

Identifikation von metastasierungs- relevanten Genen in humanen und murinen Tumorzellen



Dissertation zur Erlangung des Grades

Doktor der Naturwissenschaften

vorgelegt von

Benjamin Otto

beim Department Biologie

der Fakultät für Mathematik, Informatik und Naturwissenschaften

Universität Hamburg

Hamburg 2008

Genehmigt vom Department Biologie
der Fakultät für Mathematik, Informatik und Naturwissenschaften
an der Universität Hamburg
auf Antrag von Professor Dr. C. WAGENER
Weiterer Gutachter der Dissertation:
Herr Professor Dr. W. SCHÄFER
Tag der Disputation: 30. Januar 2009

Hamburg, den 16. Januar 2009



Jörg Ganzhorn

Professor Dr. Jörg Ganzhorn
Leiter des Departments Biologie

Danksagung

Der vorliegenden Dissertationsschrift liegt nicht nur viel Arbeit zugrunde, sondern auch eine riesige Menge an Unterstützung, sowohl fachliche als auch moralische. All jenen, die es mir dadurch überhaupt erst ermöglicht haben, diese Arbeit erfolgreich abzuschließen möchte ich hiermit meinen Dank aussprechen.

Herrn Prof. Dr. C. Wagener möchte ich daher für die Gelegenheit zur Durchführung dieser Arbeit danken, für fachlichen Rat und die fortwährende Unterstützung.

Herrn Prof. Dr. W. Schäfer möchte ich für seine Unterstützung danken und dafür, daß er es mir ermöglicht hat, diese Arbeit im Department Biologie zu veröffentlichen.

Dr. Thomas Streichert danke ich für seine Anleitung, fachliche und moralische Unterstützung, Anregungen, Kritik und Hilfestellung.

Wer „Microarrays am Uniklinikum Hamburg-Eppendorf“ sagt, muss auch „Kristin Reeck“ sagen! Danke Kristin, ohne Deine Hilfe hätte ich wirklich alt ausgesehen.

Prof. W. Deppert, Dr. Genrich Tolstonog und der Gruppe um Prof. U. Schumacher und Dr. Valentiner möchte ich für die Kooperation in den beiden Projekten und die anregenden Gespräche danken, die diese Arbeit ermöglicht haben.

Und schließlich ein Dank an meinem Arbeitskreis und all die guten Geister mit ihrer moralischen Unterstützung. Es gibt vermutlich nicht viele Arbeitskreise, in denen man so gut aufgehoben ist, wie bei Euch. Auch wenn ihr mir mit Euren Fragen, Kritiken und Korrekturen mitunter Kopfzerbrechen bereitet haben mögt, danke . . . danke . . . es hat geholfen; heute wissen wir ja nun endlich auch, daß der Lebertran im Kindesalter gut für den Knochenaufbau war.

Zuletzt und ohne viele Worte: *Danke Muddi, Danke Walied und ...*

Lina . . .

Was ich an dir mag, ist unbeschreibbar,
ist ein ungeschriebenes Gedicht
Besser ich beschränke mich aufs Staunen,
mag dich weiter und erklär dich nicht.

Was ich an dir mag, ist das Geheime,
jedes Wort zuviel ist schon Gefahr,
denn so schnell verfällt ins Allgemeine,
was zuvor so ganz besonders war.

Ist es deine Stimme, sind es deine Hände,
ach - dein ganzes Wesen fesselt mich,
um dich zu beschreiben, bräucht es Bände,
besser sag ich schlicht: Ich liebe dich.

Was ich an dir mag, ist ohne Frage
auch ein Teil von dem, was in mir ist,
was ich mir erhoffe, auch was ich beklage,
alles, was man an sich selbst vermisst.

Manchmal ist's dein Gang
und manches Mal dein Lachen,
manchmal deine Frechheit, deine Wut.
Eins ist nicht genug - was soll ich machen,
was ich an dir mag: Du tust mir gut.

Was ich an dir mag, ist das Geheime,
jedes Wort zuviel ist schon Gefahr,
denn so schnell verfällt ins Allgemeine,
was zuvor so ganz besonders war.

Konstantin Wecker, 1993

Danke für Deine Geduld!

Mouse's Tale

Fury said to
a mouse, That
he met
in the
house,
'Let us
both go
to law:
I will
prosecute
you. —
Come, I'll
take no
denial;
We must
have a
trial:
For
really
this
morning
I've
nothing
to do.'
Said the
mouse to
the cur,
Such a
trial,
dear sir,
With no
jury or
judge,
would be
wasting
our breath.'
'I'll be
judge,
I'll be
jury.'
Said
cunning
old Fury;
'I'll try
the whole
cause
and
condemn
you
to
death.'

Inhaltsverzeichnis

1	Einleitung	1
1.1	Tumoren und Progression	2
1.1.1	Konzepte der Differenzierung	2
1.1.2	Konzepte der Metastasierung	3
1.2	Mausmodelle	4
1.2.1	WAP-T Mausmodell der Brusttumore	5
1.2.2	Mausmodell der Neuroblastome	7
1.3	Aufgabenstellung	9
2	Material und Methoden	11
2.1	Materialien und Bezugsquellen	11
2.2	Laboranalytische Methoden	15
2.2.1	Sterilisierung der Arbeitsmaterialien	15
2.2.2	Photometrische Ermittlung von RNA-Konzentrationen	16
2.2.3	Ermittlung von RNA-Qualität mittels Kapillargelelektrophorese	16
2.2.4	Expressionsarrays	17
2.2.5	SNP Arrays	24
2.2.6	MAPK Transkriptionsfaktor-Arrays	29
2.3	Programme und statistische Methoden	33
2.3.1	Programme und Umgebungen	33
2.3.2	Scannen der Chips und Bildtransformation	35
2.3.3	Hintergrundkorrektur, Normalisierung und Segmentierung	35
2.3.4	Visualisierungsverfahren	40
2.4	Supervidierte und unsupervidierte Testverfahren	42
2.4.1	Hypothesentests und Korrektur für multiples Testen	42
2.4.2	Clustering-Verfahren	43
2.4.3	Korrelationsverfahren	44

2.4.4	Signaltransduktionswege und Funktionsanalytik	48
2.4.5	Funktionsaufrufe und Flags	48
2.5	Dateiformate	49
3	Ergebnisse	51
3.1	Expressionsanalytik	51
3.1.1	Brusttumor	52
3.1.2	Neuroblastome	64
3.2	Copy-Number-Variations & Genotypisierung	80
3.2.1	Brusttumor: CGH-Arrays	80
3.2.2	Neuroblastome: SNP-Arrays	85
3.3	Korrelation	94
3.3.1	Brusttumor	94
3.3.2	Neuroblastome	100
3.4	MAPK-Arrays	106
4	Diskussion	109
4.1	Brusttumor	109
4.2	Neuroblastome	120
4.3	Zusammenfassung	134
	Einheiten & Abkürzungen	136
	Literatur	137
	A Brusttumor	153
	B Neuroblastome	161

Kapitel 1

Einleitung

In Deutschland erkranken derzeit jährlich rund 206.000 Frauen an Krebs [74]. Davon stellt Brustkrebs, mit 57.000 Neuerkrankungen pro Jahr, die häufigste Krebserkrankung dar. Unter Kindern sind derzeit rund 18.000 Krebsneuerkrankungen pro Jahr zu beobachten [75]. Den größten Anteil davon bilden mit etwa 4.000 Neuerkrankungen jährlich die ZNS-Tumore. Als einer der wichtigsten Vertreter dieser Tumorklasse ist das Neuroblastom eine der häufigsten bösartigen Krebserkrankungen im Kindesalter. Sowohl vom medizinischen Standpunkt, als auch volkswirtschaftlich sind Möglichkeiten zur Prognose und Therapie dieser beiden Krebsarten sehr wünschenswert.

Im Zentrum dieser Arbeit standen zwei Modelle zur Beschreibung von zwei Eigenschaften, die die Aggressivität von Tumoren mitbestimmen. Mittels Genexpressionsanalysen und genomischen Analysen wurden in Neuroblastomen Determinanten der Metastasierung und in Brusttumoren Determinanten der Differenzierung untersucht. Modellübergreifend sollte dabei das Potential einer Integration der genomischen und transkriptomischen Daten evaluiert werden.

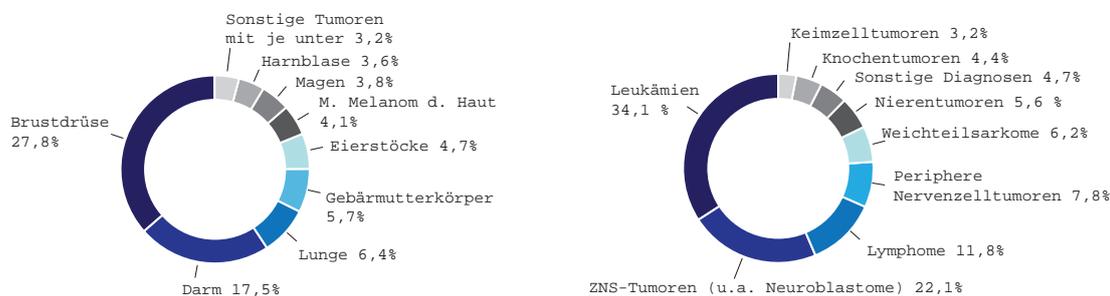


Abbildung 1.1: Häufigkeit der wichtigsten Tumore bei Frauen (links) und im Kindesalter (rechts). Die Daten sind den aktuellen Broschüren des Robert Koch Instituts [74, 75] entnommen. Zeitraum der Datenerhebung: 1997-2006 (Kinder) und 2004 (Frauen).

1.1 Tumoren und Progression

Verschiedene Befunde sprechen dafür, daß Tumore durch genetische Veränderung entstehen. Sowohl die Aktivierung als auch die Inaktivierung kritischer Gene können demgemäß Tumorwachstum verursachen. Verbreitet ist das Modell des klonalen Wachstums. Dabei entstehen Tumore aus einer einzelnen maligne entarteten Zelle. Die Eigenschaften der Malignität gehen von der Mutter- zur Tochterzelle über. Da mitunter mehrere Mutationen oder Aberrationen für eine vollständige Ausprägung des malignen Phänotyps verantwortlich sind, allgemein auch als *Mehrschritt-Karzinogenese* bezeichnet, geht das Modell der klonalen Evolution davon aus, daß die mutagenen Ereignisse den Zellklonen einen Überlebensvorteil bieten.

Die Ausbreitung transformierter Zellen im Organismus wird allgemein als Tumorprogression bezeichnet. Zu Beginn weisen die Zellen noch Zeichen der Differenzierung auf. Die Tumorprogression ist durch eine zunehmende Entdifferenzierung oder Fehldifferenzierung gekennzeichnet. Der Grad der Differenzierung bzw. Fehldifferenzierung ist ein Charakteristikum für die Malignität. Besonders aggressive Tumore können im weiteren Verlauf die Zell-Zell-Adhäsion verlieren. Einzelne Zellen oder kleinere Zellgruppen können dann den Zellverband verlassen. Während der folgenden Metastasierung durchdringen diese Zellen die Blut- oder Lymphgefäße und gelangen in ferne Körperorgane, wo sie sich einnisten und neue Geschwülste bilden können.

1.1.1 Konzepte der Differenzierung

Die Differenzierung ist ein Maß für den Ausreifungsgrad der Tumorzellen. Normalerweise unterliegen Zellteilung, Differenzierung und Apoptose der Zellen genauen homöostatischen Kontrollmechanismen. Im neoplastischen Gewebe ist die Homöostase zulasten der Differenzierung und Apoptose in Richtung Proliferation verschoben und die Apoptose blockiert. In gut differenzierten Geweben kann die Gewebestruktur morphologisch verändert, aber insgesamt erhalten sein. Geringer differenzierte Tumore weisen Tumorareale mit und ohne dem Ursprungsgewebe ähnlicher Struktur auf. In entdifferenzierten Karzinomen ist oft nicht mehr erkennbar, aus welchem Gewebe der Tumor entstand. Der Grad der Differenzierung wird in den Stufen G1 (differenziert) bis G4 (undifferenziert) angegeben. Für viele Tumore ist ein enger Zusammenhang zwischen Tumordifferenzierung und klinischer Prognose nachgewiesen worden [175].

1.1. TUMOREN UND PROGRESSION

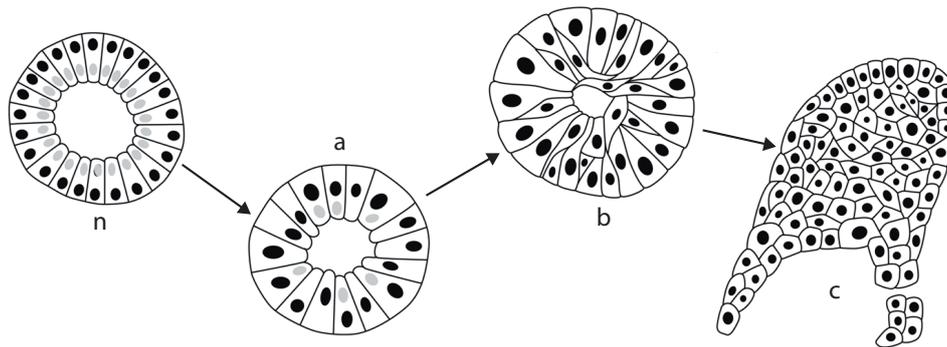


Abbildung 1.2: Verschiedene Differenzierungsgrade am Beispiel eines einschichtigen, schleimbildenden Zylinderepithels.

n, normaler Drüsenschlauch; a, geringgradige Entdifferenzierung, Schleimbildung z.T. noch erhalten; b, hochgradige Entdifferenzierung, Drüsenschlauchstruktur noch erhalten; c, vollständige Entdifferenzierung mit Verlust der Drüsenschlauchstruktur (nach Wagener, 1999; nach Zollinger, 1974).

Durch die Verzahnung der Prozesse ist der spezifische Einfluss von Onkogenen auf die einzelnen Abläufe wie Differenzierung, Invasion, Angiogenese oder Metastasierung oft nur sehr schlecht zu trennen. Häufig als Wächter des Genoms bezeichnet [90], spielt p53 in der Tumorentwicklung eine wichtige Rolle. Es kontrolliert Reparaturmechanismen bei DNA-Schäden und Gene, die an der Kontrolle von Zellzyklus und Apoptose beteiligt sind. Weitere Gene, denen mehrfach wichtige Rollen nachgewiesen wurden, kodieren für Bindungspartner von p53, z.B. MDM2, Zelladhäsionsmoleküle der CEA-Familie oder CD44 sowie Transkriptionsfaktoren wie MYC. Speziell in Brusttumoren sind die beiden Gene *BRCA1* und *BRCA2* für ca. 80% der familiären, früh auftretenden Brustkrebsfälle verantwortlich [175].

1.1.2 Konzepte der Metastasierung

Die Vorgänge während der Metastasierung sind sehr organisiert, äußerst komplex und bei weitem noch nicht vollständig verstanden. Sie erfolgen in mehreren verzahnten Stufen, während derer Tumorzellen den Primärtumor verlassen, in den Blutkreislauf gelangen, an einen weiter entfernten Ort des Körpers gelangen, sich dort einnisten und erneut wachsen. In der ersten Stufe, der Invasion, ist der Tumor durch lockere Zell-Zell-Adhäsion, erhöhte Proliferation und Motilität sowie eine herabgesetzte Alterung der malignen Zellen gekennzeichnet [175]. Während der Intravasation penetrieren die Tumorzellen das Endothel der Blut- oder Lymphgefäße. Zum Überleben müssen sie eine erhöhte Resistenz gegenüber dem

Immunsystem, z.B. Zytokinen, und eine herabgesetzte Apoptose aufweisen [17]. Über die Gefäßsysteme werden sie in entfernte Körperorgane transportiert. Gelangen sie dabei zu einer Nische, die ihnen gute Wachstumsbedingungen bietet, nisten sie sich dort ein. Ob die metastatische Nische aktiv geschaffen wird oder die Einnistung passiv erfolgt, ist jedoch noch nicht geklärt.

Die grundlegenden Mechanismen werden kontrovers diskutiert. Ein Modell geht davon aus, daß sich einige disseminierte Tumorzellen vom Primärtumor absiedeln und erst nach der Extravasation einige Zellen die zusätzliche Fähigkeit zur Metastasenbildung erlangen [34]. Demgegenüber steht die Theorie, dergemäß bereits während der Etablierung des Primärtumors Zellen mit erhöhtem metastatischem Potential entstehen [58, 22]. Hierbei handelt es sich nicht um eine Akkumulation, sondern um primäre genetische Veränderungen, die das Schicksal der Zelle von Beginn an bestimmen. Jüngst werden einige dieser Zellen mit sogenannten Krebs-Stammzellen in Verbindung gebracht, die häufig eine Resistenz gegen klassische Therapien aufweisen [50].

1.2 Mausmodelle

Ein Verständnis der Biologie von Tumoren ist nur dann adäquat möglich, wenn die Untersuchungen *in vivo* durchgeführt werden. Es ist kaum möglich in Zellkultur das äußerst komplexe Umfeld eines kompletten Organismus zu simulieren. Humane Patientenproben bieten den Vorteil einer guten Wiedergabe der biologischen Situation. Andererseits bieten sie lediglich eine Momentaufnahme eines Prozesses, der sich von Patient zu Patient unterscheidet. Nicht zuletzt unterscheidet sich der biologische und medizinische Werdegang, der zur Tumorentstehung und Progression geführt hat. Diese Faktoren erschweren eine Analyse der betroffenen Faktoren, insbesondere bei konkreten Fragestellungen.

Tiermodelle hingegen erlauben Verlaufsuntersuchungen und weitestgehend standardisierte, kontrollierte Bedingungen. Diese sollen die Vergleichbarkeit der Prozesse während der Tumorgenese erhöhen. Es bleibt allerdings stets zu beachten, daß die kontrollierten Bedingungen bei induzierten Tumoren in Tiermodellen die komplexe Biologie, die zur spontanen Tumorentstehung im Menschen führen, nur in Annäherung wiedergeben.

1.2. MAUSMODELLE

1.2.1 WAP-T Mausmodell der Brusttumore

Brustkrebs ist die häufigste weibliche Krebsart und häufigste Todesursache bei Krebserkrankungen unter Frauen. Derzeit erkranken rund 57.000 Frauen in Deutschland pro Jahr an Brustkrebs [74]. Rund zehn Prozent der Frauen mit Brustkrebs weisen weitere Erkrankungsfälle in der Familie auf [166]. Dennoch sind familiär vererbte, genetische Veränderungen, die Brustkrebs begünstigen, selten. Die höchste Wahrscheinlichkeit einer genetischen Prädisposition liegt bei Mutationen der Gene *BRCA1* und *BRCA2* vor. Weitere bekannte genetische Veränderungen, die Brusttumoren propagieren, sind Mutationen der Gene *TP53* oder *PTEN* [175].

Die Klassifikation der Tumore erfolgt nach dem TNM-System. Der Grad der Tumordifferenzierung spiegelt dabei häufig die Aggressivität der Tumore wider. Er wird hierbei in vier Stufen von G1 (deutlich differenziert) bis G4 (undifferenziert) angegeben.

Zur Untersuchung von Determinanten der Differenzierung in Brusttumoren und des Einflusses von mutiertem p53 wurde von Schulze-Garg et al. [142] ein transgenes BALB/c Mausmodell mit der Bezeichnung WAP-T etabliert. Ein weiteres darauf aufbauendes Modell wurde von Krepulat et al. [88] etabliert. Diese beiden werden im Folgenden als NP8 Modell bzw. T1 Modell und zusammengefasst als WAP-T Modell bezeichnet.

WAP-T Mäuse tragen ein Transgen, das aus der frühen Genregion des *simian Virus 40* SV40 unter Kontrolle des *murinen whey acidic protein WAP*-Promotors besteht. Zum Ende der Schwangerschaft und während der Laktation wird der Promotor in Mammaepithelzellen durch laktotrophe Hormone wie Prolaktin und Östrogen induziert. In Folge wird das *SV40 T-Antigen* lokal und temporär in den Zellen des Brustdrüsengewebes transkribiert.

Die frühe Genregion des SV40 Virus kodiert für das große T-Antigen (LT), das kleine T-Antigen (st) und das 17kT Protein. Das große T-Antigen, im Folgenden *SV40 T-Antigen* genannt, wird aufgrund seiner Fähigkeit, Zellen unter bestimmten Bedingungen in Kultur zu transformieren, als Onkogen eingestuft. Die transformierende Aktivität basiert zu großen Teilen auf Protein-Protein Interaktionen mit zellulären Proteinen. Für das Mausmodell relevant ist seine Interaktion mit dem Tumorsupressorprotein p53. Dabei wird p53 komplexiert und inaktiviert [142]. Desweiteren inaktiviert das SV40 T-Antigen pRb, das Mitglieder der E2F Transkriptionsfaktorfamilie durch Bindung inhibiert. Durch Freigabe der E2F Faktoren kann der Eintritt der Zellen in die S-Phase gefördert werden [142, 88]. Es folgt eine erhöhte Proliferation der Zellen. Das kleine T-Antigen (st) kann pp2A binden,

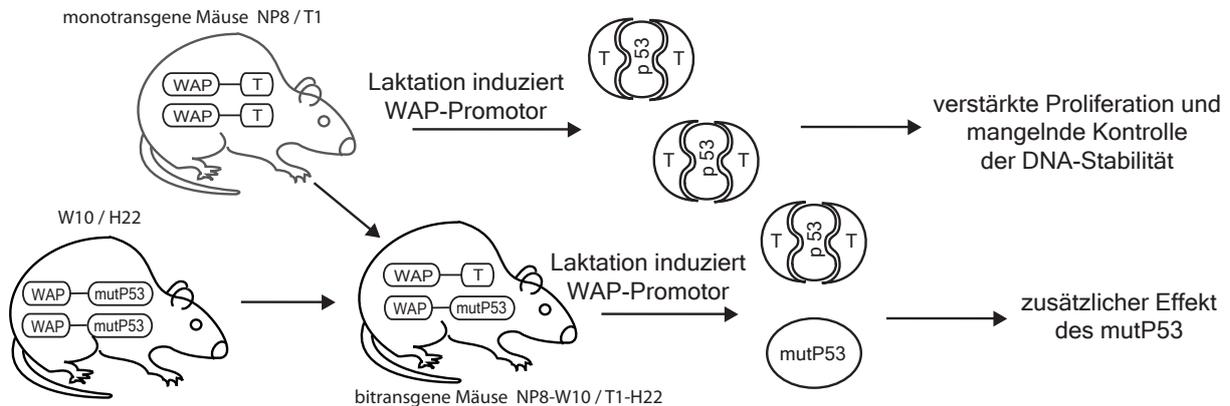


Abbildung 1.3: WAP-T Mausmodell zur Analyse von Differenzierungsdeterminanten in Brusttumoren. Monotransgene Mäuse tragen ein Konstrukt aus *WAP*-Promotor und *SV40 T-Antigen*, das während der Laktation exprimiert wird. Das T-Antigen hemmt die Wildtyp-p53 Aktivität. Bitransgene Mäuse exprimieren zusätzlich eine mutierte Variante des p53. Dadurch wird eine *mutP53 gain of function* Situation simuliert.

die Phosphatase-Aktivität inhibieren und zu einer konstitutiven Aktivierung des WNT-Signaltransduktionsweges führen [40]. In vielen humanen Mammakarzinomen werden genetische Veränderungen in Form von Mutationen in *ARF* und *pRB*, sowie Repression von *p16* und Hochregulation von Zyklinen und Zyklin-abhängigen Kinasen detektiert [152]. Dadurch wird die Funktion von *p53* und *pRB* inhibiert. Diese Situation wird im WAP-T Mausmodell simuliert. Im Gegensatz zu den T1 Mäusen tragen Mäuse des NP8 Modells zusätzlich zum *SV40 T-Antigen* ein *MHC class I (H-2d) restricted T-cell* Epitop des LCM Virus im WAP-T-Konstrukt.

Die soeben beschriebenen WAP-T Mäuse werden zukünftig als *monotransgene* Mäuse bezeichnet. Zur Simulation einer *gain of function* Situation eines mutierten *Trp53* Gens tragen *H22* transgene Mäuse ein *WAP-mutP53* Konstrukt. Die *mutP53* Sequenz kodiert für ein *p53* Protein mit einer R270H Mutation. Demgegenüber stehen die *W10* transgenen Mäuse. Sie tragen eine R245W Mutation des *p53*. Aus der Kreuzung von NP8 Mäusen mit *W10* Mäusen gingen NP8-*W10* Mäuse hervor. Aus der Kreuzung von T1 Mäusen mit *H22* Mäusen gingen T1-*H22* Mäuse hervor. Diese beiden Kreuzungen werden zukünftig als *bitransgene* Mäuse bezeichnet. NP8 und NP8-*W10* Mäuse werden zur Unterscheidung in der folgenden Arbeit als NP8 Modell bezeichnet, T1 und T1-*H22* Mäuse als T1 Modell.

1.2. MAUSMODELLE

1.2.2 Mausmodell der Neuroblastome

Unter der Bezeichnung Neuroblastome wird in der Regel ein Spektrum von neuroblastischen Tumoren zusammengefasst, die sich aus primitiven sympathischen Ganglionzellen entwickeln [167]. Neuroblastome sind klinisch heterogen. Sie variieren in der Lokalisation, im histologischen Phänotyp und den biologischen Eigenschaften. Das außerordentlich breite Spektrum an klinischem Verhalten erstreckt sich von spontaner Rückbildung, über die Reifung zu benignen Ganglioneuromen, bis hin zu einer aggressiven Erkrankung mit metastatischer, oft tödlicher, Ausbreitung. Die klinische Diversität korreliert gut mit zahlreichen klinischen und biologischen Faktoren. Dennoch sind die molekularen Hintergründe zu großen Teilen nach wie vor ungeklärt.

Zur Analyse von metastatischen Determinanten in Neuroblastomen sowie Korrelationen zwischen genetischem und transkriptomischem Habitus wurde auf ein Modell immundefizienter BALB/c SCID Mäuse [169] zurückgegriffen. Bisher ist eine Vielzahl von Mutationen bekannt, die die Funktion des Immunsystems in Mäusen beeinflussen. Das ursprünglich von Bosma et al. [27, 28] etablierte BALB/c SCID Mausmodell, formal *severe combined immunodeficiency* Modell, weist kombinierte Gendefekte auf, die Anzahl und Funktion von T- und B-Zellen deutlich reduzieren. Die rezessive SCID Mutation auf Chromosom 16 hemmt die Rekombination von Antigenrezeptorgenen und die Differenzierung von Stammzellen zu reifen Lymphozyten. Die meisten homozygoten SCID Tiere weisen keine nachweisbaren Expressionen von *IgM*, *IgG1*, *IgG2a*, *IgG2b*, *IgG3* oder *IgA* auf. Bei Transfektionen kann somit eine Immunreaktion in der Maus als Wirtsorganismus umgangen werden. Immundefiziente Mausmodelle eignen sich daher gut zum Verständnis von Tumorwachstum und Progression.

Das hier verwendete, von Valentiner et. al. etablierte [169], BALB/c SCID Mausmodell beschreibt das Tumorwachstumsverhalten und die Expression von CD44 für mehrere Neuroblastomzelllinien. CD44 ist eine Gruppe transmembraner Glycoproteine, die als Rezeptor für Hyaluronsäure dienen. In der Regel wird mit CD44 eine Standardform assoziiert, die keine zusätzlichen Exondomänen aufweist. CD44 fördert Zell-Zell- und Zell-Matrix-Adhäsion sowie Migration [175]. Für CD44 ist bereits eine wichtige Rolle in der Invasion von Neuroblastomen beschrieben [17]. Im Gegensatz zu anderen Tumoren charakterisiert die Abwesenheit von CD44 in Neuroblastomen die Aggressivität des Tumors.

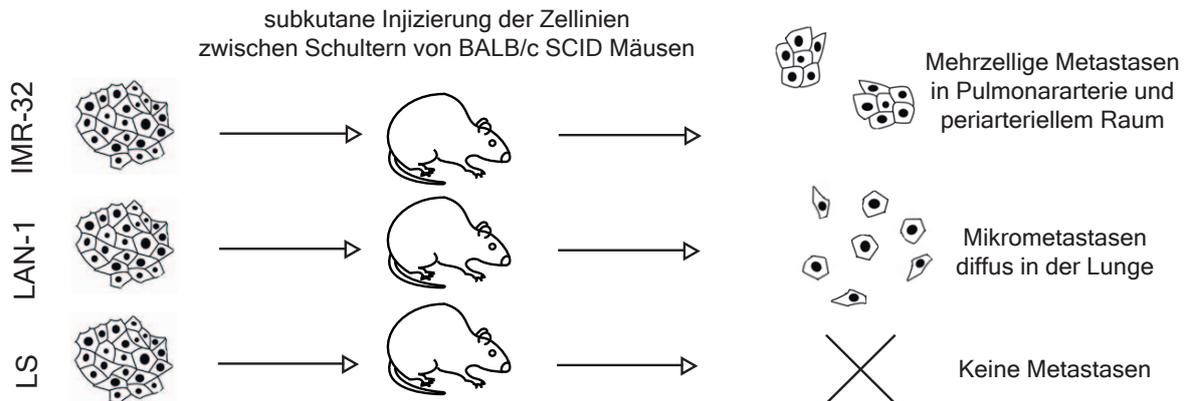


Abbildung 1.4: Mausmodell zur Analyse von Determinanten der Metastasierung in Neuroblastomen. BALB/c SCID Mäusen wurden Zellen dreier verschiedener Neuroblastomzelllinien von Valentiner et al. subkutan zwischen die Schultern injiziert. Die sich daraufhin entwickelnden Neuroblastome wurden hinsichtlich ihres Metastasierungsverhaltens untersucht.

Die im Rahmen dieser Arbeit erhobenen Daten beschreiben das Tumorwachstum von drei Neuroblastomzelllinien. Acht bis zehn Wochen alten BALB/c SCID Mäusen wurden Zellen der Zelllinien subkutan zwischen die Schulterblätter injiziert. Die Tumorsektion erfolgte nachdem die Tumore maximale Größe (bis zu 20% des ursprünglichen Körpergewichts) erreicht hatten. Die hier untersuchten und in SCID Mäuse injizierten humanen Neuroblastomzelllinien IMR-32, LAN-1 und LS zeigten in immunhistochemischen Analysen [169] unterschiedliche CD44 Expressionen. IMR-32 [163] wies mit einem Anteil von über 50% an CD44-positiven Zellen eine starke Immunreaktivität mit dem Antikörper auf. Sowohl LAN-1 [143] als auch LS-Zellen [136] waren CD44-negativ.

Im hier verwendeten Modell wiesen die drei Zelllinien ein unterschiedliches Metastasierungsverhalten auf. IMR-32 induzierte Neuroblastome entwickelten mehrzellige, primär in der Pulmonararterie und im periarteriellen Raum der Lunge lokalisierte Metastasen. LAN-1 induzierte Tumore entwickelten im verwendeten Modell Mikrometastasen, die die gesamte Lunge diffus infiltrierte. LS induzierte Tumore hingegen wiesen kein metastatisches Verhalten auf. Die drei Zelllinien eigneten sich somit gut zur Untersuchung von potentiell metastasierungsrelevanten Faktoren und des Einflusses von CD44 auf die Metastasierung.

1.3 Aufgabenstellung

Nach wie vor ist unklar welche Mechanismen Tumore zu einer Metastasierung befähigen oder die Differenzierung beeinflussen. Insbesondere stellt sich die Frage, inwieweit der Primarius schon hinsichtlich einer Metastasierung genetisch determiniert ist.

Ein Ansatz zur Untersuchung dieser möglichen Korrelationen zwischen Genotyp und Phänotyp liegt im Einsatz von microarray-basierten Technologien. Hier können primär zwei Strategien verfolgt werden: Erstens ermöglichen High-Throughput-Genotypisierungen und Expressionsanalysen tiefere Einblicke in Regulationen bei limitiertem biologischem Hintergrundwissen. Zweitens ist eine Identifikation von Genexpressions-Signaturen möglich, die den Aktivierungszustand von potentiell onkogenen Pathways reflektieren bzw. die Verwendung von a priori definierten Gen-Sets, die eine biologische Funktion oder genetische Information kodieren. Durch Analyse der Copy Number Variation (CNV) zusammen mit dem Expressionssignal der korrespondierenden Transkripte soll der Einfluss der genetischen Variation oder Läsion auf die Metastasierung und den Differenzierungsgrad abgeschätzt werden. Dieser Ansatz könnte dabei helfen, die Notwendigkeit einer vollständigen Sequenzierung des Tumorgenomes zu umgehen. Obwohl bereits erste Methoden zur Verknüpfung von Expressionssignaturen und genomischen Daten veröffentlicht wurden, ist die Kombination dieser Daten nach wie vor komplex, insbesondere da die Daten auf unterschiedlichen Plattformen generiert werden und in der Folge sehr inhomogen sind.

Im Rahmen dieser Arbeit sollten Tumormodelle untersucht werden, die sich in ihrem Metastasierungsverhalten bzw. der Differenzierung unterscheiden. Zur Analyse von Metastasierungs-faktoren wurden zuvor drei unterschiedlich stark metastasierende Neuroblastomzelllinien (stark metastasierend, Mikrometastasen bildend, keine Metastasierung) in SCID-Mäuse transfiziert, um anschließend Proben aus den sich entwickelnden Primärtumoren isolieren zu können. Zur Untersuchung von Differenzierungsfaktoren wurde ein Mausmodell für Mammarkarzinome mit mutiertem Trp53-Gen verwendet, das eine verstärkte Karzinomentwicklung aufwies. In diesem Modell sollten Primärtumore unterschiedlichen Gradings untersucht werden.

Mittels Microarray-Technologie (Affymetrix) sollte das Expressionsniveau in den Geweben ermittelt werden. Zusätzlich sollte die CNV der Gene mittels array-CGH (Agilent) für das Mausmodell bzw. über SNP-Arrays (Affymetrix) für die Neuroblastome ermittelt

werden. In den Expressionsdaten sollten sodann Vergleiche durchgeführt werden, um differenziell exprimierte Gene zwischen den Gruppen (metastasierend/ nicht-metastasierend, bzw. differenziert G1/G2 und undifferenziert G3/G4) zu identifizieren. Aus den SNP-Arrays sollte die CNV berechnet werden. Integrativ sollten die Resultate im Anschluß auf Korrelationen untersucht werden. Aufbauend auf aktuelle Verfahren sollte eine Methode entwickelt werden, die eine verbesserte Bestimmung der Korrelationsgüte sowohl für einzelne Gene als auch über das gesamte Genom einer Probe ermöglicht. Anhand von ontologischen Annotationen sollten die korrelierenden Gene hinsichtlich der molekularen Funktion oder als Vertreter biologischer Prozesse eingeordnet werden. Gegebenfalls sollten verstärkt betroffene Pathways identifiziert werden. Abschließend sollten die so ermittelten Daten in einen medizinisch-biologischen Rahmen eingebettet werden, der helfen könnte mögliche Kandidatengene zu identifizieren.

Kapitel 2

Material und Methoden

Diese Arbeit kann primär in zwei Abschnitte unterteilt werden. Die erste Phase umfasste die biologischen und laboranalytischen Schritte, ausgehend von der Aufbereitung des Probenmaterials bis zur Hybridisierung auf die Microarrays. Die zweite Phase umfasste die bioinformatische Dateninterpretation und statistische Auswertung, angefangen mit dem Scannen der Chips bis hin zum Entwurf einer biologischen Hypothese. In den folgenden Abschnitten werden die in dieser Arbeit verwendeten Materialien aufgeführt. Anschließend wird auf die verwendeten laboranalytischen Techniken sowie statistisch-bioinformatischen Methoden eingegangen.

2.1 Materialien und Bezugsquellen

Im folgenden sind die in dieser Arbeit verwendeten Materialien, Reagenzien und Geräte sowie ihre Bezugsquellen tabellarisch gelistet.

Tabelle 2.1: Reagenzien und Bezugsquellen

Reagenz	Hersteller
2-Propanol	Merck KGaA, Darmstadt
Agarose	Biozym, Hessisch Oldendorf
All Purpose Hi-Lo DNA Marker	Bionexus, Oakland, CA, USA
Aqua iniectabilia	Baxter, Unterschleißheim
Biotinylated Antibody	Vector Laboratories (Linaris), Wertheim-Bettingen
Bovine Serum Albumin (BSA)	Invitrogen Life Technologies, Karlsruhe
Buffer EB	Qiagen, Hilden
Chloroform	Merck KGaA, Darmstadt

Fortsetzung auf der nächsten Seite

Tabelle 2.1: Reagenzien und Bezugsquellen – Fortsetzung

Reagenz	Hersteller
Denhardt's Solution	Sigma-Aldrich, München
DEPC-Wasser	Ambion (Applied Biosystems), Darmstadt
DMSO	Sigma-Aldrich, München
EDTA Disodium Salt	Sigma-Aldrich, München
Essigsäure	Merck KGaA, Darmstadt
Ethanol	Merck KGaA, Darmstadt
Ethidiumbromid	Merck KGaA, Darmstadt
Gel Loading Buffer	Ambion (Applied Biosystems), Darmstadt
GeneChip Sample Cleanup Module	Affymetrix, Inc., Santa Clara, USA
Genome-Wide Human SNP Nsp/Sty Assay Kit 5.0/6.0	Affymetrix, Inc., Santa Clara, USA
Goat IgG	Sigma-Aldrich, München
Herring Sperm DNA	Promega Corporation, Mannheim
Human Cot-1 DNA	Invitrogen Life Technologies, Karlsruhe
Hybridization Control Kit	Affymetrix, Inc., Santa Clara, USA
IVT Labeling Kit	Affymetrix, Inc., Santa Clara, USA
Magnetic Beads	Agencourt (Beckman Coulter GmbH), Krefeld
MES hydrate SigmaUltra	Sigma-Aldrich, München
MES Sodium Salt	Sigma-Aldrich, München
N-Laurosarkosine	Sigma-Aldrich, München
Natriumchlorid (NaCl)	Ambion (Applied Biosystems), Darmstadt
Natriumhypochloritlösung	Roth, Karlsruhe
Nsp I	New England Biolabs, Frankfurt a.M.
NuPage 4-12% Bis-Tris Gel	Invitrogen Life Technologies, Karlsruhe
One Cycle cDNA Synthesis Kit	Affymetrix, Inc., Santa Clara, USA
Pellet Paint	Novagen (Merck), Darmstadt
Phenol-Chloroform-Isoamylalkohol	Fluka (Sigma-Aldrich), München
Prestained Protein Ladder, 10-180kDa	Fermentas Life Sciences, USA
Proteinase K	Roche, Mannheim
RNeasy Mini Kit	Qiagen, Hilden
Salzsäure	Merck KGaA, Darmstadt
Sample Cleanup Module	Affymetrix, Inc., Santa Clara, USA
Silver Staining Kit, PlusOne	Amersham Biosciences, USA
SSPE	Accugene (Lonza), Verviers, Belgium
Streptavidin Phycoerythrin	Molecular Probes (Invitrogen), Karlsruhe
Sty I	New England Biolabs, Frankfurt a.M.

Fortsetzung auf der nächsten Seite

2.1. MATERIALIEN UND BEZUGSQUELLEN

Tabelle 2.1: Reagenzien und Bezugsquellen – Fortsetzung

Reagenz	Hersteller
Surfactant-Amps 20	Pierce Chemical (Perbio Science), Bonn
T4 DNA Ligase	New England Biolabs, Frankfurt a.M.
Tetramethyl Ammonium Chlorid (TMACL)	Sigma-Aldrich, München
TF Chip MAPK kit	Eppendorf AG, Hamburg
Titanium DNA Amplification Kit	Clontech, St-Germain-en-Laye, France
Titriplex (EDTA)	Sigma-Aldrich, München
Trizma Base	Sigma-Aldrich, München
Trizol	Invitrogen Life Technologies, Karlsruhe
Urea Powder	Sigma-Aldrich, München

Ende Tabelle 2.1

Material	Hersteller
Deep Well Storage Plate, 2,4ml (Pooling Plate)	Greiner Bio One, Frickenhausen
Filter Tips 2,5µl , 10µl , 20µl , 1000µl	Sarstedt, Nümbrecht
Filter Tips 200µl	Eppendorf AG, Hamburg
Genome-Wide Human SNP Array 6.0	Affymetrix, Inc., Santa Clara, USA
Handschuhe, Latex	Hartmann, Heidenheim
Handschuhe, Nitril	Ansell, München
Human Genome U133 Plus 2.0	Affymetrix, Inc., Santa Clara, USA
Konische Röhrchen 15ml , 50ml	Greiner Bio One, Frickenhausen
Mouse Genome 430 2.0	Affymetrix, Inc., Santa Clara, USA
PCR-Gefäße 0,5ml	Sarstedt, Nümbrecht
Plate, 2ml , 96 Well Format Filterplate	Seahorse Bioscience, North Billerica, MA, USA
Polaroid-Filme 664 schwarz-weiß	Polaroid, Dreieich-Sprendlingen
Safe-Lock Tubes 0,5ml , 1,5ml , 2ml	Eppendorf AG, Hamburg
Sterilfiltrationssystem 0,5µm	Nalgene (Fisher Scientific), Schwerte

Tabelle 2.2: Materialien und Bezugsquellen

Gerät	Typ/Modell	Hersteller
Bioanalyzer	2100	Agilent Technologies, Santa Clara, USA
Eismaschine	AF-10	Scotsman, Milan, Italy
Gefrierschrank	-20°C	Liebherr, Biberach an der Riss
Gefrierschrank	-80°C	Kryotec, Hamburg
Gelkammer		PEQLAB, Erlangen
Heizblock	ThermoStat plus	Eppendorf AG, Hamburg
Hybridisierofen	GeneChip Hybridization Oven 640	Affymetrix, Inc., Santa Clara, USA
Mixer	Reax top	Heidolph, Schwabach
Kamera	MP 4+	Polaroid, Dreieich-Sprendlingen
Kühlschrank		Liebherr, Biberach an der Riss
Mikrowelle	Dimension 4	Panasonic, Hamburg
NanoDrop	ND-1000	PEQLAB Biotechn. GmbH, Erlangen
PCR Cycler	PTC 200 Peltier Thermal Cycler	MJ Research, USA
Picofuge		Stratagene, Amsterdam, Netherlands
Pipetten	2,5µl, 10µl, 20µl, 200µl, 1000µl	Eppendorf AG, Hamburg
Pipettierhilfe	Pipetus Akku	Hirschmann, Eberstadt
Power Supply	GPS 200/400	Pharmacia LKB, Freiburg
Reinheitsbank	LaminAir HBB243	Heraeus Instruments, Osterode
Scanner	GeneChip Scanner 3000 7G	Affymetrix, Inc., Santa Clara, USA
UV Tisch	UVT 2035	Herolab, Wiesloch
Vacuum Manifold	MultiScreenHTS	Millipore, Schwalbach/Ts
Waage	MXX 612	Denver Instrument, Göttingen
Waschstation	GeneChip Fluidics Station 450	Affymetrix, Inc., Santa Clara, USA
Zentrifuge	MiniSpin	Eppendorf AG, Hamburg
Zentrifuge	RC5C	Sorvall Instruments

Tabelle 2.3: Geräte, Modelle und Hersteller

2.2 Laboranalytische Methoden

Der erste Teil dieser Arbeit umfasste Techniken und Methoden im Laborbereich. Zur Untersuchung möglicher Korrelationen zwischen Genotyp und Phänotyp wurden für alle Tumorproben zwei Strategien verfolgt. Einerseits wurden Expressionsprofile durch den Einsatz von Expressionsarrays der Firma Affymetrix erstellt. Andererseits wurden zur Bestimmung von genetischen Variationen SNP-Arrays der Firma Affymetrix und CGH-Arrays der Firma Agilent eingesetzt. Zwecks Ergänzung der Strategie im Tumormodell zur Analyse von Determinanten der Metastasierung in Neuroblastomen, wurden hier zusätzlich Expressionsprofile auf Proteinebene im Bereich des MAPK Signaltransduktionsweges untersucht. Hierfür wurden Arrays der Firma Eppendorf eingesetzt.

Die RNA-Proben der in Abschnitt 1.2 beschriebenen Tumormodelle wurden bereits zuvor von den Kooperationspartnern isoliert und aufgereinigt. In den folgenden Schritten wurden diese Proben auf Qualität und Konzentration überprüft. Die Reinheit und Konzentration der isolierten RNA wurde sowohl photometrisch mit dem NanoDrop Photometer, siehe Abschnitt 2.2.2, als auch mittels Kapillargelelektrophorese auf dem Bioanalyzer 2100 von Agilent bestimmt, siehe Abschnitt 2.2.3. Das Probenmaterial wurde anschließend amplifiziert und für die Hybridisierung auf den Microarrays vorbereitet. Für die Expressionsarrays, SNP-Technologie und Proteinarrays wurden dafür jeweils unterschiedliche, zuvor festgelegte Protokolle verwendet.

Bei einigen Proben, z.B. bei der CGH-Analytik, wurden die laboranalytischen Schritte aus technischen Gründen von Kooperationspartnern in Amsterdam, Berlin und Oslo durchgeführt. Die histologische Einstufung der Tumore, sofern relevant, erfolgte durch Prof. Dr. Udo Schumacher, Institut für Anatomie, Universitätsklinikum Hamburg-Eppendorf. Eine Übersicht aller Proben der beiden Projekte sowie der verwendeten Array-Technik findet sich in Tabellen 3.6 und 3.1. In den folgenden Abschnitten werden die durchgeführten Schritte und die verwendeten Protokolle beschrieben.

2.2.1 Sterilisierung der Arbeitsmaterialien

Eppendorf-Tubes, Pipettenspitzen und allgemeine Materialien aus Polypropylen wurden bei 121°C und 2bar für 4T autoklaviert. Flüssige Substanzen, z.B. Puffer, wurden entsprechend bei einer Dauer von 6T behandelt. Temperaturempfindliche Lösungen wurden mit 0,2µm Filtern sterilisiert. Glasgeräte wurden bei trockener Hitze von 180°C für 16h sterilisiert.

2.2.2 Photometrische Ermittlung von RNA-Konzentrationen

Photometrische Bestimmungen von RNA-Konzentrationen erfolgten mit dem NanoDrop-Photometer von Agilent. Dabei absorbieren die einzelsträngigen Ribonukleinsäuren die UV-Strahlung im Bereich von $\lambda=260\text{nm}$ während DNA, Proteine oder aromatische Substanzen (z.B. Phenol) die Strahlung um $\lambda=280\text{nm}$ absorbieren. Da die Menge der absorbierten UV-Strahlung dem Gehalt der Probe direkt proportional ist und $\text{OD}=1$, bei geeichtem Gerät sowie $\lambda=260\text{nm}$, einer RNA-Konzentration von $40\text{ng}/\mu\text{l}$ entspricht, kann die RNA-Konzentration aus dem A_{260} -Wert gemäß Formel 1 ermittelt werden. Aus dem A_{260}/A_{280} Verhältnis kann andererseits ein Rückschluss auf die Reinheit der Probe getroffen werden. Dieses Verhältnis sollte idealerweise zwischen 1,8 und 2,1 liegen, ist jedoch vom pH-Wert und der Ionenstärke abhängig. Ist der Koeffizient kleiner als 1,8 ist die Präparation stark kontaminiert.

$$\text{Konzentration [ng}/\mu\text{l}] = E_{260} \times \text{Verduennung} \times 40$$

Gleichung 1: Berechnungsformel der RNA-Konzentration

2.2.3 Ermittlung von RNA-Qualität mittels Kapillargelelektrophorese

Bei der im BioAnalyzer 2100 von Agilent umgesetzten Kapillargelelektrophorese findet die Trennung der Ribonukleinsäuren in einem dünnen Kapillarrohr in einer Elektrolytlösung statt. Die Probenvolumina sind mit $1\mu\text{l}$ deutlich geringer als bei normalen Agarosegelen. Abgesehen von einem digital berechneten konventionellen Gelbild werden Elektropherogramme ausgegeben. Hier wird die Exktinktion über die Zeit gemessen als Graph dargestellt und somit die Wanderungsgeschwindigkeit der vorliegenden Moleküle wiedergegeben. Für normale nicht degradierte RNA werden zwei markante Peaks bei ca. 14s und 39s erwartet.

2.2. LABORANALYTISCHE METHODEN

2.2.4 Expressionsarrays

Für die Neuroblastome wurden humane Whole Genome Expressionsarrays von Affymetrix (Human Genome U133 Plus 2.0) verwendet. Bei den Brusttumoren wurden murine Whole Genome Arrays (Mouse Genome 430 2.0) eingesetzt. Die folgenden Schritte erfolgten gemäß dem One-Cycle-Protokoll¹ mit dem One-Cycle cDNA Synthesis Kit² der Firma Affymetrix .

Gesamt-RNA	RT-Volumen
5,0 – 8,0 µg	1.0 µl
8,1 – 16,0 µg	2.0 µl
16,1 – 20,0 µg	3.0 µl

Tabelle 2.4: Einzusetzendes Superscript II Volumen für Erststrangsynthese

cDNA Synthese Zur Aufbereitung der RNA für die Expressionsarrays von Affymetrix wurde eine cDNA-Synthese mit vorangehender photometrischer Bestimmung der RNA-Konzentration durchgeführt. Das A260/A280-Verhältnis sollte bei ungefähr 2,0 liegen. Es wurde ein HPLC gereinigter T7-(dT)24 Primer der folgenden Sequenz verwendet:



Je 8µg Gesamt-RNA wurden 1µl SuperScript II Reverse Transcriptase (200U/µl) eingesetzt, siehe Tabelle 2.4.

Erststrang cDNA Bei der Erststrangsynthese wurden Primer und RNA hybridisiert. Abhängig von der Konzentration wurden hierzu zwischen 5-20µg Gesamt-RNA mit 2µl 50µmol T7-(dT)24 Primer sowie RNase-freiem Wasser, das spätere Endvolumen sollte 20µl betragen, versetzt und für 10min bei 70°C inkubiert. Im Anschluß wurde der gesamte Ansatz auf 4°C abgekühlt und 4µl 5x Erststrang cDNA Puffer, 2µl 0,1M DTT und 1µl 10mM dNTP Mix hinzupipettiert. Danach erfolgte ein weiterer Inkubationsschritt für 2min bei 42°C mit anschließender Zugabe von 1µl/8µg RNA SuperScript II RT (200U/µl). Dieser Ansatz wurde 1h bei 42°C inkubiert.

¹Alternative Protocol for One-Cycle cDNA Synthesis from Total RNA, GeneChip Expression Analysis Technical Manual, 701021 Rev. 5, Affymetrix

²siehe Tabelle 2.1

Gesamt-RNA	cDNA-Volumen
1,0 – 8,0 µg	12 µl
8,1 – 15,0 µg	6 µl

Tabelle 2.5: Einzusetzendes cDNA-Volumen für die IVT

Zweitstrang cDNA Zur Zweitstrangsynthese wurde der Ansatz der Erststrangsynthese auf Eis gekühlt und danach kurz zentrifugiert. Nach Hinzugabe von 91µl RNase-freiem Wasser, 30µl 5x Zweitstrangreaktionspuffer, 3µl 10mM dNTP Mix (jeweils 200µM), 1µl 10U/µl E.coli DNA Ligase, 4µl 10U/µl E.coli DNA Polymerase I und 1µl 2U/µl E.coli RNase H wurde der Gesamtansatz im Eppendorfreagenzgefäß gemischt, kurz zentrifugiert und für 2h bei 16°C inkubiert. Anschließend wurden 2µl 10U T4 DNA Polymerase hinzugegeben und weitere 5min bei gleicher Temperatur gekühlt. Zum Stoppen der Reaktion wurden 10µl 0,5M EDTA hinzugegeben.

Aufreinigung Die Aufreinigung erfolgte mit dem GeneChip Sample Cleanup Module³ und den cDNA Cleanup Spin Columns nach folgendem Protokoll.

Zum Volumen der cDNA-Synthese wurden 600µl Binding Buffer gegeben. 500µl der Probe wurden auf die cDNA Cleanup Spin Columns transferiert und bei 8000g für 1min zentrifugiert. Der Durchfluss wurde verworfen, die restliche Probe auf die Säule pipettiert, wiederum 1min bei 8000g zentrifugiert und der Durchfluss erneut verworfen. Danach wurde die Säule mit 750µl cDNA Waschpuffer versetzt. Nach weiterer Zentrifugation bei 8000g für 1min wurde der Durchfluss abermals verworfen und die Probe für 5min bei maximaler Geschwindigkeit (12.000g) zentrifugiert. Die Säule wurde in ein frisches 1,5ml Tube transferiert und die cDNA unter Zugabe von 14µl cDNA Elution Buffer, 1min bei maximaler Geschwindigkeit eluiert. Für ein Kontrollgel wurde 1µl der cDNA entnommen.

Synthese der Biotin-markierten cRNA Zur Synthese von Biotin-markierter Gegenstrang-cRNA aus der hergestellten cDNA wurde eine *in vitro* Transkription mit dem IVT Labeling Kit⁴ durchgeführt. Da mindestens 1µg cDNA benötigt wurde richtete sich das eingesetzte Volumen nach der RNA-Ausgangsmenge und wurde aus Tabelle 2.5 ermittelt.

³Affymetrix, Inc., Santa Clara, USA

⁴siehe Tabelle 2.1

2.2. LABORANALYTISCHE METHODEN

***in vitro* Transkription** Das aus Tabelle 2.5 ermittelte cDNA-Volumen wurde in ein steriles Eppendorfreagenzgefäß pipettiert und 4µl 10x IVT Labeling Buffer, 12µl IVT Labeling NTP Mix sowie 4µl IVT Labeling Enzyme Mix hinzugegeben. Der Ansatz wurde mit RNase-freiem Wasser auf ein Endvolumen von 20µl aufgefüllt, vorsichtig gemischt und für 16h bei 37°C inkubiert. Im Anschluß wurde ein 1µl Aliquot der Probe zur Qualitätskontrolle entnommen.

Aufreinigung Die Aufreinigung erfolgte mit dem GeneChip Sample Cleanup Module und den cRNA Cleanup Spin Columns nach folgendem Protokoll.

Zum Volumen der cRNA-Synthese wurden 60µl RNase-freies Wasser, 350µl IVT cRNA Binding Buffer und 250µl 100%iges Ethanol gegeben. Das Gesamtvolumen der Probe wurde auf die cRNA Cleanup Spin Column transferiert und bei 8000g für 15s zentrifugiert. Anschließend wurde der Durchfluss erneut auf die Säule pipettiert und bei 8000g für 1min zentrifugiert. Der Durchfluss wurde verworfen und die Säule mit 500µl cRNA Waschpuffer versetzt. Nach erneuter Zentrifugation bei 8000g für 1min wurde der Durchfluss abermals verworfen, die Probe mit 500µl 80%igem Ethanol gewaschen, 15s bei 8000g und anschließend nochmals für 5min bei maximaler Geschwindigkeit (12.100g) zentrifugiert. Die Säule wurde in ein frisches 1,5ml Tube gesteckt und die cRNA einmal mit 11µl und einmal mit 10µl RNase-freiem Wasser 1min bei maximaler Geschwindigkeit (12.000g) eluiert.

Ermittlung RNA-Konzentration Die RNA-Konzentration des Eluats wurde photometrisch ermittelt, siehe Abschnitt 2.2.2.

Quantifizierung des IVT-Produktes Da das Eluat anteilig noch Gesamt-RNA enthält musste vor der Fragmentierung der wahre Konzentrationsanteil der cRNA berechnet werden. Eine entsprechende Berechnung wurde mittels Formel 2 durchgeführt. Für die Fragmentierung war eine minimale korrigierte cRNA-Konzentration von 0,6µg/µl notwendig.

$$c_{cRNAkorr} = \left(\frac{n_{cRNAkorr}}{V_{cRNAelu}} \right)$$

mit

$$n_{cRNAkorr} = n_{cRNAges} - n_{totalRNA} * \left(\frac{V_{cDNAivt}}{V_{cDNAelu}} \right)$$

- $c_{cRNAkorr}$ = korrigierte cRNA-Konzentration [$\mu\text{g}/\mu\text{l}$]
- $n_{cRNAkorr}$ = korrigierte cRNA-Menge [μg]
- $V_{cRNAelu}$ = Eluationsvolumen nach IVT [μl]
- $n_{cRNAges}$ = cRNA-Menge gemessen nach der IVT [μg]
- $n_{totalRNA}$ = total RNA-Startmenge für die cDNA-Synthese [μg]
- $V_{cDNAivt}$ = Einsatzvolumen für die IVT [μl]
- $V_{cDNAelu}$ = Eluationsvolumen nach der cDNA-Synthese [μl]

Gleichung 2: Berechnungsformel der korrigierten cRNA-Menge

Fragmentierung der cRNA und Produktkontrolle Für die Fragmentierung war eine minimale korrigierte cRNA-Konzentration gemäß Formel 2 von $0,6\mu\text{g}/\mu\text{l}$ notwendig. Dem entsprechenden RNA-Volumen wurden $8\mu\text{l}$ Fragmentierungspuffer zugegeben und die gesamte Probe mit RNase-freiem Wasser auf $40\mu\text{l}$ Endvolumen gebracht. Anschließend wurde die Probe bei 94°C für 35min inkubiert. Der Puffer zerkleinert die cRNA während der Inkubation durch metallinduzierte Hydrolyse in 35-200 Basen große Fragmente. Im Anschluß wurde die fragmentierte Probe auf 4°C heruntergekühlt. Zur späteren Gelkontrolle wurde erneut ein Aliquot von $1\mu\text{l}$ abgenommen.

Der Grad der Fragmentierung, die Größenverteilung der Fragmente und somit Degradationsgrad und Qualität des IVT-Produktes wurden per Kappilargelelektrophorese auf dem Bioanalyzer von Agilent bestimmt. Hier wird die Fluoreszenzstärke in Abhängigkeit von der Wanderungsgeschwindigkeit gemessen und als Verteilungsfunktion wiedergegeben.

Hybridisierung der Ziel-RNA Die Zusammensetzung der für die Hybridisierung der Ziel-RNA verwendeten Lösung kann Tabelle 2.6 entnommen werden.

Aufgrund der von Affymetrix verwendeten photolithographischen Wafer-Technologie bei der Herstellung ihrer Chips ist die Gitterstruktur der Spots in sich sehr homogen und

2.2. LABORANALYTISCHE METHODEN

konsistent. Dennoch sind minimale Ausrichtungsfehler sowie eine Verschiebung des Gesamtgitters nicht auszuschließen. Die hier verwendeten B2-Kontrolloligonukleotide dienen daher als Marker unterschiedlicher Intensität, anhand derer beim späteren Scannen die genaue Lage des Lesegitters bestimmt werden kann.

Bei den Hybridisierungskontrollen *bioB*, *bioD* und *bioC* handelt es sich um Gene des Biotinsyntheseweges von *Escherichia coli*; *cre* ist das Rekombinase-Gen von P1 Bakteriophagen. Diese Gene werden als positive Hybridisierungskontrollen für Eukaryotische Chips eingesetzt, da die Gefahr von Kreuzhybridisierungen mit Eukaryontengen sehr niedrig ist. Anhand der verwendeten unterschiedlichen Endkonzentrationen⁵ lässt sich die Qualität der Hybridisierung leicht abschätzen.

Die Standardvolumina der eingesetzten Komponenten ergaben sich aus der benötigten Menge von 15µg fragmentierter cRNA. Somit wurden 5µl Kontrolloligonukleotid B2, 15µl 20-facher, für 5min bei 65°C vortemperierter, Eukaryotic-Hybridisierungscocktail, 3µl Heringssperma, 3µl acetyliertes BSA, 150µl 2x Hybridisierungspuffer und 30µl DMSO eingesetzt. Das Endvolumen wurde mit RNase-freiem Wasser auf 300µl aufgefüllt.

Vor der Hybridisierung der Arrays erfolgte ein Vorinkubationsschritt mit 200µl 1x Hybridisierungspuffer für 10min bei 45°C im Rotationsofen. In der Zwischenzeit wurde der Hybridisierungsansatz für 5min auf 99°C erwärmt mit anschließenden 5min bei 45°C. Vor dem Austausch des Hybridisierungspuffers gegen den Hybridisierungsansatz in der Chipkammer wurde der Hybridisierungscocktail für 5min bei maximaler Geschwindigkeit (12.100g) zentrifugiert. Die endgültige Hybridisierung fand anschließend für 16h bei 45°C und 60 rpm im Rotationsofen statt.

⁵bioB = 1,5pM; bioD = 5pM; bioC = 25pM und cre = 100pM

Reagenz	Menge / Volumen
Fragmentierte cRNA	15µg
Kontroll-Oligonukleotide B2 (3 nM)	5µl
20x Eukaryotic Hybridisierungskontroll-Cocktail mit Biotin markierten RNAs bioB, bioC, bioD und cre	15µl
Hering Sperm DNA (10 mg/ml)	3µl
Acetyliertes BSA (50 mg/ml)	3µl
2x Hybridisierungspuffer	150µl
DMSO	30µl
RNase freies Wasser	auf 300µl aufgefüllt

Tabelle 2.6: Zusammensetzung des Hybridisierungscocktails pro Array für die Expressionsarrays von Affymetrix

Waschen, Färben, Scannen Nach 16h im Rotationsofen ist der Hybridisierungsvorgang beendet. Der Hybridisierungscocktail wurde durch Abpipettieren aus der Chipkammer entfernt und der Array mit 200µl Non-Stringent Waschpuffer befüllt. Für die folgende Färbung wurden zwei Lösungen benötigt, ein Streptavidin-Phycoerythrin-(SAPE)-Mix sowie eine Antikörperlösung. Aufgrund der hohen Lichtempfindlichkeit von SAPE, sollte die Lagerung im Dunkeln bei 4°C stattfinden. Die Zusammensetzung des SAPE-Färbemix ist in Tabelle 2.7 angegeben. Der Mix wurde auf zwei Aliquots zu je 600µl für zwei unterschiedliche Färbeschritte aufgeteilt. Die Zusammensetzung der Waschlösungen und des Antikörper-Mix kann Tabelle 2.7 entnommen werden.

Die Wasch- und Färbeschritte wurden halbautomatisch in der Fluidics-Station 450 (Affymetrix, Inc., Santa Clara, USA) und die einzelnen Schritte gemäß dem One-Cycle-Protokoll des Herstellers für Standardarrays [4] durchgeführt. Nach Abschluss aller Wasch- und Färbeschritte wurden die Arrays mit dem GeneChip Scanner 3000 7G (Affymetrix, Inc., Santa Clara, USA) eingelesen. Die Bedienung des Gerätes erfolgte über die Affymetrix GeneChip Operating Software⁶. Beim Scanvorgang verwendete Wellenlänge betrug 570nm, die Pixelgröße 1,56µm .

⁶Version 1.4

2.2. LABORANALYTISCHE METHODEN

Reagenz	Volumen
<i>2x Stain Buffer</i>	
MES Stock Buffer (12x)	41,7ml
NaCl (5M)	92,5ml
Tween-20 (10%)	2,5ml
Aqua ad iniectabilia	113,3ml
<i>SAPe Solution Mix</i>	
2x MES Färbepuffer	600µl
50mg/ml Acetyliertes BSA	48µl
1mg/ml Streptavidin Phycoerythrin	12µl
Aqua ad iniectabilia	540µl
<i>Antibody Solution Mix</i>	
2x MES Färbepuffer	300µl
50mg/ml acetyliertes BSA	24µl
10mg/ml Normal Ziegen IgG	6µl
0,5mg/ml biotinierte Antikörper	3,6µl
Aqua ad iniectabilia	266,4µl
<i>Array Holding Buffer</i>	
MES Stock (12x)	8,3 ml
NaCl (5M)	18,5 ml
Tween-20 (10%)	0,1 ml
Aqua ad iniectabilia	73,1 ml
<i>MES Stock (12x)</i>	
MES hydrate	74,4 g
MES Sodium Salt	193,3 g
Aqua ad iniectabilia	1000 ml
<i>Wash A: Non-Stringent Wash Buffer</i>	
SSPE (20x)	300 ml
Tween-20 (10%)	1 ml
Aqua ad iniectabilia	699 ml
<i>Wash B: Stringent Wash Buffer</i>	
MES Stock (12x)	83,3ml
NaCl (5M)	5,2ml
Tween-20 (10%)	1,0ml
Aqua ad iniectabilia	910,5 ml

Tabelle 2.7: Zusammensetzung der Wasch und Färbelösungen für die Expression von Affymetrix. Die Volumina der ersten vier Lösungen sind als Volumen pro Array angegeben. Array Holding Buffer, MES Stock-Lösung sowie die Waschlösungen wurden jeweils steril filtriert.

2.2.5 SNP Arrays

DNA-Isolierung Für die DNA Isolierung wurde ein Mastermix aus 4,5ml 1x Lysispuffer, 4,5ml Aqua dest sowie 1ml Proteinase K mit einer Konzentration von 10mg/ml vorbereitet. Die Zusammensetzung des Lysepuffer kann Tabelle 2.8 entnommen werden. Die Zellkulturpellets bzw. Gewebeproben wurden mit 500µl Mastermix/Probe auf dem Schüttelblock bei 55°C für ca. 40h verdaut. Anschließend wurde die DNA mit 500µl Phenol-Chloroform-Isoamylalkohol 25:24:1 für 2min bei 13000rpm extrahiert und der Überstand mit 1000µl 100%igem Ethanol (5min bei 13000rpm) gefällt. Das Pellet wurde mit 500µl 70%igem Ethanol gewaschen und 5min bei 13000rpm zentrifugiert. Der Überstand wurde verworfen, das Pellet luftgetrocknet und in 50µl TE-Puffer aufgenommen. Die Bestimmung der Reinheit erfolgte auf einem 0,6%igem Agarosegel, die Konzentrationsbestimmung auf dem NanoDrop Photometer.

Restriktionsenzymverdau Die aufgereinigte DNA wurde in zwei Ansätzen einem Restriktionsenzymverdau mit Sty I bzw. Nsp I unterzogen. Eingesetzt wurden 5µl genomische DNA bei einer Konzentration von 50ng/µl. Die Zusammensetzung des jeweiligen Mastermix ist Tabelle 2.9 zu entnehmen. Beide Ansätze wurden erst 2h bei 37°C sowie 20min bei 65°C inkubiert und anschließend auf 4°C abgekühlt.

Ligation Für die verdauten Proben wurde jeweils ein restriktionsenzymabhängiger Ligationvorgang durchgeführt. Den Sty I verdauten Proben wurden 2,5µl T4 DNA Ligase Puffer DNA, 0,75µl Adaptor Sty I sowie 2µl T4 DNA Ligase zugegeben. Die Nsp I verdauten Proben wurden mit 2,5µl T4 DNA Ligase Puffer DNA, 0,75µl Adaptor Nsp I sowie

Reagenz	Konzentration	Menge
Urea Powder	8M	240g
EDTA (0,5M , pH 8,0)	20mM	20ml
N-Laurosarkosine, 30%	1%	16,5ml
Tris-Cl (1M , pH 8,0)	0,2M	100ml
NaCl (5M)	0,4M	40ml
Gesamtvolumen		0,5l

Tabelle 2.8: Zusammensetzung des 2x Lysepuffer, für 0,5 l Gesamtvolumen, zur DNA Isolierung für die SNP-Arrays. Für die Isolation wurde der Puffer mit Aqua Dest. im Verhältnis 1:1 verdünnt.

2.2. LABORANALYTISCHE METHODEN

Reagenz	Volumen/Probe Sty I	Volumen/Probe Nsp I
Aqua dest	11.55 µl	11.55 µl
NE Puffer 2	—	2µl
NE Puffer 3	2µl	—
BSA	0,2µl	0,2µl
Sty I	1µl	—
Nsp I	—	1µl
Gesamtvolumen	14,75 µl	14.75 µl

Tabelle 2.9: Zusammensetzung Mastermix für Restriktionsenzymverdau für die SNP-Arrays

2µl T4 DNA Ligase versetzt. Die Ansätze wurden jeweils 3h bei 16°C sowie im Anschluß 20min bei 70°C inkubiert und danach auf 4°C gekühlt. Danach wurden die Proben mit 75µl Wasser verdünnt.

PCR Amplifikation Zur Amplifikation der DNA wurden drei PCR Reaktionen (Sty I) bzw. vier Reaktionen (Nsp I) je ligierte Probe für eine PCR Amplifikation durchgeführt. Die Zusammensetzung des PCR Mastermix kann Tabelle 2.11 entnommen werden. Die jeweiligen Ansätze wurden im Thermocycler gemäß dem in Tabelle 2.10 aufgeführten Ablauf inkubiert.

Aufreinigung des PCR-Produkts Für die nachfolgenden Schritte wurden jeweils alle PCR-Ansätze einer Probe gepoolt und mit 1ml Magnetic Beads versetzt. Nach einer Inkubationszeit von 10min bei Raumtemperatur wurden die Proben auf eine Filterplatte überführt und die Platte versiegelt. Unter Anlage von Vakuum wurden die Beads vom Puffer abgetrennt. Die Proben wurden einmal mit 75%igem Ethanol gewaschen und das Vakuum aufrecht erhalten bis alle Wells trocken gefallen waren. Zum Eluieren wurden die einzelnen Wells mit 55µl Buffer EB versetzt und unter Schütteln für 10min inkubiert. Die Filterplatte wurde auf eine Sammelplatte übertragen und unter erneutem Anlegen von Vakuum filtriert.

Fragmentierung Das Fragmentation Reagent musste zu Beginn auf eine Konzentration von 0,1U/µl verdünnt werden. Die Zusammensetzung des verdünnten Reagenz bestimmte sich dabei aus Tabelle 2.12. Jeder Probe wurden 5µl 10x Fragmentation Buffer und 5µl verdünntes Fragmentation Reagent hinzugegeben. Der Ansatz wurde im Thermocycler

Temperatur	Zeit	Zyklen
94°C	3 min	1
94°C	30 s	} 30
60°C	30 s	
68°C	15 s	
68°C	7 min	1

Tabelle 2.10: PCR Programm für den Thermocycler

erst für 35min bei 37°C inkubiert, anschließend 15min bei 95°C und wurde zuletzt auf 4°C heruntergekühlt. Je 1,5µl jeder fragmentierten Probe wurde zur Qualitätskontrolle der Fragmentierung auf ein 2%iges Agarosegel geladen.

Labeling und Hybridisierung Vor der Hybridisierung wurden die fragmentierten Proben Biotin-markiert. Jede Probe wurde dafür mit 14µl 5x TdT-Buffer, 2µl DNA Labeling Reagent und 3,5µl TdT-Enzyme versetzt. Der Ansatz wurde im Thermocycler erst für 4h bei 37°C inkubiert, anschließend für 15min bei 95°C und zuletzt auf 4°C heruntergekühlt.

Die gelabelten Proben wurden mit 190µl Hybridisierungscocktail, siehe Tabelle 2.13, versetzt und 10min bei 95°C denaturiert. Auf 50°C abgekühlt wurden 200µl jeder Probe auf einen Microarray geladen und unmittelbar darauf im auf 50°C vorgeheizten Hybridisierofen für 16h bei 50°C und 60rpm hybridisiert.

Waschen, Färben, Scannen Mit vollendeter Hybridisierung wurde der Hybridisierungscocktail aus dem Array entfernt und durch 270µl Array Holding Buffer ersetzt. Die Färbung erfolgte gemäß dem Herstellerprotokoll für Genome Wide SNP 6.0 Arrays halb-automatisch in der Fluidics-Station 450 von Affymetrix . Vier Färbe- und Waschlösungen wurden dabei eingesetzt. Die Zusammensetzungen der vier dabei eingesetzten Lösungen (Stain Buffer, Streptavidin Phycoerythrin-(SAPE) Solution Mix, Antibody Solution Mix und Array Holding Buffer) sind in Tabelle 2.14 angegeben. Da SAPE lichtempfindlich ist, erfolgte die Lagerung im Dunkeln bei 4°C. Nach Abschluss der Waschschritte und Färbungen wurde der Array mit dem GeneChip Scanner 3000 7G von Affymetrix bei einer Wellenlänge von 570nm und einer Pixelgröße von 1,56µm ausgelesen.

2.2. LABORANALYTISCHE METHODEN

Reagenz	Volumen je Ansatz
Aqua dest	39,5 µl
Titanium Taq PCR Puffer	10 µl
GC-Melt	20 µl
dNTP	14 µl
PCR Primer 002	4,5 µl
Titanium Taq DNA Polymerase	2 µl
Gesamtvolumen	90 µl

Tabelle 2.11: Zusammensetzung des PCR Mastermix für den Amplifikationsschritt der ligierten DNA für die SNP Arrays.

Reagenz	Urspr. Konzentration d. Fragmentation Reagent				
	2 U/µl	2,25 U/µl	2,5 U/µl	2,75 U/µl	3 U/µl
Wasser	306 µl	308 µl	309,6 µl	310,9 µl	312 µl
10x Fragmentation Buffer	36 µl	36 µl	36 µl	36 µl	36 µl
Fragmentation Reagent	18 µl	16 µl	14,4 µl	13,1 µl	12 µl
Gesamtvolumen	360 µl	360 µl	360 µl	360 µl	360 µl

Tabelle 2.12: Zusammensetzung der Verdünnung des Fragmentation Reagent abhängig von der ursprünglichen Konzentration.

Reagenz	Volumen pro Array
MES (12x; 1,25 M)	12 µl
Denhardt's Solution (50x)	13 µl
EDTA (0,5 M)	3 µl
Herring Sperm DNA (10 mg/ml)	3 µl
Oligo Control Reagent	2 µl
Human Cot-1 DNA (1 mg/ml)	3 µl
Tween-20 (3%)	1 µl
DMSO (100%)	13 µl
TMACL (5 M)	140 µl
Gesamtvolumen	190 µl

Tabelle 2.13: Zusammensetzung des Hybridisierungscocktails pro Array für die SNP 6.0 Arrays von Affymetrix

KAPITEL 2. MATERIAL UND METHODEN

Reagenz	Volumen
<i>Stain Buffer</i>	
SSPE (20x)	360 µl
Tween-20 (3%)	3,96 µl
Denhardt's Solution (50x)	24 µl
Aqua ad iniectabilia	800,04 µl
<i>SAPe Solution Mix</i>	
Stain Buffer	594 µl
Streptavidin Phycoerythrin	6 µl
<i>Antibody Solution Mix</i>	
Stain Buffer	594 µl
Biotinylated Antibody	6 µl
<i>Array Holding Buffer</i>	
MES Stock (12x)	8,3 ml
NaCl (5M)	18,5 ml
Tween-20 (10%)	0,1 ml
Aqua ad iniectabilia	73,1 ml
<i>MES Stock (12x)</i>	
MES hydrate	74,4 g
MES Sodium Salt	193,3 g
Aqua ad iniectabilia	1000 ml
<i>Wash A: Non-Stringent Wash Buffer</i>	
SSPE (20x)	300 ml
Tween-20 (10%)	1 ml
Aqua ad iniectabilia	699 ml
<i>Wash B: Stringent Wash Buffer</i>	
SSPE (20x)	30 ml
Tween-20 (10%)	1 ml
Aqua ad iniectabilia	969 ml

Tabelle 2.14: Zusammensetzung der Wasch und Färbelösungen für die SNP 6.0 Arrays von Affymetrix. Die Volumina der ersten vier Lösungen sind als Volumen pro Array angegeben. Array Holding Buffer, MES Stock-Lösung sowie die Waschlösungen wurden jeweils steril filtriert.

2.2. LABORANALYTISCHE METHODEN

2.2.6 MAPK Transkriptionsfaktor-Arrays

Ausgehend von Beobachtungen in den vorangehenden Versuchen den MAPK-Signaltransduktionsweg betreffend wurden Proteinarrays der Firma Eppendorf zur Analyse von 8 MAPK-involvierten Transkriptionsfaktoren eingesetzt. In je 5 konfluenten 175cm² Zellkulturflaschen wurden die drei Zelllinien kultiviert. Im Anschluß wurden die Proteine mit dem *Nuclear Extract Kit* der Firma Active Motif extrahiert und auf dem MAPK-TF-Chip hybridisiert. Die Hybridisierung und Färbung erfolgte gemäß dem *TF Chip MAPK Kit* Protokoll von Eppendorf, die verwendeten Reagenzien entstammten ebenfalls dem Kit.

Isolierung Die Extraktion aus der Zellkultur erfolgte gemäß dem *Nuclear Extract Kit* Protokoll der Firma Active Motif. Die Zusammensetzung der verwendeten Lösungen kann Tabelle 2.15 entnommen werden. Die Zellen wurden mit 10ml eisgekühlter PBS/Phosphatase Inhibitorlösung gewaschen, die Lösung aspiriert und nochmals 6ml Inhibitorlösung hinzugegeben. Die Zellen wurden in 15ml Falcontubes übertragen und die Suspension für 5min bei 500rpm und 4°C abzentrifugiert. Der Überstand wurde verworfen.

Zur Isolierung der zytoplasmatischen Fraktion wurden die Zellpellets in 500µl 1x Hypotonic Buffer resuspendiert. Nach Inkubation der Suspension für 15min in microcentrifuge tubes wurden 25µl Detergenz hinzupipettiert und 10min lang gevortext. Danach erfolgte ein

Reagenz	Volumen
<i>PBS/Phosphatase Inhibitor</i>	
10x PBS	1,6ml
Aqua dest	13,6ml
Phosphatase Inhibitor	0,8ml
<i>1x Hypotonic Buffer</i>	
10x Hypotonic Buffer	100,0µl
Aqua dest	0,9ml
<i>Complete Lysis Buffer</i>	
10 mM DDT	10,0µl
Lysis Buffer AM1	98,0µl
Protease Inhibitor Cocktail	1,0µl

Tabelle 2.15: Übersicht der Zusammensetzung bei der Proteinisolierung verwendeter Lösungen. Die aus den Neuroblastomzellkulturen isolierten Proteine wurden für die MAPK-TF-Chips der Firma Eppendorf verwendet. Die Volumina sind pro Zellkulturflasche angegeben.

Zentrifugationsschritt für 30s bei 14.000g und 4°C. Der Überstand (Zytoplasmatische Fraktion) wurde für spätere Applikationen entnommen, die Pellets zur Extraktion der nuklearen Fraktion in 50µl Complete Lysis Buffer resuspendiert. Im Anschluß an eine 30minütige Inkubation auf Eis im Schüttler bei 150rpm wurde die Lösung 30 Sekunden lang gevortext und danach 10 Minuten lang bei 14.000g und 4°C zentrifugiert. Der Überstand (nukleare Fraktion) wurde entnommen und die Proteinkonzentration mittels Bradford Assay bestimmt.

Silberfärbung Um bereits vor der Hybridisierung die Proteinextrakte auf Anreicherungen spezieller Proteinfractionen zu untersuchen wurde eine Silberfärbung mit dem *PlusOne Silver Staining Kit* der Firma Amersham Biosciences auf einem Bis-Tris Gradientengel (4-12%) durchgeführt. Aufgetragen wurden pro Probe je 100ng, 200ng und 400ng.

Hybridisierung Vor der Hybridisierung der Samples auf den Arrays erfolgte eine Blockierung der Arraytranskripte mit 100µl Blocking Lösung 2 für 1h bei Raumtemperatur. Nach dem Waschen mit der TF Washing Lösung wurden 30µg nuklearer Extrakt (100µl Samplecocktail) in die Arraykammern pipettiert und die Arrayslices im Thermomixer über Nacht 16h lang bei Raumtemperatur und 600rpm inkubiert. Die Volumina der nuklearen Extrakte für die Samplelösung ergaben sich aus den zuvor gemessenen Proteinkonzentrationen. Die Zusammensetzungen der verwendeten Lösungen, sofern nicht anwendungsbereit im Kit enthalten, sind in Tabelle 2.16 angegeben.

Waschen, Färben, Scannen Nach der Inkubation wurden die Samples aspiriert. Die Arrays wurden 3fach für je eine Minute mit 100µl TF Washing Lösung gewaschen. Anschließend wurden die Arraykammern mit 100µl der Primary Antibody Lösung befüllt und 1h lang inkubiert. Die Primary Antibody Lösung wurde aspiriert und die Kammern mit 100µl TF Washing Lösung gewaschen.

Die im Kit mitgelieferte Silverquant Box, eine verschließbare Hybridisierungskammer für Objektträger, wurde mit 15ml TF Washing Lösung befüllt. Die Hybridisierungsrahmen der Arrays wurden von den Objektträgern entfernt, die Slides in die Kammer eingebettet und Silverquant Box geschlossen. Nach einer 1 minütigen Waschung erfolgte ein Austausch der Waschlösung gegen 15ml Secondary Antibody Lösung und eine Inkubation bei Raumtemperatur für 10min. Im Anschluß wurden die Slides erst 3fach je für 1min mit SQ Wash Buffer gewaschen und danach in 15ml verdünnter Gold-Konjugat-Lösung für 45min bei Raumtemperatur und 300rpm im Thermomixer inkubiert.

2.2. LABORANALYTISCHE METHODEN

Reagenz	Volumen/Menge
<i>TF Washing Solution</i>	
Washing Buffer A 5ml	
Aqua dest	95ml
<i>Samplecocktail</i>	
Nuklearer Extrakt	30µg
Sample Solution	auf 40µl aufgefüllt
Binding Solution	60µl
<i>Primary Antibody Solution</i>	
Primary Antibody Cocktail	4µl
Blocking Solution 1	600µl
<i>Secondary Antibody Solution</i>	
Secondary Biotinylated Antibody Mix	6µl
Blocking Solution 1	15ml
<i>SQ Wash Buffer</i>	
UniBuffer	5ml
Tween 20	200µl
Aqua dest	195ml
<i>Pre-Blocking Buffer</i>	
UniBuffer	ml
Aqua dest 6 15ml	
<i>Verdünnte Gold-Konjugat-Lösung</i>	
Pre-Blocking Buffer	2ml
Aqua dest	18ml
Anti-Biotin-Gold Konjugat	40µl
<i>Rinsing Buffer</i>	
Silverquant Buffer	2ml
Aqua dest 18ml	

Tabelle 2.16: MAPK-TF-Chips: Zusammensetzung der verwendeten Lösung für die MAPK Transkriptionsfaktor Arrays der Firma Eppendorf. Die hier verwendeten Reagenzien entstammen dem TF Chip MAPK Kit.

KAPITEL 2. MATERIAL UND METHODEN

Erneut erfolgte ein 4facher Waschschrift für je 1min mit SQ Wash Buffer, gefolgt von einer Waschung für 1min mit Rinsing Buffer. Die Silverquant Box wurde mit 7,5ml Silverquant Solution B befüllt und die Slides 5min lang bei Raumtemperatur und 300rpm inkubiert. Zuletzt wurden die Arrays zweifach für je 30s mit destilliertem Wasser gewaschen und 5min lang bei 600g zentrifugiert.

Das Scannen erfolgte mit dem Eppendorf Silverquant Scanner und die Detektion sowie Auswertung mit der Silverquant Analysis Software.

2.3 Programme und statistische Methoden

Ausgehend von einer vollendeten Hybridisierung der Samples auf den Microarrays war eine Vielzahl von technischen und bioinformatischen Schritten notwendig, um aussagekräftige Resultate in einen biologischen Rahmen zu setzen. Diese Schritte können in drei Stufen unterteilt werden, die Generation semi-quantitativer Hybridisierungssignale, die statistische Auswertung der Signale und schließlich die biologische Interpretation der Daten. Die Generation der Hybridisierungssignale umfasste das technische Scannen der Chips, eine Umwandlung der Bilddateien in numerische Werte sowie anschließende Korrekturen der erhaltenen Signalwerte. Bei der bioinformatischen Auswertung kamen diverse Visualisierungstechniken, Testverfahren, Annotationen und Korrelationsverfahren zum Einsatz. Diese sollten die potentiellen Unterschiede und Zusammenhänge der untersuchten Modelle aufzeigen. Während der biologischen Auswertung wurden schließlich die Ergebnisse der vorherigen Schritte in einen biologischen Rahmen gesetzt. Dabei sollte ein Modell gefunden werden, daß die Beobachtungen beschreibt und mögliche Ansätze zur Systemkontrolle in Aussicht stellt. In den folgenden Abschnitten werden die durchgeführten Schritte und die dabei verwendeten Programme und Methoden beschrieben. Bei Arrays von Affymetrix wird üblicherweise zwischen *Proben*, einzelne 25 basenpaarlange Oligonucleotide, und *Probesets*, aus mehreren *Proben* bestehende, genspezifische Transkripte, unterschieden. Zur Vermeidung von Irritationen werden *Tumorproben* daher zukünftig in den bioinformatischen Abschnitten mit *Sample* referenziert.

2.3.1 Programme und Umgebungen

Der Großteil der bioinformatischen Arbeit erstreckte sich über drei Plattformen. Die ersten Schritte erfolgten mit der GeneChip Operating Software. Die nachfolgenden Schritte wurden in der Programmierumgebung R und unter Einsatz von Online-Ressourcen durchgeführt.

GeneChip Operating Software Die Software zur Steuerung der Waschstation und des Scanners von Affymetrix sowie zur primären Analyse der Ergebnisse trägt den Namen GeneChip Operating Software. Sie ist zur Zeit in der Version 1.4 auf der Affymetrix Website erhältlich. Im Rahmen dieser Arbeit wurde die Software zur Bedienung der Waschstation und des Scanners, sowie zur Generierung von CEL-Dateien verwendet. Alle weitergehenden Analyseschritte wurden mit externen Programmen durchgeführt.

Programmiersprache R R ist eine freie open source Sprache für statistische Analysen und Visualisierungen. Gleichzeitig stellt R eine Entwicklerumgebung bereit, in der Berechnungen interaktiv durchgeführt werden können. Aufgrund des modularen Aufbaus des Systems können Funktionen, Methoden und Datensätze in Form von sogenannten *Packages* jederzeit bereitgestellt und nachträglich installiert werden. Wegen der hohen Flexibilität und den guten Schnittstellen zu anderen Sprachen erfreut sich die Sprache großer Beliebtheit. Im Bereich der Microarraybasierten Forschung hat sich eine Entwickler- und Forschungsgemeinde etabliert, die laufend neue Verfahren und Techniken in Form von R-Packages veröffentlicht.

Im Rahmen der statistischen Auswertung in dieser Arbeit ist die Programmiersprache R in der Version 2.6.2 zum Einsatz gekommen. Zur Zeit ist Version 2.7.2 unter <http://www.r-project.org> verfügbar.

Bioconductor Bei Bioconductor [62] handelt es sich um ein freies open source Software Projekt auf der Basis der Programmiersprache R. Die Module dienen speziell der Analyse und Verarbeitung von genomischen Daten. Ein besonderer Schwerpunkt liegt dabei auf der Verarbeitung von Microarray-Experimenten. Der Zugriff auf die öffentlichen Packages kann jederzeit online⁷ erfolgen. Anpassungen der Module an eigene Bedürfnisse sind meist problemlos möglich. Im Rahmen dieser Arbeit sind diverse Module des Bioconductor Projektes eingesetzt worden. Eine Auflistung aller verwendeter Pakete und ihrer Versionen ist in Tabelle 2.17 zu finden. In den folgenden Abschnitten werden die jeweils verwendeten Packages explizit ausgewiesen.

GEO *The Gene Expression Omnibus*[19, 54] ist eine Online-Datenbank zur Speicherung sowie Bereitstellung der Datensätze und Ergebnisse von Microarrayexperimenten. Die Datenbank wird von NCBI betrieben und ist frei zugänglich. Publikationen von Microarray-ergebnissen in einschlägigen Fachzeitschriften erfordern mittlerweile eine Publikation der zugrunde liegenden Datensätze in einer der anerkannten Datenbanken, wie *GEO*⁸ oder der vom EBI betriebenen *ArrayExpress*⁹ Datenbank[124].

⁷Bioconductor: <http://www.bioconductor.org>

⁸<http://www.ncbi.nlm.nih.gov/geo/>

⁹<http://www.ebi.ac.uk/microarray-as/ae/>

2.3.2 Scannen der Chips und Bildtransformation

Die hybridisierten Affymetrix -Chips wurden mit dem GeneChip Scanner 3000 7G von Affymetrix gescannt. Die Ausgabe besteht aus einer Bilddatei je Chip in JPG und DAT-Format. Um daraus verwertbare Signalwerte zu generieren, die den Expressionszustand der einzelnen Gene, die genomische Kopienzahl oder den Genotypen beschreiben, sind theoretisch mehrere Einflüsse zu beachten. So registriert der Scanner auch einen gewissen Anteil an Hintergrundlicht, das von allgemeiner Reflektion herrühren kann, Signale von potentiellen Artefakten oder nichtspezifischen Hybridisierungen. Die Verarbeitung dieser Daten kann daher generell in drei Schritte untergliedert werden, die Adressierung, auch als Gridding bekannt, Segmentierung und Signalextraktion[184]. Während der Adressierung wird die genaue Position der einzelnen Spots lokalisiert und einem speziellen Gentranskript zugeordnet. Dies erfolgt durch die Projektion und Anpassung eines Lesegitters, bei dem jede Zelle einem Spot entspricht. Die Segmentierung erfasst die Grenzen der einzelnen Spots und unterscheidet somit zwischen Signal- und Hintergrundbereich. Bei der Signalextraktion werden schliesslich die gemessenen Pixelintensitäten sowie die Hintergrundsignale erfasst und in einer Datei mit CEL-Format gespeichert. Um eindeutige Werte zu erhalten, die die Menge der hybridisierten cDNA beschreiben, folgen zuletzt noch Korrekturen anhand der gemessenen Hintergrundsignale und der globalen Verteilung der erfassten Signale auf den einzelnen Chips.

Die Verarbeitung der Daten bis hin zu den CEL-Dateien erfolgte mit der GeneChip Operating Software, Version 1.4, von Affymetrix .

2.3.3 Hintergrundkorrektur, Normalisierung und Segmentierung

Die vom Scanner erfassten und von der Affymetrix Software in den CEL-Dateien gespeicherten Signale sind typischerweise fehlerbehaftet. Größten Einfluss auf die Signale haben präanalytische Parameter und systematische Abweichungen. Hinzu kommen technische Fehler. Diese mussten bei der Auswertung berücksichtigt und soweit möglich korrigiert werden. Diese Verfahren umfassten Schritte zur Hintergrundkorrektur, zur Berechnung der Signale aus einzelnen Sondensignalen sowie zur Normalisierung der Signalbandbreite über die Gesamtheit der gemessenen Chips. Im Falle der array-CGH Chips mussten die Daten segmentiert werden, um Amplifikationen und Deletionen zu bestimmen. Für die SNP-Arrays wurden zusätzlich die Genotypen bestimmt. Im folgenden wird auf die einzelnen Schritte etwas näher eingegangen.

Hintergrundkorrektur umfasst typischerweise eine Reihe von verschiedenen Verfahren, die einzelne oder auch mehrere der folgenden Aufgaben leisten:

- Eine Korrektur von Hintergrundrauschen und Prozessierungseffekten.
- Eine Berücksichtigung von Kreuzhybridisierungen durch nichtspezifische DNA-Bindungen auf dem Array.
- Eine Anpassung der Signalschätzungen an eine sinnvolle Skala, bspw. derart, daß sie einem linearen Konzentrationsgradienten folgen.

Bei vielen Systemen kann die Umgebung der einzelnen Spots zur Berechnung der Hintergrundsignale herangezogen werden. Aufgrund der dichten Packung der Spots auf den Arrays bei Affymetrix musste in diesem Fall der Hintergrund aus den Signalen selber geschätzt werden. Zur Zeit werden für Expressionsarrays bei Forschergruppen primär zwei verschiedene Verfahren favorisiert.

Das von Affymetrix veröffentlichte und in der eigenen Software implementierte MAS 5.0 Verfahren [4, 3, 9, 8] verwendet die eigens dafür konstruierten Mismatchproben auf dem Chip. Jeder Transkriptprobe mit komplementärer Zielsequenz (*PM = Perfect Match*) steht dabei eine Probe mit einer Basenmutation an Position 13 von 25 (*MM = MisMatch*) gegenüber. Die für die MM-Proben gemessenen Signale geben theoretisch den Anteil der unspezifischen Bindungen an den PM-Proben an. Eine Korrektur kann somit durch Subtraktion der MM-Signale von den PM-Signalen erfolgen. Zur Berücksichtigung von Hintergrundlichteinflüssen wird der Chip in mehrere Regionen unterteilt. Für jede dieser Regionen wird der Hintergrund separat bestimmt, über den Chip gewichtet und anschließend mit den einzelnen Signalen verrechnet.

Die Effektivität der MM-Proben ist allerdings seit längerer Zeit umstritten, da die tatsächlichen Vorgänge und Affinitäten bei der Hybridisierung zu komplex sind, um sie über einen einzelnen Basenaustausch abzuschätzen [115, 77]. Tatsächlich können mitunter Signale der MM-Proben beobachtet werden, die über jenen der zugehörigen PM-Proben liegen. Hier überlagern die Effekte der Kreuzhybridisierungen jene Signale der spezifischen Bindungen. Das von Irizarry et al. publizierte *Robust-Multiarray-Average Verfahren* RMA [77, 25, 76] verwirft daher die Signale der MM-Proben und berechnet eine Schätzung des Hintergrundes sowie der unspezifischen Hybridisierungen aus der Verteilung der PM-Signale. Mittlerweile hat sich RMA als Standardverfahren etabliert und ist auch von

2.3. PROGRAMME UND STATISTISCHE METHODEN

Affymetrix in der neuesten Analysesoftware als Alternative zum MAS 5.0 Verfahren implementiert [2].

Die Ergebnisse der Hintergrundkorrektur durch RMA liegen üblicherweise in logarithmischer Form zur Basis 2 vor. Die gemessenen Signalintensitäten der hybridisierten Transkripte tendieren dazu, sich über eine exponentielle Skala zu erstrecken. Die Grundlage dafür bilden exponentielle Funktionen der sequenzabhängigen Affinitäten und der Energetik der Basenpaarung. Eine logarithmische Transformation bietet somit die Möglichkeit die Hybridisierungsskala adäquat anzupassen [153].

Desweiteren bietet eine logarithmische Skala den Vorteil einer besseren Bandbreite der Extremwerte sowie einer leichteren Lesbarkeit. Nicht zuletzt bietet die Transformation von Multiplikationsrechnungen auf Additionsrechnungen einen deutlichen Geschwindigkeitsvorteil bei rechenintensiven Aufgaben, da Multiplikationen auf Computerebene deutlich aufwendiger sind.

Die Basenzusammensetzung der einzelnen Proben führt zu unterschiedlichen Affinitäten [116]. Einerseits spielt dabei der GC-Gehalt der Proben eine Rolle. Andererseits sind nur zwei der vier Basen (Uracil und Cytosin) biotinmarkiert. Die Biotinmarkierung jedoch beeinflusst die Bindungsaffinität. Desweiteren ist die Größe der ausgetauschten Base in der MM-Probe von Bedeutung, da der sterische Einfluss bei zwei sich gegenüberliegenden Purinen größer ist als bei Beteiligung von Pyrimidinen. Einige Korrekturverfahren wie *gcRMA*[182], das auf dem *RMA-Verfahren* aufbaut, versuchen diese Effekte zusätzlich zu modellieren.

Im Rahmen dieser Arbeit wurde RMA als Korrekturverfahren eingesetzt. Die verwendete Funktion *justRMA* stammt aus dem Bioconductor-Package *affy*. Die Ausgabe bestand aus einer Tabelle, die die Expressionssignale enthält. Die einzelnen Transkripte des Microarrays sind zeilenweise angeordnet, die Samples spaltenweise. Die Werte lagen in logarithmischer Form zur Basis 2 vor und wurden als TXT-Datei abgelegt.

Im Falle der array-CGH Daten wird gleichzeitig mit der Tumorprobe eine Referenzprobe auf dem gleichen Array hybridisiert. Die beiden Proben werden mit unterschiedlichen Farbstoffen markiert und anschließend die beiden Signale voneinander subtrahiert. Hier dient die Referenz der Bestimmung von Normalkopienzahl und einer zusätzlichen Korrektur der Hintergrundsignale.

Normalisierung Hierbei handelt es sich um ein Verfahren, in dessen Verlauf unerwünschte, nicht biologische Varianzen verrechnet werden, die zwischen mehreren Microarrays auftauchen. Prinzipiell können Unterschiede zwischen Microarrayergebnissen als biologisch interessant oder verschleiern klassifiziert werden[68]. Die verschleiernde Varianz ist jener Anteil, der im Prozess der Normalisierung entfernt werden soll. Diese Varianzen können z.B. in Unterschieden in der eingesetzten RNA-Menge, in minimalen Einstellungen des Scanners oder anderen Faktoren begründet sein. Gängige Verfahren sind das *Quantile Normalisation Verfahren* [25], *Contrast Normalisation Verfahren* [15], bzw. das *Cyclic Loess Verfahren* [184] oder auch ein einfaches *Scaling Verfahren* [4, 3, 9, 8], wie es Affymetrix ursprünglich verwendete.

Die Expressionsarrays sind mittels der im Bioconductor Package *affy* umgesetzten *Quantile Normalisation* prozessiert worden. Für die SNP-Arrays wurde das *SNP 6.0 Probe Level Access Tool* SPLAT [7] verwendet. Im Anschluß wurde ein Abgleich gegen eine interne Referenz durchgeführt. Die array-CGH Daten wurden mit dem *Loess*-Verfahren normalisiert.

Die Ergebnisse der Normalisierungsverfahren bilden die Grundlage für alle weiteren Auswertungen und statistischen Verfahren.

AffyProbeMiner Das Design der Arrays von Affymetrix beruht auf einer Gliederung in *Proben* und *Probesets*. Die Probesets setzen sich aus mehreren einzelnen Proben, 25 basenpaarlangen Oligonucleotiden, zusammen. Diese sind entlang der Transkriptsequenz eines Gens 3'lastig verteilt. Während der Signalberechnung, werden die Einzelsignale aller Proben eines Probesets nach der Hintergrundkorrektur und Normalisierung zu einem gemeinsamen Signal aufsummiert. Diese Summe gibt das später verwendete Expressionssignal des Probesets an.

Im Zeitraum zwischen dem Design der Microarrays und der Durchführung der Analyse hat sich das Wissen über die humanen und murinen Gene und deren Transkripte stark erweitert. Auf den Arrays von Affymetrix wurden zur Entwicklungszeit etliche EST-Sequenzen integriert. Zudem wurden für viele Gene alternative Transkripte gleichzeitig auf dem Chip abgebildet. Diese sollten mögliche Splice-Varianten der Gene abbilden. In Kontrast zu dem Begriff *Gen* wird daher im weiteren Verlauf dieser Arbeit für die auf den Arrays integrierten Transkripte der Begriff *Probeset* verwendet.

2.3. PROGRAMME UND STATISTISCHE METHODEN

Ende 2007 wurde von einer Gruppe von Mitarbeitern des NIH und der University of Maryland eine Revision der auf einigen Chips befindlichen Probesets und den zugrunde liegenden Transkripten durchgeführt [98]. Im Zuge der Arbeit wurden viele Probesets, die jeweils Transkripte des gleichen Gens abbildeten, zu einem einzigen Probeset zusammengefasst. Damit ging eine Reduzierung der Datenmenge und eine Steigerung der Datenverständlichkeit einher. Die durch die Revision entstandenen neuen Annotationen wurden unter dem Namen *AffyProbeMiner* veröffentlicht¹⁰. Im Rahmen dieser Arbeit wurde die AffyProbeMiner-Annotation für die murinen Arrays der Brusttumore eingesetzt. Bei Ergebnissen werden in diesem Fall sowohl die original Probeset-Bezeichnungen von Affymetrix als auch die neuen Bezeichnungen angegeben.

Segmentierung Genomische DNA erfährt während der Tumorprogression häufig lokale Veränderungen in der Kopienzahl. Im Gegensatz zu Expressionsarrays ist hier das Verhältnis der Signalniveaus der einzelnen Gentranskripte von entscheidender Bedeutung. Das Verhältnis der Transkriptsignale untereinander bestimmt die Kopienzahl von chromosomalen Abschnitten, der Segmente. Segmentierungsverfahren, häufig auch *Break-Point-Detection Verfahren* genannt, bestimmen die Grenzen der einzelnen Segmente. Jedem Segment wird ein gewichteter Mittelwert der von ihm zusammengefassten Transkriptsignale zugeordnet. Im Rahmen dieser Arbeit wurde sowohl für die SNP-Arrays als auch die array-CGH Chips das *Circular Binary Segmentation Verfahren* [121, 172] angewendet. Dieses ist im Bioconductor Package *DNAcopy* umgesetzt.

Bei den SNP 6.0 Arrays der Neuroblastome wurden sowohl die Signale der SNP als auch der CN Proben zur Bestimmung von Amplifikationen und Deletionen herangezogen. Diese wurden zu Beginn jeweils separat mit einer internen Referenz standardisiert und anschließend segmentiert. Nach der Bestimmung und Klassifizierung der Aberrationen wurden die Ergebnisse zusammengeführt.

Ausgehend von den mit *DNAcopy* berechneten Segmenten wurde bei genomischen Aberrationen zwischen *Gain*, *Amplifikation*, *Loss* und *Deletion* unterschieden. Zur Klassifikation wurde vorangehend die Nulllinie der Signale bestimmt. Hier erfolgte vorerst eine Zentrierung der Segmente um den Modalwert der Segmentmittelwertverteilung jedes Samples. Für jedes einzelne Sample wurde die Standardabweichung des 50% Quantils der

¹⁰AffyProbeMiner: <http://gauss.dbb.georgetown.edu/liblab/affyprobeminer/home.html>

unsegmentierten Signalwerte ermittelt. Anschließend wurden Aberrationen als *Gains* bzw. *Losses* klassifiziert, sofern der Segmentmittelwert größer bzw. kleiner als die (negative) x_{low} -fache Standardabweichung war. Für *Amplifikationen* bzw. *Deletionen* wurde ein Segmentmittelwert vorausgesetzt, der die (negative) x_{high} -fache Standardabweichung des 90% Quantils über- bzw. unterschritt. Die Faktoren x_{low} und x_{high} waren modellabhängig.

Genotypisierung Zur Bestimmung der Genotypen für die SNP 6.0 Arrays wurde die Genotyping Console Version 2.1 von Affymetrix verwendet. Diese verwendet das *Birdseed Verfahren* [6, 129, 5]. Die Ausgabe besteht aus einer TXT-Datei, die die Genotypen der einzelnen SNPs für alle Gewebeproben speichert.

Mittlere Absolute Abweichung Die Mittlere Absolute Abweichung oder auch Mean Absolute Deviation, kurz MAD-WERT, ist ein robuster statistischer Schätzer für die Standardabweichung. Er wird häufig zur Abschätzung der Stärke des Hintergrundsignals bei array-CGH Experimenten verwendet. Je kleiner der Wert ist, umso geringer wird das Hintergrundsignal eingeschätzt und umso höher ist die Qualität der weiterverarbeiteten Signale. Der MAD-Wert errechnet gemäß Gleichung 3.

$$MAD = median|x_i - \bar{x}|$$

Gleichung 3: Berechnung der Mittleren Absoluten Abweichung (Mean Absolute Deviation). Sie wird häufig als Schätzer für die Stärke des Hintergrundsignals bei array-CGH Experimenten verwendet. Je kleiner der Wert ausfällt, umso geringer ist das Hintergrundsignal.

2.3.4 Visualisierungsverfahren

Zur Visualisierung der Datenqualität und Ergebnisse wurden mehrere Standardverfahren zur Anwendung gebracht. Diese sollen im folgenden Abschnitt beschrieben werden.

Density-Plot Die Dichtefunktion der Signalwerte eines Arrays gibt Aufschluss über Verteilung und Zusammensetzung der Signale. Häufig ist eine Trennung zwischen spezifischen Hybridisierungen und Transkripten mit vorwiegend nichtspezifischem Signal als bimodale Verteilung erkennbar. Desweiteren kann die Signalverteilung als Maß für die Vergleichbarkeit mehrerer Arrays untereinander herangezogen werden. Die Berechnung der Dichtefunktionen in dieser Arbeit erfolgte mittels der Funktion *density* aus dem R-Package *stats*.

2.3. PROGRAMME UND STATISTISCHE METHODEN

Box-Whisker Plots Ein Box-Whisker Plot [161], auch kurz Boxplot genannt, ist ein Diagramm, das zur einfachen graphischen Darstellung der Verteilung von Daten und ihrer Quantile eingesetzt wird. Er fasst verschiedene Maße der zentralen Tendenz, Streuung und Schiefe in einem Diagramm zusammen. Alle Werte der Fünf-Punkte-Zusammenfassung, also der Median, die zwei Quartile und die beiden Extremwerte, sind dargestellt. Die Länge der Whisker beträgt maximal das 1,5-fache des Interquartilabstands. Werte, die diese Grenze überschreiten, werden als Ausreißer gewertet und als einzelne Punkte abgebildet. Für diese Arbeit erstellte Boxplots wurden mithilfe der Funktion *graphics* des R-Packages *graphics* generiert.

Heatmaps stellen eine Anzahl gemessener Signalwerte in farblicher Kodierung dar. Diese Strategie ermöglicht einen schnellen Überblick über den Signalverlauf in verschiedenen Samples oder Genen. Dabei sind üblicherweise die Samples in Spalten und Gene in Zeilen angeordnet. Häufig wird diese Technik mit einem vorangehenden Cluster-Verfahren (siehe Abschnitt 2.4.2) kombiniert. In diesem Fall sind die Samples und/oder Gene nach Ähnlichkeit sortiert. Dendrogramme ergänzen in diesem Fall die Übersicht zur Visualisierung der Verwandtschaftsverhältnisse. Im Rahmen dieser Arbeit wurden Heatmaps mithilfe der *heatmap.2* Funktion des R-Package *gplots* erstellt. Für erweiterte Farbskalen ist das R-Package *RColorBrewer* eingesetzt worden.

Venn-Diagramme Venn-Diagramme sind Mengendiagramme, die alle möglichen Relationen zwischen zwei oder mehr Mengen anhand von Kreisen oder Ellipsen darstellen. Die Schnittmengen werden durch überlappende Regionen der geometrischen Formen symbolisiert. Im Rahmen dieser Arbeit wurden Venn-Diagramme zur Darstellung der Schnittmengen mehrerer Genlisten eingesetzt. Die dafür verwendeten Funktionen *vennCounts* und *vennDiagram* stammen aus dem Bioconductor-Package *limma*.

Segmentation Plots Um Änderungen in der Kopienzahl von DNA Samples schnell erfassen zu können werden die Ergebnisse einer Segmentierung von array-CGH Daten üblicherweise graphisch dargestellt. Dabei kann sich der Rahmen/Umfang auf einzelne Chromosomen beschränken oder über das gesamte Genom erstrecken. Die Signale der einzelnen Transkripte werden als Punkte dargestellt. Die Ordinate bildet dabei den Transkript-LogRatio, das Verhältnis des Tumorsignals zum Referenzsignal, ab. Die einzelnen Transkripte sind gemäß der Rangfolge ihrer chromosomalen Position entlang der Abszisse angeordnet. Zusätzlich zu den ursprünglichen LogRatios werden die Segmente als Kopienstu-

fen abgebildet. Einige Graphiken in dieser Arbeit wurden mittels der *plot* Funktion des BioConductor-Packages *DNAcopy* erstellt. Andere wurden manuell über den *plot* Befehl des R-Packages *graphics* generiert.

2.4 Supervidierte und unsupervidierte Testverfahren

Zur Untersuchung möglicher in der Tumorgenese involvierter Gene können primär zwei Strategien verfolgt werden. Erstens können sogenannte supervidierte Verfahren eingesetzt werden. Dabei wird a priori Wissen über die phänotypische Zusammensetzung der Samples oder den funktionellen Zusammenhang bestimmter Transkripte für eine gezielte Gruppierung verwendet. Supervidierte Verfahren identifizieren Gene, Regionen oder Samplegruppen, die in Hinsicht auf die getroffene Gruppierung einen signifikanten Unterschied aufweisen. Zweitens ist der Einsatz von nichtsupervidierten Verfahren möglich. In diesem Fall werden Gruppen von Samples oder Genen ausschließlich auf der Basis der gemessenen Signale bestimmt, die hinsichtlich ihrer Signalwerte in verschiedene Klassen eingeordnet werden können. Der folgende Abschnitt soll eine Übersicht über die verwendeten Verfahren geben.

2.4.1 Hypothesentests und Korrektur für multiples Testen

Um statistisch signifikante Unterschiede in den Expressionsniveaus der Gruppen zu bestimmen, wurden verschiedene Testverfahren eingesetzt. Das Konzept der statistischen Signifikanz beruht dabei auf der Annahme, daß die Wahrscheinlichkeit die observierten Unterschiede zufällig zu beobachten äußerst gering ist. Das Signifikanzniveau ist dabei ein Maß für die Wahrscheinlichkeit die Nullhypothese abzulehnen, obwohl sie in Wirklichkeit zutrifft. Zu diesem Zweck wird ein P-Wert berechnet. Ist dieser Wert kleiner als das Signifikanzniveau wird die Nullhypothese abgelehnt, andernfalls angenommen. Je kleiner der P-Wert ist, umso signifikanter ist der Unterschied. Abhängig davon, ob man von einer Normalverteilung der Daten ausgeht oder nicht, bezeichnet man die Testverfahren als parametrisch oder nichtparametrisch.

Student's t-Test Der *Student's t-Test* [64] ist ein parametrisches Verfahren, bei dem die Mittelwerte zweier Gruppen verglichen werden. Die primär berechnete Teststatistik wird im Anschluß in einen P-Wert transformiert. Der Test geht von einer Normalverteilung der Daten und gleicher Varianz in beiden Gruppen aus. Die Teststatistik T wird anhand

2.4. SUPERVIDIERTE UND UNSUPERVIDIERTE TESTVERFAHREN

einer Formel berechnet, die die Differenz der Mittelwerte im Zähler trägt. Der Nenner wird durch den Standardfehler der Differenz gebildet, der mit sinkender Varianz innerhalb der Gruppen oder steigender Gruppengröße schrumpft. Steigt die Differenz im Zähler oder schrumpft der Standardfehler im Nenner so wächst T. Die T-Statistik steht in umgekehrter Relation zum P-Wert. Je größer T ist, umso kleiner ist demnach der P-Wert.

Welch t-Test Der Nachteil des *Student's t-Test* ist die Voraussetzung gleicher Varianzen in den betrachteten Gruppen. Die Bedingung ist häufig erst bei Gruppengrößen von 10.000 Werten erfüllt. Selbst die umfangreichsten Microarray-Studien erreichen jedoch nicht annähernd solche Größen. In diesem Fall kann der *Welch t-Test* [177] eingesetzt werden. Es handelt sich dabei um eine Abwandlung des *Student's t-Test*, die von potentiell unterschiedlich großen Varianzen innerhalb der Gruppen ausgeht. Im Rahmen dieser Arbeit wurde die Funktion *t.test(var.equal=FALSE)* aus dem R-Package *stats* aufgerufen, um Welch-P-Werte zu berechnen.

Bonferroni Korrektur Bei gleichzeitigen multiplen Vergleichen von Gruppen steigt die Wahrscheinlichkeit per Zufall das Signifikanzniveau zu unterschreiten. So existiert bei 100 Vergleichen und einem Signifikanzniveau von 5% bereits eine reelle Chance, rein zufällig 5 als signifikant bewertete Ergebnisse zu erhalten. Korrekturverfahren für multiples Testen bereinigen Resultate aus statistischen Tests um jene Zufallsergebnisse. Zu den recht konservativen und stringenten Methoden zählt die *Bonferroni Korrektur* [181]. Sie kontrolliert die Anzahl der falsch-positiven Ergebnisse. In dieser Arbeit erfolgte die Berechnung der *Bonferroni Korrektur* über die Funktion *p.adjust(method="Bonferroni")* des R-Packages *stats*.

Benjamini-Hochberg Korrektur Im Gegensatz zu der *Bonferroni Korrektur* kontrolliert die *Benjamini-Hochberg Korrektur* [20] die Anzahl der falsch-negativen Ergebnisse. Das Verfahren ist damit weniger stringent als ersteres und damit für manche Experimente besser geeignet. In dieser Arbeit erfolgte die Berechnung der *Benjamini-Hochberg Korrektur* über die Funktion *p.adjust(method="BH")* des R-Packages *stats*.

2.4.2 Clustering-Verfahren

Clusteringverfahren dienen im Allgemeinen der nichtsupervidierten Gruppierung von Daten. Ein *Cluster* stellt dabei eine Gruppe von Datenpunkten dar, die sich gegenseitig äh-

licher sind als den übrigen Daten. Die berechneten Gruppierungen können als Bestätigung vorgefasster Hypothesen aufgefasst werden, zur Visualisierung von Ähnlichkeiten oder als Ausgangspunkt für weitere supervidierte Testverfahren, so z.B. zur Bestimmung der Gene, die die determinierte Ähnlichkeit am besten beschreiben.

Hierarchical Clustering Unter Clustering versteht man die Klassifizierung von Daten in mehrere Gruppen auf der Grundlage von definierten Distanzmaßen. Das Distanzmaß bestimmt die Ähnlichkeit zweier Werte und somit die Relation der Daten zueinander. Je kleiner die Distanz ist, umso ähnlicher sind die Daten. Häufig verwendete Distanzmaße sind die *Euklidische Distanz*, *Manhattan Distanz* und das *Pearson Produkt-Moment-Korrelationsmaß* [26]. Die Gruppierung der Daten anhand der berechneten Distanzen kann nach verschiedenen Verfahren erfolgen. Beim Hierarchischen Cluster Verfahren wird in einem iterativen Prozess ein Datenpunktpaar zu einem Cluster zusammengefasst und die Distanz der übrigen Datenpunkte zu diesem Cluster neu berechnet. Die Auswahl des in einer Iteration zusammengefassten Paares erfolgt gemäß einem zuvor definierten Linkage-Verfahren.

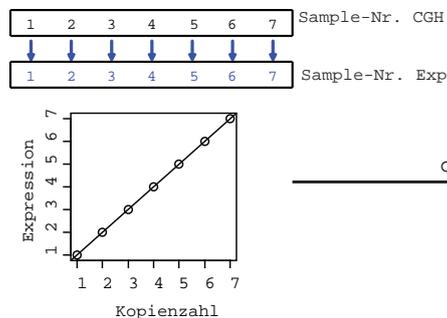
Als Distanzmaß wurde im Rahmen dieser Arbeit die Euklidische Distanz verwendet. Als Linkage-Verfahren wurde das *Ward Verfahren* eingesetzt. Die Distanzberechnung und Clustering Prozeduren wurden über die Funktionen *dist* und *hclust* des R-Packages *stats* durchgeführt.

2.4.3 Korrelationsverfahren

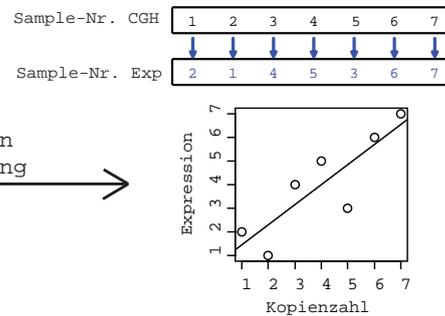
Ein wichtiges Ziel des Projektes war die Korrelation der aus den Expressionsarrays gewonnenen Daten mit den Ergebnissen der SNP-Analytik bzw. der array-CGH Chips. Dieser Ansatz sollte helfen, den Einfluss der genetischen Variationen oder Läsionen auf die Metastasierung und den Differenzierungsgrad abzuschätzen. Obwohl bereits erste Methoden zur Verknüpfung von Expressionssignaturen und genomischen Daten veröffentlicht wurden [186, 173, 41, 10, 157], ist die Kombination dieser Daten nach wie vor komplex, insbesondere da die Daten auf unterschiedlichen Plattformen generiert wurden und in der Folge sehr inhomogen sind. Zwei Korrelationsansätze wurden untersucht und werden im folgenden näher beschrieben.

2.4. SUPERVIDIERTE UND UNSUPERVIDIERTE TESTVERFAHREN

native / korrekte Zuordnung von
CGH-Daten zu Expressionsdaten



permutierte Zuordnung der Daten



Permutation
der Zuordnung

Bei guter Korrelation der nativen Datenzuordnung liegen deren Werte dichter an der Diagonalen als bei den permutierten Zuordnungen!

Abbildung 2.1: Permutation der Samplezuordnung beim Korrelationsverfahren nach B. Stranger et al. [157]. Bei einer Permutation wird eine „falsche“ Zuordnung zwischen Expressions- und CGH- bzw. SNP-Arrays vorgenommen. Die Korrelationsgüte der originalen, „korrekten“ Zuordnung im Vergleich zu den P-Werten der permutierten Zuordnungen entscheidet über die Signifikanz.

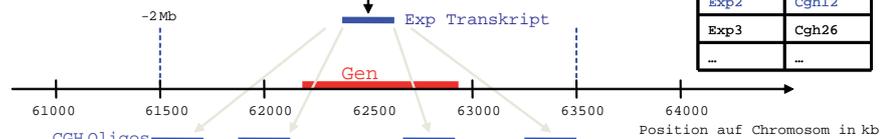
Verfahren nach B. Stranger In ihrer Arbeit verwendeten B. Stranger et al. [157] erstmals eine Korrelationsstrategie, die Übereinstimmungen in Expressionsdaten und Copy-number Variations mittels linearer Regressionen bewertet. Im Laufe der vorliegenden Arbeit wurde dieses Verfahren adaptiert und für spätere Berechnungen auf einem Server des Zentrum für Bioinformatik Hamburg¹¹ zugänglich gemacht. Zu Beginn des Verfahrens wird eine 1:n Tabelle erstellt, in der jedem Probeset des Expressionschips mehrere Transkripte des CGH- oder SNP-Arrays, innerhalb eines definierten Basenpaarfensters up- und downstream vom Expressionstranskript, zugeordnet werden, siehe Abbildung 2.2. In den Folgeschritten werden für die Zuordnungen lineare Regressionen zur Bestimmung der Übereinstimmungsgüte von Expression und genomischer Kopienzahl durchgeführt. Um statistische Fehler durch multiples Testen zu berücksichtigen, werden für jedes Probeset des Expressionschips, mittels 10.000 randomisierter Permutationen, definierte P-Wert-Quantilgrenzen bestimmt, siehe Abbildungen 2.1 und 2.2. Fällt der native P-Wert einer Regression unter eine gewünschte Quantilgrenze der P-Wert-Verteilung, wird die Korrelation als signifikant anerkannt.

¹¹Zentrum für Bioinformatik Hamburg: <http://www.zbh.uni-hamburg.de>

A)

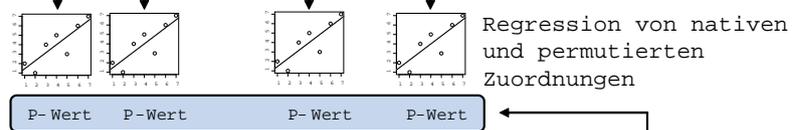
Erstellung von 1:n Tabelle, die jedem Gen auf Expressionsebene mehrere CGH-Oligos innerhalb definierter kb Grenzen up- und downstream vom Gen zuordnet.

Exp. ID	Cgh ID
Exp1	Cgh1
Exp1	Cgh2
Exp1	Cgh3
Exp2	Cgh9
Exp2	Cgh10
Exp2	Cgh11
Exp2	Cgh12
Exp3	Cgh26
...	...



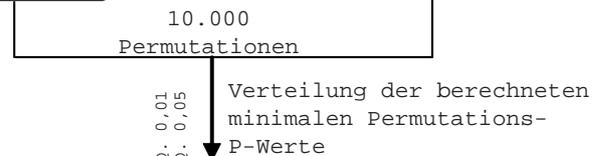
B)

Für jede Exp-CGH Kombination wird eine Regression durchgeführt und der Minimalwert der Kombinationen bestimmt.



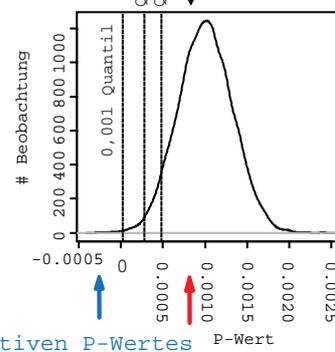
C)

Permutation der Samplezuordnung und Neu-Kalkulation der P-Werte und ihres Minimums



D)

In Abhängigkeit von der Lage des nativen P-Wertes (der korrekten Zuordnung) bezüglich der Verteilung, wird das Gen als gut (blau) oder schlecht korrelierend (rot) eingestuft.



Lages des nativen P-Wertes

Abbildung 2.2: Das Korrelationsverfahren für einzelne Gene nach B. Stranger [157] wurde adaptiert und zusätzlich für SNP 6.0 Arrays von Affymetrix angepasst. Jedem Gen werden auf Expressionsebene mehrere SNP/CNV- oder CGH-Probes zugeordnet (A), die innerhalb eines definierten Bereiches up- und downstream vom Transkript liegen. Über eine Reihe von Regressionen (B) und Permutationen (C), siehe auch Abbildung 2.1, wird eine Korrelations-P-Wert-Verteilung berechnet. Die Lage des nativen P-Wertes innerhalb dieser Verteilung bestimmt über seine Signifikanz (D).

2.4. SUPERVIDIERTE UND UNSUPERVIDIERTE TESTVERFAHREN

Genomweite Korrelation Trägt man die Signale eines Expressions-Microarrays in einem Segmentation Plot auf, so sind die Auswirkungen von Änderungen in der Kopienzahl eines chromosomalen Abschnitts normalerweise kaum zu erkennen. Diese Beobachtung wurzelt in der Heterogenität der *Normalniveaus* der Transkriptsignale. Auf DNA Ebene liegt das Normalniveau der Kopienzahl für alle chromosomalen Abschnitte, bei diploiden Organismen, natürlicherweise bei zwei Kopien, sofern keine Polyploidie vorliegt. Bei der Expression von mRNA hingegen kann der Normalzustand von Gen zu Gen unterschiedlich sein. Um dennoch starke Abweichungen von diesem Normalzustand abbilden zu können wurde im Rahmen dieser Arbeit eine *z-Score* ähnliche Transformation der Werte über mehrere Chips berechnet. Für jedes Probeset bzw. Gen wurde ein mittelwertähnlicher, robuster Tukey's-Biweight Schätzer \bar{x}_T über alle Samples sowie die Standardabweichung σ bestimmt. Der transformierte Signalwert ergab sich aus der Differenz zwischen Originalwert und robustem Mittelwert dividiert durch die Standardabweichung, siehe auch Gleichung 4. Diese Werte wurden anschließend als Segmentation Plot aufgetragen.

$$x_{scaled} = \frac{x - \bar{x}_T}{\sigma}$$

x_{scaled}	=	Transformierter Signalwert eines Probesets
x	=	Ausgangssignalwert eines Probesets
\bar{x}_T	=	Mittelwertähnlicher, Robuster Tukey-Biweight Schätzer
σ	=	Standardabweichung des Probesets über alle Samples

Gleichung 4: Berechnung des modifizierten *z-scores* je Probeset zur globalen Korrelation der Expressionsdaten mit den genomischen Kopienänderungen.

Eine hinreichend gute Referenz ließ sich hier allerdings nur mit ausreichend vielen, transkriptionell verschiedenen Samples berechnen. Bei den Brusttumoren konnten die 80 verfügbaren Samples herangezogen werden. Im Falle der Neuroblastome mussten die 9 Samples zur Standardisierung um weitere ergänzt werden. Zu diesem Zweck wurden 72 Samples der GEO-Datenbank¹² eingesetzt.

¹²siehe Abschnitte 2.3.1 und 3.3.2

2.4.4 Signaltransduktionswege und Funktionsanalytik

GenMapp GenMapp¹³ ist ein frei verfügbares Programm zur Projektion von Expressionsdaten auf bekannte und publizierte Signaltransduktionswege. Größtenteils handelt es sich hierbei um KEGG- und BioCarta-Maps. Einerseits können Gene farblich hervorgehoben werden, die ein definiertes Expressionsmuster aufweisen. Andererseits können anhand von dem Modul *Pathfinder* signifikant betroffene Transduktionswege ermittelt werden.

EXPANDER Das Programm *EXPANDER*¹⁴ des Weizmann Institutes in Israel stellt verschiedene Funktionen für Expressionsarrays zur Verfügung. Hier wurden primär die beiden Module *TANGO* und *PRIMA* eingesetzt. Ersteres bestimmt innerhalb von übergebenen Genlisten die Signifikanz von Anreicherungen bestimmter Genfunktionsklassen gemäß den GeneOntology-Annotationen. Bei *PRIMA* handelt es sich um eine Methode, die spezielle Anreicherungen von Transkriptionsfaktorbindungsstellen vorhersagt. Die Vorhersagen beruhen auf Bindungsmotiven der Datenbanken TRANSFAC in Form von positionsgewichteten Matrizen. Sie sind rein statistischer Natur und beinhalten noch keine Aussage über den tatsächlichen Aktivierungszustand der Transkriptionsfaktoren aus.

2.4.5 Funktionsaufrufe und Flags

Methode	R-Package	Funktionsaufruf
t-Test	<i>stats</i>	<code>t.test(x, y, var.equal=FALSE)</code>
Benjamini-Hochberg	<i>stats</i>	<code>p.adjust(x, method="BH")</code>
Heatmaps	<i>gplots</i>	<code>heatmap.2(x, scale="Row")</code>
Clustering	<i>stats</i>	<code>hclust(x, method="ward")</code>
Euklidische Distanz	<i>stats</i>	<code>dist(x)</code>
Normalisierung Expressionsarrays	<i>affy</i>	<code>justRMA (»files«)</code>
Normalisierung CGH-Arrays	<i>limma</i>	
Quantil-Berechnung	<i>stats</i>	<code>quantile(x, probs=c(0.05, 0.25, 0.75, 0.95))</code>
Segmentierung	<i>DNAcopy</i>	<code>segment(smooth.CNA(CNA(x)))</code>
Tukey-Biweight	<i>affy</i>	<code>tukey.biweight(x)</code>
Varianz / Mittelwert	<i>stats</i>	<code>var(x) / mean(x)</code>

¹³<http://www.genmapp.org/>

¹⁴EXPANDER: <http://acgt.cs.tau.ac.il/expander/expander.html>

2.5. DATEIFORMATE

Package	Version
affy	1.16.0
DNAcopy	1.12
gplots	2.3.2
graphics	2.6.2
limma	2.12.10
RColorBrewer	1.0-2
simpleaffy	1.16.0
stats	2.6.2

Tabelle 2.17: Namen und Versionen der verwendeten Bioconductor Packages

2.5 Dateiformate

Im Zusammenhang mit dieser Arbeit sind einige gängige und einige weniger verbreitete Dateiformate zu nennen. Sie betreffen die auf dem beiliegenden Datenträger veröffentlichten Daten. Diese sollen im folgenden Abschnitt etwas näher beschrieben werden.

DAT Format *Data File*, Affymetrix Microarray Datei, die das eingelesene Bild des Microarrays enthält.

CEL Format *Cel Intensity File*, Affymetrix Microarray Datei, die einen einzelnen Intensitätswert pro Gitterzelle enthält und von der GeneChip Operating Software aus den DAT-Dateien erstellt wird. Diese Dateien werden üblicherweise als Roh-Ergebnis eines Microarrayexperimentes angesehen und mittlerweile bei Veröffentlichungen von Datenbanken wie GEO¹⁵ oder ArrayExpress¹⁶ eingefordert. Ausgehend von diesen Dateien sollten alle weiteren Ergebnisse bei nachvollziehbarer Auswertungsstrategie reproduzierbar sein.

TXT Format *Text File*, ein gängiges Format zur Textspeicherung in Klarformat ohne besondere Formatierung. Im Rahmen dieser Arbeit werden damit spezielle *Tab Delimited Files* bezeichnet, die Tabellen speichern. Jede Textzeile entspricht einer Tabellenzeile, die einzelnen Spalten sind mittels Tabulatorzeichen voneinander getrennt.

XLS Format *Excel File*, gängiges Format der Firma Microsoft zur kodierten Speicherung von Tabellen und ihren Formatierungen.

¹⁵<http://www.ncbi.nlm.nih.gov/geo>[19, 54]

¹⁶<http://www.ebi.ac.uk/arrayexpress>[124]

JPG Format *JPG* oder auch *JPEG*, ist ein Format zur verlustbehafteten Komprimierung und Speicherung von Bildern. Es handelt sich hier um die gebräuchliche Bezeichnung für die Norm ISO 10918-1. Das Format wurde sowohl für die Speicherung der Scansbilder der Microarrays als auch für erstellte Grafiken verwendet.

PS Format *PostScript* ist ein Grafikformat, das sich über Jahre zu einem Standard in der Druckindustrie entwickelt hat. Die Dateien können nahezu ohne Verlust von Informationen in PDF-Dateien eingebettet werden. Das Format wurde zusammen mit dem *JPG*-Format zur Erstellung von Grafiken eingesetzt.

Kapitel 3

Ergebnisse

Im Rahmen dieser Arbeit wurden 9 Neuroblastome des SCID Mausmodells sowie 80 Mammakarzinome des BALB/c Mausmodells untersucht. Alle 9 Neuroblastome wurden sowohl mit SNP-Arrays auf genetische Aberrationen als auch mit Expressionsarrays hinsichtlich Änderungen auf Transkriptionsebene untersucht. Bei den Mammakarzinomen hingegen erfolgte die Diagnostik auf genomischer Ebene mittels array-CGH Technik. Hier wurden 56 der 80 Proben sowohl phänotypisch als auch genotypisch analysiert. Die Mengen des jeweils eingesetzten Materials, die Resultate der Konzentrations- und Qualitätskontrollen sowie der Arrayhybridisierungen werden im folgenden Abschnitt beschrieben. Anschließend werden die aus den statistischen Verfahren gewonnen Ergebnisse beschrieben.

3.1 Expressionsanalytik

In diesem Abschnitt werden die Resultate der Expressionsanalytik beschrieben. Zur Vermeidung von Mehrdeutigkeiten und zum besseren Verständnis werden in diesem und den folgenden Abschnitten die angegebenen Konventionen eingehalten: Expressionswerte werden generell in logarithmischer Form zur Basis 2 angegeben. Die daraus resultierenden Änderungen im Expressionsniveau berechnen sich als Differenz der Mittelwerte in logarithmischer Form und werden als Signal Log Ratio (SLR) bezeichnet. Zur Darstellung der Änderungen in linearem Maßstab, z.B. 2fache Änderung, wird die Bezeichnung *Fold Change* (FC) verwendet. Eine SLR von 1 entspricht somit einem 2fachen FC.

3.1.1 Brusttumore

Die Untersuchungen von Brusttumoren in Mäusen umfassten eine Reihe von unterschiedlichen Fragestellungen, die eine bessere Beschreibung der Tumorgenese ermöglichen sollten. Zum einen sollte der Effekt von Mutationen im *Trp53*-Gen (nachfolgend mit *mutP53* abgekürzt) während der Laktationsphase erfasst werden. Hierzu wurden sowohl monotransgene Mäuse mit viralem *SV40 T-Antigen* als auch bitransgene Mäuse mit *SV40 T-Antigen* und *mutP53* untersucht. Zum anderen sollte ein potentieller transkriptioneller Unterschied zwischen differenzierten und undifferenzierten Tumoren identifiziert werden. Es wurden daher Tumore der Differenzierungsgrade G1 bis G4 untersucht. Wildtyp BALB/c Mäuse sollten dabei einen Abgleich der Tumorsignaturen mit dem Normalzustand ermöglichen. Tabelle 3.1 weist die Zuordnung der einzelnen Samples zu den verschiedenen Gruppen aus. Außerdem können der Tabelle jene Samples entnommen werden, für die gleichzeitig eine array-CGH Untersuchung durchgeführt wurde.

Tabelle 3.1: Liste der untersuchten Brustkrebs-Samples und ihre Eingruppierung. Bei den Samples mit der Bezeichnung „MaCa“ in der Tumorgradingspalte handelt es sich um spontane Mammakarzinome, für die kein Grading durchgeführt wurde. Tumore der Stufen G1 und G2 wurden als differenziert eingestuft, Tumore der Stufen G3 und G4 als undifferenziert. Mäuse der Linien NP8 und T1 exprimierten lediglich das SV40 T-Antigen. Mäuse der Linien NP8-W10 und T1-H22 exprimierten zusätzlich eine mutierte Variante des *Trp53* Gens (R245W bzw. R270H).

Array Name	interne ID	Tumorgrading	Mauslinie	Zusätzl. CGH
<i>Tumore</i>				
00803.CEL	12030-4	G1	NP8	X
00804.CEL	12271-5	G1	NP8	X
00805.CEL	12159-6	G1	NP8	X
00806.CEL	10667-2	G2	NP8	X
00807.CEL	10811-2	G2	NP8	X
00808.CEL	11885-2	G2	NP8	X
00809.CEL	11885-5	G2	NP8	X
00810.CEL	12108-5	G3	NP8	X
00812.CEL	12192-8	G3	NP8	X
00813.CEL	12212-2	G3	NP8	X
00814.CEL	11304-3	G3	NP8	X
00853.CEL	12109-1	G3	NP8	X
00815.CEL	10870-6	G4	NP8	X
00816.CEL	12192-2	G4	NP8	X
00817.CEL	11093-8	G4	NP8	X
00821.CEL	11063-5	G2	NP8-W10	X
00857.CEL	11254-2	G3	NP8-W10	X
00822.CEL	11254-7	G2	NP8-W10	X
00858.CEL	11461-2	G4	NP8-W10	X
00823.CEL	11519-6	G2	NP8-W10	X
00824.CEL	11947-2	G2	NP8-W10	X
00825.CEL	12525-1	G2	NP8-W10	X
00826.CEL	12569-8	G2	NP8-W10	X
00827.CEL	11560-1	G3	NP8-W10	X
00829.CEL	11445-2	G3	NP8-W10	X
00830.CEL	11485-7	G3	NP8-W10	X
00831.CEL	11551-2	G3	NP8-W10	X

Fortsetzung auf der nächsten Seite

3.1. EXPRESSIONSANALYTIK

Übersicht der Brustkrebssamples – Fortsetzung

Array Name	interne ID	Tumorgrading	Mauslinie	Zusätzl. CGH
00832.CEL	11445-7	G4	NP8-W10	X
00834.CEL	11551-6	G4	NP8-W10	X
10864-6.CEL	10864-6	G1	T1	X
11084-6.CEL	11084-6	G1	T1	X
11096-3.CEL	11096-3	G1	T1	X
10911-7.CEL	10911-7	G2	T1	X
11155-4.CEL	11155-4	G2	T1	X
11155-7.CEL	11155-7	G2	T1	X
10760-8.CEL	10760-8	G3	T1	X
11089-1.CEL	11089-1	G3	T1	X
11191-2.CEL	11191-2	G3	T1	X
11043-4.CEL	11043-4	G4	T1	X
11383-3.CEL	11383-3	G1	T1-H22	X
10682-3.CEL	10682-3	G3	T1-H22	X
11696-5.CEL	11696-5	G3	T1-H22	X
11389-1.CEL	11389-1	G1	T1-H22	X
10682-6.CEL	10682-6	G3	T1-H22	X
11389-3.CEL	11389-3	G2	T1-H22	X
10696-4.CEL	10696-4	G4	T1-H22	X
10699-7.CEL	10699-7	G1	T1-H22	X
11389-8.CEL	11389-8	G2	T1-H22	X
11390-3.CEL	11390-3	G2	T1-H22	X
10726-5.CEL	10726-5	G4	T1-H22	X
<i>spontane Mammakarzinome</i>				
00859.CEL	11085-2	MaCa	H22	X
00849.CEL	11721-1	MaCa	H8	X
00851.CEL	11866-7	MaCa	W10	X
00852.CEL	10590-1	MaCa	W10	X
00848.CEL	11536-8	MaCa	WT	X
00856.CEL	11618-1	MaCa	WT	X
<i>Transgene Mäuse vor der Tumorentwicklung</i>				
hei1_12198_6-7.CEL	12198-6/7	–	T1 T130PW	
hei1_12199_6-7.CEL	12199-6/7	–	T1 T130PW	
hei1_12200_6-7.CEL	12200-6/7	–	T1 T130PW	
00818.CEL	11433-3/7	–	NP8	
00819.CEL	11466-3/7	–	NP8	
00855.CEL	11637-3/7	–	NP8	
00835.CEL	12545-3/7	–	NP8-W10	
00836.CEL	12555-3/7	–	NP8-W10	
00837.CEL	12449-3/7	–	NP8-W10	
00844.CEL	12603-6/7	–	H8	
00845.CEL	12602-6/7	–	H8	
00846.CEL	12605-6/7	–	H8	
00841.CEL	12600-6/7	–	W10	
00842.CEL	12601-6/7	–	W10	
00843.CEL	12609-6/7	–	W10	
<i>Wildtypen</i>				
00838.CEL	12615-6/7	–	WT	
00839.CEL	12613-6/7	–	WT	
00840.CEL	12614-6/7	–	WT	
11554-2.3.CEL	11554-2+3	–	WT	
WT1_Mouse430_2.CEL	WT1	–	WT	
WT2_Mouse430_2.CEL	WT2	–	WT	
hei1_12066_6-7.CEL	12066-6/7	–	WT T130PW	
hei1_12132_6-7.CEL	12132-6/7	–	WT T130PW	
hei1_12133_6-7.CEL	12133-6/7	–	WT T130PW	

Ende Tabelle 3.1

Laboranalytik Zur Untersuchung der Tumorgenese in Mammakarzinomen wurde eine Expressionsanalytik für 80 Samples durchgeführt. Davon erfolgten die laboranalytischen Schritte für 24 Samples des T1-Modells nach der Isolierung der RNA¹ bis zur Hybridisierung im eigenen Labor. Die übrigen 56 Samples des NP8-Modells wurden an der Universität von Oslo² aufgereinigt, amplifiziert und hybridisiert.

Die photometrische Konzentrationsbestimmung der 24 Samples nach der Isolierung ergab Konzentrationen zwischen 2,3µg/µl und 7,4µg/µl. Die Verhältnisse der A260/A280 Extinktionen lagen bei allen Chips bei einem Wert von 2,2. Daraus wurde auf eine ausreichende Reinheit der Proben geschlossen. Die vollständige Liste der Extinktionsverhältnisse und Konzentrationen ist in Tabelle 3.2 angegeben.

Für die cDNA-Synthese wurden jeweils 15µg total RNA eingesetzt. Die Qualität der cDNA wurde auf einem 1%igem EtBr-Agarosegel überprüft. Die in Abbildung 3.1 erkennbaren Banden der ehemals 28S und 18S ribosomalen RNA waren repräsentativ für alle Proben und wiesen auf eine erfolgreiche cDNA Synthese hin. Die Gelbilder der übrigen Proben sind in Abbildung A.1 im Anhang abgebildet. Die nach der *in vitro* Transkription erfolgte Kontrolle auf dem Bioanalyser wies allgemein, bis auf eine Ausnahme, eine starke Extinktion bei 34 Sekunden bei Sample 10682-6, einen markanten Spitzenpeak nach 39 Sekunden (18S Peak) auf. Ein repräsentatives Elektropherogramm findet sich in Abbildung 3.1, die Elektropherogramme der übrigen Samples im Anhang in Abbildung A.2 und A.3. Die Banden auf einem 1%igem EtBr-Agarosegel ergaben charakteristische Banden. Es konnte daher von intaktem nicht degradierten Material ausgegangen werden. Die photometrisch ermittelten A260/A280-Verhältnisse nach der Aufreinigung lagen im Schnitt bei 1,7. Die gemessenen Konzentrationen waren für die nachfolgenden Schritte ausreichend und lagen nach der Bereinigung um den Gesamt-RNA Anteil zwischen 5,0µg/µl und 7,7µg/µl. Die Werte sind zusammen mit den um die Gesamt-RNA bereinigten Konzentrationen in Tabelle 3.2 aufgeführt. Die Agarosegel-Kontrollen der fragmentierten cRNA ergaben einzelne Banden bei rund 100 Basenpaaren und wiesen auf eine erfolgreiche Fragmentierung hin.

¹Die Isolierung der RNA wurde bereits in der Kooperationsgruppe von Dr. Christina Heinlein, Heinrich-Pette-Institut Hamburg, Abteilung für Tumorstudiologie, durchgeführt.

²Microarray Core Facility, Norwegian Microarray Consortium, Rikshospitalet-Radiumhospitalet HF, University of Oslo

3.1. EXPRESSIONSANALYTIK

Nr.	Sample	Konz.	A260/A280	Konz. n. IVT	korr. Konz.	A260/A280
1	10864-6	5,0	2,16	7,53	7,03	1,72
2	11084-6	3,6	2,15	7,61	7,11	1,73
3	11096-3	3,2	2,14	5,78	5,28	1,72
4	11383-3	6,5	2,15	7,42	6,93	1,73
5	11389-1	6,8	2,14	6,13	5,63	1,71
6	10699-7	2,6	2,16	6,14	5,64	1,71
7	10911-7	3,7	2,20	5,72	5,22	1,63
8	11155-4	3,5	2,17	6,86	6,36	1,65
9	11155-7	4,1	2,22	6,98	6,48	1,65
10	11389-3	2,9	2,19	6,98	6,48	1,70
11	11389-8	2,3	2,20	8,2	7,70	1,66
12	11390-3	3,8	2,19	5,54	5,04	1,68
13	10760-8	3,8	2,14	7,47	6,97	1,69
14	11098-1	2,8	2,14	7,20	6,70	1,64
15	11191-2	2,9	2,14	6,21	5,71	1,65
16	11696-5	5,2	2,14	6,66	6,16	1,66
17	10682-3	4,3	2,14	6,12	5,62	1,71
18	10682-6	2,6	2,14	4,02	3,52	1,76
19	11043-4	7,4	2,16	4,39	3,89	1,76
20	10696-4	5,8	2,16	6,78	6,28	1,83
21	10726-5	5,6	2,16	5,67	5,17	1,76
22	11554-2+3	2,7	2,19	6,60	6,10	1,80
23	WT1	3,1	2,14	6,05	5,55	1,80
24	WT2	2,8	2,13	7,73	7,23	1,72

Tabelle 3.2: Übersicht der Konzentrationen und Reinheit nach der Isolierung der RNA aus den Brusttumoren. Desweiteren sind die Konzentrationen und A260/A280-Verhältnisse nach der *in vitro*-Transkription aufgeführt. Angegeben sind die Konzentrationen in $\mu\text{g}/\mu\text{l}$ und die Reinheit als Quotient der photometrischen Extinktionen bei 260 nm und 280 nm. Das A260/A280-Verhältnis sollte idealerweise zwischen 1,8 und 2,0 liegen. Die Samplenummern, angegeben in der ersten Spalte, dienen der Zuordnung der Samples zu den Agarosegelbanden.

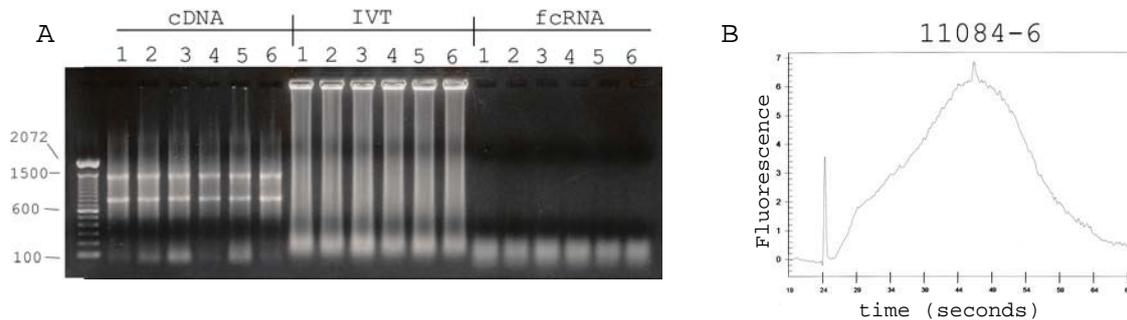


Abbildung 3.1: Links: Repräsentative 1%ige EtBr-Agarosegele der Brusttumorsamples. Aufgetragen wurden die Kontrollaliquots nach cDNA-Synthese, nach der *in vitro*-Transkription und nach erfolgter Fragmentierung. Die Zuordnung der Samples kann Tabelle 3.2 entnommen werden.

Rechts: Elektropherogramm der Tumorprobe 11084-6 auf dem Bioanalytiker nach erfolgter *in vitro*-Transkription. Gut erkennbar ist der nach 39 Sekunden auftretende Spitzenpeak der Verteilung.

Qualitätskontrolle nach Scan Die wichtigsten Parameter zur Qualitätskontrolle der Expressionsarrays sind in Abbildung 3.2 zusammengefasst. Die 3'/5' Verhältnisse der Housekeeping Gene (GAPDH < 1,25; β -Actin < 3) lagen alle im Toleranzbereich. Eine zu starke Degradation der Proben wurde daher ausgeschlossen. Die Bandbreiten der Signalintensitäten sowie die Signalverteilungen waren ebenfalls gut vergleichbar. Die Anzahl der jeweils als „Present“ gewerteten Probesets ließ außerdem auf akzeptable Hintergrundwerte schließen. Insgesamt wurden die Signale der Arrays daher als gut vergleichbar eingestuft.

Differentielle Expression Mittels eines Welch t-Tests und anschließender P-Wert Korrektur nach dem *Benjamini-Hochberg-Verfahren* wurden Vergleiche zwischen verschiedenen Gruppen durchgeführt. Um potentielle Gene aufzuzeigen, die den Differenzierungsprozess beeinflussen, wurden differenzierte und undifferenzierte Tumore verglichen (Vergleich 1). Um den Effekt des *mutP53* zu untersuchen, wurden mono- und bitransgene Samples verglichen (Vergleich 2). Um die Differenz zwischen *mutP53*-transgenen und *mutP53*-freien Tumoren statistisch zu validieren, wurde ein zusätzlicher Vergleich in drei Schritten durchgeführt. Erstens wurden die allgemeinen Determinanten der Tumorgenese bestimmt (Vergleich 3). Dazu wurden alle Tumore mit den Wildtyp-Samples verglichen. Zweitens wurden die durch SV40 T-Antigen, MHC class I T-Zellepitop und gehemmtem p53-Protein bedingten Expressionsänderungen ermittelt (Vergleich 5). Hierzu wurden die monotransgenen Samples mit den Wildtyp-Samples verglichen. Drittens wurde der zusätzliche Effekt

3.1. EXPRESSIONSANALYTIK

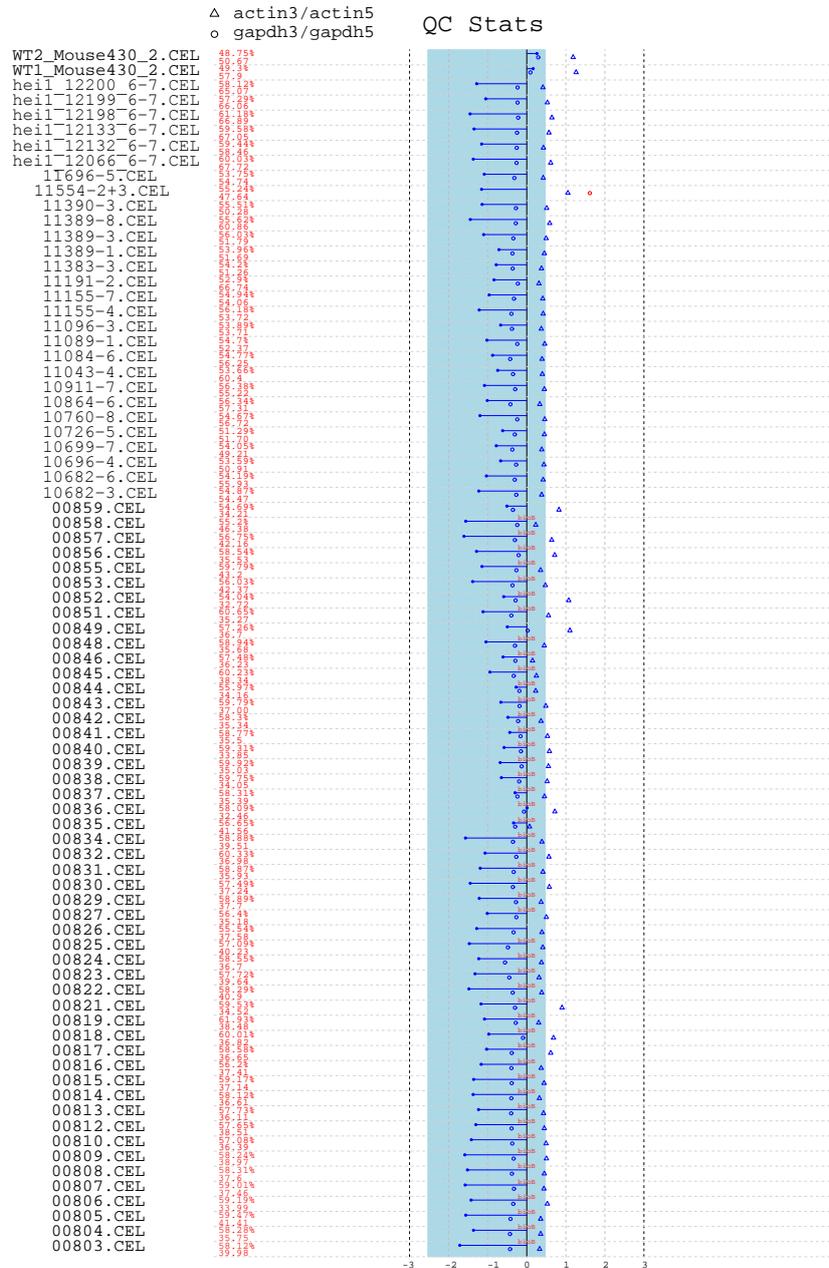


Abbildung 3.2: Qualitätskontroll-Plot der Brustkrebsamples. Dargestellt sind die Abweichungen der logarithmierten GAPDH (\circ) und β -Actin (Δ) 3'/5' Verhältnisse vom Idealwert 0. Werte die die Toleranzgrenzen von 1,25 bzw. 3 nicht überschreiten sind blau dargestellt, Überschreitungen rot. Desweiteren sind die Abweichungen der mittleren Signalintensitäten der Chips vom gemeinsamen Mittelwert als horizontale Balken dargestellt. Die Bandbreite der Abweichungen ist kleiner als 3 (blaue Fläche), die Samples also gut vergleichbar. Weiterhin sind der Anteil der als "Present" gewerteten Probesets in Prozent sowie die mittleren Hintergrundsignale in linearer Form angegeben.

Nr.	Gruppe 1	Gruppe 2	Kurzbeschreibung d. Ziels
1	Diff. Tumore	Undiff. Tumore	Determinanten der Differenzierung
2	Monotransgene Tumore	Bitransgene Tumore	Einfluss des <i>mutP53</i>
3	Alle Tumore	Wildtypen	Allg. Determinanten d. Tumorgenese
4	Bitransgene Tumore	Wildtypen	<i>SV40 T-Antigen</i> + <i>mutP53</i> Einfluss auf Tumorgenese
5	Monotransgene Tumore	Wildtypen	<i>SV40 T-Antigen</i> Einfluss auf Tumorgenese
6	Vergleich 4 - Vergleich 5		Validation Einfluss des <i>mutP53</i>

Tabelle 3.3: Übersicht der durchgeführten Vergleiche mit Kurzbeschreibung des Hintergrundes.

des *mutP53* bestimmt, indem die bitransgenen Samples mit den Wildtyp-Samples verglichen wurden (Vergleich 4). Der zusätzliche Einfluss des mutierten *Trp53* wurde zuletzt durch die Differenzen in den drei Listen an differentiell exprimierten Genen beschrieben. Eine kurze Übersicht aller Vergleiche findet sich in Tabelle 3.3.

Für alle Vergleiche wurde ein Signifikanzniveau von 0,05 nach der *Benjamini-Hochberg*-Korrektur angesetzt. Zwischen differenzierten und undifferenzierten Samples (Vergleich 1) ergaben sich so 2870 differentiell exprimierte Gene. Dabei zeigten 1450 Gene eine Hochregulation und 1420 Gene eine Herunterregulation.

Beim Vergleich zwischen mono- und bitransgenen Samples (Vergleich 2) lagen alle korrigierten P-Werte bei einem Wert von 1, was dafür sprach, daß selbst eine Benjamini-Hochberg-Korrektur in diesem Fall zu hart bzw. die Unterschiede sehr gering waren. Um den Einfluss des *mutP53* näher zu untersuchen wurde daher die Liste der vom p53-Transkriptionsfaktor regulierten Gene³ einem Clustering-Verfahren unterzogen, siehe Abbildung 3.4. Sowohl Samples als auch Gene wurden bei Euklidischer Distanz mit dem *Ward-Verfahren* hierarchisch geclustert. Die p53-Signatur konnte weder mono- und bitransgene Samples noch differenzierte und undifferenzierte Tumore in Gruppen auflösen. Dahingegen wurde eine Gruppierung in die beiden Mausmodelle (T1-Modell bzw. NP8-Modell) sichtbar.

³p53Knowledgebase: <http://p53.bii.a-star.edu.sg>

3.1. EXPRESSIONSANALYTIK

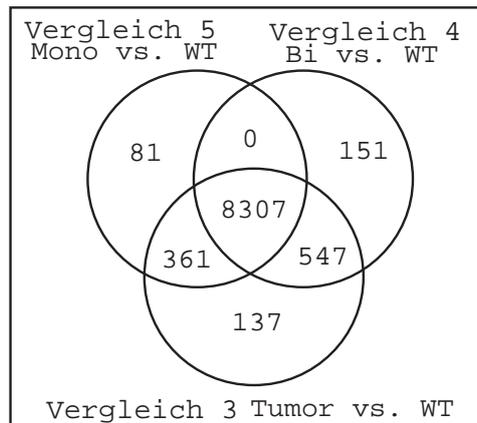


Abbildung 3.3: Brusttumore: Schnittmengen der Genlisten aus den Vergleiche 3 bis 5 (Tumorsamples vs. Wildtypsamples / bitransgene Samples vs. Wildtypsamples / monotransgene Samples vs. Wildtypsamples). Die Listen stimmen zu 97% (8307 Gene) überein. Ein signifikanter Effekt des *mutP53* ist anhand dieser Resultate nicht nachweisbar.

Über eine nähere Evaluation der einzelnen Mausmodelle konnten lediglich 8 signifikant regulierte Gene mit äußerst geringen Regulationsniveaus (SLR=-0,8 bis SLR=0,5) zwischen mono- und bitransgenen NP8-Samples bestimmt werden. *Trp53* erwies sich als einziges Gen stärker reguliert (SLR=3,5). Das T1-Modell wies nach wie vor keine Unterschiede auf.

Im Vergleich aller Tumorsamples mit den Wildtypsamples (Vergleich 3) ergab sich mit 9352 differentiell exprimierten Genen die mit Abstand größte Gruppe. Fast 57% aller erfassten Gene zeigten hier eine differentielle Expression. 5097 Gene zeigten eine Hochregulation, 4255 Gene eine Herunterregulation. Der Einfluss des *SV40 T-Antigen* zusammen mit dem *mutP53* (Vergleich 4) ergab 8749 differentiell exprimierte Gene. Davon waren 4759 Gene hochreguliert und 3990 Gene runterreguliert. Der Effekt des *SV40 T-Antigen* (Vergleich 5) konnte durch 9005 differentiell exprimierte Gene beschrieben werden, von denen 4918 hochreguliert und 4087 runterreguliert waren. Insgesamt deckten sich die Gensets aus den Vergleichen 3-5 mit 97% Übereinstimmung nahezu vollständig. Einen Überblick über die Schnittmengen der letztgenannten Vergleiche gibt Abbildung 3.3. Die Übereinstimmung zeigte sich auch in der Stärke der Genregulationen. Eine Auswahl der jeweils gemäß SLR am stärksten regulierten Gene ist in Tabelle 3.4 aufgeführt.

Funktionsanalyse Die differenzierungsspezifischen Gene, die in Vergleich 1 differentiell exprimiert waren, konnten vor allem mit translatorischen Prozessen in Verbindung gebracht werden. Die GO-Klasse *Translation* war ebenso vertreten wie ribosomenkonstituierende Ge-

Symbol	Affy-ID	CDF-ID	SLR	P-Wert
<i>Vergleich 1 (differenzierte vs. undifferenzierte Tumore)</i>				
Igfbp2	1454159_a_at	16008	3,8	1,00E-05
Glycam1	1424825_a_at	14663	3,5	4,00E-04
St7	1418967_a_at	64213	2,4	1,20E-03
Met	1422990_at	17295	2,4	3,20E-03
Cyp2d10	1418113_at	13101	-2,7	1,0E-04
Slco1a5	1443247_at ; 1449203_at	108096	-2,7	1,0E-06
<i>Vergleich 2 (monotransgene vs. bitransgene Tumore)</i>				
Keine signifikant regulierten Gene				
<i>Vergleich 3 (Tumore vs. Wildtypen)</i>				
Ect2	1419513_a_at	13605	5,0	8,19E-07
Bub1	1416558_at	17279	4,9	2,57E-09
Pbk	1448627_s_at	52033	4,8	8,19E-07
Rbp4	1426225_at	19662	-7,5	1,55E-31
Ces3	1435370_a_at ; 1435371_x_at 1449081_at	104158	-7,5	3,47E-34
<i>Vergleich 4 (bitransgene Tumore vs. Wildtypen)</i>				
Ect2	1419513_a_at	13605	4,6	1,66E-10
Bub1	1416558_at	17279	4,9	9,30E-11
Pbk	1448627_s_at	52033	4,8	4,95E-08
Rbp4	1426225_at	19662	-7,7	9,60E-25
Ces3	1435370_a_at ; 1435371_x_at	104158	-7,7	1,75E-22
<i>Vergleich 5 (monotransgene Tumore vs. Wildtypen)</i>				
Ect2	1419513_a_at	13605	5,1	8,29E-14
Bub1	1416558_at	17279	4,9	5,49E-10
Pbk	1448627_s_at	52033	4,8	3,96E-07
Rbp4	1426225_at	19662	-7,3	3,81E-26
Ces3	1435370_a_at ; 1435371_x_at	104158	-7,3	3,09E-22

Tabelle 3.4: Brusttumore: Auswahl der jeweils am stärksten regulierten Gene in den Gruppenvergleichen der Brusttumore. Angegeben sind die Benjamini-Hochberg korrigierten P-Werte der Gruppenvergleiche sowie die *Signal Log Ratios* (SLR) der Vergleiche. Eine SLR von ± 1 entspricht einer ± 2 fachen Änderung, eine SLR von ± 2 einer ± 4 fachen Änderung. Die Affy-ID bezieht sich auf die offizielle Bezeichnung der betroffenen Transkripte von Affymetrix. Bei der CDF-ID handelt es sich um die interne Bezeichnung der neu zusammengestellten ProbeSets durch AffyProbeMiner. Sie ist mit der Entrez GID der zugeordneten Gene identisch.

3.1. EXPRESSIONSANALYTIK

GO Name	GO Identifier	Korr. P-Wert	Raw P-Wert
structural constituent of ribosome	GO:0003735	0,001	2,379E-16
ribosome biogenesis and assembly	GO:0042254	0,002	5,447E-08
macromolecule biosynthetic process	GO:0009059	0,001	6,031E-14
cellular macromolecule metabolic process	GO:0044260	0,002	4,133E-07
translation	GO:0006412	0,001	1,400E-13

GO Name	# enth. Gene	Freq. in Genliste
structural constituent of ribosome	51	1,78 %
ribosome biogenesis and assembly	31	1,08 %
macromolecule biosynthetic process	125	4,37 %
cellular macromolecule metabolic process	358	12,52 %
translation	91	3,18 %

Tabelle 3.5: Resultate der Funktionsanalyse, mit dem Programm EXPANDER, auf Basis von GO-Annotationen für die differenzierungsspezifischen Gene der Mammakarzinome. Verwendet wurde ein Signifikanzniveau von 0,05 des unkorrigierten P-Wertes. Eine Anreicherung spezifischer Funktionen konnte für den Vergleich zwischen mono- und bi-transgenen Tumoren nicht nachgewiesen werden.

ne der Klassen *Structural constituent of ribosome* und *Ribosomal biogenesis and assembly*. Die GO-Klassen *Macromolecule biosynthetic process* und *Cellular macromolecule metabolic process* machten sogar 4% bzw. 13% der Genliste aus. Eine Übersicht der in dem Vergleich als signifikant angereichert identifizierten Klassen und ihren Signifikanzwerten findet sich in Tabelle 3.5.

Die Funktionsanalyse der Gene, die zwischen den Tumoren und Wildtypen reguliert waren ergab über 30 Funktionsgruppen, die von der Transkription, über Morphogenese und Zelladhäsion bis Immunantworten fast alle Klassen von Regulationsmechanismen umfassten. Die Differenzen in den Genlisten der Vergleiche 3 bis 5 wiesen keinerlei signifikante Funktionsgruppen auf.

Signaturen Für eine Untersuchung eines potentiellen Einflusses des *mutP53* auf die Differenzierung der Tumore wurde ein Clustering der Transkriptionszielgene durchgeführt. Die Liste der p53-regulierten Gene wurde der p53Knowledgebase⁴ entnommen. Das Clustering

⁴p53Knowledgebase: <http://p53.bii.a-star.edu.sg>

der Genexpressionen, siehe Abbildung 3.4, zeigte keine eindeutige Gliederung der Samples in differenzierte und undifferenzierte oder monotransgene bzw. bitransgene Mäuse. Hingegen konnte zwischen den beiden Mausmodellen (T1 bzw. NP8-Modell) unterschieden werden. Deutlich differentiell exprimiert waren die Gene in den Wildtyp-Mäusen bzw. Transgenen, die noch keine Tumore aufwiesen. Eine einheitliche Gliederung der Gene in Gruppen gemäß der Aktivator-/Repressorfunktion des p53 war ebenfalls nicht zu erkennen.

3.1. EXPRESSIONSANALYTIK

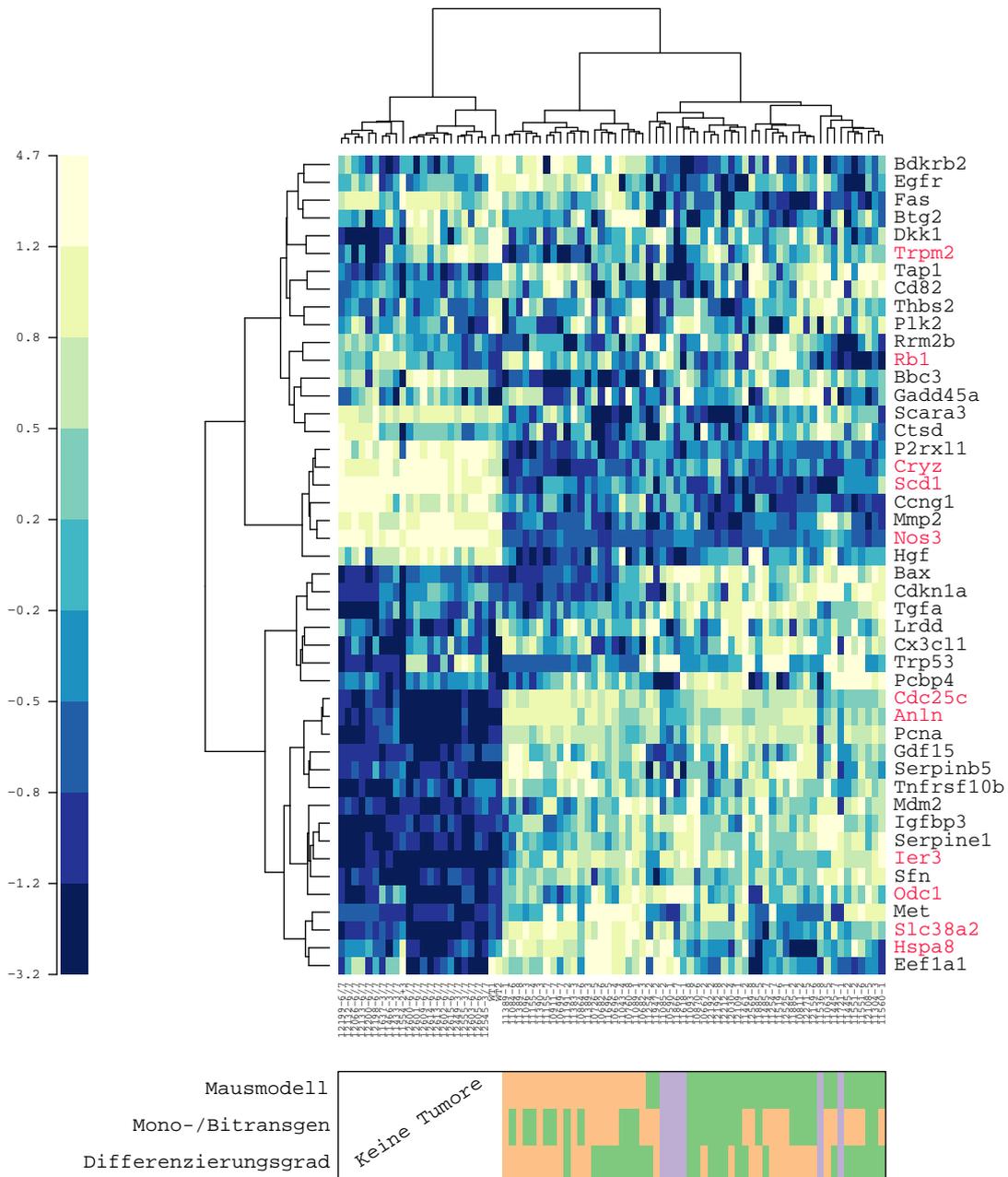


Abbildung 3.4: Brusttumore: p53-Signatur. Bei den Genen der p53-Signatur handelt es sich um die Transkriptionsziele des p53 gemäß der p53Knowledgebase. Die Expressionssignale sind für die Darstellung als z-score standardisiert. Die Werte wurden hierarchisch nach dem Ward-Verfahren geclustert.

Farbkodierung der Gene: Schwarz, p53 aktiviert; Rot, p53 inhibiert.

Farbkodierung der Gruppen: Orange, differenzierte Tumore, monotransgene Tumore, T1-Modell; Grün, undifferenzierte Tumoren, bitransgene Tumore, NP8-Modell; Lila, spontane Mammakarzinome; Weiß Wildtypen bzw. noch keine Tumorentwicklung.

Links: Farbskala für die Kodierung der Expressionswerte.

3.1.2 Neuroblastome

Zur Untersuchung von metastasierungsrelevanten Genen in Neuroblastomen wurden drei Neuroblastomzelllinien mit unterschiedlichem Metastasierungsverhalten in SCID-Mäuse transfiziert. Je drei Tumore wurden pro Zelllinie untersucht. Eine Übersicht über die einzelnen Samples kann Tabelle 3.6 entnommen werden.

Laboranalytik Die Konzentration der eingesetzten Gesamt-RNA lag bei 2µg/µl. Die Qualitätskontrollen wurden analog zu den Expressionsarrays der Brusttumore durchgeführt. Sie ließen durchgehend auf eine gute Qualität der Proben schließen.

Qualitätskontrolle nach Scan Die grundlegenden Parameter der Qualitätskontrolle für die Expressionsarrays der Neuroblastome sind in Abbildung 3.5 zusammengefasst. Die 3'/5' Verhältnisse von *GAPDH* und *β-Actin* lagen alle im Toleranzbereich von maximal 1,25 bzw. 3. Es konnte daher von einer adäquaten RNA-Qualität ausgegangen werden. Die Bandbreite der mittleren logarithmierten Signalintensitäten der Arrays vor der Normalisierung, siehe auch Abbildungen 3.6 und 3.5, war relativ klein. Sie lag unter dem gewünschten Maximalwert von 3. Die durchschnittlichen Hintergrundsignale und der Prozentanteil der als "Present" gewerteten Probesets erfüllten ebenfalls allesamt die gewünschten Kriterien. Es konnte daher von einem suffizienten Hybridisierungsvorgang mit mäßigem Hintergrund sowie guter Vergleichbarkeit der Arrays ausgegangen werden.

Array Name	interne ID	Zelllinie
IMR-208-2.CEL	IMR-208-2	IMR-32
IMR-208-3.CEL	IMR-208-3	IMR-32
IMR-208-6.CEL	IMR-208-6	IMR-32
LAN 1_212-2.CEL	LAN 1_212-2	LAN-1
LAN 1_212-3.CEL	LAN 1_212-3	LAN-1
LAN 1_217-1.CEL	LAN 1_217-1	LAN-1
LS_218-1.CEL	LS_218-1	LS
LS_218-2.CEL	LS_218-2	LS
LS_218-3.CEL	LS_218-3	LS

Tabelle 3.6: Liste der untersuchten Neuroblastom-Samples und ihre Zuordnung zu den Zelllinien.

3.1. EXPRESSIONSANALYTIK

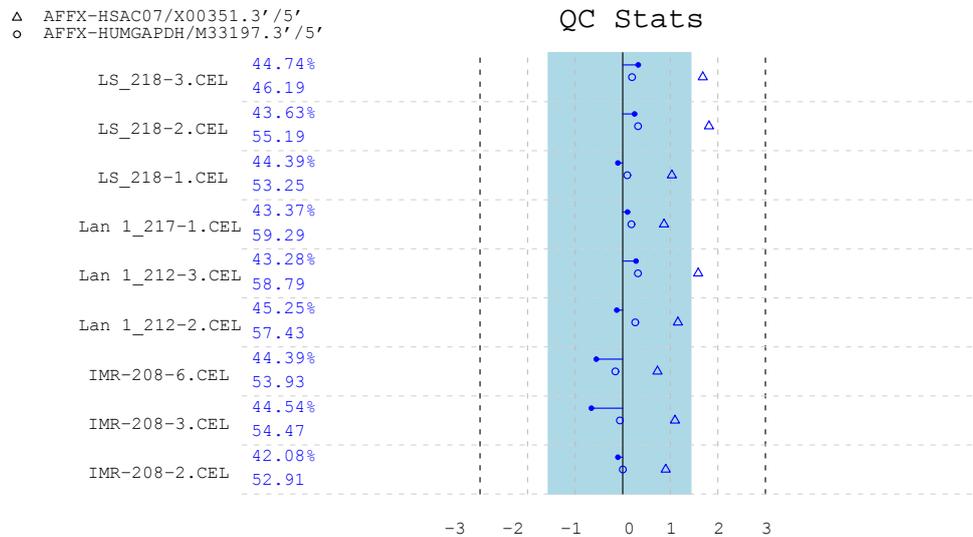


Abbildung 3.5: Qualitätskontrollplot der Neuroblastom-Expressionsarrays. Dargestellt sind die Abweichungen der logarithmierten *GAPDH* (O) und *β-Actin* (Δ) 3'/5' Verhältnisse vom Idealwert 0. Diese Werte sollten einen Wert von 1,25 bzw. 3 nicht überschreiten. Desweiteren sind die Abweichungen der mittleren Signalintensitäten der Chips vom gemeinsamen Mittelwert als horizontale Balken dargestellt. Für eine gute Vergleichbarkeit sollte das maximale Verhältnis der Einzelwerte einen Grenzwert von 3 (blaue Fläche) nicht überschreiten. Für die einzelnen Chips sind der Anteil der als "Present" gewerteten Probesets in Prozent sowie die mittleren Hintergrundsignale in linearer Form angegeben.

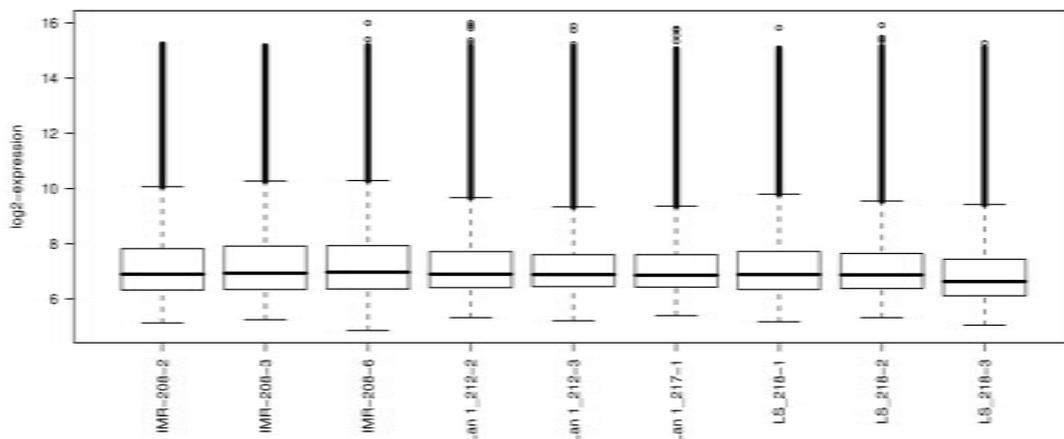


Abbildung 3.6: Neuroblastome: Boxplots der logarithmierten Expressionssignalwerte der Neuroblastomarrays vor der Normalisierung. Die Signalwerte sind bereits vor der Normalisierung gut vergleichbar. Dies sprach für eine gute Signalqualität und geringe technische Varianz.

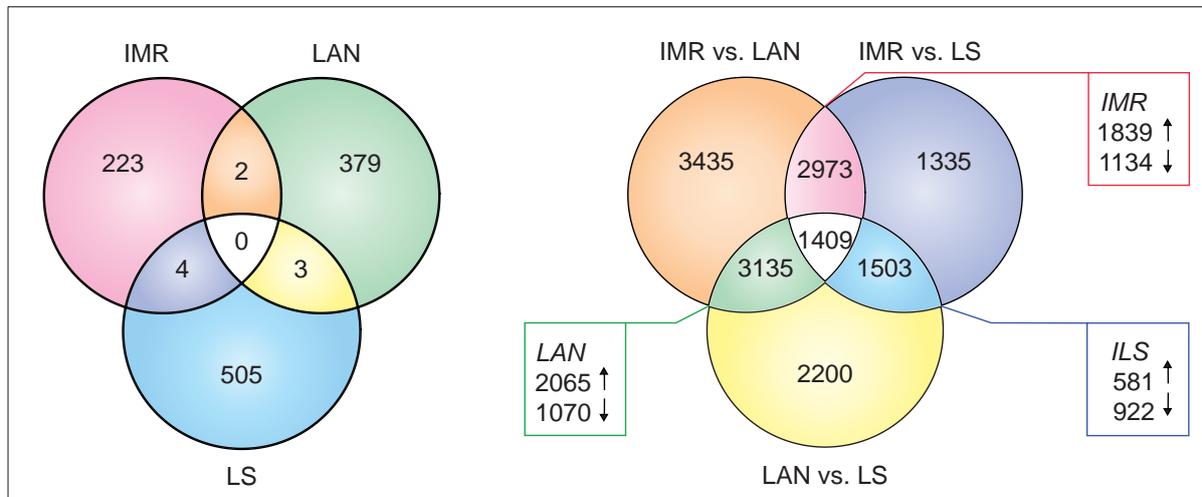


Abbildung 3.7: Links: Venn-Diagramm der Probesets, die innerhalb der einzelnen Neuroblastomlinien auf Expressionsebene eine Varianz von Null aufweisen. Eine Bestimmung von differentieller Expression kann hier mittels t-Test nicht erfolgen, da dieser mathematisch eine Varianz größer Null voraussetzt. Differentielle Expression wurde daher in diesem Fall über einen minimalen Fold Change von 2 festgelegt.

Rechts: Anzahl der Probesets, die mittels Welch t-Test und anschließendem Benjamini-Hochberg Korrektur-Korrekturverfahren bei einem Signifikanzniveau von 0,05 als differentiell exprimiert identifiziert wurden. Die Rosetten bildenden Schnittmengen geben die zelllinienspezifischen Regulationen an. Anhängig findet sich eine Aufteilung der spezifischen Schnittmengen in Hoch- und Herunterregulationen.

Differentielle Expression Zur Identifikation von differentiell exprimierten Genen zwischen den drei Neuroblastomlinien wurde ein Welch t-Test mit anschließender P-Wert Korrektur nach dem *Benjamini-Hochberg-Verfahren* durchgeführt. Einige der Probesets wiesen innerhalb einzelner Gruppen keine Varianz auf. Da bei der Berechnung der t-Statistik die Varianz in den Nenner eingeht und Null-Werte dabei nicht zulässig sind musste eine Analyse dieser 1029 Probesets daher anderweitig erfolgen. Eine Bestimmung putativ unterschiedlich regulierter Gene in dieser Untergruppe wurde anhand eines Grenzwertes für eine Mindeständerung durchgeführt.

Für alle Vergleiche des t-Tests wurde ein Signifikanzniveau von 0,05 nach der Korrektur angesetzt. Beim Vergleich der IMR und der LAN Proben konnten damit 8852 Probesets als differentiell exprimiert identifiziert werden. Der Vergleich zwischen den IMR und den LS Proben ergab 7220 differentiell exprimierte Probesets. Zwischen LAN und LS wurden 8274 Probesets als unterschiedlich stark reguliert identifiziert. Die Schnittmenge dieser drei Genlisten wurde von 1409 Probesets gebildet. Diese Gene hatten ein unterschiedlich

3.1. EXPRESSIONSANALYTIK

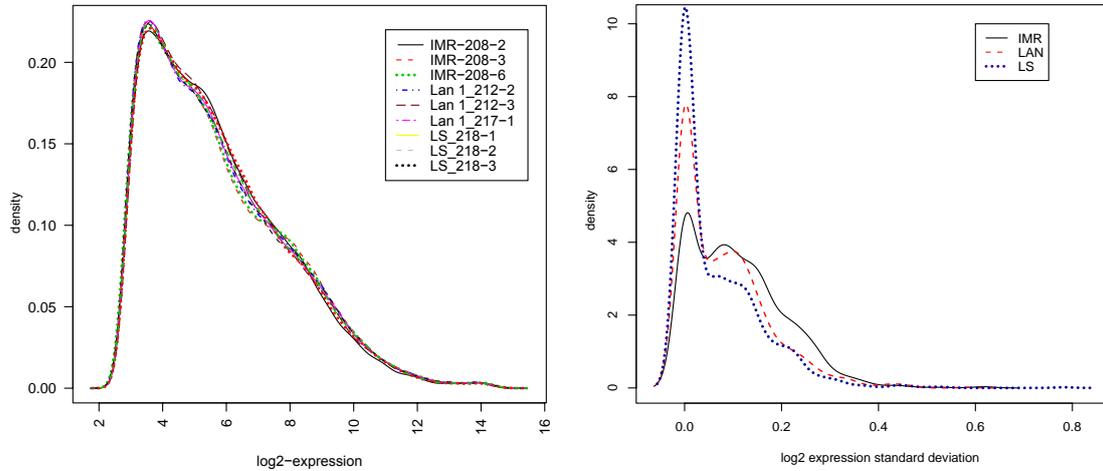


Abbildung 3.8: Neuroblastome: Links: Verteilungen der Expressionswerte nach der Normalisierung mit dem RMA-Verfahren. Rechts: Verteilung der Standardabweichungen innerhalb der drei Gruppen IMR, LAN und LS.

hohes Transkriptionsniveau in allen drei Zelllinien. Desweiteren wurden hier 2973 Probesets identifiziert, die in der IMR-Zelllinie ein unterschiedliches Transkriptionsniveau aufwiesen in Bezug auf einen Vergleich sowohl mit der LAN-Linie als auch der LS-Linie, für die jedoch kein Unterschied zwischen letztgenannten Gruppen nachgewiesen werden konnte. Diese Gruppe von Genen wurde als spezifisch für die IMR-Neuroblastome gewertet. Davon waren 1839 Probesets hoch- und 1134 Probesets runterreguliert. Analog wurden für die LAN bzw. LS Neuroblastome spezifische Genlisten identifiziert. Diese bestanden aus 3135 Probesets bzw. 1503 Probesets respektive. Von den LAN-1 spezifischen Probesets wiesen 2065 eine Hochregulation und 1070 eine Herunterregulation auf. Die LS spezifische Liste bestand aus 581 hochregulierten und 922 runterregulierten Probesets. Eine Übersicht der Größen der Probesetlisten und deren Schnittmengen findet sich in Abbildung 3.7.

Zur Bestimmung der transkriptionellen Unterschiede bei den Genen, die mit einem t-Test nicht untersucht werden konnten, wurde ein 2x FC (SLR=1) im Mittel der Gruppenexpressionswerte als Grenzwert für eine signifikante Änderung festgesetzt. Um die Wahl eines zu kleinen Grenzwertes zu vermeiden, wurde die maximal auftretende Standardabweichung innerhalb der Gruppen berechnet und die Verteilung der Standardabweichungen (für alle Gene) bei diesen Probesets, siehe Abbildung 3.8, herangezogen. Der Maximalwert lag bei 0.78 sowie der Großteil der Standardabweichungen unter 0.2 und damit deutlich unter der Grenze des SLR von 1. Der Vergleich zwischen den IMR und den LS Proben ergab 66 differentiell exprimierte Probesets. Die Vergleiche zwischen IMR und LS bzw. LAN und

Symbol	AffyID	P-Wert IMRvsLAN	P-Wert IMRvsLS	P-Wert LANvsLS	SLR IMRvsLAN	SLR IMRvsLS	SLR LANvsLS
<i>IMR spezifisch</i>							
SPANXA1	220922_s_at	0,0006	0,0010	0,5550	9,6	9,5	-0,1
CD44	212063_at	0,0013	0,0018	0,9131	8,8	8,7	0,0
DKK1	204602_at	0,0155	0,0118	0,5751	7,7	7,8	0,1
MYH10	212372_at	0,0119	0,0131	0,1265	-6,1	-5,8	0,3
PKIB	223551_at	0,0059	0,0045	0,0554	-4,5	-6,1	-1,5
SLAIN1	225619_at	0,0067	0,0013	0,5805	-6,6	-6,4	0,1
<i>LAN spezifisch</i>							
ASCL1	209988_s_at	0,0015	0,3308	0,0094	-8,2	-0,3	7,8
RGS5	218353_at	0,0034	0,7590	0,0018	-7,3	0,1	7,4
MYCN	209757_s_at	0,0354	0,9569	0,0289	-6,9	-0,1	6,8
PRNP	201300_s_at	0,0010	0,0583	0,0011	8,6	0,8	-7,8
ANXA1	201012_at	0,0027	0,2015	0,0032	7,7	-0,5	-8,2
CAV1	212097_at	0,0007	0,1149	0,0022	8,9	-0,5	-9,4
<i>LS spezifisch</i>							
DDAH1	209094_at	0,3035	0,0079	0,0217	0,3	6,0	5,7
FAIM	220643_s_at	0,5039	0,0016	0,0037	-0,2	5,9	6,1
GPD2	225447_at	0,1111	0,0037	0,0057	0,6	5,5	4,9
APCDD1	225016_at	0,2830	0,0039	0,0024	-0,2	-8,3	-8,0
EDIL3	225275_at	0,9171	0,0019	0,0024	0,0	-8,7	-8,7
FABP7	205030_at	0,7234	0,0158	0,0053	-0,1	-8,9	-8,9
<i>Gene, die alle drei Linien unterscheiden</i>							
STEAP1	205542_at	0,0026	0,0334	0,0022	6,7	-2,1	-8,8
CAV2	203324_s_at	0,0013	0,0077	0,0006	6,9	-1,6	-8,5
HMGA2	208025_s_at	0,0060	0,0060	0,0174	9,2	3,3	-5,9
HLA-C	208812_x_at	0,0004	0,0244	0,0073	7,3	2,2	-5,2
RGS4	204338_s_at	0,0084	0,0035	0,0018	-2,4	5,7	8,1
KCTD12	212192_at	0,0248	0,0308	0,0020	-3,6	3,1	6,7
RGS4	204337_at	0,0134	0,0485	0,0375	-1,5	6,5	8,0

Tabelle 3.7: Neuroblastome: Auswahl der jeweils am stärksten regulierten Gene in den Vergleichen zwischen den Neuroblastomzelllinien. Angegeben sind die Benjamini-Hochberg korrigierten P-Werte der Gruppenvergleiche sowie die *Signal Log Ratios* (SLR) der Vergleiche. Eine SLR von ± 1 entspricht einer ± 2 fachen Änderung, eine SLR von ± 2 einer ± 4 fachen Änderung.

3.1. EXPRESSIONSANALYTIK

LS ergaben 45 bzw. 26 Probesets. Die Ergebnisse der beiden Testverfahren wurden für alle weiteren Schritte zusammengefasst. Einige der in den Gruppenvergleichen am stärksten regulierten Gene sind in Tabelle 3.7 aufgeführt.

Funktionsanalyse Bei einer funktionellen Analyse der für die Neuroblastomzelllinien spezifischen Genlisten konnte bei einem Signifikanzniveau von 0.05 eine Anreicherung mehrerer Geneontology-Klassen beobachtet werden. So traten in den LAN-Neuroblastomen vermehrt transkriptionsfunktionelle Gene auf, die den GO-Klassen *RNA binding*, *Nucleic acid binding*, *Nucleotide binding*, *RNA metabolic process* und *RNA Processing* zugeordnet werden konnten.

Die IMR-Neuroblastome wiesen besonders differentielle Expressionen bei Genen auf, die im Zusammenhang mit transkriptionellen Aktivitäten, Zellorganisation sowie Zellzyklus und -metabolismus standen. Hinsichtlich der Transkriptionsaktivitäten waren hier die GO-Klassen *Nucleotide binding*, *Nucleic acid binding*, *DNA binding*, *Transcription*, *Transcription from RNA polymerase II promoter*, *Transcription factor activity* und *RNA metabolic process* vertreten. Die Zellmorphologie und Entwicklung wurde durch die GO-Klassen *Multicellular organismal development*, *Cellular developmental process*, *Cellular structure morphogenesis*, *Anatomical structure development*, *Anatomical structure morphogenesis* und *Cellular component organization and biogenesis* repräsentiert. Die vermehrte Anreicherung von Genen der GO-Klassen *Positive regulation of cellular process*, *Negative regulation of cellular process*, *Regulation of cell cycle* und *Regulation of metabolic process* wiesen auf verstärkte Zellzyklusprozesse gegenüber der LAN und LS-Linie hin. Ein geringer Anteil der IMR spezifischen Gene konnte zusätzlich noch dem sekretorischen Transkriptionsweg, *Secretory pathway*, zugeordnet werden.

Ein verstärktes Vorkommen spezieller Funktionsklassen in den für LS spezifischen Genen konnte nicht nachgewiesen werden. Unter jenen Genen, die in allen drei Zelllinien ein unterschiedlich hohes Expressionsniveau aufwiesen, konnten allerdings die GO-Klassen *Cytoskeletal protein binding*, *Transcription from RNA polymerase II promoter* und *multicellular organismal development* vermehrt beobachtet werden. Eine Gesamtübersicht der beobachteten Geneontology-Klassen und ihrer Signifikanzen findet sich in Tabelle 3.8.

Tabelle 3.8: Resultate der Funktionsanalyse, mit dem Programm EXPANDER, auf Basis von GO-Annotationen für die Neuroblastom-Genlisten. Verwendet wurde ein Signifikanzniveau von 0.05 des unkorrigierten P-Wertes. Eine Anreicherung spezifischer Funktionen konnte für die LS-Neuroblastome nicht nachgewiesen werden.

GO Name	GO Identifier	Korr. P-Wert	Raw P-Wert	# enth. Gene	Freq. in Genliste
<i>IMR spezifisch</i>					
nucleotide binding	GO:0000166	0,015	4,46E-06	649	7 %
nucleic acid binding	GO:0003676	0,002	1,11E-06	978	11 %
DNA binding	GO:0003677	0,037	1,02E-05	672	7 %
transcription	GO:0006350	0,002	3,49E-07	757	8 %
transcription from RNA polymerase II promoter	GO:0006366	0,002	3,72E-07	234	3 %
transcription factor activity	GO:0003700	0,003	5,58E-07	322	4 %
RNA metabolic process	GO:0016070	0,011	3,23E-06	847	9 %
multicellular organismal development	GO:0007275	0,001	1,07E-08	704	8 %
cellular developmental process	GO:0048869	0,001	3,31E-09	693	8 %
cellular structure morphogenesis	GO:0032989	0,003	9,84E-07	163	2 %
anatomical structure development	GO:0048856	0,001	1,79E-07	636	7 %
anatomical structure morphogenesis	GO:0009653	0,001	1,72E-08	351	4 %
cellular component organization and biogenesis	GO:0016043	0,001	6,65E-09	797	9 %
nervous system development	GO:0007399	0,002	3,26E-07	252	3 %
cytoskeletal protein binding	GO:0008092	0,002	2,91E-07	162	2 %
positive regulation of cellular process	GO:0048522	0,008	2,71E-06	312	3 %
negative regulation of cellular process	GO:0048523	0,001	3,17E-08	364	4 %

Fortsetzung auf der nächsten Seite

3.1. EXPRESSIONSANALYTIK

Resultate der Neuroblastom Funktionsanalyse – Fortsetzung

GO Name	GO Identifier	Korr. P-Wert	Raw P-Wert	# enth. Gene	Freq. in Genliste
regulation of cell cycle	GO:0051726	0,038	1,05E-05	194	2 %
regulation of metabolic process	GO:0019222	0,002	2,95E-07	835	9 %
transferase activity	GO:0016740	0,001	3,38E-08	572	6 %
kinase activity	GO:0016301	0,005	2,01E-06	290	3 %
protein kinase cascade	GO:0007243	0,001	8,75E-09	154	2 %
protein transport	GO:0015031	0,020	5,91E-06	232	3 %
protein modification process	GO:0006464	0,001	9,44E-08	545	6 %
secretory pathway	GO:0045045	0,021	5,96E-06	96	1 %
cell death	GO:0008219	0,003	7,09E-07	273	3 %
<i>LAN spezifisch</i>					
RNA processing	GO:0006396	0,050	1,24E-05	75	2 %
RNA binding	GO:0003723	0,001	2,46E-09	117	4 %
nucleotide binding	GO:0000166	0,005	2,26E-06	275	9 %
nucleic acid binding	GO:0003676	0,001	5,10E-08	415	13 %
mRNA metabolic process	GO:0016071	0,034	8,73E-06	54	2 %
RNA metabolic process	GO:0016070	0,022	6,13E-06	349	11 %
oligosaccharide metabolic process	GO:0009311	0,005	2,21E-06	11	0,3 %
<i>Gene, die alle drei Limien unterscheiden</i>					
cytoskeletal protein binding	GO:0008092	0,016	4,58E-06	45	3 %
transcription from RNA polymerase II promoter	GO:0006366	0,002	6,24E-06	61	4 %
multicellular organismal development	GO:0007275	0,025	6,59E-06	158	11 %

Ende Tabelle 3.8

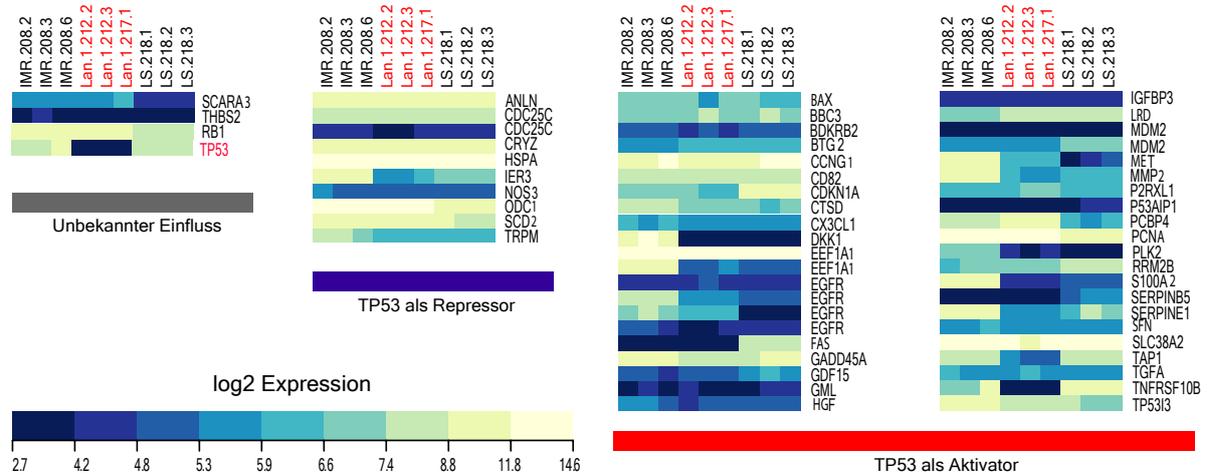


Abbildung 3.9: Neuroblastome: p53 Transkriptionsziele gemäß p53Knowledgebase. Dargestellt sind die log₂ Expressionswerte für alle 9 Neuroblastomsamples. Die LAN-1 Neuroblastome wiesen keine Expression des *TP53* auf. In der Tendenz zeigten diese Tumore niedrigere Expressionswerte bei p53 aktivierten Genen als bei p53 inhibierten Genen. Dennoch wiesen einige p53 aktivierte Gene, wie *PCNA* oder *SLC38A2*, eine deutliche Expression auf. p53 konnte demnach keinen begrenzenden Faktor für die Transkription dieser Gene darstellen.

Signaturen Ausgehend von der differentiellen Expression und späteren Ergebnissen der CNV-Analytik sowie der MAPK-TF-Chips konnte ein deutlich niedrigeres Expressionsniveau des Tumorsuppressorgenes *TP53* in LAN-Zellen gegenüber IMR oder LS-Zellen beobachtet werden. Anhand der vom p53 Transkriptionsfaktor regulierten Gene⁵ wurde die p53-Signatur untersucht, siehe Abbildung 3.9. Die LAN-1 Neuroblastome wiesen keine Expression des p53 auf. Generell zeigten diese Tumore niedrigere Expressionswerte bei p53 aktivierten Genen als bei p53 inhibierten Genen. Allerdings wiesen einige p53 aktivierte Gene, wie *PCNA* oder *SLC38A2*, eine hohe Expression auf. Desweiteren waren lediglich 19 der 50 Signaturgene zwischen LAN und IMR bzw. LS signifikant reguliert, z.B. *RB1*, *FAS* oder *EGFR*. Die aktivierende Funktion des p53 Faktors musste bei diesen Genen demnach von anderen Faktoren übernommen worden sein. Als alternative Erklärung könnte p53 bei diesen Genen eher als Repressor wirken, anstatt, wie von der p53Knowledgebase beschrieben, als Aktivator.

⁵Quelle: p53Knowledgebase, <http://p53.bii.a-star.edu.sg>

3.1. EXPRESSIONSANALYTIK

GenMapp Ergänzend zu der vorangehenden Funktionsanalyse wurden für die IMR und LAN Zelllinien einige Karten von Signaltransduktionswegen und molekularen Funktionen mit dem Programm GenMapp erstellt. Dargestellt wurden Gene, die im Vergleich zur LS Zelllinie als signifikant reguliert identifiziert wurden. Dadurch konnten zusätzlich auffällige Unterschiede zwischen den beiden Zelllinien im *MAPK signaling pathway*, *TGF-Beta signaling pathway* und *TNF alpha/NF kappaB pathway* ausgemacht werden. Weiterhin bestanden Unterschiede in den Klassen *RNA splicing*, *Response to DNA damage stimulus*, *Protein folding*, *Proteasome degradation*, *mRNA processing* sowie *Oxidative stress*.

Beim *Proteasome Degradation* Pathway waren die MHC Klasse I Gene sowie die Gene *PSME1*, *PSMB8* und *PSMB9* in IMR Zellen im Vergleich zu LS hochreguliert. In LAN Zellen waren sie runterreguliert. Die *Oxidativem Stress* verwandten Gene *CYBA*, *XDH*, *NFIX* und *MT1X* waren in IMR Zellen hochreguliert und in LAN Zellen runterreguliert. Weiterhin unterschieden sich die Zellen unter anderem in der Expression von *FOS*, *GPX3*, *HMOX1*, *GSR* und *NQO1*.

Beim *TGF-Beta signaling pathway* waren in den IMR Zellen die Gene *NOG*, *TGFB1*, *TGFBR3*, *TGFBR2*, *TGFBR1*, *FKBP1A*, *SMAD2*, *SMAD3*, *RUNX3*, *SERPINE1* sowie *TFE3* hochreguliert. Bis auf *TGFBR1* waren diese Gene in den LAN Zellen im Vergleich zu LS nicht reguliert oder runterreguliert. So war bei *TGFB1* und *TGFBR2* eine Herunterregulation zu beobachten, ebenso wie für *EGF*, *ENG*, *SMAD6*, *MAPK9*, *RUNX3*, *LEF1*, *STAT3*, *JUN* und *FOS*.

Hochregulationen in IMR Zellen im Vergleich zu LS konnten beim *MAPK signaling pathway* bei *RASGRF2*, *PRKCD*, *PTPRR*, *MAP3K14*, *MAP4K4*, *MAP2K2*, *TGFBR1,2*, *PTPRR*, *RAC2*, *DUSP4*, *DUSP6*, *DUSP7* und *MINK1* sowie *ELK1* und *AKT1* beobachtet werden. Die Gene *EGF*, *NRAS*, *PRKCB1*, *BRAF*, *TMEM37*, *MAPT*, *MYC*, *HSPA5*, *MAPK7*, *MAPK8*, *ATF2*, *CASP1*, *CASP6*, *CASP7*, *FAS*, *MAP2K5*, *MAP2K6* und *NR4A1* wiesen ein niedrigeres Expressionsniveau auf als in LS. IMR und LAN gemeinsam waren dabei die Regulationen der Gene *BDNF*, *DUSP4*, *EGFR* und *PRKCB1*, *AKT*, *ELK1*, *FAS* sowie der Caspasen 1, 6 und 7. Hingegen waren die Gene *JUN*, *FOS*, *TP53*, *DUSP5*, *DUSP6*, *MAPK9*, *HSPA5*, *FLNA*, *NTF5*, *FAS*, *TGFBR2*, *RASA1* und *MAP3K12* runterreguliert. Die Gene *CDC25B* und *MAP2K1IP1* waren hochreguliert. Insgesamt konnten damit verstärkte pro-proliferative und anti-apoptotische Regulationen in den IMR-32 Neuroblastomen beobachtet werden. Die Regulationen in den LAN-1 Neuroblastomen deuteten auf reduzierte Zellapoptose hin. Die erzeugten GenMapp-Karten finden sich in den Abbildungen 3.10 bis 3.15.

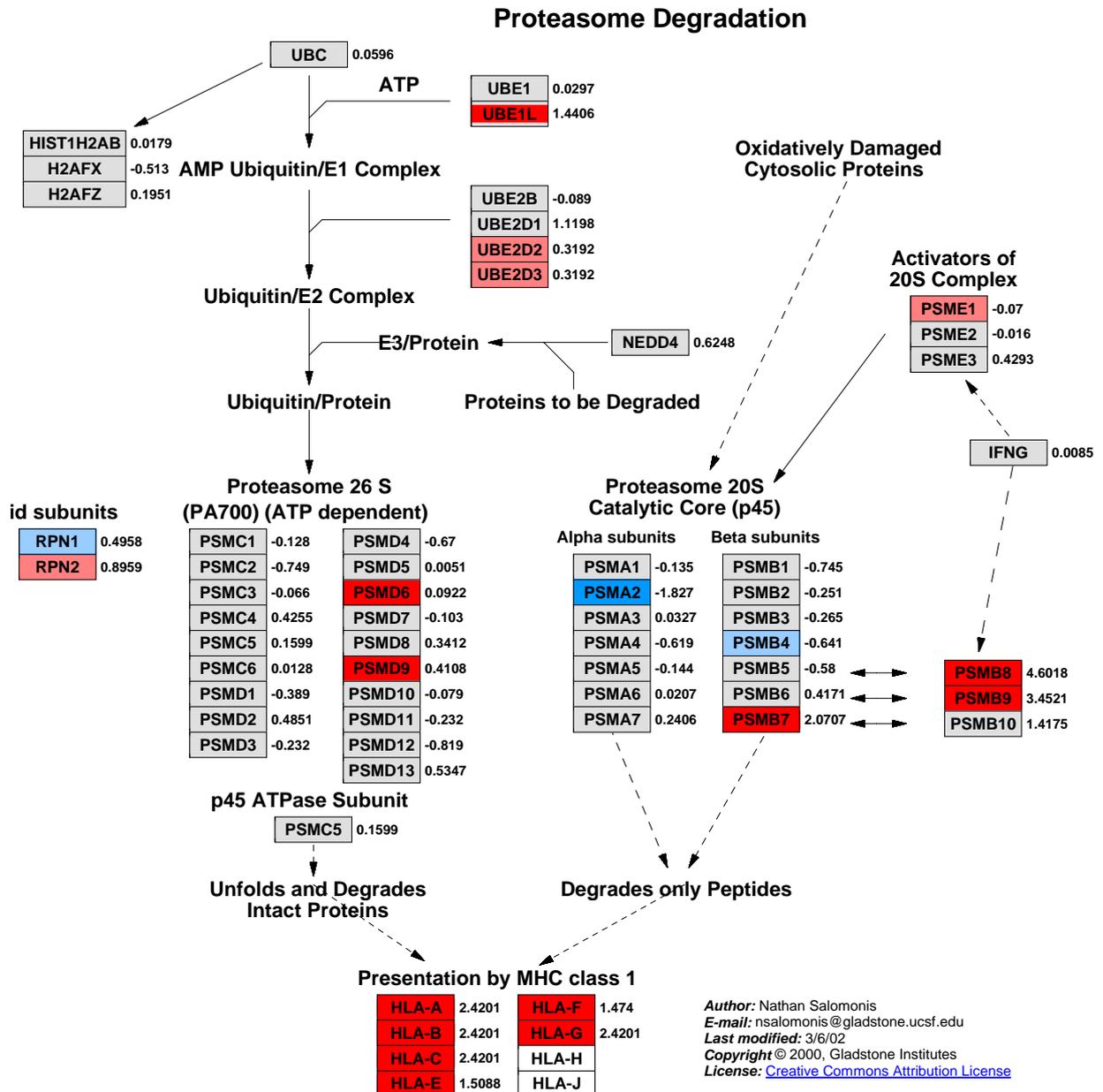


Abbildung 3.10: IMR 32: Mit GenMapp erstellte Karte der Proteasomdegradation regulierenden Gene. Hervorgehoben sind Gene, die im Vergleich zu LS signifikant (P-Wert < 0,05) reguliert sind.

Farbkodierung: Rot, Hochregulation mit SLR > 1; Hellrot, Hochregulation mit SLR < 1; Blau, Herunterregulation mit SLR > 1; Hellblau, Herunterregulation mit SLR < 1.

3.1. EXPRESSIONSANALYTIK

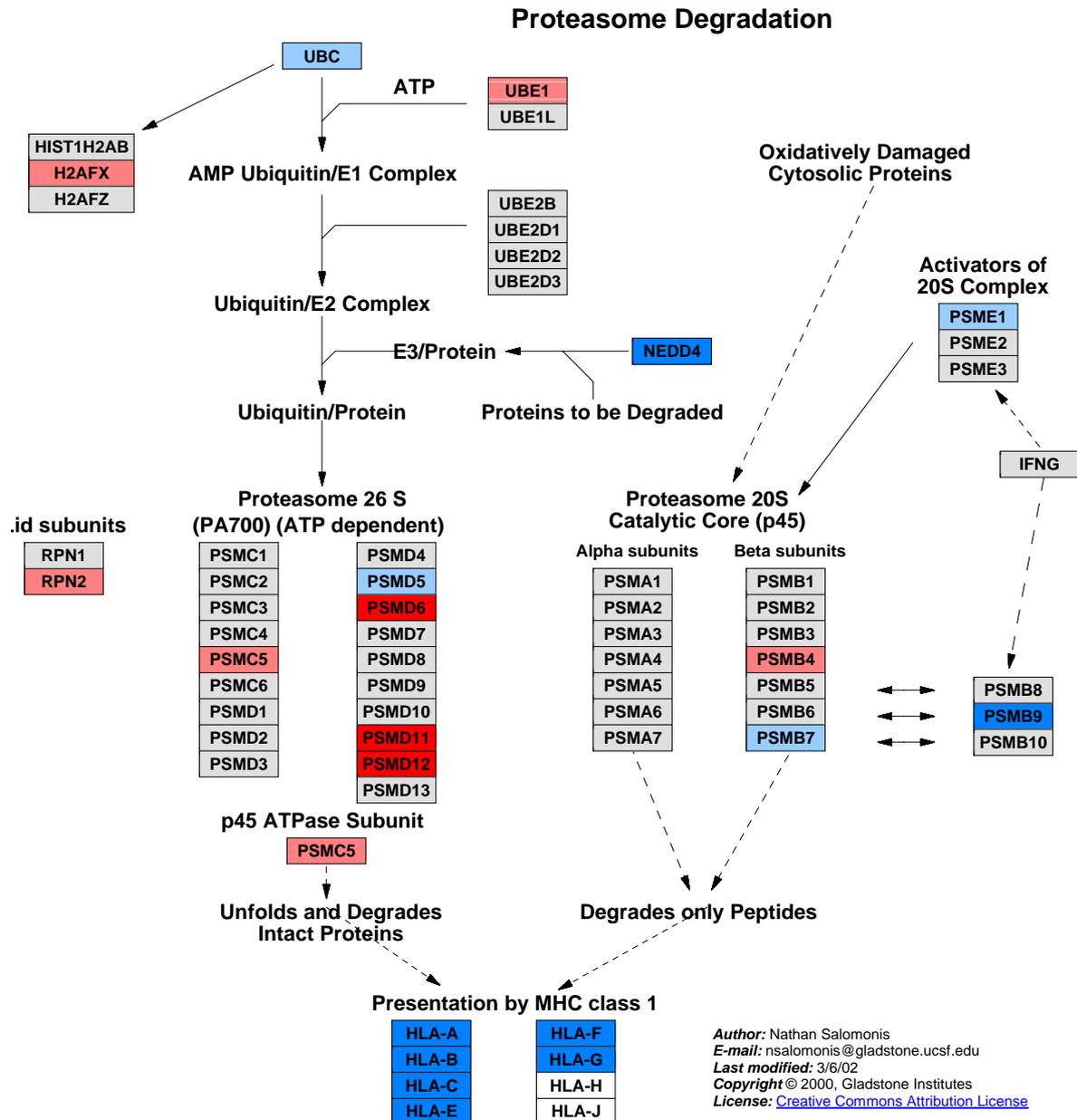


Abbildung 3.11: LAN 1: Mit GenMapp erstellte Karte der Proteasomdegradation regulierenden Gene. Hervorgehoben sind Gene, die im Vergleich zu LS signifikant (P-Wert < 0,05) reguliert sind.

Farbkodierung: Rot, Hochregulation mit SLR > 1; Hellrot, Hochregulation mit SLR < 1; Blau, Herunterregulation mit SLR > 1; Hellblau, Herunterregulation mit SLR < 1.

Author: Adapted from KEGG by Sebastian Burel
 E-mail: sburel@neurocrine.com
 Converted to human by GenMAP.org
 License: Creative Commons Attribution License

MAPK signaling pathway

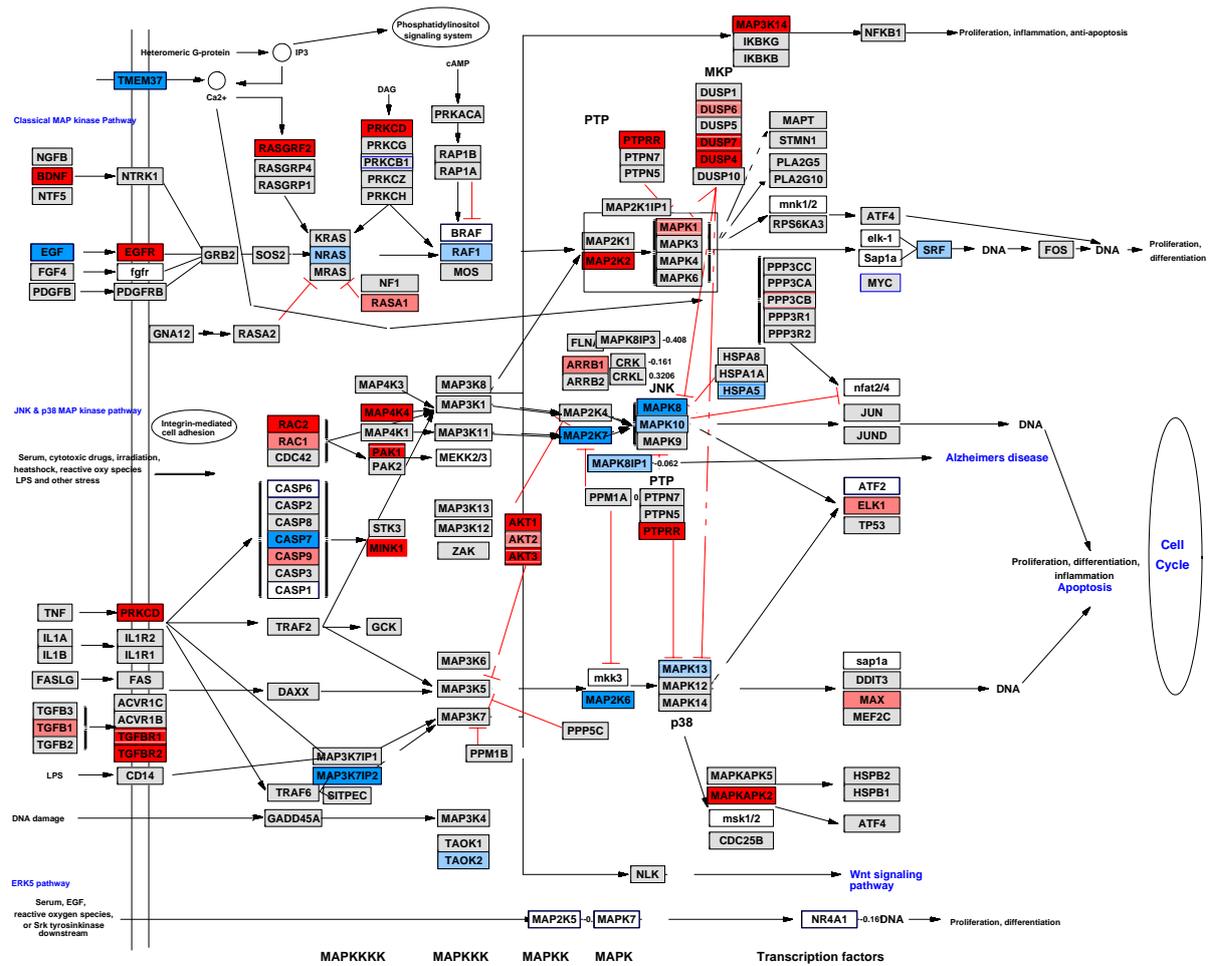


Abbildung 3.12: IMR 32: Mit GenMapp erstellte Karte des MAPK Signaling Pathways. Hervorgehoben sind Gene, die im Vergleich zu LS signifikant ($P\text{-Wert} < 0,05$) reguliert sind.

Farbkodierung: Rot, Hochregulation mit $SLR > 1$; Hellrot, Hochregulation mit $SLR < 1$; Blau, Herunterregulation mit $SLR > 1$; Hellblau, Herunterregulation mit $SLR < 1$.

3.1. EXPRESSIONSANALYTIK

Author: Adapted from KEGG by Sebastian Burel
 E-mail: sburel@neurocrine.com
 Converted to human by GenMAP.org
 License: Creative Commons Attribution License

MAPK signaling pathway

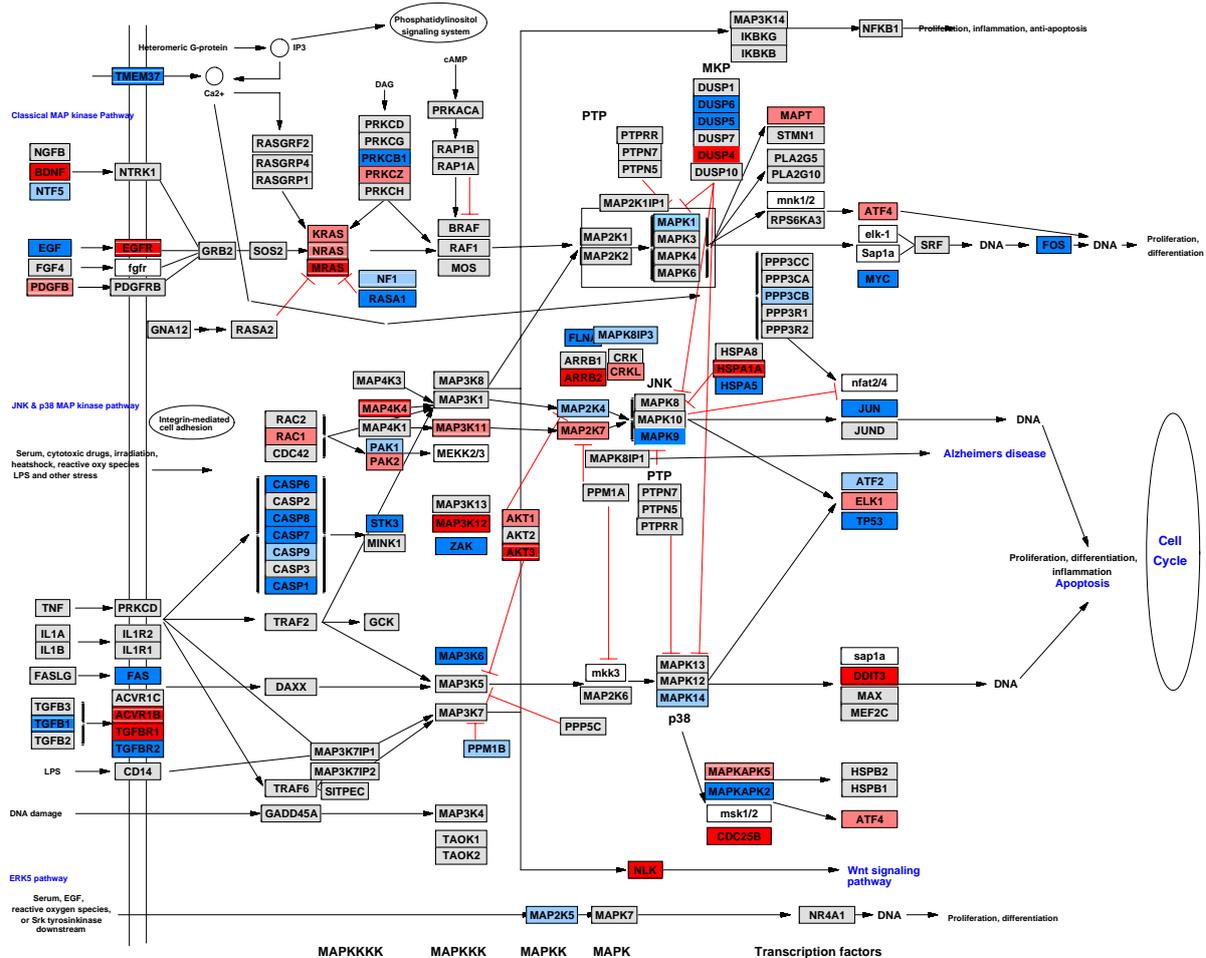


Abbildung 3.13: LAN 1: Mit GenMapp erstellte Karte des MAPK Signaling Pathways. Hervorgehoben sind Gene, die im Vergleich zu LS signifikant ($P\text{-Wert} < 0,05$) reguliert sind.

Farbkodierung: Rot, Hochregulation mit $SLR > 1$; Hellrot, Hochregulation mit $SLR < 1$; Blau, Herunterregulation mit $SLR > 1$; Hellblau, Herunterregulation mit $SLR < 1$.

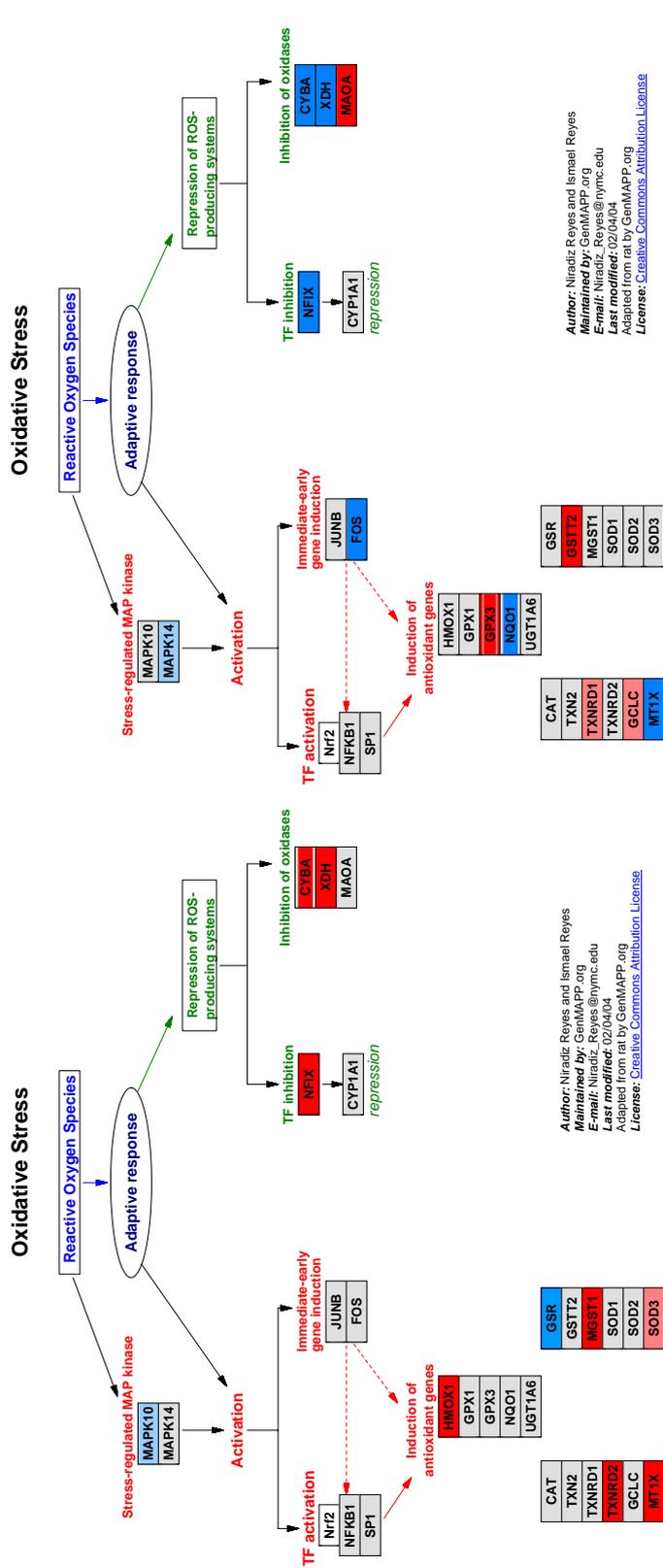


Abbildung 3.14: IMR 32 links, LAN 1 rechts: Mit GenMapp erstellte Karte der oxidativem Stress verwandten Gene. Hervorgehoben sind Gene, die im Vergleich zu LS signifikant (P-Wert < 0,05) reguliert sind. Farbkodierung: Rot, Hochregulation mit SLR > 1; Hellrot, Hochregulation mit SLR < 1; Blau, Herunterregulation mit SLR > 1; Hellblau, Herunterregulation mit SLR < 1.

3.1. EXPRESSIONSANALYTIK

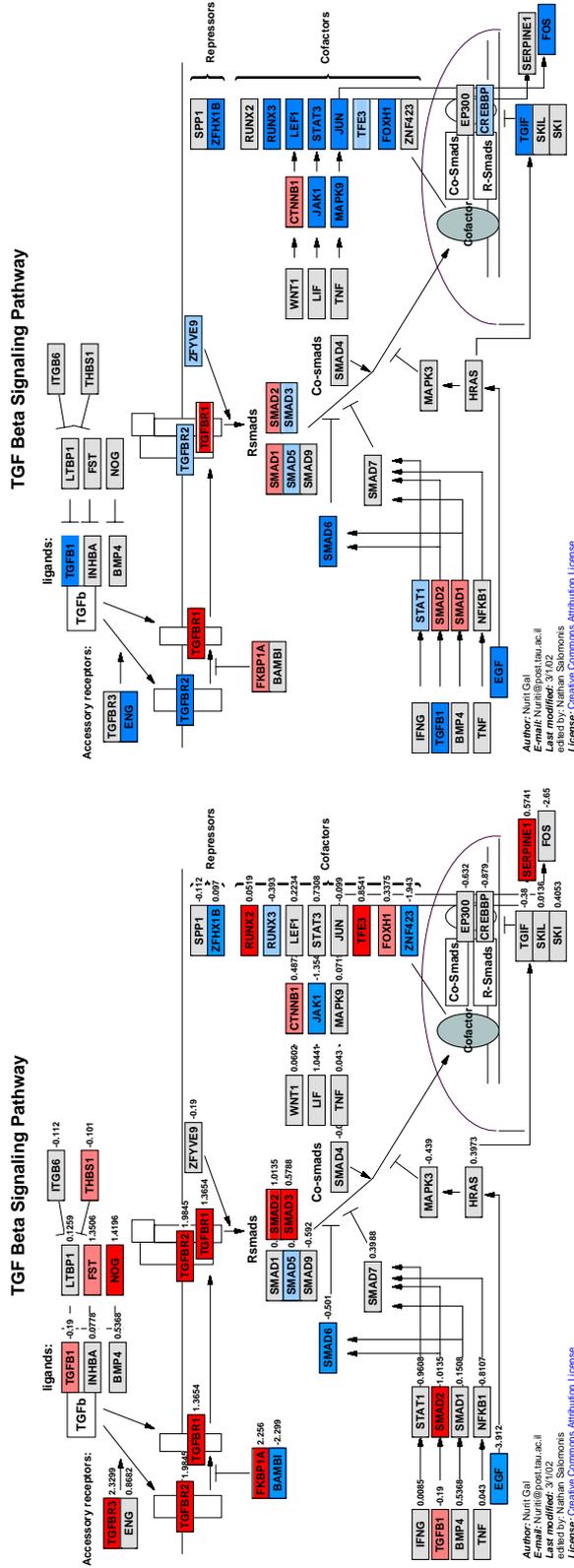


Abbildung 3.15: IMR 32 links, LAN 1 rechts: Mit GenMapp erstellte Karte des TGF Beta Signaling Pathways. Hervorgehoben sind Gene, die im Vergleich zu LS signifikant (P-Wert < 0,05) reguliert sind. Farbkodierung: Rot, Hochregulation mit SLR > 1; Hellrot, Hochregulation mit SLR < 1; Hellblau, Herunterregulation mit SLR > 1; Hellblau, Herunterregulation mit SLR < 1.

3.2 Copy-Number-Variations & Genotypisierung

Die Bestimmung von genomischen Aberrationen in den SNP-Arrays und CGH-Arrays sollte die Expressionsanalytik der beiden Tumormodelle ergänzen. Im Folgenden werden Probenqualität und die identifizierten Aberrationen beschrieben.

3.2.1 Brusttumore: CGH-Arrays

Zur Identifizierung von Amplifikationen und Deletionen in den Mammakarzinomen wurden 56 der 80 Sampleproben mittels array-CGH Technik untersucht. Die 21 Samples des T1-Mausmodells wurden auf Agilent-Chips einer älteren Generation in Amsterdam hybridisiert. Die 35 Samples des NP8-Modells und der spontanen Mammakarzinome wurden auf Agilents 44k Arrays in Oslo hybridisiert.

Qualitätskontrolle nach Scan Die Qualität der Amsterdamer array-CGH Chips präsentierte sich schlechter als jene der neueren 44k Arrays von Agilent. Etliche der auf dem Chip befindlichen Oligos wurden von der in Amsterdam verwendeten BlueFuse-Software aufgrund zu schlechter Signalqualität maskiert. Diese standen für die weiteren Analyse-schritte nicht mehr zur Verfügung. Die Raten der nach der Qualitätsfilterung noch verwendeten Sonden lag zwischen 65% und 81%. Die ausführliche Liste der Raten findet sich in Tabelle 3.9. Bei den 35 Osloer Samples auf den 44k Arrays konnten hingegen alle Oligos verwendet werden.

Sample	Anteil %	Sample	Anteil %	Sample	Anteil %
11390-3	75,58	10682-6	79,38	10911-7	75,61
11389-3	65,8	10696-5	71,44	11155-4	64,73
11389-9	79,29	10696-4	75,72	11155-7	68,56
11383-3	68,87	10726-5	71,39	10760-8	68,73
10699-7	71,13	10864-6	78,79	11089-1	72,55
11389-1	74,67	11084-6	72,74	11191-2	66,01
10682-3	72,02	11096-3	80,59	11043-4	75,16

Tabelle 3.9: Anteil der Oligos mit verwertbarem Signal auf den 21 array-CGH-Chips aus Amsterdam (T1-Modell). Ausschließlich Oligos mit verwertbarem Signal flossen in die weiteren Auswertungsschritte ein. Oligos mit nichtverwertbarem Signal wurden von der BlueFuse-Software automatisch erkannt und maskiert.

3.2. COPY-NUMBER-VARIATIONS & GENOTYPISIERUNG

	Differenziert	Undifferenziert
Amplifiziert	8	23
Nicht amplifiziert	19	2

Tabelle 3.10: Kontingenzmatrix der Amplifikationshäufigkeiten des *Met*-Locus in den Brusttumoren. Ein Exakter Fisher-Test ergab einen P-Wert von $3,9 \cdot 10^{-6}$.

Als Qualitätskontrollmaß der Signalbandbreiten wurde für jede Sampleprobe der MAD-Werte (Mean Absolute Deviation⁶) bestimmt. Die MAD-Werte aller Sampleproben können Tabelle 3.11 entnommen werden. Die Werte lagen im Schnitt bei 0,1944 und wurden daher durchgehend als gut eingestuft. Die Modalwerte der Segmentverteilungen sowie die Standardabweichungen der Signale sind ebenfalls in Tabelle 3.11 aufgeführt.

Aberrationsbestimmung Zur Bestimmung von Aberrationen wurden die Segmente um die Modalwerte der Segmentmittelwertverteilungen zentriert. Diese Werte lagen zwischen -0,1160 und 0,1317. Die Standardabweichungen der 50% Quantile lagen zwischen 0,0466 und 0,1645, die der 90% Quantile zwischen 0,0564 und 0,2125. Eine Klassifikation als Gain bzw. Loss erfolgte sofern der Segmentmittelwert die ± 2 fache Standardabweichung des 50% Quantils der unsegmentierten log₂-Signale über- bzw. unterschritt. Als Amplifikation bzw. Deletion wurden Segmente eingestuft, deren Mittelwert die ± 3 fache Standardabweichung (90% Quantil) über- bzw. unterschritt.

Aberrationen Nach der Segmentierung und Bestimmung von Gains und Losses zeigte sich kein einheitliches Aberrationsprofil. Vielmehr konnten starke Schwankungen zwischen den einzelnen Proben beobachtet werden. Vergleichsweise homogen zeigte sich allerdings eine starke Amplifikation des *Met*-Locus in den undifferenzierten Tumoren (Stufen G3 und G4), die in den differenzierten Tumoren (Stufen G1 und G2) deutlich seltener zu beobachten waren. Ein Exakter Fisher-Test ergab hier einen signifikanten P-Wert von $3,9 \cdot 10^{-6}$. Tabelle 3.10 zeigt die Frequenz der *Met*-Locus-Amplifikationen in den differenzierten und undifferenzierten Tumorproben.

⁶Zur Berechnung siehe Abschnitt 2.3

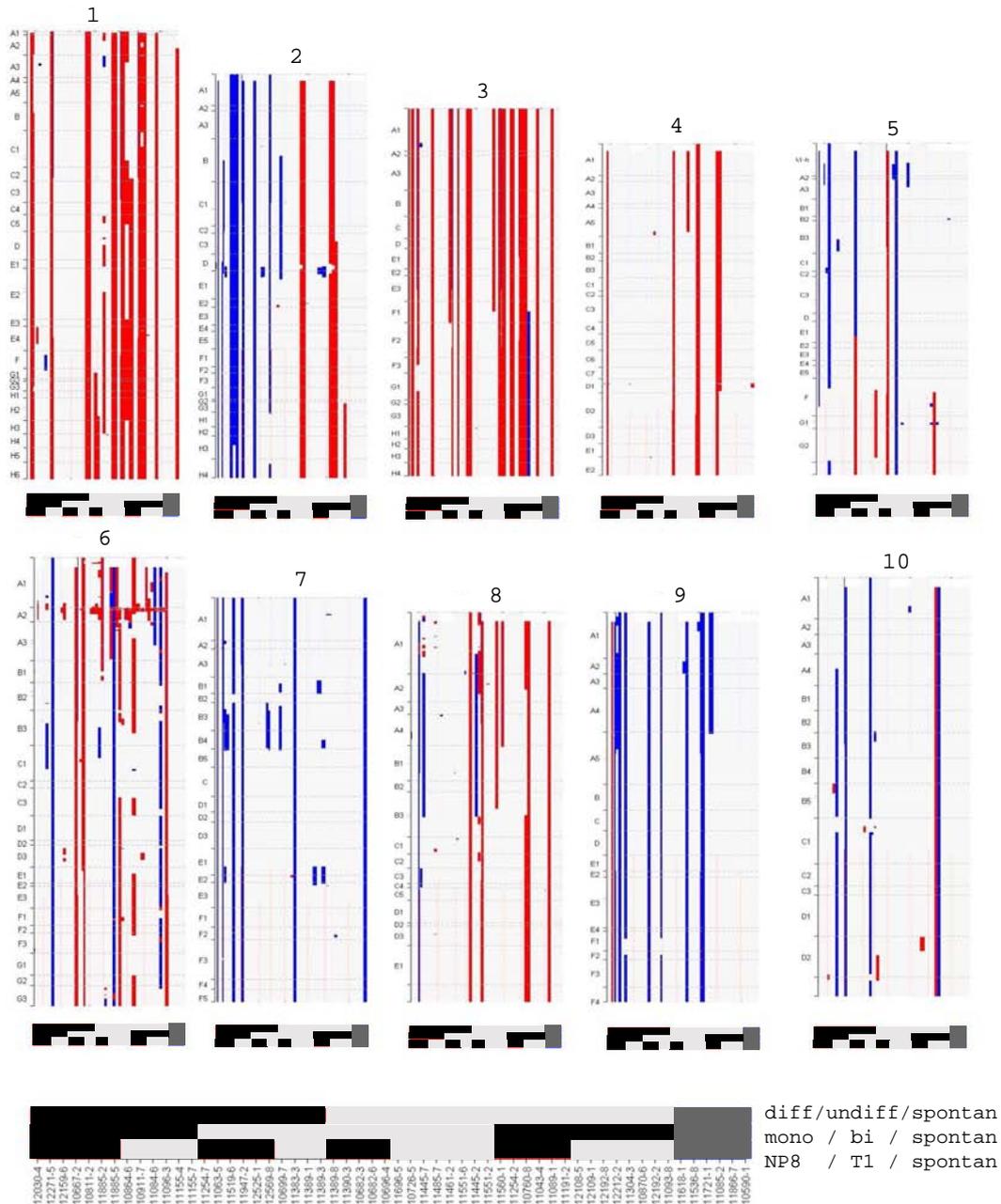


Abbildung 3.16: Aberrationsprofile von 56 Brustkrebsamples. Abgebildet sind Chromosomen 1 bis 10.

Farbkodierung Aberrationen: Rot, Gains; Blau, Losses; Grau keine Aberration.

Farbkodierung Differenzierung: Schwarz, differenzierte Tumore; Grau, undifferenziert.

Farbkodierung mutP53-Expression: Schwarz, monotransgen; Grau, bitransgen.

Farbkodierung Mausmodell: Schwarz, NP8-Modell; Grau, T1-Modell. Die spontanen Mammakarzinome sind jeweils immer dunkelgrau markiert.

3.2. COPY-NUMBER-VARIATIONS & GENOTYPISIERUNG

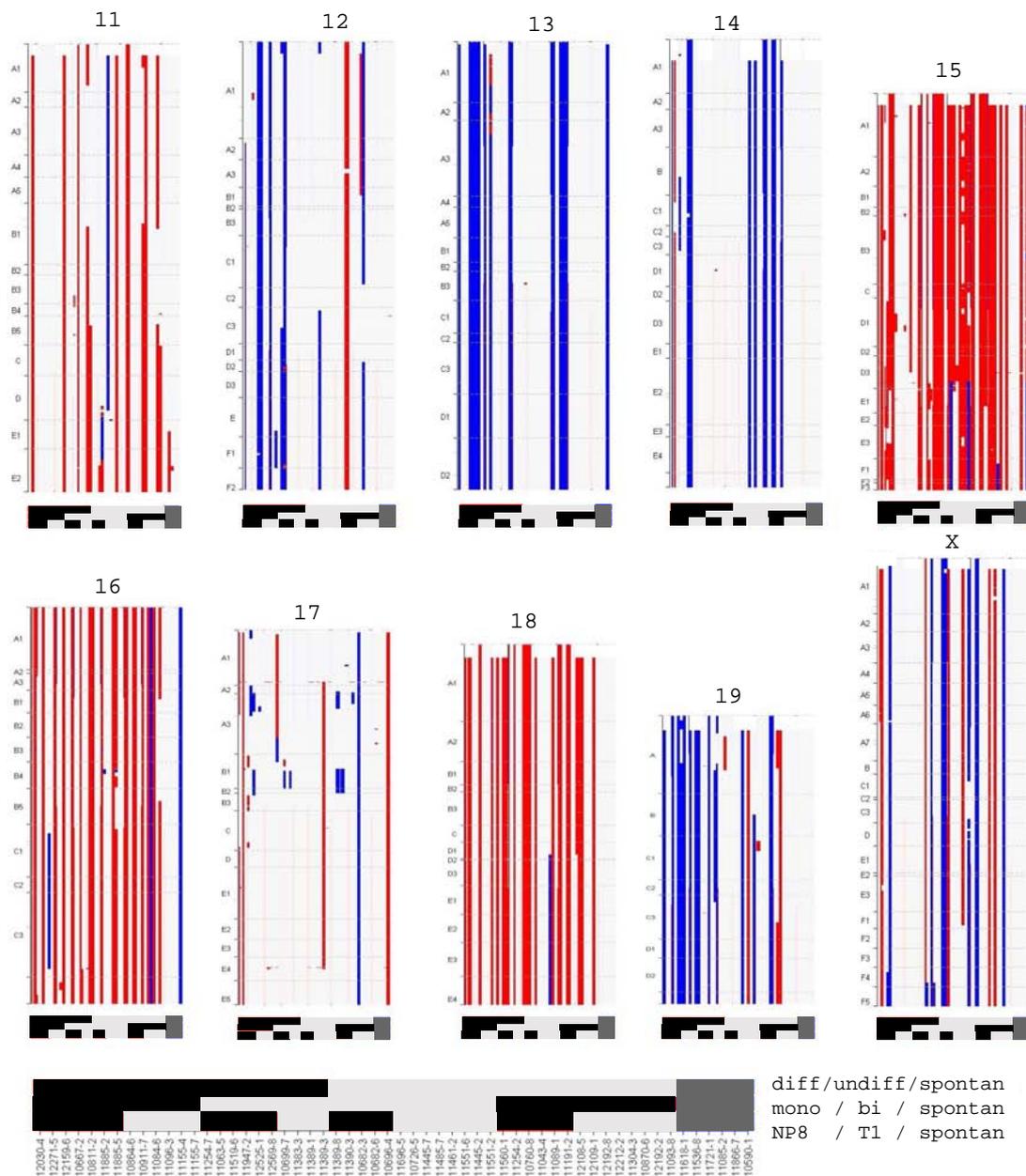


Abbildung 3.17: Aberrationsprofile von 56 Brustkrebssamples. Abgebildet sind Chromosomen 11 bis 19 und ChrX.

Farbkodierung Aberrationen: Rot, Gains; Blau, Losses; Grau keine Aberration.

Farbkodierung Differenzierung: Schwarz, differenzierte Tumore; Grau, undifferenziert.

Farbkodierung mutP53-Expression: Schwarz, monotransgen; Grau, bitransgen.

Farbkodierung Mausmodell: Schwarz, NP8-Modell; Grau, T1-Modell. Die spontanen Mammakarzinome sind jeweils immer dunkelgrau markiert.

Sample	MAD-Wert	SD 50% Q.	SD 90% Q.	Modalwert
10590-1	0,1164	0,0637	0,0828	0,0053
10667-2	0,2279	0,1264	0,162	-0,0204
10682-3	0,1774	0,0992	0,1219	0,0124
10682-6	0,1865	0,1039	0,1326	0,1317
10696-4	0,2218	0,1235	0,1539	0,0467
10699-7	0,2967	0,1645	0,2125	0,0755
10726-5	0,2831	0,1556	0,2084	0,0472
10760-8	0,201	0,1129	0,14	0,0537
10811-2	0,2287	0,1287	0,1485	0,0096
10864-6	0,1894	0,106	0,1362	0,0009
10870-6	0,2042	0,1148	0,1341	-0,0391
10911-7	0,166	0,0928	0,1082	-0,0381
11043-4	0,2183	0,1223	0,1547	-0,0158
11063-5	0,2005	0,1121	0,133	0,0173
11084-6	0,2079	0,1153	0,1359	-0,0523
11085-2	0,0856	0,0477	0,0592	-0,0005
11089-1	0,216	0,1213	0,1452	0,0398
11093-8	0,1494	0,0819	0,1178	-0,0066
11096-3	0,2041	0,1163	0,114	-0,0141
11155-4	0,2438	0,1369	0,1475	-0,0163
11155-7	0,1793	0,101	0,1058	0,0001
11191-2	0,193	0,1065	0,1328	-0,0323
11254-2	0,1956	0,1084	0,1362	-0,0162
11254-7	0,2141	0,1204	0,1322	0,0948
11304-3	0,1589	0,0876	0,1171	-0,0003
11383-3	0,1553	0,0874	0,1013	-0,0001
11389-1	0,1827	0,1025	0,1218	0,0512
11389-3	0,1932	0,1086	0,1357	0,1252
11389-8	0,224	0,1255	0,1513	-0,0253
11390-3	0,2862	0,1608	0,1859	-0,0428
11445-2	0,1325	0,0744	0,0894	-0,0204
11445-7	0,2502	0,1399	0,1749	-0,0988
11461-2	0,1667	0,0898	0,1402	-0,0151
11485-7	0,2055	0,1142	0,1408	0,0026
11519-6	0,1868	0,1052	0,1222	-0,0215
11536-8	0,1306	0,0725	0,0903	0,0107
11551-2	0,2025	0,1126	0,1364	0,021
11551-6	0,164	0,0919	0,1056	-0,0094
11560-1	0,1683	0,0911	0,129	-0,0947
11618-1	0,0835	0,0466	0,0564	-0,0011
11696-5	0,2711	0,1526	0,1851	-0,116
11721-1	0,086	0,0481	0,0602	-0,0166
11866-7	0,1397	0,0774	0,1004	-0,0151
11885-2	0,1735	0,0971	0,1126	-0,0017
11885-5	0,1909	0,1068	0,1243	-0,0068
11947-2	0,1825	0,1018	0,1307	-0,0224
12030-4	0,1393	0,0776	0,0994	0,0037
12108-5	0,2767	0,1536	0,1696	0,077
12109-1	0,2193	0,1223	0,1522	-0,0159
12159-6	0,2025	0,1134	0,1241	0,0633
12192-2	0,214	0,1158	0,1657	-0,1095
12192-8	0,2341	0,1322	0,1532	-0,0637
12212-2	0,2619	0,1459	0,1812	-0,07
12271-5	0,2445	0,1367	0,1504	-0,0027
12525-1	0,1757	0,0987	0,1083	-0,0072
12569-8	0,1976	0,1105	0,1322	-0,0036

Tabelle 3.11: Modalwerte der Segmentmittelwerte und Standardabweichungen der Signale für die 56 array-CGH Samples. Standardabweichungen wurden für das 50%-Quantil sowie das 90%-Quantil berechnet. Die MAD-Werte wurden zur Bestimmung der Nulllinie und die Standardabweichungen zur Bestimmung von Aberrationen verwendet.

3.2. COPY-NUMBER-VARIATIONS & GENOTYPISIERUNG

Die Häufigkeit kleinerer Gains und Losses war im Vergleich zu Änderungen in der Kopienzahl ganzer Chromosomen deutlich geringer. So wiesen die Chromosomen 1, 3, 15, 16 und 18 häufig Steigerungen der Kopienzahl auf. Verluste von ganzen Chromosomen oder Großteilen konnten hingegen bei Chromosom 2 (7 Samples) und 19 (13 Samples) beobachtet werden. Mit Ausnahme der spontanen Mammakarzinome, die vergleichsweise wenige Aberrationen aufwiesen, konnte kein weiterer spezifischer Zusammenhang zwischen den Aberrationen und den Differenzierungsstadien bzw. den Mausmodellen hergestellt werden.

Eine grafische Übersicht der Aberrationen ist in den Abbildungen 3.16 und 3.17 gegeben. Die Segmentprofile der einzelnen Proben sowie die kompletten zugehörigen Tabellen finden sich auf dem beiliegenden Datenträger.

3.2.2 Neuroblastome: SNP-Arrays

Die SNP 6.0 Arrays von Affymetrix umfassten 906.368 SNP-Probesets zur Identifikation von Polymorphismen und CNVs sowie weitere 945.806 CN-Probesets zur ausschließlichen CNV-Bestimmung. Eine Bestimmung der Signalqualität erfolgte mittels interner Kontrollen. Die Bestimmung von genomischen Aberrationen erfolgte wie in Abschnitt 2.3.3 beschrieben mit DNACopy. Neben den 9 Neuroblastomproben wurde je Zelllinie eine Zellkulturprobe auf den SNP-Arrays untersucht. Das Ergebnis der IMR-32 Zelllinie wurde in einem zweiten, biologisch unabhängigen Versuch validiert. Über einen Vergleich der Tumorsamples mit den Zellkultursamples sollte eine bessere Charakterisierung der Aberrationen ermöglicht werden, die nach der Transfektion in den Mäusen auftraten.

Laboranalytik Nach der Isolierung der DNA aus den Neuroblastomproben erfolgte die Bestimmung von Konzentration und Qualität des Materials auf einem 0,6%igen Agarosegel und im NanoDrop-Photometer. Eine Aufnahme des Gels findet sich in Abbildung 3.18. Hier wiesen alle Proben eine ausreichende DNA-Qualität auf. Die Quotienten der photometrischen A260/A280 Exktinktionen lagen zwischen 1,94 und 2,08. Die ermittelten Konzentrationen lagen im Bereich zwischen 695ng/µl und 3565ng/µl, siehe Tabelle 3.12.

Nach dem Restriktionsenzymverdau mit Sty I bzw. Nsp I und der anschließenden Ligation bzw. PCR Amplifikation wurde eine Kontrolle auf einem 2%igem Agarosegel durchgeführt, siehe Abbildung 3.19. Die Kontrolle ergab charakteristische Banden, was auf intaktes und wenig degradiertes Material hinwies. Die Kontrolle der Fragmentierung auf dem 2%igem Agarosegel wies allgemein Banden bei unter 100 Basenpaaren auf, siehe Abbildung 3.18. Die Kontrolle zeigte eine erfolgreiche Fragmentierung an.

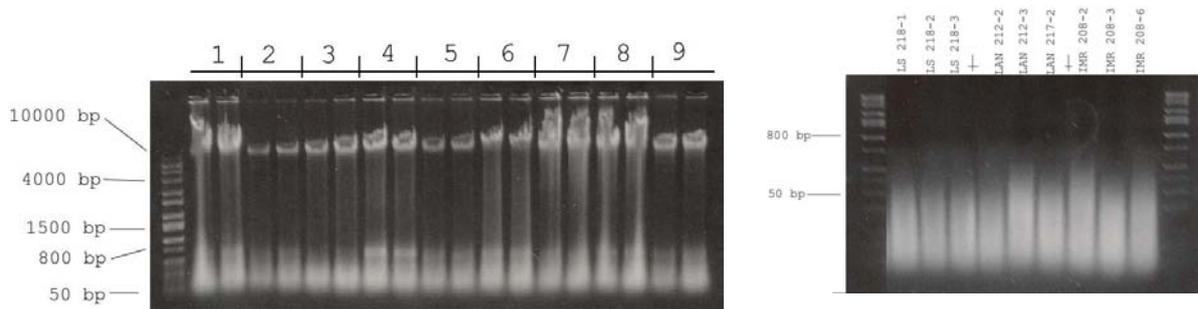


Abbildung 3.18: Agarosegel zur Bestimmung der Qualität des isolierten DNA-Materials aus den Neuroblastomproben (links) bzw. zur Kontrolle der Fragmentierung (rechts). Die Gelkonzentration betrug 0,6% (Isolierung) bzw. 2% (Fragmentierung). Aufgrund des großen Volumens von jeweils 2ml pro Probe wurden diese bei der Isolierung auf je 2 Tubes aufgeteilt. Je zwei benachbarte Bahnen stellen daher hier die Ergebnisse einer Probe dar.

Qualitätskontrolle nach Scan In Tabelle 3.13 sind die „Kontrast QC“ und „Intensitäts QC Call Raten“ für die Neuroblastom SNP-Arrays aufgeführt. Die „Intensity QC Call Rate“ gibt den Anteil der Polymorphismenproben an, für die ein eindeutiger Genotyp erfolgreich zugeordnet werden konnte. Gemäß der Empfehlung des Herstellers sollte diese Rate über einem Wert von 86% liegen. Sie lag bei den 9 Arrays zwischen 97,2% und 99,4%. Die Kontrast QC-Rate gibt das Potential eines Array-Experimentes an, die gemessenen SNP-Signale in 3 Genotyp-Klassen aufzulösen. Sie sollte einen Mindestwert von 0,4 überschreiten und lag bei den Neuroblastomarrays im Minimum bei einem Wert von 1,8.

Als weiteres Kriterium zur Qualitätskontrolle wurde die Konkordanz der Genotypen zwischen den Proben herangezogen. Innerhalb der einzelnen Gruppen (Neuroblastomzelllinien) sollte die Konkordanz möglichst hoch sein, da spontane Mutationen nach der Transfektion nicht erwartet wurden. Tabelle 3.14 zeigt, daß die Konkordanz zwischen Proben der gleichen Gruppe generell über 99% liegt und somit sehr hoch ist. Die Werte bei Vergleichen von Proben aus verschiedenen Gruppen deutlich niedriger. Diese Werte bewegten sich zwischen 62% und 66%.

Genotypisierung Über die SNP-Probesets der SNP 6.0 Arrays von Affymetrix konnte eine Genotypisierung der Neuroblastomsamples durchgeführt werden. Innerhalb der drei Zelllinien lag die Konkordanz der Genotypen zwischen 97,05% und 99,46%, siehe auch Tabelle 3.14. Die Samples der einzelnen Zelllinien wiesen somit erwartungsgemäß nur geringe Unterschiede auf. Differente Genotypen konnten beim Vergleich zwischen den IMR und LAN Samples in 305.814 SNPs (34%) beobachtet werden. Davon waren 98.564 SNPs

3.2. COPY-NUMBER-VARIATIONS & GENOTYPISIERUNG

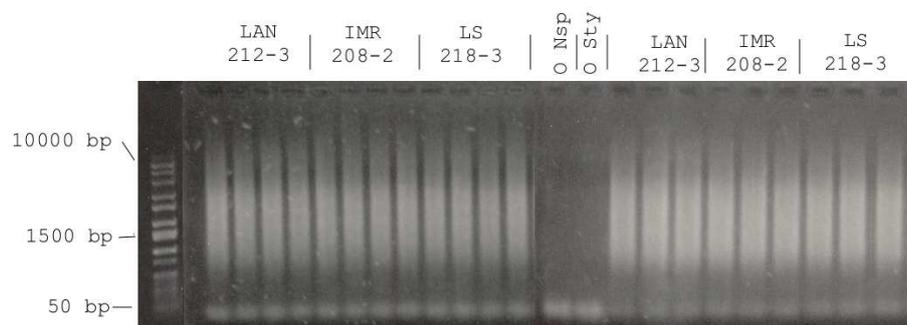


Abbildung 3.19: Neuroblastome: Kontrollen auf 2%igem EtBr-Agarosegel nach erfolgter Restriktion mit Nsp I (links von Negativkontrollen) und Sty I (rechts von Negativkontrollen). Abgebildet sind 3 der 9 Samples. Die Gelbilder der übrigen Samples finden sich im Anhang in Abbildung B.23.

Sample	ng/ μ l	A260	A280	A260/A280
Lan1 217-2_1	3195,73	63,915	32,454	1,97
Lan1 217-2_2	1638,54	32,771	16,053	2,04
Lan1 212-2_1	695,34	13,907	6,730	2,07
Lan1 212-2_2	1127,27	22,545	10,848	2,08
Lan1 212-3_1	1651,71	33,034	16,101	2,05
Lan1 212-3_2	1578,70	31,574	15,225	2,07
IMR 208-3_1	1971,62	39,432	19,203	2,05
IMR 208-3_2	1721,17	34,423	16,742	2,06
IMR 208-6_1	1027,31	20,546	9,952	2,06
IMR 208-6_2	1156,02	23,120	11,175	2,07
IMR 208-2_1	1541,47	30,829	15,121	2,04
IMR 208-2_2	1723,99	34,480	16,952	2,03
LS 218-1_1	2023,37	40,467	20,514	1,97
LS 218-1_2	2666,24	53,325	27,120	1,97
LS 218-2_1	2483,41	49,668	25,474	1,95
LS 218-3_2	3565,40	71,308	36,718	1,94
LS 218-3_1	978,80	19,576	9,460	2,07
LS 218-3_2	1510,07	30,201	14,641	2,06

Tabelle 3.12: Übersicht der Konzentrationen und Reinheit nach der Isolierung der DNA aus dem Neuroblastomprobenmaterial. Angegeben sind die Konzentrationen in ng/ μ l und die Reinheit als Quotient der photometrischen Extinktionen bei 260 nm und 280 nm. Das A260/A280-Verhältnis sollte idealerweise zwischen 1,8 und 2,0 liegen. Aufgrund des großen Volumens von jeweils 2ml pro Probe wurden diese bei der Aufreinigung auf je 2 Tubes aufgeteilt. Es finden sich somit zwei Messreihen pro Probe in der Tabelle.

Sample	Innerhalb der Grenzen	Kontrast QC	Call Rate	Geschlecht
ZL IMR	Ja	2,86	97,52	M
IMR 208-2	Ja	2,40	99,09	M
IMR 208-3	Ja	2,47	99,09	M
IMR 208-6	Ja	1,24	97,23	M
ZL Lan	Ja	1,99	99,35	W
Lan 212-2	Ja	1,80	99,28	W
Lan 212-3	Ja	1,80	99,37	W
Lan 217-2	Ja	2,18	99,34	W
ZL LS	Ja	2,04	98,85	M
LS 218-1	Ja	2,40	99,34	M
LS 218-2	Ja	2,01	99,28	M
LS 218-3	Ja	2,14	99,38	M

Tabelle 3.13: Übersicht der Qualitätskontrollen für die SNP 6.0 Chips der Neuroblastome. Die Kontrast QC-Rate gibt das Potential eines Array-Experimentes an, die gemessenen SNP-Signale in 3 Genotyp-Klassen aufzulösen. Sie sollte einen Mindestwert von 0,4 überschreiten. Die Call-Rate gibt den Anteil der Probesets an, für die aus den gemessenen Signalen erfolgreich ein Genotyp ermittelt werden konnte. Sie sollte einen Mindestwert von 86% erreichen.

Sample	IMR 208-2	IMR 208-3	IMR 208-6	LAN 212-2	LAN 212-3	LAN 217-2	LS 218-1	LS 218-2	LS 218-3
IMR 208-2	100	99,28	97,05	62,57	62,53	62,48	65,33	65,26	65,40
IMR 208-3		100	97,21	62,56	62,52	62,48	65,29	65,25	65,35
IMR 208-6			100	61,78	61,77	61,69	64,31	64,28	64,38
LAN 212-2				100	99,43	99,27	62,35	62,30	62,45
LAN 212-3					100	99,46	62,32	62,28	62,43
LAN 217-2						100	62,31	62,28	62,39
LS 218-1							100	99,37	99,41
LS 218-2								100	99,29
LS 218-3									100

Tabelle 3.14: Übersicht der Konkordanz (in %) der Genotypen zwischen den SNP 6.0 Chips der Neuroblastome. Die Konkordanz gibt den Anteil der identischen Genotypen zwischen zwei Chips in % an. Per Definition liegt sie beim Vergleich eines Chips mit sich selbst bei 100%.

3.2. COPY-NUMBER-VARIATIONS & GENOTYPISIERUNG

in IMR heterozygot und in LAN homozygot bzw. 139.227 SNPs in LAN heterozygot und in IMR homozygot. Zwischen IMR und LS waren Unterschiede in 281.341 SNPs (31%) zu beobachten. Von diesen SNPs wiesen 104.980 in IMR einen heterozygoten Genotyp auf während 97.762 dieser SNPs in LS heterozygot vorlagen. Der Vergleich zwischen LAN und LS ergab Unterschiede in 306.360 SNPs (34%), von denen in LAN 140.994 SNPs heterozygot waren und 93.113 SNPs eine Heterozygotie in LS aufwiesen. Eine grafische Übersicht der Genotypen findet sich in Abbildung 3.20.

Aberrationsbestimmung Die Signale der SNP und CN Proben der SNP6.0 Chips wurden zur Bestimmung von Amplifikationen und Deletionen herangezogen. Die Modalwerte der Segmentmittelwerte lag bei den SNP Probesets zwischen -0,047 und 0,034, bei den CN Probesets zwischen -0,133 und 0,074. Die Standardabweichungen der 50% Quantile lagen bei den SNP Probesets zwischen 0,067 und 0,110 sowie bei den CN Probesets im Bereich 0,076 bis 0,128. Die Standardabweichungen der 90% Quantile lagen zwischen 0,089 und 0,177 bei den SNP Probesets. Bei den CN Probesets erstreckten sich die Werte von 0,098 bis 0,159. Eine Übersicht der Werte findet sich in Tabelle 3.15.

Zur Klassifikation der Aberrationen wurden für Gains eine Signalgrenze in Höhe der 1,8fachen, für Losses der negativen 2fachen Standardabweichung (50% Quantil) festgesetzt. Für Amplifikationen und Deletionen wurde ein Grenzwert in Höhe der ± 3 fachen Standardabweichung (90% Quantil) verwendet. Als Ausnahme wurde für die dritte IMR Probe, IMR 208-6, ein Grenzwert von 1,6facher Standardabweichung (50%Quantil) zur Bestimmung von Gains festgelegt.

Eine grafische Übersicht der Aberrationen ist in Abbildung 3.21 gegeben. Die Segmentprofile der einzelnen Proben sowie die vollständigen zugehörigen Tabellen finden sich auf dem beiliegenden Datenträger.

Aberrationen Größere gemeinsame Gains konnten bei den IMR-Tumoren in 1p36.33-36.13, 3p11-q11, 9p24.3-22.1, 9p13.3-q12, 10p12.33-q22.3 und Xp11-q13 inklusive des ersten Abschnittes von Xq21 beobachtet werden. Außerdem zeigten sich nahezu die kompletten Arme 14q, 16q und 19q amplifiziert, sowie ein Großteil des p-Arms von Chromosom 18 und 18q21-22. Chromosom 20 wies durchgehend mehr als 2 Kopien auf. Desweiteren waren etliche kleinere Gains zu beobachten, insbesondere in 6p22.3. Gemeinsame Losses fanden sich in 5q11.2-13.2, 9p21.3-21.1 und 17q21.31-24.2. Teile von 9p21 wurden dabei als vollständig deletiert eingestuft.

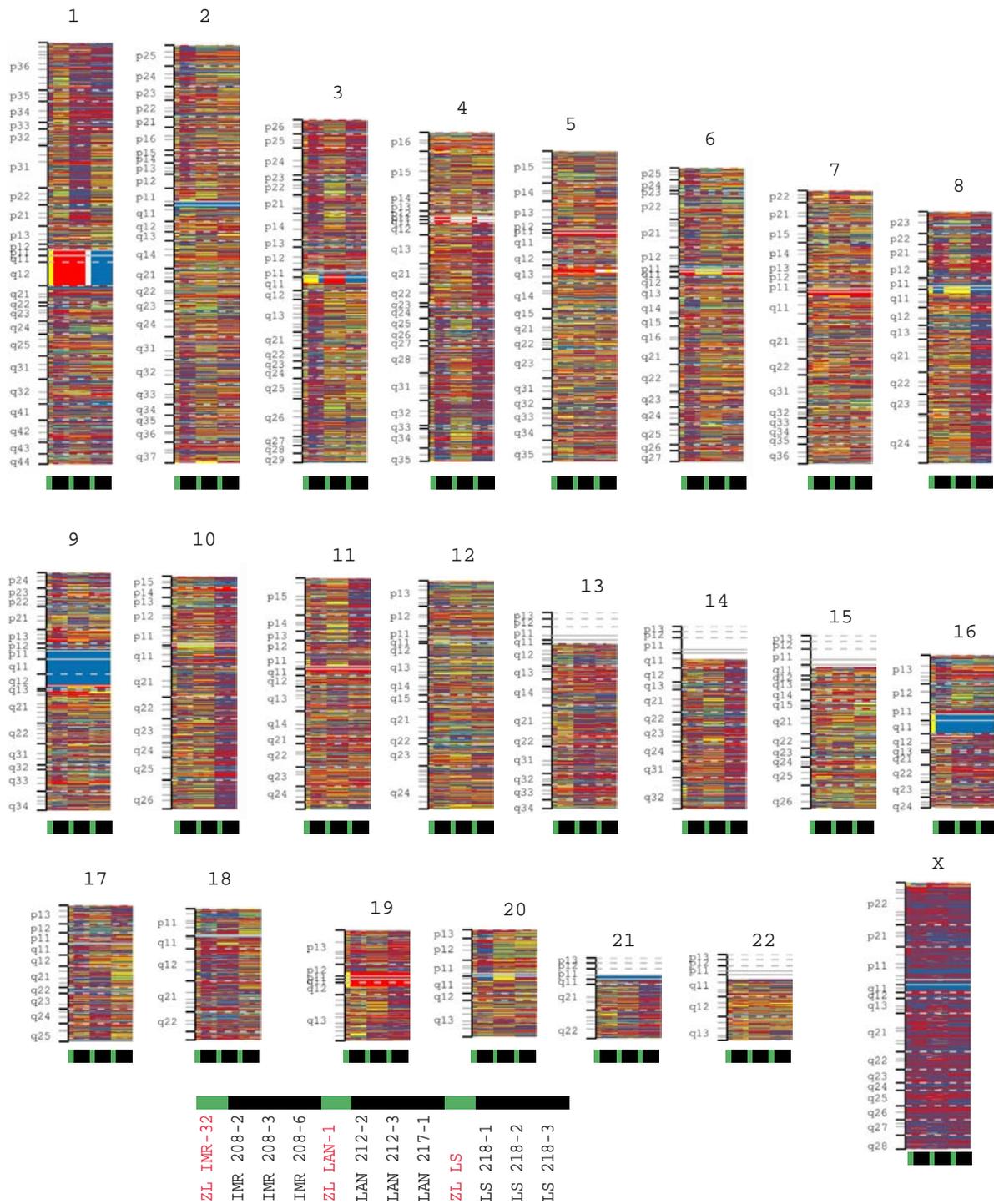


Abbildung 3.20: Genotypen der Neuroblastome, die mittels der SNP 6.0 Arrays von Affymetrix ermittelt wurden. Genotyp-Farbkodierung: „AA“ - blau, „BB“ - rot, „AB“ - gelb, „Nicht ermittelbar“ - grau. Dabei stehen „AA“ und „BB“ für homozygote Genotypen und „AB“ für einen heterozygoten Genotypen.

3.2. COPY-NUMBER-VARIATIONS & GENOTYPISIERUNG

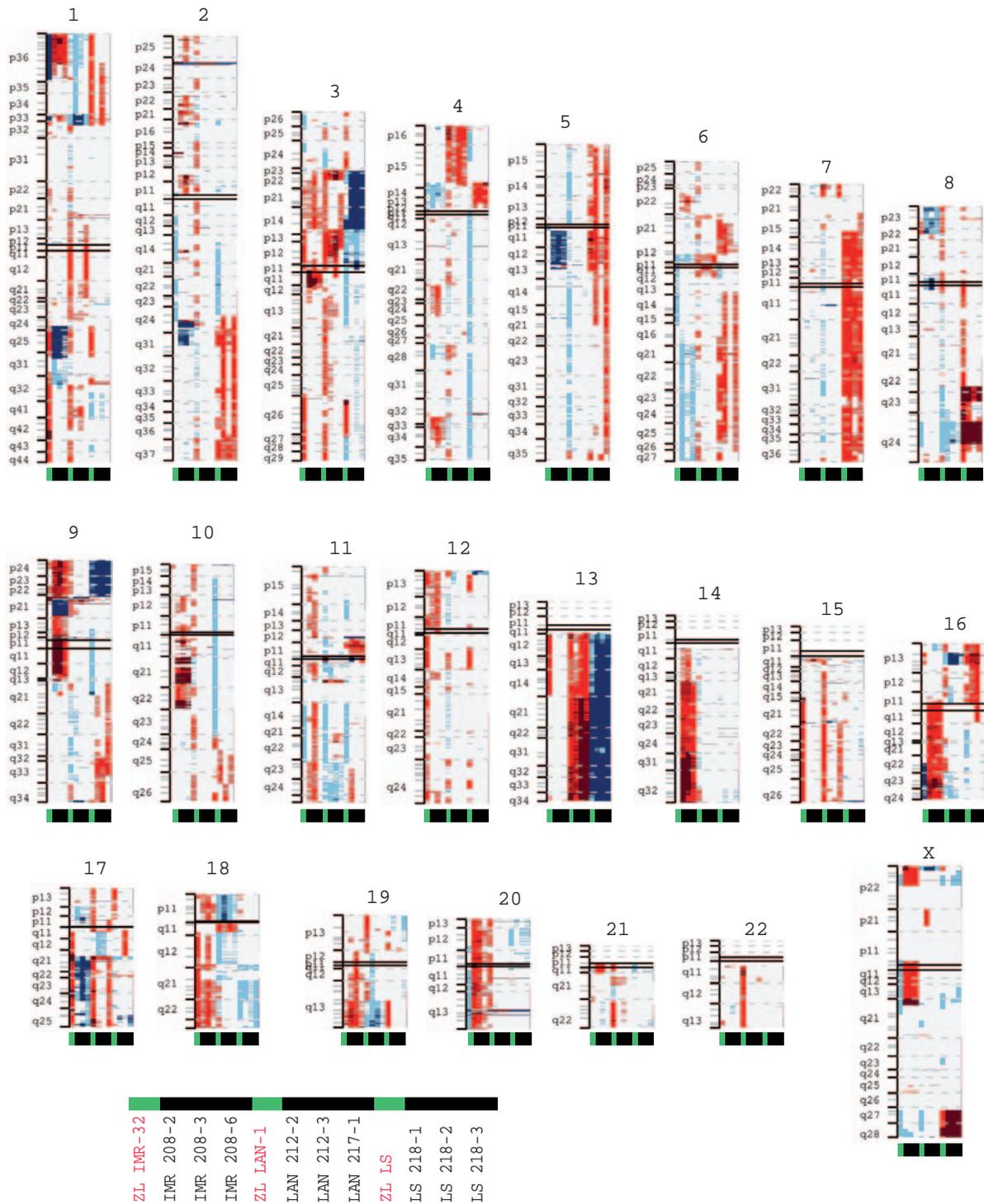


Abbildung 3.21: Aberrationsprofile der Neuroblastomsamples.

Farbkodierung: Rot, Gain; Dunkelrot, Amplifikation, Gewinn mehrerer Kopien; Blau, Loss; Dunkelblau, Deletion, Verlust beider Kopien; Grau, keine Aberration.

Sample	Modalwert CN	Modalwert SNP	SD 50% Q. CN	SD 50% Q. SNP	SD 90% Q. CN	SD 90% Q. SNP
IMR 208-2	0,037	-0,020	0,098	0,088	0,130	0,126
IMR 208-3	-0,028	0,009	0,095	0,083	0,123	0,114
IMR 208-6	-0,011	-0,032	0,128	0,108	0,159	0,139
Lan 212-2	0,074	0,027	0,086	0,074	0,111	0,103
Lan 212-3	0,038	0,032	0,076	0,067	0,098	0,089
Lan 217-2	0,030	0,034	0,102	0,080	0,151	0,177
LS 218-1	-0,033	-0,047	0,106	0,088	0,132	0,112
LS 218-2	-0,029	-0,046	0,099	0,083	0,124	0,102
LS 218-3	-0,133	0,006	0,099	0,083	0,110	0,104

Tabelle 3.15: Modalwerte der Segmentmittelwerte nach Segmentierung der SNP und CN Probesets für die Neuroblastomproben. Zur Bestimmung von Gains, Losses, Amplifikationen und Deletionen wurden die Segmente zuvor um die berechneten Modalwerte zentriert.

Einige Gains und Losses konnten bei den angelegten Klassifikationskriterien nicht in allen drei Proben beobachtet werden. Hierzu zählten unter anderem Gains in 2p22, 2p12, 3p22-14, 4q33-34, 12p sowie 14q11-12. Die Bande 1q24 wies sich in den beiden Proben, IMR.208.2 und IMR.208.3 ein Loss auf.

Bei den LAN Proben konnten größere gemeinsame Gains in 4p16-15, 3p14.1-12.1, 13q und in 18q11.2 beobachtet werden. Desweiteren fanden sich etliche kleinere und punktuelle Gains, z.B. entlang der Gene *NAG*, *DDX1* und *MYCN* in 2p24.3 oder bei *DPP3* in 11q13.1. Gemeinsame Losses fanden sich in 16p13 und in 17p13, die unter anderem das Gen *TP53* umfassten und in Xp22.33.

In 3p13-12 konnten einige größere Gains beobachtet werden, die bei den gewählten Kriterien mitunter nur in zwei der drei Proben auftraten.

In den LS Proben zeigten sich Gains im gesamten p-Arm von Chromosom 16, in 2q37, 4p14-13, 5p15-q15, 6p12.1, 7p14-q36, 8q22.3-q23.1, 8q24.13-24.23, 9q33 und 10q25. Deletionen konnten in 9p24-22, 3p23-p14 und dem gesamten q-Arm von Chromosom 13 beobachtet werden.

Etliche Regionen wiesen nur in zwei von drei Fällen Gains auf. Dazu zählten die Banden 2q24-q36, 6q13-q26, 8q23.2-24.12 und 9q13-31.

Die Vergleiche mit den Zelllinien zeigten größtenteils gute Übereinstimmungen mit den Tumorsamples. Teilweise wurde bei den Zelllinien Gewinne oder Verluste von einer Kopie

3.2. COPY-NUMBER-VARIATIONS & GENOTYPISIERUNG

in Regionen detektiert, in denen die Tumorsamples keine Änderungen aufwiesen. So wurden in der LAN-1 Zelllinie z.B. kleinere Regionen auf Chromosom 2 als Gain eingestuft und große Teile von Chromosom 5 als Loss. Die LS Zelllinie wies Losses auf Chromosom 4 und 10 auf, die in den Tumoren nicht erkennbar waren. Unterbrochen wurde der Loss auf Chromosom 4 von jener Region, die in den Tumoren einen Gain aufwies. Andererseits konnten die Gains großer Regionen von Chromosom 7 (LS) und 13q (LAN-1) sowie die Deletionen in 19p24-p22 und 13q (LS) sowohl in der Zelllinie als auch den Tumoren beobachtet werden.

Bei der IMR-32 Zelllinie konnten einige Gains beobachtet werden, die bei den Tumoren Losses aufwiesen. Dies war bei 1q25 und 17q der Fall. Desweiteren konnten Losses und Deletionen in Regionen beobachtet werden, die bei den Tumoren Gains oder Amplifikationen aufwiesen. Dies war bei 1p36, 16q23-q24 und 20q der Fall. Die in den Tumoren aberranten Regionen 5q11-q12, 10p12.33-q22.3 und 14q zeigten in der IMR-32 Zelllinie keine Veränderungen. Andererseits zeigte sich eine gute Übereinstimmung in den Aberrationen in 3p11-q11, 19q13 und 18q. Diese Ergebnisse konnten in einer zweiten, biologisch unabhängigen Probe bestätigt werden. Hierzu wurden die IMR-32 Zellen einige Monate später neu kultiviert und das genomische Profil untersucht. Die beiden Proben wiesen im Vergleich identische Aberrationsprofile auf.

3.3 Korrelation

Mittels einer Integration von Expressionsdaten und CNVs sollte eine Eingrenzung der bisherigen Ergebnisse auf aussagekräftige Kandidaten ermöglicht werden. Weiterhin galt das Interesse der Fragestellung, wie unmittelbar sich Aberrationen auf das Expressionsniveau der betroffenen Gene auswirken. Die Ergebnisse dieser Korrelationen werden im folgenden Abschnitt beschrieben.

3.3.1 Brusttumore

Die aus der Expressionsanalytik und CNV-Analytik gewonnenen Daten wurden mit dem Verfahren von Stranger et al. [157] auf Korrelationen untersucht. Aufgrund der zwei unterschiedlichen array-CGH-Plattformen, die bei den Mammakarzinomen zum Einsatz kamen, wurde die Korrelation für die NP8 und BALB/c-Mausproben separat durchgeführt. Für die weiteren Untersuchungen wurde die Schnittmenge der signifikant korrelierenden Gene aus beiden Bestimmungen verwendet. Als signifikant wurden Gene eingestuft, deren Korrelations-P-Wert ein Signifikanzniveau von mindestens 0,001 aufwies. Die BALB/c-Samples wiesen dabei 136 signifikant korrelierende Gene auf, die NP8-Samples 1317. Die Schnittmenge wurde von 87 Genen gebildet, die primär im *Met*-Locus auf Chromosom 6, sowie auf den Chromosomen 3, 15 und 19 zu finden waren. Unter diesen Genen befanden sich unter anderem die Gene *Tes*, *Cav1*, *Met*, *Capza2* und *St7* des *Met*-Locus. Auf Chromosom 15 fanden sich unter anderem die Gene *Rbx1* und *Atxn10*. Gute Korrelationen auf Chromosom 19 konnten unter anderem bei den Genen *Fau*, *Cox8a*, *Stx5a*, und *Ddb1* beobachtet werden und auf Chromosom 3 bei *Psmc4*.

Eine Auftragung der Expressionswerte gegen die zugehörigen Segmentmittelwerte für 11 dieser Gene findet sich in Abbildung 3.23. Eines der beiden bestkorrelierenden Gene war *St7* mit einem R^2 -Wert von 0,891 (NP8-Mäuse) bzw. 0,755 (Balbc-Mäuse). Zusammen mit den Genen des *Met*-Locus wies es bereits in der Expressionsanalyse eine signifikante differentielle Regulation zwischen differenzierten und undifferenzierten Tumoren auf. Zudem konnte in der CNV-Analytik zuvor eine häufigere Amplifikation der Region in undifferenzierten Tumoren beobachtet werden. Die Korrelationsfaktoren, P-Werte, Zuordnungen von Expressions-Probesets zu CGH-Identifiern und die chromosomalen Locationen der 11 Gene sind in Tabelle 3.16 aufgeführt. Die entsprechenden Daten aller 87 Gene finden sich in Tabelle A.1 im Anhang.

3.3. KORRELATION

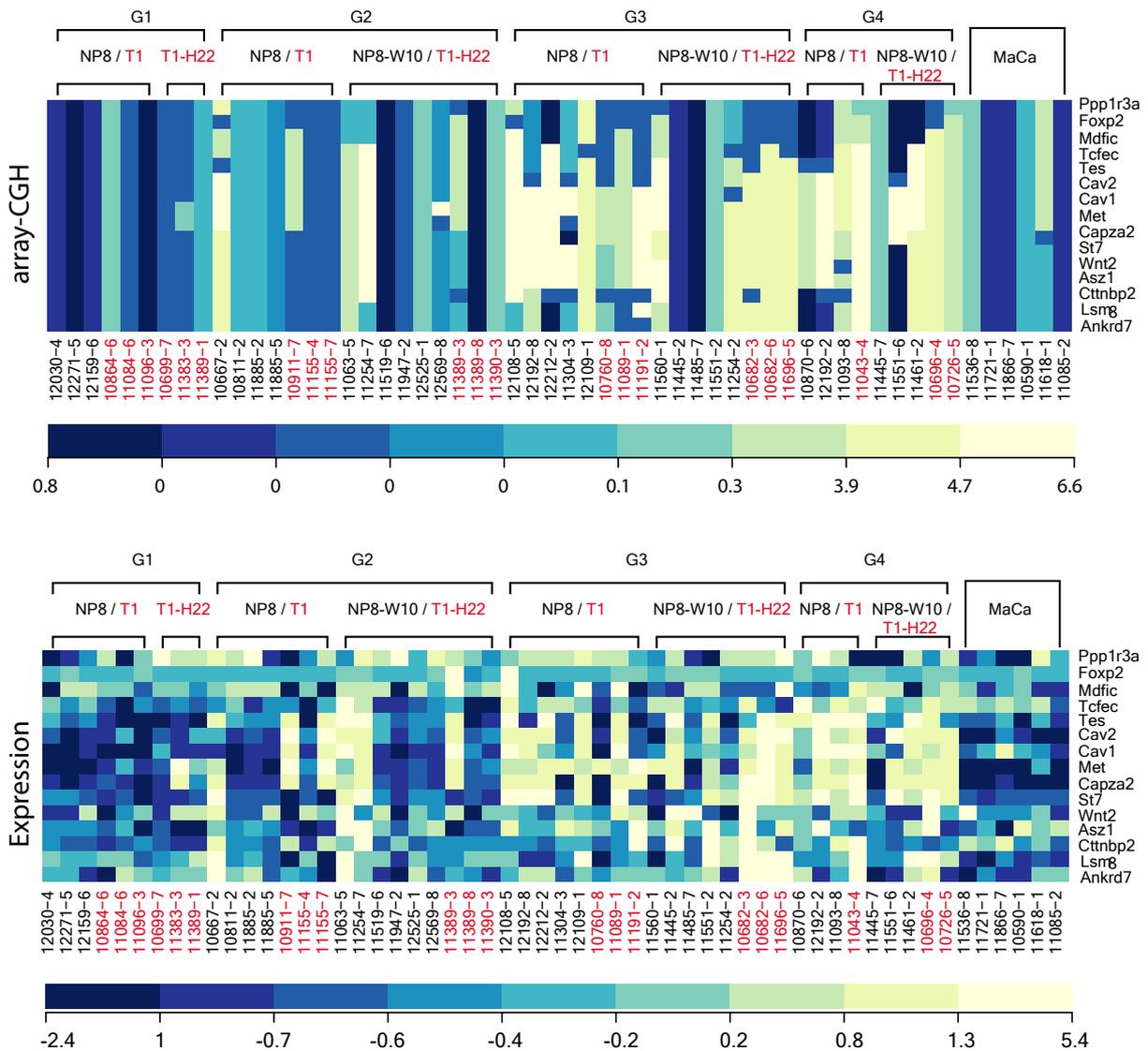


Abbildung 3.22: Brusttumore: Expressionswerte und array-CGH Daten des *Met*-Locus im Vergleich. Die Expressionswerte sind als z-score pro Gen über die Samples standardisiert. Die array-CGH Daten entsprechen den \log_2 ratios nach der Segmentierung. Erkennbar ist, daß der *Met*-Locus in undifferenzierten Tumoren (G3 & G4) deutlich häufiger amplifiziert ist. Die deutliche Asymmetrie der Werteskala hebt die Stärke der Amplifikationen hervor. Die Expressionswerte folgen den Amplifikationen. Die spontanen Mammakarzinome (am äußersten rechten Rand) gleichen hinsichtlich des *Met*-Locus eher den differenzierten Tumoren.

Oben, array-CGH Daten; unten, Expressionswerte.

Zur Evaluation des generellen Einflusses von Aberrationen auf die Genexpression in Brusttumoren wurde zusätzlich die genomweite Korrelationsmethode verwendet. Die notwendige Referenz wurde aus allen 80 Samples der Expressionsanalytik berechnet. Es bestand ein deutlicher Zusammenhang zwischen den Aberrationen und Expressionsniveaus über das gesamte Genom. Eine Gegenüberstellung von segmentierten Expressionswerten und array-CGH Daten für 4 der Brusttumorsamples findet sich in Abbildung 3.24. Deutlich erkennbar waren steigende Expressionstendenzen in Regionen, die bei der array-CGH Gains aufwiesen. Entsprechend konnten auch deutlich niedrigere Expressionssignale in Loss-Regionen beobachtet werden. Sowohl Gains als auch Losses, z.B. in Sample 11461-2, waren auf Transkriptomebene klar erkennbar. Bei Sample 12192-2 konnten selbst graduelle Unterschiede in der Kopienzahl von Chromosom 11 auf Transkriptomebene dargestellt werden.

3.3. KORRELATION

Symbol	EXP ID	CGH ID	Raw P-Wert	R^2	Q 0,01	Q 0,001	Chr.	Position
<i>Arrays aus Oslo</i>								
Tes	21753	A_53_P122283	14,264	0,829	2,454	3,394	6	17035507
Cav1	12389	A_53_P104726	5,343	0,456	2,437	3,467	6	17273898
Met	17295	A_53_P121226	15,176	0,848	2,344	3,717	6	17468890
Capza2	12343	A_53_P157131	13,959	0,91	2,291	3,34	6	17601821
St7	64213	A_53_P121352	13,787	0,891	2,399	3,392	6	17768484
Lsm8	76522	A_53_P103703	12,892	0,796	2,332	4,196	6	18801343
Myo10	17909	A_53_P155317	3,934	0,35	1,98	2,809	15	25647854
Rbx1	56438	A_53_P142789	6,687	0,542	2,069	3,169	15	81301772
Atxn10	54138	A_53_P116218	5,497	0,467	2,243	3,557	15	85230224
Rbm4	19653	A_53_P136800	4,253	0,375	1,917	2,855	19	4789088
Fau	14109	A_53_P136738	6,851	0,552	2,375	3,592	19	6058735
Cox8a	12868	A_53_P124485	6,535	0,533	2,101	3,243	19	7290854
Stx5a	56389	A_53_P117603	9,187	0,669	1,933	3,148	19	8823032
Ddb1	13194	A_53_P115728	8,092	0,618	2,221	3,09	19	10692182
Trim8	93679	A_53_P146014	6,353	0,522	2,024	2,912	19	46583801
<i>Arrays aus Amsterdam</i>								
Tes	21753	sc10001061,1_50	6,296	0,743	2,519	3,597	6	17035507
Cav1	12389	sc10001108,1_2	5,682	0,703	2,487	3,601	6	17273898
Met	17295	sc130355,22,1_19	5,26	0,672	2,097	3,495	6	17468890
Capza2	12343	sc1012343,12_1	6,13	0,772	2,444	3,351	6	17601821
St7	64213	sc10001099,1_185	6,493	0,755	2,45	3,347	6	17768484
Lsm8	76522	sc100319202,1_96	8,981	0,865	1,879	3,46	6	18801343
Myo10	17909	sc10002440,1_24	3,913	0,549	1,914	2,665	15	25647854
Rbx1	56438	sc10002434,1_7	4,268	0,585	1,974	2,747	15	81301772
Atxn10	54138	sc147668,10_45	3,799	0,537	2,007	2,858	15	85230224
Rbm4	19653	sc1081909,8_96	4,343	0,592	2,024	2,978	19	4789088
Fau	14109	sc1056275,3_18	5,842	0,714	1,964	2,943	19	6058735
Cox8a	12868	sc1016973,22_185	5,366	0,68	2,036	3,133	19	7290854
Stx5a	56389	sc1068157,2_24	5,635	0,7	1,963	2,984	19	8823032
Ddb1	13194	sc1060361,6_3	4,067	0,565	1,909	2,995	19	10692182
Trim8	93679	sc1057344,13_241	4,303	0,588	1,93	2,686	19	46583801

Tabelle 3.16: Brsuttumore: Die aufgeführten Gene stellen eine Auswahl der Top-Kandidaten dar, die eine gute Korrelation zwischen Expressionsniveau und CNV-Analytik aufweisen. Angeführt sind die Identifier der Gene auf dem Expressionsarray sowie die Identifier der am dichtesten gelegenen Probesets auf dem SNP-Array. $Q\ 0,01$ und $Q\ 0,001$ bezeichnen die P-Werte, die die 0,01 bzw. 0,001 Quantile denotieren. Beim *Raw P-Wert* handelt es sich um den unkorrigierten P-Wert. Die P-Werte sind in negativer, dekadisch-logarithmischer Form angegeben. R^2 gibt den Korrelationskoeffizienten an, der im Idealfall (absolute Übereinstimmung) bei 1 liegt. Die komplette Liste der als gut korrelierend eingestuften Gene umfasst 87 Probesets und ist in Tabelle A.1 im Anhang aufgeführt.

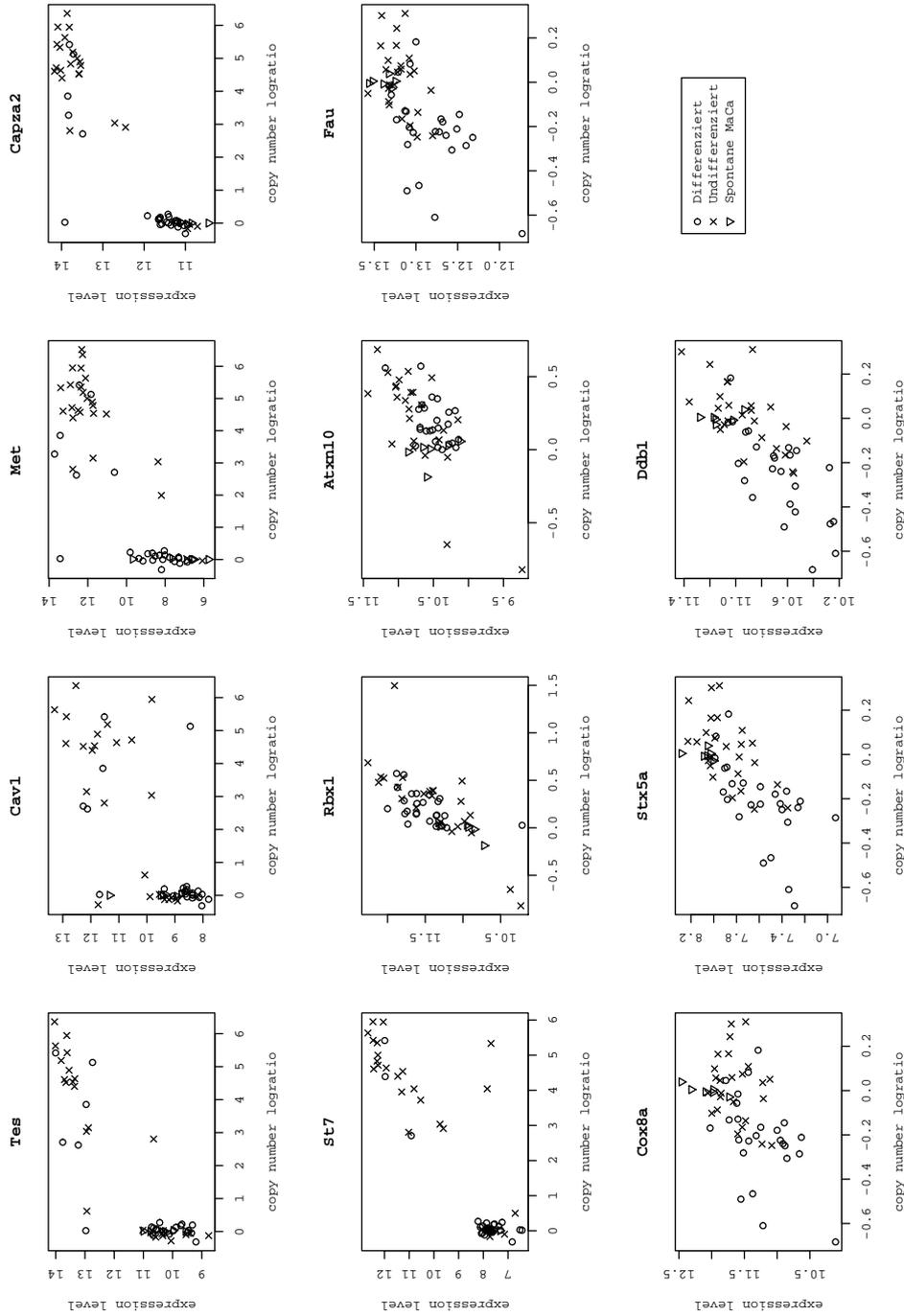


Abbildung 3.23: Brusttumore: Gegenüberstellung von Expressionsniveaus und segmentierter logratios der CGH-Analytik für die 11 top korrelierenden Gene der Brusttumore. Die Gene des Met-Locus (*Tes*, *Cav1*, *Met*, *Capza2*, *St7*) weisen zusätzlich eine differentielle Expression zwischen differenzierten Tumoren (G1, G2) und undifferenzierten Tumoren (G3, G4) auf.

3.3. KORRELATION

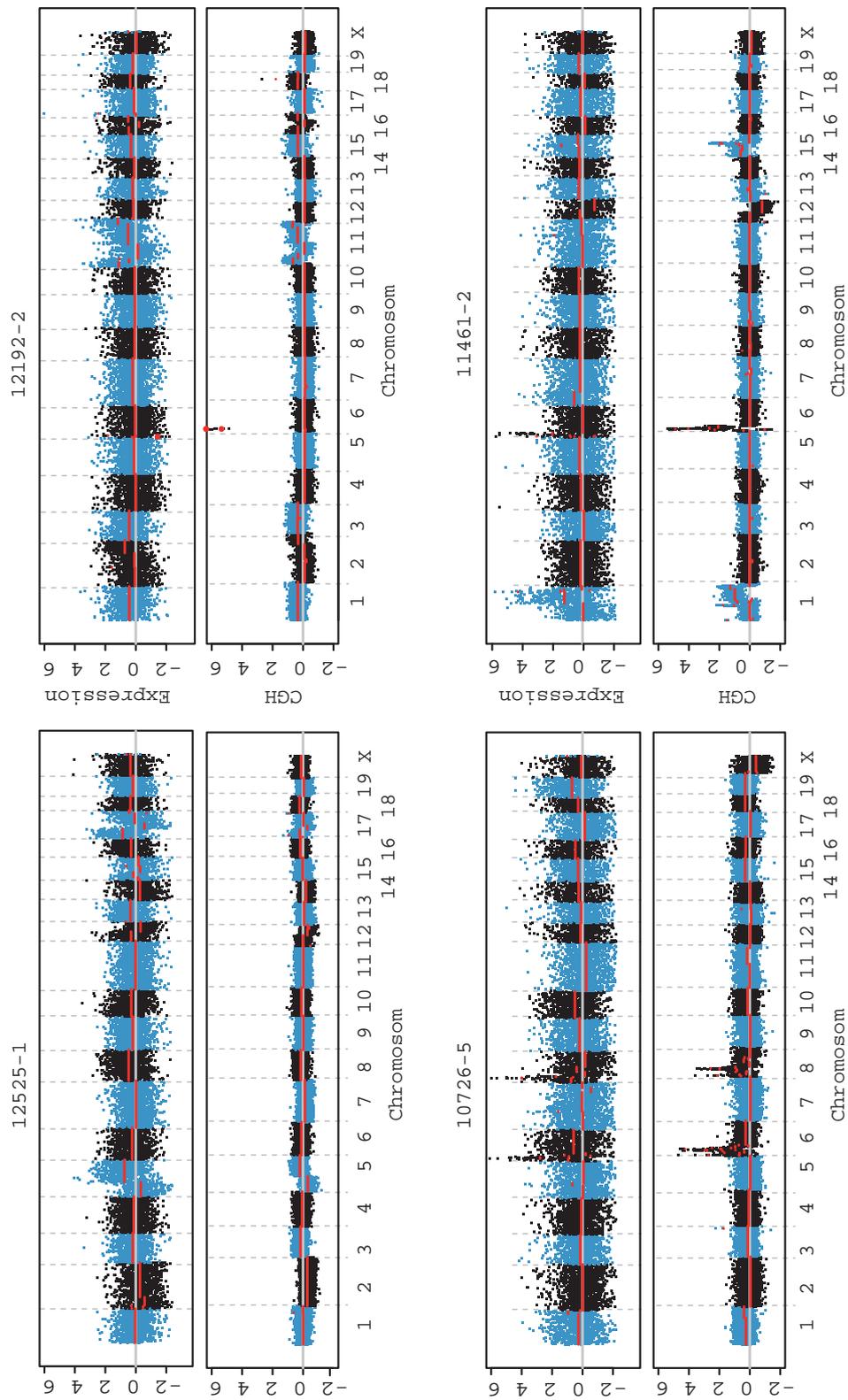


Abbildung 3.24: Brusttumore: Gegenüberstellung der segmentierten Expressionsprofile nach Standardisierung und CGH von vier Brusttumorsamples. Die CGH und Expressionsprofile sind jeweils paarweise übereinander abgebildet. Die blau und schwarz gefärbten Punkte stellen die logRatios (Verhältnis Tumorsignal zu Referenzsignal) der einzelnen CGH-Oligos bzw. Expressions-ProbeSets dar. Hierbei sind die einzelnen Chromosomen, zur Unterscheidung, alternierend blau und schwarz dargestellt. Die roten Linien geben die berechneten Mittelwerte der Segmente wieder. Ein positives Tumor-Referenz-Verhältnis beschreibt einen Hinzugewinn chromosomaler Kopien. Kopienverluste spiegeln sich in negativen logRatios wider.

3.3.2 Neuroblastome

Zur Korrelation der Ergebnisse aus Expressionsanalytik und der SNP-Arrays wurde das Verfahren von Stranger et al. [157] eingesetzt. Bei der darauf folgenden Auswertung der Resultate wurden die spezifischen Genlisten auf jene Gene reduziert, die bei der Korrelationsanalyse ein Signifikanzniveau von 0,001 unterschritten. Die Größe der IMR spezifischen Genliste reduzierte sich dadurch auf 857 Gene, die LAN spezifische Genliste auf 952 Gene und die LS-Liste auf 537 Gene. Von der Liste der Gene, die alle drei Gruppen unterschieden erwiesen sich beim Korrelationsverfahren 440 Gene als signifikant.

Zu den gut korrelierenden IMR-spezifischen Genen gehörten unter anderem *FBXO9*, *WATP* und *DKK1*. Die Liste der LAN-spezifischen Gene umfasste unter anderem die Gene *IGSF9*, *CNTN2*, *MYCN* und *TP53*. Unter den LS-spezifischen Genen mit guter Korrelation fanden sich *CALM1*, *SESN1* und *KLF12*. Die Gene *MED28* und *AKT1* unterschieden alle drei Zelllinien und wiesen auch eine gute Korrelation auf. Eine Auswahl der Korrelationsfaktoren und P-Werte für einige der nach der Korrelation verbliebenen Gene findet sich in Tabelle 3.18. Abbildungen 3.26 und 3.27 zeigen eine Auftragung der Expressionswerte der in Tabelle 3.18 aufgeführten Gene gegenüber den zugehörigen Segmentmittelwerten. Abbildungen B.1 bis B.22 im Anhang zeigen eine Projektion der Expressionswerte der signifikant regulierten Gene auf die Segmentprofile der einzelnen Chromosomen.

Für eine genomweite Korrelation aller Genexpressionen jeweils eines Samples mit den zugehörigen Segmentmittelwerten wurde die genomweite Korrelation zur Anwendung gebracht. Zur Bereitstellung einer hinreichend guten Referenz wurden bereits publizierte Neuroblastom-Datensätze mit insgesamt 72 Samples der GEO-Datenbank entnommen. Diese wurden zusammen mit dem eigenen Datensatz standardisiert. Datenbankidentifizier und Publikationsreferenzen der Datensätze können Tabelle 3.17 entnommen werden. Nach der Standardisierung erfolgte eine Segmentierung mit DNACopy. Eine Übersicht der chromosomalen Regionen mit besonders hohen bzw. niedrigen Expressionswerten nach der Standardisierung findet sich in Abbildung 3.25. Die zusammen mit den Signalen der Proben aus der GEO-Datenbank standardisierten Expressionsprofile wiesen eine gute Übereinstimmung mit den Profilen der Copynumber-Analyse auf. So konnten in den IMR Samples lediglich die Gains in Chromosom 9 und 10 nur teilweise rekonstruiert werden. Die Losses in den Chromosomen 1, 3 und 5 fanden sich ebenfalls wieder. Die Gains in den LAN Samples auf Chromosom 1, 3,4 und 18 waren teilweise auch hier deutlich höher exprimiert. Die Losses in Chromosom 16, 17 und 18 zeigten sich verhältnismäßig niedrig exprimiert.

3.3. KORRELATION

Accession	Anzahl Samples	PMID	Contributor
GSE4600	12	16682435	Peddada S. et. al.
GSE4773	20		nicht vorhanden
GSE6451	6	17360324	Liu Y. et. al.
GSE8066	16	17643814	Koppen A. et. al.
GSE8596	18	11836553	Miyagawa Y. et. al.

Tabelle 3.17: Neuroblastome: Accession-Nummern der für die genomweite Korrelation als Referenz verwendeten GEO-Datensätze, siehe Abschnitt 3.3.2. Desweiteren sind die Anzahl der jeweils enthaltenen Samples aufgeführt sowie die PubMed IDs der Publikationen, soweit vorhanden. Für Datensatz GSE4773 ist bei GEO keine Publikation angegeben⁸.

In den LS Samples konnten die Gains und Losses in den Chromosomen 2, 3, 4, 5, 6, 10 und 13 sehr gut erkennen, während die Aberrationen in Chromosom 7, 8 und 9 nur teilweise oder gar nicht erkennen ließen.

⁸WebLink: <http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE4773>

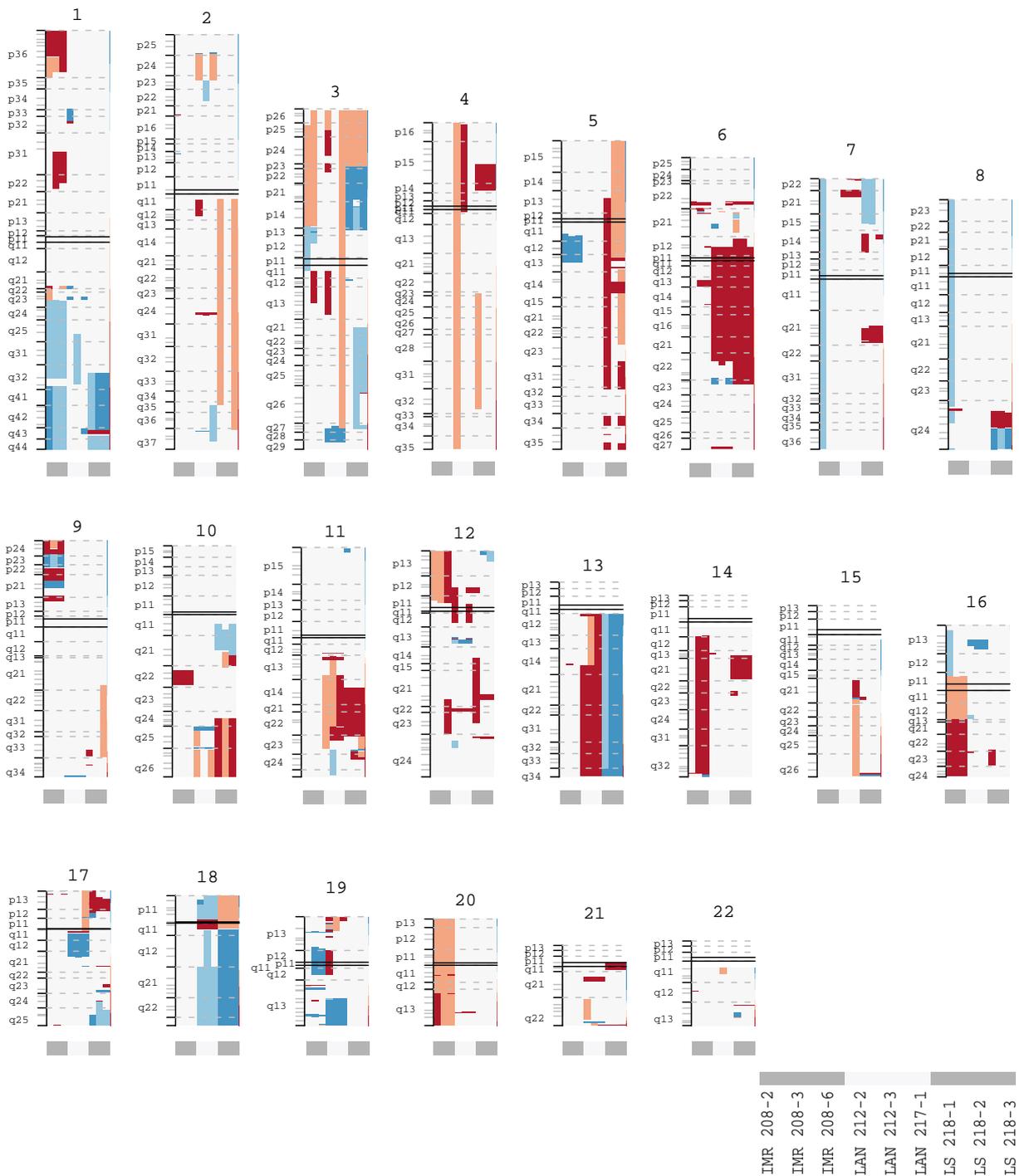


Abbildung 3.25: Neuroblastome: Besonders hohe und niedrige Expressionswerte nach der Standardisierung der Expressionswerte mit 5 in GEO veröffentlichten Datensätzen. Farbkodierung: Sehr hoch - dunkelrot, hoch - rot, niedrig - blau, sehr niedrig - dunkelblau.

3.3. KORRELATION

Symbol	EXP ID	SNP/CN ID	Raw P-Wert	R^2	Q 0,01	Q 0,001	Chr.	Position
<i>LS spezifisch</i>								
FAIM	220643_s_at	SNP_A-2131205	5,641	0,965	1,245	5,946	3	139810231
CAMK2N2	230706_s_at	SNP_A-4263948	3,336	0,844	2,268	3,012	3	185459696
SESN1	218346_s_at	SNP_A-2059510	5,201	0,954	2,274	4,618	6	109414339
PKIA	226864_at	CN_374786	5,631	0,965	2,689	5,833	8	79590890
CAMK2G	212669_at	SNP_A-8386882	2,815	0,783	2,211	3,69	10	75242264
KLF12	229881_at	SNP_A-2082448	5,997	0,973	2,165	5,288	13	73158150
CALM1	209563_x_at	SNP_A-8559964	2,598	0,75	2,349	3,81	14	89933125
MAGEC1	206609_at	SNP_A-8504436	7,103	0,987	1,033	7,381	23	140819345
ATP2B3	207026_s_at	SNP_A-8399751	4,368	0,92	1,603	5,036	23	152454773
<i>LAN spezifisch</i>								
IGSF9	229276_at	SNP_A-8370689	5,776	0,968	5,122	6,549	1	158163455
RGS5	218353_at	SNP_A-2128736	6,334	0,978	6,026	7,021	1	161378721
CNTN2	230045_at	SNP_A-2245056	6,878	0,985	4,264	5,781	1	203278962
RRP15	214764_at	SNP_A-1886610	5,573	0,964	3,288	5,281	1	216525251
CHML	226350_at	SNP_A-8428684	5,288	0,956	3,276	5,009	1	239858790
MYCN	209756_s_at	CN_819701	5,361	0,958	4,507	5,708	2	15998133
PCOLCE2	219295_s_at	SNP_A-1873067	6,612	0,982	2,232	3,887	3	144019401
MED28	222636_at	SNP_A-2009619	5,273	0,956	3,953	5,64	4	17225370
TP53	211300_s_at	SNP_A-8484217	5,07	0,95	3,549	5,514	17	7512444
<i>IMR spezifisch</i>								
FBXO9	212991_at	SNP_A-8592647	5,718	0,967	2,586	3,525	6	53043729
PKIB	223551_at	SNP_A-1926493	5,255	0,955	2,677	4,335	6	122973075
WTAP	210285_x_at	SNP_A-4253618	5,813	0,969	3,11	4,825	6	160068141
MCPH1	228778_at	SNP_A-8474975	5,67	0,966	4,07	5,623	8	6251528
DKK1	204602_at	SNP_A-2104364	5,737	0,967	5,06	5,675	10	53744046
MMD	203414_at	SNP_A-8395221	6,304	0,978	2,875	4,359	17	50824972
<i>Gene, die alle drei Linien unterscheiden</i>								
GORASP1	215749_s_at	SNP_A-1916117	6,539	0,981	2,534	5,174	3	39113094
MED28	222635_s_at	SNP_A-2009619	5,288	0,956	3,944	5,314	4	17225370
KCNMA1	221583_s_at	SNP_A-2029301	6,339	0,978	4,104	5,412	10	78299369
KCTD12	212192_at	SNP_A-1882800	6,646	0,982	2,315	6,01	13	76352304
RAP2A	225585_at	SNP_A-8412923	5,14	0,952	2,286	5,073	13	96884476
BDKRB1	207510_at	SNP_A-8319774	6,014	0,973	3,162	4,679	14	95792311
AKT1	207163_s_at	SNP_A-1870603	5,092	0,95	2,46	3,85	14	104306733

Tabelle 3.18: Neuroblastome: Die aufgeführten Gene stellen eine Auswahl der Top-Kandidaten dar, die eine gute Korrelation zwischen Expressionsniveau und CNV-Analytik aufweisen. Angeführt sind die Identifier der Gene auf dem Expressionsarray sowie die Identifier der am dichtesten gelegenen Probesets auf dem SNP-Array. $Q\ 0,01$ und $Q\ 0,001$ bezeichnen die P-Werte, die die 0,01 bzw. 0,001 Quantile denotieren. Beim *Raw P-Wert* handelt es sich um den unkorrigierten P-Wert. Die P-Werte sind in negativer, dekadisch-logarithmischer Form angegeben. R^2 gibt den Korrelationskoeffizienten an, der im Idealfall (absolute Übereinstimmung) bei 1 liegt. Die komplette Liste der als gut korrelierend eingestuften Gene umfasst 2139 Probesets und wird aufgrund ihres Umfangs auf dem beiliegenden Datenträger veröffentlicht.

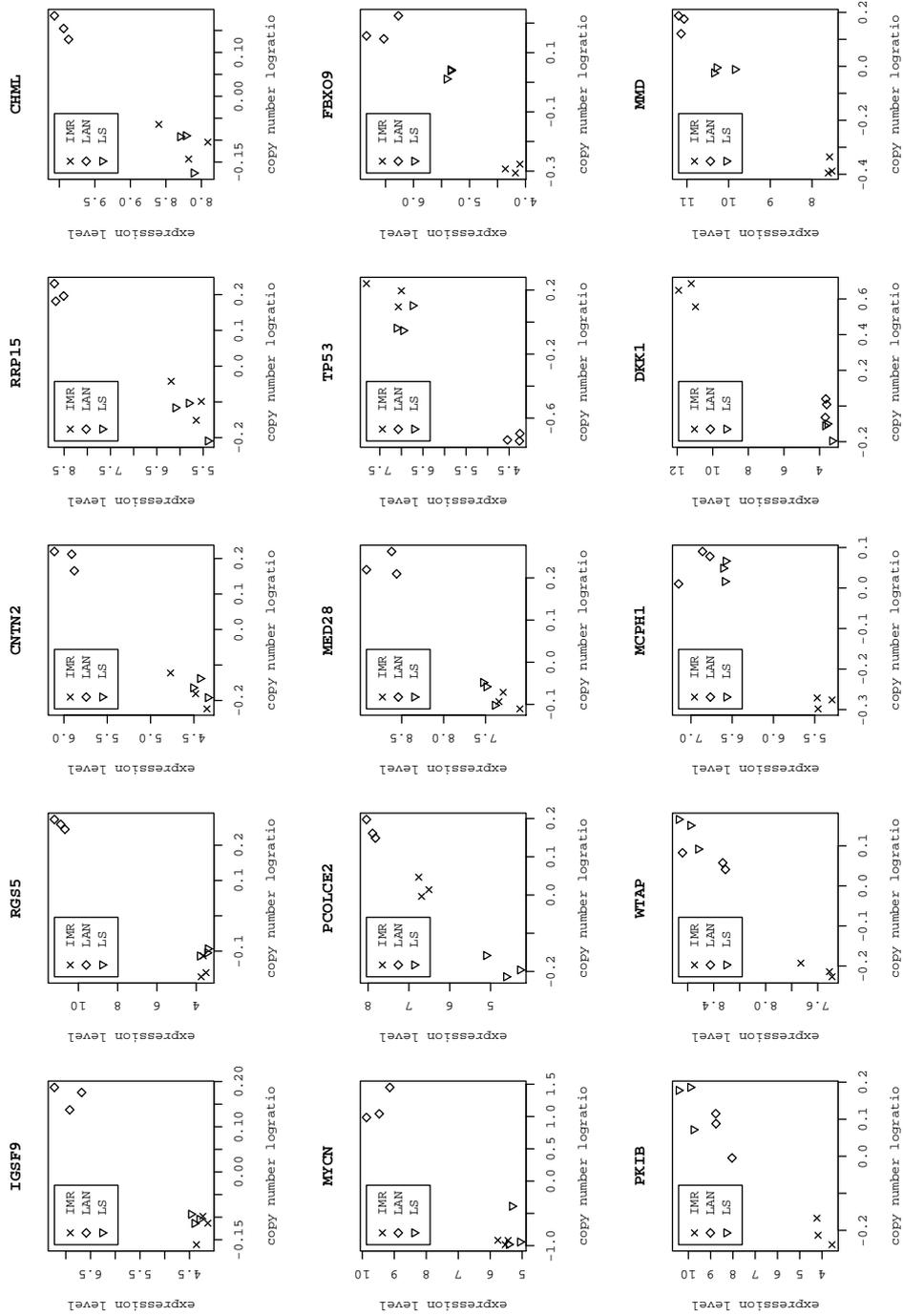


Abbildung 3.26: IMR-32 und LAN-1 Neuroblastome: Gegenüberstellung von Expressionsniveaus und segmentierter logarithmischer SNP-Analytik für die 15 top korrelierenden Gene, die sich für die IMR 32 und LAN 5 Zelllinie als spezifisch gezeigt hatten. Siehe auch Tabelle 3.18.

3.3. KORRELATION

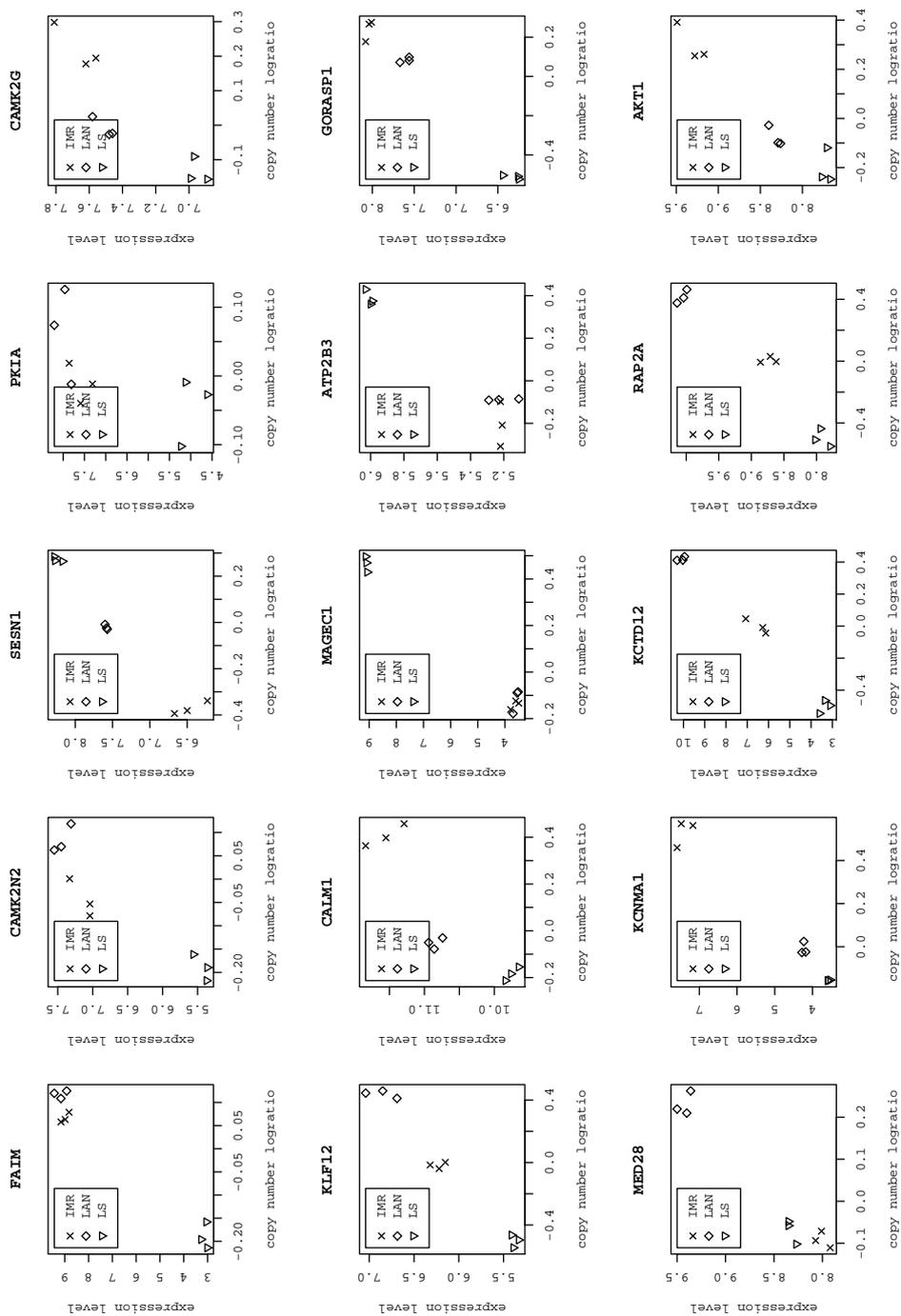


Abbildung 3.27: LS Neuroblastome: Gegenüberstellung von Expressionsniveaus und segmentierter logratios der SNP-Analytik für die 15 top korrelierenden Gene, die sich für die LS Zelllinie als spezifisch gezeigt hatten bzw. alle drei Zelllinien unterschieden. Siehe auch Tabelle 3.18.

3.4 MAPK-Arrays

Zur Bestimmung der Aktivität von 8 Transkriptionsfaktoren des MAPK-Transduktionsweges wurden Proteinarrays der Firma Eppendorf eingesetzt. Die mittels Bradfordtest ermittelten Proteinkonzentration lagen zwischen $2,9\mu\text{g}/\mu\text{l}$ und $4,1\mu\text{g}/\mu\text{l}$, siehe Tabelle 3.19.

Mittels einer Silberfärbung auf einem Bis-Tris Gradientengel (4-12%) sollte bereits im Vorraus überprüft werden, ob sich Anreicherungen spezieller Proteinbanden erkennen lassen. Das Ergebnis der Silberfärbung ist in Abbildung 3.28 wiedergegeben. Es zeigte sich, daß keine besonderen Größenverteilungen zu diesem Zeitpunkt nachgewiesen werden konnten. Die Variation der eingesetzten Proteinmenge hatte hier keinen Einfluss auf das Ergebnis. Die jeweils drei Banden mit unterschiedlicher Einsatzmenge (100ng, 200ng, 400ng) wiesen vergleichbare Bandenmuster auf.

Aktivitäten Auf den MAPK-TF-Chips konnten einige Transkriptionsfaktoren nachgewiesen werden. In allen drei Zelllinien wurden cMyc und MEF2 detektiert. cMyc bewegte sich dabei in allen drei Zelllinien auf nahezu gleich niedrigem Niveau. In der LS-Zelllinie überstieg das MEF2-Signal jene aus den LAN und IMR-Zelllinien um mehr als das Dreifache. Die LAN-Zelllinie wies im Vergleich zu den IMR-Zellen ein rund anderthalbfaches MEF2-Niveau auf. Das Tumorsupressorprotein p53 konnte in den IMR und LS-Zelllinien nachgewiesen werden, nicht jedoch in der LAN-Zelllinie. Zudem lag das Signalniveau des p53-Proteins bei den IMR-Zellen deutlich höher als in den LS-Zellen.

Probe (2 μl)	Konzentration $\mu\text{g}/\mu\text{l}$	Extinktion
LS	2,97	0,389
LAN 5	4,10	0,519
IMR 32	3,67	0,468

Tabelle 3.19: Extinktionen und berechnete Konzentrationen der Proteinextrakte aus den Neuroblastomzellkulturen. Für die Messung wurden jeweils $2\mu\text{l}$ Extrakt eingesetzt. Die Proteinextrakte wurden später für die Analyse der Aktivitäten von 8 Transkriptionsfaktoren des MAPK-Signaltransduktionsweges auf Proteinarrays der Firma Eppendorf verwendet. Für die Arrays wurden später $30\mu\text{g}$ Protein eingesetzt.

3.4. MAPK-ARRAYS

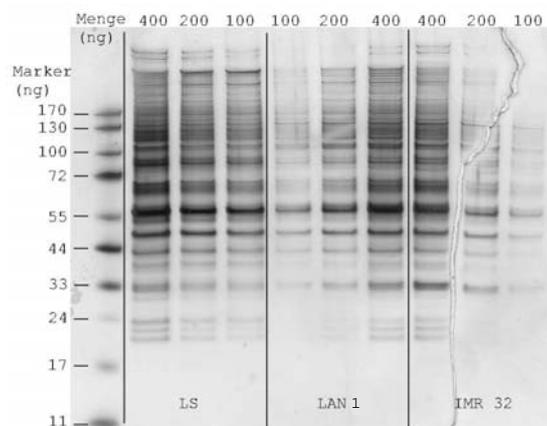


Abbildung 3.28: Gelelektrophorese (NuPage 4-12% Bis-Tris Gel) mit Silberfärbung der Proteinextrakte aus den Neuroblastomzelllinien mit jeweils drei verschiedenen Einsatzmengen. Es sollte vor der Hybridisierung auf den TF-Chips geprüft werden, ob sich Anreicherungen spezieller Proteinfractionen zeigen.

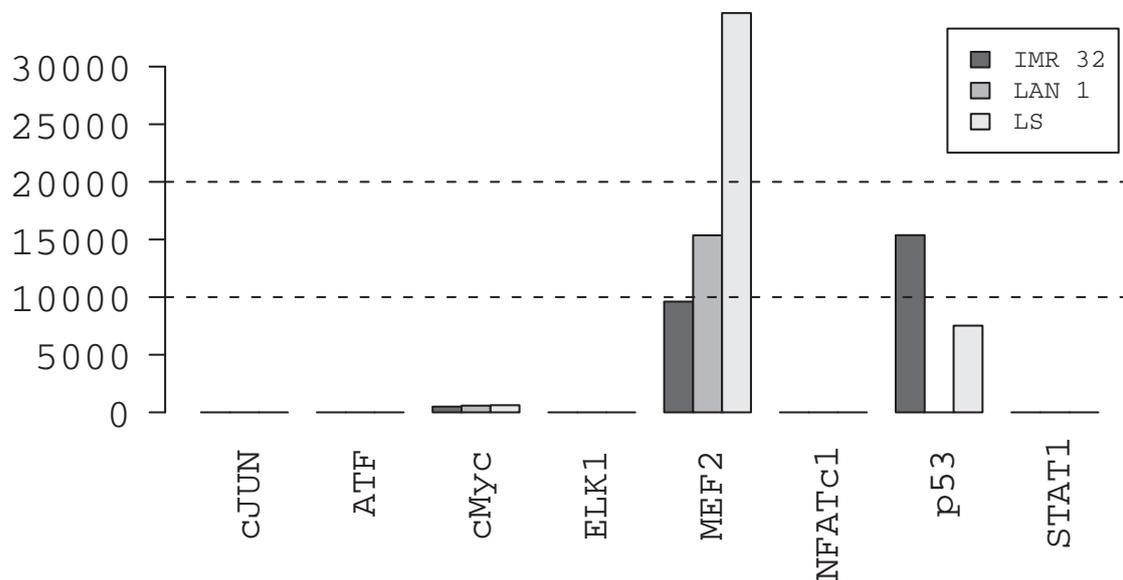


Abbildung 3.29: Signale der Transkriptionsfaktoren des MAPK-Arrays der Firma Eppendorf für alle drei Neuroblastomzelllinien. Drei der Transkriptionsfaktoren konnten nachgewiesen werden. Gut erkennbar ist die Abwesenheit des p53 Tumorsuppressorproteins in der LAN-Zelllinie.

Kapitel 4

Diskussion

In der vorliegenden Arbeit sollten Determinanten der Differenzierung und Metastasierung von Tumoren untersucht werden. Ein besonderes Augenmerk lag dabei auf dem Potential der Integration von Genexpressionen und Copy-Number-Variationen. Um modellbehaf-tete und entitätsbedingte Einflüsse bei der Bewertung der Korrelationsmöglichkeiten zu reduzieren, wurden zwei unterschiedliche Tumormodelle analysiert. Die Untersuchungen behandelten zwei sowohl medizinisch als auch volkswirtschaftlich wesentliche Tumorarten, Mammakarzinome und Neuroblastome. Die häufigste Krebserkrankung bei Frauen in Deutschland ist der Brustkrebs [74]. Im Kindesalter hingegen stellt das Neuroblastom nach der akuten lymphatischen Leukämie eine der häufigsten bösartigen Krebserkrankungen [75] dar. Nach wie vor sind viele Determinanten der Differenzierung in Brusttumoren und des Metastasierungsverhaltens von Neuroblastomen unbekannt.

Für die tumorrelevanten Gene in den beiden Modellen ergaben sich inhaltlich, wie vor dem Hintergrund der biologisch sehr verschiedenen Tumormodelle zu erwarten war, sehr unterschiedliche Ergebnisse. Die Integration von Genexpressionen und genomischen Daten wurde als vielversprechende Methode zur weitergehenden Analyse und Interpretation identifiziert. In der folgenden Diskussion werden die Beobachtungen einzelner Abschnitte zusammengefasst und in einen entsprechenden biologischen Kontext gesetzt.

4.1 Brusttumore

Das hier verwendete WAP-T Mausmodell [142, 88] wurde zur Analyse des Einflusses von mutiertem p53 auf den Differenzierungsgrad in Brusttumoren verwendet. Es handelt sich hierbei um ein transgenes BALB/c-Mausmodell, in dem mittels einer Induzierung des WAP-Promotors eine Expression der frühen Genregion des simian Virus 40, des *SV40*

T-Antigen, gezielt hervorgerufen werden kann. Das SV40 T-Antigen inhibiert die Funktion des Wildtyp-p53. Zusätzlich zum *SV40 T-Antigen* kann in bitransgenen Mäusen die Expression einer mutierten *Trp53* Variante, *mutP53*, hervorgerufen werden.

Die hier untersuchten Mammakarzinome unterschieden sich hinsichtlich Tumorgrading, der Expression von *mutP53* und dem Mausmodell. Es sollte untersucht werden, welche Faktoren den Differenzierungsgrad der Tumore in definierten Tumormodellen charakterisieren. Weiterhin waren die Einflüsse des mutP53 Proteins und SV40 T-Antigen auf die Differenzierung von Interesse. Einerseits konnten die Ergebnisse der Expressionsarrays, CNV-Bestimmungen und Korrelationen gut in Einklang mit der Literatur gebracht werden. Von zentraler Bedeutung war hierbei der *Met*-Locus auf Chromosom 6. Andererseits zeigte sich eine teils große Heterogenität, die die Tumore als singulare Entitäten erscheinen ließen.

Expressionsdaten Die Vergleiche der verschiedenen Gruppen ergaben eine Vielzahl von signifikant differentiell exprimierten Genen. Bis zu 9352 Gene charakterisierten die unterschiedlichen Gruppen. Im Folgenden werden daher auszugsweise die am stärksten regulierten Gene behandelt.

Determinanten der Differenzierung Der Vergleich der differenzierten Tumore (G1 & G2) mit den undifferenzierten (G3 & G4) zeigte eine signifikant differenzielle Expression in 2870 Genen. Mit am stärksten hochreguliert in den undifferenzierten Tumoren waren die Gene *insulin-like growth factor binding protein 2 (Igfbp2)*, *glycosylation dependent cell adhesion molecule 1 (Glycam1)*, *Suppression of tumorigenicity 7 (St7)* und *met proto-oncogene (Met)*. Vergleichsweise am niedrigsten exprimiert waren die beiden Gene *cytochrome P450, family 2, subfamily d, polypeptide 10 (Cyp2d10)* und *solute carrier organic anion transporter family, member 1a5 (Slco1a5)*. Weitere niedrigere Expression in den höhergradigen Tumoren wiesen das *Breast cancer 1* Gen (*Brca1*, SLR=-0,45) und das *Breast Cancer 2* Gen (*Brca2*, SLR=-0,30) auf.

Igfbp2 spielt eine Rolle in der Zellproliferation. Die hier beobachtete höhere Expression von *Igfbp2* in undifferenzierten Tumoren im Vergleich zu differenzierten Tumoren bestätigt bisherige Publikationen. So konnte während der Differenzierungsphase von undifferenzierten Epithelzellen der Brust eine Herunterregulation des Sekretionsspiegels von *Igfbp2* beobachtet werden [125]. Prolactin kann die Sekretion des Proteins fördern [125]. Ebenfalls

4.1. BRUSTTUMORE

auf Prolactin reagieren die IFN- γ -aktivierten Promotorelemente von *Glycam1* in Brustepithelzellen [73, 72]. Prolactin induziert dabei die *Glycam1* Expression über den STAT5-Signaltransduktionsweg. Glycam1 ist ein L-Selektin Ligand, der in der Laktationsphase als hocheprimiert beobachtet wurde [65, 118] und desweiteren in Reaktion auf inflammatorische Stimuli verstärkt im Zellplasma beobachtet werden konnte [158]. L-Selektine werden hierbei an der Oberfläche von Leukozyten exprimiert. Sie fördern deren Bindungsfähigkeit an Endothelzellen.

St7 wurde in Brustkrebs, Darmkrebs und Prostatakrebs ein tumorsupprimierender Effekt zugeschrieben [188]. In unabhängigen Arbeiten wurde hingegen eine Überexpression des Proteins mit der Malignität der Tumore korreliert [185]. Das murine Gen liegt in direkter Nachbarschaft von *Wnt2* und *Met* auf Chromosom 6qA2. Den eigenen Resultaten zufolge scheint die Überexpression von *St7* eher mit einer steigenden Malignität der Tumore einherzugehen.

Met ist ein Rezeptor für HGF, formal *Hepatocyte Growth Factor*. Ihm wird immense Relevanz in der Biologie von Tumoren zugeschrieben. Mittels Caspase-Hemmung kann *Met* negativ in die Apoptoseregulation eingreifen [57]. Das Molekül fördert den PI3K/AKT-Signaltransduktionsweg [113] und damit die Proliferation. *Met* ist über seine ERK, PI3K/AKT, und EAC1/p38 Modulation für die corticale Neuronmigration notwendig [144]. In Kooperation mit CD44 soll *Met* in der Synapsenentwicklung und Myelinisierung von Axonen im zentralen und peripheren Nervensystem beteiligt sein [104]. Dem Rezeptor konnte ein Einfluss auf die Differenzierung von Monozyten zu Makrophagen nachgewiesen werden [38]. Unter Stresstimulationen soll *Met* von einem Förderer des Zellüberlebens zum proapoptotischen Faktor werden [162]. Die Überexpression der Wildtypvariante kann zur Entwicklung nichtprogressiver Neoplasien führen, die sich bei zusätzlicher Überexpression von *Myc* zu Mammakarzinomen entwickeln können [179]. Amplifikationen in Brusttumoren sind bereits beschrieben [155]. In der vorliegenden Arbeit konnte die Hochregulation von *Myc* ebenfalls gezeigt werden. Während der Fold Change von *Met* bei rund 5fach lag, war *Myc* 1,7fach reguliert. Das kodierende Gen ist eines der Gene, die von p53 reguliert werden.

Cyp2d10 gehört zur Familie der Cytochrome P450. Diese Cytochrome sind primär am oxidativen Abbau von Substanzen und teilweise der Biosynthese von Steroiden und Prostaglandinen beteiligt. Die genaue Funktion des *Cyp2d10* ist noch weitestgehend unbekannt. *Slco1a5* ist ein Anion-getriebenes Transportermolekül, das vermutlich an der Absorption

von Gallensäuren beteiligt ist [176]. Neben *Slco1a5* waren weitere *solute carrier organic anion transporter* Gene, wie *Slc2a6*, *Slc5a8*, *Slc25a13*, *Slc44a2*, *Slc44a3*, *Slc39a11* oder *Slc9a3r2* runterreguliert.

Brcal und *Brcal2* sind Tumorsuppressorgene, die für ca. 80% der familiären, früh auftretenden Brustkrebsfälle verantwortlich sind [175]. Sie sind an der Reparatur von DNA-Doppelstrang-Brüchen [175, 53], der Transkriptionsregulation [33, 178], Zellzyklusregulation [114] und an p53-abhängiger Apoptose beteiligt. Mittels einer Phosphorylierung bzw. direkter Interaktion mit p53 [55] übt *Brcal* weiteren Einfluss auf Zellzykluskontrolle und Reparaturmechanismen aus. In den Genen wurde eine Vielzahl von Mutationen nachgewiesen, die das Risiko an Brustkrebs zu erkranken stark erhöhen. Gemäß jüngeren Forschungen können *Brcal* defiziente Brusttumore den Verlust durch eine Höherregulation von *Rad51* und zwei seiner, mit Verzögerung agierenden, Co-Faktoren *Rad54* und *Rad51ap1* ausgleichen [102]. Daraus würde eine fehlerhafte homologe Rekombination folgen, die zur Tumorgeneese führen könnte. Eine signifikante Regulation von *Rad51* konnte in dieser Arbeit nicht beobachtet werden. *Rad54* hingegen war in differenzierten Tumoren rund 1,5fach höher reguliert. Da *Rad54* später als *Rad51* exprimiert wird [102] und die Tumore sich bereits in einem Stadium hoher genomischer Instabilität befanden, könnte die signifikante Regulation von *Rad51* zu einem früheren Zeitpunkt stattgefunden haben.

Zusammenfassend sprachen die Resultate differenzierungsspezifischer Determinanten für einen deutlichen Einfluss der Amplifikation des *Met*-Locus. Neben *Met* und *St7*, wiesen die chromosomal nahegelegenen *Cav1*, *Cav2* und *Capza2* Hochregulationen auf. Es kann vermutet werden, daß *Met*, *Myc* und *St7* im Zusammenhang mit der Entwicklung von höhergradigen Tumoren stehen. Die gemäß der Funktionsanalyse vermehrte Transkriptionstätigkeit könnte auf das verstärkte Tumorwachstum zurückzuführen sein. Nicht zuletzt sind auch für *Cav1* und *Cav2* tumorfördernde Wirkungen in Tumoren fortgeschrittenen Stadiums beschrieben [147, 51].

Einfluss des mutP53 Die bitransgenen Mäuse sollten im Vergleich zu den monotransgenen zusätzlich zum *SV40 T-Antigen* die mutierten *Trp53* Gene exprimieren. Durch einen Vergleich der beiden Gruppen sollte daher der Einfluss der mutierten *Trp53* Gene auf die epitheliale Differenzierung untersucht werden. Im Rahmen dieser Arbeit konnte keine signifikante differentielle Expression zwischen *mutP53* exprimierenden und nichtexprimierenden Tumoren festgestellt werden. Bei der dreistufigen Validation der Unterschiede (Vergleiche

4.1. BRUSTTUMORE

3-5) lag die Übereinstimmung der drei Genlisten bei 97 %. Da in den nicht übereinstimmenden Genlisten zudem kein funktioneller Zusammenhang identifiziert werden konnte, wurden die geringen Unterschiede als technische Varianzen eingestuft.

Die Expression von *Trp53* innerhalb der Gruppen war heterogen. Das Clustering der vom p53 regulierten Gene konnte weder zwischen mono- und bitransgenen Samples noch zwischen differenzierten und undifferenzierten Tumoren unterscheiden. Auch hier war somit ein Einfluss des mutP53 Proteins auf die Differenzierung der Tumore nicht nachvollziehbar. Innerhalb der NP8-Tumore betrachtet erwies sich *Trp53* mit rund 11fach höherem Signal als einziges Gen stärker zwischen mono- und bitransgenen Samples differentiell exprimiert. *Bmp8a*, *Rnf10*, *Dnajc17*, *Cpne3*, *Cep350*, *Gatad2a* und *Psmc7* wiesen zwar einen signifikanten P-Wert auf, die Regulationsstärken bewegten sich jedoch lediglich zwischen 1,7facher Herunterregulation und 1,4facher Hochregulation. Desweiteren scheinen die Gene in keinem direkten, biologischen Zusammenhang zu stehen. Zusammenfassend musste aus den Resultaten daher der Schluss gezogen werden, daß ein Einfluss des mutP53 auf die Tumordifferenzierung zum erfolgten Zeitpunkt der Tumorsektion nicht nachweisbar ist.

Die vorliegenden Beobachtungen lassen mehrere Deutungen zu. Das Wildtyp-p53 Protein wird erst nach der Translation auf Proteinebene vom SV40 T-Antigen gebunden, damit am Eintritt in den Nukleus gehindert und in der Folge seine Aktivität gehemmt. Die Genexpression bleibt dabei unbeeinflusst. Es war daher in allen Samples eine vergleichbar hohe Expression des Wildtyps zu erwarten. Unterschiede auf Expressionsebene wären demnach auf die zusätzliche Expression des mutierten *Trp53* zurückzuführen. Dies konnte im NP8-Modell beobachtet werden, nicht jedoch im T1-Modell. Für die mangelnde differentielle Expression stellte sich die Frage, ob das *mutP53*-Transgen im T1-Modell überhaupt durchgehend exprimiert wird.

Im NP8-Modell wurde eine differentielle Expression des *Trp53* beobachtet, für die jedoch kein Einfluss auf die weitere Genexpression nachweisbar war. Die Repression des Wildtyp-p53 könnte hierbei durch das SV40 T-Antigen in den monotransgenen Mäusen nur teilweise erfolgt sein. In diesem Fall fände eine Überlagerung zwischen den Wirkungen des mutierten und des Wildtypgenes statt. Die Einzeleffekte wären dann schwer zu unterscheiden. Diese Hypothese könnte mit einer bereits für das NP8-Modell publizierten Beobachtung gestützt werden, derzufolge lediglich in 50% der NP8-Tumorzellen das SV40 T-Antigen exprimiert wird [142]. In jenen Zellen, die kein SV40 T-Antigen exprimieren,

fände auch keine Hemmung des Wildtyp-p53 statt. Eine weitere Erklärung wäre ein SV40-mutP53 Komplex. Hierbei könnte das SV40 T-Antigen das mutP53 einzeln oder zusammen mit dem Wildtyp-p53 in größeren Komplexen binden. Es würde damit gleichzeitig mit der Wildtyp-p53 Funktion auch die Wirkung des mutP53 hemmen.

Eine alternative, putative Erklärung bestünde in der Funktionslosigkeit des mutP53. Dagegen spricht allerdings, daß in größeren Tumorkollektiven des NP8-Modells bereits eine erhöhte Inzidenz an höhergradigen Tumoren (G3 & G4) in bitransgenen Mäusen beschrieben wurde [69]. Dieser Zusammenhang wird von den Ergebnissen dieser Arbeit nicht unterstützt. Allerdings könnte der Zeitpunkt der Tumorprobenentnahme hier die entscheidende Rolle spielen. Setzt die Wirkung des mutP53 Proteins kurzzeitig nach der Laktation ein, so würde diese zum Zeitpunkt der Tumorprobenentnahme bereits von den nachfolgenden Expressionsdysregulationen überlagert sein. Die allgemein in sämtlichen Tumoren stattfindenden Prozesse, wie verstärkte Proliferation (z.B. Igfbp2, Met), reduzierte Apoptose (z.B. Met, Myc), Zellzykluskontrollen (z.B. Brca1, Brca2) oder anti-inflammatorische Reaktionen (z.B. Glycam1), wären dominant ausgeprägt und ließen zu diesem Zeitpunkt den mutP53-Effekt nicht mehr erkennen. Die äußerst hohe Rate an Dysregulationen (57%) würde diese Theorie untermauern. Sie zeigt, wie deutlich sich die Prozesse in Tumoren von jenen im Normalgewebe unterscheiden. Insbesondere der Einfluss des zu diesem Zeitpunkt hochregulierten *Met*-Locus in den undifferenzierten Tumoren kann als sehr dominant eingestuft werden. Als Schlüsselprotein bei der Steuerung von DNA-Reparaturmechanismen und Regulator des *Met* Gens wäre ein Zusammenhang des p53 mit der Amplifikation des *Met*-Locus durchaus denkbar. Ob das mutierte Protein ursprünglich zu einer erhöhten Amplifikationsrate des *Met*-Locus führt, müsste noch validiert werden.

Funktionsanalyse Erwartungsgemäß konnten in der differenzierungsspezifischen Genliste verstärkt translatorische Funktionsklassen beobachtet werden. Es konnten vermehrt in der Translation involvierte Gene, ribosomenkonstituierenden Gene der Klassen *Structural constituent of ribosome* und *Ribosomal biogenesis and assembly* sowie Gene der GO-Klassen *Macromolecule biosynthetic process* und *Cellular macromolecule metabolic process* beobachtet werden. Diese Ergebnisse reflektieren das vermehrte Zellwachstum in den undifferenzierten Tumoren und die damit verbundene, verstärkte RNA- und Protein Synthese. Diese Analyse identifizierte zwar keine potentiellen onkogenen Einzelkandidaten, konnte

4.1. BRUSTTUMORE

jedoch zur Strukturierung der Resultate und zum Verständnis der involvierten Prozesse beitragen.

Beim Vergleich der Tumore mit den Wildtypen wurden mehr als 30 Funktionsklassen identifiziert, die von der Transkription, über Morphogenese und Zelladhäsion bis zu Immunantworten sehr verschieden waren. Diese Bandbreite an Funktionsklassen kann vermutlich auf die außerordentlich große Anzahl an regulierten Genen zurückgeführt werden. Bei einer Regulation von rund 57% aller erfassten Gene auf einem *Whole Genome Chip* steigt bereits aus statistischer Sicht die Wahrscheinlichkeit stark an, eine der Funktionsklassen zufällig zu identifizieren. Die Funktionsbandbreite hebt das Ausmaß der Abnormität der in Brusttumoren ablaufenden Prozesse, im Vergleich zu den Normalgeweben, noch deutlicher hervor.

Copy-Number-Variationen Die Analyse von genomischen Aberrationen in den Brusttumoren sollte die Expressionsanalytik ergänzen. Die bereits auf Expressionsebene beobachtete Regulation des *Met*-Locus konnte hierbei sehr gut auf Genomebene nachvollzogen werden. Die Resultate der durchgeführten Korrelationen bestätigten einen direkten Zusammenhang zwischen Aberrationen und Expressionsniveau.

Aberrationen sind in Brusttumoren weit verbreitet. Häufigere Amplifikationen wurden bei den Genen *Fgfr1*, *Fgfr2*, *ErbB2* oder *Myc* beobachtet. *Brca1* ist nicht selten deletiert [175]. Bei den Tumoren des in dieser Arbeit verwendeten Mausmodells handelte es sich zumeist um tetraploide Zellen [69]. Der genomische Hintergrund der Samples befand sich damit bereits nicht mehr im Normalzustand.

Die Aberrationen in den einzelnen Tumoren waren sehr heterogen und ließen keine einheitlichen Profile erkennen. Zumeist lagen Gains oder Losses ganzer Chromosomen vor. Aberrationen einzelner Loci traten seltener auf. Sollten die Aberrationen die Entwicklung der Tumore widerspiegeln, so erweckten die Samples den Eindruck singulärer Entitäten, die sehr unterschiedliche Wege der Entwicklung durchlaufen hatten. Da das Mausmodell einen Verlust des Wildtyp-p53 simuliert, das eine entscheidende Rolle für die Genomstabilität besitzt, ist die außerordentlich hohe Anzahl an Aberrationen nicht verwunderlich. Desweiteren müsste der einfache oder sogar zweifache Loss bzw. Gain von Chromosomen bei den tetraploiden Zellen nicht zwangsläufig Konsequenzen haben. Homo- und Heterozygotien, die im Falle eines zweifachen oder dreifachen Verlustes zum Tragen kommen könnten, waren anhand der array-CGH Technik nicht zu bestimmen. Ein größerer Einfluss wäre daher

erst bei totalem Verlust einer Region oder massiven Amplifikationen nachvollziehbar.

Vor diesem Hintergrund wurden zwei Beobachtungen als die wichtigsten Resultate dieser Analyse identifiziert. Erstens wiesen die spontanen Mammakarzinome auffällig wenige Aberrationen im Vergleich zum Gesamtkollektiv auf. Die spontanen Brusttumore haben sich vor der Laktation ohne Einfluss des SV40 T-Antigen und *mutP53* oder sogar in Wildtypmäusen entwickelt. Die deutlich geringere Anzahl an genomischen Aberrationen machte deutlich, daß die Aberrationsdichte in den induzierten Tumoren auf die Induzierung über den *WAP*-Promotor zurückgeführt werden mussten. Da die *mutP53* exprimierenden Tumore keine spezifischen Aberrationen aufwiesen, konnte davon ausgegangen werden, daß die massiven Aberrationen allein durch den Verlust des funktionellen Wildtyp-p53 und den Reiz durch das SV40 T-Antigen verursacht wurden.

Zweitens konnte eine Amplifikation des *Met*-Locus in 6qA2 beobachtet werden. Die *Met*-Amplifikation korrelierte sehr gut mit den undifferenzierten, höhergradigen Tumoren. Der Anteil der Beobachtungen einer Amplifikation des *Met*-Locus in undifferenzierten Tumoren (23 von 25) war signifikant größer als in den differenzierten Tumoren (8 von 27). Ein Exakter Fisher-Test ergab einen P-Wert von $3,9 \cdot 10^{-6}$. Abbildung 3.22 zeigt, daß neben einer Amplifikation des *Met* Gens häufig auch eine Amplifikation der benachbarten Gene, *Tes*, *Cav2*, *Cav1*, *Capza2*, *St7* und *Wnt*, zu beobachten war. Die Amplikongrenzen konnten sich von Tumor zu Tumor deutlich unterscheiden.

Amplifikationen des Locus im Zusammenhang mit der Aggressivität von Darmkrebs, Lebermetastasen, Magenkrebs aber auch Brusttumoren wurden bereits mehrfach beschrieben [187, 60, 155, 146]. Es konnte also vermutet werden, daß die Amplifikation dieser Region maßgeblichen Anteil an der Entwicklung zu höhergradigeren Tumoren hatte. Es bliebe allerdings zu klären, ob es sich hierbei um die Ursache oder lediglich eine Folge der Entwicklungsstufe handelt.

Korrelation Korrelationen von genomischen Daten und Genexpressionen wurden bisher in verschiedensten Formen durchgeführt und beschrieben. So wurden bereits Clustering-Verfahren [186], Pearson und Spearman Korrelationen [173] oder Wilcoxon Hypothesentests [41] in Brusttumoren, gewichtete Mittelwertdifferenzen [10] in Pankreaskarzinomen

4.1. BRUSTTUMORE

oder Regressionsverfahren [157] eingesetzt. Gemein ist diesen Verfahren, daß bisher die Korrelation einzelner Gene betrachtet wurde. Eine Darstellung der genomweiten Zusammenhänge wurde bisher nicht adäquat beschrieben. Während einfache Strategien, wie Pearson, den Vorteil einer kurzen Kalkulationszeit bieten, wird in dem hier verwendeten Verfahren eine erhöhte Aussagekraft zur Signifikanz erzielt. Die Permutationen erübrigen damit nicht nur weitere Korrekturverfahren sondern integrieren auch, gleichzeitig mit der P-Wert-Berechnung, eine Bestimmung potentieller Korrelationen mit dicht benachbarten, genomischen Loci.

Zusammenfassend konnte eine sehr gute Korrelation der genomischen Resultate mit den gemessenen Expressionswerten beobachtet werden. Insbesondere die Gene des stark amplifizierten *Met*-Locus wiesen eine sehr gute Korrelation mit den hohen Expressionswerten auf. Trotz der modellbedingten, genetischen Heterogenität wiesen sowohl mono- als auch bitransgene Tumore mit dem *Met*-Locus eine modellübergreifende genomische Aberration auf, anhand derer differenzierte und undifferenzierte Tumore klar unterschieden werden konnten. Die Amplifikation wurde von einer signifikant erhöhten Genexpression begleitet, siehe Abbildung 3.22. Dies legte den Schluß nahe, daß der Entwicklung von höhergradigen Tumoren (G3 & G4) gemeinsame Mechanismen zugrunde lagen. Diese mussten in direktem Zusammenhang mit den Aberrationen in 6qA2 stehen.

Mit *Met*, *St7*, *Capza2* und den Caveolinen als Topkandidaten zur Unterscheidung zwischen differenzierten und undifferenzierten Tumoren kam diesen Resultaten eine besondere Bedeutung zu. Obwohl *Met*-Amplifikationen in humanen Brusttumoren selten zu beobachten sind [155], wurden *Met*-Überexpression bereits seit längerem mit der Tumorprogression sowie höherer Invasivität und Metastasierung in humanen Tumoren, einschließlich Brusttumoren, in Verbindung gebracht [24, 23]. Trotz unterschiedlicher molekularer Mechanismen im hier verwendeten Mausmodell und in der humanen Karzinogenese, konnte demnach ein ähnlicher Phänotyp, eine *Met*-Überexpression beobachtet werden, die mit der Progression von Mammakarzinomen assoziiert ist. Ebenfalls auffällig war dabei die Beobachtung eines Tumors (Sample 11155-7), der bei mangelnder *Met*-Amplifikation dennoch eine deutliche *Met*-Überexpression aufwies. Dies könnte ein Hinweis darauf sein, daß eine Aktivierung von *Met* als Onkogen nicht zwangsläufig durch eine Locus-Amplifikation erfolgt. Die Beobachtung könnte aber auch darauf hinweisen, daß ein möglicher Mechanismus für chromosomale

Amplifikationen auf DNA-Fragilität beruht, die durch verstärkte Transkription induziert wird [49].

Zusätzlich zu *Met* konnte in undifferenzierten Tumoren eine Co-Amplifikation einiger benachbarter Gene, wie *St7*, *Tes* oder *Capza2*, beobachtet werden, die eine ähnlich gute Korrelation und somit ein vergleichbares Expressionsprofil aufwiesen. Während die biologische Funktion dieser Gene nach wie vor nicht endgültig geklärt ist, könnte die Co-Expression mit *Met* auf eine Beteiligung an der Tumorprogression hindeuten. Interessanterweise schien eine Co-Amplifikation nicht immer mit einer erhöhten Expression einherzugehen. So konnte z.B. für *Wnt2* keine signifikante Korrelation (P-Wert = 0,23) nachgewiesen werden. In SV40 transformierten Zellen wird der Wnt-Pathway konstitutiv durch das kleine SV40 T-Antigen aktiviert [168]. Es ist daher anzunehmen, daß die konstitutive down-stream-Aktivierung zu einer transkriptionellen Herunterregulation von *Wnt2* führen kann. Die Ergebnisse zeigten somit auch hier, daß die Integration von genomischen Daten und Genexpressionen eine vielversprechende Möglichkeit bietet, um die großen Listen von regulierten Genen auf eine kleinere Anzahl, mit größerer Aussagekraft, zu reduzieren.

Bei den meisten Tumoren dieser Modelle liegt eine Tetraploidie vor [69]. Es stellte sich also die Frage, inwiefern sich der einfache oder mehrfache Verlust bzw. Gewinn von Chromosomen sich auf transkriptioneller Ebene bemerkbar machten. Die genomweite Korrelation von Genexpressionen und genomischen Signalen sollte daher zeigen, ob selbst leichte Änderungen in der Kopienzahl sich auf die gesamte betroffene Region oder lediglich auf einzelne Gene auswirken. Bei den vorliegenden Brusttumorsamples manifestierte sich eine extrem große, genomweite Übereinstimmung zwischen den beiden Ebenen, siehe Abbildung 3.24. Sowohl starke als auch schwache Änderungen in der Kopienzahl wurden auf Expressionsebene sichtbar, siehe auch Sample 10726-5 Chr6, Chr8 und Chr19. Hierbei schien es keine Rolle zu spielen, ob ganze Chromosomen betroffen waren oder lediglich kleinere Regionen. Mehrere graduelle Amplifikationsstufen innerhalb eines Chromosoms (Sample 12192-2 Chr11) konnten genauso hervorgehoben werden, wie die Verhältnisse zwischen den verschiedenen Chromosomen, siehe auch Sample 12525-1. Dieser Effekt war bei allen betrachteten Samples gut erkennbar, unabhängig vom Differenzierungsgrad und davon, ob es sich um induzierte oder spontane Tumore handelte. Er spricht für eine hohe Sensitivität des Transkriptionsniveaus gegenüber DNA Kopienänderungen.

4.1. BRUSTTUMORE

Die genomweite Korrelationsstrategie erwies sich damit bei den Brusttumoren als sehr vielversprechende Ergänzung zur Expressionsanalytik. Wenn auch, wie bei *Wnt2* beobachtet, die Mikrosteuerung der Expression hiermit nicht erfasst werden kann, so konnte doch deutlicher als bisher beschrieben gezeigt werden, daß Änderungen in der genomischen Kopienzahl sich nicht nur auf einzelne, sondern auf einen Großteil der betroffenen Gene auswirken können. Wie unter anderem die graduellen Änderungen bei Sample 12192-2 in Chromosom 11 vermuten ließen, scheint das Expressionssignal hierbei proportional zur Aberration in der Kopienzahl zu variieren. Die Tumore wiesen trotz der unterschiedlichen Aberrationsprofile innerhalb der Gruppen von differenzierten und undifferenzierten Tumoren teilweise homogene Expressionssignaturen auf. Es findet demnach eine zusätzliche Regulationssteuerung auf Expressionsebene statt, die mit der hier beschriebenen Methode nicht erfasst werden kann.

Nachteilig wirkt sich bei der hier angewandten Methode die für die Referenzberechnung notwendige Anzahl der Samples aus. Frühere Versuche der gleichen Strategie scheiterten bei 30 Samples daran, ein repräsentatives Normalniveau für die Signale der einzelnen Gene zu bestimmen. Die wachsende Anzahl an publizierten und frei verfügbaren Datensätzen in Datenbanken, wie GEO oder ArrayExpress, dürfte diese Hürde zukünftig senken, wie es in dieser Arbeit am Neuroblastomdatensatz gezeigt werden konnte. Von großer Bedeutung andererseits war, daß die Visualisierung von genomischen Aberrationen unabhängig von CGH-Arrays oder SNP-Daten wurde. Eine Vorhersage etlicher Amplifikationen und Deletionen könnte somit bei guter Referenzgruppe bereits anhand der Expressionsdaten möglich werden.

4.2 Neuroblastome

Im Rahmen dieser Arbeit wurde das Metastasierungsverhalten von drei Neuroblastomzelllinien untersucht. Hierzu wurden die Zellen subkutan zwischen die Schulterblätter von BALB/c SCID Mäusen transplantiert. Die sich daraufhin entwickelnden Tumore unterschieden sich stark hinsichtlich ihres Metastasierungsverhaltens. Die Auswahl der Zelllinien sollte dadurch eine Untersuchung von Determinanten der Metastasierung erleichtern.

Mäuse des in dieser Arbeit verwendeten SCID Mausmodells, formal *severe combined immunodeficiency* [27, 28], weisen Mutationen in Chromosom 16 auf, die eine Differenzierung von Stammzellen zu Lymphozyten verhindern. Die Mäuse besitzen dadurch eine mangelnde Immunreaktion und das Immunsystem wird bei der Transfektion der Neuroblastomzellen umgangen.

Die IMR-32 Neuroblastomzelllinie wurde von Tumulowicz et al. etabliert [163]. Es handelt sich um eine hyperploide humane Zelllinie, die zwei morphologisch unterschiedliche Zelltypen, gleichen klonalen Ursprungs, umfasst. Beide Zelltypen entstammen demselben Zellenstamm und ließen sich von Tumulowicz et al. nicht in zwei separate Zelllinien splitten. Die aus IMR-32 Zellen entstandenen Tumore entwickelten große mehrzellige Metastasen. Die Metastasen waren primär in der Pulmonararterie und im periarteriellen Raum der Lunge lokalisiert. LAN-1 Zellen, von Seeger et al. etabliert [143], wiesen etliche kleinere Mikrometastasen auf, welche die ganze Lunge diffus infiltrierte. Die LS Neuroblastomzelllinie wurde von Rudolph et al. etabliert [136]. LS Neuroblastome wiesen keine nachweisbaren Metastasen auf. Die Ergebnisse der Expressionsarrays konnten gut mit den bestimmten Kopienänderungen in Einklang gebracht werden. Weiterhin konnten einige neue Faktoren ermittelt werden, die potentiell eine große Rolle im Metastasierungsablauf spielen. Zuletzt konnte gezeigt werden, daß nach der Transfektion in der IMR-32 Zelllinie eine Verschiebung des Gleichgewichts zwischen den zwei konstituierenden Subpopulationen stattfindet.

Expressionsdaten Im Zuge der Expressionsanalytik wurden Gene identifiziert, die im Vergleich der drei Zelllinien jeweils spezifisch signifikant reguliert waren. Es werden im Folgenden auszugsweise die am stärksten regulierten Gene dieser Listen behandelt.

IMR-32 Bei den mit IMR-32 Zellen infizierten Mäusen konnten besonders hohe Expressionen des *sperm protein associated with the nucleus, X-linked, family member A1* (*SPANXA1*, SLR=9,6), *CD44 Antigens* (*CD44*, SLR=8,8) und des *Dickkopf homolog 1*

4.2. NEUROBLASTOME

(*DKK1*, SLR=7,8) beobachtet werden. Die Gene *myosin, heavy chain 10* (*MYH10*, SLR=-6,1), *protein kinase inhibitor beta* (*PKIB*, SLR=-6,1) und *SLAIN motif family, member 1* (*SLAIN1*, SLR=-6,6) hingegen waren hier besonders niedrig exprimiert.

Tatsächlich wurde ein Zusammenhang zwischen der Überexpression von *SPANXA1* und der Wirkung als potentielles Onkogen in nichtlymphatischen metastatischen Melanomen [180] bzw. embryonalen Karzinomen [138] beschrieben. Die Expression von *CD44* in IMR-32 Zellen und fehlende Expression in LAN-1 Zellen wurde bereits beschrieben [169] sowie eine Korrelation mit Metastasierung postuliert [169, 91, 80]. Die beobachtete Überexpression in einer der beiden metastasierenden Zelllinien steht somit gut in Einklang mit der bereits publizierten Literatur. Auf *CD44* und *DKK1*, für das Einflüsse auf Prostatakarzinome, gastrointestinale Tumorgenese, Myelome und die Bildung von Knochenmetastasen beschrieben wurden, wird später, auf Seite 4.2 und 4.2, noch näher eingegangen.

Die Herunterregulation von *MYH10* wird in Zusammenhang mit einer Relokalisation von APP beschrieben [103], einem Schlüsselprotein bei Morbus Alzheimer. *PKIB* inhibiert die *cAMP-dependent protein kinase* [170]. Eine Herunterregulation von *PKIB* hätte demnach eine höhere Aktivität der *cAMP-dependent protein kinase* zur Folge. Für die 1α -Untereinheit dieser Kinase sind putative Funktionen als Tumorsuppressoren [29, 139] und ein Einfluss auf Zellmetabolismus und Proliferation, z.B. in Melanomen [132, 133, 100], beschrieben. Ein Ungleichgewicht in der Expression der 1α und 2β -Untereinheiten zugunsten der 1α -Einheit soll die Proliferation von ausdifferenzierten adrenocorticalen Zellen fördern [101]. Bei den IMR-32 Samples konnte ein 26faches Expressionsverhältnis der beiden Untereinheiten (SLR=4,7) zugunsten der 1α -Einheit beobachtet werden, während bei den LAN-1 und LS Samples der Quotient lediglich bei 3,4fach bzw. 4,5fach (SLR=1,8 bzw. SLR=2,2) lag. *SLAIN1* ist ein primär während der Differenzierungsphase von embryonischen Stammzellen exprimiertes Gen [70], das erst im Jahre 2006 charakterisiert wurde. Bisher ist ausschließlich die Erstbeschreibung bei PubMed erfasst.

LAN-1 Besonders hoch exprimiert waren bei den LAN-1 infizierten Mäusen das *achaete-scute complex homolog 1* (*ASCL1*, SLR=8,2), *regulator of G-protein signaling 5* (*RGS5*, SLR=7,4) und *v-myc myelocytomatosis viral related oncogene, neuroblastoma derived* (*MYCN*, SLR=6,9). Vergleichsweise niedrig exprimiert waren hingegen unter anderem das *prion protein* (*PRNP*, SLR=-8,6), *annexin A1* (*ANXA1*, SLR=-8,2) und *caveolin 1* (*CAV1*, SLR=-9,4).

ASCL1 ist ein Mitglied der *basic helix-loop-helix* Familie der Transkriptionsfaktoren. Es ist an der neuronalen Differenzierung beteiligt und wurde bei Überexpression mit der Tumorgenese von endokrinem Pankreas, gastrointestinalen neuroendokrinen Karzinomen und der Differenzierung von Prostatakarzinomen in Verbindung gebracht [79, 130, 149]. Das Expressionsniveau von *ASCL1* wurde weiterhin als diagnostischer Marker zur Unterscheidung von hochgradigen Esthesioneuroblastomen von Tumoren mit niedrigem Differenzierungsgrad vorgeschlagen [107]. Erst im März 2008 wurde eine mögliche Repression von *DKK1* durch *ASCL1* publiziert [122]. Möglicherweise spielt in den verwendeten LAN-1 Zellen das Molekül eine vergleichbare Rolle. *RGS5* wiederum ist ein Signaltransduktionsmolekül, das als GTPase-Aktivator für G-Proteine wirkt. Einflüsse einer höheren Expression in Pericyten auf die Remodellierung und Entwicklung von Blutgefäßen [21, 1] wurden ebenso beschrieben, wie Zusammenhänge mit einer Anfälligkeit für Schizophrenie [112, 35]. *RGS5* könnte bei den Neuroblastomen an der Tumorangiogenese beteiligt sein. Bei *MYCN* handelt es sich um ein Gen, das ein nuklear lokalisiertes Protein kodiert. Amplifikationen dieses Gens werden mit einer Reihe von Tumoren in Verbindung gebracht, allen voran Neuroblastomen. Es wird später, auf Seite 4.2, noch näher auf *MYCN* eingegangen.

PRNP wird als pathogen primär mit Creutzfeldt-Jakob, Huntington und Gerstmann-Straussler Erkrankungen in Verbindung gebracht [134, 47, 37, 110]. *ANXA1* kodiert für ein $\text{Ca}(2+)$ -abhängiges, Phospholipide bindendes Protein mit *phospholipase A2* inhibierender Funktion. Antiinflammatorische Wirkungen wurden ebenso beschrieben, wie hemmende Wirkungen auf Proliferation und Cytokinproduktion [83, 141, 61]. Niedrige Expressionen des Gens wurden außerdem bereits im Zusammenhang mit der Entwicklung von etlichen Tumoren wie Brustkrebs, Prostatakrebs und Blasenkrebs gebracht [148, 92, 183]. Eine Herunterregulation des Gens soll mit einer Metastasensuppression in Brustkrebs einhergehen [45]. Die GenMapp Analyse, siehe Abbildung 3.14, zeigte eine herabgesetzte Reaktion der LAN-1 Neuroblastome auf oxidativen Stress. Dieser steht in direktem Zusammenhang mit dem $\text{Ca}(2+)$ Haushalt der Zellen und könnte somit eine entscheidende Rolle bei der Expression von *ANXA1* gespielt haben.

CAV1 gilt als Tumorsupressorgen mit einem negativen Regulationseffekt auf den MAPK-Signaltransduktionsweg. [63, 59] und endotheliale Zellproliferation [56]. Herunterregulation dieses Gens führen potentiell zur Metastasierung in Lungenadenokarzinomen und Magenkrebs [99, 71, 150]. Eine *CAV1* Expression soll Zell-Zell-Adhäsionen in Eierstockkarzinomen fördern [108]. Weiterhin wird das Gen mit Brusttumoren und Darmkrebs assoziiert

4.2. NEUROBLASTOME

[123, 39] Nicht zuletzt wird ihm eine modulierende Rolle bei neuronalen Ca(2+) Kanalaktivitäten zugeschrieben [31, 160]. In dieser Funktion wäre eine Assoziation mit der niedrigen Expression von *ANXA1* denkbar.

LS In den LS induzierten Tumoren konnten besonders hohe Expressionen der Gene *adenomatosis polyposis coli down-regulated 1* (*APCDD1*, SLR=8,3), *EGF-like repeats and discoidin I-like domains 3* (*EDIL3*, SLR=8,7) und *fatty acid binding protein 7* (*FABP7*, SLR=8,9) beobachtet werden. Die Gene *dimethylarginine dimethylaminohydrolase 1* (*DDAH1*, SLR=-6,0), *Fas apoptotic inhibitory molecule* (*FAIM*, SLR=-6,1) und *glycerol-3-phosphate dehydrogenase 2* (*GPD2*, SLR=-5,5) wiesen ein vergleichsweise sehr niedriges Expressionsniveau auf.

APCDD1 ist ein vom *Beta-Catenin/Tcf*-Komplex reguliertes Gen, das in höherer Expression bereits in Darmkrebs beobachtet wurde [159]. In Wilms-Tumoren wurden Hochregulationen dieses Genes beschrieben [189], die potentiell zur vermehrten Proliferationsfähigkeit führen. *EDIL3* ist ein Integrin-Ligand mit Angiogenese vermittelnder Wirkung und Einfluss auf endotheliales Zellverhalten. Es führt potentiell zur Bildung und dem Wachstum solider Tumore über eine erhöhte Vaskularisation, erhöhte Zelladhäsion und reduzierte Apoptose [16, 131, 119]. Weiterhin soll es potentiell die Migrationsfähigkeit der Zellen erhöhen [16]. Diese Eigenschaften könnten *EDIL3* zu einem wichtigen Kandidaten bei der Entwicklung der metastasenfreen LS Tumore qualifizieren. Während reduzierte Apoptose und erhöhte Migrationsfähigkeit die Tumorprogression fördern, würde eine erhöhte Zelladhäsion die Metastasenbildung reduzieren.

FABP7 kodiert für ein zytoplasmisches Protein, das an Aufnahme, Transport und Metabolismus von langkettigen Fettsäuren beteiligt ist. Interessanterweise wurden höhere Expressionen im Zusammenhang mit erhöhter Tumorzellmotilität, Migration und schlechteren Prognosen bei Glioblastomen [97, 96, 82, 109], sowie beim Down-Syndrom [156] beobachtet. Hier konnten somit einerseits erhöhte Expressionen bei Genen beobachtet werden, die die Zellmigration erhöhen und im Zusammenhang mit schlechteren Prognosen stehen. Andererseits konnten Hochregulationen beobachtet werden, für die zelladhäsionsfördernde Charakteristika beschrieben wurden. Tatsächlich wurden in Tumoren von LS Zellen keine Metastasenbildungen beobachtet.

DDAH1 ist an der Regulation des Stickstoffmonoxidhaushaltes der Zelle beteiligt [93, 84] und somit an der postnatalen Angiogenese und Arteriogenese [78]. Höhere Konzentrationen des Proteins wurden, unter anderem, in Prostatakrebs nachgewiesen [165]. Stickstoff-

monoxid hat potentiell einen negativen Regulationseffekt auf MYCN und hemmt somit die Proliferation [44]. Eine negative Regulation von *DDAH1* hätte demzufolge eine proliferative Wirkung zur Folge. FAIM weist eine apoptoseinhibitorische Wirkung auf. Insbesondere die Expression der *long*-Isoform resultiert im Schutz der Neuronen vor zytotoxischen Angriffen [145]. GPD2 ist an der äußeren Oberfläche der inneren mitochondrialen Membran lokalisiert und katalysiert die Umwandlung von *Glycerol-3-Phosphate* zu *Dihydroxyaceton-Phosphate* und führt somit indirekt zu einer Reoxidation von, während der Glykolyse produziertem, NADH. Höhere Expressionen des Gens werden im Zusammenhang mit der Tumorgenese von Prostatakrebs erwähnt [42, 43]. Letztendlich müssen hier alle drei Fälle vermutlich als erhöhte Expression in IMR-32 und LAN-1 Zellen aufgefasst werden, während die Gene in LS Zellen nicht exprimiert werden. Die beiden letztgenannten Gene würden demnach das Tumorstadium in IMR-32 und LAN-1 Zellen fördern und gleichzeitig die Zellen vor Apoptosemechanismen und Cytokinen schützen.

Funktionsanalyse Anhand der mit EXPANDER¹ ermittelten Anreicherungen spezieller Funktionen konnte eine aufschlussreiche Klassifikation der regulierten Gene durchgeführt werden. Um die vorliegenden Ergebnisse zu erklären, muss man sich nochmals vor Augen führen, daß den verwendeten Genregulationen keine Normalgewebereferenz zu Grunde liegt. Es handelt sich um Vergleiche dreier Neuroblastomlinien, die alle von verstärktem Tumorstadium betroffen sind. Bei gleichem Tumorstadium in den drei Zelllinien könnten daher keine speziellen Funktionsanreicherungen entdeckt werden. Bei den ermittelten Funktionsklassen handelt es sich daher um Gruppen, die in den Zelllinien IMR-32, LAN-1 und LS trotz des grundlegenden cancerogenen Phänotyps unterschiedlich stark ausgeprägt sind. Sie sollten damit die unterschiedlichen Phänotypen beschreiben.

Während in Tumoren von mit LAN-1 Zellen infizierten Mäusen lediglich Funktionsgruppen identifiziert wurden, die auf erhöhten RNA-Metabolismus hindeuteten, konnten in IMR-32 induzierten Neuroblastomen zusätzlich eine Reihe von Zellstruktur und -organisation definierenden Klassen ermittelt werden. Dies war insofern von Relevanz, als LAN-1 Zellen kleine Mikrometastasen ausbilden, während IMR-32 Neuroblastome größere multizelluläre Metastasen entwickelten. Eine verstärkte Organisation von Zellverbänden, wie sie beispielsweise die GO-Klassen *multicellular organismal development* oder *anatomical structure development* beschreiben, könnte hier in direktem Zusammenhang mit den zellverbandsähnlichen Metastasen stehen.

¹<http://acgt.cs.tau.ac.il/>

4.2. NEUROBLASTOME

Obwohl auch in LS Neuroblastomen Tumorwachstum stattfindet und somit eine verstärkte Transkription zu erwarten ist, konnten hier keine speziellen Funktionsanreicherungen identifiziert werden. Eine mögliche Erklärung für die verstärkte Ausprägung von transkriptionsverwandten Prozessen in IMR-32 und LAN-1 Neuroblastomen könnte daher die dort stattfindende Metastasierung sein. Beschrieben durch GO-Klassen wie *transcription*, *transcription factor activity* oder *RNA metabolic process*, könnte die Ausbildung von Metastasen einen erhöhten RNA-Metabolismus im Vergleich zum ausschließlichen Tumorwachstum in LS Neuroblastomen erfordern.

Dem in IMR-32 identifizierten *secretory pathway* (GO:0045045) gehören Gene an, die in der Exozytose (z.B. *CAPS1*, *ANK1*, *Secernin 1*, *CCR3*), dem Transport zwischen Endoplasmatischem Retikulum und Golgiapparat (z.B. *SCAMP3*, *AP3B1*, *STX16*) oder im Golgivesikel vermitteltem Transport (z.B. *Beta-COP*, *SNAP-alpha*) beteiligt sind. Eine Regulation von Genen dieser Klasse könnte eine höhere Suszeptibilität der Zellen gegenüber Chemokinen zur Folge haben. Sie könnte allerdings auch erhöhte Sekretionsaktivitäten darstellen, die, wie später eingehender behandelt, eine Sezernierung von DKK1 fördern könnten.

Genotypisierung Zur Ergänzung der Expressionsanalytik wurden genomische Aberrationen in den Neuroblastomen mit SNP-Arrays bestimmt. Diese erlaubten gleichzeitig eine Genotypisierung der Samples.

Die verwendeten Zelllinien sind seit vielen Jahren etablierte Linien. Die Genotypisierung zeigte deutliche genetische Unterschiede zwischen den drei Zelllinien. Die Anzahl der Variationen, die zwischen 31% und 34% lagen, erschwerten die Identifizierung von spezifisch auslösenden oder tumorfördernden Faktoren. Die IMR-32 Zelllinie stammt aus dem Jahre 1970 von einem 13 Monate alten, kaukasischen Jungen [163]. Der ethnische Hintergrund der Spender der andern beiden Zelllinien kann der Literatur nicht entnommen werden. Die LS-Zellen wurden in Deutschland gespendet [136], die LAN-1 Zellen in den Vereinigten Staaten [143]. Mit hoher Wahrscheinlichkeit handelt es sich also um Europäer oder Afroamerikaner. Die hohe genetische Variabilität wird daher dem ethnischen Hintergrund zuzuschreiben sein.

Aberrationen Chromosomale Aberrationen sind in Neuroblastomen weit verbreitet [30] und werden oft als Indikatoren für Metastasierungsverhalten angesehen [17]. Gains werden dabei meist mit moderatem Krankheitsbild und spontanem Rückgang der Erkrankung in Verbindung gebracht, während Losses häufig als Zeichen eines metastasierenden Verhaltens

gedeutet werden [105, 111]. Andererseits werden einige Gains wie von *MYCN* in 2p24, z.B. in LAN-1, in direkte Korrelation mit Metastasierungen gesetzt [17]. Deletionen oder Losses in Chromosom 1p, 11q, und 16q sowie Gains in 1p, 1q, 8q, Chromosom 12 und 17q wurden für IMR-32 bereits beschrieben [36, 111]. Weiterhin sind Amplifikationen des *MYCN* Onkogens in 2p24 und Losses in 14q in Neuroblastomen wie LAN-1 bekannt. Gains in 1q, Chromosom 4 und 12p sowie Losses in 3p und Chromosom 15 wurden für LS beschrieben [36].

Die Aberrationen der IMR-32 Zelllinie wiesen gute Übereinstimmung mit der Literatur auf. Die beschriebenen Verluste in 1p, 11q14-q22, 16q23-q24 sowie die Gains in 1q25, 1q32-q44, 17q und Chromosom 12 konnten auch hier beobachtet werden. Das beobachtete Profil konnte in einer biologisch unabhängigen Probe bestätigt werden. Da die Zellen für die Validation einige Monate später neu kultiviert wurden, konnte von einem stabilen Aberrationsprofil ausgegangen werden. Diese Aberrationen wurden in den IMR-32 induzierten Tumoren hingegen nicht beobachtet. Eine Übereinstimmung mit der Literatur lag bei den Tumoren nur für den Gain in 16q vor. Eine Deutung dieser Resultate folgt auf Seite 128.

Für LS ist die Deletion in 3p bereits beschrieben. Eine Beschreibung der Deletionen in 9p24-22 und 13q in der Literatur lag nicht vor. Eine deutliche Amplifikation des *MYCN* in 2p24 konnte auch im Rahmen dieser Arbeit bei LAN-1 Zellen beobachtet werden, nicht jedoch in IMR-32 oder LS induzierten Tumoren. Weiterhin zeigte sich eine Deletion des *TP53* Genes in LAN-1. Zwar ist in der Literatur eine Inaktivierung des p53-Pathways in Neuroblastomen [14, 86] beschrieben, eine Deletion des Locus jedoch nicht.

Insgesamt konnte eine deutliche Heterogenität im Vergleich der Zelllinien untereinander einerseits, und im Vergleich mit der Literatur andererseits beobachtet werden. LS und LAN-1 Samples wiesen im Gegensatz zu IMR-32 meist Gains oder Losses von größeren Regionen, oft sogar ganzen Chromosomenarmen, auf. Auffällig in diesem Sinne war auch der Loss von 9p21 in IMR-32 Neuroblastomen inmitten eines flankierenden Gains. Aus diesen Beobachtungen konnte auf eine deutlich reduzierte genomische Stabilität in IMR-32 induzierten Neuroblastomen geschlossen werden. Desweiteren besteht hier der Verdacht, daß nach der Transfektion der IMR-32 Zellen eine Verschiebung im Gleichgewicht der Subpopulationen stattgefunden hat.

Innerhalb der einzelnen Zelllinien zeigte sich eine recht hohe Homogenität. Die in den Zellen erfolgenden Abläufe dürften sich innerhalb der Zelllinien also kaum unterscheiden. Sie führen zum gleichen Phänotyp.

4.2. NEUROBLASTOME

Korrelation Insgesamt konnte im Rahmen dieser Arbeit eine gute Korrelation der genomischen Resultate mit den Genexpressionen aufgezeigt werden. Die am stärksten regulierten Gene wiesen zusätzlich sehr gute Korrelationsfaktoren auf. Sofern keine Repression durch Transkriptionsfaktoren stattfindet, kann vermutlich davon ausgegangen werden, daß Gains und Amplifikationen im direktem Zusammenhang mit höheren Expressionen stehen. Inwieweit eine Zelle auf den Verlust einer einzelnen Kopie reagiert ist noch nicht ganz klar. Komplette Deletionen jedoch haben zwangsläufig einen Verlust der Funktion der betroffenen Gene zur Folge.

Mittels der genomweiten Korrelationsstrategie konnte erstmalig deutlich gezeigt werden, daß Veränderungen in der genomischen Kopienzahl nicht nur punktuelle Einzeleffekte haben. Die Veränderungen scheinen sich direkt im großem Maßstab auf die Expression aller betroffenen Gene auszuwirken. Insbesondere Amplifikationen von *MYCN* und *DKK1* konnten auf Expressionsebene als extrem hohe Expression wahrgenommen werden. Bei normaler Kopienzahl wiesen die Gene kaum eine Expression auf. Es kann von einer normalen Kopienzahl somit nicht auf ein definiertes Expressionsniveau geschlossen werden. Umgekehrt scheinen Rückschlüsse von Aberrationen auf hohe oder niedrige Expressionswerte möglich zu sein. Die Integration von genomischen Daten und Expressionsanalytik kann daher als vielversprechende Methode zur weitergehenden Analyse und Interpretation angesehen werden.

MAPK-TF-Arrays Drei Transkriptionsfaktoren, namentlich MEF2, c-MYC und p53, wurden in allen drei Zelllinien als aktiviert erkannt. Lediglich in den LAN-1 Zellen konnte keine Aktivität des p53 Proteins erkannt werden. Dieses Ergebnis ließ sich mit der zuvor identifizierten Deletion des Locus und der fehlenden Expression sehr gut in Einklang bringen. Interessanterweise wurde bei der Analyse der p53 Signatur festgestellt, daß einige durch p53 aktivierte Gene wie *SLC38A2*, *CCNG1* oder *EEF1A1* in LAN-1 Tumoren dennoch hohe Expressionswerte aufwiesen. Für diese Gene stellt p53 demnach keinen begrenzenden Faktor für die Regulation dar. Bei der Aktivierung kann p53 substituiert werden.

Als Transkriptionsfaktor ist c-MYC an einer Vielzahl von Zellprozessen beteiligt. Diese umfassen über eine Regulation von MAPK-Signaltransduktionsweg und Transkription das Zellwachstum, Differenzierung oder auch Apoptose. Die gemessene Aktivität von c-Myc bewegte sich in allen Neuroblastomen auf einem äußerst niedrigen Niveau. Der Aktivität von c-Myc in den Neuroblastomen ist daher schwer einer speziellen Funktion zuzuordnen.

Die *myocyte enhancer factor-2* MEF2 Proteine umfassen eine Familie von Transkriptionsfaktoren, die Calcium-abhängig die Zelldifferenzierung und den MAPK-Pathway kontrollieren sowie Stressreaktionen vermitteln [106]. Der MAPK-TF-Chip erfasst alle MEF2 Varianten. Eine Einschränkung der Aktivität auf eine Variante war somit nicht möglich. Andererseits sind die funktionellen Spezialisierungen der Familienmitglieder noch nicht bekannt. In LS konnte eine mehr als 3fach stärkere MEF2 Aktivität gemessen werden. Da es sich hierbei um die am wenigsten aggressive der drei Neuroblastomzelllinien handelt, schien MEF2 eine putativ wichtige Rolle bei der Differenzierung der Tumorzellen und Repression von Metastasen zu spielen.

Die GenMapp-Analyse, siehe Abbildungen 3.12 und 3.13, schien im Vergleich zu LS Neuroblastomen auf verstärkte pro-proliferative und anti-apoptotische Reaktionen in IMR-32 Neuroblastomen und reduzierte Kontrollen der Proliferation, Differenzierung und Apoptose in LAN-1 Neuroblastomen hinzudeuten. Insbesondere die niedrigen Expressionen der Caspasen in LAN-1 Neuroblastomen ließen eine deutlich reduzierte Apoptose erahnen. Diese Beobachtungen könnten potentiell im Zusammenhang mit der mangelnden p53 Aktivität in LAN-1 induzierten Neuroblastomen und der niedrigen Aktivität von MEF2 Faktoren in IMR-32 und LAN-1 Neuroblastomen stehen.

Shift in Subpopulationen von IMR-32 Interessanterweise wiesen einige der beschriebenen Regionen in den Tumoren eine gegenläufige Kopienänderung auf. Die Verluste der Banden 1q25, 5q11-q12, 9p21 und 17q21-q23 müssten demnach erst nach der Transfektion und Tumorprogression aufgetreten sein. Ebenso dürften die Gains in 9q, 10q und 16q im Verlauf der Tumorprogression aufgetreten sein. Sollte es sich bei der 1p36 Bande der IMR-32 Zelllinie um einen Verlust sämtlicher Kopien handeln, bliebe vorerst ungeklärt, wie ein späterer Hinzugewinn dieser Region möglich war. Wie die Korrelationsergebnisse zeigten, wiesen die in den IMR-32 Neuroblastomen beobachteten aberranten Regionen konforme Expressionsmuster der lokalisierten Gene auf. Die Expressionsdaten bestätigten die genomischen Daten eindeutig.

Es musste also davon ausgegangen werden, daß lediglich ein partieller Kopienverlust vorlag oder die Zelllinie aus zwei Subpopulationen bestand. Der äußerst niedrige Signalwert der 1p36 Bande sprach deutlich für eine vollständige Deletion und damit gegen die erste Hypothese. Tatsächlich sind für die IMR-32 Zelllinie zwei morphologisch unterschiedliche Subpopulationen beschrieben [163]. Die Zelllinie wird dabei in Zellkultur zu über

4.2. NEUROBLASTOME

90% von einer der Subpopulationen dominiert. Im vorliegenden Experiment könnten die Subpopulationen in Zellkultur unterschiedliche Aberrationsmuster aufweisen. Bei der SNP Analyse wäre das Profil der dominanten, häufig beobachteten Zellpopulation, zu erkennen. Sollten die Wachstumsbedingungen nach der Transfektion zu einer Verlagerung des Gleichgewichts zwischen den Subpopulationen führen, würde sich damit das beobachtete Aberrationsmuster ändern. Eine experimentelle Bestätigung dieser Hypothese wäre von Bedeutung. Die Dominanz einer zuvor mit unter 10% repräsentierten Subpopulation würde die Komplexität der Tumoriologie unterstreichen. Diese Resultate würden nicht nur den immensen Einfluss der Umgebung auf die Tumorprogression hervorheben. Der Umstand, daß die Subpopulationen sich regenerativ aus der IMR-32 Zelllinie entwickelten [163], würde auch die potentielle Anpassungsfähigkeit von Tumorstammzellen an die jeweiligen Umweltbedingungen widerspiegeln. Die Bestätigung der Aberrationen über eine Verknüpfung mit den Expressionsprofilen erwies sich hier als sehr nützlich.

Metastasierung Metastasierungs Vorgänge sind äußerst komplexe und organisierte Prozesse, die ein Zusammenspiel von vielen Faktoren erfordern. Sie erfolgen in mehreren verzahnten Stufen, während derer Tumorzellen den Primärtumor verlassen, in den Blutkreislauf gelangen, an einen entfernten Ort im Körper gelangen, sich dort einnisten und erneut wachsen. Die erste Stufe, die Invasion, setzt lockere Zell-Zell-Adhäsionen und Motilität der malignen Zellen voraus. Während der nächsten Stufe, der Intravasation, penetrieren die Tumorzellen das Endothel der Blut- oder Lymphgefäße. Um im Kreislaufsystem überhaupt überleben zu können, müssen sie eine erhöhte Resistenz gegen Apoptose gewinnen. Nach der Wanderung beginnt die letzte Stufe, die Einnistung und das Wachstum. Noch nicht endgültig geklärt ist, ob Metastasen sich passiv dort einnisten, wo ihnen eine günstige Wachstums Umgebung geboten wird oder ob sie sich diese Mikrosphäre aktiv selber erschaffen.

Invasion Während IMR-32 Zellen große, mehrzellige Metastasen ausbilden, weist die LAN-1 Zelllinie kleine Mikrometastasen auf. LS bleibt Metastasenfrei. Es musste also geklärt werden, welche Faktoren bei IMR-32 und LAN-1 im Vergleich zu LS zur Metastasierung führen konnten und wie der Metastasenphänotyp beeinflusst wird. Die beiden auffälligsten Unterschiede zwischen den IMR-32 und LAN-1 Zellen bestanden in der ausschließlichen Expression von *CD44* in IMR-32 Zellen und von *MYCN* in LAN-1 Zellen. Die Wirkungsweise dieser beiden Faktoren soll daher näher beschrieben werden.

CD44 ist eine Gruppe transmembraner Glycoproteine, die als Rezeptoren für Hyaluronsäure dienen. In der Regel wird mit CD44 eine Standardform assoziiert, die keine zusätzlichen Exondomänen aufweist. CD44 fördert Zell-Zell- und Zell-Matrix-Adhäsion sowie Migration [175]. Seine zelladhäsive Wirkung ist insbesondere während der Mitose ausgeprägt [120]. Das Glycoprotein spielt eine wichtige Rolle in der Invasion von Neuroblastomen [17]. In etlichen malignen Tumoren wurde seine Expression mit Tumorwachstum und metastatischem Verhalten assoziiert [175, 137, 117]. In Neuroblastomen hingegen soll CD44 mit einem nichtmetastatischen Phänotyp einhergehen [48, 151]. Als wichtiger Bestandteil des Bindegewebes könnte Hyaluronsäure die Adhäsion und Migration invadierender Tumorzellen unterstützen.

Die hier untersuchten und in SCID Mäuse injizierten humanen Neuroblastomzelllinien IMR-32, LAN-1 und LS zeigten in immunhistochemischen Analysen [169] unterschiedliche CD44 Expressionen. IMR-32 wies mit einem Anteil von über 50% an CD44-positiven Zellen eine starke Immunreaktivität mit dem Antikörper auf. Sowohl LAN-1 als auch LS-Zellen waren CD44-negativ. Insofern kann der in mitunter beschriebene Zusammenhang zwischen CD44 und fehlender Metastasierung [17, 48, 151] in Neuroblastomen nicht grundsätzlich bestätigt werden. Unter Umständen hängt der metastatische Charakter von der CD44-Isoform ab.

Auf Expressionsebene wurde im Vergleich der drei Zelllinien eine deutliche Spezifität von *CD44* für IMR-32 festgestellt. Tatsächlich wies das Gen bei IMR-32 ein 445fach höheres Signal (SLR=8,8) im Vergleich zu LAN-1 und LS auf. Dabei handelte es sich nach *SPANXA1* (SLR=9,6) um die zweitstärkste Hochregulation in dieser Gruppe gefolgt von *DKK1*. Es konnte hier von einer fehlenden Expression des Genes in den letztgenannten Linien ausgegangen werden. In den IMR-32 Zellen hingegen könnte CD44 das metastatische Verhalten, wie in Brustkrebs, Pankreas- und Darmkrebs [175, 13] einerseits und die beobachtete erhöhte Adhäsion zwischen den metastatischen Zellen andererseits, beeinflusst haben.

MYCN ist ein im Nukleus lokalisiertes Mitglied der MYC-Familie. Diese ist an der Transkriptionsregulation von etlichen metastasierungsrelevanten Genen beteiligt. MYCN wird eine zentrale Rolle in der Entwicklung von Neuroblastomen und der Metastasierung zugesprochen [17, 18, 12, 11, 154, 67]. In einigen Neuroblastomen, so z.B. in der LAN-1 Zelllinie, ist eine Amplifikation des *MYCN* Genes bereits beschrieben [143]. Hinweise auf Amplifikationen des Gens in IMR-32 und LS Zellen fehlen bislang. In den untersuchten

4.2. NEUROBLASTOME

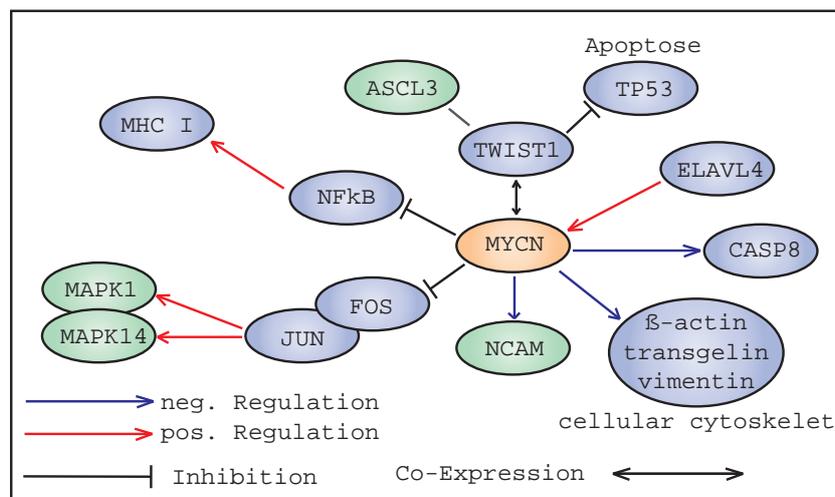


Abbildung 4.1: Interaktionspartner des MYCN Onkogens, formal *v-myc myelocytomatosis viral related oncogene, neuroblastoma derived*. LAN-1 induzierte Tumore wiesen eine Amplifikation des *MYCN* Gens sowie ein erhöhtes Expressionsniveau auf. Im Vergleich zu den LS und IMR-32 Tumoren wiesen LAN-1 Tumore eine niedrige Expression von *CASP8*, *VIM*, *FOS*, *JUN* und der MHC Klasse I Gene sowie eine hohe Expression von *ELAVL4* auf. *TP53* wies einen Verlust mindestens einer Kopie sowie keine Expression auf. Die MAPK-TF-Chips konnten keine Aktivität des p53 Proteins nachweisen. Die hier dargestellten Interaktionen standen somit gut im Einklang mit den Resultaten dieser Arbeit.

Samples konnte ein 120fach höheres Expressionsniveau (SLR=6,9) von *MYCN* in LAN-1 im Vergleich zu IMR-32 und LS beobachtet werden. Weiterhin erhöht zeigte sich das Expressionsniveau von *ELAVL4* im Vergleich zu IMR-32 (SLR=4), einem neuronalen Antigenprotein kodierenden Gen, bei dem eine potentielle Beteiligung in der Prozessierung und dem Splicing von *MYCN* vermutet wird [135].

MYCN kann zelluläre cytoskelettale Proteine wie *beta actin*, *vimentin* oder *transgelin* runterregulieren [17]. Somit wäre ein Einfluss auf die höhere Proliferationsrate der Zellen im Vergleich zu IMR-32 möglich. In den LAN-1 induzierten Tumoren war *vimentin* gegenüber IMR-32/LS 13fach bzw. 4fach (SLR=-3,7/-1,9) niedriger exprimiert. Eine Repression der *CASP8* Protease [17], die in LAN-1 im Vergleich zu IMR-32 und LS Zellen überhaupt nicht exprimiert wurde könnte die Apoptose der Zellen suprimieren.

MYCN kann die Expression von *FOS* und der p50 Untereinheit von *NFκB* runterregulieren. Dies resultiert in einer Inhibition der *NFκB*-Aktivität [171]. In LAN-1 Zellen konnte eine Repression der p50 Untereinheit im Vergleich zu IMR-32 Zellen beobachtet werden, nicht jedoch gegenüber den LS Zellen. Ob diese Regulation *MYCN* zuzuschreiben

ist, bleibt unklar. Es könnte sich alternativ um eine höhere Expression in den IMR-32 Zellen handeln. *FOS* war gegenüber den IMR-32 Zellen 212fach niedriger exprimiert und gegenüber LS Zellen 23fach niedriger. Es kann vermutet werden, daß die Apoptose stimulierende Wirkung dieser Regulationen durch den Verlust des *TP53* Gens in LAN-1 Zellen ausgeglichen wurde.

MYCN werden Regulationen der *phosphoinositide-3-kinase*, PI3K, gesteuerten *VEGF*-Expression in Neuroblastomen [85], eine negative Regulation des *DKK1* Genes, [87] sowie Assoziationen zur Herunterregulation von *NCAM* zugeschrieben [17]. Im Kontrast dazu ist *NCAM* allerdings in LAN-1 Zellen deutlich höher exprimiert als in IMR-32 (SLR=5,6) und LS Zellen (SLR=5,3). *DKK1* hingegen ist, in der Tat, in IMR-32 Zellen sehr hoch und in LAN-1 Zellen kaum exprimiert. In LS Zellen mit mangelnder *MYCN* Expression konnte jedoch ebenfalls keine *DKK1* Expression beobachtet werden. Es war daher nicht nachvollziehbar, ob die niedrige *DKK1* Expression in LAN-1 Zellen aus dem hohen *MYCN* Spiegel resultiert.

Intravasation & Migration In Brusttumorzellen konnte eine CD44-anhängige Chemotaxis entlang Hyaluronangradienten beobachtet werden [164]. Eine Vermittlung in der Fernmetastasierung ist bereits beschrieben [175]. Im SCID-Mausmodell konnten die CD44-positiven Neuroblastome durch die Pulmonararterien transmigrieren [169]. Der Rezeptor könnte hier an einer Adhäsion zwischen Neuroblastom- und Endothelzellen bzw. auf dem Weg zur Einnistung und/oder der Auffindung einer geeigneten Wachstums Umgebung beteiligt sein.

MYCN unterdrückt die Expression von *major histocompatibility complex class I* (MHC I) Genen mittels Bindungsreduzierung von H2TF1 an den *MHC class I gene enhancer*, einer frühen genetischen Aberration zu Beginn der Entwicklung des metastatischen Phänotyps [17, 94]. Suppressionen bzw. Herunterregulationen der MHC I Gene konnten anhand der GenMapp-Analyse, in der Tat, in LAN-1 Neuroblastomen beobachtet werden, siehe auch Abbildung 3.11. Das Molekül könnte damit eine entscheidende Rolle bei der Maskierung der Metastasen gegenüber dem Immunsystem spielen.

Einnistung *DKK1* kodiert für ein sezerniertes Protein, das in der embryonalen Entwicklung beteiligt ist. Das Protein kann den WNT Signaltransduktionsweg inhibieren [140, 126]. In dieser Funktion wird ihm Schlüsselrolle während der Osteoblastendifferenzierung [127] zugeschrieben. Indem *DKK1* den WNT Pathway inhibiert und die Expression von *OPG* herabsetzt, wird das Expressionsverhältnis zwischen *OPG* und *RANKL* verschoben [126].

4.2. NEUROBLASTOME

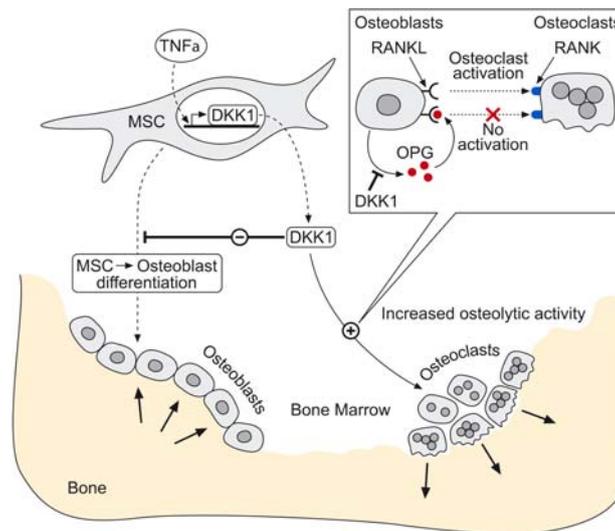


Abbildung 4.2: DKK1 wirkt im Knochen sowohl anabolisch als auch katabolisch. Es inhibiert die Osteoblastenbildung, setzt die Expression von *OPG* herab und führt damit zu reduziertem Knochenwachstum. Im Gegenzug erhöht DKK1 die Osteoklastenaktivität und führt damit zu verstärkter Knochenresorption. (Nach Pinzone et al., 2008 [126])

Infolge wird die Osteoblastendifferenzierung und mit ihr die Knochenbildung herabgesetzt. Im Gegenzug wird die Osteoklastenbildung und damit die Knochenresorption beschleunigt. Oxidativer Stress scheint eine regulative Wirkung auf *DKK1* in Multiplen Myelomen zu haben [46].

DKK1 soll somit eine wichtige Rolle bei der Entstehung osteolytischer Läsionen [52, 95, 128] spielen. In Multiplen Myelomen wurde gezeigt, daß die Serumkonzentrationen von *DKK1* mit dem Ausmaß der lytischen Knochenläsionen korrelieren [81]. Weiterhin wird das Protein mit der Bildung von Knochenmetastasen in Prostatakarzinomen [66], Brusttumoren [174, 32] und malignen Melanomen [89] assoziiert. Präklinische Studien haben gezeigt, daß eine Neutralisierung von *DKK1* und/oder eine Propagierung des Wnt/ β -Catenin Pathways die Behandlung von Knochenerkrankungen verbessern könnten [126].

In der vorliegenden Arbeit konnte eine sehr hohe Expression von *DKK1* in IMR-32 induzierten Tumoren beobachtet werden. In LAN-1 und LS induzierten Tumoren hingegen war eine Expression nicht nachweisbar. Der negative Einfluss von *DKK1* auf die Osteoblastendifferenzierung könnte ein Schlüsselereignis bei der Einnistung der Metastasen darstellen. Von den Metastasen im Knochenmark sekretiert, könnte *DKK1* aktiv eine lokal verstärkte Auflösung des Knochens bewirken und den Metastasen so eine Wachstumsnische schaffen.

4.3 Zusammenfassung

Mammakarzinome stellen in der westlichen Welt die häufigste Krebserkrankung bei Frauen dar. Im Rahmen dieser Arbeit wurden in einem Mausmodell 80 Brusttumore auf Expressionsebene untersucht. Davon wurden 56 Tumore gleichzeitig auf genomischer Ebene analysiert. Die Tumore wurden hinsichtlich Determinanten der Differenzierung analysiert, insbesondere im Hinblick auf den Effekt einer Mutation im *Trp53* Gen.

Im vorliegenden Mausmodell der Brusttumore konnte ein Zusammenhang zwischen dem Differenzierungsgrad und einer Amplifikation des *Met*-Locus hergestellt werden. G3 und G4 Tumore wiesen dabei eine deutlich höhere Expression der im Locus lokalisierten Gene auf. Ein direkter Einfluss des mutierten *Trp53* Genes war jedoch nicht nachvollziehbar. Der Mechanismus nach dem das *mutP53* die Tumorprogression beeinflusst ist nach wie vor unbekannt.

Neuroblastome sind in Deutschland nach der akuten lymphatischen Leukämie eine der häufigsten Erkrankungen im Kindesalter. Im Rahmen dieser Arbeit wurden drei Neuroblastomzelllinien auf metastasierungsrelevante Gene hin untersucht. Es konnte ein prominenter Einfluss von *MYCN* und *CD44* auf das metastatische Verhalten beschrieben werden. *MYCN* war in LAN-1 induzierten Tumoren deutlich amplifiziert und charakterisierte verstärktes mikrometastatisches Verhalten. Demgegenüber charakterisierte *CD44* Neuroblastome der IMR-32 Zelllinie mit wenigen mehrzelligen Metastasen die über verstärkte Zelladhäsion und die Fähigkeit zur Transmigration durch die Pulmonararterien verfügten. Potentiell findet hier ein Shift in den IMR-32 Subpopulationen nach der Transfektion statt. Desweiteren konnte die Hypothese formuliert werden, daß DKK1 bei der aktiven Schaffung einer Wachstumsnische für die Metastasen beteiligt sein könnte.

Modellübergreifend sollte das Potential einer Integration von genomischen und transkriptomischen Daten zur Beschreibung von tumorgenetischen Einflüssen evaluiert werden. Zur Integration von genomischen und transkriptomischen Daten wurden zwei unterschiedliche Methoden eingesetzt. Die direkte Korrelation von CNV Daten mit Expressionssignalen pro Gen mittels linearer Regressionen ermöglichte eine Eingrenzung der Ergebnisse auf eine geringere Anzahl signifikanter Determinanten. Die Methode der Korrelation über das gesamte Genom hob den Einfluss von genomischen Aberrationen auf das Genexpressionsniveau hervor. Die Ergebnisse zeigten deutlich, daß Änderungen der chromosomalen Kopienzahl sich direkt funktionell auf die Transkription auswirken. Insbesondere die Bestäti-

4.3. ZUSAMMENFASSUNG

gung der beobachteten Aberrationen in den IMR-32 Neuroblastomen auf Expressionsebene ermöglichte eine Hypothesenformulierung, dergemäß ein Shift in den Subpopulationen nach der Transfektion erfolgt sein musste.

Die beiden untersuchten Modelle waren deutlich verschieden und inhaltlich schwer zu vergleichen. Nach wie vor bleibt die Verknüpfung der Daten sehr komplex. Modellübergreifend können die Strategien zur Integration von genomischen und transkriptomischen Daten dennoch als vielversprechende Analytik eingestuft werden.

Einheiten

µg	Mikrogramm.
µl	Mikroliter.
µmol	Mikromol.
h	Stunden.
mg	Milligramm.
min	Minuten.
ml	Milliliter.
ng	Nanogramm.
nm	Nanometer.
rpm	Rotationen pro Minute.
sec	Sekunden.
U	Units (Enzymeinheiten).

Abkürzungen

array-CGH array Comparative Genomic Hybridization.

CN Copy Number, Kopienzahl.

CNV Copy-Number-Variation, Änderung in der genomischen Kopienzahl.

DNA Desoxyribonukleinsäure.

FC Fold Change. Quotient zweier Expressionssignale.

HPLC High performance liquid chromatography.

IMR-32 Neuroblastomzelllinie, die von Tumilowicz et al. etabliert wurde.

IVT In Vitro Transkription.

LAN-1 Neuroblastomzelllinie, die von Seeger et al. etabliert wurde.

LS Neuroblastomzelllinie, die von Rudolph et al. etabliert wurde.

MM Mis-Match, Terminologie Microarrays, Affymetrix..

mutP53 Mutiertes P53 Gen bzw. Protein.

Abkürzungen

p53 P53 Protein.

PM Perfect Match, Terminologie Microarrays, Affymetrix.

Probe Laboranalytik: Gewebe-/DNA-/RNA-Probe. Bioinformatik: 25basenpaarlange Oligonucleotide, Bestandteil eines Probesets auf Affymetrix Arrays.

Probeset Bioinformatik: Satz aus mehreren 25basenpaarlangen Oligonucleotid-Proben.

RNA Ribonukleinsäure.

Sample Bioinformatik: Gewebe-/DNA-/RNA-Probe nach Hybridisierung und Scan.

SAPE Streptavidin-Phycoerythrin.

SCID Severe combined immunodeficiency;.

SLR/logRatio Verhältnis zwischen Testsignal und Referenzsignal (logarithmisch, Basis 2).

SNP Single Nucleotide Polymorphism.

SV40 T-Antigen Großes T-Antigen des simian Virus 40.

TP53 humanes P53 Gen.

Trp53 murines P53 Gen.

WAP-T BALB/c Mausmodell zur Untersuchung des mutP53 Einflusses.

ZNS Zentrales Nervensystem.

Literatur

- [1] L. Adams, R. Geary, J. Li, A. Rossini, and S. Schwartz. Expression profiling identifies smooth muscle cell diversity within human intima and plaque fibrous cap: loss of RGS5 distinguishes the cap. *Arterioscler Thromb Vasc Biol*, 26(2):319–25, Feb 2006. PMID: 16293795.
- [2] Affymetrix. *Affymetrix Expression Console Software Version 1.0*. Affymetrix, Inc., Santa Clara, USA.
- [3] Affymetrix. *Affymetrix Microarray Suite User's Guide version 5.0*. Affymetrix, Inc., Santa Clara, USA, 2001.
- [4] Affymetrix. *GeneChip[®] Expression Analysis Technical Manual*. Affymetrix, Inc., Santa Clara, USA, 2004.
- [5] Affymetrix. *BRLMM: An Improved Genotype Calling Method for the Mapping 500K Array Set*. Affymetrix, Inc., Santa Clara, USA, 2006.
- [6] Affymetrix. *Genotyping Console 2.1 User Manual*. Affymetrix, Inc., Santa Clara, USA, 2007.
- [7] Affymetrix. *SNP 6.0 Probe Level Access Tool User Manual*. Affymetrix, Inc., Santa Clara, USA, 2007.
- [8] Affymetrix. *Statistical algorithms description document*. Affymetrix, Inc., Santa Clara, USA, 2007.
- [9] Affymetrix. *Statistical algorithms reference guide*. Affymetrix, Inc., Santa Clara, USA, 2007.
- [10] A. J. Aguirre, C. Brennan, G. Bailey, R. Sinha, B. Feng, C. Leo, Y. Zhang, J. Zhang, J. D. Gans, N. Bardeesy, C. Cauwels, C. Cordon-Cardo, M. S. Redston, R. A. DePinho, and L. Chin. High-resolution characterization of the pancreatic adenocarcinoma genome. *Proc Natl Acad Sci U S A*, 101(24):9067–9072, Jun 2004. PMID: 15199222.
- [11] M. Alaminos, J. Mora, N.-K. V. Cheung, A. Smith, J. Qin, L. Chen, and W. L. Gerald. Genome-wide analysis of gene expression associated with MYCN in human neuroblastoma. *Cancer Res*, 63(15):4538–4546, Aug 2003. PMID: 12907629.
- [12] O. Altungoz, N. Aygun, S. Tumer, E. Ozer, N. Olgun, and M. Sakizli. Correlation of modified Shimada classification with MYCN and 1p36 status detected by fluorescence in situ hybridization in neuroblastoma. *Cancer Genet Cytogenet*, 172(2):113–119, Jan 2007. PMID: 17213019.
- [13] C. S. Alves, M. M. Burdick, S. N. Thomas, P. Pawar, and K. Konstantopoulos. The dual role of CD44 as a functional P-selectin ligand and fibrin receptor in colon carcinoma cell adhesion. *Am J Physiol Cell Physiol*, 294(4):C907–C916, Apr 2008. PMID: 18234849.
- [14] K. Ando, T. Ozaki, H. Yamamoto, K. Furuya, M. Hosoda, S. Hayashi, M. Fukuzawa, and A. Nakagawara. Polo-like kinase 1 (Plk1) inhibits p53 function by physical interaction and phosphorylation. *J Biol Chem*, 279(24):25549–25561, Jun 2004. PMID: 15024021.
- [15] M. Antrand. Contrast Normalization of Oligonucleotide Arrays. *J Comput Biol*, 10(1):95–102, February 2003.

- [16] Y. Aoka, F. Johnson, K. Penta, K. K. Hirata, C. Hidai, R. Schatzman, J. Varner, and T. Quertermous. The embryonic angiogenic factor Dell1 accelerates tumor growth by enhancing vascular formation. *Microvasc Res*, 64(1):148–61, Jul 2002. PMID: 12074641.
- [17] T. Ara and Y. A. DeClerck. Mechanisms of invasion and metastasis in human neuroblastoma. *Cancer Metastasis Rev*, 25(4):645–657, Dec 2006. PMID: 17160711.
- [18] S. Asgharzadeh, R. Pique-Regi, R. Sposto, H. Wang, Y. Yang, H. Shimada, K. Matthay, J. Buckley, A. Ortega, and R. C. Seeger. Prognostic significance of gene expression profiles of metastatic neuroblastomas lacking MYCN gene amplification. *J Natl Cancer Inst*, 98(17):1193–1203, Sep 2006. PMID: 16954472.
- [19] T. Barrett, D. Troup, S. Wilhite, P. Ledoux, D. Rudnev, C. Evangelista, I. Kim, A. Soboleva, M. Tomashevsky, and R. Edgar. NCBI GEO: mining tens of millions of expression profiles—database and tools update. *Nucleic Acids Res*, 35:760–765, January 2007. PMID: 17099226.
- [20] Y. Benjamini and Y. Hochberg. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J Royal Stat Soc Series B*, 57:289–300, 1995.
- [21] M. Berger, G. Bergers, B. Arnold, G. Hämmerling, and R. Ganss. Regulator of G-protein signaling-5 induction in pericytes coincides with active vessel remodeling during neovascularization. *Blood*, 105(3):1094–101, Feb 2005. PMID: 15459006.
- [22] R. Bernards and R. A. Weinberg. A progression puzzle. *Nature*, 418(6900):823, Aug 2002. PMID: 12192390.
- [23] C. Birchmeier, W. Birchmeier, E. Gherardi, and G. F. V. Woude. Met, metastasis, motility and more. *Nat Rev Mol Cell Biol*, 4(12):915–925, Dec 2003.
- [24] C. Boccaccio and P. M. Comoglio. Invasive growth: a MET-driven genetic programme for cancer and stem cells. *Nat Rev Cancer*, 6(8):637–645, Aug 2006.
- [25] B. Bolstad, R. Irizarry, M. Astrand, and T. Speed. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics*, 19(2):185–193, 2003.
- [26] I. Borg and P. Groenen. *Modern Multidimensional Scaling: Theory and Applications*. Springer-Verlag, Berlin, 1997.
- [27] G. C. Bosma, R. P. Custer, and M. J. Bosma. A severe combined immunodeficiency mutation in the mouse. *Nature*, 301(5900):527–530, Feb 1983. PMID: 6823332.
- [28] M. J. Bosma and A. M. Carroll. The SCID mouse mutant: definition, characterization, and potential uses. *Annu Rev Immunol*, 9:323–350, 1991. PMID: 1910681.
- [29] I. Bossis and C. Stratakis. Minireview: PRKAR1A: normal and abnormal functions. *Endocrinology*, 145(12):5452–8, Dec 2004. PMID: 15331577.
- [30] N. Bown. Neuroblastoma tumour genetics: clinical and biological aspects. *J Clin Pathol*, 54(12):897–910, Dec 2001. PMID: 11729208.
- [31] S. Brazer, B. Singh, X. Liu, W. Swaim, and I. Ambudkar. Caveolin-1 contributes to assembly of store-operated Ca²⁺ influx channels by regulating plasma membrane localization of TRPC1. *J Biol Chem*, 278(29):27208–15, Jul 2003. PMID: 12732636.
- [32] G. Bu, W. Lu, C.-C. Liu, K. Selander, T. Yoneda, C. Hall, E. T. Keller, and Y. Li. Breast cancer-derived Dickkopf1 inhibits osteoblast differentiation and osteoprotegerin expression: implication for breast cancer osteolytic bone metastases. *Int J Cancer*, 123(5):1034–1042, Sep 2008. PMID: 18546262.
- [33] P. Cabart, H. K. Chew, and S. Murphy. BRCA1 cooperates with NUFIP and P-TEFb to activate transcription by RNA polymerase II. *Oncogene*, 23(31):5316–5329, Jul 2004. PMID: 15107825.

LITERATUR

- [34] M. D. Cameron, E. E. Schmidt, N. Kerkvliet, K. V. Nadkarni, V. L. Morris, A. C. Groom, A. F. Chambers, and I. C. MacDonald. Temporal progression of metastasis in lung: cell survival, dormancy, and location dependence of metastatic inefficiency. *Cancer Res*, 60(9):2541–2546, May 2000. PMID: 10811137.
- [35] D. Campbell, L. Lange, T. Skelly, J. Lieberman, P. Levitt, and P. Sullivan. Association of RGS2 and RGS5 variants with schizophrenia symptom severity. *Schizophr Res*, 101(1-3):67–75, Apr 2008. PMID: 18262772.
- [36] J. Carr, N. P. Bown, M. C. Case, A. G. Hall, J. Lunec, and D. A. Tweddle. High-resolution analysis of allelic imbalance in neuroblastoma cell lines by single nucleotide polymorphism arrays. *Cancer Genet Cytogenet*, 172(2):127–138, Jan 2007. PMID: 17213021.
- [37] L. Cervenakova, O. Yakovleva, and C. McKenzie. Protease-resistant prion protein in lymphoreticular tumors of variant Creutzfeldt-Jakob disease mice. *Emerg Infect Dis*, 12(3):511–3, Mar 2006. PMID: 16704797.
- [38] Q. Chen, M. C. DeFrances, and R. Zarnegar. Induction of met proto-oncogene (hepatocyte growth factor receptor) expression during human monocyte-macrophage differentiation. *Cell Growth Differ*, 7(6):821–832, Jun 1996. PMID: 8780895.
- [39] S. Chen, S. Lin, K. Yeh, S. Kuo, W. Chan, Y. Chu, and J. Chang. Mutational, epigenetic and expressional analyses of caveolin-1 gene in breast cancers. *Int J Mol Med*, 14(4):577–82, Oct 2004. PMID: 15375584.
- [40] W. Chen, R. Possemato, K. T. Campbell, C. A. Plattner, D. C. Pallas, and W. C. Hahn. Identification of specific PP2A complexes involved in human cell transformation. *Cancer Cell*, 5(2):127–136, Feb 2004. PMID: 14998489.
- [41] S. F. Chin, A. E. Teschendorff, J. C. Marioni, Y. Wang, N. L. Barbosa-Morais, N. P. Thorne, J. L. Costa, S. E. Pinder, M. A. van de Wiel, A. R. Green, I. O. Ellis, P. L. Porter, S. Tavaré, J. D. Brenton, B. Ylstra, and C. Caldas. High-resolution aCGH and expression profiling identifies a novel genomic subtype of ER negative breast cancer. *Genome Biol*, 8(10):R215, 2007. PMID: 17925008.
- [42] S. Chowdhury, A. Gemin, and G. Singh. High activity of mitochondrial glycerophosphate dehydrogenase and glycerophosphate-dependent ROS production in prostate cancer cell lines. *Biochem Biophys Res Commun*, 333(4):1139–45, Aug 2005. PMID: 15967408.
- [43] S. Chowdhury, S. Raha, M. Tarnopolsky, and G. Singh. Increased expression of mitochondrial glycerophosphate dehydrogenase and antioxidant enzymes in prostate cancer cell lines/cancer. *Free Radic Res*, 41(10):1116–24, Oct 2007. PMID: 17886033.
- [44] E. Ciani, S. Severi, A. Contestabile, R. Bartesaghi, and A. Contestabile. Nitric oxide negatively regulates proliferation and promotes neuronal differentiation through N-Myc downregulation. *J Cell Sci*, 117(Pt 20):4727–4737, Sep 2004. PMID: 15331636.
- [45] M. Cicek, R. Samant, M. Kinter, D. Welch, and G. Casey. Identification of metastasis-associated proteins through protein analysis of metastatic MDA-MB-435 and metastasis-suppressed BRMS1 transfected-MDA-MB-435 cells. *Clin Exp Metastasis*, 21(2):149–57, 2004. PMID: 15168732.
- [46] S. Colla, F. Zhan, W. Xiong, X. Wu, H. Xu, O. Stephens, S. Yaccoby, J. Epstein, B. Barlogie, and J. D. Shaughnessy. The oxidative stress response regulates DKK1 expression through the JNK signaling cascade in multiple myeloma plasma cells. *Blood*, 109(10):4470–4477, May 2007. PMID: 17255354.
- [47] M. Colucci, F. Moleres, Z. Xie, A. Ray-Chaudhury, S. Gutti, C. Butefisch, L. Cervenakova, W. Wang, L. Goldfarb, Q. Kong, B. Ghetti, S. Chen, and P. Gambetti. Gerstmann-Sträussler-Scheinker: a new phenotype with 'curly' PrP deposits. *J Neuropathol Exp Neurol*, 65(7):642–51, Jul 2006. PMID: 16825951.

- [48] V. Combaret, N. Gross, C. Lasset, D. Frappaz, C. Beretta-Brogna, T. Philip, D. Beck, and M. C. Favrot. Clinical relevance of CD44 cell surface expression and MYCN gene amplification in neuroblastoma. *Eur J Cancer*, 33(12):2101–2105, Oct 1997. PMID: 9516862.
- [49] A. Coquelle, E. Pipiras, F. Toledo, G. Buttin, and M. Debatisse. Expression of fragile sites triggers intrachromosomal mammalian gene amplification and sets boundaries to early amplicons. *Cell*, 89(2):215–225, Apr 1997.
- [50] M. Dean, T. Fojo, and S. Bates. Tumour stem cells and drug resistance. *Nat Rev Cancer*, 5(4):275–284, Apr 2005. PMID: 15803154.
- [51] G. G. V. den Eynden, S. J. V. Laere, I. V. der Auwera, S. D. Merajver, E. A. V. Marck, P. van Dam, P. B. Vermeulen, L. Y. Dirix, and K. L. van Golen. Overexpression of caveolin-1 and -2 in cell lines and in human samples of inflammatory breast cancer. *Breast Cancer Res Treat*, 95(3):219–228, Feb 2006. PMID: 16244790.
- [52] D. Diarra, M. Stolina, K. Polzer, J. Zwerina, M. S. Ominsky, D. Dwyer, A. Korb, J. Smolen, M. Hoffmann, C. Scheinecker, D. van der Heide, R. Landewe, D. Lacey, W. G. Richards, and G. Schett. Dickkopf-1 is a master regulator of joint remodeling. *Nat Med*, 13(2):156–163, Feb 2007. PMID: 17237793.
- [53] S. T. Durant and J. A. Nickoloff. Good timing in the cell cycle for precise DNA repair by BRCA1. *Cell Cycle*, 4(9):1216–1222, Sep 2005. PMID: 16103751.
- [54] R. Edgar, M. Domrachev, and A. Lash. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res*, 30(1):207–210, January 2002.
- [55] M. Fabbro, K. Savage, K. Hobson, A. J. Deans, S. N. Powell, G. A. McArthur, and K. K. Khanna. BRCA1-BARD1 complexes are required for p53Ser-15 phosphorylation and a G1/S arrest following ionizing radiation-induced DNA damage. *J Biol Chem*, 279(30):31251–31258, Jul 2004. PMID: 15159397.
- [56] K. Fang, W. Fu, A. Beardsley, X. Sun, M. Lisanti, and J. Liu. Overexpression of caveolin-1 inhibits endothelial cell proliferation by arresting the cell cycle at G0/G1 phase. *Cell Cycle*, 6(2):199–204, Jan 2007. PMID: 17245131.
- [57] D. Ferraro, S. Corso, E. Fasano, E. Panieri, R. Santangelo, S. Borrello, S. Giordano, G. Pani, and T. Galeotti. Pro-metastatic signaling by c-Met through RAC-1 and reactive oxygen species (ROS). *Oncogene*, 25(26):3689–3698, Jun 2006. PMID: 16462764.
- [58] I. J. Fidler and I. R. Hart. Biological diversity in metastatic neoplasms: origins and implications. *Science*, 217(4564):998–1003, Sep 1982. PMID: 7112116.
- [59] Y. Fujita, S. Maruyama, H. Kogo, S. Matsuo, and T. Fujimoto. Caveolin-1 in mesangial cells suppresses MAP kinase activation and cell proliferation induced by bFGF and PDGF. *Kidney Int*, 66(5):1794, Nov 2004. 804.
- [60] S. Garcia, J.-P. Dalès, E. Charafe-Jauffret, S. Carpentier-Meunier, L. Andrac-Meyer, J. Jacquemier, C. Andonian, M.-N. Lavaut, C. Allasia, P. Bonnier, and C. Charpin. Poor prognosis in breast carcinomas correlates with increased expression of targetable CD146 and c-Met and with proteomic basal-like phenotype. *Hum Pathol*, 38(6):830–841, Jun 2007. PMID: 17316758.
- [61] F. Gavins, J. Dalli, R. Flower, D. Granger, and M. Perretti. Activation of the annexin 1 counter-regulatory circuit affords protection in the mouse brain microcirculation. *FASEB J*, 21(8):1751–8, Feb 2007. PMID: 17317721.

LITERATUR

- [62] R. Gentleman, V. Carey, D. Bates, B. Bolstad, M. Dettling, S. Dudoit, B. Ellis, L. Gautier, Y. Ge, J. Gentry, K. Hornik, T. Hothorn, W. Huber, S. Iacus, R. Irizarry, F. L. C. Li, M. Maechler, A. Ros-sini, G. Sawitzki, C. Smith, G. Smyth, L. Tierney, J. Y. H. Yang, and J. Zhang. Bioconductor: Open software development for computational biology and bioinformatics. *Genome Biology*, 5(10):R80, 2004. PMID: 15461798.
- [63] R. Gosens, G. Stelmack, G. Dueck, K. McNeill, A. Yamasaki, W. Gerthoffer, H. Unruh, A. Gounni, J. Zaagsma, and A. Halayko. Role of caveolin-1 in p42/p44 MAP kinase activation and proliferation of human airway smooth muscle. *Am J Physiol Lung Cell Mol Physiol*, 291(3):L523–34, Sep 2006. PMID: 16617096.
- [64] W. Gosset. The probable error of mean. *Biometrika*, 6(1):1–25, 1908.
- [65] M. A. Groenen, R. J. Dijkhof, and J. J. van der Poel. Characterization of a GlyCAM1-like gene (glycosylation-dependent cell adhesion molecule 1) which is highly and specifically expressed in the lactating bovine mammary gland. *Gene*, 158(2):189–195, Jun 1995. PMID: 7607540.
- [66] C. L. Hall, S. D. Daignault, R. B. Shah, K. J. Pienta, and E. T. Keller. Dickkopf-1 expression increases early in prostate cancer development and decreases during progression from primary tumor to metastasis. *Prostate*, 68(13):1396–1404, Sep 2008. PMID: 18561248.
- [67] L. M. Hansford, W. D. Thomas, J. M. Keating, C. A. Burkhardt, A. E. Peaston, M. D. Norris, M. Haber, P. J. Armati, W. A. Weiss, and G. M. Marshall. Mechanisms of embryonal tumor initiation: distinct roles for MycN expression and MYCN amplification. *Proc Natl Acad Sci U S A*, 101(34):12664–12669, Aug 2004. PMID: 15314226.
- [68] A. Hartemink, D. Gi, o Jaakkola, and R. Young. Maximum likelihood estimation of optimal scaling factors for expression array normalizations. *SPIE Bios*, 2001.
- [69] C. Heinlein, F. Krepulat, J. Löhler, D. Speidel, W. Deppert, and G. V. Tolstonog. Mutant p53(R270H) gain of function phenotype in a mouse model for oncogene-induced mammary carcinogenesis. *Int J Cancer*, 122(8):1701–1709, Apr 2008. PMID: 18092324.
- [70] C. Hirst, E. Ng, L. Azzola, A. Voss, T. Thomas, E. Stanley, and A. Elefanty. Transcriptional profiling of mouse and human ES cells identifies SLAIN1, a novel stem cell gene. *Dev Biol*, 293(1):90–103, May 2006. PMID: 16546155.
- [71] C. Ho, P. Huang, H. Huang, Y. Chen, P. Yang, and S. Hsu. Up-regulated caveolin-1 accentuates the metastasis capability of lung adenocarcinoma by inducing filopodia formation. *Am J Pathol*, 161(5): 1647–56, Nov 2002. PMID: 12414512.
- [72] Z. Hou, J. P. Bailey, A. J. Vomachka, M. Matsuda, J. A. Lockfefer, and N. D. Horseman. Glycosylation-dependent cell adhesion molecule 1 (GlyCAM 1) is induced by prolactin and sup-pressed by progesterone in mammary epithelium. *Endocrinology*, 141(11):4278–4283, Nov 2000. PMID: 11089563.
- [73] Z. Hou, S. Srivastava, M. J. Mistry, M. P. Herbst, J. P. Bailey, and N. D. Horseman. Two tandemly linked interferon-gamma-activated sequence elements in the promoter of glycosylation-dependent cell adhesion molecule 1 gene synergistically respond to prolactin in mouse mammary epithelial cells. *Mol Endocrinol*, 17(10):1910–1920, Oct 2003. PMID: 12869589.
- [74] R. K. Institut. Broschüre Krebs in Deutschland, Brustdrüse der Frau, ICD-10 C50. Broschüre, 2008.
- [75] R. K. Institut. Broschüre Krebs in Deutschland, Krebs bei Kindern. Broschüre, 2008.
- [76] R. A. Irizarry, B. M. Bolstad, F. Collin, L. M. Cope, B. Hobbs, and T. P. Speed. Summaries of Affymetrix GeneChip probe level data. *Nucl Acids Res*, 31(4):e15–, 2003.

- [77] R. A. Irizarry, B. Hobbs, F. Collin, Y. D. Beazer-Barclay, K. J. Antonellis, U. Scherf, and T. P. Speed. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostat*, 4(2):249–264, 2003.
- [78] J. Jacobi, K. Sydow, G. von Degenfeld, Y. Zhang, H. Dayoub, B. Wang, A. Patterson, M. Kimoto, H. Blau, and J. Cooke. Overexpression of dimethylarginine dimethylaminohydrolase reduces tissue asymmetric dimethylarginine levels and enhances angiogenesis. *Circulation*, 111(11):1431–8, Mar 2005. PMID: 15781754.
- [79] T. Johansson, M. Lejonklou, S. Ekeblad, P. Stålberg, and B. Skogseid. Lack of nuclear expression of hairy and enhancer of split-1 (HES1) in pancreatic endocrine tumors. *Horm Metab Res*, 40(5):354–9, May 2008. PMID: 18491256.
- [80] S. Jothy. CD44 and its partners in metastasis. *Clin Exp Metastasis*, 20:195–201, 2003. PMID: .
- [81] M. Kaiser, M. Mieth, P. Liebisch, R. Oberländer, J. Rademacher, C. Jakob, L. Kleeberg, C. Fleissner, E. Braendle, M. Peters, D. Stover, O. Sezer, and U. Heider. Serum concentrations of DKK-1 correlate with the extent of bone disease in patients with multiple myeloma. *Eur J Haematol*, 80(6):490–494, Jun 2008. PMID: 18331598.
- [82] G. Kaloshi, K. Mokhtari, C. Carpentier, S. Taillibert, J. Lejeune, Y. Marie, J. Delattre, R. Godbout, and M. Sanson. FABP7 expression in glioblastomas: relation to prognosis, invasion and EGFR status. *J Neurooncol*, 84(3):245–8, Sep 2007. PMID: 17415524.
- [83] A. Kamal, S. Smith, W. M. De Silva, E. Solito, and C. Corrigan. An annexin 1 (ANXA1)-derived peptide inhibits prototype antigen-driven human T cell Th1 and Th2 responses in vitro. *Clin Exp Allergy*, 31(7):1116–25, Jul 2001. PMID: 11468004.
- [84] E. Kang, T. Cates, D. Harper, T. Chiang, L. Myers, S. Acchiardo, and M. Kimoto. An enzyme hydrolyzing methylated inhibitors of nitric oxide synthase is present in circulating human red blood cells. *Free Radic Res*, 35(6):693–707, Dec 2001. PMID: 11811522.
- [85] J. Kang, P. G. Rychahou, T. A. Ishola, J. M. Mourot, B. M. Evers, and D. H. Chung. N-myc is a novel regulator of PI3K-mediated VEGF expression in neuroblastoma. *Oncogene*, 27(28):3999–4007, Jun 2008. PMID: 18278068.
- [86] N. Keshelava, J. J. Zuo, N. S. Waidyaratne, T. J. Triche, and C. P. Reynolds. p53 mutations and loss of p53 function confer multidrug resistance in neuroblastoma. *Med Pediatr Oncol*, 35(6):563–568, Dec 2000. PMID: 11107118.
- [87] A. Koppen, R. Ait-Aissa, S. Hopman, J. Koster, F. Haneveld, R. Versteeg, and L. J. Valentijn. Dickkopf-1 is down-regulated by MYCN and inhibits neuroblastoma cell proliferation. *Cancer Lett*, 256(2):218–228, Oct 2007. PMID: 17643814.
- [88] F. Krepulat, J. Löhler, C. Heinlein, A. Hermannstädter, G. V. Tolstonog, and W. Deppert. Epigenetic mechanisms affect mutant p53 transgene expression in WAP-mutp53 transgenic mice. *Oncogene*, 24(29):4645–4659, Jul 2005. PMID: 15870706.
- [89] S. Kuphal, S. Lodermeier, F. Bataille, M. Schuierer, B. H. Hoang, and A. K. Bosserhoff. Expression of Dickkopf genes is strongly reduced in malignant melanoma. *Oncogene*, 25(36):5027–5036, Aug 2006. PMID: 16568085.
- [90] D. P. Lane. Cancer. p53, guardian of the genome. *Nature*, 358(6381):15–16, Jul 1992. PMID: 1614522.
- [91] S. Lee, K. Lee, H. Chang, M. Choi, M. Cho, S. Min, H. Lee, Y. Mun, E. Nam, C. Seong, and S. Lee. Prognostic significance of CD44s expression in biliary tract cancers. *Ann Surg Oncol*, 15:1155–1160, 2008. PMID: .

LITERATUR

- [92] U. Lehnigk, U. Zimmermann, C. Woenckhaus, and J. Giebel. Localization of annexins I, II, IV and VII in whole prostate sections from radical prostatectomy patients. *Histol Histopathol*, 20(3):673–80, Jul 2005. PMID: 15944914.
- [93] J. Leiper, M. Nandi, B. Torondel, J. Murray-Rust, M. Malaki, B. O’Hara, S. Rossiter, S. Anthony, M. Madhani, D. Selwood, C. Smith, B. Wojciak-Stothard, A. Rudiger, R. Stidwill, N. McDonald, and P. Vallance. Disruption of methylarginine metabolism impairs vascular homeostasis. *Nat Med*, 13(2):198–203, Feb 2007. PMID: 17273169.
- [94] M. Leonardo, A. Rustgi, A. Schievella, and R. Bernards. Suppression of MHC class I gene expression by N-myc through enhancer inactivation. *EMBO Journal*, 8:3351–3355, 1989.
- [95] Z. G. Li, J. Yang, E. S. Vazquez, D. Rose, F. Vakar-Lopez, P. Mathew, A. Lopez, C. J. Logothetis, S.-H. Lin, and N. M. Navone. Low-density lipoprotein receptor-related protein 5 (LRP5) mediates the prostate cancer-induced formation of new bone. *Oncogene*, 27(5):596–603, Jan 2008. PMID: 17700537.
- [96] Y. Liang, A. Bollen, K. Aldape, and N. Gupta. Nuclear FABP7 immunoreactivity is preferentially expressed in infiltrative glioma and is associated with poor prognosis in EGFR-overexpressing glioblastoma. *BMC Cancer*, 6:97, Apr 2006. PMID: 16623952.
- [97] Y. Liang, M. Diehn, N. Watson, A. Bollen, K. Aldape, M. Nicholas, K. Lamborn, M. Berger, D. Bostein, P. Brown, and M. Israel. Gene expression profiling reveals molecularly and clinically distinct subtypes of glioblastoma multiforme. *Proc Natl Acad Sci U S A*, 102(16):5814–9, Apr 2005. PMID: 15827123.
- [98] H. Liu, B. R. Zeeberg, G. Qu, A. G. Koru, A. Ferrucci, A. Kahn, M. C. Ryan, A. Nuhanovic, P. J. Munson, W. C. Reinhold, D. W. Kane, and J. N. Weinstein. AffyProbeMiner: a web resource for computing or retrieving accurately redefined Affymetrix probe sets. *Bioinformatics*, 23(18):2385–2390, 2007.
- [99] Z. Lu, S. Ghosh, Z. Wang, and T. Hunter. Downregulation of caveolin-1 function by EGF leads to the loss of E-cadherin, increased transcriptional activity of beta-catenin, and enhanced tumor cell invasion. *Cancer Cell*, 4(6):499–515, Dec 2003. PMID: 14706341.
- [100] G. Mantovani, S. Bondioni, A. Lania, M. Rodolfo, E. Peverelli, N. Polentarutti, R. T. Veliz, S. Ferrero, S. Bosari, P. Beck-Peccoz, and A. Spada. High expression of PKA regulatory subunit 1A protein is related to proliferation of human melanoma cells. *Oncogene*, 27(13):1834–43, Mar 2008. PMID: 17906691.
- [101] G. Mantovani, A. Lania, S. Bondioni, E. Peverelli, C. Pedroni, S. Ferrero, C. Pellegrini, L. Vicentini, G. Arnaldi, S. Bosari, P. Beck-Peccoz, and A. Spada. Different expression of protein kinase A (PKA) regulatory subunits in cortisol-secreting adrenocortical tumors: relationship with cell proliferation. *Exp Cell Res*, 314(1):123–30, Jan 2008. PMID: 17904549.
- [102] R. W. Martin, B. J. Orelli, M. Yamazoe, A. J. Minn, S. Takeda, and D. K. Bishop. RAD51 up-regulation bypasses BRCA1 function and is a common feature of BRCA1-deficient breast tumors. *Cancer Res*, 67(20):9658–9665, Oct 2007. PMID: 17942895.
- [103] S. Massone, F. Argellati, M. Passalacqua, A. Armirotti, L. Melone, C. d’Abramo, U. Marinari, C. Domenicotti, M. Pronzato, and R. Ricciarelli. Downregulation of myosin II-B by siRNA alters the subcellular localization of the amyloid precursor protein and increases amyloid-beta deposition in N2a cells. *Biochem Biophys Res Commun*, 362(3):633–8, Oct 2007. PMID: 17727819.
- [104] A. Matzke, V. Sargsyan, B. Holtmann, G. Aramuni, E. Asan, M. Sendtner, G. Pace, N. Howells, W. Zhang, H. Ponta, and V. Orian-Rousseau. Haploinsufficiency of c-Met in CD44^{-/-} mice identifies a collaboration of CD44 and c-Met in vivo. *Mol Cell Biol*, 27(24):8797–8806, Dec 2007. PMID: 17923692.

- [105] L. McArdle, M. McDermott, R. Purcell, D. Grehan, A. O'Meara, F. Breatnach, D. Catchpoole, A. C. Culhane, I. Jeffery, W. M. Gallagher, and R. L. Stallings. Oligonucleotide microarray analysis of gene expression in neuroblastoma displaying loss of chromosome 11q. *Carcinogenesis*, 25(9):1599–1609, Sep 2004. PMID: 15090470.
- [106] T. McKinsey, C. Zhang, and E. Olson. MEF2: a calcium-dependent regulator of cell division, differentiation and death. *Trends in Biochemical Sciences*, 27:40–47, January 2002.
- [107] P. Mhawech, M. Berczy, M. Assaly, F. Herrmann, H. Bouzourene, A. Allal, P. Dulguerov, and J. Schwaller. Human achaete-scute homologue (hASH1) mRNA level as a diagnostic marker to distinguish esthesioneuroblastoma from poorly differentiated tumors arising in the sinonasal tract. *Am J Clin Pathol*, 122(1):100–5, Jul 2004. PMID: 15272537.
- [108] S. Miotti, A. Tomassetti, I. Facetti, E. Sanna, V. Berno, and S. Canevari. Simultaneous expression of caveolin-1 and E-cadherin in ovarian carcinoma cells stabilizes adherens junctions through inhibition of src-related kinases. *Am J Pathol*, 167(5):1411–27, Nov 2005. PMID: 16251425.
- [109] R. Mita, J. Coles, D. Glubrecht, R. Sung, X. Sun, and R. Godbout. B-FABP-expressing radial glial cells: the malignant glioma cell of origin. *Neoplasia*, 9(0):734–44, Sep 2007. PMID: 17898869.
- [110] R. Moore, F. Xiang, J. Monaghan, D. Han, Z. Zhang, L. Edström, M. Anvret, and S. Prusiner. Huntington disease phenocopy is a familial prion disease. *Am J Hum Genet*, 69(6):1385–8, Dec 2001. PMID: 11593450.
- [111] M. Morowitz, S. Shusterman, Y. Mosse, G. Hii, C. L. Winter, D. Khazi, Q. Wang, R. King, and J. M. Maris. Detection of single-copy chromosome 17q gain in human neuroblastomas using real-time quantitative polymerase chain reaction. *Mod Pathol*, 16(12):1248–1256, Dec 2003. PMID: 14681326.
- [112] D. Morris, A. Rodgers, K. McGhee, S. Schwaiger, P. Scully, J. Quinn, D. Meagher, J. Waddington, M. Gill, and A. Corvin. Confirming RGS4 as a susceptibility gene for schizophrenia. *Am J Med Genet B Neuropsychiatr Genet*, 125B(1):50–3, Feb 2004. PMID: 14755443.
- [113] A. Moumen, A. Ieraci, S. Patané, C. Solé, J. X. Comella, R. Dono, and F. Maina. Met signals hepatocyte survival by preventing Fas-triggered FLIP degradation in a PI3k-Akt-dependent manner. *Hepatology*, 45(5):1210–1217, May 2007. PMID: 17464994.
- [114] P. B. Mullan, J. E. Quinn, and D. P. Harkin. The role of BRCA1 in transcriptional regulation and cell cycle control. *Oncogene*, 25(43):5854–5863, Sep 2006. PMID: 16998500.
- [115] F. Naef, D. A. Lim, N. Patil, and M. A. Magnasco. From features to expression: High-density oligonucleotide array analysis revisited. Downloaded from <http://arxiv.org/abs/physics/0102010v2>, 2001.
- [116] F. Naef and M. M. O. Solving the riddle of the bright mismatches: Labeling and effective binding in oligonucleotide arrays. *Phys Rev E Stat Nonlin Soft Matter Phys*, 68(1 Pt 1):011906, July 2003. PMID: 12935175.
- [117] D. Naor, R. V. Sionov, and D. Ish-Shalom. CD44: structure, function, and association with the malignant process. *Adv Cancer Res*, 71:241–319, 1997. PMID: 9111868.
- [118] T. Nishimura, N. Takeshita, H. Satow, and K. Kohmoto. Expression of the mC26 gene encoding GlyCAM-1 in the lactating mouse mammary gland. *J Biochem*, 114(4):567–569, Oct 1993. PMID: 8276769.
- [119] J. Niu, W. Zhang, L. Ye, L. Wu, G. Zhu, Z. Yang, G. Grau, and J. Lou. The role of adhesion molecules, alpha v beta 3, alpha v beta 5 and their ligands in the tumor cell and endothelial cell adhesion. *Eur J Cancer Prev*, 16(6):517–27, Dec 2007. PMID: 18090124.

LITERATUR

- [120] A. Oertl, J. Castein, T. Engl, W.-D. Beecken, D. Jonas, R. Melamed, and R. A. Blaheta. Endothelial adhesion of synchronized gastric tumor cells changes during cell cycle transit and correlates with the expression level of CD44 splice variants. *World J Gastroenterol*, 11(40):6243–6248, Oct 2005. PMID: 16419149.
- [121] A. B. Olshen, E. S. Venkatraman, R. Lucito, and M. Wigler. Circular binary segmentation for the analysis of array-based DNA copy number data. *Biostat*, 5(4):557–572, 2004.
- [122] H. Osada, S. Tomida, Y. Yatabe, Y. Tatematsu, T. Takeuchi, H. Murakami, Y. Kondo, Y. Sekido, and T. Takahashi. Roles of achaete-scute homologue 1 in DKK1 and E-cadherin repression and neuroendocrine differentiation in lung cancer. *Cancer Res*, 68(6):1647–55, Mar 2008. PMID: 18339843.
- [123] S. Park, J. Kim, Y. Kim, Y. Kim, and S. Kim. Caveolin-1 is down-regulated and inversely correlated with HER2 and EGFR expression status in invasive ductal carcinoma of the breast. *Histopathology*, 47(6):625–30, Dec 2005. PMID: 16324201.
- [124] H. Parkinson, M. Kapushesky, M. Shojatalab, N. Abeygunawardena, R. Coulson, A. Farne, E. Holloway, N. Kolesnykov, P. Lilja, M. Lukk, R. Mani, T. Rayner, A. Sharma, E. William, U. Sarkans, and A. Brazma. ArrayExpress—a public database of microarray experiments and gene expression profiles. *Nucleic Acids Res*, 35:747–750, January 2007. PMID: 17132828.
- [125] K. Phillips, M. A. Park, L. H. Quarrie, M. Boutinaud, J. D. Lochrie, D. J. Flint, G. J. Allan, and J. Beattie. Hormonal control of IGF-binding protein (IGFBP)-5 and IGFBP-2 secretion during differentiation of the HC11 mouse mammary epithelial cell line. *J Mol Endocrinol*, 31(1):197–208, Aug 2003. PMID: 12914536.
- [126] J. J. Pinzone, B. M. Hall, N. K. Thudi, M. Vonau, Y.-W. Qiang, T. J. Rosol, and J. D. Shaughnessy Jr. The role of Dickkopf-1 in bone development, homeostasis and disease. *Blood*, Prepublished online October 3, Oct 2008.
- [127] Y.-W. Qiang, B. Barlogie, S. Rudikoff, and J. D. Shaughnessy. Dkk1-induced inhibition of Wnt signaling in osteoblast differentiation is an underlying mechanism of bone loss in multiple myeloma. *Bone*, 42(4):669–680, Apr 2008. PMID: 18294945.
- [128] Y.-W. Qiang, Y. Chen, O. Stephens, N. Brown, B. Chen, J. Epstein, B. Barlogie, and J. D. Shaughnessy. Myeloma-derived Dickkopf-1 disrupts Wnt-regulated osteoprotegerin and RANKL production by osteoblasts: a potential mechanism underlying osteolytic bone lesions in multiple myeloma. *Blood*, 112(1):196–207, Jul 2008. PMID: 18305214.
- [129] N. Rabbee and T. P. Speed. A genotype calling algorithm for affymetrix SNP arrays. *Bioinformatics*, 22(1):7–12, 2006.
- [130] I. Rapa, P. Ceppi, E. Bollito, R. Rosas, S. Cappia, E. Bacillo, F. Porpiglia, A. Berruti, M. Papotti, and M. Volante. Human ASH1 expression in prostate cancer with neuroendocrine differentiation. *Mod Pathol*, 21(6):700–7, Jun 2008. PMID: 18311112.
- [131] M. Rezaee, K. Penta, and T. Quertermous. D β 1 mediates VSMC adhesion, migration, and proliferation through interaction with integrin α (v) β (3). *Am J Physiol Heart Circ Physiol*, 282(5):H1924–32, May 2002. PMID: 11959660.
- [132] A. Robinson-White, T. Hundley, M. Shiferaw, J. Bertherat, F. Sandrini, and C. Stratakis. Protein kinase-A activity in PRKAR1A-mutant cells, and regulation of mitogen-activated protein kinases ERK1/2. *12*, 13(1475-84), Jul 2003. PMID: 12812976.
- [133] A. Robinson-White, W. Leitner, E. Aleem, P. Kaldis, I. Bossis, and C. Stratakis. PRKAR1A inactivation leads to increased proliferation and decreased apoptosis in human B lymphocytes. *Cancer Res*, 66(21):10603–12, Nov 2006. PMID: 17079485.

- [134] S. Roeber, E. Grasbon-Frodl, O. Windl, B. Krebs, W. Xiang, C. Vollmert, T. Illig, A. Schröter, T. Arzberger, P. Weber, I. Zerr, and H. Kretzschmar. Evidence for a pathogenic role of different mutations at codon 188 of PRNP. *PLoS ONE*, 3(5):e2147, May 2008. PMID: 18478114.
- [135] R. A. Ross, D. L. Lazarova, G. T. Manley, P. S. Smitt, B. A. Spengler, J. B. Posner, and J. L. Biedler. HuD, a neuronal-specific RNA-binding protein, is a potential regulator of MYCN expression in human neuroblastoma cells. *Eur J Cancer*, 33(12):2071–2074, Oct 1997. PMID: 9516855.
- [136] G. Rudolph, K. Schilbach-Stückle, R. Handgretinger, P. Kaiser, and H. Hameister. Cytogenetic and molecular characterization of a newly established neuroblastoma cell line LS. *Hum Genet*, 86(6):562–566, Apr 1991. PMID: 2026421.
- [137] W. Rudy, M. Hofmann, R. Schwartz-Albiez, M. Zöller, K. H. Heider, H. Ponta, and P. Herrlich. The two major CD44 proteins expressed on a metastatic rat tumor cell line are derived from different splice variants: each one individually suffices to confer metastatic behavior. *Cancer Res*, 53(6):1262–1268, Mar 1993. PMID: 8443806.
- [138] M. Salemi, A. Calogero, R. Castiglione, D. Tricoli, P. Asero, R. Rosa, G. Rappazzo, and E. Vicari. Expression of SpanX proteins in normal testes and in testicular germ cell tumours. *Int J Androl*, 29(2):368–73, Apr 2006. PMID: 16390498.
- [139] F. Sandrini, L. Matyakhina, N. Sarlis, L. Kirschner, C. Farmakidis, O. Gimm, and C. Stratakis. Regulatory subunit type I-alpha of protein kinase A (PRKAR1A): a tumor-suppressor gene for sporadic thyroid cancer. *Genes Chromosomes Cancer*, 35(2):182–92, Oct 2002. PMID: 12203783.
- [140] C. Scali, F. Caraci, M. Gianfriddo, E. Diodato, R. Roncarati, G. Pollio, G. Gaviraghi, A. Copani, F. Nicoletti, G. C. Terstappen, and A. Caricasole. Inhibition of Wnt signaling, modulation of Tau phosphorylation and induction of neuronal cell death by DKK1. *Neurobiol Dis*, 24(2):254–265, Nov 2006. PMID: 16919965.
- [141] M. Scannell, M. Flanagan, A. deStefani, K. Wynne, G. Cagney, C. Godson, and P. Maderna. Annexin-1 and peptide derivatives are released by apoptotic cells and stimulate phagocytosis of apoptotic neutrophils by macrophages. *J Immunol*, 178(7):4595–605, Apr 2007. PMID: 17372018.
- [142] C. Schulze-Garg, J. Löhler, A. Gocht, and W. Deppert. A transgenic mouse model for the ductal carcinoma in situ (DCIS) of the mammary gland. *Oncogene*, 19(8):1028–1037, Feb 2000. PMID: 10713686.
- [143] R. C. Seeger, S. A. Rayner, A. Banerjee, H. Chung, W. E. Laug, H. B. Neustein, and W. F. Benedict. Morphology, growth, chromosomal pattern and fibrinolytic activity of two new human neuroblastoma cell lines. *Cancer Res*, 37(5):1364–1371, May 1977. PMID: 856461.
- [144] J. Segarra, L. Balenci, T. Drenth, F. Maina, and F. Lamballe. Combined signaling through ERK, PI3K/AKT, and RAC1/p38 is required for met-triggered cortical neuron migration. *J Biol Chem*, 281(8):4771–4778, Feb 2006. PMID: 16361255.
- [145] M. Segura, C. Sole, M. Pascual, R. Moubarak, M. Perez-Garcia, R. Gozzelino, V. Iglesias, N. Badiola, J. Bayascas, N. Llecha, J. Rodriguez-Alvarez, E. Soriano, V. Yuste, and J. Comella. The long form of Fas apoptotic inhibitory molecule is expressed specifically in neurons and protects them against death receptor-triggered apoptosis. *J Neurosci*, 27(42):11228–41, Oct 2007. PMID: 17942717.
- [146] R. Seruca, R. F. Suijkerbuijk, F. Gärtner, B. Criado, I. Veiga, D. Olde-Weghuis, L. David, S. Castedo, and M. Sobrinho-Simões. Increasing levels of MYC and MET co-amplification during tumor progression of a case of gastric cancer. *Cancer Genet Cytogenet*, 82(2):140–145, Jul 1995. PMID: 7664243.
- [147] M. Shatz and M. Liscovitch. Caveolin-1: a tumor-promoting role in human cancer. *Int J Radiat Biol*, 84(3):177–189, Mar 2008. PMID: 18300018.

LITERATUR

- [148] D. Shen, F. Nooraie, Y. Elshimali, V. Lonsberry, J. He, S. Bose, D. Chia, D. Seligson, H. Chang, and L. Goodglick. Decreased expression of annexin A1 is correlated with breast cancer development and progression as determined by a tissue microarray analysis. *Hum Pathol*, 37(12):1583–91, Dec 2006. PMID: 16949910.
- [149] T. Shida, M. Furuya, T. Nikaido, T. Kishimoto, K. Koda, K. Oda, Y. Nakatani, M. Miyazaki, and H. Ishikura. Aberrant expression of human achaete-scute homologue gene 1 in the gastrointestinal neuroendocrine carcinomas. *Clin Cancer Res*, 11(2 Pt 1):450–8, Jan 2005. PMID: 15701827.
- [150] J. Shin, J. Kim, B. Ryu, S. Chi, and H. Park. Caveolin-1 is associated with VCAM-1 dependent adhesion of gastric cancer cells to endothelial cells. *Cell Physiol Biochem*, 17(5-6):211–20, 2006. PMID: 16790997.
- [151] E. Shtivelman and J. M. Bishop. Expression of CD44 is repressed in neuroblastoma cells. *Mol Cell Biol*, 11(11):5446–5453, Nov 1991. PMID: 1922057.
- [152] J. Silva, J. M. Silva, G. Domínguez, J. M. García, B. Cantos, R. Rodríguez, F. J. Larrondo, M. Provençio, P. España, and F. Bonilla. Concomitant expression of p16INK4a and p14ARF in primary breast cancer and analysis of inactivation mechanisms. *J Pathol*, 199(3):289–297, Mar 2003. PMID: 12579530.
- [153] R. Simon, E. Korn, L. McShane, M. Radmacher, G. Wright, and Y. Zhao. *Design and Analysis of DNA Microarray Investigations*. Statistics for Biology and Health. Springer Verlag, Berlin, 2003.
- [154] A. Slack, Z. Chen, R. Tonelli, M. Pule, L. Hunt, A. Pession, and J. M. Shohet. The p53 regulatory gene MDM2 is a direct transcriptional target of MYCN in neuroblastoma. *Proc Natl Acad Sci U S A*, 102(3):731–736, Jan 2005. PMID: 15644444.
- [155] G. A. Smolen, B. Muir, G. Mohapatra, A. Barmettler, W. J. Kim, M. N. Rivera, S. M. Haserlat, R. A. Okimoto, E. Kwak, S. Dahiya, J. E. Garber, D. W. Bell, D. C. Sgroi, L. Chin, C.-X. Deng, and D. A. Haber. Frequent met oncogene amplification in a Brca1/Trp53 mouse model of mammary tumorigenesis. *Cancer Res*, 66(7):3452–3455, Apr 2006. PMID: 16585167.
- [156] M. Sánchez-Font, A. Bosch-Comas, R. González-Duarte, and G. Marfany. Overexpression of FABP7 in Down syndrome fetal brains is associated with PKNOX1 gene-dosage imbalance. *Nucleic Acids Res*, 31(11):2769–77, Jun 2003. PMID: 12771203.
- [157] B. E. Stranger, M. S. Forrest, M. Dunning, C. E. Ingle, C. Beazley, N. Thorne, R. Redon, C. P. Bird, A. de Grassi, C. Lee, C. Tyler-Smith, N. Carter, S. W. Scherer, S. Tavaré, P. Deloukas, M. E. Hurles, and E. T. Dermitzakis. Relative impact of nucleotide and copy number variation on gene expression phenotypes. *Science*, 315(5813):848–853, Feb 2007. PMID: 17289997.
- [158] T. Suguri, A. Kikuta, H. Iwagaki, T. Yoshino, N. Tanaka, and K. Orita. Increased plasma GlyCAM-1, a mouse L-selectin ligand, in response to an inflammatory stimulus. *J Leukoc Biol*, 60(5):593–597, Nov 1996. PMID: 8929549.
- [159] M. Takahashi, M. Fujita, Y. Furukawa, R. Hamamoto, T. Shimokawa, N. Miwa, M. Ogawa, and Y. Nakamura. Isolation of a novel human gene, APCDD1, as a direct target of the beta-Catenin/T-cell factor 4 complex with probable involvement in colorectal carcinogenesis. *Cancer Res*, 62(20):5651–6, Oct 2002. PMID: 12384519.
- [160] M. Toselli, G. Biella, V. Taglietti, E. Cazzaniga, and M. Parenti. Caveolin-1 expression and membrane cholesterol content modulate N-type calcium channel activity in NG108-15 cells. *Biophys J*, 89(4):2443–57, Oct 2005. PMID: 16040758.
- [161] J. W. Tukey. *Exploratory Data Analysis*. Addison-Wesley, 1977.

- [162] D. Tulasne, J. Deheuninck, F. C. Lourenco, F. Lamballe, Z. Ji, C. Leroy, E. Puchois, A. Moumen, F. Maina, P. Mehlen, and V. Fafeur. Proapoptotic function of the MET tyrosine kinase receptor through caspase cleavage. *Mol Cell Biol*, 24(23):10328–10339, Dec 2004. PMID: 15542841.
- [163] J. J. Tumilowicz, W. W. Nichols, J. J. Cholon, and A. E. Greene. Definition of a continuous human cell line derived from neuroblastoma. *Cancer Res*, 30(8):2110–2118, Aug 1970. PMID: 5459762.
- [164] G. Tzircotis, R. F. Thorne, and C. M. Isacke. Chemotaxis towards hyaluronan is dependent on CD44 expression and modulated by cell type variation in CD44-hyaluronan binding. *J Cell Sci*, 118(Pt 21):5119–5128, Nov 2005. PMID: 16234326.
- [165] R. Ummanni, H. Junker, U. Zimmermann, S. Venz, S. Teller, J. Giebel, C. Scharf, C. Woenckhaus, F. Dombrowski, and R. Walther. Prohibitin identified by proteomic analysis of prostate biopsies distinguishes hyperplasia and cancer. *Cancer Lett*, 266(2):171–85, Aug 2008. PMID: 18384941.
- [166] UpToDate Online. An overview of treatment for locally advanced, recurrent, and metastatic breast cancer, May 2008. Last literature review version 16.2, May 2008.
- [167] UpToDate Online. Treatment and prognosis of neuroblastoma, March 2008. Last literature review version 16.2, May 2008.
- [168] A. Uren, S. Fallen, H. Yuan, A. Usubütün, T. Küçükali, R. Schlegel, and J. A. Toretsky. Activation of the canonical Wnt pathway during genital keratinocyte transformation: a model for cervical cancer progression. *Cancer Res*, 65(14):6199–6206, Jul 2005.
- [169] U. Valentiner, F. Valentiner, and U. Schumacher. Expression of CD44 is associated with a metastatic pattern of human neuroblastoma cells in a SCID mouse xenograft model. *Tumour Biol*, 29(3):152–60, 2008. PMID: 18612220.
- [170] S. Van Patten, D. Ng, J. Th'ng, K. Angelos, A. Smith, and D. Walsh. Molecular cloning of a rat testis form of the inhibitor protein of cAMP-dependent protein kinase. *Proc Natl Acad Sci U S A*, 88(12):5383–7, Jun 1991. PMID: 2052616.
- [171] L. J. van 't Veer, R. L. Beijersbergen, and R. Bernards. N-myc suppresses major histocompatibility complex class I gene expression through down-regulation of the p50 subunit of NF-kappa B. *EMBO J*, 12(1):195–200, Jan 1993. PMID: 8428579.
- [172] E. S. Venkatraman and A. B. Olshen. A faster circular binary segmentation algorithm for the analysis of array CGH data. *Bioinformatics*, 23(6):657–663, 2007.
- [173] A. Vincent-Salomon, C. Lucchesi, N. Gruel, V. Raynal, G. Pierron, R. Goudefroye, F. Reyat, F. Radvanyi, R. Salmon, J.-P. Thiery, X. Sastre-Garau, B. Sigal-Zafrani, A. Fourquet, O. Delattre, and breast cancer study group of the Institut Curie. Integrated genomic and transcriptomic analysis of ductal carcinoma in situ of the breast. *Clin Cancer Res*, 14(7):1956–1965, Apr 2008. PMID: 18381933.
- [174] N. Voorzanger-Rousselot, D. Goehrig, F. Journe, V. Doriath, J. J. Body, P. Clézardin, and P. Garnero. Increased Dickkopf-1 expression in breast cancer bone metastases. *Br J Cancer*, 97(7):964–970, Oct 2007. PMID: 17876334.
- [175] C. Wagener. *Molekulare Onkologie, Entstehung und Progression maligner Tumoren*. G. Thieme Verlag, Stuttgart, 2 edition, 1999.
- [176] H. C. Walters, A. L. Craddock, H. Fusegawa, M. C. Willingham, and P. A. Dawson. Expression, transport properties, and chromosomal location of organic anion transporter subtype 3. *Am J Physiol Gastrointest Liver Physiol*, 279(6):G1188–G1200, Dec 2000. PMID: 11093941.
- [177] B. Welch. The generalization of Student's "problem" when several different population variances are involved. *Biometrika*, 34:28–35, 1947.

LITERATUR

- [178] P. L. Welch, M. K. Lee, R. M. Gonzalez-Hernandez, D. J. Black, M. Mahadevappa, E. M. Swisher, J. A. Warrington, and M.-C. King. BRCA1 transcriptionally regulates genes involved in breast tumorigenesis. *Proc Natl Acad Sci U S A*, 99(11):7560–7565, May 2002. PMID: 12032322.
- [179] A. L. Welm, S. Kim, B. E. Welm, and J. M. Bishop. MET and MYC cooperate in mammary tumorigenesis. *Proc Natl Acad Sci U S A*, 102(12):4324–4329, Mar 2005. PMID: 15738393.
- [180] V. Westbrook, P. Schoppee, A. Diekman, K. Klotz, M. Allietta, K. Hogan, C. Slingluff, J. Patterson, H. Frierson, W. J. Irvin, C. Flickinger, M. Coppola, and J. Herr. Genomic organization, incidence, and localization of the SPAN-x family of cancer-testis antigens in melanoma tumors and cell lines. *Clin Cancer Res*, 10:101–112, Jan 2004. PMID: 14734458.
- [181] S. Wright. Adjusted P-values for simultaneous inference. *Biometrics*, 48:1005–1013, 1992.
- [182] Z. Wu, R. A. Irizarry, R. Gentleman, F. Martinez-Murillo, and F. Spencer. A Model Based Background Adjustment for Oligonucleotide Expression Arrays. *Johns Hopkins University, Dept. of Biostatistics Working Papers*, 2004. Working Paper 1.
- [183] G. Xiao, Y. Jin, Q. Lu, Z. Zhang, A. Beldegrun, R. Figlin, A. Pantuck, Y. Yen, F. Li, and J. Rao. Annexin-I as a potential target for green tea extract induced actin remodeling. *Int J Cancer*, 120(1):111–20, Jan 2007. PMID: 17019707.
- [184] Y. H. Yang, S. Dudoit, P. Luu, D. M. Lin, V. Peng, J. Ngai, and T. P. Speed. Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. *Nucl Acids Res*, 30(4):e15, 2002.
- [185] B. Yatsula, C. Galvao, M. McCrann, and A. S. Perkins. Assessment of F-MuLV-induced tumorigenesis reveals new candidate tumor genes including Pecam1, St7, and Prim2. *Leukemia*, 20(1):162–165, Jan 2006. PMID: 16307020.
- [186] Y. Yi, J. Mirosevich, Y. Shyr, R. Matusik, and A. L. George. Coupled analysis of gene expression and chromosomal location. *Genomics*, 85(3):401–412, Mar 2005. PMID: 15718107.
- [187] Z.-S. Zeng, M. R. Weiser, E. Kuntz, C.-T. Chen, S. A. Khan, A. Forslund, G. M. Nash, M. Gimbel, Y. Yamaguchi, A. T. Culliford, M. D’Alessio, F. Barany, and P. B. Paty. c-Met gene amplification is associated with advanced stage colorectal cancer and liver metastases. *Cancer Lett*, 265(2):258–269, Jul 2008. PMID: 18395971.
- [188] J. C. Zenklusen, C. J. Conti, and E. D. Green. Mutational and functional analyses reveal that ST7 is a highly conserved tumor-suppressor gene on human chromosome 7q31. *Nat Genet*, 27(4):392–398, Apr 2001. PMID: 11279520.
- [189] B. Zirn, B. Samans, S. Wittmann, T. Pietsch, I. Leuschner, N. Graf, and M. Gessler. Target genes of the WNT/beta-catenin pathway in Wilms tumors. *Genes Chromosomes Cancer*, 45(6):565–74, Jun 2006. PMID: 16575872.

Anhang A

Brusttumore

Auf den folgenden Seiten werden einige ergänzende Abbildungen und Tabellen des Brusttumorprojektes aufgeführt. Sie sollen einer Vervollständigung der in dieser Arbeit vorgestellten Daten dienen. Aufgeführt werden die EtBr-Agarosegelbilder aller Samples für die cDNA Transkription und nach der *in vitro* Transkription. Desweiteren findet sich eine tabellarische Auflistung aller im Korrelationsverfahren als signifikant identifizierten Gene, siehe Tabelle A.1.

ANHANG A. BRUSTTUMORE

Tabelle A.1: Brusttumore: Signifikant korrelierende Gene zwischen CNV-Analytik und Expressions-Analytik. Angeführt sind die Identifier der Gene auf dem Expressionsarray sowie die Identifier der am dichtesten gelegenen Probesets auf dem SNP-Array. $Q\ 0,01$ und $Q\ 0,001$ bezeichnen die P-Werte, die die 0,01 bzw. 0,001 Quantile denotieren. Beim *Raw P-Wert* handelt es sich um den unkorrigierten P-Wert. R^2 gibt den Korrelationskoeffizienten an, der im Idealfall (absolute Übereinstimmung) bei 1 liegt.

Symbol	EXP ID	CGH ID	Raw P-Wert	R^2	Q 0,01	Q 0,001	Chr.	Position
<i>Arrays aus Oslo</i>								
Mrpl9	78523	A_53_P102566	6,506	0,531	2,246	3,65	3	94250735
Psm4	19185	A_53_P144113	4,309	0,38	2,273	3,268	3	94841576
Cdc42se1	57912	A_53_P142129	4,632	0,405	2,184	3,226	3	95035768
Zfp364	67845	A_53_P130764	4,586	0,401	1,993	2,804	3	96563211
Psm5	26442	A_53_P112045	5,924	0,495	2,004	3,037	3	108071357
Shfm1	20422	A_53_P170734	3,757	0,335	2,056	2,914	6	6518460
Clgalt1	94192	A_53_P130866	7,929	0,61	3,389	5,231	6	7808690
Tmem168	101118	A_53_P109340	5,732	0,482	1,986	3,027	6	13544375
B630005N14Rik	101148	A_53_P109340	4,246	0,375	1,988	3,011	6	13603306
Tes	21753	A_53_P122283	14,264	0,829	2,454	3,394	6	17035507
Cav2	12390	A_53_P154945	10,347	0,715	2,522	3,531	6	17235165
Cav1	12389	A_53_P104726	5,343	0,456	2,437	3,467	6	17273898
Met	17295	A_53_P121226	15,176	0,848	2,344	3,717	6	17468890
Capza2	12343	A_53_P157131	13,959	0,91	2,291	3,34	6	17601821
St7	64213	A_53_P121352	13,787	0,891	2,399	3,392	6	17768484
Lsm8	76522	A_53_P103703	12,892	0,796	2,332	4,196	6	18801343
Calu	12321	A_53_P176635	7,289	0,576	4,07	5,453	6	29312410
Atp6v1f	66144	A_53_P179080	6,021	0,501	2,964	3,938	6	29419154
Ssbp1	381760	A_53_P141142	4,096	0,363	1,932	2,789	6	40426618
D11Wsu99e	28081	A_53_P157705	8,994	0,66	4,039	5,667	11	113534954
Rnasen	14000	A_53_P150067	11,047	0,74	4,966	6,853	15	12809835
Myo10	17909	A_53_P155317	3,934	0,35	1,98	2,809	15	25647854
Zfp622	52521	A_53_P168532	5,892	0,493	2,182	3,318	15	25920691
6-Mar	223455	A_53_P164496	8,809	0,652	4,43	5,881	15	31423223
Wdsof1	223499	A_53_P115059	9,581	0,685	2,244	3,908	15	38961438
E430025E21Rik	223593	A_53_P136541	5,743	0,483	2,268	3,358	15	59184637
Vps28	66914	A_53_P167739	7,442	0,585	3,208	4,569	15	76454521
Commd5	66398	A_53_P167739	8,622	0,643	2,727	4,569	15	76731053
Rbx1	56438	A_53_P142789	6,687	0,542	2,069	3,169	15	81301772
Poldip3	73826	A_53_P142447	9,324	0,675	2,427	3,469	15	82968066
Arfgap3	66251	A_53_P102365	3,636	0,325	2,004	3,002	15	83155423
Ttll1	319953	A_53_P125971	7,532	0,589	2,068	3,088	15	83327762
Atxn10	54138	A_53_P116218	5,497	0,467	2,243	3,557	15	85230224
Tbc1d22a	223754	A_53_P153756	6,529	0,533	2,027	3,129	15	86186903
Pfdn5	56612	A_53_P163891	7,564	0,591	2,166	3,24	15	102159264
Ube2l3	22195	A_53_P100087	5,143	0,442	1,944	2,885	16	17177424
Sod1	20655	A_53_P171158	4,438	0,39	1,969	2,983	16	90223806
Traf7	224619	A_53_P174585	4,372	0,385	2,003	3,076	17	24655285
Ier3ip1	66191	A_53_P174936	3,759	0,335	2,058	3,044	18	77189872
Mrpl21	353242	A_53_P138874	8,79	0,651	2,109	3,165	19	3287790
Chka	12660	A_53_P119828	4,857	0,421	1,994	2,941	19	3872700
Ndufs8	225887	A_53_P105009	7,698	0,598	2,136	3,579	19	3910810
Ndufv1	17995	A_53_P128679	8,509	0,638	2,06	3,022	19	4010105
Gstp1;Gstp2	14869;14870	A_53_P105009	7,234	0,573	2,241	3,459	19	4041283
Rbm4	19653	A_53_P136800	4,253	0,375	1,917	2,855	19	4789088
Rbm4	19653	A_53_P179232	4,253	0,375	1,917	2,855	19	4789088
Mrpl11	66419	A_53_P140319	8,678	0,646	2,024	2,961	19	4963649
Brms1	107392	A_53_P154906	6,152	0,51	1,987	2,816	19	5045652
Yif1a	68090	A_53_P179232	9,717	0,691	2,008	3,095	19	5090524
Sf3b2	319322	A_53_P125735	5,671	0,478	1,955	2,907	19	5284673
Banf1	23825	A_53_P149998	6,395	0,525	2,115	3,285	19	5365636
2010003J03Rik	69860	A_53_P149998	4,977	0,43	2,003	2,727	19	5368628
Sart1	20227	A_53_P126437	4,683	0,408	1,962	3,046	19	5383113
Fibp	58249	A_53_P152973	6,464	0,529	2,072	3,115	19	5462844
Cfl1	12631	A_53_P101264	4,119	0,365	2,193	3,495	19	5492259

Fortsetzung auf der nächsten Seite

Resultate der Brustkrebs Korrelationsanalyse – Fortsetzung

Symbol	EXP ID	CGH ID	Raw P-Wert	R ²	Q 0,01	Q 0,001	Chr.	Position
Mtvr2;Sssca1	17826;56390	A_53_P140371	5,31	0,454	2,14	3,443	19	5728867
Fau	14109	A_53_P136738	6,851	0,552	2,375	3,592	19	6058735
Snx15	69024	A_53_P105339	4,167	0,368	2,04	3,02	19	6123810
Cox8a	12868	A_53_P124485	6,535	0,533	2,101	3,243	19	7290854
Stx5a	56389	A_53_P117603	9,187	0,669	1,933	3,148	19	8823032
1810059G22Rik	67706	A_53_P117603	8,137	0,62	2,095	3,055	19	8847981
Polr2g	67710	A_53_P117603	5,064	0,436	2,196	3,165	19	8870321
D19Ertd721e	225896	A_53_P155534	10,752	0,73	2,071	3,134	19	8948156
AI462493	107197	A_53_P155534	6,983	0,559	2,033	2,952	19	8954978
1810009A15Rik	66276	A_53_P149682	10,681	0,727	1,965	2,91	19	8964320
Ints5	109077	A_53_P149682	5,191	0,445	1,971	2,97	19	8969938
Fth1	14319	A_53_P172214	5,858	0,491	2,203	3,281	19	10058388
1810006K21Rik	69038	A_53_P112363	5,169	0,444	1,93	2,722	19	10280506
0610038F07Rik	66072	A_53_P101617	9,208	0,67	2,495	3,495	19	10588238
Ddb1	13194	A_53_P115728	8,092	0,618	2,221	3,09	19	10692182
Tmem109	68539	A_53_P178963	10,106	0,706	2,074	2,988	19	10950689
Prpf19	28000	A_53_P120721	10,653	0,726	2,04	2,893	19	10974876
Mrpl16	94063	A_53_P160033	10,038	0,703	1,946	3,242	19	11847083
AV312086	225929	A_53_P110186	4,403	0,387	2,052	3,141	19	12003202
Cep78	208518	A_53_P126096	4,796	0,417	1,896	2,753	19	16044871
Gnaq	14682	A_53_P107982	8,089	0,618	2,01	2,94	19	16334632
Cbwd1	226043	A_53_P160716	7,472	0,586	1,948	2,793	19	25015472
Exosc1	66583	A_53_P123592	6,559	0,535	2,193	3,075	19	42002921
Mms19l;Zdhhc16	72199;74168	A_53_P160508	5,969	0,498	2,134	3,219	19	42036911
Avpil	69534	A_53_P160508	5,705	0,481	2,009	3,175	19	42200706
Ndufb8	67264	A_53_P118831	5,831	0,489	2,184	3,516	19	44627318
Mrpl43	94067	A_53_P123604	11,202	0,745	2,105	3,176	19	45080218
5330431N19Rik	226162	A_53_P145771	9,71	0,69	2,048	2,933	19	45643672
Trim8	93679	A_53_P146014	6,353	0,522	2,024	2,912	19	46583801
Usmg5	66477	A_53_P124014	6,311	0,519	2,305	3,377	19	47161540
Zdhhc6	66980	A_53_P147132	6,419	0,526	1,984	3,232	19	55377301
Prdx3	11757	A_53_P155292	4,43	0,389	1,969	3,14	19	60945193
<i>Arrays aus Amsterdam</i>								
Mrpl9	78523	sc122918,3,1_7	5,212	0,668	2,603	3,91	3	94250735
Psmid4	19185	sc122908,1,21_29	4,554	0,612	2,108	3,059	3	94841576
Cdc42se1	57912	sc10002021,1_5	6,369	0,748	2,725	4,223	3	95035768
Zfp364	67845	sc10002046,1_230	4,137	0,572	2,24	3,457	3	96563211
Psma5	26442	sc121678,15_29	4,001	0,558	2,016	2,914	3	108071357
Shfm1	20422	sc129298,3_504	4,625	0,618	2,141	3,317	6	6518460
C1galt1	94192	sc10094192,2_160	4,696	0,625	2,83	3,755	6	7808690
Tmem168	101118	sc129272,8_468	6,783	0,771	3,489	5,136	6	13544375
B630005N14Rik	101148	sc10064450,1_199	6,955	0,781	4,626	5,499	6	13603306
Tes	21753	sc10001061,1_50	6,296	0,743	2,519	3,597	6	17035507
Cav2	12390	sc10001108,1_2	5,222	0,669	2,621	3,759	6	17235165
Cav1	12389	sc10001108,1_2	5,682	0,703	2,487	3,601	6	17273898
Met	17295	sc130355,22,1_19	5,26	0,672	2,097	3,495	6	17468890
Capza2	12343	sc1012343,12_1	6,13	0,772	2,444	3,351	6	17601821
St7	64213	sc10001099,1_185	6,493	0,755	2,45	3,347	6	17768484
Lsm8	76522	sc100319202,1_96	8,981	0,865	1,879	3,46	6	18801343
Calu	12321	sc129213,18,1_30	8,309	0,841	5,373	6,51	6	29312410
Atp6v1f	66144	sc10319452,1_137	7,733	0,818	5,413	6,5	6	29419154
Ssbp1	381760	sc130159,2_384	4,824	0,636	2,433	3,745	6	40426618
D11Wsu99e	28081	sc139372,21_84	9,312	0,875	7,003	8,345	11	113534954
Rnasen	14000	sc10077877,1_184	3,871	0,544	1,993	2,945	15	12809835
Myo10	17909	sc10002440,1_24	3,913	0,549	1,914	2,665	15	25647854
Zfp622	52521	sc10017909,2_304	3,556	0,51	2,027	2,976	15	25920691
6-Mar	223455	sc148021,6_11	4,55	0,612	1,894	2,827	15	31423223
Wdsof1	223499	sc100118452,2_120	4,328	0,591	1,953	2,932	15	38961438
E430025E21Rik	223593	sc147890,1_269	9,752	0,888	5,024	6,648	15	59184637
Vps28	66914	sc147806,1,1_223	4,338	0,592	1,956	3,023	15	76454521

Fortsetzung auf der nächsten Seite

ANHANG A. BRUSTTUMORE

Resultate der Brustkrebs Korrelationsanalyse – Fortsetzung

Symbol	EXP ID	CGH ID	Raw P-Wert	R ²	Q 0,01	Q 0,001	Chr.	Position
Commd5	66398	scl47801,3_265	3,992	0,557	1,955	3,067	15	76731053
Rbx1	56438	scl0002434,1_7	4,268	0,585	1,974	2,747	15	81301772
Poldip3	73826	scl46917,6,1_157	5,16	0,664	1,895	2,741	15	82968066
Arfgap3	66251	scl46913,3_62	3,591	0,514	1,965	2,921	15	83155423
Ttll1	319953	scl0021411,1_1	5,372	0,68	1,883	2,753	15	83327762
Atxn10	54138	scl47668,10_45	3,799	0,537	2,007	2,858	15	85230224
Tbc1d22a	223754	scl47650,13_510	3,923	0,55	2,016	2,924	15	86186903
Pfdn5	56612	scl0053622,1_278	4,172	0,575	1,972	2,833	15	102159264
Ube2l3	22195	scl49376,15_6	4,016	0,56	1,964	3,015	16	17177424
Sod1	20655	scl48230,1,2_87	3,97	0,555	2,052	3,086	16	90223806
Traf7	224619	scl0001721,1_31	4,037	0,562	2,049	3,171	17	24655285
Ier3ip1	66191	scl0002235,1_6	3,564	0,511	1,991	2,947	18	77189872
Mrpl21	353242	scl054683,1_165	5,855	0,715	1,99	3,152	19	3287790
Chka	12660	scl000519,1_15	6,604	0,761	2,13	3,349	19	3872700
Ndufs8	225887	scl0109168,12_22	5,086	0,658	2,066	3,151	19	3910810
Ndufv1	17995	scl0069888,1_155	3,511	0,505	2,015	2,88	19	4010105
Gstp1;Gstp2	14869;14870	scl0069888,1_155	3,704	0,526	2,091	3,007	19	4041283
Rbm4	19653	scl081909,8_96	4,343	0,592	2,024	2,978	19	4789088
Mrpl11	66419	scl019708,12_324	4,833	0,637	2,005	2,972	19	4963649
Brms1	107392	scl53499,1,4_164	5,012	0,652	2,046	2,896	19	5045652
Yif1a	68090	scl53499,1,4_164	3,748	0,531	2,091	3,137	19	5090524
Sf3b2	319322	scl53492,15,1_3	4,201	0,578	1,981	2,917	19	5284673
Banf1	23825	scl53499,1,4_164	3,774	0,534	1,987	2,939	19	5365636
2010003J03Rik	69860	scl53499,1,4_164	4,133	0,572	1,876	2,581	19	5368628
Sart1	20227	scl0012631,1_129	4,368	0,595	1,894	2,802	19	5383113
Fibp	58249	scl0053611,1_163	5,036	0,654	1,984	3,061	19	5462844
Cfl1	12631	scl53492,15,1_3	5,608	0,698	2,024	3,057	19	5492259
Mtvr2;Sssca1	17826;56390	scl0053611,1_163	4,519	0,609	2,032	3,086	19	5728867
Fau	14109	scl056275,3_18	5,842	0,714	1,964	2,943	19	6058735
Snx15	69024	scl020743,1_114	3,977	0,556	2,062	3,136	19	6123810
0610038D11Rik	67674	scl52792,14_158	4,033	0,561	1,943	2,973	19	6985084
Cox8a	12868	scl016973,22_185	5,366	0,68	2,036	3,133	19	7290854
Stx5a	56389	scl068157,2_24	5,635	0,7	1,963	2,984	19	8823032
1810059G22Rik	67706	scl068157,2_24	5,645	0,7	1,986	2,988	19	8847981
Polr2g	67710	scl068157,2_24	6,297	0,743	2,032	2,966	19	8870321
D19Ertd721e	225896	scl068157,2_24	6,044	0,727	1,969	2,811	19	8948156
AI462493	107197	scl068157,2_24	4,557	0,612	1,989	3,079	19	8954978
1810009A15Rik	66276	scl068157,2_24	4,159	0,574	2,04	2,942	19	8964320
Ints5	109077	scl068157,2_24	4,558	0,612	1,912	2,976	19	8969938
Fth1	14319	scl52731,7,1_8	5,121	0,661	2,172	3,173	19	10058388
1810006K21Rik	69038	scl52731,7,1_8	5,401	0,682	2,084	3,077	19	10280506
0610038F07Rik	66072	scl52727,6_19	5,595	0,697	1,972	3,029	19	10588238
Ddb1	13194	scl060361,6_3	4,067	0,565	1,909	2,995	19	10692182
Tmem109	68539	scl52721,8,1_207	4,533	0,61	2,037	3,091	19	10950689
Prpf19	28000	scl52720,7,1_78	3,651	0,52	1,983	3,195	19	10974876
Mrpl16	94063	scl53342,2_24	4,735	0,628	2,03	3,022	19	11847083
AV312086	225929	scl00107146,2_299	6,426	0,751	2,004	2,972	19	12003202
Cep78	208518	scl53313,7_541	4,295	0,588	2,063	2,872	19	16044871
Gnaq	14682	scl52684,1,1_316	5,659	0,701	1,939	3,154	19	16334632
Cbwd1	226043	scl53254,5_108	4,403	0,598	1,974	3,025	19	25015472
Exosc1	66583	scl53113,11_103	4,246	0,583	1,921	2,779	19	42002921
Mms19l;Zdhhc16	72199;74168	scl53113,11_103	4,553	0,612	1,947	3,167	19	42036911
Avp1l	69534	scl53109,1,1_278	3,528	0,507	2,041	2,821	19	42200706
Ndufb8	67264	scl52438,3,1_120	4,489	0,606	2,007	3,036	19	44627318
Mrpl43	94067	scl52428,1_1	4,12	0,57	1,985	2,99	19	45080218
5330431N19Rik	226162	scl52418,9,1_3	5,778	0,71	1,943	2,859	19	45643672
Trim8	93679	scl057344,13_241	4,303	0,588	1,93	2,686	19	46583801
Usmg5	66477	scl52384,15,1_68	3,632	0,518	1,966	2,724	19	47161540
Zdhhc6	66980	scl52991,1,1_77	5,932	0,72	1,941	2,942	19	55377301
Prdx3	11757	scl52967,8,1_33	4,333	0,591	2,052	2,96	19	60945193

Ende Tabelle A.1

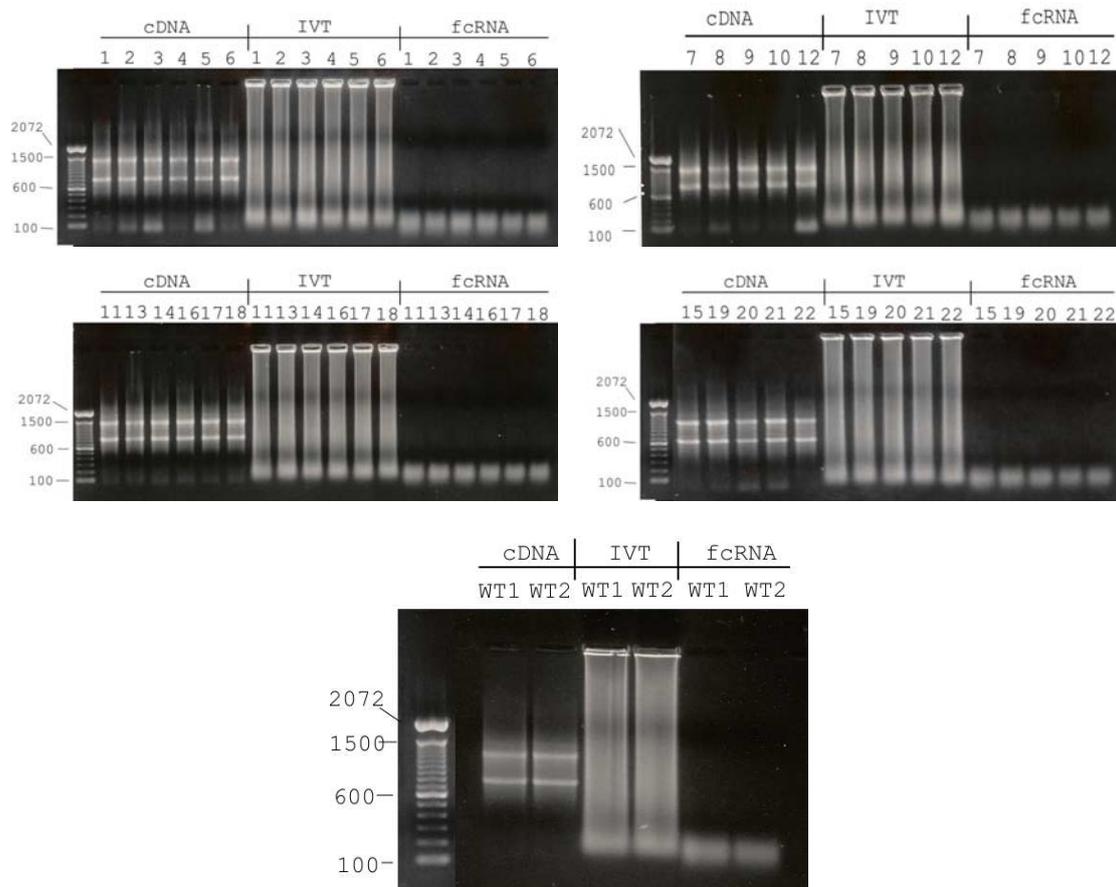


Abbildung A.1: 1%ige EtBr-Agarosegele der Brusttumorsamples. Aufgetragen wurden die Kontrollaliquots nach cDNA-Synthese, nach der *in vitro*-Transkription und nach erfolgter Fragmentierung.

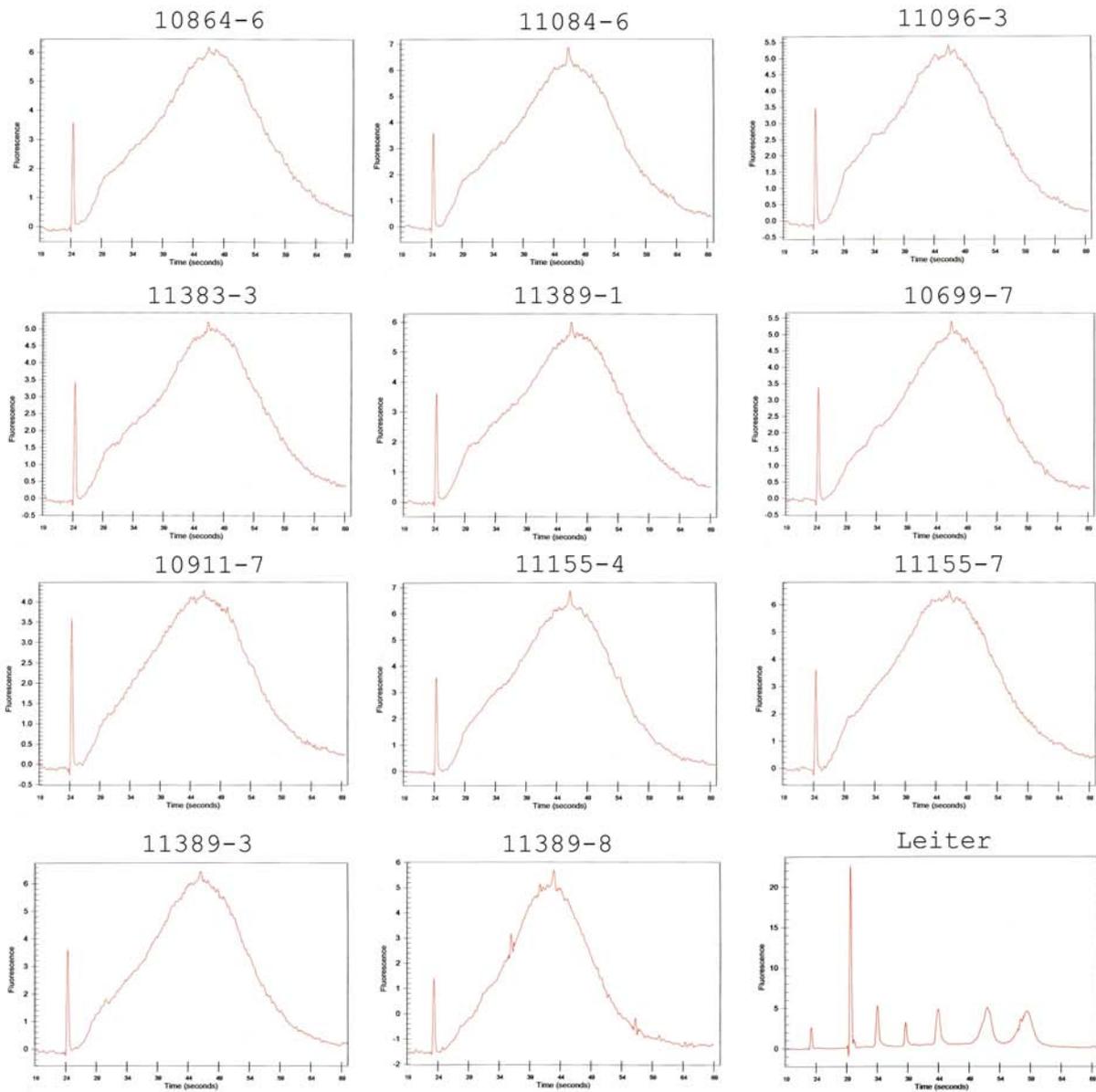


Abbildung A.2: Brusttumore: Elektropherogramme der Kontrollen, Sample 1 bis 11, auf dem Bioanalyser nach erfolgter *in vitro*-Transkription. Der Spitzenpeak der Verteilung sollte nach 39 Sekunden auftreten. Die numerische Zuordnung der Samples kann Tabelle 3.2 entnommen werden.

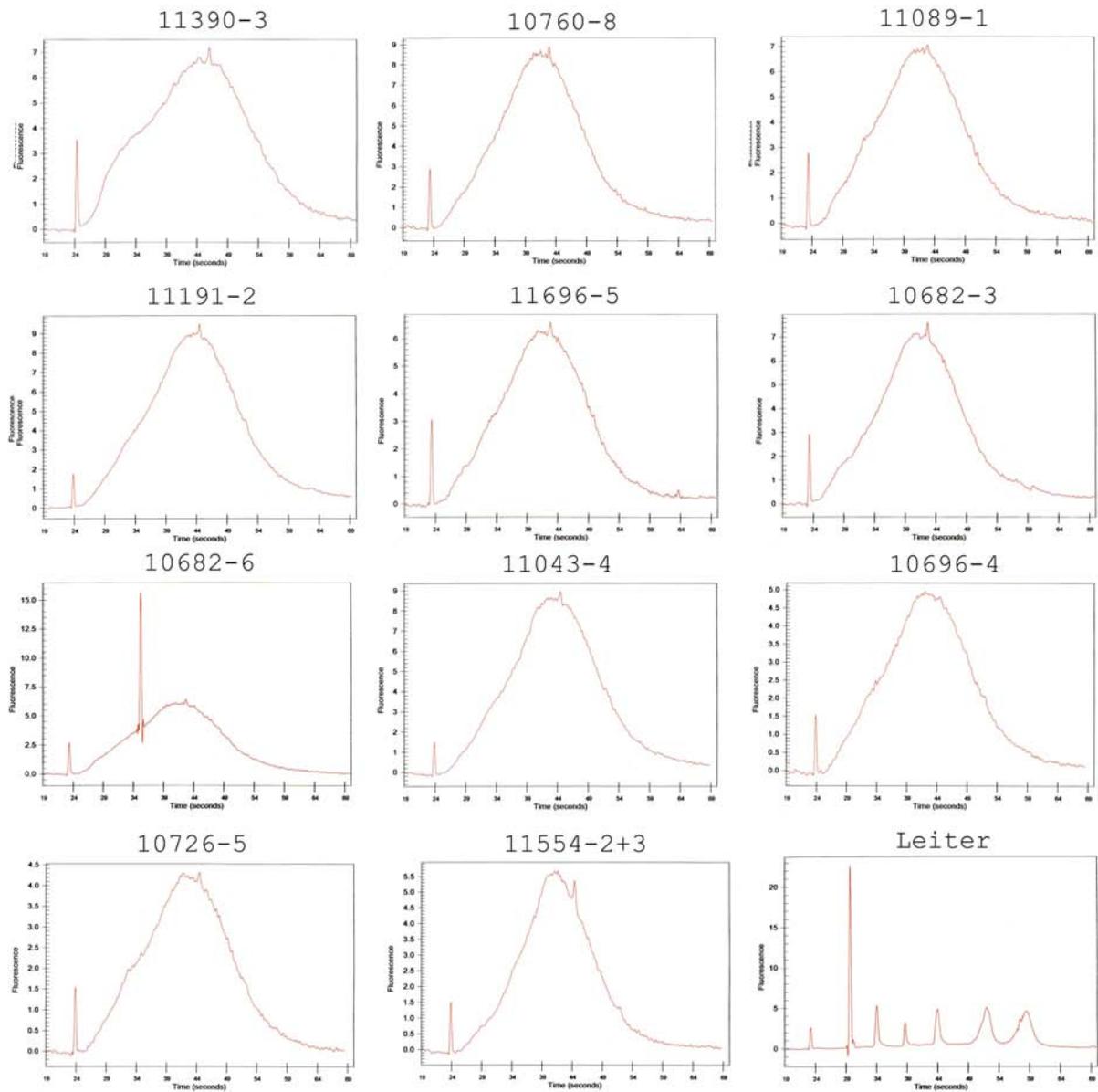


Abbildung A.3: Brusttumore: Elektropherogramme der Kontrollen, Sample 12 bis 22, auf dem Bioanalyser nach erfolgter *in vitro*-Transkription. Der Spitzenpeak der Verteilung sollte nach 39 Sekunden auftreten. Die numerische Zuordnung der Samples kann Tabelle 3.2 entnommen werden.

Anhang B

Neuroblastome

Auf den folgenden Seiten werden einige ergänzende Abbildungen und Tabellen des Neuroblastomprojektes aufgeführt. Sie sollen einer Vervollständigung der in dieser Arbeit vorgestellten Daten dienen. Aufgeführt werden die Projektionen der Expressionswerte auf die genomischen Segmente. Desweiteren finden sich die EtBr-Agarosegelbilder nach der Restriktion. Auf eine tabellarische Auflistung aller im Korrelationsverfahren als signifikant identifizierten Gene, analog zu Tabelle A.1 (Brusttumore), wird aufgrund der Datenmenge verzichtet. Die Tabelle wird auf dem beiliegenden Datenträger veröffentlicht.

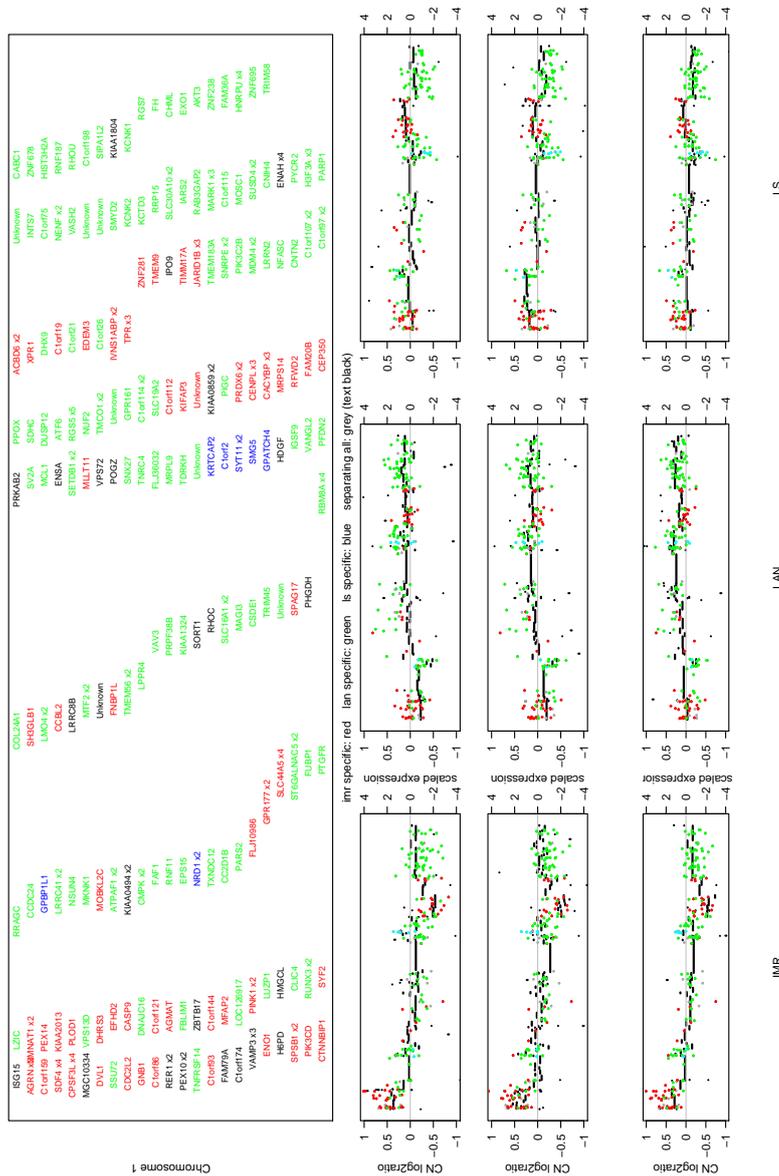


Abbildung B.1: Neuroblastome, Chromosom 1: Projektion der Expressionswerte für die gut korrelierenden Gene (siehe Abschnitt 3.3.2) auf die Segmente der SNP-Analytik entlang Chromosom 1. Die Neuroblastomzelllinien sind in der Reihenfolge IMR 32, LAN5, LS von links nach rechts angeordnet, die drei Samples einer jeden Zelllinie von oben nach unten. In der oberen Hälfte der Grafik findet sich die Annotation der Gene. IMR-spezifische Gene (siehe Abschnitt 3.1.2) sind rot koloriert, LAN-spezifische Gene grün, LS-spezifische Gene blau und Gene, die in allen drei Zelllinien differentiell exprimiert waren schwarz (Annotation) bzw. grau (Expressions-Spots).

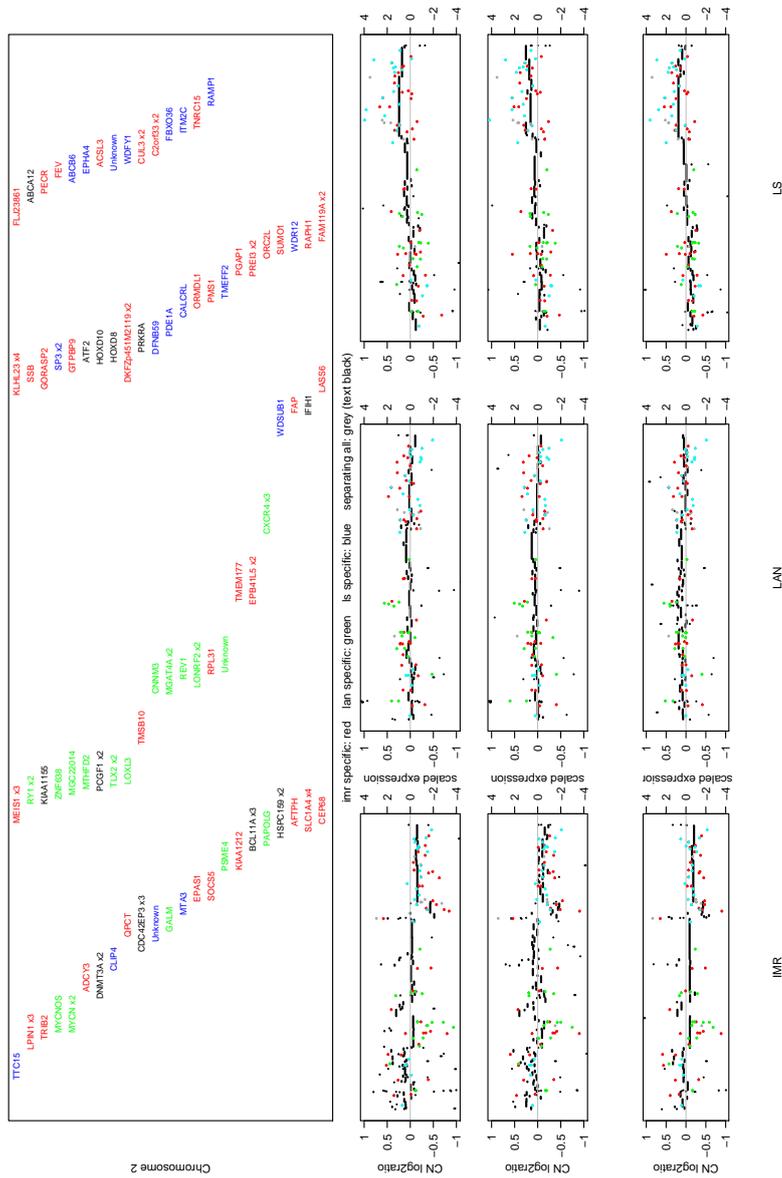


Abbildung B.2: Neuroblastome, Chromosom 2: Projektion der Expressionswerte für die gut korrelierenden Gene (siehe Abschnitt 3.3.2) auf die Segmente der SNP-Analytik entlang Chromosom 2. Die Neuroblastomzelllinien sind in der Reihenfolge IMR 32, LAN5, LS von links nach rechts angeordnet, die drei Samples einer jeden Zelllinie von oben nach unten. In der oberen Hälfte der Grafik findet sich die Annotation der Gene. IMR-spezifische Gene (siehe Abschnitt 3.1.2) sind rot koloriert, LAN-spezifische Gene grün, LS-spezifische Gene blau und Gene, die in allen drei Zelllinien differentiell exprimiert waren schwarz (Annotation) bzw. grau (Expressions-Spots).

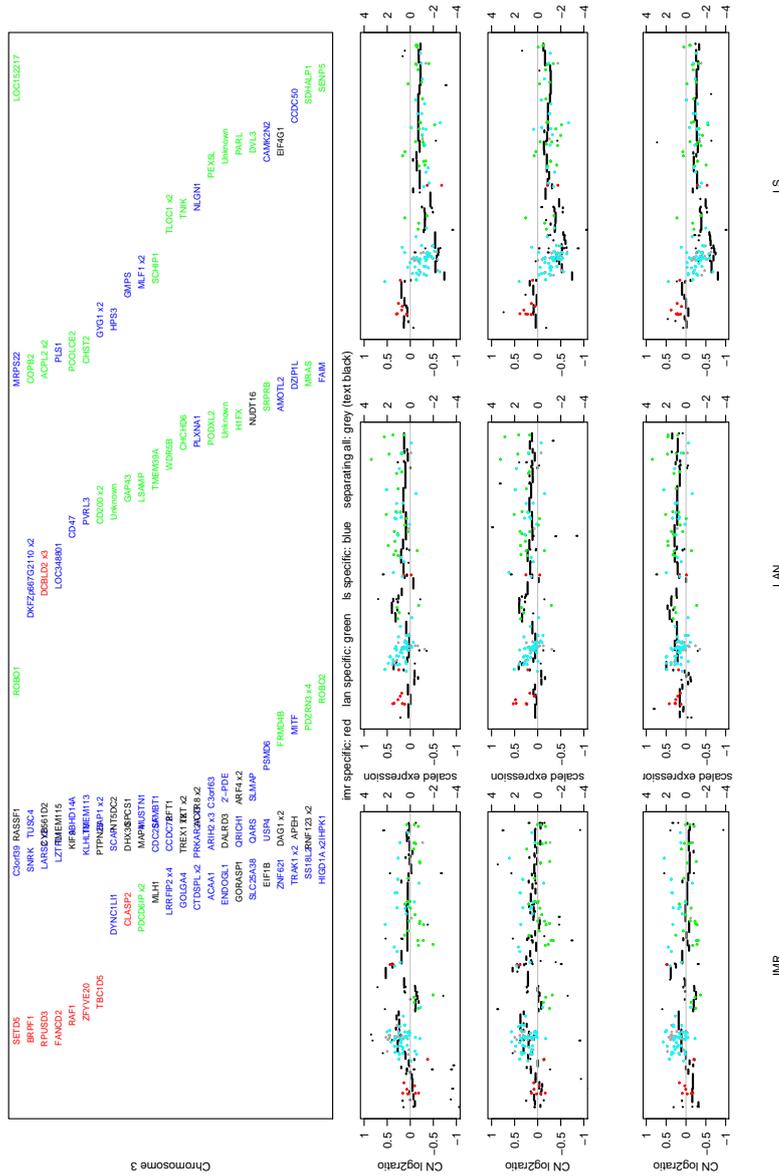


Abbildung B.3: Neuroblastome, Chromosom 3: Projektion der Expressionswerte für die gut korrelierenden Gene (siehe Abschnitt 3.3.2) auf die Segmente der SNP-Analytik entlang Chromosom 3. Die Neuroblastomzelllinien sind in der Reihenfolge IMR 32, LAN5, LS von links nach rechts angeordnet, die drei Samples einer jeden Zelllinie von oben nach unten. In der oberen Hälfte der Grafik findet sich die Annotation der Gene. IMR-spezifische Gene (siehe Abschnitt 3.1.2) sind rot koloriert, LAN-spezifische Gene grün, LS-spezifische Gene blau und Gene, die in allen drei Zelllinien differentiell exprimiert waren schwarz (Annotation) bzw. grau (Expressions-Spots).

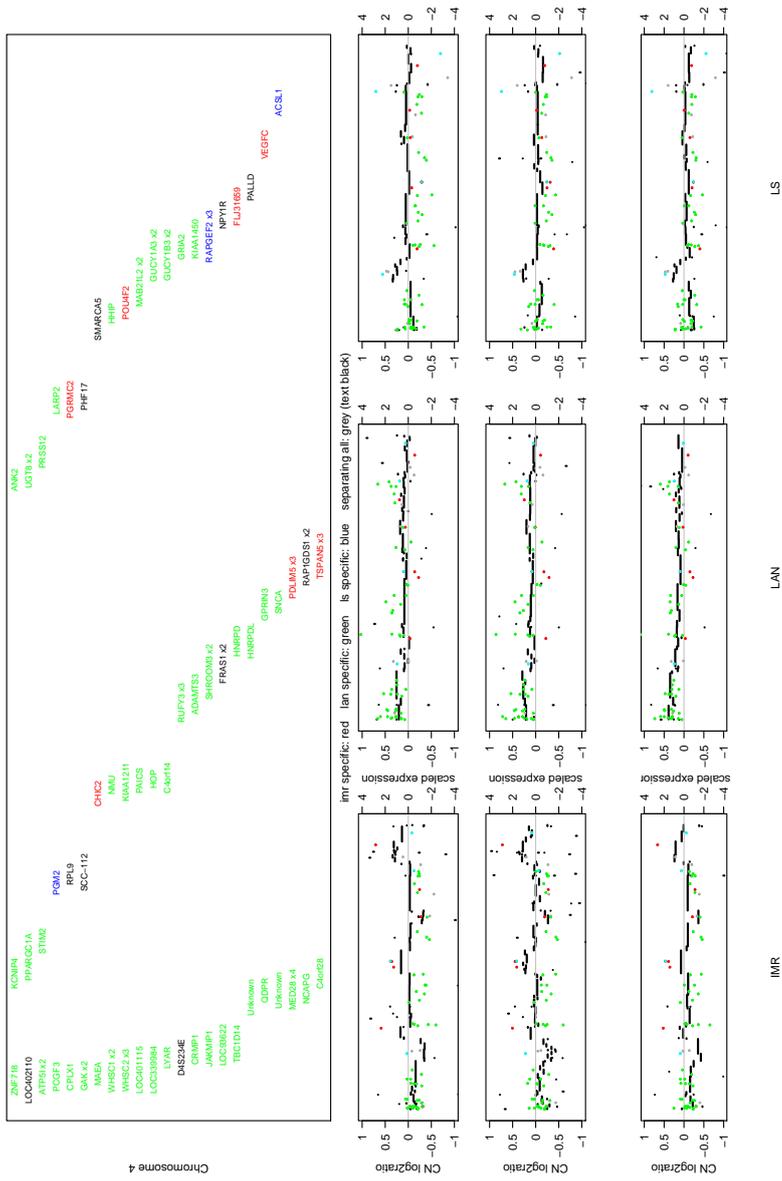


Abbildung B.4: Neuroblastome, Chromosom 4: Projektion der Expressionswerte für die gut korrelierenden Gene (siehe Abschnitt 3.3.2) auf die Segmente der SNP-Analytik entlang Chromosom 4. Die Neuroblastomzelllinien sind in der Reihenfolge IMR 32, LAN5, LS von links nach rechts angeordnet, die drei Samples einer jeden Zelllinie von oben nach unten. In der oberen Hälfte der Grafik findet sich die Annotation der Gene. IMR-spezifische Gene (siehe Abschnitt 3.1.2) sind rot koloriert, LAN-spezifische Gene grün, LS-spezifische Gene blau und Gene, die in allen drei Zelllinien differentiell exprimiert waren schwarz (Annotation) bzw. grau (Expressions-Spots).

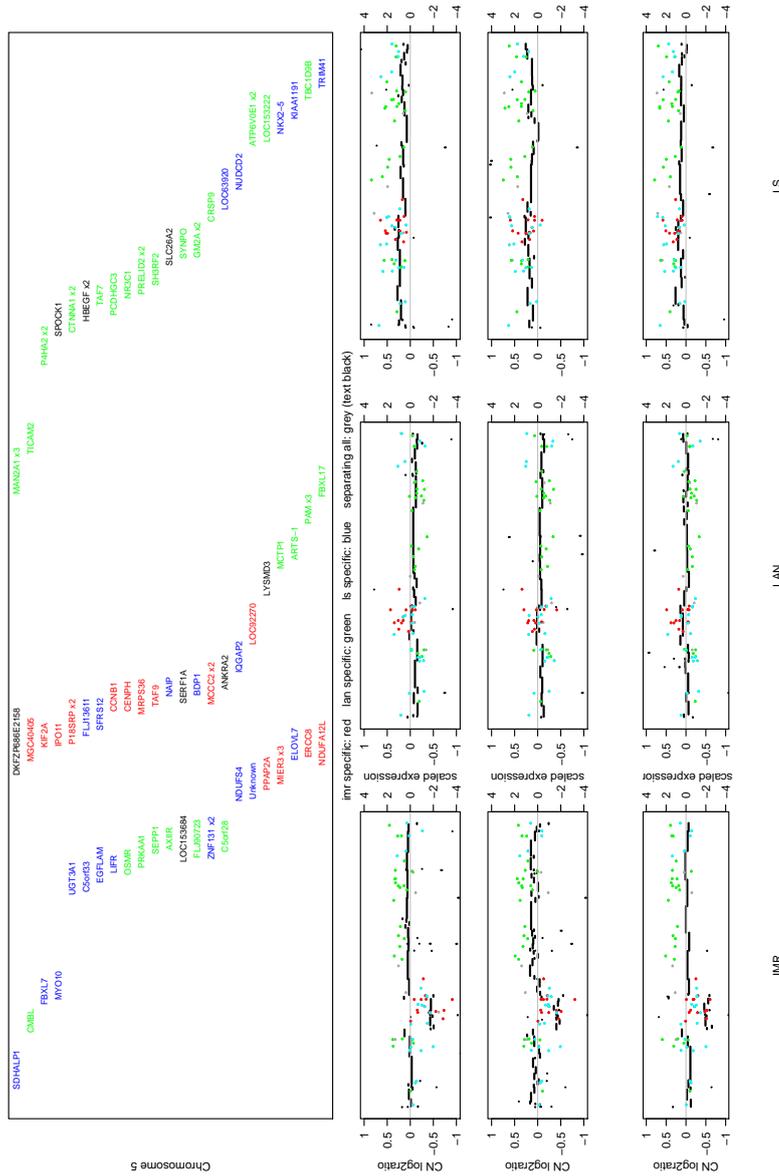


Abbildung B.5: Neuroblastome, Chromosom 5: Projektion der Expressionswerte für die gut korrelierenden Gene (siehe Abschnitt 3.3.2) auf die Segmente der SNP-Analytik entlang Chromosom 5. Die Neuroblastomzelllinien sind in der Reihenfolge IMR 32, LAN5, LS von links nach rechts angeordnet, die drei Samples einer jeden Zelllinie von oben nach unten. In der oberen Hälfte der Grafik findet sich die Annotation der Gene. IMR-spezifische Gene (siehe Abschnitt 3.1.2) sind rot koloriert, LAN-spezifische Gene grün, LS-spezifische Gene blau und Gene, die in allen drei Zelllinien differentiell exprimiert waren schwarz (Annotation) bzw. grau (Expressions-Spots).

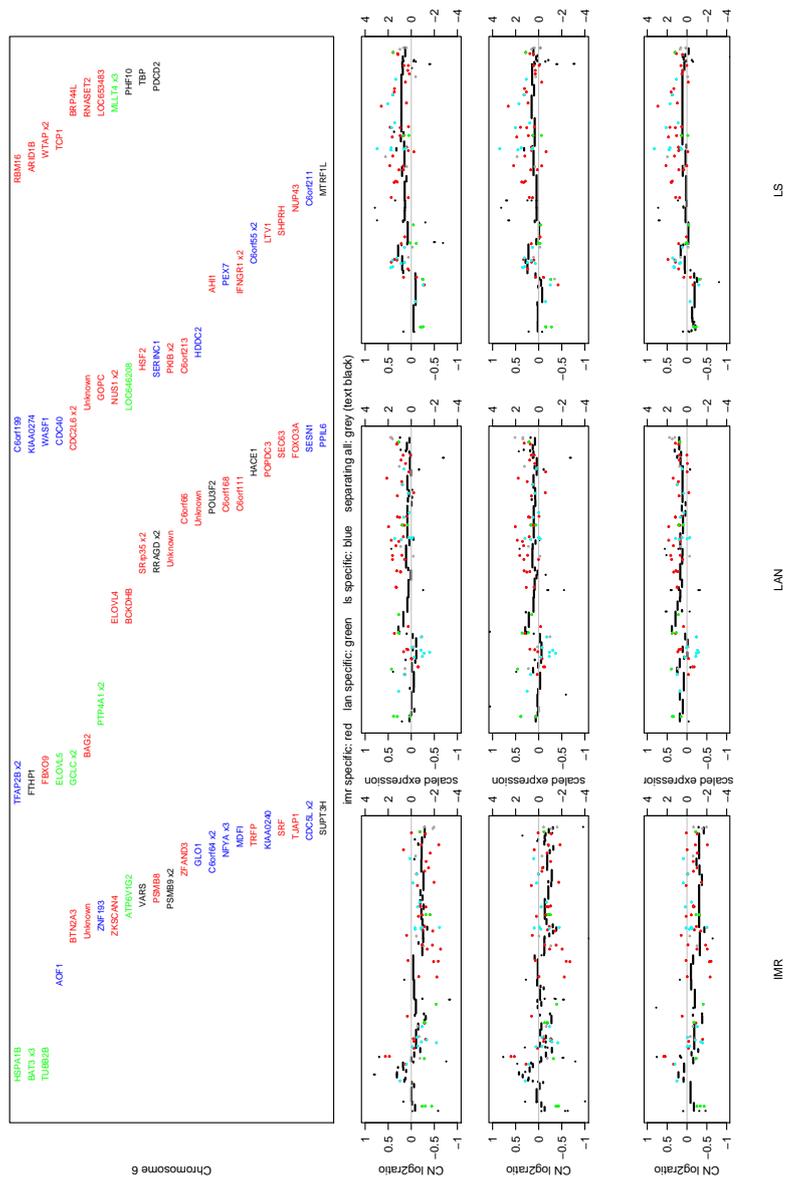


Abbildung B.6: Neuroblastome, Chromosom 6: Projektion der Expressionswerte für die gut korrelierenden Gene (siehe Abschnitt 3.3.2) auf die Segmente der SNP-Analytik entlang Chromosom 6. Die Neuroblastomzelllinien sind in der Reihenfolge IMR 32, LAN5, LS von links nach rechts angeordnet, die drei Samples einer jeden Zelllinie von oben nach unten. In der oberen Hälfte der Grafik findet sich die Annotation der Gene. IMR-spezifische Gene (siehe Abschnitt 3.1.2) sind rot koloriert, LAN-spezifische Gene grün, LS-spezifische Gene blau und Gene, die in allen drei Zelllinien differentiell exprimiert waren schwarz (Annotation) bzw. grau (Expressions-Spots).

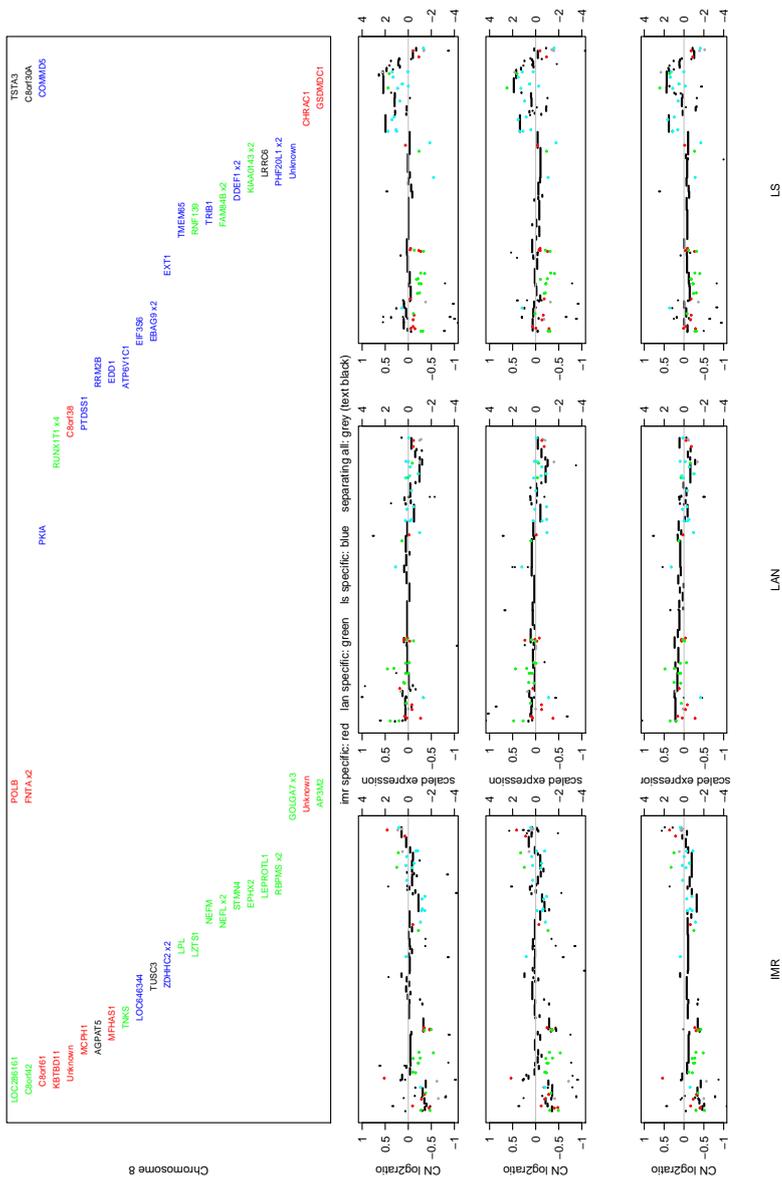


Abbildung B.8: Neuroblastome, Chromosom 8: Projektion der Expressionswerte für die gut korrelierenden Gene (siehe Abschnitt 3.3.2) auf die Segmente der SNP-Analytik entlang Chromosom 8. Die Neuroblastomzelllinien sind in der Reihenfolge IMR 32, LAN5, LS von links nach rechts angeordnet, die drei Samples einer jeden Zelllinie von oben nach unten. In der oberen Hälfte der Grafik findet sich die Annotation der Gene. IMR-spezifische Gene (siehe Abschnitt 3.1.2) sind rot koloriert, LAN-spezifische Gene grün, LS-spezifische Gene blau und Gene, die in allen drei Zelllinien differentiell exprimiert waren schwarz (Annotation) bzw. grau (Expressions-Spots).

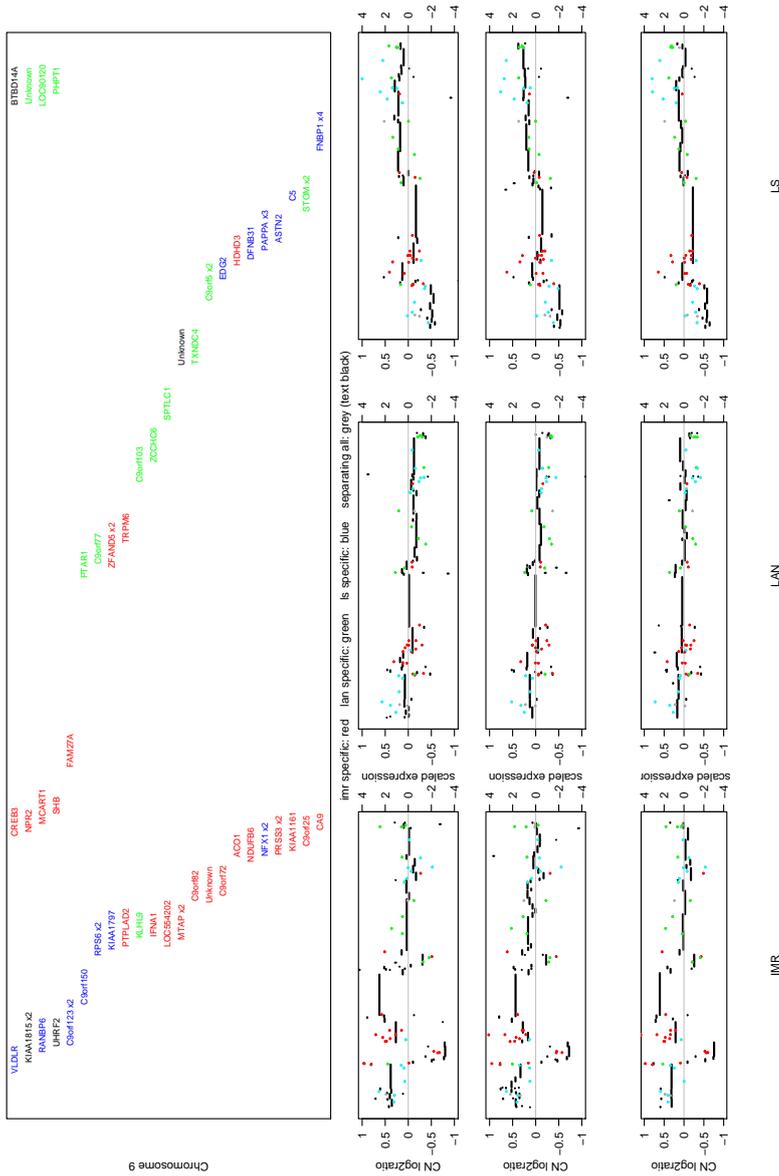


Abbildung B.9: Neuroblastome, Chromosom 9: Projektion der Expressionswerte für die gut korrelierenden Gene (siehe Abschnitt 3.3.2) auf die Segmente der SNP-Analytik entlang Chromosom 9. Die Neuroblastomzelllinien sind in der Reihenfolge IMR 32, LAN5, LS von links nach rechts angeordnet, die drei Samples einer jeden Zelllinie von oben nach unten. In der oberen Hälfte der Grafik findet sich die Annotation der Gene. IMR-spezifische Gene (siehe Abschnitt 3.1.2) sind rot koloriert, LAN-spezifische Gene grün, LS-spezifische Gene blau und Gene, die in allen drei Zelllinien differentiell exprimiert waren schwarz (Annotation) bzw. grau (Expressions-Spots).

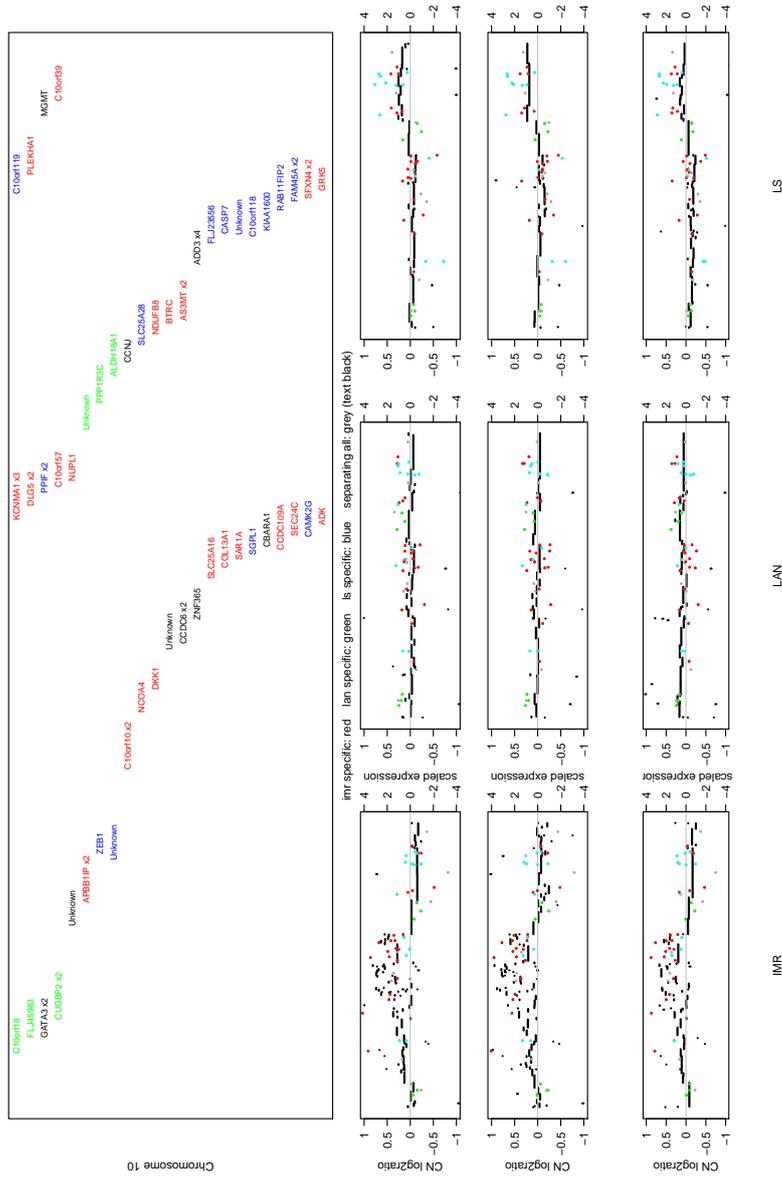


Abbildung B.10: Neuroblastome, Chromosom 10: Projektion der Expressionswerte für die gut korrelierenden Gene (siehe Abschnitt 3.3.2) auf die Segmente der SNP-Analytik entlang Chromosom 10. Die Neuroblastomzelllinien sind in der Reihenfolge IMR 32, LAN5, LS von links nach rechts angeordnet, die drei Samples einer jeden Zelllinie von oben nach unten. In der oberen Hälfte der Grafik findet sich die Annotation der Gene. IMR-spezifische Gene (siehe Abschnitt 3.1.2) sind rot koloriert, LAN-spezifische Gene grün, LS-spezifische Gene blau und Gene, die in allen drei Zelllinien differentiell exprimiert waren schwarz (Annotation) bzw. grau (Expressions-Spots).

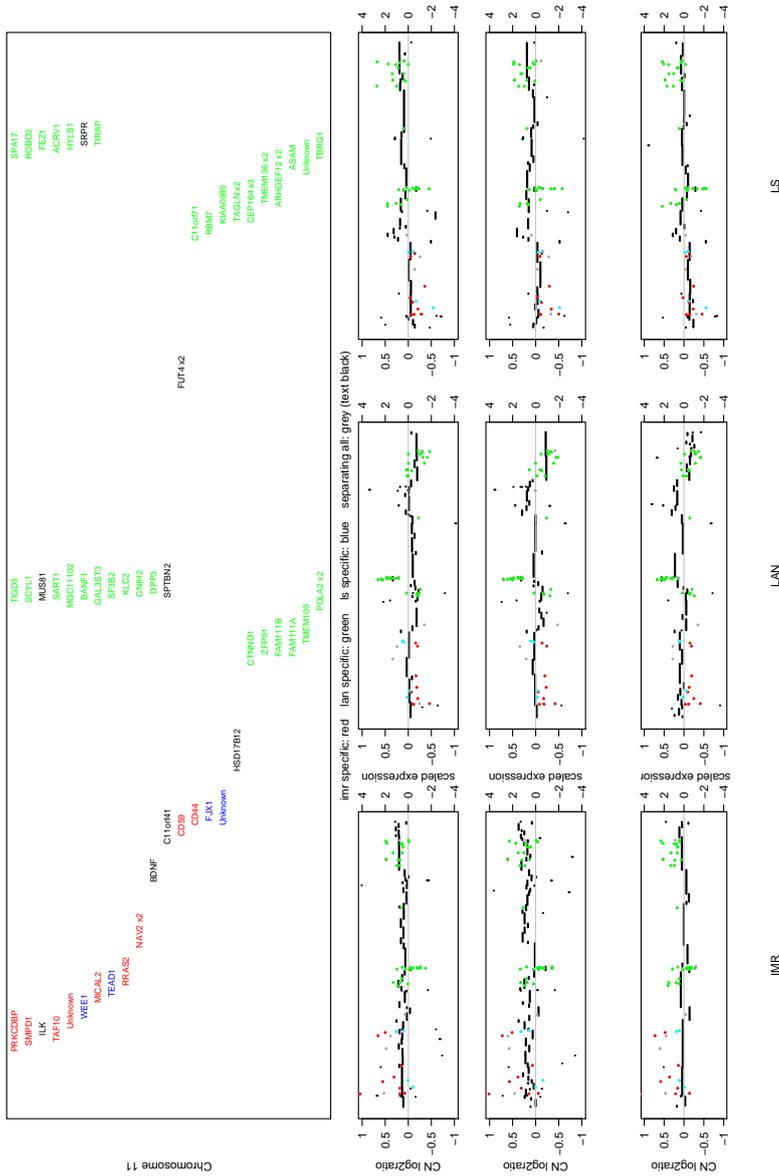


Abbildung B.11: Neuroblastome, Chromosom 11: Projektion der Expressionswerte für die gut korrelierenden Gene (siehe Abschnitt 3.3.2) auf die Segmente der SNP-Analytik entlang Chromosom 11. Die Neuroblastomzelllinien sind in der Reihenfolge IMR 32, LAN5, LS von links nach rechts angeordnet, die drei Samples einer jeden Zelllinie von oben nach unten. In der oberen Hälfte der Grafik findet sich die Annotation der Gene. IMR-spezifische Gene blau und Gene, die in allen drei Zelllinien differentiell exprimiert waren schwarz (Annotation) bzw. grau (Expressions-Spots).

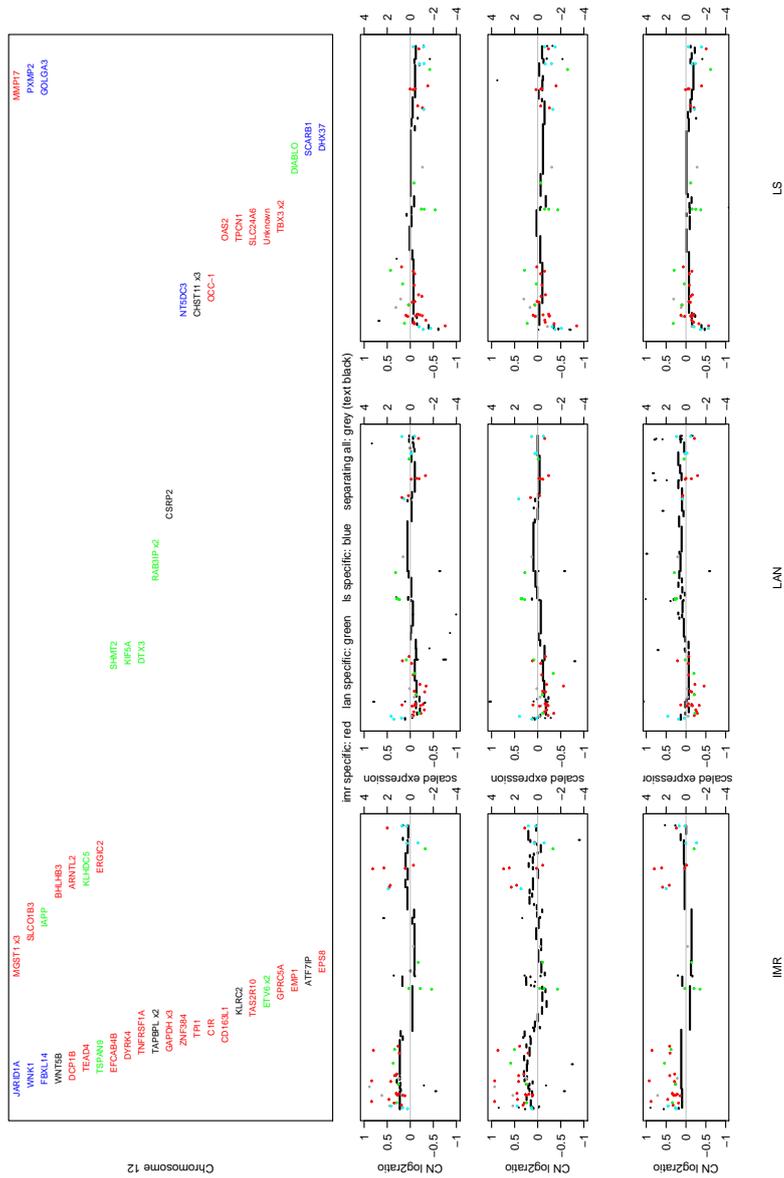


Abbildung B.12: Neuroblastome, Chromosom 12: Projektion der Expressionswerte für die gut korrelierenden Gene (siehe Abschnitt 3.3.2) auf die Segmente der SNP-Analytik entlang Chromosom 12. Die Neuroblastomzelllinien sind in der Reihenfolge IMR 32, LAN5, LS von links nach rechts angeordnet, die drei Samples einer jeden Zelllinie von oben nach unten. In der oberen Hälfte der Grafik findet sich die Annotation der Gene. IMR-spezifische Gene blau und Gene, die in allen drei Zelllinien differentiell exprimiert waren schwarz (Annotation) bzw. grau (Expressions-Spots).

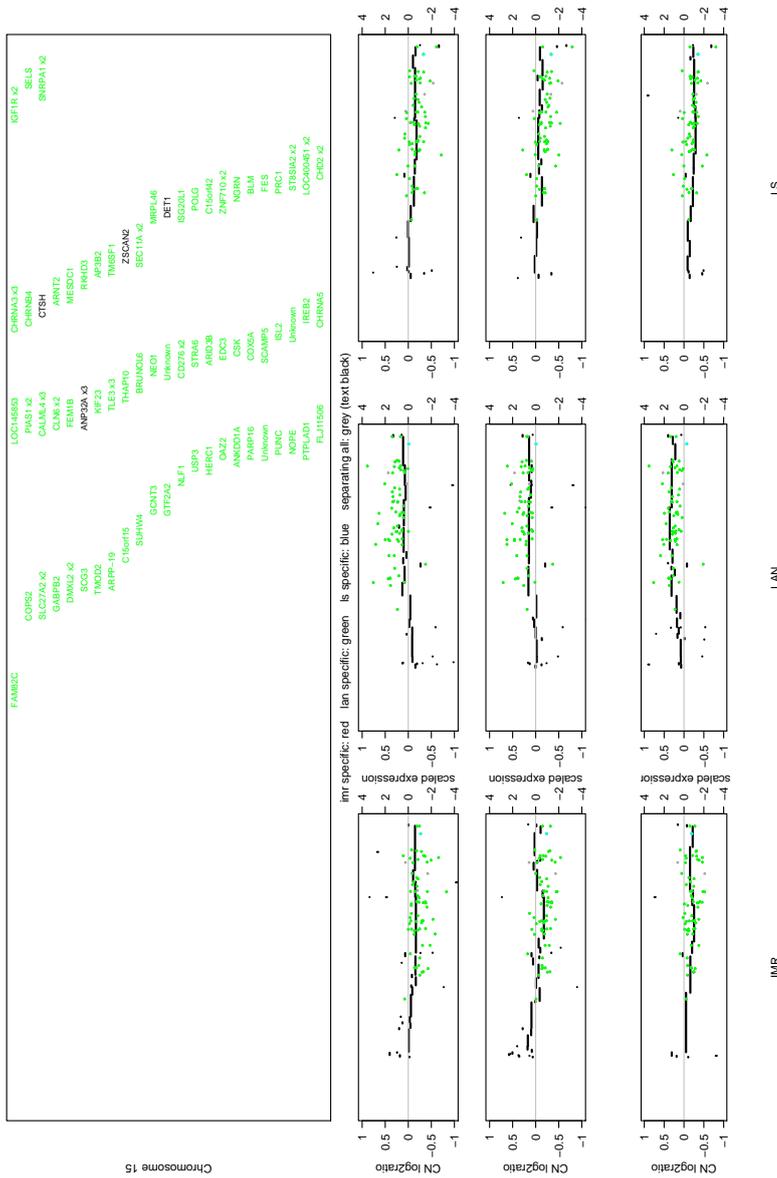


Abbildung B.15: Neuroblastome, Chromosom 15: Projektion der Expressionswerte für die gut korrelierenden Gene (siehe Abschnitt 3.3.2) auf die Segmente der SNP-Analytik entlang Chromosom 15. Die Neuroblastomzelllinien sind in der Reihenfolge IMR 32, LAN5, LS von links nach rechts angeordnet, die drei Samples einer jeden Zelllinie von oben nach unten. In der oberen Hälfte der Grafik findet sich die Annotation der Gene. IMR-spezifische Gene blau und Gene, die in allen drei Zelllinien differentiell exprimiert waren schwarz (Annotation) bzw. grau (Expressions-Spots).

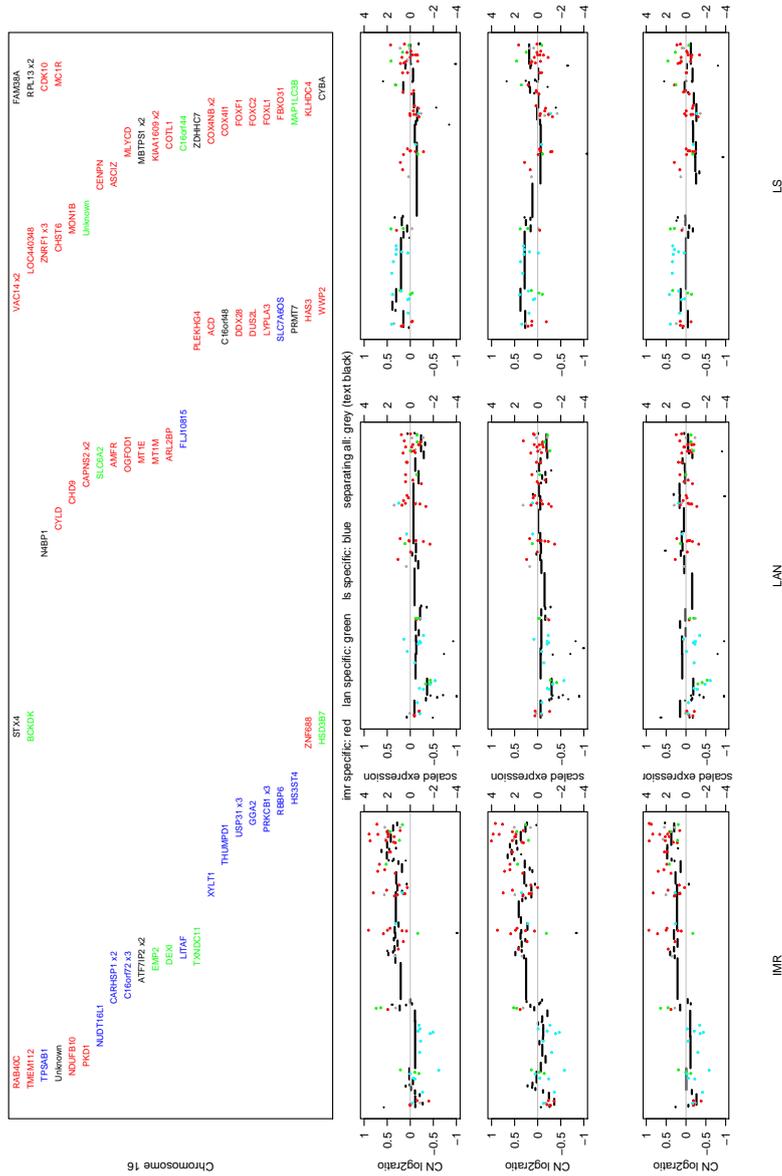


Abbildung B.16: Neuroblastome, Chromosom 16: Projektion der Expressionswerte für die gut korrelierenden Gene (siehe Abschnitt 3.3.2) auf die Segmente der SNP-Analytik entlang Chromosom 16. Die Neuroblastomzelllinien sind in der Reihenfolge IMR 32, LAN5, LS von links nach rechts angeordnet, die drei Samples einer jeden Zelllinie von oben nach unten. In der oberen Hälfte der Grafik findet sich die Annotation der Gene. IMR-spezifische Gene (siehe Abschnitt 3.1.2) sind rot koloriert, LAN-spezifische Gene grün, LS-spezifische Gene blau und Gene, die in allen drei Zelllinien differentiell exprimiert waren schwarz (Annotation) bzw. grau (Expressions-Spots).

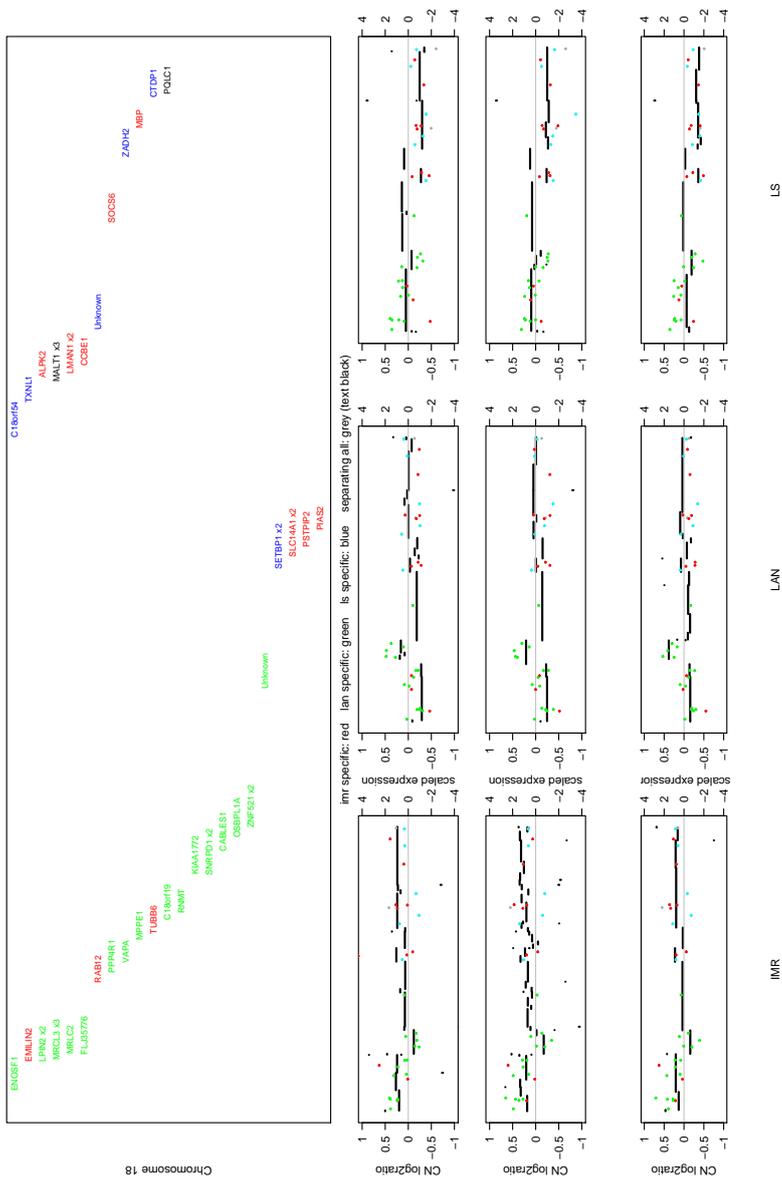


Abbildung B.18: Neuroblastome, Chromosom 18: Projektion der Expressionswerte für die gut korrelierenden Gene (siehe Abschnitt 3.3.2) auf die Segmente der SNP-Analytik entlang Chromosom 18. Die Neuroblastomzelllinien sind in der Reihenfolge IMR 32, LAN5, LS von links nach rechts angeordnet, die drei Samples einer jeden Zelllinie von oben nach unten. In der oberen Hälfte der Grafik findet sich die Annotation der Gene. IMR-spezifische Gene (siehe Abschnitt 3.1.2) sind rot koloriert, LAN-spezifische Gene grün, LS-spezifische Gene blau und Gene, die in allen drei Zelllinien differentiell exprimiert waren schwarz (Annotation) bzw. grau (Expressions-Spots).

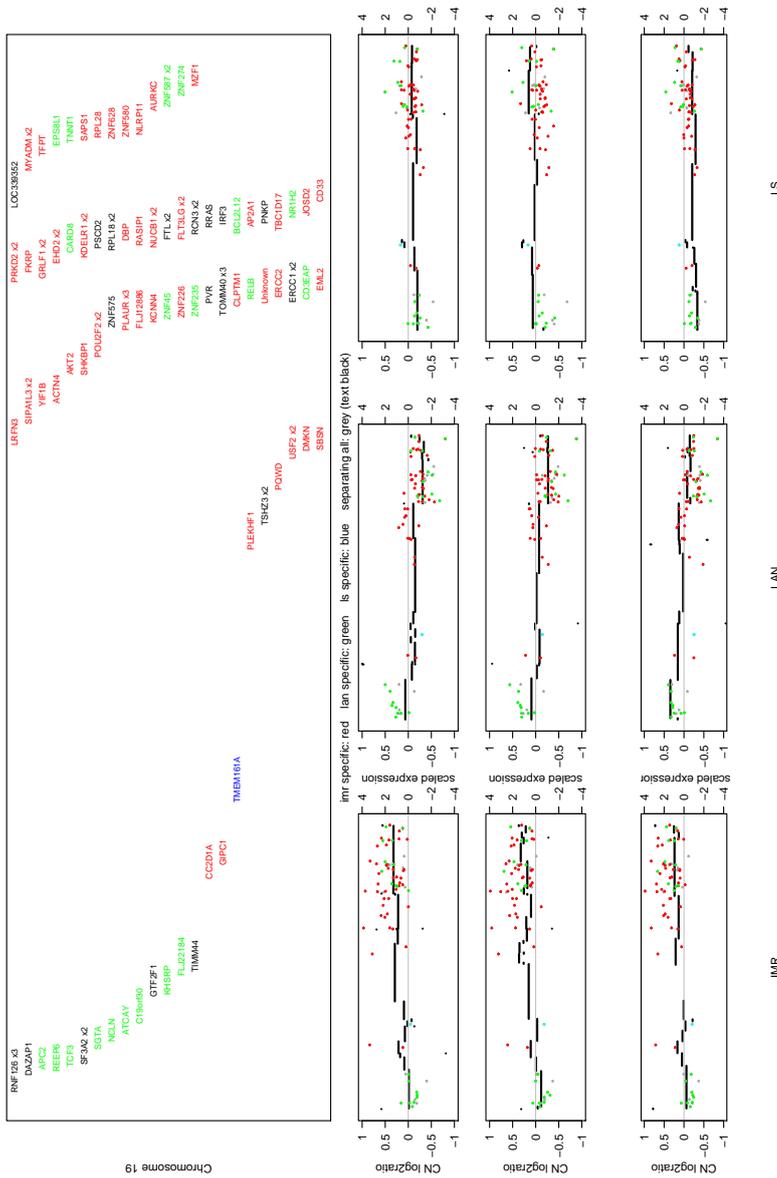


Abbildung B.19: Neuroblastome, Chromosom 19: Projektion der Expressionswerte für die gut korrelierenden Gene (siehe Abschnitt 3.3.2) auf die Segmente der SNP-Analytik entlang Chromosom 19. Die Neuroblastomzelllinien sind in der Reihenfolge IMR 32, LAN5, LS von links nach rechts angeordnet, die drei Samples einer jeden Zelllinie von oben nach unten. In der oberen Hälfte der Grafik findet sich die Annotation der Gene. IMR-spezifische Gene blau und Gene, die in allen drei Zelllinien differentiell exprimiert waren schwarz (Annotation) bzw. grau (Expressions-Spots).

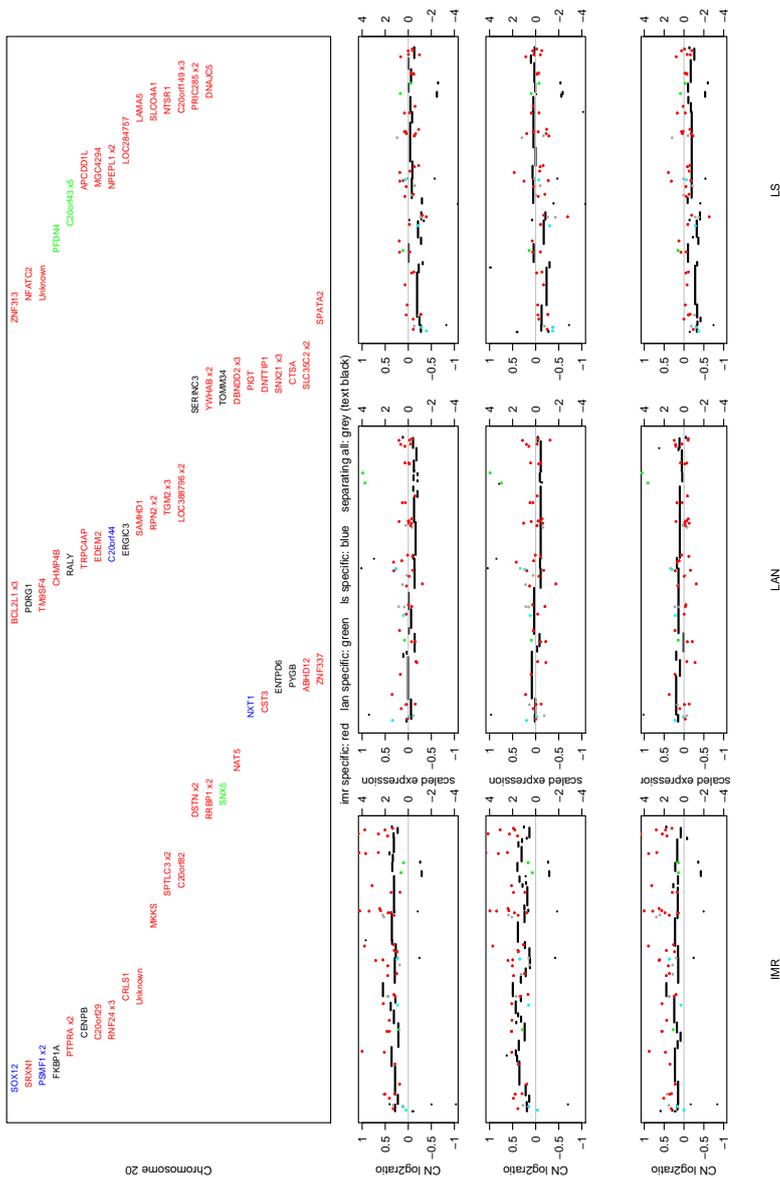


Abbildung B.20: Neuroblastome, Chromosom 20: Projektion der Expressionswerte für die gut korrelierenden Gene (siehe Abschnitt 3.3.2) auf die Segmente der SNP-Analytik entlang Chromosom 20. Die Neuroblastomzelllinien sind in der Reihenfolge IMR 32, LAN5, LS von links nach rechts angeordnet, die drei Samples einer jeden Zelllinie von oben nach unten. In der oberen Hälfte der Grafik findet sich die Annotation der Gene. IMR-spezifische Gene blau und Gene, die in allen drei Zelllinien differentiell exprimiert waren schwarz (Annotation) bzw. grau (Expressions-Spots).

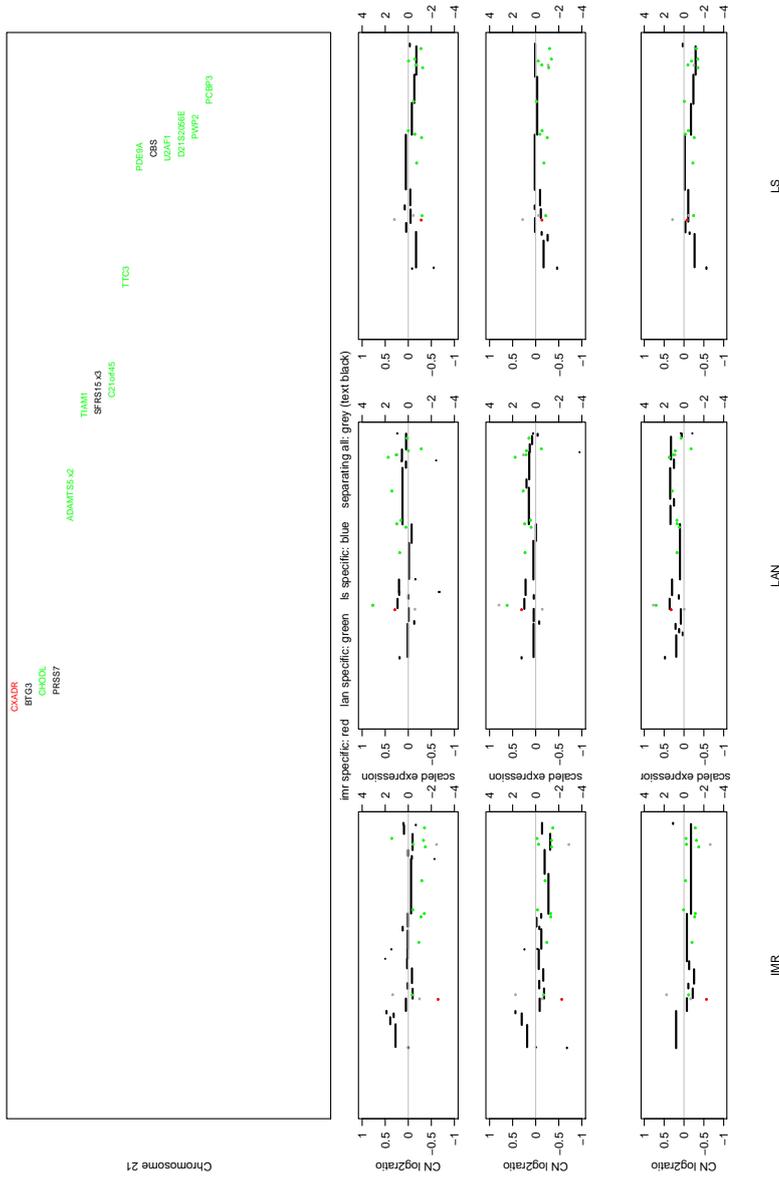


Abbildung B.21: Neuroblastome, Chromosom 21: Projektion der Expressionswerte für die gut korrelierenden Gene (siehe Abschnitt 3.3.2) auf die Segmente der SNP-Analytik entlang Chromosom 21. Die Neuroblastomzelllinien sind in der Reihenfolge IMR 32, LAN5, LS von links nach rechts angeordnet, die drei Samples einer jeden Zelllinie von oben nach unten. In der oberen Hälfte der Grafik findet sich die Annotation der Gene. IMR-spezifische Gene blau und Gene, die in allen drei Zelllinien differentiell exprimiert waren schwarz (Annotation) bzw. grau (Expressions-Spots).

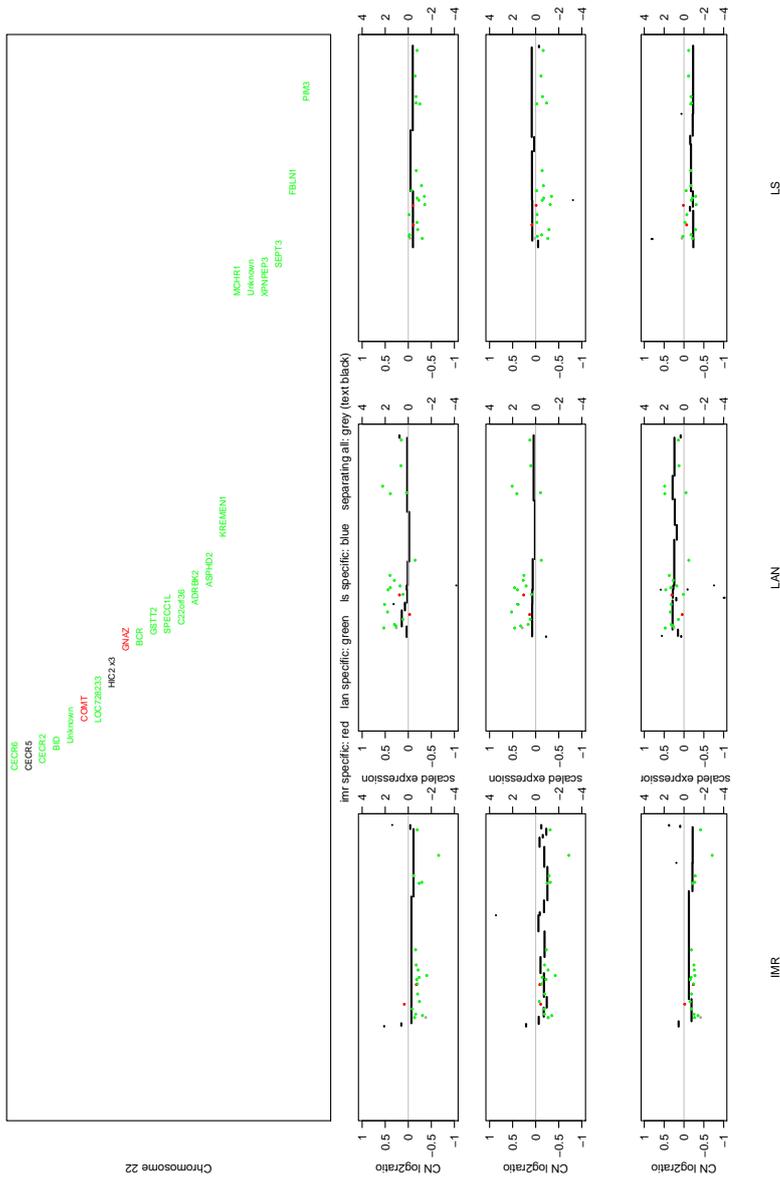


Abbildung B.22: Neuroblastome, Chromosom 22: Projektion der Expressionswerte für die gut korrelierenden Gene (siehe Abschnitt 3.3.2) auf die Segmente der SNP-Analytik entlang Chromosom 22. Die Neuroblastomzelllinien sind in der Reihenfolge IMR 32, LAN5, LS von links nach rechts angeordnet, die drei Samples einer jeden Zelllinie von oben nach unten. In der oberen Hälfte der Grafik findet sich die Annotation der Gene. IMR-spezifische Gene (siehe Abschnitt 3.1.2) sind rot koloriert, LAN-spezifische Gene grün, LS-spezifische Gene blau und Gene, die in allen drei Zelllinien differentiell exprimiert waren schwarz (Annotation) bzw. grau (Expressions-Spots).

ANHANG B. NEUROBLASTOME

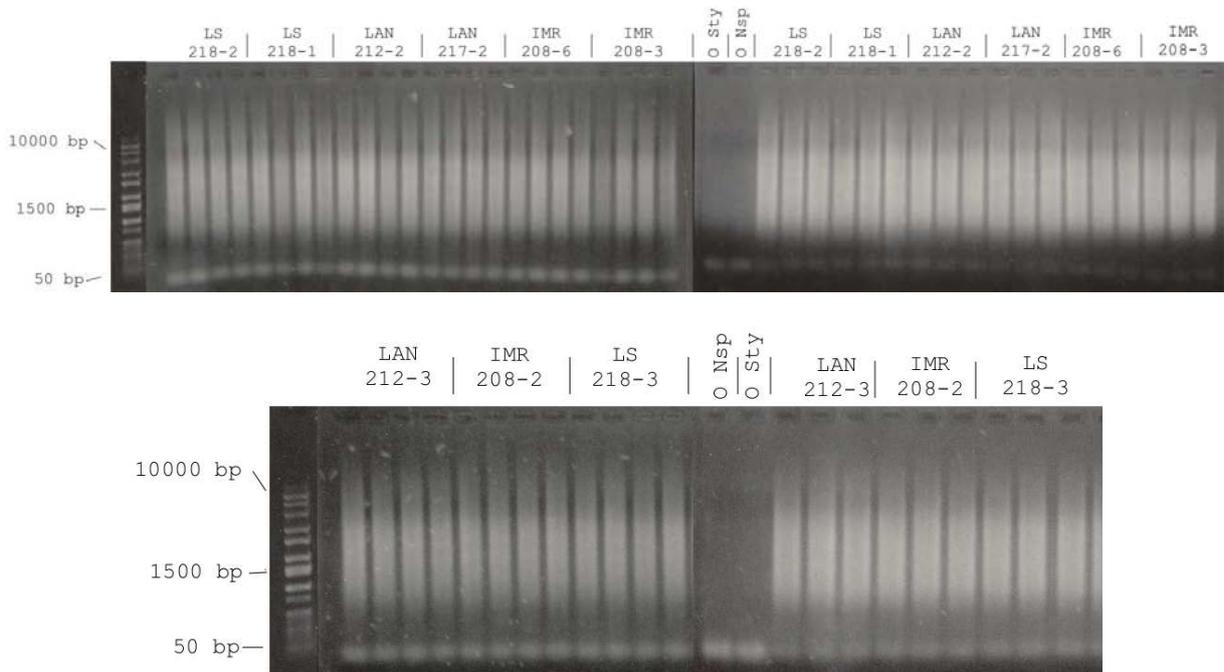


Abbildung B.23: Neuroblastome: Kontrollen auf 2%igem EtBr-Agarosegel nach erfolgter Restriktion mit Nsp I (links von Negativkontrollen) und Sty I (rechts von Negativkontrollen).

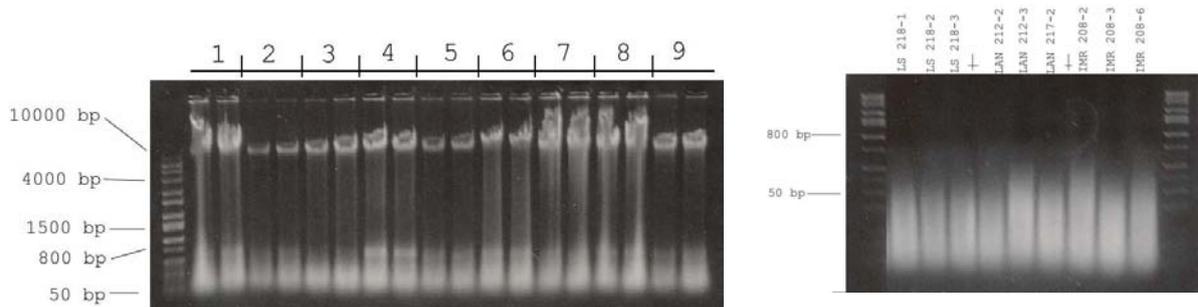


Abbildung B.24: Agarosegel zur Bestimmung der Qualität des isolierten DNA-Materials aus den Neuroblastomproben (links) bzw. zur Kontrolle der Fragmentierung (rechts). Die Gelkonzentration betrug 0,6% (Isolierung) bzw. 2%. (Fragmentierung). Aufgrund des großen Volumens von jeweils 2ml pro Probe wurden diese bei der Isolierung auf je 2 Tubes aufgeteilt. Je zwei benachbarte Bahnen stellen daher hier die Ergebnisse einer Probe dar.

Originalitätserklärung

Hiermit versichere ich ausdrücklich, daß ich die vorliegende Arbeit selbstständig verfaßt und keine anderen als die angegebenen Hilfsmittel und Quellen verwendet habe. Ich erkläre hiermit, daß ich diese Dissertation nicht einem Fachvertreter an einer anderen Hochschule zur Überprüfung vorgelegt oder mich anderweitig um Zulassung zur Promotion beworben habe.

Hamburg, den

Benjamin Otto