

**Identifizierung von Targetgenen des MLL Proteins zur
Aufklärung des Pathomechanismus von Hochrisiko-Leukämien**

Dissertation
zur Erlangung des Doktorgrades
der Naturwissenschaften

vorgelegt beim Fachbereich
Chemische und Pharmazeutische Wissenschaften
der Johann Wolfgang Goethe-Universität
Frankfurt am Main

von
Dorothe Schraets
aus Kerken

Frankfurt am Main, 2003

(DF1)

Vom Fachbereich
Chemische und Pharmazeutische Wissenschaften der
Johann Wolfgang Goethe – Universität in Frankfurt am Main
als Dissertation angenommen.

Dekan: Prof. Dr. W. Müller
1. Gutachter: Prof. Dr. R. Marschalek
2. Gutachter: Prof. Dr. Th. Dingermann

Datum der Disputation: 03.06.2003

meinen Eltern

I Inhaltsverzeichnis

I	Inhaltsverzeichnis	I
II	Abbildungsverzeichnis	V
1	Zusammenfassung	1
2	Einleitung	3
2.1	MLL Struktur und Genprodukt	4
2.1.1	Intrazelluläre Lokalisation	5
2.2	MLL als Regulatorprotein bei Differenzierungsvorgängen	6
2.2.1	Trithorax-Proteinkomplexe (TrxG) im <i>Drosophila</i> System	6
2.2.2	Polycomb-Proteinkomplexe (PcG) im <i>Drosophila</i> System	7
2.2.3	Enhancer of Trithorax and Polycomb (ETP)	9
2.2.4	TrxG/PcG-Antagonismus im murinen System	10
2.2.5	TrxG und PcG Targetgene im murinen System	11
2.2.6	Neueste Daten zur Funktion von MLL in Säugerzellen	12
2.3	MLL und chromosomale Translokationen	16
2.3.1	Hochrisiko-Leukämien	16
2.3.2	t(4;11) Translokation	16
2.3.3	<i>AF-4</i> Gen	20
2.3.3.1	Rekombinationsmechanismen	20
2.4	Biologisches Testsystem	21
2.5	DNA-Mikrochip-Arrays	22
2.6	Proteomanalyse	24
2.6.1	Zweidimensionale Gelelektrophorese	25
2.7	Zielsetzung dieser Arbeit	26
3	Ergebnisse	27
3.1	Etablierung neuer Zelllinien	27
3.1.1	Tetracyclin-induzierbares Expressionssystem in MLL ^{-/-} Zellen	27
3.1.2	Konstitutives Expressionssystem in MLL ^{+/+} und MLL ^{-/-} Zellen	32

3.2	MLL als Differenzierungsprotein in der Embryonalentwicklung / Identifizierung von MLL Targetgenen	34
3.2.1	Morphologische Charakteristika von MLL ^{+/+} und MLL ^{-/-} Zellen	34
3.2.2	Intrazelluläre Lokalisation	36
3.2.3	Subtraktive Kloniertechniken	37
3.2.3.1	Differentielle Genexpression in MLL ^{+/+} Zellen	42
3.2.3.2	Differentielle Genexpression in MLL ^{-/-} Zellen	45
3.2.3.3	Verifizierung identifizierter Gene durch Dot-Blot Hybridisierung	46
3.2.4	DNA-Microchip-Arrays	48
3.2.4.1	Scatter-Blots	50
3.2.4.2	GeneSpring [®] Auswertung	51
3.2.4.3	Differenzielle Genexpression der Wildtyp-Zellen	56
3.2.4.4	Differentielle Genexpression der <i>Mll knock-out</i> Zellen	61
3.3	Vergleichende Proteomanalyse der Zelllinien MLL^{+/+} und MLL^{-/-}	67
3.4	Komplementationsanalyse	74
3.5	Identifizierung von Targetgenen der Derivatproteine	78
3.6	Identifizierung eines artifiziellen Promotors im <i>Mll</i> Gen	81
3.6.1	Charakterisierung des Transkriptionsstarts	81
3.6.2	Intrazelluläre Lokalisation	84
3.6.3	Funktionelle Analyse des Promotorkonstruktes (Intron 11 - Exon 12)	86
3.6.4	Funktionelle Analyse des Promotorkonstruktes (Intron 11 - Exon 13)	88
4	Diskussion	91
4.1	Zellsysteme und Transfektionsexperimente	91
4.2	MLL als Differenzierungsprotein in der Embryonalentwicklung	93
4.2.1	Morphologische Charakteristika	93
4.2.2	Intrazelluläre Lokalisation	94
4.2.3	Identifizierung von Targetgenen des MLL Proteins	95
4.2.3.1	Durch MLL transkriptionell aktivierte Gene	97
4.2.3.2	Durch MLL reprimierte Gene	101
4.3	Proteomanalyse	107
4.4	Komplementationsstudie	110

4.5	Identifizierung der Targetgene der Derivatproteine	112
4.5.1	Hypothetischer Pathomechanismus der t(4;11) Translokation	113
4.6	Identifizierung eines artifiziellen Promotors im <i>Mll</i> Gen	118
5	Material und Methoden	120
5.1	Biologisches Material	120
5.1.1	Zelllinien, Charakteristika und Kulturbedingungen	120
5.1.2	Bakterienstämme	121
5.2	Nukleinsäuren	122
5.2.1	Oligonukleotide	122
5.3	Enzyme und Feinchemikalien	122
5.3.1	Enzyme	122
5.3.2	Feinchemikalien	122
5.4	Methoden	123
5.4.1	Transfektion / Selektion	123
5.4.2	Molekularbiologische Methoden	123
5.4.2.1	PCR-Techniken und Programme	123
5.4.2.2	DNA-Sequenzierung	124
5.4.2.3	Isolierung von mRNA und gesamt-RNA	125
5.4.3	Expressionsanalyse	126
5.4.3.1	Subtraktive Klonierung	126
5.4.3.2	Dot-Blot Hybridisierungsmethode	126
5.4.3.3	Radioaktiv markierte cDNA-Synthese und Hybridisierung	127
5.4.3.4	cRNA-Synthese, DNA-Microchip-Hybridisierung und Datenanalyse	128
5.4.4	Proteomanalyse / 2D-Gele / MALDI-TOF MS	129
5.4.4.1	Isolierung von Plasma- und Kern-Proteinfractionen	129
5.4.4.2	Bestimmung der Proteinkonzentration mittels BCA	130
5.4.4.3	Isoelektrische Fokussierung / Erste Dimension	131
5.4.4.4	SDS-Polyacrylamid-Gelelektrophorese / Zweite Dimension	132
5.4.4.5	Färbemethode	134
5.4.4.6	Detektion differentieller Protein-Spots	134
5.4.4.7	Proteinidentifizierung / MALDI-TOF MS	135
5.4.5	Immunohistochemische Methoden	135
5.4.5.1	Konfokales Laserscanmikroskop	136

6	Anhang	137
6.1	Abkürzungsverzeichnis	137
6.2	Primer-Liste	139
6.2.1	Allgemeine Primer	139
6.2.2	Spezifische murine und humane Primer	140
6.2.3	Primer zur Verifizierung der DNA-Microchip-Daten	141
6.2.4	Primer zur Klonierung der Promotorkonstrukte	141
6.3	Literaturverzeichnis	142
6.4	Expressionsprofil der MLL^{+/+} Zellen	
6.5	Expressionsprofil der MLL^{-/-} Zellen	
6.6	Publikation	
6.7	Lebenslauf	
6.8	Danksagung	

II Abbildungsverzeichnis

- Abb. 2.1 *MLL* Genstruktur
- Abb. 2.2 Schematische Darstellung der homologen Proteine MLL und Trx
- Abb. 2.3 Zusammensetzung und Funktionsweise der TrxG-, PcG_i- und PcG_m-Komplexe
- Tab. 2.4 Zusammenfassung der an den chromatinregulierenden Komplexen beteiligten Proteine
- Abb. 2.5 Darstellung des proteolytisch gespaltenen MLL Proteins
- Abb. 2.6 *MLL* und *AF-4* Genstrukturen
- Abb. 2.7 Proteinstatus einer arretierten pro-B Zelle bei einer ALL (t(4;11))
- Abb. 2.8 Zerstörung eines *Mll* Allels durch homologe Rekombination
- Abb. 2.9 Synthese biotinylierter antisense-RNA
- Abb. 3.1 Schematische Darstellung des Linkers UHD mit seinen Schnittstellen
- Abb. 3.2 Restriktionsverdau der induzierbaren Vektoren
- Abb. 3.3 Expressionskontrolle des Transaktivators von *MLL*^{-/-}::tTA, Klon 3
- Tab. 3.4 Induzierbare Zelllinien und deren Analyse
- Abb. 3.5 RT-PCR eines uninduzierten *MLL*^{-/-}tTA::der11 Klons
- Tab. 3.6 Konstitutiv exprimierende Zelllinien und deren Analyse
- Abb. 3.7 Morphologische Charakteristika von *MLL*^{+/+} und *MLL*^{-/-} Zellen
- Abb. 3.8 Immunfluoreszenzfärbungen verschiedener Zytoskelettproteine
- Abb. 3.9 Nukleäre Lokalisation vom N-Terminus des MLL Proteins
- Abb. 3.10 Schematische Darstellung der subtraktiven Kloniertechnik von *MLL*^{+/+} vs. *MLL*^{-/-}
- Abb. 3.11 Kontroll-PCR zur Bestimmung der Ligationseffizienz
- Tab. 3.12 Zusammenfassung der Zielgene des murinen MLL Proteins
- Tab. 3.13 Zusammenfassung der durch das MLL Protein reprimierten Zielgene
- Abb. 3.14 Ergebnisse der Dot-Blot Hybridisierung
- Abb. 3.15 Integrität und Expression von *Mll* Wildtyp und *knock-out* Allelen
- Abb. 3.16 Scatter-Blot der Expressionsprofile von *MLL*^{+/+} und *MLL*^{-/-} Zellen
- Tab. 3.17 GeneSpring[®] Auswertung der *MLL*^{+/+} und *MLL*^{-/-} Expressionsmuster
- Tab. 3.18 *Heatmaps* von Genexpressionsprofilen der *MLL*^{+/+} und *MLL*^{-/-} Zellen

-
- Abb. 3.19 Expressionsprofil der *Mll* Wildtyp-Zellen
- Abb. 3.20 Expressionsprofil der *Mll knock-out* Zellen
- Abb. 3.21 Notch-1 Signalkaskade in Säugerzellen
- Abb. 3.22 RT-PCR zur Ergebnis-Verifizierung der DNA-Microchip-Arrays
- Abb. 3.23 Zweidimensionale Gelbilder von MLL^{+/+} und MLL^{-/-} Plasma-Proteinfraktionen
- Abb. 3.24 *Master-Gel* (Plasma-Proteinfraktion)
- Abb. 3.25 Zweidimensionale Gelbilder von MLL^{+/+} und MLL^{-/-} Kern-Proteinfraktionen
- Abb. 3.26 Überprüfung der RNA-Qualität
- Abb. 3.27 Graphische Darstellung der Komplementationsergebnisse
- Abb. 3.28 Partielle morphologische Komplementation
- Abb. 3.29 Expressionsanalyse der Transgene in den stabil transfizierten Zelllinien
- Abb. 3.30 Gegenüberstellung der Expressionsmuster von Transfektanden mit dem der Negativkontrolle
- Abb. 3.31 RT-PCR Analysen zur Charakterisierung eines geninternen Transkriptionsstartes
- Abb. 3.32 Charakterisierung des internen Transkriptionsstartes
- Abb. 3.33 Intranukleäre Lokalisation der C-terminalen Region des MLL Proteins
- Abb. 3.34 HE-Linker *in-frame* zum ATG des *EGFP* Gens
- Abb. 3.35 Klonierungsstrategie zur funktionellen Promotoranalyse
- Abb. 3.36 Homologie-Vergleich murines/humanes *Mll/MLL* Gen
- Abb. 3.37 KS-Linker *in-frame* zum ATG des *EGFP* Gens
- Abb. 3.38 Klonierungsstrategie B
- Abb. 4.1 Schematische Darstellung des Wildtyp MLL MPSC
- Abb. 4.2 Hypothetische Funktionsweise des der11 Proteins
- Abb. 4.3 Hypothetische Funktionsweise des reziproken Derivatproteins
- Tab. 5.1 Stabil transfizierte Zelllinien und ihre entsprechenden Selektionsmedien
- Abb. 5.2 Strukturformel RuBPS

1 Zusammenfassung

Die Mehrzahl der akuten B-Zell Leukämien (B-ALL) im Kindesalter kann heutzutage geheilt werden (ca. 80%). Es gibt jedoch eine Untergruppe, die sog. Hochrisiko-Leukämien, der ein anderer Pathomechanismus zu Grunde liegt und für die keine effektive Therapie zur Verfügung steht. Diese Form tritt fast ausschliesslich bei Kleinkindern im ersten Lebensjahr und bei älteren Patienten als Sekundärleukämie nach Chemotherapie auf. Diese akuten Hochrisiko-Leukämien sind zu 80% mit Translokationen des *MLL* Gens, Chromosom 11, Bande q23, assoziiert. Die reziproke Translokation t(4;11), bei der das *MLL* Gen mit dem *AF-4* Gen fusioniert wird, hat die Expression der zwei funktionellen Derivatproteine *MLL*•*AF-4* und *AF-4*•*MLL* und gleichzeitig eine Dosisreduktion des nativen *MLL* Proteins um 50% zur Folge. Das Zusammenspiel dieser Faktoren scheint die Grundvoraussetzung für die pathologische klonale Expansion leukämischer Blasten zu sein.

Aus Untersuchungen in *Drosophila melanogaster* und im Maussystem ist bereits bekannt, dass das Wildtyp *MLL* Protein eine essentielle Funktion in der Steuerung von Genexpression durch Histon- und Chromatinregulation ausübt. Daraus stellte sich als Gegenstand dieser Arbeit die Frage nach der bislang überwiegend unbekanntem Wildtyp-Funktion des *MLL* Proteins, und inwieweit das native Expressionsmuster einer Zelle durch die *MLL* Dosisreduktion beeinflusst, bzw. verändert wird. Zunächst wurden die *MLL* Targetgene identifiziert, und zwar anhand von je zwei DNA-Microchip-Hybridisierungen mit cRNA aus den *MLL*^{+/+} und *MLL*^{-/-} Fibroblasten-Zelllinien. Der Expressionsvergleich dieser Datensätze ergab insgesamt 197 differentiell exprimierte Gene, die sich in der Expressionsstärke um mind. den 2,5-fachen Wert unterscheiden. Davon wurden 136 Gene bei völliger Abwesenheit des *MLL* Proteins, im Vergleich zum Normalzustand der Wildtyp-Zellen, um mind. den 2,5-fachen Wert transkriptionell aktiviert, die übrigen 61 Targetgene um mind. diese Stärke transkriptionell deaktiviert. Die Entdeckung dieser transkriptionell reprimierenden Eigenschaften des *MLL* Proteins, von dem bislang ausschliesslich aktivierende und transkriptionsaufrechterhaltende Eigenschaften bekannt waren, ist eines der wesentlichen Ergebnisse der vorliegenden Arbeit. Ein Teil der, durch die Abwesenheit des *MLL* Proteins hochregulierten 136 Gene ist bereits als Tumormarker bekannt oder an folgenden onkogenen Mechanismen beteiligt: Proliferation durch Fehlsteuerung des Zellzyklus mit gesteigerter Nukleotid-Biosynthese, erhöhte Migrationsaktivität durch veränderte extrazelluläre Matrix mit der Folge von Metastasierung/Organinfiltration, erhöhter Schutz vor pro-

teolytischem Abbau nukleärer (Onko-) Proteine, und der Generation von Spleiss-Varianten mit z.T. negativem Einfluss auf essentielle Differenzierungswege. Zu diesen, in der Summe das Krebsrisiko erhöhenden Effekten, kommt noch hinzu, dass eine Gegenregulation durch MLL induzierte Expression von Tumorsuppressoren (verschiedene Zellzyklus-Inhibitoren) fehlt. Da bei Leukämie-Zellen die MLL Proteindosis reduziert ist, liegt der Schluss nahe, den oben genannten 136 identifizierten Genen eine mögliche direkte Beteiligung an der Leukämogenese beizumessen.

Die meisten der 61 Gene, die durch das MLL Protein transkriptionell aktiviert wurden, kodieren für Faktoren, die zum größten Teil in embryonale Differenzierungsprozesse involviert sind. Dabei spielt die Entwicklung von meso- und ektodermalen Geweben eine besondere Rolle. Das MLL Protein ist somit für die Organogenese von Herz, Leber, Nieren, sensorischen Organen, hämatopoietischen Zellen und ebenso für Knochen und Muskeln essentiell.

Die gewonnenen Daten wurden durch verschiedene Experimente (subtraktive Klonierung und RT-PCR) verifiziert und die wenigen, bereits publizierten MLL Targetgene konnten durch diese Arbeit bestätigt werden.

Neben den Targetgenen des MLL Proteins sollten auch diejenigen der beiden Derivatproteine MLL•AF-4 und AF-4•MLL identifiziert werden. MLL^{+/+} Zellen wurden mit den humanen Derivatkonstrukten stabil transfiziert und im Vergleich zur Leervektor-Kontrolle per DNA-Microchips analysiert. Wieder Erwarten wurden keine unterschiedlichen Transkriptionsmuster erhalten. Daraufhin wurden Komplementationsexperimente zur Eignungsüberprüfung des Testsystems durchgeführt. Es stellte sich heraus, dass das humane *MLL* Expressionskonstrukt die MLL^{-/-} Zellen funktionell nicht komplementieren konnte. Dafür gibt es verschiedene Erklärungsansätze, wobei jedoch die Hypothese, dass das „transkriptionelle Gedächtnis“ in den murinen embryonalen Fibroblasten (Entwicklungsstatus Tag 10,5 p.c.), bereits stabil etabliert und nur noch marginal veränderbar ist, die Wahrscheinlichste ist. Sollte sich herausstellen, dass die epigenetische Programmierung der Zellen verantwortlich für die hier erhaltenen Ergebnisse ist, hätte das einen dramatischen Einfluss auf unser Verständnis vom Leukämie-Pathomechanismus. Es würde nämlich bedeuten, dass MLL und davon abgeleitete reziproke Derivatproteine nur in einem engen Zeitfenster Einfluss auf Genexpressionsmuster haben und nach einen „hit and run“ Mechanismus die Leukämie auslösen.

2 Einleitung

Leukämien sind die am häufigsten vorkommenden Krebserkrankungen im Kindesalter. Durch den Einsatz von Chemotherapie und die ständige Weiterentwicklung der Behandlungsprotokolle haben sich die Überlebensaussichten für Kinder mit einer akuten Leukämie in den vergangenen Jahren entscheidend verbessert. Es gibt jedoch eine Untergruppe der akuten B-Zell Leukämien (B-ALL) mit äußerst geringen Heilungschancen. Diese sogenannten „Hochrisiko-Leukämien“ treten überwiegend bei Kindern unter einem Jahr auf und zeigen einen anderen Pathomechanismus als die gut heilbaren Formen. In den meisten Fällen kann eine erste Remission erreicht werden, auf die jedoch schnell ein Rezidiv folgt. Die leukämischen Blasten, die das Rezidiv verursachen, sind refraktär gegen weitere Therapiekonzepte und die Prognose ist fatal. Zytogenetische Untersuchungen der arretierten pro-B Zellen haben zu der Erkenntnis geführt, dass es einen Zusammenhang zwischen einer Chromosomenaberration des *MLL* Gens auf dem langen Arm des Chromosoms 11, Bande q23 und dem Auftreten von Neoplasien des hämatopoietischen Systems gibt.

Nachdem die Korrelation zwischen den genetischen Veränderungen im *MLL* Gen und dem Phänotyp der Hochrisiko-Leukämie erkannt wurde, stieg das Interesse an diesem Gen. Northern-Blot Analysen zeigten die Expression von *MLL* in fast allen adulten Geweben: Knochenmark, Leber, Niere, Herz, Muskeln, Lunge, Gehirn, Plazenta und Hoden, wobei der Expressionslevel im Gehirn weitaus am höchsten liegt. Bisher ist jedoch über die Wildtyp-Funktion des *MLL* Proteins auf molekularer Ebene wenig bekannt. Der Funktionsverlust beider *Mll* Allele (*Mll*^{-/-}) erwies sich im transgenen Mausexperiment als embryonal letal, während ein heterozygot vorliegender Phänotyp (*Mll*^{+/-}) zu einer Veränderung des Expressionsmusters verschiedener Homeobox Gene führte, die sich durch Skelettdeformationen, Wachstumsretardierung und eine beeinträchtigte Hämatopoiese auszeichnete (Yu et al., 1995). Das *MLL* Protein gehört zu der Familie der Trithorax-Proteine und ist an der Aufrechterhaltung von transkriptionellen Vorgängen in der Zelle beteiligt („transkriptionelles Gedächtnis“). *MLL* scheint somit erheblichen Einfluss auf Chromatin-Modulation und ebenfalls Transkriptionsaktivierung in Entwicklungs- und Differenzierungsprozessen während der Embryogenese auszuüben.

2.1 *MLL* Struktur und Genprodukt

Der offene Leserahmen des *MLL* Gens (~ 100 kb) auf Chromosom 11, Bande q23, wird durch 37 Exons kodiert (Nilson et al., 1996). Neben dem am häufigsten vorkommenden Transkript mit einer Länge von 15,5 kb können weitere Spleiss-Produkte mit Längen von 13,5 kb und 5 kb detektiert werden, deren Funktion jedoch noch weitgehend unbekannt ist.

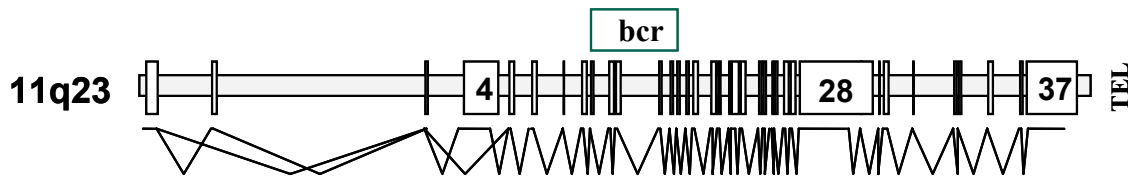


Abb. 2.1: *MLL* Genstruktur

Exon-Intron Struktur des *MLL* Gens mit möglichen Spleiss-Varianten; bcr: Bruchpunkt-Cluster-Region; TEL: Telomer-Seite.

Das *MLL* Gen kodiert für ein Protein mit maximal 4005 Aminosäuren und damit einem Molekulargewicht von 435 kDa (Nilson et al., 1996). Die Proteinsequenz zeigt eine Identität von 24% zu Trithorax (Trx), einem wichtigen zur Trithorax Proteingruppe (TrxG) zählenden Differenzierungsprotein aus *Drosophila melanogaster*. Neben den Aufgaben zur Chromatin-Modulation konnte gezeigt werden, dass bestimmte Domänen des *MLL* Proteins in transienten Transfektionsexperimenten die Transkription von Reporterkonstrukten beeinflussen (Prasad et al., 1995).

Das *MLL* Protein (Mixed Lineage Leukemia, Ziemin-van der Poel et al., 1991; HRX: human Trithorax, Tkachuk et al., 1992; ALL-1: Acute Lymphoblastic Leukemia-1, Gu et al., 1992) besteht aus einer Vielzahl konservierter Domänen und ist modular aufgebaut. Dazu gehören neben einer Transkriptionsaktivator-Domäne und Transkriptionsrepressor-Domäne verschiedene DNA-Bindungsmotive (AT-Haken und eine Methyltransferase-Domäne) die auf eine mögliche Funktion als DNA-bindendes Protein hinweisen (Ennas et al., 1997; Zeleznik-Le et al., 1994). Ein Cystein-/Histidin-reicher Proteinabschnitt kann in 10 klassische Zinkfinger gefaltet werden, oder in eine Struktur, die als PHD (Plant Homeobox Domain) bezeichnet wird. Diese ist an Kondensations- und Dekondensationsprozessen des Chromatins beteiligt (Aasland et al., 1995) und wurde zuerst bei verschiedenen Pflanzenproteinen entdeckt (Schindler et al., 1993). In immunohistochemischen Experimenten konnte die Kernlokalisation, die sich ebenfalls aus der Sequenz (snLS) ableiten läßt, bestätigt werden (Ennas et al., 1997; Yano et al., 1997).

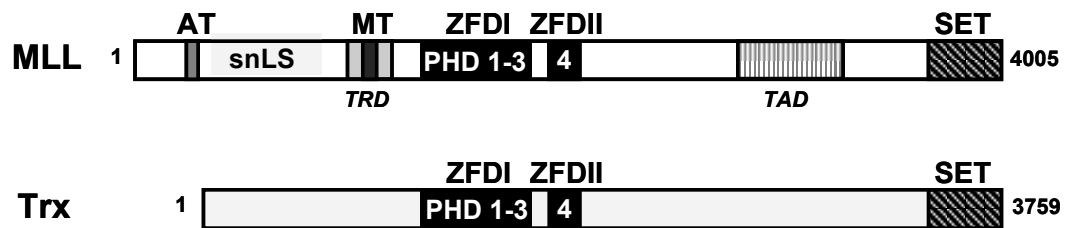


Abb. 2.2: Schematische Darstellung der homologen Proteine MLL (human) und Trx (*Drosophila*).

AT: AT-Haken, snLS: subnukleäre Lokalisationssequenz, MT: Methyltransferase-Domäne, TRD: Transkriptionsrepressor-Domäne, ZFD: Zinkfinger-Domäne, PHD: Plant Homeobox Domain, TAD: Transkriptionsaktivator-Domäne, SET: SET-Domäne.

Für einige Proteinmotive konnte bereits die Fähigkeit zur Interaktion mit anderen Kernproteinen nachgewiesen werden (Adler et al., 1997; Nakamura et al., 2002). Am C-Terminus befindet sich eine stark konservierte Homologie-Domäne mit dem Trx Protein, die sogenannte SET-Domäne (Mozer and Dawid, 1989). Sie wurde ebenfalls in den chromatinmodifizierenden Proteinen Suppressor of variegation (Su(var)), Enhancer of zeste (E(z)) und Trithorax (Trx) entdeckt, die in Transkriptionsaktivierung und –aufrechterhaltung u.a. der Homeobox Gen-Cluster während der Embryonalentwicklung von *D. melanogaster* involviert sind. Für diesen Bereich des MLL Proteins wurden u.a. Proteine des SWI/SNF-Komplexes als Interaktionspartner gefunden, die für Chromatin-Remodulierungsprozesse verantwortlich sind (Rozenblatt-Rosen et al., 1998). Darüber hinaus konnte eine enzymatische Aktivität der SET-Domäne beobachtet werden. Die Aminosäuresequenz im Bereich AS 3745-3969 besitzt eine Histon-Methylaseaktivität, die spezifisch den Lysin-Rest an Position 4 des Histons 3 methyliert (H3K4; Nakamura et al., 2002).

2.1.1 Intrazelluläre Lokalisation

Immunohistochemische Experimente mit polyklonalem E28 Serum (Epitop: Exon 28) bestätigten die durch die Proteinstuktur (snLS) vermutete Kernlokalisierung des MLL Proteins, das gleichmäßig über den Kern verteilt vorliegt (Ennas et al., 1997). In ruhenden Zellen wurde MLL in der inneren Hemisphäre des Zellkerns nachgewiesen. Da sich dort aktivierte Chromatinbereiche befinden, stützt dieses Erkenntnis die erwartete und auch schon publizierte Funktion des MLL Proteins als transkriptionsregulierenden Faktor. In der

Mitose findet man das MLL Protein direkt an den Metaphase-Chromosomen in einem Kompartiment, das als Chromosomenhülle („*chromosomal envelope*“) bezeichnet wurde. MLL ist vermutlich unstringent mit dem Chromatin assoziiert und nicht fest an die DNA gebunden (Ennas et al., 1997).

2.2 MLL als Regulatorprotein bei Differenzierungsvorgängen

Die Identität verschiedener Zellspezies wird durch früh definierte Muster von aktiven und inaktiven Chromatinabschnitten charakterisiert. Sie führen zu spezifischen Expressionsmustern, die Zellspezifität während Mitose (und Meiose) sicherstellen und in stabiler Form an die Tochterzellen weiter vererbt werden. Dieses Phänomen wird als „transkriptionelles Gedächtnis“ bezeichnet und über spezifische Histon- und DNA-Methylierungsmuster definiert. Verschiedene Zustände des Chromatins werden dabei durch zwei antagonistisch agierende Proteinfamilien aufrechterhalten. Die Proteine der Trithorax-Gruppe (TrxG), zu denen auch das Trx/MLL Protein zählt, gehören generell zu den transkriptionsaktivierenden und -aufrechterhaltenden Proteinen, während die Proteine der Polycomb-Gruppe (PcG) für die stabile Vererbung von reprimierten Chromatinabschnitten verantwortlich sind. Diese, zu jedem Zeitpunkt in der Entwicklung spezifischen Expressionsmuster müssen fein reguliert werden, da schon ein geringes Ungleichgewicht zwischen den Transkriptionssteuerungskomplexen (TrxG und PcG) zu fatalen Entwicklungsstörungen führt (Hanson et al., 1999; Yu et al., 1995). Das Verständnis über die Interaktionen zwischen TrxG- und PcG-Proteinkomplexen scheint bei der Suche nach Grundlagen von krankheitsassoziierten Veränderungen von Genexpressionsmustern relevant zu sein.

2.2.1 Trithorax-Proteinkomplexe (TrxG) im *Drosophila* System

Durch Forschungsarbeiten an *D. melanogaster* konnte die Funktion des Trx Proteins (*Drosophila* Homolog zu MLL) teilweise aufgeklärt werden. Die bis zu 2 MDa großen Trithorax-Proteinkomplexe sind aus vielen Proteinen mit unterschiedlichsten Funktionen zusammengesetzt. Darunter befinden sich z.B. DNA-bindende Komponenten (Trx, GAGA Faktor und Pleiohomeotic), Histon-Acetyltransferasen und chromatinmodifizierende Proteine (z.B. Snr1), die der stark konservierten SWI/SNF-Transkriptionsmaschinerie angehören. TrxG-Komplexe weisen jedoch nicht immer die gleiche Zusammensetzung auf. Die

Zusammensetzung dieser Komplexe ist entscheidend für die Bindung an die unterschiedlichen Chromatinabschnitte. Diese Komplexe binden an ca. 100 im *D. melanogaster* Genom vorliegende TREs (Trithorax Response-Elements), die durch immunohistochemische Experimente am Polyän-Chromosom lokalisiert wurden (Tillib et al., 1999).

2.2.2 Polycomb-Proteinkomplexe (PcG) im *Drosophila* System

An der Ausbildung von multimeren PcG-Proteinkomplexen (2-5 MDa) sind viele stark konservierte Proteine beteiligt, die neben ihrer Funktion als DNA-bindende und chromatinmodulierende Komponenten auch für die strukturelle Integrität der Komplexe von Bedeutung sind. Mit Hilfe von Experimenten an *D. melanogaster* und Säugerzellen konnte nachgewiesen werden, dass zwei unterschiedliche Arten von Komplexen mit verschiedenen Funktionen existieren.

Der PcG_i-Komplex (Polycomb-Group Initiation Complex) besteht hauptsächlich aus den Proteinen Extra sex combs (*esc*), Enhancer of zeste (*E(z)*), GAGA-Faktor, Pleiohomeotic (*pho*) und Histon-Deacetyltransferasen und wird von dem frühen Repressor Hunchback (*Hb*) in der Repressionsinitiationsphase rekrutiert. Durch eine lokale Histon-Deacetylierung wird die Bindung des zweiten Komplexes (PcG_m; Polycomb-Group Maintenance Complex) ermöglicht. Dieser wird u.a. aus den Proteinen Polycomb (*pc*), Posterior sex combs (*psc*) und Polyhomeotic (*ph*) gebildet. Er löst PcG_i in seiner Funktion ab und sichert eine stabile Aufrechterhaltung inaktiver Chromatinstrukturen über viele Zellzyklen, oder bis zur Verdrängung durch TrxG-Komplexe. Mutationen in den für PcG_i kodierenden Genen sind embryonal letal, während sie in Genen, die für PcG_m kodieren, (nur) zu Fehlbildungen im Körperbauplan führen.

Beide Komplexe binden an die gleichen PREs (Polycomb Response-Elements). Hierbei handelt es sich um spezifische Sequenzen, die ca. 30-40 bp neben den TREs und bis zu 20 kb vor den entsprechenden Promotorsequenzen liegen. Aufgrund der dualen Funktion dieses gesamten Bereiches wurde vorgeschlagen, diese Region als „Maintenance-Element“ (ME) zu bezeichnen. Antigene Strukturen verschiedener Proteine aus dem PcG-Komplex wurden mit denen der TrxG kolokalisiert (Castelli-Gair and Garcia-Bellido, 1990). Diese Daten lassen vermuten, dass beide Proteingruppen über einen koordinierten Mechanismus den Zustand des Chromatins und damit die zelluläre Transkription von vielen verschiedenen Zielgenen steuern können.

Abbildung 2.3 zeigt schematisch die Zusammensetzung von *D. melanogaster* TrxG-, PcG_i- und PcG_m-Komplexen. Der in Vertebraten vorkommende MLL MPSC (Multiprotein-Superkomplex; Nakamura et al., 2002) beinhaltet andere Faktoren und eine Vielzahl weiterer Proteine, die in Kap. 2.2.6 noch genauer betrachtet werden.

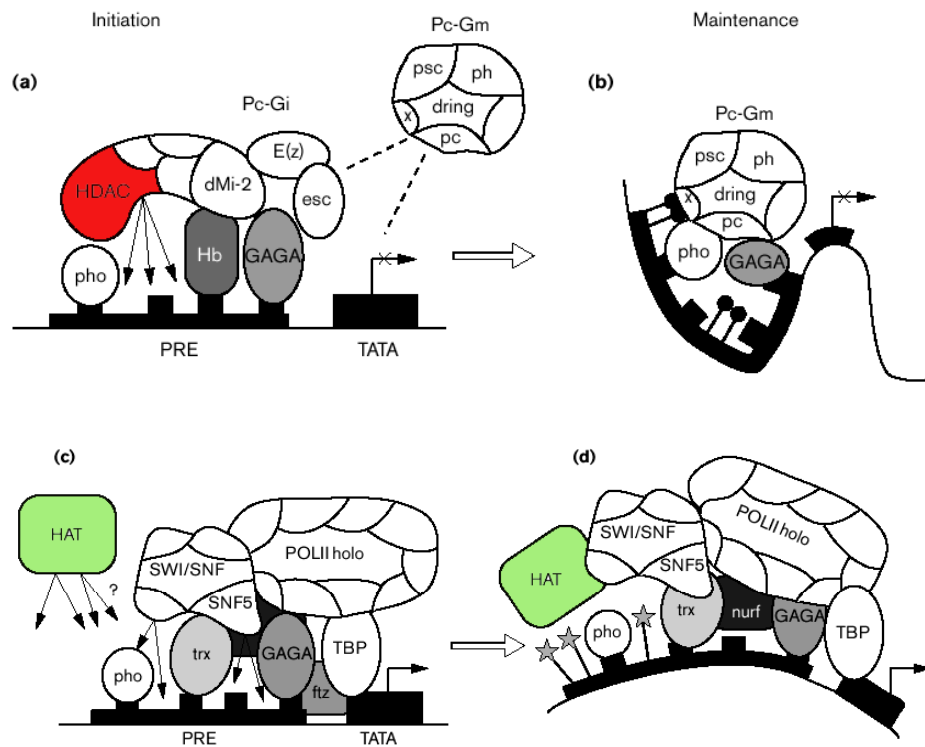


Abb. 2.3: Zusammensetzung und Funktionsweise der TrxG-, PcG_i- und PcG_m-Komplexe

Hb: Hunchback, pho: Pleiohomeotic, GAGA: DNA-bindender Faktor, E(z): Enhancer of zeste, esc: Extra sex combs, dMi-2: Drosophila Homolog von Mi-2, HDAC: Histon-Deacetyltransferase, psc: Posterior sex combs, ph: Polyhomeotic, pc: Polycomb, HAT: Histon-Acetyltransferase, SWI/SNF: Chromatin-Remodulierungsmaschinerie, trx: Trithorax. **(a)**: Repressionsinitiation, Hb rekrutiert HDAC/dMi-2 und im Anschluß weitere PcG_i-Faktoren, nach lokaler Histon-Deacetylierung bindet PcG_m. **(b)**: PcG_m Bindung an PREs führt zur Chromatin-Konformationsänderung und stabilisiert den reprimierten Gen-Lokus. **(c)**: Bindung von HAT und TrxG an „ehemaliges“ PRE, Verdrängung des PcG_m-Komplexes, Histon-Acetylierung und Rekrutierung weiterer Trx-Faktoren. **(d)**: Transkriptionsmöglichkeit durch Histon-Acetylierung und folgender Chromatin-Remodulierung. Darstellung: Transkriptionssteuerung in *D. melanogaster* (van Lohuizen, 1999).

2.2.3 Enhancer of Trithorax and Polycomb (ETP)

Neueste Forschungsergebnisse haben gezeigt, dass einige Mitglieder der PcG auch zur Aufrechterhaltung von aktivierten Genabschnitten wichtig sind: Posterior sex combs (*psc*) wird z.B. für die Transkriptionsaktivierung des *Polyhomeotic (ph)* Gens benötigt. Sie binden an entsprechende Promotor-/Enhancer-Elemente, ohne ihre charakteristischen reprimierenden Eigenschaften auszuüben. Im transgenen Mausmodell konnte gezeigt werden, dass Mutationen in sechs verschiedenen PcG Genen (z.B. *Bmi-1*, *Enx1*) eine reduzierte Homeobox Genexpression und somit den Phänotyp von *Trx* Mutanten zur Folge hatten. Diese Ergebnisse lassen die Vermutung zu, dass ganz bestimmte PcG Proteine an der Aktivierung und Repression mancher Hox Gene beteiligt sind. Aufgrund ihrer dualen Funktion wurden sie als „Enhancer of Trithorax and Polycomb“ (ETP) definiert und sind in Tabelle 2.4 zusammengefasst (Brock and van Lohuizen, 2001). Dazu wurde eine Modellvorstellung entwickelt bei der davon ausgegangen wird, dass die ETP Proteine zusammen eine Untereinheit bilden, deren Expressionslevel sehr fein reguliert sein muss. Der vermutete Diskriminierungsmechanismus könnte durch die dosisabhängige Interaktion des ETP-Komplexes entweder mit *TrxG* oder PcG zustande kommen.

Proteinkomplex	Drosophila Gene	humane Homologe	murine Homologe
PcG	<i>Extra sex combs (esc)</i>	<i>EED</i>	<i>eed</i>
	<i>Pleiohomeotic (pho)</i>	<i>hYY1</i>	<i>mYY1</i>
	<i>Polycomb (pc)</i>	<i>M33 (HPC1), HPC2/3</i>	<i>M 33, mPc2</i>
	<i>Polyhomeotic (ph)</i>	<i>HPH1, HPH2</i>	<i>Mph1/rae28, Mph2</i>
	<i>Multi sex combs (msc)</i>	/	/
ETP	<i>Posterior sex combs (psc)</i>	<i>BMI-1, MEL 18</i>	<i>Bmi-1, Mel 18</i>
	<i>Enhancer of zeste ((E)z)</i>	<i>EZH1, EZH2</i>	<i>Enx1, Enx2</i>
	<i>Additional sex combs</i>	<i>ASXL1, ASXL2</i>	<i>Ascl1, Asxl2</i>
	<i>Enhancer of Polycomb</i>	<i>EPC1, EPC2</i>	<i>Epc1, Epc2</i>
	<i>Sex combs on midleg</i>	<i>SCML1, SCML2</i>	<i>Scml1, Scml2</i>
	<i>Suppressor of zeste 2</i>	<i>BMI-1, MEL 18</i>	<i>Bmi-1, Mel 18</i>
TrxG	<i>Trithorax (trx)</i>	<i>MLL, ALL-1, HRX</i>	<i>Mll</i>
	<i>Brahma (brm)</i>	<i>hBRM, BRG1</i>	<i>mbrm1, mBrg1</i>
	<i>Snr1</i>	<i>hSNF5</i>	<i>mSNF5</i>

Tab. 2.4: Zusammenfassung der an den chromatinregulierenden Komplexen beteiligten Proteine

Bisher ist nur das PcG Protein Pleiohomeotic in TrxG-Komplexen identifiziert worden. Neben diesem scheint das TrxG Protein Trithoraxlike (Trl) nicht ausschliesslich für Transkriptionsaktivierung, sondern ebenfalls für die Aufrechterhaltung von reprimierten Gen-Loci eine Rolle zu spielen. *Trl* kodiert für den allgemeinen Transkriptionsfaktor GAGA, der als Genaktivator vieler *Drosophila* Gene, darunter auch Homeobox Gene, von Bedeutung und mit PREs assoziiert ist (*in vivo* Daten). Anhand dieser Ergebnisse müsste *Trl* ebenfalls als ETP klassifiziert werden.

2.2.4 TrxG/PcG-Antagonismus im murinen System

Im transgenen Maussystem konnte die antagonistische Wirkweise der TrxG- und PcG-Komplexe experimentell belegt werden. Heterozygote *Mll* Mutanten (*Mll*^{+/-}) führten ebenso wie *Bmi-1*^{+/-} Mutanten (PcG Protein) durch deregulierte Hox Genexpression zu einem stark veränderten Phänotyp: abnormale Segmentidentität (Skelettdeformationen), Wachstumsretardierung und eine beeinträchtigte Hämatopoiese (Fidanza et al., 1996). Mäuse, die gleichzeitig beide Mutationen aufwiesen, zeigten als Doppel-Heterozygoten (*Mll*^{+/-} und *Bmi-1*^{+/-}) keinerlei auffälligen Phänotyp und schienen gesund zu sein (Hanson et al., 1999). Aus diesen Daten lässt sich die Aussage ableiten, dass es sich bei der Steuerung von Genexpression um dosisabhängige Effekte der Regulatorkomplexe an den Zielstrukturen handelt. Ein Ungleichgewicht zwischen TrxG und PcG hat fatale Folgen für den Organismus und führt u.a. zu erheblichen Missbildungen im Körperbauplan während der Embryonalentwicklung. Dieser Mechanismus des „Gendosis-Effekts“ zeigt, dass minutiöse Veränderungen in der Expression dieser Proteine dramatische Folgen haben können.

Die kombinierte Forschung an TrxG- und PcG-Komplexen in *D. melanogaster*, Mäusen, transformierten Zelllinien und in Tumorzellen ist zur Funktionsaufklärung und für das Verständnis von Transkriptionsmechanismen/-regulation und der Ausbildung eines „transkriptionellen Gedächtnisses“ von größter Bedeutung. Aus diesen Erkenntnissen lassen sich möglicherweise Korrelationen zwischen Chromatin-Deregulation und der Entstehung und Entwicklung verschiedenster Erkrankungen ableiten. Die bisherigen Daten legen die Vermutung nahe, dass aus einer Gleichgewichtsverschiebung zwischen TrxG- und PcG-Proteinkomplexen verschiedene potentielle Erkrankungen oder Fehlbildungen resultieren, bzw. dass sie direkt zur Tumorentwicklung führen könnten.

2.2.5 TrxG und PcG Targetgene im murinen System

Aufgrund der bisher vorliegenden Forschungsergebnisse scheint das MLL Protein eine zentrale Funktion in der Ausbildung von TrxG-Komplexen und folglich auf Entwicklungsprozesse zu haben. Die Identifizierung von MLL Zielgenen (direkte und indirekte) ist deshalb für das Verständnis von Differenzierungsmechanismen von großer Bedeutung. Durch Kenntnis der MLL Zielgene würden möglicherweise auch weiterführende Hinweise gewonnen, die zur Aufklärung des leukämogenen Transformationsprozesses beitragen könnten. Hierzu gibt es zwei prinzipielle Denkansätze: durch chromosomale Translokationen des *MLL* Gens entsteht eine Gendosis-Erniedrigung. Zum anderen darf man potentielle Interaktionen und damit Interferenz aberranter MLL Fusionsproteine mit den Funktionen der TrxG- und PcG-Komplexe nicht ausser Acht lassen.

Evolutionär konservierte Hox Gene wurden bereits als MLL Targetgene identifiziert. Die 37 Hox Gene in Säugern sind in 4 Cluster (A, B, C und D) eingeteilt. Gene, die am 3'-Ende eines jeden Clusters kodieren, sind während der Embryonalentwicklung für die rostrale Region des Embryos wichtig. Die zentralen Gene sind für den Thorax-Bereich und die Gene am 5'-Ende für den kaudalen Bereich des Embryos relevant. Alle Hox Gene kontrollieren die Segmentidentität entlang der Körperachse, wobei die Etablierung der Expressionsmuster zum richtigen Entwicklungszeitpunkt im richtigen Segment des Embryos für die korrekte Ausbildung des Körperbauplans und der Organe wichtig sind. Durch die Dosisreduktion des MLL Proteins in den heterozygoten Mäusen kommt es zu einer aberranten Hox Genexpression. Da die Targetgene des MLL Proteins den Gruppen 7-9 angehören, sind sie maßgeblich an der thorakalen Stukturausbildung beteiligt. Ein Nachweis wurde durch Untersuchungen an *Mll*^{+/-} Mäusen erbracht, deren Skelett verschiedene Deformationen oder Fehler aufwies (deformiertes Brustbein, fehlender Thorax-Wirbel oder zusätzlicher Lendenwirbel; Yu et al., 1995).

Die Transkriptionsaktivierung muriner *hox a7*, *hox c8* und *hox c9* Gene erfolgt initial zwischen den Tagen 7.5 und 8.5 p.c., gefolgt von einer späteren Phase der gewebespezifischen Aufrechterhaltung, die durch TrxG-Komplexe bewirkt wird. Um die Frage zu beantworten, ob MLL eher an Transkriptionsaktivierung oder -aufrechterhaltung beteiligt ist, wurde das *hox a7* Expressionsmuster analysiert (Yu et al., 1998). Bis zum Tag 9 p.c. wurde kein Unterschied in der Expressionsstärke zwischen *Mll*^{+/+} und *Mll*^{-/-} Mäuseembryos im Neuroektoderm und Mesoderm detektiert. Ab Tag 9.5 p.c. konnte jedoch in den *knock-out*

Zellen kein Hox a7 mehr nachgewiesen werden, während in den *Mll*^{+/+} Embryos das korrekte Expressionsmuster des *hox a7* Gens manifestiert wurde. Durch diese Versuche wurde der Zeitpunkt identifiziert, ab dem die *hox a7* Genexpression absolut von der Anwesenheit des MLL Proteins abhängig wird. Dieser Zeitpunkt reflektiert eine dynamische Periode der Chromatin-Reorganisation, die entweder zur Genaktivierung oder –reprimierung führt.

Durch funktionelle Analysen von *Bmi-1 knock-out* Mäusen (*Bmi-1*^{-/-}) wurde neben den Hox Genen der *INK4a* Locus als weiteres Zielgen entdeckt, der für die beiden Tumorsuppressoren und Zellzyklus-Inhibitoren p16 und p19^{ARF} kodiert. *Bmi-1* Mangel (homolog zu Posterior Sex Combs (Psc); van Lohuizen et al., 1991) induziert eine p16 und p19^{ARF} Überexpression, während *Bmi-1* Überexpression zum Mangel an den Tumorsuppressoren p16 und p19^{ARF} und in transgenen Mäusen zu Lymphomen führt (Jacobs et al., 1999). Daraus resultiert die Fragestellung, ob durch eine MLL Translokation, bei der die Zerstörung eines *MLL* Allels zu einer Dosisreduktion an nativem Protein führt, das Gleichgewicht in Richtung PcG-Komplexe verschoben wird. Dies würde einer *BMI-1* Überexpression entsprechen und zu einer transkriptionellen Inaktivierung des *INK4a* Locus führen. Da p16 und p19^{ARF} als wichtige Regulatoren des Zellzyklus bekannt sind und der p53 Tumorsuppressor-Kaskade übergeordnet sind, scheinen PcG-Faktoren in der Lage zu sein, den Zellzyklus aktiv zu steuern. Sie können somit zusätzlich zu ihrer Funktion als Genrepressoren als Kontrollglied zwischen Zellproliferation/-differenzierung und Apoptose dienen.

2.2.6 Neueste Daten zur Funktion von MLL in Säugerzellen

Kürzlich wurden experimentelle Daten publiziert, die eine proteolytische Spaltung des MLL Proteins belegen. Diese Spaltung ist bereits 10 Minuten nach Translationsbeginn detektierbar. Hsieh et al. (2003) und Yokoyama et al. (2002) konnten mit ihren Experimenten zwei unterschiedlich große Peptide auf einem Proteingel nachweisen, die als MLL-N (p300, AS 1-2666) und als MLL-C (p180, AS 2719-3679) bezeichnet wurden. Das MLL-C Peptid wird schon nach kürzester Zeit posttranslational modifiziert. Eine Phosphorylierung konnte mittels [³²P] nachgewiesen werden (Yokoyama et al., 2002). Innerhalb des MLL Proteins wurden zwei Protease-Erkennungssequenzen identifiziert, die in konservierten Aminosäure-Motiven liegen (*Cleavage site 1* (CS1): QxD•GADD; CS2: QxD•GVDD) und

die Interaktionsmöglichkeit zwischen Histonen und chromatinassoziierten Proteinen definieren.

Die im MLL MPSC identifizierten Proteine gehören z.T. der nukleosomenremodulierenden SWI/SNF-Familie an (BRM, BAF 170/155 und INI 1), dem NuRD-Komplex (Mi2, HDAD1/2, RbAp46/48), dem transkriptionsreprimierenden Sin3A-Komplex (SAP30/18, HDAD1/2 und Sin3A) oder der TAFII-Familie (TAF250/80/31/20 und TBP), die ihre Bedeutung in der direkten Verbindung des MLL MPSC mit der Transkriptionsmaschinerie (Pol-II Komplex) hat. Neben dem MLL-C Peptid sind weitere Faktoren im MLL MPSC gefunden worden, die eine SET-Domäne besitzen und damit Histone an spezifischen Lysin-Resten durch Methylierung modifizieren können (RbBP5 und WDRP5). Zusätzlich ist das CPSF-Protein im Komplex vertreten, das zur Promotorerkennung beiträgt und durch RNA-Polyadenylierung aktiv an wichtigen posttranskriptionellen Prozessierungsvorgängen beteiligt ist. HSNF2H ist ein wichtiger Faktor, der neben den beiden MLL Peptiden die Komplex-Assemblierung stabilisiert. Mittels RNAi Technologie wurde gezeigt, dass eine Blockade der MLL Expression dazu führt, dass kein Multiprotein-Superkomplex gebildet werden kann (Nakamura et al., 2002).

Überraschenderweise wurde durch die zuvor erwähnten Untersuchungen eine Vielzahl von Proteinen identifiziert, denen durch Histon-Deacetylierungs- und auch spezifische Histon-Methylierungsaktivitäten, generell transkriptionsreprimierende Eigenschaften zugesprochen werden können. Trotzdem wurden bisher ausschliesslich transkriptionsaktivierende Prozesse des MLL MPSC beobachtet (z.B. Hox Genregulation). Durch Experimente im transgenen Mausmodell konnte weiterhin nachgewiesen werden, dass den MLL MPSC an definierten Targetgenen weniger eine transkriptionsinitiierende, als eine –aufrechterhaltende Funktion in der Steuerung der Genexpression zukommt (s. Kapitel 2.2.5).

Im Gegensatz zu den TrxG-Komplexen in *D. melanogaster*, die an TREs lokalisiert wurden, konnte in humanen HeLa-Zellen durch Chromatin-Immunoprecipitation eine direkte Bindung des MLL MPSC an verschiedene Hox Gen-Promotoren (*hox a9* und *hox c8*) nachgewiesen werden (Milne et al., 2002; Nakamura et al., 2002). Diese aktivierten Promotoren waren mit Histonen assoziiert, die am Lysin-Rest (Position 4) des Histons H3 methyliert waren (H3K4). Diese Modifikation wird durch die SET-Domäne des MLL-C Peptids erreicht. Es ist jedoch nicht auszuschliessen, dass der MLL MPSC ebenfalls initial mit unterschiedlichen Enhancer-Elementen assoziiert ist, in Folge jedoch direkt an proxi-

male Sequenzen von Targetgen-Promotoren bindet und dort aktivierend oder reprimierend wirkt.

Erkenntnisse über essentielle Grundlagen aus den Arbeiten von Milne et al. (2002) und Nakamura et al. (2002) tragen zur Aufklärung des molekularbiologischen Mechanismus vom Ant- und/oder Agonismus der MLL MPSC und PcG-Komplexe bei. TAFII250, ein Mitglied des MLL MPSC, acetyliert die Aminosäuren K9 und K14 des Histons H3. Diese Acetylierungen scheinen die spezifische und essentielle Histon-Methylaseaktivität (H3K4 Methylierung) der SET-Domäne des MLL-C Peptids zu erhöhen. HDAC-1, ein Protein aus dem Sin3A- und NuRD-Komplex mit chromatinreprimierenden Eigenschaften, konnte bereits in aktivierten MLL/Trx-abhängigen Promotoren (*Abd-B*) identifiziert werden. Deshalb liegt die Vermutung nahe, dass neben spezifischen Acetylierungen ebenso definierte Histon-Deacetylierungen notwendig zur essentiellen H3K4 Methylierung sind. Möglicherweise verhindert die H3K4 Methylierung im Anschluss eine Assoziation inhibitorischer PcG-Komplexe (Nishioka et al., 2002). In weiteren Experimenten konnte gezeigt werden, dass mit mutierten MLL Peptiden (deletierte SET-Domäne) die H3K4 Methylierung ausblieb und der *hox c8* Promotor nicht aktiviert wurde. Wenn im Histon H3 die Aminosäure K9 nicht acetyliert, sondern methyliert ist, kann das entsprechende Gen nicht transkribiert werden.

Hsieh et al. (2003) konnte mit immunohistochemischen Experimenten nachweisen, dass beide MLL Peptide, sofern keine Mutationen im Bereich der Dimerisierungsdomänen vorlagen, in dem schon beschriebenen punktierten Muster im Zellkern kolokalisierten. Bei einer Deletion der FYRC-/SET-Domäne des MLL-C Peptids, wurde dieses diffus im Kern detektiert. Das MLL-N Peptid konnte kaum nachgewiesen werden. Es scheint, dass die Prozessierung des MLL Proteins nicht nur die Stabilität beider Peptide reguliert, sondern maßgeblich an intranukleären Lokalisierungsmechanismen und folglich an der Funktionalität des MLL MPSC beteiligt ist.

2.3 MLL und chromosomale Translokationen

2.3.1 Hochrisiko-Leukämien

Reziproke chromosomale Translokationen sind ein typisches Kennzeichen von Leukämien und Lymphomen. Die reziproke Translokation t(4;11) wird vorwiegend bei Kleinkindern mit akuten lymphatischen Leukämien, aber auch bei älteren Patienten mit Sekundärleukämie gefunden. Diese treten besonders häufig nach einer Tumorbehandlung mit Topoisomerase-II Hemmstoffen (Etoposide) auf (Bower et al., 1994; Cortes et al., 1994; Domer et al., 1995). Da beide Patientengruppen eine ausserordentlich schlechte Prognose aufweisen und meist gegen konventionelle Therapien resistent sind, wird dieser Phänotyp als „Hochrisiko-Leukämie“ bezeichnet.

Obwohl die Zellen, die eine t(4;11) Translokation tragen, morphologisch zumeist dem Krankheitsbild der ALL (akute lymphatische Leukämie) zugeordnet werden, zeigen sie bei einer Untersuchung mit molekulargenetischen und immunologischen Methoden eine parallele Expression von lymphoiden (CD19) und myeloischen (CD13, CD15 und CD33) Oberflächenmarkern. Dieser Biphänotyp führte auch zu der Bezeichnung „*Mixed Lineage Leukemia*“ Gen (*MLL*), das durch die Translokation heterozygot zerstört wird und durch Dosisreduktion und Generierung von Derivatproteinen für die Entstehung dieses Phänotyps ausschlaggebend ist (*MLL*; Ziemin-van der Poel et al., 1991). Es ist deshalb davon auszugehen, dass die transformierten Zellen von primitiven Vorläuferzellen mit dem Potential zur lymphoiden und myeloischen Differenzierung abstammen. Sie sind in ihrer Entwicklung auf der Stufe der pro-B Zellen arretiert und proliferieren als unreife Blasten. Klinische Merkmale sind das junge Alter der Patienten (meist unter einem Lebensjahr) zum Zeitpunkt der Diagnose, Hyperleukozytose, Hepatosplenomegalie und ein hoher Anteil an ZNS-Infiltrationen.

2.3.2 t(4;11) Translokation

Die an dieser Translokation beteiligten Partnergene sind zum einen das *MLL* Gen auf Chromosom 11q23, und zum anderen das *AF-4* Gen, das auf Chromosom 4q21 lokalisiert wurde. Bei der Translokation finden die Chromosomenbrüche in sogenannten Bruchpunkt-Cluster-Regionen (bcr) statt, die in Relation zur Gengröße ungewöhnlich klein sind und sich zum größten Teil in Intronsequenzen befinden. Die Clusterung der genomischen

Bruchpunkte ist auffällig und deutet darauf hin, dass bei einer reziproken Translokation die Chromosomen nicht beliebig miteinander fusionieren können. Eine Gesetzmäßigkeit, der die Produkte einer Chromosomenumlagerung unterworfen sind, scheint die notwendige Aufrechterhaltung eines durchgehenden Leserasters zu sein. Auf genomischer Ebene liegen fast alle untersuchten Bruchpunkte verschiedener 11q23 Translokationen innerhalb des *MLL* Gens in einem 8.3 kb großen *Bam*H1-Fragment (Gu et al., 1994; Marschalek et al., 1995), das im Exon 8 beginnt und im Exon 14 endet.

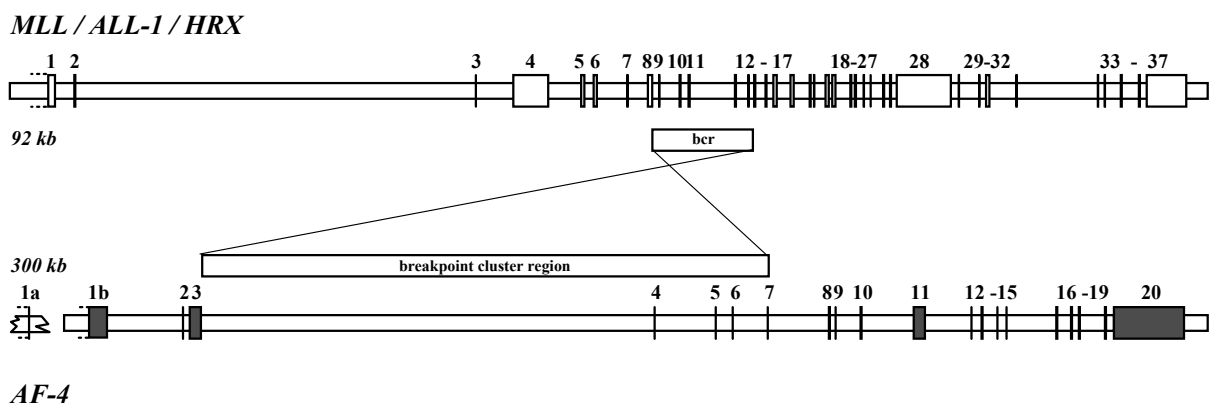


Abb. 2.6: *MLL* und *AF-4* Genstrukturen

Darstellung beider Gene mit ihrer Intron-Exon Struktur. Die Bruchpunkt-Cluster-Regionen (bcr) sind als weiße Boxen über, bzw. unter dem entsprechenden Gen markiert.

Die beiden Partnerchromosomen werden so miteinander fusioniert, dass auf beiden Derivatchromosomen ein durchgehender Leseraster für zwei funktionsfähige Derivatproteine entsteht: *MLL*•*AF-4* (der11) und *AF-4*•*MLL* (der4). In leukämischen Blasten werden in der Regel die Transkripte beider Fusionsgene gefunden. Es bleibt weiterhin zu klären, nach welchem Mechanismus die chimären Proteine einen selektierbaren Proliferationsvorteil entwickeln, der folglich zur klonalen Expansion der leukämischen Blasten führt.



Abb. 2.7: Proteinstatus einer arretierten pro-B Zelle bei einer ALL (t(4;11))

Schematische Darstellung der MLL und AF-4 Proteine mit den entsprechenden Fusionsproteinen MLL•AF-4 (der11) und AF-4•MLL (der4).

AT: AT-Haken, MT: Methyltransferase-Domäne, TRD: Transkriptionsrepressor-Domäne, ZFD: Zinkfinger-Domäne, PHD: Plant Homeobox Domain, TAD: Transkriptionsaktivator-Domäne, SET: SET-Domäne, NHD: N-terminale Homologie-Domäne, NLS: Kernlokalisationssequenz, CHD: C-terminale Homologie-Domäne, pSER: poly-Serin Bereich, ALF: konservierte Domäne in der AF4/LAF4/FMR2 Proteinfamilie.

Das *MLL* Gen ist in ca. 50 illegitime Rekombinationen involviert (Ayton and Cleary, 2001), wobei die meisten dieser Translokationen nicht den bereits beschriebenen Phänotyp einer pro-B ALL, sondern weitaus häufiger mit dem Krankheitsbild einer akuten myeloiden Leukämie (AML) assoziiert sind. Generell weisen AML eine höhere Sensitivität gegenüber Chemotherapeutika auf.

Die Fusionspartner sind in ihrer Struktur sehr unterschiedlich und weisen auch in ihren funktionellen Domänen, soweit bekannt, kaum Gemeinsamkeiten auf. Trotzdem wurde die These formuliert, dass die chimären Proteine die für die maligne Transformation verantwortlichen Onkoproteine sind. Jedoch stellt nach dieser Hypothese nur der N-terminale Bereich des MLL Proteins, als kleinster gemeinsamer Nenner aller bekannten Fusionsproteine, die relevante onkogene Domäne dar. Es konnte jedoch gezeigt werden, dass allein die Verkürzung des MLL Proteins nicht zur Entstehung von Neoplasien im hämatopoietischen System führt. In den wenigen Fällen, die man bisher detaillierter untersucht hat,

spielt der C-terminale Anteil des Partnergens für den Prozess der Transformation eine wichtige Rolle (Corral et al., 1996; DiMartino et al., 2000; Lavau et al., 1997).

Für verschiedene Translokationen konnte im Tiermodell das onkogene Potential der Fusionsproteine nachgewiesen werden. Die Translokation t(9;11) wurde im Mausmodell durch *knock-in* auf chromosomaler Ebene rekonstruiert. Die Expression des Fusionsproduktes führte zur Ausprägung einer Leukämie, die durch die Expression eines verkürzten MLL Proteins und eines MLL•myc Fusionsproteins, nicht ausgelöst werden konnte (Corral et al., 1996). Für die Translokation t(11;19) konnte durch retrovirale Transduktion des *MLL•ENL* Fusionskonstruktes in Knochenmarkszellen gezeigt werden, dass das entsprechende Fusionsprotein in der Lage war, hämatopoietische Vorläuferzellen zu immortalisieren. Weder der N-Terminus des MLL noch ENL allein zeigte diesen onkogenen Effekt (Slany et al., 1998). Interessanterweise haben Knochenmarkstransplantationsassays gezeigt, dass MLL Fusionsproteine nur dann ein Potential zur Transformation primärer hämatopoietischer Zellen aufweisen, wenn das in die Translokation involvierte Partnergen für ein nukleäres Protein kodiert.

Bislang fehlen jedoch publizierte funktionelle Studien, die das transformierende Potential der beiden chimären Proteine MLL•AF-4 und AF-4•MLL in einem biologischen System eindeutig belegt hätten.

Die aus der Translokation resultierende MLL Wildtyp Dosisreduktion scheint zusammen mit der Expression der zwei Fusionsproteine MLL•AF-4 und AF-4•MLL ein verändertes Expressionsmuster in hämatopoietischen Stammzellen oder Vorläuferzellen zu etablieren. Zellen, die *MLL* Translokationen tragen, zeigen ein charakteristisches Genexpressionsprofil, das sie eindeutig von konventionellen akuten lymphatischen und akuten myeloischen Leukämien unterscheidet (Armstrong et al., 2002). Es ist deshalb anzunehmen, dass das spezifische Expressionsmuster der Hochrisiko-Blasten auf die Reduktion der MLL Proteindosis in der Zelle und auf der Anwesenheit eines oder beider Fusionsproteine zurückzuführen ist.

Dieser Zusammenhang lässt sich möglicherweise durch die Wildtyp-Funktion des MLL Proteins erklären: Der N-Terminus besitzt zwei unterschiedliche DNA-Bindungsdomänen (AT-Haken und MT-Domäne) und eine Lokalisierungssequenz (snLS) für bestimmte funktionelle Kernbereiche (Yano et al., 1997). Das chimäre Protein MLL•AF-4 wird nun wahrscheinlich ebenfalls in das nukleäre Subkompartiment transportiert und kompetiert mögli-

cherweise um Bindestellen und stört damit die Funktion des MLL MPSC. Die C-terminale SET-Domäne sowie das PHD-Motiv sind Protein-Protein Interaktionsdomänen, die MLL MPSC, PcG oder andere Proteine binden können. Zusätzlich gibt es im C-Terminus des MLL Proteins eine nukleäre Lokalisationssequenz für Kernimport. Es kann deshalb davon ausgegangen werden, dass das Fusionsprotein AF-4•MLL ebenfalls in den Zellkern gelangt und dort direkt mit dem chromatinregulierenden MLL MPSC interagieren kann.

2.3.3 *AF-4* Gen

Das *AF-4* Gen gehört, neben den *LAF-4*, *FMR-2* und *AF-5* Genen, zur ALF-Familie (Nilson et al., 1997). Es ist auf Chromosom 4, Bande q21 lokalisiert (Chen et al., 1993) und kodiert für ein Serin/Prolin-reiches Protein mit einem Molekulargewicht von 131 kDa. Durch einen Sequenzvergleich wurden 5 hochkonservierte Domänen definiert: die N-terminale Homologie Domäne (NHD), die ALF-Domäne (ALF), der poly-Serin Bereich (pSER), die nukleäre Lokalisationssequenz (NLS) und die C-terminale Homologie-Domäne. Der N-terminale Bereich des AF-4 Proteins (50 kDa) besitzt eine transkriptionsaktivierende Domäne (Baskaran et al., 1997). Die Funktionen der einzelnen Domänen sind z.T. noch unbekannt und Gegenstand aktueller Forschung. Homozygote *knock-out* Experimente konnten jedoch bereits zeigen, dass das *AF-4* Gen nicht essentiell für die Embryonalentwicklung ist, aber einen wesentlichen Einfluss auf die Entwicklung von lymphoiden Vorläuferzellen, kardiovaskulären und neuronalen Strukturen und Skelettformation aufweist. In reifen B-Lymphozyten ist das *AF-4* Gen transkriptionell inaktiviert (Baskaran et al., 1997).

2.3.3.1 Rekombinationsmechanismen

Auffällig ist die extreme Häufung von repetitiven Alu-Elementen in der Bruchpunkt-Cluster-Region, die als Ursache einer homologen Rekombination zwischen zwei Alu-Elementen diskutiert wurde. Dies konnte jedoch widerlegt werden (Gillert et al., 1999; Reichel et al., 1998).

Neben der Alu-Alu Rekombinationstheorie gibt es einen weiteren diskutierten Mechanismus der Translokationsentstehung. Dieser beruht auf der Beobachtung, dass häufig in Folge einer Behandlung mit Topoisomerase-II Hemmstoffen ein rearrangiertes *MLL* Gen gefunden wurde. Dabei lag der genomische Bruchpunkt überwiegend direkt neben einer

Topoisomerase-II Konsensus-Bindestelle. Ein Chromosomenbruch an dieser spezifischen Sequenz konnte ebenfalls mit anderen genotoxischen Agentien ausgelöst werden. Insofern besteht der Verdacht, dass der Bruch des DNA-Stranges im *MLL* Gen durch eine Fragmentierung von übergeordneten Chromatinstrukturen nach Induktion von Apoptose auftritt. Neuere Untersuchungen lassen vermuten, dass der Zeitpunkt des chromosomalen Rearrangements im ersten Trimenon der Schwangerschaft liegt. Vor diesem Hintergrund entstand die Vorstellung, dass natürliche Stoffe oder Medikamente, die von der Mutter während der Schwangerschaft aufgenommen werden, zu Zell-Giften metabolisiert werden und über die Plazenta den Fötus erreichen (Alexander et al., 2001).

2.4 Biologisches Testsystem

Zur Funktionsaufklärung und Targetgen-Identifizierung des nativen Wildtyp *MLL* Proteins und der zwei t(4;11)-Fusionsproteine wurden zwei isogene murine embryonale Fibroblasten-Zelllinien verwendet.

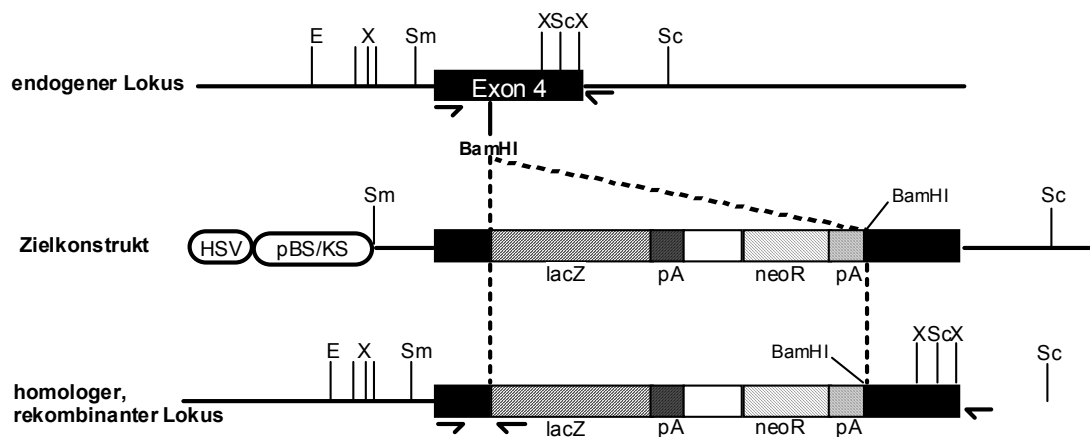


Abb. 2.8: Zerstörung eines *MLL* Allels durch homologe Rekombination

Erkennungssequenzen der folgenden Restriktionsendonukleasen: E: *EcoRI*, X: *Xho*, Sm: *SmaI*, pA: poly-A Sequenz, *neo^R*: Neomycin-Resistenzgen, *lacZ*: *LacZ* Gen, pBS/KS: Vektor pBluescript.

Die murine Fibroblasten-Zelllinie MLL^{-/-} wurde aus Hautzellen eines absterbenden, homozygoten *Mll knock-out* Embryos am Tag 10,5 p.c. etabliert. In das Exon 4 des *Mll* Gens wurde durch homologe Rekombination eine LacZ/neo^R-Kassette inseriert (Yu et al., 1995). Durch diesen „*knock-in*“ (jedoch im Folgenden als Gen „*knock-out*“ bezeichnet) sollte die Transkription und Translation ab Exon 4 komplett inhibiert werden. Die isogene Zelllinie mit dem homozygot intakten *Mll* Gen (MLL^{+/+}) diente als Kontrolle. Beide Zelllinien wurden durch Infektion mit dem T-Antigen des Polyoma-Virus immortalisiert.

2.5 DNA-Mikrochip-Arrays

Mit der DNA-Microchip-Hybridisierung steht heutzutage eine innovative Technologie zur Verfügung, mit der das gesamte Transkriptom einer Zelllinie oder Gewebes in kürzester Zeit analysiert werden kann.

Auf Mikrochips aufgespottete Oligonukleotide werden mit biotinylierten cRNAs hybridisiert und anschliessend mit Streptavidin-markierten Fluoreszenz-Farbstoffen versetzt. Ein GeneArray Scanner wertet die Signalstärken aus, die den Expressionsstärken der einzelnen Gene entsprechen. Diese Methode zeichnet sich durch die schnelle und reproduzierbare Quantifizierung von Genexpressionen aus, wobei viele tausend bis mehrere zehntausend Gene gleichzeitig analysiert und auch zum Vergleich zu anderen Zelllinien oder Geweben herangezogen werden können. Durch diese Entwicklung wurde ein neues Instrument geschaffen, das die Möglichkeiten bietet, unterschiedliche Expressionsmuster zu detektieren und Genfunktionen sowie Zielgene aufzudecken. Die DNA-Mikrochip-Analyse erfordert geringen Zeitaufwand und ist im Gegensatz zu herkömmlichen Methoden, wie z.B. der subtraktiven Klonierung, relativ zeit- und kostengünstig.

Für die Hybridisierung wird hochqualitative und absolut DNA-freie RNA aus Zelllinien oder Geweben eingesetzt. Nach dem „*standard-one-round amplification protocol*“ (s. Abb. 2.9) wird diese in biotinylierte cRNA umgeschrieben und anschliessend mit einem Testchip zur Qualitätsbestimmung hybridisiert. Im nächsten Schritt wird der Analysechip hybridisiert, dann Streptavidin-markiertem Phycoerythrin zugegeben und die Signalstärke detektiert. Auf diesem Chip befinden sich verschiedene „Proben-Sets“, sowie entsprechende Sets für interne Kontrollen.

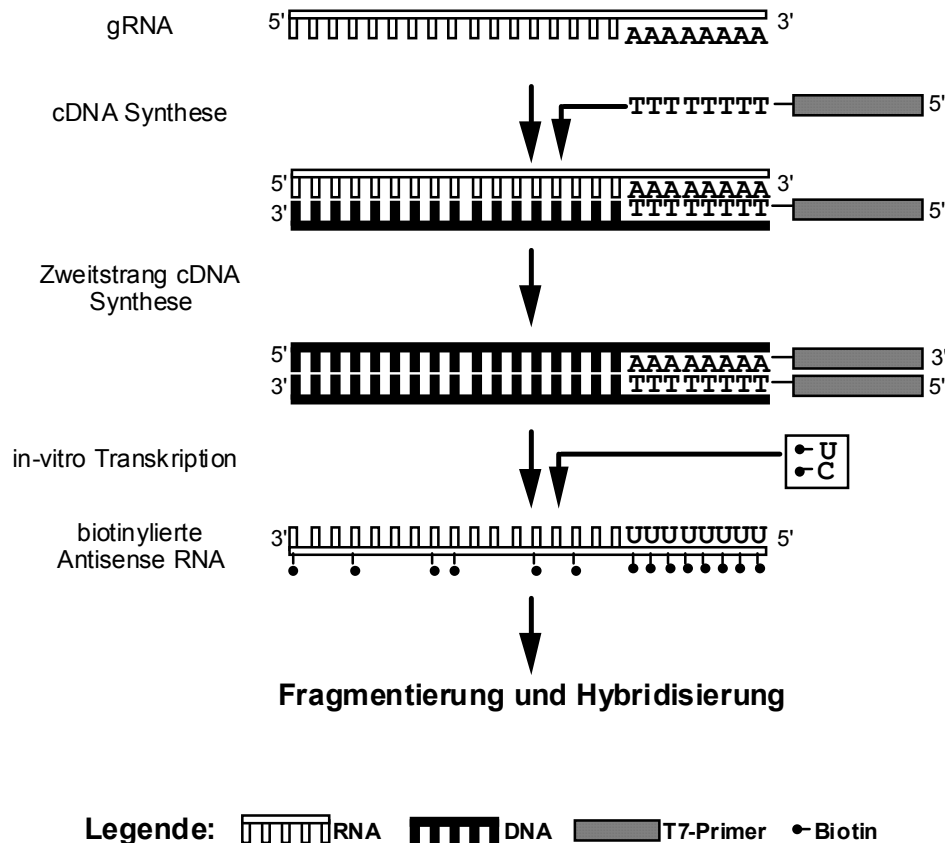


Abb. 2.9: Synthese biotinylierter antisense-RNA („standard-one-round amplification protocol“)
gRNA: gesamt-RNA

Jedes Gen wird durch ein Proben-Set, bestehend aus 20 verschiedenen Proben-Paaren, repräsentiert. Ein einzelnes Proben-Paar besteht aus einem 25mer Oligonukleotid eines Gens (PM; Perfect Match) und dem gleichen Oligonukleotid, indem jedoch die Base an Position 13 mutiert wurde (MM; Mismatch). Das Signal des MM stellt ein Maß für die unspezifische Hybridisierung dar und wird von dem des PM subtrahiert. Die Differenz dieser beiden Signalstärken gibt Auskunft über die Anwesenheit (PM>MM) oder die Abwesenheit (PM<MM) des entsprechenden Transkriptes. Aus diesen 20 Proben-Paaren wird nun der Mittelwert gebildet, der dann in eine vergleichende und statistische Analyse eingeht und dort als „Signal log₂ ratio“ bezeichnet wird und für die Auswertung von großer Bedeutung ist.

Zur Fehlerminimierung und zur Qualitätsbestimmung sind auf jedem Chip Kontrollgene aufgebracht, die eine Überprüfung der erfolgreichen cRNA-Synthese und Hybridisierung möglich machen. Mit der sogenannten „*polyA*⁺-*spike*“ Kontrolle wird ermittelt, ob die cRNA-Synthese erfolgreich war. Hierzu werden mRNA Transkripte der Gene *DapX*, *LysX*, *PheX*, *ThrX* und *TrpnX* dem „*standard-one-round amplification protocol*“- Probenansatz zugegeben und müssen nach der Hybridisierung eine vorher definierte Signalstärke ergeben. Die vier cRNA Transkripte der aufgespotteten Gene *BioB*, *BioC*, *BioD* und *CreX*, werden dem Hybridisierungsansatz beigelegt. Anhand deren Signalstärken lässt sich eine Aussage über die Hybridisierungsqualität machen. Wie auf den Testchips befinden sich auch auf dem Analysechip *housekeeping* Gene (Oligonukleotide vom 3'- und 5'-Ende), die die eingesetzte RNA-Qualität definieren.

Werden nun die Expressionsmuster und -stärken zweier DNA-Microchips miteinander verglichen, muss zunächst eine Normierung durchgeführt werden, bei der durch verschiedene Skalierungsverfahren die Signalintensität der DNA-Microchips rechnerisch angeglichen wird. Das Hintergrund-Rauschen wird durch Messung der 20 schwächsten Pixel aus 16 Chip-Feldern bestimmt.

2.6 Proteomanalyse

Die Proteomanalyse ist eine wichtige Ergänzung zur Transkriptionsanalyse, da sich das komplexe Netzwerk der Zellregulation nicht allein durch das Entschlüsseln von RNA Expressionsmustern einer Zelle charakterisieren lässt. Untersuchungen der RNA Expression liefern zwar ein realistisches Bild der tatsächlichen Vorgänge in einer Zelle, jedoch korrelieren die RNA Level nicht unbedingt mit den Mengen entsprechender Genprodukte. Bereits auf der Transkriptionsebene setzen die ersten Mechanismen zur Regulation der Genaktivität ein: alternatives Spleissen, Degradation und unterschiedliche Translationsraten. Zur Absicherung von differentiellen RNA Analysen (z.B. subtraktive Klonierung und DNA-Microchip-Hybridisierung) und zur Ermittlung des tatsächlichen intrazellulären Geschehens ist eine Analyse der verschiedenen Proteinmoleküle und ihrer modifizierten Isoformen notwendig. Die Anzahl der Proteinspezies ist z.B. durch posttranslationale Modifikation (z.B. Hydroxylierung, Glykosylierung und Phosphorylierung) um ein Vielfaches größer als die Anzahl der differentiell exprimierten mRNA Moleküle. Im Gegensatz zum

statischen Genom verändern sich Transkriptom und Proteom während der Entwicklung und als Reaktion auf unterschiedliche Einflüsse dynamisch.

Die Proteomanalyse ist somit ein wichtiges Instrument um die Fragen nach posttranslationaler Modifikation, Isoformen, räumlicher Verteilung und Expressionsstärken unter veränderten Bedingungen zu beantworten. Mit Hilfe der zweidimensionalen Gelelektrophorese lassen sich mehrere tausend Proteine einer Zelle trennen, auf einem Gel darstellen und durch Kopplung an hochauflösende massenspektrometrische Verfahren charakterisieren. Die daraus gewonnenen Erkenntnisse lassen sich zur Struktur- und Funktionsaufklärung der einzelnen Proteine in der Zelle nutzen. Proteomics stellt somit ein Ansatz dar, der die klassischen Methoden der DNA- und RNA-Forschung sinnvoll erweitert, der aber auch limitiert ist, weil nur diejenigen Proteine erfasst werden können, die in ausreichenden Mengen in einer Zellpopulation vorhanden sind.

2.6.1 Zweidimensionale Gelelektrophorese

Die zweidimensionale Gelelektrophorese ist derzeit die leistungsfähigste Methode, mit der die reproduzierbare Auftrennung von komplexen Proteingemischen unter Beibehaltung der quantitativen Verhältnisse durchgeführt werden kann. Der Vorteil ist, dass die Auftrennung der Proteine nach zwei unterschiedlichen physikalisch-chemischen Kriterien, nämlich Trennung aufgrund von Ladung (isoelektrischer Punkt, 1. Dimension) und Masse (Molekulargewicht, 2. Dimension) erfolgt und die Gesamtheit der Proteine innerhalb bestimmter Grenzen auf einem Gel sichtbar gemacht werden können. Im Anschluss können mit Hilfe von computergestützten Bildauswertungsverfahren und massenspektrometrischen Methoden einzelne Proteine analysiert, bzw. identifiziert werden.

2.7 Zielsetzung dieser Arbeit

Aufgrund der Bedeutung des *MLL* Gens bei der Chromatinregulation während embryonaler Differenzierungsvorgänge und seiner Involvierung in zahlreiche Translokationen sowie der damit verbundenen pathogenetischen Bedeutung, ist das *MLL* Gen seit seiner Entdeckung Gegenstand intensiver Forschung.

1. Mit Hilfe eines biologischen Testsystems, das aus zwei isogenen embryonalen Mausfibroblasten-Zelllinien (*MLL*^{-/-} und *MLL*^{+/+}) bestand, sollten die wichtigsten Funktionen des nativen *MLL* Proteins durch die Charakterisierung seiner Zielgene aufgeklärt werden. Dazu sollten verschiedene Analysemethoden angewendet werden. Die Erstellung eines *MLL*-abhängigen Expressionsprofils sollte zuerst auf RNA-Ebene durch subtraktive Kloniertechniken und DNA-Microchip-Arrays erfolgen. Diese Ergebnisse sollten durch massenspektrometrische Verfahren (MALDI-TOF MS) nach zweidimensionaler Gelelektrophorese auf Protein-Ebene verifiziert werden.

2. In der vorliegenden Arbeit sollte zusätzlich eine Komplementationsanalyse Aufschluss darüber geben, ob das humane *MLL* Protein das fehlende murine Homolog in den *MLL*^{-/-} Zellen in seiner Funktion komplementieren kann. Dazu sollten zwei weitere *MLL*^{-/-} Zelllinien (*MLL*^{-/-}::h*MLL* und *MLL*^{-/-}::BC230A-Leervektor) etabliert werden. Aus dem Vergleich beider Expressionsmuster mittels DNA-Microchip-Arrays sollten Rückschlüsse gezogen werden, inwieweit das murine System zur Untersuchung des Einflusses humaner Proteine auf Transkriptionsprozesse (auch im Hinblick auf der unter Punkt 3 beschriebenen Versuche) geeignet ist.

3. Bei der t(4;11) Translokation entstehen durch illegitime Rekombination die zwei Fusionsgene *MLL*•*AF-4* und *AF-4*•*MLL*, die für funktionsfähige Derivatproteine kodieren. Diese Proteine scheinen ein onkogenes Potential aufzuweisen, das nach einem bisher unbekanntem Mechanismus zur klonalen Expansion der leukämischen Blasten führt. Die cDNA Konstrukte der chimären Gene sollten ebenso wie der Leervektor (pBC230A) in *Mll* Wildtyp-Zellen stabil transfiziert werden. Mittels DNA-Microchip-Technologie sollten in diesem Versuchsansatz die Zielgene der einzelnen Fusionsproteine und ihrer Kombination identifiziert werden. Die Kenntnis darüber sollte zum Verständnis des pathogenetischen Mechanismus der Leukämogenese beitragen.

3 Ergebnisse

3.1 Etablierung neuer Zelllinien

3.1.1 Tetracyclin-induzierbares Expressionssystem in MLL^{-/-} Zellen

Das Hauptziel dieser Arbeit war die Identifizierung von Targetgenen des humanen MLL Proteins. Daneben sollte versucht werden, Targetgene zu identifizieren, deren Transkription durch Fusionsproteine beeinflusst wird, die durch die Translokation t(4;11) entstehen (MLL•AF-4 und AF-4•MLL). Dazu sollten induzierbare Expressionskonstrukte stabil in *Mll knock-out* Zellen stabil transfiziert werden, da durch das Fehlen des endogenen MLL Proteins ein direkter Einfluss auf die Genexpression detektierbar sein sollte. Im Anschluss an die Induktion sollte nach 4, 8, 12, 24 und 48 h RNA aus den Zellen isoliert werden und diese mittels subtraktiven Kloniermethoden und DNA-Microchip-Arrays auf unterschiedliche Transkriptionsmuster analysiert werden.

Mit der Entwicklung des *tet-off* Systems wurde von Gossen et al. (1994) ein induzierbares Expressionssystem für Säugerzellen geschaffen. Es besteht aus einem Vektor, der die Sequenz des Transaktivatorproteins (tTA) enthält und einem weiteren, induzierbaren Expressionsvektor, in den das induzierbare Transgen einkloniert wird. Das Transaktivatorprotein ist ein Fusionsprotein aus dem N-terminalen Bereich (AA 1-207) des prokaryontischen Tet-Repressors und den C-terminalen 127 Aminosäuren der HSV-VP16 Aktivierungsdomäne. Der VP16 Bereich konvertiert den Tet-Repressor vom Transkriptionsrepressor in einen Transkriptionsaktivator. Der induzierbare Vektor zeichnet sich durch 7 hintereinander geschaltete Tet-Operator-Sequenzen aus, die unter dem Einfluss eines Minimal-CMV-Promotors stehen. Durch das Fehlen der assoziierten starken Promotor-Enhancer Elemente kann das einklonierte Transgen nur dann effizient exprimiert werden, wenn alle Tet-Operatoren durch die Transaktivatorproteine TetR::VP16 (tTA) gebunden werden. Anstelle des Tetracyclins wird Doxycyclin verwendet, da es eine längere Halbwertszeit im Medium aufweist (ca. 24 h) und somit die Kultivierung der Zellen erleichtert. Doxycyclin bindet an den Transaktivator und führt zu einer Konformationsänderung, die folglich eine Bindung an die Tet-Operons verhindert und die Transkription des Transgens inhibiert. Aufgrund der Vorteile: exakte Regulationsmöglichkeit, Fehlen von pleiotropen tTA Effekten, Erreichen höherer Expressionslevel als durch den CMV-Promotor und den gut charakterisierten Doxycyclin-Eigenschaften, schien dieses Expressionssystem für die Durchführung unserer Experimente gut geeignet zu sein.

Zunächst mussten hierzu jedoch die Ausgangsvektoren für die Expressionsexperimente kloniert und getestet werden.

Als induzierbaren Vektor wurde pUHD10-3 gewählt, der sich durch ein *tet-responsive-element* (TRE, 7 P_{MinCMV}-gesteuerte Tet-Operons) vor der multiplen Klonierstelle (*mcs*) und der Abwesenheit einer Resistenzkassette auszeichnet. In die *mcs* wurde über *EcoRI*- und *BamHI*-Schnittstellen der *SfiI*-Linker UHD•3/•5 einkloniert, in den die vorhandenen *SfiI*-Kassetten der Expressionskonstrukte für *hMLL* (12,1 kb), *MLL•AF-4* (6,9 kb) und *AF-4•MLL* (9,1 kb) eingebracht wurden. Die *ClaI*-Schnittstelle diente zur Überprüfung der erfolgreichen Einklonierung des Linkers.

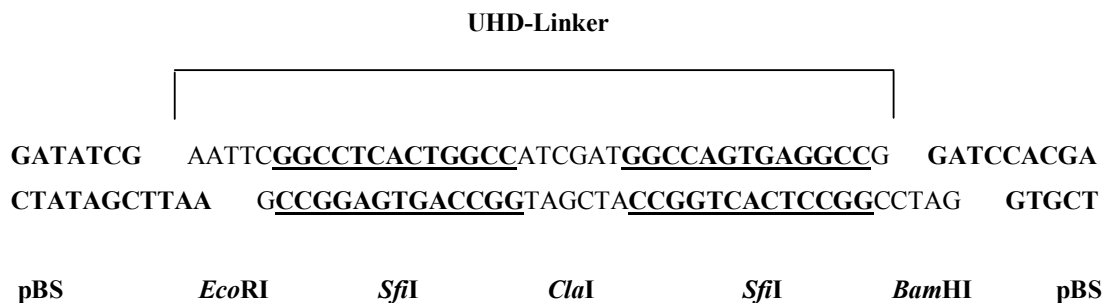


Abb. 3.1: Schematische Darstellung des Linkers UHD mit seinen Schnittstellen

In der Mitte der Abbildung ist der ds Linker UHD dargestellt (UHD•3/UHD•5). Die Überhänge wurden so gewählt, dass sie mit dem *EcoRI*- und *BamHI*-geschnittenen Vektor pBluescript ligiert werden konnten. Durch Restriktionsverdau mit *ClaI* liess sich die erfolgreiche Ligation überprüfen. Anschliessender *SfiI*-Verdau machte das gerichtete Einklonieren der Expressionskonstrukte möglich.

Das Fehlen der Resistenzkassette sollte dazu führen, dass die Kopienanzahl der transgenen Expressionskonstrukte in der Zelle nicht durch Selektionsdruck erhöht wurde. Es wird ein, im Idealfall mindestens siebenfacher, Transaktivator-Überschuss für die transkriptionelle Kontrollierbarkeit eines Transgens benötigt. Zur Selektionserleichterung der transfizierten Zellen wurden die klonierten Vektoren pUHD::hMLL, pUHD::der4 und pUHD::der11 mit dem Vektor pHMR272 (Hygromycin-Resistenz) kotransfiziert.

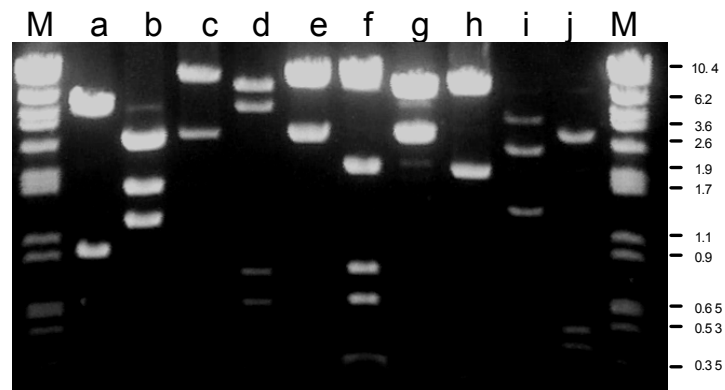


Abb. 3.2: Restriktionsverdau der induzierbaren Vektoren

Bei allen Konstrukten sind die cDNAs als *Sfi*I-Kassetten mobilisierbar. M: Marker (λ *Cla*I), Größenangaben in kb, a: pcDNA::tTA (*Eco*RI/*Bam*HI), b: pcDNA::tTA (*Eco*RI/*Pvu*II) c: pUHD::hMLL (*Sfi*I), d: pUHD::hMLL (*Eco*RI), e: pUHD::der4 (*Sfi*I), f: pUHD::der4 (*Eco*RI), g: pUHD::der11 (*Sfi*I), h: pUHD::der11 (*Eco*RI), i: pHMR272 (*Bam*HI/*Hind*III), j: pHMR272 (*Eco*RI). Die Restriktionskarten stimmen in allen Fällen mit den hier gezeigten Fragmentgrößen überein.

Der die Transaktivatorsequenz enthaltende Vektor pUHD15-1 besitzt eine Neomycin-Resistenz. Da jedoch bei der Herstellung der *knock-out* Zellen u.a. ein Neo^R-Konstrukt durch homologe Rekombination ins Genom eingeführt wurde, war eine Selektion mit G418 nicht mehr möglich. Aus diesem Grund wurde die Transaktivatorsequenz für das Protein TetR::VP16 (tTA) durch einen Restriktionsverdau mit den Enzymen *Eco*RI und *Bam*HI aus pUHD15-1 ausgeschnitten. Nach einem Aufreinigungsschritt aus dem Agarosegel wurde dieses Fragment in die entsprechenden Schnittstellen des Vektors pcDNA3.1, der eine Zeocin-Resistenzkassette besitzt, einkloniert.

Vor Beginn der Zelltransfektion mussten zunächst Wachstumskurven mit MLL^{-/-} Zellen bei unterschiedlichen Konzentrationen an Selektionsantibiotika aufgenommen werden. Es wurde diejenige Konzentration bestimmt, bei der die nichttransfizierten Zellen nach 10 Tagen komplett abgestorben waren. Für Zeocin ergab sich daraus eine Konzentration von 50 μ g/ml und für Hygromycin 120 μ g/ml.

Die $MLL^{-/-}$ Zellen wurden zuerst mit dem Vektor pcDNA::tTA stabil transfiziert, der für den Tet-R::VP16 Transaktivator kodiert. Nach erfolgreicher Selektion mit Zeocin wurde mittels PCR an genomischer DNA die Integration des Transaktivator-Fragmentes bestätigt und zusätzlich durch RT-PCR und Southern-Blots getestet, welche der verschiedenen $MLL^{-/-}$::tTA Klone konstitutiv eine ausreichende Menge des Transaktivators transkribierte (s. Abb 3.3). Der Klon mit der höchsten RNA-Transkriptionsrate ($MLL^{-/-}$::tTA; Klon 3) wurde für alle weiteren Experimente verwendet. Im Anschluss an diese Scening-Experimente wurden die induzierbaren *MLL* Expressionskonstrukte und deren Kombination in Klon 3 stabil transfiziert.

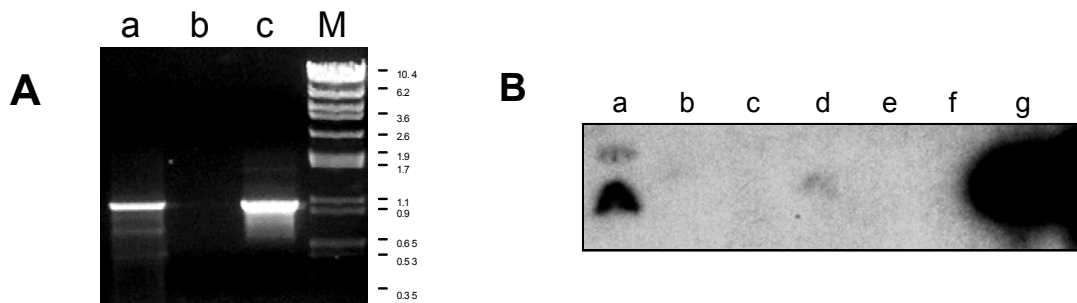


Abb. 3.3: Expressionskontrolle des Transaktivators von $MLL^{-/-}$::tTA, Klon 3

A: Aus den $MLL^{-/-}$::tTA Zellen (Klon 3) wurde RNA präpariert und diese revers transkribiert (a: mit RT, b: ohne RT, als Test auf DNA-Kontamination), c: tTA-Positivkontrolle (PCR am Ausgangsvektor); M: Marker (λ *ClaI*), Größenangaben in kb. Alle RT-PCR-Reaktionen wurden mit dem Oligonukleotid-Paar tTA*3/*5 durchgeführt; erwartete Fragmentgröße: 1025 bp, **B:** Southern Blot, a-f: $MLL^{-/-}$::tTA Klon 3 bis Klon 8, g: Positivkontrolle; Hybridisierung mit radioaktiv markiertem tTA-PCR Fragment (Abb. A, Spalte c).

Klon 3 (jeweils Spalte a) zeigte in verschiedenen Experimenten die stärkste tTA-Expression und wurde für weitere Transfektionen verwendet. $MLL^{-/-}$::tTA Klon 6 (Spalte d) führte im Southern Blot zu einem schwachen Signal.

Die Zweit-Transfektionsexperimente wurden in Gegenwart von 10 μ g/ml Doxycyclin durchgeführt, da unter diesen Bedingungen der TetT::VP16 Transaktivator nicht an seine Erkennungssequenz binden und somit die Transkription einklonierter Transgene inhibiert sein sollte. Insgesamt wurden die folgenden 4 induzierbaren Zelllinien etabliert: $MLL^{-/-}$ tTA::hMML, $MLL^{-/-}$ tTA::der4, $MLL^{-/-}$ tTA::der11 und $MLL^{-/-}$ tTA::der4/der11, die jeweils eine Zeocin- und Hygromycin-Resistenz aufwiesen (s. Tab. 3.4).

Ausgangszelllinie	Transaktivator	induzierb. Vektor + Selektionsvektor	Generierte Zelllinie	PCR an gDNA	RT-PCR (+ Doxy.)	RT-PCR (- Doxy.)
MLL ^{-/-}	pcDNA::tTA	pUHD::hMLL + pHMR272-Hyg	MLL ^{-/-} tTA::hMLL	pos	pos	pos
MLL ^{-/-}	pcDNA::tTA	pUHD::der4 + pHMR272-Hyg	MLL ^{-/-} tTA::der4	pos	pos	pos
MLL ^{-/-}	pcDNA::tTA	pUHD::der11 + pHMR272-Hyg	MLL ^{-/-} tTA::der11	pos	pos	pos
MLL ^{-/-}	pcDNA::tTA	pUHD::der4/der11 + pHMR272-Hyg	MLL ^{-/-} tTA::der4/11	pos	pos	pos

Tab. 3.4: Induzierbare Zelllinien und deren Analyse

Die RT-PCR-Reaktionen mit cDNA aus den Zelllinien, die in Doxycyclin-haltigem Medium kultiviert wurden, hätten bei einem funktionierenden *tet-off* System kein positives Ergebnis zeigen dürfen. tTA: Transaktivator-Protein, Doxy: Doxycyclin (10µg/ml), Hyg: Hygromycin-Resistenz.

Durch RT-PCR Analysen konnte jedoch gezeigt werden, dass trotz der Anwesenheit der zehnfachen (10 µg/ml) der empfohlenen Doxycyclin-Konzentration, eine erhebliche Expression aller Transgene stattfand, die für *der11* in Abb. 3.5 beispielhaft gezeigt wird. Das *tet-off* System schien, obwohl tTA in ausreichenden Mengen exprimiert wurde, durchlässig zu sein und wurde den experimentellen Erfordernissen nicht gerecht.

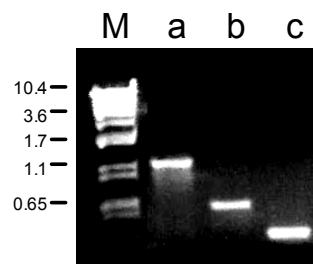


Abb. 3.5: RT-PCR eines uninduzierten MLL^{-/-}tTA::der11 Klons.

a: Transaktivator (tTA•3/•5), b: *der11* (MLL•3/AF-4•5), c: Positivkontrolle (β -Actin), M: Marker (λ *Clal*), Größenangaben in kb. Trotz der Anwesenheit von Doxycyclin wird *der11* exprimiert.

3.1.2 Konstitutives Expressionssystem in $MLL^{+/+}$ und $MLL^{-/-}$ Zellen

In Ermangelung eines induzierbaren Expressionssystems wurden Experimente in einem konstitutiv exprimierenden Zellsystem geplant und durchgeführt.

Zur Etablierung dieser neuen Zelllinien wurden die Expressionskassetten *hMLL*, *der4* und *der11* in die *SfiI*-Schnittstelle des eukaryontischen Vektors pBC230A (Hyg^R) einkloniert. Durch die im Vektor vorhandene EBNA1-Kodiersequenz lag der Vektor nach stabiler Transfektion in den Zellen episomal vor, weshalb keine Einzelklone isoliert werden mussten. Während der Klonierungen wurden einzelne Schritte jeweils über Restriktionsverdau und Sequenzierung überprüft. Nach Abschluss der Klonierungen wurden alle Konstrukte nochmals auf Größe und Orientierung getestet und durch partielle Sequenzanalysen auf Fehlerfreiheit der Expressionskassetten überprüft.

Die Wildtyp-Zellen wurden mit den Vektoren pBC230::*der4*, pBC230::*der11* und dem Originalvektor pBC230A (Negativkontrolle) stabil transfiziert und mit Hygromycin (120 µg/ml) selektiert. Für eine weitere Kotransfektion wurden Zellen verwendet, die bereits das *der11* Konstrukt im Vektor pBC230A trugen. In diese Zellen wurde zusätzlich das Expressionskonstrukt pBC364A/His::*der4* transfiziert. Diese Kotransfektante wurde mit Hygromycin und Histidinol (2 mM) selektiert und konnte nach 3 Wochen stabil kultiviert werden.

Alle hier verwendeten Transgene und deren konstitutive Transkription wurden durch PCR- und RT-PCR-Experimente analysiert und verifiziert, wobei für alle Ansätze positive Ergebnisse erzielt werden konnten.

Ausgangszelllinie	Expressionsvektor	Zelllinie	RT-PCR	Sequenzierung
$MLL^{+/+}$	pBC230A	$MLL^{+/+}::BC230A$	pos	ok
$MLL^{+/+}$	pBC230:: <i>der4</i>	$MLL^{+/+}::der4$	pos	ok
$MLL^{+/+}$	pBC230:: <i>der11</i>	$MLL^{+/+}::der11$	pos	ok
$MLL^{+/+}$	pBC230:: <i>der11</i> pBC364A/His:: <i>der4</i>	$MLL^{+/+}::der4/11$	pos	ok

Als Ausgangsvektor zur Durchführung der Komplementationsanalyse wurden der konstitutiv exprimierende Vektor pBC230::hMLL und der Originalvektor pBC230A in die *knock-out* Zellen stabil transfiziert.

Ausgangszelllinie	Expressionsvektor	Zelllinie	RT-PCR	Sequenzierung
MLL ^{-/-}	pBC230A	MLL ^{-/-} ::BC230A	pos	ok
MLL ^{-/-}	pBC230::hMLL	MLL ^{-/-} ::hMLL	pos	ok

Tab. 3.6: Konstitutiv exprimierende Zelllinien und deren Analyse

Bei der RT-PCR wurden jeweils Oligonukleotide eingesetzt, die an Erkennungssequenzen innerhalb der Expressionskonstrukte *hMLL*, *MLL•AF-4* und *AF-4•MLL* gebunden haben. Bei den *mock*-transfizierten Zelllinien wurden EBNA1-spezifische Primer verwendet.

3.2 MLL als Differenzierungsprotein in der Embryonalentwicklung / Identifizierung von MLL Targetgenen

3.2.1 Morphologische Charakteristika von MLL^{+/+} und MLL^{-/-} Zellen

Bei der Kultivierung von *Mll* Wildtyp und *knock-out* Zellen fiel auf, dass die Zelllinien eine stark unterschiedliche Morphologie zeigen. *Mll* Wildtyp-Zellen lassen sich durch ihre zahlreichen und langen Filopodien deutlich von den MLL^{-/-} Zellen unterscheiden. MLL^{-/-} Zellen sind größer, haben einen größeren Zellkern und weisen charakteristische Lamellipodien und sog. „ruffles“ auf.

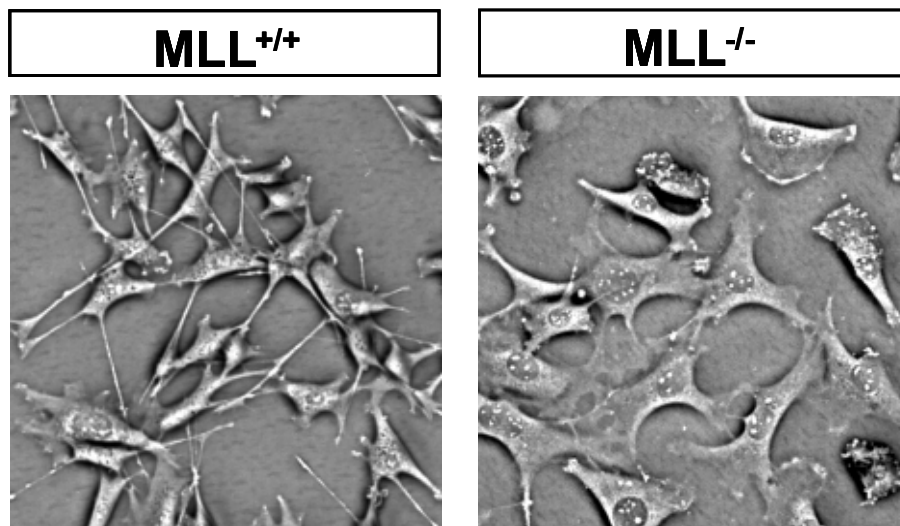


Abb. 3.7: Morphologische Charakteristika von MLL^{+/+} und MLL^{-/-} Zellen
Phasenkontrastaufnahme bei 100-facher Vergrößerung.

In einer Kooperationsarbeit mit Herrn Prof. Dr. M. Presta (Inst. für Biotechnologie, Universität Brescia, Italien) konnten folgende Immunfluoreszenz-Aufnahmen generiert werden. Diese zeigen extreme Unterschiede im Zytoskelettaufbau der verschiedenen Zelllinien.

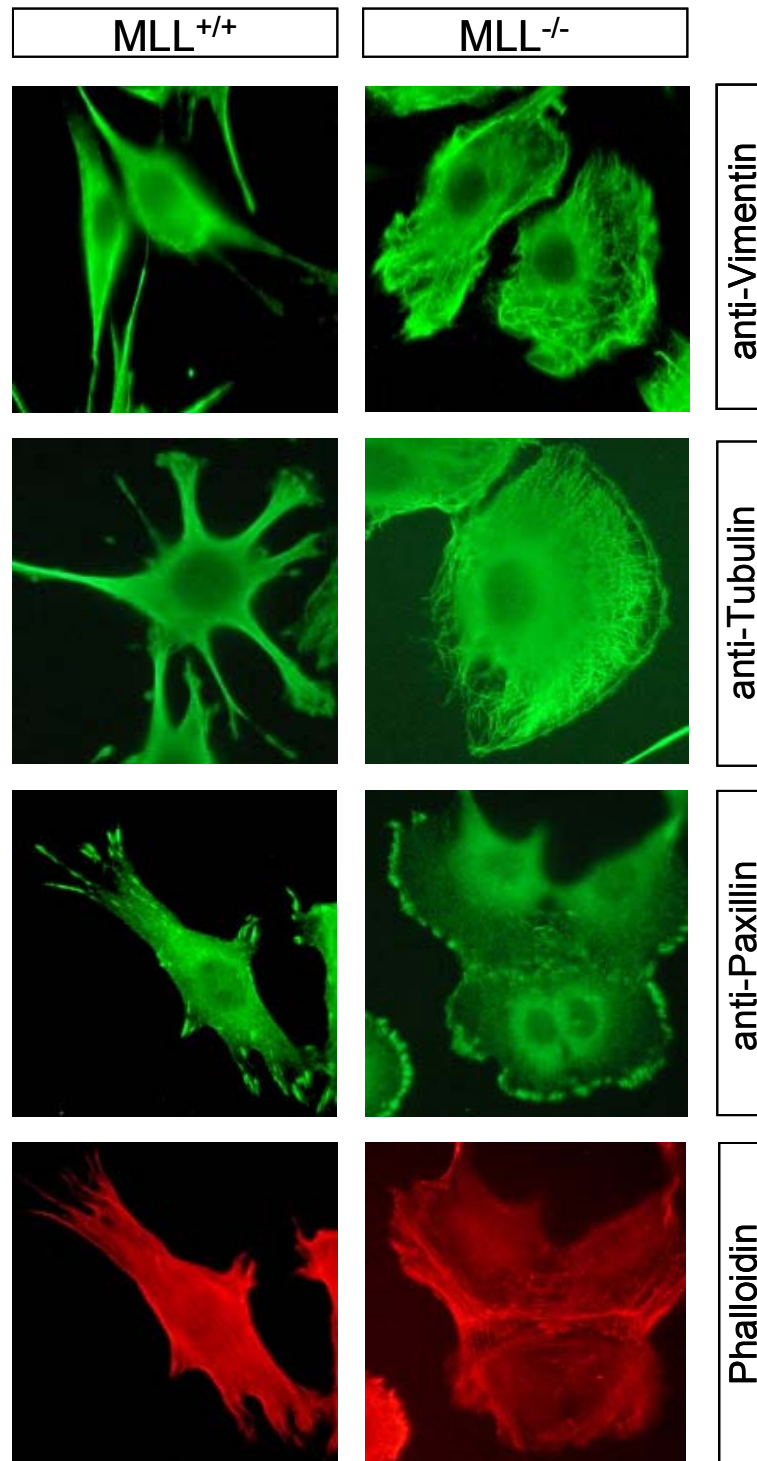


Abb. 3.8: Immunfluoreszenzfärbungen verschiedener Zytoskelettproteine

3.2.2 Intrazelluläre Lokalisation

Immunohistochemische Experimente bieten Lokalisationsmöglichkeiten definierter Proteine innerhalb einer Zelle. In diesem Versuch sollte die intranukleäre Lokalisation des endogenen murinen MLL Proteins bestätigt werden, die für die Funktion als transkriptionsregulierender Faktor zwingend notwendig ist. Dazu wurde ein monoklonaler Antikörper verwendet, der gegen ein spezifisches Epitop des Exon 3 gerichtet ist (MAK E3). Nach Markierung des MLL Antikörpers durch einen FITC-markierten Zweit-Antikörper konnten folgende Bilder mit Hilfe des konfokalen Laserscannmikroskops (CLSM) aufgenommen werden.

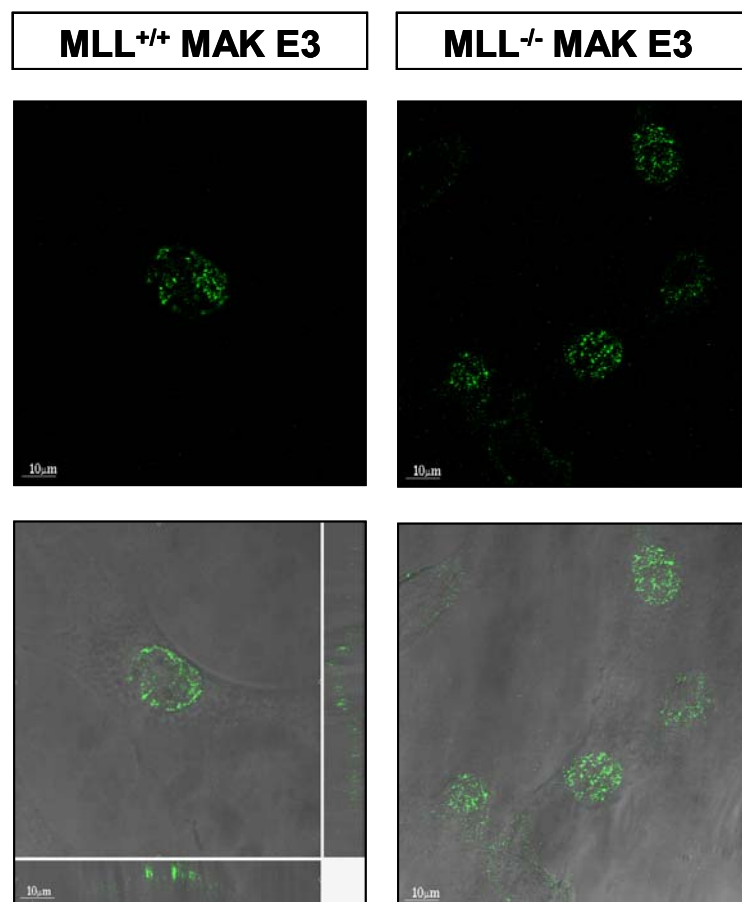


Abb. 3.9: Nukleäre Lokalisation vom N-Terminus des MLL Proteins in MLL^{+/+} und MLL^{-/-} Zellen

Der verwendete murine monoklonale Antikörper ist gegen Epitope in Exon 3 des MLL Proteins gerichtet. Als Zweit-Antikörper wurde ein FITC-markierter anti-Maus Antikörper verwendet. Nach erfolgreicher Präparation wurden Bilder der Zellen mit Hilfe des CLSM aufgenommen.

In der linken Spalte sind Bilder der Wildtyp-Zellen und in der rechten Spalte der *knock-out* Zellen abgebildet. Bei den Bildern in der oberen Reihe handelt es sich um reine Fluoreszenz-Aufnahmen, bei den unteren Darstellungen wurden diese Fluoreszenz-Bilder mit den entsprechenden Transmissionsbildern kombiniert (Bildbearbeitungsprogramm IMARIS). Diese „overlays“ ermöglichen eine exaktere Darstellung der intrazellulären Lokalisation. Die Fluoreszenz-Verteilung in den einzelnen Zellen bestätigt die publizierte nukleäre Lokalisation des murinen MLL Proteins, mit ihrer charakteristischen gleichmäßig punktierten Distribution (Ennas et al., 1997; Yano et al., 1997). Fluoreszenz ist hauptsächlich im Kern und nur schwach an einzelnen Stellen im Zytoplasma detektierbar. Das endogene Wildtyp MLL Protein ist in seinen MLL MPSC gleichmäßig über den gesamten Nukleus verteilt. In den *knock-out* Zellen, in denen das MLL Protein nur bis zur Insertionskassette in Exon 4 exprimiert wird, ist eine vergleichbare Fluoreszenz-Verteilung vorhanden, da sich die subnukleäre Lokalisierungssequenz vor Exon 4 befindet und somit in diesem Peptidfragment ebenfalls vorhanden ist (Yano et al., 1997).

3.2.3 Subtraktive Kloniertechniken

Die Tatsache, dass es sich bei MLL um ein nukleäres, chromatinregulierendes Protein handelt, ermöglicht die Untersuchung seines Einflusses auf Transkriptionsmuster und Zielgene. Diese Untersuchung sollte in den folgenden Experimenten durch Vergleich der Expressionsmuster von Wildtyp und *knock-out* Zellen mittels unterschiedlicher Methoden erfolgen.

Mittels subtraktiver Klonierung ist es prinzipiell möglich, differentiell exprimierte Gene im Vergleich zweier Zellpopulationen zu identifizieren. In einem ersten Versuchsansatz sollten deshalb die Zielgene des nativen murinen MLL Proteins detektiert werden, d.h. das Transkriptom der MLL^{-/-} Zellen wurde von dem der MLL^{+/+} Zellen „subtrahiert“. Bei dieser Subtraktion stellte die Wildtyp Zelllinie den sog. „tester“ und die *knock-out* Linie den „driver“ (im 100-fachen Überschuss) dar. Eine schematische Darstellung dieser Methode wird in Abb. 3.10 gezeigt.

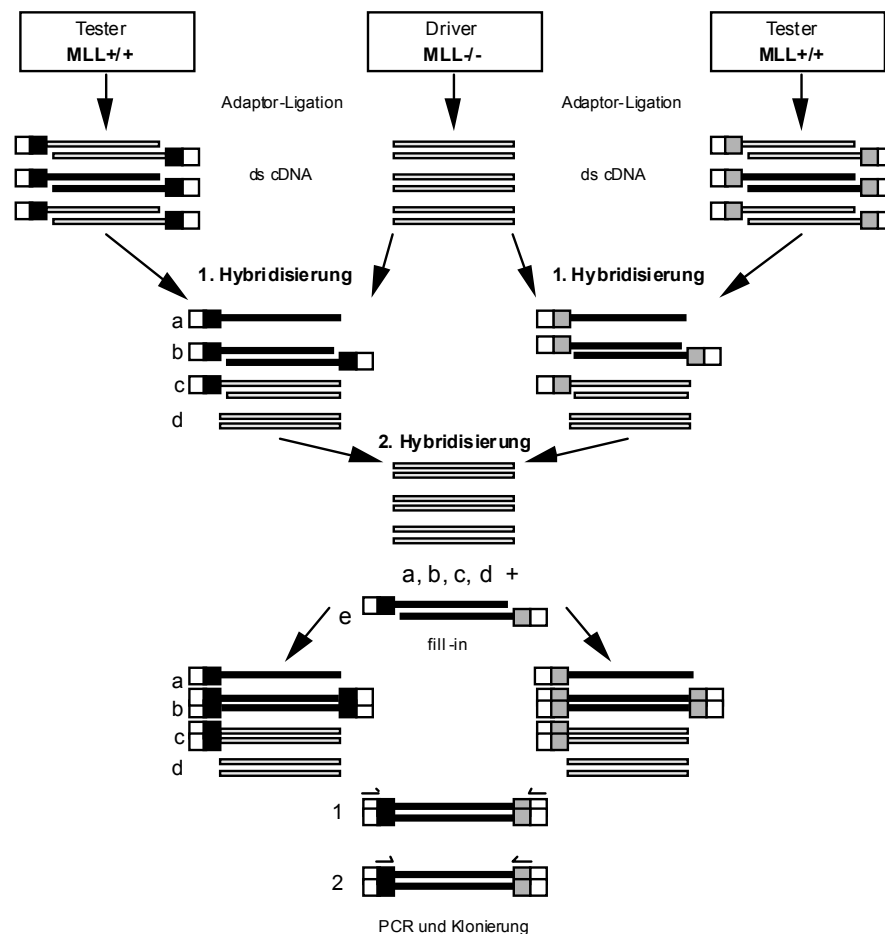


Abb. 3.10: Subtraktive Kloniertechnik von $MLL^{+/+}$ vs. $MLL^{-/-}$

Schwarze Balken stellen Transkripte dar, die ausschliesslich in den Wildtyp-Zellen exprimiert und weiße Balken, die in beiden Zelllinien exprimiert werden. Weiße Boxen beider Adaptoren sind sequenzidentisch, schwarze und graue Bereiche sind nicht sequenzidentisch.

Nachdem aus den isolierten mRNA-Fractionen der $MLL^{+/+}$ und $MLL^{-/-}$ Zelllinien doppelsträngige cDNA synthetisiert wurde, schloss sich ein *RsaI* Restriktionsverdau an. An die glatten Enden der *tester* cDNA ($MLL^{+/+}$) wurden in unterschiedlichen Reaktionsansätzen zwei verschiedene Adaptoren anligiert, die in nachfolgenden Hybridisierungsschritten die Anreicherung differentieller Transkripte durch suppressive und *nested*-PCR Schritte ermöglichte.

Die quantitative Adapterligation ist für den Erfolg dieser Kloniermethode von größter Bedeutung, da die Amplifizierung der differentiell exprimierten Transkripte von ihr abhängig ist. Aus diesem Grund ist eine Kontroll-PCR des *housekeeping* Gens *GAPDH* zur Bestimmung der Ligationseffizienz unerlässlich. Von beiden *tester*-Fractionen, die unter-

schiedliche Adaptoren enthalten sollten, wurden je zwei PCR-Reaktionen durchgeführt. Für die eine wurden zwei *GAPDH* Oligonukleotide, und für die andere je ein *GAPDH*-spezifischer und ein Adaptor-spezifischer Primer eingesetzt. Beide Banden sollten sich auf dem Agarosegel nicht mehr als um die dreifache Intensität unterscheiden, was für eine Adaptorligation mit mindestens 33%iger Ligationseffizienz spräche. Alle durchgeführten subtraktiven Klonierungen wurden diesen Anforderung gerecht. Beispielhaft ist dies für die Subtraktion $MLL^{+/+}$ vs. $MLL^{-/-}$ in Abb. 3.11 dargestellt.

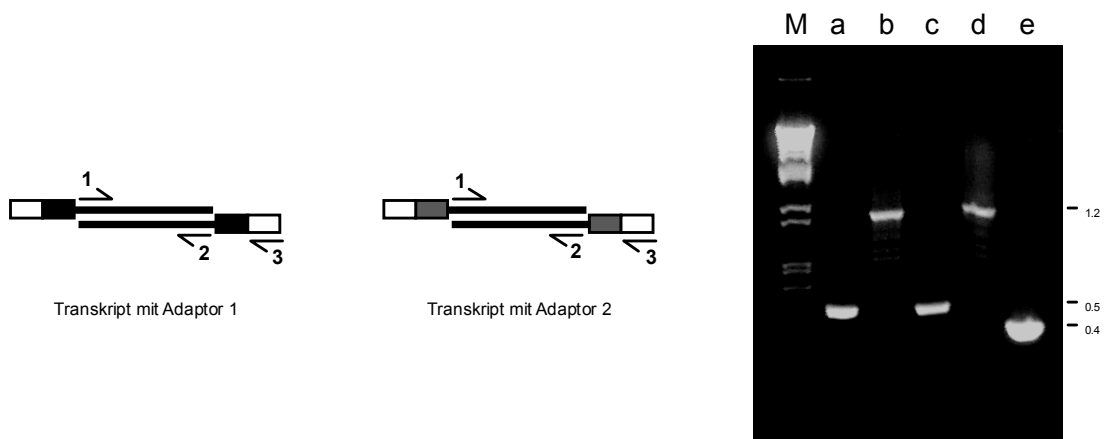


Abb. 3.11: Kontroll-PCR zur Bestimmung der Ligationseffizienz

a.: *tester* mit Adaptor 1 (*GAPDH*•3•5 (1/2)), b: *tester* mit Adaptor 1 (*GAPDH*•3/PCR-Primer (1/3)), c: *tester* mit Adaptor 2 (*GAPDH*•3•5 (1/2)), d: *tester* mit Adaptor 2 (*GAPDH*•3/PCR-Primer (1/3)), e: Positivkontrolle, M: Marker (λ ClaI), Größenangaben in kb.

Durch die Hybridisierungsschritte mit einem Überschuss an *driver* cDNA werden Transkripte, die in beiden Zelllinien vorkommen, aus dem Reaktionsansatz „entfernt“. Übrig bleiben die differentiell exprimierte cDNA Spezies, die durch kein komplementäres Transkript der *knock-out* Zellen durch Hybridisierung an der Amplifizierung gehindert werden.

Die Adapterligation mit zwei verschiedenen Linkern, deren Sequenz im äusseren Bereich (s. Abb. 3.11) identisch ist, hat folgende Vorteile: Nach den zwei Hybridisierungen entstehen die Produkte a-e, wobei ausschließlich das doppelsträngige cDNA Molekül e beide Adaptoren aufweist und exponentiell amplifiziert werden kann. Nach einer *fill-in* Reaktion kann mit Oligonukleotiden, die im äusseren Bereich beider Linker binden, eine PCR Reaktion durchgeführt werden. Transkript a und d können aufgrund der fehlenden zweiten Primerbindungsstelle nicht amplifiziert werden. Moleküle des Typs c werden nur linear amplifiziert. Theoretisch könnte auch Transkript b amplifiziert werden. Die Entste-

hung einer *pan-like* Struktur von Transkript b wird begünstigt, da die Adaptoren eines DNA-Stranges durch die *fill-in* Reaktion komplementär werden und die Adaptoren dieses DNA-Stranges mit sich selbst hybridisieren, also folglich nicht amplifiziert werden können. Mit dieser Strategie wurde durch die suppressive PCR-Reaktion eine Verminderung der Amplifikation von falsch positiven Transkripten erreicht. Es wird nun im zweiten Schritt eine *nested*-PCR angeschlossen, deren Primer spezifisch an die unterschiedlichen Sequenzen der zwei Adaptoren binden. Dadurch wird eine unspezifische Amplifikation, z.B. von Molekül b, weiter unterdrückt. Die Amplifikate der cDNA Spezies e wurden in den Vektor pGEM-T kloniert, der zur effizienten Ligation von PCR-Produkten mit einem terminalen Überhang einer einzelnen Thymin-Base ausgestattet ist (Fa. Promega).

Die subtraktive Klonierung wurde zudem in umgekehrter Richtung durchgeführt. Durch diese Subtraktion wurden Gene identifiziert, die durch die Anwesenheit des MLL Proteins reprimiert werden.

Nach Transformation der entsprechenden Ligationsansätze in den *E.coli* Stamm Sure™ konnten bei der Subtraktion von MLL^{+/+} vs. MLL^{-/-} ca. 165 Klone auf der Agar-Platte detektiert werden. Bei der Subtraktion von MLL^{-/-} vs. MLL^{+/+} waren es 70 Klone. Bei einer Wiederholung der Experimente resultierten 145 differentiell exprimierte MLL^{+/+} und 168 MLL^{-/-} Klone.

Von den meisten Klonen wurden Mini-Lysate angefertigt und diese sequenziert. Durch Anfrage am BLAST-Server (www.ncbi.nlm.nih.gov/blast) konnten viele Klone identifiziert werden. Diese sind in der Tabellen 3.12 und 3.13 zusammenfassend dargestellt.

ACC #	Zielgene (hochreguliert in MLL ^{+/+} Zellen)	Häufigkeit
XM_011188.2	Actin filament associated protein; AFAP-110	3x
NM_022310.1	Heat-shock protein 70; protein 5, HSP 70	3x
NM_010481.1	Heat-shock protein 74; HSP 74	3x
NM_010123	Eucaryotic translation initiation factor 3 (Eif3)	2x
AF095728	GDP-dissociation inhibitor	3x
J01420	Mouse mitochondrion, complete genome	2x
NM_010687	Glycosyltransferase	2x
BC013165.1	Ribosomal protein L9; Rpl9	2x
BC002167.1	Heat-shock protein 84; HSP 84	2x
NM_027086	Chromosom 9, nicht identifiziert	2x
AF364051.1	CLAST5, SHARP-2; hairy-related protein	1x
NM_009369	TGF-beta induced protein, (Tgfb1), beta(ig-h3)	1x
AF021335	T-cell-receptor gamma locus (TCR gamma-2/4)	1x
NM_019919	Latent TGF beta binding protein; Ltbp1	1x
NM_009270.1	Squalen epoxidase	1x
NM_013562	Interferon-related developmental regulator 1	1x
BC003893	Cell division cycle 5 protein	1x
AB022209	Ribonucleoporin F	1x
NM_010728	Lysyl oxidase; Lox	1x
BC005549	SAR-1 protein	1x
NM_010232	Flavin-containing monooxygenase 5 (Fmo5)	1x
AY026252	ADP-ribose pyrophosphatase; NUDT9	1x
NM_020092	S/R-rich splicing regulatory protein 86 (Srrp86)	1x
NM_008122	Gap junction membrane channel protein alpha 7 (Gja7)	1x
M11718	Procollagen, type V, alpha II	1x
NM_009933.1	Collagen, type VI, alpha I; Col6a1	1x
BC004725.1	Aldo-keto reductase family B1	1x
AF026481.1	Eucaryotic translation initiation factor 1a (Eif1a)	1x
NM_013647.1	Ribosomal protein S16	1x
AC021642.14	Ribosomal protein 23, containing myf5/myf6	1x
NP_008117.1	NADH ubiquinon reductase	1x
U23921	Osmotic stress protein 94 (Osp94), APG-1	1x
NM_011293	RNA Polymerase II 4 (14 kDa subunit)	1x
XM_030096	Malic enzyme (ME2)	1x
XM_144249	Vacuole membrane protein	1x
NM_020408	CGI-203 protein	1x
AK013434	Chromosom X, nicht identifiziert	1x
AK003172	Chromosom 15, nicht identifiziert	1x
AK013471	Chromosom 4, nicht identifiziert	1x
BC005559	Chromosom 6, nicht identifiziert	1x
BC004013	Chromosom 5, nicht identifiziert	1x
AK009972	Chromosom 10, nicht identifiziert	1x
AK009025	Chromosom 9, nicht identifiziert	1x
AK007636	Chromosom 5, nicht identifiziert	1x

Tab. 3.12: Zusammenfassung der Zielgene des murinen MLL Proteins

Charakterisierung durch subtraktive Kloniermethodik und anschließender Sequenzierung. Fettgedruckt dargestellt sind die Gene, die ebenfalls durch Microchip-Arrays identifiziert wurden. „Häufigkeit“: bedeutet Anzahl identifizierter Klone.

ACC #	Zielgene (hochreguliert in MLL ^{-/-} Zellen)	Häufigkeit
X93167	Fibronectin	5x
BC003825	Tubulin-β 5 (Tubb5)	3x
NM_009025	Ras-p21 activator-3 (Rasa-3)	3x
U68542	CASP (cux)	3x
NM_011619	Troponin T2 (cardiac; Tnnt2)	2x
BC006592	Rho GTPase activating protein	2x
AK005387	Chromosom 10, nicht identifiziert	2x
AK003401	Chromosom 14, nicht identifiziert	2x
XM_002669.2	NADH ubiquinon reductase	2x
BC003292	Apoptosis inducing factor (AIF)	1x
NM_011580	Thrombospondin (THBS1)	1x
XM_124321	Smooth muscle γ-actin, enteric	1x
XM_030096	Malic enzyme	1x
NM_011258	Replication factor (Recc1)	1x
NM_009983	Cathepsin D	1x
NM_009906	Ceroid-lipofuscinosis neuronal protein (Cln2)	1x
AJ276796	Cysteinyl-transferase (CysRS)	1x
AAH04756	Nuclear receptor binding protein	1x
NP_002152	Isoleucine-tRNA synthetase	2x
AV351990	Serine-threonine protein kinase (TAO1)	1x
AB053477	BPOZ	1x
NM_011125	Phospholipid transfer protein (Pltp)	1x
NM_015650	Microtubule-interacting prot. that assoc. with TRAF3	1x
NM_019094	Nudix (Nudt4)	1x
NM_015928	Androgen-induced prostate proliferation shut-off (AS3)	1x
NM_013786	Hydroxysteroid-17-β-dehydrogenase	1x
AK011390	Protein kinase, cAMP-dependent, type II α	1x
AK004221	Chromosom 8, nicht identifiziert	1x

Tab. 3.13: Zusammenfassung der durch das MLL Protein reprimierten Zielgene

Charakterisierung durch subtraktive Kloniermethodik und anschliessender Sequenzierung. Fettgedruckt dargestellt sind die Gene, die ebenfalls durch Microchip-Arrays identifiziert wurden. „Häufigkeit“ bedeutet Anzahl identifizierter Klone.

3.2.3.1 Differentielle Genexpression in MLL^{+/+} Zellen

Im Folgenden wird kurz auf interessante MLL Targetgene eingegangen. Die am häufigsten vorkommenden Targetgene, die durch die Anwesenheit des MLL Proteins in ihrer Transkription aktiviert werden, kodieren für verschiedene **Hitzeschock-Proteine** (HSPs). HSP70 und HSP74 wurden je 3 mal, HSP84 2 mal identifiziert. Zusätzlich konnte das Gen für *osmotic-stress protein 94*, das in seiner Funktion den HSPs ähnelt, gefunden werden.

Hitzeschock-Proteine sind bezüglich Geninduktion und Eigenschaften in allen bisher untersuchten Organismen hochkonserviert. Neben ihrer transienten Schutzfunktion für wichtige Proteine nach Hitzeeinwirkung, osmotischem Stress oder Schwermetallexposition, wird die Mehrzahl der HSPs ebenfalls in definierten Entwicklungsphasen konstitutiv synthetisiert und übt essentielle Funktionen in der Zelle aus (Nover, 1991; Pechan, 1991). Dazu gehören Proteinfaltungs- und Translationsprozesse (Miernyk and Hayman, 1996), Reorganisation des Zytoskelettes und die Autoregulation der HSP-Geninduktion. Die Vertreter der Proteinfamilie HSP70 bestehen aus einer N-terminalen ATP-bindenden Domäne und einer C-terminalen Wechselwirkungsdomäne, die eine Vielzahl ungefalteter oder denaturierter Proteine binden kann (Bukau and Horwich, 1998; Miernyk and Hayman, 1996). Die Regulation der Expression von Hitzeschock-Proteinen wird durch einen Autoregulationsmechanismus kontrolliert (Morimoto, 1993). Hitzeschock-Transkriptionsfaktoren sind an HSP70 Proteine gebunden und bilden eine monomere, inaktive Einheit. Durch Stress-einwirkung werden die Transkriptionsfaktoren freigesetzt und in den Zellkern transportiert, wo sie die HSP Transkription durch Bindung an entsprechende Promotorregionen initiieren. Die Herunterregulation der HSP Expression erfolgt, sobald ausreichende HSP Mengen zur Bindung der Transkriptionsfaktoren vorhanden sind.

In MLL^{+/+} Zellen wird das **Actin filament associated protein (AFAP 110)** stark hochreguliert. Die konservierte prolinreiche SH3-Domäne ist ein Substrat des Src-Protoonkogens (Reynolds et al., 1989) und führt nach seiner Bindung zu einer Konformationsänderung (Yu, 1994). Erhöhte intrazelluläre Src-Konzentrationen führen zu einer verstärkten Tyrosin-Phosphorylierung des Actin filament associated protein und Remodulierung des Aktins in eine rosettenähnliche Struktur (Flynn et al., 1993). Die im AFAP-Protein vorhandenen pleckstrin-homology (PH) Domänen sind charakteristische Merkmale von Zytoskelett-Proteinen (Gibson et al., 1994; Musacchio et al., 1993) und interagieren möglicherweise mit der Plasmamembran. Zwischen den beiden PH-Domänen liegt eine Serin-/Threonin-reiche Region, die als Phosphorylierungssubstrat für Proteinkinase C postuliert wird. Am C-terminalen Ende befindet sich ein Leucin-Zipper Motiv und zwei Aktin-bindende Domänen (ABD1 und ABD2; (Qian et al., 1998). Mit Hilfe von Immunfluoreszenz-Aufnahmen konnte bewiesen werden, dass AFAP110 tatsächlich mit Aktinfilamenten assoziiert ist (Flynn et al., 1992). Aktinfilament-bindende Proteine spielen bei der Stabilisierung und Reorganisation des Zytoskeletts eine ausserordentlich wichtige Rolle.

CLAST5/SHARP2/Hairy-related protein/DEC1 ist ein *basic-helix-loop-helix* Protein der HES (Hairy and Enhancer of Split) Proteinfamilie. Das Hairy-related protein ist bei *D. melanogaster* in der frühen Entwicklung für die Neurogenese und die Ausbildung sensorischer Organe von großer Bedeutung, da es über die direkte Transkriptionsrepression des proneuralen Gens *achaete* regulierend eingreift (Shen et al., 1997). Ebenfalls konnte durch Northern-Blot Analysen gezeigt werden, dass dieses Protein in unterschiedlichen Geweben (Lunge, Milz, Intestinaltrakt und Knorpel) exprimiert wird. Es wird vermutet, dass Hairy-related protein an der Kontrolle von Zelldifferenzierungsprozessen beteiligt ist und somit eine essentielle Rolle während der Embryonalentwicklung spielt.

Als weiteres Gen, das in *Mll* Wildtyp-Zellen differentiell exprimiert wird, wurde **Latent TGF- β binding protein (LTBP1)** identifiziert. LTBP1 besitzt Matrix- und TGF- β -Bindedomänen und ist, durch kovalente Bindung an Strukturen der extrazellulären Matrix, in der Lage, TGF- β gezielt seinen Rezeptoren zuzuführen (Taipale et al., 1994). Dabei wird das inaktive, gebundene TGF- β durch Proteolyse aus der „Speicherform“ entlassen und in einen aktiven Zustand überführt. Es wird vermutet, dass LTBP1 zusätzliche Schutzfunktionen vor proteolytischem Abbau durch Matrix-Proteasen ausübt, während der Proteinkomplex mit der Zelloberfläche assoziiert ist (Flaumenhaft et al., 1993).

Durch die Anwesenheit des MLL Proteins werden bei Wildtyp-Zellen, im Gegensatz zu den *knock-out* Zellen, verstärkt Proteine exprimiert, die an der Bildung und/oder Aufrechterhaltung des Zytoskeletts verantwortlich sind. Dazu gehören **Actin filament associated protein; Gap junction membrane channel protein α ; Procollagen type V, α II; Procollagen type VI, α I und TGF- β induced protein (Beta ig-h3). TGF- β induced protein** bindet Typ I, II und IV Kollagen und spielt als extrazelluläres Adhäsionsprotein eine wichtige Rolle bei Zell-Kollagen-Interaktionen. Es kann mit Mikrofibrillen und Zelloberflächen assoziiert sein. Für die Knochenbildung scheint **Beta ig-h3** ebenfalls von Bedeutung zu sein und wurde zusätzlich durch DNA-Microchip-Hybridisierung in MLL^{+/+} Zellen (ebenso wie Col6a3) identifiziert.

3.2.3.2 Differentielle Genexpression in MLL^{-/-} Zellen

Die MLL^{-/-} Zellen wiesen eine stark erhöhte Expressionsrate an **Fibronectin** (FN) auf. Die subtraktive Klonierung führte zur Identifizierung von 5 Klonen, die den Vektor mit der Sequenz für dieses Glykoprotein trugen. FN wurde auf Zelloberflächen als Teil der extrazellulären Matrix und im gelösten Zustand im Zytoplasma gefunden. Dort interagiert es mit einer Vielzahl von Makromolekülen (Fibrin, Kollagen, Heparin, Integrine u.v.a.m.), die ebenfalls für die Aufrechterhaltung der extrazellulären Matrix (ECM), für das Komplementsystem und für die Stabilität von Oberflächen-Rezeptoren bedeutsam sind (Hynes, 1990, in „Fibronectins“, NY, Springer-Verlag). FN ist folglich in Wundheilungsprozesse, Embryogenese, Blutgerinnung und hauptsächlich in Zellmigration/-adhäsion involviert. Die Wichtigkeit Fibronectins während der Embryogenese konnte bereits nachgewiesen werden. Es spielt eine besonders bedeutende Rolle bei der mesodermalen Zellmigration in der Gastrulationsphase, der Migration präkardialer Zellen, bei der Zelladhäsion von Knorpel- und Knochenzellen, sowie bei der Neurogenese. *Integrin $\alpha 7$* (Fibronectin-Rezeptor) wurde mittels DNA-Microchip-Hybridisierungen in MLL^{-/-} Zellen identifiziert, in denen es um den Faktor 4,6 hochreguliert vorlag.

Bei dem als **CASP** bezeichneten Protein handelt es sich um ein alternatives Spleiss-Produkt des evolutionär konservierten *CDP* Gens (CCAAT displacing protein). Dieses gehört zu den Homeobox-enthaltenden Transkriptionsrepressoren der *Cux* Familie. Forschungsarbeiten an *D. melanogaster* haben mittels *knock-out* Experimenten gezeigt, dass Cut (CDP-Homolog) für die Lungen- und Haarfollikelentwicklung essentiell ist. Zusätzlich wurde seine Involvierung in die Prozesse der terminalen Differenzierung und Kontrolle der *nuclear matrix attachment regions* bestätigt (Ellis et al., 2001). Es wurde eine verstärkte Expression während der S-Phase detektiert, die zu der Vermutung führt, dass Cut eine wichtige Rolle in der Zellzyklus-Progression spielt und für DNA-Replikation essentiell ist (Martin-Soudant et al., 2000). Bei einer Vielzahl von Krebserkrankungen ist dieses Gen rearrangiert (Nepveu, 2001).

Rho-GTPase activating protein reguliert u.a. die Organisation von Aktinfilamenten, die sich durch den Einfluss der Rho-GTPase in Form von Lamellipodien/“*ruffles*“ formieren.

Thrombospondin-1 (THBS-1), ein hochmolekulares Glykoprotein, ist ein Hauptbestandteil der extrazellulären Matrix von Endothel- und Tumorzellen. *THBS-1* wird in malignen Geweben stark überexprimiert und ist in hohen Konzentrationen im Plasma von Krebspatienten nachweisbar. Zusätzlich fördert es Adhäsion und Invasion von metastasierenden Zellen (Qian and Tuszynski, 1996).

3.2.3.3 Verifizierung identifizierter Gene durch Dot-Blot Hybridisierung

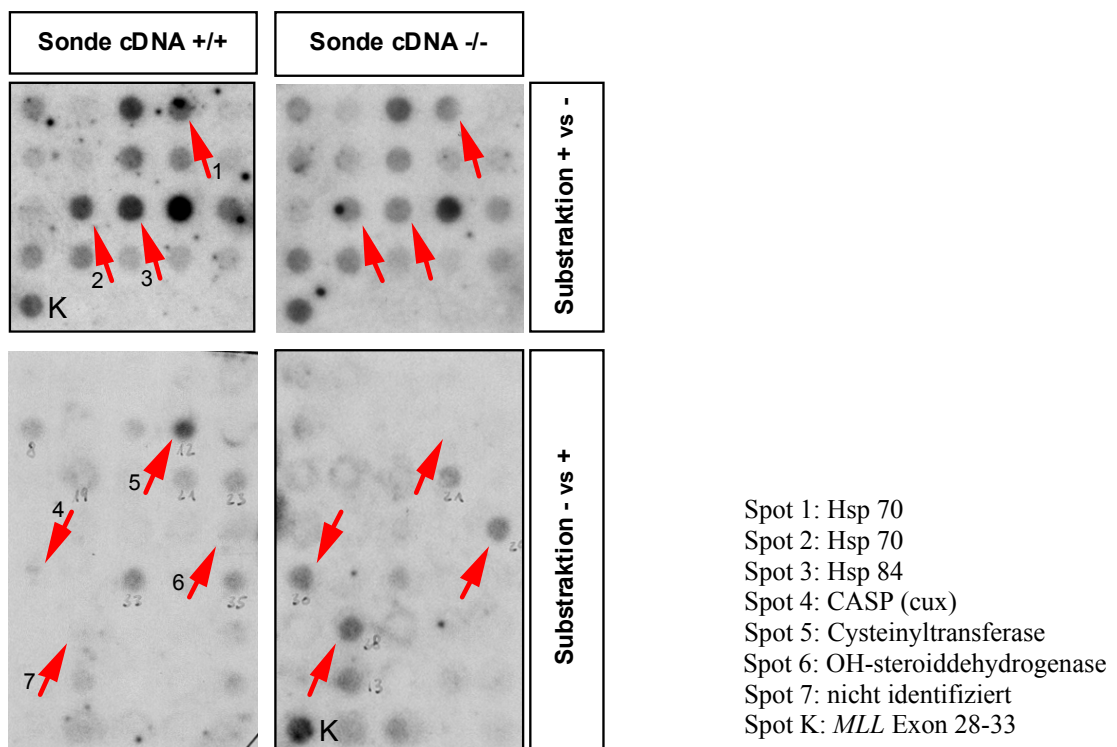


Abb. 3.14: Ergebnisse der Dot-Blot Hybridisierung

Die Pfeile markieren Spots mit unterschiedlicher Signalstärke. Die Filter der ersten Spalte wurden mit radioaktiv markierter cDNA der Wildtyp-Zellen und die in der zweiten Spalte mit markierter cDNA der *knock-out* Zellen hybridisiert. Die oberen kleinen Filter repräsentieren Transkripte, die ausschliesslich in den *MLL*^{+/+} Zellen exprimiert werden. Auf die unteren Arrays wurden cDNAs der differentiell exprimierten Transkripte der *MLL*^{-/-} Zellen aufgespottet. Die Positivkontrolle K sollte ausschliesslich auf den Filtern, hybridisiert mit der cDNA *MLL*^{+/+} Sonde, ein Signal ergeben.

Zur Verifizierung der identifizierten Targetgene wurden Dot-Blot Experimente durchgeführt. Als vorteilhaft erweist sich im Gegensatz zur Kolonie-Hybridisierung, dass bei der Dot-Blot Methode ausschliesslich DNA und kein Bakterienlysat auf der Nylon-Membran vorhanden ist. Insofern sollten unspezifische Bindungen und folglich Hintergrund-Signale weiterhin unterbunden werden. Die Durchführung ist zudem wesentlich zeitsparender als Northern- oder Southern-Blots, da der Gelelektrophoreseschritt entfällt. Ein Vorteil dieser Methode besteht in der Begrenzung der aufgetragenen Proben auf streng definierte Bereiche der Membran, die einen besseren Vergleich der Intensitäten von Hybridisierungssignalen ermöglicht. Auch bei dieser Hybridisierungsmethode wurden ausschliesslich radioaktiv markierte cDNA Sonden eingesetzt, die eine vergleichbare Intensität aufwiesen. Die zwei identischen Filter wurden entweder mit der radioaktiven MLL^{+/+} cDNA oder MLL^{-/-} cDNA hybridisiert. Anhand der Resultate, dargestellt in Abb. 3.14, wird ersichtlich, dass die Ergebnisse in keinsten Weise den Erwartungen entsprachen. Optimalerweise hätten ausschliesslich Signale auf dem oberen linken und unteren rechten Filter detektiert werden sollen. Bei vereinzelt Spots gab es Intensitätsunterschiede. Mehrere Wiederholungen der Hybridisierungsexperimente erbrachten kein eindeutig auswertbares Ergebnis.

Eine hervorzuhebende Besonderheit zeigte sich in der Kontrolle. Es wurde ein PCR-Produkt von Exon 28-33 des *Mll* Gens aufgespottet, das bei der Hybridisierung mit der MLL^{+/+} Sonde als Positivkontrolle und mit der MLL^{-/-} Sonde als Negativkontrolle dienen sollte. Bei der Herstellung der MLL^{-/-} Zellen wurde in das *Mll* Gen eine Kasette inseriert, die die Transkription in Exon 4 terminieren sollte. Demnach dürfte in den MLL^{-/-} Zellen kein Transkript aus dem hinteren Bereich des Gens existieren. Erstaunlicherweise ergab die Hybridisierung mit markierter cDNA der MLL^{-/-} Zellen ebenfalls ein Signal, dessen Intensität die der eigentlichen Positivkontrolle z.T. weit übertraf. Vermutungen, es könnte sich um eine Verwechslung der Sonden handeln, wurden durch weitere PCR-Analysen der cDNA Sonden widerlegt. Der Test auf gegenseitige Kontamination beider Zelllinien fiel ebenfalls negativ aus (s. Abb. 3.15). Dieses Phänomen, das zugleich die Ergebnisse der Dot-Blots in Frage stellte, wurde in weiteren Experimenten untersucht und aufgeklärt (s. Kapitel 3.6).

3.2.4 DNA-Microchip-Arrays

Aufgrund der Ergebnisse dieser subtraktiven Klonierungsstrategie wurden DNA-Microchip-Hybridisierungen durchgeführt, die ebenfalls zu einer Charakterisierung der Transkriptome beider Zelllinien führen sollte.

Bevor jedoch die erste DNA-Microchip-Hybridisierung durchgeführt wurde, wurden die Ausgangszelllinien $MLL^{+/+}$ und $MLL^{-/-}$ erneut auf ihren Phänotyp und *Mll* Expression getestet. Die Ergebnisse der RT-PCR und PCR-Reaktionen an gDNA sind in Abb. 3.15 zusammengefasst. Die Integrität der RNA wurde auf einem denaturierenden Agarosegel überprüft (s. Abb. 3.26, Kapitel 3.4).

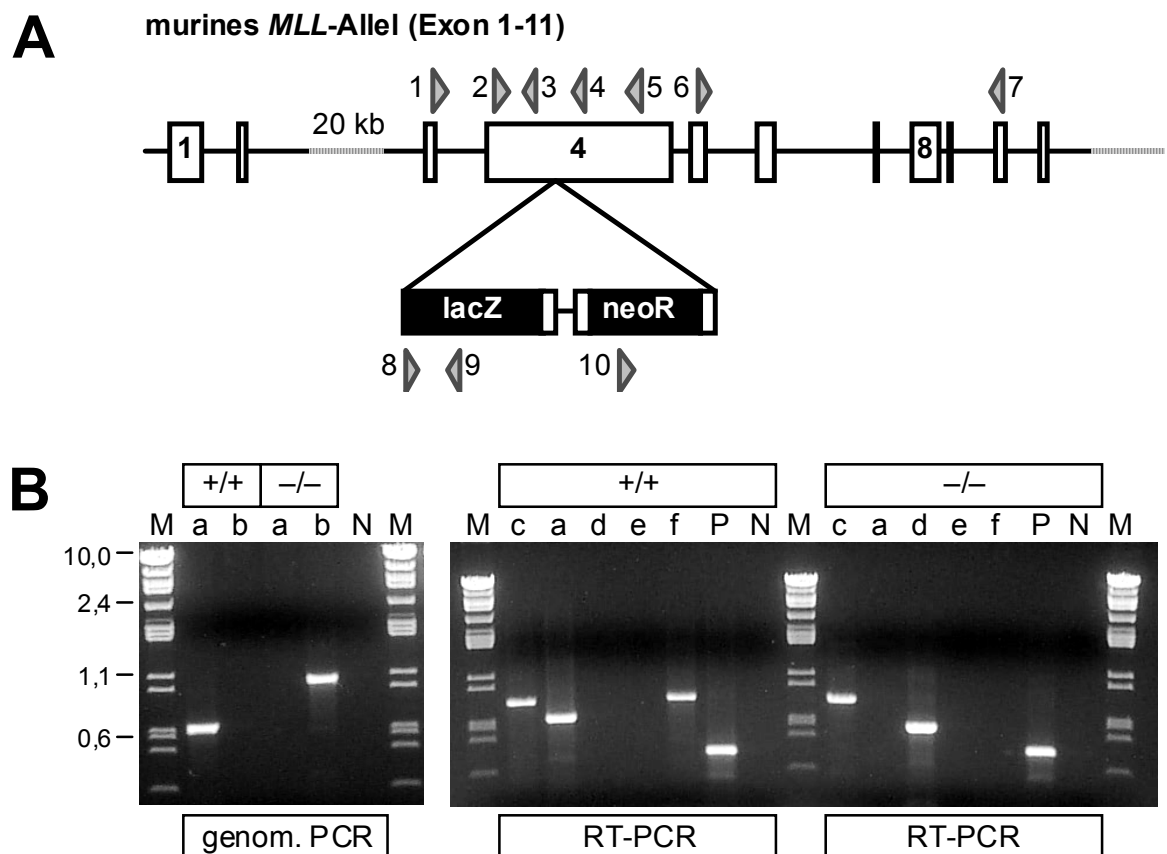


Abb. 3.15: Integrität und Expression von *Mll* Wildtyp und *knock-out* Allelen

A: Schema des *Mll* Gens Exon 1-11. Durch Insertion des *LacZ* Gens in Exon 4 entstand die *Mll knock-out* Zelllinie. Eine Serie von Oligonukleotiden (1-10) wird durch Dreiecke dargestellt. **B:** genomische und RT-PCR Analysen. Linkes Bild: PCR an genomischer DNA von Wildtyp und *knock-out* Zellen; rechtes Bild: RT-PCR Analysen. Die Bande in Spur d weist die insertierte *LacZ*-Kassette nach, während die Bande in Spur f die endogene *Mll* Expression von Exon 5-10 zeigt und für $MLL^{+/+}$ Zellen spezifisch ist. Spalte a: Primer 2/4 (689 bp), b: 2/9 (1060 bp), c: 1/3 (837 bp), d: 8/9 (550 bp), e: 10/5, f: 6/7 (833 bp), P: GAPDH•3/•5 (432 bp), N: Wasserkontrolle, M: Marker, Größenangabe in kb.

Als Ausgangsmaterial für die DNA-Microchip-Hybridisierung dienten RNA Fraktionen der Zelllinien MLL^{+/+} und MLL^{-/-} (Experiment 1). Zur Sicherstellung eines normalisierten Genexpressionsprofils wurden je drei unabhängige RNA Präparationen verschiedener Zeitpunkte zur einer Fraktion vereinigt. Nach 6 Monaten wurde das Experiment mit *ge-poolter* RNA aus Leervektor-transfizierten MLL^{+/+} und MLL^{-/-} Zellen wiederholt (Experiment 2). Der Leervektor pBC230A liegt nach Transfektion, aufgrund der EBNA-1 Sequenz, episomal vor und sollte theoretisch keinen Einfluss auf das Transkriptionsmuster nehmen. Nach dem „*standard-one-round amplification protocol*“ (s. Kapitel 2.5) wurde die RNA in biotinylierte *antisense*-RNA umgeschrieben und mit den murinen DNA-Microchips *MG_U744v2* (Fa. Affymetrix), die ca. 12.000 Gene enthalten, hybridisiert. Die nachfolgenden Daten beruhen auf den Ergebnissen des ersten Hybridisierungsschritts, d.h. nachdem die biotinylierte cRNA zum ersten Mal mit Streptavidin-markiertem Phycocerythrin versetzt und vermessen wurde.

Für die Hybridisierung mit MLL^{+/+} und MLL^{-/-} cDNA wurde je ein DNA-Microchip verwendet. Aufgrund der Methodik können ausschliesslich Gene identifiziert werden, die ein ausreichendes Fluoreszenz-Signal ausstrahlen. D.h., es besteht nur die Möglichkeit, transkribierte Gene zu detektieren. Die erhaltenen Datensätze von MLL^{+/+} und MLL^{-/-} Zellen wurden anschliessend miteinander verglichen und Differenzen aus den Signalintensitäten gebildet. Gene, die differentiell in MLL^{+/+} Zellen vorkommen (stärker, oder ausschliesslich in dieser Zelllinie exprimiert werden), werden bezüglich der *knock-out* Zellen als hochreguliert bezeichnet. Für MLL^{-/-} gilt das gleiche *vice versa*.

Die DNA-Microchip-Hybridisierungen wurden von Dr. L. Klein-Hitpass, Universitätsklinikum Essen, durchgeführt. Zur Datenanalyse wurden diejenigen Gene herangezogen, die sowohl im ersten als auch im zweiten Experiment differentiell exprimiert vorlagen. Aus dem *signal log₂ ratio* der Gene beider Experimente wurden Mittelwerte errechnet, die für das *ranking* (s. Tab. 3.16 und 3.17) ausschlaggebend waren. Nach dem Anlegen höchster Signifikanz-Kriterien (*increase/decrease* (A), *signal log₂ ratio* > 1,32 oder < -1,32 (B) und *P-value* > 0,999 oder < 0.001 (C)) wurden 197 differentiell transkribierte Gene identifiziert. Davon wurden 136 Gene in MLL^{-/-} Zellen und 61 Targetgene im *Mll* Wildtyp-Zellen hochreguliert. Eine Liste der 197 Gene befindet sich im Anhang.

3.2.4.1 Scatter-Blots

Die Ergebnisse der DNA-Microchip-Hybridisierungen sind in Scatter-Blots graphisch dargestellt (Abb. 3.16). In Abb. A sind auf der Ordinate die Signalstärken aus der Hybridisierung mit Phycoerythrin-markierter *antisense*-RNA der $MLL^{-/-}$ Zellen und auf der Abszisse die der $MLL^{+/+}$ Zellen aufgetragen. In Abb. B handelt es sich um die markierten Proben der Leervektor-transfizierten Wildtyp- und *knock-out* Zelllinien. Die einzelnen farbigen Punkte stellen je ein detektiertes Gen der Chips *MG_U74Av2* dar. Die diagonal angeordneten Geraden grenzen die Bereiche der 2,5-, 3-, 5- und 10-fachen Hochregulation ab.

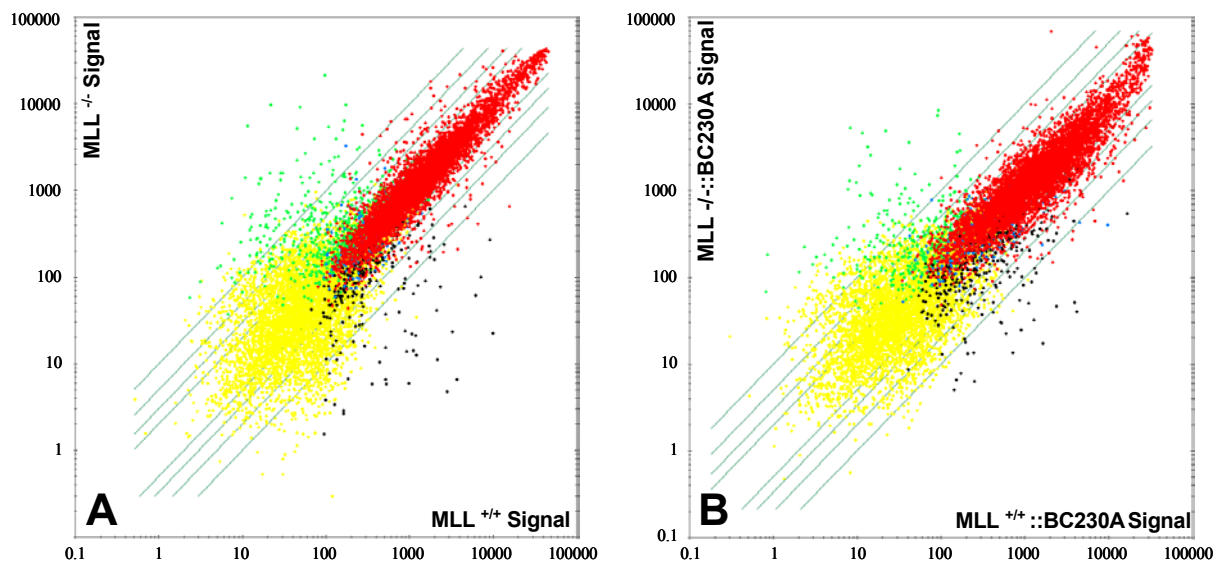


Abb. 3.16: Scatter-Blots ($MLL^{+/+}$ vs. $MLL^{-/-}$, bzw. $MLL^{+/+}::BC230A$ vs. $MLL^{-/-}::BC230A$)

Die grün markierten Punkte stehen für überexprimierte Gene der *knock-out* Zellen, die schwarzen für überexprimierte Gene der Wildtyp-Zellen. Die Expressionsstärken zahlreicher Gene zeigen erhebliche Unterschiede.

Der Pearson-Korrelationskoeffizient, der aus den *signal log₂ ratios* berechnet wurde, beträgt für den Vergleich der $MLL^{+/+}$ und $MLL^{-/-}$ Expressionsstärken aller Gene $r = 0,8859$ und für den Expressionsvergleich der leervektortransfizierten $MLL^{+/+}$ und $MLL^{-/-}$ Zellen $r = 0,8853$. Hierbei handelt es sich um einen relativ kleinen Wert, was bedeutet, dass sich die Genexpressionsmuster beider Zelllinien signifikant unterscheiden. Dies wird besonders an der weiten Streuung einer großen Anzahl der abgebildeten „Gen-Punkte“ um die Re-

gressionsgerade, die mittig zwischen den beiden 2,5-fachen Begrenzungslinien liegt, deutlich.

Die gelb markierten, und die anderen im Bereich unter Signalstärke 200 liegenden Gene, wurden von der Analyse ausgeschlossen. Entsprechende Transkripte waren entweder in beiden Analysen nicht oder nur marginal vorhanden. Es besteht ebenfalls die Möglichkeit, dass sich beide Transkripte in ihrer Expressionsstärke erheblich (durch einen hohen *signal log₂ ratio* Wert) unterscheiden, sich die Intensität jedoch zu nah am Rausch-Signal (Bereich < 200) befindet. In diesem Fall ist keine gesicherte Aussage über Fluoreszenzintensitäten möglich. Die als rote Spots dargestellten Gene waren in beiden Transkriptomen vorhanden, unterschieden sich jedoch zum größten Teil kaum in ihrer Expressionsstärke. Die meisten dieser Gene wurden bis maximal um den 2,5-fachen Wert differentiell exprimiert. Dieser Wert wurde von uns als unterste Signifikanzgrenze definiert.

Von großer Bedeutung sind die Gene, die grün und schwarz markiert sind. Die grün markierten wurden ausschliesslich in den *knock-out* Zellen, und die schwarz markierten in den Wildtyp-Zellen exprimiert. Wenn man z.B. die grün markierten Gene betrachtet, liegen sie auf der Abszisse in einem Bereich < 200, d.h., dass das Signal durch Hybridisierung mit *MLL^{+/+} antisense-RNA* im Bereich des Hintergrund-Rauschens liegt und das Gen somit als nicht vorhanden definiert werden kann. Je weiter sie von der Regressionsgerade entfernt liegen, umso höher ist ihr *signal log₂ ratio* Wert.

Aus dieser Graphik ist nun abzuleiten, dass ca. 60 Gene (grün) in *MLL^{-/-}* Zellen und ca. 35 Gene (schwarz) in *MLL^{+/+}* Zellen exklusiv exprimiert werden.

Diese Art der Darstellung ermöglicht einen schnellen Überblick über Expressionsunterschiede zweier Transkriptome, ohne jedoch die einzelnen Gene zu identifizieren.

3.2.4.2 GeneSpring[®] Auswertung

Neben der Scatter-Blot Darstellung gibt es eine weitere Möglichkeit, Ergebnisse einer DNA-Microchip-Hybridisierung graphisch zu präsentieren. Die Auswertung erfolgt mittels Verarbeitung der Hybridisierungsrohdaten durch die Software GeneSpring[®] (Fa. Silicon-Genetics). Dabei wird jedes einzelne der 12.000 Gene als Gerade dargestellt. Die Steigung und Farbe der Geraden ist ein Mass für die Stärke der Über- oder Unterexpression. Die unter der Abbildung gezeigte Farbskala bezieht sich auf die Expression der Wildtyp-Zellen. Demnach werden alle in den *MLL^{+/+}* Zellen überexprimierten Gene als rote, alle

unterexprimierten als blaue und die nicht beeinflussten Gene als gelbe Linien dargestellt. Auf der Ordinate sind die Signalstärken als Exponentialfunktion, auf der Abszisse sind die beiden untersuchten Zelllinien dargestellt.

Die gezeigten Gene wurden einer Normierung unterzogen. D.h., es wurden software-spezifische mathematische Berechnungen (Statistikberechnungen, Algorithmen, u.ä.) durchgeführt, die sich positiv auf die Signifikanz auswirken.

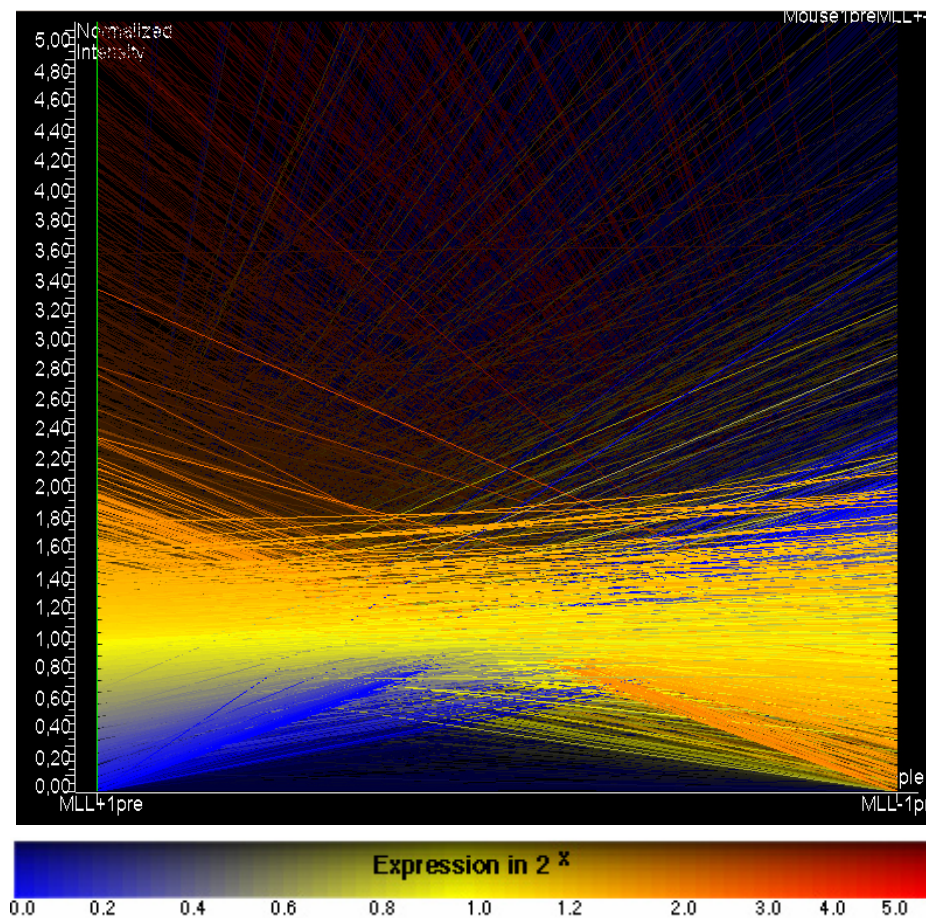


Abb. 3.17: GeneSpring® Auswertung des Expressionsmuster-Vergleichs von MLL^{+/+} und MLL^{-/-} Zellen
 Auf der linken Seite werden die Signalstärken (\log_2) der Wildtyp und auf der rechten Seite die der *knock-out* Zellen dargestellt. Die farbige Expressionsskala bezieht sich auf die Expressionsstärke der MLL^{+/+} Zellen.

Die Graphik zeigt deutlich, dass es zahlreiche rote und blaue Geraden gibt, die durch ihre extreme Steigung im „Gesamtpool“ auffallen. Eine Identifikation dieser Gene ist mittels der Darstellung nicht möglich. Das Resultat stimmt mit den erhaltenen Hybridisierungsdaten (und der quantitativen Verteilung der Gene im Scatter-Blot) überein.

In der folgenden Abbildung sind die Resultate beider DNA-Microchip-Hybridisierungen als *heatmap* graphisch dargestellt. Die jeweils linke Spalte repräsentiert das Ergebnis aus dem Expressionsvergleich $MLL^{+/+}$ vs. $MLL^{-/-}$, die rechte $MLL^{+/+}::BC230A$ vs. $MLL^{-/-}::BC230A$. Anhand der Farbintensitäten ist ersichtlich, dass die Resultate des ersten Experimentes verifiziert werden konnten und sich die Expressionsstärken der einzelnen Gene nur marginal unterscheiden. Dargestellt sind jeweils die 40 Gene, die entweder durch die Abwesenheit des MLL Proteins am stärksten (um den Faktor 2,5 oder höher) hochreguliert wurden (rot), oder durch seine Anwesenheit (grün).

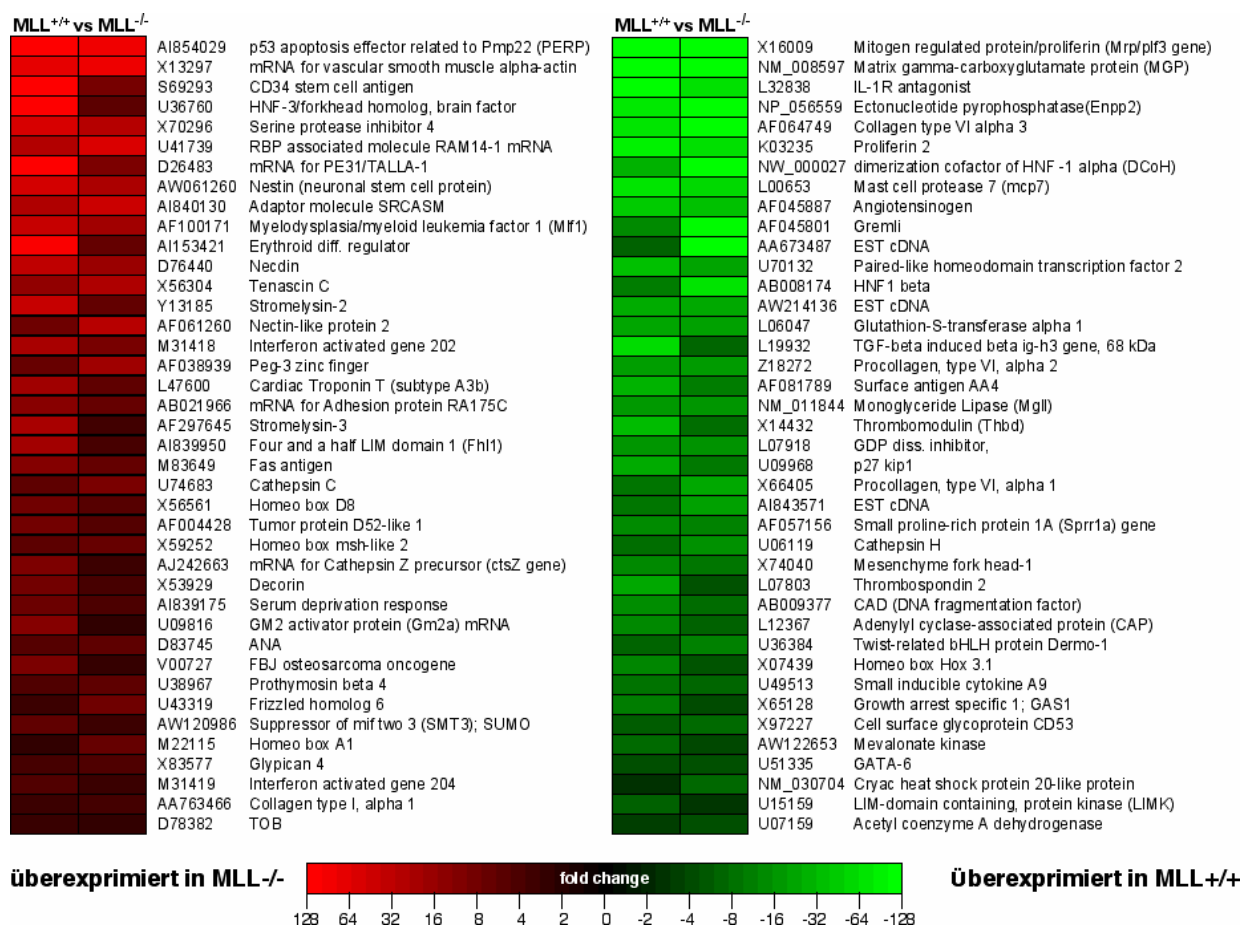


Abb. 3.18: *Heatmaps* der Genexpressionsprofile von $MLL^{+/+}$ und $MLL^{-/-}$ Zellen

Die Top 40 exprimierten Gene beider Zelllinien aus dem 1. und 2. Hybridisierungsexperiment werden in *heatmaps* dargestellt. Die mit der roten Skala versehenen Gene werden in $MLL^{-/-}$ Zellen, die mit der grünen Skala in $MLL^{+/+}$ Zellen hochreguliert. Gen-Accession-Codes und Genbezeichnungen identifizieren die entsprechenden Kästchen. Der Farbcode repräsentiert 8 \log_2 Schritte.

AAC #	Zielgene (hochreguliert in MLL ^{+/+} Zellen)	sig log ₂ ratios	
X16009	Mitogen regulated protein/proliferin (Mrp/plf3 gene)	9,23	6,61
NM_008597	Matrix γ -carboxyglutamate protein (MGP)	6,34	6,56
L32838	IL-1R antagonist	9,10	3,26
NP_056559	Ectonucleotide pyrophosphatase/phosphodiesterase 2 (Enpp2)	8,85	2,53
AF064749	Collagen type VI α 3 (Col6A3)	5,92	4,99
K03235	Proliferin	4,89	6,01
NW_000027	Dimerization cofactor of HNF -1 α (DCoH)	7,22	3,43
L00653	Mast cell protease 7 (mcp7)	5,81	4,70
AF045887	Angiotensinogen	4,83	5,58
AF045801	Gremlin	5,45	4,43
AA673487	EST cDNA	6,87	2,75
U70132	Paired-like homeodomain transcription factor 2/RIEG/ARP1/Ptx2	5,25	4,23
AB008174	HNF-1 β	4,00	4,78
AW214136	EST cDNA	5,45	2,70
L06047	Glutathion-S-transferase α 1	3,03	5,04
L19932	TGF-beta induced Beta ig-h3 , 68 kDa	4,62	3,31
Z18272	Procollagen, type VI, α 2	2,83	4,37
AF081789	Surface antigen AA4	4,40	2,62
NM_011844	Monoglyceride lipase (Mgll)	3,75	2,79
X14432	Thrombomodulin (Thbd)	4,62	1,79
L07918	GDP diss. inhibitor, preferentially expr. in hematopoietic cells	4,45	1,95
U09968	p27 kip1	3,63	2,76
X66405	Procollagen, type VI, α 1	2,59	3,35
AI843571	EST cDNA	3,10	2,37
AF057156	Small proline-rich protein 1A (Sprr1a)	3,13	2,32
U06119	Cathepsin H	2,53	2,81
X74040	Mesenchyme forkhead-1	3,44	1,66
L07803	Thrombospondin 2	3,13	1,95
AB009377	CAD (DNA fragmentation factor)	2,95	2,09
L12367	Adenylyl cyclase-associated protein (CAP)	3,67	1,34
U36384	Twist-related bHLH protein Dermo-1	2,36	2,60
X07439	Homeo box Hox 3.1	3,40	1,47
U49513	Small inducible Zytokine A9	2,21	2,60
X65128	Growth arrest specific 1; GAS1	1,62	3,05
X97227	Cell surface glycoprotein CD53	2,66	1,66
AW122653	Mevalonate kinase	1,37	2,78
U51335	GATA-6	1,93	2,17
NM_030704	Cryac heat shock protein 20-like protein	2,23	1,61
U15159	LIM-domain containing, protein kinase (LIMK)	1,64	1,88
U07159	Acetyl coenzyme A dehydrogenase	1,55	1,39

Abb. 3.19: Genexpressionsmuster der Mll Wildtyp-Zellen

Zusammenfassung der Top 40 Gene, die durch das MLL Protein aktiviert werden, incl. Gen-Accession-Code, Genbezeichnung und signal log₂ ratios beider Experimente.

AAC #	Zielgene (hochreguliert in MLL ^{-/-} Zellen)	sig log ₂ ratios	
AI854029	p53 apoptosis effector related to Pmp22 (PERP)	8,57	9,11
X13297	vascular Smooth muscle α -actin (vSMA)	7,86	7,97
S69293	CD34 stem cell antigen	7,49	6,22
U36760	HNF-3/forkhead homolog, Brain factor 1(BF-1)	6,45	7,21
X70296	Serine protease inhibitor 4 (PN-1)	6,32	7,31
U41739	RBP associated molecule RAM14-1	6,68	6,18
D26483	PE31/TALLA-1(A15)	4,92	7,87
AW061260	Nestin (neuronal stem cell protein)	6,33	6,01
AI840130	Adaptor molecule SRCASM (Src activating signaling molecule)	5,62	5,33
AF100171	Myelodysplasia/Myeloid leukemia factor 1 (Mlf1)	3,83	6,94
AI153421	Erythroid differentiation regulator	2,73	7,18
D76440	Necdin	5,32	4,54
X56304	Tenascin C	3,47	6,29
Y13185	Stromelysin-2 (MMP10)	4,75	4,67
AF061260	Nectin-like protein 2	4,49	4,45
M31418	Interferon activated gene 202	6,09	2,83
AF038939	Peg-3 zinc finger	4,38	4,29
L47600	Cardiac Troponin T (subtype A3b)	4,95	3,48
AB021966	Adhesion protein RA175C	4,23	4,15
AF297645	Stromelysin-3 (MMP11)	5,20	3,06
AI839950	Four and a half LIM domain 1 (Fhl1)	4,17	4,08
M83649	Fas antigen	4,74	3,37
U74683	Cathepsin C	3,24	4,65
X56561	Homeo box Hox D8	3,29	4,47
AF004428	Tumor protein D52-like 1	3,86	3,60
X59252	Homeo box msh-like 2 (hox 8.1)	3,07	4,02
AJ242663	Cathepsin Z precursor (ctsZ gene)	3,81	3,27
X53929	Decorin	4,64	2,28
AI839175	Serum deprivation response	3,90	3,00
U09816	GM2 activator protein (Gm2a)	3,79	2,73
D83745	ANA	2,79	3,58
V00727	FBJ osteosarcoma oncogene	3,71	2,34
U38967	Prothymosin β 4	3,21	2,83
U43319	Frizzled homolog 6	3,47	2,09
AW120986	Suppressor of mif two 3 (SMT3); SUMO	2,58	2,96
M22115	Homeo box Hox A1	2,96	2,03
X83577	Glypican 4	2,21	2,23
M31419	Interferon activated gene 204	2,42	1,42
AA763466	Collagen type I, α 1	2,71	1,53
D78382	TOB	1,74	2,20

Abb. 3.20: Genexpressionsmuster der Mll knock-out Zellen

Zusammenfassung der Top 40 Gene, die durch das MLL Protein reprimiert werden, incl. Gen-Accession-Code, Genbezeichnung und signal log₂ ratios beider Experimente.

3.2.4.3 Differenzielle Genexpression der Wildtyp-Zellen

Die am höchsten in den MLL^{+/+} Zellen überexprimierten Gene kodieren für Mrp/plf3 Proliferin, Matrix γ -carboxyglutamate protein (MGP), IL-1R antagonist, Ectonucleotide pyrophosphatase, Collagen type VI α 3, Dimerization cofactor of hepatocyte nuclear factor 1 α (DCoH), Mast cell protease 7 (mcp7) und Angiotensinogen.

Proliferin gehört der Prolactin-/Wachstumshormon-Familie an und fungiert u.a. als Wachstumsfaktor und proangiogenetisches Placentahormon. Es wird natürlicherweise während Wundheilungsprozessen und im Haarfollikel-Zyklus exprimiert. Kürzlich wurde von Toft et al. gezeigt, dass Proliferin im Endstadium von Fibrosarkomen überexprimiert vorlag. Dies weist möglicherweise darauf hin, dass Proliferin eine potentielle Rolle als angiogenetisches Hormon für transformierte Tumorzellen spielt (Toft et al., 2001).

Matrix γ -carboxyglutamate protein (MGP) wurde identifiziert als Schlüsselregulator-Molekül bei osteoblastischen Knochenmineralisierungsprozessen nach PTH-vermittelter Inhibition (Gopalakrishnan et al., 2001). Eine MGP-Überexpression während der Entwicklung hat eine Inhibition von Knochenmineralisierung und eine verzögerte Knorpelentwicklung zur Folge. Es scheint ein regulatorisches Schlüsselenzym bei der Knochen- und Knorpelmineralisation zu sein. Zusätzlich wird MGP in vaskulärer, glatter Muskulatur konstitutiv exprimiert und wirkt dort vaskulärer Calcifizierung entgegen (Yagami et al., 1999).

Interleukin-1 Rezeptor-Antagonist (IL-1 Ra) ist ein kompetitiver Inhibitor von IL-1 in Bezug auf die Bindung an den IL-Rezeptor (IL-R). IL-1 Ra ist ein stark wirksamer anti-inflammatorischer Faktor, der Entzündungsprozessen, einem septischen Schock und hämatopoietischen Defekten entgegen wirkt (Irikura et al., 2002). Mäuse, denen das *IL-1R antagonist* Gen fehlt, zeigen abnormale Entwicklung und Hämostase, sowie eine veränderte Immunantwort auf inflammatorische Stimuli. Er spielt während der Entwicklung eine wichtige Rolle für die Embryoimplantation und andere Vorgänge (z.B. Blastomer-Teilungsvorgänge) während der Embryogenese (Spandorfer et al., 2000).

Ectonucleotide pyrophosphatase (Enpp2) gehört zu einer Familie von membrangebundenen Enzymen, die an der Degradation von ATP und ATP-Derivaten beteiligt sind.

Somit greift Enpp2 in Entzündungsprozesse, Apoptose-Mechanismen, Zellmotilität und Signalkaskaden ein. Es ist ebenfalls bekannt als Marker von aktivierten basophilen Granulozyten (CD203a) nach allergischer Induktion und kann durch TGF- β transkriptionell aktiviert werden (Bollen et al., 2000).

Collagen type VI α 3 (COL6A3) ist ein elementarer Bestandteil der extrazellulären Matrix (ECM) und zeigt spezifische Interaktionen mit anderen Matrixproteinen, wie z.B. Decorin, Hyaluronan, Heparansulfat und dem Proteoglykan NG-2 (Burg et al., 1996). Erwähnenswert ist, dass das Oberflächenprotein NG-2 in Patienten mit *MLL* Translokationen identifiziert wurde (Armstrong et al., 2002). COL6A3 spielt eine Rolle während der späten Phase der Neuralrohr-Entwicklung, indem es die Migration der Neuronen fördert. TGF- β führt hier ebenfalls zu einer erhöhten Transkriptionsrate (Perris et al., 1993).

Der **Dimerization cofactor of hepatocyte nuclear factor 1 α** , das bifunktionelle Protein **DcoH**, stabilisiert HNF-1 α -Dimere und verstärkt ihre transkriptionelle Aktivität (Mendel et al., 1991). Daneben agiert es als Pterin-4- α -carbinolamindehydratase und trägt zur Regeneration des Co-Faktors Tetrahydrobiopterin bei. *HNF-1 α* (*TCF-1*) wurde in hepatischen Adenomen im homozygot inaktiven Zustand vorgefunden (Bluteau et al., 2002) und hat eine bedeutende Funktion als nukleäres Target der Wnt/ β -Catenin-Signalkaskade. Die korrekte murine Entwicklung (Extremitätenausknospung, Lungenformation, Nephrogenese und Thymusentwicklung) ist stark Wnt- und HNF-1 α -abhängig. Erwähnenswert ist zusätzlich, dass ebenfalls **HNF-1 β** (*TCF-2*) in den *Mll* Wildtyp-Zellen differentiell höher exprimiert vorlag.

Neben den bisher genannten überexprimierten Faktoren wurden in *MLL*^{+/+} Zellen weitere identifiziert, die als spezifische Transkriptionsfaktoren agieren und hauptsächlich in ekto- und mesodermale Entwicklung während der Organogenese involviert sind. Dazu gehört z.B. das **Mesenchyme forkhead-1 (MFH-1)** Protein (winged helix/forkhead-Transkriptionsfaktor), das als positiver Transaktivator fungiert und mit der HNF-3 Binde-domäne interagiert. MFH-1 wird temporär während der Embryonalentwicklung exprimiert. Zuerst wird es im nicht-notochordalen Mesoderm, dann in der zukünftigen Hirnregion des Neuralrohrs und anschliessend im mesodermalen Gewebe der prechordalen Region exprimiert. BMP-2 (bone morphogenetic protein-2) induziert die MFH-1 Expression in sich entwickelndem Knorpelgewebe, Nieren und Arterien. MFH-1 ist dabei essentiell für die

normale Entwicklung des Skeletts und die Bildung des Aortenbogens in Mäusen. MFH-1 *knock-out* Mäuse sterben embryonal oder perinatal und weisen erhebliche Skelettveränderungen, vaskuläre und neuronale Defekte auf (Iida et al., 1997).

Twist-related bHLH protein Dermo-1 ist ein multifunktionaler *helix-loop-helix* Transkriptionsfaktor, der an E-Box Konsensus-Sequenzen (CANNTG) bindet und Gentranskription und Apoptose auf potente Weise negativ reguliert. Dermo-1 ist das Homolog zu Twist (*D. melanogaster*) und wird in einer ganzen Reihe von meso- und ekto-dermal abgeleiteten Geweben exprimiert. Dermo-1 reprimiert dosisabhängig den aktivierenden Transkriptionseffekt des Heterodimers MyoD/MEF-2 (myocyte enhancer factor 2), indem es direkt an MEF-2 bindet. Es wird vermutet, dass die im Dermo-1 Protein vorhandene HDAC-Aktivität für bestimmte repressorische Funktionen verantwortlich ist. Dermo-1 ist für die Aufrechterhaltung von unreifen osteoblastische Vorläuferzellen verantwortlich und inhibiert dadurch Osteoblasten-Differenzierung (Gong and Li, 2002).

Verschiedene transkriptionell aktivierte Hox Gene wurden ebenfalls in MLL^{+/+} Zellen identifiziert. Die Homeobox Gene *hox a7*, *hox b9*, *hox c9*, *hox 3.1 (hox c8)*, *iroquois 3* und der *paired-like homeodomain transcription factor 2 (RIEG, Ptx2, ARPI)* werden um den Faktor 14.8, 5.4, 9.0, 5.4, 3.9 und 26.7 in MLL^{+/+} Zellen überexprimiert. Interessanterweise wurden die Gene *hox a7*, *hox c8*, *hox c9*, *iroquois-3* und *ARPI* bereits als MLL Targetgene identifiziert (Arakawa et al., 1998; Netter et al., 1998; Yu et al., 1998). *Hox b9* Expression ist mit myeloischen und erythropoietischen Differenzierungsprozessen korreliert, während eine *hox c8* Expression hauptsächlich während der Neuronalentwicklung detektiert wurde. Ein homozygoter *hox c9 knock-out* führt zu Skelettanomalien u.a. mit einem zusätzlichen Rippenpaar (Le Mouellic et al., 1992). Das Homeobox Protein Iroquois-3 spielt in Vertebraten eine essentielle Rolle bei der Spezifikation des Neuroektoderms und ist zusammen mit weiteren Transkriptionsfaktoren für die Ausbildung der anterioren/posterioren bzw. der dorso/lateralen Achsenausbildung notwendig (Cavodeassi et al., 2001). Das *ARPI* Gen wird in Knochenmarkszellen hoch exprimiert, ist jedoch völlig abwesend in leukämischen Blasten, die eine *MLL* Translokation tragen. Es ist in Hirnentwicklungsprozesse (Diencephalon, Mesencephalon und Rhombencephalon), sowie in Augenentwicklung involviert. Mutationen führen zu Erkrankungen wie z.B. dem Rieger Syndrom (Alward, 2000).

Eine Anzahl von Genen, die für sezernierbare Proteine kodieren, wurden in MLL^{+/+} Zellen überexprimiert vorgefunden. **Gremlin** ist ein Mitglied der *Dan* Genfamilie von BMP-Antagonisten, die eine wichtige Rolle in der Embryonalentwicklung spielen. Gremlin kann an BMP2, BMP4 und BMP7 binden und verhindert ihre Interaktion mit spezifischen Rezeptoren. Diese Inhibition beeinflusst mesodermale Entwicklung, Extremitätenwachstum und Lungenausbildung (Capdevila et al., 1999). Gremlin wird in transformierten Zellen, wie auch in den MLL^{-/-} Zellen, herunterreguliert, während eine Überexpression in Tumorzelllinien einen suppressiven Effekt durch die Induktion des p21 Proteins zur Folge hat.

Beta ig-h3 wird von einer ganzen Reihe verschiedener Zellen nach TGF- β Behandlung sezerniert (Skonier et al., 1992) und wurde ebenfalls durch die Methode der subtraktiven Klonierung in MLL^{+/+} Zellen identifiziert. Humane Fibroblasten sezernieren Beta ig-h3 in die extrazelluläre Matrix, wo es an Fibronectin (FN) und Kollagen, Typ1 (Col I) bindet. Es ist zusätzlich für die Ausbreitung von Chondrozyten und Fibroblasten (über Integrin $\alpha 1\beta 1$) verantwortlich. Es ist ebenfalls in die Ausbildung von Blutgefäßen während der Angiogenese involviert; eine weitere Rolle spielt es während der Embryonalentwicklung, bei der Wundheilung und beim Tumorwachstum. In transformierten Zellen wird die Transkription des *beta ig-h3* Gens normalerweise herunterreguliert.

Der **GDP-dissociation inhibitor (D4)** ist homolog zu RhoGDI. Beide Faktoren zeigen GDI-Aktivität bei den Ras-verwandten GTPasen der Rho-Familie (CDC42Hs, rho und rac I Proteine). Während der murinen Embryogenese konnten Transkripte des GDP-dissociation inhibitors (D4) in Dottersack-Zellen nachgewiesen werden, wo die ersten hämatopoietischen Vorläuferzellen gebildet werden. Wenn diese proliferieren und ausdifferenzieren, wird eine dramatische Expressionserhöhung detektiert. Mit Hilfe von *in vitro* Experimenten konnte gezeigt werden, dass bei Differenzierung totipotenter Zellen in hämatopoietische Zellen, die Transkription des *GDP-dissociation inhibitor (D4)* Gens zu Beginn der Hämatopoese hochreguliert und während weiterer Entwicklungsstufen in seiner Stärke moduliert wird (Lelias et al., 1993).

Sprr1 Genüberexpression wurde in verletzten Neuronen detektiert, in denen es um den 60-fachen Wert hochreguliert vorlag. Überexprimiertes Sprr1 Protein kolokalisiert mit F-Aktin und ist essentiell für eine erfolgreiche Regeneration beschädigter Neuronen, indem es axonales Wachstum fördert (Bonilla et al., 2002).

Growth arrest specific 1 (GAS-1) wurde in *Mll* Wildtyp-Zellen identifiziert. Dieses Protein ist als Tumorsuppressor bekannt und inhibiert Wachstumstransformation durch Zellzyklus-Blockade im G₀/G₁- und/oder S-Stadium. Dieser p53-abhängige Mechanismus konnte in p53-mutierten Zellen oder MDM-2 überexprimierenden Zellen nicht detektiert werden. GAS-1 wird während Entwicklungsprozessen in definierten räumlich-temporären Expressionsmustern exprimiert. Seine Expression wurde in 8,5-14,5 Tage alten Embryos nachgewiesen, wobei die Stärke in den verschiedenen Organen variierte (Gehirn, Herz, Nieren, Extremitäten, Lunge und Gonaden; Lee et al., 2001).

Das Oberflächenprotein **CD53** ist ein Indikator für die Anwesenheit des Transkriptionsfaktors Pu.1 (*ets* Genfamilie), dessen cDNA zur Überprüfung seines Expressionsstatus leider nicht auf den verwendeten murinen DNA-Microchips vorhanden war. CD53 wurde als Pu.1 Targetgen identifiziert, indem *Pu.1 knock-out* Mäuse mit einem Pu.1 exprimierenden Retrovirus transduziert wurden (Henkel et al., 2002). Spezifische myeloische Gene, die für CD11a (n.p. = nicht präsent auf dem Chip), CD18 (n.p.), c-fes (1,6 fach überexprimiert), sowie die für B-Zell-spezifische Gene kodieren (Btk (n.p.), CD27 (n.p.) und vav-1 (1,8 fach überexprimiert)) sind Zielgene des Pu.1 Proteins. *Pu.1* wird schon früh in myeloischen und lymphatischen Vorläuferzellen exprimiert. Heterozygote *Pu.1* Mäuse zeigten Reifungsdefekte in allen hämatopoietischen Zelllinien, was für die notwendige Anwesenheit von Pu.1 zur Entwicklung funktionsfähiger myeloischer und lymphatischer Linien spricht. Erste Mutationen des *Pu.1* Gens konnten in AML-Patienten bereits nachgewiesen werden (Lamandin et al., 2002).

Das **GATA-6** Protein wird schon in den ersten Endoderm-Strukturen, sowie in meso- und endodermal abgeleiteten Geweben (Herz, Leber, Lunge, Gonaden und Intestinaltrakt) exprimiert. Dort spielt es zusammen mit den Faktoren GATA-4 und GATA-5 eine bedeutende Rolle bei der Regulation gewebespezifischer Genexpression (Morrisey et al., 1996). Ein homozygoter *GATA-6 knock-out* führt zur embryonalen Letalität (Tag 5,5-7,5 p.c.), wobei der Embryo schwere Defekte im Bereich des viszeralen Endoderms aufweist (Koutsourakis et al., 1999; Morrisey et al., 1998). GATA-6 hat ebenfalls Bedeutung bei der Wachstumsregulation von vaskulären glatten Muskelzellen aufgrund mitotischer oder mechanischer Stimuli.

Die **LIM-domain containing protein kinase (LIMK)** spielt eine signifikante Rolle in der Hirnentwicklung, der Synapsenformation und fungiert zusätzlich als Organisator von Aktin-Filamenten. Dabei kann LIMK die Cofilin-induzierte Aktin-Depolymerisation reverteren (Yang et al., 1998). Eine abnormale Expression ist mit dem Auftreten des Williams Syndrom korreliert (Meng et al., 2002). Zwischen Tag 7,5 und 8,5 p.c. wird *LIMK* in drei breiten Domänen des Embryos exprimiert. Dazu gehören Neuroektoderm (vordere und mittlere Hirnregion), kardiales Mesoderm und die endodermale Region des Gastrointestinal-Traktes. Bei Erwachsenen konnte LIMK hauptsächlich in bestimmten Hirnregionen detektiert werden.

3.2.4.4 Differentielle Genexpression der *Mll knock-out* Zellen

Im Folgenden soll nun kurz auf die bisher bekannten Funktionen und physiologischen Eigenschaften der interessantesten Transkripte eingegangen werden.

Die am stärksten durch die Abwesenheit des MLL Proteins aktivierten Genen kodieren für p53 apoptosis effector related tp pmp22 (*PERP*), vascular Smooth muscle α -actin, CD34 stem cell antigen, HNF-3/forkhead homolog (Brain factor 1), Serine protease inhibitor 4 (PN-1), RBP associated molecule RAM14-1, PE31/TALLA-1, Nestin, Adaptor molecule SRCASM und Myelodysplasia/Myeloid leukemia factor 1.

Das ***PERP*** Gen wurde als direktes p53-Targetgen nach Apoptoseinduktion in murinen embryonalen Fibroblasten (MEFs) identifiziert. Die *PERP* Transkription korreliert dabei mit der p53-abhängigen Apoptose, nicht jedoch mit dem p53-abhängigen Zellzyklus-Arrest. *PERP* wurde vorzugsweise in Membranen des endoplasmatischen Retikulums (ER) und des Golgi-Apparats lokalisiert und weist eine gewisse Sequenz-Ähnlichkeit zu pmp22/GAS-3 (Peripheral myelin protein 22/Growth arrest specific protein 3) auf. Dieses Membranprotein liegt in erblich bedingten peripheren Neuropathien (z.B. Charcot-Marie-Tooth) häufig mutiert vor. *PERP* scheint eine wesentliche Rolle in Regulationsmechanismen von Zellproliferation und Apoptose zu spielen (Attardi et al., 2000).

Vascular Smooth muscle alpha-actin (vSMA) ist ein Protein, das in Wundheilungsprozesse und der Differenzierung von Zellen in Myofibroblasten involviert ist (Cogan et al., 2002). Die *vSMA* Überexpression, die u.a. zu unterschiedlicher Zell-Morphologie von

$MLL^{+/+}$ und $MLL^{-/-}$ Zellen führt, wurde durch immunohistologische Experimente verifiziert (bisher unveröffentlichte Daten).

Über das murine **CD34** Antigen ist bisher wenig bekannt. Interessanterweise wird dieses Transkript in $MLL^{-/-}$ Zellen wesentlich stärker exprimiert, während es in den Wildtyp-Zellen fast komplett herunterreguliert wurde. Dieses Phänomen spricht für die Hypothese, dass in hämatopoietischen Stammzellen, die einen hohen Level an CD34 aufweisen, kaum oder gar keine *MLL* Expression vorliegt. Während der hämatopoietischen Differenzierung wird jedoch das *MLL* Gen verstärkt transkribiert, die CD34 Produktion nimmt unterdessen ab und wird schliesslich völlig herunterreguliert.

HNF-3/forkhead homolog (Brain factor 1) gehört zu der *winged-helix* Genfamilie, die für transkriptionelle Repressoren kodiert, und wurde hauptsächlich in neuronalen Vorläuferzellen gefunden. BF1 führt zu pleiotropen Effekten während der neuronalen Entwicklung, besonders in der cerebralen Hemisphäre. Mäuse, die einen BF-1 Mangel aufwiesen, zeigten morphologische Defekte im Neocortex und Telencephalon (Hanashima et al., 2002). BF-1 interagiert über das FAST-2 Protein (forkhead activin signal transducer-2) mit verschiedenen SMAD Proteinen und verhindert ihre Bindung an spezifische *DNA responsive-elements*, die in vielen Promotoren von TGF- β -induzierbaren Genen lokalisiert sind. BF-1 scheint somit einen negativen Einfluss auf die TGF- β Signalkaskade auszuüben, die bei zahlreichen Entwicklungsprozessen sowie für die Hämatopoiese eine wichtige Rolle spielt (Dou et al., 2000).

Serine protease inhibitor 4 (Protease Nexin 1, PN-1) gehört in die Klasse der Serpine und fungiert als Inaktivator verschiedener Proteasen, wie z.B. Thrombin, Urokinase, Plasminogenaktivatoren und Plasmin. PN-1 ist in Wundheilungsprozesse, Gewebemodulierung, Zellinvasionsvorgänge, Matrix-Degradation und Tumorwachstum involviert (Kim et al., 2001). PN-1 wird von Astrozyten in großen Mengen sezerniert. Die Überexpression in diesen Zellen hat eine neuronale und motorische Dysfunktionen, aufgrund von veränderten postsynaptischen Transmissionen, zur Folge. Eine Überexpression in Fibroblasten führt durch eine verstärkte Kollagen-Promotoraktivität zu einer pathogenetischen Bindegewebe-Verhärtung (Meins et al., 2001).

RBP associated molecule RAM14-1 (Kyo-T2) ist eine interessante Spleiss-Variante des *Kyo-T* Gens. RAM14-1 weist zwei LIM-Domänen und eine spezifische Interaktionsdomäne für das RBP-J κ Protein auf. Durch spezifische Bindung an RBP-J κ wird dessen Interaktion mit DNA-Targetsequenzen inhibiert und führt zu einer Hemmung der Notch-1 IC/ RBP-J κ -Signalkaskade (Taniguchi et al., 1998). Durch die Abwesenheit des MLL Proteins werden deshalb Gene (z.B. *hairy/enhancer of split (hes)*, *HERP1/2* und *NF κ B2*), die für somatische Rekombination, MHC-Expression und besonders für neuronale/sensorische Organentwicklung wichtig sind, in ihrer Expression inhibiert (de la Pompa et al., 1997). Die Folge ist eine Blockierung der Embryonalentwicklung. Die Bedeutung beider Proteine wird durch *knock-out* Experimente deutlich: *Notch-1* und *RBP-J κ* *knock-out* Mäuse sind embryonal letal und sterben am Tag 10,5 p.c. *in utero*, zum gleichen Zeitpunkt wie auch *MLL*^{-/-} Embryos (Yu et al., 1995).

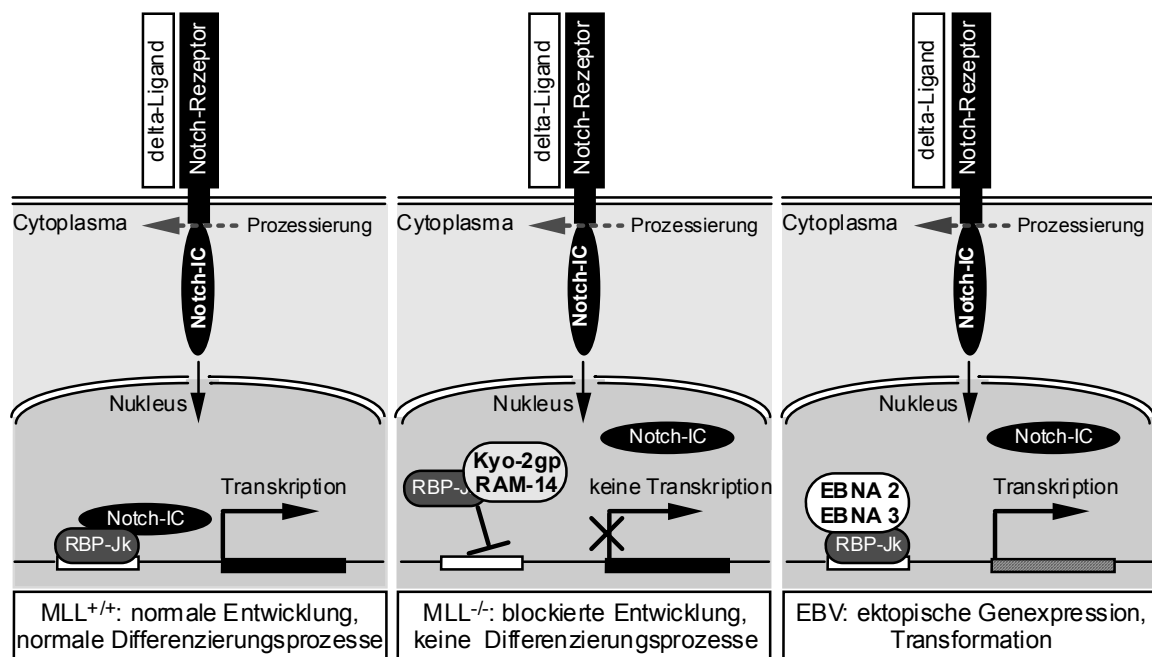


Abb. 3.21: Notch-1 Signalkaskade in Säugerzellen

Links: Nach Bindung des Liganden an den Notch-Rezeptor wird intrazelluläres notch-Protein prozessiert, wandert in den Kern und heterodimerisiert mit RBP-J κ zur Aktivierung von Gentranskription. **Mitte:** In Anwesenheit von RAM14-1 wird die Bindung von notch an RBP-J κ inhibiert. Das Heterodimer aus RAM14-1 und RBP-J κ kann nicht mehr an die spezifischen *DNA-response-elements* binden und blockiert Gentranskription. **Rechts:** ektopische Genexpression durch EBNA/RBP-J κ -Heterodimere nach EBV-Infektion.

Das *four and a half LIM domain* Transkript (*FLH1*) wurde ebenfalls in den *MLL^{-/-}* Zellen transkriptionell aktiviert. Dieses Transkript repräsentiert neben *Kyo-T2* (*RAM14-1*, *FLH2*) eine weitere Spleiss-Variante des *Kyo-T* Gens und wird hauptsächlich in gestreifter Muskulatur exprimiert und stellt einen wichtigen Faktor für die Myogenese dar (Chu et al., 2000). Durch Bindung an CREB- und CREM-Transkriptionsfaktoren spielt FLH1 als Co-Aktivatoren eine wesentliche Rolle, da es diese, unabhängig von ihrem Phosphorylierungsstatus aktiviert, indem es eine Heterodimerisierung mit CBP fördert.

PE31/TALLA-1 (A15), sowie *Tspan-6* (4,4-fach hochreguliert) sind Mitglieder der *transmembrane 4* Gen-Superfamilie (TM4SF). TALLA-1 wurde in einer cDNA-Bank unreifer T-Zellen identifiziert (*T-ALL-associated antigen 1*). Dieses Antigen wird ausschliesslich von T-ALL Zelllinien, nicht aber von peripheren Blutzellen, incl. aktivierter T-Zellen, exprimiert (Takagi et al., 1995). Northern-Blot Analysen von RNA Proben verschiedenster muriner Gewebe liessen verstärkte TALLA-1 Expressionen in Hirn, Kolon, Muskeln, Herz, Nieren und Milz erkennen. *In situ* Hybridisierungen des murinen Gehirns führen zu der Annahme, dass TALLA-1 eine fundamentale Rolle im ZNS spielt (Hosokawa et al., 1999).

Nestin, ein embryonales Intermediärfilament, repräsentiert einen neuronalen Stammzell-Marker, der mit der Zell-Differenzierung in Richtung Astroglia-Zellen korreliert. Dem Nestin Protein kommt möglicherweise eine wichtige Rolle bei der Remodulierung und Reparatur postnataler und adulter ZNS-Schäden zu. Nestin ist ebenfalls in Myogenese involviert (Wei et al., 2002).

SRCASM ist ein Adaptorprotein, das mit der Tyrosinkinase Fyn (der Src-Familie zugehörig) interagiert und von dieser phosphoryliert wird. Anschliessend kann SRCASM mit Grb-2 und der regulatorischen Untereinheit der Phosphoinositol-3-Kinase interagieren, was zu einer Aktivierung verschiedener Signalkaskaden führt (Seykora et al., 2002).

Myelodysplasia/Myeloid leukemia factor 1, ein nukleäres Proto-Onkoprotein, ist in die Translokation t(3;5)(q25.1;q34) (NPM•MLF1 Fusionsprotein) involviert, die mit dem Phänotyp einer akuten myeloischen Leukämie assoziiert ist (Yoneda-Kato et al., 1996). *Mlf1* Expression konnte ausschliesslich in glatter Muskulatur, Herz- und Skelettmuskulatur, Gonaden, Nasenschleimhaut und im Choroid Plexus detektiert werden. *Mlf1*

Transkripte konnten bisher nicht in gesunden lymphatischen/hämatopoietischen Organen nachgewiesen werden.

Einige Hox Gene wurden auch in Abwesenheit des MLL Proteins hochreguliert. Dazu gehören *hox a1*, *hox d8* und *msh-like (hox 8.1)*. Das Hox a1 Protein ist in die Neuralrohr-Entwicklung involviert. Zerstörung dieser Allele führt zu Abnormalitäten der Augen (Pasqualetti et al., 2001) und zu Fehlentwicklungen des Hirns. Das Hox d8 Protein wird in meso- und neuroektodermalen Geweben exprimiert, sowie in den Extremitäten und Gonaden (Izpisua-Belmonte et al., 1990).

SUMO Proteine sind kleine Ubiquitin-ähnliche Polypeptide, die reversibel an verschiedene Kernproteine konjugiert werden. Eine SUMOylierung von Targetproteinen führt jedoch nicht zu deren Degradation, sondern scheint eher für Protein-Protein Interaktionen und spezifische Kernlokalisation verantwortlich zu sein und schützt gleichzeitig vor proteolytischem Abbau. Untersuchungen von „*PML nuclear bodies*“, dessen Komponenten SUMOyliert sind haben gezeigt, dass diese Modifikation essentiell für funktionale und strukturelle Integrität ist (Seeler and Dejean, 2001).

Cathepsin Z precursor ist das Vorläufer-Protein des Cathepsin Z Proteins, einer Cystein-Proteinase aus der Papain Familie. Es wird in einer Vielzahl von Krebs-Zelllinien und primären Tumoren gefunden und könnte wie andere Cathepsine eine Rolle in der Entwicklung von Tumoren spielen (Santamaria et al., 1998).

ANA und **TOB** gehören der BTG-Familie an, die sich durch antiproliferative Eigenschaften auszeichnet. Dazu gehören ebenfalls Proteine wie Pc3/Btg2, Btg1, TOB2 und Pc3k. Unphosphoryliertes TOB Protein inhibiert Osteoblasten- und T-Zell-Proliferation durch Inhibition des Zellzyklus beim Übergang von der G₀/G₁- zur S-Phase. Eine Aktivierung der ras/raf-Kaskade führt zu einer ERK1/2-abhängigen TOB Phosphorylierung, die eine Verminderung der Aktivität als Zellzyklus-Inhibitor zur Folge hat (Suzuki et al., 2002).

Verifizierung

Zur Verifizierung der DNA-Microchip Daten wurden Experimente auf RT-PCR Basis, sowie auf Protein-Ebene (s. Kapitel 3.3) durchgeführt. Anhand der beispielhaft ausgewählten RT-PCR Ergebnisse in Abb. 3.21 wird gezeigt, dass die Daten der DNA-Microchip-Hybridisierung mittels weiterer Experimente reproduziert werden konnten. Semi-quantitative RT-PCR Reaktionen ermöglichten die Detektion unterschiedlicher Expressionsstärken. Die Transkripte der Gene, die bei den Arrays hochreguliert vorlagen, führten auch in diesem Versuchsansatz zu wesentlich stärkeren Banden. In Abb. A wurden beispielhaft zwei Targetgene gewählt, die in *knock-out* Zellen hochreguliert wurden, in Abb. B handelt es sich dagegen um zwei aus dem Pool der Wildtyp-Zellen. In einer Diplomarbeit wurde der Nachweis für die Übereinstimmung mit den bereits bestehenden Daten für 35 weitere Gene erbracht (Tanja Lehmann).

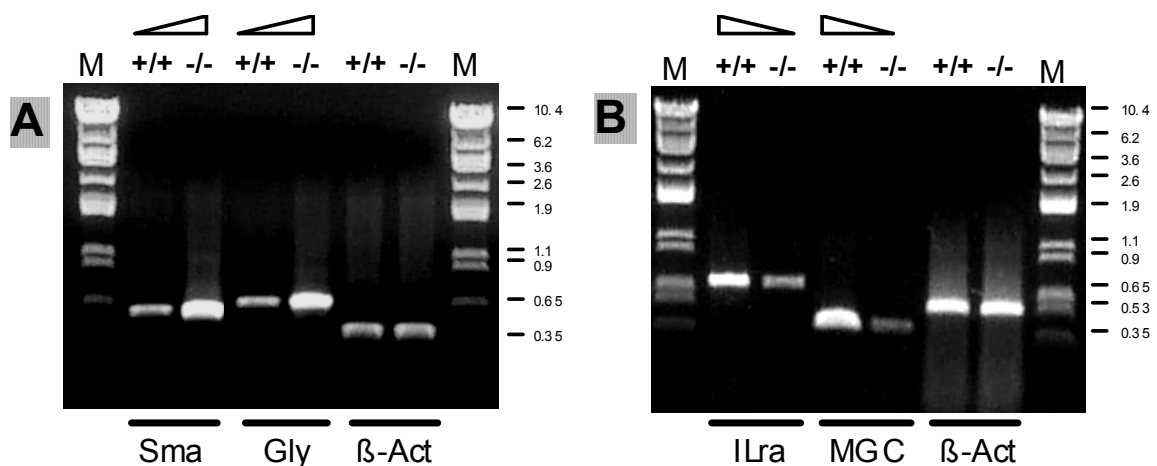


Abb. 3.22: RT-PCR zur Ergebnisverifizierung der DNA-Microchip-Hybridisierungen

Sma: vascular Smooth muscle alpha actin, Gly: Glypican, Ilra: IL-1R antagonist, MGC: Matrix γ -carboxyglutamate protein, β -Act: β -Actin (Positivkontrolle), M: Marker (λ Clal) Größenangabe in kb, +/+ : cDNA der Wildtyp Zelllinie, -/- : cDNA der *knock-out* Zelllinie. A: hier wurden RT-PCR Reaktionen (30 Zyklen) mit Oligonukleotid-Paaren von Genen durchgeführt, die in *MLL*^{-/-} Zellen und in B: die in Wildtyp-Zellen stark überexprimiert wurden.

3.3 Vergleichende Proteomanalyse der Zelllinien MLL^{+/+} und MLL^{-/-}

Zweidimensionale Gelelektrophorese

Die zweidimensionale Gelelektrophorese ist derzeit die leistungsfähigste Methode, mit der man die reproduzierbare Auftrennung komplexer Proteingemische unter Beibehaltung der quantitativen Verhältnisse durchführen kann. Ein Vorteil ist, dass die Auftrennung der Proteine nach zwei unterschiedlichen physikalisch-chemischen Kriterien, nämlich Trennung aufgrund von Ladung (isoelektrischer Punkt, 1. Dimension) und Masse (Molekulargewicht, 2. Dimension) erfolgt und die Gesamtheit der Proteine innerhalb bestimmter Grenzen auf einem Gel sichtbar gemacht werden können. Im Anschluss können mit Hilfe von computergestützten Bildauswertungsverfahren und massenspektrometrischen Methoden (z.B. MALDI-TOF MS) einzelne Proteinspots analysiert, bzw. identifiziert werden.

Die Grundvoraussetzung zum Erhalt guter 2D-Gele ist neben nukleinsäurefreien Proteinfraktionen das Auftragen der richtigen Proteinmenge. Es darf maximal soviel aufgetragen werden, dass die einzelnen Proteinspots noch sauber detektierbar und isolierbar sind. Eine Kontamination mit Proteinen aus „Nachbar-Spots“ muss unbedingt vermieden werden. Auf der anderen Seite sollten maximale Proteinmengen auf das Gel aufgebracht werden, so dass nach Verdau und Elution eine ausreichende Peptidmenge zur Identifizierung zur Verfügung steht. Aus diesen Gründen wurden Kernproteine von der Zytoplasma-Proteinfraktion abgetrennt. Dadurch konnte die Anzahl der Proteinspots auf einem Gel verringert und gleichzeitig die Auftragsmenge erhöht werden. Dies spielt besonders bei Kernproteinen eine große Rolle, da sie meist nur in sehr geringen Mengen vorliegen (z.B. Transkriptionsfaktoren) und nur schwer detektierbar sind.

Durch die vergleichende Proteomanalyse der Zelllinien MLL^{+/+} und MLL^{-/-} sollten die Ergebnisse der DNA-Mikrochip-Arrays partiell verifiziert werden.

Kern- und Plasma-Proteinfraktionen wurden isoliert und durch einen Nukleinsäure-Verdau mittels Benzonase[®] von DNA und RNA befreit. 275 µg Plasmaproteine oder 350 µg Kernproteine wurden durch eine aktive Rehydrierung auf die Gel-Strips aufgebracht, die im Gegensatz zur passiven eine höhere Reproduzierbarkeit und quantitativere Proteinaufnahme mit schon beginnender Ausrichtung der Proteine ermöglicht. Das Prinzip der isoelektrischen Fokussierung (IEF) beruht auf der Trennung der Proteine aufgrund ihres

unterschiedlichen Ladungszustandes. Unter Einfluss eines elektrischen Feldes wandern Proteine durch einen pH-Gradienten zu der Stelle, die ihren isoelektrischen Punkten entsprechen. Ein immobilisierter pH-Gradient wird durch Ko-Polymerisation von Acrylamidmonomeren und Acrylamidderivaten mit puffernden Gruppen erzeugt. Da ein komplexes Proteingemisch viele Proteine enthält die ihren isoelektrischen Punkt im neutralen Bereich haben, wurden in dieser Arbeit *DryStrips* mit einem nicht-linearen Gradienten von pH 3-10 für die folgenden Versuche eingesetzt, um eine ideale Auftrennung der Proteine durch eine Spreitung des Bereichs pH 6-8 zu gewährleisten.

Die gelelektrophoretische Auftrennung in der 2. Dimension wurde in Gegenwart des anionischen Detergenz Natriumdodecylsulfat (engl. *sodium dodecyl sulphate*, SDS) durchgeführt. Durch die Anlagerung von SDS-Molekülen entstehen anionische Mizellen mit konstanter Nettoladung pro Masseinheit mit einer linearen Beziehung zwischen dem Logarithmus des jeweiligen Molekulargewichtes und der relativen Wanderungsstrecke. Gradientengele weisen einen stärkeren Siebeffekt auf und erzielten deshalb verbesserte Trennschärfen.

Die Protein-Spots der zweidimensionalen Gele wurden mit dem Fluoreszenz-Farbstoff Ruthenium-II tris(bathophenanthrolindisulfonat) (RuBPS) angefärbt und anschliessend mit dem Fluoreszenz-Imager FLA-3000 eingescannt. Es folgte die Auswertung der Gele mit der Software PDQuest[®].

In der Abb. 3.23 ist je ein charakteristisches Zytoplasma-Gel der Proteinfractionen aus Wildtyp und *knock-out* Zellen dargestellt. Die Abszisse stellt den pH-Gradienten dar, wobei ganz links im Bild der pH-Wert bei 3 liegt und rechts bei maximal 10. Das Molekulargewicht nimmt von oben (150 kDa) nach unten (20 kDa) kontinuierlich ab.

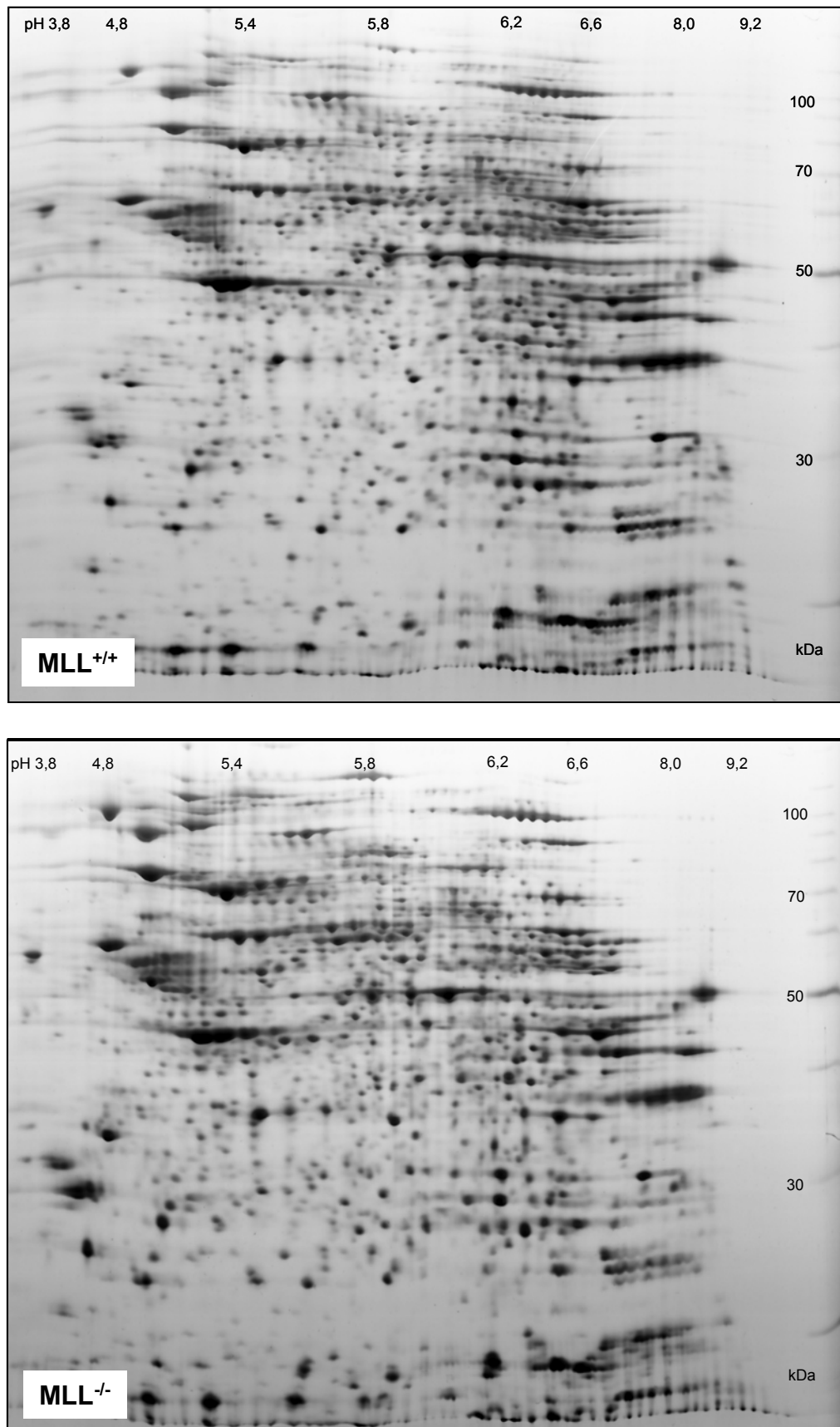


Abb. 3.23: Zweidimensionale Gelbilder von MLL^{+/+} und MLL^{-/-} Plasma-Proteinfraktionen

Zur Auswertung der Plasmaprotein-Proben wurden je 6 zweidimensionale MLL^{+/+} und MLL^{-/-} Gele miteinander verglichen. Die Erstellung eines *Match-Sets* beginnt mit der Erstellung des *Master-Gels*. Hierbei wird das Gel mit den meisten gut aufgelösten Protein-Spots ausgewählt und durch die Software als synthetisches Bild dargestellt. In diesem Fall handelte es sich um ein Gel, das die Plasmafraktion der Wildtyp Zelllinie enthielt. Im Anschluss werden visuell eindeutig identische Spots (*Landmarks*), die auf allen Gelen vorhanden sein sollten, speziell markiert. Die automatische Zuordnung umliegender Spots (markiert durch grüne Buchstaben) erfolgt über Interpolation. Protein-Spots, die nicht auf dem *Master-Gel*, aber auf anderen Gelen des *Match-Sets* detektierbar waren, wurden dem synthetischen Bild manuell zugefügt. Das resultierende *Master-Gel* repräsentiert somit jeden einzelnen Spot, der im gesamten *Match-Set* vorhanden ist. Scheinbar differentiell exprimierte Proteine werden markiert und einer zusätzlichen manuellen Überprüfung unterzogen. Dabei können noch viele vermeindliche Spots von der Analyse ausgeschlossen werden, die letztendlich nur „Schatten“ sind. Die Software PDQuest[®] ermöglicht neben der Definition differentiell exprimierter Proteine eine statistische Auswertung (z.B. Student t-Test). Nach Anlegen spezieller Filter können Relationen in Bezug auf Proteinmengen verschiedener Spots aufgezeigt werden. Bei der Auswertung der Plasma-Gele wurden Proteinspots als überexprimiert definiert, deren Proteinmengen sich mindestens um den doppelten Wert unterschieden.

Zelllinie	Proteinfraktion	exklusive Protein-Spots	mind. 2-fache Überexpression
MLL ^{+/+}	Zytoplasma	8	6
MLL ^{-/-}	Zytoplasma	12	6

Mittels DNA-Microchip-Arrays konnten 197 differentiell exprimierte Gene identifiziert werden. Das Ergebnis der zweidimensionalen Gelelektrophorese sind 8, bzw. 12 exklusiv exprimierte Protein-Spots. Je 6 unterschieden sich lediglich in ihrer Expressionsstärke. In der folgenden Abbildung sind die interessanten Protein-Spots markiert, die bisher jedoch noch nicht identifiziert wurden.

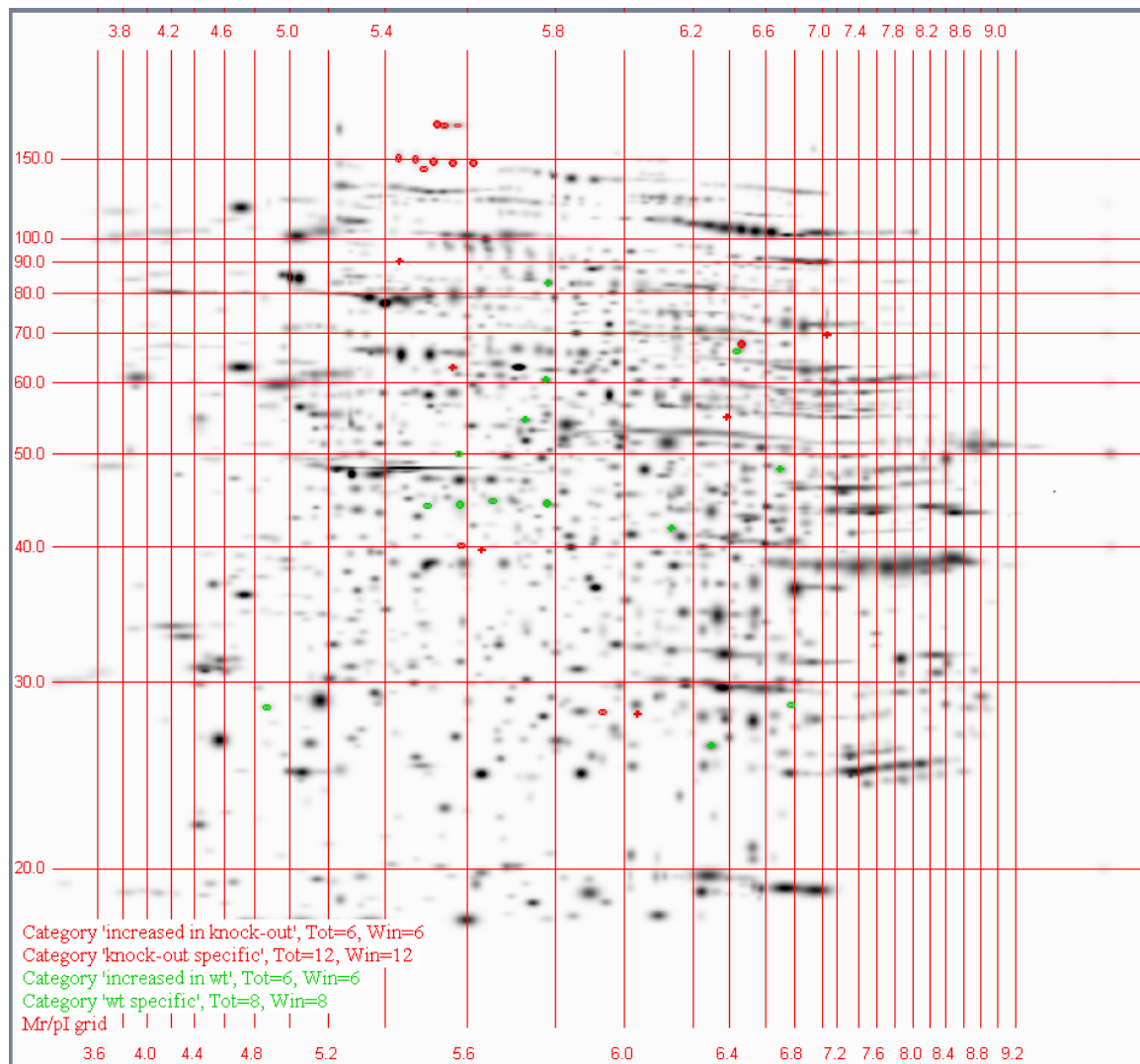


Abb. 3.24: Master-Gel (Zytoplasma-Proteinfraktion)

Bei dieser Abbildung handelt es sich um das synthetische Gel des *Match-Sets*: Plasma-Proteinfraktion $MLL^{+/+}$ vs. $MLL^{-/-}$. Rote Kreise markieren Proteine, die ausschliesslich in $MLL^{-/-}$ exprimiert werden; rote Kreuze definieren Proteine, deren Spot auf $MLL^{-/-}$ Gelen mind. doppelt so groß ist wie auf dem $MLL^{+/+}$ Gel. Für die grün markierten Spots gilt das gleiche bezüglich der Wildtyp-Zellen.

Die interessanten Protein-Spots wurden aus dem Gel ausgeschnitten und werden im Rahmen einer anderen Dissertationsarbeit (M. Hampel; AK Prof. Dr. M. Karas) nach Trypsin-Verdau und Elution massenspektrometrisch analysiert und identifiziert.

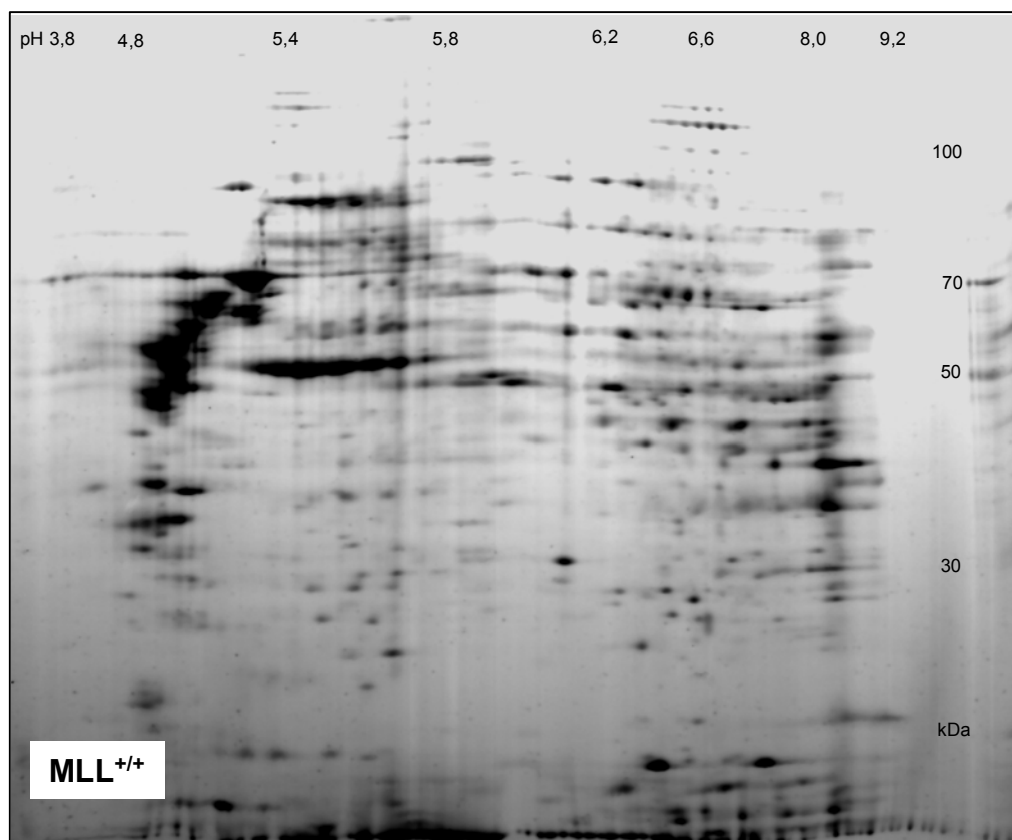
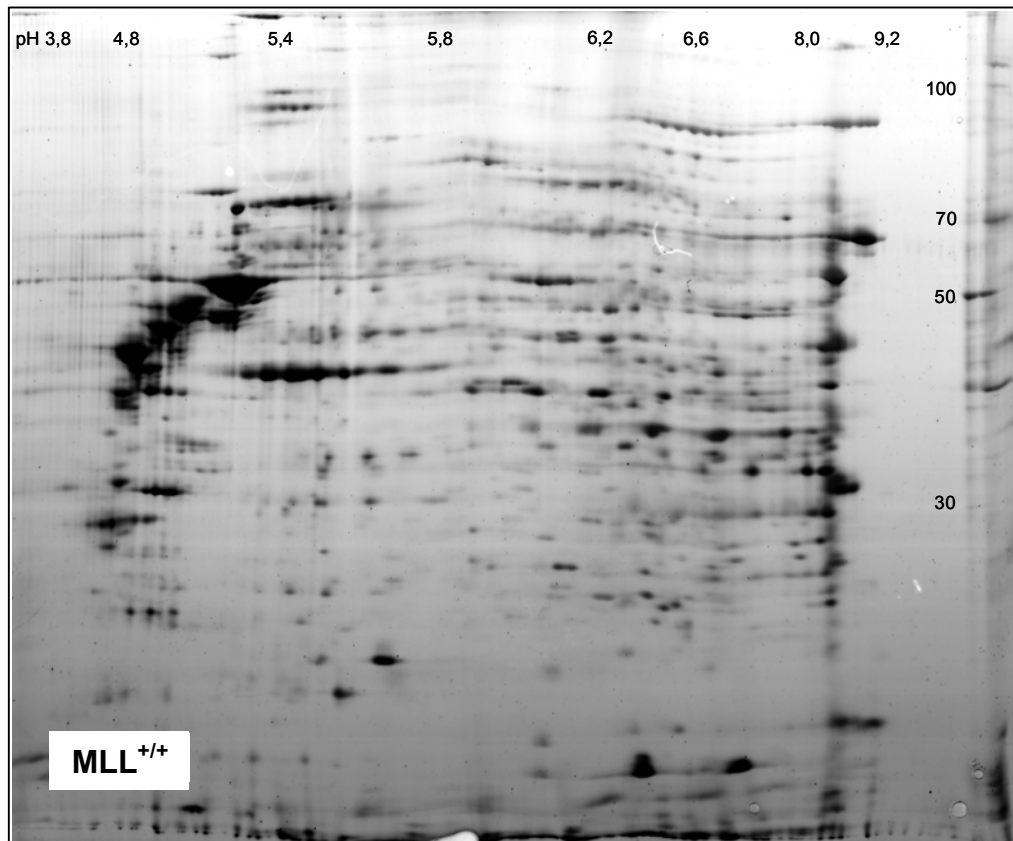


Abb. 3.25: Zweidimensionale Gelbilder von $MLL^{+/+}$ und $MLL^{-/-}$ Kern-Proteinfraktionen

Zur Erstellung des *Match-Sets* wurden je 3 MLL^{+/+} und MLL^{-/-} Kern-Proteingele miteinander verglichen. Mit Hilfe der Software PDQuest konnte, nach der zuvor beschriebenen Auswertungsmethode, nur jeweils 1 Protein-Spot gefunden werden, der ausschliesslich in einer Zellpopulation exprimiert wurde.

Zelllinie	Proteinfraktion	exklusive Protein-Spots	mind. 2-fache Überexpression
MLL ^{+/+}	Kerne	1	-
MLL ^{-/-}	Kerne	1	-

3.4 Komplementationsanalyse

In der vorliegenden Arbeit sollte eine Komplementationsanalyse Aufschluss darüber geben, ob das humane MLL Protein das fehlende murine Homolog in den $MLL^{-/-}$ Zellen in seiner Funktion komplementieren kann. Für dieses Experiment wurde RNA aus den Zelllinien $MLL^{-/-}::BC230A$ und $MLL^{-/-}::hMLL$ isoliert und auf einem denaturierenden Agarosegel analysiert (s. Abb. 3.26).

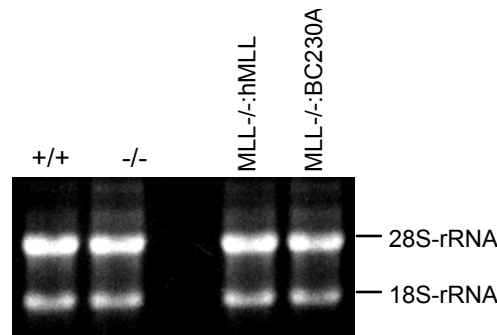
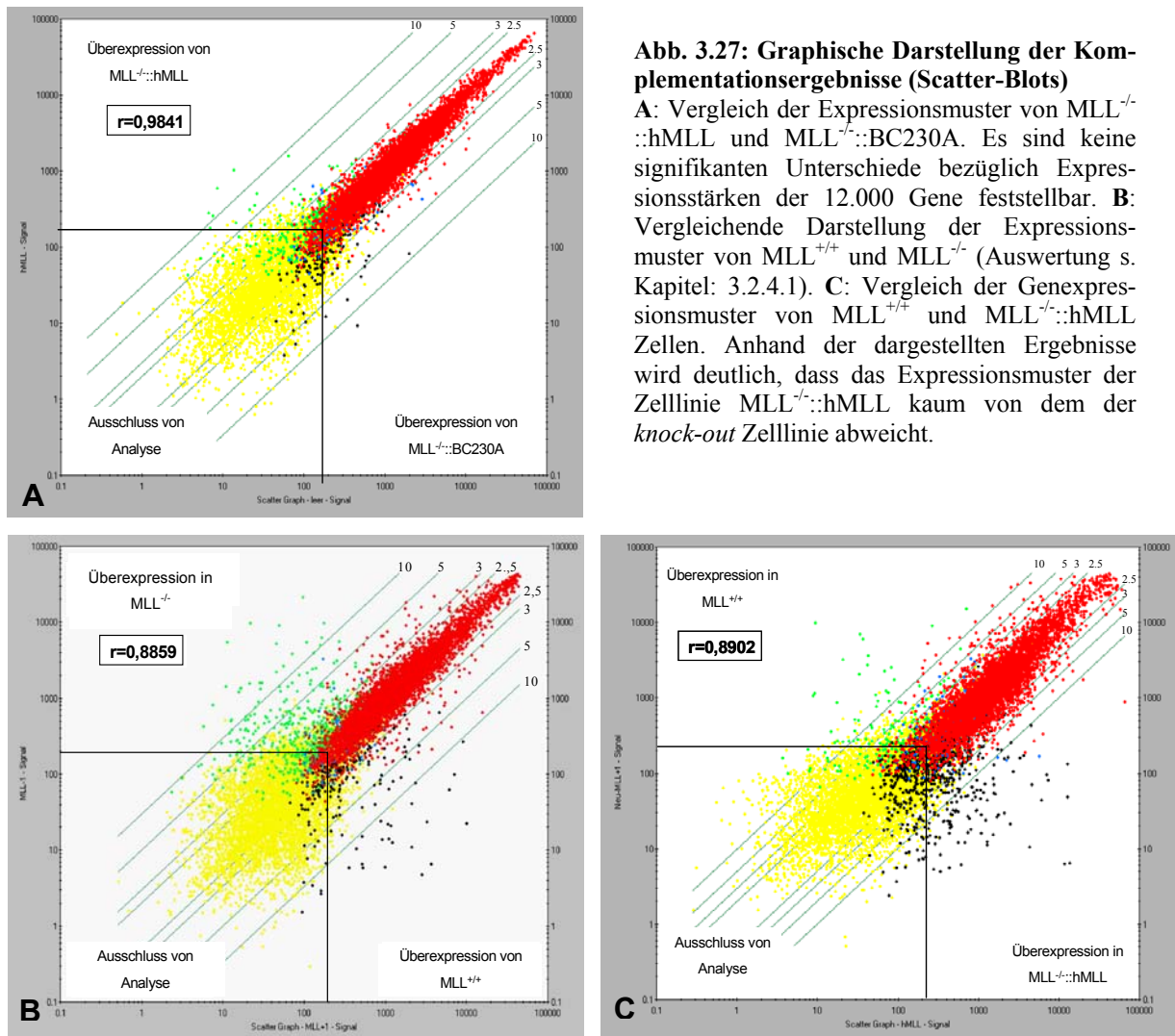


Abb. 3.26: Überprüfung der RNA-Qualität

Nach erfolgreicher cRNA-Synthese wurden anschliessend wieder DNA-Microchips hybridisiert. Die Ergebnisse sind in Abb. 3.27 als Scatter-Blot dargestellt.

Auf Ordinate und Abszisse sind die Signalstärken aus der Hybridisierung mit Phycoerythrin-markierter *antisense*-RNA verschiedener Zelllinien aufgetragen. In Abbildung A werden die Expressionsstärken der $MLL^{-/-}$ und $MLL^{-/-}::hMLL$ Zellen vergleichend graphisch dargestellt. Die einzelnen farbigen Punkte stellen je ein detektiertes Gen des Chips *MG_U74Av2* dar. Die diagonal angeordneten Geraden grenzen die Bereiche der 2,5-, 3-, 5- und 10-fachen Überexpression ab. Weitere Erläuterungen zur Farbmarkierung einzelner Gene finden sich in Kapitel 3.2.4.1.



In Abbildung A liegen die „Gen-Punkte“ fast alle in unmittelbarer Umgebung der Regressionsgraden (rot), die zwischen den beiden Begrenzungslinien der 2,5-fachen Überexpression liegt. Diese Gene werden in beiden Zelllinien mit fast gleicher Stärke exprimiert (unterscheiden sich maximal um den Faktor 2,5). Grüne und schwarze Punkte, die exklusiv exprimierte Gene repräsentieren, sind wider Erwarten kaum vorhanden (9 überexprimierte $MLL^{-/-}::hMLL$ Gene, 7 Gene überexprimiert in $MLL^{-/-}::BC230A$ (Leervektor)).

Diese Ergebnisse zeigen, dass sich die Expressionsmuster beider Zelllinien, in Bezug auf die 12.000 aufgebrachtene Gene der DNA-Microchips, kaum unterscheiden. Dieses Ergebnis wird durch den Pearson-Korrelationskoeffizienten bestätigt, der mit einem Wert von $r=0,9841$ (Vergleich $MLL^{-/-}::hMLL$ vs. $MLL^{-/-}::BC230A$) relativ hoch liegt. Je stärker sich der Korrelationskoeffizient dem Wert 1 nähert, umso geringer sind die Unterschiede der zwei untersuchten RNA Populationen.

Abbildung B zeigt die graphische Gegenüberstellung der Expressionsmuster von Wildtyp und *knock-out* Zellen, die schon in Kapitel 3.2.4.1 beschrieben wurde. Ein sehr ähnliches Expressionsmuster wird bei dem Vergleich von $MLL^{-/-}::hMLL$ mit $MLL^{+/+}$ erhalten (Abb.C). Die Anzahl und das dargestellte Scatter-Blot Muster der transkriptionell aktivierten Gene von $MLL^{-/-}$ Zellen entspricht ungefähr dem der $MLL^{-/-}::hMLL$ -Zellen. Das Resultat verdeutlicht, dass in diesem Zellsystem eine Komplementation der murinen $MLL^{-/-}$ Zellen mit Hilfe des humanen *MLL* Konstruktes unter diesen Bedingungen nicht möglich ist. Bei einem positiven Ergebnis hätte das Expressionsmuster der $MLL^{-/-}::hMLL$ Zelllinie zum größten Teil dem der *MLL* Wildtyp-Zellen entsprechen sollen.

In einem weiteren Versuchsansatz sollte überprüft werden, ob die Transfektion dieses Konstruktes einen Einfluss auf die Zell-Morphologie besitzt. Durch stabile Transfektion der $MLL^{-/-}$ Zellen mit den Vektoren pBC230::hMLL und dem Leervektor pBC230A (als Negativkontrolle) konnten zwei stabile Zelllinien etabliert werden. Die *mock*-transfizierte Zelllinie weist keine morphologischen Unterschiede zur *knock-out* Zelllinie auf. Beide Linien lassen sich durch grossflächige Zellkörper, gut sichtbare Zellkerne und breite Lamellipodien charakterisieren. Diese Merkmale grenzen sich ganz deutlich von denen der Wildtyp-Zellen (wie schon in Kapitel 3.2.1 beschrieben) ab. Die Zelllinie $MLL^{-/-}::hMLL$ weist noch die etwas größeren Zellkörper der *knock-out* Zellen auf, doch dominiert der $MLL^{+/+}$ Phänotyp durch die zahlreichen (ca. 7-8) und extrem langen dünnen Filopodien. Das transfizierte humane *MLL* Konstrukt scheint die *knock-out* Zellen in ihrer Morphologie zu verändern. Es findet möglicherweise eine partielle Komplementation statt (s. Abb 3.28).

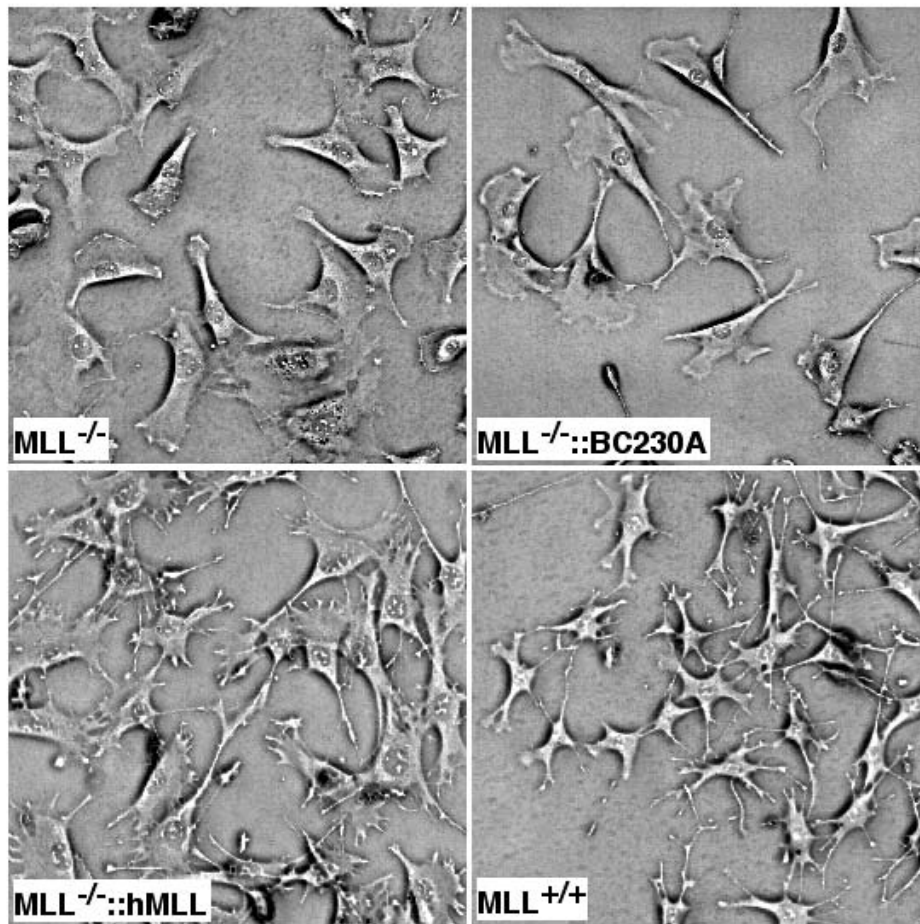


Abb. 3.28: Partielle morphologische Komplementation

Vergleichende Darstellung der unterschiedlichen Morphologien. Gezeigt werden lichtmikroskopische Zell-Bilder der Wildtyp und *knock-out* Zellen, sowie der *mock*- und *hMMLL* transfizierten Zellen bei 100-facher Vergrößerung. Das humane *MLL* cDNA Konstrukt scheint die murinen *MLL^{-/-}* Zellen in ihrer Morphologie teilweise zu komplementieren.

3.5 Identifizierung von Targetgenen der Derivatproteine MLL•AF4 und AF-4•MLL

Die Chromosomen-Translokation t(4;11) führt zur Bildung von zwei Fusionsgenen, die für die funktionellen Fusionsproteine MLL•AF-4 (der11) und AF-4•MLL (der4) kodieren. Kenntnisse über die Targetgene dieser Proteine sind zur Aufklärung des pathogenetischen Mechanismus und zur Erstellung einer therapeutischen Strategie von großer Bedeutung.

In DNA-Microchip-Hybridisierungsexperimenten sollte versucht werden, die entsprechenden Targetgene zu identifizieren. Dazu wurde die MLL^{+/+} Zelllinie stabil mit Expressionskonstrukten für MLL•AF-4 und AF-4•MLL transfiziert (Wildtyp MLL Protein wird wie in Leukämie-Zellen exprimiert). Zunächst wurde mit RT-PCR Experimenten die Transgen-Expression der stabil transfizierten Zelllinien überprüft (MLL^{+/+}::der4, MLL^{+/+}::der11 und MLL^{+/+}::der4/11). Mit den in Abb. 3.29 dargestellten Ergebnissen wurde nachgewiesen, dass eine ausreichende Menge von Transkripten der Derivatgene vorlag, die zu einer Beeinflussung der Transkriptionsmuster führen könnten (eine quantitative Translation in funktionsfähige Derivatproteine vorausgesetzt).

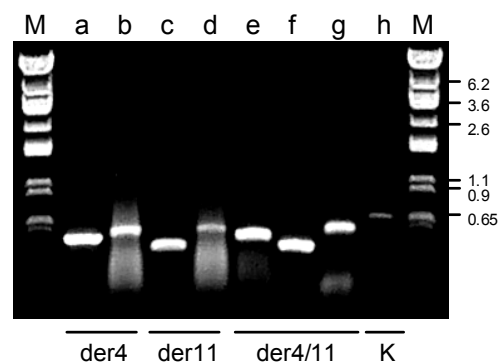


Abb. 3.29: Expressionsanalyse der Transgene in den stabil transfizierten Zelllinien (RT-PCR)

M: Marker (λ ClaI) Größenangaben in kb, Spalten b, d und g: Positivkontrolle der unterschiedlichen cDNA Populationen (β -Actin), a: AF4•3/MLL•5 (der4), c: MLL•3/AF4•5 (der11), e: AF4•3/MLL•5 (der4), f: MLL•3/AF4•5 (der11), h: Positivkontrolle: GAPDH. Diese Ergebnisse sprechen für eine ausreichende Transgen-Expression.

Im Anschluss an RNA Isolierung und DNase-Verdau wurde die RNA als Ausgangsmaterial auf ihre Qualität überprüft (denaturierendes Agarosegel). Die extrahierte RNA entsprach den Qualitätsanforderungen und konnte folglich zur cRNA-Synthese und DNA-Microchip-Hybridisierung eingesetzt werden.

Für die Analysen wurden jeweils die Expressionsmuster der stabil transfizierten Zelllinien ($MLL^{+/+}::der4$, $MLL^{+/+}::der11$ und $MLL^{+/+}::der4/11$) mit dem der Negativkontrolle ($MLL^{+/+}::BC230A$) verglichen. Diese Ergebnisse sind im Scatter-Blot graphisch dargestellt.

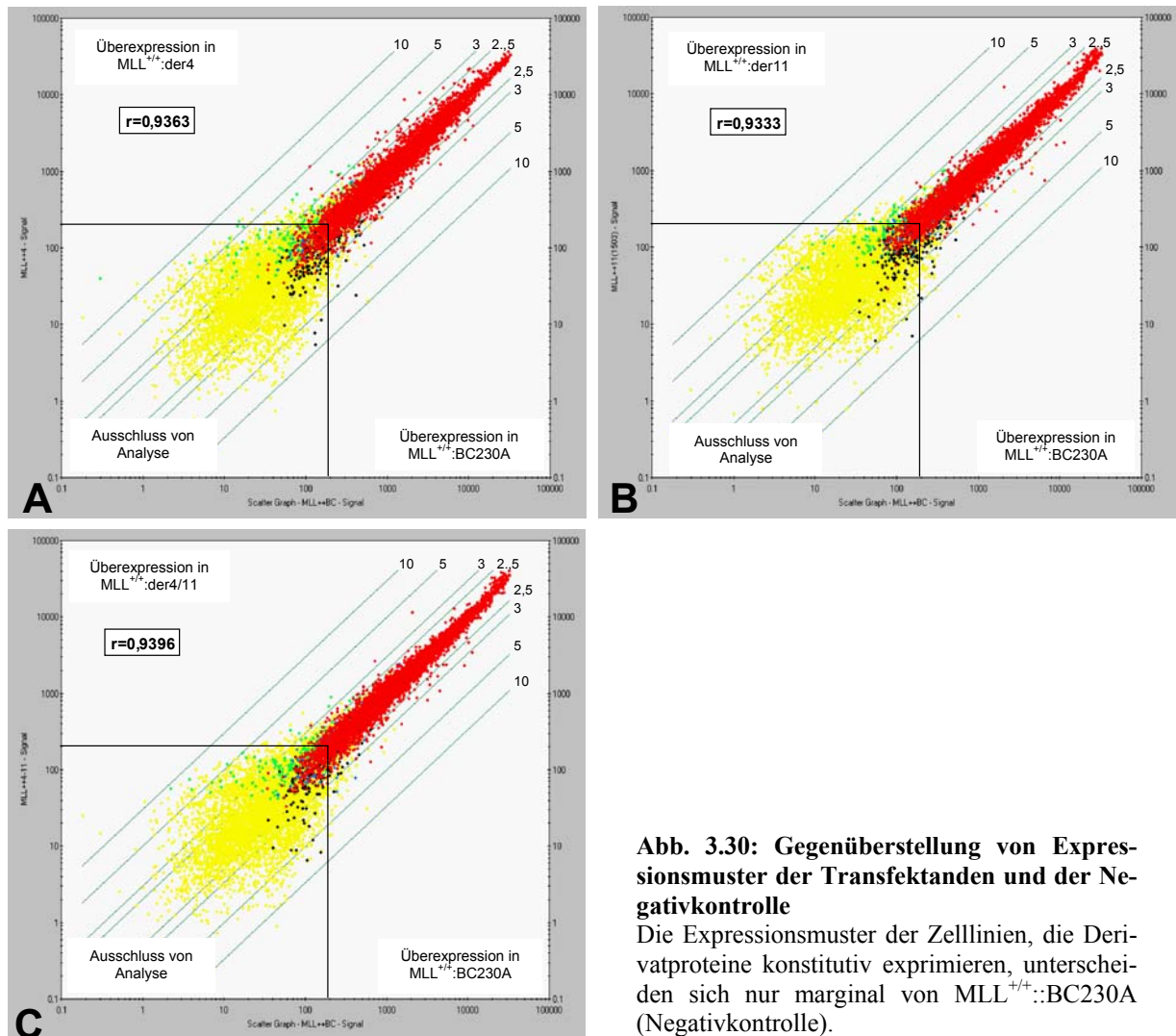


Abb. 3.30: Gegenüberstellung von Expressionsmuster der Transfektanden und der Negativkontrolle

Die Expressionsmuster der Zelllinien, die Derivatproteine konstitutiv exprimieren, unterscheiden sich nur marginal von $MLL^{+/+}::BC230A$ (Negativkontrolle).

Wie schon in den Kapiteln 3.2.4.1 und 3.4 beschrieben, sind die als grün und schwarz markierten „Gen-Punkte“ für die Auswertung besonders aussagekräftig. Dabei handelt es sich bei den Grünen jeweils um die differentiell exprimierten Gene der Transfektanden und bei den Schwarzen um die der Negativkontrolle ($MLL^{+/+}$ mock-transfiziert). Wie in den Scatter-Blots der Abb. 3.30 ersichtlich ist, unterscheiden sich die Expressionsmuster der Transfektanden von dem der Negativkontrolle nur marginal. Abgesehen von den Genen,

die aufgrund ihres geringen Fluoreszenzsignals (< 200) von der Analyse ausgeschlossen wurden, befinden sich fast alle Gene in dem Bereich bis zu einer 2,5-fachen differentiellen Expression. Eine 2,5-fache Hochregulation wurde zur Signifikanzsteigerung für die Auswertung als unterstes Limit definiert. Die Pearson-Korrelationskoeffizienten dieser drei Analysen mit den Werten $r = 0,9363$ (A), $r = 0,9333$ (B) und $r = 0,9396$ (C) bestätigen die visuell offensichtlich geringen Unterschiede zwischen den Expressionsmustern von Transfektanten und Negativkontrolle. Je weiter sich diese Korrelationskoeffizienten dem Wert 1 annähern, desto geringer sind die Abweichungen von der Regressionsgerade, die mittig zwischen den beiden 2,5-fachen Begrenzungslinien liegt. Wider Erwarten konnte bei fast allen 12.000 Genen keine signifikant unterschiedliche Expression, beeinflusst durch die Derivatproteine, detektiert werden.

3.6 Identifizierung eines kryptischen Promotors im *Mll* Gen

3.6.1 Charakterisierung des Transkriptionsstarts

Zur Verifizierung differentiell exprimierter Gene mittels Dot-Blot-Hybridisierung wurde das PCR-Fragment Exon 28-33 des *Mll* Gens auf die Membran aufgespottet (s. Kap. 3.2.3.3). Dieses sollte als Positivkontrolle für die Hybridisierung mit radioaktiv markierter $MLL^{+/+}$ cDNA und als Negativkontrolle für die Hybridisierung mit $MLL^{-/-}$ cDNA gelten. Da in den *knock-out* Zellen das *Mll* Gen ab Exon 4 nicht mehr transkribiert werden sollte, dürfte bei dieser Hybridisierung kein Signal detektierbar sein. Wider Erwarten konnte jedoch ein relativ starkes Signal identifiziert werden.

Vermutungen, es könne sich um eine Verwechslung der Sonden oder um einen Zellmischklon aus $MLL^{+/+}$ und $MLL^{-/-}$ Zellen handeln, wurden durch RT-PCR Analysen widerlegt (s. Abb. 3.15; Spalten 1,2 und 3). Diese Ergebnisse führen zu dem Schluss, dass in den *knock-out* Zellen ebenfalls ein *Mll* Transkript im 5'-Bereich exprimiert wird. In der Abbildung 3.31 werden *Mll* Wildtyp und *knock-out* Allel mit ihren Oligonukleotid-Paaren schematisch dargestellt, sowie die entsprechenden RT-PCR Ergebnisse.

Zuerst wurde mit neu entworfenen Oligonukleotiden (g/h) überprüft, ob das *Mll* Transkript in den *knock-out* Zellen durch die poly-A Sequenz, der ins Exon 4 insertierten Neo^R -Kassette, überhaupt terminiert wurde. Durch das Fehlen einer Bande in Spur 4, durch Oligonukleotide, die den Bereich Neo^R bis in den 5'-Teil des Exon 4 überspannen, musste davon ausgegangen werden, dass die Transkription definitiv gestoppt wurde, die poly-A Sequenz somit funktionsfähig war und nicht durch alternative Spleiss-Vorgänge eliminiert wurde. Es wurden im Anschluss weitere Oligonukleotid-Paare synthetisiert, die überlappend das gesamte murine *Mll* Transkript abdeckten. Mit den folgenden Primer-Paaren i/k und l/n, die den Bereich von Exon 5 bis Exon 18 überspannten, konnte in den $MLL^{-/-}$ Zellen kein Transkript nachgewiesen werden (Spuren e und f). Ab Exon 12 wurde jedoch mit den Oligonukleotiden m-v ein *Mll* Transkript detektiert (Spuren g bis l). In den Wildtyp-Zellen konnte selbstverständlich mit allen *Mll* Oligonukleotiden, bis auf die $MLL^{-/-}$ spezifischen (z.B. LacZ-Primer), ein positives Ergebnis erzielt werden.

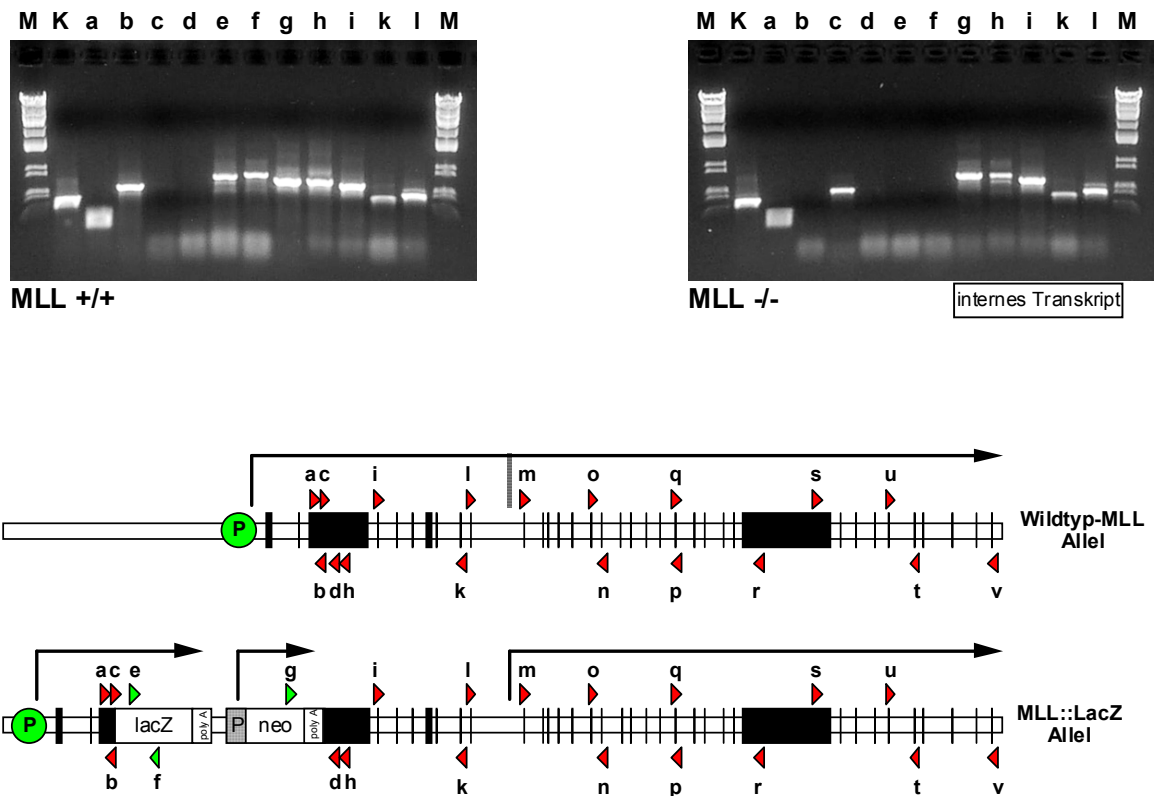


Abb. 3.31: RT-PCR Analysen zur Charakterisierung eines geninternen Transkriptionsstarts

Links: RT-PCR von cDNA der Wildtyp-Zellen ergab mit jedem *Mll* spezifischen Primerpaar eine Bande. Rechts: mit dem LacZ-spezifischen Oligonukleotidpaar (Spalte 3) und dem negativen PCR-Ergebnis in Spalte 2 lässt sich die *MLL*^{-/-} Zelllinie identifizieren. Ein Transkript wurde von Exon 12 bis Exon 37 nachgewiesen.

K: Kontrolle (GAPDH•3/•5), M: Marker (λ -*Cla*I), Größenangaben in kb, Spur a: mMLL4•3/ mMLL4•5 (a/b), Spur b: MLL•3/MLL•5 (c/d), Spur c: LacZ•3/•5 (e/f), Spur d: neo•3/mMLL4•51 (g/h), Spur e: mMLL5•3/mMLL10•5 (i/k), Spur f: mMLL11•3/18•5 (l/n), Spur g: mMLL12•3/ mMLL23•5 (m/n), Spur h: mMLL17•3/ mMLL23•5 (o/p), Spur i: mMLL23•3/mMLL28•5 (q/r), Spur k: mMLL28•3/mMLL33•5 (s/t), Spur l: mMLL32•3/mMLL37•5 (u/v).

Zur näheren Charakterisierung des Transkriptionsstartes in dem Bereich zwischen den Bindungsstellen der Oligonukleotide mMLL11•3 (Exon 11) und mMLL12•3 (Exon 12), wurden weitere Primer für eine RT-PCR Analyse entworfen.

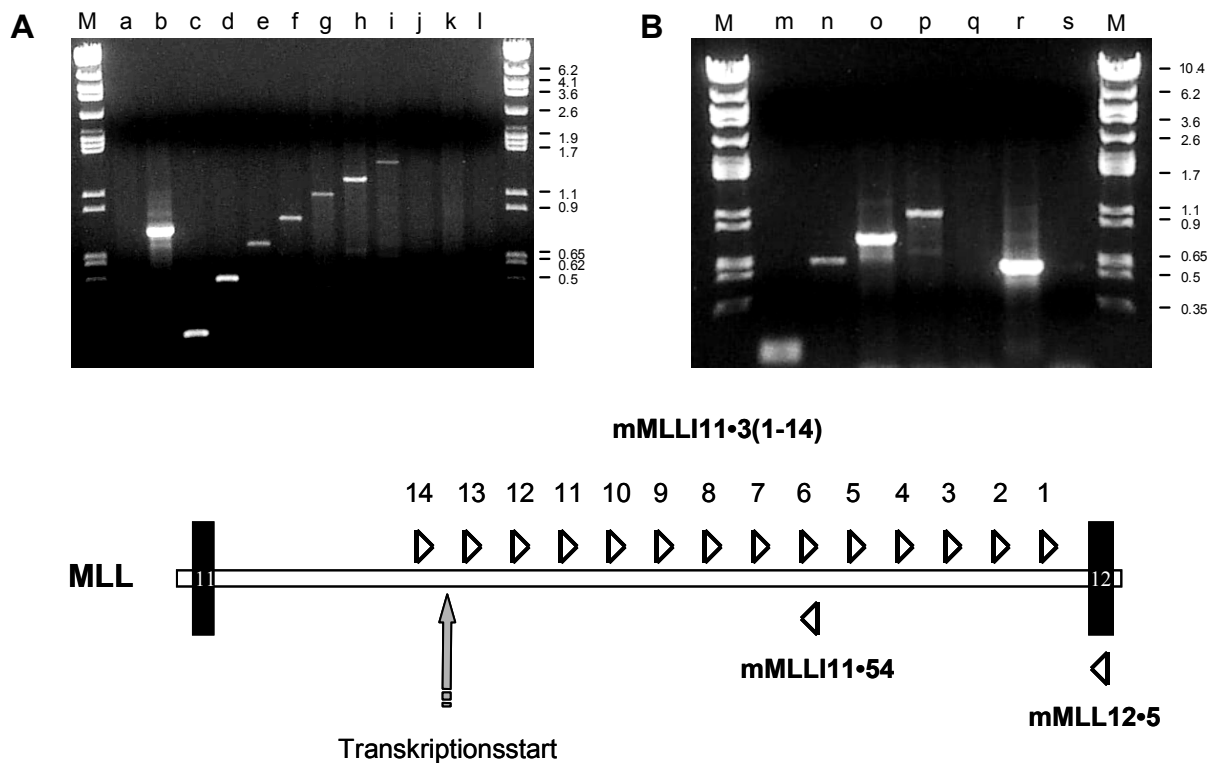


Abb. 3.32: Charakterisierung des internen Transkriptionsstarts zwischen Exon 11 und 12 in $MLL^{-/-}$ Zellen

A: a: Negativkontrolle (mMLL10•3/mMLL11•5), b: Positivkontrolle (mMLL12•3/mMLL18•5), für alle weiteren PCR-Reaktionen wurde der Primer mMLL12•5 mit den entsprechenden 3'-Primern eingesetzt. c: mMLLI11•31, d: mMLLI11•32, e: mMLLI11•33, f: mMLLI11•34, g: mMLLI11•35, h: mMLLI11•36, i: mMLLI11•37, j: mMLLI11•38, k: mMLLI11•39, l: mMLLI11•311. **B:** bei allen PCR-Reaktionen wurde der Primer mMLLI11•54 mit den entsprechenden 3'-Primern eingesetzt. m: mMLLI11•39, n: mMLLI11•311, o: mMLLI11•312, p: mMLLI11•313, q: mMLLI11•314, r: Positivkontrolle (β -Actin), s: Wasserkontrolle, M: Marker (λ -Clal), Größenangaben in kb.

Wie in Abbildung 3.32 A ersichtlich, fand durch die Verwendung verschiedener Oligonukleotid-Paare eine schrittweise Annäherung an den Transkriptionsstart statt. Ab Spalte j, d.h. mit dem Paar mMLLI11•38/mMLL12•5, wurde kein Amplifikat mehr detektiert. Der Grund dafür hätte zum einen sein können, dass sich in diesem Bereich der wahre Transkriptionsstart befindet, oder dass die Größe des zu amplifizierenden Fragments nicht mit dem verwendeten PCR-Programm zu erreichen sind. Um dies abzuklären wurde ein neuer Primer entworfen (mMLLI11•54), der näher am potentiellen Startpunkt liegt und daher zu kürzeren Amplimeren führte. Abbildung 3.32 B zeigt, dass die Annäherung in mehreren Schritten fortgeführt werden konnte und das negative Ergebnis in Spalte j, k und l nicht mit der Transkriptionsinitiierungsstelle in Verbindung stand. Aus der Betrachtung

der Spalten p und q wird ersichtlich, dass der potentielle Transkriptionsstart zwischen den Bindungsstellen der Oligonukleotide mMLLI11•313 und mMLLI11•314 liegen muss.

Es wurden anschliessend zwei weitere Oligonukleotide entworfen (mMLLI11•55 und mMLLI11•56), deren Bindungssequenzen in der näheren Umgebung des vermuteten Transkriptionsstarts liegen und für *Primer-Extension* Experimente, im Rahmen einer weiteren Doktorarbeit, zur Verfügung stehen.

3.6.2 Intrazelluläre Lokalisation

Die Detektion des internen Transkripts (Intron 11-Exon 37) warf die Frage nach dessen Translation in das entsprechende C-terminale MLL Peptid auf. Dieses sollte durch immunohistochemische Experimente in MLL^{-/-} Zellen nachgewiesen werden. Für diese Experimente wurde ein monoklonaler Antikörper verwendet, der gegen Epitope aus dem Exon 28 gerichtet ist (MAK E28). In Kombination mit einem FITC-markierten Zweit-Antikörper wurden die resultierenden CLSM-Bilder erzeugt (oben: Fluoreszenz-Bilder, unten: Kombination von Fluoreszenz- und Phasenkontrastbildern; Abb.3.34). Alle Aufnahmen wurden bei identischen Geräteeinstellungen vorgenommen.

Die intrazelluläre Verteilung des MLL Proteins in den Wildtyp-Zellen, die durch Inkubation mit MAK E28 detektiert wurde, ist der Verteilung des N-terminalen Peptids (s. Abb. 3.9) sehr ähnlich. Hierbei handelt es sich ebenfalls um eine Verteilung von kleineren Proteinkomplexen, die sich hauptsächlich im Nukleus und auch an der Kernmembran lokalisieren lassen. Die deutlich vorhandene Fluoreszenz in den MLL^{-/-} Zellen bestätigt, dass trotz Insertion der LacZ/Neo^R-Kassette in Exon 4, der C-terminale Bereich des MLL Proteins ab Intron 11 transkribiert und voraussichtlich ab Exon 17 (erstes ATG nach Transkriptionsstart) translatiert wird. Die Verteilung des verkürzten MLL Proteins unterscheidet sich jedoch erheblich von der des endogenen Wildtyp Proteins. Während das native MLL Protein eher eine gleichmäßige Verteilung von kleineren Proteinkomplexen aufweist, wird bei dem verkürzten C-terminalen Protein eine Konzentration in größere Komplexe im Zentrum des Zellkerns festgestellt.

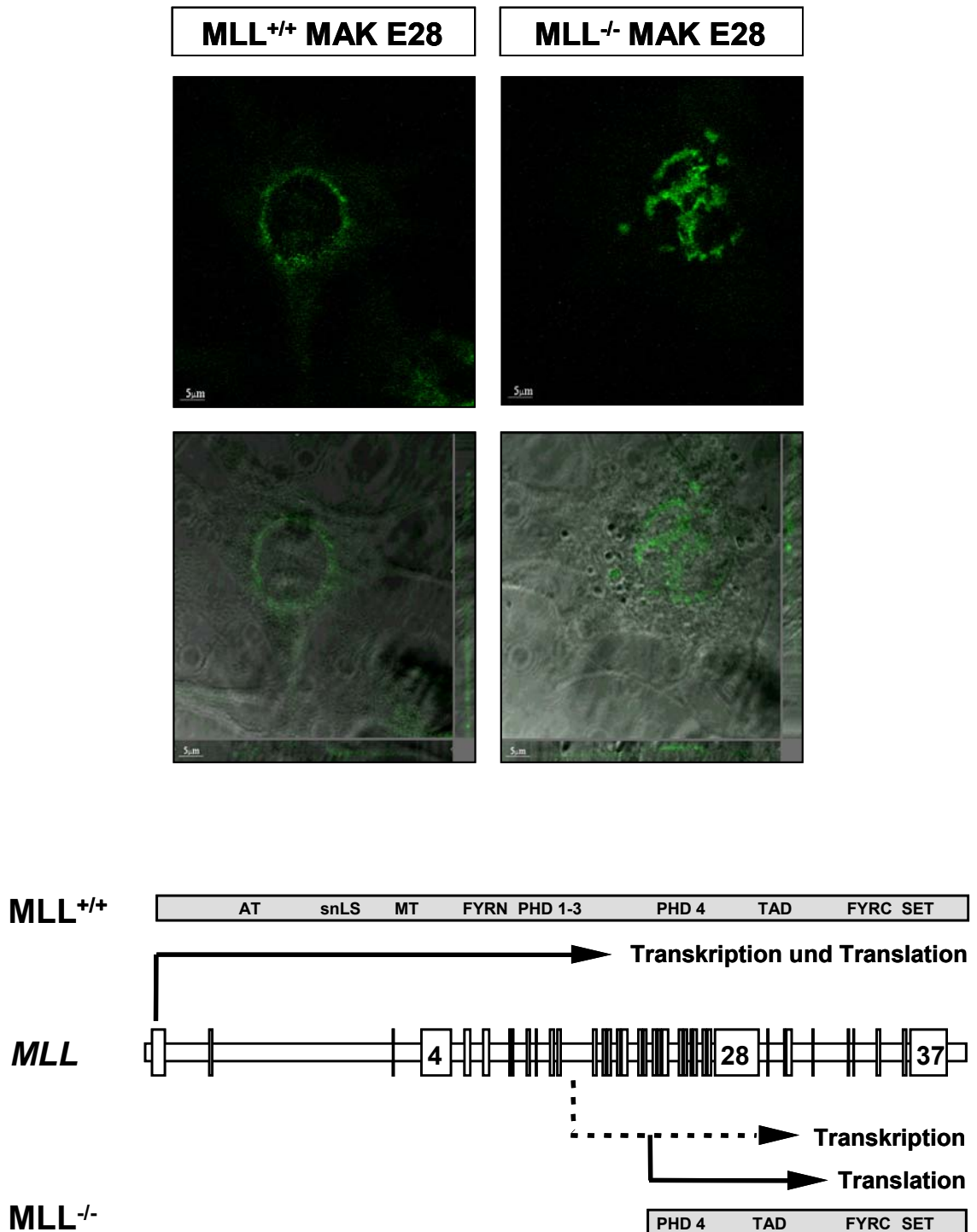


Abb. 3.33: Intranukleäre Lokalisation der C-terminalen Region des MLL Proteins

MLL^{+/+} und MLL^{-/-} Zellen wurden mit dem monoklonalen E28 Antikörper inkubiert. **Oben:** Fluoreszenz-Bilder, **unten:** Kombination von Fluoreszenz- und Phasenkontrastbildern; CLSM-Aufnahmen. **Schematische Darstellung:** **oben:** endogenes MLL Protein, exprimiert in MLL^{+/+} Zellen; **Mitte:** *Mll* Allel (Exon-Intron Struktur); **unten:** C-terminales MLL Peptid, exprimiert in MLL^{-/-} Zellen, Transkriptionsinitiierung durch internen Promotor im Intron 11, Translation voraussichtlich ab Exon 17.

3.6.3 Funktionelle Analyse des Promotorkonstruktes (Intron 11 - Exon 12)

Zur funktionellen Analyse sollte der Bereich in dem der „interne *Mll* Promotor“ vermutet wurde (Intron 11 - Exon 12), vor ein promotorloses *GFP* Gen kloniert und in 293 Zellen transient transfiziert werden.

Als Ausgangsmaterial dafür diente ein „Promotor-enthaltendes“ PCR-Fragment, das durch Amplifizierung mit den Oligonukleotiden mMLLI11•3 (Intron 11) und mMLL12•5 (Exon 12) an genomischer DNA von *MLL*^{+/+} Zellen gewonnen wurde. Nach Aufreinigung aus dem Agarosegel folgten ein *fill-in* und ein Kinasierungsschritt mit anschließender Fällung. Dieses Konstrukt wurde mit glatten Enden in den Vektor pBluescript subkloniert, der zuvor mit dem Restriktionsenzym *EcoRV* verdaut und dephosphoryliert worden war.

Zur funktionellen Überprüfung des potentiellen Promotorkonstruktes wurde der eukaryontische Vektor pEGFP-C1 ausgewählt, der für ein grün fluoreszierendes Protein kodiert. Die Modifizierung dieses Vektors ist in Abb. 3.35 schematisch dargestellt. Zuerst wurde die multiple Klonierstelle (mcs), zur Entfernung störender Schnittstellen, über die Enzyme *XhoI* und *BamHI* eliminiert (pEGFP-mcs). Im Anschluss daran musste zunächst der EGFP-spezifische hCMV-Promotor durch *AseI/NheI*-Restriktionsverdau entfernt werden. In diese Schnittstellen wurde der Linker HE-Linker im Leserahmen zum ATG der EGFP-Sequenz inkloniert, der sich durch eine *HindIII*- und *EcoRI*-Restriktionsschnittstelle auszeichnete (pEGFP+HE-mcs).

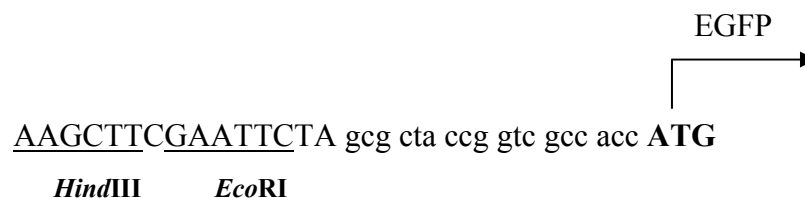


Abb. 3.34: HE-Linker *in-frame* zum ATG des *EGFP* Gens

Das potentielle murine Promotorfragment (Intron 11 - Exon 12) konnte nun mit den Enzymen *HindIII* und *EcoRI* aus dem Vektor pBluescript (pBSmProm11-12) ausgeschnitten und in die entsprechenden Schnittstellen des Linkers in pEGFP+HE-mcs inkloniert werden. Parallel zu diesen Experimenten wurde die entsprechende Klonierung mit dem humanen „Promotorfragment“ durchgeführt.

Als Positivkontrolle diente der Originalvektor pEGFP-C1 und als Negativkontrolle der Vektor pEGFP+HE-mcs, der keine Promotorsequenz enthielt.

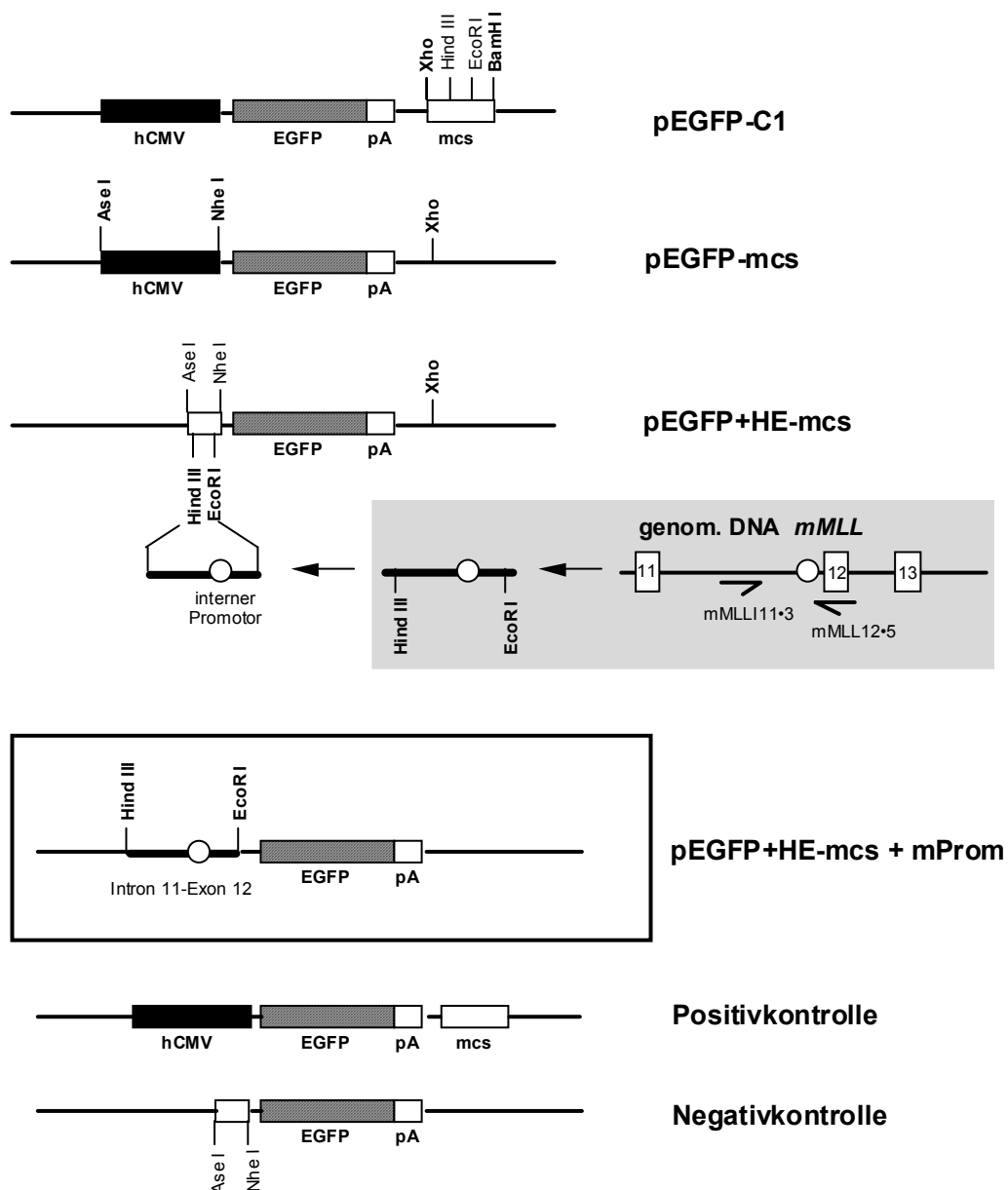


Abb. 3.35: Klonierungsstrategie zur funktionellen Promotoranalyse

hCMV: humaner Zytomegalie-Virus Promotor, EGFP: enhanced green fluorescent protein, pA: poly-A, mcs: multiple cloning site, HE: Linker mit *Hind*III und *Eco*RI-Schnittstelle, kleiner Kreis: potentieller Promotor.

Alle drei Vektoren wurden in 293 Zellen transfiziert mit dem Ergebnis, das nach 24 h und 48 h im Fluoreszenz-Mikroskop, ausschliesslich die Positivkontrolle grün fluoreszierte. Die Transfektionsrate der Positivkontrolle lag über 90%. Mehrfache Wiederholungen der Transfektionsexperimente brachten identische Resultate.

3.6.4 Funktionelle Analyse des Promotorkonstruktes (Intron 11 – Exon 13)

Klonierungsstrategie A

Aufgrund der nur beschränkt aussagekräftigen Resultate aus den Experimenten mit dem Promotorkonstrukt Intron 11 – Exon 12 wurde eine zweite Strategie zur funktionellen Analyse entwickelt. Das murine *Mll* Gen und das humane *MLL* Gen wurden auf genomischer DNA Ebene einem Homologie-Vergleich mittels der Software *strider 3.1* unterzogen.

Auf der Ebene von genomischer DNA wurde mit Hilfe der Software *strider 3.1* ein Homologie-Vergleich von murinen und humanen Sequenzen des *Mll/MLL* Gens (Exon 11 – 13) durchgeführt. Durch Erstellung einer Matrix konnten mehrere homologe Bereiche identifiziert werden.

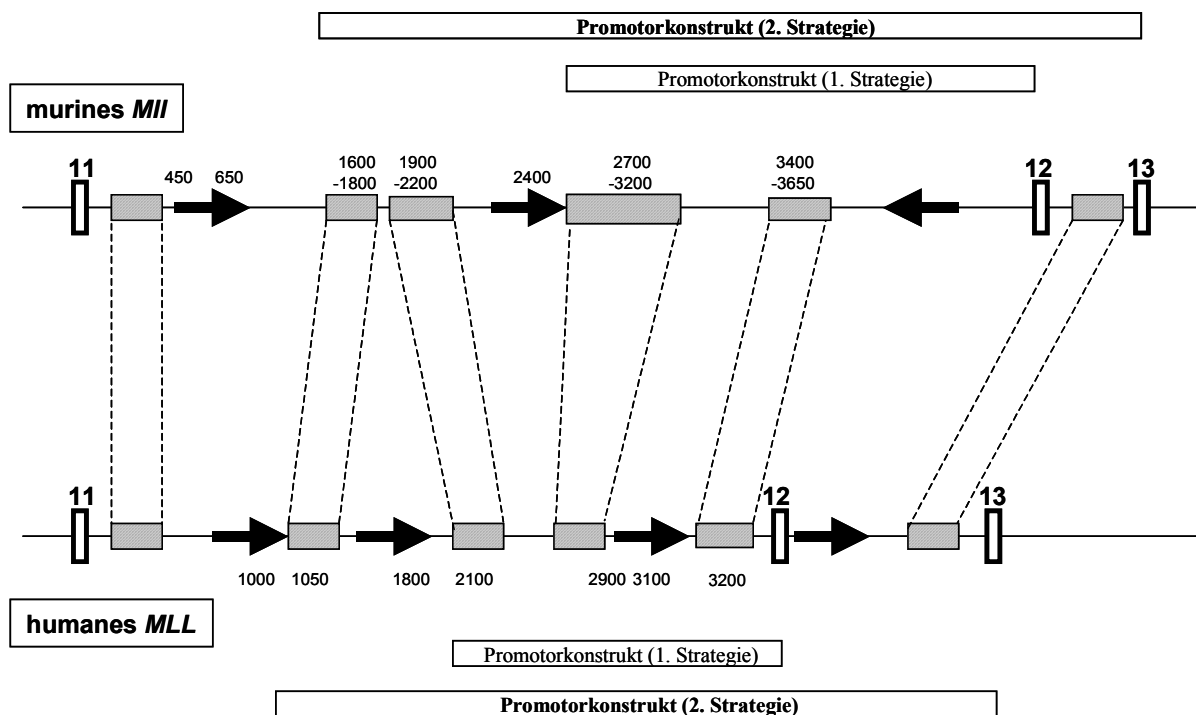


Abb. 3.36: Homologie-Vergleich murines/humanes *Mll/MLL* Gen

Schwarze Pfeile markieren Alu-Elemente, graue Boxen homologe Sequenzen, weiße Boxen markieren Exons, in Balken dargestellt: amplifizierte und klonierte Konstrukte.

Möglicherweise waren fehlende, aber essentielle „Enhancer-Elemente“ für das negative Ergebnis der funktionellen Analyse des potentiellen ersten Promotorkonstruktes verantwortlich. Aus diesem Grund wurde das entsprechende Fragment vergrößert. Wie die Abbildung 3.36 verdeutlicht, wurde in 5'-Richtung das Konstrukt um ca. 1000 bp, zwei weiterer homologer Sequenzen und ein zusätzliches Alu-Element, erweitert. Ebenfalls wurde das Fragment im 3'-Bereich um die Sequenz des Exons 13, und somit auch um einen weiteren homologen Bereich, verlängert. Dieses Fragment konnte im Gegensatz zu dem kürzeren Konstrukt der ersten Strategie nicht in den HE-Linker inkloniert werden, da durch die Erweiterung der Sequenz nun die *HindIII*- und *EcoRI*-Restriktionsschnittstellen präsent waren. Aus diesem Grund wurde in den vorhandenen Vektor pEGFP+HE-mcs ein *KpnI*/*SalI*-Linker (KS-Linker) in die *HindIII*/*EcoRI*-Seite, im Leseraster zum ATG der EGFP-Sequenz, inkloniert.

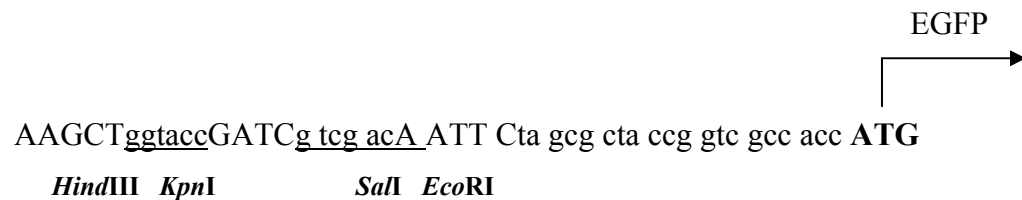


Abb. 3.37: KS-Linker *in-frame* zum ATG des *EGFP* Gens

Zur PCR-Amplifizierung dieses murinen Bereiches (3.3 kb) wurden neue Oligonukleotide entwickelt. Der Primer mProm•3 wurde zusätzlich mit einer *KpnI*-Seite und Prom•5 mit einer *SalI* Restriktionsschnittstelle versehen. Nach Amplifikation an muriner gDNA wurde dieses zweite Promotorfragment aus dem Agarosegel aufgereinigt und ein *fill-in* mit anschließender Kinasierung durchgeführt. Dieses Konstrukt wurde in dem zuvor *EcoRV*-verdauten und dephosphorylierten Vektor pBluescript subkloniert. Alle Klonierungen wurden mittels Sure™ durchgeführt und mit Hilfe von Restriktionsanalysen überprüft, wobei für alle Experimente ein positives Resultat erzielt wurde.

Das Fragment Intron 11 – Exon 13 sollte im Anschluss mit den Enzymen *KpnI* und *SalI* aus pBluescript (pBSmPromI11-E13) eliminiert werden und in die entsprechenden Schnittstellen des Expressionsvektors pEGFP+KS-mcs inkloniert werden. Dieser Klonierungsschritt erwies sich jedoch als unmöglich, da nach Ligation und Transformation des Promotorkonstruktes in den *E.coli* Stamm Top10 keine Klone erhalten wurden, bzw. wenn Klone vorhanden waren, keine intakte Plasmid-DNA gewonnen werden konnte. Der *recB* positive

Stamm Sure™ konnte nicht verwendet werden, da eine Kanamycin-Resistenz im Bakterienstamm bereits vorhanden ist. Das Einklonieren einer zusätzlichen Ampicillin-Resistenzkassette in pEGFP+KS-mcs schlug fehl, da in diesem Vektor kleinere homologe Amp^R-Sequenzen vorhanden sind, die ebenfalls zu intramolekularen Rekombinationen führten.

Klonierungsstrategie B

Das Einklonieren des Promotorkonstruktes in pEGFP+KS-mcs war aufgrund der beobachteten Rekombinationen nicht möglich. Daraufhin sollte nun die promotorlose EGFP-Sequenz *in-frame* hinter das bereits klonierte Promotorfragment (Intron 11 – Exon 13) eingebracht werden (Vektor pBSmPromI11-13).

Bevor mit den Oligonukleotiden LT•3 und LT•5 das EGFP-Fragment amplifiziert werden konnte, musste die mcs, die sich zwischen der EGFP- und der poly-A-Sequenz befindet und störende Restriktionsschnittstellen aufweist, entfernt werden. Dies geschah durch einen *Bam*HI/*Bgl*II-Restriktionsverdau mit anschließender Religation des Vektors. Der Primer LT•3 wurde mit einer *Sal*I- und der Primer LT•5 mit einer *Xho*I-Schnittstelle ausgestattet. Nach erfolgreicher Amplifikation wurde das Fragment mit diesen Enzymen verdaut.

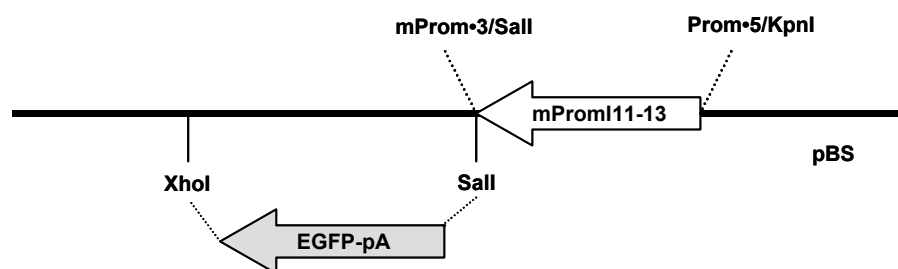


Abb. 3.38: Klonierungsstrategie B

Zur funktionellen Analyse wurde das EGFP-Konstrukt (incl. poly-A) *in-frame* hinter das murine Promotor-Konstrukt kloniert.

Der das „Promotorkonstrukt“ enthaltende Vektor pBSmPromI11-E13 wurde ebenfalls mit den Restriktionsendonukleasen *Sal*I und *Xho*I behandelt. Mit diesen Experimenten wurde die Grundlage für die Einklonierung der EGFP-Kassette und die anschließende Transformation in Säugerzellen geschaffen.

4 Diskussion

Die Häufigkeit von Leukämie-Erkrankungen im Kindesalter und die prognostisch extrem schlechten Heilungschancen von akuten lymphatischen Hochrisiko-Leukämien unterstreichen die Bedeutung intensiver Forschungsaktivität. Diese Arbeit befasst sich mit Grundlagen zum Pathomechanismus der t(4;11) Leukämie (pro-B ALL).

Obwohl der Zusammenhang zwischen dem gehäuften Auftreten von akuten lymphatischen Leukämien und der Chromosomen-Abberation des *MLL* Gens auf Chromosom 11 seit längerem bekannt ist, sind viele Zusammenhänge in Bezug auf die *MLL* Wildtyp-Funktion und die Involvierung der Derivatproteine in die Leukämogenese weiterhin unklar. Das Wissen über den molekularen Mechanismus ermöglicht die Entwicklung adäquater neuer Therapien.

4.1 Zellsysteme und Transfektionsexperimente

Das in dieser Arbeit verwendete Zellsystem, das aus den isogenen murinen Fibroblasten-Linien *MLL*^{+/+} und *MLL*^{-/-} bestand, eignete sich zur vergleichenden Genexpressionsanalyse, in der die Auswirkungen des *MLL* Proteins auf Entwicklungs- und Differenzierungsgene deutlich wurden. Der homozygote *knock-out* des *Mll* Gens eröffnete die Möglichkeit, die Expression transfizierter Transgene und deren Auswirkung auf die Zellen, ohne störenden Hintergrund durch endogenes *MLL*, zu testen. Zusätzlich konnten mittels der *MLL*^{-/-} Zelllinie Komplementationsstudien mit einem humanen *MLL* Konstrukt durchgeführt werden (die Identität zwischen murinem und humanem *MLL* Protein beträgt 92%).

Experimente mit einer humanen *MLL knock-out* Zelllinie wären aussagekräftiger gewesen, da es sich bei den Expressionskonstrukten ebenfalls um humane cDNAs handelt. Bisher war es jedoch nicht möglich, humane *MLL*^{-/-} Zellen zu generieren.

Die Etablierung stabil transfizierter *MLL*^{-/-} Zelllinien, die *MLL*•*AF-4*, *AF-4*•*MLL* oder beide Transgene exprimieren, stellte die Grundlage aller weiteren Experimente dar. Ein Ziel dieser Arbeit war die Identifizierung von Targetgenen dieser Derivatproteine, die durch vergleichende Analyse ihrer Transkriptions- und Expressionsmuster gegenüber der *mock*-transfizierten Zelllinie erhalten werden sollten. Die *Mll knock-out* Zelllinie wies für

diese Experimente im Gegensatz zu den Wildtyp-Zellen den Vorteil auf, dass durch das Fehlen des endogenen MLL Proteins, der ungestörte Einfluss der Derivatproteine detektierbar sein sollte. Informationen über Funktionalität dieser Gene sollten zum besseren Verständnis des pathogenetischen Mechanismus der Leukämogenese beitragen.

Durch die Entwicklung des *tet-off* Systems wurden scheinbar ideale Voraussetzungen für die Durchführung unserer Versuche geschaffen. Das Einklonieren der Expressionskonstrukte (*der4*, *der11* und *der4/11*) in den induzierbaren Vektor pUHD10-3 hätte theoretisch die Möglichkeit eröffnet, die Transkription der Transgene gezielt steuern zu können. Somit wäre eine zeitabhängige Detektion der Transkriptom-Beeinflussung durch die Expressionskonstrukte nach RNA-Extraktion zu verschiedenen Zeitpunkten und anschließenden DNA-Microchip-Hybridisierungen möglich geworden.

Die mangelhafte transkriptionelle Kontrolle der Transgene in den etablierten Zelllinien liess jedoch kein verwertbares Experiment zu, da trotz hoher Doxycyclin-Konzentrationen die Transkription nicht vollständig inhibiert werden konnte. Deshalb wurde die Strategie geändert und auf ein konstitutiv exprimierendes Zellsystem zurückgegriffen. Damit verbunden ist der Nachteil der fehlenden Möglichkeit zur Identifizierung primärer Targetgene durch die kontinuierliche Expression der Transgene. Durch Induktion und RNA-Isolierung zu verschiedenen Zeitpunkten wäre dies mit Hilfe des induzierbaren Zellsystems möglich gewesen. Andererseits spiegeln konstitutiv exprimierende Zellen die Situation in leukämischen B-Zellen jedoch realistischer wieder.

Im Gegensatz zu der ersten Strategie, bei der die induzierbaren Expressionskonstrukte in die *knock-out* Zellen transfiziert wurden, wurden hier die entsprechenden Vektoren in MLL^{+/+} Zellen eingebracht. Grund dafür waren Ergebnisse aus Experimenten, die im Rahmen einer anderen Dissertationsarbeit erzielt wurden (Ute Bangerter). Dort wurde mit einer Transfektion des *der4* in NIH-3T3 Zellen (Mutation im *p16* Tumorsuppressorgen), im *Focus Formation Assay* ein onkogenes Potential durch Ausbildung eines Phänotyps mit Foci-Bildung beobachtet, der jedoch bei gleichen Bedingungen in MLL^{-/-} Zellen nicht detektierbar war. Dieses Ergebnis spricht möglicherweise für die notwendige Anwesenheit des nativen MLL Proteins zur Etablierung eines wachstumstransformierenden Phänotyps.

Mit der erfolgreichen Etablierung der konstitutiv exprimierenden Zelllinien wurde die Grundlage für die nachfolgenden Experimente geschaffen.

4.2 MLL als Differenzierungsprotein in der Embryonalentwicklung

4.2.1 Morphologische Charakteristika

Die morphologischen Unterschiede zwischen den *Mll* Wildtyp und *knock-out* Zelllinien repräsentieren stabile Phänotypen, die während der gesamten Kultivierung vorhanden waren und im direkten Zusammenhang mit der An- oder Abwesenheit des MLL Proteins stehen. Die MLL-Abhängigkeit des Phänotyps wurde desweiteren durch das Komplementationsexperiment bestätigt. Es konnte damit gezeigt werden, dass ein humanes *MLL* Konstrukt, stabil in *MLL*^{-/-} Zellen transfiziert, die charakteristische Morphologie des *knock-out* Phänotyps in den der Wildtyp-Zellen revertieren konnte. Das Resultat der Komplementationsanalyse machte visuell deutlich, dass das hMLL Protein einen Einfluss auf die Zellmorphologie hat. Mittels DNA-Microchip-Hybridisierung wurde jedoch belegt, dass die Anwesenheit des hMLL Proteins praktisch keinen Einfluss auf das murine Transkriptionsmuster besitzt. Vermutlich kommt die morphologische Komplementation durch Einfluss auf der Protein-Ebene zustande. Das humane MLL Protein besitzt Protein-Bindungsdomänen und interagiert möglicherweise mit murinen Proteinen wichtiger Signalkaskaden.

Die durch die An- oder Abwesenheit des murinen MLL Proteins veränderte intrazelluläre Situation, die schliesslich zur morphologischen Veränderung führt, lässt sich z.T. aus den Ergebnissen der Affymetrix DNA-Microchip-Hybridisierungen und der subtraktiven Klonierungen herleiten. Das in *MLL*^{-/-} Zellen identifizierte Gen *Rho GTPase activating protein* (durch subtraktive Klonierung identifiziert, „sK“) kodiert für ein Protein, das die Umstrukturierung von F-Aktin fördert und eine Lamellipodien-Formation („ruffles“) induziert, die durch mikroskopische Aufnahmen (s. Kap. 3.2.1) bestätigt wurde. Die Veränderung, bzw. die Zerstörung der Aktin-Zytoskelett-Ordnung, die durch Immunfluoreszenz-Bilder in *MLL*^{-/-} Zellen mittels Phalloidin nachgewiesen wurde (s. Kap. 3.2.1), steht möglicherweise mit der Herunterregulation des *Actin filament binding protein* (sK) im Zusammenhang. Der Verlust der Aktinfilament-Organisation ist ein Charakteristikum von Tumorzellen, da die bessere Verformbarkeit die Migration in umliegendes Gewebe erleichtert. In den Wildtyp-Zellen, in denen AFAP hoch exprimiert wurde, sind gleichmäßig angeordnete Aktinfilamente sichtbar. Im Vergleich beider Zelllinien wurde in den *knock-out* Zellen, durch Verwendung eines anti-Tubulin Antikörpers, eine verstärkte Tubulin Expression detektiert, die mit den Ergebnissen der subtraktiven Kloniermethodik konform ist.

4.2.2 Intrazelluläre Lokalisation

Die intrazelluläre Lokalisation bestimmter Proteine ist nur dann eindeutig zuzuordnen, wenn die Zelle in dreidimensionaler Struktur betrachtet werden kann. Dies kann mit Hilfe des konfokalen Laserscanmikroskops (CLSM) durch optische Schnitte erreicht werden. Da sich aufgrund der intrazellulären Lokalisation Schlüsse auf mögliche Funktionen ziehen lassen, wurde die Lokalisation der MLL Peptide durch immunohistochemische Experimente näher charakterisiert.

In der vorliegenden Arbeit konnte belegt werden, dass die intrazelluläre Lokalisation beider MLL Peptide in MLL^{+/+} Zellen (MLL-N (p300) und MLL-C (p180)) mit den bisher bekannten Lokalisationssmustern identisch ist (Hsieh et al., 2003). Mit den monoklonalen Antikörpern MAK E3 (gegen Exon 3 in MLL-N gerichtet) und MAK E28 (gegen Exon 28 in MLL-C gerichtet) wurde die bisher beschriebene gleichmäßig punktierte Fluoreszenzverteilung im Zellkern detektiert (s. Kap. 3.2.2 und 3.6.2). Hsieh et al. zeigte in seiner neuesten Arbeit, dass die Expression mit anschließender Prozessierung des MLL Proteins mit deletierter FYRC- und SET-Domäne zu einem veränderten Kernlokalisationsmuster führte. Die FYRC- und SET-Domäne des MLL-C Peptids sind für die Dimerisierung beider MLL Peptide essentiell. Durch fehlende Interaktionsmöglichkeit kommt es folglich zum proteolytischen Abbau des p300 (MLL-N), da seine Stabilität durch intaktes p180 (MLL-C) reguliert wird. Mittels monoklonaler Antikörper konnte p300 kaum noch nachgewiesen werden. Das um die beiden Dimerisierungsdomänen verkürzte MLL-C Peptid konnte dagegen leicht detektiert werden, wies jedoch ein sehr diffuses Lokalisationsmuster im Zellkern auf. Diese Ergebnisse decken sich mit den Resultaten, die in dieser Arbeit erzielt wurden. In Kapitel 3.6.1 wird beschrieben, dass in *Mll knock-out* Zellen ein MLL-N Peptid bis Exon 4 und ein MLL-C Peptid von Exon 17 bis Exon 37 exprimiert wird. Damit liegt hier eine ähnliche Situation wie oben dargestellt vor. Das Fehlen der FYRN-Dimerisierungsdomäne im stark verkürzten MLL-N Peptid (Exon 1-4) verhindert die für das gleichmäßig punktierte Lokalisationsmuster essentielle Interaktion mit dem MLL-C Peptid. Dieses wird wahrscheinlich an der CS1 und CS2 proteolytisch gespalten und generiert das Wildtyp p180. Somit konnte, wie zuvor schon von Hsieh et al. beschrieben, durch fehlende Interaktion und fehlender subnukleärer Lokalisierungssequenz des p300, aber bei gleichzeitigem Vorhandensein des MLL-C Peptids, mit dem MAK E28 ein vergleichbares diffuses Kernlokalisationsmuster detektiert werden. Mit diesen Experimenten wurde belegt, dass die Interak-

tion beider intakter MLL Peptide die zwingend notwendige Voraussetzung für ihre proteolytische Stabilität und die korrekte subnukleäre Lokalisation ist.

4.2.3 Identifizierung von Targetgenen des MLL Proteins

Bisherige Untersuchungen haben gezeigt, dass das MLL Protein an Histon- und Chromatin-Modulierungen beteiligt ist und damit eine Beeinflussung von Transkriptionsprozessen herbeiführen kann. Die Identifizierung von MLL Targetgenen ist von hoher Relevanz, da reziproke Translokationen zur Zerstörung eines *MLL* Allels, zur Generation von zwei funktionsfähigen Derivatproteinen und letztendlich zum fatalen Phänotyp einer Hochrisiko-Leukämie führen. Hinweise über diese Targetgene sind deshalb von Bedeutung, da sie möglicherweise durch verringerte MLL Dosis (Reduktion auf ein intaktes Allel) in ihrer normalen Transkriptionstärke verändert werden. Zellen, die eine Translokation im *MLL* Gen tragen, weisen im Gegensatz zu herkömmlichen ALL-Formen ein charakteristisches Expressionsprofil auf, das vermutlich auf die Reduktion der MLL Dosis zurückzuführen ist (Armstrong et al., 2002). Ebenfalls können die Derivatproteine, die jeweils den N- oder C-Terminus des nativen MLL Proteins tragen, mit dem MLL MPSC um dessen Targetgene kompetieren. Aus den Erkenntnissen über Zielstrukturen und ihre Hoch- oder Herunterregulation, werden Hinweise auf einen möglichen Pathomechanismus erhofft, die als Grundlage für die rationale Entwicklung neuer Therapie-Strategien dienen könnten.

Die Durchführung von DNA-Microchip-Hybridisierungen ist heutzutage die „state-of-the-art“-Methode zur Detektion unterschiedlicher Expressionsmuster. Durch Mitführung verschiedenster Kontrollen (s. Kap. 3.2.3) wird die Aussagekraft gesteigert. Die Glaubwürdigkeit der in dieser Arbeit gezeigten Ergebnisse wird durch bereits publizierte Gene bestätigt. So konnten alle in der Literatur bisher als MLL Targetgene identifizierten Gene durch die hier durchgeführten Experimente ebenfalls nachgewiesen werden. Dazu gehören die verschiedenen Hox Gene: *hox a7*, *hox c8*, *hox c9* (Milne et al., 2002; Yu et al., 1998), *iroquois-3* (Netter et al., 1998) und *paired-like homeodomain transcription factor 2* (*RIEG/Ptx2/ARPI*; Arakawa et al., 1998).

Aufgrund der Vielzahl der identifizierten Gene werden im Folgenden nahezu ausschliesslich die jeweils 40 am stärksten transkriptionell aktivierten Gene der *MLL*^{+/+} und *MLL*^{-/-} Zellen zur Auswertung herangezogen. Die Ergebnisse der subtraktiven Kloniermethodik konnten zwar mittels Dot-Blot Hybridisierung nicht verifiziert werden (was ver-

mutlich auf methodische Probleme zurückzuführen ist), scheinen jedoch repräsentativ zu sein, da sie zum größten Teil mit den Resultaten der DNA-Microchip-Arrays konform sind, darüber hinaus immunohistochemisch bestätigt werden konnten und sich z.T. sinnvoll ergänzen (z.B. Fibronectin (sK) und Integrin $\alpha 7$; verstärkte Adhäsion gegenüber der Wildtyp-Zellen während der Zellkultur). Manche dieser Gene sind sogar durch beide von einander unabhängigen Methoden identifiziert worden (*TGF- β induced protein Beta ig-h3*, *Collagen type VI $\alpha 3$* , *GDP dissociation inhibitor*, *cardiac Troponin (Tnnt2)*).

Gene, die durch die subtraktive Klonierstrategie identifiziert wurden, werden im Folgenden durch die Abkürzung „sK“ gekennzeichnet.

In der vorliegenden Arbeit wurden DNA-Microchip-Hybridisierungen und subtraktive Klonierungen zur Auffindung von MLL Targetgenen durchgeführt, denen ein biologisches Testsystem, bestehend aus den zwei isogenen Zelllinien MLL^{+/+} und MLL^{-/-}, zu Grunde lag. Die sich daraus ergebenden Daten geben jedoch keinen Hinweis darauf, ob es sich bei den identifizierten Genen um primäre oder sekundäre Targets handelt. Zusätzlich wird bei den Resultaten, die in dieser Arbeit erzielt wurden, kein Anspruch auf Vollständigkeit der MLL Targetgene erhoben.

Die Vereinigung von RNA Fraktionen für die DNA-Microchip-Hybridisierungen, die zu unterschiedlichen Zeitpunkten aus den Zelllinien isoliert wurden, diente zur Sicherstellung eines generalisierten *steady-state* Zustandes. Zur Verifizierung der Daten wurde sechs Monate später eine zweite DNA-Microchip-Hybridisierung durchgeführt. Zur Auswertung wurden beide unabhängigen Datensätze herangezogen, wobei aus diesen ein Mittelwert gebildet wurde und anschliessend nur die Gene betrachtet wurden, die mindestens um den 2,5-fachen Wert differentiell exprimiert wurden.

Dieses Verfahren führte zur Identifizierung von 197 MLL Targetgenen. Interessanterweise wurden in MLL Abwesenheit (MLL^{-/-} Zellen) wesentlich mehr Gene transkriptionell aktiviert (136), als bei MLL Anwesenheit in den *Mill* Wildtyp-Zellen (61). Dieses Ergebnis spricht für eine Involvierung des MLL MPSC nicht nur in die bisher publizierten chromatinaktivierenden, sondern verstärkt in chromatinreprimierende Prozesse. Aufgrund der neuesten Daten aus der Arbeit von Nakamura et al. (2002) kann diese Funktion durch die am MLL MPSC beteiligten Faktoren erklärt werden. Neben transkriptionsaktivierenden sind zahlreiche reprimierende Proteine im Komplex assembliert, wie z.B. Faktoren der

NuRD- und Sin3A-Komplexe, die Histon-Deacetylaseaktivität besitzen und folglich mit transkriptioneller Inhibition assoziiert sind.

4.2.3.1 Durch MLL transkriptionell aktivierte Gene

Differenzierungsprozesse/Organogenese

Gene, die durch das MLL Protein transkriptionell aktiviert wurden, kodieren für Faktoren, die zum größten Teil in embryonale Differenzierungsprozesse involviert sind. Dabei spielt die Entwicklung von meso- und ektodermalen Geweben eine besondere Rolle. Collagen type VI $\alpha 3$ (43-fach hochreguliert), Dimerization cofactor of HNF-1 α (40-fach), Gremlin (30-fach), TGF- β induced protein Beta ig-h3 (15-fach), GDP dissociation inhibitor D4 (9-fach), Mesenchym forkhead-1 (6-fach), Iroquois-3 (6-fach im 1. Experiment), Dermo-1 (6-fach), verschiedene Hox Gene, GATA-6 (4-fach) und LIM-domain containing protein kinase (3-fach) sind für die Organogenese von Herz, Leber, Nieren, sensorischen Organen, hämatopoietischen Zellen und ebenso für Knochen und Muskeln essentiell. Dies gilt ebenso für CLAST5/Hairy-related protein und Interferon-related developmental regulator 1, die durch die subtraktive Kloniermethodik (sK) identifiziert wurden. Zu den aktivierten Genen zählen auch solche, deren Expression in Tumorgewebe bekanntermaßen herunterreguliert ist (*Gremlin* und *TGF- β induced protein Beta ig-h3*).

Besonders hervorzuheben ist die Beteiligung vieler, der durch das MLL Protein induzierten Faktoren, an der Neuronalentwicklung. Neben der Differenzierung zu unterschiedlichen Hirnarealen durch Collagen type VI $\alpha 3$, Mesenchym forkhead-1 und LIM-domain containing protein kinase, wurden Gene identifiziert, deren exprimierte Proteine maßgeblich an Regenerationsprozessen beschädigter Neurone beteiligt sind (Small proline-rich protein 1A (6-fach), Small inducible Cytokine A9 (5-fach)). Die Involvierung des MLL Proteins in neuronaler Entwicklung konnte bereits durch *in situ* Hybridisierungsexperimente belegt werden. Ein MLL•LacZ Fusionsprotein wurde in heterozygoten Mäusen hauptsächlich in ZNS, Herz und Leber exprimiert (Ayton and Cleary, 2001; Yu et al., 1995). Diese Daten sind konform mit Ergebnissen aus Untersuchungen im Maussystem und an humanen Geweben (Butler et al., 1997), wo neben den oben genannten Organen der Expressionslevel des MLL Proteins im Gehirn am höchsten ist.

Der Interleukin-1 Rezeptor-Antagonist wird in MLL^{+/+} Zellen, im Gegensatz zu den MLL^{-/-} Zellen, um den Faktor 74 hochreguliert. Dieses Protein ist neben seiner Involvie-

rung in inflammatorische Prozesse und die Immunantwort essentiell für die Embryoimplantation in den Uterus und die korrekte Blastomer-Teilung.

Zytoskelettorganisation

Neben den differenzierungsspezifischen Genen konnten weitere gefunden werden, die für die Aufrechterhaltung eines stabilen Zytoskeletts wichtig sind. Die Organisation des Zytoskeletts ist für alle intrazellulären Mechanismen von Bedeutung: von Vesikeltransport, über Abläufe von Signalkaskaden bis hin zur mitotischen Teilung. Es wurden verschiedene Hitzeschock-Proteine (HSP 70, HSP 74, HSP 84 und OSP94) und das Actin filament associated protein (AFAP-110) identifiziert (alle sK), die an der Zytoskelett-Organisation beteiligt sind.

Tumorsuppressoren

Das MLL Protein induziert ebenfalls Gentranskription von Genen, die für Tumorsuppressoren kodieren. Die Tatsache, dass bei einer Translokation des *MLL* Gens nur ein intaktes Allel verbleibt, was zu einer MLL Protein-Dosisreduktion führt, könnte diesbezüglich ein Teilaspekt für die pathologische Entartung hämatopoietischer Zellen sein. Identifiziert wurden Gene, die für die Zellzyklus-Inhibitoren p27kip1, Growth arrest specific 1, Ribosomal protein L9 (sK) und Lysyl oxidase (sK) kodieren.

Wundheilungsprozesse

MLL scheint generell Mechanismen zu beeinflussen, die eine Rolle bei Wachstum und Entwicklung spielen. Dazu gehören auch Wundheilungsprozesse, deren beteiligte Faktoren zum größten Teil durch MLL und/oder TGF- β induziert werden. TGF- β ist ein essentieller Faktor bei Proliferations- und Differenzierungsprozessen, ganz besonders jedoch bei den unterschiedlichen Stadien der Wundheilung. MLL scheint in einer bisher unbekannte Weise TGF- β zu regulieren, da in MLL^{+/+} Zellen viele Proteine transkriptionell aktiviert vorliegen, die bekanntermassen durch TGF- β induziert werden. Dazu gehören Ectonucleotide pyrophosphatase/phosphodiesterase 2 (52-fach hochreguliert), Collagen type VI α 3 (43-fach), HSP 70 (sK), TGF- β induced Beta ig-h3 (15-fach) und Lysyl oxidase (Tumorsuppressor, sK). Zusätzlich wird der Faktor Latent TGF- β binding protein 1 in MLL^{+/+} Zellen identifiziert, der die TGF- β Signalkaskade verstärkt, indem er TGF- β seinem Rezeptor zuführt und vor proteolytischem Abbau schützt. Die Wundheilung lässt sich in mehrere Stadien gliedern. Zuerst erfolgt die Hämostase mit Kontraktion der Wundränder. Hier

agiert der Faktor Thrombomodulin (10-fach hochreguliert), der als Co-Faktor die Thrombin-Wirkung der Gerinnungskaskade verstärkt. Im nächsten Schritt wird durch den 74-fach hochregulierten Interleukin-1 Rezeptor-Antagonist eine nachfolgende Entzündungsreaktion weitestgehend verhindert, bzw. gemildert. An der Bindegewebssynthese und der Organisation/Remodulierung neuer Zellen, sind viele der identifizierten Gene beteiligt: *Mitogen regulated protein/proliferin* (240-fach hochreguliert), das für den Wachstumsfaktor Proliferin kodiert, *TGF- β induced Beta ig-h3* (15-fach), Dermo-1 (6-fach, Dermatozyten-Differenzierung), *Glycosyltransferase* (sK, Zell-Zell Verbindung über ECM Proteine) und *Lysyl oxidase* (sK, fördert *cross-linking* von Kollagen und Elastin für die ECM). Zusätzlich fördern manche der genannten Faktoren die Neovaskularisierung, die für die Entstehung und Funktionalität des neuen Gewebes essentiell ist: Mitogen regulated protein/proliferin, TGF- β induced Beta ig-h3 und Latent TGF- β binding protein-1. Dabei darf jedoch nicht ausser Acht gelassen werden, dass ebenfalls ein Angiogenese-Inhibitor (Thrombospondin 2) und die schon bereits genannten Zellzyklus-Inhibitoren identifiziert wurden. Für die Aufrechterhaltung einer normalen Funktionalität der Zelle müssen aktivierende und reprimierende Faktoren vorhanden sein, jedoch darf das vorherrschende Gleichgewicht nicht in eine Richtung verschoben werden.

Demgegenüber wurden in Abwesenheit des MLL Proteins keine Gene hochreguliert, die TGF- β abhängig sind. Das Genprodukt von *HNF-3/forkhead homolog* (Brain factor 1; 110-fach überexprimiert) blockiert sogar die TGF- β Signalkaskade. Auch Decorin (11-fach) hemmt die TGF- β Aktivität. An diesem Beispiel wird die antagonistische Gegenregulation bei An- oder Abwesenheit des MLL Proteins deutlich.

Hämatopoiese

Die Tatsache, dass eine Translokation in das *MLL* Gen zu einer pathologischen Proliferation von Lymphoblasten führt, wirft die Frage auf, inwieweit MLL in die Hämatopoiese involviert ist. Mittels der DNA-Microchip-Arrays konnten Gene identifiziert werden, die durch das MLL Protein induziert wurden und die Hämatopoiese initiieren, bzw. fördern. *GDP dissociation inhibitor D4* (preferentially expressed in hematopoietic cells) wurde durch die Anwesenheit des MLL Proteins, im Vergleich zum Expressionsmuster der *Mll knock-out* Zellen, um den 10-fachen Wert hochreguliert. Dieser Faktor wird im Dottersack während der frühesten Hämatopoiese exprimiert. Es ist anzunehmen, dass Entwicklungsstörungen in der Blutbildung resultieren, wenn dieses Protein zu jenem spezifischen Zeitpunkt nicht in ausreichenden Mengen exprimiert wird. Gleiches gilt auch für Interferon-

related developmental regulator 1 (sK). Dieser Faktor wurde nach der Gastrulation im Leber-Frühstadium detektiert, wo im Anschluss an die frühe Hämatopoiese im Dottersack die erste „richtige“ Blutbildung stattfindet.

Das Oberflächen-Antigen CD53 wird in den Wildtyp-Zellen ebenfalls durch MLL hochreguliert. Es ist jedoch bekannt, dass *CD53* ein Targetgen des Transkriptionsfaktors Pu.1 ist (Henkel et al., 2002), der in die Synthese und Differenzierung hämatopoietischer Zellen involviert ist. Dessen cDNA ist leider auf den Affymetrix Arrays nicht vorhanden ist, so dass keine Aussage über seinen Expressionsstatus gemacht werden kann. Wahrscheinlich ist *CD53* kein primäres, sondern ein sekundäres MLL Target. Sollte MLL die *Pu.1* Transkription aktivieren, können Rückschlüsse über einen Zusammenhang zwischen MLL An-/Abwesenheit und Leukämien gezogen werden. In Experimenten wurde bereits gezeigt, dass heterozygote *Pu.1*^{+/-} Mäuse hämatopoietische Reifungsdefekte aufweisen (Henkel et al., 2002), was für die notwendige Anwesenheit von Pu.1 zur Entwicklung funktionsfähiger myeloischer und lymphatischer Linien spricht. Dieses Erkenntnis lässt die Vermutung zu, dass ein heterozygot vorliegendes *MLL*^{+/-} Allel möglicherweise mit einem heterozygoten *PU.1*^{+/-} Allel vergleichbar ist. Das wäre ein hypothetischer Erklärungsansatz für auftretende Reifungsdefekte von Lymphozyten mit *MLL* Translokationen (*MLL*^{+/-} Genotyp). Die Hypothese wird gestützt durch die Tatsache, dass in manchen AML-Patienten ein mutiertes *PU.1* Gen nachgewiesen wurde (Lamandin et al., 2002).

In der Literatur wird beschrieben, dass die meisten Hox Gene der Cluster A, B, C und D an Differenzierungsvorgängen und der Hämatopoiese während der Embryonalentwicklung beteiligt sind (Magli et al., 1997). In *MLL*^{+/+} Zellen liegen *hox a7*, *hox b9*, *hox c8* und *hox c9* verstärkt exprimiert vor.

Die meisten Gene, die durch die subtraktive Kloniermethodik identifiziert wurden, sind leider nicht auf den Affymetrix DNA-Microchips vorhanden. Daher bleibt es spekulativ anzunehmen, ob nicht z.B. die verschiedenen Mitglieder der Hitzeschock-Proteinfamilie (HSP 70, 74, 84 und OSP 94) dort ebenfalls identifiziert worden wären. Hybridisierungen von *MLL*^{+/+} cDNA mit Filter Arrays (Fa. Clontech) bestätigen jedoch die differentielle Expression von HSP 84 (Ute Bangerter, bisher unveröffentlichte Daten).

Verifizierung

Arbeiten aus der eigenen Arbeitsgruppe führten zu einer Verifizierung der DNA-Microchip-Array Daten. Für fünf der am stärksten hochregulierten Gene (IL-1R antagonist, Collagen type VI $\alpha 3$, Matrix γ -carboxyglutamate protein, Angiotensinogen, TGF- β induced protein Beta ig-h3) wurden semi-quantitative RT-PCR Experimente durchgeführt, die aufgrund ihrer Bandenstärke im Agarosegel die differentielle Expression, nach 25, 30 und 35 Zyklen, deutlich bestätigten.

Des Weiteren wurden RT-PCR Reaktionen mit RNA aus SEM Zellen durchgeführt. Hierbei handelt es sich um humane t(4;11) Zellen, die aus einem Patienten mit einer ALL isoliert wurden (Greil et al., 1994). Mittels dieser Experimente sollte eine Aussage über die Übertragbarkeit der aus den murinen Zellen gewonnenen Daten auf humane Leukämie-Zellen mit *MLL* Translokation gewonnen werden. Gene, die in *MLL*^{+/+} Zellen extrem hochreguliert wurden, sollten theoretisch in den SEM Zellen weniger stark (durch das Fehlen eines *MLL* Allels bedingt) exprimiert werden. Die Gene *Gremlin*, *GDP dissociation inhibitor D4*, *Adenyl cyclase associated protein*, *Growth arrest specific 1*, *CD53* und noch weitere, wurden mit SEM cDNA deutlich amplifiziert (Tanja Lehmann, Daten nicht gezeigt). Zum Vergleich stand keine gesunde B-Zelllinie zur Verfügung. Durch das in SEM vorliegende intakte *MLL* Allel ist die Beeinflussung der Genexpression durch Mangel an *MLL* Protein vermutlich schwächer ausgeprägt als in *MLL*^{-/-} Zellen.

Für 7 weitere Gene, die in *MLL*^{+/+} Zellen hochreguliert wurden, konnte in SEM Zellen keine Expression nachgewiesen werden. Diese Tatsache könnte die Hypothese stützen, dass ihre Expression in t(4;11) Zellen durch die Zerstörung eines *MLL* Allels stark reduziert, bzw. nicht vorhanden ist. Hierbei handelt es sich um die wichtigsten embryonalen Differenzierungsgene, die in dieser Arbeit identifiziert wurden (*Mesenchym forkhead 1*, *Dermo-1*, *hox A7*, *TGF- β induced protein Beta ig-h3*, *LIM-domain containing protein kinase*, *Brain factor 2*, *Small proline-rich protein 1A*).

4.2.3.2 Durch *MLL* reprimierte Gene

Wie schon eingangs in Kap. 4.2.3 erwähnt, werden gemäß den mit dieser Arbeit gewonnenen Erkenntnissen durch das *MLL* Protein wesentlich mehr Gene in ihrer Transkription reprimiert als aktiviert ($\frac{2}{3}$ vs. $\frac{1}{3}$). Diese Tatsache ist bemerkenswert, weil dem *MLL* MPSC bisher eine fast ausschliesslich transkriptionsaufrechterhaltende Aktivität zugespro-

chen wurde. Scheinbar üben die transkriptionsreprimierenden Komplex-Faktoren mit Histon-Deacetylaseaktivität einen stärkeren Einfluss aus, als bisher angenommen wurde.

Die Abwesenheit des MLL Proteins ist verbunden mit der Expression vieler interessanter Targetgene, die möglicherweise einen funktionellen Beitrag zur Leukämogenese liefern. Für einige dieser Gene wird bereits eine Funktion in der Tumorentstehung vermutet, andere werden bereits als Tumormarker in der Diagnostik verwendet.

Tumorgenese/Tumormarker

Serine protease inhibitor 4 (PNI, Protease Nexin 1), in MLL^{-/-} Zellen 110-fach hochreguliert, spielt eine Rolle in der Tumorentstehung, bei Gewebeveränderungen und in der zellulären Invasivität (Kim et al., 2001). Dieses Gen wird ebenfalls in der humanen t(4;11) Zelllinie SEM extrem stark exprimiert, was aus der MLL Dosisreduktion resultieren könnte. Für das Protein FBJ osteosarcoma oncogene konnte bereits transformierendes Potential bewiesen werden, das letztendlich zu Osteosarkomen und Lymphomen führt. TALLA-1 ist ein Tumormarker für T-ALL Zellen (Takagi et al., 1995). Ebenso sind die Proteine Tumor protein D52-like (auch in Leukämie-Zellen überexprimiert) und Tenascin C Tumormarker, die für die Diagnostik von Mamma- und Ovarialkarzinomen von enormer Bedeutung sind (Wilson et al., 1999). Tenascin C protagiert über einen pro-angiogenetischen Mechanismus die Neovaskularisierung im Tumorgewebe. In vielen primären Tumoren und Krebszelllinien wurde Cathepsin-Z-precursor detektiert, der in die Entwicklung von Karzinomen involviert sein soll (Santamaria et al., 1998). Das gleiche gilt auch für Necdin-like protein 2.

Durch die subtraktive Kloniermethodik wurden weitere Gene identifiziert, die mit Tumorentstehung in Verbindung gebracht werden können. Eine Überexpression von *CASP* (*cux*, Hox Gen) und *Thrombospondin1 (THBS1)* wurde in verschiedenen Tumoren bereits beobachtet, wo sie zur verstärkten Invasion maligner Zellen beitragen. *Cathepsin D*, eine estrogenabhängige Aspartat-Protease, wird bei Mammakarzinomen hochreguliert. BPOZ, ein Wachstumsinhibitor, der einen Arrest in G₀/G₁ herbeiführt, ist in die Entwicklung von Leukämien involviert.

Proliferation

Einige weitere in *Mll knock-out* Zellen hochregulierte Gene, die auch in t(4;11) Leukämie-Zellen verstärkt exprimiert wurden, sind in Signaltransduktionskaskaden und der Steuerung des Zellzyklus involviert. TOB und ANA sind Komponenten des CCR4 Transkriptionsfaktor-Komplexes, der u.a. für die Steuerung des Zellzyklus verantwortlich

ist (Yoshida et al., 2001). GM2 activator protein ist ebenfalls eine Zellzyklus Komponente. Eine Störung der Komplex-Zusammensetzung durch Überexpression dieser Proteine könnte negativen Einfluss auf die Mitose-Steuerung haben, zu einem Ausbruch aus dem normalen Zellzyklus und folglich zur unkontrollierten Proliferation führen. Gleizeitig werden Gene überexprimiert, die durch Förderung des Nukleotid-Metabolismus (*Nudt4*, *CASP*, beide sK) und der Membran-Biosynthese (*Phospholipid transfer protein*, sK), Proliferationsprozesse unterstützen. Im Zusammenspiel mit der Hochregulation anderer Proto-Onkogene kann diese entkoppelte Zellzyklus-Kontrolle die Grundlage onkogener Mechanismen darstellen.

Ebenso kann eine Überexpression von *SRCASM* (*Src activating signaling molecule*), durch Störung der regulatorischen Untereinheit der Phosphoinositol-3-Kinase, zu einer Störung zahlreicher zellulärer Signalkaskaden führen (Seykora et al., 2002).

Zytoskelettorganisation/Zellmigration

Hervorzuheben sind die zahlreichen Faktoren, die durch Veränderungen der extrazellulären Matrix (ECM) und Zytoskelett-Organisation, Zellmigration fördern. Bei gleichzeitiger Transformation der Zelle durch Hochregulation von onkogenen Faktoren, kann eine Umstrukturierung der ECM in Kombination mit Aktin-Modulation, zu fatalen Folgen hinsichtlich Organ-Infiltrierung und Metastasierung führen.

Fibronectin (*FN*, sK) und *Stromal cell-derived factor-1* (*SDF-1*) gehören zu den Targetgenen, die durch Mangel an MLL Protein positiv reguliert werden. FN ist einer der wichtigsten Komponenten der extrazellulären Matrix und maßgeblich an Migrationsmechanismen beteiligt (s. Kap. 3.2.3.2). Daneben wird SDF-1 von Stroma-Zellen im Knochenmark exprimiert, dient als Wachstumsfaktor und fördert dort, aufgrund seiner chemotaktischen Eigenschaften, das *homing* von hämatopoietischen Vorläufer-Zellen. In drei unabhängigen Arbeiten wurde gezeigt, dass Fibronectin (auf Filter aufgebracht oder in löslicher Form) den chemotaktischen Einfluss des SDF-1 auf CD34-positive Knochenmarkszellen von AML-Patienten, Vorläufer-Zellen aus CML-Patienten und pre-B Zellen aus ALL-Patienten, erheblich steigert und deren Migration fördert (Bhatia et al., 2001; Sbaa-Ketata et al., 2001; Voermans et al., 2002). Interessanterweise wurde ebenfalls der Stammzell-Marker CD34 in *MLL*^{-/-} verstärkt transkribiert (um den Faktor 110), der evtl. bei der FN/SDF-1 vermittelten Migration eine bislang unbekannte Rolle spielt. Die verstärkte Infiltration leukämischer t(4;11) Blasten ins ZNS und andere Organe, könnte möglicherweise durch einen solchen Mechanismus begründet sein. Mittels der subtraktiven

Kloniermethodik wurde das Gen *Rho GTPase activating protein* in $MLL^{-/-}$ Zellen als überexprimiert identifiziert. Dessen Genprodukt fördert die Polymerisation von F-Aktin und induziert dadurch eine Lamellipodien-Formation in $MLL^{-/-}$ Zellen (morphologisch eindeutig charakterisierbar, s. Kap. 3.2.1), die ebenfalls zu erhöhter Zell-Mobilität führt. Weitere Faktoren, die durch MLL Mangel hochreguliert werden und durch Strukturveränderungen der ECM das Migrationsverhalten von Zellen fördern, sind Serine protease inhibitor 4 (PN-1), die beiden Matrix-Metalloproteinasen Stromelysin 2 (MMP10) und Stromelysin 3 (MMP11; Rosenberg, 2002), das Proteoglykan Decorin und Glypican 4.

Inhibition von Proteindegradation

Mit der 7-fachen Überexpression des Gens, das für Suppressor of mit two 3 (SUMO) kodiert, wird ein Faktor hochreguliert, der durch Konjugation andere Proteine vor proteolytischem Abbau schützt. Liegt dieser Faktor aufgrund verstärkter transkriptioneller Induktion in größeren Mengen im Zellkern vor kann nicht ausgeschlossen werden, dass zusätzlich „Nicht-Target-Proteine“ vor Abbau durch das Proteasom geschützt werden. Dabei könnte es sich theoretisch auch um hochregulierte Proteine mit onkogenem Potential handeln.

Inhibition von Differenzierungsprozessen

Die Transkripte für das Protein RBP associated molecule RAM14-1, eine Spleiss-Variante des *Kyo-T* Gens, wurde in den *Mll knock-out* Zellen um den Faktor 80 hochreguliert. Durch Blockade der Notch-1 Signalkaskade nimmt dieser Faktor erheblichen negativen Einfluss auf wichtigste Entwicklungsprozesse (s. Kap. 3.2.4.2). Hierdurch werden z.B. *hes* Gene in ihrer Transkription gehemmt. Im Gegensatz dazu wurde bei Anwesenheit des MLL Proteins ($MLL^{+/+}$ Zellen) eines der *hes* Gene (*CLAST5/Hairy-related protein*, sK) verstärkt transkribiert. Die Tatsache, dass *Notch-1* und *RBP-Jκ knock-out* Mäuse zum exakt gleichen Zeitpunkt wie *Mll^{-/-}* Mäuse (Tag 10,5 p.c.) sterben, lässt Schlüsse auf Zusammenhänge zwischen diesen drei Faktoren zu. Möglicherweise sind die Blockade der Notch-1 Kaskade und die dadurch gehemmte Embryonalentwicklung für die Letalität der *Mll knock-out* Mäuse ursächlich.

Das humane Homolog zu RAM14-1, FHL-1C sowie die ungespleisste Form, wurden in $MLL^{-/-}$ Zellen und der t(4;11) Zelllinie SEM (ein intaktes *MLL* Allel) sehr stark exprimiert. In $MLL^{+/+}$ Zellen wurde nur eine abgeschwächte Expression detektiert. Im Gegensatz dazu konnten beide Transkripte in der T-Zelllinie Jurkat (zwei intakte *MLL* Allele) nicht nach-

gewiesen werden (Daten nicht gezeigt, Tanja Lehmann). Diese Tatsache könnte allein durch die MLL Dosisreduktion begründet sein. Eine mögliche Beeinflussung durch die Derivatproteine darf jedoch nicht ausser Acht gelassen werden.

Im Hinblick auf die Expression von T-Zell Markern in der pro-B Zelllinie ist dieses Ergebnis von Bedeutung. SEM Zellen exprimieren lymphoide sowie myeloische Marker. Mittels RT-PCR Reaktionen konnte eine GATA-3 Expression nachgewiesen werden (Ute Bangerter, unpublizierte Daten), die normalerweise ausschliesslich bei T-Zellen vorkommt und bisher nicht in der normalen B-Zell Entwicklung nachgewiesen werden konnte (Oosterwegel et al., 1992).

RAM14-1 hemmt den für die Lymphopoiese wichtigen Notch-1 Signalweg und für das humane Homolog wird die gleiche Funktion postuliert. Der Notch-1 Signalweg dirigiert lymphoide Vorläuferzellen in Richtung T-Zell-Entwicklung und inhibiert die Differenzierung von Monozyten und B-Zellen (De Smedt et al., 2002). Erwartungsgemäss wurden deshalb keine FHL-1C Transkripte in Jurkat Zellen (T-Zellen) nachgewiesen. In murinen Lymphknoten und normalen B-Zelllinien wurde bisher in Experimenten keine Expression der gespleissten und ungespleissten Form detektiert (Taniguchi et al., 1998). Die Inhibierung des Notch-1 Signalweges in SEM Zellen durch FHL-1C, ausgelöst durch Mangel an MLL Protein, könnte eine Erklärung sein, warum sich SEM Zellen trotz GATA-3 Expression nicht zu T-Zellen entwickeln können. Da eine gestörte Regulation der Notch-Signalkaskade in humanen und murinen T-Zellen zu Leukämien führt (Izon et al., 2002), wäre eine Involvierung von FHL-1C in den Pathomechanismus von Hochrisiko-Leukämien durchaus denkbar. Um weitere Schlüsse ziehen zu können, müsste zunächst abgesichert werden, dass FHL-1C in gesunden humanen B-Zellen nicht exprimiert wird und die Expression in SEM Zellen somit ein Charakteristikum der transformierten Zelllinie darstellt.

Nakamura et al. (2002) entdeckte das Protein Symplekin als Komponente im MLL MPSC. Dieser Faktor interagiert mit Huntingtin und ist an RNA Prozessierungsmechanismen beteiligt. In *Mll* Wildtyp-Zellen, in denen die Assemblierung des MLL MPSC durch Dimerisierung beider MLL Peptide induziert wird, ist die gespleisste Variante des *Kyo-T* Gens (RAM14-1) nur abgeschwächt detektierbar. Da die beiden intakten MLL Peptide die zwingend notwendigen Grundkomponenten zur Bildung des MLL MPSC sind, kann eine Assemblierung dieses multimeren Komplexes in *MLL*^{-/-} Zellen nicht erfolgen. Es ist nun anzunehmen, dass das unkomplexiert vorliegende Symplekin zusammen mit Huntingtin die Prozessierung verschiedenster Transkripte (z.B. FHL-1 zur RAM14-1) fördert. Zusätz-

lich wurde mit Hilfe der subtraktiven Klonierung in $MLL^{-/-}$ Zellen CASP, eine Spleiss-Variante des *CDP* Gens, identifiziert.

Oberflächenproteine

Über die Funktion und die Expressionsmuster des murinen Stammzell-Antigens CD34 ist bisher wenig bekannt. Dieses Gen wurde in Abwesenheit des MLL Proteins um den Faktor 110 hochreguliert. Eine Modellvorstellung, in der Stammzellen mit hohen Expressionsraten von CD34-Oberflächenprotein folglich ihr *Mll* Gen abgeschaltet haben, könnte aus diesen Erkenntnissen formuliert werden. Während hämatopoietischer Differenzierung wird das *Mll* Gen aktiviert und die *CD34* Expression reprimiert. Sollte sich dieser Mechanismus auf humane hämatopoietische Stammzellen übertragen lassen, hätte man z.B. durch RNAi Techniken die Möglichkeit, eine Langzeit-Kultivierung dieser Zellen ohne Verlust ihres Differenzierungspotentials durchführen zu können.

Verifizierung

Durch semi-quantitative RT-PCR Analysen mit cDNA aus der humanen t(4;11) Zelllinie SEM sollte eine Aussage über die Übertragbarkeit der murinen Daten auf humane Leukämie-Zellen gewonnen werden. Die folgenden Gene wurden, wahrscheinlich bedingt durch die MLL Dosisreduktion, in SEM Zellen ebenfalls stark bis sehr stark exprimiert: *vascular Smooth muscle α actin*, *Serine protease inhibitor 4*, *RBP associated molecule RAM 14-1*, *TALLA-1*, *SRCASM*, *Nectin-like protein 2*, *Adhesion protein RA175C*, *Four and a half LIM-domain 1*, *Cathepsin Z precursor*, *GM2 activator protein* und *TOB* (RT-PCR Experimente, Tanja Lehmann). Bei diesen Genen handelt es sich um solche, die zum größten Teil bewiesenermaßen an onkogenen Prozessen beteiligt sind. Es standen jedoch keine Daten aus gesunden B-Zellen für einen Expressionsvergleich zur Verfügung, weshalb die Beurteilung dieser Ergebnisse zu Zeit nicht möglich ist. Allerdings werden diese Untersuchungen momentan in der Arbeitsgruppe durchgeführt.

Bei der Betrachtung der Funktionen identifizierter MLL Targetgene wurde eine Reihe von Hinweisen erhalten, dass diese Genprodukte z.T. aktiv an der Entstehung von Leukämien beteiligt sein könnten. Zahlreiche Zusammenhänge zwischen einzelnen Faktoren und bereits publizierten onkogenen Prozessen tragen möglicherweise zur Aufklärung des bisher unbekanntes Pathomechanismus von Hochrisiko-Leukämien mit *MLL* Translokation bei.

Die Reduktion der MLL Proteindosis führt über die Aktivierung bestimmter Mechanismen zur Zelltransformation mit der Folge der Tumorentstehung. Dazu gehören verstärkte Proliferation durch Fehlsteuerung des Zellzyklus und gesteigerter Nukleotid-Biosynthese, erhöhte Migrationsaktivität durch veränderte ECM mit der Folge von Metastasierung/Organinfiltration, erhöhter Schutz vor proteolytischem Abbau nukleärer (Onko-) Proteine, und der Generation von Spleiss-Varianten mit z.T. negativem Einfluss auf verschiedene Differenzierungsprozesse. Der zusätzliche Einfluss der reziproken Derivatproteine auf pathologische Prozesse muss ebenfalls Beachtung finden und wird im Rahmen einer anderen Dissertation untersucht werden. Für manche dieser Fusionsproteine wurde ein onkogenes Potential bereits nachgewiesen. Zu diesen kumulativen Effekten mit negativer Ausprägung hinsichtlich der Leukämogenese kommt noch hinzu, dass eine Gegenregulation durch MLL induzierte Expression von Tumorsuppressoren (verschiedene Zellzyklus-Inhibitoren) fehlt. Daher ist es nicht verwunderlich, dass aus der Kombination dieser hoch- und herunterregulierten Faktoren eine Transformation resultieren kann. Diese Ergebnisse werden durch die Tatsache gestützt, dass für viele der in dieser Arbeit identifizierten Gene in MLL^{-/-} Zellen, eine Involvierung in Tumorerkrankungen bereits belegt wurde. Für eine Übertragbarkeit der erhaltenen murinen Expressionsdaten auf das humane System, müssen allerdings noch weitere Experimente durchgeführt werden.

4.3 Proteomanalyse

Mit Hilfe der zweidimensionalen Gelelektrophorese lassen sich mehrere tausend Proteine einer Zelle trennen, auf einem Gel darstellen und durch Kopplung an hochauflösende massenspektrometrische Verfahren charakterisieren. Die daraus gewonnenen Erkenntnisse lassen sich zur Struktur- und Funktionsaufklärung der einzelnen Proteine in der Zelle nutzen. Proteomics stellt somit einen neuen Ansatz dar, der die molekularen DNA- und RNA-Methoden sinnvoll komplementiert, der aber auch limitiert ist, weil nur diejenigen Proteine erfasst werden können, die in ausreichenden Mengen in einer Zellpopulation vorhanden sind. Das Ziel dieser Experimente bestand in der Verifizierung der zuvor durch DNA-Microchip-Arrays gewonnenen Daten differentieller Gentranskription auf der Protein-Ebene.

Die Trennung von Kern- und Plasmaproteinfraktionen erbrachte zusammen mit der etablierten Durchführung der 2D-Gelelektrophorese mit anschliessender Fluoreszenzfärbung gute Resultate. Zur Detektion der Proteine im zweidimensionalen Polyacrylamidgel wurde der Fluoreszenz-Farbstoff Ruthenium-II tris(bathophenan-throindisulfonat) (RuBPS) verwendet, der sich für eine schnelle und unproblematische Fluoreszenz-Bestimmung eignet. Dieser metallorganische Chelatkomplex bindet nichtkovalent an Proteinstrukturen und zeigt eine Sensitivität bis in den Nanogrammbereich (~100 ng). Vorteile gegenüber herkömmlichen Silberfärbungen sind zum einen das Fehlen von unspezifischen Hintergrundfärbungen und Überfärbungen, sowie zum anderen die bessere Quantifizierbarkeit, Reproduzierbarkeit (durch „*steady-state staining*“) und MALDI-TOF MS Kompatibilität. Quantifizierungen von Proteinen mit Hilfe von Silberfärbungen (z.B. nach (Shevchenko et al., 1996) sind wesentlich ungenauer, da der lineare Bereich nur sehr gering ist und für manche Proteine gar keine Linearität besteht. Die Nachweisgrenze für Proteine mit Silberfärbung ist jedoch weitaus niedriger.

Die jeweils 6 produzierten 2D-Gele beider Zelllinien bezüglich der Plasmaproteinfraktionen waren absolut reproduzierbar und liessen die Detektion differentiell exprimierter Proteine zu. Diese wurden nach einem Trypsin-Verdau aus dem Gel eluiert und massenspektrometrisch vermessen, konnten bisher jedoch noch nicht identifiziert werden. Ursache dafür könnte zum einen die noch nicht optimierte Protein-Elution, sowie die Art der Probenaufbereitung für die MALDI-TOF massenspektrometrische Untersuchung sein. Zunächst müsste durch weitere Vorversuche geklärt werden, welche Matrix (DHB oder HCCA) für diese Proteine am besten geeignet ist. Andererseits kann die Peptid-Quantität den limitierenden Faktor darstellen. Durch *poolen* gleicher Spots aus mehreren Gelen könnte hier Abhilfe geschaffen werden, vorausgesetzt, dass die Isolierung identischer Spots sichergestellt werden kann. Nachdem durch die Erstellung des Match-Sets die Position der differentiellen Proteine bekannt ist, wäre die Verwendung von *DryStrips*, die nur einen kleinen pH-Bereich um deren IEP aufweisen, die vielversprechendste Lösung. Es könnte dadurch eine wesentlich größere Proteinmenge aufgetragen werden, von der ausschliesslich die Proteine mit dem entsprechenden isoelektrischen Punkt in das Gel wandern und schliesslich in der 2. Dimension aufgetrennt werden. Durch diese Methode würde die Proteinmenge für folgende Experimente erhöht und gleichzeitig die Spotanzahl auf dem Gel verringert werden. Im Anschluss sollte dann eine Protein-Identifizierung mittels MALDI-TOF MS möglich sein.

Der Proteom-Vergleich von MLL^{+/+} und MLL^{-/-} Zellen (bezüglich der Zytoplasmaproteine) ergab, dass die Fraktionen der Wildtyp-Zellen 8 und die der *knock-out* Zellen 12 exklusiv exprimierte Proteine aufwiesen. Bei beiden Zelllinien wurden zusätzlich je 6 Proteine um mindestens das Zweifache über-, bzw. unterexprimiert. In Relation zu den Ergebnissen der Transkriptom-Vergleiche (DNA-Microchip-Arrays), konnte mit der zweidimensionalen Gelelektrophorese nur ein Bruchteil detektiert werden. Die Relation zwischen der Anzahl von transkriptionell aktivierten Genen der MLL^{-/-} Zellen und der MLL^{+/+} Zellen konnte jedoch auf Protein-Ebene durch die 2D-Gelelektrophorese bestätigt werden. Die DNA-Microchip-Arrays zeigten, dass wesentlich mehr Gene in den *knock-out* Zellen hochreguliert wurden. Parallel dazu konnten auch mehr differentielle Protein-Spots auf dem MLL^{-/-} Gel identifiziert werden.

Der limitierende Faktor ist die Proteinmenge, da mittels Fluoreszenzfärbung nur Protein-Spots größer als 100 ng detektiert werden konnten. Stark überexprimierte Struktur- und Matrixproteine, sowie am Zell-Stoffwechsel beteiligte Enzyme sollten somit theoretisch auf einem 2D-Gel sichtbar sein.

Zur Detektion von Kernproteinen ist die Methode der zweidimensionalen Gelelektrophorese scheinbar nicht geeignet. Obwohl gut auswertbare 2D-Gele gewonnen wurden, konnte bei den Kern-Proteinfractionen beider Zelllinien nur je ein überexprimiertes Protein detektiert werden. Die meisten Kernproteine, die in diesem Fall von Bedeutung waren, sind Transkriptionsfaktoren. Deren Vorkommen in der Zelle ist so gering, dass auch eine 10-fache Überexpression eines solchen Faktors nicht ausreicht, um ihn unter Verwendung dieser Methode detektieren und identifizieren zu können.

Zur Erhöhung der Nachweisgrenze könnte auf eine empfindlichere Färbemethode, zum Beispiel die Silberfärbung nach Shevchenko, zurückgegriffen werden, die jedoch in anderen Bereichen Nachteile (z.B. schlechtere MALDI Kompatibilität) aufweist. Ebenfalls könnten in weiteren Experimenten *DryStrips* mit einem geringeren pH-Bereich gewählt werden, die das Auftragen größerer Proteinmengen und dadurch eine bessere Detektion erlauben. Trotz dieser Möglichkeiten zur Optimierung bleibt die Detektion differentiell exprimierter Kernproteine fraglich.

4.4 Komplementationstudie

Das Komplementationsexperiment wurde u.a. vor dem Hintergrund durchgeführt, die Eignung des Testsystems zu überprüfen. In verschiedenen Experimenten (DNA-Microchip-Arrays) wurde der Einfluss humaner Proteine (hMLL und die Derivatproteine der4, der11 und der4/11) auf das murine Genom untersucht. Daraufhin musste analysiert werden, ob diesbezüglich die Beibehaltung der Funktionalität bei der Überschreitung der Artengrenze überhaupt möglich und auswertbar ist. Für diese Analyse eignete sich die stabil transfizierte Zelllinie MLL^{-/-}::hMLL im Vergleich zur Wildtyp Linie MLL^{+/+}.

Es sollte überprüft werden, ob das humane MLL Protein die endogene Funktionalität des murinen Proteins in den MLL^{-/-} Zellen komplementieren kann. Bei einem positiven Ergebnis hätte das Transkriptionsmuster der Transfektante im Wesentlichen dem der murinen Wildtyp-Zellen entsprechen müssen. Die Scatter-Blot Darstellung des Vergleichs von MLL^{-/-} mit MLL^{-/-}::hMLL hätte dem von MLL^{+/+} vs. MLL^{-/-} stark ähneln und ebenfalls einen relativ niedrigen Pearson-Korrelationskoeffizienten aufweisen sollen. Wider Erwarten war dies jedoch nicht der Fall. Dafür könnte es verschiedene Gründe geben. Ein Erklärungsansatz wäre, dass das humane MLL Protein tatsächlich keinen Einfluss auf das murine Genom hat. Dies scheint bei einer 92%ige Identität zwischen dem humanen und murinen Protein jedoch unwahrscheinlich zu sein, zumal die Identität in wichtigen Domänen sogar bei 98% liegt. Zusätzlich spricht dagegen, dass schon in verschiedenen Experimenten gezeigt wurde, dass humane Proteine das murine Genom beeinflussen können (Milne et al., 2002).

Eine weitere Erklärungsmöglichkeit für fehlende Komplementation könnte die Interaktion des endogenen murinen MLL-C Peptids mit den humanen MLL Peptiden sein. Wie zuvor in Kap. 3.6.2 beschrieben, wird in den *Mll knock-out* Zellen der C-terminale Teil des MLL Proteins (Exon 17-37) exprimiert. Dieser wird an CS1 und CS2 proteolytisch gespalten und generiert das murine MLL-C Peptid, das aufgrund der FYRC-Dimerisierungsdomäne mit einem MLL-N Peptid und den weiteren Faktoren des MLL MPSC assemblieren könnte. Ein endogenes MLL-N Peptid mit FYRN-Domäne ist jedoch in den *knock-out* Zellen nicht vorhanden. In der Literatur wurde bereits beschrieben, dass das humane MLL Protein ebenfalls in murinen Zellen proteolytisch gespalten wird (Milne et al., 2002). Aufgrund der Annahme, dass auch in diesem Fall beide humanen MLL Peptide in der Zelle vorliegen, könnte es mögliche Interaktionen mit dem endogenen murinen MLL-C Peptid geben, die evtl. eine gestörte Komplex-Assemblierung zur Folge haben. Möglicherweise

besitzt der murine MLL MPSC andere Faktoren, die nicht mit den humanen MLL Peptiden assemblieren können. Milne et al. konnte in seiner Arbeit nachweisen, dass eine Komplementation der gleichen MLL^{-/-} Zellen mit der humanen MLL Konstrukt nur zu einem geringen Teil möglich ist. Er verglich in seinen Experimenten die hMLL abhängige Promotor-Aktivierung des murinen *hox c8* Gens. MLL^{-/-} Zellen zeigten keine *hox c8* Genexpression, während in den, mit dem humanen MLL Konstrukt, transfizierten Zellen eine geringe Expression nachweisbar war. Diese reichte jedoch bei weitem nicht an die Stärke der Wildtyp-Zellen heran.

Wahrscheinlicher erscheint jedoch die Hypothese, dass bestimmte Eigenschaften des verwendeten Zellsystems zu diesem unerwarteten Resultat führten. Die aus den transgenen Mäusen entnommenen Zellen weisen einen charakteristischen Entwicklungsstatus auf, der dem Tag 10,5 p.c. entspricht. Theoretisch müsste sich das Transkriptom zu diesem Zeitpunkt noch flexibel verändern lassen, da zu diesem Zeitpunkt die Organogenese beginnt, für die ein entsprechendes spezifisches Transkriptionsmuster über Änderung des „Histon-Codes“ neu definiert wird. Möglicherweise ist aber zu diesem Zeitpunkt bereits das sogenannte „transkriptionelle Gedächtnis“, durch spezifische MLL MPSC und PcG Proteinkomplexe etabliert, doch schon so stabil, dass es durch äussere Einflüsse nur noch marginal beeinflussbar ist. Das würde für die Tatsache sprechen, dass MLL und davon abgeleitete reziproke Derivatproteine nur in einem engen Zeitfenster (d.h. bei der Maus zwischen Tag 8- 10,5 p.c.) Einfluss auf Genexpressionsmuster haben.

Das Problem könnte ebenfalls mit dem Alter der Zelllinie begründet sein. Die Zelllinien wurden 1996 etabliert und es ist nicht auszuschliessen, dass durch Alter, häufiges Passagieren oder übergeordnete Mechanismen die transkriptionelle Situation in einer unbekannt Art und Weise fixiert wurde. Zur Klärung dieses Problems müsste man eine „Reprogrammierung“ der MLL^{+/+} und MLL^{-/-} Zellen durchführen, indem man sie mit bestimmten Substanzen behandelt, die das epigenetische Programm der Zellen eliminieren. Anschliessend würde das Transkriptionsmuster unter Einfluss verschiedener Expressionskonstrukte (*hMLL*, *der4*, *der11* oder *der4/11*), neu definiertbar sein. Spätestens zu diesem Zeitpunkt sollten erhebliche Unterschiede im Expressionsmuster zu erwarten sein.

Die sinnvollste Überprüfung des biologischen Testsystems wäre durch die stabile Transfektion des murinen *Mll* Konstrukts in die MLL^{-/-} Zelllinie erreichbar. Sollte die Komplementation der murinen MLL^{-/-} Zellen durch das entsprechende murine Expressionskonstrukt ebenfalls negativ ausfallen und nicht exakt das Expressionsmuster der *Mll*

Wildtyp-Zellen widerspiegeln, wäre sichergestellt, dass der Grund für das Scheitern bei der epigenetischen Fixierung der $MLL^{-/-}$ Zelllinie zu suchen ist. Damit könnte gezeigt werden, dass die heterologe Komplementation Mensch-Maus nicht das eigentliche Problem darstellt. Ein murines *Mll* Expressionskonstrukt stand im Rahmen dieser Arbeit nicht zur Verfügung.

4.5 Identifizierung der Targetgene der Derivatproteine MLL•AF-4 und AF-4•MLL

Sollte sich herausstellen, dass die epigenetische Programmierung der Zellen verantwortlich für die hier erhaltenen Ergebnisse ist, hätte das einen dramatischen Einfluss auf unser Verständnis vom Leukämie-Pathomechanismus. Es würde nämlich bedeuten, dass MLL oder seine Derivate über einen „hit and run“ Mechanismus arbeiten. Diese Annahme hat eine enorme Bedeutung für die Auswertung der weiteren Versuche. Neben $MLL^{-/-}::hMLL$ wurden weitere Zelllinien durch stabile Transfektion etabliert ($MLL^{+/+}::der4$, $MLL^{+/+}::der11$, $MLL^{+/+}::der4/11$ und $MLL^{+/+}::BC230A$), die ebenfalls durch Expressionsvergleiche charakterisiert wurden. Die Erkenntnisse über Targetgene der Derivatproteine der4, der11 und der4/11 und deren Einfluss auf das Transkriptionsmuster sollten zum Verständnis und zur Aufklärung des Pathomechanismus der t(4;11) Leukämie beitragen. Diese Fragestellung war ein wichtiger Teil dieser Arbeit. Die Daten, die aus verschiedenen Affymetrix DNA-Microchip-Hybridisierungen gewonnen wurden, zeigen kaum unterscheidbare Expressionsmuster zu den verwendeten Kontroll-Zellen. Dies lässt vermuten, dass auch die $MLL^{+/+}$ Zelllinie epigenetisch fixiert ist. Ein Einfluss der Transgene wäre voraussichtlich nur nach „Reprogrammierung“ der Zellen messbar.

Die Resultate, die durch die DNA-Microchip-Hybridisierungen erhalten wurden, sind in Kap. 3.5 in Scatter-Blots dargestellt. Anhand der graphischen Auswertung ist ersichtlich, dass sich die Expressionsmuster der stabil transfizierten Zelllinien nur marginal von den *mock*-transfizierten Wildtyp-Zellen unterscheiden. Die kotransfizierte Zelllinie stellt im weitesten Sinne die Situation in einer arretierten, leukämischen B-Zelle dar. Neben dem Wildtyp MLL Protein liegen beide Derivatproteine in ausreichenden Mengen in der Zelle vor. Da dieser Proteinstatus in leukämischen Blasten der Auslöser eines onkogenen Pathomechanismus ist, wurden hier aussagekräftige Expressionsunterschiede erwartet. Aus früheren Arbeiten war bekannt, dass der4 onkogenes Potential besitzt, das durch Transfor-

mation von NIH 3T3 Zellen nachgewiesen wurde (Ute Bangerter; Focus Formation Assay, unveröffentlichte Daten).

Wahrscheinlich ist das den Experimenten zugrunde liegende biologische Testsystem mit den Ausgangs-Zelllinien MLL^{-/-} und MLL^{+/+} für die unerwarteten und wahrscheinlich nicht realistischen Resultate verantwortlich. Die Ergebnisse der durchgeführten Komplementationsanalyse (s. Kap. 3.4) belegen, dass das hier verwendete Testsystem diesbezüglich problematisch ist. Die Vermutung, dass humane Proteine keinen Einfluss auf das murine Genom haben könnten, ist unwahrscheinlich.

Wahrscheinlicher ist das zuvor in Kap. 4.4 beschriebene starre Transkriptionsmuster, das evtl. durch Alterung der Zelllinien, zu häufiges Passagieren oder wahrscheinlicher durch epigenetische Mechanismen zustande kam und in diesem Stadium nur noch marginal beeinflussbar ist. Hier könnte die bereits angesprochene „Reprogrammierung“ der Zellen mit anschließender Transfektion, Abhilfe schaffen.

Aufgrund der Unsicherheit über die Eignung des verwendeten biologischen Testsystems wird hier nicht näher auf die Ergebnisse eingegangen. Die DNA-Microchip-Hybridisierungen werden im Rahmen einer weiteren Dissertationsarbeit, nach „Zell-Reprogrammierung“ und anschließender Transfektion, wiederholt. Daneben wäre denkbar, RNA aus Zellen zu isolieren und für DNA-Microchip-Hybridisierungen zu nutzen, die aufgrund von Transgenen (z.B. der4) Foci bilden. Der Expressionsvergleich dieser transformierten Zellen mit untransformierten müsste mit hoher Wahrscheinlichkeit zur Identifizierung differentiell exprimierter Gene mit onkogenem Potential führen.

4.5.1 Hypothetischer Pathomechanismus der t(4;11) Translokation

In den *Mll knock-out* Zellen konnten im Vergleich zu den MLL^{+/+} Zellen mittels DNA-Microchip-Arrays 136 verstärkt transkriptionell aktivierte Gene identifiziert werden, von denen ein Großteil onkogenes Potential besitzt und bereits in transformierten Zellen in überexprimiertem Zustand detektiert wurde. Der Zusammenhang zwischen der Abwesenheit des MLL Proteins und der Hochregulation dieser interessanten Targetgene ist möglicherweise für den pathogenetischen Mechanismus der Leukämogenese von großer Relevanz.

Die folgenden Abbildungen geben einen schematischen Überblick über den hypothetischen Ablauf der Transkriptionsregulation in einer t(4;11) Zelle, wobei in Abb. 4.1 zunächst die Wildtyp Situation gezeigt wird.

Aufgrund der bekannten Fakten lassen sich möglicherweise Rückschlüsse auf die Leukämogenese ziehen. Das Derivatprotein der11 stellt den „kleinsten gemeinsamen Nenner“ aller t(11q23) Translokationen dar und leistet vermutlich einen erheblichen Beitrag zur Entstehung einer Leukämie. Das Partnergen beeinflusst wahrscheinlich die Art der auftretenden Leukämie. Die meisten Brüche bei den Translokationen im *MLL* Gen finden in der Bruchpunkt-Cluster-Region (bcr; Exon 8–14; gehäuft im Intron 9) statt. Das resultierende der11 Protein besitzt nur noch den N-terminalen Bereich des MLL Proteins, bestehend aus dem AT-Haken, der MT-Domäne und der subnukleären Lokalisierungssequenz (snLS). Die für die Interaktion mit dem MLL-C Peptid notwendige Dimerisierungsdomäne FYRN ist im Derivatprotein nicht mehr vorhanden. Die Dimerisierung beider Peptide ist jedoch für die Komplex-Assemblierung, Lokalisation und Ausübung der Wildtyp-Funktion essentiell. Das der11 Protein stellt somit keine MLL MPSC Komponente mehr dar, kann aber aufgrund der vorhandenen snLS an die gleichen Targetgene binden. Es konnte bereits nachgewiesen werden, dass verschiedene Derivatproteine (MLL•AF-4, MLL•AF-9, MLL•ENL und MLL•AF-17) das gleiche Kernlokalisationsmuster wie endogenes MLL Protein aufweisen (Butler et al., 1997; Li et al., 1998; Yano et al., 1997). Zu dem gleichen Ergebnis kamen wir durch eigene Arbeiten, bei denen mittels des monoklonalen Antikörpers E3 (gegen Epitope im Exon 3 gerichtet), ein MLL•LacZ Protein im gleichen Verteilungsmuster detektiert wurde (s. Kap. 3.2.2).

Aufgrund der Bindungsmöglichkeit an gleiche Targetgene ist eine Konkurrenzreaktion des der11 zum endogenen MLL MPSC (durch intaktes *MLL* Allel generiert) zu erwarten. Möglicherweise könnte das der11 Protein quantitativer an die Targets binden, weil es kleiner als der MLL MPSC ist und deshalb sterische Vorteile aufweist. Das an die Targetsequenzen gebundene der11 Protein könnte nun die Bindung des MLL MPSC blockieren und folglich zu einer Situation, die einem *MLL knock-out* Status gleicht, führen. Aus diesem Grund sind die in den *MLL*^{-/-} Zellen hochregulierten Gene von besonderem Interesse, da diese möglicherweise durch die o.g. Blockade auch in dieser Situation transkriptionell aktiviert werden. Wie zuvor beschrieben sind viele dieser hochregulierten Gene mit onkogenen Mechanismen assoziiert, weshalb auch eine Beteiligung an der Leukämie-Entstehung diskutiert werden kann. Der N-terminalen MLL Region im Derivatprotein fehlt die transkriptionsaktivierende Domäne, die im MLL-C Peptids des Wildtyp Proteins lokalisiert ist. Falls jedoch das Partnergen eine solche Sequenz besitzt, kann nach Kernlokalisierung und DNA-Bindung (durch MLL-N Anteil) eine ektopische Genexpression stattfinden.

In der Literatur wurde beschrieben, dass die Methylierung des Lysin-Restes K4 des Histons H3 durch die SET-Domäne essentiell für die Transkriptionsaktivierung ist und zugleich die Bindung „genrepressiver“ Proteine (z.B. PcG-Komplexe) inhibiert (Nishioka et al., 2002). Das der11 Protein (ohne SET-Domäne) kann somit über unterschiedliche Mechanismen die endogene MLL MPSC vermittelte MLL Wildtyp-Funktion stören. Zum einen über die Blockade der Targetgene (fehlende Aktivierung) und zum anderen erhöht es dadurch indirekt die Wahrscheinlichkeit der PcG-Komplex Bindung (aktive Repression).

Das reziproke Fusionsprotein einer t(11q23) Translokation besteht aus dem N-terminalen Bereich des entsprechenden Partnergens und der C-terminalen Region (meist ab Exon 10) des MLL Proteins. Dieses Derivatprotein besitzt neben den beiden Dimerisierungsdomänen FYRC und FYRN, ebenfalls die zwei Protease-Erkennungssequenzen CS1/CS2. Es ist anzunehmen, dass dieses Protein ebenso wie das Wildtyp Protein in zwei Peptide gespalten wird. Aus dieser proteolytischen Spaltung resultiert zum einen das Wildtyp identische MLL-C Peptid und zum anderen ein N-terminales Fragment, das neben der DNA-bindenden PHD-Region die FYRC-Interaktionsdomäne besitzt. Die Assemblierung des MLL MPSC könnte durch den hohen Überschuss an MLL-C Peptid gestört werden, ebenso wie durch die Assoziation mit dem Derivat Peptid. Über die FYRC-Domäne kann es mit MLL-C interagieren und die Assemblierung der 27 weiteren Faktoren stören oder gar inhibieren.

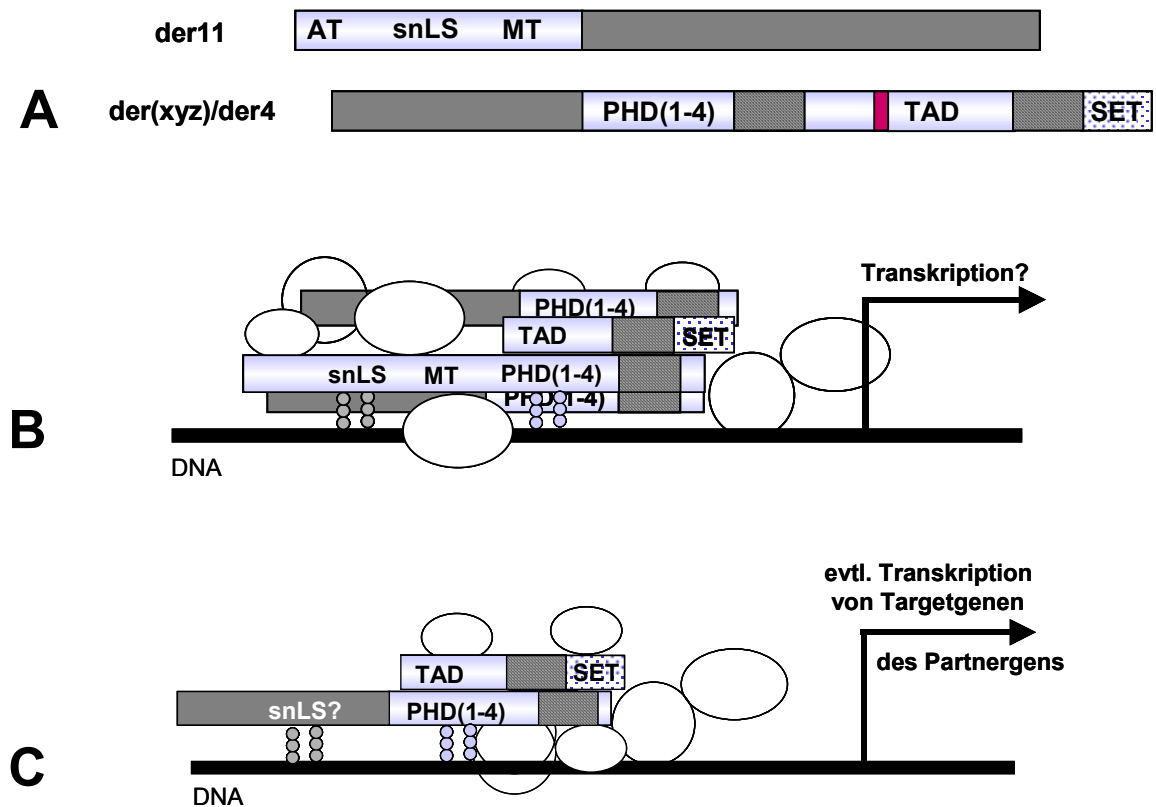


Abb. 4.3: Hypothetische Funktionsweise des reziproken Proteins

A: Darstellung der beiden reziproken Derivatproteine, resultierend aus einer MLL Translokation, **B:** hypothetische Funktionsweise des reziproken Derivatproteins; gestörte Komplex-Assemblierung durch Derivat-N Peptid, Konkurrenz-Reaktion. **C:** Mechanismus einer ektopten Genaktivierung, initiiert durch das Partnergen. AT: AT-Haken, snLS: subnukleäre Lokalisierungssequenz, MT: Methyltransferase-Domäne, PHD: Zinkfinger-Domäne, gestreifte Boxen: Dimerisierungsdomänen, SET: Methyltransferase-Domäne, TAD: Transkriptionsaktivator-Domäne, große graue Boxen: Proteinanteil des Partnergens, helle Boxen: MLL Proteinanteil, weisse Kreise: MLL MPSC Faktoren.

Die im reziproken Derivatprotein vorhandene SET-Domäne ist vielleicht trotz veränderter Komplex-Zusammensetzung in der Lage, den Lysin-Rest K4 des Histons H3 zu methylieren und aktiv die Transkription zu steuern. Durch eine Kernlokalisierungssequenz und evtl. vorhandene DNA-Bindungsdomäne des Partnergens, könnte der Proteinkomplex an Targetgene des Partnergens binden und über die TAD-Domäne dort ektope Transkription verursachen. Für das reziproke Fusionsprotein der4 kann dieser Mechanismus vermutet werden. In Blackwell-Weintraub Experimenten (DNA-Protein Bindungsstudie; Ralf Greim, unveröffentlichte Daten) konnte gezeigt werden, dass der N-Terminus des AF-4

Proteins eine DNA-Bindungsdomäne besitzt. Diese ist in dem Bereich der AS 60-297, zwischen der NHD-Domäne und der Bruchpunkt-Cluster-Region, lokalisiert.

Diese Hypothese würde z.B. erklären, warum *MLL* Translokationen mit verschiedenen Partnergenen zu unterschiedlichen Formen von Leukämien führen und ebenso unterschiedliche Heilungschancen aufweisen.

4.6 Identifizierung eines artifiziellen Promotors im *Mll* Gen

Die Hybridisierung des PCR-Fragmentes Exon 28-33 mit einer radioaktiv markierten Sonde aus *MLL*^{-/-} Zellen ergab, dass trotz LacZ/neo^R-Insertion in Exon 4, der C-terminale Bereich des *Mll* Gens transkribiert wurde. Verwechslung und Zell-Kontamination mit *MLL*^{+/+} Zellen konnten aufgrund von RT-PCR Experimenten ausgeschlossen werden. Zusätzlich wurde bestätigt, dass die *MLL*^{-/-} Zellen tatsächlich einen homozygoten *knock-out* beider *Mll* Allele trugen.

Zur Aufklärung dieses Phänomens wurden Experimente auf RNA- und auf Proteinebene durchgeführt. Mittels RT-PCR Reaktionen konnte die potentielle Transkriptionsinitiationsstelle bis auf einen Bereich von ca. 150 bp näher eingegrenzt werden, die im ersten Drittel des Introns 11 lokalisiert ist. Eine exakte Bestimmung wird, im Rahmen einer anderen Dissertationsarbeit in unserer Arbeitsgruppe, durch *Primer-Extension* Experimente durchgeführt werden.

Ein Nachweis über die Translation dieses C-terminalen Transkriptes wurde durch immunohistochemische Untersuchungen mit dem monoklonalen Antikörper gegen Exon 28 erbracht. Die eindeutige Lokalisation dieses Peptids bestätigt die Vermutung, dass im *Mll* Gen ein artifizieller Promotor vorhanden sein muss. Voraussichtlich kann die RNA erst ab Exon 17 translatiert werden (erstes ATG nach Intron 11) und das resultierende Peptid wird wahrscheinlich an CS1/CS2 proteolytisch gespalten. Das würde bedeuten, dass ein Wildtyp *MLL*-C Peptid synthetisiert wurde, dessen diffuse Kernlokalisierung mit aktuellen publizierten Daten korreliert (Hsieh et al., 2003). Aufgrund dieser Daten stellt sich die Frage, unter welchen Voraussetzungen an diesem artifiziellen Promotor Transkription initiiert wird. Der funktionelle Nachweis des alternativen Promotors konnte bisher noch nicht erbracht werden.

Die Existenz eines kryptischen Promotors, der nur bei Ausfall des eigentlichen Promotors aktiv wird, ist bisher z.B. aus dem *c-myc* Gen bekannt. Dort findet am Ende des Exons

1 eine „Terminationsreaktion“ statt, da die RNA-Polymerase nicht über das Ende dieses Exons hinaus lesen kann. Vor Exon 2 liegt ein kryptischer Promotor, der artifizielle Enhancer-Elemente zur Transkriptions-Initiierung benötigt. Dieses verkürzte Transkript wurde in allen *c-myc* exprimierenden Zellen nachgewiesen (Bentley und Groudine 86) und ist stark onkogen.

Aufgrund der Homologie zwischen der murinen und humanen „Promotor-Region“ geben die in dieser Arbeit gewonnenen Erkenntnisse unter Umständen Hinweise auf die Ursache der Translokationsentstehung im *MLL* Gen. Der potentielle kryptische Promotor ist in der Bruchpunkt-Cluster-Region lokalisiert. Zur Bindung des gesamten Pol-II Komplexes und zur nachfolgenden Transkription, muss die DNA entspiralisiert werden und einzelsträngig vorliegen. In diesem Zustand ist DNA extrem empfindlich und wird leicht hydrolysiert. Eventuell lassen sich die gehäuften DNA-Brüche in dieser Region durch die Eigenschaften dieses neu entdeckten Promotors erklären.

5 Material und Methoden

5.1 Biologisches Material

5.1.1 Zelllinien, Charakteristika und Kulturbedingungen

Zusammensetzung des Mediums: (Medium A)	10%	Dulbeccos Modified Eagle Medium
	2 mM	FCS
	100 U/ml	Glutamax
	50 μ M	Penicillin/Streptomycin
		β -Mercaptoethanol

Zusammensetzung des Mediums: (Medium B)	10%	RPMI
	2 mM	FCS
	100 U/ml	Glutamax
	50 μ M	Penicillin/Streptomycin
		β -Mercaptoethanol

Kulturbedingungen: 5% CO₂; 95% rel. Luftfeuchtigkeit, 37°C

MLL^{+/+} und MLL^{-/-} Zellen

Bei den isogenen Zelllinien MLL^{+/+} und MLL^{-/-} handelt es sich um murine embryonale Fibroblasten (s. Kapitel 2.4), die sich ausschliesslich durch die An- bzw. Abwesenheit eines funktionellen *Mll* Gens unterscheiden. Sie wurden am Tag 10,5 aus einem absterbenden Mäuse-Embryo isoliert und durch Infektion mit dem T-Antigen des Polyoma-Virus immortalisiert. Die Zelllinien MLL^{+/+} und MLL^{-/-} werden mit Medium A kultiviert.

Zum Passagieren wurden die adhärenenten Zellen mit 1x PBS gewaschen und mit 1x Trypsin/EDTA von der Petrischale abgelöst. Im Anschluss wurden die Zellen bei 700 UpM für 4 min pelletiert und in einer Verdünnung von ca. 1:5 wieder ausgebracht.

Ca. 5×10^6 Zellen (eine 80-90% konfluente Platte) wurden zum Anlegen von Stocks abgelöst und pelletiert. Das Pellet wurde in 900 μ l FCS resuspendiert, mit 100 μ l DMSO versetzt und sofort im Stickstoffdampf eingefroren. Am nächsten Tag erfolgte die dauerhafte Lagerung durch Eintauchen in flüssigen Stickstoff.

Transfizierte Zelllinien

In der folgenden Tabelle werden die stabil transfizierten Zelllinien mit ihren entsprechenden Selektionsmedien zusammengefasst. Erklärungen über die Klonierungsstrategien der Vektorkonstrukte und den Transfektionsexperimenten sind im Kapitel 3.1 zu finden.

Zelllinie	Medium A	Zeocin	Doxycyclin	Hygromycin	Histidinol
MLL ^{-/-} ::tTA	X	50 µg/ml	-	-	-
MLL ^{-/-} ::tTA::hMLL MLL ^{-/-} ::tTA::der4 MLL ^{-/-} ::tTA::der11 MLL ^{-/-} ::tTA::der4/der11	X	50 µg/ml	10 µg/ml	120 µg/ml	-
MLL ^{-/-} ::BC230A MLL ^{-/-} ::BC230A-hMLL	X	-	-	120 µg/ml	-
MLL ^{+/+} ::BC230A MLL ^{+/+} ::BC230A-der4 MLL ^{+/+} ::BC230A-der11	X	-	-	120 µg/ml	-
MLL ^{+/+} ::BC230A-der4/der11	X	-	-	120 µg/ml	2 mM

Tab. 5.1: Stabil transfizierte Zelllinien und ihre entsprechenden Selektionsmedien

293 Zellen

Die zur Transfektion des Promotorkonstruktes/EGFP verwendeten 293 Zellen sind humane embryonale Nierenkarzinom-Zellen und wurden in Medium B kultiviert.

Das Passagieren und Stock-Anlegen wurde nach der unter „MLL^{+/+}“ und „MLL^{-/-}“ beschriebenen Methode durchgeführt.

5.1.2 Bakterienstämme

Sure™ e14-(mcrA), Δ(mcrCB-hsdSMR-mrr)171, endA1, supE44, thi-1, gyrA96, relA1, lac, recB, recJ, sbcC, umu::Tn5, kanR, uvrC, [F', proAB, lacIqZΔM15, Tn10 (tetR)]

Top10 E.coliF' (tetR), mcrA Δ (mrr-hsdRMS-mcrBC), phi 80 Δ lac Δ M15, Δ lacX74deoR, recA1, araD139, Δ(ara, leu) 7697, galU, galK, lambda⁻, rpSL (str^r) endA1, nupG

5.2 Nukleinsäuren

5.2.1 Oligonukleotide

Die Oligonukleotide wurden von den Firmen Sigma, Pharmacia und BioSpring bezogen. Sie dienten als spezifische PCR-Primer, zur Einführung benötigter Schnittstellen für Restriktionsendonukleasen und als Sequenzierprimer. Eine Liste der in dieser Arbeit verwendeten Oligonukleotide befindet sich im Anhang.

5.3 Enzyme und Feinchemikalien

5.3.1 Enzyme

Restriktionsendonukleasen, Polymerasen, DNA-Ligasen, Gold-Star Taq-Polymerase, T₄-Polynukleotidkinase, rNasin[®] RNase Inhibitor, Reverse Transkriptase (Superscript[™] II), alkalische Phosphatase, Benzonase[®] und pankreatische RNase wurden von den Firmen Roche Diagnostics, Invitrogen, New England Biolabs (NEB), Eurogentec, Promega und Amersham Biosciences bezogen.

5.3.2 Feinchemikalien

Alle gebräuchlichen Chemikalien waren vom Reinheitsgrad „pro analysi“ und wurden von den Firmen Merck, Roth, und Sigma bezogen. Medien und Medienbestandteile für die Zellkultur (DMEM, Glutamax, Penicillin/Steptomycin, β-Mercaptoethanol, Trizol) wurden von Invitrogen, FCS Gold von PAA Laboratories bezogen. Das Radioisotop [$\alpha^{32}\text{P}$]-dATP (10mCi/ml, 3000 Ci/mmol) stammte von der Fa. Amersham Biosciences. Chemikalien, die für die zweidimensionale Gelelektrophorese verwendet wurden, wurden von der Fa. SERVA bezogen.

5.4 Methoden

5.4.1 Transfektion / Selektion

Die stabile Transfektion der adhären Zellen MLL^{+/+} und MLL^{-/-} wurde mittels FuGENE™ 6 Transfektionsreagenz durchgeführt (Lipofektion; Roche Diagnostics).

Für die Liposomenbildung wurden zunächst 272 µl serumfreies Medium vorgelegt, in welchem anschliessend 18 µl des FuGENE™ 6 Transfektionsreagenzes pipettiert wurden. Dem Ansatz wurden 6 µl DNA-Lösung (1µg/µl; Verhältnis FuGENE:DNA von 3:1) zugegeben. Diese Lösung wurde nach einer Inkubationszeit von 15 min bei RT tröpfchenweise auf die mit frischem Medium versorgten Zellen pipettiert. Die optimale Zelldichte für eine Transfektion liegt bei 60-70%iger Konfluenz.

Nach 48 stündiger Inkubation bei 37°C/5% CO₂/95% Luftfeuchtigkeit wurde entsprechend der Resistenzkassetten der unterschiedlichen Vektoren mit der Selektion begonnen. Eine tabellarische Auflistung der verwendeten Selektionsantibiotika mit den Konzentrationsangaben befindet sich im Kapitel 5.1.1.

5.4.2 Molekularbiologische Methoden

Die allgemein gebräuchlichen molekularbiologischen Methoden wurden, soweit nicht im einzelnen erwähnt, nach Maniatis (Sambrook et al., 1989) durchgeführt und sind daher hier nicht oder nur kurz beschrieben.

5.4.2.1 PCR-Techniken und Programme

PCR-Reaktionen wurden mit dem Gerät GeneAmp PCR System 9600 (Fa. Perkin Elmer) im 50 µl Maßstab durchgeführt. Für jede Reaktion wurde sowohl eine Positiv- als auch Negativkontrolle mitgeführt. Als Polymerase wurde GoldStar Taq (Eurogentec) mit dem entsprechenden Puffern verwendet. Für Amplifikate mit einer Länge größer als 2000 bp wurden die Reagenzien des *long-range Kits Expand 20kb^{PLUS} PCR System* (Fa. Roche Diagnostics) nach Angaben des Herstellers eingesetzt. Für die Amplifikation von cDNA wurde das *touch-down* Programm verwendet. Die erhaltenen Amplimere wurden mittels Gelelektrophorese analysiert und bei Bedarf mit dem *Quiagen PCR Purification Kit* aufgereinigt.

<u>Standardansatz:</u>	100 ng	Template (DNA)
	5 µl	2 pmol/µl je Oligonukleotid
	5 µl	2,5 mM dNTP-Lösung
	4 µl	25 mM MgCl ₂ -Lösung
	2 U	Taq-Polymerase
	ad 50 µl	Wasser

Standard-Programm: 94°C/2' + [94°C/20'' + 58°C/30'' + 72°C/1']_{35x} + 72°C/3'

long-range Programm: [94°C/2']_{1x}
 [92°C/10'' + 62°C/30'' + 68°C/18'']_{10x}
 [92°C/10'' + 62°C/30'' + 68°C/18'+10'' pro Zyklus]_{25x}
 [68°C/7']_{1x}

touch-down Programm: [94°C/2']_{1x}
 [94°C/30'' + 65°C/1' -0,5°C pro Zyklus + 72°C/90'']_{26x}
 [94°C/30'' + 52°C/1' + 72°C/90'']_{14x}
 [72°C/3']_{1x}

5.4.2.2 DNA-Sequenzierung

Die DNA-Sequenzierungen wurden mit dem Gerät ABI Prism 310 Genetic Analyzer (Fa. ABI, Applied Biosystems) im 10 µl Maßstab durchgeführt. 300 ng Template wurden mit 3 µl Lösung aus dem *Big Dye Terminator Cycle Sequencing Kit* (Fa. Perkin Elmer) und 6 pmol Primer gemischt und mit dem folgenden Programm im PCR System 9600 (Fa. Perkin Elmer) amplifiziert:

Sequenzier-Programm: [96°C/10'' + 56°C/10'' + 60°C/4']_{25x}

Im Anschluß daran wurde das Amplifikat mit Wasser auf 100 µl aufgefüllt und mit 1/10 Vol. 3 M NaOAc (pH 4,7) und 2,5 Vol. Ethanol (96%) gefällt. Nach dem Zentrifugationsschritt wurde das Nukleinsäure-Pellet mit 70%igem Ethanol gewaschen. Nach erneuter Zentrifugation wurde das Pellet getrocknet und anschliessend in 13 µl TSR-Puffer resuspendiert. Die Identifizierung der resultierenden Sequenz erfolgte durch Anfrage beim BLAST-Server unter der Web-Adresse: www.ncbi.nlm.nih.gov/blast.

5.4.2.3 Isolierung von mRNA und gesamt-RNA

mRNA Präparation für Subtraktive Klonierungen

Für die erfolgreiche Durchführung der Subtraktiven Klonierung ist hochqualitative und absolut DNA-freie mRNA erforderlich. Diese wurde mittels des *mRNA Isolation Kits* nach Angaben des Herstellers Roche Diagnostics gewonnen.

Die Methode basiert auf der Hybridisierung eines biotinylierten (dT)₂₀-Oligonukleotids an die polyadenylierte mRNA. Durch Zugabe von Streptavidin-markierten Magnetpartikeln wird mit Hilfe eines Magneten eine Aufreinigung von mRNA möglich. Durch intensive Waschschrte kann die mRNA mit Wasser eluiert und die Konzentration durch eine UV-photometrische Messung bei 260 nm in einer Mikroküvette (Gesamt-Vol. 50 µl) bestimmt werden.

Gesamt-RNA Präparation für cDNA- und cRNA-Synthese

Für die Isolierung von gesamt-RNA wurde 1 ml Trizol (Fa. Invitrogen) auf die Petrischale, mit den zuvor mit PBS gespülten adhärenen Zellen (ca. 2-5x10⁶ Zellen), gegeben. Dieses Lysat wurde anschliessend mit 200 µl Chlorophorm versetzt, 20 s gevortext und 3 min bei RT inkubiert. Nach einem Zentrifugationsschritt von 15 min/14.000 UpM (bei 4°C) wurde die obere, wäßrige Phase abpipettiert und mit 0,5 ml Isopropanol versetzt. Nach einer Inkubationszeit von 10 min bei RT wurde die RNA 30 min bei 14.000 UpM/4°C gefällt, das Pellet mit Ethanol 70% gewaschen und anschliessend bei RT getrocknet. Resuspension der isolierten RNA erfolgte mit 100 µl RNase-freiem Wasser.

Als Ausgangsmaterial für die cDNA-Synthese wurde die Konzentration der RNA UV-photometrisch bei 260 nm vermessen und 1 µg für die Synthese mit N₆-Primern, rRNasin[®] und Superscript[™] II eingesetzt.

Für die DNA-Microchip-Hybridisierung war jedoch eine RNA von besonders hoher Reinheit und absoluter DNA-Abwesenheit erforderlich. Dazu wurden 100 µg der durch die oben genannte Trizol-Methode gewonnenen gesamt-RNA nach Angaben des Herstellers auf die *RNeasy[®] Mini Kit* Säule gegeben und auf dieser mit Hilfe des *RNase-Free DNase Sets* (beide Fa. Qiagen) von DNA befreit. Nach der anschliessenden Elution mit RNase-freiem Wasser wurde die Qualität der gesamt-RNA auf einem denaturierenden Agarosegel

und in einer Standard-PCR-Reaktion mit β -Aktin Oligonukleotiden auf DNA-Abwesenheit überprüft.

Für die DNA-Microchip-Hybridisierung wurde die aufgereinigte RNA auf eine Konzentration von 1,5 $\mu\text{g}/\mu\text{l}$ eingestellt und im Institut für Zellbiologie (Tumorforschung) des Universitätsklinikums Essen von Hr. PD Dr. Klein-Hitpass zur cRNA-Synthese und anschließender DNA-Microchip-Hybridisierung verwendet.

5.4.3 Expressionsanalyse

5.4.3.1 Subtraktive Klonierung

Die Subtraktive Klonierung stellt eine Technik dar, die eine Identifizierung von differentiell exprimierten Genen ermöglicht. Die zu vergleichenden RNA-Populationen werden in cDNA umgeschrieben und mehrfach miteinander hybridisiert. Die Transkripte, die nur in einer Fraktion vorkommen, bleiben übrig und können mit Hilfe suppressiver und *nested*-PCR amplifiziert, kloniert und anschließend sequenziert werden.

Zur Durchführung dieser Methode wurde der *PCR-Select cDNA Subtraction Kit* und für die PCR-Reaktionen der *Advantage[®] cDNA PCR Kit* (beide Fa. Clontech) nach Angaben des Herstellers verwendet.

5.4.3.2 Dot-Blot Hybridisierungsmethode

Die Dot-Blot Technik ist eine schnelle Methode zur Untersuchung einer großen Zahl von cDNA Proben auf das Vorkommen bestimmter Spezies. Die Durchführung ist wesentlich zeitsparender als Northern oder Southern Blots, da der Gelelektrophoreseschritt entfällt. Der Vorteil dieser Methode besteht in der Begrenzung der aufgetragenen Proben auf streng definierte Bereiche der Membran, die einen besseren Vergleich der Intensitäten von Hybridisierungssignalen ermöglicht.

Zur Denaturierung wurden 2 μg Plasmid-DNA mit 180 μl MilliQ-Wasser und 20 μl 1 M NaOH versetzt und 1 h bei 68°C inkubiert. Nach der anschließenden Neutralisierung mit 20 μl 1.5 M NaH_2PO_4 wurde die Probe auf die Nylonmembran aufgebracht.

Die Apparatur besteht aus verschiedenen Plexiglasblöcken, die durch Schrauben zusammen gehalten werden. Zwischen die oberen beiden Blöcke, die eine regelmäßige An-

ordnung von 2x40 runden Bohrungen besitzen, wurde die in 2x SSC equilibrierte Nylonmembran (Porablot NY plus, Fa. Macherey-Nagel) und ein angefeuchtetes 3MM-Filterpapier eingeklemmt. In jede Bohrung wurden 110 µl (1 µg) cDNA-Lösung pipettiert und, bedingt durch den Hohlraum im untersten Block, ein Vakuum (Wasserstrahlpumpe) angelegt.

Jede aufbereitete Probe wurde auf 2 Bohrungen verteilt (an die gleiche Position in der 40 Loch Anordnung), so daß daraus zwei Membranen mit einer identischen Serie von Proben resultierten, die zum Vergleich getrennt mit unterschiedlichen radioaktiv markierten Sonden hybridisiert wurden. Die Membranen wurden 2 h bei 80°C inkubiert, um die applizierten Nukleinsäuren kovalent an die Membran zu koppeln. Anschliessend wurde die Membran über Nacht in Prähybridisierungspuffer bei 42°C vorhybridisiert. Informationen zur Hybridisierung mit radioaktiv markierten Sonden siehe Kap. 5.4.3.3.

5.4.3.3 Radioaktiv markierte cDNA-Synthese und Hybridisierung

<u>LowTE:</u>	10 mM	Tris (mit HCl auf pH 8)
	0,1 M	EDTA
<u>Hybridisierungspuffer:</u>	200 ml	Formamid
	100 ml	20x SSC
	20 ml	20% SDS
	8 ml	50x Dennhardts
	72 ml	Wasser

Für die Synthese radioaktiv markierter Sonden wurden 0,8 µg mRNA oder 1 µg gesamt-RNA eingesetzt. Nach Zugabe von 200 pmol N₆-Primer und dem Auffüllen mit RNase-freiem Wasser auf ein Volumen von 7 µl wurde dieser Ansatz 10 min bei 70°C denaturiert. Im Anschluss daran wurde er 5 min auf Eis abgekühlt. Dann wurden:

2 µl	DTT
3 µl	dNTPs (3.3 mM each, ohne dCTP)
4 µl	[α ³² P] dCTP
4 µl	5x Puffer
1 µl	Superscript-II

zugegeben und insgesamt 1 h bei 42°C inkubiert. Nach 30 min wurden 2 µl dCTP (3.3 mM) zugegeben. Im Anschluss an die Inkubation wurde der Ansatz mit LowTE-Puffer auf

200 µl aufgefüllt und über eine Sephadex 50-Säule aufgereinigt. Die Aktivität der Sonden wurde nachfolgend durch eine Čerenkov-Messung (Wallac 1409, Liquid Scintillation Counter, Fa. Wallac) bestimmt. Es wurden nur Sonden mit einer Mindest-Aktivität von 10 Mio cpm eingesetzt. Nach 5 minütiger Denaturierung bei 95°C wurde die cDNA in den Hybridisierungspuffer gegeben. In dieser Lösung wurde die zu hybridisierende Membran über Nacht bei 42°C inkubiert. Das Waschen der Membranen erfolgte zuerst 20 min bei RT mit 2x SSC und anschliessend 20 min in der Regel bei 58°C mit 2x SSC/0,1% SDS. Die Darstellung der Signale erfolgte durch die Verwendung spezifischer Filme (BioMax MS-1 sowie „Intensifying Screen“ für BioMax MS-1, Fa. Kodak).

5.4.3.4 cRNA-Synthese, DNA-Microchip-Hybridisierung und Datenanalyse

In dieser Arbeit wurden die Affymetrix Chips „*MurineGenome U74Av2*“ für die Hybridisierungen verwendet, die alle bis zum Entwicklungszeitpunkt funktional charakterisierten murine Gene (~ 6000; aus der Maus-UniGene Datenbank) und weitere ~6000 ESTs enthalten. Die meisten ESTs konnten mittlerweile durch die ständig aktualisierte Datenbank identifiziert werden.

Für die cDNA-Synthese wurden 9 µl (13,5 µg) RNA mit 1 µl einer Mischung aus 3 verschiedenen polyadenylierten Kontroll-RNAs, 1 µl 100 µM T7-oligo-d(T)₂₄ Primer [5'-GGCCAGTGAATTGTAATACGACTCACTATAGGGAGGCGG-(dT₂₄)-3'] gemischt und nach einer Inkubation von 10 min bei 70°C auf Eis gehalten. Anschliessend wurden

4 µl	5x Erststrang-Synthesepuffer
2 µl	0,1 M DTT
1 µl	10 mM dNTPs

zugegeben und der Ansatz 2 min bei 42°C präinkubiert. 2 µl (200 U) Superscript II (Life Technologies) wurden addiert und die Inkubation bei 42°C für eine Stunde fortgesetzt. Für die Zweitstrangsynthese wurden:

30 µl	5x Zweitstrang-Synthesepuffer
91 µl	RNase-freies Wasser
3 µl	10 mM dNTPs
4 µl	<i>E.coli</i> DNA Polymerase I (40 U)
1 µl	<i>E.coli</i> DNA-Ligase (12 U)
1 µl	RNase H (2 U)

zugegeben und dieser Ansatz 2 h bei 16°C inkubiert. Nach der Zugabe von 2.5 µl (10 U) T4 polymerase I (TaKaRa) wurde für weitere 5 min inkubiert. Die Reaktion wurde durch das Hinzupipettieren von 10 µl 0.5 M EDTA gestoppt und die ds cDNA durch einen Phenolisierungsschritt aufgereinigt. Nach Fällung wurde die cDNA in 12 µl Wasser resuspendiert.

5 µl ds cDNA wurden für die Synthese von biotinylierter *antisense*-RNA mit Hilfe des *BioArray High Yield RNA Transcript Labeling Kit* (Fa. Enzo Diagnostics) verwendet. Die biotinylierte cRNA wurde über *RNeasy Säulen* (Fa. Qiagen) aufgereinigt. Proben-Fragmentierung, Hybridisierung mit *MG_U74Av2* MicroArrays (Fa. Affymetrix), Waschschritte, Färbung und Scannen mit dem GeneArray Scanner (Fa. Agilent) wurden nach dem Standard-Protokoll von Affymetrix (*Gene Expression Technical Manual*) durchgeführt. Zur Rohdaten-Analyse und zur Erstellung der Vergleichsdaten wurde die Affymetrix GeneChip Analysis Suite 4.0 Software verwendet.

Datenanalyse

Nach dem Einscannen der DNA-Microchips wurden zur Datenauswertung die Signalintensitäten mit Hilfe der MAS 5.0 Software (Fa. Affymetrix) ermittelt. Für die zwei zu vergleichenden Datensätze wurden nach einem Normierungsschritt folgende Filter-Kriterien angewendet: (A) *increase/decrease*, (B) *signal log₂ ratio* > 1,32 oder < -1,32 (entspricht einer 2,5-fachen Über- oder Unterexpression), (C) p-value > 0,999 oder < 1,001.

5.4.4 Proteomanalyse / 2D-Gele / MALDI-TOF MS

5.4.4.1 Isolierung von Plasma- und Kern-Proteinfraktionen

<u>RSB-Puffer + Triton</u>	10 mM	Tris-HCl pH 8,0
	10 mM	NaCl
	3 mM	MgCl ₂
	0,5 %	Triton X 100
<u>Benzonasepuffer</u>	150 mM	Tris-HCl pH 8.0
	0,2 mM	MgCl ₂

<u>Proteasehemmer-Mix (EK)</u>	1 mM	PMSF
	10 μ M	Pepstatin
	50 μ g/ml	TLCK
	100 μ g/ml	TPCK

Für die Präparation von Plasma- und Kernproteinfraktionen wurden die zuvor mit PBS gespülten adhärenen Zellen (ca. $1-3 \times 10^7$) mit einem Zellspatel abgelöst und in flüssigem Stickstoff eingefroren, sofern nicht sofort mit der Protein-Isolierung fortgefahren wurde. Ansonsten wurde das Zellpellet anschliessend in 1 ml RSB-Puffer/Triton X 100 und 2 μ l Proteasehemmer-Mix resuspendiert und für 4 min auf Eis inkubiert. Nach einem Zentrifugationsschritt (3 min bei 14.000 UpM) wurde der Überstand (Plasmafraktion) abpipettiert und erneut zentrifugiert, damit eine quantitative Abtrennung der Zellkerne gewährleistet wurde. Nach zweimaligem Waschen des Kernpellets mit RSB-Puffer (ohne Triton X 100) wurde bei 14.000 UpM für 3 min zentrifugiert.

Da Nukleinsäuren eine korrekte Auftrennung der einzelnen Proteine im SDS-Gel der 2. Dimension behindern und das Ergebnis verfälschen können, wurde stets ein RNA- und DNA-Verdau mittels Benzonase[®] nach Angaben des Herstellers durchgeführt.

Das Kernpellet wurde in 261 μ l Benzonasepuffer resuspendiert, mit 2 μ l Proteasehemmer-Mix, 30 μ l LDAO (Lauryldimethylaminoxid, 10 %) und 7 μ l Benzonase[®] versetzt und anschliessend 1 h bei 37°C inkubiert.

Die vollständige Lyse der Kerne wurde mittels Zugabe von Trypanblau kontrolliert. Dabei wurden je 5 μ l Kernsuspension und Trypanblau vermischt, auf einen Objektträger aufgebracht und unter dem Lichtmikroskop betrachtet. Intakte Zellkerne waren als blaue-färbte Strukturen erkennbar.

5.4.4.2 Bestimmung der Proteinkonzentration mittels BCA

Der BCA-Assay wurde mit dem *BCA Protein Assay Kit* (Fa. Pierce) durchgeführt und beruht auf der Reduktion von Cu^{2+} zu Cu^+ durch Proteine im alkalischen Medium (Biuret Reaktion) und der kolorimetrischen Bestimmung des Kupfer-Kations durch Chelatierung mit zwei Bicinchoninsäure-Molekülen. Der gebildete, wasserlösliche Komplex besitzt ein Absorptionsmaximum bei 562 nm, wobei jedoch auch Wellenlängen zwischen 540-590 nm zur Messung verwendet werden können.

Je 20 µl Probe bzw. 20 µl einer entsprechenden Verdünnung werden mit 200 µl frisch hergestelltem BCA-Reagenz (1:50) in je einer Vertiefung einer Microtiterplatte versetzt. Nach einer 30-minütigen Inkubation bei 37°C werden die Lösungen am ELISA-Reader MR5000 (Fa. Dynatech) bei 550 nm vermessen und die Proteinkonzentration mittels einer mitgeführten BSA-Eichgerade berechnet.

Zur Vorbereitung für 2D-gelelektrophoretische Experimente wurde das entsprechende Volumen für 275 µg Plasmaproteine und 350 µg Kernproteine wurde mit 1 Volumen TCA (10%, eiskalt) versetzt und 30 min auf Eis gefällt. Nach einem Zentrifugationsschritt von 30 min/14.000 UpM wurde mit Ethanol 96% (eiskalt) gewaschen, das Pellet anschliessend in der *speed-vac* getrocknet und bei -80°C eingefroren.

5.4.4.3 Isoelektrische Fokussierung/Erste Dimension

<u>Rehydrierungslösung</u>	7 M	Harnstoff
	2 M	Thioharnstoff
	20 mM	Tris-HCl pH 8,0
	4 %	CHAPS
	125 µl	Bromphenolblaulösung (1%) auf 25 ml Lösung
<u>Equilibrierungslösung</u>	6 M	Harnstoff
	50 mM	Tris-HCl pH 8,0
	2 %	SDS
	30 %	Glycerol
	2,5 ml	Bromphenolblaulösung (1%) auf 500 ml Lösung

Für die isoelektrische Fokussierung wurden 275 µg Zytoplasmaproteine oder 350 µg Kernproteine in 350 µl Rehydrationslösung gelöst und auf *DryStrips* (18 cm, Fa. Amersham Biosciences) mit dem Gerät IPGphor (Fa. Amersham Biosciences) mit folgendem Spannungs-Zeit-Programm fokussiert. Es wurde eine aktive Rehydrierung über 12 h bei 30 V durchgeführt, die im Gegensatz zur passiven eine höhere Reproduzierbarkeit und quantitativere Proteinaufnahme mit schon beginnender Ausrichtung der Proteine ermöglicht.

Im Anschluß daran findet die eigentliche Fokussierung bei 20°C mit einem Stromlimit von 50 µA statt. Binnen einer Stunde wird die Spannung von 30 V auf 1000 V erhöht. In der folgenden halben Stunde auf 2000 V, in der nächsten auf 3000 V, in weiteren 30 min auf 4000 V und dann auf 8000 V.

Diese wurden über 8 h gehalten, um ein Volt-Stunden-Integral von 64 kVh zu erhalten. Anschliessend wurden die Strips bei -80°C gelagert.

Zeit [h]	Spannung [V]	Modus
12	30	Step-and-hold
1	1000	Gradient
0,5	2000	Gradient
0,5	3000	Gradient
0,5	4000	Gradient
0,5	8000	Gradient
8	8000	Step-and-hold

Zur Reduktion vorhandener Disulfid-Brücken wurde eine frische Lösung aus 200 mg DTT und 20 ml Equilibrierungslösung hergestellt, mit der die aufgetauten Strips 20 min bei RT inkubiert wurden. Für die anschliessende Alkylierung der freien reaktiven SH-Gruppen wurden 20 ml Equilibrierungslösung mit 900 mg Iodacetamid (IAA) versetzt. Nach einer Inkubation der Strips für ebenfalls 20 min bei RT wurden sie zum Abspülen der IAA-Lösung kurz in Elektrodenpuffer eingetaucht und konnten nun für die zweite Dimension eingesetzt werden.

Die für die IEF und SDS-Polyacrylamid-Gelelektrophorese benötigten Reagenzien wurden von der Fa. SERVA bezogen.

5.4.4.4 SDS-Polyacrylamid-Gelelektrophorese/Zweite Dimension

<u>Sammelgel (4%)</u>	1,3 ml	Acrylamid-Stocklsg.
	2,5 ml	Sammelgelpuffer
	6,1 ml	Wasser
	0,01 ml	TEMED
	0,1 ml	APOD (10%)

<u>Trenngel</u>	<u>10%</u>	<u>15%</u>	
	8,0 ml	5,3 ml	Acrylamid-Stocklsg.
	4,1 ml	4,1 ml	Trenngelpuffer
	4,3 ml	7,0 ml	Wasser
	0,01 ml	0,01 ml	TEMED
	0,075 ml	0,075 ml	APOD (10%)
<u>Sammelgelpuffer (4x, pH 6,8)</u>		0,5 M 0,4 %	Tris-HCl SDS
<u>Trenngelpuffer (4x, pH 8,8)</u>		1,5 M 0,4 %	Tris-HCl SDS
<u>Elektrodenpuffer (10x)</u>		1 % 0,25 M 1,92 M	SDS Tris Glycin
<u>Acrylamid-Stocklösung</u>		300 g 8 g ad 100 ml	Acrylamid Bisacrylamid Wasser

Prinzip nach Lämmli

Zur Auftrennung der Proteine in der zweiten Dimension wurden in der Protean II Multi-Gel-Casting Chamber (Fa. BIO-RAD) bis zu 10 Polyacrylamid-Gele gleichzeitig mit einem linearen Gradienten von 10-15% mittels eines Gradientenmischers und einer Peristaltikpumpe gegossen. Der equilibrierte Strip wurde anschliessend vorsichtig zwischen die zwei Glasplatten auf das Gel geschoben und mit einer Lösung aus 0,5% Agarose in Elektrodenpuffer fixiert. Nach dem Auspolymerisieren der Agarose wurden 10 µl einer 1:5 verdünnten Stammlösung des Proteinmarkers *BOA protein marker* (Fa. MoBiTec), der rekombinante Proteine in der Größe von 20-150 kDa enthält, auf ein kleines Stück Filterpapier aufgetragen und neben den Strip plaziert. In dem Protean II xl System (Fa. BIO-RAD) konnten 2 Gele gleichzeitig bei 25 mA/pro Gel (Sammelgel) und bei 50 mA/pro Gel (Trenngel) laufen, wobei ein Umwälzkryostat für Kühlung der Gele sorgte.

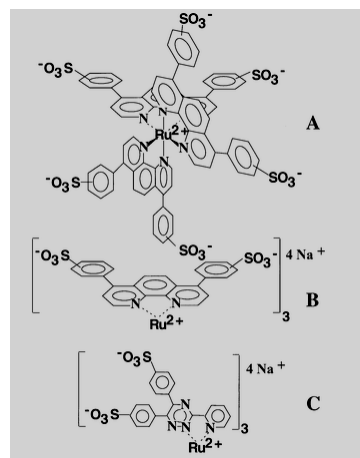
5.4.4.5 Färbemethode

<u>Fixierlösung</u>	30%	Ethanol
	10%	Essigsäure

Die Proteingele wurden in 250 ml Fixierlösung 4x30 min inkubiert und anschliessend in 20%igem Ethanol 4x30 min gewaschen. Die Färbelösung wurde nach (Rabilloud, 2000) hergestellt und in einer Konzentration von 10 µl/l Ethanol (20%) verwendet. Jedes Gel wurde in 400 ml RuBPS-Färbelösung über Nacht inkubiert und im Anschluß 2x10 min mit destilliertem Wasser gewaschen. Alle Behandlungsschritte wurden bei Raumtemperatur durchgeführt.

Die anschliessende Detektion wurde mit dem Fluoreszenz-Imager FLA-3000 (Anregungswellenlänge: $\lambda=473$ nm, Sperrfilter: $\lambda=520$ nm; 100 Pixel/mm²; Fa. Raytest) durchgeführt. Zur Datenaufnahme und Konvertierung ins Tiff-Format wurde die Software BASReader 3.01 und AIDA 2.41 (beide Fa. Raytest) verwendet.

Abb. 5.2 :Strukturformel: RuBPS
(Quelle: Rabilloud, 2000)



5.4.4.6 Detektion differentieller Protein-Spots

Da auf guten zweidimensionalen Gelen bis zu 2000 Proteinspots detektiert werden können, ist die anschliessende Suche nach differentiell vorkommenden Proteinen bei einem Vergleich mehrerer Gele nur computergestützt durchführbar. Dazu wurde die Software der Fa. BioRad (PDQuest 6.2) verwendet. Informationen zur Erstellung des Match-Sets, Definition des „Master-Gels“ und Herstellung des synthetischen Bildes werden in Kapitel 3.3 gegeben.

5.4.4.7 Proteinidentifizierung / MALDI-TOF MS

Die aus den zweidimensionalen Gelen ausgeschnittenen differentiellen Proteinspots werden im Rahmen einer anderen Dissertationsarbeit aufgearbeitet und identifiziert. Dazu gehören Trypsin-Verdau im Gelfragment, Elutionsschritte und Probenaufbereitung für die massenspektrometrische Bestimmung mittels MALDI-TOF MS (matrix-assisted laser desorption ionisation, time-of-flight Massenspektrometrie).

5.4.5 Immunohistochemische Methoden

Für die immunohistochemischen Experimente wurden von den Zell-Linien MLL^{+/+} und MLL^{-/-} je 5x10³ Zellen / 400 µl pro well in *Chamberslides* (Lab-Tec® Chambered Coverglass System, Fa. Nalge Nunc International) ausgebracht. Bei einer erreichten Konfluenz von ca. 70% wurden die Zellen zur Fixierung und Dehydratisierung mit Methanol (eiskalt) versetzt und für 8 min bei -20°C aufbewahrt. Nach Entfernung des Methanols wurden die Zellen bei RT getrocknet und konnten anschliessend bei -20°C gelagert werden. Die Rehydratisierung erfolgte durch Zugabe von PBS für 2-3 min bei RT. Nach dem Entfernen des PBS wurden die Zellen mit PBS/Triton X 100 (0,1%) 3 min permeabilisiert. Die Lösung wurde abpipettiert und die Blocking-Solution (10% Normal goat serum in 1x PBS, Fa. Sigma) in 1x PBS) bei RT für 30 min zugegeben. Nach der Entfernung dieser Lösung wurden die Zellen mit dem Erstantikörper versetzt. Der murine monoklonale Antikörper MAK E28 ist gegen Exon 28 des humanen *MLL* Gens gerichtet und wurde in unserem Labor von Dr. I. Zündorf hergestellt und von J. Zech aufgereinigt. Der monoklonale Antikörper MAK E3 erkennt Exon 3 des *MLL* Gens. Nach Inkubation bei 37°C für die Dauer von 1 h wurde die Erstantikörperlösung abpipettiert. Im Anschluss wurde 3x mit PBS/Triton X (0,1%) für jeweils 5 min gewaschen und der Zweit-Antikörper zugegeben. Hierfür wurde eine 1/250-fache Verdünnung des Antikörpers Anti-Maus-FITC (Ziege; Fa. ICN), mit PBS/1% NGS, hergestellt. Nach der Inkubation von 30 min bei RT unter Lichtausschluss wurde erneut 3x mit PBS/Triton X (0,1%) für jeweils 5 min gewaschen. Zur Equilibrierung wurden in jede Kammer 3-4 Tropfen der „Lösung C“ getropft und nach 5 min „Komponente A“ zugegeben (beide Lösungen: *SlowFade®Light Antifade Kit*, Fa. Molecular Probes). Direkt im Anschluss wurden die präparierten Zellen im CLSM untersucht.

Als Negativkontrolle wurden Zellen an Stelle des Erst-Antikörpers mit einer 1%igen Normal-goat-serum/PBS Lösung inkubiert und in allen weiteren Schritten ebenfalls wie die Proben behandelt.

5.4.5.1 Konfokales Laserscanmikroskop

Die Verwendung des konfokalen Laserscanmikroskops (CLSM, Fa. Leica Lasertechnik) bietet die Möglichkeit, 3-dimensionale Bilder aufzunehmen (Pawley 1995), die eine Aussage über intrazelluläre Lokalisation bestimmter Proteine erlauben. Die gespeicherten Informationen werden mit Hilfe der Imaris- und Selima-Software als 3D-Bild dargestellt. Die Kombination aus Argon-Krypton-Laser und den vorhandenen Filtern ermöglicht die Anregung bei drei diskreten Wellenlängen ($\lambda = 488 \text{ nm}$, 568 nm und 647 nm).

Bei dem FITC-markierten anti-Maus Zweitantikörper handelt es sich um das Fluorochrom Fluorescein, mit einer Excitation von 485 nm und einer Emission von 520 nm .

6 Anhang

6.1 Abkürzungsverzeichnis

Abb.	Abbildung
AF-4	mit ALL-1 (MLL) fusioniertes Gen auf Chromosom 4
ALL/ <i>ALL</i>	akute lymphatische Leukämie, veraltete Bezeichnung für <i>MLL</i>
AML	akute myeloische Leukämie
Amp ^R	Ampicillin-Resistenz
APOD	Ammoniumperoxodisulfat
AS	Aminosäuren
BCA	Bicinchoninsäure
bcr	breakpoint cluster region (Bruchpunkt-Cluster-Region)
<i>BioB</i>	<i>Biotin-Synthase</i> Gen (<i>E.coli</i>)
bp	Basenpaare
bzw.	beziehungsweise
cDNA	komplementäre DNA
CHAPS	3-[(3-Cholamidopropyl)dimethylammonio]-1-propan sulfonat
CLSM	Konfokales Laserscanmikroskop
Cpm	Zerfälle/min im Čerenkov-Kanal
<i>CreX</i>	<i>cre recombinase</i> Gen (Bacteriophage P1)
cRNA	komplementäre/antisense RNA
<i>DapX</i>	<i>E.coli</i> Gen
der4	Derivatchromosom 4 (AF-4•MLL)
der11	Derivatchromosom 11 (MLL•AF-4)
DHB	Dihydroxybenzoesäure
DNA	Desoxyribonukleinsäure
dNTP	2'Desoxynucleosid-5'-Triphosphat
ds	doppelsträngig
DTT	Dithiothreitol
EBNA	Epstein-Barr Virus nukleäres Antigen
ECM	extrazelluläre Matrix
<i>E. coli</i>	<i>Escherichia coli</i>
EFS	event-free-survival
EK	Endkonzentration
ETP	Enhancer of Trithorax and Polycomb
evtl.	eventuell
Fa.	Firma
FCS	fetal calf serum, fötales Kälberserum
FITC	Fluoresceinisothiocyanat
FYRC	Dimerisierungsdomäne in MLL-N
FYRN	Dimerisierungsdomäne in MLL-C
g	Gramm
G418	Neomycin
GAPDH	Glycerinaldehyd-3-phosphatdehydrogenase
gDNA	genomische DNA
h	Stunde
H3	Histon 3

HAT	Histon-Acetyltransferase
HCCA	α -Cyano-4-hydroxymethylsäure
HDAC	Histon-Deacetyltransferase
HERP	helix-loop-helix Protein, Ähnlichkeit zu HES Proteinen
<i>hes</i>	<i>hairy and enhancer of split</i> Gene
his	Histidinol
hMLL	humanes MLL Protein
<i>hox</i>	Homeobox-Gene
Hyg	Hygromycin
kb	Kilobasenpaare
kDa	Kilodalton
l	Liter
<i>LysX</i>	<i>diaminopimelate decarboxylase</i> Gen (<i>B. subtilis</i>)
M	molar
MAK E3	monoklonaler Antikörper gegen Epitop im Exon 3
MAK E28	monoklonaler Antikörper gegen Epitop im Exon 28
MALDI-TOF	Matrix-assisted-laser-desorption time-of-flight Massenspektrometrie
mcs	Multiple Klonierstelle
ME	Maintenance Element
MEF	murine embryonale Fibroblasten
mg	Milligramm
μ g	Mikrogramm
min	Minute
mind.	mindestens
ml	Milliliter
μ l	Mikroliter
<i>MLL/Mll</i>	<i>Mixed lineage leukemia</i> Gen (human/murin)
MLL MPSC	MLL Multiprotein-Superkomplex
mM	Millimol
mRNA	messenger-RNA
NCBI	National Center for Biotechnology Information
neo ^R	Neomycin-Resistenz
nmol	Nanomol
pBC230A	Expressionsvektor
pBS/KS	Vektor pBluescript
PAGE	Polyacrylamid-Gelelektrophorese
PcG _{i/m}	Polycomb-Proteingruppe (Initiationskomplex/Maintenance Komplex)
PCR	Polymerase Chain Reaction
<i>PheX</i>	<i>phe</i> Gen (<i>B. subtilis</i>)
PRE	Polycomb Response Element
q	lange Arm eines Chromosoms
RNA	Ribonukleinsäure
RNAse	Ribonuklease
UpM	Umdrehungen pro Minute
RT	Raumtemperatur
RT-PCR	Reverse Transkriptase PCR
s	Sekunde
SDS	Natriumdodecylsulfat
snLS	subnukleäre Lokalisationssequenz
t(4;11)	chromosomale Translokation (<i>MLL/AF-4</i>)

Tab.	Tabelle
TAD	Transaktivierungsdomäne
TCA	Trichloressigsäure
TEMED	Tetramethylethyldiamin
<i>ThrX</i>	<i>thr</i> Gen (<i>B. subtilis</i>)
TLCK	Trypsin-Inhibitor (1-Chloro-3-tosylamido-7-amino-2-heptanon)
TNF- β	Tumornekrose-Faktor beta
TPCK	Chymotrypsin-Inhibitor (1-Chloro-3-tosylamido-4-phenylbutanon)
TRD	Transkriptionsrepressor-Domäne
TRE	Trithorax Resoponse Element
Tris	Trishydroxymethylaminomethan
Triton X 100	t-Octylphenoxyethanol
<i>TrpnX</i>	<i>trpn</i> Gen (<i>B. subtilis</i>)
Trx	Trithorax Protein (<i>Drosophila</i> Homolog zu MLL)
TrxG	Trithorax-Proteingruppe
tTA	Transaktivator Protein
u.a.	unter anderem
UpM	Umdrehungen pro Minute
Vol.	Volumen
vs.	versus/gegen
WT	Wildtyp
z.B.	zum Beispiel

Aminosäure-Code

A	Alanin
D	Asparaginsäure
G	Glycin
K	Lysin
Q	Glutamin
V	Valin
x	beliebige AS

6.2 Primer-Liste

Die aufgeführten Primer sind in 5' \rightarrow 3' Richtung dargestellt.

Bsp.: MLL•3: Primer liegt im Exon; MLLI•3: Primer bindet an Intronsequenz

6.2.1 Allgemeine Primer

T3	ATTAACCCTCACTAAAGG
T7	AATACGACTCACTATAGG
β -Act•3	TCGAGCACGGCATCGTCAC
β -Act•5	TCATTGCCAATGGTGATGACC
GAPDH•3	CTTCACCACCATGGAGAAGG

GAPDH•5	CCTGCTTCACCACCTTCTTG
lacZ•3	CGTCGTTTTACAACGTCGTGAC
lacZ•5	CTCATCCGCCACATATCCTGA
Neo•3	ATGATTGAACAAGATGGATTGCA
tTA•3	GTTCCGCGTTACATAACTTACG
tTA•5	CCCACAGCGCTGAGTGCAT
AF-4•3	GTTGCAATGCAGCAGAAGCC
AF-4•5	ACTGTCACTGTCCTCACT
EBNA1•3	TGGTTTGGG AGCATCGTGGTC
EBNA1•5	GAGGCAAATCTACTCCATCGTC

6.2.2 Spezifische murine und humane Primer

MLL1•31	AGACTCGAGAACATGGCGCACAGCTGTGC
MLL2•5	ATTTAAGCACCAACTCCATGC
MLL•3	GGCCTCCATCTACAGAACGGA
MLL•5	CAGACATCGAATACCGTCTGC
mMLL4•3	AAGACAGTTGTCAGACAAAGC
mMLL4•5	AATCCTCATCCTCTATGAACC
mMLL4•51	GCATAAGCCAAGGTGTGTGG
mMLL5•3	CCTTACCGTGGGAAGAACG
mMLL10•5	GTTCAAAGTGCCTGCATTCTC
mMLL11•3	CTCTGTTCCAGCAGTGAGCA
mMLL11•3	CAGTCTGCACTATACTGCTAC
mMLL11•31	CTGTAAGATGCAAGACTCACG
mMLL11•32	CAGATTCACAGAGTGCCACC
mMLL11•33	TTGAGACCATAAATGTTTTAGG
mMLL11•34	CCAAGTGCTGGAATCAGACG
mMLL11•35	GGAAATGTAGCTCATTCTCG
mMLL11•36	TTTAATCACGTGTAGGTGTC
mMLL11•37	CCTCTGTATCGTCATCAAGC
mMLL11•38	CCAAATGTGAATAGAGGTCC
mMLL11•39	CACAGTCTGCACTATACTGC
mMLL11•311	TGGTTCAGTGAAGTCCAGC
mMLL11•312	ATGACACAGCTTTCATCTCC
mMLL11•313	TCTGTGAGTAAGCAGATTTGG
mMLL11•314	TTGCCTCTCGTGTGCAAGG
mMLL11•54	ACTCCCAGATAGGGACACC
mMLL11•55	TTCTGAGATCGCTATGCACG
mMLL11•56	AAGCACTTAGGTTTCATACAGG

mMLL12•3	GGACCAGCTGGAAAACCTGG
mMLL12•5	GGAAGGGCTCACAAACAGAC
mMLL17•3	CAGCAAGCAGGATGAACAGC
mMLL18•5	GAATGATCTTCACAATATCATCG
mMLL23•3	GCCTCACATCTTGCACCAGC
mMLL23•5	CAGGAAGACACAGTTCTTGG
mMLL28•3	CTGGTCTTCGGAGCATTGGC
mMLL28•5	GCCAATGCTCCGAAGACCAG
mMLL32•3	CTGAAGGAGCTCTCATTTGC
mMLL37•5	CATCAATTCGGAACATGTAGC

6.2.3 Primer zur Verifizierung der DNA-Microchip-Daten

GLY•3	GAGTACTTGGAATGTGTGAGC
GLY•5	GAGGACCAGAACTTCTTAGC
ILRA•3	TGTAACCATGTTCTGAGACC
ILRA•5	GCTTCCTGTCTAGCTCAGC
MGC•3	AAGAGCCTGCTCCCTCTGG
MGC•5	AGATCATGTAAGGAGCAAGC
SMA•3	CAGCCAGTCGCTGTCAGG
SMA•5	GACGCATGATGGCATGAGG

6.2.4 Primer zur Klonierung der Promotorkonstrukte

Amp•3	TTCACACCGCATCAGGTGGCAC
Amp•5	GCAAGCAGCAGATTACGCGCAG
GFP•3	GAAGCGCGATCACATGGTCC
GFP•5	GGTCAGCTTGCCGTAGGTG
Prom•3	CAGGTACCGTATTTAAAACCTCAAGAGTTTAAATT AGTATAGAGCC
Prom•5	GCTAGTCGACGTGGGTTTGGTGGGGTAGTTTGGTCCCAGGCACTC
LT•3	CGACGTCGACGTCAGATCCGCTAGCGCTACCG
LT•5	CGCTCTCGAGAATGCAGTGAAAAAATGCTTTATTTG

6.3 Literaturverzeichnis

- Aasland, R., Gibson, T.J. and Stewart, A.F. (1995) The PHD finger: implications for chromatin-mediated transcriptional regulation. *Trends Biochem Sci*, **20**, 56-59.
- Adler, H.T., Nallaseth, F.S., Walter, G. and Tkachuk, D.C. (1997) HRX leukemic fusion proteins form a heterocomplex with the leukemia-associated protein SET and protein phosphatase 2A. *J Biol Chem*, **272**, 28407-28414.
- Alexander, F.E., Patheal, S.L., Biondi, A., Brandalise, S., Cabrera, M.E., Chan, L.C., Chen, Z., Cimino, G., Cordoba, J.C., Gu, L.J., Hussein, H., Ishii, E., Kamel, A.M., Labra, S., Magalhaes, I.Q., Mizutani, S., Petridou, E., de Oliveira, M.P., Yuen, P., Wiemels, J.L. and Greaves, M.F. (2001) Transplacental chemical exposure and risk of infant leukemia with MLL gene fusion. *Cancer Res*, **61**, 2542-2546.
- Alward, W.L. (2000) Frequency doubling technology perimetry for the detection of glaucomatous visual field loss. *Am J Ophthalmol*, **129**, 376-378.
- Arakawa, H., Nakamura, T., Zhadanov, A.B., Fidanza, V., Yano, T., Bullrich, F., Shimizu, M., Blechman, J., Mazo, A., Canaani, E. and Croce, C.M. (1998) Identification and characterization of the ARP1 gene, a target for the human acute leukemia ALL1 gene. *Proc Natl Acad Sci U S A*, **95**, 4573-4578.
- Armstrong, S.A., Staunton, J.E., Silverman, L.B., Pieters, R., den Boer, M.L., Minden, M.D., Sallan, S.E., Lander, E.S., Golub, T.R. and Korsmeyer, S.J. (2002) MLL translocations specify a distinct gene expression profile that distinguishes a unique leukemia. *Nat Genet*, **30**, 41-47.
- Attardi, L.D., Reczek, E.E., Cosmas, C., Demicco, E.G., McCurrach, M.E., Lowe, S.W. and Jacks, T. (2000) PERP, an apoptosis-associated target of p53, is a novel member of the PMP-22/gas3 family. *Genes Dev*, **14**, 704-718.
- Ayton, P.M. and Cleary, M.L. (2001) Molecular mechanisms of leukemogenesis mediated by MLL fusion proteins. *Oncogene*, **20**, 5695-5707.
- Baskaran, K., Erfurth, F., Taborn, G., Copeland, N.G., Gilbert, D.J., Jenkins, N.A., Iannaccone, P.M. and Domer, P.H. (1997) Cloning and developmental expression of the murine homolog of the acute leukemia proto-oncogene AF4. *Oncogene*, **15**, 1967-1978.
- Bhatia, R., Munthe, H.A. and Forman, S.J. (2001) Abnormal growth factor modulation of beta1-integrin-mediated adhesion in chronic myelogenous leukaemia haematopoietic progenitors. *Br J Haematol*, **115**, 845-853.
- Bluteau, O., Jeannot, E., Bioulac-Sage, P., Marques, J.M., Blanc, J.F., Bui, H., Beaudoin, J.C., Franco, D., Balabaud, C., Laurent-Puig, P. and Zucman-Rossi, J. (2002) Bi-allelic inactivation of TCF1 in hepatic adenomas. *Nat Genet*.
- Bollen, M., Gijssbers, R., Ceulemans, H., Stalmans, W. and Stefan, C. (2000) Nucleotide pyrophosphatases/phosphodiesterases on the move. *Crit Rev Biochem Mol Biol*, **35**, 393-432.
- Bonilla, I.E., Tanabe, K. and Strittmatter, S.M. (2002) Small proline-rich repeat protein 1A is expressed by axotomized neurons and promotes axonal outgrowth. *J Neurosci*, **22**, 1303-1315.
- Bower, M., Parry, P., Gibbons, B., Amess, J., Lister, T.A., Young, B.D. and Evans, G.A. (1994) Human trithorax gene rearrangements in therapy-related acute leukaemia after etoposide treatment. *Leukemia*, **8**, 226-229.
- Brock, H.W. and van Lohuizen, M. (2001) The Polycomb group--no longer an exclusive club? *Curr Opin Genet Dev*, **11**, 175-181.
- Bukau, B. and Horwich, A.L. (1998) The Hsp70 and Hsp60 chaperone machines. *Cell*, **92**, 351-366.

- Burg, M.A., Tillet, E., Timpl, R. and Stallcup, W.B. (1996) Binding of the NG2 proteoglycan to type VI collagen and other extracellular matrix molecules. *J Biol Chem*, **271**, 26110-26116.
- Butler, L.H., Slany, R., Cui, X., Cleary, M.L. and Mason, D.Y. (1997) The HRX proto-oncogene product is widely expressed in human tissues and localizes to nuclear structures. *Blood*, **89**, 3361-3370.
- Capdevila, J., Tsukui, T., Rodriguez Esteban, C., Zappavigna, V. and Izpisua Belmonte, J.C. (1999) Control of vertebrate limb outgrowth by the proximal factor Meis2 and distal antagonism of BMPs by Gremlin. *Mol Cell*, **4**, 839-849.
- Castelli-Gair, J.E. and Garcia-Bellido, A. (1990) Interactions of Polycomb and trithorax with cis regulatory regions of Ultrabithorax during the development of *Drosophila melanogaster*. *Embo J*, **9**, 4267-4275.
- Cavodeassi, F., Modolell, J. and Gomez-Skarmeta, J.L. (2001) The Iroquois family of genes: from body building to neural patterning. *Development*, **128**, 2847-2855.
- Chen, C.S., Hilden, J.M., Frestedt, J., Domer, P.H., Moore, R., Korsmeyer, S.J. and Kersey, J.H. (1993) The chromosome 4q21 gene (AF-4/FEL) is widely expressed in normal tissues and shows breakpoint diversity in t(4;11)(q21;q23) acute leukemia. *Blood*, **82**, 1080-1085.
- Chu, P.H., Ruiz-Lozano, P., Zhou, Q., Cai, C. and Chen, J. (2000) Expression patterns of FHL/S LIM family members suggest important functional roles in skeletal muscle and cardiovascular system. *Mech Dev*, **95**, 259-265.
- Cogan, J.G., Subramanian, S.V., Polikandriotis, J.A., Kelm, R.J., Jr. and Strauch, A.R. (2002) Vascular smooth muscle alpha-actin gene transcription during myofibroblast differentiation requires Sp1/3 protein binding proximal to the MCAT enhancer. *J Biol Chem*, **277**, 36433-36442.
- Corral, J., Lavenir, I., Impey, H., Warren, A.J., Forster, A., Larson, T.A., Bell, S., McKenzie, A.N., King, G. and Rabbitts, T.H. (1996) An MLL-AF9 fusion gene made by homologous recombination causes acute leukemia in chimeric mice: a method to create fusion oncogenes. *Cell*, **85**, 853-861.
- Cortes, J., O'Brien, S., Kantarjian, H., Cork, A., Stass, S., Freireich, E.J., Keating, M., Pierce, S. and Estey, E. (1994) Abnormalities in the long arm of chromosome 11 (11q) in patients with de novo and secondary acute myelogenous leukemias and myelodysplastic syndromes. *Leukemia*, **8**, 2174-2178.
- de la Pompa, J.L., Wakeham, A., Correia, K.M., Samper, E., Brown, S., Aguilera, R.J., Nakano, T., Honjo, T., Mak, T.W., Rossant, J. and Conlon, R.A. (1997) Conservation of the Notch signalling pathway in mammalian neurogenesis. *Development*, **124**, 1139-1148.
- De Smedt, M., Reynvoet, K., Kerre, T., Taghon, T., Verhasselt, B., Vandekerckhove, B., Leclercq, G. and Plum, J. (2002) Active form of Notch imposes T cell fate in human progenitor cells. *J Immunol*, **169**, 3021-3029.
- DiMartino, J.F., Miller, T., Ayton, P.M., Landewe, T., Hess, J.L., Cleary, M.L. and Shilatifard, A. (2000) A carboxy-terminal domain of ELL is required and sufficient for immortalization of myeloid progenitors by MLL-ELL. *Blood*, **96**, 3887-3893.
- Domer, P.H., Head, D.R., Renganathan, N., Raimondi, S.C., Yang, E. and Atlas, M. (1995) Molecular analysis of 13 cases of MLL/11q23 secondary acute leukemia and identification of topoisomerase II consensus-binding sequences near the chromosomal breakpoint of a secondary leukemia with the t(4;11). *Leukemia*, **9**, 1305-1312.
- Dou, C., Lee, J., Liu, B., Liu, F., Massague, J., Xuan, S. and Lai, E. (2000) BF-1 interferes with transforming growth factor beta signaling by associating with Smad partners. *Mol Cell Biol*, **20**, 6201-6211.
- Ellis, T., Gambardella, L., Horcher, M., Tschanz, S., Capol, J., Bertram, P., Jochum, W., Barrandon, Y. and Busslinger, M. (2001) The transcriptional repressor CDP (Cutl1) is essential for epithelial cell differentiation of the lung and the hair follicle. *Genes Dev*, **15**, 2307-2319.

- Ennas, M.G., Sorio, C., Greim, R., Nieddu, M., Scarpa, A., Orlandini, S., Croce, C.M., Fey, G.H. and Marschalek, R. (1997) The human ALL-1/MLL/HRX antigen is predominantly localized in the nucleus of resting and proliferating peripheral blood mononuclear cells. *Cancer Res*, **57**, 2035-2041.
- Fidanza, V., Melotti, P., Yano, T., Nakamura, T., Bradley, A., Canaani, E., Calabretta, B. and Croce, C.M. (1996) Double knockout of the ALL-1 gene blocks hematopoietic differentiation in vitro. *Cancer Res*, **56**, 1179-1183.
- Flaumenhaft, R., Abe, M., Sato, Y., Miyazono, K., Harpel, J., Heldin, C.H. and Rifkin, D.B. (1993) Role of the latent TGF-beta binding protein in the activation of latent TGF-beta by co-cultures of endothelial and smooth muscle cells. *J Cell Biol*, **120**, 995-1002.
- Flynn, D.C., Leu, T.H., Reynolds, A.B. and Parsons, J.T. (1993) Identification and sequence analysis of cDNAs encoding a 110-kilodalton actin filament-associated pp60src substrate. *Mol Cell Biol*, **13**, 7892-7900.
- Flynn, D.C., Schaller, M.D. and Parsons, J.T. (1992) Tyrosine phosphorylation of a 120,000 dalton membrane-associated protein by the neural form of pp60c-src, pp60c-src+. *Oncogene*, **7**, 579-583.
- Gibson, T.J., Hyvonen, M., Musacchio, A., Saraste, M. and Birney, E. (1994) PH domain: the first anniversary. *Trends Biochem Sci*, **19**, 349-353.
- Gillert, E., Leis, T., Repp, R., Reichel, M., Hosch, A., Breitenlohner, I., Angermuller, S., Borkhardt, A., Harbott, J., Lampert, F., Griesinger, F., Greil, J., Fey, G.H. and Marschalek, R. (1999) A DNA damage repair mechanism is involved in the origin of chromosomal translocations t(4;11) in primary leukemic cells. *Oncogene*, **18**, 4663-4671.
- Gong, X.Q. and Li, L. (2002) Dermo-1, a multifunctional basic helix-loop-helix protein, represses MyoD transactivation via the HLH domain, MEF2 interaction, and chromatin deacetylation. *J Biol Chem*, **277**, 12310-12317.
- Gopalakrishnan, R., Ouyang, H., Somerman, M.J., McCauley, L.K. and Franceschi, R.T. (2001) Matrix gamma-carboxyglutamic acid protein is a key regulator of PTH-mediated inhibition of mineralization in MC3T3-E1 osteoblast-like cells. *Endocrinology*, **142**, 4379-4388.
- Greil, J., Gramatzki, M., Burger, R., Marschalek, R., Peltner, M., Trautmann, U., Hansen-Hagge, T.E., Bartram, C.R., Fey, G.H., Stehr, K. and et al. (1994) The acute lymphoblastic leukaemia cell line SEM with t(4;11) chromosomal rearrangement is biphenotypic and responsive to interleukin-7. *Br J Haematol*, **86**, 275-283.
- Gu, Y., Alder, H., Nakamura, T., Schichman, S.A., Prasad, R., Canaani, O., Saito, H., Croce, C.M. and Canaani, E. (1994) Sequence analysis of the breakpoint cluster region in the ALL-1 gene involved in acute leukemia. *Cancer Res*, **54**, 2326-2330.
- Gu, Y., Nakamura, T., Alder, H., Prasad, R., Canaani, O., Cimino, G., Croce, C.M. and Canaani, E. (1992) The t(4;11) chromosome translocation of human acute leukemias fuses the ALL-1 gene, related to *Drosophila trithorax*, to the AF-4 gene. *Cell*, **71**, 701-708.
- Hanashima, C., Shen, L., Li, S.C. and Lai, E. (2002) Brain factor-1 controls the proliferation and differentiation of neocortical progenitor cells through independent mechanisms. *J Neurosci*, **22**, 6526-6536.
- Hanson, R.D., Hess, J.L., Yu, B.D., Ernst, P., van Lohuizen, M., Berns, A., van der Lugt, N.M., Shashikant, C.S., Ruddle, F.H., Seto, M. and Korsmeyer, S.J. (1999) Mammalian Trithorax and polycomb-group homologues are antagonistic regulators of homeotic development. *Proc Natl Acad Sci U S A*, **96**, 14372-14377.
- Henkel, G.W., McKercher, S.R. and Maki, R.A. (2002) Identification of three genes up-regulated in PU.1 rescued monocytic precursor cells. *Int Immunol*, **14**, 723-732.

- Hosokawa, Y., Ueyama, E., Morikawa, Y., Maeda, Y., Seto, M. and Senba, E. (1999) Molecular cloning of a cDNA encoding mouse A15, a member of the transmembrane 4 superfamily, and its preferential expression in brain neurons. *Neurosci Res*, **35**, 281-290.
- Hsieh, J.J., Ernst, P., Erdjument-Bromage, H., Tempst, P. and Korsmeyer, S.J. (2003) Proteolytic Cleavage of MLL Generates a Complex of N- and C-Terminal Fragments That Confers Protein Stability and Subnuclear Localization. *Mol Cell Biol*, **23**, 186-194.
- Iida, K., Koseki, H., Kakinuma, H., Kato, N., Mizutani-Koseki, Y., Ohuchi, H., Yoshioka, H., Noji, S., Kawamura, K., Kataoka, Y., Ueno, F., Taniguchi, M., Yoshida, N., Sugiyama, T. and Miura, N. (1997) Essential roles of the winged helix transcription factor MFH-1 in aortic arch patterning and skeletogenesis. *Development*, **124**, 4627-4638.
- Irikura, V.M., Lagraoui, M. and Hirsh, D. (2002) The epistatic interrelationships of IL-1, IL-1 receptor antagonist, and the type I IL-1 receptor. *J Immunol*, **169**, 393-398.
- Izon, D.J., Aster, J.C., He, Y., Weng, A., Karnell, F.G., Patriub, V., Xu, L., Bakkour, S., Rodriguez, C., Allman, D. and Pear, W.S. (2002) Deltex1 redirects lymphoid progenitors to the B cell lineage by antagonizing Notch1. *Immunity*, **16**, 231-243.
- Izpisua-Belmonte, J.C., Dolle, P., Renucci, A., Zappavigna, V., Falkenstein, H. and Duboule, D. (1990) Primary structure and embryonic expression pattern of the mouse Hox-4.3 homeobox gene. *Development*, **110**, 733-745.
- Jacobs, J.J., Scheijen, B., Voncken, J.W., Kieboom, K., Berns, A. and van Lohuizen, M. (1999) Bmi-1 collaborates with c-Myc in tumorigenesis by inhibiting c-Myc-induced apoptosis via INK4a/ARF. *Genes Dev*, **13**, 2678-2690.
- Kim, N.K., Choi, M.J., Chung, H.M., Sohn, T.J., Hwang, S.G., Oh, D., Lee, H.H., Lee, Y.H. and Ko, J.J. (2001) Increased expression and localization of a serine protease inhibitor, protease nexin-1 (PN-1), in the ovary and uterus during implantation in rat. *Thromb Res*, **103**, 135-142.
- Koutsourakis, M., Langeveld, A., Patient, R., Beddington, R. and Grosveld, F. (1999) The transcription factor GATA6 is essential for early extraembryonic development. *Development*, **126**, 723-732.
- Lamandin, C., Sagot, C., Roumier, C., Lepelley, P., De Botton, S., Cosson, A., Fenaux, P. and Preudhomme, C. (2002) Are PU.1 mutations frequent genetic events in acute myeloid leukemia (AML)? *Blood*, **100**, 4680-4681.
- Lavau, C., Szilvassy, S.J., Slany, R. and Cleary, M.L. (1997) Immortalization and leukemic transformation of a myelomonocytic precursor by retrovirally transduced HRX-ENL. *Embo J*, **16**, 4226-4237.
- Le Mouellic, H., Lallemand, Y. and Brulet, P. (1992) Homeosis in the mouse induced by a null mutation in the Hox-3.1 gene. *Cell*, **69**, 251-264.
- Lee, K.K., Leung, A.K., Tang, M.K., Cai, D.Q., Schneider, C., Brancolini, C. and Chow, P.H. (2001) Functions of the growth arrest specific 1 gene in the development of the mouse embryo. *Dev Biol*, **234**, 188-203.
- Lelias, J.M., Adra, C.N., Wulf, G.M., Guillemot, J.C., Khagad, M., Caput, D. and Lim, B. (1993) cDNA cloning of a human mRNA preferentially expressed in hematopoietic cells and with homology to a GDP-dissociation inhibitor for the rho GTP-binding proteins. *Proc Natl Acad Sci U S A*, **90**, 1479-1483.
- Li, Q., Frestedt, J.L. and Kersey, J.H. (1998) AF4 encodes a ubiquitous protein that in both native and MLL-AF4 fusion types localizes to subnuclear compartments. *Blood*, **92**, 3841-3847.
- Magli, M.C., Largman, C. and Lawrence, H.J. (1997) Effects of HOX homeobox genes in blood cell differentiation. *J Cell Physiol*, **173**, 168-177.

- Marschalek, R., Greil, J., Lochner, K., Nilson, I., Siegler, G., Zweckbronner, I., Beck, J.D. and Fey, G.H. (1995) Molecular analysis of the chromosomal breakpoint and fusion transcripts in the acute lymphoblastic SEM cell line with chromosomal translocation t(4;11). *Br J Haematol*, **90**, 308-320.
- Martin-Soudant, N., Drachman, J.G., Kaushansky, K. and Nepveu, A. (2000) CDP/Cut DNA binding activity is down-modulated in granulocytes, macrophages and erythrocytes but remains elevated in differentiating megakaryocytes. *Leukemia*, **14**, 863-873.
- Meins, M., Piosik, P., Schaeren-Wiemers, N., Franzoni, S., Troncoso, E., Kiss, J.Z., Brosamle, C., Schwab, M.E., Molnar, Z. and Monard, D. (2001) Progressive neuronal and motor dysfunction in mice over-expressing the serine protease inhibitor protease nexin-1 in postmitotic neurons. *J Neurosci*, **21**, 8830-8841.
- Mendel, D.B., Khavari, P.A., Conley, P.B., Graves, M.K., Hansen, L.P., Admon, A. and Crabtree, G.R. (1991) Characterization of a cofactor that regulates dimerization of a mammalian homeodomain protein. *Science*, **254**, 1762-1767.
- Meng, Y., Zhang, Y., Tregoubov, V., Janus, C., Cruz, L., Jackson, M., Lu, W.Y., MacDonald, J.F., Wang, J.Y., Falls, D.L. and Jia, Z. (2002) Abnormal spine morphology and enhanced LTP in LIMK-1 knockout mice. *Neuron*, **35**, 121-133.
- Miernyk, J.A. and Hayman, T.G. (1996) ATPase activity and molecular chaperone function of the stress70 proteins. *Plant Physiol*, **110**, 419-424.
- Milne, T.A., Briggs, S.D., Brock, H.W., Martin, M.E., Gibbs, D., Allis, C.D. and Hess, J.L. (2002) MLL targets SET domain methyltransferase activity to Hox gene promoters. *Mol Cell*, **10**, 1107-1117.
- Morimoto, R.I. (1993) Cells in stress: transcriptional activation of heat shock genes. *Science*, **259**, 1409-1410.
- Morrisey, E.E., Ip, H.S., Lu, M.M. and Parmacek, M.S. (1996) GATA-6: a zinc finger transcription factor that is expressed in multiple cell lineages derived from lateral mesoderm. *Dev Biol*, **177**, 309-322.
- Morrisey, E.E., Tang, Z., Sigrist, K., Lu, M.M., Jiang, F., Ip, H.S. and Parmacek, M.S. (1998) GATA6 regulates HNF4 and is required for differentiation of visceral endoderm in the mouse embryo. *Genes Dev*, **12**, 3579-3590.
- Mozer, B.A. and Dawid, I.B. (1989) Cloning and molecular characterization of the trithorax locus of *Drosophila melanogaster*. *Proc Natl Acad Sci U S A*, **86**, 3738-3742.
- Musacchio, A., Gibson, T., Rice, P., Thompson, J. and Saraste, M. (1993) The PH domain: a common piece in the structural patchwork of signalling proteins. *Trends Biochem Sci*, **18**, 343-348.
- Nakamura, T., Mori, T., Tada, S., Krajewski, W., Rozovskaia, T., Wassell, R., Dubois, G., Mazo, A., Croce, C.M. and Canaani, E. (2002) ALL-1 is a histone methyltransferase that assembles a supercomplex of proteins involved in transcriptional regulation. *Mol Cell*, **10**, 1119-1128.
- Nepveu, A. (2001) Role of the multifunctional CDP/Cut/Cux homeodomain transcription factor in regulating differentiation, cell growth and development. *Gene*, **270**, 1-15.
- Netter, S., Fauvarque, M.O., Diez del Corral, R., Dura, J.M. and Coen, D. (1998) white⁺ transgene insertions presenting a dorsal/ventral pattern define a single cluster of homeobox genes that is silenced by the polycomb-group proteins in *Drosophila melanogaster*. *Genetics*, **149**, 257-275.
- Nilson, I., Lochner, K., Siegler, G., Greil, J., Beck, J.D., Fey, G.H. and Marschalek, R. (1996) Exon/intron structure of the human ALL-1 (MLL) gene involved in translocations to chromosomal region 11q23 and acute leukaemias. *Br J Haematol*, **93**, 966-972.
- Nilson, I., Reichel, M., Ennas, M.G., Greim, R., Knorr, C., Siegler, G., Greil, J., Fey, G.H. and Marschalek, R. (1997) Exon/intron structure of the human AF-4 gene, a member of the AF-4/LAF-4/FMR-2

- gene family coding for a nuclear protein with structural alterations in acute leukaemia. *Br J Haematol*, **98**, 157-169.
- Nishioka, K., Chuikov, S., Sarma, K., Erdjument-Bromage, H., Allis, C.D., Tempst, P. and Reinberg, D. (2002) Set9, a novel histone H3 methyltransferase that facilitates transcription by precluding histone tail modifications required for heterochromatin formation. *Genes Dev*, **16**, 479-489.
- Nover, L. (1991) HSFs and HSPs--a stressful program on transcription factors and chaperones. Stress Proteins and the Heat Shock Response, sponsored by Cold Spring Harbor Laboratory, Cold Spring Harbor, NY USA, April 29-May 2, 1991. *New Biol*, **3**, 855-859.
- Oosterwegel, M., Timmerman, J., Leiden, J. and Clevers, H. (1992) Expression of GATA-3 during lymphocyte differentiation and mouse embryogenesis. *Dev Immunol*, **3**, 1-11.
- Pasqualetti, M., Neun, R., Davenne, M. and Rijli, F.M. (2001) Retinoic acid rescues inner ear defects in Hoxa1 deficient mice. *Nat Genet*, **29**, 34-39.
- Pechan, P.M. (1991) Heat shock proteins and cell proliferation. *FEBS Lett*, **280**, 1-4.
- Perris, R., Kuo, H.J., Glanville, R.W., Leibold, S. and Bronner-Fraser, M. (1993) Neural crest cell interaction with type VI collagen is mediated by multiple cooperative binding sites within triple-helix and globular domains. *Exp Cell Res*, **209**, 103-117.
- Prasad, R., Yano, T., Sorio, C., Nakamura, T., Rallapalli, R., Gu, Y., Leshkowitz, D., Croce, C.M. and Canaani, E. (1995) Domains with transcriptional regulatory activity within the ALL1 and AF4 proteins involved in acute leukemia. *Proc Natl Acad Sci U S A*, **92**, 12160-12164.
- Qian, X. and Tuszynski, G.P. (1996) Expression of thrombospondin-1 in cancer: a role in tumor progression. *Proc Soc Exp Biol Med*, **212**, 199-207.
- Qian, Y., Baisden, J.M., Westin, E.H., Guappone, A.C., Koay, T.C. and Flynn, D.C. (1998) Src can regulate carboxy terminal interactions with AFAP-110, which influence self-association, cell localization and actin filament integrity. *Oncogene*, **16**, 2185-2195.
- Rabilloud, T. (2000) Detecting proteins separated by 2-D gel electrophoresis. *Anal Chem*, **72**, 48A-55A.
- Reichel, M., Gillert, E., Nilson, I., Siegler, G., Greil, J., Fey, G.H. and Marschalek, R. (1998) Fine structure of translocation breakpoints in leukemic blasts with chromosomal translocation t(4;11): the DNA damage-repair model of translocation. *Oncogene*, **17**, 3035-3044.
- Reynolds, A.B., Roesel, D.J., Kanner, S.B. and Parsons, J.T. (1989) Transformation-specific tyrosine phosphorylation of a novel cellular protein in chicken cells expressing oncogenic variants of the avian cellular src gene. *Mol Cell Biol*, **9**, 629-638.
- Rosenberg, G.A. (2002) Matrix metalloproteinases and neuroinflammation in multiple sclerosis. *Neuroscientist*, **8**, 586-595.
- Rozenblatt-Rosen, O., Rozovskaia, T., Burakov, D., Sedkov, Y., Tillib, S., Blechman, J., Nakamura, T., Croce, C.M., Mazo, A. and Canaani, E. (1998) The C-terminal SET domains of ALL-1 and TRITHORAX interact with the INI1 and SNR1 proteins, components of the SWI/SNF complex. *Proc Natl Acad Sci U S A*, **95**, 4152-4157.
- Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) *Molecular cloning: A laboratory manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
- Santamaria, I., Velasco, G., Pendas, A.M., Fueyo, A. and Lopez-Otin, C. (1998) Cathepsin Z, a novel human cysteine proteinase with a short propeptide domain and a unique chromosomal location. *J Biol Chem*, **273**, 16816-16823.

- Sbaa-Ketata, E., Vasse, M., Lenormand, B., Schneider, P., Soria, C. and Vannier, J.P. (2001) Fibronectin increases the migration induced by stromal cell-derived factor-1 alpha (SDF-1 alpha) in pre-B acute lymphoblastic leukemia cells. *Eur Cytokine Netw*, **12**, 223-230.
- Schindler, U., Beckmann, H. and Cashmore, A.R. (1993) HAT3.1, a novel Arabidopsis homeodomain protein containing a conserved cysteine-rich region. *Plant J*, **4**, 137-150.
- Seeler, J.S. and Dejean, A. (2001) SUMO: of branched proteins and nuclear bodies. *Oncogene*, **20**, 7243-7249.
- Seykora, J.T., Mei, L., Dotto, G.P. and Stein, P.L. (2002) 'Srcasm: a novel Src activating and signaling molecule. *J Biol Chem*, **277**, 2812-2822.
- Shen, M., Kawamoto, T., Yan, W., Nakamasu, K., Tamagami, M., Koyano, Y., Noshiro, M. and Kato, Y. (1997) Molecular characterization of the novel basic helix-loop-helix protein DEC1 expressed in differentiated human embryo chondrocytes. *Biochem Biophys Res Commun*, **236**, 294-298.
- Shevchenko, A., Wilm, M., Vorm, O. and Mann, M. (1996) Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. *Anal Chem*, **68**, 850-858.
- Skonier, J., Neubauer, M., Madisen, L., Bennett, K., Plowman, G.D. and Purchio, A.F. (1992) cDNA cloning and sequence analysis of beta ig-h3, a novel gene induced in a human adenocarcinoma cell line after treatment with transforming growth factor-beta. *DNA Cell Biol*, **11**, 511-522.
- Slany, R.K., Lavau, C. and Cleary, M.L. (1998) The oncogenic capacity of HRX-ENL requires the transcriptional transactivation activity of ENL and the DNA binding motifs of HRX. *Mol Cell Biol*, **18**, 122-129.
- Spandorfer, S.D., Neuer, A., Liu, H.C., Bivis, L., Clarke, R., Veeck, L., Witkin, S.S. and Rosenwaks, Z. (2000) Interleukin-1 levels in the supernatant of conditioned media of embryos grown in autologous endometrial coculture: correlation with outcome after in vitro fertilization. *Am J Reprod Immunol*, **43**, 6-11.
- Suzuki, T., J, K.T., Ajima, R., Nakamura, T., Yoshida, Y. and Yamamoto, T. (2002) Phosphorylation of three regulatory serines of Tob by Erk1 and Erk2 is required for Ras-mediated cell proliferation and transformation. *Genes Dev*, **16**, 1356-1370.
- Taipale, J., Miyazono, K., Heldin, C.H. and Keski-Oja, J. (1994) Latent transforming growth factor-beta 1 associates to fibroblast extracellular matrix via latent TGF-beta binding protein. *J Cell Biol*, **124**, 171-181.
- Takagi, S., Fujikawa, K., Imai, T., Fukuhara, N., Fukudome, K., Minegishi, M., Tsuchiya, S., Konno, T., Hinuma, Y. and Yoshie, O. (1995) Identification of a highly specific surface marker of T-cell acute lymphoblastic leukemia and neuroblastoma as a new member of the transmembrane 4 superfamily. *Int J Cancer*, **61**, 706-715.
- Taniguchi, Y., Furukawa, T., Tun, T., Han, H. and Honjo, T. (1998) LIM protein KyoT2 negatively regulates transcription by association with the RBP-J DNA-binding protein. *Mol Cell Biol*, **18**, 644-654.
- Tillib, S., Petruk, S., Sedkov, Y., Kuzin, A., Fujioka, M., Goto, T. and Mazo, A. (1999) Trithorax- and Polycomb-group response elements within an Ultrabithorax transcription maintenance unit consist of closely situated but separable sequences. *Mol Cell Biol*, **19**, 5189-5202.
- Tkachuk, D.C., Kohler, S. and Cleary, M.L. (1992) Involvement of a homolog of Drosophila trithorax by 11q23 chromosomal translocations in acute leukemias. *Cell*, **71**, 691-700.
- Toft, D.J., Rosenberg, S.B., Bergers, G., Volpert, O. and Linzer, D.I. (2001) Reactivation of proliferin gene expression is associated with increased angiogenesis in a cell culture model of fibrosarcoma tumor progression. *Proc Natl Acad Sci U S A*, **98**, 13055-13059.

- van Lohuizen, M. (1999) The trithorax-group and polycomb-group chromatin modifiers: implications for disease. *Curr Opin Genet Dev*, **9**, 355-361.
- van Lohuizen, M., Frasch, M., Wientjens, E. and Berns, A. (1991) Sequence similarity between the mammalian bmi-1 proto-oncogene and the Drosophila regulatory genes Psc and Su(z)2. *Nature*, **353**, 353-355.
- Voermans, C., van Heese, W.P., de Jong, I., Gerritsen, W.R. and van Der Schoot, C.E. (2002) Migratory behavior of leukemic cells from acute myeloid leukemia patients. *Leukemia*, **16**, 650-657.
- Wei, L.C., Shi, M., Chen, L.W., Cao, R., Zhang, P. and Chan, Y.S. (2002) Nestin-containing cells express glial fibrillary acidic protein in the proliferative regions of central nervous system of postnatal developing and adult mice. *Brain Res Dev Brain Res*, **139**, 9-17.
- Wilson, K.E., Bartlett, J.M., Miller, E.P., Smyth, J.F., Mullen, P., Miller, W.R. and Langdon, S.P. (1999) Regulation and function of the extracellular matrix protein tenascin-C in ovarian cancer cell lines. *Br J Cancer*, **80**, 685-692.
- Yagami, K., Suh, J.Y., Enomoto-Iwamoto, M., Koyama, E., Abrams, W.R., Shapiro, I.M., Pacifici, M. and Iwamoto, M. (1999) Matrix GLA protein is a developmental regulator of chondrocyte mineralization and, when constitutively expressed, blocks endochondral and intramembranous ossification in the limb. *J Cell Biol*, **147**, 1097-1108.
- Yang, N., Higuchi, O., Ohashi, K., Nagata, K., Wada, A., Kangawa, K., Nishida, E. and Mizuno, K. (1998) Cofilin phosphorylation by LIM-kinase 1 and its role in Rac-mediated actin reorganization. *Nature*, **393**, 809-812.
- Yano, T., Nakamura, T., Blechman, J., Sorio, C., Dang, C.V., Geiger, B. and Canaani, E. (1997) Nuclear punctate distribution of ALL-1 is conferred by distinct elements at the N terminus of the protein. *Proc Natl Acad Sci U S A*, **94**, 7286-7291.
- Yokoyama, A., Kitabayashi, I., Ayton, P.M., Cleary, M.L. and Ohki, M. (2002) Leukemia proto-oncoprotein MLL is proteolytically processed into 2 fragments with opposite transcriptional properties. *Blood*, **100**, 3710-3718.
- Yoneda-Kato, N., Look, A.T., Kirstein, M.N., Valentine, M.B., Raimondi, S.C., Cohen, K.J., Carroll, A.J. and Morris, S.W. (1996) The t(3;5)(q25.1;q34) of myelodysplastic syndrome and acute myeloid leukemia produces a novel fusion gene, NPM-MLF1. *Oncogene*, **12**, 265-275.
- Yoshida, Y., Hosoda, E., Nakamura, T. and Yamamoto, T. (2001) Association of ANA, a member of the antiproliferative Tob family proteins, with a Caf1 component of the CCR4 transcriptional regulatory complex. *Jpn J Cancer Res*, **92**, 592-596.
- Yu, B.D., Hanson, R.D., Hess, J.L., Horning, S.E. and Korsmeyer, S.J. (1998) MLL, a mammalian trithorax-group gene, functions as a transcriptional maintenance factor in morphogenesis. *Proc Natl Acad Sci U S A*, **95**, 10632-10636.
- Yu, B.D., Hess, J.L., Horning, S.E., Brown, G.A. and Korsmeyer, S.J. (1995) Altered Hox expression and segmental identity in Mll-mutant mice. *Nature*, **378**, 505-508.
- Zeleznik-Le, N.J., Harden, A.M. and Rowley, J.D. (1994) 11q23 translocations split the "AT-hook" cruciform DNA-binding region and the transcriptional repression domain from the activation domain of the mixed-lineage leukemia (MLL) gene. *Proc Natl Acad Sci U S A*, **91**, 10610-10614.
- Ziemin-van der Poel, S., McCabe, N.R., Gill, H.J., Espinosa, R., III, Patel, Y., Harden, A., Rubinelli, P., Smith, S.D., LeBeau, M.M., Rowley, J.D. and et al. (1991) Identification of a gene, MLL, that spans the breakpoint in 11q23 translocations associated with human leukemias. *Proc Natl Acad Sci U S A*, **88**, 10735-10739.

Genes overexpressed 2^{|SigLogRatio|}-fold in MLL^{+/+} (cells compared to MLL^{-/-} cells)

SigLogRatio	p-value	SigLogRatio	p-value	average	GAC#	Descriptions
9,23	0	6,61	0	7,92	X16009	Mouse mrp/plf3 gene for mitogen regulated protein/proliferin (MRP/PLF), exon 1
6,34	0	6,56	0	6,45	D00613	Matrix gamma-carboxyglutamate (gla) protein
9,1	0	3,26	0	6,18	L32838	Mouse germline interleukin 1 receptor antagonist (IL-1rn) gene
8,85	0	2,53	0	5,69	BC003264	ectonucleoside pyrophosphatase 2
5,92	0	4,99	0	5,455	AF064749	Mus musculus type VI collagen alpha 3 subunit mRNA
4,89	0	6,01	0	5,45	K03235	Proliferin
7,22	0	3,43	0	5,325	BC024354	cofactor of hepatocyte nucl. Factor 1 alpha
5,81	0,00001	4,7	0,000019	5,255	L00653	Mast cell protease 7
4,83	0,000001	5,58	0	5,205	AF045887	Angiotensinogen
5,45	0,000012	4,43	0	4,94	AF045801	Mus musculus gremlin mRNA
6,87	0,000001	2,75	0,000001	4,81	AA673487	Mus musculus cDNA, 3 end
5,25	0,000004	4,23	0,000068	4,74	U70132	Paired-like homeodomain transcription factor 2
4	0,000333	4,78	0,000649	4,39	AB008174	Mus musculus HNF1beta mRNA for hepatic nuclear factor 1-beta long form
5,45	0,000005	2,7	0,000001	4,075	AW214136	Mus musculus cDNA, 3 end
3,03	0	5,04	0	4,035	L06047	Glutathione-S-transferase, alpha 1 (Ya)
4,62	0	3,31	0	3,965	L19932	Transforming growth factor, beta induced, 68 kDa
3,81	0,000031	3,81	0	3,81	X70392	Inter-alpha trypsin inhibitor, heavy chain 2
3,97	0,00057	3,27	0,000025	3,62	M89797	Wingless-related MMTV integration site 4
2,83	0	4,37	0	3,6	Z18272	Procollagen, type VI, alpha 2
4,4	0,000007	2,62	0,000001	3,51	AF081789	Mus musculus cell surface antigen AA4 (AA4) mRNA
4,71	0,000009	2,07	0	3,39	M90388	Mouse protein tyrosine phosphatase (70zpep) mRNA
3,75	0,000003	2,79	0,000001	3,27	XM_124758	Monoglyceride Lipase (Mgll)
4,62	0,000012	1,79	0,000002	3,205	X14432	Thrombomodulin
4,45	0	1,95	0	3,2	L07918	Mus musculus GDP-dissociation inhibitor mRNA, preferentially expressed in hematopoietic cells
3,63	0,00004	2,76	0,000739	3,195	U09968	Mus musculus cyclin-dependent kinase inhibitor p27kip1 mRNA, complete cds
2,6	0	3,46	0	3,03	NM_009933	Procollagen, type VI, alpha 1
2,72	0,000649	3,28	0,000011	3	AF033034	Mus musculus aldehyde dehydrogenase 3 (aldh3) gene
2,59	0,000001	3,35	0,000001	2,97	X66405	Procollagen, type VI, alpha 1
3,1	0,00005	2,37	0,000789	2,735	AI843571	Mus musculus cDNA, 3 end
3,13	0	2,32	0	2,725	AF057156	Mus musculus small proline-rich protein 1A (Sprr1a) gene
2,53	0,000002	2,81	0,000079	2,67	U06119	Cathepsin H
3,44	0,000236	1,66	0,000099	2,55	X74040	M.musculus mRNA for mesenchyme fork head-1 protein
3,13	0,000011	1,95	0	2,54	L07803	Thrombospondin 2
2,95	0,000011	2,09	0,00004	2,52	AB009377	Mus musculus mRNA for CAD
3,67	0	1,34	0,000004	2,505	L12367	Mus musculus adenylyl cyclase-associated protein (CAP) mRNA
3,28	0,000003	1,7	0,000043	2,49	AV251191	Mus musculus cDNA, 3 end
2,36	0	2,6	0	2,48	U36384	Mus musculus twist-related bHLH protein Dermo-1 mRNA
3,4	0,000003	1,47	0,000001	2,435	X07439	Mouse mRNA for Hox-3.1 homeo gene
2,21	0,000001	2,6	0,000001	2,405	U49513	Small inducible cytokine A9
1,62	0	3,05	0	2,335	X65128	Growth arrest specific 1
2,35	0,000007	1,99	0,000001	2,17	D13695	Lymphocyte antigen 84
2,66	0,000001	1,66	0,000001	2,16	X97227	M.musculus mRNA for cell surface glycoprotein CD53
1,37	0,000023	2,78	0	2,075	XM_137955	similar to apolipoprotein B-100 precursor
1,93	0,000177	2,17	0,000074	2,05	U51335	GATA-binding protein 6
1,55	0,000011	2,44	0,000007	1,995	AI787317	Mus musculus cDNA, 3 end
2,31	0,000841	1,6	0,000007	1,955	Y12650	M.musculus gene encoding hereditary haemochromatosis-like protein
2,23	0	1,61	0	1,92	NM_030704	Cryac heat shock protein 20-like protein

2,12	0,000133	1,63	0,000003	1,875	AI837786	Mus musculus cDNA, 3 end
2,05	0,000005	1,5	0,000001	1,775	AF016482	Mus musculus rho7 (rho7) mRNA
1,64	0,000037	1,88	0	1,76	U15159	LIM-domain containing, protein kinase
1,81	0,000015	1,69	0,000002	1,75	AF064748	Mus musculus S3-12 mRNA
1,92	0,000015	1,56	0,000002	1,74	U83509	Mus musculus angiotensin-1 mRNA
1,49	0,000001	1,97	0,000001	1,73	XM_130285	pyruvate dehydrogenase kinase, isoenzyme 1
1,48	0,000107	1,9	0,000007	1,69	X04367	Platelet derived growth factor receptor, beta polypeptide
1,77	0,000007	1,49	0,00005	1,63	M81591	Membrane metallo endopeptidase
1,55	0,000001	1,39	0,000001	1,47	U07159	Acetyl coenzyme A dehydrogenase, medium chain
1,55	0	1,37	0	1,46	AA691492	Mus musculus cDNA, 5 end
1,35	0,000001	1,49	0	1,42	XM_130285	pyruvate dehydrogenase kinase, isoenzyme 1
1,33	0,000001	1,48	0,000002	1,405	AW121600	Mus musculus cDNA, 3 end
1,45	0,000019	1,36	0,000029	1,405	X99572	C-fos induced growth factor
1,41	0	1,39	0,000001	1,4	Z31362	M.musculus (Balb/C) Tx01 mRNA

Genes overexpressed 2^{|SigLogRatio|}-fold in MLL^{-/-} (cells compared to MLL^{+/+} cells)

SigLogRatio	p-value	SigLogRatio	p-value	average	GAC#	Descriptions
-8,57	1	-9,11	1	-8,84	Al854029	p53 apoptosis effector related to Pmp22 (Perp-pending)
-7,86	0,999991	-7,97	1	-7,915	X13297	Mouse mRNA for vascular smooth muscle alpha-actin
-7,49	1	-6,22	0,999999	-6,855	S69293	CD34 stem cell antigen
-6,45	0,999998	-7,21	0,999999	-6,83	U36760	HNF-3/forkhead homolog, brain factor 1
-6,32	1	-7,31	1	-6,815	X70296	Serine protease inhibitor 4
-6,68	1	-6,18	1	-6,43	U41739	Mus musculus RBP associated molecule RAM14-1 mRNA
-4,92	1	-7,87	1	-6,395	D26483	Mouse mRNA for PE31/TALLA
-4,61	0,999991	-7,98	0,999997	-6,295	AW120874	Mus musculus cDNA, 3 end
-6,33	1	-6,01	1	-6,17	AW061260	Nestin (neuronal stem cell protein)
-5,62	0,999999	-5,33	1	-5,475	AF395837	adaptor molecule SRCASM (Src activating signaling molecule)
-3,83	0,999991	-6,94	0,999999	-5,385	AF100171	Mus musculus myelodysplasia/myeloid leukemia factor 1 (Mlf1) mRNA
-5,14	1	-5,34	0,999993	-5,24	Al835060	Mus musculus cDNA, 3 end
-4,94	1	-5,32	1	-5,13	X70922	M.musculus Ly-6F.1 gene for neurotoxin homologue, exons 1-2
-5,13	0,999999	-5,08	0,999997	-5,105	S69293	CD34 stem cell antigen
-4,1	0,999667	-6	0,999533	-5,05	U56724	Mus musculus dystrophin gene, exon 23 and partial cds
-2,73	0,999999	-7,18	0,999999	-4,955	NM_133362.1	Erythroid diff. regulator
-5,32	1	-4,54	0,999997	-4,93	D76440	Necdin
-3,47	0,999999	-6,29	1	-4,88	X56304	Tenascin C
-5,92	0,999999	-3,61	0,999885	-4,765	BC026326	alpha inhibiting activity polypeptide 1
-4,75	1	-4,67	0,999999	-4,71	Y13185	Mus musculus mRNA for stromelysin-2
-5,3	0,999392	-3,94	0,999104	-4,62	AW122851	Mus musculus cDNA, 3 end
-4,59	1	-4,45	1	-4,52	AF061260	Mus musculus immunosuperfamily protein BI2 mRNA
-6,09	0,999932	-2,83	0,999998	-4,46	M31418	Interferon activated gene 202
-3,27	0,999977	-5,57	1	-4,42	Al604345	Mus musculus cDNA, 3 end
-2,79	0,999915	-6,04	0,999977	-4,415	AV230686	Mus musculus cDNA, 3 end
-4,32	0,999643	-4,49	0,999977	-4,405	Al847230	Mus musculus cDNA, 3 end
-4,38	0,999885	-4,29	0,999999	-4,335	AF038939	Mus musculus zinc finger protein (Peg3) mRNA
-4,52	0,999995	-4,07	0,999941	-4,295	AF061260	Mus musculus immunosuperfamily protein BI2 mRNA
-3,98	0,999921	-4,61	0,999997	-4,295	AW049373	Mus musculus cDNA, 3 end
-2,92	0,999963	-5,63	0,999996	-4,275	M63630	Mus musculus predicted GTP binding protein (IRG-47) mRNA
-5,29	0,999999	-3,19	1	-4,24	AB006758	Mus musculus Pcdh7 mRNA for BH-protocadherin-a
-4,95	1	-3,48	1	-4,215	L47600	Mus musculus cardiac troponin T isoform A3b mRNA
-4,23	1	-4,15	0,999999	-4,19	AB021966	Mus musculus mRNA for adhesion protein RA175C
-3,22	0,999973	-5,11	0,999993	-4,165	M83985	Mouse phosphoprotein (F1-20) mRNA
-5,21	1	-3,06	1	-4,135	X66402	Stromelysin-3; Matrix metalloproteinase 3
-5,2	1	-3,06	0,999999	-4,13	AF297645.1	stromelysin-3 (Mmp 11)
-4,17	1	-4,08	1	-4,125	NM_010211	four and a half LIM domain 1 (Fhl1)
-4,78	0,999989	-3,37	0,999997	-4,075	AW122260	Mus musculus cDNA, 3 end
-4,56	0,999996	-3,55	0,999997	-4,055	X16995	Hormone receptor
-4,74	0,999986	-3,37	0,999876	-4,055	M83649	Mus musculus Fas antigen mRNA
-2,69	0,999992	-5,34	0,999999	-4,015	U10420	N-methylpurine-DNA glycosylase
-3,06	0,999501	-4,86	0,999997	-3,96	M89798	Wingless-related MMTV integration site 5A
-3,24	0,999998	-4,65	0,999915	-3,945	U74683	Cathepsin C
-3,29	0,999997	-4,47	0,999908	-3,88	X56561	Homeo box D8
-4,14	0,999954	-3,59	0,999611	-3,865	M83649	Mus musculus Fas antigen mRNA

-3,55	0,999954	-4,14	0,999981	-3,845	AI854793	Mus musculus cDNA, 3 end
-3,69	0,999921	-3,99	0,999995	-3,84	XM_144949	Similar to H1 histone family, member X
-3,77	0,999999	-3,85	0,999995	-3,81	AA692074	Mus musculus cDNA, 5 end
-4,46	0,999999	-3,1	0,999999	-3,78	AJ007909	Mus musculus mRNA for erythroid differentiation regulator
-3,86	1	-3,6	1	-3,73	AF004428	Tumor protein D52-like 1
-3,13	0,999999	-4,26	1	-3,695	AI839082	Mus musculus cDNA, 3 end
-3,75	0,999981	-3,43	0,99995	-3,59	D85028	Mouse collapsin-1 mRNA
-3,07	0,99978	-4,02	0,999045	-3,545	X59252	Homeo box, msh-like 2
-3,81	1	-3,27	1	-3,54	AJ242663	Mus musculus mRNA for cathepsin Z precursor (ctsZ gene)
-3,49	0,999835	-3,55	0,999995	-3,52	AW124627	Mus musculus cDNA, 3 end
-3,83	0,999969	-3,16	0,999926	-3,495	U31966	Carbonyl reductase
-4,64	1	-2,28	0,99999	-3,46	X53929	Decorin
-3,9	0,999995	-3	0,999999	-3,45	BC027005.1	serum deprivation response
-3,2	0,99978	-3,67	0,999999	-3,435	U67187	Mus musculus G protein signaling regulator RGS2 (rgs2) mRNA
-3,79	0,99995	-2,96	0,999998	-3,375	AI852760	Mus musculus cDNA, 3 end
-2,63	0,999981	-3,98	0,999981	-3,305	AB006714	Mus musculus mRNA for dihydropyrimidinase related protein 1
-3,79	0,999999	-2,73	0,999999	-3,26	U09816	Mus musculus GM2 activator protein (Gm2a) mRNA
-3,7	1	-2,76	1	-3,23	XM_124980	protease, serine, 11 (Igf binding) (Prss11)
-2,79	1	-3,58	1	-3,185	D83745	Mus musculus mRNA for ANA
-3,42	0,999988	-2,9	0,999997	-3,16	L25274	Mus musculus transmembrane glycoprotein (DM-GRASP) mRNA
-4,76	0,999996	-1,51	0,999501	-3,135	U10090	Killer cell lectin-like receptor, subfamily A, member 4
-4,22	0,999969	-1,92	0,999765	-3,07	AW228162	Mus musculus cDNA, 3 end
-4,04	0,99996	-2,06	0,999915	-3,05	AV112006	Mus musculus cDNA
-4,27	0,999999	-1,79	0,999932	-3,03	AA285446	Mus musculus cDNA, 5 end
-3,71	1	-2,34	0,999592	-3,025	V00727	FBJ osteosarcoma oncogene
-3,21	0,999996	-2,83	1	-3,02	U38967	Prothymosin beta 4
-2,24	0,99971	-3,73	0,999045	-2,985	AI647003	Mus musculus cDNA, 3 end
-2,05	0,999994	-3,65	0,999351	-2,85	X59846	M.musculus GAS 6 mRNA associated with growth-arrest
-3,33	1	-2,35	0,999995	-2,84	M98530	Neuron specific gene family member 1
-3,47	0,999885	-2,09	0,999885	-2,78	U43319	Frizzled homolog 6, (Drosophila)
-1,79	0,999993	-3,76	0,999618	-2,775	AB017202	Mus musculus mRNA for entactin-2
-2,58	1	-2,96	1	-2,77	NM_019929	suppressor of mif two 3 (SMT3); SUMO
-3,36	0,999989	-2,08	0,999915	-2,72	NM_019813	drebrin 1 (Dbrn1)
-1,67	0,999467	-3,67	0,999999	-2,67	AW121625	Mus musculus cDNA, 3 end
-3,91	0,999979	-1,39	0,999932	-2,65	X00246	Mouse mRNA with a Set 1 repetitive element for a class I major histocompatibility complex(MHC) antigen
-2,95	0,999996	-2,03	0,999997	-2,49	M22115	Homeo box A1
-2,71	0,999969	-2,21	0,999999	-2,46	AA260005	Mus musculus cDNA, 5 end
-1,34	0,999995	-3,55	0,99978	-2,445	AI843108	Mus musculus cDNA, 3 end
-2,02	0,999835	-2,86	0,999995	-2,44	U52524	Hyaluronan synthase 2
-1,78	0,999667	-3,08	0,999045	-2,43	AF067806	Mus musculus cAMP-specific cyclic nucleotide phosphodiesterase PDE8 mRNA
-1,59	0,999997	-3,26	0,999954	-2,425	D45203	Mouse pentylenetetrazol-related mRNA PTZ-17 (3 UTR of E3.1)
-1,44	0,999999	-3,27	1	-2,355	M70642	Fibroblast inducible secreted protein
-3,06	0,999995	-1,63	0,99971	-2,345	AF047838	Mus musculus calcium-sensitive chloride conductance protein-1 (mCLCA1) mRNA
-2,7	0,999998	-1,98	0,999996	-2,34	L12029	Stromal cell derived factor 1
-1,91	0,999995	-2,65	0,999997	-2,28	AF004666	Solute carrier family 8 (sodium/calcium exchanger), member 1
-2,3	0,999986	-2,25	0,999937	-2,275	L23423	Integrin alpha 7
-2,18	0,999533	-2,3	0,999998	-2,24	L35261	Myeloblastosis oncogene-like 1
-3,14	0,999996	-1,32	0,999957	-2,23	AJ250491	Mus musculus mRNA for receptor activity modifying protein 3 (Ramp3 gene)
-2,21	0,999995	-2,23	0,999987	-2,22	X83577	Glypican 4

-2,2	1	-2,16	1	-2,18	L31397	Dynamin
-1,66	0,999999	-2,64	0,999921	-2,15	M32490	Insulin-like growth factor binding protein 10
-2,75	0,999999	-1,55	0,999747	-2,15	U35330	Histocompatibility 2, class II, locus Mb1
-1,38	0,999957	-2,9	0,999966	-2,14	XM_129595	similar to Interferon-activatable protein 204 (Ifi-204) (Interferon-inducible protein p204)
-2,71	0,999999	-1,53	1	-2,12	U50767	collagen alpha-1, type 1
-2	1	-2,19	1	-2,095	U26437	Mus musculus tissue inhibitor of metalloproteinases-3 (TIMP-3) gene
-1,93	0,999991	-2,22	0,999901	-2,075	AF052453	3 -phosphoadenosine 5 -phosphosulfate synthase 2
-1,85	0,999822	-2,25	0,999997	-2,05	X15052	Neural cell adhesion molecule
-2,45	0,999618	-1,65	0,999991	-2,05	Y13439	Mus musculus mRNA for stress activated protein kinase-3
-1,67	0,999975	-2,42	0,999998	-2,045	AV093331	Mus musculus cDNA
-2,23	0,999987	-1,79	0,999996	-2,01	AI853444	Mus musculus cDNA, 3 end
-1,69	0,999045	-2,3	0,999991	-1,995	AF028123	untitled
-1,61	1	-2,36	0,999999	-1,985	M58661	CD24a antigen
-1,44	1	-2,51	0,999999	-1,975	AF053454	Mus musculus tetraspan TM4SF (Tspan-6) mRNA
-1,74	1	-2,2	1	-1,97	D78382	Mouse DNA for tob family
-1,71	1	-2,23	1	-1,97	X78989	M.musculus mRNA for testin
-1,89	0,999999	-2,03	0,999999	-1,96	U89915	Mus musculus junctional adhesion molecule (Jam) mRNA
-2,11	0,999921	-1,77	0,99999	-1,94	AF061972	Mus musculus Tat-interacting protein TIP30 mRNA
-2,27	0,999809	-1,6	0,99971	-1,935	AF006071	Mus musculus alpha-catenin related protein (ACRP) mRNA
-2,13	0,999966	-1,73	0,999996	-1,93	AW123921	Mus musculus cDNA, 3 end
-2,42	0,999993	-1,42	0,999982	-1,92	M31419	Interferon activated gene 204
-1,36	0,999981	-2,46	0,999747	-1,91	U68535	Mus musculus aldo-keto reductase mRNA
-1,63	1	-2,19	1	-1,91	X15591	Cytotoxic T lymphocyte-associated protein 2 alpha
-2,03	0,999987	-1,78	1	-1,905	NM_005103.3	fasciculation and elongation protein zeta-1 (zygin-1, FEZ-1)
-1,74	1	-2,04	1	-1,89	AI853375	Mus musculus cDNA, 3 end
-2,16	0,999867	-1,52	0,999908	-1,84	AV341723	Mus musculus cDNA, 3 end
-1,41	1	-2,26	1	-1,835	U30244	Mus musculus Eph family ligand ELF-2 precursor mRNA
-1,68	1	-1,99	1	-1,835	AA833425	Mus musculus cDNA, 5 end
-1,97	0,999969	-1,69	0,999995	-1,83	D14883	CD82 antigen
-1,77	0,999973	-1,65	0,999999	-1,71	L31397	Dynamin
-1,59	0,999999	-1,83	0,999999	-1,71	XM_132155	similar to genethonin 1
-1,58	1	-1,79	1	-1,685	AI843399	Mus musculus cDNA, 3 end
-1,7	0,999996	-1,65	0,99999	-1,675	AW125272	Mus musculus cDNA, 3 end
-1,61	0,999999	-1,63	0,999997	-1,62	M90365	Junction plakoglobin
-1,76	0,999795	-1,43	0,999995	-1,595	M31419	Interferon activated gene 204
-1,54	0,999997	-1,59	1	-1,565	M27266	Fyn proto-oncogene
-1,44	0,999999	-1,69	1	-1,565	M88242	Mouse glucocorticoid-regulated inflammatory prostaglandin G/H synthase (griPGHS) mRNA
-1,38	0,999966	-1,75	0,99971	-1,565	D28818	Mouse mRNA for NBL4
-1,49	0,999989	-1,48	0,999993	-1,485	AA407794	Mus musculus cDNA, 3 end
-1,34	0,999993	-1,59	0,999994	-1,465	U31993	Interleukin 17 receptor
-1,47	0,999998	-1,35	1	-1,41	X66084	CD44 antigen
-1,4	0,999999	-1,36	1	-1,38	AW228316	Mus musculus cDNA, 5 end



ONCOGENOMICS

MLL-mediated transcriptional gene regulation investigated by gene expression profiling

Dorothe Schraets¹, Tanja Lehmann¹, Theo Dingermann¹ and Rolf Marschalek^{*1}

¹Institute of Pharmaceutical Biology, Biocenter N230, University of Frankfurt/Main, Marie-Curie-Str. 9, D-60439 Frankfurt/Main, Germany

The human mixed lineage leukemia (*MLL*) gene is involved in about 50 different chromosomal translocations, associated with the disease phenotype of acute leukemia. However, the normal function of *MLL* is less understood. Homozygous knockouts of murine *Mll* were embryonal lethal, while heterozygous disruption led to aberrant *hox* gene expression associated with skeletal malformations, growth retardation, and impaired hematopoiesis. To understand *MLL* functions on the molecular level, gene expression profiling experiments were performed with a pair of murine cell lines (*MLL*^{+/+} and *MLL*^{-/-}). Microarray hybridization experiments revealed 197 potential target genes that are differentially expressed, providing new and important clues about *MLL* functions.

Oncogene (2003) 0, 000–000. doi:10.1038/sj.onc.1206438

Keywords: *MLL*; expression profiling; development; differential gene expression

1. Introduction

MLL translocations are identified in a subset of acute leukemia patients (Cimino *et al.*, 1998; DiMartino and Cleary, 1999), and in patients with therapy-induced secondary leukemia (Felix, 1998). Unlike many other types of leukemia, the presence of an *MLL* translocation predicts early relapse and poor prognosis (for a review, see Cimino *et al.*, 1998; DiMartino and Cleary, 1999). Frequently, leukemic blasts are characterized by the expression of lymphatic and myeloid surface markers, defining *MLL* or *mixed lineage leukemia* gene (Ziemin van der Poel *et al.*, 1991). Recently, patients carrying *MLL* translocations were shown to exhibit a specific gene expression profile of activated and inactivated genes, suggesting a novel and specific disease phenotype (Armstrong *et al.*, 2002).

MLL protein belongs to the Trithorax protein family and is part of a nuclear regulatory mechanism that establishes an epigenetic transcriptional memory system (Francis and Kingston, 2001). This system is based on

high molecular weight protein complexes that exert histone/chromatin-modifying/remodeling and transactivating/repressing activities. In general, 'chromatin-modifier complexes' are composed of different combinations of Polycomb- and Trithorax-group proteins, as well as general transcription factors (e.g. GATA proteins) in combination with protein complexes that exert chromatin remodeling and histone-modifying activities. As a result, specific gene expression patterns will be established and maintained throughout subsequent mitotic cell cycles, which in turn will allow cells to cope with their cellular fate, or their specific developmental or differentiation pathways.

The importance of *MLL* protein for *hox* gene expression has been demonstrated in the embryos of heterozygous and homozygous *Mll* knockout mice. Expression of *MLL* protein was not necessary for turning on transcription of certain *hox* genes, but for the maintenance of their transcription (Yu *et al.*, 1998). Insufficient Hox protein concentrations led to impaired development (*Mll*^{+/−}), while the homozygous knockout of the *Mll* gene was embryonal lethal at day 10.5 p.c. (Yu *et al.*, 1995).

The human *MLL* gene became of interest because of its genetic alterations that are strictly correlated with human acute leukemias (acute myeloid leukemia or acute lymphoblastic leukemia). An illegitimate recombination event between the *MLL* gene and any of its many translocation partner genes results in reciprocal *MLL* fusion genes. Comparative twin analysis and quantitative analysis of leukemic cell clones on Guthrie test cards suggested that *MLL* translocations occur between the second and third month of pregnancy. By that time point, organogenesis has already taken place and fetal liver hematopoiesis has been activated (Wiemels *et al.*, 1999). Obviously, *MLL* translocations are only permissive in cells predisposed to develop in hematopoietic cells, while they were never detected in any other tissue of the human body.

Based on our assumptions of *MLL* protein function, the disruption of an *MLL* wild-type allele and the additional expression of aberrant *MLL* fusion gene products are likely to result in the establishment of incorrect gene expression patterns during development of hematopoietic stem cells or progenitors.

Recently, experimental data became available that explained *MLL* protein function in more detail: Firstly,

*Correspondence: R. Marschalek;
E-mail: Rolf.Marschalek@em.uni-frankfurt.de
Received 14 October 2002; revised 29 January 2003; accepted 30 January 2003



the MLL protein is rapidly proteolyzed into two large peptides, named MLL-N (amino acids 1 2666; p300) and MLL-C (amino acids 2719 3679; p180); the proteolytical sites were identified within two conserved amino-acid motifs (CS1: Q × D · GADD; CS2: Q × D · GVDD) and proteolytical cleavage occurs post-translationally between amino acids 2666/2667 and 2718/2719, respectively (Nakamura *et al.*, 2002; Yokoyama *et al.*, 2002; Hsieh *et al.*, 2003). The resulting MLL peptides interact via dimerization domains located in MLL-N (amino acids 1979 2130, FYRN) and MLL-C (amino acids 3656 3969; FYRC and SET domain). The two complex MLL peptides build the basic structure required for an MLL multiprotein super-complex (MLL MPSC) that consists of both MLL peptides and at least 27 additional proteins (Nakamura *et al.*, 2002). The MLL MPSC exerts chromatin remodeling and histone-modifying activities, as concluded from the composition of the MLL MPSC and from experimental data (Nakamura *et al.*, 2002). The post-translational cleavage and binding of MLL-N to MLL-C is protecting the MLL-N peptide from degradation (Hsieh *et al.*, 2003); Secondly, MLL translocations occur predominantly between the exons 9 and 12, thus disrupting the wild-type MLL protein between amino-acid positions 1362 and 1444, respectively. As a result of reciprocal translocations, the N-terminal fusion portion of MLL exhibits two different DNA binding domains (AT-hooks and MT-domain; Zeleznik-Le *et al.*, 1994; Birke *et al.*, 2002) and the critical regions for correct subnuclear localization (Yano *et al.*, 1997). Therefore, fusions of the N-terminal portion of the MLL protein (amino acids 1 1362/1444) with various partner protein sequences (derivative 11 proteins) should target the same subnuclear compartment, as evidenced from independent studies that have demonstrated a nuclear punctate distribution in immunohistochemical studies (Yano *et al.*, 1997; Caslini *et al.*, 2000). However, all derivative 11 proteins are missing the critical dimerization domain FYRN. This may indicate that derivative 11 proteins may not enter the MLL MPSC, but may compete for DNA binding (AT-hook, MT domain) at control regions or promoters of MLL MPSC target genes. Finally, fusion proteins containing the C-terminal portion of the MLL protein (derivative Z protein) contain both the FYRN and FYRC/SET interaction domain, as well as the proteolytical cleavage sites CS1 and CS2. Therefore, two cleaved peptides can be expected; one containing partner protein sequences fused to the PHD domain and the FYRN dimerization domain, while the other should be identical to the MLL-C peptide encoded by the wild-type MLL protein. The latter one exhibits the transcriptional transactivation domain, the FYRC dimerization domain, and the conserved SET domain. Thus, both post-translation cleavage products from a given derivative Z protein are functionally competent to compete for assembly into the MLL MPSC.

In conclusion, reciprocal fusion proteins of a given *MLL* translocation may either be assembled into the MLL MPSC, or are competing for DNA binding sites at

MLL MPSC target genes, and thus, they are able to alter or disable functions of the MLL MPSC, which in turn may result in altered or ectopic gene expression patterns.

Apart from *MLL* translocations in acute leukemia and the recent findings, our knowledge of wild-type MLL protein function is limited, in particular, our insights into potential target genes affected by the MLL MPSC. For this reason, we used an established cell system to learn more about the pivotal role of the MLL protein in a 'normal' cellular environment. Expression profile experiments were performed using murine fibroblast lines that carry either wild-type or homozygously Disrupted *Mll* alleles to identify potential target genes that are affected by the absence or presence of MLL protein.

Results

Hybridization experiments and data analysis

First, the two fibroblast cell lines were analysed for their genotype and transcriptional properties of wild-type and disrupted *Mll* alleles (see Figure 1). In a second step, three independent mRNA preparations (one a week) were isolated from the two lines and pooled to generate a normalized probe for the chip hybridization experiments. The complete experiment was repeated after 6 months of continuing cell culture to assure the significance of the data and to prevent arguments against genetic instability of the maintained lines. Both data sets were compared and the results indicated that both cell lines exhibit extremely divergent gene expression profiles, with many genes up- and downregulated (Pearson correlation coefficient r of experiment 1 0.8859; experiment 2 0.8853; see Figure 2). Using different filter set criteria (increase/decrease (A), signal \log_2 ratio > 1.32 or < -1.32 (B) and change P -value > 0.999 or < 0.001 (C)), a total of 197 differentially expressed candidate genes were consistently identified in both experiments. In *Mll* knockout cells, 136 genes were overexpressed, while in *Mll* wild-type cells 61 candidate genes were overexpressed. The top 40 gene lists of both cell lines are summarized in Tables 1 and 2 and the \log_2 expression rates of these genes were visualized in heat maps (Figure 3).

Differential gene expression in *Mll* knockout cells

The most highly activated genes identified in the *Mll* knockout cells code for the p53 apoptosis effector related to *Pmp22* (*PERP*), vascular smooth muscle α -actin (*VSMA*), CD34 stem cell antigen, HNF-3/forkhead homolog (Brain factor 1), serine protease inhibitor 4 (*PN-1*), RBP-associated molecule *RAM 14-1*, *PE31/TALLA-1*, *Nestin*, adaptor molecule *SRCASM*, and the Myelodysplasia/Myeloid leukemia factor 1.

The *PERP* gene has been identified as a direct p53 target gene upon induction of apoptosis in murine embryonal fibroblasts (MEFs). *PERP* gene transcrip-

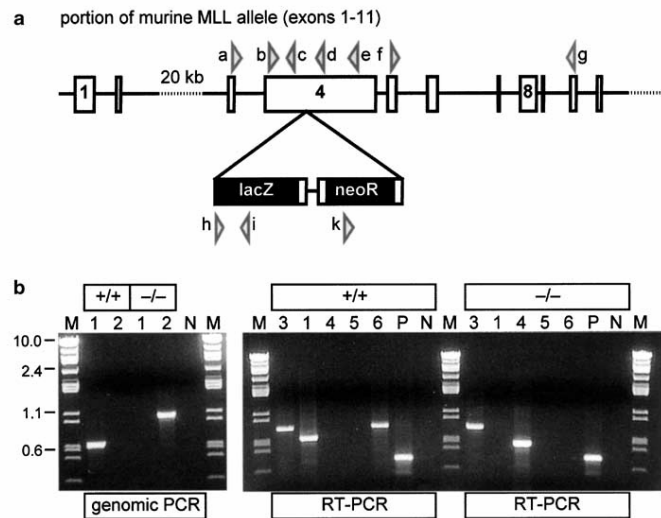


Figure 1 Integrity and expression of wildtype and knockout murine *MLL* alleles. (a) Scheme of the 5'-end of the murine *MLL* gene (exons 1–11). The *MLL* knockout cell line carries the in-frame-insertion of the LacZ open reading frame into *MLL* exon 4. A series of oligonucleotides (a–k) is shown by triangles. (b) Genomic and RT-PCR analyses. Genomic DNA and isolated RNA from the *MLL*^{+/+} and *MLL*^{-/-} cell lines were used for PCR analyses. Left panel: genomic PCR using the oligonucleotide combinations a–c (3: 837 bp), b–d (1: 689 bp) and b–i (2: 1060 bp). Right panel: RT-PCR using the oligonucleotide combinations a–c (3: 837 bp), b–d (689 bp), h–i (4: 550 bp), k–e, f–g (6: 833 bp), positive (P: GAPDH: 432 bp), and negative controls, respectively. A PCR fragment in reaction 4 is indicative for the genomic LacZ cassette insertion, while reaction 6 verifies the expression of the *MLL* gene beyond the insertion site of the LacZ/NeoR cassette.

tion correlates with p53-dependent apoptosis but not with p53-mediated cell cycle arrest. PERP predominantly localizes to the endoplasmic reticulum and Golgi apparatus, and displays sequence similarity to PMP-22/GAS-3 (peripheral myelin protein 22/growth arrest specific 3), a tetraspan membrane protein that is commonly mutated in human hereditary peripheral neuropathies (i.e. Charcot-Marie-Tooth). PERP appears to function in part by regulating cell proliferation and apoptosis (Attardi *et al.*, 2000).

vSMA is an important protein for smooth muscle contractions and involved in the process of wound healing and differentiation into myofibroblasts (Cogan *et al.*, 2002). The overexpression of vSMA was verified by immunohistological experiments and led to an alteration in the cellular morphology of *MLU*^{-/-} and *MLL*^{+/+} cells (unpublished data).

Relatively, little is known about the murine CD34 stem cell antigen. The murine counterpart of the human CD34 stem cell antigen has less importance for developmental processes of the hematopoietic system than in the human system. However, it is interesting to note that the CD34 antigen is highly expressed in *MLI*^{-/-} fibroblasts, while transcription of CD34 is shut down in fibroblasts expressing the MLL protein (First experiment >100-fold; second experiment 74.5-fold). This may argue for a model in which stem cells that

express high levels of the CD34 antigen have their *Mll* gene turned off. During hematopoietic differentiation, the *Mll* gene is transcriptionally turned on, and thus, CD34 production is lowered and finally shut down. If this is also true for human cells, long-term cultivation of hematopoietic stem cells, without losing differentiation potential, may be performed by blocking *MLL* gene expression using RNAi technologies.

HNF-3/forkhead homolog (Brain factor 1; BF-1) belongs to the *winged helix* gene family that encodes transcriptional repressor proteins. BF1 plays a pleiotropic role during development of the cerebral hemispheres of the brain. Mice lacking BF1 have defects in the morphogenesis of the dorsal telencephalon (e.g. neocortex) and the ventral telencephalon (Hanashima *et al.*, 2002). BF1 interacts with several SMAD proteins via the FAST-2 protein (forkhead activin signal transducer-2) protein. This prevents SMAD2/4 proteins to bind to their cognate DNA response elements, located in several promoters of TGF- β inducible genes. Moreover, BF-1 antagonizes TGF- β signaling (Dou *et al.*, 2000).

Serine protease inhibitor 4 (protease-nexin 1, PN-1) is a serine protease inhibitor (serpin) that inactivates several proteases, including thrombin, urokinase, plasminogen activators (PA), and plasmin. It also plays a role in regulating proteolytic activity generated by the

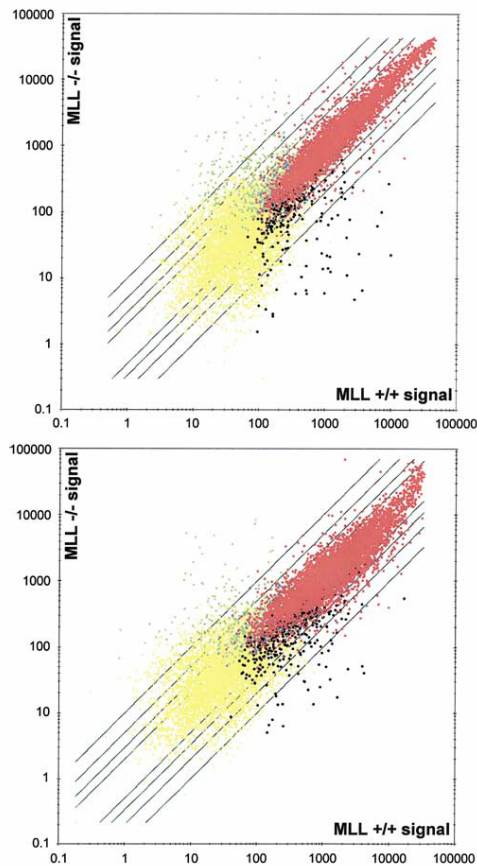


Figure 2 Scatter data of the chip hybridization experiment. Scatter blots showing the raw data of the first and second chip hybridization experiment. The second experiment was performed 6 months after the first experiment. Red dots: present in both population but unaffected; yellow dots: absent or marginal in both populations; green dots: genes only present in MLL^{-/-} cells; black dots: only present in MLL^{+/-} cells. The Pearson correlation coefficients were 0.8859 and 0.8853, respectively. The diagonal lines indicate two-, three-, five-, and 10-fold overexpression for both subsets

PA system. PN-1 is known to be involved in tissue remodeling, cellular invasiveness, matrix degradation, and tumor growth. PN-1 is highly secreted by astrocytes in the brain but overexpression of PN-1 leads to progressive neuronal and motor dysfunction in post-mitotic neurons because of altered postsynaptic transmission (Meins et al., 2001).

RBP-Jk associated molecule RAM 14-1 (Kyo-T2) mRNA represents a splice variant of the *Kyo-T* gene. It encodes two LIM domains and a specific interaction

domain for the RBP-Jk protein. RAM 14-1 protein was identified as a specific inhibitor of the Notch-1 pathway by preventing the Notch-1 IC/RBP-Jk heterodimeric protein from binding to its cognate DNA target site (Taniguchi et al., 1998). Thus, the Notch-1 and subsequent differentiation pathways are blocked. The currently known target genes affected by the Notch-1 pathway are the *hairy/enhancer of split (hes)* genes, the *HERP1* and *HERP2* genes, and the *NFkB2* gene. They are important for somatic recombination, MHC expression and are involved in the processes of neuronal and sensory organ development (de la Pompa et al., 1997). RBP-Jk protein heterodimerizes also with EBNA2 and EBNA3A-C, and thus establishes ectopic gene expression patterns leading to cellular transformation after EBV infection (summarized in Figure 4). More interestingly, both the *notch-1* and the *RBP-Jk* knockouts are embryonal lethal at day 10.5 p.c., the same day of development when homozygous *Mll* knockout mice die *in utero*. This argues for a link of the *Mll* knockout phenotype with impaired Notch-1 or RBP-Jk functions.

The four and a half LIM domain 1 gene transcript (*FLH1*; *Kyo-T1*) was also found to be activated. This transcript represents another splice form of the *Kyo-T* gene that is predominantly found in the cardiovascular system and in striated muscles (Chu et al., 2000). *FLH1* expression is linked to key events in myogenesis and functions as transcriptional coactivator for CREB or CREM transcription factors.

PE31/TALLA-1 (A15) is a member of the *transmembrane 4* gene superfamily and was isolated by differential screening of the cDNAs that are preferentially expressed on immature T-cells. TALLA-1 antigen has been found to be expressed exclusively on T-ALL cell lines, but was completely absent on peripheral blood mononuclear cells, including fresh and activated T cells (Takagi et al., 1995). *A15* encodes 244 amino acids with four hydrophobic domains. Northern blot analysis of the RNA samples from various mouse tissues disclosed that the *A15* transcripts are expressed most strongly in the brain, and are detectable in the colon, muscle, heart, kidney, and spleen. *In situ* hybridization of the mouse brain with an *A15* ribo-probe suggests that A15 protein may have a fundamental role within the higher nervous system (Hosokawa et al., 1999).

Nestin is an embryonic intermediate filament protein that represents a neural precursor/embryonal stem cell marker and is correlated with a differentiation fate to astroglial cells. Nestin may play an important role in remodeling and repairing the postnatal and adult central nervous system. Nestin has also been implicated in myogenesis (Wei et al., 2002).

The SRCASM protein is an adapter protein that interacts with and is phosphorylated by Fyn, a Src family tyrosine kinase. Phosphorylated SRCASM protein is able to interact with Grb-2 and the regulatory subunit of phosphoinositol 3-kinase, p85, and thus links a cell surface receptor to several signal cascades (Seykora et al., 2002).

Myelodysplasia/myeloid leukemia factor 1 (Mlf1) is a nuclear proto-oncoprotein involved in the

Table 1 Expression profile of *MLL* knockout cells. Top 40 list includes Genbank accession codes (GAC#), gene names (target gene), and \log_2 factors of differential gene expression (signal \log_2 ratios)

GAC#	Target gene	Sig \log_2 ratios	
A1854029	p53 apoptosis effector related to Pmp22 (PERP)	8.57	9.11
X13297	mRNA for vascular smooth muscle α -actin	7.86	7.97
S69293	CD34 stem cell antigen	7.49	6.22
U36760	HNF-3/forkhead homolog, Brain factor 1	6.45	7.21
X70296	Serine protease inhibitor 4	6.32	7.31
U41739	RBP-associated molecule RAM 14-1 mRNA	6.68	6.18
D26483	mRNA for PE31/TALLA-1	4.92	7.87
AW061260	Nestin (neuronal stem cell protein)	6.33	6.01
A1840130	Adaptor molecule SRCASM (Src activating signaling molecule)	5.62	5.33
AF100171	Myelodysplasia/myeloid leukemia factor 1 (Mlf1)	3.83	6.94
A1153421	Erythroid diff. regulator	2.73	7.18
D76440	Nedrin	5.32	4.54
X56304	Tenascin C	3.47	6.29
Y13185	Stromelysin-2	4.75	4.67
AF061260	Nectin-like protein 2	4.49	4.45
M31418	Interferon-activated gene 202	6.09	2.83
AF038939	Peg-3 zinc-finger	4.38	4.29
L47600	Cardiac Troponin T (subtype A3b)	4.95	3.48
AB021966	mRNA for Adhesion protein RAI175C	4.23	4.15
AF297645	Stromelysin-3	5.20	3.06
A1839950	Four and a half LIM domain 1 (Fhl1)	4.17	4.08
M83649	Fas antigen	4.74	3.37
U74683	Cathepsin C	3.24	4.65
X56561	Homeo box D8	3.29	4.47
AF004428	Tumor protein D52-like 1	3.86	3.60
X59252	Homeo box msh-like 2	3.07	4.02
AJ242663	mRNA for Cathepsin Z precursor (ctsZ gene)	3.81	3.27
X53929	Decorin	4.64	2.28
A1839175	Serum deprivation response	3.90	3.00
U09816	Gm2 activator protein (Gm2a) mRNA	3.79	2.73
D83745	ANA	2.79	3.58
V00727	FBJ osteosarcoma oncogene	3.71	2.34
U38967	Prothymosin beta 4	3.21	2.83
U43319	Frizzled homolog 6	3.47	2.09
AW120986	Suppressor of mif two 3 (SMT3); SUMO	2.58	2.96
M22115	Homeo box A 1	2.96	2.03
X83577	Glypican 4	2.21	2.23
M31419	Interferon activated gene 204	2.42	1.42
AA763466	Collagen type I, alpha 1	2.71	1.53
D78382	TOB	1.74	2.20

t(3;5)(q25.1;q34) translocation (NPM MLL1 fusion protein) associated with acute myeloid leukemia (Yoneda-Kato *et al.*, 1996). Mlf1 tissue distribution is restricted during development and postnatal life, with high levels present only in skeletal, cardiac, and selected smooth muscle, gonadal tissues, and rare epithelial tissues including the nasal mucosa and the ependyma/choroid plexus in the brain. Mlf1 transcripts were undetectable in the lympho-hematopoietic organs (embryonic and adult mouse), suggesting that the NPM-MLL1 fusion protein contributes to MDS/AML by ectopic overexpression of parts of the MLL1 protein in hematopoietic tissues. MLL1 protein is able to interact with the myeloid nuclear differentiation factor (MND1) that is also a nuclear protein expressed specifically in the developing cells of the myelomonocytic lineage, including the end-stage monocytes/macrophages and granulocytes. Moreover, elevated MLL1 expression is correlated with malignant progression from myelodysplastic syndrome.

Some *hox* genes were identified to be transcriptionally activated in *MLL*^{-/-} cells: the homeo box genes *hoxA1*, *hoxD8*, and *msh-like 2* (*hox8.1*) were found to be upregulated 5.6-fold, 14.8-fold, and 11.7-fold, respectively. HoxA1 protein is involved in patterning the hindbrain and neuronal crest cells. Targeted disruption of *hoxA1* led to severe abnormalities of the optic capsule and membranous labyrinth (Pasqualetti *et al.*, 2001) and to malformations of the rhombomeric organization of the hindbrain. HoxD8 protein was found to be expressed in mesoderm and neuroectodermal structures, as well as in limbs and gonads (Izpisua-Belmonte *et al.*, 1990). Hox8.1 protein is expressed in the surface ectoderm and in the optical vesicle before invagination (Monaghan *et al.*, 1991). SUMO proteins are small ubiquitin-related polypeptides that are reversibly conjugated to many nuclear proteins. SUMOylation of target proteins does not lead to degradation but instead appears to regulate protein-protein interactions, intracellular localization and protects some modified targets from ubiquitin-dependent degradation. SUMO alters


Table 2 Expression profile of *MLL* wild-type cells. Top 40 list includes Genbank accession codes (GAC#), gene names (target gene), and \log_2 factors of differential gene expression (signal \log_2 ratios)

GAC#	Target gene		Sig \log_2 ratios
X16009	Mitogen regulated protein/proliferin (Mrp/plf3 gene) exon 1	9.23	6.61
NM_008597	Matrix gamma-carboxyglutamate protein (MGP)	6.34	6.56
L32838	IL-1R antagonist	9.10	3.26
NP_056559	Ectonucleotide pyrophosphatase/phosphodiesterase 2 (Enpp2)	8.85	2.53
AF064749	Collagen type VI α 3	5.92	4.99
K03235	Proliferin	4.89	6.01
NW_000027	Dimerization cofactor of HNF-1 α (DCoH)	7.22	3.43
L00653	Mast cell protease 7 (mcp7)	5.81	4.70
AF045887	Angiotensinogen	4.83	5.58
AF045801	Gremlin	5.45	4.43
AA673487	EST cDNA	6.87	2.75
U70132	Paired-like homeodomain transcription factor 2/RIEG/ARPI/Ptx2	5.25	4.23
AB008174	HNF1 beta	4.00	4.78
AW214136	EST cDNA	5.45	2.70
L06047	Glutathion-S-transferase α 1	3.03	5.04
L19932	TGF- β induced β ig-h3 gene, 68kDa	4.62	3.31
Z18272	Procollagen, type VI, α 2	2.83	4.37
AF081789	Surface antigen AA4	4.40	2.62
NM_011844	Monoglyceride lipase (Mgl)	3.75	2.79
X14432	Thrombomodulin (Thbd)	4.62	1.79
L07918	GDP diss. inhibitor, preferentially expr. in hematopoietic cells	4.45	1.95
U09968	p27 kipl	3.63	2.76
X66405	Procollagen, type VI, α 1	2.59	3.35
AI843571	EST cDNA	3.10	2.37
AF057156	Small proline-rich protein 1A (Sprr1a) gene	3.13	2.32
U06119	Cathepsin H	2.53	2.81
X74040	Mesenchyme fork head-1	3.44	1.66
L07803	Thrombospondin 2	3.13	1.95
AB009377	CAD (DNA fragmentation factor)	2.95	2.09
L12367	Adenylyl cyclase-associated protein (CAP)	3.67	1.34
U36384	Twist-related bHLH protein Dermo-1	2.36	2.60
X07439	Homeo box Hox 3.1	3.40	1.47
U49513	Small inducible cytokine A9	2.21	2.60
X65128	Growth arrest specific 1; GAS1	1.62	3.05
X97227	Cell surface glycoprotein CD53	2.66	1.66
AW122653	Mevalonate kinase	1.37	2.78
U51335	GATA-6	1.93	2.17
NM_030704	Cryac heat-shock protein 20-like protein	2.23	1.61
U15159	LIM-domain containing, protein kinase (LIMK)	1.64	1.88
U07159	Acetyl coenzyme A dehydrogenase	1.55	1.39

the interaction properties of its targets and affects their subcellular localization. Examination of the PML nuclear bodies, whose components are SUMOylated, has revealed this modification to be essential for structural and functional integrity (Seeler and Dejean, 2001).

TOB and ANA are both members of the TOB/BTG antiproliferative protein family. The *BTG* gene family includes *Pc3/Btg2*, *Btg1*, *TOB*, *TOB2*, *ANA/Btg3*, *Pc3k*, and others. Unphosphorylated TOB negatively regulates proliferation of osteoblasts and T cells by inhibiting G0/G1 to S transitions. Activation of the ras/raf cascade leads to ERK1/2-dependent phosphorylation of TOB (Suzuki *et al.*, 2002) that reduces its ability to block cell cycle progression (affects Rb, cyclin D1).

Differential gene expression in Mll wild-type cells

The most activated genes identified in the *Mll* wild-type cells code for Mrp/plf3 proliferin, matrix γ -carboxyglutamate protein (MGP), IL-1R antagonist, ectonu-

cleotide pyrophosphatase, collagen type VI alpha 3 (COL6A3), dimerization cofactor of hepatocyte nuclear factor 1 α (DCoH), mast cell protease 7, and angiotensinogen.

Proliferin has been described as a growth factor of the prolactin/growth hormone family and as an angiogenic placental hormone. It is normally expressed during cutaneous wound healing processes and is temporally expressed during the hair follicle cycle. Recently, it has been identified to be overexpressed in the end stage of fibrosarcomas, indicating a potential role as angiogenic hormone for malignantly transformed tumor cells (Toft *et al.*, 2001).

MGP has been identified as the key regulatory molecule after PTH-mediated inhibition of the bone mineralization process by osteoblasts (Gopalakrishnan *et al.*, 2001). MGP overexpression in developing limbs inhibits cartilage mineralization, delays chondrocyte maturation, and blocks endochondral ossification. MGP is a powerful and developmentally regulated inhibitor of cartilage mineralization, and appears to

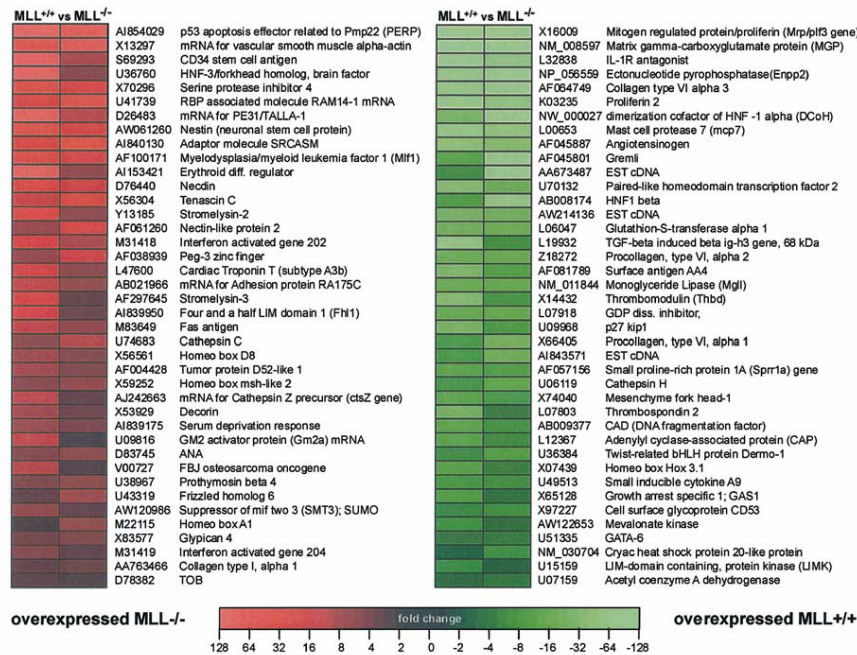


Figure 3 Heat maps of gene expression changes in MLL wildtype versus MLL knock-out cells. The top 40 genes expressed in MLL^{-/-} and MLL^{+/+} cells of both hybridization experiments are shown as heat maps. Color-codes: red-scale = overexpressed in MLL^{-/-} cells; green-scale = overexpressed in MLL^{+/+} cells. Genbank accession codes and gene names are given next to colored rectangles. Color scales represent eight log₂ steps

have a role in regulating chondrocyte maturation and ossification processes (Yagami *et al.*, 1999). In addition, MGP mRNA is constitutively expressed in normal vascular smooth muscle cells and plays a role in preventing vascular calcification processes.

IL-1R antagonist is a competitive inhibitor of the binding of IL-1 to the IL-1 receptor (IL-1R). Thus, IL-1R antagonist is a powerful inflammatory inhibitor that counteracts IL-1-mediated inflammation, septic shocks, and hematopoietic defects. Mice lacking the *IL-1R antagonist (IL-1Ra)* gene show abnormal development and homeostasis as well as altered responses to infectious and inflammatory stimuli (Irikura *et al.*, 2002). Absence of IL-1 signaling appears to assist the early embryonic blastomer cleavage (Spandorfer, 2000). In addition, IL-1R antagonist is highly expressed in embryos during implantation into the endometrial surface.

Ectonucleotide pyrophosphatase belongs to a family of membrane bound enzymes that degrade extracellular ATP and ATP derivatives, which are known to influence cell signaling, cell motility, inflammation, and apoptosis. It is also known to be an activation marker of basophil granulocytes (CD203c) after allergic response and can be transcriptionally upregulated by TGF- β . Over-

expression of ectonucleotide pyrophosphatase serves to prevent inflammation and influences cellular communication (for a review, see Bollen *et al.*, 2000).

COL6A3 is a primary constituent of the extracellular matrix and disruption of the *COL6A3* gene has been implicated in muscular dystrophies. It also plays an important role at later stages of neural crest development and promotes neural crest cell migration and attachment (Perris *et al.*, 1993). *COL6A3* shows specific interaction with decorin as well as interactions with other matrix proteins, for example, hyaluronan, heparan sulfate, and NG-2 proteoglycans (Burg *et al.*, 1996). Significantly, the NG-2 surface protein has been identified in MLL translocation patients (Armstrong *et al.*, 2002). *COL6A3* can be transcriptionally upregulated by TGF- β .

The dimerization cofactor of hepatocyte nuclear factor 1 alpha, the bifunctional DCoH protein, stabilizes HNF-1 alpha dimers and enhances their transcriptional activity (Mendel *et al.*, 1991). In addition, DCoH protein is a pterin-4- α -carbinolamine dehydratase that is involved in the regeneration of the cofactor tetrahydrobiopterin. HNF-1 α (TCF-1) has been identified to be biallelic inactivated in hepatic adenomas (Bluteau *et al.*, 2002) and plays an important role as nuclear

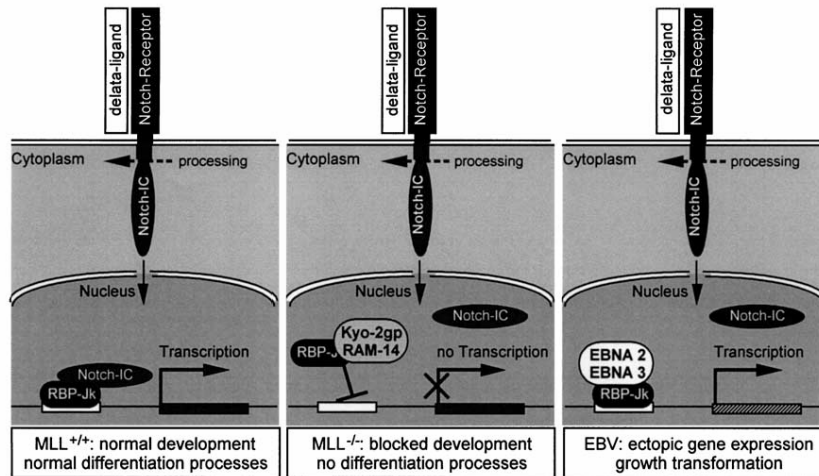


Figure 4 Notch-1/RBP-Jk pathways in mammalian cells. Left: intracellular Notch-1 (Notch-IC) is cleaved after ligand binding to the Notch-1 receptor. The cleaved intracellular domain enters the nucleus and heterodimerizes with RBP-Jk to activate specifically target gene transcription. Middle: in the presence of RAM 14-1 protein, Notch-IC is prevented from binding to RBP-Jk; the RBP-Jk/RAM14-1 heterodimer is unable to bind to its cognate DNA response element, and thus, transcription of target genes is blocked. Right: after EBV infection, EBNA proteins compete with Notch-IC for RBP-Jk binding; the heterodimers of different EBNA proteins and RBP-Jk lead to ectopic transcription of genes that are important for the EBV life cycle and subsequent transformation

target protein of the Wnt/ β -catenin signaling pathway. During mouse development, limb bud formation, lung formation, nephrogenesis and thymocyte development are critically dependent on Wnt signaling and TCF-1. Interestingly, *HNF-1 β* (*TCF-2*) has also been identified as a differentially regulated candidate gene in *MLL^{+/+}* cells.

Besides these, several transcription factors have been identified in *Mll* wild-type cells, which are involved in ecto- and mesodermal development during organogenesis.

Mesenchyme fork head-1 (MFH-1) is a winged helix/forkhead transcription factor that binds to the HNF-3 binding site and acts as positive transactivator. MFH-1 is expressed temporally in developing embryos, first in the non-notochordal mesoderm and later in areas of mesenchymal condensation in the prospective trunk region, in cephalic neural crest and cephalic mesoderm-derived mesenchymal cells of the prechordal region in early embryos. Moreover, BMP-2 induces the expression of MFH-1 in developing cartilaginous tissues, kidney, and arch arteries, and is essential for the normal development of the axial skeleton and aortic arch formation of mice. MFH-1-deficient mice die embryonically and perinatally, and exhibit an interrupted aortic arch, skeletal defects, defects in the neurocranium and the vertebral column (Iida *et al.*, 1997).

Twist-related bHLH protein Dermo-1 is a multi-functional basic helix-loop-helix transcription factor binding to the E-box consensus sequence (CANNTG) and has been shown to be a potent negative regulator

for gene transcription and apoptosis. Dermo-1 is a homolog to *Drosophila* Twist and is expressed in a subset of mesodermally and ectodermally derived tissues. Dermo-1 represses in a dose-dependent fashion the transcriptional activation mediated by the heterodimer MyoD/MEF-2 (myocyte enhancer factor 2; Gong and Li, 2002).

A set of *hox* genes was identified to be transcriptionally activated in *MLL^{+/+}* cells. The genes *hoxA7*, *hoxB9*, *hoxC9*, *hox 3.1* (*hoxC8*), *iroquois*, and the *paired-like homeo-domain transcription factor 2* (*RIEG*, *Ptx2*, *ARPI*) were upregulated 14.8-, 5.4-, 9.0-, 5.4-, 3.9-, and 26.7-fold, respectively. Significantly, the four genes *hoxa7*, *hox C8*, *hoxC9*, and *ARPI* have already been identified as *MLL* target genes (Arakawa *et al.*, 1998; Yu *et al.*, 1998; Milne *et al.*, 2002).

HoxB9 expression is correlated with myeloid and erythroid differentiation processes, while *HoxC8* is found to be mainly involved in neuronal development. Homozygous knockout of *hoxC9* is correlated with anterior transformation of several skeletal vertebrae and an additional pair of ribs (Le Mouellic *et al.*, 1992). The Iroquois homeobox protein 3 (*IRX-3*) plays a pivotal function in the initial specification of the vertebrate neuroectoderm, and in concert with other transcription factors for later subdivision of the anterior-posterior and dorso-ventral axis of the neuroectoderm. Iroquois proteins are effective early in development to specify the identity of diverse territories of the body, such as the dorsal head and dorsal mesothorax of *Drosophila* and the neural plate (Cavodeassi *et al.*, 2001). In addition,

iroquois genes are known targets of Pc-G/Trx-G complexes, an additional hint that an MLL target gene has been identified. ARPI is expressed highly in bone marrow cells, but is completely absent in leukemic blasts carrying an MLL translocation. The latter protein is also involved in developing processes of the brain (diencephalon, mesencephalon, and rhombencephalon), in postmitotic neuron development, as well as in the development of eyes, teeth and the umbilicus. Mutations lead to malformations similar to those described in Rieger syndrome (for a review, see Alward, 2000).

23

A number of genes were found to be activated in MLL-expressing cells that encode secreted proteins. Gremlin is a member of the Dan family of BMP antagonists that functions early in development. Gremlin is able to bind to BMP2, BMP4, and BMP7 and prevent them from binding to their cognate receptors. Binding to these BMPs influences many developmental processes, for example, mesoderm development, control of limb outgrowth, and lung development (Capdevila *et al.*, 1999). Gremlin is downregulated in transformed cells, while overexpression of Gremlin in tumor cell lines has a tumor suppressive effect.

β ig-h3 protein is secreted from many different cells after TGF- β treatment (Skonier *et al.*, 1992). Human fibroblasts secrete β ig-h3 into the extracellular matrix where it binds to fibronectin (FN) and type I collagen (Col I). It is involved in the tube formation process of blood vessels during angiogenesis, an important mechanism for embryonic development, wound healing, and tumor growth. In addition, transformed cells downregulate β ig-h3 protein.

The GDP dissociation inhibitor (D4) is a homolog of RhoGDI; both exhibit GDI activity against the Ras-related Rho family GTPases (CDC42Hs, rho and rac I proteins). During murine embryogenesis, GDP dissociation inhibitor (D4) transcripts are detected in yolk-sac cells, where the earliest hematopoietic precursors are found. When these precursors undergo proliferation and differentiation, a dramatic increase in GDP dissociation inhibitor (D4) expression is observed. Another argument for the significance of the GDP dissociation inhibitor (D4) arises from *in vitro* differentiation experiments; when totipotent murine embryonic stem cells develop into hematopoietic cells, the *GDP dissociation inhibitor (D4)* gene is activated with the onset of hematopoiesis and modulated during further development (Lelias *et al.*, 1993).

Sprr1a gene overexpression has been identified in injured neurons, where it is upregulated up to 60-fold. Overexpressed *Sprr1a* colocalizes with F-actin in membrane ruffles and augments axonal outgrowth and is essential for successful nerve regeneration (Bonilla *et al.*, 2002).

Growth arrest specific 1 (GAS-1) was identified in *Mll* wild-type cells. The gene product is known as a tumor suppressor and is able to inhibit growth transformation of tumor cells by blocking G0/G1 transition and DNA synthesis. The mechanism is p53-dependent and cannot be observed in p53-mutated cells or MDM-2-overexpressing cells. Besides its effects on cell cycle and

proliferation, GAS proteins are developmentally expressed in spatial temporal expression patterns. GAS-1 expression was identified in 8.5- to 14.5-day-old embryos where it is expressed in different abundance in most organ systems including the brain, heart, kidney, limb, lung, and gonad (Lee *et al.*, 2001).

The surface protein CD53 is an indicator for the presence of the ets family transcription factor Pu.1. CD53 has been identified as one of the target genes that have been identified in Pu.1 knockout cells when stably transduced with a PU.1-expressing retrovirus (Henkel *et al.*, 2002). In our study, the myeloid-specific genes coding for CD11a, CD18, GM-CSFR- α , c-fes (1.6-fold) as well as the B-cell specific genes like Btk, CD27, vav-1 (1.8-fold) are downstream targets of Pu.1 complexes that contain several other transcription factors (e.g. Sp1, C/EBP, AML1, Oct). Pu.1 is expressed very early in precursors for myeloid and lymphatic lineages. An invariant consequence of *Pu.1* mutations is a multi-lineage defect in the generation of progenitors for B and T lymphocytes, monocytes, and granulocytes.

GATA-6 protein is expressed in primitive endoderm structures and various mesoderm- and endoderm-derived tissues such as heart, liver, lung, gonad and gut where it plays a critical role together with GATA-4 and GATA-5 protein in regulating tissue-specific gene expression (Morrisey *et al.*, 1996). The homozygous knockout of *GATA-6* led to embryonal lethality (days 5.5-7.5 p.c.), with defects in visceral endoderm function and subsequent extraembryonic development (Morrisey *et al.*, 1998; Koutsourakis *et al.*, 1999). Interestingly, GATA-6 protein plays a role in vascular smooth muscle cells by regulating proliferation in response to mitogenic or mechanical stimulation.

LIM-kinase (LimK-1) plays a significant role in brain development, synapse formation and functions by acting on Rac-mediated actin reorganization and reverses cofilin-induced actin depolymerization (Yang *et al.*, 1998). Abnormal expression of LimK-1 is correlated with Williams syndrome, a mental disorder (Meng *et al.*, 2002). Between days 7.5 and 8.5 p.c, LimK-1 is expressed in three broad domains within the embryo, the neuroectodermal of the prospective forebrain and midbrain regions, the cardiac mesoderm, and the newly formed definitive endodermal derivatives, the foregut and hindgut. By day 10.0, LimK-1 remains prominently expressed in the ventromedial regions of the developing forebrain and midbrain, with continued expression in the hindgut. In adults, LimK-1 is expressed most prominently in the brain.

Verification of the obtained chip hybridization data by RT-PCR experiments

A series of oligonucleotides was designed and used with total RNA (DNA free) to perform RT PCR experiments to verify the hybridization data. In a first step, a total of 48 genes were tested and 41 reactions resulted in a specific PCR product. In a second step, 25 of these target genes were amplified with 25, 30, and 35 cycles to semiquantitatively reproduce the observed gene expres-

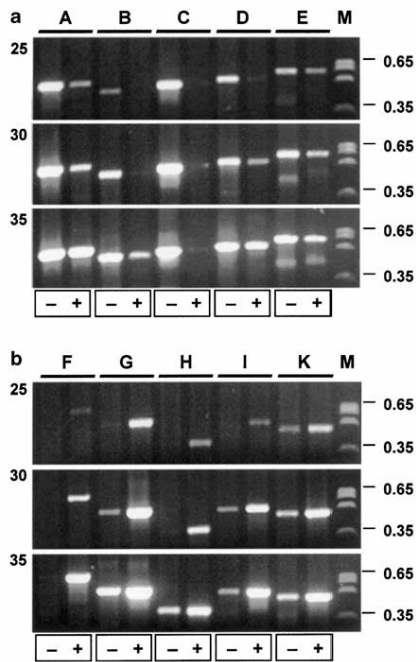


Figure 5 Verification of differential gene expression by RT-PCR. Several identified target genes were tested by semiquantitative RT-PCR experiment using 25–35 cycles. (a) RT-PCR of potential target genes overexpressed in *MLL*^{-/-} cells: A: mRNA for vascular smooth muscle α -actin (#2 of 40); B: RBP associated molecule RAM14-1 mRNA (#6 of 40); C: Serine protease inhibitor 4 (#4 of 40); D: Four and a half LIM domain 1 (#21 of 40); E: TOB (#40 of 40); M: DNA size marker. The number of cycles are indicated on the left. (b) RT-PCR of potential target genes overexpressed in *MLL*^{+/+} cells: F: IL-1R antagonist (#3 of 40); G: Collagen type VI alpha 3 (#5 of 40); H: Matrix γ -carboxyglutamate protein (#2 of 40); I: Angiotensinogen (#9 of 40); K: TGF- β induced beta ig-h3 gene, 68 kDa (#16 of 40); M: DNA size marker. The number of cycles are indicated on the left. Sizes are given on the right.

sion differences in *MLL*^{-/-} and *MLL*^{+/+} cells (all experiments were performed in triplicate). Representative results from 10 genes are shown in Figure 5. Although this type of analysis does not allow us to draw any accurate conclusions, the observed results of the RT-PCR experiment were in good agreement with the data obtained from the chip hybridization experiments. Specific oligonucleotides were designed for the human homolog of the interesting candidate gene *KyoT*, named *FHL-1*, enabling the analysis of cDNA derived from human cell lines. As shown in Figure 6, the Notch-1 IC inhibitor RAM14-1, named FHL-1C in the human system, was not found in the human T-cell line Jurkat, but was strongly expressed in the human t(4;11) cell line SEM (Greil *et al.*, 1994; Marschalek *et al.*, 1995). This demonstrates that the *KyoT/FHL-1* gene is not only

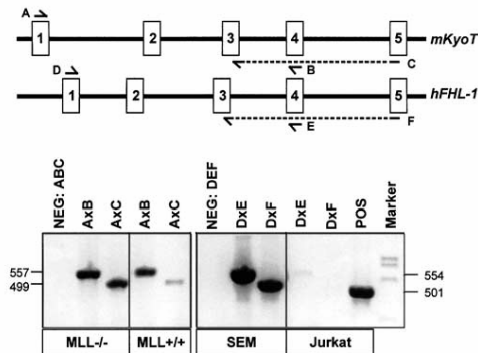


Figure 6 RAM14-1/FHL-1C is expressed in t(4;11) but not in Jurkat cells. RT-PCR experiments were performed on cDNA derived from *MLL*^{-/-}, *MLL*^{+/+}, SEM, and Jurkat cells. To distinguish between the splice variants of the mouse *KyoT* and the human *FHL-1* gene, oligonucleotides were designed to detect transcripts either containing exon 4 (primer combinations A, B, D, E) or missing exon 4 (primer combinations A, C, D, F). Top: scheme of the homolog genes *KyoT* and *FHL-1*. Bottom: RT-PCR analyses showing specific PCR amplicons after 35 cycles. Sizes of amplicons are given on the left (murine) and right (human).

highly transcribed in *MLL*^{-/-} cells, but also alternatively spliced. Thus, alternative splicing and the steady-state amount of the *KyoT/FHL-1* RNA in both the mouse and human system seemed to be affected by the gene dose provided by the endogenous *Mll/MLL* alleles.

Discussion

Potential MLL target genes have been unveiled by gene profiling experiments in a biological model system. Comparison of two independent experimental data sets led to the discovery of 197 potential target genes of the MLL protein, constituting the MLL MPSC. Interestingly, in the absence of MLL protein, more target genes were identified than in the presence of MLL (136 versus 61 target genes). This suggests that the MLL multi-protein supercomplexes are not only involved in the process of chromatin activation, but also in the process of chromatin inactivation. This is in agreement with the published data about the composition of the MLL MPSC (Nakamura *et al.*, 2002): the supercomplex is constituted in part by central components of the NuRD and Sin3A complex, indicating that histone deacetylase activity – associated with transcriptional downregulation – is presumably a central function of the MLL MPSC at specific target genes.

Several *hox* genes were transcriptionally influenced by the absence or presence of the *Mll* alleles. The three genes *hoxA7*, *hoxC8*, and *hoxC9* were described before as being transcriptionally maintained by MLL expression (Yu *et al.*, 1998; Milne *et al.*, 2002), thus verifying

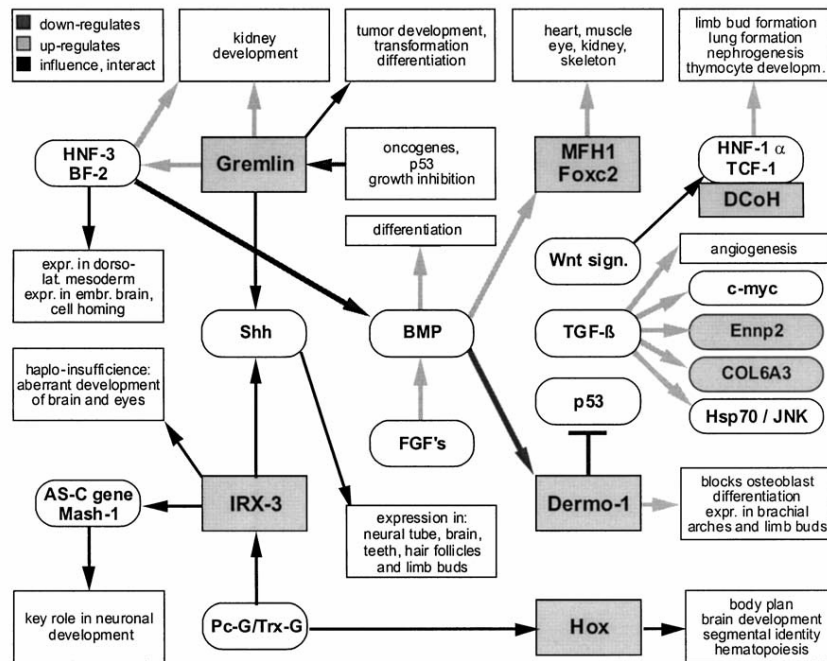


Figure 7 Interaction scheme of potential target gene products of MLL^{+/+} cells. Scheme of known interactions between various proteins that have been identified in the MLL^{+/+} cell line. Light gray rectangles: genes identified. White Ovals: genes affected or controlled by the light gray-marked genes. White rectangles: downstream effects of the light gray-marked or white oval-marked genes. Light gray arrows: 'upregulating'. Dark gray arrows: 'downregulating'. Black arrows: 'influencing' or 'directly interacting with'. Abbreviations: BMP, bone morphing proteins; TGF- β , tumor growth factor-beta; Shh, Sonic hedgehog; Pc-G/Trx-G, Polycomb-group/Trithorax-group; Hox, *hox* genes influenced by Pc-G/Trx-G; AS-C, Antennapedia/bicoid complex; HNF-3/BF-2, Hepatocyte nuclear factor-3/brain factor 2; C-myc, c-myc proto-oncoprotein; Hsp70/JNK, heatshock protein 70/Jun N-terminal kinase

the versatility of the biological system used in this study. The same is true for the *paired-like homeodomain transcription factor 2* gene (*RIEG*, *Ptx2*, *ARPI*) that has been identified to be transcriptionally activated by the MLL protein (Arakawa *et al.*, 1998).

Genes found to be activated in the presence of MLL protein code for factors involved in differentiation pathways, specially in the development of mesodermal and ectodermal tissues or organogenesis (e.g. *COL6A3*, *DCoH*, *gremlin*, β *ig-h3*, *GDID4*, *Sprr1*, *Mfh-1*, *irx-3*, *dermo-1*, several *hox* genes, *GATA-6*, LIMK). These genes affect muscle, heart, kidney, liver, limbs, bone, and the development of nerve cells, sensory organs and hematopoietic cells (summarized in Figure 7). This is in agreement with data from heterozygous *MLL* knockout experiments, where the expression of a LacZ reporter protein demonstrated MLL expression predominantly in the central nervous system, spinal ganglia, epithelial cells, endothelial cells, heart and liver (Yu *et al.*, 1995; Ayton *et al.*, 2001). In summary, potential target genes identified in wild-type MLL cells are involved in differentiation processes, tissue repair, or tissue remo-

deling. In addition, p27kip1 and GAS-1 are known tumor suppressor proteins involved in cell cycle regulation.

The absence of endogenous MLL protein is linked to the expression of a variety of interesting candidate genes that might be important for leukemogenesis. The overexpression of the *KyoT* splice variant, coding for the RAM 14-1 protein, interferes with correct differentiation by inhibiting the Notch-1 pathway. The human homolog of the *KyoT* splice variant, FHL-1C, is overexpressed in the human t(4;11) cell line SEM, but was completely absent in the human T-cell line Jurkat. This argues for the possibility that the reduction of the *MLL* gene dose (haplo-insufficiency) is mimicking the *MLL* knockout situation, and moreover, in the absence of at least one *MLL/MLL* allele, splicing of certain mRNA species can be altered.

Other interesting candidate genes identified in the *MLL* knockout cell line, including *HNF-3/BF-1*, *Mif1*, *FBJ osteosarcoma oncogene*, *Tenascin C*, *PE31/TALLA-1*, and *Tumor protein D52-like* gene, are all known to be overexpressed in malignant cells. Some other genes code

for proteins that modify the extracellular matrix and are important for tumor cell invasion (e.g. *Serin protease inhibitor 4*, *Stromelysin-2/3*). Thus, it seems that *Mll* knockout cells have upregulated genes that are linked to an oncogenic phenotype.

The fact that genes with oncogenic properties have been identified in *Mll* knockout cells has to be discussed also in the context of the available information of potential functions of MLL and MLL fusion proteins.

First, the common denominator of *MLL* translocations is the derivative 11 protein. As a result of a chromosomal translocation that occurs between *MLL* exons 9-12, only the AT hooks, the subnuclear localization sequences, and the MT domain are present in a given derivative 11 protein. All derivative 11 proteins are missing from the FYRN dimerization domain, and thus, should be unable to enter the MLL MPSC. However, derivative 11 proteins should target the same subnuclear compartment as the MLL MPSCs because of their subnuclear localization sequences (Yano et al., 1997). Therefore, derivative 11 proteins may compete with MLL MPSCs for binding sites at control regions or promoters of MLL target genes. As a direct result, potential MLL target genes may be transcriptionally deregulated which mimic the *MLL* knockout situation to an extent. If this is true, the identified target genes in the *Mll* knockout cells become highly interesting candidates for further analysis and may be important for leukemogenesis.

Conversely, reciprocal derivative Z fusion proteins should be processed into an N-terminal peptide (comprising partner protein sequences, the PHD domain, and the dimerization domain FYRN) and an MLL-C peptide that is identical to the MLL-C peptide derived from the wild-type MLL protein. The N-terminal peptide may enter the MLL MPSC (through FYRN) and compete with wild-type MLL-N peptides for binding to MLL-C peptides. If this is true, the assembly of an MLL MPSC will be altered or prevented. Moreover, if a given partner protein offers domains with DNA binding activities, for example AF4 (unpublished data), an altered MPSC will be targeted to partner protein specific DNA binding sites elsewhere in the genome, leading to ectopic gene activations or inactivations. In conclusion, the presence of a specific chromosomal translocation of the *MLL* gene correlated with human acute leukemias results in the expression of a derivative 11 and/or Z protein. The presence of derivative proteins may not only affect certain *hox* genes in their transcriptional activation or maintenance, but also may lead to broader changes in the genome. Such changes completely alter the gene expression pattern, similar to that observed in *Mll* knockout cells.

The results presented here do not claim that all possible MLL target genes were identified in this biological system. Moreover, we are aware of the fact that most genes identified in this study are presumably secondary and not primary targets of the MLL protein. However, the results of this study have helped to identify potential target genes affected by the presence or absence of MLL protein. A complete list of all 197

identified target genes, their log expression data, and the heatmap can be retrieved from our web site (MLL+.pdf; MLL-.pdf; Table1.pdf, Table2.pdf, Heat-Map.pdf, Scatter.pdf; Oligo.pdf): http://www.biozentrum.uni-frankfurt.de/Pharmazeutische_Biologie/Mitarbeiter/Marschalek/Main.html. Many of the identified target genes are currently being investigated in the human system for their transcriptional activity to unravel their functional importance.

Materials and methods

Establishment of cell lines and cell culture

Fibroblasts from isogenic MLL^{+/+} and MLL^{-/-} mice were isolated from dissected murine embryos at day 10.5 p.c. Cells were placed into cell culture and stable lines were obtained after transfection with an expression construct coding for the Polyoma large T antigen. Stable cell lines were cultured in Dulbecco's MEM (25 mm HEPES; Gibco-BRL), supplemented with 10% FCS Gold (PAA, Austria), 100 U/ml penicillin, 100 U/ml streptomycin, 2 mM glutamax and 50 μ M β -mercaptoethanol (all Gibco-BRL). Both cell lines were incubated in a humidified atmosphere (95%) with 5% CO₂ at 37°C.

Chip hybridization: RNA purification, labeling and hybridization

We used 1×10^7 cells to prepare total RNA using the Trizol (Gibco-BRL) purification method. This generally yielded between 80 and 100 μ g total RNA. This RNA was further purified and DNase digested (RNase-Free DNase Set, Qiagen) on RNeasy columns (Qiagen) and the quality was examined by gel electrophoresis. If the ribosomal bands were intact, the RNA was determined to be of good quality and used for subsequent synthesis of biotinylated antisense RNA. In addition, RNA was tested for the absence of genomic DNA by specific PCR reactions using sets of GAPDH oligonucleotides.

For first-strand cDNA synthesis, 9 μ l (13.5 μ g) of total RNA were mixed with 1 μ l of a mixture of three polyadenylated control RNAs, 1 μ l 100 μ M T7-oligo d(T)24 primer (5'-GGCCAGTGAATTGTAATACGACTCACTATAGGGAG-GCGG-(dT24)-3'), incubated at 70°C for 10 min and put on ice. Next, 4 μ l of 5 \times first-strand buffer, 2 μ l 0.1 M DTT and 1 μ l 10 mM dNTPs were added and the reaction was preincubated at 42°C for 2 min. Then, 2 μ l (200 U) Superscript II (Life Technologies) were added and incubation continued at 42°C for 1 h.

For second-strand synthesis, 30 μ l of 5 \times second-strand buffer, 91 μ l of RNase-free water, 3 μ l 10 mM dNTPs, 4 μ l (40 U) *Escherichia coli* DNA polymerase I (Life Technologies), 1 μ l (12 U) *E. coli* DNA ligase (TaKaRa), 1 μ l (2 U) RNase H (TaKaRa) were added and the reaction was incubated at 16°C for 2 h. Then, 2.5 μ l (10 U) T4 DNA polymerase I (TaKaRa) were added at 16°C for 5 min. The reaction was stopped by adding 10 μ l 0.5 M EDTA, and the double-stranded cDNA was extracted with phenol/chloroform and the aqueous phase was recovered by phase lock gel separation (Eppendorf). After precipitation, the cDNA was restored in 12 μ l of RNase-free water.

Double-stranded cDNA (5 μ l) was used to synthesize biotinylated cRNA using the BioArray High Yield RNA Transcript Labeling Kit (Enzo Diagnostics, Santa Clara). Labeled cRNA was purified using the RNeasy mini kit (Qiagen). Fragmentation of cRNA, hybridization to

MG_U74Av2 microarrays (Affymetrix), washing and staining as well as scanning of the arrays in a GeneArray scanner (Agilent) were performed as recommended in the Affymetrix Gene Expression Analysis Technical Manual. Absolute and comparison analyses were done using the Affymetrix GeneChip Analysis Suite 5.0 software. A scaling across all probe sets of a given array was included to compensate for variations in the amount and quality of the cRNA samples and other experimental variables.

Data analysis

After scanning, absolute and comparison analyses were carried out using the Affymetrix GeneChip Analysis Suite 5.0 software. A scaling across all probe sets of a given array was included to compensate for variations in the amount and quality of the cRNA samples and other experimental variables. To isolate differentially expressed genes, comparison files were further filtered using the Affymetrix Data Mining Tool 3.0 software. Filter criteria for robustly up- and downregulated

genes included Change = 'P' and Signal Log Ratio > 1.32 and Change *P*-value > 0.001 or Change = 'D' and Signal Log Ratio < -1.32 and Change *P*-value > 0.999.

RT-PCR analyses

RT-PCR analyses were performed as described (Löchner *et al.*, 1996). All oligonucleotides used for this study are summarized in a document (Oligo.pdf) that can be obtained from our web sites (see above).

Acknowledgements

We thank Dr Jay Hess for providing the MLL wild-type and knockout cell lines, Dr Peter Oliver for a critical reading of the manuscript, and Dr Ludger Klein-Hitpass for Affymetrix chip hybridization and computer analysis. This work was funded in part by the DFG-Grant MA 1876/3-1 and the BMBF cancer network Grant 01GS0104/K-S04T07.

References

- Alward WL. (2000). *Am. J. Ophthalmol.* **130**, 107–115.
- Arakawa H, Nakamura T, Zhadanov AB, Fidanza V, Yano T, Bullrich F, Shimizu M, Blechman J, Mazo A, Canaani E and Croce CM. (1998). *Proc. Natl. Acad. Sci. USA*, **95**, 4573–4578.
- Armstrong SA, Stouton JE, Silverman LB, Pieters R, den Boer ML, Minden MD, Sallan SE, Lander ES, Golub TR and Korsmeyer SJ. (2002). *Nat. Genet.*, **30**, 41–47.
- Attardi LD, Reczek EE, Cosmas C, Demicco EG, McCurrach ME, Lowe SW and Jacks T. (2000). *Genes Dev.*, **14**, 704–718.
- Ayton P, Snaedon SF, Palmer DB, Rosewell IR, Owen MJ, Young B, Presley R and Subramanian V. (2001). *Genesis*, **30**, 201–212.
- Birke M, Schreiner S, Garcia-Cuellar MP, Mahr K, Titgemeyer F and Slany RK. (2002). *Nucleic Acids Res.*, **30**, 958–965.
- Bluteau O, Jeannot E, Bioulac-Sage P, Marques JM, Blanc JF, Bui H, Beaudoin JC, Franco D, Balabaud C, Laurent-Puig P and Zucman-Rossi J. (2002). *Nat. Genet.*, **32**, 312–315.
- Bollen M, Gijssbers R, Ceulemans H, Stalmans W and Stefan C. (2000). *Crit. Rev. Biochem. Mol. Biol.*, **35**, 393–432.
- Bonilla IE, Tanabe K and Strittmatter SM. (2002). *J. Neurosci.*, **22**, 1303–1315.
- Burg MA, Tillet E, Timpl R and Stallcup WB. (1996). *J. Biol. Chem.*, **271**, 26110–26116.
- Capdevila J, Tsukui T, Rodriquez Esteban C, Zappavigna V and Izpisua Belmonte JC. (1999). *Mol. Cell*, **4**, 839–849.
- Caslini C, Shilatifard A, Yang L and Hess JL. (2000). *Proc. Natl. Acad. Sci. USA*, **91**, 2797–2802.
- Cavodeassi F, Modolell J and Gomez-Skarmeta JL. (2001). *Development*, **128**, 2847–2855.
- Chu PH, Ruiz-Lozano P, Zhou Q, Cai C and Chen J. (2000). *Mech. Dev.*, **95**, 259–265.
- Cimino G, Rapanotti MC, Sprovieri T and Elia L. (1998). *Haematologica*, **83**, 350–357.
- Cogan JG, Subramanian SV, Polikandriotis JA, Kelm Jr RJ and Strauch AR. (2002). *J. Biol. Chem.*, **277**, 36433–36442.
- de la Pompa JL, Wakeham A, Correia KM, Samper E, Brown S, Aguilera RJ, Nakano T, Honjo T, Mak TW, Rossant J and Conlon RA. (1997). *Development*, **124**, 1139–1148.
- DiMartino JF and Cleary ML. (1999). *Br. J. Haematol.*, **106**, 614–626.
- Dou C, Lee J, Liu B, Liu F, Massague J, Xuan S and Lai E. (2000). *Mol. Cell Biol.*, **20**, 6201–6211.
- Felix CA. (1998). *Biochim. Biophys. Acta*, **1400**, 233–255.
- Francis NJ and Kingston RE. (2001). *Nat. Rev. Mol. Cell Biol.*, **2**, 409–421.
- Gong XQ and Li L. (2002). *J. Biol. Chem.*, **277**, 12310–12317.
- Gopalakrishnan R, Ouyang H, Somerman MJ, McCauley LK and Franceschi RT. (2001). *Endocrinology*, **142**, 4379–4388.
- Greil J, Gramatzki M, Burger R, Marschalek R, Peltner M, Trautmann U, Hansen-Hagge TE, Bartram CR, Fey GH, Stehr K and Beck J. (1994). *Br. J. Haematol.*, **86**, 275–283.
- Hanashima C, Shen L, Li SC and Lai E. (2002). *J. Neurosci.*, **22**, 6526–6536.
- Henkel GW, McKercher SR and Maki RA. (2002). *Int. Immunol.*, **14**, 723–732.
- Hosokawa Y, Ueyama E, Morikawa Y, Maeda Y, Seto M and Senba E. (1999). *Neurosci. Res.*, **35**, 281–290.
- Hsieh JJ, Ernst P, Erdjument-Bromage H, Tempst P and Korsmeyer SJ. (2003). *Mol. Cell Biol.*, **23**, 186–194.
- Iida K, Koseki H, Kakinuma H, Kato N, Mizutani-Koseki Y, Ohuchi H, Yoshioka H, Noji S, Kawamura K, Kataoka Y, Ueno F, Taniguchi M, Yoshida N, Sugiyama T and Miura N. (1997). *Development*, **124**, 4627–4638.
- Irikura VM, Lagraoui M and Hirsh D. (2002). *J. Immunol.*, **169**, 393–398.
- Izpisua-Belmonte JC, Dolle P, Renucci A, Zappavigna V, Falkenstein H and Duboule D. (1990). *Development*, **110**, 733–745.
- Koutsourakis M, Langeveld A, Patient R, Beddington R and Grosveld F. (1999). *Development*, **126**, 723–732.
- Lee KK, Leung AK, Tang MK, Cai DQ, Schneider C, Brancolini C and Chow PH. (2001). *Dev. Biol.*, **234**, 188–203.
- Lelias JM, Adra CN, Wulf GM, Guillemot JC, Khagad M, Caput D and Lim B. (1993). *Proc. Natl. Acad. Sci. USA*, **90**, 1479–1483.
- Le Mouellie H, Lallemand Y and Brulet P. (1992). *Cell*, **69**, 251–264.
- Löchner K, Siegler G, Führer M, Greil J, Beck JD, Fey GH and Marschalek R. (1996). *Cancer Res.*, **56**, 2171–2177.
- Marschalek R, Greil J, Löchner K, Nilson I, Siegler G, Zweckbronner I, Beck JD and Fey GH. (1995). *Br. J. Haematol.*, **90**, 308–320.



- Meins M, Piosik P, Schaeren-Wiemers N, Franzoni S, Troncoso E, Kiss JZ, Brosamle C, Schwab ME, Molnar Z and Monard D. (2001). *J. Neurosci.*, **21**, 8830–8841.
- Mendel DB, Khavari PA, Conley PB, Graves MK, Hansen LP, Admon A and Crabtree GR. (1991). *Science*, **254**, 1762–1767.
- Meng Y, Zhang Y, Tregoubov V, Janus C, Cruz L, Jackson M, Lu WY, MacDonald JF, Wang JY, Falls DL and Jia Z. (2002). *Neuron*, **35**, 121–133.
- Milne TA, Briggs SD, Brock HW, Martin ME, Gibbs D, Allis CD and Hess JL. (2002). *Mol. Cell*, **10**, 1107–1117.
- Monaghan AP, Davidson DR, Sime C, Graham E, Baldock R, Bhattacharya SS and Hill RE. (1991). *Development*, **112**, 1053–1061.
- Morrissey EE, Ip HS, Lu MM and Parnacek MS. (1996). *Dev. Biol.*, **177**, 309–322.
- Morrissey EE, Tang Z, Sigrist K, Lu MM, Jiang F, Ip HS and Parnacek MS. (1998). *Genes Dev.*, **12**, 3579–3590.
- Nakamura T, Mori T, Tada S, Krajewski W, Rozovskaia T, Wassell R, Dubois G, Mazo A, Croce CM and Canaani E. (2002). *Mol. Cell*, **10**, 1119–1128.
- Pasqualetti M, Neun R, Davenne M and Rijli FM. (2001). *Nat. Genet.*, **29**, 34–39.
- Perris R, Kuo HJ, Gланville RW and Bronner-Fraser M. (1993). *Dev. Dyn.*, **198**, 135–149.
- Seeler JS and Dejean A. (2001). *Oncogene*, **20**, 7243–7249.
- Seykora JT, Mei L, Dotto GP and Stein PL. (2002). *J. Biol. Chem.*, **277**, 2812–2822.
- Skonier J, Neubauer M, Madisen L, Bennett K, Plowman GD and Purchio AF. (1992). *Cell. Biol.*, **11**, 511–522.
- So CW and Cleary ML. (2002). *Mol. Cell Biol.*, **22**, 6542–6552.
- Spandorfer SD. (2000). *Am. J. Reprod. Immunol.*, **43**, 6–11.
- Suzuki T, K-Tsuzuku J, Ajima R, Nakamura T, Yoshida Y and Yamamoto T. (2002). *Genes Dev.*, **16**, 1356–1370.
- Takagi S, Fujikawa K, Imai T, Fukuhara N, Fukudome K, Minegishi M, Tsuchiya S, Konno T, Hinuma Y and Yoshie O. (1995). *Int. J. Cancer*, **61**, 706–715.
- Taniguchi Y, Furukawa T, Tun T, Han H and Honjo T. (1998). *Mol. Cell Biol.*, **18**, 644–654.
- Toft DJ, Rosenberg SB, Bergers G, Volpert O and Linzer DI. (2001). *Proc. Natl. Acad. Sci. USA*, **98**, 13055–13059.
- Wei LC, Shi M, Chen LW, Cao R, Zhang P and Chan YS. (2002). *Brain Res. Dev. Brain Res.*, **139**, 9–17.
- Wiemels JL, Cazzaniga G, Daniotti M, Eden OB, Addison GM, Maser A, Saha V, Biondi A and Greaves MF. (1999). *Lancet*, **354**, 1499–1503.
- Yagami K, Suh JY, Enomoto-Iwamoto M, Koyama E, Abrams WR, Shapiro IM, Pacifici M and Iwamoto M. (1999). *J. Cell Biol.*, **147**, 1097–1108.
- Yang N, Higuchi O, Ohashi K, Nagata K, Wada A, Kangawa K, Nishida E and Mizuno K. (1998). *Nature*, **393**, 809–812.
- Yano T, Nakamura T, Blechman J, Sorio C, Dang CV, Geiger B and Canaani E. (1997). *Proc. Natl. Acad. Sci. USA*, **94**, 7286–7291.
- Yokoyama A, Kitabayashi I, Ayton PM, Cleary ML and Ohki M. (2002). *Blood*, **100**, 3710–3718.
- Yoneda-Kato N, Look AT, KIRSTEIN MN, Valentine MB, Raimondi SC, Cohen KJ, Carroll AJ and Morris SW. (1996). *Oncogene*, **12**, 265–275.
- Yu BD, Hanson RD, Hess JL, Horning SE and Korsmeyer SJ. (1998). *Proc. Natl. Acad. Sci. USA*, **95**, 10632–10636.
- Yu BD, Hess JL, Horning SE, Brown GA and Korsmeyer SJ. (1995). *Nature*, **378**, 505–508.
- Zeleznik-Le NJ, Harden AM and Rowley JD. (1994). *Proc. Natl. Acad. Sci. USA*, **91**, 10610–10614.
- Ziemin van der Poel S, McCabe NR, Gill HJ, Espinosa III R, Patel Y, Harden A, Rubinielli P, Smith SD, LeBeau MM, Rowley JD and Diaz MO. (1991). *Proc. Natl. Acad. Sci. USA*, **88**, 10735–10739.

Lebenslauf

Persönliche Daten Dorothe Schraets
geb. am 23.06.1974 in Geldern (Kreis Kleve)
deutsche Staatsangehörigkeit

Schulbildung

07/80-06/84 St. Peter und Paul Grundschule in Kerken (Kreis Kleve)
07/84-06/93 Liebfrauegymnasium Mülhausen (Kreis Viersen), Abschluss: Abitur

Studium

10/93-05/98 Pharmaziestudium an der J.W. Goethe - Universität Frankfurt/Main
04/97-07/97 Forschungsarbeit am Kings College in London; Dept. of Pharmacy
09/95 1. Staatsexamen
06/98 2. Staatsexamen
07/99 3. Staatsexamen und Erlangung der Approbation

Promotion

08/99-02/03 Anstellung als wissenschaftliche Mitarbeiterin am Institut für Pharmazeutische Biologie der J.W. Goethe - Universität Frankfurt/Main unter der Leitung von Prof. Dr. R. Marschalek

Praktische Tätigkeiten

06/98-11/98 Pharmaziepraktikum in der Apotheke am Weissen Stein, Frankfurt
12/98-05/99 Pharmaziepraktikum in der Labor- und Qualitätsservice GmbH (LQS); Eschborn
07/99-10/02 Tätigkeiten in der Apotheke am Weissen Stein, Frankfurt/Main
08/99-12/02 Tätigkeiten in der Kettenhof-Apotheke, Frankfurt/Main; Notdienste

Akademische Lehrer

Prof. Dr. H. Blume, Prof. Dr. Th. Dingermann, Prof. Dr. E. Ehlers, Prof. Dr. H. Hoffmann, Prof. Dr. H. Linde (†), Prof. Dr. Dr. Dr. h.c. E. Mutschler, Prof. Dr. J. Kreuter, Prof. Dr. W. Müller, Prof. Dr. C. Noe, Prof. Dr. D. Steinhilber

Publikation

Dorothe Schraets, Tanja Lehmann, Theo Dingermann & Rolf Marschalek

“MLL mediated transcriptional gene regulation investigated by gene expression profiling”

2003, **Oncogene** *in press*

Vorträge und Konferenzteilnahmen

Signalling Pathways in Cellular Differentiation

Kongress in Ulm, Teilnahme im Oktober 2001

“Identification of the target genes of the human MLL protein”

Vortrag und Posterbeitrag auf dem DPhG European Graduate Student Meeting in Frankfurt/Main; Februar 2002

“Involvement of human MLL in akute lymphoblastic leukemia in early childhood”

Vortrag auf der Jahrestagung der Kind-Phillip Stiftung für Leukämieforschung in Wilse-de, Juni 2002

“Identification of target genes of the human MLL protein”

Posterbeitrag auf dem *Chip-Meeting* im Georg-Speyer-Haus, Frankfurt/ Main; Juni 2002

Danksagung

Ich danke Allen ganz herzlich, die zum Gelingen dieser Arbeit beigetragen haben:

Meinem Doktorvater Prof. Dr. Rolf Marschalek für die Überlassung des interessanten und abwechslungsreichen Themas und für seine intensive Unterstützung und ständige Diskussionsbereitschaft.

Prof. Dr. Theo Dingermann für die Begutachtung dieser Arbeit und die vielen Anregungen bei den Arbeitsseminaren.

Dr. Ludger Klein-Hitpass, Universitätsklinikum Essen, für die Durchführung der DNA-Microchip Hybridisierungen und die ausserordentliche Geduld!

Dr. Michael Walther und Martin Hampel für die tatkräftige Unterstützung bei der Durchführung der 2D-Gelelektrophorese.

Dr. Ilse Zündorf für die Herstellung der monoklonalen Antikörper und Jürgen Zech für deren Aufreinigung und Charakterisierung.

Tanja Lehmann für die wertvollen Beiträge zu der Publikation.

Unseren „Super-TA’s“: Martina Berg, Melanie Seidel, Christin Uttner und Chrissi Tschepke, die mir durch ihre Unterstützung die Arbeit und den Alltag im Labor erleichtert haben.

Meinen Kollegen für Tipps, Tricks, den Spass im Labor und die vielen netten Kaffeepausen: Sven Moritz, Jürgen Zech, Anne Gaussmann, Claus Meyer, Heidi Bursen, Sebastian Scharf, Peter Beck, Monika Baik, Ingo Jennes, Oliver Siol, Christian Klein, Jan Keller und den „Ehemaligen“: Susann Henke, Sascha Hövelmann, Uli Hentschel, Thorsten Kohl, Karola Dannat und Sabine Smietana. Vielen Dank auch an HD Dr. Thomas Winckler und Dr. Ilse Zündorf, die immer für praktische Hilfen zur Verfügung standen.

Besonders danke ich Anke Rohmann, Anke Friese, Christin Utner, Kirstin Berger, Kristina Ruf und Martina Berg für die schöne Zeit, die gemeinsamen Stunden bei gutem Essen und einem Gläschen Wein und ganz besonders für die Freundschaft!

Mein herzlichster Dank gilt meiner Familie und Dirk, die mir viel Geduld und Verständnis während meiner Arbeit entgegengebracht haben, mich das ganze Studium und darüber hinaus unterstützten und immer für mich da waren. Danke für die vielen aufmunternden Worte!