

Vergleichende Analyse aktivierter Proviren der humanen endogenen
Retrovirus-Familie HERV-K/HML 2 und Erarbeitung eines Modells zur
Steuerung ihrer Expression in Melanom- und Keimzelltumorlinien.

Dissertation zur Erlangung des Doktorgrades der Naturwissenschaften

vorgelegt beim Fachbereich
Biochemie, Chemie und Pharmazie der Johann Wolfgang Goethe –Universität
in Frankfurt am Main

von
Nina Verena Fuchs
aus Bielefeld

Frankfurt am Main 2009
(D 30)

vom Fachbereich Biochemie, Chemie und Pharmazie der
Johann Wolfgang Goethe - Universität als Dissertation angenommen.

Dekan: Prof. Dr. D. Steinhilber

Gutachter : Prof. Dr. R. Marschalek
Prof. Dr. J. Löwer

Datum der Disputation : 01.12.2010

Inhaltsverzeichnis

1. EINLEITUNG	1
1.1. Retroelemente	1
1.2. Retroviren	2
1.2.1. Klassifikation und Taxonomie der Retroviren	3
1.2.2. Morphologie der Retroviren	3
1.3. Endogen Retroviren (ERVs)	6
1.3.1. HERV-K HML 2	8
1.3.1.1. Genomische Organisation und Genprodukte von HERV-K	9
1.3.1.2. Die biologische Bedeutung humaner endogener Retroviren	11
1.4. Transkription	14
1.4.1. Core-Promotor-Elemente	16
1.4.2. Transkriptionskontrolle und Initiation retroviraler Gene	18
1.4.3. Transkriptionsregulatoren	20
1.4.4. Transkriptionsinitiation und Regulation der HERV-K LTR	22
1.5. Ziele dieser Arbeit	23
2. ERGEBNISSE	24
2.1. Analyse der HERV-K Expression in der Melanomlinie UKRV Mel 2-C9 ...	24
2.1.1. Etablierung einer Melanomlinie mit starker HERV-K Expression	24
2.1.2. Nachweis der HERV-K Proteine Gag, Env und Rec mit spezifischen Antikörpern in der Immunfluoreszenz und im Immunoblot	25
2.2. Vergleich transkriptionell aktiver Proviren in der Melanomlinie UKRV Mel 2 und Keimzelltumorlinien	29
2.2.1. Analyse möglicher HERV-K Polymorphismen	29
2.2.2. Identifikation aktiv transkribierter HERV-K Proviren in der Melanomlinie UKRV Mel 2-C9, und den Keimzelltumorlinien NCCIT und GH	31
2.3. Die Melanomlinie UKRV Mel 2-C9 bildet Retroviruspartikel	33
2.3.1. Analyse der Partikelmorphologie im Elektronenmikroskop	35
2.3.2. HERV-K RNA wird in HERV-K Partikel verpackt	36
2.3.3. HERV-K 108 der Linie UKRV Mel 2-C9 besitzt keine funktionale Reverse Transkriptase	38
2.4. Analyse der HERV-K LTR	39
2.4.1. Bestimmung von Transkriptionsstart und Polyadenylierungsstelle .. (5' und 3' RACE)	39
2.4.2. Vergleich der HERV-K LTR Promotoraktivität in UKRV Mel 2-C9 und GH	42
2.4.3. Besitzt das TATA-Motiv Relevanz für die LTR-Aktivität von HERV-K in Melanomen?	44
2.5. <i>In silico</i> Suche nach an der HERV-K Transkriptionsinitiation beteiligten Faktoren und Identifikation putativer Transkriptionsfaktorbindestellen	46
2.5.1. Identifikation putativer Transkriptionsfaktorbindestellen in der HERV-K LTR durch <i>in silico</i> Datenbankanalysen	47
2.5.1.1. Analyse der Expression von Sp1 und Sp3 in den untersuchten Zelllinien im Immunoblot	48
2.5.1.2. Bindung von Sp1 und Sp3 in räumlicher Nähe zum Transkriptionsstart (<i>in vitro</i> EMSA-Analysen)	50

2.5.1.3.	Mutation der Sp1/Sp3-Bindungsstelle in der GC-Box 1 reduziert <i>in-vivo</i> signifikant die LTR Promotoraktivität.....	53
2.5.1.4.	<i>siRNA</i> vermittelter <i>Knockdown</i> von Sp1 and Sp3 reduziert die HERV-K Promotoraktivität	54
2.5.1.5.	Bindung von Sp1 und Sp3 an den HERV-K Promotor	56
	(<i>in vivo</i> ChIP-Assay)	56
3.	DISKUSSION.....	59
3.1.	Expression von HERV-K in der Melanomlinie UKRV Mel 2-C9.....	59
3.2.	Transkriptionsregulation von HERV-K in Melanomen im Vergleich zu Teratokarzinomen.....	65
4.	MATERIAL UND METHODEN.....	74
4.1.	Allgemeine Materialien und Geräte.....	74
4.2.	Zellbiologische Methoden.....	74
4.2.1.	Geräte, Medien und Chemikalien	74
4.2.2.	Zelllinien und Kulturbedingungen.....	75
4.2.3.	Einfrieren und Auftauen von Zellen	76
4.2.4.	Bestimmung der Zellvitalität und Zellzahl mit Trypanblau	76
4.2.5.	Mycoplasmentest.....	77
4.2.6.	Subklonierung eukaryotischer Zellen	77
4.3.	Bakterienkultur	77
4.3.1.	Medien und Chemikalien.....	77
4.3.2.	Bakterienstämme	78
4.3.3.	Anlegen einer Übernachtskultur.....	78
4.3.4.	Erzeugung kompetenter Bakterien	78
4.3.5.	Transformation von Bakterien mit rekombinanter DNA	78
4.3.6.	Selektionsprinzip.....	79
4.3.7.	Plasmidpräparation (Mini -, Midi- und Maxipräparation).....	79
4.3.8.	Einfrieren und Lagern von Bakterien.....	80
4.4.	Molekularbiologische Methoden.....	80
4.4.1.	Nukleinsäure-Techniken.....	80
4.4.1.1.	Photometrische Bestimmung von Nukleinsäurekonzentrationen	80
4.4.1.2.	Phenol-Chloroform-Extraktion von Nukleinsäuren	81
4.4.1.3.	Fällung von Nukleinsäuren	81
4.4.2.	Ribonukleinsäure-Techniken	82
4.4.2.1.	Isolierung zytoplasmatischer RNA	82
4.4.2.2.	Poly-A-Trakt- <i>mRNA</i> Aufreinigung.....	82
4.4.2.3.	Elektrophorese von RNA in denaturierenden Agarosegelen.....	82
4.4.2.4.	Verdau von DNA (DNaseI-Behandlung)	83
4.4.2.5.	cDNA-Synthese (Reverse Transkription).....	83
4.4.3.	Desoxyribonukleinsäure-Technik.....	84
4.4.3.1.	Isolierung von genomischer DNA aus eukaryotischen Zellen	84
4.4.3.2.	Restriktionsenzymatische Spaltung von DNA.....	84
4.4.3.3.	Auftrennung der DNA-Fragmente im Agarosegel.....	85
4.4.3.4.	DNA Extraktion aus Agarosegelen.....	86
4.4.3.5.	Ligation von DNA in das pGEM [®] -T Easy Vektor System.....	86
4.4.3.6.	Polymerase-Kettenreaktion (PCR) (Saiki <i>et al.</i> , 1985).....	86
4.5.	Electrophoretic Mobility Shift Assay (EMSA) und Supershift Assay	87
4.5.1.	Radioaktive Markierung von Nukleinsäuren	87

4.5.2.	Oligonukleotid-Hybridisierung	88
4.5.3.	Protein-Nukleinsäure Bindungsreaktion	88
4.5.4.	native Polyacrylamid Gel Elektrophorese	89
4.6.	Bestimmung des proviralen Transkriptionsstart und -endpunktes.....	89
4.7.	Proteinchemische Methoden	90
4.7.1.	Proteinextraktion aus eukaryotischen Zellen	90
4.7.2.	Fraktionierte Proteinpräparation aus Zytoplasma und Kern	91
4.7.3.	Proteinbestimmung (nach Bradford)	92
4.7.4.	SDS-Polyacrylamidgel-Elektrophorese (SDS-PAGE)	93
4.7.5.	Luziferase-Assay	94
4.7.5.1.	Dualer Luziferase-Assay	94
4.7.5.2.	Luziferase-Assay mit siRNA- <i>Knockdown</i>	96
4.8.	Immunologische Methoden	96
4.8.1.	Proteinnachweis durch Antikörper auf Trägermembran (Immunoblot)	96
4.8.2.	Immunopräzipitation.....	98
4.8.3.	Immunfluoreszenzfärbung	99
4.8.4.	Chromatinimmunopräzipitation	100
4.9.	Analyse retroviraler Partikel	102
4.9.1.	Isolierung von RNA und Protein aus retroviralen Partikeln	102
4.9.2.	Elektronenmikroskopie retroviraler Partikel.....	103
4.10.	Sequenzanalysen und Datenbankrecherchen	104
5.	ZUSAMMENFASSUNG.....	105
6.	LITERATURVERZEICHNIS.....	106
7.	ABKÜRZUNGSVERZEICHNIS	118
8.	ANHANG	120
8.1.	Primerliste.....	120
8.2.	Sequenzen	121
8.3.	Datenbankauswertung.....	152
8.4.	Veröffentlichungen und Kongressbeiträge	157
	LEBENS LAUF	158
	DANKSAGUNG	159
	ERKLÄRUNG	160

1. EINLEITUNG

1.1. Retroelemente

Die beinahe vollständige Entschlüsselung des menschlichen Genoms im Rahmen des Humanen Genom Projekts (*International Human Genome Sequencing Consortium*, 2001) zeigten, dass nur ein geringer Teil von etwa 1,5 % des Genoms aus kodierenden Exons besteht (Abb.: 1). Dagegen werden 45 % des menschlichen Genoms von transponierbaren mobilen Elementen gebildet (Lander *et al.*, 2001). Die transposablen Elemente lassen sich in vier Kategorien einteilen, LINES (*long interspersed elements*), SINEs (*short interspersed elements*), LTR-Retrotransposons und DNA-Transposons (Abb.: 1).

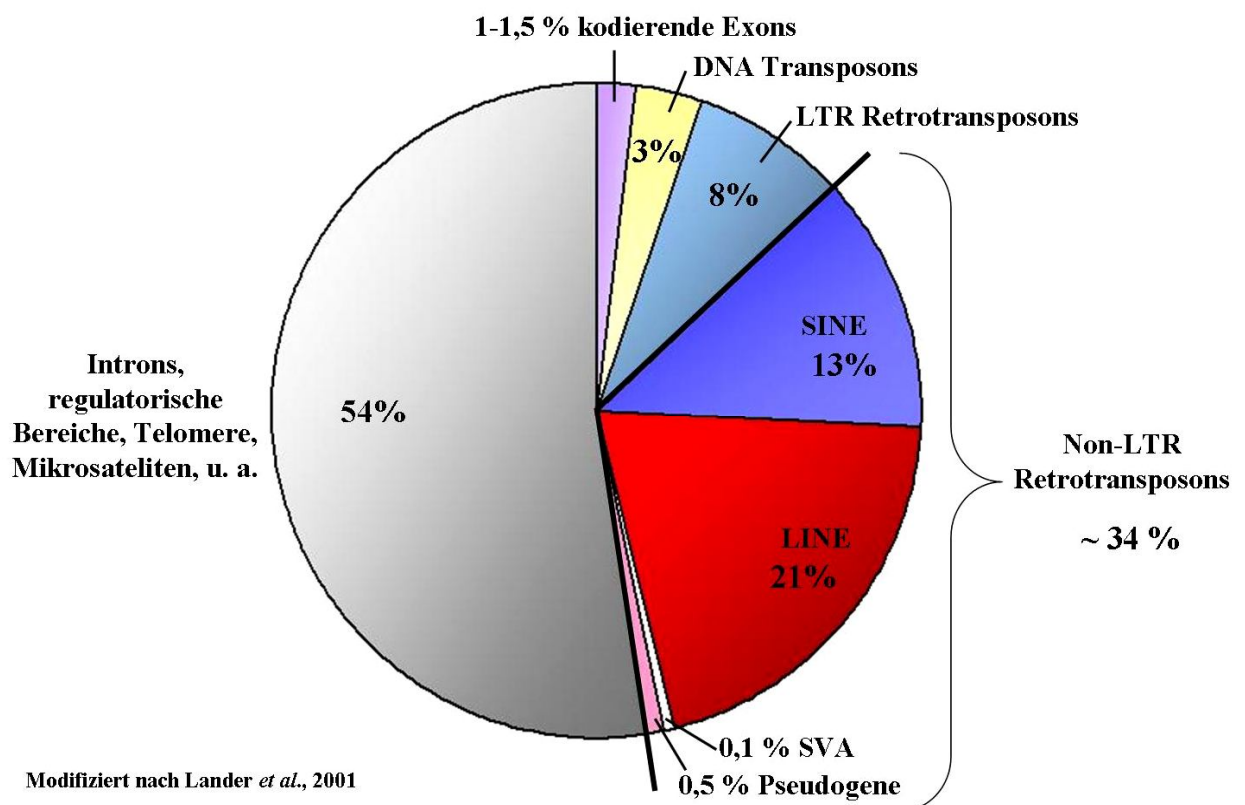


Abbildung 1: Zusammensetzung des menschlichen Genoms

Die DNA-Transposons sind in der Lage, ihren genomischen Lokus über eine DNA-Zwischenstufe zu verändern, LTR-Retrotransposons, SINEs und LINEs vollziehen ihre Replikation über ein RNA-Zwischenprodukt (Wilkinson *et al.*, 1994). Der Replikationszyklus gliedert sich dabei in folgende Schritte: Transkription der DNA durch eine RNA-Polymerase des Wirtsorganismus, reverse Transkription der mRNA durch eine von diesen Elementen codierte RNA abhängige DNA-Polymerase (Reverse Transkriptase)

und (Re) Integration der synthetisierten DNA in das Genom. In Folge des Integrationsmechanismus wird jedes Retroelement von kurzen Sequenzwiederholungen eingeschlossen (*direct repeats*). Die Klassifikation der Retroelemente erfolgt auf Grund von strukturellen und funktionalen Merkmalen. Man unterscheidet Elemente mit LTR (*long terminal repeat*) und ohne LTR. Die Non-LTR-Elemente stellen mit ~34 % einen großen Anteil des menschlichen Genoms dar (Lander *et al.*, 2001). Sie werden weiter unterteilt in autonome Elemente, die in der Lage sind, sich selbst im Genom zu verbreiten (LINEs), und in nichtautonome Elemente (SINEs, SVA, Pseudogen), deren Verbreitung wahrscheinlich über LINE-Proteine vermittelt wird. Die Gruppe der LTR-Retrotransposons umfasst die endogenen Retroviren und Retrovirus ähnlichen Elemente (Lander *et al.*, 2001). Ungefähr 8 % des menschlichen Genoms bestehen aus endogenen Retroviren und abgeleiteten Sequenzen.

1.2. Retroviren

In Abhängigkeit des Übertragungsmodus unterscheidet man exogene und endogene Retroviren. Die endogenen Retroviren sind im Genom des Wirtsorganismus fixiert und werden „vertikal“ durch Vererbung über die Keimbahn des Wirtes übertragen. Die exogenen Retroviren werden dagegen „horizontal“ durch Infektion übertragen.

Die Familie der Retroviren (*Retroviridae*) besteht aus RNA-Viren, die gemeinsame Merkmale in der Struktur der Virionen und in der Art der Replikation aufweisen (Levy, 1992). Alle Retroviren zeichnen sich durch das Vorkommen des Enzyms Reverse Transkriptase aus, das erstmalig 1970 von Temin und Baltimore unabhängig voneinander beschrieben wurde. Die unter Wirbeltieren verbreitete Virusfamilie wurde Anfang des 20. Jahrhunderts durch P. Rous entdeckt. Er zeigte, dass Sarkome durch zellfreie Extrakte von erkrankten auf gesunde Hühne übertragen werden konnten (Rous, 1911). Nach ihm wurde der virale Erreger der Erkrankung, als Rous-Sarkom-Virus (RSV) benannt. Die Virus-Wirt-Interaktionen sind ausgesprochen vielfältig: Infektionen mit Retroviren können beim Wirtsorganismus schwere Erkrankungen wie Neoplasien hervorrufen, ebenso kann eine pathogene Wirkung aber auch ausbleiben. In den Fokus des öffentlichen Interesses rückten die Retroviren in der Mitte der 80er Jahre des 20. Jahrhunderts durch die globale Verbreitung des HI-Virus (*human immunodeficiency virus*), dem Erreger des erworbenen Immundefekt-Syndroms (AIDS). Neben dem HI-Virus ist noch ein weiteres humanpathogenes Retrovirus bekannt, das humane T-Zell Leukämievirus (HTLV). Eine

HTLV-Infektion verursacht beim Menschen beispielsweise eine adulte T-Zell Leukämie. Diese Krankheit kann bei akuter Verlaufsform innerhalb von sechs Monaten zum Tode führen. In der biomedizinischen Forschung wird den Retroviren vor allem Aufmerksamkeit in Bereichen der therapeutischen Anwendung als Vektoren für den Gentransfer zuteil. Zudem ist ihre Rolle bei der Entstehung von malignen Tumorerkrankungen Gegenstand von Untersuchungen.

1.2.1. Klassifikation und Taxonomie der Retroviren

Die Klassifikation des „*International Committee on Taxonomy of Viruses*“ umfasst sieben Gattungen (Alpha-, Beta-, Gamma-, Delta- und Epsilon-Retroviren sowie Lenti- und Spumaviren), die sich ihrerseits in Subgenera und Spezies unterteilen. HIV-1 und HIV-2 sind Vertreter der Lentiviren, HTLV-1 und HTLV-2 gehören der Gruppe der Delta-Retroviren an. Verwandtschaftsgrade zwischen den Retroviren lassen sich anhand des relativ gut konservierten Polymerase-Genes ermitteln (Doolittle *et al.*; 1990).

Die veraltete Einteilung der Retroviren in die Klassen A, B, C, und D geschah anhand morphologischer Unterschiede der Viruspartikel im elektronenmikroskopischen Bild (Bernhard, 1960; Fine und Schochetman, 1978).

1.2.2. Morphologie der Retroviren

Die retroviralen Viruspartikel (Abb.: 2) haben einen Durchmesser von ca. 80-130 nm und besitzen einen typischen Aufbau. Die Virionen bestehen aus einem viralen Kapsid (*core*), welches das virale Genom enthält. Es wird von einer Lipoproteinmembran (*envelope*) umgeben, in die die viralen Hüllproteine eingelagert sind. Das Viruskapsid besitzt eine sphärisch-ikosaedrische (Alpha-, Beta-, Gamma-, Delta-Retroviren und Spumaviren) oder konische (einige Beta-Retroviren und Lentiviren) Form.

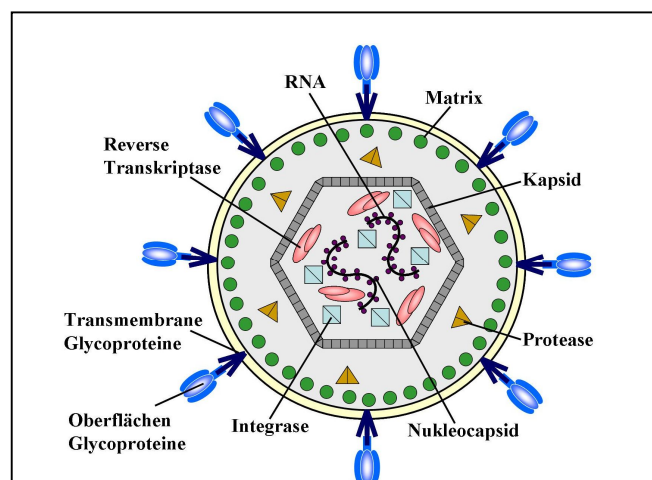


Abbildung 2: Schematische Darstellung eines Virions (modifiziert nach Coffin, 1992)

Die Erbinformation liegt in doppelter Kopie als einzelsträngige (+) RNA vor und wird von Nukleokapsidproteinen umschlossen.

Durch die 5'Cap-Struktur und die 3'-Polyadenylierung besitzt die virale Erbinformation alle Charakteristika einer eukaryotischen mRNA. Neben dem Genom enthält das virale Kapsid auch die viralen Enzyme Protease, Integrase und Reverse Transkriptase. Zwischen dem Kapsid und der umgebenden Lipidmembran werden netzartig miteinander verbundene Matrixproteine eingelagert, diese sind häufig über aminoterminal angefügte Myristinsäurereste in der Membran verankert. Die virale Lipiddoppelschicht entstammt der Wirtszelle. In die Virusmembran wird das transmembrane Hüllprotein (TM-Protein, *trans membrane*) eingelagert. An den extraviralen Bereich des TM-Proteins ist über nicht kovalente Bindung das virale Oberflächen-Protein (SU-Protein, *surface unit*) gekoppelt. Die Glykoproteine liegen als trimere Komplexe vor und vermitteln den Kontakt und die Fusion mit der Zielzelle.

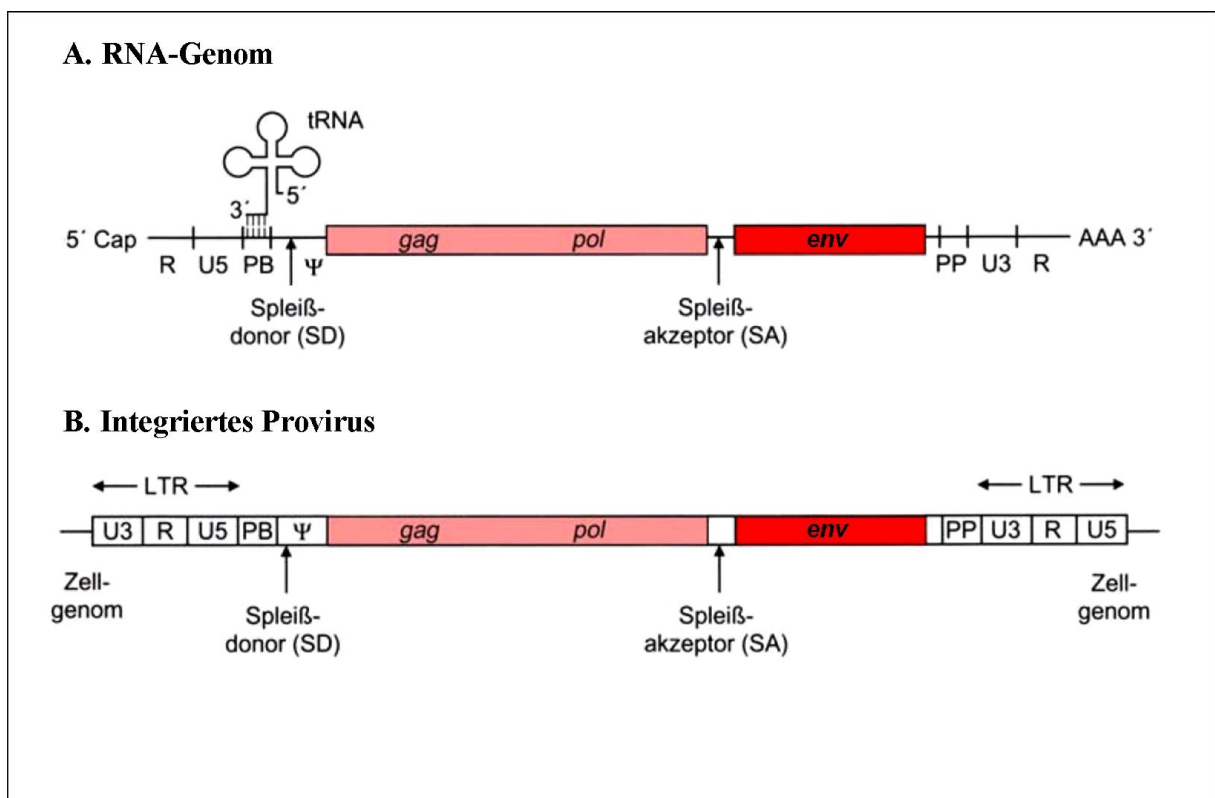


Abbildung 3: Genomorganisation der Retroviren. Vereinfachte schematische Darstellung (A) des viralen RNA-Genoms und (B) des Provirus nach Integration in das Wirtsgenom. Die abgebildeten Merkmale finden sich in allen Retroviren. Ψ : Verpackungssignal, AAA: Polyadenylierung, 5'Cap: 7-Methyl.Guanosin-5'-triphosphat, *gag*: gruppenspezifisches Antigen, *pol*: Polymerase-Gen, *env*: Envelope-Gen, LTR: *long terminal repeat*, PB: Primerbindungsstelle, PP: Poly-Purin-Tract, R: Redundanz-Regionen, U3: Unique 3'-End, U5: Unique 5'-End, SD und SA: Spleißdonor- und Spleißakzeptor (nach Modrow und Falke, 1997).

Die kodierende Region der Retroviren ist zwischen 7 und 12 kb lang und umfasst die Gene *gag*, *pol* und *env* (Abb.: 3). Diese Gene sind für die virale Struktur und Vermehrung nach reverser Transkription notwendig. Das *gag*-Gen (gruppenspezifische Antigene) kodiert für ein Polyprotein, das während des Reifungsprozesses durch die viruseigene Protease in die viralen *Core*proteine (Matrix, Kapsid und Nukleokapsid) gespalten wird. *pol* trägt die genetische Information für mehrere Enzyme, die Protease im 5'-Bereich, dann die Reverse Transkriptase und die Integrase. Sie werden durch Übergehen des Gag-Stopcodons zunächst als Vorläuferproteine synthetisiert und anschließend durch die virale Protease prozessiert wird. Wenn die *gag*, *protease* und *polymerase* Gene in unterschiedlichen Leserahmen liegen, sind zur Bildung der Gag/Protease/Pol-Fusionsproteine Leserahmenverschiebungen nötig.

Das *env* Gen kodiert für die Glykoproteine der viralen Hülle. Die Oberflächen- und Transmembranproteine werden ebenfalls als gemeinsames Vorläuferprotein von einer einfach gespleißten *mRNA* translatiert. Weitere akzessorische Gene, oft mit regulatorischen Eigenschaften, unterscheiden sich von Virus zu Virus. Diese Proteine werden meist von mehrfach gespleißten Transkripten translatiert.

Die kodierenden Regionen des Virusgenoms werden am 5'- und am 3'-Ende von regulatorischen Sequenzen flankiert, die bei allen Retroviren gleich aufgebaut sind (Abb.: 3). Die R-Region (R = redundant) ist für die reverse Transkription und Integrations ins Wirtsgenom von Bedeutung und liegt in identischer Basenabfolge und Orientierung unmittelbar nach der 5'-Cap und vor dem Poly-A-Schwanz am 3'-Ende. An die R-Region am 5'-Ende schließt sich der U5-Bereich an (U = unique), gefolgt von der achtzehn Nukleotide umfassenden Primerbindungsstelle (PBS), die über Basenpaarung an das 3'-Ende einer zellulären t-RNA hybridisiert ist. Das t-RNA-Molekül dient als Primer für die Erststrang-DNA-Synthese während der reversen Transkription und wird mit dem RNA-Genom in die viralen Partikel verpackt. Der Genomabschnitt zwischen der PBS und dem Beginn des *gag*-Gens setzt sich aus Spleißdonor (SD) und ψ -Stelle zusammen und wird als Leader-Sequenz bezeichnet. Dieser Spleißdonor wird für die Herstellung aller gespleißten viralen *mRNAs* verwendet. Die ψ -Stelle dient als Verpackungssignal des viralen RNA-Genoms. Im Anschluss an die kodierenden Gene (3'- des *env*-Gens) befindet sich der Polypurinbereich (PP). Dieser Bereich ist eine Basenfolge von mindestens neun Adenosin- oder Guaninbasen, denen eine Schlüsselrolle bei der Zweitstrangsynthese der reversen Transkription zukommt. Es folgt die U3-Region, die Enhancer und Promotorelemente enthält. An die U3-Region schließt sich abermals die R-Region an, in ihr ist das

Polyadenylierungssignal enthalten, das üblicherweise mehr als zwanzig Nukleotide stromaufwärts der Polyadenylierungsstelle liegt. Bei sehr kurzen R-Regionen kann das Polyadenylierungssignal aber auch im U3-Bereich liegen. Das Ende des retroviralen Genoms bildet der Poly-(A)-Schwanz, er besteht üblicherweise aus 200-300 Adenosinresten.

Durch das Umschreiben der viralen Erbinformation von einzelsträngiger RNA in doppelsträngige DNA kommt es zur Entstehung von Sequenzwiederholungen an beiden Enden des Provirusgenoms, man bezeichnet diese als LTRs. Eine LTR setzt sich aus den Sequenzabschnitten U3, R und U5 zusammen und beinhaltet alle *cis*-aktiven Elemente, die für die Regulation der viralen Genexpression verantwortlich sind (Abb.: 3B). Nach Integration der viralen Erbinformation in das Genom des Wirts, wird das Virus wie ein zelluläres Gen transkribiert. Die Transkripte werden von den zellulären Ribosomen in Proteine translatiert.

1.3. Endogen Retroviren (ERVs)

Endogene Retroviren (ERVs) kommen sowohl im menschlichen Genom als auch im Genom anderer Wirbeltiere vor. ERVs durchlaufen keinen vollständigen Replikationszyklus mehr, sie werden als Teil des Wirtsgenoms in Form von Proviren von Generation zu Generation weitergegeben. Man nimmt an, dass endogene Retroviren durch Infektionen von Keimbahnzellen in das Genom des Wirtes gelangten und dann an die nachfolgende Generation weitergegeben wurden. Endogenisierte Retroviren bleiben nur über eine geringe Anzahl von Wirtsgenerationen hinweg infektiös. Sobald ein Retrovirus in einer Population fixiert ist, sollte auf Grund von Rezeptorinterferenz eine Infektion mit der exogenen Form ausgeschlossen sein. Durch Deletionen und Mutationen, die sich im Laufe von Millionen Jahren ereignet haben, wurden die offenen Leserahmen, die für virale Proteine kodieren, zerstört. Ein Großteil dieser Viren ist daher nicht mehr replikationskompetent oder wird nicht mehr transkribiert. Die Mutationsrate kann durch antiretrovirale Wirtsfaktoren wie beispielsweise Mitglieder der APOBEC3-Familie erhöht werden (Chiu und Green, 2008; Lee *et al.*, 2008). Durch homologe Rekombination zwischen 5'- und 3'-LTR, können die kodierenden Bereiche verloren gehen, es entstehen solitäre LTRs (Wilkinson *et al.*, 1994).

Der Ursprung der meisten der gegenwärtig im humanen Genom vorhandenen endogenen Retroviren, liegt in der Fixierung einst infektiöser exogener Retroviren im Genom von

Primatenvorläufern seit 30 Millionen Jahren (Löwer *et al.*, 1996). Es gibt aber auch einige humanspezifischen Integrate. Humane endogene Retroviren (HERVs) werden aus historischen Gründen nach der t-RNA benannt, die an der Primerbindungsstelle die reverse Transkription initiieren würde. Dabei wird zur Bezeichnung der Einbuchstabencode der Aminosäure der betreffenden t-RNA verwendet. Es sind bereits etwa dreißig sogenannte „Familien“ der HERV beschrieben (Holmes, 2007). Sie können taxonomisch auf Grund von Sequenzhomologien den Genera exogener Retroviren zugeordnet werden. Die Mitglieder der HERV-K „Familie“ zum Beispiel gehören zu den Betaretroviren. Über Sequenzhomologien im Bereich des *pol*-Gens erfolgt eine weitere Einteilung in die Gruppen HML-1 bis HML-10 (Medstrand und Blomberg, 1993; Seifarth *et al.*, 1998). Unsere Arbeitsgruppe beschäftigt sich mit der evolutionär jüngsten Gruppe, HML-2, welche ausschließlich im Genom von Catarrhines (Altwelt-Affen und Hominiden) zu finden ist. Obwohl humane endogene Retroviren nicht mehr infektiös sind, besitzt eine begrenzte Anzahl das Potential, virale Produkte und virusähnliche Partikel zu produzieren (Nelson *et al.*, 2003). Darüber hinaus können sich Auswirkungen endogenisierter Retroviren auf den Wirtsorganismus in Abhängigkeit vom Integrationsort der viralen Sequenz zeigen. Proviren oder abgeleitete Sequenzen können durch enthaltene Promotor- oder Enhancer-Elemente zur Aktivierung zellulärer Gene führen, ebenso ist der Funktionsverlust eines Gens durch ein provirales Integrat möglich (Kobayashi *et al.*, 1998; Ostertag *et al.*, 2003). Die Integration eines HERV-E Elements in die Promotor-Region des Amylasegens führte beispielsweise dazu, dass Amylase nicht nur in der Pankreas, sondern auch in der Speicheldrüse exprimiert wird (Alarcon, 2006).

Replikationskompetente endogene Retroviren sind in einigen Säugern mit Krebsformen assoziiert, wie das murine Leukämievirus (MMTV) in Mäusen oder das JSRV (*jaagsiekte sheep retrovirus*) in Schafen. Diese Viren sind sowohl als exogene als auch als endogene Agenzien in ihren Wirtsspezies verbreitet (Golovkina *et al.*, 1992; de Martini *et al.*, 2003). Durch neue Verfahren wie *MicroArray*-Analysen lassen sich die Expressionsmuster endogener Retroviren sehr gut untersuchen. Dabei zeigen sich häufig gewebsspezifische Expressionsmuster (Wilkinson *et al.*, 1994; Löwer *et al.*, 1996, Rote *et al.*, 2004). HERVs sind überwiegend in Keimbahnzellen, embryonalem Gewebe und in der Plazenta aktiv. Das HERV-W Env Protein wird in humanen Trophoblasten stark exprimiert und ist an der Trophoblasten-Zellfusion und Differenzierung zum Syncytiotrophoblasten beteiligt (Fredo *et al.*, 2003). HERV-H wird häufig in Tumoren des Magen-Darm-Traktes exprimiert (Stauffer *et al.*, 2004).

1.3.1. HERV-K HML 2

HERV-K HML-2 besteht aus 10000 bis 25000 solitären LTRs, die über das gesamte Genom verteilt sind und etwa dreißig Proviren, von denen einige größere Deletionen aufweisen (Seifart *et al.*, 1998). Den Proviren werden zur Kennzeichnung Nummern zugeordnet. Humanspezifische provirale Integrate zeigen dass Vertreter dieser Gruppe auch nach der Trennung von Mensch und Affe, vor circa fünf Millionen Jahren, noch aktiv waren (Barbulescu *et al.*, 1999; Medstrand und Mager, 1998; Turner *et al.*, 2001).

HERV-K-HML-2 Proviren besitzen als einzige HERV Familie vollständige offene Leserahmen für alle viralen Gene (Löwer *et al.*, 1993a; Müller-Lantzsch *et al.*, 1993) und kodieren für virale Partikel (Boller *et al.*, 1993; Löwer *et al.*, 1993a). Im weiteren Verlauf der Arbeit wird HERV-K-HML-2 nur noch als HERV-K bezeichnet.

Das erste HERV-K Provirus wurde 1986 von Ono mit einer Sonde gegen eine konservierte Region des *pol*-Gens im Southern Blot entdeckt und vollständig sequenziert (Ono *et al.*, 1986). HERV-K Partikel konnten erstmals in Teratokarzinom-Zelllinien nachgewiesen werden (Boller *et al.*, 1983; Löwer *et al.*, 1984) und werden daher auch als HTDV Partikel (*human teratocarcinoma derived virus*) bezeichnet. Es können zwei Subtypen von HERV-K Proviren differenziert werden, Typ 1 und Typ 2 (Löwer *et al.*, 1993a) (Abb.: 4). Das von Ono identifizierte Provirus HERV-K 10 und verwandte Proviren weisen eine charakteristische Deletion von 292 bp im Bereich des *env*-Gens auf. Diese werden als Typ 1 bezeichnet.

Die Aktivität von HERV-K zeigt sich auch in dem Auftreten von Polymorphismen (Turner *et al.*, 2001). Das jüngste Integrat weist noch keine Unterschiede in der 5' und 3' LTR auf und ist schätzungsweise weniger als 20000 Jahre alt. Es handelt sich um das auf Chromosom 19p13.11 lokalisierte Provirus HERV-K 113, ID (GenBank) AY037928, das mit einer Frequenz von 10 % in der humanen Population nachgewiesen werden kann. HERV-K 113 ist ein Vollängenintegrat (Typ 2) mit intakten offenen Leserahmen für alle viralen Proteine (Barbulescu *et al.*, 1999). Es ist in der Lage, virale Partikel zu produzieren (Boller *et al.*, 2008), besitzt aber keine funktionale Reverse Transkriptase (Boller *et al.*, 2008) und kein funktionales Env (Dewannieux *et al.*, 2005).

Das zweite polymorphe HERV-K Integrat ist das Provirus HERV-K 115, ID (GenBank) AC134684, welches sich auf Chromosom 8p23.1 befindet. Es zeigt eine Basendeletion 92 bp stromaufwärts des Stop-Codons im *gag*-ORF, daraus ergibt sich ein veränderter Translationsrahmen. Im Gegensatz zu HERV-K 113 zeigen die beiden LTRs von HERV-K 115 bereits Unterschiede. Es gibt zwischen der 5' und 3' LTR fünfzehn

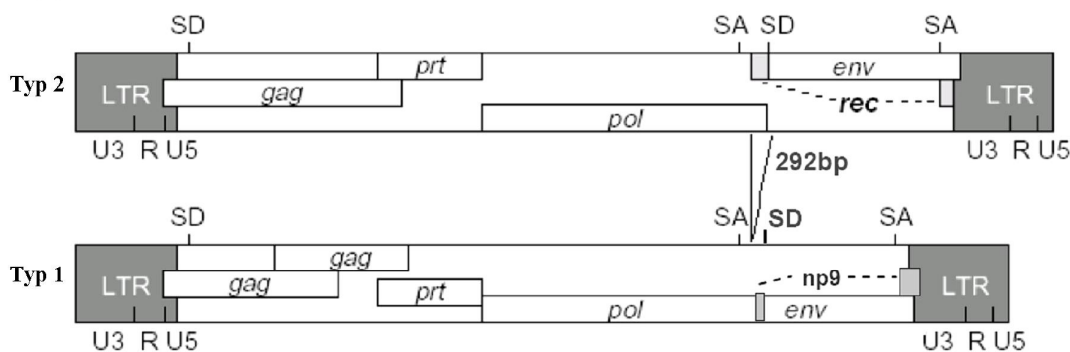
Basensubstitutionen und acht Deletionen. Man nimmt an, dass diese Unterschiede durch ein Rekombinationsereignis zweier Proviren entstanden, da einige der Mutationen auch in den LTRs anderer Proviren präsent sind (Barbulescu *et al.*, 1999).

1.3.1.1. Genomische Organisation und Genprodukte von HERV-K

HERV-K spezifische Volllängen- und gespleißte *mRNA* konnten in Teratokarzinom- und Melanomzelllinien nachgewiesen werden (Löwer *et al.*, 1993; Muster *et al.*, 2003). Ein prototypisches Provirus besitzt offene Leserahmen für die klassischen retroviralen Proteine Gag, Pro, Pol und Env (Abb.: 4). Zusätzlich zur Volllängen-*mRNA* und der einfach gespleißten *env-mRNA* werden noch zwei doppelt gespleißte Transkripte gebildet, die für zwei akzessorische Proteine Rec und Np9 kodieren (Löwer *et al.*, 1995; Armbruster *et al.*, 2002). Des Weiteren wird ein einfach gespleißtes Transkript von 1,5 kb gebildet, mit bisher unbekannter Bedeutung.

Die *gag*, *prt* und *pol* Gene liegen in drei unterschiedlichen Leserahmen. Durch jeweils eine –1 Leserasterverschiebung werden neben dem Gag Polyprotein das Gag-Protease und das Gag-Protease-Polymerase Polyproteine gebildet. Die viruseigene Protease spaltet alle diese Vorläuferproteine in die strukturellen und enzymatischen Einzelkomponenten, Gag zum Beispiel in die Matrix (20 kDa), Kapsid (30 kDa) und Nukleokapsidproteine (16 kDa). Die proteolytische Aktivität der HERV-K Protease (17 kDa) konnte bereits mehrfach gezeigt werden (Boller *et al.*, 1993; Müller-Lantsch *et al.*, 1993; Schommer *et al.*, 1996). Die Aktivität der Reversen Transkriptase (160 kDa) wurde sowohl in Teratokarzinomzelllinien als auch in Melanomen nachgewiesen (Tönjes *et al.*, 1996; Muster *et al.*, 2003). HERV-K kodiert zusätzlich im 5' Bereich des *prt* Gens für eine UTPase (Harris *et al.*, 1997). Dies ist typisch für einige der Betaretroviridae. Von der einfach gespleißten *env-mRNA* wird zuerst ein Env-Vorläuferprotein gebildet. Die Env Vorläuferproteine werden über das endoplasmatische Retikulum und den Golgi-Apparat an die Zelloberfläche transportiert, und dort von einer zellulären Protease in den externen und den transmembranen Teil gespalten. Es konnte gezeigt werden, dass sowohl Env Vorläuferproteine, als auch prozessierte TM und SU-Proteine in den genannten Zelllinien gebildet werden (Tönjes *et al.*, 1997; Muster *et al.*, 2003; Dewannieux *et al.*, 2005).

A. Provirus



B. HERV-K Transkripte

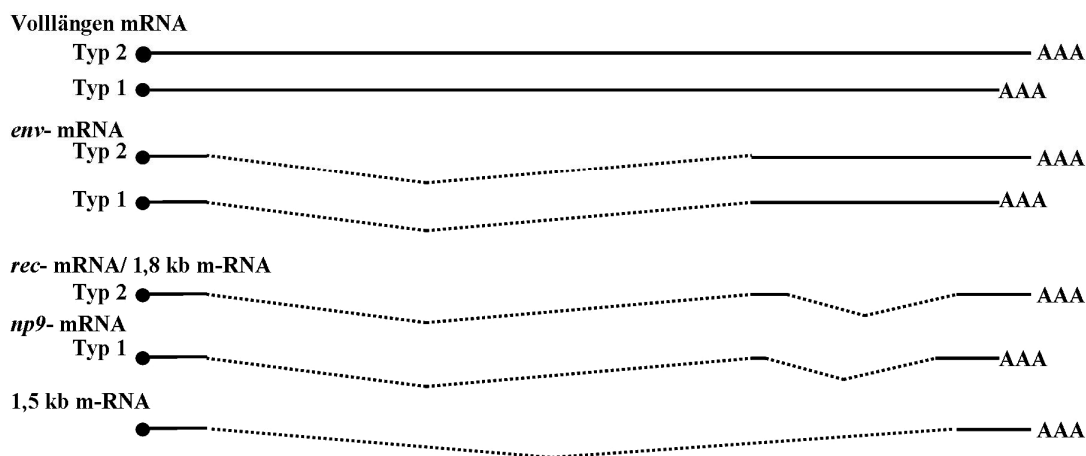


Abbildung 4: Genomische Organisation von HERV-K und Darstellung der gebildeten Transkripte. A. Organisation von HERV-K Proviren der Subtypen 1 und 2. B. Gezeigt sind alle gebildeten mRNA-Spezies: Vollängen mRNA, gespleißte *env*, *rec*, *np9* und 1,5 kb Transkripte (modifiziert nach Löwer *et al* 1995).

Nur HERV-K Proviren vom Typ 1 kodieren für das akzessorische Protein Np9 (Armbruster *et al.*, 2002). Np9 hat ein Molekulargewicht von 8,7 kDa und ist im Kern lokalisiert, es besitzt drei potentielle Kernlokalisationssignale. Proviren vom Typ 2 kodieren dagegen für das Rec Protein (Löwer *et al.*, 1995). Die Bildung von Rec nutzt den gleichen Spleißdonor (SD) der zur Herstellung der *env*-mRNA verwendet wird, daher sind das erste Exon von Rec und das 5'Ende von Env identisch. In einem zweiten Spleißereignis wird das zweite Rec Exon angefügt. Das zweite Exon liegt in einem anderen Leserahmen als Env.

Durch die fehlenden 292 Basen in den Typ 1 Proviren ist der ursprüngliche zweite Spleißdonor nicht mehr vorhanden, zur Entstehung des zweiten Np9 Exons kommt es durch die Bildung eines neuen Spleißdonors. Zur Translation des zweiten Exons von Np9 wird ein zu Env und Rec verschiedener Leserahmen verwendet. Die ersten fünfzehn Aminosäuren sind sowohl in Env, Rec als auch Np9 gleich (Armbruster *et al.*, 2002;

Löwer *et al.*, 1995). Die Expression von Rec und Np9 findet sich vor allem in transformiertem Gewebe (Armbruster *et al.*, 2004).

Das mit einem Molekulargewicht von 12 kDa (Löwer *et al.*, 1995) relativ kleine Rec Protein ist ein RNA Exportadapter (Magin *et al.*, 1999) und in seiner Funktion homolog zum HIV-Rev und HTLV-Rex. Diese Proteine vermitteln den Export viraler Vollängen- und nicht vollständig gespleißter Transkripte aus dem Kern ins Zytoplasma durch Bindung an spezifische Sequenzen bzw. Struktureinheiten (z. B. an das *Rec responsiv element* RcRE, Magin *et al.*, 1999; Magin-Lachmann *et al.*, 2001) der mRNA. Das im Zytoplasma translatierte Rec Protein gelangt durch seine Kernlokalisationssequenz (NLS *nuclear localisation signal*) (Löwer *et al.*, 1995; Magin *et al.*, 1999) zurück in den Kern, um dort an das RcRE zu binden. Dieser Komplex bindet auf Grund des Kernexportsignals (NES *nuclear export signal*) an den zellulären Exportrezeptor Crm1 und gelangt auf diesem Weg ins Zytoplasma (Magin *et al.*, 1999; Boese *et al.*, 2000). Die Lokalisation von Rec ist nukleolär.

1.3.1.2. Die biologische Bedeutung humaner endogener Retroviren

Viele Überreste einst aktiver Viren spielen im Wirtsorganismus wahrscheinlich heute keine Rolle mehr. Da die Expression endogener Sequenzen häufig in Reproduktionsgeweben auftritt (Löwer *et al.*, 1999; Rote *et al.*, 2004) stellt sich die Frage, ob es sich lediglich um Relikte handelt oder ob diese Elemente dem Wirtsorganismus Vorteile bieten. Sollten beispielsweise nach der Reproduktion profitable Effekte durch HERV Aktivitäten in Erscheinung treten, wäre eine funktionale Erhaltung dieser Elemente im Wirtsgenom wahrscheinlich. Durch die Expression einer aktiven Reversen Transkriptase oder Integrase, könnten diese endogenen Elemente einen Beitrag zur Modellierung und Arrangierung des Genoms leisten. Die Erhaltung der Plastizität des Wirtsgenoms stellt einen evolutionären Nutzen für die Wirtspopulation dar. Pathogene Effekte einzelner Individuen würden ausbalanciert werden (Blomberg *et al.*, 2007).

Hinweise auf einen möglichen Zusammenhang zwischen endogenen Retroviren und Autoimmunerkrankungen wie Diabetes mellitus (Conrad *et al.*, 1997), multiple Sklerose (Sugimoto *et al.*, 2001) oder bei systemischem Lupus erythematosus (Ogasawara *et al.*, 2001) werden kontrovers diskutiert.

Die fusiogenen Eigenschaften der retroviralen Env Proteine können vorteilhaft vom Wirt genutzt werden. Ein Beispiel ist das HERV-W Env Protein (Syncytin-1). Zusammen mit

dem Env Protein Syncytin-2 des endogenen Retrovirus HERV-FRD ist Syncytin-1 an der Formation von Syncytiotrophoblasten beteiligt (Blond *et al.*, 2000; Mallet *et al.*, 2004). Die zusätzliche immunsuppressive Wirkung des Syncytin-2 könnte ursächlich sein für die Toleranz der Mutter gegenüber dem Fötus während der Embryonalentwicklung (Blaise *et al.*, 2003; Mangeney *et al.*, 2007). In einigen humanen Brustkrebsgeweben und endometriischen Karzinomen konnte ebenfalls die Expression von Syncytin nachgewiesen werden. In diesen Fällen könnte eine Syncytin vermittelte Zellfusion das Tumorwachstum propagieren (Duelli und Lazebnik, 2003).

Die Expression von Superantigenen konnte sowohl für endogene als auch für exogene Retroviren gezeigt werden (Fleischer, 1991; Meylan *et al.*, 2005). Superantigene Verbinden die Rezeptorbindungsstellen von Klasse-II Haupthistokompatibilitäts-Molekülen auf Antigen-präsentierende Zellen mit bestimmten T-Zellrezeptoren CD4-positiver T-Lymphozyten, was zur Stimulation dieser T-Zellen führt. Die superantigene Wirkung resultiert in der Aktivierung von 10^3 bis 10^5 mal mehr T-Zellen, als die Stimulation durch konventionelle Antigene bewirkt (Sundberg *et al.*, 2002). Dadurch kann der Organismus oft auf weitere Antigene nicht reagieren (Nelson *et al.*, 2004). Unabhängig von der Antigenspezifität der T-Lymphozyten kommt es durch ein Superantigen zu einer unkoordinierten Hyperproliferation und zur massiven Freisetzung von Zytokinen, die eine starke Entzündungsreaktion auslösen und gegebenenfalls zu Autoimmunerkrankungen führen können. Unter anderem auf Grund dieser Erkenntnisse wird eine Beteiligung von HERVs an Autoimmunerkrankungen und bei der Entstehung von Tumoren diskutiert (Nakagawa und Harris, 1996; Conrad *et al.*, 1997; Löwer *et al.*, 1998; Posnet und Yarilina, 2001).

Im Wirtsgenom enthaltene ERV Integrate könnten dem Schutz vor Infektionen verwandter exogener Viren auf Grund von Rezeptorinterferenz dienen (Goff, 2004; Spencer *et al.*, 2003). Einige Mäusestämme exprimieren ein *env* Gen (Fv-4) dessen Genprodukt, die Zellen vor einer MLV-Infektion schützt. Das Protein bindet den MLV-Rezeptor und verhindert auf diese Weise die Infektion (Limjoco *et al.*, 1993). Die Expression des HERV-W Env Proteins vermittelt eine Resistenz gegen die Infektion mit dem *spleen necrosis virus* (SNV) (Ponferrada *et al.*, 2003). Das Fehlen exogener Pendanten zu den zahlreichen endogenen HERV Familien kann als Hinweis auf den erfolgreichen Schutz durch die endogenisierten Viren gedeutet werden (Löwer *et al.*, 1999).

Regulatorische Elemente der LTR können gewebsspezifisch aktive Enhancer, Promotor oder Polyadenylierungssignale für zelluläre Gene bereitstellen (de Parseval und Heidmann,

2005). Ein Drittel der im humanen Genom vorhandenen Bindestellen für den Transkriptionsfaktor und Tumorsuppressor p53 ist in ERV LTR Regionen lokalisiert. Diese LTRs sind Teil des p53 Regulationsnetzwerkes und damit indirekt an der Steuerung von p53 Zielgenen beteiligt (Wang *et al.*, 2007). Die Aktivierung von Protoonkogenen durch HERV Promotorsequenzen ist denkbar und könnte zu einer bösartigen Entartung der betroffenen Zellen beitragen.

Da die Familie HERV-K die einzige endogene Retrovirus-Familie ist, die noch intakte Leserahmen für alle retroviralen Proteine besitzt, scheint ihr eine bedeutende biologische Relevanz zuzukommen. Obwohl es im menschlichen Genom zum jetzigen Zeitpunkt keine infektiösen Proviren mehr gibt, konnten zwei Arbeitsgruppen solche künstlich rekonstruieren. Diese Konstrukte erzeugen lediglich Partikel mit schwacher Infektiosität (Dewannieux *et al.*, 2006; Lee *et al.*, 2007). Das könnte darauf hinweisen, dass die Endogenisierung tatsächlich zum Verschwinden der infektiösen exogenen Verwandten geführt hat.

In Krebspatienten induziert die Reaktivierung von HERV-K Proviren durch die damit verbundene Expression der viralen Proteine eine humorale und zelluläre Immunreaktion. Das konnte von verschiedenen Gruppen mit Hilfe von HERV-K Partikelpräparationen oder rekombinanten HERV-K Proteinen sowie synthetischen Peptiden festgestellt werden (z. B. Boller *et al.*, 1997; Denner *et al.*, 1995; Sauter *et al.*, 1995; Sauter *et al.*, 1996; Herbst *et al.*, 1998; Hahn *et al.*, 2008). Antikörperantworten gegen das HERV-K Gag Protein eignen sich als Surrogat-Parameter für Hodentumore. Bei über sechzig Prozent dieser Tumorpatienten kommt es zur Bildung von Antikörpern, in der Vergleichsgruppe der gesunden Menschen kommt es nur in drei Prozent der untersuchten Individuen zu einer Antikörperantwort (Sauter *et al.*, 1995). Einige Seminompatienten zeigen bereits mehr als sechs Monate vor der Diagnose des Tumors eine Antikörperantwort gegen HERV-K (Goeddert *et al.*, 1999). Diese Antikörperspiegel sinken nach erfolgreicher Therapie der Erkrankung wieder ab (Sauter *et al.*, 1996). Die prognostische Relevanz einer Antikörperantwort gegen HERV-K konnte ebenso für Melanompatienten gezeigt werden. Patienten mit Immunantwort haben signifikant reduzierte Überlebenschancen (Hahn *et al.*, 2008). Dies deutet darauf hin, dass die Expression von HERV-K Genen das Tumorgeschehen negativ beeinflusst, obwohl beim Menschen bislang kein eindeutiger Zusammenhang zwischen der Expression retroviraler Sequenzen und der Tumorentstehung nachgewiesen werden konnte. Im Mausmodell konnte ein Zusammenhang zwischen der Expression des HERV-K Rec Proteins und der Entstehung von Keimzelltumoren gezeigt

werden (Galli *et al.*, 2005). HERV-K Rec interagiert mit dem Stammzell-Regulator und Tumorsuppressor PLZF. Durch diese Bindung kann die PLZF bedingte transkriptionelle Repression von c-Myc aufgehoben werden. c-Myc wird verstärkt exprimiert und dies wiederum hat Auswirkungen auf die Expression weiterer Gene und führt phänotypisch zu erhöhter Proliferation und verminderter Apoptose (Denne *et al.*, 2007).

Die physiologische Bedeutung des NP9 Proteins ist noch unbekannt es wird aber vermutet, dass NP9 durch Interaktion mit dem Numb/Notch-Signalweg eine Rolle bei der malignen Entartung von Zellen spielen könnte (Armbruster *et al.*, 2004) oder ebenfalls mit dem Tumorsuppressor PLZF interagiert (Denne *et al.*, 2007). Die Proteine Rec und Np9 werden auch in Melanomen exprimiert (Büscher *et al.*, 2006). Kürzlich konnte gezeigt werden, dass ein Mangel an PLZF Protein in Melanomen zur Expression der microRNA-221/222 führt, die wiederum wichtige Tumorsuppressorproteine ausschaltet und so zur Tumorprogression beiträgt (Felicetti *et al.*, 2008).

1.4. Transkription

Alle Zellen, von Bakterien bis zu menschlichen Zellen exprimieren ihre genetische Information auf die gleiche Weise. Der Strom der genetischen Information in der Zelle verläuft von der DNA über RNA zum Protein (Crick, 1958). Dieser Mechanismus ist so grundlegend, dass er als das zentrale Dogma der Molekularbiologie bezeichnet wird. Durch die Entdeckung anderer Wege, die der Weitergabe genetischer Informationen dienen, musste das zentrale Dogma erweitert werden. Der Informationsfluss kann, wie bereits oben beschrieben, im Fall einer retroviralen Infektion von RNA zu DNA führen (Baltimore, 1970; Temin und Mizutani, 1970). Auch die Weitergabe von Informationen von Protein zu Protein (Prusiner *et al.*, 1982) durch Konformationsänderungen ist möglich. Als Transkription bezeichnet man die Produktion von RNA anhand einer DNA-Matrize. Dieser genauestens koordinierte Prozess wird durch die RNA-Polymerase vermittelt. Die enzymatische Aktivität der RNA-Polymerase wurde 1959 von Weiss und Gladstone in Kernextrakten aus Rattenleberzellen entdeckt. Die universelle Bedeutung der RNA Polymerase zeigte sich, als auch in *Escherichia coli* eine ähnliche RNA-Polymerasaktivität nachgewiesen werden konnte (Hurwitz *et al.*, 1960; Stevens, 1960; Chamberlin und Berg, 1962).

Heute sind in höheren Eukaryoten vier RNA Polymerasen bekannt. Die eukaryotischen RNA-Polymerasen I, II und III wurden in unterschiedlichen chromatographischen Fraktionen von Seeigelembryozellkernen identifiziert (Roeder und Rutter, 1969).

Roeder und Rutter entdeckten 1970, dass RNA-Polymerase I (Pol I) hauptsächlich verantwortlich ist für die Transkription der ribosomalen 18S- und 28S-RNAs. RNA-Polymerase II (Pol II) transkribiert zelluläre mRNA und RNA-Polymerase III (Pol III) transkribiert 5SrRNA und tRNA (Roeder und Ruttner, 1970; Gniazdowski *et al.*, 1970; Kedinger *et al.*, 1970; Weinmann und Roeder, 1974; Weinmann *et al.*, 1974). Eine vierte RNA-Polymerase wurde in Pflanzen identifiziert (Herr *et al.*, 2005; Kanno *et al.*, 2005; Onodera *et al.*, 2005).

Alle RNA-Polymerasen sind auf weitere Proteine angewiesen, um die Transkription gezielt am Promotor zu initiieren und auf diese Weise einen Transkriptionsinitiationskomplex zu bilden (Woychik, 1998). Daher wird die RNA-Polymerase durch ein Set von akzessorischen Faktoren ergänzt. Dieses Set bilden die generellen Transkriptionsfaktoren (GTFs) (Wasylyk, 1988; Zawel und Reinberg, 1993; Orphanides *et al.*, 1996). Die wichtigsten dieser Faktoren sind TFIIA, TFIIB, TFIID, TFII E, TFII F und TFII H. Zur Benennung der Faktoren bedient man sich folgender Nomenklatur: TF bezeichnet Transkriptionsfaktor, die römische Nummerierung II zeigt einen Pol II abhängigen Transkriptionsvorgang und durch den Buchstaben wird die chromatographische Fraktion angegeben, in welcher der Faktor vorliegt und entdeckt wurde. Das Proteinarrangement und die funktionalen Eigenschaften der GTFs sind unterschiedlich. Nur wenn alle GTFs (eine Ausnahme bildet TFIIA) gemeinsam vorliegen, kann die Transkription beginnen.

Die Assemblierung dieses so genannten Präinitiationskomplexes wurde ursprünglich als schrittweise ablaufender Prozess beschrieben (Orphanides *et al.*, 1996). TFIID bindet zuerst an die Promotorregion, beispielsweise an eine TATA-Box gefolgt von TFIIA und TFIIB, die die Bindung von TFIID stabilisieren. Anschließend werden Pol II/TFII F rekrutiert. Nach der Bildung eines stabilen TFIID-TFIIA-TFIIB-Pol II/TFII F-Promotorkomplexes wird TFII E rekrutiert und nachfolgend wird der Komplex von TFII H ergänzt (Greenblatt, 1997).

Neuere Ergebnisse deuten allerdings auf eine alternative Erklärungsmöglichkeit der Entstehung eines Präinitiationskomplexes in Form des Pol II-Holoenzymmodells (Kim *et al.*, 1994; Koleske und Young, 1994) hin. Die Zusammensetzung der isolierten Holoenzymkomplexe ist allerdings von der Art und den Bedingungen der Aufreinigung abhängig. Ob die eukaryotische RNA-Polymerase II als Holoenzymkomplex in der Zelle

vorliegt oder der Präinitiationskomplex schrittweise aufgebaut wird, blieb ungeklärt. Wahrscheinlich werden beide Wege zur Initiation der Transkription genutzt. Eine Schlüsselfunktion bei der Initiation des aktiven Transkriptionsvorganges fällt dem Transkriptionsfaktor TFIID zu. Auf Grund seiner Helikaseaktivität wird der Matrizenstrang für die Polymerase zugänglich. Die Pol II beginnt mit der Synthese kurzer RNA-Moleküle. Die Freisetzung der Polymerase aus dem Präinitiationskomplex erfolgt erst nach der Übertragung von Phosphatgruppen auf die C-terminale-Domäne des Enzyms. Diese Phosphorylierung wird ebenfalls von TFIID katalysiert. Die Polymerase kann sich aus dem Präinitiationskomplex lösen, erfährt weitere Konformationsänderungen, verbunden mit der Anlagerung verschiedener Faktoren, die die Elongation des RNA-Strangs begleiten, und vollzieht die Transkription des Gens. Transkription, die durch dieses minimale Set aus Proteinen initiiert wird, bezeichnet man als basale Transkription.

In Eukaryoten gibt es drei verschiedenartige, *cis*-regulierenden DNA-Sequenzelemente, die Einfluss auf die Transkriptionsinitiation durch die Pol II nehmen können: den *Core*-Promotor, die Promotor-proximalen-Elemente und die Promotor-distalen-Elemente, auch als Enhancer bezeichnet (Orphanides *et al.*, 1996). Die *cis*-regulierenden DNA-Sequenzelemente, welche nicht dem *Core*-Promotor angehören, werden oft auch als *upstream activating sequences* (UAS) bezeichnet.

1.4.1. *Core*-Promotor-Elemente

Durch Studien an Eukaryoten konnten bisher sieben *Core*- (oder Minimal-) Promotor-Elemente identifiziert werden (Abb.: 5) (Hahn *et al.*, 1989; Singer *et al.*, 1990; Wobbe und Struhl, 1990; Zhou und Chiang, 2000; Deng und Roberts, 2005). Sie zeichnen sich durch charakteristische Sequenzen aus. Durch *Core*-Promotor-Elemente werden die exakte Assemblierung, sowie die richtige Orientierung des Präinitiationskomplexes gewährleistet. Ein wichtiges *Core*-Promotor-Element ist die TATA-Box. Das Motiv ist in den meisten eukaryotischen Organismen konserviert (Hahn *et al.*, 1989; Singer *et al.*, 1990; Wobbe und Struhl, 1990). Die spezifische Sequenzabfolge der TATA-Box wird vom TATA-binding-protein (TBP), einer Untereinheit des TFIID, erkannt und gebunden (Hernandez, 1993; Burley und Roeder, 1996). Ein weiteres *Core*-Promotor-Element ist das Initiatorelement (Inr). Es enthält eine pyrimidinreiche Sequenz, die den Transkriptionsstart umgibt. Das Inr ist in der Lage, die Transkriptionsinitiation allein zu vermitteln oder zusammen mit einer TATA-Box oder anderen *Core*-Promotor-Elementen zu agieren. Die TAF1/TAF2-

Komponenten des TFIID sind an der Erkennung des Inr beteiligt (Abb.: 5). Das *downstream promotor element* (DPE) stellt das dritte *Core*-Promotor-Element dar (Zhou und Chiang, 2000; Kutach und Kadonaga, 2000). Man nimmt an, das TFIID das DPE erkennt und so die Bindung des Präinitiationskomplexes vermittelt (Shao *et al.*, 2005).

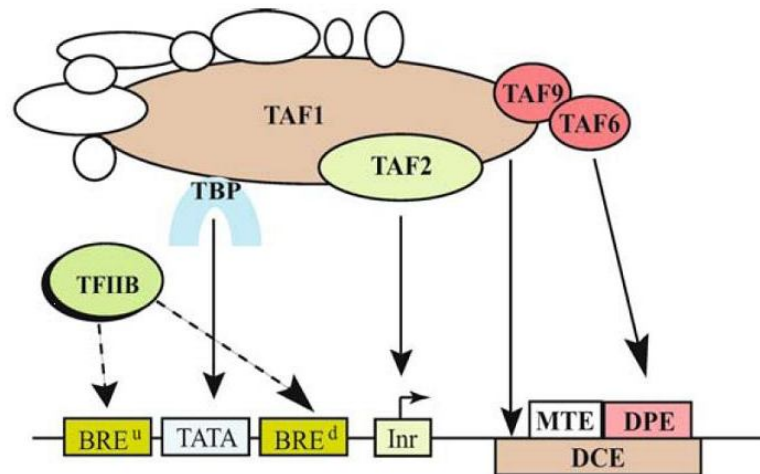


Abbildung 5: Erkennung der *Core*-Promotor-Elemente durch TFIID und TFIIIB. Gezeigt ist die Interaktion zwischen TFIID und TFIIIB mit den sieben bekannten *Core*-Promotor-Elementen, der TATA-Box (TATA), dem Initiatorelement (Inr), dem *upstream TFIIIB*-*recognition element* (BRE^u), sowie dem *downstream TFIIIB*-*recognition element* (BRE^d), dem *downstream promotor element* (DPE), dem *motif ten element* (MTE) und dem *downstream core element* (DCE). TFIID besteht aus dem TATA-Binding-Protein (TBP) und den TBP-assoziierten Faktoren (TAF) TAF1 bis TAF14. Die verschiedenen Komponenten sind dabei spezialisiert auf die Erkennung bestimmter *Core*-Promotor-Elemente. Das TBP bindet die TATA-Box, TAF1 und TAF2 sind involviert in die Erkennung des Inr. Der dimere TAF9/TAF6-Komplex ist in die Erkennung des DPE involviert und TAF1 ist in der Lage mit dem DCE zu interagieren. TFIIIB ist ein einzelnes Protein, welches die BRE-elemente binden kann. Abbildung aus Thomas und Chian 2006.

Neben dem DPE sind noch zwei weitere *Core*-Promotor-Elemente bekannt, die *downstream* des Transkriptionsstarts liegen, das MTE (*motif ten element*, Ohler *et al.*, 2002; Lim *et al.*, 2004) und das DCE (*downstream core element*, Lee *et al.*, 2005). Normalerweise fungiert das MTE in Verbindung mit dem Inr, um auf diese Weise die Transkriptionsrate zu erhöhen. Es kann allerdings auch die TATA-Box oder das DPE ersetzen oder synergistisch mit TATA-Box oder DPE in Abhängigkeit des Inr die Promotoraktivität erhöhen (Lim *et al.*, 2004). Die Faktoren, die die Transkriptionsinitiation am MTE vermitteln, sind noch unbekannt.

Im Gegensatz zur Sequenzkontinuität der bereits erwähnten *Core*-Promotor-Elemente besteht das DCE aus drei diskontinuierlichen Untereinheiten. Wie Promotoranalysen des humanen Genoms gezeigt haben, schließen sich DCE und DPE gegenseitig in ihrer

Anwesenheit aus. TAF1 als Komponente des TFIID ist in der Lage, jede der drei Untereinheiten zu binden.

Die beiden verbleibenden Promotorelemente werden von einem anderen GTF erkannt, dem TFIIB. Das stromaufwärts der TATA-Box gelegene Bindemotiv für TFIIB wird als TFIIB-*recognition element* (BRE^u) bezeichnet, das stromabwärts gelegene als BRE^d (Lagrange *et al.*, 1998; Tsai und Sigler, 2000; Deng und Roberts, 2005).

1.4.2. Transkriptionskontrolle und Initiation retroviraler Gene

Nachdem die Zelle von einem Retrovirus infiziert wurde und dieses seine Erbinformation in das Erbgut des Wirtes integriert hat, ist das Provirus in seiner Genexpression den zellulären Regulationsmechanismen unterworfen. Der Wirtsorganismus hat verschiedene Möglichkeiten die Expression des Virus zu unterdrücken.

Transposable Elemente in der genomischen DNA des Wirtes werden normalerweise methyliert und damit inaktiviert. Die Cytosin-Methylierung wird dabei auch als Verteidigungsmechanismus gegen intragenomische Parasiten bezeichnet. Durch weitere epigenetische Veränderungen wie Histonmodifikationen kann der DNA-Bereich, in welchem die Integration liegt, umgestaltet und durch eine dichte Chromatinverpackung für die Transkriptionsmaschinerie verschlossen werden.

Liegt ein unmethyliertes Provirus in einem offenen Chromatinbereich, so kann sich an der viralen LTR ein Transkriptionsinitiationskomplex ausbilden. Unmethylierte Proviren sind jedoch nicht zwingend transkriptionell aktiv. Dies kann auf das Fehlen notwendiger Transkriptionsfaktoren zurückzuführen sein und ist für HERV-K bereits gezeigt worden (Lavie *et al.*, 2005). Sowohl zellulärer Kontext als auch epigenetischer Status des Provirus bedingen die Transkription.

Wie jeder eukaryotische Promotor besteht auch die virale LTR aus basalen Strukturelementen und solchen, die zur Modulation der Transkriptionsaktivität beitragen. Wichtige regulatorische Bereiche sind ebenso wie der Core-Promotor im U3-Bereich der viralen LTR lokalisiert. Die Bindestelle für den Holoenzymeinkomplex einschließlich der RNA-Polymerase II und damit der Ort des Transkriptionsstarts definiert den Übergang von U3 zu R. Die gebildeten proviralen Transkripte werden ebenso wie die zelluläre RNA mit einer 5'Cap Struktur und einem Poly-A-Schwanz versehen (Furuichi und Shatkin, 2000; Shatkin und Manley, 2000; Wickens, 1990). In der 5'LTR des Provirus sind alle *cis*-

aktivierenden Sequenzen lokalisiert, die von *trans*-aktiven viralen oder zellulären Faktoren für die Steuerung der basalen und der induzierten viralen Genexpression benötigt werden. Viele exogene Retroviren wie beispielsweise das HI-Virus nutzen eine TATA-Box zur Initiation der Transkription (Jones und Peterlin, 1994). Die Interaktion sowohl mit zellulären als auch mit viralen Faktoren kann sich aktivierend oder reprimierend auf die HIV-Transkription auswirken. Der HIV *Core*-Promotor (1.4.1.) bestehend aus der TATA-Box und drei Bindungsstellen für den Transaktivator Sp1 (1.4.3.). Dieser Bereich ist für die Anlagerung des Präinitiationskomplexes und damit für die basale Transkription wesentlich. Zur Aktivierung der Transkription ist die Bindung T-Zell-spezifischer Transkriptionsaktivatoren wie NF- κ B, NF-AT und ETS an entsprechende *cis*-regulatorische Elemente der LTR erforderlich (Gaynor *et al.*, 1992; Holzmeister *et al.*, 1993; Verdin *et al.*, 1995; Al-Harhi *et al.*, 1998). Stromaufwärts des *Core*-Promotors findet sich ein Enhancerbereich mit zwei Bindestellen für den T-Zell-spezifischen aktivierenden Transkriptionsfaktor NF- κ B (Jones *et al.*, 1986; Nabel und Baltimore, 1987). Im 5' regulatorischen Bereich und stromabwärts des basalen Promotors im 5' untranslatierten Bereich (5'UTR) liegen regulatorische Elemente mit Bindestellen für weitere Transkriptionsfaktoren. Zusammen mit dem HIV-1-spezifischen TAR-Element, welches als RNA-Struktur vorliegt und daher nicht zum eigentlichen LTR-Promotor gehört, regulieren die LTR Elemente die basale wie auch die transaktivierte Transkription von HIV-1. Die Bindung des viralen transaktivierenden Tat-Proteins an die TAR-RNA steigert die Transkriptionsrate auf ein effektives Niveau und ist daher essentiell für eine maximale Aktivierung der viralen Genexpression.

Auch die Transkription von MMTV wird über eine TATA-Box initiiert, der Promotor unterliegt zudem einer sehr starken Induktion durch Steroidhormone (Truss *et al.*, 1992). Die Initiation der Transkription endogener Retroviren erfolgt meist über ein Initiatorelement in Zusammenspiel mit Transkriptionsaktivatoren. ERV9 ist ein humanes endogenes Retrovirus, das in vergleichsweise geringer Anzahl an proviralen Insertionen im menschlichen Genom vertreten ist. ERV9 ist hauptsächlich in undifferenzierten embryonalen Karzinomzellen NT2/D1 transkriptionell aktiv (Lania *et al.*, 1992). Die Transkriptionsinitiation wird wahrscheinlich über ein Inr und eine Sp1-, Sp3-Bindestelle vermittelt (La Mantia *et al.*, 1992; Strazzullo *et al.*, 1994). Das im viralen Genom enthaltene TATA-Motiv erfüllt hier lediglich regulatorische Aufgaben, ist aber nicht direkt an der Transkriptionsinitiation beteiligt. Wird das TATA-Motiv entfernt sinkt die Promotoraktivität signifikant aber nicht drastisch ab. Ebenso verhält es sich mit der 5' LTR

von HERV-W. Auch diese enthält eine TATA-Box die nicht direkt an der Initiation der Transkription beteiligt ist (Lee *et al.*, 2003).

Ferner wird der *Core*-Promotor des humanen endogenen Retrovirus H (HERV-H) durch ein Initiatorelement und eine Sp1-Bindestelle gebildet. Die Transkriptionsinitiation erfolgt hier völlig unabhängig von der ebenfalls vorhandenen TATA-Box (Sjøttem *et al.*, 1996). Diese Feststellungen deuten auf einen großen Einfluss der Sp-Proteine bei der Initiation der Transkription endogener Retroviren hin. Neben den ubiquitär vorkommenden Proteinen Sp1 und Sp3 müssen aber auch gewebsspezifische Transkriptionsfaktoren in der Regulation der Expression eine Rolle spielen, da viele HERV-Familien nur in spezifischen Zelltypen exprimiert werden.

1.4.3. Transkriptionsregulatoren

Eukaryotische Transkriptionsregulatoren zeigen typischerweise einen modularen Aufbau. Sie enthalten mindestens zwei Domänen: Eine DNA-Bindedomäne, durch welche der Transkriptionsregulator sequenzspezifisch an die DNA binden kann, und die Aktivierungsdomäne. Mit Hilfe der Aktivierungsdomäne können DNA-gebundene Transkriptionsaktivatoren direkt und/oder indirekt die Transkriptionsmaschinerie rekrutieren und so die Assemblierung eines Präinitiationskomplexes am Promotor unterstützen. Auch eine aktivatorvermittelte Modifikation der Chromatinstruktur, beispielsweise durch die Rekrutierung chromatinmodifizierender Moleküle, kann bei der Bildung des Präinitiationskomplexes von Bedeutung sein.

Manche Transkriptionsregulatoren erfüllen vielfältige Aufgaben und sind an der Expression und Regulation von Haushaltsgenen, gewebsspezifisch exprimierten Genen, viralen Genen und zellzyklusspezifischen Genen beteiligt. Zu dieser Gruppe gehören beispielsweise die Sp-Proteine (Bouwman und Philipsen, 2002; Philipsen und Suske, 1999; Suske, 1999). Sp-Proteine erkennen und binden GC reiche Promotorelemente. Die Sp-Proteinfamilie ist gekennzeichnet durch drei Cys2His2-Zinkfinger, die eine sequenzspezifische DNA-Bindung vermitteln (Suske, 1999). Sp1 wurde identifiziert als ein in HeLa Zellen exprimierter Kofaktor, der die Eigenschaft besitzt, *in vitro* die Expression des SV40-Promotors zu steigern. Der Faktor wurde entsprechend der Aufreinigungsprozedur über Sepharyl-Phosphozellulosesäulen Sp (*specificity protein*) benannt (Dyner und Tjian, 1983a; 1983b). Die Familie umfasst vier Mitglieder Sp1, Sp2, Sp3 und Sp4, die alle einen ähnlichen modularen Aufbau zeigen. Lediglich Sp1 und Sp3 werden in Säugetieren ubiquitär exprimiert und binden an die gleiche Konsensussequenz (GGGGCGGGG). Sp1

und Sp3 zeigen mehr als 90 % Sequenzhomologie (Abb.: 6) in ihrer DNA-Bindedomäne, erfüllen aber unterschiedliche Funktionen im Organismus, wie Versuche mit *Knockoutmäusen* gezeigt haben (Marin *et al.*, 1997; Bouwman *et al.*, 2000; Göllner *et al.*, 2001; Van Loo *et al.*, 2003). Die beiden Proteine Sp1 und Sp3 lassen sich in verschiedene funktionale Bereiche gliedern (Abb.: 6) (Kadonaga *et al.*, 1988; Pascal und Tjian, 1991; Courey *et al.*, 1989; Courey und Tjian, 1988).

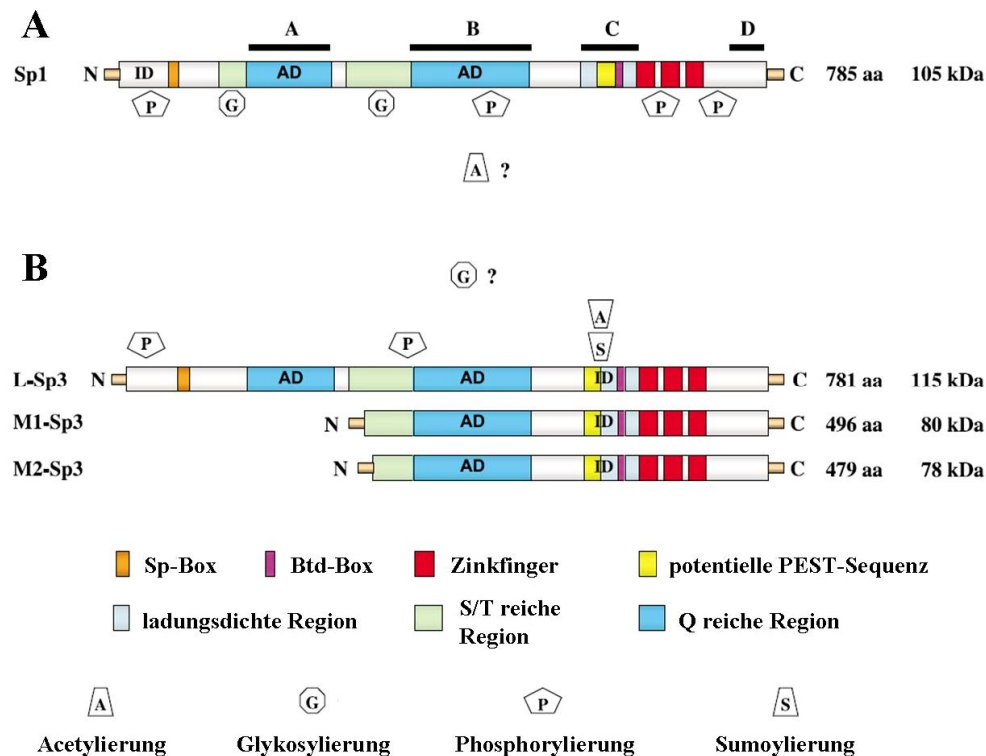


Abbildung 6: Strukturelle Domänen von Sp1 und Sp3. Die Anzahl der Aminosäuren und das Molekulargewicht sind auf der rechten Seite angegeben. Die Legende zur Farbgebung der Sp-Box, der Buttonhead (Btd) Box, der Zinkfingerdomäne, der potenziellen PEST-Sequenz sowie Bereiche hoher Ladungsdichte, (S/T)-reiche und (Q)-reiche Regionen findet sich am unteren Ende. Modifikationsstellen zur Acetylierung, Glykosylierung, Phosphorylierung und Sumoylierung sind ebenfalls gekennzeichnet. Acetylierungsstellen für Sp1 sind bisher nicht bekannt. **A)** Sp1. Die Balken A, B, C und D geben die Untereinheiten an. **B)** Sp3. Gezeigt sind die lange Isoform L und die beiden kurzen Isoformen (M1, M2). AD bezeichnet die Aktivierungsdomäne, ID dagegen die Inhibierungsdomäne. (Abbildung verändert nach Li *et al.*, 2004)

Die Transaktivierungsdomäne besteht aus zwei Untereinheiten (A und B). Jede Untereinheit kann die Transkription anregen, wenn sie mittels Bindedomäne an die DNA gebunden ist. Diese aktivierenden Regionen sind reich an Glutamin. Die Serin/Threoninreichen Domänen, die an die Aktivierungsdomänen grenzen, sind wahrscheinlich in post-translationalen Modifikationsmechanismen involviert (Roos *et al.*, 1997; Yang *et al.*, 2001). Die C-Domäne ist reich an stark geladenen Aminosäuren, das enthaltene *buttonhead element* ist bedeutsam für den synergistischen Aktivierungseffekt in

Zusammenhang mit dem *sterol regulatory element binding protein* (Athaniyar *et al.*, 1997; Yieh *et al.*, 1995; Harrison *et al.*, 2000). Der carboxyterminale Bereich der C-Domäne enthält die drei für die Familie charakteristischen Zinkfinger, die der Bindung der DNA dienen. Domäne D kann synergistisch mit den Aktivierungsdomänen A und B wirken (Pascal und Tjian, 1991). Der N-terminus der Sp-Proteine enthält eine konservierte Sp-Box. Im stark geladenen Teil der Proteine befindet sich eine putative PEST-Sequenz (Mortensen *et al.*, 1997). Diese Sequenz findet sich oft in Proteinen, die einem starken Umsatz unterliegen und somit eine kurze Halbwertszeit in der Zelle zeigen. Die PEST-Sequenz dient wahrscheinlich als Signalpeptid für die Degradation.

Die Lokalisation der Inhibierungsdomäne ist der größte Unterschied zwischen den Proteinen Sp1 und Sp3. Im Sp1 Protein ist sie N-terminal lokalisiert, während sie bei den Sp3 Isoformen unmittelbar vor den Zinkfingern liegt. Man nimmt an, dass dies der Grund für die unterschiedlichen Funktionen der beiden Proteine ist (Dennig *et al.*, 1996; Suske, 1999). Obwohl Sp1 und Sp3 die gleichen strukturellen Eigenschaften besitzen und an das gleiche DNA-Motiv binden, zeigen sie deutliche Unterschiede im Bindungsverhalten und in ihrer regulatorischen Funktion. Diese Funktion hängt vom Promotorkontext und dem zellulären Hintergrund ab (Majello *et al.*, 1997; Yu *et al.*, 2003). An vielen Promotoren wirkt Sp1 als starker Aktivator, während Sp3 meist inaktivierend oder nur leicht anregend wirkt. Die kurzen Isoformen von Sp3 resultieren nicht aus alternativen Spleißereignissen, sondern entstehen auf translationaler Ebene durch unterschiedliche Initiation.

1.4.4. Transkriptionsinitiation und Regulation der HERV-K LTR

In der HERV-K LTR konnten zwei putative *Core*-Promotorsequenzen mit typischen eukaryotischen Merkmalen (Strachan und Read, 1999) identifiziert werden, eine TATA-Box und ein Inr. Die Bestimmung des Transkriptionsstarts in Keimzelltumorklinen und in testikulärem Parenchym ergaben, dass die TATA-Box 200 Nukleotide stromaufwärts des Transkriptionsstarts lokalisiert ist, das Inr dagegen direkt am Transkriptionsstart liegt (M. Kraft, 2004; Kovalskaya *et al.*, 2005). In der Nähe des Inr liegen putative Bindestellen für die Transkriptionsfaktoren Sp1 und Sp3.

Unsere Arbeitsgruppe konnte zeigen, dass sich die Bindung des Transkriptionsfaktors YY1 (YingYang) an die Nukleotide 62-83 der HERV-K LTR positiv auf die transkriptionelle Aktivität auswirkt. Wird diese Bindestelle zerstört, sinkt die Aktivität der LTR im Reporterassay drastisch ab (Knössel *et al.*, 1999).

1.5. Ziele dieser Arbeit

Die Proteinexpression und Formation retroviraler Partikel des endogenen Retrovirus K konnte in Teratokarzinom- und Melanomzelllinien gezeigt werden. Ziel der vorliegenden Arbeit war es, die Expression von HERV-K in Melanomlinien im Vergleich zu Teratokarzinomzelllinien zu untersuchen.

Vorraussetzung für diese vergleichenden Analysen waren gut charakterisierte stark HERV-K produzierende Teratokarzinom- und Melanomzelllinien. Zu Beginn der Arbeit stand keine geeignete Melanomlinie zur Verfügung.

Daher sollte in einem ersten Schritt eine uniform HERV-K exprimierende Melanomlinie etabliert werden und in ihren Eigenschaften bezüglich der HERV-K Expression analysiert werden. Sowohl die Expression und Prozessierung der viralen Proteine, wie auch die Produktion und Morphologie der viralen Partikel, sollte beschrieben werden. In einem parallelen Ansatz sollten die reaktivierten, transkriptionel aktiven HERV-K Proviren bestimmt werden.

Des Weiteren sollte untersucht werden, ob die transkriptionelle Aktivität der HERV-K Proviren in Melanomen anderen Regulationsmechanismen unterliegt als in den vergleichsweise gut untersuchten Keimzelltumoren. Hierzu sollte zunächst die Aktivität, des proviralen Promotors vergleichend in Melanomen und Teratokarzinomen in einem Luziferaseassay untersucht werden. In folgenden Schritten sollten der Transkriptionsstart und die Polyadenylierungsstelle der HERV-K Transkripte bestimmt werden. Weitere Analysen zur Transkriptionsinitiation am *Core*-Promotor sollten folgen. Im Zuge dessen sollte die Funktionalität des in der viralen LTR lokalisierten TATA-Motivs untersucht werden. Durch weitergehende Analysen sollten Hinweise auf den zur basalen Transkriptionsinitiation benutzen *Core*-Promotor gewonnen werden.

2. ERGEBNISSE

2.1. Analyse der HERV-K Expression in der Melanomlinie UKRV Mel 2-C9

2.1.1. Etablierung einer Melanomlinie mit starker HERV-K Expression

Bei den meisten Keimzelltumorkulturen ist in fast allen Zellen der Kultur ein konstant hohes Expressionsniveau der HERV-K Proteine zu beobachten. Dagegen zeigten drei Melanomzelllinien, die bereits als HERV-K positiv publiziert wurden (Muster *et al.*, 2003), in unserem Labor häufig keine oder nur schwache Signale. Dies weist daraufhin, dass die HERV-K Expression in Melanomen fluktuieren kann. Ziel war deshalb, eine Linie mit gleichmäßig hohem Expressionsniveau zu etablieren. Deshalb wurden weitere Melanomlinien, die von Prof. Dirk Schadendorf, Heidelberg, zur Verfügung gestellt wurden, auf ihre HERV-K Expression untersucht (siehe Material und Methoden 4.2.). Als Indikator wurde das akzessorische Protein Rec gewählt, das von Typ 2 Vollängenproviren kodiert wird. Da Rec in seiner Funktion als Shuttle-Protein sowohl die ungespleißten als auch die einfachgespleißten *env* Transkripte aus dem Kern exportiert, ist ohne Rec keine Translation viraler Proteine zu erwarten. Rec ist überwiegend in den Nukleoli lokalisiert.

Die HERV-K Rec Expression wurde mittels Immunfluoreszenz (siehe 4.8.3.) nachgewiesen (Abb.: 7). Dazu wurden Zellen der zu untersuchenden Linien auf Deckgläschen ausgebracht und am folgenden Tag mit Paraformaldehyd fixiert, permeabilisiert und mit einem anti-Rec Antikörper inkubiert. Über einen FITC – gekoppelten Zweitantikörper wurde das Rec Protein in den Nukleoli visualisiert (grüne Fluoreszenz). Um die Kompartimentierung der Zelle zu verdeutlichen, wurden die Zellkerne mit Dapi gefärbt (blaue Fluoreszenz). Lediglich die Linie UKRV Mel 2 zeigte in ungefähr fünf Prozent der Zellen deutliche Rec Signale in den Nukleoli, die Linien Ma Mel 11, WM 98.2, UKRV Mel 14a, Ma Mel 21 und HCl Mel 19 waren negativ (Abb.: 7). Zum Nachweis der Signalspezifität wurden Isotyp-Kontrollen durchgeführt, das Signal ließ sich mit rekombinantem Rec Protein erfolgreich kompetitieren (Daten nicht gezeigt).

Da nur fünf Prozent der Zellen in der Melanomlinie UKRV Mel 2 eine deutliche HERV-K Expression zeigten, wurden durch Einzelzellklonierung subklonale Linien erzeugt (siehe 4.6.2.) und ebenfalls in der Immunfluoreszenz auf Expression des Rec Proteins hin untersucht. Die subklonale Linie UKRV Mel 2-C9 zeigte nach einem Zeitraum von zwölf

Wochen eine stabile starke HERV-K Rec Expression in über neunzig Prozent der Zellen. Dieses Expressionsniveau ist vergleichbar mit dem der in dieser Arbeit eingesetzten Keimzelltumorlinien GH, 2102EP und NCCIT. Die Linie UKRV Mel 2-C9 wurde für alle nachfolgenden Experimente verwendet. Die Linie Ma Mel 19 diente als Negativkontrolle.

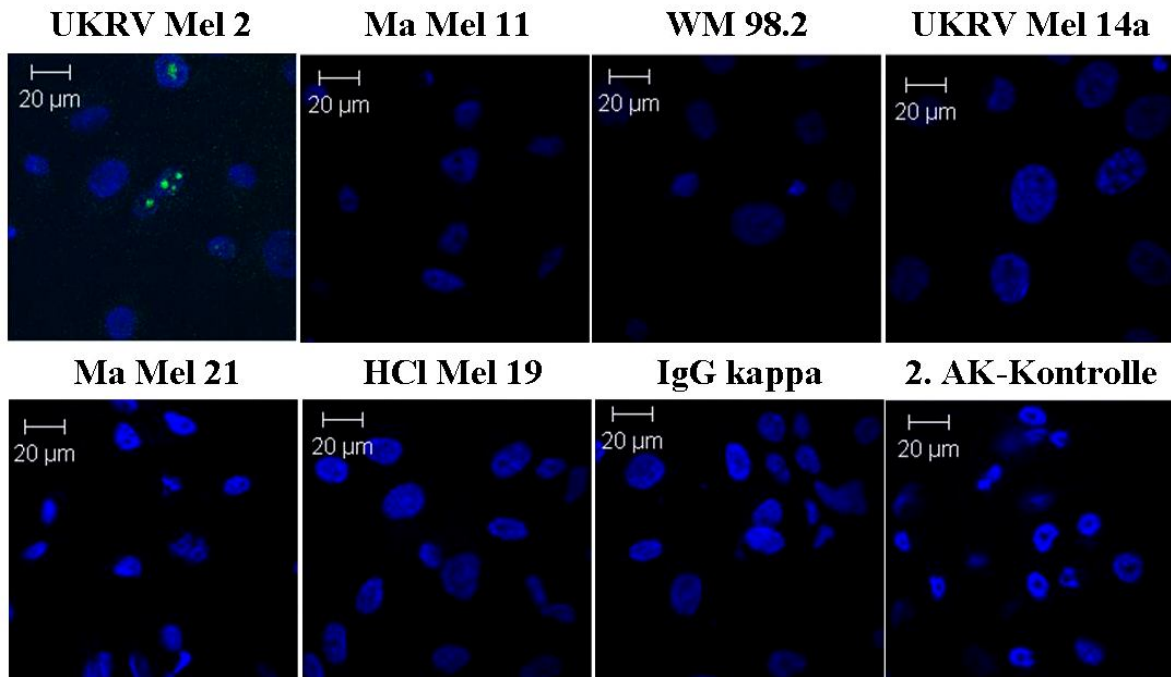


Abbildung 7: Immunfluoreszenz zum Nachweis von HERV-K Rec in verschiedenen Melanomzelllinien. Die Detektion des Proteins erfolgte unter Verwendung des Primärantikörpers anti-HERV-K Rec aus Kaninchen (1:500 verdünnt) und des FITC gekoppelten Sekundärantikörpers anti-Kaninchen aus Esel (1:300 verdünnt). Um die subzelluläre Lokalisation des Proteins zu verdeutlichen, wurden die Kerne mit DAPI gefärbt. Nur in 5 % der Zellen der Linie UKRV Mel 2 konnte Rec in den Nukleoli detektiert werden.

2.1.2. Nachweis der HERV-K Proteine Gag, Env und Rec mit spezifischen Antikörpern in der Immunfluoreszenz und im Immunoblot

Da das virale RNA Export-Adaptorprotein Rec in der Immunfluoreszenz nachgewiesen werden konnte, ist die essentielle Voraussetzung für den Export der einfach gespleißten HERV-K *env* sowie der ungespleißten HERV-K Vollängentranskripte gewährleistet. Als nächstes wurde deshalb untersucht, ob in der Melanomlinie UKRV Mel 2-C9 neben Rec weitere HERV-K Proteine gebildet werden. Als Vergleich dienten die HERV-K

exprimierende Keimzelltumormlinie GH sowie die nicht HERV-K exprimierenden Linien HeLa (Zervixkarzinom) und HCl Mel 19 (Melanom).

Das virale HERV-K Gag-Protein wurde zunächst durch Immunfluoreszenz mit einem spezifischen monoklonalen anti-HERV-K Gag Antikörper nachgewiesen. Die Analyse wurde an mit Paraformaldehyd fixierten und dann permeabilisierten Zellen durchgeführt. Typischerweise lagern sich funktionale retrovirale Gag-Proteine zu Virus-ähnlichen Partikeln zusammen. Diese heben sich im konfokalen mikroskopischen Bild als Cluster von Punkten mit sehr hoher Signalintensität an der Zelloberfläche hervor. Eine charakteristische Gag Färbung zeigte sich sowohl in der Melanomlinie UKRV Mel 2-C9 als auch, deutlich intensiver, in der Positivkontrolle GH (Abb.: 8, grüne Fluoreszenz). Die Isotypkontrolle und die Zweitantikörperkontrolle (Daten nicht gezeigt) ergaben keine Signale, ebenso wie die HERV-K negative Melanomlinie HCl Mel 19.

Wie bereits im vorherigen Kapitel beschrieben, zeigte UKRV Mel 2-C9 in der Immunfluoreszenz eine sehr starke Expression des HERV-K Rec-Proteins in über neunzig Prozent der Zellen (Abb.: 8, rote Fluoreszenz), ebenso wie die Linie GH, während HCl Mel 19 Rec negativ war.

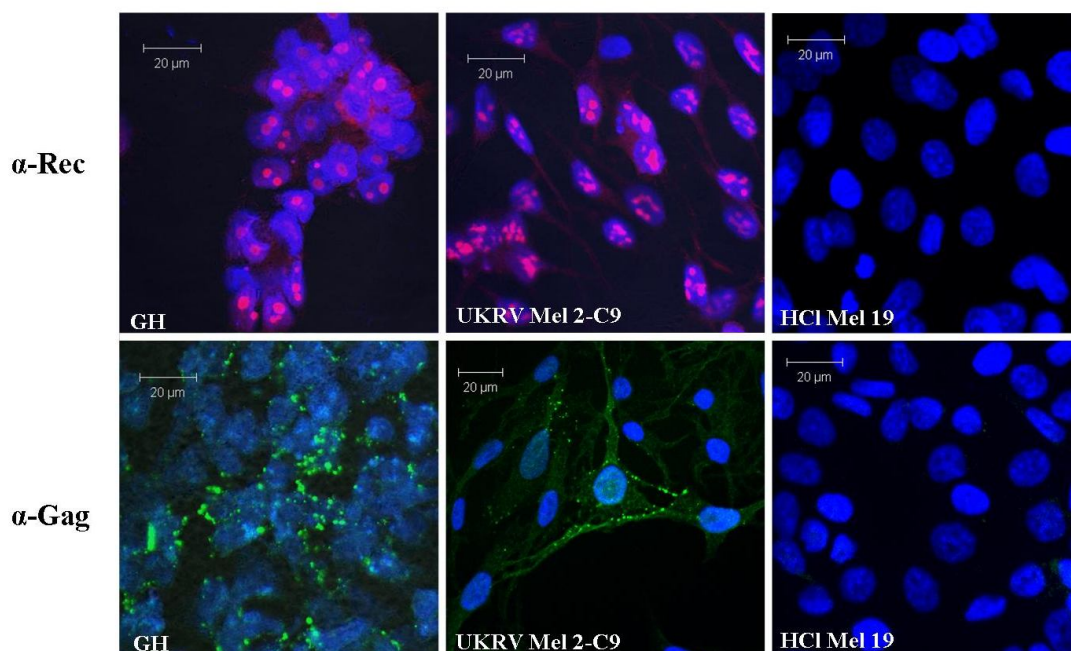


Abbildung 8: Immunfluoreszenzanalysen zur HERV-K Proteinexpression. Nachweis des viralen Proteins Rec (rot) durch den Primärantikörper Kaninchen anti-Rec (1:500 verdünnt) und den Rhodamin gekoppelten Sekundärantikörper anti-Kaninchen aus Esel (1:800). Nachweis des Gag Proteins (grün) durch den monoklonalen anti-Gag Antikörper HERMA 4 (Hybridomaüberstand) und den FITC gekoppelten Sekundärantikörper anti-Maus aus Esel (1:300 verdünnt) in der Keimzelltumormlinie GH, der HERV-K positiven Melanomlinie UKRV Mel 2-C9 und der HERV-K negativen Melanomlinie HCl Mel 19. Zellkernfärbung mit Dapi (blau).

Ferner wurde die Expression von HERV-K Gag und Env Proteinen im Gesamtproteinextrakt von UKRV Mel 2-C9 Zellen mittels Immunoblot analysiert. Extrakte der Linie GH dienten dabei als Positivkontrolle, die Linien HeLa und HCl Mel 19 wurden als Negativkontrolle mitgeführt. Die zu untersuchenden Linien wurden in T75-Zellkulturflaschen bis zu einer Konfluenz von ~75 % expandiert, mit PBS⁻ gewaschen und anschließend in Tripple-Lysis-Buffer (4.7.1.) mit Proteinaseinhibitor aufgenommen (4.7.1.). Um genomische DNA aus den Extrakten zu entfernen, wurden die Proben mit Ultraschall behandelt. Die Nukleinsäuren wurden durch Zentrifugation pelletiert. Der proteinhaltige Überstand wurde zur Konzentrationsbestimmung mittels Bradford-Proteinassay (4.7.3) eingesetzt. Für die Immunoblotanalysen wurden gleiche Mengen an Protein aus den Extrakten auf ein 12 % SDS-Polyacrylamidgel geladen und elektrophoretisch aufgetrennt. (4.7.4.) Anschließend wurden die Proteine aus dem Gel auf eine Nitrozellulosemembran übertragen und die Immunreaktionen durchgeführt (4.8.1.). Der Nachweis von β -Aktin diente als Ladekontrolle, um zu überprüfen, dass äquivalente Proteinmengen eingesetzt und auf die Membran übertragen worden waren (Abb.: 9A und 9B). Nach der Immunreaktion mit einem monoklonalen, spezifischen anti-HERV-K Env Antikörper konnte in den Proteinextrakten von GH und UKRV Mel 2-C9 sowohl das Env Vorläuferprotein (90 kDa) als auch die prozessierte Surface-Untereinheit (SU, 60 kDa) detektiert werden (Abb.: 9A). Die unspezifische Bande bei ~34 kDa kann ebenso wie Aktin zur Abschätzung der transferierten Proteinmenge dienen. Nach Vergleich der spezifischen Banden und der Aktinbande fällt auf, dass im Vergleich zu den anderen Proben eine größere Menge GH-Extrakt geladen und transferiert wurde. Daraus ergibt sich die Schlussfolgerung, dass das Expressionsniveau von HERV-K Env scheinbar in der Linie UKRV Mel 2-C9 und GH gleich hoch ist. Die HERV-K negativen Linien HeLa und HCl Mel 19 zeigten keine Signale.

Ebenso konnten mit einem monoklonalen anti-HERV-K Gag spezifischen Antikörper in GH und UKRV Mel 2-C9 Zellextrakten Signale für das Gag Vorläuferprotein (80 kDa) und für die vollständig prozessierte Untereinheit Kapsid (30 kDa) nachgewiesen werden (Abb.: 9B). Zusätzlich wurden Intermediate unvollständig prozessierter Gag Proteine detektiert. Da die Gag Untereinheiten während der Partikelreifung von der viralen Protease aus dem Vorläuferprotein freigesetzt werden, weisen die Intermediate und das Kapsid darauf hin, dass mit dem verwendeten Lysepuffer auch Viruspartikel aufgeschlossen wurden, die im Knospungsstadium noch an den Zelloberflächen anheften. Die andersgearteten Bandenmuster in den Linien GH und UKRV Mel 2-C9 könnten auf die

Aktivität unterschiedlicher Proviren zurück zu führen sein (siehe auch 2.3.), deren Gag unterschiedlich prozessiert wird. GH Zellen produzieren wesentlich mehr Gag Proteine als UKRV Mel 2-C9 Zellen. Die Linie HCl Mel 19 zeigte erwartungsgemäß keine HERV-K Gag Expression.

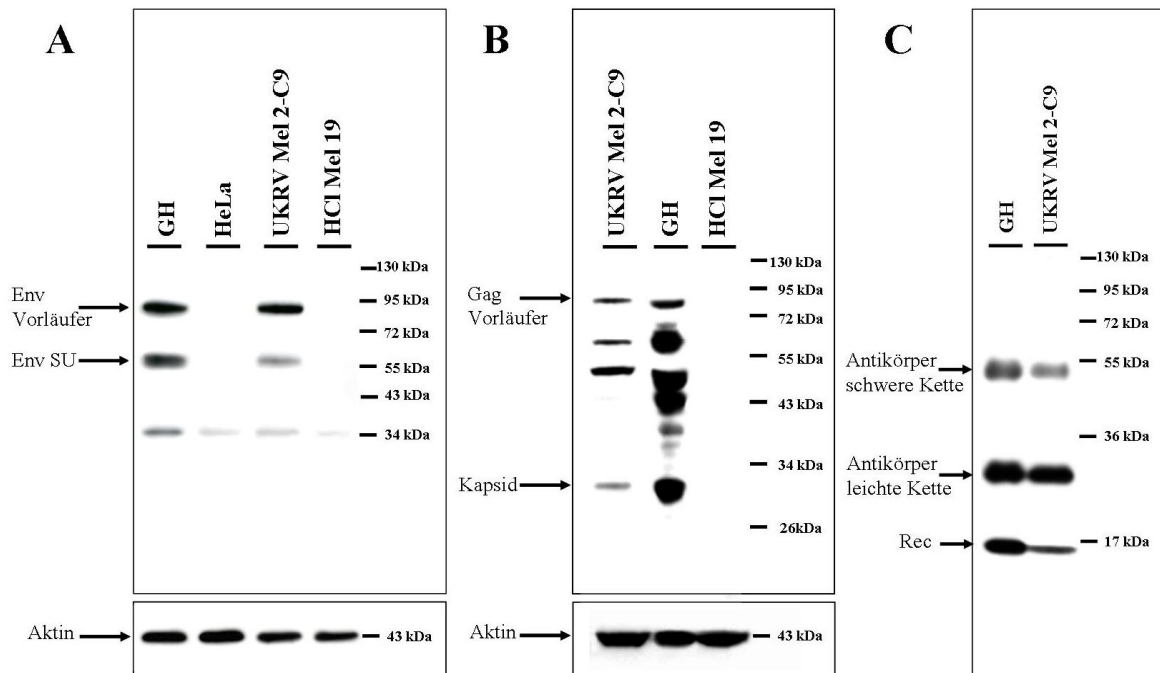


Abbildung 9: Immunoblotanalyse der Zellextrakte von UKRV Mel 2-C9, GH, HeLa und HCl Mel 19.

Zum Nachweis von HERV-K Env (A) und Gag (B) wurden jeweils 20 µg Gesamtproteinextrakt der jeweiligen Zelllinie über ein 12 %iges SDS-Gel separiert und auf die Nitrozellulosemembran übertragen. Die gelelektrophoretische Auftrennung zum Nachweis von Rec (C) erfolgte über ein 4-12 %ige Bis-Tris-Gel. A) Immunoblot belegt mit Primärantikörper Kaninchen-anti-Env 8655 in einer Verdünnung von 1:4000 und dem peroxidasegekoppelten Sekundärantikörper anti-Kaninchen aus Ziege in einer Verdünnung von 1:10000. Das Env-Vorläuferprotein konnte mit einem Molekulargewicht von ~90 kDa detektiert werden und die prozessierte Surface-Untereinheit mit ~60 kDa. B) Immunoblot belegt mit Primärantikörper Maus-anti-HERV-K Gag HERMA 7 (unverdünnt) und dem peroxidasegekoppelten Sekundärantikörper anti-Maus aus Ziege, 1:10000. Es wurden Signale für das Gag-Vorläuferprotein mit einem Molekulargewicht von 80 kDa detektiert, sowie für das prozessierte Kapsid (~30 kDa). C) Der Nachweis von Rec erfolgte über eine Immunopräzipitation mit einem Kaninchen anti-Rec-Antikörper. Die nachfolgende Immunreaktion wurde mit einem Maus anti-Rec-Antikörper, in einer Verdünnung von 1:3000 durchgeführt. Als Sekundärantikörper wurde ein peroxidasegekoppelter anti-Maus-Antikörper aus Ziege in einer Verdünnung von 1:10000 verwendet, in die Präzipitation wurde das Äquivalent des Gesamtproteinextrakt von circa $1,6 \times 10^6$ Zellen eingesetzt.

Da der Nachweis des sehr kleinen Rec Proteins (14 kDa) aus Gesamtzellextrakt schwierig ist, wurde vor dem Immunoblot eine Immunopräzipitation (4.8.2.) durchgeführt. Das Rec Protein wurde mit 5 µl des polyklonalen Kaninchen anti-HERV-K Rec Antikörpers und Protein-G-Sepharose aus Gesamtproteinextrakt präzipitiert. Das gewonnene Präzipitat wurde elektrophoretisch aufgetrennt und auf eine Membran übertragen. Der Nachweis erfolgte mit einem monoklonalen anti-Rec Antikörper. Neben Signalen der leichten und

schweren Kette des Antikörpers lassen sich sowohl im Präzipitat aus UKRV Mel 2-C9 als auch im Präzipitat aus GH Zellen deutliche Rec Signale detektieren (Abb.: 9C). Ein quantitativer Vergleich ist hier nicht möglich. Wurde anstelle des anti-Rec Antikörpers das korrespondierende Präimmunserum eingesetzt, so konnte nach der Immunreaktion kein Rec Protein detektiert werden (Kontrolle nicht gezeigt).

2.2. Vergleich transkriptionell aktiver Proviren in der Melanomlinie UKRV Mel 2 und Keimzelltumorklinien

2.2.1. Analyse möglicher HERV-K Polymorphismen

Da das Bandenmuster an unvollständig prozessierten Gag Proteinen in den Linien UKRV Mel 2-C9 und GH unterschiedlich war, lag der Verdacht nahe, dass in beiden Zelllinien verschiedene Proviren transkribiert werden. Zunächst wurde mit einer PCR Strategie (Abb.: 10) analysiert, ob sich die in der Einleitung beschriebenen polymorphen Proviren HERV-K 113 und HERV-K 115, die noch nicht im humanen Genom fixiert sind, in den hier untersuchten Zelllinien nachweisen lassen.

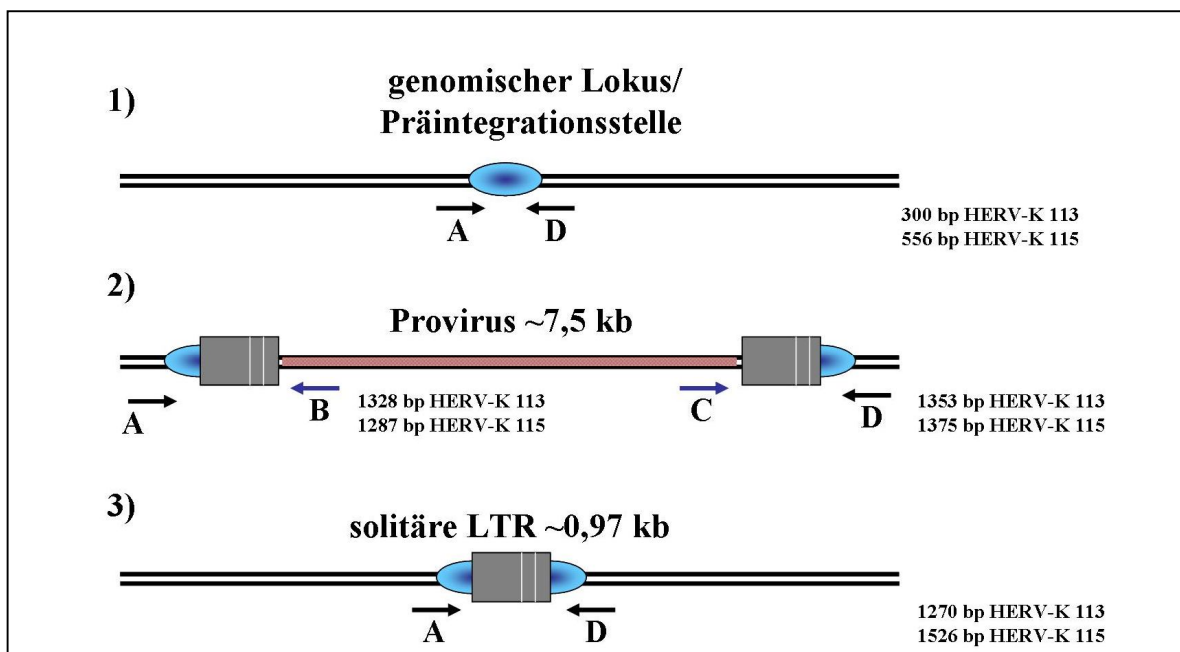


Abbildung 10: PCR-Strategie zur Detektion von HERV-K 113 und HERV-K 115 Proviren oder entsprechender Präintegrationsstellen im Genom. Lokalisation und Orientierung der verwendeten Primerpaare sind dargestellt, ebenso die Größe der erwarteten Amplifikate. **1)** Präintegrationsort ohne provirale Insertion.. **2)** Insertion eines Vollängenprovirus mit einem kodierenden Bereich von ~7,5 kb und **3)** Insertion einer solitären LTR mit einer Größe von ~ 0,97 kb.

Für diese PCR Analyse wurden Primerpaare (Abb.: 10, 2) verwendet, die im Provirus (B, C) und den jeweiligen flankierenden genomischen Bereichen (A, D) binden. Unter Verwendung des Primerpaares A und D kann der entsprechende genomische Präintegrationsort amplifiziert werden. Für das Amplifikat des Präintegrationsortes von HERV-K 115 (Chromosom 8p23.1) ergibt sich eine Größe von 300 bp, für den Präintegrationsort des Provirus HERV-K 113 (Chromosom 19p13.11) dagegen eine Größe von 556 bp. Sollte an der untersuchten Stelle durch Rekombination zwischen den viralen LTRs lediglich eine solitäre LTR vorhanden sein, ergibt sich ein um ~970 bp größeres Produkt. Liegt jedoch eine provirale Insertion vor, wird unter den gewählten PCR-Bedingungen mit den Primern A und D kein Produkt amplifiziert.

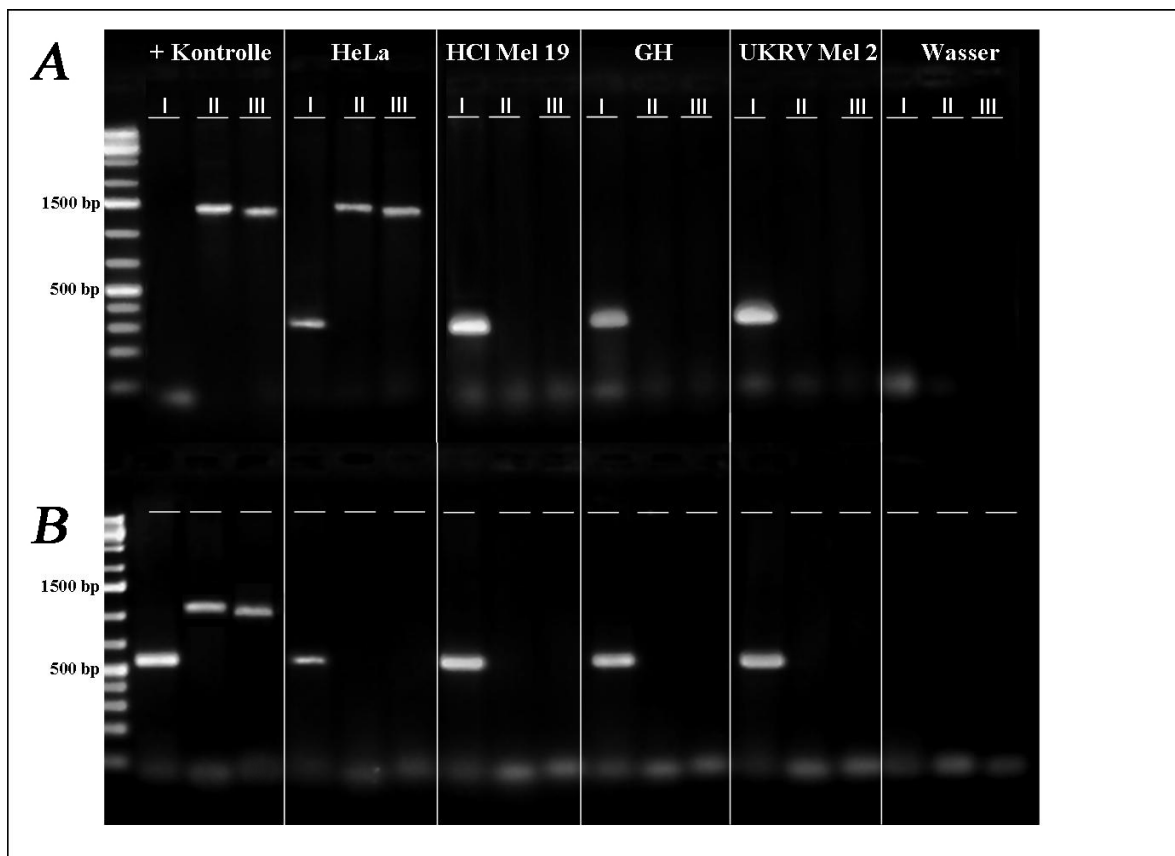


Abbildung 11: Untersuchungen zur Anwesenheit der polymorphen Proviren HERV-K 115 und HERV-K 113 **A)** PCR-Ergebnisse nach PCR mit einem Primerset zur Detektion von HERV-K 115 aus genomischer DNA. **B)** Amplifikate bei Verwendung eines spezifischen HERV-K 113-Primersets, DNA-Größenstandard (links), Positivkontrolle genomische DNA eines HERV-K 113 oder 115 positiv getesteten menschlichen Individuums. Detektion der Präintegrationsstelle in der Positivkontrolle zeigt Heterozygotie. **I)** PCR-Produkt der Primer A und D zur Amplifikation des Präintegrationsorts **II)** PCR-Produkt unter Verwendung der Primer C und D zur Amplifikation eines Bereichs aus der genomischen DNA im Übergang ins provirale Genom **III)** Amplikon bei Verwendung des Primerpaares A und B. Amplifikation eines Bereichs aus der genomischen DNA in die virale LTR.

Dafür ist die Amplifikation von Produkten aus dem Provirus in die flankierende Sequenz mit den Primerpaaren A und B beziehungsweise C und D möglich. Die Amplifikate werden über ein 1,5 %iges Agarosegel (4.4.3.3.) elektrophoretisch aufgetrennt.

Treten sowohl positive Ergebnisse bei der Detektion des Präintegrationsortes als auch für die Insertion auf, so ist das Individuum heterozygot für das untersuchte Provirus. Die PCR Analyse (4.4.3.6.) mit isolierter genomischer DNA (4.4.3.1.) der verwendeten Zelllinien ergab, dass die Linie UKRV Mel 2-C9 ebenso wie die Linien GH und HCl Mel 19 keines der beiden polymorphen Proviren enthält (Abb.: 11). Die Linie HeLa ist heterozygot für HERV-K 115.

2.2.2. Identifikation aktiv transkribierter HERV-K Proviren in der Melanomlinie UKRV Mel 2-C9, und den Keimzelltumorklinen NCCIT und GH

Die folgenden Ergebnisse wurden zum Teil in Zusammenarbeit mit der Praktikantin Jeanette Meister erarbeitet. Um zu bestimmen, welche HERV-K Proviren in den Zelllinien UKRV Mel 2-C9 und GH transkriptionell aktiv sind, wurde polyadenylierte RNA (4.4.2.2.) aus zuvor isolierter zytoplasmatischer RNA (4.4.2.1.) aufgereinigt (Daten nicht gezeigt). Diese aufwendige RNA-Extraktions-Methode wurde gewählt, um unreife Transkripte ausschließen zu können. Unvollständig prozessierte *mRNA* wird nicht zur Proteinsynthese aus dem Kern ins Zytoplasma transportiert, sondern einem Abbauprozess zugeführt. Als weitere HERV-K positive Probe wurde die Keimzelltumorklinie NCCIT parallel mitgeführt. Aus den aufgereinigten RNA Präparationen wurde mit Hilfe von Oligo(dT)-Primern cDNA synthetisiert (4.4.2.5.). Mittels PCR wurden aus den jeweiligen cDNAs Bereiche aus den HERV-K Genen *gag* (2190 bp), *gag-prt* (2900 bp) und mindestens 560 bp aus *env* amplifiziert (4.4.3.6. und 8.2.). Die Amplifikate wurden in 1,5 % Agarosegelen elektrophoretisch von den Primern abgetrennt (Daten nicht gezeigt). Banden der korrekten Größe wurden ausgeschnitten, aus dem Gel extrahiert (4.4.3.4) und in den Vektor pGEM-Teasy ligiert (4.4.3.5.) (Daten nicht gezeigt). Mit dem Ligationsansatz wurden kompetente *E. Coli* DH5 α -Bakterienzellen transformiert. Transformierte Klone wurden durch die plasmidvermittelte Ampicillinresistenz selektioniert (4.3.6.). Von jedem PCR-Amplifikat wurden mehrere einzelne Klone vermehrt. Die transgene Plasmid-DNA wurde isoliert (4.3.7.) und von der Firma MWG-Eurofins sequenziert.

Durch vergleichende Sequenzanalysen wurden die Klone proviralen HERV-K Loci zugeordnet. Dazu wurde ein Abgleich mit der Datenbank für das Humangenom-Projekt am *National Center for Biotechnology Information* (NCBI, *human genome assembly*) unter Verwendung des Analyseprogramms BLAT und durch Vergleiche mit bekannten HERV-K Loci unter Verwendung des Analyseprogramms Vektor-NTI durchgeführt. Entscheidend für die Zuordnung der gefundenen Sequenzen zu proviralen HERV-K Loci waren die charakteristischen Nukleotidunterschiede zwischen den Proviren. Die Sequenzen der untersuchten PCR-Klone unterschieden sich zusätzlich vereinzelt in einigen Nukleotiden. Diese Abweichungen stammen wahrscheinlich aus Fehlern, die während der cDNA-Synthese oder der PCR-Amplifikation auftraten und wurden für die Analyse vernachlässigt. Die in den Zelllinien aktivierten HERV-K Proviren sind in Tabelle 1 aufgelistet. Die Sequenzen finden sich im Anhang (8.2.).

In der Linie UKRV Mel 2-C9 konnten aus den 16 sequenzierten Klonen nur Transkripte des Provirus HERV-K 108, ID (GenBank) AC072054, nachgewiesen werden. Dieses Provirus ist auf Chromosom 7p22.2 lokalisiert und gehört zu den wenigen Vollängenproviren (Typ2), die offene Leserahmen (ORF) für alle HERV-K Gene tragen. Es liegt normalerweise als Dublette vor, wobei die 3'LTR des ersten zugleich die 5' LTR des zweiten Provirus ist.

Wie erwartet zeigten sich Unterschiede im Muster der exprimierten HERV-K Proviren zwischen der Melanomlinie UKRV Mel 2-C9 und den Keimzelltumorkolonien. In den parallel mitgeführten Keimzelltumorkolonien NCCIT und GH wurden neben dem in beiden Linien ebenfalls aktiven HERV-K 108 weitere transkriptionell aktive Proviren identifiziert. In den 21 sequenzierten Klonen der Linie NCCIT konnten am häufigsten Transkripte des Lokus HERV-K 102, ID (GenBank) AC044819, nachgewiesen werden. HERV-K 102 ist ein Typ1 Provirus mit der charakteristischen 292 bp Deletion im Bereich des *rec/env* Gens. und könnte einen offenen Leserahmen im *gag* besitzen. Die Eintragung der NCBI-Datenbank zeigt für HERV-K 102 ein vorzeitiges Stop-Codon im Gag Leserahmen. Doch ist bereits bekannt, dass in einigen Individuen ein intakter Leserahmen vorkommen kann (Ruprecht *et al.*, 2008). Weiterhin wurden mit niedriger Frequenz Transkripte der Typ1 Loci HERV-K 10 (auch als HERV-K 107 bezeichnet), ID (GenBank) AC016577; HERV-K 33 ID, (GenBank) AL392086 und des Typ 2 Lokus HERV-K 68, (auch als HERV-K 106 bezeichnet) ID (GenBank) AC078785, detektiert.

In den 14 sequenzierten Klonen der Linie GH konnten neben Transkripten von HERV-K 108 mit niedriger Frequenz auch die Typ 1 Loci HERV-K 10 und HERV-K II, ID

(GenBank) AC084198, als transkriptionell aktiv nachgewiesen werden. Am häufigsten wurden jedoch Transkripte des Typ 1 Locus HERV-K 101, ID (GenBank) AC007326, mit offenem Gag Leserahmen detektiert.

Tabelle 1: Transkriptionell aktive HERV-K Proviren aus UKRV Mel 2-C9, GH und NCCIT. Die durch Sequenzierung von cDNA-Klonen identifizierten HERV-K Loci sind angegeben mit bibliographischen Namen und GenBank-Identifikationsnummern. Typ 1 oder 2 bezeichnet die Anwesenheit (Typ 2) oder die Abwesenheit (Typ 1) einer 292 bp langen Sequenz am Übergang von *pol* zu *env*. Vorhandene offene Leserahmen für Gag, Env und Rec sind jeweils mit einem + vermerkt. Provirus HERV-K 102 auf Chromosom 1q22 trägt laut Human Genome Browser (March, 2006) ein vorzeitiges Stop-Codon (+) im Gag, ebenso wie HERV-K 33 und HERV-K II.

Zelllinie	Anzahl Klone	Amplifikat	Provirus	Chromosomale Lokalisation	ID (GenBank)	Typ	ORFs		
							Gag	Env	Rec
UKRV Mel 2 -C9	7	<i>gag/prot</i>	HERV-K 108	7p22.1	AC072054	2	+	+	+
	9	<i>env</i>	HERV-K 108	7p22.1	AC072054	2	+	+	+
NCCIT	9	<i>gag/prot</i>	HERV-K 102	1q22	AC044819	1	(+)		
	3	<i>gag/prot</i>	HERV-K 108	7p22.1	AC072054	2	+	+	+
	2	<i>gag/prot</i>	HERV-K 10, HERV-K 107	5q33	AC016577	1	+		
	1	<i>gag/prot</i>	HERV-K 33	10p24	AL392086	1	(+)		
	1	<i>gag/prot</i>	HERV-K 68, HERV-K 106	3q13.2	AC078785	2	+	+	+
	2	<i>env</i>	HERV-K 102	1q22	AC044819	1	(+)		
	3	<i>env</i>	HERV-K 108	7p22.1	AC072054	2	+	+	+
	2	<i>env</i>	HERV-K 108	7p22.1	AC072054	2	+	+	+
GH	1	<i>gag/prot</i>	HERV-K 108	7p22.1	AC072054	2	+	+	+
	2	<i>gag/prot</i>	HERV-K 10, HERV-K 107	5q33	AC016577	1	+		
	6	<i>gag/prot</i>	HERV-K 101	22q11.21	AC007326	1	+		
	4	<i>env</i>	HERV-K 108	7p22.1	AC072054	2	+	+	+
	1	<i>env</i>	HERV-K II	3q12.3	AC084198	1	(+)		

2.3. Die Melanomlinie UKRV Mel 2-C9 bildet Retroviruspartikel

Die Produktion viraler Partikel konnte für GH Zellen und andere Teratokarzinomzelllinien bereits zweifelsfrei gezeigt werden (Boller *et al.*, 1983; Löwer *et al.*, 1984; Boller *et al.*, 1993; Löwer *et al.*, 1993). Auch für einige Melanomlinien gibt es Hinweise auf die Bildung viraler HERV-K Partikel (Muster *et al.*, 2003). Um die Melanomlinie UKRV Mel 2-C9 auf die Produktion viraler Partikel zu untersuchen, wurden zwei Tage nach Teilung der Kulturen Zellkulturüberstände abgenommen. Daraus wurden mittels Ultrazentrifugation Viruspartikel aufgereinigt (siehe 4.9.). Zum Vergleich wurden Kulturüberstände einer stabil mit einem rekombinanten Vollängen HERV-K Genom transfizierten HeLa-Linie (HeLa_{HKges}) untersucht. Da der HERV-K Promotor in HeLa Zellen nicht aktiv ist, war er in dem rekombinanten HERV-K Konstrukt durch einen CMV

Promotor ersetzt worden (M. Magin, 1989; Hahn *et al.*, 2008). Kulturüberstände der nicht HERV-K exprimierenden Linie HCl Mel 19 wurden als negative Kontrolle mitgeführt.

Die Viruspräparationen wurden lysiert und im Immunoblot mit Hilfe von spezifischen Antikörpern auf die Anwesenheit von HERV-K Gag und Env analysiert. In den Partikel der Linien UKRV Mel 2-C9 und HeLa_{HKges} war nur das virale Kapsid (30 kD), nicht aber das Gag Vorläuferprotein (80 kD) nachzuweisen (Abb.: 12B). Dies deutet auf eine vollständige Prozessierung des Gag Proteins hin. Da die Prozessierung des Gag-Vorläuferproteins bei der Partikelreifung von der viralen Protease durchgeführt wird, impliziert dieses Ergebnis auch, dass die Proteasen des Provirus HERV-K 108 (UKRV Mel 2-C9) und des genomischen Klons HK_{ges} aktiv sein müssen. Viruspräparationen der nicht HERV-K exprimierenden Linie HCl Mel 19 zeigten keine Signale. In Partikeln aus der transgenen Linie HeLa_{HKges} konnten sowohl das Env Vorläuferprotein (~90 kDa) wie auch die prozessierte Surface-Untereinheit (~60 kDa) nachgewiesen werden (Abb.: 12A). Dies könnte an einem Defekt in der Prozessierung von Env liegen. Da aber die Sequenzen der *env* Gene in HERV-K 108 und dem transgenen Konstrukt HK_{ges} identisch sind und das Env Protein während des Transports durch den sekretorischen Exportweg prozessiert und gespalten wird, erscheint es wahrscheinlicher, dass die HeLa_{HKges} Viruspräparation auch partikuläres zelluläres Material mit nicht gespaltenem Env enthielten. In Partikelpräparationen aus UKRV Mel 2-C9 konnte nur ein Signal für die Surface-Untereinheit (~60 kDa) detektiert werden. Die Präparation scheint reiner zu sein. In den Viruspräparationen von HCl Mel 19 Zellkulturüberständen konnte kein Env nachgewiesen werden.

Zusammenfassend lässt sich sagen, dass sowohl von der Melanomlinie UKRV Mel 2-C9 wie auch von der transgenen Linie HeLa_{HKges} virale Partikel in das umgebende Nährmedium abgegeben werden. Vergleicht man, unter Beachtung der Menge des aufgearbeiteten Zellkulturüberstandes, die Stärke der jeweiligen Gag Immunoblotssignale, wird deutlich, dass die Produktion viraler Partikel in der transgenen Linie HeLa_{HKges} um ein Vielfaches höher ist als in der Linie UKRV Mel 2-C9. Dies könnte daran liegen, dass der CMV-Promotor des rekombinanten HERV-K Genoms in den HeLa-Zellen stärker ist als der authentische Promotor des HERV-K 108 Provirus in den UKRV Mel 2-C9-Zellen. In die Partikel der Linie UKRV Mel 2-C9 wird mehr Env Protein eingelagert als in die Partikel der Linie HeLa_{HKges}.

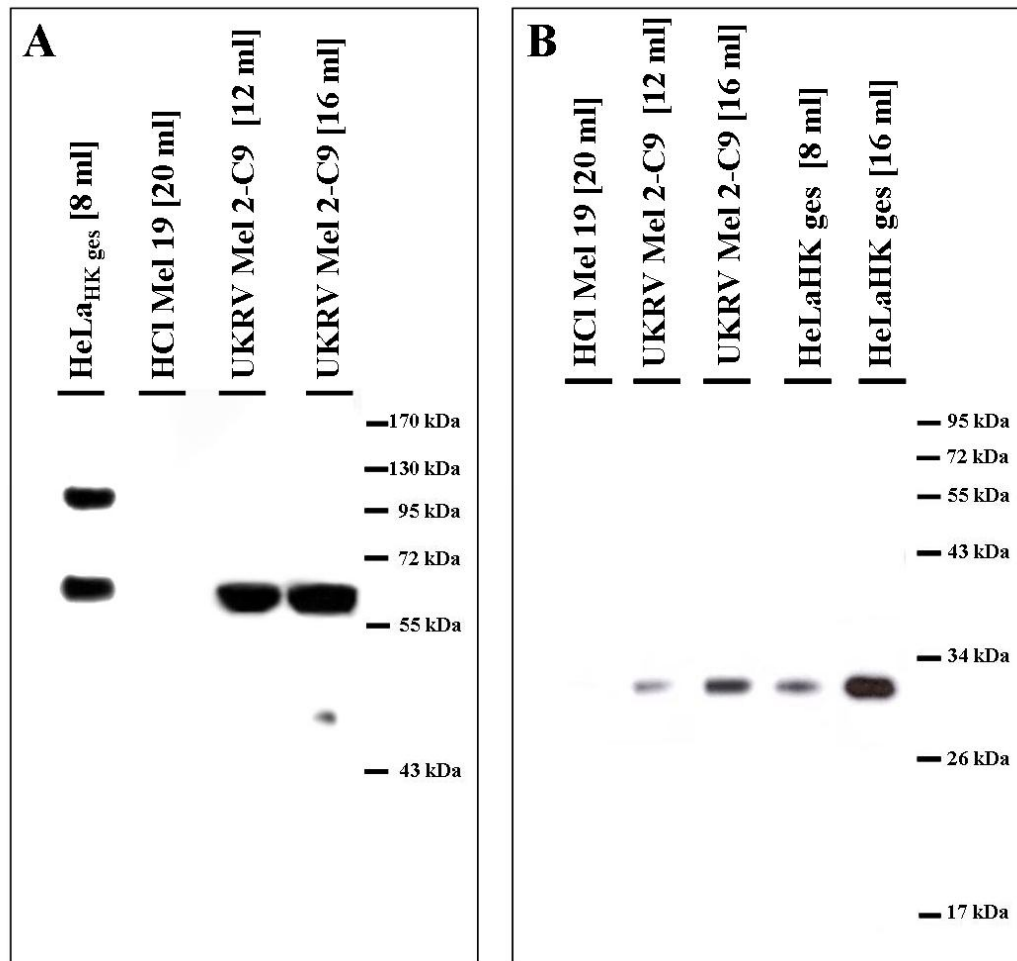


Abbildung 12: Immunoblotanalyse zur Untersuchung von partikulär verpacktem HERV-K Gag und Env-Protein in UKRV Mel 2-C9 und HeLa_{HK_ges}. Die durch Ultrazentrifugation von Zellkulturüberständen pelletierten Viruspartikel wurden lysiert und über ein 12 %iges SDS-Gel aufgetrennt. In Klammern angegeben ist jeweils das Äquivalent der Zellkulturüberstände, aus dem aufgereinigt wurde. **A)** Das Env-Vorläuferprotein (~90 kDa) und die prozessierte Surface-Untereinheit (~60 kDa) konnten detektiert werden. Es wurde der Primärantikörper Kaninchen anti-HERV-K Env 8655 in einer Verdünnung von 1:4000 eingesetzt. Zur Detektion der Signale wurde ein peroxidasegekoppelter Sekundärantikörper anti-Kaninchen aus Ziege in einer Verdünnung von 1:10000 verwendet. **B)** Mit spezifischen Antikörpern für HERV-K Gag konnte vollständig prozessiertes Kapsid mit einem Molekulargewicht von 30 kDa detektiert werden. Es wurde der monoklonale Mausantikörper HERMA 7 (Hybridomaüberstand) verwendet. Als Sekundärantikörper diente ein peroxidasegekoppelter Mausantikörper aus Ziege in einer Verdünnung von 1:10000.

2.3.1. Analyse der Partikelmorphologie im Elektronenmikroskop

Für die elektronenoptischen Untersuchungen wurden UKRV Mel 2-C9 Zellen mit Glutaraldehyd fixiert und in Paraffin eingebettet. Mit dem Mikrotom hergestellte Ultradünnschnitte wurden kontrastiert und im Elektronenmikroskop analysiert (4.9.2.). In UKRV Mel 2-C9 Zellen konnten virale Strukturen (Abb.: 13) mit morphologischer Ähnlichkeit zu publizierten HERV-K Partikeln identifiziert werden. Sie traten allerdings nicht in den für Keimzelltumoren typischen Gruppen auf, sondern meist vereinzelt. Die

viralen Partikel wiesen prominente *Spike*-Strukturen an der Oberfläche auf. Diese werden von den viralen Hüllproteinen gebildet und sind ein Hinweis auf intakte Env-Proteine. Damit zeigten die von der Linie UKRV Mel 2-C9 produzierten Partikel eine starke morphologische Ähnlichkeit zu den HERV-K Partikeln, die von der Keimzelltumormlinie 2102EP gebildet werden (Bieda *et al.*, 2001). Die untersuchten Partikel besaßen in den meisten Fällen ein unvollständig kollabiertes *Core*, es konnten jedoch auch einzelne Partikel mit kondensierten *Core*strukturen detektiert werden. Letztere sind charakteristisch für vollständig ausgereifte Partikel und eine Voraussetzung für Infektiosität. Diese Ergebnisse wurden in Zusammenarbeit mit dem Fachgebiet Morphologie des Paul-Ehrlich-Instituts unter der Leitung von Dr. Klaus Boller erzielt.

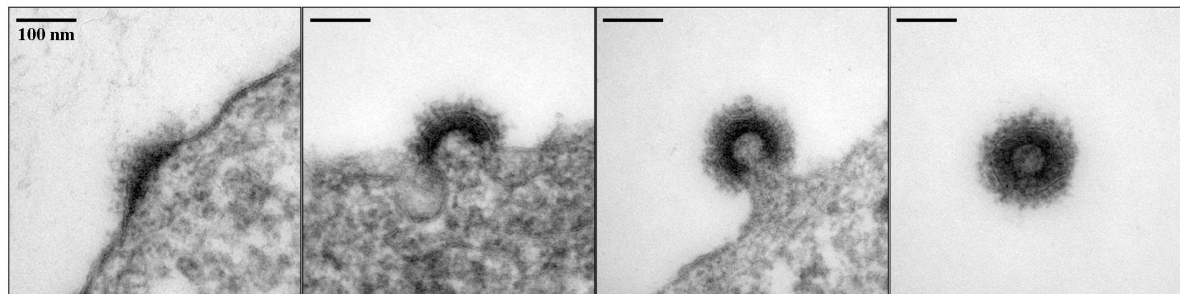


Abbildung 13: Elektronenmikroskopische Aufnahmen unterschiedlicher Knospungsstadien von HERV-K Partikeln in Dünnschnitt-Präparaten von UKRV Mel 2-C9. Die Viruspartikel zeigen prominente *Spike*-Strukturen an ihrer Oberfläche. Reife Partikel mit kollabiertem *Core* sind sehr selten. Maßbalken 100 nm.

2.3.2. HERV-K RNA wird in HERV-K Partikel verpackt

Intakte retrovirale Partikel beherbergen die virale Erbinformation in Form von zwei +RNA-Strängen. Um zu analysieren, ob von der Melanomlinie UKRV Mel 2-C9 gebildete Partikel verpackte HERV-K RNA Genome enthalten, wurde RNA aus Partikeln extrahiert und untersucht (4.9.1.). Dazu wurden erneut virale Partikel aus Zellkulturüberständen durch Ultrazentrifugation ankonzentriert. Viruspräparationen aus Zellkulturüberständen der Keimzelltumormlinie GH wurden parallel als Positivkontrolle mitgeführt. Als Negativkontrolle dienten Partikelpräparationen aus Kulturüberständen der nicht HERV-K exprimierenden Melanomlinie HCl Mel 19. Durch eine RNase-Behandlung vor dem Aufschließen der Partikel wurden Verunreinigungen mit zytoplasmatischer RNA ausgeschlossen. Die Partikel wurden lysiert und die enthaltene RNA extrahiert. Vor der cDNA-Synthese wurde zusätzlich ein DNase-Verdau der RNA durchgeführt, um

Kontaminationen mit zellulärer DNA auszuschließen. Anschließend wurden ein 637 bp großer Bereich aus dem *gag* Gen und ein 567 bp großer Bereich aus dem *env* Gen mittels PCR amplifiziert. Die verwendeten Primer (4.9.1.) binden in konservierten Sequenzbereichen und sollten daher die Amplifikation vieler HERV-K Proviren ermöglichen. Zur Kontrolle der PCR wurde genomische HeLa DNA eingesetzt. Abbildung 14 zeigt, dass aus den cDNA Präparationen von GH und UKRV Mel 2-C9 Partikeln sowie aus der genomischen HeLa DNA die erwarteten *gag* und *env* Produkte amplifiziert wurden. In den Kontrollen der Aufreinigung der Linie HCl Mel 19 und den Ansätzen ohne Zugabe von Reverser Transkriptase, waren keine Amplifikate nachweisbar.

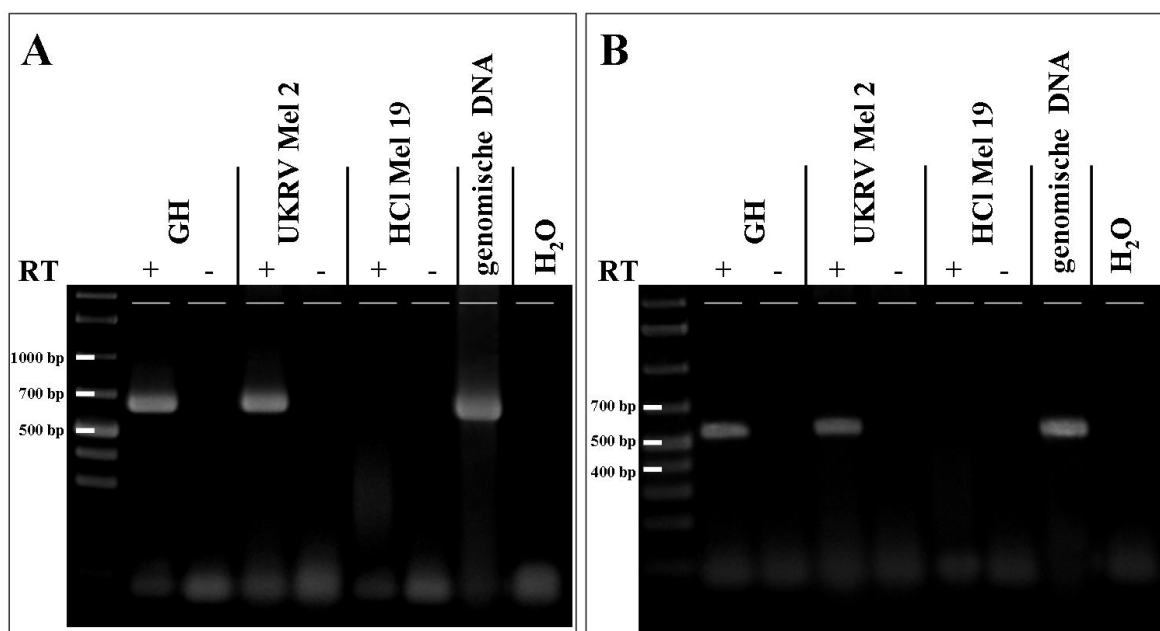


Abbildung 14: Analyse der Transkripte in viralen Partikeln aus den Tumorklinien UKRV Mel 2-C9 und GH. Gezeigt sind Amplifikate aus HERV-K *gag* und *env* Bereichen nach cDNA-Synthese und PCR. Als Template diente RNA aus aufgereinigten viralen Partikeln mit (+) oder ohne (-) Zugabe von Reverser Transkriptase. **A)** Amplifikate aus dem Bereich von *gag* (637 bp), in den Linien UKRV Mel 2-C9 und GH, HCl Mel 19 als Negativkontrolle, H₂O-Kontrolle, genomische HeLa-DNA zur Positivkontrolle, DNA-Marker **B)** Amplifikate aus dem Bereich von *env* (567 bp).

Die zu untersuchenden PCR Amplifikate aus cDNA wurden in den Vektor pGEM-Teasy kloniert (Vergleich Kapitel 2.2.2.). 46 Klone aus der UKRV Mel 2-C9 Analyse und 33 Klone aus der GH Analyse wurden sequenziert. Die Auswertung der Sequenzierergebnisse und die Zuordnung der Sequenzen zu bestimmten HERV-K Loci erfolgte analog zu 2.2.2. in vergleichenden Analysen unter Verwendung der Analysesoftware BLAT und Vektor-NTI. Auch hier konnten in den untersuchten Klonen vereinzelt Nukleotidunterschiede detektiert werden, die aus Fehlern der cDNA-Synthese oder PCR-Amplifikation resultierten und deshalb nicht für die Zuordnung zu HERV-K Proviren gewertet wurden. In

den Klonen aus der Viruspräparation der Linie UKRV Mel 2-C9 wurden mit einer Häufigkeit von 96 % Sequenzen des HERV-K 108 (Typ 2 Provirusgenom) detektiert (Tabelle 2). HERV-K 101 (Typ 1 Provirusgenom) wurde mit einer Frequenz von 4 % verpackt. Die von der Keimzelltumorklinie GH gebildeten retroviralen Partikel enthalten RNA-Sequenzen verschiedener Proviren. HERV-K 101 wurde mit einer Frequenz von 64 %, HERV-K 108 mit einer Frequenz von 33 % und HERV-K 102 mit einer Frequenz von 3 % verpackt. Die Ergebnisse der Sequenzierung finden sich im Anhang (8.2).

Tabelle 2: Zuordnung partikulärer HERV-K RNA zu den korrespondierenden chromosomalen Loci. Aus Partikeln der Linie UKRV Mel 2-C9 und GH isolierte RNA wurde in cDNA umgeschrieben. Mittels PCR wurden definierte Bereiche aus *gag* und *env* amplifiziert, sequenziert und identifiziert. Angegeben ist jeweils die Anzahl untersuchter Klone und die zugeordneten Proviren. Die HERV-K Loci sind durch ihre bibliographischen Namen und die GenBank-Identifikationsnummern beschrieben. Typ 1 oder 2 zeigt die Anwesenheit (Typ 2) oder die Abwesenheit (Typ 1) einer 292 bp langen Sequenz am Übergang von *pol* zu *env* an. Vorhandene offene Leserahmen für Gag, Env und Rec sind jeweils mit einem + vermerkt. Provirus 102 auf Chromosom 1q22 trägt laut Human Genome Browser (March, 2006) ein vorzeitiges Stop-Codon (+) im Gag.

Zelllinie	Anzahl Klone	Amplifikat	Provirus	Chromosomale Lokalisation	ID (GenBank)	Typ	ORFs		
							Gag	Env	Rec
UKRV Mel 2-C9	25	<i>gag</i>	HERV-K 108	7p22.1	AC072054	2	+	+	+
	19	<i>env</i>	HERV-K 108	7p22.1	AC072054	2	+	+	+
	2	<i>env</i>	HERV-K 101	22q11.21	AC007326	1	+		
GH	13	<i>gag</i>	HERV-K 101	22q11.21	AC007326	1	+		
	2	<i>gag</i>	HERV-K 108	7p22.1	AC072054	2	+	+	+
	8	<i>env</i>	HERV-K 101	22q11.21	AC007326	1	+		
	1	<i>env</i>	HERV-K 102	1q22	AC044819	1	(+)		
	9	<i>env</i>	HERV-K 108	7p22.1	AC072054	2	+	+	+

2.3.3. HERV-K 108 der Linie UKRV Mel 2-C9 besitzt keine funktionale Reverse Transkriptase

Eine wesentliche Eigenschaft des retroviralen Replikationszyklus ist die Insertion des revers transkribierten viralen Genoms in die Erbinformation des Wirtes. Eine aktive Reverse Transkriptase trägt spezifische Aminosäuresequenzen in Form von konservierten Motiven. Die katalytisch wichtigen Motive LPQG und YIDD sind eine Voraussetzung für eine funktionale Reverse Transkriptase. Um die Frage zu klären, ob von der Linie UKRV Mel 2-C9 eine aktive Reverse Transkriptase gebildet werden kann, wurde der fragliche Bereich aus partikulär verpackten Vollängentranskripten amplifiziert und sequenziert. Die Untersuchungen ergaben, dass das in der Melanomlinie UKRV Mel 2-C9 reaktivierte Provirus HERV-K 108 kein intaktes YIDD-Motiv trägt (Abb.: 15). Eine von diesem Provirus gebildete Reverse Transkriptase ist wahrscheinlich nicht funktional. Zum

Vergleich wurden Proben aus GH Zellen untersucht. Die Linie GH exprimiert sehr stark das Provirus HERV-K 102. Laut Datenbankeinträgen zeigt dieses Provirus ein YIDD-Motiv, das konnte auch im Experiment bestätigt werden.

HERV-K 108 (Datenbankeintrag)	AEQDCEKFAFTI PAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVREKFSDCYI IHCIDDTLCAAETKDKLIDCYTF LQAEVANA
UKRV Mel 2-C9-1	AEQDCEKFAFTI PAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVREKFSDCYI IHCIDDTLCAAETKDKLIDCYTF LQAEVANA
UKRV Mel 2-C9-2	AEQDCEKFAFTI PAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVREKFSDCYI IHCIDDTLCAAETKDKLIDCYTF LQAEVANA
UKRV Mel 2-C9-3	AEQDCEKFAFTI PAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVREKFSDCYI IHCIDDTLCAAETKDKLIDCYTF LQAEVANA
UKRV Mel 2-C9-4	AEQDCEKFAFTI PAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVREKFSDCYI IHCIDDTLCAAETKDKLIDCYTF LQAEVANA
UKRV Mel 2-C9-5	AEQDCEKFAFTI PAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVREKFSDCYI IHCIDDTLCAAETKDKLIDCYTF LQAEVANA
GH-1	AEQDCEKFAFTI PAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVREKFSDCYI IHCIDDTLCAAETKDKLIDCYTF LQAEVANA
GH-2	AEQDCEKFAFTI PAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVREKFSDCYI IHCIDDTLCAAETKDKLIDCYTF LQAEVANA
GH-3	AEQDCEKFAFTI PAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVREKFSDCYI IHCIDDTLCAAETKDKLIDCYTF LQAEVANA
GH-4	AEQDCEKFAFTI PAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVREKFSDCYI IHCIDDTLCAAETKDKLIDCYTF LQAEVANA
GH-5	AEQDCEKFAFTI PAINNKEPATRFQWKVLPQGMLNSPTICQTFVGRALQPVREKFSDCYI IHCIDDTLCAAETKDKLIDCYTF LQAEVANA

Abbildung 15: Aminosäurealignment aus dem Bereich des aktiven Zentrums der Reversen Transkriptase von HERV-K. Gezeigt sind die Sequenzierergebnisse unterschiedlicher Klone der *pol*-Amplifikate aus den Linien UKRV Mel 2-C9 und GH. Die Amplifikate sind generiert worden mittels spezifischer Primer aus dem Bereich der Reversen Transkriptase, als Template diente cDNA aus partikulär verpackter viraler RNA. Der Datenbankeintrag für HERV-K 108 ist zum Vergleich mit aufgeführt.

Zusammenfassend konnte gezeigt werden, dass die von der Zelllinie UKRV Mel 2-C9 und GH produzierten viralen Partikel HERV-K RNA Genome enthalten. Das in UKRV Mel 2-C9 aktivierte Provirus HERV-K 108 kodiert für intakte Hüllproteine, die *Spikes* bilden und für eine intakte Protease, die die Gag Proteine so prozessiert, dass zumindest das Kapsid freigesetzt wird. Die Reverse Transkriptase allerdings scheint inaktiv zu sein. UKRV Mel 2-C9 Partikel können aus diesem Grunde nicht infektiös sein.

2.4. Analyse der HERV-K LTR

2.4.1. Bestimmung von Transkriptionsstart und Polyadenylierungsstelle (5' und 3' RACE)

Da in der Melanomlinie UKRV Mel 2-C9 ähnliche Proviren reaktiviert werden, wie in der Teratokarzinomlinie GH, wurden sowohl der Transkriptionsstart in der 5' LTR wie die Polyadenylierungsstelle in der 3' LTR parallel in beiden Zelllinien bestimmt. Diese Analysen wurden an endogen exprimierten Proviren vorgenommen. Das verwendete Kit basiert auf der Kopplung eines RNA-Linkeroligonukleotids an zuvor aus der Zelle isolierte mRNA-Moleküle. Eine genaue Beschreibung und ein schematischer Überblick über die verwendete Methode findet sich in 4.6.. Durch Phosphatasebehandlung wurden die nicht durch ein 5'-CAP geschützten RNA-Spezies dephosphoryliert. Im darauf folgenden Schritt wurde durch die Pyrophosphatase das 5'-CAP der vollständigen mRNAs entfernt, so dass das 5' Phosphat für die Ligation des RNA-Linkers verfügbar wurde. Es folgten eine

cDNA-Synthese und eine PCR Amplifikation in einem *Nested*-Ansatz mit jeweils einem spezifischem Primer für das angefügte Oligo und einem spezifischem proviralen Primer. Die verwendeten HERV-K Primer liegen in einer konservierten Region des Provirus, der Primerbindungsstelle. Durch die Lokalisation der PCR-Primer in diesem Bereich ist die Amplifikation vieler bekannter Proviren möglich. Die PCR Amplifikate wurden kloniert, sequenziert und ausgewertet. Da der Transkriptionsstart in GH Zellen bereits bekannt war (M. Kraft, 2004), diente die aus GH Zellen isolierte mRNA als zusätzliche Kontrolle.

Die Sequenzierung der durch PCR generierten Amplifikate und der Vergleich dieser Sequenzierungsergebnisse mit HERV-K Datenbankeinträgen zeigte einen übereinstimmend favorisierten Transkriptionsstartpunkt in GH Zellen und UKRV Mel 2-C9. Die Transkripte beginnen bei nt 794 (Abb.: 16). In Einzelfällen kann es zu einer verfrühten Initiation der Transkription bis zu zehn Nukleotiden kommen. Dies wurde in der Melanomzelllinie UKRV Mel 2-C9 häufiger beobachtet als in GH Zellen. Ausschließlich in den Melanomzellen konnte ein bis mehr als dreißig Nukleotide verspätet einsetzender Transkriptionsbeginn ermittelt werden. Der Transkriptionsstartpunkt in GH Zellen scheint definierter zu sein als in Melanomen. Da in der strukturellen Aufteilung der LTR der Transkriptionsstart zugleich den Übergang von U3 zu R markiert, beginnt R somit bei Position nt 794 (Abb.: 17).

Transkriptionsstart	
RACE-Produkte	+1 ↓
GH- 1	ACAGATGATCAGTAAATACTAAGGGAACCTCAGAGGCTGGCGGGATCCTCCATATGCTGA
GH- 2	-----GGTAAATACTAAGGGAACCTCAGAGGCTGGCGGGATCCTCCATATGCTGA
GH- 3	-----GTAAATACTAAGGGAACCTCAGAGGCTGGCGGGATCCTCCACATGCTGA
GH- 4	-----AGTAAATACTAAGGGAACCTCAGAGGCTGGCGGGATCCTCCATATGCTGA
GH- 5	-----AGTAAATACTAAGGGAACCTCAGAGGCTGGCGGGATCCTCCATATGCTGA
GH- 6	-----AGTAAATACTAAGGGAACCTCAGAGGCTGGCGGGATCCTCCATATGCTGA
Mel- 1	-----AAATACTAAGGGAACCTCAGAGGCTGGCGGGATCCTCCATATGCTGA
Mel- 2	-----GTAAATACTAAGGGAACCTCAGAGGCTGGCGGGATCCTCCATATGCTGA
Mel- 3	-----GTAAATACTAAGGGAACCTCAGAGGCTGGCGGGATCCTCCATATGCTGA
Mel- 4	-----AATACTAAGGGAACCTCAGAGGCTGGCGGGATCCTCCATATGCTGA
Mel- 5	ACAGATGATCAGTAAATACTAAGGGAACCTCAGAGGCTGGCGGGATCCTCCATATGCTGA
Mel- 6	ACAGATGATCAGTAAATACTAAGGGAACCTCAGAGGCTGGCGGGATCCTCCATATGCTGA
Mel- 7	-----AGTAAATACTAAGGGAACCTCAGAGGCTGGCGGGATCCTCCATATGCTGA
Mel- 8	----ATGATCAGTAAATACTAAGGGAACCTCAGAGGCTGGCGGGATCCTCCATATGCTGA
Mel- 9	-----AGTAAATACTAAGGGAACCTCAGAGGCTGGCGGGATCCTCCATATGCTGA
Mel- 10	-----ATCCTCCATATGCTGA
Mel- 11	-----TCTCCATATGCTGA
pcK30	ACAGATGATCAATAATACTAAGGGAACCTCAGAGGCTGGCGGGATCCTCCATATGCTGA
	↓ 794 nt

Abbildung16: Sequenzalignment der RACE-Produkte aus GH und UKRV Mel 2-C9 mit der aktiven LTR pcK30. Der Transkriptionsstartpunkt liegt an Position nt 794 in der 5' LTR.

Der Übergang von R zu U3 ist definiert durch die Polyadenylierungsstelle in der 3' LTR. Das Polyadenylierungssignal (AGTAAA) fällt auf die Position des in GH und UKRV Mel 2-C9 bestimmten Transkriptionsstarts. In GH Zellen konnte gezeigt werden, dass die Polyadenylierungsstelle 84 nt downstream an Position nt 878 liegt (Löwer *et al.*, 1993a). Dieser Abstand ist etwa 60 nt länger als gewöhnlich für *mRNAs* beobachtet, bei denen er 20 bis 30 nt beträgt. Für die Melanomlinie UKRV Mel 2-C9 wurde mittels einer 3'RACE die Polyadenylierungsstelle bestimmt. Es konnten zwei unterschiedliche Positionen als Transkriptionsenden definiert werden. Eine Polyadenylierungsstelle liegt an Position 878 nt wie vermutet und bereits in GH Zellen gezeigt. Eine weitere findet sich an Position 813 nt und entspricht damit in ihrem Abstand zum Polyadenylierungssignal eher der Regel. Wie für den Start der Transkription bereits festgestellt, scheint in GH Zellen auch das Ende der Transkription determinierter zu sein als in der Melanomlinie.

Die Struktureinheiten der HERV-K LTR definieren sich in GH Zellen und Melanomen identisch wie folgt (Abb.: 17). Der U3-Bereich umfasst nt 1 bis nt 793. Innerhalb dieses Bereiches gibt es eine putative TATA-Box (532-538 nt), die allerdings weit stromaufwärts des Transkriptionsstarts liegt. Die R Region beginnt in der Regel bei nt 794 und endet bei nt 879, in der untersuchten Melanomlinie alternativ auch bei nt 814. Sie ist mit 85 Nukleotiden (alternativ 25 Nukleotiden) sehr kurz. An ihrem 5' Ende befindet sich das Polyadenylierungssignal AGTAAA. Der U5-Bereich beginnt bei nt 880 bzw. nt 820 und endet bei nt 968. Er umfasst damit 88 beziehungsweise 148 Nukleotide.

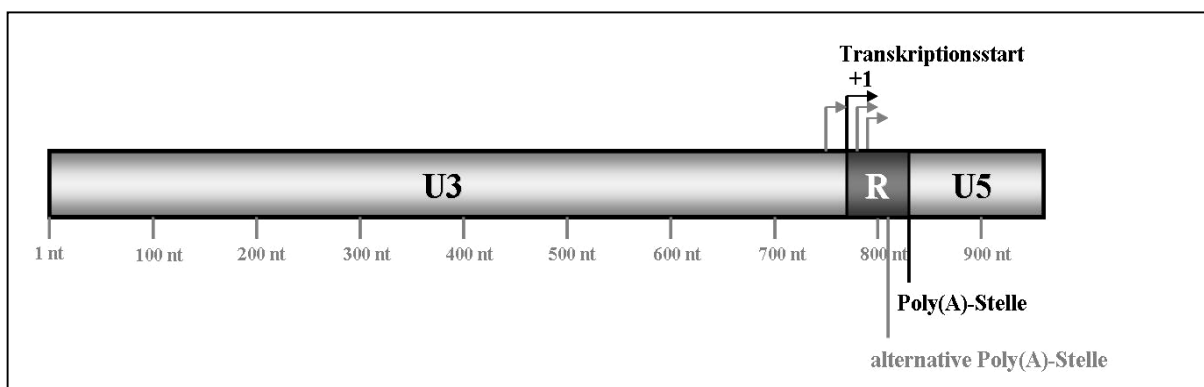


Abbildung 17: Schematische Darstellung der HERV-K LTR. Die Strukturbereiche U3/R/U5 der HERV-K LTR sind dargestellt, der Transkriptionsstart ist durch +1 gekennzeichnet, alternative Transkriptionsstartpunkte sind durch graue Pfeile gekennzeichnet. Die Poly(A)-Stelle in der 3' LTR ist eingetragen, alternative Poly(A)-Stellen sind ebenfalls durch graue Pfeile kenntlich gemacht.

2.4.2. Vergleich der HERV-K LTR Promotoraktivität in UKRV Mel 2-C9 und GH

Um einen Eindruck der Promotoraktivität der HERV-K LTR in den untersuchten Zelllinien zu bekommen, wurde das duale Luziferasereportersystem von Promega eingesetzt (4.7.5.). Hierbei wurden die zu testenden Zelllinien mit Konstrukten transfiziert, in denen die HERV-K Promotoren vor ein Luziferasegen kloniert worden waren. Je aktiver ein Promotor umso effizienter kann das Luziferasegen vom Plasmid abgelesen werden. Das Prinzip des Assays ist in Abb.: 18 dargestellt.

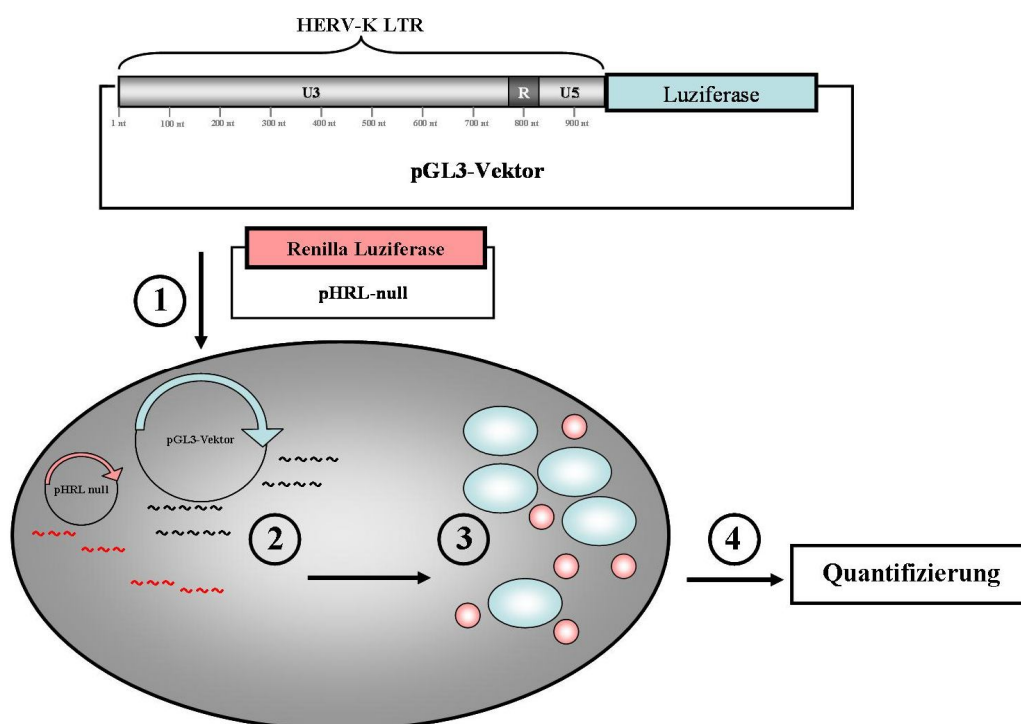


Abbildung 18: Funktionsweise des Luziferasereporterassays. Der zu untersuchende Promotor, zum Beispiel die HERV-K LTR, befindet sich vor einem Glühwürmchen-Luziferasegen (blau). **1)** Das Promotorkonstrukt wird durch Transfektion in die Zelle eingebracht. Zur Normierung wird ein interner Standard kotransfiziert. Dieses Konstrukt enthält keinen aktiven Promotor vor dem Luziferasegen der Weichkoralle *Renilla reniformis* (rot) und zeigt damit nur basale Aktivität (pHRL null). **2)** Je nach Aktivität des Promotors wird das jeweilige Luziferasegen abgelesen. **3)** Von den *mRNAs* wird die jeweilige Luziferase synthetisiert. **4)** Nach der Lyse der Zellen kann die Menge gebildeter Luziferase im Photomultiplier dadurch quantifiziert werden, dass das von dem jeweiligen Enzym umgesetzte Substrat Photonen aussendet. Zur Normierung werden die gemessenen Lichteinheiten der Glühwürmchen-Luziferase durch die der Weichkorallen-Luziferase dividiert.

Verwendet wurden die Konstrukte pGL3-pcK30, pGL3-LTR 21, pGL3control und pGL3basic. Der Vektor pGL3basic enthält keine Promotor- und Enhancer-Sequenzen und dient als Negativkontrolle. Dagegen bewirken die SV40 Promotor- und Enhancer-Sequenzen im Vektor pGL3control eine sehr starke Luziferase-Expression in sehr vielen

Zelltypen. Der promotorlose Vektor pHRL-null dient als interne Kontrolle, auf die normiert werden kann, pGL3-pcK30 enthält eine HERV-K LTR (Magin *et al.*, 1999). Die pcK30-LTR wurde aus cDNA-Sequenzen der Zelllinie GH zusammengesetzt. Die größte Homologie zwischen der pcK30 und einer endogen vorkommenden LTR zeigt die virale LTR des Provirus HERV-K 108. Zwischen der pcK30 und der 5' LTR des Provirus HERV-K 108 gibt es sechs Basenunterschiede (Tabelle 3). Das Konstrukt pGL3-LTR 21 enthält eine LTR, die aus genomischer DNA amplifiziert wurde. Ihre Sequenz weist sehr große Ähnlichkeit zur LTR des Provirus HERV-K 18 auf. Lediglich neun Punktmutationen unterscheiden die LTR 21 von der endogen vorkommenden HERV-K 18 LTR (Tabelle 3). Die LTR 21 zeigte in Keimzelltumorkonstrukten keine Aktivität (B. Hasenmeier, 1995). Beide HERV-K LTR Konstrukte enthalten neben den eigentlichen LTR Sequenzen noch einen kurzen Sequenzabschnitt der 5'UTR mit der Primerbindestelle. Die Sequenzunterschiede zwischen den natürlich vorkommenden LTRs und den artifiziell hergestellten Promotorkonstrukten könnten in Folge von PCR-Fehlern entstanden sein.

Tabelle 3: Sequenzunterschiede zwischen der im Luziferaseassay verwendeten pcK30 und der 5' LTR des Provirus HERV-K 108. Und zwischen der im Luziferaseassay als Negativkontrolle eingesetzten LTR 21 und der LTR des Provirus HERV-K 18.

Position in LTR [nt]	HERV-K 108	pcK30	Position in LTR [nt]	HERV-K 18	LTR 21
175	G	A	16	C	G
179	A	G	301	G	C
341	C	G	359	G	T
358	T	A	419	A	G
361	C	T	464	C	T
794	G	C	503	A	G
			511	T	C
			771	C	A
			772	A	G

Vierundzwanzig Stunden nach der Transfektion wurden die Zellen mit Promega-Lysepuffer behandelt. Die Luziferaseaktivität wurde quantifiziert und auf die kotransfizierte interne Renillakontrolle (pHRLnull) normiert. Die LTR 21 zeigte in keiner Zelllinie Promotoraktivität (Abb.: 19, roter Balken) und bewegte sich damit auf dem gleichen Niveau wie die promotorlose Negativkontrolle (gelber Balken). Dagegen wies die LTR pcK30 (schwarzer Balken) in der Keimzelltumorkonstruktion GH eine um das dreifache erhöhte Aktivität im Vergleich zum SV40 Promotor auf. Dieses Ergebnis war zu erwarten, da bekannt ist, dass der SV40 Promotor zwar in vielen Geweben stark unterstützt wird, jedoch in Keimzelltumoren oft einer Hemmung unterliegt. In der Melanomlinie

UKRV Mel2-C9 zeigten die HERV-K LTR pcK30 und der SV40 Promotor gleich starke Aktivitäten (Abb.: 19). Die Linie HeLa zeigte lediglich Luziferaseaktivität bei Verwendung des SV40-Promotors. Das Fehlen von HERV-K Transkripten in HeLa Zellen scheint auf transkriptioneller Inhibition zu beruhen. Das Ergebnis legt nahe, dass HERV-K Proviren durch zelltypspezifische Faktoren aktiviert werden. Diese Faktoren scheinen in HeLa Zellen nicht oder nicht in ausreichendem Maße vorhanden zu sein. In Teratokarzinomen und den HERV-K positiven Melanomlinien wirken sie sich dagegen positiv auf die HERV-K Expression aus.

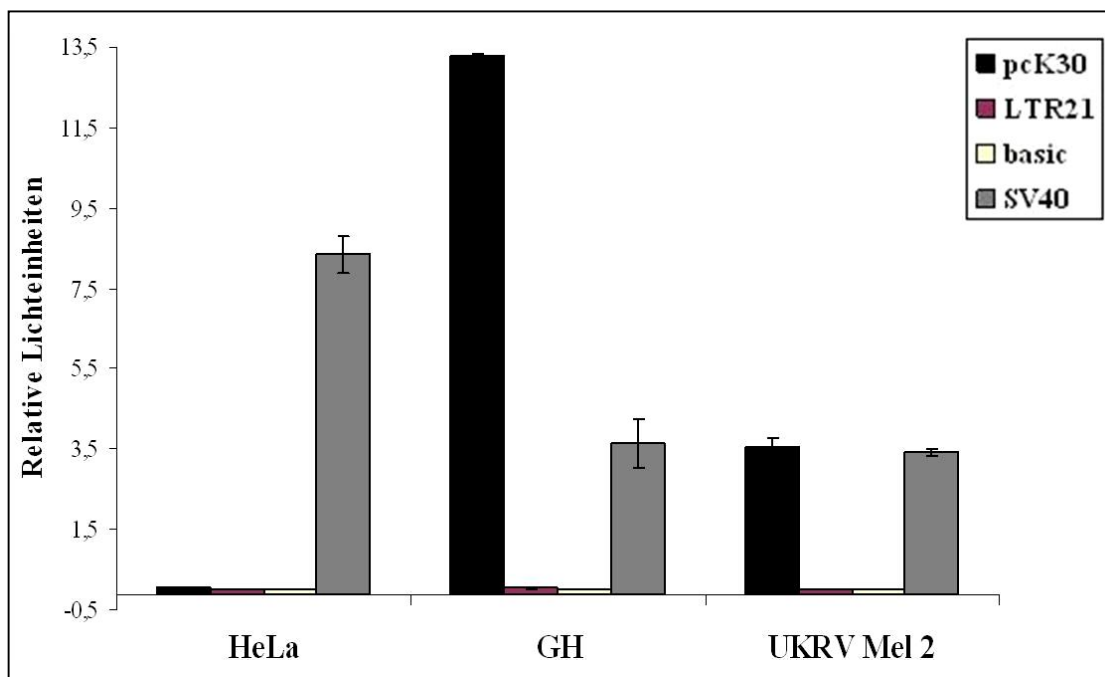


Abbildung 19: Promotoraktivität der HERV-K LTR in unterschiedlichen Zelllinien. HeLa, GH und UKRV Mel 2-C9 Zellen wurden transfiziert mit einem Luziferasereporter gen unter der Kontrolle des HERV-K 108 Promotors pcK30 (schwarze Balken) im Vergleich zu dem inaktiven Promotor LTR 21 (roter Balken), dem SV40 Promotor (grauer Balken) und einem promotorlosen Konstrukt (gelber Balken). Die Luziferaseaktivität als Maß für die Promotorstärke wurde 24 Stunden nach der Transfektion bestimmt und ist in Form relativer Lichteinheiten angegeben. Normiert wurde auf die interne Renillakontrolle. Dargestellt sind die Mittelwerte mit Standardabweichung aus vier parallelen Ansätzen eines repräsentativen Experiments.

2.4.3. Besitzt das TATA-Motiv Relevanz für die LTR-Aktivität von HERV-K in Melanomen?

Bei der Untersuchung der Struktureinheiten der HERV-K LTR in Kapitel 2.4.1. fiel auf, dass sich die putative TATA-Box 262 Nukleotide vor dem Transkriptionsstart befindet. Als nächstes sollte deshalb untersucht werden, ob dem TATA-Motiv bei 532-538 nt überhaupt eine Rolle in der Regulation der HERV-K LTR zukommt. Bereits von M. Kraft

(M. Kraft, 2004) war ein Luziferase-Reporterkonstrukt kloniert worden, in dem das TATAAAA-Motiv der pcK30 LTR gegen die Thymidinfolge TTTTTTT ausgetauscht worden war. Dieses Konstrukt hatte in Keimzelltumorzelllinien unverändert Promotoraktivität gezeigt.

UKRV Mel 2-C9 und GH Zellen wurden mit Luziferase-Reporterkonstrukten (4.7.5.) transfiziert, die entweder die Wildtyp pcK30 LTR oder die mutierte LTR als Promotor enthielten. Nach 48 Stunden wurde die Luziferaseaktivität als Maß für die Promotorstärke bestimmt (Abb.: 20). Die Mutation des TATA-Motivs führte in der Zelllinie GH zu einem Anstieg der Promotoraktivität auf ~280 % im Vergleich zum Wildtyp. In der Melanomzelllinie UKRV Mel 2-C9 wurde mit der Mutante ein Anstieg der Promotoraktivität auf 144 % erreicht. Die Unterschiede in der Luziferaseaktivität zwischen Wildtyppromotor und Mutationskonstrukt waren für die beiden getesteten Zelllinien statistisch signifikant. Dies Ergebnis beweist, dass das TATAAAA-Motiv in keiner Zelllinie als Initiator der Transkription dient. Es deutet eher darauf hin, dass Inhibitoren an dieser Stelle binden.

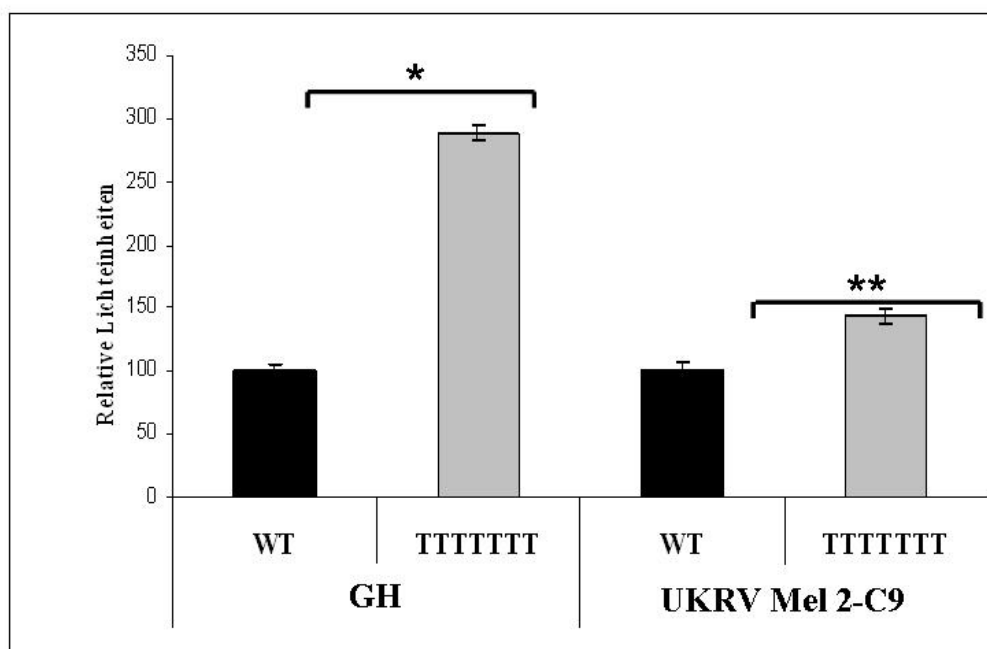


Abbildung 20: Promotoraktivität des TATA-Box-Mutationskonstruktes im Vergleich zum Wildtyp in GH und UKRV Mel 2-C9. GH und UKRV Mel 2-C9 Zellen wurden transfiziert mit einem Luziferasereporter unter der Kontrolle des HERV-K Wildtyp-Promotors (schwarze Balken) im Vergleich zu dem Mutationskonstrukt [TATAAAA>TTTTTTT (532-538 nt)] (graue Balken). Die Luziferaseaktivität als Maß für die Promotorstärke wurde 48 Stunden nach der Transfektion bestimmt. Alle Werte wurden auf die kotransfizierte interne Renillakontrolle normiert und sind als relative Lichteinheiten dargestellt. Die Promotoraktivität des Wildtyps wurde individuell für jede Zelllinie als 100 % gesetzt. Die gezeigten Daten sind die Mittelwerte mit Standardabweichung aus sechs parallelen Ansätzen eines repräsentativen Experiments. Das Mutationskonstrukt zeigte signifikante Unterschiede im Vergleich zum Wildtyp [p* und p**<0,001, Student's t-test].

2.5. In silico Suche nach an der HERV-K Transkriptionsinitiation beteiligten Faktoren und Identifikation putativer Transkriptionsfaktorbindestellen

Wie in vorangegangenen Experimenten gezeigt wurde, handelt es sich im Falle der LTR von HERV-K um einen TATA-unabhängigen Promotor. Die Transkriptionsinitiation an solchen Promotoren wird häufig durch ein Initiatorelement (Inr) im Zusammenspiel mit bestimmten Transkriptionsfaktoren vermittelt. Die Funktion der TATA-Box wird vom Inr übernommen. Die LTR von HERV-K 108 zeigt ein putatives Inr am Transkriptionsstartpunkt (TCAGTAA, 791-798 nt), allerdings mit minimalen Abweichungen von der Konsensussequenz (siehe Alignment). Die Konsensussequenz für ein funktionales Inr lautet Py Py A⁺N T Py Py, wobei die Transkription mit Adenin beginnt. Die LTR von HERV-K 108 weist nach dem Thymin nicht zwei Pyrimidin-, sondern zwei Purinbasen auf.

HERV-K108	TGTGGGGAAAAGCAAGAGAGATCAGATTGTTACTGTGTCTGTGTAGAAAGAAGTAGACAT	60
pcK30	60
HERV-K18C.....G.....A.....	60
LTR21C.....G.....A.....	60
HERV-K108	AGGAGACTCCATTTTGTATGTACTAAGAAAAATTCTCTGCCTTGAGATTCTGTTAATC	120
pcK30	120
HERV-K18C.....A.....G.....	120
LTR21C.....A.....G.....	120
HERV-K108	TATGACCTTACCCCAACCCCGTCTCTCTGAAACATGTGCTGTGTCCA-CTCAGGGTTA	179
pcK30-.....A...G	179
50180		
LTR21A.A.....	180
HERV-K108	AATGGATTAAGGGCGGTGCAGGATGTGCTTTGTTAAACAGATGCTTGAAGGCAGCATGCT	239
pcK30	239
HERV-K18A.....	240
LTR21A.....	240
HERV-K108	CCTTAAGAGTCATCACCCTCCCTAATCTCAAGTACCCAGGGACACAAAACTGCGGAAG	299
pcK30	299
HERV-K18	.A.....	300
LTR21	.A.....	300
HERV-K108	GCCGCAGGGACCTCTGCCTAGGAAAGCCAGGTATTGTCCAACGTTTCTCCCATGTGATA	359
pcK30G.....A.....	359
HERV-K18	..T.....G.....G.....G.....	360
LTR21	..T.....G.....G.....G.....	360
HERV-K108	GCCTGAAATATGGCCTCGTGGAAGGAAAGACCTGACCGTCCCCAGCCCGACACCCGT	419
pcK30	.T.....	419
HERV-K18	GT.....A.....	420
LTR21	GT.....A.....	420
HERV-K108	AAAGGGTCTGTGCTGAGGAGGATTAGTAAAAGAGGAAGGAATGCCTCTTGAGATTGAGAC	479
pcK30	479
HERV-K18T.....A.C.....	480
LTR21T.....A.C.....	480

HERV-K108	AAGAGGAAGGCATCTGTCTCCTGCCTGTCCCTGGGCAATGGAATGTCTCGGTATAAAACC	539
pcK30	539
HERV-K18T...CA..CA...T.....	540
LTR21T...CA..CA...T.....	540
HERV-K108	CGATTGTATGCTCCATCTACTGAGATAGGGAAAAACCGCCTTAGGGCTGGAGGTGGGACC	599
pcK30	599
HERV-K18C.T....C.....G.....A.....A	600
LTR21C.T....C.....G.....A.....A	600
HERV-K108	TGCGGGCAGCAATACTGCTTTGTAAAGCACTGAGATGTTTATGTGTATGCATATCTAAAA	659
pcK30	659
HERV-K18	...A.....T.....	660
LTR21	...A.....T.....	660
HERV-K108	GCACAGCACTTAATCCTTTACATTGTCTATGATGCAAAGACCTTTGTTACATGTTTGTC	719
pcK30	719
HERV-K18T.....C.....G.....	720
LTR21T.....C.....G.....	720
HERV-K108	TGCTGACCCTCTCCCACAATTGTCTTGTGACCCTGACACATCCCCCTCTTCGAGAAACA	779
pcK30	779
HERV-K18	...C.....T.....T.....CA.....	780
LTR21	...C.....T.....T.....CA.....	780
HERV-K108	CCCACAGATGATCAGTAAATACTAAGGGAACCTCAGAGGCTGGCGGGATCCTCCATATGCT	839
pcK30A.....	839
HERV-K18GA.....A.....G.....T.....	840
LTR21GA.....A.....G.....T.....	840
HERV-K108	GAACGCTGGTTCCCCGGGTCCCCCTTCTTTCTTTCTTACTTTGTCTCTGTGTCTTTTT	899
pcK30	899
HERV-K18T.....C.....A.....	900
LTR21T.....C.....A.....	900
HERV-K108	CTTTTCCAAATCTCTCGTCCCACCTTACGAGAAACACCCACAGGTGTGTAGGGGCAACCC	959
pcK30	959
HERV-K18G.....A.TG.....G.....	960
LTR21G.....A.TG.....G.....	960
HERV-K108	ACCCCTACA	968
pcK30	968
HERV-K18T..	969
LTR21T..	969

Alignment: Sequenzvergleich der im Luziferaseassay verwendeten HERV-K LTR pcK30 und LTR 21, sowie den LTRs der Proviren HERV-K 108 und HERV-K 18. Übereinstimmungen mit der Sequenz des Provirus 108 sind als Punkt dargestellt, Abweichungen durch den Einbuchstabencode der entsprechenden Basen. Striche kennzeichnen das Fehlen von Basenpositionen, die mindestens in einer anderen Sequenz vorkommen.

2.5.1. Identifikation putativer Transkriptionsfaktorbindestellen in der HERV-K LTR durch *in silico* Datenbankanalysen

Da die Initiation der Transkription und die Aktivität eines Promotors durch spezifische Transkriptionsfaktoren vermittelt wird, wurden putative Transkriptionsfaktorbindestellen durch eine online Suchanfrage beim „*Transcription Element Search System (TESS)*“ identifiziert, das auf die TRANSFACv6.0 Datenbank zugreift. In dieser *in silico* Analyse wurden die im Luziferaseassay aktive LTR pcK30 und die inaktive LTR 21 auf

Unterschiede in putativen Transkriptionsfaktorbindestellen hin untersucht. Die Auswertung findet sich im Anhang (7.3.). Die Sequenzen der LTR 21 und der pcK30 sind dem Alignment in 2.5. zu entnehmen. Besonderes Interesse galt dabei den Sequenzen in der Nähe des Transkriptionsstarts. Hier konnte eine hoch konservierte GC/GT-Box identifiziert werden, die nur in der aktiven LTR pcK30 vorliegt. An solche GC-Boxen binden bevorzugt Sp (*specific protein*) Proteine. Die Sp-Proteine sind Zinkfingerproteine und gehören der Familie der SP/XKLF Transkriptionsfaktoren an. Die ubiquitär exprimierten Transkriptionsfaktoren Sp1 und Sp3 binden an das gleiche Sequenzmotiv. Da in der Regulation der Transkriptionsinitiation durch Sp1 und Sp3 auch das Arrangement und die Anzahl der Erkennungsstellen auf der LTR eine Rolle spielt, darf die zweite putative Sp1/Sp3-Bindestelle in direkter Nähe zum Transkriptionsstart nicht unbeachtet bleiben. Diese zweite Bindestelle zeigt eine nicht ganz konservierte GC-Box und ist in der aktiven LTR pcK30 ebenso wie in der inaktiven LTR 21 präsent.

	GC-Box1	GC-Box2
pcK30	CCCTGACACATCCCCTCTTCG	GAGAAACACCCACAGATGATCAGTAAATA
LTR 21	CCCTGACACATCTCCCTCTAGG	GAGAAACACCCACAGATGATCAGTAAATA

Sequenzvergleich der pcK30 und LTR 21 im Bereich der GC-Boxen.

2.5.1.1. Analyse der Expression von Sp1 und Sp3 in den untersuchten Zelllinien im Immunoblot

Um die Rolle der Transkriptionsfaktoren Sp1 und Sp3 bei der Transkriptionssteuerung von HERV-K zu untersuchen, wurden die hier verwendeten Zelllinien zunächst in der Expression der Sp1 und Sp3 Proteine verglichen. Dazu wurden fraktionierte Proteinextrakte der Zelllinien UKRV Mel 2-C9 und GH sowie der nicht HERV-K exprimierenden Linien HeLa und HCl Mel 19 hergestellt (4.7.2.). Sp1 und Sp3 sind als Transkriptionsfaktoren im Zellkern lokalisiert. In Immunoblotanalysen konnte Sp1 mit einem Molekulargewicht von 105 kDa in allen Kernproteinaufreinigungen detektiert werden (Abb.: 21). In Extrakten des Zytoplasmas konnten keine Signale nachgewiesen werden. Dies zeigt eine gute Trennung zwischen den Kern- und Zytoplasma-Proteinfractionen während der Aufreinigung. Als Ladekontrolle und als weiteres Merkmal für erfolgreiche Trennung der Kompartimente wurde der Immunoblot mit Lamin A belegt. Lamin A ist ein struktureller Bestandteil des Zellkerns und kommt nur in diesem

Kompartiment vor. Auch Signale für Lamin konnten jeweils nur in den Kernfraktionen detektiert werden.

Sp3 Proteine kommen in drei Isoformen vor, die ebenfalls alle in den jeweiligen Kernproteinextrakten nachgewiesen werden konnten (Abb.: 21). Die lange Isoform mit einem Molekulargewicht von 115 kDa, die auch als Transkriptionsaktivator fungieren kann, wird in GH Zellen stärker exprimiert als die beiden kürzeren Isoformen, denen die zweite Aktivierungsdomäne fehlt. Die kurzen Isoformen haben ein Molekulargewicht von 80 kDa und 78 kDa. In allen anderen Linien scheint das Verhältnis der Isoformen zueinander ausgeglichen (Abb.: 21).

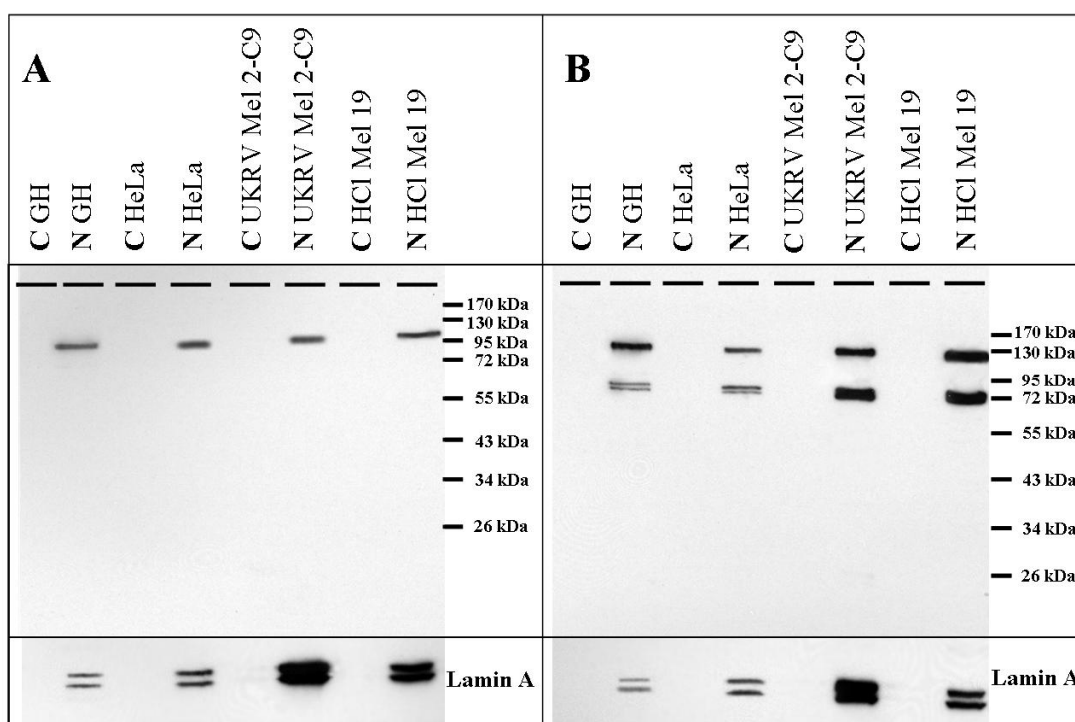


Abbildung 21: Immunoblotanalysen zur Expression von Sp1 und Sp3 in unterschiedlichen Zelllinien. A) Nachweis von Sp1 in den Zytoplasma- [C] und Kernfraktionen [N] der Zelllinien GH, HeLa, UKRV Mel 2-C9 und HCl Mel 19. Verwendet wurde ein polyklonaler Kaninchen anti-Sp1 Antikörper in einer Verdünnung von 1:3000. Sp1 findet sich nur im Nukleus [N]. Zur Kontrolle einer erfolgreichen Fraktionierung wurde auf Anwesenheit des nukleären Proteins Lamin A getestet. Hierzu wurde ein Kaninchen Lamin-A Antikörper in einer Verdünnung von 1:4000 eingesetzt. B) Nachweis von Sp3 Proteinen in den Zytoplasma- [C] und Kernextrakten [N] mit einem polyklonalen Kaninchen anti-Sp3 Antikörper in einer Verdünnung von 1:4000. Als Sekundärantikörper diente in allen Fällen ein peroxidasegekoppelter anti-Kaninchen Antikörper in einer Verdünnung von 1:10000.

2.5.1.2. Bindung von Sp1 und Sp3 in räumlicher Nähe zum Transkriptionsstart (*in vitro* EMSA-Analysen)

Um zu untersuchen, ob die in der HERV-K LTR identifizierten putativen GC-Boxen Bindungsstellen für die Transkriptionsfaktoren Sp1 und Sp3 darstellen, wurden EMSA- (*Electrophoretic Mobility Shift Assays*) Analysen durchgeführt (4.5.). Kernproteinextrakte von GH und UKRV Mel 2-C9 wurden mit einer radioaktiv markierten DNA-Sonde inkubiert (Abb.: 22A). Die Kernproteinextrakte wurden zuvor auf die Anwesenheit von Sp1 und Sp3 untersucht. Als Sonde diente ein 66 bp langes Stück aus der pK30 LTR (738-804 nt), das sowohl die hoch konservierte Sp-Bindestelle (GC-Box1) als auch die weniger konservierte zweite GC-Box enthielt (Abb.: 22A). Die Sonde wurde über PCR mit den Primern MKr21 und MKr28 (8.1.) generiert, anschließend über ein 3 % Agarosegel (4.4.3.2.) aufgereinigt, aus dem Gel eluiert und anschließend radioaktiv markiert (4.5.).

Um die experimentellen Bedingungen zu überprüfen (Abb.: 22B), wurde die Sonde (Spur 1) mit rekombinantem Sp3 Proteinen inkubiert. Der Sonden/SP3 Komplex zeigte in der nativen Gelelektrophorese ein deutlich verzögertes Laufverhalten, einen Shift (Spur 3). Dieser stellte sich als Doppelbande dar, die entweder auf eine Bindung des Proteins als Monomer oder Dimer oder auf unterschiedliche Besetzung der beiden GC-Boxen zurückzuführen sein könnte. Bei Zugabe eines spezifischen Sp3 Antikörpers konnte ein Supershift beider Banden erzeugt werden (Spur 2). Die experimentellen Bedingungen sollten demnach eine Komplexbildung zwischen der Sonde und den Sp-Proteinen der Kernextrakte sowie einen Supershift bei Zugabe spezifischer Antikörper ermöglichen. In der Tat konnten nach Inkubation der Kernproteinextrakte aus GH Zellen mit der DNA-Sonde drei Shift-Banden in der nativen Gelelektrophorese identifiziert werden (Abb.: 22C, Spur 2). Um zu klären, ob es sich um spezifische Komplexe handelt, wurde geprüft, ob die Komplexe durch Zugabe von nicht radioaktiv markierter DNA-Sonde zu kompetitieren sind. Durch einen 50- beziehungsweise 100-fachen Überschuss an nicht markierter Sonde konnte die hochmolekulare Bande spezifisch kompetitiert werden (Abb.: 22C, Spur 3 und 4). Eine vollständige Kompetition mit einem nicht markierten korrespondierenden DNA-Fragment aus der inaktiven LTR 21 war nicht möglich (Abb.: 22C, Spur 5 und 6). Da sich die beiden DNA-Fragmente nur durch die Mutationen in der GC-Box1 unterscheiden (Abb.: 22A), bedeutet dies, dass sich der spezifische hochmolekulare Komplex vor allem an der GC-Box1 bildet.

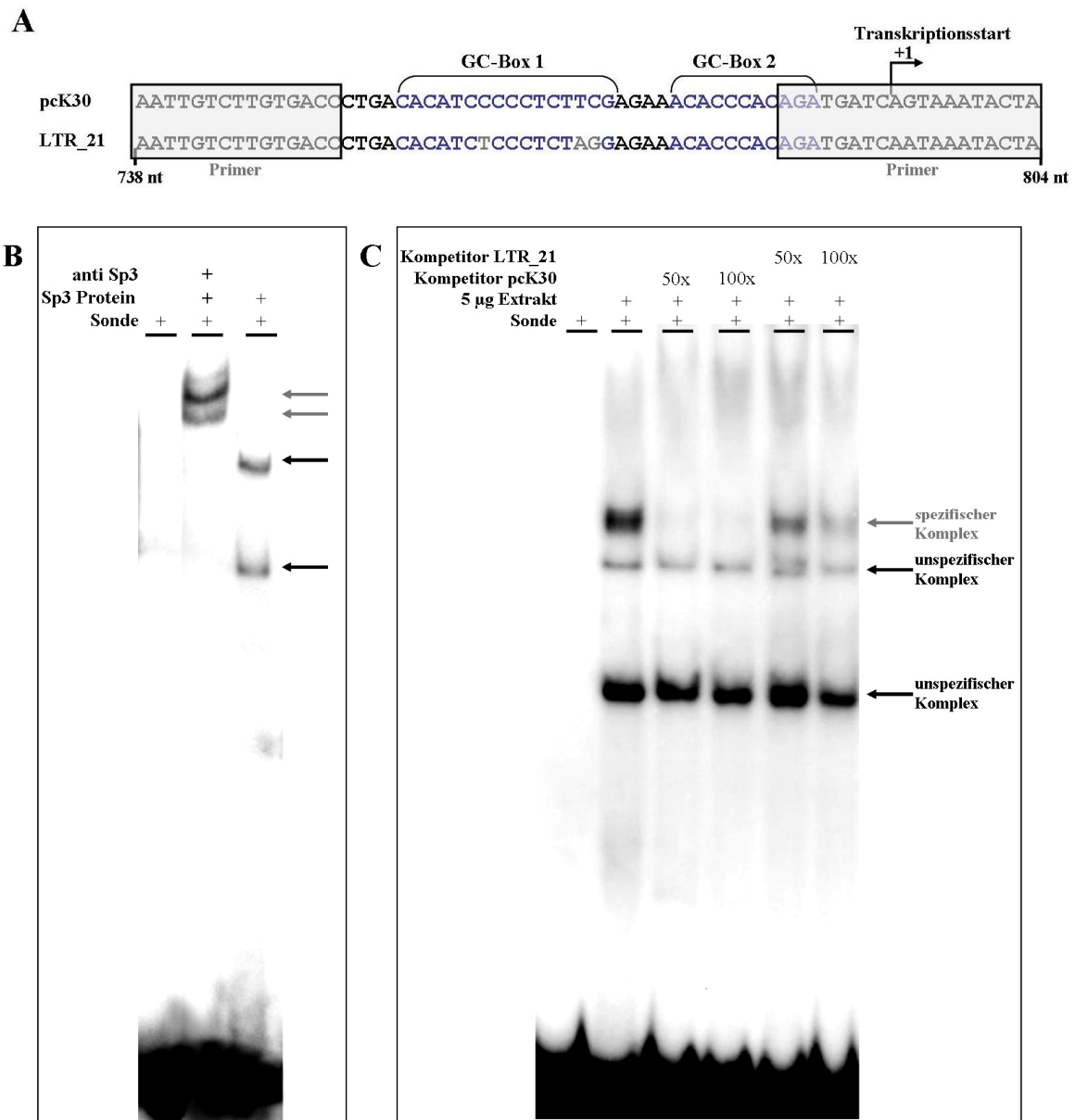


Abbildung 22: EMSA-Analyse des Transkriptionsinitiationsbereichs der HERV-K LTR. A) Schematische Darstellung der verwendeten Sonde mit einer Länge von 66 bp. Die putative GC-Box 1 ist in der LTR 21 durch drei Punktmutationen in ihrem Motiv verändert. Die zweite GC-Box ist sowohl in der pcK30 als auch in der LTR 21 vorhanden. Die zur Amplifikation eingesetzten Primer sind farblich markiert (siehe auch 8.1) B) Spur 1: radioaktiv markierte pcK30 Sonde, Spur 2: Inkubation der Sonde mit 2 µl rekombinatem Sp3-Protein (ALX-201-106, Alexis) erzeugt zwei Shift Banden (schwarze Pfeile), Spur 3: Zugabe von 2 µg spezifischem Sp3 Antikörper zu den Komplexen aus Sonde und rekombinatem Sp3 Protein induziert einen deutlichen Supershift der beiden Banden (graue Pfeile). C) Spur 1 radioaktiv markierte pcK30 Sonde, Spur 2: Inkubation der Sonde mit 5 µg Kernextrakt aus GH Zellen, Spur 3 bis Spur 6: Konkurrenz mit 50-fachem beziehungsweise 100-fachem Überschuss unmarkierter Sonde pcK30 oder LTR 21.

Die beiden anderen Komplexe ließen sich nicht kompetieren und sind somit als unspezifisch zu bezeichnen. Dieses Wettbewerbsverhalten stellte sich für GH und UKRV Mel 2-C9 Proteinextrakt identisch dar. Durch Zugabe von spezifischen Sp1

Antikörper während der Inkubation der Kernproteinextrakte aus GH Zellen mit der markierten DNA-Sonde konnte bei dem hochmolekularen spezifischen Komplex ein deutlicher Supershift in der nativen Gelelektrophorese erzielt werden (Abb.: 23A, Spur 4). Dies war durch Zugabe eines anti-Aktin Antikörpers nicht möglich (Spur 3). Die unspezifischen Komplexe veränderten sich nicht. Bei Zugabe von Sp3 spezifischen Antikörpern bildeten sich sogar zwei Shiftbanden aus (Spur 5). Das Ergebnis deutet darauf hin, dass der hochmolekulare spezifische Komplex sowohl Sp1 als auch Sp3 Protein enthält.

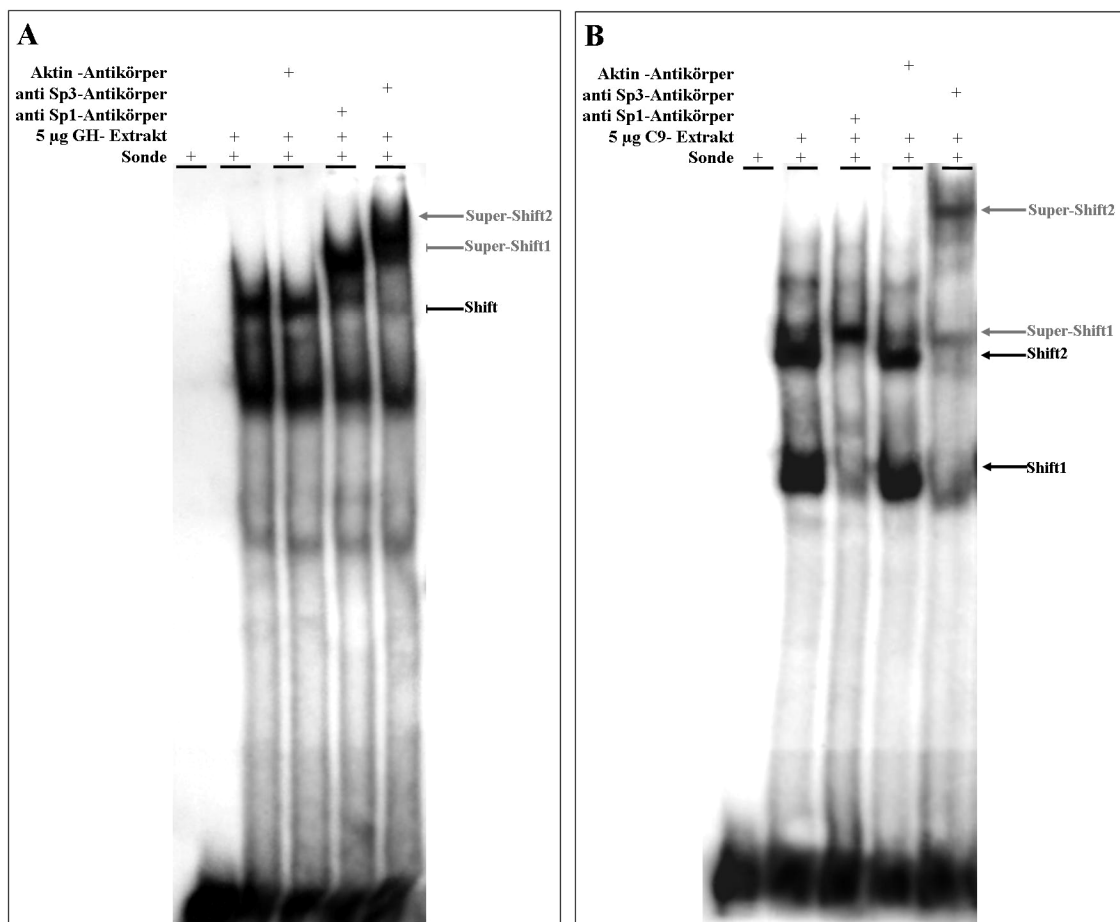


Abbildung 23: Ergebnisse der EMSA-Analysen mit nukleären Extrakten aus Keimzelltumor und Melanomzelllinien. **A)** Spur 1: radioaktiv markierte pcK30 Sonde, Spur 2: Inkubation der Sonde mit 5 µg Kernextrakt aus GH Zellen, Spur 5 und 4: Bei Inkubation der Komplexe aus Sonde und Kernextrakt mit 2 µg spezifischem Antikörper gegen Sp1 beziehungsweise Sp3 konnte ein Supershift des hochmolekularen Komplexes detektiert werden (graue Pfeile), Spur 3: Inkubation der Komplexe aus Sonde und Kernextrakt mit 2 µg Aktinantikörper. **B)** Spur 1: radioaktiv markierte pcK30 Sonde, Spur 2: Inkubation der Sonde mit 5 µg Kernextrakt aus UKRV Mel 2-C9 Zellen, Spur 4: Inkubation der Komplexe aus Sonde und Kernextrakt mit 2 µg Aktinantikörper. Spur 3 und Spur 5: Bei Inkubation der Komplexe aus Sonde und Kernextrakt mit 2 µg spezifischem Antikörper gegen Sp1 beziehungsweise Sp3 konnte ein Supershift des hochmolekularen Komplexes detektiert werden (graue Pfeile), Der niedermolekulare Komplex wird durch die Antikörper kompetitiert.

Bei Inkubation von nukleärem Melanomextrakt mit der radioaktiv markierten DNA-Sonde zeigten sich in der nativen Gelelektrophorese zwei Komplexe (Abb.: 23B, Spur 2, Shift 1 und 2). Bei Zugabe von spezifischem Sp1 Antikörper verschwanden beide Shifts und es bildete sich ein deutlich retardierter Komplex (Spur 3, grauer Pfeil Supershift 1). Bei Zugabe von spezifischem Sp3 Antikörper zeigten sich zwei Supershiftbanden (Abb.: 23B, Spur 5, graue Pfeile). Wieder scheint es, dass beide Komplexe sowohl Sp1 als auch Sp3 Proteine enthalten. Bei Zugabe von Aktinantikörper konnte kein Super-Shift beobachtet werden (Spur 4).

Die Resultate der EMSA-Studien zeigen, dass die hochkonservierte GC-Box 1 in der Nähe des Transkriptionsstarts der HERV-K LTR als Bindungsstelle für die Transkriptionsfaktoren Sp1 und Sp3 dient.

2.5.1.3. Mutation der Sp1/Sp3-Bindungsstelle in der GC-Box 1 reduziert *in-vivo* signifikant die LTR Promotoraktivität

Durch die EMSA-Studien konnte bestätigt werden, dass Sp1 und Sp3 in der Nähe des Transkriptionsstarts an die HERV-K LTR binden. Im Luziferasereporterassay (4.7.5.) wurde daher getestet, ob die Konvertierung der Sequenz der GC-Box1 der aktiven pcK30 (ACACATCCCCCTCTTCG) in die Sequenz der nicht aktiven LTR 21 (ACACATCTCCCTCTAGG) eine Veränderung der Promotorstärke bewirkt. Die Mutation dieser Sp1/Sp3 Bindestelle führte sowohl in GH als auch in UKRV Mel 2-C9 Zellen zu einem drastischen Abfall der Promotoraktivität (Abb.: 24). In GH Zellen sank die Aktivität um mehr als die Hälfte im Vergleich zum Wildtyp (pcK30). In der Melanomlinie war dieser Effekt um ein Vielfaches ausgeprägter. Die Aktivität fiel auf etwa 20 % ab.

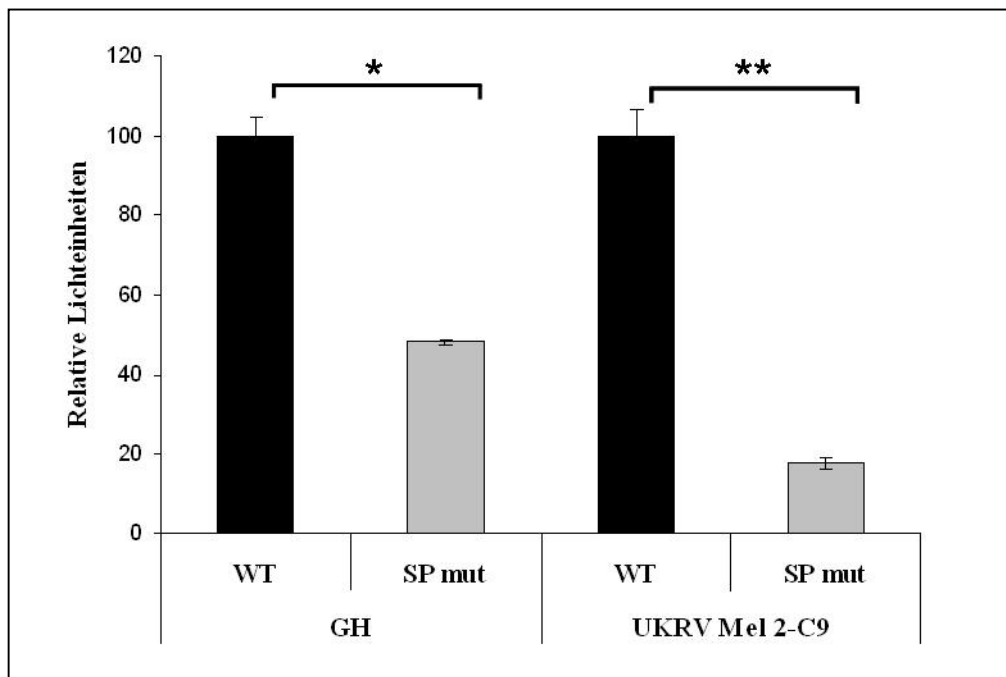


Abbildung 24: Relative Luziferaseaktivität einer LTR Mutante mit verändertem Sp-Bindemotiv (Sp mut) im Vergleich zum Wildtyp (WT). GH und UKRV Mel 2-C9 Zellen wurden entweder mit einem Luziferasereporter gen unter der Kontrolle des HERV-K Wildtyp-Promotors (schwarze Balken) oder mit einem Luziferasereporter gen unter der Kontrolle eines HERV-K Promotors, in dem die GC-Box 1 zur Sequenz der inaktiven LTR 21 mutiert worden war (graue Balken), transfiziert. Die relativen Lichteinheiten, die ein Maß für die Promotoraktivität sind, wurden 48 Stunden nach der Transfektion bestimmt und auf die kotransfizierte interne Renilla Kontrolle normalisiert. Die Promotoraktivität des Wildtyps wurde für jede Zelllinie als 100 % gesetzt. Die aufgetragenen Daten sind die Mittelwerte mit Standardabweichung aus sechs parallelen Ansätzen eines repräsentativen Experiments. Das Mutationskonstrukt zeigt signifikante Unterschiede im Vergleich zum Wildtyp [p^* und $p^{**}<0,001$, Student's t-test].

2.5.1.4. siRNA vermittelter Knockdown von Sp1 and Sp3 reduziert die HERV-K Promotoraktivität

Um zu bestätigen, dass die Transkriptionsfaktoren Sp1 und Sp3 Auswirkungen auf die transkriptionelle Aktivität des HERV-K Promotors haben, wurde die Translation jeweils eines der beiden Faktoren mittels transfizierter siRNA in den zu untersuchenden Zellen unterdrückt (Durchführung des Assays 4.7.5.2.). Die Funktionalität der verwendeten siRNA-Oligos wurde durch Immunoblotanalysen bestätigt (Abb.: 25A und 25B). Bei Verwendung von 150 bzw. 300 ng siRNA pro $0,5 \times 10^6$ GH oder UKRV Mel 2-C9 Zellen konnten sowohl der Gehalt an Sp1 als auch das Niveau der langen und kurzen Isoformen von Sp3 in der Linie UKRV Mel 2-C9 deutlich reduziert werden (Abb.: 25A und B). Als Kontrolle dienten mit Transfektionsreagenz behandelte Zellen. Um die Auswirkungen des jeweiligen *Knockdowns* auf die Promotoraktivität zu bestimmen, wurde jeweils eine der beiden siRNAs mit dem pcK30 LTR-Luziferasekonstrukt in GH und UKRV Mel 2-C9 Zellen kotransfiziert.

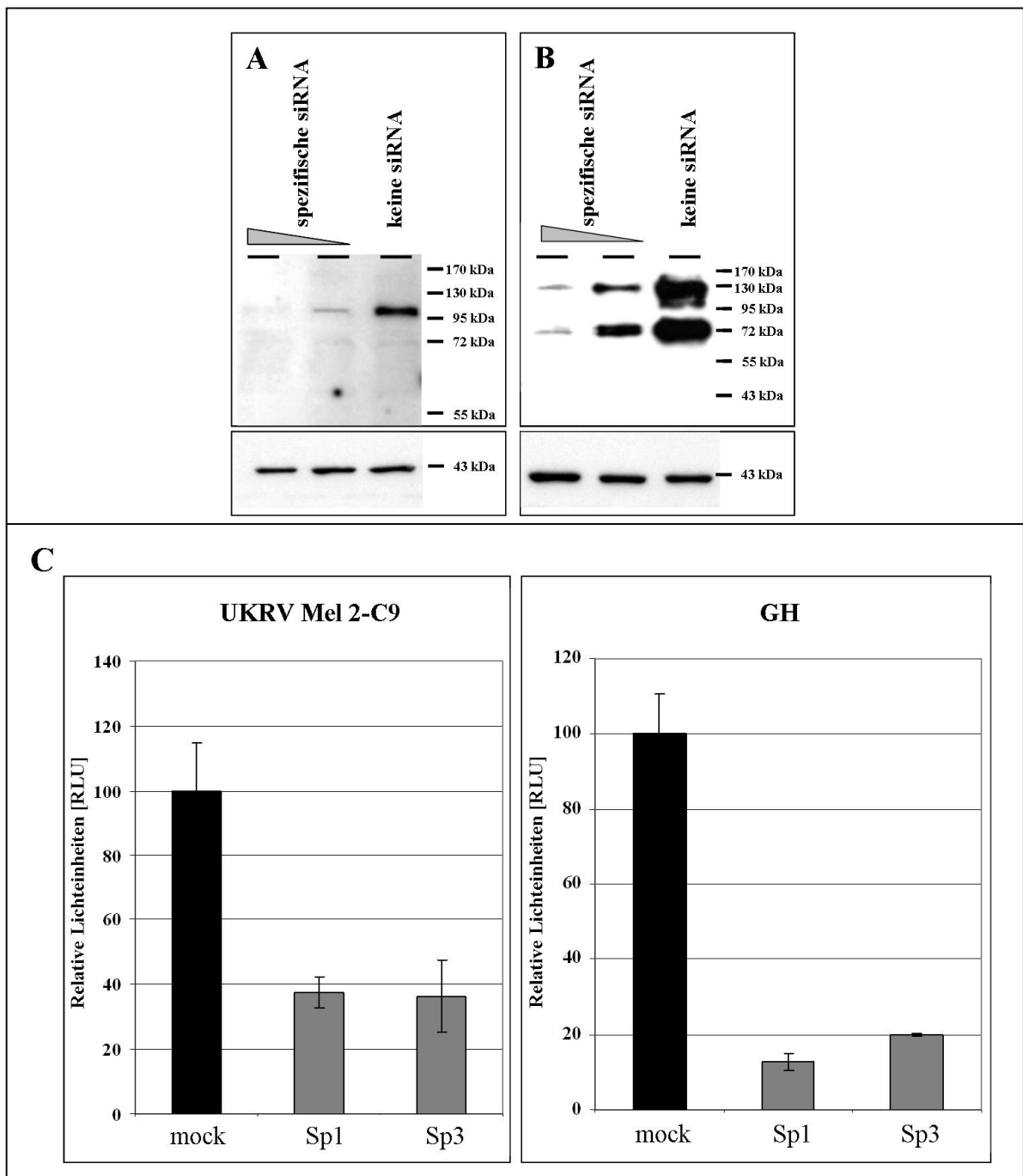


Abbildung 25: siRNA vermittelter Sp1 und Sp3 Knockdown vermindert die Promotoraktivität der HERV-K LTR A und B) 300 ng oder 150 ng Sp3 spezifischer siRNA wurden in $0,5 \times 10^6$ UKRV Mel 2-C9 Zellen transfiziert. Nach 48 Stunden wurden Lysate mit dem Promega Lysepuffer hergestellt. Sp1 und Sp3 wurden im Immunoblot mit spezifischen Antikörpern nachgewiesen. C) $0,5 \times 10^6$ UKRV Mel 2-C9 oder GH Zellen wurden mit 1 μ g pcK30-Luziferasevektor (mock, Spur 1) beziehungsweise mit 1 μ g pcK30-Luziferasevektor und 300 ng Sp1 (Spur 2) oder 300 ng Sp3 (Spur 3) spezifischer siRNA transfiziert. Nach 48 Stunden wurden die Zellen mit Promega Lysepuffer lysiert (4.7.1.) und die Luciferaseaktivität des jeweiligen Lysats als Maß für die Promotorstärke bestimmt. Alle Werte wurden auf den Proteingehalt normiert und sind als relative Lichteinheiten dargestellt. Die mock-Kontrolle wurde als 100 % gesetzt. Gezeigt sind die Mittelwerte mit Standardabweichung aus drei parallelen Ansätzen eines repräsentativen Experimentes.

Als Kontrolle dienten Zellen, die analog behandelt wurden, aber nur mit dem Luziferasekonstrukt transfiziert wurden („mock“). Da die Kotransfektion eines Renilla-Kontrollplasmids den Assay störte, wurden die ermittelten Luziferasewerte auf den jeweiligen Proteingehalt des Lysats normiert. Es konnte ein deutlicher Abfall der Promotoraktivität in allen siRNA-Ansätzen beobachtet werden (Abb.: 25C). Im Vergleich zur „mock“-Kontrolle sank die Aktivität der pcK30-LTR in der Linie UKRV Mel 2-C9 bei Verwendung der Sp1-siRNA ebenso wie bei Einsatz der Sp3-siRNA auf ~ 40 % ab. In GH Zellen war der Abfall der Promotoraktivität sogar noch wesentlich deutlicher. Bei Verwendung von Sp1 und Sp3 spezifischer siRNA wurden gegenüber dem „mock“-Wert nur noch 15 – 20 % der relativen Lichteinheiten gemessen. Dieses Experiment bestätigt, dass sowohl Sp1 als auch Sp3 an der Regulation der transkriptionellen Aktivität des HERV-K Promotors in UKRV Mel 2-C9 und GH Zellen beteiligt sind.

2.5.1.5. Bindung von Sp1 und Sp3 an den HERV-K Promotor (*in vivo* ChIP-Assay)

Die Bindung von Sp1 und Sp3 an die HERV-K LTR sollte zur Bestätigung der EMSA-Analysen mittels Chromatinimmunopräzipitation (ChIP) *in vivo* nachgewiesen werden (4.8.4.). Es wurden $0,5 \times 10^6$ GH, UKRV Mel 2-C9 und HCl Mel 19 Zellen pro Ansatz eingesetzt. Durch zwanzigminütige Inkubation in einer 1 %igen Formaldehydlösung in Vollmedium wurden DNA und Proteine vernetzt. Die Zellen wurden anschließend in SDS-Lysepuffer aufgenommen wodurch der DNA-Proteinkomplex aus dem Zellkern freigesetzt wurde.

Die Analyse definierter Proviren war leider nicht möglich. Ein lokusspezifischer Primer, zur Detektion der LTR nach der Präzipitation, kann auf Grund des proviralen Integrationsortes oft erst 2 kb stromaufwärts definiert werden. Eine effiziente Chromatin-IP mit DNA-Fragmenten dieser Größe ist technisch fast unmöglich. Die DNA wurde daher mittels spezifischer Restriktionsenzyme geschnitten (Abb.: 25A). Da die Sequenzen der HERV-K Proviren bekannt waren (2.2.2.), konnten Restriktionsenzyme unter Berücksichtigung der spezifischen Basenunterschiede zwischen den Proviren gewählt werden. Jede LTR enthält stromaufwärts des Transkriptionsstarts entweder eine Erkennungsstelle für das Restriktionsenzym NspI oder durch Abweichung eines Nukleotids an gleicher Stelle eine Erkennungssequenz für AflIII. HindIII schneidet in allen

Proviren in der 5' untranslatierten Region vor dem *gag* Gen. Bei Zugabe aller drei Enzyme in den Ansatz entstanden auf diese Weise Fragmente mit einer Größe von ~370 bp.

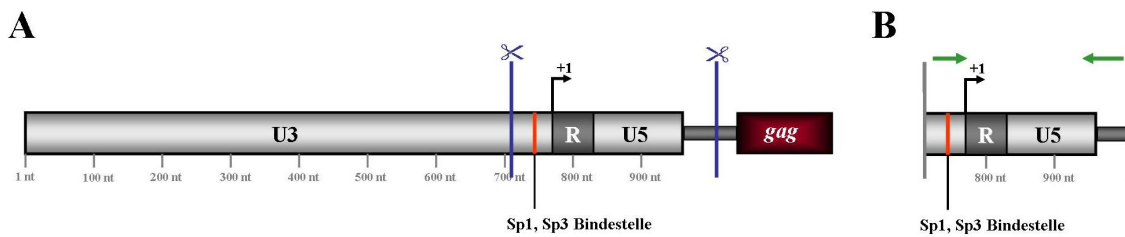


Abbildung 26: Schema zur Chromatin Immunpräzipitation der HERV-K LTR mittels spezifischer Antikörper. A) Das HERV-K Genom wurde mit Hilfe von spezifischen Restriktionsenzymen geschnitten; eine Restriktionsschnittstelle ist in der LTR lokalisiert, die zweite befindet sich in der Nähe der Primerbindungsstelle. B) Die Amplifikation aus dem Präzipitat erfolgte mittels spezifischer Primer (grüne Pfeile, nicht maßstabsgetreu).

Da das menschliche Genom eine Vielzahl von solitären LTRs enthält wurden die Primer zur Amplifikation der präzipitierten DNA so gewählt, dass nur ein Produkt entsteht, wenn eine Primerbindungsstelle vorhanden ist (Abb.: 26B). Auf diese Weise wurden Signale von solitären LTRs ausgeschlossen. Die durch die Restriktion entstandenen Fragmente der HERV-K LTR konnten mittels spezifischer Antikörper gegen Sp1 und Sp3 in allen verwendeten Zelllinien präzipitiert und mittels PCR nachgewiesen werden (Abb.: 27). In den HERV-K positiven Linien UKRV Mel 2-C9 und GH bindet deutlich mehr Sp3 als Sp1 an den Promotor. In der Melanomlinie HCl Mel 19 sind keine HERV-K Proviren transkriptionel aktiv. Hier besteht ein ausgewogenes Verhältnis in der Bindung von Sp1 und Sp3 an die LTR. Bei Verwendung eines Aktinantikörpers konnten dagegen keine Signale nachgewiesen werden (Abb.: 27). Ebenso traten in Ansätzen ohne Antikörperzugabe keine Signale auf. Dies zeigt, dass die Signale bei Verwendung der Sp1 und Sp3 Antikörpern spezifisch sind. Zur Kontrolle wurde am Ausgangsmaterial der Immunpräzipitation eine PCR durchgeführt, in der wie erwartet deutliche Signale auftraten. Bei Verwendung eines Primerpaares aus dem Bereich von HERV-K *env* konnten lediglich im Eingangsmaterial und nicht in den Präzipitaten Signale detektiert werden (Daten nicht gezeigt). Die Chromatinimmunpräzipitation zeigt, dass Sp1 und Sp3 *in vivo* an den HERV-K Promotor binden.

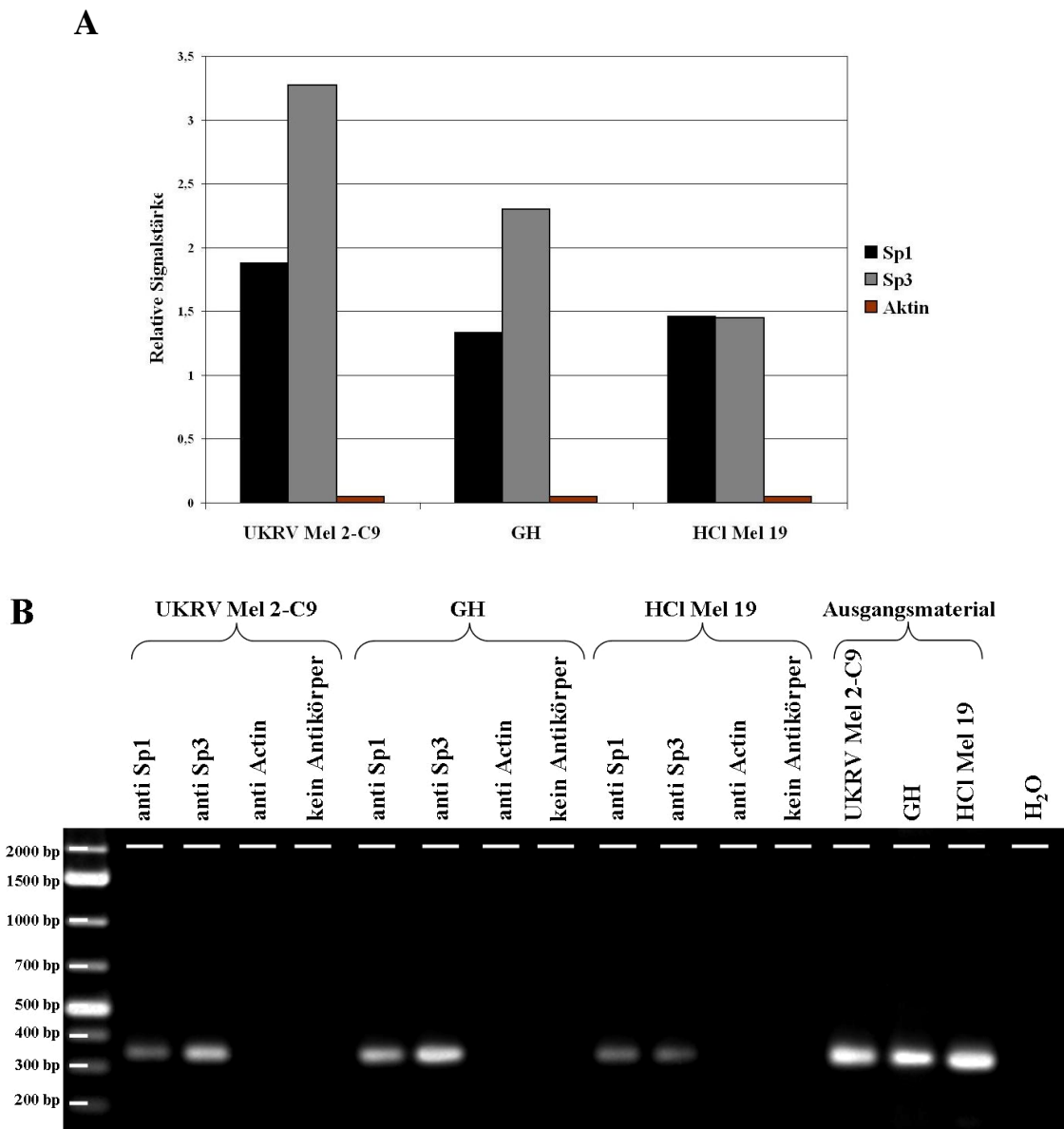


Abbildung 27: Ergebnis und Auswertung der Chromatinimmunopräzipitation A) Quantifizierung der PCR-Signale nach Chromatin Immunopräzipitation der HERV-K LTR. Normiert wurde auf die Signalstärke des geschnittenen Eingangsmaterials. B) Amplifikate aus dem Bereich der HERV-K LTR nach der Chromatin Immunopräzipitation mit jeweils 2 µg spezifischem Antikörper gegen Sp1, Sp3 und Aktin.

3. DISKUSSION

3.1. Expression von HERV-K in der Melanomlinie UKRV Mel 2-C9

Eine erhöhte transkriptionelle HERV-K Aktivität war für Keimzelltumoren (Herbst *et al.*, 1996), Keimzelltumorklinien (Löwer *et al.*, 1993; Sauter *et al.*, 1995) und einige Melanomlinien (Muster *et al.*, 2003) bereits bekannt. Die in dieser Arbeit gewonnenen Ergebnisse zeigen, dass durch Einzelzellklonierung eine Melanomlinie, UKRV Mel 2-C9, mit hoher HERV-K Expression isoliert werden konnte. Ein Vergleich der Signale der Immunfluoreszenzanalysen (Abb.: 8) macht jedoch deutlich, dass die Keimzelltumorklinie GH die viralen Gag Proteine deutlich stärker als die Melanomlinie exprimiert und auch GH vermehrte Virusproduktion zeigt. Dies konnte durch Immunoblotanalysen bestätigt werden. Aus gleichen Mengen von Gesamtproteinextrakten konnte deutlich mehr Gag Protein in GH Zellen detektiert werden als in Extrakten der Melanomlinie (2.1.2.). Virale Partikel werden sowohl von der Linie UKRV Mel 2-C9 als auch von der Linie GH produziert und in das Nährmedium abgegeben. Das konnte durch Immunoblotanalysen der jeweiligen Zellkulturüberstände (2.3.) gezeigt werden. Die HERV-K Expression der Linie UKRV Mel 2-C9 ist beinahe ausschließlich auf die Aktivität des Provirus HERV-K 108 zurück zu führen, damit bietet diese Linie die Möglichkeit, den Phänotyp dieses spezifischen Provirus zu charakterisieren.

Das Gag Vorläuferprotein wird in der Melanomlinie UKRV Mel 2 –C9 ebenso wie in der Keimzelltumorklinie GH proteolytisch gespalten. In beiden Zelllinien wird also eine aktive HERV-K Protease exprimiert, die in der Lage ist, das Gag Vorläuferprotein zu prozessieren. Das prozessierte Kapsid kann sowohl im Gesamtzellextrakt als auch in Partikelpräparationen detektiert werden. Die virale Protease wird erst bei der Reifung der viralen Partikel aktiviert. Dass auch im Gesamtzellextrakt Signale für die gespaltenen Kapsidproteine detektiert werden können, ist auf virale Partikel zurückzuführen, die sich zum Zeitpunkt der Zelllyse noch im Knospungsstadium an der Zelloberfläche befanden und bei der Herstellung der Proteinextrakte gemeinsam mit den Zellen lysiert wurden. Im Zellextrakt aus UKRV Mel 2-C9 und GH Zellen lassen sich neben dem Kapsid (30 kDa) und dem Gag Vorläuferprotein (80 kDa) auch einige Spaltintermediate detektieren (Abb.: 9). Im Vergleich zu GH Zellen ergibt sich ein deutlich abweichendes Bandenmuster in der untersuchten Melanomlinie. In GH Zellen sind neben Vorläuferprotein und Kapsid mindestens fünf Intermediate detektierbar, in der Linie UKRV Mel 2-C9 dagegen nur drei.

Das Bandenmuster der Melanomlinie UKRV Mel 2-C9 wird vor allem durch die Aktivität des Provirus HERV-K 108 bedingt. In GH Zellen dagegen sind neben HERV-K 108 drei weitere Proviren (HERV-K 107, HERV-K 101 und HERV-K II) aktiv. Diese heterogene Virenpopulation enthält somit provirale Proteine unterschiedlicher Sequenz, die eventuell nur unvollständig prozessiert werden, beispielsweise auf Grund veränderter Proteaseschnittstellen. Die von Typ 1 Proviren (HERV-K 101, HERV-K II) gebildeten Gag Proteine können auch verkürzt sein, denn diese Sequenzen weisen oft vorzeitige Stop-Codons auf. Das Verbleiben der viralen Partikel im Knospungsstadium, das vor allem in GH Zellen zu beobachten ist, und die überwiegende Anwesenheit von Partikeln mit unvollständig kollabierten *Core*strukturen deutet ebenfalls auf eine heterogene Population von exprimierten Proviren hin.

In Partikelpräparationen aus Zellkulturüberstand von UKRV Mel 2-C9 konnte lediglich gespaltenes Kapsid detektiert werden, was typisch für virale Partikel ist, in denen Gag vollständig prozessiert wird. Dies könnte bedeuten, dass das Provirus 108 für ein intaktes *gag* Gen kodiert. Da sich aber auch hier die Partikel nicht recht von der Membran lösen, müssen andere Defekte in Betracht gezogen werden. Dies ist zu einem späteren Zeitpunkt Gegenstand der Diskussion.

Die viralen Hüllproteine sind verantwortlich für die Erkennung zellulärer Rezeptoren und vermitteln die Bindung des Virus an diese Oberflächenmoleküle. Auf diese Art ermöglichen sie den Eintritt des Virus in die Zelle. Die Infektiosität der retroviralen Partikel wird daher wesentlich von den Env Proteinen bestimmt.

Der Nachweis des Env Vorläuferproteins (90 kDa) und der gespaltenen SU-Untereinheit (60 kDa) mit einem HERV-K Env-SU-Antikörper war aus Zellextrakten von UKRV Mel 2-C9 und GH Zellen möglich (Abb.: 9). Das Env Vorläuferprotein wird von einer zellulären Protease in SU- und TM-Untereinheit gespalten. Sowohl in der Melanomlinie als auch in der Teratokarzinomlinie GH wird Env offensichtlich von der Zelle korrekt prozessiert. Verglichen mit der Aktin-Ladkontrolle und der unspezifischen Bande bei 30 kDa waren bei UKRV Mel 2-C9 die Env Signale des Zellextrakts stärker als in GH Zellen. In den Melanomzellen scheint ein ausgewogenes Verhältnis zwischen der Produktion viraler HERV-K Gag und Env Proteine zu bestehen, während in GH Zellen deutlich mehr Gag als Env produziert wird. Dieses wenig ausgewogene Verhältnis in GH Zellen ist auf eine vermehrte Aktivität proviraler Loci, die nur einen offenen Gag Leserahmen besitzen, zurück zu führen (2.2.2.). HERV-K 108 Partikel der Melanomlinie scheinen, wie der Immunoblot (2.3.) aus Zellkulturüberstand zeigt, vollständig prozessierte

Env Proteine zu enthalten. Auch weist die Partikelmorphologie im Elektronenmikroskop (2.3.1.) deutliche *Spike*-Strukturen auf, wie sie für Betaretroviren üblich sind. Vergleichbare Analysen retroviraler Partikel aus GH Zellen zeigten, dass dort vergleichsweise wenig *Spikes* auf der Oberfläche auftreten, die gebildeten Partikel enthalten weniger Env Glykoproteine (Bieda *et al.*, 2001). Dieser Befund ist konsistent mit den Resultaten der Immunoblotanalysen. Der verminderte Einbau von Env Proteinen in GH Partikel beruht auf der Expression verschiedener HERV-K Proviren, während die von der Linie UKRV Mel 2-C9 gebildeten Partikeln den Phänotyp des Provirus HERV-K 108 zeigen, bei dem Gag und Env in gleichem Umfang translatiert werden.

In mehreren Keimzelltumorkolonien konnte die Expression von Typ 1 und Typ 2 HERV-K Proviren gezeigt werden (Löwer *et al.*, 1993). Diese Ergebnisse konnten für die Keimzelltumorkolonien NCCIT und GH im Rahmen der vorliegenden Arbeit bestätigt werden. Eine Zuordnung der HERV-K Transkripte zu aktiven Loci in den jeweiligen Zelllinien erfolgte in den meisten vorangegangenen Untersuchungen nicht und wurde im Rahmen dieser Arbeit erstmals umfangreich durchgeführt. Einige der HERV-K HML-2 Proviren zeigen Basenunterschiede in der Sequenz der zur Amplifikation eingesetzten *env*-Primer. Die von einer Mutation betroffenen Proviren wären daher in der PCR möglicherweise nicht ausreichend vervielfältigt und somit nicht berücksichtigt worden. Alle HERV-K HML-2 Proviren mit einem intakten ORF für Gag Proteine zeigen keinerlei Basenunterschiede in der Sequenz der verwendeten *gag*-Primer und sollten daher vollständig in der Analyse enthalten sein.

Die Untersuchung der partikulär verpackten RNA-Genome (2.3.2.) aus UKRV Mel 2-C9 ergab, dass neben HERV-K 108 zu einem sehr geringen Anteil auch Provirus HERV-K 101 transkribiert wird. Da zur Analyse der partikulär verpackten Genome vergleichsweise deutlich mehr Klone analysiert wurden als zur Identifikation reaktiver Proviren in der Analyse von zytoplasmatischer Poly-(A)plus RNA, sind schwache Signale transkriptionel wenig aktiver HERV-K Proviren nicht aus zu schließen. Da aber eine so geringe Aktivität kaum Einfluss nehmen wird auf die Gesamtheit der gebildeten HERV-K Proteine, können solch geringe Transkriptionsniveaus vernachlässigt werden. Selektives Verpacken der retroviralen RNA wird vermittelt durch Interaktion spezifischer viraler RNA-Segmente des viralen Genoms (Verpackungssignal, ψ -Motiv), mit Bereichen des Nukleokapsids während der Aggregation der Gag Polyproteine. Das Verpackungssignal der HML-2 Familie ist noch nicht bekannt, scheint aber im Provirus 108 funktional zu sein. Die von der Linie GH produzierten Partikel enthalten hauptsächlich (64 %) RNA-Genome

des Provirus 102 neben Genomen des Provirus 108 (33 %). Von der Teratokarzinomlinie Tera 1 produzierte Partikel enthalten ebenfalls hauptsächlich RNA-Genome von HERV-K 102 (Ruprecht *et al.*, 2008). Dieses Provirus hat offensichtlich auch ein funktionales Verpackungssignal. In all diesen Analysen ist nicht aus zu schließen, dass die viralen Partikel im Zuge der Aufreinigung mit Ribosomen und anderen Mikrovesikeln verunreinigt wurden. Daher könnten auch Membranfragmente mit viralen Bestandteilen, die nicht aus vollständigen Partikeln stammen, zu einem geringen Anteil in die Analyse eingegangen sein.

Nach den Informationen der NCBI-Datenbank weist das HERV-K Provirus 108 offene Leserahmen für alle retroviralen Proteine auf und zeigt lediglich eine Mutation im konservierten YXDD-Motiv der Reversen Transkriptase, welche die enzymatische Aktivität des Enzyms verhindert (Barbulescu *et al.*, 1999; Mayer *et al.*, 1999; Tönjes *et al.*, 1999). In einigen Individuen konnte jedoch eine intakte Reverse Transkriptase-Domäne für das Provirus identifiziert werden (Reus *et al.*, 2001).

Die Untersuchung der Reversen Transkriptase-Domäne von HERV-K 108 in der Linie UKRV Mel 2-C9 ergab, dass im Vergleich zum Datenbankeintrag keine Sequenzveränderung vorliegt. Das aktive Zentrum des Enzyms ist inaktiv. Dagegen weist das in sehr geringe Mengen transkribierte Provirus HERV-K 101 ein intaktes YXDD-Motiv auf. Die Expression in der betreffenden Melanomlinie führt wahrscheinlich nicht zur Produktion ausreichender Mengen an funktionaler Reverser Transkriptase um eine Infektiosität vermitteln zu können. Mit einem Enzymaktivitätstest soll diese Fragestellung zukünftig abgeklärt werden.

HERV-K 108 liegt im Wirtsorganismus gewöhnlich als Dublette zweier aufeinander folgender Proviren in gleicher Orientierung mit einer gemeinsamen LTR in der Mitte vor (Reus *et al.*, 2001). Das Provirus kann allerdings auch als vollständiges Einzelintegrat vorkommen (Macfarlane und Simmonds, 2004; Mayer *et al.*, 2005). Erste durch PCR gewonnene Erkenntnisse (Daten nicht gezeigt) legen nahe, dass aus den Linien UKRV Mel 2-C9 und GH die Amplifikation eines Produktes von *env* nach *gag* möglich ist. Das entstehende Produkt weist die für ein Tandemintegrat mit einer zentralen LTR erwartete Größe auf. Es muss allerdings in weiterführenden Untersuchungen festgestellt werden, ob diese Signale wirklich dem Provirus 108 zuzuordnen sind.

Bis zum jetzigen Zeitpunkt konnten *in vivo* keine infektiösen HERV-K Partikel identifiziert werden. Alle von Teratokarzinom- und Melanomlinien produzierten HML-2 Viruspartikel scheinen Defekte auf zu weisen (Löwer *et al.*, 1984; Löwer *et al.*, 1996;

Bieda *et al.*, 2001; Boller *et al.*, 2008). Es gibt verschiedene Erklärungsmöglichkeiten für den Verlust der viralen Eigenschaften. Über die Möglichkeit eines Defekts im Polymerasegen wurde bereits weiter oben diskutiert. Dass HERV-K 108 für ein intaktes Env kodiert, wurde von der Heidmann Gruppe gezeigt. Das Env Protein war in der Lage, lentivirale Vektoren zu pseudotypisieren (Dewannieux *et al.*, 2005). Die fehlende Infektiosität der UKRV Mel 2-C9 Partikel wird daher wahrscheinlich nicht durch einen Defekt im Env Protein bedingt.

GH Partikel könnten auf Grund eines Defektes in der Prozessierung von Env oder ausgelöst durch Mutationen im Bereich von Gag weniger Glykoproteine enthalten. Für HIV konnte bereits gezeigt werden, dass Mutationen im Matrixprotein den Einbau von Env Proteinen in die viralen Partikel behindern können (Yu *et al.*, 1992). Auch die Zusammenlagerung von intakten und defekten Gag Proteinen könnte ein Grund für die fehlende Infektiosität sein. Der Einbau eines geringen Anteils nicht funktionaler Gag-Proteine führt in HI-Viren zur Bildung von nicht infektiösen viralen Partikeln (Zybarth *et al.*, 1994). Veränderungen im Bereich von Gag können sich auch auf die Morphologie der viralen Partikel auswirken. Gag Proteine von C-Typ Retroviren wie *moloney murine leukemia virus* (MoMuLV) assemblieren und knospen an der Plasmamembran der Wirtszelle. Bei B und D-Typ Viren wie *mason pfizner monkey virus* (MPMV) und *mouse mammary tumor virus* (MMTV) assemblieren die Gag Polyproteine bereits im Zytoplasma zu Kapsiden und werden dann zur Zellmembran transportiert, von der sie sich abschnüren. Es konnte gezeigt werden, dass für die Assemblierung im Zytoplasma ein 18 Aminosäuren umfassendes Sequenzmotiv verantwortlich ist, das in MPMV und MMTV konserviert ist (Choi *et al.*, 1999). Eine Mutation von Arg55 im Bereich der Matrix von MPMV führt zu einem veränderten Buddingverhalten. Die Partikel bilden nicht länger *Core*strukturen im Zytoplasma, sondern assemblieren erst an der Plasmamembran. HERV-K zeigt auf Grund seiner genomischen Organisation deutliche Parallelen zu D-Typ Viren und wird heute wie MMTV als Betaretrovirus klassifiziert. Es wäre daher zu erwarten, dass bei HERV-K bereits im Zytoplasma der Zelle eine Zusammenlagerung der Gag Polyproteine stattfindet. Dieser Phänotyp konnte jedoch bisher nicht beobachtet werden. Es ist denkbar, dass HERV-K die Fähigkeit zur Ausbildung von *Core*strukturen im Zytoplasma durch die Akkumulation von Mutationen im Verlauf der Zeit verloren hat. Offensichtlich ist das Abschnüren der Partikel von der Zellmembran der Wirtszelle verzögert, besonders deutlich wird dies in elektronenmikroskopischen Schnitten von GH Zellen (Bieda *et al.*, 2001). Zudem enthalten die viralen Partikel selten vollständig kollabierte *Core*strukturen. Ebenso

besitzen nur sehr wenige der von der Linie UKRV Mel 2-C9 produzierten Partikel eine ausgereifte vollständig kollabierte *Core*struktur.

Die Freisetzung viraler Partikel aus der Wirtszelle ist oft abhängig von bestimmten Aminosäuremotiven im Bereich von Gag. Für MPMV konnte gezeigt werden, dass die Aminosäuresequenz PPPY essentiell für die Freisetzung der Partikel ist (Yasuda und Hunter, 1998). Dieses Motiv liegt konserviert in *gag*-Genen einer Reihe von Retroviren vor. Wird dieses Motiv durch Mutationen verändert, können die viralen Partikel sich nicht korrekt von der Zelle lösen. Neben dem PPPY-Motiv kann auch ein PTAP-Motiv involviert sein zur korrekten Freisetzung viraler Partikel. HTLV-1 kodiert neben einem PPPY-Motiv im C-terminalen Bereich auch für ein PTAP Motiv im Bereich des Matrixproteins. Beide Motive vermitteln in die Abschnürung der Partikel von der Zellmembran (Wang *et al.*, 2004). HERV-K besitzt kein PPPY-Motiv im Bereich des *gag*-Gens. Weder die in dieser Arbeit als aktiv identifizierten Proviren, noch das reaktivierte HERV-K Phönix (Dewannieux *et al.*, 2006) hat eine solche Aminosäuresequenz. Ein PTAP-Motiv ist allerdings bei HERV-K vorhanden. Eventuell haben hier Mutationsereignisse zu einem Verlust des PPPY-Motivs geführt, wodurch die Freisetzung der viralen HERV-K Partikel zwar nicht komplett unterbunden aber doch stark verzögert werden könnte. Das HERV-K Phönixkonstrukt produziert auch nur sehr geringe Virustiter. Interessant ist zudem die Fragestellung warum der provirale Locus HERV-K 108 in fast jeder Zelle des Subklons UKRV Mel 2-C9 aktiv ist, während in der parentalen Zelllinie UKRV Mel 2 nur in fünf Prozent der Zellen HERV-K Expression auftritt. Bisher ist nicht untersucht worden, ob die Aktivität in diesem geringen Anteil der Zellpopulation ebenfalls von HERV-K 108 bedingt wird. Worin der Unterschied zwischen Ausgangslinie und Subklon liegt, muss in weiten Analysen geklärt werden. Entweder exprimiert die subklonale Linie einen Faktor der sich positiv auf die Expression von HERV-K 108 auswirkt, oder ein reprimierender Faktor, ist nicht mehr vorhanden oder in seiner Expression herabgesetzt. Diese Fragestellung könnte durch vergleichende *MikroArray*-Analysen zwischen parentaler Linie und Subklon beantwortet werden. Dabei sind vor allem Unterschiede in der Expression von Transkriptionsfaktoren und Signalmolekülen interessant.

Des Weiteren müssen epigenetische Unterschiede analysiert werden, die Einfluss auf die Aktivität des proviralen Locus nehmen könnten. Über die Regulation der Transkription von HERV-K ist bisher wenig bekannt. Um erste Erkenntnisse bezüglich der proviralen

Transkriptionssteuerung in Melanomen zu erhalten, wurden die LTR Bereiche in der subklonalen Linie im Vergleich zu Keimzelltumorklinen analysiert.

3.2. Transkriptionsregulation von HERV-K in Melanomen im Vergleich zu Teratokarzinomen

Die LTR des in Keimzelltumoren und Melanomen aktiven Provirus HERV-K 108 weist deutliche Ähnlichkeit zu den LTRs der in Keimzelltumoren reaktivierten Proviren (HERV-K 107, HERV-K 106, HERV-K 33, HERV-K 101, HERV-K 102, HERV-K II) auf. Stammbaumanalysen der LTRs aktiver Proviren zeigen eine nahe Verwandtschaftsbeziehung und unterscheiden sich deutlich von der im Luziferaseassay inaktiven LTR 21 (HERV-K 18). In Analysen zur HERV-K Expression konnte weder aus Keimzelltumoren noch aus Melanomen ein Transkript des Provirus HERV-K 18 detektiert werden. Es ist davon auszugehen, dass der Promotor auf Grund der Mutationen *in vivo* keine Aktivität in den untersuchten Linien zeigt. Akkumulationen von Punktmutationen können zu transkriptioneller Inaktivität der proviralen Promotoren führen. Die aktiven LTRs unterscheiden sich von der inaktiven HERV-K 18 LTR durch Punktmutationen, die in bestimmten Bereichen als Cluster auftreten (Alignent 2.5.).

Ob einige dieser Mutationen möglicherweise eine Rolle bei der Initiation der basalen HERV-K Transkription spielen könnten, wurde in dieser Arbeit untersucht und wird weiter unten diskutiert. Die Identifikation von Transkriptionsstart und Transkriptionseende mittels 5' und 3'RACE ergab, dass die Strukturbereiche der HERV-K LTR in der untersuchten Melanomlinie UKRV Mel 2-C9 gleichermaßen definiert sind wie in der Keimzelltumorklinie GH.

Die HERV-K LTR besteht aus einem vergleichsweise sehr großen U3-Bereich (Abb.: 17). Der Export der Vollängen und unvollständig gespleißten viralen Transkripte wird durch die Bindung viraler RNA-Adaptor-Proteine an charakteristische Bereiche der mRNA reguliert. Wie *in vivo* Funktionsanalysen gezeigt haben, benötigt das HERV-K Rec Protein zur Bindung an das *Rec responsive element* (RcRE) der viralen Transkripte, einen weitaus größeren Sequenzbereich als andere Export Adaptoren (Magin-Lachmann *et al.*, 2001). Zur Bindung benötigen die Proteine Rev und Rex lediglich zwanzig bis dreißig Nukleotide (Malim *et al.*, 1990; Gröne *et al.*, 1994). Rec benötigt zur Bindung an das RcRE zweihundertzehn Nukleotide. Dieser große Sequenzabschnitt ist im U3-Bereich der proviralen LTR lokalisiert.

Der U3-Bereich enthält am 5'-Ende außerdem eine Enhancersequenz mit der Bindungsstelle für den Transkriptionsfaktor YY1 (Knöbl *et al.*, 1999). Aus weiteren Analysen (M. Kraft, 2004) ist bekannt, dass Deletionen oder Mutationen im 5'-Bereich der U3-Region (nt 1 - nt 236) die transkriptionelle Aktivität der LTR im Luziferaseassay negativ beeinflussen können. Zu einem deutlichen Abfall der Promotoraktivität führt neben der Deletion auch die Substitution dieses Bereichs durch Sequenzen der inaktiven LTR 21, die große Homologie zur LTR des Provirus HERV-K 18 aufweist. Dieser Bereich könnte Bindestellen für weitere regulatorische Transkriptionsfaktoren tragen. Auch Substitutionen des Bereichs nt 238 – nt 580 durch LTR 21-Sequenzen zeigen einen Aktivitätsverlust von 25 % (M. Kraft, 2004).

Die R-Region wird durch den Transkriptionsstart in der 5'LTR und die Poly-(A)-Stelle in der 3'LTR begrenzt. Da die 5' und 3' LTR identisch sind und gleichermaßen Motive zur Transkriptionsinitiation und Termination enthalten, muss sichergestellt werden, dass die Transkription in der 5' LTR initiiert und in der 3' LTR terminiert wird. Im Falle von HERV-K ist das Poly-A-Signal (AAUAAA) an der gleichen Stelle in der LTR lokalisiert wie der Transkriptionsstart. In vielen Retroviren wie beispielsweise MMTV, HTLV und ALV ist das Poly-A-Signal stromaufwärts vom Transkriptionsstart lokalisiert. Die Transkripte enthalten daher nur das Polyadenylierungssignal der 3' LTR. Liegt jedoch das Poly-A-Signal stromabwärts des Transkriptionsstarts, können die Transkripte auf Grund des 5' oder 3' Poly-A-Signals terminiert werden. Bei HIV, MMTV, MLV und HERV-K sind daher Regulationsmechanismen notwendig, die eine vorzeitige Polyadenylierung verhindern. Diese Regulation könnte über Konformationsänderungen des DNA-Strangs bei der Transkriptionsinitiation verursacht werden, die ein Überlesen des Poly-A-Signals verursachen oder das Entkoppeln der Polymerase vom entstehenden RNA-Strang unterbinden, damit die Transkription fortgesetzt wird.

Durch Promotorinterferenz agiert die 3' LTR meist nicht als funktionaler Promotor. Wird allerdings dieser Regulationsmechanismus außer Kraft gesetzt, so kann auch eine Transkriptionsinitiation in der 3'LTR erfolgen (Cullen *et al.*, 1984). Das in der Melanomlinie überwiegend aktive HERV-K Provirus 108 weist an der Stelle des Poly-A-Signals einen Basenunterschied zwischen 5' und 3' LTR auf.

5' LTR : AGTAAAT
3' LTR : AATAAAT

Abbildung 28: Vergleich der Poly-(A)-Signale in der 3' und 5' LTR von HERV-K 108

Durch das Guanin an zweiter Position des Motivs in der 5'LTR wird das Poly-A-Signal in seiner Effektivität drastisch reduziert. Es wird spekuliert, dass auch die flankierenden DNA-Bereiche Einfluss auf die spezifische Funktionalität der 5' und 3' LTR haben. Diese Auswirkungen könnten auf Grund von spezifischen Sequenzmotiven, Chromatinstrukturen oder durch die Bindung von Transkriptionsfaktoren vermittelt werden (Guntaka, 1993).

Der U5-Bereich könnte ebenso wie der U3-Bereich eine Vielzahl putativ wichtiger Transkriptionsfaktorbindestellen enthalten. Es ist bereits bekannt, dass der U5-Bereich der HERV-K LTR regulatorische Bereiche enthält. Deletiert man die letzten 60 Basenpaare, resultiert dies in einem deutlichen Anstieg der Promotoraktivität (Domansky et al., 2000; M. Kraft, 2004). Solche in U5 lokalisierten Repressoren sind auch für das *primaten foamy virus* (PFV) (Yang et al., 1997) und HTLV (Okumura et al., 1996) bekannt.

Die Transkriptionsinitiation und Termination (Abb.: 17) erfolgen in der untersuchten Melanomlinie undefinierter als in der Keimzelltumormlinie GH. Dies könnte auf Unterschiede in der Zusammensetzung des Transkriptionsinitiationskomplexes deuten. Die Existenz multipler Transkriptionsstartpunkte, wie für die HERV-K Transkripte in der Melanomlinie UKRV Mel 2-C9 festgestellt, ist meist ein Verweis darauf, dass die Initiation auf Grund des Fehlens eindeutiger Kernpromotorelemente unpräzise abläuft (Weis und Reinberg, 1992; Geng und Johnson, 1993).

Tabelle 4: Anwesenheit und Lokalisation typischer eukaryotischer Core-Promotor-Elemente in der HERV-K-LTR. Aufgeführt sind jeweils die Konsensusmotive der sieben für Eukaryoten bekannten Core-Promotor-Elemente **TATA-Box**, Initiatorelement (**Inr**), *downstream promotor element* (**DPE**), *motif ten element* (**MTE**), *downstream core element* (**DCE**), *upstream* und *downstream* TFIIB-recognition element (**BRE^u** **BRE^d**) und bei Anwesenheit die Lokalisation in der HERV-K LTR.

Core-Promotor-Element	Konsensusmotiv	Lokalisation [nt]
TATA-Box	TATA(A/T)A(A/T)(A/G)	531-537
Inr	Py, Py, A ⁺¹ , N, T, Py, Py	791-798
DPE	(A/G)G(A/T)CGTG	-
MTE	C(G/C)A(A/G)C(G/C)(G/C)AACG(G/C)	-
DCE	CTTC, CTGT, AGC	-
BRE ^u	(G/C)(G/C)(G/A)CGCC	-
BRE ^d	(G/A)T(T/G/A)(T/G)(G/T)(T/G)(T/G)	-

Die HERV-K LTR enthält verschiedene Sequenzmotive, die eine Transkriptionsinitiation vermitteln könnten. Welche dieser Core-Promotor-Elemente zum Tragen kommen, kann abhängig vom Typ und Differenzierungsgrad des Gewebes sein. Der U3-Bereich weist ein TATA-Box-Motiv auf (nt 531 – 537). Ein Inr (nt 791 - 798) ist in der R-Region lokalisiert

(Tabelle 4). Durch die Bestimmung des Transkriptionsstarts bei nt 792 und unter Berücksichtigung der typischen Abstände möglicher Promotorsequenzmotive zum Startpunkt der Transkription lassen sich erste Rückschlüsse auf die tatsächlich genutzten *Core*-Elemente ziehen. Die im U3-Bereich enthaltene TATA-Box liegt zu weit vom Transkriptionsstart entfernt, als dass sie zur Initiation der basalen HERV-K Transkription genutzt werden könnte. Mit Hilfe eines Promotorkonstrukts, das statt der TATA-Box eine Thymidinabfolge enthielt, konnte gezeigt werden, dass in der Melanomlinie UKRV Mel 2-C9 ebenso wie in der Keimzelltumormlinie GH bei Veränderung des Motivs kein Aktivitätsverlust des Promotors auftritt. Die Initiation der Transkription erfolgt über eine TATA-Box unabhängige Regulation. In anderen Geweben oder bestimmten Differenzierungszuständen könnte alternativ durchaus eine TATA vermittelte Transkriptionsinitiation stattfinden. Für MMTV konnte in der U3-Region ein alternativer Promotor identifiziert werden, der lediglich in B-Zellen zur Transkriptionsinitiation benutzt wird (Günzburg *et al.*, 1993).

Das Inr im R-Bereich der proviralen LTR zeigt eine typische Lokalisation direkt am Transkriptionsstart und ist damit ein vielversprechender Kandidat. Ein Inr-Element kann in Analogie zur TATA-Box eine exakte und effiziente Transkriptionsinitiation dirigieren. Durch den Befund alternativer Transkriptionsinitiationsstellen bleibt allerdings fraglich, ob lediglich das Inr für die Transkriptionsinitiation verantwortlich ist oder ob synergistisch mehrere Motive zur Rekrutierung des Präinitiationskomplexes genutzt werden. Hier könnten auch Bindestellen für gewebsspezifische Transkriptionsfaktoren wichtigen Einfluss nehmen. Um zu bestätigen, dass das Initiatorelement die Initiation der Transkription an der HERV-K LTR vermittelt, sind Luziferasestudien mit Mutationskonstrukten durch zu führen. Für die anderen in der Tabelle aufgeführten *Core*-Promotor-Elemente gibt es in der HERV-K LTR keine Sequenzentsprechung.

Es könnten zusätzliche regulatorische Motive eine Rolle bei der Transkriptionsinitiation spielen. Eine CAAT-Box ist ein häufiges Element in eukaryotischen Promotoren und in TATA-losen Promotoren meist in reverser Orientierung etwa 41 bis 80 Nukleotide stromaufwärts des Startpunkt der Transkription zu finden (Mantovani, 1998). Die HERV-K LTR weist eine solche CAAT-Box in passendem Abstand an der Positionen nt 739 bis nt 743 auf. Es gibt mehrere Proteine, die an die CAAT-Konsensussequenz binden können. Potenzielle Bindepartner sind *CCAAT/enhancer binding protein* (C/EBP), *CCAAT transcription factor/nuclear factor-I* (CTF/NF-I), *CCAAT displacement protein* (CDP) und NF-Y (alternative Namen: CBF oder CP1) (Alonso *et al.*, 1996; Mantovani, 1998).

NF-Y ist ein ubiquitär exprimiertes Proteinheteromer aus drei Untereinheiten (NF-YA, NF-YB, NF-YC). Alle Untereinheiten werden für die DNA-Bindung benötigt (Sinha *et al.*, 1995). Durch EMSA-Analysen und Chromatinimmunopräzipitation kann eine Interaktion einer dieser Faktoren mit dem putativen Bindemotiv der LTR untersucht werden. Durch Luziferaseassays mit Mutationen des Bindemotivs und durch siRNA-*Knockdown* einzelner Proteinkomponenten können Hinweise auf die Bedeutung dieser Transkriptionsfaktoren für die Steuerung der HERV-K Expression gewonnen werden.

Bei Genen, deren Promotor keine TATA-Box aufweist, kann eine Sp1, Sp3 Bindungsstelle die kritische Determinante zur transkriptionellen Aktivierung darstellen und durch Interaktion der SP-Proteine mit TFIID den Transkriptionsstartpunkt positionieren (Pugh und Tjian, 1990; Smale *et al.*, 1990; Pugh und Tjian, 1991; Lu *et al.*, 1994; Emami *et al.*, 1995; Körner *et al.*, 1997; Saluja *et al.*, 1998). Die Eigenschaften der Sp-Proteine werden durch die Anwesenheit weiterer Transkriptionsfaktoren zusätzlich beeinflusst.

In der HERV-K LTR wurden durch *in silico* Analysen (2.5.1.) zwei putative Sp 1, Sp 3 Bindestellen in unmittelbarer Nähe zum prominenten Transkriptionsstart identifiziert. GC-Box1 weist Unterschiede zwischen der aktiven und der inaktiven LTR auf, die in der inaktiven LTR zu einem Verlust des typischen Sp1, Sp3 Bindemotivs führen. Die zweite GC-Box schließt sich nach einem *Spacer* von vier Basenpaaren an. Hier gibt es keine Sequenzunterschiede zwischen aktiver und inaktiver LTR. GC-Box1 scheint zur erfolgreichen Transkriptionsinitiation unabdingbar zu sein.

Erste Hinweise lieferten *Knockdown* Experimente mit siRNA (2.5.1.4.). Diese haben gezeigt, dass Sp1 und Sp3 die HERV-K Transkription beeinflussen. Sp1 regulierte Promotoren weisen häufig multiple Bindungsstellen für den Transkriptionsfaktor auf, wobei die wichtigste Stelle diejenige ist, welche 40 bis 70 Nukleotide 5' vom Transkriptionsstartpunkt lokalisiert ist (Kadonaga *et al.*, 1986). In diesem Abstand zum Transkriptionsstart befindet sich die oben beschriebene GC-Box1. Das GC-Box1 ein zur Transkriptionsinitiation entscheidendes Bindemotiv ist, konnte durch die Substitution der GC-Box1-Sequenz in der pK30 LTR, durch die Sequenz der LTR 21 im Luziferaseassay (2.5.1.3.) gezeigt werden. Die veränderte LTR zeigte im betreffenden Assay einen deutlichen Aktivitätsabfall. Sp1 und Sp3 binden als Transkriptionsfaktoren an die HERV-K LTR in unmittelbarer Nähe zum Initiatorelement wie durch EMSA (2.5.1.2.) und Chromatinimmunopräzipitation (2.5.1.5.) nachgewiesen werden konnte. Die spezifischen Shift-Signale im EMSA konnten nur durch eine nicht radioaktiv markierte Sonde, die das

GC-Motiv der aktiven pK30 LTR enthielt, nicht aber durch eine Sonde, die das LTR 21 GC-Box1 Motiv besitzt, kompetitiert werden.

Offensichtlich kann die GC-Box1 zusammen mit dem Inr den basalen Transkriptionsinitiationskomplex rekrutieren und somit den Transkriptionsstart der HERV-K LTR festlegen. Wie bereits erwähnt soll die Rolle des Initiatorelementes in folgenden Experimenten durch den Einsatz von Mutationskonstrukten genauer untersucht werden. Die Mutation des GC-Box1 Motivs führte zu einer deutlichen Reduktion der Promotoraktivität jedoch nicht zu einem totalen Verlust der transkriptionellen Aktivität. Wenn Inr und GC-Box1 synergistisch wirken, sollte die Mutation beider Motive zu einem vollständigen Funktionsverlust des Promotors im Luziferaseassay führen.

Die Beteiligung der Sp-Proteine an der Initiation der Transkription könnte der Grund für die diffusen Transkriptionsstartpunkte in der Linie UKRV Mel 2-C9 sein. Da jedoch in GH Zellen bei Vorhandensein der gleichen Konsensusmotive eine genauere Initiation der Transkription ermöglicht wird, müssen zudem unterschiedliche Modifikationen der Sp-Proteine eine Rolle spielen, ebenso wie die Ausstattung der Zelle mit weiteren Transkriptionsfaktoren, die zur Interaktion mit den Sp-Proteinen zur Verfügung stehen. Die leicht unterschiedlichen Shift-Banden im EMSA könnten auf das Vorhandensein unterschiedlicher Proteine in den hochmolekularen Multiprotein-Shift-Komplex von GH und UKRV Mel 2-C9 deuten. Beispiele für eine Sp-Protein vermittelte Transkriptionsinitiation finden sich zum Beispiel im Gen der Pyruvatkinase der Retikulozyten. Hier reguliert Sp1 synergistisch mit Mitgliedern der GATA-Familie die Genexpression (Gregory *et al.*, 1996). Für GATA konnten auch im HERV-K Promotor Konsensusstellen nachgewiesen werden. Die Expression des Gens der humanen DNA Polymerase α (Lin *et al.*, 1996) wird durch Kooperation der Sp-Proteine mit E2F reguliert. An vielen Promotoren wirkt Sp1 als leistungsfähiger Aktivator, während Sp3 inaktivierend oder nur leicht stimulierend wirkt. Den beiden kurzen Isoformen von Sp3 fehlt die Transaktivierungsdomäne A (Abb.: 6), diese Isoformen könnten für eine reprimierende Wirkung von Sp3 verantwortlich sein. In den Immunoblotanalysen konnten die drei Isoformen von Sp3 in allen untersuchten Zelllinien detektiert werden. Das Verhältnis von Sp1 zu Sp3 und seinen Isoformen könnte sich ebenfalls auf die Regulation des HERV-K Promotors auswirken. Eine quantitative Analyse der jeweilig exprimierten Proteine ist daher ratsam.

Zudem konnte gezeigt werden, dass Sp3 post-translational am Lysin551 in der inhibitorischen Domäne sumoyliert werden kann und nach der Modifikation als

transkriptioneller Repressor fungiert (Sapetschnig *et al.*, 2002, Ross *et al.*, 2002; Sapetschnig *et al.*, 2004). Viele durch post-translationale Modifikationen ausgelöste Auswirkungen auf die Funktionsweise der Sp-Proteine sind noch unaufgeklärt. Modifizierte Sp-Proteine können mit jeweils anderen Transkriptionsfaktoren interagieren und auf diese Weise zusätzliche regulatorische Eigenschaften ausüben. Um tiefere Einblicke in das komplexe regulative Netz der Sp1 und Sp3-Proteine an der HERV-K LTR zu erhalten, ist der Einsatz modifikationsspezifischer Antikörper und gezielter Mutationskonstrukte notwendig. Sp-Proteine sind zudem in der Lage die Transkription aus großer Distanz zum Startpunkt mittels Protein/Protein-Interaktionen und DNA-Faltung zu verstärken. Durch diese Veränderung in der DNA-Struktur können distal gelegene Aktivatoren in die Nähe des basalen Transkriptionskomplexes gelangen (Fry und Farnham, 1999). In zukünftigen Untersuchungen sollen daher weitere putative Sp1 und Sp3 Bindestellen in anderen Bereichen der HERV-K LTR berücksichtigt werden.

Neben der Ausstattung der Zelle mit allen erforderlichen Transkriptionsfaktoren muß zur erfolgreichen Transkriptionsinitiation auch das Chromatin am proviralen Locus die Transkription erlauben. Daher ist die Frage zu klären ob in der parentalen Melanomzelllinie UKRV Mel 2 der Locus 108 durch epigenetische Mechanismen inaktiviert vorliegt. Sollte der Locus epigenetisch offen sein, ist die transkriptionelle Aktivität des Subklons UKRV Mel 2-C9 eindeutig auf Unterschiede in der Ausstattung der Zelle mit Transkriptionsfaktoren zurück zu führen. Für HERV-K konnte gezeigt werden, dass auch epigenetisch offene Proviren nicht zwangsläufig transkriptionel aktiv sind (Lavie *et al.*, 2005).

Bei sogenannten *housekeeping*-Genen, die ständig transkribiert werden, findet häufig keine Methylierung der CpG-Dinukleotide in der Promotorsequenz statt (Bird, 1986; Gardiner-Garden und Frommer, 1987; Larsen *et al.*, 1992). Der Einfluss der Promotor-Methylierung auf die Transkription konnte beispielsweise in einem episomalen System, basierend auf dem Epstein Barr Virus (EBV), gezeigt werden. Hier ist die Methylierung eines einzigen bestimmten CpG-Dinukleotids für die Inaktivierung des Promotors verantwortlich (Robertson *et al.*, 1995). Wenn 7% der CpG-Dinukleotide des Promotors methyliert sind, kommt es im EBV ebenfalls zu einer transkriptionellen Repression (Hsieh, 1994). Auch die Methylierung einzelner CpG-Dinukleotide der Thymidin-Kinase beim Herpes Simplex Virus-1 (HSV) führt zur Herunterregulation der Promotoraktivität *in vivo* (Ben-Hattar und Jiricny, 1988). Die Methylierung von CpG-Dinukleotiden kann ebenfalls Einfluss auf die Bindungseigenschaften des Transkriptionsapparates an der Doppelhelix und die

Bindungseigenschaften der Proteine, welche die Chromatinstruktur beeinflussen, nehmen (Nan *et al.*, 1997; Kass *et al.*, 1997). 5-Methylcytosine haben zwar keinen Einfluss auf die Basenpaarung, beeinträchtigen aber die Protein/DNA Interaktion, da sie in die große Furche der Doppelhelix reichen. Auf diese Weise könnten sie die Bindung eines die Transkription stimulierenden Faktors beeinträchtigen (Razin und Riggs, 1980).

Einige Promotoren benötigen beispielsweise für die Transkription ganz bestimmte Faktoren, wie E2F, CREB und USF, die an methylierte CpG-Dinukleotide nicht gut binden können, und dadurch wird die Transkription inhibiert (Kovesdi *et al.*, 1987; Watt und Molloy, 1988; Faisst und Meyer, 1992; Nevins, 1992; Campanero *et al.*, 2000). Andererseits gibt es Faktoren, wie Sp1, CTF und TCR-ATF, bei denen die Bindung durch Methylierung nicht beeinflusst wird (Harrington *et al.*, 1988; Holler *et al.*, 1988; Ben-Hattar *et al.*, 1989). Die sonst im Luziferaseassay aktive HERV-K LTR pcK30 zeigt nach Methylierung aller CpG-Dinukleotide nahezu keine Promotoraktivität (M. Kraft, 2004). Die bisher als Bindungspartner der HERV-K LTR identifizierten Proteine Sp1 und YY1 sind in ihren Bindungseigenschaften methylierungsinsensitiv. Die Reduktion der Aktivität bei Verwendung des methylierten Konstruktes im Luziferaseassay muss daher auf andere Effekte zurück zu führen sein. Bei der transkriptionellen Repression kommt es durch einen sequenzunabhängigen Prozess zu einer Änderung der Chromatinstruktur und des Acetylungsgrades der Histone (Kass *et al.*, 1997; Bestor, 1998). Der Promotor rekrutiert nicht ausreichend Transkriptionsfaktoren, und in Folge dessen können an das aufgelockerte Heterochromatin Methyl-Cytosin-bindende Proteine unter Rekrutierung eines Ko-Repressor-Komplexes, der auch Histonacetylasen enthält, gebunden werden (Wade *et al.*, 1999; Ng *et al.*, 1999; Nan *et al.*, 1998; Jones *et al.*, 1998). Dies könnte die Bindung von Sp1 und Sp3 beeinflussen. Nach Deacetylierung der Histone kann es zu einer Verdichtung des Heterochromatins und zum fast vollständigen Erliegen der Transkriptionsaktivität kommen (Jones und Laird, 1999).

Die Methyl-Cytosin-bindenden Proteine spielen eine entscheidende Rolle zwischen Methylierung und dem sogenannten transkriptionellen *Silencing*. MeCP2 wurde als erstes Mitglied dieser Genfamilie kloniert und charakterisiert (Henderich und Bird, 1998). Durch Trichostatin A, einen Histondeacetylase-Hemmer kann die stabile Repression der Chromatinstruktur aufgehoben werden. In der HERV-K negativen Linien HCl Mel 19 und HeLa wird allerdings durch die Trichostatin A Behandlung keine HERV-K Transkription initiiert. Zur Initiation fehlen wahrscheinlich die erforderlichen Faktoren. Die Beteiligung von Histonmodifikationen bei der Transkriptionsregulation von HERV-K Promotoren kann

durch Chromatinimmunopräzipitationen mit modifikationsspezifischen Histon-Antikörpern untersucht werden. Zu Veränderungen der epigenetischen Signaturen kommt es vor allem während der Embryonalentwicklung. Erste Studien an humanen embryonalen Stammzellen, die unsere Arbeitsgruppe in Zusammenarbeit mit einem Kooperationspartner in den USA unternommen hat, sowie Ergebnisse aus der Arbeitsgruppe von J. Denner (C.Büscher, 2006), deuten auf eine vermehrte HERV-K Aktivität nur in den undifferenzierten Zellen. Die HERV-K Expression geht mit zunehmender Differenzierung zurück. Zu klären bleibt, ob die Expression von HERV-K in diesen Zellen einen funktionalen Zweck erfüllt oder lediglich ein Nebeneffekt ist.

Zusammenfassend kann ein Modell zur Initiation der basalen HERV-K Transkription in Melanom- und Keimzelltumorlinien aufgestellt werden (Abb.: 29). Die Transkription erfolgt unabhängig von der TATA Box. Die Rekrutierung des transkriptionellen Präinitiationskomplexes erfolgt an einem Initiatorelement (Inr) in Zusammenspiel mit den Transkriptionsfaktoren SP1 und SP3. Hierbei spielen posttranslationale Modifikationen der Sp-Proteine (*) sowie das quantitative Verhältnis von Sp1 und Sp3 Isoformen und die Interaktion mit Koaktivatoren oder Korepressoren eine wichtige Rolle. Im 5'-Bereich der LTR wirkt modellierend der Transkriptionsfaktor YY1 als Enhancer (Knössel et al., 1999). Im 3'-Bereich der LTR wirkt ein unbekannter Faktor als Repressor. Ob an der Promotorsteuerung epigenetische Signaturen sowie weitere, evtl. zelltypspezifische Transkriptionsfaktoren beteiligt sind, die in Wechselwirkung mit den hier genannten Faktoren stehen oder mit diesen synergistisch wirken, ist zu vermuten und soll in weiteren Experimenten aufgeklärt werden.

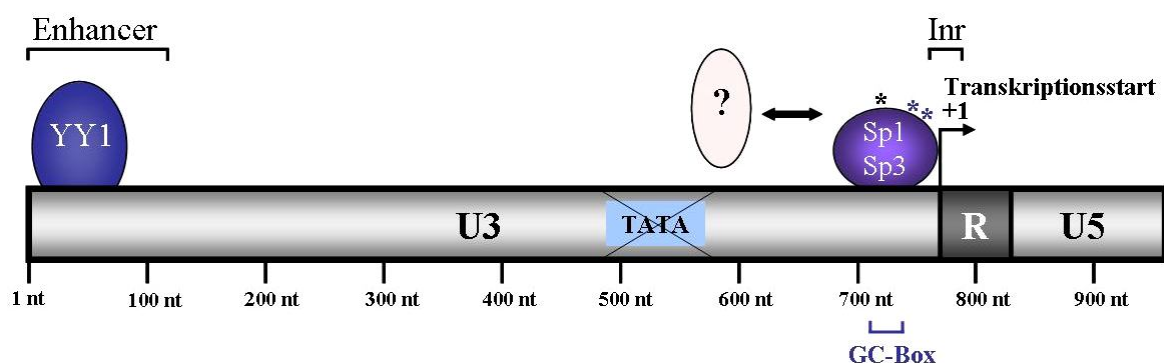


Abbildung 29: Modell zur Initiation der basalen Transkription an der 5' LTR von HERV-K. Die Initiation der Transkription erfolgt durch das Initiatorelement (Inr) und an die GC-Box1 gebundene Sp-Proteine. Dargestellt sind der U3, R und U5-Bereich der HERV-K LTR. Der Transkriptionsstart ist durch +1 gekennzeichnet. Durch * sind mögliche Proteinmodifikationen dargestellt. Der Pfeil deutet die Möglichkeiten zur Interaktion mit weiteren, noch unbekanntem Proteinen an. An den Enhancer bindet der Transkriptionsfaktor YY1.

4. MATERIAL UND METHODEN

4.1. Allgemeine Materialien und Geräte

Alle molekularbiologischen Arbeiten wurden, soweit nicht anders angegeben nach Standard-Protokollen durchgeführt. Diese Protokolle wurden den Standardwerken „Molecular Cloning“ (Sambrook *et al.*, 1998) und „Current Protocols in Molecular Biology“ (Ausubel *et al.*, 1992) entnommen. Umdrehungszahlen der Zentrifugationsschritte sind in rpm (rounds per minute) angegeben. Die Zentrifugation größerer Volumina (>50 ml) wurde in der Sorvall[®] RC26 Plus mit dem SLA-1500-Rotor durchgeführt. Volumina zwischen 2 und 50 ml (Greinergefäße) wurden in der Multifuge 4KR (Heraeus) zentrifugiert. Standard Plastikreaktionsgefäße (Eppendorf) mit Volumina von 1,5 und 2 ml wurden in der Fresco 17 (Heraeus) zentrifugiert. Die Ultrazentrifugationsschritte wurden in einer Optima[™] L-80 XP Ultrazentrifuge in einem Ti45, SW32 Ti oder SW60 Rotor durchgeführt.

4.2. Zellbiologische Methoden

4.2.1. Geräte, Medien und Chemikalien

Brutschrank:	Heraeus Typ BBD6220
Mikroskop:	DM IL (Leica)
Mundschutz:	Coldex
Sterilbank:	Steril Gard Hood (The Baker Company, inc.; Sanford Maine)
Einfrierampullen:	Greiner
Stickstofftonne:	BT55/BT40 (L'air liquide)
Zellkulturflaschen:	80 cm ² , straight neck, Nunclon Flasche (Nalge Nunc International)
DMEM:	Dulbecco's Minimal Essential Medium; mit 4,5 g Glukose pro Liter (Biochrom KG, Berlin)
RPMI 1640:	Biochrom KG, Berlin

CMRL 1969:	1 Liter: 1 g Glukose; 7,8 g NaCl; 0,45 KCl, 0,22 g MgSO ₄ *7H ₂ O; 0,08 g NaH ₂ PO ₄ *H ₂ O; 0,2 g NaHPO ₄ ; 0,21 g CaCl ₂ *2H ₂ O; 0,022 g Phenolrot, 0,001 g Folsäure, Amminosäuren in geeigneten Konzentrationen
DMSO:	Dimethylsulfoxid (Sigma)
FKS:	Fötale Kälberserum (Biowest; South America Origin)
L-Glutamin:	200 mM Biochrom KG, Berlin
Pen/Strep:	Streptomycin 10 mg/ml, Penicillin 10.000 U/ml (Gibco/Invitrogen)
PBS ⁻ :	Phosphat-gepufferte Salzlösung ohne Ca ²⁺ , Mg ²⁺ (Biochrom KG, Berlin)
Trypsin/Versen:	10 ml 2,5 % Trypsin, 40 ml Versen, 50 ml PBS ⁻
Versen:	159 mM NaCl, 2,7 mM KCl, 6,5 mM Na ₂ HPO ₄ , 1,5 mM KH ₂ PO ₄ , 2,7 mM Versen

4.2.2. Zelllinien und Kulturbedingungen

Tabelle 5: Übersicht über die verwendeten humanen Zelllinien und die jeweiligen Kulturbedingungen.

Zellen	Zellart/Herkunft	Medium	wöchentliche Teilungsrate
GH	Teratokarzinom (Löwer <i>et al.</i> , 1984)	CMRL	1:2
2102 EP	Teratokarzinom (Andrews <i>et al.</i> , 1980)	DMEM	1:4
NCCIT	Teratokarzinom	DMEM	1:4
HeLa	Zervixkarzinom (ATCC)	DMEM	1:10
UKRV Mel 2	malignes Melanom	RPMI	1:5
HCl Mel 19	malignes Melanom	RPMI	1:5
Ma Mel 11	malignes Melanom	RPMI	1:5
WM 98.2	malignes Melanom	RPMI	1:5
UKRV Mel 14a	malignes Melanom	RPMI	1:5
Ma Mel 21	malignes Melanom	RPMI	1:5
UKRV Mel 2-C9	malignes Melanom, Subklon	RPMI	1:5
HeLa HK _{ges}	HeLa stabil HERV-K transfiziert, Zusatz von Neomycin zur Selektion	DMEM	1:8

Die Medien DMEM 4,5 g/l Glucose, RPMI und CMRL wurden jeweils mit 10 % (v/v) FKS, 2 mM L-Glutamin, 40 U/ml Penicillin und 40 µg/ml Streptomycin versetzt. Alle verwendeten Linien wurden in der Grundhaltung in 80 cm² Kulturflaschen (Nunc) mit

~20 ml Medium bei 37 °C, 5 % CO₂ und 80 % Luftfeuchtigkeit kultiviert. Alternierend wurde pro Woche je einmal das Medium gewechselt und die Zellen im Verhältnis 1:2 - 1:10 passagiert (Tabelle 5). Zum Passagieren der Zellen wurde das Medium aus den Kulturflaschen komplett entfernt und die Zellen mit PBS⁻ gewaschen. Zum Ablösen der Zellen wurden diese mit 2 ml Trypsin/Versen für circa fünf Minuten bei 37 °C inkubiert und dann in ca. 20 ml frischem Medium resuspendiert und auf neue Kulturflaschen aufgeteilt. Zur fraktionierten Proteinaufreinigung (4.7.2.) und Isolation viraler Partikel (4.9.) aus dem Kulturüberstand wurden die Zellen in T175-Flaschen (Nunc) ausgebracht.

4.2.3. Einfrieren und Auftauen von Zellen

Die Lagerung von Zellen über einen langfristigen Zeitraum erfolgte bei -196 °C in flüssigem Stickstoff. Um Schäden auf Grund von Eiskristallbildung während des Einfrierprozesses zu minimieren wurden die Zellen mit Dimethylsulfoxid (DMSO) versetzt. Die Zellen wurden mechanisch mit einem Zellschaber vom Boden der 80 cm² Flasche gelöst, in 10 ml PBS⁻ resuspendiert, anschließend in 10 ml Falcontubes überführt und bei ~800 rpm für fünf Minuten pelletiert. Der Überstand wurde verworfen und das Pellet in 4 ml Kulturmedium mit 20 % FKS und 10 % DMSO aufgenommen und in 2 ml Kryogefäße gegeben. Die Kryoröhrchen wurden über Nacht in einer Styroporbox bei minus 80 °C gelagert und am nächsten Tag in einen Kryotank zur Lagerung überführt. Zum Auftauen der Zellen wurden die Kryoröhrchen bei 37 °C im Wasserbad aufgetaut und die Zellsuspension anschließend in vorgewärmtes Vollmedium gegeben.

4.2.4. Bestimmung der Zellvitalität und Zellzahl mit Trypanblau

Trypanblau: 0,37 % (Merck) in PBS⁻

Zählkammer: Fuchs Rosenthal

Mit Hilfe von Trypanblau, einem Derivat des Azofarbstoffes Benzopurpurin, wurden lebende und tote Zellen unterschieden. Das Anion des sauren Farbstoffs bindet Zellproteine. Es kann durch defekte Zellmembranen toter Zellen in deren Zytoplasma eindringen und diese Zellen tiefblau färben. Mit Hilfe der Fuchs Rosenthal Zählkammer wurde auf diese Weise die Gesamtzellzahl ermittelt und gleichzeitig die Vitalität überprüft bevor die Zellen in den Dualen Luziferaseassay eingesetzt wurden (4.7.5.).

4.2.5. Mycoplasmentest

Kit: „PCR Mycoplasma Test Kit“ (AppliChem)

Um eine Kontamination der eukaryotischen Zellkultur mit Mycoplasmen auszuschließen wurden in regelmäßigen Abständen Mycoplasmentestungen durchgeführt. Neben dem mikroskopischen Testverfahren durch Anfärben möglicher Mycoplasmen mit Bisbenzimid (Hoechst), wurde ein „PCR Mycoplasma Test Kit“ (AppliChem) nach Angaben der Gebrauchsanweisung verwendet.

4.2.6. Subklonierung eukaryotischer Zellen

Klonierungsplatte: 384 Loch-Platte (Greiner)

Zellkulturplatten: 24 und 6 Loch –Platte (Nunc)

Die Zellen wurden mit Trypsin vereinzelt und in Kulturmedium verdünnt, die Verdünnung wurde so gewählt, dass statistisch in jedes Loch der Klonierungsplatte eine Zelle gesät wurde. Nach zwei Tagen Inkubation unter Standardbedingungen wurde das Medium gewechselt. Die Klone wurden auf 24 Loch- und anschließend auf 6 Loch-Platten expandiert.

4.3. Bakterienkultur

4.3.1. Medien und Chemikalien

LB-Medium: Luria-Bertani-Medium

1 % Bacto-Trypton, 0,5 % Bacto Hefe-Extrakt, 1 % NaCl, pH [7,0]; autoklaviert

LB-Amp-Medium: LB-Medium mit 100 µg/ml Ampicillin

LB-Agar-Platten: 1,5 % Bacto-Agar in LB-Medium, pH [7,0]; autoklaviert

LB-Amp-Platten: LB-Agar-Platten mit 100 µg/ml Ampicillin

LB-Amp-X-Gal-Platten: LB-Agar-Platten mit 100 µg/ml Ampicillin und 80 µg/ml X-Gal

LB-Kan-Platten: LB-Agar-Platten mit 30 µg/ml Kanamycin

TSB:	5 % [v/v] Dimethylsulfoxid, 10 mM MgCl ₂ , 10 mM MgSO ₄ , 10 % [w/v] PEG 6000 in LB-Medium
5-fach KCM:	500 mM KCl, 150 mM CaCl ₂ , 250 mM MgCl ₂
Bakterienschüttler:	Orbital-Shaker (Thermo)
Photometer:	Spectronic 1201 (Milton Roy)

4.3.2. Bakterienstämme

<i>Escherichia coli</i> DH5 α	Invitrogen F ⁻ ϕ 80dlacZ Δ M15 Δ (lacZYA-argF) U169 recA1 endA1 hsdR17(r _k ⁻ , m _k ⁺) phoA supE44 λ thi-1 gyrA96 relA1
--------------------------------------	---

4.3.3. Anlegen einer Übernachtskultur

Eingefrorene Bakterien oder Kolonien von einer Klonierungsplatte wurden in 5 ml LB-Medium mit geeignetem Antibiotikum über Nacht im Schüttelinkubator bei 37 °C und 225 rpm herangezogen. Aus den Übernachtskulturen wurden die gewünschten Plasmide aufgereinigt (4.3.7.).

4.3.4. Erzeugung kompetenter Bakterien

100 ml LB-Medium wurden mit 1 ml einer frischen Übernachtskultur beimpft. Die Bakterien wurden bis zu einer optischen Dichte von 0,5 - 0,7 bei 37 °C im Schüttelinkubator herangezogen. Die Bestimmung der optischen Dichte erfolgte bei einer Wellenlänge von 600 nm im Spektralphotometer. Anschließend wurden die Zellen fünfzehn Minuten bei 3500 rpm und 4 °C abzentrifugiert. Alle folgenden Schritte wurden im Kühlraum bei 4 °C mit vorgekühlten Geräten durchgeführt. Das Bakterienpellet wurde in 7 ml TSB resuspendiert. Aliquots a 100 μ l wurden in flüssigem Stickstoff schockgefroren. Die Aliquots wurden bis zur Verwendung bei -80 °C gelagert.

4.3.5. Transformation von Bakterien mit rekombinanter DNA

100 μ l kompetenter Bakterienzellen wurden auf Eis aufgetaut und mit 80 μ l DNA (2-10 ng Plasmid-DNA oder 10 μ l Ligationsansatz) (4.4.3.5.) versetzt. Nach der Zugabe von 20 μ l

5-fach KCM und dem Durchmischen des Ansatzes erfolgte eine zwanzigminütige Inkubation auf Eis. Anschließend wurden die Bakterien für zehn Minuten bei Raumtemperatur (RT) inkubiert. Es bilden sich Poren in der Bakterienwand, durch welche das Plasmid in das Innere des Bakteriums gelangen kann. Nach Zugabe von 1 ml LB-Medium ohne Antibiotikum wuchsen die Bakterien eine Stunde bei 37 °C im Schüttelinkubator. Nach Ablauf der Inkubationszeit wurden 50-100 µl der Bakteriensuspension auf Agarplatten mit geeignetem Selektionszusatz ausgestrichen und bei 37 °C über Nacht inkubiert. Anschließend wurden einzelne Kolonien mit einer Impföse in Medium mit entsprechendem Antibiotikum überführt und zur Präparation vermehrt.

4.3.6. Selektionsprinzip

LB-Amp-X-Gal-Platten: LB-Agar-Platten mit 100 µg/ml Ampicillin, 80 µg/ml X-Gal

Die Selektion transgener Bakterien erfolgte durch den jeweiligen Selektionsmarker, der sich auf dem verwendeten Plasmid befindet. In den meisten Fällen handelte es sich um eine spezifische Antibiotikaresistenz. Eine weitere Selektion kann mittels blau-weiß Selektion erreicht werden, die auf der sogenannten α -Komplementation beruht. Bakterienstämme, die eine Mutation im 5'-Bereich des lacZ-Gens aufweisen, können keine aktiven Tetramere der β -Galaktosidase bilden. Komplementation mit dem plasmidkodierten lacZ α -Peptid (26 AS N-terminales lacZ-Fragment) führt zur Wiederherstellung der enzymatischen Aktivität. Das chromogene Substrat X-Gal wird umgesetzt, diese Kolonien zeigen eine blaue Färbung. Wird der Leserahmen des α -Peptids im Plasmid-Polylinker durch eine transgene Insertion unterbrochen, können die rekombinanten Bakterien anhand mangelnder Enzymaktivitäten als weiße Klone identifiziert werden.

4.3.7. Plasmidpräparation (Mini -, Midi- und Maxipräparation)

Die in Bakterien eingeschleuste rekombinante DNA wurde mit Hilfe der Plasmidpräparationstechnik gewonnen. Für eine Plasmidisolierung im kleinen Maßstab ,z.B. für Sequenzierreaktionen, wurden 2 ml einer Übernachtskultur verwendet. Mit Hilfe des „Qiaprep Spin Miniprep Kit“ der Firma Qiagen wurde nach den Herstellerangaben die Plasmid-DNA isoliert. Die Qiagen-Säulen enthalten ein Anionen-Austauscher-Material auf

Silicabasis. Die in dem Zellysate enthaltene DNA bindet auf Grund ihrer negativen Ladung an das Anionen-Austauscher-Material in der Säule. Hierbei kommt es zur Interaktion zwischen dem negativ geladenen Sauerstoffatom des Phosphates der DNA und dem positiv geladenen Stickstoffatom des Säulenmaterials. Nach der Elution der DNA wurde die Konzentration photometrisch bestimmt (4.4.1.1.). Durch eine Midi- bzw. Maxipräparation wurden Plasmid-DNA-Präparationen in einem größeren Maßstab durchgeführt.

4.3.8. Einfrieren und Lagern von Bakterien

Eine frische Übernachtskultur wurde mit ca. 20-30 % sterilem Glycerol in Wasser versetzt. Die Bakteriensuspension wurde anschließend bei -80 °C gelagert.

4.4. Molekularbiologische Methoden

4.4.1. Nukleinsäure-Techniken

4.4.1.1. Photometrische Bestimmung von Nukleinsäurekonzentrationen

Bei der photometrischen Bestimmung der Nukleinsäurekonzentrationen nutzt man das charakteristische Absorptionsspektrum der Moleküle. Die Stärke der Absorption ist von der Konzentration der Moleküle abhängig.

Für die Absorption gilt das Lambert-Beersche Gesetz:

$$I = I_0 e^{-kd}, \text{ wobei gilt}$$

I= Intensität der Strahlung nach durchlaufen der Schichtdicke d

I₀= Intensität des eingestrahnten Lichts

d= Schichtdicke

k= Absorptionskoeffizient des Moleküls

Das Absorptionsspektrum von Nukleinsäuren besitzt ein Maximum bei 260 nm. Es unterscheidet sich damit deutlich von dem der Proteine bei 280 nm, so dass auf diese Weise Verunreinigungen durch Überlagerungsspektren sichtbar werden. Nach der Messung der Extinktion wurde die Konzentration der Nukleinsäure-Lösung mit folgender Formel bestimmt:

$\text{Nukleinsäurekonzentration} = [\text{Extinktion}_{260 \text{ nm}} \times \text{Verdünnungsfaktor} \times \text{Faktor}] / 1.000$
--

Zur Berechnung der DNA-Konzentration wurde der Faktor 50 verwendet, für Konzentrationsberechnungen von RNA wurde der Faktor 40 eingesetzt.

4.4.1.2. Phenol-Chloroform-Extraktion von Nukleinsäuren

Phenol:	Roti-Phenol (Roth)
Phenol-Chloroform:	Phenol, Chloroform, Isoamylalkohol Verhältnis 25:24:1
Chloroform-Isoamylalkohol:	Chloroform, Isoamylalkohol Verhältnis 24:1

Die wässrige DNA oder RNA-Lösung wurde mit dem gleichen Volumen einer Phenol-Chloroform-Lösung versetzt und ungefähr eine Minute in einem Mischer (Vortex) geschüttelt. Anschließend wurden die Phasen durch Zentrifugation bei 13000 rpm für fünf Minuten separiert. Die obere, wässrige Phase, wurde in ein neues Eppendorf-Gefäß überführt. Die untere organische Phase wurde verworfen. Verbliebene Phenolreste wurden durch Zugabe gleichen Volumens einer Chloroform-Isoamylalkohol-Lösung entfernt. Dazu wurden die Chloroform-Isoamylalkohol-Lösung und die wässrige Phase eine Minute im Vortexer gemischt und für fünf Minuten bei 13000 rpm zur Phasentrennung zentrifugiert. Wieder wurde die wässrige Phase abgenommen und der organische Rest verworfen. Die Nukleinsäure wurde anschließend gefällt (4.4.1.3.).

4.4.1.3. Fällung von Nukleinsäuren

TE-Puffer:	10 mM Tris HCl pH [7,4], 1 mM Ethylendiamintetraessigsäure
Natriumacetat:	3 M Natriumacetat pH [6]

Die gelösten Nukleinsäuren wurden mit 0,1 Volumenteil einer 3 M Natriumacetatlösung versetzt und kurz durchmischt. Es folgte die Zugabe von 2 - 2,5 Volumenteilen Ethanol 96 % und erneutes Durchmischen der Lösung. Die Nukleinsäure wurde dreißig Minuten bei $-20\text{ }^{\circ}\text{C}$ präzipitiert und anschließend bei 14000 rpm für dreißig Minuten bei $4\text{ }^{\circ}\text{C}$ abzentrifugiert. Der Überstand wurde verworfen, während das Pellet mit ca. 1 ml 70 % Ethanol gewaschen wurde um verbliebene Salze zu entfernen. Erneut wurde die DNA bei 14000 rpm und $4\text{ }^{\circ}\text{C}$ für fünfzehn Minuten pelletiert. Der Überstand wurde verworfen und das Pellet bei RT für zehn bis fünfzehn Minuten getrocknet. Die Nukleinsäure wurde in

einem der gewünschten Konzentration entsprechenden Volumen Aqua bidest oder TE-Puffer über Nacht gelöst.

4.4.2. Ribonukleinsäure-Techniken

4.4.2.1. Isolierung zytoplasmatischer RNA

Kit: RNeasy Midi Kit (Qiagen)

Die Isolierung zytoplasmatischer RNA aus humanen Zelllinien wurde mit dem „RNAeasy Midi-Kit“ der Firma Qiagen entsprechend der Herstellerangaben durchgeführt. Um die Kapazität der verwendeten Säulen optimal zu nutzen, wurden die Zellen zuvor ausgezählt. Die Quantität und Qualität der aufgereinigten RNA wurde vor der Lagerung bei -80 °C photometrisch bestimmt (4.4.1.1.).

4.4.2.2. Poly-A-Trakt-*m*RNA Aufreinigung

Kit: PolyATtract® (Roche)

Die Gewinnung von *m*RNA erfolgte mit Hilfe des „PolyATtract® *m*RNA Isolation Systems“ (Promega). Aus zuvor gewonnener zytoplasmatischer RNA (4.4.2.1.) wurde *m*RNA nach den Angaben des Herstellers aufgereinigt.

4.4.2.3. Elektrophorese von RNA in denaturierenden Agarosegelen

Ladepuffer: 5 Ficoll 400, 0,25 % Bromphenolblau, 0,25 % Xylenxyanol; 50 mM Ethylendiamintetraessigsäure

FF-Lösung: Zusammengesetzt aus 40 µl 5-fach (Morpholino)Propansulfonsäure in 0,1 % Diethylen Pyrocarbonat Wasser, 70 µl Formaldehyd und 220 µl Formamid

Laufpuffer: 1 fach TAE in 0,1 % Diethylen Pyrocarbonat Wasser

Gel: 2 % SeaKem-Agarose (FMC, BioProducts) in 1-fach TAE in 0,1 % Diethylen Pyrocarbonat Wasser

Um die Qualität der RNA-Präparationen zu testen, wurden Proben isolierter RNA über ein RNA-Gel entsprechend ihrer Größe aufgetrennt. Um die Bildung von Sekundärstrukturen in der RNA durch intramolekulare Wechselwirkungen zu vermeiden, wurde die Auftrennung unter denaturierenden Bedingungen durch Zugabe von Formaldehyd und Formamid durchgeführt. Dazu wurden die Proben im Verhältnis 1:2 mit FF-Lösung versetzt und fünfzehn Minuten bei 55 °C inkubiert, dann mit 5-fach Laufpuffer versetzt und auf das Gel aufgetragen.

4.4.2.4. Verdau von DNA (DNaseI-Behandlung)

Kit: DNaseI (Invitrogen)

Um verbliebene DNA Kontaminationen in zytoplasmatischen RNA-Präparationen zu beseitigen wurden ~ 5 µg RNA mit 5 µl 10-fach DNaseI-Puffer (Invitrogen) und 5 µl DNaseI (1U/µl) (Invitrogen) in einem Volumen von 50 µl fünfzehn Minuten bei RT inkubiert. Anschließend wurde der Ansatz mit 5 µl 25 mM Ethylendiamintetraessigsäure (Invitrogen) versetzt und zehn Minuten bei 65 °C inkubiert. Die RNA wurde gefällt (4.4.1.3.), in Diethylen Pyrocarbonat Wasser gelöst und bei -80 °C bis zur weiteren Verwendung gelagert.

4.4.2.5. cDNA-Synthese (Reverse Transkription)

Kit: SuperScript™ II Reverse Transcriptase (Invitrogen)
Oligo-dT₁₂₋₁₈: Oligo-dT-Primer (Fermentas), (2 pmol/µl)
dNTP-Mix: 2,5 mM Desoxyadenosin-, 2,5 mM Desoxycytidin, 2,5 mM Desoxyguanosin, 2,5 mM Desoxythymidin (Applied-Biosystems)

Das Enzym Reverse Transkriptase (Superscript II; Invitrogen) synthetisiert cDNA ausgehend von einer einzelsträngigen RNA Matrize. 10 pg – 500 ng polyA-Trakt gereinigte mRNA (4.4.2.2.) wurden in einem Gesamtvolumen von 39 µl mit 3 µl Oligo-dT₁₂₋₁₈-Primer (2 pmol/µl) und 3 µl dNTPs (je 2,5 mmol) versetzt. Anschließend fünf Minuten bei 65 °C im Heizblock inkubiert und dann abgekühlt. Sekundärstrukturen lösen sich durch die Behandlung auf und die Bindung des Primers wird ermöglicht. Nach Zugabe von 12 µl First Strand Buffer, 3 µl Dithiothreitol (0,1 M), 3 µl RNaseOut sowie 3 µl

Superscript II folgte eine Inkubation bei 42 °C für 1 Stunde. Das Enzym wurde durch eine 15 minütige Inkubation bei 70 °C inaktiviert. Die gewonnene cDNA wurde aliquotiert und bei -20 °C gelagert.

4.4.3. Desoxyribonukleinsäure-Technik

4.4.3.1. Isolierung von genomischer DNA aus eukaryotischen Zellen

Lysepuffer: 10 mM Tris HCl pH [8], 10 mM Ethylendiamintetraessigsäure,
 200 mM NaCl, 0,5 % Sarkosyl

Als Ausgangsmaterial zur Isolierung genomischer DNA wurde eine ~ 80 % konfluente T75 Kultur-Flasche benutzt. Die Zellen wurden mit PBS⁻ gewaschen und mit einem Zellschaber in 10 ml PBS⁻ vom Boden der Flasche mechanisch gelöst. Die resuspendierten Zellen wurden in 15 ml Falcontubes überführt und für zehn Minuten bei 600 rpm pelletiert. Der Überstand wurde verworfen, das Pellet in ein 2 ml Plastikreaktionsgefäß überführt und mit 750 µl Lysepuffer und 7,5 µl Proteinase-K 20 mg/ml (Invitrogen) versetzt. Nach der Inkubation über Nacht bei 55 °C wurde die DNA Phenol/Chloroform (4.4.1.2.) extrahiert und präzipitiert (4.4.1.3.).

4.4.3.2. Restriktionsenzymatische Spaltung von DNA

Restriktionsenzyme: NEB
Puffer: 10-fach Puffer, vom Hersteller jeweils mitgeliefert
BSA: 100-fach, 10 mg/ml bovines Serum-Albumin

Restriktionsendonukleasen erkennen spezifische kurze DNA-Sequenzmotive von 4 bis ungefähr 10 bp (meist 6). Der DNA-Strang wird innerhalb des Erkennungsmotivs hydrolytisch gespalten, wobei die Schnittstelle je nach Spezifität der Endonuklease glatt oder 5'- bzw. 3'- überhängend sein kann. Die Reaktionsbedingungen (Temperatur, Puffer und gegebenenfalls BSA-Zusatz) wurden nach Angaben des Herstellers gewählt. Pro µg DNA wurden 1-2 Units Enzym zugesetzt. Die Restriktionsenzyme zeigten bei Zugabe des spezifischen Puffers 100 % Aktivität. Bei Verwendung von zwei unterschiedlichen Enzymen in einem Restriktionsansatz wurde ein Puffer eingesetzt, der für beide Enzyme geeignet war. Dabei richtet sich das Volumen des Restriktionsansatzes nach der

eingesetzten DNA-Menge, die nicht mehr als 1 µg/10 µl Ansatz betragen sollte. Kontrollrestriktionen von Mini-, Midi- und Maxipräparationen wurden für circa zwei Stunden in 10-20 µl Gesamtvolumen inkubiert. DNA-Material für Chromatinimmunopräzipitation wurde über Nacht in einem größeren Volumen inkubiert.

4.4.3.3. Auftrennung der DNA-Fragmente im Agarosegel

Agarose:	Seakem LE Agarose FMC (Invitrogen)
TAE-Puffer:	40 mM Tris-Acetat; 1 mM Ethylendiamintetraessigsäure
Ethidiumbromid-Lösung:	10 mg Ethidiumbromid/ml H ₂ O
DNA-Probenpuffer (6×):	Orange Loading Dye (Fermentas)
Agarosegelkammer:	diverse Gelapparaturen (NEB, Owl Scientific)
Transformator:	Power PAC 300 (BIO-RAD)
DNA-Größenstandard:	Gene Ruler TM ultra low range (Fermentas) Gene Ruler TM 1 kb DNA ladder plus (Fermentas)
Transilluminator:	UV-Star (Biometra)

Die mit Probenpuffer versetzten DNA-Fragmente aus PCR oder Restriktionsansätzen wurden nach ihrer Molekülgröße in 1 % bis 3 % (w/v) Agarosegelen elektrophoretisch aufgetrennt. Zur Herstellung eines Gels wurden 1-3 g Agarose in 100 ml Tris-Acetat-Puffer durch Erhitzen gelöst. Auf Grund der negativen Ladung der Phosphatreste in der DNA wandert diese im elektrischen Feld der Gelkammer zur positiven Anode. Die DNA-Fragmente wurden üblicherweise bei einer Spannung von 120 V für dreißig bis sechzig Minuten in Tris-Acetat-Laufpuffer getrennt. Die Wanderungsgeschwindigkeit ist bei konstanter Feldstärke umgekehrt proportional zur Molekülgröße. Nach der Größenauftrennung wurde das Gel für fünf Minuten in einem Ethidiumbromidbad gefärbt. Das zugesetzte Ethidiumbromid interkaliert mit der doppelsträngigen DNA und ermöglicht so ihre Detektion unter UV-Licht ($\lambda = 245 \text{ nm}$). Die Dokumentation der Komplexe aus Ethidiumbromid und doppelsträngiger DNA erfolgte an einer UV-Durchlichtapparatur mit Software und Kamera der Firma INTAS.

4.4.3.4. DNA Extraktion aus Agarosegelen

Kit: MinElute® (Qiagen)

Nach der Auftrennung über ein Agarosegel (4.4.3.3.) wurden die spezifischen Banden der zu verwendenden DNA eng aus dem Gel ausgeschnitten. Mit Hilfe des „MinElute® Kits“ (Qiagen) wurde die gewünschte DNA nach Herstellerangaben in 10 µl Elutionspuffer (vorgewärmt auf 50 °C!) gewonnen. Die DNA-Konzentration wurde photometrisch bestimmt (4.4.1.1.).

4.4.3.5. Ligation von DNA in das pGEM®-T Easy Vektor System

Enzym: T4 Ligase (NEB)

Puffer: 5-fach Puffer (NEB)

Kit: pGEM-T Easy Vector System (Promega)

Die Reaktion wurde in einem Gesamtvolumen von 20 µl durchgeführt. Die Menge an linearisiertem Vektor betrug 50-100 ng. Das zu klonierende Fragment wurde in einem 2-5 fachen molaren Überschuss zugegeben. Nach Zugabe von 0,5-1 U T4-Ligase wurde der Ansatz über Nacht bei 4 °C inkubiert und anschließend in Bakterien transformiert (4.3.5. – 4.3.8.). Die Insertmenge kann bei vorgegebener Vektormenge durch folgende Formel berechnet werden:

$$\text{Insert [ng]} = \frac{\text{Vektor [ng]} * \text{Insertgröße [Kb]}}{\text{Vektorgröße [Kb]}} * \text{molares Verhältnis} \quad \text{Insert / Vektor}$$

4.4.3.6. Polymerase-Kettenreaktion (PCR) (Saiki *et al.*, 1985)

Sterilbank: Lamin Air HLB 2448 GS (Heraeus)

Cycler: Bio-Rad iCycler

Polymerase: Ampli Taq, Ampli Taq Gold DNA Polymerase
(Applied Biosystems)

Puffer: 10x Puffer (Applied Biosystems)

dNTPs: 2,5 mM Desoxyadenosin-, 2,5 mM Desoxycytidin, 2,5 m
Desoxyguanosin, 2,5 mM Desoxythymidin
(Applied Biosystems)

Zusatz: 5 mM MgCl₂ (Applied Biosystems)

Mit Hilfe der PCR wurden definierte DNA-Bereiche amplifiziert. Zur Erzielung reproduzierbarer PCR-Ergebnisse ist es nötig, die verwendeten Primerpaare zuvor bei unterschiedlichen Annealingtemperaturen zu testen. Alle Parameter müssen für die jeweilige Reaktion optimiert werden. Eine Tabelle der verwendeten PCR-Primer befindet sich im Anhang (8.1.). Abbildung 30 gibt einen Überblick über ein repräsentatives PCR Programm (Abb.: 31A) und einen typischen 25 µl PCR-Ansatz (Abb.: 31B).

A				B	
Zyklen	Reaktion	Temperatur [°C]	Zeit [min]	PCR-Komponente	Volumen [µl]
1	Denaturierung	94	1	10 x PCR-Puffer	5
35	Denaturierung	94	1	25 mM MgCl ₂	3
	Primer-Annealing	spezifisch	1	dNTP-Mix [je 2,5 mM]	4
	Elongation	72	1min/1kb	Primer I [10 pmol/µl]	3
1	Primerextension	72	5	Primer II [10 pmol/µl]	3
				TAQ-Polymerase	1
				steriles Wasser	29
				DNA-Template	2

Abbildung 30: **A)** Übersicht über die typischen Ablauf einer Polymerasekettenreaktion mit Angabe der gewöhnlich verwendeten Zyklenzahl . **B)** Pipetierschema für einen 25 µl PCR-Ansatz.

4.5. Electrophoretic Mobility Shift Assay (EMSA) und Supershift Assay

4.5.1. Radioaktive Markierung von Nukleinsäuren

Kinase:	T4 Polynukleotid Kinase 10 U/µl (NEB)
Puffer:	10-fach (NEB)
Radioaktivität:	Adenosin 5'-[γ32P]-triphosphat, 10 µCi/µl (Hartmann Analytics)
Gelfiltration:	Nick Column (Pharmacia)
Szintillationszähler:	LS 6500; (Beckman Coulter)
TE-Puffer:	10 mM Tris-HCl pH [8], 1 mM Ethylendiamintetraessigsäure

Zur Endmarkierung mittels Kinasierung wurden 3,5 pmol DNA mit 1 µl 10-fach Puffer, 10 U T4 Polynukleotidkinase und 10 µCi [α 32P]ATP in einem Gesamtvolumen von 10 µl für dreißig Minuten bei 37 °C inkubiert. Die markierte DNA wurde über eine Nick-Säule entsprechend der Herstellerangaben gereinigt. Die Cerenkov-Strahlung jeder erhaltenen Fraktion wurde im Szintillationszähler bestimmt.

4.5.2. Oligonukleotid-Hybridisierung

Die von der Firma Eurofins MWG Operon (Ebersberg) bezogenen jeweils komplementären einzelsträngigen Oligonukleotide wurden, bevor sie zur Kompetition im EMSA (4.5.3.) eingesetzt werden konnten, zu einem Doppelstrang hybridisiert. Gleiche Mengen in Wasser gelöster komplementärer Oligos ($\sim 100 \mu\text{M}$) wurden im Thermocycler dreißig Sekunden auf 95°C erhitzt um Sekundärstrukturen zu lösen und anschließend jeweils zwei Minuten bei 72°C , 37°C und 25°C inkubiert. Die entstandene doppelsträngige DNA wurde anschließend auf Eis gestellt, aliquotiert und bei -20°C gelagert.

4.5.3. Protein-Nukleinsäure Bindungsreaktion

10-fach EMSA-Puffer:	100 mM Tris-HCl pH [7,5], 500 mM NaCl, 10 mM MgCl_2 , 5 mM Ethylendiamintetraessigsäure, 5 mM Dithiothreitol, 40 % Glycerin, 0,5 $\mu\text{g}/\mu\text{l}$ Poly(dI:dC), 5 mM ZnCl_2
10-fach Gel-Ladepuffer:	250 mM Tris-HCl pH [7,5], 0,2 % Bromphenolblau, 0,2 % Xylencyanol, 40 % Glycerin
Rekombinante Proteine:	humanes recombinantes Sp1 Protein (AXXORA) humanes recombinantes Sp3 Protein (AXXORA)

In einen Reaktionsansatz von $20 \mu\text{l}$ gingen 20000 cpm ($\sim 0,035 \text{ pmol}$ Sonde) (4.5.1.) und $5 \mu\text{g}$ nukleärer Proteinextrakt (4.7.2.) in 1-fach EMSA-Puffer ein. Bei Verwendung rekombinanter Proteine wurden jeweils 10 ng eingesetzt. Die Ansätze wurden für dreißig Minuten bei RT inkubiert. Kompetitive-Ansätze enthielten zusätzlich spezifische unmarkierte Kompetitor-DNA in variablen Mengen, meist in 50- bis 100-fachen molaren Überschuss. Diese wurden vor Zugabe der markierten Sonden zwanzig Minuten bei RT mit den nukleären Extrakten vorinkubiert. Die Reaktion wurde mit $2 \mu\text{l}$ 10-fach Gel-Ladepuffer gestoppt und anschließend auf ein natives Polyacrylamid-Gel aufgetragen (Fried und Crothers, 1981; Garner und Revzin, 1981; Lane *et al.*, 1992) (4.5.4.). Für Supershift-Assays wurde die Reaktion nach Zugabe der markierten Sonden-DNA mit $2 \mu\text{g}$ Antiserum versetzt und vor dem Auftragen auf das native Gel zwanzig Minuten bei RT inkubiert.

4.5.4. native Polyacrylamid Gel Elektrophorese

Geltrockner:	Gel Dryer Model 583 (Biorad)
Gelapparatur:	Protean II xi System (BioRad)
Wasserkühlung:	MGW LAUDA RMT / RM6
TBE-Puffer:	90 mM Tris-Borat pH [7,2], 2 mM Ethylendiamintetraessigsäure
Laufpuffer:	0,5-fach TBE-Puffer
APS:	10 % Ammonium Persulfat (Sigma)
TEMED:	Tetramethylethylendiamin (SERVA)
Film:	Medizinische Röntgenfilme Super RX Fuji

Die EMSA-Ansätze (4.5.3.) wurden in einem 5 %igem nativen Polyacrylamidgel elektrophoretisch aufgetrennt. Für 80 ml Gelansatz wurden 10,3 ml Rothiphores Gel A (Roth), 5,3 ml Rothiphores Gel B (Roth), 4 ml 10-fach TBE, 2,5 ml 80 % Glycerin in Wasser, 57,4 ml Wasser, 450 µl APS und 90 µl TEMED verwendet. Vor dem Auftragen der Proben wurden die Gele auf 12 °C gekühlt und zwanzig Minuten bei 15 V/cm gereinigt. Die Elektrophorese fand bei 10 V/cm und 4 °C statt. Das Gel wurde auf 3 mm Whatmanfilterpapier überführt und bei 80 °C für mindestens sechzig Minuten getrocknet. Zur Detektion der Signale wurde über Nacht ein Röntgenfilm aufgelegt.

4.6. Bestimmung des proviralen Transkriptionsstart und -endpunktes

Kit: GeneRacer® Kit with SuperScript™ III RT (Invitrogen)

Die Bestimmung des Transkriptionsstart und -endpunktes der HERV-K Proviren wurde mit Hilfe der *RNA ligase-mediated rapid amplification of 5' and 3' cDNA ends* (RLM-RACE) Technik ermöglicht. Die Prozedur wurde nach Herstellerangaben durchgeführt. In die Reaktion wurden 2 µg messenger RNA eingesetzt, die zuvor einer DNaseI-Behandlung unterzogen worden waren (4.4.2.4.). Einen Überblick über die einzelnen Schritte gibt Abbildung 31. Nach erfolgreicher PCR (Primerliste siehe 8.1.) unter Verwendung folgender Primerpaare wurden die PCR-Produkte sequenziert:

I Bestimmung des Transkriptionsstarts.

PCR	GeneRacer™ 5' Primer & B2 (HERV-K PBS)
Nested PCR	GeneRacer™ 5' Nested Primer & Lys1,2 (HERV-K PBS)

II Bestimmung des Transkriptionsendpunktes:

PCR GeneRacer™ 3' Primer & Env-sense

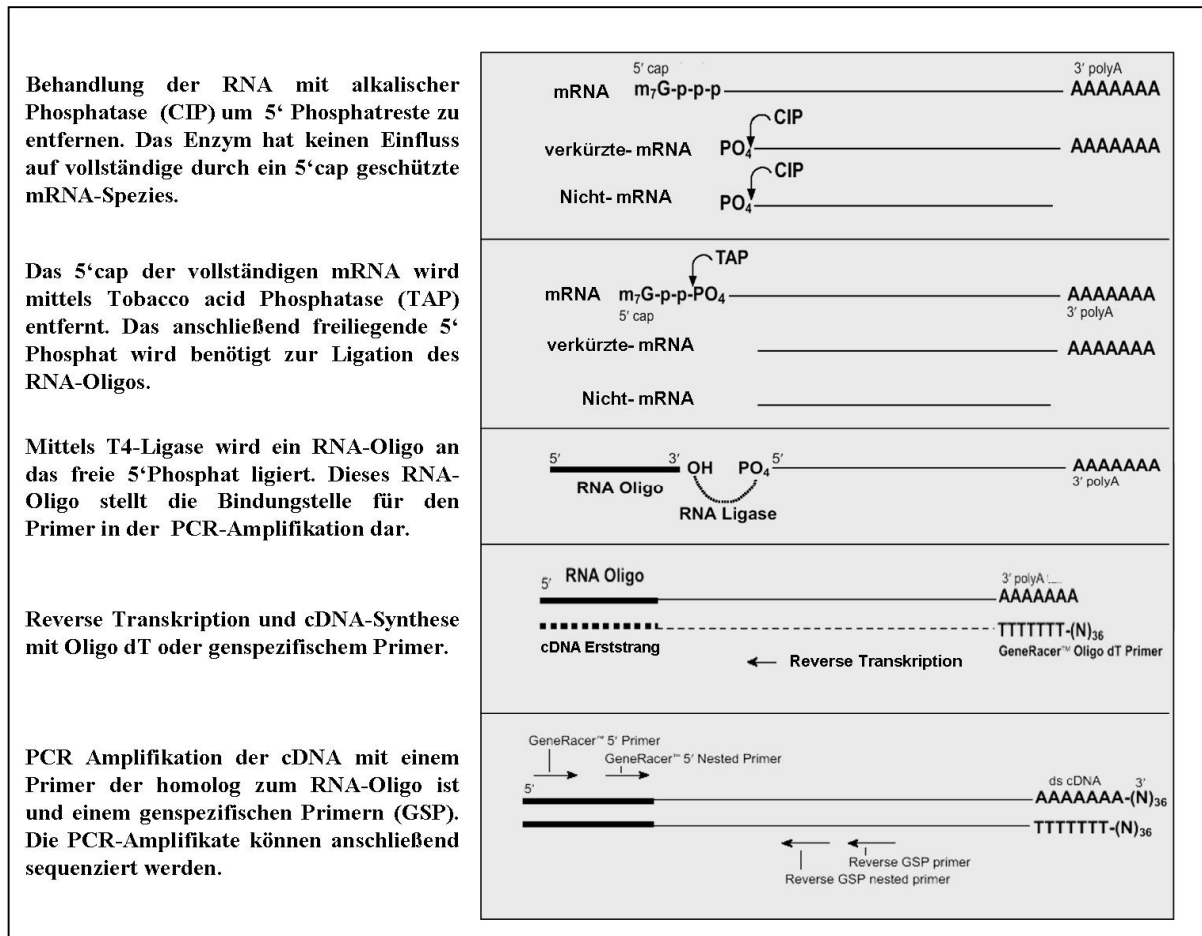


Abbildung 31: Schematische Übersicht über die verschiedenen Schritte der 5'RACE

4.7. Proteinchemische Methoden

4.7.1. Proteinextraktion aus eukaryotischen Zellen

- Proteinaseinhibitoren: Complete Protease Inhibitor Cocktail Tablets (Roche), Phenylmethylsulfonylfluorid [PMSF] (Fluka)
- Tripple-Lysis-Buffer: 20 mM Tris-HCl pH [7,5], 150 mM NaCl, 10 mM Ethylendiamintetraessigsäure, 0,1 % SDS, 1 % Triton X-100, 1 % Desoxycholat, Proteaseinhibitoren: 1 mM PMSF, 1/25 Complete, frisch zugeben!

Vor der Lyse wurden die Zellen zweifach mit PBS⁻ gewaschen, dann mit Lysepuffer mit Proteaseinhibitor belegt (200 µl für ein Loch einer 6-Loch-Platte, 1 ml für eine T75

Flasche) und unter Schwenken einige Minuten inkubiert. Das Lysat wurde in 2 ml Plastikreaktionsgefäße überführt und ab diesem Zeitpunkt auf Eis gehalten. Die Zellen wurden mit Ultraschall behandelt (Duty Cycle30, Output Contr. 1) und eine Stunde auf Eis inkubiert. Zellbruchstücke wurden anschließend durch Zentrifugation bei 13000 rpm für drei bis fünf Minuten bei 4 °C entfernt. Der Überstand wurde zur Proteinbestimmung eingesetzt, aliquotiert und bei 80 °C gelagert.

4.7.2. Fraktionierte Proteinpräparation aus Zytoplasma und Kern

Proteinaseinhibitor:	Complete Protease Inhibitor Cocktail Tablets (Roche)
Photometer:	Ultrospec 3100pico (Amersham Biosciences)
Dialyse-Kassetten:	Slide-A-Lyzer Dialysis Cassette, 3,500 MWCO (Pierce)
PE-Puffer A:	10 mM 2-(4-(2-Hydroxyethyl)-1-piperazinyl) ethansulfonsäure (HEPES) pH [7,9], 0,3 M Saccharose, 1,5 mM MgCl ₂ , 10 mM KCl, 0,1 mM Ethylenglykoltetraessigsäure (EGTA), frisch zugeben!: 0,5 mM Dithiothreitol, 1/25 Complete-Proteinaseinhibitor für GH Zellen zusätzlich 50 mM NaCl, 0,2 % NP40
PE-Puffer B:	20 mM HEPES pH [7,9], 420 mM NaCl, 1,5 mM MgCl ₂ , 0,1 mM EGTA, 25 % Glycerin (v/v), frisch zugeben!: 0,5 mM Dithiothreitol, 1/25 Complete-Proteinaseinhibitor
PE-Puffer C:	300 mM HEPES pH [7,9], 1,4 M KCl, 30 mM MgCl ₂ , frisch zugeben!: 0,5 mM Dithiothreitol, 1/25 Complete-Proteinaseinhibitor
PE-Puffer D:	20 mM HEPES pH [7,9], 75 mM NaCl, 0,1 mM Ethylendiamintetraessigsäure, 20 % Glycerin (v/v), frisch zugeben!: 0,5 mM Dithiothreitol, 1/25 Complete-Proteinaseinhibitor

Das Protokoll zur fraktionierten Proteinaufreinigung wurde nach Dignam (Dignam *et al.*, 1983) leicht modifiziert. Die in T175 Flaschen (Nunc) herangezogenen Zellen wurden zweimal mit kaltem PBS⁻ gewaschen und mit einem Zellschaber in 10 ml PBS⁻ pro Kulturflaschen vom Flaschenboden mechanisch gelöst. Die Zellen wurden für zehn Minuten bei 900 rpm, 4 °C pelletiert und im 5-fachen Zellvolumen PBS⁻ (4 °C)

resuspendiert. Nach erneuter Zentrifugation wurden sie im drei- (GH) bis fünffachen Zellvolumen PE-Puffer A (4 °C) resuspendiert und zehn Minuten auf Eis inkubiert. Im Glas-Homogenisator (Braun) wurden die Zellen mechanisch mit zehn bis zwanzig (GH) Stempelhüben aufgeschlossen. Der erfolgreiche Aufschluss wurde mikroskopisch kontrolliert. Die intakten Zellkerne wurden für zehn Minuten bei 2500 rpm, 4 °C pelletiert. Zur Präparation des zytoplasmatischen Protein-Extraktes wurde der Überstand mit 0,11 Volumen PE-Puffer C versetzt und für fünfzig Minuten bei 50000 rpm (SW-60), 4 °C abzentrifugiert. Der Überstand wurde über Nacht gegen das 20-fache Volumen PE-Puffer D bei 4 °C dialysiert. Der gewonnene zytoplasmatische Protein-Extrakt (= Dialysat) wurde für zwei Minuten bei 14000 rpm 4 °C zentrifugiert, der Überstand aliquotiert, und in flüssigem Stickstoff schockgefroren. Zur Präparation des nukleären Protein-Extraktes wurde das Zellkernpellet mit dem 2,5-fachen Volumen PE-Puffer B versetzt und im Glas-Homogenisator mit mindestens zwanzig Stempelhüben mechanisch aufgeschlossen (mikroskopische Kontrolle!). Nach leichter Agitation auf einem Magnetrührer für dreißig Minuten 4 °C wurde die Suspension für fünfzig Minuten bei 50000 rpm (SW-60), 4 °C zentrifugiert. Der Überstand wurde über Nacht gegen das 50-fache Volumen PE-Puffer D bei 4 °C dialysiert. Das Dialysat wurde für zwei Minuten bei 14000 rpm, 4 °C zentrifugiert. Der nukleäre Protein-Extrakt (= Überstand) wurde aliquotiert, in flüssigem Stickstoff gefroren und bei -80 °C gelagert. Der Proteingehalt der zytoplasmatischen und nukleären Extrakte wurde mit dem BioRad Protein-Assay (BioRad) entsprechend der Herstellerangaben quantifiziert (4.7.3.).

4.7.3. Proteinbestimmung (nach Bradford)

Färbe-Reagenz:	5-fach dye reagent concentrate (Bio Rad)
BSA:	10 mg/ml bovines Serum Albumin
Küvetten:	1,5 ml halbmikro Einmalküvetten (Brand)
Spektralphotometer:	Spectronic 1201 (Milton Roy)

Die colorimetrische Konzentrationsbestimmung von Proteinen mit dem BioRad-Reagenz beruht auf der pH-Wert abhängigen Bindung des Farbstoffes Coomassie Brilliantblau G 250. In saurem Milieu bindet G250 an basische und aromatische Aminosäurereste. In Folge der Komplexbildung mit dem Protein kommt es zur Stabilisation der unprotonierten, anionischen Sulfonatform des Farbstoffes und somit zu einer Verschiebung des Absorptionsmaximums von 465 nm zu 595 nm. Zur Standardisierung des Proteingehalts wurde eine Eichgrade mit Lösungen bekannter Proteingehalte (0-50 mg/ml BSA) erstellt.

4.7.4. SDS-Polyacrylamidgel-Elektrophorese (SDS-PAGE)

Gelapparatur:	Mighty small II SE 250 (Hoefer Scientific Instruments)
Transformator:	Electrophoresis Power Supply-EPS 301 (Amersham Pharmacia Biotech)
DTT:	30-fach; Dithiothreitol (NEB)
Laufpuffer:	25 mM Tris-HCl pH [8,3], 250 mM Glycin, 0,1 % SDS [w/v]
Auftragspuffer:	3-fach SDS-Auftragspuffer (NEB)
Marker:	PageRuler™ Prestained Protein Ladder (Fermentas) PageRuler™ Unstained Protein Ladder (Fermentas)
Lösung A:	30 % Acrylamid; 0,8 % Bisacrylamid
Lösung B:	1,5 M Tris; 0,4 % SDS, pH [8,8]
Lösung C:	0,5 M Tris; 0,4 % SDS; pH [6,8]
APS:	10 % Ammonium Persulfat (Sigma)
TEMED:	Tetramethylethyldiamin (SERVA)

Über die SDS-Polyacrylamid-Gelelektrophorese lassen sich Proteine effizient entsprechend ihres apparenten Molekulargewichts auftrennen. Polyacrylamid bildet durch Kopolymerisation mit dem quernetzenden N'N'-Methylenbisacrylamid eine Gelmatrix mit definierter Porengröße. Bei der Elektrophorese werden die Proteine entsprechend ihrer Größe und Ladung in der Polyacrylamid (PAA)-Matrix zurückgehalten. Die Porengröße lässt sich dem gewünschten Trennbereich entsprechend anpassen, so ist ein 8 %iges Gel geeignet, um Proteine von 50 bis 200 kDa zu trennen, ein 15 %iges für Proteine zwischen 10 und 60 kDa. Ein 10 %iges Gel eignet sich für Molekulargewichte von 30 bis 120 kDa. Das im Ladepuffer enthaltene SDS (Sodiumdodecylsulfat) ist ein anionisches Detergenz, das an die hydrophoben Regionen des Proteins bindet und diese annähernd proportional zu ihrer Größe mit negativen Ladungen versieht. Pro Gramm Protein binden circa 1,4 g SDS. Die Proben wurden mit einem entsprechenden Volumen an 3-fach Ladepuffer versetzt und für zehn Minuten bei 97 °C inkubiert. Diese Behandlung denaturiert die Proteine durch Auflösung der Tertiär- und Sekundärstrukturen sowie der Wasserstoffbrücken. Die Proteine liegen anschließend in gestreckter Form vor. Durch Zugabe von Dithiothreitol oder β -Mercaptoethanol werden Sulfhydrylgruppen reduziert.

Die Auftrennung der Proteine erfolgte in 12,5 % (w/v)-Acrylamidgelen unter Verwendung von 1-fach SDS-Laufpuffer. Durch Verwendung eines vorgeschalteten 5 % (w/v)-Sammelgels, mit einem pH-Wert von 6,8 wurden schärfere Banden im Gel erzielt.

Zusammensetzung siehe Tabelle 6. Zur Detektion des HERV-K Rec Proteins wurden ausschließlich 4-12 % NuPage-Gele (Invitrogen) mit einer Dicke von 1 mm mit 1-fach MOPS-Puffer (Invitrogen) verwendet.

Die Trennung der Proteine erfolgte bei einer konstant angelegten Stromstärke von 60 Milliampere. Die Proteine wandern auf Grund ihrer negativen Ladung zur Anode. Die Abschätzung der Molekulargewichte nach der Elektrophorese wurde durch Vergleich mit aufgetragenen Markerproteinen möglich.

Tabelle 6: Zusammensetzung der SDS-PAGE

	12,5 % Trenngel	5 % Sammelgel
Wasser	2,48 ml	2,9 ml
Lsg. A	3,15 ml	0,83 ml
Lsg. B	1,88 ml	-
Lsg. C	-	1,25 ml
APS 10%	75 µl	40 µl
Temed	4 µl	4 µl

4.7.5. Luziferase-Assay

Lysispuffer:	Passive Lysis Buffer, 5x (Promega)
Pipettenspitzen:	Biozym budget line Safe Seal Tips (Biozym)
Transfektionsreagenz:	Reporter: Lipofectamine (Invitrogen) siRNA+ Reporter: jetSI TM -ENDO (Polyplus)
6-Loch-Platten:	Nunclon MicroWell Platte (Nalge Nunc International)
96-Loch-Platten (weiß):	LumiNunc Platte (Nalge Nunc International)
Luminometer:	MicroLumat Plus (EG&G Berthold)
Reporter-Assay-Kit	Dual Luciferase Assay (Promega)
siRNA:	Hs_SP1_1_HP, Hs_SP3_1_HP (Qiagen).

4.7.5.1. Dualer Luziferase-Assay

Die transiente Transfektion eukaryotischer Zellen für den dualen Luziferaseassay erfolgte mit Lipofectamin nach den Angaben des Herstellers. Um DNA in Zellen einzubringen, wird deren negative Ladung durch Komplexbildung mit kationischen Lipiden maskiert. In dieser Form können die DNA-Lipidkomplexe sich an die Zellmembran anlagern und diese passieren. Am Vortag der Transfektion wurden 6-Loch-Platten angelegt mit $0,5 \times 10^6$

eukaryotische Zellen pro Loch. In die Transfektion wurden 900 ng Reporter Luziferasevektor (pGL3), und 100 ng Renilla-Vektor (pHRLnull) in FKS freiem Medium eingesetzt. Die verwendeten Konstrukte sind in Abbildung 32 aufgeführt. Nach dreistündiger Inkubation der Transfektionsansätze wurde das Transfektionsmedium gegen FKS-haltiges Vollmedium ausgetauscht. 24 oder 48 Stunden nach Transfektion wurden die Zellen mit PBS⁻ gewaschen und mit Lysepuffer behandelt und entsprechend des Protokolls in den dualen Luziferase-Assay eingesetzt. Pro Loch einer 6-Loch-Platte wurden 500 µl 1-fach Passive-Lysis-Puffer (Promega) eingesetzt. Zellen und Lysepuffer wurden für fünfzehn Minuten inkubiert. Das Lysat wurde in ein 1,5 ml Plastikreaktionsgefäße überführt und zwei Minuten bei 14000 rpm und 4 °C zentrifugiert. 20 µl des Überstandes wurden für die Messung der Luziferaseaktivität eingesetzt. Die Messung der Luziferaseaktivität erfolgte im Luminometer nach folgendem Schema: Zugabe von 50 µl Luziferasesubstrat, zwei Sekunden Verzögerung gefolgt von der Messung der Photonenemission für zehn Sekunden. Hieraus ergab sich der Wert für die Firefly – Luziferase. Anschließend erfolgte die Zugabe von 50 µl Stop&Glow Reagent mit einer erneuten zehnhündigen Messung der Photonenemission nach einer Verzögerung von zwei Sekunden. Auf diese Weise wurde der Wert für die kotransfizierte Renilla Luziferase bestimmt. Der Quotient aus beiden Werten ergab einen normalisierten Wert, der Fehlerquellen wie unterschiedliche Zellzahlen oder unterschiedliche Transfektionseffizienzen minimierte.

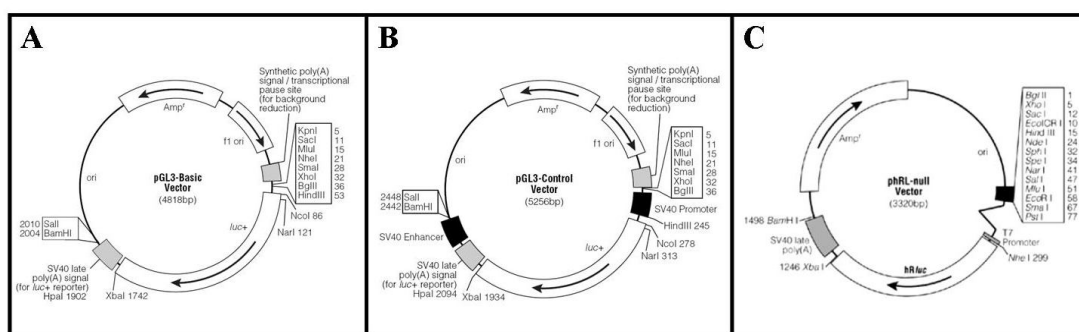


Abbildung 32: Übersicht über die im Luziferaseassay verwendeten Konstrukte. A) Der PGL3-basic-Vektor (Promega) enthält keine eukaryotischen Promotor oder Enhancer Sequenzen vor dem Markergen (Glühwürmchen Luziferase). Die HERV-K LTRs pcK30 und LTR 21 wurden von M. Kraft (Kraft, 2004) über HindIII aus dem pSVOAL Vektor in den pGL3 basic ligiert. Die Mutationskonstrukte der TATA und GC-Box wurden ebenfalls von M. Kraft erzeugt. B) Der pGL3-control Vektor (Promega) enthält einem SV40-Promotor und entsprechende Enhancer-Sequenzen, die eine starke Expression des nachgeschalteten Luziferasegens in vielen Geweben verursachen. C) Das Plasmid pHRL-null enthält keine Promotor oder Enhancer Sequenzen vor dem Luziferasegen der Weichkoralle *Renilla reniformis*, dieses Plasmid wurde in einigen Experimenten zur Normierung kotransfiziert.

4.7.5.2. Luziferase-Assay mit siRNA-*Knockdown*

Die kleinen doppelsträngigen siRNA werden mittels Transfektionsreagenz in die Zielzelle eingebracht und anschließend in einen Proteinkomplex RISC (RNA-induced silencing complex) eingefügt. Mit Hilfe der inkorporierten siRNA-Fragmente bindet RISC komplementär an spezifische Ziel-*mRNA*. Die Ziel-*mRNA* wird durch RNA-Helicase Nuklease-Aktivitäten des RISC-Komplexes entwunden und gespalten. Diese führt zum Abbau der Ziel-*mRNA* durch intrazelluläre Nukleasen.

Zur Kotransfektion von siRNA und Reporterkonstrukten wurden 1 µg Reporter und 150-400 ng siRNA mit dem Transfektionsreagenz jetSITM-ENDO in $0,5 \times 10^6$ Zellen eingebracht. Dabei wurden die Angaben des Herstellers beachtet. Mit siRNA transfizierte Zellen wurden nach 48 Stunden lysiert auf die gleiche Weise behandelt wie die Proben für den dualen Luziferaseassay (4.7.5.1). 20 µl des Überstandes wurden für die Messung der Luziferaseaktivität eingesetzt. Die Messung der Luziferaseaktivität erfolgte im Luminometer nach folgendem Schema: Zugabe von 50 µl Luziferasesubstrat, zwei Sekunden Verzögerung gefolgt von der Messung der Photonenemission für zehn Sekunden. Hieraus ergab sich der Wert für die Firefly –Luziferase. Da kein Renilla-Plasmid zur Normalisierung kotransfiziert werden konnte wurden die Messwerte auf den Gesamtproteingehalt der Lysate normiert.

In Paralleleanalysen wurden Extrakte zur Verwendung im Immunoblot hergestellt, dazu wurden die Zellen mit 200 µl Trippel-Lysepuffer belegt (4.7.1.), mit Ultraschall behandelt und der Proteingehalt bestimmt (4.7.3.).

4.8. Immunologische Methoden

4.8.1. Proteinnachweis durch Antikörper auf Trägermembran (Immunoblot)

Blotkammer:	TE Series Transfer Electrophoresis Unit (Hoefer Scientific Instruments)
Blotmembran:	Protran; Nitrozellulose Transfer Membrane, 0,2 µm Porengröße (Whatman)
Blotpuffer:	25 mM Tris-HCl pH [8,6], 192 mM Glycin, 15 % Methanol
Filterpapier:	3 mm (Whatman)

Färbelösung:	Ponceau S in 5 Essigsäure (Sigma)
Blockierlösung:	4 % Magermilchpulver (Merck) in PBS ⁻ mit 0,05 % Tween 20 (Sigma) (PBST)
Waschlösung:	0,05 % Tween 20 (Sigma) in (PBST)
anti-HERV-K-Rec:	aus Kaninchen (Löwer <i>et al.</i> , 1995), 1:2000 in 5 % Magermilch/PBST verdünnt
anti-HERV-K-Rec:	aus Maus (monoklonal), 1:10000 bereitgestellt von Dr. Thomas Muster (Green Hills Biotechnology)
anti-HERV-K-Env:	aus Maus (monoklonal) (Austral Biologicals), 1:5000 in 5 % Magermilch/PBST
anti-HERV-K-Gag:	aus Maus HERMA 6, HERMA 7 (Boller <i>et al.</i> ; 2008) unverdünnt
anti-β-Aktin:	aus Maus (Sigma), 1:15000 in 5 % Magermilch/PBST
anti-LaminA:	aus Kaninchen, Biozol, 1:3000 in 5 % Magermilch/PBST
anti-Kaninchen:	aus Ziege, Peroxidase gekoppelt (Sigma), 1:20000 in 5 % Magermilch/PBST verdünnt
anti-Maus:	aus Ziege, Peroxidase gekoppelt (Sigma), 1:20000 in 5 % Magermilch/PBST verdünnt
Detektion:	ECL plus Western Blotting Detection System (Amersham)
Filme:	Hyperfilm ECL (Amersham)

Das Blotverfahren beruht auf dem Transfer der Proteine aus der Polyacrylamid-Gelmatrix auf eine Trägermembran (Nitrocellulose) mittels eines elektrischen Feldes. Auf die Anode wurde dazu ein in Transferpuffer äquilibrierter Schwamm und ein ebenfalls getränktes Filterpapier (Whatman) in Gelgröße (8,5 mm x 7,5 mm) gelegt, gefolgt von der Membran, dem Gel, einem weiteren feuchten Filterpapier und abschließend einem weiteren Schwamm. Der Blot wurde in den mit Transferpuffer gefüllten Tank eingesetzt und die Spannung angelegt. Die Überführungszeit der Proteine vom Gel auf die Membran betrug zwei Stunden bei 4 °C, einer Spannung von 50 V und maximaler Stromstärke. Der Übertragungserfolg auf die Nitrozellulosemembran wurde mittels Ponceau-Rot überprüft. Der Farbstoff färbt die geblotteten Proteine mit einer Empfindlichkeit von 50 ng an und fixiert diese zusätzlich auf der Membran durch die enthaltene Essigsäure. Da der Farbstoff reversibel bindet, ist eine völlige Entfernung mit PBS⁻ + 0,05 % Tween 20 (PBST) möglich.

Zum Nachweis spezifischer Proteine mittels Antikörper wurde die Nitrozellulosemembran zunächst über Nacht bei 4 °C mit Blockier-Lösung inkubiert, um unspezifische Bindungsstellen auf der Membran abzusättigen. Danach wurde die Membran ein Mal mit PBST gewaschen und anschließend mit einer entsprechenden Verdünnung des primären Antikörpers belegt. Die Inkubation erfolgte für eine Stunde bei RT unter leichtem Schwenken. Während dieser Zeit können die auf der Membran immobilisierten Proteine mit dem primären Antikörper reagieren. Danach wurde die Membran vier Mal für je fünf Minuten mit PBST gewaschen. Anschließend erfolgte für eine Stunde die Inkubation mit dem sekundären Antikörper. Um nicht gebundenen Sekundäreantikörper zu entfernen, wurde die Membran vier Mal für je fünf Minuten mit PBST gewaschen. An den sekundären Antikörper ist das Enzym Meerrettich-Peroxidase gekoppelt. Die Peroxidase katalysiert die Umsetzung des Substrates (Luminol). Bei dieser Reaktion werden Photonen emittiert, die mittels eines Films detektiert werden können. Die Expositionszeit betrug zwischen 20 Sekunden und 24 Stunden.

4.8.2. Immunopräzipitation

Sepharose:	Protein-G-Sepharose 4 Fast Flow (Amersham)
Ladepuffer:	3-fach SDS Sample Puffer (NEB)
Präimmunserum:	Kaninchen
α -Rec:	aus Kaninchen (Löwer <i>et al.</i> , 1995)

Die Immunopräzipitation ist eine Methode, bei der mittels eines Antikörpers ein Antigen aus einer Lösung aufkonzentriert wird. Um unspezifische Wechselwirkungen zwischen Protein und eingesetzter Sepharose sowie zwischen Protein und Antikörperlösung zu verhindern, wurde zunächst ein geeignetes Volumen Zellysate (500 μ l entspricht einem Zelläquivalent von $\frac{1}{2}$ T75-Flasche) in ein Eppendorf-Gefäß überführt und mit 100 μ l Protein G Sepharose bei 4 °C für eine Stunde auf dem Rollrad vorinkubiert. Die Sepharose wurde zuvor drei Mal mit Proteinlysepuffer gewaschen.

Nach der Inkubation von Extrakt und Sepharose wurde der Ansatz bei 1000 rpm für zwanzig Sekunden abzentrifugiert. Der Überstand wurde in ein neues Eppendorfgefäß überführt und mit einem geeigneten Volumen des entsprechenden Antikörpers (α -Rec: 6 μ l) oder Präimmunserum versetzt und zwei Stunden unter Rotation bei 4 °C inkubiert. Nach Ablauf der Inkubationszeit wurden 50 μ l Protein G Sepharose zugegeben und der

Ansatz für eine weitere Stunde bei 4 °C unter Rotation inkubiert. Die Sepharose wurde bei 1000 rpm für zwanzig Sekunden abzentrifugiert. Der Überstand enthielt die ungebundenen Proteine und wurde als Kontrolle bei -20 °C eingefroren. Die Sepharose mit dem gebundenen Antigen-Antikörper-Komplex wurde vier Mal mit je 1 ml Proteinlysepuffer gewaschen. An die einzelnen Waschschrte schloss sich jeweils ein 20-sekündiger Zentrifugationsschritt bei 1000 rpm an. Anschließend wurde die Sepharose mit einem geeigneten Volumen (30-50 µl) Ladepuffer versetzt und für zehn Minuten bei 95 °C aufgeköcht. Während dieses Vorgangs wird der Komplex gelöst und die Proteine werden von der Sepharose durch Zentrifugation (2000 rpm, zwanzig Sekunden) getrennt. Der Überstand kann bei -20 °C gelagert oder direkt im Immunoblot analysiert werden (4.8.1.). Das Sepharosepellet wird verworfen.

4.8.3. Immunfluoreszenzfärbung

PBS-Dulbecco:	137 mM NaCl, 3 mM KCl, 16,5 mM Na ₂ HPO ₄ , 1,5 mM KH ₂ PO ₄ , 90 mM CaCl ₂ , 50 mM MgCl ₂
Triton X 100:	0,5 % (Merck) in PBS Dulbecco
Fixiermittel:	2 % Paraformaldehyd in PBS Dulbecco
BSA:	2 % Bovines Serum Albumin, Fraktion V (frei von Immunglobulinen; Boehringer Mannheim) in PBS Dulbecco
anti-HERV-K-Rec:	aus Kaninchen (Löwer <i>et al.</i> , 1995), 1:200 bis 1:500
anti-HERV-K-Rec:	aus Maus (monoklonal), 1:3000 bereitgestellt von Dr. Thomas Muster (Green Hills Biotechnology)
anti-HERV-K-Gag:	aus Maus (monoklonal), HERMA 6 & HERMA 7, unverdünnt
anti-HERV-K-Env:	aus Maus (monoklonal), Austral Biologicals, 1:1000
anti-Kaninchen:	aus Esel (Rhodamin) 1:800 (Dianova)
anti-Kaninchen:	aus Esel (FITC) 1:200 (Dianova)
anti-Maus:	aus Esel (FITC) 1:300 (Dianova)
anti-Maus:	aus Esel (Rhodamin) 1:200 (Dianova)
Dapi-Lösung:	1mg/ml, 4',6-Diamidino-2-phenylindole-dihydrochlorid (Sigma)
Objektträger:	Superfrost (Menzel)

Mowiol-Lösung:	2,4 g Mowiol 40-88 (Aldrich!) in 6 g Glycerol suspendieren, 6 ml Aqua bidest zugeben, zwei Stunden bei RT quellen lassen, 12 ml 0,2 M Tris/HCl pH [8,5] zugeben, über Nacht unter Rühren lösen (evtl. bei 50 °C), bei 4000 rpm, fünfzehn Minuten abzentrifugieren, bei -20 °C lagern
Mikroskop:	Axiophot, LSM 510 META (Zeiss)

Zellen wurden auf Deckgläschen ausgebracht und etwa 24 Stunden unter Standardbedingungen im Brutschrank gehalten (4.2.2.). Nach zwanzig Minuten Fixierung mit 2 % Paraformaldehyd bei RT wurden die Zellen für fünfzehn Minuten in 0,5 % Triton X-100 permeabilisiert. Unspezifische Bindestellen wurden durch zwanzigminütige Inkubation mit BSA blockiert. Die Zellen wurden drei Mal mit je 5 ml PBS-Dulbecco gewaschen und mit 100 µl einer entsprechenden Verdünnung der Antikörper belegt. Der Primärantikörper wurde für eine Stunde bei 37 °C inkubiert. Danach wurden die Zellen drei Mal mit PBS-Dulbecco gewaschen und anschließend eine Stunde mit 100 µl einer Verdünnung der fluoreszenzmarkierten Antikörper im Dunkeln inkubiert. Hinterher wurde zwei Mal mit PBS-Dulbecco gewaschen. Zur Darstellung der Nukleinsäure wurden die Zellen zwei Minuten mit Dapi-Lösung inkubiert und anschließend drei Mal mit PBS-Dulbecco gewaschen. Die Zellen auf den Deckgläschen wurden mit Dapi-Lösung belegt um das Kernkompartiment sichtbar zu machen. Nach erneutem mehrmaligem Waschen mit PBS-Dulbecco wurden die Deckgläser mit Mowiol auf den Objektträgern eingedeckt. Die Präparate wurden nach 24 Stunden am Fluoreszenzmikroskop unter Verwendung entsprechender Filter bei 630-facher Vergrößerung ausgewertet. Bilder wurden durch Verwendung des Laserscan-Mikroskopes (LSM 510 META- Carl Zeiss) erstellt.

4.8.4. Chromatinimmunopräzipitation

Kit:	Chromatin Immunoprecipitation Assay Kit (Upstate)
anti-Sp1:	polyklonaler Antikörper aus Kaninchen, 1 µg/µl (Millipore)
anti-Sp3:	polyklonaler Antikörper aus Kaninchen, 200 µg/0.1 ml (Santa Cruz Biotechnology)
anti-YY1:	polyklonaler Antikörper aus Kaninchen, 200 µg/0.1 ml (Santa Cruz Biotechnology)

Die Bindung von Proteinen an die HERV-K LTR wurde mit Hilfe des Chromatin Immunoprecipitation Assay Kits (Upstate) nach den Angaben des Herstellers durchgeführt. Abbildung 32 gibt einen Überblick über die einzelnen Schritte der Methode. Zur Vernetzung von Protein und DNA wurden $0,5 \times 10^6$ Zellen für zehn Minuten in Kulturmedium mit 1 % Formaldehyd inkubiert. Nach Waschen mit PBS⁻ und Lyse der Zellen mit SDS-Puffer (Upstate) wurde die DNA durch Zugabe von Restriktionsenzymen (4.4.3.2.) (jeweils 10 U HindIII, NspI, AflI in Puffer 2 mit BSA) über Nacht bei 37 °C verdaut. Der Verdau wurde auf einem Agarosegel überprüft. Anschließend erfolgte die Präzipitation mit spezifischen Antikörpern. Zur Auswertung der Ergebnisse wurden PCR-Amplifikationen aus Bereichen der HERV-K LTR und des HERV-K *env* Gens durchgeführt. Primersequenzen befinden sich im Anhang (8.1.).

Verwendete Primer: ChIP-LTR-FW & ChIP-LTR-RV
ChIP-Env-FW & ChIP-Env-RV

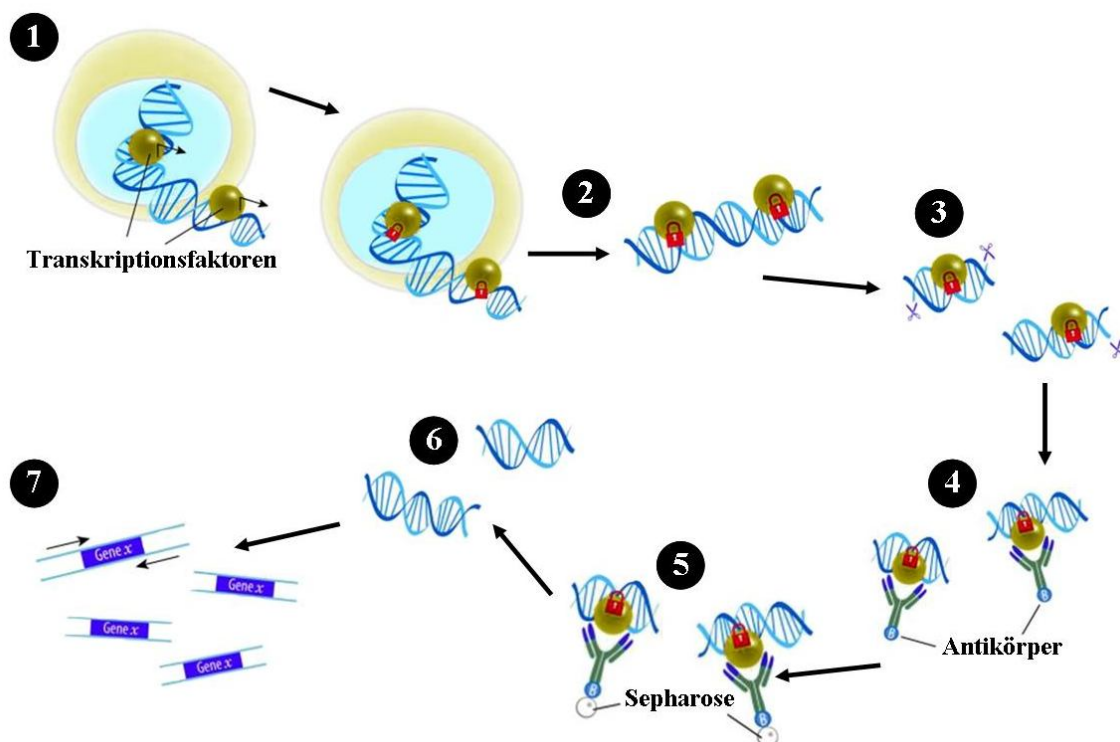


Abbildung 32: Übersicht über die verschiedenen Schritte der Chromatinimmunopräzipitation.

1) Vernetzen von Chromatin und Protein. 2) Aufschluss der Zellen und Isolation der Protein/Chromatin-Komplexe. 3) Fragmentierung der Protein/Chromatin-Komplexe mit Restriktionsenzymen. 4) Immunopräzipitation der Protein/Chromatin-Komplexe mit spezifischen Antikörpern. 5) Bindung der Antikörper an Sepharose und Aufreinigung der Sepharose/Antikörper/Protein/Chromatin-Komplexe. 6) Auflösen des Protein/Chromatin-Komplexes und Verdau der Proteine mit anschließender Aufreinigung der DNA. 7) PCR-Analyse der präzipitierten und aufgereinigten DNA mittels definierter Primerpaare.

4.9. Analyse retroviraler Partikel

Zur Isolierung retroviraler Partikel wurden die Zellen auf T75-Falschen zwei bis drei Tage lang bis zu einer Konfluenz von ~75 % herangezogen. Die Zellkulturüberstände wurden gesammelt und Zelltrümmer durch Zentrifugation bei 9000 rpm (Sorvall Rotor: SLA1500) für fünfundzwanzig Minuten bei 4 °C entfernt. Die Überstände wurden durch eine Membran mit einer Porengröße von 0,45 µm gefiltert und anschließend über 5 ml einer 30 %igen Glycerinlösung in PBS⁻ geschichtet. Die viralen Partikel wurden in einem SW 32 Ti-Rotor für 200 Minuten bei 27500 rpm und 4 °C pelletiert. Die Pellets wurden zur Analyse im Immunoblot (4.8.1.) direkt mit 3-fach SDS-Probenpuffer (NEB) versetzt und bis zur weiteren Verwendung bei -80 °C gelagert.

4.9.1. Isolierung von RNA und Protein aus retroviralen Partikeln

Kit: RNase ONETM (Promega)
 QIAamp viral RNA mini kit (Qiagen)
 Turbo DNA-free (Ambion Inc.)
Ladepuffer: 3-fach SDS Sample Puffer (NEB)

Um virale RNA aus den Partikelpellets (4.9.) zu gewinnen wurden diese in PBS⁻ aufgenommen und um Kontaminationen mit zellulärer RNA zu vermeiden für fünfundvierzig Minuten bei 37 °C einem RNA-Verdau unterzogen (RNase One , Promega). Die Partikel wurden anschließend mit Lysepuffer (Qiagen) versetzt. Die virale RNA wurde mit Hilfe des QIAamp viral RNA mini kits (Qiagen) nach Angaben des Herstellers in 50 µl H₂O eluiert und bei -80 °C gelagert oder direkt nach einem DNase-Verdau (Turbo-DNA-Kit, Ambion Inc.) für die cDNA-Synthese eingesetzt. 40 µl Partikel-RNA wurden mit 2 Units DNase versetzt und dreißig Minuten bei 37 °C inkubiert, anschließend wurden zwei weitere Units Enzym zugeben und weitere dreißig Minuten bei 37 °C inkubiert. Die Reaktion wurde durch Zugabe von 10 µl DNase-removal gestoppt. In die cDNA-Synthese (4.4.2.5.) wurden 10 µl aus dem DNase Verdau eingesetzt und 25 µM random hexamer Primer verwendet. Die gewünschten HERV-K Gene wurden in der PCR unter Verwendung unten aufgeführter Primer amplifiziert. Die PCR-Produkte in den pGEMT-Easy-Vektor (Promega) nach Herstellerangaben kloniert (4.4.3.5.) und sequenziert. Die Sequenzierung wurde von der Firma Eurofins MWG Operon

durchgeführt. Sequenzierergebnisse wurden durch Datenbankvergleiche ausgewertet (4.10).

Primer (Ruprecht <i>et al.</i> , 2008)	HERV-K gag +, HERV-K gag - HERV-K env 8146 + HERV-K env 8665
Primer (Laderoute <i>et al.</i> , 2007)	pol FW + pol RV

4.9.2. Elektronenmikroskopie retroviraler Partikel

Mikroskop:	Zeiss EM 109
Glutaraledhyd:	25 % Solution Grade I (Sigmar-Adrich)
Osmium Tetroxid:	1 % Osmiumtetroxid (Sigma-Aldrich) in PBS
Tannin-Lösung:	0,1 % Tannin (Roth) in 20 mM HEPES
Uranylacetat-Lösung:	2 % Uranylacetat-Dihydrat (Merck) in Wasser
PBS:	Phosphat-gepufferte Salzlösung ohne Ca^{2+} , Mg^{2+} (Biochrom KG, Berlin)
Natriumsulfatlösung:	1 % (Roth) in 0,5 M Hepes
Zellkulturschale:	10 cm –Platten (Nunc)

Zellen wurden bis zur Konfluenz auf 10 cm Petrischalen (Nunc) gehalten und mit 2,5 % Glutaraldehyd im Nährmedium (!) für vierzig Minuten fixiert. Die fixierten Zellen wurden zwei Mal mit PBS- gewaschen und mit einem Zellschaber vom Untergrund gelöst. Anschließend wurden die Zellen bei 3000 rpm für zehn Minuten pelletiert, der Überstand wurde verworfen. Das Zellpellet wurde zu gleichen Anteilen mit 3 %iger flüssiger Agarose gemischt und zehn Minuten auf Eis inkubiert. Der abgekühlte Agarosekegel wurde in 1-2 mm³ große Würfel zerteilt. Die Agarosewürfel wurden in einer 1 %igen Osmium Tetroxid-Lösung eine Stunde nachfixiert. Nach dem Waschen mit PBS⁻ erfolgte eine weitere einstündige Inkubation in einer 0,1%igen Tannin-Lösung zur Darstellung viraler Glykoproteine. Nach erneutem Waschen mit einer Natriumsulfatlösung und destilliertem Wasser folgte eine Inkubation für eine Stunde in 2 %iger Uranylacetat-Lösung. Nachdem überschüssiges Uranylacetat durch Waschen mit destilliertem Wasser entfernt wurde, folgte eine aufsteigende Alkoholreihe (30 % - 96 %). Der Alkohol wurde anschließend durch Epoxidharz (Fluka) über Nacht ausgetauscht. Das Aushärten des Harzes erfolgte in Gelatine kapseln über einen Zeitraum von 48 Stunden bei 60 °C. Aus den eingebetteten

Proben wurden Dünnschnitte hergestellt, die vor der Elektronenmikroskopie mit 2 % Uranylacetat und 1 % Bleiacetat kontrastiert wurden.

4.10. Sequenzanalysen und Datenbankrecherchen

Zur Sequenz- und Expressionsanalyse sowie als Informationsquelle wurden die folgenden Internetangebote verwendet:

National Center for Biotechnology Information (NCBI)	http://ncbi.nlm.nih.gov
Human Genome Browser	http://genome.ucsc.edu
ClustalW	http://www.ebi.ac.uk/clustalw/
ExPASy Molekular Biology Server	http://www.expasy.org/
TRANSFAC 7.0 (2005)	http://www.gene-regulation.com
MatInspector	http://www.genomatix.de

5. ZUSAMMENFASSUNG

Humane endogene Retroviren (HERVs) sind Relikte von Infektionen der Keimzellen von Primatenvorläufern mit exogenen Retroviren vor 40 Millionen Jahren. Die meisten HERV Elemente sind defekt durch die Akkumulation von Mutationen, Deletionen und Rearrangements. Lediglich einige Proviren der HERV-K/HML-2 Familie enthalte noch offenen Leserahmen für alle retroviralen Proteine. Die Expression repetitiver endogener Elemente wird in somatischen Zellen unterdrückt. Transkription und Translation von HERV-K Genprodukten tritt in Keimzelltumoren, Melanomen und Eierstockkrebs auf.

In der vorliegenden Arbeit wurde die HERV-K exprimierende Melanomlinie UKRV Mel 2-C9 etabliert und charakterisiert. Es konnte gezeigt werden, dass die HERV-K Expressionen durch die dominante Aktivität des Provirus 108 bedingt ist. Die Expression und korrekte Prozessierung der viralen Proteine konnte nachgewiesen werden. Durch Elektronenmikroskopie wurde bestätigt, dass HERV-K 108 virale Partikel mit charakteristischen Strukturen der Hüllproteine produziert. Diese Partikel enthalten virale Vollängen RNA-Genome. Da HERV-K 108 in der Linie UKRV Mel 2-C9 eine Mutation im aktiven Zentrum der Reversen Transkriptase zeigt, sind die Partikel nicht infektiös.

Um Einblicke in die der HERV-K Expression in Keimzelltumoren und Melanomen zu Grunde liegenden transkriptionellen Regulationsmechanismen zu erhalten, wurden mittels 5' und 3' RACE (rapid amplification of cDNA ends) der Transkriptionsstart und die Terminationsstelle bestimmt. Es konnte gezeigt werden, dass der HERV-K Promotor in den untersuchten Zelllinien TATA-unabhängig agiert obwohl die HERV-K LTR (long terminal repeat) ein TATA-Motiv aufweist. Die transkriptionelle Aktivität des HERV-K Promotors ist abhängig von einer GC-Box in unmittelbarer Nähe zu einem Initiatorelement am Startpunkt der Transkription. Das GC-Sequenzmotiv ist die Bindestelle für die Transkriptionsfaktoren SP1 (specific protein) und SP3. Mutationen des Bindemotivs führten zu einem drastischen Abfall der Promotoraktivität im Luziferaseassay. Durch Proteinknockdown mittels spezifischer siRNA gegen SP1 und SP3 konnte die Promotoraktivität ebenso deutlich reduziert werden. Elektrophoretic mobility shift assays und Chromatinimmunopräzipitationen bestätigten eine Bindung von SP1 und SP3 an die HERV-K LTR. Die Transkriptionsfaktoren SP1 und SP3 sind an der Initiation der basalen Transkription von HERV-K maßgeblich beteiligt.

6. LITERATURVERZEICHNIS

Alarcon G. S., (2000). Methotrexate use in rheumatoid arthritis. A Clinician's perspective. *Immunopharmacology*; 47: 259-271.

Alonso C. R., Pesce C. G., Kornblihtt A. R., (1996). The CCAAT-binding proteins CP1 and NF-I cooperate with ATF-2 in the transcription of the fibronectin gene. *J Biol Chem.*; 271(36): 22271-9.

Andrews P. W., Bronson D. L., Benham F., Strickland S., Knowles B. B., (1980). A comparative study of eight cell lines derived from human testicular teratocarcinoma. *International Journal of Cancer*; 26: 269-280.

Al-Harathi L., Spear G. T., Hashemi F. B., Landay A., Sha B. E., Roebuck K. A., (1998). A human immunodeficiency virus (HIV)-inducing factor from the female genital tract activates HIV-1 gene expression through the kappaB enhancer. *J Infect Dis.*; 178(5): 1343-51.

Armbruester V., Sauter M., Krautkraemer E., Meese E., Kleiman A., Best B., Roemer K., Mueller-Lantsch N., (2002). A novel gene from the human endogenous retrovirus K expressed in transformed cells. *Clin Cancer Res.*; 8(6): 1800-7.

Armbruester V., Sauter M., Roemer K., Best B., Hahn S., Nty A., Schmid A., Philipp S., Mueller A., Mueller-Lantsch N., (2004). Np9 protein of human endogenous retrovirus K interacts with ligand of numb protein X. *J Virol.*; 78(19): 10310-9.

Athanikar J. N., Sanchez, H. B., Osborne, T. F., (1997). Promoter selective transcriptional synergy mediated by sterol regulatory element binding protein and Sp1: a critical role for the Btd domain of Sp1. *Mol. Cell Biol.*; 17: 5193-5200.

Baltimore D., (1970). RNA-dépendant DNA Polymerase in Virions of RNA tumor viruses. *Nature*; 226: 1209-1211.

Barbulescu M., Turner G., Seaman S.I., et al., (1999). Many human endogenous retrovirus K (HERV-K) proviruses are unique to humans. *Curr Biol.*; 9(16): 861-868.

Ben-Hattar J. und **Jiricny J.**, (1988). Methylation of single CpG dinucleotides within a promoter element of the Herpes simplex virus tk gene reduces its transcription in vivo. *Gene*; 65(2): 219-27.

Ben-Hattar J., Beard P., Jiricny J., (1989). Cytosine methylation in CTF and Sp1 recognition sites of an HSV tk promoter: effects on transcription in vivo and on factor binding in vitro. *Nucleic Acids Res.* ; 17(24): 10179-90.

Bernhard W., (1960). The detection and study of tumor viruses with the electron microscope. *Cancer Res.*; 18: 491-509.

Bieda K., Hoffmann A., Boller K., (2001). Phenotypic heterogeneity of human endogenous retrovirus particles produced by teratocarcinoma cell lines. *J Gen Virol.*; 82(Pt 3): 591-6.

Bird A.P., (1986). CpG-rich islands and the function of DNA methylation. *Nature*; 321(6067): 209-13.

Blaise S., de Parseval N., Bénit 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(22): 13013-8.

Blomberg J., Ushameckis D., Jern P., (2005). Evolutionary aspects of human endogenous retroviral sequences (HERVs) and disease. *Retroviruses and Primate Genome Evolution 204-238*, Svedlov, E.D. (ed.), Eurekah.com

Blond J.L., Lavillette D., Cheynet V., Bouton O., Oriol G., Chapel-Fernandes S., Mandrand B., Mallet F., Cosset F. L., (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(7): 3321-9.

- Boese A.**, Sauter M., Mueller-Lantzsch N., (2000). A rev-like NES mediates cytoplasmic localization of HERV-K cORF. *FEBS Lett.*; 468(1): 65-7.
- Boller K.**, Frank H., Löwer J., Löwer R., Kurth R., (1983). Structural organization of unique retrovirus-like particles budding from human teratocarcinoma cell lines. *J. Gen Virol.*; 64 (Pt 12): 2549-2559.
- Boller K.**, König H., Sauter M., (1993). Evidence that HERV-K is the endogenous retrovirus sequence that codes for the human teratocarcinoma-derived retrovirus HTDV. *Virology*; 196(1): 349-353.
- Boller K.**, Janssen O., Schuldes H., Tonjes R. R., Kurth R., (1997). Characterization of the antibody response specific for the human endogenous retrovirus HTDV/HERV-K. *J. Virol.* 71: 4581–4588.
- Boller K.**, Schönfeld K., Lischer S., Fischer N., Hoffmann A., Kurth R., Tönjes R. R., (2008). Human endogenous retrovirus HERV-K113 is capable of producing intact viral particles. *J. Gen Virol.*; 89(Pt 2): 567-72.
- Bouwman P.**, Göllner H., Elsässer H. P., Eckhoff G., Karis A., Grosveld F., Philipsen S., Suske G., (2000). Transcription factor Sp3 is essential for post-natal survival and late tooth development. *EMBO J* ; 19: 655-661.
- Bouwman P.** und **Philipsen S.**, (2002). Regulation of the activity of Sp1-related transcription factors. *Mol Cell Endocrinol.*; 195: 27-38.
- Büscher K.**, Hahn S., Hofmann M., Trefzer U., Ozel M., Sterry W., Löwer J., Löwer R., Kurth R., Denner J., (2006). Expression of the human endogenous retrovirus-K transmembrane envelope, Rec and Np9 proteins in melanomas and melanoma cell lines. *Melanoma Res.*; 16(3): 223-34.
- Burley S. K.** und **Roeder R. G.**, (1996). Biochemistry and structural biology of transcription factor IID (TFIID). *Annu. Rev. Biochem.*; 65: 769–799.
- Campanero M. R.**, Armstrong M. I., Flemington E. K., (2000). CpG methylation as a mechanism for the regulation of E2F activity. *Proc Natl Acad Sci U S A.*; 97(12): 6481-6.
- Chamberlin M.** und **Berg P.**, (1962). Deoxyribonucleic acid-directed synthesis of ribonucleic acid by an enzym from *Escherichia coli*. *Proc Natl Acad Sci USA*; 48: 81-94.
- Chiu Y. L.**, Greene W. C., (2008). The APOBEC3 cytidine deaminases: an innate defensive network opposing exogenous retroviruses and endogenous retroelements. *Annu Rev Immunol.*; 26: 317-53.
- Choi G.**, Park S., Choi B., Hong S., Lee J., Hunter E., Rhee S. S., (1999). Identification of a cytoplasmic targeting/retention signal in a retroviral Gag polyprotein. *J Virol.*; 73(7): 5431-7.
- Conrad B.**, Weissmahr R. N., Boni J., Arcari R., Schupbach J., Match B., (1997). A human endogenous retroviral superantigen as candidate autoimmune gene in type I diabetes. *Cell*; (2): 303-13.
- Courey A. J.** und **Tjian R.**, (1988). Analysis of Sp1 in vivo reveals multiple transcriptional domains, including a novel glutaminerich activation motif. *Cell*; 55: 887–898.
- Courey A. J.**, Holtzman D. A., Jackson S. P., Tjian R., (1989). Synergistic activation by the glutamine-rich domains of human transcription factor Sp1. *Cell*; 59: 827–836.
- Crick F. H. C.**, (1958). On Protein Synthesis. *Symp Soc Exp Biol.*; 12: 138-163.
- Cullen B. R.**, Lomedico P. T., Ju G., (1984). Transcriptional interference in avian retroviruses--implications for the promoter insertion model of leukaemogenesis. *Nature*; 25;307(5948): 241-5.
- de Martini, J. C.**, Carlson J. O., Leroux C., Spencer T., Palmarini M., (2003). Endogenous retroviruses related to jaagsiekte sheep retrovirus. *Curr Top Microbiol Immunol.*; 275: 117-137.
- de Parseval N.** und **Heidmann T.**, (2005). Human endogenous retroviruses: from infectious elements to human genes. *Cytogenet Genome Res.*; 110(1-4): 318-32.

- Deng W.** und **Roberts S. G. E.**, (2005). A core promoter element downstream of the TATA box that is recognized by TFIIB. *Genes Dev.*; 19: 2418–2423.
- Denne M.**, Sauter M., Armbruster V., Licht J. D., Roemer K., Mueller-Lantzsch N., (2007). Physical and functional interactions of human endogenous retrovirus proteins Np9 and rec with the promyelocytic leukemia zinc finger protein. *J. Virol.*; 81(11): 5607-16.
- Dennig J.**, Beato M., Suske G., (1996). An inhibitor domain in Sp3 regulates its glutamine-rich activation domains. *EMBO J.*; 15: 5659–5667.
- Dewannieux M.**, Blaise S, Heidmann T., (2005). Identification of a functional envelope protein from the HERV-K family of human endogenous retroviruses. *J Virol.*; 79(24): 15573-7.
- Dewannieux M.**, Harper F., Richaud A., Letzelter C., Ribet D., Pierron G., Heidmann T., (2006). Identification of an infectious progenitor for the multiple-copy HERV-K human endogenous retroelements. *Genome Res.*; 16: 1548-1556.
- Doolittle R. F.**, Feng D. F., McClure M. A., Johnso M. S., (1990). Retrovirus phylogeny and evolution. *Curr.Top. Microbiol. Immunol.*; 157: 1-18.
- Duelli D.** und **Lazebnik Y.**, (2003). Cell fusion: a hidden enemy? *Cancer Cell.*; 3(5): 445-8.
- Dynan W. S.** und **Tjian R.**, (1983a). Isolation of transcription factors that discriminate between different promoters recognized by RNA polymerase II. *Cell*; 35(1): 79-87.
- Dynan W. S.** und **Tjian R.**, (1983b). Isolation of transcription factors that discriminate between different promoters recognized by RNA polymerase II. *Cell*; 32(3): 669-80.
- Faisst S.** und **Meyer S.**, (1992). Compilation of vertebrate-encoded transcription factors. *Nucleic Acids Res.*; 20(1): 3-26.
- Felicetti F.**, Errico M. C., Bottero L., Segnalini P., Stoppacciaro A., Biffoni M., Felli N., Mattia G., Petrini M., Colombo M. P., Peschle C., Carè A., (2008). The promyelocytic leukemia zinc finger-microRNA-221/-222 pathway controls melanoma progression through multiple oncogenic mechanisms. *Cancer Res.*; 68(8):2745-54.
- Fine D.** und **Schochetman G.**, (1978). Type D primate retroviruses: A review. *Cancer Res.*; 38: 3123-3139.
- Fleischer B.**, (1991). Stimulation of human T cells by microbial 'superantigens'. *Immunol Res*, 10, 349-55.
- Frendo J. L.**, Olivier D., Cheynet V., Blond J. L., Bouton O., Vidaud M., Rabreau M., Evain-Brion D., Mallet F., (2003). Direct involvement of HERV-W Env glycoprotein in human trophoblast cell fusion and differentiation. *Mol Cell Biol.*; 23, 3566-3574.
- Furuichi Y.** und **Shatkin A. J.**, (2000). Viral and cellular mRNA capping: past and prospects. *Adv Virus Res.*; 55: 135-84.
- Gardiner-Garden M.** und **Frommer M.**, (1987). CpG islands in vertebrate genomes. *J Mol Biol.*; 196(2): 261-82.
- Gaynor R.** (1992). Cellular transcription factors involved in the regulation of HIV-1 gene expression. *AIDS*; 6(4): 347-63.
- Geng Y.** und **Johnson L. F.**, (1993). Lack of an initiator element is responsible for multiple transcriptional initiation sites of the TATA-less mouse thymidylate synthase promoter. *Mol Cell Biol.*; 13(8): 4894-903.
- Gniazdowski M.**, Mandel J. L. Jr., Gissinger F., Kedinger C., Chambon P., (1970). Calf thymus RNA polymerases exhibit template specificity. *Biochem Biophys Res Commun.*; 38: 1033–1040.
- Goeddert J. J.**, Sauter M. E., Jacobson L. P., Vessella R. L., Hilgartner M. W., Leitmann S. F., Fraser M. C., Mueller-Lantzsch N. G., (1999). High prevalence of antibodies against HERV-K10 in patients with testicular cancer but not with AIDS. *Cancer Epidemiol Biomarkers Prev.*; 8: 293-296.

- Göllner H.**, Dani C., Phillips B., Philipsen S., Suske G., (2001). Impaired ossification in mice lacking the transcription factor Sp3. *Mech Dev.*; 106, 77-83.
- Golovkina T. V.**, Chervonsky A., Dudley J. P., Ross S. R., (1992). Transgenic mouse mammary tumor virus superantigen expression prevents viral infection. *Cell*; 69, 637-645.
- Goff S. P.**, (2004). Retrovirus restriction factors. *Mol Cell*.; 16(6): 849-59.
- Greenblatt J.**, (1997). RNA polymerase II holoenzyme and transcriptional regulation. *Curr. Opin. Cell Biol.*; 9: 310–319.
- Gröne M.**, Hoffmann E., Berchtold S., Cullen B. R., Grassmann R., (1994). A single stem-loop structure within the HTLV-1 Rex response element is sufficient to mediate Rex activity in vivo. *Virology*; 204(1): 144-52.
- Guntaka R. V.**, (1993). Transcription termination and polyadenylation in retroviruses. *Microbiol Rev.*; 57(3): 511-21.
- Hahn S.**, Buratowski S., Sharp P. A., Guarente L., (1989). Yeast TATA-binding protein TFIID binds to TATA elements with both consensus and nonconsensus sequences. *Proc. Natl. Acad. Sci. USA*; 86: 5718–5722.
- Hahn S.**, Ugurel S., Hanschmann K. M., Strobel H., Tondera C., Schadendorf D., Löwer J., Löwer R., (2008). Serological response to human endogenous retrovirus K in melanoma patients correlates with survival probability. *AIDS Res Hum Retroviruses*; 24(5): 717-23.
- Harrington M. A.**, Jones P. A., Imagawa M., Karin M., (1988). Cytosine methylation does not affect binding of transcription factor Sp1. *Proc Natl Acad Sci USA*; 85(7): 2066-70.
- Harris J. M.**, Haynes R. H., McIntosh E. M., (1997). A consensus sequence for a functional human endogenous retrovirus K (HERV-K) dUTPase. *Biochem Cell Biol.*; 75(2): 143-51.
- Harrison S. M.**, Houzelstein D., Dunwoodie S. L., Beddington, R. S., (2000). Sp5, a new member of the Sp1 family, is dynamically expressed during development and genetically interacts with Brachyury. *Dev. Biol.*; 227: 358–372.
- Herbst H.**, Sauter M., Mueller-Lantzsch N., (1996). Expression of human endogenous retrovirus K elements in germ cell and trophoblastic tumors. *Am J Pathol.*; 149(5): 1727-35.
- Herbst H.**, Sauter M., Kuhler-Obbarius C., Loning T., Mueller- Lantzsch N., (1998). Human endogenous retrovirus (HERV)-K transcripts in germ cell and trophoblastic tumours. *APMIS.*: 106: 216–220.
- Hernandez N.**, (1993). TBP, a universal eukaryotic transcription factor? *Genes Dev.*, 7: 1291–1308.
- Herr A. J.**, Jensen M. B., Dalmay T., Baulcombe D. C., (2005). RNA polymerase IV directs silencing of endogenous DNA. *Science.*; 308: 118–120.
- Höller M.**, Westin G., Jiricny J., Schaffner W., (1988). Sp1 transcription factor binds DNA and activates transcription even when the binding site is CpG methylated. *Genes Dev.*; 2(9): 1127-35.
- Holmes E. C.**, (2007). Ancient lentiviruses leave their mark. *Proc Natl Acad Sci USA*; 104(15): 6095-6.
- Holzmeister J.**, Ludewig B., Pauli G., Simon D., (1993). Sequence specific binding of the transcription factor c-Ets1 to the human immunodeficiency virus type I long terminal repeat. *Biochem Biophys Res Commun.*; 197(3): 1229-33.
- Hsieh C. L.**, (1994). Dependence of transcriptional repression on CpG methylation density. *Mol Cell Biol.*; 14(8): 5487-94.
- Hurwitz J.**, Bresler A., Diring R., (1960). The enzymic incorporation of ribonucleotides into polyribonucleotides and the effect of DNA. *Biochem Biophys Res Commun.*; 3: 15-18.

- Jones K. A., Kadonaga J. T., Luciw P. A., Tjian R., (1986).** Activation of the AIDS retrovirus promoter by the cellular transcription factor, Sp1. *Science*; 232(4751): 755-9.
- Jones K. A. und Peterlin B. M., (1994).** Control of RNA initiation and elongation at the HIV-1 promoter. *Annu Rev Biochem.*; 63: 717-43.
- Kadonaga J. T., Courey A. J., Ladika J., Tjian R., (1988).** Distinct regions of Sp1 modulate DNA binding and transcriptional activation. *Science*; 242: 1566–1570.
- Kanno, T., Huettel, B., Mette, M.F., Aufsatz, W., Jaligot, E., Daxinger, L., Kreil, D.P., Matzke, M., and Matzke, A.J.M. 2005.** Atypical RNA polymerase subunits required for RNA-directed DNA methylation. *Nature Genetics* 37:761–765.
- Kass S.U., Pruss D., Wolffe A.P. (1997).** How does DNA methylation repress transcription? *Trends Genet.*; 13(11): 444-9.
- Kedinger C., Gniazdowski M., Mandel J. L. Jr, Gissinger F., Chambon P., (1970).** Alpha-amanitin: a specific inhibitor of one of two DNA-pendent RNA polymerase activities from calf thymus. *Biochem Biophys Res Commun.*; 38: 165–171.
- Kim Y. J., Bjorklund S., Li Y., Sayre M. H., Kornberg R. D., (1994).** A multiprotein mediator of transcriptional activation and its interaction with the C-terminal repeat domain of RNA polymerase II. *Cell*; 77: 599–608.
- Knössl M., Löwer R., Löwer J., (1999).** Expression of the human endogenous retrovirus HTDV/HERV-K is enhanced by cellular transcription factor YY1. *J Virol.*; 73(2): 1254-61.
- Kobayashi K., Nakahori Y., Miyake M., Matsumura K., Kondo-Iida E., Nomura Y., Segawa M., Yoshioka M., Saito K., Osawa M., (1998).** An ancient retrotransposal insertion causes Fukuyama-type congenital muscular dystrophy. *Nature*; 394: 388-392.
- Koleske A. J. und Young R. A., (1994).** An RNA polymerase II holoenzyme responsive to activators. *Nature*; 368: 466–469.
- Kovalskaya E., Buzdin A., Gogvadze E., Vinogradova T., Sverdlov E., (2006).** Functional human endogenous retroviral LTR transcription start sites are located between the R and U5 regions. *Virology*; 346(2): 373-8.
- Kovesdi I., Reichel R., Nevins J. R., (1987).** Role of an adenovirus E2 promoter binding factor in E1A-mediated coordinate gene control. *Proc Natl Acad Sci USA.*; 84(8): 2180-4.
- Kutach A. und Kadonaga J. T., (2000).** The downstream promoter element DPE appears to be as widely used as the TATA box in Drosophila core promoters. *Mol Cell Biol.*; 20: 4754–4764.
- Laderoute M. P., Giulivi A., Larocque L., Bellfroy D., Hou Y., Wu H. X., Fowke K., Wu J., Diaz-Mitoma F., (2007).** The replicative activity of human endogenous retrovirus K102 (HERV-K102) with HIV viremia. *AIDS*; 21(18): 2417-24.
- La Mantia G., Maglione D., Pengue G., Di Cristofano A., Simeone A., Lanfrancone L., Lania L., (1991).** Identification and characterization of novel human endogenous retroviral sequences preferentially expressed in undifferentiated embryonal carcinoma cells. *Nucleic Acids Res.*; 19(7): 1513-20.
- La Mantia G., Majello B., Di Cristofano A., Strazzullo M., Minchiotti G., Lania L., (1992).** Identification of regulatory elements within the minimal promoter region of the human endogenous ERV9 proviruses: accurate transcription initiation is controlled by an Inr-like element. *Nucleic Acids Res.*; 20(16): 4129-36.
- Lania L, Di Cristofano A, Strazzullo M, Pengue G, Majello B, La Mantia G. (1992)** Structural and functional organization of the human endogenous retroviral ERV9 sequences. *Virology.*; 191(1): 464-8.

- Lander** E. S., Linton L. M., Birren B., Nusbaum C., Zody M. C., Baldwin J., Devon K., Dewar K., Doyle M., FitzHugh W., Funke R., Gage D., Harris K., Heaford A., Howland J., Kann L., Lehoczký J., LeVine R., McEwan P., Mc Kernan K., Meldrim J., Mesirov J. P., Miranda C., Morris W., Naylor J., Raymond C., Rosetti M., Santos R., Sheridan A., Sougnez C., Stange-Thomann N., Stojanovic N., Subramanian, Wyman D., Rogers J., Sulston J., Ainscough R., Beck S., Bentley D., Burton J., Clee C., Carter N., Coulson, Deadman R., Deloukas P., Dunham A., Dunham I., Durbin R., French L., Grafham D., Gregory S., Hubbard T., Humphray S., Hunt A., Jones M., Lloyd C., McMurray A., Matthews L., Mercer S., Milne S., Mullikin J. C., Mungall A., Plumb R., Ross M., Shownkeen R., Sims S., Waterston R. H., Wilson R. K., Hillier L. W., McPherson J. D., Marra M. A., Mardis E. R., Fulton L. A., Chinwalla A. T., Pepin K. H., Gish W. R., Chissoe S. L., Wendl M. C., Delehaunty K. D., Miner T. L., Delehaunty A., Kramer J. B., Cook L. L., Fulton R. S., Johnson D. L., Minx P. J., Clifton S. W., Hawkins T., Branscomb E., Predki P., Richardson P., Wenning S., Slezak T., Doggett N., Cheng J. F., Olsen A., Lucas S., Elkin C., Uberbacher E., Frazier M. et al. (2001). Initial sequencing and analysis of the human genome. *Nature*; 409: 860-921.
- Lagrange** T., Kapanidis A. N., Tang H., Reinberg D., Ebright R. H., (1998). New core promoter element in RNA polymerase II-dependent transcription: sequence-specific DNA binding by transcription factor IIB. *Genes Dev.*; 12: 34-44.
- Larsen** F., Gundersen G., Prydz H., (1992). Choice of enzymes for mapping based on CpG islands in the human genome. *Genet Anal Tech Appl.*; 9(3): 80-5.
- Lavie** L., Kitova M., Maldener E., Meese E., Mayer J., (2005). CpG methylation directly regulates transcriptional activity of the human endogenous retrovirus family HERV-K(HML-2). *J Virol.*; 79(2): 876-83.
- Lee** W. J., Kwun H. J., Jang K. L., (2003). Analysis of transcriptional regulatory sequences in the human endogenous retrovirus W long terminal repeat. *J Gen Virol.*; 84(Pt 8): 2229-35.
- Lee** D.-H., Gershenson N., Gupta M., Ioshikhes I. P., Reinberg D., Lewis B. A., (2005). Functional characterization of core promoter elements: the downstream core element is recognized by TAF1. *Mol Cell Biol.*; 25: 9674-9686.
- Lee** Y. N. und **Bieniasz** P. D., (2007). Reconstitution of an infectious human endogenous retrovirus. *PLoS Pathog.*; 3(1): e10.
- Lee** Y. N., Malim M. H., Bieniasz P. D., (2008). Hypermutation of an ancient human retrovirus by APOBEC3G. *J Virol.*; 82(17): 8762-70.
- Levy** J., (1992). *The Retroviridae*. Plenum Press New York
- Li** L., He S., Sun J. M., Davie J. R., (2004). Gene regulation by Sp1 and Sp3. *Biochem Cell Biol.*; 82(4): 460-71.
- Lim** C. Y., Santoso B., Boulay T., Dong E., Ohler U., Kadonaga, J. T., (2004). The MTE, a new core promoter element for transcription by RNA polymerase II. *Genes Dev.*; 18: 1606-1617.
- Limjoco** T. I., Dickie P., Ikeda H., Silver J., (1993). Transgenic Fv-4 mice resistant to Friend virus. *J Virol.*; 67(7): 4163-8.
- Löwer** R., Löwer J., Frank H., Harzmann R., Kurth R., (1984). Human teratocarcinomas cultured in vitro produce unique retrovirus-like viruses. *J Gen Virol.*; 65 (5): 887-898.
- Löwer** R., Boller K., Hasenmayer B., Korbmacher C., Mueller-Lantzsch N., Loewer J., Kurth R., (1993a). Identification of human endogenous retrovirus with complex mRNA expression and particle formation. *Proc Natl Acad Sci USA*; 90 (10): 4480-4.
- Löwer** R., Tonjes R. R., Korbmacher C., et al., (1995). Identification of a Rev-related protein by analysis of spliced transcripts of the human endogenous retroviruses HTDV/HERV-K. *J Virol.*; 69(1): 141-149.
- Löwer** R., Lower J., Kurth R., (1996). The viruses in all of us: characteristics and biological significance of human endogenous retrovirus sequences. *Proc Natl Acad Sci USA*; 93(11): 5177-5184.

- Macfarlane C.** und **Simmonds P.**, (2004). Allelic variation of HERV-K(HML-2) endogenous retroviral elements in human populations. *J Mol Evol.*; 59(5): 642-56.
- Malim M. H.**, **Tiley L. S.**, **McCarn D. F.**, **Rusche J. R.**, **Hauber J.**, **Cullen B. R.**, (1990). HIV-1 structural gene expression requires binding of the Rev trans-activator to its RNA target sequence. *Cell*; 60(4): 675-83.
- Magin C.**, **Lower R.**, **Lower J.**, (1999). cORF and RcRE, the Rev/Rex and RRE/RxRE homologues of the human endogenous retrovirus family HTDV/HERV-K. *J Virol.*; 73(11): 9496-9507.
- Magin-Lachmann C.**, **Hahn S.**, **Strobel H.**, **Held U.**, **Löwer J.**, **Löwer R.**, (2001). Rec (formerly Corf) function requires interaction with a complex, folded RNA structure within its responsive element rather than binding to a discrete specific binding site. *J Virol.*; 75(21): 10359-71.
- Majello B.**, **De Luca P.**, **Lania L.**, (1997). Sp3 is a bifunctional transcription regulator with modular independent activation and repression domains. *J. Biol. Chem.*; 272: 4021-4026.
- Mangeny M.**, **Renard M.**, **Schlecht-Louf G.**, **Bouallaga I.**, **Heidmann O.**, **Letzelter C.**, **Richaud A.**, **Ducos B.**, **Heidmann T.**, (2007). Placental syncytins: Genetic disjunction between the fusogenic and immunosuppressive activity of retroviral envelope proteins. *Proc Natl Acad Sci USA*; 104(51): 20534-9.
- Mantovani R.**, (1999). The molecular biology of the CCAAT-binding factor NF-Y. *Gene*; 239(1): 15-27.
- Mallet F.**, **Bouton O.**, **Prudhomme S.**, **Cheyne V.**, **Oriol G.**, **Bonnaud B.**, **Lucotte G.**, **Duret L.**, **Mandrand B.**, (2004). The endogenous retroviral locus ERVWE1 is a bona fide gene involved in hominoid placental physiology. *Proc Natl Acad Sci USA*; 101(6): 1731-6.
- Mayer J.**, **Sauter M.**, **Rácz A.**, **Scherer D.**, **Mueller-Lantzsch N.**, **Meese E.**, (1999). An almost-intact human endogenous retrovirus K on human chromosome 7. *Nat Genet.*; 21(3): 257-8.
- Mayer J.** und **Meese E. U.**, (2003). Presence of dUTPase in the various human endogenous retrovirus K (HERV-K) families. *J Mol Evol.*; 57(6): 642-9.
- Mayer J.**, **Stuhr T.**, **Reus K.**, **Maldener E.**, **Kitova M.**, **Asmus F.**, **Meese E.**, (2005). Haplotype analysis of the human endogenous retrovirus locus HERV-K(HML-2.HOM) and its evolutionary implications. *J Mol Evol.*; 61(5): 706-15.
- Marin M.**, **Karis A.**, **Visser P.**, **Grosveld F.**, **Philipsen S.**, (1997). Transcription factor Sp1 is essential for early development but dispensable for cell growth and differentiation. *Cell*; 89: 619-628.
- Medstrand P.** and **Blomberg 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.
- Medstrand P.** and **Mager D.L.**, (1998). Human-specific integrations of the HERV-K endogenous retrovirus family. *J Virol.*; 72(12): 9782-9787.
- Meylan, F.**, **de Smedt, M.**, **Leclercq, G.**, **Plum, J.**, **Leupin, O.**, **Marguerat, S.**, **Conrad, B.** (2005) Negative thymocyte selection to HERV-K18 superantigens in humans. *Blood*, 105, 4377-82.
- Modrow S.**, **Falke D.**, (1997). *Molekulare Virology. Spektrum Lehrbuch*, Springer Verlag
- Mortensen E. R.**, **Marks P. A.**, **Shiotani A.**, **Merchant J. L.** (1997). Epidermal growth factor and okadaic acid stimulate Sp1 proteolysis. *J. Biol. Chem.*; 272: 16 540 – 16 547.
- Mueller-Lantzsch N.**, **Sauter M.**, **Weiskircher A.**, et al. (1993). Human endogenous retroviral element K10 (HERV-K10) encodes a full-length gag homologous 73-kDa protein and a functional protease. *AIDS Res Hum Retroviruses.*; 9(4): 343-350.
- Muster T.**, **Waltenberger A.**, **Grassauer A.**, et al., (2003). An endogenous retrovirus derived from human melanoma cells. *Cancer Res.*; 63(24): 8735-8741.

- Nabel G.**, Baltimore D., (1987). An inducible transcription factor activates expression of human immunodeficiency virus in T cells. *Nature*; 326(6114): 711-3.
- Nakagawa K.** und **Harrison L. C.**, (1996). The potential roles of endogenous retroviruses in autoimmunity. *Immunol Rev.*; 152: 193-236.
- Nan X.**, Campoy F.J., Bird A., (1997). MeCP2 is a transcriptional repressor with abundant binding sites in genomic chromatin. *Cell*; 88(4): 471-81.
- Nelson P. N.**, Carnegie P. R., Martin J., Davari Ejtehadi H., Hooley P., Roden D., Rowland-Jones S., Warren P., Astley J., Murray P. G., (2003). Demystified...Human endogenous retroviruses. *J Clin Pathol: Mol Pathol.*; 56: 11-18.
- Nelson P. N.**, Hooley P., Roden D., Davari Ejtehadi H., Rylance P., Warren P., Martin J., Murray P. G., (2004). Human endogenous retroviruses: transposable elements with potential? *Clin Exp Immunol.*; 138: 1-9.
- Nevins J. R.**, (1992). Transcriptional regulation. A closer look at E2F. *Nature*; 358(6385): 375-6.
- O'Connell C.**, O'Brien S., Nash W. G., Cohen M., (1984). ERV3, a full-length human endogenous provirus: chromosomal localization and evolutionary relationships. *Virology*; 138(2): 225-35.
- Ogasawara H.**, Naito T., Kaneko H., Hishikawa T., Sekigawa I., Hashimoto H., Kaneko Y., Yamamoto N., Maruyama N., Yamamoto N., (2001). Quantitative analyses of messenger RNA of human endogenous retrovirus in patients with systemic lupus erythematosus. *J Rheumatol.*; 28(3): 533-8.
- Ohler U.**, Liao G. C., Niemann H., Rubin G. M., (2002). Computational analysis of core promoters in the *Drosophila* genome. *Genome Biology.*; 3(12): research0087
- Okumura K.**, Sakaguchi G., Takagi S., Naito K., Mimori T., Igarashi H., (1996). Sp1 family proteins recognize the U5 repressive element of the long terminal repeat of human T cell leukemia virus type I through binding to the CACCC core motif. *J Biol Chem.*; 271(22): 12944-50.
- Ono M.**, (1986). Molecular cloning and long terminal repeat sequences of human endogenous retrovirus genes related to types A and B retrovirus genes. *J Virol.*; 58: 937-944.
- Onodera Y.**, Haag J. R., Ream T., Nunes P. C., Pontes O., Pikaard C. S., (2005). Plant nuclear RNA polymerase IV mediates siRNA and DNA methylation-dependent heterochromatin formation. *Cell*; 120: 613-622.
- Ostertag E. M.**, Goodier J. L., Zhang Y., Kazazian H. H., Jr., (2003). SVA elements are nonautonomous retrotransposons that cause disease in humans. *Am J Hum Genet.*; 73: 1444-1451.
- Orphanides G.**, Lagrange T., Reinberg D., (1996). The general transcription factors of RNA polymerase II. *Genes Dev.*; 10: 2657-2683.
- Pascal E.**, und **Tjian R.**, (1991). Different activation domains of Sp1 govern formation of multimers and mediate transcriptional synergism. *Genes Dev.*; 5: 1646-1656.
- Ponferrada V.G.**, Mauck B.S., Wooley D.P., (2003). The envelope glycoprotein of human endogenous retrovirus HERV-W induces cellular resistance to spleen necrosis virus. *Arch Virol.*; 148(4): 659-75.
- Posnett D. N.** und **Yarilina A. A.**, (2001). Sleeping with the enemy-endogenous superantigens in humans. *Immunity.*; 15(4): 503-506.
- Philipsen S.**, und **Suske G.**, (1999). A tale of three fingers: the family of mammalian Sp/XKLF transcription factors. *Nucleic Acids Res.*; 27: 2991-3000.
- Ptashne M.** und **Gann A.**, (1997). Transcriptional activation by recruitment. *Nature*; 386(6625): 569-77.
- Razin A.** und **Riggs A.D.**, (1980). DNA methylation and gene function. *Science*; 210(4470): 604-10.

- Repaske R.**, O'Neill R. R., Steele P. E., Martin M. A., (1983). Characterization and partial nucleotide sequence of endogenous type C retrovirus segments in human chromosomal DNA. *Proc Natl Acad Sci U S A* 80(3): 678-82.
- Reus K**, Mayer J, Sauter M, Scherer D, Müller-Lantzsch N, Meese E. (2001) Genomic organization of the human endogenous retrovirus HERV-K(HML-2.HOM) (ERVVK6) on chromosome 7. *Genomics.*; 72(3): 314-20.
- Robertson K. D.**, Hayward S. D., Ling P. D., Samid D., Ambinder R. F., (1995). Transcriptional activation of the Epstein-Barr virus latency C promoter after 5-azacytidine treatment: evidence that demethylation at a single CpG site is crucial. *Mol Cell Biol.*; 15(11): 6150-9.
- Roeder R. G. und Rutter W. J.**, (1969). Multiple forms of DNA-dependent RNA polymerase in eukaryotic organisms. *Nature*; 224: 234–237.
- Roeder R. G. und Rutter W.J.**, (1970). Specific nucleolar and nucleoplasmic RNA polymerases. *Proc Natl Acad Sci USA*; 65: 675–682.
- Roos M. D.**, Su K., Baker J. R., Kudlow J. E., (1997). O glycosylation of an Sp1-derived peptide blocks known Sp1 protein interactions. *Mol. Cell Biol.*; 17: 6472–6480.
- Ross S.**, Best J. L., Zon L. I., Gill G., (2002). SUMO-1 modification represses Sp3 transcriptional activation and modulates its subnuclear localization. *Mol Cell.*; 10: 831-842.
- Rote N.S.**, Chakrabarti S., Stetzer B.P., (2004). The role of human endogenous retroviruses in trophoblast differentiation and placental development. *Placenta*; 25(8-9): 673-683.
- Rous P.**, (1911). A sarcoma of the fowl transmissible by an agent separable from the tumor cells. *J. Exp. Med.*; 13:397-411.
- Ruprecht K.**, Ferreira H., Flockerzi A., Wahl S., Sauter M., Mayer J., Mueller-Lantzsch N., (2008). Human endogenous retrovirus family HERV-K(HML-2) RNA transcripts are selectively packaged into retroviral particles produced by the human germ cell tumor line Tera-1 and originate mainly from a provirus on chromosome 22q11.21. *J Virol.*; 82(20):10008-16.
- Sapetschnig A.**, Rischitor G., Braun H., Doll A., Schergaut M., Melchior F., Suske G., (2002). Transcription factor Sp3 is silenced through SUMO modification by PIAS1. *EMBO J.*; 21: 5206-5215.
- Sapetschnig A.**, Koch F., Rischitor G., Mennenga T., Suske G., (2004). Complexity of translationally controlled transcription factor Sp3 isoform expression. *J Biol Chem.*; 279: 42095-42105.
- Sauter M.**, Schommer S., Kremmer E., Remberger K., Dolken G., Lemm I., Buck M., Best B., Neumann-Haefelin D., Mueller-Lantzsch N., (1995). Human endogenous retrovirus K10: expression of Gag protein and detection of antibodies in patients with seminomas. *J Virol.*; 69: 414-421.
- Sauter M.**, Roemer K., Best B., Afting M., Schommer S., Seitz G., Hartmann M., Mueller-Lantzsch N., (1996). Specificity of antibodies directed against Env-protein of human endogenous retroviruses in patients with germ cell tumors. *Cancer Res.*; 56(19): 4362-5.
- Shatkin A. J. und Manley J. L.**, (2000). The ends of the affair: capping and polyadenylation. *Nat Struct Biol.*; 7(10): 838-42.
- Schommer S.**, Sauter M., Krausslich H. G., et al. (1996). Characterization of the human endogenous retrovirus K proteinase. *J Gen Virol.*; 77(2): 375-379.
- Shao H.**, Revach M., Moshonov S., Tzuman Y., Gazit K., Albeck S., Unger T., Dikstein R., (2005). Core promoter binding by histonelike TAF complexes. *Mol Cell Biol.*; 25: 206–219.
- Singer V. L.**, Wobbe C. R., Struhl K., (1990). A wide variety of DNA sequences can functionally replace a yeast TATA element for transcriptional activation. *Genes Dev.* ; 4: 636–645.

- Sinha S.**, Maity S. N., Lu J., de Crombrughe B., (1995). Recombinant rat CBF-C, the third subunit of CBF/NFY, allows formation of a protein-DNA complex with CBF-A and CBF-B and with yeast HAP2 and HAP3. *Proc Natl Acad Sci USA.*; 92(5): 1624-8.
- 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(1): 188-98.
- Seifarth W.**, Baust C., Murr A., et al., (1998). Proviral structure, chromosomal location, and expression of HERV-K-T47D, a novel human endogenous retrovirus derived from T47D particles. *J Virol.*; 72(10): 8384-8391.
- Spencer T. E.**, Mura M., Gray C. A., Griebel P. J., Palmarini M., (2003). Receptor usage and fetal expression of ovine endogenous betaretroviruses: implications for coevolution of endogenous and exogenous retroviruses. *J Virol.*; 77(1): 749-53.
- Strachan T.** und **Read A. P.**, (1999). (2nd edR). *Hum. Mol. Genet.*; vol. 2. BIOS Scientific Publishers Ltd., Oxford, UK.
- Stauffer Y.**, Theiler G., Sperisen P., et al., (2004). Digital expression profiles of human endogenous retroviral families in normal and cancerous tissues. *Cancer Immun.*; 4: 2.
- Stevens A.**, (1960). Incorporation of the adenine ribonucleotide into RNA by cell fractions from *E. coli* B. *Biochem Biophys Res Commun.*; 3: 92-96.
- Strazzullo M.**, Majello B., Lania L., La Mantia G., (1994). Mutational analysis of the human endogenous ERV9 proviruses promoter region. *Virology*; 200(2): 686-95.
- Sugimoto J.**, Matsuura N., Kinjo Y., Takasu N., Oda T., Jinno Y., (2001). Transcriptionally active HERV-K genes: Identification, isolation, and chromosomal mapping. *Genomics*; 72(2): 137-44.
- Sundberg E. J.**, Li Y., Mariuzza R. A., (2002). So many ways of getting in the way: diversity in the molecular architecture of superantigen-dependent Tcell signaling complexes. *Curr Opin Immunol.*; 14: 36-44.
- Suske G.**, (1999). The Sp-family of transcription factors. *Gene*; 238: 291-300.
- Temin H. M.** und **Mizutani S.**, (1970). RNA-dependant DNA Polymerase in Virions of Rous Sarcoma Virus. *Nature*; 226: 1211-1213.
- Thomas M. C.** und **Chian C. M.**, (2006). The General Transcription Machinery and General Cofactors *Crit Rev Biochem Mol Biol.*; 41(3): 105-78.
- Truss M.**, Chalepakis G., Piña B., Baretino D., Brüggemeier U., Kalff M., Slater E. P., Beato M., (1992). Transcriptional control by steroid hormones. *J Steroid Biochem Mol Biol.*; 41(3-8): 241-8.
- Tsai F. T. F.** und **Sigler P. B.**, (2000). Structural basis of preinitiation complex assembly on human pol II promoters. *EMBO J.*; 19: 5-36.
- Tönjes R. R.**, Limbach C., Löwer R., Kurth R., (1997). Expression of human endogenous retrovirus type K envelope glycoprotein in insect and mammalian cells. *J Virol.*; 71(4): 2747-56.
- Turner G.**, Barbulescu M., Su M., et al., (2001). Insertional polymorphisms of full-length endogenous retroviruses in humans. *Curr Biol.*; 11(19): 1531-1535.
- Van Loo P. F.**, Bouwman P., Ling K. W., Middendorp S., Suske G., Grosveld F., Dzierzak E., Philipsen S., Hendriks R. W., (2003). Impaired hematopoiesis in mice lacking the transcription factor Sp3. *Blood*; 102: 858-866.
- Verdin E.**, Van Lint C., (1995). Internal transcriptional regulatory elements in HIV-1 and other retroviruses. *Cell Mol Biol (Noisy-le-grand)*; 41(3): 365-9.

- Wang T.**, Zeng J., Lowe C. B., Sellers R. G., Salama S. R., Yang M., Burgess S. M., Brachmann R. K., Haussler D., (2007). Species-specific endogenous retroviruses shape the transcriptional network of the human tumor suppressor protein p53. *Proc Natl Acad Sci USA*; 104(47): 18613-8.
- Wang H.**, Machesky N.J., Mansky L.M., (2004). Both the PPPY and PTAP motifs are involved in human T-cell leukemia virus type 1 particle release. *J Virol.*; 78(3): 1503-12.
- Wasylyk B.**, (1988). Transcription elements and factors of RNA polymerase B promoters of higher eukaryotes. *CRC Crit Rev Biochem.*; 23: 77-120.
- Watt F.** und **Molloy P. L.**, (1988). Cytosine methylation prevents binding to DNA of a HeLa cell transcription factor required for optimal expression of the adenovirus major late promoter. *Genes Dev.*; 2(9): 1136-43.
- Weinmann R.**, Raskas H. J., Roeder R. G., (1974). Role of DNA-dependent RNA polymerases II and III in transcription of the adenovirus genome late in productive infection. *Proc Natl Acad Sci USA*; 71: 3426-3439.
- Weinmann R.** and **Roeder R. G.**, (1974). Role of DNA-dependent RNA polymerase III in the transcription of the tRNA and 5S RNA genes. *Proc Natl Acad Sci USA*; 71: 1790-1794.
- Weiss S.** und **Gladstone L.**, (1959). A mammalian system for the incorporation of cytidine triphosphate into ribonucleic acid. *J Am Chem Soc.*; 81: 4118-4119.
- Weis L.** und **Reinberg D.**, (1992). Transcription by RNA polymerase II: initiator-directed formation of transcription-competent complexes. *FASEB J.*; 6(14): 3300-9.
- Wickens M.**, (1990). How the messenger got its tail: addition of poly(A) in the nucleus. *Trends Biochem Sci.*; 15(7): 277-81.
- Wilkinson D. A.**, Mager D. L., Leong J. C., (1994). Endogenous human retrovirus. In *The Retroviridae*, J.A.Levy, ed. (New York: Plenum Press) 465-535.
- Wobbe C. R.** und **Struhl K.**, (1990). Yeast and human TATA-binding proteins have nearly identical DNA sequence requirements for transcription in vitro. *Mol. Cell. Biol.*; 10: 3859-3867.
- Wu S. Y.** und **Chiang C. M.**, (1998). Properties of PC4 and an RNA polymerase II complex in directing activated and basal transcription in vitro. *J Biol Chem.*; 273: 12492-12498.
- Wu S. Y.**, Thomas M. C., Hou S. Y., Likhite V., Chiang C. M., (1999). Isolation of mouse TFIID and functional characterization of TBP and TFIID in mediating estrogen receptor and chromatin transcription. *J Biol Chem.*; 274: 23480-23490.
- Yang X.**, Su K., Roos M. D., Chang Q., Paterson A. J., Kudlow J. E., (2001). O-linkage of N-acetylglucosamine to Sp1 activation domain inhibits its transcriptional capability. *Proc. Natl. Acad. Sci. USA*; 98: 6611-6616.
- Yang P.**, Zemba M., Aboud M., Flugel R. M., Lochelt M., (1997). Deletion analysis of both the long terminal repeat and the internal promoters of the human foamy virus. *Virus Genes*; 15(1): 17-23.
- Yasuda J.**, und Hunter E., (1998). A proline-rich motif (PPPY) in the Gag polyprotein of Mason-Pfizer monkey virus plays a maturation-independent role in virion release. *J Virol.*; 72(5): 4095-103.
- Yieh L.**, Sanchez H. B., Osborne T. F., (1995). Domains of transcription factor Sp1 required for synergistic activation with sterol regulatory element binding protein 1 of low density lipoprotein receptor promoter. *Proc. Natl. Acad. Sci. USA*; 92: 6102-6106.
- Yu X.**, Yuan X., Matsuda Z., Lee T. H., Essex M., (1992). The matrix protein of human immunodeficiency virus type 1 is required for incorporation of viral envelope protein into mature virions. *J Virol.*; 66(8): 4966-71.

Yu B., Datta P. K., Bagchi S., (2003). Stability of the Sp3-DNA complex is promoter-specific: Sp3 efficiently competes with Sp1 for binding to promoters containing multiple Sp-sites. *Nucleic Acids Res.*; 31: 5368–5376.

Zawel L. und **Reinberg D. R.**, (1993). Initiation of transcription by RNA polymerase II: a multi-step process. *Prog. Nucl. Acids Res. Mol. Biol.*, 44: 67–108.

Zhou T., und **Chiang C. M.** (2000). The intronless and TATA-less human TAFII55 gene contains a functional initiator and a downstream promoter element. *J Biol Chem.*; 276: 25503–25511.

Zybarth G., Kräusslich H.G., Partin K., Carter C., (1994). Proteolytic activity of novel human immunodeficiency virus type 1 proteinase proteins from a precursor with a blocking mutation at the N terminus of the PR domain. *J Virol.*; 68(1): 240-50.

7. ABKÜRZUNGSVERZEICHNIS

ψ	Verpackungssignal
AIDS	Aquired Immune Deficiency Syndrome
ALV	Aviäres Leukämie-Virus
bp	Basenpaar
BRE	TFIIB recognition element
CA	Capsid Protein
cDNA	komplementäre Desoxyribonukleinsäure
CDP	CCAAT <i>displacement protein</i>
C/EBP	CCAAT/ <i>enhancer binding protein</i>
CTF/NF-I	CCAAT <i>transcription factor/nuclear factor-I</i>
DCE	Downstream core element
DNA	Desoxyribonukleinsäure
DPE	Downstream Promoter Element
EMSA	Electrophoretic Mobility Shift Assay
Env	Envelope Protein
ERV	endogenes Retrovirus
Gag	Group Specific Antigen
GC/GT-Box	Sp1/Sp3-Konsensussequenz
GH	humane Teratokarzinomzelllinie
GTF	Genereller Transkriptionsfaktor
HeLa	humane Cervixkarzinomzelllinie
HERV	Humanes endogenes Retrovirus
HIV	Humanes Immundefizienzvirus
HTDV	Human Teratocarcinoma Derived Virus
HTLV	Humanes T-Zell Leukämie Virus
IN	Integrase Protein
Inr	Initiator
JSRV	Jaagsiekte-Schaf-Retrovirus
kDA	Kilodalton
kB	Kilobase
LINE	Long Interspersed Element
LTR	Long Terminal Repeat
MA	Matrix Protein
MMTV	Maus-Mamma-Tumor-Virus
mRNA	Messenger RNA
MTE	Motif ten element
nt	Nukleotid (A, C, T, G)
NC	Nukleocapsid Protein
NF-κB	Nukl. Faktor der Ig κ leichten Kette von B-Zellen
ORF	Open Reading Frame
PBS	Primer Binding Site
PCR	Polymerasekettenreaktion
PVF	primaten foamy virus
Pol	Polymerase / Reverse Transkriptase Protein
PolII, PolIII, PolIII	RNA-Polymerase I, II, III
Poly(A)-	Polyadenylierungs-
PPT	Poly-Purin Tract

PRT	Protease Protein
Pu	Purin
Py	Pyrimidin
R	Terminal Redundancy
RACE	Rapid Amplification of cDNA Ends
Rec	funktionales HERV-K-Äquivalent zu Rev (HIV)
Rev	Regulator of Expression of Virion Proteins (HIV)
RNA	Ribonukleinsäure
rRNA	Ribosomale RNA
RSV	Rous-Sarkom-Virus
RT	Reverse Transkriptase Protein
SA	Spleißakzeptor
SD	Spleißdonor
SINE	Short Interspersed Element
sLTR	solitäre LTR
SNV	Spleen necrosis virus
Sp1, Sp3	Specificity Protein 1, 3
SRY	Sex-determining Region Y
SU	Surface Protein
SUMO	Small Ubiquitin-like Modifier
SV40	Simian Virus 40
TAF	TBP-assoziierter Faktor
Tat	Transactivator of Transcription (Protein)
TATA-Box	TATA-Binde Protein-Konsensussequenz
TBP	TATA-Binde Protein
TFIID	Transkriptionsfaktor IID
TM	Transmembran Protein
tRNA	Transfer RNA
U	Unit
U3	Unique 3'-End
U5	Unique 5'-End
UAS	Upstream activating sequences
YY1	Yin-Yang 1 (Protein)

Abkürzungen, die nur einmal verwendet wurden, sind an Ort und Stelle erläutert. Aminosäuren und Nukleotide werden im Einbuchstabencode abgekürzt. Gen-Namen sind kursiv gedruckt. Physikalische, chemische und biologische Größen sowie gängige Abkürzungen der deutschen Sprache sind hier nicht aufgeführt.

8. ANHANG

8.1. Primerliste

Primer	Sequenz	
HK A 113	5'GCA TGG GGA GAT TCA GAA CC3'	5' flank.
HK B 113	5'CTT TTA GTT TGC CCC ATT ATC3'	<i>gag</i>
HK C 113	5' ATG GCC TCG CAA ATC TTA3'	<i>env</i>
HK D 113	5'CAT GTT TCC TGT CTG CCC AC3'	3' flank
HK A 115	5'CCG CAC CTA GTC AAC TTA GC3'	5' flank
HK B 115	5'GCT CGA GCG TAC CTT CAC C3'	<i>gag</i>
HK C 115	5'GTG TGC CTG TTT TGT CTG TTG TTA G3'	<i>env</i>
HK D 115	5' TCA GTT CCC GAT TTT CTG CC 3'	3' flank
GeneRacer™ 5' Primer	5'-CGACTGGAGCACGAGGACACTGA-3'	RACE
B2	5'-GAAAAGCCTCCACGTTGGGCACCA-3'	PBS
GeneRacer™ 5' Nested Primer	5'-GGACACTGACATGGACTGAAGGAGTA-3'	RACE
Lys1,2	5'-GCCCCAGGTTGGGCGCCA-3	PBS
GeneRacer™ 3' Primer	5'-GCTGTCAACGATACGCTACGTAACG-3'	RACE
Env-sense	5'-CCAGACTCCCAGACTATAACCTGTG-3'	<i>env</i>
ChIP-LTR-FW	5'-TCCTTTACATTGTCTATGATGCAAAGAC-3'	LTR
ChIP-LTR-RV	5'-GGAAAAGAAAAAGACACAGAGACAAAGT-3'	LTR
ChIP-Env-FW	5'-GTGCACAAGTGAGTCCAGCTGTT-3'	<i>env</i>
ChIP-Env-RV	5'-CATATAAGGGGGCTTTACGCAA-3'	<i>env</i>
HERV-K gag +	Ruprecht et al. 2008	<i>gag</i>
HERV-K gag -	Ruprecht et al. 2008	<i>gag</i>
HERV-K env 8146	Ruprecht et al. 2008	<i>env</i>
HERV-K env 8665	Ruprecht et al. 2008	<i>env</i>
pol-FW	Laderoute et al. 2007	<i>pol</i>
pol-RV	Laderoute et al. 2007	<i>pol</i>
MKr_21	5'-AAGCTTTTATTGATC-3'	LTR
MKr_27	5-'AAGCTTCTTGTGACC-3'	LTR
JM1-gag	5'CTAAAAGTAAAATTTAAAAGTAAATATGCCTC-3'	<i>gag</i>
JM2-gag/prot		<i>prot</i>

8.2. Sequenzen

Nukleotidalignment aus dem *env*-Bereich, Klone - Partikel GH

	1	50
GH_E_P14	(1)	AATGAGTCTGAGCATCACTGGGACATGGTTAGACGCCATCTACAGGGAAG
GH_P_E11 rc	(1)	AATGAGTCTGAGCATCACTGGGACATGGTTAGACGCCATCTACAGGGAAG
GH_P_E7 rc	(1)	AATGAGTCTGAGCATCACTGGGACATGGTTAGACGCCATCTACAGGGAAG
GH_P_E12	(1)	AATGAGTCTGAGCATCACTGGGACATGGTTAGACGCCATCTACAGGGAAG
GH_P_E4	(1)	AATGAGTCTGAGCATCACTGGGACATGGTTAGACGCCATCTACAGGGAAG
GH_P_E18	(1)	AATGAGTCTGAGCATCACTGGGACATGGTTAGACGCCATCTACAGGGAAG
GH_E_P13	(1)	AATGAGTCTGAGCATCACTGGGACATGGTTAGACGCCATCTACAGGGAAG
GH_P_E6	(1)	AATGAGTCTGAGCATCACTGGGACATGGTTAGACGCCATCTACAGGGAAG
GH_P_E8	(1)	AATGAGTCTGAGCATCACTGGGACATGGTTAGACACCATCTACAGGGAAG
GH_P_E9	(1)	AATGAGTCTGAGCATCACTGGGACATGGTTAGACACCATCTACAGGGAAG
GH_P_E17 rc	(1)	AATGAGTCTGAGCATCACTGGGACATGGTTAGACACCATCTACAGGGAAG
GH_P_E10	(1)	AATGAGTCTGAGCATCACTGGGACATGGTTAGACACCATCTACAGGGAAG
GH_P_E1 rc	(1)	AATGAGTCTGAGCATCACTGGGACATGGTTAGACACCATCTACAGGGAAG
GH_P_E15 rc	(1)	AATGAGTCTGAGCATCACTGGGACATGGTTAGACACCATCTACAGGGAAG
GH_P_E5 rc	(1)	AATGAGTCTGAGCATCACTGGGACATGGTTAGACACCATCTACAGGGAAG
GH_P_E16	(1)	AATGAGTCTGAGCATCACTGGGACATGGTTAGACACCATCTACAGGGAAG
GH_P_E20	(1)	AATGAGTCTGAGCATCACTGGGACATGGTTAGACACCATCTACAGGGAAG
GH_P_E19 rc	(1)	AATGAGTCTGAGCATCACTGGGACATGGTTAGACACCATCTACAGGGAAG
	51	100
GH_E_P14	(51)	AGAAGATAATCTCACTTTAGACATTTCCAAATTTAAAAGAACAATTTTTCG
GH_P_E11 rc	(51)	AGAAGATAATCTCACTTTAGACATTTCCAAATTTAAAAGAACAATTTTTCG
GH_P_E7 rc	(51)	AGAAGATAATCTCACTTTAGACATTTCCAAATTTAAAAGAACAATTTTTCG
GH_P_E12	(51)	AGAAGATAATCTCACTTTAGACATTTCCAAATTTAAAAGAACAATTTTTCG
GH_P_E4	(51)	AGAAGATAATCTCACTTTAGACATTTCCAAATTTAAAAGAACAATTTTTCG
GH_P_E18	(51)	AGAAGATAATCTCACTTTAGACATTTCCAAATTTAAAAGAACAATTTTTCG
GH_E_P13	(51)	AGAAGATAATCTCACTTTAGACATTTCCAAATTTAAAAGAACAATTTTTCG
GH_P_E6	(51)	AGAAGATAATCTCACTTTAGACATTTCCAAATTTAAAAGAACAATTTTTCG
GH_P_E8	(51)	AGAAGATAATCTCACTTTAGACATTTCCAAATTTAAAAGAACAATTTTTCG
GH_P_E9	(51)	AGAAGATAATCTCACTTTAGACATTTCCAAATTTAAAAGAACAATTTTTCG
GH_P_E17 rc	(51)	AGAAGATAATCTCACTTTAGACATTTCCAAATTTAAAAGAACAATTTTTCG
GH_P_E10	(51)	AGAAGATAATCTCACTTTAGACATTTCCAAATTTAAAAGAACAATTTTTCG
GH_P_E1 rc	(51)	AGAAGATAATCTCACTTTAGACATTTCCAAATTTAAAAGAACAATTTTTCG
GH_P_E15 rc	(51)	AGAAGATAATCTCACTTTAGACATTTCCAAATTTAAAAGAACAATTTTTCG
GH_P_E5 rc	(51)	AGAAGATAATCTCACTTTAGACATTTCCAAATTTAAAAGAACAATTTTTCG
GH_P_E16	(51)	AGAAGATAATCTCACTTTAGACATTTCCAAATTTAAAAGAACAATTTTTCG
GH_P_E20	(51)	AGAAGATAATCTCACTTTAGACATTTCCAAATTTAAAAGAACAATTTTTCG
GH_P_E19 rc	(51)	AGAAGATAATCTCACTTTAGACATTTCCAAATTTAAAAGAACAATTTTTCG
	101	150
GH_E_P14	(101)	AAGCATCAAAAGCCCATTTAAATTTGGTGCCAGGAACTGAGGCAATTGCA
GH_P_E11 rc	(101)	AAGCATCAAAAGCCCATTTAAATTTGGTGCCAGGAACTGAGGCAATTGCA
GH_P_E7 rc	(101)	AAGCATCAAAAGCCCATTTAAATTTGGTGCCAGGAACTGAGGCAATTGCA
GH_P_E12	(101)	AAGCATCAAAAGCCCATTTAAATTTGGTGCCAGGAACTGAGGCAATTGCA
GH_P_E4	(101)	AAGCATCAAAAGCCCATTTAAATTTGGTGCCAGGAACTGAGGCAATTGCA
GH_P_E18	(101)	AAGCATCAAAAGCCCATTTAAATTTGGTGCCAGGAACTGAGGCAATTGCA
GH_E_P13	(101)	AAGCATCAAAAGCCCATTTAAATTTGGTGCCAGGAACTGAGGCAATTGCA
GH_P_E6	(101)	AAGCATCAAAAGCCCATTTAAATTTGGTGCCAGGAACTGAGGCAATTGCA
GH_P_E8	(101)	AAGCATCAAAAGCCCATTTAAATTTGGTGCCAGGAACTGAGGCAATTGCA
GH_P_E9	(101)	AAGCATCAAAAGCCCATTTAAATTTGGTGCCAGGAACTGAGGCAATTGCA
GH_P_E17 rc	(101)	AAGCATCAAAAGCCCATTTAAATTTGGTGCCAGGAACTGAGGCAATTGCA
GH_P_E10	(101)	AAGCATCAAAAGCCCATTTAAATTTGGTGCCAGGAACTGAGGCAATTGCA
GH_P_E1 rc	(101)	AAGCATCAAAAGCCCATTTAAATTTGGTGCCAGGAACTGAGGCAATTGCA
GH_P_E15 rc	(101)	AAGCATCAAAAGCCCATTTAAATTTGGTGCCAGGAACTGAGGCAATTGCA
GH_P_E5 rc	(101)	AAGCATCAAAAGCCCATTTAAATTTGGTGCCAGGAACTGAGGCAATTGCA
GH_P_E16	(101)	AAGCATCAAAAGCCCATTTAAATTTGGTGCCAGGAACTGAGGCAATTGCA
GH_P_E20	(101)	AAGCATCAAAAGCCCATTTAAATTTGGTGCCAGGAACTGAGGCAATTGCA
GH_P_E19 rc	(101)	AAGCATCAAAAGCCCATTTAAATTTGGTGCCAGGAACTGAGGCAATTGCA

	151	200
GH_E_P14	(151)	GGAGTTGCTGATGGCCTCGCAAATCTTAACCCGTGCACTTGGGTAAAGAC
GH_P_E11 rc	(151)	GGAGTTGCTGATGGCCTCGCAAATCTTAACCCGTGCACTTGGGTAAAGAC
GH_P_E7 rc	(151)	GGAGTTGCTGATGGCCTCGCAAATCTTAACCCGTGCACTTGGGTAAAGAC
GH_P_E12	(151)	GGAGTTGCTGTTGGCCTCGCAAATCTTAACCCGTGCACTTGGGTAAAGAC
GH_P_E4	(151)	GGAGTTGCTGATGGCCTCGCAAATCTTAACCCGTGCACTTGGGTAAAGAC
GH_P_E18	(151)	GGAGTTGCTGATGGCCTCGCAAATCTTAACCCGTGCACTTGGGTAAAGAC
GH_E_P13	(151)	GGAGTTGCTGATGGCCTCGCAAATCTTAACCCGTGCACTTGGGTAAAGAC
GH_P_E6	(151)	GGAGTTGCTGATGGCCTCGCAAATCTTAACCCGTGCACTTGGGTAAAGAC
GH_P_E8	(151)	GGAGTTGCTGATGGCCTCGCAAATCTTAACCCGTGCACTTGGGTAAAGAC
GH_P_E9	(151)	GGAGTTGCTGATGGCCTCGCAAATCTTAACCCGTGCACTTGGGTAAAGAC
GH_P_E17 rc	(151)	GGAGTTGCTGATGGCCTCGCAAATCTTAACCCGTGCACTTGGGTAAAGAC
GH_P_E10	(151)	GGAGTTGCTGATGGCCTCGCAAATCTTAACCCGTGCACTTGGGTAAAGAC
GH_P_E1 rc	(151)	GGAGTTGCTGATGGCCTCGCAAATCTTAACCCGTGCACTTGGGTAAAGAC
GH_P_E15 rc	(151)	GGAGTTGCTGATGGCCTCGCAAATCTTAACCCGTGCACTTGGGTAAAGAC
GH_P_E5 rc	(151)	GGAGTTGCTGATGGCCTCGCAAATCTTAACCCGTGCACTTGGGTAAAGAC
GH_P_E16	(151)	GGAGTTGCTGATGGCCTCGCAAATCTTAACCCGTGCACTTGGGTAAAGAC
GH_P_E20	(151)	GGAGTTGCTGATGGCCTCGCAAATCTTAACCCGTGCACTTGGGTAAAGAC
GH_P_E19 rc	(151)	GGAGTTGCTGATGGCCTCGCAAATCTTAACCCGTGCACTTGGGTAAAGAC

	201	250
GH_E_P14	(201)	CATTGGAAGTACTACGATTATAAATCTCATATTAATCCTTGTGTGCCTGT
GH_P_E11 rc	(201)	CATTGGAAGTACTACGATTATAAATCTCATATTAATCCTTGTGTGCCTGT
GH_P_E7 rc	(201)	CATTGGAAGTACTACGATTATAAATCTCATATTAATCCTTGTGTGCCTGT
GH_P_E12	(201)	CATTGGAAGTACTACGATTATAAATCTCATATTAATCCTTGTGTGCCTGT
GH_P_E4	(201)	CATTGGAAGTACTACGATTATAAATCTCATATTAATCCTTGTGTGCCTGT
GH_P_E18	(201)	CATTGGAAGTACTACGATTATAAATCTCATATTAATCCTTGTGTGCCTGT
GH_E_P13	(201)	CATTGGAAGTACTACGATTATAAATCTCATATTAATCCTTGTGTGCCTGT
GH_P_E6	(201)	CATTGGAAGTACTACGATTATAAATCTCATATTAATCCTTGTGTGCCTGT
GH_P_E8	(201)	CATTGGAAGTACTACGATTATAAATCTCATATTAATCCTTGTGTGCCTGT
GH_P_E9	(201)	CATTGGAAGTACTACGATTATAAATCTCATATTAATCCTTGTGTGCCTGT
GH_P_E17 rc	(201)	CATTGGAAGTACTACGATTATAAATCTCATATTAATCCTTGTGTGCCTGT
GH_P_E10	(201)	CATTGGAAGTACTACGATTATAAATCTCATATTAATCCTTGTGTGCCTGT
GH_P_E1 rc	(201)	CATTGGAAGTACTACGATTATAAATCTCATATTAATCCTTGTGTGCCTGT
GH_P_E15 rc	(201)	CATTGGAAGTACTACGATTATAAATCTCATATTAATCCTTGTGTGCCTGT
GH_P_E5 rc	(201)	CATTGGAAGTACTACGATTATAAATCTCATATTAATCCTTGTGTGCCTGT
GH_P_E16	(201)	CATTGGAAGTACTACGATTATAAATCTCATATTAATCCTTGTGTGCCTGT
GH_P_E20	(201)	CATTGGAAGTACTACGATTATAAATCTCATATTAATCCTTGTGTGCCTGT
GH_P_E19 rc	(201)	CATTGGAAGTACTACGATTATAAATCTCATATTAATCCTTGTGTGCCTGT

	251	300
GH_E_P14	(251)	TTTGTCTGTTGTTAGTCTGCAGGTGTACCCAACAGCTCCGAAGAGACAGC
GH_P_E11 rc	(251)	TTTGTCTGTTGTTAGTCTGCAGGTGTACCCAACAGCTCCGAAGAGACAGC
GH_P_E7 rc	(251)	TTTGTCTGTTGTTAGTCTGCAGGTGTACCCAACAGCTCCGAAGAGACAGC
GH_P_E12	(251)	TTTGTCTGTTGTTAGTCTGCAGGTGTACCCAACAGCTCCGAAGAGACAGC
GH_P_E4	(251)	TTTGTCTGTTGTTAGTCTGCAGGTGTACCCAACAGCTCCGAAGAGACAGC
GH_P_E18	(251)	TTTGTCTGTTGTTAGTCTGCAGGTGTACCCAACAGCTCCGAAGAGACAGC
GH_E_P13	(251)	TTTGTCTGTTGTTAGTCTGCAGGTGTACCCAACAGCTCCGAAGAGACAGC
GH_P_E6	(251)	TTTGTCTGTTGTTAGTCTGCAGGTGTACCCAACAGCTCCGAAGAGACAGC
GH_P_E8	(251)	TTTGTCTGTTGTTAGTCTGCAGGTGTACCCAACAGCTCCGAAGAGACAGC
GH_P_E9	(251)	TTTGTCTGTTGTTAGTCTGCAGGTGTACCCAACAGCTCCGAAGAGACAGC
GH_P_E17 rc	(251)	TTTGTCTGTTGTTAGTCTGCAGGTGTACCCAACAGCTCCGAAGAGACAGC
GH_P_E10	(251)	TTTGTCTGTTGTTAGTCTGCAGGTGTACCCAACAGCTCCGAAGAGACAGC
GH_P_E1 rc	(251)	TTTGTCTGTTGTTAGTCTGCAGGTGTACCCAACAGCTCCGAAGAGACAGC
GH_P_E15 rc	(251)	TTTGTCTGTTGTTAGTCTGCAGGTGTACCCAACAGCTCCGAAGAGACAGC
GH_P_E5 rc	(251)	TTTGTCTGTTGTTAGTCTGCAGGTGTACCCAACAGCTCCGAAGAGACAGC
GH_P_E16	(251)	TTTGTCTGTTGTTAGTCTGCAGGTGTACCCAACAGCTCCGAAGAGACAGC
GH_P_E20	(251)	TTTGTCTGTTGTTAGTCTGCAGGTGTACCCAACAGCTCCGAAGAGACAGC
GH_P_E19 rc	(251)	TTTGTCTGTTGTTAGTCTGCAGGTGTACCCAACAGCTCCGAAGAGACAGC

	301	350
GH_E_P14	(301)	GACCATCGAGAACGGGCCATGATGACGATGGCGGTTTTGTGCGAAAAGAAA
GH_P_E11 rc	(301)	GACCATCGAGAACGGGCCATGATGACGATGGCGGTTTTGTGCGAAAAGAAA
GH_P_E7 rc	(301)	GACCATCGAGAACGGGCCATGATGACGATGGCGGTTTTGTGCGAAAAGAAA
GH_P_E12	(301)	GACCATCGAGAACGGGCCATGATGACGACGGTGGTTTTGTGCGAAAAGAAA
GH_P_E4	(301)	GACCATCGAGAACGGGCCATGATGACGATGGCGGTTTTGTGCGAAAAGAAA
GH_P_E18	(301)	GACCATCGAGAACGGGCCATGATGACGATGGCGGTTTTGTGCGAAAAGAAA
GH_E_P13	(301)	GACCATCGAGAACGGGCCATGATGACGATGGCGGTTTTGTGCGAAAAGAAA
GH_P_E6	(301)	GACCATCGAGAACGGGCCATGATGACGATGGCGGTTTTGTGCGAAAAGAAA
GH_P_E8	(301)	GACCATCGAGAACGGGCCATGATGACGATGGCGGTTTTGTGCGAAAAGAAA
GH_P_E9	(301)	GACCATCGAGAACGGGCCATGATGACGATGGCGGTTTTGTGCGAAAAGAAA
GH_P_E17 rc	(301)	GACCATCGAGAACGGGCCATGATGACGATGGCGGTTTTGTGCGAAAAGAAA
GH_P_E10	(301)	GACCATCGAGAACGGGCCATGATGACGATGGCGGTTTTGTGCGAAAAGAAA
GH_P_E1 rc	(301)	GACCATCGAGAACGGGCCATGATGACGATGGCGGTTTTGTGCGAAAAGAAA
GH_P_E15 rc	(301)	GACCATCGAGAACGGGCCATGATGACGATGGCGGTTTTGTGCGAAAAGAAA
GH_P_E5 rc	(301)	GACCATCGAGAACGGGCCATGATGACGATGGCGGTTTTGTGCGAAAAGAAA
GH_P_E16	(301)	GACCATCGAGAACGGGCCATGATGACGATGGCGGTTTTGTGCGAAAAGAAA
GH_P_E20	(301)	GACCATCGAGAACGGGCCATGATGACGATGGCGGTTTTGTGCGAAAAGAAA
GH_P_E19 rc	(301)	GACCATCGAGAACGGGCCATGATGACGATGGCGGTTTTGTGCGAAAAGAAA

	351	400
GH_E_P14	(351)	AGGGGGAAATGTGGGGAAAAGCAAGAGAGATCAGATTGTTACTGTGTCTG
GH_P_E11 rc	(351)	AGGGGGAAATGTGGGGAAAAGCAAGAGAGATCAAATTGTTACTGTGTCTG
GH_P_E7 rc	(351)	AGGGGGAAATGTGGGGAAAAGCAAGAGAGATCAAATTGTTACTGTGTCTG
GH_P_E12	(351)	AGGGGGAAATGTGGGGAAAAGCAAGAGAGATCAAATTGTTACTGTGTCTG
GH_P_E4	(351)	AGGGGGAAATGTGGGGAAAAGCAAGAGAGATCAGATTGTTACTGTGTCTG
GH_P_E18	(351)	AGGGGGAAATGTGGGGAAAAGCAAGAGAGATCAGATTGTTACTGTGTCTG
GH_E_P13	(351)	AGGGGGAAATGTGGGGAAAAGCAAGAGAGATCAGATTGTTACTGTGTCTG
GH_P_E6	(351)	AGGGGGAAATGTGGGGAAAAGCAAGAGAGATCAGATTGTTACTGTGTCTG
GH_P_E8	(351)	AGGGGGAAATGTGGGGAAAAGCAAGAGAGATCAGATTGTTACTGTGTCTG
GH_P_E9	(351)	AGGGGGAAATGTGGGGAAAAGCAAGAGAGATCAGATTGTTACTGTGTCTG
GH_P_E17 rc	(351)	AGGGGGAAATGTGGGGAAAAGCAAGAGAGATCAGATTGTTACTGTGTCTG
GH_P_E10	(351)	AGGGGGAAATGTGGGGAAAAGCAAGAGAGATCAGATTGTTACTGTGTCTG
GH_P_E1 rc	(351)	AGGGGGAAATGTGGGGAAAAGCAAGAGAGATCAGATTGTTACTGTGTCTG
GH_P_E15 rc	(351)	AGGGGGAAATGTGGGGAAAAGCAAGAGAGATCAGATTGTTACTGTGTCTG
GH_P_E5 rc	(351)	AGGGGGAAATGTGGGGAAAAGCAAGAGAGATCAGATTGTTACTGTGTCTG
GH_P_E16	(351)	AGGGGGAAATGTGGGGAAAAGCAAGAGAGATCAGATTGTTACTGTGTCTG
GH_P_E20	(351)	AGGGGGAAATGTGGGGAAAAGCAAGAGAGATCAGATTGTTACTGTGTCTG
GH_P_E19 rc	(351)	AGGGGGAAATGTGGGGAAAAGCAAGAGAGATCAGATTGTTACTGTGTCTG

	401	450
GH_E_P14	(401)	TGTAGAAAGAAGTAGACATAGGAGACTCCATTTTGTATGTGCTAAGAAA
GH_P_E11 rc	(401)	TGTAGAAAGAAGTAGACATAGGAGACTCCATTTTGTATGTGCTAAGAAA
GH_P_E7 rc	(401)	TGTAGAAAGAAGTAGACATAGGAGACTCCATTTTGTATGTGCTAAGAAA
GH_P_E12	(401)	TGTAGAAAGAAGTAGACATAGGAGACTCCATTTTGTATGTGCTAAGAAA
GH_P_E4	(401)	TGTAGAAAGAAGTAGACATGGGAGACTCCATTTTGTATGTGTTAAGAAA
GH_P_E18	(401)	TGTAGAAAGAAGTAGACATAGGAGACTCCATTTTGTATGTACTAAGAAA
GH_E_P13	(401)	TGTAGAAAGAAGTAGACATAGGAGACTCCATTTTGTATGTACTAAGAAA
GH_P_E6	(401)	TGTAGAAAGAAGTAGACATAGGAGACTCCATTTTGTATGTACTAAGATA
GH_P_E8	(401)	TGTAGAAAGAAGTAGACATAGGAGACTCCATTTTGTATGTGCTAAGAAA
GH_P_E9	(401)	TGTAGAAAGAAGTAGACATAGGAGACTCCATTTTGTATGTGCTAAGAAA
GH_P_E17 rc	(401)	TGTAGAAAGAAGTAGACATAGGAGACTCCATTTTGTATGTGCTAAGAAA
GH_P_E10	(401)	TGTAGAAAGAAGTAGACATAGGAGACTCCATTTTGTATGTGCTAAGAAA
GH_P_E1 rc	(401)	TGTAGAAAGAAGTAGACATAGGAGACTCCATTTTGTATGTGCTAAGAAA
GH_P_E15 rc	(401)	TGTAGAAAGAAGTAGACATAGGAGACTCCATTTTGTATGTACTAAGAAA
GH_P_E5 rc	(401)	TGTAGAAAGAAGTAGACATAGGAGACTCCATTTTGTATGTACTAAGAAA
GH_P_E16	(401)	TGTAGAAAGAAGTAGACATAGGAGACTCCATTTTGTATGTGCTAAGAAA
GH_P_E20	(401)	TGTAGAAAGAAGTAGACATAGGAGACTCCATTTTGTATGTGCTAAGAAA
GH_P_E19 rc	(401)	TGTAGAAAGAAGTAGACATAGGAGACTCCATTTTGTATGTGCTAAGAAA

		451		500
GH_E_P14	(451)	AATTCTTCTGCCTTGAGATTCTGTTAATCTATGACCTTACCCCCAACCC		
GH_P_E11 rc	(451)	AATTCTTCTGCCTTGAGATTCTGTTAATCTATGACCTTACCCCCAACCC		
GH_P_E7 rc	(451)	AATTCTTCTGCCTTGAGATTCTGTTAATCTATGACCTTACCCCCAACCC		
GH_P_E12	(451)	AATTCTTCTGCCTTGAGATTCTGTTAATCTATAACCTTACCCCCAACCC		
GH_P_E4	(451)	AATTCTTCTGCCTTGAGATTCTGTTAATCTATGACCTTACCCCCAACCC		
GH_P_E18	(451)	AATTCTTCTGCCTTGAGATTCTGTTAATCTATGACCTTACCCCCAACCC		
GH_E_P13	(451)	AATTCTTCTGCCTTGAGATTCTGTTAATCTATGACCTTACCCCCAACCC		
GH_P_E6	(451)	AATTCTTCTGCCTTGAGATTCTGT-----GACCTTACCCCCAACCC		
GH_P_E8	(451)	AATTCTTCTGCCTTGAGATTCTGTTAATCTATAACCTTACCCCCAACCC		
GH_P_E9	(451)	AATTCTTCTGCCTTGAGATTCTGTTAATCTATAACCTTACCCCCAACCC		
GH_P_E17 rc	(451)	AATTCTTCTGCCTTGAGATTCTGTTAATCTATAACCTTACCCCCAACCC		
GH_P_E10	(451)	AATTCTTCTGCCTTGAGATTCTGTTAATCTATGACCTTACCCCCAACCC		
GH_P_E1 rc	(451)	AATTCTTCTGCCTTGAGATTCTGTTAATCTATAACCTTACCCCCAACCC		
GH_P_E15 rc	(451)	AATTCTTCTGCCTTGAGATTCTGTTAATCTATGACCTTACCCCCAACCC		
GH_P_E5 rc	(451)	AATTCTTCTGCCTTGAGATTCTGTTAATCTATGACCTTACCCCCAACCC		
GH_P_E16	(451)	AATTCTTCTGCCTCGAGATTCTGTTAATCTATAACCTTACCCCCAACCC		
GH_P_E20	(451)	AATTCTTCTGCCTTGAGATTCTGTTAATCTATAACCTTACCCCCAACCC		
GH_P_E19 rc	(451)	AATTCTTCTGCCTTGAGATTCTGTTAATCTATAACCTTACCCCCAACCC		
		501		543
GH_E_P14	(501)	GTGCTCTCTGAAACATGTGCTGTGTCCACTCAGAGTTGAATGG		
GH_P_E11 rc	(501)	GTGCTCTCTGAAACATGTGCTGTGTCCACTCAGAGTTGAATGG		
GH_P_E7 rc	(501)	GTGCTCTCTGAAACATGTGCTGTGTCCACTCAGAGTTGAATGG		
GH_P_E12	(501)	GTGCTCTCTGAAACATGTGCTGTGTCCACTCAGAGTTGAATGG		
GH_P_E4	(501)	GTGCTCTCTGAAACATGTGCTGTGTCCACTCAGAGTTGAATGG		
GH_P_E18	(501)	GTGCTCTCTGAAACATGTGCTGTGTCCACTCAGAGTTGAATGG		
GH_E_P13	(501)	GTGCTCTCTGAAACATGTGCTGTGTCCACTCAGAGTTGAATGG		
GH_P_E6	(493)	GTGCTCTCTGAAACATGTGCTGTGTCCACTCAGAGTTGAATGG		
GH_P_E8	(501)	GTGCTCTCTGAAACATGTGCTGTGTCCACTCAGAGTTGAATGG		
GH_P_E9	(501)	GTGCTCTCTGAAACATGTGCTGTGTCCACTCAGAGTTGAATGG		
GH_P_E17 rc	(501)	GTGCTCTCTGAAACATGTGCTGTGTCCACTCAGAGTTGAATGG		
GH_P_E10	(501)	GTGCTCTCTGAAACATGTGCTGTGTCCACTCAGAGTTGAATGG		
GH_P_E1 rc	(501)	GTGCTCTCTGAAACATGTGCTGTGTCCACTCAGAGTTGAATGG		
GH_P_E15 rc	(501)	GTGCTCTCTGAAACATGTGCTGTGTCCACTCAGAGTTGAATGG		
GH_P_E5 rc	(501)	GTGCTCTCTGAAACATGTGCTGTGTCCACTCAGAGTTGAATGG		
GH_P_E16	(501)	GTGCTCTCTGAAACATGTGCTGTGTCCACTCAGAGTTGAATGG		
GH_P_E20	(501)	GTGCTCTCTGAAACATGTGCTGTGTCCACTCAGAGTTGAATGG		
GH_P_E19 rc	(501)	GTGCTCTCTGAAACATGTGCTGTGTCCACTCAGAGTTGAATGG		

Nukleotidalignment aus dem gag-Bereich, Klone - Partikeln GH

		1		50
GH_P_1	(1)	GCAGCCCTATTTCTTCGACCTGTTCTTGTACCCCATCAATCCACCAAGT		
GH_P_2	(1)	-----GCCATCAGAGTCTAAACCA-CGAGGGAC-AAGT		
GH_P_9 rc	(1)	-----GCCATCAGAGTCTAAACCA-CGAGGGAC-AAGT		
GH_P_3 rc	(1)	-----GCCATCAGAGTCTAAACCA-CGAGGGAC-AAGT		
GH_P_4	(1)	-----GCCATCAGAGTCTAAACCA-CGAGGGAC-AAGT		
GH_P_8 rc	(1)	-----GCCATCAGAGTCTAAACCA-CGAGGGAC-AAGT		
GH_P_6 rc	(1)	-----GCCATCAGAGTCTAAACCA-CGAGGGAC-AAGT		
GH_P_12	(1)	GCAGCCCTATTTCTTCGACCTGTTCTTGTACCCCATCAATCCACCAAGT		
GH_P_9	(1)	GCAGCCCTATTTCTTCGACCTGTTCTTGTACCCCATCAATCCACCAAGT		
GH_P_3	(1)	GCAGCCCTATTTCTTCGACCTGTTCTTGTACCCCATCAATCCACCAAGT		
GH_P_6	(1)	GCAGCCCTATTTCTTCGACCTGTTCTTGTACCCCATCAATCCACCAAGT		
GH_P_11	(1)	GCAGCCCTATTTCTTCGACCTGTTCTTGTACCCCATCAATCCACCAAGT		
GH_P_13	(1)	GCAGCCCTATTTCTTCGACCTGTTCTTGTACCCCATCAATCCACCAAGT		
GH_P_14	(1)	GCAGCCCTATTTCTTCGACCTGTTCTTGTACCCCATCAATCCACCAAGT		
GH_P_5	(1)	GCAGCCCTATTTCTTCGACCTGTTCTTGTACCCCATCAATCCACCAAGT		
GH_P_7	(1)	GCAGCCCTATTTCTTCGACCTGTTCTTGTACCCCATCAATCCACCAAGT		
GH_P_8	(1)	GCAGCCCTATTTCTTCGACCTGTTCTTGTACCCCATCAATCCACCAAGT		

51 100

GH_P_1 (51) CTTAAATTGTAAAATTGAGAGG GTGAGAGAGACGATT TTGCCAGAACT

GH_P_2 (32) CCTC--TTCCAGCAGGTCAG---GTGCCTGTAAC-ATT-----ACAACCT

GH_P_9 rc (33) CATC--TTCCAGCAGGTCAG---GTGCCCGTAAC-ATT-----ACAACCT

GH_P_3 rc (33) CCTC--TTCCAGCAGGTCAG---GTGCCCGTAAC-ATT-----ACAACCT

GH_P_4 (32) CCTC--TTCCAGCAGGTCAG---GTGCCCGTAAC-ATT-----ACAACCT

GH_P_8 rc (33) CCTC--TTCCAGCAGGTCAG---GTGCCCGTAAC-ATT-----ACAACCT

GH_P_6 rc (33) CCTC--TTCCAGCAGGTCAG---GTGCCCGTAAC-ATT-----ACAACCT

GH_P_12 (51) CTTAAATTGTAAAATTGAGAGG GTGAGAGAGACGATT TTGCCAGAACT

GH_P_9 (51) CTTAAATTGTAAAATTGAGAGG GTGAGAGAGACGATT TTGCCAGAACT

GH_P_3 (51) CTTAAATTGTAAAATTGAGAGG GTGAGAGAGACGATT TTGCCAGAACT

GH_P_6 (51) CTTAAATTGTAAAATTGAGAGG GTGAGAGAGACGATT TTGCCAGAACT

GH_P_11 (51) CTTAAATTGTAAAATTGAGAGG GTGAGAGAGACGATT TTGCCAGAACT

GH_P_13 (51) CTTAAATTGTAAAATTGAGAGG GTGAGAGAGACGATT TTGCCAGAACT

GH_P_14 (51) CTTAAATTGTAAAATTGAGAGG GTGAGAGAGACGATT TTGCCAGAACT

GH_P_5 (51) CTTAAATTGTAAAATTGAGAGG GTGAGAGAGACGATT TTGCCAGAACT

GH_P_7 (51) CTTAAATTGTAAAATTGAGAGG GTGAGAGAGACGATT TTGCCAGAACT

GH_P_8 (51) CTTAAATTGTAAAATTGAGAGG GTGAGAGAGACGATT TTGCCAGAACT

101 150

GH_P_1 (101) CCCAATCATAAGGAATGAGTCTA TGTCCATAAGCAATGGAATCTAATAAT

GH_P_2 (71) CAAAAGCAGGTT-AAAGAAAATAAGACC-----CAACCG---CCAGTAGC

GH_P_9 rc (72) CAAAAGCAGGTT-AAAGAAAATAAGACC-----CAACCG---CCAGTAGC

GH_P_3 rc (72) CAAAAGCAGGTT-AAAGAAAATAAGACC-----CAACCG---CCAGTAGC

GH_P_4 (71) CAAAAGCAGGTT-AAAGAAAATAAGACC-----CAACCG---CCAGTAGC

GH_P_8 rc (72) CAAAAGCAGGTT-AAAGAAAATAAGACC-----CAACCG---CCAGTAGC

GH_P_6 rc (72) CAAAAGCAGGTT-AAAGAAAATAAGACC-----CAACCG---CCAGTAGC

GH_P_12 (101) CCCAATCATAAGGAATGAGTCTA TGTCCATAAGCAATGGAATCTAATAAT

GH_P_9 (101) CCCAATCATAAGGAATGAGTCTA TGTCCATAAGCAATGGAATCTAATAAT

GH_P_3 (101) CCCAATCATAAGGAATGAGTCTA TGTCCATAAGCAATGGAATCTAATAAT

GH_P_6 (101) CCCAATCATAAGGAATGAGTCTA TGTCCATAAGCAATGGAATCTAATAAT

GH_P_11 (101) CCCAATCATAAGGAATGAGTCTA TGTCCATAAGCAATGGAATCTAATAAT

GH_P_13 (101) CCCAATCATAAGGAATGAGTCTA TGTCCATAAGCAATGGAATCTAATAAT

GH_P_14 (101) CCCAATCATAAGGAATGAGTCTA TGTCCATAAGCAATGGAATCTAATAAT

GH_P_5 (101) CCCAATCATAAGGAATGAGTCTA TGTCCATAAGCAATGGAATCTAATAAT

GH_P_7 (101) CCCAATCATAAGGAATGAGTCTA TGTCCATAAGCAATGGAATCTAATAAT

GH_P_8 (101) CCCAATCATAAGGAATGAGTCTA TGTCCATAAGCAATGGAATCTAATAAT

151 200

GH_P_1 (151) GTCCTCATATAAGGGGAGTTGGGTCCA-TACTGTTT TACTCCCTCTTTCA

GH_P_2 (112) CTATCAAT--ACTGGCCTCCGGCTGAACTTCAGTAT--CGGCCACCCCA

GH_P_9 rc (113) CTATCAAT--ACTGGCCTCCGGCTGAACTTCAGTAT--CGGCCACCCCA

GH_P_3 rc (113) CTATCAAT--ACTGGCCTCCGGCTGAACTTCAGTAT--CGGCCACCCCA

GH_P_4 (112) CTATCAAT--ACTGGCCTCCGGCTGAACTTCAGTAT--CGGCCACCCCA

GH_P_8 rc (113) CTATCAAT--ACTGGCCTCCGGCTGAACTTCAGTAT--CGGCCACCCCA

GH_P_6 rc (113) CTATCAAT--ACTGGCCTCCGGCTGAACTTCAGTAT--CGGCCACCCCA

GH_P_12 (151) GTCCTCATATAAGGGGAGTTGGGTCCA-TACTGTTT TACTCCCTCTTTCA

GH_P_9 (151) GTCCTCATATAAGGGGAGTTGGGTCCA-TACTGTTT TACTCCCTCTTTCA

GH_P_3 (151) GTCCTCATATAAGGGGAGTTGGGTCCA-TACTGTTT TACTCCCTCTTTCA

GH_P_6 (151) GTCCTCATATAAGGGGAGTTGGGTCCA-TACTGTTT TACTCCCTCTTTCA

GH_P_11 (151) GTCCTCATATAAGGGGAGTTGGGTCCA-TACTGTTT TACTCCCTCTTTCA

GH_P_13 (151) GTCCTCATATAAGGGGAGTTGGGTCCA-TACTGTTT TACTCCCTCTTTCA

GH_P_14 (151) GTCCTCATATAAGGGGAGTTGGGTCCA-TACTGTTT TACTCCCTCTTTCA

GH_P_5 (151) GTCCTCATATAAGGGGAGTTGGGTCCA-TACTGTTT TACTCCCTCTTTCA

GH_P_7 (151) GTCCTCATATAAGGGGAGTTGGGTCCA-TACTGTTT TACTCCCTCTTTCA

GH_P_8 (151) GTCCTCATATAAGGGGAGTTGGGTCCA-TACTGTTT TACTCCCTCTTTCA

201 250

GH_P_1 (200) TATCTTTTAGCTTTTATCGAAAAAGACTTGTATCTGGCCTCAACTGTG
 GH_P_2 (158) GAAAGTC-AGTATGGATATCCAGGAATGCCCCAGCA--CCACAGGGCAG
 GH_P_9 rc (159) GAAAGTC-AGTATGGATATCCAGGAATGCCCCAGCA--CCACAGGGCAG
 GH_P_3 rc (159) GAAAGTC-AGTATGGATATCCAGGAATGCCCCAGCA--CCACAGGGCAG
 GH_P_4 (158) GAAAGTC-AGTATGGATATCCAGGAATGCCCCAGCA--CCACAGGGCAG
 GH_P_8 rc (159) GAAAGTC-AGTATGGATATCCAGGAATGCCCCAGCA--CCACAGGGCAG
 GH_P_6 rc (159) GAAAGTC-AGTATGGATATCCAGGAATGCCCCAGCA--CCACAGGGCAG
 GH_P_12 (200) TATCTTTTAGCATTTTATCGAAAAAGACTTGTATCTGGCCTCAACTGTG
 GH_P_9 (200) TATCTTTTAGCATTTTATCGAAAAAGACTTGTATCTGGCCTCAACTGTG
 GH_P_3 (200) TATCTTTTAGCATTTTATCGAAAAAGACTTGTATCTGGCCTCAACTGTG
 GH_P_6 (200) TATCTTTTAGCATTTTATCGAAAAAGACTTGTATCTGGCCTCAACTGTG
 GH_P_11 (200) TATCTTTTAGCATTTTATCGAAAAAGACTTGTATCTGGCCTCAACTGTG
 GH_P_13 (200) TATCTTTTAGCATTTTATCGAAAAAGACTTGTATCTGGCCTCAACTGTG
 GH_P_14 (200) TATCTTTTAGCATTTTATCGAAAAAGACTTGTATCTGGCCTCAACTGTG
 GH_P_5 (200) TATCTTTTAGCATTTTATCGAAAAAGACTTGTATCTGGCCTCAACTGTG
 GH_P_7 (200) TATCTTTTAGCATTTTATCGAAAAAGACTTGTATCTGGCCTCAACTGTG
 GH_P_8 (200) TATCTTTTAGCATTTTATCGAAAAAGACTTGTATCTGGCCTCAACTGTG

251 300

GH_P_1 (250) GGAGGCTCTCCCTCTTGGGCTCCTTCTCCAGGTGGCATCGGTTCTAACGT
 GH_P_2 (205) GCGCCATACCCTC--AGCCGCCACT--AGGAGAC-TTAATCCTACGGC
 GH_P_9 rc (206) GCGCCATACCCTC--AGCCGCCACT--AGGAGAC-TTAATCCTACGGC
 GH_P_3 rc (206) GCGCCATACCCTC--AGCCGCCACT--AGGAGAC-TTAATCCTACGGC
 GH_P_4 (205) GCGCCATACCCTC--AGCCGCCACT--AGGAGAC-TTAATCCTACGGC
 GH_P_8 rc (206) GCGCCATACCCTC--AGCCGCCACT--AGGAGAC-TTAATCCTACGGC
 GH_P_6 rc (206) GCGCCATACCCTC--AGCCGCCACT--AGGAGAC-TTAATCCTACGGC
 GH_P_12 (250) GGAGGCTCTCCCTCTTGGGCTCCTTCTCCAGGTGGCATCGGTTCTAACGT
 GH_P_9 (250) GGAGGCTCTCCCTCTTGGGCTCCTTCTCCAGGTGGCATCGGTTCTAACGT
 GH_P_3 (250) GGAGGCTCTCCCTCTTGGGCTCCTTCTCCAGGTGGCATCGGTTCTAACGT
 GH_P_6 (250) GGAGGCTCTCCCTCTTGGGCTCCTTCTCCAGGTGGCATCGGTTCTAACGT
 GH_P_11 (250) GGAGGCTCTCCCTCTTGGGCTCCTTCTCCAGGTGGCATCGGTTCTAACGT
 GH_P_13 (250) GGAGGCTCTCCCTCTTGGGCTCCTTCTCCAGGTGGCATCGGTTCTAACGT
 GH_P_14 (250) GGAGGCTCTCCCTCTTGGGCTCCTTCTCCAGGTGGCATCGGTTCTAACGT
 GH_P_5 (250) GGAGGCTCTCCCTCTTGGGCTCCTTCTCCAGGTGGCATCGGTTCTAACGT
 GH_P_7 (250) GGAGGCTCTCCCTCTTGGGCTCCTTCTCCAGGTGGCATCGGTTCTAACGT
 GH_P_8 (250) GGAGGCTCTCCCTCTTGGGCTCCTTCTCCAGGTGGCATCGGTTCTAACGT

301 350

GH_P_1 (300) TACTGGGAATTGCCATGCCTCAGTATCTCCTTCCTTTCTTGATTTATCAA
 GH_P_2 (250) ACCACCTAGTAGACAGGGTAGTGAATTACATGAAATATTGATAAATCAA
 GH_P_9 rc (251) ACCACCTAGTAGACAGGGTAGTGAATTACATGAAATATTGATAAATCAA
 GH_P_3 rc (251) ACCACCTAGTAGACAGGGTAGTGAATTACATGAAATATTGATAAATCAA
 GH_P_4 (250) ACCACCTAGTAGACAGGGTAGTGAATTACATGAAATATTGATAAATCAA
 GH_P_8 rc (251) ACCACCTAGTAGACAGGGTAGTGAATTACATGAAATATTGATAAATCAA
 GH_P_6 rc (251) ACCACCTAGTAGACAGGGTAGTGAATTACATGAAATATTGATAAATCAA
 GH_P_12 (300) TACTGGGAATTGCCATGCCTCAGTATCTCCTTCCTTTCTTGATTTATCAA
 GH_P_9 (300) TACTGGGAATTGCCATGCCTCAGTATCTCCTTCCTTTCTTGATTTATCAA
 GH_P_3 (300) TACTGGGAATTGCCATGCCTCAGTATCTCCTTCCTTTCTTGATTTATCAA
 GH_P_6 (300) TACTGGGAATTGCCATGCCTCAGTATCTCCTTCCTTTCTTGATTTATCAA
 GH_P_11 (300) TACTGGGAATTGCCATGCCTCAGTATCTCCTTCCTTTCTTGATTTATCAA
 GH_P_13 (300) TACTGGGAATTGCCATGCCTCAGTATCTCCTTCCTTTCTTGATTTATCAA
 GH_P_14 (300) TACTGGGAATTGCCATGCCTCAGTATCTCCTTCCTTTCTTGATTTATCAA
 GH_P_5 (300) TACTGGGAATTGCCATGCCTCAGTATCTCCTTCCTTTCTTGATTTATCAA
 GH_P_7 (300) TACTGGGAATTGCCATGCCTCAGTATCTCCTTCCTTTCTTGATTTATCAA
 GH_P_8 (300) TACTGGGAATTGCCATGCCTCAGTATCTCCTTCCTTTCTTGATTTATCAA

```

351                                     400
GH_P_1 (350) TAAATTCATGTAATTTACTACCCTGTCTACTAGGTGGTGCCGTAGGATTA
GH_P_2 (300) GAAAGGAAGGAGATACTGAGGCATGGCAATCCCAGTAACGTTAGAACCG
GH_P_9 rc (301) GAAAGGAAGGAGATACTGAGGCATGGCAATCCCAGTAACGTTAGAACCG
GH_P_3 rc (301) GAAAGGAAGGAGATACTGAGGCATGGCAATCCCAGTAACGTTAGAACCG
GH_P_4 (300) GAAAGGAAGGAGATACTGAGGCATGGCAATCCCAGTAACGTTAGAACCG
GH_P_8 rc (301) GAAAGGAAGGAGATACTGAGGCATGGCAATCCCAGTAACGTTAGAACCG
GH_P_6 rc (301) GAAAGGAAGGAGATACTGAGGCATGGCAATCCCAGTAACGTTAGAACCG
GH_P_12 (350) TAAATTCATGTAATTTACTACCCTGTCTACTAGGTGGTGCCGTAGGATTA
GH_P_9 (350) TAAATTCATGTAATTTACTACCCTGTCTACTAGGTGGTGCCGTAGGATTA
GH_P_3 (350) TAAATTCATGTAATTTACTACCCTGTCTACTAGGTGGTGCCGTAGGATTA
GH_P_6 (350) TAAATTCATGTAATTTACTACCCTGTCTACTAGGTGGTGCCGTAGGATTA
GH_P_11 (350) TAAATTCATGTAATTTACTACCCTGTCTACTAGGTGGTGCCGTAGGATTA
GH_P_13 (350) TAAATTCATGTAATTTACTACCCTGTCTACTAGGTGGTGCCGTAGGATTA
GH_P_14 (350) TAAATTCATGTAATTTACTACCCTGTCTACTAGGTGGTGCCGTAGGATTA
GH_P_5 (350) TAAATTCATGTAATTTACTACCCTGTCTACTAGGTGGTGCCGTAGGATTA
GH_P_7 (350) TAAATTCATGTAATTTACTACCCTGTCTACTAGGTGGTGCCGTAGGATTA
GH_P_8 (350) TAAATTCATGTAATTTACTACCCTGTCTACTAGGTGGTGCCGTAGGATTA

```

```

401                                     450
GH_P_1 (400) A-GTCTCCTAGTGG--GGCGGCT--GAGAGTATGGCCGCCCTGCCCTGTGGT
GH_P_2 (350) ATGCCACCTGGAGAAGGAGCCCAAGAGGGAGAGCCTCCCACAGTTGAGGC
GH_P_9 rc (351) ATGCCACCTGGAGAAGGAGCCCAAGAGGGAGAGCCTCCCACAGTTGAGGC
GH_P_3 rc (351) ATGCCACCTGGAGAAGGAGCCCAAGAGGGAGAGCCTCCCACAGTTGAGGC
GH_P_4 (350) ATGCCACCTGGAGAAGGAGCCCAAGAGGGAGAGCCTCCCACAGTTGAGGC
GH_P_8 rc (351) ATGCCACCTGGAGAAGGAGCCCAAGAGGGAGAGCCTCCCACAGTTGAGGC
GH_P_6 rc (351) ATGCCACCTGGAGAAGGAGCCCAAGAGGGAGAGCCTCCCACAGTTGAGGC
GH_P_12 (400) A-GTCTCCTAGTGG--GGCGGCT--GAGGGTATGGCCGCCCTGCCCTGTGGT
GH_P_9 (400) A-GTCTCCTAGTGG--GGCGGCT--GAGGGTATGGCCGCCCTGCCCTGTGGT
GH_P_3 (400) A-GTCTCCTAGTGG--GGCGGCT--GAGGGTATGGCCGCCCTGCCCTGTGGT
GH_P_6 (400) A-GTCTCCTAGTGG--GGCGGCT--GAGGGTATGGCCGCCCTGCCCTGTGGT
GH_P_11 (400) A-GTCTCCTAGTGG--GGCGGCT--GAGGGTATGGCCGCCCTGCCCTGTGGT
GH_P_13 (400) A-GTCTCCTAGTGG--GGCGGCT--GAGGGTATGGCCGCCCTGCCCTGTGGT
GH_P_14 (400) A-GTCTCCTAGTGG--GGCGGCT--GAGGGTATGGCCGCCCTGCCCTGTGGT
GH_P_5 (400) A-GTCTCCTAGTGG--GGCGGCT--GAGGGTATGGCCGCCCTGCCCTGTGGT
GH_P_7 (400) A-GTCTCCTAGTGG--GGCGGCT--GAGGGTATGGCCGCCCTGCCCTGTGGT
GH_P_8 (400) A-GTCTCCTAGTGG--GGCGGCT--GAGGGTATGGCCGCCCTGCCCTGTGGT

```

```

451                                     500
GH_P_1 (445) --GCTGGGGGCATTCTCGGATAATCCATACTGACTT-TCTGGGGGTGGCCG
GH_P_2 (400) CAGATACAAGTCTTTTTCGATAAAAATGCTAAAAGATATGAAAGAGGGAG
GH_P_9 rc (401) CAGATACAAGTCTTTTTCGATAAAAATGCTAAAAGATATGAAAGAGGGAG
GH_P_3 rc (401) CAGATACAAGTCTTTTTCGATAAAAATGCTAAAAGATATGAAAGAGGGAG
GH_P_4 (400) CAGATACAAGTCTTTTTCGATAAAAATGCTAAAAGATATGAAAGAGGGAG
GH_P_8 rc (401) CAGATACAAGTCTTTTTCGATAAAAATGCTAAAAGATATGAAAGAGGGAG
GH_P_6 rc (401) CAGATACAAGTCTTTTTCGATAAAAATGCTAAAAGATATGAAAGAGGGAG
GH_P_12 (445) --GCTGGGGGCATTCTCGGATAATCCATACTGACTT-TCTGGGGGTGGCCG
GH_P_9 (445) --GCTGGGGGCATTCTCGGATAATCCATACTGACTT-TCTGGGGGTGGCCG
GH_P_3 (445) --GCTGGGGGCATTCTCGGATAATCCATACTGACTT-TCTGGGGGTGGCCG
GH_P_6 (445) --GCTGGGGGCATTCTCGGATAATCCATACTGACTT-TCTGGGGGTGGCCG
GH_P_11 (445) --GCTGGGGGCATTCTCGGATAATCCATACTGACTT-TCTGGGGGTGGCCG
GH_P_13 (445) --GCTGGGGGCATTCTCGGATAATCCATACTGACTT-TCTGGGGGTGGCCG
GH_P_14 (445) --GCTGGGGGCATTCTCGGATAATCCATACTGACTT-TCTGGGGGTGGCCG
GH_P_5 (445) --GCTGGGGGCATTCTCGGATAATCCATACTGACTT-TCTGGGGGTGGCCG
GH_P_7 (445) --GCTGGGGGCATTCTCGGATAATCCATACTGACTT-TCTGGGGGTGGCCG
GH_P_8 (445) --GCTGGGGGCATTCTCGGATAATCCATACTGACTT-TCTGGGGGTGGCCG

```



```

501
GH_P_1 (492) ATACTGAAGTTCAGCCGGAGGCCAGT-----ATTGATAGGCTAC
GH_P_2 (450) TAAAACAGTATGGACCCAACCTCCCTTATATGAGGACATTATTAGATTCC
GH_P_9 rc (451) TAAAACAGTATGGACCCAACCTCCCTTATATGAGGACATTATTAGATTCC
GH_P_3 rc (451) TAAAACAGTATGGACCCAACCTCCCTTATATGAGGACATTATTAGATTCC
GH_P_4 (450) TAAAACAGTATGGACCCAACCTCCCTTATATGAGGACATTATTAGATTCC
GH_P_8 rc (451) TAAAACAGTATGGACCCAACCTCCCTTATATGAGGACATTATTAGATTCC
GH_P_6 rc (451) TAAAACAGTATGGACCCAACCTCCCTTATATGGGGACATTATTAGATTCC
GH_P_12 (492) ATACTGAAGTTCAGCCGGAGGCCAGT-----ATTGATAGGCTAC
GH_P_9 (492) ATACTGAAGTTCAGCCGGAGGCCAGT-----ATTGATAGGCTAC
GH_P_3 (492) ATACTGAAGTTCAGCCGGAGGCCAGT-----ATTGATAGGCTAC
GH_P_6 (492) ATACTGAAGTTCAGCCGGAGGCCAGT-----ATTGATAGGCTAC
GH_P_11 (492) ATACTGAAGTTCAGCCGGAGGCCAGT-----ATTGATAGGCTAC
GH_P_13 (492) ATACTGAAGTTCAGCCGGAGGCCAGT-----ATTGATAGGCTAC
GH_P_14 (492) ATACTGAAGTTCAGCCGGAGGCCAGT-----ATTGATAGGCTAC
GH_P_5 (492) ATACTGAAGTTCAGCCGGAGGCCAGT-----ATTGATAGGCTAC
GH_P_7 (492) ATACTGAAGTTCAGCCGGAGGCCAGT-----ATTGATAGGCTAC
GH_P_8 (492) ATACTGAAGTTCAGCCGGAGGCCAGT-----ATTGATAGGCTAC

```

```

551
GH_P_1 (531) -TGGCGGTTGGGTCTTATTTTCTTTAACCTGCTTTTGGAGTTGT-----A
GH_P_2 (500) ATTGCTTATGGACATAGACTCATTCCTTATGATTGGGAGATTCTGGCAAA
GH_P_9 rc (501) ATTGCTTATGGACATAGACTCATTCCTTATGATTGGGAGATTCTGGCAAA
GH_P_3 rc (501) ATTGCTTATGGACATAGACTCATTCCTTATGATTGGGAGATTCTGGCAAA
GH_P_4 (500) ATTGCTCATGGACATAGACTCATTCCTTATGATTGGGAGATTCTGGCAAA
GH_P_8 rc (501) ATTGCTTATGGACATAGACTCATTCCTTATGATTGGGAGATTCTGGCAAA
GH_P_6 rc (501) ATTGCTTATGGACATAGACTCATTCCTTATGATTGGGAGATTCTGGCAAA
GH_P_12 (531) -TGGCGGTTGGGTCTTATTTTCTTTAACCTGCTTTTGGAGTTGT-----A
GH_P_9 (531) -TGGCGGTTGGGTCTTATTTTCTTTAACCTGCTTTTGGAGTTGT-----A
GH_P_3 (531) -TGGCGGTTGGGTCTTATTTTCTTTAACCTGCTTTTGGAGTTGT-----A
GH_P_6 (531) -TGGCGGTTGGGTCTTATTTTCTTTAACCTGCTTTTGGAGTTGT-----A
GH_P_11 (531) -TGGCGGTTGGGTCTTATTTTCTTTAACCTGCTTTTGGAGTTGT-----A
GH_P_13 (531) -TGGCGGTTGGGTCTTATTTTCTTTAACCTGCTTTTGGAGTTGT-----A
GH_P_14 (531) -TGGCGGTTGGGTCTTATTTTCTTTAACCTGCTTTTGGAGTTGT-----A
GH_P_5 (531) -TGGCGGTTGGGTCTTATTTTCTTTAACCTGCTTTTGGAGTTGT-----A
GH_P_7 (531) -TGGCGGTTGGGTCTTATTTTCTTTAACCTGCTTTTGGAGTTGT-----A
GH_P_8 (531) -TGGCGGTTGGGTCTTATTTTCTTTAACCTGCTTTTGGAGTTGT-----A

```

```

601
GH_P_1 (575) AT-GTTACGGGCACC---TGACCTGCTGGAAG--AGGACTTG-TGCCTCG
GH_P_2 (550) ATCGTCTCTCTCACCCCTCTCAATTTTACAATTTAAGACTTGGTGATTG
GH_P_9 rc (551) ATCGTCTCTCTCACCCCTCTCAATTTTACAATTTAAGACTTGGTGATTG
GH_P_3 rc (551) ATCGTCTCTCTCACCCCTCTCAATTTTACAATTTAAGACTTGGTGATTG
GH_P_4 (550) ATCGTCTCTCTCACCCCTCTCAATTTTACAATTTAAGACTTGGTGATTG
GH_P_8 rc (551) ATCGTCTCTCTCACCCCTCTCAATTTTACAATTTAAGACTTGGTGATTG
GH_P_6 rc (551) ATCGTCTCTCTCACCCCTCTCAATTTTACAATTTAAGACTTGGTGATTG
GH_P_12 (575) AT-GTTACGGGCACC---TGACCTGCTGGAAG--ATGACTTG-TGCCTCG
GH_P_9 (575) AT-GTTACGGGCACC---TGACCTGCTGGAAG--ATGACTTG-TGCCTCG
GH_P_3 (575) AT-GTTACGGGCACC---TGACCTGCTGGAAG--AGGACTTG-TCCCTCG
GH_P_6 (575) AT-GTTACGGGCACC---TGACCTGCTGGAAG--AGGACTTG-TCCCTCG
GH_P_11 (575) AT-GTTACGGGCACC---TGACCTGCTGGAAG--AGGACTTG-TCCCTCG
GH_P_13 (575) AT-GTTACGGGCACC---TGACCTGCTGGAAG--AGGACTTG-TCCCTCG
GH_P_14 (575) AT-GTTACGGGCACC---TGACCTGCTGGAAG--AGGACTTG-TCCCTCG
GH_P_5 (575) AT-GTTACGGGCACC---TGACCTGCTGGAAG--AGGACTTG-TCCCTCG
GH_P_7 (575) AT-GTTACGGGCACC---TGACCTGCTGGAAG--AGGACTTG-TCCCTCG
GH_P_8 (575) AT-GTTACGGGCACC---TGACCTGCTGGAAG--AGGACTTG-TCCCTCG

```

		651		688
	GH_P_1	(618)	-TGGTTTACTCTGATGGCC-----	
	GH_P_2	(600)	ATGGGGTACAAGAACAGGTCCGAAGAAATAGGGCTGCA	
	GH_P_9 rc	(601)	ATGGGGTACAAGAACAGGTCCGAAGAAATAGGGCTGC-	
	GH_P_3 rc	(601)	ATGGGGTACAAGAACAGGTCCGAAGAAATAGGGCTGC-	
	GH_P_4	(600)	ATGGGGTACAAGAACAGGTCCGAAGAAATAGGGCTGCA	
	GH_P_8 rc	(601)	ATGGGGTACAAGAACAGGTCCGAAGAAATAGGGCTGC-	
	GH_P_6 rc	(601)	ATGGGGTACAAGAACAGGTCCGAAGAAATAGGGCTGC-	
	GH_P_12	(618)	-TGGTTTACTCTGATGGCC-----	
	GH_P_9	(618)	-TGGTTTACTCTGATGGCC-----	
	GH_P_3	(618)	-TGGTTTACTCTGATGGCC-----	
	GH_P_6	(618)	-TGGTTTACTCTGATGGCC-----	
	GH_P_11	(618)	-TGGTTTACTCTGATGGCC-----	
	GH_P_13	(618)	-TGGTTTACTCTGATGGCC-----	
	GH_P_14	(618)	-TGGTTTACTCTGATGGCC-----	
	GH_P_5	(618)	-TGGTTTACTCTGATGGCC-----	
	GH_P_7	(618)	-TGGTTTACTCTGATGGCC-----	
	GH_P_8	(618)	-TGGTTTACTCTGATGGCC-----	

Nukleotidalignment aus dem *env*-Bereich, Klone - Partikeln UKRV Mel 2-C9

		1		50
	C9_P_E2 rc	(1)	AATGAGTCTGAGCATCACTGGGACATGGTTAGACGCCATCTACAGGGAAG	
	UKRV Mel2_P_E3 rc	(1)	AATGAGTCTGAGCATCACTGGGACATGGTTAGACGCCATCTACAGGGAAG	
	C9_P_E4	(1)	AATGACTCTGAGCATCGCTGGGACATGGTTAGACGCCATCTACAGGGAAG	
	C9_P_E5 rc	(1)	AATGAGTCTGAGCATCACTGGGACATGGTTAGACGCCATCTACAGGGAAG	
	C9_P_E6	(1)	AATGAGTCTGAGCATCACTGGGACATGGTTAGACGCCATCTACAGGGAAG	
	UKRV Mel 2_P_E7 rc	(1)	AATGAGTCTGAGCATCACTGGGACATGGTTAGACGCCATCTACAGGGAAG	
	C9_P_E8	(1)	AATGAGTCTGAGCATCACTGGGACATGGTTAGACGCCATCTACAGGGAAG	
	C9_P_E9	(1)	AATGAGTCTGAGCATCACTGGGACATGGTTAGACGCCATCTACAGGGAAG	
	UKRV Mel 2_P_E10	(1)	AATGAGTCTGAGCATCACTGGGACATGGTTAGACGCCATCTACAGGGAAG	
	C9_P_E11	(1)	AATGAGTCTGAGCATCACTGGGACATGGTTAGACGCCATCTACAGGGAAG	
	C9_P_E12	(1)	AATGACTCTGAGCATCGGAGGGACATGGCAAGACGCCATCGACCGGGAAG	
	C9_P_E13 rc	(1)	AATGAGTTTTAGCATCACTGGGACCTGGTTTACGCCATTACAGGGAAG	
	UKRV Mel 2_P_E14 rc	(1)	AATGAGTCTGAGCATCACTGGGACATGGTTAGACGCCATCTACAGGGAAG	
	C9_P_E15 rc	(1)	AATGAGTCTGAGCATCACTGGGACATGGTTAGACGCCATCTACAGGGAAG	
	UKRV Mel 2_P_E16	(1)	AATGAGTCTGAGCATCACTGGGACATGGTTAGACGCCATCTACAGGGAAG	
	UKRV Mel 2_P_E17 rc	(1)	AATGAGTTTTAGCATCACTGGGACATGGTTAGACGCCATCTACAGGGAAG	
	UKRV Mel 2_P_E19 rc	(1)	AATGAGTCTGAGCATCACTGGGACATGGTTAGACGCCATCTACAGGGAAG	
	UKRV Mel 2_P_E20	(1)	AATGAGTCTGAGCATCACTGGGACATGGTTAGACGCCATCTACAGGGAAG	
	UKRV Mel 2_P_E21 rc	(1)	AATGAGTCTGAGCATCACTGGGACATGGTTAGACGCCATCTACAGGGAAG	
	C9_P_E1	(1)	AATGAGTCTGAGCATCACTGGGACATGGTTAGACGCCATCTACAGGGAAG	
	UKRV Mel 2_P_E18 rc	(1)	AATGAGTCTGAGCATCACTGGGACATGGTTAGACGCCATCTACAGGGAAG	
		51		100
	C9_P_E2 rc	(51)	AGAAGATAATCTCACTTTAGACATTTCCAAATTAAGAAACAAATTTTCG	
	UKRV Mel2_P_E3 rc	(51)	AGAAGATAATCTCACTTTAGACATTTCCAAATTAAGAAACAAATTTTCG	
	C9_P_E4	(51)	AGAGTATGGTCTCACTTTAAACATTTCCAAATTAAGAAACAAATTTTC	
	C9_P_E5 rc	(51)	AGAAGATAATCTCACTTTAGACATTTCCAAATTAAGAAACAAATTTTCG	
	C9_P_E6	(51)	AGAAGATAATCTCACTTTAGACATTTCCAAATTAAGAAACAAATTTTCG	
	UKRV Mel 2_P_E7 rc	(51)	AGAAGATAATCTCACTTTAGACATTTCCAAATTAAGAAACAAATTTTCG	
	C9_P_E8	(51)	AGAAGATAATCTCACTTTAGACATTTCCAAATTAAGAAACAAATTTTCG	
	C9_P_E9	(51)	AGAAGATAATCTCACTTTAGACATTTCCAAATTAAGAAACAAATTTTCG	
	UKRV Mel 2_P_E10	(51)	AGAAGATAATCTCACTTTAGACATTTCCAAATTAAGAAACAAATTTTCG	
	C9_P_E11	(51)	AGAAGATAATCTCACTTTAGACATTTCCAAATTAAGAAACAAATTTTCG	
	C9_P_E12	(51)	AGAAGATAATCTCCCTTTGCACCTTTGCAAATTAAGAAACAAATTTTC	
	C9_P_E13 rc	(51)	AGAAGATAATCTCACTTTAGACATTTCCAAATTAAGAAACAAATTTTCG	
	UKRV Mel 2_P_E14 rc	(51)	AGAAGATAATCTCACTTTAGACATTTCCAAATTAAGAAACAAATTTTCG	
	C9_P_E15 rc	(51)	AGAAGATAATCTCACTTTAGACATTTCCAAATTAAGAAACAAATTTTCG	
	UKRV Mel 2_P_E16	(51)	AGAAGATAATCTCACTTTAGACATTTCCAAATTAAGAAACAAATTTTCG	
	UKRV Mel 2_P_E17 rc	(51)	AGAAGATAATCTCACTTTAGACATTTCCAAATTAAGAAACAAATTTTCG	
	UKRV Mel 2_P_E19 rc	(51)	AGAAGATAATCTCACTTTAGACATTTCCAAATTAAGAAACAAATTTTCG	
	UKRV Mel 2_P_E20	(51)	AGAAGATAATCTCACTTTAGACATTTCCAAATTAAGAAACAAATTTTCG	
	UKRV Mel 2_P_E21 rc	(51)	AGAAGATAATCTCACTTTAGACATTTCCAAATTAAGAAACAAATTTTCG	
	C9_P_E1	(51)	AGAAGATAATCTCACTTTAGACATTTCCAAATTAAGAAACAAATTTTCG	
	UKRV Mel 2_P_E18 rc	(51)	AGAAGATAATCTCACTTTAGACATTTCCAAATTAAGAAACAAATTTTCG	

	101	150
C9_P_E2 rc	(101)	AAGCATCAAAAAGCCCATTTAAATTTGGTGCCAGGAACTGAGGCAATTGCA
UKRV Mel2_P_E3 rc	(101)	AAGCATCAAAAAGCCCATTTAAATTTGGTGCCAGGAACTGAGGCAATTGCA
C9_P_E4	(101)	AAGCATCAAAAAGCCCATTTAAATTTGGTGCCCTTAACCTGAGGCAATTGCC
C9_P_E5 rc	(101)	AAGCATCAAAAAGCCCATTTAAATTTGGTGCCAGGAACTGAGGCAATTGCA
C9_P_E6	(101)	AAGCATCAAAAAGCCCATTTAAATTTGGTGCCAGGAACTGAGGCAATTGCA
UKRV Mel 2_P_E7 rc	(101)	AAGCATCAAAAAGCCCATTTAAATTTGGTGCCAGGAACTGAGGCAATTGCA
C9_P_E8	(101)	AAGCATCAAAAAGCCCATTTAAATTTGGTGCCAGGAACTGAGGCAATTGCA
C9_P_E9	(101)	AAGCATCAAAAAGCCCATTTAAATTTGGTGCCAGGAACTGAGGCAATTGCA
UKRV Mel 2_P_E10	(101)	AAGCATCAAAAAGCCCATTTAAATTTGGTGCCAGGAACTGAGGCAATTGCA
C9_P_E11	(101)	AAGCATCAAAAAGCCCATTTAAATTTGGTGCCAGGAACTGAGGCAATTGCA
C9_P_E12	(101)	AAGCCTCATAAGCCCATTTAAATTTGGTGCCAGGAACTGAGGCAATTGCA
C9_P_E13 rc	(101)	GAGCATCAAAAACCCCATTTAAATTTGGTGCCCGGAACTGAGGCAATTGCC
UKRV Mel 2_P_E14 rc	(101)	AAGCATCAAAAAGCCCATTTAAATTTGGTGCCAGGAACTGAGGCAATTGCA
C9_P_E15 rc	(101)	AAGCATCAAAAAGCCCATTTAAATTTGGTGCCAGGAACTGAGGCAATTGCA
UKRV Mel 2_P_E16	(101)	AAGCATCAAAAAGCCCATTTAAATTTGGTGCCAGGAACTGAGGCAATTGCA
UKRV Mel 2_P_E17 rc	(101)	AAGCATCAAAAAGCCCATTTAAATTTGGTGCCAGGAACTGAGGCAATTGCA
UKRV Mel 2_P_E19 rc	(101)	AAGCATCAAAAAGCCCATTTAAATTTGGTGCCAGGAACTGAGGCAATTGCA
UKRV Mel 2_P_E20	(101)	AAGCATCAAAAAGCCCATTTAAATTTGGTGCCCGGAACTGAGGCAATTGCA
UKRV Mel 2_P_E21 rc	(101)	AAGCATCAAAAAGCCCATTTAAATTTGGTGCCAGGAACTGAGGCAATTGCA
C9_P_E1	(101)	AAGCATCAAAAAGCCCATTTAAATTTGGTGCCAGGAACTGAGGCAATTGCA
UKRV Mel 2_P_E18 rc	(101)	AAGCATCAAAAAGCCCATTTAAATTTGGTGCCAGGAACTGAGGCAATTGCA

	151	200
C9_P_E2 rc	(151)	GGAGTTGCTGATGGCCTCGCAAATCTTAACCCTGTCACTTGGGTAAAGAC
UKRV Mel2_P_E3 rc	(151)	GGAGTTGCTGATGGCCTCGCAAATCTTAACCCTGTCACTTGGGTAAAGAC
C9_P_E4	(151)	GGAGTTGCTGATGGCCTCCAGATCTTAACCCTGTCACTTGGGTAAAGAC
C9_P_E5 rc	(151)	GGAGTTGCTGATGGCCTCGCAAATCTTAACCCTGTCACTTGGGTAAAGAC
C9_P_E6	(151)	GGAGTTGCTGATGGCCTCGCAAATCTTAACCCTGTCACTTGGGTAAAGAC
UKRV Mel 2_P_E7 rc	(151)	GGAGTTGCTGATGGCCTCGCAAATCTTAACCCTGTCACTTGGGTAAAGAC
C9_P_E8	(151)	GGAGTTGCTGATGGCCTCGCAAATCTTAACCCTGTCACTTGGGTAAAGAC
C9_P_E9	(151)	GGAGTTGCTGATGGCCTCGCAAATCTTAACCCTGTCACTTGGGTAAAGAC
UKRV Mel 2_P_E10	(151)	GGAGTTGCTGATGGCCTCGCAAATCTTAACCCTGTCACTTGGGTAAAGAC
C9_P_E11	(151)	GGAGTTGCTGATGGCCTCGCAAATCTTAACCCTGTCACTTGGGTAAAGAC
C9_P_E12	(151)	TGAGTTGCTGATGGCCTCGGAAATCTTAACCCTGTCTTGGGTAAAGAC
C9_P_E13 rc	(151)	GGAGTTGCTGATGGCCTCGCAAATTTAACCCTGTCACTTGGGTAAAGAC
UKRV Mel 2_P_E14 rc	(151)	GGAGTTGCTGATGGCCTCGCAAATCTTAACCCTGTCACTTGGGTAAAGAC
C9_P_E15 rc	(151)	GGAGTTGCTGATGGCCTCGCAAATCTTAACCCTGTCACTTGGGTAAAGAC
UKRV Mel 2_P_E16	(151)	GGAGTTGCTGATGGCCTCGCAAATCTTAACCCTGTCACTTGGGTAAAGAC
UKRV Mel 2_P_E17 rc	(151)	GGAGTTGCTGATGGCCTCGCAAATCTTAACCCTGTCACTTGGGTAAAGCC
UKRV Mel 2_P_E19 rc	(151)	GGAGTTGCTGATGGCCTCGCAAATCTTAACCCTGTCACTTGGGTAAAGAC
UKRV Mel 2_P_E20	(151)	GGAGTTGCTGATGGCCTCGCAAATCTTAACCCTGTTCCCGGTAAAGAC
UKRV Mel 2_P_E21 rc	(151)	GGAGTTGCTGATGGCCTCGCAAATCTTAACCCTGTCACTTGGGTAAAGAC
C9_P_E1	(151)	GGAGTTGCTGATGGCCTCGCAAATCTTAACCCTGTCACTTGGGTAAAGAC
UKRV Mel 2_P_E18 rc	(151)	GGAGTTGCTGATGGCCTCGCAAATCTTAACCCTGTCACTTGGGTAAAGAC

	201	250
C9_P_E2 rc	(201)	CATTGGAAGTACTACGATTATAAATCTCATATTAATCCTTGTGTGCCTGT
UKRV Mel2_P_E3 rc	(201)	CATTGGAAGTACTACGATTATAAATCTCATATTAATCCTTGTGTGCCTGT
C9_P_E4	(201)	CATTGGAAGTACTACGATTATAAATCTCATATTAATCCTTGTGTGCCTGT
C9_P_E5 rc	(201)	CATTGGAAGTACTACGATTATAAATCTCATATTAATCCTTGTGTGCCTGT
C9_P_E6	(201)	CATTGGAAGTACTACGATTATAAATCTCATATTAATCCTTGTGTGCCTGT
UKRV Mel 2_P_E7 rc	(201)	CATTGGAAGTACTACGATTATAAATCTCATATTAATCCTTGTGTGCCTGT
C9_P_E8	(201)	CATTGGAAGTACTACGATTATAAATCTCATATTAATCCTTGTGTGCCTGT
C9_P_E9	(201)	CATTGGAAGTACTACGATTATAAATCTCATATTAATCCTTGTGTGCCTGT
UKRV Mel 2_P_E10	(201)	CATTGGAAGTACTACGATTATAAATCTCATATTAATCCTTGTGTGCCTGT
C9_P_E11	(201)	CATTGGAAGTACTACGATTATAAATCTCATATTAATCCTTGTGTGCCTGT
C9_P_E12	(200)	CATTGGAAGTACTACGATTATAAATCTCATATTAATCCTTGTGTGCCTGT
C9_P_E13 rc	(201)	CCTTGAAGTACTACGATTATAAATCTCCTATTAATCCTTGTGGCCCTGT
UKRV Mel 2_P_E14 rc	(201)	CATTGGAAGTACTACGATTATAAATCTCATATTAATCCTTGTGTGCCTGT
C9_P_E15 rc	(201)	CATTGGAAGTACTACGATTATAAATCTCATATTAATCCTTGTGTGCCTGT
UKRV Mel 2_P_E16	(201)	CATTGGAAGTACTACGATTATAAATCTCATATTAATCCTTGTGTGCCTGT
UKRV Mel 2_P_E17 rc	(201)	CATTGGAAGTACTACGATTATAAATCTCATATTAATCCTTGTGTGCCTGT
UKRV Mel 2_P_E19 rc	(201)	CATTGGAAGTACTACGATTATAAATCTCATATTAATCCTTGTGTGCCTGT
UKRV Mel 2_P_E20	(201)	CATTGGAAGTACTACGATTATAAATCTCATATTAATCCTGGTGTGCCTGT
UKRV Mel 2_P_E21 rc	(201)	CATTGGAAGTACTACGATTATAAATCTCATATTAATCCTTGTGTGCCTGT
C9_P_E1	(201)	CATTGGAAGTACTACGATTATAAATCTCATATTAATCCTTGTGTGCCTGT
UKRV Mel 2_P_E18 rc	(201)	CATTGGAAGTACTACGATTATAAATCTCATATTAATCCTTGTGTGCCTGT

		251		300
	C9_P_E2 rc	(251)	TTTGTCTGTTGTTAGTCTGCAGGTGTACCCAACAGCTCCGAAGAGACAGC	
	UKRV Mel2_P_E3 rc	(251)	TTTGTCTGTTGTTAGTCTGCAGGTGTACCCAACAGCTCCGAAGAGACAGC	
	C9_P_E4	(251)	TTTGTCTGTTGTTAGTCTGCCGGGTACCCAACAGCTCCGAAGAGACAGC	
	C9_P_E5 rc	(251)	TTTGTCTGTTGTTAGTCTGCAGGTGTACCCAACAGCTCCGAAGAGACAGC	
	C9_P_E6	(251)	TTTGTCTGTTGTTAGTCTGCAGGTGTACCCAACAGCTCCGAAGAGACAGC	
	UKRV Mel 2_P_E7 rc	(251)	TTTGTCTGTTGTTAGTCTGCAGGTGTACCCAACAGCTCCGAAGAGACAGC	
	C9_P_E8	(251)	TTTGTCTGTTGTTAGTCTGCAGGTGTACCCAACAGCTCCGAAGAGACAGC	
	C9_P_E9	(251)	TTTGTCTGTTGTTAGTCTGCAGGTGTACCCAACAGCTCCGAAGAGACAGC	
	UKRV Mel 2_P_E10	(251)	TTTGTCTGTTGTTAGTCTGCAGGTGTACCCAACAGCTCCGAAGAGACAGC	
	C9_P_E11	(251)	TTTGTCTGTTGTTAGTCTGCAGGTGTACCCAACAGCTCCGAAGAGACAGC	
	C9_P_E12	(250)	TTTGTCTGTTGTTAGTCTGCAGGTGTACCCAACAGCTCCGAAGAGACAGC	
	C9_P_E13 rc	(251)	TTTGTCTGTTGTTAGTTTGCCTGGGGTACCCAACAGCTCCGAAGAGACAGC	
	UKRV Mel 2_P_E14 rc	(251)	TTTGTCTGTTGTTAGTCTGCAGGTGTACCCAACAGCTCCGAAGAGACAGC	
	C9_P_E15 rc	(251)	TTTGTCTGTTGTTAGTCTGCAGGTGTACCCAACAGCTCCGAAGAGACAGC	
	UKRV Mel 2_P_E16	(251)	TTTGTCTGTTGTTAGTCTGCAGGTGTACCCAACAGCTCCGAAGAGACAGC	
	UKRV Mel 2_P_E17 rc	(251)	TTTGTCTGTTGTTAGTCTGCAGGTGTACCCAACAGCTCCGAAGAGACAGC	
	UKRV Mel 2_P_E19 rc	(251)	TTTGTCTGTTGTTAGTCTGCAGGTGTACCCAACAGCTCCGAAGAGACAGC	
	UKRV Mel 2_P_E20	(251)	TTTGTCTGTTGTTAGTCTGCAGGTGTACCCAACAGCTCCGAAGAGACAGC	
	UKRV Mel 2_P_E21 rc	(251)	TTTGTCTGTTGTTAGTCTGCAGGTGTACCCAACAGCTCCGAAGAGACAGC	
	C9_P_E1	(251)	TTTGTCTGTTGTTAGTCTGCAGGTGTACCCAACAGCTCCGAAGAGACAGC	
	UKRV Mel 2_P_E18 rc	(251)	TTTGTCTGTTGTTAGTCTGCAGGTGTACCCAACAGCTCCGAAGAGACAGC	
		301		350
	C9_P_E2 rc	(301)	GACCATCGAGAACGGGCCATGATGACGATGGCGGTTTTGTTCGAAAAGAAA	
	UKRV Mel2_P_E3 rc	(301)	GACCATCGAGAACGGGCCATGATGACGATGGCGGTTTTGTTCGAAAAGAAA	
	C9_P_E4	(301)	GACCATGAAGAACGGGCCATGATGACGATGGCGGTTTTGTTCGAAAAGAAA	
	C9_P_E5 rc	(301)	GACCATCGAGAACGGGCCATGATGACGATGGCGGTTTTGTTCGAAAAGAAA	
	C9_P_E6	(301)	GACCATCGAGAACGGGCCATGATGACGATGGCGGTTTTGTTCGAAAAGAAA	
	UKRV Mel 2_P_E7 rc	(301)	GACCATCGAGAACGGGCCATGATGACGATGGCGGTTTTGTTCGAAAAGAAA	
	C9_P_E8	(301)	GACCATCGAGAACGGGCCATGATGACGATGGCGGTTTTGTTCGAAAAGAAA	
	C9_P_E9	(301)	GACCATCGAGAACGGGCCATGATGACGATGGCGGTTTTGTTCGAAAAGAAA	
	UKRV Mel 2_P_E10	(301)	GACCATCGAGAACGGGCCATGATGACGATGGCGGTTTTGTTCGAAAAGAAA	
	C9_P_E11	(301)	GACCATCGAGAACGGGCCATGATGACGATGGCGGTTTTGTTCGAAAAGAAA	
	C9_P_E12	(300)	GACCATCGAGAACGGTCCATGATGACGATGGGGGGTTTTGCCAAAAGAAAG	
	C9_P_E13 rc	(301)	GACCATCGAGAACGGGCCATGAAGACGAAGGCGTTTTGTTCGAAAAGAAA	
	UKRV Mel 2_P_E14 rc	(301)	GACCATCGAGAACGGGCCATGATGACGATGGCGGTTTTGTTCGAAAAGAAA	
	C9_P_E15 rc	(301)	GACCATCGAGAACGGGCCATGATGACGATGGCGGTTTTGTTCGAAAAGAAA	
	UKRV Mel 2_P_E16	(301)	GACCATCGAGAACGGGCCATGATGACGATGGCGGTTTTGTTCGAAAAGAAA	
	UKRV Mel 2_P_E17 rc	(301)	GACCTTCGAGAACGGGCCATGATGACGATGGCGGTTTTGTTCGAAAAGAAA	
	UKRV Mel 2_P_E19 rc	(301)	GACCATCGAGAACGGGCCATGATGACGATGGCGGTTTTGTTCGAAAAGAAA	
	UKRV Mel 2_P_E20	(301)	GACCATGGAGAACGGGCCATGAAGACGAGGGGGTTTTGTAGAAAAGAAA	
	UKRV Mel 2_P_E21 rc	(301)	GACCATCGAGAACGGGCCATGATGACGATGGCGGTTTTGTTCGAAAAGAAA	
	C9_P_E1	(301)	GACCATCGAGAACGGGCCATGATGACGATGGCGGTTTTGTTCGAAAAGAAA	
	UKRV Mel 2_P_E18 rc	(301)	GACCATCGAGAACGGGCCATGATGACGATGGCGGTTTTGTTCGAAAAGAAA	
		351		400
	C9_P_E2 rc	(351)	AGGGGAAAATGTGGGAAAAGCAAGAGAGATCAGATTGTTACTGTGTCTG	
	UKRV Mel2_P_E3 rc	(351)	AGGGGAAAATGTGGGAAAAGCAAGAGAGATCAGATTGTTACTGTGTCTG	
	C9_P_E4	(351)	AGGGGAAAATGTGGGAAAAGCAAGAGAGATCAGATTGTTACTGTGTCTG	
	C9_P_E5 rc	(351)	AGGGGAAAATGTGGGAAAAGCAAGAGAGATCAGATTGTTACTGTGTCTG	
	C9_P_E6	(351)	AGGGGAAAATGTGGGAAAAGCAAGAGAGATCAGATTGTTACTGTGTCTG	
	UKRV Mel 2_P_E7 rc	(351)	AGGGGAAAATGTGGGAAAAGCAAGAGAGATCAGATTGTTACTGTGTCTG	
	C9_P_E8	(351)	AGGGGAAAATGTGGGAAAAGCAAGAGAGATCAGATTGTTACTGTGTCTG	
	C9_P_E9	(351)	AGGGGAAAATGTGGGAAAAGCAAGAGAGATCAGATTGTTACTGTGTCTG	
	UKRV Mel 2_P_E10	(351)	AGGGGAAAATGTGGGAAAAGCAAGAGAGATCAGATTGTTACTGTGTCTG	
	C9_P_E11	(351)	AGGGGAAAATGTGGGAAAAGCAAGAGAGATCAGATTGTTACTGTGTCTG	
	C9_P_E12	(350)	ACGGGAAAATGTGGGAAAAGCAAGAGAGATCAGATTGTTACTGTGTCTG	
	C9_P_E13 rc	(351)	AGGGGAAAATGGGGAAAAGCAAGAGAGATCAGATTGTTTCTGGGTCTG	
	UKRV Mel 2_P_E14 rc	(351)	AGGGGAAAATGTGGGAAAAGCAAGAGAGATCAGATTGTTACTGTGTCTG	
	C9_P_E15 rc	(351)	AGGGGAAAATGTGGGAAAAGCAAGAGAGATCAGATTGTTACTGTGTCTG	
	UKRV Mel 2_P_E16	(351)	AGGGGAAAATGTGGGAAAAGCAAGAGAGATCAGATTGTTACTGTGTCTG	
	UKRV Mel 2_P_E17 rc	(351)	AGGGGAAAATGTGGGAAAAGCAAGAGAGATCAGATTGTTACTGTGTCTG	
	UKRV Mel 2_P_E19 rc	(351)	AGGGGAAAATGTGGGAAAAGCAAGAGAGATCAGATTGTTACTGTGTCTG	
	UKRV Mel 2_P_E20	(351)	CGGGGAAAACGGGGAAAAGCAAGAGAGATCAGATTGTTTCTGTGTATG	
	UKRV Mel 2_P_E21 rc	(351)	AGGGGAAAATGTGGGAAAAGCAAGAGAGATCAGATTGTTACTGTGTCTG	
	C9_P_E1	(351)	AGGGGAAAATGTGGGAAAAGCAAGAGAGATCAGATTGTTACTGTGTCTG	
	UKRV Mel 2_P_E18 rc	(351)	AGGGGAAAATGTGGGAAAAGCAAGAGAGATCAGATTGTTACTGTGTCTG	

		401		450
	C9_P_E2 rc	(401)	TGTAGAAAGAAGTAGACATAGGAGACTCCATTTTGTATGTACTAAGAAA	
	UKRV Mel2_P_E3 rc	(401)	TGTAGAAAGAAGTAGACATAGGAGACTCCATTTTGTATGTACTAAGAAA	
	C9_P_E4	(401)	TGTAGAAAGAAGTAGACCTAGGAGACTCCATTTTGTATGTACTAAGAAA	
	C9_P_E5 rc	(401)	TGTAGAAAGAAGTAGACATAGGAGACTCCATTTTGTATGTACTAAGAAA	
	C9_P_E6	(401)	TGTAGAAAGAAGTAGACATAGGAGACTCCATTTTGTATGTACTAAGAAA	
	UKRV Mel 2_P_E7 rc	(401)	TGTAGAAAGAAGTAGACATAGGAGACTCCATTTTGTATGTACTAAGAAA	
	C9_P_E8	(401)	TGTAGAAAGAAGTAGACATAGGAGACTCCATTTTGTATGTACTAAGAAA	
	C9_P_E9	(401)	TGTAGAAAGAAGTAGACATAGGAGACTCCATTTTGTATGTACTAAGAAA	
	UKRV Mel 2_P_E10	(401)	TGTAGAAAGAAGTAGACATAGGAGACTCCATTTTGTATGTACTAAGAAA	
	C9_P_E11	(401)	TGTAGAAAGAAGTAGACATAGGAGACTCCATTTTGTATGTACTAAGAAA	
	C9_P_E12	(400)	TGTAGAAAGAAGTAGACATACGAGACTCCCTTTTGTATGTACTAAGAAA	
	C9_P_E13 rc	(401)	TGTAGAAAGAAGTAGACATAGGAGACTCCATTTTGTATGTACTAAGAAA	
	UKRV Mel 2_P_E14 rc	(401)	TGTAGAAAGAAGTAGACATAGGAGACTCCATTTTGTATGTACTAAGAAA	
	C9_P_E15 rc	(401)	TGTAGAAAGAAGTAGACATAGGAGACTCCATTTTGTATGTACTAAGAAA	
	UKRV Mel 2_P_E16	(401)	TGTAGAAAGAAGTAGACATAGGAGACTCCATTTTGTATGTACTAAGAAA	
	UKRV Mel 2_P_E17 rc	(401)	TGTAGAAAGAAGTAGACATAGGAGACTCCATTTTGTATGTACTAAGAAA	
	UKRV Mel 2_P_E19 rc	(401)	TGTAGAAAGAAGTAGACATAGGAGACTCCATTTTGTATGTACTAAGAAA	
	UKRV Mel 2_P_E20	(401)	TGTAGAAAGAAGTAGACATAGGAGACTCCATTTTGTATGTACTAAGAAA	
	UKRV Mel 2_P_E21 rc	(401)	TGTAGAAAGAAGTAGACATAGGAGACTCCATTTTGTATGTACTAAGAAA	
	C9_P_E1	(401)	TGTAGAAAGAAGTAGACATAGGAGACTCCATTTTGTATGTACTAAGAAA	
	UKRV Mel 2_P_E18 rc	(401)	TGTAGAAAGAAGTAGACATAGGAGACTCCATTTTGTATGTACTAAGAAA	
		451		500
	C9_P_E2 rc	(451)	AATTCCTTCTGCCTTGAGATTCTGTTAATCTATGACCTTACCCCCAACCCC	
	UKRV Mel2_P_E3 rc	(451)	AATTCCTTCTGCCTTGAGATCCTGTTAATCTATGACCTTACCCCCAACCCC	
	C9_P_E4	(451)	AATTCCTTCTGCCTTGAGATTCTGTTAATCTATGACCTTACCCCCAACCCC	
	C9_P_E5 rc	(451)	AATTCCTTCTGCCTTGAGATTCTGTTAATCTATGACCTTACCCCCAACCCC	
	C9_P_E6	(451)	AATTCCTTCTGCCTTGAGATTCTGTTAATCTATGACCTTACCCCCAACCCC	
	UKRV Mel 2_P_E7 rc	(451)	AATTCCTTCTGCCTTGAGATTCTGTTAATCTATGACCTTACCCCCAACCCC	
	C9_P_E8	(451)	AATTCCTTCTGCCTTGAGATTCTGTTAATCTATGACCTTACCCCCAACCCC	
	C9_P_E9	(451)	AATTCCTTCTGCCTTGAGATTCTGTTAATCTATGACCTTACCCCCAACCCC	
	UKRV Mel 2_P_E10	(451)	AATTCCTTCTGCCTTGAGATTCTGTTAATCTATGACCTTACCCCCAACCCC	
	C9_P_E11	(451)	AATTCCTTCTGCCTTGAGATTCTGTTAATCTATGACCTTACCCCCAACCCC	
	C9_P_E12	(450)	AATTCCTTCTGCCTTGACATTCTGTTAATCTATGACCTTACCCCCAACCCC	
	C9_P_E13 rc	(451)	AATTTTTTCTGCCTTGGGATTTTGTAAATCTATGACCTTACCCCCAACCCC	
	UKRV Mel 2_P_E14 rc	(451)	AATTCCTTCTGCCTTGAGATTCTGTTAATCTATGACCTTACCCCCAACCCC	
	C9_P_E15 rc	(451)	AATTCCTTCTGCCTTGAGATTCTGTTAATCTATGACCTTACCCCCAACCCC	
	UKRV Mel 2_P_E16	(451)	AATTCCTTCTGCCTTGAGATTCTGTTAATCTATGACCTTACCCCCAACCCC	
	UKRV Mel 2_P_E17 rc	(451)	AATTTTTTCTGCCTTGAGATTCTGTTAATCTATGACCTTACCCCCAACCCC	
	UKRV Mel 2_P_E19 rc	(451)	AATTCCTTCTGCCTTGAGATTCTGTTAATCTATGACCTTACCCCCAACCCC	
	UKRV Mel 2_P_E20	(451)	AATTTTTTCTGCCTTGAGATTCTGTTAATCTATGACCTTACCCCCAACCCC	
	UKRV Mel 2_P_E21 rc	(451)	AATTCCTTCTGCCTTGAGATTCTGTTAATCTATGACCTTACCCCCAACCCC	
	C9_P_E1	(451)	AATTCCTTCTGCCTTGAGATTCTGTTAATCTATAACCTTACCCCCAACCCC	
	UKRV Mel 2_P_E18 rc	(451)	AATTCCTTCTGCCTTGAGATTCTGTTAATCTATAACCTTACCCCCAACCCC	
		501		543
	C9_P_E2 rc	(501)	GTGCTCTCTGAAACATGTGCTGTGTCCACTCAGAGTTGAATGG	
	UKRV Mel2_P_E3 rc	(501)	GTGCTCTCTGAAACATGTGCTGTGTCCACTCAGAGTTGAATGG	
	C9_P_E4	(501)	GTGCTCTCTGAAACATGTGGTGTGTCCACTCCTAGTTGAATGG	
	C9_P_E5 rc	(501)	GTGCTCTCTGAAACATGTGCTGTGTCCACTCAGAGTTGAATGG	
	C9_P_E6	(501)	GTGCTCTCTGAAACATGTGCTGTGTCCACTCAGAGTTGAATGG	
	UKRV Mel 2_P_E7 rc	(501)	GTGCTCTCTGAAACATGTGCTGTGTCCACTCAGAGTTGAATGG	
	C9_P_E8	(501)	GTGCTCTCTGAAACATGTGCTGTGTCC-CTCAGAGTTGAATGG	
	C9_P_E9	(501)	GTGCTCTCTGAAACATGTGCTGTGTCCACTCAGAGTTGAATGG	
	UKRV Mel 2_P_E10	(501)	GTGCTCTCTGAAACATGTGCTGTGTCCACTCAGAGTTGAATGG	
	C9_P_E11	(501)	GTGCTCTCTGAAACATGTGCTGTGTCCACTCAGAGTTGAATGG	
	C9_P_E12	(500)	GTGCTCTCTGAAACATGTGCTGTGTCCACTGGGCCATGAATGG	
	C9_P_E13 rc	(501)	GTGCTTTTTGAAACATGTGCTGTGTCCAATCAGAGTTGAATGG	
	UKRV Mel 2_P_E14 rc	(501)	GTGCTCTCTGAAACATGTGCTGTGTCCACTCAGAGTTGAATGG	
	C9_P_E15 rc	(501)	GTGCTCTCTGAAACATGTGCTGTGTCCACTCAGAGTTGAATGG	
	UKRV Mel 2_P_E16	(501)	GTGCTCTCTGAAACATGTGCTGTGTCCACTCAGAGTTGAATGG	
	UKRV Mel 2_P_E17 rc	(501)	GTGCTCTCTGAAACATGTGCTGTGTCCACT-AGAGTTGAATGG	
	UKRV Mel 2_P_E19 rc	(501)	GTGCTCTCTGAAACATGTGCTGTGTCCACTCAGAGTTGAATGG	
	UKRV Mel 2_P_E20	(501)	GTGCTCTCTGAAACATGTGCTGTGTCCAATCAGAGTTGAATGG	
	UKRV Mel 2_P_E21 rc	(501)	GTGCTCTCTGAAACATGTGCTGTGTCCACTCAGAGTTGAATGG	
	C9_P_E1	(501)	GTGCTCTCTGAAACATGTGCTGTGTCCACTCAGAGTTGAATGG	
	UKRV Mel 2_P_E18 rc	(501)	GTGCTCTCTGAAACATGTGCTGTGTCCACTCAGAGTTGAATGG	

Nukleotidalignment aus dem *gag*-Bereich, Klone - Partikeln UKRV Mel 2-C9

```

1
UKRV Mel 2_P_1 (1) -GCAGCCCTATTTCTTCGGACCTGTTCTTGTACCCCATCAATCCACCAAG
UKRV Mel 2_P_2 (1) -GCAGCCCTATTTCTTCGGACCTGTTCTTGTACCCCATCAATCCACCAAG
UKRV Mel 2_P_3 (1) -GCAGCCCTATTTCTTCGGACCTGTTCTTGTACCCCATCAATCCACCAAG
UKRV Mel 2_P_5 (1) -GCAGCCCTATTTCTTCGGACCTGTTCTTGTACCCCATCAATCCACCAAG
UKRV Mel 2_P_6 (1) TGCAGCCCTATTTCTTCGGACCTGTTCTTGTACCCCATCAATCCACCAAG
UKRV Mel 2_P_8 (1) -GCAGCCCTATTTCTTCGGACCTGTTCTTGTACCCCATCAATCCACCAAG
UKRV Mel 2_P_11 (1) -GCAGCCCTATTTCTTCGGACCTGTTCTTGTACCCCATCAATCCACCAAG
UKRV Mel 2_P_12 (1) -GCAGCCCTATTTCTTCGGACCTGTTCTTGTACCCCATCAATCCACCAAG
UKRV Mel 2_P_13 (1) -GCAGCCCTATTTCTTCGGACCTGTTCTTGTACCCCATCAATCCACCAAG
UKRV Mel 2_P_14 (1) -GCAGCCCTATTTCTTCGGACCTGTTCTTGTACCCCATCAATCCACCAAG
UKRV Mel 2_P_15 (1) -GCAGCCCTATTTCTTCGGACCTGTTCTTGTACCCCATCAATCCACCAAG
UKRV Mel 2_P_16 rc (1) -GCAGCCCTATTTCTTCGGACCTGTTCTTGTACCCCATCAATCCACCAAG
UKRV Mel 2_P_18 (1) -GCAGCCCTATTTCTTCGGACCTGTTCTTGTACCCCATCAATCCACCAAG
UKRV Mel 2_P_20 (1) -----TTCTTCGGACCTGTTCTTGTACCCCATCAATCCACCAAG
UKRV Mel 2_P_22 (1) -GCAGCCCTATTTCTTCGGACCTGTTCTTGTACCCCATCAATCCACCAAG
UKRV Mel 2_P_23 rc (1) GGCAGCCCTATTTCTTCGGACCTGTTCTTGTACCCCATCAATCCACCAAG
UKRV Mel 2_P_25 (1) -GCAGCCCTATTTCTTCGGACCTGTTCTTGTACCCCATCAATCCACCAAG
UKRV Mel 2_P_26 (1) -GCAGCCCTATTTCTTCGGACCTGTTCTTGTACCCCATCAATCCACCAAG
UKRV Mel 2_P_27 (1) -GCAGCCCTATTTCTTCGGACCTGTTCTTGTACCCCATCAATCCACCAAG
UKRV Mel 2_P_28 (1) -GCAGCCCTATTTCTTCGGACCTGTTCTTGTACCCCATCAATCCACCAAG
UKRV Mel 2_P_29 (1) -GCAGCCCTATTTCTTCGGACCTGTTCTTGTACCCCATCAATCCACCAAG
UKRV Mel 2_P_30 (1) -GCAGCCCTATTTCTTCGGACCTGTTCTTGTACCCCATCAATCCACCAAG

```

```

51
UKRV Mel 2_P_1 (50) TCTTAAATTGTAAAAATTGACAGGGTGAGAGAGACGATTTTGCCAGAATC
UKRV Mel 2_P_2 (50) TCTTAAATTGTAAAAATTGACAGGGTGAGAGAGACGATTTTGCCAGAATC
UKRV Mel 2_P_3 (50) TCTTAAATTGTAAAAATTGACAGGGTGAGAGAGACGATTTTGCCAGAATC
UKRV Mel 2_P_5 (50) TCTTAAATTGTAAAAATTGACAGGGTGAGAGAGACGATTTTGCCAGAATC
UKRV Mel 2_P_6 (51) TCTTAAATTGTAAAAATTGACAGGGTGAGAGAGACGATTTTGCCAGAATC
UKRV Mel 2_P_8 (50) TCTTAAATTGTAAAAATTGACAGGGTGAGAGAGACGATTTTGCCAGAATC
UKRV Mel 2_P_11 (50) TCTTAAATTGTAAAAATTGACAGGGTGAGAGAGACGATTTTGCCAGAATC
UKRV Mel 2_P_12 (50) TCTTAAATTGTAAAAATTGACAGGGTGAGAGAGACGATTTTGCCAGAATC
UKRV Mel 2_P_13 (50) TCTTAAATTGTAAAAATTGACAGGGTGAGAGAGACGATTTTGCCAGAATC
UKRV Mel 2_P_14 (50) TCTTAAATTGTAAAAATTGACAGGGTGAGAGAGACGATTTTGCCAGAATC
UKRV Mel 2_P_15 (50) TCTTAAATTGTAAAAATTGACAGGGTGAGAGAGACGATTTTGCCAGAATC
UKRV Mel 2_P_16 rc (50) TCTTAAATTGTAAAAATTGACAGGGTGAGAGAGACGATTTTGCCAGAATC
UKRV Mel 2_P_18 (50) TCTTAAATTGTAAAAATTGACAGGGTGAGAGAGACGATTTTGCCAGAATC
UKRV Mel 2_P_20 (40) TCTTAAATTGTAAAAATTGACAGGGTGAGAGAGACGATTTTGCCAGAATC
UKRV Mel 2_P_22 (50) TCTTAAATTGTAAAAATTGACAGGGTGAGAGAGACGATTTTGCCAGAATC
UKRV Mel 2_P_23 rc (51) TCTTAAATTGTAAAAGTTGACAGGGTGAGAGAGACGATTTTGCCAGAATC
UKRV Mel 2_P_25 (50) TCTTAAATTGTAAAAATTGACAGGGTGAGAGAGACGATTTTGCCAGAATC
UKRV Mel 2_P_26 (50) TCTTAAATTGTAAAAATTGACAGGGTGAGAGAGACGATTTTGCCAGAATC
UKRV Mel 2_P_27 (50) TCTTAAATTGTAAAAATTGACAGGGTGAGAGAGACGATTTTGCCAGAATC
UKRV Mel 2_P_28 (50) TCTTAAATTGTAAAAATTGAAAGGGGGAGAGAGACGATTTTGCCAGAATC
UKRV Mel 2_P_29 (50) TCTTAAATTGTAAAAATTGACAGGGTGAGAGAGACGATTTTGCCAGAATC
UKRV Mel 2_P_30 (50) TCTTAAATTGTAAAAATTGACAGGGTGAGAGAGACGATTTTGCCAGAATC

```


551 600

UKRV Mel 2_P_1 (550) TCTTTAACCTGCCTTTTGAGGTTGTAATGTTACAGGCACCTGACCTGCTGG
 UKRV Mel 2_P_2 (550) TCTTTAACCTGCCTTTTGAGGTTGTAATGTTACAGGCACCTGACCTGCTGG
 UKRV Mel 2_P_3 (550) TCTTTAACCTGCCTTTTGAGGTTGTAATGTTACAGGCACCTGACCTGCTGG
 UKRV Mel 2_P_5 (550) TCTTTAACCTGCCTTTTGAGGTTGTAATGTTACAGGCACCTGACCTGCTGG
 UKRV Mel 2_P_6 (551) TCTTTAACCTGCCTTTTGAGGTTGTAATGTTACAGGCACCTGACCTGCTGG
 UKRV Mel 2_P_8 (550) TCTTTAACCTGCCTTTTGAGGTTGTAATGTTACAGGCACCTGACCTGCTGG
 UKRV Mel 2_P_11 (550) TCTTTAACCTGCCTTTTGAGGTTGTAATGTTACAGGCACCTGACCTGCTGG
 UKRV Mel 2_P_12 (550) TCTTTAACCTGCCTTTTGAGGTTGTAATGTTACAGGCACCTGACCTGCTGG
 UKRV Mel 2_P_13 (550) TCTTTAACCTGCCTTTTGAGGTTGTAATGTTACAGGCACCTGACCTGCTGG
 UKRV Mel 2_P_14 (550) TCTTTAACCTGCCTTTTGAGGTTGTAATGTTACAGGCACCTGACCTGCTGG
 UKRV Mel 2_P_15 (550) TCTTTAACCTGCCTTTTGAGGTTGTAATGTTACAGGCACCTGACCTGCTGG
 UKRV Mel 2_P_16 rc (550) TCTTTAACCTGCCTTTTGAGGTTGTAATGTTACAGGCACCTGACCTGCTGG
 UKRV Mel 2_P_18 (550) TCTTTAACCTGCCTTTTGAGGTTGTAATGTTACAGGCACCTGACCTGCTGG
 UKRV Mel 2_P_20 (540) TCTTTAACCTGCCTTTTGAGGTTGTAATGTTACAGGCACCTGACCTGCTGG
 UKRV Mel 2_P_22 (550) TCTTTAACCTGCCTTTTGAGGTTGTAATGTTACAGGCACCTGACCTGCTGG
 UKRV Mel 2_P_23 rc (551) TCTTTAACCTGCCTTTTGAGGTTGTAATGTTACAGGCACCTGACCTGCTGG
 UKRV Mel 2_P_25 (550) TCTTTAACCTGCCTTTTGAGGTTGTAATGTTACAGGCACCTGACCTGCTGG
 UKRV Mel 2_P_26 (550) TCTTTAACCTGCCTTTTGAGGTTGTAATGTTACAGGCACCTGACCTGCTGG
 UKRV Mel 2_P_27 (550) TCTTTAACCTGCCTTTTGAGGTTGTAATGTTACAGGCACCTGACCTGCTGG
 UKRV Mel 2_P_28 (550) TCTTTAACCTGCCTTTTGAGGTTGTAATGTTACAGGCACCTGACCTGCTGG
 UKRV Mel 2_P_29 (550) TCTTTAACCTGCCTTTTGAGGTTGTAATGTTACAGGCACCTGACCTGCTGG
 UKRV Mel 2_P_30 (550) TCTTTAACCTGCCTTTTGAGGTTGTAATGTTACAGGCACCTGACCTGCTGG

601 638

UKRV Mel 2_P_1 (600) AAGAGGACTTGTGCCTCGTGGTTTACTGCTGATGGCC
 UKRV Mel 2_P_2 (600) AAGAGGACTTGTGCCTCGTGGTTTACTGCTGATGGCC
 UKRV Mel 2_P_3 (600) AAGAGGACTTGTGCCTCGTGGTTTACTGCTGATGGCC
 UKRV Mel 2_P_5 (600) AAGAGGACTTGTGCCTCGTGGTTTACTGCTGATGGCC
 UKRV Mel 2_P_6 (601) AAGAGGACTTGTGCCTCGTGGTTTACTGCTGATGGCC
 UKRV Mel 2_P_8 (600) AAGAGGACTTGTGCCTCGTGGTTTACTGCTGATGGCC
 UKRV Mel 2_P_11 (600) AAGAGGACTTGTGCCTCGTGGTTTACTGCTGATGGCC
 UKRV Mel 2_P_12 (600) AAGAGGACTTGTGCCTCGTGGTTTACTGCTGATGGCC
 UKRV Mel 2_P_13 (600) AAGAGGACTTGTGCCTCGTGGTTTACTGCTGATGGCC
 UKRV Mel 2_P_14 (600) AAGAGGACTTGTGCCTCGTGGTTTACTGCTGATGGCC
 UKRV Mel 2_P_15 (600) AAGAGGACTTGTGCCTCGTGGTTTACTGCTGATGGCC
 UKRV Mel 2_P_16 rc (600) AAGAGGACTTGTGCCTCGTGGTTTACTGCTGATGGCC
 UKRV Mel 2_P_18 (600) AAGAGGACTTGTGCCTCGTGGTTTACTGCTGATGGCC
 UKRV Mel 2_P_20 (590) AAGAGGACTTGTGCCTCGTGGTTTACTGCTGATGGCC
 UKRV Mel 2_P_22 (600) AAGAGGACTTGTGCCTCGTGGTTTACTGCTGATGGCC
 UKRV Mel 2_P_23 rc (601) AAGAGGACTTGTGCCTCGTGGTTTACTGCTGATGGCC
 UKRV Mel 2_P_25 (600) AAGAGGACTTGTGCCTCGTGGTTTACTGCTGATGGCC
 UKRV Mel 2_P_26 (600) AAGAGGACTTGTGCCTCGTGGTTTACTGCTGATGGCC
 UKRV Mel 2_P_27 (600) AAGAGGACTTGTGCCTCGTGGTTTACTGCTGATGGCC
 UKRV Mel 2_P_28 (600) AAGAGGACTTGTGCCTCGTGGTTTACTGCTGATGGCC
 UKRV Mel 2_P_29 (600) AAGAGGACTTGTGCCTCGTGGTTTACTGCTGATGGCC
 UKRV Mel 2_P_30 (600) AAGAGGACTTGTGCCTCGTGGTTTACTGCTGATGGCC

Aminosäurealignment aus dem *gag/prot* –Bereich zytoplasmatischer *mRNA*-Klone

		1	50
NCCIT-1	(1)	-----	-----MVDLQAAAN
NCCIT-2	(1)	-----	-----MVDLQAAAN
NCCIT-3	(1)	-----	-----
NCCIT-5	(1)	-----	-----MVDLQAAAN
NCCIT-6	(1)	-----	-----
NCCIT-14	(1)	-----	-----MAHVGNEQGAAQPTTTGAFHQPFVLRVSGT
NCCIT-15	(1)	-----	-----MVDLQAAAN
NCCIT-16	(1)	-----	-----MHHALGALPYGRPAGGRE
NCCIT-17	(1)	-----	-----
NCCIT-19	(1)	-----	-----
NCCIT-18	(1)	-----	-----
NCCIT-20	(1)	-----	-----
NCCIT-21	(1)	-----	-----MAHVGNEQGAAQPTTTGAFHQPFVLRVSGT
NCCIT-22	(1)	-----	-----
NCCIT-24	(1)	-----	-----
Mel 2-73	(1)	-----	-----
Mel 2-74	(1)	MSRCKSKHGLVNVVYLIKWAACRKRARGSSAPQQTGHPNSAICSSGFSGT	
Mel 2-80	(1)	-----	-----
Mel 2-81	(1)	-----	-----
Mel 2-82	(1)	-----	-----
Mel 2-78	(1)	-----	-----
Mel 2-95	(1)	-----	-----
GH-36	(1)	-----	-----
GH-39	(1)	-----	-----
GH-40	(1)	-----	-----
GH-41	(1)	-----	-----

		51	100
NCCIT-1	(10)	SLVILARSEEARVIMGQTKSKIKSKYASYLSFIKILLKRGGVKVKSTKNLI	
NCCIT-2	(10)	SLVILARSEEARVIMGQTKSKIKSKYASYLSFIKILLKRGGVKVSANLI	
NCCIT-3	(1)	-----	
NCCIT-5	(10)	SLVILARSEEARVIMGQTKSKIKSKYASYLSFIKILLKRGGVKVSANLI	
NCCIT-6	(1)	-----VIMGQTKSKIKSKYASYLSFIKILLKRGGVKVKSTKNLI	
NCCIT-14	(31)	TTPHCPKCFRNKPVTTIQQLSPATSGSAAVDLCTIQAVSLLPGEPQKTP	
NCCIT-15	(10)	SLVILARSEEARVIMGQTKSKIKSKYASYLSFIKILLKRGGVKVKSTKNLI	
NCCIT-16	(19)	FTSDSCAPEEARVIMGQTKSKIKSKYASYLSFIKILLKRGGVKVKSTKNLI	
NCCIT-17	(1)	-----MGQTKSKIKSKYASYLSFIKILLKRGGVKVKSTKNLI	
NCCIT-19	(1)	-----	
NCCIT-18	(1)	-----VIMGQTKSKIKSKYASYLSFIKILLKRGGVKVKSTKNLI	
NCCIT-20	(1)	-----	
NCCIT-21	(31)	TTPHCPKCFRNKPVTTIQQLSPATSGSAAVDLCTIQAVSLLPGEPQKTP	
NCCIT-22	(1)	-----	
NCCIT-24	(1)	-----	
Mel 2-73	(1)	-----	
Mel 2-74	(51)	TTPTVPSVSGNKPVTTIQQLSPATSGSAAVDLCTIQAVSLLPGEPQKTP	
Mel 2-80	(1)	-----MGQTKSKIKSKYASYLSFIKILLKRGGVKVKSTKNLI	
Mel 2-81	(1)	-----MGQTKSKIKSKYASYLSFIKILLKRGGVKVKSTKNLI	
Mel 2-82	(1)	-----MGQTKSKIKSKYASYLSFIKILLKRGGVKVKSTKNLI	
Mel 2-78	(1)	-----MGQTKSKIKSKYASYLSFIKILLKRGGVKVKSTKNLI	
Mel 2-95	(1)	-----MGQTKSKIKSKYASYLSFIKILLKRGGVKVKSTKNLI	
GH-36	(1)	-----	
GH-39	(1)	-----	
GH-40	(1)	-----	
GH-41	(1)	-----	

	101	150
NCCIT-1	(60)	KLFQIEQFCPWFPEQGTLDLKDWKRIKELKQAGRKGNIIPITVWNDWA
NCCIT-2	(60)	KLFQIEQFCPWFPEQGTLDLKDWKRIKELKQAGRKGNIIPITVWNDWA
NCCIT-3	(1)	-----VVDSYKGEIQVISSSIP
NCCIT-5	(60)	KLFQIEQFCPWFPEQGTLDLKDWKRIKELKQAGRKGNIIPITVWNDWA
NCCIT-6	(39)	KLFQIEQFCPWFPEQGTLDLKDWKRIKELKQAGRKGNIIPITVWNDWA
NCCIT-14	(81)	TGVYGPLPKGTVGLLILGRSSLNLKGVQIHTSVVDSYKGEIQVISSSIP
NCCIT-15	(60)	KLFQIEQFCPWFPEQGTLDLKDWKRIKELKQAGRKGNIIPITVWNDWA
NCCIT-16	(69)	KLFQIEQFCPWFPEQGTLDLKDWKRIKELKQAGRKGNIIPITVWNDWA
NCCIT-17	(37)	KLFQIEQFCPWFPEQGTLDLKDWKRIKELKQAGRKGNIIPITVWNDWA
NCCIT-19	(1)	-----VVDSYKGEIQVISSSIP
NCCIT-18	(39)	KLFQIEQFCPWFPEQGTLDLKDWKRIKELKQAGRKGNIIPITVWNDWA
NCCIT-20	(1)	-----VVDSYKGEIQVISSSIP
NCCIT-21	(81)	TGVYGPLPKGTVGLLILGRSSLNLKGVQIHTSVVDSYKGEIQVISSSIP
NCCIT-22	(1)	-----VVDSYKGEIQVISSSIP
NCCIT-24	(1)	-----VVDSYKGEIQVISSSIP
Mel 2-73	(1)	-----VVDSYKGEIQVISSSIP
Mel 2-74	(101)	TGVYGPLPKGTVGLLILGRSSLNLKGVQIHTSVVDSYKGEIQVISSSIP
Mel 2-80	(37)	KLFQIEQFCPWFPEQGTLDLKDWKRIKELKQAGRKGNIIPITVWNDWA
Mel 2-81	(37)	KLFQIEQFCPWFPEQGTLDLKDWKRIKELKQAGRKGNIIPITVWNDWA
Mel 2-82	(37)	KLFQIEQFCPWFPEQGTLDLKDWKRIKELKQAGRKGNIIPITVWNDWA
Mel 2-78	(37)	KLFQIEQFCPWFPEQGTLDLKDWKRIKELKQAGRKGNIIPITVWNDWA
Mel 2-95	(37)	KLFQIEQFCPWFPEQGTLDLKDWKRIKELKQAGRKGNIIPITVWNDWA
GH-36	(1)	-----VVDSYKGEIQVISSSIP
GH-39	(1)	-----VVDSYKGEIQVISSSIP
GH-40	(1)	-----VVDSYKGEIQVISSSIP
GH-41	(1)	-----VVDSYKGEIQVISSSIP

	151	200
NCCIT-1	(110)	I I K A A L E P F Q T E E D S V S V S D A L G S C I I D C N E N T R K K S Q K E T E G L H C E Y V A
NCCIT-2	(110)	I I K A A L E P F Q T E E D S V S V S D A L G S C I T D C N E N T R K K S Q K E T E G L H C E Y V A
NCCIT-3	(20)	W S A S P G D R I A Q L L L L P Y I K G G N S E I K R I G G L G S T D P T G K A A Y W A S Q V S E N
NCCIT-5	(110)	I I K A A L E P F Q T E E D S V S V S D A L G S C I T D C N E N T R K K S Q K E T E G L H C E Y V A
NCCIT-6	(89)	I I K A A L E P F Q T E E D S V S V S D A L G S C I I D C N E N T R K K S Q K E T E G L H C E Y V A
NCCIT-14	(131)	W S A S P R D R I A Q L L L L P Y I K G G N S E I K R I G G L G S T D P T G K A A Y W A S Q V S E N
NCCIT-15	(110)	I I K A A L E P F Q T E E D S V S V S D A L G S C I I D C N E N T R K K S Q K E T E G L H C E Y V A
NCCIT-16	(119)	I I K A A L E P F Q T E E D S V S V S D A L G S C I I D C N E N T R K K S Q K E T E G L H C E Y V A
NCCIT-17	(87)	I I K A A L E P F Q T E E D S V S V S D A P G S C I I D C N E N T R K K S Q K E T E G L H C E Y V A
NCCIT-19	(20)	W S A S P G D R I A Q L L L L P Y I K G G N S E I K R I G G L G S T D P T G K A A Y W A S Q V S E N
NCCIT-18	(89)	I I K A A L E P F Q T E E D S V S V S D A L G S C I I D C N E N T R K K S Q K E T E G L H C E Y V A
NCCIT-20	(20)	W S A S P G D R I A Q L L L L P Y I K G G N S E I K R I G G L G S T D P T G K A A Y W A S Q V S E N
NCCIT-21	(131)	W S A S P R D R I A Q L L L L P Y I K G G N S E I K R I G G L G S T D P T G K A A Y W A S Q V S E N
NCCIT-22	(20)	W S A S P G D R I A Q L L L L P Y I K G G N S E I K R I G G L G S T D P T G K A A Y W A S Q V S E N
NCCIT-24	(20)	W S A S P G D R I A Q L L L L P Y I K G G N S E I K R I G G L G S T D P T G K A A Y W A S Q V S E N
Mel 2-73	(20)	W S A S P R D R I A Q L L L L P Y I K G G N S E I K R I G G L G S T D P T G K A A Y W A S Q V S E N
Mel 2-74	(151)	W S A S P R D R I A Q L L L L P Y I K G G N S E I K R I G G L G S T D P T G K A A Y W A S Q V S E N
Mel 2-80	(87)	I I K A A L E P F Q T E E D S V S V S D A P G S C I I D C N E N T R K K S Q K E T E G L H C E Y V A
Mel 2-81	(87)	I I K A A L E P F Q T E E D S V S V S D A P G S C I I D C N E N T R K K S Q K E T E G L H C E Y V A
Mel 2-82	(87)	I I K A A L E P F Q T E E D S V S V S D A P G S C I I D C N E N T R K K S Q K E T E G L H C E Y V A
Mel 2-78	(87)	I I K A A L E P F Q T E E D S V S V S D A P G S C I I D C N E N T R K K S Q K E T E G L H C E Y V A
Mel 2-95	(87)	I I K A A L E P F Q T E E D S V S V S D A P G S C I I D C N E N T R K K S Q K E T E G L H C E Y V A
GH-36	(20)	W S A S P G D R I A Q L L L L P Y I K G G N S E I K R I G G L G S T D P T G K A A Y W A S Q V S E N
GH-39	(20)	W S A S P G D R I A Q L L L L P Y I K G G N S E I K R I G G L G S T D P T G K A A Y W A S Q V S E N
GH-40	(20)	W S A S P G D R I A Q L L L L P Y I K G G N S E I K R I G G L G S T D P T G K A A Y W A S Q V S E N
GH-41	(20)	W S A S P G D R I A Q L L L L P Y I K G G N S E I K R I G G L G S T D P T G K A A Y W A S Q V S E N

		201	250
NCCIT-1	(160)	EPVMAQSTONVDYNQLQEVYYPETLKLEGKGP	ELVGPSESKPRGTSHLPA
NCCIT-2	(160)	EPVMAQSTONVDYNQLQEVYYPETLKLEGKGP	ELVGPSESKPRGTSHLPA
NCCIT-3	(70)	RPVCKAIIQGGKQFEGIVDTGADVSI IALNQWP	KNWPKQKAVTGLVGIGTA
NCCIT-5	(160)	EPVMAQSTONVDYNQLQEVYYPETLKLEGKGP	ELVGPSESKPRGTSHLPA
NCCIT-6	(139)	EPVMAQSTONVDYNQLQEVYYPETLKLEGKGP	ELVGPSESKPRGTSHLPA
NCCIT-14	(181)	RPVCKAIIQGGKQFEGIVDTGADVSI IALNQWP	KNWPKQKAVTGLVGIGTA
NCCIT-15	(160)	EPVMAQSTONVDYNQLQEVYYPETLKLEGKGP	ELVGPSESKPRGTSHLPA
NCCIT-16	(169)	EPVMAQSTONVDYNQLQEVYYPETLKLEGKGP	ELVGPSESKPRGTSHLPA
NCCIT-17	(137)	EPVMAQSTONVDYNQLQEVYIYLETLKLEGKGP	ELVGPSESKPRGTSPLPA
NCCIT-19	(70)	RPVCKAIIQGGKQFEGIVDTGADVSI IALNQWP	KNWPKQKAVTGLVGIGTA
NCCIT-18	(139)	EPVMAQSTONVDYNQLQEVYYPETLKLEGKGP	ELVGPSESKPRGTSHLPA
NCCIT-20	(70)	RPVCKAIIQGGKQFEGIVDTGADVSI IALNQWP	KNWPKQKAVTGLVGIGTA
NCCIT-21	(181)	RPVCKAIIQGGKQFEGIVDTGADVSI IALNQWP	KNWPKQKAVTGLVGIGTA
NCCIT-22	(70)	RPVCKAIIQGGKQFEGIVDTGADVSI IALNQWP	KNWPKQKAVTGLVGIGTA
NCCIT-24	(70)	RPVCKAIIQGGKQFEGIVDTGADVSI IALNQWP	KNWPKQKAVTGLVGIGTA
Mel 2-73	(70)	RPVCKAIIQGGKQFEGIVDTGADVSI IALNQWP	KNWPKQKAVTGLVGIGTA
Mel 2-74	(201)	RPVCKAIIQGGKQFEGIVDTGADVSI IALNQWP	KNWAKQKAVTGLVGIGTA
Mel 2-80	(137)	EPVMAQSTONVDYNQLQEVYIYLETLKLEGKGP	ELVGPSESKPRGTSPLPA
Mel 2-81	(137)	EPVMAQSTONVDYNQLQEVYIYLETLKLEGKGP	ELVGPSESKPRGTSPLPA
Mel 2-82	(137)	EPVMAQSTONVDYNQLQEVYIYLETLKLEGKGP	ELVGPSESKPRGTSPLPA
Mel 2-78	(137)	EPVMAQSTONVDYNQLQEVYIYLETLKLEGKGP	ELVGPSESKPRGTSPLPA
Mel 2-95	(137)	EPVMAQSTONVDYNQLQEVYYPETLKLEGKGP	ELVGPSESKPRGTSPLPA
GH-36	(70)	RPVCKAIIQGGKQFEGIVDTGADVSI IALNQWP	KNWPKQKAVTGLVGIGTA
GH-39	(70)	RPVCKAIIQGGKQFEGIVDTGADVSI IALNQWP	KNWPKQKAVTGLVGIGTA
GH-40	(70)	RPVCKAIIQGGKQFEGIVDTGADVSI IALNQWP	KNWPKQKAVTGLVGIGTA
GH-41	(70)	RPVCKAIIQGGKQFEGIVDTGADVSI IALNQWP	KNWPKQKAVTGLVGIGTA

		251	300
NCCIT-1	(210)	GOVPVTLQPQKQVKENKTQPPVAYQYWPPAELQYRPPPE	OYGYPGMPSA
NCCIT-2	(210)	GOVPVTLQPQKQVKENKTQPPVAYQYWPPAELQYRPPPE	OYGYPGMPPA
NCCIT-3	(120)	SEVYQSTEILHCLGPDNQESTVQPMITSIPLNLWGRDLLQ	QWCAEITMPA
NCCIT-5	(210)	GOVPVTLQPQKQVKENKTQPPVAYQYWPPAELQYRPPPE	OYGYPGMPPA
NCCIT-6	(189)	GOVPVTLQPQKQVKENKTQPPVAYQYWPPAELQYRPPPE	OYGYPGMPPA
NCCIT-14	(231)	SEVYQSTEILHCLGPDNQESTVQPMITSIPLNLWGRDLLQ	QWCAEITMPA
NCCIT-15	(210)	GOVPVTLQPQKQVKENKTQPPVAYQYWPPAELQYRPPPE	OYGYPGMPSA
NCCIT-16	(219)	GOVPVTLQPQKQVKENKTQPPVAYQYWPPAELQYRPPPE	OYGYPGMPPA
NCCIT-17	(187)	GOVPVTLQPQKQVKENKTQPPVAYQYWPPAELQYRPPPE	OYGYPGMPPA
NCCIT-19	(120)	SEVYQSTEILHCLGPDNQESTVQPMITSIPLNLWGRDLLQ	QWCAEITMPA
NCCIT-18	(189)	GOVPVTLQPQKQVKENKTQPPVAYQYWPPAELQYRPPPE	OYGYPGMPPA
NCCIT-20	(120)	SEVYQSTEILHCLGPDNQESTVQPMITSIPLNLWGRDLLQ	QWCAEITMPA
NCCIT-21	(231)	SEVYQSTEILHCLGPDNQESTVQPMITSIPLNLWGRDLLQ	QWCAEITMPA
NCCIT-22	(120)	SEVYQSTEILHCLGPDNQESTVQPMITSIPLNLWGRDLLQ	QWCAEITMPA
NCCIT-24	(120)	SEVYQSTEILHCLGPDNQESTVQPMITSIPLNLWGRDLLQ	QWCAEITMPA
Mel 2-73	(120)	SEVYQSTEILHCLGPDNQESTVQPMITSIPLNLWGRDLLQ	QWCAEITMPA
Mel 2-74	(251)	SEVYQSTEILHCLGPDNQESTVQPMITSIPLNLWGRDLLQ	QWCAEVTMPA
Mel 2-80	(187)	GOVPVTLQPQKQVKENKTQPPVAYQYWPPAELQYRPPPE	OYGYPGMPPA
Mel 2-81	(187)	GOVPVTLQPQKQVKENKTQPPVAYQYWPPAELQYRPPPE	OYGYPGMPPA
Mel 2-82	(187)	GOVPVTLQPQKQVKENKTQPPVAYQYWPPAELQYRPPPE	OYGYPGMPPA
Mel 2-78	(187)	GOVPVTLQPQKQVKENKTQPPVAYQYWPPAELQYRPPPE	OYGYPGMPPA
Mel 2-95	(187)	GOVPVTLQPQKQVKENKTQPPVAYQYWPPAELQYRPPPE	OYGYPGMPPA
GH-36	(120)	SEVYQSTEILHCLGPDNQESTVQPMITSIPLNLWGRDLLQ	QWCAEITMPA
GH-39	(120)	SEVYQSTEILHCLGPDNQESTVQPMITSIPLNLWGRDLLQ	QWCAEITMPA
GH-40	(120)	SEVYQSTEILHCLGPDNQESTVQPMITSIPLNLWGRDLLQ	QWCAEITMPA
GH-41	(120)	SEVYQSTEILHCLGPDNQESTVQPMITSIPLNLWGRDLLQ	QWCAEITMPA

301 350

NCCIT-1 (260) P QGRAPYPQPP T RRLNPTAPPSRQGSSELHEIIDKSRKEGDTEAWQFPVTL
NCCIT-2 (260) P QGRAPYPQPP T RRLNPTAPPSRQGSSELHEIIDKSRKEGDTEAWQFPVTL
NCCIT-3 (170) P LYSPTSQKIM T KMGYIPGKGLGKNEDGIKVPVEAKINQEREGIGYPF--
NCCIT-5 (260) P QGRAPYPQPP T RRLNPTAPPSRQGSSELHEIIDKSRKEGDTEAWQFPVTL
NCCIT-6 (239) P QGRAPYPQPP T RRLNPTAPPSRQGSSELHEIIDKSRKERDTEAWQFPVTL
NCCIT-14 (281) P SYSPTSQKIM T KMGYIPGKGLGKNEDGIKIPVEAKINQEREGIGNPC--
NCCIT-15 (260) P QGRAPYPQPP T RRLNPTAPPSRQGSSELHEIIDKSRKEGDTEAWQFPVTL
NCCIT-16 (269) P QGRAPYPQPP T RRLNPTAPPSRQGSSELHEIIDKSRKEGDTEAWQFPVTL
NCCIT-17 (237) P QGRAPYPQPP T RRLNPTAPPSRQGSKLHEIIDESRKEGDTEAWQFPVTL
NCCIT-19 (170) P LYSPTSQKIM T KMGYIPGKGLGKNEDGIKVPVEAKINQEREGIGYPF--
NCCIT-18 (239) P QGRAPYPQPP T RRLNPTAPPSRQGSSELHEIIDKSRKERRY-----
NCCIT-20 (170) P LYSPTSQKIM T KMGYIPGKGLGKNEDGIKVPVEAKINQEREGIGYPF--
NCCIT-21 (281) P SYSPTSQKIM T KMGYIPGKGLGKNEDGIKIPVEAKINQEREGIGNPC--
NCCIT-22 (170) P LYSPTSQKIM T KMGYIPGKGLGKNEDGIKVPVEAKINQEREGIGYPF--
NCCIT-24 (170) P LYSPTSQKIM T KMGYIPGKGLGKNEDGIKVPVEAKINQEREGIGYPF--
Mel 2-73 (170) P SYSPTSQKIM T KMGYIPGKGLGKNEDGIKIPVEAKINQEREGIGNPC--
Mel 2-74 (301) P SYSPTSQKIM T KMGYIPGKGLGKNEDGIKIPVEAKINQEREGIGNPC--
Mel 2-80 (237) P QGRAPYPQPP T RRLNPTAPPSRQGSKLHEIIDESRKEGDTEAWQFPVTL
Mel 2-81 (237) P QGRAPYPQPP T RRLNPTAPPSRQGSKLHEIIDESRKEGDTEAWQFPVTL
Mel 2-82 (237) P QGRAPYPQPP T RRLNPTAPPSRQGSKLHEIIDESRKEGDTEAWQFPVTL
Mel 2-78 (237) P QGRAPYPQPP T RRLNPTAPPSRQGSKLHEIIDESRKEGDTEAWQFPVTL
Mel 2-95 (237) P QGRAPYPQPP T RRLNPTAPPSRQGSKLH-----
GH-36 (170) P LYSPTSQKIM T KMGYIPGKGLGKNEDGIKVPVEAKINQEREGIGYPF--
GH-39 (170) P LYSPTSQKIM T KMGYIPGKGLGKNEDGIKVPVEAKINQEREGIGYPF--
GH-40 (170) P LYSPTSQKIM T KMGYIPGKGLGKNEDGIKVPVEAKINQEREGIGYPF--
GH-41 (170) P LYSPTSQKIM T KMGYIPGKGLGKNEDGIKVPVEAKINQEREGIGYPF--

351 400

NCCIT-1 (310) EPMPPEGE-EPKRESLPQLSQQVFSIKMLKDMKEGVKQYGPNSPYMRHYR
NCCIT-2 (310) EPMPPEK-EPKRESLPQLSQQSFR--NAKRYEEGVKQYGPNSLI--D--
NCCIT-3 (218) -----
NCCIT-5 (310) EPMPPEK-EPKRESLPQLSQQSFR--NAKRYEEGVKQYGPNSLI--D--
NCCIT-6 (289) EPMPPEK-SPRGEPPPTVEPDTSLFR-----
NCCIT-14 (329) -----
NCCIT-15 (310) EPMPPEK-EPKRESLPQLSQQSFR-----
NCCIT-16 (319) EPMPPEGEGAQEGEPPPTVEARYSLFSIKMLKDMKEGVKQYGPPTPLI-----
NCCIT-17 (287) EPMPPEKSPRGRASHS-----
NCCIT-19 (218) -----
NCCIT-18 (280) -----
NCCIT-20 (218) -----
NCCIT-21 (329) -----
NCCIT-22 (218) -----
NCCIT-24 (218) -----
Mel 2-73 (218) -----
Mel 2-74 (349) -----
Mel 2-80 (287) EPMPPEKSPRGRASHS-----
Mel 2-81 (287) EPMPPEKSPRGRASHS-----
Mel 2-82 (287) EPMPPEKSPRGRASHS-----
Mel 2-78 (287) EPMPPEKSPRGRASHS-----
Mel 2-95 (266) -----
GH-36 (218) -----
GH-39 (218) -----
GH-40 (218) -----
GH-41 (218) -----

	401	450
NCCIT-1	(359)	FIAHDIDSSHGEILAIVFSHPHSLQLGLR--RVQERSENAATLGYHRMA
NCCIT-2	(353)	--IIMLMTTHLWEILAIVSLTSHLH---TGEWFRTVRDAASRNLCDHYRN
NCCIT-3	(218)	-----
NCCIT-5	(353)	--IIMLMTTHLWEILAIVSLTSHLH---TGEWFRTVRDAASRNLCDHYRN
NCCIT-6	(313)	-----
NCCIT-14	(329)	-----
NCCIT-15	(334)	-----
NCCIT-16	(364)	-----
NCCIT-17	(303)	-----
NCCIT-19	(218)	-----
NCCIT-18	(280)	-----
NCCIT-20	(218)	-----
NCCIT-21	(329)	-----
NCCIT-22	(218)	-----
NCCIT-24	(218)	-----
Mel 2-73	(218)	-----
Mel 2-74	(349)	-----
Mel 2-80	(303)	-----
Mel 2-81	(303)	-----
Mel 2-82	(303)	-----
Mel 2-78	(303)	-----
Mel 2-95	(266)	-----
GH-36	(218)	-----
GH-39	(218)	-----
GH-40	(218)	-----
GH-41	(218)	-----

	451	464
NCCIT-1	(407)	TSPLGNSVHKLLGG
NCCIT-2	(398)	WSIWGSYYV-----
NCCIT-3	(218)	-----
NCCIT-5	(398)	WSIWGSYYV-----
NCCIT-6	(313)	-----
NCCIT-14	(329)	-----
NCCIT-15	(334)	-----
NCCIT-16	(364)	-----
NCCIT-17	(303)	-----
NCCIT-19	(218)	-----
NCCIT-18	(280)	-----
NCCIT-20	(218)	-----
NCCIT-21	(329)	-----
NCCIT-22	(218)	-----
NCCIT-24	(218)	-----
Mel 2-73	(218)	-----
Mel 2-74	(349)	-----
Mel 2-80	(303)	-----
Mel 2-81	(303)	-----
Mel 2-82	(303)	-----
Mel 2-78	(303)	-----
Mel 2-95	(266)	-----
GH-36	(218)	-----
GH-39	(218)	-----
GH-40	(218)	-----
GH-41	(218)	-----

		301		350
C9_Env-1-2	(301)	CACAAGGATTAATATGAGATTTATAATCGTAGT	ACTTCCAATGGTCTTAA	
C9_Env-1-3	(301)	CACAAGGATTAATATGAGATTTATAATCGTAGT	CCTTCCAATGGTCTTAA	
C9_Env-1-1	(301)	CACAAGGATTAATATGAGATTTATAATCGTAGT	ACTTCCAATGGTCTTAA	
GH_Env-4	(300)	CACAAGGATTAATATGAGATTTATAATCGTAGT	ACTTCCAATGGTCTTAA	
GH_Env-12	(300)	CACAAGGATTAATATGAGATTTATAATCGTAGT	ACTTCCAATGGTCTTAA	
GH_Env-3	(301)	CACAAGGATTAATATGAGATTTATAATCGTAGT	ACTTCCAATGGTCTTAA	
GH_Env-11	(301)	CACAAGGATTAATATGAGATTTATAATCGTAGT	ACTTCCAATGGTCTTAA	
GH_Env-5	(301)	CACAAGGATTAATATGAGATTTATAATCGTAGT	ACTTCCAATGGTCTTAA	
GH_Env-8	(72)	CACAAGGATTAATATGAGATTTATAATCGTAGT	ACTTCCAATGGTCTTAA	
		351		400
C9_Env-1-2	(351)	CCCAAGTGACAGGGTTAAGATTTGCGAGGCCA	CAGCAACTCCTGCAATT	
C9_Env-1-3	(351)	CCCAAGTGACAGGGTTAAGATTTGCGAGGCCA	CAGCAACTCCTGCAATT	
C9_Env-1-1	(351)	CCCAAGTGACAGGGTTAAGATTTGCGAGGCCA	CAGCAACTCCTGCAATT	
GH_Env-4	(350)	CCCAAGTGACAGGGTTAAGATTTGCGAGGCCA	CAGCAACTCCTGCAATT	
GH_Env-12	(350)	CCCAAGTGACAGGGTTAAGATTTGCGAGGCCA	CAGCAACTCCTGCAATT	
GH_Env-3	(351)	CCCAAGTGACAGGGTTAAGATTTGCGAGGCCA	CAGCAACTCCTGCAATT	
GH_Env-11	(351)	CCCAAGTGACAGGGTTAAGATTTGCGAGGCCA	CAGCAACTCCTGCAACT	
GH_Env-5	(351)	CCCAAGTGACAGGGTTAAGATTTGCGAGGCCA	CAGCAACTCCTGCAATT	
GH_Env-8	(122)	CCCAAGTGACAGGGTTAAGATTTGCGAGGCCA	CAGCAACTCGTCAATT	
		401		450
C9_Env-1-2	(401)	GCCTCAGTTCCTGGCA--CCAAATTTAAATGGGCTTTTGATGCTTCGAAA		
C9_Env-1-3	(401)	GCCTCAGTTCCTGG--CACCAAATTTAAATGGGCTTTTGATGCTTCGAAA		
C9_Env-1-1	(401)	GCCTCAGTTCCTGG--CACCAAATTTAAATGGGCTTTTGATGCTTCGAAA		
GH_Env-4	(400)	GCCTCAGTTCCTGG--CACCAAATTTAAATGGGCTTTTGATGCTTCGAAA		
GH_Env-12	(400)	GCCTCAGTTCCTGG--CACCAAATTTAAATGGGCTTTTGATGCTTCGAAA		
GH_Env-3	(401)	GCCTCAGTTCCTGG--CACCAAATTTAAATGGGCTTTTGATGCTTCGAAA		
GH_Env-11	(401)	GCCTCAGTTCCTGG--CACCAAATTTAAATGGGCTTTTGATGCTTCGAAA		
GH_Env-5	(401)	GCCTCAGTTCCTGG--CACCAAATTTAAATGGGCTTTTGATGCTTCGAAA		
GH_Env-8	(172)	GCCTCAGTTCCTGG--CACCAAATTTAAATGGGCTTTTGATGCTTCGAAA		
		451		500
C9_Env-1-2	(449)	ATTTGTTCTTTTAATTTGGAAATGTC	TAAAGTGAGATTATCTTCTCTTCC	
C9_Env-1-3	(449)	ATTTGTTCTTTTAATTTGGAAATGTC	TAAAGTGAGATTATCTTCTCTTCC	
C9_Env-1-1	(449)	ATTTGTTCTTTTAATTTGGAAATGTC	TAAAGTGAGATTATCTTCTCTTCC	
GH_Env-4	(448)	ATTTGTTCTTTTAATTTGGAAATGTC	TAAAGTGAGATTATCTTCTCTTCC	
GH_Env-12	(448)	ATTTGTTCTTTTAATTTGGAAATGTC	TAAAGTGAGATTATCTTCTCTTCC	
GH_Env-3	(449)	ATTTGTTCTTTTAATTTGGAAATGTC	TAAAGTGAGATTATCTTCTCTTCC	
GH_Env-11	(449)	ATTTGTTCTTTTAATTTGGAAATGTC	TAAAGTGAGATTATCTTCTCTTCC	
GH_Env-5	(449)	ATTTGTTCTTTTAATTTGGAAATGTC	TAAAGTGAGATTATCTTCTCTTCC	
GH_Env-8	(220)	ATTTGTTCTTTTAATTTGGAAATGTT	TAAAGTGAGATTATCTTCTCTTCC	
		501		545
C9_Env-1-2	(499)	CTGTAGATGGCGTCTAACCATGTCCCAGTGATGCTCAGACTCATT		
C9_Env-1-3	(499)	CTGTAGATGGCGTCTAACCATGTCCCAGTGATGCTCAGACTCATT		
C9_Env-1-1	(499)	CTGTAGATGGCGTCTAACCATGTCCCAGTGATGCTCAGACTCATT		
GH_Env-4	(498)	CTGTAGATGGTGTCTAACCATGTCCCAGTGATGCTCAGACTCATT		
GH_Env-12	(498)	CTGTAGATGGTGTCTAACCATGTCCCAGTGATGCTCAGACTCATT		
GH_Env-3	(499)	CTGTAGATGGTGTCTAACCATGTCCCAGTGATGCTCAGACTCATT		
GH_Env-11	(499)	CTGTAGATGGTGTCTAACCATGTCCCAGTGATGCTCAGACTCATT		
GH_Env-5	(499)	CTGTAGATGGTGTCTAACCATGTCCCAGTGATGCTCAGACTCATT		
GH_Env-8	(270)	CTGTAGATGGTGTCTAACCATGTCCCAGTGATGCTCAGACTCATT		

Nukleotidalignment 2 aus dem *env* –Bereich zytoplasmatischer *mRNA*-Klone

		1		50
GH_3-18	(1)	---CTGAAATGCATCC--ACGCGTTGGGAGCTCTCCCATATGGTCGACCTG		
Mel_2_3-2	(1)	-----GCATC--ACGCGTTGGGAGCTCTCCCATATGGTCGACCTG		
Mel_2_3-3	(1)	---TGGACAGCATCC--ACGCGTTGGGAGCTCTCCCATATGGTCGACCTG		
NCCIT_3-25	(1)	--CCATACAGCATCC--ACGCGTTGGGAGCTCTCCCATATGGTCGACCTG		
Mel_2_3-5	(1)	CCCGAAATGCATCC--ACGCGTTGGGAGCTCTCCCATATGGTCGACCTG		
Mel_2_3-6	(1)	CCTTACCATGGCATCC--ACGCGTTGGGAGCTCTCCCATATGGTCGACCTG		
Mel_2_3-4	(1)	--TCTCCATGCATCC--ACGCGTTGGGAGCTCTCCCATATGGTCGACCTG		
NCCIT_3-26	(1)	-CCCTGAATGCATCCAACGCGTTGGGAGCTCTCCCATATGGTCGACCTG		
NCCIT_3-27	(1)	-CTTTAAATGCATCC--ACGCGTTGGGAGCTCTCCCATATGGTCGACCTG		

		51	100
		(47)	CAGGCGGCCGCGAATTCAGTAGTGATTGCTGTCAACGATACGCTAC---G
		(39)	CAGGCGGCCGCGAATTCAGTAGTGATTCCAGACTCCAGACTATAAACC
		(47)	CAGGCGGCCGCGAATTCAGTAGTGATTCCAGACTCCAGACTATAAACC
		(48)	CAGGCGGCCGCGAATTCAGTAGTGATTCCAGACTCCAGACTATAAACC
		(50)	CAGGCGGCCGCGAATTCAGTAGTGATTCCAGACTCCAGACTATAAACC
		(50)	CAGGCGGCCGCGAATTCAGTAGTGATTCCAGACTCCAGACTATAAACC
		(48)	CAGGCGGCCGCGAATTCAGTAGTGATTCCAGACTCCAGACTATAAACC
		(50)	CAGGCGGCCGCGAATTCAGTAGTGATTCCAGACTCCAGACTATAAACC
		(48)	CAGGCGGCCGCGAATTCAGTAGTGATTCCAGACTCCAGACTATAAACC
		101	150
		(94)	TAACGGCATGACAGTGT-----TTTTTTTAAATTTTAAA-
		(89)	TGAAAATCTGTAGATTGCTTACTTGCATTGATTCAACTTTTAAATGGCAAC
		(97)	TGAAAATCTGTAGATTGCTTACTTGCATTGATTCAACTTTTAAATGGCAAC
		(98)	TGAAAATCTGTAGATTGCTTACTTGCATTGATTCAACTTTTAAATGGCAAC
		(100)	TGAAAATCTGTAGATTGCTTACTTGCATTGATTCAACTTTTAAATGGCAAC
		(100)	TGAAAATCTGTAGATTGCTTACTTGCATTGATTCAACTTTTAAATGGCAAC
		(98)	TGAAAATCTGTAGATTGCTTACTTGCATTGATTCAACTTTTAAATGGCAAC
		(100)	TGAAAATCTGTAGATTGCTTACTTGCATTGATTCAACTTTTAAATGGCAAC
		(98)	TGAAAATCTGTAGATTGCTTACTTGCATTGATTCAACTTTTAAATGGCAAC
		151	200
		(143)	----TTTTATTAAAGT-ACAG--AGTTAACAAGTTTGTAGTTTTTTATAT
		(139)	ACCGTATTCTGCTGGTGAGAGCAAGAGAGGGCGTGTGGATCCCTGTGTCC
		(147)	ACCGTATTCTGCTGGTGAGAGCAAGAGAGGGCGTGTGGATCCCTGTGTCC
		(148)	ACCGTATTCTGCTGGTGAGAGCAAGAGAGGGCGTGTGGATCCCTGTGTCC
		(150)	ACCGTATTCTGCTGGTGAGAGCAAGAGAGGGCGTGTGGATCCCTGTGTCC
		(150)	ACCGTATTCTGCTGGTGAGAGCAAGAGAGGGCGTGTGGATCCCTGTGTCC
		(148)	ACCGTATTCTGCTGGTGAGAGCAAGAGAGGGCGTGTGGATCCCTGTGTCC
		(150)	ACCGTATTCTGCTGGTGAGAGCAAGAGAGGGCGTGTGGATCCCTGTGTCC
		(148)	ACCGTATTCTGCTGGTGAGAGCAAGAGAGGGCGTGTGGATCCCTGTGTCC
		201	250
		(186)	A-GGAAAAGCCTAGTCAA--TTCAGATGCTTTCTAGAAA--AATTAACA
		(189)	ATGGACCGACCGTGGGAGGCCTCACCATCCGTCATATTTTGACTGAAGT
		(197)	ATGGACCGACCGTGGGAGGCCTCGCCATCCGTCATATTTTGACTGAAGT
		(198)	ATGGACCGACCGTGGGAGGCCTCACCATCCGTCATATTTTGACTGAAGT
		(200)	ATGGACCGACCGTGGGAGGCCTCGCCATCCGTCATATTTTGACTGAAGT
		(200)	ATGGACCGACCGTGGGAGGCCTCGCCATCCGTCATATTTTGACTGAAGT
		(198)	ATGGACCGACCGTGGGAGGCCTCGCCATCCGTCATATTTTGACTGAAGT
		(200)	ATGGACCGACCGTGGGAGGCCTCACCATCCGTCATATTTTGACTGAAGT
		(198)	ATGGACCGACCGTGGGAGGCCTCACCATCCGTCATATTTTGACTGAAGT
		251	300
		(230)	TTAAAAAACAAATAGAAATCCATGACTAAAGGGGAAAAATAACTTTCAA
		(239)	ATTAAAAGGTGTTTTAAATAGATC-CAAAAGATTCATTTTACTTT-AAT
		(247)	ATTAAAAGGTGTTTTAAATAGATC-CAAAAGATTCATTTTACTTT-AAT
		(248)	ATTAAAAGGTGTTTTAAATAGATC-CAAAAGATTCATTTTACTTT-AAT
		(250)	ATTAAAAGGTGTTTTAAATAGATC-CAAAAGATTCATTTTACTTT-AAT
		(250)	ATTAAAAGGTGTTTTAAATAGATC-CAAAAGATTCATTTTACTTT-AAT
		(248)	ATTAAAAGGTGTTTTAAATAGATC-CAAAAGATTCATTTTACTTT-AAT
		(250)	ATTAAAAGGTGTTTTAAATAGATC-CAAAAGATTCATTTTACTTT-AAT
		(248)	ATTAAAAGGTGTTTTAAATAGATC-CAAAAGATTCATTTTACTTT-AAT
		301	350
		(280)	AGTTACCAAATTCGAATCATATCAGAGACCATTATAAATT-TCAAACAG
		(287)	TGCAGTGATTATGGGATTAATTGCAGTCACAGCTACGGGTGCTGTAGCAG
		(295)	TGCAGTGATTATGGGATTAATTGCAGTCACAGCTACGGGTGCTGTAGCAG
		(296)	TGCAGTGATTATGGGATTAATTGCAGTCACAGCTACGGGTGCTGTAGCAG
		(298)	TGCAGTGATTATGGGATTAATTGCAGTCACAGCTACGGGTGCTGTAGCAG
		(298)	TGCAGTGATTATGGGATTAATTGCAGTCACAGCTACGGGTGCTGTAGCAG
		(296)	TGCAGTGATTATGGGATTAATTGCAGTCACAGCTACGGGTGCTGTAGCAG
		(298)	TGCAGTGATTATGGGATTAATTGCAGTCACAGCTACGGGTGCTGTAGCAG
		(296)	TGCAGTGATTATGGGATTAATTGCAGTCACAGCTACGGGTGCTGTAGCAG

		351	400
		TAGAT--TTACCACACATAT-TGCATTT--TCAAATTC--TAATG--TAG	
	GH_3-18	(329)	GAGTTGCATTGCACTCTTCTGTTTCAGTCAGTAACTTTGTTAATGATTGG
	Mel 2_3-2	(337)	GAGTTGCATTGCACTCTTCTGTTTCAGTCAGTAACTTTGTTAATGATTGG
	Mel 2_3-3	(345)	GAGTTGCATTGCACTCTTCTGTTTCAGTCAGTAACTTTGTTAATGATTGG
	NCCIT_3-25	(346)	GAGTTGCATTGCACTCTTCTGTTTCAGTCAGTAACTTTGTTAATGATTGG
	Mel 2_3-5	(348)	GAGTTGCATTGCACTCTTCTGTTTCAGTCAGTAACTTTGTTAATGATTGG
	Mel 2_3-6	(348)	GAGTTGCATTGCACTCTTCTGTTTCAGTCAGTAACTTTGTTAATGATTGG
	Mel 2_3-4	(346)	GAGTTGCATTGCACTCTTCTGTTTCAGTCAGTAACTTTGTTAATGATTGG
	NCCIT_3-26	(348)	GAGTTGCATTGCACTCTTCTGTTTCAGTCAGTAACTTTGTTAATGATTGG
	NCCIT_3-27	(346)	GAGTTGCATTGCACTCTTCTGTTTCAGTCAGTAACTTTGTTAATGATTGG
		401	450
		CAAAACGTAACCACATAATT-TGG---CTACAGCTA--ATCGTTTCAGA	
	GH_3-18	(370)	CAAAAAAATTTCTACAAGATTGTGGAATTCACAATCTAGTATTGATCAAAA
	Mel 2_3-2	(387)	CAAAAAAATTTCTACAAGATTGTGGAATTCACAATCTAGTATTGATCAAAA
	Mel 2_3-3	(395)	CAAAAAAATTTCTACAAGATTGTGGAATTCACAATCTAGTATTGATCAAAA
	NCCIT_3-25	(396)	CAAAAAAATTTCTACAAGATTGTGGAATTCACAATCTAGTATTGATCAAAA
	Mel 2_3-5	(398)	CAAAAAAATTTCTACAAGATTGTGGAATTCACAATCTAGTATTGATCAAAA
	Mel 2_3-6	(398)	CAAAAAAATTTCTACAAGATTGTGGAATTCACAATCTAGTATTGATCAAAA
	Mel 2_3-4	(396)	CAAAAAAATTTCTACAAGATTGTGGAATTCACAATCTAGTATTGATCAAAA
	NCCIT_3-26	(398)	CAAAAGAATTTCTACAAGATTGTGGAATTCACAATCTAGTATTGATCAAAA
	NCCIT_3-27	(396)	CAAAAGAATTTCTACAAGATTGTGGAATTCACAATCTAGTATTGATCAAAA
		451	500
		A-----AA-----GTTTAAAAAATTAG-CAAA--GTTATATCTATA----	
	GH_3-18	(413)	ATTGGCAAATCAAATTAATGATCTTAGACAAACTGTCATTTGGATGGGAG
	Mel 2_3-2	(437)	ATTGGCAAATCAAATTAATGATCTTAGACAAACTGTCATTTGGATGGGAG
	Mel 2_3-3	(445)	ATTGGCAAATCAAATTAATGATCTTAGACAAACTGTCATTTGGATGGGAG
	NCCIT_3-25	(446)	ATTGGCAAATCAAATTAATGATCTTAGACAAACTGTCATTTGGATGGGAG
	Mel 2_3-5	(448)	ATTGGCAAATCAAATTAATGATCTTAGACAAACTGTCATTTGGATGGGAG
	Mel 2_3-6	(448)	ATTGGCAAATCAAATTAATGATCTTAGACAAACTGTCATTTGGATGGGAG
	Mel 2_3-4	(446)	ATTGGCAAATCAAATTAATGATCTTAGACAAACTGTCATTTGGATGGGAG
	NCCIT_3-26	(448)	ATTGGCAAATCAAATTAATGATCTTAGACAAACTGTCATTTGGATGGGAG
	NCCIT_3-27	(446)	ATTGGCAAATCAAATTAATGATCTTAGACAAACTGTCATTTGGATGGGAG
		501	550
		--AAACTTTTG---TAGT-----TTTCTTTTTCGAAAGT-----AAA	
	GH_3-18	(446)	ACAGACTCATGAGCTTAGAACATCGTTTCAGTTACAATGTGACTGGAAT
	Mel 2_3-2	(487)	ACAGACTCATGAGCTTAGAACATCGTTTCAGTTACAATGTGACTGGAAT
	Mel 2_3-3	(495)	ACAGACTCATGAGCTTAGAACATCGTTTCAGTTACAATGTGACTGGAAT
	NCCIT_3-25	(496)	ACAGACTCATGAGCTTAGAACATCGTTTCAGTTACAATGTGACTGGAAT
	Mel 2_3-5	(498)	ACAGACTCATGAGCTTAGAACATCGTTTCAGTTACAATGTGACTGGAAT
	Mel 2_3-6	(498)	ACAGACTCATGAGCTTAGAACATCGTTTCAGTTACAATGTGACTGGAAT
	Mel 2_3-4	(496)	ACAGACTCATGAGCTTAGAACATCGTTTCAGTTACAATGTGACTGGAAT
	NCCIT_3-26	(498)	ACAGACTCATGAGCTTAGAACATCGTTTCAGTTACAATGTGACTGGAAT
	NCCIT_3-27	(496)	ACAGACTCATGAGCTTAGAACATCGTTTCAGTTACAATGTGACTGGAAT
		551	600
		AAGGCTTAAATCTTTAATAAAGGAAAACAAAACAATCC-TCTTAA-ATTT	
	GH_3-18	(478)	ACGTCAGATTTTGTATTACACCCCAAATTTATAATGAGTCTGAGCATCA
	Mel 2_3-2	(537)	ACGTCAGATTTTGTATTACACCCCAAATTTATAATGAGTCTGAGCATCA
	Mel 2_3-3	(545)	ACGTCAGATTTTGTATTACACCCCAAATTTATAATGAGTCTGAGCATCA
	NCCIT_3-25	(546)	ACGTCAGATTTTGTATTACACCCCAAATTTATAATGAGTCTGAGCATCA
	Mel 2_3-5	(548)	ACGTCAGATTTTGTATTACACCCCAAATTTATAATGAGTCTGAGCATCA
	Mel 2_3-6	(548)	ACGTCAGATTTTGTATTACACCCCAAATTTATAATGAGTCTGAGCATCA
	Mel 2_3-4	(546)	ACGTCAGATTTTGTATTACACCCCAAATTTATAATGAGTCTGAGCATCA
	NCCIT_3-26	(548)	ACGTCAGATTTTGTATTACACCCCAAATTTATAATGAGTCTGAGCATCA
	NCCIT_3-27	(546)	ACGTCAGATTTTGTATTACACCCCAAATTTATAATGAGTCTGAGCATCA
		601	650
		CTTATAAATAGCT---CTCCAG--ACATATATTACAA----ATCTG-CTG	
	GH_3-18	(526)	CTGGGACATGGTTAGACCCATCTACAGGGAAGAGAAGATAATCTCACTT
	Mel 2_3-2	(587)	CTGGGACATGGTTAGACCCATCTACAGGGAAGAGAAGATAATCTCACTT
	Mel 2_3-3	(595)	CTGGGACATGGTTAGACCCATCTACAGGGAAGAGAAGATAATCTCACTT
	NCCIT_3-25	(596)	CTGGGACATGGTTAGACCCATCTACAGGGAAGAGAAGATAATCTCACTT
	Mel 2_3-5	(598)	CTGGGACATGGTTAGACCCATCTACAGGGAAGAGAAGATAATCTCACTT
	Mel 2_3-6	(598)	CTGGGACATGGTTAGACCCATCTACAGGGAAGAGAAGATAATCTCACTT
	Mel 2_3-4	(596)	CTGGGACATGGTTAGACCCATCTACAGGGAAGAGAAGATAATCTCACTT
	NCCIT_3-26	(598)	CTGGGACATGGTTAGACCCATCTACAGGGAAGAGAAGATAATCTCACTT
	NCCIT_3-27	(596)	CTGGGACATGGTTAGACCCATCTACAGGGAAGAGAAGATAATCTCACTT

		651	700
	GH_3-18	(566)	TAAGCTTTCCTTACCTGAGAGAAC---TTCCAGGATCCTTTATCCCAA
	Mel 2_3-2	(637)	TAGACATTTCCAAATTAAGAAGAACAAATTTTCGAAGCATCAAAGCCCAT
	Mel 2_3-3	(645)	TAGACATTTCCAAATTAAGAAGAACAAATTTTCGAAGCATCAAAGCCCAT
	NCCIT_3-25	(646)	TAGACATTTCCAAATTAAGAAGAACAAATTTTCGAAGCATCAAAGCCCAT
	Mel 2_3-5	(648)	TAGACATTTCCAAATTAAGAAGAACAAATTTTCGAAGCATCAAAGCCCAT
	Mel 2_3-6	(648)	TAGACATTTCCAAATTAAGAAGAACAAATTTTCGAAGCATCAAAGCCCAT
	Mel 2_3-4	(646)	TAGACATTTCCAAATTAAGAAGAACAAATTTTCGAAGCATCAAAGCCCAT
	NCCIT_3-26	(648)	TAGACATTTCCAAATTAAGAAGAACAAATTTTCGAAGCATCAAAGCCCAT
	NCCIT_3-27	(646)	TAGACATTTCCAAATTAAGAAGAACAAATTTTCGAAGCATCAAAGCCCAT
		701	750
	GH_3-18	(612)	AGGATTACCTTAAAGAGTTCTTCCATCATT-----TACTCATG---T
	Mel 2_3-2	(687)	TTAAATTTGGTGCCAGGAACTGAGGCAATTGCAGGAGTTGCTGATGGCCT
	Mel 2_3-3	(695)	TTAAATTTGGTGCCAGGAACTGAGGCAATTGCAGGAGTTGCTGATGGCCT
	NCCIT_3-25	(696)	TTAAATTTGGTGCCAGGAACTGAGGCAATTGCAGGAGTTGCTGATGGCCT
	Mel 2_3-5	(698)	TTAAATTTGGTGCCAGGAACTGAGGCAATTGCAGGAGTTGCTGATGGCCT
	Mel 2_3-6	(698)	TTAAATTTGGTGCCAGGAACTGAGGCAATTGCAGGAGTTGCTGATGGCCT
	Mel 2_3-4	(696)	TTAAATTTGGTGCCAGGAACTGAGGCAATTGCAGGAGTTGCTGATGGCCT
	NCCIT_3-26	(698)	TTAAATTTGGTGCCAGGAACTGAGGCAATTGCAGGAGTTGCTGATGGCCT
	NCCIT_3-27	(696)	TTAAATTTGGTGCCAGGAACTGAGGCAATTGCAGGAGTTGCTGATGGCCT
		751	800
	GH_3-18	(652)	GAATATGATTAAAC---TCCTATAG---AAGTGGATTGGGACATATGC-A
	Mel 2_3-2	(737)	CGCAAATCTTAACCTGTCACTTGGTTAAGACCATTGGAAGTACTACGA
	Mel 2_3-3	(745)	CGCAAATCTTAACCTGTCACTTGGTTAAGACCATTGGAAGTACTACGA
	NCCIT_3-25	(746)	CGCAAATCTTAACCTGTCACTTGGTTAAGACCATTGGAAGTACTACGA
	Mel 2_3-5	(748)	CGCAAATCTTAACCTGTCACTTGGTTAAGACCATTGGAAGTACTACGA
	Mel 2_3-6	(748)	CGCAAATCTTAACCTGTCACTTGGTTAAGACCATTGGAAGTACTACGA
	Mel 2_3-4	(746)	CGCAA-TCCTAACCTGTCACTTGGTTAAGACCATTGGAAGTACTACGA
	NCCIT_3-26	(748)	CGCAAATCTTAACCTGTCACTTGGTTAAGACCATTGGAAGTACTACGA
	NCCIT_3-27	(746)	CGCAAATCTTAACCTGTCACTTGGTTAAGACCATTGGAAGTACTACGA
		801	850
	GH_3-18	(695)	TTCTTAATCTGCCCTTCC-CCATTTGTTTCT-TTCTGAAAGGATTTTG-C
	Mel 2_3-2	(787)	TTATAAATCTCATATTAATCCTTGTGTGCCTGTTTGTCTGTTGTTAGTC
	Mel 2_3-3	(795)	TTATAAATCTCATATTAATCCTTGTGTGCCTGTTTGTCTGTTGTTAGTC
	NCCIT_3-25	(796)	TTATAAATCTCATATTAATCCTTGTGTGCCTGTTTGTCTGTTGTTAGTC
	Mel 2_3-5	(798)	TTATAAATCTCATATTAATCCTTGTGTGCCTGTTTGTCTGTTGTTAGTC
	Mel 2_3-6	(798)	TTATAAATCTCATATTAATCCTTGTGTGCCTGTTTGTCTGTTGTTAGTC
	Mel 2_3-4	(795)	TTATAAATCTCATATTAATCCTTGTGTGCCTGTTTGTCTGTTGTTAGTC
	NCCIT_3-26	(798)	TTATAAATCTCATATTAATCCTTGTGTGCCTGTTTGTCTGTTGTTAGTC
	NCCIT_3-27	(796)	TTATAAATCTCATATTAATCCTTGTGTGCCTGTTTGTCTGTTGTTAGTC
		851	900
	GH_3-18	(742)	TTAAGG--AAAAAAAAAGCTCTTGGTAAAGGCCAAATATTTCAACCTTT
	Mel 2_3-2	(837)	TGCAGGTGTAACCAACAGCTCCGAAGAGACAGCGACCATCGAGAACGGGC
	Mel 2_3-3	(845)	TGCAGGTGTAACCAACAGCTCCGAAGAGACAGCGACCATCGAGAACGGGC
	NCCIT_3-25	(846)	TGCAGGTGTAACCAACAGCTCCGAAGAGACAGCGACCATCGAGAACGGGC
	Mel 2_3-5	(848)	TGCAGGTGTAACCAACAGCTCCGAAGAGACAGCGACCATCGAGAACGGGC
	Mel 2_3-6	(848)	TGCAGGTGTAACCAACAGCTCCGAAGAGACAGCGACCATCGAGAACGGGC
	Mel 2_3-4	(845)	TGCAGGTGTAACCAACAGCTCCGAAGAGACAGCGACCATCGAGAACGGGC
	NCCIT_3-26	(848)	TGCAGGTGTAACCAACAGCTCCGAAGAGACAGCGACCATCGAGAACGGGC
	NCCIT_3-27	(846)	TGCAGGTGTAACCAACAGCTCCGAAGAGACAGCGACCATCGAGAACGGGC
		901	950
	GH_3-18	(790)	CAAAATGAC--TG-CC--TCTGT--GAAAGAGTTGTTGAGAAAGAAGAAA
	Mel 2_3-2	(887)	CATGATGACGATGGCGGTTTGTTCGAAAAGAAAAGGGGAAAATGTGGGGA
	Mel 2_3-3	(895)	CATGATGACGATGGCGGTTTGTTCGAAAAGAAAAGGGGAAAATGTGGGGA
	NCCIT_3-25	(896)	CATGATGACGATGGCGGTTTGTTCGAAAAGAAAAGGGG-AAAATGTGGGGA
	Mel 2_3-5	(898)	CATGATGACGATGGCGGTTTGTTCGAAAAGAAAAGGGGAAAATGTGGGGA
	Mel 2_3-6	(898)	CATGATGACGATGGCGGTTTGTTCGAAAAGAAAAGGGG-AAAATGTGGGGA
	Mel 2_3-4	(895)	CATGATGACGATGGCGGTTTGTTCGAAAAGAAAAGGGGAAAATGTGGGGA
	NCCIT_3-26	(898)	CATGATGACGATGGCGGTTTGTTCGAAAAGAAAAGGGGAAAATGTGGGGA
	NCCIT_3-27	(896)	CATGATGACGATGGCGGTTTGTTCGAAAAGAAAAGGGGAAAATGTGGGGA

		951	1000
		AGAGAGAGAGAGAGAAAGGTCTAAACATCTGTGTGAACAGCTCTCCAGTA	
		AAAGCAAGAGAGATCAGATTGTTACTGTGTCTGTGTAGAAAGAAGTAGAC	
		AA-GCAAGAGAGATCAGATTGTTACTGTGTCTGTGTAGAAAGAAGTAGAC	
		AAAGCAAGAGAGATCAGATTGTTGCTGTGTCTGTGTAGAAAGAAGTAGAC	
		AAAGCAAGAGAGATCAGATTGTTACTGTGTCTGTGTAGAAAGGAGTAGAC	
		AAAGCAAGAGAGATCAGATTGTTACTGTGTCTGTGTAGAAAGAAGTAGAC	
		AAAGCAAGAGAGATCAGATTGTTAC-----TGTAGAAAGAAGTAGAC	
		AAAGCAAGAGAGATCAGATTGTTACTGTGTCTGTGTAGAAAGAAGTAGAC	
		AAAGCAAGAGAGATCAGATTGTTACTGTGTCTGTGTAGAAAGAAGTAGAC	

		1001	1050
		CTGTGA-AGTCAAAGGCCAACATTACAGAGCGCA---CCTCTGCCT-GA	
		ATAGGAGACTCCATTTT-GTTATGTGCTAAGAAAAATTCTTCTGCCT-GA	
		ATAGGAGACTCCATTTT-GTTATGTACTAAGAAAAATTCTTCTGCCTTGA	
		ATAGGAGACTCCATTTT-GTTATGTACTAAGAAAAATTCTTCTGCCTTGA	
		ATAGGAGACTCCATTTT-GTTATGTACTAAGAAAAATTCTTCTGCCTTGA	
		ATAGGAGACTCCATTTT-GTTATGTACTAAGAAAAATTCTTCTGCCTTGA	
		ATAGGAGACTCCATTTT-GTAAATGTACTAAGAAAA-TTCTTCTGCCTTGA	
		ATGGGAGACTCCATTTTGTATGTGTTA-GAAAAATTCTTCTGCCTTGA	
		ATGGGAGACTCCATTTT-GTTATGTGTAAAGAAAAATTCTTCTGCCTTGA	

		1051	1100
		AATAC---AAACTA--AGTTAACTAGCAA-----GTTACAGGAAA--T	
		GAT-CTGTAA-TCTATAACCTTACCCCCACCCCGGTGCTCTCTGAAACGT	
		GATTCTGTTA-TCTATGACCTTACCCCCAACCTCGTGCTCTCTGAAACAT	
		GATTCTGTTA-TCTATGACCTTACCCCCAACCCCGTGCTCTCTGAAACAT	
		GATTCTGTTAATCTATGACCTTACCCCCAACCCCGTGCTCTCTGAAACAT	
		GATTCTGTTAATCTATGACCTTACCCCCAACCCCGTGCTCTCTGAAACAT	
		GATTCTGTTAATCTATGACCTTACCCCCAACCCCGTGCTCTCTGAAACGT	
		GATTCTGTTAATCTATGACCTTACCCCCAACCCCGTGCTCTCTGAAACGT	

		1101	1150
		AGCCT-----TCAG---TAAATTCCACAAGCCAAGTG--GCTACTGC	
		GTGCTGTGTCAACTCAGAGTA-AATG-AT--AGGCTGTGCAG-AT-TGC	
		GTGCTGTGTCCAACCTCAGAGTTGAATG----ATAGGCGGTGCAG-ATGTGC	
		GTGCTGTGTCTACTCAGAGTTGAATGGATTAAGGGCGGTGCAGGATGTGC	
		GTGCTGTGTCCAACCTCAGAGTTGAATGGATTAAGGGCGGTGCAGGATGTGC	
		GTGCTGTGTCCAACCTCAGAGTTGAATGGATTAAGGGCGGTGCAGGATGTGC	
		GTGCTGTGTCCAACCTCAGGTTGAATGGATTAAGGGCGGTGCAGGATGTGC	
		GTGCTGTGTCCAACCTCAGGTTGAATGGATTAAGGGCGGTGCAGGATGTGC	

		1151	1200
		ATTGT--CCCTGAAGAAGGAGGCCCCAGTGTCTTTCTGGGTGTGTAAGG	
		TTTGT-AGGCAGATGCCTGGAG-CAGC-TGCTC--TTAGG--TCATCACC	
		TTTGTAAACAGATGCTTGAAGGCAGCATGCTCC-TTAAGAGTCATCACC	
		TTTGTAAACAGATGCTTGAAGGCAGCATGCTCC-TTAAGAGTCATCACC	
		TTTGTAAACAGATGCTTGAAGGCAGCATGCTCC-TTAAGAGTCATCACC	
		TTTGTAAACAGATGCTTGAAGGCAGCATGCTCC-TTAAGAGTCATCACC	
		TTTGTAAACAGATGCTTGAAGGCAGCATGCTCC-TTAAGAGTCATCACC	
		TTTGTAAACAGATGCTTGAAGGCAGCATGCTCC-TTAAGAGTCATCACC	
		TTTGTAAACAGATGCTTGAAGGCAGCATGCTCC-TTAAGAGTCATCACC	

		1201	1250
		TCTTAC-TTAGTT-CAAGGTTT-----GT--CCCT---TCTTACTT--GT	
		ACTCCTATCAGGTACAAGGACCAAACTGGAAAGCCTGGACCTGCCTAGAGG	
		ACTCCC-TAATCT-CAAGTACCCA---GGGACACAAAAA CTGCGGA-AGG	
		ACTCCC-TAATCT-CAAGTACCCA---GGGACACAAAAA CTGCGGA-AGG	
		ACTCCC-TAATCT-CAAGTACCCA---GGGACACAAAAA CTGCGGA-AGG	
		ACTCCC-TAATCT-CAAGTACCCA---GGGACACAAAAA CTGCGGA-AGG	
		ACTCCC-TAATCT-CAAGTACCCA---GG-ACACAAAAA CTGCGGA-AGG	
		ACTCCC-TAATCT-CAAGTACCCA---GG-ACACAAAAA CTGCGGA-AGG	

```

1251                                     1300
GH_3-18 (1086) AGGTTGGGCTCTCTTTCGTTTCCAAG--AGGTATAGGACACTAGAGTACA
Mel 2_3-2 (1220) CAGATGTTCACTCTGCCTTAGAAAAGCCAGG-ATTGTCCA--AG--TTTT
Mel 2_3-3 (1230) CCGCAGGGACCTCTGCCTAGGAAA-GCCAGGTATTGTCCA--AGGTTTCT
NCCIT_3-25 (1238) CCGCAGGGACCTCTGCCTAGGAAAAGCCAGGTATTGTCCA--AGGTTTCT
Mel 2_3-5 (1240) CCGCAGGGACCTCTGCCTAGGAAA-GCCAGGTATTGTCCA--AGGTTTCT
Mel 2_3-6 (1239) CCGCAGGGACCTCTGCCTAGGAAA-GCCAGGTATTGTCCA--AGGTTTCT
Mel 2_3-4 (1228) CCGCAGGGACCTCTGCCTAGGAAA-GCCAGGTATTGTCCA--AGGTTTCT
NCCIT_3-26 (1239) CCGCAGGGACCTCTGCCTAGGAAA-GCCAGGTATTGTCCA--AGGTTTCT
NCCIT_3-27 (1237) CCGCAGGGACCTCTGCCTAGGAAA-GCCAGGTATTGTCCA--AGGTTTCT

```

```

1301                                     1350
GH_3-18 (1134) CCAA-ATGAGATACT----A----T-----TTGGAAGGCAATAAAGGG
Mel 2_3-2 (1265) CCCC-ATGGGATAGT--TAAA--ATAGGCCCTGGGAAGAGAAAGAACTG
Mel 2_3-3 (1277) CCCC-ATGTGAAAGTCTGAAA-TATGGCCTCGTGAGAAGAGAAAGACCTG
NCCIT_3-25 (1286) CCCCATGTGAAAGTCTGAAAATATGGCCTCGTGGGAGGGGAAAGACCTG
Mel 2_3-5 (1287) CCCC-ATGTGAAAGTCTGAAA-TATGGCCTCGTGGGAAGGGAAAGACCTG
Mel 2_3-6 (1286) CCCC-ATGTGAAAGTCTGAAA-TATGGCCTCGTGGGAAGGGAAAGACCTG
Mel 2_3-4 (1275) CCCC-ATGTGAAAGTCTGAAA-TATGGCCTCGTGGGAAGGGAAAGACCTG
NCCIT_3-26 (1286) CCCC-ATGTGATAGTCTGAAA-TATGGCCTCGTGGGAAGGGAAAGACCTG
NCCIT_3-27 (1284) CCCC-ATGTGATAGTCTGAAA-TATGGCCTCGTGGGAAGGGAAAGACCTG

```

```

1351                                     1400
GH_3-18 (1169) AGTGT-----AAAAGCCT-----AGGTAGCTTAA
Mel 2_3-2 (1310) ACGTT---TCAG--CGACAAC-GTAAAGGGTCTGTG-TGGGGAGGATT-G
Mel 2_3-3 (1325) ACCGTCCCCCAGCCCAGACCCCGTAAAGGGTCTGTGCTGAGGAGGATTAG
NCCIT_3-25 (1336) ACCGTCCCCCAGCCCAGACCCCGTAAAGGGTCTGTGCTGAGGAGGATTAG
Mel 2_3-5 (1335) ACCGTCCCCCAGCCCAGACCCCGTAAAGGGTCTGTGCTGAGGAGGATTAG
Mel 2_3-6 (1334) ACCGTCCCCCAGCCCAGACCCCGTAAAGGGTCTGTGCTGAGGAGGATTAG
Mel 2_3-4 (1323) ACCGTCCCCCAGCCCAGACCCCGTAAAGGGTCTGTGCTGAGGAGGATTAG
NCCIT_3-26 (1334) ACCGTCCCCCAGCCCAGACCCCGTAAAGGGTCTGTGCTGAGGAGGATTAG
NCCIT_3-27 (1332) ACCGTCCCCCAGCCCAGACCCCGTAAAGGGTCTGTGCTGAGGAGGATTAG

```

```

1401                                     1450
GH_3-18 (1193) ATACAGGCTCTGGAATTTCCCTCGCAG--G---AAGGTGAAG-CACCAA
Mel 2_3-2 (1352) TAAAAGAGAAAGGAAT-CCTCTTGCAGTGG-GACAAGAGGAAGGCTCTG
Mel 2_3-3 (1375) TAAAAGAGGAAGGAATGCCCTTGCAGTTGAGACAAGAGGAAGGCATCTG
NCCIT_3-25 (1386) TAAAAGAGGAAGGAATGCCCTTGCAGTTGAGACAAGAGGAAGGCATCTG
Mel 2_3-5 (1385) TAAAAGAGGAAGGAATGCCCTTGCAGTTGAGACAAGAGGAAGGCATCTG
Mel 2_3-6 (1384) TAAAAGAGGAAGGAATGCCCTTGCAGTTGAGACAAGAGGAAGGCATCTG
Mel 2_3-4 (1373) TAAAAGAGGAAGGAATGCCCTTGCAGTTGAGACAAGAGGAAGGCATCTG
NCCIT_3-26 (1384) TATAAGAGGAAGGAATGCCCTTGCAGTTGAGACAAGAGGAAGGCATCTG
NCCIT_3-27 (1382) TATAAGAGGAAGGAATGCCCTTGCAGTTGAGACAAGAGGAAGGCATCTG

```

```

1451                                     1500
GH_3-18 (1236) TGA--TGGCAAAAATTAATAA--ATACATCATCTTAAAAAGGCAG-AGG
Mel 2_3-2 (1400) TCTCCTGCTGTTCCTGGGCAATGGAATGTCCTGTATAAAAACCGGATG
Mel 2_3-3 (1425) TCTCCTGCTGTCCCTGGGCAATGGAATGTCCTCGGTATAAAAACCG-ATT
NCCIT_3-25 (1436) TCTCCTGCTGTCCCTGGGCAATGGAATGTCCTCGGTATAAAAACCG-ATT
Mel 2_3-5 (1435) TCTCCTGCTGTCCCTGGGCAATGGAATGTCCTCGGTATAAAAACCG-ATT
Mel 2_3-6 (1434) TCTCCTGCTGTCCCTGGGCAATGGAATGTCCTCGGTATAAAAACCG-ATT
Mel 2_3-4 (1423) TCTCCTGCTGTCCCTGGGCAATGGAATGTCCTCGGTATAAAAACCG-ATT
NCCIT_3-26 (1434) TCTCCTGCTGTCCCTGGGCAATGGAATGTCCTCGGTATAAAAACCG-ATT
NCCIT_3-27 (1432) TCTCCTGCTGTCCCTGGGCAATGGAATGTCCTCGGTATAAAAACCG-ATT

```

```

1501                                     1550
GH_3-18 (1281) GTAGGCTGCGTGCAGTGGTTTACGCCTATAATCCC--AGCACTTGGGAA
Mel 2_3-2 (1450) GTATGTTCCATCTACTGAGATAGGGAAAAACCCCTTAGGGCTGGAGGTG
Mel 2_3-3 (1474) GTATGTTCCATCTACTGAGATAGGGAAAAACCCCTTAGGGCTGGAGGTG
NCCIT_3-25 (1485) GTATGTTCCATCTACTGAGATAGGGAAAAACCCCTTAGGGCTGGAGGTG
Mel 2_3-5 (1484) GTATGTTCCATCTACTGAGATAGGGAAAAACCCCTTAGGGCTGGAGGTG
Mel 2_3-6 (1483) GTATGTTCCATCTACTGAGATAGGGAAAAACCCCTTAGGGCTGGAGGTG
Mel 2_3-4 (1472) GTATGTTCCATCTACTGAGATAGGGAAAAACCCCTTAGGGCTGGAGGTG
NCCIT_3-26 (1483) GTATGTTCCATCTACTGAGATAGGGAAAAACCCCTTAGGGCTGGAGGTG
NCCIT_3-27 (1481) GTATGTTCCATCTACTGAGATAGGGAAAAACCCCTTAGGGCTGGAGGTG

```

		1551		1600
	GH_3-18	(1329)	GCCAAGGCGGGC---ACATTGCTT-----GAGCTCAGGAGTTTGAGAC	
	Mel 2_3-2	(1500)	G-ACCTGCGGGCCGCAATATTGCTTTGTAAAGCACTGAGAAGTTTATGTG	
	Mel 2_3-3	(1524)	GGACCTGCGGGCAGCAATACTGCTTTGTAAAGCACTGAGATGTTTATGTG	
	NCCIT_3-25	(1535)	GGACCTGCGGGCAGCAATACTGCTTTGTAAAGCACTGAGATGTTTATGTG	
	Mel 2_3-5	(1534)	GGACCTGCGGGCAGCAATACTGCTTTGTAAAGCACTGAGATGTTTATGTG	
	Mel 2_3-6	(1533)	GGACCTGCGGGCAGCAATACTGCTTTGTAAAGCACTGAGATGTTTATGTG	
	Mel 2_3-4	(1522)	GGACCTGCGGGCAGCAATACTGCTTTGTAAAGCACTGAGATGTTTATGTG	
	NCCIT_3-26	(1533)	GGACCTGCGGGCAGCAGTACTGCTTTGTAAAGCACTGAGATGTTTATGTG	
	NCCIT_3-27	(1531)	GGACCTGCGGGCAGCAATACTGCTCTGTAAAGCACTGAGATGTTTATGTG	
		1601		1650
	GH_3-18	(1369)	CAGCC---TGGACAACCCCATCTCTACAAAATAAAATCGGCCAGGT	
	Mel 2_3-2	(1549)	TATGCATTTT--AAAGCCCAGCCCT---ATCCCTTACATTGTTTATGAT	
	Mel 2_3-3	(1574)	TATGCATATTAAAGAGCCCAGCACTT--AATCCTTACATTGTTTATGAT	
	NCCIT_3-25	(1585)	TATGCATATTAAAGAGCACAGCACTT--AATCCTTACATTGTTTATGAT	
	Mel 2_3-5	(1584)	TATGCATATCTAAAAGCACAGCACTT--AATCCTTACATTGTTTATGAT	
	Mel 2_3-6	(1583)	TATGCATATCTAAAAGCACAGCACTT--AATCCTTACATTGTTTATGAT	
	Mel 2_3-4	(1572)	TATGCATATCTAAAAGCACAGCACTT--AATCCTTACATTGTTTATGAT	
	NCCIT_3-26	(1583)	TATGCATATCTAAAAGCACAGCACTT--AATCCTTACATTGTTTATGAT	
	NCCIT_3-27	(1581)	TATGCATATCTAAAAGCACAGCACTT--AATCCTTACATTGTTTATGAT	
		1651		1700
	GH_3-18	(1416)	GTGATG-----GGGCAAACTGTGG-----TCTCAGCT-ACTTG	
	Mel 2_3-2	(1593)	-----CTCCCCCAATTG	
	Mel 2_3-3	(1622)	GCAAAGACCTTTGTTTCACATGTTTGTCTGCTGACCCTCTCCCCCAATTG	
	NCCIT_3-25	(1633)	GCAAAGACCTTTGTTTCACATGTTTGTCTGCTGACCCTCTCCCCCAATTG	
	Mel 2_3-5	(1632)	GCAAAGACCTTTGTTTCACATGTTTGTCTGCTGACCCTCTCCCCCAATTG	
	Mel 2_3-6	(1631)	GCAAAGACCTTTGTTTCACATGTTTGTCTGCTGACCCTCTCCCCCAATTG	
	Mel 2_3-4	(1620)	GCAAAGACCTTTGTTTCACATGTTTGTCTGCTGACCCTCTCCCCCAATTG	
	NCCIT_3-26	(1631)	GCAAAGACCTTTGTTTCACATGTTTGTCTGCTGACCCTCTCCCCCAATTG	
	NCCIT_3-27	(1629)	GCAAAGACCTTTGTTTCACATGTTTGTCTGCTGACCCTCTCCCCCAATTG	
		1701		1750
	GH_3-18	(1449)	GGAGGTGGTCCCTT--GAGCCCA----GGAGGTCAAACCTGCCATGAGT	
	Mel 2_3-2	(1606)	T-TTGTGGCCCG-CCCCTCCCCTTTTCGGG---CCCCCCGATGA-T	
	Mel 2_3-3	(1672)	TTTTGTGACCTTGACACATCCCCTTTTCGAGAAACCCCCAGATGA-T	
	NCCIT_3-25	(1683)	TTTTGTGACCTTGACACATCCCCTTTTCGAGAAACCCCCAGATGA-T	
	Mel 2_3-5	(1682)	TCTTGTGACCTTGACACATCCCCTTTCGAGAAACACCCACAGATGA-T	
	Mel 2_3-6	(1681)	TCTTGTGACCTTGACACATCCCCTTTCGAGAAACACCCACAGATGA-T	
	Mel 2_3-4	(1670)	TCTTGTGACCTTGACACATCCCCTTTCGAGAAACACCCACAGATGA-T	
	NCCIT_3-26	(1681)	TCTTGTGACCTTGACACATCCCCTTTCGAGAAACACCCACAGATGA-T	
	NCCIT_3-27	(1679)	TCTTGTGACCTTGACACATCCCCTTTCGAGAAACACCCACAGATGA-T	
		1751		1800
	GH_3-18	(1491)	CAAGATCATGTCACTGCACTCC-----AGCGTG----GCTGA	
	Mel 2_3-2	(1650)	CAGTC--TTCTAAGGGAAGCTCAGAGCGGGTCGGGATC-TGCATA-GCTGA	
	Mel 2_3-3	(1721)	CAATAAATAATTAAGGGAAGCTCAGGGGCGGGGGGATCCCCCATATGTTGA	
	NCCIT_3-25	(1732)	CAATAAATAATTAAGGGAAGCTCAGGGGCGGGGGGATCCTCCATATGTTGA	
	Mel 2_3-5	(1731)	CAATAAATACTAAGGGAAGCTCA-----AA-----	
	Mel 2_3-6	(1730)	CAATAAATACTAAGGGAAGCTCA-----AA-----	
	Mel 2_3-4	(1719)	CAATAAATACTAAGGGAAGCTCAGAGGCTGGCGGGATCCTCCATATGCTGA	
	NCCIT_3-26	(1730)	CAATAAATACTAAGGGAAGCTCAGAGGCTGGCGGGATCCTCCATATGCTGA	
	NCCIT_3-27	(1728)	CAATAAATACTAAGGGAAGCTCAGAGGCTGGCGGGATCCTCCATATGCTGA	
		1801		1850
	GH_3-18	(1524)	CAGA-----GTGAGCCCG--T-----CTCCAAAAA	
	Mel 2_3-2	(1696)	ACGATTTTGATAGGATTCACATACAGCGAGAGCATTGAAACTAAAAAAA	
	Mel 2_3-3	(1771)	ACGCGGGTTCCCCGGGTCCCCTTTT-----TTTTTTTTTAAAAAAA	
	NCCIT_3-25	(1782)	ACGCGGGTTCCCCGGGTCCCCTTTT-----TTTTTTTTTAAAAAAA	
	Mel 2_3-5	(1755)	-----AAAAAA	
	Mel 2_3-6	(1754)	-----AAAAAA	
	Mel 2_3-4	(1769)	ACGCTGGTTCCCCGGGTCCCCTTCT-----TTCTTCTCTAAAAAA	
	NCCIT_3-26	(1780)	ACGCTGGTTCCCCGGGTCCCCTTAT-----TTCTTCTCTAAAAAA	
	NCCIT_3-27	(1778)	ACGCTGGTTCCCCGGGTCCCCTTAT-----TTCTTCTCTAAAAAA	

		1851		1900
	GH_3-18	(1548)	AAAA----AAAAAAAAAAAAA	-----
	Mel 2_3-2	(1746)	AAAAAAAAAAAAAAAAAAAAA	AAAAAAAAAAAAAAAAAAAAA
	Mel 2_3-3	(1812)	AAAAAAAAAAAAAAAAAAAAA	AAAAAAAAAAAAAAAAAAAAA
	NCCIT_3-25	(1823)	AAAAAAAAAAAAAAAAAAAAA	AAAAAAAAAAAAAAAAAAAAA
	Mel 2_3-5	(1761)	AAAA----AAAAAAAAAAAAA	-----
	Mel 2_3-6	(1760)	AAAA----AAAAAAAAAAAAA	-----
	Mel 2_3-4	(1810)	AAAACAACAAAAAAAAAAAAA	GAAAAA-----
	NCCIT_3-26	(1821)	AAAA----AAAAAAAAAAAAA	-----
	NCCIT_3-27	(1819)	AAAA----AAAAAAAAAAAAA	-----

		1901		1950
	GH_3-18	(1566)	-----	CACTGTCATG-CCGTTACGTAGCGTATCG
	Mel 2_3-2	(1796)	AAAAAAAAAAAAA	CACTGTCATGCGCGTTACGTAGCGTATCG
	Mel 2_3-3	(1862)	AAAAAAAAAAAAA	CACTGTCATG-CCGTTACGTAGCGTATCG
	NCCIT_3-25	(1867)	-----	CACTGTCATG-CCGTTACGTAGCGTATCG
	Mel 2_3-5	(1778)	-----	CACTGTCATG-CCGTTACGTAGCGTATCG
	Mel 2_3-6	(1772)	-----	CACTGTCATG-CCGTTACGTAGCGTATCG
	Mel 2_3-4	(1837)	-----	CACTGTCATG-CCGTTACGTAGCGTATCG
	NCCIT_3-26	(1838)	-----	CACTGTCATG-CCGTTACGTAGCGTATCG
	NCCIT_3-27	(1836)	-----	CACTGTCATG-CCGTTACGTAGCGTATCG

		1951		2000
	GH_3-18	(1594)	TTGACAGCAATCGAAT	TCCCGCGCCGCCATGGCGCGG-AGCATGCGA
	Mel 2_3-2	(1846)	TTGACAGCAATCGAAT	CCC CGCGCCGCCATGGCT-CCGG-AGCATGCGA
	Mel 2_3-3	(1903)	TTGACAGCAATCGAAT	TCCCGCGCCGCCATGGCGCGG-AGCA-GCGA
	NCCIT_3-25	(1898)	TTGACAGCAATCGAAT	TCCCGCGCCGCCATGGCGCGG-AGCATGCGA
	Mel 2_3-5	(1809)	TTGACAGCAATCGAAT	TCCCGCGCCGCCATGGCGCGG-AGCATGCGA
	Mel 2_3-6	(1800)	TTGACAGCAATCGAAT	TCCCGCGCCGCCATGGCGCGG-AGCATGCGA
	Mel 2_3-4	(1868)	TTGACAGCAATCGAAT	TCCCGCGCCGCCATGGCGCGG-AGCATGCGA
	NCCIT_3-26	(1868)	TTGACAGCAATCGAAT	TCCCGCGCCGCCATGGCGCGG-AGCATGCGA
	NCCIT_3-27	(1867)	TTGACAGCAATCGAAT	TCCCGCGCCGCCATGGCGCGG-AGCATGCGA

		2001	2012
	GH_3-18	(1643)	-CTACGCCAGC
	Mel 2_3-2	(1893)	CTCACAG-----
	Mel 2_3-3	(1951)	-CTACGACA---
	NCCIT_3-25	(1947)	--CACACC----
	Mel 2_3-5	(1858)	-CTACGCCCGCC
	Mel 2_3-6	(1850)	-CTACGCCACT-
	Mel 2_3-4	(1917)	-CT-CGCCACT-
	NCCIT_3-26	(1917)	-CTCCGCCCTC
	NCCIT_3-27	(1917)	TCTACGCCAACT

8.3. Datenbankauswertung

Datenbankanalyse der im Luziferaseassay aktiven pck30

Detailed Family Information	Opt. ?	from	to	Strand	Core sim. ?	Matrix sim. ?	(red: ci-value ? > 60 CAPITALS: core sequence ?)
Myeloid zinc finger 1 factors	0.99	2	12	(+)	1.000	1.000	gtGGGGaaaag
SOX/SRY-sex/testis determinig and related HMG box factors	0.87	15	37	(-)	1.000	0.983	cacagtaaCAATctgtatctct
GATA binding factors	0.91	16	28	(+)	1.000	0.950	gagAGATcagatt
DM domain-containing transcription factors	0.80	22	42	(+)	1.000	0.965	tcagattGTTActgtgtgt
Vertebrate SMAD family of transcription factors	0.96	36	44	(+)	1.000	0.966	tGCTGTgt
Downstream Immunoglobulin Control Element, critical for B cell activity and specificity	0.80	38	52	(-)	0.892	0.809	ttctTCTAcacaga
MEF2, myocyte-specific enhancer binding factor	0.84	47	69	(-)	1.000	0.867	gagctcCTATgtctactcttt
Krueppel like transcription factors	0.93	60	76	(-)	1.000	0.934	caaaatGGAGtctccta
Activator/repressor binding to transcription initiation site	0.96	64	84	(+)	1.000	0.980	agactCCATttgtatgtac
PAR/bZIP family	0.84	72	88	(+)	1.000	0.859	ttttTTATgtactaag
Brn-5 POU domain factors	0.77	98	120	(+)	0.750	0.832	tctgcttgagATTCTgttaatc
Hepatic Nuclear Factor 1	0.84	106	122	(+)	1.000	0.853	gagattctGTTAatcta
Cellular and viral myb-like transcriptional regulators	0.88	107	119	(-)	0.817	0.904	attAACGagaact
Brn-5 POU domain factors	0.74	108	130	(-)	1.000	0.757	taaggtCATAgattaacagaaac
Brn POU domain factors	0.78	110	128	(-)	1.000	0.792	aggtcatagATTAcagaa
Nuclear receptor subfamily 2 factors	0.79	110	134	(-)	0.805	0.853	ggggtaaggTCATagattaacagaa
RXR heterodimer binding sites	0.69	112	136	(-)	1.000	0.770	tggggtaaaGGTCatagattaacag
Hepatic Nuclear Factor 1	0.80	113	129	(+)	1.000	0.877	tGTTAatctatgacct
Peroxisome proliferator-activated receptor	0.76	113	135	(-)	0.785	0.789	gggggtaaggtaoATAGTaaaca
CLOX and CLOX homology (CDP) factors	0.81	114	132	(+)	0.807	0.888	gttaATCTatgacctacc
v-ERB and RAR-related orphan receptor alpha	0.92	114	136	(-)	1.000	0.921	tgggggtaaaGGTCatagattaac
Estrogen response elements	0.87	118	136	(-)	1.000	0.931	tgggggtaAGGTcatagat
NGFI-B response elements, nur subfamily of nuclear receptors	0.86	119	133	(-)	1.000	0.934	gggtAAGGtcataga
RXR heterodimer binding sites	0.78	119	143	(-)	1.000	0.815	acggggtgggggtaAGGTcataga
Peroxisome proliferator-activated receptor	0.76	120	142	(-)	1.000	0.800	cgggggtgggggtaAGGTcatag
snRNA-activating protein complex	0.73	122	140	(+)	0.857	0.732	atgacCTTAcoccccaaccc
Zinc binding protein factors	0.91	124	146	(+)	1.000	0.948	gaccttaCCCcaaccccgctgt
GLI zinc finger family	0.88	127	141	(+)	1.000	0.883	cttaCCCcaacccc
E-box binding factors	0.84	136	148	(-)	1.000	0.864	agagCACGggggt
Twist subfamily of class B bHLH transcription factors	0.86	147	167	(-)	0.882	0.929	gacacAGCAcatgtttcagag
Two-handed zinc finger homeodomain transcription factors	0.98	148	160	(-)	1.000	0.986	cacatGTTTcaga
Iroquois homeobox transcription factors	0.86	150	162	(-)	1.000	0.913	agcaCATGtttca
Microphthalmia transcription factor	0.92	150	164	(+)	1.000	0.955	tgaacaCATGtctgt
Iroquois homeobox transcription factors	0.86	151	163	(+)	1.000	0.920	gaaCATGtctgt
Nuclear receptor subfamily 2 factors	0.83	151	175	(-)	0.778	0.861	tctgagtgaCACAGcaactgttct
E-box binding factors	0.92	152	164	(-)	0.860	0.936	acagcaCATGttt
NKX homeodomain factors	0.96	160	178	(-)	1.000	0.969	aactctgAGTGgacacagc
GLI zinc finger family	0.88	163	177	(+)	1.000	0.887	gtgtCCAActcagagt
SOX/SRY-sex/testis determinig and related HMG box factors	0.86	170	192	(+)	1.000	0.861	ctcagagtgAATGgattaaggg
CLOX and CLOX homology (CDP) factors	0.81	172	190	(-)	0.804	0.881	cttaATCCattcaactctg
Paralog hox genes 1-8 from the four hox clusters A, B, C, D	0.94	173	191	(+)	1.000	0.960	agagtgAATGgattaaggg
Cart-1 (cartilage homeoprotein 1)	0.87	173	193	(-)	1.000	0.902	gcctTAATccattcaactct
CCAAT binding factors	0.90	174	188	(-)	0.856	0.908	taatCCATcaactc
Selenocysteine tRNA activating factor	0.77	174	196	(-)	0.784	0.771	acgcCCCTtaatccattcaactc
RXR heterodimer binding sites	0.74	175	199	(+)	0.765	0.780	agttgaaGGATaaaggcgggtgca
Bicoid-like homeodomain transcription factors	0.96	177	193	(-)	1.000	0.970	gcctTAATccattca
Brn POU domain factors	0.89	177	195	(-)	1.000	0.891	cgcctTAATccattca
GC-Box factors SP1/GC	0.91	187	201	(+)	1.000	0.911	taaGGGCGgtgcagg
Glucocorticoid responsive and related elements	0.87	192	210	(+)	0.869	0.911	gcggtgcaggatGTGCttt
Human and murine ETS1 factors	0.93	192	212	(+)	1.000	0.979	gcggtgcaGGATgtgctttgt
SOX/SRY-sex/testis determinig and related HMG box factors	0.92	197	219	(-)	1.000	0.939	ctgtttaACAAgacacatcctgc
Nuclear receptor subfamily 2 factors	0.78	197	221	(-)	1.000	0.831	atctgtttaaCAAAGcaactcctgc
LEF1/TCF	0.86	202	218	(-)	1.000	0.907	tglttaaCAAAGcacat
Hepatic Nuclear Factor 1	0.84	203	219	(+)	1.000	0.865	tggtcttGTTAaacag
Fork head domain factors	0.84	207	223	(+)	1.000	0.882	ctttgtTAAAcagatgc
Hepatic Nuclear Factor 1	0.79	208	224	(-)	0.756	0.795	agcatctGTTTAcaaa
Twist subfamily of class B bHLH transcription factors	0.81	210	230	(+)	1.000	0.958	tglttaacaGATGcttgaagg
PAX-4/PAX-6 paired domain binding sites	0.75	212	230	(-)	0.774	0.803	ccctcAAGCcatctgttaa
NeuroD, Beta2, HLH domain	0.98	214	226	(-)	1.000	0.989	caagCATCTgtttt
PAX-2/5/8 binding sites	0.92	216	230	(-)	0.962	0.930	ccTCAageactctgt
Homeodomain transcription factors	0.83	232	250	(+)	1.000	0.909	agcatctcCTTAagagtc
Homeodomain transcription factors	0.83	237	255	(-)	1.000	0.943	gtgatgactCTTAaggagc
MAF and AP1 related factors	0.82	239	259	(+)	0.750	0.850	tccttaAGAGtcatcaccact
RXR heterodimer binding sites	0.75	248	272	(-)	0.750	0.760	cttgagattagggagTGGTgatgac
Krueppel like transcription factors	0.93	251	267	(-)	1.000	0.948	gattagGGAGgtggtgat
Bicoid-like homeodomain transcription factors	0.94	258	274	(+)	1.000	0.962	ctcccTAATcacaagta
Cart-1 (cartilage homeoprotein 1)	0.87	258	278	(+)	1.000	0.889	ctcccTAATcacaagta
X-box binding factors	0.90	269	287	(-)	0.750	0.905	ttgtTCCCctgggtactctg
Zinc finger transcription factor RU49, zinc finger proliferation 1 - Zipr1	0.98	270	276	(+)	1.000	1.000	aAGTAcc
X-box binding factors	0.90	270	288	(+)	0.875	0.922	aagtaCCCAgggacacaaa
Basonuclen rDNA transcription factor (PoI)	0.85	276	294	(-)	1.000	0.867	acaatttttGTCCctca

Human and murine ETS1 factors	0.96	287	307	(+)	1.000	0.962	aaaactgcGAAAgccgcagg
Neuron-specific-olfactory factor	0.88	293	315	(-)	1.000	0.919	cagaggTCCctgcggcctccgc
Signal transducer and activator of transcription	0.84	309	327	(-)	1.000	0.856	ggctTTCctaggcagaggt
POZ domain zinc finger expressed in B-Cells	0.76	310	326	(-)	1.000	0.778	gctTTCctaggcagagg
Signal transducer and activator of transcription	0.84	310	328	(+)	0.759	0.867	cctcTGCctaggaaagcca
SOX/SRY-sex/testis determinig and related HMG box factors	0.77	316	338	(+)	1.000	0.776	cctaggaaagccaggtatTGTcc
SOX/SRY-sex/testis determinig and related HMG box factors	0.87	321	343	(-)	1.000	0.875	accttggcaCAATaccctggcttc
Vertebrate steroidogenic factor	0.94	334	346	(+)	1.000	0.970	tgtcCAAGgttc
Twist subfamily of class B bHLH transcription factors	0.94	343	363	(-)	0.952	0.961	agactttCACAtggggagaaaa
snRNA-activating protein complex	0.75	344	362	(+)	0.839	0.751	tTCTCcccatgtgaaagtc
SWI/SNF related nucleophosphoproteins with a RING finger DNA binding motif	0.96	348	358	(+)	1.000	0.961	ccCCATgtgaa
E-box binding factors	0.90	348	360	(-)	0.915	0.905	ctttCACatgggg
Positive regulatory domain 1 binding factor	0.81	350	368	(+)	1.000	0.840	ccatgtGAAAgctctgaaat
Neuron-specific-olfactory factor	0.88	366	388	(-)	1.000	0.890	ttccctTCCcagcagggccatatt
E-box binding factors	0.92	370	382	(+)	1.000	0.938	tggcctCGTGgga
Human and murine ETS1 factors	0.86	372	392	(+)	1.000	0.863	gcctctgtGGAAGggaaagac
RBPJ - kappa	0.94	374	388	(+)	1.000	0.957	ctcgTGGGaaagggaa
Nuclear factor kappa B/c-rel	0.91	384	396	(-)	1.000	0.949	tcaggctTTCc
Zinc binding protein factors	0.91	394	416	(+)	1.000	0.919	tgaccgtCCCcagcccgacacc
Zinc binding protein factors	0.73	397	419	(+)	1.000	0.793	ccgtccCCCAGcccgacaccctg
Ras-responsive element binding protein	0.80	402	416	(+)	1.000	0.801	cCCAGcccgacacc
Abdominal-B type homeodomain transcription factors	0.81	411	427	(+)	1.000	0.824	gacaccggTAAAggggtc
Retinoblastoma-binding proteins with demethylase activity	0.96	424	432	(-)	1.000	0.972	GCACagacc
Bicoid-like homeodomain transcription factors	0.94	433	449	(-)	1.000	0.970	tttacTAATcctctca
Vertebrate caudal related homeodomain protein	0.94	438	456	(-)	1.000	0.978	ttccttTTTtAataatcct
Barbiturate-inducible element box from pro-eukaryotic genes	0.88	444	458	(+)	1.000	0.906	agtaAAAGaggaag
Human and murine ETS1 factors	0.88	445	465	(+)	1.000	0.973	gtaaaagaGGAAGgaatgcct
TEA/ATTS DNA binding domain factors	0.90	453	465	(-)	1.000	0.939	aggCATTccttc
Cellular and viral myb-like transcriptional regulators	0.90	465	477	(-)	1.000	0.904	ctCAACtgcaaga
EV11-myleoid transforming protein	0.72	470	486	(+)	0.765	0.735	cagttGAGAcagagga
EV11-myleoid transforming protein	0.83	475	491	(+)	1.000	0.838	gagacAAGAggaagga
Human and murine ETS1 factors	0.94	476	496	(+)	1.000	0.964	agacaagAGGAagcactgt
Carbohydrate response elements, consist of two E box motifs separated by 5 bp	0.82	479	495	(+)	0.800	0.824	CAAGaggaaggaactcg
Twist subfamily of class B bHLH transcription factors	0.85	482	502	(-)	1.000	0.875	caggagaCAGAtgcctctc
NeuroD, Beta2, HLH domain	0.98	486	498	(+)	1.000	1.000	aaggCATCtctct
CP2-erythrocyte Factor related to drosophila Elf1	0.84	498	516	(+)	0.833	0.866	tCCTGcctgtccctgggca
Activator/repressor binding to transcription initiation site	0.82	505	525	(-)	1.000	0.832	acattCCATtgcccagggaca
Octamer binding protein	0.89	510	526	(-)	0.900	0.894	gaCATTccattgccag
TEA/ATTS DNA binding domain factors	0.90	515	527	(-)	1.000	0.940	agaCATTccattg
PAX-2 binding sites	0.78	518	540	(+)	1.000	0.834	tggagtctctcggtataAACcc
Vertebrate caudal related homeodomain protein	0.84	525	543	(-)	1.000	0.851	atcgggtTTTAtaccgaga
Vertebrate TATA binding protein factor	0.90	528	544	(+)	1.000	0.981	cggtTAAAaccggatt
Grainyhead-like transcription factors	0.82	533	545	(-)	1.000	0.839	caatcgGTTTta
Activator/repressor binding to transcription initiation site	0.94	547	567	(+)	1.000	0.958	atgctCCATctactgagatag
GATA binding factors	0.96	559	571	(+)	1.000	0.992	ctgaGATAgggaa
Nuclear factor kappa B/c-rel	0.91	567	579	(-)	1.000	0.961	ggcggtttTTCc
CAS interating zinc finger protein	0.97	568	578	(+)	1.000	0.985	ggAAAAaccgc
Cellular and viral myb-like transcriptional regulators	0.96	570	582	(+)	0.898	0.961	aaAAAacgcctta
Activator protein 2	0.92	574	588	(+)	1.000	0.984	accGCTtagggctg
Zinc binding protein factors	0.73	580	602	(-)	1.000	0.737	gcaggtCCCacctccagccctaa
Neuron-restrictive silencer factor	0.67	602	632	(+)	1.000	0.689	cgggCAGCaatactgctttgaaagcactga
Zinc finger protein ZNF35	0.96	604	616	(+)	1.000	0.962	ggcaagAATactg
Vertebrate TATA binding protein factor	0.84	617	633	(+)	1.000	0.867	ctttgTAAAgcactgag
NKX homeodomain factors	0.80	618	636	(+)	0.855	0.802	tttgtaaaGCATgagatg
Mouse Krueppel like factor	0.74	619	639	(-)	0.750	0.750	aaacatcagtgCTTTacaa
Octamer binding protein	0.89	622	638	(-)	0.950	0.910	aaCATCctagcttcta
HOX - MEIS1 heterodimers	0.78	630	644	(+)	0.750	0.832	TGAGatgtttatgtg
Fork head domain factors	0.96	630	646	(-)	1.000	0.991	taacaTAAAcctctca
Octamer binding protein	0.88	638	654	(-)	1.000	0.945	gatatGCATacacataa
Octamer binding protein	0.88	643	659	(+)	1.000	0.970	tgtatGCATatctaaaa
GDNF-inducible zinc finger gene 1	0.73	652	664	(-)	0.750	0.766	TGTGcttttaat
NKX homeodomain factors	0.96	658	676	(-)	1.000	0.971	aggattaAGTgctgtctt
Retinoblastoma-binding proteins with demethylase activity	0.96	660	668	(+)	1.000	0.989	GCACagcac
Homeodomain transcription factors	0.81	660	678	(-)	0.793	0.886	aaaggatTAAGtctgtgc
Bicoid-like homeodomain transcription factors	0.96	665	681	(+)	1.000	0.977	gcactTAATcctttaca
Hepatic Nuclear Factor 1	0.80	667	683	(+)	0.767	0.812	aCTTAatcctttacatt
C2H2 zinc finger protein PLZF	0.86	668	682	(-)	0.958	0.885	atgTAAAggattaag
Octamer binding protein	0.80	669	685	(-)	0.846	0.815	acAATGtaaaggattaa
DM domain-containing transcription factors	0.75	669	689	(+)	1.000	0.753	ttaatcctTACAttgtctat
SOX/SRY-sex/testis determinig and related HMG box factors	0.87	670	692	(-)	1.000	0.872	atcatagaCAATgtaaaggatta
Octamer binding protein	0.88	683	699	(-)	1.000	0.932	ftcttGCATcatagaca
cAMP-responsive element binding proteins	0.71	683	703	(+)	0.750	0.710	tgcttaTGATgcaaagacctt
PAR/bZIP family	0.84	684	700	(+)	0.808	0.857	gtctaTGATgcaaagac
Octamer binding protein	0.85	688	704	(+)	1.000	0.951	atgATGCaagacctt

AARE binding factors	0.95	689	697	(-)	0.981	0.976	tTGCatca
SOX/SRY-sex/testis determining and related HMG box factors	0.92	691	713	(-)	1.000	0.982	catgtgaACAAaggtcttgcac
Nuclear receptor subfamily 2 factors	0.82	691	715	(-)	1.000	0.879	aacatgtgaaCAAaggtcttgcac
Peroxisome proliferator-activated receptor LEF1/TCF	0.76	694	716	(-)	1.000	0.762	aaacatgtgaaacaaAGGTctttg
Fork head domain factors	0.86	696	712	(-)	1.000	0.900	atgtgaaCAAaggtctt
Microphthalmia transcription factor	0.83	699	715	(-)	1.000	0.888	aacatgtgAACaaaggt
Iroquois homeobox transcription factors	0.92	704	718	(-)	1.000	0.941	acaaaCATGTgaaca
Iroquois homeobox transcription factors	0.86	705	717	(-)	1.000	0.925	caaaCATGTgaac
MAF and AP1 related factors	0.86	706	718	(+)	1.000	0.918	ttoaCATGttgt
Peroxisome proliferator-activated receptor	0.97	713	733	(+)	1.000	0.998	gtttgtctGCTGaccctctcc
Sterol regulatory element binding proteins	0.76	720	742	(-)	0.769	0.761	caattgtgggagaGGGTcaagca
Myeloid zinc finger 1 factors	0.80	726	740	(+)	0.750	0.838	cccTCTCcccaaat
EVII1-myleoid transforming protein	0.99	727	737	(-)	1.000	1.000	gtGGGgagagg
RXR heterodimer binding sites	0.83	735	751	(-)	0.750	0.875	gtcaacaAGACaattgtg
DM domain-containing transcription factors	0.69	737	761	(-)	1.000	0.743	atgtgtcaagGCTCacaagacaattg
MAF and AP1 related factors	0.80	740	760	(+)	0.809	0.802	ttgtcttTGAccctgacaca
cAMP-responsive element binding proteins	0.89	746	766	(-)	0.813	0.914	gggggaTGTGtcagggtcaaca
Calsenilin, presenilin binding protein, EF hand transcription factor	0.71	748	768	(+)	1.000	0.806	tgaccTGACacatccccctc
GC-Box factors SP1/GC	0.95	749	759	(-)	1.000	0.979	gtGTCAGgtgtc
Heat shock factors	0.83	752	766	(-)	0.750	0.879	ggGGGAtgtgtcagg
Human acute myelogenous leukemia factors	0.75	761	785	(+)	1.000	0.774	tccccctcttCGAaaacaccacaca
Twist subfamily of class B bHLH transcription factors	0.84	774	788	(-)	1.000	0.842	atctGTGGgtgttc
NeuroD, Beta2, HLH domain	0.87	777	797	(+)	1.000	0.932	acacccaCAGATgatacaataa
CLOX and CLOX homology (CDP) factors	0.98	781	793	(-)	1.000	0.987	tgatCATCTgtgtg
Onecut homeodomain factor HNF6	0.94	781	799	(-)	1.000	0.944	atttattGATCactgtgtg
Abdominal-B type homeodomain transcription factors	0.82	783	799	(+)	0.750	0.846	acagatGATCaataaat
CLOX and CLOX homology (CDP) factors	0.91	786	802	(+)	1.000	0.910	gatgatcaATAAact
Homeodomain transcription factors	0.81	786	804	(+)	0.874	0.911	gatgATCAataaataactaa
GHF-1 pituitary specific pou domain transcription factor	0.81	786	804	(-)	0.756	0.829	ttagtaTTTAttgatcacc
Selenocysteine tRNA activating factor	0.92	788	802	(-)	1.000	0.929	agtaTTTAttgatca
Activator-, mediator- and TBP-dependent core promoter element for RNA polymerase II transcription from TATA-less promoters	0.76	807	829	(-)	0.905	0.777	ggatCCCgccacgtctgtgagttc
Nuclear factor kappa B/c-rel	0.80	819	829	(+)	1.000	0.830	tgCGGGgatcc
Twist subfamily of class B bHLH transcription factors	0.89	821	833	(+)	1.000	0.904	gcGGGAtctctcca
Twist subfamily of class B bHLH transcription factors	0.91	824	844	(-)	1.000	0.948	cgttcagCATAtgaggatccc
Activator/repressor binding to transcription initiation site	0.91	825	845	(+)	1.000	0.915	gatctccCATAtgctgaacgc
NeuroD, Beta2, HLH domain	0.94	826	846	(+)	1.000	0.944	atctCCATatgtctgaacgct
SWI/SNF related nucleophosphoproteins with a RING finger DNA binding motif	0.92	828	840	(+)	1.000	0.931	cctCCATatgtctg
X-box binding factors	0.96	829	839	(+)	1.000	0.991	ctCCATatgtct
Krüppel like transcription factors	0.90	845	863	(-)	0.882	0.909	aggggaccggGGAAaccag
GA-boxes	0.86	856	872	(-)	1.000	0.973	aaagaagaAGGGgacc
CAS interacting zinc finger protein	0.78	857	881	(-)	1.000	0.782	agtatAGAGaaagaagaaggggac
EVII1-myleoid transforming protein	0.97	891	901	(-)	1.000	0.972	agAAAagaca
Ccoat/Enhancer Binding Protein	0.83	891	907	(-)	1.000	0.883	tggaaAGAAAAagaca
GC-Box factors SP1/GC	0.94	899	913	(-)	1.000	0.945	gagatttgGAAAaga
PAX-3 binding sites	0.80	913	927	(-)	1.000	0.819	gtaaggtgGGACgag
Heat shock factors	0.76	913	931	(+)	1.000	0.858	cTCGTcccaaccttaagaga
Signal transducer and activator of transcription	0.75	917	941	(+)	1.000	0.768	tcccaccttaCGAaaacaccacaca
Human acute myelogenous leukemia factors	0.87	920	938	(+)	0.808	0.896	caaccttaagAGAAacacccc
PAX-5 B-cell-specific activator protein	0.84	930	944	(-)	1.000	0.843	acctGTGGgtgttc
Olfactory associated zinc finger protein	0.79	930	958	(+)	0.857	0.815	gaaacaCCCACaggtgtgtaggggcaacc
Brachyury gene, mesoderm developmental factor	0.73	931	947	(+)	0.750	0.785	aaACACccacaggtgtg
Two-handed zinc finger homeodomain transcription factors	0.73	932	948	(-)	0.750	0.765	acACACctgtgtgtgtt
EGR/nerve growth factor induced protein C & related factors	0.66	932	952	(+)	1.000	0.660	aacacccacAGGTgtgttaggg
THAP domain containing protein	0.98	937	949	(-)	1.000	0.987	tacacACCTgtgtg
Krüppel-like C2H2 zinc finger factors hypermethylated in cancer	0.79	941	957	(+)	0.777	0.809	aggtGTGTAggggcaac
GC-Box factors SP1/GC	0.90	946	956	(+)	1.000	0.912	tgtaggGGCAa
Testis-specific bHLH-Zip transcription factors	0.93	948	960	(-)	1.000	0.961	tgggtTGCccta
GC-Box factors SP1/GC	0.83	952	966	(-)	1.000	0.879	taGGGGTgggtgtcc
Testis-specific bHLH-Zip transcription factors	0.94	953	963	(-)	0.896	0.943	gGGTggtgttc

Datenbankanalyse der im Luziferaseassay inaktiven LTR 21

Detailed Family Information	Opt. ?	from	to	Strand	Core sim. ?	Matrix sim. ?	(red: qi-value ? > 60 CAPITALS: core sequence ?)
Myeloid zinc finger 1 factors	0.99	2	12	(+)	1.000	1.000	gtGGGGaaaaag
v-ERB and RAR-related orphan receptor alpha	0.88	11	33	(+)	1.000	0.895	agcaagagagGTCAGattgttac
DM domain-containing transcription factors	0.77	14	34	(+)	1.000	0.771	aagagaggtcagaTTGTtact
SOX/SRY-sex/testis determinig and related HMG box factors	0.87	15	37	(-)	1.000	0.983	cacagtaaCAATctgacctctct
DM domain-containing transcription factors	0.80	22	42	(+)	1.000	0.965	tcagattGTTActgtgtctgt
GDNF-inducible zinc finger gene 1	0.73	34	46	(+)	0.750	0.761	TGTGctgtatag
Vertebrate SMAD family of transcription factors	0.96	36	44	(+)	1.000	0.970	tGTCtgtat
MEF2, myocyte-specific enhancer binding factor	0.84	47	69	(-)	1.000	0.867	gagctctCTATgtctctcttt
Kruppel like transcription factors	0.93	60	76	(-)	1.000	0.934	caaaaatGGAGtctccta
Glucocorticoid responsive and related elements	0.89	64	82	(+)	1.000	0.901	agactccattttGTTCTgt
Activator/repressor binding to transcription initiation site	0.96	64	84	(+)	1.000	0.980	agactCCATttttgtctgtac
Brn-5 POU domain factors	0.74	82	104	(-)	0.750	0.787	aggcagAATAattttcttagta
Homeodomain transcription factors	0.89	84	102	(+)	1.000	0.895	ctaagaaaAATTattctgc
Homeodomain transcription factors	0.88	85	103	(-)	1.000	0.882	ggcagaatAATTttctta
Homeobox transcription factors	0.85	88	106	(-)	1.000	0.886	caagccagaaTAATttttc
Fork head domain factors	0.83	90	106	(-)	0.818	0.882	caaggcAGAAtaatttt
Cellular and viral myb-like transcriptional regulators	0.88	107	119	(-)	0.817	0.885	attAACAgcatct
Brn-5 POU domain factors	0.74	108	130	(-)	1.000	0.757	taaggtCATAgattaacagcatc
PBX1 - MEIS1 complexes	0.76	109	125	(-)	0.750	0.767	tcatagatTAACagcat
Brn POU domain factors	0.78	110	128	(-)	1.000	0.792	aggtcatagATTAacagca
Nuclear receptor subfamily 2 factors	0.79	110	134	(-)	0.805	0.853	ggggtaaggtCATagattaacagca
RXR heterodimer binding sites	0.69	112	136	(-)	1.000	0.770	tgggggtaaGGTCatagattaacag
Hepatic Nuclear Factor 1	0.80	113	129	(+)	1.000	0.877	tGTTAatctatgacct
Peroxisome proliferator-activated receptor	0.76	113	135	(-)	0.785	0.789	gggggtaaggtcatAGATtaaca
CLOX and CLOX homology (CDP) factors	0.81	114	132	(+)	0.807	0.888	gttaATCTatgacctacc
v-ERB and RAR-related orphan receptor alpha	0.92	114	136	(-)	1.000	0.921	tgggggtaaGGTCatagattaac
Estrogen response elements	0.87	118	136	(-)	1.000	0.931	tgggggtaAGGTcatagat
NGFI-B response elements, nur subfamily of nuclear receptors	0.86	119	133	(-)	1.000	0.934	gggtAAGTcataga
RXR heterodimer binding sites	0.78	119	143	(-)	1.000	0.815	acggggttgggggtaAGGTcataga
Peroxisome proliferator-activated receptor	0.76	120	142	(-)	1.000	0.800	cggggtgggggtaAGGTcatag
snRNA-activating protein complex	0.73	122	140	(+)	0.857	0.732	atgacCTTAaccoccaacc
Zinc binding protein factors	0.91	124	146	(+)	1.000	0.948	gaccttaCCCCoaaccocgtct
GLI zinc finger family	0.88	127	141	(+)	1.000	0.883	cttaCCCCoaaccoc
E-box binding factors	0.84	136	148	(-)	1.000	0.864	agagCACGgggtt
Twist subfamily of class B bHLH transcription factors	0.86	147	167	(-)	0.882	0.929	gacacAGCcatgtttcagag
Two-handed zinc finger homeodomain transcription factors	0.98	148	160	(-)	1.000	0.986	cacatGTTTcaga
Iroquois homeobox transcription factors	0.86	150	162	(-)	1.000	0.913	agcaCATGtttca
Microphthalmia transcription factor	0.92	150	164	(+)	1.000	0.955	tgaacCATGtgcctg
Iroquois homeobox transcription factors	0.86	151	163	(+)	1.000	0.920	gaaacCATGtgcctg
E-box binding factors	0.92	152	164	(-)	0.860	0.936	acagcaCATGttt
MAF and AP1 related factors	0.81	154	174	(-)	1.000	0.914	tgagttGACacagcacatgt
MAF and AP1 related factors	0.82	156	176	(+)	1.000	0.890	atgtGCTGtgcacaactcagg
TALE homeodomain class recognizing TG motifs	1.00	157	173	(+)	1.000	1.000	tgtgtgtGTCAaactc
CLOX and CLOX homology (CDP) factors	0.81	173	191	(-)	0.804	0.891	cttaATCCattttaaccctg
NKX homeodomain factors	0.84	174	192	(+)	0.751	0.842	agggttAAATgggtaaagg
Cart-1 (cartilage homeoprotein 1)	0.87	174	194	(-)	1.000	0.902	gccctTAATccattttaaccct
Selenocysteine tRNA activating factor	0.77	175	197	(-)	0.784	0.774	accgCCCtTaatccattttaacc
SWI/SNF related nucleophosphoproteins with a RING finger DNA binding motif	0.96	177	187	(-)	1.000	0.967	atCCATtaac
Bicoid-like homeodomain transcription factors	0.96	178	194	(-)	1.000	0.970	gccctTAATccatttaa
Brn POU domain factors	0.89	178	196	(-)	1.000	0.891	ccgccctTAATccatttaa
GC-Box factors SP1/GC	0.91	188	202	(+)	1.000	0.911	taaAGGCgggtcagag
Glucocorticoid responsive and related elements	0.87	193	211	(+)	0.869	0.907	gcgggtcaagatGTGCttt
SOX/SRY-sex/testis determinig and related HMG box factors	0.77	193	215	(+)	1.000	0.862	gcgggtcaagatgtgctTTGTta
EVII-myleoid transforming protein	0.90	195	211	(+)	1.000	0.904	gggtcAAGAtgtgcttt
SOX/SRY-sex/testis determinig and related HMG box factors	0.92	198	220	(-)	1.000	0.939	ctgtttaACAAagcacactcttc
Nuclear receptor subfamily 2 factors	0.78	198	222	(-)	1.000	0.809	atctgttttaacAAAGcaactcttc
FAST-1 SMAD interacting proteins	0.81	199	215	(+)	1.000	0.815	caagaTGTGctttgta
LEF1/TCF	0.86	203	219	(-)	1.000	0.907	tgttttaacAAAGcaact
Hepatic Nuclear Factor 1	0.84	204	220	(+)	1.000	0.865	tgtgtctttGTTAaagag
Fork head domain factors	0.84	208	224	(+)	1.000	0.882	ctttgtTAAAcagatgc
Hepatic Nuclear Factor 1	0.79	209	225	(-)	0.756	0.795	agcatctgTTTaaacaaa
Twist subfamily of class B bHLH transcription factors	0.81	211	231	(+)	1.000	0.958	tgttaaacGATGcttgaagg
PAX-4/PAX-6 paired domain binding sites	0.75	213	231	(-)	0.774	0.803	ccctcAAGCactctgtttaa
NeuroD, Beta2, HLH domain	0.98	215	227	(-)	1.000	0.989	caagCATCtgttt
PAX-2/5/8 binding sites	0.92	217	231	(-)	0.962	0.930	ccTTCAagcatctgt
Brn-5 POU domain factors	0.83	231	253	(+)	1.000	0.962	gcagcatgctATTAaagatcat
Lim homeodomain factors	0.84	231	253	(+)	0.777	0.840	gcagcatgctCATTAaagatcat
Paralog hox genes 1-8 from the four hox clusters A, B, C, D	0.84	232	250	(-)	0.784	0.847	actcttAATGagcatgctg
Octamer binding protein	0.81	233	249	(+)	0.750	0.819	agcatgctCATTaagag
Brn POU domain factors	0.83	234	252	(-)	1.000	0.851	tgactctTAATGagcatgc
Homeobox transcription factors	0.85	234	252	(+)	1.000	0.874	gcagctcATTAaagatgca
Abdominal-B type homeodomain transcription factors	0.83	235	251	(+)	1.000	0.884	catgctcATTAaagatgc
NK6 homeobox transcription factors	0.85	236	250	(-)	1.000	0.865	actctTAATGagcat
Homeodomain transcription factors	0.92	238	256	(-)	1.000	0.921	gtgatgactCTTAATgagc
MAF and AP1 related factors	0.82	240	260	(+)	0.750	0.850	tcattaAGAGtoatccaccct
RXR heterodimer binding sites	0.75	249	273	(-)	0.750	0.760	cttgagattaggagTGGTgatgac

Bicoid-like homeodomain transcription factors	0.94	259	275	(+)	1.000	0.962	ctcccTAATctcaagta
Cart-1 (cartilage homeoprotein 1)	0.87	259	279	(+)	1.000	0.889	ctcccTAATctcaagtaacca
X-box binding factors	0.90	270	288	(-)	0.750	0.905	ttgtgTCCctgggtacttg
Zinc finger transcription factor RU49, zinc finger proliferation 1 - Zipro1	0.98	271	277	(+)	1.000	1.000	aAGTAcc
X-box binding factors	0.90	271	289	(+)	0.875	0.922	aagtaCCCAGgggacacaaa
Basonuclein rDNA transcription factor (PolI)	0.85	277	295	(-)	1.000	0.867	gcagttttgTGTcctgg
Human and murine ETS1 factors	0.96	288	308	(+)	1.000	0.965	aaaactgcGGAAGcctgcagg
Signal transducer and activator of transcription	0.84	310	328	(-)	1.000	0.856	ggctTTCctaggcagaggc
POZ domain zinc finger expressed in B-Cells	0.76	311	327	(-)	1.000	0.778	gctTTCctaggcagagg
Signal transducer and activator of transcription	0.84	311	329	(+)	0.759	0.867	cctcTGCCtaggaaagcca
SOX/SRY-sex/testis determining and related HMG box factors	0.77	317	339	(+)	1.000	0.776	cctaggaagccaggtatTGTcc
SOX/SRY-sex/testis determining and related HMG box factors	0.87	322	344	(-)	1.000	0.875	acctggaCAATacctgcttcc
Vertebrate steroidogenic factor	0.94	335	347	(+)	1.000	0.970	tgtcCAAGgttcc
Twist subfamily of class B bHLH transcription factors	0.94	344	364	(-)	0.952	0.970	agactatCACATggggagaaa
Microphthalmia transcription factor	0.92	347	361	(+)	1.000	0.940	ctcccCATGtgatag
E-box binding factors	0.92	349	361	(-)	0.860	0.922	ctatcaCATGggg
Nuclear receptor subfamily 2 factors	0.75	358	382	(-)	0.778	0.754	cccacgAGGCcatattcagactat
Neuron-specific-oligofactor	0.88	367	389	(-)	1.000	0.890	ttccctTCCcagcagggccatatt
E-box binding factors	0.92	371	383	(+)	1.000	0.938	tggtctCGTGgga
Human and murine ETS1 factors	0.86	373	393	(+)	1.000	0.863	gcctcgtgGGAAGggaaagac
RBPJ - kappa	0.94	375	389	(+)	1.000	0.957	ctcgTGGGaaagggaa
Nuclear factor kappa B/c-rel	0.91	385	397	(-)	1.000	0.949	tcaggtctTTCcc
Zinc binding protein factors	0.91	395	417	(+)	1.000	0.919	tgaccgtCCCcagcccgacacc
Zinc binding protein factors	0.73	398	420	(+)	1.000	0.793	ccgtccCCCAGcccgacaccogt
Ras-responsive element binding protein	0.80	403	417	(+)	1.000	0.801	cCCCAGcccgacacc
Abdominal-B type homeodomain transcription factors	0.81	412	428	(+)	1.000	0.824	gacaccogTAAAGggtc
Retinoblastoma-binding proteins with demethylase activity	0.96	425	433	(-)	1.000	0.972	GCACagacc
Bicoid-like homeodomain transcription factors	0.94	434	450	(-)	1.000	0.989	tatacTAATctctctca
Octamer binding protein	0.80	436	452	(-)	0.755	0.823	cttATACtaatctctct
Brn-5 POU domain factors	0.77	441	463	(-)	0.750	0.781	catgctttctCTTAtactaatc
Vertebrate TATA binding protein factor	0.82	444	460	(+)	1.000	0.828	tagTATAagaggaagac
p53 tumor suppressor	0.92	445	467	(+)	1.000	0.942	agtataagaggaagCATGtctc
Human and murine ETS1 factors	0.88	446	466	(+)	1.000	0.881	gtataagaGGAAGcagctgtc
Nuclear factor of activated T-cells	0.95	449	467	(+)	1.000	0.965	taagaGGAAGcagctgtc
Cellular and viral myb-like transcriptional regulators	0.90	466	478	(-)	1.000	0.904	ctCACTgcaaga
EVI1-myleoid transforming protein	0.72	471	487	(+)	0.765	0.735	cagttGAGAcagagga
EVI1-myleoid transforming protein	0.83	476	492	(+)	1.000	0.838	gagacAAGAGgaagagca
Human and murine ETS1 factors	0.94	477	497	(+)	1.000	0.964	agacaagAGGAagcctctgt
Carbohydrate response elements, consist of two E box motifs separated by 5 bp	0.82	480	496	(+)	0.800	0.824	CAAGaggaagcactgt
Twist subfamily of class B bHLH transcription factors	0.91	483	503	(-)	0.857	0.957	cgggaaaCAGAtgcctctc
NeuroD, Beta2, HLH domain	0.98	487	499	(+)	1.000	0.989	aaggCATCgttt
Two-handed zinc finger homeodomain transcription factors	0.98	491	503	(+)	1.000	0.988	catctGTTTcccg
E2F-myc activator/cell cycle regulator	0.81	493	509	(-)	1.000	0.966	gatgGGCGggaacacaga
GC-Box factors SP1/GC	0.91	495	509	(-)	1.000	0.930	gatGGCGggaacaaa
Activator-, mediator- and TBP-dependent core promoter element for RNA polymerase II transcription from TATA-less promoters	0.80	496	506	(-)	1.000	0.859	ggCGCGggaac
E2F-myc activator/cell cycle regulator	0.71	496	512	(+)	0.796	0.775	gtttcCCGCccatccct
Selenocysteine tRNA activating factor	0.76	496	518	(+)	0.905	0.813	gtttCCCgcctccctgggcaaa
Activator/repressor binding to transcription initiation site	0.82	506	526	(-)	1.000	0.836	acattCCATgcccaggggatg
Octamer binding protein	0.89	511	527	(-)	0.900	0.894	gaCATTccattgcccag
TEA/ATTS DNA binding domain factors	0.90	516	528	(-)	1.000	0.940	agaCATTccattg
PAX-2 binding sites	0.78	519	541	(+)	1.000	0.834	tggaatgtctcgtgataAAACcc
Vertebrate caudal related homeodomain protein	0.84	526	544	(-)	1.000	0.851	atcgggtTTTAtaccgaga
Vertebrate TATA binding protein factor	0.90	529	545	(+)	1.000	0.981	cggtTAAAaccogatt
Grainyhead-like transcription factors	0.82	534	546	(-)	1.000	0.839	caatcgGGTTta
GATA binding factors	0.96	560	572	(+)	1.000	0.992	ctgaGATAggggag
Human acute myelogenous leukemia factors	0.84	569	583	(-)	1.000	0.871	taagGTGtctcc
Zinc binding protein factors	0.73	581	603	(-)	1.000	0.733	gcctgtCCCACctccagccctaa
GC-Box factors SP1/GC	0.80	588	602	(+)	1.000	0.813	tggaaggtGGACatg
Zinc finger protein ZNF35	0.96	605	617	(+)	1.000	0.962	ggcagcAATActg
Vertebrate TATA binding protein factor	0.84	618	634	(+)	1.000	0.852	ctttgTAAAgcattgag
Octamer binding protein	0.89	623	639	(-)	0.950	0.921	aaCATCtcaatgttta
HOX - MEIS1 heterodimers	0.78	631	645	(+)	0.750	0.832	TGAGatgtttatgtg
Fork head domain factors	0.96	631	647	(-)	1.000	0.991	tacacaTAAAcacttca
Octamer binding protein	0.88	639	655	(-)	1.000	0.945	gatatGCATcacataa
Octamer binding protein	0.88	644	660	(+)	1.000	0.970	tgtatGCATatctaaaa
GDNF-inducible zinc finger gene 1	0.73	653	665	(-)	0.750	0.766	TGTGcttttagat
Homeodomain transcription factors	0.81	661	679	(-)	0.767	0.874	aaaggtTAAAtgctgtgc
Bicoid-like homeodomain transcription factors	0.96	666	682	(+)	1.000	0.994	gcattTAATCctttacc
Hepatic Nuclear Factor 1	0.80	668	684	(+)	0.784	0.812	aTTTAatcctttaccct
C2H2 zinc finger protein PLZF	0.86	669	683	(-)	0.958	0.895	aggTAAAggattaaa
OVO homolog-like transcription factors	0.80	671	685	(+)	0.750	0.807	taatccTTTAccttg
EVI1-myleoid transforming protein	0.83	673	689	(-)	0.750	0.876	tagacAAGGtaaaggat
Vertebrate steroidogenic factor	0.95	677	689	(-)	1.000	0.961	tagaCAAGGtaaa
Octamer binding protein	0.88	684	700	(-)	1.000	0.932	tctttGCATcatagaca
cAMP-responsive element binding proteins	0.71	684	704	(+)	0.750	0.710	tgtctaTGATgcaagaccct
PAR/bZIP family	0.84	685	701	(+)	0.808	0.857	gtctaTGATgcaagacc

PAR/bZIP family	0.84	685	701	(+)	0.808	0.857	gtctaTGATgcaaagac
Octamer binding protein	0.85	689	705	(+)	1.000	0.951	atgATGCaaagaccttt
AARE binding factors	0.95	690	698	(-)	0.981	0.976	tTGCatca
SOX/SRY-sex/testis determinig and related HMG box factors	0.92	692	714	(-)	1.000	0.982	cacgtgaCAAaggtcttgcac
Nuclear receptor subfamily 2 factors	0.82	692	716	(-)	1.000	0.879	aacacgtgaaCAAaggtcttgcac
Peroxisome proliferator-activated receptor	0.76	695	717	(-)	1.000	0.776	aaacacgtgaaacaaAGGTctttg
LEF1/TCF	0.86	697	713	(-)	1.000	0.916	acgtgaaCAAaggtctt
Fork head domain factors	0.83	700	716	(-)	1.000	0.885	aacacgtgAACaaggt
cAMP-responsive element binding proteins	0.88	702	722	(+)	1.000	0.938	ctttgttcACGTgtttgtctg
Hypoxia inducible factor, bHLH/PAS protein family	0.89	703	719	(+)	1.000	0.962	tttgttcaCGTgtttgt
Vertebrate homologues of enhancer of split complex	0.96	704	718	(-)	1.000	0.973	caaacaCGTgaacaa
Hypoxia inducible factor, bHLH/PAS protein family	0.89	704	720	(-)	1.000	0.944	gacaaacaCGTgaacaa
E-box binding factors	0.87	705	717	(+)	1.000	0.981	tgttCACGgtttt
Vertebrate homologues of enhancer of split complex	0.85	705	719	(+)	1.000	0.871	tgttCACGgtttgt
E-box binding factors	0.92	706	718	(-)	1.000	0.984	caaacaCGTgaac
Sterol regulatory element binding proteins	0.80	727	741	(+)	0.750	0.838	cccTCTCcccaactat
Myeloid zinc finger 1 factors	0.99	728	738	(-)	1.000	1.000	gtGGGGagagg
SOX/SRY-sex/testis determinig and related HMG box factors	0.90	729	751	(-)	1.000	0.908	tcacaagaCAATagtggtggagag
EV11-myleoid transforming protein	0.83	736	752	(-)	0.750	0.875	gtcacaAGACaatagtg
RXR heterodimer binding sites	0.69	738	762	(-)	1.000	0.734	atgtgtcagGGTCacaagacaatag
DM domain-containing transcription factors	0.80	741	761	(+)	0.809	0.802	ttgtcttGTGAcctgacaca
v-ERB and RAR-related orphan receptor alpha	0.79	746	768	(-)	1.000	0.809	agggagatgtGTCAgggtcaca
MAF and AP1 related factors	0.89	747	767	(-)	0.813	0.914	gggagaTGTGtcagggtcaca
cAMP-responsive element binding proteins	0.71	749	769	(+)	1.000	0.796	tgacctGACacatctcoctc
Calsenilin, presenilin binding protein, EF hand transcription factor	0.95	750	760	(-)	1.000	0.979	gtGTCAgggtc
Heat shock factors	0.95	773	797	(+)	0.899	0.953	gagaacacccaCGAAtgatcaata

8.4. Veröffentlichungen und Kongressbeiträge

N. Fuchs, J. Löwer, R. Löwer “**Deciphering a Promoter**“, Retreat Paul-Ehrlich-Institut, Elmstein – Harzofen (2006).

N. Fuchs, S. Hahn, H. Schmitz, H. Strobel, J. Löwer, R. Löwer, “**Transcriptional Regulation of HERV-K Expression in Germ Cell Tumor and Melanoma Cell Lines**“, Third European Congress of Virology, Nürnberg (2007).

N. Fuchs, S. Hahn, H. Schmitz, H. Strobel, J. Löwer, R. Löwer, “**Transcriptional Regulation of HERV-K Expression in Germ Cell Tumor and Melanoma Cell Lines**“, Retreat Paul-Ehrlich-Institut, Kröckelbach (2007).

N. Fuchs, S. Hahn, H. Strobel, J. Löwer, R. Löwer, “**Revised definition of the functional domains of Rec, the RNA export adaptor encoded by the human endogenous retrovirus K**“, Workshop on Post-Transcription Regulation of Viral Gene Expression, Virginia, USA (2008).

K. Boller, N. Fuchs, R. Eberle, J. Löwer, R. Löwer “**Ultrastructural analysis of HERV-K virus particles produced by human melanoma cells**“, Microscopy Conference 2009, Graz (2009).

LEBENS LAUF

Persönliche Daten:

Name: Nina Verena Fuchs
 Geburtsdatum/-ort: 12.12.1980/ Bielefeld
 Familienstand: ledig
 Staatsangehörigkeit: deutsch

Schulbildung:

08./1987 - 07./1991 Grundschole Leopoldshöhe
 08./1991 - 07./2000 Gymnasiums im Schulzentrum Aspe, Bad Salzuflen
 06./2000 Abitur

Studium:

10./2000 – 08./2005 Studium der Biologie an der Universität Bielefeld
 10./2000 – 10./2002 Grundstudium
 10./2002 – 07./2005 Hauptstudium
 07./2004 – 07./2005 Diplomarbeit am Lehrstuhl für Biochemie und Physiologie der Pflanzen unter der Leitung von Prof. Dr. Karl Josef Dietz

Thema: „Die Bedeutung mitochondrialer Redoxproteine im Redoxstoffwechsel und in der Entwicklung von *Arabidopsis thaliana*.“

07./2005 Abschluss des Studiums mit dem akademischen Grad:
 Diplom Biologin

Promotion:

seit 08./2005 Promotion am Paul-Ehrlich-Institut (Bundesinstitut für Impfstoffe und biomedizinische Arzneimittel) unter der Leitung von Herrn Prof. Dr. J. Löwer (in der Arbeitsgruppe von Frau Dr. R. Löwer) in Zusammenarbeit mit Herrn Prof. Dr. R. Marschalek (Fachbereich Biochemie, Chemie und Pharmazie der Johann Wolfgang Goethe-Universität in Frankfurt am Main).

Thema: „Charakterisierung der Expression des humanen endogenen Retrovirus K in der Melanomlinie UKRV Mel 2-C9 und Untersuchungen zur Regulation der Transkription in Melanom- und Keimzelltumormlinien.“

DANKSAGUNG

An dieser Stelle möchte ich mich bei allen herzlich bedanken, die zum Gelingen dieser Arbeit beigetragen haben. Vielen Dank für die Unterstützung während meiner Doktorandenzeit:

Prof. Dr. J. Löwer danke ich für die Vergabe des Themas sowie das stete Interesse am Fortgang der Arbeit, sowie für die sehr produktive Diskussionsbereitschaft und die vielen guten Anregungen.

Frau Dr. R. Löwer danke ich sehr herzlich für die hervorragende Betreuung und die großzügig gewährte Freiheit bei der Durchführung der Arbeit. Ich bin außerordentlich dankbar für die stete hilfreiche Diskussionsbereitschaft.

Bei Herrn Prof. Dr. R. Marschalek bedanke ich mich für die Betreuung seitens des Fachbereiches Biochemie, Chemie und Pharmazie der Johann Wolfgang Goethe-Universität in Frankfurt am Main.

Mein Dank gilt auch Herrn Prof. Dr. G. Schumann für viele hilfreiche Ideen und anregende Diskussionen.

Ebenso herzlich möchte ich mich bei Frau Dr. S. Hahn für die tatkräftige Unterstützung im Labor und die zahlreichen nützlichen Tipps bedanken.

Bei der Durchführung der Arbeiten am Fluoreszenz- und Elektronenmikroskop waren mir Dr. K. Boller und Frau R. Eberle eine wertvolle Hilfe.

Frau C. Tondera danke ich für die ausgezeichnete Zusammenarbeit und die produktive Unterstützung bei der Erzeugung experimenteller Daten.

Frau H. Schmitz danke ich ganz besonders für die fürsorgliche Pflege der Zellkultur.

Ich möchte mich auch bei meiner Mentorin PD Dr. S. Weßler bedanken, die mir im Rahmen der Doktorandenausbildung am Paul-Ehrlich-Institut stets mit konstruktiver Kritik zur Seite stand.

Ein großes Dankeschön geht an alle jetzigen und ehemaligen Mitarbeiter des Sachgebietes PR2 für viele anregende Diskussionen und fortwährende Hilfsbereitschaft. Danke an Heike, Jeannette, Axel, Julija, Sinja, Ulrike, Lili und Sabine für die sehr gute Arbeitsatmosphäre in den vergangenen Jahren.

Den Mitarbeiter der Arbeitsgruppen 2/2 und NG3 des Paul-Ehrlich-Instituts danke ich herzlich für stete Bereitschaft mit Kleinigkeiten und Ratschlägen aus zu helfen.

ERKLÄRUNG

Ich versichere, dass ich meine Dissertation

„Charakterisierung der Expression des humanen endogenen Retrovirus K in der Melanomlinie UKRV Mel 2-C9 und Untersuchungen zur Regulation der Transkription in Melanom- und Keimzelltumorklinien.“

unter der Leitung von Herrn Prof. Dr. J. Löwer am Paul-Ehrlich-Institut (Bundesinstitut für Impfstoffe und biomedizinische Arzneimittel) in Langen in Zusammenarbeit mit Herrn Prof. Dr. R. Marschalek (Fachbereich Biochemie, Chemie, Pharmazie der Johann Wolfgang Goethe-Universität in Frankfurt am Main) selbständig angefertigt habe, insbesondere, dass aus Schriften Entlehnungen, soweit sie in der Dissertation nicht ausdrücklich als solche mit Angabe der betreffenden Schrift bezeichnet sind, nicht stattgefunden haben.

Frankfurt am Main, 01.August 2009
