

Dissertation zur Erlangung des Doktorgrades
der Fakultät für Chemie und Pharmazie
der Ludwig-Maximilians-Universität München



Grundlagen für zielgerichtete adjuvante Therapien solider Tumore:

* * *

Eine Methode zur Analyse der globalen Genexpression einzelner metastatischer Vorläuferzellen

Claudia Helena Hartmann
aus Heilbronn

2006

Erklärung

Diese Dissertation wurde in Sinne von § 13 Abs. 3 bzw. 4 der Promotionsordnung vom 29. Januar 1998 von Herrn PD Dr. Christoph Klein betreut und von Herrn Prof. Dr. Martin Biel vor der Fakultät für Chemie und Pharmazie vertreten.

Ehrenwörtliche Versicherung

Diese Dissertation wurde selbständig, ohne unerlaubte Hilfe erarbeitet.

München, den 20.06.2006

C. Hartmann

Dissertation eingereicht am: 02.05.2006
Erster Gutachter: PD Dr. Christoph Klein
Zweiter Gutachter: Prof. Dr. Martin Biel
Tag der mündlichen Prüfung: 14.06.2006

**You can't depend on your eyes,
when your imagination is out of focus.**

Mark Twain

Für Alex.

1	<u>EINLEITUNG</u>	1
1.1	DAS MAMMAKARZINOM	1
1.1.1	Klassifikation und Prognose des Mammakarzinoms	1
1.1.2	Therapie des Mammakarzinoms	3
1.2	DIE MINIMAL RESIDUALE KREBSERKRANKUNG (MRD)	4
1.2.1	Begriffsbestimmung	4
1.2.2	Prognostische Bedeutung der MRD	5
1.3	DAS EPITHELIALE ADHÄSIONSMOLEKÜL EpCAM	5
1.4	MOLEKULARE ANALYSE DISSEMINIERTER TUMORZELLEN	7
1.4.1	Charakteristika disseminierter Tumorzellen	7
1.4.2	Untersuchung von disseminierten Tumorzellen beim Mammakarzinom	7
1.5	METHODEN ZUR UNTERSUCHUNG DES TRANSKRIPTOMS VON EINZELZELLEN	8
1.6	GENEXPRESSIONSANALYSE MITTELS ARRAYHYBRIDISIERUNG	8
1.6.1	Historische Entwicklung	8
1.6.2	Technologie der Hybridisierung	10
1.6.3	Design von Mikroarrayexperimenten	10
1.6.4	Bioinformatische Methoden zur Auswertung von Genexpressionsarrays	12
1.6.5	Genexpressionsanalysen beim Mammakarzinom	13
1.7	URSACHEN GESTÖRTER GENEXPRESSION BEI TUMOREN	13
1.8	ZIELSETZUNG	15
2	<u>MATERIAL UND METHODEN</u>	16
2.1	METHODEN	16
2.1.1	Isolierung und Amplifikation von mRNA mit dem PNA-Protokoll	16
2.1.2	Isolierung und Amplifikation von mRNA mit dem Oligo dT-Bead-Protokoll	17
2.1.3	Aufarbeitung der genomischen DNA parallel zur mRNA	17
2.1.4	Quantitative real time PCR	18
2.1.5	Reamplifikations- und Markierungs-PCR	19
2.1.6	Nachträglicher Einbau eines T7-Promotors in das cDNA-Primäramplifikat	20
2.1.7	Effizienztest des T7-Promotor in einer <i>in vitro</i> -Transkription	21
2.1.8	Markierung über den T7-Promotor durch <i>in vitro</i> -Transkription	21
2.1.9	Markierung von Gesamt-RNA während der cDNA-Synthese	21
2.1.10	Herstellung von <i>E.coli</i> -DNA zum Blockieren bei Hybridisierungen	21
2.1.11	Hybridisierung auf murine Operon-Arrays	22
2.1.12	Hybridisierung auf humane Operon-Arrays	23
2.1.13	Auswertung der Operon-Hybridisierungen in GeneSpring	23

2.1.14	Hybridisierung auf Nylonfilter	23
2.1.15	Komparative genomische Hybridisierung (CGH)	24
2.1.16	Kultur von murinen dendritischen Zellen aus Knochenmark	25
2.1.17	Zellsortierung mit dem MACS-System	26
2.1.18	Aufarbeitung von humanem Knochenmark	26
2.2	VERWENDETE STANDARPROTOKOLLE	26
2.2.1	Zellkultur	26
2.2.2	Genspezifische PCR	26
2.2.3	Agarose Gelelektrophorese	26
2.2.4	Konzentrationsbestimmung von Nukleinsäuren	26
2.2.5	Aufreinigung von PCR-Produkten	26
2.2.6	Isolierung von Gesamt-RNA	26
2.2.7	Alternative Markierungsreaktionen für PCR-Produkte	26
2.3	VERZEICHNIS DER VERWENDETEN ABKÜRZUNGEN	27
2.4	MATERIAL, GERÄTE, SOFTWARE, PUFFER UND LÖSUNGEN	27
2.5	OLIGONUKLEOTIDE UND PRIMER	30
2.6	PATIENTENKOLLEKTIV	32
2.7	KOOPERATIONSPARTNER	33
3	<u>ERGEBNISSE: ETABLIERUNG EINES PROTOKOLLS ZUR UNTERSUCHUNG DER GENEXPRESSION VON EINZELZELLEN</u>	34
3.1	OPTIMIERUNG EINES PROTOKOLLS ZUR ANALYSE VON mRNA AUS EINZELZELLEN	34
3.1.1	Etablierung des PNA-Protokolls	34
3.1.2	Lagerung von Proben vor der Aufarbeitung	39
3.1.3	Validierung des Amplifikationsprotokolls	40
3.2	AUSWAHL EINES ARRAYS ZUR ANALYSE DER GENEXPRESSION	42
3.2.1	Affymetrix GeneChip-Arrays	42
3.2.2	Febit-Arrays	48
3.2.3	Memorec PIQOR-Arrays	49
3.3	HYBRIDISIERUNG AUF MURINE OPERON SIEBZIGMER OLIGONUKLEOTIDARRAYS	49
3.3.1	Etablierung des Protokolls	49
3.3.2	Ursachen verbleibender falsch-positiver Signale	53
3.3.3	Reproduzierbarkeit und statistische Trennschärfe des Verfahrens	54
3.3.4	Effektivität der Methode bei verschiedenen Mengen Ausgangs-RNA	55
3.3.5	Methoden zur Markierung von PCR-Produkten für Hybridisierungen	56
3.3.6	Maßnahmen zur Steigerung der Sensitivität	58

3.4	VALIDIERUNG DES HYBRIDISIERUNGSPROTOKOLLS	59
3.4.1	LPS-Stimulation von murinen dendritischen Zellen und Hybridisierung	59
3.4.2	Validierung mittels quantitativer real time PCR	60
3.4.3	Biologische Aussagekraft und Sensitivität des Verfahrens	61
3.5	HERSTELLUNG VON OPERON-ARRAYS AUS LYOPHILISIERTEN OLIGONUKLEOTIDEN	63
3.5.1	Wahl der geeigneten Beschichtung des Objektträgers	64
3.5.2	Wahl des Spottingpuffers	65
3.5.3	Etablierung des Hybridisierungsprotokolls für humane Operon-Arrays	67
3.5.4	Auswahl der optimalen Herstellung humaner Operon-Arrays und Validierung	68
4	<u>ERGEBNISSE: UNTERSUCHUNG DISSEMINIERTER TUMORZELLEN</u>	70
4.1	ETABLIERUNG DER LINEAGE DEPLETION AUS HUMANEM PROBENMATERIAL	70
4.1.1	Auswahl des Depletionsverfahrens zur Anreicherung epithelialer Zellen	70
4.1.2	Optimierung der Dichtegradientenzentrifugation	72
4.1.3	Wiederfindungsrate epithelialer Zellen	73
4.1.4	Schwierigkeiten bei der Untersuchung von Knochenmark	73
4.2	ANALYSE VON KNOCHENMARKSASPIRATEN BEI MAMMAKARZINOMPATIENTINNEN	76
4.2.1	Detektionsrate EpCAM-positiver Zellen und Isolation der Nukleinsäuren	76
4.2.2	Doppelfärbung mit dem Farbstoff Hoechst 33342 und epithelialen Markern	77
4.2.3	Genexpressionsanalyse auf Nylonarrays	77
4.2.4	Genexpressionsanalyse auf Oligonukleotidarrays	79
4.2.5	Untersuchung der genomischen DNA	83
5	<u>DISKUSSION</u>	85
6	<u>ZUSAMMENFASSUNG</u>	99
7	<u>LITERATURVERZEICHNIS</u>	100
8	<u>DANKSAGUNG</u>	109
9	<u>LEBENS LAUF</u>	110
10	<u>VERÖFFENTLICHUNGEN UND STIPENDIEN</u>	111
11	<u>ANHANG</u>	112

1 Einleitung

1.1 Das Mammakarzinom

Trotz fallender Inzidenzraten im Lauf der letzten Jahre stellt das Mammakarzinom bei Frauen mit 24,4% noch immer die häufigste Krebserkrankung dar. In Deutschland erkrankt etwa jede zehnte Frau im Laufe ihres Lebens an Brustkrebs. Jährlich treten etwa 47500 Neuerkrankungen auf, was einer Inzidenz von 87,8 Fällen pro 100000 Frauen entspricht. Altersspezifisch steigen die Inzidenzraten etwa ab dem 35. Lebensjahr und erreichen ein Plateau bei 55-60 Jahren. Das mittlere Alter bei Diagnosestellung beträgt 61,4 Jahre, etwa 40% der Brustkrebspatientinnen sind jedoch jünger als sechzig¹. Bei Männern ist das Mammakarzinom mit 400 Neuerkrankungen pro Jahr eine seltene Erkrankung.

Im Gegensatz zu vielen anderen Tumoren ist die Mortalität beim Mammakarzinom in den letzten Jahren nicht gesunken und beträgt 28,3 pro 100000 Frauen. Dieser Wert ist seit Einführung der deutschen Krebsstatistik 1967 nahezu konstant. Damit stellt das Mammakarzinom die zweithäufigste krebsbedingte (17,9% der Todesfälle) und insgesamt die vierthäufigste (3,8%) Todesursache bei Frauen dar². Auch das Fünfjahresüberleben konnte trotz intensiver Therapieanstrengungen nicht wesentlich gesteigert werden und beträgt heute im Schnitt 76% gegenüber 73% in den achtziger Jahren³.

Die Situation von Patientinnen mit Mammakarzinom hat sich also trotz umfangreicher Anstrengungen hinsichtlich Forschung und Therapie in den letzten Jahren nicht wesentlich verbessert. Von Standardtherapien profitieren nach wie vor nur etwa 2-10% der erkrankten Frauen^{3,4}.

1.1.1 Klassifikation und Prognose des Mammakarzinoms

Um prognostische Aussagen in der Klinik zu erleichtern, erfolgt eine Einteilung des Mammakarzinoms nach unterschiedlichen Kriterien. Als Hauptmarker dient dabei das TNM-Stadium der UICC⁵, das endgültig postoperativ erhoben und dann mit dem Präfix „p“ gekennzeichnet wird. Dabei werden Tumorgöße, Art und Anzahl der befallenen Lymphknoten und das Vorliegen von Fernmetastasen klassifiziert. Einen Überblick gibt Tabelle 1.1.

Neben diesen Kriterien werden noch etwaige Einbrüche in Venen und Lymphgefäße sowie der Malignitätsgrad anhand des Ausmaßes der Differenzierung sowie der Resektionsrand beurteilt.

In Europa ist neben der TNM-Klassifikation auch das so genannte „Grading“ gebräuchlich. Dabei erfolgt eine rein histologisch-morphologische Einteilung⁶. In den USA wird die TNM-Klassifikation zur Vereinfachung in Stadien überführt (Stage I-IV). TNM T₁, N₀, M₀ entspricht dabei zum Beispiel Stage I.

Tabelle 1.1 TNM-Klassifikation des Mammakarzinoms (UICC 2002)

T: Ausdehnung des Primärtumors	
T _x	Primärtumor nicht bestimmbar
T ₀	kein Hinweis auf Primärtumor
T _{is}	Carcinoma in situ, keine nachweisbare Tumormasse
T ₁	Tumor ≤ 2 cm
T ₂	2 cm < Tumor ≤ 5 cm
T ₃	Tumor > 5 cm
T ₄	Tumor jeglicher Größe mit direkter Ausbreitung auf Haut oder Brustwand; inflammatorische Karzinome, Hautulzerationen
N: Regionäre Lymphknotenmetastasen	
N _x	regionäre Lymphknotenmetastasen nicht bestimmbar
N ₀	keine regionären Lymphknotenmetastasen
N _{mi}	Mikrometastasen (maximal 2 mm)
N ₁	Metastasen in 1-3 axillären Lymphknoten derselben Seite
N ₂	Metastasen in 4-9 axillären Lymphknoten derselben Seite
N ₃	Metastase in Lymphknoten derselben Seite unterhalb des Brustbeins; zehn und mehr befallene axilläre Lymphknoten
M: Fernmetastasen	
M _x	Vorhandensein von Fernmetastasen kann nicht bestimmt werden
M ₀	keine Fernmetastasen
M _{mi}	Mikrometastasen (maximal 2 mm)
M _i	isolierte Tumorzellen in Knochenmark oder parenchymatösen Organen
M ₁	Fernmetastasen vorhanden

Die Prognose eines Mammakarzinoms ist abhängig vom TNM-Stadium. Dabei sinkt das Zehnjahresüberleben deutlich bei größeren Tumoren (81,7% für pT₁ gegenüber 30,7% für pT₄) und vor allem in Abhängigkeit von der Anzahl befallener Lymphknoten (86,4% ohne Lymphknotenmetastasierung gegenüber 27,3% bei 10 oder mehr befallenen Lymphknoten) oder beim Vorliegen von Fernmetastasen⁷. Die häufigsten Orte der Fernmetastasierung sind Knochen (70%), Lunge (60%) und Leber (50%).

Neben den TNM-Kriterien spielen für Prognose und Therapie weitere Faktoren eine Rolle. Dazu gehören die (Über-) Expression verschiedener Rezeptoren wie Estrogen-, Progesteron- oder Her2-Rezeptor. Kommen diese Moleküle auf der Zelloberfläche vor, hat dies zum Teil auch Konsequenzen für die Therapie. Zum Beispiel existiert mit Trastuzumab ein monoklonaler Antikörper gegen den Her2-Rezeptor, von dem Patientinnen mit überexprimiertem Protein profitieren^{8, 9}. Ebenso sind endokrine Therapien nur bei Hormonrezeptor positiven Tumoren effektiv.

In den letzten Jahren wurden weitere molekularbiologische Verfahren zur Charakterisierung der Tumore entwickelt. Aus den gewonnenen Daten lassen sich ebenfalls prognostische Aussagen ableiten. So sind Mutationen in Tumorsuppressorgenen wie *TP53* mit einem reduzierten

Fünfjahresüberleben assoziiert¹⁰. Fortschritte auf diesem Gebiet wurden auch durch Genexpressionsanalysen der Primärtumore auf Mikroarrays erreicht. So lassen sich Gengruppen identifizieren, aus deren Expressionsmuster individuelle Prognosen für die Patientinnen sowie Aussagen über eine spätere Metastasierung abgeleitet werden können¹¹⁻¹³. Eine detailliertere Darstellung dieser Untersuchungen findet sich in Abschnitt 1.6.5.

1.1.2 Therapie des Mammakarzinoms

Therapie lokal begrenzter Karzinome

Mittel der Wahl ist wie bei allen soliden Tumoren die chirurgische Entfernung der Neoplasie. 70% der Patientinnen werden heute brusterhaltend operiert. Gleichzeitig werden axilläre oder, bei prognostisch günstigeren Karzinomen, nur der erste drainierende, so genannte Wächter-Lymphknoten zur Bestimmung des Nodalstatus entfernt. Nach der Operation wird für alle Patientinnen eine Bestrahlung zur Reduktion des Lokalrezidivrisikos empfohlen⁷.

Für die meisten Patientinnen kommt daneben eine adjuvante Hormon- oder Chemotherapie in Frage, um Mikrometastasen oder vermutete Tumorreste zu beseitigen. Die Therapieentscheidung ist dabei individuell und wird durch die Heterogenität der Erkrankung zusätzlich verkompliziert. Zu beachten sind neben TNM-Status, Grading und Tumortyp das Alter bzw. der Menopausenstatus der Patientin sowie der Rezeptorstatus des Tumors. Patientinnen mit negativem Estrogenrezeptorstatus sprechen eher auf eine Chemotherapie an, wohingegen Frauen mit estrogenrezeptorpositiven Tumoren von endokrinen Therapien profitieren. Dazu gehören neben selektiven Estrogenrezeptormodulatoren wie Tamoxifen und Raloxifen Aromatasehemmer wie Anastrozol und GnRH-Superagonisten. Letztere eignen sich nur für prämenopausale Patientinnen mit intakter Ovarfunktion. Wird eine Tamoxifentherapie begonnen, sollte nach derzeitigem Therapiestandard nach fünf Jahren auf einen Aromatasehemmer umgestellt werden¹⁴. Eine fünfjährige Tamoxifentherapie reduziert die jährliche Todesrate um 31%⁴. Bei einer Kombination von zytostatischen und endokrin wirksamen Substanzen sollte die Hormontherapie im Anschluss an die Chemotherapie durchgeführt werden. Exakte Risikoreduktionen für diese kombinierten Therapien sind jedoch noch nicht bekannt. Die empfohlenen Zuordnungen der Patientinnen zu entsprechenden Therapien nach der Konsensuskonferenz in Sankt Gallen 2003 sind in Tabelle 1.2 dargestellt.

Die Standardchemotherapie besteht für Patientinnen mit niedrigerem Risiko derzeit aus Cyclophosphamid, Methotrexat und 5-Fluorouracil (CMF-Schema) oder Cyclophosphamid und Anthracyclin (Doxorubicin oder Epirubicin; AC- bzw. EC-Schema). Bei höherem Risiko werden anthracyclinhaltige Schemata mit mindestens drei Substanzen oder Taxane in Zweitlinien-Therapien eingesetzt, da sie dem CMF-Schema überlegen zu sein scheinen. Hochdosischemotherapien mit autologer Stammzelltransplantation werden im Moment generell nicht empfohlen⁷.

Tabelle 1.2 Sankt Gallen-Einteilung adjuvanter Therapien 2003

Risikogruppe	hormonrezeptorpositiv		hormonrezeptor-negativ
	prämenopausal	postmenopausal	
niedrig: T ₁ , N ₀ , G ₁ , HR ^a , über 35 Jahre	Tamoxifen oder keine Therapie	Tamoxifen oder keine Therapie	-
alle anderen Konstellationen, über 35 Jahre	GnRH-Superagonisten + Tamoxifen oder Chemotherapie + Tamoxifen	N ₀ : Tamoxifen oder Chemotherapie + Tamoxifen N ₁ : Chemotherapie + Tamoxifen	Chemotherapie
alle anderen Konstellationen, unter 35 Jahre	Chemotherapie + GnRH- Superagonisten + Tamoxifen	-	Chemotherapie

^a: Hormonrezeptor positiv

Die derzeit verwendeten Therapien reduzieren die Rezidivrate nach 15 Jahren um 23%, die Mortalitätsrate um 17%. Anthracyclinbasierte Dreierkombinationen mit 5-Fluorouracil und Cyclophosphamid reduzieren im Vergleich zu CMF-Therapien die Rezidivrate nochmals um 11%, die Mortalitätsrate um 16%. Die größten Unterschiede treten für alle Therapien vor allem bei jüngeren Patientinnen sowie in den ersten fünf Jahren nach Diagnosestellung auf, danach sind die Verläufe für therapierte und unbehandelte Patientinnen gleich. Je nach Alter profitieren von der Behandlung letztlich jedoch nur drei bis zehn Prozent der Patientinnen, unabhängig von allen tumorassoziierten Risikofaktoren⁴. Vor allem die große Gruppe der älteren Frauen und solche mit Tumoren mit schlechter Prognose, die besonders auf effektive Therapien angewiesen wären, sind daher benachteiligt. Die Bestrebungen, Patientinnen oder Tumortypen zu identifizieren, die von einer adjuvanten Therapie profitieren werden und warum, sind nach wie vor groß, um Unter- aber auch Übertherapien zu vermeiden. Außerdem fehlen offensichtlich neue Angriffspunkte für Arzneimittel, die eine effektivere Therapie erlauben würden.

Therapie bei eingetretener Metastasierung

Bei manifester Metastasierung haben sämtliche Therapien nur noch palliativen Charakter, da eine Heilung metastasierter Mammakarzinome derzeit nicht möglich ist. In Frage kommen alle therapeutischen Optionen lokal begrenzter Karzinome wie Operation, Radio- oder Chemotherapie, soweit der Nutzen die Belastung für die Patientin übersteigt⁷. Vinca-Alkaloide und Trastuzumab besitzen bisher nur eine Zulassung zur Therapie bei eingetretener Metastasierung¹⁵.

1.2 Die minimal residuale Krebserkrankung (MRD)

1.2.1 Begriffsbestimmung

Definiert wurde der Begriff der minimal residualen Erkrankung ursprünglich für Leukämien und Lymphome. Er bezeichnet das Vorhandensein weniger Tumorzellen nach einer Therapie mit kurativer Absicht^{16, 17}. Bei soliden Tumoren wurden als Äquivalent zur MRD der Leukämien und Lymphome

einzelne disseminierte Zellen epithelialen Ursprungs im Knochenmark der Patienten erstmals in den achtziger Jahren beschrieben¹⁸⁻²⁰. Diese Zellen unterscheiden sich von den mesenchymalen Zellen des Knochenmarks durch die Expression epithelialer Marker, die mittels immunocytochemischer Färbungen oder PCR nach reverser Transkription der mRNA (RT-PCR) nachgewiesen werden können²⁰⁻²³. Der Immunocytochemie sollte wegen geringerer Raten falschpositiver Ergebnisse und der Möglichkeit zu weiterführenden Analysen der Vorzug gegeben werden²⁴. Die Detektion epithelialer Zellen im Knochenmark von Patientinnen mit Mammakarzinom ist seit den neunziger Jahren unter der Bezeichnung „M_i“ Teil der TNM-Klassifikation^{25,26}.

1.2.2 Prognostische Bedeutung der MRD

Die MRD ist bei soliden Tumoren kein seltenes Ereignis. Wiederholt wurde gezeigt, dass auch nach Resektion des Primärtumors und adjuvanter Therapie disseminierte Tumorzellen im Knochenmark des Patienten verbleiben²⁷⁻³¹. Die am häufigsten eingesetzten Nachweismethoden für diese Zellen beruhen auf der Detektion epithelialer Marker wie Zytokeratin 8, 18 oder 19 oder dem Epithelialen Adhäsionsmolekül EpCAM. Beim Mammakarzinom finden auch brustspezifische Marker wie Mammaglobin A zum Nachweis Verwendung³². In der überwiegenden Mehrheit der Studien korreliert das Auftreten einer MRD mit einer schlechteren Prognose hinsichtlich krankheitsfreiem und Gesamtüberleben^{30, 33-36}. Die Datenlage zur unabhängigen prognostischen Bedeutung der MRD ist jedoch uneinheitlich. Zurückzuführen ist dies einerseits auf die Verwendung unterschiedlicher Marker zur Detektion epithelialer Zellen im Knochenmark der Patienten. Andererseits stellen disseminierte Tumorzellen ein sehr seltenes Ereignis dar, das je nach verwendeter Analyseverfahren gefunden oder übersehen werden kann. Studien mit den Markern MUC-1 oder Zytokeratin 18 (Antikörper CK2) an Knochenmarkaspiraten beschreiben beispielsweise die MRD oft nicht als unabhängigen prognostischen Faktor im Vergleich zu Tumorgröße oder Lymphknotenstatus^{37, 38}. Untersuchungen mit dem Panzytokeratinantikörper A45-B/B3 und anderen Markern zeigen hingegen, dass die MRD einen unabhängigen prognostischen Faktor darstellen kann, was die Heterogenität disseminierter Tumorzellen und den Einfluss des Nachweisverfahrens demonstriert^{28, 34, 39}.

1.3 Das Epitheliale Adhäsionsmolekül EpCAM

Neben den Zytokeratinen ist das Epitheliale Zelladhäsionsmolekül (Epithelial Cell Adhesion Molecule; EpCAM) ein weiteres oft verwendetes Molekül zur Detektion disseminierter Tumorzellen^{20, 40}. Auch für diesen Marker konnte ein unabhängiger prognostischer Zusammenhang mit dem Gesamtüberleben bei zahlreichen Tumoren gezeigt werden. Daten für das Mammakarzinom fehlen jedoch bisher⁴¹⁻⁴³.

Humanes EpCAM ist ein hoch konserviertes Glykoprotein mit einem Molekulargewicht von 40 kDa. Die 314 Aminosäuren verteilen sich auf eine große Extrazellulardomäne sowie auf kurze

Transmembran- und Intrazellulärbereiche. Das Protein wird fast ausschließlich von einfachen, vor allem glandulären, sowie pseudo-stratifizierten und transitorischen Epithelien exprimiert. Daher findet sich eine physiologische starke Expression in adulten Geweben vor allem in der Mucosa des Intestinaltraktes, der Bronchien und duktalem Gewebe verschiedener Drüsen. In embryonalen Geweben ist eine EpCAM-Expression jedoch weit verbreitet, das Molekül scheint eine bedeutende Rolle bei der Entwicklung und Differenzierung zu spielen. Auch Vorläuferzellen, die an der Leberregeneration bei Erwachsenen beteiligt sind, zeigen eine EpCAM-Expression und verlieren diese mit zunehmender Differenzierung⁴⁴.

Die physiologische Funktion des Moleküls liegt vermutlich in der Vermittlung homotypischer Zell-Zell-Adhäsion, die auch unter Beteiligung des Zytoskeletts verläuft. Die EpCAM-vermittelten Bindungen sind dabei deutlich schwächer als die anderer Adhäsionsmoleküle wie Cadherine oder Integrine. Cadherin-Bindungen werden durch EpCAM negativ reguliert. EpCAM-vermittelte Bindungen verleihen Geweben also Plastizität, die für Entwicklungsvorgänge notwendig ist.

Die Tumorigenese der meisten Karzinome geht mit einer starken EpCAM-Expression einher. Dabei ist die verstärkte oder *de novo*-Expression mit einer zunehmenden Proliferation und abnehmenden Differenzierung korreliert. EpCAM ist daher zum einen ein früher Marker der Tumorentstehung, zum anderen ein Anzeichen für erhöhten Malignitätsgrad. Beim Mammakarzinom wurden 100-fach bis 1000-fach erhöhte Expressionslevel gefunden⁴⁵. Sie sind als unabhängiger Risikofaktor mit einer schlechteren Prognose korreliert⁴⁶. Auf diesen Erkenntnissen beruhen Versuche, anti-EpCAM Antikörper in der Krebstherapie zu verwenden. Wegen der physiologischen Expression eignen sich nur niedrigaffine Antikörper, bei denen die Nebenwirkungsrate tolerierbar bleibt⁴⁴. Erfolge konnten in der adjuvanten Therapie vor allem beim Kolon- und Mammakarzinom in der MRD erzielt werden, da solide Tumore für Antikörper schlecht zugänglich sind⁴⁷⁻⁵⁰. Für die Indikation „Kolo-rektales Karzinom, Stadium Dukes C“ war ein monoklonaler Antikörper gegen EpCAM als Monotherapie bis 2000 zugelassen (Edrecolomab; 17-1A, Panorex[®]). Dann wurde die Zulassung wegen fehlender Verbesserungen im Vergleich zur Standardtherapie widerrufen. Weitere Studien mit anderen Antikörpern oder in Kombination mit Chemotherapie wurden und werden durchgeführt. Die Ergebnisse sind vielversprechend^{51, 52}.

EpCAM ist jedoch nicht nur als eigenständiges Zielmolekül einer adjuvanten Therapie von Interesse sondern dient auch der Detektion disseminierter Tumorzellen in mesenchymalen Organen zu deren weiterer Charakterisierung in der Grundlagenforschung. Hierbei ist EpCAM unter anderem von Bedeutung, da es im Gegensatz zu den intrazellulären Zytokeratinen ein Oberflächenprotein ist. Eine Färbung ist daher auch an nichtpermeabilisierten lebenden Zellen möglich, was eine unabdingbare Voraussetzung für Genexpressionsanalysen ist. Es konnte allerdings gezeigt werden, dass nicht alle Zytokeratin-positiven disseminierten Tumorzellen auch EpCAM exprimieren und umgekehrt⁵³, was ein weiteres Indiz für die Heterogenität disseminierter Tumorzellen ist^{54, 55}.

1.4 Molekulare Analyse disseminierter Tumorzellen

1.4.1 Charakteristika disseminierter Tumorzellen

Je nach verwendeter Methode und Tumorstadium werden bei 20-60% der Patientinnen mit Mammakarzinom epitheliale Zellen im Knochenmark detektiert^{27, 29, 33}. Meist werden dabei eine bis zehn Zellen pro 10⁶ Knochenmarkszellen gefunden. Geeignete Verfahren zur Anreicherung dieser seltenen Zellen haben daher für die Grundlagenforschung einen hohen Stellenwert. Neben Dichtegradientenzentrifugation werden vermehrt magnet- oder fluoreszenzbasierte immunologische Verfahren eingesetzt, um die Detektionsraten zu erhöhen^{53, 56, 57}.

Um einerseits den Ursprung der epithelialen Zellen im Knochenmark aus dem Tumor zu beweisen und andererseits diese Zellen näher zu charakterisieren, wurden zahlreiche Verfahren zur weiteren Untersuchung der Zellen angewandt. Beispielsweise konnten durch Oberflächenfärbungen und Fluoreszenz-*in situ*-Hybridisierung Amplifikationen von Onkogenen wie HER2 beim Mammakarzinom^{58, 59} oder mittels PCR Mutationen von *KI-RAS* bei gastrointestinalen Tumoren nachgewiesen werden⁶⁰. Zelllinien, die z.T. durch SV40-Transfektion aus disseminierten Tumorzellen gewonnen wurden, zeigten genomische Aberrationen und tumorigenes Potenzial^{61, 62}. Wichtig für die nähere Charakterisierung disseminierter Tumorzellen war die Entwicklung von Techniken zur Isolierung und molekularbiologischen Untersuchung einzelner Zellen^{63, 64}. Durch die sonst unvermeidliche Kontamination der disseminierten Zellen mit mesenchymalen Zellen des Knochenmarks oder die Undurchführbarkeit vieler Verfahren mangels ausreichender Mengen Ausgangsmaterial sind eingehende Analysen dieser Zellen sonst erschwert oder unmöglich.

1.4.2 Untersuchung von disseminierten Tumorzellen beim Mammakarzinom

Bisherige Analysen stützen sich auf Immuncytochemie und Analysen des Genoms. Zielmoleküle der Immuncytochemie waren vor allem Proteine, die in der Diagnose und Prognose des Primärtumors eine Rolle spielen sowie Therapieentscheidungen beeinflussen. Sowohl Studien zum uPA-Rezeptor⁶⁵ als auch zum Estrogenrezeptor⁶⁶ fanden nur geringe Übereinstimmung zwischen Primärtumor und disseminierten Zellen, und zwar sowohl hinsichtlich Qualität als auch Quantität der Expression. Ähnliche Ergebnisse ergaben genetische Untersuchungen. Disseminierte Zellen von Patientinnen ohne manifeste Metastasen unterschieden sich in ihren genomischen Aberrationen deutlich vom jeweiligen Primärtumor, wohingegen sich die Muster bei Patientinnen mit soliden Metastasen ähnelten⁶⁷. Auch stellte sich die Situation für Amplifikationen des *HER2*-Gens ähnlich dar wie bei den Proteinexpressionsstudien zum Estrogen- und uPA-Rezeptor. Hier kann ebenfalls von den Ergebnissen des Primärtumors nicht auf die Situation bei den disseminierten Zellen geschlossen werden⁶⁸.

Wie aus diesen Untersuchungen mit therapierelevanten Molekülen deutlich wurde, können sich disseminierte Tumorzellen als die eigentlichen Ziele adjuvanter Therapien vom Primärtumor deutlich unterscheiden. Die Ergebnisse werfen die Frage auf, inwieweit Analysen am Primärtumor geeignet

sind, prädiktive Aussagen für einzelne Patientinnen zu treffen sowie neue Zielstrukturen für Therapien zu definieren und verdeutlichen die Notwendigkeit einer detaillierteren Charakterisierung der disseminierten Zellen.

1.5 Methoden zur Untersuchung des Transkriptoms von Einzelzellen

Eine einzelne Zelle enthält etwa 0,1-6 pg mRNA, was 1-5% der Gesamt-RNA der Zelle ausmacht. Einige wenige Transkripte liegen in sehr hohen Kopienzahlen vor, die Masse jedoch nur mit 10-15 Kopien pro Zelle⁶⁹. Methoden zur Untersuchung der Genexpression in Einzelzellen müssen daher sehr sensitiv sein. Meist ist vor der eigentlichen Analyse eine Amplifikation der revers transkribierten mRNA erforderlich. Dies kann entweder durch so genannte lineare Verfahren, die auf *in vitro*-Transkription basieren^{70, 71}, oder exponentiell durch PCR geschehen^{64, 72}. Bei allen Amplifikationsverfahren stellt sich die Frage, ob ursprünglich vorhandene Transkriptverhältnisse beibehalten werden oder bestimmte cDNAs aufgrund ihrer Sequenz oder Länge bevorzugt amplifiziert werden⁷³⁻⁷⁵.

Aus revers transkribierter und gegebenenfalls amplifizierter cDNA können Genexpressionsanalysen auf verschiedene Weise durchgeführt werden. Zum einen kann mittels konventioneller oder quantitativer PCR die Expression individueller Transkripte bestimmt werden. Aus Einzelzellen sind ohne vorherige Amplifikation etwa zehn Bestimmungen möglich, da 95% aller Transkripte nur in dieser Anzahl pro Zelle existieren und die Probe daher bestenfalls in zehn identische Ansätze aufgeteilt werden kann. Auch nach einer Amplifikation ist die einzelne Analyse vieler Transkripte für größere Fragestellungen wegen der Vielzahl durchzuführender PCR-Reaktionen jedoch nicht praktikabel. Northern Blots als zweites klassisches Verfahren sind aus Einzelzellen wegen der geringen RNA-Mengen ebenfalls nicht durchführbar. Für umfassende Analysen eignet sich somit nach einer Amplifikation der cDNA die Genexpressionsanalyse mittels Arrayhybridisierung.

1.6 Genexpressionsanalyse mittels Arrayhybridisierung

1.6.1 Historische Entwicklung

Technologien zur Analyse der Genexpression mittels Arrays bieten den Vorteil der gleichzeitigen Bestimmung zahlreicher Transkripte. Anfangs wurden, abgeleitet vom Southernblot, cDNA-Sonden auf positiv geladene Nylonmembranen aufgebracht. Die Detektion nach der Hybridisierung erfolgte mittels radioaktiv markierter Proben oder durch Enzymreaktionen unter Emission von Licht, wodurch Röntgenfilme geschwärzt werden. Die Stärke der Schwärzung wird mit der gebundenen Menge Probe korreliert. Dieses Verhältnis ist jedoch nicht linear, was eine echte Quantifizierung mit solchen Systemen unmöglich macht. Außerdem neigen Sonden aus klonierter oder PCR-amplifizierter cDNA wegen ihrer Länge und Präparation in bakteriellen Systemen relativ stark zu Kreuzhybridisierungen.

Um nicht-radioaktive, echte Quantifizierungen zu ermöglichen, wurden fluoreszenzbasierte Detektionsverfahren entwickelt. Die ersten Glasarrays, die Fluoreszenz zur Detektion benutzen, wurden Mitte der neunziger Jahre beschrieben⁷⁶⁻⁷⁸. Das Problem der Kreuzhybridisierungen wurde durch verändertes Sondendesign angegangen. Transkriptomarrays neuerer Generation benutzen keine langen cDNA-Sonden mehr sondern relativ kurze Oligonukleotide unterschiedlicher Länge. Die Sonden werden spezifisch für einzelne Transkripte entworfen und erlauben je nach Lokalisation innerhalb des Transkripts z.B. auch eine Unterscheidung von Genen einer Familie mit hoher Sequenzhomologie oder von Splicevarianten. Die ersten kommerziellen Oligonukleotidarrays wurden in den neunziger Jahren von der Firma Affymetrix auf den Markt gebracht. Diese Arrays bestehen aus etwa zehn bis 15 Sonden von 25 Nukleotiden Länge pro Gen. Hinzu kommen so genannte „Mismatch-Oligos“, die der Korrektur unspezifischer Hybridisierungen dienen. Eine unabdingbare Voraussetzung für das Affymetrix-System ist das Vorhandensein eines T7-Promotors im cDNA-Template, der üblicherweise während der cDNA-Synthese über den Primer eingefügt wird. Über diesen Promotor erfolgt die Markierungsreaktion im Rahmen einer *in vitro*-Transkription. Hybridisiert wird cRNA (für copy), was zur Bildung thermodynamisch sehr stabiler DNA/RNA-Hybride mit den Sonden des Arrays führt. Bei geringen Mengen Ausgangs-RNA kann das Produkt der *in vitro*-Transkription erneut revers transkribiert und erst dann endgültig markiert werden, um die Amplifikation zu verstärken. Alle diese Verfahren beruhen letztlich auf dem so genannten „Eberwine-Protokoll“⁷⁰ bzw. dessen Modifikation⁷¹ und werden nach wie vor nicht für RNA-Mengen aus weniger als 100 Zellen empfohlen (www.affymetrix.com; Jan. 2006). Zur Untersuchung von Einzelzellen scheinen sie nicht geeignet^{79,80}.

Die Entwicklung weiterer Oligonukleotidarrays führte zur Verwendung längerer Sonden. Als optimal stellten sich Moleküle mit einer Länge von 70 Basenpaaren, so genannte Siebzigmere heraus⁸¹. Diese Arrays mit einem einzelnen längeren Oligonukleotid pro Gen sind mittlerweile in vielen Varianten verwendet und auch validiert worden⁸²⁻⁸⁴. Die Ergebnisse der Oligonukleotidarrays sind dabei den cDNA-Arrays hinsichtlich Übereinstimmung zu Northern Blot, SAGE oder quantitativer PCR überlegen⁸⁵⁻⁸⁷. Long Oligonucleotide-Arrays werden mittlerweile von verschiedenen Herstellern in zahlreichen Varianten angeboten. Sie sind meist preisgünstiger und flexibler als das System der Firma Affymetrix. Zwischen einzelnen Plattformen und Labors können jedoch starke Unterschiede auftreten, was die Notwendigkeit einer guten Validierung, Planung der Experimente und vor allem auch bioinformatischen Auswertung der Daten verdeutlicht⁸⁸⁻⁹³.

1.6.2 Technologie der Hybridisierung

Sämtliche Hybridisierungstechniken beruhen auf der Eigenschaft von Nukleinsäuren, aus zwei komplementären Einzelsträngen einen Doppelstrang ausbilden zu können. Bei Arrayhybridisierungen handelt es sich zudem um ein zweiphasiges System: Ein Einzelstrang wird als Sonde immobilisiert (feste Phase), der komplementäre Einzelstrang befindet sich in der Probe (mobile Phase). Mehrere Sonden, die nach einem bekannten Muster auf einem Träger angeordnet sind, bilden den Array.

Die Hybridisierung selbst ist ein thermodynamischer Prozess. Einfluss auf das Ergebnis nimmt neben der Temperatur auch die Konzentration von Probe, Salzen und anderen Zusätzen wie Formamid in der Hybridisierungslösung und den nachfolgenden Waschpuffern, da die Schmelztemperatur T_m eines DNA-Doppelstranges von diesen Faktoren abhängt. Sie kann für Doppelstränge von über 50 bp Länge nach folgender Formel berechnet werden:

$$T_m = 81,5 + (41 * (yG+zC)/(wA+xT+yG+zC)) - (500/(wA+xT+yG+zC)) + 16,6 * \log_{10}([Na^+]) - 0,62F,$$

wobei w , x , y und z die Anzahl der jeweiligen Nukleotide und F der prozentuale Gehalt an Formamid sind. Da die Natriumionenkonzentration den größten Einfluss auf die Stabilität von DNA-Doppelsträngen besitzen, können andere Kationen vernachlässigt werden⁹⁴.

Über die Stringenz der Waschschritte im Anschluss an die Hybridisierung kann deren Spezifität maßgeblich beeinflusst werden. Die Stringenz nimmt bei steigender Temperatur, abnehmender Salzkonzentration und durch Tenside zu. Die Bedingungen für Hybridisierung und Waschschritte sind so einzustellen, dass bei möglichst hoher Sensitivität eine ausreichende Spezifität erreicht wird. Unspezifische Kreuzhybridisierungen müssen unterdrückt oder durch Waschen wieder aufgehoben werden, spezifische Bindungen sollen erhalten bleiben.

Neben Quantität und Qualität der hybridisierten cDNA nimmt aber auch der Array Einfluss auf das Ergebnis. Je nach verwendeter Sonde werden grundsätzlich Oligonukleotid- und cDNA-Arrays unterschieden. Ausreichende Mengen dieser Sonden müssen fest an der Glasoberfläche verankert werden, was abhängig von der Oberflächenbeschichtung des Objektträgers durch elektrostatische oder kovalente Bindungen geschehen kann. Weitere Einflussfaktoren auf die Hybridisierungseigenschaften liegen in den Bedingungen bei der Herstellung der Arrays wie beispielsweise Puffer, Luftfeuchtigkeit oder Nachbehandlung, die jeweils optimiert werden müssen.

1.6.3 Design von Mikroarrayexperimenten

Eine zentrale Fragestellung ist die Auswahl eines geeigneten Arrays. Es werden so genannte „General Purpose Arrays“ von solchen mit eingeschränkter Verwendung unterschieden. Die Allzweck-Arrays enthalten normalerweise alle bekannten Gene und/oder Expressed Sequence Tags (ESTs) einer Spezies. Sie eignen sich für datengetriebene Fragestellungen, bei denen nicht im Voraus bekannt ist, welche Gene von Interesse sein werden. Arrays mit eingeschränkter Verwendung enthalten vorausgewählte Teile der General Purpose Arrays hinsichtlich Genfunktion, Erkrankung oder Art der

Sequenz. Es muss vor der Durchführung des Experimentes bereits eine Hypothese vorliegen, die mit den Genen des Arrays getestet werden kann. Dafür sind solche Experimente wegen der reduzierten Genzahl meist preisgünstiger und leichter auszuwerten. Für beide Arten von Arrays existieren Plattformen zahlreicher Hersteller mit entsprechenden Varianten, Vor- und Nachteilen, von denen die geeignetste ausgewählt werden muss.

Des Weiteren wird zwischen Ein- und Zweifarbentechnologien unterschieden. Bei Zweifarbenhybridisierungen wird neben der eigentlichen Probe eine mit einer anderen Farbe markierte Referenz kohybridisiert. Das Verfahren dient der Normalisierung und erlaubt eine interne Standardisierung der Ergebnisse, wenn für alle Experimente eine universelle Referenz verwendet wird. Meist werden dazu Mischungen von cDNA aus verschiedenen Geweben eingesetzt, da in der Referenz möglichst viele Gene exprimiert sein sollen^{89, 95}. Einen Überblick über das Vorgehen bei einem Zweifarbenexperiment gibt Abbildung 1.1.

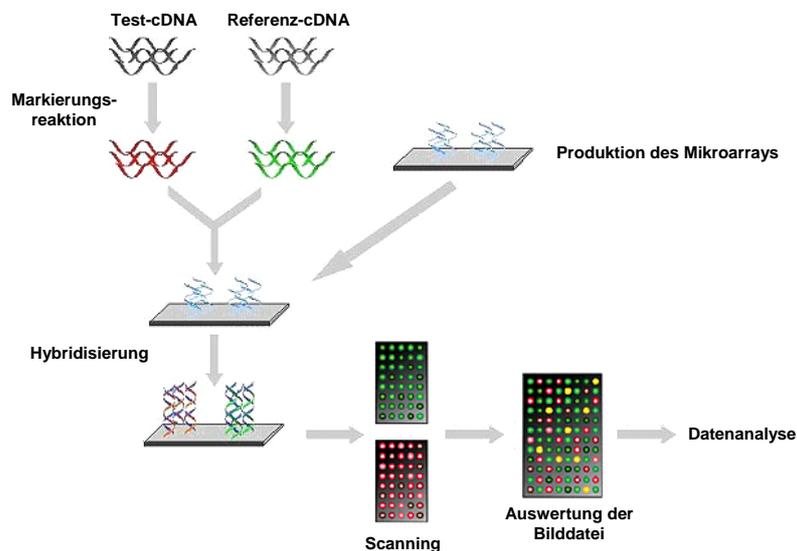


Abbildung 1.1 Das Prinzip der Zweifarbenhybridisierung auf Glasarrays.

Getrennt isolierte und mit unterschiedlichen Farben markierte Test- und Referenz-cDNA werden gemeinsam auf einen Array hybridisiert. Nach anschließenden Waschschrinen zur Steuerung der Spezifität werden beide Farben getrennt gescannt. Die überlagerte Bilddatei wird in eine Tabelle mit Intensitätswerten überführt und analysiert, wobei die Referenz-cDNA der Standardisierung dient.

Die Anzahl der benötigten Replikate des Experiments kann über eine Trennschärfenanalyse berechnet werden. Eingang in die Berechnung finden neben der Anzahl der untersuchten Gruppen und der technischen und biologischen Variabilität des Verfahrens auch das angestrebte Signifikanzniveau sowie die Größe der zu beobachtenden Unterschiede. Biologischen Replikaten sollte dabei wegen der Abdeckung von biologischer und technischer Variabilität der Vorzug vor technischen Replikaten gegeben werden^{96, 97}. Wichtig ist auch die sorgfältige Auswahl geeigneter bioinformatischer Methoden zur Auswertung der erhobenen Rohdaten⁹⁸.

1.6.4 Bioinformatische Methoden zur Auswertung von Genexpressionsarrays

General Purpose Arrays zur Analyse der Genexpression enthalten alle bekannten Gene und/oder ESTs einer Spezies. Die Anzahl gleichzeitig bestimmter Moleküle erreicht daher Größenordnungen einiger Zehntausende. Das humane Oligoset Version 4.0 zur Herstellung eines Arrays der Firma Operon umfasst beispielsweise 35035 Sonden, die 25100 humane Gene und 39600 Transkripte abdecken. Aus diesen Zahlen wird deutlich, dass bei der Verwendung dieser Systeme sehr große Mengen an Daten erhoben werden. Gängigerweise wird davon ausgegangen, dass mindestens 80% der gemessenen Gene zwischen verschiedenen Gruppen eines Experiments nicht reguliert sind. Von den verbleibenden 20% gilt es diejenigen zu finden, die tatsächlich durch die experimentellen Bedingungen reguliert werden und für die jeweilige Fragestellung von Interesse sind. Die Herausforderung bioinformatischer Methoden zur Auswertung von Genexpressionsarrays liegt also in der Extraktion der wenigen relevanten Effekte aus der Masse nicht veränderter Expressionswerte und in der Abtrennung von Rauschen und zufälliger Streuung⁹⁹.

Zur Lösung dieser Aufgaben ist ein mehrstufiges Verfahren gebräuchlich. Zuerst werden systematische Fehler, die sich aus dem Verfahren ergeben, durch Normalisierung entfernt. Dazu gehören zum Beispiel Abhängigkeit eines Messwertes von der Intensität bedingt durch nichtlineare Systeme in Anregung und Messung von Fluoreszenzen. Für die Zweifarbentechnologie ist als fortschrittlichstes Verfahren die LOESS-Normalisierung etabliert. Dabei wird eine lokal gewichtete polynomiale Regression durchgeführt, die unter anderem räumliche und intensitätsabhängige Effekte auf den Arrays entfernt^{100, 101}. Um in weiterführenden statistischen Verfahren die Anzahl falsch-positiver Resultate zu minimieren, muss neben speziellen Verfahren zur Reduzierung dieser so genannten „False Discovery Rate“ die Anzahl zu testender Gene reduziert werden^{102, 103}. Dazu werden unzuverlässige Messwerte und uninteressante Gene aus den Listen entfernt. Beispielsweise eignet sich eine Beschränkung auf regulierte Gene, die mindestens einen bestimmten Messwert erreichen. Die Grenzen für diese Kriterien müssen zuvor festgelegt werden^{103, 104}.

Mit den aus diesem Prozess resultierenden so genannten „gefilterten“ Genlisten lassen sich die eigentlichen Verfahren zur Auswertung durchführen. Meist wird dabei nach Ähnlichkeiten in den Expressionsprofilen zwischen verschiedenen Proben sowie nach Unterschieden in Abhängigkeit vom Zustand der Probe gesucht. Gängige Verfahren umfassen neben klassischen statistischen Methoden wie t-Test oder ANOVA auch Clusteranalysen und komplexere überwachte und nichtüberwachte Verfahren¹⁰⁵⁻¹⁰⁷. Dazu gehören hierarchisches, k-means und Quality-threshold Clustering, selbstorganisierende Karten, Support Vector Machines, Principal Component Analyse und viele andere¹⁰⁸⁻¹¹¹. Das Feld der Bioinformatik ist nach wie vor stark in der Entwicklung begriffen. Standardvorgehensweisen haben sich noch nicht durchgesetzt, was die Vergleichbarkeit der Ergebnisse verschiedener Studien nur unter Einbeziehung der verwendeten Verfahren erlaubt.

1.6.5 Genexpressionsanalysen beim Mammakarzinom

Da die Überlebensraten beim Mammakarzinom in den letzten Jahren nicht verbessert werden konnten und auch viele Patientinnen trotz adjuvanter Therapie Rezidive oder Metastasen erleiden, wird mit molekularbiologischen Methoden versucht, die Karzinome weiter zu klassifizieren. Ziel ist eine bessere Abschätzung des individuellen Risikos sowie der voraussichtlichen Effizienz adjuvanter Therapien. Seit kurzem werden dazu auch Genexpressionsanalysen der Primärtumore auf General Purpose-Arrays durchgeführt. Aus den gewonnenen Rohdaten werden so genannte „Signaturen“ der relevanten Gene erstellt. Anhand der Expressionsniveaus der Gene einer Signatur sind Aussagen über spätere Metastasierung, Effizienz von Therapien und Überleben möglich. Beispielsweise kann das Therapieansprechen auf Grundlage solcher Daten mit einer Sensitivität von 80% bei einer Spezifität von 90% vorausgesagt werden¹¹², die Prädiktion einer späteren Metastasierung bei Patientinnen ohne Lymphknotenbefall gelingt mit einer Sensitivität von über 90% und einer Spezifität von bis zu 73%^{13,113}.

Die meisten Signaturen umfassen etwa 70 bis 90 Gene und können unabhängige prognostische Faktoren darstellen¹². Probleme treten aber nach wie vor bei der Erstellung der Signaturen auf. Durch das Fehlen standardisierter bioinformatischer Verfahren können je nach verwendeter Routine aus identischen Rohdaten unterschiedliche Signaturen erhalten werden. Jede für sich ist dabei in der Lage, eine Prognose zu erstellen^{114,115}. Außerdem wurden auch für deutlich kleinere Signaturen bis hinunter zu vier Genen Aussagen gleicher Qualität erhalten¹¹⁶. Ursachen dieser Diskrepanzen sind zum einen die Verwendung unterschiedlicher Arrays, Protokolle zur Probenaufarbeitung und bioinformatischer Methoden. Aber auch die oft unzureichende Studiengröße und die Auswahl und Einteilung der Patienten spielen eine wichtige Rolle^{117,118}. So werden beispielsweise sowohl Genexpressionsprofile als auch Therapie und Prognose beim Mammakarzinom stark vom Estrogenrezeptorstatus des Tumors geprägt, was bei der Auswertung von Genexpressionsstudien berücksichtigt werden muss^{12,119}. Bedingt durch solche Faktoren fehlen bisher Überschneidungen zwischen den Ergebnissen verschiedener Studien, und es ist schwierig, einzelne Moleküle für ein gezieltes Screening zu definieren. Größere Metaanalysen und verfeinerte bioinformatische Methoden könnten das Problem jedoch in Zukunft lösen¹²⁰.

1.7 Ursachen gestörter Genexpression bei Tumoren

Von Bedeutung für die Entstehung von Tumoren sind Faktoren, die der Tumorzelle einen Wachstums- und Selektionsvorteil vermitteln. Die zellulären Funktionen und damit die Genexpression unterscheiden sich von gesundem Gewebe. Zu den Ursachen solcher Fehlregulationen gehören genomische Veränderungen, die mit veränderter Expression der aberranten DNA-Abschnitte einhergehen. Außer bei Tumoren wurde dieser Effekt auch bei der Trisomie 21 des Down-Syndroms beobachtet¹²¹⁻¹²⁴. Genomische Veränderungen werden mit der Zellteilung weitergegeben und wirken

sich daher auch auf Tochtergenerationen aus. Neben abweichenden Kopienzahlen von Genen oder DNA-Bereichen durch Amplifikationen oder Deletionen sind auch Translokationen, Polymorphismen und Punktmutationen, die zu einer Störung der Genregulation führen können, von Bedeutung. Alle diese Veränderungen sind durch spezifische Verfahren nachweisbar. Die in dieser Arbeit benutzte Methode der Komparativen Genomischen Hybridisierung (CGH) vergleicht das komplette Genom der Tumorzelle mit einem normalen diploiden Chromosomensatz. Für Einzelzelluntersuchungen wird aus beiden Proben die genomische DNA isoliert, global amplifiziert und mit verschiedenen Farben markiert. Beide Genomrepräsentationen werden gemeinsam auf normale humane Metaphasen hybridisiert und anschließend jede Farbe getrennt digital fotografiert. Mit Hilfe digitaler Datenverarbeitung lassen sich aus den gewonnenen Bildern über Intensitätsunterschiede die relativen Probenmengen im jeweiligen Bereich des Chromosoms berechnen und graphisch darstellen (Abbildung 1.2). Die Methode kann gleichzeitig sämtliche numerische Aberrationen der Tumorzelle abbilden ohne Metaphasen gewinnen zu müssen, nicht jedoch Veränderungen ohne Einfluss auf die Kopienzahl eines DNA-Bereiches und solche unterhalb einer Ausdehnung von etwa 10-20 Mb^{63, 125-127}.

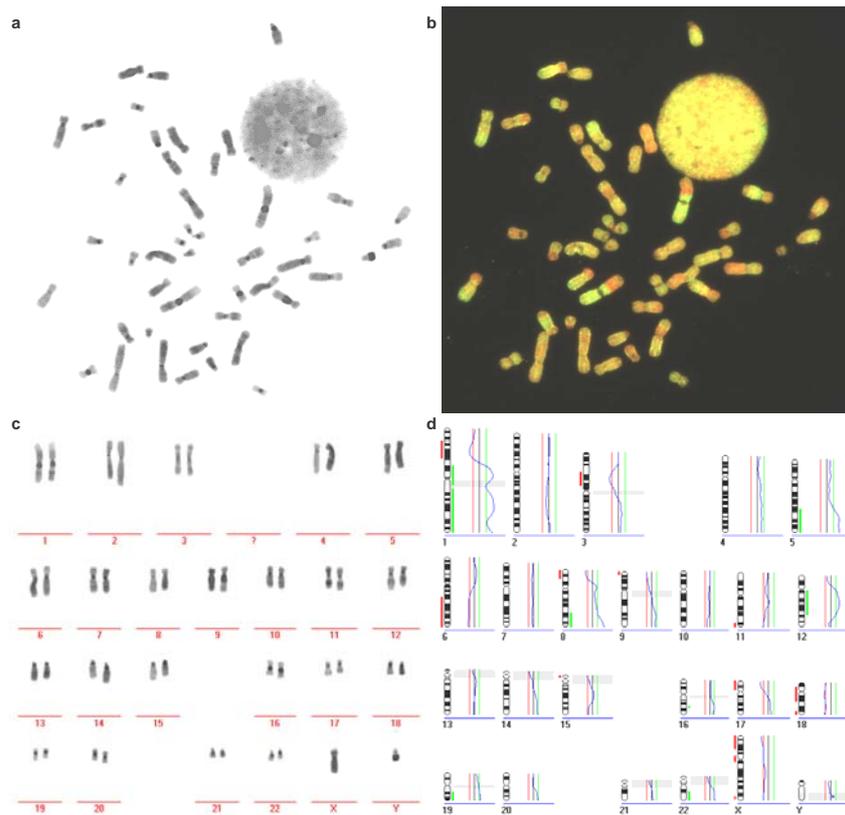


Abbildung 1.2 Das Prinzip der CGH.

Auf Metaphasen eines gesunden Spenders (a) werden grün bzw. rot markierte Test- und Referenz-DNA hybridisiert (b). Nach einer Gegenfärbung mit DAPI können die Chromosomen anhand ihrer Bänderung in ein Karyogramm sortiert werden (c). Das CGH-Profil zur Darstellung chromosomaler Veränderungen wird durch eine Software aus den Intensitätsunterschieden beider Farben berechnet. Signifikante Veränderungen werden durch Balken neben den Chromosomen symbolisiert. Amplifikationen sind dabei durch einen grünen Balken am rechten Rand des Chromosoms dargestellt, Deletionen durch einen roten Balken am linken Rand (d).

In Untersuchungen zeigen allerdings nur etwa 40-60% der stark amplifizierten Gene auch eine Überexpression und etwa 10% der überexprimierten Gene sind auf Genomebene amplifiziert¹²¹⁻¹²³. Als weitere Ursache gestörter Genexpression bei Tumoren spielen daher auch nichtgenomische Faktoren eine Rolle. Dazu gehören epigenetische Veränderungen, die durch Störungen der DNA-Methylierung, des Imprintings und einer Veränderungen der Chromatinstruktur durch Histonmodifikationen Einfluss auf die Genexpression nehmen¹²⁸. Epigenetische Regulation findet auch bei vielen Entwicklungsvorgängen statt, und Ungleichgewichte wurden für zahlreiche tumorrelevante Gene beschrieben. Sie gelten als frühes Ereignis in der Tumorigenese, das genomischen Veränderungen vorausgeht¹²⁹. Epigenetische Veränderungen sind wie nicht-numerische und subchromosomale Aberrationen mit der CGH ebenfalls nicht detektierbar.

1.8 Zielsetzung

Ziel der vorliegenden Dissertation war die Erarbeitung einer Methode zur Untersuchung der Genexpression von disseminierten Tumorzellen auf Oligonukleotidarrays, die das gesamte Genom abdecken. Der Nachweis solcher Zellen korreliert mit einer schlechteren Prognose für die Patienten, gleichzeitig stellen disseminierte Tumorzellen die Ziele adjuvanter Therapien dar. Über die Biologie dieser Zellen ist jedoch noch wenig bekannt, so dass die Untersuchung der Genexpression zum Verständnis der systemischen Krankheitsprogression beitragen sowie wichtige Anhaltspunkte für die Entwicklung neuer Therapien geben kann.

Die Aufgabenstellung der Arbeit gliederte sich in mehrere Teilbereiche. Erstens musste ein bestehendes Verfahren zur Amplifikation von cDNA aus Einzelzellen soweit optimiert werden, dass eine umfassende und quantitative Untersuchung des Transkriptoms auf Genexpressionsarrays möglich wurde. Dazu musste zweitens eine geeignete Plattform ausgewählt und die Hybridisierungsbedingungen etabliert werden. Drittens sollte die Isolierung disseminierter Tumorzellen aus Knochenmarkaspiraten verbessert werden, da mit den Standardverfahren nur sehr wenige Zellen in einem geringen Prozentsatz der Proben gefunden werden. Abschließend sollten die etablierten Techniken an Knochenmarkpunktionen von Patientinnen mit Mammakarzinom angewandt, disseminierte Tumorzellen isoliert und ihr Genom sowie die Genexpression untersucht werden.

2 Material und Methoden

2.1 Methoden

2.1.1 Isolierung und Amplifikation von mRNA mit dem PNA-Protokoll

Pro Ansatz werden 5 µl Lysispuffer, 1 µg Protease, 1 µl biotinkoppelte Oligo-dT-PNAs (mTrap Midi-Kit, rekonstituiert in 400 µl Wasser) und 10 ng tRNA vorbereitet. Sollen die gepickten Zellen bei -80°C gelagert und erst später aufgearbeitet werden, werden nur 4 µl Lysispuffer mit 10 ng tRNA vorgelegt und erst bei der Aufarbeitung PNAs und Protease zugegeben. Einzelne Zellen werden in einem Volumen von 1 µl isoliert und die Proben zum Aufschluss und zur Inaktivierung von RNasen in der PCR-Maschine einem Temperaturprogramm unterworfen (10 min 45°C, 1 min 70°C, 15 min 22°C zum Annealing der PNAs). Dann werden pro Ansatz 4 µl Streptavidin-Beads hinzugefügt und die Proben 45 min bei Raumtemperatur gerollert.

Zu jeder Probe werden 10 µl cDNA-Igepal-Waschpuffer gegeben, die Beads an einem Magnetständer an der Gefäßwand immobilisiert und der Überstand abgenommen. Dieser wird zur Aufarbeitung der genomischen DNA in ein silanisiertes Reaktionsgefäß mit 0,8 µl Polyacrylamid-Carrier überführt (vgl. Absatz 2.1.3). Der Waschschrift wird mit 20 µl cDNA-Tween-Waschpuffer und anschließend mit 20 µl cDNA-Igepal-Puffer wiederholt. Beide Überstände werden ebenfalls gesammelt und mit dem ersten vereinigt. Dann werden die Beads in 2 µl 5x Reverse Transcription-Puffer, 1 µl DTT 0,1 M, 0,5 µl Igepal 10%, 0,5 µl Wasser und 6 µl Primer resuspendiert. Folgende Primerkombinationen werden verwendet:

- 3 µl CFL5CT15 bzw. CFL5CT24 100 µM + 3 µl CFL5C8 200 µM
- 3 µl T7TC/BLOCK 2 50 µM + 3 µl T7N8C/BLOCK 2 100 µM; der Block wird zuvor äquimolar über einen Temperaturgradienten an das Oligonukleotid angelagert (1 sec bei 94°C, dann von 65°C bis 20°C in 1°C-Schritten für jeweils eine Minute).
- 6 µl T7-Primer 50 µM, vergleiche Absätze 2.5 und 3.2.1.

Die Primer werden für 10 min bei Raumtemperatur angelagert, anschließend 2 µl 5x Reverse Transcription-Puffer, 1 µl DTT 0,1 M, 1 µl dNTPs 10 mM, 5 µl Wasser und 1 µl Superscript II Reverse Transcriptase zugegeben und die Reaktionsgefäße 45 min bei 44°C im Ofen rotiert.

Nach der cDNA-Synthese werden die Beads am Magneten in 20 µl Tailingwaschpuffer gewaschen, in 10 µl Tailing-Mix resuspendiert (4 mM MgCl₂, 0,1 mM DTT, 0,2 mM dGTP, 10 mM KH₂PO₄) und der Ansatz mit 40 µl PCR-Öl überschichtet. Eine PCR-Maschine wird auf 94°C vorgewärmt und die Proben darin 4 min denaturiert. Im Anschluss werden sie direkt auf Eis gestellt, 11,2 U Terminal-Deoxynukleotid-Transferase (TdT) zugegeben und die Tailingreaktion bei 37°C für 60 min durchgeführt. Nach Inaktivierung der TdT bei 70°C für 5 min erfolgt die Zugabe von 4 µl Expand

Long Template Puffer 1, 7,5 µl Formamid 20% und 24 µl Wasser. Die Proben werden in die PCR-Maschine gestellt und das LIB-Programm gestartet. Bei Erreichen der 78°C werden 2,4 µl CP2-Primer 24 µM, 1,75 µl dNTPs 10 mM und 1,5 µl Expand Long Template PolMix (5,25 U) zugefügt.

Tabelle 2.1 LIB-PCR-Programm für die cDNA-Primäramplifikation

1.	78°C	0:30 min
2.	94°C	0:15 min
3.	65°C	0:30 min
4.	68°C	2:00 min
5.	zurück zu 2. 19 Mal	
6.	94°C	0:15 min
7.	65°C	0:30 min
8.	68°C	2:30 min + 10 sec/Zyklus
9.	zurück zu 6. 20 Mal	
10.	68°C	7:00
11.	4°C	∞

Das PCR-Produkt ist bei -20°C stabil und dient als Ausgangsmaterial für weitere Untersuchungen.

2.1.2 Isolierung und Amplifikation von mRNA mit dem Oligo dT-Bead-Protokoll

Pro Ansatz werden 10 µl Oligo-dT-Beads (Dynal) zweimal mit 10 µl LiDS-Waschpuffer gemischt und der Überstand am Magneten abgenommen. Die Beads werden in 10 µl Lysispuffer aufgenommen und auf Eis gestellt. Nach Isolation der Zelle werden die Proben 30 min bei Raumtemperatur gerollert. Es folgen die Zugabe von 10 µl cDNA-Igepal-Waschpuffer und alle weiteren Schritte wie für das PNA-Protokoll unter Absatz 2.1.1 beschrieben.

2.1.3 Aufarbeitung der genomischen DNA parallel zur mRNA

Die genomische DNA aus den Überständen der mRNA-Waschschritte (Absatz 2.1.1) wird mit 120 µl 100%igem Ethanol über Nacht bei -20°C gefällt. Die Fällung wird für 45 min bei 14000 rpm und 4°C abzentrifugiert, die Flüssigkeit bis auf 10 µl abgenommen und durch 180 µl 70%igen Ethanol ersetzt. Die Proben werden 10 min bei Raumtemperatur geschüttelt und dann 10 min bei 14000 rpm zentrifugiert. Dieser Waschschritt wird noch zweimal wiederholt, dann wird der Alkohol komplett abgenommen, das Pellet kurz getrocknet und anschließend in 3,48 µl Wasser über Nacht auf dem Schüttler bei Raumtemperatur rückgelöst. Dann werden 0,5 µl OnePhorAll-Puffer (OPA), je 0,13 µl Tween 10% und Igepal 10% sowie 0,26 µl Proteinase K 10 mg/ml zugegeben und der Verdau für 15 h bei 42°C in der PCR-Maschine durchgeführt.

Das Produkt des Proteinase K-Verdau wird mit 0,25 µl *MseI* (12,5 U) und 0,25 µl Wasser bei 37°C für 3 h in der PCR-Maschine verdaut, dann wird das Enzym bei 65°C in 5 min inaktiviert. Während des Verdau wird der Ligationsansatz bestehend aus 0,5 µl OPA-Puffer, jeweils 0,5 µl LIB1- und

ddMse11-Oligonukleotid (beide 100 μ M) und 1,5 μ l Wasser in der PCR-Maschine einem Temperaturgradienten zur Basenpaarung unterworfen (65°C bis 15°C für je 1 min pro Temperaturschritt). Der Ansatz wird zusammen mit 5 U T4 DNA-Ligase und 0,01 mM ATP zum Produkt des *MseI*-Verdau gegeben (Endvolumen 10 μ l) und über Nacht bei 15°C ligiert. Nach der Ligation wird der Ansatz zur Primäramplifikation mit 3 μ l Expand Long Template Puffer 1, 2 μ l dNTPs 10 mM, 35 μ l Wasser und 1 μ l Expand Long Template PolMix (3,5 U) aufgefüllt.

Tabelle 2.2 PCR-Programm für die Primäramplifikation genomischer DNA

1.	68°C	3:00 min
2.	94°C	0:40 min
3.	57°C	0:30 min
4.	68°C	1:30 min + 1sec/Zyklus
5.	zurück zu 2. 14 Mal	
6.	94°C	0:40 min
7.	57°C-65°C	0:30 min + 1°C/Zyklus
8.	68°C	1:45 min + 1 sec/Zyklus
9.	zurück zu 6. 7 Mal	
10.	94°C	0:40 min
11.	65°C	0:30 min
12.	68°C	1:53 min + 1 sec/Zyklus
13.	zurück zu 10. 21 Mal	
14.	68°C	3:40 min
15.	4°C	∞

2.1.4 Quantitative real time PCR

Unter Verwendung von LightCycler-Faststart DNA Master SYBR Green I-Kits wird ein 20 μ l-Ansatz aus 2,0 μ l SYBR-Green/Polymerase-Mix, 3 mM MgCl₂, jeweils 1,6 μ M 5'- und 3'-Primer, 1 μ l Template und Wasser hergestellt. Für die Primer ACE und IL2R werden 2 mM Magnesiumchlorid als Endkonzentration eingesetzt. Die PCR wird im LightCycler nach folgendem Programm durchgeführt:

Tabelle 2.3 PCR-Programm für die real time PCR im LightCycler

Schritt	Temperatur	Zeit	Heizrate	Fluoreszenzmessung
1.	50°C	2:00 min	20°C/min	keine
2.	95°C	10:00 min	20°C/min	keine
3.	95°C	0:10 min	20°C/min	keine
4.	60°C	0:05 min	20°C/min	keine
5.	72°C	0:10 min	20°C/min	ein Mal am Ende
6.	zurück zu 3. 39 Mal			
7.	95°C	0:00 min	20°C/min	keine
8.	65°C	1:00 min	20°C/min	keine
9.	ad 95°C	0:00 min	0,1°C/sec	kontinuierlich
10.	40°C	0:30 min	20°C/min	keine

Für alle Primer werden mit einer Verdünnungsreihe Standardkurven angelegt und aus deren Steigungen Korrekturkoeffizienten für unterschiedliche PCR-Effizienzen ermittelt. In jede folgende PCR wird zudem ein Ansatz mit einer dieser Standardverdünnungen zur Bestimmung der aktuellen Effizienz integriert. Mit Hilfe einer Software wird aus den PCR-Verläufen jeder Probe für das jeweilige Test- und ein Haushalts-Referenzgen nach Effizienz-Korrektur die relative Menge des Testgens im Verhältnis zum Referenzgen ermittelt. Eine Analyse der Schmelzkurven der PCR-Produkte am Ende der Reaktion dient der Kontrolle auf Spezifität.

2.1.5 Reamplifikations- und Markierungs-PCR

Durch den Einbau von Nukleotidderivaten während einer Reamplifikations-PCR werden die Primäramplifikate zur Detektion nach einer Hybridisierung markiert. Test-DNA wird dabei immer Digoxigenin-, Referenz-DNA Biotin-markiert. Als Referenz für genomische Hybridisierungen dient diploide DNA, cDNA-Hybridisierungen auf Oligonukleotidarrays benutzen Pools aus cDNA verschiedener Zelllinien und Gewebe.

Tabelle 2.4 Ansätze für die Markierung von cDNA (links) und genomischer DNA (rechts)

5 µl	Expand Long Template Puffer 1	4 µl	Expand Long Template Puffer 1
7,5 µl	Formamid 20%	8 µl	LIB1 10 µM
5 µl	CP2(-BGL) 24 µM	1,5 µl	7/8 dNTP-Mix
1,75µl	dNTP-Mix, s.u.	1,75 µl	markiertes dUTP 1 mM
x µl	markierte Nukleotide	1 µl	Taq-Polymerase 5 U/µl
0,75 µl	Taq-Polymerase 5 U/µl	24,25 µl	Wasser
ad 49 µl	Wasser	1 µl	Template
1 µl	Template		

Für cDNA-Markierungen werden folgende Nukleotidderivate und dNTP-Mixe verwendet:

- Nylon-Filter: Dig-dUTP alkalilabil 0,05 mM und 1/7 dNTP-Mix
- Oligonukleotidarray; Einfachmarkierung: Dig-dUTP bzw. Bio-dUTP 0,05 mM, 1/7-Mix; Verwendung des CP2-BGL-Primer
- Oligonukleotidarray; Doppelmarkierung: Dig-dUTP und Aminodig-dCTP bzw. Bio-dUTP und Bio-dCTP je 0,05 mM; 2/7-Mix; Verwendung des CP2-BGL-Primers
- Labelling-Tests: Aminoallyl-dUTP in Anteilen von 60-100% am dTTP des Nukleotidmixes; Cy5-dUTP, Cy5-dCTP, Cy3-dUTP, Cy3-dCTP alle 0,05 mM

Tabelle 2.5 Zusammensetzung der dNTP-Mixe für Markierungsreaktionen

7/8-Mix	dATP, dCTP, dGTP alle 10 mM, dTTP 8,75 mM
1/7-Mix	dATP, dCTP, dGTP alle 10 mM, dTTP 8,4 mM
2/7-Mix	dATP und dGTP 10 mM, dTTP und dCTP 8,4 mM
Reamplifikation ohne Markierung	dATP, dCTP, dTTP, dGTP alle 10 mM

Tabelle 2.6 PCR-Programme für die Markierung von cDNA (links) und genomischer DNA (rechts)

1.	94°C	2:00 min	1.	94°C	1:00 min
2.	68°C	4:00 min	2.	60°C	0:30 min
3.	94°C	0:15 min	3.	72°C	2:00 min
4.	68°C	4:00 min	4.	94°C	0:30 min
5.	zurück zu 3. 10 Mal		5.	60°C	0:30 min
6.	94°C	0:15 min	6.	72°C	2:00 min + 0:20 min/Zyklus
7.	68°C	4:00 min + 0:10 min/Zyklus	7.	zurück zu 4. 10 Mal	
8.	zurück zu 6. 2 Mal		8.	4°C	∞
9.	68°C	7:00 min			
10.	4°C	∞			

Im Anschluss an die PCR werden für die Hybridisierung genomischer DNA und für cDNA-Hybridisierungen auf Oligonukleotidarrays die Primer durch einen Restriktionsverdau abgetrennt. Genomischen Reamplifikaten werden dazu 2 µl Puffer R⁺ und 50 U *TruI* hinzugefügt und bei 65°C für 3 h inkubiert. Bei cDNA kann nur der CP2-BGL-Primer durch einen Verdau des PCR-Produktes mit 30 U *BglI* für 3 h bei 37°C abgetrennt werden.

2.1.6 Nachträglicher Einbau eines T7-Promotors in das cDNA-Primäramplifikat

Reamplifikation mit einem Primer, der den T7-Promotor enthält

Es wird eine Reamplifikations-PCR wie unter Absatz 2.1.5 beschrieben durchgeführt. Als Primer wird jedoch T7TE-CP2 24 µM verwendet.

Annealing und Auffüllreaktion mit einem Primer, der den T7-Promotor enthält

25 µl eines Reamplikats nach Absatz 2.1.5 werden erneut mit dem Reamplifikationsansatz für cDNA jedoch ohne Wasser und mit 5,2 µl T7TE-CP2 als Oligonukleotid versetzt. Die PCR-Produkte werden in der PCR-Maschine in 2 min bei 94°C denaturiert, dann 30 sec bei 65°C der Primer angelagert und dieser in 7 min bei 68°C zum Doppelstrang aufgefüllt.

Direkte Ligation eines T7-Promotorkonstruktes an das Reamplifikat

25 µl eines Reamplikats nach Absatz 2.1.5 werden mit 0,625 µl T7TE-CP2 200 µM, 0,625 µl PO₄-T7TE-BLOCK 200 µM und 4 µl Wasser einem Temperaturgradienten von 94°C bis 15°C mit einer Absenkungsrate von 1°C/min unterworfen. Dann werden 1,75 µl ATP 10 mM und 1,75 µl T4-DNA-Ligase zugegeben und der Ansatz über Nacht bei 15°C ligiert. Eine Verknüpfung beider Stränge wird durch die 5'-Phosphatgruppe am PO₄-T7TE-BLOCK-Oligonukleotid gewährleistet.

Ligation eines T7-Promotorkonstruktes nach Restriktionsverdau

20 µl eines 8 Zyklen-Reamplikats nach Absatz 2.1.5 werden mit 1,5 µl *XbaI* 3 h bei 37°C verdaut und das Enzym anschließend für 15 min bei 65°C inaktiviert. Während des Verdau werden 1,1 µl T7-

ADAPTER 200 μM und 1,1 μl ddT7 200 μM einem Temperaturgradienten von 71°C bis 15°C mit einer Absenkungsrate von 1°C/min unterworfen. Das Produkt des Verdau wird über eine Qiagen MinElute-Säule aufgereinigt und in 10 μl Wasser pH 7,7 eluiert. Der basengepaarte Adaptor, 1,6 μl Puffer B⁺, 1,6 μl ATP 10 mM und 1,6 μl T4-DNA-Ligase werden mit dem Säuleneluat vereinigt und über Nacht bei 15°C ligiert. Zum Produkt der Ligation werden 0,5 μl KCl 200 mM, 2,0 μl dNTPs 10 mM, 0,3 μl Taq-Polymerase und 1,2 μl Wasser gegeben und der Adaptor während 10 min bei 72°C zum Doppelstrang aufgefüllt.

2.1.7 Effizienztest des T7-Promotor in einer *in vitro*-Transkription

Unter Verwendung der Reagenzien aus dem Ampliscribe T7-Transcription-Kit wird ein Ansatz aus 2 μl 10x T7 Puffer, 2 μl DTT 100 mM, 1,4 μl NTPs 10 mM, 2 μl T7-RNA-Polymerase, 11,6 μl Wasser und 1 μl cDNA-Template hergestellt und 3 h bei 37°C *in vitro*-transkribiert. Es folgt ein Verdau von 4 μl des Produkts mit 1 μl DNase I für 15 min bei 37°C. Nach einer Konzentrationsbestimmung lässt sich die Promotoreffizienz als Verhältnis der Menge cRNA aus der *in vitro*-Transkription zur Menge der eingesetzten cDNA berechnen.

2.1.8 Markierung über den T7-Promotor durch *in vitro*-Transkription

Die Reaktion verläuft analog der *in vitro*-Transkription aus Absatz 2.1.7. Es werden jedoch 1,4 μl 6/7-Nukleotidmix (8,4 mM UTP, alle anderen Nukleotide 10 mM) statt NTPs 10 mM und zusätzlich 2 μl Dig-UTP 1 mM eingesetzt. Als Template wird 1 μg cDNA verwendet. Das einzusetzende Volumen wird durch eine Konzentrationsbestimmung ermittelt, kann aber maximal 11,6 μl betragen, die benötigte Wassermenge wird entsprechend angepasst. Der DNase-Verdau entfällt.

2.1.9 Markierung von Gesamt-RNA während der cDNA-Synthese

Eine Gesamt-RNA-Präparation wird 3 min bei 14000 rpm/4°C zentrifugiert. Zu 50 μg aus dem Überstand werden 5 μg CFL5CT24 gegeben und mit Wasser auf 11 μl aufgefüllt. Nach einer Inkubation bei 70°C für 10 min werden die Proben direkt auf Eis gestellt und der Primer anschließend für 10 min bei Raumtemperatur angelagert. Dann werden 6 μl 5x Reverse Transcription-Puffer, 3 μl DTT 100 mM, 1,5 μl dNTP-Mix (2.1.5), 3 μl markiertes Nukleotid 1 mM, 2 μl Superscript II Reverse Transcriptase und Wasser ad 30 μl zugegeben und die Reaktion bei 44°C für 2 h durchgeführt. Die RNA wird durch Zugabe von 15 μl 0,1 N NaOH bei 70°C in 10 min zerstört, der Ansatz mit 15 μl 0,1 N HCl neutralisiert und über eine Qiagen PCR Purification Säule aufgereinigt.

2.1.10 Herstellung von *E.coli*-DNA zum Blockieren bei Hybridisierungen

Über Nacht werden Kulturen von pBS-transfizierten *E.coli* in LB-Medium + (0,1 mg/100 ml) Ampicillin angelegt. Die Bakterien werden geerntet und in 10 ml Puffer R (50 mM EDTA, 50 mM Tris, 5% Tween 20, 5% Triton X-100) resuspendiert. Mit 2 mg RNase, 20 mg Lysozym und 9 mg

Proteinase K wird über Nacht bei 37°C verdaut. Die DNA wird mit Phenol-Chloroform extrahiert und mit Ultraschall geschert. Nach einer Fällung mit Natriumacetat und Isopropanol erfolgt eine erneute Scherung mit Ultraschall bis zu einer maximalen Fragmentgröße von 1 kb sowie eine Konzentrationsbestimmung.

2.1.11 Hybridisierung auf murine Operon-Arrays

Paare aus Dig-markierter-Test- und Biotin-markierter Referenz-DNA aus Absatz 2.1.5 werden über Qiagen PCR Purification Säulen aufgereinigt und mit 50 µg Cot-1 DNA, 15 µg fragmentierter pBS-*E.coli*-DNA, 1/10 Volumenteil 3 M NaOAc pH 5,2 und 2,5 Volumenteilen 100% EtOH über Nacht bei -20 C gefällt. Die Fällung wird 45 min bei 14000 rpm/4°C zentrifugiert, das Pellet zweimal mit 700 µl EtOH 70% gewaschen und luftgetrocknet. Die DNA wird in 5 µl Wasser resuspendiert und für mindestens 30-45 min bei 37°C rückgelöst.

Währenddessen wird ein Operon-Array 45 min in Prähybridisierungspuffer (5x SSC + 0,1% SDS, 0,1% BSA Fraktion V) bei 42°C inkubiert, anschließend 20 Mal in Aqua dem. getaucht und durch Zentrifugation für 3 min bei 1200 rpm getrocknet. Der Array wird in die Arraybooster-Hybridisierungsstation eingelegt, eine AdvaCard mit 96%igem Ethanol gereinigt und über die Hybridisierungsfläche gelegt. Die Reservoirs der Hybridisierungskammer werden mit 2x SSC + 25% Formamid gefüllt und das Gerät auf 42°C vorgewärmt.

Die Probe wird 5 min bei 95°C denaturiert, mit einem geeigneten Volumen SlideHyb-Puffer 1 gemischt und auf den Array aufgetragen. Die Hybridisierung im Arraybooster verläuft bei 42°C über Nacht in 10 s-Zyklen mit einer Mixingrate von 24 dBm für 3 s gefolgt von 7 s Pause.

Nach Ende der Hybridisierung wird der Array aus der Hybridisierungsstation entnommen und sofort bei 50°C in folgenden Puffern gewaschen: 1x SSC + 0,1%SDS zweimal 10 min auf dem Schüttler, 0,5x SSC + 0,1%SDS zweimal 10 min auf dem Schüttler, 0,1x SSC zweimal 30 min im Hybridisierungssofen. Anschließend werden unspezifische Bindungen mit 1%iger Blocklösung in PBS für 30 min abgesättigt. Dann werden die Arrays mit 16 µg/ml α -Dig-Cy5 und 18 µg/ml Streptavidin-Cy3 in 1% Block/PBS 45 min unter Lichtschutz gefärbt. Überschüssiges Färbereagenz wird durch zweimaliges Waschen in 4x SSC + 0,2% Tween für 10 min entfernt, die Arrays kurz in 0,1x SSC gewaschen und durch Zentrifugation bei 1200 rpm für 2 min getrocknet. Dann werden sie in einem Axon-Reader gescannt.

Mit Hilfe der GenePix-Software können die Scans ausgewertet werden. Dazu wird ein Raster über das Bild gelegt, das alle Spots mit zugehörigen Gennamen enthält. Das Raster wird mit den Spots zur Deckung gebracht und fehlerhafte Spots markiert. Die Software überführt anschließend die Fluoreszenzintensitäten für jeden Spot, die Intensität des lokalen Hintergrunds und Aussagen über die Qualität der Signale in eine Tabelle.

2.1.12 Hybridisierung auf humane Operon-Arrays

Die Hybridisierung verläuft analog zum Verfahren für murine Operon-Arrays (Absatz 2.1.11) mit folgenden Modifikationen:

- Bei der Fällung werden weder Cot-1 DNA noch fragmentierte pBS-*E.coli*-DNA zugegeben.
- Nach der Hybridisierung wird folgendermaßen gewaschen: zweimal 5 min in 1x SSC + 0,1%SDS, 42°C, zweimal 10 min in 0,1x SSC + 0,1%SDS, RT, zweimal 2,5 min in 0,1x SSC, RT

2.1.13 Auswertung der Operon-Hybridisierungen in GeneSpring

Normalisierung

Es wird die für Zweifarben-Arrays übliche LOESS-Normalisierung verwendet. Zuvor können alle Messwerte kleiner 0,01 gleich 0,01 gesetzt werden, um Fehler bei der Bildung von Logarithmen zu vermeiden.

Filtern von Genlisten

Aus der normalisierten Liste aller Gene werden folgende entfernt:

- Gene, die nicht in mindestens einem der Arrays eines Experiments als „present“ eingestuft sind, also einen Messwert aufweisen.
- Gene, deren Fluoreszenzintensität für den Referenzkanal unter dem durch das Fehlermodell der Software für das jeweilige Experiment bestimmten Wert liegt, die also nicht zuverlässig gemessen werden konnten.

Je nach Experiment wird die Schnittmenge dieser Listen oder die Liste der „present genes“ für die weitere Auswertung verwendet.

Weitere Auswertungen

Je nach Fragestellung des Experiments werden weitere Auswertungen durchgeführt. Dazu gehören Scatter Plots, Clusteranalysen oder Berechnungen des Present Call. Die jeweils verwendeten Verfahren sind aus dem Ergebnisteil ersichtlich.

2.1.14 Hybridisierung auf Nylonfilter

Hybridisierung von cDNA-Proben

Die Membran wird in 2x SSC äquilibriert und dann in eine Hybridisierungsröhre mit 5 ml Dig Easy überführt. 500 µg fragmentierte pBS-*E.coli*-DNA werden für 5 min bei 94°C denaturiert, zugegeben und die Membran bei 45°C für mindestens 45 min prähybridisiert. 5 µg einer Dig-markierten Probe aus einer Labellingreaktion nach Absatz 2.1.5 werden zusammen mit 100 µg Heringsperm-DNA für

5 min bei 94°C denaturiert und der Mix zu der Prähybridisierungslösung hinzugefügt. Der Filter wird bei 45°C über Nacht hybridisiert.

Die Waschschritte nach der Hybridisierung erfolgen bei 68°C. Der Hybridisierungsofen wird aufgeheizt und der Filter zuerst in 2x SSC + 0,1%SDS gefolgt von 1x SSC + 0,1%SDS kurz geschwenkt. Dann wird in 0,5x SSC + 0,1%SDS für 15 min und anschließend zweimal 30 min in 0,1x SSC + 0,1%SDS im Ofen rotierend gewaschen. Die Membran wird in eine Schale mit Puffer 1 überführt und 30 min in 25 ml 1%iger Blocklösung geschwenkt. Dann werden 2,5 µl αDig-Fab-Alkalische Phosphatase-Konjugat zugegeben und der Filter 30 min gefärbt. Überschüssiger Antikörper wird zweimal 15 min in Puffer 1 + 0,003% Tween entfernt, der Filter zweimal kurz in Puffer 3 (0,01 M Tris-HCl, pH 9,5, 0,01 M NaCl) umgepuffert, von überschüssiger Flüssigkeit befreit und in eine Folie gelegt. 750 µl einer 1:10-Verdünnung von CSPD in Puffer 3 werden auf der Membran verteilt. Nach 5 min wird der Filter vorsichtig abgetrocknet, in eine neue Folie überführt und 15 min bei 37°C im Hybridisierungsofen inkubiert. Auf den Filter wird ein Film aufgelegt und dieser nach Expositionszeiten von 5 bis 120 min entwickelt.

Hybridisierung von cRNA-Proben

Die Hybridisierung verläuft wie bei cDNA-Proben allerdings bei einer Temperatur von 50°C und ohne die Denaturierung der markierten Probe. Außerdem ist auf die Verwendung RNase-freier, z.B. DEPC-behandelter Lösungen zu achten. Dazu werden 100 ml der zu behandelnden Lösung mit 100 µl DEPC versetzt, wiederholt gut geschüttelt und über Nacht stehen gelassen. Das überschüssige DEPC wird durch Autoklavieren zerstört.

2.1.15 Komparative genomische Hybridisierung (CGH)

Test-DNA und Referenz-DNA werden wie in Absatz 2.1.5 beschrieben markiert und zusammen mit 100 µg Cot-1 DNA, 100 µg Heringsperm-DNA, 1/10 Volumenteil 3 M NaOAc pH 5,2 und 2,5 Volumenteilen 100% EtOH über Nacht bei -20°C gefällt. Die Fällung wird für 45 min bei 14000 rpm/4°C abzentrifugiert, das Pellet mit 700 µl EtOH 70% gewaschen, luftgetrocknet und in 6 µl 100% Formamid bei 37°C 1 h unter Schütteln gelöst. Mit 6 µl 30%igem Master-Mix (4x SSC, 30% Dextran Sulfat Mr > 500000) wird weitere 30 min bei 37°C geschüttelt, dann wird die DNA bei 78°C für 6 min denaturiert und die Cot-1 DNA 30-60 min bei 37°C mit den repetitiven Anteilen der DNA prähybridisiert.

Parallel wird ein Objektträger mit Lymphozyten-Metaphasen eines gesunden männlichen Spenders (hergestellt nach Speicher *et al*¹³⁰) nach einem Äquilibrierschritt in 2x SSC für 1 h bei 37°C mit 200 µl RNase A (0,1 µg/µl) verdaut und diese dann durch dreimaliges Waschen in 2x SSC entfernt. Es folgt ein Verdau mit 50 mg Pepsin in 100 ml 10 mM HCl pH 2,0 bei 37°C für 2-4 min. Der Verdau wird durch dreimaliges Waschen in PBS gestoppt, der Objektträger in einer aufsteigenden Alkoholreihe (70%, 85% und 100% Ethanol) dehydriert und auf einer Heizplatte bei 37°C getrocknet. In einem

Wasserbad werden 100 ml 70% Formamid in 2x SSC, pH 7,2 auf 70°C erhitzt und der Objektträger darin für 1:45-2 min denaturiert. Die Denaturierungszeit muss wie auch die Zeit für den Pepsinverdau für jede Charge Metaphasen-Objektträger neu ermittelt werden. Der Objektträger wird erneut getrocknet, der DNA-Hybridisierungsansatz aufgetragen und das Hybridisierungsfeld durch ein Deckglas abgedeckt und mit Gummikleber verschlossen. Die Hybridisierung verläuft über zwei Nächte bei 37°C.

Nach der Hybridisierung wird der Objektträger dreimal in 4x SSC + 0,2% Tween/42°C gefolgt von 1x SSC/60°C gewaschen. Nach einem Umpufferschritt in 1x PBS + 0,2% Tween werden die unspezifischen Bindungsstellen für 30 min mit 1 ml 3%iger Block-Lösung (1x PBS pH 7,4 + 0,2% Tween, 3% BSA Fraktion V, 5% FCS) abgesättigt. Die überschüssige Block-Lösung wird mit 1x PBS + 0,2% Tween entfernt und der Objektträger mit 20 µl Anti-Digoxigenin-FITC und 2 µl Avidin-Cy 3.5 plus 180 µl 1%iger Block-Lösung (1x PBS + 0,2% Tween pH 7,4, 1% BSA Fraktion V, 5% FCS) 60 min bei 37°C unter Lichtschutz gefärbt. Überschüssiger Antikörper wird in drei Waschschritten bei 42°C in 4x SSC + 0,2% Tween entfernt und die DNA 2 min mit 1 ml DAPI-Lösung (1 µg/ml in 4x SSC + 0,2% Tween) angefärbt. Der Objektträger wird kurz mit destilliertem Wasser gewaschen, im Dunkeln getrocknet und mit Vectashield H-1000 eingedeckelt.

Zur Auswertung der CGH wird an einem Fluoreszenzmikroskops von 12 bis 15 geeigneten Metaphasen jeweils ein Bild mit den Fluoreszenzfiltern für DAPI, FITC und Cy3.5 aufgenommen. Die Chromosomen der digitalisierten Metaphasen können mit geeigneter Software in ein Karyogramm sortiert werden. Die Software vergleicht nach Normalisierung die integrierte Fluoreszenzintensität der Kanäle FITC und Cy 3.5. Das durchschnittliche Verhältnis aller ausgewerteten Metaphasen einer Hybridisierung wird graphisch in einem CGH Profil dargestellt. Ist das Intensitätsverhältnis FITC/Cy 3.5 größer als 1,25 bzw. kleiner als 0,75, liegt ein Gewinn oder Verlust im entsprechenden Chromosomenabschnitt vor.

2.1.16 Kultur von murinen dendritischen Zellen aus Knochenmark

Aus Femur und Tibia von C57/Bl6-Mäusen wird durch Ausspülen das Knochenmark gewonnen. Die Zellen werden in RPMI 1640 supplementiert mit 30 ng/ml GM-CSF, 5% FCS, 200 U/ml Penicillin, 200 U/ml Streptomycin und 50 mM β-Mercaptoethanol in sechs Tagen zu dendritischen Zellen (DCs) differenziert. Die Zellen werden mit 0,005 mg/ml αCD16/32 gefolgt von αCD86-Biotin (1:500) gefärbt und mit Streptavidin Beads des MACS-Systems markiert. Die CD86-positiven Zellen werden entfernt, die verbleibenden Zellen mit CD11c-Beads gefärbt und die positiven Zellen selektiert. Nach Färbung mit αCD11c-FITC (1:200) werden unter dem Mikroskop einzelne immature DCs isoliert. Zur Generierung maturer DCs werden die verbleibenden Kulturen über Nacht mit 1 µg/ml LPS stimuliert und weitere zwei Tage kultiviert. Mit αCD86-Biotin und Streptavidin Beads werden positive Zellen selektiert, diese mit αCD11c-FITC gefärbt und einzelne mature DCs isoliert. Durch FACS-Färbung wird die Reinheit jeder Population der Aufreinigung bestimmt.

2.1.17 Zellsortierung mit dem MACS-System

Es wird nach dem Standardprotokoll der Firma Miltenyi verfahren: Je $1 \cdot 10^7$ Zellen werden mit 20 μl beadgekoppeltem Antikörper versetzt, mit MACS-Puffer auf 100 μl aufgefüllt und 15 min bei 4°C gefärbt. Überschüssige Beads werden durch Waschen mit 10-20fachem Volumen MACS-Puffer entfernt. Die Proben werden in 500 μl MACS-Puffer aufgenommen und die beadmarkierten Zellen auf einer Säule in einem Magnetfeld immobilisiert. Nach dreimaligem Waschen mit MACS-Puffer kann die gebundene Fraktion abseits vom Magneten eluiert werden.

Sollen andere Antikörper als die von der Firma Miltenyi an Beads gekoppelte verwendet werden, stehen Streptavidin-, α Maus-, α FITC-, α PE-, und α APC-Beads zur Verfügung. Diese können nach einer Färbung mit dem entsprechenden Antikörper wie Direktkonjugate verwendet werden.

2.1.18 Aufarbeitung von humanem Knochenmark

Das Knochenmark wird mit derselben Menge Hank's Salzlösung versetzt und 10 min bei 850 rpm/4°C zentrifugiert. Der Überstand wird abgenommen und die Zellen auf 7 ml 70% Percoll in isotoner NaCl-Lösung geschichtet. Die Dichtegradientenzentrifugation erfolgt bei 2050 rpm/4°C für 20 min.

Die Zellen der Interphase werden entnommen und in 50 ml 1x PBS gewaschen. Nach einer Zentrifugation bei 1500 rpm/4°C für 10 min wird die Zellzahl bestimmt.

Die Zellen werden pro $1 \cdot 10^7$ Zellen mit 100 μl α CD33-APC 1:20 und α CD45-APC 1:20 für 15 min bei 4°C gefärbt. Die Depletion positiver Zellen erfolgt nach Absatz 2.1.17 mit α APC- und α CD235a-Beads. Die Negativfraktion der Depletion wird gegebenenfalls zuerst mit 5 $\mu\text{g/ml}$ Hoechst33342 in DMEM-Medium mit 2% FCS für 30 min bei 37°C gefärbt (nach Goodell *et al*¹³¹). Anschließend erfolgt eine Färbung mit α EpCAM-Biotin 1:200 und 1,8 $\mu\text{g/ml}$ Streptavidin-Cy3 (nach Klein *et al*⁶⁴).

2.2 Verwendete Standardprotokolle

2.2.1 Zellkultur: TUBO-Zelllinie: adhären in DMEM + 20% FCS

2.2.2 Genspezifische PCR: Es wird nach Schardt *et al*⁶⁸ verfahren.

2.2.3 Agarose Gelelektrophorese: 1,5% Agarose in 1x TBE nach „Molecular Cloning“¹³²

2.2.4 Konzentrationsbestimmung von Nukleinsäuren: photometrische Bestimmung¹³²

2.2.5 Aufreinigung von PCR-Produkten: Es wird nach dem Protokoll der Firma Qiagen für MinElute und PCR-Purification Kit verfahren.

2.2.6 Isolierung von Gesamt-RNA: Es wird nach Chomczynski, P. und Sacchi, N.¹³³ verfahren.

2.2.7 Alternative Markierungsreaktionen für PCR-Produkte

1. Aminoallyl-Markierung: Das Protokoll findet sich unter http://pga.tigr.org/sop/M004_1a.pdf
2. TdT-Endmarkierung: Es wird analog der Tailing-Reaktion aus Absatz 2.1.1 verfahren.
3. Randompriming: Es wird nach „Molecular Cloning“¹³² verfahren

2.3 Verzeichnis der verwendeten Abkürzungen

Formelgrößen wurden mit dem international gebräuchlichen SI-Einheiten und chemische Elemente gemäß ihrer Bezeichnung im Periodensystem abgekürzt. Für die Abkürzung gebräuchlicher Wörter der deutschen Sprache wurde gemäß Duden verfahren.

A	Adenin
α	anti
Bio	Biotin
bp	Basenpaar
C	Cytosin
dd	dideoxy
Dig	Digoxygenin
DNA	Desoxyribonukleinsäure
dNTP	2'-Desoxynucleosid-5'-triphosphat
DTT	1,4-Dithiothreitol
EST	Expressed Sequence Tag
FITC	Fluorescein
G	Guanin
mRNA	Messenger Ribonukleinsäure
PCR	Polymerasekettenreaktion
RNA	Ribonukleinsäure
SDS	Natriumdodecylsulfat
T	Thymin
U	Uracil

2.4 Material, Geräte, Software, Puffer und Lösungen

Enzyme, Antikörper und Konjugate

α CD11c-FITC	BD Pharmingen, Heidelberg
α CD16/32	BD Pharmingen, Heidelberg
α CD33-APC	Caltag Laboratories, Hamburg
α CD45-APC	Caltag Laboratories, Hamburg
α CD86-Biotin	BD Pharmingen, Heidelberg
α Dig-Cy5	Jackson Laboratories, Bar Harbor, ME
α Dig-FITC	Roche, Mannheim
α EpCAM (3B10)-Biotin	Micromet, München
Avidin-Cy3.5	Roche, Mannheim
<i>Bgl</i> I	MBI Fermentas, St. Leon-Rot
Klenow-Enzym/DNA-Polymerase I (<i>E. coli</i>)	New England Biolabs, Frankfurt
<i>Mse</i> I 50 U/ μ l	New England Biolabs, Frankfurt
Pepsin	Roche, Mannheim
Proteinase K	Roche, Mannheim
RNase A	Qiagen, Hilden
Streptavidin-Cy3	Jackson Laboratories, Bar Harbor, ME
Superscript II Reverse Transcriptase	Invitrogen, Karlsruhe
T4 DNA-Ligase	Roche, Mannheim
Taq-Polymerase	Roche, Mannheim
Terminal Deoxynukleotid Trynsferase	Amersham Biosciences, Freiburg
<i>Tru</i> I	MBI Fermentas, St. Leon-Rot
<i>Xba</i> I	MBI Fermentas, St. Leon-Rot

Chemikalien

Aminoallyl-dUTP	Sigma, Deisenhofen
Aminodigoxigenin-dCTP	NEN Perkin Elmer, Rodgau-Jügesheim
ATP	Roche, Mannheim
Biotin-dCTP	Invitrogen, Karlsruhe
Biotin-dUTP	Roche, Mannheim
BSA Fraktion V	Paesel + Lorei, Hanau
BSA für PCR	Roche, Mannheim
Cot-1 DNA	Roche, Mannheim
CSPD	Roche, Mannheim
Cy3-dUTP, Cy3-dCTP, Cy5-dUTP, Cy5-dCTP	NEN Perkin Elmer, Rodgau-Jügesheim
DAPI	Roche, Mannheim
dATP, dCTP, dGTP, dTTP	Amersham Biosciences, Freiburg
DEPC	Sigma, Deisenhofen
Dextran Sulfat	Sigma, Deisenhofen
Dig-dUTP	Roche, Mannheim
Dig-UTP	Roche, Mannheim
Heringsperm-DNA	Invitrogen, Karlsruhe
Igepal	Sigma, Deisenhofen
OnePhorAll Puffer	Amersham Biosciences, Freiburg
PCR-Öl	Sigma, Deisenhofen
Perkoll	Amersham Biosciences, Freiburg
Puffer B ⁺	MBI Fermentas, St. Leon-Rot
Puffer R ⁺	MBI Fermentas, St. Leon-Rot
Reaktiver Cy3-Ester	Amersham Biosciences, Freiburg
Reaktiver Cy5- Ester	Amersham Biosciences, Freiburg
Reverse Transcription Puffer 5x	Invitrogen, Karlsruhe
Röntgenfilm (Biomax Light)	Kodak
SDS	Sigma, Deisenhofen
Slide Hyb Puffer	Ambion, Huntingdon, GB
tRNA	Roche, Mannheim
Tween 20	Sigma, Deisenhofen
Vectashield H-1000	Linearis, Hamburg
Wasser (LiChrosolv für Chromatographie)	Merk, Darmstadt

Kits

Ampliscribe T7 High Yield Transcription Kit	Epicentre, Madison, WI
Dig Luminescent Detection Kit	Roche, Mannheim
Dynalbeads mRNA DIRECT KIT	Dynal, Hamburg
Dynalbeads zur Zellseparation	Dynal, Hamburg
Expand Long Template Kit	Roche, Mannheim
LightCycler-Faststart DNA Master SYBR Green I	Roche, Mannheim
MACS-Antikörper, -Beads und -Säulen	Miltenyi, Bergisch Gladbach
MinElute Kit	Qiagen, Hilden
mTrap mRNA Isolation Kit	Active Motif, Rixensart, B
PCR Purification Kit	Qiagen, Hilden
StemSep Cell Separation	Stem Cell Technologies, Vancouver, CAN

Arrays

Atlas Human Cancer Array (Nylon).....	Clontech, Palo Alto, CA
Corning Epoxide Coated Slide	Corning, Schiphol-Rijk, NL
Custom Standard Array (Nylon)	AG Klein, Institut für Immunologie, München
Febit Array	Febit AG, Mannheim
GAPS II-Slides	Corning, Schiphol-Rijk, NL
GeneChip U95A	Affymetrix, Santa Clara, CA
Memorec PIQOR cDNA Array System	Miltenyi, Bergisch Gladbach
Nexterion Slide E	Schott Nexterion, Jena
OpArray human Version 4	Operon, Köln
Operon Arrayready Oligo-Set Human Version 3	Operon, Köln
Operon Arrayready Oligo-Set Human Version 4	Operon, Köln
Operon Arrayready Oligo-Set Mouse Version 2	Operon, Köln
Operon Human Sample Set Version 3	Operon, Köln
Operon Human Sample Set Version 4	Operon, Köln
UltraGAPS-Slides	Corning, Schiphol-Rijk, NL

Geräte

AdvaCard Größe 2 und 3.....	Implen, München
Arraybooster Hybridisierungsstation	Implen, München
Floureszenzmikroskop Leica DMXA-RF8	Leica Microsystems, Bensheim
Hybridisierungsöfen Shake `n` Stak.....	Hybaid, Heidelberg
Hybridisierungsröhren.....	Biometra, Göttingen
Invertmikroskop Leica DM IRB	Leica Microsystems, Bensheim
LightCycler.....	Roche, Mannheim
Microarray Scanner Genepix 4000A.....	Axon Instruments, Union City, CA
MJR Research Peltier Thermal Cycler PTC-200	Bio-Rad, Waltham, MA

Software

GenePix Pro 4.1.....	Axon Instruments, Union City, CA
GeneSpring.....	Silicon Genetics, Redwood City, CA
QCGH.....	Leica Microsystems, Bensheim
QFISH	Leica Microsystems, Bensheim
R.....	The R-Project, http://www.r-project.org/
RelQuant.....	Roche, Mannheim
Statistik.....	http://faculty.vassar.edu/lowry/VasserStats.html

Puffer und Lösungen

PBS (Phosphate Buffered Saline).....	8,5 mM Na ₂ HPO ₄ , 2 mM KH ₂ PO ₄ , 150 mM NaCl; pH 7,4
SSC (Saline Sodiumcitrat).....	15 mM NaCitrat, 150 mM NaCl ; pH 7,4
Polyacrylamid-Carrier	5% Acrylamid, 40 mM Tris-HCl pH 7-8, 20 mM Natrium-acetat pH 7,8, 1 mM EDTA pH 7,8-8,0, 1% Ammonium-persulfat, 0,1% TEMED
cDNA-Igepal-Waschpuffer	50 mM Tris-HCl pH 8,0, 75 mM KCl, 10 mM DTT, 0,25% Igepal
cDNA-Tween-Waschpuffer	50 mM Tris-HCl pH 8,0, 75 mM KCl, 10 mM DTT, 0,5% Tween

Tailing-Waschpuffer	50 mM KH ₂ PO ₄ pH 7,0, 1 mM DTT, 0,25% Igepal
Puffer 1	100 mM Maleinsäure, 150 mM NaCl; pH 7,5
PCR-Puffer + dNTPs.....	10 mM MgCl ₂ , 100 mM Tris-HCl pH 8,5, 500 mM KCl, 1 mM dNTP's
MACS-Puffer	1x PBS, 0,5% BSA, 2 mM EDTA

2.5 Oligonukleotide und Primer

Oligonukleotide und Primer für humane genomische DNA

LIB1	AGT GGG ATT CCT GCT GTC AGT
ddMse11	TAA CTG ACA G-ddC

Gen	5'-Primer	3'-Primer
TP53 EX 2/3	GAA GCG TCT CAT GCT GGA TC	CAG CCC AAC CCT TGT CCT TA
CK 19	GAA GAT CCG CGA CTG GTA C	TTC ATG CTC AGC TGT GAC TG

Primer für genspezifische PCR, humane cDNA

Gen	5'-Primer	3'-Primer
EF1 α	TGC CCC AGG ACA CAG AGA AC	CTG TGT CGG GGT TGT AGC CA
β -AKTIN	CTC CTT AAT GTC ACG CAC GAT TTC	GTG GGG CGC CCC AGG CAC CA
p53_1	TCT GTC CCT TCC CAG AAA ACC	GTC AGA GCC AAC CTC AGG C
p53_2	GGA TGA CAG AAA CAC TTT TCG	TCA GCT CTC GGA ACA TCT C

Primer für quantitative real time PCR, murine cDNA

Gen	5'-Primer	3'-Primer
GAPDH	GCT GAG TAT GTC GTG GAG TC	AGT TGG TGG TGC AGG ATG C
KRT2-8	CGT CTG TGG TGC TGT CTA TG	GCG GTT GAT GTT GCG GTT C
CATNB	GGT GCT GAC TTT GCT TGC TT	ATC GGA CCA TTT CTA TCT GTC
V-ERB-B2	CAC AAA GCA GGG AGA AGG TC	GGC AGT CAG GGT CTG TAA TC
MAPK1	CTT GAT GGG CTC TCC TCA TG	CAC CGC ACT GGG ATA ACA C
SPAG9	TGT GGC AAG TGA TGT GTG GC	CTC GCT CCT CTG TAT TGT TAC
LTA4H	CAT TTG TGG ACG GTT GTT TG	GGG ATG GAG GAA TAG GCT ACA TCG
IL1 β	GAG CCC ATC CTC TGT GAC TC	GTT GCT TGG TTC TCC TTG TAC
ACE	AGG GTC TTT GAT GGA AGC ATC	CCT CGT GGA ACT GGA ACT G
IL2R	ATC TGG GAT GAG GGA ACC TC	GCC ACG GAA GAA AGC CAT GC
PHLDA1	ACC TCC AAC TCT GCC TGA AG	GTC GGG TGA AGT CCA AAC TAC
CSF1R	TGA GTC AGA AGC CCT TCG AC	CAA CAG CAG TAG CAC CAG C
LTA4H	CAT TTG TGG ACG GTT GTT TG	GGG ATG GAG GAA TAG GCT ACA TCG
RARA	AGC CAG ACA AGG TGG ACA TG	AGA ACG AAT GAG CAG GAA GC
EMID1	GAG GAC ACC CAA CCA ACT AC	TCA CAG AGG GAA CCA ACC AG
SIGNR1	TCC ACA GAA GCC AAG ATG C	AGC AGG AGC CCA GCC AAG

Oligonukleotide und Primer für cDNA-Primäramplifikation und Einbau eines T7-Promotors

BLOCK 1	CTA TAG TGA GTC GTA TT-ddC
BLOCK 2	TAG TGA GTC GTA TT-ddC
CFL5C10	CCC CCC CCC CCC CCC GTC TAG ANN NNN NNN NN
CFL5C6	CCC CCC CCC CCC CCC GTC TAG ANN NNNN
CFL5C8	CCC CCC CCC CCC CCC GTC TAG ANN NNN NNN
CFL5C8-NO XBA.....	CCC CCC CCC CCC CCC GTC TAC ANN NNN NNN
CFL5CT15	CCC CCC CCC CCC CCC GTC TAG ATT TTT TTT TTT TVN
CFL5CT15-NO XBA	CCC CCC CCC CCC CCC GTC TAC ATT TTT TTT TTT TVN
CFL5CT24	CCC CCC CCC CCC CCC GTC TAG ATT TTT TTT TTT TTT TTT TVN
CFL5CT24-NO XBA	CCC CCC CCC CCC CCC GTC TAC ATT TTT TTT TTT TTT TTT TVN
CP2	TAG GAA TTC ATG CCC CCC CCC CCC
CP2-BGL	TCA GAA TTC ATG CCG CCC CCC CGG CCC
ddT7	CTA GCC GCC T-ddC
G9-BLOCK.....	GGG GGG GGG-ddC
LIB-N8.....	AGT GGG ATT CCT GCT GTC AGT NNN NNN NN
LIB-T7TE-T24-VN	AGT GGG ATT CCT GCT GTC AGT TAA TAC GAC TCA CTA TAG GGA GGC TTT TTT TTT TTT TTT TTT VN
OLIGO dC-N6	CCC CCC CCC CCC CCC NNN NNN
OLIGO dC-N8	CCC CCC CCC CCC CCC NNN NNN NN
OLIGO dC-N10.....	CCC CCC CCC CCC CCC NNN NNN NNN N
T7-ADAPTER	GGC CAG TGA ATT GTA ATA CGA CTC ACT ATA GGG AGG CGG
T7N8C	CCC CCC CCC CCC CCC TAA TAC GAC TCA CTA TAG GGN NNN NNN N
T7TA-T24-VN.....	CCC CCC CCC CCC CCC TGG TAA TAC GAC TCA CTA TAG GGA GGC GGT CTA GAT TTT TTT TTT TTT TTT TTT TTVN
T7TB-T24-VN.....	CCC CCC CCC CCC CCC GGC CAG TGA ATT GTA ATA CGA CTC ACT ATA GGG AGG TTT TTT TTT TTT TTT TTT TTT VN
T7TC-T15-VN.....	CCC CCC CCC CCC CCC TAA TAC GAC TCA CTA TAG GGT TTT TTT TTT TTT TTVN
T7TC-T24-VN.....	CCC CCC CCC CCC CCC TAA TAC GAC TCA CTA TAG GGT TTT TTT TTT TTT TTT TTT TTVN
T7TD-T24-VN.....	CCC CCC CCC CCC CCC GGC CAG TGA ATT GTA ATA CGA CTC ACT ATA GGG AGG CGG TTT TTT TTT TTT TTT TTT TTT VN
T7TE-BLOCK	PO ₄ -GCC TCC CTA TAG TGA GTC GTA TTA
T7TE-CP2.....	TAA TAC GAC TCA CTA TAG GGA GGC TCA GAA TTC ATG CCC CCC CCC CCC
T7TE-T24-VN	CCC CCC CCC CCC CCC GGC CAG TAA TAC GAC TCA CTA TAG GGA GGC TTT TTT TTT TTT TTT TTT TTT VN

Alle Oligonukleotide wurden von Metabion, München, synthetisiert.

2.6 Patientenkollektiv

Soweit zum Zeitpunkt der Erstellung der Dissertation bekannt, enthält Tabelle 2.7 die Daten zum Kollektiv der Mammakarzinompatienten, von denen in der vorliegenden Arbeit Knochenmarkaspirate untersucht wurden.

Tabelle 2.7 Daten zum Patientenkollektiv

Kriterium		Anzahl
Geschlecht des Patienten	weiblich	38
	männlich	1
Alter des Patienten	<30 Jahre	0
	31-40 Jahre	3
	41-50 Jahre	8
	51-60 Jahre	7
	61-70 Jahre	8
	71-80 Jahre	8
	>80 Jahre	2
Ausdehnung des Primärtumors	pT ₁	10
	pT ₂	10
	pT ₃	3
	pT ₄	3
	DCIS	1
Nodalstatus	N ₀	11
	N ₁₋₃	16
Vorhandensein von Fernmetastasen	M ₀	25
	M ₁	6
Karzinomtyp	duktal	20
	lobulär	4
Estrogenrezeptor Status des Primärtumors	positiv	19
	negativ	7
Progesteronrezeptor Status des Primärtumors	positiv	13
	negativ	13
HER2-rezeptor Status des Primärtumors	positiv	4
	negativ	21

Tabelle 2.8 enthält detailliertere Angaben zu den vier Patientinnen, bei denen die Genexpression EpCAM-positiver Zellen aus dem Knochenmark außer auf Nylonarrays auch auf humanen Operon-Oligonukleotidarrays analysiert wurde.

Tabelle 2.8 Detaillierte Patientendaten zu den Hybridisierungen auf Operon-Arrays

Patientennummer	050125	050713	050727	050728	
Geschlecht	weiblich	weiblich	weiblich	weiblich	
Tumor	Mamma-Ca	zweiherdiges Mamma-Ca	Mamma-Ca	Mamma-Ca	
Alter bei Diagnosestellung	67 Jahre	75 Jahre	57 Jahre	64 Jahre	
Histologie	k.A. ^a	invasiv; lobulär und duktolobulär	invasiv; duktal	invasiv; duktal	
TNM	initial pT2, pN0, cM0, mittlerweile M1	pT1c, cN0, cM0	pT2, pN0 (0/11), cM0	pT3, pN1a (1/18), cM0	
Grading	G3	G2	G3	G3	
neoadjuvante Therapie	nein	nein	nein	nein	
Estrogenrezeptorstatus	0/12	9/12	4/12	12/12	
Progesteronrezeptorstatus	0/12	6/12	2/12	12/12	
HER2/neu-Status	negativ	negativ	negativ	negativ	
Isolierte Zelle	050125-03	050713- 06	050713- 07	050727-08	050728-03
EpCAM	+	+	+	+	+
Hoechst 33342	+	n.a. ^b	n.a.	n.a.	n.a.
ERI ^c (Array)	0,9	- ^d	1,3	1,5	5,6
HER2 (Array)	3,3	-	1,0	0,5	0,5

^a: keine Angaben^b: nicht analysiert^c: Estrogenrezeptor α ^d: nicht nachweisbar

2.7 Kooperationspartner

Humane Knochenmarkaspirate wurden von Dr. Daniel Oruzio, II. Medizinische Klinik, Zentral-klinikum Augsburg sowie von Thomas Blankenstein, Frauenklinik des Klinikums Innenstadt der LMU München, zur Verfügung gestellt. Die Proben wurden bei Raumtemperatur schnellstmöglich zur Aufarbeitung ins Institut für Immunologie gebracht.

Für die Herstellung von Genexpressionsarrays aus lyophilisierten Oligonukleotiden der Firma Operon bestanden Kooperationen mit Dr. Jan Vijg, STCBM San Antonio, Texas; Dr. Peter Lichter und Dr. Bernhard Radlwimmer, DKFZ Heidelberg; und Dr. Markus Panhuysen, MPI für Psychiatrie, München. Affymetrix GeneChip-Arrays wurden in Kooperation mit Dr. Michael Hummel, Charité Berlin hybridisiert.

3 Ergebnisse: Etablierung eines Protokolls zur Untersuchung der Genexpression von Einzelzellen

3.1 Optimierung eines Protokolls zur Analyse von mRNA aus Einzelzellen

Zu Beginn der Arbeit lag bereits ein Protokoll zur Erstellung von cDNA-Bibliotheken einzelner Zellen und Analyse der Genexpression in der Arbeitsgruppe vor⁶⁴. Das Protokoll benutzte die Extraktion der mRNA über den Poly A-Schwanz durch eine Bindung an Oligo dT-Beads gefolgt von einer reversen Transkription der mRNA. Die cDNA wurde anschließend durch eine PCR primär amplifiziert, wobei ein einziger Primer zur Vervielfältigung aller Transkripte unabhängig von deren Sequenz eingesetzt werden konnte (Abbildung 3.1 links). Die Primäramplifikate wurden auf verschiedene Nylon-Filter, die selbst hergestellt oder z.B. von der Firma Clontech bezogen wurden, hybridisiert.

Parallel bestand die Möglichkeit zur Isolation der genomischen DNA derselben Zelle. Dazu wurde die mRNA während der Extraktion an den Beads in einem Magnetfeld immobilisiert, wonach die übrigen zellulären Bestandteile inklusive der genomische DNA abgetrennt werden konnten. Die genomische DNA wurde aus der Lösung ausgefällt und nach dem Rücklösen ebenfalls einem Amplifikationsprotokoll unterworfen.

3.1.1 Etablierung des PNA-Protokolls

Eine Schwierigkeit des bestehenden Protokolls lag in der unvollständigen Isolation der mRNA, was zum Verlust selten exprimierter Transkripte führte. Außerdem zeigten sich Limitationen bei allgemein geringem RNA-Gehalt der Zelle sowie bei der parallelen Isolierung der genomischen DNA (Abbildung 3.2 und Abbildung 3.4). Daher wurde im Rahmen der vorliegenden Arbeit versucht, das bestehende Protokoll zu optimieren sowie um die Möglichkeit der Hybridisierung auf Oligonukleotidarrays mit gesamter Genomabdeckung zu erweitern.

Da murine Zellen nur etwa 10% der RNA-Menge humaner Zellen enthalten (Daten nicht gezeigt) wurden Teile der Experimente mit Mauseinzelzellen durchgeführt, was Aussagen über die Leistungsfähigkeit der Methode bei Zellen mit geringem RNA-Gehalt erleichterte. Untersuchungen einzelner muriner Zellen mit dem bestehenden Amplifikations-Protokoll waren nicht sinnvoll möglich, da die mRNA durch die Oligo dT-Beads nicht quantitativ aus dem Zellysate extrahiert wurde. Um die Isolation zu verbessern, wurde das Immobilisationsprinzip verändert. Statt Oligo dT-Beads wurden in einem ersten Schritt biotinylierte Oligo dT-PNAs an den Poly A-Schwanz der mRNA angelagert und diese anschließend mit Streptavidin-Beads gekoppelt (Abbildung 3.1 rechts). Bei PNAs handelt es sich um Peptide Nucleic Acid Analoga, in denen das Zucker-Phosphat-Rückgrat der DNA durch ein Peptidgerüst ersetzt ist. PNAs binden mit einer wesentlich höheren Affinität an Nukleinsäuren als Oligonukleotide.

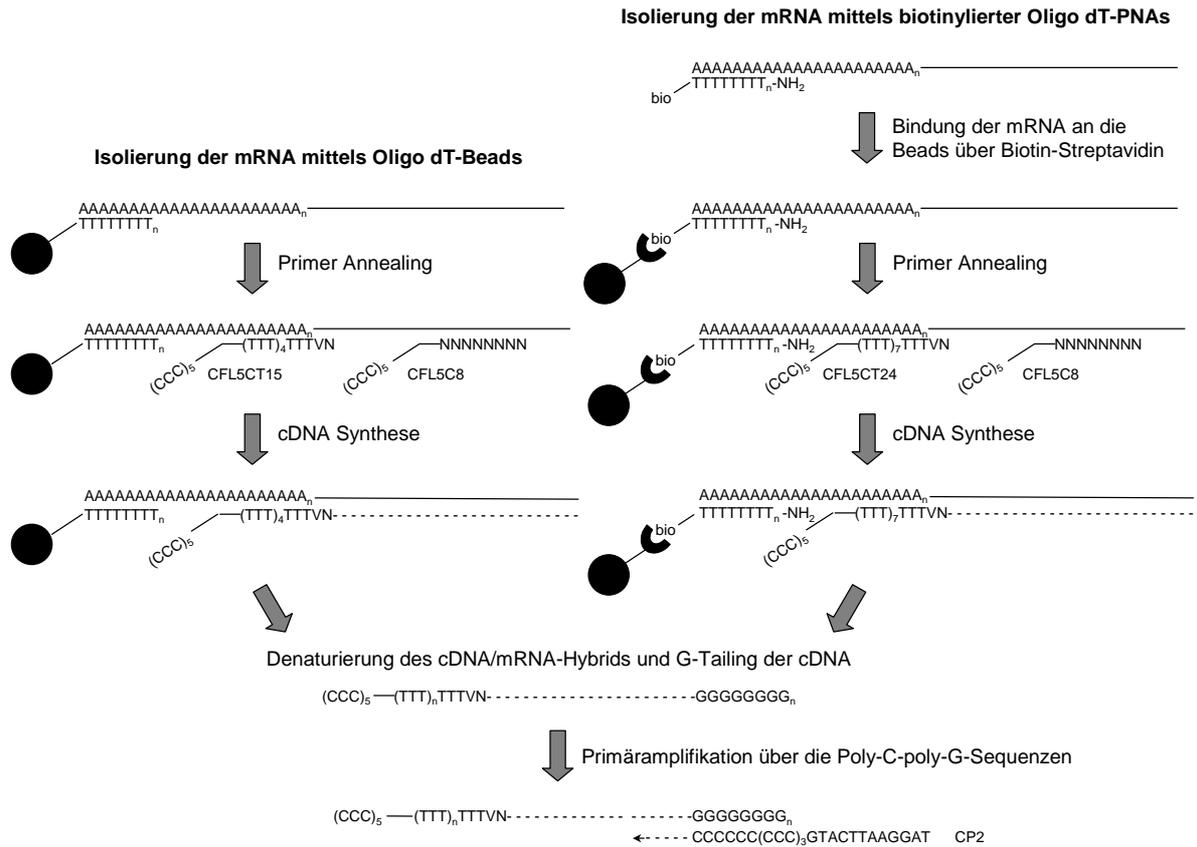


Abbildung 3.1 Schema beider Protokolle zur Isolation und Amplifikation von mRNA aus Einzelzellen

Das bei Beginn der Arbeit vorliegende Oligo dT-Beadprotokoll ist links dargestellt, das neu erarbeitete PNA-Protokoll rechts. Nach der Isolation der mRNA erfolgt die cDNA-Synthese mit einem Gemisch aus Poly T- (CFL5CT) und Zufallsprimer (CFL5C8) zur Fragmentierung der entstehenden cDNA. Beide Primer enthalten eine Poly C-Sequenz als erster Primerbindungsstelle für die spätere PCR zur Primäramplifikation. Das entstandene cDNA/mRNA-Hybrid wird denaturiert und ein G-Tailing am 3'-Ende der cDNA durchgeführt, wobei eine zweite Primerbindungsstelle für die nachfolgende Amplifikation mittels PCR gebildet wird. In der PCR kann ein ubiquitärer Poly G-Primer (CP2) zur Amplifikation aller cDNA-Fragmente unabhängig von ihrer Sequenz verwendet werden.

Die Versuche zur Optimierung der Isolation der mRNA wurden wie auch alle anderen Experimente zur Veränderung des Isolations- und Amplifikationsprotokoll nach denselben Kriterien ausgewertet. Der Amplifikationserfolg wurde durch genspezifische PCR-Reaktionen auf hoch exprimierte Transkripte (*EF1-α* und *β-AKTIN*) kontrolliert. Als Maß für Ausbeute der cDNA sowie quantitative Isolation und Fragmentierungsgrad der mRNA diente eine direkte Gelelektrophorese des Primäramplifikats. Optimal war dabei ein homogener Schmier zwischen 200 bp und 2000 bp mit der höchsten Intensität bei etwa 500 bp. Banden oder das Fehlen von Teilen oder des gesamten Schmiere deuteten auf unvollständige Extraktion, Fehler in der Amplifikation beziehungsweise Fragmentierung der mRNA hin. Die Ausbeute ließ sich zudem durch genspezifische PCR-Reaktionen auf gering exprimierte Transkripte wie *TP53* überprüfen.

Abbildung 3.2 zeigt einen Vergleich der beiden Protokolle nach der Isolation und Amplifikation von mRNA aus Mauseinzelzellen. Besonders deutlich wurden die Unterschiede auch nach Hybridisierung

der Proben auf verschiedene Arrays. Als Referenzwert diente dabei eine Hybridisierung von nicht amplifizierter Gesamt-RNA nach einem Standardprotokoll, zu der Korrelationskoeffizienten der amplifizierten Proben berechnet wurden.

Kriterium		PNA-Protokoll	Oligo dT-Bead-Protokoll
Isolation/Amplifikation	PCR (hoch exprimiert)	40/40 (100 %)	16/16 (100 %)
	Schmier gut	39/40 (98 %)	0/8 (0 %)
Hybridisierung	Arrayhybridisierung	möglich	nicht möglich
	Korrelation zur Gesamt-RNA	0,29	-0,12

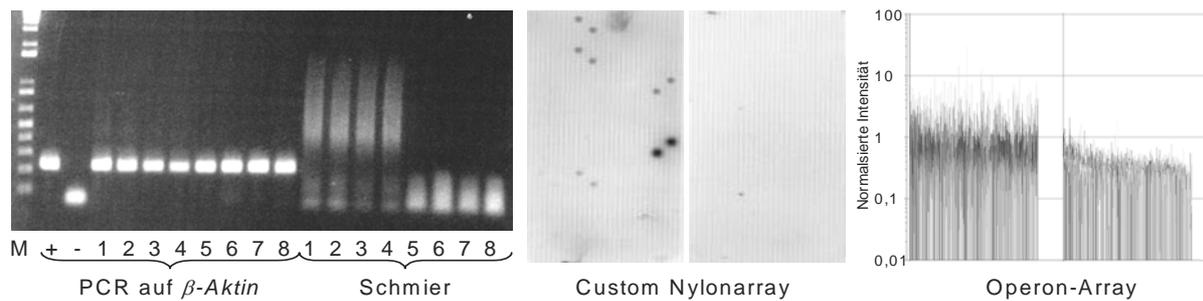


Abbildung 3.2 Vergleich von cDNA-Amplifikaten aus Mauseinzelzellen mit beiden Amplifikationsprotokollen

Einzelzellen wurden mit dem PNA-Protokoll (1-4 bzw. linker Array) und dem Oligo dT-Bead-Protokoll (5-8 bzw. rechter Array) bearbeitet. Links sind die Ergebnisse der genspezifischen PCR auf β -Aktin sowie die direkte Gelelektrophorese des Primäramplifikats gezeigt, in der Mitte die Hybridisierung auf Nylonarrays, rechts auf Operon-Oligonukleotidarrays. Hier sind vor allem schwarz dargestellte Signale mit einer normalisierten Intensität größer eins als gültig einzustufen.

Mit dem Oligo dT-Bead-Protokoll war eine Amplifikation von cDNA aus Mauseinzelzellen zwar grundsätzlich möglich (positive β -Aktin-PCR), ein Fehlen des Schmiere wies jedoch auf mangelhafte Isolation, Ausbeute und Repräsentation der cDNA-Amplifikate hin. Dies bestätigte sich nach der Hybridisierung, da die Oligo dT-Probe auf Nylonarrays keinerlei Signale, auf Operon-Oligonukleotidarrays nur Intensitäten im Bereich des Hintergrundrauschens aufwies und in ihrem Hybridisierungsergebnis zur Gesamt-RNA antikorreliert war. Das PNA-Protokoll hingegen brachte einen deutlichen Sensitivitätsgewinn, vermutlich durch eine höhere Affinität zwischen mRNA und PNAs im Vergleich zu mRNA und Oligo dT-Beads, und führte zu sinnvollen Hybridisierungsergebnissen.

Neben der Veränderung des Immobilisationsprinzip für mRNA wurden weitere Schritte im bestehenden Protokoll auf die Möglichkeit der Optimierung geprüft. Um die Kinetik der Isolation der mRNA näher zu charakterisieren, wurden verschieden lange Zeiten zur Bindung der PNA-mRNA-Hybride an die Beads zwischen 30 und 60 min getestet. Zu lange Zeiten gingen mit Verlusten bei seltenen Transkripten durch Fragmentierung einher. Das Gleiche galt für eine zu lange Inkubation bei 78°C während des Hotstarts der Primäramplifikation oder eine PCR ganz ohne Hotstart. Weitere Experimente, auch in Zusammenhang mit den Versuchen zum Einbau eines T7-Promotors in die cDNA-Primäramplifikate (Absatz 3.2.1), dienten der Identifizierung optimaler cDNA-Synthesepriemer. Die Erfolgsraten verschiedener Primer hinsichtlich genspezifischer PCR-Reaktionen und des Schmiere

des Primäramplifikats sowie die Korrelation zu einer Gesamt-RNA-Präparation nach Hybridisierung auf einen Oligonukleotidarray sind in Abbildung 3.3 dargestellt.

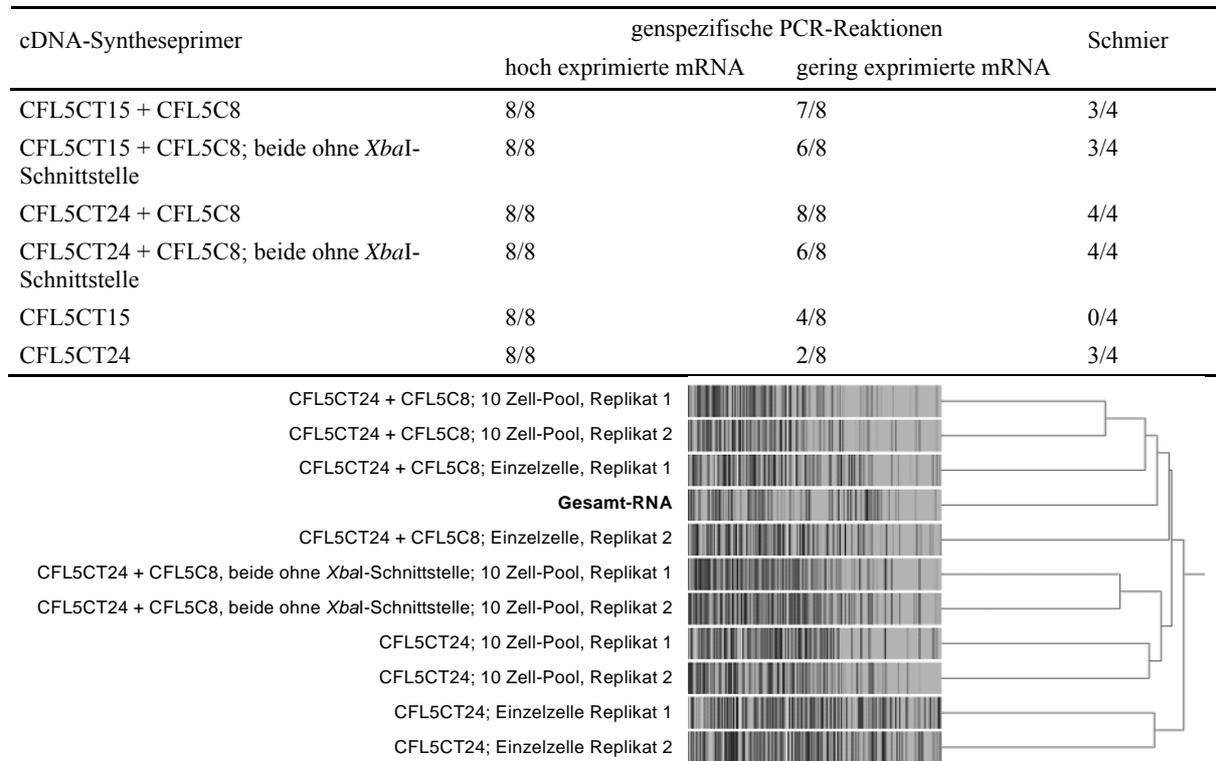


Abbildung 3.3 Vergleich verschiedener cDNA-Syntheseprimer

Zur cDNA-Synthese wurden Oligo dT-Primer mit unterschiedlicher Anzahl Thymidinnukleotide (CFL5CT) allein oder in Kombination mit einem Zufallsprimer (CFL5C8) verwendet. Bei manchen Primern war durch den Austausch einer Base die Schnittstelle für das Restriktionsenzym *XbaI* entfernt worden. Der Erfolg der cDNA-Amplifikation wurde über den Anteil positiver genspezifischer PCR-Reaktionen sowie die Elektrophorese des Primäramplifikats (Schmier) beurteilt. Die Qualität der cDNA-Amplifikate wurde nach der Hybridisierung auf einen Oligonukleotidarray durch einen Vergleich mit einer Gesamt-RNA-Präparation in einer hierarchischen Clusteranalyse ermittelt. Ähnliche Proben kommen bei diesem Verfahren in einem gemeinsamen Arm des Dendrogramms zu liegen. Für den Vergleich nach der Hybridisierung wurden nur Primer mit 24 Thymidin-Nukleotiden verwendet, da sie den T15-Primern nach den PCR-Ergebnissen überlegen schienen.

Optimal war die Kombination eines Primers mit 24 Thymidin-Nukleotiden mit einem Zufalls-Oktamer-Primer. Bereits die Modifikation nur einer Base, wie sie bei der Elimination der *XbaI*-Schnittstelle auftrat, verschlechterte das Ergebnis. Das Gleiche galt für die Verwendung eines Primers mit nur 15 Thymidin-Nukleotiden oder den Verzicht auf den Zufallsprimer. Bei Verwendung von Primern mit 24 Thymidin-Nukleotiden wurde zudem die Plateauphase der PCR zur Primäramplifikation im Schnitt elf Zyklen früher erreicht als bei Primern mit 15 Thymidin-Nukleotiden (24 vs. 35 Zyklen), was auf eine höhere Templatmenge in der PCR bedingt durch effektivere Isolation der mRNA hinweist.

Vom Hersteller der PNAs wird zur Verbesserung der mRNA-Isolation die Verwendung einer Protease zum Aufschluss und zur Inaktivierung von RNasen empfohlen. Da PNAs eine hohe Affinität nicht nur zu RNA sondern auch zu DNA besitzen, bestand die Befürchtung, dass bei Verwendung der Protease die Kernmembran zerstört werden könnte. Das hätte zu einem Verlust der genomischen DNA geführt,

die RNA-Präparation mit DNA verunreinigt und eine parallele Untersuchung von mRNA und genomischer DNA einer Zelle unmöglich gemacht. Daher wurde genomische DNA aus Einzelzellen, deren mRNA mit den beiden Protokollen bearbeitet worden war, isoliert und amplifiziert. Als Kriterium für eine erfolgreiche Gewinnung dienten genspezifische PCR-Reaktionen und die Hybridisierung der DNA in einer CGH (Komparative Genomische Hybridisierung). Wie Abbildung 3.4 zeigt, war eine CGH parallel zur mRNA-Isolation mit PNAs nicht nur möglich sondern homogener und damit qualitativ besser als beim Oligo dT-Bead-Protokoll. Die genomische DNA konnte auch aus einem höheren Prozentsatz der Proben erfolgreich isoliert werden. Der Zusatz der Protease erwies sich also als nicht störend für die parallele Untersuchung der genomischen DNA. Für die quantitative Isolation der mRNA war die Aktivität der Protease hingegen notwendig, da gering exprimierte Transkripte wie *TP53* sonst nach der Amplifikation nicht mehr nachweisbar waren. Proben, die nach der Isolierung ohne vorhergehende Aktivierung der Protease bei -80°C gelagert worden waren, ließen sich zum Teil überhaupt nicht mehr amplifizieren (Absatz 3.1.2).

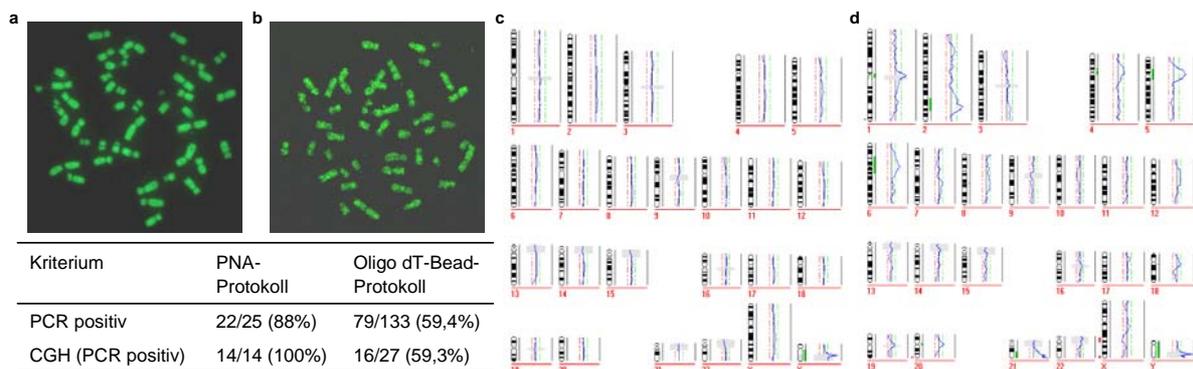


Abbildung 3.4 Isolation der genomischen DNA und CGH-Hybridisierung parallel zur Isolation und Amplifikation von mRNA mit beiden Protokollen

Genomische DNA aus dem PNA-Protokoll (a) zeigt in der CGH ein deutlich homogeneres Hybridisierungsverhalten als solche aus dem Oligo dT-Bead-Protokoll (b). Zudem werden aus diploiden Zellen mit dem PNA-Protokoll stabile balancierte CGH-Profile erhalten (c), wohingegen bei genomischer DNA aus dem Oligo dT-Bead-Protokoll die Profile unruhig sind (d). Die parallele Isolation von genomischer DNA und mRNA ist mit dem PNA-Protokoll außerdem häufiger möglich und eine erfolgreiche CGH-Hybridisierung aus den Ergebnissen genspezifischer PCR besser vorhersagbar.

Die Ursache für die Unterschiede hinsichtlich Qualität und Erfolgsraten bei der parallelen Isolation der genomischen DNA lag vermutlich in der unterschiedlichen Zusammensetzung der Puffer zur Zelllyse beider Protokolle. Die genomische DNA musste zur weiteren Aufarbeitung aus den Überständen, die bei der Immobilisation der mRNA-Bead-Konstrukte in einem Magnetfeld entstehen, ausgefällt und dann rückgelöst werden. Probleme können dabei durch nicht quantitative Fällung und/oder Rücklösung entstehen. Die Fällung kann durch den Zusatz eines Polyacrylamid (PAA)-Carriers oder von Salzen verbessert werden, das Rücklösen durch effizientes Waschen in 70%igem Ethanol zur Entfernung der Salze sowie durch längere Lösungszeiten. Einen Überblick über den Einfluss dieser Variablen auf die parallele Isolation von genomischer DNA gibt Tabelle 3.1.

Tabelle 3.1 Aufarbeitung der genomischen DNA nach Isolierung der mRNA

Verfahren zur parallelen Isolation der genomischen DNA	Positive PCR-Reaktionen
Fällung mit PAA-Carrier und 2,5 x Vol. Ethanol	6/12 (50,0 %)
Fällung mit PAA-Carrier und 2,5 x Vol. Ethanol; Rücklösen über Nacht in Wasser	9/12 (75,0 %)
Fällung mit PAA-Carrier, 0,01 M MgCl ₂ , 0,3 M Natriumacetat und 2,5 x Vol. Ethanol	8/12 (66,6%)
Fällung mit PAA-Carrier, 0,01 M MgCl ₂ , 0,3 M Natriumacetat und 2,5 x Vol. Ethanol; über Nacht bei 4°C in 70% Ethanol waschen	4/12 (33,3 %)
Fällung mit PAA-Carrier, 0,01 M MgCl ₂ , 0,3 M Natriumacetat und 1 x Vol. Isopropanol	9/12 (75,0 %)

Parallel zur Isolation der mRNA ließ sich die genomische DNA am besten aufarbeiten, wenn sie nur in Gegenwart des Carriers mit Ethanol ausgefällt und das Pellet dann über Nacht in Wasser rückgelöst wurde. Der Zusatz von Salzen erschwerte vor allem bei längeren Lagerzeiten die Rücklösung und führte dann ebenso wie zu langes Waschen in 70%igem Ethanol zu Verlusten.

3.1.2 Lagerung von Proben vor der Aufarbeitung

Das Protokoll bietet die Möglichkeit, Proben nach der Isolierung bis zur eigentlichen Amplifikation der cDNA bei -80°C zu lagern. Dies bedeutet insbesondere bei umfangreichen Verfahren zur Probenaufarbeitung eine deutliche Erleichterung. Es wurden verschiedene Stadien im Protokoll als Zeitpunkt für die Unterbrechung geprüft. Die Proben wurden bei -80°C gelagert und das Protokoll nach dem Auftauen der Probe an gleicher Stelle fortgesetzt. Die Ergebnisse sind in Abbildung 3.5 dargestellt.

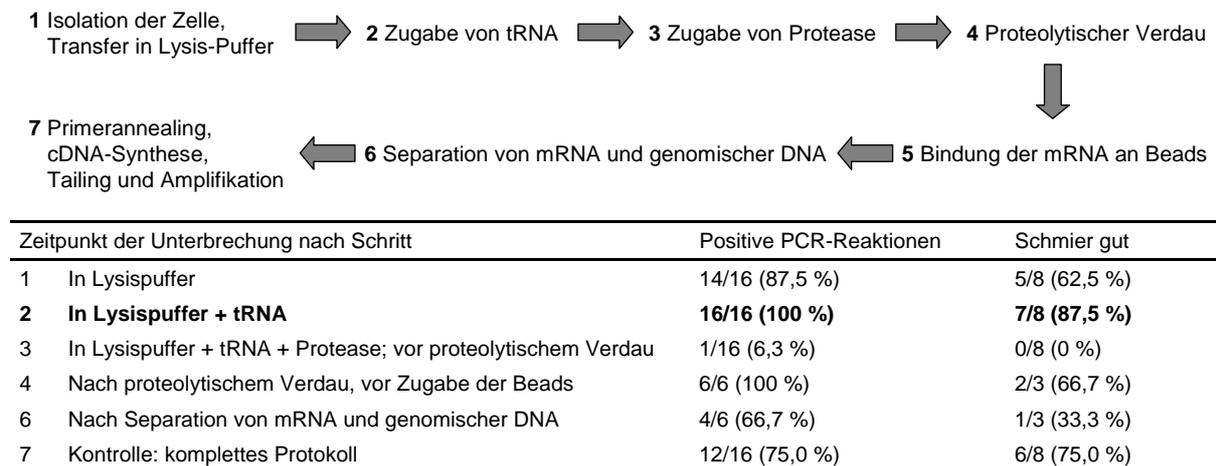


Abbildung 3.5 Lagerung von Proben bei -80°C in verschiedenen Stadien des Protokolls

Die Stadien 1-7 markieren Schritte vor allem zu Beginn des Protokolls (Abbildung 3.1), nach denen der Prozess unterbrochen und die Proben bis zur Fortsetzung bei -80°C gelagert wurden. Nach Vollendung des Protokolls wurden genspezifische PCR-Reaktionen und eine unspezifische Gelelektrophorese des Primäramplifikats ausgewertet.

Der kritische Schritt bestand offensichtlich im proteolytischen Verdau, der direkt nach der Zugabe der Protease erfolgen musste, um deren Aktivität nicht durch Einfrieren zu reduzieren und eine reproduzierbare erfolgreiche Amplifikation der mRNA zu ermöglichen. Mittels quantitativer real time

PCR (qPCR) wurde geprüft, ob während des Einfrierens, der Lagerung und des Auftauens mRNA zerstört wird. Dazu wurden jeweils acht Proben pro Ansatz auf die Expression von sieben Genen relativ zum Haushaltsgen *Gapdh* untersucht. Proben, die in Lysispuffer gelagert worden waren, erreichten im Durchschnitt 84,3% der Expression der Kontrolle. Mit Zusatz von tRNA erhöhte sich der Mittelwert auf 102,2% der Kontrolle, nach dem proteolytischen Verdau wurden 109,8% und nach der Separation von mRNA und genomischer DNA 84,2% gemessen.

Isolierte Zellen konnten also am besten in Lysispuffer mit tRNA oder nach dem proteolytischen Verdau einige Tage bei -80°C gelagert und später weiter verarbeitet werden. Längere Lagerzeiten waren wegen der dann einsetzenden Degradation der mRNA nicht empfehlenswert. Für die parallele Isolation der genomischen DNA bestand kein signifikanter Unterschied zwischen den Zeitpunkten der Unterbrechung. Um auch die Untersuchung von Proben aus fremden Labors mit einem einheitlichen Protokoll zu ermöglichen, wurde die Lagerung in Lysispuffer mit tRNA gewählt.

3.1.3 Validierung des Amplifikationsprotokolls

Eine wichtige Frage im Zusammenhang mit der exponentiellen Primäramplifikation der cDNA war, ob sich während der PCR die relativen Verhältnisse zwischen einzelnen Transkripten verschieben. Dieser systematische Fehler wurde für alle Amplifikationsprotokolle immer wieder angenommen und für die so genannte „lineare Amplifikation“ nach dem Eberwine Verfahren gezeigt^{75, 134, 135}, auch wenn andere Publikationen weder für lineare noch für exponentielle Amplifikation Verzerrungen feststellten^{136, 137}.

Um etwaige Verschiebungen der relativen Transkriptmengen während der Primär-amplifikation möglichst auszuschließen, wurde versucht, die PCR noch vor Erreichen der Plateauphase zu beenden. In der exponentiellen Phase sollte die Gefahr der Verzerrung, die beispielsweise durch Substratmangel und eine dadurch bedingte bevorzugte Amplifikation bestimmter Transkripte im Plateau auftreten könnte, reduziert sein. Die exponentielle Phase wurde für jede Probe experimentell bestimmt, indem in regelmäßigen Abständen nach der Elongationsphase des jeweiligen PCR-Zyklus ein Teil des Produktes entnommen wurde. Nach einer Gelelektrophorese war die exponentielle Phase als der Bereich definiert, in dem der Schmier der PCR-Produkte noch mit jedem Messpunkt stärker wurde. Von einer Probe wurden zudem ein Amplifikat der exponentiellen als auch der Plateauphase auf Oligonukleotidarrays hybridisiert und die Korrelationskoeffizienten zu einer Hybridisierung von nicht amplifizierter Gesamt-RNA berechnet (Tabelle 3.2).

Tabelle 3.2 Korrelation von Proben verschiedener Phasen der Primäramplifikation zur Gesamt-RNA

Probentyp	Exponentielle Probe	Probe aus der Plateauphase
Einzelzelle	0,498	0,502
Pool aus 10 Zellen	0,579	0,586
Pool aus 100 Zellen	0,645	0,656

Da die Korrelation der Plateauproben in allen Fällen geringfügig höher war als die der exponentiellen Proben, wurde die globale Amplifikation fortan im Plateau beendet.

Zur näheren Überprüfung der Transkriptverhältnisse während der Amplifikation, einer etwaigen Abhängigkeit von der Menge der Ausgangs-RNA, sowie der Reproduzierbarkeit des Protokolls wurden verschiedene Proben typen mittels qPCR untersucht. Einzelzellen, Pools aus zehn Zellen und Einzelzelläquivalente zur Bestimmung der technischen Variabilität wurden dem Amplifikationsprotokoll bis in die Plateauphase unterworfen. Parallele Ansätze wurden nur revers transkribiert und nicht global amplifiziert. Aus allen Produkten wurden genspezifische qPCR-Reaktionen durchgeführt und die Transkriptmengen relativ zu *Gapdh* bestimmt. Bei nicht amplifizierten Proben waren maximal zehn Bestimmungen möglich, da die überwiegende Mehrheit der Transkripte nur in fünf bis zehn Kopien pro Zelle vorliegt und ein Zelllysate daher nur in zehn identische Ansätze aufgeteilt werden kann⁶⁹.

Es zeigte sich, dass während der globalen Amplifikation keine Verschiebung relativer Transkriptverhältnisse stattfand, kein Unterschied in der Sensitivität zwischen einer und zehn Zellen bestand und die Reproduzierbarkeit hoch war (Abbildung 3.6). Sämtliche Messwerte waren dem Mittelwert vergleichbar, unabhängig von Proben typ und Methode. Die Einzelzellen zeigten wegen der zellulären Heterogenität erwartungsgemäß die höchste Variabilität. Dieser Effekt schien durch die Amplifikation bei einigen Genen verstärkt zu werden (*Catnb* und *v-Erb-b2*). Außerdem führte die Amplifikation zu einem Sensitivitätsgewinn, da einige Transkripte in den nichtamplifizierten Proben auch in 10 Zell-Pools nicht nachweisbar waren, wohl aber in einzelnen amplifizierten Zellen.

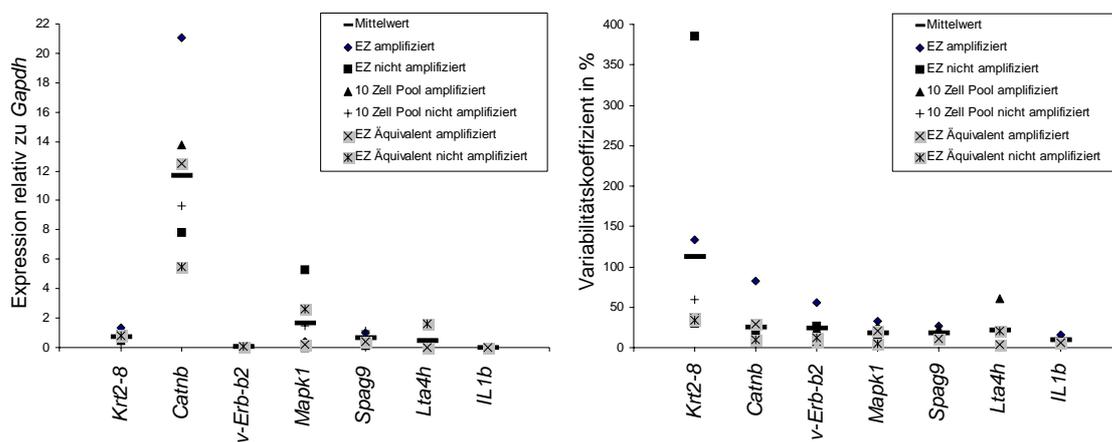


Abbildung 3.6 Reproduzierbarkeit und Variabilität des Amplifikationsprotokolls im Vergleich zu nicht amplifizierten Proben

Dargestellt sind die durch qPCR bestimmten relativen Expressionswerte aller untersuchten Gene zum Haushaltsgen *Gapdh* für sämtliche Proben typen (links), sowie die Variabilitätskoeffizienten der Replikate (rechts). Neben dem Mittelwert aller Proben sind jeweils die Einzelmesswerte aufgetragen. Von den gezeigten Genen waren *Spag9*, *Lta4h* und *IL1b* in den nicht amplifizierten Proben überhaupt nicht oder nur teilweise in den Zehn-Zell-Pools nachweisbar. EZ: Einzelzelle

3.2 Auswahl eines Arrays zur Analyse der Genexpression

Nächstes Ziel dieser Arbeit war die Etablierung eines Protokolls zur fluoreszenzbasierten Hybridisierung der Einzelzell-cDNA-Amplifikate auf Genexpressionsarrays für eine quantitative Analyse aller Transkripte. Es sollte ein Protokoll für die Hybridisierung auf Oligonukleotidarrays erarbeitet werden, die zum Zeitpunkt der Arbeit die fortschrittlichste Technologie darstellten^{83, 85}. Zudem neigen Arrays älterer Generation, deren Sonden aus klonierter cDNA bestehen, vor allem bei der Untersuchung von Einzelzellamplifikaten verstärkt zu Fehlhybridisierungen. Die bei der Amplifikation der cDNA verwendeten Enzympräparationen werden meist durch rekombinante DNA-Technologien in Bakterien hergestellt und enthalten daher nicht unerhebliche Mengen bakterieller DNA als Verunreinigung. Diese Sequenzen werden im Verlauf des Protokolls ebenfalls amplifiziert und markiert und führen durch Hybridisierungen mit den bakteriellen Sequenzen der Sonden des Arrays zu falsch-positiven Ergebnissen.

3.2.1 Affymetrix GeneChip-Arrays

Bei Beginn der Arbeit kamen nur GeneChip-Arrays der Firma Affymetrix in Betracht. Es handelte sich um ein Oligonukleotid-System, das auf jeweils zwei Arrays das komplette menschliche Genom erfasste. Um GeneChip-Arrays verwenden zu können, muss jedoch in jede cDNA ein T7-Promotor integriert werden. Die Protokolle von Affymetrix sehen eine Hybridisierung von cRNA vor, die im Rahmen der Markierungsreaktion über eine *in vitro*-Transkription mit der T7 RNA-Polymerase gebildet wird.

Einbau des T7-Promotors über den Primer während der cDNA-Synthese

Die nahe liegende Überlegung war, in Anlehnung an das Standardverfahren, bei dem der Promotor über den cDNA-Syntheseprimer eingeführt wird⁷⁰, für das laboreigene Amplifikationsprotokoll ebenfalls einen cDNA-Syntheseprimer mit T7-Promotorsequenz zu verwenden, der dann in der PCR mit amplifiziert würde. Die Promotorsequenz besteht aus einer Bindungsdomäne für die RNA-Polymerase im 5'-Bereich und der Transkriptionsinitiationsdomäne im 3'-Bereich, deren Längen variiert wurden, um einen optimalen Primer hinsichtlich Einbau in die cDNA und Effektivität in der *in vitro*-Transkription zu erhalten. Neben diesem eigentlichen T7-Promotoranteil enthielten sämtliche getesteten cDNA-Syntheseprimer die Oligo dT-Sequenz zur Bindung an den Poly A-Schwanz der mRNA und die Oligo dC-Sequenz als späterer Primerbindungsstelle für die PCR aus dem konventionellen Protokoll. Es wurden Experimente mit zwei bis drei technischen Replikaten für jede Protokollvariation angelegt. Neben den in Absatz 3.1.1 beschriebenen Kontrollen für Isolation und Amplifikation wurde eine *in vitro*-Transkription zur Bestimmung der Aktivität des Promotors durchgeführt und die Produkte auf Nylonarrays hybridisiert.

Als Schwierigkeiten wurden im Besonderen erkannt:

- Um eine ausreichende Effizienz des Promotors in der *in vitro*-Transkription zu erreichen, mussten über die Core-Sequenz TAA TAC GAC TCA CTA TA hinaus vor allem noch weitere Nukleotide in der Transkriptionsinitiationsdomäne im 3'-Bereich vorhanden sein.
- Zusammen mit den obligatorischen Oligo dT- und Oligo dC-Sequenzen erreichten diese T7-Primer Längen von bis zu 80 Basen. Das führte zu schlechten Primingraten bei der cDNA-Synthese durch Bildung von Sekundärstrukturen und einem erschwerten Zugang zum Poly A-Schwanz. Daraus resultierten keine oder nicht reproduzierbare Ergebnisse nach der Amplifikation von cDNA aus Einzelzellen. Bei größeren Mengen Ausgangs-RNA (ab etwa 10-100 Zellen) waren die Ergebnisse besser, aber auch nicht zufrieden stellend.

Um Abhilfe zu schaffen, wurden Primer-Modifikationen untersucht (Abbildung 3.7):

- Die Anlagerung von kurzen Block-Oligonukleotiden an den T7-Primer in einem vorgelagerten Schritt sollte die Bildung von Sekundärstrukturen unterdrücken. Es wurden komplementäre Oligonukleotide zur Oligo dC-Sequenz (G9-BLOCK) und zur Core-Promotorsequenz (BLOCK 1 und 2) allein und in Kombination getestet. Die Ausbeuten an cDNA in diesen Experimenten gingen jedoch noch weiter zurück. Von den geblockten Primern wurde daher Abstand genommen (V2).
- Es wurde geprüft, ob Zufallsbasen (VN) am Ende der cDNA-Synthesepriemer notwendig sind oder zur verstärkten Bildung von Sekundärstrukturen beitragen. Die Basen stellten sich als essentiell für eine repräsentative Amplifikation mit guter Ausbeute heraus und wurden daher beibehalten (V3).
- Als ubiquitäre Primerbindungsstelle für die PCR zur Primäramplifikation wurde die LIB-Sequenz statt der Oligo dC-Sequenz verwendet (V4).

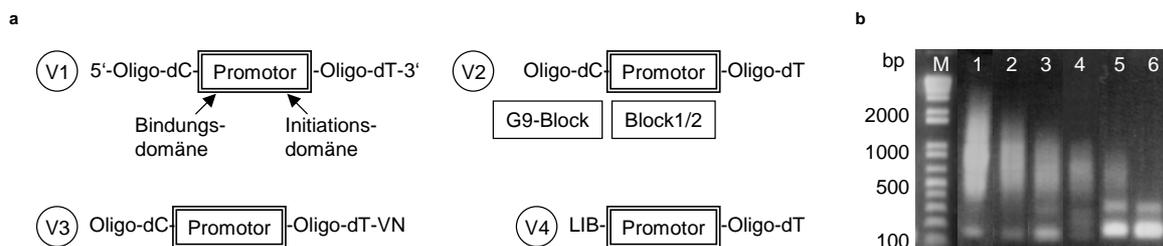


Abbildung 3.7 Variation der cDNA-Synthesepriemer mit T7-Promotor und Beurteilung der Amplifikation anhand einer Gelelektrophorese des Primäramplikats

a Die cDNA-Synthesepriemer mit T7-Promotor wurden variiert: V1: Im Promorteil, um ein ausgewogenes Verhältnis von Primerlänge und den Bindungs- und Initiationsdomänen zu finden. V2: Mit Block-Oligonukleotiden für verschiedene Primerbereiche zur Unterdrückung von Sekundärstrukturen. V3: Mit Zufallsbasen am 3'-Ende. V4: Mit einer abweichenden Sequenz für die ubiquitäre Primerbindungsstelle. **b** Die Qualität der cDNA-Amplifikate wurde anhand einer Gelelektrophorese des Produktes von 1 (sehr gut) über 2 (gut), 3 (befriedigend), 4 (ausreichend), 5 (mangelhaft) bis 6 (ungenügend) beurteilt (vgl. Tabelle 3.3).

Die Ergebnisse einer Auswahl der Experimente zeigt Tabelle 3.3, die Primersequenzen sind in Abschnitt 2.5 des Methodenteils aufgeführt. Der Einbau der Primer während der cDNA-Synthese

wurde über den Anteil positiver genspezifischer PCR-Reaktionen sowie den positiven Anteil und die Qualität der Elektrophorese des primären Amplifikationsproduktes beurteilt. Die Bewertung dieser Ergebnisse erfolgte auf einer sechsstufigen Skala von „sehr gut“ bis „ungenügend“ analog zu Abbildung 3.7 b. Für die genspezifischen PCR-Reaktionen wurden unterschiedlich stark exprimierte Transkripte sowie unterschiedliche Amplikonlängen (EF1- α : 280 bp, β -AKTIN: 500 bp) benutzt. Außerdem wurde die Effizienz des Promotors in der *in vitro*-Transkription als Verhältnis der gebildeten Menge cRNA-Produkt zur eingesetzten Menge cDNA-Template bestimmt.

Tabelle 3.3 Einbau und Effizienz verschiedener T7-Promotorprimer

cDNA-Syntheseprimer	hoch exprimiert		niedrig exprimiert		Amplifikation; Qualität	Effizienz in der IVT ^a
	EF1- α	β -AKTIN	TP53_1	TP53_2		
T7TA-T24-VN	8/12 (67 %)	4/12 (33 %)	0/9 (0 %)	0/9 (0 %)	3/12 (25 %); ungenügend	53-fach
T7TA-T24-VN + BLOCK2	3/3 (100 %)	0/3 (0 %)	0/3 (0 %)	0/3 (0 %)	0/3 (0 %)	n.a. ^b
T7TA-T24-VN + G9-BLOCK	3/3 (100 %)	3/3 (100 %)	0/3 (0 %)	1/3 (33 %)	2/3 (67 %); mangelhaft	n.a.
T7TA-T24-VN + G9-BLOCK + BLOCK2	2/6 (33 %)	3/6 (50 %)	0/6 (0 %)	0/6 (0 %)	0/6 (0 %)	n.a.
T7TC-T24	1/3 (33 %)	0/3 (0 %)	0/3 (0 %)	0/3 (0 %)	0/3 (0 %)	n.a.
T7TC-T24 + BLOCK2	3/3 (100 %)	0/3 (0 %)	0/3 (0 %)	0/3 (0 %)	2/3 (67 %); mangelhaft	3-fach
T7TC-T24-VN	28/40 (70 %)	25/40 (63 %)	16/38 (42 %)	11/40 (28 %)	23/40 (56%); ausreichend	7,2-fach
T7TC-T24-VN + BLOCK2	3/3 (100 %)	3/3 (100 %)	0/3 (0 %)	0/3 (0 %)	3/3 (100 %); ausreichend	2,5-fach
T7TC-T24-VN + G9-BLOCK	8/9 (89 %)	6/9 (67 %)	0/9 (0 %)	0/9 (0 %)	5/9 (56 %); ausreichend	n.a.
T7TC-T24-VN + G9-BLOCK + BLOCK2	3/6 (50 %)	3/6 (50 %)	0/6 (0 %)	1/6 (17 %)	0/6 (0 %)	n.a.
T7TE-T24-VN	13/18 (72 %)	10/18 (56 %)	8/18 (44 %)	3/18 (33 %)	1/18 (6 %) ungenügend	14-fach

^a *in vitro*-Transkription
^b nicht analysiert

Der relativ kurze Primer T7TC-T24-VN ließ sich am besten während der cDNA-Synthese einführen und in der PCR amplifizieren, erreichte aber unzureichende Effizienzen in der *in vitro*-Transkription. Lange Primer wie T7TA-T24-VN, die in der *in vitro*-Transkription eine hohe Effizienz zeigten, störten jedoch die reverse Transkription und Amplifikation von mRNA aus Einzelzellen. Dies äußerte sich in negativen PCR-Reaktionen vor allem bei gering exprimierten Genen und längeren Amplika sowie nicht zufrieden stellenden Ergebnissen in der Elektrophorese des primären Amplifikationsproduktes. Daher gelang es nicht, einen Primer zu finden, der sich reproduzierbar zu cDNA von guter Qualität amplifizieren ließ und gleichzeitig ausreichend cRNA während der *in vitro*-Transkription generierte. Das Protokoll ohne T7-Sequenz im cDNA-Syntheseprimer blieb in ähnlichen Versuchen mit PCR- und Schmier-Ergebnissen von fast 100% und daher sehr guten Einbauraten ohne Fragmentierung

deutlich überlegen. Diese Ergebnisse wurden auch durch Filterhybridisierungen bestätigt. Die T7-Proben lieferten nur selten Signale, unabhängig davon ob nach einer *in vitro*-Transkription cRNA oder nach einem PCR-Labeling über den ubiquitären Primer cDNA hybridisiert wurde.

Weiterreichende Veränderungen am gesamten Protokoll wurden mit dem Ziel untersucht, trotzdem eine ausreichende Amplifikation mit den T7-Primern zu erreichen. Dazu wurde geprüft, ob sich ein Schritt des Standardprotokolls durch Optimierung effizienter gestalten ließ (Abbildung 3.8):

- Der mRNA-Isolationsschritt wurde verändert. Die PNAs wurden zuerst 30 min an die Beads gekoppelt und dann die Zelle im Lysispuffer/Protease-Mix zugegeben. Nach dem proteolytischen Verdau zum Aufschluss der Sekundärstrukturen konnten die Bead-PNA-Konstrukte an die mRNA binden (V5).
- Die Dauer der Bindung der mRNA an die Beads variierte zwischen 30 min und 12 h (V6).
- Die Primer für die cDNA-Synthese wurden bereits während der Bindung der mRNA an die Beads oder im Rahmen eines Temperaturgradienten in der PCR-Maschine angelagert (V7).
- Statt der Tailingreaktion wurde zur Schaffung einer zweiten Primerbindungsstelle für die PCR zur Primäramplifikation eine Zweitstrangsynthese mit Klenow-Enzym oder Sequenase mit und ohne anschließenden Verdau einzelsträngiger cDNA mit Mungbean-Nuklease durchgeführt. Als Primer für die Zweitstrangsynthese dienten 6er, 8er und 10er Zufallsoligonukleotide verschiedener Konzentration. Diese wurden bei Raumtemperatur oder während eines Temperaturgradienten in der PCR-Maschine an den ersten Strang cDNA angelagert (V8).
- Die Primer für cDNA-Synthese und Zweitstrangsynthese enthielten statt der Oligo dC- die LIB-Sequenz als universelle Primerbindungsstelle für die PCR zur Primäramplifikation (V9).

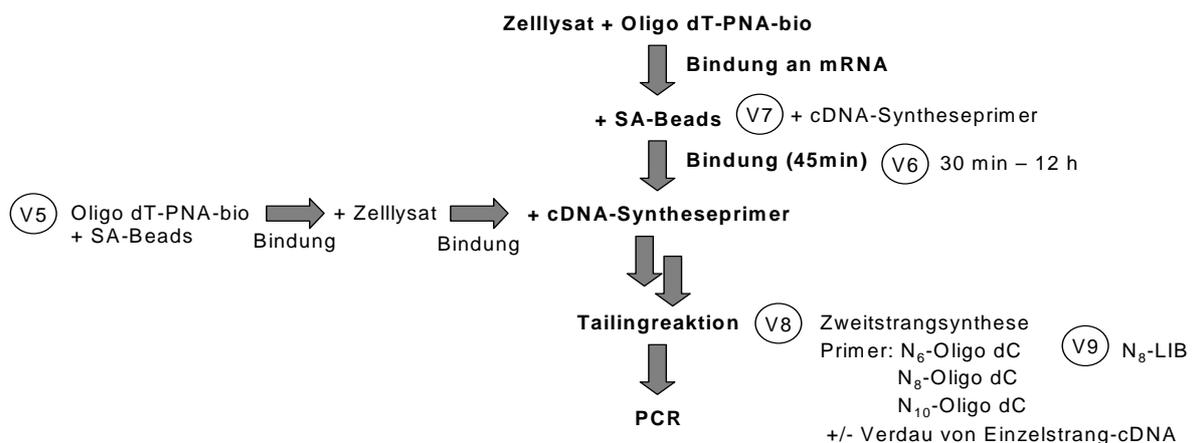


Abbildung 3.8 Varianten des Protokolls zur Amplifikation von mRNA aus Einzelzellen

Fett dargestellt ist der Ablauf des Standardprotokolls. Rechts und links sind in normaler Schrift die Varianten der entsprechenden Schritte im Protokoll angegeben, die mit dem Ziel der Optimierung des Verfahrens geprüft wurden. V5 variierte den Isolationsschritt der mRNA, V6 die Dauer der Isolation, V7 den Zeitpunkt der Zugabe der cDNA-Syntheseprimer, V8 die Bildung der zweiten Primerbindungsstelle für die PCR und V9 deren Sequenz. bio: Biotin, SA: Streptavidin

Alle diese Versuche brachten nicht den gewünschten Erfolg hinsichtlich des Einbaus eines effizienten T7-Promotors, und die Modifikationen schienen dem ursprünglichen Protokoll allgemein nicht überlegen zu sein (Tabelle 3.4). Der direkte Einbau eines T7-Promotors während der cDNA-Synthese wurde daher nicht weiter verfolgt, am Standardprotokoll wurden keine Modifikationen vorgenommen.

Tabelle 3.4 Ergebnisse der Protokollvarianten V5bis V9 (Abbildung 3.8)

Variante	hoch exprimiert		niedrig exprimiert		Amplifikation
	<i>EF1-α</i>	<i>β-AKTIN</i>	<i>TP53_1</i>	<i>TP53_2</i>	
V5	8/8 (100 %)	8/8 (100 %)	3/8 (38 %)	2/8 (25 %)	0/8 (0 %)
V6	11/15 (73 %)	3/15 (20 %)	2/3 (67 %)	3/3 (100 %)	0/3 (0 %)
V7	4/4 (100 %)	2/4 (50 %)	1/4 (25 %)	2/4 (50 %)	0/4 (0 %)
V8	N6-Oligo dC	1/1 (100 %)	0/1 (0 %)	0/1 (0 %)	0/1 (0 %)
	N8-Oligo dC	8/8 (100 %)	2/8 (25 %)	1/8 (13 %)	0/8 (0 %)
	N10-Oligo dC	1/1 (100 %)	1/1 (100 %)	0/1 (0 %)	0/1 (0 %)
V9	3/4 (75 %)	0/4 (0 %)	0/4 (0 %)	0/4 (0 %)	0/4 (0%)
Standardprotokoll	8/8 (100%)	8/8 (100%)	8/8 (100%)	7/8 (88%)	8/8 (100%)

Nachträglicher Einbau des T7-Promotors in Primäramplifikate

Eine Alternative zur Einführung des T7-Promotors während der cDNA-Synthese bestand in dessen Einbau in fertige cDNA-Primäramplifikate aus dem konventionellen Protokoll. Die hierzu untersuchten Möglichkeiten waren (Abbildung 3.9):

- Eine Reamplifikations-PCR mit einem Primer, der neben der Oligo dC-Sequenz die Promotorsequenz enthielt (V10)
- Annealing eines Oligonukleotids, das neben der Oligo dC-Sequenz die Promotorsequenz enthielt, an das Amplifikationsprodukt gefolgt von einer Auffüllreaktion, um den aktiven doppelsträngigen Promotor zu bilden (V11),
- die direkte Ligation eines doppelsträngigen T7-Promotorkonstruktes an das Amplifikationsprodukt (V12) und
- die Ligation eines einzelsträngigen T7-Promotorkonstruktes an das Amplifikationsprodukt nach Entfernen des ubiquitären Primers durch Restriktionsverdau mit *XbaI* gefolgt von einer Auffüllreaktion. Besitzt nur der Oligo dT-Primer für die cDNA-Synthese die Schnittstelle für das Restriktionsenzym, der Zufallsprimer jedoch nicht, enthalten nur cDNA-Fragmente aus dem 3'-Bereich der mRNA am Ende einen Promotor. Dies entspricht der Situation nach Verwendung der Protokolle der Firma Affymetrix bzw. dem Einbau des Promotors während der cDNA-Synthese unter Verwendung eines Oligo dT-Primers mit Promotorsequenz (V13).

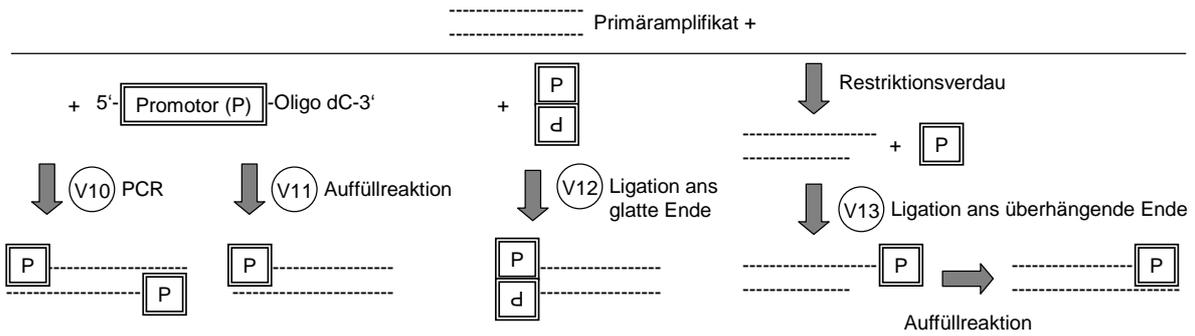


Abbildung 3.9 Möglichkeiten zur Einführung eines T7-Promotors in fertige cDNA-Primäramplifikate

Ausgehend von Primäramplifikaten aus dem konventionellen Protokoll bieten sich unterschiedliche Möglichkeiten, einen T7-Promotor einzuführen: V10: Durch Reamplifikation mit einem Primer, der die Promotorsequenz enthält. V11: durch Anlagerung eines solchen Primers gefolgt von einer Auffüllreaktion zur Bildung eines aktiven doppelsträngigen Promotors. V12: Durch Ligation eines doppelsträngigen Promotors an das Primäramplifikat. V13: Durch Ligation einer einzelsträngigen Promotorsequenz an das Primäramplifikat nach Restriktionsverdau, gefolgt von einer Auffüllreaktion.

Nach ersten Filterhybridisierungen schien die letzte Methode (V13) am viel versprechendsten. Sie bot außerdem den Vorteil, dass sich der Promotor nur in einem Strang der cDNA befand und während der Markierung nur von diesem Strang *in vitro*-transkribiert wurde. Anderenfalls wären mit dem für Affymetrix GeneChip-Arrays ungewöhnlichen Gemisch aus sense und antisense cRNA-Strängen verstärkt Kreuzhybridisierungen zu befürchten gewesen. Durch die Auffüllreaktion wurde zudem gewährleistet, dass sich keine freien doppelsträngigen und damit aktiven Promotorsequenzen im Ansatz befanden, die die Effizienz der *in vitro*-Transkription hätten vermindern können (V12).

Das Protokoll wurde zum im Methodenteil beschriebenen Verfahren optimiert und auf Proben angewendet, die in Kooperation mit dem Labor von Dr. Michael Hummel (Berlin) auf GeneChip-Arrays hybridisiert wurden. Die Ergebnisse zweier technischer Replikate sind in Abbildung 3.10 dargestellt.

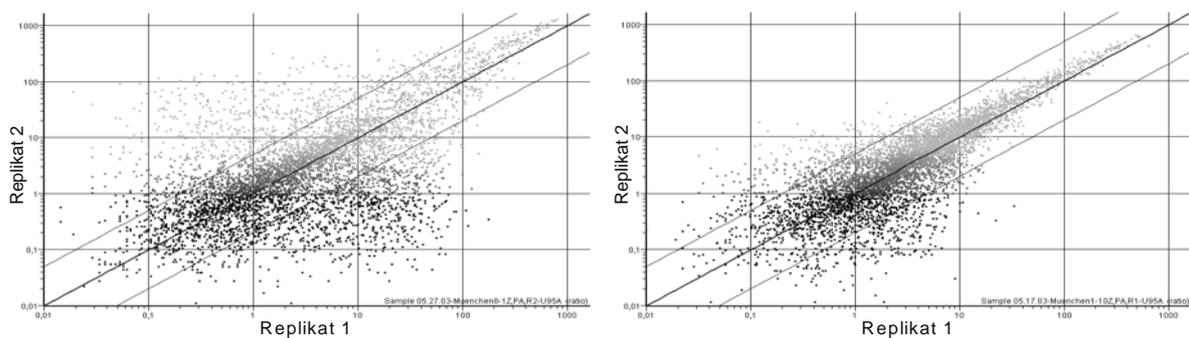


Abbildung 3.10 Scatterplots von Hybridisierungen auf Affymetrix GeneChip-Arrays

Dargestellt sind technische Replikate des Gesamtprotokolls von Einzelzellen (links) und 10 Zell-Pools (rechts). Die Signalintensitäten jeweils eines der Replikate sind auf der x- bzw. y-Achse abgetragen. Je besser die Korrelation der Replikate, desto mehr nähert sich die Punktelcke der Winkelhalbierenden.

Vor allem im Vergleich mit den Experimenten zur Einführung des T7-Promotors während der cDNA-Synthese aber auch mit den damals gültigen Empfehlungen der Firma Affymetrix zur Verwendung

reiner T7-basierter Systeme ohne PCR-Amplifikation nicht unter 100-1000 Zellen, waren diese Ergebnisse viel versprechend. Bei der Hybridisierung von 10 Zell-Pools war auch der für GeneChip-Arrays typische Effekt, dass geringe Signalintensitäten mit großen Streuungen einhergehen, wohingegen sich die Punktwolke bei höheren Intensitäten verschmälert gut zu beobachten. Trotzdem wurde auf die Hybridisierung von GeneChip-Arrays und die weitere Optimierung des Protokolls für die Verwendung eines T7-basierten Systems verzichtet, da sich mittlerweile Alternativen anderer Hersteller ergeben hatten. Diese Arrays arbeiten mit cDNA und benötigen daher keinen T7-Promotor im Template, was zu einer Vereinfachung des Verfahrens und dadurch zu einer Reduktion der Variabilität beitrug, außerdem konnten sie ohne Kooperationspartner hybridisiert werden.

3.2.2 Febit-Arrays

Ein System zur Hybridisierung von cDNA-Amplifikaten ohne T7-Promotor kam von der Firma Febit. Dabei handelte es sich um *in situ*-synthetisierte Oligonukleotidarrays, bei denen jeder Array speziell für das jeweilige Experiment zusammengestellt und selbst synthetisiert werden musste. Das System war daher sehr flexibel, konnte allerdings nicht alle bekannten Transkripte in einer einzigen Hybridisierung erfassen.

Zu Testzwecken wurden Einzelzell-cDNA-Amplifikate hergestellt, vom Unternehmen hybridisiert und die Rohdaten übermittelt. Bei der Untersuchung von Replikaten stellte sich heraus, dass bei den Arrays keine Chargenkonstanz gegeben war. Vier Proben wurden auf Arrays aus zwei verschiedenen Produktionen hybridisiert, Probe eins und drei auf Charge eins, Probe zwei und vier auf Charge zwei. Nach der Hybridisierung unterschieden sich die Intensitätsverteilungen der normalisierten Expressionswerte aller Sonden des Arrays je nach verwendeter Charge deutlich, obwohl bei cDNA aus Replikaten eigentlich sehr ähnliche Ergebnisse zu erwarten gewesen wären (Abbildung 3.11). Der Array nahm also in starker Weise Einfluss auf das Ergebnis der Hybridisierung, was nicht wünschenswert ist.

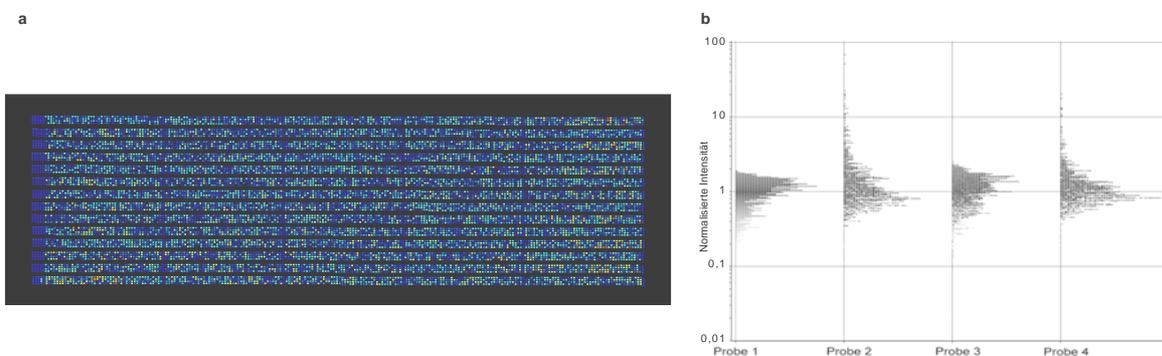


Abbildung 3.11 Hybridisierung von Einzelzellamplifikaten auf Febit-Arrays

a Scan eines Febit-Arrays nach Hybridisierung eines Einzelzellamplikats. **b** Intensitätsverteilung der normalisierten Expressionswerte aller erfassten Gene. Probe 1 und 3 wurden auf Arrays einer Charge hybridisiert, Probe 2 und 4 auf Arrays einer zweiten.

Zur weiteren Bewertung der Ergebnisse sollte ein Vergleich der Hybridisierungsergebnisse mit den Resultaten genspezifischer PCR-Reaktionen als unabhängiger Methode dienen. Dazu wurden für jede der vier Proben 25 Transkripte untersucht. Der Vergleich gestaltete sich jedoch als schwierig, da der Bereich zuverlässiger Messwerte für die Hybridisierung unbekannt war. Aufgrund von Autofluoreszenz, Hintergrund und ähnlichen Effekten waren nach der Hybridisierung für alle Gene Fluoreszenzsignale messbar, die jedoch erst ab einer gewissen Intensität als „gültig“ einzustufen sind (Abbildung 3.11 a). Es wurden verschiedene Schwellenwerte anhand von mittleren Signalintensitäten und der Streuung der Messwerte berechnet. Im optimalen Fall ergaben sich übereinstimmende Ergebnisse zwischen Hybridisierung und PCR in 58,3% der Untersuchungen, positive Ergebnisse auf dem Array bei negativer PCR („Array falsch-positiv“) in 12,5% und negative Hybridisierungssignale bei positiver PCR („Array falsch-negativ“) in 29,2% der Fälle. Das System wurde daher wegen fehlender Sensitivität und Spezifität, geringer Chargenkonstanz und mangelhafter Anwenderfreundlichkeit als nicht geeignet eingestuft und als Kandidat für den Routineeinsatz ausgeschlossen.

3.2.3 Memorec PIQOR-Arrays

Beim dritten System im Test handelte es sich um kommerzielle kleinere cDNA-Arrays, die je nach gewähltem Typ unterschiedliche zelluläre Funktionen abdeckten, also ebenfalls keine Aussage über alle bekannten Transkripte in einem Experiment erlaubten. Die ersten Hybridisierungsergebnisse waren nicht zufrieden stellend, da der Hintergrund auf den Arrays sehr hoch war. Zusätzlich stellten cDNA-Arrays wegen verstärkter Neigung zu Fehlhybridisierungen durch kontaminierende bakterielle und Plamidsequenzen tendenziell eine veraltete Technologie und kaum eine Verbesserung im Vergleich mit cDNA-Nylonfiltern dar. Daher wurde auch diese Option verworfen.

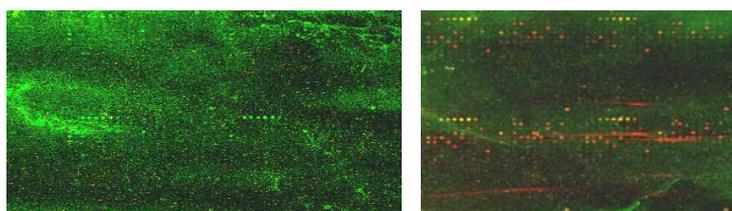


Abbildung 3.12 Hybridisierung von zwei Einzelzellamplifikaten auf Memorec PIQOR-Arrays

Beide Hybridisierungen weisen starke und zudem nicht konstante Hintergrundeffekte auf. Das Verhältnis von Spotsignal zu Hintergrund reicht für eine quantitative Auswertung der Hybridisierung nicht aus.

3.3 Hybridisierung auf murine Operon Siebziger Oligonukleotidarrays

3.3.1 Etablierung des Protokolls

Ein weiteres mögliches Hybridisierungssystem waren Siebziger Oligonukleotide der Firma Operon, die vom Anwender auf einem beschichteten Objektträger eigener Wahl immobilisiert werden müssen. Es existieren Oligonukleotid-Sätze für zahlreiche Spezies - für die Fragestellungen der Arbeitsgruppe

sind humane und murine Sequenzen von Interesse - die von Operon nach den gleichen Kriterien entworfen und produziert werden. Daher schien es unerheblich zu sein, ob eine Etablierung der Hybridisierungsbedingungen am humanen oder Maussystem erfolgen würde. Anfangs waren murine Arrays für Etablierung und Validierung des Protokolls verfügbar, die im Labor von Jan Vijg in San Antonio, Texas hergestellt wurden und etwa 17'000 Maussequenzen umfassten.

Genexpressionsanalysen von Mauseinzelzellen stellen wegen des geringeren RNA-Gehaltes muriner Zellen im Vergleich zu humanen erhöhte Anforderungen an die Sensitivität des Verfahrens. Die Etablierung der Bedingungen auf Mausarrays ermöglichte daher direkt Aussagen über die Leistungsfähigkeit von Array und Protokoll bei anspruchsvollen Proben. Der Probenpool für die Testexperimente musste allerdings von human auf murin umgestellt werden, der Vergleich mit dem Febit-, Memorec- und Affymetrixsystem konnte nicht mehr wie bisher mit identischen Proben erfolgen.

Auf Operon-Arrays können im Gegensatz zu den Systemen der Firmen Affymetrix und Febit zwei unterschiedlich markierte Proben in Konkurrenz zueinander hybridisiert werden. Diese so genannte Zweifarben-Technik wurde unter Verwendung einer universellen Referenz aus gemischten cDNA-Amplifikaten verschiedener Gewebe wie Mamma, Lunge und Leber sowie cDNA aus Zelllinien benutzt, was eine interne Standardisierung jeder Hybridisierung erlaubt und gleichzeitig verschiedene Experimente leichter vergleichbar macht. Nach ersten Hybridisierungen wurde offensichtlich, dass bei Einzelzellamplifikaten auch auf Oligonukleotidarrays mit Standardhybridisierungsprotokollen unspezifische Signale auftreten können. Das Phänomen beruht auf der Kontamination der cDNA-Amplifikate mit bakterieller DNA, die durch Enzympräparationen eingeschleppt, amplifiziert und markiert wird. Besonders deutlich wurde der Effekt als eine Negativkontrolle, d.h. eine Probe, die das komplette Verfahren von Amplifikation, Markierung und Hybridisierung durchlaufen hatte, nur dass anfangs keine Zelle zugegeben worden war, und die daher ausschließlich kontaminierende Sequenzen enthielt, gegen die Referenz als Positivkontrolle hybridisiert wurde. Der Anteil positiver Signale auf dem Array betrug für beide Proben 100%, die Hybridisierung war vollständig unspezifisch (Abbildung 3.13).

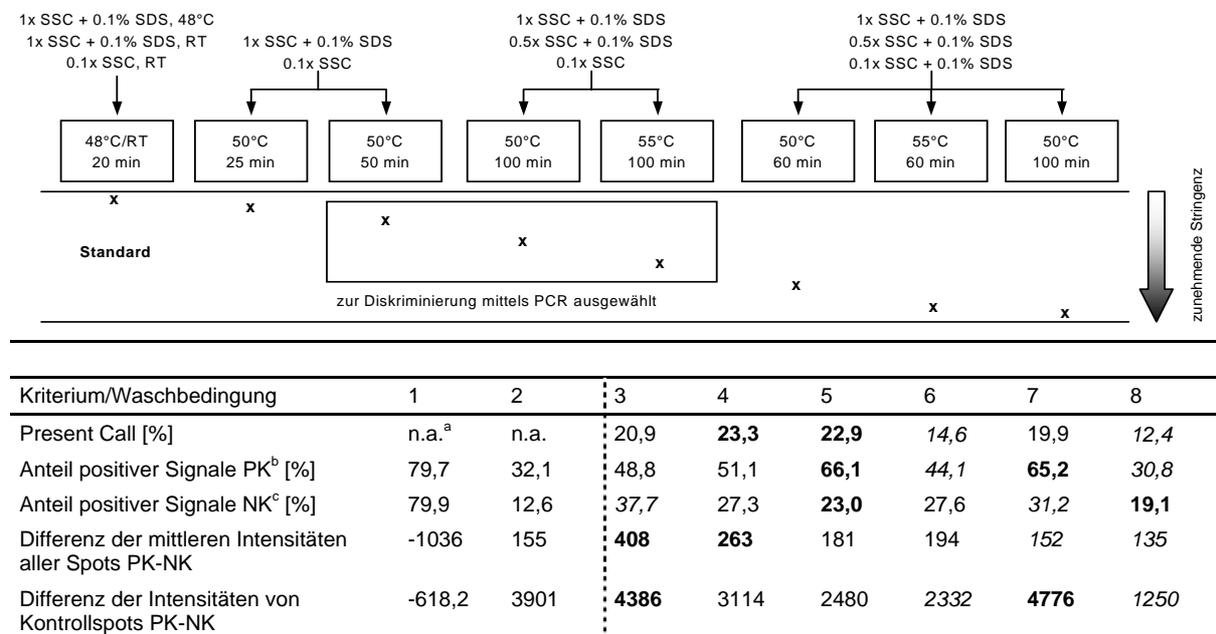


Abbildung 3.13 Unspezifische Hybridisierung einer Negativkontrolle auf einem Oligonukleotidarray

Auf einen Testarray mit fehlerhafter Spotmorphologie wurde nach einem Standardprotokoll eine Negativkontrolle des Amplifikationsprotokolls zusammen mit der universellen Referenz hybridisiert. Gezeigt ist ein Scan der Hybridisierung der Negativkontrolle.

Es bestand also die Notwendigkeit, die Hybridisierungsbedingungen für cDNA-Amplifikate aus dem beschriebenen Protokoll neu zu etablieren. Die Spezifität einer Hybridisierung wird vor allem durch die Stringenz der Waschschrte nach der Bindung der Probe an die Sonden des Arrays gesteuert. Weitere Einflussfaktoren sind der Hybridisierungspuffer, die Hybridisierungstemperatur und der Zusatz von Blockierungsreagenzien entweder zur Probe selbst oder in Form einer Prähybridisierung. Alle diese Faktoren wurden systematisch untersucht und optimiert. Den bedeutendsten Einfluss auf das Ergebnis hatte die Stringenz der Waschschrte.

Eine Auswahl der am ehesten geeigneten Waschverfahren ist in Abbildung 3.14 dargestellt. Bewertet wurden die Verfahren nach dem prozentualen Anteil positiver Signale allgemein (Present Call) sowie in Positiv- und Negativkontrollen (Sensitivität bzw. Spezifität), Signalintensitäten, Hintergrund und Intensitäten von Kontrollspots.



^a nicht analysiert
^b: Positivkontrolle
^c: Negativkontrolle

Abbildung 3.14 Waschbedingungen zur Prozessierung von murinen Operon-Arrays in der engeren Wahl sowie Ergebnisse der Bewertungskriterien

Die Abbildung zeigt einen Ausschnitt aller 35 durchgeführten Experimente. Im oberen Teil sind die getesteten Waschbedingungen geordnet nach zunehmender Stringenz gezeigt, sie werden von links nach rechts den acht Spalten der Tabelle zugeordnet. Bewertet wurden verschiedene Kriterien zur Abschätzung von Sensitivität und Spezifität der Hybridisierung. Der Present Call unterscheidet nicht zwischen beiden Proben der Hybridisierung und sollte ebenso wie der Anteil positiver Signale in der Positivkontrolle und die Differenzen zwischen den Messwerten der Positiv- und Negativkontrolle möglichst groß sein, der Anteil messbarer Signale aus der Negativkontrolle hingegen möglichst klein. Da Bedingung 1 und 2 ohne Hybridisierungsstation durchgeführt wurden, die übrigen Bedingungen mit Hybridisierungsstation, wurde Bedingung 1 und 2 nicht gewertet. Die jeweils besten beiden Bedingungen für jedes Bewertungskriterium sind fett markiert, die beiden schlechtesten kursiv.

Nach den gewählten Kriterien am ehesten geeignet schienen Waschschritte mittlerer Stringenz zu sein, da sie Sensitivität und Spezifität am besten balancierten. Mittels PCR für 44 spezifische Gensequenzen

wurde anschließend aus den Protokollen der engeren Wahl dasjenige bestimmt, das die größte Übereinstimmung der Hybridisierungsergebnisse zur PCR hinsichtlich positiver und negativer Signale lieferte (s. Absatz 3.3.2).

Durch Verwendung von Block-DNA aus *E.coli* zur Absättigung bakterieller Sequenzen und von Cot1-DNA für repetitive Sequenzen konnten die unspezifischen Signale weiter unterdrückt werden. Am effektivsten war die Blockierung in Form einer Prähybridisierung über Nacht, bevor die eigentliche Probe auf den Array aufgetragen wurde. Dieses Vorgehen ist jedoch bei Verwendung der Hybridisierungsstation nicht möglich, da nur eine Probenaufgabe erfolgen kann. Durch entsprechend größere Mengen Block-DNA konnten jedoch vergleichbar effektive Ergebnisse erzielt werden. Wichtig bei diesem Vorgehen war, dass die Block-DNA kein zu hohes Molekulargewicht besitzen durfte, da sonst der Hintergrund bei der Hybridisierung stark zunahm. Dasselbe galt für die maximal mögliche DNA-Menge, die auf den Array aufgetragen werden konnte.

Einen Überblick über die Etablierungsexperimente gibt folgende Auflistung. Die jeweils optimale Variante ist fett dargestellt. Das komplette Protokoll ist in Absatz 2.1.11 des Methodenteils beschrieben.

- Entfernen der ubiquitären Primer mittels Restriktionsverdau: **ja** – nein
- Aufreinigung der markierten Probe: keine - Phenol/Chloroform-Extraktion – **Qiagen-Aufreinigungssäule für PCR-Produkte**
- Zusatz von Cot1-DNA [μg]: 5 – **50** – 100
- Zusatz von *E.coli*-DNA [μg] als Prähybridisierung über Nacht : 5 – **6** – 7 – 20 – 50 – 75 – 100
- Zusatz von *E.coli*-DNA [μg] mit der Probe: 0.25 – 5 – 6 – 10 – **15** – 30 – 50
- Vorbehandlung der *E.coli*-DNA: DNase Verdau – Hitzefragmentierung in 0,4 M NaOH bei 95°C – **20 min Scherung durch Ultraschall**
- Zusatz von Heringsperm-DNA [μg]: **0** – 100
- Hybridisierungspuffer Ambion Slide Hyb Nr.: **1** – 2 – 3 – 4
- Hybridisierungstemperatur [°C]: 38 – 40 – **42** – 44 – 47
- Hybridisierungsstation: nein – **ja**
- Waschprotokoll: **50°C, 100 min; 1 x SSC + 0,1% SDS, 0,5 x SSC + 0,1% SDS, 0,1 x SSC**

Die Verwendung der Hybridisierungsstation („Arraybooster“) verbesserte den Anteil positiver Signale im Vergleich zur manuellen Hybridisierung um 33%. Das Gerät bewegt die Probenlösung während der Hybridisierung über dem Array, was die Wahrscheinlichkeit des Zusammentreffens einer bestimmten cDNA in der Probe mit der komplementären Sonde erhöht. Daten zur Reproduzierbarkeit wurden nicht erhoben, sie dürfte jedoch mit Hybridisierungsstation ebenfalls höher liegen als ohne. Ein Gerät zum automatischen Waschen nach der Hybridisierung brachte keine Verbesserungen, auch nicht hinsichtlich der Reproduzierbarkeit (Daten nicht gezeigt).

3.3.2 Ursachen verbleibender falsch-positiver Signale

Mit dem optimalen Protokoll lieferten Hybridisierung und PCR übereinstimmende Ergebnisse in 321/456 Fällen (70,4%) und 40 positive PCR-Reaktionen bei negativem Hybridisierungssignal (8,8%). Außerdem verblieben 95 falsch-positive Signale auf dem Array (20,8%). Eine mögliche Ursache liegt in nicht eindeutigem Sondendesign. Für 50mere sind in der Literatur Kreuzhybridisierungen bei 75% Sequenzähnlichkeit oder 15 identischen Basen in Folge zwischen zwei unterschiedlichen cDNAs beschrieben⁸². Eine BLAST-Suche mit den Sondensequenzen der für die PCR verwendeten Transkripte bestätigte zweideutiges Sondendesign als mögliche Ursache falsch-positiver Signale. Es fanden sich Homologien von über 15 Basen zu anderen als den erwünschten cDNAs in 21 von 44 untersuchten Sonden (48 %). Von diesen zeigten 13 Oligonukleotide falsch-positive Hybridisierungsergebnisse (61,9%). Exemplarisch wurde die Situation für die *Maged2*-Sonde näher untersucht. Unter den zusätzlichen Treffern in BLAST waren *Maged1* und *Maged3* mit jeweils 43 identischen Basen in Folge. Primer, die zwischen den Genen der Familie diskriminierten, zeigten eine Expression verschiedener *Maged*-Gene, die alle von der *Maged2*-Sonde erkannt werden (Abbildung 3.15). Das Hybridisierungssignal stellte in diesem Fall eine Summe der Expressionen verschiedener Gene dar. Bei Proben, die ausschließlich andere Gene als *Maged2* exprimierten, resultierten falschpositive Ergebnisse auf dem Array. Dies war für vier von elf Proben der Fall.

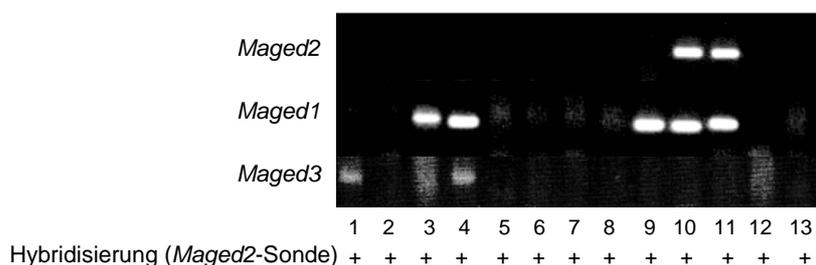


Abbildung 3.15 Ergebnisse spezifischer PCR-Reaktionen auf verschiedene Gene der *Maged*-Familie

Die *Maged2*-Sonde des Arrays erkennt neben *Maged2* auch *Maged1* und *Maged3*. Probe 1, 3, 4 und 9 exprimieren ausschließlich andere Mitglieder der *Maged*-Familie als *Maged2* und ergeben dadurch falsch-positive Hybridisierungssignale. Bei Probe 10 und 11 setzt sich das Hybridisierungssignal aus den Intensitäten mehrerer Gene zusammen. Für Probe 2, 5, 6, 7, 8, 12 und 13 konnte durch dieses Experiment keine Ursache der falsch-positiven Signale gefunden werden.

Eine weitere Ursache nicht übereinstimmender Ergebnisse zwischen Hybridisierung und PCR konnte in der Lage der PCR-Primer relativ zur Operon-Sonde liegen. Beim Design der Primer war die Sondensequenz und damit ihre Lage auf der cDNA noch nicht bekannt. Daher erkennen viele Primer nicht den exakt gleichen Bereich der cDNA wie die Arraysonde, auch wenn in beiden Fällen der terminale 3'-Bereich der mRNA gewählt wurde. Statistische Tests zeigten keinen signifikanten Zusammenhang zwischen der Lage der Primer relativ zur Sonde und dem Anteil übereinstimmender Ergebnisse (χ^2 -Test; $p=0,1$). Bei ausgewählten Sonden konnte der Anteil abweichender Ergebnisse

durch neue Primer exakt im Bereich der Sonde nicht erhöht werden. Die Ergebnisse der PCR-Reaktionen in Abhängigkeit von der Primerlokalisierung sind in Tabelle 3.5 zusammengefasst.

Tabelle 3.5 Ergebnisse der PCR-Reaktionen in Abhängigkeit von der Lage der Primer relativ zur Sonde des Arrays

	Array falsch-negativ	Array und PCR übereinstimmend	Array falsch-positiv
Primer im Vergleich zur Sonde weiter 3'	10	59	11
Lokalisation übereinstimmend	9	93	20
Primer im Vergleich zur Sonde weiter 5'	19	137	52

Neben nicht übereinstimmenden Ergebnissen, die z.B. durch BLAST-Suche erklärt werden konnten, gab es weiterhin „echte“ falsch-positive und falsch-negative Ereignisse, die in Kontamination, Verwechslung der Sonde bei der Herstellung der Arrays und in der Sensitivität der jeweiligen Technologie für die einzelne cDNA begründet sein können. Beispielsweise müssen auf Arrays sämtliche Sonden unter den gleichen Bedingungen hybridisieren, bei den PCR-Reaktionen können Primer unterschiedlich effizient sein. Dieser Anteil abweichender Ergebnisse kann nicht weiter reduziert werden und stellt eine immanente Eigenschaft des jeweiligen Verfahrens dar.

3.3.3 Reproduzierbarkeit und statistische Trennschärfe des Verfahrens

Für Zweifarbenarrays mit Probe und universeller Referenz werden für große Zellzahlen und die Hybridisierung von Gesamt-RNA-Präparationen meist Korrelationen zwischen Replikaten von $\geq 0,8$ angestrebt. Das beschriebene Verfahren erreicht diese Werte fast auch für Einzelzellamplifikate und überschreitet sie für die Gesamt-RNA-Präparation. Die Korrelationskoeffizienten wurden dabei für technische Replikate mit identischer Ausgangs-RNA bestimmt, die nach dem Zellaufschluss getrennt amplifiziert und hybridisiert wurden. Der Wert beinhaltet also sowohl die Streuung der Amplifikation als auch der Hybridisierung.

Tabelle 3.6 Reproduzierbarkeit des Gesamtprotokolls aus Amplifikation und Hybridisierung

Probe	Korrelation zum Replikat
Einzelzelle	0,73
Pool aus 10 Zellen	0,79
Pool aus 100 Zellen	0,81
Gesamt-RNA	0,85

Mittels einer Trennschärfenanalyse wurde anschließend berechnet, wie viele biologische Replikate in einer Untersuchung mit dem beschriebenen Verfahren durchgeführt werden müssen, um statistische zuverlässige Aussagen zu ermöglichen. Es wurde dabei eine Populationsvariabilität von 50% sowie ein zweiseitiges ungepaartes Experiment angenommen. Die Populationsvariabilität wurde aus dem

Experiment mit dendritischen Zellen (Abschnitt 3.4.1) abgeleitet, bei dem sie 44% betragen hatte, und aus Sicherheitsgründen auf 50% aufgerundet. Die Trennschärfenanalyse wurde mit Hilfe der Software R mit den Funktionen „power.anova.test“ und „power.t.test“ durchgeführt. In die Berechnung gehen neben der Art des Tests die Irrtumswahrscheinlichkeit (Wahrscheinlichkeit für Fehler 1. Art), die Trennschärfe oder Power (1-Wahrscheinlichkeit für Fehler 2.Art), die Populationsvariabilität, die Größe der zu beobachtenden Unterschiede sowie die Reproduzierbarkeit des Verfahrens ein. Tabelle 3.7 gibt die Anzahl benötigter Replikate pro Gruppe in Abhängigkeit von der angestrebten Trennschärfe, Irrtumswahrscheinlichkeit und der zu messenden Regulation der Genexpression zwischen den beiden Gruppen an. Deutlich wird, dass die Anzahl vor allem von der zu messenden Genregulation abhängt. Sollen Aussagen über um den Faktor zwei (logarithmiert zur Basis zwei) differenziell exprimierte Gene getroffen werden, sind etwa 25-35 Replikate nötig, bei Faktor drei 12-16 und bei Faktor fünf etwa neun.

Tabelle 3.7 Replikatzahl pro Gruppe bei verschiedenen statistischen Bedingungen

Trennschärfe (%)	0,95		0,98		0,99	
Irrtumswahrscheinlichkeit	0,0005	0,0001	0,0005	0,0001	0,0005	0,0001
Regulation um Faktor zwei	24	28	27	32	29	34
Regulation um Faktor drei	12	14	13	15	13	16
Regulation um Faktor fünf	7	9	8	9	8	10

3.3.4 Effektivität der Methode bei verschiedenen Mengen Ausgangs-RNA

Von einer Gesamt-RNA-Präparation wurden Verdünnungen angelegt, deren RNA-Mengen mit verschiedenen Zellzahlen korrelierten. Dabei wurde von einem Gesamt-RNA-Gehalt einer Einzelzelle von etwa 0,12 ng ausgegangen, was bei einem mRNA-Anteil von 5% 6 pg mRNA entspricht⁶⁹. Die Verdünnungen wurden mit dem beschriebenen Protokoll amplifiziert, hybridisiert und mit der Hybridisierung der unverdünnten Gesamt-RNA nach einem Standardprotokoll verglichen. Ein großer Sprung in den Korrelationskoeffizienten der Verdünnungen zur Gesamt-RNA trat erst zwischen 6 pg und 0,6 pg mRNA auf, was das Detektionslimit der Methode darstellen könnte.

Tabelle 3.8 Korrelation verschiedener Gesamt-RNA-Verdünnungen zur unverdünnten Gesamt-RNA

mRNA	Zelläquivalent	Korrelation zur Gesamt-RNA
600 pg	100 Zellen	0,496
60 pg	10 Zellen	0,473
6 pg	1 Zelle	0,379
0,6 pg	0,1 Zelle	0,159

3.3.5 Methoden zur Markierung von PCR-Produkten für Hybridisierungen

Prinzipiell werden zur Fluoreszenzmarkierung von DNA direkte und indirekte Verfahren unterschieden. Bei direkten Verfahren erfolgt sofort der Einbau fluoreszenzmarkierter Nucleotidderivate in die DNA, bei indirekten Verfahren der Einbau von Derivaten, die in nachfolgenden Schritten mit Fluorochromen gekoppelt werden. Beide Varianten bieten Vor- und Nachteile. Direkte Verfahren sind meist schneller und verlangen weniger Arbeitsaufwand, dafür sind sie sehr kostenintensiv und auch nicht immer ausreichend effektiv. Die gängigen Fluorochrome stellen im Vergleich zu nichtderivatisierten Nucleotiden sehr große Moleküle dar, die von Enzymen nicht oder nicht reproduzierbar in Nucleinsäuren eingebaut werden. Daher ist die Labellingeffizienz oft unzureichend oder nicht konstant. Im Gegensatz dazu werden bei indirekten Verfahren kleinere Nucleotidderivate verwendet, die effektiv und reproduzierbar eingebaut werden. Dafür sind der Arbeitsaufwand und eventuell auch die Variabilität durch zusätzliche Kopplungsreaktionen höher. Klassische Verfahren der indirekten Markierung arbeiten mit Aminoallyl-, Digoxigenin- und Biotinmarkierten Nucleotiden.

Die Beurteilung einer Labellingreaktion kann spektroskopisch erfolgen. Mit dem Photometer lassen sich sowohl Menge der in der Reaktion gebildeten Nucleinsäure als auch Fluoreszenzintensität der eingebauten Fluorochrome bestimmen. Aus beiden Größen wird die Einbaurate derivatisierter Nucleotide pro 1000 Nucleotide berechnet (FOI; Frequency of Incorporation = eingebaute Menge Fluorochrom [pmol] / gebildete Menge cDNA [ng] * 234,5; zu berechnen aus A_{550} für Cy3 bzw. A_{649} für Cy5 und A_{260} für die gebildete Menge cDNA). Sowohl die FOI als auch die Produktmenge müssen Schwellenwerte erreichen, damit die Probe erfolgreich hybridisiert werden kann. Für die FOI gilt dabei ein Wert von zehn als Grenze, markierte DNA sollten mehrere Mikrogramm vorhanden sein.

Verglichen wurden einfache und doppelte Direktmarkierungen, die durch PCR mit verschiedenen Polymerasen (PolMix, Taq-Polymerase, Thermosequenzase), als Verlängerung von Zufallsprimern durch Klenowenzym oder in Form einer Endmarkierung, in der mittels Terminaldeoxynucleotidtransferase (TdT) derivatisierte Nucleotide ans 3'-Ende der cDNA angehängt werden, durchgeführt wurden. Außerdem wurden indirekte Markierungen mit Aminoallyl-dUTP getestet. Dabei machte das derivatisierte Nucleotid Anteile am dTTP/dUTP-Gemisch von 60-100% aus. Die PCR-Produkte wurden anschließend über die Aminogruppen mit einem cyaninkonjugierten reaktiven Ester gekoppelt. Die letzte getestete Möglichkeit bestand im Einbau biotinylierter Nucleotide während einer PCR und einer Kopplung mit Streptavidin-Cy3. Vor der Messung im Photometer wurden bei allen Reaktionen überschüssige Fluorochrome über eine Qiagen Aufreinigungssäule für PCR-Produkte, die DNA spezifisch bindet, abgetrennt. Abbildung 3.16 gibt einen Überblick über die Ergebnisse der Experimente.

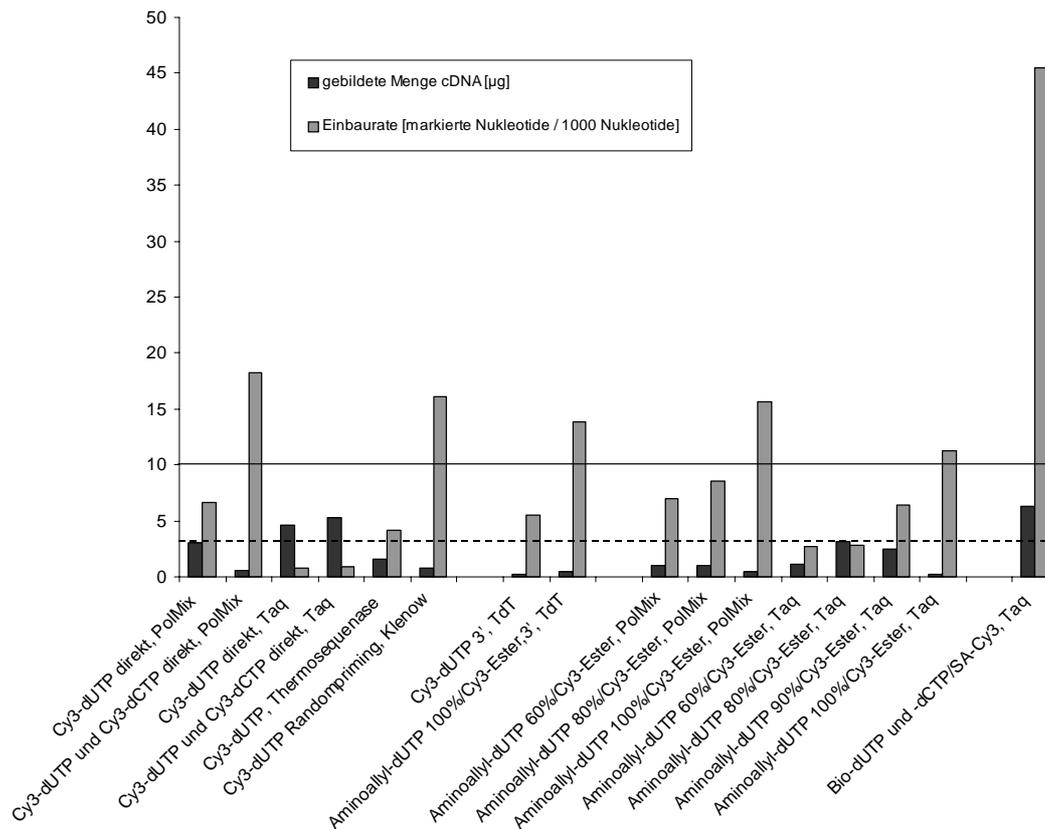


Abbildung 3.16 Menge cDNA am Ende der Markierungsreaktion und Einbaurrate markierter Nucleotide in Abhängigkeit vom verwendeten Verfahren

Dargestellt ist die gebildete Menge cDNA sowie die Einbaurrate derivatisierter Nucleotide nach verschiedenen Markierungsverfahren. Links finden sich Direktmarkierungen mit verschiedenen Enzymen, die in unterschiedlichem Ausmaß in der Lage sind, cyaninkonjugierte Nucleotide in Nucleinsäuren einzubauen. In der Mitte sind Endmarkierungen durch TdT und rechts indirekte Verfahren mit Aminoallyl-dUTP und Biotin-dUTP (Bio-dUTP) gezeigt. Der Mindestwert für die Einbaurrate beträgt zehn derivatisierte Nucleotide pro 1000 Nucleotide (durchgezogene horizontale Linie). Die gebildete Menge cDNA sollte mehrere Mikrogramm erreichen (gestrichelte Linie).SA: Streptavidin

Nur wenige Verfahren führten zu ausreichenden Einbauraten markierter Nucleotide, darunter drei Verfahren der Aminoallylmarkierung. Ein Nachteil dieser Technik ist der hohe Arbeitsaufwand, der sich über drei Tage erstreckt, da die DNA zur Kopplung mit dem Fluorochrom und für die Hybridisierung jeweils über Nacht gefällt werden muss. Die Direktmarkierung war nicht reproduzierbar und sehr teuer. Die besten Ergebnisse hinsichtlich Einbauraten, gebildeter Menge cDNA und Arbeitsaufwand lieferte die indirekte Markierung über Biotin und Streptavidin-Cy3. Sie stellte die einzige Möglichkeit dar, gute Einbauraten mit ausreichenden Produktmengen zu vereinbaren. Mit allen anderen Methoden war jeweils nur das eine auf Kosten des anderen zu erreichen. Diese Art der indirekten Markierung hat sich auch bei anderen Hybridisierungen bewährt und wird in ähnlicher Form von der Firma Affymetrix eingesetzt. Auch nach Hybridisierungen auf Operonarrays war das Verfahren die Methode mit der höchsten Reproduzierbarkeit (0,51 gegen 0,22 für Direktmarkierung derselben biologischen Replikate) und wurde daher verwendet.

3.3.6 Maßnahmen zur Steigerung der Sensitivität

Zur Verbesserung der Sensitivität des bestehenden Isolations- und Amplifikationsprotokolls von mRNA aus Einzelzellen sowie der Hybridisierung auf Oligonukleotidarrays wurden verschiedene Größen auf ihren Einfluss untersucht und zum Teil bereits detaillierter beschrieben. Dazu gehörten die

- mRNA-Extraktion mit PNAs statt Oligo dT-Beads (vgl. Absatz 3.1.1)
- Anzahl der Thymidin-Nukleotide im cDNA-Syntheseprimer (vgl. Absatz 3.1.1),
- Zyklenzahl bei der Labelling-PCR für die Hybridisierung bis ins Plateau (16 Zyklen) oder nur bis in die exponentielle Phase (10 Zyklen),
- Doppelmarkierung für die Hybridisierung mit markiertem dUTP und dCTP statt dUTP allein,
- Verwendung der Hybridisierungsstation im Vergleich zur manuellen Hybridisierung.

Die genannten Aspekte wurden in verschiedenen Experimenten (A bis D) verglichen und jeweils die Korrelationen der Hybridisierungen zu Ergebnissen mit nicht amplifizierter Gesamt-RNA bestimmt. Für jedes Experiment konnte dadurch eine überlegene Variante definiert werden (Tabelle 3.9).

Tabelle 3.9 Korrelationskoeffizienten verschiedener Maßnahmen zur Erhöhung der Sensitivität im Vergleich mit Gesamt-RNA-Hybridisierungen

Experiment	Varianten	Korrelation zur Gesamt-RNA
A	mRNA-Extraktion mit PNAs	0,29
	mRNA-Extraktion mit Oligo dT-Beads	-0,12
B	cDNA-Syntheseprimer mit 15 Thymidin-Nukleotiden	0,35
	cDNA-Syntheseprimer mit 24 Thymidin-Nukleotiden	0,48
C	10 Zyklen Labellingreaktion	0,38
	16 Zyklen Labellingreaktion	0,47
D	Einfachmarkierung ohne Hybridisierungsstation	0,46
	Doppelmarkierung ohne Hybridisierungsstation	0,50
	Einfachmarkierung mit Hybridisierungsstation	0,60
	Doppelmarkierung mit Hybridisierungsstation	0,60

Die größte Verbesserung brachte also die Umstellung vom Oligo dT-Bead- auf das PNA-Protokoll. Die anderen Maßnahmen ergaben ebenfalls Verbesserungen unterschiedlichen Ausmaßes. Zusammengefasst führte die Extraktion der mRNA mit PNAs, ein cDNA-Syntheseprimer mit 24 Thymidin-Nukleotiden, eine Labellingreaktion mit markiertem dUTP und dCTP und 16 Zyklen PCR sowie die Nutzung der Hybridisierungsstation zu den besten Ergebnissen.

3.4 Validierung des Hybridisierungsprotokolls

3.4.1 LPS-Stimulation von murinen dendritischen Zellen und Hybridisierung

Zur Validierung des Hybridisierungsprotokolls an einer biologischen Fragestellung wurden aus murinem Knochenmark Vorläuferzellen isoliert und in Gegenwart von GM-CSF *in vitro* zu immaturren dendritischen Zellen (DCs) differenziert. Immature DCs sind durch Expression von CD11c gekennzeichnet und konnten über diesen Marker selektiert werden. Ein Teil der Kulturen wurde anschließend mit LPS stimuliert, was die Reifung der dendritischen Zellen auslöst¹³⁸. Mature DCs exprimieren neben CD11c Aktivierungsmarker wie beispielsweise CD86 und wurden so von anderen Populationen in der Kultur unterschieden (Abbildung 3.17). Genexpressionsanalysen von Einzelzellen aus sowohl stimulierten als auch unstimulierten Ansätzen sollten Aufschluss über Sensitivität und Reproduzierbarkeit der Hybridisierungen geben. Das System ist für solche Tests besonders geeignet, da immature und mature DCs sich nur im Aktivierungszustand unterscheiden, was bei der zu erwartenden hohen Variabilität der Genexpression zwischen einzelnen Zellen, die sich beispielsweise in unterschiedlichen Differenzierungs- und Zellzykluszustände befinden können, einen relativ geringen Unterschied darstellt.

Nach der Hybridisierung von vier Einzelzellen jeder Gruppe auf Operon-Arrays wurden die Messwerte einem hierarchischen Clustering unterworfen. Dabei werden in einem n-dimensionalen Raum bestehend aus den Messwerten aller Proben für n Gene die Abstände der Proben zueinander berechnet. Die beiden ähnlichsten Proben werden zu einer Gruppe zusammengefasst und wieder alle Abstände berechnet. Das Verfahren wird so lange wiederholt, bis alle Proben nach ihrer Ähnlichkeit gruppiert worden sind. Die graphische Darstellung erfolgt in einem Dendrogramm, bei dem die jeweils zueinander ähnlichsten Proben oder Gruppen von Proben in einem Ast zu liegen kommen. Die Länge der senkrechten Strecken zwischen den Knoten innerhalb der Baumstruktur ist ein Maß für die Distanz zweier Proben, macht also eine Aussage über die Ähnlichkeit ihrer Genexpressionsmuster¹⁰⁶. Bei der Analyse der DC-Daten war der Algorithmus in der Lage, anhand der Genexpression der Einzelzellen die Proben in die stimulierte beziehungsweise unstimulierte Gruppe einzuteilen. Daneben enthält das Dendrogramm auch zwei Einzelzell-Äquivalente der epithelialen Zelllinie TUBO, die von den dendritischen Zellen deutlich als erstes abgetrennt werden. Die Sensitivität des Protokolls erlaubt es also, auch geringe Unterschiede in der Genexpression neben der interzellulären Variabilität abzubilden.

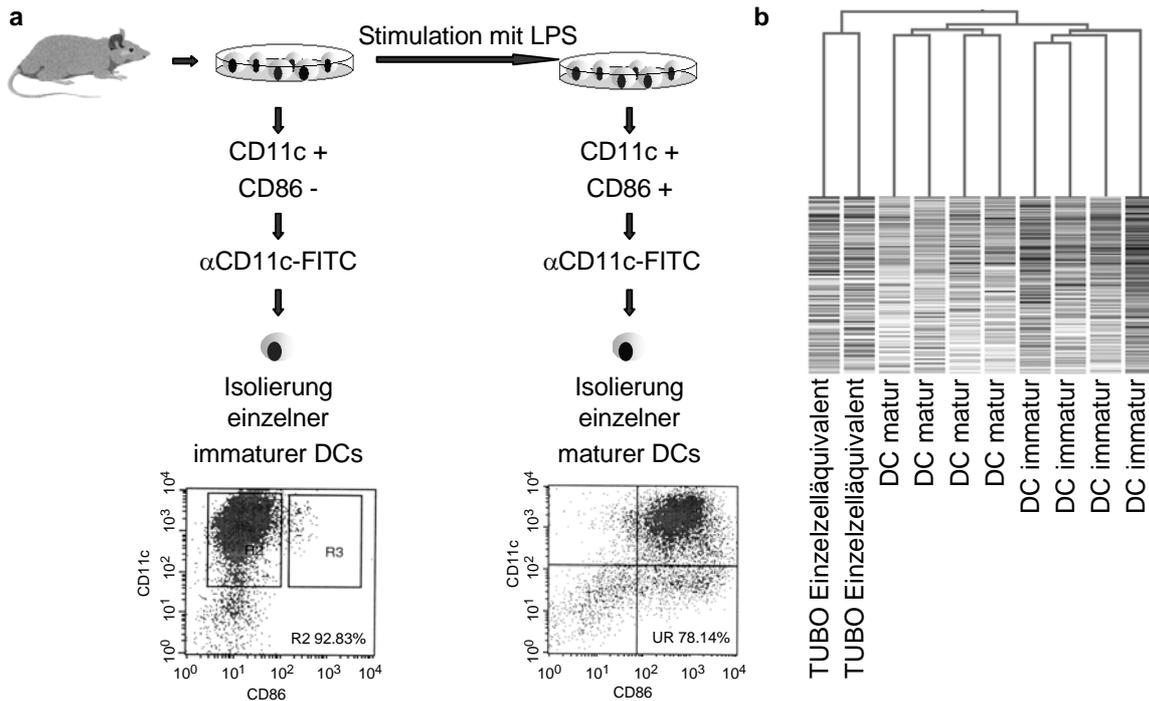


Abbildung 3.17 Durchführung und Ergebnis des DC Experiments

a Aus murinem Knochenmark wurden dendritische Zellen (DCs) generiert. Sowohl immature als auch mature DCs exprimieren CD11c, mature DCs, die durch Stimulation mit LPS erhalten wurden, sind zusätzlich durch die Expression von CD86 gekennzeichnet. Beide Populationen können über eine Färbung gegen diese Marker mit beadgekoppelten Antikörpern in einem Magnetfeld aus den Kulturen isoliert werden. Durch geeignete Kombination von Selektion und Depletion werden mature bzw. immature DCs angereichert, wie aus den FACS-Untersuchungen hervorgeht. Nach einer Färbung mit α CD11c-FITC wurden aus beiden Ansätzen einzelne fluoreszenzmarkierte Zellen unter dem Mikroskop isoliert und die Genexpression analysiert. **b** Ein hierarchischer Clusteralgorithmus trennt nicht nur einzelne epitheliale (TUBO) von dendritischen Zellen, sondern auch stimulierte von unstimulierten DCs anhand ihrer Genexpressionsprofile.

3.4.2 Validierung mittels quantitativer real time PCR

Als unabhängige Methode zur Überprüfung der Hybridisierungsergebnisse diente die quantitative real time PCR (qPCR). Aus dem DC-Experiment wurden zufällige Gene ausgewählt, die im stimulierten Zustand relativ zum unstimulierten auf dem Array hoch-, nieder- oder nicht reguliert waren. Bei der Untersuchung der Proben mittels qPCR konnten die relativen Expressionswerte der stimulierten zu den unstimulierten Proben aus der Hybridisierung bestätigt werden (Abbildung 3.18). Alle Gene wurden mit beiden Verfahren derselben Gruppe (hoch-, nieder- oder nicht reguliert) zugeordnet. Die Unterschiede in der gemessenen Genregulation zwischen beiden Verfahren lagen bei Faktor 6,5 (2,7-fach nach Logarithmierung zur Basis zwei) und bewegten sich damit im Rahmen publizierter Daten aus Gesamt-RNA-Experimenten. Bei verschiedenen Oligonukleotidarrays werden in diesen Untersuchungen 50-96% der betrachteten Gene in der PCR der gleichen Gruppe zugeordnet, die logarithmierten Expressionswerte schwanken bis um den Faktor 512 (9-fach nach \log_2), für Operon-Arrays um den Faktor 8,6 (3,4-fach nach \log_2). Wie bei den Einzelzellamplifikaten wird auch in diesen

Experimenten die Genregulation von allen Arrays eher unterschätzt^{87, 139, 140}. Die Hybridisierung von Einzelzellamplifikaten nach dem beschriebenen Protokoll ist also mit der gleichen Güte möglich wie die Untersuchung von Gesamt-RNA-Präparationen nach Standardprotokollen.

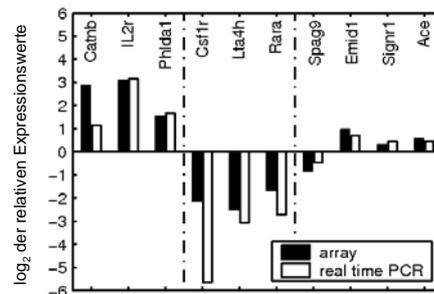


Abbildung 3.18 Vergleich der Ergebnisse von Hybridisierung und qPCR

Die Genexpressionswerte wurden relativ zum Haushaltsgen *Gapdh* bestimmt und dann das Verhältnis der Mittelwerte der stimulierten Proben zu dem der unstimulierten Proben gebildet. Jedes Gen fällt dabei in eine der drei Gruppen der hochregulierten (links), niederregulierten (Mitte) und nicht regulierten Gene (rechts). Kriterium war eine zweifache Hochregulation, d.h. eins bei logarithmischer Skalierung zur Basis zwei, zweifache Niederregulation, d.h. minus eins bei logarithmischer Skalierung zur Basis zwei oder keine Regulation zwischen diesen Grenzen.

3.4.3 Biologische Aussagekraft und Sensitivität des Verfahrens

Durch die qPCR- und Clusterverfahren konnte gezeigt werden, dass die Ergebnisse der Hybridisierungen valide Messwerte darstellten. Ob auch biologisch sinnvolle Ergebnisse zu erhalten waren, wurde durch eine Betrachtung des Tlr4-Signalweges, dessen Ligand LPS ist, untersucht^{141, 142}. Zur besseren Übersicht wurde eine graphische Darstellung gewählt. Dazu wurden die Expressionswerte der stimulierten und unstimulierten DCs jeweils gemittelt und in ein Schema des Signalweges eingetragen, das vom KEGG-Konsortium erstellt wurde und als Download zur Verfügung steht (www.genome.jp/kegg/). Die Aktivierung des Signalweges konnte auf Einzelzellebene dargestellt werden und folgte dem klassischen MyD88-abhängigen Signalweg^{143, 144}. Es fand eine Hochregulation von Genen innerhalb des Signalkaskade, beispielsweise *Irak4* oder *Tab1*, wie auch eine erhöhte Expression von Effektoren des Signalweges wie *CD80*, *IL6* oder dem Chemokin *Rantes* statt (Abbildung 3.19).

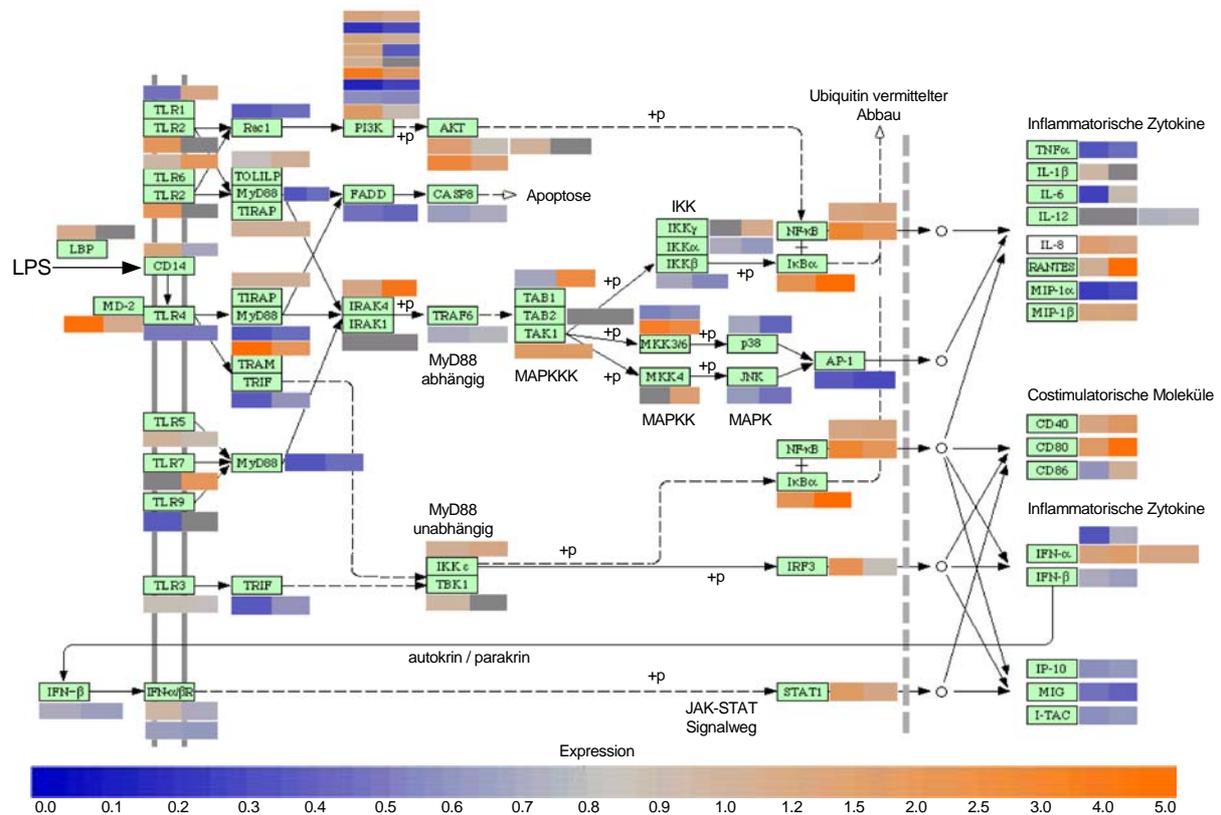


Abbildung 3.19 Graphische Darstellung der Aktivierung des Tlr-Signalweges durch LPS auf Einzelzell-
ebene

In die Vorlage des Tlr-Signalweges des KEGG-Konsortiums wurden die Expressionswerte aller Gene als Mittelwerte von acht immaturren DCs auf der linken Seite jedes Balkens sowie von acht maturaen DCs auf der rechten Seite des Balkens neben dem entsprechenden Gen eingetragen. Die Farbe gibt dabei die Expressionsstärke an und nimmt von blau nach rot zu.

Die Sensitivität des Verfahrens war also so hoch, dass aus den Hybridisierungen der einzelnen Zellen biologisch sinnvolle Ergebnisse zu erhalten waren. Dies konnte nach der Untersuchung weiterer Proben bestätigt werden. Dazu wurde das Gesamtprotokoll auf zahlreiche murine Zellen unterschiedlichen histogenetischen Ursprungs und Probenotyps angewendet, darunter DCs, hämatopoietische Vorläuferzellen und epitheliale Zellen. Die Hybridisierungsergebnisse wurden gemeinsam einem hierarchischen Clustering unterworfen. (Abbildung 3.20). In der Baumstruktur wurden zuerst die undifferenzierten hämatopoietischen Zellen (CD34+/-, Hoechst 33342+/-) von den differenzierten Zellen abgetrennt. Die einzige CD34-negative Zelle verhielt sich wiederum anders als die übrigen CD34-positiven Zellen. Als nächstes wurden epitheliale (TUBO) von dendritischen Zellen getrennt. Letztere gliederten sich in eine eher homogene Gruppe der immaturren DCs und eine heterogener Gruppe maturaer DCs, was eventuell unterschiedliche Aktivierungszustände reflektiert. Die epithelialen Zellen spalteten sich in Gruppen von Zell-Äquivalenten, die aus Gesamt-RNA verdünnt worden waren, sowie Zellpools untergruppiert nach Zellzahl. Auch die 0,1 Zell-Äquivalente wurden korrekt mit den epithelialen Zellen gruppiert, wenn sie auch die am wenigsten verwandten Proben darstellten. Die Einzelzellen waren über die gesamte Gruppe der epithelialen Zellen verteilt,

was ihre Heterogenität reflektiert. Zwei Zellen wurden je zweimal hybridisiert, wobei die Replikate nebeneinander zu liegen kamen.

Alle Proben wurden also gemäß ihrem histogenetischen Ursprung bzw. nach Probentypen gruppiert, was die hohe Sensitivität und Reproduzierbarkeit des Verfahrens zum Ausdruck bringt.

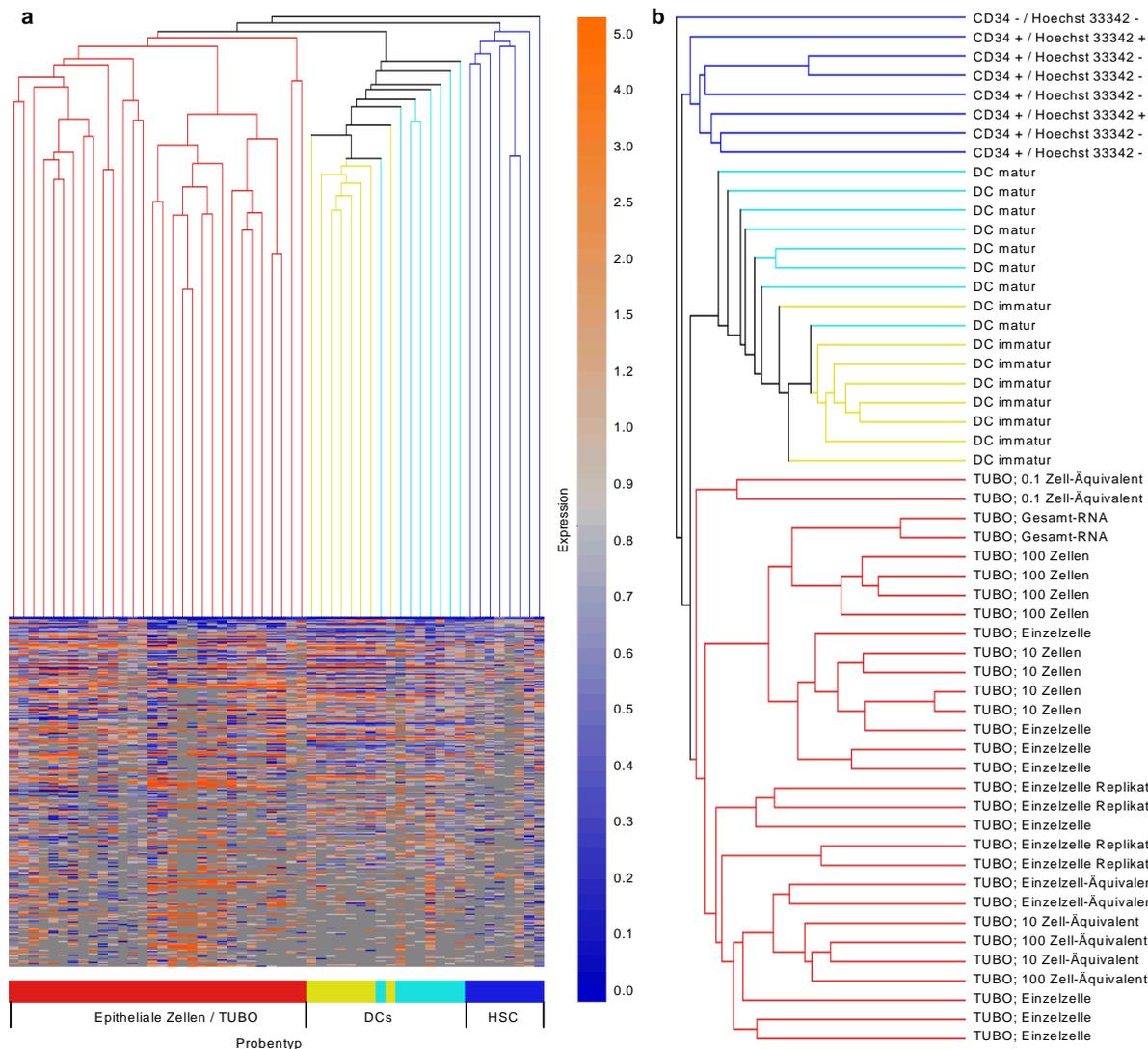


Abbildung 3.20 Ergebnis der hierarchischen Clusteranalyse sämtlicher muriner Proben

a Im Gesamtbild des Clusters sind alle Proben gemäß ihrem biologischen Ursprung farbig markiert. Die genaue Probenbezeichnung ist der Vergrößerung der Baumstruktur zu entnehmen (**b**). Alle Zellen werden vom Clusteralgorithmus gemäß ihrem histogenetischen Ursprung, nach Probentypen und Zellzahlen gruppiert. DC: Dendritische Zelle, HSC: Hämatopoietische Stammzelle

3.5 Herstellung von Operon-Arrays aus lyophilisierten Oligonukleotiden

Nach den erfolgreichen Versuchen mit den murinen Arrays sollte für Proben menschlichen Ursprungs ebenfalls das Operon-System verwendet werden. Es wurden fünf mögliche Quellen für humane Arrays miteinander verglichen:

1. Eigene Herstellung aus einem humanen Oligoset der Firma Operon im Labor von Dr. Jan Vijg, San Antonio, Texas. Von diesem Kooperationspartner stammten die Mausarrays.
2. Eigene Herstellung aus einem humanen Oligoset der Firma Operon im Labor von Dr. Peter Lichter, DKFZ Heidelberg. Mit diesem Labor bestand eine Kooperation zur Herstellung genomischer Arrays.
3. Eigene Herstellung aus einem humanen Oligoset der Firma Operon im Labor von Dr. Markus Panhuysen, MPI für Psychiatrie, München.
4. Bezug fertiger Arrays aus der laufenden Produktion im Labor von Dr. Ake Borg, Lund, Schweden.
5. Bezug fertiger Arrays der Firma Operon.

Da anfangs nur die Möglichkeit bestand, selbst Arrays zu produzieren und das zugleich die wirtschaftlichste Alternative darstellt, wurden Vorversuche zur Beantwortung technischer Fragen und eventuell weiterer Verbesserungen des Systems hinsichtlich der Hybridisierungseigenschaften durchgeführt.

3.5.1 Wahl der geeigneten Beschichtung des Objektträgers

Entscheidenden Einfluss auf die Hybridisierungseigenschaften nimmt der Objektträger, auf den die Sonden aufgebracht werden. Normalerweise werden modifizierte Glasobjektträger benutzt, wobei je nach verwendeter Oberflächenbeschichtung elektrostatische oder kovalente Bindungen zwischen Oligonukleotid und Objektträger möglich sind. Unterstützt werden diese Bindungen durch UV-Strahlung und Hitze.

Getestet wurden die beiden gängigen Oberflächen aus Epoxid- und Aminosilanbeschichtung (Gamma-amino Polysilan, GAPS). Epoxide binden aminomodifizierte Oligonukleotide wie die der Firma Operon kovalent, was ein Vorteil gegenüber den Silanoberflächen darstellen soll. Sie waren jedoch bei selbst hergestellten Arrays mit dem Hybridisierungsprotokoll nicht kompatibel, da der Hintergrund sehr stark und nicht vorhersagbar zunahm (Abbildung 3.21). Hintergrundeffekte waren auch bei der Hybridisierung von Gesamt-RNA zu beobachten und traten nicht nur bei Einzelzellen auf. Zusätzlich war die Spotgröße gering, innerhalb einer Charge nicht konstant, und die Ablagerung der DNA erfolgte oft nur in der Peripherie des Spots, wohingegen das Zentrum nicht hybridisierte. Da diese Effekte bei epoxidmodifizierten Objektträgern verschiedener Hersteller wiederholt auftraten, wurden aminosilanmodifizierte Objektträger verwendet, die konstantere Hybridisierungsergebnisse lieferten. Die UltraGAPS- brachten dabei noch etwas bessere Resultate als die GAPS II-Slides.

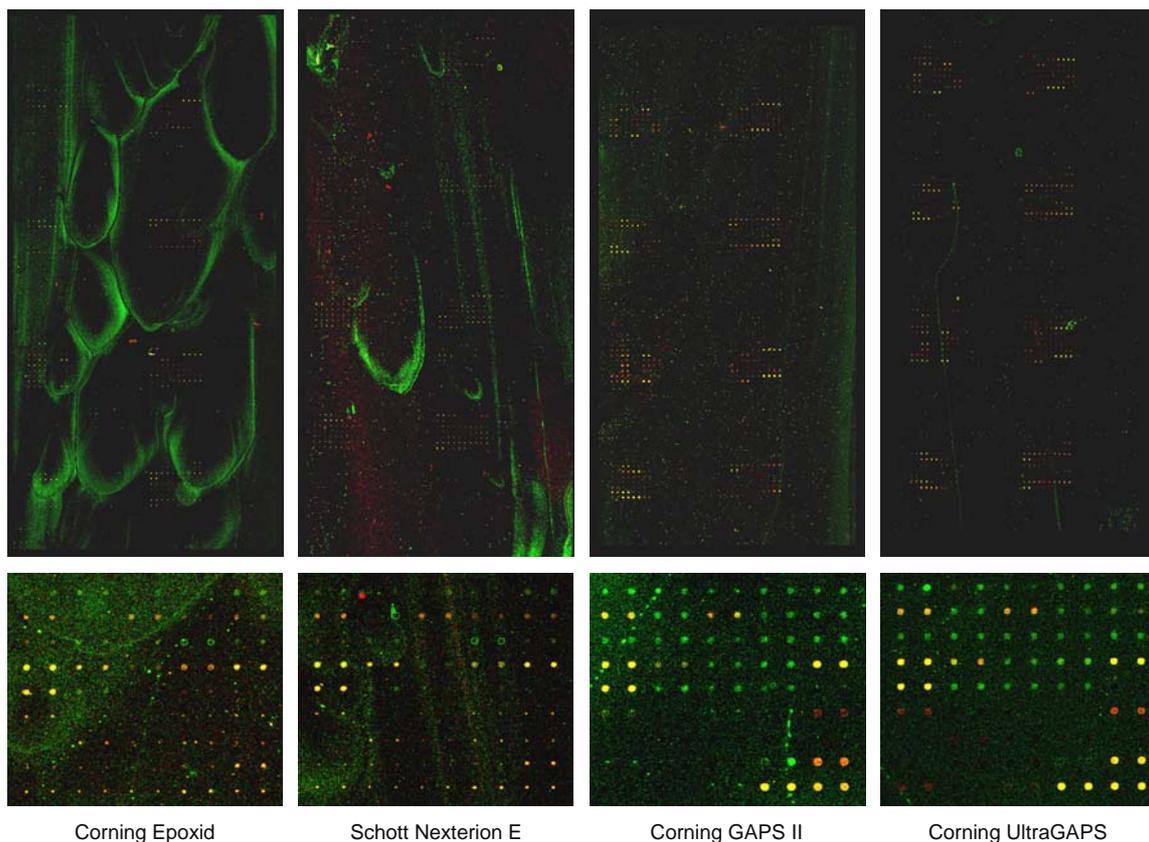


Abbildung 3.21 Einfluss der Oberflächenbeschichtung von Objektträgern zur Herstellung von Mikroarrays auf das Hybridisierungsverhalten

Identische Oligonukleotide wurden auf Epoxid-Oberflächen von zwei verschiedenen Herstellern (Corning Epoxid und Schott Nexterion E) sowie zwei Generationen von Aminosilanbeschichtungen (GAPS II und UltraGAPS) aufgebracht. Nach der Hybridisierung identischer cDNA-Proben sind auf den Epoxid-Oberflächen Hintergrundeffekte zu beobachten. Außerdem variierte die Spotgröße stark, wie in der Vergrößerung zu erkennen ist. Die Epoxid-Objektträger der Firma Corning erreichen Spotgrößen von 50-80 μm , die der Firma Schott von 60-80 μm . Auf GAPS II- und UltraGAPS-Oberflächen werden Größen von 120-150 μm erreicht. Für den verwendeten Scanner zur Detektion der Hybridisierungssignale mit einer Auflösung von 10 μm pro Pixel sind Spotgrößen über 80 μm wünschenswert.

3.5.2 Wahl des Spottingpuffers

Die Wahl des Puffers, in dem die lyophilisierten Oligonukleotide zum Spotten aufgenommen werden, hat einen entscheidenden Einfluss auf Größe, Morphologie und Hybridisierungsverhalten der Spots. Der Standardpuffer, der auch für die Mausarrays verwendet wurde, ist 3x SSC. Dieser Puffer hat allerdings den Nachteil, dass er verdunstet. Bei langen Herstellungsprozessen, wie sie bei Arrays mit Tausenden von Spots auftreten, führt dies zu Schwankungen in der Oligonukleotidkonzentration über die Zeit. Außerdem muss der Puffer nach dem Spotten abgezogen und die Oligonukleotide für den nächsten Lauf wieder im entsprechenden Volumen Wasser aufgenommen werden, was mit Ungenauigkeiten z.B. durch Pipettierfehler einhergeht. Wünschenswert wäre also ein nicht evaporierender Puffer, der durch verschiedene Zusätze, die die Verdunstung herabsetzen, wie DMSO,

Ethylenglycol, Betain oder Tenside, erhalten wird. Viele dieser Substanzen, besonders Tenside, neigen jedoch vor allem in der Kälte zu Ausfällungen, die unter schonenden Bedingungen nicht wieder in Lösung gebracht werden können. Zu Testzwecken wurden in verschiedenen Puffern, die prinzipiell geeignet erschienen, identische Oligonukleotide zu einer einheitlichen Konzentration gelöst und in Replikaten auf einen einzigen Objektträger aufgebracht. Nach der Hybridisierung wurden Morphologie und Konsistenz von Spots und Replikaten beurteilt (Abbildung 3.22).

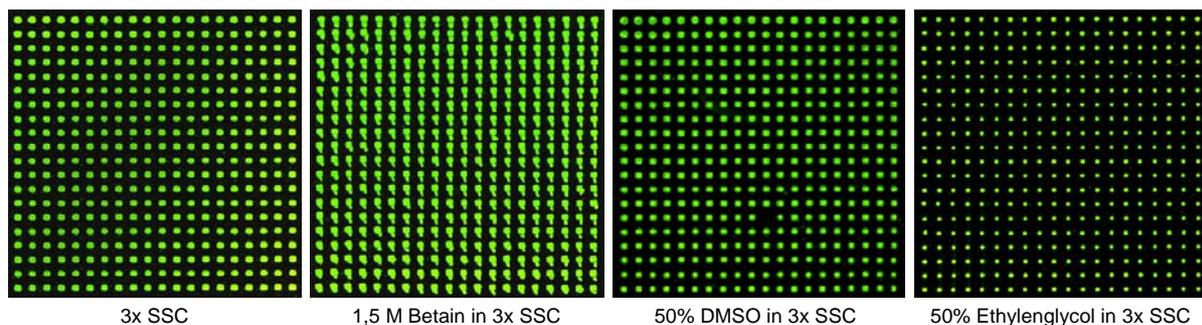


Abbildung 3.22 Test verschiedener Spottingpuffer

Dasselbe Oligonukleotid wurde in vier verschiedenen Puffern gelöst und jeweils 400 Spots aufgebracht. Dem Standardpuffer 3x SSC wurden verschiedene Substanzen zugesetzt, die die Verdunstung herabsetzen. Die Größe der Spots beträgt für 3x SSC 120 μm , für 1,5 M Betain in 3x SSC 100 μm , für 50% DMSO in 3x SSC 150-80 μm und für 50% Ethylenglycol in 3x SSC 60 μm .

Kriterium für einen geeigneten Puffer ist eine konstante, ausreichende Größe der Spots bei gleich bleibend runder Morphologie und Intensität. Diese Anforderungen werden im gezeigten Beispiel am ehesten von 3x SSC und 50 % DMSO in 3x SSC erfüllt. Der DMSO-haltige Puffer reduzierte jedoch die Sensitivität des Systems, da gering exprimierte Gene nicht mehr nachweisbar waren. Außerdem ist die Größe der Spots über die 400 Replikate nicht konstant. Aus diesen Experimenten blieb letztlich 3x SSC als der am besten geeignete Puffer übrig.

Neben dem Puffer spielt auch die Konzentration der Oligonukleotide eine entscheidende Rolle für das spätere Hybridisierungsverhalten. Einzelzell-Amplifikate konnten auf Operon-Arrays, die mit einer Konzentration von 40 μM Oligonukleotid hergestellt wurden, erfolgreich hybridisiert werden, wohingegen bei Arrays, die mit 24 μM Spottingkonzentration produziert worden waren, deutliche Sensitivitätsverluste auftraten. Daher wurde die optimale Spottingkonzentration experimentell bestimmt, indem von 96 Oligonukleotiden Verdünnungsreihen in 3x SSC von 10 μM bis 80 μM angelegt und auf Corning UltraGAPS-Slides aufgebracht wurden. Nach der Hybridisierung von Einzelzell-Amplifikaten wurde die mittlere Signalintensität aller Oligonukleotide für beide Fluoreszenzen Cy5 und Cy3 ermittelt (Abbildung 3.23). Optimale Ergebnisse ergaben Konzentrationen von 40 μM Oligonukleotid und darüber. Für Cy3 wurde bereits bei 40 μM ein Plateau erreicht, wohingegen für Cy5 die Messwerte auch bei 80 μM Spottingkonzentration noch stiegen, dann allerdings über dem Maximalwert für Cy3 lagen. Aus Gründen der Vergleichbarkeit der

Intensitätswerte beider Kanäle und der Wirtschaftlichkeit der Array-Produktion lag die optimale Oligonukleotidkonzentration daher bei 40 µM.

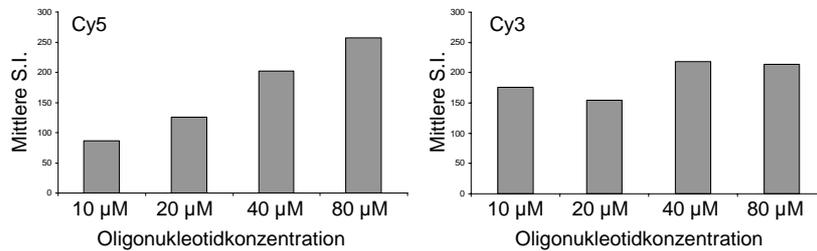


Abbildung 3.23 Bestimmung der optimalen Oligokonzentration zur Herstellung von Operon-Arrays

Nach der Hybridisierung von Einzelzellamplifikaten (Cy5) gegen einen Referenzpool (Cy3) wurden die mittleren Signalintensitäten (S.I.) beider Kanäle für alle Oligonukleotide des Arrays bestimmt. Im Cy5-Kanal nimmt der Messwert mit steigender Konzentration kontinuierlich zu. Für den Cy3-Kanal scheint bei 40 µM ein Schwellenwert zu liegen, ober- und unterhalb werden jeweils konstante Intensitäten gemessen. Bei einer Konzentration von 40 µM erreichen beide Kanäle vergleichbare Intensitätswerte.

3.5.3 Etablierung des Hybridisierungsprotokolls für humane Operon-Arrays

Zu Testzwecken wurden in San Antonio aus Kontrolloligonukleotiden in 3x SSC (40 µM) auf UltraGAPS-Slides kleinere humane Arrays hergestellt. In ersten Experimenten stellte sich die Annahme, cDNA-Amplifikate aus humanen Zellen müssten sich auf identisch produzierten Arrays nach den gleichen Bedingungen hybridisieren lassen wie die murinen Primäramplifikate, als falsch heraus. Mit den für Mausarrays definierten Bedingungen waren aus humanen Proben keinerlei Signale detektierbar. Dies betraf Arrays von drei verschiedenen Herstellern und machte eine erneute Einstellung der Hybridisierungsbedingungen erforderlich. Dazu wurde analog zur Etablierung des Maussystems verfahren. Optimale Ergebnisse resultierten aus einem Protokoll ohne Zusatz konkurrierender Block-DNA und mit deutlich verkürzten Waschschritten bei reduzierter Temperatur (Abbildung 3.24).

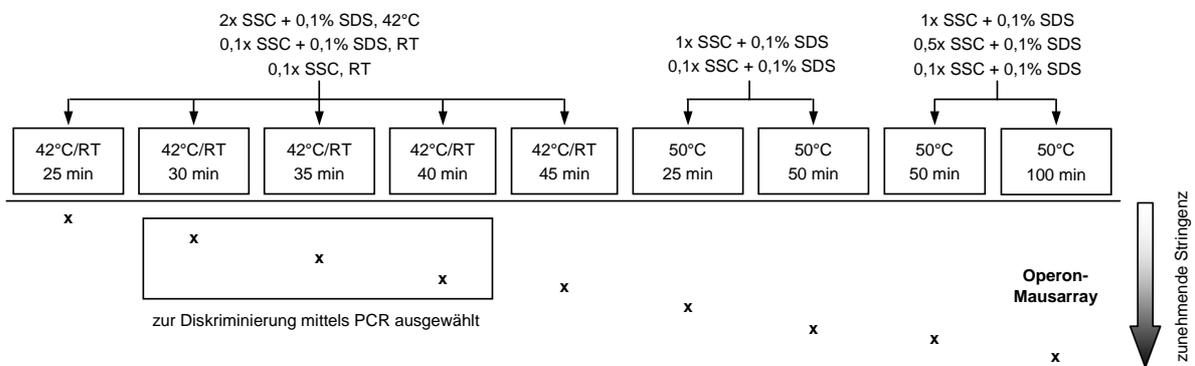


Abbildung 3.24 Etablierung der Hybridisierung auf humanen Operonarrays

Ganz rechts finden sich die Bedingungen des etablierten Mausarrays. Da diese Bedingungen deutlich zu stringent waren, wurden die Waschschritte erneut titriert und mittels PCR validiert.

Optimal waren Waschschritte in 2x SSC/SDS, 0,1x SSC/SDS und 0,1x SSC bei 42°C beziehungsweise Raumtemperatur für 35 Minuten. Mit diesen Bedingungen ließen sich humane Arrays aller Hersteller spezifisch hybridisieren.

Aus dem Labor von Ake Borg in Lund, von dem humane Arrays bezogen wurden, wurde ebenfalls eine Charge Mausarrays hybridisiert. Auch diese Arrays lieferten mit dem etablierten Mausprotokoll für Arrays aus San Antonio keine Signale. Unterschiede bestanden im verwendeten Spottingpuffer, der Feinbeschaffenheit der UltraGAPS-Oberfläche, der Oligonukleotidkonzentration, im Herstellungsprozess und in der Version des Oligosets. Da für humane Arrays beider Hersteller die gleichen Hybridisierungsbedingungen benutzt werden konnten, schienen die Unterschiede im Herstellungsverfahren als Ursache für das unterschiedliche Verhalten der Mausarrays unwahrscheinlich zu sein. Ob sie in der Version des Oligosets als zweiter möglicher Ursache zu suchen waren, ließ sich bisher nicht klären.

3.5.4 Auswahl der optimalen Herstellung humaner Operon-Arrays und Validierung

Mit dem neu eingestellten Hybridisierungsprotokoll wurden selbst hergestellte humane Arrays vom DKFZ in Heidelberg, aus San Antonio, Lund und vom MPI in München sowie von der Firma Operon fertig bezogene Arrays hybridisiert und verglichen (Abbildung 3.25).

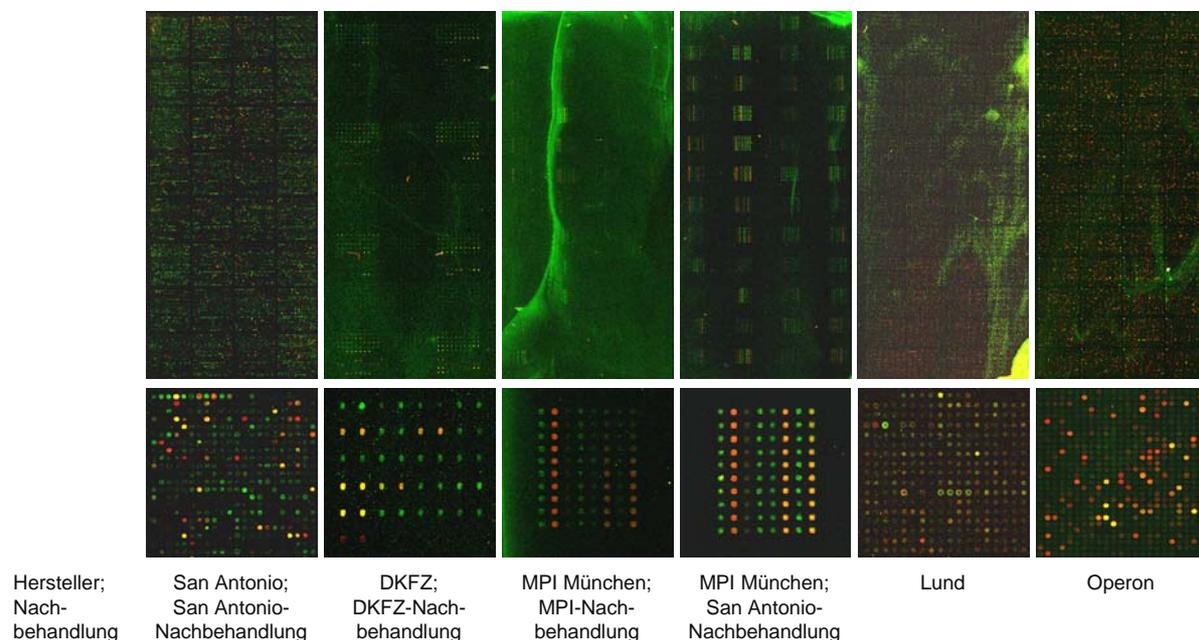


Abbildung 3.25 Vergleich von Operonarrays verschiedener Hersteller

Einzelzell-cDNA-Amplifikate wurden auf Arrays hybridisiert, die in verschiedenen Labors aus lyophilisierten Oligonukleotiden selbst hergestellt oder von der Firma Operon fertig bezogen wurden. Es traten Unterschiede hinsichtlich Hintergrund sowie Morphologie und Intensität der Spots auf. An den beiden Arrays aus dem MPI München ist auch der Einfluss der Nachbehandlung nach dem Spotten durch UV-Strahlung, Hitze und z.T. chemische Passivierung deutlich zu erkennen.

Gute Resultate lieferten neben den Arrays aus San Antonio die fertigen Slides der Firma Operon und solche aus dem MPI in München, die durch Ruhenlassen über Nacht, 600 mJ UV-Strahlung und 80°C für 4 h („San Antonio“) nachbehandelt worden waren. Nicht geeignet zu sein schienen Arrays aus Lund, eine Produktion am DKFZ in Heidelberg, sowie eine Nachbehandlung durch kurzes Erhitzen auf 100°C gefolgt von 200 mJ UV-Strahlung oder chemisches Passivieren der Oberfläche mit Succinanhydrid.

Die Eigenherstellung von qualitativ guten Arrays wäre somit in San Antonio oder am MPI in München möglich gewesen. Fertig gespottete Arrays zu verwenden bot hingegen den Vorteil, Qualitätskontrolle, Chargenkonstanz und die Etablierung und Einhaltung optimaler Bedingungen in fremde Hände legen zu können. Die fertigen Arrays der Firma Operon wurden zudem als einzige in einem Verfahren analog dem Tintenstrahldrucken hergestellt, wohingegen alle anderen Arrays mit Nadeln im Kontaktverfahren produziert wurden. Durch kontaktfreie Verfahren ist eine gleichmäßigere Morphologie der Spots bei besserem Hybridisierungsverhalten zu erreichen. Der Nachteil fertiger Arrays lag in der geringeren Wirtschaftlichkeit und dem fehlenden Einfluss auf die Herstellungsbedingungen hinsichtlich Oberfläche des Objektträgers, Spottingpuffer und -konzentration. Beispielsweise wurden die Arrays aus Lund mit einer Oligonukleotidkonzentration von nur 24 µM produziert, was im Vergleich zu einer Spottingkonzentration von 40 µM zu geringeren Signalintensitäten und dadurch verringerter Sensitivität und im Vergleich stärkerem Hintergrund führte und die nicht zufrieden stellenden Resultate erklärte. Die fertigen Arrays der Firma Operon wiederum wurden auf Schott Nexterion E- anstelle der bisher verwendeten UltraGAPS-Objektträger aufgebracht, waren mit dem Hybridisierungsprotokoll jedoch im Gegensatz zu auf dieser Oberfläche selbst hergestellten Arrays kompatibel.

Anhand der vergleichenden Experimenten wurde beschlossen, humane Arrays für den gesamten Genatz von Operon direkt zu erwerben, Mausarrays aus San Antonio zu verwenden und kleinere Arrays falls notwendig wegen der günstigen Erreichbarkeit am MPI für Psychiatrie in München selbst zu produzieren.

Die Ergebnisse der Hybridisierung humaner Proben auf den fertigen Arrays der Firma Operon wurden mittels genspezifischer PCR-Reaktionen validiert. Mit dem optimalen Hybridisierungsprotokoll betrug die Übereinstimmung zur PCR 18/24 Fälle (75,0%) falsch positive Hybridisierungsergebnisse traten in vier Fällen (16,7%) falsch negative in zwei Fällen auf (8,3%). Die Produkte dieser genspezifischen PCR-Reaktionen waren mit meist über 400 bp relativ lang. Durch die Bedingungen des Amplifikationsprotokolls mit dem Einsatz des fragmentierenden Zufallsprimer in der cDNA-Synthese nimmt die Wahrscheinlichkeit positiver PCR-Ergebnisse mit steigender Produktlänge ab. Daher ist die Rate falsch-positiver Hybridisierungen unter Umständen artifiziell hoch und könnte durch kürzere Produktlängen in der genspezifischen PCR (wie z.B. 280 bp für das Maussystem) weiter reduziert werden. Das humane System war also den Mausarrays in Sensitivität und Spezifität ebenbürtig.

4 Ergebnisse: Untersuchung disseminierter Tumorzellen

Ein zweiter Teil dieser Dissertation beschäftigte sich mit der Charakterisierung des Transkriptoms disseminierter Tumorzellen aus dem Knochenmark von Patientinnen mit Mammakarzinom unter Verwendung der Technologien, die im ersten Teil der Arbeit entwickelt und beschrieben wurden. Zur Detektion von Zellen epithelialen Ursprungs wurde eine Färbung gegen das Epitheliale Zell-Adhäsionsmolekül (EpCAM) benutzt. Von besonderem Interesse sind bei dieser Untersuchung Zellen, die in frühen Erkrankungsstadien gefunden werden, da sie die Ziele adjuvanter Therapien darstellen und Einfluss auf den weiteren Verlauf der Erkrankung haben. Ein Hindernis bei der Analyse solcher Zellen ist ihre extreme Seltenheit und Heterogenität. So wird mit Standardverfahren bei Brustkrebspatientinnen im Stadium M₀ nur in etwa jedem zehnten Knochenmark überhaupt eine EpCAM-positive disseminierte Tumorzelle detektiert.

4.1 Etablierung der Lineage Depletion aus humanem Probenmaterial

4.1.1 Auswahl des Depletionsverfahrens zur Anreicherung epithelialer Zellen

Die hier beschriebene Methode dient der Anreicherung epithelialer Tumorzellen in Knochenmarkspunktaten über die Standardverfahren hinaus. Sie basiert auf einer Färbung mit Antikörpern, die an paramagnetische Beads gekoppelt sind. In einem Magnetfeld können die markierten Zellen immobilisiert werden, die nicht markierten Zellen werden abgetrennt. Auf diese Weise lassen sich verschiedene Zellpopulationen voneinander separieren.

Grundsätzlich gibt es zwei Möglichkeiten der Selektion mittels dieser Technik. Bei der Positivselektion werden die Zielzellen mit Antikörpern markiert und aus dem Ansatz isoliert. Im Gegensatz dazu werden bei der Negativselektion alle nicht erwünschten Zellen markiert und entfernt, die Zielzellen indes bleiben unangetastet, was z.B. die Beeinflussung der Genexpression durch Antikörperbindung oder Scherkräfte während der Isolation minimiert. Als weiterer Vorteil müssen für die Negativselektion keine Marker bekannt sein, die von allen Zielzellen exprimiert werden. Die *Lineage Depletion* ist ein Verfahren der Negativselektion und eliminiert alle Zellen, die sich bereits in eine erythrozytäre oder leukozytäre Zellreihe differenziert haben. Das klassische Einsatzgebiet ist die Untersuchung undifferenzierter hämatopoietischer Zellen. Im Rahmen dieser Arbeit wurde versucht, die Methode für die Anreicherung von Tumorzellen zu modifizieren. Dazu mussten Oberflächenmarker zur Depletion ausgewählt werden, die auf hämatopoietischen Zellen und ihren Vorläufern, nicht aber auf Tumorzellen exprimiert sind.

Für humane Zellen wurde anfangs ein Antikörpercocktail aus α CD45, das auf allen Leukozyten exprimiert ist, und α CD235a als Oberflächenmarker der Erythrozyten sowie ein kommerzielles System zur Anreicherung von Stammzellen getestet. Die Antikörper wurden mit

Anreicherungssystemen verschiedener Hersteller (Dynal, Miltenyi und Stem Cell Technologies) verwendet. Dazu wurden CFSE-markierte epitheliale Tumorzellen der Mammakarzinom Zelllinie MCF-7 in peripheres Blut eingemischt. Zur ersten Aufreinigung des Probenmaterials vor der eigentlichen Anreicherung wurde ein Leukozytenfilm („Buffycoat“) präpariert oder eine Dichtegradientenzentrifugation durchgeführt und anschließend die Gesamtzahl der zur Anreicherung verwendeten Zellen bestimmt. Die Depletionssysteme von Dynal und Miltenyi entfernten aus diesen vorfraktionierten Proben CD45- und CD235a-positive Zellen, das System von Stem Cell Technologies CD2-, CD3-, CD14-, CD16-, CD19-, CD24-, CD56-, CD66b- und CD235a-positive Zellen. Nach der Anreicherung wurde die Gesamtzellzahl sowie unter dem Fluoreszenzmikroskop die Anzahl CFSE-markierter Zellen in den verschiedenen Fraktionen bestimmt (Tabelle 4.1). Die Gesamtwiederfindungsrate des Systems wurde aus der eingesetzten Zellzahl (Leukozyten + epitheliale Zellen) und der kumulativen Zellzahl aller Fraktionen nach der Anreicherung ermittelt. Die Wiederfindungsrate epithelialer Zellen pro Fraktion berechnete sich als prozentualer Anteil der CFSE-positiven Zellen an der eingesetzten Anzahl markierter Zellen. Idealerweise sollten während der Anreicherung weder Leukozyten noch Tumorzellen verloren gehen, also die Gesamtwiederfindungsrate 100 % betragen, sowie alle epithelialen Zellen sich in der Zielzell-Fraktion B befinden.

Tabelle 4.1 Systeme zur Depletion Lineagemarker-positiver Zellen aus humanem Probenmaterial

	Dynalsystem; Buffycoat		Dynalsystem; Dichtegradient 50% Perkoll		Miltenyisystem, Dichtegradient 50% Perkoll		Stem Cell Technologiessystem; Dichtegradient 50% Perkoll	
Zeitbedarf [h:min]	3:05		4:10		3:00		3:35	
Gesamtwiederfindungsrate	32,1 %		88,6%		105,5%		86,4%	
Fraktion	A ^a	B ^b	A	B	A	B	A	B
Wiederfindungsrate epithelialer Zellen	0,55%	0,02%	0,59%	17,4%	0,96%	15,2%	8,8%	0,28%

^a: Depletierte Fraktion, Depletionsmarker-positiv

^b: Zielzell-Fraktion, Depletionsmarker-negativ

Bei beiden Dynal-Ansätzen befanden sich auch in den Depletionsmarker-negativen Fraktionen viele Beads, die bei der weiteren Untersuchung am Mikroskop störten. Beim System von Stem Cell Technologies fanden sich die epithelialen Zellen in der falschen Fraktion. Mindestens einer der verwendeten Oberflächenmarker wurde offensichtlich auch von MCF-7 exprimiert. Am besten, auch was den Zeitbedarf anging, schnitt das System von Miltenyi ab. Dieses System arbeitet mit Säulen, an denen in einem Magnetfeld die markierten Zellen immobilisiert werden. Es gibt zwei Sorten von Säulen, Selektions- und Depletionssäulen. Obwohl es sich um ein Depletionsverfahren handelt, waren die Depletionssäulen nicht geeignet, da sich zu viele nicht mit Antikörpern markierte Zellen in der depletierten Fraktion befanden und daher verloren gingen (Abbildung 4.1).

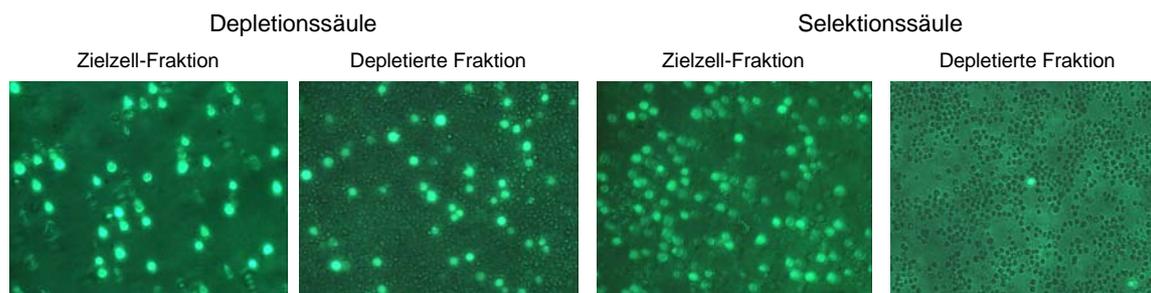


Abbildung 4.1 Vergleich verschiedener Säulen des Miltenyi-Systems

CFSE-markierte epitheliale Zellen (hell) wurden peripherem Blut zugesetzt. Nach der Depletion hämatopoietischer Zellen mit α CD45 und α CD235a sollen sich epitheliale Zellen (CD45-/CD235a-) in der Zielzell-Fraktion befinden. Die depletierte Fraktion wird im eigentlichen Experiment verworfen, epitheliale Zellen, die sich in dieser Fraktion befinden, gehen verloren.

4.1.2 Optimierung der Dichtegradientenzentrifugation

Da eine Dichtegradientenzentrifugation als erster Aufreinigungsschritt dem Buffycoat überlegen war, wurde dieses Verfahren genauer charakterisiert. Ziel ist eine möglichst weitgehende Abtrennung von Erythrozyten bei gleichzeitiger Anreicherung der epithelialen Zellen. In Abhängigkeit von der Dichte der Lösung sammeln sich während der Zentrifugation verschiedene Zelltypen in der Interphase, andere werden ins Pellet abzentrifugiert. Gängigerweise wird im Labor Ficoll oder eine 50%ige Perkolllösung mit einer Dichte von 1,065 g/cm³ verwendet. Die Dichte kann jedoch über den prozentualen Anteil an Perkoll reguliert werden. Granulozyten als Hauptfraktion der Knochenmarkzellen besitzen eine Dichte von 1,080-1,095 g/cm³, Erythrozyten eine von 1,09-1,11 g/cm³. Ficoll sowie verschiedene Perkollkonzentrationen unterschiedlicher Dichte wurden auf ihre Fähigkeit zur Anreicherung epithelialer Zellen untersucht. Dazu wurden wiederum peripheres Blut und CFSE-markierte MCF-7 Zellen verwendet (Tabelle 4.2).

Tabelle 4.2 Verfahren zur ersten Anreicherung epithelialer Zellen

Dichtegradient	Dichte [g/cm ³]	Gesamtzellzahl (Leukozyten + epitheliale Zellen) pro 1 ml Ausgangsmaterial		Wiederfindung epithelialer Zellen	
		Interphase	Pellet	Interphase	Pellet
Perkoll 40 %	1,053	20000	3500	2,9 %	0,05 %
Perkoll 50 %	1,065	56000	200	4,5 %	0,002 %
Perkoll 60 %	1,077	440000	100	12,5 %	0,0008 %
Perkoll 65 %	1,083	490000	3000	12,7 %	0,01 %
Perkoll 70 %	1,089	990000	2000	21,0 %	0,007 %
Perkoll 75 %	1,095	250000	400	3,0 %	0,002 %
Perkoll 80 %	1,100	0	1000	0,3 %	0,01 %
Perkoll 85 %	1,106	0	3000	0,2 %	0,007 %
Ficoll	1,077	150000	3000	2,6 %	0,04 %
Buffy Coat	-	2000	n.d. ^a	0,04 %	n.d.

^a: nicht durchgeführt

Eine optimale Ausbildung der Interphase mit Anreicherung der Tumorzellen, bestimmt durch die Gesamtzellzahl sowie die Wiederfindungsrate epithelialer Zellen in dieser Fraktion im Vergleich zum Pellet, trat bei 70% Perkoll auf. Allerdings waren in der Fraktion auch zahlreiche Erythrozyten enthalten, die durch Depletion jedoch quantitativ entfernt werden konnten. Bei Konzentrationen über 75% Perkoll wurden die Erythrozyten nicht mehr abgetrennt, geringere Konzentrationen führten zum Verlust epithelialer Zellen. Probleme traten allgemein durch die Erythrozyten auf, die ein Auszählen der Zellen bei geringen Perkollkonzentrationen im Pellet und bei hohen Konzentrationen in der Interphase erschwerten.

4.1.3 Wiederfindungsrate epithelialer Zellen

Von 100 zugegebenen epithelialen Tumorzellen in 7 ml peripherem Blut konnten nach Dichtegradientenzentrifugation mit 70%iger Perkolllösung und Depletion von CD45- und CD235a-positiven Zellen 52 epitheliale Zellen wieder gefunden werden. Vier Zellen befanden sich in der depletierten Fraktion, waren aber dem Aussehen nach nicht mehr vital. Die übrigen Zellen gingen verloren.

4.1.4 Schwierigkeiten bei der Untersuchung von Knochenmark

Sämtliche bisherigen Experimente wurden aus Verfügbarkeitsgründen mit peripherem Blut durchgeführt. Bei der Umstellung auf Knochenmark ergaben sich unerwartete Probleme. Die Depletion CD45- und CD235a-positiver Zellen aus peripherem Blut erreichte eine Abreicherung der Leukozyten um zwei Zehnerpotenzen, bei Knochenmarksproben wurden jedoch nur etwa 40% der Zellen entfernt. Von den verbliebenen Zellen waren in FACS-Analysen 55% weiterhin CD45 positiv. Diese Population wurde mit weiteren Färbungen im FACS näher charakterisiert. Es stellte sich heraus, dass die Lymphozytenpopulation (CD3+ oder CD19+) durch Depletion mit CD45 vollständig entfernt wurde, die Granulozyten (CD33+) aber zum Großteil im Ansatz verblieben (Abbildung 4.2). Da Granulozyten die Hauptfraktion der Knochenmarkszellenausmachen, stellt das ein nicht zu vernachlässigendes Problem dar.

Mögliche Ursachen können unterschiedliche Subtypen von CD45 auf peripheren Granulozyten und Vorstufen im Knochenmark, die von verschiedenen Antikörpern (nicht) erkannt werden, oder unterschiedlich starke Expression eines gemeinsamen Subtyps sein^{145, 146}. Die Färbeintensität für CD45 lag im FACS für die Mehrheit der Lymphozyten knapp eine Zehnerpotenz über der der Granulozyten. Da zudem auch der für die FACS-Analysen verwendete α CD45-Antikörper in der Depletion keine Verbesserung erbrachte, war die Ursache für die mangelnde Effizienz offensichtlich in einer für die Bindung an die Säule zu geringen Expression des Selektionsmarkers zu suchen.

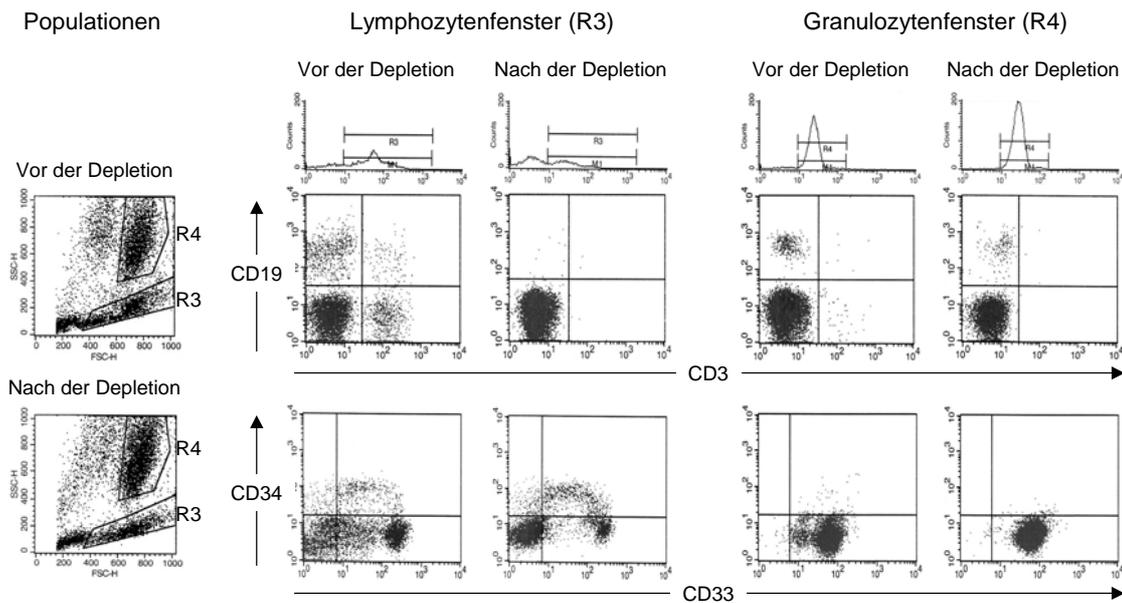


Abbildung 4.2 Probleme bei der Depletion aus Knochenmark

Vor und nach der Depletion CD45-positiver Zellen wurden Knochenmarksproben im FACS auf die Expression von Oberflächenmarkern zur Unterscheidung verschiedener hämatopoietischer Zelltypen untersucht. Lymphozyten (CD3+ bzw. CD19+) werden im Gegensatz zu Granulozyten (CD33+) nahezu vollständig entfernt.

Dem Problem der nicht ausreichenden Markierung und dadurch bedingten unvollständigen Depletion von Zellen kann mit einem hohen Überschuss an Antikörper, zwei aufeinander folgenden Depletionsrunden oder einer Kombination von Antikörpern gegen verschiedene Epitope eines Antigens begegnet werden. Eine weitere Möglichkeit zur Verstärkung von Färbung und Depletion besteht in der Verwendung PE- oder APC-konjugierter Antikörper anstelle unmarkierter Antikörper, Biotin- oder FITC-Konjugate kombiniert mit Beads, die gegen das Fluorochrom gerichtet sind^{147, 148}. Außerdem können zusätzliche Oberflächenmarker für die Depletion verwendet werden. In Frage kam hierzu nach den FACS-Analysen CD33. Das Molekül wurde auf epithelialen Tumorzellen noch nicht beschrieben, ist in frühen Entwicklungsstadien der Granulozyten hoch exprimiert und wird später herunterreguliert^{149, 150}. Späte Stadien weisen jedoch eine höhere Expression von CD45 auf, was in der Kombination mit CD33 eine gute Abdeckung sämtlicher Differenzierungsstufen erlauben sollte.

Die genannten Möglichkeiten wurden auf ihr Potenzial zur Optimierung der Depletion hin untersucht und die Ergebnisse in Tabelle 4.3 zusammengefasst. Von Interesse ist dabei die Zellzahl der Negativfraktion der Depletion, da sich epitheliale Tumorzellen dort befinden. Als Maß für die Effizienz der Depletion diente der Prozentsatz depletierter Zellen, der sich aus der Differenz der Ausgangszellzahl und der Zellzahl der Negativfraktion berechnen lässt. Bei mehreren Experimenten unter den gleichen Bedingungen ist hier zusätzlich die Standardabweichung angegeben.

Tabelle 4.3 Depletion von $1 \cdot 10^7$ Knochenmarkzellen durch verschiedener Antikörperkonjugate

Depletionsantikörper	Anzahl der Experimente	Mittlere Zellzahl Negativfraktion	Deletiert wurden
α CD45-Beads	4	$4,3 \cdot 10^6$	57% +/- 14%
α CD45-Beads; 2 Runden	3	$2,9 \cdot 10^6$	70% +/- 9%
α CD45-FITC / α FITC-Beads	1	$1,4 \cdot 10^6$	86%
α CD45-Beads + α CD45-FITC / α FITC-Beads	2	$2,0 \cdot 10^6$	80% +/- 4%
α CD45-FITC / α FITC-Beads + α CD33-PE / α PE-Beads	1	80'000	99%
α CD45-PE + α CD33-PE / α PE-Beads	1	200'000	98%
αCD45-APC + αCD33-APC / αAPC-Beads	3	310'000	97% +/- 2%

Da die Negativfraktion mittels weiterer Färbungen auf das Vorhandensein von Tumorzellen untersucht werden sollte, durfte die Zellzahl nicht zu hoch sein, um möglichst alle Zellen analysieren zu können. Zu geringe Zellzahlen waren jedoch ebenfalls nicht wünschenswert, da sonst die unvermeidlichen Verluste während der weiteren Schritte sehr stark ins Gewicht fallen würden. Als optimal konnten Zellzahlen von $1-3 \cdot 10^6$ Zellen in der Negativfraktion angesehen werden. Bei der Aufarbeitung von etwa 3 ml Knochenmark über einen Dichtegradienten mit 70% Percoll wurden im Schnitt $1-2 \cdot 10^8$ Leukozyten erhalten, es mussten daher 97-99,5 % der Zellen durch Depletion entfernt werden. Dies war vor allem durch die Verwendung PE-konjugierter Antikörper zu gewährleisten. Da jedoch für die nachfolgende Färbung zur Detektion epithelialer Zellen ein Cy3-markierter α EpCAM-Antikörper benutzt wurde, verbot sich die Depletion mit PE-konjugierten Antikörpern. Weil die Emissionswellenlängen von Cy3 und PE mit 570 nm bzw. 578 nm zu nahe beieinander liegen, wäre es nicht möglich gewesen, zwischen nicht vollständig entfernten PE-markierten hämatopoietischen Zellen und Cy3-markierten epithelialen Zellen zu unterscheiden.

Als Alternative boten sich APC-konjugierte Depletionsantikörper, die jedoch etwas weniger effektiv waren, oder die Verwendung eines weiteren Oberflächenmarkers an. Aus Hämatoxylin/Eosin-Färbungen von Ausstrichen der Negativfraktion der Depletion wurde deutlich, dass auch nach Entfernung CD33- und CD45-positiver Zellen Granulozyten die verbliebene Hauptfraktion ausmachten. Es wurden nach der Untersuchung von acht Knochenmarkspunktaten 67,8% segmentkernige neutrophile Granulozyten, 2,0% eosinophile Granulozyten, 2,5% basophile Granulozyten, 15,2% Lymphozyten und 12,4% immature Granulozyten sowie andere Zellen gefunden.

Für alle potentiellen zusätzlichen Marker zur Depletion, deren Expression in den entsprechenden Differenzierungsstadien der Granulozyten als hoch beschrieben war, beispielsweise CD15, CD55, CD59, CD162 oder CD43^{149, 151, 152}, lagen Berichte für die Expression auf Tumorzellen vor. Andere Marker wie CD14, CD16, oder CD56 wurden in FACS-Analysen als auf den relevanten Zellen nicht oder zu gering exprimiert gefunden. Um weitere mögliche Oberflächenmarker für die Depletion nicht-epithelialer Zellen zu definieren, wurden daher beschlossen, zunächst disseminierte Tumorzellen in Knochenmarkpunktaten von Brustkrebspatientinnen nach Depletion von CD45- und

CD33-positiven Zellen mit APC-konjugierten Antikörpern zu isolieren und das Transkriptom zu analysieren.

4.2 Analyse von Knochenmarksaspiraten bei Mammakarzinompatientinnen

Im Zeitraum von Januar 2005 bis November 2005 wurden 43 Knochenmarkaspirate von Patientinnen mit Mammakarzinom untersucht. 31 Proben stammten aus der II. Medizinischen Klinik bzw. der Gynäkologie des Zentralklinikums Augsburg, 12 Proben aus der Universitätsfrauenklinik des Klinikums Innenstadt der LMU München. Eine Übersicht des Kollektivs, soweit bei Abschluss der Dissertation bekannt, befindet sich in Absatz 2.6 des Methodenteils.

4.2.1 Detektionsrate EpCAM-positiver Zellen und Isolation der Nukleinsäuren

Die Proben wurden dem beschriebenen Verfahren von Depletion CD33-, CD45- und CD235a-positiver Zellen nach Anreicherung über einen Dichtegradienten mit 70% Percoll unterworfen. Anschließend erfolgte eine Färbung gegen EpCAM zur Detektion epithelialer Tumorzellen. Sieben Präparate wurden zusätzlich mit dem Hoechst 33342-Farbstoff-gefärbt (0). Einzelne EpCAM-positive Zellen wurden isoliert und das Transkriptom sowie das Genom untersucht. Positive Zellen wurden in 38/43 (88,4%) aller Aspirate gefunden. Durchschnittlich fanden sich vier EpCAM-positiv pro 10^6 Zellen (SD=4, min=1, max=16, median=2). Eine Isolation und Amplifikation der mRNA gelang bei 90/142 Zellen (63,4%), der genomischen DNA (gDNA) bei 104/142 Zellen (73,2%) und beider Nukleinsäuren aus derselben Zelle in 73/142 Zellen (51,4%). Die Erfolgsraten waren abhängig davon, ob eine Färbung auch mit dem Hoechst-Farbstoff durchgeführt worden war. Da diese Färbung aktive zelluläre Prozesse voraussetzt, muss sie in einem Kulturmedium bei 37°C und über längere Zeiträume von bis zu zwei Stunden durchgeführt werden¹³¹. Diese Bedingungen beeinträchtigen offensichtlich die Vitalität der Zellen, so dass von vielen Proben keine mRNA mehr isoliert werden konnte (Tabelle 4.4).

Tabelle 4.4 Erfolgsraten der Isolierung von Nukleinsäuren aus einzelnen EpCAM-positiven Zellen

	Zellen mit Hoechst-Färbung	Zellen ohne Hoechst-Färbung	Alle Zellen
Nur mRNA-Isolierung erfolgreich	3/36 (8,3%)	14/106 (13,2%)	17/142 (12,0%)
Nur gDNA-Isolierung erfolgreich	15/36 (41,7%)	16/106 (15,1%)	31/142 (21,8%)
Beide Isolierungen erfolgreich	7/36 (19,4%)	66/106 (62,3%)	73/142 (51,4%)
Keine Isolierung erfolgreich	11/36 (30,6%)	10/106 (9,4%)	21/142 (14,8%)

4.2.2 Doppelfärbung mit dem Farbstoff Hoechst 33342 und epithelialen Markern

Hoechst 33342 ist ein DNA-bindender Farbstoff, der von Zellen aktiv über den ABCG2-Transporter ausgeschleust werden kann, welcher von Stamm- und Vorläuferzellen exprimiert wird. Die Färbung kann daher zur Identifizierung dieser Zellen benutzt werden, da sie im Gegensatz zu differenzierten Zellen ungefärbt erscheinen. Außerdem wurde der Transporter auf gegen mehrere Zytostatika (Multi Drug)-resistenten Tumorzellen beschrieben.

Bei Doppelfärbungen wurde zuerst der Hoechst-Farbstoff verwendet und dann die Immunfluoreszenz durchgeführt. Wurde die Probe dabei konsequent kühl gehalten, diffundierte der Hoechstfarbstoff auch bei intrazellulären Färbungen wie z.B. gegen epitheliale Zytokeratine nicht passiv aus der Zelle. Abbildung 4.3 zeigt das Ergebnis einer Doppelfärbung von humanem Knochenmark mit dem Hoechst-Farbstoff kombiniert mit einer Färbung gegen EpCAM. Die Emissionswellenlänge für den Hoechstfarbstoff liegt bei 461 nm, für EpCAM (Cy3) bei 570 nm. Zu erkennen ist eine einzelne EpCAM-positive disseminierte Tumorzelle, die gleichzeitig Hoechst-negativ ist.

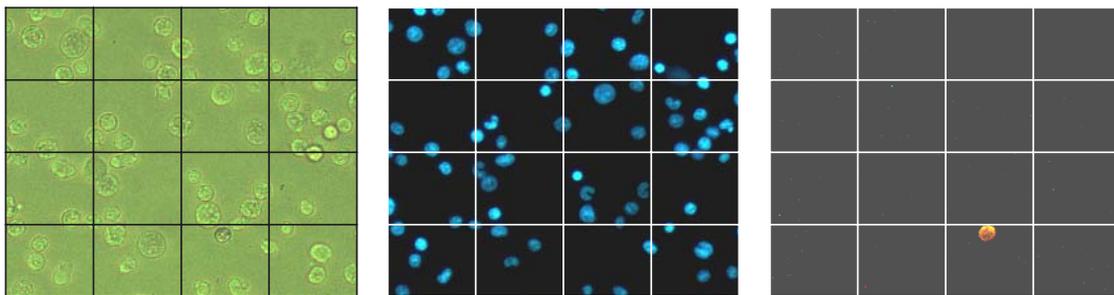


Abbildung 4.3 Doppelfärbung mit Hoechst 33342 und α EpCAM-Cy3

Humanes Knochenmark einer Tumorpatientin wurde nach Depletion CD33- und CD45-positiver Zellen zuerst mit dem Hoechst-Farbstoff und anschließend gegen EpCAM gefärbt. Links das Bild der Färbung im Durchlicht, in der Mitte die Hoechst-Färbung. Rechts ist eine EpCAM-positive Zelle zu erkennen, die im Vergleich mit den anderen Ansichten Hoechst-negativ ist.

Offensichtlich existieren also im Knochenmark von Brustkrebspatientinnen EpCAM-positive Zellen, die den Hoechst-Farbstoff ausschleusen und daher möglicherweise Vorläuferzellcharakter besitzen. Da jedoch die Rate erfolgreicher Isolationen des Transkriptoms nach der Hoechst-Färbung signifikant erniedrigt war (Absatz 4.2.1; $p < 0,0001$), wurde auf die Färbung wieder verzichtet.

4.2.3 Genexpressionsanalyse auf Nylonarrays

Um einen ersten Eindruck von der Genexpression der isolierten Zellpopulation zu gewinnen, wurden die 90 EpCAM-positiven Zellen, aus denen erfolgreich mRNA isoliert werden konnte, auf Nylonfilter hybridisiert. Die 181 Sonden des Arrays umfassten Gene zur Identifizierung des histologischen Ursprungs der Zelle, tumorassoziierte Gene, Wachstumsfaktoren sowie zahlreiche Proteasen und deren Inhibitoren (Abbildung 4.4).

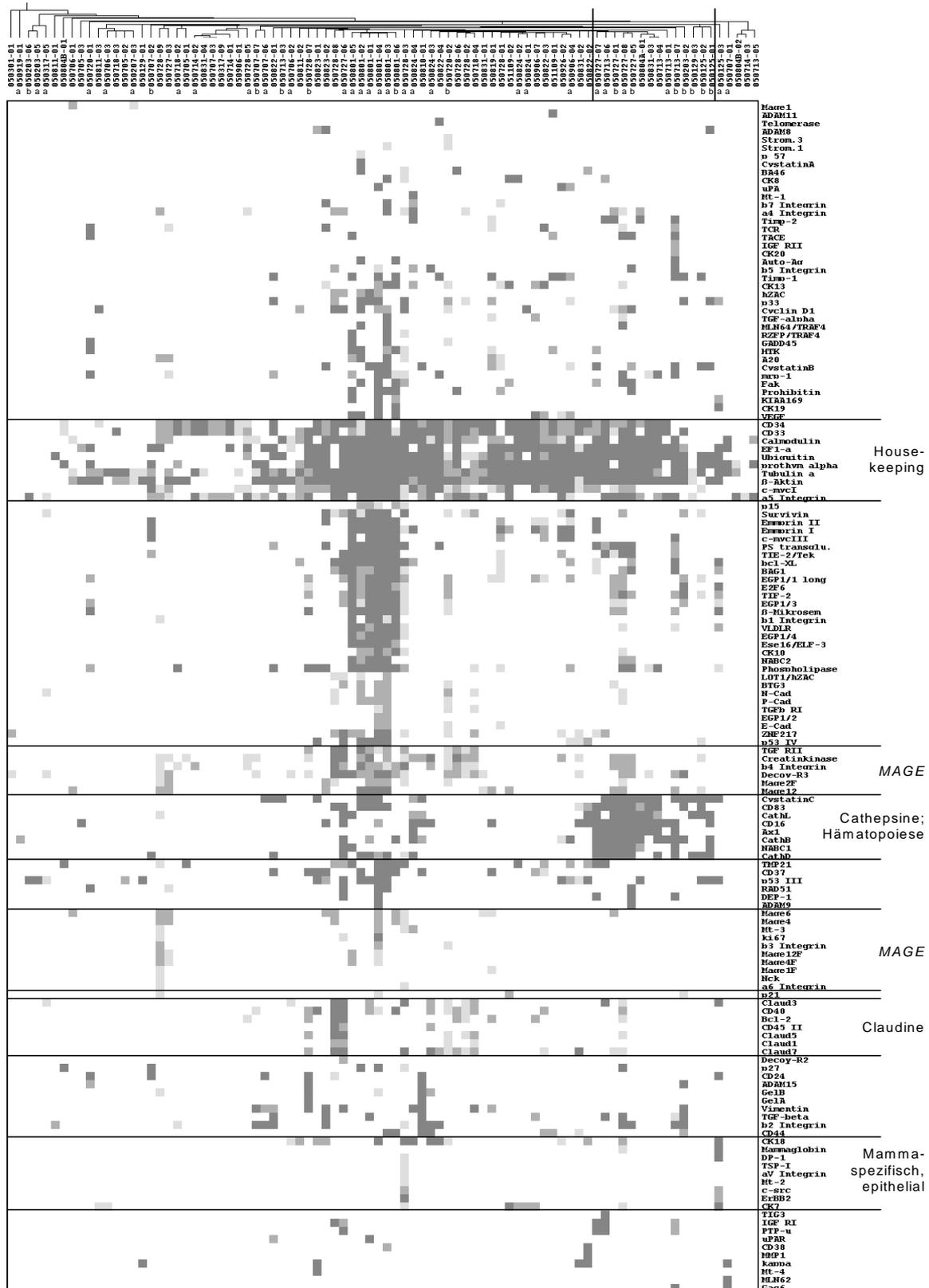


Abbildung 4.4 Cluster der Genexpressionsdaten EpCAM-positiver Zellen auf Nylonarrays

86 Zellen (horizontal) weisen für mindestens eines von 146 Genen des Nylonfilters (vertikal) einen Messwert auf. 35 Gene und vier Zellen zeigten kein Signal. Je dunkler die Färbung des Kästchens, desto stärker die Expression auf einer Skala von 0 bis 3. Koexprimierte Gene sind am Rand zu Gruppen zusammengefasst. Zellen mit aberrantem CGH-Profil sind mit einem „a“ gekennzeichnet, Zellen mit balanciertem Profil mit „b“. Von Zellen ohne Markierung existiert kein CGH-Profil.

Bestimmte Gruppen von Genen waren untereinander koexprimiert und sind am Rand der Abbildung markiert und mit dem jeweiligen Charakteristikum bzw. der hauptsächlich vertretenen Genfamilie bezeichnet. Darunter fanden sich neben Housekeeping-Genen auch solche, die mit epithelialen Tumoren assoziiert sind, wie die *MAGE*-Familie und mammaspezifische Marker sowie verschiedene an der Bildung der Tight Junctions beteiligte Claudine. Die Gengruppe „Cathepsin-Proteasen, Hämatopoese“ wurde hauptsächlich von Zellen mit balanciertem CGH-Profil exprimiert. Ob es sich bei diesen Zellen ebenfalls um Tumorzellen handelt, muss durch umfangreichere Genexpressionsanalysen oder Untersuchungen des Genoms im subchromosomalen Bereich beantwortet werden, da CGH-Veränderungen kein notwendiges Charakteristikum von Tumorzellen darstellen (vgl. auch Absatz 5/ Diskussion). Bereits am reduzierten Gensatz des Nylonarrays und bei der nur semi-quantitativ vorgenommenen Auswertung der Hybridisierungssignale fiel jedoch auf, dass auch als eigentlich hämatopoietisch einzustufende Gene wie *CD33*, *CD34*, *CD16* und *CD83* von Zellen mit aberrantem CGH-Profil, bei denen es sich also sicher um Tumorzellen handelt, zum Teil stark exprimiert werden. Auch bei den Zytokeratinen fand keine klassische organspezifische Expression statt, da die Keratine der terminalen epithelialen Differenzierung *CK8* und *CK18* nur von 19 Zellen (21,3%), das Paar *CK8/CK18* sogar nur von einer Zelle exprimiert wurde. Dagegen fand eine Expression von Zytokeratinen wie *CK10*, die von glandulären Epithelien typischerweise nicht exprimiert werden sollten, in Zellen mit CGH-Aberrationen statt. Auch wenn es sich bei vielen untersuchten Zellen also sicher um Tumorzellen handelte, stellten die Zellen in ihrer Genexpression meist keine typischen terminal differenzierten epithelialen Zellen dar.

Diskrepanzen bestanden auch zwischen der Expression eines Gens auf mRNA-Ebene und der Proteinmenge auf der Zelloberfläche. Sämtliche Zellen wurden beispielsweise aus der Fraktion isoliert, in der CD33-positive Zellen depletiert worden waren. Dennoch wiesen 56 Zellen eine CD33 mRNA-Expression auf. Im Gegensatz dazu waren alle auf Proteinebene EpCAM-positiv, die mRNA wurde jedoch nur in 22 Zellen transkribiert. Da für die Depletion jedoch relativ große Mengen des Proteins auf der Zelloberfläche vorhanden sein müssen, wie an den Problemen bei der Etablierung des Depletions-Protokolls aus Knochenmark deutlich wurde, und Proteine eine sehr lange Halbwertszeit besitzen können, so dass das Gen selten transkribiert werden muss, erscheint dies kein absoluter Widerspruch zu sein.

4.2.4 Genexpressionsanalyse auf Oligonukleotidarrays

Fünf Zellen wurden außer auf den Nylonarray auch auf humane Oligonukleotidarrays hybridisiert, die 38880 Sonden umfassten. Es handelte sich um Proben verschiedener Patientinnen, Tumorstadien (M_0 und M_1) und -typen (duktal und lobulär). Auswertungen dieser Experimente sind wegen der geringen Probenzahl nur bedingt möglich und mit Vorsicht zu betrachten, einige Beobachtungen waren jedoch trotzdem interessant.

Beispielsweise zeigte sich in einer hierarchischen Clusteranalyse eine Abhängigkeit der Genexpressionsprofile vom Tumorstadium (Abbildung 4.5).

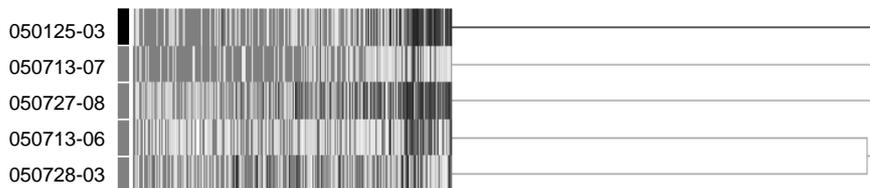


Abbildung 4.5 Clusteranalyse der Genexpression EpCAM-positiver Zellen auf Oligonukleotidarrays

Es wurden fünf Zellen von vier Patientinnen auf humane Oligonukleotidarrays hybridisiert. Zelle 050125-03, gekennzeichnet mit einem schwarzen Balken, stammte aus dem Knochenmark einer Patientin im Tumorstadium M_1 , die übrigen Zellen (graue Balken) von M_0 -Patientinnen.

Eine geringe Übereinstimmung zu Aussagen bezüglich des Tumorstadiums wurde jedoch zu einer publizierten Untersuchung gefunden, in der für individuelle Patientinnen eine Metastasierung innerhalb der nächsten fünf Jahre aus der Expression einer Signatur mit 70 Genen in Primärtumor vorhergesagt wurde¹³. 39 dieser 70 Gene konnten eindeutig einem Transkript und damit einer Sonde auf dem Operon-Array zugeordnet und ihre Expression in den disseminierten Zellen bestimmt werden (Abbildung 4.6).

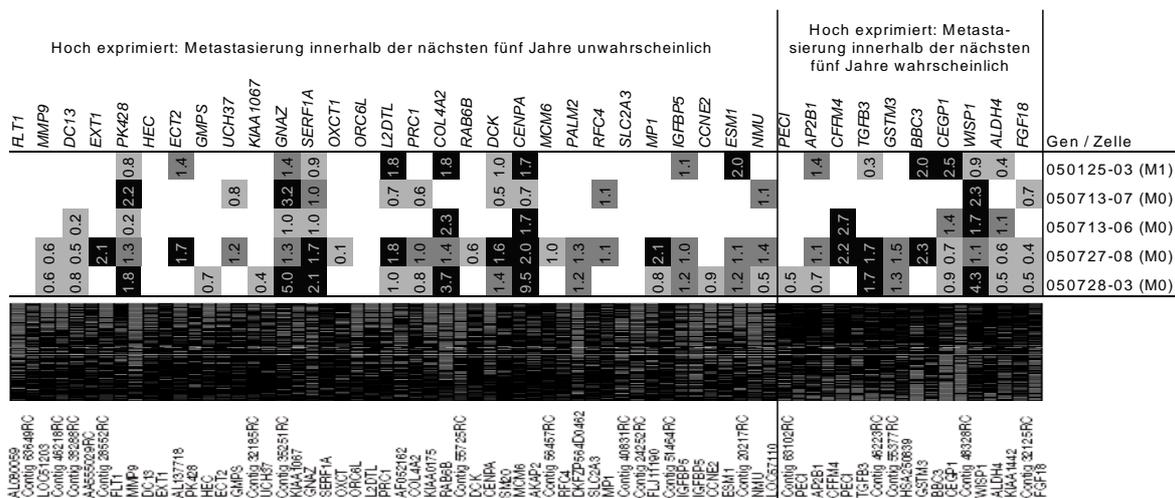


Abbildung 4.6 Expression einer publizierten Signatur zur Vorhersage der Metastasierung in EpCAM-positiven Zellen

Aus der Arbeit von Laura van't Veer et al¹³ wurde die publizierte 70 Gen-Signatur zur Vorhersage einer Metastasierung innerhalb der nächsten fünf Jahre sowie die Expressionsmatrix der zur Erstellung der Signatur verwendeten Primärtumore entnommen. Die Expression der 39 eindeutig zuzuordnenden Gene in den EpCAM-positiven Zellen befindet sich darüber. Je dunkler jeweils die Graustufe, desto höher die Expression des zugehörigen Gens, nicht nachweisbare Messwerte sind weiß dargestellt.

Für die Zelle aus dem Knochenmark der Patientin, bei der eine Metastasierung bereits eingetreten war, konnte dieses Ereignis aus der Expressionssignatur nicht vorhergesagt werden, da die Zelle zwei der zehn M_1 -typischen Gene und vier der 29 M_0 -typischen Gene stark exprimiert. Auch bei den Patientinnen im Tumorstadium M_0 schien eine Vorhersage schwierig, da hohe Expressionswerte

sowohl bei Genen, die für eine zukünftige systemische Progression gesprochen hätten als auch bei solchen, bei denen dieses Ereignis unwahrscheinlich wäre, gleich verteilt vorkamen.

Eine Abhängigkeit des Genexpressionsprofils der EpCAM-positiven Zellen war hingegen vom Tumortyp zu beobachten. 33 Gene waren mit einem p-Wert von unter 0,01 mit diesem Kriterium korreliert. Anhand dieser Genliste würde der Primärtumor, dessen Typ zum Zeitpunkt der Untersuchung unbekannt war, in einer hierarchischen Clusteranalyse den nicht-duktalem Karzinomen zugeordnet. Diese Eingruppierung konnte mittlerweile bestätigt werden, da es sich nach Aussage des beurteilenden Pathologen bei dem Tumor der Patientin 050125 um ein Mammakarzinom vom lobulären Typ handelte.

Unter den 20 der 33 relevanten Gene mit bekannter Funktion waren fünf an der DNA-Konformation und Transkription, vier an der Mitose, drei an der Zelladhäsion und eines an der Differenzierung beteiligt. Vier kodierten Transportproteine, drei Membranproteine (Abbildung 4.7).

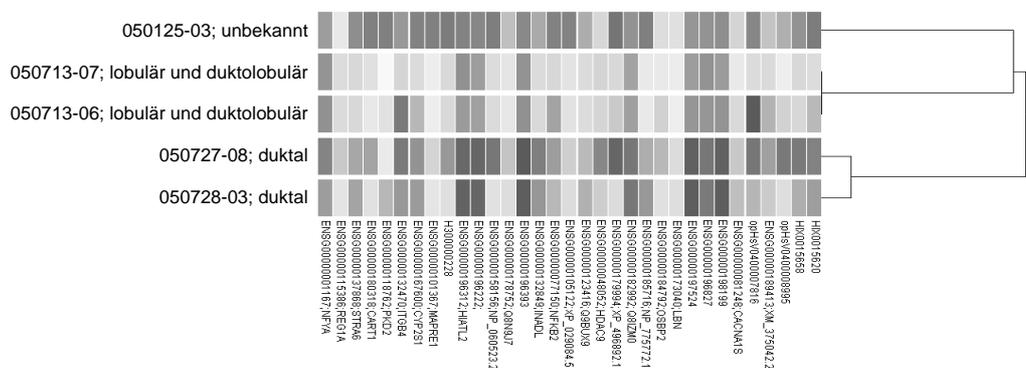


Abbildung 4.7 Assoziation von Genexpression und Tumortyp in EpCAM-positiven Zellen

Die Expressionswerte von 33 Genen erlauben eine Gruppierung der disseminierten Zellen nach dem Typ des Primärtumors. Je dunkler dabei die Graustufe, desto stärker die Expression des jeweiligen Gens.

Weder die Abhängigkeit der Genexpressionsprofile vom Tumorstadium noch vom Tumortyp wurden durch den Estrogenrezeptorstatus des Primärtumors überlagert, obwohl dieser als einer der größten Einflussfaktoren auf die Genexpression gilt¹². Sowohl beim Primärtumor auf Proteinebene als auch in den EpCAM-positiven Zellen auf mRNA-Ebene war die Expression des Estrogenrezeptors heterogen verteilt. Zellen, die auf dem Array ein deutliches Signal für *ERI* zeigten (050713-06, 050727-08, 050728-03), exprimierten auch mehrere der in der Literatur genannten estrogenregulierten Gene wie *CDH3*, *EGFR*, *TIMP3*, *CYP2B*, *CEACAM5*, *TFF3* und *HDAC1*¹⁵³⁻¹⁵⁵. Zu den Genen deren Expression mit dem Estrogenrezeptorstatus assoziiert war, gehörten außerdem Sonic hedgehog (*SHH*) und Indian hedgehog (*IHH*), zwei Liganden des Hedgehog-Signalweges. Tabelle 4.5 zeigt ihre Expression zusammen mit der weiteren Gene des Signalwegs sowie die Funktion der Gene in der Kaskade. Ein Expressionswert über 1,5 wurde als dabei als „hoch exprimiert“ gewertet.

Da der Hedgehog-Signalweg normalerweise nur während der Embryonalentwicklung und in Stammzellen adulter Gewebe aktiv ist, war die Detektion dieser Transkripte ungewöhnlich. Neben diesen Genen waren jedoch auch andere, deren Expression in der Literatur eher für Stamm- und

Progenitorzellen als für differenzierte Zellen der Mamma beschrieben ist, in den EpCAM-positiven Zellen exprimiert (Tabelle 4.6), Stammzelleigenschaften bei diesen Zellen scheinen daher möglich.

Tabelle 4.5 Expression verschiedener Gene des Hedgehog-Signalwegs in EpCAM-positiven Zellen

Gen	Funktion	050125-03	050713-07	050713-06	050727-08	050728-03
<i>SHH</i>	Ligand	1,3	4,1	11,7	1,1	3,6
<i>IHH</i>	Ligand	1,5	4,5	13,6	1,9	7,2
<i>DHH</i>	Ligand	3,5	- ^a	-	1,3	-
<i>PTCH1</i>	Rezeptor, Negativregulator	2,2	-	-	1,7	-
<i>PTCH2</i>	Rezeptor, Negativregulator	-	2,1	-	1,1	0,6
<i>HIP</i>	Rezeptor, Negativregulator	-	-	-	-	-
<i>SMO</i>	Inhibitor	-	-	-	-	-
<i>GLI1</i>	Effektor	-	-	-	-	-
<i>GLI2</i>	Effektor	-	-	-	-	-
<i>GLI3</i>	Effektor	-	-	-	2,3	-
<i>GLI4</i>	Effektor	3,7	-	-	2,0	1,0
<i>FU</i>	Inhibitor	2,4	-	-	1,0	0,8
<i>SUFU</i>	Inhibitor	-	-	-	3,4	-

^a: nicht nachweisbar

Tabelle 4.6 Expression potentieller Stammzellmarker in EpCAM-positiven Zellen

Gen/Zelle/ Literatur	050125- 03	050713- 07	050713- 06	050727- 08	050728- 03	Al- Hajj ¹⁵⁶	Shackle- ton ¹⁵⁷	Stingl ¹⁵⁸	Sonstige ^{159, 160}
Stammzellmarker									
<i>CD44</i>	2,4 ^a	- ^b	7,4	-	2,1	+			
<i>CD24</i>	1,7	-	-	-	-	-/low		medium	
<i>EpCAM (Esa)</i>	0,1	0,9	0,1	0,7	0,3	+/-			
<i>CD29</i>	1,3	-	1,9	1,6	5,3		+		
<i>ITGA6</i>	-	-	-	-	1,0		+	+	
<i>LY6H (Sca-1)</i>	-	-	-	0,1	2,3		+	+	
<i>ER^c</i>	0,9	-	1,3	1,5	5,6				+
<i>MUC-1</i>	1,5	-	-	1,0	-				+
<i>VIM</i>	2,0	1,3	0,6	0,7	0,7				+
<i>ACTA2 (Sma)</i>	0,5	1,7	0,3	0,9	2,0				+
<i>KRT5</i>	0,7	1,3	0,6	0,4	0,6				+
<i>KRT6</i>	1,4	1,1	-	2,1	0,8			+	+
<i>KRT19</i>	1,5	-	24,3	1,5	6,7			+	+
<i>KRT14</i>	1,8	1,8	0,5	1,1	1,8		+	+	+
Differenzierungsmarker									
<i>KI67</i>	2,4	-	0,4	0,7	2,2				-/Tumor+
<i>KRT8</i>	1,0	1,0	1,1	1,1	3,7				-
<i>KRT18</i>	0,7	4,1	0,8	1,0	2,0		-	selten	-
<i>KRT1</i>	-	-	-	0,8	0,6				-
<i>KRT10</i>	1,3	-	-	1,3	0,9				-

^a: Werte über 1,5 wurden als erhöhte Expression gewertet

^b: nicht nachweisbar

^c: Estrogenrezeptor

4.2.5 Untersuchung der genomischen DNA

Bisher wurde das Genom von 46 der 73 Zellen, von denen auch die mRNA isoliert wurde, untersucht. Die Hybridisierung in der CGH gelang bei 45 Zellen (97,8%). 15 der 45 Zellen zeigten balancierte CGH-Profile, 30 Zellen Aberrationen. Die Gewinne und Verluste chromosomalen Materials dieser Zellen sind in Abbildung 4.8 zusammengestellt.

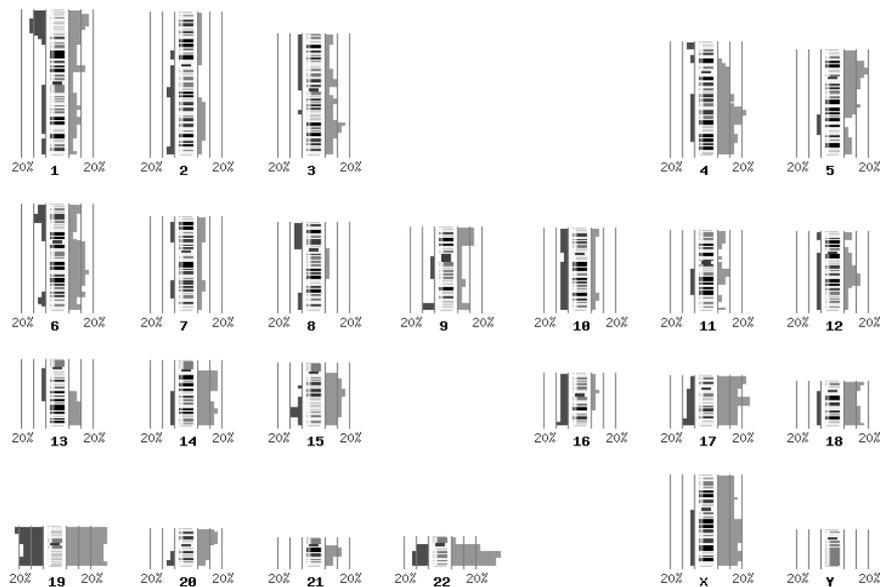


Abbildung 4.8 Amplifikationen und Deletionen der 30 Zellen mit CGH-Veränderungen

Für jedes Chromosom 1 bis X sind die Deletionen im jeweiligen Bereich als schwarzer Balken am linken Rand, die Amplifikationen als grauer Balken am rechten Rand als Summe aller Zellen dargestellt. Heterochromatinbereiche wie die p-Arme der Chromosomen 13-15 sowie Chromosom Y wurden nicht ausgewertet. Die einzelnen CGH-Profile aller Zellen befinden sich im Anhang der Dissertation.

Häufige Gewinne fanden sich auf Chromosom 1p, 4q, 5p, 6q, 12q, 15q, 17, 18, 19, 22 und X. Verluste traten vor allem auf Chromosom 1p, 19 und 20 auf.

Insgesamt wurden mehr Amplifikationen als Deletionen beobachtet. Besonders häufig war eine Vermehrung chromosomalen Materials im Bereich des *HER2*-Gens auf Chromosom 17q21. Von den acht betroffenen Zellen zeigten drei einen Gewinn des gesamten Chromosoms, drei eine Amplifikation nur im Bereich 17q21 und zwei Zellen eine Mischform. Allerdings wies nur eine Zelle mit Amplifikation auch eine deutliche Expression von *HER2* auf mRNA-Ebene auf. Die Ursache könnte einerseits in den Hybridisierungseigenschaften des verwendeten Nylonarrays liegen, andererseits könnte die mRNA auch wegen der Langlebigkeit vieler Proteine auf der Zelloberfläche zum Zeitpunkt der Zellyse tatsächlich nicht transkribiert worden sein. Dies bleibt durch weitere Untersuchungen, beispielsweise die Hybridisierung auf Oligonukleotidarrays zu klären.

Abschließend wurden die Zellen anhand der vorliegenden genomischen Aberrationen auf ihre Verwandtschaft untersucht. Sie ließen sich durch ein hierarchisches Clusterverfahren in zwei Gruppen

aufteilen: Zellen mit Deletionen auf Chromosom 1p, 17, 22 und 19, dafür mit Gewinnen in zahlreichen anderen Regionen des Genoms, sowie Zellen mit Amplifikationen auf Chromosom 1p, 17, 22 und 19, die seltener andere Veränderungen trugen (Abbildung 4.9). Wurden von einer Patientin mehrere Zellen analysiert, fielen diese meist in eine gemeinsame Gruppe.

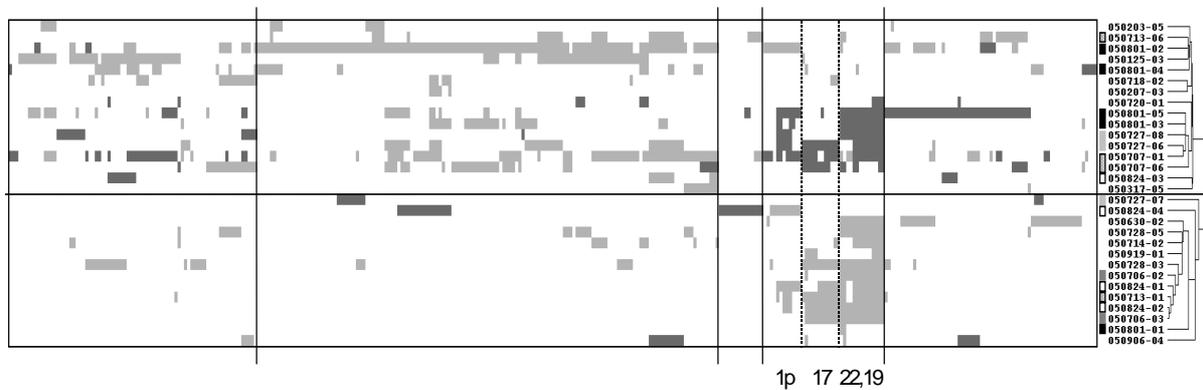


Abbildung 4.9 Hierarchische Clusteranalyse der CGH-Veränderungen aller Zellen

Die 30 Zellen mit CGH-Veränderungen trennen sich nach den Aberrationen auf Chromosom 1p, 17, 22 und 19 in zwei Gruppen auf. Zellen mit Deletionen in diesen Bereichen (schwarz) zeigen häufiger Amplifikationen (grau) in anderen Regionen des Genoms, wohingegen Zellen mit Amplifikationen in diesen Bereichen seltener andere Veränderungen aufweisen. Mehrere Zellen einer Patientin sind am Rand mit einem gleichen Balken gekennzeichnet.

5 Diskussion

Isolation und Amplifikation von mRNA aus Einzelzellen

Das in dieser Dissertation benutzte Protokoll zur Isolation und Amplifikation von mRNA aus Einzelzellen sowie dessen Vorgänger⁶⁴ basieren auf einer beadgekoppelten gezielten Isolation der mRNA über den Poly A-Schwanz. Andere RNA-Spezies sowie Proteine und auch die genomische DNA werden aus dem Ansatz entfernt. Dieses Vorgehen erlaubt zum einen die parallele Isolierung der genomischen DNA der Zelle, zum anderen entsteht eine ausschließlich aus mRNA bestehende und damit reine Population als Ausgangsmaterial für die Amplifikation. Diese beginnt mit einer cDNA-Synthese, wobei ein Primergemisch aus einem Oligo dT-Primer und einem Zufallsoktamer verwendet wird. Der Zufallsprimer dient dabei der Fragmentierung der entstehenden cDNA auf eine für PCR geeignete Länge. Beide Primer besitzen eine flankierende Region aus 15 Cytidinnukleotiden. Nach der cDNA-Synthese erfolgt eine Tailingreaktion, bei der eine zweite Primerbindungsstelle für die abschließende PCR zur globalen Amplifikation gebildet wird. Durch die Terminal Deoxynukleotid Transferase werden in diesem Schritt 15-30 Nukleotide am 3'-Ende der einzelsträngigen cDNA angehängt. Um eine der C15-Region komplementäre Sequenz zu schaffen, wird dGTP als alleiniges Nukleotid verwendet. Die abschließende globale Amplifikation aller Transkripte in einer PCR kann dann mit einem einzigen Primer, unabhängig von ihrer Sequenz, erfolgen da alle cDNA-Fragmente identische flankierende Regionen besitzen.

Validierung des Amplifikationsprotokolls

Die beschriebene Methode ist zusammen mit einem weiteren publizierten Protokoll den exponentiellen, PCR-basierten Amplifikationsverfahren für cDNA zuzuordnen⁷². Als methodische Alternative zu dieser Art der Amplifikation existieren in der Literatur Protokolle, bei denen die Vervielfältigung der mRNA durch *in vitro*-Transkription und damit linear erfolgt^{70, 71}. Von Kritikern der PCR-basierten Verfahren wird häufig vorgebracht, es handle sich dabei um eine nichtlineare, außerdem sequenzabhängige Reaktion und die Konstanz verschiedener Transkriptverhältnisse zueinander sei daher nicht zu gewährleisten, was zu systematischen Fehlern führe^{71, 161}.

Grundsätzlich ist bei einer PCR zwischen exponentieller und Plateauphase zu unterscheiden. Während der exponentiellen Phase werden bei idealer Effizienz in jedem Zyklus aus einem Template-Molekül zwei Produkte gebildet: $N(c) = N(0) \cdot 2^c$; wobei c die Zykluszahl, $N(0)$ die Anzahl Transkripte vor Beginn der Amplifikation und $N(c)$ die Anzahl Transkripte nach c Zyklen darstellt. Für ein beliebiges Transkriptverhältnis vor der Amplifikation von $N(0)/N(0)'$ ergibt sich folgendes Verhältnis nach c Zyklen: $N(c)/N(c)' = N(0) \cdot 2^c / N(0)' \cdot 2^c = N(0)/N(0)'$. Relative Transkriptverhältnisse bleiben also unverändert für den Fall, dass die PCR sich in der exponentiellen Phase befindet und die Amplifikationseffizienzen für alle Sequenzen gleich sind.

Die reale Effizienz einer PCR ist meist kleiner als die Grundzahl zwei und wird unter anderem beeinflusst von der Wahrscheinlichkeit eines Primingereignisses, der Länge des zu amplifizierenden Produkts und der Salz-, Nukleotid- und Magnesiumkonzentration des Puffers. Da im beschriebenen Verfahren sämtliche Sequenzen über einen identischen Primer in einem Ansatz amplifiziert werden, unterscheiden sich die Primingrate ebenso wie die Reaktionsbedingungen nicht. Die Länge der zu amplifizierenden Sequenzen wird durch die Wahl des Zufallsprimers während der cDNA-Synthese beeinflusst, wobei die Fragmentlänge mit abnehmender Zahl der Zufallsnukleotide sinkt. Letzter Einflussfaktor bleibt also lediglich eine etwaige sequenzabhängige Polymerisation während der PCR. Durch die gewählten stringenten Bedingungen werden jedoch Haarnadelstrukturen und ähnliche Hindernisse für eine gleichmäßige Polymerisation aufgelöst, eine sequenzabhängige Amplifikationseffizienz ist unwahrscheinlich.

Zur Untersuchung einer möglichen Verschiebung der Transkriptverhältnisse während der globalen Amplifikation sowie zur Bestimmung der Variabilität des Verfahrens wurden genspezifische quantitative real time PCR-Reaktionen durchgeführt. Eine Hälfte der analysierten Proben war global amplifiziert, die andere Hälfte nur revers transkribiert worden. Die relativen Expressionswerte zum Referenzgen *Gapdh* waren in diesem Test für alle untersuchten Probentypen (Einzelzelle, 10 Zell Pool und technische Replikate für Einzelzellen) vergleichbar, gleichgültig ob die cDNA zuvor global amplifiziert worden war oder nicht. Bei Hybridisierungen von Proben aus exponentieller und Plateauphase der PCR zur Primäramplifikation auf Oligonukleotidarrays waren die Proben der Plateauphase zudem enger zu Gesamt-RNA-Präparationen korreliert als die exponentiellen Proben. Die Ursache liegt vermutlich in der reduzierten Produktmenge während der exponentiellen Phase, was zu einer höheren Streuung bei der Hybridisierung führt. Ein systematischer Fehler während der PCR-basierten Amplifikation konnte also nicht beobachtet werden, die relativen Transkriptverhältnisse entsprachen der Situation vor der Amplifikation. Diese Daten sind in Übereinstimmung mit publizierten Untersuchungen, die ebenfalls keine systematische Verfälschung durch PCR-Amplifikation feststellten^{137, 162}. Quantitative Analysen der Genexpression sind also auch aus amplifizierten Proben möglich, eine Hybridisierung der Einzelzell-cDNA-Amplifikate auf Oligonukleotidarrays erscheint sinnvoll.

Hybridisierung von Einzelzellamplifikaten auf Oligonukleotidarrays

Auswahl einer geeigneten Plattform

Ein generelles Problem bei der Hybridisierung von Einzelzellamplifikaten ist die starke Verunreinigung der Proben mit bakterieller DNA, die mit den zahlreichen zugegebenen Enzympräparationen eingeschleppt, amplifiziert und markiert wird. Bei der Hybridisierung führt diese Kontamination zu falschpositiven Signalen mit Sonden, die Komplementarität zu bakteriellen Sequenzen besitzen. Oligonukleotidarrays sollten für diese Art der Fehlhybridisierung deutlich weniger anfällig sein als beispielsweise cDNA-Arrays, da hochspezifische kurze Oligonukleotide

speziell für einzigartige Zielsequenzen der jeweiligen Transkripte entworfen werden, wohingegen lange cDNA-Sonden, die zusätzlich oft in Plasmidsystemen präpariert werden, eher die Gefahr der Komplementarität zu verschiedenen Transkripten und bakteriellen Sequenzen bergen. Eine Zielsetzung dieser Dissertation war daher die Etablierung eines Oligonukleotidarrays zur Genexpressionsanalyse von Einzelzellamplifikaten, der die in der Arbeitsgruppe verwendeten cDNA-Nylonfilter ersetzen sollte.

Die ersten kommerziellen Oligonukleotidarrays wurden Mitte der neunziger Jahre von der Firma Affymetrix entwickelt. Das Angebot erweitert(e) sich jedoch im Laufe der Zeit deutlich, so dass mittlerweile von zahlreichen Herstellern Arrays in unterschiedlichen Ausführungen erhältlich sind. Im Rahmen der vorliegenden Dissertation wurden vier Plattformen getestet.

Affymetrix GeneChip-Arrays sind am längsten auf dem Markt. Das Unternehmen besitzt einen großen Erfahrungsschatz sowie zahlreiche unterstützende Geräte und Programme für Probenaufarbeitung und Datenauswertung. Durch den hohen Standardisierungsgrad sind die Ergebnisse zuverlässig, reproduzierbar und validiert^{93, 140}, dafür sind die Arrays sehr teuer und können nur mit der gesamten Technik der Firma Affymetrix benutzt werden. Viele Schritte sind den Anwendern durch Firmengeheimnis verborgen, so dass keine Möglichkeit besteht, den Prozess zu beeinflussen. Die Vorbereitung von Einzelzellpräparationen zur Hybridisierung auf GeneChip-Arrays war zudem kompliziert und somit tendenziell störanfällig, die Hybridisierung selbst nur in Kooperation möglich. Daher wurden weitere Systeme getestet. Die Arrays der Firmen Febit und Memorec zeichneten sich jedoch durch fehlende Chargenkonstanz bzw. Verwendung von cDNA-Sonden aus, was sie für den Routineeinsatz ungeeignet machte.

Hybridisierung auf Operon-Oligonukleotidarrays

Sehr gute Ergebnisse wurden hingegen auf Oligonukleotidarrays der Firma Operon erzielt. Diese verwenden Sonden von 70 Nukleotiden Länge, so genannte Siebzigmere, die in verschiedenen Versionen des Sondendesigns und für zahlreiche Spezies erhältlich sind. Die Oligonukleotide müssen selbst auf einen Objektträger eigener Wahl aufgebracht werden. Daher war es nötig, im Rahmen dieser Arbeit auch technische Fragen zu beantworten. Beispielsweise wurde das Hybridisierungsverhalten stark vom verwendeten Puffer, in dem die Oligonukleotide zum Spotten aufgenommen wurden, ihrer Konzentration, der Oberflächenbeschaffenheit des Objektträgers und der Nachbehandlung zur Fixierung der Sonden auf der Oberfläche beeinflusst.

Obwohl es sich bei den Sonden um speziell für einzelne Transkripte entworfene Siebzigmere handelte, traten bei ersten Experimenten mit Einzelzell-cDNA-Amplifikaten unerwartet Fehlhybridisierungen bakterieller Sequenzen auf. Es gelang jedoch, diese Signale durch ein stringentes Hybridisierungsprotokoll zu eliminieren. Wichtigster Aspekt dabei waren die Waschschritte im Anschluss an die eigentliche Hybridisierung, die im Vergleich zu Standardprotokollen für Gesamt-RNA-Präparationen zum Teil sehr stringent gewählt werden mussten. Für Arrays verschiedener

Hersteller und Oligonukleotidversionen waren jedoch unterschiedliche Protokolle optimal. Worin die Gründe für diese differierenden Hybridisierungseigenschaften liegen, konnte nicht endgültig geklärt werden. Viel spricht jedoch für verändertes Sondendesign der Firma Operon zwischen Spezies und Oligonukleotidversionen, das zu unterschiedlichen Hybridisierungseigenschaften führt.

Mit den optimierten Protokollen wurden zwischen Hybridisierung und willkürlichen genspezifischen PCR-Reaktionen 70-75% übereinstimmende Ergebnisse erzielt. Ein Drittel der abweichenden Ergebnisse konnte jeweils durch das Sondendesign und die speziellen Eigenschaften des jeweiligen Verfahrens erklärt werden. Für PCR-Reaktionen auf Gene, die in einem Hybridisierungsexperiment biologische Relevanz gezeigt hatten, betrug die Übereinstimmung 100%, die gemessene Genregulation erreichte mit beiden Verfahren ähnliche Werte. Die Daten waren vergleichbar oder besser als ähnliche Untersuchung mit Gesamt-RNA^{87, 139, 140}. Genexpressionsanalysen an Einzelzellen sind unter Verwendung der beschriebenen Protokolle also mit ähnlicher Qualität möglich wie Gesamt-RNA-Untersuchungen nach Standardverfahren.

Validierung des Hybridisierungsprotokolls

Das Amplifikations- und Hybridisierungsprotokoll wurde anschließend an einer biologischen Fragestellung validiert. Dazu wurden Vorläuferzellen aus murinem Knochenmark in Kultur zu immaturren dendritischen Zellen (DCs) differenziert, die durch Stimulation mit LPS zu maturaen DCs ausgereift werden können¹³⁸. Die LPS-Effekte werden dabei durch den Tlr4-Signalweg vermittelt^{141, 142}, dessen Aktivierung auch auf Einzelzellebene in der Genexpression verfolgt werden konnte, nachdem einzelne immature und mature DCs nach dem beschriebenen Protokoll hybridisiert worden waren. Es wurden Gene innerhalb der Kaskade sowie inflammatorische Zytokine und kostimulatorische Moleküle an deren Ende hochreguliert. Die Daten decken sich mit publizierten Arbeiten über die LPS-induzierte Aktivierung des MyD88-abhängigen Signalwegs und die Hochregulation bestimmter Chemokine wie *IL6* und *Rantes*^{143, 144, 163}. Zum Teil ist ein Vergleich mit Literaturdaten jedoch schwierig, da die Ausreifung stimulus- und zeitabhängig verläuft^{138, 164}.

Einzelzellen aus beiden Aktivierungszuständen konnten durch ein hierarchisches Clusterverfahren anhand ihrer Genexpressionswerte in die beiden Gruppen der immaturren und maturaen DCs aufgeteilt werden. Es handelt sich dabei um ein gängiges iteratives Verfahren zur Gruppierung von Proben nach ihrer Ähnlichkeit. Der Algorithmus berechnet zunächst Distanzen zwischen sämtlichen Messwerten aller Proben und fasst die beiden ähnlichsten Proben zu einer Gruppe zusammen. In den folgenden Schritten werden Distanzen zwischen den verbleibenden Proben sowie den bisher entstandenen Gruppen berechnet, bis alle Proben eingeordnet wurden. Die graphische Darstellung des Ergebnisses erfolgt in Form eines Dendrogramms, wie es von entwicklungsgeschichtlichen Stammbäumen bekannt ist¹⁰⁶. Das Verfahren kann zur Klasseneinteilung und -vorhersage benutzt werden und hat sich seit seiner ersten Verwendung in den neunziger Jahren als gängiges Instrument zur Auswertung komplexer Datensätze etabliert¹⁰⁷. Mit einem hierarchischen Clustering gelang es auch, zusätzlich zu den DCs

hybridisierte Proben verschiedenen Typs sowie Einzelzellen unterschiedlichen histogenetischen Ursprungs, darunter epitheliale Zellen und hämatopoietische Progenitoren, korrekt zu ordnen.

Die Hybridisierungsergebnisse konnten also nicht nur mit PCR als einer unabhängigen Methode validiert werden, sondern es war auch möglich, biologische Fragen zu beantworten und geringe Unterschiede in der Genexpression, wie sie zwischen immaturren und maturaen DCs zu erwarten sind, abzubilden. Dies ist umso bemerkenswerter, als die Expressionsprofile einzelner Zellen stark von Faktoren geprägt werden, die nichts mit der Fragestellung des jeweiligen Experiments zu tun haben. Dazu gehören die stochastische Natur der Genexpression ebenso wie unterschiedliche zelluläre Funktionen und (nicht identifizierte) Differenzierungsstadien¹⁶⁵. Trotz dieser Einflussfaktoren war die Methode geeignet, in Einzelzellen mit hoher Sensitivität und Reproduzierbarkeit valide Daten zu erheben. Sie sollte daher auf disseminierte Tumorzellen aus dem Knochenmark von Brustkrebspatientinnen angewendet werden, um Erkenntnisse über die Biologie dieser Zellen zu gewinnen, die für die Prognose der Patienten relevant sind, außerdem die eigentlichen Zielzellen adjuvanter Therapien darstellen, über die aber wegen ihrer extremen Seltenheit bisher nur wenige Daten existieren. Zur Detektion disseminierter Tumorzellen wurde eine Färbung gegen das epitheliale Adhäsionsmolekül EpCAM verwendet.

Detektion EpCAM-positiver Zellen im Knochenmark von Brustkrebspatientinnen

EpCAM ist für die Untersuchung disseminierter Tumorzellen aus verschiedenen Gründen von Bedeutung. Zum einen geht die Tumorigenese der meisten Karzinome mit einer verstärkten oder *de novo*-Expression dieses Proteins einher, die mit zunehmender Proliferation, abnehmender Differenzierung und einer schlechten Prognose korreliert ist⁴⁶. Zum anderen handelt es sich um ein Molekül, das im adulten Organismus fast ausschließlich von einfachen, vor allem glandulären, sowie pseudo-stratifizierten und transitorischen Epithelien exprimiert wird. Plattenepithel und nichtepitheliale Gewebe sind normalerweise EpCAM-negativ⁴⁴. EpCAM eignet sich daher zur Detektion von Zellen epithelialen Ursprungs in mesenchymalen Geweben^{20, 40}.

Besondere Bedeutung für diese Arbeit besitzt EpCAM, da es sich im Gegensatz zu anderen klassischen Markern disseminierter Tumorzellen wie epithelialen Zytokeratinen um ein Oberflächenmolekül handelt. Eine Färbung ist daher auch an nichtpermeabilisierten lebenden Zellen möglich, was eine Voraussetzung für Genexpressionsanalysen ist. In Kombination mit dem in dieser Arbeit entwickelten Anreicherungsverfahren für epitheliale Zellen aus Knochenmarkaspiraten, bestehend aus einer Dichtegradientenzentrifugation mit 70% Percoll gefolgt von einer Depletion hämatopoietischer Zellen, wurden in 88,4% aller Punktate EpCAM-positive Zellen gefunden. Dies stellt einen deutlichen Fortschritt im Vergleich zu den bisher benutzten Verfahren dar, mit denen nur in 10-30% der Aspiarte disseminierte Zellen gefunden werden^{29, 33}. Von den 142 isolierten Zellen gelang eine Amplifikation des Transkriptoms in über 70% der Fälle, aus etwa der Hälfte der Zellen konnte das Genom parallel zur mRNA amplifiziert werden.

Untersuchung EpCAM-positiver Zellen auf genomische Aberrationen mittels CGH

45 EpCAM-positive Zellen wurden mittels Komparativer Genomischer Hybridisierung (CGH) auf chromosomale Veränderungen untersucht. 30 Zellen zeigten aberrante CGH-Profile und konnten daher eindeutig als Tumorzellen klassifiziert werden. Bei den 15 Zellen mit balanciertem CGH-Profil kann es sich um seltene hämatopoietische Zellen mit EpCAM-Expression handeln. Eine solche Expression wurde beispielsweise für immature erythroide Progenitorzellen beschrieben¹⁶⁶. Ein balanciertes CGH-Profil ist aber auch bei Tumorzellen aus verschiedenen Gründen nicht ausgeschlossen. Zum einen ist die CGH nicht geeignet, chromosomale Aberrationen zu detektieren, die ohne Veränderung der Kopienzahl einhergehen, wie zum Beispiel balancierte Translokationen, oder solche unterhalb der Auflösungsgrenze des Verfahrens von 10-20 Mb^{125, 127}. So konnte durch LOH-Analysen gezeigt werden, dass auch Zellen ohne CGH-Aberrationen Tumorzellen sein können⁶⁸. Im Rahmen der Tumorprogression nehmen genomischen Veränderungen vom normalen diploiden Stadium bis hin zur genomischen Krise mit intrachromosomalen Brüchen zu. CGH-normale Zellen mit subchromosomalen Aberrationen befinden sich daher früh in dieser Entwicklung und sind für die Suche nach tumorinitiierenden Veränderungen besonders interessant^{67, 167}. Neben mit der CGH nicht detektierbaren genomischen Aberrationen kommen nichtgenomische, epigenetische Veränderungen als Charakteristikum von Tumorzellen und Ursache gestörter Genexpression vor allem in frühen Stadien der Tumorigenese in Betracht¹²⁹. So gehört beispielsweise die Hypermethylierung von *p16^{INK4a}* zu einem typischen Ereignis in präkanzerösen Läsionen und frühen Tumorstadien¹⁶⁸, aber auch andere Gene, darunter *HOXA5*, *p21^{WAF}* und E-Cadherin, sind von epigenetischen Veränderungen beim Mammakarzinom betroffen¹²⁸. Bei den Zellen ohne CGH-Veränderungen könnte es sich also um Tumorzellen handeln, die solche Defekte aufweisen.

Häufige CGH-Aberrationen in EpCAM-positiven Zellen

In den 30 Zellen mit CGH-Aberrationen waren Amplifikationen häufiger zu beobachten als Deletionen. Besonders oft waren Chromosom 1p, 4q, 5p, 6q, 12q, 15q, 17, 18, 19, 22 und X amplifiziert, Verluste traten vor allem auf Chromosom 1p, 19 und 20 auf. Besonders die Veränderungen an Chromosom 1p 19 müssen jedoch mit Vorsicht betrachtet werden. Auf Grund der CG-reichen Sequenz ist eine Amplifikation der DNA mit dem verwendeten Protokoll erschwert, da zur Fragmentierung der DNA *MseI* mit der Schnittstelle TTAA verwendet wird und CG-reiche Sequenzbereiche dadurch unterrepräsentiert sein können. Nichtsdestotrotz zeigten alle 15 Zellen mit balanciertem CGH-Profil Hybridisierungen auch auf Chromosom 19. Außerdem traten in den Zellen mit chromosomalen Veränderungen in diesem Bereich sowohl Amplifikationen als auch Deletionen auf, was ein generelles Versagen des Verfahrens bei Einzelzellen im Vergleich zum Referenzpool, der einen Mittelwert vieler Zellen darstellt, unwahrscheinlich macht.

Amplifikation der Region 17q21

Besonders häufig war eine Vermehrung chromosomalen Materials im Bereich des *HER2*-Gens auf Chromosom 17q21. Von den acht betroffenen Zellen zeigten drei einen Gewinn des gesamten Chromosoms, drei eine Amplifikation nur im Bereich 17q21 und zwei Zellen eine Mischform. Eine Amplifikation des *HER2*-Gens geht oft mit einer Überexpression des Proteins einher und ist assoziiert mit einem aggressiveren Krankheitsverlauf und meist fortgeschrittenem Erkrankungsstadium⁹. Durch eine Untersuchung mittels quantitativer PCR zur Bestimmung der Gendosis von *HER2* in Zytokeratin-positiven Zellen aus dem Knochenmark von Brustkrebspatientinnen konnte jedoch gezeigt werden, dass eine Amplifikation bereits bei etwa 17% der Patientinnen im Stadium M0 vorliegt und kein Zusammenhang zwischen der Proteinexpression im Primärtumor und der Genamplifikation in den disseminierten Zellen besteht⁶⁸. Die hier gezeigten Daten finden eine Amplifikation der Region 17q21 bei 30% (3/10) der Patientinnen im Stadium M0, was aber wegen der geringen Probenzahl und der möglichen Beteiligung anderer Gene der Region als *HER2* weiter untersucht werden muss. Sollte es sich bei der Amplifikation jedoch tatsächlich um *HER2* handeln, könnte das zukünftig Konsequenzen für die Therapie der Patientinnen haben, da mit Trastuzumab ein effektives Therapeutikum bei *HER2*-überexprimierenden Tumoren existiert. Der Antikörper besitzt bisher keine Zulassung für das nicht-metastasierte Erkrankungsstadium. Er könnte aber gerade bei Patientinnen ohne solide Metastasen, allerdings mit *HER2*-positiven Zellen im Knochenmark, von großem Nutzen zur Eradikation der Mikrometastasen sein. Eine Studie zur Verwendung von Trastuzumab in frühen Erkrankungsstadien zeigte Ende 2005 schon bei einer Nachbeobachtungszeit von einem Jahr ein signifikant verlängertes erkrankungsfreies Überleben für die Patientinnen der Behandlungsgruppe¹⁶⁹. Eine baldige Zulassungserweiterung scheint daher möglich.

Amplifikation des X-Chromosoms

Die ebenfalls häufige Amplifikation des X-Chromosoms ist in sofern interessant, als dieses Chromosom zwar im Vergleich mit dem Rest des Genoms genarm aber reich an Genen der Cancer-Testis-Antigenengruppe (CT-Antigene) ist. Zu diesen Genen, die ausschließlich im Hoden und in Tumoren exprimiert werden, gehört beispielsweise die *MAGE*-Familie. Nach der kompletten Sequenzierung des X-Chromosoms wurden 99 Gene und damit 10% der Genzahl des Chromosoms als CT-Antigene vorhergesagt. Im gesamten übrigen Genom werden 84 CT-Antigene Gene gelistet. Von den 35 Vertretern der *MAGE*-Familie liegen 32 auf dem X-Chromosom¹⁷⁰. Eine Amplifikation des Chromosoms könnte daher Einfluss auf die Expression von CT-Antigenen nehmen. Die biologische Funktion und die Rolle bei der Tumorentstehung ist für die meisten dieser Gene nach wie vor unbekannt, zum Teil bestehen Assoziationen mit ungünstigen klinischen Parametern¹⁷¹. Zwischen der Expression der *MAGE*-Gene des Nylonarrays und der Amplifikation des X-Chromosoms gab es allerdings keinen Zusammenhang. Von den 21 *MAGE*-exprimierenden Zellen wiesen zwei eine X-

Amplifikation auf, fünf ein balanciertes CGH-Profil, und von acht Zellen existierte keine CGH. Eine Aktivierung der *MAGE*-Expression in Tumorzellen wurde allerdings auch durch Promotorhypomethylierung beschrieben¹⁷², was mit der CGH nicht detektierbar ist. Zudem wurden keine Expressionsdaten weiterer CT-Antigene des X-Chromosoms erhoben, die auf dem Nylon-Array nicht repräsentiert sind. Auf den Oligonukleotidarrays exprimierten alle fünf untersuchten Zellen mindestens einen Vertreter der *MAGEA*-Familie, zwei der Zellen wiesen eine X-Amplifikation auf.

Vergleich der CGH-Aberrationen mit publizierten Daten

Generell unterschied sich das Aberrationsmuster der untersuchten Zellen von publizierten Daten über Zytokeratin-positive disseminierte Tumorzellen im Knochenmark von Brustkrebspatientinnen sowohl im Stadium M₀ als auch im Stadium M₁⁶⁷. Dies liegt vermutlich darin begründet, dass sich die untersuchten Zellpopulationen unterscheiden. Zum einen exprimieren nicht alle EpCAM-positive Zellen auch Zytokeratine und umgekehrt⁵³, zum anderen wurden die Zellen dieses Kollektivs durch Dichtegradientenzentrifugation mit 70%iger Percoll-Lösung gefolgt von der Depletion hämatopoetischer Zellen angereichert, wohingegen in der Literatur ein konventionelles Anreicherungsverfahren über Ficoll benutzt wurde. Der Begriff „disseminierte Tumorzelle“ beschreibt also offensichtlich eine extrem heterogene Zellpopulation. Ihre genauere Charakterisierung auch über das Genom hinaus scheint daher dringend erforderlich.

Genexpressionsanalysen EpCAM-positiver Zellen aus dem Knochenmark von Brustkrebspatientinnen

Genexpressionsstudien und Progression des Mammakarzinoms

Das gängige Modell zur Progression maligner Erkrankungen orientiert sich an der Adenom-Karzinom-Sequenz des Kolonkarzinoms¹⁷³. Eine analoge lineare Tumorprogression von der normalen Brustdrüse über die atypische duktaile Hyperplasie, das duktaile Carcinoma in situ zum invasiven Karzinom und später zur Metastase wurde auch für das Mammakarzinom vorgeschlagen¹⁷⁴. Die Entwicklung eines Karzinoms beruht dabei auf einer für den Tumor günstigen Ansammlung von Mutationen in Onkogenen und Tumorsuppressorgenen über lange Zeiträume¹⁷⁵. Die Metastasierung stellt den finalen Schritt dar, der erst durch weitere ausschlaggebende Mutationen möglich wird. Die Wahrscheinlichkeit des Eintretens einer solchen Mutation steigt mit zunehmender Zellzahl des Primärtumors, die Mutation selbst trifft eine zufällige Zelle.

In den letzten Jahren haben gerade Genexpressionsstudien zu einer Gegenhypothese geführt. Es zeigte sich, dass Expressionssignaturen von Primärtumor und Metastase desselben Patienten eine hohe Ähnlichkeit besitzen und zueinander ähnlicher sind als Primärtumoren oder Metastasen verschiedener Patienten untereinander¹⁷⁶⁻¹⁷⁸. Außerdem ließ sich eine spätere Metastasierung aus der Genexpression des Primärtumors vorhersagen¹³. Daraus wurde gefolgert, dass auch in frühen Tumorstadien schon die spätere Metastasierung im gesamten Tumor angelegt und nicht einzelnen Zellen oder Subgruppen

eines großen Tumors nach entsprechenden Mutationen zuzuschreiben ist. Wieder andere Studien beschrieben jedoch Genexpressionsprofile, die sich im Laufe der Tumorprogression veränderten¹⁷⁹. Während der malignen Transformation waren Gene des Gleichgewichts zwischen Proliferation und Zelltod betroffen, in späteren Stadien Invasions- und Mobilitätsgene sowie solche, die an Umbauvorgängen der extrazellulären Matrix beteiligt sind¹⁸⁰.

Häufiges Problem solcher Studien, das auch die abweichenden Ergebnisse erklären könnte, ist die Probengewinnung. Manche Ansätze untersuchen Areale eines Tumors, denen unterschiedliche histologische Stufen zugeordnet werden¹⁷⁸, andere ganze Tumore von Patienten in unterschiedlichen Krankheitsstadien. Beide Ansätze bieten Vor- und Nachteile. Histologische Einteilungen sind beispielsweise stark vom untersuchten Tumoreal, dem jeweiligen Gewebsschnitt und dem beurteilenden Pathologen abhängig. Beim Vergleich verschiedener Patientinnen hingegen können differierende Genexpressionsprofile aus Unterschieden in Einflussfaktoren wie dem Estrogenrezeptorstatus, die das Genexpressionsprofil stark prägen, resultieren, so dass eine Zuordnung zu klinischen Parametern, Tumorstadien oder Prognosen überlagert oder unmöglich wird^{180, 181}. Zudem stellen alle diese Untersuchungen immer die Analyse eines mittleren Genexpressionsprofils zahlreicher Zellen und Zelltypen, darunter auch Immun- und Stromazellen, dar, was die Genexpression relevanter Zellen, die nur einen geringen Anteil des gesamten Tumormaterials ausmachen, maskieren kann.

Bedeutung der Untersuchung einzelner disseminierter Tumorzellen

Die Untersuchung einzelner disseminierter Tumorzellen bietet in diesem Zusammenhang die Möglichkeit zur Gewinnung völlig neuer Erkenntnisse. Zum einen nehmen disseminierte Tumorzellen eine Mittelstellung zwischen Primärtumor und manifesten Metastasen ein. Ihre Untersuchung gibt also möglicherweise Aufschlüsse über den Zeitpunkt und die auslösenden Faktoren am Beginn der Metastasierung. Durch einen Vergleich mit Daten aus Primärtumorstudien kann so ein Beitrag zum Verständnis der Tumorprogression und zur Entwicklung neuer adjuvanter Therapien geleistet werden. Zum anderen stellt eine einzelne Zelle eine maximal reine Population dar. Bei entsprechender Güte des Verfahrens wird die gemessene Genexpression daher nicht durch Einflüsse wie unterschiedliche Zellpopulationen oder -funktionen verfälscht. Die fünf EpCAM-positiven Zellen aus dem Knochenmark von Brustkrebspatientinnen, die auf humane Operonarrays hybridisiert wurden, befanden sich beispielsweise in verschiedenen Stadien des Zellzyklus, mit entsprechendem Einfluss auf die Genexpressionsprofile (Abbildung 5.1). Während Zelle 4 und 5 vor allem Gene der Mitosephase exprimieren, zeigt Zelle 3 eine Expression von Cyclin D1 und anderen Genen der G1/S-Phase. Zelle 2 exprimiert kaum Zellzyklusgene und befindet sich wahrscheinlich in der G0-Phase. Alle Zellen exprimieren neben Zellzykluspromotoren aber auch zahlreiche Gene, die bei Kontrollpunkten des Zellzyklus, beispielsweise auf DNA-Schäden, eine Rolle spielen.

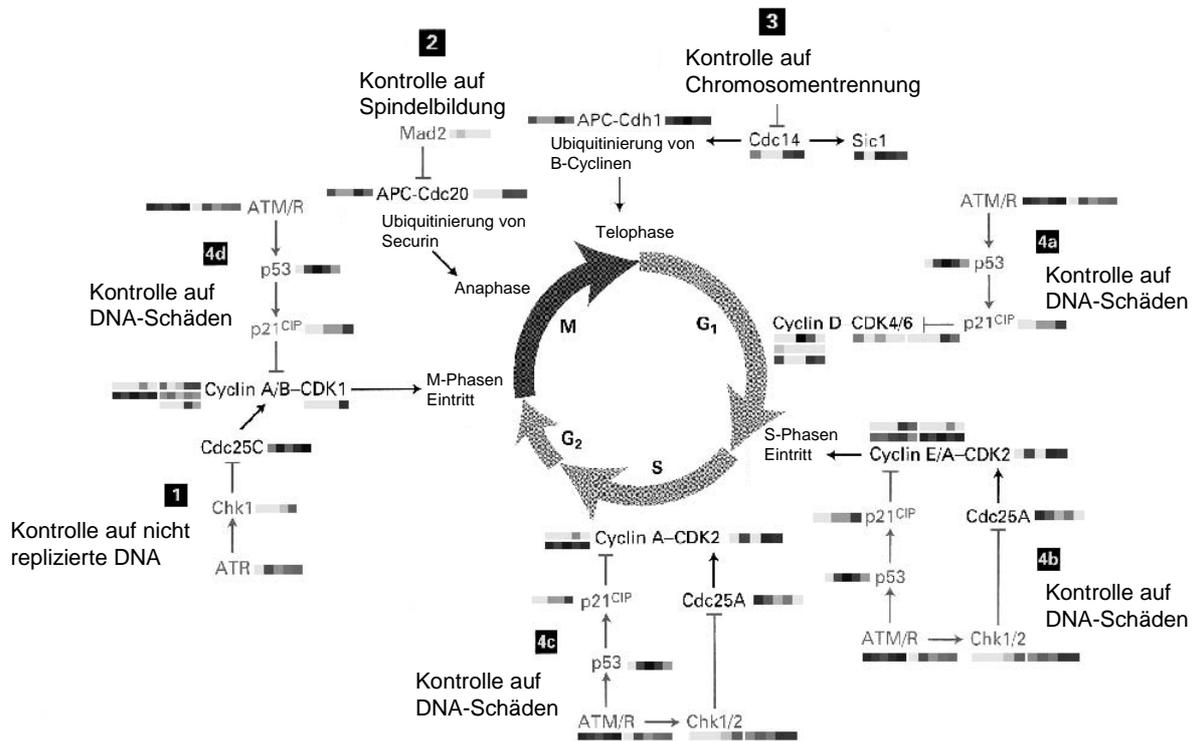


Abbildung 5.1 Expression von Zellzyklusgenen in fünf EpCAM-positiven Zellen aus dem Knochenmark von Patientinnen mit Mammakarzinom.

Die Abbildung des Zellzyklus wurde dem Buch „Molecular Cell Biology“¹⁸² entnommen und mit den Expressionswerten der aufgeführten Gene versehen. Die Expression in jeder der fünf untersuchten Zellen ist von links nach rechts durch ein Quadrat in der Leiste neben dem jeweiligen Gen dargestellt. Je dunkler dabei die Färbung, desto höher die Expression im Vergleich mit der Referenzprobe.

Vergleich der Genexpression EpCAM-positiver Zellen mit publizierten Daten

Die Genexpressionsprofile EpCAM-positiver Zellen auf Oligonukleotidarrays aus dieser Arbeit wurden mit publizierten Daten zum Mammakarzinom verglichen. Zum einen handelte es sich dabei um Untersuchungen, welche die Genexpression während der Entwicklung des Primärtumors von präinvasiven Stadien zum Carcinoma *in situ* verfolgten^{178, 180}. Zum anderen wurden Arbeiten gewählt, die einen Vergleich von Primärtumor und Metastase bzw. die Vorhersage einer späteren Metastasierung zum Inhalt hatten^{12, 13, 183}. Mit den Vergleichen sollte erstens geprüft werden, ob in disseminierten Zellen ebenfalls Gene exprimiert sind, die in anderen Studien zum Mammakarzinom beschrieben wurden, die Daten also in den Kontext der Literatur zu dieser Erkrankung passen. Zweitens sollte untersucht werden, ob die Zellen einem bestimmten Tumorstadium zuzuordnen sind oder prognostische Information enthalten.

Insgesamt war die Übereinstimmung zu den publizierten Daten gering, zwischen den ausgewählten Arbeiten war sie allerdings ebenfalls minimal. Neben Unterschieden in der Fragestellung, den eingeschlossenen Patientinnen und dem verwendeten Verfahren zur Bestimmung der Genexpression spielt dabei auch das verwendete Probenmaterial eine bedeutende Rolle, so dass eigentlich kein anderes Ergebnis zu erwarten war¹¹⁵ (vgl. auch Absatz 1.6.5). Ein Viertel der relevanten Gene einer

Studie waren beispielsweise Komponenten der extrazellulären Matrix, die von Bindegewebszellen synthetisiert werden. Dies reflektiert das Problem der Untersuchung des gesamten Tumorgewebes an Stelle der eigentlichen Tumorzellen.

Die EpCAM-positiven Zellen waren in diesen Vergleichen weder eindeutig als (nicht)metastatisch zu klassifizieren noch ließ sich die Zelle, die aus dem Knochenmark der Patientin isoliert wurde, bei der dieses Ereignis bereits eingetreten war der Gruppe zuordnen, bei der eine Metastasierung innerhalb der nächsten fünf Jahre vorausgesagt wurde. Nichtsdestotrotz waren auch bei der geringen Probenzahl in den in dieser Arbeit erhobenen Daten Zusammenhänge zwischen dem Tumorstadium oder dem Tumortyp und der Genexpression der EpCAM-positiven Zellen zu beobachten. Die Analyse weiterer disseminierter Tumorzellen auf Oligonukleotidarrays sowie die weitere Verfolgung des Krankheitsverlaufs der Patientinnen wird sicher interessante Ergebnisse liefern, vor allem da die disseminierten Zellen eine Mittelstellung zwischen lokaler Erkrankung und manifester Metastasierung einzunehmen scheinen.

Analyse der Genexpression EpCAM-positiver Zellen auf Nylonarrays

Die Analyse der Genexpression aller EpCAM-positiven Zellen auf Nylonarrays führte ebenfalls zu einigen interessanten Ergebnissen. Zum einen konnte im Vergleich mit den CGH-Untersuchungen gezeigt werden, dass eine strenge gewebstypische Genexpression bei disseminierten Tumorzellen offensichtlich nicht stattfindet. Zellen mit CGH-Aberrationen, also eindeutige Tumorzellen, exprimierten auch typische hämatopoietische Gene wie *CD33*, *CD16* oder *CD83*. Zum anderen wurde deutlich, dass es sich bei der großen Mehrzahl der isolierten Zellen offensichtlich nicht um terminal differenzierte epitheliale Zellen der Mamma handelt, da sowohl die Expression epithelialer wie auch brustspezifischer Marker nur in einem geringen Teil der Zellen zu beobachten war. Welchem Zelltyp oder Differenzierungsstadium die große Mehrheit der Zellen zuzuordnen ist, ließ sich anhand des beschränkten Gensatzes des Nylonarrays jedoch nicht feststellen. Durch chromosomale Aberrationen in der CGH oder die Expression von epithelialen Markern und Genen der Cancer-Testis-Gruppe waren viele dieser Zellen aber sicher als Tumorzellen zu klassifizieren.

Die Tumorstammzelltheorie im Zusammenhang mit EpCAM-positiven Zellen

In den letzten Jahren wird vermehrt die so genannte Tumorstammzelltheorie diskutiert¹⁸⁴. Dabei wird davon ausgegangen, dass eine begrenzte Anzahl maligne transformierter, tumorinitiierender Zellen mit Stammzeleigenschaften wie Immortalität und geringer Teilungsrates bei unbegrenzter Teilungsfähigkeit existiert. Diese Tumorstammzellen bringen eine große Anzahl Progenitor- bzw. differenzierter Zellen hervor, die den Primärtumor bilden. Im Vergleich dazu gehen Tumoren nach dem konventionellen Modell aus maligne transformierten differenzierten Zellen hervor, die alle tumorinitiierend sein können (Abbildung 5.2).

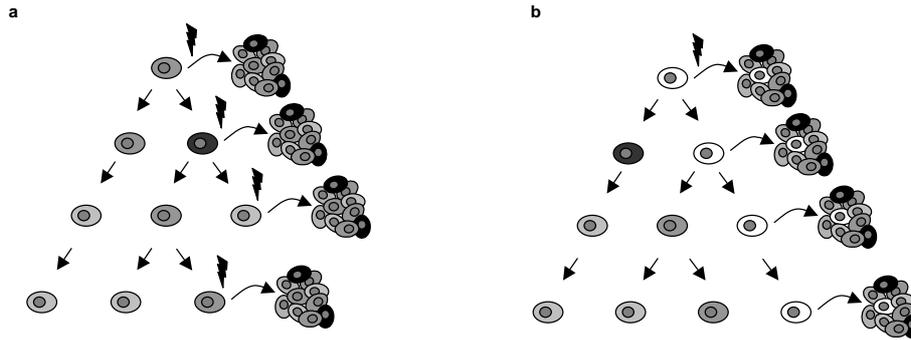


Abbildung 5.2 Vergleich der konventionellen Theorie zur Tumorentstehung (a) mit der Tumorstammzelltheorie (b)

a Nach der klassischen Theorie bestehen Tumore aus heterogenen Zellen, von denen jedoch prinzipiell alle tumorinitiierend sind. **b** Die Tumorstammzelltheorie unterscheidet die tumorinitiierenden Tumorstammzellen von der Masse des eigentlichen Primärtumors oder der Metastase, die aus differenzierteren heterogenen Tochterzellen hervorgehen, die nicht tumorigen sind. Die Tumorstammzellen sind die eigentlichen Zielzellen adjuvanter Therapien, da sie auch bei Eradikation der Tumormasse in der Lage sind, einen neuen Tumor zu bilden. Die Abbildung wurde der Publikation von Reya, Morrison et al 2001¹⁸⁴ entnommen und modifiziert.

Bei hämatopoietischen Tumoren konnte bereits gezeigt werden, dass beispielsweise die akute myeloische Leukämie durch Mutationen in Progenitorzellen entsteht¹⁸⁵⁻¹⁸⁷. Die Tumorstammzellhypothese war Anfang der 70er Jahre für diese Neoplasien formuliert worden¹⁸⁸, seit Ende der 70er Jahre wird ihre Übertragung auf solide Tumore diskutiert¹⁸⁹. Der Beweis steht nach wie vor aus, in letzter Zeit gelang jedoch die nähere Charakterisierung der Population tumoriger Zellen aus Mammakarzinomen¹⁵⁶. Ein Problem ist, dass Marker zur Identifizierung und Charakterisierung der Tumorstammzellen noch fehlen. Dennoch sind Tumorstammzellen, so sie existieren, die eigentlichen Zielzellen adjuvanter Therapien. Möglicherweise sind sie jedoch wegen stammzelltypische Eigenschaften wie geringer Teilungsrate, fehlender Markerexpression oder Expression bestimmter Transporter, die potentiell toxische Substanzen aus der Zelle ausschleusen, schwer zu eliminieren. Zu diesen Transportern gehört ABCG2, das unter anderem den Hoechst 33342-Farbstoff transportiert^{190, 191}. Eine Färbung mit diesem Farbstoff ist daher geeignet, Zellen mit Stammzeleigenschaften nachzuweisen. Da das Ausschleusen des Farbstoffs aus der Zelle einen aktiven biologischen Prozess darstellt, muss die Färbung unter physiologischen Bedingungen durchgeführt werden und erhöhte in dieser Arbeit den Anteil apoptotischer Zellen, aus denen keine mRNA mehr isoliert werden konnte. Aus diesem Grund und da der ABCG2-Transporter neben Stammzellen auch auf multi Drug-resistenten Tumorzellen exprimiert wird, wurde auf die Färbung wieder verzichtet.

Expression potentieller Stammzellmarker in EpCAM-positiven Zellen

Im Zusammenhang mit disseminierten Tumorzellen ist die Tumorstammzelltheorie von besonderem Interesse, da disseminierte Tumorzellen grundsätzlich in der Lage sind, manifeste Metastasen hervorzubringen, also nahezu unbegrenzt zu proliferieren. Außerdem schaffen sie sich analog der Stammzellnische in ektopen Organen günstige Überlebensbedingungen. Besonders das Knochenmark

ist dabei ein Organ, das das Wachstum und Überleben zahlreicher Stamm- und Progenitorzellen unterstützt. Eine nicht unbedeutende Rolle kommt in diesem Zusammenhang dem Stroma zu. Ein Signalweg, der bei der Interaktion epithelialer Zellen mit dem Stroma auch während der Brustentwicklung eine Rolle spielt, ist die Hedgehog-Kaskade. Mäuse, die haploinsuffizient für den Hedgehog-Rezeptor *Ptch1* sind, zeigen beispielsweise eine Hyper- und Dysplasie der Brust¹⁹². Interessant war in diesem Zusammenhang die Expression der Liganden Sonic hedgehog (*SHH*), Desert Hedgehog (*DHH*) und vor allem Indian hedgehog (*IHH*) sowie verschiedener nachgeschalteter Gene des Signalweges in den EpCAM-positiven Zellen. Der Signalweg ist physiologischerweise nur während der Embryonalentwicklung und in Stammzellen adulter Gewebe aktiv. Eine unphysiologische Aktivierung wurde in einigen Tumoren gefunden, die von Stammzellen auszugehen schienen. Neben der Rolle des Signalweges bei der Migration und Differenzierung während der Organogenese wird auch eine Beteiligung bei der Zellzykluskontrolle und Metastasierung diskutiert. Es existieren Repressoren für verschiedene Ebenen des Hedgehog-Signalwegs, darunter monoklonale Antikörper und Inhibitoren wie Cyclopamin, ein Alkaloid aus *Veratum californicum*, die als experimentelle Therapeutika verwendet werden. Da der Signalweg nur in einem geringen Prozentsatz adulter Zellen aktiv ist, haben diese Moleküle kaum bis keine toxischen Nebenwirkungen¹⁹³. Die Koregulation von *SHH* und *IHH* in drei EpCAM-positiven Zellen war nicht auf Sequenzhomologien und fehlende Differenzierung seitens der Arraysonde, sondern auf echte Expression der beiden Transkripte zurückzuführen. Indian hedgehog war in unterschiedlichem Ausmaß in allen fünf Zellen erhöht exprimiert. Es gilt als wahrscheinlicher Ligand des Signalwegs in der Brustdrüse¹⁹². Die Expression von Genen der Hedgehog-Kaskade in disseminierten Tumorzellen könnte daher auf eine mögliche Aktivierung stammzellspezifischer Signalwege in diesen Zellen hindeuten.

Der Beweis der Existenz einer Bruststammzelle ist ebenso wie ihre Charakterisierung noch Gegenstand der Forschung¹⁵⁹. Verschiedene Gruppen konnten jedoch Brustzellen mit Progenitoreigenschaften durch Kombinationen verschiedener Oberflächenmarker identifizieren. EpCAM-positivere Zellen zeigten verschieden starke Übereinstimmungen in ihrer Genexpression mit drei dieser Arbeiten sowie einer Reihe weiterer möglicher Marker von Bruststammzellen, die in der humanen und murinen Literatur häufiger genannt werden¹⁵⁶⁻¹⁶⁰. Der Detektionsmarker EpCAM selbst spielt ebenfalls eine bedeutende Rolle bei der Entwicklung und Differenzierung während der Embryogenese und wurde vor allem in noch nicht komplett ausdifferenzierten Zellen nachgewiesen^{194, 195}. Diese Beobachtungen legen ebenso wie die Aktivität des Hedgehog-Signalweges die Hypothese nahe, dass zumindest ein Teil der disseminierten Zellen Stammzeleigenschaften besitzen könnte. Von den fünf untersuchten Zellen trifft dies möglicherweise für Zelle 050125-03, 050727-08 und 050728-03, also für drei von fünf Patientinnen, nach mehreren Kriterien zu (Tabelle 5.1). Da sie zudem chromosomale Veränderungen in der CGH zeigen, handelt es auch eindeutig um Tumorzellen.

Tabelle 5.1 EpCAM-positiven Zellen mit möglichen Stammzeleigenschaften

Kriterium		050125-03	050713-07	050713-06	050727-08	050728-03
Hedgehog-Signalweg (Tabelle 4.5)	Expression von Ligand und Effektor	<i>IHH, DHH, GLI4</i>	<i>SHH, IHH</i>	<i>SHH, IHH</i>	<i>IHH, GLI3, GLI4</i>	<i>SHH, IHH</i>
Stammzeleigenschaften entsprechend den Kriterien von: (Tabelle 4.6)	Al Hajj ¹⁵⁶	1/2	1/2	2/2	1/2	2/2
	Shackleton ¹⁵⁷	0/3	0/3	1/3	1/3	2/3
	Stingl ¹⁵⁸	3/4	1/4	1/4	2/4	2/4
	Sonstige ^{159, 160}	6/12	3/12	3/12	6/12	6/12
CGH-Profil		aberrant	balanciert	aberrant	aberrant	aberrant

Sollte es sich also tatsächlich zumindest bei einem Teil der EpCAM-positiven Zellen um disseminierte Tumorzellen mit Stammzeleigenschaften handeln, hätte das auch Konsequenzen für die Therapie. Sämtliche adjuvanten Therapeutika werden zurzeit am Primärtumor bzw. an aus diesem generierten Zelllinien getestet. Ebenso werden Therapieentscheidungen für einzelne Patienten anhand des Status des Primärtumors getroffen, also an der Masse der Tumorzellen. Sollten sich darunter einige wenige Tumorstammzellen befinden, werden diese nicht berücksichtigt. Da die Expression von Zielmolekülen wahrscheinlich zwischen Tumorzellen und Tumorstammzellen differiert, sind Arzneimittel gegen spezielle Moleküle unter Umständen nicht in der Lage, Tumorstammzellen zu eliminieren¹⁹⁶. Zytostatika könnten jedoch ebenfalls wirkungslos sein, da Stammzellen generell eine geringe Teilungsrate aufweisen und zudem Transportproteine exprimieren können, die zytotoxische Substanzen aus der Zelle ausschleusen. Zusammengenommen bedeutet dies, dass sämtliche Therapeutika zwar in der Lage wären, die Tumormasse drastisch zu reduzieren, die eigentlich tumorinitiierenden Zellen jedoch nicht erfasst würden. Ein erneutes Tumorwachstum wäre vorprogrammiert. Die detaillierte Charakterisierung disseminierter Tumorzellen als Zielzellen adjuvanter Therapien scheint also dringend notwendig.

6 Zusammenfassung

Trotz der frühzeitigen Operation von Brustkrebspatientinnen in kurativer Absicht und der breiten Anwendung adjuvanter Therapien entwickeln viele Patientinnen nach mehreren Jahren Fernmetastasen und versterben an der Erkrankung. Ausgangspunkt der systemischen Krankheitsprogression sind disseminierte Tumorzellen, die nach der Operation im Körper der Patientin verbleiben. Ihr Nachweis im Knochenmark ist mit einer schlechten Prognose für die Betroffenen assoziiert, gleichzeitig stellen sie die eigentlichen Zielzellen adjuvanter Therapien dar. Sowohl Arzneimittelforschung als auch Therapieentscheidungen werden jedoch derzeit am Primärtumor durchgeführt. Die Ineffektivität der adjuvanter Therapie bei vielen Patientinnen verdeutlicht die Notwendigkeit einer genaueren Charakterisierung der disseminierten Tumorzellen.

Ziel der vorliegenden Dissertation war daher die Erarbeitung einer Methode zur Untersuchung der Genexpression einzelner Zellen auf genomweiten Oligonukleotidarrays und die Anwendung des Verfahrens an disseminierten Tumorzellen, die aus dem Knochenmark von Patientinnen mit Mammakarzinom isoliert werden können. Dadurch sollten wichtige Einblicke in die Biologie der Zellen ermöglicht und potentieller Marker für Diagnose und Therapie gefunden werden.

Die entwickelte Methode besteht aus Isolation und Amplifikation der mRNA einzelner Zellen gefolgt von einer Hybridisierung der Primäramplifikate auf Oligonukleotidarrays, die etwa 39000 Sonden umfassen. Beide Teile des Protokolls wurden durch unabhängige Verfahren auf Reproduzierbarkeit und systematische Fehler untersucht. Zur Validierung der Amplifikation dienten quantitative PCR-Reaktionen in einem Vergleich mit nicht amplifizierten Proben. Eine Verfälschung relativer Transkriptverhältnisse durch die Amplifikation wurde dabei nicht beobachtet. Die Hybridisierung wurde durch genspezifische qualitative und quantitative PCR überprüft. Es wurde eine hohe Übereinstimmung hinsichtlich Sensitivität, Spezifität und gemessener Genregulation festgestellt. Die Reproduzierbarkeit von sowohl Amplifikation als auch Hybridisierung war hoch.

Anschließende Genexpressionsuntersuchungen disseminierter Tumorzellen mit dem beschriebenen Verfahren auf Nylon- und Oligonukleotidarrays zeigten eine sehr heterogene Zellpopulation. Die meisten Zellen exprimierten keine Marker der terminalen epithelialen Differenzierung. Bei vielen dieser Zellen handelte es sich jedoch sicher um Tumorzellen, da parallel zum Transkriptom die genomische DNA der Zellen untersucht und chromosomale Veränderungen gefunden wurden. Einige Zellen wiesen eine Expression von Markern auf, die mit Stamm- und Vorläuferzellen der Brustentwicklung in Verbindung gebracht worden sind. Sollte daher zumindest ein Teil der disseminierten Tumorzellen im Gegensatz zum Primärtumor Stamm- und Progenitoreigenschaften besitzen, hätte das deutliche Konsequenzen für die adjuvante Therapie. Eine Untersuchung weiterer disseminierter Tumorzellen über den Rahmen dieser Arbeit hinaus sollte zur Beantwortung dieser Frage beitragen sowie mögliche Marker als Angriffspunkte für neue Therapeutika definieren helfen.

7 Literaturverzeichnis

1. Robert Koch Institut (www.rki.de, 2005).
2. Statistisches Bundesamt (www.destatis.de, 2005).
3. CochraneDatabase. Multi-agent chemotherapy for early breast cancer. *Cochrane Database Syst Rev*, CD000487 (2002).
4. EBCTCG. Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. *Lancet* 365, 1687-717 (2005).
5. Sobin, L. H., Wittekind, C. *TNM Classification of Malignant Tumours* (Wiley-Liss, New York, 2002).
6. Elston, C. W. & Ellis, I. O. Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up. *Histopathology* 19, 403-10 (1991).
7. Sauer, H. *Tumormanual Mammakarzinome* (W. Zuckschwerdt Verlag, Tumorzentrum München, 2003).
8. Pegram, M. D. et al. Phase II study of receptor-enhanced chemosensitivity using recombinant humanized anti-p185HER2/neu monoclonal antibody plus cisplatin in patients with HER2/neu-overexpressing metastatic breast cancer refractory to chemotherapy treatment. *J Clin Oncol* 16, 2659-71 (1998).
9. Slamon, D. J. et al. Studies of the HER-2/neu proto-oncogene in human breast and ovarian cancer. *Science* 244, 707-12 (1989).
10. Bergh, J., Norberg, T., Sjogren, S., Lindgren, A. & Holmberg, L. Complete sequencing of the p53 gene provides prognostic information in breast cancer patients, particularly in relation to adjuvant systemic therapy and radiotherapy. *Nat Med* 1, 1029-34 (1995).
11. Sorlie, T. et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci U S A* 98, 10869-74 (2001).
12. Wang, Y. et al. Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer. *Lancet* 365, 671-9 (2005).
13. van 't Veer, L. J. et al. Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 415, 530-6 (2002).
14. Goss, P. E. Changing clinical practice: extending the benefits of adjuvant endocrine therapy in breast cancer. *Semin Oncol* 31, 15-22 (2004).
15. Rote Liste (www.rote-liste.de, 2005).
16. Emmerich, B. *Tumormanual Maligne Lymphome* (W. Zuckschwerdt Verlag, Tumorzentrum München, 2003).
17. Hiddemann, W. & Haferlach, T. *Tumormanual Leukämien und myelodysplastische Syndrome* (W. Zuckschwerdt Verlag, Tumorzentrum München, 2003).
18. Dearnaley, D. P. et al. Increased detection of mammary carcinoma cells in marrow smears using antisera to epithelial membrane antigen. *Br J Cancer* 44, 85-90 (1981).
19. Mansi, J. L. et al. Micrometastases in bone marrow in patients with primary breast cancer: evaluation as an early predictor of bone metastases. *Br Med J (Clin Res Ed)* 295, 1093-6 (1987).
20. Schlimok, G. et al. Micrometastatic cancer cells in bone marrow: in vitro detection with anti-cytokeratin and in vivo labeling with anti-17-1A monoclonal antibodies. *Proc Natl Acad Sci U S A* 84, 8672-6 (1987).
21. Pantel, K., Felber, E. & Schlimok, G. Detection and characterization of residual disease in breast cancer. *J Hematother* 3, 315-22 (1994).

22. Benoy, I. H. et al. Real-time RT-PCR correlates with immunocytochemistry for the detection of disseminated epithelial cells in bone marrow aspirates of patients with breast cancer. *Br J Cancer* 91, 1813-20 (2004).
23. Slade, M. J., Smith, B. M., Sinnett, H. D., Cross, N. C. & Coombes, R. C. Quantitative polymerase chain reaction for the detection of micrometastases in patients with breast cancer. *J Clin Oncol* 17, 870-9 (1999).
24. Hermanek, P., Hutter, R. V., Sobin, L. H. & Wittekind, C. International Union Against Cancer. Classification of isolated tumor cells and micrometastasis. *Cancer* 86, 2668-73 (1999).
25. Wittekind, C., Wagner, G. *TNM-Klassifikation maligner Tumoren* (Springer Verlag, Heidelberg, 1997).
26. Hermanek, P. What's new in TNM? *Pathol Res Pract* 190, 97-102 (1994).
27. Slade, M. J. et al. Persistence of bone marrow micrometastases in patients receiving adjuvant therapy for breast cancer: results at 4 years. *Int J Cancer* 114, 94-100 (2005).
28. Janni, W. et al. Prognostic significance of an increased number of micrometastatic tumor cells in the bone marrow of patients with first recurrence of breast carcinoma. *Cancer* 88, 2252-9 (2000).
29. Kasimir-Bauer, S., Oberhoff, C., Schindler, A. E. & Seeber, S. A summary of two clinical studies on tumor cell dissemination in primary and metastatic breast cancer: methods, prognostic significance and implication for alternative treatment protocols (Review). *Int J Oncol* 20, 1027-34 (2002).
30. Braun, S. et al. Lack of effect of adjuvant chemotherapy on the elimination of single dormant tumor cells in bone marrow of high-risk breast cancer patients. *J Clin Oncol* 18, 80-6 (2000).
31. Tsavellas, G., Patel, H. & Allen-Mersh, T. G. Detection and clinical significance of occult tumour cells in colorectal cancer. *Br J Surg* 88, 1307-20 (2001).
32. Janku, F. et al. Mammaglobin A, a novel marker of minimal residual disease in early stages breast cancer. *Neoplasma* 51, 204-8 (2004).
33. Roggel, F. et al. Minimal residual disease in breast cancer and gynecological malignancies: phenotype and clinical relevance. *Recent Results Cancer Res* 162, 89-100 (2003).
34. Diel, I. J. et al. Micrometastatic breast cancer cells in bone marrow at primary surgery: prognostic value in comparison with nodal status. *J Natl Cancer Inst* 88, 1652-8 (1996).
35. Pierga, J. Y. et al. Clinical significance of immunocytochemical detection of tumor cells using digital microscopy in peripheral blood and bone marrow of breast cancer patients. *Clin Cancer Res* 10, 1392-400 (2004).
36. Braun, S. et al. A pooled analysis of bone marrow micrometastasis in breast cancer. *N Engl J Med* 353, 793-802 (2005).
37. Cote, R. J., Rosen, P. P., Lesser, M. L., Old, L. J. & Osborne, M. P. Prediction of early relapse in patients with operable breast cancer by detection of occult bone marrow micrometastases. *J Clin Oncol* 9, 1749-56 (1991).
38. Mansi, J. L. et al. Outcome of primary-breast-cancer patients with micrometastases: a long-term follow-up study. *Lancet* 354, 197-202 (1999).
39. Braun, S. et al. Cytokeratin-positive cells in the bone marrow and survival of patients with stage I, II, or III breast cancer. *N Engl J Med* 342, 525-33 (2000).
40. Woelfle, U. et al. Bi-specific immunomagnetic enrichment of micrometastatic tumour cell clusters from bone marrow of cancer patients. *J Immunol Methods* 300, 136-45 (2005).
41. Izbicki, J. R. et al. Prognostic value of immunohistochemically identifiable tumor cells in lymph nodes of patients with completely resected esophageal cancer. *N Engl J Med* 337, 1188-94 (1997).
42. Kubuschok, B., Passlick, B., Izbicki, J. R., Thetter, O. & Pantel, K. Disseminated tumor cells in lymph nodes as a determinant for survival in surgically resected non-small-cell lung cancer. *J Clin Oncol* 17, 19-24 (1999).
43. Rosenberg, R. et al. Prognostic evaluation and review of immunohistochemically detected disseminated tumor cells in peritumoral lymph nodes of patients with pN0 colorectal cancer. *Int J Colorectal Dis* 19, 430-7 (2004).

44. Balzar, M., Winter, M. J., de Boer, C. J. & Litvinov, S. V. The biology of the 17-1A antigen (Ep-CAM). *J Mol Med* 77, 699-712 (1999).
45. Osta, W. A. et al. EpCAM is overexpressed in breast cancer and is a potential target for breast cancer gene therapy. *Cancer Res* 64, 5818-24 (2004).
46. Spizzo, G. et al. Prognostic significance of Ep-CAM AND Her-2/neu overexpression in invasive breast cancer. *Int J Cancer* 98, 883-8 (2002).
47. Riethmuller, G. et al. Monoclonal antibody therapy for resected Dukes' C colorectal cancer: seven-year outcome of a multicenter randomized trial. *J Clin Oncol* 16, 1788-94 (1998).
48. Riethmuller, G. et al. Randomised trial of monoclonal antibody for adjuvant therapy of resected Dukes' C colorectal carcinoma. German Cancer Aid 17-1A Study Group. *Lancet* 343, 1177-83 (1994).
49. Braun, S. et al. Monoclonal antibody therapy with edrecolomab in breast cancer patients: monitoring of elimination of disseminated cytokeratin-positive tumor cells in bone marrow. *Clin Cancer Res* 5, 3999-4004 (1999).
50. Hempel, P. et al. Combination of high-dose chemotherapy and monoclonal antibody in breast-cancer patients: a pilot trial to monitor treatment effects on disseminated tumor cells. *Cytotherapy* 2, 287-95 (2000).
51. Kirchner, E. M., Gerhards, R. & Voigtmann, R. Sequential immunochemotherapy and edrecolomab in the adjuvant therapy of breast cancer: reduction of 17-1A-positive disseminated tumour cells. *Ann Oncol* 13, 1044-8 (2002).
52. Schwartzberg, L. S. Clinical experience with edrecolomab: a monoclonal antibody therapy for colorectal carcinoma. *Crit Rev Oncol Hematol* 40, 17-24 (2001).
53. Thurm, H. et al. Rare expression of epithelial cell adhesion molecule on residual micrometastatic breast cancer cells after adjuvant chemotherapy. *Clin Cancer Res* 9, 2598-604 (2003).
54. Kasimir-Bauer, S. et al. Rare expression of target antigens for immunotherapy on disseminated tumor cells in breast cancer patients without overt metastases. *Int J Mol Med* 12, 969-75 (2003).
55. Braun, S., Hepp, F., Sommer, H. L. & Pantel, K. Tumor-antigen heterogeneity of disseminated breast cancer cells: implications for immunotherapy of minimal residual disease. *Int J Cancer* 84, 1-5 (1999).
56. Weihrauch, M. R. et al. Immunomagnetic enrichment and detection of isolated tumor cells in bone marrow of patients with epithelial malignancies. *Clin Exp Metastasis* 19, 617-21 (2002).
57. Zhong, X. Y., Kaul, S., Lin, Y. S., Eichler, A. & Bastert, G. Sensitive detection of micrometastases in bone marrow from patients with breast cancer using immunomagnetic isolation of tumor cells in combination with reverse transcriptase/polymerase chain reaction for cytokeratin-19. *J Cancer Res Clin Oncol* 126, 212-8 (2000).
58. Muller, P., Weckermann, D., Riethmuller, G. & Schlimok, G. Detection of genetic alterations in micrometastatic cells in bone marrow of cancer patients by fluorescence in situ hybridization. *Cancer Genet Cytogenet* 88, 8-16 (1996).
59. Braun, S. et al. ErbB2 overexpression on occult metastatic cells in bone marrow predicts poor clinical outcome of stage I-III breast cancer patients. *Cancer Res* 61, 1890-5 (2001).
60. Dietmaier, W. et al. Multiple mutation analyses in single tumor cells with improved whole genome amplification. *Am J Pathol* 154, 83-95 (1999).
61. Hosch, S. et al. Malignant potential and cytogenetic characteristics of occult disseminated tumor cells in esophageal cancer. *Cancer Res* 60, 6836-40 (2000).
62. Pantel, K. et al. Establishment of micrometastatic carcinoma cell lines: a novel source of tumor cell vaccines. *J Natl Cancer Inst* 87, 1162-8 (1995).
63. Klein, C. A. et al. Comparative genomic hybridization, loss of heterozygosity, and DNA sequence analysis of single cells. *Proc Natl Acad Sci U S A* 96, 4494-9 (1999).
64. Klein, C. A. et al. Combined transcriptome and genome analysis of single micrometastatic cells. *Nat Biotechnol* 20, 387-92 (2002).

65. Hensen, A. et al. Comparative evaluation of urokinase-type plasminogen activator receptor expression in primary breast carcinomas and on metastatic tumor cells. *Int J Cancer* 107, 903-9 (2003).
66. Ditsch, N. et al. Estrogen receptor expression profile of disseminated epithelial tumor cells in bone marrow of breast cancer patients. *Recent Results Cancer Res* 162, 141-7 (2003).
67. Schmidt-Kittler, O. et al. From latent disseminated cells to overt metastasis: genetic analysis of systemic breast cancer progression. *Proc Natl Acad Sci U S A* 100, 7737-42 (2003).
68. Schardt, J. A. et al. Genomic analysis of single cytokeratin-positive cells from bone marrow reveals early mutational events in breast cancer. *Cancer Cell* 8, 227-39 (2005).
69. Alberts, B., Johnson, A., Lewis, J. et al. *Molecular Biology of the Cell* (Garland Publishing, New York, 2002).
70. Van Gelder, R. N. et al. Amplified RNA synthesized from limited quantities of heterogeneous cDNA. *Proc Natl Acad Sci U S A* 87, 1663-7 (1990).
71. Baugh, L. R., Hill, A. A., Brown, E. L. & Hunter, C. P. Quantitative analysis of mRNA amplification by in vitro transcription. *Nucleic Acids Res* 29, E29 (2001).
72. Brady, G. & Iscove, N. N. Construction of cDNA libraries from single cells. *Methods Enzymol* 225, 611-23 (1993).
73. Todd, R. & Margolin, D. H. Challenges of single-cell diagnostics: analysis of gene expression. *Trends Mol Med* 8, 254-7 (2002).
74. Phillips, J. K. & Lipski, J. Single-cell RT-PCR as a tool to study gene expression in central and peripheral autonomic neurones. *Auton Neurosci* 86, 1-12 (2000).
75. Li, J., Adams, L., Schwartz, S. M. & Bumgamer, R. E. RNA amplification, fidelity and reproducibility of expression profiling. *C R Biol* 326, 1021-30 (2003).
76. Lashkari, D. A. et al. Yeast microarrays for genome wide parallel genetic and gene expression analysis. *Proc Natl Acad Sci U S A* 94, 13057-62 (1997).
77. Schena, M., Shalon, D., Davis, R. W. & Brown, P. O. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* 270, 467-70 (1995).
78. DeRisi, J. et al. Use of a cDNA microarray to analyse gene expression patterns in human cancer. *Nat Genet* 14, 457-60 (1996).
79. Klur, S., Toy, K., Williams, M. P. & Certa, U. Evaluation of procedures for amplification of small-size samples for hybridization on microarrays. *Genomics* 83, 508-17 (2004).
80. Player, A., Barrett, J. C. & Kawasaki, E. S. Laser capture microdissection, microarrays and the precise definition of a cancer cell. *Expert Rev Mol Diagn* 4, 831-40 (2004).
81. Jordan, R., Somiari, S.B. et al. Performanve and Bioinformatic Evaluation of Overlapping Variable Length Oligonucleotide Probes Used in Spotted Microarrays. *Transactions of the Integrated Biomedical Informatics & Enabling Technologies Symposium Journal* 2004 1, 15-24 (2004).
82. Kane, M. D. et al. Assessment of the sensitivity and specificity of oligonucleotide (50mer) microarrays. *Nucleic Acids Res* 28, 4552-7 (2000).
83. Barczak, A. et al. Spotted long oligonucleotide arrays for human gene expression analysis. *Genome Res* 13, 1775-85 (2003).
84. Wang, H. Y. et al. Assessing unmodified 70-mer oligonucleotide probe performance on glass-slide microarrays. *Genome Biol* 4, R5 (2003).
85. Kothapalli, R., Yoder, S. J., Mane, S. & Loughran, T. P., Jr. Microarray results: how accurate are they? *BMC Bioinformatics* 3, 22 (2002).

86. Hughes, T. R. et al. Expression profiling using microarrays fabricated by an ink-jet oligonucleotide synthesizer. *Nat Biotechnol* 19, 342-7 (2001).
87. Rogojina, A. T., Orr, W. E., Song, B. K. & Geisert, E. E., Jr. Comparing the use of Affymetrix to spotted oligonucleotide microarrays using two retinal pigment epithelium cell lines. *Mol Vis* 9, 482-96 (2003).
88. Chuaqui, R. F. et al. Post-analysis follow-up and validation of microarray experiments. *Nat Genet* 32 Suppl, 509-14 (2002).
89. Holloway, A. J., van Laar, R. K., Tothill, R. W. & Bowtell, D. D. Options available--from start to finish--for obtaining data from DNA microarrays II. *Nat Genet* 32 Suppl, 481-9 (2002).
90. Lee, J. K. et al. Comparing cDNA and oligonucleotide array data: concordance of gene expression across platforms for the NCI-60 cancer cells. *Genome Biol* 4, R82 (2003).
91. Tan, P. K. et al. Evaluation of gene expression measurements from commercial microarray platforms. *Nucleic Acids Res* 31, 5676-84 (2003).
92. Miklos, G. L. & Maleszka, R. Microarray reality checks in the context of a complex disease. *Nat Biotechnol* 22, 615-21 (2004).
93. Irizarry, R. A. et al. Multiple-laboratory comparison of microarray platforms. *Nat Methods* 2, 345-50 (2005).
94. Howley, P. M., Israel, M. A., Law, M. F. & Martin, M. A. A rapid method for detecting and mapping homology between heterologous DNAs. Evaluation of polyomavirus genomes. *J Biol Chem* 254, 4876-83 (1979).
95. Kerr, M. K. & Churchill, G. A. Statistical design and the analysis of gene expression microarray data. *Genet Res* 77, 123-8 (2001).
96. Churchill, G. A. Fundamentals of experimental design for cDNA microarrays. *Nat Genet* 32 Suppl, 490-5 (2002).
97. Yang, M. C., Yang, J. J., McIndoe, R. A. & She, J. X. Microarray experimental design: power and sample size considerations. *Physiol Genomics* 16, 24-8 (2003).
98. Holzman, T. & Kolker, E. Statistical analysis of global gene expression data: some practical considerations. *Curr Opin Biotechnol* 15, 52-7 (2004).
99. Tusher, V. G., Tibshirani, R. & Chu, G. Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci U S A* 98, 5116-21 (2001).
100. Smyth, G. K. & Speed, T. Normalization of cDNA microarray data. *Methods* 31, 265-73 (2003).
101. Yang, Y. H. et al. Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. *Nucleic Acids Res* 30, e15 (2002).
102. Storey, J. D. & Tibshirani, R. Statistical significance for genomewide studies. *Proc Natl Acad Sci U S A* 100, 9440-5 (2003).
103. Heyer, L. J., Kruglyak, S. & Yooseph, S. Exploring expression data: identification and analysis of coexpressed genes. *Genome Res* 9, 1106-15 (1999).
104. Asyali, M. H., Shoukri, M. M., Demirkaya, O. & Khabar, K. S. Assessment of reliability of microarray data and estimation of signal thresholds using mixture modeling. *Nucleic Acids Res* 32, 2323-35 (2004).
105. Slonim, D. K. From patterns to pathways: gene expression data analysis comes of age. *Nat Genet* 32 Suppl, 502-8 (2002).
106. Eisen, M. B., Spellman, P. T., Brown, P. O. & Botstein, D. Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci U S A* 95, 14863-8 (1998).
107. Golub, T. R. et al. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* 286, 531-7 (1999).
108. Kohonen, T. The self-organizing map. *Proceedings of the IEEE* 78, 1464-1480 (1990).

109. Yeung, K. Y. & Ruzzo, W. L. Principal component analysis for clustering gene expression data. *Bioinformatics* 17, 763-74 (2001).
110. Muller. An Introduction to Kernel Based Learning Algorithms. *IEEE Transactions on Neural Networks* 12, 181-202 (2001).
111. Burgess. A Tutorial on Support Vector Machines for Pattern Recognition. *Data Mining and Knowledge Discovery* 2, 121-167 (1998).
112. Chang, J. C. et al. Gene expression profiling for the prediction of therapeutic response to docetaxel in patients with breast cancer. *Lancet* 362, 362-9 (2003).
113. van de Vijver, M. J. et al. A gene-expression signature as a predictor of survival in breast cancer. *N Engl J Med* 347, 1999-2009 (2002).
114. Michiels, S., Koscielny, S. & Hill, C. Prediction of cancer outcome with microarrays: a multiple random validation strategy. *Lancet* 365, 488-92 (2005).
115. Ein-Dor, L., Kela, I., Getz, G., Givol, D. & Domany, E. Outcome signature genes in breast cancer: is there a unique set? *Bioinformatics* 21, 171-8 (2005).
116. Glinsky, G. V., Higashiyama, T. & Glinskii, A. B. Classification of human breast cancer using gene expression profiling as a component of the survival predictor algorithm. *Clin Cancer Res* 10, 2272-83 (2004).
117. Janssen, T. K. & Hovig, E. Gene-expression profiling in breast cancer. *Lancet* 365, 634-5 (2005).
118. Shi, L. et al. QA/QC: challenges and pitfalls facing the microarray community and regulatory agencies. *Expert Rev Mol Diagn* 4, 761-77 (2004).
119. Gruvberger, S. K. et al. Expression profiling to predict outcome in breast cancer: the influence of sample selection. *Breast Cancer Res* 5, 23-6 (2003).
120. Shen, R., Ghosh, D. & Chinnaiyan, A. M. Prognostic meta-signature of breast cancer developed by two-stage mixture modeling of microarray data. *BMC Genomics* 5, 94 (2004).
121. Heidenblad, M. et al. Microarray analyses reveal strong influence of DNA copy number alterations on the transcriptional patterns in pancreatic cancer: implications for the interpretation of genomic amplifications. *Oncogene* 24, 1794-801 (2005).
122. Pollack, J. R. et al. Microarray analysis reveals a major direct role of DNA copy number alteration in the transcriptional program of human breast tumors. *Proc Natl Acad Sci U S A* 99, 12963-8 (2002).
123. Hyman, E. et al. Impact of DNA amplification on gene expression patterns in breast cancer. *Cancer Res* 62, 6240-5 (2002).
124. Mao, R., Zielke, C. L., Zielke, H. R. & Pevsner, J. Global up-regulation of chromosome 21 gene expression in the developing Down syndrome brain. *Genomics* 81, 457-67 (2003).
125. du Manoir, S. et al. Detection of complete and partial chromosome gains and losses by comparative genomic in situ hybridization. *Hum Genet* 90, 590-610 (1993).
126. Kallioniemi, O. P. et al. Optimizing comparative genomic hybridization for analysis of DNA sequence copy number changes in solid tumors. *Genes Chromosomes Cancer* 10, 231-43 (1994).
127. Kallioniemi, A. et al. Comparative genomic hybridization for molecular cytogenetic analysis of solid tumors. *Science* 258, 818-21 (1992).
128. Mielnicki, L. M., Asch, H. L. & Asch, B. B. Genes, chromatin, and breast cancer: an epigenetic tale. *J Mammary Gland Biol Neoplasia* 6, 169-82 (2001).
129. Feinberg, A. P. & Tycko, B. The history of cancer epigenetics. *Nat Rev Cancer* 4, 143-53 (2004).
130. Speicher, M. R. et al. Molecular cytogenetic analysis of formalin-fixed, paraffin-embedded solid tumors by comparative genomic hybridization after universal DNA-amplification. *Hum Mol Genet* 2, 1907-14 (1993).

131. Goodell, M. A., Brose, K., Paradis, G., Conner, A. S. & Mulligan, R. C. Isolation and functional properties of murine hematopoietic stem cells that are replicating in vivo. *J Exp Med* 183, 1797-806 (1996).
132. *Molecular Cloning* (eds. Sambrook, J. & Russel, D.) (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 2001).
133. Chomczynski, P. & Sacchi, N. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal Biochem* 162, 156-9 (1987).
134. Heil, S. G., Kluijtmans, L. A., Spiegelstein, O., Finnell, R. H. & Blom, H. J. Gene-specific monitoring of T7-based RNA amplification by real-time quantitative PCR. *Biotechniques* 35, 502-4, 506-8 (2003).
135. Nygaard, V. et al. Effects of mRNA amplification on gene expression ratios in cDNA experiments estimated by analysis of variance. *BMC Genomics* 4, 11 (2003).
136. Feldman, A. L. et al. Advantages of mRNA amplification for microarray analysis. *Biotechniques* 33, 906-12, 914 (2002).
137. Iscove, N. N. et al. Representation is faithfully preserved in global cDNA amplified exponentially from sub-picogram quantities of mRNA. *Nat Biotechnol* 20, 940-3 (2002).
138. Granucci, F. et al. Inducible IL-2 production by dendritic cells revealed by global gene expression analysis. *Nat Immunol* 2, 882-8 (2001).
139. Yuen, T., Wurmbach, E., Pfeffer, R. L., Ebersole, B. J. & Sealfon, S. C. Accuracy and calibration of commercial oligonucleotide and custom cDNA microarrays. *Nucleic Acids Res* 30, e48 (2002).
140. Park, P. J. et al. Current issues for DNA microarrays: platform comparison, double linear amplification, and universal RNA reference. *J Biotechnol* 112, 225-45 (2004).
141. Iwasaki, A. & Medzhitov, R. Toll-like receptor control of the adaptive immune responses. *Nat Immunol* 5, 987-95 (2004).
142. Poltorak, A. et al. Defective LPS signaling in C3H/HeJ and C57BL/10ScCr mice: mutations in Tlr4 gene. *Science* 282, 2085-8 (1998).
143. Suzuki, N. et al. Severe impairment of interleukin-1 and Toll-like receptor signalling in mice lacking IRAK-4. *Nature* 416, 750-6 (2002).
144. Suzuki, N. et al. IL-1R-associated kinase 4 is required for lipopolysaccharide-induced activation of APC. *J Immunol* 171, 6065-71 (2003).
145. Caldwell, C. W., Patterson, W. P., Toalson, B. D. & Yesus, Y. W. Surface and cytoplasmic expression of CD45 antigen isoforms in normal and malignant myeloid cell differentiation. *Am J Clin Pathol* 95, 180-7 (1991).
146. Craig, W., Poppema, S., Little, M. T., Dragowska, W. & Lansdorp, P. M. CD45 isoform expression on human haemopoietic cells at different stages of development. *Br J Haematol* 88, 24-30 (1994).
147. Firat, H. et al. Determination of myeloid antigen expression on childhood acute lymphoblastic leukaemia cells: discrepancies using different monoclonal antibody clones. *Leuk Lymphoma* 42, 75-82 (2001).
148. Garnache-Ottou, F. et al. Expression of the myeloid-associated marker CD33 is not an exclusive factor for leukemic plasmacytoid dendritic cells. *Blood* 105, 1256-64 (2005).
149. Elghetany, M. T. Surface antigen changes during normal neutrophilic development: a critical review. *Blood Cells Mol Dis* 28, 260-74 (2002).
150. Edvardsson, L., Dykes, J., Olsson, M. L. & Olofsson, T. Clonogenicity, gene expression and phenotype during neutrophil versus erythroid differentiation of cytokine-stimulated CD34+ human marrow cells in vitro. *Br J Haematol* 127, 451-63 (2004).
151. Shah, V. O., Civin, C. I. & Loken, M. R. Flow cytometric analysis of human bone marrow. IV. Differential quantitative expression of T-200 common leukocyte antigen during normal hemopoiesis. *J Immunol* 140, 1861-7 (1988).

152. Fujimoto, H. et al. Flow cytometric method for enumeration and classification of reactive immature granulocyte populations. *Cytometry* 42, 371-8 (2000).
153. Dressman, M. A. et al. Genes that co-cluster with estrogen receptor alpha in microarray analysis of breast biopsies. *Pharmacogenomics J* 1, 135-41 (2001).
154. Grubberger, S. et al. Estrogen receptor status in breast cancer is associated with remarkably distinct gene expression patterns. *Cancer Res* 61, 5979-84 (2001).
155. Nagai, M. A. et al. Gene expression profiles in breast tumors regarding the presence or absence of estrogen and progesterone receptors. *Int J Cancer* 111, 892-9 (2004).
156. Al-Hajj, M., Wicha, M. S., Benito-Hernandez, A., Morrison, S. J. & Clarke, M. F. Prospective identification of tumorigenic breast cancer cells. *Proc Natl Acad Sci U S A* 100, 3983-8 (2003).
157. Shackleton, M. et al. Generation of a functional mammary gland from a single stem cell. *Nature* 439, 84-8 (2006).
158. Stingl, J. et al. Purification and unique properties of mammary epithelial stem cells. *Nature* (2006).
159. Smalley, M. & Ashworth, A. Stem cells and breast cancer: A field in transit. *Nat Rev Cancer* 3, 832-44 (2003).
160. Clarke, R. B. et al. A putative human breast stem cell population is enriched for steroid receptor-positive cells. *Dev Biol* 277, 443-56 (2005).
161. Eberwine, J. et al. mRNA expression analysis of tissue sections and single cells. *J Neurosci* 21, 8310-4 (2001).
162. Bahar, R., Hartmann, C. H., Rodriguez, K. A. & Denny, A. D. Increased cell-to-cell variation in gene expression in aging mouse heart. (submitted).
163. Chen, Z., Gordon, J. R., Zhang, X. & Xiang, J. Analysis of the gene expression profiles of immature versus mature bone marrow-derived dendritic cells using DNA arrays. *Biochem Biophys Res Commun* 290, 66-72 (2002).
164. Messmer, D., Messmer, B. & Chiorazzi, N. The global transcriptional maturation program and stimuli-specific gene expression profiles of human myeloid dendritic cells. *Int Immunol* 15, 491-503 (2003).
165. Kaern, M., Elston, T. C., Blake, W. J. & Collins, J. J. Stochasticity in gene expression: from theories to phenotypes. *Nat Rev Genet* 6, 451-64 (2005).
166. Lammers, R. et al. Monoclonal antibody 9C4 recognizes epithelial cellular adhesion molecule, a cell surface antigen expressed in early steps of erythropoiesis. *Exp Hematol* 30, 537-45 (2002).
167. Albertson, D. G., Collins, C., McCormick, F. & Gray, J. W. Chromosome aberrations in solid tumors. *Nat Genet* 34, 369-76 (2003).
168. Tlsty, T. D. et al. Genetic and epigenetic changes in mammary epithelial cells may mimic early events in carcinogenesis. *J Mammary Gland Biol Neoplasia* 9, 263-74 (2004).
169. Piccart-Gebhart, M. J. et al. Trastuzumab after adjuvant chemotherapy in HER2-positive breast cancer. *N Engl J Med* 353, 1659-72 (2005).
170. Ross, M. T. et al. The DNA sequence of the human X chromosome. *Nature* 434, 325-37 (2005).
171. Mischo, A. et al. Prospective study on the expression of cancer testis genes and antibody responses in 100 consecutive patients with primary breast cancer. *Int J Cancer* 118, 696-703 (2006).
172. Xiao, J. & Chen, H. S. Biological functions of melanoma-associated antigens. *World J Gastroenterol* 10, 1849-53 (2004).
173. Vogelstein, B. et al. Genetic alterations during colorectal-tumor development. *N Engl J Med* 319, 525-32 (1988).
174. Allred, D. C., Mohsin, S. K. & Fuqua, S. A. Histological and biological evolution of human premalignant breast disease. *Endocr Relat Cancer* 8, 47-61 (2001).

175. Hanahan, D. & Weinberg, R. A. The hallmarks of cancer. *Cell* 100, 57-70 (2000).
176. Ramaswamy, S., Ross, K. N., Lander, E. S. & Golub, T. R. A molecular signature of metastasis in primary solid tumors. *Nat Genet* 33, 49-54 (2003).
177. Weigelt, B. et al. Gene expression profiles of primary breast tumors maintained in distant metastases. *Proc Natl Acad Sci U S A* 100, 15901-5 (2003).
178. Ma, X. J. et al. Gene expression profiles of human breast cancer progression. *Proc Natl Acad Sci U S A* 100, 5974-9 (2003).
179. Nishidate, T. et al. Genome-wide gene-expression profiles of breast-cancer cells purified with laser microbeam microdissection: identification of genes associated with progression and metastasis. *Int J Oncol* 25, 797-819 (2004).
180. Abba, M. C. et al. Transcriptomic changes in human breast cancer progression as determined by serial analysis of gene expression. *Breast Cancer Res* 6, R499-513 (2004).
181. Sotiriou, C. et al. Breast cancer classification and prognosis based on gene expression profiles from a population-based study. *Proc Natl Acad Sci U S A* 100, 10393-8 (2003).
182. Lodish, H., Berk, A. & al., e. *Molecular Cell Biology* (W.H. Freeman and Company, New York, 2003).
183. Porter, D. A. et al. A SAGE (serial analysis of gene expression) view of breast tumor progression. *Cancer Res* 61, 5697-702 (2001).
184. Reya, T., Morrison, S. J., Clarke, M. F. & Weissman, I. L. Stem cells, cancer, and cancer stem cells. *Nature* 414, 105-11 (2001).
185. Bonnet, D. & Dick, J. E. Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. *Nat Med* 3, 730-7 (1997).
186. Miyamoto, T., Weissman, I. L. & Akashi, K. AML1/ETO-expressing nonleukemic stem cells in acute myelogenous leukemia with 8;21 chromosomal translocation. *Proc Natl Acad Sci U S A* 97, 7521-6 (2000).
187. Lapidot, T. et al. A cell initiating human acute myeloid leukaemia after transplantation into SCID mice. *Nature* 367, 645-8 (1994).
188. Park, C. H., Bergsagel, D. E. & McCulloch, E. A. Mouse myeloma tumor stem cells: a primary cell culture assay. *J Natl Cancer Inst* 46, 411-22 (1971).
189. Hamburger, A. W. & Salmon, S. E. Primary bioassay of human tumor stem cells. *Science* 197, 461-3 (1977).
190. Han, B. & Zhang, J. T. Multidrug resistance in cancer chemotherapy and xenobiotic protection mediated by the half ATP-binding cassette transporter ABCG2. *Curr Med Chem Anti-Canc Agents* 4, 31-42 (2004).
191. Kim, M. et al. The multidrug resistance transporter ABCG2 (breast cancer resistance protein 1) effluxes Hoechst 33342 and is overexpressed in hematopoietic stem cells. *Clin Cancer Res* 8, 22-8 (2002).
192. Wiseman, B. S. & Werb, Z. Stromal effects on mammary gland development and breast cancer. *Science* 296, 1046-9 (2002).
193. Pasca di Magliano, M. & Hebrok, M. Hedgehog signalling in cancer formation and maintenance. *Nat Rev Cancer* 3, 903-11 (2003).
194. de Boer, C. J., van Krieken, J. H., Janssen-van Rhijn, C. M. & Litvinov, S. V. Expression of Ep-CAM in normal, regenerating, metaplastic, and neoplastic liver. *J Pathol* 188, 201-6 (1999).
195. Cirulli, V. et al. KSA antigen Ep-CAM mediates cell-cell adhesion of pancreatic epithelial cells: morphoregulatory roles in pancreatic islet development. *J Cell Biol* 140, 1519-34 (1998).
196. Al-Hajj, M., Becker, M. W., Wicha, M., Weissman, I. & Clarke, M. F. Therapeutic implications of cancer stem cells. *Curr Opin Genet Dev* 14, 43-7 (2004).

8 Danksagung

PD Dr. Christoph Klein danke ich für die hervorragende Betreuung dieser Arbeit. Seine Anregungen und Ideen sowie sein umfassendes Know-how in fachlichen und praktischen Fragen waren stets eine große Hilfe. Ich danke ihm außerdem für die Freiräume zum selbständigen Arbeiten, für seinen Einsatz zur Finanzierung des Projekts und meiner Stelle und nicht zuletzt für Motivation und freundschaftliche Zusammenarbeit.

Herrn Prof. Dr. Martin Biel möchte ich danken für die Bereitschaft sich als Doktorvater dieser Dissertation zur Verfügung zu stellen und meine Promotion vor der Fakultät für Chemie und Pharmazie zu vertreten, sowie für sein Interesse an den Fortschritten der Arbeit während der letzten Jahre.

Für Rat und Motivation danke ich Sara Mangialaio und Céline Pallaud.

Daneben danke ich allen Patientinnen, die mit ihrer Einwilligung zur Verwendung ihrer Knochenmarkpunktate einen Teil dieser Dissertation erst möglich gemacht haben, sowie den klinischen Kooperationspartnern, die mir neben dem Material auch die relevanten Patientendaten zur Verfügung gestellt haben.

Ich danke allen Mitgliedern der AG Klein am Institut für Immunologie für die tolle Zeit während der Doktorarbeit und die viele Hilfe auf allen Ebenen. Besonderer Dank für Rat und Tat gilt Isabell Blochberger, Elke Burghart und Kerstin Möhr für ihre Unterstützung bei zahlreichen Libraries und der Aufarbeitung der Knochenmarksproben. Für Einarbeitung, gute Diskussionen, einen kritischen Blick auf diese Arbeit und viel Spaß danke ich besonders Bernhard Polzer, Christine Fuhrmann und Nikolas Stoecklein, ohne die viele Dinge unmöglich oder nicht so lustig gewesen wären.

Mein besonderer Dank gilt meiner Familie - meinen Schwestern und vor allem meinen Eltern - ohne deren uneingeschränkte Unterstützung ich heute sicher nicht da wäre, wo ich bin.

Und nicht zuletzt und vor allem - danke Alex.

9 Lebenslauf

Claudia Helena Hartmann

geboren am 01.03.1977

in Heilbronn

- 1983-1987 Besuch der Hölderlin-Grundschule in Lauffen am Neckar
- 1987-1996 Besuch des Hölderlin-Gymnasiums in Lauffen am Neckar
1996 Allgemeine Hochschulreife
- 1996-2000 Pharmaziestudium an der Universität Regensburg
1998 Erstes Staatsexamen
2000 Zweites Staatsexamen
- 2000-2001 Sechsmonatiges Pharmaziepraktikum in der Elb-Apotheke, Hamburg
2001 Sechsmonatiges Pharmaziepraktikum bei der Novartis Pharma AG, Basel
- 2001 Drittes Staatsexamen
Pharmazeutische Prüfung
- 2002 Approbation als Apothekerin
- 2002 Beginn der vorliegenden Dissertation an der Fakultät für Chemie und Pharmazie der Ludwig-Maximilians-Universität München betreut von Prof. Dr. Martin Biel. Die praktischen Arbeiten wurden am Institut für Immunologie der Ludwig-Maximilians-Universität München im Labor von PD Dr. Christoph Klein durchgeführt.

10 Veröffentlichungen und Stipendien

Urbanowska T, Mangialaio S, Hartmann C, Legay F. (2003). Development of protein microarray technology to monitor biomarkers of rheumatoid arthritis disease. *Cell Biol Toxicol* **19** (3): 198-202

Schardt J¹, Meyer M¹, Hartmann CH¹, Schubert F, Schmidt-Kittler O, Fuhrmann C, Polzer B, Petronio M, Eils R, Klein CA. (2005). Genomic analysis of single cytokeratin-positive cells from bone marrow reveals early mutational events in breast cancer. *Cancer Cell* **8** (3) 227-39

¹: equal contributors

Hartmann CH, Klein CA. Gene expression profiling of single cells on large-scale oligonucleotide arrays. (*eingereicht*)

Bahar R, Hartmann CH, Rodriguez KA, Denny AD, Busuttill RA, Dollé MET, Calder RB, Chisholm GB, Pollock BH, Klein CA, Vijg J. Increased cell-to-cell variation in gene expression in aging mouse heart. (*akzeptiert zur Publikation in Nature*)

Stoecklein NH, Hosch SB, Stern F, Hartmann CH, Siegmund A, Scheunemann P, Knoefel WT, Verde PE, Erbersdobler A, Izbicki JR, Klein CA. Direct target evaluation on disseminated cancer cells: HER2 is a target for adjuvant therapy in a subgroup of oesophageal cancer patients. (*eingereicht*)

Posterpräsentationen

Christoph A. Klein, Claudia H. Hartmann, Andrea Eberhardt, Susanne Steinhaus, Julie Waldispühl. (2002, 2003, 2004, 2005, 2006). Struktur und Funktion der Genome einzelner mikrometastatischer Tumorzellen: Ein neuer Zugang zur klonalen Evolution von Zellpopulationen in menschlichen Karzinomen. *Jahrestagung des BMBF-Wettbewerbs „BioFuture“*, **Berlin**

Claudia H. Hartmann & Christoph A. Klein. (2004). Gene expression analysis of single cells using high-density oligonucleotide arrays. *AACR Special conference “Chromatin, Chromosomes, and Cancer Epigenetics”*, **Waikoloa, Hawaii**

Claudia H. Hartmann, Daniel Oruzio, Christoph A. Klein. (2006). Gene expression profiling of single cells on high-density oligonucleotide arrays: A novel approach to study early micrometastatic cancer cells. *Accepted for the AACR 97th Annual Meeting 2006*, **Washington, DC**

Stipendium

AACR Pfizer Award 2004: Abstractpreis in Form eines Reisestipendiums zur Teilnahme an der *AACR Special conference “Chromatin, Chromosomes, and Cancer Epigenetics”*, **Waikoloa, Hawaii**

11 Anhang

CGH-Profile aller untersuchten EpCAM-positiven Zellen

