amplifiziert werden (3).

Die Produkte dieser sekundären PCR wurden auf ein Gel aufgetragen.

Die wenigen sichtbaren Banden traten sowohl in der Lovo- (2) und LNCAP- (3) Probe als auch in der Negativkontrolle (1, cDNA-Synthese auf Wasser) auf und wurden als unspezifische Produkte gewertet. M-Marker. 2 % Agarosegel, Sybr Gold.

Magnetische Anreicherung und cDNA-Synthese mit dem Primer 288-Fish-Bio



Diese beiden PCR-Strategien wurden auf Proben der drei Zelllinien BT20, LNCAP und PC3 angewendet

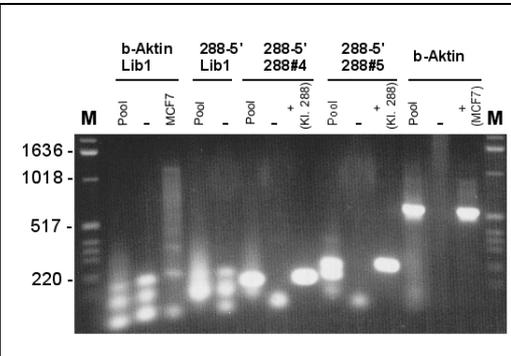
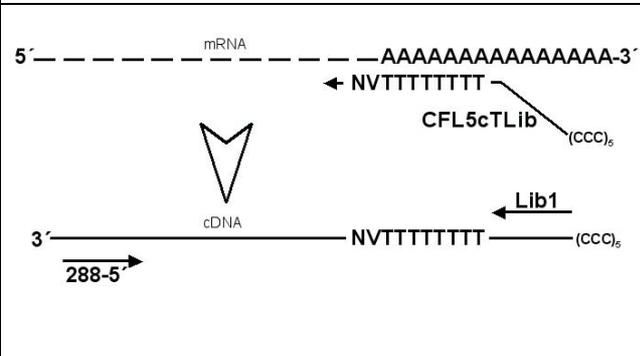
Möglicherweise stört eine hohe DNA-Konzentration während der Bindung des biotinylierten Primers 288-Fish-Bio an die Streptavidin-Beads und führt zur Amplifikation vieler unspezifischer Produkte. Es wurde versucht, den spezifischen biotinylierten Primer 288-Fish-Bio zuerst an die Streptavidin-Beads zu binden (1) und erst im Anschluß mit dem cDNA-Pool zu hybridisieren. Es wird gründlich gewaschen um die ungebundenen und unspezifischen cDNA Fragmente zu entfernen. Die Sequenz wird entweder mit den 288-Fish-Nest und CP2 Primern (2a) oder mit dem Primer CP2 allein (2b) amplifiziert.

Diese Versuche blieben erfolglos, aufgrund negativer oder unspezifischer Amplifikationen.

Da jede mit dem Primer CP2 durchgeführte PCR unspezifische Produkte erzeugte, wurden andere Primerkombinationen getestet.

Der CFL5cTLib Primer wurde bereits bei anderen cDNA-Synthesen für andere Projekte im Labor erfolgreich eingesetzt.

cDNA-Synthese mit dem Primer CFL5cTLib

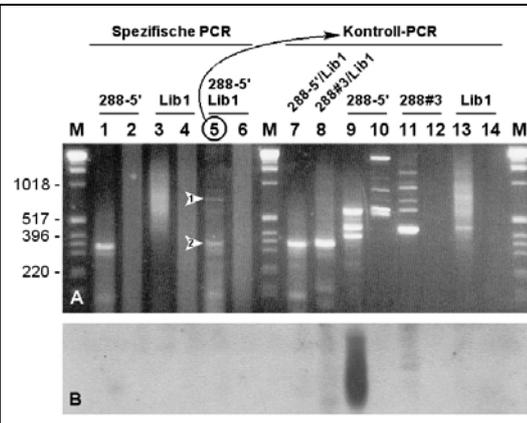
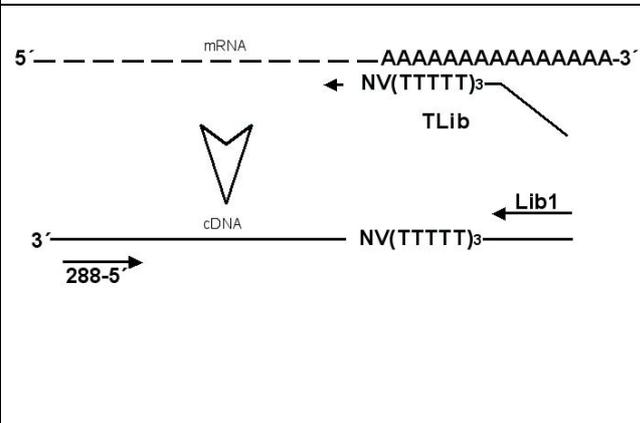


Dank seines poly-T 3'-Endes kann der Primer CFL5cTlib an den Poly-A Schwanz der mRNA binden und die cDNA-Synthese von 3'-Ende aus initiieren. Mit den Primern 288-5' und Lib1 kann in der folgenden PCR die Sequenz #288 spezifisch amplifiziert werden.
Es wurde versucht, die Sequenz aus einem cDNA-Pool aus drei Zelllinien (BT20, Lovo, PC3) zu isolieren. Als Positivkontrolle der Methode wurde mit entsprechenden Primern die β -Aktin Sequenz isoliert. Die erwartete Produktgröße betrug mit den β -Aktin-5' und Lib1-Primern 1030 bp. Die spezifische Amplifikation von der Sequenz #288 wurde mit den Primern 288-5'/Lib1 durchgeführt.

Erwartet wurde bei der PCR mit den Primern 288-5' und Lib1 einen Schmier von PCR Produkten mit einer spezifischen Bande. Die β -Aktin Kontrolle ergab die 1030 bp erwartete Bande (β -Aktin/Lib1, MCF7). Die Anwesenheit des Exons 1 der Sequenz #288 wurde nach der cDNA-Synthese mit dem Primerpaar 288-5'/288-4 und die des Exons 1+2 mit dem Primerpaar 288-5'/288-5 überprüft (Positiv Kontrolle: Kl. 288 = Klon #288). Ebenso wurde auch die Anwesenheit von β -Aktin geprüft (Positiv Kontrolle: MCF7 cDNA). Negativkontrolle (-): Wasser, M-Marker, 1,5 % Agarosegel, Sybr Gold.

► Im Zuge der hier durchzuführenden cDNA-Synthese besteht jedoch das Risiko, dass die 15 Cytidinreste am 5'-Ende des CFL5cTLib Primers unerwünschte Sekundärstrukturen ausbilden. Deshalb wurde für die folgenden cDNA-Synthesen ein Primer ohne Cytidin-15mer verwendet: der TLib-Primer

cDNA-Synthese mit dem Primer TLib



Dank seinem poly-T 3'-Endes kann der TLib Primer an den Poly-A Schwanz der mRNA binden und die cDNA-Synthese initiieren. Mit den Primern 288-5' und Lib1 kann in der folgenden PCR die Sequenz #288 spezifisch amplifiziert werden.
Die cDNA-Synthese wurde auf einen Pool von Gesamt-RNA aus BT20, SKBr3 und PC3 Zelllinien durchgeführt. Danach wurde eine spezifische PCR mit den Primern 288-5' und Lib1 um die cDNA #288 zu amplifizieren. Um die Spezifität der PCR zu überprüfen, wurden auch PCR mit jedem einzelnen Primer 288-5' oder Lib1 durchgeführt.

A- Mit den Primern 288-5' und Lib1 wurde versucht die cDNA #288 zu amplifizieren.
- Die Produkte der spezifischen PCR mit den Primern 288-5' und Lib1 (Spur 5, Pfeile 1 und 2) wurden wiederum überprüft durch eine Kontroll-PCR in der die Primerkombinationen 288-5'/Lib1, 288#3/Lib1 und jeweils nur einer dieser drei Primer eingesetzt wurden.
- Die Produkte von spezifischer- und Kontroll-PCR wurden auf ein einziges 1,5 % Agarosegel aufgetragen um den Vergleich zu erleichtern.
B- Die in Teil A aufgetrennten Produkte wurden auf einer Nylonmembran immobilisiert und mit der Dig-markierten Sequenz #288 hybridisiert (Southern Blot).

Die Tatsache, dass ein positiver Nachweis der Sequenz #288 nie erbracht werden konnte, deutete auf ein sehr niedriges Expressionsniveau in den untersuchten Zelllinien hin. Es wurde daraufhin beschlossen, die Sequenz in einer kommerziellen cDNA-Bank zu suchen.

3.3.3 Suche nach der Sequenz #288 in einer cDNA Bank

Von den zwei Möglichkeiten, entweder einen Microarray zu hybridisieren oder eine 96-Well-Platte mittels PCR nach der Sequenz #288 zu durchsuchen, wurde die PCR Lösung gewählt. Zum einen ist eine PCR sensitiver als eine Hybridisierung, zum anderen wurde das bisher bekannte 200 bp bis 400 bp große Fragment für zu kurz angesehen, um ein befriedigendes Hybridisierungssignal zu ergeben. Das „Cloning Matrix™ System“ der Firma Stratagene enthält etwa 300 000 Klone, die aus 29 verschiedenen menschlichen Geweben stammen. Dank einer Normalisierung ist die Zahl repetitiver Sequenzen reduziert, wobei die Fragmente eine mittlere Größe von 1,7 kb besitzen. Die 768 Mikrotiterplatten (384 einzelne Klone pro Platte) sind räumlich organisiert, da die Platten als Pools auf einer einzelnen 96-Well Platte auf drei Achsen abgelegt sind (C, R und S, für *Column*, *Raw* und *Surface*). Durch positive PCR-Reaktionen an bestimmten Punkten der drei Achsen kann jede von den 768 Platten individuell identifiziert werden. Eine weitere PCR ist erforderlich, um den positiven Klon in der ausgewählten Mikrotiterplatte zu identifizieren. Stratagene stellt die positiven Klone für weitergehende Analysen zur Verfügung.

3.3.3.1 Screening-PCR

Die Screening-PCR wurde mit den Primern 288-5' und 288#4 (Primäre PCR) durchgeführt, doch die PCR-Ergebnisse des drei Achsen-cDNA-Pools ergaben keinen klaren Schnittpunkt. Die erhaltenen Banden wurden isoliert, kloniert und sequenziert. Es handelte sich zwar meist um die gesuchte Sequenz #288, jedoch vereitelten einige unspezifische Produkte die sichere Selektion einer potentiell positiven Mikrotiterplatte. Um die positiven PCR-Reaktionen zu bestätigen, wurden zwei Nested-PCR-Varianten getestet, mit den Primern 288-5' und 288#12 bzw. mit den Primern 288#12 und 288#13 (**Abbildung 31** und auch 3.3.5.2), wobei letztere Kombination die besseren Ergebnisse zeigte. Leider waren auch die Nested-PCR-Reaktionen nicht reproduzierbar. Wenn auch zwei PCR-Ansätze auf der gleichen Primären-PCR dieselben Produkte lieferten, ließen sich die Ergebnisse aus zwei verschiedenen Primären-PCR nicht reproduzieren (vergleiche in **Abbildung 25** a-b mit c-d). Daher ist davon auszugehen, dass der kritische Faktor die Primäre-PCR ist.

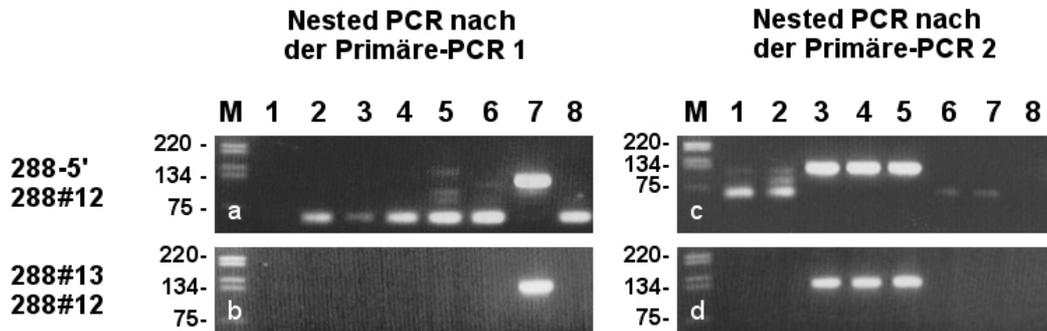


Abbildung 25. Screening der Pool 1-8 der Ebene "R" des "Cloning Matrix™ Systems": Nested-PCR.

Die Primäre-PCR mit den Primern 288-5' und 288#4 wurde zweimal wiederholt (Primäre-PCR 1 und 2). Auf jede Primäre-PCR wurden zwei Nested-PCR's durchgeführt, einmal mit dem Primerpaar 288-5'/288#12 (a und c) und einmal mit dem Primerpaar 288#13/288#12 (b und d). Nach den zwei Nested-PCR auf die Primäre-PCR 1 ergab der Pool 7 positive Banden (a und b), aber nach den zwei Nested-PCR auf die Primäre-PCR 2 ergaben die Pool 3, 4 und 5 positive Banden (c und d). 3 % Agarosegel, EtBr.

Es wurde daraufhin ein anderer primärer PCR-Ansatz gewählt, dessen PCR-Produkt beide Exons der Sequenz #288 überspannt. Zwei Primer-Kombinationen wurden dafür getestet: Primer 288#11 oder 288#10 auf Exon 2 gepaart mit jeweils dem gleichen Primer 288-5' auf Exon 1 (**Abbildung 31**). Primer 288#10 amplifizierte viele unspezifische Produkte, die mit der anschließenden Nested-PCR nicht bestätigt werden konnten. Der Primer 288#11 erzielte im Vergleich dazu bessere Ergebnisse (**Abbildung 26-a**), die in einer unabhängigen Wiederholung der PCR bestätigt werden konnten (**Abbildung 26-b**). Eine Nested-PCR mit den Primern 288#12 und 288#13 konnte die möglicherweise positive Bande der Probe #20 nicht verifizieren (**Abbildung 26-c**).

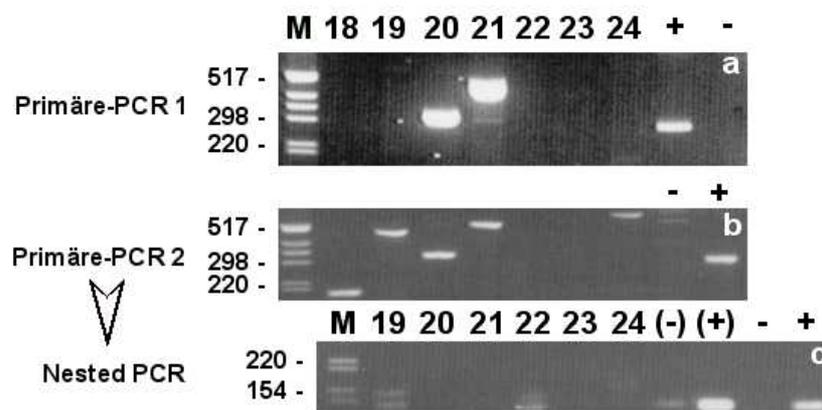


Abbildung 26. Screening-PCR der Pool 18-24 von der Ebene "C" des "Cloning Matrix™ Systems".

Die Primäre-PCR mit den Primern 288-5' und 288#11 wurde zweimal wiederholt (a und b; erwartete Produktlänge: 277 bp) und eine Nested-PCR mit dem Primerpaar 288#12/288#13 (erwartete Produktlänge: 119 bp) auf Grundlage der zweiten Primären-PCR durchgeführt (c). Positivkontrolle: SW480 cDNA, Negativkontrolle: Wasser. (+) und (-) bei (c): Nested-PCR auf die Positiv- und Negativkontrolle der Primären-PCR (b). M-Marker. 3 % Agarosegel, EtBr.

3.3.3.2 Kontrolle der Nested-PCR durch einen EarI Verdau.

Es wurde versucht, die PCR-Ergebnisse der Nested-PCR mit den Primern 288#12 und 288#13 mittels Restriktionsverdau zu überprüfen, indem die Banden aus dem Agarosegel ausgeschnitten, gereinigt und mit dem Restriktionsenzym Ear I geschnitten wurden. Ear I besitzt eine Schnittstelle innerhalb des Exon 1 und spaltet das 119 bp lange PCR-Produkt der Primer 288#12 und 288#13 in ein 86 und ein 33 bp großes Fragment. Wie aus dem Beispiel in **Abbildung 27** hervorgeht, war es dadurch möglich die falsch positiven PCR-Produkte zu identifizieren. Nach der Screening-PCR wurden die Produkte auf einem Agarosegel aufgetrennt (A), die positive Banden isoliert und mit Ear I verdaut (B). Die PCR-Produkte #15 und #25 zeigten eine partielle Spaltung, im Gegensatz dazu handelte es sich bei dem PCR-Produkt #2 um ein kontaminierendes, unverdautes Produkt, das offensichtlich in allen drei PCR-Reaktionen auftrat. Es fiel auf, dass ein vollständiger Verdau der PCR-Produkte weder durch eine verlängerte Reaktionszeit noch durch höhere Enzymkonzentrationen erreicht werden konnte.

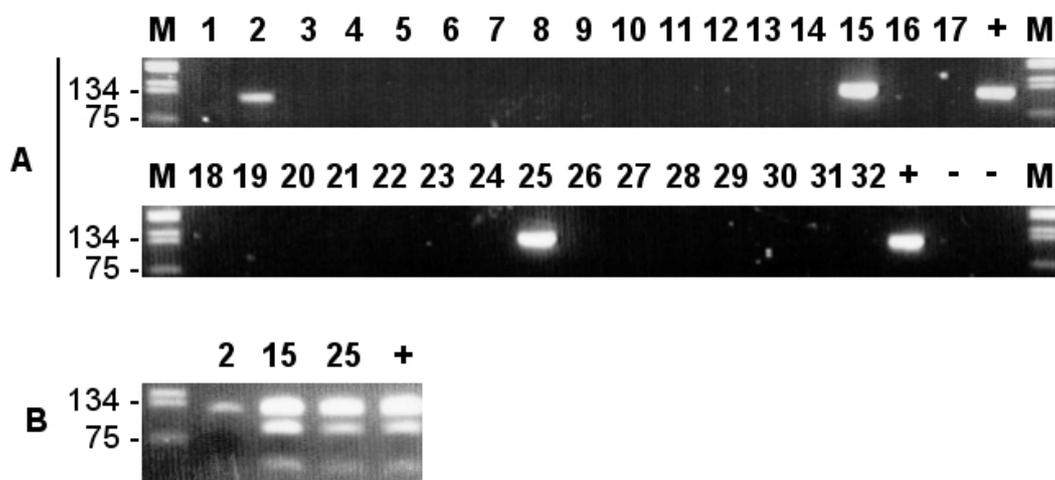


Abbildung 27. Screening-PCR der Poolen 1-32 von der Ebene "S" des "Cloning Matrix™ Systems".

A- Die Primäre-PCR wurde mit den Primern 288-5' und 288#12 durchgeführt. Die erhaltenen Produkte dienen als Ausgangsmaterial für eine Nested-PCR mit den Primern 288#12 und 288#13. **B-** Die Banden der Proben 2, 15, 25 und der Positivkontrolle, wurden ausgeschnitten, gereinigt und mit dem Ear I Restriktionsenzym verdaut. Die PCR-Produkte #15 und #25 zeigten eine partielle Spaltung, im Gegensatz dazu handelte es sich bei dem PCR-Produkt #2 um ein unverdautes kontaminierendes Produkt, das offensichtlich in allen drei PCR-Reaktionen auftrat. M-Marker. 3 % Agarosegel, EtBr.

Durch die Kombination der Nested-PCR mit dem Ear I-Verdau wurde trotz allem eine Mikrotiterplatte selektioniert, deren 384 Klone wiederum mittels PCR und Restriktionsverdau analysiert wurden. Zwei Klone zeigten positive Resultate, die jedoch nicht reproduziert werden konnten. Nach diesen Ergebnissen wurde das Screening beendet, da die Kontaminationsprobleme ein eindeutiges Ergebnis verhinderten. Eine weitere Möglichkeit ist

die sehr niedrige Expression der Sequenz #288 in der cDNA Bank, die die Detektion erschwert. Eine unterschiedliche, gewebespezifische Expression könnte erklären, warum die Sequenz #288 in verschiedenen Geweben unterschiedlich gut zu detektieren ist.

3.3.4 Expressionsmuster der Sequenz #288 in verschiedenen Geweben

Die Expression des Transkriptes #288 in verschiedenen Geweben wurde mit Hilfe einer Hybridisierung auf einen kommerziellen Nylon Array (Human Multiple Tissue Expression (MTE) Array 2, Clontech) überprüft. Da mit der Dig-UTP markierten Sonde kein Signal zu beobachten war, wurde das Experiment mit einer radioaktiv markierten Sonde wiederholt. Trotzdem konnte keine Hybridisierung mit der mRNA (isoliert von 57 Gewebearten, 8 Tumorzelllinien, 7 fetalen Gewebearten und 8 unterschiedlichen Negativkontrollen) auf der Nylonmembran beobachtet werden, und selbst die aus den bisherigen Ergebnissen erwartete Hybridisierung der spezifischen Sonde #288 mit der mRNA der Zelllinie SW480 (s. **Abbildung 21**, Seite 70) konnte nicht gefunden werden. Eine lange Expositionszeit (bis zu einer Woche) verbesserte die Ergebnisse nicht.

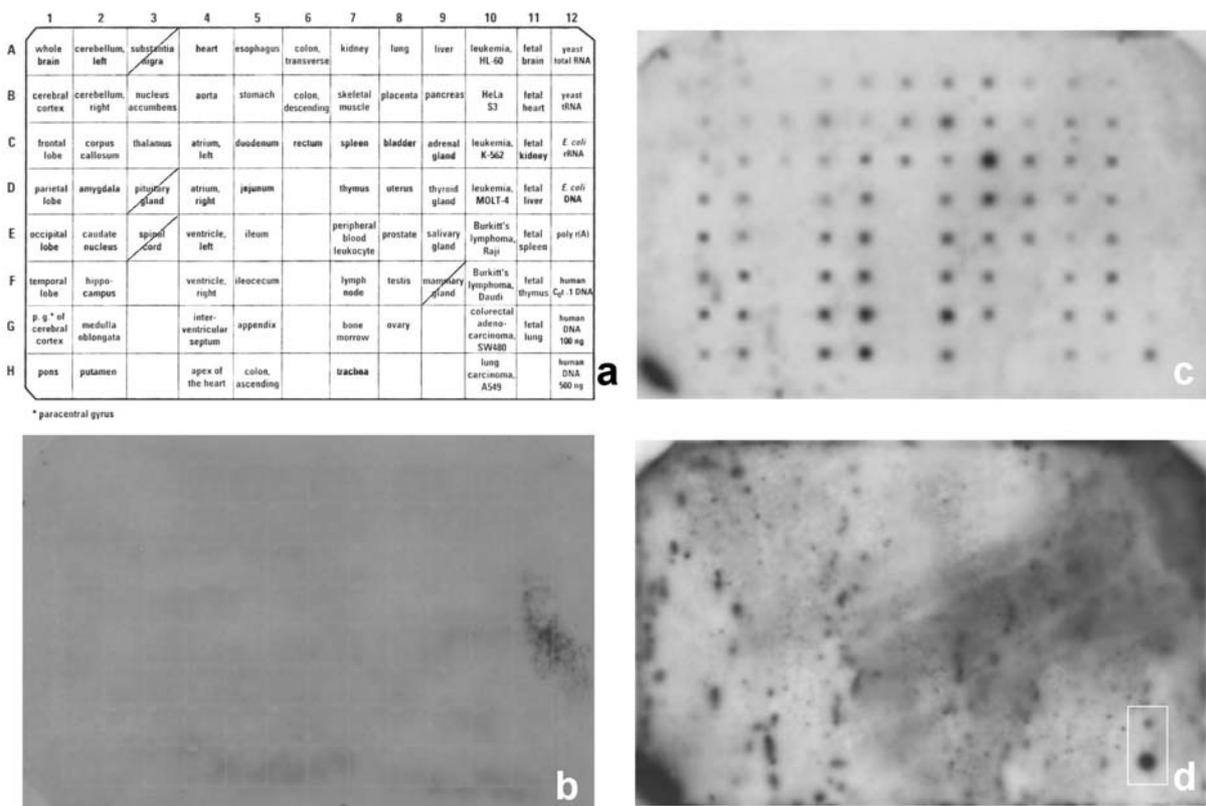


Abbildung 28. Hybridisierung des Transkriptes #288 auf den MTE Array.

Der Human Multiple Tissue Expression (MTE) Array wurde mit einer 136 bp langen Dig-markierten Sequenz #288 (Primer: 288-5'/288#4). (a) Matrix des Arrays: Die durchgestrichenen Positionen sind auf dem Filter nicht vorhanden. (b): zu erkennen ist kein Signal. **Die Arrays c und d sind aus der Diplomarbeit von Kerstin Bockler.** (c): Hybridisierung mit einer radioaktiv-markierten β -Aktin Probe: als Haushaltgene ergibt β -Aktin positive Signale in allen Geweben. (d): Hybridisierung mit einer 382 bp langen radioaktiv-markierten Sequenz #288 (Primer: NT4/288#13, s. **Abbildung 31**). Diese Probe ergibt ein Signal nur mit der gesamten Human DNA (weisses Kästchen). Je höher die DNA Konzentration, desto stärker das Signal.

Es wurde vermutet, dass die relativ geringe Konzentration (zwischen 40 und 300 ng) der verschiedenen fixierten mRNA Spots nicht genügt, um seltene Transkripte zu detektieren. Diese Vermutung kann durch die für Northern Blot empfohlene RNA-Konzentrationen von 3 µg für Poly(A)⁺mRNA weiter bekräftigt werden (Strazzullo *et al.*, 1998).

3.3.5 Elektronische Analyse

3.3.5.1 EST Analyse

Die genomische Umgebung der Sequenz #288 wurde mit dem Programm BLAT (<http://www.genome.ucsc.edu/cgi-bin/hgBlat>) auf das Vorhandensein von *expressed sequence tag* (ESTs) analysiert. Die dabei gefundenen ungespleißten humanen ESTs, die um die Sequenz lokalisiert sind (z.B. AI910205, AI910206, BF749373, AW975312, BI460663) oder einen Abschnitt aus der Sequenz #288 beinhalten (z.B. BF748881, BU618468), waren fast alle aus Mammakarzinomgewebe isoliert worden (Dias Neto *et al.*, 2000). Die zwei Ausnahmen waren aus der human Fibrosarkomzelllinie HT1080 (BG183639, (Harrington *et al.*, 2001) und aus "Large cell carcinoma" (BU175589) isoliert worden.

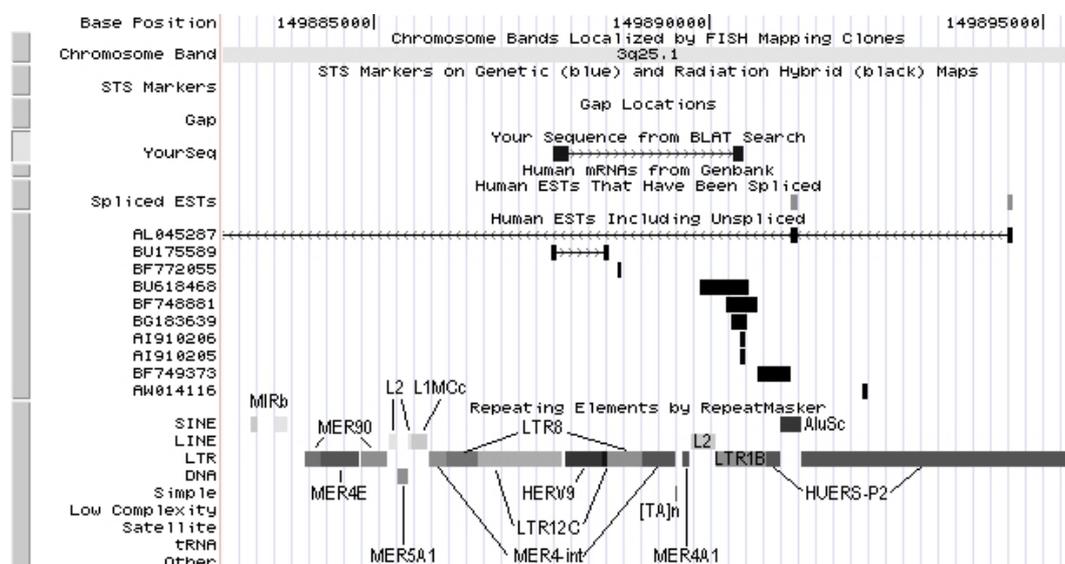


Abbildung 29. Alignment des Transkriptes #288 durch BLAT.

Die Sequenz #288 (Your Seq) ist eine gespleißte mRNA. (<http://www.genome.ucsc.edu/cgi-bin/hgBlat?> - freeze date: November 2002)

3.3.5.2 Ist das Transkript #288 eine Spleißvariante der cDNA AL045287 ?

Die ursprünglich klonierte Sequenz #288 ist 199 Basen lang. Die korrespondierende cDNA liegt auf dem Negativ-Strang des Chromosomes 3 (s. 3.3.1), in der chromosomalen Bande 3q25.1. Die cDNA besteht aus Exon 1 (170 bp) und Exon 2 (29 bp), die von einem Intron mit 2438 bp getrennt werden.

Diese zwei Exons befinden sich zwischen den Exons II und III der cDNA AL045287. Es wurde daher überprüft, ob die zwei Exons der Sequenz #288 eine alternative Spleißvariante der cDNA AL045287 sind. Dazu wurden Primer synthetisiert, die für die fünf Exons der cDNA AL045287 sowie für die zwei Exons des Transkriptes #288 spezifisch sind (**Abbildung 30**). Verschiedene Kombinationen dieser Primer wurden in PCRs auf zwei Mammakarzinom Zelllinien (BT20 und MCF7), zwei Knochenmarksproben von gesunden Patienten und auf jede einzelne Komponente des Testers und des Drivers getestet. Die PCR-Ergebnisse ergaben keine Anhaltspunkte dafür, dass das Transkript #288 eine alternative Spleißvariante von AL045287 ist. Durch die Kombination des Primers 288-5' mit 288#9, 288#10 oder 288#11 wurden jedoch längere Transkripte amplifiziert. Sie erlaubten die bekannte Sequenz des Exons 2 auf 82 bp zu erweitern.

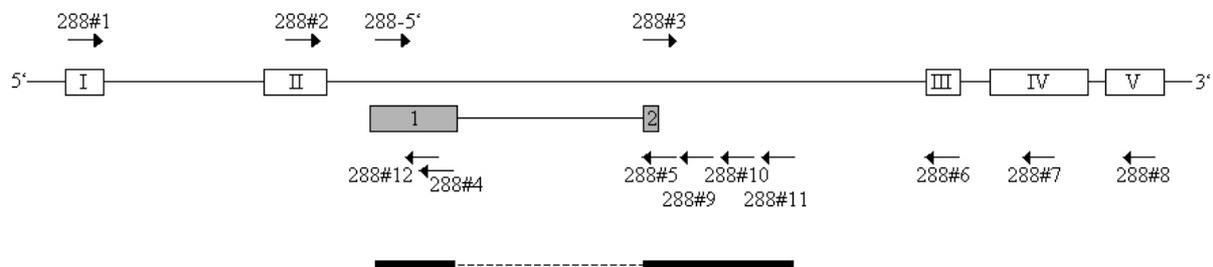


Abbildung 30. Position der verwendeten Primer in Exon- und Intronbereichen der cDNA AL045287.

Um zu testen, ob das Transkript #288 zu einer alternativen Spleißvariante des Transkriptes AL045287 gehört, wurden 13 Primer (Pfeile), die in 5'- und 3'- Richtung von der isolierten Sequenz #288 lagen, in verschiedenen Kombinationen getestet. Die cDNA AL045287 besteht aus fünf Exons, die als weiße Rechtecke dargestellt sind. Die zwei Exons des Transkriptes #288 sind unten als grau markierte Rechtecke verzeichnet. Die schwarzen Linien unten stellen das mit der Primerkombination 288-5'/288#11 amplifizierte Transkript dar.

3.3.5.3 Ist das Transkript #288 die vorhergesagte cDNA NT_005616.25 ?

Nach dem UCSC Genome Browser von Dezember 2001 wird im Bereich der Sequenz #288 auf dem positiven Strang des Chromosomes 3 ein proteinkodierender Bereich vorhergesagt. Diese vorhergesagte cDNA ist 1395 Basen lang und enthält ein offenes Leseraster (ORF) von 464 Aminosäuren. Das vorhergesagte Protein-Sequenzstück zeigt (N-Terminal) hohe Ähnlichkeit zu einem "unbekannten" Protein AAL55829 (Genbank), das mit "*Novel human cDNA clones with function of inhibiting cancer cell growth*" annotiert ist. Diese vorhergesagte cDNA mit der Referenznummer NT_005616.25 beinhaltet neun Exons. Interessanterweise korreliert das Exon 1 der cDNA-Sequenz #288 zu 100 % mit dem Exon VI der vorhergesagten NT_005616.25 cDNA. Außerdem korreliert das mit der Primerkombination 288-5'/288#11 amplifizierte Exon 2 (3.3.5.2) teilweise mit dem Exon IV der vorhergesagten

NT_005616.25 cDNA. Es wurde mit verschiedenen Primerkombinationen getestet, ob die Sequenz #288 mit der vorhergesagten cDNA NT_005616.25 übereinstimmt. Die Lokalisation der Primer ist in **Abbildung 31** gezeigt.

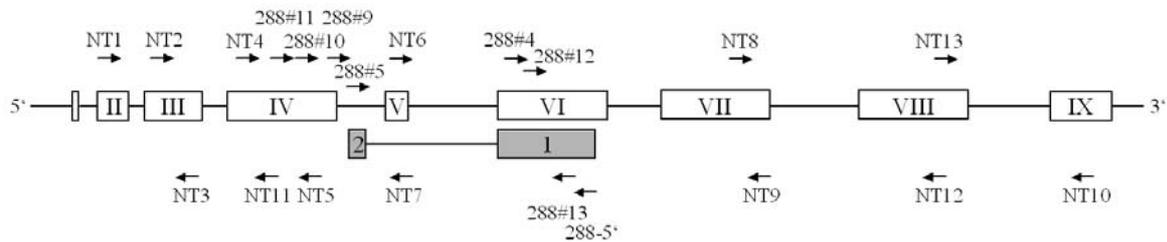


Abbildung 31. Position der verwendeten Primer in Exonbereichen des vorhergesagten Gens NT_005616.25.

Auf dem positiven Strang des Chromosombereiches 3p25.1 liegt die vorhergesagte cDNA NT_005616.25 (die neun Exons sind als weiße Rechtecke dargestellt). Da das Exon 1 des Transkriptes #288 (grau markiert) zu 100 % mit dem Exon VI der vorhergesagten NT_005616.25 cDNA korreliert, wurden Primer (Pfeile), die in 5'- und 3'- Richtung von der isolierten Sequenz #288 und in Exonbereichen von NT_005616.25 lagen, in verschiedenen Kombinationen getestet.

Die PCR-Experimente wurden mit cDNA aus der Zelllinie SW480 durchgeführt, die eine relativ hohe Expression der Sequenz #288 aufweist. Nur die Kombination der Primer NT4 und 288#13 ergab zwei PCR-Produkte mit einer Länge von ungefähr 400 bp (**Abbildung 32**).

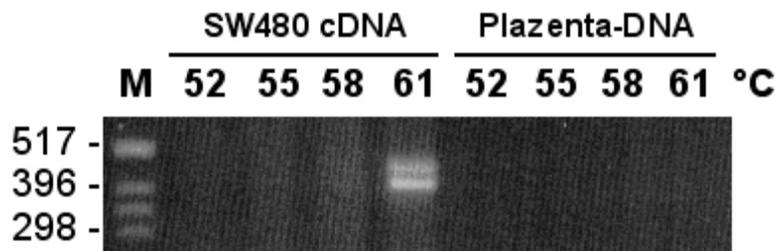


Abbildung 32. Amplifikation von zwei Produkten aus der SW480 Kolonkarzinom Zelllinien mit den Primern NT4 und 288#13.

Die PCR wurde mit SW480 cDNA und mit einem Temperaturgradienten (52, 55, 58 und 61°C) durchgeführt. Als Negativkontrolle diente genomische Plazenta-DNA. 40 Zyklen; 2 % Agarosegel, Sybr Gold.

Die mit diesen Primern gefundenen PCR-Produkte wurden beide sequenziert. Es handelte sich um zwei Größenvarianten der Sequenz #288, die sich in Exon 1 um 31 Basen unterschieden (370 bp bzw. 401 bp). Exon 1 entspricht größtenteils dem Exon VI des vorhergesagten Gens NT_005616.25, Exon 2 entspricht dem Exon IV jedoch mit einer abweichenden Exon-Intron-Grenze (**Abbildung 33**). Falls es sich also bei der Sequenz #288 tatsächlich um die vorhergesagte cDNA NT_005616.25 handelt, ist das vorhergesagte Exon V nicht in der exprimierten mRNA enthalten. Darüber hinaus ergaben sich keine Hinweise, dass NT_005616.25 der Sequenz #288 entspricht.

Beide Spleißvarianten wurden mit den Primern 288#13 und NT4 aus verschiedenen Zelllinien (SW480, BT20, LNCAP, PC3) mehrmals amplifiziert und sequenziert.

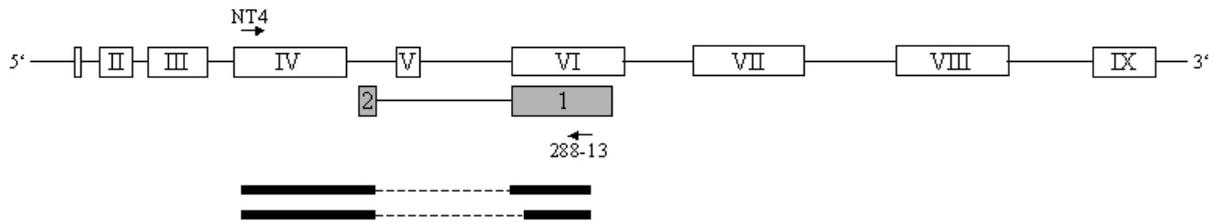


Abbildung 33. Position der Sequenz #288 und der zwei Spleißvarianten in Relation zum vorhergesagten Gen NT_005616.25.

Mit den Primern NT4 und 288#13 (Pfeile) wurden zwei Spleißvarianten des Transkripts #288 isoliert (schwarze Linien). Die ursprünglich isolierte Sequenz #288 ist grau markiert.

3.3.5.4 Suche nach Leserahmen

Als Transkript-Referenz für die *in silico* Analysen diente das durch die PCR mit den Primern NT4 und 288#13 amplifizierte Fragment. Von diesem Transkript #288 sind zwei Spleißvarianten vorhanden, auf denen mehrere offene Leserahmen (ORF) mit dem Programm ORF Finder (<http://www.ncbi.nlm.nih.gov/gorf/gorf.html>) identifiziert werden konnten (Abbildung 34).

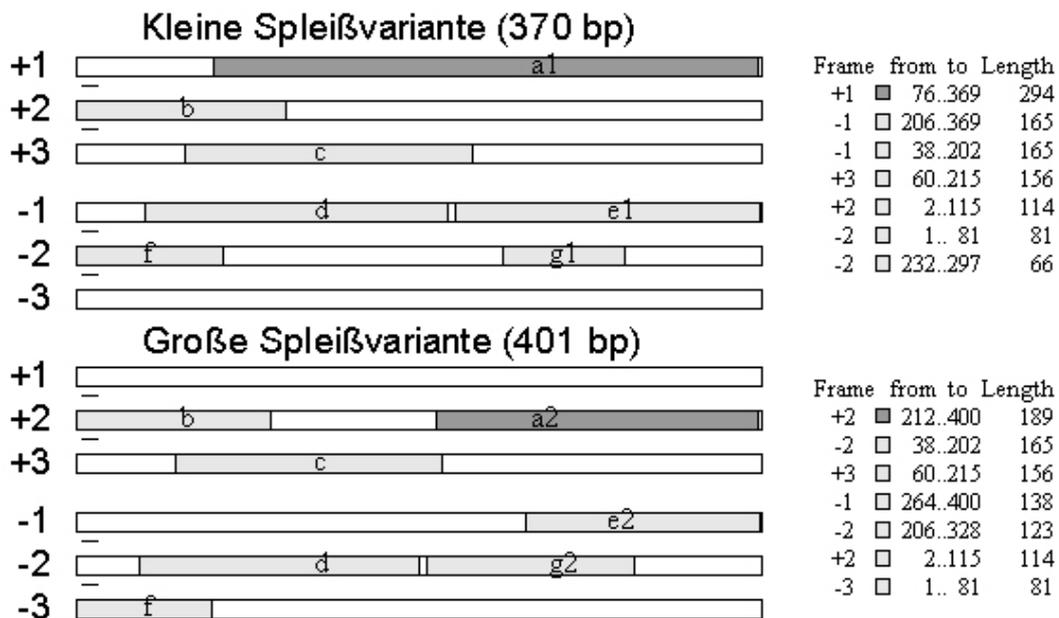


Abbildung 34. Offene Leserahmen (ORF) beider Spleißvarianten des Transkripts #288.

Die Analyse wurde mit dem Programm ORF Finder von NCBI durchgeführt (<http://www.ncbi.nlm.nih.gov/gorf/gorf.html>). Als Referenz wurde das mit den Primern NT4 und 288#13 amplifizierte Fragment genommen. + und - Zeichen entsprechen der Leserahmenorientierung und 1, 2 und 3 stellen die drei möglichen Leserahmen für jede Orientierung dar. Dunkelgrau markiert sind die ORF, deren Aminosäuresequenzen Übereinstimmungen mit bereits bekannten Aminosäuresequenzen zeigen.

Nur bei der Aminosäuresequenz eines einzigen ORFs kann die Analyse mit dem Programm BLAST Übereinstimmungen mit bereits bekannten Aminosäuresequenzen finden. Der längste und gleichzeitig interessanteste ORF wird mit der kleinen alternativen Spleißvariante erhalten; Es handelt sich um den ORF "a1", dessen Aminosäuresequenz in **Abbildung 35** dargestellt ist.

NH₂-LQHSLRVSAASFLKS^{VRRR}THQFQTHFGDLDGTIAYHRVVRPSP^{IAERR}DNHLS^{PSNKQ}MNGETIQ
QRKREEEHLNAERSSAEGGRRGLATQQEDYLP-COOH

Abbildung 35. Aminosäuresequenz des ORFs "a1" der kleinen Spleißvariante der Transkript #288 (NT4/288#13).

M: potentielles Start Codon.

Eine Suche mit dem BLAST Programm (Altschul *et al.*, 1997) zeigt, dass zwei Bereiche dieser potentiellen Proteine Homologien zu bekannten Proteinen haben (**Abbildung 36**). Das Protein-Sequenzstück zeigt (N-Terminal) eine Ähnlichkeit zu der DNA-Polymerase NP041148 eines Herpesvirus und (C-Terminal) zu dem hypothetischen Protein AAA81016, das mit "*novel transcript; similar to transcription factors activation domains*" bezeichnet ist. Die Homologie betrifft das carboxyterminale Ende eines potentiellen Fusionsproteins, das bei der Entstehung eines Lipoms beteiligt sein kann. Nach einer Translokation ist der DNA-Bindungsbereich des *Architectural Factor HMGI-C* mit einer sauren Transaktivierungsdomäne verbunden, die auch bei unserem potentiellen Protein zu finden ist. Der betroffene Bereich trägt viele saure Aminosäuren (besonders Glutaminsäure, E) sowie Serin- und Threonin-Reste (S und T) (Ashar *et al.*, 1995). Solche Bereiche spielen oft in der Transkriptionsaktivierung eine Rolle und können die Transkription sowohl von fern der TATA-Box gelegenen Positionen als auch von nahe gelegenen Positionen anregen (Mitchell & Tjian, 1989; Seipel *et al.*, 1992).

AAA81016: novel transcript; similar to transcription factors activation domains; linked at 5' end to AT hook motif of HMGI-C

Query: 55 PSNKQMNGETIQQRKR-EEEHLNAERSSAEGGRRGL 89
P+++Q++ +T++QRK+ EEEHLN ERSSA GG RG+
Sbjct: 10 PADQQVDSQTM^{RQRK}EEEEHLN^{TERSS}SAGGGWRGV 45

NP 041148: DNA polymerase [Ictalurid herpesvirus 1]

Query: 2 QHSLRVSAASFLKS^{VRRR}THQFQTHFGDLDGTIAYHRVVRPSP^{IAERR}DNHLS^{PSNKQ}MN 61
QH +++ F+K++ R H ++GD D T+ YH PS +E + + +M
Sbjct: 772 QHKIKL-VNEFIKTLNRTGHS^{LFPNYGDT}DSTMLYH---PSDESETQLED^{MVTLEDEM}R 826

Abbildung 36. Protein Homologien.

Als "Query" dient der ORF "a1", dessen Aminosäuresequenz in **Abbildung 35** dargestellt ist. Bei der potentiellen Transaktivierungsdomäne des AAA81016 Proteins sind die sauren Aminosäuren (Glutaminsäure - E) unterstrichen und die Aminosäuren Serin und Threonin (S, T) fett gedruckt.

3.3.5.5 RepeatMasker Analyse

Die Sequenz #288 ist in einem genomischen Bereich lokalisiert, in dem viele Sequenzwiederholungen zu finden sind, insbesondere *Long Terminal Repeat* (LTR) Sequenzen von der Endogenen Retrovirus Familie 1 (ERV1) (**Abbildung 29**, Seite 80). Nach einem Suchlauf mit dem Programm RepeatMasker (<http://woody.embl-heidelberg.de/repeatmask/>), wird eines der Exons der Sequenz als pTR5 Element erkannt. Das pTR5 Element ist eine repetitive Sequenz, die in embryonalen Karzinomzelllinien entdeckt wurde (La Mantia *et al.*, 1989) und als Transkript des menschlichen Endogenen Retrovirus HERV9 identifiziert wurde (Lania *et al.*, 1992).

3.3.6 Die ERV9-LTR-Hypothese

3.3.6.1 Welche Orientierung hat die Sequenz #288 ?

Die Isolierung der vollständigen mRNA #288 aus Gesamt-RNA ist erfolglos geblieben. Alle verwendeten Methoden stützen sich auf die Transkriptions-Orientierung der potentiellen mRNA, die aus der Position des Cfl5c8 Primers, der für die globale cDNA-Amplifikation benutzt wurde, hergeleitet werden kann. Dennoch wird die theoretische Orientierung auf dem komplementären Strang durch zwei Befunde in Frage gestellt:

- Zum einen von der Korrelation der Sequenz #288 mit der vorhergesagten NT_005616.25 mRNA, die auf dem positiven Strang des Chromosoms 3 liegt,
- zum anderen von der Gegenwart einer langen terminalen Sequenzwiederholung (*Long Terminal Repeat*, LTR) des *Human Endogenous Retrovirus* ERV9 auf eben diesem positiven Strang, der die Transkription von mRNA steuern könnte. Mit Hilfe des Programms RepeatMasker (<http://woody.embl-heidelberg.de/repeatmask/>) wurde festgestellt, dass eines der beiden Exons der Sequenz #288 eine 100 %ige Homologie mit einem pTR5 Repetitiven Element (La Mantia *et al.*, 1989) aufweist, das ein Transkript des humanen endogenen Retrovirus ERV9 ist (Lania *et al.*, 1992).

In Anbetracht dieser Hypothesen wurde ein neuer Anlauf im Labor unternommen, um die gesamte Sequenz #288 zu isolieren (**Diplomarbeit von Kerstin Bockler**). Ausgehend von der These, dass sie auf dem positiven Strang des Chromosoms 3 lokalisiert ist, wurde der Poly(A)-Schwanz mit neuen Primern in der anderen Richtung gesucht. Der Poly(A)-Schwanz wurde ~80 bp stromaufwärts von der Bindungsstelle des Primers 288#13 gefunden. Fünfzehn Nukleotide stromaufwärts dieser Poly-A-Stelle liegt ein konserviertes Polyadenylierungssignal AATAAA (**Abbildung 39**, Seite 90). Außerdem wurden PCR-

Produkte nur mit Primern, die stromabwärts von der wahrscheinlichen Transkriptionstelle liegen, erhalten.

Dieses Ergebnis bestätigt, dass die Sequenz #288 auf dem positiven Strang des Chromosoms 3 liegt und erklärt, warum sämtliche Versuche, den Poly(A)-Schwanz in der anderen Richtung zu finden, erfolglos geblieben sind. Die Position des Cfl5c8-Primers sollte die Orientierung des Fragments ohne Zweifel aufklären können, wie in **Abbildung 4** (Seite 31) zu sehen ist. Leider ist wahrscheinlich im Verlauf der globalen cDNA-Amplifikation eine unerwartete Reaktion geschehen. Die Reverse Transkriptase (RT) hat die Fähigkeit, ein DNA-Fragment als Matrize zu benutzen, wenn ein Primer zu Verfügung steht. Man kann daher nicht ausschliessen, dass während der reversen Transkription ein ungebundener Cfl5c8-Primer eine unspezifische Bindung mit einer neu synthetisierten cDNA eingeht und so eine Amplifikation ermöglicht. Dieses ist möglich, da der Cfl5c8-Primer im Überschuss zugegeben wird und da seine Sequenz acht *wobble*-Basen am 3'-Ende trägt. Eine solche Reaktion ist mit dem Cfl5cT-Primer unwahrscheinlicher, da das T-Homopolymer-3'-Ende weniger Möglichkeiten für eine zufällige Bindung des Primers bietet.

3.3.6.2 Die ERV9 LTR Struktur

Das Exon 1 der Sequenz #288 befindet sich in einer langen terminalen Sequenzwiederholung (*Long Terminal Repeat*, LTR) des *Human Endogenous Retrovirus* ERV9. Endogene Retroviren und retrovirale Elemente sind im Genom vieler Wirbeltiere zu finden. Man schätzt, dass bis zu 8 % des menschlichen Genoms aus ERV bestehen. Für das ERV9 wurden annähernd 120 Loci auf fast allen Chromosomen gefunden (Svensson *et al.*, 2001). Außerdem wurden auch bis zu 4000 einzelne (*solitary*) ERV9 LTRs detektiert.

Die allgemeine Struktur von viralen Retrotransposons ist in **Abbildung 37** dargestellt. Die LTR sind ein besonderes Merkmal der viralen Retrotransposons und flankieren den zentralen proteinkodierenden Bereich. LTR können die Initiation der Transkription oder die Spaltung des Primärtranskripts und das Anhängen einer Poly(A)-Schwanz steuern, je nach ihrer Position strangauf- oder strangabwärts des zentralen Bereichs. Alle ERV LTRs sind nach dem gleichen Muster, U3-R-U5, organisiert.



Abbildung 37. Allgemeine Struktur eukaryotischer viraler Retrotransposons.

Die zentrale proteinkodierende Region ist von zwei sequenzspezifischen langen terminalen direkten Sequenzwiederholungen (LTRs) flankiert. Die linke LTR-Sequenz steuert die Initiation der Transkription durch die zelluläre RNA-Polymerase II am ersten Nukleotid des linken R-Bereichs. Das Primärtranskript reicht über die rechte LTR-Sequenz hinaus, welche die Spaltung des Primärtranskripts am letzten Nukleotid des rechten R-Bereichs und das Anhängen eines Poly(A)-Schwanzes durch zelluläre Enzyme steuert (Die verschiedenen Bereiche sind nicht im selben Maßstab dargestellt; der proteinkodierende Bereich nimmt mindestens 80 Prozent eines Retrotransposons ein).

Im Vergleich mit LTRs von anderen ERV-Familien haben die U3-Segmente der ERV9 LTRs eine kennzeichnende Sequenz: Sie bestehen aus einem charakteristischen Bereich und 5 bis 17 *tandem repeats* (auch E-Elemente genannt) aus 37 bis 41 Basen, die sich in 4 Subtypen (1, 2, 3 und 4) klassifizieren lassen. Das U5-Segment enthält 80-bp lange B-Elemente, ebenfalls in variierender Anzahl, so dass die ERV9 LTR-Sequenzen sehr unterschiedlich sind (La Mantia *et al.*, 1991; Long *et al.*, 1998). Im Durchschnitt sind die ERV9 LTRs 1,8 kb lang (Costas & Naveira, 2000). Zwischen den zwei U3- und U5-Segmenten befindet sich eine Promotor-Sequenz und ein R-Segment, das den Transkriptionsstartpunkt (+1) trägt. Eventuell kann auch die Position +4 als sekundärer Transkriptionsstartpunkt benutzt werden (La Mantia *et al.*, 1992; Strazzullo *et al.*, 1994).

3.3.6.3 Ist die Sequenz #288 von einem einzelnen ERV9 LTR gesteuert?

Wenn man die Sequenz #288 mit dem UCSC Genome Browser von November 2002 (Karolchik *et al.*, 2003) suchen lässt, wird sie in den Bereich 3q25.1 lokalisiert. Die Sequenz liegt geteilt auf zwei Exons vor. Laut Sequenz-Orientierung nach der globalen cDNA-Amplifikation liegen die zwei Exons auf dem negativen Strang zwischen den Exons II und III der cDNA AL045287; sie sind aber keine alternativen Spleißprodukte von der cDNA AL045287 (3.3.5.2). Alle Versuche, die cDNA mittels des Poly-(A) Schwanzes und eines spezifischen Primers zu isolieren, blieben erfolglos. Interessanter Weise korreliert ein Exon der Sequenz #288 zu 100 % mit einem Exon der vorhergesagten cDNA NT_005616.25, welche auf dem positiven Strang liegt. Obwohl unsere Ergebnisse vermuten lassen, dass die vorhergesagte cDNA NT_005616.25 der Sequenz #288 nicht entspricht, stellt sich die Frage, ob die Sequenz #288 nicht auf dem positiven Strang gelegen ist.

Eine Analyse durch das Programm RepeatMasker (Jurka, 2000) zeigt, dass der Chromosomenbereich 3q25.1 viele repetitive Elemente enthält, die meisten davon sind LTRs der ERV1-Familie. Das Exon 2 der Sequenz #288 hat eine 100 %ige Homologie mit einem pTR5 Repetitiven Element (La Mantia *et al.*, 1989), welches ein Transkript des humanen endogenen Retrovirus ERV9 ist (Lania *et al.*, 1992).

Eine genauere Analyse der die Sequenz #288 flankierenden chromosomalen Bereiche zeigt, dass der positive Strang einen kompletten ERV9 LTR von 1244 bp Länge (unter LTR12C im Repbase registriert) mit dem typischen U3-R-U5 Aufbau beinhaltet (**Abbildung 38**). Der U3-Bereich besteht aus einer nicht-repetitiven Sequenz von 641 bp und einem *Tandem Repeat* Bereich von fünf E-Elementen. Diese E-Elemente lassen sich sehr gut in vier Subtypen klassifizieren und zeigen eine typische 1-2-3-4-1 Reihenfolge (Long *et al.*, 1998). Die nicht-repetitive 641 bp lange Sequenz enthält sechs CCACC Motive, der *Tandem Repeat* Bereich beinhaltet (mit Ausnahme von Subtyp 4) pro E-Element je ein CCAAT Motive.

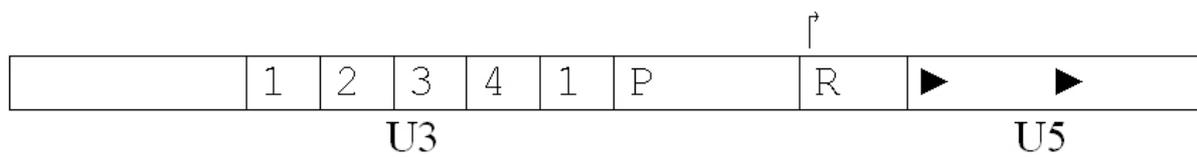


Abbildung 38. Struktur des ERV9 LTR.

Der U3-Bereich besteht aus einer nicht-repetitiven Sequenz und aus fünf Sequenzwiederholungen (E-Elemente). Diese E-Elemente können in vier Subtypen klassifiziert werden, und zeigen eine typische 1-2-3-4-1 Reihenfolge (Long *et al.*, 1998). P: Promotor. Winkelförmiger Pfeil: Transkriptionsinitiationspunkt. Die schwarzen Pfeile stellen zwei lange Sequenzwiederholungen (B-Elemente) in dem U5-Bereich dar.

Der Promotor ist 97 bp lang und ist direkt stromaufwärts der R-Region lokalisiert, deren 5'-Ende die Transkriptionsinitiationsstelle (+1) für die retrovirale RNA-Synthese ist. Der Promotor trägt 23 Basen vor dem Initiationspunkt der Transkription eine TATA-like Sequenz AATAAA, und außerdem potentielle Bindungsstellen für andere Transkriptionsfaktoren wie AMLI (ACCAC), CCAAT, SP1 und GATA. Die R-Region ist 98 bp lang. Sie enthält eine konservierte Hind III-Schnittstelle und ein potentielles Polyadenylierungssignal, 35 beziehungsweise 59 Basen stromabwärts des Transkriptionsinitiationspunktes (Lania *et al.*, 1992). Der U5-Bereich ist 207 bp lang und enthält zwei B-Elemente. Mit dem RepeatMasker Programm kann direkt strangabwärts davon eine 600 bp lange innere Sequenz vom ERV9 beobachtet werden, gefolgt von einem einzigen B-Element eines LTR12C. Allgemein enthält dieser Chromosomen-Bereich viele repetitive Elemente von LTRs, aber kein komplettes ERV9. Der hier entdeckte LTR gehört zu den einzelnen ERV9 LTRs (*solitary LTR*), die im Genom verstreut vorkommen.

Im UCSC Genome Browser (*freeze date*: November 2002) entspricht die Position 149886571 von Chromosom 3 der ersten Base dieses ERV9 LTR. Das 5'-Ende des isolierten mRNA-Fragments #288 befindet sich an Position 149887669, in der Mitte des U5-Segments. Außerdem wurde ~80 bp stromabwärts des 288#13 Primers ein Poly-(A) Schwanz gefunden, dessen Anheftung wahrscheinlich von einem 15 bp stromaufwärts liegenden konservierten Polyadenylierungssignal (AATAAA) gesteuert wird. Zuletzt wurde eine gespleißte mRNA mittels einer PCR mit den Primern F1 und 288#13 aus verschiedenen Tumorzelllinien isoliert. Allgemein konnten die erwarteten Produkte mit dem 288#13 Primer und einem zweiten Primer, der stromabwärts des Transkriptionsinitiationspunkts liegt, immer amplifiziert werden. Dies war allerdings nicht mehr der Fall, wenn als zweiter Primer ein stromaufwärts des Transkriptionsinitiationspunkts liegendes Oligonukleotid (wie z.B. NT2) benutzt wurde (**Diplomarbeit von Kerstin Bockler**). Diese Ergebnisse stimmen mit der Tatsache überein, dass ein proviraler LTR drei funktionale Bereiche (U3, R und U5) enthält. Dabei liegen die Kontrollelemente der Transkription im U3-Bereich, während der Transkriptionsinitiationspunkt den Anfang des R-Bereichs festlegt. Deswegen sollte das 5'-Ende eines Transkriptes, das von einem einzelnen LTR gesteuert wird, keine Sequenz des U3-Bereichs enthalten.

Es ist daher sehr wahrscheinlich, dass die Sequenz #288 ein Bruchstück eines von einem ERV9 LTR gesteuerten Transkriptes ist.

	> chr3:149886511-149890710		
	taatttggca ggtaggggct tggaaagtag ggagtggtga ttggccaagt	149886560	
U3	tggagatggt gagaggtgac agcgtgctgg cagtctcgc agccctcact	149886610	U3 erste Base
	cgctcttggc gctcctctg cctgggttcc cactttggca gcacttgagg	149886660	
	agccattcag cccgcccgtg cacttggtga gccccttct gggctggcca	149886710	
	aggccagagc cggctcctc agcttgagg gagtggtga gggagaggcg	149886760	
	cgggtgggaa cctggctgt gctgggtgt tgtgggccag cacgagttc	149886810	
	cggtgggtgt ggctcgggtg ggcccacac tgggagcagc cggcaggccc	149886860	
	caccgggtccc gggcagtgag gggctttaac acctgggcca gcagctgctg	149886910	
	cgctcaattt ctcaacgggc cttagctgcc ttccctcggg gcagggtctct	149886960	
	ggactgcag cccgcatgc ctgagctccc cccaaccccc tgcgcccgt	149887010	
	gggtcctgt gggcccgag cctcccgag agcgccacc cctgctccac	149887060	
	AGCACCCAGT CCGATCGACC ACCCAAGGGC TGAGGAGTGT GGGCGCAGGG	149887110	NT_005616.25
	CACAGGACTG GCAGGCAGCT CCACCTGCGG CCCTGGTGC GGATCCACCG	149887160	
	GTTGAAGCCA GCTGGGCTCC TGAGTCTGgt ggggacttgg agaacttta	149887210	
	tgtctagcta ggggattgta aatacaccaa tcagcactct gtatctagct	149887260	Element E
	caaggtttgt aaacacacca atcagcacc tgtgtctaac tcagGGTTG	149887310	
TGAATGCACC AATCGACACT CTGTATCTAG CTACTCTGGT GGGACTTGG	149887360	NT2>	
AGAACCTTTA TGTCTAGCTC AGGGATTGTA AACTCACCAA TCAGCGCCCT	149887410		
P	GTCAAAACAG ACCACTCGGC TCTCTGTAAA ATGGACCAAT CAGCAGGATG	149887460	
	TGGgtggggc cagataagaa aataaaagca ggctgcctga gccagcagtg	149887510	SP1 Binding tata box
R	gcaacctgct tgggtccctt tccacaccgc ggaagctttg ttcttttget	149887560	[R] (+1, +4)
	ctttgcaata aatcttctgt ctgctcactc tttgggtcca cactgccttt	149887610	Hind III
U5	atgagctgta acaccgcgaa ggtccacagC TTCACTCTG AGCCAGCGAG	149887660	
	ACCACGAACC CCACCAGAAG GAAGAATCTG TGAACACATC CGAACATCAG	149887710	NT4>
	AAGGAACAAA CTCTAGACAT GCCGCCTTTA AGAACTGCAA CACTCACTGC	149887760	Element B
	GAGTGTCGC GGCTTCATC TTGAAGTCAG TGAGACGAAG AACCCACCAA	149887810	<NT5
	TCCAGACAC ATTTTGGCGA CCTAGATGGG ACAATCGCCT ATCACCGAGT	149887860	Sequenz #288
Ggtgagacca tcgctatag ctgagcggcg agacaatcac ctatgcgcaa	149887910		
gcagtgagta ccacgggacc ccttccgctt gctattctgt cctgtctttg	149887960		
..... tgaggagacg cctctattcc gtgtcgttta gAATGAGCAA	149890360	Spleiß-Variante	
ACAGTTGTCT TCACCATCCG AGACAAGCAG ATGAATGGTG AAACGATACA	149890410		
GCAGAGAAG AGAGAAGAGG AACATTGAA CGCCGAGAGG AGTTCAGCCG	149890460		
AGGGTGGTTCG GAGAGGTTTG GCCACCCAGC AGGAAGATTA TCTTCCCGTG	149890510	<288#13	
CCATCCCOCG TTCCAGCTCT GCATCCATCC TTCAGAGAAC CACCTCCACC	149890560	3'Ende	
ACTCAATAAA ACCCTGCATT CAACCTTCAA GCCCATgtgt ggccctgattc	149890610	PolyA Signal	
ttccaggatg ctgggcaaga gcttgggata caaaaagtgg tcacattggc	149890660		

Abbildung 39. Die ERV9 LTR Sequenz: Struktur und exprimierte mRNA.

Der U3-Bereich besteht aus einer einzigen Sequenz von 641 bp (deren erstes T rot markiert ist) und aus einem repetitiven Bereich von fünf E-Elementen, die mit punktierten Pfeilen unterstrichen sind. Der Promotor beginnt direkt nach dem letzten E-Element und ist 97 bp lang. Er trägt unter anderem eine SP1-Bindungsstelle (kursive Schrift) und eine TATA-like Sequenz AATAAA (fette unterstrichene Schrift), die sich 23 Basen vor dem Initiationspunkt der Transkription befindet. Dieser Initiationspunkt definiert den Anfang der R-Region und enthält eine konservierte Hind III Schnittstelle (doppelt unterstrichen) und ein potentielles Polyadenylierungssignal (fett unterstrichen). Der U5-Bereich enthält zwei repetitive B-Elemente von 77 bzw. 75 Basen (unterstrichene Sequenzen). Orange, braun und rot markiert sind potentielle Bindungsstellen für Transkriptionsfaktoren. Die blau markierte Sequenz steht für die im Laufe dieser Arbeit isolierte Sequenz (die alternativ gespleißte Sequenz ist punktiert unterstrichen). Das potentielle 3'-Ende der mRNA mit dem Polyadenylierungssignal ist grün markiert. Der hellgraue Kasten zeigt einige Primersequenzen (>: Strangabwärts, <: Strängaufwärts). Die vorhergesagte cDNA NT_005616.25 ist mit Großbuchstaben gekennzeichnet.

3.4 Zusammenfassende Übersicht der Ergebnisse

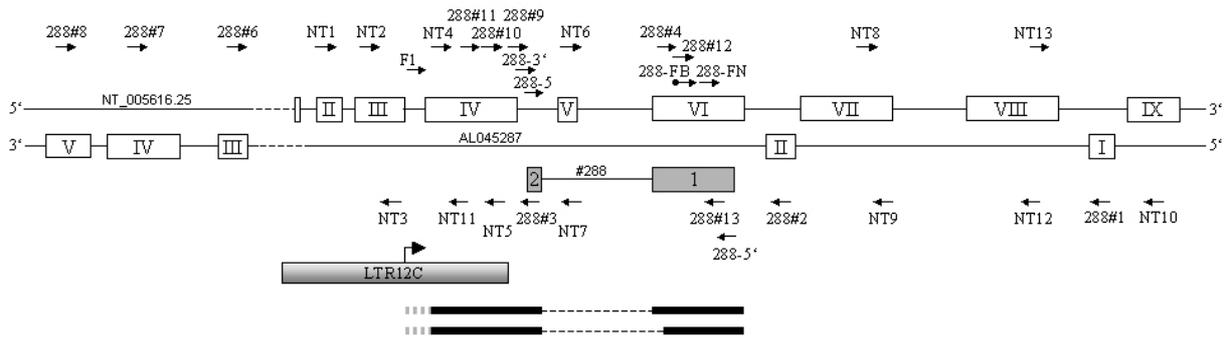


Abbildung 40. Übersicht über die verwendeten Primer und Sequenzen.

Die weißen Rechtecke zeigen die Exons der vorhergesagten cDNA NT_005616.25 auf dem positiven Strang und der cDNA AL045287 auf dem negativen Strang des Chromosoms 3 und ihre relative Position zueinander. Ebenso wird die Position der zwei Exons des Transkripts #288 (graue Rechtecke) und des LTR12C (grau-schattiertes Rechteck) dargestellt (Winkelförmiger Pfeil: Transkriptionsinitiationspunkt). Die schwarzen Pfeile geben die Position und die Orientierung der Primer an, die in dieser Arbeit benutzt wurden. Die schwarzen Linien unten stellen die zwei kompletten Spleißvarianten dar.

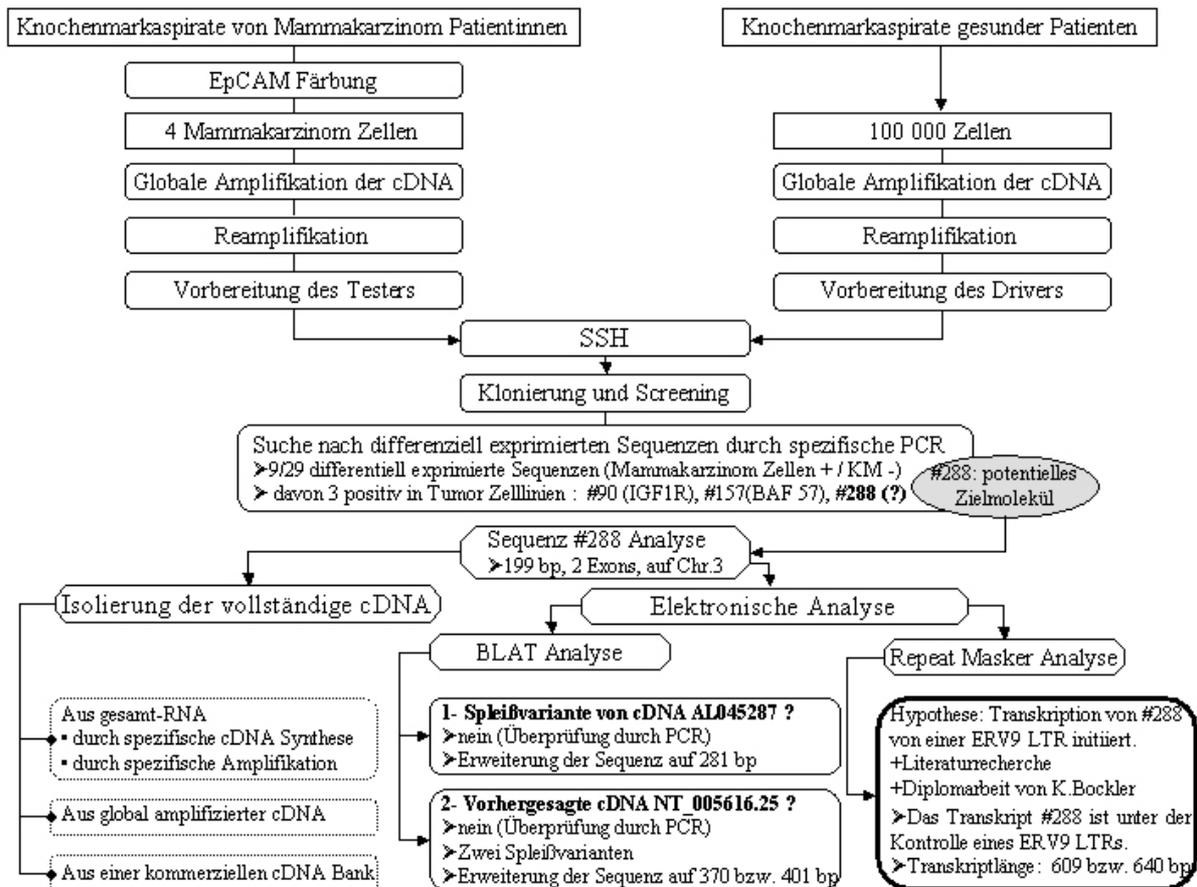


Abbildung 41. Übersicht der Vorgehensweise und Ergebnisse der Arbeit.

Das verwendete Material ist in Rechtecken wiedergegeben, die verwendete Methoden in abgerundeten Rechtecken (gepunktete Linien zeigen Methoden, die keine Ergebnisse gebracht haben). Die erhaltenen Ergebnisse sind mit Pfeilen (>) angeführt.

4 Diskussion

Bei der Krebserkrankung hängen die Heilungschancen entscheidend von einer frühzeitigen Diagnose ab. Aus diesem Grund ist es von hoher Relevanz, dass die Diagnose schnell und sicher erfolgt. Wenn eine Krebsgeschwulst entdeckt wird, wird von manchen Forschungsgruppen seit einiger Zeit das Knochenmark nach einzelnen, bereits disseminierten Tumorzellen untersucht, um eine Abschätzung der systemischen Ausbreitung der Krebserkrankung zu erhalten. Diese Untersuchung beruht auf Methoden, die Zellen unterschiedlichen Ursprungs (z.B. hämatopoietischen Zellen versus epithelialen Zellen) differenzieren können. Die meisten Studien benutzen immunzytochemische Methoden. Um Tumore epithelialen Ursprungs in mesenchymalen Organen wie Knochenmark nachzuweisen, haben sich Antigene wie zum Beispiel Zytokeratine (CK) bewährt. Der Nachweis dieser Zytokeratin positiven Zellen im Knochenmark korreliert mit einer geringeren Überlebensrate der Patienten (Braun *et al.*, 2000).

Abgesehen von den klinischen Aspekten ist der Nachweis und die Untersuchung der disseminierten Tumorzellen auch für die Beantwortung biologischer Fragen wichtig. Welcher Mechanismus verleiht den Zellen die Möglichkeit, in einer fremden Umgebung zu überleben? Werden spezielle Gene exprimiert? Um solche Fragen zu beantworten, kann die mRNA, die die Genaktivität widerspiegelt, für weitere Analysen isoliert werden. Um ein möglichst unverfälschtes Bild der Genexpression zu erhalten, muss die Zelle bei der Isolierung unbeschädigt sein, was die üblichen routinemäßigen intrazellulären Färbungen mit Zytokeratin-Antikörpern ausschließt. Um also eine möglichst gute mRNA-Qualität zu erhalten, ist eine Oberflächenfärbung der Zelle nötig.

Unseres Wissens gibt es zur Zeit keine bekannten Oberflächenmoleküle, die eine sichere differentielle Expression zwischen epithelialen und hematopoetischen Zellen zeigen. Die Identifizierung differentiell exprimierter Moleküle ist von großer Bedeutung, denn (i) überexprimierte Gene können als Marker für einige biologische Prozesse wie Invasion oder Metastasierung dienen, (ii) solche Moleküle können als Ziel für eine antikörperbasierte Immuntherapie (Scott & Welt, 1997) benutzt werden, (iii) das Verständnis der phenotypischen Änderungen, die mit der Überexpression einiger Gene assoziiert sind, ermöglicht, Signalwege besser zu verstehen und neue therapeutische Ziele zu definieren.

Das Ziel dieser Arbeit war es, durch eine Genexpressionsanalyse eines oder mehrere differentiell exprimierte Moleküle zu finden.

4.1 **Auswahl der Methode und des Materials**

Heute wird für eine Genexpressionsanalyse häufig die Chip-Technologie angewandt. Die Chip/Microarray-Technologie (wie auch die SAGE Methode) ist adäquat, um globale Expressionsprofile zu generieren. Microarray-Profile tragen sogar prognostische Informationen, wie Van de Vijver *et al.* durch eine Brustkrebs-"Signature" zeigen konnte (van de Vijver *et al.*, 2002). Einer der großen Vorteile der Methode ist, dass sie die Expression vieler Gene schnell und gleichzeitig wiedergibt. Die Analyse ist allerdings auf die bereits verfügbaren klonierten Gene begrenzt. Diese Beschränkung auf bereits bekannte Sequenzen macht die Array-Chip Methode für das Ziel dieser Arbeit ungeeignet. Außerdem begann diese Arbeit, bevor ein umfangreicher Chip zur Verfügung stand. Im Laufe des *Human-Genome-Project* ist es klar geworden, dass noch viele Gene zu entdecken sind. Die systematische Sequenzierung ist ein erster Schritt in diese Richtung. Mit bioinformatischen Programmen kann man zwar weitere Gene vorhersagen, nichtsdestoweniger können diese die direkte Isolierung aus der Zelle nicht ersetzen.

Das Vorkommen von bisher noch unbekanntem differentiell exprimierten Genen mit hoher Transkriptionsrate ist relativ unwahrscheinlich. Dagegen ist es aber anzunehmen, dass differentielle Transkripte in niedriger Kopienzahl im Transkriptom vorliegen. Daher kamen Methoden nicht in Frage, bei denen stark redundante Transkripte nicht eliminiert werden können (SAGE, Random Sampling...). Vorgezogen wurden Methoden, die nur eine Subtraktionsrunde benötigen, was einen mühsamen Eins-zu-Eins-Vergleich von zwei Populationen vermeidet und die Isolierung differentiell exprimierter Moleküle fördert. Von den zwei Methoden RDA (*Representational Difference Analyse*) und SSH (*Suppression Subtractive Hybridization*), die in Frage kommen, weil sie die stark exprimierten Gene durch eine Subtraktion aus dem Probenansatz entfernen können, wurde die SSH gewählt. Diese Methode hat den Vorteil, dass sie mittels eines Normalisierungsschrittes die Redundanz von Transkripten aufhebt, ohne dass seltene Transkripte verloren gehen. Die SSH wird für die Etablierung von Expressionsprofilen sehr häufig benutzt (Difilippantonio *et al.*, 2003; Hufton *et al.*, 1999; Li *et al.*, 2003). Obwohl sowohl die RDA als auch die SSH wenig Ausgangsmaterial benötigen, sind für die RDA meistens mehrere Subtraktionsrunden nötig. Dies wirkt sich negativ auf die Sensitivität der Methode aus, was für die vorliegende Arbeit, die extrem wenige Zellen verwendet, kritisch ist. Mit ihrem einzigen Subtraktionsschritt ist die SSH hierfür besser geeignet.

In der vorliegenden Arbeit wurde mit Zellen von Patienten und nicht von Zelllinien gearbeitet. Wie Zhang *et al.* (1997) zeigen konnten, sind viele Gene, die im

Kolorektalkarzinom differentiell exprimiert sind, in Zelllinien von Kolorektalkarzinomen ebenfalls differentiell exprimiert, aber durchaus nicht alle. Die Autoren weisen darauf hin, dass Zelllinien keine perfekte Abbildung eines Tumors in seiner natürlichen Umgebung darstellen. Sicherlich exprimiert eine metastatische Tumorzelle im KM nicht genau die gleichen Gene, wie unter Kulturbedingungen, und möglicherweise sind diese Gene besonders wichtig. Deswegen wurden 4 Einzelzellen von 4 Patienten ausgewählt. Durch die Vorarbeit mit Nylonarrays, auf die verschiedene Tumormarker aufgespottet sind, konnte gezeigt werden, dass es sich mit Gewißheit um Tumorzellen handelt.

Die geringe Anzahl der Zellen ergibt sich in erster Linie daraus, dass die disseminierten Einzelzellen selten zu finden sind. Die Frequenz dieser Zellen liegt bei 10^{-5} - 10^{-6} bei jedem dritten Krebspatienten (Pantel *et al.*, 1999). Außerdem kommen noch drei Schwierigkeiten hinzu:

- es muss gewährleistet sein, dass Tumorzellen isoliert werden, nicht jedoch normale Zellen. Normale Zellen können auch das Oberflächenprotein EpCAM tragen.
- die Zellen müssen intakt sein und ihre RNA noch enthalten.
- Zuletzt muss die Probenpräparation gelingen.

Da es schwierig ist, allen diesen Voraussetzungen mit Sicherheit jedesmal zu entsprechen, ist die Anzahl der zu studierenden Zellen relativ gering. Wir entschlossen uns nur mit vier Zellen zu arbeiten, für die es aber sicher bestätigt wurde, dass es sich um Tumorzellen handelt.

Die SSH Methode ist im Vergleich zu anderen Methoden mit relativ wenig Ausgangsmaterial (1-2 μ g mRNA) durchführbar. Zwei Beispiele beweisen, dass die SSH Methode noch für zehnfach verringertes Ausgangsmaterial (100 ng) geeignet ist. So konnte spezifische cDNA aus 50 Oozyten (d.h. aus 92 ng mRNA) isoliert werden (Robert *et al.*, 2000) und mit der gleichen Technologie cDNA aus kleinen Biopsien (i.e. aus 100 ng gesamte RNA) amplifiziert werden (Zanders *et al.*, 2000). Die vorhandene Menge an Ausgangsmaterial aus vier Zellen liegt aber annähernd bei 20 pg, d.h. 105-fach weniger als für die originale Methode angegeben. Eine so geringen Menge an Material fordert eine vorhergehende Amplifikation. Dies wird mit einer speziellen globalen Amplifikationsmethode erreicht, dessen Anwendung eine Anpassung der SSH verlangt.

4.2 **Entwicklung einer SSH für Einzelzellen**

Gleich nach der Isolierung der mRNA jeder einzelnen Zelle wurde eine cDNA-Synthese durchgeführt und die cDNA unverzüglich global amplifiziert (Klein *et al.*, 2002). Nach der globalen cDNA-Amplifikation trägt die cDNA durch den speziellen Primer eine besondere

CP2-Sequenz an beiden Enden, die in den weiteren Schritten der SSH neue Adaptoren und Primer benötigt.

Die erste wichtige Modifikation betrifft die Sequenzen der Adaptoren. Dabei wird die Tatsache genutzt, dass die CP2-Sequenz eine EcoRI-Schnittstelle trägt, wodurch eine Ligation von Adaptoren mit kohäsiven Enden möglich wird. Eine solche Ligation ist bis zu 100 mal schneller als eine Ligation mit glatten Enden (Sugino *et al.*, 1977). Die Ligation ist einer der wichtigsten Schritte der Subtraktion. Ihr Erfolg kann zum Beispiel eine Rolle bei der Entstehung von Hintergrundsignalen spielen: Lange Oligonukleotide aus unligierten Adaptoren können während der subtraktiven Hybridisierung unspezifisch auf cDNA-Moleküle hybridisieren, die eine ähnliche Sequenz haben. Nach der DNA-Auffüllung können solche Moleküle als Matrize für sowohl die erste als auch die zweite PCR der SSH dienen (Rebrikov *et al.*, 2000). Bei dem Versuch die unligierten Adaptoren mittels einer Microcon-Säule zu entfernen, wurde zu viel Material verloren. Da dieser Reinigungsschritt nicht zwingend notwendig für das Gelingen der SSH ist, wurde er weglassen.

Die zweite wichtige Modifikation betrifft die Primer, die die Suppressive PCR ermöglichen. Die Benutzung der CP2 Primer (5'-TCA GAA TTC ATG (CCC)₅-3') während der cDNA-Synthese fügt (Cp)₁₅ und (Gp)₁₅ Sequenzen im 5'- bzw. 3'-Bereich der Fragmente ein. Dadurch wird die Bildung von Haarnadelstrukturen gefördert, die die Anlagerung von Primern behindert. Um thermodynamisch die Bindung der Primer zu unterstützen, wurden zum einen Primer von 30 Basen anstatt 22 Basen in der PCR1 benutzt. Zum anderen wurde der Driver mit einem CPBgl-Primer (5'-TCA GAA TTC ATG CCG CCC CCC CGG CCC-3') amplifiziert. Das CPBgl Oligonukleotid beinhaltet eine Restriktionsschnittstelle für Bgl I (GCCNNNN[▼]NGGC) sechs Basen vor seinem 3'-Ende. Nach der Amplifikation erlaubt der Verdau mit Bgl I 21 Basen des Primers abzuschneiden, so dass der Großteil der (Cp)₁₅-Sequenz entfernt werden kann.

4.3 SSH von Tumorzellen und Knochenmark

4.3.1 Anzahl der humanen Sequenzen

Das theoretische Modell der SSH besagt, dass der kritische Faktor für die Effizienz der SSH das Konzentrationsverhältnis (R) von einem Gen zwischen zwei cDNA Proben ist. Bevorzugt angereichert werden "all or nothing" differentiell exprimierte Gene ($R \rightarrow \infty$) oder Gene mit einem hohen R-Wert. Deshalb muss ein Gen, um angereichert zu werden, mindestens eine Konzentration von 0,1 % im Gesamt-mRNA Material haben, und R muss bei mindestens 5 liegen. Niedrig exprimierte Gene, wie zum Beispiel Gene für Transkriptionsfaktoren,

Zytokine und Rezeptoren, die bei vielen Erkrankungen eine Schlüsselrolle spielen, werden daher mit der SSH nicht detektiert. Außerdem sind differentiell exprimierte Gene nur R^3 -mal häufiger in einer SSH Library als in einer unsubtrahierten Library angereichert. Folglich sollte es nicht überraschen, wenn eine kleine SSH Library kein einziges differentiell exprimiertes Gen enthält (Ji *et al.*, 2002).

Trotz dieser strengen Einschränkungen der SSH Methode zeigten 9 von den 29 klonierten menschlichen Sequenzen eine differentielle Expression im Tester (Zellen von Mammakarzinompatienten) und nicht im Driver (KM Zellen) (**Abbildung 42**).

weiße Klone (386)	
Sequenzierung (110)	Leer oder sehr kleine Inserts (276) ►keine Sequenzierung
35	keine Ergebnisse
13	Mischsequenzen
33	bakterielle ribosomale Sequenzen
29	humane Sequenzen
9	differentiell exprimierte Sequenzen im Mammakarzinomzellen (Tester) und nicht in KM Zellen (Driver)

Abbildung 42. Gesamtüberblick über die klonierten Sequenzen.

Ein sehr grosse Anzahl der Sequenzen war nicht zu analysieren. Von den 110 sequenzierten Inserts waren nur 29 humanen Ursprungs, neun davon differentiell exprimiert.

4.3.2 Bakterielle Sequenzen

Eines der größten Probleme bei dieser Arbeit ist die Herkunft der isolierten Sequenzen. Über die Hälfte (33/62) sind Sequenzen von bakteriellen 16S-Ribosomen. Die Kontamination findet während der globalen Amplifikation statt, wie es der Schmier bei der Negativkontrolle zeigt. Die beste Erklärung hierfür ist eine Kontamination durch bakterielle DNA, die in den verwendeten Enzymen vorhanden ist. Während der Herstellung von Enzymen werden Nukleinsäuren, einschließlich der ribosomalen DNA-Sequenzen, entfernt. Es wurde trotzdem mehrmals gezeigt, dass das Enzym *Taq*-Polymerase als Folge des Herstellungsverfahrens eine der wichtigsten Quellen der kontaminierenden DNA ist (Bottger, 1990; Corless *et al.*, 2000; Hughes *et al.*, 1994). Es handelt sich dabei um DNA von verschiedenen bakteriellen Spezies, aber nicht von *E. coli* (oder bei der *Taq* Polymerase von *T. aquaticus*), und man findet viele Sequenzen des Ribosoms 16S. Es gibt im Genom von Eubakterien viele Kopien dieses Gens. Sowohl die *Taq*-Polymerase als auch die Reverse Transkriptase und die Terminale Desoxynucleotidyl-Transferase sind oft als rekombinante Proteine in *E. coli* exprimiert und alle Enzympräparationen können kontaminierende DNA-Moleküle enthalten. Fast alle hier isolierten Sequenzen tragen das Cfl5c8 Oligonukleotid, was ein Hinweis darauf ist, dass die bakterielle Kontamination zwischen der Isolierung der Zellen und der cDNA-Synthese mit

den Cfl5c-8 und -T-Primern auftritt und wahrscheinlich aus der Präparation der Reversen Transkriptase stammen. Die Reverse Transkriptase kann einzelsträngige DNA als Matrize benutzen, wenn ihr ein Primer zur Verfügung steht. Da diese Kontaminationen schon im Ausgangsmaterial (sowohl im Tester als auch im Driver) vorhanden sind, könnte man annehmen, dass sie subtrahiert werden. Die ersten Versuche haben leider gezeigt, dass dies nicht der Fall ist.

Der Ursprung des Problems liegt wahrscheinlich an der Anzahl von verarbeiteten Zellen bei der globalen Amplifikation der cDNA. Unabhängig von der Anzahl isolierter Zellen wird die gleiche Menge an Reverse Transkriptase benutzt. Angenommen, dass eine Zelle 3-6 pg mRNA enthält und dass x pg kontaminierende Sequenzen mit dem Enzym in die Präparation kommen, dann ist das Verhältnis von Kontamination zu zellulärem Material bei der Isolierung einer Einzelzelle x pg / 3 pg. Das Verhältnis bei der Isolierung 100 000 KM Zellen ist dagegen x pg / 300 000 pg $\rightarrow 0$ und wird damit verschwindend gering.

Deswegen wurde der Driver mit kontaminierenden bakteriellen Sequenzen angereichert, in der Hoffnung diese Sequenzen vom Tester besser zu subtrahieren. Dieser Versuch blieb erfolglos, obwohl eine ähnliche Anreicherung bei einem anderen Subtraktionsansatz erfolgreich getestet wurde (Hubank & Schatz, 1994). Eine mögliche Ursache dafür ist, dass die endogene Kontamination nicht ganz identisch mit den zugegebenen Sequenzen ist, so dass keine Subtraktion möglich ist. Eine andere Möglichkeit ist, dass die kontaminierenden Sequenzen im Tester so häufig vorkommen, dass einige bakterielle Sequenzen nach der Subtraktion weiter vorhanden bleiben. Die PCR kann eine einzelne Kopie einer Vorlage bis zu 10^6 - 10^7 -mal amplifizieren (Saiki *et al.*, 1988). Eine geringe Kontamination kann nach zwei PCR-Amplifikationen, wie es bei der SSH der Fall ist, ein schwerwiegendes Problem werden. Solange keine ultrareinigten Enzyme entwickelt werden, wird die bakterielle Kontamination für viele Methoden eines der Hauptprobleme bleiben.

In der vorliegenden Untersuchung allerdings ist die starke Anwesenheit der bakteriellen Sequenzen ein guter Hinweis auf den Erfolg der Methode, da die bakteriellen Amplifikate hier als differentiell exprimierte Sequenzen betrachtet werden können.

4.3.3 Identität und Expression der Sequenzen

Drei der isolierten Sequenzen stammen von Genen, die möglicherweise eine Rolle bei der Tumorprogression haben:

- BAF57, eine Untereinheit des menschlichen SWI/SNF Komplexes, der bei der Steuerung der Genexpression beteiligt ist. BAF57 ist in vielen Geweben exprimiert, aber die Expression im KM wurde nicht nachgeprüft (Wang *et al.*, 1998). BAF57 ist

in vielen Mammakarzinom-Zelllinien exprimiert (Decristofaro *et al.*, 2001). Wahrscheinlich spielt der SWI/SNF Komplex eine Rolle bei der Tumorchemmung (Klochender-Yeivin *et al.*, 2002).

- IGF1R (Insulin-Like Growth Factor I Receptor), spielt eine kritische Rolle bei der malignen Transformation. Er wird in den meisten bösartigen Geweben stark exprimiert, in denen er anti-apoptotisch wirkt und folglich das Überleben der Zelle fördert. Die Inaktivierung von IGF1R kann von mutiertem p53 (Werner *et al.*, 1996) oder mutiertem BRCA1 (Maor *et al.*, 2000) aufgehoben werden, was eine wichtige Rolle bei der Tumorgenese haben könnte. Interessanterweise hat eine Arbeitsgruppe (All-Ericsson *et al.*, 2002) die Expression von IGF1R untersucht und hat sich auf dessen Rolle beim Zellwachstum in Aderhautmelanomen konzentriert. Ihre Analyse zeigt eine deutliche Verknüpfung zwischen einer hohen IGF1R Expression und der Sterblichkeit an einer Streuung des Tumors.
- LPP (Lipoma Preferred Partner), dessen cDNA bei der Translokation in gutartigen Lipomen (Schoenmakers *et al.*, 1995) oder einer Akuten Myeloischen Leukämie (Daheron *et al.*, 2001) beteiligt ist.

Da es das Ziel dieser Arbeit ist einen Marker zu finden, der einzelne disseminierte Tumorzellen im KM besser detektierbar macht, wurde die Expression der im Tester differentiell exprimierten Sequenzen zwischen elf Tumorzelllinien und KM geprüft. Drei Sequenzen können nur in Zelllinien und nicht im KM nachgewiesen werden. Es handelt sich um den IGF1-Rezeptor (Sequenz #90), um die BAF57 Untereinheit (#157) und um ein Fragment des Chromosoms 3 (#288), dem kein bekanntes Gen entspricht (**Abbildung 21**, Seite 70).

Die differentielle Expression von IGF1R in Zelllinien stimmt mit den bis jetzt veröffentlichten Ergebnissen überein (s.o.). BAF57 wurde in dieser Arbeit nur in den zwei Zelllinien HT29 und H69 (Colon- bzw. Lungenkarzinom) gefunden, aber nicht in den drei getesteten Mammakarzinom-Zelllinien MCF7, Skbr3 und BT20. Dieses Ergebnis stimmt mit den Ergebnissen von Decristofaro nicht überein (Decristofaro *et al.*, 2001). Alle Zelllinien, die diese Arbeitsgruppe für die Expression von BAF57 getestet hat (21 Mammakarzinome ZL und 19 andere Karzinome), waren positiv für BAF57, mit Ausnahme von BT549, welche die Autoren als erste BAF57-negative Zelllinie bezeichnen.

4.4 Die Sequenz #288

4.4.1 Analyse des chromosomalen Bereichs 3q25.1

Die Analyse der *expressed sequence tags* (ESTs) kann darauf hinweisen, dass der Bereich 3q25.1 eine Rolle beim Prozess der Tumorentstehung spielen könnte. Tatsächlich wurden fast alle ungespleißten humanen ESTs, die die Sequenz #288 flankieren oder sogar ein Stück der Sequenz enthalten, aus Mammakarzinom-Gewebe isoliert (Dias Neto *et al.*, 2000). Zwei ESTs wurden aus der humanen Fibrosarkom-Zelllinie HT1080 bzw. aus einem "Large cell carcinoma" isoliert (Harrington *et al.*, 2001).

Außerdem stimmt die Sequenz #288 mit zwei Exons der vorhergesagten cDNA NT_005616.25 überein, die ein offenes Leseraster (ORF) für ein interessantes Protein-Sequenzstück enthält. Das vorhergesagte Protein zeigt (N-Terminal) hohe Ähnlichkeit zu einem "unbekannten" Protein AAL55829 (Genbank), das mit "*Novel human cDNA clones with function of inhibiting cancer cell growth*" annotiert ist.

Interessanterweise zeigt ein ORF der kleinen Spleißvariante der Sequenz #288 eine Homologie zu einer sauren Transaktivierungsdomäne, die für die Transkriptionsaktivierung eine Rolle spielen könnte. Chimäre Transkripte, die ebenfalls so eine Domäne tragen, wurden schon aus Lipomgeweben isoliert ((Ashar *et al.*, 1995) (3.3.5.4).

4.4.2 Die ERV9-LTR-Hypothese

Eine Recherche mit der Sequenz #288 im UCSC Genome Browser zeigt dass die Sequenz #288 auf dem Chromosome 3 im Bereich 3q25.1 lokalisiert ist. Sie ist auf zwei Exons aufgeteilt. Das Exon 1 befindet sich in einer langen terminalen Sequenzwiederholung (*Long Terminal Repeat*, LTR) des humanen endogenen Retrovirus ERV9. Das Exons 2 hat eine 100 %ige Homologie mit einem pTR5 Repetitiven Element (La Mantia *et al.*, 1989), welches ein Transkript des humanen ERV9 ist (Lania *et al.*, 1992). Eine detaillierte Analyse dieses Bereiches zeigt, dass ein kompletter ERV9 LTR von 1244 bp Länge mit dem typischen U3-R-U5 Aufbau auf dem positiven Strang liegt. Dieses Element ist in der Repbase Datenbank unter LTR12C registriert und gehört zu den solitären ERV9 LTRs (*solitary LTR*), die bis zu 4000 mal im Genom verstreut vorkommen. Das 5'-Ende des isolierten mRNA-Fragments #288 befindet sich in der Mitte des U5-Segments, das 3'-Ende circa 50 bp vor ein Polyadenylierungssignal. So lange der 288-13 Primer (3') mit einem 5'-Primer, der stromabwärts des Transkriptionsinitiationspunktes liegt, kombiniert wird (wie 288-13 mit NT4 in dieser Arbeit, oder 288-13 mit F1 in der Diplomarbeit von Kerstin Bockler), können die erwarteten Produkte amplifiziert werden. Sobald der zweite Primer stromaufwärts des Transkriptionsinitiationspunktes liegt, ist keine Amplifizierung möglich. Dieses Ergebnis ist

ein entscheidender Hinweis, dass es sich hier sehr wahrscheinlich um ein Transkript handelt, das von einem ERV9 LTR gesteuert ist.

4.4.3 Die Promotoraktivität des ERV9 LTR

Dass ein neues von einem ERV9 LTR gesteuertes Transkript entdeckt wird, überrascht nicht, wenn man bedenkt, dass die ERV9 Familie aus ungefähr 50 integrierten Proviren und mindestens 4000 einzelnen LTRs besteht. Diese sind über verschiedene genomische Loci verteilt und können als potentielle Promotoren und Enhancer dienen.

Es wurden bereits verschiedene zelluläre mRNA-Transkripte, die von ERV9 LTRs gesteuert werden, beschrieben. Die ersten Transkripte wurden aus einer humanen Teratokarzinom-Zelllinie (NTERA-2/D1) isoliert (La Mantia *et al.*, 1989; Lania *et al.*, 1992). Neuere Northern Blot-Analysen zeigen, dass eine geringe Kopienzahl von ERV9 LTRs, die zu mRNA von unterschiedlicher Größe führen, auch in vielen normalen Geweben exprimiert wird. Die Expression ist meistens gering, außer im Testis- und Nebennierengewebe (Svensson *et al.*, 2001). Andere von ERV9 LTRs gesteuerte mRNAs wurden auch aus der T-Lymphom-Zelllinie Peer isoliert und charakterisiert (Strazzullo *et al.*, 1998). Die ERV9 mRNA-Transkripte zeigen verschiedene Längen, meistens zwischen 800 bp und 1500 bp lang. Es wurde auch gezeigt, dass ERV9 LTRs die Expression eines Reportergens im Verlaufe einer transienten Transfektion steuern können (La Mantia *et al.*, 1992). Das menschliche ZNF80-Gen, das für ein Zink-Finger-Protein kodiert, ist von einem ERV9 LTR, der strangaufwärts des Promotorbereichs liegt, reguliert (Di Cristofano *et al.*, 1995).

In dieser Arbeit finden wir eine differentielle Expression der Sequenz #288 zwischen Tumorzelllinien und Knochenmark. Dieser Befund deutet darauf hin, dass ERV9 LTRs die Transkription in spezifischen Geweben beeinflussen können. Auf diese Weise ist zum Beispiel die Transkription des humanen β -Globin-Gens in erythroiden Zellen reguliert (Long *et al.*, 1998). In H4IIE-C3 Hepatom-Zellen wird die Expression von der humanen Alkoholdehydrogenase 1C (ADH1C) durch einen ERV9 LTR gesteuert und ist dort sechsmal höher als in HeLa Zellen (Chen *et al.*, 2002).

Es wird diskutiert, ob die Gewebespezifität durch das U3-Segment verursacht wird, da dieser Bereich besonders in den *Tandem Repeats* viele potentielle Transkriptionsfaktor-Bindungsmotive trägt. Die GATA, CCAAT (5) und CCACC (6) Bindungsmotive, die in dieser Region zu finden sind, würden die Enhancer-Aktivität der ERV9 LTRs in embryonalen und hämatopoietischen Zellen fördern (Ling *et al.*, 2002; Tenen *et al.*, 1997).

Eine jüngst veröffentlichte Arbeit von H.J. Edenberg über die humane Alkoholdehydrogenase 1C (ADH1C) zeigt, dass die von dem U3-Segment getragenen CCAAT Bindungsmotive die Gewebespezifität der ADH1C steuern können (Chen *et al.*, 2002). Das vom NY-FC Faktor erkannte CCAAT Bindungsmotiv ist unveränderlich von anderen Transkriptionsfaktor-Bindungsmotiven flankiert. Der NY-F Faktor könnte die Bindungsaffinität der benachbarten Transkriptionsfaktoren erhöhen. Wenn diese CCAAT Motive für die Expression des Transkripts #288 wichtig sind, ist wahrscheinlich ein Synergie-Effekt des NY-F Faktors mit anderen Faktoren auch hier an der Gewebespezifität beteiligt, da wir keine Expression der Sequenz #288 in hämatopoietischen Zellen (d.h. in Knochenmark) feststellen konnten (3.2.2).

Die TATA-Box und Sp1 Bindungsmotive steuern die optimale Transkription des ERV9 LTR (La Mantia *et al.*, 1992; Strazzullo *et al.*, 1994), aber es ist nicht ausgeschlossen, dass das Sp1 Bindungsmotiv auch bei der Gewebespezifität eine Rolle spielt. Häufig sind "die" Sp1 Bindungsmotive (*Sp1 sites*, d.h. die GC-reiche und CACCC Sequenzen) an der Genexpression in Geweben beteiligt. Übereinstimmend mit der Beobachtung, dass Mitglieder der Sp1/KLF Familie an der Wachstumskontrolle beteiligt sind, können einige ihrer Mitglieder auch auf Onkogene und Tumorsuppressoren (z.B. p53) wirken (Black *et al.*, 2001). Nach der Arbeit von La Mantia kann man vermuten, dass das Sp1 Bindungsmotiv eine Rolle bei der Expression des Transkripts #288 spielt (La Mantia *et al.*, 1992). Dieses Sp1 Bindungsmotiv befindet sich in der minimalen Promotorsequenz und ist für die maximale Aktivität zuständig, was auch mit seiner Position als promotornahes Element übereinstimmt. Die Transkription anderer HERV-LTRs ist auch von Sp1 Bindungsmotiven abhängig, z.B. der HERV-H LTR (Sjottem *et al.*, 1996) oder ein HERV-E LTR (Landry & Mager, 2003). Allerdings ist das Sp1-Bindungsmotiv nicht allein an der Aktivierung der Transkription beteiligt und seine genaue Rolle ist noch nicht geklärt. So könnte im HERV-W LTR die Anwesenheit des Sp1-Bindungsmotivs mit der Unterdrückung der Transkription korrelieren (Schon *et al.*, 2001), was jedoch von einer neuen Arbeit bestritten wird (Lee *et al.*, 2003). Auch andere Bindungsmotive spielen eine wichtige Rolle und können die Gewebespezifität beeinflussen (Landry & Mager, 2003).

4.4.4 Schlußfolgerungen

Lange hat man geglaubt, dass DNA ein inertes und unmodifiziertes Molekül ist, mit Ausnahme des Auftretens von Punktmutationen (10^{-9} Aminosäure/Jahr). Während der letzten 15 Jahre ist jedoch klar geworden, dass viele andere Mechanismen an genomischer Instabilität

beteiligt sind. Insbesondere die beweglichen DNA-Elemente wie Transposons und Retrotransposons könnten eine wichtige Rolle spielen.

Transposons können in Form von DNA direkt an andere Stellen im Genom springen, Retrotransposons müssen zuerst in eine RNA-Kopie transkribiert werden, und werden dann in DNA umgeschrieben. Ungefähr 42% des menschlichen Genoms bestehen aus Retrotransposons (Lander *et al.*, 2001), und obwohl die meisten inaktiv sind, können einige retrotransposieren. Retrotransposons lassen sich in nicht-autonome (Alu) und autonome Elemente teilen, je nachdem ob sie Enzyme der Zelle für ihre Beweglichkeit brauchen oder nicht. Die autonomen Elemente lassen sich ebenfalls in zwei Gruppen teilen: LTR-Retrotransposons (*Long Terminal Repeat*, besitzen lange terminale Sequenzwiederholungen), und nicht-LTR Retrotransposons. Die LTR-Sequenzen sind das Kennzeichen für den viralen Ursprung dieser Elemente. LTR-Retrotransposons haben folglich eine ähnliche Struktur wie Retroviren, aber sie tragen kein funktionelles *env* Gen, so dass sie andere Zellen nicht infizieren können. Zu dieser Gruppe gehören auch die humanen endogenen Retroviren (HERV), obwohl es unwahrscheinlich ist, dass sie autonome aktive Elemente besitzen. Bis zu 8% des menschlichen Genoms besteht aus defekten ERVs und *solitären* LTRs (Lander *et al.*, 2001). Die nicht-LTR Retrotransposon Gruppe umfaßt die LINE und SINE (*Long* bzw. *Short interspersed nucleotide elements*) Sequenzen. Die häufigsten SINE Elemente beim Menschen sind die Alu-Sequenzen; sie machen etwa 11% der Gesamt-DNA aus. Zu den LINE Sequenzen gehören die L1 Elemente, die die häufigsten Retrotransposons im Genom der Wirbeltiere sind.

All dies macht deutlich, dass die beweglichen DNA-Elemente bei verschiedenen biologischen Prozessen in der Zelle eine Bedeutung haben können.

Die Transkription von Genen kann zum Beispiel durch einen LTR erleichtert oder sogar gesteuert werden, da diese Elemente oft eine starke Aktivität als Promotor und Enhancer haben (Di Cristofano *et al.*, 1995; Kato *et al.*, 1990; Liu & Abraham, 1991). Auch L1 Elemente ohne LTR können eine solche Rolle übernehmen, wie es für das Apolipoprotein(a) Gen der Fall ist (Yang *et al.*, 1998). Ihre regulative Funktion drückt sich auch dadurch aus, dass LTR-Sequenzen die Expression in spezifischen Geweben beeinflussen können (Chen *et al.*, 2002; Ling *et al.*, 2002), oder dass ein L1 Element an der Regulierung der Gerinnung durch den Faktor IX beteiligt ist (Kurachi *et al.*, 1999).

Viele Retrovirale Elemente können eigene proteinkodierende (für Reverse Transkriptase (RT) und Endonuklease) und auch nichtkodierende RNA synthetisieren. Nichtkodierende Regionen

werden heute noch oft als unwichtig ("Junk") betrachtet. Diese Ansicht ändert sich aber mit der Feststellung, dass viele nichtkodierende RNAs (*noncoding RNAs*, *ncRNAs*) transkribiert werden und es immer offensichtlicher wird, dass RNA regulierende Netzwerke kontrollieren kann (Mattick & Makunin, 2005). Es wurden bereits verschiedene ncRNAs beschrieben, zu denen *small nucleolar RNAs* (snoRNAs), *microRNAs* (miRNAs) und *short interfering RNAs* (siRNAs) gehören. **Small nucleolar RNAs (snoRNAs)** sind - wie der Name impliziert - meist in den Nucleoli anzutreffen, wo sie an der Prozessierung und Modifikation anderer Ribonukleinsäuren - insbesondere ribosomaler RNA (rRNAs) - beteiligt sind. Die Modifikationen, die in die ribosomalen RNAs eingeführt werden, sind essentiell für die Funktion des Ribosoms und die in den *small nuclear RNAs* (snRNAs) essentiell für das Splicing. **MicroRNA (miRNAs)** dienen zur Feinregulierung des Wachstums. Sie binden spezifisch an messenger RNA (mRNAs) und verhindern deren Aktivität in der Translation, so dass z.B. Wachstumsfaktoren nicht in übermäßiger Menge hergestellt werden. Die **small interfering RNA (siRNA)** sind Fragmente von ca. 22 Nukleotiden Länge und werden in den Enzymkomplex RISC (RNA-induced silencing complex) eingebaut. Mithilfe der inkorporierten RNA-Fragmente bindet RISC komplementär an DNA, z.B. Genbereiche, oder mRNA und kann diese damit "abschalten". Dies dient einerseits dem Abbau von Fremd-RNA, z.B. der von Viren, aber auch der zellinternen Regulation. Dieser Signalweg der Zelle wird als **RNA Interference (RNAi)** zusammengefasst.

An den ncRNAs wird aktuell intensiv geforscht, darüber hinaus kommen viele andere ncRNAs in Tieren und Pflanzen vor. Sie bestehen aus verschiedenen Transkripten, wie zum Beispiel transkribierten Introns oder antisense Transkripten und bei einigen können eine Cap-Gruppe und ein Poly-A Schwanz angehängt sein. Neue Studien zeigen, dass es wahrscheinlich mehr ncRNA, als bis heute bekannt gibt (Washietl *et al.*, 2005). Jede neue ncRNA Entdeckung unterstützt die Vermutung, dass es komplexe Kontrollnetze gibt, die unter anderem durch miRNA oder verschiedene andere, teilweise noch unbekannte, RNA Signale reguliert werden. Daher kann man nicht ausschließen, dass ncRNAs, deren Transkription zum Beispiel von LTRs gesteuert wird, auch eine Rolle in der Zellregulation haben, wie zum Beispiel bei *post transcriptional gene silencing* (Ling *et al.*, 2002; Ling *et al.*, 2003; Prak & Kazazian, 2000). So hätten miRNAs oder siRNAs, die zum Beispiel von transkribierten HERV Sequenzen stammen, tausende potentielle Zielsequenzen, da HERV Sequenzen in vielen Kopien über das ganze Genom verstreut sind. Obwohl es bis jetzt noch keine Beweise gibt, sollte man diese Möglichkeit gegenwärtig nicht ausschließen.

Die beweglichen Elemente können auch chromosomale Veränderungen, wie Translokationen, Deletionen oder Inversion, vermitteln. Sie können somit für Mutationen, die beim Menschen zu genetisch bedingten Erkrankungen führen, verantwortlich sein. L1 und Alu-Elemente sind hierbei besonders häufig beteiligt, zum Beispiel bei der Duchenne-Muskeldystrophie, Hämophilie, oder bei Brustkrebs (Kolomietz *et al.*, 2002; Ostertag & Kazazian, 2001).

Die differentielle Expression des Transkriptes #288 bei Tumorzelllinien und Patienten wirft die Frage nach der Bedeutung der Expression solcher Transkripte auf. Einerseits kann die Expression der Transkripte unspezifisch durch eine Veränderung der Transkriptionsaktivität des Tumors ausgelöst werden. Eine Arbeit über Alu-Sequenzen spricht zum Beispiel für einen solchen Mechanismus: die Expression von Alu-Sequenzen ändert sich mit der Transformation der Zelle (Chu *et al.*, 1998). Dies unterstreicht die Wichtigkeit mit Zellen verschiedenen Ursprungs (d.h. sowohl mit Zelllinien als auch mit primären Tumorzellen) zu arbeiten.

Andererseits kann das Transkript angeschaltet werden, weil es eine funktionelle Bedeutung für den Tumor besitzt. Bestimmte Funktionen der HERV für die Entwicklung der Plazenta könnten eine Bedeutung bei der Tumorprogression nahelegen. Die Plazenta zeigt biologische Eigenschaften wie Gewebeinvasion, Fusion der Trophoblastenzellen, Suppression des mütterlichen Immunabwehr, die diesem Organ auch die Bezeichnung "Pseudotumor" einbrachten (Harris, 1998).

Das pathogene Potential nicht defekter endogener Retroviren ist bis jetzt nur in Mäusen demonstriert worden, in denen ERV Tumore und immunologische Störungen verursachen können (Asch *et al.*, 1993; Natoli *et al.*, 1996). Immer mehr Arbeiten interessieren sich jedoch für die Expression in verschiedenen Geweben und für ihre mögliche Rolle bei der Krebserkrankung des Menschen (Schon *et al.*, 2001; Seifarth *et al.*, 2003). Man konnte schon oft eine höhere Transkriptionsaktivität in Tumorgeweben als in Normalgeweben feststellen. Einige Autoren sehen ERV-Sequenzen als eine vielversprechende Quelle neuer Tumormarker (Wang-Johanning *et al.*, 2003). Zwar sind bisher wenige Transkripte bekannt, die in Proteine übersetzt werden (Frendo *et al.*, 2003; Lower *et al.*, 1995; Schiavetti *et al.*, 2002). Dennoch könnten gerade diese Proteine als Tumorantigene für Immuntherapien dienen. Vor kurzem wurde eine HERV-K *env* Sequenz beschrieben, die ein Antigen kodiert, das von zytotoxischen T Lymphozyten erkannt wird und die Abstoßung von Melanomzellen bewirkt. Das Antigenpeptid wurde durch einen sehr kurzen Leserahmen kodiert. Diese Ergebnisse legen nahe, dass HERV *env* Proteine oder ihre Peptide eine Quelle für Tumorantigene sein könnten und folglich für Schutzimpfung verwendet werden könnten (Schiavetti *et al.*, 2002).

5 Zusammenfassung

Das Auftreten einer Metastasierung ist von großer prognostischer Relevanz für Patienten mit einer Tumorerkrankung und hat einen bedeutenden Einfluß auf die Heilungsrate. Bisher gibt es keine Therapie, die die Ausbreitung der Krebserkrankung verhindern oder stoppen kann, unter anderem, weil der Prozess der Metastasierung noch weitgehend unklar ist. Deswegen ist es wichtig, die molekularen Mechanismen der Tumorzell dissemination zu verstehen: Gibt es bestimmte Gene, die in disseminierten Zellen unter- oder überexprimiert sind? Hat ihre differentielle Expression einen Einfluss auf die Disseminierung? An welchen Signalwegen sind die dazu gehörigen Proteine beteiligt? Um Antworten auf diese Fragestellungen zu erhalten, ist es von großem Interesse, disseminierte Tumorzellen direkt zu studieren. Die Detektion solcher Zellen ist allerdings schwierig, da die Zellen extrem selten sind und es an spezifischen Markern mangelt.

Das Ziel dieser Arbeit war es, neue Marker für die Detektion epithelialer Zellen in mesenchyalem Gewebe zu finden. Dafür wurde die differentielle Genexpression zwischen vier metastatischen Zellen von Brustkrebspatienten (Tester) und Knochenmarkzellen gesunder Patienten (Driver) mit der Suppression Subtractive Hybridization (SSH) Methode analysiert. Da das Ausgangsmaterial aus global amplifizierter cDNA einzelner Zellen besondere Adaptersequenzen trägt, musste die SSH zuerst daran angepasst werden.

Neun der gewonnenen Sequenzen zeigten eine differentielle Expression, die positiv im Tester und negativ im Driver war. Drei der identifizierten Sequenzen (BAF57, IGF1R und LPP) sind für ihre potentielle Beteiligung bei der Tumorentstehung bereits bekannt. Die Expression der neun Sequenzen wurde in elf Tumorzelllinien und Knochenmark geprüft. Drei Sequenzen sind nur in Zelllinien und nicht in Knochenmark nachzuweisen: IGF1R, BAF57 und eine Sequenz von uns #288 genannt, die keinem bereits bekannten Gen entspricht. Sie konnte als Fragment eines von ERV9 Long Terminal Repeat (LTR) gesteuerten Transkriptes identifiziert werden.

LTRs von humanen endogenen Retroviren (HERV) enthalten Virale Promotoren, Enhancer und Polyadenylierungssignale und können die Genexpression regulieren. Es wurden bereits verschiedene zelluläre mRNA-Transkripte gefunden, die von ERV9 LTRs gesteuert werden. Die differentielle Expression der Sequenz #288 zwischen Karzinomzelllinien und Knochenmark deutet darauf hin, dass der ERV9 LTR für eine Gewebespezifität verantwortlich zeichnen könnte. Da in Krebsgewebe auch eine höhere Transkriptionsaktivität für nicht kodierende Transkripte festgestellt wurde, stellt sich die Frage, ob die Sequenz #288 entweder einen neuen Tumormarker darstellen könnte oder an der Regulation anderer Gene beteiligt ist.

6 Literaturverzeichnis

- Abati A., and Liotta L. A. (1996). Looking forward in diagnostic pathology: the molecular superhighway. *Cancer* **78**: 1-3.
- All-Ericsson C., Girmila L., Seregard S., Bartolazzi A., Jager M. J., and Larsson O. (2002). Insulin-like growth factor-1 receptor in uveal melanoma: a predictor for metastatic disease and a potential therapeutic target. *Invest Ophthalmol Vis Sci* **43**: 1-8.
- Altschul S. F., Madden T. L., Schaffer A. A., Zhang J., Zhang Z., Miller W., and Lipman D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**: 3389-402.
- Asch B. B., Asch H. L., Stoler D. L., and Anderson G. R. (1993). De-regulation of endogenous retrotransposons in mouse mammary carcinomas of diverse etiologies. *Int J Cancer* **54**: 813-9.
- Ashar H. R., Fejzo M. S., Tkachenko A., Zhou X., Fletcher J. A., Weremowicz S., Morton C. C., and Chada K. (1995). Disruption of the architectural factor HMGI-C: DNA-binding AT hook motifs fused in lipomas to distinct transcriptional regulatory domains. *Cell* **82**: 57-65.
- Balzar M., Winter M. J., de Boer C. J., and Litvinov S. V. (1999). The biology of the 17-1A antigen (Ep-CAM). *J Mol Med* **77**: 699-712.
- Birnboim H. C., and Doly J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res* **7**: 1513-23.
- Black A. R., Black J. D., and Azizkhan-Clifford J. (2001). Sp1 and kruppel-like factor family of transcription factors in cell growth regulation and cancer. *J Cell Physiol* **188**: 143-60.
- Bottger E. C. (1990). Frequent contamination of Taq polymerase with DNA. *Clin Chem* **36**: 1258-9.
- Braun S., Pantel K., Muller P., Janni W., Hepp F., Kentenich C. R., Gastroph S., Wischnik A., Dimpfl T., Kindermann G., Riethmuller G., and Schlimok G. (2000). Cytokeratin-positive cells in the bone marrow and survival of patients with stage I, II, or III breast cancer. *N Engl J Med* **342**: 525-33.
- Braun S., Schindlbeck C., Hepp F., Janni W., Kentenich C., Riethmuller G., and Pantel K. (2001). Occult tumor cells in bone marrow of patients with locoregionally restricted ovarian cancer predict early distant metastatic relapse. *J Clin Oncol* **19**: 368-75.
- Braun S., Vogl F. D., Naume B., Janni W., Osborne M. P., Coombes R. C., Schlimok G., Diel I. J., Gerber B., Gebauer G., Pierga J. Y., Marth C., Oruzio D., Wiedswang G., Solomayer E. F., Kundt G., Strobl B., Fehm T., Wong G. Y., Bliss J., Vincent-Salomon A., and Pantel K. (2005). A pooled analysis of bone marrow micrometastasis in breast cancer. *N Engl J Med* **353**: 793-802.
- Brown P. O., and Botstein D. (1999). Exploring the new world of the genome with DNA microarrays. *Nat Genet* **21**: 33-7.
- Brugger W., Buhring H. J., Grunebach F., Vogel W., Kaul S., Muller R., Brummendorf T. H., Ziegler B. L., Rappold I., Brossart P., Scheduling S., and Kanz L. (1999). Expression of MUC-1 epitopes on normal bone marrow: implications for the detection of micrometastatic tumor cells. *J Clin Oncol* **17**: 1535-44.
- Buhring H. J., Muller T., Herbst R., Cole S., Rappold I., Schuller W., Zhu X., Fritsch U., Faul C., Armeanu S., Ullrich A., Klein G., and Schmidt H. (1996). The adhesion molecule E-cadherin and a surface antigen recognized by the antibody 9C4 are selectively expressed on erythroid cells of defined maturational stages. *Leukemia* **10**: 106-16.
- Carter B. A., Jensen R. A., Simpson J. F., and Page D. L. (2000). Benign transport of breast epithelium into axillary lymph nodes after biopsy. *Am J Clin Pathol* **113**: 259-65.

- Chen H. J., Carr K., Jerome R. E., and Edenberg H. J. (2002). A retroviral repetitive element confers tissue-specificity to the human alcohol dehydrogenase 1C (ADH1C) gene. *DNA Cell Biol* **21**: 793-801.
- Chu W. M., Ballard R., Carpick B. W., Williams B. R., and Schmid C. W. (1998). Potential Alu function: regulation of the activity of double-stranded RNA-activated kinase PKR. *Mol Cell Biol* **18**: 58-68.
- Clark J. M. (1988). Novel non-templated nucleotide addition reactions catalyzed by procaryotic and eucaryotic DNA polymerases. *Nucleic Acids Res* **16**: 9677-86.
- Consortium I. H. G. S. (2004). Finishing the euchromatic sequence of the human genome. *Nature* **431**: 931-45.
- Corless C. E., Guiver M., Borrow R., Edwards-Jones V., Kaczmarek E. B., and Fox A. J. (2000). Contamination and sensitivity issues with a real-time universal 16S rRNA PCR. *J Clin Microbiol* **38**: 1747-52.
- Costas J., and Naveira H. (2000). Evolutionary history of the human endogenous retrovirus family ERV9. *Mol Biol Evol* **17**: 320-30.
- Daheron L., Veinstein A., Brizard F., Drabkin H., Lacotte L., Guilhot F., Larsen C. J., Brizard A., and Roche J. (2001). Human LPP gene is fused to MLL in a secondary acute leukemia with a t(3;11) (q28;q23). *Genes Chromosomes Cancer* **31**: 382-9.
- Debies M. T., and Welch D. R. (2001). Genetic basis of human breast cancer metastasis. *J Mammary Gland Biol Neoplasia* **6**: 441-51.
- Debus E., Weber K., and Osborn M. (1982). Monoclonal cytokeratin antibodies that distinguish simple from stratified squamous epithelia: characterization on human tissues. *Embo J* **1**: 1641-7.
- Decristofaro M. F., Betz B. L., Rorie C. J., Reisman D. N., Wang W., and Weissman B. E. (2001). Characterization of SWI/SNF protein expression in human breast cancer cell lines and other malignancies. *J Cell Physiol* **186**: 136-45.
- Di Cristofano A., Strazullo M., Longo L., and La Mantia G. (1995). Characterization and genomic mapping of the ZNF80 locus: expression of this zinc-finger gene is driven by a solitary LTR of ERV9 endogenous retroviral family. *Nucleic Acids Res* **23**: 2823-30.
- Dias Neto E., Correa R. G., Verjovski-Almeida S., Briones M. R., Nagai M. A., da Silva W., Jr., Zago M. A., Bordin S., Costa F. F., Goldman G. H., Carvalho A. F., Matsukuma A., Baia G. S., Simpson D. H., Brunstein A., de Oliveira P. S., Bucher P., Jongeneel C. V., O'Hare M. J., Soares F., Brentani R. R., Reis L. F., de Souza S. J., and Simpson A. J. (2000). Shotgun sequencing of the human transcriptome with ORF expressed sequence tags. *Proc Natl Acad Sci U S A* **97**: 3491-6.
- Diatchenko L., Lau Y. F., Campbell A. P., Chenchik A., Moqadam F., Huang B., Lukyanov S., Lukyanov K., Gurskaya N., Sverdlov E. D., and Siebert P. D. (1996). Suppression subtractive hybridization: a method for generating differentially regulated or tissue-specific cDNA probes and libraries. *Proc Natl Acad Sci U S A* **93**: 6025-30.
- Diatchenko L., Lukyanov S., Lau Y. F., and Siebert P. D. (1999). Suppression subtractive hybridization: a versatile method for identifying differentially expressed genes. *Methods Enzymol* **303**: 349-80.
- Dickson D. (1999). Gene estimate rises as US and UK discuss freedom of access. *Nature* **401**: 311.
- Difilippantonio S., Chen Y., Pietas A., Schluns K., Pacyna-Gengelbach M., Deutschmann N., Padilla-Nash H. M., Ried T., and Petersen I. (2003). Gene expression profiles in human non-small and small-cell lung cancers. *Eur J Cancer* **39**: 1936-47.
- Engel J., Schubert-Fritschle G., and Hölzel D. (2001). Epidemiologie. In "Mammakarzinome" (H. Sauer, Ed.), pp. 1-8, W. Zuckschwerdt Verlag München, München.
- Feinberg A. P., and Vogelstein B. (1983). A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal Biochem* **132**: 6-13.

- Frendo J. L., Olivier D., Cheynet V., Blond J. L., Bouton O., Vidaud M., Rabreau M., Evain-Brion D., and Mallet F. (2003). Direct involvement of HERV-W Env glycoprotein in human trophoblast cell fusion and differentiation. *Mol Cell Biol* **23**: 3566-74.
- Frigerio J. M., Berthezene P., Garrido P., Ortiz E., Barthelley S., Vasseur S., Sastre B., Seleznieff I., Dagorn J. C., and Iovanna J. L. (1995). Analysis of 2166 clones from a human colorectal cancer cDNA library by partial sequencing. *Hum Mol Genet* **4**: 37-43.
- Funke I., and Schraut W. (1998). Meta-analyses of studies on bone marrow micrometastases: an independent prognostic impact remains to be substantiated. *J Clin Oncol* **16**: 557-66.
- Gerber B., Krause A., Muller H., Richter D., Reimer T., Makovitzky J., Herrnring C., Jeschke U., Kundt G., and Friese K. (2001). Simultaneous immunohistochemical detection of tumor cells in lymph nodes and bone marrow aspirates in breast cancer and its correlation with other prognostic factors. *J Clin Oncol* **19**: 960-71.
- Ghossein R. A., and Rosai J. (1996). Polymerase chain reaction in the detection of micrometastases and circulating tumor cells. *Cancer* **78**: 10-6.
- Gottlinger H. G., Funke I., Johnson J. P., Gokel J. M., and Riethmuller G. (1986). The epithelial cell surface antigen 17-1A, a target for antibody-mediated tumor therapy: its biochemical nature, tissue distribution and recognition by different monoclonal antibodies. *Int J Cancer* **38**: 47-53.
- Hanahan D., and Weinberg R. A. (2000). The hallmarks of cancer. *Cell* **100**: 57-70.
- Harrington J. J., Sherf B., Rundlett S., Jackson P. D., Perry R., Cain S., Leventhal C., Thornton M., Ramachandran R., Whittington J., Lerner L., Costanzo D., McElligott K., Boozer S., Mays R., Smith E., Veloso N., Klika A., Hess J., Cothren K., Lo K., Offenbacher J., Danzig J., and Ducar M. (2001). Creation of genome-wide protein expression libraries using random activation of gene expression. *Nat Biotechnol* **19**: 440-5.
- Harris J. R. (1998). Placental endogenous retrovirus (ERV): structural, functional, and evolutionary significance. *Bioessays* **20**: 307-16.
- Hastie N. D., and Bishop J. O. (1976). The expression of three abundance classes of messenger RNA in mouse tissues. *Cell* **9**: 761-74.
- Hermanek P. (1994). What's new in TNM? *Pathol Res Pract* **190**: 97-102.
- Hermanek P., Hutter R. V., Sobin L. H., and Wittekind C. (1999). International Union Against Cancer. Classification of isolated tumor cells and micrometastasis. *Cancer* **86**: 2668-73.
- Hillier L. D., Lennon G., Becker M., Bonaldo M. F., Chiapelli B., Chisoe S., Dietrich N., DuBuque T., Favello A., Gish W., Hawkins M., Hultman M., Kucaba T., Lacy M., Le M., Le N., Mardis E., Moore B., Morris M., Parsons J., Prange C., Rifkin L., Rohlfing T., Schellenberg K., Marra M., and et al. (1996). Generation and analysis of 280,000 human expressed sequence tags. *Genome Res* **6**: 807-28.
- Holton T. A., and Graham M. W. (1991). A simple and efficient method for direct cloning of PCR products using ddT-tailed vectors. *Nucleic Acids Res* **19**: 1156.
- Howard B. A., and Gusterson B. A. (2000). Human breast development. *J Mammary Gland Biol Neoplasia* **5**: 119-37.
- Hubank M., and Schatz D. G. (1994). Identifying differences in mRNA expression by representational difference analysis of cDNA. *Nucleic Acids Res* **22**: 5640-8.
- Hufton S. E., Moerkerk P. T., Brandwijk R., de Bruine A. P., Arends J. W., and Hoogenboom H. R. (1999). A profile of differentially expressed genes in primary colorectal cancer using suppression subtractive hybridization. *FEBS Lett* **463**: 77-82.
- Hughes M. S., Beck L. A., and Skuce R. A. (1994). Identification and elimination of DNA sequences in Taq DNA polymerase. *J Clin Microbiol* **32**: 2007-8.

- Jacob K., Webber M., Benayahu D., and Kleinman H. K. (1999). Osteonectin promotes prostate cancer cell migration and invasion: a possible mechanism for metastasis to bone. *Cancer Res* **59**: 4453-7.
- Ji W., Wright M. B., Cai L., Flament A., and Lindpaintner K. (2002). Efficacy of SSH PCR in isolating differentially expressed genes. *BMC Genomics* **3**: 12.
- Jurka J. (2000). Repbase update: a database and an electronic journal of repetitive elements. *Trends Genet* **16**: 418-20.
- Karolchik D., Baertsch R., Diekhans M., Furey T. S., Hinrichs A., Lu Y. T., Roskin K. M., Schwartz M., Sugnet C. W., Thomas D. J., Weber R. J., Haussler D., and Kent W. J. (2003). The UCSC Genome Browser Database. *Nucleic Acids Res* **31**: 51-4.
- Kato N., Shimotohno K., VanLeeuwen D., and Cohen M. (1990). Human proviral mRNAs down regulated in choriocarcinoma encode a zinc finger protein related to Kruppel. *Mol Cell Biol* **10**: 4401-5.
- Klein C. A. (2000). The biology and analysis of single disseminated tumour cells. *Trends Cell Biol* **10**: 489-93.
- Klein C. A. (2003). The systemic progression of human cancer: a focus on the individual disseminated cancer cell--the unit of selection. *Adv Cancer Res* **89**: 35-67.
- Klein C. A., Schmidt-Kittler O., Schardt J. A., Pantel K., Speicher M. R., and Riethmuller G. (1999). Comparative genomic hybridization, loss of heterozygosity, and DNA sequence analysis of single cells. *Proc Natl Acad Sci U S A* **96**: 4494-9.
- Klein C. A., Seidl S., Petat-Dutter K., Offner S., Geigl J. B., Schmidt-Kittler O., Wendler N., Passlick B., Huber R. M., Schlimok G., Baeuerle P. A., and Riethmuller G. (2002). Combined transcriptome and genome analysis of single micrometastatic cells. *Nat Biotechnol* **20**: 387-92.
- Klochender-Yeivin A., Muchardt C., and Yaniv M. (2002). SWI/SNF chromatin remodeling and cancer. *Curr Opin Genet Dev* **12**: 73-9.
- Kolomietz E., Meyn M. S., Pandita A., and Squire J. A. (2002). The role of Alu repeat clusters as mediators of recurrent chromosomal aberrations in tumors. *Genes Chromosomes Cancer* **35**: 97-112.
- Krag D. N., Ashikaga T., Moss T. J., Kusminsky R. E., Feldman S., Carp N. Z., Moffat F. L., Beitsch P. D., Frazier T. G., Gaskin T. A., Shook J. W., Harlow S. P., and Weaver D. L. (1999). Breast Cancer Cells in the Blood: A Pilot Study. *Breast J* **5**: 354-358.
- Kuang W. W., Thompson D. A., Hoch R. V., and Weigel R. J. (1998). Differential screening and suppression subtractive hybridization identified genes differentially expressed in an estrogen receptor- positive breast carcinoma cell line. *Nucleic Acids Res* **26**: 1116-23.
- Kurachi S., Deyashiki Y., Takeshita J., and Kurachi K. (1999). Genetic mechanisms of age regulation of human blood coagulation factor IX. *Science* **285**: 739-43.
- La Mantia G., Maglione D., Pengue G., Di Cristofano A., Simeone A., Lanfrancone L., and Lania L. (1991). Identification and characterization of novel human endogenous retroviral sequences preferentially expressed in undifferentiated embryonal carcinoma cells. *Nucleic Acids Res* **19**: 1513-20.
- La Mantia G., Majello B., Di Cristofano A., Strazzullo M., Minchiotti G., and Lania L. (1992). Identification of regulatory elements within the minimal promoter region of the human endogenous ERV9 proviruses: accurate transcription initiation is controlled by an Inr-like element. *Nucleic Acids Res* **20**: 4129-36.
- La Mantia G., Pengue G., Maglione D., Pannuti A., Pascucci A., and Lania L. (1989). Identification of new human repetitive sequences: characterization of the corresponding cDNAs and their expression in embryonal carcinoma cells. *Nucleic Acids Res* **17**: 5913-22.

- Lacey J. V., Jr., Devesa S. S., and Brinton L. A. (2002). Recent trends in breast cancer incidence and mortality. *Environ Mol Mutagen* **39**: 82-8.
- Lammers R., Giesert C., Grunebach F., Marxer A., Vogel W., and Buhring H. J. (2002). Monoclonal antibody 9C4 recognizes epithelial cellular adhesion molecule, a cell surface antigen expressed in early steps of erythropoiesis. *Exp Hematol* **30**: 537-45.
- Lander E. S., Linton L. M., Birren B., Nusbaum C., Zody M. C., Baldwin J., Devon K., Dewar K., Doyle M., FitzHugh W., Funke R., Gage D., Harris K., Heaford A., Howland J., Kann L., Lehoczky J., LeVine R., McEwan P., McKernan K., Meldrim J., Mesirov J. P., Miranda C., Morris W., Naylor J., Raymond C., Rosetti M., Santos R., Sheridan A., Sougnez C., Stange-Thomann N., Stojanovic N., Subramanian A., Wyman D., Rogers J., Sulston J., Ainscough R., Beck S., Bentley D., Burton J., Clee C., Carter N., Coulson A., Deadman R., Deloukas P., Dunham A., Dunham I., Durbin R., French L., Grafham D., Gregory S., Hubbard T., Humphray S., Hunt A., Jones M., Lloyd C., McMurray A., Matthews L., Mercer S., Milne S., Mullikin J. C., Mungall A., Plumb R., Ross M., Shownkeen R., Sims S., Waterston R. H., Wilson R. K., Hillier L. W., McPherson J. D., Marra M. A., Mardis E. R., Fulton L. A., Chinwalla A. T., Pepin K. H., Gish W. R., Chissoe S. L., Wendl M. C., Delehaunty K. D., Miner T. L., Delehaunty A., Kramer J. B., Cook L. L., Fulton R. S., Johnson D. L., Minx P. J., Clifton S. W., Hawkins T., Branscomb E., Predki P., Richardson P., Wenning S., Slezak T., Doggett N., Cheng J. F., Olsen A., Lucas S., Elkin C., Uberbacher E., Frazier M., et al. (2001). Initial sequencing and analysis of the human genome. *Nature* **409**: 860-921.
- Landry J. R., and Mager D. L. (2003). Functional analysis of the endogenous retroviral promoter of the human endothelin B receptor gene. *J Virol* **77**: 7459-66.
- Lania L., Di Cristofano A., Strazzullo M., Pengue G., Majello B., and La Mantia G. (1992). Structural and functional organization of the human endogenous retroviral ERV9 sequences. *Virology* **191**: 464-8.
- Lebeau A., Nathrath W., and Permanetter W. (2001). Pathomorphologie des Mammakarzinoms. In "Mammakarzinome" (H. Sauer, Ed.), pp. 18-33, W. Zuckschwerdt Verlag München, München.
- Lee W. J., Kwun H. J., and Jang K. L. (2003). Analysis of transcriptional regulatory sequences in the human endogenous retrovirus W long terminal repeat. *J Gen Virol* **84**: 2229-35.
- Levi F., Lucchini F., Negri E., and La Vecchia C. (2001). The fall in breast cancer mortality in Europe. *Eur J Cancer* **37**: 1409-12.
- Lewin B. (1975). Units of transcription and translation: the relationship between heterogeneous nuclear RNA and messenger RNA. *Cell* **4**: 11-20.
- Li S. R., Dorudi S., and Bustin S. A. (2003). Identification of differentially expressed genes associated with colorectal cancer liver metastasis. *Eur Surg Res* **35**: 327-36.
- Ling J., Pi W., Bollag R., Zeng S., Keskinetepe M., Saliman H., Krantz S., Whitney B., and Tuan D. (2002). The solitary long terminal repeats of ERV-9 endogenous retrovirus are conserved during primate evolution and possess enhancer activities in embryonic and hematopoietic cells. *J Virol* **76**: 2410-23.
- Ling J., Pi W., Yu X., Bengra C., Long Q., Jin H., Seyfang A., and Tuan D. (2003). The ERV-9 LTR enhancer is not blocked by the HS5 insulator and synthesizes through the HS5 site non-coding, long RNAs that regulate LTR enhancer function. *Nucleic Acids Res* **31**: 4582-96.
- Liu A. Y., and Abraham B. A. (1991). Subtractive cloning of a hybrid human endogenous retrovirus and calbindin gene in the prostate cell line PC3. *Cancer Res* **51**: 4107-10.

- Long Q., Bengra C., Li C., Kutlar F., and Tuan D. (1998). A long terminal repeat of the human endogenous retrovirus ERV-9 is located in the 5' boundary area of the human beta-globin locus control region. *Genomics* **54**: 542-55.
- Lower R., Tonjes R. R., Korbmacher C., Kurth R., and Lower J. (1995). Identification of a Rev-related protein by analysis of spliced transcripts of the human endogenous retroviruses HTDV/HERV-K. *J Virol* **69**: 141-9.
- MacDougall J. R., and Matrisian L. M. (2000). Targets of extinction: identification of genes whose expression is repressed as a consequence of somatic fusion between cells representing basal and luminal mammary epithelial phenotypes. *J Cell Sci* **113**: 409-23.
- Mallon E., Osin P., Nasiri N., Blain I., Howard B., and Gusterson B. (2000). The basic pathology of human breast cancer. *J Mammary Gland Biol Neoplasia* **5**: 139-63.
- Mansi J. L., Easton D., Berger U., Gazet J. C., Ford H. T., Dearnaley D., and Coombes R. C. (1991). Bone marrow micrometastases in primary breast cancer: prognostic significance after 6 years' follow-up. *Eur J Cancer* **27**: 1552-5.
- Mansi J. L., Gogas H., Bliss J. M., Gazet J. C., Berger U., and Coombes R. C. (1999). Outcome of primary-breast-cancer patients with micrometastases: a long-term follow-up study. *Lancet* **354**: 197-202.
- Maor S. B., Abramovitch S., Erdos M. R., Brody L. C., and Werner H. (2000). BRCA1 suppresses insulin-like growth factor-I receptor promoter activity: potential interaction between BRCA1 and Sp1. *Mol Genet Metab* **69**: 130-6.
- Marchuk D., Drumm M., Saulino A., and Collins F. S. (1991). Construction of T-vectors, a rapid and general system for direct cloning of unmodified PCR products. *Nucleic Acids Res* **19**: 1154.
- Marth C., Kisic J., Kaern J., Trope C., and Fodstad O. (2002). Circulating tumor cells in the peripheral blood and bone marrow of patients with ovarian carcinoma do not predict prognosis. *Cancer* **94**: 707-12.
- Mattick J. S., and Makunin I. V. (2005). Small regulatory RNAs in mammals. *Hum Mol Genet* **14 Spec No 1**: R121-32.
- Mauro L., and Surmacz E. (2004). IGF-I receptor, cell-cell adhesion, tumour development and progression. *J Mol Histol* **35**: 247-53.
- Mitchell P. J., and Tjian R. (1989). Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. *Science* **245**: 371-8.
- Moll R., Franke W. W., Schiller D. L., Geiger B., and Krepler R. (1982). The catalog of human cytokeratins: patterns of expression in normal epithelia, tumors and cultured cells. *Cell* **31**: 11-24.
- Müller A., Homey B., Soto H., Ge N., Catron D., Buchanan M. E., McClanahan T., Murphy E., Yuan W., Wagner S. N., Barrera J. L., Mohar A., Verastegui E., and Zlotnik A. (2001). Involvement of chemokine receptors in breast cancer metastasis. *Nature* **410**: 50-6.
- Natoli F., Crowley M. R., Asch H. L., Stoler D. L., and Asch B. B. (1996). Mutations involving the endogenous ecotropic murine leukemia virus in primary mammary carcinomas of BALB/c mice. *Cancer Lett* **99**: 121-7.
- Ostertag E. M., and Kazazian H. H., Jr. (2001). Biology of mammalian L1 retrotransposons. *Annu Rev Genet* **35**: 501-38.
- Pantel K., Cote R. J., and Fodstad O. (1999). Detection and clinical importance of micrometastatic disease. *J Natl Cancer Inst* **91**: 1113-24.
- Perou C. M., Sorlie T., Eisen M. B., van de Rijn M., Jeffrey S. S., Rees C. A., Pollack J. R., Ross D. T., Johnsen H., Akslen L. A., Fluge O., Pergamenschikov A., Williams C., Zhu S. X., Lonning P. E., Borresen-Dale A. L., Brown P. O., and Botstein D. (2000). Molecular portraits of human breast tumours. *Nature* **406**: 747-52.

- Pitelka D. R. (1988). The mammary gland. In "Cell and tissue biology" (L. Weiss, Ed.), pp. 879-898, Urban and Schwarzenberg, Munich.
- Prak E. T., and Kazazian H. H., Jr. (2000). Mobile elements and the human genome. *Nat Rev Genet* **1**: 134-44.
- Rebrikov D. V., Britanova O. V., Gurskaya N. G., Lukyanov K. A., Tarabykin V. S., and Lukyanov S. A. (2000). Mirror orientation selection (MOS): a method for eliminating false positive clones from libraries generated by suppression subtractive hybridization. *Nucleic Acids Res* **28**: E90.
- Robert C., Barnes F. L., Hue I., and Sirard M. A. (2000). Subtractive hybridization used to identify mRNA associated with the maturation of bovine oocytes. *Mol Reprod Dev* **57**: 167-75.
- Saiki R. K., Gelfand D. H., Stoffel S., Scharf S. J., Higuchi R., Horn G. T., Mullis K. B., and Erlich H. A. (1988). Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* **239**: 487-91.
- Sambrook J., and Russel D. W. (2001). "Molecular Cloning," Cold Spring Harbor Laboratory Press, New York.
- Schiavetti F., Thonnard J., Colau D., Boon T., and Coulie P. G. (2002). A human endogenous retroviral sequence encoding an antigen recognized on melanoma by cytolytic T lymphocytes. *Cancer Res* **62**: 5510-6.
- Schlimok G., Funke I., Holzmann B., Gottlinger G., Schmidt G., Hauser H., Swierkot S., Warnecke H. H., Schneider B., Koprowski H., and et al. (1987). Micrometastatic cancer cells in bone marrow: in vitro detection with anti-cytokeratin and in vivo labeling with anti-17-1A monoclonal antibodies. *Proc Natl Acad Sci U S A* **84**: 8672-6.
- Schoenmakers E. F., Wanschura S., Mols R., Bullerdiek J., Van den Berghe H., and Van de Ven W. J. (1995). Recurrent rearrangements in the high mobility group protein gene, HMGI-C, in benign mesenchymal tumours. *Nat Genet* **10**: 436-44.
- Schon U., Seifarth W., Baust C., Hohenadl C., Erfle V., and Leib-Mosch C. (2001). Cell type-specific expression and promoter activity of human endogenous retroviral long terminal repeats. *Virology* **279**: 280-91.
- Scott A. M., and Welt S. (1997). Antibody-based immunological therapies. *Curr Opin Immunol* **9**: 717-22.
- Seifarth W., Spiess B., Zeilfelder U., Speth C., Hehlmann R., and Leib-Mosch C. (2003). Assessment of retroviral activity using a universal retrovirus chip. *J Virol Methods* **112**: 79-91.
- Seipel K., Georgiev O., and Schaffner W. (1992). Different activation domains stimulate transcription from remote ('enhancer') and proximal ('promoter') positions. *Embo J* **11**: 4961-8.
- Siebert P. D., Chenchik A., Kellogg D. E., Lukyanov K. A., and Lukyanov S. A. (1995). An improved PCR method for walking in uncloned genomic DNA. *Nucleic Acids Res* **23**: 1087-8.
- Sjottem E., Anderssen S., and Johansen T. (1996). The promoter activity of long terminal repeats of the HERV-H family of human retrovirus-like elements is critically dependent on Sp1 family proteins interacting with a GC/GT box located immediately 3' to the TATA box. *J Virol* **70**: 188-98.
- Solomayer E. F., Diel I. J., Salanti G., Hahn M., Gollan C., Schutz F., and Bastert G. (2001). Time independence of the prognostic impact of tumor cell detection in the bone marrow of primary breast cancer patients. *Clin Cancer Res* **7**: 4102-8.
- Spizzo G., Obrist P., Ensinger C., Theurl I., Dunser M., Ramoni A., Gunsilius E., Eibl G., Mikuz G., and Gastl G. (2002). Prognostic significance of Ep-CAM AND Her-2/neu overexpression in invasive breast cancer. *Int J Cancer* **98**: 883-8.

- Sporn M. B. (1996). The war on cancer. *Lancet* **347**: 1377-81.
- Strazzullo M., Majello B., Lania L., and La Mantia G. (1994). Mutational analysis of the human endogenous ERV9 proviruses promoter region. *Virology* **200**: 686-95.
- Strazzullo M., Parisi T., Di Cristofano A., Rocchi M., and La Mantia G. (1998). Characterization and genomic mapping of chimeric ERV9 endogenous retroviruses-host gene transcripts. *Gene* **206**: 77-83.
- Sugino A., Goodman H. M., Heyneker H. L., Shine J., Boyer H. W., and Cozzarelli N. R. (1977). Interaction of bacteriophage T4 RNA and DNA ligases in joining of duplex DNA at base-paired ends. *J Biol Chem* **252**: 3987-94.
- Svensson A. C., Raudsepp T., Larsson C., Di Cristofano A., Chowdhary B., La Mantia G., Rask L., and Andersson G. (2001). Chromosomal distribution, localization and expression of the human endogenous retrovirus ERV9. *Cytogenet Cell Genet* **92**: 89-96.
- Tenen D. G., Hromas R., Licht J. D., and Zhang D. E. (1997). Transcription factors, normal myeloid development, and leukemia. *Blood* **90**: 489-519.
- van de Vijver M. J., He Y. D., van't Veer L. J., Dai H., Hart A. A., Voskuil D. W., Schreiber G. J., Peterse J. L., Roberts C., Marton M. J., Parrish M., Atsma D., Witteveen A., Glas A., Delahaye L., van der Velde T., Bartelink H., Rodenhuis S., Rutgers E. T., Friend S. H., and Bernards R. (2002). A gene-expression signature as a predictor of survival in breast cancer. *N Engl J Med* **347**: 1999-2009.
- Velculescu V. E., Zhang L., Vogelstein B., and Kinzler K. W. (1995). Serial analysis of gene expression. *Science* **270**: 484-7.
- Venter J. C., Adams M. D., Myers E. W., Li P. W., and Mural R. J. (2001). The sequence of the human genome. *Science* **291**: 1304-51.
- Wang W., Chi T., Xue Y., Zhou S., Kuo A., and Crabtree G. R. (1998). Architectural DNA binding by a high-mobility-group/kinesin-like subunit in mammalian SWI/SNF-related complexes. *Proc Natl Acad Sci U S A* **95**: 492-8.
- Wang-Johanning F., Frost A. R., Jian B., Epp L., Lu D. W., and Johanning G. L. (2003). Quantitation of HERV-K env gene expression and splicing in human breast cancer. *Oncogene* **22**: 1528-35.
- Washietl S., Hofacker I. L., and Stadler P. F. (2005). Fast and reliable prediction of noncoding RNAs. *Proc Natl Acad Sci U S A* **102**: 2454-9.
- Wells C. A., Heryet A., Brochier J., Gatter K. C., and Mason D. Y. (1984). The immunocytochemical detection of axillary micrometastases in breast cancer. *Br J Cancer* **50**: 193-7.
- Werner H., Karnieli E., Rauscher F. J., and LeRoith D. (1996). Wild-type and mutant p53 differentially regulate transcription of the insulin-like growth factor I receptor gene. *Proc Natl Acad Sci U S A* **93**: 8318-23.
- Wittekind C., and Wagner G. (1997). "TNM-Klassifikation Maligner Tumoren," Springer-Verlag, Heidelberg.
- Xie D., Jauch A., Miller C. W., Bartram C. R., and Koeffler H. P. (2002). Discovery of over-expressed genes and genetic alterations in breast cancer cells using a combination of suppression subtractive hybridization, multiplex FISH and comparative genomic hybridization. *Int J Oncol* **21**: 499-507.
- Yang Z., Boffelli D., Boonmark N., Schwartz K., and Lawn R. (1998). Apolipoprotein(a) gene enhancer resides within a LINE element. *J Biol Chem* **273**: 891-7.
- Zanders E. D., Goulden M. G., Kennedy T. C., and Kempell K. E. (2000). Analysis of immune system gene expression in small rheumatoid arthritis biopsies using a combination of subtractive hybridization and high-density cDNA arrays. *J Immunol Methods* **233**: 131-40.

Zhang L., Zhou W., Velculescu V. E., Kern S. E., Hruban R. H., Hamilton S. R., Vogelstein B., and Kinzler K. W. (1997). Gene expression profiles in normal and cancer cells. *Science* **276**: 1268-72.

7 Danksagung

Ich danke Prof. Gert Riethmüller und Prof. Thomas Brocker für die Möglichkeit meine Doktorarbeit an ihrem Institut durchzuführen.

Besonderer Dank gilt PD Dr. Christoph Klein für die gute Betreuung, die vielen Ideen zu Methodik und Durchführung der Versuche und seine optimistische Einstellung, mit der er mich immer wieder motivieren konnte. Er hat immer verstanden, dass eine Frau auch mit kleinen Kindern berufstätig bleiben will. Er hat immer an mich und meine Arbeit geglaubt und mir alle Freiheiten bei der Durchführung meines Projektes gelassen.

Oleg Schmidt-Kittler, Thomas Blankenstein, Nicholas Stöcklein und Jochen Geigl danke ich für die sorgfältige Korrektur meiner Arbeit.

Großer Dank gilt Kerstin Bockler und Susanne Steinhaus, die meinen französischen Satzbau auch für deutsche Leser verständlich gemacht haben.

Vielen Dank allen Mitarbeitern der Arbeitsgruppe für die tolle Stimmung und die gute Zusammenarbeit.

8 Erklärung

Hiermit erkläre ich, Karina Petat-Dutter, an Eides statt, dass ich die vorgelegte Doktorarbeit selbständig und nur mit den angegebenen Quellen angefertigt habe.

München, den _____

9 Lebenslauf

Name: Karina PETAT-DUTTER
Geburtsdatum: 10.01.1972
Geburtsort: Saint Martin d'Hères (Frankreich)
Familienstand: verheiratet, drei Kinder
Staatsangehörigkeit: französisch

Studium

- 1990 : Baccalauréat C (Abitur / Oyonnax, Frankreich). Abschluß mit der Gesamtnote "befriedigend". Hauptfächer: Biologie, Mathematik, Physik, Chemie.
- 1992 : DEUG B (Abschlußprüfung des allgemeinen Universitätsstudiums nach zwei Studienjahren / Vordiplom), Claude Bernard Universität Lyon (Lyon, Frankreich). Abschluß mit der Gesamtnote "gut". Hauptfach: Geologie.
- 1995 : Diplom Ingenieurin, CUST, Hochschule für Ingenieurwesen (Clermont-Ferrand, Frankreich).
Kompetenzbereich: Mikrobiologie, Verfahrenstechniken, Biotechnologie.
- Seit 2000: Promotion zum Doktor der Humanbiologie am Institut für Immunologie der LMU unter der Leitung von PD Dr. C. Klein mit dem Thema "Analyse der Genexpression von disseminierten Mammakarzinomzellen aus dem Knochenmark".

Berufserfahrungen

- 1994
4 Monate KRAFT JACOBS SUCHARD (Praktikum / Banbury, England)
- Verbesserung des Herstellungsverfahrens des Instantkaffees CICONA®.
- 1995
9 Monate BIOMERIEUX (Abschluß der Diplomarbeit / Lyon, Frankreich)
- Verwendungsmöglichkeiten der RNA-Polymerase des Bakteriophagen T7 für die Synthese eines RNA-Primers zur Charakterisierung des *Transcription Initiation Complex*.
- Anwendungsmöglichkeiten einer neuen molekularen Technik zur Vermehrung von Genmaterial.
- 1996
6 Monate PHARMACIA BIOTECH (Mutterschaftsvertretung / Paris, Frankreich)
- Kundendienst einer Kampagne für ein Molekularbiologiesortiment.
- 1997
7 Monate BIOMERIEUX (Mutterschaftsvertretung / Lyon, Frankreich).
- Erforschung der *template*-spezifischen Besonderheiten der RNA-Polymerase des Bakteriophagen T7 durch kontrollierte Genmutationen.