

Institut für Entwicklungsgenetik der Technischen Universität München

# **Zelltypspezifische Regulation humaner endogener Retroviren**

**Stephan Weinhardt**

Vollständiger Abdruck der von der Fakultät Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt der Technischen Universität München zur Erlangung des akademischen Grades eines

**Doktors der Naturwissenschaften**

genehmigten Dissertation

Vorsitzender: Univ.-Prof. Dr. Siegfried Scherer

Prüfer der Dissertation:

1. Univ.-Prof. Dr. Wolfgang Wurst
2. apl. Prof. Dr. Christine Leib-Mösch,  
Ruprecht-Karls-Universität Heidelberg

Die Dissertation wurde am 14.12.2005 bei der Technischen Universität München eingereicht und durch die Fakultät Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt am 25.01.2006 angenommen.

<b>1</b>	<b>Einleitung</b>	<b>6</b>
1.1	Endogene Retroviren als Retroelemente	6
1.2	Retrovirale Genomorganisation	7
1.3	Der retrovirale Lebenszyklus	10
1.4	Retrovirale Taxonomie	12
1.5	Humane Endogene Retroviren (HERVs)	12
1.6	Einfluss von HERVs auf zelluläre Gene	18
1.7	Struktur und Aktivität der HERV-LTRs	22
1.7.1	HERV-R	25
1.7.2	HERV-E	25
1.7.3	HERV-I	26
1.7.4	HERV-H	26
1.7.5	ERV9	28
1.7.6	HERV-W	29
1.7.7	HERV-K (HML-1 bis 10)	30
1.8	HERV-L	32
<b>2</b>	<b>Material</b>	<b>37</b>
2.1	Puffer und Lösungen	37
2.2	Oligonukleotide	40
2.2.1	Primer für „Site Directed Mutagenesis“	40
2.2.2	Oligonukleotide für „Electrophoretic Mobility Shift Assay“ (EMSA)	41
2.2.3	PCR-Primer	42
2.3	Reagenzien	43
2.4	Sonstige Labor Materialien	44
2.5	Gebrauchsfertige Reaktionssysteme (Kits)	44
2.6	Medien	45

2.6.1	Bakterienmedien	45
2.6.2	Zellkulturmedium	46
2.6.3	Medium zum Einfrieren von Zellen	46
2.7	Antikörper	46
2.8	Plasmide	47
2.9	Zelllinien	47
2.10	Bakterienstämme	48
2.11	Geräte	49
2.12	Verwendete Software und Datenbanken	50
2.13	HERVd-Datenbank: Accession-Nummern	51
<b>3</b>	<b>Methoden</b>	<b>52</b>
3.1	Kultivierung adhärenter Zellen	52
3.2	Zellzahlbestimmung	52
3.3	Auftauen von Zellen	52
3.4	Einfrieren von Zellen	53
3.5	Transfektion eukaryotischer Zellen mit Plasmid-DNA	53
3.5.1	Kalziumphosphat-Kopräzipitation	53
3.5.2	Transfektion mit „Effectene“	54
3.5.3	Transfektion mit „Superfect“	54
3.6	DNA-Isolierung aus Geweben	55
3.7	Isolierung von Plasmid-DNA aus Bakterien	56
3.7.1	Plasmidmaxipräparationen (Maxipräp)	56
3.7.2	Plasmidminipräparation (Minipräp)	57
3.8	Standard-PCR	57
3.9	Restriktionsverdau	59
3.10	Agarosegelelektrophorese	60

3.11	DNA-Isolierung aus Agarosegelen	60
3.12	Ligation	61
3.12.1	TOPO TA Klonierung	61
3.12.2	Transformation von <i>E. coli</i> s mit Plasmid-DNA	62
3.13	„Site directed mutagenesis“	63
3.13.1	Synthese der mutierten Plasmide	65
3.13.2	Transformation der XLII Blue superkompetenten Zellen	65
3.14	Sequenzierung	66
3.15	RNA-Extraktion	66
3.16	cDNA-Synthese	67
3.17	Proteinkonzentrationsbestimmung mit Bicinchioninsäure	69
3.18	„Dual Luziferase® Reporter Assay“	69
3.19	Kernextraktpräparation	71
3.20	„Electrophoretic Mobility Shift Assay“ (EMSA)	71
3.20.1	Hybridisierung der Einzelstrang Oligonukleotide	73
3.20.2	Radioaktive Markierung der Doppelstrang-Oligonukleotide (Blunt End-Markierung)	73
3.20.3	Reaktionsansätze und Polyacrylamidgelelektrophorese	74
3.21	Computeranalysen	75
3.21.1	Matinspector	75
3.21.2	Sequence Shaper	76
3.21.3	Frameworker	76
<b>4</b>	<b>Ergebnisse</b>	<b>78</b>
4.1	Isolierung und Charakterisierung von HERV-L-LTRs	78
4.1.1	Phylogenetische Analyse der HERV-L-LTRs	78
4.1.2	Klonierung der HERV-L-LTRs in den pBL-Expressionsvektor	84

4.2	Promotoraktivität von HERV-L-LTR-Sequenzen in humanen Zellen _____	85
4.2.1	Optimierung der Transfektion von HaCaT-Zellen _____	85
4.2.2	Vergleich der Promotoraktivität von HERV-L-LTRs in verschiedenen Zelllinien _____	88
4.2.3	Promotoraktivität der 5'- und der 3'-HERV-L-LTR(Cord) _____	94
4.2.4	Promotoraktivität von HERV-LTRs in transgenen Mäusen _____	95
4.2.5	Untersuchung HERV-H-H6-EGFP-transgener Mäuse _____	96
4.2.6	Untersuchung 5'-HERV-L-LTR(Cord)-Luziferase-transgener Mäuse _____	104
4.3	Identifizierung potentieller Transkriptionsfaktorbindungsstellen durch Computeranalyse _____	106
4.3.1	Sequenzvergleich solitärer und proviraler HERV-L-LTRs _____	106
4.3.2	Identifizierung potentieller Transkriptionsfaktorbindungsstellen mit Hilfe des Programms „Matinspector“ _____	109
4.3.3	Promotor Module _____	110
4.3.4	Identifizierung von TF-Bindungsstellen durch Sequenzvergleich der 5'-HERV-L-LTR(Cord) mit der inaktiven 3'-HERV-L-LTR(Cord) _____	112
4.4	Überprüfung der Funktionalität der identifizierten TF-Bindungsstellen durch „Site directed Mutagenesis“ _____	116
4.5	Identifizierung funktioneller Bindungsstellen mittels „Electrophoretic Mobility Shift Assay“ (EMSA) _____	120
<b>5</b>	<b>Diskussion</b> _____	<b>141</b>
5.1	HERVs und Gewebespezifität _____	141
5.2	5'-HERV-L-LTR(Cord): Zelltypspezifische Promotoraktivität _____	144
5.3	Lässt sich gewebespezifische Expression der 5'-HERV-L-LTR(Cord) und der 5'-HERV-H-H6-LTR in transgenen Mäusen bestätigen? _____	145
5.4	Untersuchung des Promotors der 5'-HERV-L-LTR(Cord) _____	147
5.5	Ausblick _____	156

<b>6</b>	<b>Zusammenfassung</b>	<b>157</b>
<b>7</b>	<b>Literaturverzeichnis</b>	<b>161</b>
<b>8</b>	<b>Abkürzungen</b>	<b>182</b>

### 1 Einleitung

Das menschliche Genom besteht zu ca. 9% aus retroviralen Elementen (Lander *et al.*, 2001; Venter *et al.*, 2001; Li *et al.*, 2001), das entspricht in etwa dem Fünffachen der Protein-codierenden Sequenzen. Der hohe Anteil der viralen DNA geht auf Infektionen von Keimbahnzellen vor 40 bis 70 Millionen Jahren durch exogene Retroviren zurück (Greenwood *et al.*, 2005). Seitdem konnten sie sich im Genom durch Reinfektion und Retrotransposition vermehren (Belshaw *et al.*, 2004; Belshaw *et al.*, 2005). Ein Teil dieser endogenen Retroviren wurde im Laufe der Zeit in die Regulation zellulärer Vorgänge der Wirtszelle integriert und hat damit ganz wesentlich zur Evolution und genetischen Vielfalt der Primaten beigetragen (van de Lagemaat *et al.* 2003; Review: Leib-Mösch *et al.*, 2005). Diese regulatorischen Eigenschaften retroviraler Elemente sind Gegenstand meiner Doktorarbeit.

#### 1.1 Endogene Retroviren als Retroelemente

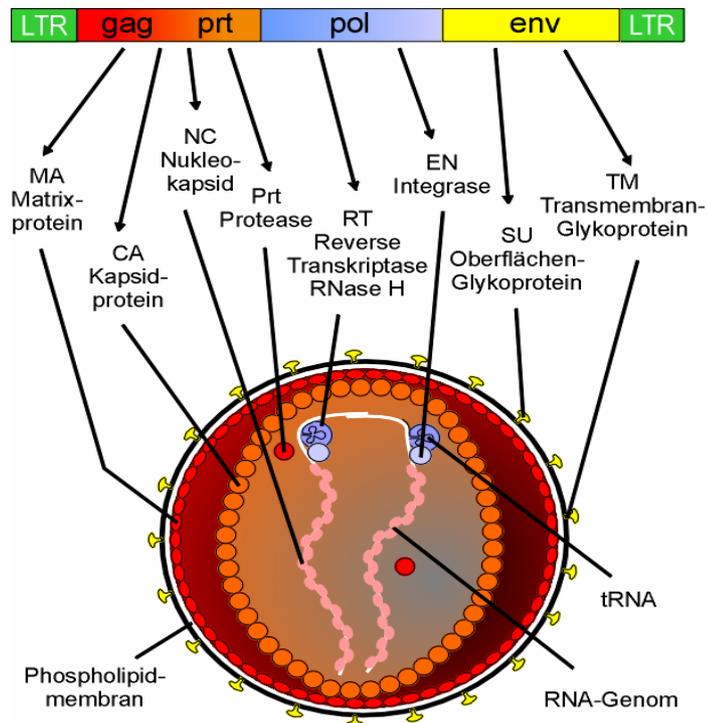
Endogene Retroviren werden zusammen mit LINE-Elementen (Eickbush 1992; Singer *et al.*, 1993) und SINE-Elementen (Deininger, 1989) in die Gruppe der Retroelemente eingeordnet. Retroelemente sind in allen Eukaryoten vorhanden und konnten auch in vielen Prokaryoten nachgewiesen werden. Charakteristisch für alle diese Elemente ist die Eigenschaft des Umschreibens von RNA in DNA durch die reverse Transkriptase (RT). Allerdings codieren nicht alle Retroelemente ihre eigene RT, z.B. nutzen nicht autonome Retroelemente wie SINEs die reversen Transkriptasen anderer Retroelemente.

Die reverse Transkriptase hat vermutlich schon in der frühen RNA-Welt eine Rolle gespielt und den Wechsel vom Speichermedium RNA zu DNA mitbewirkt (Boeke *et al.*, 2003). Über den Ursprung der reversen Transkriptase gibt es zwei viel diskutierte Theorien. Die erste geht davon aus, dass alle bekannten RTs auf ein ursprüngliches

Enzym zurückgehen (Temin, 1980). Die zweite Theorie bezieht sich auf die große Sequenzdiversität reverser Transkriptasen und postuliert mehrfache RT-Entstehung im Laufe der Evolution (Boeke und Stoye, 1997).

### **1.2 Retrovirale Genomorganisation**

Infektiöse retrovirale Partikel besitzen zwei einzelsträngige RNA-Moleküle mit einer Länge von 7-10 kb (Abb. 1). Das RNA-Genom besteht aus den internen Regionen *gag*, *pol* und *env*, die die Protein-codierenden Bereiche enthalten. Die *gag*-Region codiert für mindestens drei virale Strukturproteine, die auch als gruppenspezifische Antigene bezeichnet werden, da sie für jede retrovirale Gruppe spezifisch sind. Die *pol*-Region codiert für eine reverse Transkriptase (RT), die RNA als Matrize verwendet und den DNA-Einzelstrang (cDNA) synthetisiert. Die Reverse Transkriptase besitzt außerdem eine RNaseH-Funktion, die den Abbau des RNA-Stranges in RNA-DNA-Hybriden bewirkt (Review: Kupiec & Sonigo, 1996). Die Integrase (IN) besitzt die Fähigkeit zur DNA-Spaltung und bewerkstelligt die Integration der viralen DNA in das Wirtsgenom. Die Protease (PR) spaltet das Gag und Gag-Pol-Vorläuferprotein (Review: Oroszlan & Luftig, 1990). Die *env*-Region codiert für zwei Glykoproteine, sie dienen der Bindung des Virus an bestimmte Rezeptoren der Wirtszelle, wodurch die Aufnahme des Virus erfolgen kann.

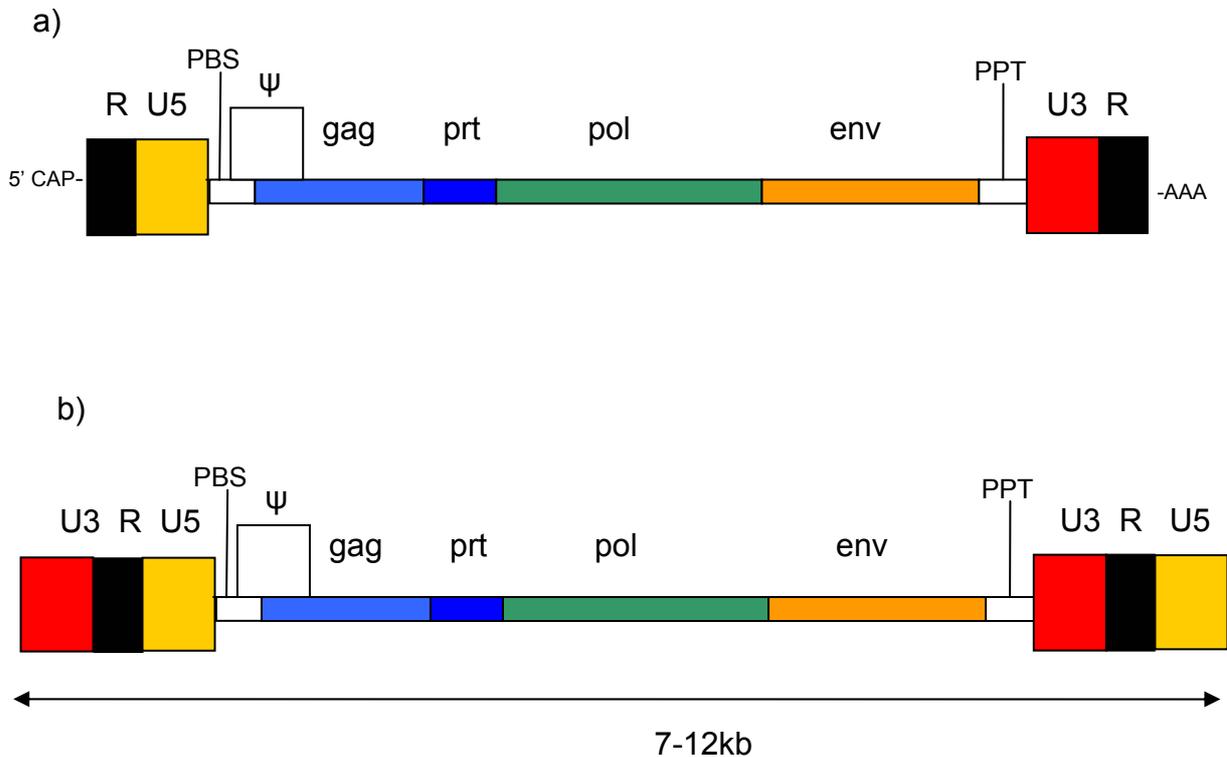


**Abb. 1:** Schematischer Aufbau eines integrierten Provirus und eines retroviralen Partikels. Das provirale Genom wird an beiden Enden von LTRs (Long Terminal Repeats) begrenzt. Das *gag*-Gen codiert für das Matrixprotein (MA), das Kapsidprotein (CA) und das Nukleokapsidprotein (NC). Das *prt*-Gen codiert für die Protease PRT und das *pol*-Gen für die Reverse Transkriptase RT, die RNase H (EN) sowie die Integrase (EN). Das *env*-Gen codiert für das Oberflächenglykoprotein (SU) und das Transmembranglykoprotein (TM).

An den Enden des proviralen RNA-Genoms befinden sich die terminalen "Repeat"-Regionen, die am 5'-Ende mit einer „Cap“-Struktur versehen und am 3'-Ende polyadenyliert sind (Abb. 2a). Die „Cap“-Struktur und die Polyadenylierung sind erforderlich, da das virale RNA-Genom auch als mRNA fungiert und von der Wirtszelle in die viralen Proteine translatiert wird. Die 10 bis 250 Nukleotide lange R-Region ist an beiden Enden des RNA-Genoms vorhanden. In der 5'-LTR schließt sich an die R-Region die U5-Region mit einer Länge von 35 bis 450 Nukleotiden an, gefolgt von einer „Leader“-Region, die die Primer-Bindungsstelle (PBS) enthält. Diese spielt eine Rolle bei der Einleitung der reversen Transkription, da sie spezifisch zelluläre tRNA bindet, die als Primer für die RT fungiert. Die tRNAs werden in vielen Fällen zur taxonomischen Klassifizierung der HERV-Familien herangezogen, wobei

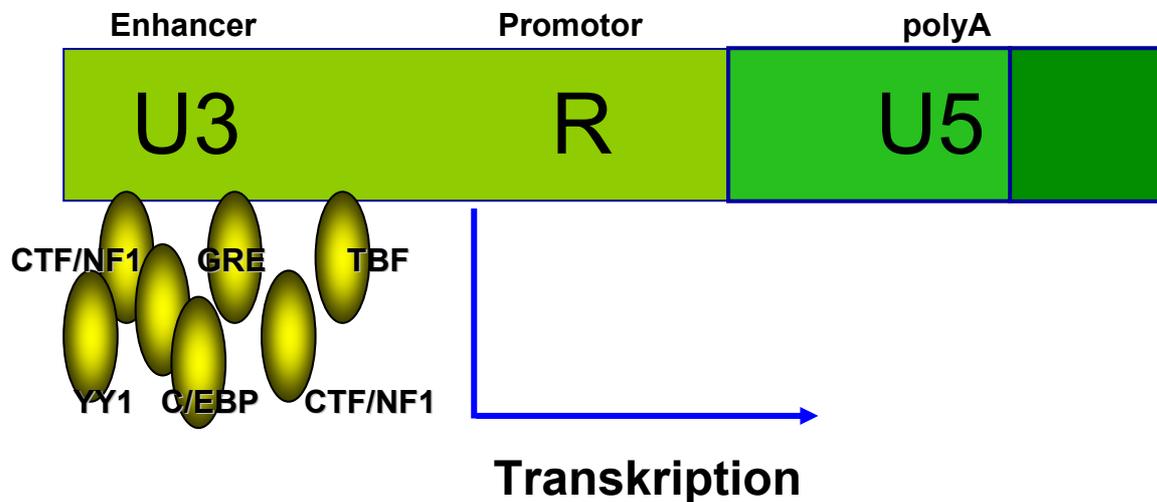
## 1. Einleitung

der „single letter aminoacid code“ für die von der tRNA getragene Aminosäure namensgebend ist. Das 3'-Ende des RNA-Genoms beinhaltet für die Synthese des zweiten cDNA-Stranges eine weitere PBS, den Polypurintrakt, die U3- und R-Region.



**Abb. 2:** a) Schematische Darstellung der genomischen RNA endogener Retroviren. Es handelt sich um ein einzelsträngiges Molekül bestehend aus: 5'-CAP; „Redundant“-Region (R); U5-Region; Primer Bindung Site (PBS); Packaging Sequence ( $\psi$ ); Polypurintrakt (PPT); Polyadenylierungssignal (PA); U3-Region; Poly(A)Tail (AAA). b) Schematische Darstellung der doppelsträngigen proviralen DNA die in das Genom der Wirtszelle integriert wird. Es sind sowohl am 5'-, als auch am 3'-Ende des Elementes „Long Terminal Repeats“ (LTRs) vorhanden. U3-Region; „Redundant“-Region (R); U5-Region; Primer Bindung Site (PBS); Packaging Sequence ( $\psi$ ); Polypurintrakt (PPT).

Der Aufbau der proviralen DNA unterscheidet sich vom RNA-Genom durch das Vorhandensein von „Long Terminal Repeats“ (LTRs), die die provirale Sequenz sowohl am 5'- als auch am 3'-Ende begrenzen (Abb. 2b). Diese Elemente bestehen aus der U3-, R- und U5-Region und entstehen während der reversen Transkription (Abb. 3).



**Abb. 3:** Schematische Darstellung der retroviralen LTR eines DNA Provirus mit der U3-, R- und U5-Region. In der U3-R-Region befinden sich die Transkriptionsfaktorbindungsstellen. Der Transkriptionsstartpunkt und die Polyadenylierungsstelle befinden sich in der U5-Region.

### 1.3 Der retrovirale Lebenszyklus

Um in die Wirtszelle eindringen zu können heften sich retrovirale Virionen an spezifische Rezeptoren der Wirtszelle (Varmus und Coffin, 1982; Varmus und Brown, 1989). Durch Fusion von Virushülle und Wirtszellmembran gelangt die virale RNA ins Cytoplasma, wo die Virushülle abgebaut wird („uncoating“). Anschließend erfolgt das Umschreiben der retroviralen RNA in Plus- und Minus-Strang-DNA durch die reverse Transkriptase (Coffin, Retroviruses, 1996). Gleichzeitig wird durch die RNaseH-Funktion der RNA-Strang abgebaut. Die Initiation der reversen Transkription beginnt an einer aus der Wirtszelle stammenden tRNA, die im Viruscapsid vorhanden ist und an die Primerbindungsstelle (PBS) der viralen RNA bindet. Die doppelsträngige virale DNA wird dann durch die virale Integrase ins Wirtsgenom eingebaut und von nun an als Provirus bezeichnet. Die in der 5'-LTR vorhandenen regulatorischen Sequenzen steuern zusammen mit zellulären Transkriptionsfaktoren die Transkription der viralen

Gene. Die synthetisierte RNA wird zur Bildung von neuem, viralem Genom, bzw. zum Teil gespleißt, als Matrize für die Translation viraler Gene verwendet. Die genomische RNA wird danach zusammen mit den glykosylierten und phosphorylierten viralen Proteinen zum Nukleocapsid zusammengebaut. Bei der anschließenden Partikelbildung, dem sogenannte „Virusbudding“, verwenden die entstehenden Virionen die Lipidmembran des Wirtes und schnüren sich von der Wirtszelle ab. Sie durchlaufen eine Reifungsphase, nach der sie vollständig infektiös sind. Retroviren infizieren normalerweise somatische Zellen, können aber in seltenen Fällen auch Keimbahnzellen infizieren, was zu einer vertikalen Übertragung auf die nachfolgenden Wirtsgenerationen führt. Im Genom von Keimbahnzellen integrierte Proviren werden als endogene Retroviren (ERVs) bezeichnet (Coffin, *Retroviruses*, 1996). Eine Integration viraler DNA ins Wirtsgenom kann in vielen Fällen negative Auswirkungen haben, vor allem, wenn sie in, oder in der Nähe von Genen erfolgt und diese dadurch zerstört, bzw. an- oder abschaltet. Daher ist es wahrscheinlich, dass viele dieser Keimbahninfektionen durch den Selektionsdruck nicht lange im Wirtsgenom der nachfolgenden Population verbleiben. Ist der Einfluss des Provirus nicht negativ, bleibt die virale DNA im Wirtsgenom erhalten. Das Virus kann sich dann durch Retrotransposition oder Reinfektion weiter im Wirtsgenom verbreiten, (Boeke und Stoye, 1997; Belshaw, 2004) und somit eine monophyletische ERV-Linie bilden. Jede Neuinfektion einer Keimbahnzelle bildet die Grundlage für eine neue unabhängige ERV-Linie (Tristem, 2000), wobei die Amplifikationsrate kurz nach der Kolonisation am höchsten ist und mit der Zeit abnimmt. Gründe hierfür sind bei negativer Auswirkung auf den Wirt u.a. negative Selektion durch Mechanismen, die eine Amplifikation oder Reinfektion verhindern, wie z.B. Methylierung (Yoder *et al.*, 1997; Waugh *et al.*, 1998) oder Verlust von Rezeptoren des Wirtes, die für die Virusaufnahme benötigt werden (Boeke und Stoye, 1997). In Primaten können die Resistenzfaktoren APOBEC3G und TRIM5 $\alpha$  retrovirale Gene durch Hypermutation inaktivieren (Sawyer *et al.*, 2004; Sawyer *et al.*, 2005; Esnault *et al.*, 2005). Außerdem akkumulieren im Laufe der Zeit Mutationen in codierenden oder regulatorischen Sequenzen, die in den meisten Fällen schließlich zur Inaktivierung der Linie führen. Dennoch sind einzelne ERVs auch nach mehreren Millionen Jahren zur Expression oder sogar zur Replikation in der Lage (Stoye, 2001; Barbulescu *et al.*, 2001; Reus *et al.*, 2001). Eine Inaktivierung

von ERVs kann durch Komplementation in *trans* aufgehoben werden. Das heißt, dass ERV-Sequenzen intakte regulatorische Bereiche wie die LTR-Regionen, die „Primer Binding Site“ (PBS) oder das Verpackungssignal besitzen, die zur Replikation oder zur Verpackung benötigten Proteine jedoch von anderen ERVs, oder auch von exogenen Retroviren geliefert werden. Ihre volle transkriptionelle Aktivität können ERVs durch Rekombination wiedererlangen. Ein Beispiel hierfür ist das endogene Gammaretrovirus BaEV, das sich aus dem *Papio cynocephalus* endogenen Retrovirus (PcEV) und dem Simian (Affen) endogenen Retrovirus (SERV) zusammensetzt (Van der Kuyl *et al.*, 1997; Mang und Van der Kuyl, 1999).

### 1.4 Retrovirale Taxonomie

Innerhalb der Familie der Retroviren unterscheidet man die folgenden sieben Gattungen: Alpharetrovirus, Betaretrovirus, Deltaretrovirus, Gammaretrovirus, Epsilonretrovirus, Spumavirus und Lentivirus (Van Regenmortel *et al.*, 2000). Diese Taxonomie basiert auf Genomsequenzdaten die größtenteils aus der relativ konservierten reversen Transkriptase Sequenz (RT) stammen (Xiong & Eickbush 1990; Doolittle *et al.*, 1989), oder aber aus dem Vergleich von LTR-Sequenzen. Das Verwenden anderer Sequenzen, wie z.B. der codierenden Region von RNaseH oder der Integrase (IN), liefert in den meisten Fällen konsistente Ergebnisse (McClure *et al.*, 1988).

Im Falle der humanen endogenen Retroviren (HERVs) unterscheidet man 26 verschiedene Linien, die alle aus separaten Kolonisationsereignissen stammen (Tristem, 2000; Benit *et al.*, 2001) und nach der HERV-Nomenklatur jeweils als Familien klassifiziert werden.

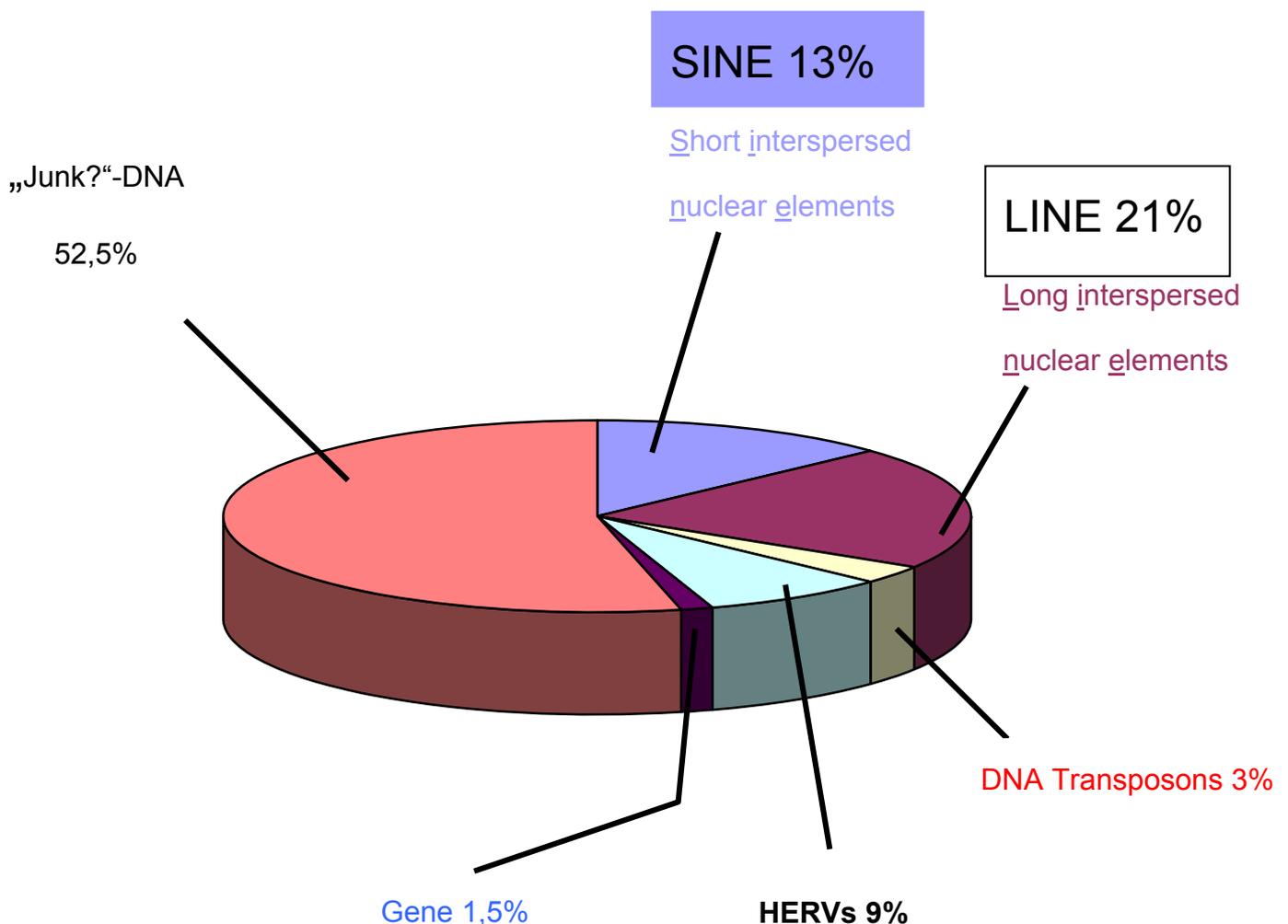
### 1.5 Humane Endogene Retroviren (HERVs)

Das menschliche Genom besteht zu ca. 9% aus HERV-Sequenzen (Lander *et al.*, 2001; Venter *et al.*, 2001; Li *et al.*, 2001; Abb. 4), die mindestens 26 phylogenetisch

## 1. Einleitung

---

unterscheidbaren Linien zuzuordnen sind (Tristem, 2000; Benit *et al.*, 2001). Der Anteil der „Short Interspersed Elements“ (SINES) liegt bei 13%, die „Long Interspersed Elements“ haben einen Anteil von 21%, DNA-Transposons von 3%. Die codierenden Bereiche der Gene (Exons) haben einen Anteil von lediglich 1,5% des menschlichen Genoms. Der überwiegende Teil der DNA, nämlich 53,5%, besteht aus sogenannter „Junk“-DNA, welcher bisher keine Funktion zugeordnet werden konnte (Abb. 4; Li *et al.*, 2001)



**Abb. 4:** Zusammensetzung des menschlichen Genoms (Li *et al.*, 2001). Der Anteil proteincodierender Sequenzen liegt lediglich bei 1,5%. „Long Interspersed Nuclear Elements“ (LINEs) stellen 21%, „Short Interspersed Elements“ (SINEs) 13%, HERV-Elemente 9% und DNA-Transposons 3%. 52,5% bildet DNA unbekannter Funktion.

In der Regel werden HERVs nach der tRNA benannt, die an ihre Primerbindungsstelle (PBS) bindet, so z.B. die HERV-L Familie, die nach tRNA-Leu benannt ist. Die

## 1. Einleitung

---

HERV-Familien werden aufgrund ihrer phylogenetischen Verwandtschaft in Klassen zusammengefasst (siehe Tabelle 1), wobei HERVs, die Verwandtschaft zu Gammaretroviren und Epsilonretroviren zeigen, zur Klasse I gerechnet werden. Solche mit Verwandtschaft zu Betaretroviren werden der Klasse II und Spumaviren verwandte HERVs der Klasse III zugeordnet (Cordonnier *et al.*, 1995; Wilkinson *et al.*, 1994).

**Tabelle 1ff:** Klassifizierung humaner endogener Retroviren (HERV). Zusätzlich sind in der Literatur verwendete alternative Familiennamen und RT-Primer angegeben.

HERV Familie	Alternativer Name	Primer
<b>Klasse I</b>		
HERV-Z69907		nicht detektiert
HERV-ADP	ADP-pol	tRNA <sup>Thr</sup>
HERV-E		tRNA <sup>Glu</sup>
HERV-F		tRNA <sup>Phe</sup>
HERV-F (Typ b)		tRNA <sup>Phe</sup>
HERV-FRD		tRNA <sup>His</sup>
HERV-H	RTLHV-H	tRNA <sup>His</sup>
HERV-H49C23		keine LTRs
HERV-I	RTLHV-I, HERV-IP	tRNA <sup>Ile</sup> , tRNA <sup>Pro</sup>
RRHERV-I		tRNA <sup>Ile</sup>

## 1. Einleitung

---

<b>HERV Familie</b>	<b>Alternativer Name</b>	<b>Primer</b>
ERV-9		tRNA <sup>Arg</sup>
HERV-F (Typ c)		tRNA <sup>Phe</sup>
HERV-P	HuRRS-P	tRNA <sup>Pro</sup>
HERV-R	ERV.R	tRNA <sup>Arg</sup>
HERV-R (Typ b)		tRNA <sup>Arg</sup>
HERV-T	HERV.S71	tRNA <sup>Thr</sup>
HERV-W	MSRV	tRNA <sup>Trp</sup> ; tRNA <sup>Arg</sup>
HERV-XA		tRNA <sup>Phe</sup>
<b>Klasse II</b>		
HERV-K(HML1)	NMWW6	tRNA <sup>Lys</sup>
HERV-K(HML2)	HERV-K10, HTDV, NMWW1	tRNA <sup>Lys</sup>
HERV-K(HML3)	HERV-K70A, HERV76; HERV50, NMWW5	tRNA <sup>Lys</sup> ; tRNA <sup>Tyr</sup>
HERV-K(HML4)	ERV-MLN, HERV-K(T47D)	tRNA <sup>Lys</sup>
HERV-K(HML5)	NMWW2	tRNA <sup>Ile</sup>
HERV-K(HML6)	NMWW4	tRNA <sup>Lys</sup>
HERV-K(HML7)	NMWW7	tRNA <sup>Lys</sup>
HERV-K(HML8)	NMWW3	nicht detektiert
HERV-K-(HML9)	NMWW9	tRNA <sup>Trp</sup>

## 1. Einleitung

---

HERV-K(C4)	HML-10	tRNA <sup>Lys</sup>
------------	--------	---------------------

Klasse III	Alternativer Name	Primer
HERV-L		tRNA <sup>Leu</sup>
HERV-S		tRNA <sup>Ser</sup>
HERV-U2		nicht detektiert
HERV-U3		nicht detektiert

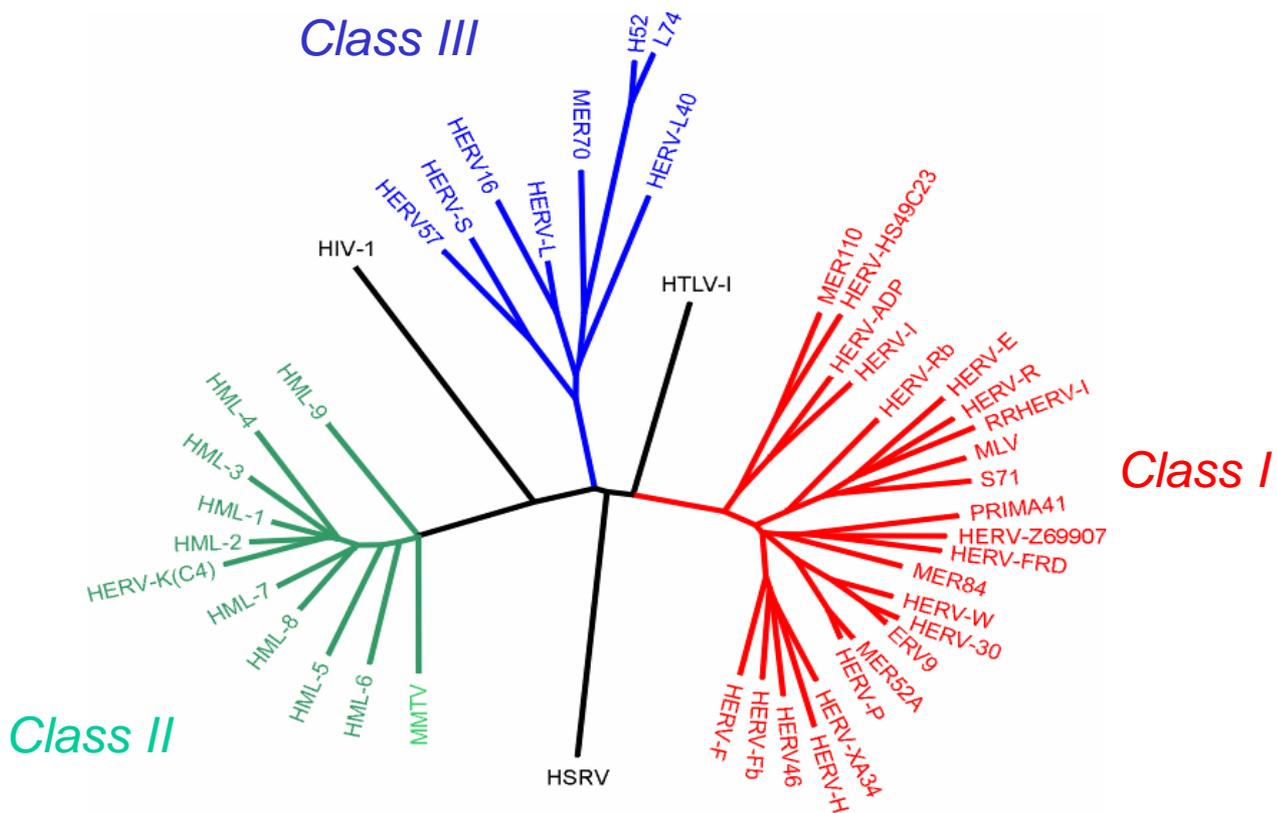
Das Alter der HERVs zeigt sich am Ausmaß der Degenerierung der „Open Reading Frames“ (ORFs). Diese enthalten in der Regel eine Vielzahl an Punktmutationen und Deletionen, wodurch zahlreiche Stopp-Codons und „Frameshifts“ entstehen (Wilkinson *et al.*, 1994; Boeke & Stoye, 1997). Bei einigen ERVs fehlt der komplette *env*-Bereich, was ihre Verbreitung auf Retrotransposition beschränkt (Cordonnier *et al.*, 1995; Mager & Henthorn, 1984; Hirose *et al.*, 1993). Andere haben durch homologe Rekombination ihrer beiden LTRs den kompletten codierenden Bereich verloren. Diese solitären LTRs sind, je nach HERV-Familie, etwa 10 bis 100 mal häufiger als die kompletten Proviren im menschlichen Genom vorhanden (Boeke & Stoye, 1997; Stoye, 2001; Bock & Stoye, 2000).

Das Vorkommen aller HERV-Familien in Altweltaffen und in einigen Fällen auch in Neuweltaffen lässt auf eine Integration ins Genom vor ca. 40 bis 70 Millionen Jahren schließen (Greenwood, 2005; Tristem, 2000; Reus *et al.*, 2001; Andersson *et al.*, 1997; Mager & Freeman, 1995; Shih *et al.*, 1991; Kim *et al.*, 1999; Mayer *et al.*, 1998).

Die Klasse I HERVs umfassen etwa 113000 Elemente und stellen die größte Anzahl an HERV-Familien (etwa 20 Familien). Klasse II HERVs stellen mit etwa 8000 Kopien die kleinste Gruppe der endogenen Retroviren im menschlichen Genom dar (Lander *et al.*, 2001). Sie werden auch als HERV-K-Superfamilie bezeichnet (Boeke & Stoye, 1997; Wilkinson *et al.*, 1994), da die meisten Proviren durch eine PBS für

## 1. Einleitung

tRNA-Lys gekennzeichnet sind. Zur Klasse III, den Spumaretrovirus-verwandten HERVs, die ca. 89000 Elemente umfassen, werden zur Zeit acht Familien gezählt (Medstrand und Mager, 2003). Die mit Abstand am zahlreichsten vertretene Familie dieser Klasse ist HERV-L, die wahrscheinlich eine Anzahl von 200-500 Exemplaren umfasst und damit nach HERV-H die zweitgrößte Familie der Humanen Endogenen Retroviren stellt (Abb. 5).



**Abb. 5:** Verwandtschaft Exogener und Endogener Retroviren basierend auf den DNA-Sequenzen der *pol*-Region (Mager und Medstrand, 2003). Zu sehen sind die Klassen I, II, III der endogenen Retroviren und die exogenen Retroviren HIV-1 (Humanes Immunodefizienzvirus), HTLV (Humanes T-Zell Leukämie Virus) und HSRV (Humanes Spumavirus).

### 1.6 Einfluss von HERVs auf zelluläre Gene

Humane endogene Retroviren können auf wenigstens drei verschiedenen Wegen Einfluss auf zelluläre Gene nehmen (Reviews: Leib-Mösch, 2004; Landry, 2003; Jordan, 2003):

1. HERV-LTRs können als Promotoren für zelluläre Gene fungieren. Einige Transkripte initiieren innerhalb einer 5'-LTR und spleißen von einer Donorstelle, die „upstream“ der *gag*-Region lokalisiert ist, in „downstream“ Bereiche, die zelluläre Gene codieren. Ein Beispiel hierfür ist das Opitzsyndrome-Gen *MID1*, das außer von seinem zellulären Promotor auch von der 5'-LTR eines HERV-E-Elements reguliert wird (Landry *et al.*, 2002; Landry und Mager 2002). Transkripte, in denen die UTRs („5'-untranslated Regions“) durch 5'-HERV-E-LTRs ersetzt wurden, konnten vor allem in Plazenta und embryonalen Nieren detektiert werden. Eine weitere Variante ist die Transkription ausgehend von einem Promotor einer 3'-, oder einer solitären LTR, direkt in das zelluläre Gen. In die Promotorsequenz des Gens *INSL4* (Insulin-Familie) ist ein HERV-K-Element integriert, das eine *env*-ähnliche Sequenz enthält und von zwei LTRs flankiert wird. Die plazentaspezifische Expression von *INSL4* wird durch die 3'-LTR des HERV-Elements reguliert und ist vermutlich an der Differenzierung des Zytotrophoblasten zum Syncytiotrophoblasten maßgeblich beteiligt (Bieche *et al.*, 2003).
2. HERV-LTRs können Enhancer-Funktion besitzen. Diese Variante ist vermutlich häufiger als die Promotorfunktion, da Enhanceraktivität unabhängig von Orientierung und Distanz zum zellulären Gen ist. Allerdings sind HERV-LTRs mit Enhancer-Funktion schwer zu detektieren, da sie nicht Teil des Transkriptes sind. Ein Beispiel für die Enhancer-Funktion einer HERV-LTR ist ein HERV-E-Element, das „upstream“ des Amylase-Gen-Clusters (*amy1A*, *1B*, *1C*, *amy2A*) lokalisiert ist und an dessen Expression in humanen Speicheldrüsen beteiligt ist. (Emi *et al.*, 1988; Samuelson *et al.*, 1988; Ting *et al.*, 1992; Meisler und Ting, 1993; Samuelson *et al.*, 1990; Samuelson *et al.*, 1996; Schulte *et al.*, 1996).

## 1. Einleitung

---

3. HERV-LTRs können durch ihre Polyadenylierungsstelle zelluläre Transkripte terminieren. Beispiele dafür sind die Gene *HHLA2* und *HHLA3*, deren Funktion bisher unbekannt ist. Beide werden vor allem im Darm, in den Nieren und der Lunge exprimiert und innerhalb einer HERV-H-LTR polyadenyliert (Mager *et al.*, 1999).

Eine Zusammenstellung von bisher identifizierten HERV-Elementen, die auf zelluläre Gene Einfluss nehmen, ist in Tabelle 2 dargestellt.

**Tabelle 2ff:** HERV-Elemente und ihr Einfluss auf zelluläre Gene. Die Tabelle zeigt die jeweilige HERV-LTR, das durch sie beeinflusste Gen, die Funktion der LTR und die Literaturquelle (Leib-Mösch *et al.*, 2004).

HERV-LTR	beeinflusstes Gen	Funktion	Referenz
HERV-E	<i>APOC1</i> (Apolipoprotein CI)	Promotor	Medstrand <i>et al.</i> , 2001
HERV-E	<i>MID1</i> (Opitz's Syndrome Gen)	Promotor	Landry <i>et al.</i> , 2002; Landry und Mager, 2002
HERV-E	<i>EBR</i> (Endothelin B Rezeptor)	Promotor	Medstrand <i>et al.</i> , 2001
HERV-E	<i>amy1A, 1B, 1C, amy2A</i> (Salivary Amylase)	Enhancer	Samuelson <i>et al.</i> , 1988; Meisler und Ting, 1993; Emi <i>et al.</i> , 1988; Samuelson <i>et al.</i> , 1990, 1996; Ting <i>et al.</i> , 1992
HERV-E	<i>PTN</i> (Pleiotrophin)	5'-UTR, Enhancer?	Schulte <i>et al.</i> , 1996; Schulte und Wellstein, 1998; Landry <i>et al.</i> , 2002
HERV-E	<i>PRDM7</i> (Zink Finger Protein)	Polyadenylierung	Landry <i>et al.</i> , 2002

## 1. Einleitung

---

<b>HERV-LTR</b>	<b>beeinflusstes Gen</b>	<b>Funktion</b>	<b>Referenz</b>
HERV-H	<i>CDC4L</i> (Cell Division Cycle Protein, $\beta$ -Transducin)	Promotor	Feuchter <i>et al.</i> , 1992
HERV-H	Calbindin (Calcium Bindeprotein)	Promotor	Liu und Abraham <i>et al.</i> , 1991
HERV-H	<i>PLA2L</i> (Phospholipase A2-like Protein)	Promotor	Feuchter <i>et al.</i> , 1993; Kowalski <i>et al.</i> , 1997, 1999; Wang <i>et al.</i> , 1998; Kowalski und Mager, 1998
HERV-H	<i>PLT</i> (Placental LTR Terminated Gene)	Polyadenylierung	Goodchild <i>et al.</i> , 1992
HERV-H	<i>HHLA2</i> (HERV-H LTR associating gene 2)	Polyadenylierung	Mager <i>et al.</i> , 1999
HERV-H	<i>HHLA3</i> (HERV-H LTR associating gene 3)	Polyadenylierung	Mager <i>et al.</i> , 1999
HERV-H	<i>HMGI-C</i> (non histone DNA binding protein)	Polyadenylierung	Kazmierczak <i>et al.</i> , 1996
HERV-I	<i>QP-C</i> (Ubichinon Bindeprotein)	Proteinbindestellen	Suzuki <i>et al.</i> , 1990
ERV-9	<i>ZNF80</i> (Zink Finger Protein)	Promotor	Di Cristofano <i>et al.</i> , 1995
ERV-9	$\beta$ -Globin Gene Cluster	Promotor/ Enhancer	Long <i>et al.</i> , 1998; Plant <i>et al.</i> , 2001; Routledge und Proudfoot, 2002

## 1. Einleitung

---

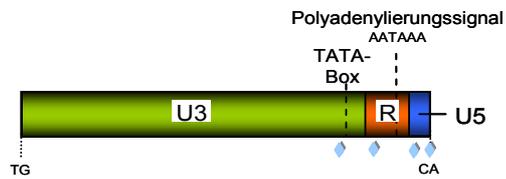
HERV-LTR	beeinflusstes Gen	Funktion	Referenz
ERV-9	<i>ADH1C</i> (Alkohol Dehydrogenase Gen)	Promotor	Chen <i>et al.</i> , 2002
HERV-F	<i>ZNF195</i> (Krüppel related zink finger protein)	Polyadenylierung	Kjellmann <i>et al.</i> , 1999
HERV-K (HML-2)	<i>LEPR</i> (leptin receptor)	Polyadenylierung	Kapitonov und Jurka, 1999
HERV-K (HML-4)	<i>FLT4</i> (Transmembran Tyrosin Kinase)	Polyadenylierung	Baust <i>et al.</i> , 2000
HERV-K(HML-6)	<i>HLA-DRB6</i> (histocompatibility complex)	5'-UTR, Promotor?	Mayer <i>et al.</i> , 1993
HERV-K(HML-8)	<i>LEP</i> (Leptin)	Enhancer	Bi <i>et al.</i> , 1997
HERV-L	<i><math>\beta</math>3Gal-T5</i>	Promotor	Dunn <i>et al.</i> , 2003; Dunn <i>et al.</i> , 2005
HERV-R (ERV-3)	<i>plk</i> (human provirus linked Krüppel gene)	Promotor	Kato <i>et al.</i> , 1987; Abrink <i>et al.</i> , 1998; Cohen <i>et al.</i> , 1988; Kato <i>et al.</i> , 1990; Larsson <i>et al.</i> , 1993; Venables <i>et al.</i> , 1995
HERV-F	<i>ZNF195</i> (Krüppel related zink finger protein)	Polyadenylierung	Kjellmann <i>et al.</i> , 1999

## **1.7 Struktur und Aktivität der HERV-LTRs**

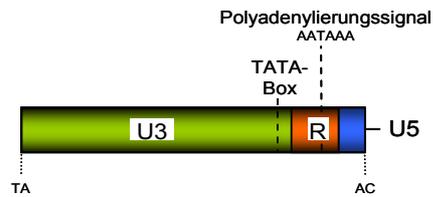
LTRs aus unterschiedlichen HERV-Familien unterscheiden sich in der Regel hinsichtlich ihrer transkriptionellen und gewebespezifischen Aktivität. Einige HERV-Elemente sind in allen Geweben aktiv, andere lediglich in bestimmten Zelltypen (Schön *et al.*, 2001). Im folgenden Abschnitt werden die wichtigsten HERV-Familien vorgestellt.

## Klasse I

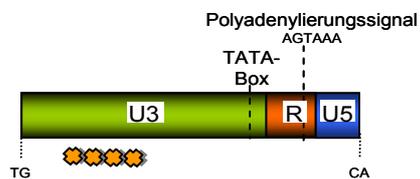
HERV-R (ERV-3)



HERV-E (4-1, ERV A)

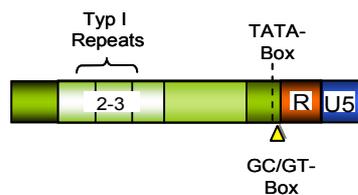


HERV-I (HERV-IP-T47D)

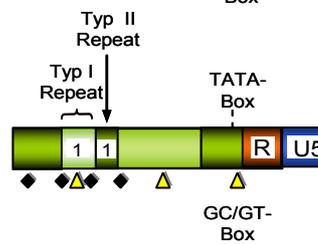


HERV-H

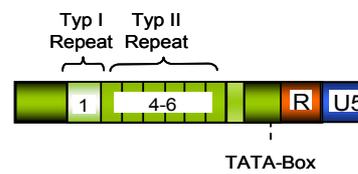
Typ I



Typ Ia



Typ II



100 bp

Abb. 6 a: Struktur und regulatorische Elemente von HERV-LTRs

# 1. Einleitung

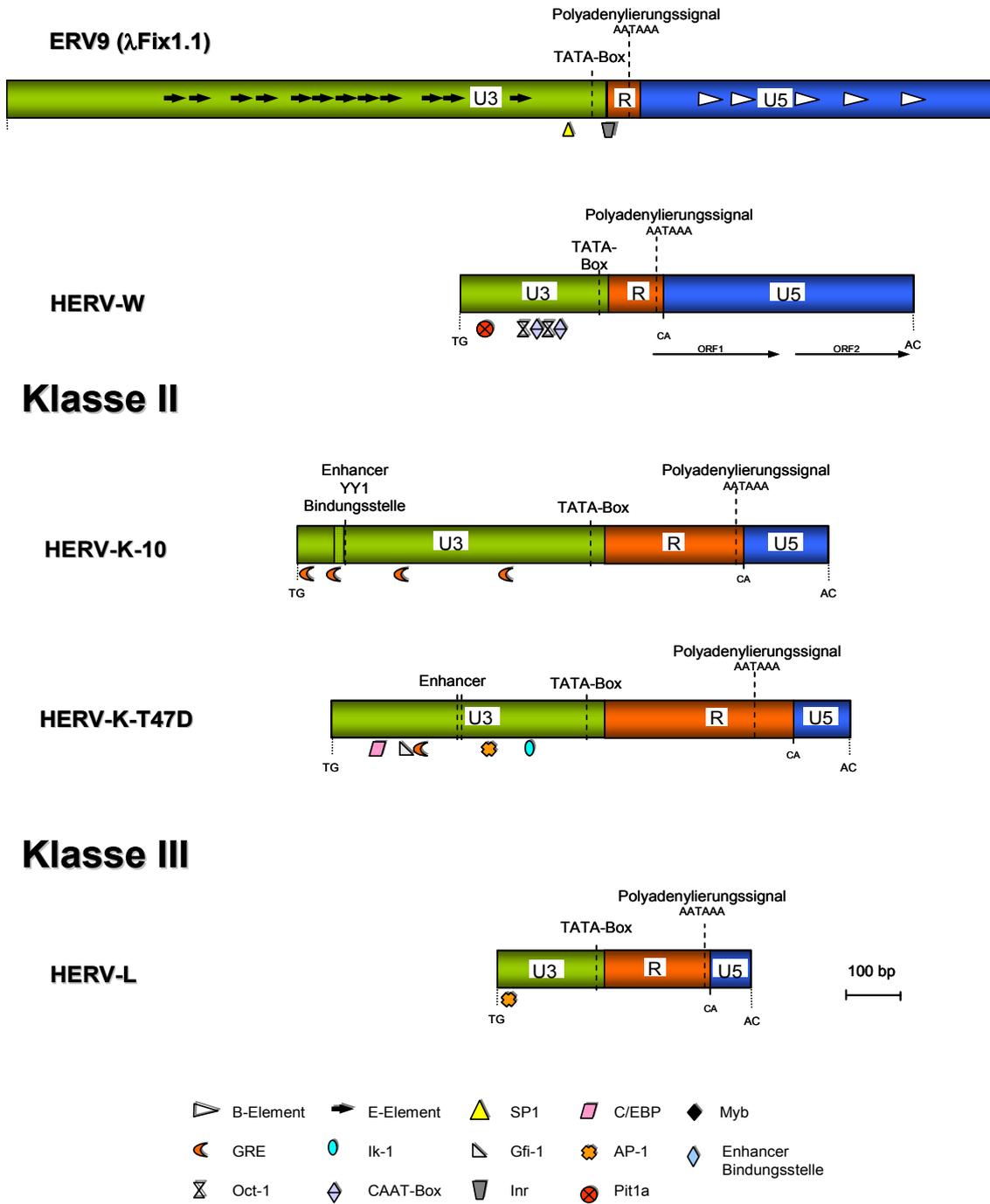


Abb. 6 b: Struktur und regulatorische Elemente von HERV-LTRs

### 1.7.1 HERV-R

HERV-R zählt zu den Klasse I Elementen, die mit den Gamma-Retroviren verwandt sind. Im menschlichen Genom sind etwa 100 provirale und 125 solitäre HERV-R-LTRs vorhanden (Mager *et al.*, 2003). Die LTR-Sequenzen haben eine Länge von 593 bp, besitzen eine TATA-Box und eine Polyadenylierungsstelle. Die U5-Region von HERV-R-Elementen ist mit 35 Basenpaaren relativ kurz (Abb. 6a). Die HERV-R-LTR enthält vier Bereiche mit Homologien zu einer Konsensus-Sequenz, die in vielen viralen Enhancern vorhanden ist. Mit Hilfe eines *pol*-spezifischen HERV-Microarrays konnten HERV-R-Transkripte in Uterus, Zervix, Plazenta, Schilddrüse und Haut nachgewiesen werden (Seifarth *et al.*, 2005). Untersuchungen an transgenen Ratten in denen eine HERV-R-LTR an ein Reporter-gen gekoppelt war, zeigten die stärkste HERV-R-LTR-Aktivität in der Harderschen Drüse und in den submandibularen Speicheldrüsen (Tanaka *et al.*, 2003).

### 1.7.2 HERV-E

Die HERV-E-Familie gehört ebenfalls zu den Klasse I Elementen und besteht aus ca. 250 proviralen und 1000 solitären LTR-Sequenzen (Mager *et al.*, 2003). Die typische HERV-E-LTR ist ca. 490 bp lang (Steele *et al.*, 1984), enthält eine potentielle TATA-Box und ein Polyadenylierungssignal (Abb. 6a). Einige HERV-E-LTRs, wurden sowohl *in vitro* als auch *in vivo* als funktionierende Promotoren und Enhancer beschrieben. Ein Beispiel hierfür sind die HERV-E-LTRs die im Amylase-Gen-Cluster lokalisiert sind, die außer der TATA-Box und dem Polyadenylierungssignal noch eine CAAT-Box und ein „Glucocorticoid Responsive Element“ (GRE) enthalten. Sie steuern die gewebespezifische Expression der menschlichen Speicheldrüsen-Amylasen (Emi *et al.*, 1988; Samuelson *et al.*, 1988; Meisler und Ting, 1993; Ting *et al.*, 1992; Samuelson *et al.*, 1990; Samuelson *et al.*, 1996). Andere HERV-E-Elemente sind in den meisten Geweben wie Blut, Gehirn, Muskel, Haut, allen inneren

## 1. Einleitung

---

Organen, Brustdrüse und Schilddrüse, sowie in Uterus, Zervix, Plazenta, Hoden und Prostata aktiv (Seifarth *et al.*, 2005).

### 1.7.3 HERV-I

Die Familie HERV-I besteht aus ungefähr 250 unterschiedlichen proviralen Elementen (Mager *et al.*, 2001) und ebenso vielen solitären LTRs. Zusätzlich sind ca. 1800 verkürzte LTR-Kopien vorhanden, denen nahezu die Hälfte des 5'-Bereiches, und in vielen Fällen weitere 60 Nukleotide am 3'-Ende, fehlen (Seifarth *et al.*, 2000). Innerhalb des verbliebenen Sequenzbereichs befindet sich eine potentielle TATA- und eine CAAT-Box (Abb. 6a; Leib-Mösch, 2005). Ein Beispiel für ein exprimiertes Element der HERV-I-Familie ist das von der menschlichen Brustzellkrebslinie T47D produzierte, aus Partikeln isolierte HERV-IP-T47D. Es gehört zu einer HERV-I-Unterfamilie, an deren Primer-Bindestelle (PBS) anstelle der Prolin-tRNA eine, Isoleucin-tRNA bindet (Seifarth *et al.*, 1995). Hier konnten mittels Computeranalyse mehrere potentielle Transkriptionsfaktorbindungsstellen wie C/EBP, GFI-1, und vier AP1-Bindungsstellen innerhalb der U3-Region identifiziert werden. Die isolierte HERV-IP-LTR erwies sich in transienten Transfektionsexperimenten als aktiv in Leber- und Nierenzellen (Schön *et al.*, 2001). HERV-I-Transkripte wurden in Haut, Muskel, Uterus, Zervix, Hoden und Schilddrüse entdeckt (Seifarth *et al.*, 2005).

### 1.7.4 HERV-H

Die HERV-H-Familie gehört ebenfalls zu den Klasse I Elementen (Mager und Henthorn, 1984). Sie besteht aus 1124 proviralen Sequenzen mit einem *pol*-Gen und zählt damit zu den größten Familien der humanen endogenen Retroviren (Jern *et al.*, 2004). Darüberhinaus enthält das menschliche Genom mehr als 1000 solitäre LTRs (Mager und Medstrand, 2003, Jern *et al.*, 2004). HERV-H-Elemente konnten auch in Neuweltaffen nachgewiesen werden und müssen daher vor über 40 Millionen Jahren in das Primaten-Genom integriert worden sein. Die hohe Kopienzahl der

HERV-H-Elemente geht auf mehrfache Amplifikationen durch Retrotransposition zurück, die vor ca. 30 bis 35 Millionen Jahren stattgefunden haben (Mager und Freeman, 1995; Andersson *et al.*, 1997). Die meisten der HERV-H-Proviren besitzen kein *env*-Gen und sind im Bereich des *pol*-Gens teilweise deletiert. Etwa 10% der HERV-H-Elemente, wie z.B. *rgh2*, sind mit einer Länge von 8,7 kb komplette Proviren. Hier sind vollständige *gag*-, *pol*- und *env*-Gene vorhanden, die allerdings durch zahlreiche Stopp-Codons unterbrochen werden (Hirose *et al.*, 1993, Jern *et al.*, 2004). Wentzensen *et al.* konnten 2004 eine hohe Expression von HERV-H-Sequenzen in Kolonkarzinomgewebe nachweisen. Mittels Microarray konnte gezeigt werden, dass die kompletten HERV-H-Proviren in Haut, Muskel, Herz, Rektum, Ovar, Uterus, Zervix und Schilddrüse exprimiert werden (Seifarth *et al.*, 2005). Sequenzvergleiche von HERV-H-LTRs zeigten, dass drei Subtypen existieren: Typ I, II und Ia (Mager, 1989). Diese Subtypen unterscheiden sich größtenteils in der U3-Region, die unterschiedliche Kopienzahlen von „Tandem Repeats“ enthält (Abb. 6a). Typ I und Typ II HERV-H-LTRs konnten bei Klonierung stromaufwärts des Human  $\beta$ -Globulin-Genpromotors (Nelson *et al.*, 1996) als negative regulatorische Elemente identifiziert (NREs) werden. Typ II HERV-H-LTRs sind vollständig inaktiv, sie beinhalten einen Typ I und vier bis sechs Typ II Repeats.

Das HERV-H-Element, für das bisher die stärkste Promotoraktivität nachgewiesen werden konnte, ist HERV-H-H6, ein Typ Ia Element. Der HERV-H-H6-Promotor ist hoch aktiv in menschlichen NTera2D1, 293, Hep2, U373 und LC5 Zellen (Anderssen *et al.*, 1997; Sjøttem *et al.*, 1996). Der Grund für diese Aktivität ist, dass in dieser LTR kaum Repressor-Elemente (NREs) vorhanden sind und mindestens zwei SP1-Transkriptionsfaktorbindungsstellen synergistisch wirken (Nelson *et al.*, 1996; Anderssen *et al.*, 1997). Alle bisher isolierten Typ I und Typ Ia LTRs enthalten eine GC-reiche Region, die sich stromabwärts der TATA-Box befindet (Andersson *et al.*, 1997; Sjøttem *et al.*, 1996) und sich als essentiell für die Promotoraktivität erwiesen hat. In dieser Region sind außerdem potentielle Transkriptionsfaktorbindungsstellen für die Faktoren SP1 und SP3 enthalten, die an der gewebespezifischen Expression von HERV-H-Elementen beteiligt sein könnten (Schön *et al.*, 2001; Nelson *et al.*, 1996; Sjøttem *et al.*, 1996). Untersuchungen an mutierten Promotorsequenzen zeigten allerdings, dass eine Inaktivierung der TATA-Box nahezu keine Auswirkung auf die Aktivität der LTR hat (Sjøttem *et al.*, 1996). Desweiteren konnte gezeigt

werden, dass mehrere der untersuchten HERV-H-LTRs eine bidirektionale Promotoraktivität (Feuchter *et al.*, 1990) aufweisen. Durch Klonierung einer HERV-H-H6-LTR in ein Reporterplasmid stromaufwärts eines  $\beta$ -Globin- oder TK-Promotors konnte Enhanceraktivität nachgewiesen werden (Feuchter *et al.*, 1990; Sjøttem *et al.*, 1996). Die Enhanceraktivität scheint von Transkriptionsfaktoren abzuhängen die „upstream“ der TATA- und der GC/GT-Box binden (Sjøttem *et al.*, 1996).

In der aktiven HERV-H-H6-LTR konnten insgesamt vier Myb-Transkriptionsfaktorbindungsstellen durch „DNase I Footprinting“ und „Electrophoretic Mobility Shift Assays“ (EMSA) nachgewiesen werden (De Parseval, 1999). Außerdem zeigte sich, dass die Promotoraktivität in humanen Teratocarcinoma Zellen bei Cotransfektion mit einem Myb-exprimierenden Vektor um das Siebenfache erhöht wird.

### 1.7.5 ERV9

Eine weitere Klasse I ( $\gamma$ -retrovirus-verwandte) HERV-Familie ist ERV9 (La Mantia *et al.*, 1991). Im menschlichen haploiden Genom sind ca. 300 provirale Elemente und mindestens 5000 solitäre LTRs vorhanden (Mager und Medstrand, 2003). Transkripte dieser HERV-Familie wurden in allen bisher untersuchten Geweben entdeckt (Seifarth *et al.*, 2005).

ERV9-LTRs sind mit bis zu 2 kb die größten bisher bekannten retroviralen LTRs. Sie besitzen unterschiedliche Kopienzahlen retroviraler Elemente in der U3-Region (41 bp E-Element) und in der U5-Region (72 bis 85 bp langen B-Elemente; Abb. 6b; Leib-Mösch, 2004). Untersuchungen der Promotoraktivität mit einem CAT-Reportergen zeigten, dass die vollständige ERV9-LTR lediglich in NT2/D1-Zellen aktiv ist, der minimale Promotor dieser LTR (-70 bis +4) sowohl in NT2/D1 als auch in der Zelllinie Hela Aktivität zeigt (Strazzullo *et al.*, 1994). Dieses Ergebnis deutet auf NREs außerhalb der minimalen Promotorregion hin.

Im ERV9-Minimalpromotor konnten zwei wichtige Regionen identifiziert werden, die für die maximale transkriptionelle Aktivität von Bedeutung sind (La Mantia *et al.*, 1991). Die erste liegt im Bereich -70 bis -39 und enthält eine Sp1-Bindungsstelle, die

zweite liegt in der Region -7 bis +6 und ähnelt einem Initiatorelement (Inr). Letztere wird von einem Inr-bindenden Protein erkannt und für einen korrekten Transkriptionsstart benötigt (La Mantia *et al.*, 1994). Strazzullo *et al.* (1994) konnten mittels Kotransfektionsexperimenten zeigen, dass die Transkriptionsfaktoren Sp1, E1A und VP16 synergistisch die Transkriptionsrate des ERV-P-Promotors erhöhen.

### 1.7.6 HERV-W

Die HERV-W-Familie (auch: „Multiple Sclerosis Associated Retroviral Element“, MSRV) ist mit der Familie ERV-9 verwandt und wurde ursprünglich aus retroviralen Partikeln amplifiziert, die aus Plasma und Serum von an multipler Sklerose erkrankten Patienten isoliert wurden (Komurian-Pradel *et al.*, 1999). Im humanen haploiden Genom wurden ca. 40 provirale Elemente und 1100 solitäre LTRs identifiziert (Mager und Medstrand, 2003). Die LTR-Sequenzen sind ca. 800 bp lang, wobei die U3-Region eine Länge von 247 bis 306 bp aufweist, und die U5-Region mit ca. 450 bp ungewöhnlich groß ist. In der U5-Region konnten zwei kurze ORFs identifiziert werden (Abb. 6b; Komurian-Pradel *et al.*, 1999; Blond *et al.*, 1999). HERV-W-*pol*-Transkripte konnten mittels Microarray-Analysen in Blut, Haut, Muskel, Herz, Hoden, Prostata, Niere, Leber, Lunge, Magen, Kolon, Mammagewebe, Uterus, Zervix, Plazenta und Schilddrüse nachgewiesen werden (Seifarh *et al.*, 2005). Isolierte HERV-W-LTRs zeigten Aktivität in einer Vielzahl an Zelllinien, dazu gehören Lungen-Fibroblasten, sowie Leber- und Pankreas-Zellen (Schön *et al.*, 2001). Ein HERV-W-Envelope-Gen codiert für das fusogene Membranprotein Syncytin, das die Fusion von Zytotrophoblasten zu Syncytiotrophoblasten in der Plazentaentstehung vermittelt (Mi *et al.*, 2000; Blond *et al.*, 2000). Außerdem konnte gezeigt werden, dass HERV-W-Env die Infektion von D-17-Zellen mit dem SN-Virus („Spleen Necrosis Virus“) verhindert (Ponferrada *et al.*, 2003).

Anhand von Computeranalysen konnten zwei potentielle Transkriptionsfaktorbindungsstellen, Pit1a und Oct-1, sowie zwei CAAT-Boxen in der Promotorregion der HERV-W-LTR identifiziert werden (Schön *et al.*, 2001; Abb. 6). Außerdem wurde eine potentielle TATA-Box in der U3-Region und eine potentielle Poly(A)Site, etwa 75 bp stromabwärts in der U3-Region, beschrieben (Schön *et al.*, 2001). Lee *et al.* (2003) konnten zeigen, dass HERV-W-Transkription durch Herpes

Simplex Typ 1 Infektion induziert werden kann. Hierfür ist das IE1-Protein von HSV-1 und die Oct-1-Bindungsstelle in der Promotorregion verantwortlich.

### 1.7.7 HERV-K (HML-1 bis 10)

Die Superfamilie HERV-K wurde ursprünglich durch Kreuzhybridisierung mit Proben des MMTV-*pol*-Gens identifiziert. Sie enthält alle  $\beta$ -Retrovirus-verwandten Klasse II Elemente. Die Abkürzung HML steht für „Human MMTV Like Virus“, und es konnten mittlerweile durch Analysen von Reverse-Transkriptase-Sequenzen zehn verschiedene Gruppen innerhalb dieser Klasse identifiziert werden (Medstrand & Blomberg, 1993; Andersson *et al.*, 1999). Die HML-2-Gruppe (Prototyp HERV-K10) umfasst ca. 60 provirale Elemente und ca. 2500 solitäre LTRs (Mager *et al.*, 2003). Die HERV-K(HML-2)-LTR ist ca. 970 bp lang (Abb. 6b). Sie enthält eine TATA-Box, ein Polyadenylierungssignal und „terminal inverted repeats“. Außerdem enthält die LTR ein partielles „Glucocorticoid Responsive Element“ (GRE) am 5'-Ende der U3-Region, welches direkt in eine Sequenz übergeht, die einer Enhancer-Core-Sequenz gleicht (Ono *et al.*, 1987). Durch die Konstellation dieser beiden Elemente weist die Sequenz Ähnlichkeit zur Bindungsstelle eines Progesteron-Bindekomplexes auf. Knössl *et al.* konnten eine 22 bp lange Sequenz (GGAGACTCCATTTTGTATGTG) als Enhancer-Komplex-Bindesequenz identifizieren die die GRE-Bindungsstelle und einen Teil des Enhancer-Kerns überlappt. Sie befindet sich zwischen den Nukleotiden 62 und 83. Außerdem wurden in der U3-Region drei weitere potentielle GREs durch Computeranalyse gefunden, die ebenfalls für die Steroidhormonstimulierbarkeit der HERV-K-Expression verantwortlich sein könnten (Knössl *et al.*, 1999). Der „Zink Finger Type Transcription Factor“ YY1 ist wahrscheinlich eine Komponente des HERV-K-Enhancer-Komplexes. Die YY1- Konsensussequenz CCATNTT ist in allen HERV-K-LTRs identisch, und es konnte gezeigt werden, dass die Inaktivierung dieser Sequenz zu signifikant reduzierter HERV-K-Expression führt (Knössl *et al.*, 1999). HERV-K-Expression wurde außerdem *in vivo* in transgenen Mäusen untersucht, wobei die HERV-K-LTR (LTRE3) an ein  $\beta$ -Galaktosidase-Reporter gen gekoppelt wurde. Es zeigte sich, dass das Reporter gen in erster Linie in undifferenzierten Spermatozyten und, in

geringerem Umfang, im Gehirn exprimiert wird (Casau *et al.*, 1999). Für einige solitäre HERV-K(HML-2)-LTRs konnte gezeigt werden, dass sie als Promotoren in beiden Orientierungen wirken können (Domansky *et al.*, 2000). Seifarth *et al.* konnten HML-2-Transkription in Gehirn, Blut, Haut, Niere, Leber, Lunge, Magen, Kolon, Rektum, Brustgewebe, Ovarien, Uterus, Zervix, Plazenta, Prostata und Schilddrüse nachweisen. Das menschliche Genom enthält nur sechs HERV-K-T478-Proviren (Seifarth *et al.*, 1998) und etwa 850 solitäre LTRs (Baust *et al.*, 2001). Transfektionsexperimente mit Luziferase als Reporter gen zeigten, dass HERV-K-T47D-LTRs in der Brustkrebszelllinie T47D aktiv sind und dass sie durch Estradiol und Progesteron stimuliert werden können (Seifarth *et al.*, 1998; Baust *et al.*, 2001). Die Steroid-abhängige Expression könnte auf einem „Glucocorticoid Response Element“, das Ähnlichkeiten zur HML-2-Unterfamilie aufweist, und auf zwei Enhancer-ähnlichen Bereichen in der U3-Region der 5'-LTR (Abb. 6) beruhen. Außerdem wurden potentielle Transkriptionsfaktorbindungsstellen der Transkriptionsfaktoren C/EBP, Gfi-1, AP1 und Ik-1 durch Computeranalyse gefunden (Seifarth *et al.*, 1998). In den bisher untersuchten HERV-K(HML-4)-LTRs wurde eine ungewöhnlich große Distanz zwischen dem Polyadenylierungssignal und den beiden alternativen Polyadenylierungsstellen festgestellt (Baust *et al.*, 2000). Eventuell werden hier die beiden Elemente durch Bildung einer verkürzenden Sekundärstruktur in die richtige Distanz gebracht. HML-4-Transkription konnte in Gehirn, Blut, Haut, Skelettmuskel, Herz, Niere, Leber, Lunge, Magen, Kolon, Rektum, Brustgewebe, Ovarien, Uterus, Zervix, Plazenta, Prostata und Schilddrüse nachgewiesen werden (Seifarth *et al.*, 2005).

### 1.8 HERV-L

Die Gruppe der Humanen endogenen Retroviren, die als Primer für die reverse Transkriptase die Leucin-tRNA verwenden, wird als HERV-L-Familie bezeichnet. Sie lässt sich durch Sequenz-Vergleiche ihrer LTRs in die Unterfamilien MLT2A1, MLT2A2, MLT2B1 und MLT2B2 einteilen („RepBase“). Im Falle HERV-L-ähnlicher Elemente aus anderen Spezies wird die Bezeichnung ERV-L verwendet. Die Familie HERV-L gehört zur Klasse III der Spumaretroviren und ist in allen bisher untersuchten Säugetierordnungen, inklusive Primaten, Rodentia, Carnivora, Lagomorpha, Perissodactyla und Artiodactyla, in der Regel in niedriger Kopienzahl (10-30 Kopien), hochkonserviert vorhanden (Benit *et al.*, 1999). Überdurchschnittlich hohe Kopienzahlen finden sich in den Ordnungen Rodentia und Primates, wobei unter den Primaten das menschliche Genom mit etwa 200 proviralen Elementen und 6000 solitären LTRs (Mager *et al.*, 2001) die höchste Kopienzahl enthält. Die erste Beschreibung eines HERV-L-Provirus erfolgte 1995 durch Cordonnier *et al.*. Dieses HERV-L-Provirus, das in der vorliegenden Arbeit als HERV-L(Cord) bezeichnet wird, hat eine Länge von 6591 bp. Die 5'-LTR ist 462 bp, die 3'-LTR 459 bp lang, beide LTRs zeigen eine Sequenzübereinstimmung von 82% und werden von den typischen „Inverted Repeats“ an beiden Enden begrenzt (Abb. 6b). Weder in der *gag*- noch in der *env*-Region des Virus sind „Open Reading Frames“ vorhanden, da hier zahlreiche Stopp-Codons vorliegen. In diesen Bereichen lässt sich auch keine phylogenetische Übereinstimmung mit exogenen Retroviren finden. Die *pol*-Region enthält mehrere unterbrochene „Open Reading Frames“ für Reverse Transkriptase, RNaseH, Integrase und dUTPase. Hier zeigt sich im Bereich der codierenden Sequenzen eine 62%-ige Aminosäurenübereinstimmung mit SFV („Simian Foamy Virus Typ1“). Außerdem lässt sich eine 53%-ige Übereinstimmung der retroviralen dUTPase-Sequenz mit dem „Mouse Mammary Tumor Virus“ feststellen (Cordonnier *et al.*, 1995). Kulski *et al.* (1999) zeigten, dass die in dem 656 kb großen Bereich der Alpha- und Betablock-Region des „Major Histocompatibility Complexes“ (MHC) vorhandenen 16 Sequenzen der P5-Multicopy-Gen-Familie mit der HERV-Familie verwandt sind. Damit liegt die Anzahl der HERV-Elemente hier 10-fach höher als im genomischen Durchschnitt. Die P5-Familie ist auf den MHC beschränkt, ihre

## 1. Einleitung

---

Funktion ist bisher nicht bekannt. Die Sequenz P5-1, die spezifisch in Lymphzellen exprimiert wird, hat eine Länge von 2,5 kb und enthält insgesamt sechs ORFs. Sequenzvergleiche zeigten eine 60%-ige Übereinstimmung auf DNA-Ebene mit der *pol*-Region von HERV-L. Nakagawa *et al.* (1997) konnten in der Gelenkflüssigkeit von Patienten mit rheumatoider Arthritis erhöhte HERV-L-Expression im Vergleich zu einer gesunden Kontrollgruppe nachweisen.

Das Mausgenom enthält ca. 100 bis 200 MuERV-L-Kopien (Murine Endogene Retroviren Typ L). Das von Benit *et al.* (1997) charakterisierte MuERV-L-Element hat eine Länge von 6471 bp. Die LTR zeigt eine Nukleotidsequenzübereinstimmung von 40% mit der 5'-HERV-L(Cord). Sowohl in der *gag*- als auch in der *pol*-Region sind intakte „Open Reading Frames“ vorhanden. In Präimplantationsembryos konnte gezeigt werden, dass MuERV-L zu Beginn des ersten Zellzyklus transkribiert wird und eine Inhibierung durch MuERV-L-Antisense-Oligonukleotide zu einer deutlich langsameren Entwicklung des Embryos nach dem 4-Zellstadium führt (Kigami *et al.*, 2003). Die *gag*-Region von MuERV-L hat eine Übereinstimmung von 43% auf Aminosäureebene zu Fv1 („Friendvirus susceptibility-1“; Benit *et al.*, 1997). Fv1 stammt wahrscheinlich von der Capsid Region (CA) eines frühen ERV-L-Elementes ab (Best *et al.*, 1996). Es kontrolliert die Replikation von MLV („Murine Leukemia Virus“) und verhindert so die MLV-assoziierte Leukämie (Jolicoeur *et al.*, 1979). Fv1 reagiert nach dem Eintritt des Virions in die Wirtszelle, bevor es zur Integration und damit zur Entstehung des Provirus kommt. Vermutlich kommt es zur Interaktion mit dem Capsidprotein (CA) des Virions, das durch das virale *gag*-Gen codiert wird, was zu einer Inaktivierung des Virions führt. Phylogenetische Untersuchungen von Vertretern der Superordnung Uranotheria (asiatischer und afrikanischer Elefant, Seekuh, Schliefer, u.a.) zeigten eine ähnlich große Anzahl an HERV-L-Elementen wie in den Ordnungen Primates und Rodentia. Untersuchungen von Mammut-DNA lieferten vergleichbare Ergebnisse, was die Konservierung dieser Sequenzen innerhalb der Superordnung Uranotheria zeigt (Greenwood *et al.*, 2004, Greenwood *et al.*, 2001).

Durch Untersuchungen der HERV-Expression in humanem Gewebe mittels HERV-Microarray-Analyse (Seifarth *et al.*, 2005), wobei HERV-*pol*-spezifische, degenerierte Primer verwendet wurden, wurde HERV-L-Expression in den Geweben Haut, Uterus,

Zervix, Plazenta und Schilddrüse detektiert. Damit zählt HERV-L zu den HERVs die am gewebespezifischsten im menschlichen Genom exprimiert werden. Die Promotoraktivität der 5'-HERV-L-LTR(Cord) wurde mittels Luziferase-Reportergen in mehreren Zelllinien getestet. Dabei konnte lediglich in der Keratinozytenzelllinie HaCaT und in den primären Zellen NHEK Aktivität detektiert werden (Schön, unveröffentlichtes Material). Computeranalysen der Sequenz lieferten eine potentielle TATA-Box (Position bp 4-13), ein potentielles Polyadenylierungssignal und eine potentielle AP1-Bindungsstellen (bp 94-114; Cordonnier *et al.*, 1995; Abb. 6b). Eine andere solitäre HERV-L-LTR dient als alternativer Promotor der  $\beta$ Gal-T5-Expression.  $\beta$ Gal-T5 ist an der Synthese von Typ1-Kohlenhydraten in gastrointestinalen und pankreatischen Geweben beteiligt. Die HERV-L-LTR(Gal) zeigt in diesem Fall die stärkste Promotoraktivität im Kolon (Dunn *et al.*, 2003). Hier sind die Transkriptionsfaktorbindungsstellen HNF-1/Cdx, eine CAAT-Box und SP1 essentiell (Dunn *et al.*, 2003; Isshiki *et al.*, 2003). In weiteren Untersuchungen konnte die Promotoraktivität der HERV-L-LTR(Gal) in Schimpansen, Gorillas, Gibbons und Pavianen nachgewiesen werden, wobei in allen Fällen die kolonspezifischen Transkriptionsfaktorbindungsstellen HNF-1/Cdx, SP1 und CAAT-Box konserviert waren (Dunn *et al.*, 2005).

In meiner Doktorarbeit habe ich mich mit der Untersuchung von HERV-L-Elementen befasst. Diese Elemente sind vor allem deshalb interessant, weil sie gewebespezifische Expressionsmuster zeigen (Seifarth *et al.*, 2005). Auf Grund dieser Gewebespezifität könnten HERV-L-LTRs Verwendung als Promotoren in retroviralen Vektoren finden und in der Gentherapie eingesetzt werden. In der vorliegenden Arbeit soll daher die HERV-L-LTR(Cord), deren Promotoraktivität bereits in der Keratinozyten-Zelllinie HaCaT gezeigt werden konnte (Schön, 2001), in weiteren Zelllinien getestet werden. Außerdem sollen weitere HERV-L-LTRs kloniert, charakterisiert und auf zelltypspezifische Promotoraktivität getestet werden, um somit weitere HERV-L-LTR-Elemente zu finden, die als spezifische Promotoren verwendet werden können. Weiterhin soll die gewebespezifische Promotoraktivität der HERV-L-LTR(Cord) in transgenen Mäusen untersucht werden. Ein weiterer wichtiger Teil der vorliegenden Arbeit ist die Identifizierung von Transkriptionsfaktoren, die an der Regulation des HERV-L-LTR(Cord)-Promotors beteiligt sind. Hier soll gezeigt werden

## 1. Einleitung

---

welche Transkriptionsfaktoren für die zelltypspezifische Expression von HERV-L(Cord) von Bedeutung sind.

## 2 Material

### 2.1 Puffer und Lösungen

Lösungen für die Zellkultur wurden autoklaviert. Hier nicht aufgeführte Puffer und Lösungen sind im Methodenabschnitt oder in den Handbüchern der einzelnen Hersteller beschrieben.

#### Puffer

#### Bestandteile

Band Shift Puffer (10x) (BSB)

10 mM Tris/HCl pH 7,5

10 mM DTT

50% Glycerin

10 mM EDTA

Ethidiumbromidlösung

Ethidiumbromid 1% (w/v)

Aqua dest.

Phosphat gepufferte Kochsalzlg.  
pH 7,4 (PBS)

140 mM NaCl

5,4 mM KCl

9,7 mM Na<sub>2</sub>HPO<sub>4</sub> x 2 H<sub>2</sub>O

2 mM KH<sub>2</sub>PO<sub>4</sub>

Aqua dest.

## 2. Material

---

Probenauftragspuffer für Agarosegelelektrophorese	Bromphenolblau 0,25% (w/v) Xylencyanol FF 0,25% (w/v) Ficoll 15% (w/v) Aqua dest.
Puffer A für Kernextraktion	10 mM Hepes (pH 7,6) 15 mM KCl 2 mM MgCl <sub>2</sub> 0,1 mM EDTA (pH 8,0)
Puffer A + NP40	Puffer A + 0,2% NP40 (v/v)
Puffer C für Kernextraktion	25 mM Hepes (pH 7,6) 50 mM KCl 0,1 mM EDTA (pH 8,0) 10% Glycerin
Puffer A / Puffer A+NP40/Puffer C	direkt vor Verwendung hergestellt: 100 µl 0,1 M PMSF 10 µl DTT 30 µl 5 mg/ml Aprotinin 10 µl 0,5 mg /ml Leupeptin
TE-Puffer (1x)	10 mM Tris/HCl 1 mM EDTA Aqua dest.
Tris Acetat EDTA Puffer (50x) (TAE)	1 M Tris Base Essigsäure 5,71% EDTA pH 8,0 Aqua dest.

## 2. Material

---

Tris Borat EDTA Puffer (10x) (TBE)	890 mM Tris Base 890 mM Borsäure 20 mM EDTA, pH 8,0 Aqua dest.
1 M Tris/HCl, (verschiedene pH-Werte)	1M Tris Base Aqua dest. pH mit HCl (konz.) einstellen
Trypanblaulösung	NaCl 0,9% (w/v) Trypanblaufarbstoff (0,5%) (w/v) Aqua dest.
Trypsin-EDTA Lösung	140 mM NaCl 5,4 mM KCl 100 mM Na <sub>2</sub> HPO <sub>4</sub> x 2 H <sub>2</sub> O 2,2 mM KH <sub>2</sub> PO <sub>4</sub> Trypsin 0,05% (v/v) EDTA 0,54% (v/v) Aqua dest.

## 2.2 Oligonukleotide

### 2.2.1 Primer für „Site Directed Mutagenesis“

>PAX8: GTCTCGACTTGATTGTATTGAGGGATGCAAAGCATTGATATTG  
>PAX8.cr: CAATATCAATGCTTTGCATCCCTCAATACAATCAAGTCGAGAC  
>AP1: CTAAGGAGATTAGCATTGTGTCCGTGGACTGGGAGAGGCAG  
>AP1.cr: CTGCCTCTCCCAGTCCACGGACACAAATGCTAATCTCCTTAG  
>NEUR: GGTGGACCCAGTCTAATCATATTCCTGCATGGCCAGAATAAAAGC  
>NEUR.cr: GCTTTTATTCTGGCCATGCAGGAATATGATTAGACTGGGTCCACC  
>RCAT: GATGCAAAGCATTGATATCGAGTGTGTCTGTGTGGGTG  
>RCAT.cr: CACCCACACAGACACACTCGATATCAATGCTTTGCATC  
>IRF: CCAGAATAAAAGCAGGCAGAAGGACGTGAGAAGACTAGACTGG  
>IFR.cr: CCAGTCTAGTCTTCTCACGTCCTTCTGCCTGCTTTTATTCTGG  
>WT: GATATTGGGTGTGTCTTTGTGGGTCTTACCTAAGGAG  
>WT.cr: CTCCTTAGGTAAGACCCACAAAGACACACCCAATATC  
>HOXF: GTGTTACCTAAGGACATTTACATTTGAGTCAGTGGACTGG  
>HOXF.cr: CCAGTCCACTGACTCAAATGTAAATGTCCTTAGGTAACAC  
>TBPF: TGTGTCTGTGTGGGTGTTACCTGATGAGATTAACATTTGAGTC  
>TBPF.cr: GACTCAAATGTTAATCTCATCAGGTAACACCCACACAGACACA  
>SMAD: GTGATGATTAATACCAACTCTCGGCTTGATTGTATTGAAGG  
>SMAD.cr: CCTTCAATACAATCAAGCCGAGAGTTGGTATTAATCATCAC

## 2. Material

---

### **2.2.2 Oligonukleotide für „Electrophoretic Mobility Shift Assay“ (EMSA)**

>RCAT/CLTR for: CAAAGCATTGATATTGGGTGTGTCTGTGTG  
>RCAT/CLTR rev: CACACAGACACACCCAATATCAATGCTTTG  
>RCAT/CLTR for mutiert: CAAAGCATTGATATCGAGTGTGTCTGTGTG  
>RCAT/CLTR rev mutiert: CACACAGACACACTCGATATCAATGCTTTG  
>EGR/WT for: TATTGGGTGTGTCTGTGTGGGTGTTACCT  
>EGR/WT for mutiert: TATTGGGTGTGTCTTTGTGGGTCTTACCT  
>EGR/WT rev: AGGTAACACCCACACAGACACACCCAATA  
>EGR/WT rev mutiert: AGGTAAGACCCACAAAGACACACCCAATA  
>HOX for: TTACCTAAGGAGATTAACATTTGAGTCAGT  
>HOX rev: ACTGACTCAAATGTTAATCTCCTTAGGTAA  
>HOX for mutiert: TTACCTAAGGACATTTACATTTGAGTCAGT  
>HOX rev mutiert: ACTGACTCAAATGTAAATGTCCTTAGGTAA  
>AP1 for: AGGAGATTAACATTTGAGTCAGTGGACTGG  
>AP1 rev: CCAGTCCACTGACTCAAATGTTAATCTCCT  
>AP1 for mutiert: AGGAGATTAGCATTTGTGTCCGTGGACTGG  
>AP1 rev mutiert: CCAGTCCACGGACACAAATGCTAATCTCCT  
>NEUR/AP4 for: GACCCAGTCTAATCATCTGCCAGCATGGCC  
>NEUR/AP4 rev: GGCCATGCTGGCAGATGATTAGACTGGGTC  
>NEUR/AP4 for mutiert: GACCCAGTCTAATCATATTCCTGCATGGCC  
>NEUR/AP4 rev mutiert: GGCCATGCAGGAATATGATTAGACTGGGTC  
>HOXF for mutiert: TTACCTAAGGAGATTAAGCACCTAGTCAGT  
>HOXF rev mutiert: ACTGACTAGGTGCTTAATCTCCTTAGGTAA  
>AP1 consensus: CGCTTGATGACTCAGCCGGAA

## 2. Material

---

>AP1 rev consensus: TTCCGGCTGAGTCATCAAGCG  
>AP1 for consensus mutiert: CGCTTGATGACTTGGCCGGAA  
>AP1 rev consensus mutiert: TTCCGGCCAAGTCATCAAGCG  
>SP1 consensus : ATTCGATCGGGGCGGGGCGAGC  
>SP1 rev consensus: GCTCGCCCCGCCCCGATCGAAT  
>SP1 consensus mutiert: ATTCGATCGGTTCTGGGGCGAGC  
>SP1 rev consensus mutiert: GCTCGCCCCGAACCGATCGAAT  
>EGR for consensus: GGATCCAGCGGGGGCGAGCGGGGGCGA  
>EGR rev consensus: TCGCCCCCGCTCGCCCCCGCTGGATCC  
>EGR for consensus mutiert: GGATCCAGCTAGGGCGAGCTAGGGCGA  
>EGR cons mut rev: TCGCCCTAGCTCGCCCTAGCTGGATCC

### 2.2.3 PCR-Primer

>MLT LTR for: GATCATTAACTGAGTC  
>MLT LTR rev: TAGTCAGGGTTCTCTAGAGGG  
>MLT LTR (*Bam*HI) for: CCGGCCGGATCCTGATCATTAACTGAGTC  
>MLT LTR (*Hind*III) rev: CGGCCGAAGCTTTAGTCAGGGTTCTCTAGAGGG  
>HERV-L pro for: TGTGATGGTTAAYAMYGAS  
>HERV-L pro rev: GTTCATGCCAAGGACTATCAGTG  
>HERV-L (*Bam*HI) pro for: CCGGCCGGATCCTGTGATGGTTAAYAMYGAS  
>HERV-L (*Hind*III) pro rev: CGGCCGAAGCTTGTTCATGCCAAGGACTATCAGTG  
>EGFP for: TCGAGCTGGACGGCGACGTA  
>EGFP rev: TAGTGGTTGTCGGGCAGCAGC  
>Luziferase for: CCGCTGGAGAGCAACTGCAT  
>Luziferase rev: TCTATGCGGAAGGGCCACAC

## 2. Material

---

>Actin for: GTCGGAAGGTGGACAGGG

>Actin rev: GCAGTGATCTCCTTCTGCATC

Die meisten der hier verwendeten Oligonukleotide wurden in der AG BIODV von Herrn Utz Linzner in der GSF, Neuherberg synthetisiert. Die in den EMSA-Untersuchungen verwendeten Konsensusoligonukleotide und die mutierten Konsensusnukleotide wurden von der Firma Santa Cruz erworben.

### 2.3 Reagenzien

Acrylamid	Biorad, München
Agarose (ME, Seakem)	Biozym, München
Ampicillin	Sigma, Deisenhofen
Antibiotika-Antimykotika Lsg. (AB/AM)	Life Technologies, Eggenstein
Aqua bidest.	Millipore, Eschborn
Bromphenolblau	Merck, Darmstadt
Chloroform	Merck, Darmstadt
dNTPs	Perkin Elmer Cetus, Überlingen
EDTA	Sigma, Deisenhofen
Ethidiumbromid	Serva, Heidelberg
Fötale Kälber Serum FKS (FCS)	Gibco, Karlsruhe
	Seromed, Berlin
Isopropanol	Merck, Darmstadt
Kälber alkalische Phosphatase (CIP)	New England Biolabs, Schwalbach
1-kb-DNS Längenstandard	Life Technologies, Eggenstein
2 log DNS Längenstandard	New England Biolabs, Schwalbach
L-Glutamin-Lsg. (100x)	Life Technologies, Eggenstein
Restriktionsenzyme	Roche Diagnostics, Mannheim
	New England Biolabs, Schwalbach
	Promega, Madison, USA
Rinderserum Albumin (BSA)	New England Biolabs, Schwalbach

## 2. Material

---

RNA Bee RNA Isolation solvent	AMS Biotechnology Europe Ltd., Wiesbaden
Roti-Aqua-Phenol	Roth, Karlsruhe
SDS (Pulver)	Merck, Darmstadt
Shrimp alkalische Phosphatase (SAP)	Roche Diagnostics, Mannheim
T4 DNS Ligase	Roche Diagnostics, Mannheim New England Biolabs, Schwalbach
T4 Polynukleotidkinase	Roche Diagnostics, Mannheim New England Biolabs, Schwalbach
Tris	Löwe Biochemika, München
Tween20	Sigma, Deisenhofen
Trypsin-EDTA Lösung	Biochrom, Berlin

### 2.4 Sonstige Labor Materialien

Filter 0,22 µm	Nalgene, USA
Filterpapier, Whatman 3 mm	Whatman, Maidstone, GB
Plastikmaterial für Zellkultur	Greiner Labortechnik, Frickenhausen, BD Falcon, Düren oder Nunc, Wiesbaden
Zählkammer Fast Read 106	Madaus Urotech, München

### 2.5 Gebrauchsfertige Reaktionssysteme (Kits)

BCA Protein Assay Reagent Kit	Pierce, Rockford, USA
CellPectTransfectionKit	Amersham Pharmacia
Dual-Luciferase® Reporter (DLR) Assay kit	Promega, Madison, USA
Gel Extraction Kit	Qiagen, Hilden
PCR-Polymerasen Kit (Taq)	Qiagen, Hilden

## 2. Material

---

PCR Purification Kit	Qiagen, Hilden
Plasmid Maxi/Midi Kit	Qiagen, Hilden
QuikChange-Mutagenesis Kit	Stratagene
Qlamp DNA Mini Kit	Qiagen, Hilden
Superscript First-Strand Synthesis RT Kit	Invitrogen, Karlsruhe
RNase-free DNase Set	Qiagen, Hilden
Rapid DNA Ligation Kit	Roche Diagnostics, Mannheim
RNeasy Kit	Qiagen, Hilden
TOPO TA Cloning Kit	Invitrogen, Karlsruhe

## 2.6 Medien

### 2.6.1 Bakterienmedien

LB-Medium (nach Luria-Bertani)	10 g Trypton 5 g Hefeextrakt 10 g NaCl ad 1000 ml mit Aqua dest., pH 7,0
Medien zum Einfrieren von Bakterien	2 Teile 80% Glycerol 1 Teil 10 mM MgCl <sub>2</sub>
SOC Vollmedium (Life Technologies, Invitrogen)	20 g Trypton 5 g Hefeextrakt 0,5 g NaCl 2,5 mM KCl 20 mM Glucose ad 1000 ml Aqua dest., pH 7,0

## 2. Material

---

### **2.6.2 Zellkulturmedium**

DMEM mit Glutamax I, ohne Na-Pyruvat, 4,5 g/l Glucose, Zugabe von 10% FKS

F12 K (Kaighn's Modification), Zugabe von 10% FKS

DMEM/F12 (1:1) mit Glutamax I, Zugabe von 10% FKS

RPMI 1640 mit 446 mg/l Glutamin

### **2.6.3 Medium zum Einfrieren von Zellen**

DMEM mit Glutamax I, ohne Na-Pyruvat, 4,5 g/l Glucose, Zugabe von 20% FKS,  
10% DMSO

FKS wurde vor Gebrauch 0,5-1 h bei 56°C hitzeinaktiviert.

## **2.7 Antikörper**

Alle Antikörper wurden von der Firma Santa Cruz Biotechnology, USA CA bezogen:

Sp3 (D20)

Sp1 (PEP2)

c-Jun (N)

c-Jun (H79)x

c-Jun AP-1 (N)x

c-Jun/AP-1 (D)x

c-Fos (4)x

FosB (102)x

c-Fos (4-10G)x

Egr-1

### 2.8 Plasmide

#### pBL

Leervektor von Felix Hoppe-Seyler (DKFZ) enthält das Firefly *P. pyralis* Luziferasegen fusioniert an das SV40 polyA Signal von pBLCAT2, kloniert in den Polylinker von pBluescript KSM13. Enthält ein Ampicillin-Resistenzgen.

#### PCR 2.1-TOPO®

PCR-Klonierungsvektor von Invitrogen, Karlsruhe. Der Vektor liegt geöffnet mit 3' überhängenden Thymidinresten und mit gebundener Topoisomerase vor.

#### pFred143

GFP-Expressionsplasmid. Das *gfp*-Gen steht unter der Kontrolle des CMV-Promotors und enthält das SV40-Polyadenylierungssignal (Dr. Markus Neumann, GSF, Neuherberg).

### 2.9 Zelllinien

#### HeLa

Humane epitheloide Zervixkarzinomzelllinie (Fibroblasten). Die Zellen wurden von ATCC bezogen (CRL-7923).

#### HaCaT

Keratinozytenzelllinie. Kultiviert in DMEM (10% FKS)

## 2. Material

---

### KLE

Humane Uterusadenokarzinomzelllinie. Bezogen von ATCC (CRL-1622). Kultiviert in DMEM/F12 (10% FKS).

### SW

Schilddrüsenadenokarzinomzelllinie. Bezogen von ATCC (HTB 107). Kultiviert in DMEM/F12 (10% FKS)

### JEG-3

Plazentazellen, Chorionkarzinom. Bezogen von ATCC (HTB-36). Kultiviert in RPMI (10% FKS).

### LoVo

Kolonadenokarzinomzelllinie. Bezogen von ATCC (CCL-229). Kultiviert in Ham's F12K Medium (10% FKS; 2 mM Glutamin; 1,5 g/L Natriumbicarbonat).

## **2.10 Bakterienstämme**

### XL-10 Epicurian Coli®

Kompetente, superkompetente und ultrakompetente „one shot“-Zellen zur chemischen Transformation ( $\text{cfu}=1 \times 10^8$ - $5 \times 10^9/\mu\text{g}$ , Stratagene, Amsterdam, Niederlande).

## 2. Material

---

### XL-10 Epicurian Coli® Ultracompetent

Ultrakompetente Zellen zur chemischen Transformation (cfu  $\geq 5 \times 10^9$  /  $\mu\text{g}$ , Stratagene, Amsterdam, Niederlande).

### **2.11 Geräte**

Herolab Clean Cab Arbeitsbench (Herolab GmbH Laborgeräte, Wiesloch)

PCR Amplifier Perkin Elmer GeneAmp PCR System 9600 (Perkin-Elmer, CT, USA)

Hettich Tischzentrifuge Mikro 20 Winkelrotor Nr. 006871 (Andreas Hettich GmbH, Tuttlingen)

Heizblock Multi-Blok Heater (Bachofer Laboratoriumsgeräte, Reutlingen)

Eppendorf Thermomixer 5436 (Eppendorf, Hamburg)

Hettich Rotanta/TR Zentrifuge (Andreas Hettich GmbH, Tuttlingen)

Wasserbad (Mettler GmbH+Co.KG, Schwabach)

Consort Microcomputer Electrophoresis Power Supply (Turnhout, Belgium)

Power Supply PowerPac 300 (BioRad, München)

Spectrafuge 24D von Abimed Labnet International, Inc. (Edison, USA)

ChemiDoc Geldokumentationsanlage (BioRad, München)

UV-visible Recording Spectrophotometer UV-160A (Shimadzu, München)

Feinwaage (Scaltec, Heiligenstadt)

Orion Microplate Luminometer (Berthold, Bad Wildbad)

Protean II xi cell (BioRad, München)

## 2.12 Verwendete Software und Datenbanken

BioEdit Version 7.0.5.2 (Tom Hall)

ClustalX Version 1.82 (Des Higgins)

Genomatix Module Library Version 3.8, Vertebrate Modules (April 2005)

Paup (Version 4; Sinauer Associates, Sunderland, MA)

Simplicity 2.0R1 (Berthold Detection Systems, Pforzheim)

Software Gel-Doc Quantity One Version 2.4.1 Build 008 (BioRad, München)

HERVd Human Endogenous Retrovirus Database <http://herv.img.cas.cz/>

Genomatix Suite <http://genomatix.de/cgi-bin/eldorado/main.pl>

Medline <http://www.ncbi.nlm.nih.gov/Pubmed>

BLAST <http://www.ncbi.nlm.nih.gov/blast/blast.cgi?Jform=1>

GenBank <http://www.ncbi.nlm.nih.gov/Entrez/nucleotide.html>

Microsoft Excel 2002 zur Auswertung der Luziferase Assays

Microsoft PowerPoint 2002

### **2.13 HERVd-Datenbank: Accession-Nummern**

1021: MLT2A1(1); 1052: MLT2A1(2); 1058: MLT2A1(3); 1069: MLT2A1(4);  
1146: MLT2A1(5); 1237, MLT2A2(1); 1062: MLT2A2(2); 1190: MLT2A2(3);  
1019: MLT2A2(4); 1212: MLT2A2(5); 1163: MLT2B1(1); 1253, MLT2B1(2);  
1275: MLT2B1(3); 1287: MLT2B1(4); 1365: MLT2B1(5); 1299: MLT2B2(1);  
1407: MLT2B2(2); 1425: MLT2B2(3); 1440: MLT2B2(4); 1468: MLT2B2(5)

### **3 Methoden**

#### **3.1 Kultivierung adhärenter Zellen**

Die Kultivierung der permanenten Zelllinien erfolgte bei 37°C in 5%-iger Kohlendioxidatmosphäre. Die verwendeten Zellkulturmedien enthielten 10% BSA. Die Zellen wurden zweimal pro Woche passagiert, wobei sie zunächst mit PBS-Puffer gewaschen und anschließend mit Trypsin-EDTA vom Boden der Kulturflaschen gelöst wurden. Je nach Wachstumsgeschwindigkeit wurden 1:2 bis 1:10 Verdünnungen hergestellt.

#### **3.2 Zellzahlbestimmung**

Einige Experimente setzen die genaue Bestimmung der eingesetzten Zellzahl voraus. Hierfür wurde eine Zählkammer (MD-Kova; Raster 10; Madaus Diagnostik) verwendet die mit 10 µl Zellsuspension beladen wurde. Anschließend wurde unter dem Lichtmikroskop die Anzahl der Zellen pro Großquadrat ermittelt, und zur Bestimmung der Zellzahl pro Milliliter, das Ergebnis mit  $10^4$  multipliziert.

#### **3.3 Auftauen von Zellen**

Die in flüssigem Stickstoff gelagerten Zellen wurden im 37°C Wasserbad aufgetaut und rasch in Kulturflaschen mit 10 ml Medium überführt. Am folgenden Tag wurde das Medium gewechselt.

#### **3.4 Einfrieren von Zellen**

Die Zellen wurden mit PBS-Puffer gewaschen, anschließend mit Trypsin-EDTA von der Kulturflasche gelöst und in 50 ml Falcon-Röhrchen überführt. Anschließend wurde 10 min bei 1200 U/min zentrifugiert und das Pellet in kaltem Einfriermedium resuspendiert. Die Zellsuspension wurde mit einer Konzentration von  $1 \times 10^6$  bis  $1 \times 10^7$  Zellen/ml in Nunc-Cryo-Röhrchen überführt und in einen auf Raumtemperatur erwärmten Kryobehälter gegeben. Dieser Behälter ist mit einer Flüssigkeit isoliert, die ein langsames Einfrieren der Proben auf  $-80^\circ\text{C}$  ermöglicht. Nach dem Einfrieren der Proben über Nacht wurden sie in flüssigem Stickstoff bei  $-196^\circ\text{C}$  gelagert.

#### **3.5 Transfektion eukaryotischer Zellen mit Plasmid-DNA**

##### **3.5.1 Kalziumphosphat-Kopräzipitation**

Als Standardmethode wurde die Kalziumphosphat-DNA-Kopräzipitationsmethode von Graham (Graham *et al.*, 1977) verwendet. Dabei wird die negativ geladene DNA durch Kalziumionen und diese durch Phosphationen gebunden, wodurch ein kristalliner Komplex ausfällt. Vermutlich wird dieser Komplex durch Endozytose in die Zelle aufgenommen. Verwendet wurde das „Cellfect Transfection Kit“ (Pharmacia). Die zu transfizierenden Zellen wurden in DMEM mit 10% FKS bis zu 80%-iger Konfluenz kultiviert. Um den für die Transfektion geeigneten pH zu erhalten wurde das Medium der zu transfizierenden Zellen 3 h vor der Transfektion erneuert. Für jeden Transfektionsansatz wurde ein Mastermix hergestellt, der 10 µg Gesamt-DNA enthielt. Anschließend wurde das Reagenz A im Verhältnis 1:1 zugegeben, und nach zehnminütiger Inkubation unter ständigem Schütteln Puffer B in doppelter Menge zupipettiert. Nach 15 Minuten konnten jeweils 480 µl des Transfektionsansatzes auf die Zellen gegeben werden. Je nach Zelltyp betrug die Einwirkzeit des Präzipitats 6

bis 24 h, danach wurde 3 mal mit PBS Puffer gewaschen und neues Medium auf die Zellen gegeben.

#### **3.5.2 Transfektion mit „Effectene“**

Die DNA-Aufnahme in die Zelle durch das „Effectene“-System (Qiagen), wird durch die Kondensation der DNA durch den Enhancer, und den Einschluss in vom „Effectene“ Reagenz gebildete Mizellen ermöglicht.

Die zu transfizierenden Zellen wurden in DMEM mit 10% FKS bis zu 80%-iger Konfluenz kultiviert. 0,4 µg DNA in TE-Puffer wurden mit dem DNA-Kondensationspuffer EC auf ein Gesamtvolumen von 100 µl gebracht. Nach Zugabe von 3,2 µl Enhancer Reagenz wurde 1 Sekunde gemischt und 2 bis 5 Minuten bei Raumtemperatur inkubiert. Danach wurden 10 µl „Effectene“ Transfektionsreagenz zugegeben, 10 sec gemischt und 5 bis 10 Minuten bei Raumtemperatur inkubiert, wobei sich der Transfektionskomplex bildete. Inzwischen wurden die Zellen einmal mit PBS gewaschen und das Medium gewechselt (1,6 ml). Nach der Inkubation wurden 0,6 ml FKS freies Zellkulturmedium zum Transfektionskomplex zugegeben und gemischt. Anschließend wurde die Mischung tropfenweise unter Schwenken auf die Zellen gegeben. Nach 6 Stunden wurde mit PBS gewaschen und das Medium gewechselt. Nach 48 h wurde die Transfektionseffizienz mittels Fluoreszenzmikroskopie bestimmt.

#### **3.5.3 Transfektion mit „Superfect“**

Das „Superfect“-Reagenz (Qiagen) besteht aus verzweigten Molekülen, die an ihren Enden positiv geladene Aminosäuren tragen. Die negativ geladene DNA wird an diese Enden gebunden. Der „Superfect“-DNA-Komplex (positive Nettoladung) bindet dann an negativ geladene Rezeptoren an der Zelloberfläche. Die zu transfizierenden Zellen wurden in DMEM-Medium mit 10% FKS bis zu 80%-iger Konfluenz kultiviert.

Die DNA (2µg) wurde durch Zugabe von Medium (ohne FKS) auf ein Volumen von 100 µl verdünnt. Danach wurden 3 µl, 7,5 µl und 15 µl „Superfect“ zugegeben und die Ansätze 10 Sekunden gemischt. Anschließend wurde 10-15 Minuten bei Raumtemperatur inkubiert, damit sich der Transfektionskomplex bilden konnte. Inzwischen wurde das Medium von den Zellen entfernt und einmal mit PBS gewaschen. Zu dem Transfektionsansatz wurden 600 µl Medium (mit FKS) zugesetzt, zweimal auf- und abpipetiert und anschließend 2-3 Stunden inkubiert. Danach wurde einmal mit PBS gewaschen und frisches Medium zugegeben. Nach 48 Stunden wurden die Zellen mittels Fluoreszenzmikroskopie untersucht.

#### **3.6 DNA-Isolierung aus Geweben**

DNA aus Gewebe wurde mittels „DNeasy Tissue Kit“ von Qiagen isoliert. Bis zu 25 mg Gewebe wurden in 180 µl Puffer ATL + 20 µl Proteinase K bei 55°C lysiert. Anschließend wurden 200 µl Puffer AL zugegeben, durchmischt und bei 70°C 10 min inkubiert. Nach Zugabe von 200 µl Ethanol wurde die DNA auf eine Anionenaustauschersäule gegeben und eine Minute zentrifugiert. Danach folgten zwei Waschschrte mit je 500 µl Puffer AW1 und AW2, und die Elution mit 200 µl Puffer AE.

### 3.7 Isolierung von Plasmid-DNA aus Bakterien

Zur Isolierung von Plasmid-DNA aus Bakterien wurden die Protokolle der Firma Qiagen angewendet. Hierfür wurden bei 37°C über Nacht gewachsenen Bakterien mittels SDS aufgeschlossen, und die Proteine anschließend durch saures Kaliumacetat gefällt. Nach Zentrifugation wurde der Überstand mit der DNA auf eine Anionenaustauschersäule (Qiagen GmbH) aufgetragen. Durch anschließende Waschschriffe wurde die DNA gereinigt und dann mittels Aqua bidest. eluiert.

#### 3.7.1 Plasmidmaxipräparationen (Maxipräp)

Hierfür wurden 3 ml LB Medium mit Ampicillin (50 µg/ml) mit einer Bakterienkultur beimpft und 6 bis 8 Stunden bei 37°C schüttelnd inkubiert. Vor Erreichen der stationären Phase wurde die Kultur in einen Kolben mit 500 ml LB-Medium + Ampicillin (50 µg/ml) überführt und über Nacht bei 37°C schüttelnd inkubiert. Am nächsten Tag wurde die Plasmidisolation nach Herstellerangaben durchgeführt. Die Bakterien suspension wurde 20 min bei 5000 g zentrifugiert und das Bakterienpellet in 10 ml Puffer 1 (enthält 100 µg/ml RNase A) resuspendiert. Aus einer 100 ml Bakterienkultur konnten zwischen 50-500 µg Plasmid-DNA isoliert werden. Anhand der OD<sub>260</sub> wurde die DNA-Konzentration ermittelt. Nach Sambrook (Sambrook, 1989) entspricht eine Absorption von 1 bei 260 nm einer DNA Konzentration von 50 µg/ml. Zusätzlich wurde zur Bestimmung der Reinheit der Probe der Quotient aus A<sub>260</sub> und A<sub>280</sub> gebildet, der bei einer Protein freien Probe zwischen 1,8 und 2 liegen sollte.

#### **3.7.2 Plasmidminipräparation (Minipräp)**

Jeweils 3 ml LB-Kulturen mit Ampicillin (50 µg/ml) wurden mit einer Bakterienkolonie angeimpft und über Nacht schüttelnd bei 37°C inkubiert. Anschließend wurden die Bakterienkulturen zentrifugiert und das Bakterienpellet durch eine alkalische SDS-Lösung aufgeschlossen. Proteine wurden durch saures Kaliumazetat ausgefällt. Nach der Zentrifugation wurde die DNA aus dem Überstand auf eine Anionenaustauschersäule aufgetragen. Durch zweimaliges Waschen wurde die gebundene Plasmid-DNA gereinigt und konnte anschließend mittels Aqua bidest. eluiert werden.

#### **3.8 Standard-PCR**

Folgende Komponenten wurden für die PCR pro Reaktionsansatz eingesetzt:

- 36 µl H<sub>2</sub>O
- 2 µl Template DNA (100ng)
- 5 µl 10xPCR-Puffer (Promega)
- 0,5 µl 25 mM dNTP-Mix
- 4 µl 25 mM MgCl<sub>2</sub>
- 2 µl Primer (120 pmol)
- 3 µl 5U/µl Taq-Polymerase (Promega)

bei der Kolonie-PCR wurde statt 2 µl Template-DNA jeweils eine Kolonie mit einer Pipettenspitze gepickt und dem Reaktionsansatz hinzugefügt.

### 3. Methoden

---

PCR-Programm für MLT-LTR und HERV-Lpro-Primer:

4 min 95°C  
30 sec bei 95°C }  
30 sec bei 68°C } 30 x  
1 min bei 72°C }  
7 min bei 72°C  
abkühlen auf 4°C

PCR-Programm für MLT-LTR und HERV-L pro-Primer mit angefügter *Bam*HI, bzw. *Hind*III-Schnittstelle:

4 min 95°C  
30 sec bei 95°C }  
30 sec bei 55°C } 10 x  
1 min bei 72°C }  
30 sec bei 95°C }  
30 sec bei 68°C } 20 x  
1 min bei 72°C }  
7 min bei 72°C  
abkühlen auf 4°C

PCR-Programm für GFP-Primer:

4 min 95°C  
30 sec bei 95°C }  
30 sec bei 68°C } 30 x  
1 min bei 72°C }  
7 min bei 72°C  
abkühlen auf 4°C

### 3. Methoden

---

PCR-Programm mit Luziferase-Primern:

4 min 95°C  
30 sec bei 95°C }  
30 sec bei 68°C } 30 x  
1 min bei 72°C }  
7 min bei 72°C  
abkühlen auf 4°C

### 3.9 Restriktionsverdau

Die Standardrestriktionsverdau wurden mit den Enzymen *Bam*HI und *Hind*III der Firma New England Biolabs (NEB) durchgeführt. Da in dieser Kombination auf Grund unterschiedlicher empfohlener Puffer die Gefahr von Sternaktivität besteht, wurden ausschließlich Einzelverdau angesetzt. Für die Restriktion mit *Bam*HI wurde der *Bam*HI-Puffer (NEB) verwendet und 2 Stunden bei 37°C inkubiert. Da vor dem zweiten Verdau ein Pufferaustausch nötig war, wurde eine Aufreinigung der DNA mittels Anionenaustauschersäulchen durchgeführt. Für die Restriktion mit dem Enzym *Bam*HI wurde der Puffer 2 (NEB) verwendet und bei 37°C über Nacht inkubiert. Nach Beendigung der Restriktion wurden die Enzyme 20 min bei 65°C hitzeinaktiviert.

Ansatz für den Restriktionsverdau (50 µl):

1 µg DNA  
5 µl 10 x Reaktionspuffer  
1 µl (10 units) Enzym  
ad 50 µl Aqua bidest.

### 3.10 Agarosegelelektrophorese

Die Agarosegelelektrophorese wurde zur Auftrennung von DNA-Fragmenten verwendet. Durch die negative Ladung der DNA lassen sich die Fragmente in einem elektrischen Feld in Agarosegelen der Größe nach auftrennen (Sambrook, 1989). Um die DNA sichtbar zu machen wurden die Gele mit Ethidiumbromid (10 µg pro 100 ml TAE-Gel) angefärbt. Ethidiumbromid interkaliert in den DNA-Doppelstrang und wird unter UV-Bestrahlung zur Fluoreszenz angeregt. Als Laufpuffer wurde TAE-Puffer verwendet. Die bei der Elektrophorese angelegte Spannung betrug 70-95 Volt. Um die DNA-Fragmentgrößen bestimmen zu können wurden als Längenstandard der 2-Log-Standard von New England Biolabs verwendet. Die untersuchten DNA-Fragmente hatten eine Größe von 0,4 bis 6 kb, daher wurden Gele mit einer Agarosekonzentration von 1,2 % (w/v) verwendet.

### 3.11 DNA-Isolierung aus Agarosegelen

Die DNA wurde unter UV-Licht mittels Skalpell ausgeschnitten, wobei die Einwirkzeit des UV-Lichts wegen der schädigenden Wirkung minimiert wurde. Die Elution der DNA aus dem Gel erfolgte nach der Vorschrift der Firma Qiagen, durch Zugabe des dreifachen Volumens eines chaotropen Puffers auf ein Gelvolumen. Inkubiert wurde bei 50°C bis zum vollständigen Lösen des Agarosegels. Nach Zugabe von Isopropanol wurde die DNA auf eine Anionenaustauschersäule gegeben und mehrfach gewaschen. Zur Elution der DNA wurden 30 µl Aqua bidest. verwendet.

### 3.12 Ligation

#### 3.12.1 TOPO TA Klonierung

Der im „TOPO TA Cloning Kit“ (Invitrogen, Karlsruhe) verwendete Vektor pCR®2.1-TOPO® liegt linearisiert vor und besitzt an den Enden jeweils einen überhängenden Thymidin (T) Rest, an den das Enzym Topoisomerase I kovalent gebunden ist. Die *Taq*-Polymerase verfügt über eine Terminaltransferase-Funktion, wodurch sie einen Adenosinrest (A) an das 3'-Ende des PCR-Produkts anhängt. Die Topoisomerase I ligiert PCR-Produkt und Vektor an den komplementären AT-Überhängen. Die Reaktion ist in Abbildung 7 dargestellt.

Reaktionsansatz:

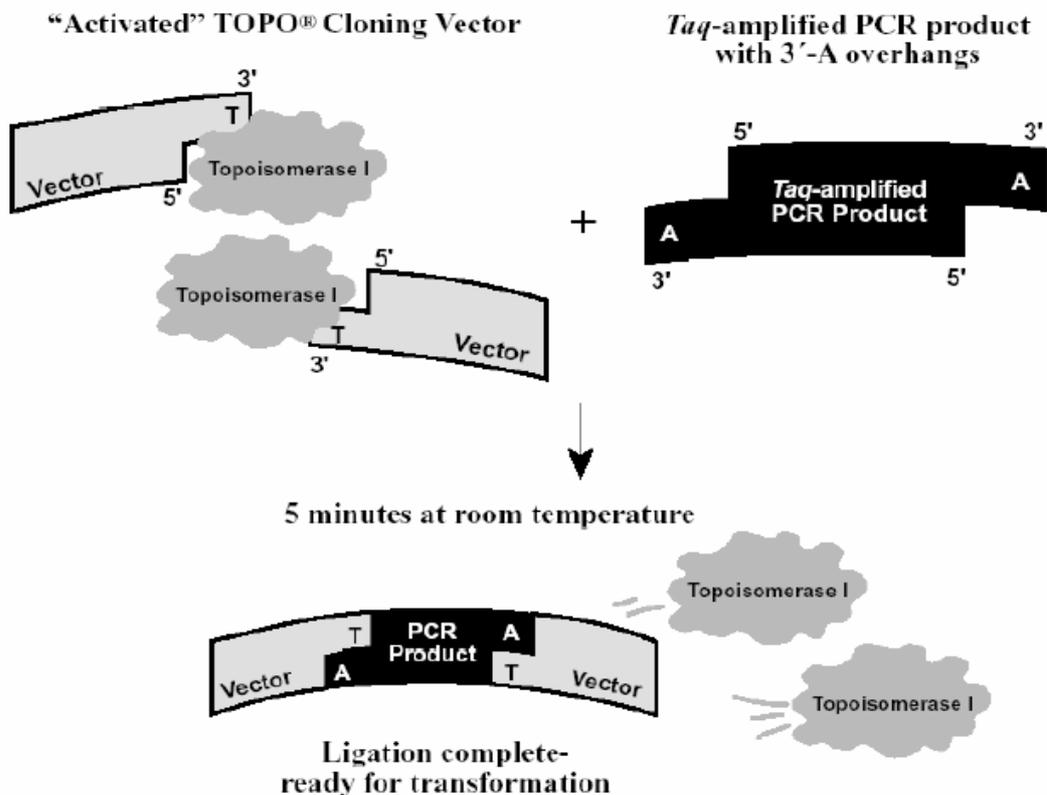
0,5-4,0 µl PCR-Produkt (mit ca. 10-100 ng DNA)

1 µl Salzlösung

ad 5,0 µl Aqua bidest.

1 µl pCR2.1-TOPO-Vektor

Der Reaktionsansatz wurde vorsichtig gemischt und 5-30 min bei RT inkubiert.



**Abb. 7** : Schematische Darstellung der TOPO-TA-Klonierung (Herstellerprotokoll; Invitrogen, Karlsruhe).

#### **3.12.2 Transformation von *E. coli* mit Plasmid-DNA**

Für die Transformation wurden ultrakompetente „XL-10 Epicurian Coli<sup>®</sup>“ (Stratagene) verwendet. Die Versuchsdurchführung erfolgte nach Herstellerangaben. Die Bakterien wurden in vorgekühlte Eppendorfgefäße zu jeweils 50 µl aliquotiert. Anschließend wurde jeweils 1 µl 1,4 molares β-Mercaptoethanol zugegeben und vorsichtig gemischt. Nach zehnminütiger Inkubation auf Eis wurde 1 µl Ligationsansatz zugegeben und 30 Minuten auf Eis inkubiert. Danach erfolgte ein Hitzeschock für 30 s bei 42°C. Nach zweiminütiger Inkubation auf Eis wurden 500 µl SOC-Medium zugegeben und bei 37°C eine Stunde inkubiert. Anschließend wurden 50 µl bis 200 µl Bakteriensuspension auf LB-Ampicillin-Platten (100 µg/ml)

ausplattiert. Nach über-Nacht-Inkubation bei 37°C wurden Kolonien gepickt, in 3 ml LB-Ampicillin-Medium (100 µg/ml) überführt und über Nacht bei 37°C schüttelnd inkubiert. Am nächsten Tag konnten die Übernachtskulturen für Miniplasmidpräparationen verwendet und die Plasmidinserts durch Restriktion oder PCR untersucht werden.

#### **3.13 „Site directed mutagenesis“**

Zur Einführung von Punktmutationen in Vektoren wurde das Protokoll der Firma Stratagene („Quick Change Site Directed Mutagenesis“) verwendet. Hierfür wurden Plasmide aus *dam*<sup>+</sup> *E. coli* Stämmen (methylierte Plasmide) verwendet. Es wurden zwei komplementäre Primer eingesetzt, die jeweils die gewünschten Punktmutationen enthielten und die Plasmide anschließend mit der *Pfu*-Polymerase amplifiziert. Die *Pfu*-Polymerase wurde verwendet, da sie im Vergleich zur *Taq*-Polymerase über eine deutlich höhere Synthesegenauigkeit verfügt. Danach wurden die methylierten, parentalen Plasmide mit dem Enzym *DpnI* verdaut. *DpnI* schneidet ausschließlich methylierte DNA. Übrig blieben die Punktmutationen tragenden, nicht ligierten Plasmide. Die Transformation erfolgte in XLII Blue superkompetenten Zellen und wurde mittels Hitzeschock durchgeführt (Abb. 8). Nach Inkubation bei 37°C über Nacht wurden die Plasmide isoliert. Um zu testen, ob die richtigen Punktmutationen vorhanden waren, wurden die Inserts der Plasmide sequenziert.

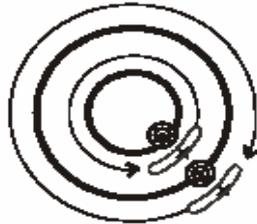
1. Plasmidpräparation



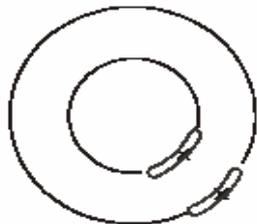
2. PCR



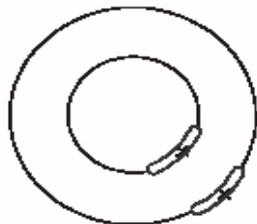
Mutation enthaltende  
Primer



3. Restriktionsverdau



4. Transformation



mutierte Plasmide  
(enthalten nichtligierte  
zirkuläre Stränge)

**Abb. 8:** Überblick über die „Quick change site-directed mutagenesis“ Methode (Stratagene).

#### **3.13.1 Synthese der mutierten Plasmide**

Für die Synthese der mutierten Plasmide wurden Oligonukleotide mit einer Länge von 25 bis 40 Basen verwendet. Die Oligonukleotide mussten zueinander komplementär sein und die Mutationen jeweils an der gleichen Stellen tragen.

Für die PCR-Reaktion wurden verwendet:

5 µl 10 x Reaktionspuffer

10 ng Plasmid-DNA

125 ng Primer 1

125 ng Primer 2

1 µl dNTP Mix

ad 50 µl H<sub>2</sub>O

2,5 µl *PfuTurbo*® DNA-Polymerase

PCR-Parameter für Punktmutationen:

95°C 30 s	} 12 x
95°C 30 s	
55°C 1 min	
68°C 6 min	

Für den *DpnI*-Verdau wurden 10 U (1 µl) des Restriktionsenzymes direkt zu den PCR-Ansätzen zugegeben. Danach wurde 1 h bei 37°C inkubiert.

#### **3.13.2 Transformation der XLII Blue superkompetenten Zellen**

Zu 50 µl der XLII Blue superkompetenten Zellen wurde jeweils 1 µl des *DpnI*-Verdau zugegeben, vorsichtig gemischt und 30 min auf Eis inkubiert. Anschließend wurde ein 45 s Hitzeschock bei 42°C durchgeführt und die Proben danach 2 min auf Eis

inkubiert (für die Transformation wurden nach Herstellerempfehlung 14 ml BD Falcon Polypropylen „Round-Bottom“ Röhrrchen verwendet). Dann wurden 0,5 ml NZY Broth (auf 42°C vorgewärmt) zugegeben und 1 h bei 37°C schüttelnd inkubiert. Anschließend wurden je 250 µl auf LB-Ampicillin-Platten ausplattiert und mindestens 16 Stunden bei 37°C inkubiert. Es wurden zwischen 10 und 1000 Kolonien erwartet.

#### **3.14 Sequenzierung**

Die Sequenzierungen wurden nach der Kettenabbruchmethode (Sanger, Nicklen und Coulson, 1977) von der Firma Sequiserve in Vaterstetten durchgeführt.

#### **3.15 RNA-Extraktion**

Auf gefrorene Gewebeproben wurden 500 bis 1000 µl RNA Bee (AMS) gegeben. Anschließend wurden die Proben mit einem elektrischen Homogenisator homogenisiert. Danach wurden 100 bis 200 µl Chloroform zugegeben und geschüttelt, anschließend wurde 15 min bei 4°C und 13000 rpm zentrifugiert. Der Überstand wurde in ein neues Röhrrchen überführt, die gleiche Menge Isopropanol zugegeben und 3 Stunden bei 4°C inkubiert. Danach wurden die Proben 5 min bei 4°C und 13000 rpm zentrifugiert. Der Überstand wurde verworfen und 200 µl 70%-iger Ethanol (in DEPC-Wasser) zugegeben. Anschließend wurden die Proben über Nacht bei -20°C gelagert. Am nächsten Tag wurden sie bei 4°C und 13000 rpm 5 min zentrifugiert. Der Überstand wurde verworfen und das Pellet in 50 µl DEPC-H<sub>2</sub>O resuspendiert.

Die Aufreinigung der RNA erfolgte gemäß der Vorgaben des Herstellers Qiagen. Es wurden ca. 10 µg Probe in einem Volumen von 100 µl mit 250 µl Ethanol und 350 µl RLT-Puffer vermischt. Dieser Ansatz (700 µl) wurde auf eine Anionenaustauschersäule aufgetragen und kurz zentrifugiert. Um die restliche vorhandene DNA zu entfernen, wurden die Säulen mit 350 µl Puffer RW1 beladen

### 3. Methoden

---

und kurz zentrifugiert. Anschließend wurden 10 µl DNase I Stammlösung mit 70 µl Puffer RDD vermischt, auf die Säule gegeben und 15 min bei Raumtemperatur inkubiert. Nach Zugabe von 350 µl Puffer RW1 und kurzer Zentrifugation wurden 500 µl RPE-Puffer hinzugefügt und zentrifugiert. Anschließend wurde mit 500 µl 80%-igem Ethanol gewaschen, zentrifugiert und die Säulen danach in ein neues 2 ml Gefäß überführt. Um den Ethanol komplett zu entfernen wurde ein weiterer fünfminütiger Zentrifugationsschritt (13000 rpm) durchgeführt. Für die Elution der RNA wurden 20 µl DEPC-Wasser auf die Säule aufgetragen und 1 min bei 13000 rpm zentrifugiert.

#### 3.16 cDNA-Synthese

Für die cDNA-Synthese wurde der „Superscript First-Strand Synthesis System for RT-PCR Kit“ (Invitrogen) verwendet. Es wurde jeweils 1 µg Probe eingesetzt. Folgende Komponenten (Tab. 3) wurden eingesetzt:

**Tabelle 3:** Pipettierschema für RT-PCR

<b>Komponente</b>	<b>Probe</b>	<b>NoRT Kontrolle</b>	<b>Kontroll RNA</b>
1 µg RNA	n µl	-	-
Kontroll RNA (50 µg/µl)	-	-	1 µl
Random Hexamers (50 µg/µl)	2 µl	2 µl	2 µl
10 mM dNTP Mix	1 µl	1 µl	1 µl
DEPC H <sub>2</sub> O	Auf 10 µl	7 µl	6 µl

### 3. Methoden

---

Die Proben wurden bei 65°C für 5 min inkubiert und anschließend 1 min auf Eis abgekühlt. Jeweils 2 µl 10x RT-Puffer, 4 µl 25 mM MgCl<sub>2</sub>, 2 µl 0,1 mM DTT, 1 µl „RNaseOUT“ wurden zugegeben, gemischt und bei 25°C 2 min inkubiert. Nach Zugabe von 1 µl (50 units) „SuperScriptII“ reverser Transkriptase, wurden die Proben weitere 10 min bei 25°C und anschließend 50 min bei 42°C inkubiert. Nachdem die Reaktion nach 15 minütigem Erhitzen auf 70°C abgestoppt worden war, wurde der Reaktionsansatz auf Raumtemperatur heruntergekühlt. Danach wurde 1 µl RNaseH zugegeben und 20 min bei 37°C inkubiert. Anschließend wurde die cDNA mittels PCR amplifiziert.

#### **3.17 Proteinkonzentrationsbestimmung mit Bicinchioninsäure**

Zur Bestimmung der Gesamtproteinkonzentration in Zelllysaten wurde der Bicinchioninsäure(BCA)-Test von Pierce verwendet. Hierbei wird die Farbreaktion (Biuret-Reaktion) die bei der Reduktion von  $\text{Cu}^{2+}$  zu  $\text{Cu}^{1+}$  im alkalischen Milieu durch Proteine katalysiert wird, ausgenutzt. Die Farbreaktion entsteht durch Chelatkomplexbildung der Proteine mit den Kupferionen. Die Chelatkomplexe absorbieren Licht der Wellenlänge 562 nm. Die Absorption ist proportional zur Proteinkonzentration. Meßbar sind 20 bis 2000  $\mu\text{g}$  Protein. Die Messungen wurden in 96-Loch-Mikrotiterplatten durchgeführt, es wurden

jeweils 20  $\mu\text{l}$  Kernlysate eingesetzt. Die Proteinkonzentration wurde anhand einer BSA-Eichkurve bestimmt (BSA-Stammlösung: 2 mg/ml).

#### **3.18 „Dual Luziferase® Reporter Assay“**

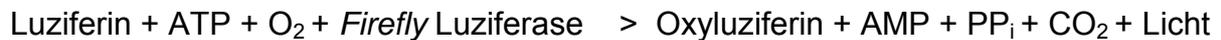
Der „Dual Luziferase® Reporter Assay“ (Promega) wurde verwendet um Promotoraktivitäten von HERV-LTRs zu untersuchen. Hierfür wurden die LTRs jeweils in den pBL-Vektor kloniert, der als Reporter das Luziferasegen von *Photinus pyralis* enthält. Als interner Standard wurde das Plasmid pRL-TK verwendet, das das aus *Renilla reniformis* stammende Luziferasegen trägt. Der pRL-TK-Vektor enthält einen konstitutiv aktiven, minimalen TK-Promotor. Die Plasmide pBL und pRL-TK wurden im Verhältnis 50:1 in die untersuchten Zellen cotransfiziert.

### 3. Methoden

---

Folgende Reaktionen werden durch die Luziferasen katalysiert:

#### Firefly Luziferase



#### Renilla Luziferase



Zur Bestimmung der Luziferaseaktivitäten in Zelllysaten, wurden die Zellen 48 h nach der Transfektion für 20 min bei Raumtemperatur in 500 µl Lysepuffer schüttelnd inkubiert. Anschließend wurden 50 µl des Lysates in weiße 96-Lochplatten (NUNC) pipettiert und die Luziferaseaktivitäten der *Photinus*- und *Renilla*-Luziferase in einem „ORION“ Luminometer (Berthold) gemessen. Hierzu wurden zu jeder Lysatprobe 100 µl des *Photinus*-Substrates LAR II zugegeben, die Luziferaseaktivität gemessen und anschließend das „Stop and Glo“-Reagenz zugemischt. Letzteres blockiert einerseits die Aktivität der *Photinus*-Luziferase, und enthält andererseits das Substrat für die *Renilla*-Luziferase. Danach wurde die *Renilla*-Luziferaseaktivität gemessen. Um die relative Luziferaseaktivität zu bestimmen, wurde der Zahlenwert der *Photinus*-Luziferaseaktivität durch den Zahlenwert der *Renilla*-Luziferaseaktivität dividiert. Außerdem wurde der Quotient aus dem Zahlenwert der relativen Luziferaseaktivität der jeweils untersuchten Promotoren und der relativen Luziferaseaktivität der Negativkontrolle (pBL-Leervektor) gebildet. Die Ergebnisse wurden mit Hilfe des Programms Excel als Säulendiagramme dargestellt und geben den Faktor an, um den die relative Luziferaseaktivität der Probe sich von dem der Negativkontrolle unterscheidet. Die Ergebnisse repräsentieren jeweils die Mittelwerte aus 4 Transfektionsansätzen die zwei unabhängigen Transfektionen entstammen.

#### **3.19 Kernextraktpräparation**

Die für die Kernextraktion verwendeten Zellen wurden in Kulturflaschen bis zu 90%-iger Konfluenz hochgezüchtet. Um das Denaturieren von Proteinen zu verhindern, wurden alle für die Kernextraktion verwendeten Puffer und Gefäße auf Eis gekühlt, und alle Arbeitsschritte auf Eis und/oder im Kühlraum durchgeführt. Die adhärennten Zellen wurden in der Kulturflasche 3 mal mit PBS gewaschen. Anschließend wurden 10 ml Puffer A zugegeben und 10 min auf Eis inkubiert. Dann wurde der Puffer entnommen und 2 ml Puffer A + NP40 zugegeben. Nach 10 minütiger Inkubation wurden die Zellen mit einem Zellschaber vom Boden der Kulturflasche entfernt und in ein vorgekühltes 10 ml Falcon-Röhrchen überführt. Nach einem weiteren Waschschrift mit 1 ml Puffer A + NP40, wurde 10 min bei 4°C zentrifugiert. Danach wurde der Überstand verworfen, das Pellet in 1 ml Puffer C resuspendiert, und anschließend 10 min bei 4 °C zentrifugiert. Nach Verwerfen des Überstandes wurde das Pellet im fünffachen Volumen Puffer C + 0,4 M NaCl resuspendiert und 30 min auf Eis inkubiert. Nach 10 minütiger Zentrifugation wurde der Überstand in vorgekühlte Eppendorfgefäße überführt.

#### **3.20 „Electrophoretic Mobility Shift Assay“ (EMSA)**

EMSAs wurden zur Untersuchung von potentiellen Transkriptionsfaktorbindungsstellen innerhalb der HERV-L-LTR eingesetzt. Für den „Electrophoretic Mobility Shift Assay“ wurden Kernextrakte aus Zelllinien verwendet in denen die untersuchte DNA-Sequenz Promotoraktivität zeigt. Ein radioaktiv markiertes Oligonukleotid, das eine potentielle Transkriptionsfaktorbindungsstelle enthält, wurde als Sonde verwendet. Bindet ein Protein an die DNA-Sonde, wandert der DNA-Proteinkomplex langsamer durch ein nichtdenaturierendes Polyacrylamidgel als die freien Oligonukleotide. Eine Protein-DNA-Bindung kann somit als „Shift“ auf einem Röntgenfilm detektiert werden.

Die EMSA-Untersuchungen enthielten jeweils folgende Reaktionsansätze: 1. Die Reaktion der DNA-Sonde mit einem Kernextrakt (alle im Folgenden beschriebenen Reaktionsansätze enthielten auch die DNA-Sonde und den Kernextrakt). 2. Zugabe eines Wildtypkompetitors. Dieser besitzt die gleiche Sequenz wie die radioaktiv markierte DNA-Sonde, ist jedoch nicht radioaktiv markiert. Der Kompetitor wird im Vergleich zur Sonde in 100-fachem Überschuß zugegeben. Handelt es sich bei einem „Shift“ um spezifisch an die DNA-Sonde bindende Proteine, werden diese durch den im Überschuß vorhandenen Kompetitor weggefangen, wodurch der „Shift“ nicht mehr detektierbar ist. 3. Zugabe (100-facher Überschuß) von mutierten, inaktiven Kompetitoren. Diese entsprechen, abgesehen von einigen Punktmutationen, die die zu untersuchende Transkriptionsfaktorbindungsstelle inaktivieren (durch die Software „Sequence Shaper“ von „Genomatix“ bestätigt), der DNA-Sonde. Durch Konkurrenz mit mutierten Kompetitoren im Überschuß kann gezeigt werden, ob die untersuchte potentielle Transkriptionsfaktorbindungsstelle für den „Shift“, und damit für die Proteinbindung verantwortlich ist. 4. Um weitere Informationen über ein spezifisch bindendes Protein zu erhalten, wurden Konkurrenzreaktionen mit kommerziell erworbenen Konsensus-Oligonukleotiden durchgeführt. Diese enthalten eine Konsensus-Sequenz, die sich von den TF-Bindungsstellen der jeweiligen Transkriptionsfaktorfamilie ableitete. Durch Konkurrenz mit diesen Konsensus-Oligonukleotiden konnte ein Protein einer bestimmten Transkriptionsfaktorfamilie zugeordnet werden. 5. Zugabe von mutierten, inaktivierten Konsensus-Oligonukleotiden (100-facher Überschuß). Ein spezifisch an eine bestimmte Konsensus-Sequenz bindender Transkriptionsfaktor bindet nicht an diesen mutierten Kompetitor (siehe 4.), der „Shift“ bleibt unbeeinflusst. 6. Konnte durch die Konkurrenzreaktionen die Bindung eines Transkriptionsfaktors an die untersuchte potentielle Transkriptionsfaktorbindungsstelle festgestellt werden, und durch die Konsensus-Konkurrenzen auch die Zugehörigkeit zu einer TF-Familie ermittelt werden, wurde, um den Faktor exakt zu identifizieren, eine „Supershift“-Untersuchung durchgeführt. Dazu wurde dem Reaktionsansatz ein für einen Transkriptionsfaktor spezifischer Antikörper hinzugefügt. Bindet der Antikörper das am „Shift“ beteiligte Protein, kommt es zum „Supershift“. Dabei wird ein weiterer „Shift“ oberhalb des Transkriptionsfaktor-DNA-Komplexes (ursprünglicher „Shift“)

detektierbar, was durch das höhere Molekulargewicht des Antikörper-Transkriptionsfaktor-Oligonukleotid-Komplexes zustande kommt. Im Falle einer sterischen Hinderung der Transkriptionsfaktor-DNA-Bindung durch Bindung des Antikörpers an die DNA-Bindedomäne des Transkriptionsfaktors sind keine „Shifts“ detektierbar. In diesem Fall handelt es sich um ein „Antikörper-Blocking“.

#### **3.20.1 Hybridisierung der Einzelstrang Oligonukleotide**

Die einzelsträngigen Oligonukleotide wurden auf eine Konzentration von 100 pmol/μl eingestellt. Die Hybridisierung wurde bei einer Salzkonzentration von 0,4 M NaCl durchgeführt.

Annealing-Ansatz:

6 μl Oligomer 1 (100 pmol/μl)  
6 μl Oligomer 2 (100 pmol/μl)  
1,6 μl NaCl (5M)  
6,4 μl H<sub>2</sub>O  
20 μl Gesamtvolumen

#### **3.20.2 Radioaktive Markierung der Doppelstrang-Oligonukleotide** **(Blunt End-Markierung)**

Oligonukleotide mit glatten Enden wurden mittels Polynukleotid-Kinase (Phage T4) durch Übertragung einer radioaktiven Phosphatgruppe markiert.

### 3. Methoden

---

Kinase-Ansatz:

2 µl doppelsträngige Oligonukleotide (3 pmol/µl)

3 µl 10 x PNK Puffer

4 µl gamma <sup>32</sup>-P-ATP

2 µl T4-PNK (10 U/µl)

19 µl H<sub>2</sub>O

30 µl Gesamtvolumen

Inkubation 1 h, 37°C

Für die Aufreinigung der Proben wurden „Nuc Trap Probe Purification Columns“ (Stratagene) verwendet. Dabei handelt es sich um Gelfiltrationssäulen, in deren Gelmatrix kleinere Moleküle eindringen können. Größere Moleküle können nicht in die Gelporen diffundieren und wandern daher schneller durch die Säule. Diese Methode ermöglicht die Trennung der freien Phosphatgruppen von den radioaktiv markierten Oligonukleotiden.

Die Säulen wurden mit 80 µl STE-Puffer (TE + 0,1M NaCl) befeuchtet. Anschließend wurden 30 µl Probe auf die Säule gegeben und danach 80 µl STE-Puffer aufgetragen. Die markierten Oligonukleotide wurden in einem Falcon-Röhrchen aufgefangen. Anschließend wurden die Proben in einem Szintillationszähler gemessen und auf 50000 counts pro µl eingestellt.

#### **3.20.3 Reaktionsansätze und Polyacrylamidgelelektrophorese**

Folgende Komponenten wurden vor dem Gellauf vorinkubiert:

2,0 µl 1 mg/ml dIdC

0,4 µl 2M NaCl

1,5 µl 10 x BSB

2,0 µl Kernextrakt (3 µg/µl)

6,9 µl H<sub>2</sub>O

Kompetitoren, bzw. Antikörper wurden anschließend zugegeben. Zuletzt wurde jeweils 1 µl der radioaktiv markierten Oligonukleotide (Hotprobe; 50000 counts) hinzugefügt. Der Reaktionsmix wurde danach entweder 1h auf Eis inkubiert, oder 30 min bei Raumtemperatur, bevor er auf ein 5%-iges, nicht-denaturierendes Polyacrylamidgel aufgetragen wurde.

nicht-denaturierende Polyacrylamidgele (60 ml, 5%):

10 ml (30:0.8) Acrylamid (30%)

3 ml 10xTBE (0,5xEndkonzentration)

47 ml Wasser

500 µl TEMED

Als Laufpuffer wurde 0,5xTBE verwendet (ca. 2 Liter). Der Gellauf wurde bei einer Stromspannung von 240 V ca. 1,5 h durchgeführt. Danach wurde das Gel auf „Whatman“-Papier übertragen und in Plastikfolie verpackt 1 h bei 80°C Vakuum getrocknet. Zur Auswertung wurde entweder ein Röntgenfilm in einer Expositions-kammer über Nacht bei -80°C exponiert, oder eine „Phosphoimager“-Platte 2 bis 7 Tage bei Raumtemperatur.

#### **3.21 Computeranalysen**

Für die bioinformatischen Untersuchungen von LTRs bezüglich ihrer Promotoraktivität wurden folgende Analyseprogramme der „GenomatixSuite“ („Genomatix Software“ GmbH) verwendet:

##### **3.21.1 Matinspector**

Dieses Computerprogramm wurde verwendet, um potentielle Transkriptionsfaktoren innerhalb einer DNA-Sequenz zu ermitteln. Dem Programm zu Grunde liegen „Weight Matrices“ einer Vielzahl von Transkriptionsfaktorbindungsstellen. Diese

„Weight Matrices“ basieren auf Alignments von funktionellen Transkriptionsfaktorbindungsstellen (Matrix). Dabei wird die Häufigkeit von Nukleotiden an einer bestimmten Position innerhalb einer Transkriptionsfaktorbindungsstelle bestimmt und als Zahlenwert ausgedrückt. Die zu untersuchende DNA-Sequenz wird mit „Weight Matrices“ aller in der Genomatix-Datenbank vorhandenen Transkriptionsfaktorbindungsstellen verglichen und bei Übereinstimmung diese angegeben. Entscheidend dabei sind vor allem die „Core Similarity“ und die „Matrix Similarity“. Die „Core Similarity“ bezieht sich auf die „Core Sequence“ einer Transkriptionsfaktorbindungsstelle, die die vier am höchsten konservierten, aufeinanderfolgenden Nukleotid-Positionen innerhalb einer Matrix enthält. Dabei beschreibt ein Zahlenwert von 1,0 die vollständige Übereinstimmung mit der „Core Sequence“ einer bestimmten Transkriptionsfaktorbindungsstelle. Die „Matrix Similarity“ besitzt den Wert 1,0, wenn jede Position der untersuchten Sequenz dem am höchsten konservierten Nucleotid der Matrix einer Transkriptionsfaktorbindungsstelle entspricht. Eine gute Übereinstimmung mit der Matrix hat einen Zahlenwert von  $>0.8$ . Abweichungen in hoch konservierten Positionen der Matrix verringern die „Matrix Similarity“ stärker als Abweichungen in weniger konservierten Bereichen.

#### **3.21.2 Sequence Shaper**

Das Programm erkennt vorhandene Transkriptionsfaktorbindungsstellen innerhalb einer untersuchten DNA-Sequenz und liefert Vorschläge, wie diese durch Einfügen von Punktmutationen inaktiviert, oder neue Transkriptionsfaktorbindungsstellen eingefügt werden können. Anhand der Ergebnisse dieses Programmes können Oligonukleotide designt werden, die z.B. als mutierte Kompetitoren in EMSAs verwendet werden.

#### **3.21.3 Frameworker**

Mit diesem Programm können in einer beliebigen Anzahl von DNA-Sequenzen übereinstimmende potentielle Transkriptionsfaktorbindungsstellen (Elemente) ermittelt werden. Dazu kann ausgewählt werden, wie hoch der Anteil der Sequenzen sein soll, in denen die gleichen Elemente vorhanden sind (zwischen 8 und 100%). Dadurch können in verwandten DNA-Sequenzen (z.B. HERV-L LTRs) konservierte,

### 3. Methoden

---

potentielle Transkriptionsfaktorbindungsstellen, bzw. Modelle von Elementen identifiziert werden. Unter Modellen sind Kombinationen von potentiellen Transkriptionsfaktorbindungsstellen zu verstehen, die aus zwei oder mehr Elementen bestehen.

## 4 Ergebnisse

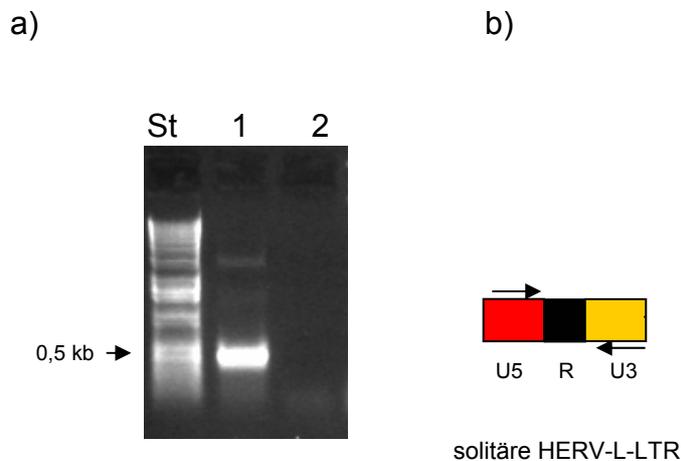
### 4.1 Isolierung und Charakterisierung von HERV-L-LTRs

Die Promotoraktivität der 5'-LTR des von Cordonnier *et al.* (1995) zuerst beschriebenen HERV-L(Cord) Provirus (NCBI Accession Nummer G895836) wurde bereits in den Zelllinien LC5 (Lungenkarzinom), U373 (Astrozytoma) , 85HG66 (Astrozytoma), MiaPaCa2 (Pankreaskarzinom), HaCaT (immortalisierte Keratinozyten), HeLa (Zervixkarzinom), MCF7(Brustkarzinom) sowie in den primären Keratinozyten NHEK getestet. Dabei konnte Promotoraktivität in der Zelllinie HaCaT und in den primären Keratinozyten NHEK nachgewiesen werden (Schön *et al.*, 2001). In der vorliegenden Arbeit wurden weitere HERV-LTRs aus dem humanen Genom isoliert und ihre Promotoraktivität und -spezifität untersucht.

#### 4.1.1 Phylogenetische Analyse der HERV-L-LTRs

##### 4.1.1.1 Solitäre HERV-L-LTRs

Für die Klonierung solitärer HERV-L-LTRs wurde DNA aus der Zelllinie HeLa isoliert und anschließend mit den HERV-L-LTR-spezifischen Primern MLT (Material 2.2.1) amplifiziert. Die Auswahl der Primer erfolgte durch Vergleich mehrerer HERV-L-LTR Sequenzen, die aus der HERVd-Datenbank (Paces *et al.*, 2002) entnommen wurden. Es wurden hierfür Sequenzen aus den HERV-L-Unterfamilien MLT2A1, MLT2A2, MLT2B1, MLT2B2 verwendet (Kapitel 1.8). Das humane Genom enthält ca. 200 provirale HERV-L-Sequenzen und etwa 6000 solitäre LTRs. Da die Primer so gewählt wurden, dass sie innerhalb der HERV-L-LTR binden, sollten damit vorzugsweise solitäre LTRs isoliert werden.



**Abb. 9:** a) Amplifikation solitärer HERV-L-LTRs aus HeLa-DNA. St: Längenstandard; 1: Probe mit HeLa-DNA; 2: Negativkontrolle (Wasser). b) Schematische Darstellung einer solitären HERV-L-LTR mit den verwendeten MLT-Primern (Kapitel 2.2.3).

Die mit den MLT-Primern (Kapitel 2.2.3) durchgeführte PCR (Kapitel 3.8) lieferte Amplifikate der erwarteten Größe von ca. 480 bp (Abb. 9). Nach Aufreinigung (PCR Purification Kit; Nucleobond) wurden die Amplifikate kloniert (TOPO TA Cloning, Invitrogen), die erhaltenen Klone mittels Kolonie-PCR getestet, und die Inserts der positiven Plasmide sequenziert. Zur Identifizierung der Sequenzen wurden sie mit der Nukleotid-Datenbank von NCBI verglichen. Um zu untersuchen, ob es sich tatsächlich um solitäre LTRs handelt, wurden die angrenzenden Bereiche der in der Datenbank gefundenen Sequenzen analysiert. Die klonierten solitären HERV-L-LTRs, ihre chromosomale Zuordnung und NCBI-Accession-Nummern sind in Tabelle 4 aufgeführt.

#### 4. Ergebnisse

---

**Tabelle 4:** Klonierte solitäre HERV-L-LTRs

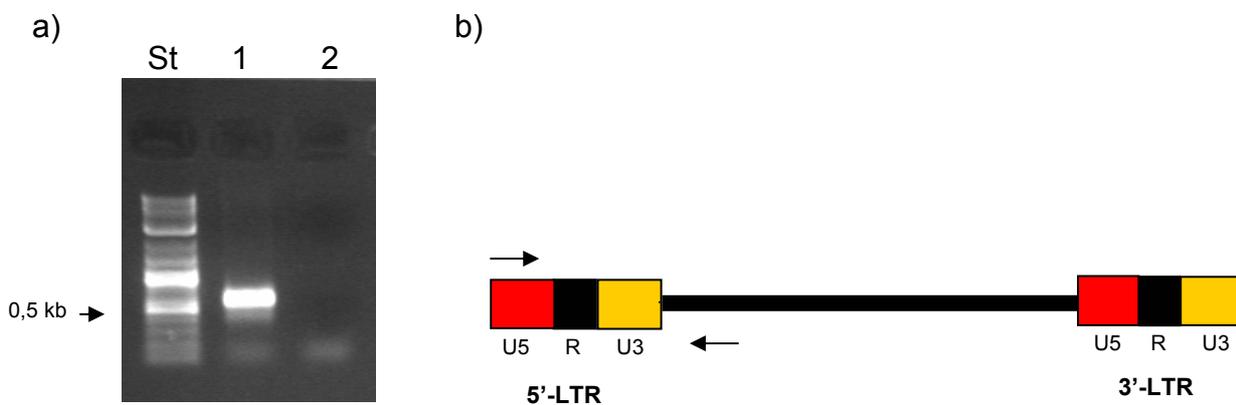
<b>solitäre HERV-L-LTR</b>	<b>Chromosom</b>	<b>NCBI-Accession-Nummer</b>
S 1	1	AL133420
S 2	X	AL049588
S 3	3	AC103549
S 4	4	AC098583
S 5	10	AL390862
S 6	3	AC103549
S 7	10	AL390862
S 8	1	AL035705
S 9	12	AL133420
S 10	10	AL390862
S 11	X	AL732596
S 13	6	AL590428
S 14	11	AC103855
S 15	2	AC079402

## 4. Ergebnisse

---

### 4.1.1.2 Provirale HERV-L-LTRs

Zur Isolierung proviraler HERV-L-LTRs wurden die Primer HERV-L pro (Kapitel 2.2.3) verwendet. Der 3'-Primer bindet im proviralen Bereich der HERV-L-Proviren. Als Template wurde Hela-DNA verwendet (Abb. 10).



**Abb. 10:** a) Amplifikation proviraler 5'-HERV-L-LTRs aus Hela-DNA. St: Längenstandard; 1: Amplifikat mit erwarteter Größe von ca. 700 bp; 2: Negativkontrolle (Wasser). b) Schematische Darstellung eines HERV-L- Provirus mit Bindungsstellen der HERV-L pro Primer (Kapitel 2.2.3).

Nach der Klonierung der ca. 700 bp langen PCR-Produkte (TOPO TA Klonierung) in *E. coli*-Zellen, wurden fünf Klone, die zuvor mittels Kolonie-PCR getestet wurden sequenziert und konnten als 5'-LTRs proviraler HERV-L-Elemente identifiziert werden (Tab. 5).

#### 4. Ergebnisse

---

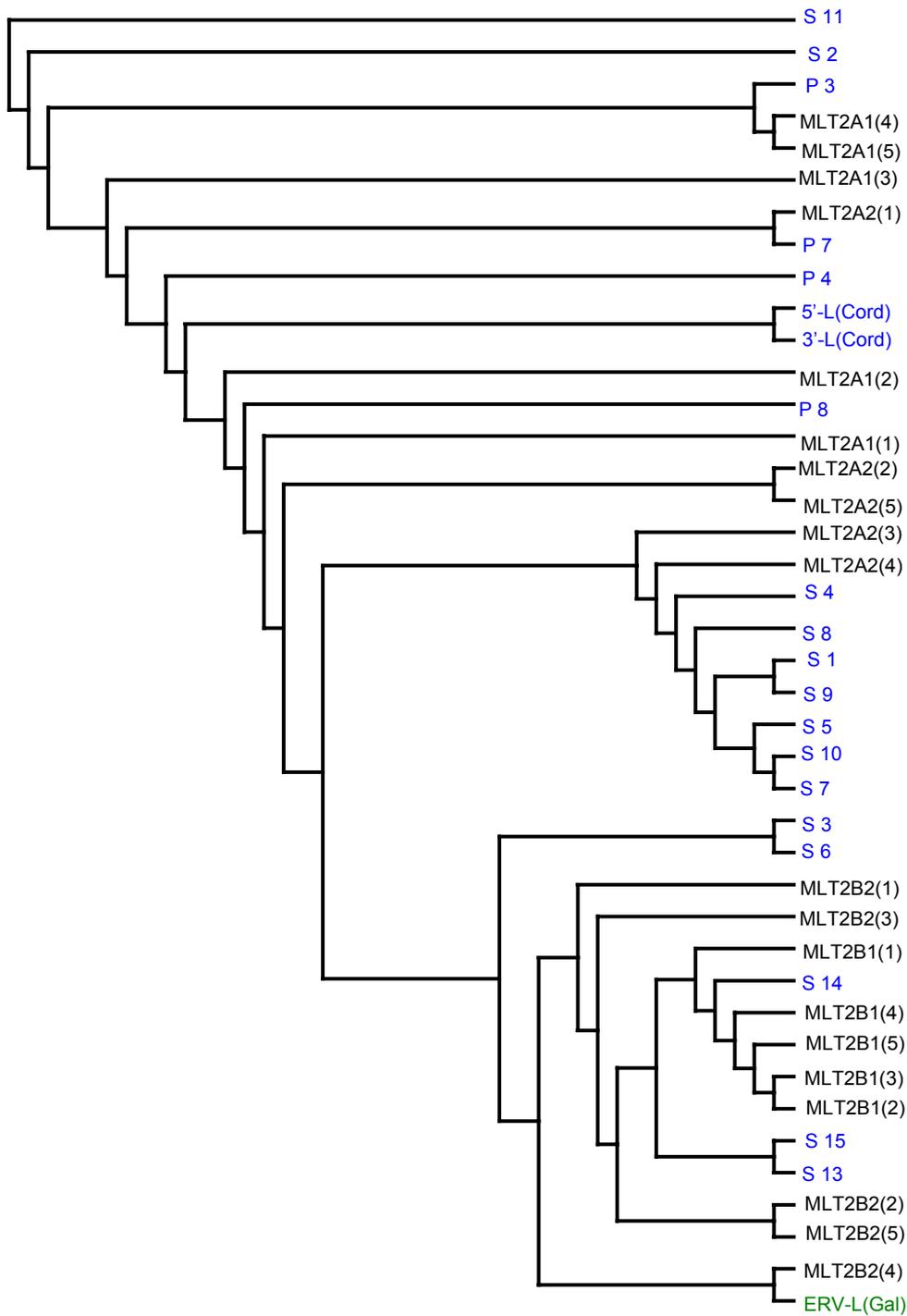
**Tabelle 5:** Klonierte provirale HERV-L-LTRs

<b>provirale HERV-L-LTR</b>	<b>Chromosom</b>	<b>NCBI-Accession-Nummer</b>
P 3	7	AC005076
P 4	16	AC092341
P 7	21	AP001604
P 8	8	AC025182
P 10	2	AC013402

Zur phylogenetischen Analyse wurde die Länge der proviralen LTR-Sequenzen an die solitären LTRs angepasst (ca. 480 bp). Als Vergleichssequenzen dienten die 5'- und die 3'-LTR des von Cordonnier *et al.* (1995) isolierten Provirus (5'-HERV-L-LTR (Cord), bzw. 3'-HERV-L-LTR(Cord)), sowie weitere aus der HERVd-Datenbank (Kapitel 2.12) entnommene HERV-L-LTRs, die den HERV-L-Subgruppen MLT2A1, MLT2A2, MLT2B1 und MLT2B2 entstammen. Ein Alignment wurde mit Hilfe der Software „ClustalX“ (Higgins und Sharp, 1988; Thompson *et al.*, 1997) erstellt (nicht gezeigt), die phylogenetische Analyse wurde mit dem Programm „Paup“ (Swafford, 1998) nach der „Neighbour Joining“-Methode (Saitou und Nei, 1986; Saitou und Nei, 1987) durchgeführt (Abb. 11).

## 4. Ergebnisse

---



**Abb. 11:** Phylogenetische Analyse der verwendeten HERV-L-LTR-Sequenzen: Enthalten sind jeweils fünf provirale LTR-Sequenzen der Familien MLT2A1 (1 bis 5), MLT2A2 (1 bis 5), MLT2B1 (1 bis 5) und MLT2B2 (1 bis 5) entnommen aus der HERVd-Datenbank, die 14 solitären HERV-L-LTRs S 1 bis S 15, die 5 proviralen 5'-LTRs P 3, P 4, P 7, P 8, P 10 und die ERV-L(Gal). Der Stammbaum wurde mit dem Programm „Paup“ nach dem „Neighbour Joining“-Verfahren erstellt. HERVd-Accession-Nummern der MLT-Sequenzen sind dem Kapitel 2.13 zu entnehmen, NCBI-Accession-Nummern der solitären LTRs (S) der Tabelle 4, der proviralen LTRs (P) der Tabelle 5.

## 4. Ergebnisse

---

Die phylogenetische Untersuchung (Abb. 11) zeigt, dass die solitären HERV-L-Elemente S 11 und S 2 am nächsten mit LTRs der Unterfamilie MLT2A1 verwandt sind. Die provirale LTR P 3 ist eindeutig der Unterfamilie MLT2A1 zuzuordnen. Die 5'-HERV-L-LTR(Cord) ist am nächsten mit der 3'-HERV-L-LTR(Cord) verwandt. Beide Elemente sind, wie auch die proviralen LTRs P 4 und P 8, der Unterfamilie MLT2A1 zuzordnen. Die Elemente P 7, S 4, S 8, S 1, S 9, S 5 und S 10 gehören zur Unterfamilie MLT2A2. Die LTRs S 3 und S 6 sind am nächsten mit der Unterfamilie MLT2B2 verwandt. ERV-L(Gal) lässt sich eindeutig in die Unterfamilie MLT2B2 einordnen. Die solitären LTRs S 13 und S 15 sind am nächsten mit der Unterfamilie MLT2B1 verwandt. Das Element S 14 ist der Unterfamilie MLT2B1 zuzuordnen.

### **4.1.2 Klonierung der HERV-L-LTRs in den pBL-Expressionsvektor**

Um zu testen, ob die klonierten HERV-L-LTRs Promotoraktivität besitzen, wurden Expressionsuntersuchungen durchgeführt. Hierfür mußten sie in den Expressionsvektor pBL, der ein Luziferase-Reportergen enthält, umkloniert werden. Zur Amplifikation der im pCR TOPO TA Vektor vorliegenden HERV-L-LTRs wurden Primer mit zusätzlich angefügter *Bam*HI- bzw. *Hind*III-Schnittstelle verwendet (für die solitären LTRs „MLT LTR *Bam*HI for / MLT LTR *Hind*III rev“; für die proviralen LTRs „HERV-L *Bam*HI for / HERV-L *Hind*III rev“). Die PCR-Produkte wurden mit den Restriktionsenzymen *Bam*HI und *Hind*III geschnitten und die Inserts nach der Klonierung mittels Kolonie-PCR überprüft. Nach mehreren Anläufen konnten die Klone mit der richtigen Insertgröße (ca. 480 bp) ermittelt werden. Mit diesen Klonen wurden dann Maxi-plasmidpräparationen durchgeführt, um die Plasmide in ausreichender Menge für die Transfektion zu erhalten. Die Inserts der Plasmide wurden sequenziert und konnten als LTRs S 1, S 4, S 5, S 11, bzw. P 3, P 4, P 7 und P 8 identifiziert werden.

### **4.2 Promotoraktivität von HERV-L-LTR-Sequenzen in humanen Zellen**

In den folgenden Versuchen wurden die isolierten 5'-HERV-L-LTRs (P) und die solitären HERV-L-LTRs (S) sowie die 5'- und 3'-HERV-L-LTR(Cord) in verschiedenen Zelllinien bezüglich ihrer zelltypspezifischen Aktivität untersucht.

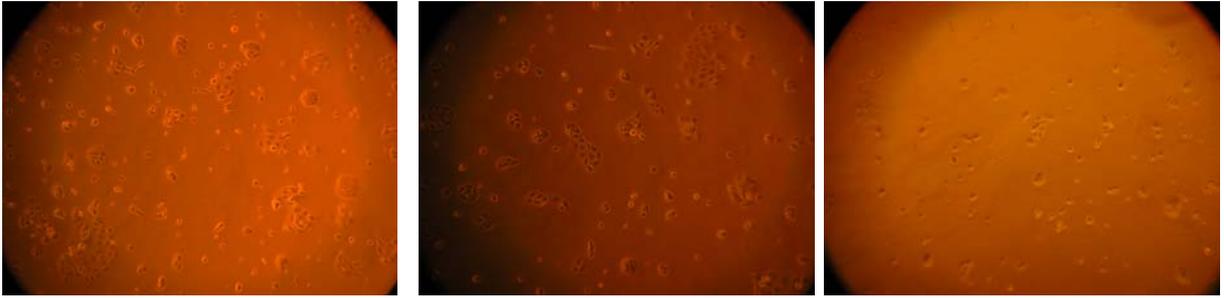
#### **4.2.1 Optimierung der Transfektion von HaCaT-Zellen**

Da für die Promotoruntersuchungen von HERV-L-LTRs u.a. die schwer zu transfizierende Keratinozytenzelllinie HaCaT verwendet werden sollte, wurden Tests mit den Transfektionsreagentien „Superfect“, „Cellpfect“ und „Effectene“ durchgeführt, um die Transfektionseffizienz zu optimieren. Die verwendeten HaCaT-Zellen wurden bei ca. 60%-iger Konfluenz transfiziert und nach 48 Stunden unter dem UV-Mikroskop untersucht. Zur Transfektion wurde das Plasmid pFred143, das als Reporter das GFP-Gen unter der Kontrolle des CMV-Promotors enthält, verwendet (Stauber *et al.*, 1998).

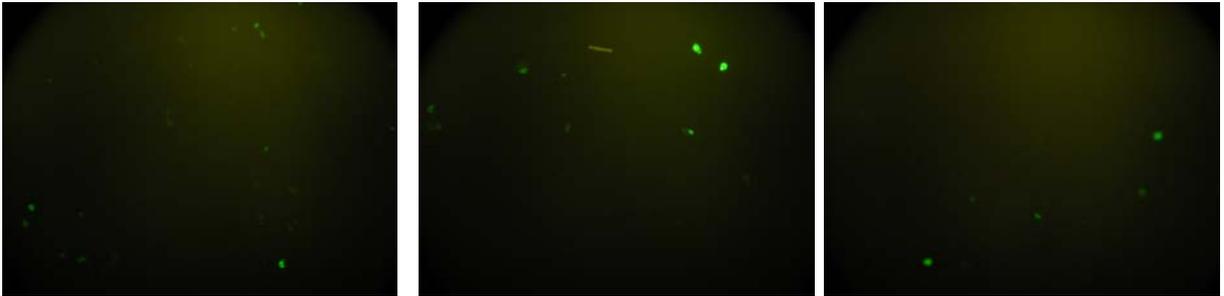
## 4. Ergebnisse

---

Phase



UV



2 µg DNA; 3 µl „Superfect“

2µg DNA; 7,5µl „Superfect“

2µg DNA; 15µl „Superfect“

**Abb. 12:** Transfektion der Zelllinie HaCaT mit „Superfect“. Es wurden jeweils 2 µg Plasmid-DNA eingesetzt und unterschiedliche Mengen des Reagenzes „Superfect“ (3 µl, 7,5 µl und 15 µl) verwendet. Zu sehen sind die tranfizierten Zellen im Phasenkontrast und unter UV-Licht.

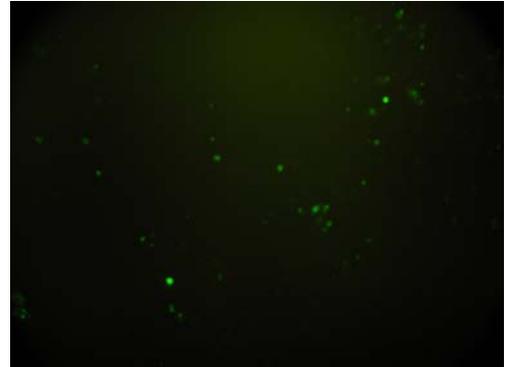
## 4. Ergebnisse

---

Phase



UV



10 µg DNA; „Cellpfect“ (Kalziumphosphatpräzipitation): 120 µl Puffer A; 240 µl Puffer B

**Abb. 13:** Transfektion der Zelllinie HaCaT mit dem Plasmid pGFP. Es wurden 10 µg DNA eingesetzt, und das „Cellpfect“ Transfektionskit (Kalzium-Phosphat-Präzipitation) verwendet: 240 µl Puffer A und 240 µl Puffer B wurden eingesetzt. Zu sehen sind die transfizierten Zellen nach 48 h im Phasenkontrast und unter UV-Licht.

Die besten Ergebnisse bezüglich der Transformation von HaCaT-Zellen konnten mit der Kalzium-Phosphat-Präzipitation („Cellpfect“) erzielt werden (Abb. 13). Deutlich ineffektiver war die Transfektion mit „Superfect“ (Abb. 12). „Effectene“ wirkte sowohl nach sechs-, als auch nach achtstündiger Inkubation stark zytotoxisch auf HaCaT-Zellen, was eine nahezu vollständige Lyse der Zellen zur Folge hatte (ohne Abb.). Daher wurde im Folgenden für alle Zelllinien die Kalzium-Phosphat-Methode verwendet.

### **4.2.2 Vergleich der Promotoraktivität von HERV-L-LTRs in verschiedenen Zelllinien**

In den folgenden Versuchen wurden die isolierten 5'- und solitären LTRs sowie die 5'- und 3'-HERV-L-LTR(Cord) in verschiedenen Zelllinien bezüglich ihrer zelltypspezifischen Aktivität untersucht. Für die Promotoruntersuchungen diente der Expressionsvektor pBL (Kapitel 2.8), der als Reporter das Luziferasegen trägt.

Aufgrund der von Seifarth *et al.* (2005) mittels Microarray-Analysen beschriebenen endogenen Aktivität von HERV-L-Elementen in Haut, Uterus, Zervix, Plazenta und Schilddrüse, wurden in der folgenden Versuchsreihe Zelllinien verwendet, die von den entsprechenden Geweben abstammen (Tabelle 6). Mit diesen Untersuchungen sollte geklärt werden, ob sich die Aktivitätsmuster der unterschiedlichen HERV-L-LTR-Gruppen voneinander unterscheiden und bestimmten Zelltypen zugeordnet werden können.

Mit Hilfe des von Seifarth *et al.* verwendeten HERV-Microarrays konnte lediglich die Transkription der ca. 200 proviralen HERV-L-Elemente, die ein *pol*-Gen enthalten, untersucht werden. In der vorliegenden Arbeit wurden jedoch auch solitäre LTRs in die Untersuchungen mit einbezogen, da diese im menschlichen Genom in etwa 6000 Kopien vorkommen und damit ein enormes Potential an transkriptionellen Regulationselementen darstellen. Für die Promotoruntersuchungen wurden die LTR-enhaltenden Plasmide (pBL-Plasmide) durch Kalzium-Phosphat-Präzipitation in die jeweiligen Zelllinien transfiziert und die Promotoraktivität anschließend, anhand des *Photinus*-Luziferase-Reportergens, ermittelt. Als interne Transfektionskontrolle wurde das Plasmid pRL-TK cotransfiziert, das als Reporter gen das *Renilla*-Luziferasegen enthält. Die Luziferasewerte wurden auf die *Renilla*-Aktivität normalisiert (*Photinus/Renilla*; Kapitel 3.18). Die Ergebnisse stammen jeweils aus zwei zeitlich unabhängigen Transfektionsexperimenten in denen jeweils zwei Transfektionsansätze untersucht wurden. Für die Untersuchungen wurden die in Tabelle 6 aufgelisteten Zelllinien verwendet, die sich von den Geweben ableiten, in denen mit dem retroviruspezifischen Microarray HERV-L-Transkripte nachgewiesen werden konnten (Seifarth *et al.*, 2005). Zusätzlich wurde die Kolon-Zelllinie LoVo in

#### 4. Ergebnisse

---

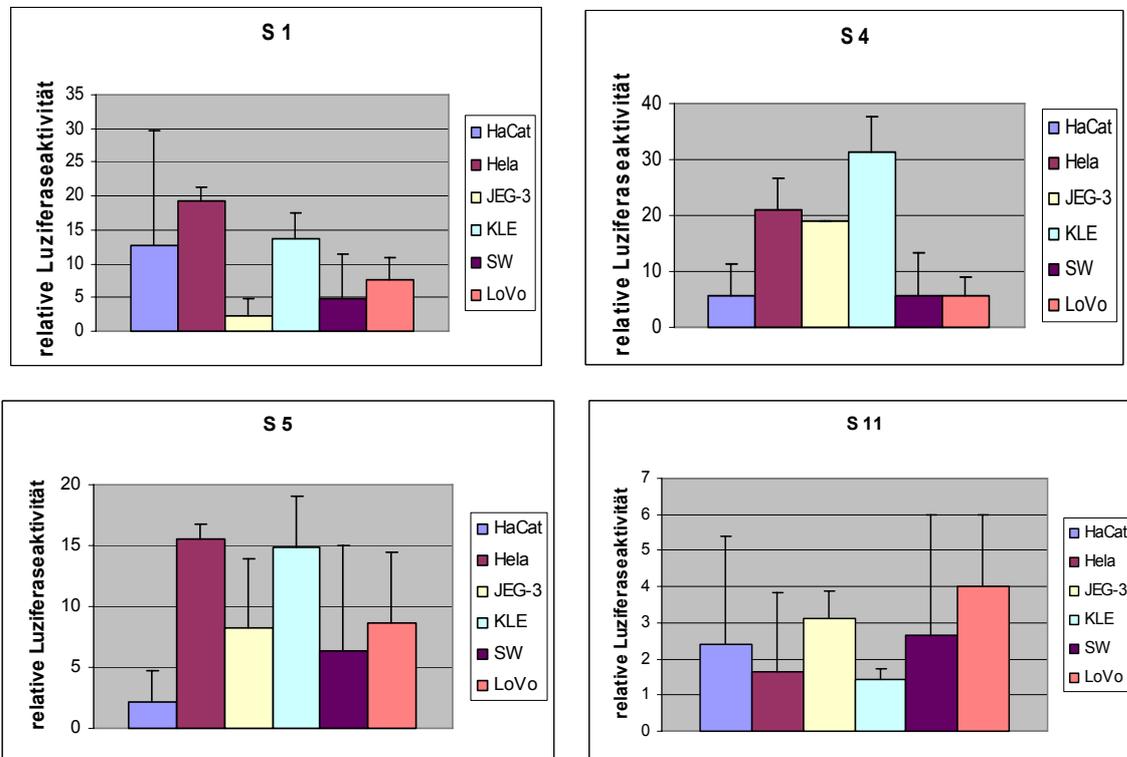
die Untersuchungen miteinbezogen, da in dieser Zelllinie eine solitäre HERV-L-LTR (Accession Number AY372061) als dominanter Promotor für das  $\beta$ 1,3-Galactosyltransferasegen fungiert (Dunn *et al.*, 2003).

**Tabelle 6:** Für die HERV-L-LTR-Promotoraktivitätsuntersuchungen verwendete Zelllinien und deren Ursprungsgewebe.

Gewebe	Zelllinie
Keratinocyten (Haut)	HaCaT
Zervix	HeLa
Plazenta	JEG-3
Uterus	KLE
Schilddrüse	SW
Kolon	LoVo

## 4. Ergebnisse

### 4.2.2.1 Promotoraktivität solitärer HERV-L-LTRs



**Abb. 14:** Promotoraktivität der solitären HERV-L-LTRs S 1, S 4, S 5, S 11 in den Zelllinien HaCaT, HeLa, JEG-3, KLE, SW und LoVo. *Photinus*-Luziferase Werte (pBL-Vektor) sind auf *Renilla*-Aktivität (interne Kontrolle pRL-TK), als Ratio *Photinus/Renilla* normalisiert. Ergebnisse sind auf die Negativkontrolle normalisiert. Die Daten stammen aus zwei unabhängigen transienten Transfektionsexperimenten (vier Transfektionsansätze).

Die HERV-L-LTR S 1 zeigte die stärkste Promotoraktivität in den HeLa-Zellen, hier lag eine ca. 19-fache Erhöhung im Vergleich zum promotorlosen pBL-Vektor (Negativkontrolle) vor, gefolgt von der Aktivität in den KLE-Zellen, die ca. 13-fach erhöht war. In den JEG-3-Zellen war nahezu keine Aktivität meßbar, in den SW-Zellen und in der Zelllinie LoVo war die Promotoraktivität sehr gering (Abb. 14).

Die LTR S 4 ist insgesamt die aktivste der hier getesteten solitären HERV-L-LTRs. Sie zeigte die höchste Promotoraktivität in der Zelllinie KLE, in der eine ca. 30-fache

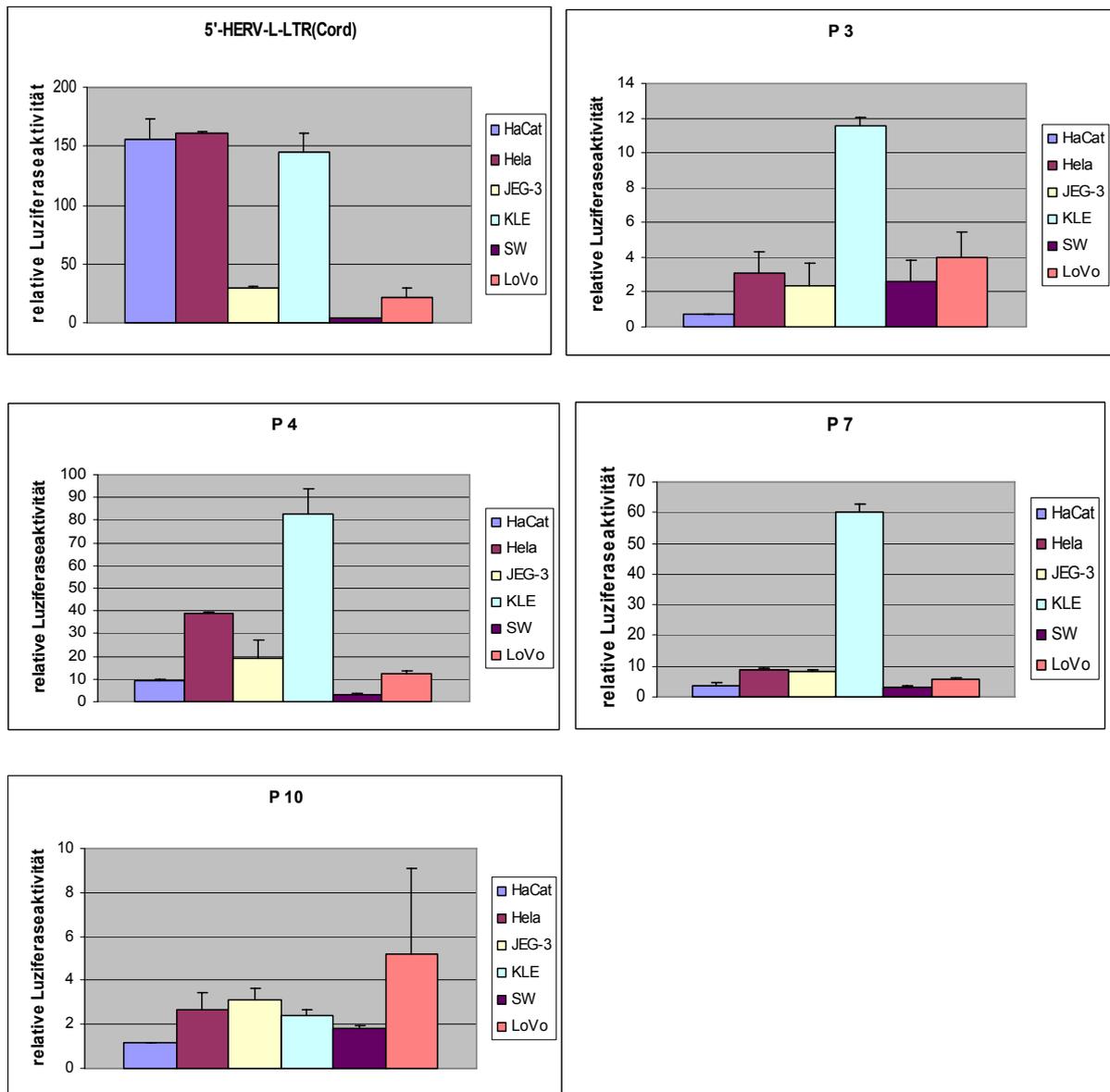
Erhöhung bezogen auf den promotorlosen pBL-Vektor (Negativkontrolle) ermittelt wurde. In den HeLa-Zellen wurde eine ca. 20-fache Aktivitätserhöhung festgestellt. Sehr schwach war die Promotoraktivität in HaCaT-Zellen und in den Zelllinien SW und LoVo. Die Aktivität lag hier nur ca. 6-fach über der Negativkontrolle (Abb. 14). Die Promotoraktivität der HERV-L-LTR S 5 war insgesamt etwas schwächer als die von S 1 und S 4. Die höchste ermittelte Promotoraktivität wurde in HeLa-Zellen und in KLE-Zellen festgestellt. Hier lag jeweils eine ca. 15-fache Aktivitätserhöhung bezogen auf den promotorlosen pBL-Vektor vor. In der Plazenta-Zelllinie JEG-3 und in der Kolon-Zelllinie LoVo wurde eine ca. 8-fache Promotoraktivitätserhöhung ermittelt. In den SW-Zellen war die Aktivität ca. 8-fach erhöht, in den HaCaT-Zellen konnte nahezu keine Erhöhung im Vergleich zur Negativkontrolle festgestellt werden (Abb. 14).

Die HERV-L-LTR S 11 war die inaktivste der getesteten solitären LTRs. Hier war in keiner der Zelllinien eine Promotoraktivitätserhöhung messbar die über dem 4-fachen der Aktivität der Negativkontrolle lag (Abb. 14).

Zusammenfassend lässt sich sagen, dass die solitäre HERV-L-LTR S 4 mit ca. 30-facher Erhöhung im Vergleich zur Negativkontrolle die stärkste Promotoraktivität unter den hier getesteten solitären HERV-L-Elementen aufweist.

## 4. Ergebnisse

### 4.2.2.2 Promotoraktivität proviraler HERV-L-LTRs



**Abb. 15:** Promotoraktivität der proviralen HERV-L-LTRs 5'-HERV-L-LTR (Cord), P 3, P 4, P 7 und P 10. Dargestellt ist die relative Luziferaseaktivität in den Zelllinien HaCaT, HeLa, JEG-3, KLE, SW und LoVo. *Photinus*-Luziferase Werte (pBL-Vektor) sind auf *Renilla*-Aktivität (interne Kontrolle pRL-TK), als Ratio *Photinus*/*Renilla* normalisiert. Ergebnisse sind auf die Negativkontrolle normalisiert. Die Daten stammen aus zwei unabhängigen transienten Transfektionsexperimenten (insgesamt vier Transfektionsansätze).

#### 4. Ergebnisse

---

Die 5'-HERV-L-LTR(Cord) zeigte die stärksten Promotoraktivitäten in den Zelllinien HaCaT, HeLa und KLE, hier lagen die Werte ca. 150-fach über der Aktivität der Negativkontrolle. Die Promotoraktivität lag in der Zelllinie JEG-3 ca. 30-fach, in der Zelllinie LoVo ca. 20-fach über der Aktivität des promotorlosen pBL-Vektors. In der Zelllinie SW war nahezu keine Promotoraktivität messbar (Abb. 15).

Die provirale HERV-L-LTR P 3 zeigte in allen getesteten Zelllinien lediglich schwache Promotoraktivität. Die höchste Aktivität wurde in der Zelllinie KLE ermittelt, hier lag sie ca. 12-fach über dem Wert der Negativkontrolle. In den Zelllinien HeLa, JEG-3, SW, LoVo und HaCaT war nur sehr geringe Aktivität vorhanden (Abb. 15).

Die HERV-L-LTR P 4 zeigte die stärkste Promotoraktivität in der Zelllinie KLE. Hier lag die Aktivität ca. 80-fach über dem Wert der Negativkontrolle. In der Zelllinie HeLa wurde eine ca. 40-fache, in JEG-3 eine ca. 20-fache, in LoVo- und HaCaT-Zellen eine ca. 10-fache Erhöhung der Aktivität im Vergleich zur Negativkontrolle festgestellt. In der Zelllinie SW war die LTR P 4 nahezu inaktiv (Abb. 15).

Die HERV-L-LTR P 7 war in der KLE-Zelllinie am aktivsten. Hier wurde ca. 60-fach höhere Aktivität als in der Negativkontrolle festgestellt. In den Zelllinien HeLa, JEG-3, SW, LoVo und HaCaT war die Promotoraktivität schwach, die höchste Aktivität, ca. 10-fach erhöht, wurde in HeLa-Zellen ermittelt (Abb. 15).

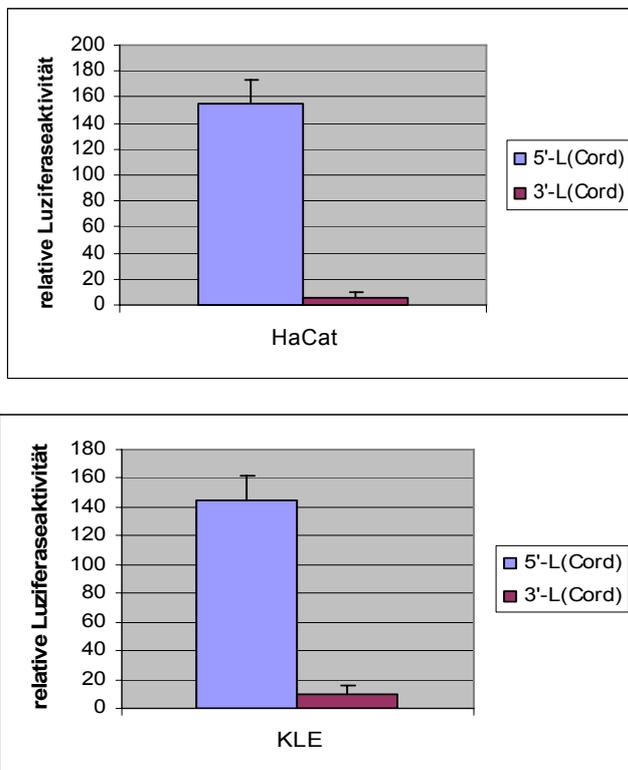
Die provirale HERV-L-LTR P 10 war in allen getesteten Zelllinien lediglich schwach aktiv. Die höchste Promotoraktivität war in der Zelllinie LoVo vorhanden (ca. 5-fach erhöht; Abb. 15).

Zusammenfassend lässt sich sagen, dass die 5'-HERV-L-LTR(Cord) von allen hier untersuchten HERV-L-LTRs die stärkste Promotoraktivität aufweist (ca. 150-fach höher als die Kontrolle). Sie war außerdem die einzige LTR, die in drei verschiedenen Zelltypen, in der Keratinozytenzelllinie HaCaT, der Uteruszelllinie KLE und der Zervixzelllinie HeLa stark aktiv war. Ebenfalls hohe Promotoraktivität, allerdings ausschließlich in der Uteruszelllinie KLE, zeigten die proviralen 5'-HERV-L-LTRs P 4 und P 7 mit 80-, bzw. 60-fach höherer Aktivität bezogen auf die promotorlose Kontrolle.

## 4. Ergebnisse

### 4.2.3 Promotoraktivität der 5'- und der 3'-HERV-L-LTR(Cord)

Untersuchungen der Promotoraktivität der 3'-HERV-L-LTR(Cord) in den Zelllinien HaCaT und KLE zeigten, dass diese LTR nahezu inaktiv ist. In Abbildung 16 sind die Ergebnisse der Luziferase-Assays dargestellt. Damit eignet sich die 3'-HERV-L-LTR(Cord) sehr gut, um auf Sequenzebene Unterschiede zur stark aktiven 5'-HERV-L-LTR(Cord) zu identifizieren, die für die Promotoraktivität- und spezifität der 5'-HERV-L-LTR(Cord) relevant sind.



**Abb. 16:** Vergleich der Promotoraktivität der 5'- und der 3'-HERV-L-LTR(Cord). Dargestellt ist jeweils die relative Luziferaseaktivität in den Zelllinien HaCaT und KLE. *Photinus*-Luziferase Werte (pBL-Vektor) wurden auf *Renilla*-Aktivität (interne Kontrolle pRL-TK) normalisiert, als Ratio *Photinus/Renilla*. Ergebnisse sind auf die Negativkontrolle normalisiert. Daten stammen aus zwei unabhängigen transienten Transfektionsexperimenten (vier Transfektionsansätze).

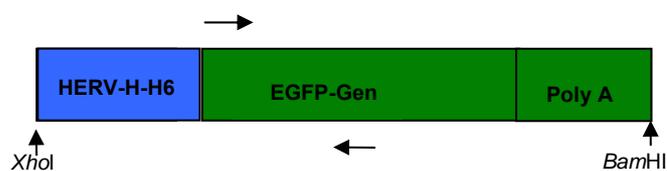
### 4.2.4 Promotoraktivität von HERV-LTRs in transgenen Mäusen

Die transgenen Mäuse wurden von Regine Schönfeld in ihrer Doktorarbeit am Institut für Molekulare Tierzucht der Ludwig-Maximilians-Universität München unter Leitung von Prof. Dr. Eckhard Wolf erzeugt (Doktorarbeit Regine Schönfeld, 2003). Es wurden drei Linien etabliert, die als Transgen das Gen für das „Enhanced Green Fluorescent Protein“ (EGFP) unter der Kontrolle des Promotors der HERV-H-H6-LTR (Linien: L5/N3, L1/N3 und L10/N4) enthalten, und eine Linie (L6/N3), die das *Photinus*-Luziferasegen unter der Kontrolle des Promotors der 5'-HERV-L-LTR(Cord) enthält (Abb. 17). Die Untersuchung der transgenen Mäusegewebe sollte Aufschluss über die Promotoraktivität der 5'-HERV-L-LTR(Cord) sowie der 5'-HERV-H-H6-LTR im Organismus der Maus liefern, da spezifische Aktivität dieser Elemente bisher ausschließlich in Zelllinien (Tumorzellen) nachgewiesen werden konnte.

a)



b)



**Abb. 17:** Schematische Darstellung der Konstrukte HERV-L-LTR(Cord)-Luziferase-Gen und HERV-H-H6-LTR-EGFP-Gen. Primer sind als Pfeile dargestellt. a) Länge des HERV-L-LTR-Luziferase-Gen-Konstrukts insgesamt: 3123 bp; Länge der HERV-L-LTR: 430 bp; Länge des Luziferase-Gens: 1714 bp; Länge des Poly(A)-Signals: 979 bp. b) Länge des HERV-H-H6-EGFP-Gen-Konstrukts insgesamt: 1455 bp; Länge der HERV-H-H6-LTR: 438 bp; Länge des EGFP-Gens: 784 bp; Länge des Poly(A)-Signals: 233 bp.

Im Rahmen meiner Doktorarbeit wurden Gewebeproben von Niere, Keimdrüse, Schilddrüse, Lunge, Nebenniere, Rektum, Milz, Zervix, Hirn, Herz, Haut, Leber, Uterus, Muskel und Brustgewebe von zuvor als transgen identifizierten Mäusen der Mauslinien L5/N3, L1/N3, L10/N4 (5'-HERV-H-H6-LTR-EGFP-Reportergen), sowie der Mauslinie L6/N3 (5'-HERV-L-LTR(Cord)-Luziferase-Reportergen) entnommen. Die Gewebeproben wurden, wie im Methodenteil zur RNA-Isolation beschrieben, behandelt und gereinigt. Nach dem anschließend durchgeführten DNase-Verdau wurden die Proben auf DNA-Kontaminationen überprüft. Dazu wurde eine Kontroll-PCR mit Primern spezifisch für das  $\beta$ -Actin-Gen durchgeführt. DNA-freie Proben wurden mittels reverser Transkriptase, unter Verwendung von Random-Hexamer-Primern, in cDNA umgeschrieben. Um zu testen, ob in den Proben noch genügend mRNA vorhanden war, wurde eine PCR mit  $\beta$ -Actin-Primern durchgeführt (erwartete Länge der Amplifikate: ca. 500 bp). Die darauf folgende PCR zur Untersuchung der LTR-Promotoraktivität wurde mit den für das zu untersuchende Reportergen spezifischen Primern (Luziferase- oder EGFP-Primer, Kapitel 2.2.3) durchgeführt. Die erwartete Länge der Luziferase-Amplifikate beträgt ca. 570 bp, die der EGFP-Amplifikate ca. 550 bp.

Von den im Rahmen dieser Arbeit untersuchten Mauslinien wurden jeweils zwei Mäuse, eine männliche und eine weibliche Maus, getestet.

### **4.2.5 Untersuchung HERV-H-H6-EGFP-transgener Mäuse**

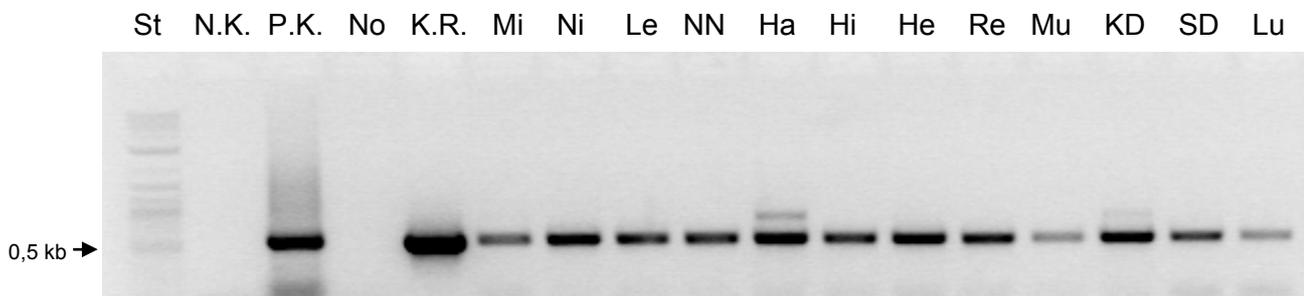
Von den HERV-H-H6-transgenen Mäusen wurden drei Linien untersucht, L5/N3, L1/N3 und L10/N4. Es wurde jeweils eine NoRT-Kontrolle verwendet, die alle Reagentien außer der reversen Transkriptase enthielt. Dadurch konnte getestet werden ob DNA-Kontaminationen in den RT-Reagentien vorhanden waren. Außerdem wurde Kontroll-RNA (Invitrogen) als Positiv-Kontrolle für die RT-PCR mitgetestet.

## 4. Ergebnisse

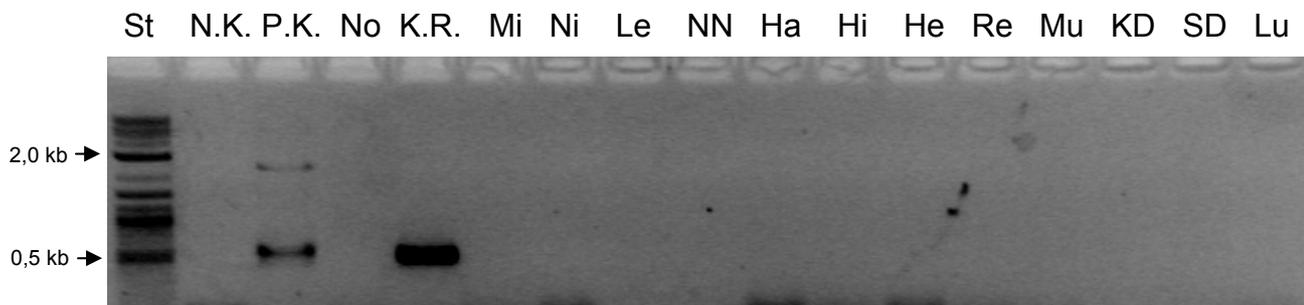
---

In Abbildung 18 a) sind die Ergebnisse der nach der cDNA-Synthese durchgeführten  $\beta$ -Actin-PCR der Mauslinie L5/N3J20 ( $\sigma$ ) gezeigt. In allen untersuchten Proben waren  $\beta$ -Actin-Amplifikate erkennbar. Somit war genügend cDNA vorhanden, um die  $\beta$ -Actin-positiven Proben nach Reportergenexpression zu untersuchen. In der anschließend durchgeführten PCR mit EGFP-Primern konnten in keiner der untersuchten Gewebeproben EGFP-Amplifikate nachgewiesen werden. Hier war also keine HERV-H-Promotoraktivität nachweisbar (Abb. 18 b).

a)



b)



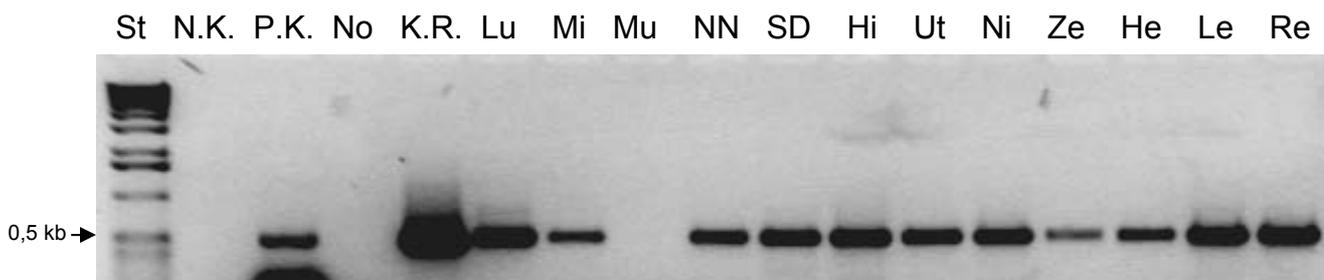
**Abb. 18:** a) PCR mit  $\beta$ -Actin-Primern nach der RT-PCR von Mauslinie L5/N3 J11 ( $\sigma$ ). b) EGFP-PCR der Mauslinie L5/N3 J11 ( $\sigma$ ). St: Längenstandard; N.K.: PCR-Negativkontrolle (Wasser); P.K.: PCR-Positivkontrolle (in  $\beta$ -Actin-PCR: Maus-DNA; in EGFP-PCR: pEGFP); No: No RT; K.R.: Kontroll-RNA; Mi: Milz; Ni: Niere; Le: Leber; NN: Nebenniere; Ha: Haut; Hi: Hirn; He: Herz; Re: Rektum; Mu: Muskel; Keimdrüse; SD: Schilddrüse; Lu: Lunge.

In Abbildung 19 sind die Ergebnisse der nach der RT-PCR durchgeführten  $\beta$ -Actin-PCR mit der Mauslinie L5/N3J20 ( $\sigma$ ) gezeigt. In allen untersuchten Proben

## 4. Ergebnisse

waren  $\beta$ -Actin-Amplifikate erkennbar (Abb. 19 a). EGFP-Expression (Abb. 19 b) konnte in Leber- und Rektumgewebe nachgewiesen werden (Spur 16 und 17). Bei den im Rektumgewebe vorhandenen PCR-Produkten, die ein größeres Molekulargewicht als das erwartete PCR-Produkt der EGFP-Primer besitzen, handelt es sich wahrscheinlich um Artefakte. In allen anderen Geweben war keine Expression detektierbar. Um auch relativ schwache Amplifikationen sichtbar zu machen wurde beim Fotografieren der Gele generell auf lange Belichtungsdauer geachtet, daher sind die Aufnahmen z.T. überbelichtet und es treten vorhandene Primer-Dimere hervor.

a)



b)



**Abb. 19:** a) PCR mit  $\beta$ -Actin-Primern nach der RT-PCR von Mauslinie L5/N3J20 (♀). b) PCR mit EGFP-Primern nach der RT-PCR von L5/N3J20 (♀). St: Standard; N. K.: PCR-Negativ Kontrolle mit  $H_2O$ ; P. K.: PCR-Positiv Kontrolle (in  $\beta$ -Actin-PCR: Maus-DNA; in EGFP-PCR: pEGFP); No: NoRT mit  $H_2O$ ; K. R.: Kontroll RNA für die RT-PCR; Lu: Lunge; Mi: Milz; Mu: Muskel; NN: Nebenniere; SD: Schilddrüse; Hi: Hirn; Ut: Uterus; Ni: Niere; Ze: Zervix; He: Herz; Le: Leber; Re: Rektum.

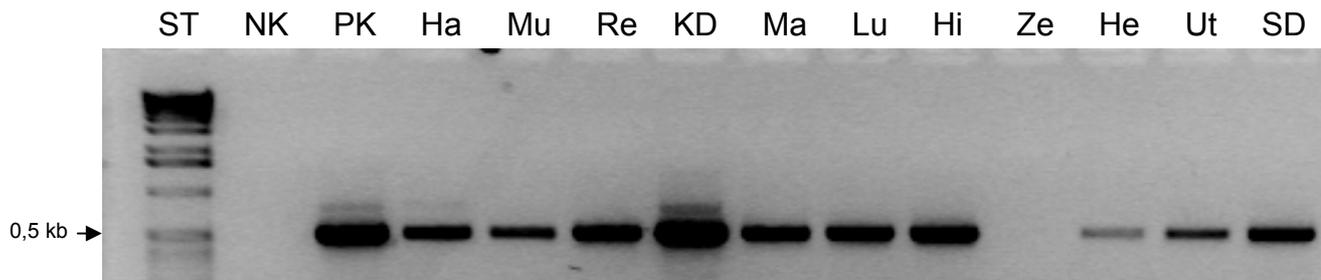
Die Ergebnisse der nach der RT-PCR durchgeführten  $\beta$ -Actin-PCR der Geweproben von L1/N3J6 (♀) sind in Abb. 20 a) gezeigt. Abgesehen von der Zervix-Probe waren

## 4. Ergebnisse

---

in allen Proben  $\beta$ -Actin-Amplifikate vorhanden. In Abbildung 20 b) sind die Ergebnisse der PCR der Gewebeproben der Mauslinie L1/N3J6 (♀) mit EGFP-Primern dargestellt. Es waren lediglich in der Hautprobe (Ha) schwache Amplifikate vorhanden.

a)



b)

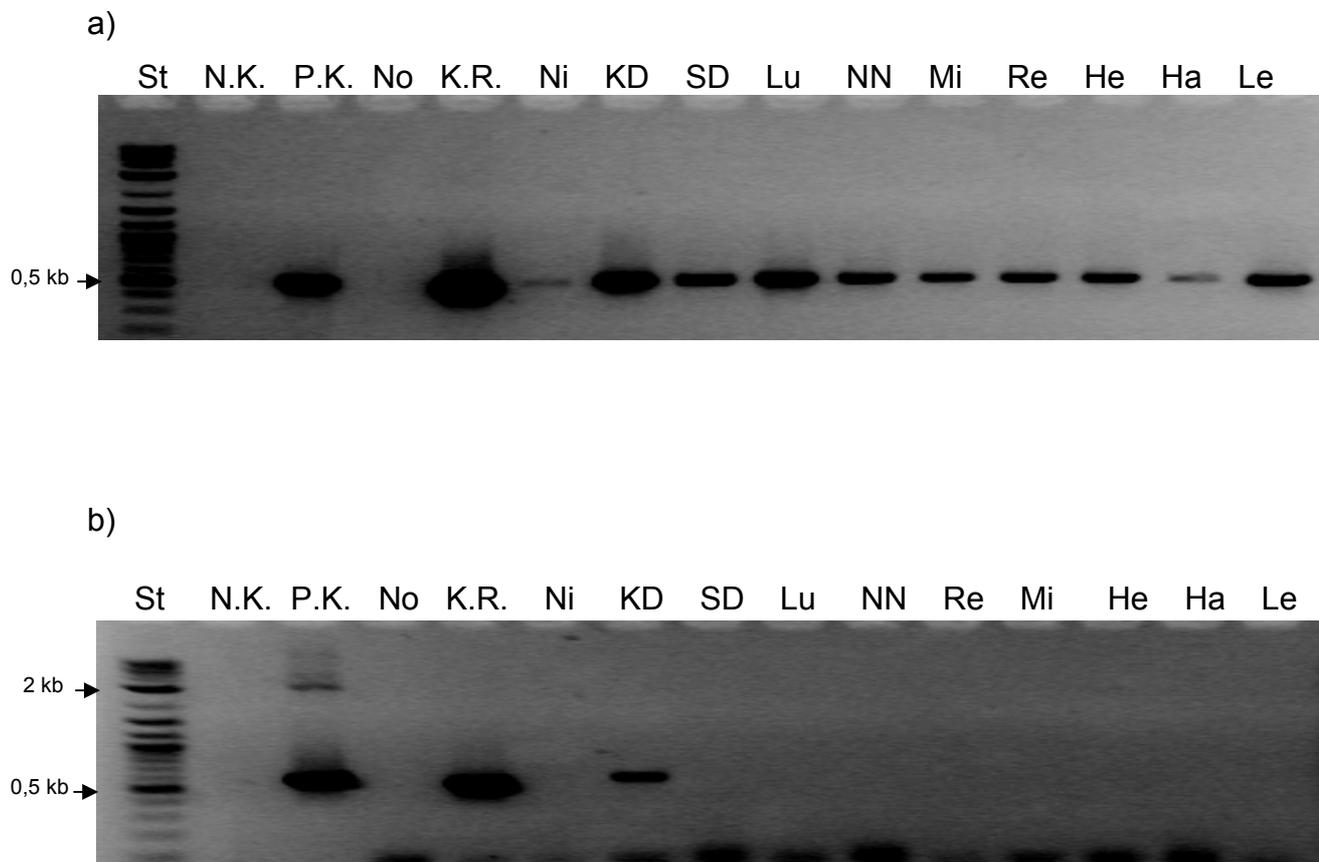


**Abb. 20:** a) PCR mit  $\beta$ -Actin-Primern nach der RT-PCR von Mauslinie L1/N3J6 (♀). b) PCR mit EGFP-Primern von L1/N3J6 (♀). St: Längenstandard; Nk: PCR-Negativkontrolle (Wasser); Pk: PCR-Positivkontrolle (in  $\beta$ -Actin-PCR: Maus-DNA; in EGFP-PCR: pEGFP); No: NoRT; K.R.: Kontroll-RNA; Re: Rektum; KD: Keimdrüse; Ma: Mammary Gland (Brustgewebe); Lu: Lunge; Hi: Hirn; Ze: Zervix; He: Herz; Ut: Uterus; Mu: Muskel; Ha: Haut.

In Abbildung 21 a) sind die Ergebnisse der nach der RT-PCR durchgeführten  $\beta$ -Actin-PCR der Linie L1/N3 J1 (♂) gezeigt. In allen untersuchten Proben waren  $\beta$ -Actin-Amplifikate erkennbar. In der PCR mit EGFP-Primern konnten im Gewebe der Keimdrüse EGFP-Amplifikate nachgewiesen werden. In der Maus L1/N3 J1 (♂) war in keiner der untersuchten Gewebe HERV-H-H6-Promotoraktivität nachweisbar (21 b).

## 4. Ergebnisse

---

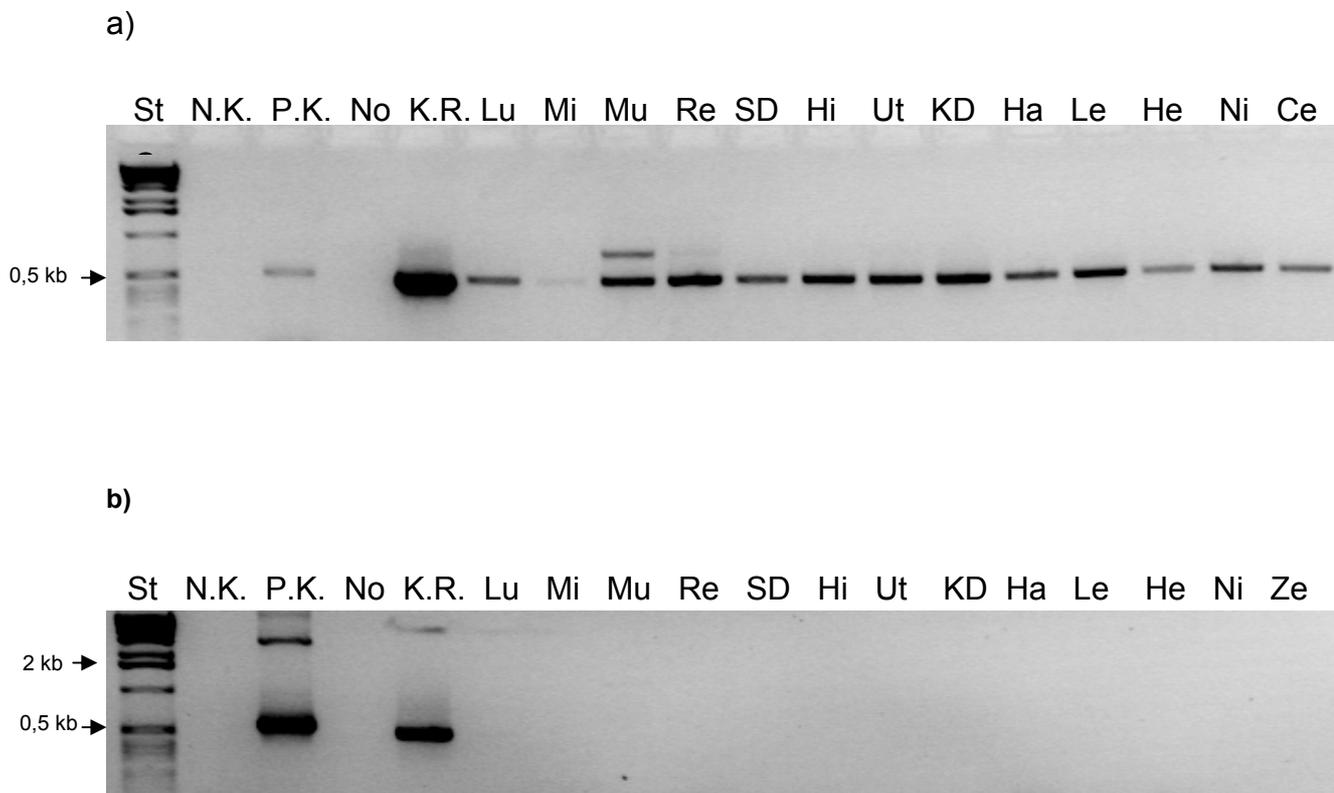


**Abb. 21:** a) PCR mit  $\beta$ -Actin-Primern nach der RT-PCR von Mauslinie L1/N3 J1 ( $\sigma$ ). b) PCR mit EGFP-Primern von L1/N3 J1 ( $\sigma$ ). St: Längenstandard; N.K.: PCR-Negativkontrolle (Wasser); P.K.: PCR-Positivkontrolle (in  $\beta$ -Actin-PCR: Maus-DNA; in EGFP-PCR: pEGFP); No: NoRT; K.R.: Kontroll-RNA; Ni: Niere; KD: Keimdrüse; SD: Schilddrüse; Lu: Lunge; NN: Nebenniere; Mi: Milz; Re: Rektum; He: Herz; Ha: Haut; Le: Leber.

Die Ergebnisse der nach der RT-PCR durchgeführten  $\beta$ -Actin-PCR der Gewebeproben von L10/N4J17 ( $\sigma$ ) sind Abbildung 22 a) dargestellt. Hier waren in allen Proben  $\beta$ -Actinamplifikate vorhanden. In der Milzprobe ist die Bande allerdings sehr schwach, hier war nur wenig cDNA vorhanden. In der PCR mit EGFP-Primern konnten in keinem der untersuchten Gewebe EGFP-Amplifikate nachgewiesen werden (Abb. 22 b). Es konnte kein Hinweis auf HERV-H-H6-Promotoraktivität gefunden werden.

## 4. Ergebnisse

---

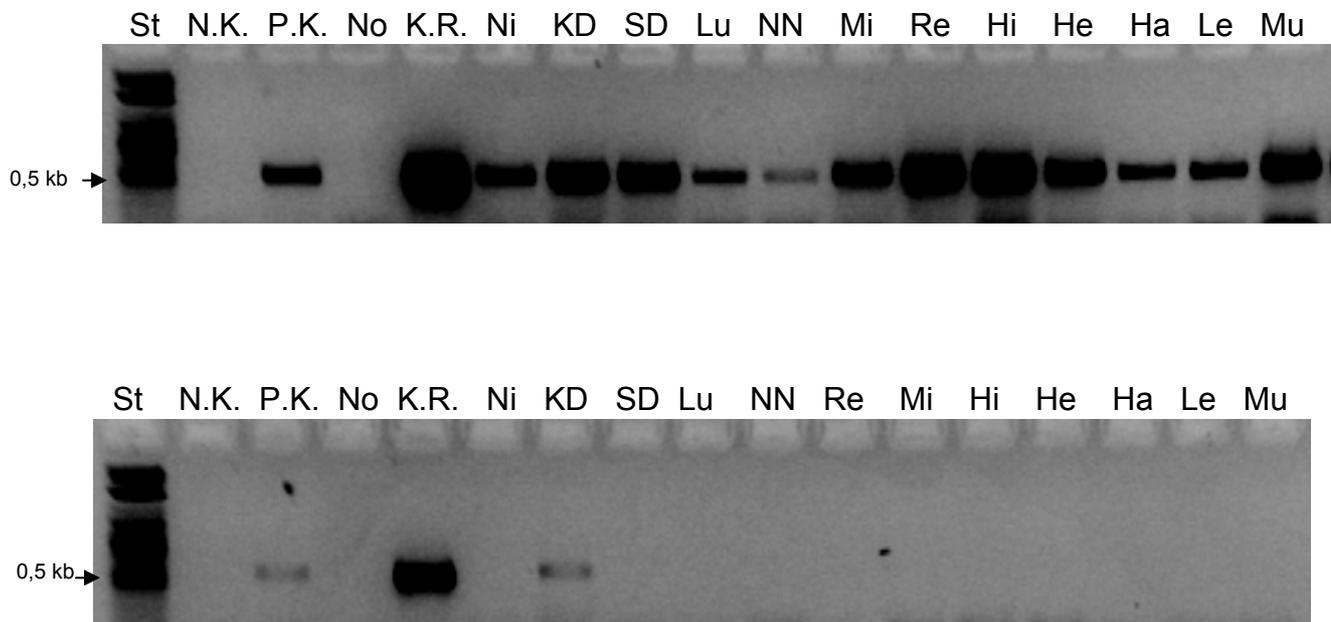


**Abb. 22:** : a) PCR mit  $\beta$ -Actin-Primern nach der RT-PCR von Mauslinie L10/N4 J17 (♀). b) PCR mit EGFP-Primern der Mauslinie L10/N4 J17 (♀). St: Standard; N.K.: PCR-Negativkontrolle (Wasser); P.K.: PCR-Positivkontrolle (in  $\beta$ -Actin-PCR: Maus-DNA; in EGFP-PCR: pEGFP); No: NoRT; K.R.: Kontroll-RNA; Lu: Lunge; Mi: Milz; Mu: Muskel; Re: Rektum; SD: Schilddrüse; Hi: Hirn; Ut: Uterus; KD: Keimdrüse; Ha: Haut; Le: Leber; He: Herz; Ni: Niere; Ze: Zervix.

In Abbildung 23 a) sind die Ergebnisse der nach der RT-PCR durchgeführten  $\beta$ -Actin-PCR der Linie L10/N4 J15 (♂) gezeigt. In allen untersuchten Proben waren  $\beta$ -Actin-Amplifikate erkennbar. Die in Abbildung 23 b) dargestellte PCR der Gewebeproben mit EGFP-Primern zeigte Amplifikate aus dem Gewebe der Keimdrüse (KD).

## 4. Ergebnisse

---



**Abb. 23:** a) PCR mit  $\beta$ -Actin-Primern nach der RT-PCR von Mauslinie L10/N4 J15 ( $\sigma$ ). b) PCR mit EGFP-Primern der Mauslinie L10/N4 J15 ( $\sigma$ ). St: Standard; N.K.: PCR-Negativkontrolle; P.K.: PCR-Positivkontrolle (in  $\beta$ -Actin-PCR: Maus-DNA; in EGFP-PCR: pEGFP); No: NoRT; K.R.: Kontroll-RNA; Ni: Niere; KD: Keimdrüse; SD: Schilddrüse; Lu: Lunge; NN: Nebenniere; Mi: Milz; Re: Rektum; Hi: Hirn; He: Herz; Ha: Haut; Le: Leber; Mu: Muskel.

Die Ergebnisse zu den drei untersuchten HERV-H-H6-LTR-transgenen Mauslinien sind in Tabelle 7 zusammengefasst.

#### 4. Ergebnisse

---

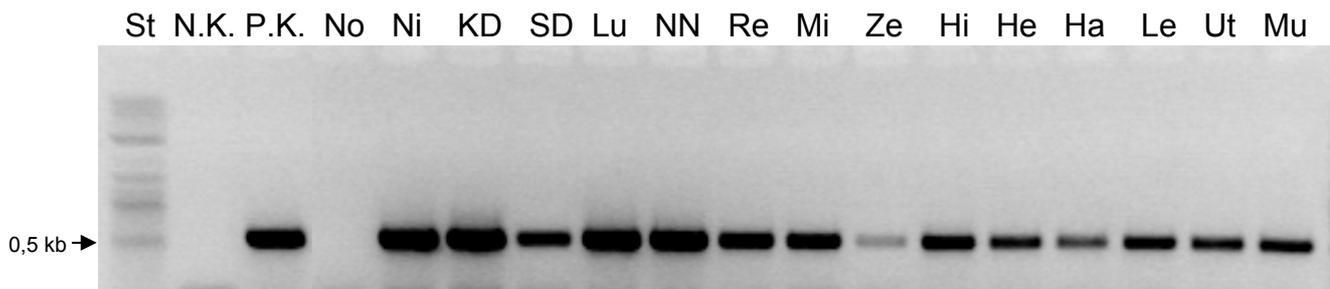
**Tabelle 7** Zusammenfassung der Ergebnisse der RT-PCR-Untersuchungen der HERV-H-H6/EGFP-transgenen Mäuse. /: kein Gewebe dieser Art vorhanden; -: keine Aktivität; +: Aktivität; -\*: keine Bande bei der  $\beta$ -Actin PCR nach der RT.

Mauslinie/ Gewebe	Linie # 1		Linie # 5		Linie # 10	
	L1/N3		L5/N3		L10/N4	
	J6 (♀)	J1 (♂)	J20 (♀)	J11 (♂)	J17 (♀)	J5 (♂)
Niere	/	-	-	-	-	-
Keimdrüse	-	+	/	-	-	+
Schilddrüse	-	-	-	-	-	-
Lunge	-	-	-	-	-	-
Nebenniere	/	-	-	-	/	-
Rektum	-	-	+	-	-	-
Milz	/	-	-	-	-	-
Zervix	-*	/	-	/	-	/
Hirn	-	/	-	-	-	-
Herz	-	-	-	-	-	-
Haut	+	-	/	-	-	-
Leber	/	-	-	-	-	-
Uterus	-	/	-	/	-	/
Muskel	-	/	-*	-	-	-
Mamma	-	/	/	/	/	/

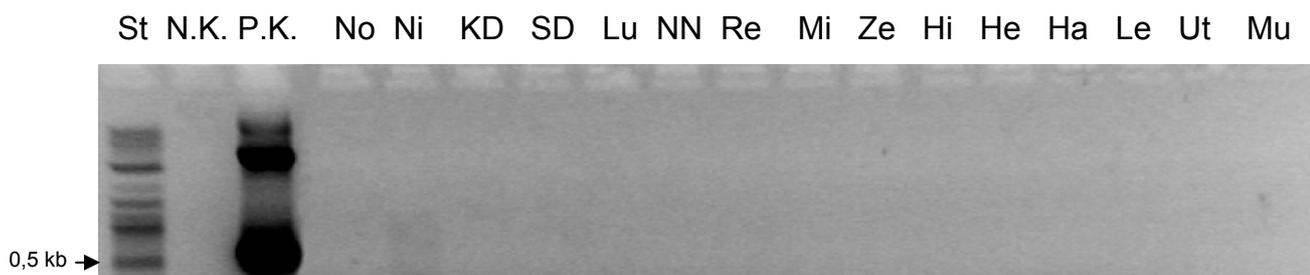
#### 4.2.6 Untersuchung 5'-HERV-L-LTR(Cord)-Luziferase-transgener Mäuse

Von den 5'-HERV-L(Cord)-transgenen Mäusen mit Luziferase als Reportergen stand lediglich die Mauslinie L6/N3 zur Verfügung (Doktorarbeit Schönfeld, 2003). In Abbildung 24 a) sind die Ergebnisse der nach der RT-PCR durchgeführten  $\beta$ -Actin-PCR der Linie L6/N3J16 (♀) gezeigt. In allen untersuchten Proben waren  $\beta$ -Actin-Amplifikate erkennbar. Die Bande in der Zervix-Probe ist allerdings deutlich schwächer als die der restlichen Proben. Hier war offensichtlich wenig cDNA vorhanden. Bei der PCR mit Luziferase-Primern konnten in keinem der Gewebe Luziferaseamplifikate nachgewiesen werden (Abb. 24 b).

a)



b)

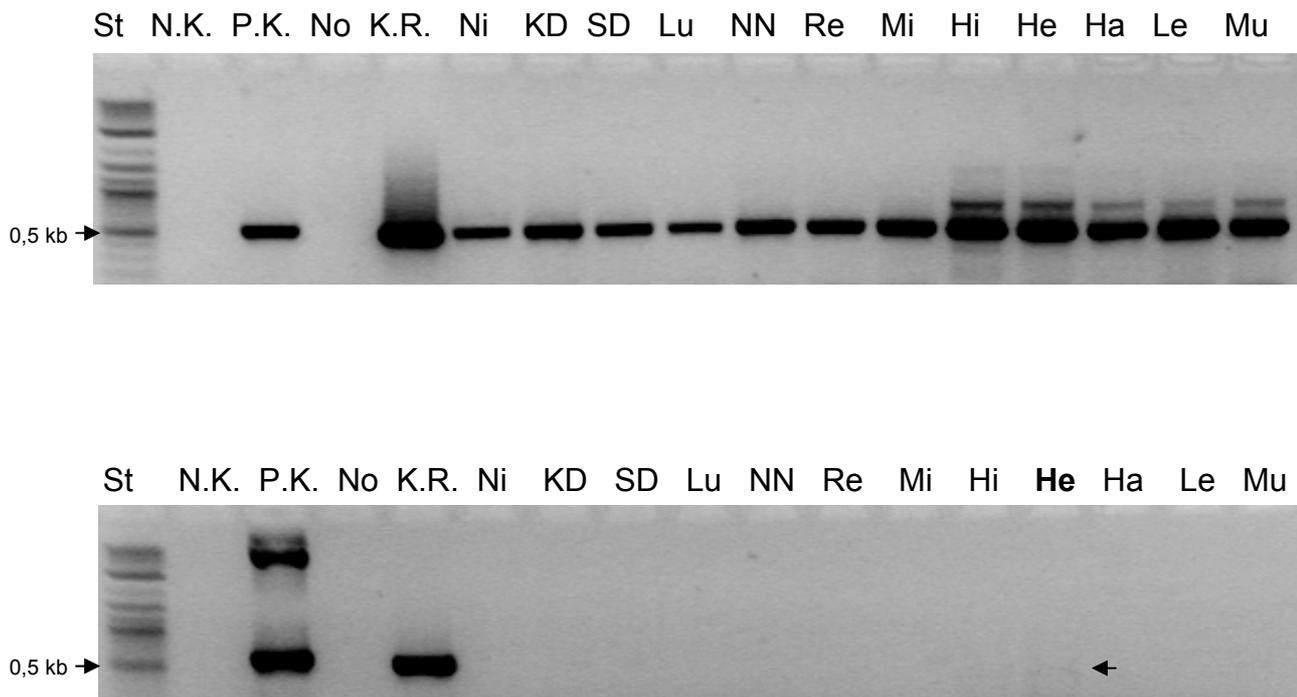


**Abb. 24:** a) PCR mit  $\beta$ -Actin-Primern nach der RT-PCR von Mauslinie L6/N3J16 (♀). b) PCR mit EGFP-Primern der Mauslinie L6/N3J16 (♀). St: Längenstandard; N.K.: PCR-Negativkontrolle (Wasser); P.K.: PCR-Positivkontrolle (in  $\beta$ -Actin-PCR: Maus-DNA; in Luziferase-PCR: pBL); No: NoRT; Ni: Niere; KD: Keimdrüse; SD: Schilddrüse; Lu: Lunge; NN: Nebenniere; Re: Rektum; Mi: Milz; Ze: Zervix; Hi: Hirn; He: Herz; Ha: Haut; Le: Leber; Ut: Uterus Mu: Muskel.

## 4. Ergebnisse

---

In Abbildung 25 a) sind die Ergebnisse der nach der RT-PCR durchgeführten  $\beta$ -Actin-PCR der Linie L6/N3 J11 ( $\sigma$ ) gezeigt. In allen untersuchten Proben waren  $\beta$ -Actin-Amplifikate erkennbar. Die PCR mit Luziferase-Primern der Gewebeproben von L6/N3 J11 ( $\sigma$ ) zeigte lediglich in der Herzprobe (He) schwache Amplifikate, die auf eine HERV-L-Promotoraktivität hinweisen (Abb. 25 b).



**Abb. 25:** a) PCR mit  $\beta$ -Actin-Primern nach der RT-PCR von Mauslinie L6/N3 J11 ( $\sigma$ ). b) PCR mit EGFP-Primern der Mauslinie L6/N3 J11 ( $\sigma$ ). St: Längenstandard; N.K.: PCR-Negativkontrolle (Wasser); P.K. PCR-Positivkontrolle (in  $\beta$ -Actin-PCR: Maus-DNA; in Luziferase-PCR: pBL); No: NoRT; K.R.: Kontroll-RNA; Ni: Niere; KD: Keimdrüse; SD: Schilddrüse; Lu: Lunge; NN: Nebenniere; Re: Rektum; Mi: Milz; Hi: Hirn; He: Herz; Ha: Haut; Le: Leber; Mu: Muskel.

### **4.3 Identifizierung potentieller Transkriptionsfaktorbindungsstellen durch Computeranalyse**

In verschiedenen Versuchen konnte gezeigt werden, dass HERV-L gewebespezifische Promotoraktivität besitzt. Einen guten Überblick über die Aktivität proviraler HERV-L-LTRs lieferten die Microarray-Untersuchungen, die mit cDNA aus Blut, Gehirn, Skelettmuskulatur, Haut, Herz, Niere, Leber, Lunge, Magen, Kolon, Rektum, Mamma, Ovar, Uterus, Zervix, Plazenta und Schilddrüse durchgeführt wurden, wobei sich Aktivität in Haut, Uterus, Zervix, Plazenta und Schilddrüsengewebe zeigte (Seifarth *et al.*, 2003 und 2005). Sowie die Untersuchungen mit der Keratinozyten-Zelllinien HaCaT, den primären Keratinozyten-Zellen NHEK (Schön, unveröffentlicht) und den Zelllinien KLE, JEG-3, SW und HeLa.

#### **4.3.1 Sequenzvergleich solitärer und proviraler HERV-L-LTRs**

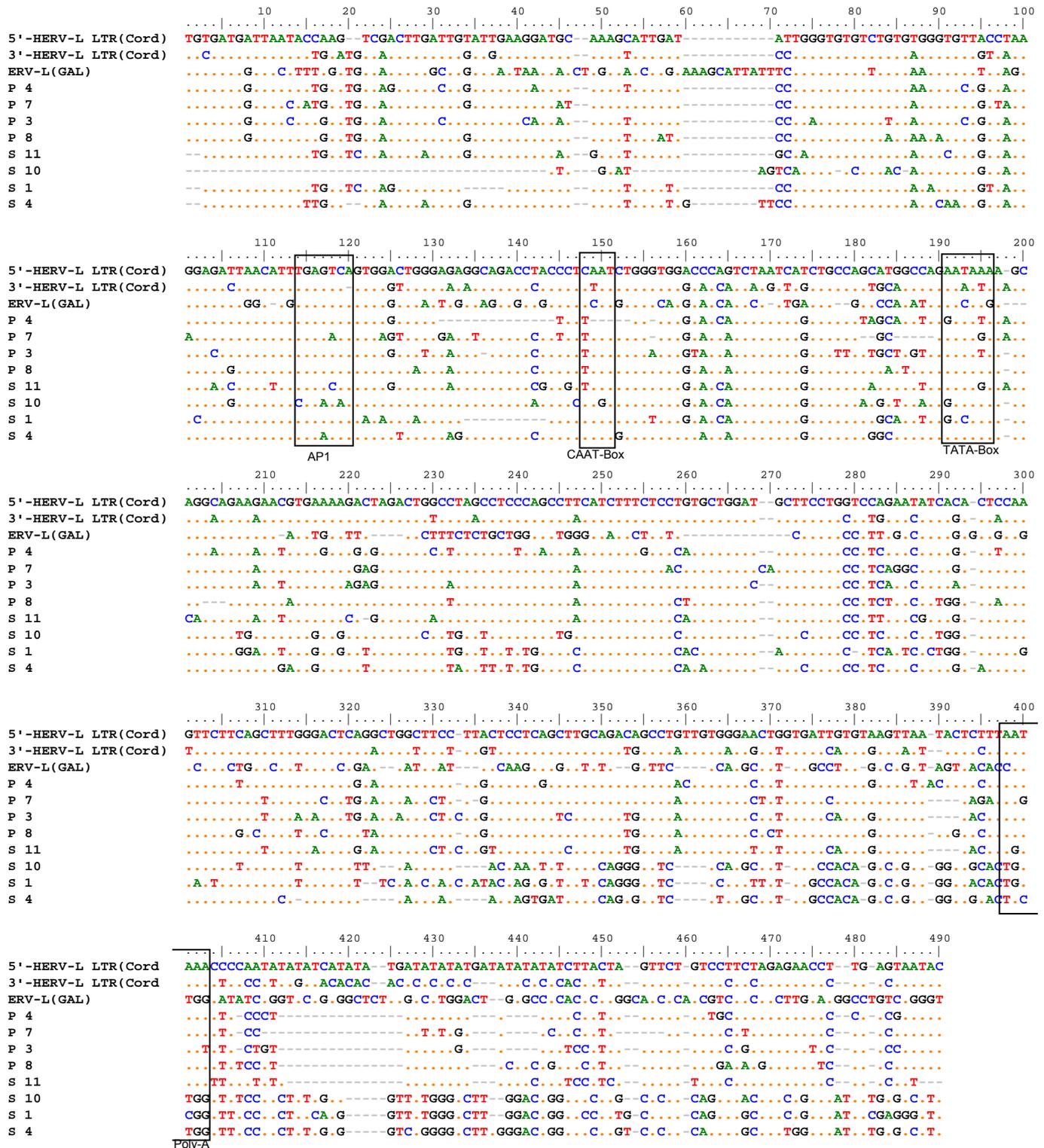
Um zu untersuchen, welche Bereiche der 462 bp langen 5'-HERV-L-LTR(Cord) für die Promotoraktivität relevant sind, wurden mit allen HERV-L-LTRs, deren Promotoraktivität bisher untersucht wurde, Alignments erstellt (Abb. 26). Die größten Sequenzhomologien sind zwischen der am stärksten aktiven 5'-LTR-L(Cord) und ihrer zugehörigen, inaktiven 3'-LTR-L(Cord) vorhanden. Beide Sequenzen sind, abgesehen von dem Bereich 392 bis 431 bp, gut konserviert. Die anderen untersuchten HERV-L-LTRs, inklusive der beiden in der Uterus-Zelllinie KLE aktiven proviralen Elemente P 4 und P 7, unterscheiden sich deutlich stärker von der Sequenz der 5'-HERV-L-LTR(Cord). In Abbildung 26 sind auch die von Cordonnier *et al.* (1995) in der 5'-HERV-L-LTR(Cord) identifizierten regulatorischen Elemente AP1, CAAT-Box, TATA-Box und Polyadenylierungssite eingezeichnet. Diese wurden durch Computeranalyse ermittelt und sind nicht experimentell bestätigt. Übereinstimmungen mit der potentiellen AP1-Bindungsstelle sind in der Sequenz

#### 4. Ergebnisse

---

ERV-L(Gal), in P 4, P 3, P 8 und S 1 vorhanden. Die CAAT-Box wäre nach Cordonnier *et al.* auch in den LTRs S 1 und S 4 zu finden. Die potentielle TATA-Box findet sich auch in der LTR P 8 und der LTR S 4, das Polyadenylierungssignal in der 3'-HERV-L-LTR(Cord), in der LTR P 4, in der LTR P 8 und der LTR S 11.

## 4. Ergebnisse



**Abb. 26:** Alignment der klonierten HERV-L-LTRs und der von Dunn *et al.* klonierten solitären ERV-L (GAL). Sequenzübereinstimmungen mit der 5'-HERV-L-LTR(Cord) sind als Punkt dargestellt (.), Lücken (Gaps) sind als Strich (-) dargestellt. Die von Cordonnier *et al.*, durch Computeranalysen in der 5'-HERV-L-LTR(Cord) identifizierten, potentiellen regulatorischen Sequenzen AP1, CAAT-Box, TATA-Box und Polyadenylierungssite sind eingerahmt.

### **4.3.2 Identifizierung potentieller Transkriptionsfaktorbindungsstellen mit Hilfe des Programms „Matinspector“**

Um für die Promotoraktivität relevante Bereiche innerhalb der 5'-HERV-L-LTR(Cord) zu identifizieren, wurde die Software „Matinspector“ (Genomatix) verwendet. Mit diesem Programm ist es möglich, potentielle Transkriptionsfaktorbindungsstellen zu bestimmen (Kapitel 3.20.1). Die Untersuchung lieferte 46 mögliche Transkriptionsfaktorbindungsstellen.

Im Folgenden sind die identifizierten Transkriptionsfaktorfamilien und ihre Position (bp) innerhalb der 5'-HERV-L-LTR(Cord) aufgelistet: HOXF (Pos. 3-19), HOXT (Pos. 5-17), SMAD (Pos. 18-26), PAX8 (Pos. 21-33), PBXF (Pos. 25-37), TBPF (Pos. 31-47), OCT1 (Pos. 40-54), RCAT (Pos. 39-63), CLOX (Pos. 46-62), SP1F (Pos. 55-69), EGRF (Pos. 69-83), CREB (Pos. 73-93), VBPF (Pos. 79-89), TBPF (Pos. 78-94), HOXF (Pos. 87-103), OCT1 (Pos. 88-102), HOXF (Pos. 93-109), AP1F (Pos. 92-112), AP1F (Pos. 96-116), GL1F (Pos. 136-150), CP2F (Pos. 139-149), HOXF (Pos. 147-163), AP4R (Pos. 153-169), NEUR (Pos. 155-167), BARB (Pos. 178-192), HIFF (Pos. 191-203), IRFF (Pos. 193-207), EVI1 (Pos. 200-216), SMAD (Pos. 205-213), RBPF (Pos. 218-232), PERO (Pos. 232-252), NRSF (Pos. 233-253), ETSF (Pos. 250-266), STAT (Pos. 337-355), CREB (Pos. 351-371), VBPF (Pos. 356-366), CEBP (Pos. 354-372), HNF1 (Pos. 365-381), SATB (Pos. 369-385), FAST (Pos. 385-399), HOXT (Pos. 388-400), EVI1 (Pos. 394-410), HOXT (Pos. 412-424), STAT (Pos. 437-455), MOKF (Pos. 438-458).

Bei diesen Ergebnissen handelt es sich lediglich um potentielle Transkriptionsfaktorbindungsstellen, die mittels Computeranalyse (Software: „Matinspector“; Kapitel 3.20.1) ermittelt wurden. Ob es sich bei den jeweiligen Sequenzmotiven tatsächlich um funktionelle Bindungsstellen handelt, kann nur experimentell verifiziert werden.

### 4.3.3 Promotor Module

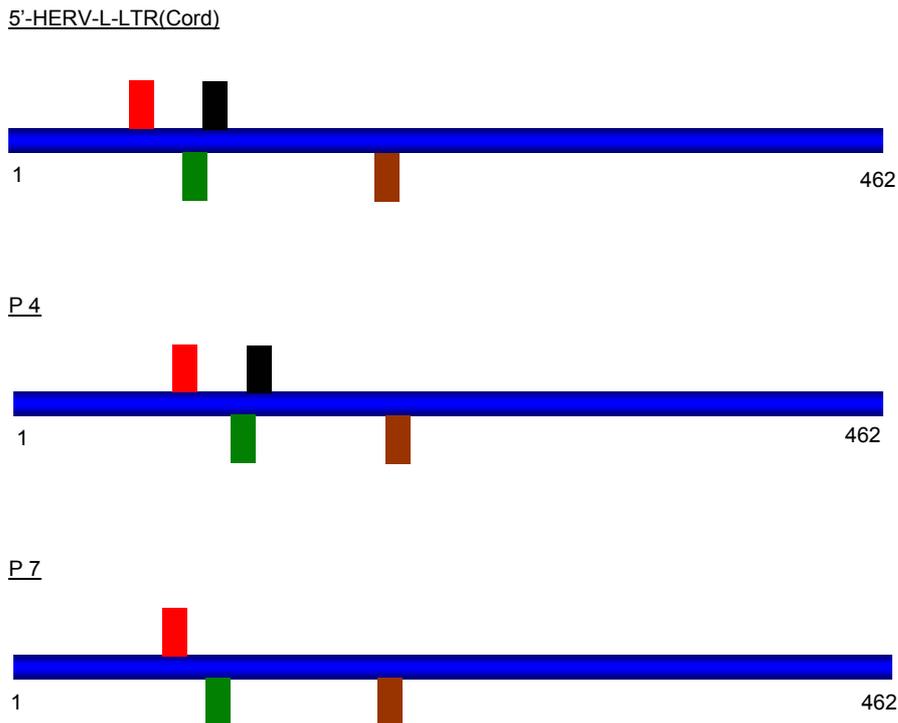
Um weitere Informationen bezüglich möglicher Funktionalität der mittels „Matinspector“ ermittelten potentiellen Transkriptionsfaktorbindungsstellen zu erhalten, wurde die „Promotor Module Library“ (Genomatix) herangezogen. Darin sind Kombinationen von Transkriptionsfaktoren aufgeführt, die, experimentell bestätigt, bereits in Promotoren nachgewiesen wurden (Module). Die dort beschriebenen Transkriptionsfaktor-Module wurden mit den in der 5'-HERV-L-LTR(Cord) ermittelten potentiellen Transkriptionsfaktorbindungsstellen verglichen. Im Folgenden sind Module beschrieben, die Transkriptionsfaktorbindungsstellen beinhalten, die auch in der 5'-HERV-L-LTR(Cord) gefunden wurden: Die Transkriptionsfaktoren SMAD und AP1F sind als synergistisch wirkend im Promotor des SMAD7-Genes der Maus beschrieben worden (Brodin *et al.*, 2000). Eine synergistische Wirkungsweise von HOXF und EGRF wurde im Promotor von LHbeta (Bovine luteneinizing hormone beta) nachgewiesen (Tremblay und Drouin, 1999), außerdem sind NEUR und HOXF zusammen im Promotor des POMC- (Rat proopiomelanocortin) Genes für dessen Aktivität essentiell (Poulin *et al.*, 1997). Bei den Transkriptionsfaktoren EGRF und AP1F konnte eine auffällige Verbindung zu SP1F festgestellt werden. Im Promotor des „Human interleukin-2-receptor beta-chain“ Genes sind die Faktoren SP1F und EGRF, und im Promotor des „Human CD11c integrin“ Genes SP1F und AP1F für die Promotoraktivität essentiell (Lin J. und Leonard W.J., 1997; Noti *et al.*, 1996).

Um weitere Hinweise darauf zu erhalten, welche der mittels „Matinspector“ bestimmten potentiellen Transkriptionsfaktorbindungsstellen biologische Relevanz besitzen, wurden die in der Uteruszelllinie KLE aktiven HERV-L-LTRs P 4, P 7 und die 5'-HERV-L-LTR(Cord) mit dem Computerprogramm „Frameworker“ (Genomatix; Kapitel 3.20.3), nach übereinstimmenden Transkriptionsfaktorbindungsstellen untersucht (Abb. 27). Dabei konnte gezeigt werden, dass alle drei HERV-LTRs sowohl eine potentielle SP1F- (Plusstrang), als auch eine potentielle HOXF-Bindungsstelle (Minusstrang) enthalten. Der Abstand dieser beiden Bindungsstellen beträgt in beiden LTRs 29 bp, in P 4 und P 7 befinden sie sich sogar exakt an der

## 4. Ergebnisse

---

gleichen Stelle der jeweiligen Sequenz. Im Falle der 5'-HERV-L-LTR(Cord) und der proviralen HERV-L-LTR P 4 ist außerdem eine potentielle AP1F-Bindungsstelle (Plusstrang) vorhanden, die in beiden LTRs im Abstand von 11 bp von der HOXF-Bindungsstelle gelegen ist.



**Abb. 27:** Schematische Darstellung übereinstimmender potentieller Transkriptionsfaktoren der mittels „Frameworker“ untersuchten promotoraktiven, proviralen HERV-L-LTRs: 5'-HERV-L-LTR(Cord), P 4 und P 7. Rotes Rechteck: **SP1F** (Plusstrang); grünes Rechteck: **HOXF**; schwarzes Rechteck (Plusstrang): **AP1F**. Braunes Rechteck: **AP4R**

Die Konservierung dieser potentiellen Transkriptionsfaktorbindungsstellen innerhalb der aktivsten drei HERV-L-LTRs ist ein guter Hinweis auf ihre biologische Relevanz hinsichtlich der Promotoraktivität dieser HERV-Elemente.

### **4.3.4 Identifizierung von TF-Bindungsstellen durch Sequenzvergleich der 5'-HERV-L-LTR(Cord) mit der inaktiven 3'-HERV-L-LTR(Cord)**

Um weitere Hinweise darauf zu erhalten, welche der mit der Software „Matinspector“-identifizierten, potentiellen Transkriptionsfaktorbindungsstellen für die Promotoraktivität der 5'-HERV-L-LTR(Cord) relevant sind, wurde auch die inaktive 3'-HERV-L-LTR(Cord) mittels „Matinspector“ untersucht und mit den Ergebnissen der 5'-HERV-L-LTR(Cord)-Untersuchung verglichen. Dabei wurde überprüft, ob jeweils eine in der 5'-LTR vorhandene, potentielle TF-Bindungsstelle im gleichen Bereich der 3'-LTR vorhanden ist. Mögliche Transkriptionsfaktorbindungsstellen, die ausschließlich in der aktiven 5'-LTR gefunden wurden, waren für weitergehende Untersuchungen von besonderem Interesse. Dabei konnten 10 potentielle Transkriptionsfaktorbindungsstellen identifiziert werden. Tabelle 8 zeigt diese potentiellen Bindungsstellen und ihre Lokalisation innerhalb der 5'-HERV-L-LTR(Cord). In dem in Abbildung 28 gezeigten Alignment der 5'-HERV-L-LTR(Cord) und der 3'-HERV-L-LTR(Cord) sind die potentiellen Transkriptionsfaktorbindungsstellen eingezeichnet.

## 4. Ergebnisse

---

**Tabelle 8ff:** Liste der mittels Computeranalysen (Software: „Matinspector“) und Vergleich der 5'- und 3'-HERV-L-LTR(Cord) identifizierten, potentiellen Transkriptionsfaktorbindungsstellen und ihre Lokalisation innerhalb der 5'-HERV-L-LTR(Cord). Rot markierte Basen beschreiben die durch „Matinspector“ ermittelte „Core similarity“, Basen, die in Großschrift dargestellt sind, die „Matrix similarity“ (Kapitel 3.20.1)

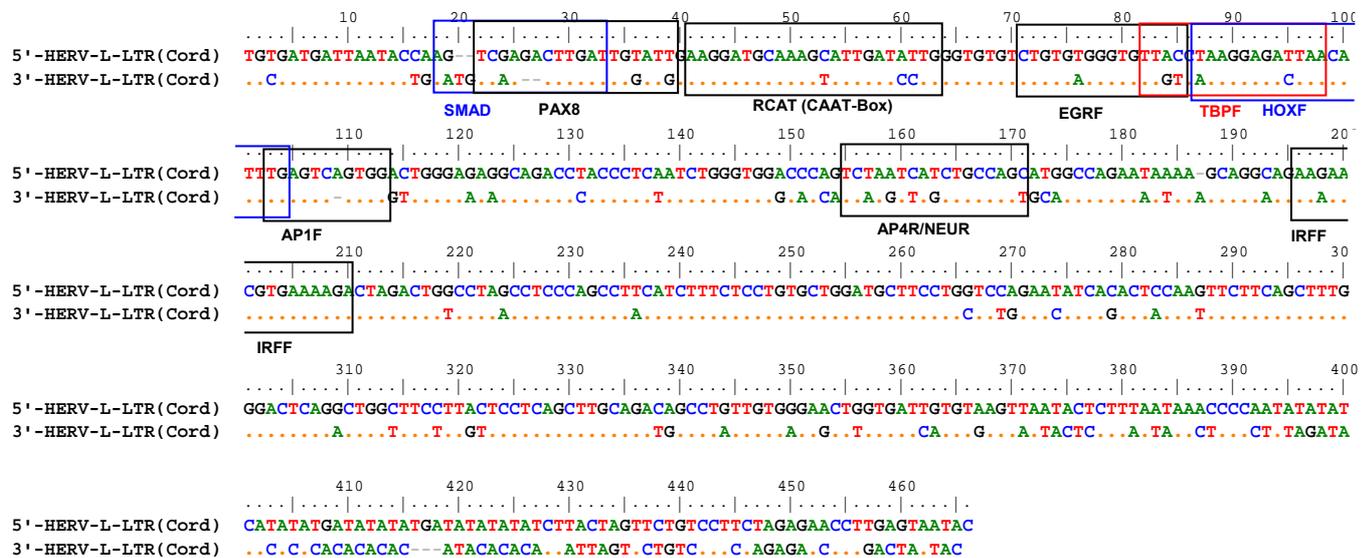
Transkriptionsfaktor (Matrix-Familie)	Information zur TF-Familie	Lokalisation	TF-Bindungsstelle innerhalb der 5'-HERV-L-LTR
SMAD	SMAD Transkriptionsfaktorfamilie	18-26 bp	GTCTcgact
PAX8	PAX 2/5/8 Binding Sites	21-33 bp	caaTCAgctctcg
RCAT	retrovirale CCAAT Binding Factors	39-63 bp	cCCAAtatcaatgctttgc atcctt
EGRF	Nerve Growth Factor induziertes Protein C und verwandte Faktoren	69-85 bp	ctgtgTGGtggttac
TBPF	TATA Binding Protein Factor	80-96 bp	tgttaccTAAGgagatt
HOXF	Faktoren mit moderater Aktivität in Homeodomänen-Konsensus-Sequenzen	85-101 bp	aggagATTAacattt
AP1F	Aktivator Protein 1	99-109 bp	ttgaGTCAgt

#### 4. Ergebnisse

---

<b>Transkriptionsfaktor (Matrix-Familie)</b>	<b>Information zur TF-Familie</b>	<b>Lokalisation</b>	<b>Position der TF-Bindungsstelle innerhalb der 5'-HERV-L LTR (Cord)</b>
AP4R	Aktivator Protein 4 und verwandte Faktoren	153-169 bp	gctgg <b>CAGAtg</b> attaga
NEUR	NEUROD1, NEUROGENIN 1 und 3 Bindungsstelle	155-167 bp	taat <b>CATCt</b> gcca
IRFF	Interferon Regulatory Factors	191-205 bp	aa <b>GAA</b> Cgt <b>gaa</b> aaga

## 4. Ergebnisse



**Abb. 28:** Alignment der 5'-HERV-L-LTR(Cord) und der 3'-HERV-L-LTR(Cord). Übereinstimmende Nukleotide sind als Punkt in der 3'-LTR dargestellt. Ausschließlich in der 5'-LTR vorhandene potentielle Transkriptionsfaktorbindungsstellen sind eingerahmt. Potentielle Transkriptionsfaktorbindungsstellen sind markiert.

Um zu entscheiden, welche der potentiellen Transkriptionsfaktorbindungsstellen für experimentelle Untersuchungen verwendet werden sollten, wurden die Ergebnisse aller Computeranalysen verglichen. Die unterschiedlichen potentiellen Transkriptionsfaktorbindungsstellen der aktiven 5'-HERV-L-LTR(Cord) und der inaktiven 3'-HERV-L-LTR(Cord) wurden als bester Hinweis auf biologische Relevanz erachtet und alle in die folgenden Untersuchungen miteinbezogen. Zusätzlich wurde auch eine potentielle SP1F-Bindungsstelle untersucht, die mit der RCAT/CLTR CAAT-Box überlappt, da sie in allen drei aktiven HERV-L-LTRs konserviert ist (Kapitel 4.3.3, Abb. 27).

### **4.4 Überprüfung der Funktionalität der identifizierten TF-Bindungsstellen durch „Site directed Mutagenesis“**

Um zu untersuchen, welche der durch Computeranalysen identifizierten, potentiellen Transkriptionsfaktorbindungsstellen tatsächlich aktiv waren, wurden die potentiellen Bindungsstellen für SMAD, PAX8, RCAT/SP1F, EGRF, TBPF, AP1F, AP4R/NEUR und IRFF mittels „Site-directed Mutagenesis“ überprüft (Stratagene; „Quick Change Site Directed Mutagenesis Kit“). Hierfür wurden die potentiellen TF-Bindungsstellen mit der Software „Sequence Shaper“ (Genomatix; Kapitel 3.20.2) nach Mutationen durchsucht, die ihre Inaktivierung zur Folge hatten. Anhand dieser Ergebnisse wurden Primer entwickelt (Abb. 29), die die entsprechenden Mutationen enthielten und damit die 5'-HERV-L-LTR(Cord) einschließlich Vektor amplifiziert. Anschließend wurden die parentalen DNA-Stränge mittels *DpnI* abgebaut und die mutierten Vektoren kloniert (Kapitel 3.12). Im Falle der potentiellen AP4R- und NEUR-Bindungsstellen sowie der RCAT/SP1F-Bindungsstelle führte das Einfügen einer Mutation in Folge der Überlappung dieser Elemente zur Inaktivierung beider TF-Bindungsstellen. Die Inserts der nach Klonierung und Plasmidpräparation erhaltenen Plasmide wurden sequenziert und die Mutationen durch Sequenzvergleich mit der Wildtyp-5'-HERV-L-LTR überprüft. Anschließend wurden die Plasmide mittels Kalzium-Phosphat-Präzipitation in die Keratinozyten-Zelllinie HaCaT und in die Uterus-Zelllinie KLE transfiziert. Danach wurde die Promotoraktivität im Luziferase-Assay bestimmt.

## 4. Ergebnisse

---

>PAX8 : GTCGAGACTTGATTGTGATTGAAGGATGCAAAGCATTGATATTG  
CTGCTCTGAACTAACATAACTTCCTACGTTTCGTAACATAAC

>PAX8 mut: GTCTCGACTTGATTGTATTGAGGGATGCAAAGCATTGATATTG  
CAGAGCTGAACTAACATAACTCCCTACGTTTCGTAACATAAC

>RCAT/SP1F: GATGCAAAGCATTGATATTGGGTGTGTCTGTGTGGG  
CATCGTTTCGTAACATAACCCACACAGACACACCC

>RCAT/SP1F MUT: GATGCAAAGCATTGATATCGAGTGTGTCTGTGTGGG  
CTACGTTTCGTAACATAAGCTCACACAGACACACCC

>EGFR: GATATTGGGTGTGCTCTGTGTGGGTGTTACCTAAGGA  
CTATAACCCACACAGACACACCCACAATGGATTCT

>EGFR MUT: GATATTGGGTGTGCTTTGTGGGTCTTACCTAAGGAG  
CTATAACCCACACAGAACACCCAGAATGGATTCTCT

>TBPF: TGGGGTGTTACCTAAGGAGATTAAACATTTGAGTC  
ACCCACAATGGATTCTCTAATTGTAAACTCAG

>TBPF MUT: TGGGGTGTACCTGATGAGATTAACATTTGAGTC  
ACCCACAATGGACTACTCTAATTGTAAACTCAG

HOXF: TGTTACCTAAGGAGATTAACATTTGAGTCAGTGGACTGG  
ACAATGGATTCTCTAATTGTAAACTCAGTCACCTGACC

HOXF MUT: TGTTACCTAAGGACATTTACATTTGAGTCAGTGGACTGG  
ACAATGGATTCTGTAAATGTAAACTCAGTCACCTGACC

>AP1F: CTAAGGAGATTAACATTTGAGTCAGTGGACTGGGAGAGGCAG  
GATTCTCTAATTGTAAACTCAGTCACCTGACCCTCTCCGTC

>AP1F MUT: CTAAGGAGATTAACATTTGTGTCGTGGACTGGGAGAGGCAG  
GATTCTCTAATCGTAAACACAGGCACCTGACCCTCTCCGTC

>AP4R/NEUR: TGGACCCAGTCTAATCATCTGCCAGCATGGCCAGAATAAAAGC  
ACCTGGGTCAGATTAGTAGACGGTCGTACCGTCTTATTTTCG

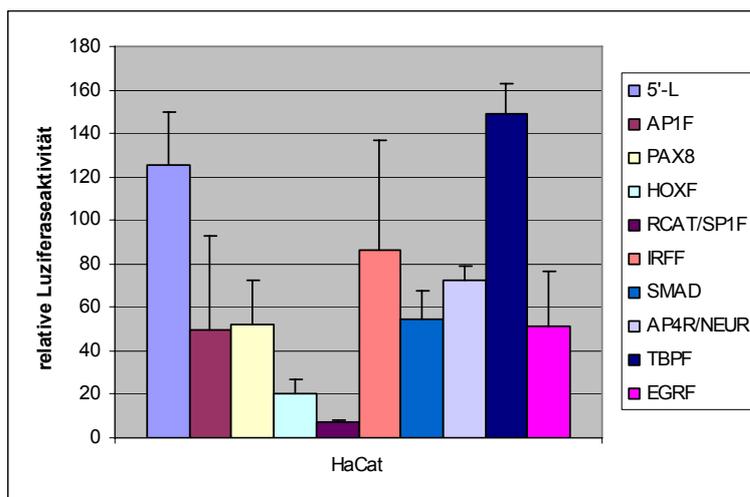
## 4. Ergebnisse

>AP4R/NEUR mut: GGTGGACCCAGTCTAATCATATTCC<sup>T</sup>GCATGGCCAGAATAAAAGC  
CCACCTGGGTCAGATTAGTATAAGG<sup>A</sup>CGTACCGGTCTTATTTTCG

>IRFF: CCAGAATAAAAGCAGGCAGAAAG<sup>A</sup>ACGTG<sup>A</sup>AAAAGACTAGACTGG  
GGTCTTATTTTCGTCCGTCTTCTTGCACTTTTCTGATCTGACC

>IRFF MUT: CCAGAATAAAAGCAGGCAGAAG<sup>G</sup>ACGTG<sup>A</sup>GAAGACTAGACTGG  
GGTCTTATTTTCGTCCGTCTTCT<sup>C</sup>TGCACT<sup>C</sup>TTCTGATCTGACC

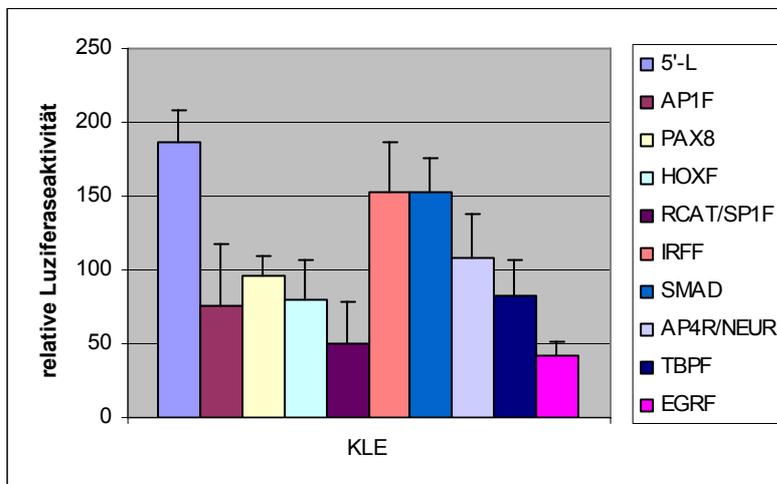
**Abb. 29.** Potentielle Transkriptionsfaktorbindungsstellen der 5'-HERV-L-LTR(Cord) und für die Mutagenese verwendete mutierte Primer (mut). Blaue Markierung: potentielle Transkriptionsfaktorbindungsstellen; Rote Markierung: In Mutagenese-Primer eingefügte Mutationen zur Inaktivierung der potentiellen TF-Bindungsstellen.



**Abb. 30:** Relative Luziferaseaktivität verschiedener 5'-HERV-L-LTR(Cord)-Mutationen in HaCaT-Zellen. Als Negativkontrolle wurde der promotorlose pBL-Vektor verwendet (nicht gezeigt). 5'-L: Positivkontrolle 5'-HERV-L-LTR(Cord); AP1F: 5'-HERV-L-LTR(Cord) mit inaktivierter potentieller AP1F-Bindungsstelle; PAX8: inaktive PAX8-Bindungsstelle; HOXF: inaktive HOXF-Bindungsstelle; RCAT/SP1F: inaktive RCAT/SP1F-Bindungsstelle; IRFF: Inaktive IRFF-Bindungsstelle; SMAD: Inaktive SMAD-Bindungsstelle; NEUR: Inaktive NEUR-Bindungsstelle; TBPf: Inaktive TBPf-Bindungsstelle; EGRF: Inaktive EGRF-Bindungsstelle. *Photinus*-Luziferase Werte (pBL-Vektor) sind auf *Renilla*-Aktivität (interne Kontrolle pRL-TK) als Ratio *Photinus/Renilla* normalisiert. Die Ergebnisse sind auf die Negativkontrolle normalisiert. Die Daten stammen aus zwei unabhängigen transienten Transfektionsexperimenten (vier Transfektionsansätze).

## 4. Ergebnisse

In Abbildung 30 sind die relativen Luziferaseaktivitäten der 5'-HERV-L-LTR(Cord)-Wildtyp-Sequenz und der mutierten LTR-Sequenzen in HaCaT-Zellen gegenübergestellt. Gezeigt sind die Aktivitäten der durch Punktmutationen inaktivierten potentiellen TF-Bindungsstellen AP1F, PAX8, HOXF, RCAT/SP1F, IRFF, SMAD, AP4R, TBPF, EGRF. Es zeigte sich, dass die Promotoraktivität durch Inaktivierung der Bindungsstellen AP1F, PAX8, SMAD, AP4R und EGRF um mehr als 50% bezogen auf die Wildtypaktivität zurückging. Die Inaktivierung der Bindungsstelle HOXF führte zu einer mehr als 6-fachen Reduzierung der Promotoraktivität. Das HERV-L-Element mit der mutierten RCAT/SP1F-Bindungsstelle war nahezu inaktiv. Demnach sind die Transkriptionsfaktorbindungsstellen AP1F, PAX8, SMAD, AP4R, EGRF und vor allem HOXF und RCAT/SP1F für die Promotoraktivität der 5'-HERV-L-LTR(Cord) essentiell.



**Abb. 31:** Relative Luziferaseaktivität von 5'-HERV-L-LTR(Cord)-Konstrukten getestet in KLE-Zellen. Als Negativkontrolle wurde der promotorlose pBL-Vektor verwendet (nicht gezeigt). 5'-L: Positivkontrolle 5'-HERV-L-LTR; AP1F: 5'-HERV-L-LTR mit inaktivierter potentieller AP1F-Bindungsstelle; TBPF (1): inaktive PAX8-Bindungsstelle; HOXF: inaktive HOXF-Bindungsstelle; RCAT/SP1F: inaktive RCAT/SP1F-Bindungsstelle; IRFF: Inaktive IRFF-Bindungsstelle; SMAD: Inaktive SMAD-Bindungsstelle; NEUR: Inaktive NEUR-Bindungsstelle; TBPF (2): Inaktive TBPF-Bindungsstelle; EGRF: Inaktive EGRF-Bindungsstelle. *Photinus*-Luziferase Werte (pBL-Vektor) sind auf *Renilla*-Aktivität (interne Kontrolle pRL-TK) als Ratio *Photinus/Renilla* normalisiert. Die Ergebnisse sind auf die Negativkontrolle normalisiert. Die Daten stammen aus zwei unabhängigen transienten Transfektionsexperimenten (vier Transfektionsansätze).

In Abbildung 31 wurden die relativen Luziferase-Aktivitäten der mutierten Konstrukte im Vergleich zur Wildtyp 5'-HERV-L-LTR(Cord) in der Uteruszelllinie KLE getestet. Ein Rückgang der Promotoraktivität um mehr als 50% ist in den HERV-Elementen, die die inaktivierten TF-Bindungsstellen AP1F, PAX8, HOXF und TBPF enthielten, zu erkennen. Die Inaktivierung der RCAT/SP1F-Bindungsstelle führte zu einem Rückgang um das 3,7-, die der EGRF-Bindungsstelle sogar um das 4,5-fache. Die HERV-L-LTR mit der inaktiven AP4R/NEUR-Bindungsstelle zeigte eine Aktivitätsverringering um das 1,7-fache der Wildtypaktivität. Die Mutation der SMAD- und IRFF-Bindungsstelle hatte lediglich einen geringen Einfluss auf die Aktivität. Als funktionelle Bindungsstellen in KLE-Zellen sind hiermit AP1F, PAX8, HOXF, TBPF und vor allem RCAT/SP1F und EGRF zu nennen.

Zusammenfassend lässt sich sagen, dass einige Übereinstimmungen funktioneller DNA-Bereiche in den Zelllinien HaCaT und KLE vorhanden sind. Eine Mutation und damit Inaktivierung der potentiellen Transkriptionsfaktorbindungsstellen AP1F, PAX8, RCAT/SP1F, AP4R/NEUR und EGRF, führt in beiden Zelllinien zu einer deutlichen Reduzierung der Promotoraktivität. Die Inaktivierung der potentiellen SMAD-Bindungsstelle ist hingegen nur in der HaCaT-Zelllinie mit einer verringerten Promotoraktivität verbunden, die Inaktivierung der TBPF-Bindungsstelle führt lediglich in den KLE-Zellen zu deutlicher Aktivitätsverminderung.

### **4.5 Identifizierung funktioneller Bindungsstellen mittels „Electrophoretic Mobility Shift Assay“ (EMSA)**

Zur Identifizierung der an die funktionellen Sequenzmotive bindenden Transkriptionsfaktoren wurden die potentiellen Bindungsstellen AP1F, RCAT/SP1F, EGRF, HOXF und NEUR/AP4R mittels EMSA untersucht. Diese potentiellen Bindungsstellen wurden ausgewählt, da ihre Inaktivierung durch Einführung von Mutationen, sowohl in der Keratinozyten-Zelllinie HaCaT, als auch in der Uterus-Zelllinie KLE, zu einem deutlichen Rückgang der Promotoraktivität führte. Es wurden Oligonukleotide ausgewählt, die jeweils einen Bereich der 5'-HERV-L-LTR(Cord)

enthielten, der sich im Luziferase Assay als funktionell erwiesen hatte. Zur Untersuchung wurden von den Zelllinien KLE und HaCaT Kernextrakte hergestellt. Enthalten die Kernextrakte Proteine, die an die Oligonukleotide binden können, kommt es nach Zugabe der Extrakte zu einer Komplexbildung, die in der Verringerung der Wanderungsgeschwindigkeit der Oligonukleotide durch das native Polyacrylamidgel resultiert. Dadurch zeigt sich eine Bande, die im Vergleich zur Bande des unbeladenen Oligonukleotids zu höheren Molekulargewichten hin verschoben ist („Shift“). Um zu testen, ob eine solche Protein-DNA-Wechselwirkung tatsächlich sequenzspezifisch ist, wurden außerdem Kompetitionsreaktionen durchgeführt in denen zusätzlich die nicht radioaktiv markierte Form des zu testenden Oligonukleotids dem Reaktionsansatz im Überschuß zugefügt wurde. Bei Wechselwirkung des unmarkierten Oligonukleotids mit dem jeweiligen Protein ist auf dem Röntgenfilm kein Protein-DNA-Komplex mehr zu erkennen. In einem weiteren Reaktionsansatz wurden zusätzlich zum markierten Oligonukleotid (Wildtyp-HERV-L-Oligonukleotid) Oligonukleotide, in denen die jeweils untersuchte Transkriptionsfaktorbindungsstelle inaktiviert war (mutierte HERV-L-Oligonukleotide), im Überschuss zugegeben. Damit kann festgestellt werden, ob die Proteinbindung spezifisch an der jeweiligen potentiellen Transkriptionsfaktorbindungsstelle erfolgt. Zur weiteren Untersuchung wurden Oligonukleotide als Konkurrenten zugegeben, die eine Konsensus-Sequenz der vermuteten Transkriptionsfaktorfamilie enthielten. Erfolgreiche Konkurrenz mit diesen Oligonukleotiden lässt eine Zuordnung des Transkriptionsfaktors zu einer bestimmten Transkriptionsfaktorfamilie zu. Außerdem wurden, um einen Transkriptionsfaktor zu identifizieren, „Supershift“-Untersuchungen mit spezifisch bindende Antikörpern durchgeführt (Kapitel 3.19).

In Abbildung 32 ist ein Ausschnitt der 5'-HERV-L-LTR(Cord) dargestellt. Die für den EMSA verwendeten Oligonukleotide zur Untersuchung der potentiellen SP1F-Bindungsstelle sind unterhalb des HERV-L-Ausschnitts dargestellt und bezüglich ihrer Position innerhalb der Sequenz angeordnet. Die mit „MatInspector“ und „Sequence Shaper“ identifizierten „Matrix-“ und „Core Similarity“-Sequenzen sind markiert. Außerdem zeigt Abbildung 32, dass die Wildtypoligonukleotide WT-HERV-L-SP1F und WT-HERV-L-EGRF teilweise überlappen, da beide potentiellen

#### 4. Ergebnisse

---

TF-Bindungsstellen nebeneinander liegen. Daher ist die „Core“-Sequenz der potentiellen SP1F-Bindungsstelle nicht nur im WT-HERV-L-SP1-Oligonukleotid, sondern auch im WT-HERV-L-EGRF-Oligonukleotid enthalten.

HERV-L: CAAAGCATTGATATTGGGTGTGTCTGTGTGGGTGTTACCTAAGGAGATT  
GTTTCGTA ACTATAACCCACACAGACACACCCACAATGGATTCTCTAA

SP1F (Core Sim.: 0,75/Matrix sim.0,878)

WT-HERV-L-SP1F: CAAAGCATTGATATTGGGTGTGTCTGTGTG  
GTTTCGTA ACTATAACCCACACAGACACAC

SP1F MUT

HERV-L-SP1F-MUT: CAAAGCATTGATATCGAGTGTGTCTGTGTG  
GTTTCGTA ACTATAGCTCACACAGACACAC

EGRF (Matrix Sim. 0,776/Core Sim. 0,857)

WT-HERV-L-EGRF: TATTGGGTGTGTCTGTGTGGGTGTTACCT  
ATAACCCACACAGACACACCCACAATGGA

EGRF MUT

HERV-L-EGRF-MUT: TATTGGGTGTGTCTTTGTGGGTCTTACCT  
ATAACCCACACAGAAACACCCAGAATGGA

**Abb. 32:** Für die EMSA-Untersuchungen verwendete Oligonukleotide WT-HERV-L-SP1F, WT-HERV-L-SP1F-MUT (potentielle SP1F-Bindungsstelle inaktiviert), WT-HERV-L-EGRF und WT-HERV-L-EGRF-MUT (potentielle EGRF-Bindungsstelle inaktiviert). Zusätzlich sind „Matrix“ und „Core similarity“ der potentiellen Transkriptionsfaktorbindungsstellen angegeben (Kapitel 3.20.1).

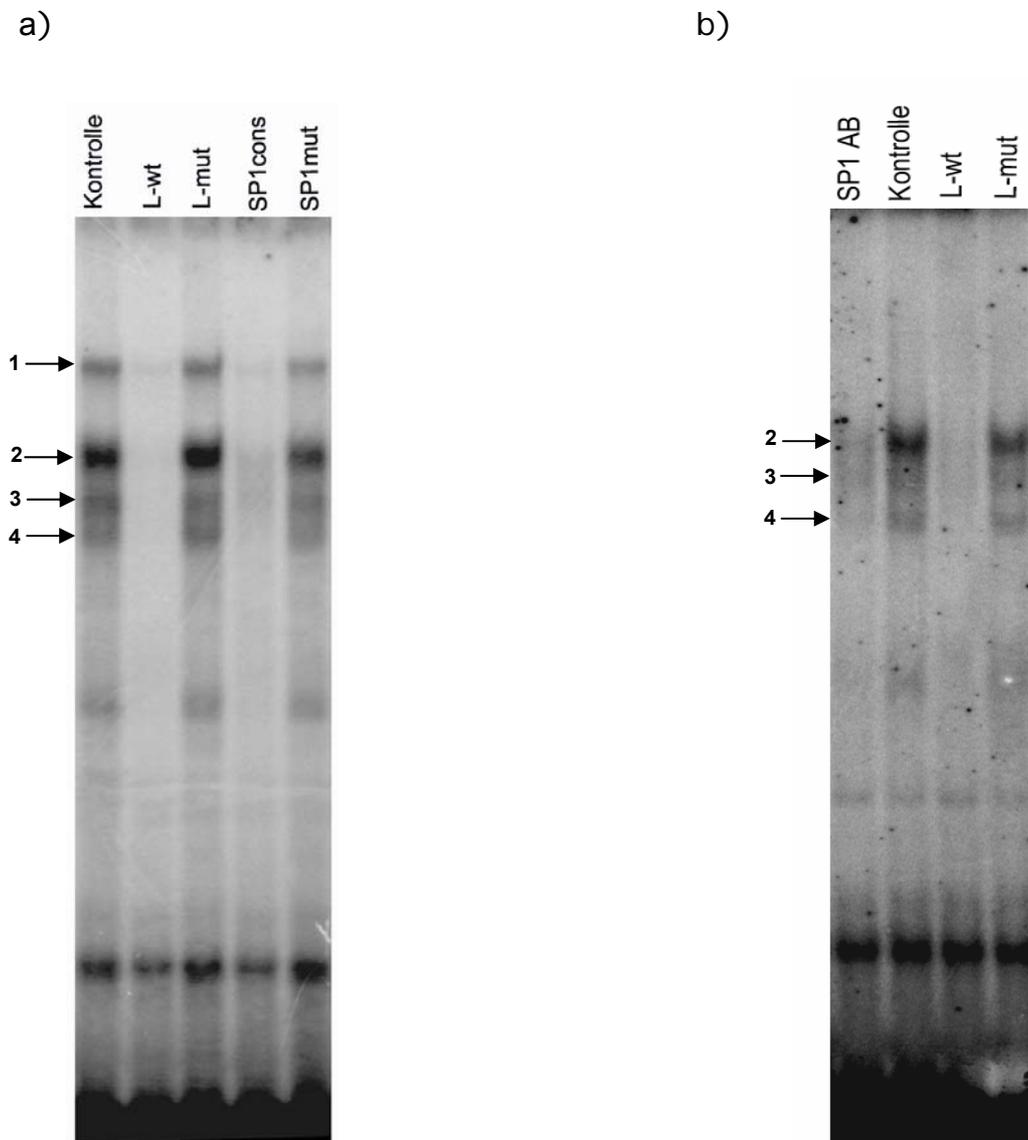
In Abbildung 33 a) ist das Ergebnis einer EMSA-Untersuchung mit dem WT-HERV-L-SP1F-Oligonukleotid gezeigt. In Spur 1 befindet sich die radioaktiv markierte Probe zusammen mit dem KLE-Kernextrakt ohne Kompetitor (Kontrolle). Es sind deutliche Verschiebungen des markierten Oligonukleotids zu höheren Molekulargewichten („Shifts“) erkennbar, die auf mehrere Protein-DNA-Komplexe hinweisen. Wie die Konkurrenzergebnisse in den Spuren 2 bis 5 zeigen, handelt es sich bei den Komplexen 1 bis 4 um sequenzspezifische Protein-DNA-Bindungen. Die in Spur 2 (L-wt) gezeigte Konkurrenz mit ca. 100-fachem Überschuß des unmarkierten Wildtypoligonukleotids führt zur vollständigen Konkurrenz, es sind keine

## 4. Ergebnisse

---

Verschiebungen mehr erkennbar. In Spur 3 (L-mut) wurde als Kompetitor die modifizierte Wildtypsequenz eingesetzt. Hier wurde unter Verwendung der Software „Sequence Shaper“ die potentielle SP1F-Bindungsstelle ausgeschaltet, und es konnte gezeigt werden, dass dieses Sequenzmotiv für die Proteinbindung essentiell ist, da keine Konkurrenz stattfindet. In Spur 4 (SP1F cons) wurden SP1F-Konsensus-Oligonukleotide (Hersteller Santa Cruz) als Kompetitoren verwendet, was zu einer fast vollständigen Konkurrenz der Bindung des radioaktiven Oligonukleotids führte. Die in Spur 5 (SP1F mut) eingesetzten Kompetitoren enthielten eine inaktivierte SP1F-Konsensus-Sequenz und bewirkten keine Konkurrenz. Die Untersuchung der potentiellen SP1F-Bindungsstelle in der 5'-HERV-L-LTR(Cord) lässt somit eine Einordnung der TF-Bindungsstelle in die Familie SP1F der Transkriptionsfaktorbindungsstellen zu. Um zu klären, welcher Transkriptionsfaktor der SP1F-Familie an diese Sequenz bindet, wurden „Supershift“-Untersuchungen mit einem Antikörper gegen den Transkriptionsfaktor Sp1 (Santa Cruz) durchgeführt. Die Inkubation der Reaktionsansätze erfolgte wegen des Antikörpers bei Raumtemperatur. Die Ergebnisse sind in Abbildung 33 b) dargestellt. In Spur 1 (SP1 AB) ist der Ansatz mit dem Sp1-Antikörper zu sehen. Spur 2 zeigt die Reaktion ohne Antikörper (Kontrolle), Spur 3 (L-wt) die Konkurrenzreaktion mit dem Wildtyp-HERV-L-SP1-Oligonukleotid und Spur 4 (L-mut) die Konkurrenzreaktion mit dem mutierten HERV-L-SP1-Oligonukleotid. Durch die Inkubation bei Raumtemperatur zeigte sich ein etwas anderes Ergebnis als bei 4°C, die Banden waren deutlich schwächer und Komplex 1 fehlte vollständig. Im Inkubationsansatz mit dem Sp1-Antikörper zeigte sich eine deutliche Reaktion des Antikörpers mit den Transkriptionsfaktoren. In diesem Fall führte die Antikörper-Transkriptionsfaktor-Bindung jedoch zu einer Blockierung der DNA-Bindungsstelle des Antikörpers („Antikörper-Blocking“), weshalb kein Komplex mehr erkennbar war. Ein solches „Antikörper-Blocking“ zeigt, genau wie ein „Supershift“, eine spezifische Transkriptionsfaktor/Antikörper-Bindung an.

Nach diesen Ergebnissen handelt es sich bei den an die potentielle SP1F-Bindungsstelle der 5'-HERV-L-LTR(Cord) bindenden Proteinen tatsächlich um den Transkriptionsfaktor Sp1.



**Abb. 33:** EMSA mit **KLE-Kernextrakt** zur Untersuchung der potentiellen **SP1F-Bindungsstelle** (alle Ansätze enthielten KLE-Kernextrakt). a) Kontrolle: <sup>32</sup>P markiertes WT-HERV-L-SP1F-Oligonukleotid („Hot Probe“, HP); L-wt: HP + Kompetitor; L-mut: HP + mutierter Kompetitor; SP1F cons: HP + Konsensus-Kompetitor; SP1F mut: HP + mutierter SP1F-Kompetitor. b) SP1 AB: <sup>32</sup>P gelabeltes SP1-Oligonukleotid + Sp1-Antikörper; Kontrolle: gelabeltes Oligonukleotid; L-wt: HP + SP1F-Wildtypkompetitor; L-mut: HP + SP1F inaktivierter Wildtypkompetitor.

#### 4. Ergebnisse

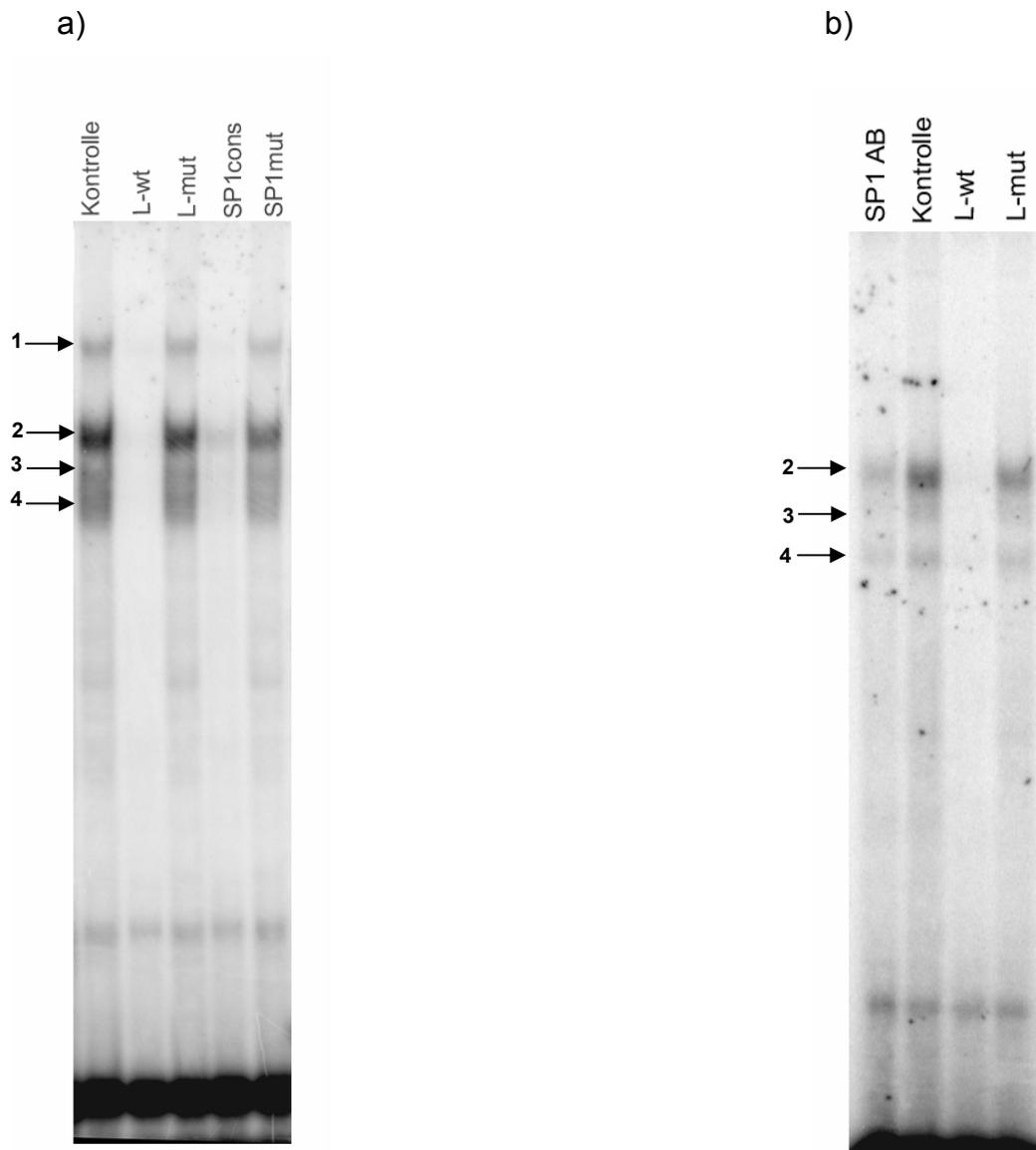
---

In Abbildung 34 a) sind die Ergebnisse des EMSA mit der radioaktiv markierten Probe („Hotprobe“) WT-HERV-L-SP1F und dem HaCaT-Kernextrakt gezeigt. In Spur 1 (Kontrolle) wurde der Reaktionsansatz mit den SP1F-Wildtyp-Oligonukleotiden aufgetragen, insgesamt sind vier Protein-DNA-Komplexe zu erkennen. In Spur 2 (L-wt) wurde zusätzlich der WT-HERV-L-SP1F-Kompetitor zugegeben, die Komplexe wurden kompetitiert. In Spur 3 (L-mut) wurde der mutierte HERV-L-SP1F-Kompetitor zugegeben, man erkennt, dass die Komplexe durch diese Kompetitoren nicht beeinflusst wurden, was bedeutet, dass hier die mutierte Sequenz für die Proteinbindung verantwortlich ist. In Spur 4 (SP1cons) wurde ein kommerziell erworbener SP1F-Konsensus-Kompetitor (Santa Cruz) zugegeben, die Komplexe wurden kompetitiert. Nach Zugabe von kommerziell erworbenen Oligonukleotiden mit mutierter SP1F-Bindungsstelle in Spur 5 (SP1mut) zeigte sich keine Auswirkung auf die Komplexe.

In Abbildung 34 b) sind die Ergebnisse der „Supershift“-Untersuchungen mit dem Antikörper Sp1 (Spur 1) dargestellt. Es zeigte sich, dass der Antikörper Sp1 mit den an die SP1F-Bindungsstelle bindenden Proteinen wechselwirkt. Die dadurch verhinderte Bindung der Transkriptionsfaktoren an die DNA verhindert das „Shiften“ der radioaktiven Probe. Als Kontrollen wurde die „Hot Probe“ ohne Kompetitor eingesetzt (Kontrolle), wobei im Gegensatz zur bei 4°C durchgeführten Inkubation (Gel A, Kontrolle) lediglich Komplex 2 vorhanden war. Die in den Spuren 4 (L-wt) und 5 (L-mut) dargestellten Kompetitionsreaktionen bestätigen die sequenzspezifische Proteinbindung an die Wildtyp-HERV-L-SP1F-Sequenz.

Die Ergebnisse zeigen, dass auch unter Verwendung des HaCaT-Kernextraktes eine Bindung des Transkriptionsfaktors Sp1 an die potentielle SP1F-Bindungsstelle der 5'-HERV-L-LTR(Cord) stattfindet.

## 4. Ergebnisse



**Abb. 34:** EMSA zur Untersuchung der potentiellen **SP1F-Bindungsstelle** in der 5'-HERV-L-LTR(Cord) mit  $^{32}\text{P}$ -markierten HERV-L-SP1F-Oligonukleotiden („Hot Probe“) und **HaCaT-Kernextrakt** (alle Ansätze enthielten HaCaT-Kernextrakt). a) Kontrolle: HP; L-wt: HP + HERV-L-SP1F-Kompetitor; L-mut: HP + mutierter (inaktiver) HERV-L-SP1F-Kompetitor; SP1 cons: HP + Konsensus-SP1-Kompetitor; SP1 mut: HP + mutierter (inaktiver) SP1F-Konsensus-Kompetitor. b) „Supershifts“: SP1 AB: HP + Sp1-Antikörper; Kontrolle: HP; L-wt: HP + HERV-L-SP1F-Kompetitor; L-mut: HP + mutierter (inaktiver) HERV-L-SP1F-Kompetitor.

#### 4. Ergebnisse

---

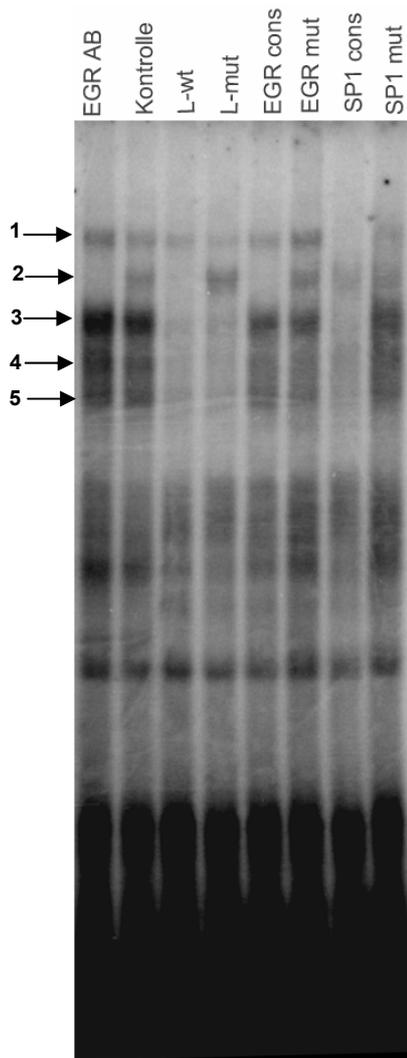
In Abbildung 35 sind die Ergebnisse der Untersuchung der potentiellen EGRF-Bindungsstelle innerhalb der 5'-HERV-L-LTR(Cord) mit dem KLE-Kernextrakt dargestellt. In Spur 1 (EGR AB) waren zusätzlich zur radioaktiv markierten Probe („Hot Probe“, HP) Egr-1-Antikörper vorhanden. Im Vergleich zu Spur 2 (Kontrolle), in der nur die radioaktive Probe und der Kernextrakt vorhanden waren, ist hier ein Protein-DNA-Komplex weniger zu sehen (siehe Markierung). In Spur 1 (EGR AB) ist also ein „Antikörper-Blocking“ erkennbar, die Egr-1-Antikörper binden an die Egr-1-Proteine, die dann nicht mehr an die DNA binden können, und somit keinen Protein-DNA-Komplex mehr verursachen. In Spur 3 (L-wt) ist die Kompetitionsreaktion mit den HERV-L-WT-EGRF-Oligonukleotiden aufgetragen, vier der in Spur 2 (Kontrolle) vorhandenen Komplexe wurden kompetitiert. Bei Konkurrenz mit den mutierten HERV-L-EGRF-Oligonukleotiden (inaktivierte EGRF-Bindungsstelle) waren Komplex 1 und Komplex 2 von der Konkurrenz nicht betroffen, während Komplex 3, 4 und 5 kompetitiert wurden. In Spur 5 (EGRcons) wurden als Konkurrenten EGRF-Konsensus-Oligonukleotide (Santa Cruz) verwendet, es wurde ausschließlich Komplex 2 kompetitiert. Als Kontrolle wurde in dem Reaktionsansatz der in Spur 6 (EGRmut) zu sehen ist, mit mutierten (inaktiven) EGRF-Oligonukleotiden kompetitiert, wobei Komplex 2 unbeeinflusst blieb.

Die hier beschriebenen Konkurrenzreaktionen und der mit dem Egr-1-Antikörper durchgeführte „Supershift“ zeigen, dass es sich bei der in Komplex 2 vorhandenen Proteinbindung um den Transkriptionsfaktor Egr-1 handelt.

Die in Spur 7 (SP1 cons) dargestellte Untersuchung mit dem SP1-Konsensus-Konkurrent, bzw. in Spur 8 (SP1 mut) mit dem mutierten (inaktiven) SP1-Konkurrent, zeigt, dass es sich bei Komplex 1 um eine SP1F-Bindung handelt. In Spur 7 (SP1 cons) ist klar zu sehen wie der SP1F-Konsensus-Konkurrent die Komplexe 1, 3, 4 und 5 kompetitiert während Komplex 2 unverändert bleibt. Es handelt sich hierbei um die in Abb. 32 und Abb. 33 bereits beschriebene Sp1-Bindungsstelle, die, da beide Bindungsstellen nah beieinander liegen, auch mit dem WT-HERV-L-EGRF-Oligonukleotid detektiert wird.

## 4. Ergebnisse

---



**Abb. 35:** EMSA mit **KLE-Kernextrakt** zur Untersuchung der potentiellen **EGRF-Bindungsstelle** in der 5'-HERV-L-LTR(Cord). Alle Reaktionsansätze enthielten die  $^{32}\text{P}$ -markierte WT-HERV-L-EGRF Probe („Hot Probe“, HP) und KLE-Kernxtrakt. EGR AB: HP + Egr-1-Antikörper („Supershift“); L-wt: HP + WT-EGRF-Oligonukleotide als Kompetitor; L-mut: HP + mutierte WT-EGRF-Oligonukleotide als Kompetitoren; EGR cons: HP + EGRF-Konsensus-Oligonukleotide als Kompetitoren (Santa Cruz); EGR mut: HP + mutierte EGRF-Konsensus-Oligonukleotide (Santa Cruz); SP1 cons: HP + SP1-Konsensus-Oligonukleotide als Kompetitoren (Santa Cruz); SP1 mut: HP + mutierte SP1-Konsensus-Oligonukleotide als Kompetitoren (Santa Cruz).

#### 4. Ergebnisse

---

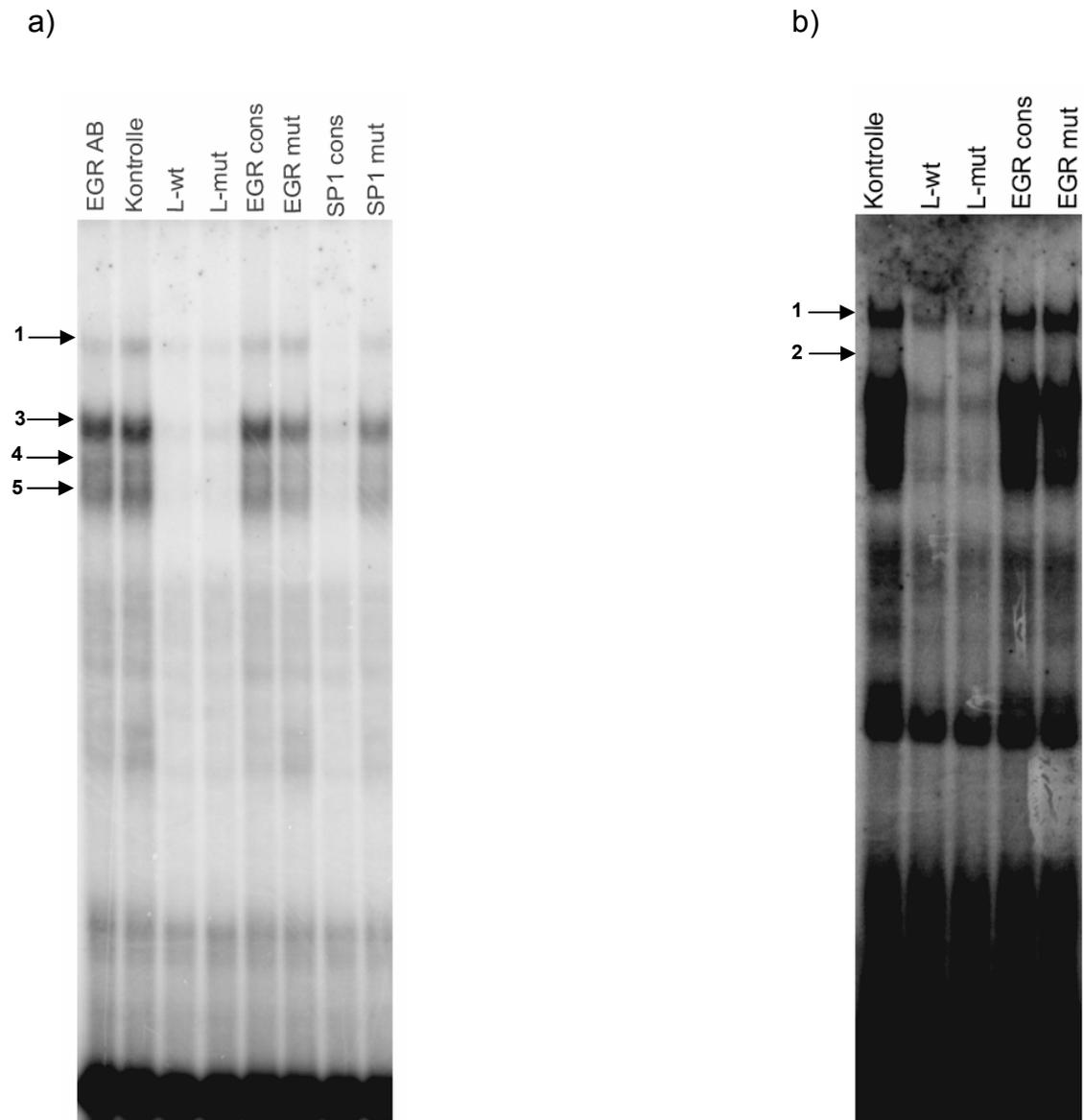
In Abbildung 36 a) sind die Ergebnisse der Untersuchung der potentiellen EGRF-Bindungsstelle der 5'-HERV-L-LTR(Cord) mit dem HaCaT-Kernextrakt dargestellt. In Spur 1 (EGR AB) wurde der „Supershift“-Reaktionsansatz aufgetragen, der die <sup>32</sup>P-markierte WT-HERV-L-EGRF-Probe („Hot Probe“) zusammen mit Egr-1-Antikörpern (Santa Cruz) enthielt. In Spur 2 (Kontrolle) war lediglich die „Hotprobe“ enthalten. Es zeigten sich in beiden Spuren mehrere Protein-DNA-Komplexe, wobei in Spur 1 (EGR AB) kein „Supershift“ vorhanden war. In Spur 3 (L-wt) aufgetragenen Reaktionsansatz waren zusätzlich zur radioaktiv markierten Probe WT-HERV-L-EGRF-Kompetitoren vorhanden, durch die die Komplexe kompetitiert werden konnten. In Spur 4 (L-mut) wurden als Kompetitoren Oligonukleotide zugegeben, deren potentielle EGRF-Bindungsstelle mutiert war (mutierte HERV-L-EGRF-Oligonukleotide). Alle in Spur 2 (Kontrolle) vorhandenen Komplexe konnten durch den mutierten HERV-L-EGRF-Kompetitor kompetitiert werden, was darauf hinweist, dass die EGRF-Bindungsstelle nicht für die hier vorliegenden Proteinbindungen verantwortlich ist. Die Untersuchungen mit den Konsensus-EGRF-Kompetitoren (Santa Cruz) in Spur 5 (EGR cons) zeigten ebenfalls keine Kompetitionsreaktion. In Spur 7 (SP1 cons) wurden SP1F-Konsensus-Kompetitoren (Santa Cruz) eingesetzt, mit denen die Komplexe kompetitiert werden konnten. Spur 8 (SP1 mut) enthielt Konsensus-Kompetitoren deren SP1F-Bindungsstelle inaktiviert war, hier hatten die Kompetitoren keinen Einfluss auf die Protein-DNA-Komplexe.

Dadurch, dass im HERV-L-EGRF-Oligonukleotid durch Überlappung ebenfalls die vorher getestete SP1F-Sequenz vorhanden war, wurden auch hier SP1F-Komplexe detektiert.

In einem anderen Versuchsdurchlauf (Abbildung 36 b) konnte zusätzlich zu den durch SP1 verursachten „Shifts“ ein weiterer Proteinkomplex nachgewiesen werden (Komplex 2), der auf einer EGRF-Bindung beruht. Dies konnte durch die entsprechenden HERV-L-EGRF- und EGRF-Konsensus-Kompetitoren bestätigt werden.

Zusammenfassend ist zu sagen, dass mit dem HaCaT-Kernextrakt lediglich ein schwacher Egr-1-„Shift“ nachgewiesen werden konnte. In den HaCaT-Zellen wird offensichtlich Egr-1 deutlich schwächer exprimiert als in den KLE-Zellen.

## 4. Ergebnisse



**Abb. 36: EMSA mit HaCaT-Kernextrakt** zur Untersuchung der potentiellen **EGRF-Bindungsstelle** in der 5'-HERV-L-LTR(Cord). Alle Reaktionsansätze enthielten die  $^{32}\text{P}$ -markierte WT-HERV-L-EGRF Probe („Hot Probe“, HP) und HaCaT-Kernextrakt. EGR AB: HP + Egr-1-Antikörper („Supershift“); Kontrolle: HP; L-wt: HP + WT-EGRF-Oligonukleotide als Kompetitor; L-mut: HP + mutierte WT-EGRF-Oligonukleotide als Kompetitoren; EGR cons: HP + EGRF-Konsensus-Oligonukleotide als Kompetitoren (Santa Cruz); EGR mut: HP + mutierte EGRF-Konsensus-Oligonukleotide (Santa Cruz); SP1 cons: HP + SP1F-Konsensus-Kompetitoren (Santa Cruz); SP1 mut: HP + mutierte SP1F-Konsensus-Kompetitoren. a) Keine EGR-Bindung erkennbar. b) Schwacher Egr-1-„Shift“ (Komplex 2) erkennbar.

## 4. Ergebnisse

---

In Abbildung 37 ist ein Ausschnitt der 5'-HERV-L-LTR(Cord) gezeigt, der die HOXF-Bindungsstelle enthält. Außerdem sind die Oligonukleotide WT-HERV-L-HOXF und WT-HERV-L-AP1F dargestellt, die zur Untersuchung der potentiellen HOXF- und AP1F-Transkriptionsfaktorbindungsstellen verwendet wurden.

5' -HERV-L-LTR: TGTTACCTAAGGAGATTAACATTTGAGTCAGTGGACTGGGAGAGGCAGACCTACC  
CTGGGTCCACCCAGACAATGGATTCTCTAATTGTAAACTCAGTCATCTGGATGG

HOXF (Core Sim.: 1,0/Matrix Sim.0,967)

WT-HERV-L-HOXF: TTACCTAAAGGAGATTAACATTTGAGTCAGT  
AATGGATTCTCTAATTGTAAACTCAGTCA

HOXF MUT

HERV-L-HOXF-MUT: TTACCTAAGGACATTTACATTTGAGTCAGT  
AATGGATTCTCTGTAATGTAAACTCAGTCA

AP1F(Matrix Sim.: 1,0/CoreSim.0,991)

WT-HERV-L-AP1F: AGGAGATTAACATTTGAGTCAGTGGACTGG  
TCCTCTAATTGTAAACTCAGTCACCTGACC

AP1F MUT

WT-HERV-L-AP1F-MUT: AGGAGATTAGCATTTGTGTCCGTGGACTGG  
TCCTCTAATCGTAAACACAGGCACCTGACC

**Abb. 37:** Ausschnitt der 5'-HERV-L-LTR(Cord) mit den zur Untersuchung der potentiellen TF-Bindungsstellen HOXF und AP1F verwendeten Oligonukleotiden WT-HERV-L-HOXF und WT-HERV-L-AP1F inklusive „Matrix“ und „Core Similarity“ (Genomatix).

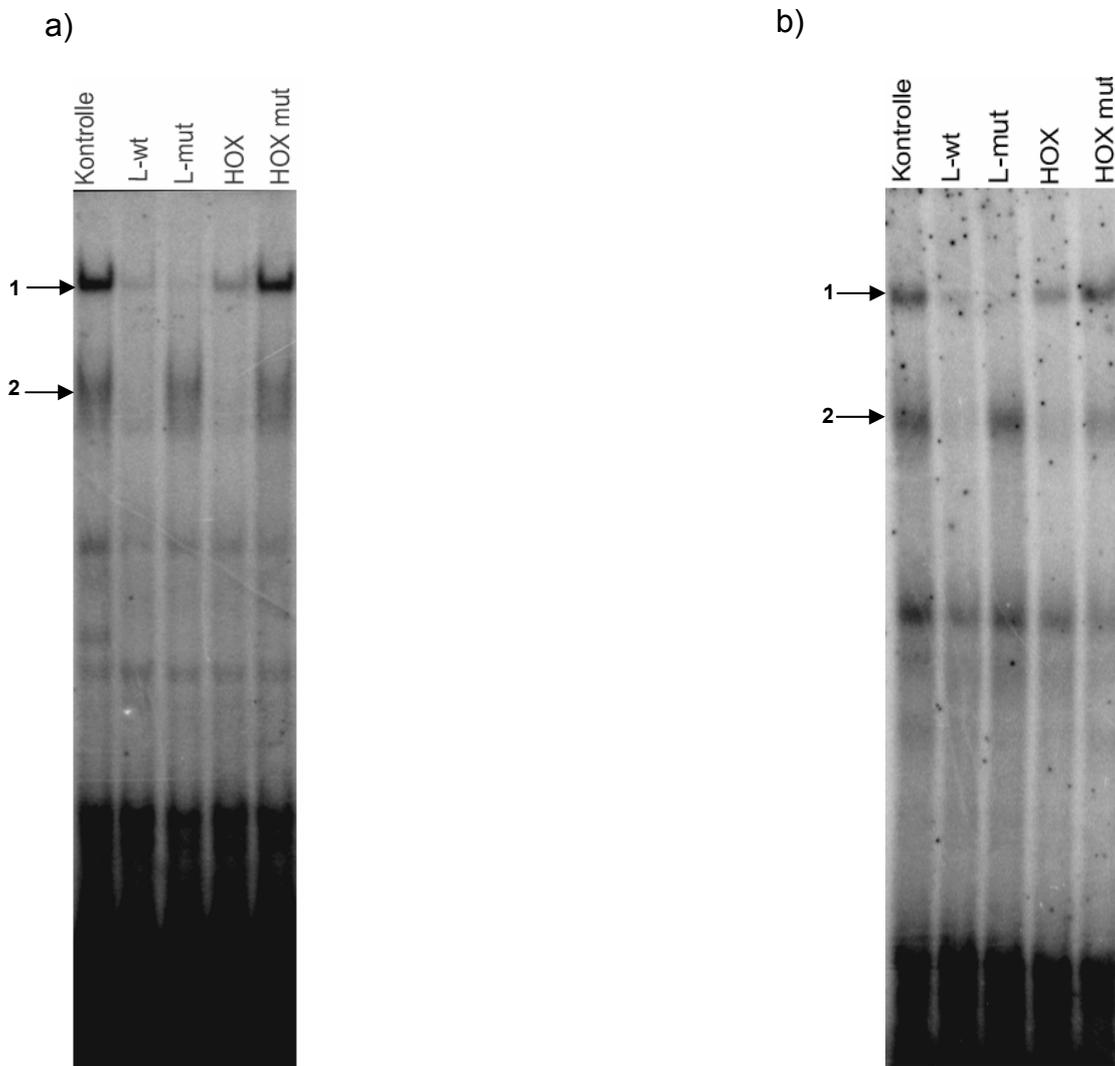
#### 4. Ergebnisse

---

Die in Abbildung 38 gezeigten EMSAs wurden zum einen mit HaCaT- (Abbildung 38 a), zum andern mit KLE-Kernextrakten (Abbildung 38 b) durchgeführt. In Spur 1 (Kontrolle) ist jeweils nur die <sup>32</sup>P-markierte Probe enthalten (WT-HERV-L-HOXF-Oligonukleotide), in Spur 2 (L-wt) wurden jeweils WT-HERV-L-HOXF-Oligonukleotide als Kompetitoren zugegeben. In Spur 3 (L-mut) ist jeweils die Kompetitionsreaktion mit den mutierten WT-HERV-L-HOXF-Kompetitoren zu sehen. In Spur 4 (HOX) wurden die HOXF-Konsensuskompetitoren eingesetzt, während in Spur 5 (HOX mut) jeweils die mutierten HOXF-Konsensus Kompetitoren zum Einsatz kamen. Es zeigte sich bei beiden Kernextrakten (HaCaT; KLE), dass hier eine spezifische HOXF-Bindung vorliegt (Komplex 2). Komplex 1 geht, wie die Kompetitionsreaktionen eindeutig zeigen, auf Proteine zurück, die nicht in die Familie HOXF einzuordnen sind und nicht im Bereich der hier untersuchten TF-Bindungsstelle binden.

## 4. Ergebnisse

---



**Abb. 38: EMSAs** zur Untersuchung der potentiellen **HOXF-Bindungsstelle** in 5'-HERV-L-LTR(Cord).  
**a) Mit HaCaT-Kernextrakt**; Kontrolle:  $^{32}$ P-markierte WT-HERV-L-HOXF-Oligonukleotide („Hot Probe“, HP); L-wt: HP + WT-HOXF-Kompetitoren; L-mut: HP + mutierte WT-HOXF-Kompetitoren; HOX: HP + HOXF-Konsensus-Kompetitoren; HOX mut: HP + mutierte HOXF-Konsensus-Kompetitoren.

**b) Mit KLE-Kernextrakt**, Auftragschema siehe a).

#### 4. Ergebnisse

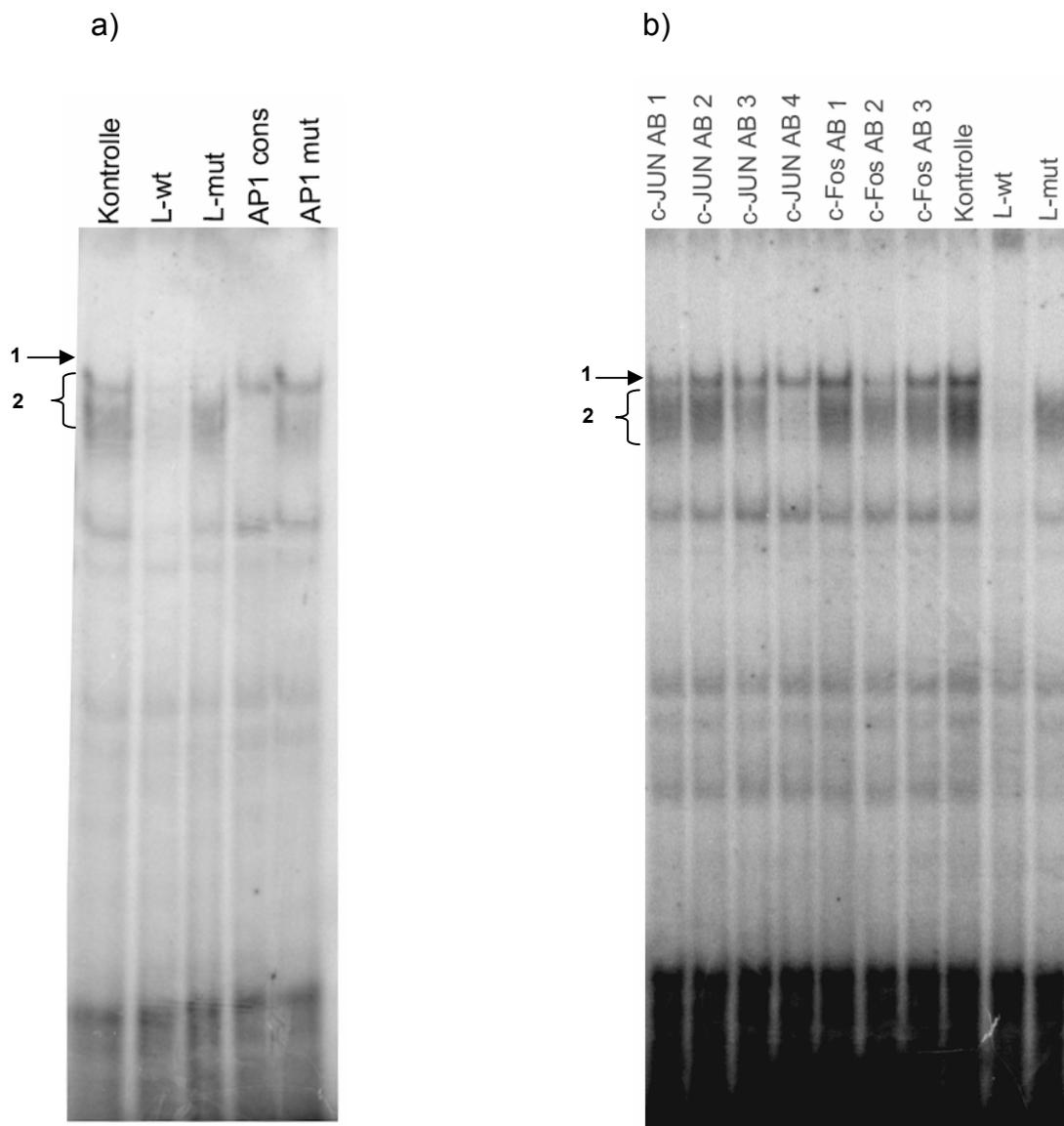
---

In den Abbildungen 39 a) und 39 b) sind die Ergebnisse der Untersuchung der potentiellen AP1F-Bindungsstelle mittels EMSA dargestellt. In Abbildung 39 a) sind die Kompetitionsreaktionen abgebildet. Komplex 1 resultiert, wie sich in der Kompetitionsreaktion mit dem mutierten HERV-L-AP1F-Oligonukleotid in Spur 3 (L-mut) zeigt, nicht von einer spezifischen Proteinbindung an die hier untersuchte potentielle HERV-L-AP1F-Sequenz. Auch die Kompetitionsreaktion mit AP1F-Konsensus-Oligonukleotiden zeigten, dass der obere Komplex nicht beeinflusst wurde. Anders hingegen Komplex 2, hier zeigt sich in der Kompetitionsreaktion mit dem inaktivierten WT-HERV-L-AP1F-Kompetitoren (Spur 3, L-mut), dass dieser Komplex nicht durch die Kompetitoren beeinflusst wurde. Außerdem lässt sich mit AP1F-Konsensus-Oligonukleotiden (Spur 4, AP1 cons) der Komplex 2 kompetitieren. In der Kompetitionsreaktion mit den inaktivierten AP1F-Konsensus-Oligonukleotiden (Spur 5, AP1 cons) zeigt sich wiederum, dass die Komplexbildung kaum beeinflusst wird. Insgesamt zeigen diese Ergebnisse, dass an der potentiellen AP1F-Bindungsstelle Transkriptionsfaktoren der AP1F-Familie binden.

In der Abbildung 39 b) sind die Ergebnisse der weiteren Untersuchung dieser Bindungsstelle dargestellt. Es wurden mit einer Reihe verschiedener c-JUN und c-FOS Antikörper (Monomere von AP1) „Supershifts“ durchgeführt (Spur 1 bis 7). Deutliche Wechselwirkungen mit AP1F wurden in Spur 3 (c-Jun AB3) und Spur 4 (c-Jun AB4) mit c-Jun bindenden Antikörpern festgestellt.

Spur 9 (Kontrolle) zeigt den AP1-Protein-DNA-Komplex ohne Kompetitoren und ohne Antikörper, Spur 10 (L-wt) nochmals die Kompetitionsreaktion mit dem WT-HERV-L-AP1F-Kompetitor und Spur 11 (L-mut) die Kompetitionsreaktion mit den inaktivierten WT-HERV-L-AP1F-Kompetitoren.

Zusammenfassend lässt sich sagen, dass das an die AP1F-Transkriptionsfaktorbindungsstelle bindende Protein mit Anti-AP1-Antikörpern reagiert. Es handelt sich hier also um den Transkriptionsfaktor AP1.



**Abb. 39: EMSA** mit potentieller **AP1F-Transkriptionsfaktorbindungsstelle** enthaltenden Oligonukleotiden aus der 5'-HERV-L-LTR(Cord), getestet mit **HaCaT-Kernextrakt**. Alle Reaktionsansätze enthielten den HaCaT-Kernextrakt. a) Kontrolle:  $^{32}\text{P}$ -markierte WT-HERV-L-AP1F-Oligonukleotide („Hot Probe“, HP); L-wt: HP + WT-AP1F-Kompetitoren; L-mut: HP + inaktivierte WT-AP1F-Kompetitoren; AP1 cons: HP + AP1F-Konsensus-Kompetitor; AP1 mut: HP + inaktivierte AP1F-Konsensus Kompetitoren; b) Spuren 1-7 „Supershifts“. Spur 1: HP + Antikörper c-Jun (N); Spur 2: HP + Antikörper c-Jun (H79); Spur 3: HP + Antikörper c-Jun/AP1 (N)x; Spur 4: HP + Antikörper c-Jun/AP1 (D)x ; Spur 5: HP + Antikörper c-Fos (4)x; Spur 6: HP + Antikörper FosB (102)x; Spur 7: HP + Antikörper c-Fos (4-10G)x ; Spur 8: HP; Spur 9: HP + WT-AP1F-Kompetitor; Spur 10: HP + inaktivierte WT-AP1F-Kompetitoren.

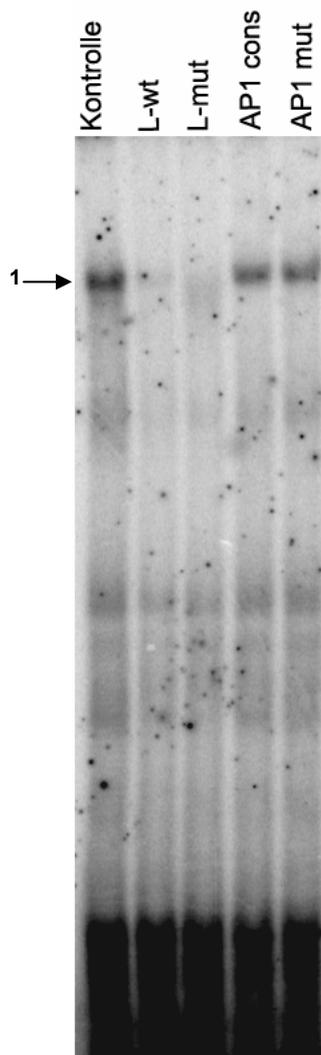
#### 4. Ergebnisse

---

In Abbildung 40 sind die Ergebnisse der Untersuchung der AP1F-Bindungsstelle mit KLE-Kernextrakten dargestellt. Der in Spur 1 (Kontrolle) erkennbare Protein-DNA-Komplex, der durch die Proteinbindung an die <sup>32</sup>P-markierten WT-HERV-L-AP1-Oligonukleotide zustande kommt, wurde in Spur 2 (L-wt) durch den WT-HERV-L-AP1F-Kompetitor kompetitiert. Die Kompetitionsreaktion mit dem mutierten HERV-L-AP1F-Kompetitor (Spur 3; L mut) führte ebenfalls zur kompletten Konkurrenz des Komplexes. Daher lässt sich sagen, dass der in Spur 1 sichtbare Komplex nicht durch Protein-Bindung an die AP1F-Bindungssequenz zustande kommt. Die Konkurrenzreaktion mit dem AP1F-Konsensus-Oligonukleotid (Spur 4, AP1 cons) bestätigte ebenfalls, dass es sich bei dem Komplex nicht um AP1 handelte. In Spur 5 (AP1 mut) ist der Reaktionsansatz mit dem mutierten AP1F-Kompetitor gezeigt, der wie erwartet den in Spur 1 (Kontrolle) und Spur 4 (AP1 cons) vorhandenen Protein-DNA-Komplex enthält.

## 4. Ergebnisse

---



**Abb. 40: EMSA** zur Untersuchung der potentiellen **AP1F-Transkriptionsfaktorbindungsstelle** der 5'-HERV-L-LTR(Cord), getestet mit **KLE-Kernextrakt**. Alle Reaktionsansätze enthielten den KLE-Kernextrakt. Kontrolle:  $^{32}\text{P}$ -markierte WT-HERV-L-AP1F-Oligonukleotide („Hot Probe“, HP); L-wt: HP + WT-HERV-L-AP1F-Kompetitoren; L-mut: HP + inaktivierte AP1F-HERV-L-Kompetitoren; AP1 cons: HP + AP1F-Konsensus-Kompetitor; AP1 mut: HP + inaktivierte AP1F-Konsensus-Kompetitoren.

Bei der Untersuchung der potentiellen AP4R/NEUR-Bindungsstelle konnte weder mit dem HaCaT- noch mit dem KLE-Kernextrakt eine spezifische Proteinbindung nachgewiesen werden (keine Abbildung). Da allerdings die Inaktivierung dieser potentiellen Bindungsstelle in der Mutageneseuntersuchung einen starken Rückgang der Promotoraktivität zur Folge hatte, und in allen untersuchten aktiven HERV-L-LTRs die potentielle AP4R-Bindungsstelle konserviert war, waren möglicherweise die

#### 4. Ergebnisse

hier verwendeten Parameter der Kernextraktion bzw. der EMSA-Durchführung für AP4R-Transkriptionsfaktoren ungeeignet.

In Tabelle 9 sind die Ergebnisse der Untersuchungen potentieller regulatorischer Elemente innerhalb der 5'- und 3'-HERV-L-LTR zusammengefasst.

**Tabelle 9f:** Übersichtstabelle der untersuchten potentiellen Transkriptionsfaktorbindungsstellen innerhalb der 5'- und der 3'-HERV-L-LTR(Cord). Angegeben ist die jeweilige Transkriptionsfaktorfamilie (Matrixfamilie, Genomatix). Bei Identifizierung mit einem spezifischen Antikörper ist der entsprechende Transkriptionsfaktor in Klammern aufgeführt. Mutagenese und EMSAs wurden ausschließlich mit der 5'-HERV-L-LTR(Cord) durchgeführt. (+) = Transkriptionsfaktorbindungsstelle gefunden, bzw. bei Mutagenese ist die Bindungsstelle für die Promotoraktivität relevant; (-) = Transkriptionsfaktor nicht gefunden, bzw. bei Mutagenese ist die Bindungsstelle für die Promotoraktivität nicht relevant; (n.t.) = nicht getestet.

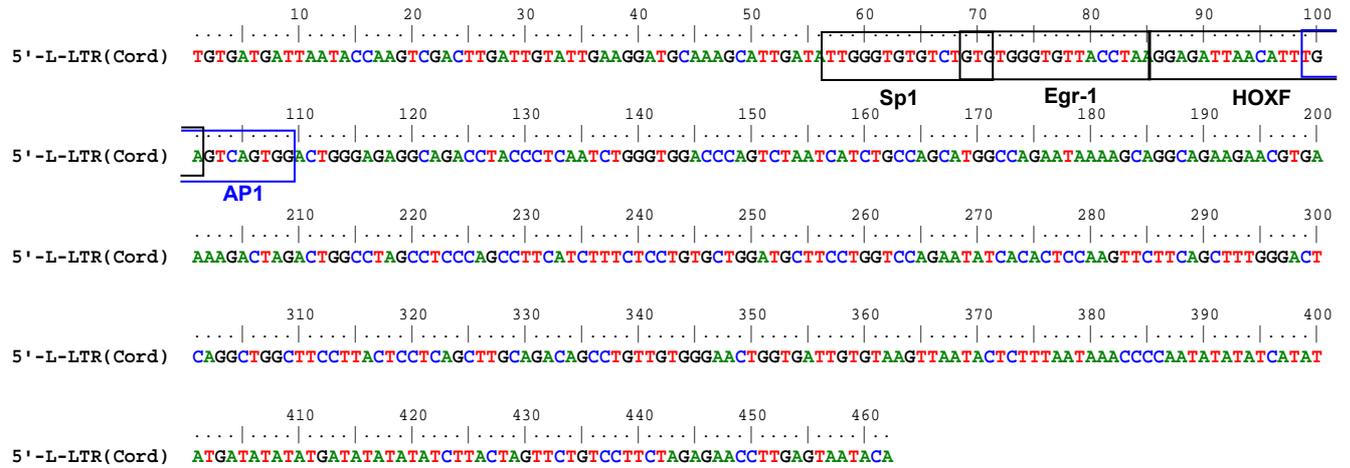
<b><u>TF</u></b> <b><u>(Matrix-</u></b> <b><u>Familie)</u></b>	<b><u>bildet u.a.</u></b> <b><u>Module mit</u></b>	<b><u>Matinspector</u></b>		<b><u>Mutagenese</u></b>		<b><u>EMSA</u></b>	
		<b><u>5'-LTR</u></b>	<b><u>3'-LTR</u></b>	<b><u>HaCaT</u></b>	<b><u>KLE</u></b>	<b><u>HaCaT</u></b>	<b><u>KLE</u></b>
SMAD	AP1F; SP1F	+	-	+	-	n.t.	n.t.
PAX8	/	+	-	+	+	n.t.	n.t.
RCAT	/	+	-	+	+	n.t.	n.t.
SP1F (Sp1)	AP1F; SMAD; EGRF	+	+	+	+	+	+
EGRF (Egr-1)	SP1F; AP1F; HOXF	+	-	+	+	+	+

#### 4. Ergebnisse

<b>TF (Matrix- Familie)</b>	<b><u>bildet u.a. Module mit</u></b>	<b><u>Matinspector</u></b>		<b><u>Mutagenese</u></b>		<b><u>EMSA</u></b>	
		<b><u>5'-LTR</u></b>	<b><u>3'-LTR</u></b>	<b><u>HaCaT</u></b>	<b><u>KLE</u></b>	<b><u>HaCaT</u></b>	<b><u>KLE</u></b>
TBPF	/	+	-	+	-	n.t.	n.t.
HOXF	SP1F; NEUR	+	-	+	+	+	+
AP1F (AP1)	SP1F; SMAD	+	-	+	+	+	-
AP4R	/	+	-	+	+	-	-
NEUR	HOXF	+	-	+	+	-	-
IRFF	/	+	-	+/-	+	n.t.	n.t.

Zusammenfassend lässt sich sagen, dass in der 5'-HERV-L-LTR(Cord) funktionelle Transkriptionsfaktorbindungsstellen für die Transkriptionsfaktoren Sp1, Egr-1, HOXF und AP1 nachgewiesen werden konnten. Die Transkriptionsfaktoren Sp1, Egr-1 und AP1 konnten mittels Antikörpern („Supershift“) eindeutig identifiziert werden. Außerdem konnte durch Kompetitionsuntersuchungen gezeigt werden, dass ein Transkriptionsfaktor der Familie HOXF für die Promotoraktivität essentiell ist. Unterschiede bezüglich der verwendeten Kernextrakte aus der Keratinozytenzelllinie HaCaT und der Uteruszelllinie KLE zeigten sich bei der Untersuchung der AP1-Bindungsstelle, die nur bei Verwendung des HaCaT-Kernextraktes AP1-Bindung zeigte. In Abbildung 41 sind die Transkriptionsfaktorbindungsstellen („Matinspector-Matrizen“) abgebildet.

## 4. Ergebnisse



**Abb. 41:** Darstellung der in der 5'-HERV-L-LTR(Cord) (462 bp) identifizierten Transkriptionsfaktorbindungsstellen („Matinspector“-Matrizen, Kapitel 3.20.1). Schwarz markiert sind die Bindungsstellen der Transkriptionsfaktoren Sp1, Egr-1 und HOXF die im Kernextrakt der Uteruszelllinie KLE und der Keratinozytenzelllinie HaCaT nachgewiesen wurden. Blau markiert ist die Bindungsstelle des Transkriptionsfaktors AP1, der nur im HaCaT-Kernextrakt nachgewiesen werden konnte

### 5 Diskussion

Das menschliche Genom besteht zu ca. 9% aus Endogenen Retroviren (Lander *et al.*, 2001; Venter *et al.*, 2001; Li *et al.*, 2001), das entspricht ungefähr dem Fünffachen der Protein-codierenden Sequenzen. Handelt es sich hierbei lediglich um genetischen Ballast der von den Zellen unnötig hohen Energieeinsatz bei der Verdoppelung der DNA abverlangt? Tatsächlich deutet einiges darauf hin, dass endogene Retroviren durch ihre Fähigkeit zur intrachromosomalen Rekombination evolutionär von Bedeutung sind. Außerdem haben sie teilweise zelluläre Funktionen übernommen, wie z.B. das HERV-W-Genprodukt Syncytin-1 (Blond *et al.*, 2000; Li *et al.*, 2000) und das von HERV-FRD exprimierte Syncytin-2 (Blaise *et al.*, 2003), die an der Syncytiotrophoblastenentstehung in der Plazenta beteiligt sind. Eine HERV-E-LTR, die als Enhancer bei Speichel-Amylasegenen (Ting *et al.*, 1992) fungiert, oder ein HERV-K(HML-10)-Element, dessen 3'-LTR als Promotor die plazentaspezifische Expression des *INSL4*-Gens steuert (Bieche *et al.*, 2003). In der Familie ERV-L konnte eine solitäre ERV-L identifiziert werden, die beim Menschen und anderen Primaten im Kolon, als alternativer Promotor die Expression von *β3Gal-T5* steuert (Dunn *et al.*, 2003; Dunn *et al.*, 2005). Insgesamt schätzt man, dass Retrotransposons, darunter überwiegend retrovirale LTRs, zu etwa 20% der humanen Gene regulative Sequenzen beisteuern (Van de Lagemaat, 2003; Jordan, 2003).

#### 5.1 HERVs und Gewebespezifität

Um der Frage nachzugehen, in welchen Geweben welche HERV-Elemente aktiv sind, untersuchten Seifarth *et al.* (2005) mittels HERV-Microarray die *pol*-Expression der Familien HERV-I, HERV-T, HERV-H, HERV-F, HERV-W, HERV-R, ERV-9, HERV-K Superfamilie und der Familie HERV-L in verschiedenen humanen Geweben. Hier konnte festgestellt werden, dass generell die Klasse I und Klasse II HERVs aktiver sind als Vertreter der Klasse III, was vermutlich mit dem Alter der

Elemente zu tun hat (Landry *et al.*, 2002). Klasse I und Klasse II Elemente integrierten erst vor 40 bis 50 Millionen Jahren ins Primatengenom, während die Vertreter der Klasse III bereits vor 70 Millionen Jahren integrierten. Sie sind daher deutlich länger der Anhäufung von inaktivierend wirkenden Mutationen ausgesetzt. Fast alle HERV-Familien, auch die Klasse III HERVs, sind in Haut, Schilddrüse, Uterus, Zervix, Hoden und Plazenta aktiv (Seifarth *et al.*, 2005). Die Aktivität in Hoden, Uterus, Plazenta und Zervix hängt sicherlich mit der generell hohen transkriptionellen Aktivität dieser Gewebe zusammen. In Uterus, Plazenta, sowie in der Haut sind aktive endogene Retroviren zudem von Vorteil, da sie vor exogenen Retroviren schützen können, die durch das mütterliche Blut bzw. über die Haut übertragen werden (Larsson und Andersson, 1998). Außerdem ist die Haut ständig anderen Umweltfaktoren ausgesetzt, die die retrovirale Expression beeinflussen können, wie z.B. UV-Licht, das wie Hohenadl *et al.* (1999) zeigen konnten, die Transkription einiger HERVs in primären epidermalen Keratinozyten aktivieren kann.

HERV-Familien, die weniger ubiquitär aktiv sind (z.B. HERV-L), sind deshalb so interessant, da sie auf Grund ihrer gewebespezifischen Expression z.B. als Promotoren in retroviralen Vektoren in der Gentherapie eingesetzt werden könnten. Der Vorteil dieser Promotoren liegt darin, dass sie durch ihre Spezifität einen gezielteren Einsatz erlauben würden. Starke HERV-L-Aktivität konnte in Haut, Uterus, Zervix, Plazenta, Hoden und Schilddrüse nachgewiesen werden, schwache Aktivität war im Skelettmuskeln und im Herz vorhanden (Seifarth *et al.*, 2005; Abb. 42). Eine solche Untersuchung bezieht allerdings alle HERV-*pol*-Elemente der jeweiligen HERV-Familien ein, da hier Bereiche der HERV-Sequenz für die Amplifikation ausgewählt werden, die hochkonserviert sind, und zudem degenerierte Primer verwendet werden. Daher lassen sich lediglich Aussagen über provirale LTRs treffen, die den ausgewählten DNA-Abschnitt nicht durch Deletion verloren haben, außerdem können keine Rückschlüsse auf die Gewebespezifität einzelner HERV-Elemente getroffen werden, da der HERV-Microarray Subfamilien detektiert, die bis zu mehr als 100 Einzelemente umfassen können. Genau das ist allerdings interessant, da die Möglichkeit besteht, dass bestimmte HERV-LTRs sehr gewebespezifisch promotoraktiv sein können. Beispielsweise könnten unter den 200



### 5.2 5'-HERV-L-LTR(Cord): Zelltypspezifische Promotoraktivität

Das zuerst von Cordonnier *et al.* (1995) beschriebene provirale HERV-L-Element (HERV-L(Cord); Länge 6591bp) wurde als erstes Element der HERV-L-Familie untersucht. Es wurde zunächst die 5'-LTR mittels Luziferase-Assays auf ihre Promotoraktivität getestet. Dafür wurden transiente Transfektionen mit der Keratinozytenzelllinie HaCaT und den primären Keratinozyten NHEK durchgeführt, wobei Promotoraktivität nachgewiesen werden konnte (Schön, unveröffentlicht). Die in diesem Kontext durchgeführten Untersuchungen nach transienten Transfektionen der Zelllinien LC5 (Lungenkarzinom), MiaPaCa2 (Pankreaskarzinom), U373 (Astrozytoma), 85HG66 (Astrozytoma), HeLa (Zervixkarzinom) und MCF7 (Brustkarzinom) zeigten dagegen keinerlei Promotoraktivität des HERV-L-Konstruktes (Schön, 2001). Damit konnte zum ersten Mal bei einer isolierten HERV-L-Sequenz eine strikt zelltypspezifische Promotoraktivität nachgewiesen werden. In der vorliegenden Arbeit wurden die Untersuchungen durch Einbeziehung weiterer unterschiedlicher HERV-L-LTRs fortgesetzt. Zuerst wurden mehrere solitäre und provirale HERV-L-LTRs kloniert und nach taxonomischer Zuordnung vier solitäre und vier provirale Elemente bezüglich ihrer Promotoraktivität getestet. Die verwendeten Zelllinien stammten von Geweben ab, in denen im HERV-Microarray (Seifarth *et al.*, 2005) HERV-L-Aktivität nachgewiesen werden konnte (Keratinozytenzelllinie HaCaT; Uteruszelllinie KLE; Plazentazelllinie JEG-3; Zervixzelllinie HeLa; Schilddrüsenzelllinie SW). Dabei zeigte sich, dass die getesteten solitären HERV-L-LTRs (S 1, S 4, S 5 und S 11) in allen Zelllinien nahezu inaktiv waren. Lediglich S 4 zeigte vor allem in der Zelllinie KLE Promotoraktivität.

Von den proviralen HERV-L-LTRs P 3, P 4, P 7, P 10 und 5'-HERV-L(Cord) war die 5'-HERV-L-LTR(Cord) in HaCaT-, HeLa- und KLE-Zellen am aktivsten. Die beiden LTRs P 4 und P 7 zeigten ebenfalls starke Promotoraktivität, allerdings ausschließlich in der Uteruszelllinie KLE.

Diese Ergebnisse zeigen, dass die 5'-HERV-L-LTR(Cord) ein HERV-L-Element ist, das in den humanen Geweben Haut, Uterus, Zervix aktiv ist. Die proviralen Elemente P 4 und P 7 sind ausschließlich im menschlichen Uterusgewebe aktiv.

Interessant ist, dass drei der fünf untersuchten proviralen LTRs hoch aktiv waren, während die vier untersuchten solitären LTRs schwache Aktivität zeigten. Sicherlich ist die Anzahl der untersuchten Proben nicht ausreichend um statistische Aussagen treffen zu können, doch ist der Aktivitätsunterschied zwischen solitären und proviralen HERV-L-LTRs zumindest auffällig. Eine mögliche Erklärung dafür ist das wahrscheinlich höhere Alter der solitären LTRs, bei denen es sich um Relikte ehemaliger Proviren handelt, die herauskombiniert wurden. Mit höherem Alter wäre auch eine größere Anzahl inaktivierend wirkender Mutationen verbunden.

### **5.3 Lässt sich gewebespezifische Expression der 5'-HERV-L-LTR(Cord) und der 5'-HERV-H-H6-LTR in transgenen Mäusen bestätigen?**

Um zu untersuchen, ob die 5'-HERV-L-LTR(Cord) auch im Organismus gewebespezifische Promotoraktivität zeigt, wurden von Regine Schönfeld (2003) transgene Mäuse erzeugt, wobei die 5'-HERV-L-LTR(Cord) an ein Luziferase-Reportergen gekoppelt war. Außerdem wurden in diesem Zusammenhang transgene Mäuse, die ein HERV-H-H6-LTR-EGFP-Konstrukt enthielten, generiert. Die HERV-H-H6-LTR konnte bereits von Schön *et al.* (2001) als promotoraktiv in diversen Zelllinien gezeigt werden.

Im Rahmen meiner Doktorarbeit wurden einige dieser transgenen Mäuse untersucht, um einen ersten Überblick zu erhalten, ob sich Parallelen zu den Ergebnissen der transienten Zellkulturuntersuchungen ergeben. Für die Untersuchung der 5'-HERV-L-LTR(Cord) stand allerdings lediglich eine Mauslinie zur Verfügung (L6/N3), von welcher ein männliches und ein weibliches Tier untersucht wurden. Dabei konnte lediglich sehr schwache Expression im Herz des männlichen Tieres nachgewiesen werden, nicht aber in den Geweben Haut, Uterus und Zervix, auf die die Zellkulturergebnisse hinweisen. Die Untersuchung der HERV-H-H6-LTR zeigte, dass Promotoraktivität in der Mauslinie L1/N3 in einem weiblichen Tier ausschließlichschließlich in der Haut vorhanden war, in einem männlichen Tier im

Hodengewebe. In der Linie L5/N3 konnte Promotoraktivität lediglich in Rektumgewebe nachgewiesen werden. Die Untersuchung der dritten HERV-H-H6-Linie (L10/N4) zeigte lediglich Aktivität in Hodengewebe. Damit gab es nur eine Übereinstimmung innerhalb der drei untersuchten HERV-H-H6-Mauslinien.

Die Tatsache, dass HERV-L-LTR(Cord)-Aktivität in der einzigen verfügbaren Linie L6/N3 fast nicht nachweisbar ist, sowie die unterschiedlichen Ergebnisse der HERV-H-H6-Untersuchung hängen vermutlich mit der Technik der DNA-Mikroinjektion zusammen. Das Hauptproblem hierbei ist, dass die DNA zufällig integriert. Dadurch können sich, abhängig vom Integrationsort, je nachdem, ob sich Promotoren in der Nähe befinden, unterschiedliche Aktivitätsmuster ergeben. Integration in den Einflussbereich eines „Silencers“ oder in Heterochromatin führen zur völligen Inaktivierung. Ein weiteres Problem der Mikroinjektion ist, dass die Anzahl der Kopien des Transgens, die ins Genom integrieren, nicht beeinflusst werden kann. So können bis zu 1000 Kopien als „Tandem Arrays“ integrieren (Lo *et al.*, 1987). Integration multipler Kopien an einer Integrationsstelle kann die Expressionsrate eines Transgens, unabhängig vom Integrationsort, negativ beeinflussen (Garrick *et al.*, 1998). Außerdem kann durch Methylierung der DNA die Expression inaktiviert werden, ein Vorgang der mit dem Phänomen des „Imprinting“ assoziiert ist. „Imprinting“ wird in Plazentasäugetern beobachtet und führt dazu, dass bestimmte maternale oder parentale Gene während der Keimzellentstehung inaktiviert oder aktiviert werden.

Die Zellkulturergebnisse bezüglich der zelltypspezifischen Promotoraktivität der 5'-HERV-L-LTR(Cord) und der 5'-HERV-H-H6-LTR lassen sich also in den entsprechenden Geweben transgener Mäuse bis jetzt nicht bestätigen. Um endgültige Aussagen über die HERV-L- und HERV-H-H6-Promotoraktivitäten treffen zu können, müssen allerdings, vor allem was HERV-L betrifft, weitere Mauslinien generiert werden.

Ein anderes Beispiel für HERV-transgene Mäuse sind die Untersuchungen von Casau *et al.* (1999) mit HERV-K-LTR-transgenen Mäusen. Hier zeigte sich die stärkste HERV-K-LTR-Aktivität in Hoden, was mit den Untersuchungen an humanem Gewebe übereinstimmt. Allerdings konnte diese Aktivität nur in zwei von vier der

etablierten Mauslinien nachgewiesen werden. Im Gegensatz zu den Untersuchungen an humanem Gewebe konnte keine HERV-K-LTR-Aktivität in Plazenta nachgewiesen werden. Für das an der gewebespezifischen Expression des menschlichen Speicheldrüsengens *AMY1C* beteiligte HERV-E-Element, konnte in Untersuchungen mit transgenen Mäusen gezeigt werden, dass die Expression in den Ohrspeicheldrüsen (*Glandula parotis*) allein auf HERV-E zurückgeht. Die Expression von *AMY1C* konnte beim Menschen in Speicheldrüsen, Lunge und in Ovar-Tumoren detektiert werden. In den HERV-E/*AMY1C*-transgenen Mäusen konnte neben der starken Expression in den Speicheldrüsen eine schwache Expression in Lunge und Ovargewebe nachgewiesen werden (Ting *et al.*, 1992). Diese Beispiele zeigen, dass sich auch in anderen Untersuchungen an HERV-transgenen Mäusen zum Teil Abweichungen zur gewebespezifischen Expression in humanem Gewebe zeigen. Unterschiede im Expressionsmuster von Mensch und Maus können z.B. im unterschiedlichen Expressionslevel von Transkriptionsfaktoren begründet sein, oder durch unterschiedliche Aktivierung von Transkriptionsfaktoren ausgelöst werden.

### **5.4 Untersuchung des Promotors der 5'-HERV-L-LTR(Cord)**

Nachdem gezeigt werden konnte, dass die provirale 5'-HERV-L-LTR(Cord) in Keratinozyten-, Uterus- und Zervixzelllinien aktiv ist, und sie sich überdies als die HERV-L-LTR mit der bisher höchsten Promotoraktivität erwiesen hatte, stellte sich als nächstes die Frage, welche regulativen Sequenzen, z.B. Transkriptionsfaktorbindungsstellen, an der beobachteten zelltypspezifischen Expression beteiligt sind. Anhand von Computerprogrammen („Genomatix; Matinspector“) können potentielle TF-Bindungsstellen identifiziert werden. Vom Programm Matinspector wurden in der im Rahmen dieser Arbeit untersuchten 5'-HERV-L-LTR(Cord), die eine Länge von 462 bp besitzt, 46 potentielle Bindungsstellen vorgeschlagen. Um weitere Informationen über diese potentiellen Bindungsstellen zu erhalten, wurden sie mit der „Module Library“ („Genomatix“) verglichen. Darin sind Kombinationen von Transkriptionsfaktorbindungsstellen aufgeführt, die, wie experimentell gezeigt wurde, in bestimmten Promotoren für die

Promotoraktivität mitverantwortlich sind. Darunter waren die Module SMAD/AP1F, HOXF/EGRF, NEUR/HOXF, EGRF/AP1F, SP1F/EGRF und SP1F/AP1F. Untersuchungen der aktiven LTRs 5'-HERV-L-LTR(Cord), P 4 und P 7 mit dem Computerprogramm „Frameworker“ („Genomatix“) zeigten zudem, dass in diesen Elementen die potentiellen Transkriptionsfaktorbindungsstellen für SP1F und HOXF und AP4R konserviert sind. In der 5'-HERV-L-LTR(Cord) und P 7 ist sogar zusätzlich eine potentielle AP1F-Bindungsstelle vorhanden. Der Abstand zwischen SP1F und HOXF und zwischen HOXF und AP1F ist jeweils exakt gleich groß. Um potentiell relevante Bereiche einzugrenzen, ist zudem der Vergleich der untersuchten Sequenz mit einer verwandten, weniger aktiven, im Idealfall inaktiven Sequenz sehr hilfreich. Daher wurde die Promotoraktivität der 3'-HERV-L-LTR(Cord) in den Zelllinien HaCaT und KLE untersucht, in denen sich die 5'-HERV-L-LTR(Cord) bereits als stark promotoraktiv gezeigt hatte. Dabei konnte festgestellt werden, dass die 3'-HERV-L-LTR(Cord) nahezu inaktiv ist. Um Informationen über mögliche regulatorische Bereiche innerhalb der 5'-HERV-L-LTR(Cord) zu erhalten, wurde sie mit der 3'-HERV-L-LTR(Cord) per Alignment verglichen. Dabei konnten einige Punktmutationen im U3/R- und deutliche Abweichungen im U5-Bereich festgestellt werden. Da die regulatorischen Bereiche einer LTR größtenteils in der U3/R-Region liegen, sollten sich durch Vergleich der entsprechenden Sequenzen von 5'- und 3'-LTR potentiell aktive Transkriptionsfaktorbindungsstellen identifizieren lassen. Die Computeranalyse (Software: „Matinspector“) der beiden LTRs zeigte folgende potentielle Transkriptionsfaktorbindungsstellen, die im gleichen Bereich nur in der aktiven 5'-HERV-L-LTR(Cord) aber nicht in der inaktiven 3'-HERV-L-LTR(Cord) identifiziert werden konnten: SMAD, PAX8, RCAT (CAAT-Box; überschneidet sich mit SP1F), EGRF, AP1F, TBPF (TATA-Box), HOXF, AP4R, NEUR.

Um die durch Computeranalysen identifizierten potentiellen Transkriptionsfaktorbindungsstellen zu überprüfen, wurden gezielt Punktmutationen durch („Site directed Mutagenesis“) in die potentiellen Transkriptionsfaktorbindungsstellen der 5'-HERV-L-LTR(Cord) eingefügt. Anschließend wurden die mutierten LTRs in die Zelllinien HaCaT und KLE transfiziert und getestet, ob diese Sequenzänderungen die Promotoraktivität der LTR beeinflussen. Die Inaktivierung der potentiellen RCAT-Bindungsstelle (CAAT-Box)

hatte gleichzeitig die Inaktivierung der SP1F-Bindungsstelle zur Folge, die in allen bisher bekannten aktiven HERV-L-LTRs (5'-HERV-L-LTR(Cord), P 4 und P 7) hochkonserviert ist. Die potentiellen Bindungsstellen AP4R und NEUR konnten ebenfalls nur gleichzeitig inaktiviert werden.

Im Luziferase-Assay konnte gezeigt werden, dass die Inaktivierung der potentiellen TF-Bindungsstellen AP1F, PAX8, HOXF, RCAT/SP1F, AP4R/NEUR, TBPF und EGRF in den Zelllinien HaCaT und KLE zu einer deutlich verringerten Promotoraktivität führte. Besonders stark wirkte sich das Ausschalten der potentiellen EGRF-Bindungsstelle auf die Aktivität in der Zelllinie HaCaT aus. Hier war nahezu keine Promotoraktivität mehr meßbar. Diese Ergebnisse zeigen, dass zwar größtenteils die gleichen regulatorischen Bereiche der 5'-HERV-L-LTR(Cord) benötigt werden, um die Promotoraktivität in den beiden Zelllinien HaCaT und KLE zu steuern, es aber dennoch Abweichungen gibt. Offensichtlich wird der Bereich, der als potentielle SMAD-Bindungsstelle identifiziert wurde, in den HaCaT-Zellen benötigt, während die potentielle TBPF-Bindungsstelle dort keine Rolle spielt. In den KLE-Zellen ist es umgekehrt, hier ist die potentielle TBPF-Bindungsstelle von Bedeutung, die Inaktivierung der potentiellen SMAD-Bindungsstelle hat allerdings kaum Auswirkungen auf die Promotoraktivität. Obwohl die Promotoraktivität der 5'-HERV-L-LTR(Cord) in den beiden Zelllinien ähnlich stark ist, zeigen sich in der Regulation der Expression leichte Unterschiede. Sollte es sich bei dem Protein, das an die potentielle TBPF-Bindungsstelle bindet, nicht um das ubiquitäre TATA-Boxbindeprotein handeln, sondern um einen anderen Transkriptionsfaktor, der nicht ubiquitär exprimiert wird, wäre eine mögliche Erklärung für dieses Ergebnis, dass für die Promotoraktivität der 5'-HERV-L-LTR(Cord) entweder die potentielle SMAD-Bindungsstelle oder die potentielle TBPF-Bindungsstelle benötigt wird. Ein Faktor kann also das Fehlen des anderen kompensieren. In HaCaT-Zellen wird das potentielle SMAD-Protein exprimiert, nicht aber das an die potentielle TBPF-Bindungsstelle bindende Protein. Beim Inaktivieren der potentiellen SMAD-Bindungsstelle sinkt die Promotoraktivität stark ab, da das zur Kompensation benötigte Protein in der Zelle nicht vorhanden ist. In der Zelllinie-KLE hingegen wird das an die potentielle TBPF-Bindungsstelle bindende Protein exprimiert, nicht aber der potentielle SMAD-Transkriptionsfaktor. Hier wirkt sich das Inaktivieren der

potentiellen TBPF-Bindungsstelle stark aktivitätsmindernd aus, da das alternativ benötigte potentielle SMAD-Protein in HaCaT-Zellen nicht exprimiert wird.

Bioinformatisch können die innerhalb einer Sequenz möglichen Transkriptionsfaktorbindungsstellen ermittelt werden. Ob einige dieser Sequenzen von funktioneller Bedeutung sind, kann mittels Mutagenese („Site directed Mutagenesis“) untersucht werden. Um aber festzustellen, welcher Transkriptionsfaktor tatsächlich an eine bestimmte Sequenz bindet, muss ein Protein an dieser Stelle gefunden und identifiziert werden. Hier kommt der „Electrophoretic Mobility Shift“, kurz EMSA, zum Einsatz. Dabei werden jeweils kurze DNA-Abschnitte, die eine der potentiellen Transkriptionsfaktorbindungsstellen enthalten, radioaktiv markiert und nach Inkubation mit Kernextrakt elektrophoretisch aufgetrennt. Bindet ein Protein an diese Sequenz wandert das Oligonukleotid langsamer durch das Gel, ein „Shift“ wird erkennbar. Durch Kompetitionsreaktionen kann festgestellt werden, ob das Protein spezifisch an einen bestimmten Bereich innerhalb einer DNA-Sequenz bindet. Unter Verwendung von Kompetitoren, die das Konsensus-Erkennungsmotiv einer bestimmten Transkriptionsfaktorfamilie enthalten, kann die Familie des bindenden Transkriptionsfaktors ermittelt werden. Die eindeutige Identifizierung kann anschließend durch einen spezifischen Antikörper erfolgen, der einen „Supershift“, oder ein „Antikörper-Blocking“ auslöst (Kapitel 3.19). In dieser Arbeit wurden die potentiellen Transkriptionsfaktorbindungsstellen AP1F, HOXF, SP1F, SMAD, AP4R, NEUR und EGRF mit Kernextrakten aus den beiden Zelllinien HaCaT und KLE untersucht. In beiden Zelllinien war die 5'-HERV-L-LTR(Cord) hoch aktiv.

Für die potentielle SP1F-Bindungsstelle konnte in beiden Zelllinien gezeigt werden, dass es sich bei dem hier bindenden Protein tatsächlich um den Transkriptionsfaktor Sp1 handelt. Dies konnte sowohl durch Kompetitionsreaktionen mit SP1F-Konsensus-Oligonukleotiden, als auch mit einem Sp1-spezifischen Antikörper nachgewiesen werden (SP1F bezieht sich auf die Transkriptionsfaktor-, bzw. Matrix-Familie: „Genomatix Matrix Family Library Version 5.0“). Der verwendete Antikörper verursachte keinen „Supershift“, aber ein „Antikörper-Blocking“. Dabei blockiert der an den Transkriptionsfaktor bindende Antikörper dessen DNA-Bindungsstelle, wodurch sich kein Protein-DNA-Komplex bilden kann. Der Transkriptionsfaktor Sp1

wird ubiquitär im Genom von Vertebraten exprimiert. Potentielle SP1F-Bindungsstellen (SP1F-Matrize, Genomatix) konnten in ca. 63% der bisher untersuchten Vertebraten-Promotoren nachgewiesen werden (Genomatix, „Transcription Factor Library“). Sp1 bindet z.B. auch an drei HERV-H-LTRs, die in Retinol stimulierten NTera2-D1-Zellen, in der Plazenta-Choriocarcinoma-Zelllinie JEG-3 und in HeLa-Zellen Promotoraktivität zeigen (Sjottem *et al.*, 1993). Für die Promotoraktivität des Elements HERV-E.PTN, ein in das Pleiotrophin-Gen (PTN) integriertes HERV-E-Element, das die Expression eines HERV-PTN-Fusionstranskripts steuert und vor allem in malignen Trophoblasten exprimiert wird, ist Sp1 ebenfalls als essentieller Transkriptionsfaktor nachgewiesen worden (Schulte, 2000). Außerdem ist Sp1 im Promotor der ERV-L-LTR(Gal) aktiv, der als alternativer Promotor für die Expression des  $\beta 3gal-T5$ -Gens fungiert. Eine synergistische Wirkung von Sp1 mit AP1 wurde z.B. im Promotor des CD11c-Integrins gezeigt (Shelley, 2002).

Die Untersuchung der potentiellen HOXF-Bindungsstelle mit HaCaT- und KLE-Kernextrakten zeigte, dass es sich bei dem hier bindenden Protein tatsächlich um einen Transkriptionsfaktor der HOXF-Familie handelt. Es wurden in diesem Fall lediglich Kompetitionsreaktionen durchgeführt und auf weitere Untersuchungen mit Antikörpern verzichtet, da die Familie HOXF aus 12 verschiedenen Transkriptionsfaktoren besteht. Transkriptionsfaktoren der HOX-Familie (HOXF) sind in vielen Promotoren von Vertebraten essentiell. Potentielle HOXF-Bindungsstellen (HOXF-Matrix, Genomatix) sind in ca. 80% der bisher untersuchten Vertebraten-Promotoren vorhanden (Genomatix). Der Faktor Crx (HOX-Familie) konnte z.B. im Promotor von mehreren Photorezeptor-spezifischen Genen, wie den Opsin-Genen, identifiziert werden (Eurukawa *et al.*, 1997). Die Faktoren Pitx1 und Pitx2 sind in die Aktivierung des proximalen humanen Prolactin-Promotors (hPRL-164luc) involviert (Quentien *et al.*, 2002). Ptx1 ist zusammen mit den Transkriptionsfaktoren Egr-1 und SF-1 in die Expression des LHbeta Gonadotropin Gens involviert (Tremblay und Drouin, 1999).

Die Untersuchung der potentiellen EGRF-Bindungsstelle lieferte mit dem KLE-Kernextrakt einen Egr-1-DNA-Komplex, was durch Konsensus-EGRF-Kompetitoren und Egr-1-Antikörper bestätigt werden konnte. Bei

Verwendung des HaCaT-Kernextraktes konnte der Komplex nur sehr schwach nachgewiesen werden. Vermutlich ist die Expression dieses Transkriptionsfaktors in HaCaT-Zellen im Vergleich zu KLE-Zellen deutlich geringer. Potentielle EGRF-Transkriptionsfaktorbindungsstellen (EGRF-Matrix; Genomatix) sind in ca. 60% der bisher untersuchten Vertebraten-Promotoren vorhanden (Genomatix). Egr-1 ist z.B. zusammen mit Sp1 an der Regulation des „Interleukin-2 IL-2 receptor beta chain“ (IL-2Rbeta) Promotors beteiligt (Lin und Leonard, 1997). Ein weiteres Beispiel ist die Hochregulierung des an der Alzheimererkrankung beteiligten Gens Presenilin-2 (*PSEN2*) in neuronalen Zellen durch Egr-1 (Renbaum *et al.*, 2003).

Die potentielle AP1F-Bindungsstelle zeigt unter Verwendung des HaCaT-Kernextraktes einen typischen AP1-„Shift“. AP1 besteht aus den Monomeren c-Jun und c-Fos, welche die Tendenz haben, verschieden kombinierte Oligomere zu bilden, wodurch Protein-DNA-Komplexe mit unterschiedlichem Molekulargewicht entstehen. Die Folge davon ist, dass kein einzelner „Shift“ sondern ein „Schmier“ entsteht. Die AP1-Bindung konnte auch durch die Verwendung eines c-Jun-Antikörpers bestätigt werden, der ein „Antikörper Blocking“ auslöste. Unter Verwendung des KLE-Kernextraktes konnte kein AP1 nachgewiesen werden, obwohl auch in den KLE-Zellen die Promotoraktivität durch das Inaktivieren der potentiellen AP1F-Bindungsstelle deutlich verringert war. Möglicherweise lag hier ein technisches Problem vor, es könnte bei der KLE-Kernextraktion Probleme mit der Isolierung von AP1 gegeben haben, wodurch das Protein gar nicht, oder nur in sehr geringer Konzentration, zur Verfügung stand.

Potentielle AP1F-Transkriptionsfaktorbindungsstellen (AP1F-Matrix, Genomatix) sind in ca. 34% der bisher untersuchten Vertebraten-Promotoren vorhanden (Genomatix). AP1 ist z.B. zusammen mit AP2 und p53 an der Regulation des HERV-K-LTR-Promotors beteiligt, der die Expression des „carbonic anhydrase family“ (CA) Genes *MN/CA9* steuert. Ein Beispiel für die kooperative Interaktion von AP1 mit Sp1 zeigt sich im Leukozyten-Integrin-Gen (*CD11c*) in der Myeloid-Zelllinie HL60. Hier sind beide Faktoren an der myeloidzellspezifischen Expression von *CD11c* beteiligt (Noti *et al.*, 1996).

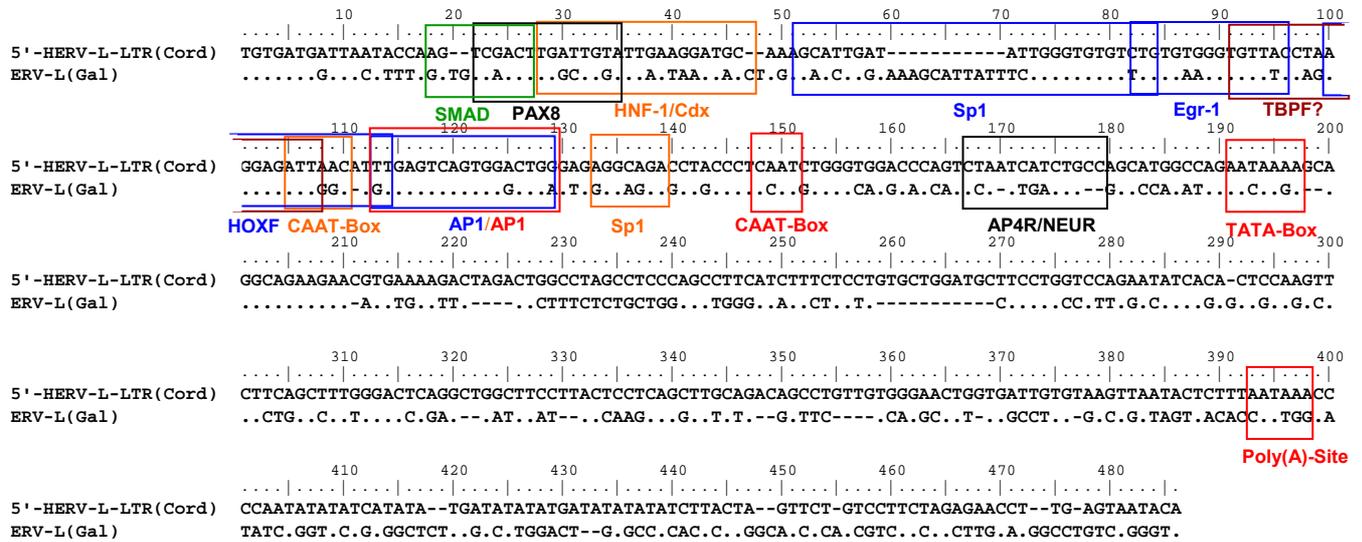
Zusammenfassend ist zu sagen, dass die 5'-HERV-L-LTR(Cord) funktionelle Bindungsstellen für AP1, HOXF, Egr-1 und Sp1 besitzt. Außerdem sind sowohl in der Keratinozytenzelllinie HaCaT als auch in der Uteruszelllinie KLE die durch Computeranalyse identifizierte und potentielle Bindungsstelle für AP4R/NEUR und eine potentielle PAX8-Bindungsstelle für die Promotoraktivität von entscheidender Bedeutung (durch Mutagenese-Untersuchung bestätigt). Die starke Promotoraktivität der 5'-HERV-L-LTR(Cord) in der Zelllinie HaCaT hängt außerdem von einer potentiellen SMAD-Bindungsstelle ab. In der Uteruszelllinie KLE ist ein in der Computeranalyse als TBPF-Bindungsstelle (TATA-Box) identifiziertes Sequenzmotiv für die Promotoraktivität der 5'-HERV-L-LTR(Cord) von entscheidender Bedeutung. Dass es sich dabei tatsächlich um eine TATA-Box handelt, ist allerdings auf Grund der Position innerhalb der LTR (Tabelle 8) und der Position der meisten anderen identifizierten Transkriptionsfaktorbindungsstellen, die sich „downstream“ befinden, unwahrscheinlich. Vermutlich bindet hier ein anderer Transkriptionsfaktor, der von der Computeranalyse nicht erfasst werden konnte (Abb. 43). In den proviralen HERV-L-LTRs P 4 und P 7, die in der Uteruszelllinie KLE aktiv sind, konnten per Computeranalyse (Genomatix) ebenfalls die potentiellen Transkriptionsfaktorbindungsstellen Sp1, HOXF, AP4R und in der LTR P 4 zusätzlich eine AP1-Bindungsstelle in fast gleicher Position wie in der 5'-HERV-L-LTR(Cord) nachgewiesen werden. Bei den Bindungsstellen für Sp1, HOXF und AP1 sind auch die Abstände konserviert. Da die LTRs P 4 und P 7 zwar in der Uteruszelllinie KLE, nicht aber in der Keratinozytenzelllinie HaCaT Promotoraktivität zeigen, lässt sich darauf schließen, dass die Transkriptionsfaktorbindungsstellen Sp1 und HOXF, sowie die potentielle AP4R-Bindungsstelle für die spezifische Aktivität in der Zelllinie KLE essentiell sind. Die zusätzlich in der 5'-HERV-L-LTR(Cord) vorhandenen Bindungsstellen AP1 und Egr-1 erhöhen die Promotoraktivität, eine außerdem vorhandene potentielle SMAD-Bindungsstelle ist für die Promotoraktivität in der Zelllinie HaCaT entscheidend.

Die im Rahmen dieser Arbeit identifizierte AP1-Bindungsstelle der 5'-HERV-L-LTR(Cord) stimmt mit der von Cordonnier *et al.* per Computeranalyse ermittelten AP1-Bindungsstelle überein. Die von Cordonnier *et al.* postulierte TATA-Box konnte

per Computeranalyse („Matinspector“) bestätigt werden, nicht aber die CAAT-Box (Abb. 43).

Vergleiche der 5'-HERV-L-LTR(Cord) mit der von Dunn *et al.* identifizierten, vor allem im Kolon von Primaten aktiven ERV-L-LTR(Gal), die als alternativer Promotor für das  $\beta$ 3gal-T5-Gen fungiert, sind in Abbildung 43 dargestellt. In diesem LTR-Promotor sind es die Transkriptionsfaktoren HNF-1/Cdx, Sp1 und die CAAT-Box, die an der kolonspezifischen Promotoraktivität beteiligt sind. Hier ist die einzige Parallele zur 5'-HERV-L-LTR(Cord) eine Sp1-Bindungsstelle. Diese befindet sich in der ERV-L-LTR(Gal) allerdings an einer anderen Position. Generell ist zu sagen, dass sich beide Sequenzen durch viele Punktmutationen unterscheiden, die auch in allen Transkriptionsfaktorbindungsstellen zum Tragen kommen. Diese Unterschiede sind allerdings nicht verwunderlich, besitzen diese beiden LTRs doch völlig unterschiedliche Promotorspezifitäten und sind damit beide gute Beispiele für die Vielfältigkeit von HERV-L-LTR-Promotoren.

## 5. Diskussion



**Abb. 43:** Sequenzvergleich der 5'-HERV-L-LTR(Cord) und der ERV-L-LTR(Gal). Dargestellt sind die Transkriptionsfaktorbindungsstellen der 5'-HERV-L-LTR(Cord) und der ERV-L-LTR(Gal). Von der ERV-L(Gal) sind nur die von der 5'-HERV-L-LTR(Cord) abweichenden Basen dargestellt. **Blau:** Im Rahmen der vorliegenden Arbeit durch EMSAs identifizierte Transkriptionsfaktorbindungsstellen in der 5'-HERV-L-LTR(Cord). **Schwarz:** Potentielle Transkriptionsfaktorbindungsstellen deren Inaktivierung durch Mutagenese sowohl in der Keratinozytenzelllinie HaCaT als auch in der Uteruszelllinie KLE zu deutlicher Verringerung der Promotoraktivität führt. **Grün:** Potentielle Transkriptionsfaktorbindungsstelle, deren Inaktivierung nur in der Zelllinie HaCaT zu starker Verringerung der Promotoraktivität führt. **Braun:** Potentielle Transkriptionsfaktorbindungsstelle, deren Inaktivierung nur in der Zelllinie KLE zu starker Verringerung der Promotoraktivität führt. **Rot:** Von Cordonnier *et al.* per Computeranalyse ermittelte potentielle Transkriptionsfaktorbindungsstellen der 5'-HERV-L-LTR(Cord). **Orange:** Durch Dunn *et al.* experimentell identifizierte Transkriptionsfaktorbindungsstellen der ERV-L-LTR(Gal).

### **Ausblick**

Es konnte gezeigt werden, dass HERV-L-LTRs in verschiedenen Zelltypen unterschiedliche Promotoraktivität besitzen. Außerdem konnte ein Teil der Promotorregulation der 5'-HERV-L-LTR(Cord) aufgeklärt werden. Die 5'-HERV-L-LTR(Cord) ist die bisher aktivste LTR aus der HERV-L-Familie, aber auch die proviralen 5'-LTRs von P 4 und P 7 sind aktiv und zeigen die bisher stärkste Spezialisierung (Uteruszelllinie KLE). Da bereits gezeigt werden konnte, dass provirale HERV-L-Elemente in den humanen Geweben Haut, Herz, Uterus, Zervix, Plazenta, Hoden und Schilddrüse (Seifarth *et al.*, 2005) aktiv sind, sollte in den zukünftigen Arbeiten mit HERV-L versucht werden, LTRs zu identifizieren, die in den Geweben aktiv sind, für die bisher noch kein entsprechendes HERV-L-Element gefunden wurde, also vor allem in Schilddrüse und Hoden. Diese sollten unter den ca. 200 im menschlichen Genom vorhandenen, proviralen HERV-L-Elementen zu finden sein. Daher sollten in Zukunft weitere, vor allem provirale, 5'-HERV-L-LTRs kloniert, und deren Promotoraktivität zunächst in von Schilddrüsen- und Hodengewebe stammenden Zelllinien getestet werden.

Retrovirale Vektoren werden aufgrund ihrer Fähigkeit stabil ins Genom zu integrieren in der Gentherapie verwendet, um defekte Gene zu ersetzen oder therapeutische Gene einzubauen. Einer der Nachteile von retroviralen Vektoren ist, dass sie bisher ubiquitär aktive Promotoren enthielten (z.B. MLV-LTRs). Dabei kann theoretisch in allen Geweben, in denen es zu einer retroviralen Integration kommt, Expression der therapeutischen Gene erfolgen, was zu unerwünschten Nebeneffekten führen kann. Durch Einbau von gewebespezifisch aktiven HERV-LTRs als Promotoren in retrovirale Vektoren könnten, je nach HERV-LTR-Typ, therapeutische Gene gezielt in bestimmten Geweben exprimiert werden. HERV-LTRs stellen auf jeden Fall ein großes Potential an gewebespezifisch aktiven Promotoren dar, die, da es sich um geschlossene Regulationssysteme handelt, nicht wie zelluläre Promotoren auf räumlich getrennte „Enhancer“ angewiesen sind.

### 6 Zusammenfassung

Das menschliche Genom besteht zu etwa 9% aus Sequenzen retroviralen Ursprungs (humane endogene Retroviren, HERVs). Ein Großteil dieser Sequenzen wird noch immer transkribiert. Mitglieder der Familie HERV-L, die aus ca. 200 proviralen Elementen und 6000 solitären LTRs besteht, sind vor allem in Haut, Uterus, Zervix, Plazenta, Schilddrüse, Herz und Hoden aktiv (Seifarth *et al.*, 2005). Durch diese gewebespezifische Promotoraktivität könnten HERV-L-LTRs als Promotoren zur Steuerung therapeutischer Gene in retroviralen Vektoren Verwendung finden. Im Rahmen dieser Arbeit wurden verschiedene provirale und solitäre HERV-L-LTRs auf ihre zelltypspezifische Aktivität untersucht, und einige der daran beteiligten zellulären Transkriptionsfaktoren identifiziert. Es konnte gezeigt werden, dass die provirale 5'-HERV-L-LTR(Cord) sowohl in der Keratinozytenzelllinie HaCaT als auch in der Uteruszelllinie KLE hoch aktiv ist. Um weitere HERV-L-Elemente zu identifizieren, die zu den gewebespezifischen Aktivitäten beitragen, wurden verschiedene HERV-L-LTRs aus humaner genomischer DNA amplifiziert und kloniert. Es wurden sowohl solitäre (S 1, S 4, S 5, S 11), als auch 5'-provirale HERV-L-LTRs (P 3, P 4, P 7, P 10) hinsichtlich ihrer Promotoraktivität untersucht. Dabei konnten zwei aktive provirale LTRs gefunden werden, die in der Uteruszelllinie KLE starke Promotoraktivität zeigten (P 4 und P 7).

Um herauszufinden, welche Transkriptionsfaktoren an der zelltypspezifischen Aktivität der 5'-HERV-L-LTR(Cord) beteiligt sind, wurde diese mit der inaktiven 3'-HERV-L-LTR(Cord) sowohl auf Sequenzebene per Alignment, als auch bezüglich ihrer potentiellen Transkriptionsfaktorbindungsstellen, verglichen (Software: „MatInspector“). Es wurden ausschließlich in der aktiven 5'-HERV-L-LTR(Cord) im U3/R-Bereich vorhandene potentielle Bindungsstellen ausgewählt, und deren Funktionalität mittels Mutagenese-Untersuchungen getestet. Dabei konnte gezeigt werden, dass das Inaktivieren der potentiellen Bindungsstellen AP1F, PAX8, HOXF, RCAT/SP1F und AP4R/NEUR in der Zelllinie HaCaT und in der Zelllinie KLE zu einer deutlichen Verringerung der 5'-HERV-L-LTR(Cord)-Promotoraktivität führt. Inaktivierung der SMAD-Transkriptionsfaktorbindungsstelle führte ausschließlich in

HaCaT-Zellen zu deutlichem Promotoraktivitätsverlust. Bei weiteren Analysen mittels „Electrophoretic Mobility Shift Assay“ (EMSA) konnten die Transkriptionsfaktoren AP1, Egr-1, HOXF und Sp1 identifiziert werden. Die Transkriptionsfaktoren AP1, Egr-1 und Sp1 wurden zusätzlich durch spezifische Antikörper bestätigt.

Vergleiche der untersuchten aktiven HERV-L-LTRs (5'-HERV-L-LTR(Cord), P 4 und P 7) mittels Computeranalyse zeigten, dass die Transkriptionsfaktorbindungsstellen Sp1, HOXF und AP4R in allen aktiven HERV-L-LTRs konserviert sind. Im Element P 4 und in der 5'-HERV-L-LTR(Cord) ist außerdem eine AP1-Bindungsstelle konserviert.

Die Ergebnisse der Promotoruntersuchungen lassen darauf schließen, dass die Transkriptionsfaktoren Sp1, HOXF (durch EMSA bestätigt) und eine potentielle AP4R-Transkriptionsfaktorbindungsstelle (durch Mutagenese bestätigt) für die Promotoraktivität in der Uteruszelllinie KLE essentiell sind. Für die Promotoraktivität in der Keratinozytenzelllinie HaCaT sind außerdem die Transkriptionsfaktoren AP1 und Egr-1 (durch EMSA bestätigt) und eine potentielle SMAD-Transkriptionsfaktorbindungsstelle (durch Mutagenese bestätigt) von Bedeutung.

Insgesamt konnte mit diesen Untersuchungen gezeigt werden, dass LTRs der HERV-L-Familie besonders klar definierte Zelltypspezifität aufweisen und sich damit gut zur Entwicklung gewebespezifischer retroviraler Vektoren eignen.

## Danksagung

Ich möchte mich hiermit bei folgenden Personen bedanken, die zu dieser Dissertation beigetragen haben:

- bei meiner Betreuerin Professor Dr. Christine Leib-Mösch für Ihre andauernde Unterstützung und für fruchtbare Diskussionen.
- Professor Dr. Wolfgang Wurst für die Freundlichkeit meine Arbeit zu begutachten.
- Professor Dr. Volker Erfle, für die Möglichkeit am Institut für Molekulare Virologie der GSF diese Doktorarbeit durchzuführen.
- Dr. Peter J. Nelson, der erheblich zum Gelingen der EMSA-Experimente beigetragen hat.
- Professor Dr. Alex Greenwood, Eva Gebefügi, Laura Leitner, Anna Stengel, Christian Bach, René Leubert und allen anderen Kollegen am IMV, die mir mit Rat und Tat zu Seite standen und mit denen ich in den letzten Jahren viel Spaß hatte.
- ganz besonders möchte ich mich bei meinen Eltern bedanken, die mich in jeglicher Hinsicht unterstützt haben.

## **Eidesstattliche Erklärung**

Hiermit versichere ich, der Unterzeichnende, dass ich die vorliegende Arbeit selbständig und nur mit den angegebenen Quellen und Hilfsmitteln angefertigt habe. Alle Ausführungen, die wörtlich oder sinngemäß übernommen wurden, sind als solche gekennzeichnet.

München, Dezember 2005

Stephan Weinhardt

## 7 Literaturverzeichnis

- Abrink M, Larsson E, Hellman L (1998). Demethylation of ERV3, an endogenous retrovirus regulating the Kruppel-related zinc finger gene H-plk, in several human cell lines arrested during early monocyte development. *DNA Cell Biol.* **17**: 27-37
- Andersson ML, Lindeskog M, Medstrand P, Westley B, May F, Blomberg J (1999). Diversity of human endogenous retrovirus class II-like sequences. *J Gen Virol.* **80**:255-260
- Andersson ML, Sjøttem E, Svineng G, and Johansen T (1997). Comparative analysis of LTRs of the ERV-H family of primate-specific retrovirus-like elements isolated from marmoset, African green monkey, and man. *Virology.* **234**:14-30
- Baust C, Seifarth W, Schon U, Hehlmann R, Leib-Moesch (2001). Functional activity of HERV-K-T47D-related long terminal repeats. *Virology.* **283**:262-272
- Baust CW, Seifarth H, Germaier R, Hehlmann R, Leib-Moesch C (2000). HERV-K-T47D-Related long terminal repeats mediate polyadenylation of cellular transcripts. *Genomics.* **66**:98-103
- Barbulescu M, Turner G, Su M, Kim R, Jensen-Seaman MI, Deinard AS, Kidd KK, Lenz J (2001). A HERV-K provirus in chimpanzees, bonobos and gorillas, but not humans. *Curr Biol.* **11**:779-83.
- Beasley BE and Hu WS (2002). Cis-acting elements important for retroviral RHA Packaging specificity. *J Virol.* **76**:4950-4960
- Belshaw R, Katzouvrakis A, Paces J, Burt A, and Tristem M (2005). High copy number in human endogenous retrovirus families is associated with copying mechanisms in addition to recognition. *Mol. Biol. Evol.* **22**: 814-817

- Belshaw R, Pereira V, Katzourakis A, Talbot G, Paces J, Burt A, Tristem (2004). Long-term reinfection of the human genome by endogenous retroviruses. *Proc Natl Acad Sci U S A*. **101**:4894-9
- Benit L, Dessen P, and Heidmann T (2001). Identification, phylogeny, and evolution of retroviral elements based on their envelope genes. *J Virol* **75**:11709-11719
- Benit L, Lallemand JB, Casella JF, Philippe H, Heidmann T (1999). ERV-L elements: A family of endogenous retrovirus-like elements active throughout the evolution of mammals. *J Virol*. **73**:3301-3308
- Benit L, De Parseval N, Casella JF, Callebaut I, Cordonnier A, Heidmann T (1997). Cloning of a new murine endogenous retrovirus, MuERV-L, with strong similarity to the human HERV-L element and with a *gag* coding sequence closely related to the *Fv1* restriction gene. *J. Virol*. **71**: 5652-7
- Best S, Le Tissier P, Towers G, Stoye JP (1996). Positional cloning of the mouse retrovirus restriction gene *Fv1*. *Nature*. **382**:762-3
- Bi S, Gavrilova O, Gong DW (1997). Identification of placental enhancer for the human leptin gene. *J Biol Chem*. **272**:30583-30588
- Bieche I, Laurent A, Laurendeau I, Duret L, Giovangrandi Y, Frenco JC, Olivi M, Fausser JC, Evain-Brion D, Vidaud M (2003). Placenta-specific *INSL4* expression is mediated by a human endogenous retrovirus element. *Biol Reprod*. **68**:1422-1429
- Blaise S, de Parseval N, Benit L, Heidmann T (2003). Genomewide screening for fusogenic human endogenous retrovirus envelopes identifies syncytin 2, a gene conserved on primate evolution. *Proc Natl Acad Sci U S A*. **100**:13013-8
- Bock M, and Stoye JP (2000). Endogenous retroviruses and the human germline. *Curr Opin Genet Dev*. **10**:651-5
- Boeke JD (2003). The unusual phylogenetic distribution of retrotransposons: A hypothesis. *Genome Research* **13**: 1975-1983

- Boeke JD, Stoye JP (1997). Retrotransposons, endogenous retroviruses, and the evolution of retroelements. In *Retroviruses*. Edited by Coffin JM, Hughes SH and Varmus HE. Cold Spring Harbor; *Cold Spring Harbor Laboratory Press* 343-435
- Brodin G, Ahgren A, ten Dijke P, Heldin CH, Heuchel R (2000). Efficient TGF-beta induction of the Smad7 gene requires cooperation between AP-1, Sp1, and Smad proteins on the mouse Smad7 promoter. *J Biol Chem.* **37**:29023-30.
- Blaise S, de Parseval N, Benit L, Heidmann T (2003). Genomewide screening for fusogenic human endogenous retrovirus envelopes identifies syncytin 2, a gene conserved on primate evolution. *Proc Natl Acad Sci U S A.* **100**:13013-8.
- Blond JL, Lavillette D, Cheynet V, Bouton O, Oriol G, Chapel-Fernandes S, Mandrand B, Mallet F, Cosset FL (2000). An envelope glycoprotein of the human endogenous retrovirus HERV-W is expressed in the human placenta and fuses cells expressing the type D mammalian retrovirus receptor. *J Virol.* **74**:3321-9.
- Blond JL, Beseme F, Duret L, Bouton O, Bedin F, Perron H, Mandrand B and Mallet F(1999). Molecular characterization and placental expression of HERV-W, an new human retrovirus family. *J Virol.* **73**:1175-1185
- Browning MT, Schmidt RD, Lew KA, Rizvi TA (2001). Primate and feline lentivirus vector RNA packaging and propagation by heterologous lentivirus virions. *J Virol* **75**: 5129-5140
- Bowerman B, Brown PO, Bishop JM, Varmus HE (1989). A nucleoprotein complex mediates the integration of retroviral DNA. *Genes Dev.* **4**:469-78
- Cartharius K, Frech K, Grote K, Klocke B, Haltmeier M, Klingenhoff A, Frisch M, Bayerlein M, Werner T (Epub 2005). MatInspector and beyond: promoter analysis based on transcription factor binding sites. *Bioinformatics.* **21**:2933-42.

- Casau AE, Vaughan JE, Lozano G, Levine AJ (1999). Germ cell expression of an isolated human endogenous retroviral long terminal repeat of the HERV-K/HTDV family in transgenic mice. *J Virol.* **73**:9976-9983
- Certo JL, Kabdulov TO, Paulson ML, Anderson JA and Hu WS (1999). The nucleocapsid is responsible for the ability of spleen necrosis virus (SNV) gag polyprotein to package both SNV and murine leukaemia virus RNA. *J Virol* **73**:9170-9177
- Chen H, Carr K, Jerome RE, Edenberg HJ (2002). A retroviral repetitive element confers tissue-specificity to the human alcohol dehydrogenase 1C (*ADH1C*) gene. *DNA AND Cell Biology.* **11**:793-801.
- Coffin JM, Hughes HH, Varmus HE (1996). *Retroviruses.* Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, USA
- Cohen M, Kato N, Larsson E (1988). ERV3 human endogenous provirus mRNAs are expressed in normal and malignant tissues and cells, but not in choriocarcinoma tumor cells. *J Cell Biochem.* **36**:121-128
- Consortium I.h.g.s, initial sequencing and analysis of the human genome. *Nature.* 409:860-921
- Cordonnier A, Casella J-P, and Heidmann T (1995). Isolation of novel human endogenous retroviral-like elements with foamy virus-related *pol* sequence. *J Virol* **69**:5890-5897
- Crittenden LB, Fadly AM, Smith EJ (1982). Effect of endogenous leukosis virus genes on response to infection with avian leukosis and reticuloendotheliosis viruses. *Avian Dis* **26**: 279-295
- Crittenden LB, Smith EJ, Fadly AM (1984). Influence of endogenous viral (ev) gene expression and strain mortality of exogenous avian leukosis virus (ALV) on mortality and ALV infection and shedding in chicken. *Avian Dis.* **28**:1037-1056
- Slagel VK, Deininger PL (1989). In vivo transcription of a cloned prosimian primate SINE sequence. *Nucleic Acids Res.***17**:8669-82.

- De Parseval N, Alkabbani H, Heidmann T (1999). The long terminal repeats of the HERV-H human endogenous retrovirus contain binding sites for transcriptional regulation by the Myb protein. *J Gen Virol.* **80**:841-845
- Deininger, P. L (1989). SINEs: Short interspersed repeated DNA elements in higher eucaryotes. In D. E. Berg & M. M. Howe (Eds.), *Mobile DNA*. American Society for Microbiology.
- Di Cristofano A, Strazullo M, Longo L. (1995). Characterization and genomic mapping of the ZNF80 locus: Expression of this zinc-finger gene is driven by a solitary LTR of ERV9 endogenous retroviral family. *Nucleic Acids Res.* **23**:2823-2830
- Di Cristofano A, Strazullo M, Parisi T, La Mantia G (1995). Mobilization of an ERV9 human endogenous retroviral element during primate evolution. *Virology.* 213:271-275
- Domansky AN, Kopantzev EP, Snezhkov EV, Lebedev YB, Leib-Moesch C, Sverdlov ED (2000). Solitary HERV-K LTRs possess bi-directional promoter activity and contain a negative regulatory element in the U5 region. *FEBS Lett.* **472**:191-195
- Doolittle RF, Feng DF, Johnson MS and McClure MA (1989). Origins and evolutionary relationships of retroviruses. *Q Rev Biol.* **64**:1-30
- Dunn CA, van de Lagemaat CN, Baillie GJ, Mager DL (2005). Endogenous retrovirus long terminal repeats as ready-to-use mobile promoters: The case of primate beta3GAL-T5. *Gene*. [Epub ahead of print]
- Dunn CA, Medstrand P, Mager DL (2003). An endogenous retroviral long terminal repeat is the dominant promoter for human beta1,3-galactosyltransferase 5 in the colon. *Proc Natl Acad Sci U S A.* **100**:12841-6.
- Emi M, Horii A, Tomita A, Nishide T, Ogawa M, Mori T, Matsubara K (1988). Overlapping two genes in human DNA: A salivary amylase gene overlaps with a gamma-actin pseudogene that carries an integrated human endogenous retroviral DNA. *Gene.* **62**:229-235

- Eickbush TJ (1992). Transposing without ends: the non-LTR retrotransposable elements. *New Biol.* **4**:430-40. Review.
- Esnault C, Heidmann O, Delebecque F, Dewannieux M, Ribet D, Hance AJ, Heidmann T, Schwartz O (2005). APOBEC3G cytidine deaminase inhibits retrotransposition of endogenous retroviruses. *Nature*. Vol. 433
- Eurokawa T, Morrow EM, Cepko CL (1997). Crx, a novel otx-like homeobox gene, shows photoreceptor-specific expression and regulates photoreceptor differentiation. *Cell*. **91**:531-41
- Feuchter A, Mager DL. (1992). SV40 large T antigen trans-activates the long terminal repeats of a large family of human endogenous retrovirus-like sequences. *Virology*. 187:242-250
- Feuchter A, Mager D (1990). Functional heterogeneity of a large family of human LTR-Like promoters and enhancers. *Nucleic Acids Res.* **18**:1261-1270
- Feuchter-Murthy AE, Freeman JD, Mager DL (1993). Splicing of a human endogenous retrovirus to a novel phospholipase A2 related gene. *Nucleic Acids Res.* **21**:135-143
- Goodchild NL, Wilkinson DA, Mager DL (1992). A human endogenous long terminal repeat provides a polyadenylation signal to a novel, alternatively spliced transcript in normal placenta. *Gene*. **121**:287-294
- Graham FL, Smiley J, Russel W, Nairn R (1977). Characteristics of a human cell line transformed by DNA from human adenovirus type 5. *J Gen Virology*. **36**:59-74
- Greenwood AD, Stengel A, Erfle V, Seifarth W, Leib-Mösch C (2005). The distribution of pol containing human endogenous retroviruses in non-human primates. *Virology* **334**: 203–213
- Greenwood AD, Englbrecht CC, MacPhee RD (2004). Characterization of an endogenous retrovirus class in elephants and their relatives. *BMC Evol Biol.* **4**:38.

- Greenwood AD, Lee F, Capelli C, DeSalle R, Tikhonov A, Marx PA, MacPhee RD (2001). Evolution of endogenous retrovirus-like elements of the woolly mammoth (*Mammuthus primigenius*) and its relatives. *Mol Biol Evol.* **18**:840-7.
- Hatada S, Grant D, Maeda N (2003). An ironic endogenous retrovirus-like sequence attenuates human haptoglobin-related gene expression in an orientation-dependent manner. *Gene.* **319**:55-63
- Higgins DG and Sharp (1988). CLUSTAL: a package for performing multiple sequence alignment on a microcomputer. *Gene.* **73**:237-244
- Hirose Y, Takamatsu M, and Harada F (1993). Presence of *env* genes in members of the RTL VH Family of human endogenous retrovirus-like elements. *Virology.* **192**: 52-61
- Isshiki S, Kudo T, Nishihara S, Ikehara Y, Togayachi A, Furuya A, Shitara K, Kubota T, Watanabe M, Kitajima M, Narimatsu H (2003). Lewis type 1 antigen synthase (beta3Gal-T5) is transcriptionally regulated by homeoproteins. *J Biol Chem.* **278**:36611-20
- Jern P, Sperber GO, Blomberg J (2004). Definition and variation of human endogenous retrovirus H. *Virology.* **327**: 93-110
- Jolicoeur P (1979). The Fv-1 gene of the mouse and its control of murine leukemia virus replication. *Curr Top Microbiol Immunol.* **86**:67-122
- Jordan I, Rogozin LB, Glazko GV, Koonin EV (2003). Origin of a substantial fraction of human regulatory sequences from transposable elements. *Trends on Genetics.* Vol. 19. No 2
- Kato N, Pfeifer-Ohlsson S, Kato M, Larsson E, Rydnert J, Ohlsson R, Cohen M (1987). Tissue-specific expression of human provirus ERV3 mRNA in human placenta: Two of three ERV3 mRNAs contain human cellular sequences. *J Virol.* **61**:2182-2191
- Kato N, Shimotohno K, Van Leeuwen D (1990). Human proviral mRNAs down regulated in choriocarcinoma encode a zinc finger protein related to Kruppel. *Mol Cell Biol.* **10**:4401-4405

- Kapitonov VV, Jurka J (1999). The long terminal repeat of an endogenous retrovirus induces alternative splicing and encodes an additional carboxy-terminal sequence in the human leptin receptor. *J Mol Evol.* **48**:248-251
- Kazmierczak B, Pohnke Y, Bullerdiek J (1996). Fusion transcripts between the HMGIC gene and RTVL-H-related sequences in mesenchymal tumors without cytogenetic aberrations. *Genomics.* **38**:223-226
- Kigami D, Minami N, Takayama H, Imai H (2003). MuERV-L is one of the earliest transcribed genes in mouse one-cell embryos. *Biol Reprod.* **68**:651-4
- Kim HS, Takenaka O, and Crow TJ (1999). Isolation and phylogeny of endogenous retrovirus sequences belonging to the HERV-W family in primates. *J Gen Virol.* **80**:2613-2619
- Kjellman C, Sjogren HO, Salford LG, and Widegreen B (1999). HERV-F(XA-34) is a full-length human endogenous retrovirus expressed in placental and fetal tissues. *Gene.* **239**:99-107
- Knossel M, Lower R, Lower J (1999). Expression of the human endogenous retrovirus HTDV/HERV-L is enhanced by cellular transcription factor YY1. *J Virol.* **73**:1254-1261
- Komurian-Pradel F, Paranhos-Bacala G, Bedin F, Ounanian-Paraz A, Sodoyer M, OH C, Rajoharison A, Garcia E, Mallet F, Mandrand B, Preean H (1999). Molecular cloning and characterization of MSRV-related sequences associated with retrovirus-like particles. *Virology.* **260**:1-9
- Kowalski PE, Freeman JD, Mager DL (1999). Intergenic splicing between a HERV-H endogenous retrovirus and two adjacent human genes. *Genomics.* **57**:371-379
- Kowalski PE, Freeman JD, Nelson DT, Mager DL (1997). Genomic structure and evolution of a novel gene (PLA2L) with duplicated phospholipase A2-Like domains. *Genomics.* **39**:38-46
- Kowalski PE, Mager DL (1998). A human endogenous retrovirus suppresses translation of an associated fusion transcript, PLA2L. *J Virol.* **72**:6164-6168

- Kupiec JJ and Sonigo P (1996). Reverse transcriptase jumps and gaps. *J Gen Virol.* **Pt9**:1987-91 Review
- Kulski JK, Gaudieri S, Martin A, Dawkins RL (1999). Coevolution of PERB11 (MIC) and HLA class I genes with HERV-16 and retroelements by extended genomic duplication. *J Mol Evol.* **49**:84-97.
- Kung HJ, Boerkoel C, Carter TH (1991), Retroviral mutagenesis of cellular oncogenes: A review with insights into the mechanisms of insertional activation. *Curr Top Microbiol Immunol* **171**: 1-25
- La Mantia G, Majello B, Di Cristofano A, Strazullo M, Minchiotti G, Lania L (1994). Identification of regulatory elements within the minimal promoter region of the human endogenous ERV9 proviruses: Accurate transcription initiation is controlled by an Inr-like element. *Nucleic Acids Res.* **20**:4129-4136
- La Mantia G, Maglione D, Pengue G. (1991). Identification and characterization of novel human endogenous retroviral sequences preferentially expressed in undifferentiated embryonal carcinoma cells. *Nucleic Acids Res.* **19**:1513-1520
- Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, Funke R, Gage D, Harris K, Heaford A, Howland J, Kann L, Lehoczky J, LeVine R, McEwan P, McKernan K, Meldrim J, Mesirov JP, Miranda C, Morris W, Naylor J, Raymond C, Rosetti M, Santos R, Sheridan A, Sougnez C, Stange-Thomann N, Stojanovic N, Subramanian A, Wyman D, Rogers J, Sulston J, Ainscough R, Beck S, Bentley D, Burton J, Clee C, Carter N, Coulson A, Deadman R, Deloukas P, Dunham A, Dunham I, Durbin R, French L, Grafham D, Gregory S, Hubbard T, Humphray S, Hunt A, Jones M, Lloyd C, McMurray A, Matthews L, Mercer S, Milne S, Mullikin JC, Mungall A, Plumb R, Ross M, Shownkeen R, Sims S, Waterston RH, Wilson RK, Hillier LW, McPherson JD, Marra MA, Mardis ER, Fulton LA, Chinwalla AT, Pepin KH, Gish WR, Chissoe SL, Wendl MC, Delehaunty KD, Miner TL, Delehaunty A, Kramer JB, Cook LL, Fulton RS, Johnson DL, Minx PJ, Clifton SW, Hawkins T, Branscomb E, Predki P, Richardson P, Wenning S, Slezak T, Doggett N, Cheng JF, Olsen A, Lucas S, Elkin C, Uberbacher E, Frazier M,

Gibbs RA, Muzny DM, Scherer SE, Bouck JB, Sodergren EJ, Worley KC, Rives CM, Gorrell JH, Metzker ML, Naylor SL, Kucherlapati RS, Nelson DL, Weinstock GM, Sakaki Y, Fujiyama A, Hattori M, Yada T, Toyoda A, Itoh T, Kawagoe C, Watanabe H, Totoki Y, Taylor T, Weissenbach J, Heilig R, Saurin W, Artiguenave F, Brottier P, Bruls T, Pelletier E, Robert C, Wincker P, Smith DR, Doucette-Stamm L, Rubenfield M, Weinstock K, Lee HM, Dubois J, Rosenthal A, Platzer M, Nyakatura G, Taudien S, Rump A, Yang H, Yu J, Wang J, Huang G, Gu J, Hood L, Rowen L, Madan A, Qin S, Davis RW, Federspiel NA, Abola AP, Proctor MJ, Myers RM, Schmutz J, Dickson M, Grimwood J, Cox DR, Olson MV, Kaul R, Raymond C, Shimizu N, Kawasaki K, Minoshima S, Evans GA, Athanasiou M, Schultz R, Roe BA, Chen F, Pan H, Ramser J, Lehrach H, Reinhardt R, McCombie WR, de la Bastide M, Dedhia N, Blocker H, Hornischer K, Nordsiek G, Agarwala R, Aravind L, Bailey JA, Bateman A, Batzoglou S, Birney E, Bork P, Brown DG, Burge CB, Cerutti L, Chen HC, Church D, Clamp M, Copley RR, Doerks T, Eddy SR, Eichler EE, Furey TS, Galagan J, Gilbert JG, Harmon C, Hayashizaki Y, Haussler D, Hermjakob H, Hokamp K, Jang W, Johnson LS, Jones TA, Kasif S, Kasprzyk A, Kennedy S, Kent WJ, Kitts P, Koonin EV, Korf I, Kulp D, Lancet D, Lowe TM, McLysaght A, Mikkelsen T, Moran JV, Mulder N, Pollara VJ, Ponting CP, Schuler G, Schultz J, Slater G, Smit AF, Stupka E, Szustakowski J, Thierry-Mieg D, Thierry-Mieg J, Wagner L, Wallis J, Wheeler R, Williams A, Wolf YI, Wolfe KH, Yang SP, Yeh RF, Collins F, Guyer MS, Peterson J, Felsenfeld A, Wetterstrand KA, Patrinos A, Morgan MJ, de Jong P, Catanese JJ, Osoegawa K, Shizuya H, Choi S, Chen YJ; International Human Genome Sequencing Consortium. (2001). Initial sequencing and analysis of the human genome. *Nature* **409**: 860-921

- Landry JR, Mager DL (2003). Functional analysis of the endogenous retroviral promoter of the human endothelin B receptor gene. *J Virol.* **13**:7459-66.
- Landry JR, Rouhi A, Medstrand P, Mager DL (2002). The opitz syndrome gene *mid1* is transcribed from a human endogenous retroviral promoter. *Mol Biol Evol.* **19**:1934-1942

- Landry JR, Mager DL (2002). Widely spaced alternative promoters, conserved between human and rodent, control expression of the Opitz syndrome gene MID1. *Genomics*. **80**:499-508
- Larsson E and Andersson G (1998). Beneficial role of human endogenous retroviruses: facts and hypotheses. *Scand J Immunol*. **4**:329-38. Review.
- Larsson E (1993). Expression of an endogenous retrovirus, HERV-R in human tissues. *J Cancer Res Clin Oncol*. **119**:S6
- Lee WJ, Kwun HJ, Kim HS, Jang KL (2003). Activation of the human endogenous retrovirus W long terminal repeat by herpes simplex virus type 1 immediate early protein 1. *Mol Cells*. **15**:75-80.
- Leib-Mösch C, Seifarth W, Schön Ulrike (2005). Influence of human endogenous retroviruses on cellular gene expression. *Retroviruses and Primate Evolution* edited by Sverdlov ED
- Li WH, Gu Z, Wang H, Nekrutenko A (2001). Evolutionary analysis of the human genome. *Nature* **409**: 847-849
- Li L, Xu B, Rote NS (2000). The cellular mechanism by which the human endogenous retrovirus ERV-3 env gene affects proliferation and differentiation in a human placental trophoblast model, BeWo. *Placenta*.**21**:73-8
- Lindeskog M, Mager DL, and Blomberg J (1999). Isolation of a human endogenous retroviral HERV-H element with an open *env* reading frame. *Virology*. **258**:441-450
- Lin J. and Leonard W.J. (1997). The immediate-early gene product Egr-1 regulates the human interleukin-2 receptor beta-chain promoter through noncanonical Egr and Sp1 binding sites. *Mol Cell Biol*. **7**:3714-22
- Lin JX, Leonard WJ (1997). The immediate-early gene product Egr-1 regulates the human interleukin-2 receptor beta-chain promoter through noncanonical Egr and Sp1 binding sites. *Mol Cell Biol*. **17**:3714-22
- Lindeskog M, Medstrand P, Cunningham AA, and Blomberg J (1998). Coamplification and dispersion of adjacent human endogenous retroviral

HERV-H and HERV-E elements: presence of spliced hybrid transcripts in normal leukocytes. *Virology*. **244**(1):219-29

- Liu AY, Abraham BA (1991). Subtractive cloning of a hybrid human endogenous retrovirus and calbindin gene in the prostate cell line PC3. *Cancer Res*. **51**:4107-4110
- Long Q, Bengra C, Li F, Kutlar F, Tuan D (1998). A long terminal repeat of the human endogenous retrovirus ERV-9 is located in the 5' boundary area of the human beta-globin locus control region. *Genomics*. **54**:542-555
- Mager DL, Henthorn PS (2003). Identification of retrovirus-like repetitive element in human DNA. *Proc Natl Acad Sci USA*. **81**:7510-4
- Mager DL, Medstrand P (2001): Retroviral repeat sequences. In: Cooper D, ed. Nature encyclopedia of the human genome. *Nature*. **409**:847-849
- Mager DL, Hunter DG, Schertzer M, Freeman JD (1999). Endogenous retroviruses provide the primary polyadenylation signal for two new genes (HHLA2 and HHLA3). *Genomics*. **59**:255-263
- Mager DL, Freeman JD (1995). HERV-H endogenous retroviruses: Presence in the New World branch but amplification in the Old World primate lineage. *Virology*. **213**:395-404
- Mager DL (1989). Polyadenylation function and sequence variability of the long terminal repeats of the human endogenous retrovirus-like family RTVL-H. *Virology*. **173**:591-599
- Mager DL, Henthorn PS (1984): Identification of a retrovirus-like repetitive element in human DNA. *Proc Natl Acad Sci USA*. **81**:7510-4
- Mang R, Goudsmit J, and Van der Kuyl AC (1999). Novel endogenous type C retrovirus in baboons: complete sequence, providing evidence for baboon endogenous virus gag-pol ancestry. *J Virol*. **73**:7021-6
- Mayer J, Meese E, and Müller-Lantzsch N (1998). Human endogenous retrovirus K homologous sequences and their coding capacity in old world primates. *J Virol*. **72**:1870-1875

- Mayer WE, O'HUigin C, Klein J (1993). Resolution of the HLA-DRB6 puzzle: A case of grafting a de novo-generated exon on an existing gene. *Proc Natl Acad Sci USA*. **21**:135-143
- McClure MA, Johnson MS, Feng D-F, and Doolittle RF (1988). Sequence comparisons of retroviral proteins: Relative rates of change and general phylogeny. *Proc Natl Acad Sci USA*. **85**: 2469-2473
- Medstrand P and Mager (2003). Human specific integrations of the HERV-K endogenous retrovirus family. *J Virol*. **72**:9782-9787
- Medstrand P, Landry JR, Mager DL (2001). Long terminal repeats are used as alternative promoters for the endothelin B receptor and apolipoprotein C-I genes in humans. *J Biol Chem*. **276**:1896-1903
- Medstrand P and Blomber J (1993). Characterization of novel reverse transcriptase encoding human endogenous retroviral sequences similar to type A and type B retroviruses: Differential transcription in normal human tissues. *J Virol*. **67**:6778-6787
- Meisler MH, Ting CN (1993). The remarkable evolutionary history of the human amylase genes. *Crit Rev Oral Med*. **4**:503-509
- Mi S, Lee X, Li X, Veldmann GM, Finnerty H, Racie L, LaVallie E, Tang XY, Edouard P, Howes S, Keith JC, McCoy JM (2000). Syncytin is a captive retroviral envelope protein involved in human placental morphogenesis. *Nature*. **403**:785-789
- Nakagawa K, Brusica V, McColl G, Harrison LC (1997). Direct evidence for the expression of multiple endogenous retroviruses in the synovial compartment in rheumatoid arthritis. *Arthritis Rheum*. **40**:627-38.
- Nelson DT, Goochild NL, Mager DL (1996). Gain of Sp1 sites and loss of repressor sequences associated with a young, transcriptionally active subset of HERV-H endogenous long terminal repeats. *Virology*. **220**:213-218

- Noti JD, Reinemann BC, Petrus MN (1996). Sp1 binds two sites in the CD11c promoter in vivo specifically in myeloid cells and cooperates with AP1 to activate transcription. *Mol Cell Biol.* **6**:2940-50.
- Paces J, Pavlicek A, Zika R, Kapitonov VV, Jurka J, Paces V (2004). HERVd: the Human Endogenous RetroViruses Database: update. *Nucleic Acids Res.* **32**(Database issue):D50.
- Paces J, Pavlicek A, Paces V (2002). HERVd: database of human endogenous retroviruses. *Nucleic Acids Res.* **30**:205-6.
- Paces J, Pavlicek A, and Paces V (2002). HERVd: Database of human endogenous retroviruses. *Nucleic Acids Res.* **30**:205-206
- Plant KE, Routledge SJE, Proudfoot NJ (2001). Intergenic transcription in the human  $\beta$  globin gene cluster. *Molecular And Cellular Biology.* **21**:6507-6514
- Ponferrada VG, Mauck BS, Wooley DP (2003). The envelope glycoprotein of human endogenous retrovirus HERV-W induces cellular resistance to spleen necrosis virus. *Arch Virol.* **148**:659-75.
- Poulin G, Lebel M, Chamberland M, Paradis FW, Drouin J (2000). Specific protein-protein interaction between basic helix-loop-helix transcription factors and homeoproteins of the Pitx family. *Mol Cell Biol.* **13**:4826-37.
- Poulin G, Turgeon B, Drouin J (1997). NeuroD1/beta2 contributes to cell-specific transcription of the proopiomelanocortin gene (1997). *Mol Cell Biol.* **11**:6673-82.
- Ono M, Kawakami M, Ushikubo H (1987). Stimulation of expression of the human endogenous retrovirus genome by female steroid hormones in human breast cancer cell line T47D. *J Virol.* **61**:2059-2062
- Oroszlan S and Luftig RB (1990). Retroviral Proteinases. *Curr Top Microbial Immunol.* **157**:153-85
- Quandt K, Frech K, Karas H, Wingender E, Werner T (1995). MatInd and MatInspector: new fast and versatile tools for detection of consensus matches in nucleotide sequence data. *Nucleic Acids Res.* **23**:4878-84

- Quantien MH, Manfroid I, Moncet D, Gunz G, Muller M, Grino M, Enjalbert A, Pellegrini I (2002). Pitx factors are involved in basal and hormone-regulated activity of the human prolactin promoter. *J Biol Chem.* **277**:44408-16
- Renbaum P, Beerli R, Gabai E, Amiel M, Gal M, Ehrenguber MU, Levy-Lahad E (2003). Egr-1 upregulates the Alzheimer's disease presenilin-2 gene in neuronal cells. *Gene.* **318**:113-24
- Reus K, Mayer J, Sauter M, Zischler H, Müller-Lantzsch, Meese E (2001). HERV-K (OLD): Ancestor sequences of the human endogenous retrovirus family HERV-K (HML-2). *J Virol.* **75**:8917-8926
- Routledge SJE and Proudfoot NJ (2002). Definition of transcriptional promoters in the human  $\beta$  globin locus control region. *J. Mol. Biol.* **323**:601-611
- Saitou N and Nei M (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol.* **4**:406-25.
- Saitou N and Nei M (1986). The number of nucleotides required to determine the branching order of three species, with special reference to the human-chimpanzee-gorilla divergence. *J Mol Evol.* **24**:189-204.
- Sambrook J, Fritsch EF, Maniatis T (1989). Molecular cloning-A Laboratory manual. *Cold Spring Harbour Press.*
- Samuelson LC, Philips RS, Swanberg LJ (1996). Amylase gene structures in primates: Retroposon insertions and promoter evolution. *Mol Biol Evol.* **13**:767-779
- Samuelson LC, Wiebauer K, Snow SK, Meisler MH (1990). Retroviral and pseudogene insertion sites reveal the lineage of human salivary and pancreatic amylase genes from a single gene during primate evolution. *Mol. Cell Biol.* **10**:2513-2520
- Samuelson LC, Wiebauer K, Gumucio DL, Meisler MH (1988). Expression of the human amylase genes: recent origin of salivary amylase promoter from an actin pseudogene. *Nucleic Acids Res.* **16**:8261-8276

- Sanger F, Nicklen S, Coulson AR (1977). DNA sequencing with chain-terminating inhibitors. *Biotechnology*. **24**:104-8
- Sawyer SL, Emermann M, Malik HS (2005). Positive selection of primate TRIM5 $\alpha$  identifies a critical species-specific retroviral restriction domain. *PNAS*. Vol. **102**: 2832-2837
- Sawyer SL, Emermann M, Malik HS (2004). Enzyme APOBEC3G. *Plos Biol*. **2**: e275
- Schön U, Seifarth W, Baust C, Hohenadl C, Erfle V, Leib-Moesch C (2001). Cell type-specific expression and promoter activity of human endogenous retroviral long terminal repeats. *Virology*. **279**:280-291
- Schulte AM, Malerczyk C, Cabal-Manzano R, Gajarsa JJ, List HJ, Tate-Riege A, Wellstein A (2000). Influence of the human endogenous retrovirus-like element HERV-E.PTN on the expression of growth factor pleiotrophin: a critical role of a retroviral Sp1-binding site. *Ocogene* . **19**:3988-3998
- Schulte AM, Lai S, Kurtz A, Czubayko F, Riegel AT, Wellstein A (1996). Human trophoblast and choriocarcinoma expression of the growth factor pleiotrophin attributable to germ-line insertion of an endogenous retrovirus. *Proc Natl Acad Sci USA*. **93**:14759-14764
- Schulte AM, Wellstein A (1998). Structure and phylogenetic analysis of an endogenous retrovirus inserted into the human growth factor gene pleiotrophin. *J Virol*. **72**:6065-6072
- Seifarth W, Baust C, Murr A, Skladny H, Krieg-Schneider F, Blusch J, Werner T, Hehlmann R, Leib-Moesch C (1998). Proviral structure, chromosomal location and expression of HERV-K-T47D, a novel human endogenous retrovirus derived from T47D particles. *J Virol*. **72**:8384-8391
- Seifarth W, Baust C, Schon U, Reichert A, Hehlmann R, Leib-Moesch C (2000). HERV-IP-T47D, a novel type C-related human endogenous retroviral sequence derived from T47D particles. *AIDS Res Hum Retroviruses*. **16**:471-480

- Seifarth W, Frank O, Zeifelder U, Spiess B, Greenwood AD, Hehlmann R, Leib-Mösch C (2005). Comprehensive analysis of human endogenous retrovirus transcriptional activity in human tissues with a retrovirus-specific microarray. *J Virol.* **79**:341-52.
- Seifarth W, Skladny H, Kief-Schneider F, Reichert A, Hehlmann R, Leib-Moesch C (1995). Retrovirus-like particles released from the human breast cancer cell line T47-D display type B- and C-related endogenous retroviral sequences. *J Virol.* **69**:6408-6416
- Shelley CS, Teodoridis JM, Park H, Farokhzad OC, Bottinger EP, Aranout MA (2002). During differentiation of the monocytic cell line U937, Pur alpha mediates induction of the CD11c beta 2 integrin gene promoter. *J Immunol.* **168**:3887-93.
- Shih A, Couvavas EE, and Rush MG (1991). Evolutionary implications of primate endogenous retroviruses. *Virology.* **182**:4x95-501
- Singer MF, Krek V, McMillan JP, Swergold GD, Thayer RE (1993). Line-1: A human transposable element. *Gene* **135**:183-188
- Sjøttem E, Anderssen S, Johansen T (1996). The promoter activity of long terminal repeats of the HERV-H family of human retrovirus-like elements is critically dependent on Sp1 family proteins interacting with a GC/GT box located immediately 3' to the TATA box. *J Virol.* **70**:188-198
- Stauber RH, Horie K, Carney P, Hudson EA, Tarasova NI, Gaitanaris GA, Pavlakis GN (1998). Development and applications of enhanced green fluorescent protein mutants. *Biotechniques.* **24**:462-6, 468-71
- Steele PE, Rabson AB, Bryan T (1984). Distinctive termini characterize two families of human endogenous retroviral sequences. *Science.* **225**:943-947
- Stoye JP (2001). Endogenous retroviruses: still active after all these years? *Curr Biol.* **11**:R914-6. Review
- Strazzullo M, Majello B, Lania L, La Mantia G (1994). Mutational analysis of the human endogenous ERV9 proviruses promoter. *Virology.* **200**:686-695

- Suzuki H, Hosokawa Y, Toda H, Nishikimi M, Ozawa T (1990). Common protein binding sites in the 5'-flanking regions of human genes for cytochrome c1 and ubiquinone-binding protein. *J Biol Chem.* **265**:8159-8163
- Tanaka S, Ikeda H, Osaka N, Yamato Y, Sugaya T, Yoshiki T (2003). Tissue specific high level expression of full length human endogenous retrovirus genome transgene, HERV-R, under control of its own promoter in rats. *TransgenicRes.* **12**:319-28
- Temin HM (1980). Origin of retroviruses from cellular moveable genetic elements. *Cell.* **21**: 599-600
- Thompson JD., Gibson TJ., Plewniak F., Jeanmougin F. and Higgins DG (1997). The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research.* **25**:4876-4882
- Ting CN, Rosenberg MP, Snow MS, Samuelson LC, Meisler MH (1992). Endogenous retroviral sequences are required for tissue-specific expression of a human salivary amylase gene. *Genes and development.* **6**:1457-1465
- Ting CN, Rosenberg MP, Snow CM, Samuelson LC, Meisler MH (1992). Endogenous retroviral sequences are required for tissue-specific expression of human salivary amylase gene. *Genes Dev.* **6**:1457-1465
- Tremblay JJ and Drouin J (1999). Egr-1 is a downstream effector of GnRH and synergizes by direct interaction with Ptx1 and SF-1 to enhance luteinizing hormone beta gene transcription. *Mol Cell Biol.* **4**:2567-76
- Tristem M (2000). Identification and characterization of novel human endogenous retrovirus families by phylogenetic screening of the human genome mapping project database. *J Virol.* **74**(8):3715-30
- Van der Kuyl AC, Mang R, Decker JC and Goudsmit J (1997). Complete nucleotide sequence of simian endogenous type D retrovirus with intact genome organization: evidence for ancestry to simian retrovirus and baboon endogenous virus. *J Virol.* **71**:3666-76

- Van de Lagemaat LN, Landry JR, Mager DL, Medstrand P (2003). Transposable elements in mammals promote regulatory variation and diversification of genes with specialized functions. *Trends Genet.* **19**: 530-6
- Van Regenmortel MHV (2000). Academic Press, San Diego. p. 1167
- Varmus HE, Brown P (1989). Retroviruses. In: Berg DE, Howe MM. Mobile DNA, 53-108, American Society of Microbiology, Washington DC, USA
- Varmus HE, Coffin JM (1982). RNA tumor viruses, Band 2, 75-134, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, USA
- Venables PJ, Brookes SM, Griffith D, Weiss RA, Boyd MT (1995). Abundance of an endogenous retroviral envelope protein in placental trophoblasts suggests a biological function. *Virology.* **211**:589-592
- Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, Smith HO, Yandell M, Evans CA, Holt RA, Gocayne JD, Amanatides P, Ballew RM, Huson DH, Wortman JR, Zhang Q, Kodira CD, Zheng XH, Chen L, Skupski M, Subramanian G, Thomas PD, Zhang J, Gabor Miklos GL, Nelson C, Broder S, Clark AG, Nadeau J, McKusick VA, Zinder N, Levine AJ, Roberts RJ, Simon M, Slayman C, Hunkapiller M, Bolanos R, Delcher A, Dew I, Fasulo D, Flanigan M, Florea L, Halpern A, Hannenhalli S, Kravitz S, Levy S, Mobarry C, Reinert K, Remington K, Abu-Threideh J, Beasley E, Biddick K, Bonazzi V, Brandon R, Cargill M, Chandramouliswaran I, Charlab R, Chaturvedi K, Deng Z, Di Francesco V, Dunn P, Eilbeck K, Evangelista C, Gabrielian AE, Gan W, Ge W, Gong F, Gu Z, Guan P, Heiman TJ, Higgins ME, Ji RR, Ke Z, Ketchum KA, Lai Z, Lei Y, Li Z, Li J, Liang Y, Lin X, Lu F, Merkulov GV, Milshina N, Moore HM, Naik AK, Narayan VA, Neelam B, Nusskern D, Rusch DB, Salzberg S, Shao W, Shue B, Sun J, Wang Z, Wang A, Wang X, Wang J, Wei M, Wides R, Xiao C, Yan C, Yao A, Ye J, Zhan M, Zhang W, Zhang H, Zhao Q, Zheng L, Zhong F, Zhong W, Zhu S, Zhao S, Gilbert D, Baumhueter S, Spier G, Carter C, Cravchik A, Woodage T, Ali F, An H, Awe A, Baldwin D, Baden H, Barnstead M, Barrow I, Beeson K, Busam D, Carver A, Center A, Cheng ML, Curry L, Danaher S, Davenport L, Desilets R, Dietz S, Dodson K, Doup L, Ferriera S, Garg N, Gluecksmann A, Hart B, Haynes J, Haynes C,

Heiner C, Hladun S, Hostin D, Houck J, Howland T, Ibegwam C, Johnson J, Kalush F, Kline L, Koduru S, Love A, Mann F, May D, McCawley S, McIntosh T, McMullen I, Moy M, Moy L, Murphy B, Nelson K, Pfannkoch C, Pratts E, Puri V, Qureshi H, Reardon M, Rodriguez R, Rogers YH, Romblad D, Ruhfel B, Scott R, Sitter C, Smallwood M, Stewart E, Strong R, Suh E, Thomas R, Tint NN, Tse S, Vech C, Wang G, Wetter J, Williams S, Williams M, Windsor S, Winn-Deen E, Wolfe K, Zaveri J, Zaveri K, Abril JF, Guigo R, Campbell MJ, Sjolander KV, Karlak B, Kejariwal A, Mi H, Lazareva B, Hatton T, Narechania A, Diemer K, Muruganujan A, Guo N, Sato S, Bafna V, Istrail S, Lippert R, Schwartz R, Walenz B, Yooseph S, Allen D, Basu A, Baxendale J, Blick L, Caminha M, Carnes-Stine J, Caulk P, Chiang YH, Coyne M, Dahlke C, Mays A, Dombroski M, Donnelly M, Ely D, Esparham S, Fosler C, Gire H, Glanowski S, Glasser K, Glodek A, Gorokhov M, Graham K, Gropman B, Harris M, Heil J, Henderson S, Hoover J, Jennings D, Jordan C, Jordan J, Kasha J, Kagan L, Kraft C, Levitsky A, Lewis M, Liu X, Lopez J, Ma D, Majoros W, McDaniel J, Murphy S, Newman M, Nguyen T, Nguyen N, Nodell M, Pan S, Peck J, Peterson M, Rowe W, Sanders R, Scott J, Simpson M, Smith T, Sprague A, Stockwell T, Turner R, Venter E, Wang M, Wen M, Wu D, Wu M, Xia A, Zandieh A, Zhu X. (2001). The sequence of the human genome. *Science* **16**: 1304-51

- Wang Y, Kowalski PE, Thalmann I, Ornitz DM, Mager DL, Thalmann R (1998). Otoconin-90, the mammalian otoconial matrix protein contains two domains of homology to secretory phospholipase A2. *Proc Natl Acad Sci USA*. **95**:15345-15350
- Waterston RH, Lindblad-Toh K, Birney E, Rogers J, Landers ES; Mouse Genome Sequencing Consortium. (2002) Initial sequencing and comparative analysis of the mouse genome. *Nature* **5**; 420 (6915): 520-62
- Waugh O' Neill RJ, Waugh O' Neill MJ, Graves JA (1998). Undermethylation associated with chromosome remodelling in an interspecific mammalian hybrid. *Nature* **392**:68-72

- Wentzensen N, Wilz B, Findeisen P, Wagner R, Dippold W, von Knebel Doeberitz M, Gebert J (2004). Identification of differentially expressed genes in colorectal adenoma compared to normal tissue by suppression subtractive hybridization. *IntJOncol.* **24**:987-94
- Wilkinson DA, Mager DL and Leong JC (1994). Endogenous human retroviruses, in Levy JA (ed.), *The Retroviridae*. Plenum Press, New York, pp. 465-535
- Wilkinson DA, Goodchild NL, Saxton TM, Wood S and Mager DL (1993). Evidence for a functional subclass of the RTL-V-H family of human endogenous retrovirus-like sequences. *J Virol.* **67**:2981-2989
- Xiong Y and Eickbush TH (1990). Origin and evolution of retroelements based upon their reverse transcriptase sequences. *EMBO J.* **9**:3353-3362
- Yang J, Bogerd HPB, Peng S, Wiegand H, Truant R and Cullen BR (1999). An ancient family of human endogenous retroviruses encodes a functional homolog of the HIV-1 Rev protein. *Proc Natl Acad Sci USA.* **96**:13404-13408
- Yoder JA, Wash CP, Bestor TH (1997). Cytosine methylation and the ecology of intragenomic parasites. *Trends Genet.* **13**:335-40. Review

### 8 Abkürzungen

A	Adenosin
AP1	Aktivatorprotein 1
AP4	Aktivatorprotein 4
ATCC	„American Type Culture Collection“, Manassas, USA
ATP	Adenosintriphosphat
BCA	Bicinchonininsäure
BLAST	„Basic Local Alignment Search Tool“
bp	Basenpaar
BSA	Bovines Serumalbumin
C	Cytidin
CIP	Alkalische Phosphatase („Calf Intestine Phosphatase“)
cDNA	Komplementäre DNA (erzeugt durch reverse Transkription eines mRNA Stranges)
cpm	Maß für Radioaktivität. Durch Szintillationszähler bestimmt („counts per minute“)
DMEM	„Dulbecco’s Modified Eagle“ Medium
DMSO	Dimethylsulfoxid
DNA	Desoxyribonukleinsäure
dNTP	Desoxynukleotidtriphosphat
DTT	Dithiothreitol
EDTA	Ethylendiamintetraacetat
EMSA	„Electrophoretic Mobility Shift Assay“; Gelshift
ERV	Endogenes Retrovirus

## 8. Abkürzungen

---

EtBr	Ethidiumbromid
EtOH	Ethanol
for	„Forward“
FKS	Fötale Kälberserum
G	Guanosin
GFP	Grün-fluoreszierendes Protein („Green Fluorescent Protein“)
h	Stunde
HP	„Hot Probe“
HERV	Humanes Endogenes Retrovirus
IPTG	Isopropyl- $\beta$ -D-thio-galactopyranosid
IRF	„InterferonRegulatory Factors“
kb, kbp	Kilobasen, Kilobasenpaare
kD, kDa	Kilodalton
LB	Bakterienmedium nach Luria-Bertani
LINE	„Long interspersed Element“
LTR	„Long Terminal Repeat“
M	Adenosin oder Cytosin
MCS	Multiple Klonierungssequenz („Multiple Cloning Site“)
min	Minute
mRNA	Boten-Ribonukleinsäure („messenger RNA“)
n	Anzahl
NCBI	„National Center for Biotechnology Information“, Bethesda, USA
ORF	Offener Leserahmen („Open Reading Frame“)
OD	Optische Dichte
PAGE	Polyacrylamidgelelektrophorese

## 8. Abkürzungen

---

PBS	Phosphat-gepufferte Kochsalzlösung („phosphate buffered saline“)
PCR	Polymerase Kettenreaktion („Polymerase Chain Reaction“)
PcEV	„Porcine Endogenous Retrovirus“
PMSF	Phenylmethylsulfonylfluorid
pot.	Potentiell
RCAT	„Retroviral CCAAT Binding Factors“
R-Region	„Redundant Region“
rev	„Reverse“
RNase	Ribonuklease
RNA	Ribonukleinsäure
rRNA	Ribosomale RNA
RT	Reverse Transkriptase
S	Cytosin oder Guanodin
SDS	Natriumdodecylsulfat („Sodiumdodecylsulfate“)
sec	Sekunde
SINE	„Short Interspersed Element“
SERV	„Simian Endogenous Retrovirus“
SOC	Bakterienvollmedium
T	Thymin
Tab.	Tabelle
<i>Taq</i>	<i>Thermus aquaticus</i>
TAE	Tris-Acetat-EDTA
TBPF	„TATA Binding Protein Factor“
TE	Tris-EDTA
TF	Transkriptionsfaktor

## 8. Abkürzungen

---

T <sub>m</sub>	Schmelzpunkt
Tris	Trishydroxyaminomethan
tRNA	Transfer Ribonukleinsäure
U	„Units“ (Enzymeinheiten)
UTR	„5' Untranslated Region“
vol	Volumenanteil
WT	Wildtyp
X-gal	5-Bromo-4chloro-3-indolyl-β-D-galactopyranosid
Y	Cytosin oder Thymidin