REGULATION OF HUMAN LINE-1 ACTIVITY AND INTEGRATION SPECIFICITY BY L1-EN MUTATIONS, DSB REPAIR PROTEINS AND THE APPLICATION OF THE TET/DOX-INDUCIBLE EXPRESSION SYSTEM

Dissertation
zur Erlangung des Doktorgrades
der Naturwissenschaften

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

von
Liliana Elisabeth Layer
aus Frankenthal (Pfalz)

Frankfurt 2016
(D 30)

Vom Fachbereich Biochemie, Chemie und Pharmazie der

Johann Wolfgang Goethe-Universität als Dissertation angenommen.

Dekan: Prof. Dr. Michael Karas, Johann Wolfgang Goethe-Universität, Frankfurt

Gutachter: Prof. Dr. Rolf Marschalek, Johann Wolfgang Goethe-Universität, Frankfurt
Prof. Dr. Gerald G. Schumann (Betreuer), Paul-Ehrlich-Institut, Langen

Datum der Disputation:

TABLE OF CONTENTS

TABLE OF FIGURES AND TABLES

SUMMARY

The human Long Interspersed Nuclear Element-1 (LINE-1, L1) is a member of the group of autonomous non-LTR retrotransposons found in almost every eukaryotic genome. L1 elements generate copies of themselves by reverse transcription of an RNA intermediate and integrate into the host genome by a process called Target Primed Reverse Transcription (TPRT). They are responsible for the generation of approximately 35% of the human genome, cover about 17% of the genome and represent the only group of active autonomous transposable elements in humans. L1 activity bears several risks for the integrity of the human genome, since the L1-encoded protein machinery generates DNA double-strand breaks (DSBs) and is capable of conducting numerous genome-destabilizing effects, e.g. causing deletions at insertion sites, disrupting or rearranging coding sequences and deregulating transcription of functional host genes. On the other side, L1 elements have had and still exert a great impact on human genome structure and evolution by increasing the genome size and rearranging and modulating gene expression. Furthermore, due to its endogenous and generally non-pathogenic nature, L1 is a promising candidate as vector for gene delivery in somatic gene therapy.

The structure of the flanking regions between *de novo* L1 integrants and the genomic sequence suggests an involvement of cellular DSB repair pathways in L1 mobilization. To elucidate the role of DSB repair proteins in L1 retrotransposition, I disabled DSB repair factors ATM, ATR, DNA-PK, p53 and Ku70 by knock down (KD) using short hairpin RNA (shRNA) expression constructs. To inhibit the function of DSB repair factors PARP and Rad51, I used dominant negative (DN) PARP and Rad51 mutants. Applying a well established L1-retrotransposition reporter assay in HeLa cells, *de novo* retrotransposition events were launched in order to test L1 for its retrotransposition activity in the context of altered DSB repair conditions. I could show that L1 retrotransposition frequency after ATM KD had increased by 3-fold, while ATR and p53 KD reduced L1 retrotransposition by approximately one third. Unfortunately, the cytotoxic effects of the DNA-PK and Ku70 shRNA expression constructs were too strong to determine potential effects of DNA-PK and Ku70 KD on L1 retrotransposition. Inhibition of PARP function by expression of the DN mutant and overexpression of wild type PARP were found to increase L1 retrotransposition by 1.8 and 1.5, respectively, while Rad51 DN had no detectable effect. Interestingly, overexpression of wild type Rad51 seemed to roughly double L1 retrotransposition frequencies.

Since in my experiments KD or expression of DN mutants was time-delayed to the onset of L1 retrotransposition after transfection into the cells, I developed a temporally controllable, tetracyclin transactivator (tTA)-dependent L1 retrotransposition reporter assay which will be of great value for future L1 retrotransposition studies that rely on temporally controllable retrotransposition. Due to a previously published hypothesis of L1 playing a role in brain development by contributing to somatic mosaicism in neuronal precursor cells, I generated a transgenic mouse (LORFUS) using the tTA-dependent L1 construct to further test this hypothesis. LORFUS harbors a bidirectional cassette driving simultaneous expression of a GFP-tagged L1 retrotransposition reporter and β-galactosidase. It was bred to another transgenic mouse line expressing tTA in the forebrain. The double transgenic offspring was used to characterize L1 expression and retrotransposition patterns in the brain at postnatal day 15 (P15). General transgene expression indicated by β-galactosidase activity was found in hippocampus, cortex and striatum, while retrotransposition events revealed by GFP expression were found in hippocampus, cortex, striatum, olfactory bulb and brainstem. These results suggested L1 retrotransposition in the granule layer of the dentate gyrus earlier than P15 and migration of cells carrying these events along the rostral migratory stream into the olfactory bulb.

To facilitate the use of L1 as gene delivery tool in gene therapy or genetic engineering, I furthermore intended to manipulate the L1 target site recognition to allow the site-specific integration into defined genomic locations. To this end, I performed crystal structure-guided mutational analysis exchanging single amino acid residues within the endonuclease (EN) domain of L1 to identify residues influencing target site recognition. However, individual point mutations did not change the nicking pattern of L1-EN, but resulted in a reduction of endonucleolytic activity reflected by a reduced retrotransposition frequency. This suggests that additional factors other than the DNA nicking specificity of L1-EN contribute to the targeted integration of non-LTR retrotransposons in the host genomes.

ZUSAMMENFASSUNG

Das menschliche Long Interspersed Nuclear Element-1 (LINE-1, L1) ist ein Mitglied der Gruppe der autonomen Non-LTR-Retrotransposons, die in fast allen eukaryotischen Genomen zu finden sind. L1-Elemente besitzen einen eigenen Promotor in ihrem 5' untranslatierten Bereich (untranslated region, UTR) und zwei offene Leserahmen (open reading frames, ORFs), von denen der erste für ein RNA-bindendes Protein kodiert und der zweite für ein multifunktionales Protein mit Endonuklease (EN)- und reverser Transkriptase (RT)-Domäne. L1-Retrotransposons generieren Kopien von sich selbst durch reverse Transkription eines RNA-Intermediats und Integration in das Wirtsgenom mittels eines Prozesses, der Target Primed Reverse Transcription (TPRT) genannt wird. Dabei spielt die EN-Domäne des ORF2 Proteins eine wichtige Rolle bei der Zielsequenzerkennung. Die bevorzugte Zielsequenz lautet AA TTTT auf dem unteren DNA-Strang (in 3' nach 5' Orientierung), der von der L1-EN zwischen dem zweiten Adenin und dem ersten Thymin gespalten wird. Die Reihe der Thymine dient dann als komplementäre Bindungssequenz für den Poly(A)-Schwanz der L1-RNA und zugleich als Primer für die reverse Transkription. Auf diese Art ist die L1-Proteinmaschinerie nicht nur in der Lage L1-Elemente zu vervielfältigen, sondern ist auch für die Mobilisierung von nicht-autonomen Retrotransposons wie *Alu*- und SVA-Sequenzen verantwortlich und für die Entstehung von Pseudogenen. Von L1 generierte Insertionen werden typischerweise von Zielsequenzverdopplungen (Target Site Duplications, TSDs) variabler Länge flankiert, die das Ergebnis eines versetzten DNA-Zweitstrangbruches während der Integration sind. L1-Elemente sind für die Entstehung von etwa 35% des humanen Genoms verantwortlich, bedecken selbst etwa 17% und stellen die einzige Gruppe aktiver, autonomer transponierbarer Elemente im Menschen dar. Die Aktivität von L1 birgt verschiedene Risiken für die Integrität des menschlichen Genoms, da die von L1 kodierte Proteinmaschinerie DNA-Doppelstrangbrüche (DSBs) verursacht und in der Lage ist, zahlreiche Genom-destabilisierende Effekte auszuüben, z.B. durch Deletionen an Integrationsstellen, Unterbrechung oder Reorganisation von kodierenden Sequenzen und durch Deregulation der Transkription funktioneller Wirtsgene. Andererseits hatten und haben L1-Elemente immer noch einen großen Einfluß auf die Struktur des menschlichen Genoms und auf unsere Evolution. Als repetitive Sequenzen bieten sie die Möglichkeit der DNA-Strang-Anlagerung für die homologe Rekombination und tragen so zur Reorganisation und Modulation der Genexpression bei. Zudem steigern sie stetig die Genomgröße. L1 stellt außerdem aufgrund seines endogenen Ursprungs und seiner grundsätzlich nicht-pathogenen Natur einen vielversprechenden Kandidaten als Vektor für den Gentransfer in der somatischen

Gentherapie dar. Besonders interessant ist dabei, daß in Retrotranspositionsversuchen genutzte L1-Elemente in der Lage sind, in ihren 3' UTRs Reportergene bei der Mobilisierung mitzunehmen, die potentiell gegen therapeutische Gene ersetzt werden könnten.

Diese Dissertation befaßt sich mit der Beantwortung von drei Fragen der L1-Forschung: 1. Spielen zelluläre Faktoren, insbesondere Proteine von DSB-Reparaturprozessen, eine Rolle bei der L1-Retrotransposition und wenn ja, welche? 2. Zur besseren Versuchsplanung ist die zeitliche und in Geweben eines Organismus örtliche Kontrolle eines L1-Retrotranspositionsreporteransatzes mittels Tetrazyklin-Transaktivator (tTA)-abhängigen Systems möglich? 3. Ist die Zielsequenz der L1-EN durch den Austausch einzelner Aminosäuren veränderbar, um L1-Retrotransposons dem Einsatz als Vektor in der Gentherapie näher zu bringen?

Zwischen dem 5' Ende einer *de novo* L1-Insertion und deren Übergang in das flankierende Genom bzw. in die TSD finden sich meist kurze Sequenzen, die sowohl vom L1-Element als auch von der genomischen Sequenz stammen könnten und nicht eindeutig einem Ursprung zuzuordnen sind. Diese werden Mikrohomologien genannt. Da man solche Mikrohomologien auch bei der Reparatur von DNA-Doppelstrangbrüchen (DSBs) durch z.B. einen Prozess namens Non-homologous End-joining (NHEJ) beobachtet, weisen diese auf eine Beteiligung von zellulären DSB-Reparaturmechanismen bei L1-Retrotranspositionsvorgängen hin. Um die Rolle von DSB-Reparaturproteinen bei der L1-Retrotransposition aufzuklären, verwendete ich sogenannte short hairpin RNA (shRNA)-Expressionskonstrukte, um verschiedene Reparaturfaktoren mittels Knock Down (KD)-Methode auszuschalten. Untersucht wurden die Faktoren ATM, ATR, DNA-PK, p53 und Ku70. Zur funktionellen Inhibition von DSB-Reparaturfaktoren benutzte ich dominant negative (DN) Proteinvarianten. Auf diese Weise inhibierte DSB-Reparaturproteine waren PARP und Rad51. Außerdem untersuchte ich auch den Effekt einer Überexpression von PARP und Rad51 mit Hilfe von Expressionsplasmiden, die für Wildtyp-PARP und -Rad51 kodierten. Durch Anwendung eines etablierten L1-Retrotranspositionsreporteransatzes in HeLa-Zellen wurden *de novo* Retrotranspositions-ereignisse zur Untersuchung der L1-Retrotransposition unter veränderten DSB-Reparaturbedingungen ausgelöst. Nur nach erfolgreicher L1-Retrotransposition von diesem L1-Retrotranspositionsreporterplasmid wird eine Neomycin-Resistenzkassette aktiviert, die den Zellen eine Neomycin- bzw. G418-Resistenz verleiht. Anschließend wurden die Zellen, die *de novo* Retrotranspositionsereignisse aufwiesen, mittels G418 selektiert und die Retrotranspositionsrate durch Anfärbung und Auszählung der G418-resistenten Kolonien bestimmt. So konnte ich zeigen, daß sich die L1-Retrotranspositionshäufigkeit nach ATM KD

etwa verdreifacht, während ATR und p53 KD eine Verringerung der L1-Retrotranspositionsrate von etwa einem Drittel zur Folge hatten. Leider waren die zytotoxischen Effekte der DNA-PK und Ku70 shRNA-Expressionskonstrukte so stark, daß Aussagen zu potentiellen Auswirkungen auf die L1-Retrotransposition nicht gemacht werden konnten. Die Inhibition von PARP durch die Expression der DN-Mutante und die Überexpression von Wildtyp-PARP erhöhten beide die Retrotranspositionshäufigkeit um etwa den Faktor 1,8 bzw. 1,5, während die Rad51 DN-Mutante zu keinem nachweisbaren Effekt führte. Interessanterweise führte die Überexpression von Wildtyp-Rad51 zu einer Verdopplung der L1-Retrotranspositionsrate. Da Rad51 den DNA-Strangaustausch während der homologen Rekombination unterstützt, könnte dies ein Hinweis darauf sein, daß auch während der L1-Retrotransposition ein DNA-Strangaustausch, z.B. während der Verknüpfung des L1 5' Endes mit der genomischen DNA, stattfindet.

In meinen Experimenten fanden der KD und die Expression von DN-Mutanten nach der Transfektion in Zellen zeitverzögert zur Initiation der L1-Retrotransposition statt. Deshalb entwickelte ich einen zeitlich kontrollierbaren, Tetrazyklin-Transaktivator (tTA)-abhängigen L1-Retrotranspositionsreporteransatz, der von großem Nutzen für künftige L1-Retrotranspositionsstudien sein könnte, die auf induzierbare Retrotranspositionsvorgänge angewiesen sind. Dazu klonierte ich den tTA-abhängigen Promotor eines kommerziell erhältlichen Plasmids anstelle des normalen CMV-Promotors eines L1-Retrotranspositionsreporterplasmids unter Weglassung des L1 5' UTRs mit dem endogenen Promotor. Der hierzu verwendete L1-Retrotranspositionsreporter war diesmal keine Neomycinresistenz sondern eine Blastizidinresistenz. Dann transfizierte ich dieses Plasmid mit dem tTA-abhängigen L1-Retrotranspositionsreporter in HeLa M2-Zellen, die den reversen tTA exprimieren und somit eine Expression des L1-Reporters nur unter Zugabe von Tetrazyklin bzw. Doxyzyklin erlauben. Anschließend wurden die transfizierten Zellen mit Blastizidin auf Retrotranspositionsereignisse selektioniert. Nur Zellen, die mit dem tTA-abhängigen L1-Retrotranspositionsreporter transfiziert wurden und Doxyzyklin erhielten bildeten Blastizidin-resistente Kolonien. Somit war die Funktionalität des tTA-abhängigen L1-Retrotranspositionsreporteransatzes bewiesen. Aufgrund einer zuvor publizierten Hypothese, wonach L1 eine Rolle in der Entwicklung des Gehirns spiele und zum somatischen Mosaizismus von neuronalen Vorläuferzellen beitrage, generierte ich mit dem tTA-abhängigen L1-Konstrukt transgene Mäuse, um diese Hypothese weiter zu untersuchen. Eine der transgenen Mäuse, LORFUS, trägt eine tTA-abhängige, bidirektionale Kassette zur simultanen Expression eines GFP-markierten L1-Retrotranspositionsreporters und einer β-

Galaktosidase. Die zweite transgene Maus, RORFUS, trägt eine identische tTA-abhängige, bidirektionale Kassette mit einem rotfluoreszierenden Protein, DsRed2, anstelle der β-Galaktosidase. Es konnten drei LORFUS- und sechs RORFUS-Founder durch Pronukleusinjektion generiert werden. Sowohl LORFUS- als auch RORFUS-Nachkommen wurden mit einer weiteren transgenen Maus gekreuzt, Kt1, die tTA unter der Kontrolle des CamKII-Promotors im Vorderhirn exprimiert. Die doppelt-transgenen Nachkommen wurden zur Charakterisierung der L1-Expressions- und Retrotranspositionsmuster im Gehirn herangezogen. Nur eine Founder-Linie von LORFUS zeigte am postnatalen Tag 15 (P15) Expression der transgenen Kassette - gekennzeichnet durch β-Galaktosidaseexpression - im Hippocampus, Cortex und Striatum der Tiere, während Retrotranspositionsereignisse - markiert durch GFP-Expression - im Hippocampus, Cortex, Striatum, Bulbus olfactorius und im Hirnstamm nachweisbar waren. Diese Ergebnisse weisen auf L1-Retrotransposition in der Körnerzellen-Schicht des Gyrus dentatus vor P15 hin und anschließende Migration der Zellen mit diesen Retrotranspositionsereignissen entlang des rostral migratorischen Stroms in den Bulbus olfactorius. RORFUS-Tiere lieferten leider keine brauchbare transgene Linie.

Um den Einsatz von L1 als Vektor in der Gentherapie oder in der Gentechnik voranzutreiben, war ein weiteres Vorhaben, die L1-Zielsequenzerkennung zur sequenzspezifischen Integration in definierte genomische Loci zu verändern. Dazu führte ich Kristallstruktur-geleitete Mutationsanalysen durch und tauschte einzelne Aminosäuren (AS) innerhalb der L1-EN-Domäne aus, um Aminosäurereste zu identifizieren, die die Zielsequenzerkennung beeinflussen. Es wurden vier AS zum Austausch ausgewählt: 1. Serin 202. Die hypothetische Funktion dieser AS liegt darin, eine Wasserstoffbrücke zum zweiten Adenin der L1-Konsensuszielsequenz zu bilden. Austausch gegen Alanin (S202A) sollte diese Wasserstoffbrückenbindung verhindern und eine Veränderung der Zielsequenz und möglicherweise eine Veränderung der Retrotranspositionsrate zur Folge haben. 2. Arginin 155. Die hypothetische Funktion dieser AS ist identisch mit der von Serin 202. Austausch gegen Alanin (R155A) sollte daher dieselbe Folge haben wie die erste EN-Punktmutante. 3. Threonin 192. Die hypothetische Funktion dieser AS ist die Spaltung der Zielsequenz und/oder die Verankerung des α11-Loops der L1-EN mittels Wasserstoffbrückenbindung innerhalb der EN. Austausch gegen Valin (T192V) sollte daher eine nicht-funktionale EN zur Folge haben. 4. Isoleucin 204. Die hypothetische Funktion dieser AS ist die Bereitstellung einer Tasche für das zweite Adenin der Zielsequenz, das sich vermutlich bei der Bindung durch das ORF2 Protein aus der DNA herausdreht und in der EN akkommodiert wird. Austausch gegen eine große AS wie Tyrosin (I204Y) sollte die Akkommodation des Adenins

in der EN verhindern und somit eine veränderte Zielsequenz und eine Veränderung der Retrotranspositionsrate zur Folge haben. Als Negativkontrolle wurde eine zuvor publizierte EN-Punktmutante verwendet (H230A), die zu einer nicht-funktionalen L1-EN führt. Alle L1-EN-Punktmutanten wurden in ein herkömmliches L1-Retrotranspositionsreporterplasmid kloniert und in HeLa-Zellen transfiziert. Danach wurden die Zellen mit G418 auf Retrotranspositionsereignisse selektioniert. Die einzelnen Punktmutationen bewirkten wider Erwarten nicht die Veränderung der Zielsequenz bzw. des Restriktionsmusters der L1-EN, sondern vielmehr nur die Erniedrigung ihrer endonukleolytischen Aktivität, die sich in verringerten Retrotranspositionsraten widerspiegelte. S202A wies nur noch etwa ein Drittel der Wildtyp-L1-EN-Retrotranspositionsaktivität auf, R155A nur etwa ein Achtel, T192V nur noch ein Zwanzigstel und I204Y hatte mit einem Hundertstel die geringste Retrotranspositionsrate von allen. Die Negativkontrolle H230A lieferte erwartungsgemäß keine G418-resistenten Kolonien. Dieses Ergebnis weist darauf hin, daß zusätzlich zur L1-EN-Zielsequenzerkennung andere Faktoren wie eventuell die dreidimensionale Struktur der genomischen Integrationsstelle für die sequenzspezifische Integration von Non-LTR-Retrotransposons im Wirtsgenom eine Rolle spielen.

Zusammenfassend lassen sich die drei in dieser Dissertation bearbeiteten Fragen wie folgt beantworten: 1. Es konnten zelluläre Faktoren von DSB-Reparaturprozessen identifiziert werden, die inhibitorische oder unterstützende Wirkung auf *de novo* L1 Integrationen haben. Deren genaue Rolle bei der L1 Retrotransposition gilt es noch aufzuklären. 2. Die L1 Retrotransposition kann durch das tTA System sowohl zeitlich als auch in Geweben eines Organismus örtlich kontrolliert werden, was die Bearbeitung vieler, weiterer Fragestellungen in der L1-Forschung vereinfacht. 3. Die Zielsequenz der L1-EN ist durch den Austausch einzelner Aminosäuren nicht derart veränderbar, um L1 ausreichend zielgerichtet in bestimmte Sequenzen integrieren zu lassen und so den Einsatz als Vektor in der Gentherapie zu ermöglichen.

1 INTRODUCTION

The human genome consists of $3x10^9$ base pairs (bp) and is organized into 23 chromosome pairs of different size ranging from ~$47x10^6$ bp (chromosome 21) up to ~$245x10^6$ bp (chromosome 1). This set of desoxyribonucleic acid (DNA) dictates every individual's properties and is passed on to following generations inheriting the traits as phenotypic appearance, physical health and mental sanity. During the process of human development and reproduction the DNA has to be transmitted and reorganized many times, always yielding a proper set of genetic information for the offspring, which can be single cells within the organism or gametes giving rise to new individuals eventually.

However, life bears many threats for the integrity of DNA molecules. Genomic mutagenesis can be caused by exogenous factors like UV irradiation, ionizing radiation and carcinogenic agents. These factors can cause chemical alterations in the DNA which in turn may lead to sequence modifications, defective or incomplete replication and single- or double-strand breaks. Endogenous threats are given by the cellular replication and cell division machinery, since none of the enzymes and factors involved is working absolutely error-free and can cause mismatches, deletions, inversions and other mistakes during DNA processing. Also reactive oxygen species generated during cell metabolism pose a risk for chemical aberration of the DNA. And even parts of the DNA itself endanger genome integrity. These include several elements formerly known as "junk DNA" located within intronic and intergenic regions and later called "selfish DNA". This kind of DNA was first described by Barbara McClintock in maize in the late 1940s, when she found that the color of corn is not passed on in a Mendelian way but rather randomly. She characterized a species of mobile DNA, which was later named DNA transposons, that inserts itself into genes coding for the corn color and thus leading to a wide range of color variations in maize fruit. Later, these elements were also described in bacteria (Hedges and Jacob 1974) and animal cells such as mouse and other higher vertebrates (Kuff et al. 1981). Until 2007, DNA transposons that move via a cut-and-paste mechanism were believed to be non-functional in vertebrates due to a series of mutations during early animal evolution rendering them inactive. Only recently, a variety of ancient transposons replicating and mobilizing themselves within the genome of the little brown bat *Myotis lucifugus* were found (Ray et al. 2007; Ray et al. 2008). Nevertheless, in the genomes of most vertebrates and mammals, including humans, DNA transposon sequences represent mere relicts in their genomes. It is another kind of mobile elements, the retrotransposable elements, which threatens the integrity of the host DNA. This group of mobile elements can be devided

into two major subgroups, the LTR-retrotransposons and the non-LTR-retrotransposons. LTR stands for long terminal repeat and denominates the flanking regions of proviruses and transposable elements of several hundreds of base pairs in length. These LTR-containing elements are thought to originate from retroviral infections of the germline and subsequently have lost a functional *env*-gene during evolution. Members of both groups of transposable elements move via a copy-and-paste mechanism either inserting stochastically or site/sequence-specifically into the host's genome and have accumulated during evolution. LTR-retrotransposons in the human genome are also termed endogenous retroviruses and comprise ~8% of the genome, while non-LTR-retrotransposons cover ~35%. The most abundant member of non-LTR-retrotransposons is the clade of long interspersed nuclear elements type 1 (LINE-1, L1). They make up as much as 17% of the human genome and are the only kind of active, autonomously replicating transposable elements within humans. Their activity bears several risks for the integrity of the human genome, since L1 elements have been shown to produce DNA double-strand breaks (DSBs) (Gasior et al. 2006) and are capable of conducting numerous deleterious effects, as for example disrupting or rearranging coding sequences and deregulating transcription of functional genes (Cordaux and Batzer 2009; Beck et al. 2011). The possible consequences of L1 activity will be discussed in more detail in chapter 1.8.

Considering all of these exogenous and endogenous genome destabilizing and shaping agents, it is conceivable that cells had to come up with a true army of damage sensing mechanisms and DNA repair pathways to escape replication fork stalling, cell death and - in the worst case scenario - inactivation of tumour supressor genes or activation of oncogenes. An ancient system of intracellular defense protecting the genome against viral infections and transposable elements will be introduced below, as well as the most common DNA repair pathways and candidate pathways potentially involved in retrotransposition of L1 elements.

1.1 RNA INTERFERENCE (RNAI)

RNAi represents an ancient and evolutionarily conserved system to protect the genome against viral infections and genomic instability caused by mobile genetic elements such as transposable and repetitive elements (Aravin et al. 2007). RNAi describes a highly conserved regulatory mechanism initiated by double-stranded RNA (dsRNA) that mediates the degradation of homologous mRNA in eukaryotic cells. "RNA-silencing" by dsRNA is a natural, widely spread cellular mechanism that was first described in the nematode *C. elegans*

(Fire et al. 1998). RNAi has subsequently been demonstrated in many other organisms including insects (Pal-Bhadra et al. 1997), planarian (Sanchez Alvarado and Newmark 1999), *Trypanosoma brucei* (Ngo et al. 1998), plants (Waterhouse et al. 1998), fungi (Catalanotto et al. 2000) and mammalian cells (Caplen et al. 2001; Elbashir et al. 2001), including mouse oocytes, embryos, and embryonic stem cells (Wianny and Zernicka-Goetz 2000).

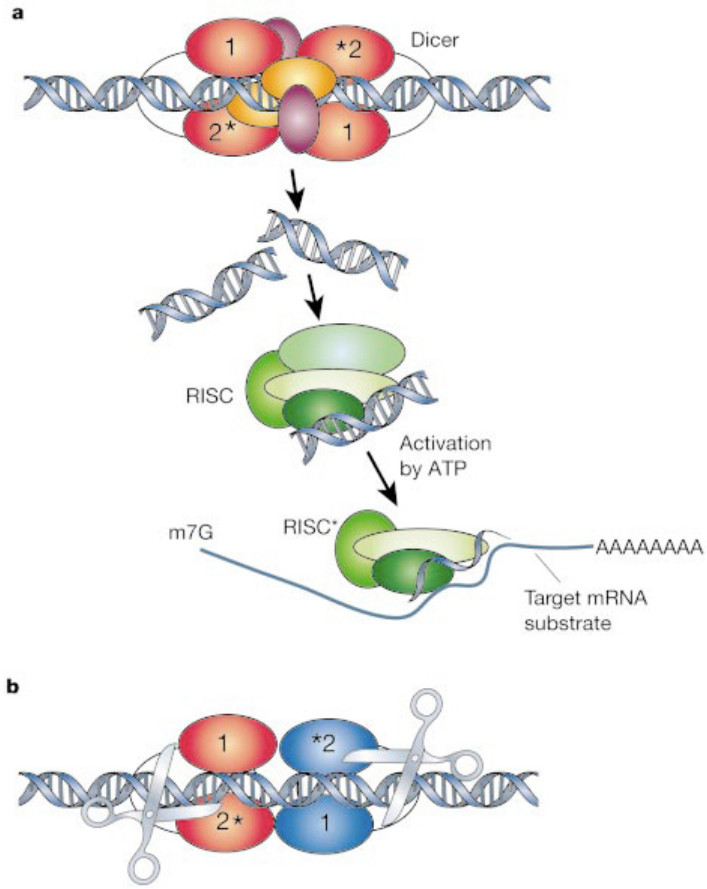


Figure 1 Model for siRNA-mediated target mRNA-degradation. a RNAi is initiated by Dicer dimers (two Dicer molecules with five domains each are shown), which processes double-stranded RNA into ~22-nucleotide small interfering RNAs. The siRNAs are incorporated into a multicomponent nuclease, RISC (green). RISC is activated from a latent form to an active form, RISC*, and unwinding of siRNAs takes place. RISC* then uses the unwound siRNA as a guide to substrate selection. **b** Diagrammatic representation of Dicer binding and cleaving dsRNA (for clarity, not all the Dicer domains are shown, and the two separate Dicer molecules are coloured differently) (Hannon 2002).

In mammalian cells, dsRNA can trigger at least two different cellular responses, 1. a general inhibition of gene expression and 2. a sequence-specific degradation of target mRNA (Hannon 2002). The first pathway involves activation of the dsRNA-dependent kinase PKR and 2'-5'-oligoadenylate synthetase normally induced by interferon. This cellular pathway is

considered a non-specific viral defense mechanism that leads to an overall shut-down of host cell protein synthesis.

In contrast, RNAi triggers sequence-specific mRNA degradation without affecting general mRNA synthesis and translation. Hereby, the length of the trigger dsRNA is critical; only dsRNA shorter than 30 bp allows differential activation of RNAi as opposed to the nonspecific mRNA degradation pathway. Such small interfering RNAs (siRNAs) can be chemically synthesized and correspond to the cleavage products from longer dsRNA.

Studies in S2 cells and *Drosophila* embryos identified 21-25 nt siRNAs generated from larger dsRNAs by the dsRNA-specific endonuclease Dicer (Bernstein et al. 2001; Nykanen et al. 2001). Dicer is evolutionary conserved in *C. elegans* and mammalian cells and has an N-terminal helicase domain, a PAZ motif, a dsRNA binding domain and two RNAse III motifs at the C-terminus (Filippov et al. 2000). siRNAs generated by Dicer serve as guide sequences to induce target-specific mRNA cleavage by the RNA induced silencing complex (RISC) (Hamilton and Baulcombe 1999; Zamore et al. 2000; Hammond et al. 2000; Yang et al. 2000). Taken together, the following mechanism for siRNA-mediated target mRNA-degradation has been proposed (Figure 1). RNAi is initiated by the enzyme Dicer which processes long dsRNA into 21-23 nt siRNA duplexes containing a 5'-phosphate and 3'-hydroxylated overhangs of 2-3 nt (Zamore et al. 2000; Bernstein et al. 2001; Elbashir et al. 2001). The resulting siRNAs are incorporated into the RISC (Nykanen et al. 2001; Elbashir et al. 2001) which in turn is activated in an ATP-dependent manner (RISC*) unwinding and separating the strands of the siRNA duplex. Then RISC* uses the antisense siRNA strand as a guide to recognize the target mRNA via base-pairing (Nykanen et al. 2001). The target mRNA is subsequently cleaved near the center of the region covered by the antisense strand siRNA (Elbashir et al. 2001; Nykanen et al. 2001). Subsequently, the cleaved target mRNA is further degraded by a cellular exonuclease (Holen et al. 2002).

1.2 DNA REPAIR MECHANISMS

Since L1 was shown to produce DSBs (Gasior et al. 2006; Belancio et al. 2010) I will focus on DSB repair pathways. Furthermore, in this thesis I will focus on mammalian DNA repair pathways, since all experiments were performed in human cell culture systems.

Cell division is under the control of several cell-cycle checkpoints in the G1, S, G2 and M phases. Activation of any of these checkpoints leads to cell-cycle arrest (Hoeijmakers 2001), which is assumed to be necessary to provide sufficient time for adequate DNA repair. Cell-

cycle arrest upon DSBs is thought to be mediated primarily by members of the phosphatidylinositol 3-kinase-like kinase (PIKK) family. These factors are activated through DNA breakage (Durocher and Jackson 2001; Khanna and Jackson 2001; Lobrich and Jeggo 2005) and activation leads to at least two PIKK signaling cascades, the ataxia telangiectasia mutated (ATM) and ataxia telangiectasia related (ATR) pathways. Those in turn lead to phosphorylation and activation of the central cell-cycle monitoring proteins p53, Chk1 and Chk2, which then are able to conduct G1 and G2 arrest via Cdk2/Cyclin E or Cdk2/Cyclin B1 (Khanna and Jackson 2001). In that way, the division of a DSB-containing cell is halted to allow the repair through several sets of enzymes.

Two major pathways that mediate DSB repair have been identified in mammalian cells: homologous recombination (HR) and non-homologous end-joining (NHEJ). So far it has not been entirely clarified which factors determine the use of HR versus NHEJ. Nevertheless, cell-cycle stage must play an important role in this decision, as the prerequisites for HR are only present during S and G2 phases, which will become clear in the following sections. Consistent with this hypothesis, it was shown that HR is nearly absent in G1, most active in S phase and declines in G2/M. In contrast, NHEJ is active throughout the cell cycle in principle and its activity increases as cells progress from G1 to G2/M (G1<S<G2/M) (Mao et al. 2008).

1.2.1 HOMOLOGOUS RECOMBINATION

During HR a homologous stretch of DNA on the corresponding sister chromatid serves as template for the repair of the broken strand. HR has different functional contexts in DSB repair, replication fork support and interstrand crosslink (ICL) repair, necessitating context-specific factors. However, the core reaction of homology search and DNA strand invasion is common to all HR reactions. HR comprises a series of related subpathways, all yielding high fidelity repair. In sum there are three variations of HR known so far: (1) the classical DSB repair (DSBR), (2) the synthesis-dependent strand annealing (SDSA) and (3) the break-induced replication (BIR) (Figure 2).

Common to all subpathways of HR is the processing of the DSB to a 3' overhang, to which the assembly of the Rad51 filament is directed. This phase is termed presynapsis and involves the 5'-3' exonuclease Exo1 (Fiorentini et al. 1997), SAE2 (Clerici et al. 2005) and the nuclear resection complex Mre11-Rad50-Nbs1 (MRN) (Mimitou and Symington 2009).

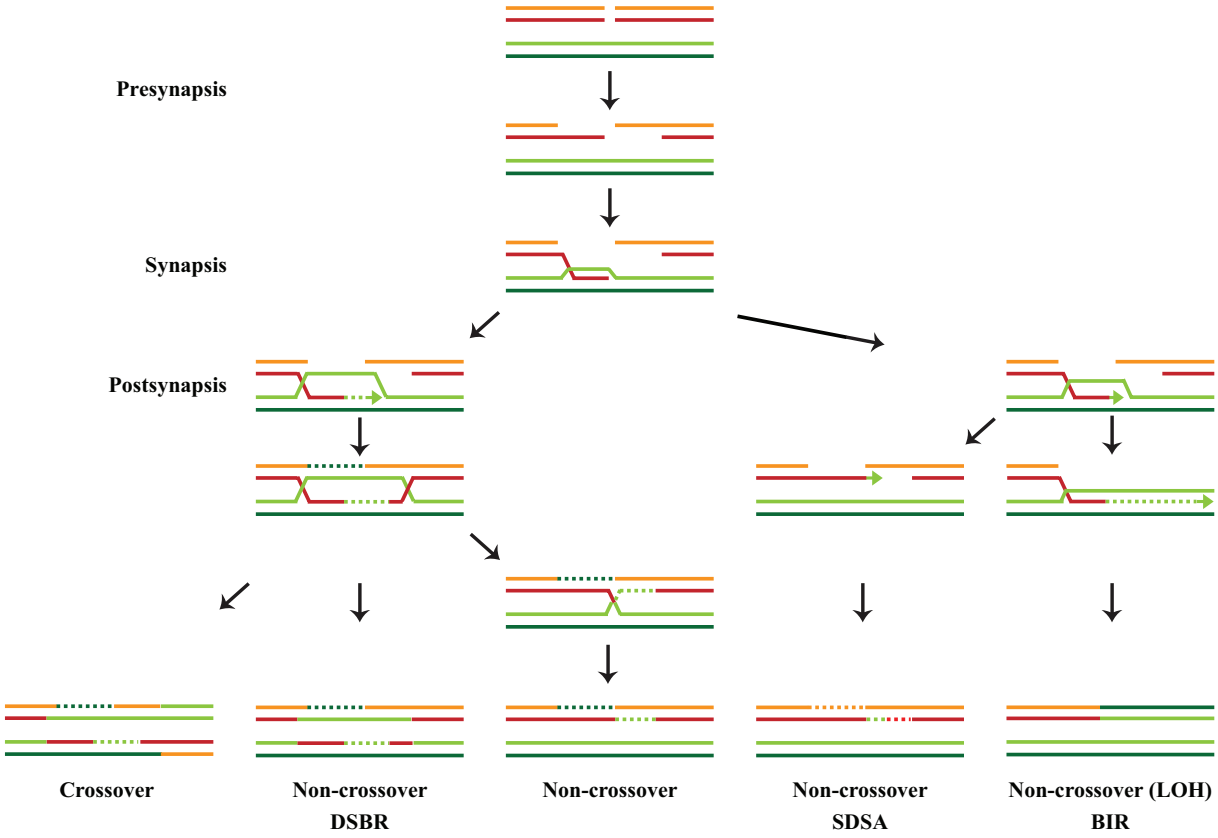


Figure 2 Pathways of homologous recombination (HR). HR can be divided into three stages: presynapsis, synapsis and postsynapsis. Presynapsis involves recognition of DSB ends and processing to single stranded 3' OH ends. During synapsis a D-loop is formed through DNA strand invasion by the Rad51-ssDNA complex. Then, at least three following pathways for strand resolution are proposed. In DSB repair (DSBR) the ends of a DSB are either exposed to independent strand invasion or to second end capture, thereby forming a double Holliday junction. This structure can either be processed by a resolvase into non-crossover or crossover products (both outer left diagrams at the bottom) or by BLM-mediated branch migration and TOPOIIIα leading to non-crossover products (center diagram at the bottom). During synthesis-dependent strand annealing (SDSA) the invading strand is released after DNA synthesis and annealed to the second end yielding localized conversion without crossover. In break-induced replication (BIR) the D-loop forms a full replication fork to copy the complete distal part of the chromosome leading to loss of heterozygosity (LOH).

SsDNA is initially bound by the ssDNA-binding protein RPA, which displays higher affinity and specificity for ssDNA than Rad51 (Wold 1997; Sung and Klein 2006), meaning that Rad51 must assemble on RPA-coated ssDNA. RPA is a heterotrimeric ssDNA-binding protein in eukaryotes and involved in all DNA metabolic processes concerning ssDNA (Wold 1997). Rad51 binds to ssDNA in a ternary complex with ATP at a stoichiometry of 1 protomer per 3-4 nucleotides (Ogawa et al. 1993; Conway et al. 2004). Furthermore, for formation of Rad51 filaments *in vivo* mediators are required. Using chromatin immunoprecipitation, the Rad55-Rad57 complex and Rad52 have been identified as key players of Rad51 filament formation in budding yeast (Gasior et al. 1998; Sugawara et al.

2003; Lisby and Rothstein 2004). There are five human Rad51 paralogs (Rad51B, Rad51C, Rad51D, Xrcc2, Xrcc3) potentially corresponding to Rad55-Rad57. But the overall low sequence conservation makes it difficult to assign which of these candidates play the role of the yeast Rad55-Rad57 in humans (Thacker 1999; Sung and Klein 2006). The mechanisms by which the Rad51 paralogs function as mediators are unknown, but all are required *in vivo* for Rad51 filament formation.

The homology search and DNA strand invasion by the Rad51-ssDNA filament are termed synapsis. During this process the D-loop intermediate is formed, positioning the invading 3' end on a template duplex DNA to initiate repair synthesis. Another protein needed in synapsis is Rad52, which forms a multimeric ring structure that binds preferentially to ssDNA on the outside of the ring through an N-terminal DNA binding domain (Shinohara et al. 1998; Singleton et al. 2002).

A mediator protein of special interest is the tumor suppressor BRCA2. Heterozygous mutations in BRCA2 predispose to breast, ovarian and other tumor types in humans, and biallelic loss of BRCA2 causes Fanconi anemia (FA) (West 2003; Pellegrini and Venkitaraman 2004; Kennedy and D'Andrea 2005). It could be shown that BRCA2 targets Rad51 filament formation to the ssDNA-dsDNA junction on RPA-coated ssDNA (Yang et al. 2002; Yang et al. 2005; San Filippo et al. 2006). The structure of BRCA2 mimics the Rad51 subunit-subunit interface, suggesting that BRCA2 might act as a point of nucleation for the Rad51 filament (Pellegrini et al. 2002). The C-terminal Rad51-binding site of BRCA2 preferentially interacts with the filamentous form of Rad51 (Davies and Pellegrini 2007; Esashi et al. 2007), suggesting that BRCA2 promotes Rad51 filament formation by nucleation and filament stabilization. This binding is negatively regulated by CDK phosphorylation on S3291, implicating that BRCA2 function in HR is regulated throughout the cell cycle (Esashi et al. 2005).

Another mediator protein is Rad54. It is a bidirectional motor protein travelling at a speed of ~300 bp/s on dsDNA powered by ATP (Amitani et al. 2006). Rad54 is found at the pairing site (Mazin et al. 2000; Van Komen et al. 2000), where it associates with and stabilizes the Rad51 presynaptic filament (Mazin et al. 2003; Wolner and Peterson 2005). But Rad54 also functions after synapsis. Rad54 dissociates Rad51 from the heteroduplex DNA, acting as a turnover factor that allows access of DNA polymerases to the invading 3'-end (Solinger et al. 2002; Li et al. 2007). Thus, Rad54 initiates the step of postsynapsis, which can be subdivided into at least three different repair pathways, as already mentioned earlier.

The aspect of crossover avoidance is one of the key features of SDSA that made this model attractive for DSB repair in somatic cells. Recombination in somatic cells is rarely associated with crossovers, which have the potential to generate genomic rearrangements and large-scale loss of heterozygosity (LOH) (Moynahan and Jasin 1997; Richardson and Jasin 2000; Elliott and Jasin 2002).

In DSBR, the branch of the intact chromosome proceeds by engaging the second end of the DSB, either by a process termed second end capture through DNA annealing or a second invasion event. Annealing of the second end is catalyzed by Rad52, which is able to anneal complementary ssDNA still bound to RPA. This results in a double Holliday junction (dHJ), which can be processed in two different ways subsequently. Either the dHJ is dissolved into non-crossover products by BLM-TOPOIIIα or it is resolved by a structure-specific endonuclease into crossover/non-crossover products (Figure 2, left). Alternative processing of dHJs takes place through the combined action of BLM DNA helicase with type I topoisomerase TOPOIIIα in cooperation with their cofactor BLAP75 (Wu and Hickson 2003; Plank et al. 2006; Wu et al. 2006) yielding a non-crossover outcome.

In SDSA, the D-loop is dissolved after DNA synthesis and the invading strand reanneals with the second end of the DSB, always forming non-crossover products (Figure 2, middle). The process of reannealing is believed to proceed similarly to second end capture in DSBR.

In BIR, the invading strand forms a replication fork-like structure to copy the entire distal arm of the template chromosome, resulting in LOH (Figure 2, right) (Malkova et al. 1996; Lydeard et al. 2007). This means the second end of the DSB is never engaged and the genetic information of the broken chromosome after the DSB is lost. The role of BIR in wild-type cells is uncertain, but in *rad51* or *mre11* mutants, BIR contributes significantly to DSB repair (Malkova et al. 1996; Nickoloff et al. 1999; Krishna et al. 2007).

1.2.2 NON-HOMOLOGOUS END-JOINING

NHEJ is a potentially less accurate form of DSB repair, in which the two DNA ends of the broken strand are directly ligated. If necessary, this process involves editing of the DNA termini to form compatible ends. This method can result in the loss of nucleotides and thus in the loss of genetic information or genomic integrity. NHEJ can be subdivided into three stages: (1) the capture of both DNA ends of the broken molecule, (2) the formation of a synapsis complex bringing both ends together and (3) the subsequent re-ligation of the DNA termini (Figure 3).

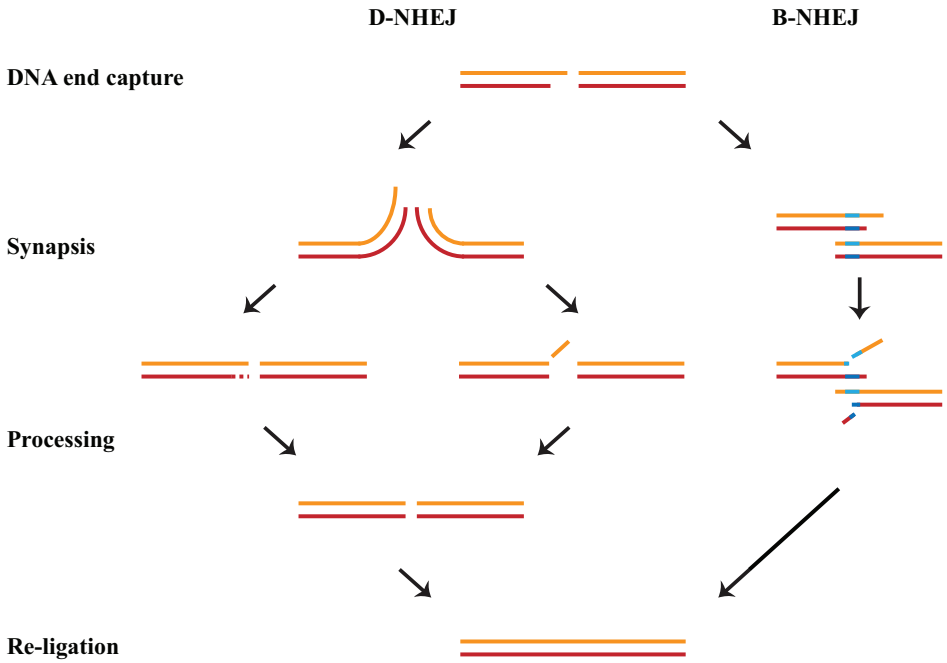


Figure 3 Pathways of non-homologous end-joining (NHEJ). During DNA end capture of D-NHEJ the Ku70/80 heterodimer associates with the two ends of the broken DNA molecule and attracts DNA-PKcs, which protects the DNA end termini from degradation. Then, the DNA-PKcs molecules on the DNA ends form a synaptic complex to tether both ends together. Subsequently, the DNA-PKcs autophosphorylation introduces a conformational change rendering the DNA termini accessible to other enzymes. Also ATM-mediated DNA-PKcs phosphorylation may play a role in this conformational change. After synapsis, non-compatible DNA ends are processed either by filling in by several polymerases or by resection by the MRN complex or Artemis before ligation can follow. Ligation is carried out by the DNA ligase IV/XRCC4 complex. Alternatively, PARP-1 may capture the DNA ends and thus direct the mode of DNA repair towards B-NHEJ. During synapsis of B-NHEJ, the ends of the broken DNA are brought into close vicinity at sites of short direct repeats. Then, both ends are processed to rejoin the molecule sparing only one copy of the microhomologies and loosing the sequence in between. Ligation is performed by ligase III.

NHEJ is initiated by the binding of the Ku70/80 heterodimer to the ends of the DNA break. Crystallographic studies revealed that the heterodimer has an open ring-shaped structure comprising one of each subunits (Cary et al. 1998; Walker et al. 2001). The center of the ring accommodates the DNA double helix and allows the Ku70/80 dimer to slide over the ends of a DSB. This provides a scaffold for the assembly of the NHEJ repair machinery.

Early during repair the DNA-Ku complex recruits the DNA-dependent protein kinase catalytic subunit (DNA-PKcs) to the DSB. DNA-PKcs exerts multiple functions, including the formation of a bridging complex bringing both DNA ends together. Electron microscopy images resolve dimeric structures consisting of two DNA ends tethered by a protein complex containing two Ku70/80 dimers and two DNA-PKcs molecules (Spagnolo et al. 2006; Llorca 2007; Rivera-Calzada et al. 2007). Furthermore, Ku is necessary for the activation of the

DNA-PKcs serine/threonine kinase, which in turn phosphorylates Ku70/80, XRCC4, p53 and itself (Smith and Jackson 1999; Chan et al. 2002; Lees-Miller and Meek 2003). Upon autophosphorylation, DNA-PKcs changes its conformation and dynamics at the synaptic complex (Weterings and Chen 2007) and allows processing and ligation of the DNA termini (Weterings et al. 2003; Block et al. 2004; Reddy et al. 2004).

Then non-ligatable DNA ends are processed before DSB repair can take place. Multiple enzymes have been identified, that account for removing or filling-in single-stranded, non-compatible overhangs. Among those are nucleases, as for instance the MRE11/RAD50/NBS1 (MRN) complex and Artemis, and polymerases, e.g. DNA polymerase μ and λ. In contrast to the nucleases, the polymerases also have functions during other cellular processes and cannot be considered NHEJ-specific. During this step of nucleolytic resection occassional loss of nucleotides can occur. Finally, NHEJ is completed by ligation of the processed DNA ends through the complex DNA ligase IV/XRCC4. The three-dimensional structure of the ligase IV/XRCC4 complex reveals two XRCC4 molecules side by side associated with one ligase IV molecule (Sibanda et al. 2001; Modesti et al. 2003). It is assumed that the amino-terminal ends of the XRCC4 molecules mediate an interaction with the DNA helix, while ligase IV repairs the DSB.

Apart from that, another indispensable NHEJ key component protein, XLF/Cernunnos, was characterized (Ahnesorg et al. 2006; Buck et al. 2006). XLF promotes the DNA ligation function of noncohesive ends by the ligase IV/XRCC4 complex (Akopiants et al. 2009; Riballo et al. 2009).

The pathway described above is referred to as D-NHEJ to indicate its dependence on DNA-PKcs (Figure 3, left). Cells with mutations in components of D-NHEJ repair the majority of DSBs utilizing a slower process, which is referred to as B-NHEJ to indicate its backup function (Figure 3, right). B-NHEJ is not sensitive to mutations in genes assigned to HR (DiBiase et al. 2000; Wang et al. 2001; Iliakis et al. 2004) and is usually suppressed by D-NHEJ (Perrault et al. 2004; Wang et al. 2006). It has been shown that DNA ligase III is involved in B-NHEJ (Audebert et al. 2004; Wang et al. 2005) as well as PARP-1 and XRCC1, which are also implicated in the repair of single strand breaks (Audebert et al. 2004). PARP-1 and Ku were shown to directly compete for the binding of free DNA ends and thus repair pathway choice (Wang et al. 2006). However, the higher affinity of Ku for DSBs limits the contribution of B-NHEJ to DSB repair in healthy cells.

Apart from the slower repair kinetics, cells deficient in components of D-NHEJ show DNA end joining with an increase in the use of microhomologies (Kabotyanski et al. 1998; Verkaik

et al. 2002) (Figure 3, right). Microhomologies are a couple of overlapping nucleotides in the sequences of the rejoined DNA ends. This microhomology-dependent DSB repair is suspected to overlap with B-NHEJ and was shown to be involved in the assembly of antigen receptor genes (V(D)J recombination) (Corneo et al. 2007; Nussenzweig and Nussenzweig 2007; Soulas-Sprauel et al. 2007; Yan et al. 2007). In *in vivo* and *in vitro* assays B-NHEJ was found to be enhanced in G2 of the cell cycle (Wu et al. 2008).

A significantly increased number of microhomologies is also found between the 5' end of L1 insertions and the flanking host DNA in *de novo* integrants in cultured cells and in transgenic mice as well as in a large number of endogenous L1s in the human and mouse genomes (Gilbert et al. 2002; Symer et al. 2002; Gilbert et al. 2005; Martin et al. 2005; Zingler et al. 2005b; An et al. 2006; Babushok et al. 2006; Suzuki et al. 2009). These findings point to the involvement of B-NHEJ in the process of 5' end attachment of L1 insertions. But L1 elements are not only relying on cellular mechanisms for their integration into the host genome, they can also mediate DSB repair themselves, as will be delineated in the following section.

1.2.3 DNA REPAIR BY LINE-1 RETROTRANSPOSITION

L1 retrotransposition events are characterized by typical structural features. These are a 3' poly-A tail, frequent 5' truncations of L1 in combination with microhomologies, and flanking target site duplications (TSDs) of the genomic sequence. Furthermore, the L1 endonuclease displays a preference for target sites following the consensus sequence 5'-TTAAAA-3', where it cleaves the bottom DNA strand 5'-TTTT ↑ AA-3' at the site indicated by the arrow. However, in this section I will focus on the role of L1 in DNA repair, that was found in 2002 and 2007 (Morrish et al. 2002; Morrish et al. 2007).

L1 elements not only mobilize themselves making use of their own endonuclease (EN) generating all the hallmarks just mentioned, they are also capable of retrotransposing in an EN-independent fashion into preexisting nicks of the DNA. This feature was found applying L1 EN null mutants in cell culture retrotransposition assays, where they retrotranspose readily, although at much lower frequencies compared to wildtype elements. Integration sites of L1 *de novo* insertions resulting from EN-independent retrotransposition events differ from the L1 EN consensus target sequence, lack TSDs and are frequently 3' truncated which means that they also lack a poly-A tail. Retrotransposition frequencies of EN-deficient L1 elements almost reach wildtype levels in NHEJ deficient chinese hamster ovary (CHO) cells. These findings show that L1s can integrate into DNA lesions, resulting in retrotransposon-mediated

DNA repair in mammalian cells (Morrish et al. 2002). In a DNA-PKcs deficient CHO cell line, EN-deficient L1 elements not only retrotranspose at almost wildtype levels, they even show a preference to integrate at telomeres, since the defect in NHEJ also leads to dysfunctional telomeres (Gilley et al. 2001; Goytisolo et al. 2001). Notably, ~30% of the observed retrotransposition events insert in an orientation-specific manner adjacent to a perfect telomere repeat (5'-TTAGGG-3'). Transient expression of a dominant-negative allele of human telomeric repeat binding factor 2 (*terf2*) in XRCC4-deficient CHO cells which disrupts telomere capping, elicits the same effect. The findings reveal similarities between the mechanisms of EN-independent retrotransposition and telomerase activity, since both processes make use of a 3' OH for priming reverse transcription at either DNA lesions or chromosome ends. Therefore, it was hypothesized that EN-independent retrotransposition might be an ancestral mechanism of RNA-mediated DNA repair that made use of non-LTR retrotransposons before an AP-endonuclease was acquired to repair DNA lesions in the host cell (Morrish et al. 2007). This hypothesis was corroborated by Kopera et al. (Kopera et al. 2011) using an *in vitro* assay to detect L1 reverse transcriptase activity to demonstrate that wildtype or EN-defective L1 ribonucleoprotein particles can use oligonucleotide adapters mimicking telomeric ends as primers to initiate the reverse transcription of L1 mRNA. They demonstrated that ORF1p is not strictly required for EN-independent retrotransposition at dysfunctional telomeres.

1.3 TRANSPOSABLE ELEMENTS IN THE HUMAN GENOME

In *Homo sapiens*, approximately 46% of the genome were formed by transposable elements, whereas only 0.5-1.5% comprise functional genes (Lander et al. 2001). The greatest portion of repetitive DNA in humans originates from retrotransposons. They can be divided into two major classes that are phylogenetically and structurally unrelated. One of them are the long terminal repeats (LTR)-containing retrotransposons, accounting for 8.3% of the human genome, the others are the non-LTR-retrotransposons covering more than 34%. Figure 4 shows the composition of the human genome.

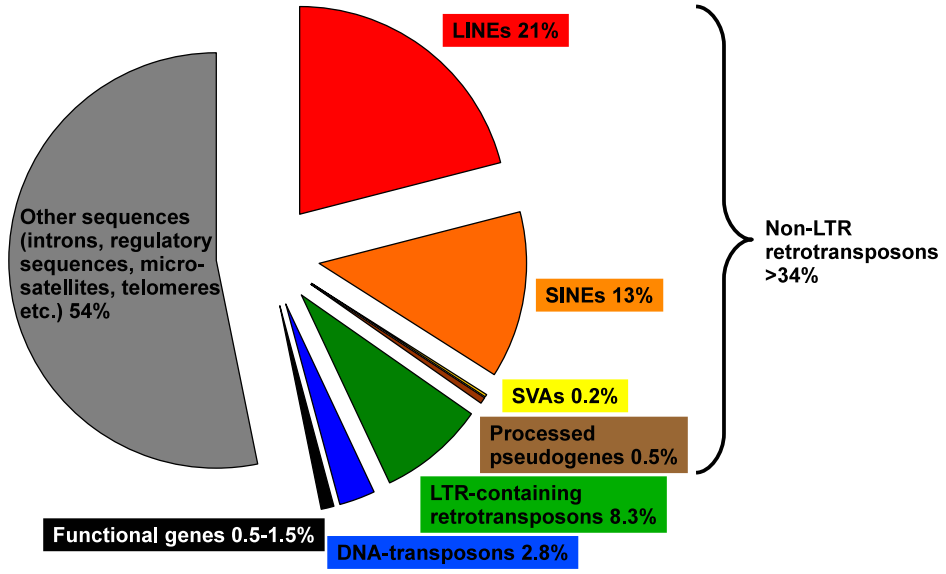


Figure 4 Composition of the human genome. 0.5-1.5% of the human genome consist of exons belonging to functional genes (black), whereas transposable elements make up more than 45%. Transposable elements include fossils of DNA-transposons (2.8%, blue), LTR-containing retrotransposons (8.3%, green) as well as non-LTR-retrotransposons (>34%). The group of non-LTR-retrotransposons can be further subdivided into LINEs (~21%, red), SINEs (~13%, orange), SVAs (~0.2%, yellow) and processed pseudogenes (~0.5%, brown). Other sequences (~54%, gray) comprise introns, regulatory sequences, microsatellites, telomeres etc.

1.3.1 DNA-TRANSPOSONS

As little as 2.8% of the human genome are derived from DNA-transposons (Figure 4) which move via a DNA intermediate in a cut-and-paste fashion mediated by an enzyme named transposase that is encoded by autonomous elements as, for instance, the Mariner transposon. Non-autonomous elements carry deletions or point mutations within the transposase coding region. All DNA-transposons are flanked by terminal inverted repeats (TIR) (Figure 5). Nevertheless, all ~294000 elements identified in the human genome are merely genomic fossils that have been inactive for at least 50 million years (Smit and Riggs 1996; Lander et al. 2001; Babushok and Kazazian 2007).

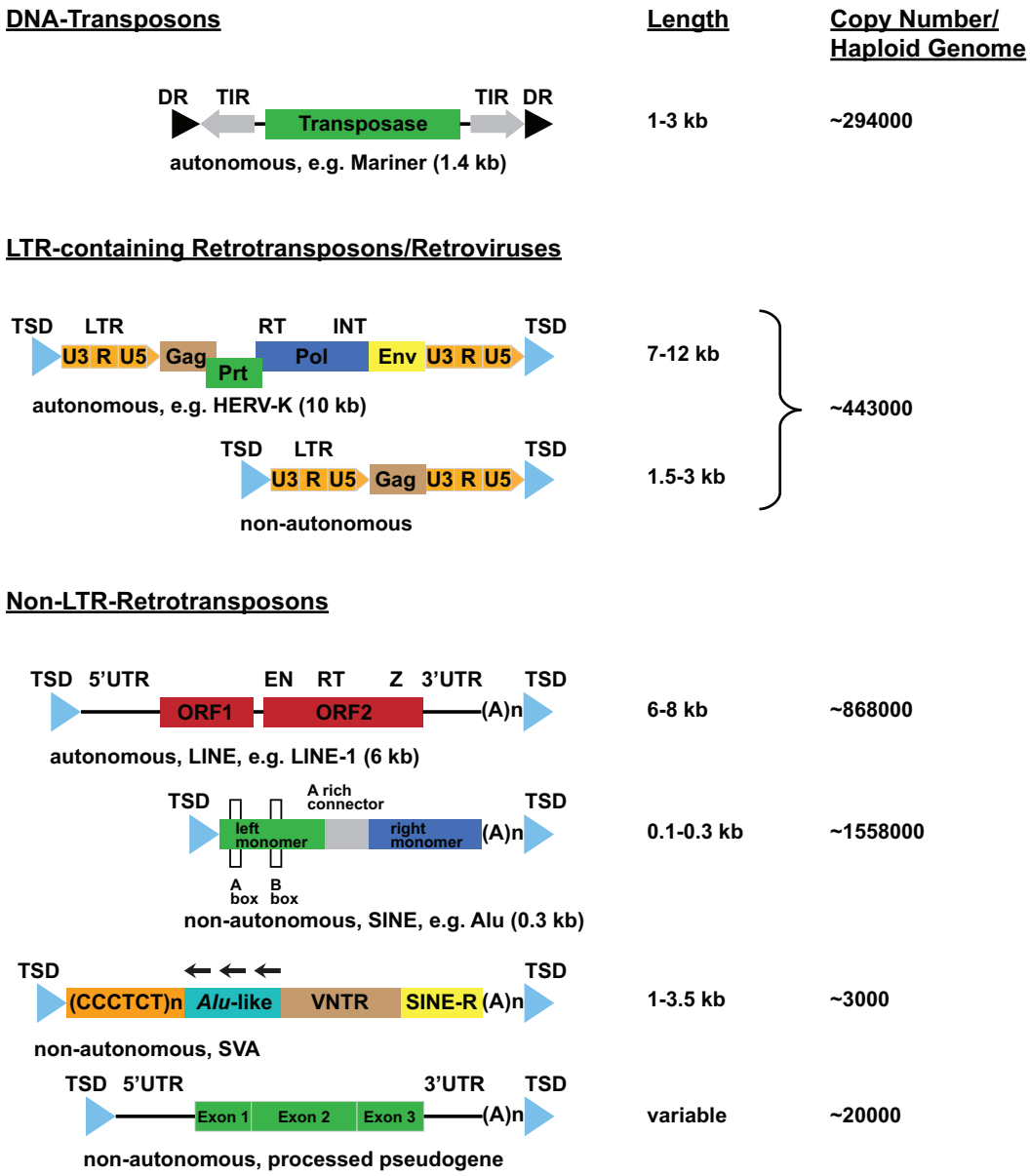


Figure 5 Transposable elements in the human genome. DNA-transposons are inactive relicts within the human genome. LTR-containing retrotransposons and LINE-1 elements are autonomous encoding all proteins needed for their retrotransposition. Non-autonomous retrotransposons like SINEs (e. g. Alus) and SVAs are mobilized by L1-encoded proteins. Processed pseudogenes are generated by L1-protein trans-mediated retrotransposition. DR, direct repeats; TIR, terminal inverted repeats; TSD, target site duplications; LTR, long terminal repeats; Gag, group specific antigene; Prt, protease; Pol, polymerase; Env, envelope; UTR, untranslated region; ORF, open reading frame; EN, endonuclease; RT, reverse transcriptase; Z, cysteine-rich region with homology to zinc fingers; A/B, A and B-box PolIII-promoter; VNTR, variable number of tandem repeats.

1.3.2 LTR-CONTAINING RETROTRANSPOSONS

Retrotransposons replicate via a copy-and-paste mechanism involving transcription of the complete element, reverse transcription of the RNA into cDNA and subsequent integration into a new genomic locus. In that way, one functional retrotransposon can generate several

copies of itself. LTR-containing retrotransposons are also called retrovirus-like elements or human endogenous retroviruses (HERVs) because they are derived from exogenous retroviruses which infected the germline several times during primate evolution. However, due to mutations, they have lost their ability to replicate and infect throughout the last 30 million years (Lower et al. 1996). Each HERV contains LTRs flanking the element that can be subdivided into three regions, U3, R and U5 (Figure 5). Promoter elements are located within the U3 region of the 5' LTR, whereas the 5' end of the R region defines the transcription starting point. A polyadenylation signal is found within the R region of the 3' LTR followed by the site of polyadenylation at the junction between R and U5 region. The 7-12 kb coding region of HERVs contains *gag*, *pol* and *env* genes. The *gag* (group-specific antigene) gene product is cleaved by a HERV-encoded protease and processed into matrix, capsid and nucleocapsid proteins. The *pol* (polymerase) gene codes for a polyprotein which is also processed by the self-encoded protease into three enzymes: protease, reverse transcriptase and integrase. The envelope proteins of the virus are transcribed from the *env* gene. Some HERVs possess additional genes coding for accessory proteins (Goodier and Kazazian 2008).

Based on the tRNA used to initialize the reverse transcription (e.g. Lysin-tRNA for HERV-K) and the sequence homologies to exogenous retroviruses, the 443000 HERVs were classified into approximately 30 HERV subgroups. Only the HERV-K members still contain intact ORFs for all functional proteins listed above (Lower et al. 1993; Mueller-Lantzsch et al. 1993; Holmes 2007) and the most recent HERV-K insertions occured approximately 20000 years ago as determined by polymorphisms (Turner et al. 2001). Expression of HERV-K proteins was found in melanoma and teratocarcinoma cell lines (Muster et al. 2003; Dewannieux et al. 2005), but no evidence for replication-competent HERV-K proviruses was found until 2015, where a group observed HERV-K102 expression, particle production and replication that were associated with foamy macrophage generation in the cultures of cord blood mononuclear cells (Laderoute et al. 2015). All other human LTR-containing retrotransposons are genomic fossils that have not been mobilized for the last 40 million years (Lander et al. 2001; Turner et al. 2001).

1.3.3 NON-LTR-RETROTRANSPOSONS

Non-LTR-retrotransposons are the dominating class of retrotransposons in the human genome and are evolutionary more ancient. Sequence alignments indicate that they have a common

origin with reverse transcriptase (RT)-bearing group II introns of bacteria and mitochondria (Xiong and Eickbush 1990; Yang et al. 1999). Comprising more than one third of the human genome (34%), non-LTR-retrotransposons had the greatest impact on our genome during evolution (Figure 4).

In general, non-LTR-retrotransposons can be classified into two subtypes, the restriction enzyme (RE)-type and the apurinic/apyrimidinic endonuclease (APE)-type non-LTR retrotransposons (Yang et al. 1999). RE-type non-LTR-retrotransposons are characterized by a single open reading frame (ORF) with an RE-like EN domain following the RT domain. This EN domain is related to type-IIS restriction endonucleases with a DNA-cleavage domain separated from a DNA-binding domain and is usually sequence-specific (Eickbush 2002). However, the human genome does not contain any RE-type elements (Lander et al. 2001), although they represent the oldest lineage of non-LTR-retrotransposons (Malik and Eickbush 2001). The human genome contains retrotransposons of the second subtype, the class of APE-type non-LTR-retrotransposons. Members of this much larger family are characterized by two ORFs and an EN domain that is distantly related to the apurinic/apyrimidinic (AP) endonucleases (Martin et al. 1995; Feng et al. 1996). The EN domain is located upstream of the RT domain in ORF2p. Based on the elements' structures and on phylogenetic analyses of their RT domains, there are four groups of APE-type non-LTR-retrotransposons, which can further be subdivided into 11 clades (Burke et al. 1999; Malik et al. 1999; Lovsin et al. 2001; Malik and Eickbush 2001; Zingler et al. 2005a).

Currently, the Long Interspersed Elements (LINEs) represent the only autonomous class of non-LTR-retrotransposons encoding the protein machinery required for their mobilization. There are three LINE-families within the human genome, termed LINE-1 (L1), LINE-2 (L2) and LINE-3 (L3). Human L1 elements belong to the L1 clade comprising three members: species comprehensive L1, *Xenopus laevis* Tx1L and *Candida albicans* Zorro3 elements (Zingler et al. 2005a). The diploid human genome contains 80-100 active L1 elements with two intact ORFs coding for functional proteins (Brouha et al. 2003). In contrast, L2 and L3 elements have accumulated numerous mutations over time yielding in the loss of autonomous retrotransposition competence 80-100 million years ago (Lander et al. 2001). LINE-1 encoded proteins display a high *cis* preference, which means that they preferentially bind to the mRNA molecule they were encoded by, assuring that mainly functional copies are replicated (Moran et al. 1996; Boeke 1997; Esnault et al. 2000; Wei et al. 2001). This *cis* preference is thought to be mediated by the region between the EN and RT domains in ORF2p at which the proteins bind the stem-loop structure of the 3' tail RNA (Hayashi et al. 2014). However, some non-L1

mRNAs can interfere with this *cis* preference and are able to recruit L1 proteins for their own proliferation. The most prominent example of such "parasitic" RNA belongs to the class of Short Interspersed Elements (SINEs), it is the non-autonomous *Alu* element.

SINEs are usually between 100-300 bp long, carry an internal PolIII-promoter and cover approximately 13% of the human genome with roughly 1558000 copies. *Alu* elements are about 300 nucleotides in length covering approximately 10.6% of the human genome with more than one million copies (Cordaux and Batzer 2009). For reverse transcription and integration, *Alu* elements depend on L1 proteins (Smit 1996; Boeke 1997; Dewannieux et al. 2003). The remarkable capability of *Alu* RNA to hijack L1 proteins for their own mobilization in *trans*, is thought to be due to its secondary structure. *Alu* elements have a dimeric structure derived from the 7SL RNA gene (Ullu and Tschudi 1984; Quentin 1992). Left and right *Alu* monomer are very similar and separated by an adenosine-rich sequence (Figure 5). 7SL RNA is the scaffold of the signal recognition particle (SRP) binding to nascent signal peptide sequences and transiently arresting translation (Siegel and Walter 1988). By mimicking this structure, *Alu* RNA may have adopted the ability to associate with ribosomes, get in close physical proximity to nascent L1 proteins and capture them for its own replication (Boeke 1997; Weichenrieder et al. 2000; Dewannieux et al. 2003).

SVA elements are the most recent active group of non-LTR-retrotransposons found in the human genome covering approximately 0.2% including roughly 3000 copies (Wang et al. 2005; Hancks and Kazazian 2012). SVAs are primate-specific and retrotranspose since approximately 25 million years. They have a modular structure (Figure 5) which is responsible for their name spelled in "antisense" orientation. At the 5' end, there is a hexameric repeat of (CCCTCT)n followed by a 355 bp *Alu*-like region comprised of three *Alu* sequences in antisense orientation and a sequence of unknown origin. Downstream of this arrangement, there is a variable number of tandem repeats (VNTR) consisting of monomers of 36-42 bp and 49-51 bp, respectively. This VNTR region spans about 3000 bp in length. An SVA's 3' end is formed by a short interspersed nuclear element of retroviral origin (SINE-R) composed of 490 bp and derived from the 3' end of the *env* gene as well as the 3' LTR of HERV-K 10 ab (Damert et al. 2009). Therefore, "SVA" is the composite of "SINE-R", "VNTR" and "Alu". Transcription of SVA elements is driven by host promoters within the 5' flanking regions and retrotransposition is achieved in *trans* by the L1 protein machinery (Wang et al. 2005; Damert et al. 2009; Raiz et al. 2012).

In rare cases, the *cis* preference of LINE-1 elements is also bypassed by spliced mRNAs of RNA Polymerase II genes. This results in so-called processed pseudogenes, an intronless and

promoterless copy of the original gene, followed by a poly-A tail and flanked by target site duplications (Figure 5) (Vanin 1985; Esnault et al. 2000).

All autonomous non-LTR-retrotransposons like LINE-1 harbor three major compounds: (1) a promoter to drive transcription of a full-length RNA, (2) a reverse transcriptase (RT) to produce cDNA of the RNA and (3) a protein machinery that mediates integration of the cDNA into a new genomic site. While in LTR retrotransposons the latter function is conducted by an element-encoded integrase (Curcio and Derbyshire 2003), in non-LTR retrotransposons this is initiated by an element-encoded endonuclease (EN).

1.4 STRUCTURE OF LINE-1 ELEMENTS AND MECHANISM OF RETROTRANSPOSITION

A complete, retrotransposition-competent L1 element is ~6 kb in length and carries two open reading frames (ORFs) (Figure 6). The 5' untranslated region (UTR) of L1 is ~900 bp in length and includes an internal RNA polymerase II (Pol II) promoter controlling L1 transcription (Kroutter et al. 2009). Driven by this L1 promoter, a single, protein-encoding, polyadenylated transcript containing several oligo-T stretches is produced. The L1 promoter exerts not only sense, but also antisense activity, meaning that L1 can drive transcription of neighboring genomic regions upstream and antisense of its own coding sequence via an antisense promoter (ASP) (Speek 2001; Roman-Gomez et al. 2005; Cruickshanks and Tufarelli 2009). L1 ORF1 is 1017 bp in length and codes for a 40 kDa protein (ORF1p). ORF1p is indispensable for the activity of APE-type retrotransposons (Moran et al. 1996) and its phosphorylation is a prerequisite (Cook et al. 2015). Furthermore, ORF1p was demonstrated to have nucleic acid chaperone activity (Martin and Bushman 2001; Martin 2006). ORF2 is separated from ORF1 by a 63 bp spacer region. It was found that translation of ORF2 from the bicistronic RNA is accomplished by an unconventional termination/reinitiation mechanism (Alisch et al. 2006). ORF2 codes for a 150 kDa polyprotein harboring an N-terminal AP-like EN as well as an RT domain (Scott et al. 1987; Mathias et al. 1991). At the C-terminal end, there is a cysteine-rich region with homologies to zinc finger domains. It has been shown that mutations in this region abolish retrotransposition in cultured cells (Moran et al. 1996). The 3' UTR spans 205 bp including a polyadenylation signal and terminates in a poly-A tail (Grimaldi et al. 1984). This portion of the L1 element is little conserved within and between species (Scott et al. 1987), and no functional role of the 3' UTR has been found so far. Interruption of this region by additional nucleotides does not

seem to have any effect on retrotransposition and therefore, the 3' UTR is frequently used to integrate reporter genes of up to 3500 bp (Moran et al. 1996; Ostertag et al. 2000; Gilbert et al. 2002; Symer et al. 2002; Xie et al. 2011).

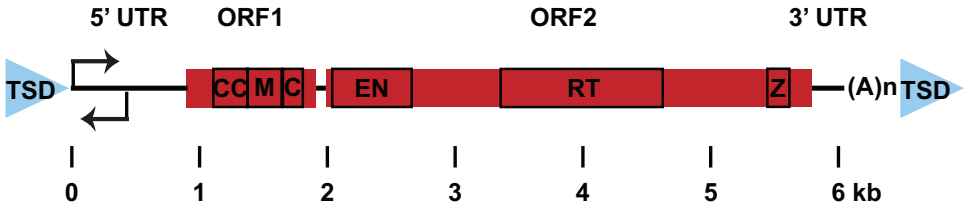


Figure 6 Structure of the human non-LTR-retrotransposon LINE-1. There are 80-100 active and retrotransposition-competent L1 elements in the human genome. 5'/3' UTR, 5'/3' untranslated region; ORF1/2, open reading frame 1/2; TSD, target site duplication; black arrows, sense and antisense promoter; CC, coiled coil domain; M, RNA-binding domain; C, conserved C-terminal domain; EN, endonuclease domain; RT, reverse transcriptase domain; Z, cysteine-rich domain with homology to zinc fingers; (A)n, poly-A tail.

All these structural hallmarks apply to full-length copies of L1. However, in humans only 5% of the endogenous L1 elements are full-length. The remaining 95% are 5' truncated and/or internally rearranged (Szak et al. 2002). Genomic L1 integrants are usually flanked by variable target site duplications (TSDs) with lengths of up to 60 bp (Szak et al. 2002) which are a consequence of the replication mechanism as explained in the following sections. L1 elements that are not flanked by TSDs may be the result of integration into blunt end nicking sites (Van Arsdell and Weiner 1984; Morrish et al. 2002) or into a staggered double strand break with a 5' instead of a 3' overhang causing a deletion of the target site (Gilbert et al. 2002).

To date, the retrotransposition mechanism of non-LTR retrotransposons is not entirely understood. However, the first steps of integration have been elucidated by biochemical studies on the site-specific RE-type retrotransposon R2 from *Bombyx mori* (Luan et al. 1993). From this, a model was developed termed Target-Primed Reverse Transcription (TPRT) (Figure 7). Cost and co-workers demonstrated that the basic mechanism of retrotransposition initiation by TPRT seems to be conserved among RE-type and APE-type elements, although they belong to different families of non-LTR retrotransposons that share only little structural similarities (Cost et al. 2002). The initial steps of L1 retrotransposition were reconstituted *in vitro*, requiring only L1 ORF2 protein, L1 RNA and a target DNA. The EN domains of both R2 and L1 were shown to initiate the integration process by nicking the target DNA. Thereby a 3' hydroxyl group is generated that serves as primer for reverse transcription of the element's RNA. As mentioned earlier, TPRT of L1 can also occur at pre-formed nicks and double strand

breaks in the target DNA. Therefore, it was concluded that nicking and reverse transcription are two steps in TPRT independent of each other (Cost et al. 2002).

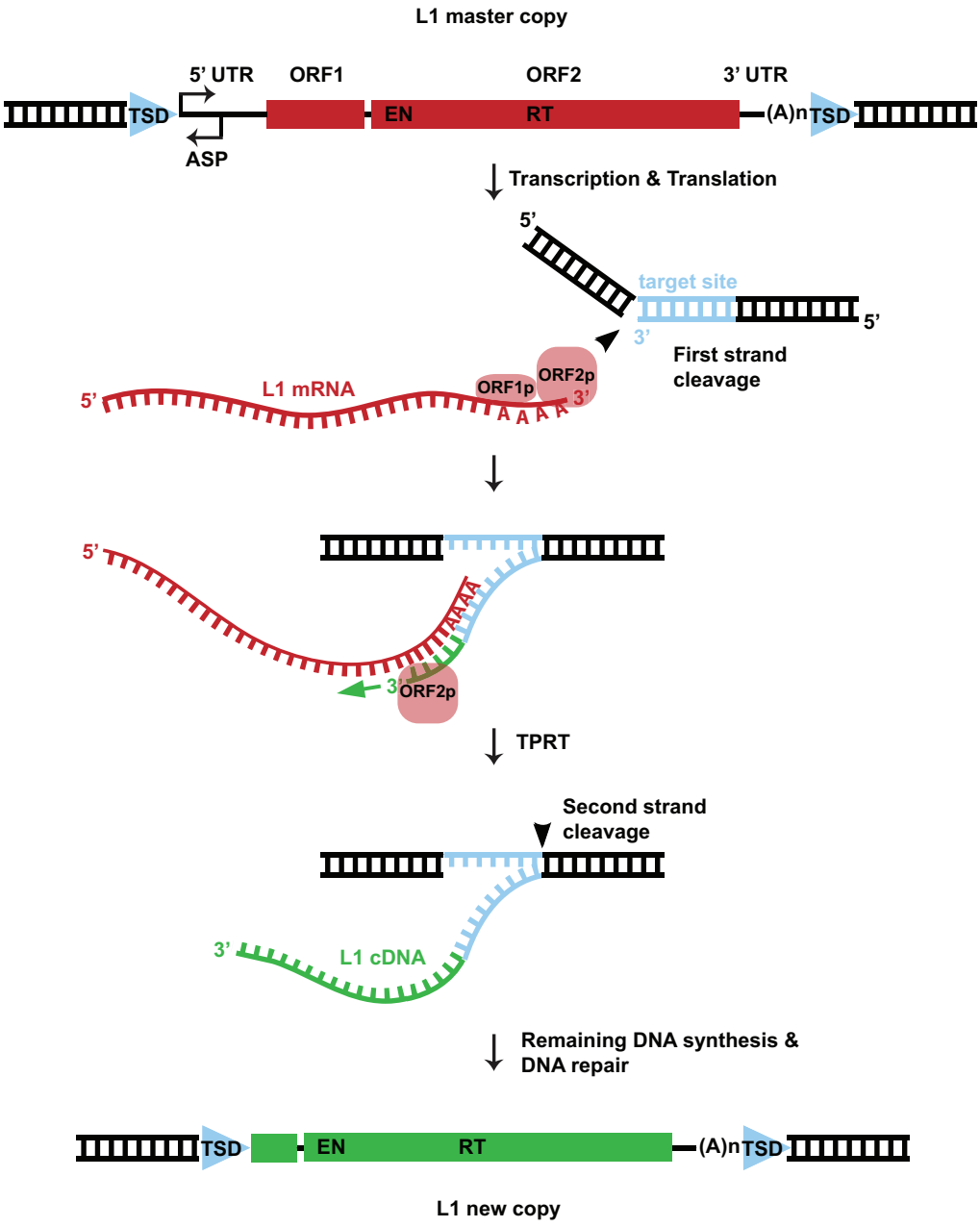


Figure 7 Schematic representation of the Target-Primed Reverse Transcription (TPRT) mechanism of L1. After L1 transcription and translation, ORF1 and ORF2 proteins associate with their own mRNA transcript. The EN domain of ORF2p initiates integration by generating a nick in the bottom strand of the genomic target DNA. Then, the RT domain uses the exposed 3' end to prime reverse transcription. After reverse transcription, cleavage of the top DNA strand occurs, creating a staggered cut. Second strand synthesis and ligation of the newly synthesised strands may be brought about either by L1-encoded enzymatic activities or by cellular DNA polymerases and repair factors. The genomic DNA is represented as black ladder with the sequence duplicated during retrotransposition (TSD) colored blue. EN, endonuclease; RT, reverse transcriptase.

So far it is unclear how second strand cleavage occurs. Depending on the position of the second nicking site relative to the first one, TPRT can yield in a deletion of genomic sequence upstream of the target site, a simple "blunt" integration, or a target site duplication (TSD) flanking the newly inserted element (Figure 8).

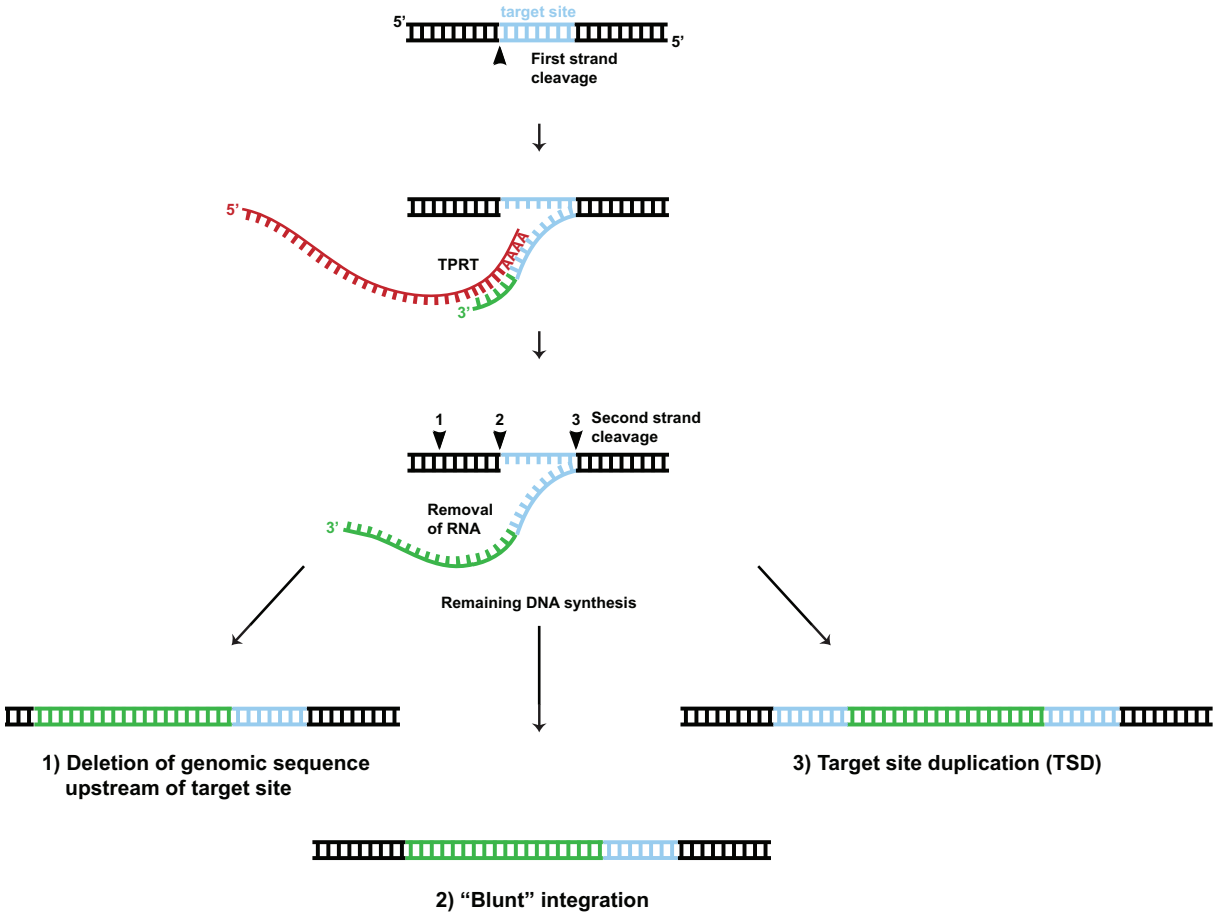


Figure 8 Alternative outcomes of TPRT depending on the position of second strand cleavage. After first strand cleavage and TPRT there are multiple positional possibilities of nicking the second strand, all resulting in one of the three depicted outcomes. (**1**) Deletion of genomic sequence upstream of the target site: A second strand cleavage upstream of the target site leads to a 5' overhang which is degraded resulting in the loss of genomic information. (**2**) "Blunt" integration: Cleavage of the second strand at the same position as the first strand leads to a simple "blunt" integration of the new element. (**3**) Target site duplication (TSD): In most cases the second strand is cleaved downstream of the first strand cleavage site. This results in TSDs flanking the newly integrated element. TSDs are variable in size depending on the distance between the two cleavage sites.

The major unanswered questions regarding the mechanism of LINE-1 retrotransposition is: What exactly happens after second strand cleavage? How is the 5' end of the newly generated element attached to the target site DNA? *In vitro* TPRT of L1 resulted in junctions between the L1 5' end and the target DNA. This indicates that the RT is able to make use of the cDNA as template for second strand synthesis, maybe by a second round of TPRT (Cost et al. 2002).

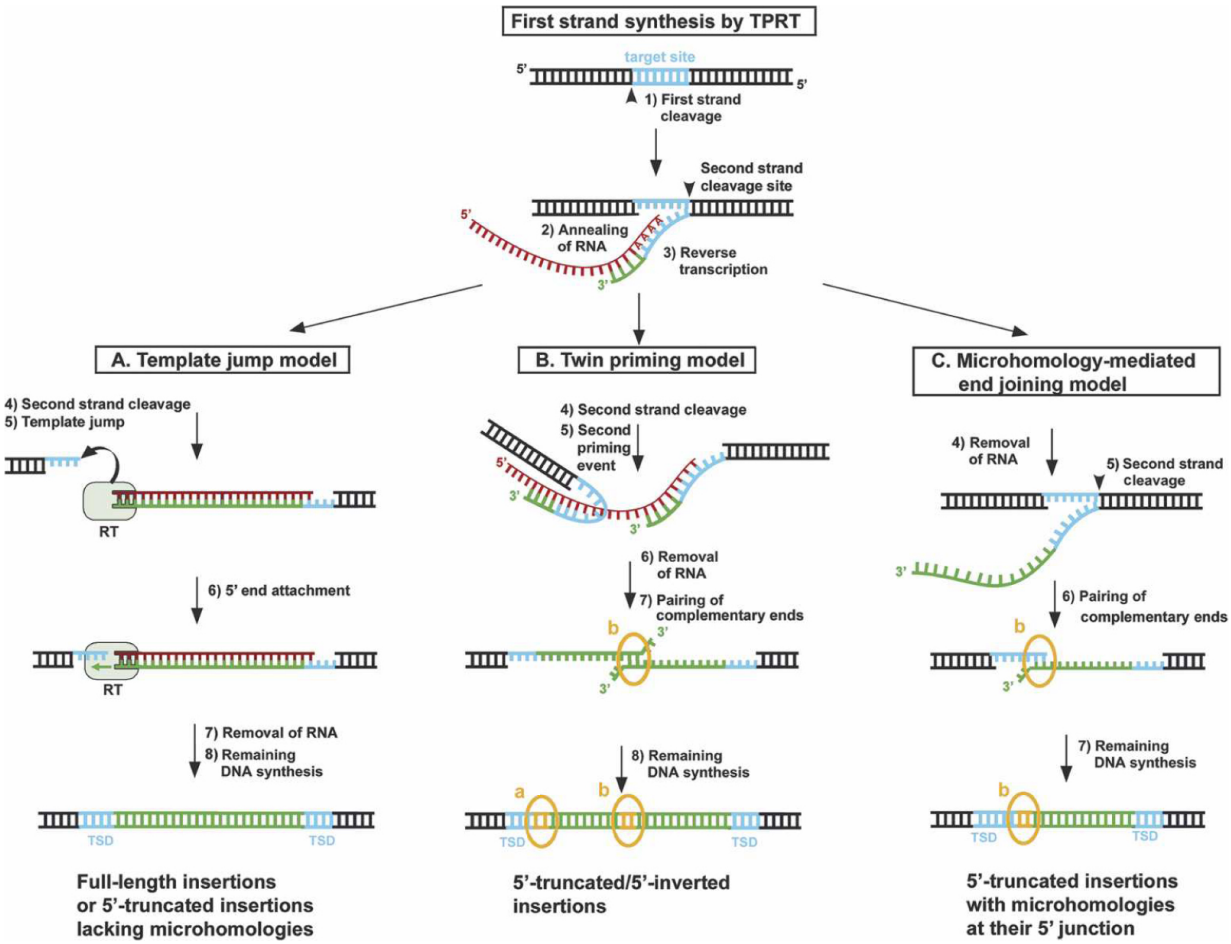


Figure 9 Alternative models for the attachment of the 3' end of the L1 cDNA to the chromosomal target DNA after second strand cleavage. According to the TPRT model, L1-encoded EN nicks the bottom strand of the target DNA (1) and exposes a 3' hydroxyl end that serves as primer for the reverse transcription of the L1 RNA (2, 3). (**A**) Template jump model: After completion of first strand synthesis, second strand cleavage occurs (4) and L1 RT jumps from the L1 RNA template onto the upstream target DNA (5). cDNA synthesis continues using the newly produced genomic 3' overhang as template. Subsequently, the 3' end of the cDNA is attached to the genomic DNA (6). In case of the RT jumping from the 5' end of a full-length RNA, full-length L1 insertions are generated. If the RT jumps from a 5' degraded L1 RNA, 5' truncated elements are generated. (**B**) Twin priming model: L1 EN cleaves the second strand (4) before reverse transcription of the first strand cDNA has been completed (3). The additional 3' end and a stretch of single-stranded DNA are used as internal primer to initiate a second round of TPRT (5) upstream of the initial TPRT and in opposite orientation. After removing the RNA from the RNA/cDNA structure (6), the single-stranded cDNAs pair at a short region of complementarity (7) and the remaining DNA synthesis is completed (8) (Ostertag and Kazazian 2001). In this model, stretches of microhomology are a consequence of RNA/DNA annealing at the 5' junction (a) and of DNA/DNA pairing at the inversion junction (b). (**C**) Microhomology-mediated end-joining model: After initiation of TPRT, reverse transcription is aborted before it has reached the 5' end of the RNA (3). Second strand cleavage (5) occurs after the termination of reverse transcription and the degradation of the L1 RNA (4). The 3' overhang of the chromosomal target site anneals to the 3' end of the L1 cDNA at a short region of complementarity (6) and primes second strand synthesis (7). This leads to the formation of a 5' truncated L1 element with microhomologies at the 5' junction (b). In all three models, the second DNA strand is joined to the target site by a cellular ligase after completion of second strand synthesis (from Zingler et al. 2005b).

However, this *in vitro* process is very inefficient and it does not necessarily reflect the natural mode of retrotransposition. Furthermore, it remains unclear how the L1-mediated damage of genomic DNA is repaired, but it is a general assumption that cellular DNA repair pathways are involved in these last steps of integration. Figure 9 offers three alternative models to answer the question of 5' end attachment which take into account naturally occuring structures of integrated elements. In most cases, newly integrated L1 elements are 5' truncated and display a short stretch of microhomology at their 5' junction to the genomic DNA (Figure 10; Zingler et al. 2005b). Microhomologies are patches of 1-12 nucleotides of micro-complementarities between the 5' end of the retrotransposon and the 3' end of the adjacent TSD derived from the genomic target site. Therefore, the origin of those overlapping nucleotides cannot be determined unambiguously. Several mechanisms have been suggested to explain the attachment of the 5' end of L1 elements to the genomic DNA generating microhomologies. The most accepted hypothesis is the microhomology-mediated end-joining model (Figure 9 C). By using bioinformatics Zingler et al. demonstrated that 5' truncated L1 elements in the human genome show a statistically significant preference for microhomologies at their 5' junctions, while full-length insertions display no distinct bias for such overlapping nucleotides at their 5' ends (Zingler et al. 2005b). Due to the structural similarity to products of B-NHEJ (Figure 3), they postulated that a cellular non-homologous DNA end-joining pathway may resolve intermediates from incomplete L1 retrotransposition events thus leading to 5' truncations, too. 3' transduction is another structural peculiarity of L1 elements. As the L1 polyadenylation signal is rather weak, it is often ignored by the RNA polymerase. In those cases a stronger poly-A signal randomly localized downstream of L1 is used. This results in retrotransposition of the L1 sequence along with its 3' flanking genomic sequence (Pickeral et al. 2000; Goodier et al. 2000; Szak et al. 2003).

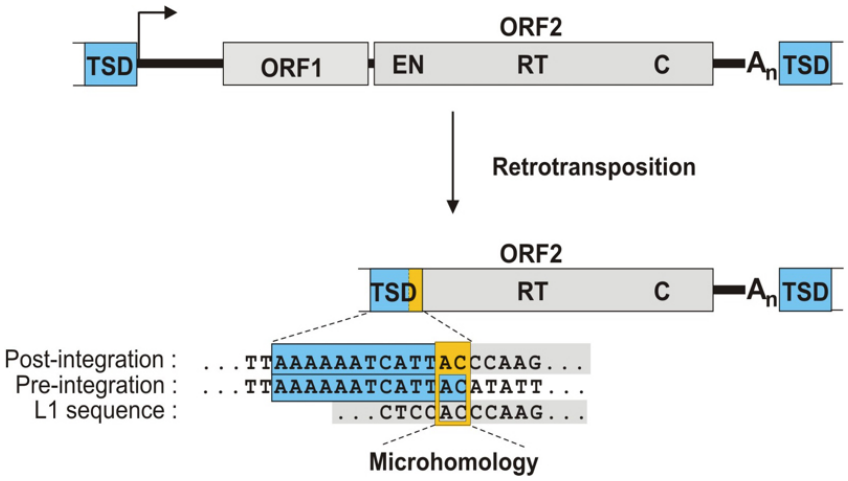


Figure 10 Model for the generation of 5' truncated L1 insertions with microhomologies. Retrotransposition of a human L1 element results frequently in 5' truncated copies of the donor element with overlapping nucleotides at the 5' junctions. These microhomologies (yellow box) between the genomic target site duplication (TSD) and the 5' end of the adjacent L1 sequence make a precise assignment of the 5' boundary of the L1 insertion ambiguous. A 13 bp TSD sequence (blue) observed after *de novo* L1 retrotransposition (Symer et al. 2002) serves as a representative example. Nucleotide sequences of the genomic pre- and post-integration sites as well as the L1 consensus sequence at the junction region are shown. ORF1/2, open reading frame 1/2; EN, endonuclease; RT, reverse transcriptase; C, cysteine-rich motif (from Zingler et al. 2005b).

1.5 IMPACT OF LINE-1 ACTIVITY ON THE HUMAN GENOME

Knowing the structure of LINE-1 retrotransposons, their mode of retrotransposition and abundance in the human genome, it becomes obvious that L1 elements must have had and still exert a great impact on human genome structure and evolution. In fact, this ranges from mere increase of genome size to genetic innovation and gene expression modulation, as well as genomic instability and rearrangements (Cordaux and Batzer 2009; Startek et al. 2015). As a result, L1 is capable of driving human evolution but also of generating a multitude of diseases caused by L1 activity-derived genomic abnormalities.

There are many ways through which L1 can generate genomic instability and one of the most prominent is insertional mutagenesis. Thereby, L1 or one of the trans-mobilized transposable elements inserts into protein-coding or regulatory regions causing genetic disorders such as haemophilia, cystic fibrosis, Apert syndrome, neurofibromatosis, Duchenne muscular dystrophy, β-thalassemia, Hypercholesterolemia and tumorigenic diseases like breast and colon cancers (Deininger and Batzer 1999; Chen et al. 2005; Belancio et al. 2008; Scott et al.

2016). 124 different cases of L1-caused genetic disorders and tumorigenic diseases have been reported (Cordaux and Batzer 2009; Hancks and Kazazian 2012, 2016).

Another way of creating genomic instability is by inflicting double strand breaks as explained in the preceding chapter, and a third one by generating microsatellites which has been studied in particular for *Alu* elements (Arcot et al. 1995; Jurka and Pethiyagoda 1995). Since the A-rich connector of these elements and the 3' poly-A tail contain homopolymeric repeats, they can lead to replication slippage and nucleotide substitutions resulting in microsatellites of variable length and composition.

Due to their high abundance, L1 and other transposable elements can provoke genomic rearrangements caused by non-allelic homologous recombination (NAHR) yielding deletions, duplications and inversions of genetic information. NAHR preferentially involves *Alu* elements and is responsible for numerous genetic diseases (Belancio et al. 2010; Beck et al. 2011).

Furthermore, genomic deletions can also be generated during retrotransposition as shown in Figure 8, but the process of L1 mobilization can as well transduce flanking sequences located either upstream or downstream of the element. 5' transduction occurs when a promoter further upstream of the retrotransposon is used to drive transcription all the way through the 5' flanking sequence and the L1 element. Following TPRT, the additional 5' flanking sequence is integrated into the new genomic locus together with the new LINE-1 copy. In contrast, 3' transduction takes place when the L1-encoded relatively weak transcription termination signal is skipped and an alternative termination signal is used tying the downstream flanking sequence to the L1 mRNA. In that case, TPRT yields a new LINE-1 copy with a copy of the sequence of the genomic origin at its 3' end. In this way, L1 is capable of mediating gene formation by exon shuffling or by simply retrotransposing pseudogenes giving rise to new functional genes eventually.

Apart from actively shaping and changing the genome, L1 can also exert regulatory effects on existing genes. With its sense and antisense promoters it can drive transcription of upstream and downstream located genes, and since it harbors a poly-A signal at its 3' UTR, it may lead to early polyadenylation of host genes, if it is located and cotranscribed in an intron.

Furthermore, L1 sequences contain many potential splice donor and acceptor sites (GT and AG dinucleotides, respectively) that can lead to exonization of retrotransposon sequences by alternative splicing during gene expression (Belancio et al. 2006, 2008).

Most of the investigated LINE-1 retrotransposition events that resulted in an obvious effect for the individual, took place in the germline making it heritable for following generations.

Originally, somatic retrotransposition events were less well studied and were of interest mainly in cases causing cancer. However, a fascinating new research area of somatic retrotransposition arose meanwhile hypothesizing an involvement of LINE-1 in the role of brain development by contributing to somatic mosaicism in neuronal precursor cells (NPCs) (Muotri et al. 2005). L1 was found to be highly expressed in NPCs of rats. Furthermore, an engineered human L1 element with an EGFP reporter was able to retrotranspose in adult rat NPCs *in vitro* influencing neuronal gene expression, and in the brains of transgenic mice *in vivo*.

1.6 LINE-1 TARGET SITE SPECIFICITY

To date, the mechanism of second strand cleavage and 5' end attachment during retrotransposition as well as the mode of target site recognition by L1 EN is largely unknown. *In vivo*, most of the genome is packed in chromatin, while a small portion is undergoing biochemical transactions such as transcription, replication or repair, which may alter the accessibility of the DNA for the L1 transposition machinery. Therefore, the effect of substrate chromatinization on the nicking activity of L1 EN was examined (Cost et al. 2001). The results pointed out that nucleosomal wrapping of DNA renders it a less-efficiently-nicked substrate. However, some phosphodiesters at specific positions in the nucleosome are nicked at an increased rate (Cost et al. 2001). While the choice of integration sites on a chromosome may be determined by the accessibility of DNA, the EN domain is the primary determinant of the consensus sequence specificity of L1 integration (Cost and Boeke 1998).

L1 elements accumulate in AT-rich regions of the genome (Lander et al. 2001) and generally transpose into the consensus sequence 5'-TT ↑ AAAA-3' in which the arrow indicates the nicking site on the complementary bottom strand (Jurka 1997; Szak et al. 2002). L1 was the first element for which a direct correlation between the insertion specificity of an APE-type retrotransposon and the nicking specificity of its EN was observed (Feng et al. 1996). *In vitro* assays showed that the specificity of purified L1 EN for the 5'-TT/AAAA-3' consensus sequence (Feng et al. 1996; Cost and Boeke 1998; Cost et al. 2001) resembles the sequence at the sites of *de novo* L1 insertions *in vivo* (Symer et al. 2002; Gilbert et al. 2002). This observation was also confirmed by *in silico* analysis of the sites of pre-existing L1 and *Alu* insertions in the human genome (Jurka 1997; Szak et al. 2002).

The TPRT model implies that the EN domain is the major determinant of target site selection, as nicking site and site of integration are identical. However, when this model was developed

initially (Luan et al. 1993), APE-type EN domains had not yet been identified in RT-bearing repetitive elements. Then, in 1995 sequence homology between the N-terminal part of ORF2p of the retrotransposon L1Tc from *Trypanosoma cruzi* and the apurinic/apyrimidinic (AP) class II endonuclease family was discovered (Martin et al. 1995).

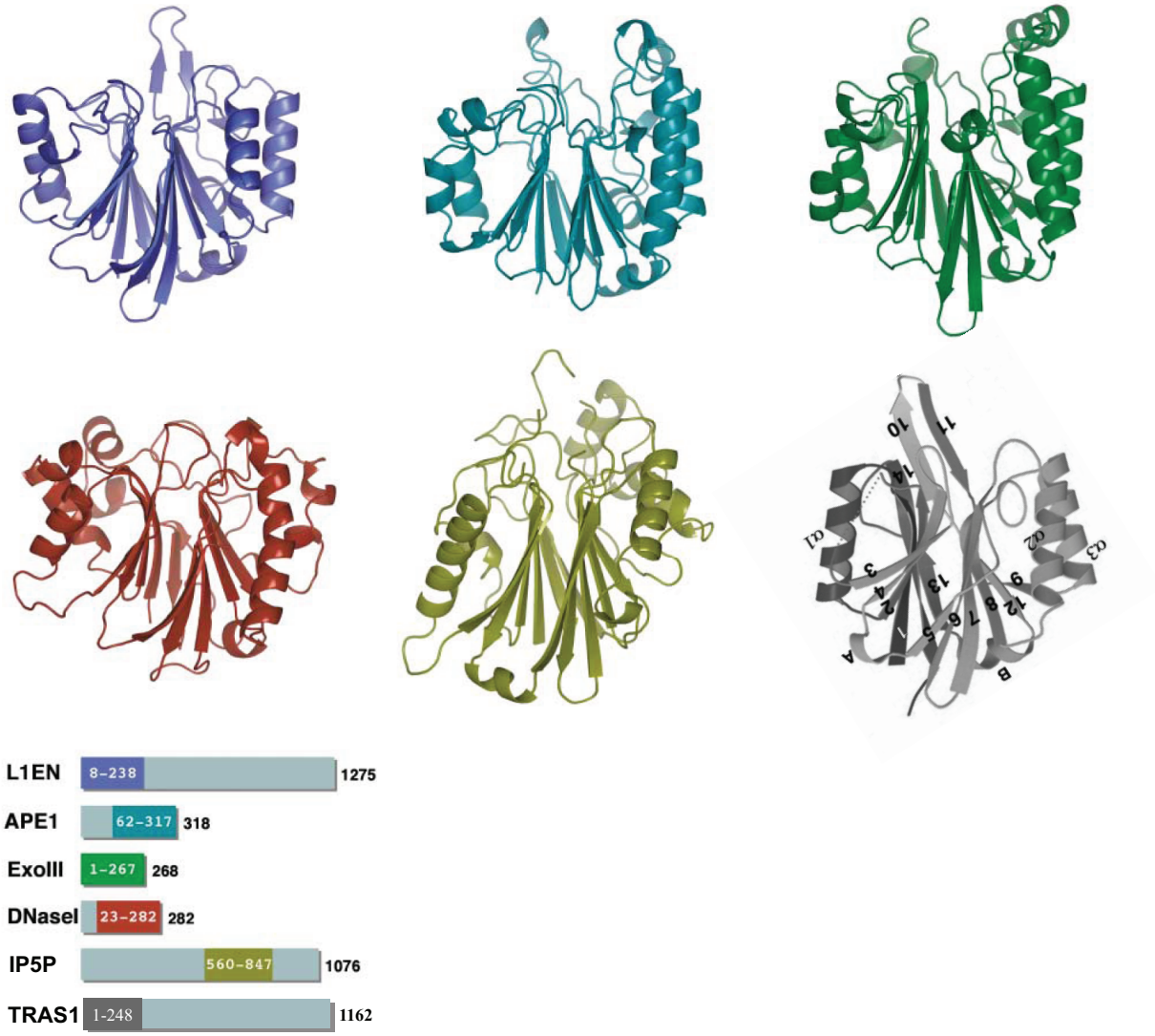


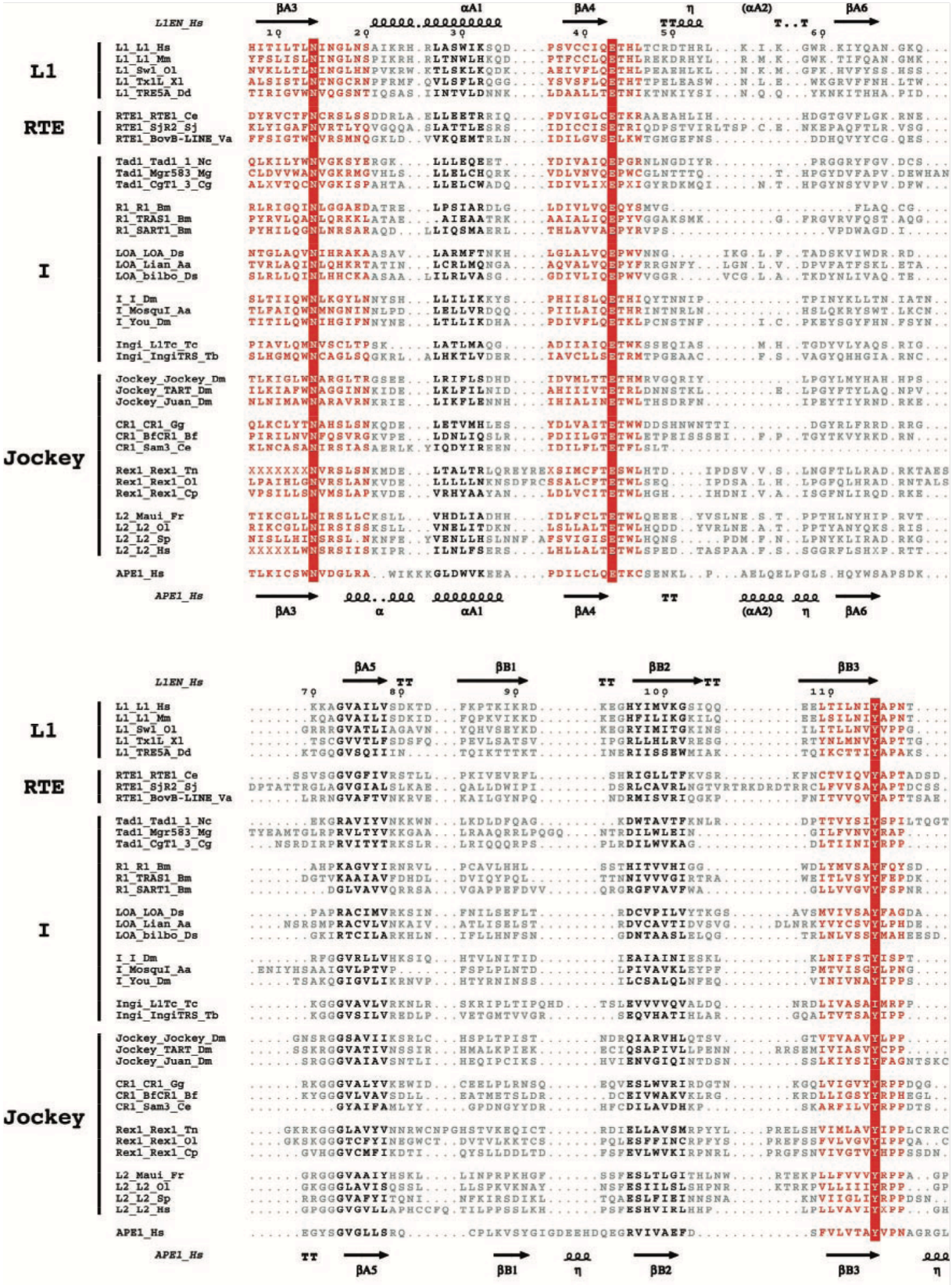
Figure 11 Crystal structures of six members of the family of AP-like ENs, APE1, ExoIII, DNaseI, IP5P, L1 EN and TRAS1 EN. The six enzymes are depicted as ribbon diagrams in the same relative orientation with the DNA binding surface on top. A common central β-sandwich is surrounded by individual α-helices and surface loops. The bars indicate the lengths of the ORFs coding for the respective enzymes and the relative position of the EN domains within (adapted from Weichenrieder et al. 2004 and Maita et al. 2004).

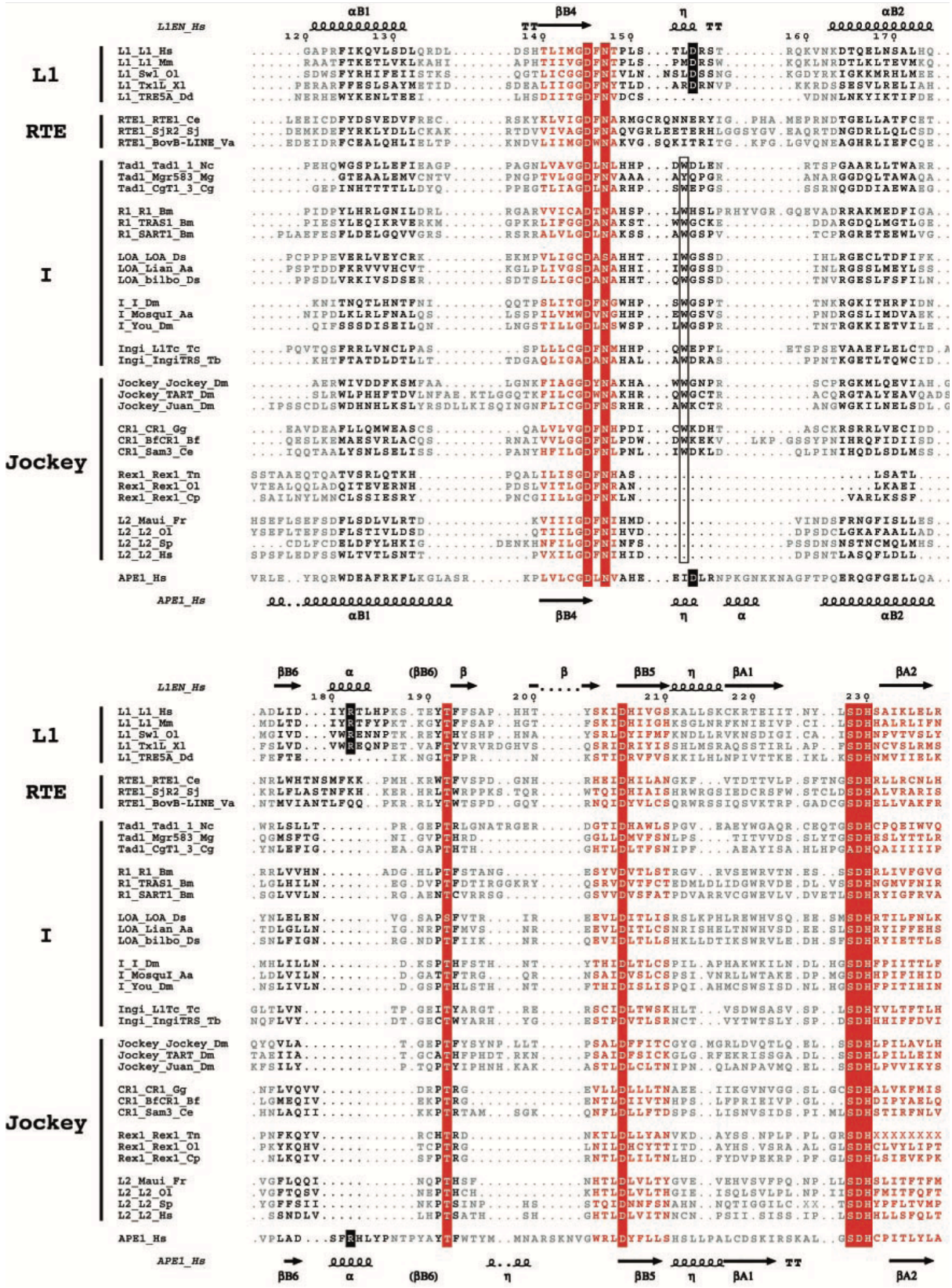
AP class II endonucleases belong to a family of highly conserved, multifunctional DNA repair enzymes found in bacteria, plants, insects, and mammals (Barzilay and Hickson 1995). They are versatile proteins which possess not only endonuclease activity, but also 3' phosphatase, 3' phosphodiesterase, RNase H and 3'→5' exonuclease activities (Demple and

Harrison 1994; Barzilay and Hickson 1995; Evans et al. 2000). They are involved in the predominant pathway for the repair of oxidative DNA damage and the resulting AP sites *in vivo*. However, in non-LTR retrotransposons the endonucleolytic cleavage activity seems to be the crucial function of the APE-like domain of ORF2 proteins (Feng et al. 1996; Feng et al. 1998; Christensen et al. 2000; Takahashi and Fujiwara 2002).

So far, six members of the family of AP-like ENs have been structurally characterized: human LINE-1 EN (L1 EN) (Weichenrieder et al. 2004), human AP endonuclease 1 (APE1, HAP1) (Gorman et al. 1997; Mol et al. 2000), *E. coli* exonuclease III (ExoIII) (Saporito et al. 1988; Mol et al. 1995), bovine pancreatic deoxyribonuclease I (DNaseI) (Lahm and Suck 1991), yeast inositol polyphosphate 5-phosphatase (IP5P) (Tsujishita et al. 2001) and *Bombyx mori* TRAS1 EN (Maita et al. 2004). Comparison of their structures reveals a common tertiary structure: the core consists of two parallel β-sheets surrounded by several α-helical structures. These structural elements are connected by flexible loops, especially on the DNA binding surface (Figure 11). It was demonstrated that the structures of L1 EN and APE1 are largely superimposable and that the active site residues and the supporting structural elements that place them into their respective positions are highly conserved (Figure 12; Weichenrieder et al. 2004). This suggests that the DNA cleavage mechanism that has been proposed for APE1 (Mol et al. 2000) applies for human L1 EN as well. From this conclusion a model was developed in which a prominent βB6-βB5 hairpin loop inserts into the DNA minor groove and is of particular importance for recognizing the DNA target (Figure 13). L1 EN was proposed to recognize the special geometry of the A tract upstream (5') of the scissile bond, while access to the DNA minor groove is supposed to be important for the hydrolysis of the phosphodiester bond at the TpA step.

Figure 12 Structure-assisted alignment of representative EN domains of APE-type non-LTR retrotransposons (on the following two pages). Two to five members of each clade were aligned using the crystal structure of human L1 EN as a guide for structure-based adjustments. The different clades were placed in the four groups L1, RTE, I and Jockey defined by Malik and Eickbush 2001. Amino acids are classified as to whether they can always be aligned automatically (red), with manual adjustments (black) or whether the alignment is not possible for all sequences (gray). Residues that are conserved among all phosphohydrolases are on a red background. Residues forming the salt-bridge that is restricted to AP DNA repair ENs and ENs encoded by elements of the L1 clade are on a black background. The conserved tryptophan that is part of a structural element restricted to the I and Jockey groups of retrotransposons is boxed. α, alpha-helix; β, beta-strand; η, 3-10 helix; TT, beta-turn (from Zingler et al. 2005a).


L1
RTE
I
Jockey
L1EN_Hs
APE1_Hs
βA3
αA1
βA4
η
(αA2)
βA6
βA5
βB1
βB2
βB3
α


L1
RTE
I
Jockey
L1EN_Hs
APE1_Hs
αB1
βB4
η
αB2
βB6
α
(βB6)
β
βB5
βA1
βA2

It is hypothesized that L1 EN accomodates a flipped, extrahelical adenine downstream of the scissile bond, as I204 is small enough for a ribose to be placed and because of the rather small serine in position 202 (S202) supplying even space for a purine base. Furthermore, S202 which is only conserved in mammalian-type L1 elements, and R155 are likely to form hydrogen bonds with the extrahelical adenine to stabilize the protein-DNA interaction (Weichenrieder et al. 2004; Zingler et al. 2005a). T192 plays an important structural role at the base of the βB6-βB5 loop as its side chain oxygen receives weak hydrogen bonds from the main chain nitrogens of I204 and D205 and thus anchors the bottom of the loop with respect to the active site. All important amino acid residues are depicted in Figure 14 in the idealized topology diagram for better understanding.

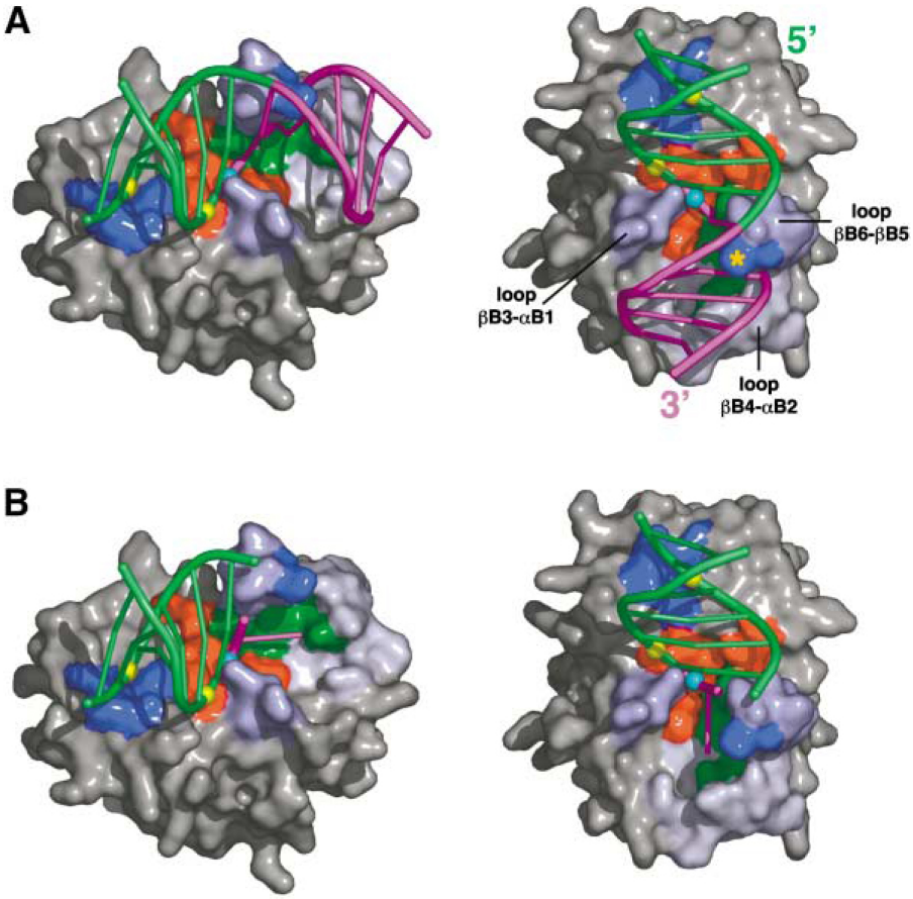


Figure 13 Model of DNA recognition by L1 EN. (**A**) Surface representation of L1 EN (colors as in the following **Figure 14**) with a docked NMR model of substrate DNA (Stefl et al. 2004) represented as ribbons. The upstream (5') and downstream (3') duplexes of the cleavage site are lime-green and magenta, respectively. Sulfate ions on the surface of L1 EN used to position backbone phosphates of the DNA are yellow and the scissile phosphate in the active site is cyan. Loop βB6-βB5 with H198 (asterisk) on its tip inserts into the wide minor groove at the TpA step. This likely bends or unwinds downstream DNA, promoting the adenine to flip. (**B**) Model including only upstream DNA and the flipped adenine downstream of the scissile bond (adapted from Weichenrieder et al. 2004).

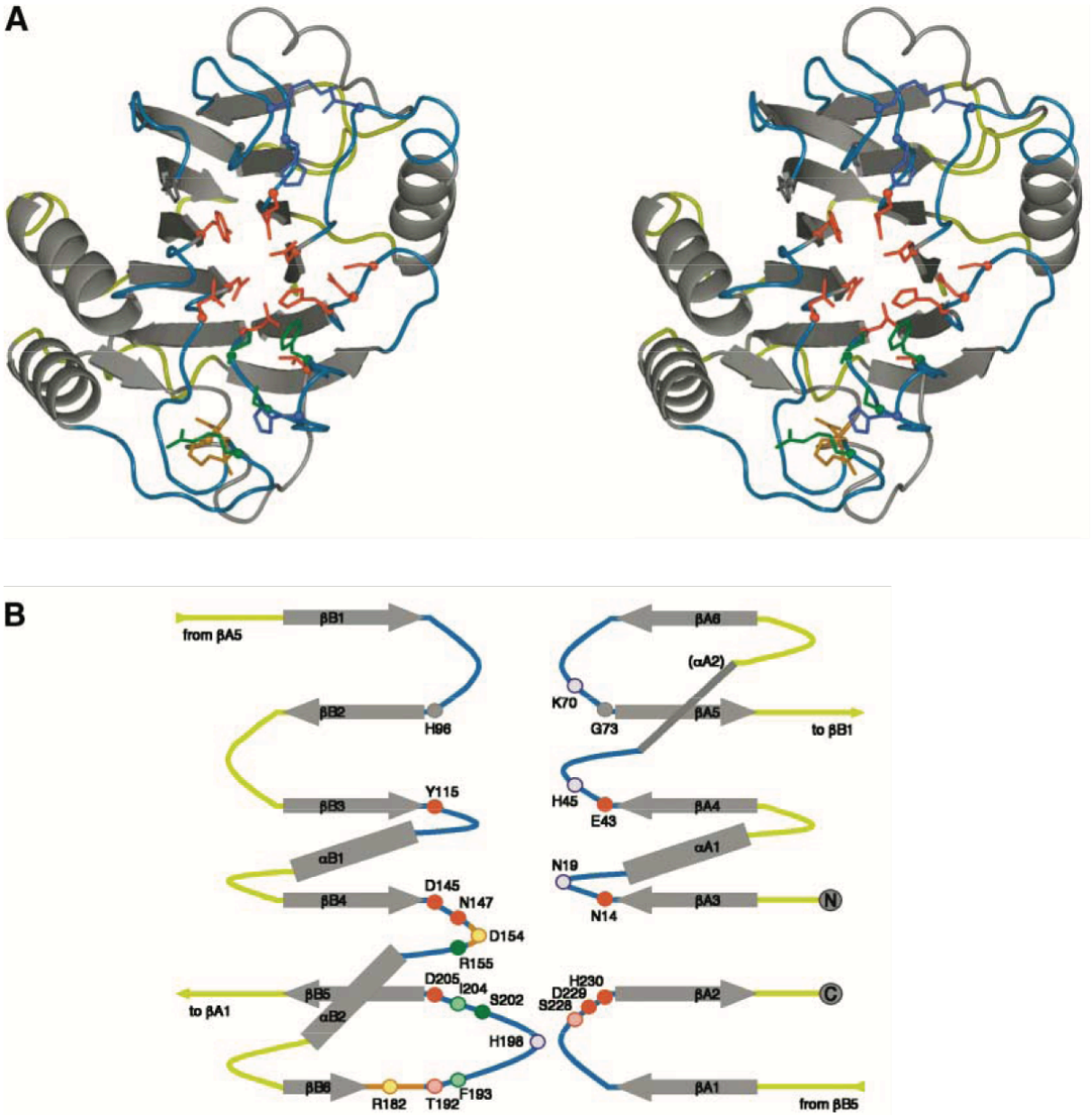


Figure 14 Structural details and topology of L1 EN. (**A**) Ribbon diagram of L1 EN with the loops on the DNA binding side in cyan and the loops on the opposite side of the molecule in yellow (stereo, top view). Selected side chains are drawn as balls-and-sticks and colored as in (B). (**B**) Idealized topology diagram of L1 EN. The diagram emphasizes the pseudo 2-fold symmetry relating the two halves of the molecule. Structural elements are labeled according to the respective half and consecutively in space, not sequence. The prominent βB6-βB5 hairpin loop is enlarged and connections to the N and C termini are indicated. Selected residues are drawn as circles at their approximate positions and color-coded. Red, residues conserved among all phosphohydrolases that are catalytically (filled) or structurally (half-filled) important; green, residues proposed to recognize the extrahelical nucleotide via the ribose (half-filled) and the base (filled); blue, putative peripheral DNA binding residues; orange, salt-bridge restricted to AP DNA repair endonucleases and mammalian-type L1 endonucleases (half-filled) (from Weichenrieder et al. 2004).

1.7 POTENTIAL OF THE HUMAN L1 RETROTRANSPOSON AS VECTOR FOR GENE DELIVERY

Functional retrotransposons are natural vehicles shuffling genetic information from one genomic locus to another. They are capable of modifying the genome similar to retroviruses. And just like retroviruses they could be harnessed as gene delivery vectors. So far, L1 has already demonstrated its potential as tool for insertional mutagenesis in mice (An et al. 2006; O'Donnell et al. 2013) and thus, employing it as gene delivery vector in transgenic animal generation or in gene therapy has become more and more a reasonable consequence.

In transgenic animal generation by pronucleus injection, a gene of interest is randomly introduced into the genome of an animal, e.g. a mouse. Thereby crucial genes can be disrupted, oncogenes can be activated or the transgene can be silenced after integrating into a transcriptional inactive locus. These side effects can only be ruled out by gene targeting via homologous recombination in embryonic stem cells or by characterizing several founder lines of transgenic animals produced by random integration following pronucleus injection. Unfortunately, both methods are very time and cost consuming. Therefore, being able to target a predefined genomic locus quickly, reliably and easily would be a great improvement.

Another field being in need of a reliable vector for gene delivery is gene therapy. It is defined as a medical intervention that changes the genetic material of a human being. During this procedure, DNA carrying a therapeutic gene is introduced into the target cells of a patient. To achieve a long-term effect, stable integration of the transgene into the host cell genome is desirable. In human gene therapy, targeting a predefined locus is of even greater importance than in the generation of transgenic animals, since integration into the wrong locus can cause severe side effects as cancer.

A popular approach for gene therapy is making use of viral vectors. Depending on the virus type, the DNA is either transiently transduced and eliminated rapidly from proliferating cells (e.g. adenoviral vectors) (Volpers and Kochanek 2004), remains episomally and replicates in synchrony with the host cell (e.g. EBV-based vectors) (Delecluse and Hammerschmidt 2000) or is integrated into the host genome (e.g. retroviral vectors) (Coffin 1996). However, since the death of a test person due to an anaphylactic shock in a clinical trial with adenovirus (Raper et al. 2003) and the incidence of several cases of leukemia probably caused by insertional mutagenesis of retroviral vectors (Hacein-Bey-Abina et al. 2003; Hacein-Bey-Abina et al. 2008), viral vectors lost much of their attractiveness.

Another approach to gene therapy was undertaken with a DNA transposon named Sleeping Beauty (SB) (Ivics and Izsvak 2010). Advantages of transposon-based gene delivery over viral vectors include:

- easy maintenance and propagation of transposon vectors as plasmids
- reduced immunogenicity (Yant et al. 2000)
- no strict limitation of the size of expression cassettes (Zayed et al. 2004)
- improved safety and toxicity profiles due to the possibility of supplying only transposase mRNA along with the transposable sequence (Wilber et al. 2006) and shielding the transcription units of the cargo with insulator sequences to prevent trans-activation of neighboring genes (Walisko et al. 2008).

Currently, the feasibility and effectiveness of using transposons in gene therapy are evaluated in clinical trials with SB (Williams 2008; Hackett et al. 2010; Singh et al. 2013; Singh et al. 2015). However, the risk of insertional mutagenesis still remains, although in the meanwhile promising attempts of targeting SB via engineered zinc finger constructs have been undertaken (Voigt et al. 2012).

Human LINE-1 retrotransposons represent an attractive alternative to viral and transposon-based gene delivery systems for several reasons:

- In patients, the danger of an immune response to the gene delivery vehicle and the modified cells is not expected, since the retroelement L1 is of human origin.
- The risk of uncontrolled, second round retrotransposition is very low, because 95% of all L1 retrotransposition events are 5' truncated resulting in L1 copies that are retrotransposition-incompetent.
- Inserted genes into the 3' UTR of L1 are readily integrated into the genomic target locus along with the retroelement during retrotransposition.
- Viruses change their host organism frequently and only have to ensure the host's survival for the virus' replicative cycle. In contrast, several non-LTR retrotransposons have developed sophisticated strategies to target specific innocuous genomic sites as well as predefined sequences in order to ensure their long-term survival in the host's genome and the survival of the host. These strategies could be applied in gene therapy to avoid dangerous side effects caused by insertional mutagenesis.
- Furthermore, there are known retrotransposons integrating into the genome in a highly site-specific manner, e.g. Tx1L from *Xenopus laevis* and R1Bm from *Bombyx mori*. Transferring a comparable site-specificity to LINE-1 would make this element a perfect vector for many gene delivery-based applications in mammals.

However, retrotransposons cannot infect cells and therefore have to be introduced by other means. Since direct physical methods like lipofection or electroporation within a patient are not feasible for obvious reasons, gene delivery could be achieved by a combination of the integrating retrotransposon vector with a viral shuttle. As proof-of-principle, stable integration and expression of transgenes delivered by an L1-adenovirus hybrid vector was previously demonstrated (Soifer et al. 2001; Soifer and Kasahara 2004; Kubo et al. 2006). Since this system involves so-called gutless adenovirus-vectors devoid of all viral genes that could be toxic or immunogenic, it provides high transduction efficiency combined with low immunogenicity and the advantages of retrotransposons mentioned above. Alternatively, isolated cells from patients could be transfected with appropriate L1 constructs. Integration of the therapeutic genes could then be carried out *in vitro*, treated cells could be selected and transferred back to the patient. In that way, a very safe and controllable system for gene therapy could be developed.

Until today, another drawback for L1 as gene delivery vector is its rather random integration into the host's genome. Therefore, I set out to address this issue and unravel basic mechanisms underlying L1 retrotransposition that could be of great importance to better understand, optimize and control retrotransposition.

1.8 AIMS OF THIS WORK

In this study, I set out to elucidate the role of DSB repair proteins in L1 retrotransposition and to identify DSB repair pathways involved in L1 mobilization. Apart from that, I intended to manipulate the L1 target site recognition to make the L1 retrotransposon a more specifically integrating transposable element that could be used as gene delivery tool. For those purposes, we applied an L1 retrotransposition reporter system that allows to monitor the fate of tagged L1 *de novo* retrotransposition events. Such a system was provided previously (Moran et al. 1996), and all experiments described here are based on this cell culture-dependent retrotransposition assay that permits determination of retrotransposition frequency and characterization of structure and sequence of integrated L1 copies (Figure 18). However, immediately after transfection, retrotransposition can already take place and cells containing an early retrotransposition event are able to proliferate and generate numerous G418-resistant colonies, while late integration events yield fewer colonies. This bias for early retrotransposition events emerges to an issue of particular interest when planning co-transfection experiments of L1 retrotransposition reporter constructs with other expression

plasmids supposed to be tested for their effect on retrotransposition. Therefore, the development of a temporally controllable L1 retrotransposition reporter assay became necessary to ensure downregulation of specific DSB repair factors at the time point of L1 retrotransposition start.

For that reason, one aim of this study was to generate the first tetracyclin (Tet)-controlled LINE-1 retrotransposition reporter for cell culture-based L1 retrotransposition assays.

Moreover, due to the recent finding that endogenous L1 mobilization takes place during brain development and contributes to somatic mosaicism (Muotri et al. 2005), I also set out to develop a bidirectional, Tet-inducible L1 reporter element in order to introduce it into mice. Those transgenic animals shall be used to characterize LINE-1 expression and retrotransposition pattern during brain development and for further behavioral analysis to elucidate the role of L1 in the brain.

The third aim of this PhD thesis of investigating the role of single DSB repair proteins in L1 retrotransposition and identifying DSB repair pathways involved in this process was based on the finding that L1 integrations that have been analyzed from L1 retrotransposition assays and from the human genome by database mining reveal microhomologies of 1 to 10 bp at their 5' junctions (Zingler et al. 2005b). This discovery suggested DSB repair by NHEJ being involved in second strand synthesis during L1 integration, as B-NHEJ is also characterized by the generation of microhomologies. In order to test the hypothesis that factors of the NHEJ pathway are involved in L1 retrotransposition, I aimed at knocking down DSB repair factors (ATM, ATR, DNA-PKcs, p53, Ku70, PARP and Rad51) in HeLa cells and monitoring the consequences on L1 retrotransposition. Applying the L1-retrotransposition reporter assay mentioned above (Moran et al. 1996), *de novo* retrotransposition events were launched in order to test LINE-1 elements for their retrotransposition activity and characterize L1 integration sites in the context of altered DSB repair conditions.

Since L1 was demonstrated to exhibit several features that render it a potential tool for gene delivery, the fourth aim of this thesis was to analyze and manipulate DNA target site recognition in order to overcome the disadvantage of rather random integration. Therefore, I performed crystal structure-guided mutational analysis exchanging single amino acids within the endonuclease domain of L1 to identify residues influencing target site recognition.

2 MATERIALS AND METHODS

The methods used in this work were carried out as described in standard laboratory manuals (Sambrook and Russell 2001; Ausubel 2002) or according to manufacturer's instructions. Therefore, only modified methods are described in detail in this chapter.

2.1 CHEMICALS

All chemicals were purchased at analytical grade from the following companies unless stated otherwise: Biochrom AG (Berlin), Invitrogen AG (Karlsruhe), Merck KGaA (Darmstadt), Roche GmbH (Mannheim), Serva (Heidelberg) and Sigma-Aldrich (München).

2.2 BUFFERS AND SOLUTIONS

Sample buffer for agarose gels
50% (v/v) Glycerol
0.25% (w/v) Xylenecyanol FF
0.25% (w/v) Bromophenol blue
in TE buffer

PBS (Phosphate Buffered Saline)
137.0 mM NaCl
2.7 mM KCl
6.5 mM Na_2HPO_4
1.5 mM K_2HPO_4
pH 7.4

TE buffer
10 mM Tris-HCl
1 mM Na_2EDTA
pH 8.0

LB broth (Luria-Bertani medium)
10 g Tryptone
5 g Yeast extract
5 g NaCl
ad 1 l dH_2O
autoclave

LB agar
10 g NaCl
5 g Tryptone
5 g Yeast extract
20 g Agar
ad 1 l H_20
autoclave

Proteinase K (10 mg/ml)
1 g Proteinase K (Boehringer, Mannheim)
300 µl NaCl (5 M)
1 ml TrisHCl (1 M, pH 7.6)
ad 100 ml H_2O
aliquot and store at -20°C, thaw at 55°C

Tens(-)
30 ml TrisHCl (1 M, pH 8.0)
120 ml EDTA (0.5 M)
12 ml NaCl (5 M)
60 ml SDS (10%)
ad 500 ml H_2O

Day 1 buffer
1% BSA
0.3% Triton X-100
in 1x PBS

Day 2 buffer
0.3% BSA
0.1% Triton X-100
in 1x PBS

X-Gal solution
12 mg/ml X-Gal in dimethylformamide
5 mM $K_4Fe(CN)_6$
5 mM $K_3Fe(CN)_6$
2 mM MgCl2
in 1x PBS

Cell lysis buffer
25 mM HEPES pH 7.4
150 mM NaCl
1% Triton X-100
add 1 tablet cOmplete protease inhibitor (Roche, Mannheim) per 25 ml

6x SDS loading buffer
3.75 ml TrisHCl pH 6.8
3 ml Glycerol
1.75 ml H_2O
6% SDS
0.5% Bromophenol blue

5x Stacking gel buffer
0.31 M Trizma HCl in H_2O
adjust pH to 6.7 with HCl

5x Resolving gel buffer
1.87 M Trizma base in H_2O
adjust pH to 8.9 with HCl

10x Running buffer
1% SDS
1.9 M Glycine
250 mM Trizma base
in H_2O

10x MOPS buffer
200 mM MOPS
50 mM Sodium acetate
10 mM EDTA

Transfer buffer
1% SDS
15% Methanol
25 mM Trizma base
192 mM Glycine
in H_2O

Stacking acrylamide gel
4% Acrylamide/bis (Bio-Rad, München)
0.05% APS
1‰ TEMED
in 1x Stacking gel buffer

Resolving acrylamide gel
8-12% Acrylamide/bis (Bio-Rad, München)
0.05% APS
0.5‰ TEMED
in 1x Resolving gel buffer

Blocking solution
0.1% TWEEN 20 (Sigma-Aldrich, München)
10% Skim milk powder (Sigma-Aldrich, München)
in 1x PBS pH 7.4

Special blocking solution
0.1% TWEEN 20 (Sigma-Aldrich, München)
5% FCS
2% BSA
in 1x PBS pH 7.4

PBS-T
0.1% TWEEN 20 (Sigma-Aldrich, München)
in 1x PBS pH 7.4

Stripping buffer
100 mM 2-Mercaptoethanol
2% SDS
62.5 mM Tris-HCl pH 6.7

2.3 ENZYMES

DNA modifying enzymes were obtained from Roche GmbH (Mannheim), Invitrogen AG (Karlsruhe), MBI Fermentas (St. Leon-Rot), New England Biolabs (Schwalbach), Stratagene (Heidelberg) and Qiagen (Hilden). Reactions were carried out according to manufacturer's instructions.

2.4 BACTERIAL CULTURE FOR CLONING AND PLASMID PREPARATION

For standard cloning, *E. coli* strains DH5α and XL1-Blue were used. Ligation reactions yielding large plasmids (>18 kb) were transformed into DH10B cells. Plasmids requiring digestion with methylation sensitive restriction enzymes were transformed into and reisolated from dcm- and dam-negative strain GM2163. For plasmid rescue experiments, ElectroMax DH10B cells (Invitrogen AG, Karlsruhe) were electroporated. Bacterial strains used in these studies are listed in Table 1.

Strain	Characteristics	Reference
DH5α	F' φ80*lac*ZΔM15 Δ(*lac*ZYA-*arg*F)U169 *deo*R *rec*A1 *end*A1 *hsd*R17(rk^-,mk^+) *pho*A *sup*E44 λ^- *thi*-1 *gyr*A96 *rel*A1	Invitrogen AG (Karlsruhe)
XL1-Blue	*rec*A1 *end*A1 *gyr*A96 *thi*-1 *hsd*R17 *sup*E44 *rel*A1 *lac* {F' *pro*AB *lac*I^qZΔM15 Tn10(Tet^R)}	Stratagene (Heidelberg)
DH10B	F^- *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) φ80d*lac*ZΔM15 Δ*lac*X74 *deo*R *rec*A1 *end*A1 *ara*D139 Δ(*ara*, *leu*)7697 *gal*U *gal*K λ^- *rps*L *nup*G	Invitrogen AG (Karlsruhe)
GM2163	F^- *ara*-14 *leu*B6 *fhu*A31 *lac*Y1 *tsx*78 *gln*V44 *gal*K2 *gal*T22 *mcr*A *dcm*-6 *his*G4 *rfb*D1 *rps*L136 *dam*13::Tn9(Cam^R) *xyl*A5 *mtl*-1 *thi*-1 *mcr*B1 *hsd*R2	New England Biolabs (Schwalbach)
Top10F'	F^- {*lac*I^q Tn10 (Tet^R)} *mcr*A Δ(*mrr hsd*RMS *mcr*BC) φ80*lac*ZΔM15 Δ*lac*X74 *rec*A1 *ara*D139 Δ(*ara*, *leu*)7697 *gal*U *gal*K *rps*L (Str^R) *end*A1 *nup*G	Invitrogen AG (Karlsruhe)

Table 1 Bacterial strains used in this work.

2.4.1 CULTIVATION AND STORAGE OF *E. COLI*

Bacteria were cultivated in Luria-Bertani (LB)-broth or on LB-agar plates at 37°C. To select for transformed bacteria, broth and plates were supplemented with 50 mg/l ampicillin or 50 mg/l kanamycin. Plated bacteria were viable for approximately one month when stored at 4°C. For long-term storage, glycerol stocks were prepared by mixing 500 µl of an overnight liquid culture with 500 µl of 15% sterile glycerol. Stocks were stored at -80°C.

2.4.2 TRANSFORMATION OF PLASMID DNA INTO *E. COLI*

Chemically competent or electro-competent *E. coli* cells were obtained from Invitrogen AG (MAX Efficiency DH10B, OneShot TOP10F' and ElectroMax DH10B) or prepared following standard CaCl2-protocols. Plasmids and ligation reactions were introduced into bacteria by heat-shock transformation according to the protocol supplied with MAX Efficiency DH10B cells or by electroporation at 1.9 kV, 200 Ω and 25 µF in a Gene Pulser Xcell electroporator (Bio-Rad, München).

2.4.3 PREPARATION OF PLASMID DNA FROM *E. COLI*

Plasmid DNA required for cloning, sequencing or transfection of HeLa cells was isolated with the commercially available "Plasmid DNA Purification" kits (Qiagen, Hilden). Extraction and purification of the DNA was achieved by ion exchange columns (QIAGEN-tip 20, 100 or 500) following the user manual. The resulting DNA pellet was dissolved in double distilled H_2O. After photometrical determination of the yield, the isolated DNA was characterized by restriction and/or sequencing.

2.5 PLASMIDS AND PLASMID CONSTRUCTION

The original retrotransposition reporter plasmids pJM101/L1.3 (Figure 15, top) (Moran et al. 1996) and pJM101/$L1_{RP}$ (Kimberland et al. 1999), the plasmid rescue vector pCEP4/L1.3mneoI400/ColE1 (Figure 15, bottom) (Gilbert et al. 2002) and the retrotransposition reporter with blasticidin resistance pTAM102/L1.3 (Figure 16, middle) (Morrish et al. 2002) were gifts from John Moran (University of Michigan, Ann Arbor).

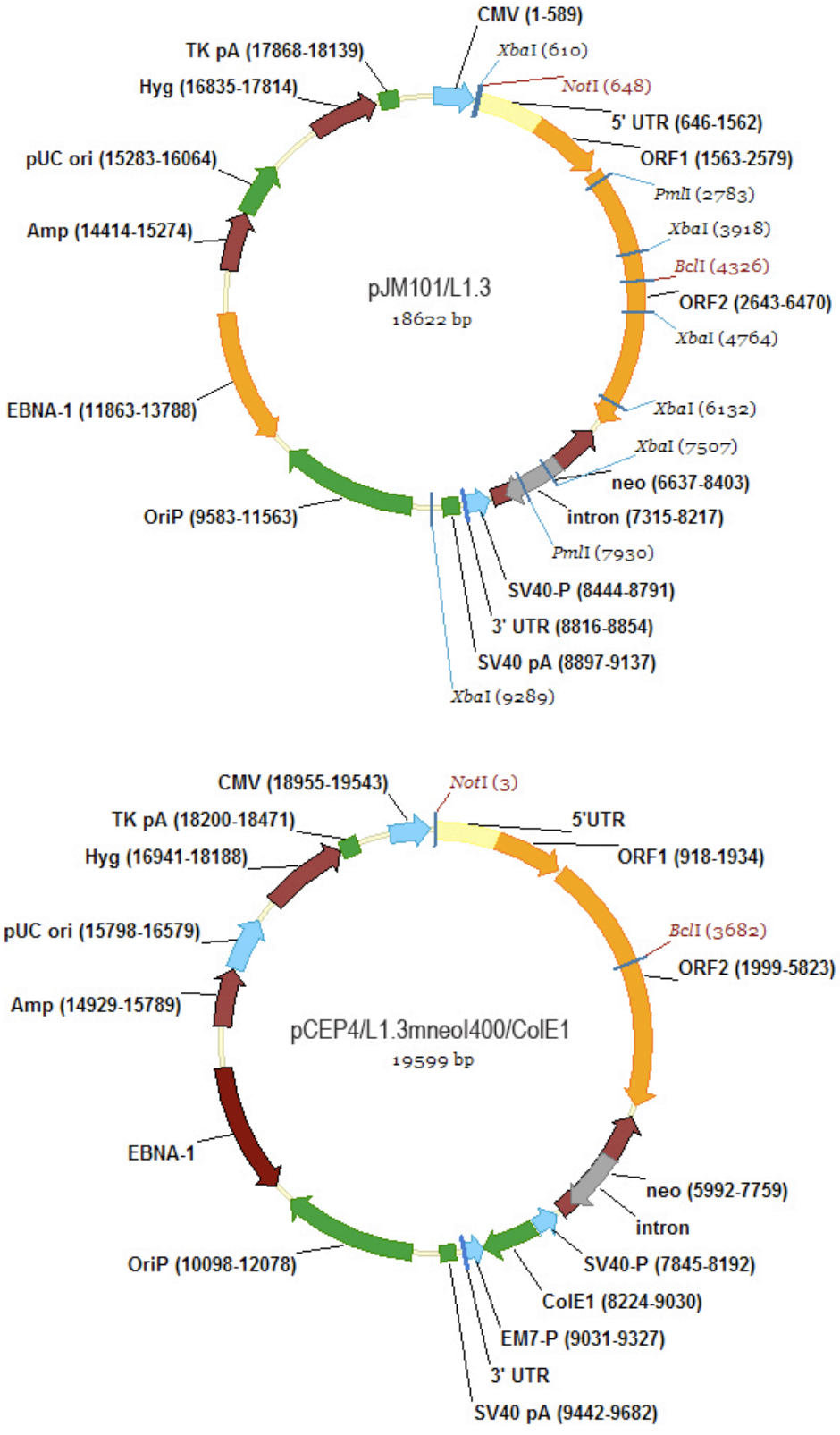


Figure 15 Retrotransposition reporter constructs pJM101/L1.3 (top) (Moran et al. 1996) and pCEP4/L1.3mneoI400/ColE1 (bottom) (Gilbert et al. 2002). Elements of the retrotransposition reporter cassette are color-coded: yellow, untranslated regions (UTRs); orange, open reading frames (ORFs); maroon, antibiotic resistance genes (neo, Amp and Hyg); gray, β-globin intron 2 (intron); light blue, SV40 promoter (SV40-P); green, SV40 polyadenylation signal (SV40 pA).

pJM101/L1$_{RP}$ is a plasmid constructed just as pJM101/L1.3, but containing a highly active L1 element published previously (Kimberland et al. 1999). The negative control retrotransposition reporter pJM101/L1.3H230A carrying a point mutation in its EN domain (H230A) was constructed previously by Nora Zingler (PhD Thesis 2004).
Plasmids containing shRNAs directed against mRNAs of ATM, ATR, DNA-PK (Collis et al. 2003) and p53 (Brummelkamp et al. 2002) were gifts from Lisa Wiesmüller (Universitätsfrauenklinik Ulm), whereas the plasmid carrying an shRNA against Ku70 was kindly provided by Zsuzsanna Izsvák (Max-Delbrück-Centrum für Molekulare Medizin, Berlin) (Figure 27 and Figure 28).
In this study, inhibition of PARP was achieved by overexpression of the DNA-binding domain of PARP (PARP DBD) (Kupper et al. 1990), whereas for interference with Rad51 function a yeast-mouse chimera of the protein (Rad51SM) was applied (Lambert and Lopez 2000) (Figure 40). Both plasmids were kindly provided by Lisa Wiesmüller.

2.5.1 CLONING PROCEDURES

Restriction digestion and ligation of plasmids and DNA fragments was done according to the enzyme supplier's instructions. When PCR products were cloned, it was ensured that there were at least 6 bp between the restriction sites used for cloning and the PCR product's end to guarantee efficient enzymatic cleavage. Vectors were dephosphorylated if necessary and controlled for their religation potential by a mock ligation without insert.

2.5.2 CLONING STRATEGIES

For the design of cloning strategies, oligos, construction of plasmid maps and sequence alignments the computer program Vector NTI (Invitrogen AG, Karlsruhe) was used. Unfortunately, this program does not allow the use of periods in plasmid denotations and other file names. Therefore, in some cases periods were replaced by hyphens in some graphical presentations of plasmids.

2.5.3 CLONING OF PTETO7CMV/L1-3BLAS

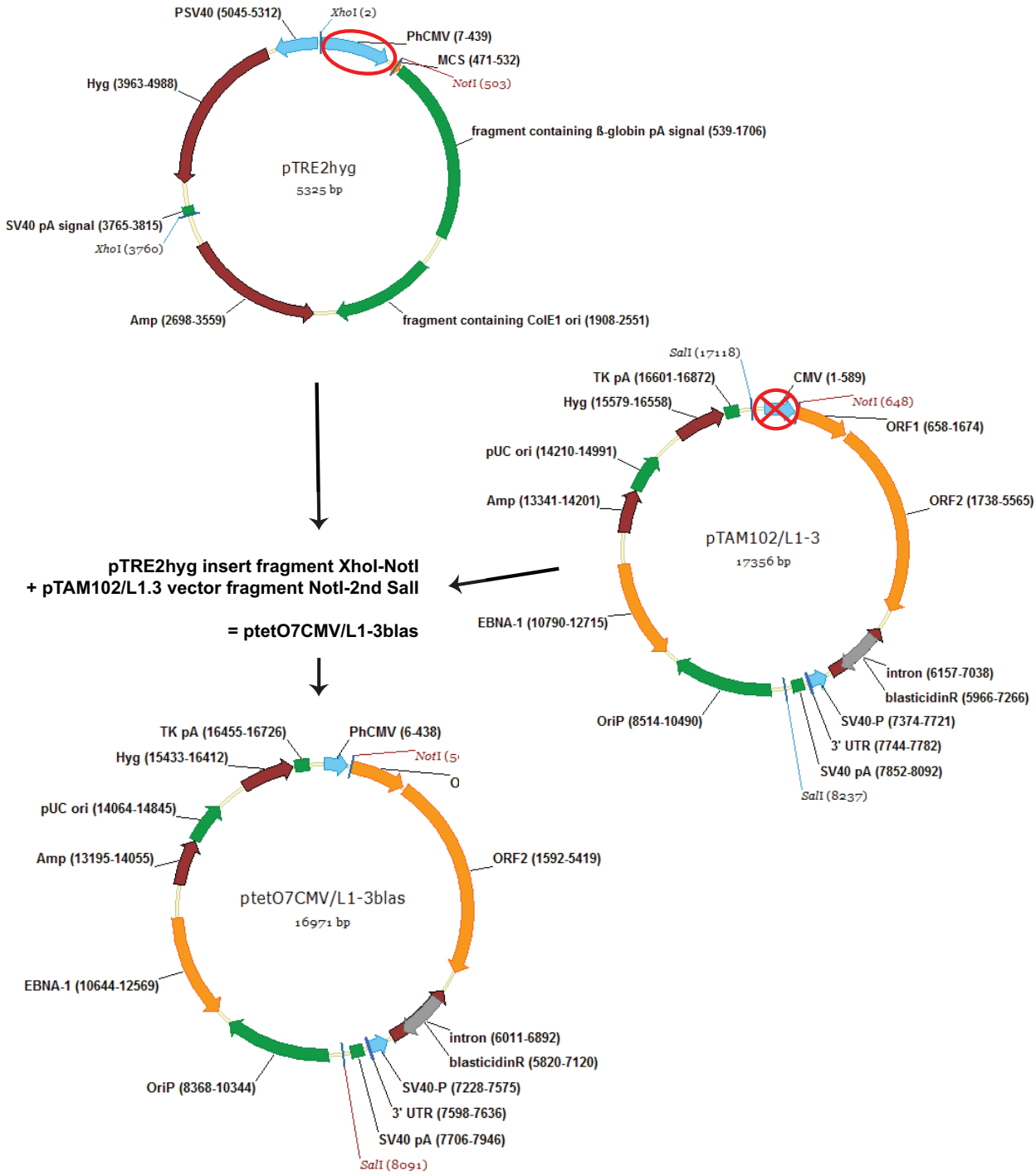


Figure 16 Cloning scheme of ptetO7CMV/L1-3blas. The fragment containing the Tet-dependent CMV promoter was excised from pTRE2hyg using *XhoI* and *NotI* and ligated into the vector fragment from pTAM102/L1.3 from which the regular CMV promoter had been released by incomplete restriction with *NotI* and *SalI*.

For the generation of ptetO7CMV/L1-3blas I used the Tet-dependent CMV promoter fragment (PhCMV) from pTRE2hyg (Clontech Laboratories, Inc., Saint-Germain-en-Laye, France) and inserted it into pTAM102/L1.3 (Morrish et al. 2002) in place of the regular CMV promoter (Figure 16). Excision of PhCMV was performed by restriction with *XhoI* and *NotI* and the removal of the regular CMV promoter by an incomplete digestion of pTAM102/L1.3 with *NotI* and *SalI*. Since *XhoI* and *SalI* produce compatible ends, no further modification of the sticky ends was necessary and the ligation could be conducted after dephosphorylation of the vector fragment. To verify the plasmid sequence of the constructed ptetO7CMV/L1-3blas, the cloning junctions were sequenced (Eurofins MWG GmbH, Ebersberg).

2.5.4 CLONING OF L1 ENDONUCLEASE POINT MUTATIONS

To introduce specific point mutations into the EN domain of L1, site-overlap extension PCR (SOE-PCR) was used (Figure 17). In the first step, two primer pairs were used to amplify two fragments that were at least 30 bp complementary to each other within the region of the mutation. One of the primers of both pairs carried the point mutation to be introduced. In the second step, assisted by the two outer primers, these two PCR products acted as megaprimers amplifying the complete fragment with the new mutation within.

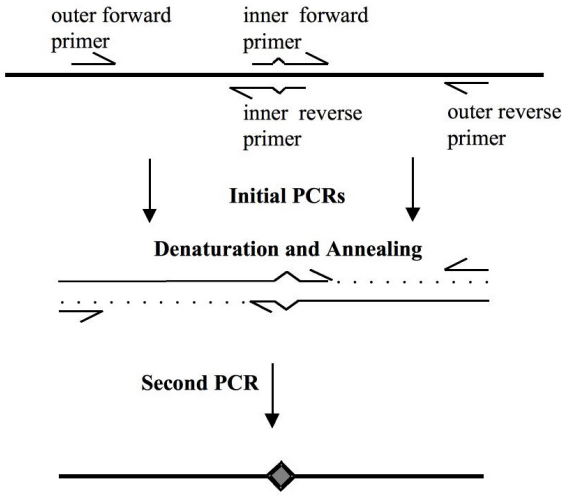


Figure 17 Principle of site-overlap extension PCR (SOE-PCR) to introduce point mutations. Two primer pairs are used in initial PCRs to amplify two fragments overlapping by 30 bp in the region of the mutation to be introduced (triangular structures). These fragments are used as megaprimers that are elongated (indicated by dotted lines) in the second PCR. The outer primers are added to increase the yield after the first overlap extension reactions have created the full-length product. Thin lines, single stranded DNA; bold lines, double stranded DNA (from Nora Zingler, PhD Thesis 2004).

Reaction mix for initial PCR:

5 µl 10x *Pfu* Polymerase buffer (Stratagene)
10 pmol of each primer
20 fmol DNA template
0.2 mM of each dNTP
1 U *Pfu* Polymerase
ad 50 µl H_2O

PCR conditions:

96°C	1 min		
96°C	20 sec	⎫	
45-65°C	15 sec	⎬	25x
72°C	30-120 sec	⎭	
72°C	7 min		
4° C	∞		

Reaction mix for second PCR:

equimolar amounts of each PCR product (about 100 fmol)
10 pmol of each outer primer
0.2 mM of each dNTP
1 U *Pfu* Polymerase
ad 50 µl H_2O

In order to enable easy modification of the L1 EN domain in a smaller plasmid, pNZ1 was used as transition vector containing the 3.7 kb *NotI-BclI* fragment of pJM101/L1.3 (Figure 15, top) in the pBluescript KS+ vector backbone (Nora Zingler, PhD Thesis 2004).

To introduce the EN point mutations, the following inner primers were used (primer sequences are listed in Section 2.10.1): for S202A primers GS334 and GS335, for R155A primers GS336 and GS337, for T192V primers GS338 and GS339 and for I204Y primers GS340 and GS341. In all four cases, GS73 and GS263 were used as outer primers. They include the restriction sites *PmlI* and *XbaI* within the *NotI-BclI* fragment of pJM101/L1.3, which were used to clone the mutated SOE-PCR products back into pNZ1. From these subclones, the EN point mutations were introduced via the restriction sites *NotI* and *BclI* into the L1 reporter pJM101/L1.3, yielding pJM101/L1.3S202A, pJM101/L1.3R155A, pJM101/L1.3T192V and pJM101/L1.3I204Y, and into the plasmid rescue vector pCEP4/L1.3mneoI400/ColE1 (Figure 15, bottom), yielding pColE1-S202A, pColE1-R155A, pColE1-T192V and pColE1-I204Y. All generated EN point mutation plasmids have the same sequence as their parent plasmid, pJM101/L1.3 or pCEP4/L1.3mneoI400/ColE1, respectively, except for the introduced point mutation in the EN domain.

2.6 TISSUE CULTURE

2.6.1 CULTIVATION OF HELA CELLS

HeLa cells (ATCC number CCL-2) were grown in an incubator at 37°C and 5% CO_2 using Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal calf serum (FCS), 2 mM glutamine, 100 U/ml penicillin and 100 µg/ml streptomycin.
HeLa M2 cells (Hampf and Gossen 2007) were grown in an incubator at 37°C and 5% CO_2 using Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal calf serum (FCS), 2 mM glutamine, 100 U/ml penicillin, 100 µg/ml streptomycin and 200 µg/ml G418. For induction of Tet-dependent expression of the L1 retrotransposition reporter ptetO7CMV/L1-3blas (Figure 16, bottom), 200 ng/ml doxycyclin were added to the medium.

2.6.2 DETERMINATION OF CELL NUMBER

Cells were trypsinized and a 20 µl aliquot was mixed with 20 µl of a 0.36% trypan blue solution. After 3 minutes of staining, only dead cells incorporated the blue dye, while living cells remain unstained. The latter were counted in a Neubauer chamber.

2.6.3 LINE-1 RETROTRANSPOSITION REPORTER ASSAY

All experiments described in this work are based on a cell culture-dependent retrotransposition reporter assay developed by John V. Moran (Moran et al. 1996) that permits determination of retrotransposition frequency and characterization of structure and sequence of integrated L1 copies (Figure 18). In this assay, drug resistance is conferred to an L1-transfected cell only after retrotransposition took place. This is achieved by introducing a reporter cassette, *mneoI* (Freeman et al. 1994), into the 3' UTR of L1 and consequently tagging it. *mneoI* consists of a selectable marker gene (*neo*) in antisense orientation preceeded by an SV40 promoter (P') and followed by an SV40 polyadenylation signal (A'). The *neo* gene is disrupted by an intron (IVS 2 of the β-globin gene) inserted in sense orientation. Therefore, L1 mRNA transcripts driven by the CMV promoter (PCMV) are spliced, but contain an antisense copy of the *neo* gene which cannot be translated at this stage. Concurrently, transcripts initiated from P' cannot be spliced and thus do not yield a functional neomycin phosphotransferase. G418 resistant colonies arise only if a transcript initiated by

the CMV promoter is spliced, reverse transcribed and reintegrated into the chromosomal DNA yielding an intact *neo* gene that can be expressed from P' and renders the host cell resistant to G418. The tagged L1 element was subcloned into the pCEP4 expression vector, a plasmid particularly suited for the expression in primate cells as it replicates extrachromosomally and is distributed to the daughter cells via an EBNA-1 (Eppstein Barr Nuclear Antigen 1) (Yates et al. 1985). Furthermore, it contains a hygromycin resistance gene (*hyg*) for the selection of transfected cells and a well characterized cytomegalovirus (CMV) immediate early promoter (Boshart et al. 1985) for eukaryotic expression of genes of interest, in this case L1.

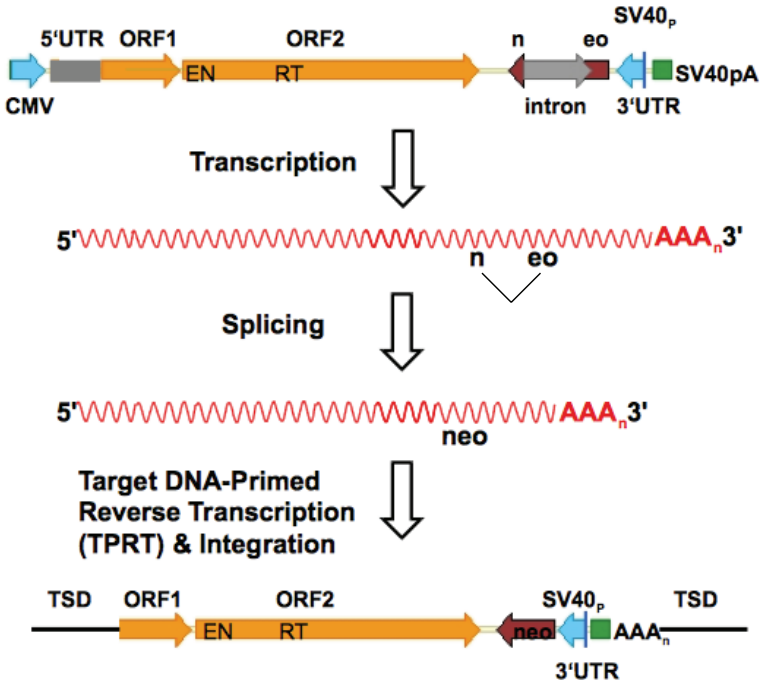


Figure 18 Schematic representation of the LINE-1 retrotransposition reporter assay. G418-resistant colonies arise only when the L1 transcript is spliced and integrated into chromosomal DNA by target primed reverse transcription (TPRT). ORF1 and ORF2 are shown in orange and the *mneoI* cassette in maroon and gray. EN, endonuclease domain; RT, reverse transcriptase domain. CMV, cytomegalovirus immediate early promoter; UTR, untranslated region; TSD, target site duplication; SV40P, SV40 promoter; SV40pA, SV40 polyadenylation signal.

After transfection of the L1 retrotransposition reporter plasmid pJM101/L1.3, the cells can either be selected with hygromycin for the presence of the episomal reporter plasmid or with G418 for retrotransposition events. A selection with hygromycin first followed by G418 enriches the transfected cell population carrying retrotransposition events and increases the number of retrotransposition events per cell (Nora Zingler, PhD Thesis 2004).

Retrotransposition assays were performed as described previously (Wei et al. 2000), except that GeneJuice transfection reagent from Novagen (VWR International GmbH, Darmstadt) was used instead of FuGENE 6. In detail, $2x10^5$ HeLa cells were plated in each well of a 6-well plate. The next day, cells were transfected using 3 µl GeneJuice transfection reagent and 1 µg DNA per well. In the case of co-transfections, 0.5 µg of each plasmid were used. At 72 h post transfection, the medium was replaced with DMEM containing either 400 µg/ml G418, 800 µg/ml hygromycin or 3 µg/ml blasticidin. G418 and hygromycin resistant colonies were detectable 10-14 days later, while the selection of blasticidin resistant colonies was completed after 8-12 days. Cell colonies were fixated and stained with Giemsas Azur-Eosin-Methylenblau (VWR International GmbH, Darmstadt) for visualization. Alternatively, genomic DNA or protein extracts were prepared for subsequent analysis.

2.6.4 QUANTIFICATION OF CELL COLONY NUMBERS AFTER ANTIBIOTIC SELECTION

Cell colonies stained with Giemsas Azur-Eosin-Methylenblau (VWR International GmbH, Darmstadt) were quantified using the automated cell colony counter ColCount (Oxford Optronics, Abingdon, UK) following the manufacturer's instructions. For each retrotransposition assay, colony numbers were expressed in % of colony numbers relative to the respective control well.

2.6.5 LINE-1 PLASMID RESCUE FROM HELA CELLS

To characterize newly integrated L1 elements, a plasmid rescue procedure published in 2002 (Symer et al. 2002; Gilbert et al. 2002) that allows direct cloning of *de novo* L1 integrants together with their flanking genomic DNA in bacteria was applied. In these plasmid rescue vectors, a bacterial origin of replication and a prokaryotic selectable marker were introduced into the 3' UTR of L1 in the retrotransposition reporter construct pJM101/L1.3. John Moran kindly supplied us with the rescue vector pCEP4/L1.3mneoI400/ColE1 (Figure 15). In this construct, the *mneoI* cassette is used as eukaryotic and prokaryotic marker by inserting the bacterial EM7 promoter and a Shine-Dalgarno-sequence upstream of the *neo* start codon. Furthermore, a ColE1 origin of replication was added. These modifications reduce the retrotransposition efficiency approximately 6-fold compared to pJM101/L1.3 (Gilbert et al. 2002), which is probably due to the increased length of the retrotransposed sequence needed

to confer G418 resistance. However, this is more than compensated by the advantage of an elegant system for recovering new retrotransposition events. To do so, genomic DNA isolated from G418 resistant cells is digested with *HindIII* (a restriction enzyme that does not cleave within the *mneoI*/ColE1 cassette), religated under strongly diluted conditions to form intramolecular circles and subsequently transformed into *E. coli* (Figure 19). To isolate *de novo* integrants, restriction by the enzyme *HindIII* was used which cleaves the L1 element at position 3667. Therefore, only L1 copies that are 5' truncated downstream of the *HindIII* cleavage site can be recovered with 5' and 3' flanking sequences at once. Full-length integrants or integrants extending beyond the *HindIII* site result in rescue plasmids carrying only the 3' part of the L1 cassette and its 3' flanking genomic sequence.

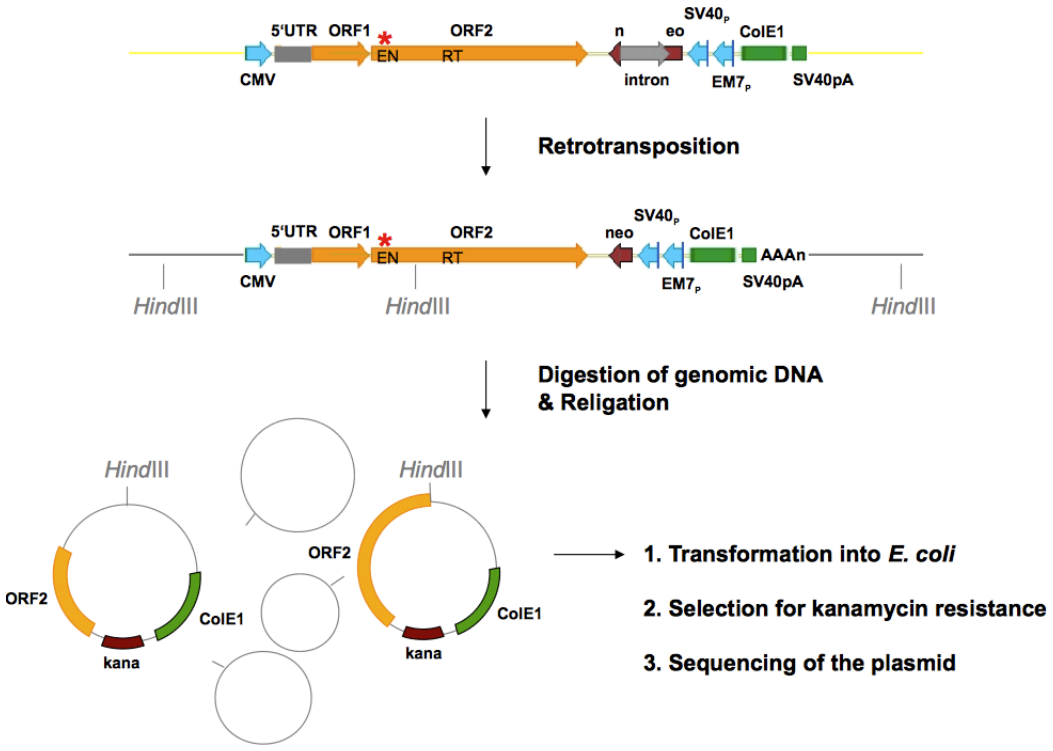


Figure 19 Schematic drawing of the procedure to rescue integrants derived from pCEP4/L1.3mneoI400/ColE1-based reporter constructs. Genomic DNA harboring a tagged L1 integrant is digested with *HindIII* and religated under dilute conditions. The ColE1 origin of replication introduced downstream of the *neo*/*kana* cassette converts the circularized genomic DNA containing a *de novo* L1 integration into a replication-competent bacterial plasmid. The prokaryotic EM7 promoter drives expression of the *kana* gene in bacterial cells. L1 copies truncated downstream of the L1-*HindIII* site can be recovered with 3' and 5' flanking sequences (rescue plasmid on the left hand side), while integrants extending beyond that position can only be recovered with their 3' junction (rescue plasmid on the right hand side). Plasmid DNA depicted in yellow lines, genomic DNA in gray lines. A red asterisk marks the site of EN point mutations.

De novo L1 integrants derived from pColE1-S202A, pColE1-R155A, pColE1-T192V and pColE1-I204Y were isolated by an adapted plasmid rescue procedure. Genomic DNA from G418 resistant HeLa cells was prepared using DNAeasy columns (Qiagen, Hilden) and restricted with *HindIII*. Fragments were then ligated under extremely dilute conditions (0.5-1 ng/µl) to favor intramolecular circularization. Typically, 300 ng restricted genomic DNA were incubated with 4 U T4 DNA ligase in a volume of 500 µl 1x ligation buffer at 16°C over night. The next day, glycogen was added to the ligation, and the mixture was then ethanol-precipitated and resuspended in 2 µl water. The entire concentrated ligation was used to transform electro-competent DH10B cells (Electromax DH10B, Invitrogen AG, Karlsruhe) by electroporation in 1 mm gap cuvettes (Bio-Rad, München). Transformed bacteria were selected on kanamycin plates. After plasmid isolation, the clones were characterized by analytical PCR and sequence analysis.

2.7 METHODS OF MOLECULAR BIOLOGY

Oligonucleotides used in this section are found in Chapter 2.10.1.

2.7.1 ANALYSIS OF LINE-1 INTEGRANTS BY MULTIPLEX PCR AND SEQUENCING

As L1 integrants are variably truncated at their 5' end, the length of the L1 copy had to be determined using a multiplex PCR with the antisense primer GS88 binding to the 5' end of the *neo* resistance cassette and a set of sense primers (GS260, GS261 and GS262) covering the L1 sequence downstream of the *HindIII* site (Figure 46). According to the appearance of various, progressively longer PCR fragments, the length of the L1 integration could be deduced. Then, using a corresponding antisense primer (GS88, GS17, GS14, GS16 or GS76), the 5' junction of the new integrant was sequenced (Figure 46).

Reaction mix:	PCR conditions:		
1 µl of a 1:20 dilution of plasmid DNA	96°C	4 min	
4 pmol of each diagnostic primer	96°C	20 sec	25x
0.2 mM of each dNTP	55°C	15 sec	
0.5 U *Taq* Polymerase	72°C	1 min	
2 µl Yellow Sub (10x)	72°C	5 min	
ad 20 µl H_2O	4° C	∞	

Yellow Sub (Geneo BioProducts GmbH, Hamburg) acts as enhancer of annealing specificity as well as loading buffer substitute at the same time.

2.7.2 PCR ON GENOMIC DNA FROM HELA CELLS

Most sequencing reactions of the rescued plasmids did not extend far enough to sequence the 3' junction between the L1 copy and the genomic DNA as well. Therefore, the obtained genomic sequences 5' of L1 were used as probes in BLAT searches in the human genome sequence to localize the genomic position of the new integrant. Then, to be able to sequence the junction of the L1 3' end, primers were designed that bind ~200 bp downstream of the presumed integration site. In cases where the L1 integrations extended upstream of the *HindIII* site, another genomic primer was designed ~200 bp upstream of the presumed integration site and used in conjunction with GS76 to amplify the 5' flanking sequence by PCR from genomic DNA of HeLa cells harboring the relevant insertions.

Reaction mix:	PCR conditions:		
1 µl of a 1:10 dilution of genomic DNA	96°C	4 min	
4 pmol of each diagnostic primer	96°C	20 sec	} 25x
0.2 mM of each dNTP	55°C	15 sec	} 25x
0.5 U *Taq* Polymerase	72°C	1 min	} 25x
2 µl Yellow Sub (10x)	72°C	5 min	
ad 20 µl H_2O	4° C	∞	

2.7.3 PREPARATION OF DNA FROM MOUSE TAIL BIOPSIES

1. Add 500 µl Tens(-) and 100 µl Proteinase K to each tail and incubate at least 2 h or better over night at 55°C in a shaker (300 rpm)
2. Centrifuge suspension for 10 min at 16600g in a table top centrifuge
3. Transfer supernatant to fresh tube and add 500 µl isopropanol, shake gently until precipitate forms
4. Centrifuge for 10 min at 16600g
5. Discard supernatant and wash pellet with 1 ml 70% ethanol (shake gently and incubate at room temperature for 5 min)
6. Centrifuge for 10 min at 16600g

7. Discard supernatant and let pellet dry with the tube up side down at room temperature for at least 10 min
8. Resuspend pellet in 200-800 µl H_2O (Millipore, autoclaved) (For regular tails from adult mice, approximately 2 mm long use 500 µl)

2.7.4 GENOTYPING OF LORFUS AND RORFUS TRANSGENIC ANIMALS

PCR mix:

	2 µl	Phusion buffer HF (5x)
+	1 µl	Yellow Sub (10x)
+	0.4 µl	dNTPs (5mM each)
+	0.2 µl	RSV3'LTRs (10 µM)
+	0.2 µl	GFP-IIas (10 µM)
+	0.1 µl	Phusion Pol. (2U/µl)
+	6.1 µl	H_2O
	10 µl	
	add 0.5 µl of tail DNA	

PCR conditions:

98°C	30 sec	
98°C	10 sec	35x
55°C	30 sec	35x
72°C	30 sec	35x
72°C	5 min	
4° C	∞	

Completed PCR reactions were loaded onto a 1% agarose gel for electrophoresis.

Expected band for transgenic animals: 1473 bp (for transgene containing intron)

Expected band for animals containing retrotransposition event: 574 bp (without intron)

2.7.5 GENOTYPING OF KT1 TRANSGENIC ANIMALS

PCR mix:

	2 µl	Phusion buffer HF (5x)
+	1 µl	Yellow Sub (10x)
+	0.4 µl	dNTPs (5mM each)
+	0.2 µl	Casli3 (10 µM)
+	0.2 µl	Ca25 (10 µM)
+	0.2 µl	Ca25as (10 µM)
+	0.1 µl	Phusion Pol. (2U/µl)
+	5.9 µl	H_2O
	10 µl	
	add 0.5 µl of tail DNA	

PCR conditions:

98°C	30 sec	
98°C	10 sec	35x
55°C	30 sec	35x
72°C	15 sec	35x
72°C	5 min	
4° C	∞	

Completed PCR reactions were loaded onto a 1% agarose gel for electrophoresis.

Expected band for transgenic animals: 500 bp

Expected band for wildtype animals: 380 bp

2.8 METHODS OF PROTEIN BIOCHEMISTRY

Antibodies used in this study are listed in Section 2.10.2.

2.8.1 PROTEIN PREPARATION FROM CELLS

First, cells were washed with an adequate volume of PBS. Then, Cell lysis buffer was added: 1 ml per confluent well of a 6-well dish and respectively less for less confluent wells. The lysis buffer was allowed to incubate at room temperature for 1-2 min. Afterwards, the cells were pipetted up and down 4-6 times for complete lysis and the well was flushed with the lysate to dislodge still adherent cells. The viscous lysates were then tranferred to a 1.5 ml tube and kept on ice for further processing. To fragment DNA, the lysates were sonicated several times until they were liquid enough to be pipetted without trouble. Then, the lysates were either stored at -70°C or protein concentration was determined immediately using the Bicinchoninic Acid Kit (Sigma-Aldrich, München) according to the manufacturer's instructions.

2.8.2 WESTERN BLOT

Depending on the experiment, 5-30 µg of protein extract were denatured in 1x SDS loading buffer by boiling for 5 min at 95°C. Subsequently, the samples were loaded onto an acrylamide gel. For Immunoblot analysis of p53, Ku70, PARP and Rad51 a 12% acrylamide resolving gel was used, and for DNA-PK an acrylamide resolving gel of 12% at the bottom and 8% on top was produced. For Western Blot analysis of ATM and ATR, I used NuPAGE 4-12% Bis-Tris Precast Gels (Invitrogen AG, Karlsruhe), since the resolution quality with self-made gels was not acceptable. Gelelectrophoresis was performed at 80-120 V in 1x Running buffer or for DNA-PK in 1x MOPS buffer yielding sharper bands in this case.
Wet transfer of the proteins onto a nitrocellulose membrane (Protran BA 85, Schleicher & Schuell, Keene, USA) was performed in Transfer buffer at 30 V and 90 mA over night. The membranes were then transferred into Blocking solution or Special blocking solution as for ATM, ATR and DNA-PK, since the skim milk powder containing Blocking solution resulted in very strong unspecific background signals after antibody incubation. After 1.5-2 h of blocking, the membranes were incubated with primary antibody solutions in PBS-T for 1 h.

Then, the membranes were washed 4 times for 10 min each in PBS-T and finally incubated in horse raddish peroxidase coupled secondary antibody at a concentration of 1:15000 in PBS-T. 45 min later, membranes were washed again 4 times for 10 min in PBS-T and subjected to ECL Plus Western Blotting Detection Reagent (GE Healthcare Europe GmbH, Freiburg) according to manufacturer's instructions. Chemiluminescence was detected using Amersham Hyperfilm ECL (GE Healthcare Europe).
Stripping of the membranes for a new antibody detection of other proteins was performed as described in the Amersham ECL Plus Western Blotting Detection Reagents instruction. Briefly, the membrane was incubated in stripping buffer at 50°C for 30 min with agitation. Then, it was washed twice at room temperature for 10 min in PBS-T and subsequently blocked again.

2.8.3 PREPARATION OF MOUSE BRAIN SECTIONS

Following deep anesthesia with Isofluran (Baxter Deutschland GmbH, Unterschleißheim) mice were sacrificed and the brains were prepared. Subsequently, they were fixated for 2 h in 4% PFA/PBS solution at 4°C and embedded in a 2% agarose/PBS gel for slicing. Sagittal sections were produced with a vibrating blade microtome (Leica Microsystems GmbH, Wetzlar) at 100 µm thickness and collected in PBS. For long-term storage at 4°C of up to 6 weeks, PBS was supplemented with sodium azide at a final concentration of 0.1%.

2.8.4 X-GAL STAINING OF MOUSE BRAIN SECTIONS

Brain sections were incubated in X-Gal solution until desired stain intensity was reached. Then, they were washed twice with 1x PBS and once with 10 mM TrisHCl pH 7.6 before mounting onto slides. Drying was performed over night at room temperature. For counter staining they were placed for 8 sec into 1% eosin/80% ethanol and washed twice in 95% ethanol. Finally, they were desiccated for 1 min in xylol, embedded with Eukitt (Sigma-Aldrich, München) and covered with a cover slip.

2.8.5 IMMUNOHISTOCHEMICAL STAINING OF MOUSE BRAIN SECTIONS

1. Place brain sections for blocking of endogenous peroxidase activity and for permeabilization for 10 min in 5% H_2O_2/PBS
2. Wash thrice in PBS, 10 min each
3. Transfer brain sections for blocking of unspecific antibody reactions 1 h into 10% normal goat serum in Day 1 buffer
4. Incubate brain sections over night with anti-GFP primary antibody 1:3000 in Day 1 buffer
5. Wash thrice in Day 2 buffer, 10 min each
6. Incubate brain sections 1 h with horse raddish peroxidase coupled secondary antibody 1:600 in Day 2 buffer
7. Wash thrice in Day 2 buffer, 10 min each
8. Transfer brain sections to 0.05% diaminobenzidine/0.01% H_2O_2/20 mM TrisHCl pH 7.6 and incubate until desired stain intensity is reached
9. Wash thrice in 20 mM TrisHCl
10. Transfer brain sections into 10 mM TrisHCl and mount onto slides
11. Dry over night
12. Desiccate for 1 min in xylol, embed with Eukitt and cover brain sections with cover slip

2.9 COMPUTATIONAL METHODS

2.9.1 HOMOLOGY SEARCHES

DNA homology searches were performed with BLAST and/or BLAT (http://blast.ncbi.nlm.nih.gov/ and http://genome.ucsc.edu/cgi-bin/hgBlat) (Altschul et al. 1990; Kent 2002). Sequence alignments were done using the computer program Vector NTI (Invitrogen AG, Karlsruhe) and, if required, further adjustments were made by hand.

2.9.2 SEQUENCE LOGOS

Sequence logos of consensus target sequences of the L1 EN variants were generated following the instructions on the website http://weblogo.berkeley.edu/logo.cgi (Crooks et al. 2004).

2.10 OLIGOS AND ANTIBODIES

2.10.1 NUCLEOTIDE SEQUENCES OF OLIGOS USED IN THIS STUDY

Oligo nucleotides were either purchased from MWG AG (Ebersberg) or Invitrogen AG (Karlsruhe).

Name	Sequence
RSV3'LTRs	CCATTTGACCATTCACCACATTGC
GFP-Is2	GTGAGCAAGGGCGAGGAGC
GFP-IIas	GGTGTTCTGCTGGTAGTGGTCGG
Casli3	TAAGCAGCTCTAATGCGCTGTTA
Ca25	GCTCAGAAGCCCCAAGCTCG
Ca25as	CAGCGCCTAACTCTGGACACC
GS10	TTTCTTCCTAGTCTCCATGGTCTTTAC
GS11	GTATCAGCCATGGAAGATGAAATGAATG
GS14	GTGTTTTGGCCATGGAGTCCTTGCCC
GS16	TTTCTTTCTGCAGTGGTTTGTAGTTCTC
GS17	GTGTCCATGTGAATTCATTGTTCAATTCC
GS51	CCCATATGCAGGATCAAATTCACACATAAC
GS52	CCGGATCCAATCCTGAGTTCTAGTTTGATTGC
GS73	GGAAACCCATCTCACGTG
GS76	CCCCAGTATAGTGACCCGATCGACCCATTCATAAAGCAAGTCCTCAG
GS88	CCTTCTATCGCCTTCTTGACGAGTTCTTC
GS90	TTCCACACCCTAACTGACACACATTCC
GS189	GCTGGTGAGGAACTGCGTTCCTTTGG
GS260	CAGGTGCTGGAGAGGATGCGGAG
GS261	CCTCAGAAATAATGCCGCATATC
GS262	CTAGAAAACCCCATCGTCTCAGC
GS263	GTGTCGAGGAATGTATCC
GS334	CCACACCACACCTATGCCAAAATTGACCACATAG
GS335	GTGGTCAATTTTGGCATAGGTGTGGTGTGGTGC
GS336	GTCAACATTAGACGCATCAACGAGACAGAAAGTC
GS337	TTCTGTCTCGTTGATGCGTCTAATGTTGACAGTGG
GS338	ATCAACAGAATATGTCTTTTTTTCAGCACCACAC
GS339	GGTGCTGAAAAAAAGACATATTCTGTTGATTTGGG
GS340	CACACCTATTCCAAATATGACCACATAGTTGGAAG
GS341	CAACTATGTGGTCATATTTGGAATAGGTGTGGTG

Table 2 Oligos used in this work. Sequences presented in 5' to 3' orientation.

2.10.2 ANTIBODIES USED IN THIS STUDY

Designation	Order data	Application
Mouse anti p53 antibody (Ab-1)	OP03, Calbiochem, Bad Soden	used at 1 µg/ml in Western Blot
Mouse anti Ku70 antibody (Ab-4)	N3H10, Neomarkers, Inc., Fremont, USA	used at 0.25 µg/ml in Western Blot
Rabbit anti ATM antibody (Ab-3)	PC116, Calbiochem, Bad Soden	used at 2 µg/ml in Western Blot
Rabbilt anti ATR antibody (Ab-2)	PC538, Calbiochem, Bad Soden	used 1:2000 in Western Blot
Mouse anti DNA-PKcs (Ab-1)	18-2, Neomarkers, Inc., Fremont, USA	used at 1µg/ml in Western Blot
Goat anti PARP-1 antibody (N-20)	sc-1561, Santa Cruz Biotechnology, Santa Cruz, USA	used 1:100 in Western Blot
Rabbit anti Rad51 antibody (H-92)	sc-8349, Santa Cruz Biotechnology, Santa Cruz, USA	used 1:200 in Western Blot
Mouse anti β-actin antibody (AC-74)	A2228, Sigma-Aldrich, München	used 1:40000 in Western Blot
Horse raddish peroxidase coupled goat anti mouse IgG antibody	A9917, Sigma-Aldrich, München	used 1:20000 in Western Blot
Horse raddish peroxidase coupled goat anti rabbit IgG antibody	A0545, Sigma-Aldrich, München	used 1:10000 in Western Blot
Horse raddish peroxidase coupled rabbit anti goat IgG antibody	A8919, Sigma-Aldrich, München	used 1:20000 in Western Blot
Rabbit anti-GFP antibody	ab6556, abcam, Cambridge, UK	used 1:3000 in Immunohistochemistry
Horse raddish peroxidase coupled goat anti rabbit IgG antibody	111-035-144, Dianova, Hamburg	used 1:600 in Immunohistochemistry

Table 3 Antibodies used in this work.

3 RESULTS

3.1 GENERATION OF A TETRACYCLIN-INDUCIBLE LINE-1 RETROTRANSPOSITION REPORTER ASSAY

The tetracyclin-dependent transactivator (tTA) system is based on the artificial fusion protein tTA that is capable of initiating transcription from a synthetic promoter in a tetracyclin-dependent fashion. The fusion protein consists of the *E. coli* repressor of the tetracyclin resistance gene (Tet repressor) and the protein binding domain of the transcription factor VP16 from *Herpes simplex* virus. In *E. coli*, the Tet repressor binds with high specificity in the absence of tetracyclin (Tet) to the Tet operator sequence to suppress the transcription of the Tet resistance gene by means of sterical inhibition. If Tet is taken up by the bacterium, the antibiotic binds to the Tet repressor which in turn dissociates from the operator sequence due to conformational changes and triggers transcription of the Tet resistance gene (Gossen and Bujard 1992).

The tTA-dependent operon harbors seven copies of the Tet operator (Tet responsive element, TRE) followed by a TATA box from the immediate early gene promoter of the human cytomegalovirus (CMV). The TATA box is needed for eukaryotic transcription initiation, it is where the RNA polymerase after activation by transcription factors, including VP16, binds to and begins with transcription. In contrast to the complete CMV promoter, the CMV minimal promoter used for Tet systems covers only approximately 100 bp and does not contain any binding sites for transcription factors. This function is conducted by the TRE sequence. Therefore, only upon binding of the repressor subunit of the tTA fusion protein to the TRE in absence of Tet and function of the VP16 domain as transcription factor, transcription is initiated at the CMV minimal promoter. Vice versa, in the presence of Tet, the affinity of the repressor for the TRE is reduced and the transcription is disrupted (Figure 20, top).

Through mutations in the Tet repressor, the tTA system was further modified to yield a reverse tTA (rtTA) system in which the repressor binds to the TRE only in presence of Tet. Therefore, in the rtTA system transcription of the gene of interest is initiated by adding Tet and disrupted upon withdrawal of the antibiotic (Gossen et al. 1995) (Figure 20, bottom).

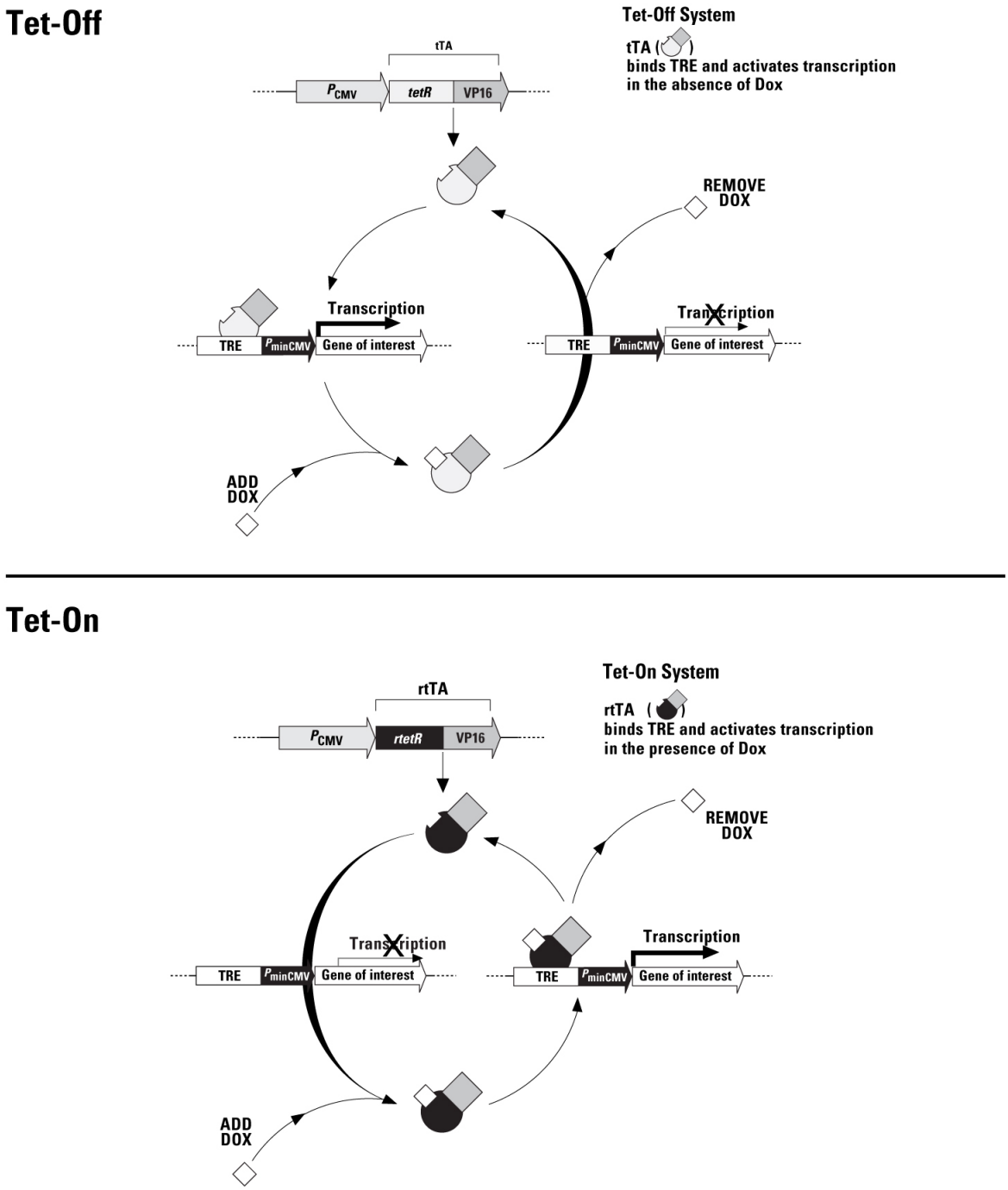


Figure 20 Schematic of gene regulation in the Tet-Off and Tet-On Systems. Tet-Off: The TRE is located upstream of the minimal immediate early promoter of cytomegalovirus (PminCMV), which is silent in the absence of activation. tTA binds the TRE - and thereby activates transcription of Gene X - in the absence of Tet or Dox. Tet-On: The "reverse" Tet repressor (rTetR) was created by four amino acid changes that reverse the protein's response to Dox. As a result of these changes, the rTetR domain of rtTA binds the TRE and activates transcription in the presence of Dox (from Clontech Laboratories, Inc., 2005).

In most cases, a derivative of tetracyclin, doxycyclin (Dox), is used for the control of the tTA (Tet-Off System) or rtTA (Tet-On System). This is mainly due to the fact that in animal

applications, Dox is not stored in bones in contrast to Tet, and therefore can be flushed out of the organism more easily allowing for quick gene regulation experiments.

In this study, I chose the application of the Tet-On System in favor of a rapid gene induction upon Dox administration. To place the L1 retrotransposon under the control of the Tet-On System, I cloned the TRE followed by the immediate early CMV promoter (PhCMV) upstream of the L1 element into a plasmid containing a blasticidin resistance gene as retrotransposition reporter (Morrish et al. 2002). At the same time, the L1 5' UTR was omitted to eliminate the element's internal promoter and exclude PhCMV-independent transcription (Figure 16 and Figure 21).

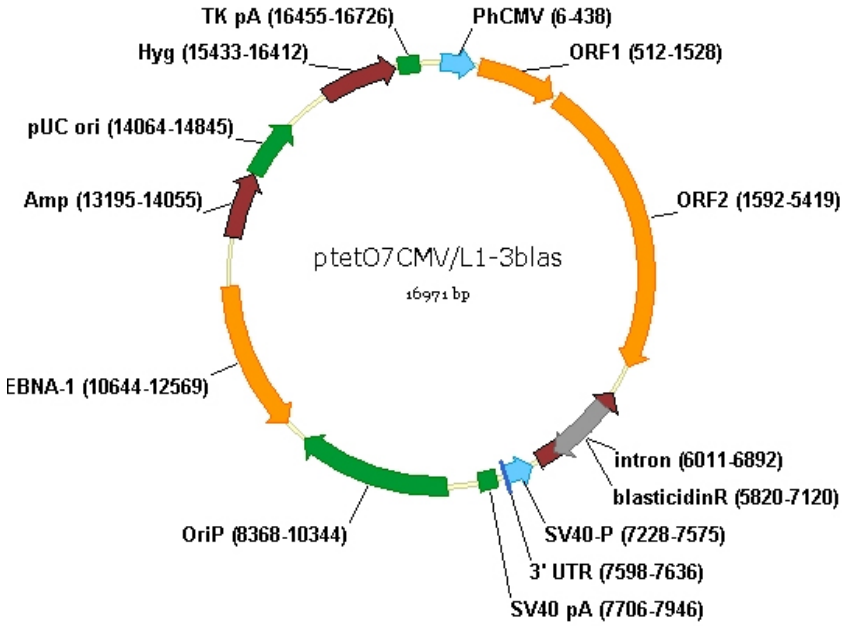


Figure 21 Plasmid map of ptetO7CMV/L1-3blas coding for the Tet/Dox-inducible L1 reporter element. From this plasmid L1 is transcribed exclusively under the control of the TRE and immediate early gene CMV promoter (PhCMV). Plasmid size and relative positions of regulatory elements as well as of ORFs are indicated. PhCMV, TRE followed by immediate early gene CMV promoter; ORF1, open reading frame 1 of L1; ORF2, open reading frame 2 of L1; blasticidinR, blasticidin resistance gene interrupted by intron; intron, beta globin intron; SV40-P, Simian virus 40 promoter; 3' UTR, 3' untranslated region of L1; SV40 pA, Simian virus 40 polyA signal; OriP, eukaryotic origin of replication; EBNA-1, Eppstein Barr Virus nuclear antigen 1; Amp, ampicillin resistance gene; pUC ori, prokaryotic origin of replication; Hyg, hygromycin resistance gene; TK pA, thymidine kinase polyA signal.

To test the cloned tTA responsive L1 reporter, I transfected the plasmid into HeLa M2 cells stably expressing the rtTA fusion protein under the control of a human elongation factor 1 promoter (Hampf and Gossen 2007). To ensure robust binding of rtTA to the Tet operator elements of the L1 reporter, Dox was added to the medium three days earlier. The results

show clearly that retrotransposition from ptetO7CMV/L1-3blas can be controlled by the addition of Dox to the medium (Figure 22). Without Dox application, a weak background expression of the Tet-inducible L1 construct can be observed by the generation of some few blasticidin resistant colonies. This "leakiness" of the PhCMV promoter was discovered already earlier in many different studies and is therefore known. However, since the extent of background expression can be considered as negligible in this assay, it was not quantified.

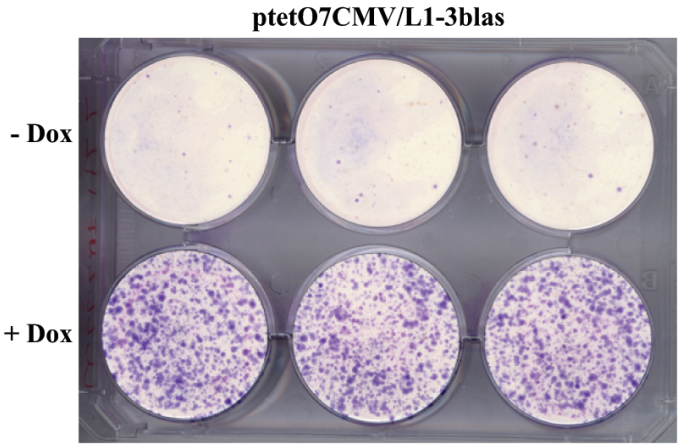


Figure 22 Proof-of-principle of the Dox-inducible L1 reporter system encoded by ptetO7CMV/L1-3blas in HeLa M2 cells. In each well, $2x10^6$ HeLa M2 cells were transfected with 1 µg of ptetO7CMV/L1-3blas. The three wells on the bottom were treated with Dox (+Dox), while the three wells on the top were incubated without Dox (-Dox). All cells were subsequently selected with blasticidin for *de novo* retrotransposition events. After addition of Dox, L1 retrotransposition is significantly increased.

This Tet/Dox inducible L1 reporter system was the basis to file an application for a patent entitled "Controlled activation of non-LTR retrotransposons in mammals" (European Patent Application, Date of publication 06.05.2009, Bulletin 2009/19, Application number 07021311.1; US Patent Application, Filing Date 10/11/2010, Application Number 12/740,943).

3.2 GENERATION OF A TRANSGENIC MOUSE LINE EXPRESSING A GFP-TAGGED LINE-1 REPORTER ELEMENT UNDER CONTROL OF THE TETRACYCLIN-DEPENDENT TRANSACTIVATOR

As I had generated and characterized transgenic mice expressing engineered glutamate receptor subunits under the control of a forebrain-specific tTA promoter system during my diploma thesis ("Funktioneller Austausch der endogenen GluR-A-Untereinheit durch GluR-

A-Varianten im Hippokampus der Maus", 2003), and due to the recently published hypothesis of LINE-1 playing a role in brain development by contributing to somatic mosaicism in neuronal precursor cells (NPCs) (Muotri et al. 2005), I decided to generate a transgenic mouse line expressing a GFP-tagged LINE-1 reporter construct enabling the analysis of L1 retrotransposition in the brain. For that purpose, I designed two constructs for pronucleus injection into fertilized mouse oocytes.

The plasmids pnlacZ-ORFeus and pRed2-ORFeus (cloned by Sinja Wendel, Paul-Ehrlich-Institut, Langen) contain a bidirectional tTA-dependent CMV promoter cassette (Baron et al. 1995) driving simultaneous expression of a GFP-tagged LINE-1 reporter and either β-galactosidase (nlacZ) or a red fluorescent protein (DsRed2), respectively (Figure 23 A and B). The L1 reporter contain a LINE-1 element codon usage optimized for expression in murine cells (ORFeus) (Han and Boeke 2004; An et al. 2006). pnlacZ-ORFeus and pRed2-ORFeus were linearized by digestion with restriction enzymes *PsiI* and *NaeI* yielding the final LORFUS and RORFUS bidirectional expression constructs (Figure 23 C and D) (linearization performed by Sinja Wendel). After purification by sucrose density gradient centrifugation, the DNA fragments were used for pronucleus injection (pronucleus injection performed by Frank Zimmermann, Zentrales Tierlabor, Heidelberg). 268 and 270 oocytes were injected with the LORFUS and RORFUS constructs, respectively. The procedure resulted in 39 LORFUS and 36 RORFUS founder mice of which 37 and 31 reached the age of weaning, respectively. After tail DNA preparation, the surviving founders were genotyped by PCR using primers GFP-Is2 and GFP-IIas amplifying the GFP region across the intron (Figure 24). Three LORFUS and seven RORFUS founder turned out to be transgenic for the injected sequences (Figure 25). Interestingly, RORFUS founder 16 seems to carry only a spliced GFP sequence. This is probably due to a retrotransposition event that occurred very early after pronucleus injection before degradation of the injected transgene itself took place and as a consequence of leakiness of the tTA-dependent CMV promoter. Although it has been reported previously that L1 is active in germ cells and during early embryogenesis (Ostertag et al. 2002; Prak et al. 2003), this is one of the earliest retrotransposition events observed and has also been reported by others (Prak et al. 2003; An et al. 2006).

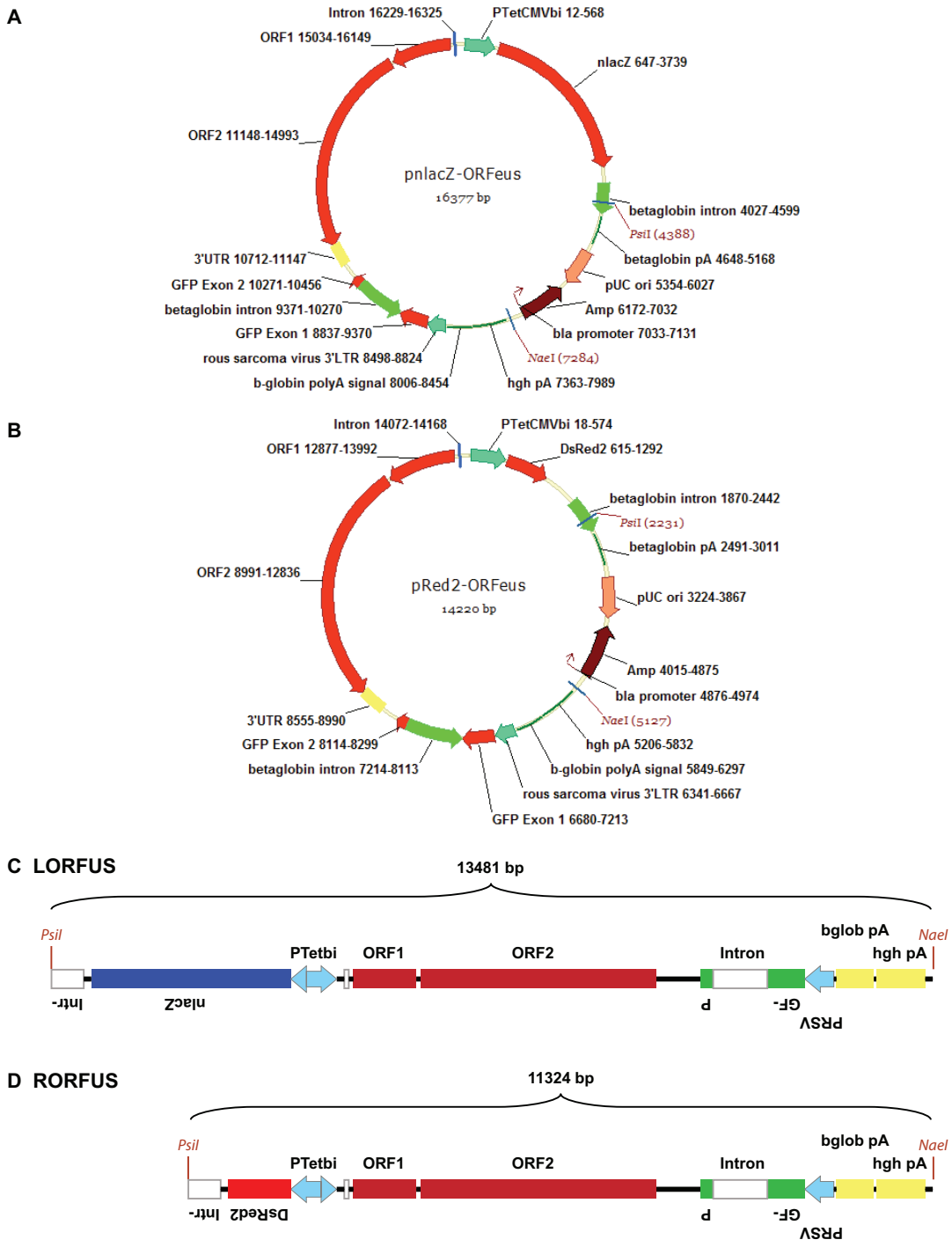


Figure 23 LORFUS and RORFUS constructs. A pnlacZ-ORFeus. Plasmid with bidirectional cassette driving expression of a LINE-1 reporter with GFP and β-galactosidase (nlacZ). **B** pRed2-ORFeus. Plasmid with bidirectional cassette driving expression of a LINE-1 reporter with GFP and a red fluorescent protein (DsRed2). **C** LORFUS construct for pronucleus injection after linearization of pnlacZ-ORFeus with *PsiI* and *NaeI*. **D** RORFUS construct for pronucleus injection after linearization of pRed2-ORFeus with *PsiI* and *NaeI*. PTetCMVbi and PTetbi, bidirectional tTA-dependent CMV promoter containing seven repeats of the Tet operator sequence; PRSV, rous sarcoma virus 3' LTR containing a promoter sequence; bglob pA, β-globin polyA signal; hgh pA, human growth hormone polyA signal.

Since the primer pair chosen is capable of amplifying almost any GFP sequence and this is a very commonly found DNA in molecular biology laboratories, I designed another sense primer for the genotyping of LORFUS and RORFUS animals to reduce the risk of identifying false positives because of PCR contaminations. Therefore, the new primer binds further upstream of the GFP reporter within the RSV3'LTR (Figure 24). This new primer pair of GFP-IIas and RSV3'LTRs prove to be very reliable in further genotyping PCRs of offspring of the LORFUS and RORFUS founder.

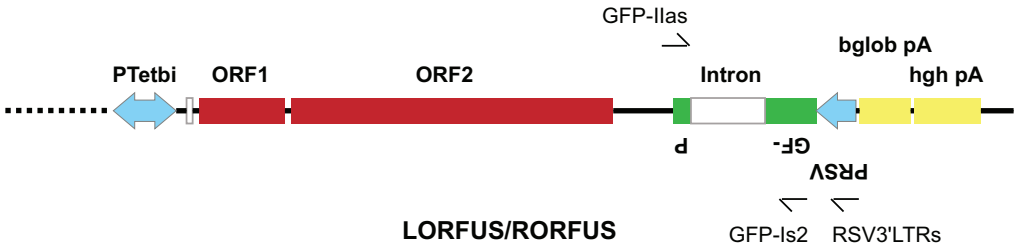


Figure 24 Schematic representation of the binding sites of primers GFP-IIas, GFP-Is2 and RSV3'LTRs used for genotyping LORFUS and RORFUS mice. For PCR either GFP-IIas and GFPIs2 yielding a 1435 bp fragment (536 bp in case of a spliced construct) or GFP-IIas and RSV3'LTRs yielding a 1473 bp fragment (574 bp in case of a spliced construct) were used. Dotted line represents either nlacZ (LORFUS) or DsRed2 (RORFUS) sequence.

To analyze expression of the different LORFUS and RORFUS lines, the transgenic founders were crossbred with Kt1 mice harboring a tTA gene under the control of the calcium-calmodulin-dependent kinase II (CaMKII) promoter (Mayford et al. 1996) which thus drives forebrain specific expression. Analysis of the double transgenic LORFUSxKt1 offspring in X-Gal stainings for *nlacZ* expression revealed the absence of any β-galactosidase expression in animals from line 4 and line 31 (data not shown). Only in a mouse of line 11 sacrificed at postnatal day 15 (P15), I found a Kt1 typical *nlacZ* expression in the forebrain including hippocampus, cortex and striatum indicating activation of the bidirectional tTA-dependent CMV promoter (Figure 26). Furthermore, in an immunohistochemical staining using an anti-GFP primary antibody, several single cells in different brain regions revealed GFP expression indicating a retrotransposition event. GFP positive cells could be detected in all hippocampus regions, cortex, striatum, olfactory bulb and brainstem (Figure 26).

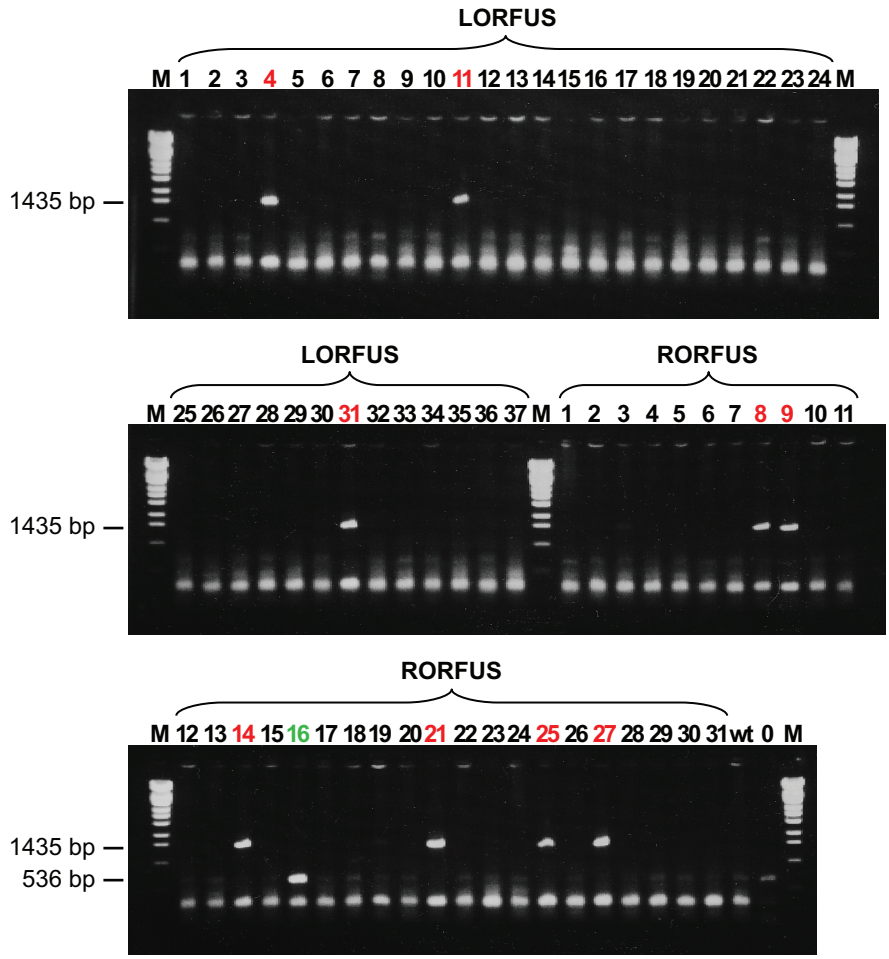


Figure 25 Genotyping PCR results of LORFUS and RORFUS founder with primers GFP-Is2 and GFP-IIas. Founder animal numbers are indicated above each lane. Founder carrying the LORFUS or the RORFUS transgene are marked in red. RORFUS founder 16 (green) harbors the spliced GFP sequence only. Although the control with the PCR master mix alone (0) shows a faint band at the size of the spliced GFP, the results were judged to be reliable. M, DNA marker made from λ Phage DNA digested with *StyI*; wt, wildtype control; 0, PCR master mix alone.

In none of the analyzed double transgenic RORFUSxKt1 offspring expression of the DsRed2 could be shown, neither by direct fluorescence nor by immunohistochemical staining using an anti-DsRed2 antibody (data not shown). Out of nine analyzed transgenic LORFUS and RORFUS mouse lines only one expressed the reporter on the opposite side of the LINE-1 cassette. Although silencing of transgenes is an often observed result of positional effects due to the integration site, the finding raised the question for quality and reliability of the integrated bidirectional expression cassettes. Therefore, I analyzed and annotated the entire nucleotide sequences of the DNAs that were used for the pronucleus injections. I found that through linearization with *PsiI* and *NaeI*, the expression cassettes for *nlacZ* and DsRed2 had been truncated within the β-globin intron and thus were incomplete lacking also the β-globin polyA signal. This explained the rather poor expression of *nlacZ* and DsRed2. Hence, the

integration of the LORFUS transgene in line 11 must have occurred at a favorable position to compensate for the lacking elements, maybe upstream of an endogenous polyA signal.

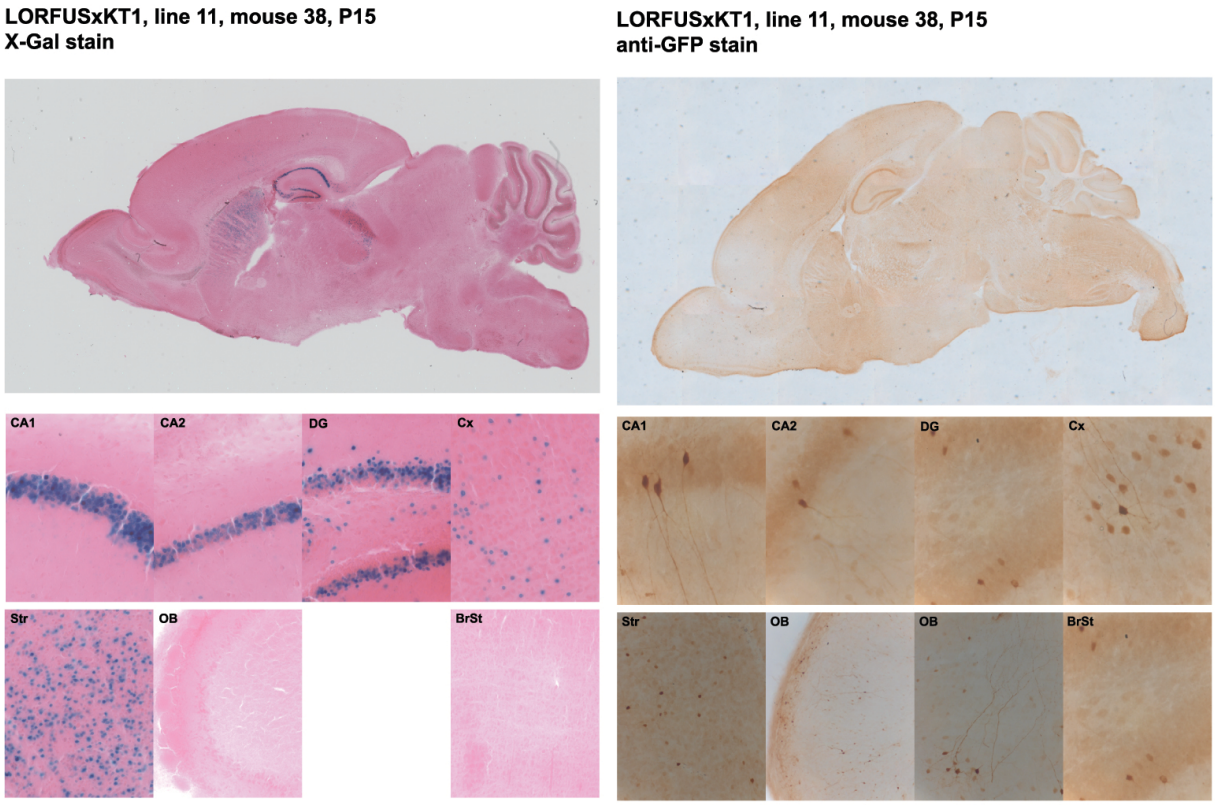


Figure 26 Sagittal brain sections of LORFUSxKt1 mouse 38 from LORFUS line 11 prepared at postnatal day 15. Left side X-Gal stain for expression of *nlacZ*/β-galactosidase activity, right side immunohistochemical stain for expression of GFP. Top pictures show overviews of the entire sections, bottom pictures show zooms into specific brain regions. General transgene expression indicated by β-galactosidase activity is found in hippocampus, cortex and striatum, while retrotransposition events indicated by GFP expression are found in hippocampus, cortex, striatum, olfactory bulb and brainstem. CA1 and CA2, cornu ammonis regions 1 and 2 of the hippocampus; DG, dentate gyrus region of the hippocampus; Cx, cortex; Str, striatum; OB, olfactory bulb; BrSt, brainstem.

3.3 IDENTIFICATION OF HOST-ENCODED DOUBLE-STRAND BREAK (DSB) REPAIR FACTORS INVOLVED IN HUMAN L1 RETROTRANSPOSITION

To analyze the role of DSB repair factors on L1 retrotransposition, I made use of the RNA interference (RNAi) method to knock down single cellular proteins. The process of siRNA triggered RNAi in mammalian cells is transient, as dsRNA is degraded. To harness this process as genetic tool, the transient nature can be circumvented by expression of stable

intracellular siRNAs. This approach may use either transcription of sense and antisense siRNAs or expression of short hairpin RNAs (shRNAs) from plasmids using Pol III promoters. Promoters for U6 snRNA (Paddison et al. 2002; Yu et al. 2002) and RNAse P RNA H1 (Brummelkamp et al. 2002) have been successfully used to transcribe shRNAs. The Pol III system is suitable for expression of small non-coding RNAs with defined 5' and 3' ends, generating roughly $4x10^5$ transcripts per cell (Weinberg and Penman 1968). Transcriptional termination occurs via a stretch of four to five consecutively transcribed uridines.

Another strategy to study the function of genes is the creation of mutants conferring a so-called "dominant-negative" phenotype in the presence of the wildtype protein (Herskowitz 1987). In the case of proteins consisting of different functional domains, it is conceivable to create a dominant-negative mutant either by overexpressing a full-length variant mutated for instance in the catalytic site or a truncated version selectively coding for a regulatory function (e. g. the DNA-binding domain). Furthermore, also trans-species dominant-negative effects have been described (Milne and Weaver 1993).

To test the effect of the knockdown (KD) or overexpression of dominant negative (DN) mutants of DSB repair factors on L1 retrotransposition, co-transfections into HeLa cells with the L1 retrotransposition reporter vector and the respective shRNA expressing construct were performed. For this study I used pJM101/$L1_{RP}$, a plasmid constructed just as pJM101/L1.3, but containing a highly active L1 element published previously (Kimberland et al. 1999). $L1_{RP}$ was isolated from a patient with X-linked retinitis pigmentosa and has approximately twice the activity of L1.3. pJM101/$L1_{RP}$ and pJM101/L1.3 both carry hygromycin resistance genes as selectable marker (Figure 15).

It is a common procedure in transfection studies to select for only one of the co-transfected plasmids, as it is accepted that cells harboring one of the plasmids also took up the other at the same time (Tuck and Crawford 1989; Wong et al. 1995). The exact mechanism underlying this practice is unknown. Additionally, in this work, I always applied the plasmid, that was not directly selected for, in molecular excess. Therefore, one day after transfection, hygromycin (Hyg) was added to select for transfected cells. Each well of transfected cells was split three days post transfection in a ratio of 1:2 and the smaller subculture was further selected with Hyg for assessment of transfection efficiency and toxicity. To determine retrotransposition efficiency, the other subculture was subsequently selected with G418.

To determine the transfection efficiency and toxicity of each transfection experiment, the obtained Hyg resistant colony numbers resulting from transfection with the L1 reporter and

with the different DSB repair interfering constructs were compared with Hyg resistant colony numbers from co-transfection experiments with the L1 reporter and the corresponding empty vectors of the DSB repair interfering constructs. The difference in colony numbers indicated the degree of transfection efficiency and toxicity.

To determine the effect of the DSB repair factor interference on L1 retrotransposition, the relative G418 resistant colony numbers (%) from each experiment were compared with the relative Hyg resistant colony numbers (%) from the same co-transfection. If the KD or DN mutant did not have any effect on L1 retrotransposition, one would expect the relative G418 colony numbers to be the same as the relative Hyg colony numbers from the same transfection experiment, both reflecting just transfection efficiency and toxicity. However, if the KD or DN mutant lead to a decrease in L1 retrotransposition frequency, one would expect the relative G418 colony number being less than the relative Hyg colony number from the same co-transfection. In contrast, if KD or DN mutant had a positive effect on L1 retrotransposition, one would expect the number of G418 colonies increasing relative to the respective Hyg colony number. For normalization, the obtained Hyg and G418 colony numbers resulting from transfection with the L1 reporter plasmid pJM101/$L1_{RP}$ and the different KD constructs were each compared to co-transfection experiments with the L1 reporter plasmid and corresponding empty vectors of the KD constructs.

3.3.1 KNOCKDOWN OF CELLULAR DSB REPAIR FACTORS VIA RNAI AND INFLUENCE ON L1 RETROTRANSPOSITION

In this study, the U6 promoter was used to drive expression of shRNAs directed against mRNAs of ATM, ATR and DNA-PK (Collis et al. 2003), whereas the H1 promoter was used for transcription of shRNAs against p53 (Brummelkamp et al. 2002) and Ku70 (Zsuzsanna Izsvák, Max-Delbrück-Centrum für Molekulare Medizin, Berlin) gene products (Figure 27 and Figure 28).

After co-transfection of the L1 retrotransposition reporter pJM101/$L1_{RP}$ and the respective shRNA expressing construct, the cells were subjected to Hyg and G418 selection as described above (Chapter 3.3). Prior to analyzing retrotransposition frequencies and toxicity of each transfection experiment, I checked the expression levels of each DSB repair factor at different time points after co-transfection of the L1 reporter plasmid and the respective shRNA construct by Western Blot analyses of total cell extracts from the Hyg selected cells.

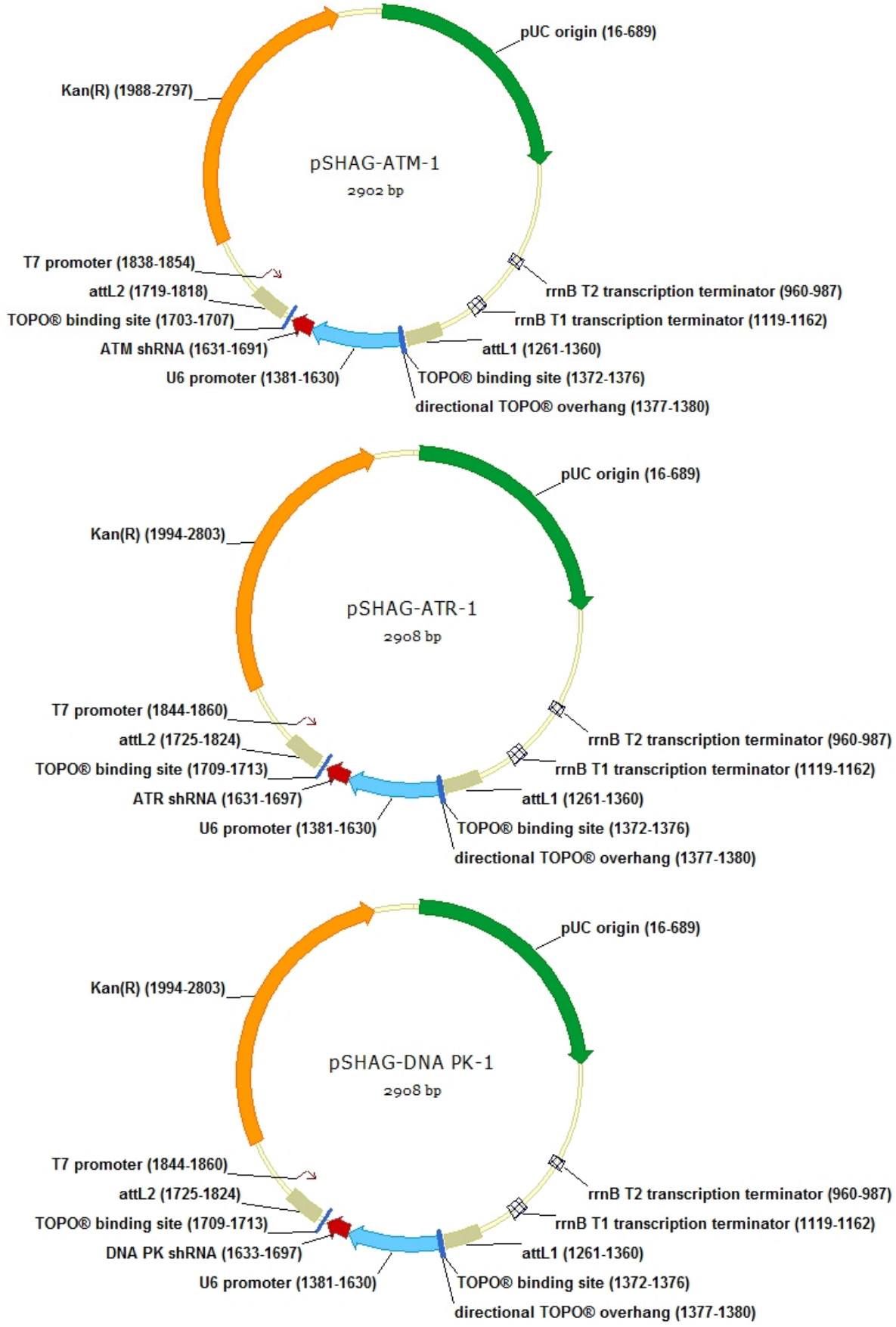


Figure 27 pSHAG based shRNA expressing vectors used in this study. Top, pSHAG-ATM-1; Middle, pSHAG-ATR-1; Bottom, pSHAG-DNA PK-1 (Collis et al. 2003). Functional plasmid elements are indicated.

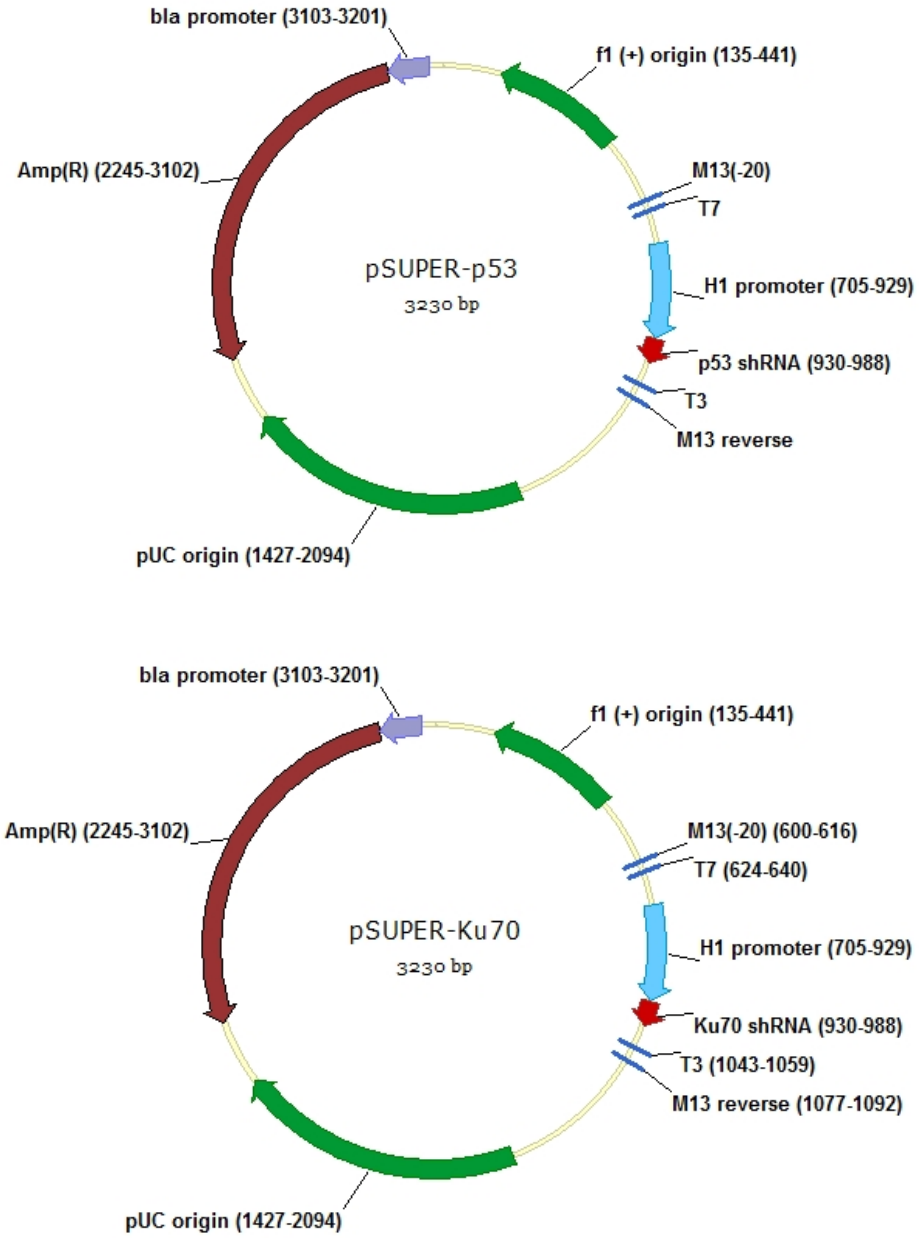


Figure 28 pSUPER based shRNA expressing vectors used in this study. Top, pSuper-p53 (Brummelkamp et al. 2002); Bottom, pSuper-Ku70 (Zsuzsanna Izsvák, Max-Delbrück-Centrum für Molekulare Medizin, Berlin). Functional plasmid elements are indicated.

Immunoblot analysis with an anti-p53 antibody uncovered that the KD was quite efficient, because there was barely any p53 detectable after 2 days after co-transfection of pJM101/$L1_{RP}$ and the shRNA expressing plasmid pSUPER-p53 (Figure 29).

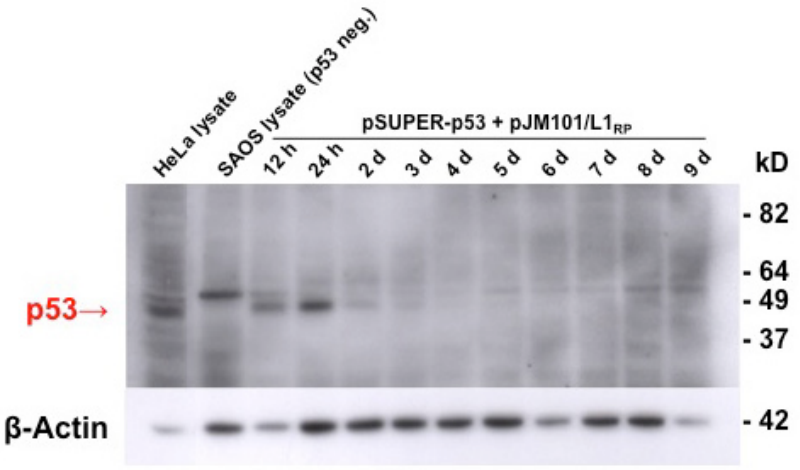


Figure 29 Western Blot showing the knockdown of p53 after co-transfection of pSUPER-p53 and pJM101/L1$_{RP}$. In each lane, 15 µg of total cell lysate were loaded. Time points at which lysates were prepared post transfection are indicated. Lysate from SAOS cells that do not express p53 was used as negative control. β-actin served as loading control. p53 KD is detectable at day 2 after transfection and later.

Although hardly any cells remained after transfection of pSUPER-Ku70 and pJM101/L1$_{RP}$, all material was collected and Western Blot was performed (Figure 30). Unexpectedly, immunoblot analysis with an anti-Ku70 antibody did not provide any evidence for an interference of the Ku70-specific shRNA with Ku70 expression.

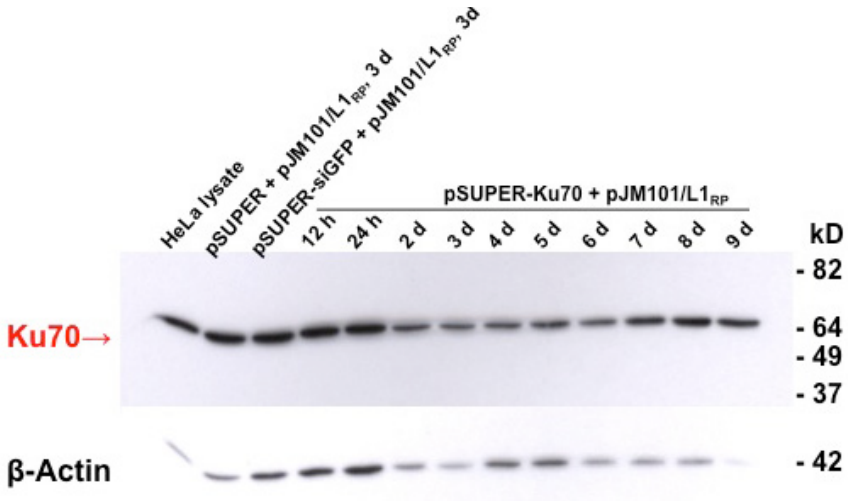


Figure 30 Western Blot showing the knockdown of Ku70 after co-transfection of pSUPER-Ku70 and pJM101/L1$_{RP}$. In each lane, 5 µg of total cell lysate were loaded. Time points at which lysates were prepared post transfection are indicated. Lysates from HeLa cells transfected with either the empty vector (pSUPER) or an unrelated control shRNA construct (pSUPER-siGFP) were used as controls. β-actin served as loading control. At no time point a Ku70 KD is detectable.

This indicates that the observed toxicity is not a consequence of Ku70 KD, but rather results from an unspecific side effect of the shRNA.

Immunoblot analysis with an anti-ATM antibody of cell extracts from HeLa cells co-transfected with pSHAG-ATM and pJM101/L1$_{RP}$ uncovered that only at day 7 post transfection, a meaningful KD of ATM expression could be observed (Figure 31). Surprisingly, this reduction of ATM expression was observed for only one day.

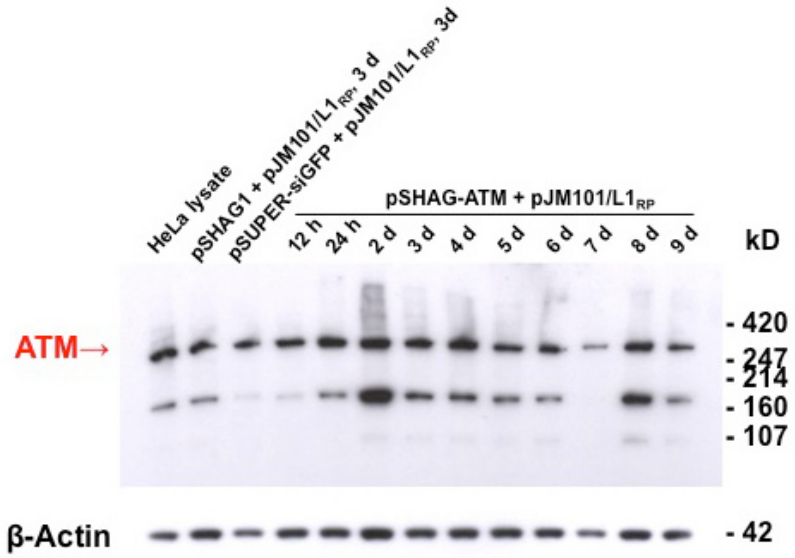


Figure 31 Western Blot showing the knockdown of ATM after co-transfection of pSHAG-ATM and pJM101/L1$_{RP}$. In each lane, 30 µg of total cell lysate were loaded. Time points at which lysates were prepared post transfection are indicated. Lysates from HeLa cells transfected with either the empty vector (pSHAG1) or an unrelated control shRNA construct (pSUPER-siGFP) were used as controls. β-actin served as loading control. A faint ATM reduction at 7 days post transfection is detectable.

ATRkd cells (Cliby et al. 1998) are transformed human fibroblasts expressing a kinase-dead ATR (ATRkd) under the control of an inducible Tet-On system. These cells were either grown in the presence of 1 µg/ml Dox for 48 hours to induce ATRkd expression (ATR+) or left untreated (ATR-). The expression of the dominant negative ATRkd leads to a competitive inhibition of ATR. Unfortunately, these cells do not support L1 retrotransposition neither with nor without Dox supplementation (data not shown). Nevertheless, they were used as controls for Western Blotting, as the ATRkd is recognized by the ATR antibody used in this study.

Immunoblot analysis demonstrates that after co-transfection of pSHAG-ATR and pJM101/L1$_{RP}$ into HeLa cells, ATR levels drop substantially only at day 6 after transfection (Figure 32). Comparison of β-actin signal intensities with ATR protein levels also show that ATR expression starts to decline slightly 2 days after transfection.

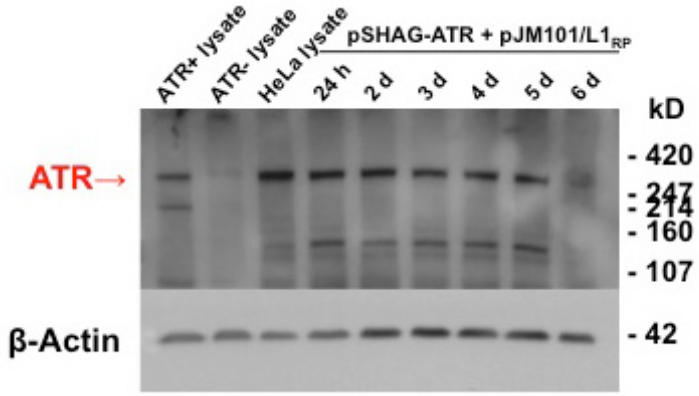


Figure 32 Western Blot showing the knockdown of ATR after co-transfection of pSHAG-ATR and pJM101/L1$_{RP}$. In each lane, 30 µg of total cell lysate were loaded. Time points at which lysates were prepared post transfection are indicated. Lysates from ATRkd cells with Dox (ATR+) and without Dox (ATR-) and from HeLa cells were used as controls. β-actin served as loading control. ATR reduction at 6 days post transfection is detectable.

Western Blot analysis with an anti-DNA-PK antibody shows that DNA-PK levels in HeLa cells transfected with pSHAG-DNA PK and pJM101/L1$_{RP}$ disappear almost entirely at 12 hours post transfection and are highly reduced from day 7 after co-transfection on (Figure 33).

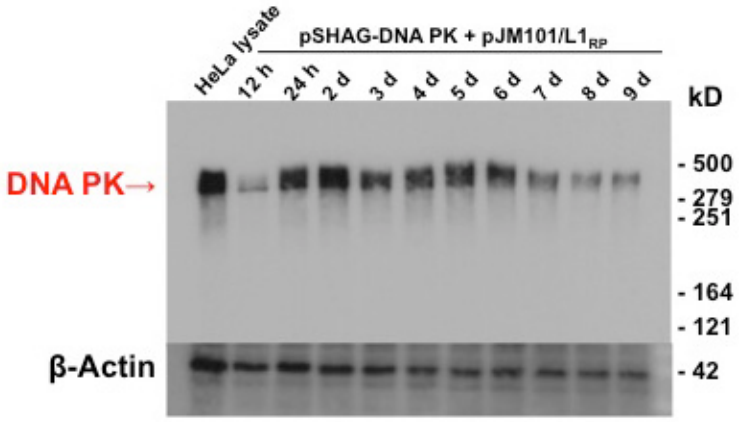


Figure 33 Western Blot showing the knockdown of DNA-PK after co-transfection of pSHAG-DNA PK and pJM101/L1$_{RP}$. In each lane, 10 µg of total cell lysate were loaded. Time points at which lysates were prepared post transfection are indicated. Lysate from untransfected HeLa cells was used as control. β-actin served as loading control. An early DNA-PK reduction at 12 hours post transfection and later at 7-9 days is detectable.

Analysis of L1 activity and toxicity by selection with G418 and Hyg, respectively, yielded very different results for the tested KD constructs (Figure 34 and Table 4).

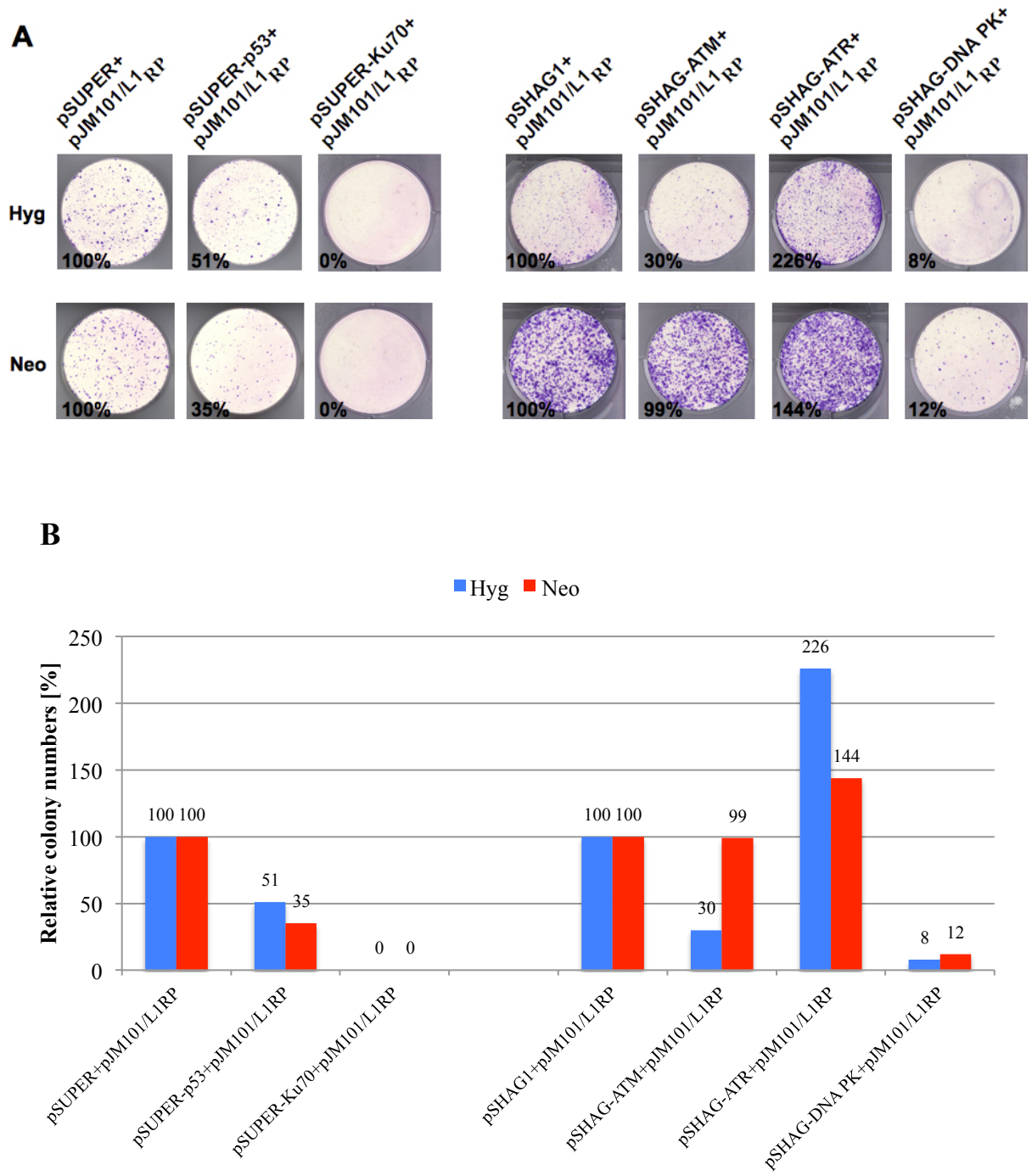


Figure 34 Effect of DSB repair factor knockdown (KD) on L1 retrotransposition and toxicity. HeLa cells were transfected with the indicated plasmids and split into subcultures three days later. Subcultures were selected either with hygromycin (Hyg) as control for transfection efficiency and toxicity or with G418 for retrotransposition events. Co-transfections of L1 reporter with empty vector controls were set as 100% and KD experiments were compared to their respective control. Relative colony numbers (%) are given. **A** Representative results performed with the indicated plasmids are shown. **B** Graphic representation of the results of the transfection and retrotransposition assays.

Plasmids co-transfected: pJM101/L1RP +	Transfection efficiency and toxicity (Hyg^r colony number)	Relative transfection efficiency and toxicity (%)	L1 activity (G418^r colony number)	Relative L1 activity (%)
pSUPER	122	100	157	100
pSUPER-p53	62	51	55	35
pSUPER-Ku70	0	0	0	0
pSHAG1	157	100	475	100
pSHAG-ATM	47	30	470	99
pSHAG-ATR	355	226	684	144
pSHAG-DNA PK	13	8	57	12

Table 4 Numbers of relative transfection efficiencies and toxicities and of relative L1 activities after DSB repair factor knockdown. Table with corresponding numbers of the transfection and retrotransposition assay results shown in **Figure 34**. Hyg^r, hygromycin resistant; G418^r, G418 resistant.

The co-transfection of pJM101/L1$_{RP}$ with pSUPER-p53 is obviously more toxic for the cells than the transfection of the L1 reporter plasmid with the corresponding empty vector pSUPER. This KD experiment yields 49% less Hyg resistant colonies than the control transfection with pSUPER and pJM101/L1$_{RP}$.

When selecting the sister subcultures with G418 for retrotransposition events, the KD experiment of p53 yields even 65% less G418 resistant colonies compared to the control transfection (Figure 34 and Table 4). If p53 depletion that was detectable after 2 days after transfection (Figure 29) had no effect on L1 retrotransposition, one would expect similar relative numbers of G418 and Hyg resistant colonies. Since in this case the G418 resistant colony numbers are reduced, this result points to a negative effect of p53 depletion on L1 retrotransposition.

Looking at the results with the Ku70 shRNA construct, one can only conclude that this seems too toxic to work with, since neither Hyg nor G418 resistant colonies can be detected (Figure 34). To exclude poor plasmid quality or a transfection mistake, e.g. not adding one or both plasmids, pSUPER-Ku70 was prepared again and this experiment was repeated yielding the same results (data not shown).

Knocking down ATM and introducing an active L1 EN into HeLa cells, results in cell death of about 70%. Only 30% of Hyg resistant colonies can be found after co-transfecting pSHAG-ATM and pJM101/L1$_{RP}$ compared to co-transfection of the empty vector pSHAG1 and pJM101/L1$_{RP}$ (Figure 34 and Table 4), although previous immunoblot analysis revealed a KD of ATM only on day 7 (Figure 31). However, after selection for retrotransposition events with G418 and quantification, I found that L1 retrotransposition frequency after ATM depletion had increased by 3-fold relative to co-transfection of pSHAG1 and pJM101/L1$_{RP}$ (Figure 34

and Table 4). If no effect on retrotransposition had occurred, only 30% of G418 colonies would be expected compared to the co-transfection of pSHAG1 and pJM101/$L1_{RP}$. Therefore, counting 99% of G418 resistant colonies after transfection of pSHAG-ATM and pJM101/$L1_{RP}$ means an increase in L1 retrotransposition by 3-fold.

Co-transfection of pSHAG-ATR and pJM101/$L1_{RP}$ and subsequent selection with Hyg as transfection and toxicity control resulted surprisingly in an increase of the number of Hyg resistant colonies up to 226% relative to the empty vector control pSHAG1 (Figure 34 and Table 4). This finding suggests either a better transfection efficiency or a reduced toxicity of L1 EN after the ATR KD that could be observed by immunoblot beginning 2 days after transfection and becoming stronger 6 days after transfection (Figure 32). If transfection efficiency and toxicity of L1 EN was reduced in the ATR KD experiment and if ATR KD had no effect on L1 retrotransposition, one would expect the same number of G418 resistant colonies after transfection with pSHAG-ATR and pJM101/$L1_{RP}$ compared to transfection with the empty vector control pSHAG1 and pJM101/$L1_{RP}$. However, the relative number of G418 resistant colonies after co-transfection of the KD construct and the L1 reporter is 144% compared to transfection with the empty vector control and the L1 reporter. This indicates that ATR KD actually inhibits L1 retrotransposition reducing it to 63% (Figure 34 and Table 4). Interestingly, this result is contrary to the result with ATM KD.

DNA-PK KD in the presence of the active L1 reporter element was highly toxic for most of the HeLa cells. Therefore, I was not able to determine any potential effect of the DNA PK KD on L1 retrotransposition. After Hyg selection of co-transfected cells with pSHAG-DNA PK and pJM101/$L1_{RP}$ only 8% of the colony numbers remain compared to the transfection experiment with empty vector control pSHAG1 and pJM101/$L1_{RP}$ (Figure 34 and Table 4). To exclude poor plasmid quality or a transfection mistake, pSHAG-DNA PK was prepared again and this experiment was repeated yielding the similar results (data not shown). Given this toxicity it is difficult to detect any changes in retrotransposition frequency, unless the level of retrotransposition would be strongly elevated.

Unfortunately, shRNA toxicity seems to be a known phenomenon and has been reported many times, e.g. (Martin et al. 2011) and (Baek et al. 2014). As the observed toxicity could be a result of synergistic effects of the respective DSB repair factor KD and L1 EN which was shown to cause DSBs (Gasior et al. 2006), I decided to also test the toxicity effects of the KD constructs with a non-functional L1 reporter (pJM101/L1.3H230A) carrying a mutation in its EN domain (H230A) that renders it inactive. The transfection efficiency and toxicity of the KD constructs for p53, Ku70, ATR and DNA-PK observed previously together with the

retrotransposition competent L1 reporter (Figure 34 and Table 4) does not seem to be influenced by the activity of L1 EN, as the co-transfection results with the inactive L1 reporter yielded similar Hyg resistant colony numbers (Figure 35 and Table 5). Interestingly, the KD of ATM appears less toxic in absence of the active L1 EN, as Hyg resistant colony numbers rise from 30% with pJM101/L1$_{RP}$ up to 80% with pJM101/L1.3H230A. This result points to an involvement of ATM in the repair of L1 induced DSBs.

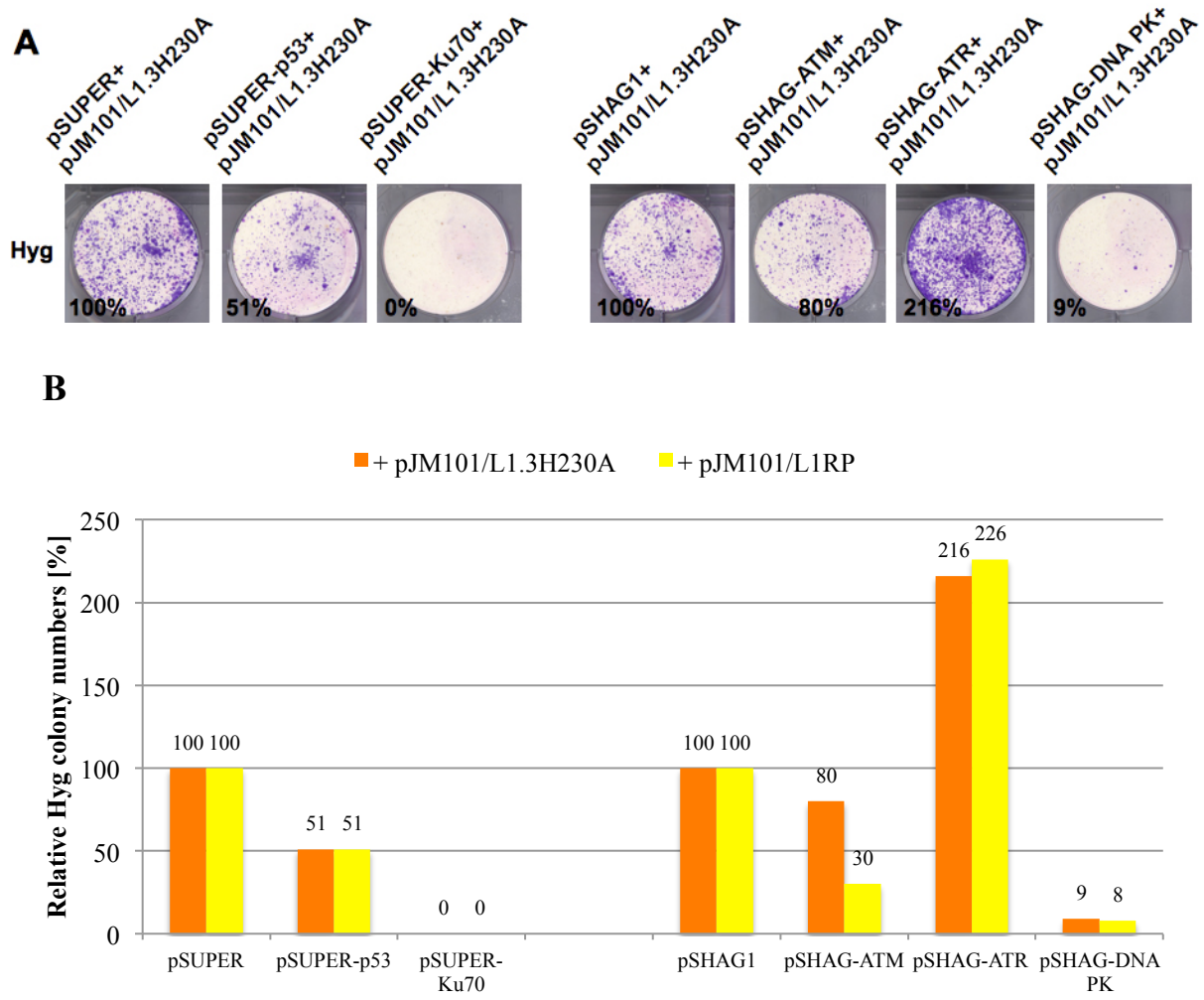


Figure 35 Transfection efficiency and toxicity of DSB repair factor knockdown (KD) with EN inactive L1 (pJM101/L1.3H230A). HeLa cells were co-transfected with the indicated plasmids and split into subcultures three days later. Subcultures were selected with Hygromycin (Hyg) as control for transfection efficiency and toxicity. Co-transfections of the L1 reporter with empty vector controls were set to 100% and KD experiments were compared to their respective control. Relative Hyg colony numbers (%) are given. **A** Representative results performed with the indicated plasmids are shown. **B** Graphic representation of transfection results of DSB repair factor KD with pJM101/L1.3H230A compared with transfection results of DSB repair factor KD with retrotransposition competent L1 (pJM101/L1$_{RP}$) from **Figure 34**.

Plasmids co-transfected	pJM101/L1$_{RP}$		pJM101/L1.3H230A	
	Transfection efficiency and toxicity (Hyg^r colony number)	Relative transfection efficiency and toxicity (%)	Transfection efficiency and toxicity (Hyg^r colony number)	Relative transfection efficiency and toxicity (%)
pSUPER	122	100	377	100
pSUPER-p53	62	51	192	51
pSUPER-Ku70	0	0	1	0
pSHAG1	157	100	357	100
pSHAG-ATM	47	30	286	80
pSHAG-ATR	355	226	771	216
pSHAG-DNA PK	13	8	32	9

Table 5 Numbers of relative transfection efficiencies and toxicities after DSB repair factor knockdown and introduction of an active or inactive L1 reporter. Table with corresponding numbers of the transfection assay results shown in **Figure 35**. Hyg^r, hygromycin resistant.

Immunoblot analyses presented in Figure 29 through Figure 33 show that knocking down the different DSB repair factors worked only for a short timespan (from 24 hours to 8 days) and therefore, the quantified L1 retrotransposition events occurred not exclusively in the absence of the respective repair factor. This is why I decided to evaluate the consequences of repair factor KD using the Dox-inducible L1 retrotransposition reporter ptetO7CMV/L1-3blas (Figure 21) whose expression period can be controlled. As the inducible L1 retrotransposition reporter assay can only be performed in cells expressing tTA or rtTA, HeLa M2 cells expressing rtTA (Hampf and Gossen 2007) were used for further studies. To check if the KD of the various DSB repair factors was possible in this cell line, I firstly analyzed the expression levels of each DSB repair factor at different time points after co-transfection of the respective shRNA expression plasmid with ptetO7CMV/L1-3blas by Western Blot analysis of total cell extracts (Figure 36 through Figure 39).

In order to knock down endogenous ATM expression, I firstly co-transfected pSHAG-ATM with ptetO7CMV/L1-3blas and used the co-transfection of pSHAG1 and ptetO7CMV/L1-3blas as negative control. Immunoblot analysis revealed the absence of any detectable ATM protein from HeLa M2 cells after 24 hours post co-transfection (Figure 36). Clearly, the observed reduction in ATM protein expression without Hyg selection cannot be the consequence of the expression of ATM-specific shRNA, because co-transfection of the negative control plasmid pSHAG1 has the same effect on ATM protein level as the co-transfection of the KD construct pSHAG-ATM. ATM depletion could be the result of cellular stress induced by the transfection procedure.

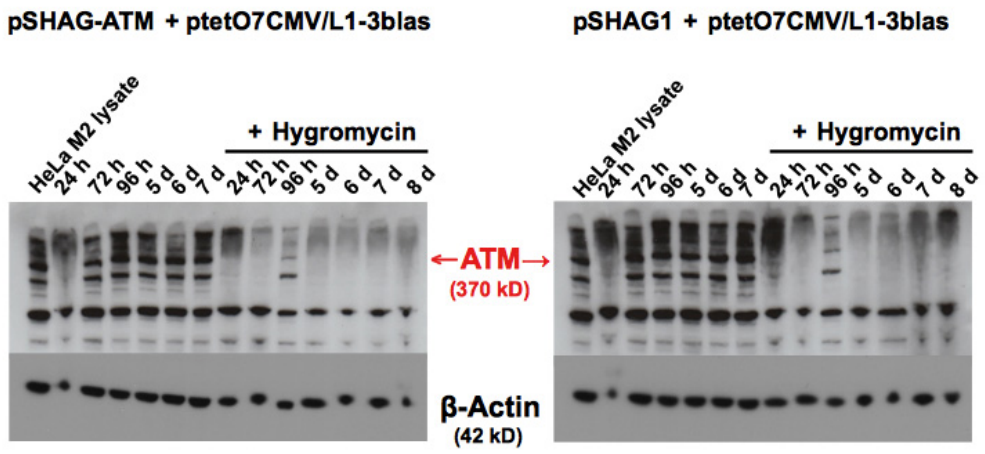


Figure 36 Western Blot showing the knockdown of ATM in HeLa M2 cells after co-transfection of pSHAG-ATM with ptetO7CMV/L1-3blas and pSHAG1 with ptetO7CMV/L1-3blas. In each lane 20 µg of total cell lysate were loaded. Cells were either treated with Hyg or left untreated, as indicated. Time points at which lysates were prepared post transfection are indicated. β-actin served as loading control. ATM reduction is detectable in both transfection experiments at 24 hours post transfection independent of Hyg treatment, and in both transfection experiments upon Hyg treatment throughout 8 days.

Figure 36 shows that during Hyg selection, cells are not only ATM-depleted after 24 hours post co-transfection, but ATM depletion extends throughout the following 8 days independent of co-transfection of the KD construct or the empty vector control. Only at 96 hours after co-transfection, a faint band at the size of ATM is visible. Therefore, I conclude that reduction of ATM expression after 72 hours after co-transfection is attributed to Hyg selection.

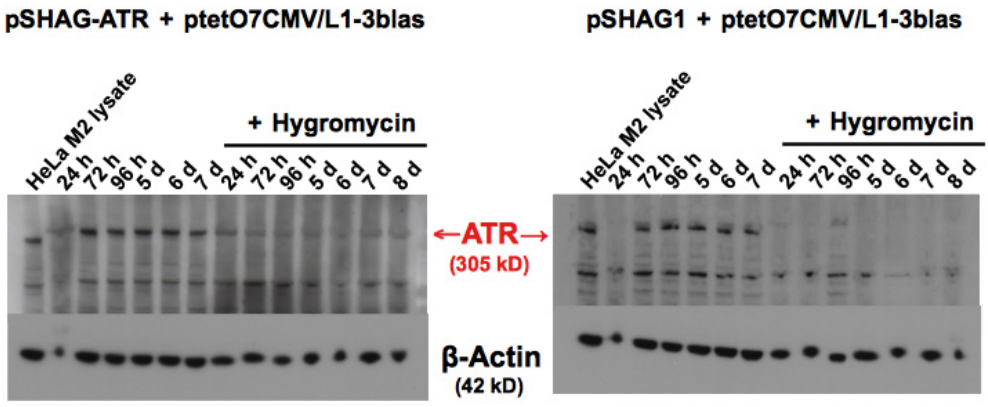


Figure 37 Western Blot showing the knockdown of ATR in HeLa M2 cells after co-transfection of pSHAG-ATR with ptetO7CMV/L1-3blas and pSHAG1 with ptetO7CMV/L1-3blas. In each lane 20 µg of total cell lysate were loaded. Cells were either treated with Hyg or left untreated, as indicated. Time points at which lysates were prepared post transfection are indicated. β-actin served as loading control. ATR reduction is detectable in both transfection experiments at 24 hours post transfection independent of Hyg treatment, and in both transfection experiments upon Hyg treatment for the entire 8 days of observation.

Next, I co-transfected pSHAG-ATR with ptetO7CMV/L1-3blas and used the co-transfection of pSHAG1 and ptetO7CMV/L1-3blas as negative control. Again, immunoblot analysis revealed a substantial reduction of detectable ATR protein from HeLa M2 cells after 24 hours post co-transfection (Figure 37). As for ATM, the Western Blot analysis shows most efficient reductions in the ATR level of M2 HeLa cells upon co-transfections of pSHAG-ATR and ptetO7CMV/L1-3blas or pSHAG1 and ptetO7CMV/L1-3blas, respectively (Figure 37). As described for ATM before, endogenous ATR protein is undetectable 24 hours post co-transfection and drastically depleted during the time course after 24 hours post co-transfection in the presence of Hyg, no matter if ATR-specific shRNA was expressed or not.

Next, the consequences of co-transfection of the DNA-PK specific shRNA expression construct pSHAG-DNA PK on endogenous DNA PK expression was analyzed. The results are shown in Figure 38.

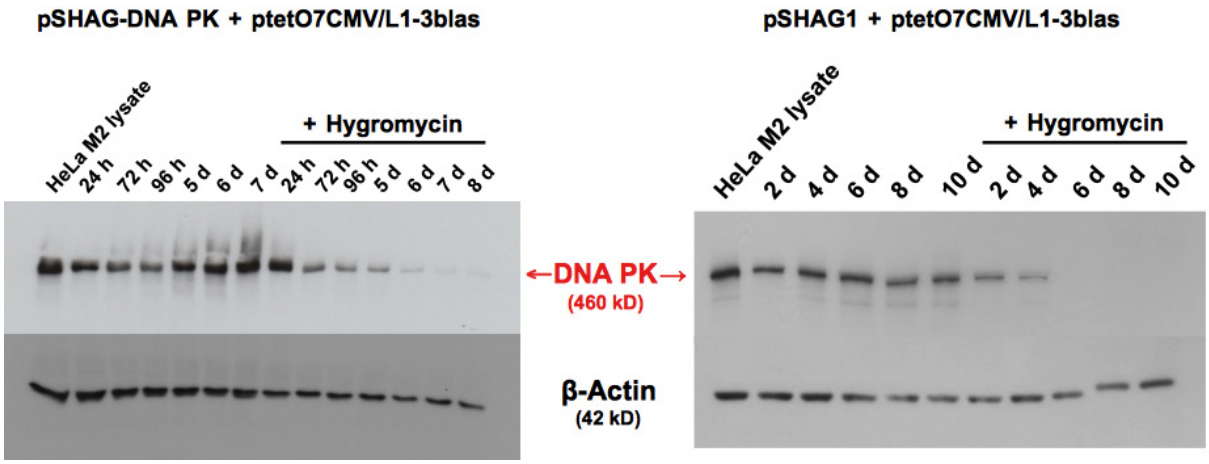


Figure 38 Western Blot showing the knockdown of DNA-PK in HeLa M2 cells after co-transfection of pSHAG-DNA PK with ptetO7CMV/L1-3blas and pSHAG1 with ptetO7CMV/L1-3blas. In each lane 15 µg of total cell lysate were loaded. Cells were either treated with Hyg or left untreated, as indicated. Time points at which lysates were prepared post transfection are indicated. β-actin served as loading control. DNA-PK reduction is detectable in both transfection experiments only upon Hyg treatment.

Immunoblot analyses of HeLa M2 cells co-transfected with pSHAG-DNA PK and ptetO7CMV/L1-3blas or with pSHAG1 and ptetO7CMV/L1-3blas show that DNA-PK expression is drastically depleted after 72 hours or 2 days following transfection, respectively, if Hyg is present. This means that in HeLa M2 cells the DNA-PK specific shRNA does not cause any DNA-PK reduction. However, the presence of Hyg causes also in this case a massive reduction of endogenous DSB repair factor expression.

Then, I tested how efficiently expression of the shRNA construct pSUPER-p53 reduced endogenous p53 levels after co-transfection of pSUPER-p53 and ptetO7CMV/L1-3blas in HeLa M2 cells by performing immunoblot analyses (Figure 39). Although I observed a drastic depletion of p53 protein levels in the presence of p53-specific shRNA, I observed a similar reduction of p53 expression after co-transfection of the L1 reporter with the empty vector control pSUPER. In these experiments, p53 protein levels are independent of presence or absence of Hyg.

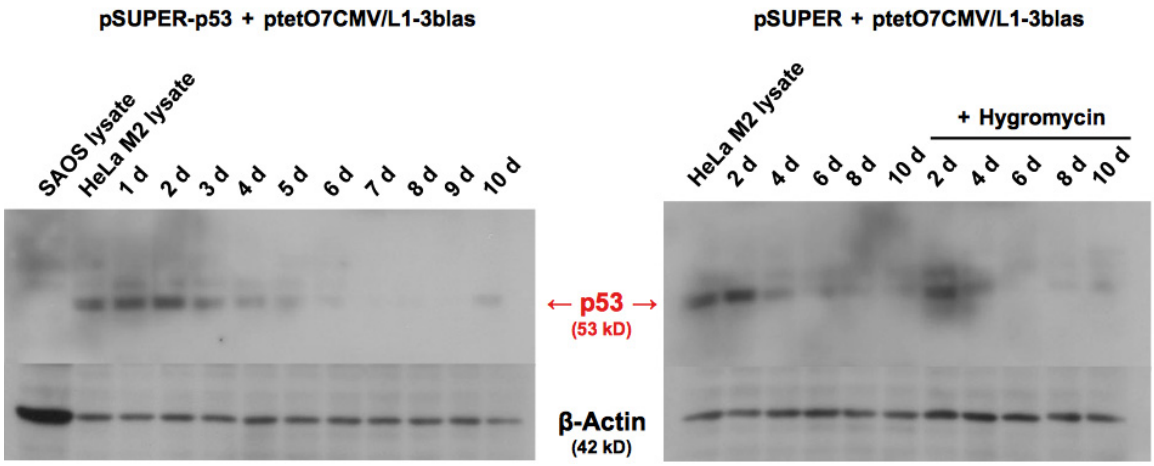


Figure 39 Western Blot showing the knockdown of p53 in HeLa M2 cells after co-transfection of pSUPER-p53 with ptetO7CMV/L1-3blas and pSUPER with ptetO7CMV/L1-3blas. In each lane 15 µg of total cell lysate were loaded. Cells were either treated with Hyg or left untreated, as indicated. Time points at which lysates were prepared post transfection are indicated. Lysate from SAOS cells that don't express p53 was used as negative control. β-actin served as loading control. p53 reduction is detectable not only after transfection of the respective shRNA construct, but also upon transfection of the empty vector control.

Taken together, the experimental system used did unfortunately not allow for expression of L1 reporter constructs exclusively during a defined time period in which the absence of any of the listed DNA repair factors could be guaranteed. Absence or at least drastic depletion of the respective repair factor during the time period tagged L1 retrotransposition events take place, is an essential prerequisite, if the effects of these repair factors on L1 retrotransposition are to be investigated. Therefore, the described experimental setup did not allow the analysis of the roles of these host factors on L1 retrotransposition. As a consequence, I was not able to quantify or characterize L1 *de novo* insertions that occurred in the absence of any of the described repair factors as originally planned. Data also suggest that the presence of hygromycin per se downregulates endogenous expression of ATM, ATR and DNA-PK, no matter if any specific shRNA is present or not.

3.3.2 COMPETITIVE INHIBITION OF DSB REPAIR FACTORS VIA EXPRESSION OF DOMINANT-NEGATIVE MUTANTS AND INFLUENCE ON L1 RETROTRANSPOSITION

In this study, overexpression of the DNA-binding domain of PARP (PARP DBD) by pPARP6 was used to inhibit the endogenous PARP protein (Kupper et al. 1990) (Figure 40). For overexpression of wildtype PARP, the same plasmid was applied carrying the coding sequence of full-length PARP instead of the PARP DBD (pPARP31) (Kupper et al. 1990).

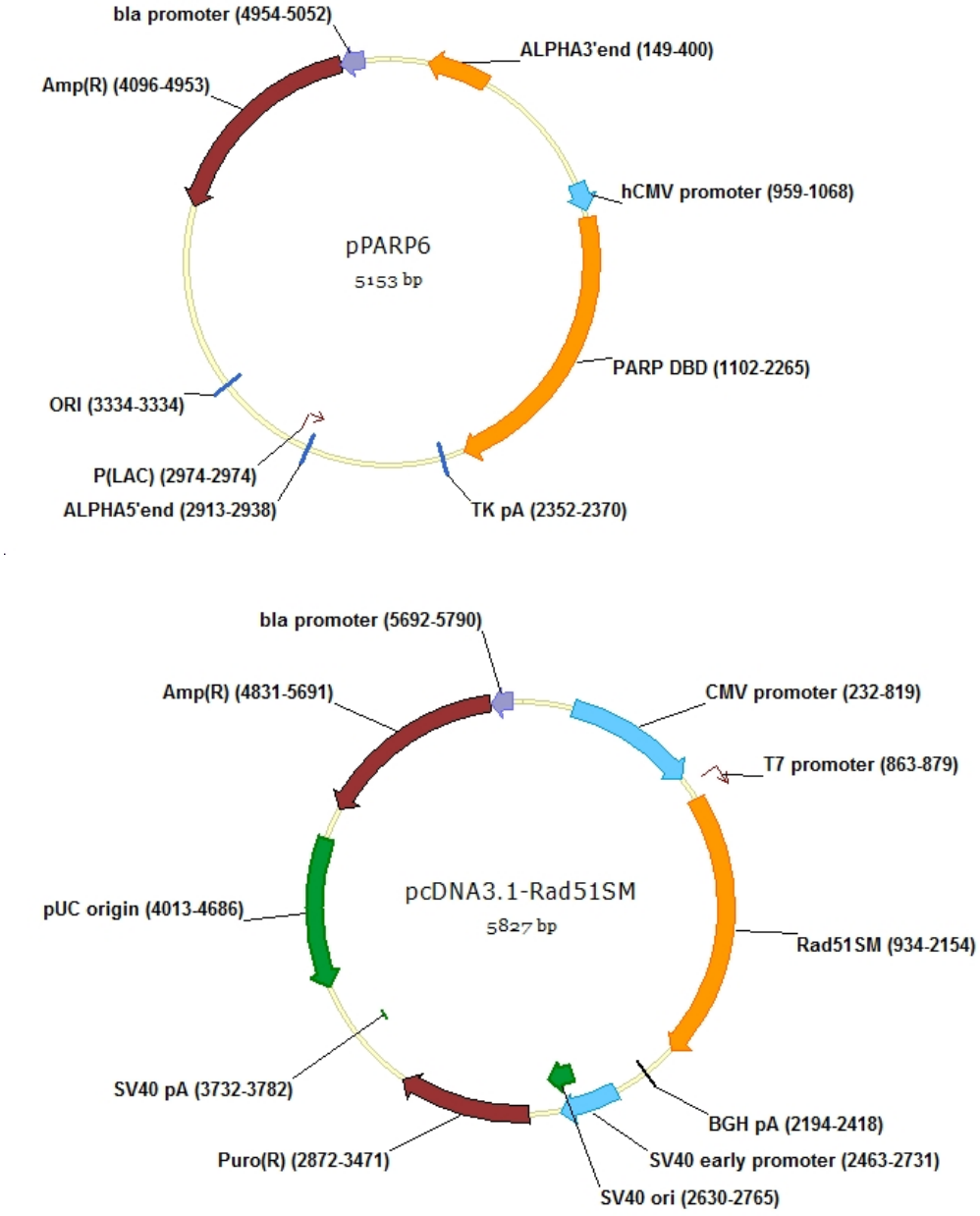


Figure 40 Dominant-negative protein expression vectors used in this study. pPARP6 (top; Kupper et al. 1990) expresses the DNA-binding domain of PARP (PARP DBD), while pcDNA3.1-Rad51SM (bottom; Lambert and Lopez 2000) carries a yeast-mouse chimera of the protein (Rad51SM).

Interference with endogenous Rad51 was achieved with a yeast-mouse chimera of the protein (Rad51SM) encoded by pcDNA3.1-Rad51SM (Lambert and Lopez 2000) (Figure 40). To overexpress wildtype Rad51, pcDNA3.1-Rad51Mm was used in which the Rad51SM sequence was replaced by the coding sequence of murine Rad51 (Rad51Mm) (Lambert and Lopez 2000).

To test the effect of the expression of both dominant negative (DN) DSB repair factor mutants and the respective wildtype proteins on L1 retrotransposition, expression constructs of the DN mutants and of the wildtype proteins were separately co-transfected with pJM101/L1$_{RP}$ into HeLa cells. In order to test for transfection efficiency and toxicity of the respective expression construct, each well containing transfected cells was split into two subcultures three days post transfection and one resulting subculture was subjected to Hyg selection. To determine retrotransposition frequency in the presence of the respective DN mutants, the second subculture was submitted to G418 selection. Before analyzing retrotransposition frequencies and toxicity of each transfection experiment, I checked the expression levels of both DN mutants and wildtype proteins at different time points after co-transfection by Western Blot analyses of total cell extracts from the Hyg selected cells.

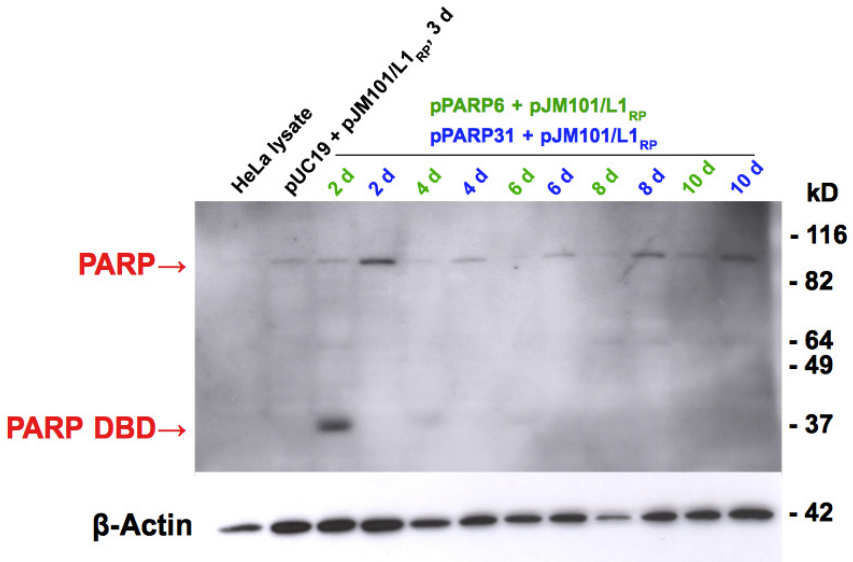


Figure 41 Western Blot showing the expression of dominant negative PARP DBD and wildtype PARP after co-transfection of pPARP6 or pPARP31 with pJM101/L1$_{RP}$. In each lane 10 µg of total cell lysate were loaded. Time points at which lysates were prepared post transfection are indicated. Lysate from untransfected HeLa cells was used as control. β-actin served as loading control. Expression of PARP DBD is detectable at two and four days post transfection, while PARP can be detected from day 2 until day 10.

Immunoblot analysis with an anti-PARP antibody revealed that the expression of PARP and PARP DBD was detectable at day 2 after transfection (Figure 41). At day 4, only a faint band for PARP DBD remains and at days 6, 8 and 10 no expression can be visualized. In the case of wildtype PARP, weak expression from day 4 until day 10 can still be detected (Figure 41). Immunoblot analysis of cell extracts from HeLa cells co-transfected with pcDNA3.1-Rad51SM or pcDNA3.1-Rad51Mm and pJM101/L1$_{RP}$ with an anti-Rad51 antibody uncovered that DN Rad51 and wildtype Rad51 were stably expressed from day 1 until day 9 after transfection (Figure 42).

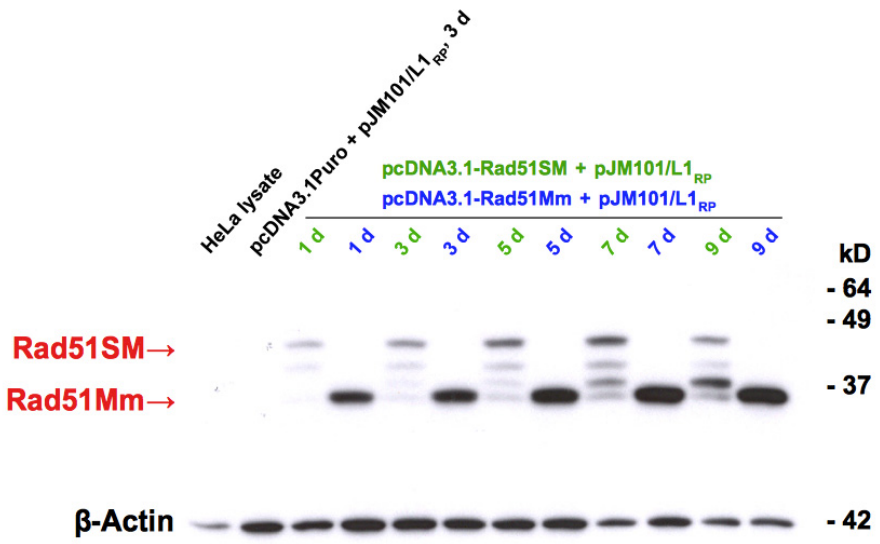


Figure 42 Western Blot showing the expression of dominant negative Rad51SM and wildtype Rad51Mm after co-transfection of pcDNA3.1-Rad51SM or pcDNA3.1-Rad51Mm with pJM101/L1$_{RP}$. In each lane 10 µg of total cell lysate were loaded. Time points at which lysates were prepared post transfection are indicated. Lysate from untransfected HeLa cells was used as control. β-actin served as loading control. Expression of Rad51SM and Rad51Mm is detectable continuously starting on day 1 after co-transfection.

The obtained colony numbers resulting from co-transfection of the L1 reporter pJM101/L1$_{RP}$ and the different DN and wildtype protein expression constructs, were compared with co-transfection experiments with the L1 reporter plasmid and the corresponding empty vectors (Figure 43 and Table 6). Transfection efficieny and toxicity evaluation by Hyg selection yielded reduced colony numbers for all of the different expression constructs. The co-transfection of pJM101/L1$_{RP}$ with either the plasmid carrying the DN mutant of PARP (pPARP6) or the PARP wildtype protein (pPARP31) is obviously less efficient or more toxic to the cells as the co-transfection of the L1 reporter plasmid together with the corresponding empty vector (pUC19) (Figure 43 and Table 6).

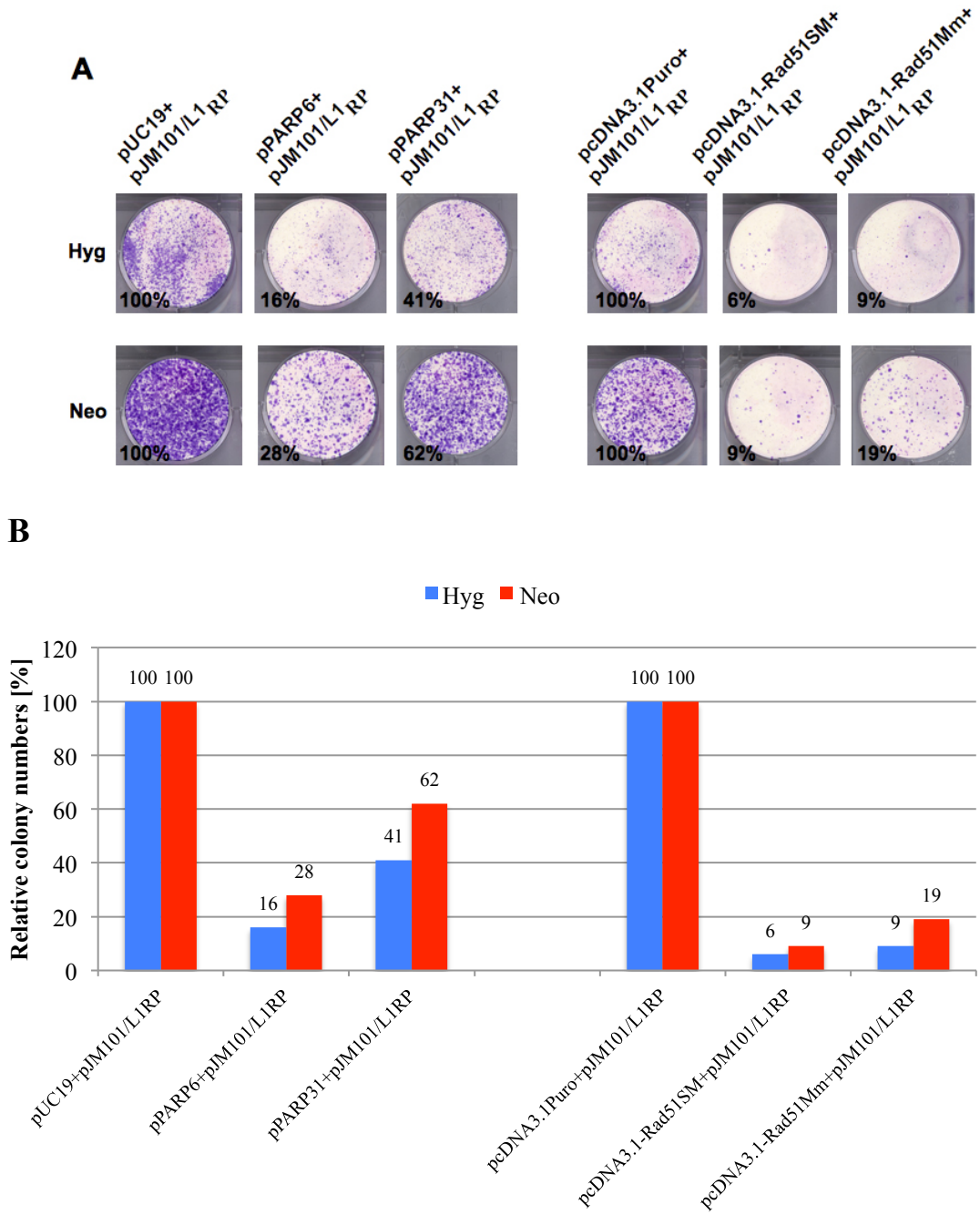


Figure 43 Effect of PARP and Rad51 inhibition with dominant negative mutants and overexpression of wildtype PARP and Rad51 on L1 retrotransposition and toxicity. HeLa cells were transfected with the indicated plasmids and split into subcultures three days later. Subcultures were selected either with hygromycin (Hyg) as control for transfection efficiency and toxicity or with G418 for retrotransposition events. Numbers of Hyg resistant colonies resulting from co-transfections of L1 reporter with empty vector controls were set as 100% and inhibition as well as overexpression experiments were compared to their respective control. Relative colony numbers (%) are given. **A** Representative results performed with the indicated plasmids are shown. **B** Graphic representation of the results of the transfection and retrotransposition assays.

Plasmids co-transfected: pJM101/L1RP +	Transfection efficiency and toxicity (Hyg^r colony number)	Relative transfection efficiency and toxicity (%)	L1 activity ($G418^r$ colony number)	Relative L1 activity (%)
pUC19	850	100	1043	100
pPARP6	136	16	292	28
pPARP31	349	41	647	62
pcDNA3.1Puro	213	100	478	100
pcDNA3.1-Rad51SM	13	6	43	9
pcDNARad51Mm	19	9	91	19

Table 6 Numbers of relative transfection efficiencies and toxicities and of relative L1 activities after expression of dominant negative mutants and overexpression of wildtype PARP and Rad51. Table with corresponding numbers of the transfection and retrotransposition assay results shown in **Figure 43**. Hyg^r, hygromycin resistant; $G418^r$, G418 resistant.

The expression of PARP6 yielded only 16% and the overexpression of PARP31 results in 41% of Hyg resistant colonies relative to the control transfection with pUC19 and pJM101/$L1_{RP}$, respectively (Figure 43 and Table 6).

In the case of co-transfection of the Rad51 plasmids with the L1 reporter, expression of the DN mutant (Rad51SM) as well as the wildtype protein (Rad51Mm) yielded reduced Hyg resistant colony numbers, 6% and 9%, respectively, compared to the empty vector control (pcDNA3.1Puro) (Figure 43 and Table 6). However, a specific mode of toxicity is so far unknown for DN mutants, wildtype PARP and Rad51. Therefore, the reduced colony numbers may also be a result of reduced transfection efficiency.

When selecting the sister subcultures with G418 for retrotransposition events, the PARP6 and PARP31 expression experiments yielded less G418 resistant colonies, 28% and 62%, respectively, than the control co-transfection of empty vector with L1 reporter plasmid (100%) (Figure 43 and Table 6). Since these G418 resistant colony numbers are slightly higher than the respective Hyg resistant colony numbers (16% and 41%, respectively), this result points towards an L1 retrotransposition supporting effect of both PARP6 and PARP31. This means that either change in endogenous PARP function (inhibition by DN mutant or overexpression of wildtype protein) seems to facilitate L1 retrotransposition.

The co-expression of the DN Rad51SM and the L1 reporter results in a number of G418 resistant colonies (9%) that is in the same range as the numbers of Hyg resistant colonies (6%) from the same co-transfection experiment (Figure 43 and Table 6). This result implies that the DN mutant of Rad51 does not interfere with L1 retrotransposition. However, when overexpressing wildtype Rad51 (Rad51Mm) together with an active L1 reporter

(pJM101/L1$_{RP}$), an increase in G418 resistant colony numbers (19%) with respect to Hyg resistant colony numbers (9%) of the same co-transfection assay is detectable. This means that the presence of Rad51Mm enhances L1 retrotransposition activity (Figure 43 and Table 6). This result argues for an involvement of Rad51 in L1 retrotransposition. As Rad51 is a protein mediating strand exchange in DSB repair, it is conceivable that it is also involved in strand exchange during TPRT.

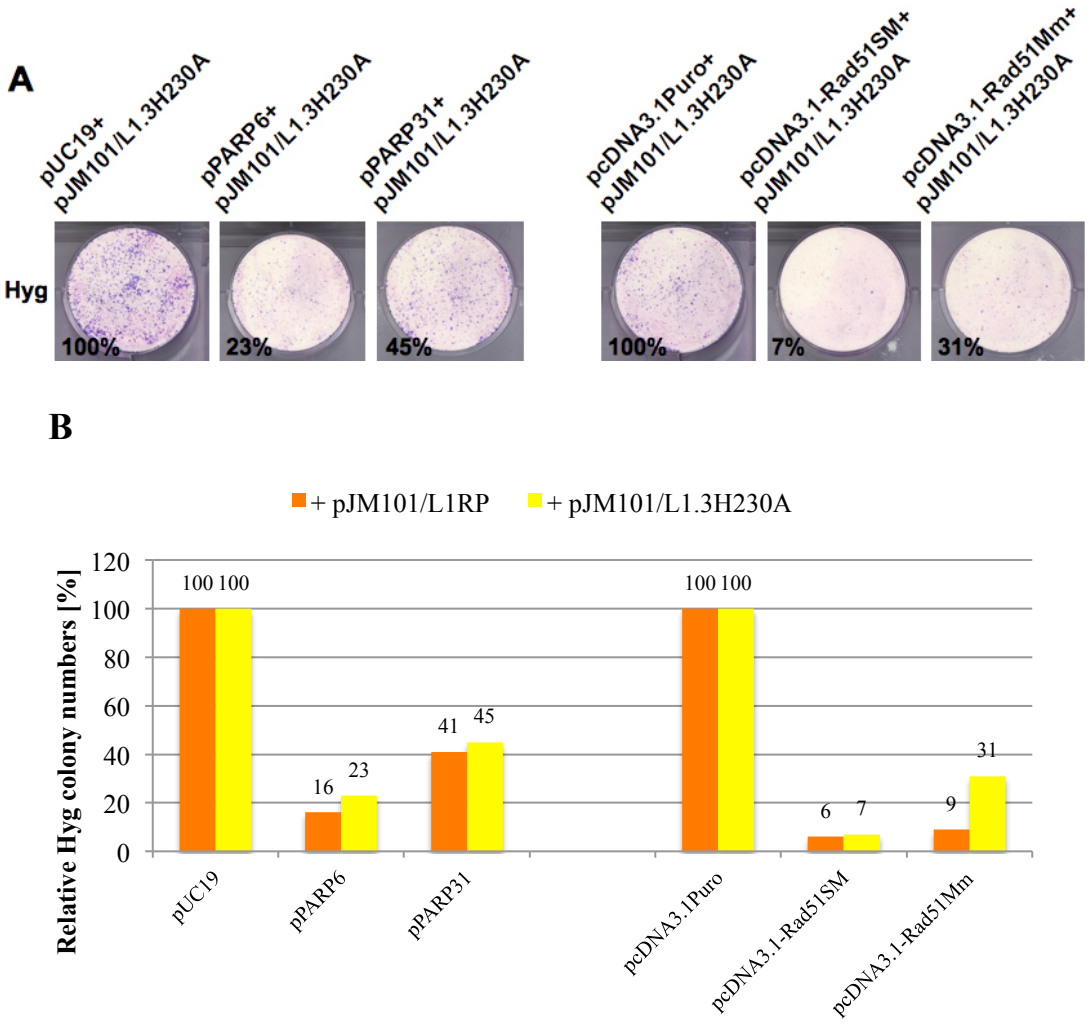


Figure 44 Transfection efficiency and toxicity resulting from PARP and Rad51 inhibition or overexpression and co-expression of an EN inactive L1 reporter encoded by pJM101/L1.3H230A. HeLa cells were co-transfected with the indicated plasmids and split into subcultures three days later. Subcultures were selected with hygromycin (Hyg) to determine transfection efficiency and toxicity. Numbers of Hyg resistant colonies resulting from co-transfections of the L1 reporter with empty vector controls were set as 100% and inhibition as well as overexpression experiments were compared to their respective control. Relative Hyg colony numbers (%) are given. **A** Representative results performed with the indicated plasmids are shown. **B** Graphic representation of transfection results of DN or wildtype PARP and Rad51 constructs with pJM101/L1.3H230A compared with transfection results of DN or wildtype PARP and Rad51 with retrotransposition competent L1 (pJM101/L1$_{RP}$) from **Figure 43**.

Plasmids co-transfected	pJM101/L1$_{RP}$		pJM101/L1.3H230A	
	Transfection efficiency and toxicity (Hyg^r colony number)	Relative transfection efficiency and toxicity (%)	Transfection efficiency and toxicity (Hyg^r colony number)	Relative transfection efficiency and toxicity (%)
pUC19	850	100	440	100
pPARP6	136	16	101	23
pPARP31	349	41	198	45
pcDNA3.1Puro	213	100	128	100
pcDNA3.1-Rad51SM	13	6	9	7
pcDNA3.1-Rad51Mm	19	9	40	31

Table 7 Numbers of relative transfection efficiencies and toxicities after PARP and Rad51 inhibition or overexpression and introduction of an active or inactive L1 reporter. Table with corresponding numbers of the transfection assay results shown in **Figure 44**. Hyg^r, hygromycin resistant.

Just as for the shRNA KD constructs, toxicity could be a result of synergistic effects of the overexpression of the respective DSB repair factor or its DN mutant and the L1 EN. For this reason, I tested the potential toxicity of the expression constructs co-transfected with a non-functional L1 reporter (pJM101/L1.3H230A) carrying a mutation in its EN domain (H230A) that renders it inactive. The transfection efficiency and toxicity of the constructs for PARP6, PARP31 and Rad51SM did not seem to be influenced particularly by the activity of L1 EN, because co-transfection with the inactive L1 EN mutant yielded very similar Hyg resistant colony numbers (Figure 44 and Table 7). Interestingly, the expression of Rad51Mm appears less toxic in absence of an active L1 EN, as colony numbers rise from 9% with retrotransposition competent L1 reporter up to 31% with EN inactive L1 reporter. This result suggests an involvement of Rad51 in the repair of L1 induced DSBs.

3.4 DETERMINANTS FOR DNA TARGET STRUCTURE SELECTIVITY OF THE HUMAN LINE-1 ENDONUCLEASE

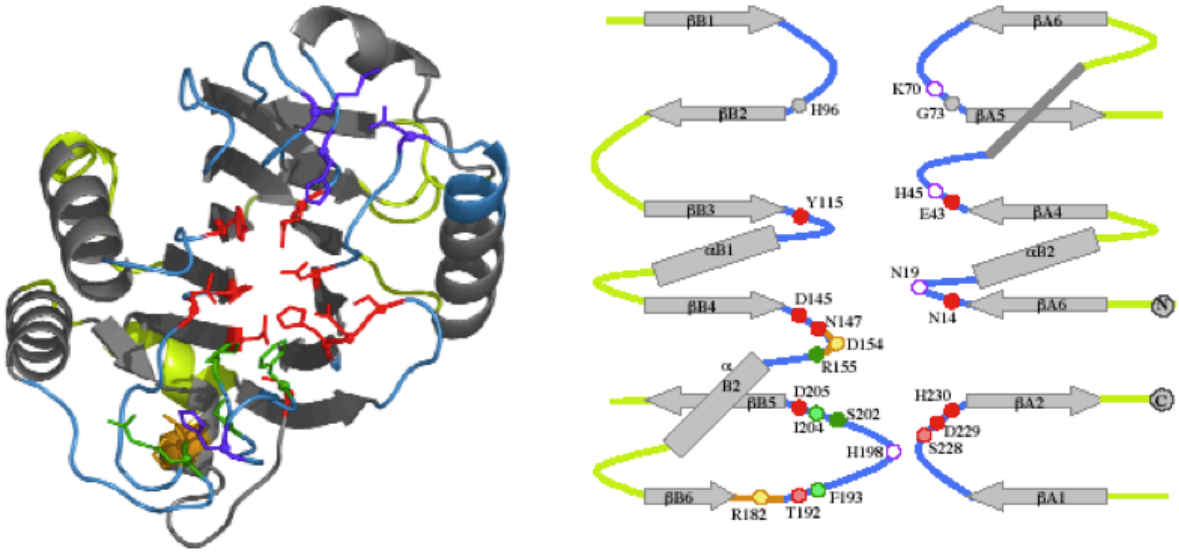
The endonuclease (EN) domain is the primary determinant of genomic target site selection during L1 integration (Cost and Boeke 1998). *In vitro* assays showed that the specificity of purified L1 EN for the 5'-TT/AAAA-3' consensus sequence (Feng et al. 1996; Cost and Boeke 1998; Cost et al. 2001) resembles the genomic integration sites of *de novo* L1 insertions *in vivo* (Symer et al. 2002; Gilbert et al. 2002). The consensus target sequence was also confirmed by *in silico* analyses of the integration sites of pre-existing L1 as well as *Alu* element insertions in the human genome (Jurka 1997; Szak et al. 2002).

Modulation of cleavage specificity of AP-like ENs is thought to be achieved mainly via variations in the surface loops contacting the DNA. Transplant experiments with ExoIII and DNaseI as well as swapping experiments with L1 EN βB6-βB5 loop and the corresponding loop in Tx1L EN support this hypothesis. In the first case, DNaseI could be converted from an unspecific endonuclease to a nicking enzyme with high selectivity for abasic sites by grafting a prominent α-helix from the AP-site-specific nuclease ExoIII onto the DNA binding surface of DNaseI (Cal et al. 1998). In the second case, a deviation from the L1 consensus was observed exchanging the βB6-βB5 loop of L1 EN by the corresponding surface loop of Tx1L EN: In 36% of the integration sites, the usually strongly conserved purine located directly upstream of the cleavage site was replaced by a cytidine (TT/AAAA→TG/AAAA). As a 5'-GAAAA-3' sequence cannot be bent by the L1 EN into the typical DNA structure occurring at cleavage, this represents a considerable alteration in sequence specificity of the hybrid EN (Nora Zingler, PhD Thesis 2004; Repanas et al. 2007).

The ability to generate sequence- or site-specific L1 retrotransposons would be a crucial step in converting non-LTR retrotransposons into genetic tools. To this end, crystal structure guided mutational analysis exchanging single amino acids within the endonuclease domain of L1 was performed in this study to identify residues influencing target site recognition.

Variants of L1 EN falling into two categories were designed (Figure 45). The first category comprises point mutations of catalytic and structurally important amino acids (T192V, H230A) that are highly conserved within the family of AP-like endonucleases. The second category consists of point mutants of moderately conserved, non-catalytic surface residues (R155A, S202A, I204Y) expected to affect the accommodation and recognition of the nucleotide downstream of the scissile bond.

A



B

Mutation in endonuclease	Proposed function of original residue	Expected result
S202A	H-bond to flipped adenine of target site consensus sequence AATTTT	Change of consensus, possibly change in retrotransposition rate
R155A	H-bond to flipped adenine of target site consensus sequence AATTTT	Change of consensus, possibly change in retrotransposition rate
T192V	Target site cleavage and/or anchores α11 loop	Non-functional endonuclease
I204Y	Provides space for flipped adenine	Change of consensus, non-functional endonuclease
H230A	Target site cleavage (confirmed)	Non-functional endonuclease as negative control

Figure 45 Structural details and topology of the targeting L1 EN and L1 EN mutations. A Ribbon diagram and idealized topology diagram of L1 EN. Residues conserved among AP endonucleases that are catalytically (filled) or structurally (half-filled) important shown in red; residues proposed to recognize extra-helical nucleotides via the base (filled) and the ribose (half-filled) shown in green; putative peripheral DNA binding residues shown in purple; salt-bridge restricted to AP DNA repair endonucleases and mammalian-type L1 endonucleases shown in orange (Weichenrieder et al. 2004). **B** L1 EN mutations with proposed function of the original residue and expected consequence of the point mutation. Newly designed L1 EN variants shown in black; previously published L1 EN variant (Feng et al. 1996) shown in green.

3.4.1 L1 EN POINT MUTATIONS AFFECT RETROTRANSPOSITION IN HELA CELLS

The L1 EN variants were tested by applying the L1 retrotransposition reporter assay described in Chapter 2.6.3 using plasmids pJM101/L1.3 (Figure 15, top), pJM101/L1.3S202A, pJM101/L1.3R155A, pJM101/L1.3T192V, pJM101/L1.3I204Y and pJM101/L1.3H230A (Chapter 2.5.4). Retrotransposition frequencies of the L1 reporter constructs carrying the different mutations in L1 EN were quantified by counting G418 resistant HeLa colonies. The results of the L1 retrotransposition assays are summarized in Table 8 and were published in *Nucleic Acids Research* (Repanas et al. 2007). Each of the tested mutations in L1 EN affected L1 retrotransposition frequencies negatively, confirming the relevance of the mutated residues.

Endonuclease mutant	L1 activity (colony number after Neo selection)	Relative L1 activity [%]
wt	125±21	100±17.1
S202A	40±10	32±7.8
R155A	15±4	12±3.3
T192V	6±4	5±3.0
I204Y	1±1	1±1.1
H230A	0	0

Table 8 Effects of point mutations in the L1 EN coding region on L1 retrotransposition frequency. The retrotransposition frequency of the L1 wildtype (wt) reporter element was set as 100% and the L1 EN point mutated reporter were compared to it. Representative results are shown. All EN variants displayed a reduced retrotransposition frequency. n=9

To test whether the point mutations had an effect on L1 target site selection, the genomic flanks of the newly integrated L1 elements were analyzed using a previously published plasmid rescue procedure (Symer et al. 2002; Gilbert et al. 2002) that allows direct cloning of *de novo* L1 integrants together with their flanking genomic DNA in bacteria (Figure 19).

For that purpose, plasmids rescue vectors pCEP4/L1.3/ColE1/mneoI400 (Figure 15, bottom), pColE1-S202A, pColE1-R155A, pColE1-T192V and pColE1-I204Y (Chapter 2.5.4) were used. After isolation of the rescue plasmids from the bacteria, sequencing was performed to analyze the genomic flanks. Unfortunately, direct sequencing of the L1 3' junctions with an mneoI/ColE1 binding primer was not possible, most likely due to polymerase slippage on the polyA tail. Therefore, the sequence located 5' of the L1 fragment was identified first. As L1 integrants are variably truncated at their 5' ends, the individual length of the L1 copy was determined using a multiplex PCR with the antisense primer GS88 binding to the 5' end of the *neo* resistance cassette and a set of sense primers (GS260, GS261 and GS262) covering the L1 sequence downstream of the *HindIII* site (Figure 46). Depending on the length of the generated PCR fragments, the length of the L1 integration could be deduced. Then, using a suitable antisense primer (GS88, GS17, GS14, GS16 or GS76), the 5' junction of the new integrant was sequenced.

Figure 46 Schematic representation of the binding sites of primers involved in the isolation of 5' and 3' junctions of *de novo* L1 integrants.

Most sequencing reactions with antisense primers did not extend far enough on the rescue plasmid to sequence the 3' junction between the L1 copy and the genomic DNA as well. Therefore, the obtained genomic flanking sequences 5' of the L1 element were used as probes in BLAT searches in the human genome sequence available at http://genome.cse.ucsc.edu/cgi-bin/hgBlat. The queries matched unique sequences present in the human genome with >99% identity, allowing to localize the genomic position of the new integrant. Then, to be able to sequence the junction of the L1 3' end, primers were designed that bind ~200 bp downstream of the presumed integration site. In cases where the integrations extended upstream of the *HindIII* site located within L1, another genomic primer was designed ~200 bp upstream of the presumed integration site and used in conjunction with GS76 to amplify the 5' flanking sequence by PCR from genomic DNA of HeLa cells harboring the relevant insertions. In this way, 35 L1 *de novo* integration events were characterized (Table 9). In all cases, the 3' flanking sequences were successfully recovered revealing the genomic integration site, while

in 23 clones both of the flanks could be characterized yielding also the size of the TSD and the number of nucleotides encompassing the microhomologies.

Genomic target sequences of L1 retrotransposition events launched from L1 reporter constructs encoding EN mutants S202A, R155A, T192V and I204Y were analyzed by aligning the upper strand sequences of the genomic target sites. Since the TPRT mechanism initiates the cDNA synthesis of L1, the sequence at the 3' junction is sufficient to identify the nicking site of the EN. Results were summarized in sequence logos (Crooks et al. 2004) identifying consensus target sequences of the different L1 ENs and compared to the consensus target sequence of wildtype L1 EN (Figure). I aligned 5 to 15 individual genomic target sequences per L1 EN mutant. Although it is questionable, if it is possible to draw conclusions from this relatively small number of target sequences, it is still remarkable that in the case of the S202A mutant, there is a strongly reduced preference for an A at the third position after the cleavage site, and that at the first position, the T of the wild type target site consensus (TT AAAA) was replaced by an A (AT AAAA). However, an A at that position may also occur in the wildtype target site as indicated in the top sequence logo, yet less frequently as in the case of the S202A variant. Genomic target sequences of the remaining L1 EN mutants R155A, T192V and I204Y did not differ strikingly from the consensus target sequence of the wildtype L1 EN. This indicates that these point mutations are clearly not sufficient to modify EN specificity remarkably, but rather result in a reduction of endonucleolytic activity reflected by a reduced transposition frequency (Table 9).

Table 9 Insertions from L1 elements with mutated endonucleases (on the following page). The data was collected from rescue plasmids generated as described in the text above. In all cases the 3' flanking sequences were recovered revealing the integration site (Putative EN nicking site), while in 23 clones both of the flanks could be characterized yielding also the size of the TSD (TSD length (bp)) and the number of bp of microhomologies (Microhomologies (bp)). In some cases there could be found genomic deletions, inversions and insertions of untemplated nucleotides at the 5' junctions (Remarks). n.d., not determined.

EN variant/rescue plasmid name	Putative EN nicking site	Remarks	TSD length (bp)	Microhomologies (bp)
S202A				
rp1.1	AT AAGA		n.d.	n.d.
rp1.2	AT AAGA		n.d.	n.d.
rp1.3	TT AAAG		47	6
rp1.4	AT ATTA		25	3
rp1.5	TT AAGA		n.d.	n.d.
rp1.6	TT CAAA	genomic deletion of 3 bp		2
rp1.8	TC AATA		2	3
rp1.9	AT AACA		22	3
rp1b	CT AAAA		16	1
rp1d	AG AGAA		1	2
rp1k	TT AGTT		6	0
rpNZ1	TT AAAA	inversion, genomic deletion of 4287 bp		6
rpNZ2	CG GAAA		14	2
rpNZ5	AT AACT	untemplated insertion of 3 bp	8	0
rpNZ9	AT AACT		n.d.	n.d.
R155A				
rpII.1	TT AAAA		n.d.	n.d.
rpII.15	AT GAAA	genomic deletion of 2 bp	206	1
rpII.17	TT AAAA		n.d.	n.d.
rpII.20	TT TAAA		n.d.	n.d.
rpII.21	TT AAAA		n.d.	n.d.
rpNZ13	AT AGGA		n.d.	n.d.
T192V				
rp3.1	TT AAGC	genomic deletion of 7 bp, untemplated insertion of 6 bp		0
rp3.8	TT AAAA	untemplated insertion of 2 bp	1	0
rp3.12	TT AAAA		3	1
rp3f	AT GAAA		1	3
rpNZ17	TT AAGA		15	2
I204Y				
rp4.1	GC AAAG		4	4
rp4.4	TC AAAA		91	1
rp.4.12	GC AAAG		33	0
rpIV.2	TT AACA	genomic deletion of 52 bp		1
rpIV.4	TT AAAA		51	1
rpIV.20	TT AAGA	untemplated insertion of 3 bp	269	5
rpIV.24	TC AAAA		n.d.	n.d.
rpIV.27	CT AGAA		n.d.	n.d.
rpIV.36	CT AAGA		n.d.	n.d.

As reported earlier (Chapter 1.4, Figure 8 and Figure 9), L1 retrotransposition events can cause genomic deletions and inversions at the L1 integration site, and are a consequence of the mechanism of L1 retrotransposition. During the analysis of *de novo* integration events, five genomic deletions in all of the four L1 EN variants driven insertions could be identified and one inversion in an S202A mutant generated integration (Table 9).

However, in four cases of S202A, T192V and I204Y L1 variants, also so-called "untemplated nucleotides" were found. These nucleotides appear at the 5' junctions between the L1 elements and the genomic DNA and their origin is unclear, as they cannot be assigned to neither the genomic sequence surrounding the integration site nor the L1 sequence. Such untemplated insertions were already described in earlier studies (Nora Zingler, PhD Thesis 2004; (Symer et al. 2002) and it was suggested to use the term "extra nucleotides" indicating that they might have been templated by other sequences than the surrounding genomic sequence and the L1 element. Since the maximum length of extra nucleotides found in this study is six, no putative parental sequence could be identified.

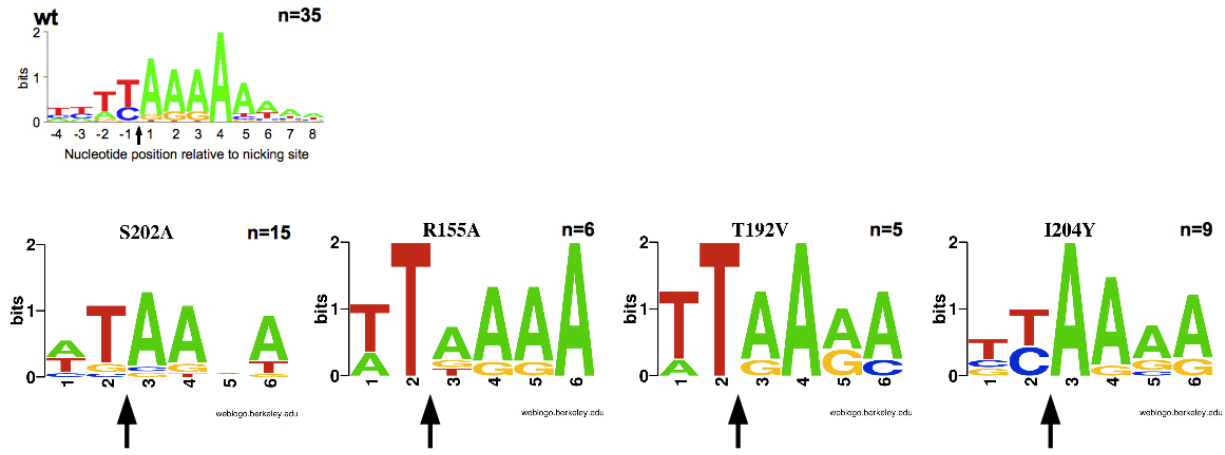


Figure 47 Sequence logos of genomic integration sites of *de novo* retrotransposition events launched from L1 reporter plasmids carrying the wildtype L1.3 element (wt) or L1 elements with EN mutants (S202A, R155A, T192V and I204Y). The frequencies of bases at each position are indicated by the relative height of letters, while the degree of sequence conservation is represented by the total height of a stack of letters given in bits of information. The total number of integration sites analyzed is given in each sequence logo. Arrows indicate the nicking sites. Sequence logo of wt L1 EN was taken from Nora Zingler, PhD Thesis 2004.

The majority of endogenous human L1 elements is flanked by TSDs ranging from 7-21 bp (Szak et al. 2002). TSDs generated after retrotransposition of the EN mutant elements displayed a much broader length distribution, starting at 1 bp up to 269 bp. Concerning very short TSDs of less than 4 bp, it has to be noted that these patches do not necessarily represent

TSDs. Since L1 retrotransposition is sometimes accompanied by the generation of untemplated nucleotides at the integration sites as described above, short patches of apparent TSDs could also resemble the insertion of such untemplated nucleotides that mirror the sequence of the integration site just by coincidence. In that case, shorter stretches of untemplated nucleotides reflecting integration site sequences are more likely than longer ones. The L1 research community has agreed on defining TSDs as sequence duplications of at least 4 nucleotides, although the probability of a coincidentally specific amino acid sequence of 3 nucleotides is only 0.016 or 1.6% (accepting that there are 4 different amino acids, the probability of 1 coincidentally specific amino acid is 0.25; therefore, the probability of 3 coincidentally specific amino acids in the proper sequence is 0.25^3=0.016). Therefore, I included also the TSDs shorter than 4 nucleotides in Table 9 to reflect the situation for the different integration sites as outright as possible. Concerning the extended TSDs longer than 21 bp, these were also isolated in other retrotransposition studies using wildtype L1 elements (Gilbert et al. 2002; Symer et al. 2002); Nora Zingler, PhD Thesis 2004) and are therefore considered not to be hallmarks of the L1 EN point mutations.

4 DISCUSSION

4.1 HOST-ENCODED FACTORS INVOLVED IN LINE-1 RETROTRANSPOSITION

4.1.1 ATM PLAYS AN AMBIVALENT ROLE IN L1 RETROTRANSPOSITION

So far, not many DSB repair factors have been shown to be involved in L1 retrotransposition. And when I started with this study, there were only strong hints pointing towards an involvement of NHEJ (Morrish et al. 2002; Zingler et al. 2005b), but no specific repair factor had been identified. In the meantime, the group around P. Deininger found two DSB repair factors being involved in L1 retrotransposition, ATM (Gasior et al. 2006) and ERCC1/XPF (Gasior et al. 2008).

Gasior et al. published in 2006 that ATM is required for L1-induced γ-H2AX foci and for L1 retrotransposition. Concerning the first point of this statement, this is in line with my findings that toxicity of the ATM KD is less severe without an active L1 EN (Figure 35). However, from my experiments I cannot conclude that ATM is required for L1 retrotransposition, since I observed rather the opposite (Figure 34). Gasior et al. first used HeLa cells to show the induction of γ-H2AX foci upon functional L1 EN introduction into the cells. As γ-H2AX foci can also be induced by apoptosis, it would have been essential to show that apoptosis was not induced upon L1 reporter transfection. Instead, the authors argue that HeLa cells do not have detectable levels of p53 protein, so an induction of foci due to L1-induced apoptosis is unlikely in that cell line. Since I was able to show very robust expression levels of p53 in HeLa cells (Figure 29) and Haoudi et al. in 2004 (Haoudi et al. 2004) were able to show that retrotransposition-competent human LINE-1 induces apoptosis in cancer cells with intact p53, it is indeed very likely that the observed γ-H2AX foci by Gasior et al. are due to L1 induced apoptosis. In fact, the argument of Gasior et al. to corroborate their assumption of HeLa cells not undergoing apoptosis, since MCF7 cells showed the same induction of γ-H2AX foci following wildtype L1 transfections in their study, is strongly in favor that HeLa cells do enter apoptosis, as Haoudi et al. have shown this in the case of MCF7 cells. However, Haoudi et al. also demonstrated, that the apoptosis induction by L1 can be prevented, if p53 is functionally inactivated by the expression of human papillomavirus (HPV) E6 protein. E6 targets p53 for ubiquitin-dependent proteolysis (Talis et al. 1998). Since HeLa cells are known to be HPV positive (Schwarz et al. 1985), it could be conceivable that apoptosis is therefore reduced in the case of Gasior et al. In my case, p53 depletion leads to reduced G418 resistant colony

numbers, meaning reduced retrotransposition events, suggesting a negative effect of p53 depletion on L1 retrotransposition. This could be explained by the findings of Harris et al. in 2009 (Harris et al. 2009) revealing a large number of p53 responsive elements or p53 DNA binding sites in L1 elements within the human genome. p53 was shown to directly bind to a short 15-nucleotide sequence within the L1 promoter. They propose a model in which p53 increases the expression of L1 mRNAs.

Furthermore, Gasior et al. used two sister cell lines of SV40 immortalized fibroblasts, one that is ATM positive and one that is ATM negative (YZ-5 and EBS7, respectively) by stable transfection with a plasmid coding for cATM or the empty vector. Recombinant expression has to be selected with Hyg, as the plasmids used are episomal and carry Hyg resistance genes. The L1 reporter plasmid used by Gasior et al. is episomal, too and carries a Hyg resitance gene, which means that the culture conditions needed by YZ-5 and EBS7 not only select for the cATM coding or empty vector, but also for the L1 reporter used. Thus, active L1 EN accumulates throughout the retrotransposition assay and is very likely causing severe toxicity, especially in the case without ATM. Therefore, the results of Gasior et al. actually confirm my findings that ATM depletion causes increased toxicity, if a functional L1 reporter element is expressed at the same time. Unfortunately, Gasior et al. do not show any control experiments to rule out increased toxicity in absence of ATM and in presence of active L1 EN. My results point to the necessity of ATM in repairing L1 induced damages and thus preventing toxicity, but at the same time ATM inhibits L1 retrotransposition. Both parts of this result are actually consistent with the idea of protecting the host cell from damages. Fortunately, this hypothesis was substantiated in 2011 by Coufal et al. showing an increase in the retrotransposition efficiency of engineered human L1s in human neural stem cells that lack or contain severely reduced levels of ATM (Coufal et al. 2011).

To conclude the discussion of contradictory results between different laboratories using HeLa cells, it must be noted that HeLa cells are an extremely variable cell line that is subject to enormous genetic drift and even was contaminated with other cell lines in the past (Masters 2002). Therefore, utmost caution has to be administered when comparing data using HeLa cells or other variable cell lines.

4.1.2 POTENTIAL ROLES OF PARP AND RAD51 IN L1 RETROTRANSPOSITION

PARP-1 (PARP) is a highly abundant chromatin-associated enzyme playing key roles in the maintenance of genomic integrity, chromatin remodeling and transcriptional control. It binds to DNA single- and double-strand breaks through an N-terminal region containing two zinc fingers. Its C-terminal catalytic domain becomes activated via an unknown mechanism, causing formation and addition of polyadenosine-ribose to acceptor proteins and autopolyadenosine-ribosylation (Eustermann et al. 2011). Ku70 and PARP compete directly for binding of DNA ends guiding the repair towards D-NHEJ or B-NHEJ, respectively (Wang et al. 2006). In this study, inhibition of PARP was achieved by overexpression of the DNA-binding domain of PARP (PARP DBD) (Kupper et al. 1990), or wildtype PARP was overexpressed expecting the enhancement of the effects of endogenous PARP. However, expression of both, PARP DBD and the wildtype PARP, led to a reduction of retrotransposition events (Figure 43). This observation could be explained, if L1 retrotransposition relied on the binding of Ku70. In that case, any PARP expression would be interfering with the binding of Ku70, thus reducing L1 retrotransposition. Unfortunately, the shRNA expressing construct directed against Ku70 expression was too toxic for the cells to work with (Figure 34 and Figure 35). Therefore, to prove this hypothesis, another experimental approach is necessary.

In a recent publication, PARP was shown to cooperate with YY1 in expression regulation of Cxcl12, where PARP has an inhibitory function on transcription, whereas YY1 has a strongly activating effect (Markovic et al. 2013). Since the L1 promoter was shown to harbor a YY1 binding site (Becker et al. 1993; Kurose et al. 1995), it is supposable that it contains a binding motif for PARP as well. Therefore, to explain the results another scenario is conceivable, in which either of the PARP constructs binds to the L1 promoter causing polyadenosine-ribosylation-independent transcriptional repression and reduced L1 retrotransposition. However, if this was the case, L1 retrotransposition with the inducible L1 reporter should be independent of PARP DBD or PARP expression, as ptetO7CMV/L1-3blas does not contain the internal L1 promoter. Since the function of PARP as transcription factor was not yet known at the time when I was performing the experiments of this study and the early expression of the PARP constructs did not demand for a verification of the results with ptetO7CMV/L1-3blas, I did not perform co-transfection experiments with the inducible L1 reporter and the PARP constructs. Therefore, this is something that could still be done to unravel the involvement of endogenous factors in L1 retrotransposition further.

Rad51 is the main catalyst of homologous recombination (HR) in all eukaryotes playing a critical role in allowing replication fork progression when encountering a DNA lesion. HR also serves as the primary mechanism for error-free repair of DNA double-strand breaks (DSBs). Rad51 was found to play a role in transposition of Sleeping Beauty in NHEJ-deficient cells (Izsvak et al. 2004) and in mitochondrial DNA replication (Sage et al. 2010). We hypothesized that Rad51 might be involved in strand transfer of L1 during TPRT. To test this hypothesis, either murine Rad51 (Rad51Mm) or a yeast-murine chimera of Rad51 (Rad51SM) was co-expressed with the L1 reporter construct. Rad51Mm was assumed to exert normal function in the cells, while Rad51SM was supposed to be a dominant negative mutant (Lambert and Lopez 2000). When Rad51Mm was co-expressed, retrotransposition rates were slightly increased suggesting a positive role of Rad51 in the retrotransposition mechanism (Figure 43). Yet, when co-expressing Rad51SM, the retrotransposition rate was not reduced as could be expected in turn. It remained more or less unaltered instead (Figure 43). However, it has to be noted that the toxicity caused by the Rad51SM expression plasmid was rather high which could mask a possible effect on L1 retrotransposition. Therefore, the results obtained after expression of Rad51SM have to be taken with a pinch of salt. The role of Rad51 in L1 retrotransposition seems ambivalent. Since Rad51Mm is of murine origin and the HeLa cells are human, it cannot be ruled out that there might be some species-dependent interference that hinders Rad51 to exert its normal function. This could also explain the rather great toxicities of both Rad51 constructs observed during the assays (assuming that endogenous Rad51 is essential to repair L1 EN-induced DSBs). Therefore, to verify the results and possibly gain further insights, the experiments should be repeated either in murine or hamster cells or with human Rad51 constructs.

4.1.3 OTHER HOST-ENCODED FACTORS INVOLVED IN L1 RETROTRANSPOSITION NOT INVESTIGATED IN THIS WORK

ERCC1 (excision repair cross-complementing rodent repair deficiency, complementation group 1) and XPF (xeroderma pigmentosum, complementation group F; also termed ERCC4) were originally found to be involved in Nucleotide Excision Repair (NER) (Cleaver 1968; De Weerd-Kastelein et al. 1972; Busch et al. 1989) and in the meantime they were shown to act on minimal DNA substrates with 3' flaps (de Laat et al. 1998). Since such 3' flaps are a structure proposed by the TPRT model as retrotransposition intermediate, Gasior et al.

(Gasior et al. 2008) had a closer look at the involvement of these factors in L1 retrotransposition. They found that even minimal reduction of XPF in HeLa cells increased L1 retrotransposition, while complementation of ERCC1 in ERCC1-deficient CHO cells reduced retrotransposition. Although again no transfection efficiency and toxicity controls are shown, the authors conclude that these NER involved factors limit L1 retrotransposition in human and hamster cells.

Not only DSB repair factors have been shown to play a critical role in L1 retrotransposition, also members of the APOBEC3 family of cytidine deaminases were previously proven to inhibit *de novo* L1 insertions (Muckenfuss et al. 2006; Kinomoto et al. 2007; Horn et al. 2014). APOBEC3 proteins are known for their involvement in anti-retroviral intrinsic immunity. In primates, there are eight paralogues, APOBEC3A, -B, -C, -D, -E, -F, -G, and -H (A3A-H), while rodents carry only one single *Apobec3* gene. Human A3G and A3F have been shown to be active against human immunodeficiency virus-1 (HIV-1) (Sheehy et al. 2002; Zheng et al. 2004; Wiegand et al. 2004; Liddament et al. 2004; Bishop et al. 2004), inducing C-to-U transitions by cytidine deamination on the viral minus strand DNA during reverse transcription (Mangeat et al. 2003; Zhang et al. 2003). Studies showed that also LTR-retrotransposons are sensitive to A3G and A3F (Schumacher et al. 2005; Dutko et al. 2005; Esnault et al. 2005). Apart from its mutagenic activity, A3G has been shown to inhibit hepatitis B virus (HBV) and human T cell leukemia virus type 1 (HTLV-1) without causing mutations in the majority of the viral genomes (Turelli et al. 2004; Rosler et al. 2004; Sasada et al. 2005; Rosler et al. 2005; Mahieux et al. 2005; Noguchi et al. 2005; Suspene et al. 2005). Interestingly, the mechanisms inhibiting L1 retrotransposition found by Muckenfuss et al. in 2006 and Kinomoto et al. 2007 where also independent of cytidine deamination. APOBEC3C was shown to restrict L1 retrotransposition by interaction with L1 ORF1 protein affecting the element's reverse transcriptase activity (Horn et al. 2014). In contrast, APOBEC3A deaminates L1 cDNA during retrotransposition when exposed single-stranded (Richardson et al. 2014).

Post-transcriptional control of LINE-1 retrotransposition by cellular host factors in somatic cells is reviewed in more detail elsewhere (Pizarro and Cristofari 2016).

Apart from the factors presumably acting at the point of reverse transcription mentioned above, several other proteins have been shown to be involved in the transcriptional regulation

of L1. For instance, Sox11, a member of the SRY family transcription factors, is a positive regulator of L1 transcription (Tchenio et al. 2000), whereas Sox2 associates with the L1 5' UTR in a pattern that correlates with the decrease in its expression observed during neural differentiation (Muotri et al. 2005; Coufal et al. 2009). Another positive regulator of L1 is the runt-domain transcription factor RUNX3 binding to nucleotides +83 to +101 of the L1 5' UTR (Yang et al. 2003). Transcription factor YY1 binds to nucleotides +13 to +26 of L1 (Becker et al. 1993; Kurose et al. 1995) and functions as a component of the L1 core promoter to direct accurate transcription initiation (Athanikar et al. 2004). Furthermore, it is capable of both, activating and repressing transcription in general. Therefore, this protein is hypothesized to play a role in L1 downregulation as well as activation in a cell type specific manner (Becker et al. 1993). Notably, the YY1 binding site seems to be conserved in all full-length L1 sequences in the human genome (Khan et al. 2006).

The Wnt gene family consists of structurally related genes, which encode secreted signaling proteins that have been implicated in oncogenesis and in several developmental processes, such as regulation of cell fate and patterning during embryogenesis. Kuwabara et al. found in 2009 that Wnt signaling, together with the removal of Sox2, triggered the simultaneous expression of NeuroD1 and LINE-1 in mice (Kuwabara et al. 2009). It was hypothesized that this regulatory mechanism could be important for adult neurogenesis and survival of neuronal progenitors.

The transcriptional regulation of L1 is clearly mediated by a heavily methylated CpG island in its 5' UTR (Woodcock et al. 1997). It is believed that this hypermethylation is a major defense mechanism to repress these elements that could be otherwise very damaging if transcribed. In a study of eight human cell lines, an inverse correlation between ORF1 protein expression and the methylation state of the L1 5' end was observed. This strongly indicates that the presence of methylation plays a role in L1 transcription (Thayer et al. 1993). In general, transcriptional repression by DNA methylation is thought to be mediated by binding of members of the methyl-CpG-binding domain (MBD) protein family, recruitment of histone deacetylases and generation of a transcriptionally inactive chromatin structure (Hendrich and Bird 1998; Bird and Wolffe 1999). The MBD family member first described is methyl-CpG-binding protein 2 (MeCP2) (Lewis et al. 1992; Weitzel et al. 1997) that was also shown to repress transcription through a histone deacetylase-independent mechanism via direct and indirect interactions with the transcription factor TFIIB (Nan et al. 1997; Yu et al. 2000; Kaludov and Wolffe 2000).

Using transient transfection assays as I did in this study, Yu et al. found in 2001 that the transcriptional repressor domain (TRD) of MeCP2 represses transcription from L1 promoter-driven luciferase constructs and that full-length MeCP2 represses transcription from a methylated version of the construct (Yu et al. 2001). Furthermore, employing a full-length neo-marked L1 reporter construct, the TRD of MeCP2 effectively reduced L1 retrotransposition in HeLa cells, and the full-length MeCP2 increased the methylation-induced repression of L1 retrotransposition. Those results suggest an involvement of MeCP2 in the host cell defense against L1 elements.
These findings were confirmed in neuronal cells and *in vivo* by Muotri et al. (Muotri et al. 2010). They show that L1 transcription and retrotransposition in mouse neuroepithelial cells is increased in the absence of MeCP2 and in the case of methylation of the L1 5' UTR. Furthermore, using neuronal progenitor cells derived from human induced pluripotent stem cells (iPSC) and human tissues processed post mortem, they reveal that patients with Rett syndrome carrying MeCP2 mutations have an increased susceptibility for L1 retrotransposition. However, the authors also point out that they cannot exclude an indirect effect of MeCP2 in regulating genes involved in L1 expression and/or in changing the chromatin structure to facilitate *de novo* L1 insertions.

In 2009, another transcription factor complex involved in L1 transcription was found (Montoya-Durango et al. 2009). The retinoblastoma tumor suppressor protein (Rb) binds to the E2F1 transcription factor preventing it from activating genes downstream of its CpG islands binding sites and associates with histone methyltransferases (HMTases) and histone deacetylases (HDACs) (Robertson et al. 2000; Gonzalo et al. 2005). In the absence of Rb, E2F1 mediates the expression of several target genes that facilitate the G1/S transition in cell cycle and S-phase. Since L1 elements are regulated by their degree of methylation, and E2F/Rb proteins are potential mediators of DNA and histone methylation, it was not surprising to find that Rb and E2F bind to the L1 promoter in human and mouse cells and that in cells lacking Rb family members, L1 expression showed marked upregulation and was coupled to reductions in histone methylation and deacetylation. L1 silencing by Rb proteins was shown to be effected through the Nucleosomal and Remodeling Deacetylase (NuRD) multiprotein complex (Montoya-Durango et al. 2016). In this context, it should also be noted that aryl hydrocarbon receptor (AHR) transcription factor is known to interact directly with Rb (Puga et al. 2000) and AHR modulates the transcriptional activity of L1 in mouse as well as human cells (Teneng et al. 2007).

4.2 TRANSFECTION EFFICIENCY AND TOXICITY CONTROLS

To test the effect of the knock down (KD) and the expression of dominant negative (DN) mutants of various double-strand break (DSB) repair factors on L1 retrotransposition, co-transfections into HeLa cells with L1 reporter and respective shRNA or DN mutant constructs were performed. One day after transfection, the cells were set on hygromycin (Hyg) selection for transfected cells. For assessment of transfection efficiency and toxicity, each well of transfected cells was split three days post transfection selecting one subculture with Hyg and the other with G418. Transfection efficiency and toxicity evaluation by quantifying Hyg resistant colonies yielded very different results for the different KD and DN mutant expressing constructs (Chapters 3.3.1 and 3.3.2), but in almost all cases of shRNA contructs, a reduced number of Hyg resistant colonies was found (Figure 34 and Figure 35). Interestingly, co-transfection of the ATR shRNA expressing plasmid along with the L1 reporter, yielded more than twice as many Hyg resistant colonies as the control with empty plasmid and L1 reporter (Figure 34). Enhanced cell survival with ATR shRNA was also observed when co-transfecting the EN inactive mutant L1 reporter instead of the wildtype one (Figure 35). This is surprising, since Western Blot analysis revealed no drastic change in ATR expression (Figure 32). Only a very late and rather weak decrease in ATR level at day 6 after transfection could be shown. At the same time, this weak ATR KD was found to reduce L1 retrotransposition level (Figure 34).

In contrast, co-transfection of the ATM shRNA plasmid with the L1 reporter caused high toxicity reducing the number of Hyg resistant colonies approximately 3-fold (Figure 34). However, the toxicity was found to be reduced when co-transfecting the EN inactive L1 reporter instead of the active one (Figure 35). This points towards a synergistic toxic effect of ATM KD and L1 EN causing DSBs. However, also in the case of ATM, Western Blot showed only a very late and weak reduction of ATM level at 7 days post transfection (Figure 31). Surprisingly, again this seemed to be sufficient to cause a 3-fold increase in L1 retrotransposition (Figure 34).

Since with most tested shRNA constructs the KD effect was rather weak and at the same time toxicity high, I propose that shRNA-non-specific modes of action account for this observation. Due to the structure of shRNA, one possible mechanism could be the activation of usually virus-induced pathways like the double-stranded RNA-dependent protein kinase (PKR) induced apoptosis (Garcia et al. 2007).

A cell viability/apoptosis assay would have given certainty on the toxicity of each transfection experiment and comparing the obtained values to the relative Hyg colony numbers would have allowed a more precise analysis of the causes for reduced Hyg resistant colony numbers. However, this was not performed, since in most cases the KD of the respective factor did not work sufficiently efficiently.

4.3 DEPLETION OF DSB REPAIR FACTORS IN HELA M2 CELLS CAUSED BY HYGROMYCIN SELECTION

There are three related protein kinases playing a central role in the repair of DNA damage, namely ATM, ATR and DNA-PK. These kinases belong to the phosphoinositide-3-like kinase kinase (PIKK) family and share similar domain structure and modes of regulation (Mordes and Cortez 2008; Lempiainen and Halazonetis 2009; Lovejoy and Cortez 2009). While DSBs activate ATM and DNA-PK, several types of DNA damage may activate ATR, including DSBs, base adducts and crosslinks (Cimprich and Cortez 2008). Once activated, these kinases preferentially phosphorylate serines and threonines followed by a glutamine in many protein substrates. Many such regulatory proteins have a relatively rapid turnover rate, which ensures that they can respond quickly to changing situations in either their rate of synthesis or rate of degradation. Unfortunately, no exact half-life numbers of these kinases are published so far.
In my experiments, ATM reduction is detectable only 24 hours after transfection of HeLa M2 cells either with pSHAG-ATM and the L1 reporter or the empty vector control and the L1 reporter (Figure 36). This knock down (KD) is therefore independent from the used shRNA targeting ATM and may be a result of stress induced in the cells by the transfection procedure. However, upon Hyg selection, ATM depletion is detectable at 24 hours post transfection and later in both co-transfection cases (Figure 36). In eukaryotic cells, the mode of action of Hyg is by inhibiting the mRNA translocation during translation. The Hyg resistance gene codes for a kinase that inactivates Hygromycin B by phosphorylation (Rao et al. 1983). Despite of the L1 reporter plasmid carrying a Hyg resistance gene, this inhibition of protein synthesis could be responsible for the enormous reduction of the ATM level, if the transfection efficiency of the cells was not sufficient to entirely remedy this effect by sufficient expression of the Hyg resistance. The same could account for ATR and DNA PK, as they display the same reaction upon Hyg selection. Why this effect does not affect the expression of proteins in general can only be hypothesized. The partial inhibition of proteinsynthesis could depend on protein size, protein turnover or even expression regulation. However, when using Hyg selection, this

effect on protein expression should be considered, as the reduction of the kinases mentioned above and of other proteins may bias the results.

4.4 EFFORTS TO MODIFY L1 INTEGRATION SPECIFICITY

MODULATION OF LINE-1 SPECIFICITY BY ENDONUCLEASE MUTATION. Engineering a target sequence specific LINE-1 retrotransposon would be of great interest as tool for genomic manipulations in the generation of transgenic or knockout animals, as well as for gene therapy treatment. Modulation of cleavage specificity is thought to be mainly achieved by variations in the surface loop of L1 endonuclease (EN) contacting the DNA (Nora Zingler, PhD Thesis 2004; Repanas et al. 2007). In more detail, it was hypothesized that L1 EN accommodates a flipped, extrahelical adenine downstream of the scissile bond, as I204 is small enough for a ribose to be placed and because of the rather small serine in position 202 supplying even space for a purine base. Furthermore, S202 and R155 are likely to form hydrogen bonds with the extrahelical adenine to stabilize the protein-DNA interaction. T192 plays an important structural role at the base of the βB6-βB5 loop as its side chain oxygen receives weak hydrogen bonds from the main chain nitrogens of I204 and D205 and thus anchors the bottom of the loop with respect to the active site (Weichenrieder et al. 2004; Zingler et al. 2005a). To prove these hypotheses, four EN point mutants were generated and characterized in this study: S202A, R155A, T192V and I204Y (Figure 45).

In our cell culture based L1 retrotransposition assay, L1 reporter elements encoding each of the four L1 EN mutants, exhibited a substantially reduced retrotransposition activity relative to the wildtype L1 reporter element (Table 8). The relative retrotransposition of the mutants compared to the wildtype element were 32% (±7.8%) for S202A, 12% (±3.3%) for R155A, 5% (±3.0%) for T192V and 1% (±1.1%) for I204Y. The reasons for these differences are not clear so far, but it is most likely due to conformational changes of the EN resulting from the exchange of single, structural relevant amino acids. Consistently with this hypothesis, exchanging T192 and I204 that anchor the bottom of the βB6-βB5 loop and thus should play the most important role in structural organization of the EN, results in almost complete inactivation of the L1 reporter element.

Interestingly, a recent study by Kines et al. demonstrated that the L1 EN domain can tolerate many mutations without significantly impacting its function, despite containing single and multiple point mutations in putative phosphorylation sites and in highly conserved residues predicted to be structurally important (others than the mutations analyzed in this thesis)

(Kines et al. 2016). This difference in results demonstrates clearly, that the function of single amino acids within the L1 EN domain is still not unequivocally elucidated.
In contrast to retrotransposition activity, target site specificity of the L1 element could not be altered relevantly. The only peculiarities observed were found in the target site consensus of the S202A variant, where at the first position, the T of the wildtype target site consensus (TT AAAA) was replaced by an A (AT AAAA) (Figure 47). However, an A may also occur in the wildtype target site at that position, although less frequently as in the S202A variant. Furthermore, the third A downstream of the nicking site is less conserved in the S202A variant as in the wildtype. Obviously the point mutations chosen are not sufficient to modify EN specificity, but rather result in a reduction of endonucleolytic activity. Therefore, additional, so far unknown specificity factors could influence the choice of nicking site or select among already nicked sites suitable ones for integration.

MODULATION OF LINE-1 SPECIFICITY BY ZINC FINGER TARGETING. Another, promising approach to target L1 to specific sequences could be the use of designed zinc finger nucleases (ZFNs). ZFNs are artificial fusion proteins that link a zinc finger DNA binding domain to a nonspecific nuclease domain (Porteus and Carroll 2005). Each zinc finger is composed of 30 amino acids folding into a ββα configuration, coordinating one Zn^{+} atom between two cysteines and two histidines and contacting 3 bp of DNA (Pavletich and Pabo 1991). By combining individual zinc fingers, the overall binding specificity of the zinc finger construct can be changed, yielding a mere unlimited number of target sequences (Porteus 2006). Zinc finger constructs have been combined with FokI type II restriction endonuclease (Kim et al. 1996) and used for instance in oocytes from *Xenopus laevis* (Smith et al. 2000; Bibikova et al. 2001), *Drosophila melanogaster* (Bibikova et al. 2002; Bibikova et al. 2003), a human embryonic kidney cell line (HEK293) (Porteus and Baltimore 2003; Porteus 2006), a transformed human erythroleukemia cell line (K562) (Urnov et al. 2005; Wienert et al. 2015), cultured primary human T cells (Urnov et al. 2005; Yuan et al. 2012), mouse spermatogonial stem cells (Fanslow et al. 2014), induced pluripotent stem cells (iPSCs) (Rahman et al. 2015) and even in the plant *Arabidopsis thaliana* (Lloyd et al. 2005; Qi et al. 2014). Furthermore, the ZFN technology has already been applied to Sleeping Beauty with reasonable success, where it resulted in retargeting of the element to the erbB-2 gene and LINE-1 sequences within the human genome, respectively (Voigt et al. 2012). Because of their small size, zinc fingers are very interesting partners for creating protein fusions as they are thought to hardly disturb protein folding and function. Therefore, it is tempting to fuse a set of sequence

specific zinc fingers to the L1 ORF2 protein harboring the L1 endonuclease (L1 EN). This could be done at several positions: (1) at the N-terminus of the L1 EN domain, (2) between L1 EN and L1 reverse transcriptase (RT) domain, (3) within the C-terminus of the L1 ORF2 protein replacing the endogenous zinc finger related motifs or (4) at the very C-terminus of the L1 ORF2 protein.

4.5 THE ROLE OF L1 RETROTRANSPOSITION IN HOST CELLS AND ORGANISMS

Scientists have long been puzzled by the question concerning a role of retroelements in the genome, since no clear function was ever published. This also led to the concept of "junk DNA", the idea that such sequences are mere evolutionary relicts causing cancer from time to time when accidentally activated. However, the finding that retrotransposons can actively reshape the genome is now challenging this hypothesis (Muotri et al. 2009). Apart from the evolutionary role of retrotransposons mentioned in the introduction, evidence in the literature points to a somatic function of L1 transcripts in cell proliferation (Kuo et al. 1998), differentiation (Mangiacasale et al. 2003) and early embryo development (Pittoggi et al. 2003), although it is still unclear how these different L1 elements act. Apart from this, L1, as well as *Alu* and SVA mobilization were also observed during reprogramming and pluripotent stem cell cultivation (Klawitter et al. 2016). It is difficult to comprehend why the genome would tolerate so many copies of a retrotransposon, if it did not bring any benefit with it.

4.5.1 SOMATIC L1 RETROTRANSPOSITION

PREVIOUS FINDINGS RELEVANT FOR THIS THESIS. Several years ago, a more special case of somatic L1 retrotransposition attracted the attention of scientists working in the field of L1 research and in neurobiology (Muotri et al. 2005). Employing the common L1 reporter construct, Muotri et al. had observed *de novo* L1 integrations in neuronal precursor cells (NPCs) derived from rat hippocampus neural stem cells. They also found that many of these events had inserted into neuronally expressed genes by conventional endonuclease-dependent retrotransposition. This ability of L1 to retrotranspose into genes was consistent with results from previous studies performed in transformed cultured cell lines (Moran et al. 1999; Gilbert et al. 2002; Symer et al. 2002). Then, to determine whether L1 retrotransposition could occur *in vivo*, they generated a transgenic mouse harboring a human L1 element ($L1_{RP}$) under the control of its endogenous promoter and tagged with the EGFP retrotransposition reporter

transcribed from the ubiquitous CMV promoter. EGFP-positive cells were detected in germ cells and in the brains of transgenic animals, including striatum, cortex, hypothalamus, hilus, cerebellum, ventricles, amygdala and hippocampus. These EGFP-positive cells co-localized only with a neuronal marker and not with oligodendrocyte or astrocyte markers, indicating that L1 retrotransposition might have occurred in neuronal precursor cells rather than glial precursor cells. At the same time, EGFP-positive cells were detected in the cephalic neural tube at embryonic development day 10.5, indicating that L1 retrotransposition takes place during embryonic and adult neurogenesis. Since in one instance, they had found retrotransposition of an L1 element into the *DLG2* gene (PSD-93) which can influence its expression and in turn influence the differentiation pattern of NPCs, these data were the first to provide evidence for somatic mosaicism in neuronal cells of the brain due to L1 retrotransposition.

To further study somatic retrotransposition, more transgenic, L1 expressing mice were generated in 2006 (An et al. 2006). Since the retrotransposition frequencies of two of the most active human L1 isolates, $L1_{RP}$ (Kimberland et al. 1999) and L1LRE3 (Brouha et al. 2002), were not satisfactory when introduced into mice ($L1_{RP}$: Ostertag et al. 2002; Prak et al. 2003; L1LRE3: Babushok et al. 2006), An et al. designed a synthetic L1 in which both ORFs of $L1_{RP}$ were codon usage optimized to mouse. This element termed ORFeus, was significantly more active when tested in cultured murine cells. After introduction of ORFeus under the control of the constitutive composite chicken β-actin promoter (CAG) into the germ line of mice, they found two real founders containing the intron-harboring reporter cassette, while six of the screened 28 mice contained the intronless reporter. This observation is consistent with my finding of only a spliced GFP gene in the case of one LORFUS founder indicating a retrotransposition event very early after pronucleus injection. After backcrossing one founder line for two generations into wildtype C57BL/6J mice, An et al. found three groups of animals: The first group contained both, the donor element and insertions making up 50.4% of the progeny; the second group had only insertions accounting for 13.9% of all animals; and the third group was negative for both. These results exhibit clearly the necessity of a conditional regulation of the L1 transgene, since after just two generations, there were no animals left containing only the donor element. Due to the accumulation of transgene retrotransposition events, these animals are prone to express GFP in more and more tissues making it more or less impossible to analyze new integration events eventually. Concerning the somatic retrotransposition events, they detected them only in a subset of six tissues, including bladder, muscle, skin, intestine, kidney and tail, and not in brain, heart, ovary, lung,

liver, eye, stomach, spleen and pancreas. Despite the high level of retrotransposition activities in germ line and somatic tissues, they were not able to observe any obvious fitness reductions in the transgenic animals.

For completeness, it should be mentioned that in 2008, a more sophisticated ORFeus transgenic mouse was published: To achieve controlled activation, the authors constructed ORFeusLSL, in which ORFeus was separated from the promoter by a loxP-β-geo-stop-loxP (LSL) cassette (An et al. 2008). By crossing ORFeusLSL to various Cre recombinase expressing lines, they were able to obtain tissue-specific ORFeus activation and otherwise maintain the ORFeusLSL transgene stable without accumulating retrotransposition events.

FINDINGS OF THIS THESIS. In this study, I generated a transgenic mouse harboring a GFP reporter tagged ORFeus and a β-galactosidase (β-Gal) simultaneously under the control of the bidirectional tTA-dependent CMV promoter (Figure 23). In this mouse named LORFUS, ORFeus and β-Gal are only expressed, if a tTA or an rtTA molecule binds to the operator sequence of the tTA-dependent CMV promoter rendering it highly controllable and tissue specific. When crossbreeding LORFUS with Kt1 mice, expression in the forebrain can be switched off by the application of doxycyclin (Dox) and can be turned back on by withdrawal of Dox. This temporally and spatially controlled expression of the transgene is a clear advantage over all other L1 transgenic animals mentioned above. By crossbreeding LORFUS with other tTA- or rtTA-expressing mouse lines, expression of ORFeus can be driven to other tissues, too. Apart from this great flexibility, the simultaneous expression of β-Gal allows the determination of tissues and cells expressing the transgene but not allowing retrotransposition for whatever reasons. Therefore, three types of scenarios can be monitored: (1) cells only positive for β-Gal and thus expressing the transgene but not allowing L1 retrotransposition; (2) cells positive for β-Gal and GFP and thus expressing the transgene and allowing L1 retrotransposition and (3) cells positive for GFP only indicating no transgene expression at that time point and that the retrotransposition event must have occurred earlier during development when the transgene was still expressed. This possibility of temporal resolution of L1 retrotransposition is unique to LORFUS and distinguishes it further from other L1 transgenic animals. However, it must be noted that mosaic expression of the two reporter genes of the bidirectional tTA-dependent CMV promoter cassette has been observed in tissues of other animals (unpublished data by Rolf Sprengel, Max-Planck-Institute for Medical Research, Heidelberg). This is probably due to silencing of just one direction of the

bidirectional promoter. Therefore, when analyzing cells for either expression of the β-Gal or the ORFeus, always cell populations should be looked at and not just single cells.

As mentioned above, when Muotri et al. in 2005 analyzed the insertion sites of their L1 *in vivo*, they found a preference of L1 to integrate into neuronally expressed genes, for instance an olfactory receptor (*Olr346*), an ion-channel-associated gene (*DLG2*) and a cadherin receptor (*rCNR*). Furthermore, the authors could show *de novo* retrotransposition events in several brain regions, including cortex, hippocampus and amygdala. The EGFP signal of the L1 co-localized only with a neuronal marker (NeuN) and not with oligodendrocyte or astrocyte markers, indicating that retrotransposition had occurred in NPCs rather than in glial or common precursor cells earlier in development. These findings are in line with my results showing ORFeus retrotransposition in hippocampus, cortex, striatum, olfactory bulb (OB) and brainstem (Figure 26). I could not detect L1 integration events in amygdala, since the slices I analyzed do not cover this region. As I found GFP-positive, but not β-Gal-positive cells in the OB, this indicates retrotransposition must have occurred earlier during development, since the transgene is not expressed at this stage (postnatal day 15). Due to the observation of β-Gal-positive cells in the granule layer of the dentate gyrus (DG) and some along the rostral migratory stream, it is likely that retrotransposition has occurred in the NPCs of the DG before or while the cells migrated into the OB.

Due to time constraints, I was not able to further analyze LORFUS. Future studies should include mapping of the genomic transgene integration, as well as a temporal and spatial characterization of L1 retrotransposition in the different brain regions throughout development. And finally, performing various behavioral assays with double transgenic animals (LORFUS/Kt1), L1 retrotransposition in the brain should be linked to a role or a phenotype.

FINDINGS AFTER MY THESIS WORK RELATING TO IT. Concerning a behavioral phenotype for L1 retrotransposition supporting transgenic animals, an environmental influence on L1 retrotransposon activity in the hippocampus of mice was reported in 2009 (Muotri et al. 2009). The authors found that neurons from mice that had access to running wheels were more likely to activate an EGFP reporter, representing L1 insertions in the brain, when compared to resting animals. The number of EGFP positive cells increased about 3-fold in the runners' DG. Nevertheless, it still remains unclear whether running can induce direct L1 retrotransposition or promote the survival of L1-inserted cells, and this is only one of the captivating questions that could be answered with the help of LORFUS.

Later, in 2010, L1 retrotransposition in neurons was confirmed to be modulated by MeCP2 just as it was found earlier in cell lines as mentioned in Chapter 4.1 (Muotri et al. 2010). Muotri et al. also found certain brain structures to be more prone to L1 retrotransposition, including the cerebellum, striatum, cortex, hippocampus and olfactory bulb (Muotri et al. 2010). Using LORFUS I was able to observe *de novo* L1 retrotransposition in the same brain regions, except for cerebellum due to forebrain restriction of Kt1 expression.
Then, in 2013, another transgenic mouse with a Dox-inducible L1 was published expressing an ORFeus element harboring a gene-trap cassette (O'Donnell et al. 2013). This L1 mouse was bred to a CMV rtTA transgenic line revealing robust retrotransposition in somatic tissues when treated with Dox. However, this L1 mouse is not suited for fundamental research on L1 retrotransposition, because it cannot visualize expression of the transgene and tagged L1 retrotransposition at the same time like LORFUS with its bidirectional cassette.
Concerning L1 mobilization in human somatic tissue, bisulfite conversion analyses on human genomic DNAs derived from matched brain and skin tissue samples revealed that the L1 5' UTR exhibited significantly less methylation in brain samples compared to the matched skin samples (Coufal et al. 2009). Furthermore, it was demonstrated that NPCs isolated from human fetal brain and NPCs derived from human embryonic stem cells (hESCs) support L1 retrotransposition *in vitro* (Coufal et al. 2009). Later, evidence for somatic L1 insertions in the human hippocampus, dentate gyrus and caudate nucleus was provided (Baillie et al. 2011; Upton et al. 2015; Kurnosov et al. 2015).

4.5.2 GERMLINE L1 RETROTRANSPOSITION AND ITS POTENTIAL AS A TOOL

For a long time, L1 retrotransposition was thought to be restricted to germ cells only and there have been many publications demonstrating L1 activity in germline (Ostertag et al. 2002; Prak et al. 2003; Ergun et al. 2004; An et al. 2008; Kano et al. 2009; Grandi et al. 2013). Apart from the disease-causing potential of L1 retrotransposition in germ cells, there have emerged several possibilities of L1 application harnessing it as tool for genetic engineering.
Due to the possibility to insert genes into the 3' UTR of L1 without interfering with the functionality of the resulting engineered L1 element, L1 could be used to increase integration efficiency of transgenes during pronucleus injection. As detailed above, several independent research groups have observed L1 retrotransposition events very early after transgene injection resulting in systemic insertions without integration of the original L1 transgene construct. As the integration of transgenes during pronucleus injection is usually a very rare

event, introducing the transgene of interest into an L1 vector could enhance insertion events substantially. If it was possible to engineer a sequence-specifically integrating L1 element, transgenes that are transported by such an L1 element could be targeted into a safe locus neither interfering with essential gene expression, nor activating oncogenes. Such a genomic site should ensure that the transgene is ubiquitously expressed in the organism and that silencing is widely ruled out. Identifying such genomic loci is one focus of current genome research and ROSA26 is one of them that had been proposed as such in the past (Soriano 1999).

Developing a mode of adaptive L1 target specificity, as could be pursued with the approach of ZFNs, would furthermore allow L1 to be used for the generation of knockout animals. So far this process requires cultivation and transformation of embryonic stem cells and is very elaborate and inefficient. If L1-ZFNs could be constructed, L1 could be targeted to the gene of interest interrupting its transcription and thus causing a functional knockout. This approach has the potential to be more efficient than targeting the gene of interest by homologous recombination, because L1 transposes readily during early embryogenesis. Furthermore, if the targeting of L1 proves to be highly efficient, the generation of knockout animals could be simplified by injecting the L1-ZFN construct into a pronucleus just as in the case of transgenic animal generation.

5 BIBLIOGRAPHY

Ahnesorg P, Smith P, Jackson SP (2006) XLF interacts with the XRCC4-DNA ligase IV complex to promote DNA nonhomologous end-joining. Cell 124 (2):301-313. doi:10.1016/j.cell.2005.12.031

Akopiants K, Zhou RZ, Mohapatra S, Valerie K, Lees-Miller SP, Lee KJ, Chen DJ, Revy P, de Villartay JP, Povirk LF (2009) Requirement for XLF/Cernunnos in alignment-based gap filling by DNA polymerases lambda and mu for nonhomologous end joining in human whole-cell extracts. Nucleic acids research 37 (12):4055-4062. doi:10.1093/nar/gkp283

Alisch RS, Garcia-Perez JL, Muotri AR, Gage FH, Moran JV (2006) Unconventional translation of mammalian LINE-1 retrotransposons. Genes & development 20 (2):210-224. doi:10.1101/gad.1380406

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. Journal of molecular biology 215 (3):403-410. doi:10.1016/S0022-2836(05)80360-2

Amitani I, Baskin RJ, Kowalczykowski SC (2006) Visualization of Rad54, a chromatin remodeling protein, translocating on single DNA molecules. Molecular cell 23 (1):143-148. doi:10.1016/j.molcel.2006.05.009

An W, Han JS, Schrum CM, Maitra A, Koentgen F, Boeke JD (2008) Conditional activation of a single-copy L1 transgene in mice by Cre. Genesis 46 (7):373-383. doi:10.1002/dvg.20407

An W, Han JS, Wheelan SJ, Davis ES, Coombes CE, Ye P, Triplett C, Boeke JD (2006) Active retrotransposition by a synthetic L1 element in mice. Proceedings of the National Academy of Sciences of the United States of America 103 (49):18662-18667. doi:10.1073/pnas.0605300103

Aravin AA, Hannon GJ, Brennecke J (2007) The Piwi-piRNA pathway provides an adaptive defense in the transposon arms race. Science 318 (5851):761-764. doi:10.1126/science.1146484

Arcot SS, Wang Z, Weber JL, Deininger PL, Batzer MA (1995) Alu repeats: a source for the genesis of primate microsatellites. Genomics 29 (1):136-144. doi:10.1006/geno.1995.1224

Athanikar JN, Badge RM, Moran JV (2004) A YY1-binding site is required for accurate human LINE-1 transcription initiation. Nucleic acids research 32 (13):3846-3855. doi:10.1093/nar/gkh698

Audebert M, Salles B, Calsou P (2004) Involvement of poly(ADP-ribose) polymerase-1 and XRCC1/DNA ligase III in an alternative route for DNA double-strand breaks rejoining. The Journal of biological chemistry 279 (53):55117-55126. doi:10.1074/jbc.M404524200

Ausubel FM (2002) Short protocols in molecular biology : a compendium of methods from Current protocols in molecular biology. 5th edn. Wiley, New York

Babushok DV, Kazazian HH, Jr. (2007) Progress in understanding the biology of the human mutagen LINE-1. Human mutation 28 (6):527-539. doi:10.1002/humu.20486

Babushok DV, Ostertag EM, Courtney CE, Choi JM, Kazazian HH, Jr. (2006) L1 integration in a transgenic mouse model. Genome research 16 (2):240-250. doi:10.1101/gr.4571606

Baek ST, Kerjan G, Bielas SL, Lee JE, Fenstermaker AG, Novarino G, Gleeson JG (2014) Off-target effect of doublecortin family shRNA on neuronal migration associated with

endogenous microRNA dysregulation. Neuron 82 (6):1255-1262. doi:10.1016/j.neuron.2014.04.036

Baillie JK, Barnett MW, Upton KR, Gerhardt DJ, Richmond TA, De Sapio F, Brennan PM, Rizzu P, Smith S, Fell M, Talbot RT, Gustincich S, Freeman TC, Mattick JS, Hume DA, Heutink P, Carninci P, Jeddeloh JA, Faulkner GJ (2011) Somatic retrotransposition alters the genetic landscape of the human brain. Nature 479 (7374):534-537. doi:10.1038/nature10531

Baron U, Freundlieb S, Gossen M, Bujard H (1995) Co-regulation of two gene activities by tetracycline via a bidirectional promoter. Nucleic acids research 23 (17):3605-3606

Barzilay G, Hickson ID (1995) Structure and function of apurinic/apyrimidinic endonucleases. BioEssays : news and reviews in molecular, cellular and developmental biology 17 (8):713-719. doi:10.1002/bies.950170808

Beck CR, Garcia-Perez JL, Badge RM, Moran JV (2011) LINE-1 elements in structural variation and disease. Annu Rev Genomics Hum Genet 12:187-215. doi:10.1146/annurev-genom-082509-141802

Becker KG, Swergold GD, Ozato K, Thayer RE (1993) Binding of the ubiquitous nuclear transcription factor YY1 to a cis regulatory sequence in the human LINE-1 transposable element. Human molecular genetics 2 (10):1697-1702

Belancio VP, Hedges DJ, Deininger P (2006) LINE-1 RNA splicing and influences on mammalian gene expression. Nucleic acids research 34 (5):1512-1521. doi:10.1093/nar/gkl027

Belancio VP, Hedges DJ, Deininger P (2008) Mammalian non-LTR retrotransposons: for better or worse, in sickness and in health. Genome research 18 (3):343-358. doi:10.1101/gr.5558208

Belancio VP, Roy-Engel AM, Deininger PL (2010) All y'all need to know 'bout retroelements in cancer. Semin Cancer Biol 20 (4):200-210. doi:10.1016/j.semcancer.2010.06.001

Bernstein E, Caudy AA, Hammond SM, Hannon GJ (2001) Role for a bidentate ribonuclease in the initiation step of RNA interference. Nature 409 (6818):363-366. doi:10.1038/35053110

Bibikova M, Beumer K, Trautman JK, Carroll D (2003) Enhancing gene targeting with designed zinc finger nucleases. Science 300 (5620):764. doi:10.1126/science.1079512

Bibikova M, Carroll D, Segal DJ, Trautman JK, Smith J, Kim YG, Chandrasegaran S (2001) Stimulation of homologous recombination through targeted cleavage by chimeric nucleases. Molecular and cellular biology 21 (1):289-297. doi:10.1128/MCB.21.1.289-297.2001

Bibikova M, Golic M, Golic KG, Carroll D (2002) Targeted chromosomal cleavage and mutagenesis in Drosophila using zinc-finger nucleases. Genetics 161 (3):1169-1175

Bird AP, Wolffe AP (1999) Methylation-induced repression--belts, braces, and chromatin. Cell 99 (5):451-454

Bishop KN, Holmes RK, Sheehy AM, Davidson NO, Cho SJ, Malim MH (2004) Cytidine deamination of retroviral DNA by diverse APOBEC proteins. Current biology : CB 14 (15):1392-1396. doi:10.1016/j.cub.2004.06.057

Block WD, Yu Y, Merkle D, Gifford JL, Ding Q, Meek K, Lees-Miller SP (2004) Autophosphorylation-dependent remodeling of the DNA-dependent protein kinase catalytic subunit regulates ligation of DNA ends. Nucleic acids research 32 (14):4351-4357. doi:10.1093/nar/gkh761

Boeke JD (1997) LINEs and Alus--the polyA connection. Nature genetics 16 (1):6-7. doi:10.1038/ng0597-6

Boshart M, Weber F, Jahn G, Dorsch-Hasler K, Fleckenstein B, Schaffner W (1985) A very strong enhancer is located upstream of an immediate early gene of human cytomegalovirus. Cell 41 (2):521-530

Brouha B, Meischl C, Ostertag E, de Boer M, Zhang Y, Neijens H, Roos D, Kazazian HH, Jr. (2002) Evidence consistent with human L1 retrotransposition in maternal meiosis I. American journal of human genetics 71 (2):327-336. doi:10.1086/341722

Brouha B, Schustak J, Badge RM, Lutz-Prigge S, Farley AH, Moran JV, Kazazian HH, Jr. (2003) Hot L1s account for the bulk of retrotransposition in the human population. Proceedings of the National Academy of Sciences of the United States of America 100 (9):5280-5285. doi:10.1073/pnas.0831042100

Brummelkamp TR, Bernards R, Agami R (2002) A system for stable expression of short interfering RNAs in mammalian cells. Science 296 (5567):550-553. doi:10.1126/science.1068999

Buck D, Malivert L, de Chasseval R, Barraud A, Fondaneche MC, Sanal O, Plebani A, Stephan JL, Hufnagel M, le Deist F, Fischer A, Durandy A, de Villartay JP, Revy P (2006) Cernunnos, a novel nonhomologous end-joining factor, is mutated in human immunodeficiency with microcephaly. Cell 124 (2):287-299. doi:10.1016/j.cell.2005.12.030

Burke WD, Malik HS, Jones JP, Eickbush TH (1999) The domain structure and retrotransposition mechanism of R2 elements are conserved throughout arthropods. Molecular biology and evolution 16 (4):502-511

Busch D, Greiner C, Lewis K, Ford R, Adair G, Thompson L (1989) Summary of complementation groups of UV-sensitive CHO cell mutants isolated by large-scale screening. Mutagenesis 4 (5):349-354

Cal S, Tan KL, McGregor A, Connolly BA (1998) Conversion of bovine pancreatic DNase I to a repair endonuclease with a high selectivity for abasic sites. The EMBO journal 17 (23):7128-7138. doi:10.1093/emboj/17.23.7128

Caplen NJ, Parrish S, Imani F, Fire A, Morgan RA (2001) Specific inhibition of gene expression by small double-stranded RNAs in invertebrate and vertebrate systems. Proceedings of the National Academy of Sciences of the United States of America 98 (17):9742-9747. doi:10.1073/pnas.171251798

Cary RB, Chen F, Shen Z, Chen DJ (1998) A central region of Ku80 mediates interaction with Ku70 in vivo. Nucleic acids research 26 (4):974-979

Catalanotto C, Azzalin G, Macino G, Cogoni C (2000) Gene silencing in worms and fungi. Nature 404 (6775):245. doi:10.1038/35005169

Chan DW, Chen BP, Prithivirajsingh S, Kurimasa A, Story MD, Qin J, Chen DJ (2002) Autophosphorylation of the DNA-dependent protein kinase catalytic subunit is required for rejoining of DNA double-strand breaks. Genes & development 16 (18):2333-2338. doi:10.1101/gad.1015202

Chen JM, Stenson PD, Cooper DN, Ferec C (2005) A systematic analysis of LINE-1 endonuclease-dependent retrotranspositional events causing human genetic disease. Human genetics 117 (5):411-427. doi:10.1007/s00439-005-1321-0

Christensen S, Pont-Kingdon G, Carroll D (2000) Comparative studies of the endonucleases from two related Xenopus laevis retrotransposons, Tx1L and Tx2L: target site specificity and evolutionary implications. Genetica 110 (3):245-256

Cimprich KA, Cortez D (2008) ATR: an essential regulator of genome integrity. Nature reviews Molecular cell biology 9 (8):616-627. doi:10.1038/nrm2450

Cleaver JE (1968) Defective repair replication of DNA in xeroderma pigmentosum. Nature 218 (5142):652-656

Clerici M, Mantiero D, Lucchini G, Longhese MP (2005) The Saccharomyces cerevisiae Sae2 protein promotes resection and bridging of double strand break ends. The Journal of biological chemistry 280 (46):38631-38638. doi:10.1074/jbc.M508339200

Cliby WA, Roberts CJ, Cimprich KA, Stringer CM, Lamb JR, Schreiber SL, Friend SH (1998) Overexpression of a kinase-inactive ATR protein causes sensitivity to DNA-

damaging agents and defects in cell cycle checkpoints. The EMBO journal 17 (1):159-169. doi:10.1093/emboj/17.1.159

Coffin JM (1996) HIV viral dynamics. Aids 10 Suppl 3:S75-84

Collis SJ, Swartz MJ, Nelson WG, DeWeese TL (2003) Enhanced radiation and chemotherapy-mediated cell killing of human cancer cells by small inhibitory RNA silencing of DNA repair factors. Cancer research 63 (7):1550-1554

Conway AB, Lynch TW, Zhang Y, Fortin GS, Fung CW, Symington LS, Rice PA (2004) Crystal structure of a Rad51 filament. Nature structural & molecular biology 11 (8):791-796. doi:10.1038/nsmb795

Cook PR, Jones CE, Furano AV (2015) Phosphorylation of ORF1p is required for L1 retrotransposition. Proceedings of the National Academy of Sciences of the United States of America 112 (14):4298-4303. doi:10.1073/pnas.1416869112

Cordaux R, Batzer MA (2009) The impact of retrotransposons on human genome evolution. Nature reviews Genetics 10 (10):691-703. doi:10.1038/nrg2640

Corneo B, Wendland RL, Deriano L, Cui X, Klein IA, Wong SY, Arnal S, Holub AJ, Weller GR, Pancake BA, Shah S, Brandt VL, Meek K, Roth DB (2007) Rag mutations reveal robust alternative end joining. Nature 449 (7161):483-486. doi:10.1038/nature06168

Cost GJ, Boeke JD (1998) Targeting of human retrotransposon integration is directed by the specificity of the L1 endonuclease for regions of unusual DNA structure. Biochemistry 37 (51):18081-18093

Cost GJ, Feng Q, Jacquier A, Boeke JD (2002) Human L1 element target-primed reverse transcription in vitro. The EMBO journal 21 (21):5899-5910

Cost GJ, Golding A, Schlissel MS, Boeke JD (2001) Target DNA chromatinization modulates nicking by L1 endonuclease. Nucleic acids research 29 (2):573-577

Coufal NG, Garcia-Perez JL, Peng GE, Marchetto MC, Muotri AR, Mu Y, Carson CT, Macia A, Moran JV, Gage FH (2011) Ataxia telangiectasia mutated (ATM) modulates long interspersed element-1 (L1) retrotransposition in human neural stem cells. Proceedings of the National Academy of Sciences of the United States of America 108 (51):20382-20387. doi:10.1073/pnas.1100273108

Coufal NG, Garcia-Perez JL, Peng GE, Yeo GW, Mu Y, Lovci MT, Morell M, O'Shea KS, Moran JV, Gage FH (2009) L1 retrotransposition in human neural progenitor cells. Nature 460 (7259):1127-1131. doi:10.1038/nature08248

Crooks GE, Hon G, Chandonia JM, Brenner SE (2004) WebLogo: a sequence logo generator. Genome research 14 (6):1188-1190. doi:10.1101/gr.849004

Cruickshanks HA, Tufarelli C (2009) Isolation of cancer-specific chimeric transcripts induced by hypomethylation of the LINE-1 antisense promoter. Genomics 94 (6):397-406. doi:10.1016/j.ygeno.2009.08.013

Curcio MJ, Derbyshire KM (2003) The outs and ins of transposition: from mu to kangaroo. Nature reviews Molecular cell biology 4 (11):865-877. doi:10.1038/nrm1241

Damert A, Raiz J, Horn AV, Lower J, Wang H, Xing J, Batzer MA, Lower R, Schumann GG (2009) 5'-Transducing SVA retrotransposon groups spread efficiently throughout the human genome. Genome research 19 (11):1992-2008. doi:10.1101/gr.093435.109

Davies OR, Pellegrini L (2007) Interaction with the BRCA2 C terminus protects RAD51-DNA filaments from disassembly by BRC repeats. Nature structural & molecular biology 14 (6):475-483. doi:10.1038/nsmb1251

de Laat WL, Appeldoorn E, Jaspers NG, Hoeijmakers JH (1998) DNA structural elements required for ERCC1-XPF endonuclease activity. The Journal of biological chemistry 273 (14):7835-7842

De Weerd-Kastelein EA, Keijzer W, Bootsma D (1972) Genetic heterogeneity of xeroderma pigmentosum demonstrated by somatic cell hybridization. Nature: New biology 238 (81):80-83

Deininger PL, Batzer MA (1999) Alu repeats and human disease. Molecular genetics and metabolism 67 (3):183-193. doi:10.1006/mgme.1999.2864

Delecluse HJ, Hammerschmidt W (2000) The genetic approach to the Epstein-Barr virus: from basic virology to gene therapy. Molecular pathology : MP 53 (5):270-279

Demple B, Harrison L (1994) Repair of oxidative damage to DNA: enzymology and biology. Annual review of biochemistry 63:915-948. doi:10.1146/annurev.bi.63.070194.004411

Dewannieux M, Blaise S, Heidmann T (2005) Identification of a functional envelope protein from the HERV-K family of human endogenous retroviruses. Journal of virology 79 (24):15573-15577. doi:10.1128/JVI.79.24.15573-15577.2005

Dewannieux M, Esnault C, Heidmann T (2003) LINE-mediated retrotransposition of marked Alu sequences. Nature genetics 35 (1):41-48. doi:10.1038/ng1223

DiBiase SJ, Zeng ZC, Chen R, Hyslop T, Curran WJ, Jr., Iliakis G (2000) DNA-dependent protein kinase stimulates an independently active, nonhomologous, end-joining apparatus. Cancer research 60 (5):1245-1253

Durocher D, Jackson SP (2001) DNA-PK, ATM and ATR as sensors of DNA damage: variations on a theme? Current opinion in cell biology 13 (2):225-231

Dutko JA, Schafer A, Kenny AE, Cullen BR, Curcio MJ (2005) Inhibition of a yeast LTR retrotransposon by human APOBEC3 cytidine deaminases. Current biology : CB 15 (7):661-666. doi:10.1016/j.cub.2005.02.051

Eickbush TH (2002) Repair by retrotransposition. Nature genetics 31 (2):126-127. doi:10.1038/ng897

Elbashir SM, Harborth J, Lendeckel W, Yalcin A, Weber K, Tuschl T (2001) Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. Nature 411 (6836):494-498. doi:10.1038/35078107

Elliott B, Jasin M (2002) Double-strand breaks and translocations in cancer. Cellular and molecular life sciences : CMLS 59 (2):373-385

Ergun S, Buschmann C, Heukeshoven J, Dammann K, Schnieders F, Lauke H, Chalajour F, Kilic N, Stratling WH, Schumann GG (2004) Cell type-specific expression of LINE-1 open reading frames 1 and 2 in fetal and adult human tissues. The Journal of biological chemistry 279 (26):27753-27763. doi:10.1074/jbc.M312985200

Esashi F, Christ N, Gannon J, Liu Y, Hunt T, Jasin M, West SC (2005) CDK-dependent phosphorylation of BRCA2 as a regulatory mechanism for recombinational repair. Nature 434 (7033):598-604. doi:10.1038/nature03404

Esashi F, Galkin VE, Yu X, Egelman EH, West SC (2007) Stabilization of RAD51 nucleoprotein filaments by the C-terminal region of BRCA2. Nature structural & molecular biology 14 (6):468-474. doi:10.1038/nsmb1245

Esnault C, Heidmann O, Delebecque F, Dewannieux M, Ribet D, Hance AJ, Heidmann T, Schwartz O (2005) APOBEC3G cytidine deaminase inhibits retrotransposition of endogenous retroviruses. Nature 433 (7024):430-433. doi:10.1038/nature03238

Esnault C, Maestre J, Heidmann T (2000) Human LINE retrotransposons generate processed pseudogenes. Nature genetics 24 (4):363-367. doi:10.1038/74184

Eustermann S, Videler H, Yang JC, Cole PT, Gruszka D, Veprintsev D, Neuhaus D (2011) The DNA-binding domain of human PARP-1 interacts with DNA single-strand breaks as a monomer through its second zinc finger. Journal of molecular biology 407 (1):149-170. doi:10.1016/j.jmb.2011.01.034

Evans AR, Limp-Foster M, Kelley MR (2000) Going APE over ref-1. Mutation research 461 (2):83-108

Fanslow DA, Wirt SE, Barker JC, Connelly JP, Porteus MH, Dann CT (2014) Genome editing in mouse spermatogonial stem/progenitor cells using engineered nucleases. PloS one 9 (11):e112652. doi:10.1371/journal.pone.0112652

Feng Q, Moran JV, Kazazian HH, Jr., Boeke JD (1996) Human L1 retrotransposon encodes a conserved endonuclease required for retrotransposition. Cell 87 (5):905-916

Feng Q, Schumann G, Boeke JD (1998) Retrotransposon R1Bm endonuclease cleaves the target sequence. Proceedings of the National Academy of Sciences of the United States of America 95 (5):2083-2088

Filippov V, Solovyev V, Filippova M, Gill SS (2000) A novel type of RNase III family proteins in eukaryotes. Gene 245 (1):213-221

Fiorentini P, Huang KN, Tishkoff DX, Kolodner RD, Symington LS (1997) Exonuclease I of Saccharomyces cerevisiae functions in mitotic recombination in vivo and in vitro. Molecular and cellular biology 17 (5):2764-2773

Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC (1998) Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. Nature 391 (6669):806-811. doi:10.1038/35888

Freeman JD, Goodchild NL, Mager DL (1994) A modified indicator gene for selection of retrotransposition events in mammalian cells. BioTechniques 17 (1):46, 48-49, 52

Garcia MA, Meurs EF, Esteban M (2007) The dsRNA protein kinase PKR: virus and cell control. Biochimie 89 (6-7):799-811. doi:10.1016/j.biochi.2007.03.001

Gasior SL, Roy-Engel AM, Deininger PL (2008) ERCC1/XPF limits L1 retrotransposition. DNA repair 7 (6):983-989. doi:10.1016/j.dnarep.2008.02.006

Gasior SL, Wakeman TP, Xu B, Deininger PL (2006) The human LINE-1 retrotransposon creates DNA double-strand breaks. Journal of molecular biology 357 (5):1383-1393. doi:10.1016/j.jmb.2006.01.089

Gasior SL, Wong AK, Kora Y, Shinohara A, Bishop DK (1998) Rad52 associates with RPA and functions with rad55 and rad57 to assemble meiotic recombination complexes. Genes & development 12 (14):2208-2221

Gilbert N, Lutz S, Morrish TA, Moran JV (2005) Multiple fates of L1 retrotransposition intermediates in cultured human cells. Molecular and cellular biology 25 (17):7780-7795. doi:10.1128/MCB.25.17.7780-7795.2005

Gilbert N, Lutz-Prigge S, Moran JV (2002) Genomic deletions created upon LINE-1 retrotransposition. Cell 110 (3):315-325

Gilley D, Tanaka H, Hande MP, Kurimasa A, Li GC, Oshimura M, Chen DJ (2001) DNA-PKcs is critical for telomere capping. Proceedings of the National Academy of Sciences of the United States of America 98 (26):15084-15088. doi:10.1073/pnas.261574698

Gonzalo S, Garcia-Cao M, Fraga MF, Schotta G, Peters AH, Cotter SE, Eguia R, Dean DC, Esteller M, Jenuwein T, Blasco MA (2005) Role of the RB1 family in stabilizing histone methylation at constitutive heterochromatin. Nature cell biology 7 (4):420-428. doi:10.1038/ncb1235

Goodier JL, Kazazian HH, Jr. (2008) Retrotransposons revisited: the restraint and rehabilitation of parasites. Cell 135 (1):23-35. doi:10.1016/j.cell.2008.09.022

Goodier JL, Ostertag EM, Kazazian HH, Jr. (2000) Transduction of 3'-flanking sequences is common in L1 retrotransposition. Human molecular genetics 9 (4):653-657

Gorman MA, Morera S, Rothwell DG, de La Fortelle E, Mol CD, Tainer JA, Hickson ID, Freemont PS (1997) The crystal structure of the human DNA repair endonuclease HAP1 suggests the recognition of extra-helical deoxyribose at DNA abasic sites. The EMBO journal 16 (21):6548-6558. doi:10.1093/emboj/16.21.6548

Gossen M, Bujard H (1992) Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. Proceedings of the National Academy of Sciences of the United States of America 89 (12):5547-5551

Gossen M, Freundlieb S, Bender G, Muller G, Hillen W, Bujard H (1995) Transcriptional activation by tetracyclines in mammalian cells. Science 268 (5218):1766-1769

Goytisolo FA, Samper E, Edmonson S, Taccioli GE, Blasco MA (2001) The absence of the dna-dependent protein kinase catalytic subunit in mice results in anaphase bridges and in increased telomeric fusions with normal telomere length and G-strand overhang. Molecular and cellular biology 21 (11):3642-3651. doi:10.1128/MCB.21.11.3642-3651.2001

Grandi FC, Rosser JM, An W (2013) LINE-1-derived poly(A) microsatellites undergo rapid shortening and create somatic and germline mosaicism in mice. Molecular biology and evolution 30 (3):503-512. doi:10.1093/molbev/mss251

Grimaldi G, Skowronski J, Singer MF (1984) Defining the beginning and end of KpnI family segments. The EMBO journal 3 (8):1753-1759

Hacein-Bey-Abina S, Garrigue A, Wang GP, Soulier J, Lim A, Morillon E, Clappier E, Caccavelli L, Delabesse E, Beldjord K, Asnafi V, MacIntyre E, Dal Cortivo L, Radford I, Brousse N, Sigaux F, Moshous D, Hauer J, Borkhardt A, Belohradsky BH, Wintergerst U, Velez MC, Leiva L, Sorensen R, Wulffraat N, Blanche S, Bushman FD, Fischer A, Cavazzana-Calvo M (2008) Insertional oncogenesis in 4 patients after retrovirus-mediated gene therapy of SCID-X1. The Journal of clinical investigation 118 (9):3132-3142. doi:10.1172/JCI35700

Hacein-Bey-Abina S, Von Kalle C, Schmidt M, McCormack MP, Wulffraat N, Leboulch P, Lim A, Osborne CS, Pawliuk R, Morillon E, Sorensen R, Forster A, Fraser P, Cohen JI, de Saint Basile G, Alexander I, Wintergerst U, Frebourg T, Aurias A, Stoppa-Lyonnet D, Romana S, Radford-Weiss I, Gross F, Valensi F, Delabesse E, Macintyre E, Sigaux F, Soulier J, Leiva LE, Wissler M, Prinz C, Rabbitts TH, Le Deist F, Fischer A, Cavazzana-Calvo M (2003) LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1. Science 302 (5644):415-419. doi:10.1126/science.1088547

Hackett PB, Largaespada DA, Cooper LJ (2010) A transposon and transposase system for human application. Molecular therapy : the journal of the American Society of Gene Therapy 18 (4):674-683. doi:10.1038/mt.2010.2

Hamilton AJ, Baulcombe DC (1999) A species of small antisense RNA in posttranscriptional gene silencing in plants. Science 286 (5441):950-952

Hammond SM, Bernstein E, Beach D, Hannon GJ (2000) An RNA-directed nuclease mediates post-transcriptional gene silencing in Drosophila cells. Nature 404 (6775):293-296. doi:10.1038/35005107

Hampf M, Gossen M (2007) Promoter crosstalk effects on gene expression. Journal of molecular biology 365 (4):911-920. doi:10.1016/j.jmb.2006.10.009

Han JS, Boeke JD (2004) A highly active synthetic mammalian retrotransposon. Nature 429 (6989):314-318. doi:10.1038/nature02535

Hancks DC, Kazazian HH, Jr. (2012) Active human retrotransposons: variation and disease. Current opinion in genetics & development 22 (3):191-203. doi:10.1016/j.gde.2012.02.006

Hancks DC, Kazazian HH, Jr. (2016) Roles for retrotransposon insertions in human disease. Mobile DNA 7:9. doi:10.1186/s13100-016-0065-9

Hannon GJ (2002) RNA interference. Nature 418 (6894):244-251. doi:10.1038/418244a

Haoudi A, Semmes OJ, Mason JM, Cannon RE (2004) Retrotransposition-Competent Human LINE-1 Induces Apoptosis in Cancer Cells With Intact p53. Journal of biomedicine & biotechnology 2004 (4):185-194. doi:10.1155/S1110724304403131

Harris CR, Dewan A, Zupnick A, Normart R, Gabriel A, Prives C, Levine AJ, Hoh J (2009) p53 responsive elements in human retrotransposons. Oncogene 28 (44):3857-3865. doi:10.1038/onc.2009.246

Hayashi Y, Kajikawa M, Matsumoto T, Okada N (2014) Mechanism by which a LINE protein recognizes its 3' tail RNA. Nucleic acids research 42 (16):10605-10617. doi:10.1093/nar/gku753

Hedges RW, Jacob AE (1974) Transposition of ampicillin resistance from RP4 to other replicons. Molecular & general genetics : MGG 132 (1):31-40

Hendrich B, Bird A (1998) Identification and characterization of a family of mammalian methyl-CpG binding proteins. Molecular and cellular biology 18 (11):6538-6547

Herskowitz I (1987) Functional inactivation of genes by dominant negative mutations. Nature 329 (6136):219-222. doi:10.1038/329219a0

Hoeijmakers JH (2001) Genome maintenance mechanisms for preventing cancer. Nature 411 (6835):366-374. doi:10.1038/35077232

Holen T, Amarzguioui M, Wiiger MT, Babaie E, Prydz H (2002) Positional effects of short interfering RNAs targeting the human coagulation trigger Tissue Factor. Nucleic acids research 30 (8):1757-1766

Holmes EC (2007) Ancient lentiviruses leave their mark. Proceedings of the National Academy of Sciences of the United States of America 104 (15):6095-6096. doi:10.1073/pnas.0701578104

Horn AV, Klawitter S, Held U, Berger A, Vasudevan AA, Bock A, Hofmann H, Hanschmann KM, Trosemeier JH, Flory E, Jabulowsky RA, Han JS, Lower J, Lower R, Munk C, Schumann GG (2014) Human LINE-1 restriction by APOBEC3C is deaminase independent and mediated by an ORF1p interaction that affects LINE reverse transcriptase activity. Nucleic acids research 42 (1):396-416. doi:10.1093/nar/gkt898

Iliakis G, Wang H, Perrault AR, Boecker W, Rosidi B, Windhofer F, Wu W, Guan J, Terzoudi G, Pantelias G (2004) Mechanisms of DNA double strand break repair and chromosome aberration formation. Cytogenetic and genome research 104 (1-4):14-20. doi:10.1159/000077461

Ivics Z, Izsvak Z (2010) The expanding universe of transposon technologies for gene and cell engineering. Mobile DNA 1 (1):25. doi:10.1186/1759-8753-1-25

Izsvak Z, Stuwe EE, Fiedler D, Katzer A, Jeggo PA, Ivics Z (2004) Healing the wounds inflicted by sleeping beauty transposition by double-strand break repair in mammalian somatic cells. Molecular cell 13 (2):279-290

Jurka J (1997) Sequence patterns indicate an enzymatic involvement in integration of mammalian retroposons. Proceedings of the National Academy of Sciences of the United States of America 94 (5):1872-1877

Jurka J, Pethiyagoda C (1995) Simple repetitive DNA sequences from primates: compilation and analysis. Journal of molecular evolution 40 (2):120-126

Kabotyanski EB, Gomelsky L, Han JO, Stamato TD, Roth DB (1998) Double-strand break repair in Ku86- and XRCC4-deficient cells. Nucleic acids research 26 (23):5333-5342

Kaludov NK, Wolffe AP (2000) MeCP2 driven transcriptional repression in vitro: selectivity for methylated DNA, action at a distance and contacts with the basal transcription machinery. Nucleic acids research 28 (9):1921-1928

Kano H, Godoy I, Courtney C, Vetter MR, Gerton GL, Ostertag EM, Kazazian HH, Jr. (2009) L1 retrotransposition occurs mainly in embryogenesis and creates somatic mosaicism. Genes & development 23 (11):1303-1312. doi:10.1101/gad.1803909

Kennedy RD, D'Andrea AD (2005) The Fanconi Anemia/BRCA pathway: new faces in the crowd. Genes & development 19 (24):2925-2940. doi:10.1101/gad.1370505

Kent WJ (2002) BLAT--the BLAST-like alignment tool. Genome research 12 (4):656-664. doi:10.1101/gr.229202. Article published online before March 2002

Khan H, Smit A, Boissinot S (2006) Molecular evolution and tempo of amplification of human LINE-1 retrotransposons since the origin of primates. Genome research 16 (1):78-87. doi:10.1101/gr.4001406

Khanna KK, Jackson SP (2001) DNA double-strand breaks: signaling, repair and the cancer connection. Nature genetics 27 (3):247-254. doi:10.1038/85798

Kim YG, Cha J, Chandrasegaran S (1996) Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. Proceedings of the National Academy of Sciences of the United States of America 93 (3):1156-1160

Kimberland ML, Divoky V, Prchal J, Schwahn U, Berger W, Kazazian HH, Jr. (1999) Full-length human L1 insertions retain the capacity for high frequency retrotransposition in cultured cells. Human molecular genetics 8 (8):1557-1560

Kines KJ, Sokolowski M, deHaro DL, Christian CM, Baddoo M, Smither ME, Belancio VP (2016) The endonuclease domain of the LINE-1 ORF2 protein can tolerate multiple mutations. Mobile DNA 7:8. doi:10.1186/s13100-016-0064-x

Kinomoto M, Kanno T, Shimura M, Ishizaka Y, Kojima A, Kurata T, Sata T, Tokunaga K (2007) All APOBEC3 family proteins differentially inhibit LINE-1 retrotransposition. Nucleic acids research 35 (9):2955-2964. doi:10.1093/nar/gkm181

Klawitter S, Fuchs NV, Upton KR, Munoz-Lopez M, Shukla R, Wang J, Garcia-Canadas M, Lopez-Ruiz C, Gerhardt DJ, Sebe A, Grabundzija I, Merkert S, Gerdes P, Pulgarin JA, Bock A, Held U, Witthuhn A, Haase A, Sarkadi B, Lower J, Wolvetang EJ, Martin U, Ivics Z, Izsvak Z, Garcia-Perez JL, Faulkner GJ, Schumann GG (2016) Reprogramming triggers endogenous L1 and Alu retrotransposition in human induced pluripotent stem cells. Nat Commun 7:10286. doi:10.1038/ncomms10286

Kopera HC, Moldovan JB, Morrish TA, Garcia-Perez JL, Moran JV (2011) Similarities between long interspersed element-1 (LINE-1) reverse transcriptase and telomerase. Proceedings of the National Academy of Sciences of the United States of America 108 (51):20345-20350. doi:10.1073/pnas.1100275108

Krishna S, Wagener BM, Liu HP, Lo YC, Sterk R, Petrini JH, Nickoloff JA (2007) Mre11 and Ku regulation of double-strand break repair by gene conversion and break-induced replication. DNA repair 6 (6):797-808. doi:10.1016/j.dnarep.2007.01.006

Kroutter EN, Belancio VP, Wagstaff BJ, Roy-Engel AM (2009) The RNA polymerase dictates ORF1 requirement and timing of LINE and SINE retrotransposition. PLoS genetics 5 (4):e1000458. doi:10.1371/journal.pgen.1000458

Kubo S, Seleme MC, Soifer HS, Perez JL, Moran JV, Kazazian HH, Jr., Kasahara N (2006) L1 retrotransposition in nondividing and primary human somatic cells. Proceedings of the National Academy of Sciences of the United States of America 103 (21):8036-8041. doi:10.1073/pnas.0601954103

Kuff EL, Smith LA, Lueders KK (1981) Intracisternal A-particle genes in Mus musculus: a conserved family of retrovirus-like elements. Molecular and cellular biology 1 (3):216-227

Kuo KW, Sheu HM, Huang YS, Leung WC (1998) Expression of transposon LINE-1 is relatively human-specific and function of the transcripts may be proliferation-essential. Biochemical and biophysical research communications 253 (3):566-570. doi:10.1006/bbrc.1998.9811

Kupper JH, de Murcia G, Burkle A (1990) Inhibition of poly(ADP-ribosyl)ation by overexpressing the poly(ADP-ribose) polymerase DNA-binding domain in mammalian cells. The Journal of biological chemistry 265 (31):18721-18724

Kurnosov AA, Ustyugova SV, Nazarov VI, Minervina AA, Komkov AY, Shugay M, Pogorelyy MV, Khodosevich KV, Mamedov IZ, Lebedev YB (2015) The evidence for increased L1 activity in the site of human adult brain neurogenesis. PloS one 10 (2):e0117854. doi:10.1371/journal.pone.0117854

Kurose K, Hata K, Hattori M, Sakaki Y (1995) RNA polymerase III dependence of the human L1 promoter and possible participation of the RNA polymerase II factor YY1 in the RNA polymerase III transcription system. Nucleic acids research 23 (18):3704-3709

Kuwabara T, Hsieh J, Muotri A, Yeo G, Warashina M, Lie DC, Moore L, Nakashima K, Asashima M, Gage FH (2009) Wnt-mediated activation of NeuroD1 and retro-elements during adult neurogenesis. Nature neuroscience 12 (9):1097-1105. doi:10.1038/nn.2360

Laderoute MP, Larocque LJ, Giulivi A, Diaz-Mitoma F (2015) Further Evidence that Human Endogenous Retrovirus K102 is a Replication Competent Foamy Virus that may Antagonize HIV-1 Replication. Open AIDS J 9:112-122. doi:10.2174/1874613601509010112

Lahm A, Suck D (1991) DNase I-induced DNA conformation. 2 A structure of a DNase I-octamer complex. Journal of molecular biology 222 (3):645-667

Lambert S, Lopez BS (2000) Characterization of mammalian RAD51 double strand break repair using non-lethal dominant-negative forms. The EMBO journal 19 (12):3090-3099. doi:10.1093/emboj/19.12.3090

Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, Funke R, Gage D, Harris K, Heaford A, Howland J, Kann L, Lehoczky J, LeVine R, McEwan P, McKernan K, Meldrim J, Mesirov JP, Miranda C, Morris W, Naylor J, Raymond C, Rosetti M, Santos R, Sheridan A, Sougnez C, Stange-Thomann N, Stojanovic N, Subramanian A, Wyman D, Rogers J, Sulston J, Ainscough R, Beck S, Bentley D, Burton J, Clee C, Carter N, Coulson A, Deadman R, Deloukas P, Dunham A, Dunham I, Durbin R, French L, Grafham D, Gregory S, Hubbard T, Humphray S, Hunt A, Jones M, Lloyd C, McMurray A, Matthews L, Mercer S, Milne S, Mullikin JC, Mungall A, Plumb R, Ross M, Shownkeen R, Sims S, Waterston RH, Wilson RK, Hillier LW, McPherson JD, Marra MA, Mardis ER, Fulton LA, Chinwalla AT, Pepin KH, Gish WR, Chissoe SL, Wendl MC, Delehaunty KD, Miner TL, Delehaunty A, Kramer JB, Cook LL, Fulton RS, Johnson DL, Minx PJ, Clifton SW, Hawkins T, Branscomb E, Predki P, Richardson P, Wenning S, Slezak T, Doggett N, Cheng JF, Olsen A, Lucas S, Elkin C, Uberbacher E, Frazier M, Gibbs RA, Muzny DM, Scherer SE, Bouck JB, Sodergren EJ, Worley KC, Rives CM, Gorrell JH, Metzker ML, Naylor SL, Kucherlapati RS, Nelson DL, Weinstock GM, Sakaki Y, Fujiyama A, Hattori M, Yada T, Toyoda A, Itoh T, Kawagoe C, Watanabe H, Totoki Y, Taylor T, Weissenbach J, Heilig R, Saurin W, Artiguenave F, Brottier P, Bruls T, Pelletier E, Robert C, Wincker P, Smith DR, Doucette-Stamm L, Rubenfield M, Weinstock K, Lee HM, Dubois J, Rosenthal A, Platzer M, Nyakatura G, Taudien S, Rump A, Yang H, Yu J, Wang J, Huang G, Gu J, Hood L, Rowen L, Madan A, Qin S, Davis RW, Federspiel NA, Abola AP, Proctor MJ, Myers RM, Schmutz J, Dickson M, Grimwood J, Cox DR, Olson MV, Kaul R, Raymond C, Shimizu N, Kawasaki K, Minoshima S, Evans GA, Athanasiou M, Schultz R, Roe BA, Chen F, Pan H, Ramser J, Lehrach H, Reinhardt R, McCombie WR, de la Bastide M, Dedhia N, Blocker H, Hornischer K, Nordsiek G, Agarwala R, Aravind L, Bailey JA, Bateman A, Batzoglou S, Birney E, Bork P, Brown DG, Burge CB, Cerutti L, Chen HC, Church D, Clamp M, Copley RR, Doerks T, Eddy SR, Eichler EE, Furey TS, Galagan J, Gilbert JG, Harmon C, Hayashizaki Y, Haussler D, Hermjakob H, Hokamp K, Jang W, Johnson LS, Jones TA, Kasif S, Kaspryzk A, Kennedy S, Kent WJ, Kitts P, Koonin EV, Korf I, Kulp D, Lancet D, Lowe TM, McLysaght A, Mikkelsen T, Moran JV, Mulder N, Pollara VJ, Ponting CP, Schuler G, Schultz J, Slater G, Smit AF, Stupka E, Szustakowski J, Thierry-Mieg D, Thierry-Mieg J, Wagner L, Wallis J, Wheeler R, Williams A, Wolf YI, Wolfe KH, Yang SP, Yeh RF, Collins F, Guyer MS, Peterson J, Felsenfeld A, Wetterstrand KA, Patrinos A, Morgan MJ, de Jong P, Catanese JJ, Osoegawa K, Shizuya H, Choi S, Chen YJ, International Human Genome Sequencing C (2001) Initial sequencing and analysis of the human genome. Nature 409 (6822):860-921. doi:10.1038/35057062

Lees-Miller SP, Meek K (2003) Repair of DNA double strand breaks by non-homologous end joining. Biochimie 85 (11):1161-1173

Lempiainen H, Halazonetis TD (2009) Emerging common themes in regulation of PIKKs and PI3Ks. The EMBO journal 28 (20):3067-3073. doi:10.1038/emboj.2009.281

Lewis JD, Meehan RR, Henzel WJ, Maurer-Fogy I, Jeppesen P, Klein F, Bird A (1992) Purification, sequence, and cellular localization of a novel chromosomal protein that binds to methylated DNA. Cell 69 (6):905-914

Li X, Zhang XP, Solinger JA, Kiianitsa K, Yu X, Egelman EH, Heyer WD (2007) Rad51 and Rad54 ATPase activities are both required to modulate Rad51-dsDNA filament dynamics. Nucleic acids research 35 (12):4124-4140. doi:10.1093/nar/gkm412

Liddament MT, Brown WL, Schumacher AJ, Harris RS (2004) APOBEC3F properties and hypermutation preferences indicate activity against HIV-1 in vivo. Current biology : CB 14 (15):1385-1391. doi:10.1016/j.cub.2004.06.050

Lisby M, Rothstein R (2004) DNA repair: keeping it together. Current biology : CB 14 (23):R994-996. doi:10.1016/j.cub.2004.11.020

Llorca O (2007) Electron microscopy reconstructions of DNA repair complexes. Current opinion in structural biology 17 (2):215-220. doi:10.1016/j.sbi.2007.03.009

Lloyd A, Plaisier CL, Carroll D, Drews GN (2005) Targeted mutagenesis using zinc-finger nucleases in Arabidopsis. Proceedings of the National Academy of Sciences of the United States of America 102 (6):2232-2237. doi:10.1073/pnas.0409339102

Lobrich M, Jeggo PA (2005) The two edges of the ATM sword: co-operation between repair and checkpoint functions. Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology 76 (2):112-118. doi:10.1016/j.radonc.2005.06.027

Lovejoy CA, Cortez D (2009) Common mechanisms of PIKK regulation. DNA repair 8 (9):1004-1008. doi:10.1016/j.dnarep.2009.04.006

Lovsin N, Gubensek F, Kordi D (2001) Evolutionary dynamics in a novel L2 clade of non-LTR retrotransposons in Deuterostomia. Molecular biology and evolution 18 (12):2213-2224

Lower R, Boller K, Hasenmaier B, Korbmacher C, Muller-Lantzsch N, Lower J, Kurth R (1993) Identification of human endogenous retroviruses with complex mRNA expression and particle formation. Proceedings of the National Academy of Sciences of the United States of America 90 (10):4480-4484

Lower R, Lower J, Kurth R (1996) The viruses in all of us: characteristics and biological significance of human endogenous retrovirus sequences. Proceedings of the National Academy of Sciences of the United States of America 93 (11):5177-5184

Luan DD, Korman MH, Jakubczak JL, Eickbush TH (1993) Reverse transcription of R2Bm RNA is primed by a nick at the chromosomal target site: a mechanism for non-LTR retrotransposition. Cell 72 (4):595-605

Lydeard JR, Jain S, Yamaguchi M, Haber JE (2007) Break-induced replication and telomerase-independent telomere maintenance require Pol32. Nature 448 (7155):820-823. doi:10.1038/nature06047

Mahieux R, Suspene R, Delebecque F, Henry M, Schwartz O, Wain-Hobson S, Vartanian JP (2005) Extensive editing of a small fraction of human T-cell leukemia virus type 1 genomes by four APOBEC3 cytidine deaminases. The Journal of general virology 86 (Pt 9):2489-2494. doi:10.1099/vir.0.80973-0

Maita N, Anzai T, Aoyagi H, Mizuno H, Fujiwara H (2004) Crystal structure of the endonuclease domain encoded by the telomere-specific long interspersed nuclear element, TRAS1. The Journal of biological chemistry 279 (39):41067-41076. doi:10.1074/jbc.M406556200

Malik HS, Burke WD, Eickbush TH (1999) The age and evolution of non-LTR retrotransposable elements. Molecular biology and evolution 16 (6):793-805

Malik HS, Eickbush TH (2001) Phylogenetic analysis of ribonuclease H domains suggests a late, chimeric origin of LTR retrotransposable elements and retroviruses. Genome research 11 (7):1187-1197. doi:10.1101/gr.185101

Malkova A, Ivanov EL, Haber JE (1996) Double-strand break repair in the absence of RAD51 in yeast: a possible role for break-induced DNA replication. Proceedings of the National Academy of Sciences of the United States of America 93 (14):7131-7136

Mangeat B, Turelli P, Caron G, Friedli M, Perrin L, Trono D (2003) Broad antiretroviral defence by human APOBEC3G through lethal editing of nascent reverse transcripts. Nature 424 (6944):99-103. doi:10.1038/nature01709

Mangiacasale R, Pittoggi C, Sciamanna I, Careddu A, Mattei E, Lorenzini R, Travaglini L, Landriscina M, Barone C, Nervi C, Lavia P, Spadafora C (2003) Exposure of normal and transformed cells to nevirapine, a reverse transcriptase inhibitor, reduces cell growth and promotes differentiation. Oncogene 22 (18):2750-2761. doi:10.1038/sj.onc.1206354

Mao Z, Bozzella M, Seluanov A, Gorbunova V (2008) DNA repair by nonhomologous end joining and homologous recombination during cell cycle in human cells. Cell cycle 7 (18):2902-2906

Markovic J, Grdovic N, Dinic S, Karan-Djurasevic T, Uskokovic A, Arambasic J, Mihailovic M, Pavlovic S, Poznanovic G, Vidakovic M (2013) PARP-1 and YY1 are important novel regulators of CXCL12 gene transcription in rat pancreatic beta cells. PloS one 8 (3):e59679. doi:10.1371/journal.pone.0059679

Martin F, Maranon C, Olivares M, Alonso C, Lopez MC (1995) Characterization of a non-long terminal repeat retrotransposon cDNA (L1Tc) from Trypanosoma cruzi: homology of the first ORF with the ape family of DNA repair enzymes. Journal of molecular biology 247 (1):49-59. doi:10.1006/jmbi.1994.0121

Martin JN, Wolken N, Brown T, Dauer WT, Ehrlich ME, Gonzalez-Alegre P (2011) Lethal toxicity caused by expression of shRNA in the mouse striatum: implications for therapeutic design. Gene Ther 18 (7):666-673. doi:10.1038/gt.2011.10

Martin SL (2006) The ORF1 protein encoded by LINE-1: structure and function during L1 retrotransposition. Journal of biomedicine & biotechnology 2006 (1):45621. doi:10.1155/JBB/2006/45621

Martin SL, Bushman FD (2001) Nucleic acid chaperone activity of the ORF1 protein from the mouse LINE-1 retrotransposon. Molecular and cellular biology 21 (2):467-475. doi:10.1128/MCB.21.2.467-475.2001

Martin SL, Li WL, Furano AV, Boissinot S (2005) The structures of mouse and human L1 elements reflect their insertion mechanism. Cytogenetic and genome research 110 (1-4):223-228. doi:10.1159/000084956

Masters JR (2002) HeLa cells 50 years on: the good, the bad and the ugly. Nature reviews Cancer 2 (4):315-319. doi:10.1038/nrc775

Mathias SL, Scott AF, Kazazian HH, Jr., Boeke JD, Gabriel A (1991) Reverse transcriptase encoded by a human transposable element. Science 254 (5039):1808-1810

Mayford M, Bach ME, Huang YY, Wang L, Hawkins RD, Kandel ER (1996) Control of memory formation through regulated expression of a CaMKII transgene. Science 274 (5293):1678-1683

Mazin AV, Alexeev AA, Kowalczykowski SC (2003) A novel function of Rad54 protein. Stabilization of the Rad51 nucleoprotein filament. The Journal of biological chemistry 278 (16):14029-14036. doi:10.1074/jbc.M212779200

Mazin AV, Bornarth CJ, Solinger JA, Heyer WD, Kowalczykowski SC (2000) Rad54 protein is targeted to pairing loci by the Rad51 nucleoprotein filament. Molecular cell 6 (3):583-592

Milne GT, Weaver DT (1993) Dominant negative alleles of RAD52 reveal a DNA repair/recombination complex including Rad51 and Rad52. Genes & development 7 (9):1755-1765

Mimitou EP, Symington LS (2009) Nucleases and helicases take center stage in homologous recombination. Trends in biochemical sciences 34 (5):264-272. doi:10.1016/j.tibs.2009.01.010

Modesti M, Junop MS, Ghirlando R, van de Rakt M, Gellert M, Yang W, Kanaar R (2003) Tetramerization and DNA ligase IV interaction of the DNA double-strand break repair protein XRCC4 are mutually exclusive. Journal of molecular biology 334 (2):215-228

Mol CD, Hosfield DJ, Tainer JA (2000) Abasic site recognition by two apurinic/apyrimidinic endonuclease families in DNA base excision repair: the 3' ends justify the means. Mutation research 460 (3-4):211-229

Mol CD, Kuo CF, Thayer MM, Cunningham RP, Tainer JA (1995) Structure and function of the multifunctional DNA-repair enzyme exonuclease III. Nature 374 (6520):381-386. doi:10.1038/374381a0

Montoya-Durango DE, Liu Y, Teneng I, Kalbfleisch T, Lacy ME, Steffen MC, Ramos KS (2009) Epigenetic control of mammalian LINE-1 retrotransposon by retinoblastoma proteins. Mutation research 665 (1-2):20-28. doi:10.1016/j.mrfmmm.2009.02.011

Montoya-Durango DE, Ramos KA, Bojang P, Ruiz L, Ramos IN, Ramos KS (2016) LINE-1 silencing by retinoblastoma proteins is effected through the nucleosomal and remodeling deacetylase multiprotein complex. BMC Cancer 16:38. doi:10.1186/s12885-016-2068-9

Moran JV, DeBerardinis RJ, Kazazian HH, Jr. (1999) Exon shuffling by L1 retrotransposition. Science 283 (5407):1530-1534

Moran JV, Holmes SE, Naas TP, DeBerardinis RJ, Boeke JD, Kazazian HH, Jr. (1996) High frequency retrotransposition in cultured mammalian cells. Cell 87 (5):917-927

Mordes DA, Cortez D (2008) Activation of ATR and related PIKKs. Cell cycle 7 (18):2809-2812

Morrish TA, Garcia-Perez JL, Stamato TD, Taccioli GE, Sekiguchi J, Moran JV (2007) Endonuclease-independent LINE-1 retrotransposition at mammalian telomeres. Nature 446 (7132):208-212. doi:10.1038/nature05560

Morrish TA, Gilbert N, Myers JS, Vincent BJ, Stamato TD, Taccioli GE, Batzer MA, Moran JV (2002) DNA repair mediated by endonuclease-independent LINE-1 retrotransposition. Nature genetics 31 (2):159-165. doi:10.1038/ng898

Moynahan ME, Jasin M (1997) Loss of heterozygosity induced by a chromosomal double-strand break. Proceedings of the National Academy of Sciences of the United States of America 94 (17):8988-8993

Muckenfuss H, Hamdorf M, Held U, Perkovic M, Lower J, Cichutek K, Flory E, Schumann GG, Munk C (2006) APOBEC3 proteins inhibit human LINE-1 retrotransposition. The Journal of biological chemistry 281 (31):22161-22172. doi:10.1074/jbc.M601716200

Mueller-Lantzsch N, Sauter M, Weiskircher A, Kramer K, Best B, Buck M, Grasser F (1993) Human endogenous retroviral element K10 (HERV-K10) encodes a full-length gag homologous 73-kDa protein and a functional protease. AIDS research and human retroviruses 9 (4):343-350

Muotri AR, Chu VT, Marchetto MC, Deng W, Moran JV, Gage FH (2005) Somatic mosaicism in neuronal precursor cells mediated by L1 retrotransposition. Nature 435 (7044):903-910. doi:10.1038/nature03663

Muotri AR, Marchetto MC, Coufal NG, Oefner R, Yeo G, Nakashima K, Gage FH (2010) L1 retrotransposition in neurons is modulated by MeCP2. Nature 468 (7322):443-446. doi:10.1038/nature09544

Muotri AR, Zhao C, Marchetto MC, Gage FH (2009) Environmental influence on L1 retrotransposons in the adult hippocampus. Hippocampus 19 (10):1002-1007. doi:10.1002/hipo.20564

Muster T, Waltenberger A, Grassauer A, Hirschl S, Caucig P, Romirer I, Fodinger D, Seppele H, Schanab O, Magin-Lachmann C, Lower R, Jansen B, Pehamberger H, Wolff K (2003) An endogenous retrovirus derived from human melanoma cells. Cancer research 63 (24):8735-8741

Nan X, Campoy FJ, Bird A (1997) MeCP2 is a transcriptional repressor with abundant binding sites in genomic chromatin. Cell 88 (4):471-481

Ngo H, Tschudi C, Gull K, Ullu E (1998) Double-stranded RNA induces mRNA degradation in Trypanosoma brucei. Proceedings of the National Academy of Sciences of the United States of America 95 (25):14687-14692

Nickoloff JA, Sweetser DB, Clikeman JA, Khalsa GJ, Wheeler SL (1999) Multiple heterologies increase mitotic double-strand break-induced allelic gene conversion tract lengths in yeast. Genetics 153 (2):665-679

Noguchi C, Ishino H, Tsuge M, Fujimoto Y, Imamura M, Takahashi S, Chayama K (2005) G to A hypermutation of hepatitis B virus. Hepatology 41 (3):626-633. doi:10.1002/hep.20580

Nussenzweig A, Nussenzweig MC (2007) A backup DNA repair pathway moves to the forefront. Cell 131 (2):223-225. doi:10.1016/j.cell.2007.10.005

Nykanen A, Haley B, Zamore PD (2001) ATP requirements and small interfering RNA structure in the RNA interference pathway. Cell 107 (3):309-321

O'Donnell KA, An W, Schrum CT, Wheelan SJ, Boeke JD (2013) Controlled insertional mutagenesis using a LINE-1 (ORFeus) gene-trap mouse model. Proceedings of the National Academy of Sciences of the United States of America 110 (29):E2706-2713. doi:10.1073/pnas.1302504110

Ogawa T, Yu X, Shinohara A, Egelman EH (1993) Similarity of the yeast RAD51 filament to the bacterial RecA filament. Science 259 (5103):1896-1899

Ostertag EM, DeBerardinis RJ, Goodier JL, Zhang Y, Yang N, Gerton GL, Kazazian HH, Jr. (2002) A mouse model of human L1 retrotransposition. Nature genetics 32 (4):655-660. doi:10.1038/ng1022

Ostertag EM, Kazazian HH, Jr. (2001) Twin priming: a proposed mechanism for the creation of inversions in L1 retrotransposition. Genome research 11 (12):2059-2065. doi:10.1101/gr.205701

Ostertag EM, Prak ET, DeBerardinis RJ, Moran JV, Kazazian HH, Jr. (2000) Determination of L1 retrotransposition kinetics in cultured cells. Nucleic acids research 28 (6):1418-1423

Paddison PJ, Caudy AA, Bernstein E, Hannon GJ, Conklin DS (2002) Short hairpin RNAs (shRNAs) induce sequence-specific silencing in mammalian cells. Genes & development 16 (8):948-958. doi:10.1101/gad.981002

Pal-Bhadra M, Bhadra U, Birchler JA (1997) Cosuppression in Drosophila: gene silencing of Alcohol dehydrogenase by white-Adh transgenes is Polycomb dependent. Cell 90 (3):479-490

Pavletich NP, Pabo CO (1991) Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 A. Science 252 (5007):809-817

Pellegrini L, Venkitaraman A (2004) Emerging functions of BRCA2 in DNA recombination. Trends in biochemical sciences 29 (6):310-316. doi:10.1016/j.tibs.2004.04.009

Pellegrini L, Yu DS, Lo T, Anand S, Lee M, Blundell TL, Venkitaraman AR (2002) Insights into DNA recombination from the structure of a RAD51-BRCA2 complex. Nature 420 (6913):287-293. doi:10.1038/nature01230

Perrault R, Wang H, Wang M, Rosidi B, Iliakis G (2004) Backup pathways of NHEJ are suppressed by DNA-PK. Journal of cellular biochemistry 92 (4):781-794. doi:10.1002/jcb.20104

Pickeral OK, Makalowski W, Boguski MS, Boeke JD (2000) Frequent human genomic DNA transduction driven by LINE-1 retrotransposition. Genome research 10 (4):411-415

Pittoggi C, Sciamanna I, Mattei E, Beraldi R, Lobascio AM, Mai A, Quaglia MG, Lorenzini R, Spadafora C (2003) Role of endogenous reverse transcriptase in murine early embryo development. Molecular reproduction and development 66 (3):225-236. doi:10.1002/mrd.10349

Pizarro JG, Cristofari G (2016) Post-Transcriptional Control of LINE-1 Retrotransposition by Cellular Host Factors in Somatic Cells. Front Cell Dev Biol 4:14. doi:10.3389/fcell.2016.00014

Plank JL, Wu J, Hsieh TS (2006) Topoisomerase IIIalpha and Bloom's helicase can resolve a mobile double Holliday junction substrate through convergent branch migration. Proceedings of the National Academy of Sciences of the United States of America 103 (30):11118-11123. doi:10.1073/pnas.0604873103

Porteus MH (2006) Mammalian gene targeting with designed zinc finger nucleases. Molecular therapy : the journal of the American Society of Gene Therapy 13 (2):438-446. doi:10.1016/j.ymthe.2005.08.003

Porteus MH, Baltimore D (2003) Chimeric nucleases stimulate gene targeting in human cells. Science 300 (5620):763. doi:10.1126/science.1078395

Porteus MH, Carroll D (2005) Gene targeting using zinc finger nucleases. Nature biotechnology 23 (8):967-973. doi:10.1038/nbt1125

Prak ET, Dodson AW, Farkash EA, Kazazian HH, Jr. (2003) Tracking an embryonic L1 retrotransposition event. Proceedings of the National Academy of Sciences of the United States of America 100 (4):1832-1837. doi:10.1073/pnas.0337627100

Puga A, Barnes SJ, Dalton TP, Chang C, Knudsen ES, Maier MA (2000) Aromatic hydrocarbon receptor interaction with the retinoblastoma protein potentiates repression of E2F-dependent transcription and cell cycle arrest. The Journal of biological chemistry 275 (4):2943-2950

Qi Y, Starker CG, Zhang F, Baltes NJ, Voytas DF (2014) Tailor-made mutations in Arabidopsis using zinc finger nucleases. Methods Mol Biol 1062:193-209. doi:10.1007/978-1-62703-580-4_10

Quentin Y (1992) Origin of the Alu family: a family of Alu-like monomers gave birth to the left and the right arms of the Alu elements. Nucleic acids research 20 (13):3397-3401

Rahman SH, Kuehle J, Reimann C, Mlambo T, Alzubi J, Maeder ML, Riedel H, Fisch P, Cantz T, Rudolph C, Mussolino C, Joung JK, Schambach A, Cathomen T (2015) Rescue of DNA-PK Signaling and T-Cell Differentiation by Targeted Genome Editing in a prkdc Deficient iPSC Disease Model. PLoS genetics 11 (5):e1005239. doi:10.1371/journal.pgen.1005239

Raiz J, Damert A, Chira S, Held U, Klawitter S, Hamdorf M, Lower J, Stratling WH, Lower R, Schumann GG (2012) The non-autonomous retrotransposon SVA is trans-mobilized by the human LINE-1 protein machinery. Nucleic acids research 40 (4):1666-1683. doi:10.1093/nar/gkr863

Rao RN, Allen NE, Hobbs JN, Jr., Alborn WE, Jr., Kirst HA, Paschal JW (1983) Genetic and enzymatic basis of hygromycin B resistance in Escherichia coli. Antimicrob Agents Chemother 24 (5):689-695

Raper SE, Chirmule N, Lee FS, Wivel NA, Bagg A, Gao GP, Wilson JM, Batshaw ML (2003) Fatal systemic inflammatory response syndrome in a ornithine transcarbamylase deficient patient following adenoviral gene transfer. Molecular genetics and metabolism 80 (1-2):148-158

Ray DA, Feschotte C, Pagan HJ, Smith JD, Pritham EJ, Arensburger P, Atkinson PW, Craig NL (2008) Multiple waves of recent DNA transposon activity in the bat, Myotis lucifugus. Genome research 18 (5):717-728. doi:10.1101/gr.071886.107

Ray DA, Pagan HJ, Thompson ML, Stevens RD (2007) Bats with hATs: evidence for recent DNA transposon activity in genus Myotis. Molecular biology and evolution 24 (3):632-639. doi:10.1093/molbev/msl192

Reddy YV, Ding Q, Lees-Miller SP, Meek K, Ramsden DA (2004) Non-homologous end joining requires that the DNA-PK complex undergo an autophosphorylation-dependent rearrangement at DNA ends. The Journal of biological chemistry 279 (38):39408-39413. doi:10.1074/jbc.M406432200

Repanas K, Zingler N, Layer LE, Schumann GG, Perrakis A, Weichenrieder O (2007) Determinants for DNA target structure selectivity of the human LINE-1 retrotransposon endonuclease. Nucleic acids research 35 (14):4914-4926. doi:10.1093/nar/gkm516

Riballo E, Woodbine L, Stiff T, Walker SA, Goodarzi AA, Jeggo PA (2009) XLF-Cernunnos promotes DNA ligase IV-XRCC4 re-adenylation following ligation. Nucleic acids research 37 (2):482-492. doi:10.1093/nar/gkn957

Richardson C, Jasin M (2000) Frequent chromosomal translocations induced by DNA double-strand breaks. Nature 405 (6787):697-700. doi:10.1038/35015097

Richardson SR, Narvaiza I, Planegger RA, Weitzman MD, Moran JV (2014) APOBEC3A deaminates transiently exposed single-strand DNA during LINE-1 retrotransposition. Elife 3:e02008. doi:10.7554/eLife.02008

Rivera-Calzada A, Spagnolo L, Pearl LH, Llorca O (2007) Structural model of full-length human Ku70-Ku80 heterodimer and its recognition of DNA and DNA-PKcs. EMBO reports 8 (1):56-62. doi:10.1038/sj.embor.7400847

Robertson KD, Ait-Si-Ali S, Yokochi T, Wade PA, Jones PL, Wolffe AP (2000) DNMT1 forms a complex with Rb, E2F1 and HDAC1 and represses transcription from E2F-responsive promoters. Nature genetics 25 (3):338-342. doi:10.1038/77124

Roman-Gomez J, Jimenez-Velasco A, Agirre X, Cervantes F, Sanchez J, Garate L, Barrios M, Castillejo JA, Navarro G, Colomer D, Prosper F, Heiniger A, Torres A (2005) Promoter hypomethylation of the LINE-1 retrotransposable elements activates sense/antisense transcription and marks the progression of chronic myeloid leukemia. Oncogene 24 (48):7213-7223. doi:10.1038/sj.onc.1208866

Rosler C, Kock J, Kann M, Malim MH, Blum HE, Baumert TF, von Weizsacker F (2005) APOBEC-mediated interference with hepadnavirus production. Hepatology 42 (2):301-309. doi:10.1002/hep.20801

Rosler C, Kock J, Malim MH, Blum HE, von Weizsacker F (2004) Comment on "Inhibition of hepatitis B virus replication by APOBEC3G". Science 305 (5689):1403; author reply 1403. doi:10.1126/science.1100464

Sage JM, Gildemeister OS, Knight KL (2010) Discovery of a novel function for human Rad51: maintenance of the mitochondrial genome. The Journal of biological chemistry 285 (25):18984-18990. doi:10.1074/jbc.M109.099846

Sambrook J, Russell DW (2001) Molecular cloning : a laboratory manual. 3rd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

San Filippo J, Chi P, Sehorn MG, Etchin J, Krejci L, Sung P (2006) Recombination mediator and Rad51 targeting activities of a human BRCA2 polypeptide. The Journal of biological chemistry 281 (17):11649-11657. doi:10.1074/jbc.M601249200

Sanchez Alvarado A, Newmark PA (1999) Double-stranded RNA specifically disrupts gene expression during planarian regeneration. Proceedings of the National Academy of Sciences of the United States of America 96 (9):5049-5054

Saporito SM, Smith-White BJ, Cunningham RP (1988) Nucleotide sequence of the xth gene of Escherichia coli K-12. Journal of bacteriology 170 (10):4542-4547

Sasada A, Takaori-Kondo A, Shirakawa K, Kobayashi M, Abudu A, Hishizawa M, Imada K, Tanaka Y, Uchiyama T (2005) APOBEC3G targets human T-cell leukemia virus type 1. Retrovirology 2:32. doi:10.1186/1742-4690-2-32

Schumacher AJ, Nissley DV, Harris RS (2005) APOBEC3G hypermutates genomic DNA and inhibits Ty1 retrotransposition in yeast. Proceedings of the National Academy of Sciences of the United States of America 102 (28):9854-9859. doi:10.1073/pnas.0501694102

Schwarz E, Freese UK, Gissmann L, Mayer W, Roggenbuck B, Stremlau A, zur Hausen H (1985) Structure and transcription of human papillomavirus sequences in cervical carcinoma cells. Nature 314 (6006):111-114

Scott AF, Schmeckpeper BJ, Abdelrazik M, Comey CT, O'Hara B, Rossiter JP, Cooley T, Heath P, Smith KD, Margolet L (1987) Origin of the human L1 elements: proposed progenitor genes deduced from a consensus DNA sequence. Genomics 1 (2):113-125

Scott EC, Gardner EJ, Masood A, Chuang NT, Vertino PM, Devine SE (2016) A hot L1 retrotransposon evades somatic repression and initiates human colorectal cancer. Genome research 26 (6):745-755. doi:10.1101/gr.201814.115

Sheehy AM, Gaddis NC, Choi JD, Malim MH (2002) Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein. Nature 418 (6898):646-650. doi:10.1038/nature00939

Shinohara A, Shinohara M, Ohta T, Matsuda S, Ogawa T (1998) Rad52 forms ring structures and co-operates with RPA in single-strand DNA annealing. Genes to cells : devoted to molecular & cellular mechanisms 3 (3):145-156

Sibanda BL, Critchlow SE, Begun J, Pei XY, Jackson SP, Blundell TL, Pellegrini L (2001) Crystal structure of an Xrcc4-DNA ligase IV complex. Nature structural biology 8 (12):1015-1019. doi:10.1038/nsb725

Siegel V, Walter P (1988) Each of the activities of signal recognition particle (SRP) is contained within a distinct domain: analysis of biochemical mutants of SRP. Cell 52 (1):39-49

Singh H, Figliola MJ, Dawson MJ, Olivares S, Zhang L, Yang G, Maiti S, Manuri P, Senyukov V, Jena B, Kebriaei P, Champlin RE, Huls H, Cooper LJ (2013) Manufacture of clinical-grade CD19-specific T cells stably expressing chimeric antigen receptor using Sleeping Beauty system and artificial antigen presenting cells. PloS one 8 (5):e64138. doi:10.1371/journal.pone.0064138

Singh H, Moyes JS, Huls MH, Cooper LJ (2015) Manufacture of T cells using the Sleeping Beauty system to enforce expression of a CD19-specific chimeric antigen receptor. Cancer Gene Ther 22 (2):95-100. doi:10.1038/cgt.2014.69

Singleton MR, Wentzell LM, Liu Y, West SC, Wigley DB (2002) Structure of the single-strand annealing domain of human RAD52 protein. Proceedings of the National Academy of Sciences of the United States of America 99 (21):13492-13497. doi:10.1073/pnas.212449899

Smit AF (1996) The origin of interspersed repeats in the human genome. Current opinion in genetics & development 6 (6):743-748

Smit AF, Riggs AD (1996) Tiggers and DNA transposon fossils in the human genome. Proceedings of the National Academy of Sciences of the United States of America 93 (4):1443-1448

Smith GC, Jackson SP (1999) The DNA-dependent protein kinase. Genes & development 13 (8):916-934

Smith J, Bibikova M, Whitby FG, Reddy AR, Chandrasegaran S, Carroll D (2000) Requirements for double-strand cleavage by chimeric restriction enzymes with zinc finger DNA-recognition domains. Nucleic acids research 28 (17):3361-3369

Soifer H, Higo C, Kazazian HH, Jr., Moran JV, Mitani K, Kasahara N (2001) Stable integration of transgenes delivered by a retrotransposon-adenovirus hybrid vector. Human gene therapy 12 (11):1417-1428. doi:10.1089/104303401750298571

Soifer HS, Kasahara N (2004) Retrotransposon-adenovirus hybrid vectors: efficient delivery and stable integration of transgenes via a two-stage mechanism. Curr Gene Ther 4 (4):373-384

Solinger JA, Kiianitsa K, Heyer WD (2002) Rad54, a Swi2/Snf2-like recombinational repair protein, disassembles Rad51:dsDNA filaments. Molecular cell 10 (5):1175-1188

Soriano P (1999) Generalized lacZ expression with the ROSA26 Cre reporter strain. Nature genetics 21 (1):70-71. doi:10.1038/5007

Soulas-Sprauel P, Le Guyader G, Rivera-Munoz P, Abramowski V, Olivier-Martin C, Goujet-Zalc C, Charneau P, de Villartay JP (2007) Role for DNA repair factor XRCC4 in immunoglobulin class switch recombination. The Journal of experimental medicine 204 (7):1717-1727. doi:10.1084/jem.20070255

Spagnolo L, Rivera-Calzada A, Pearl LH, Llorca O (2006) Three-dimensional structure of the human DNA-PKcs/Ku70/Ku80 complex assembled on DNA and its implications for DNA DSB repair. Molecular cell 22 (4):511-519. doi:10.1016/j.molcel.2006.04.013

Speek M (2001) Antisense promoter of human L1 retrotransposon drives transcription of adjacent cellular genes. Molecular and cellular biology 21 (6):1973-1985. doi:10.1128/MCB.21.6.1973-1985.2001

Startek M, Szafranski P, Gambin T, Campbell IM, Hixson P, Shaw CA, Stankiewicz P, Gambin A (2015) Genome-wide analyses of LINE-LINE-mediated nonallelic homologous recombination. Nucleic acids research 43 (4):2188-2198. doi:10.1093/nar/gku1394

Stefl R, Wu H, Ravindranathan S, Sklenar V, Feigon J (2004) DNA A-tract bending in three dimensions: solving the dA4T4 vs. dT4A4 conundrum. Proceedings of the National Academy of Sciences of the United States of America 101 (5):1177-1182. doi:10.1073/pnas.0308143100

Sugawara N, Wang X, Haber JE (2003) In vivo roles of Rad52, Rad54, and Rad55 proteins in Rad51-mediated recombination. Molecular cell 12 (1):209-219

Sung P, Klein H (2006) Mechanism of homologous recombination: mediators and helicases take on regulatory functions. Nature reviews Molecular cell biology 7 (10):739-750. doi:10. 1038/nrm2008

Suspene R, Guetard D, Henry M, Sommer P, Wain-Hobson S, Vartanian JP (2005) Extensive editing of both hepatitis B virus DNA strands by APOBEC3 cytidine deaminases in vitro and in vivo. Proceedings of the National Academy of Sciences of the United States of America 102 (23):8321-8326. doi:10.1073/pnas.0408223102

Suzuki J, Yamaguchi K, Kajikawa M, Ichiyanagi K, Adachi N, Koyama H, Takeda S, Okada N (2009) Genetic evidence that the non-homologous end-joining repair pathway is involved in LINE retrotransposition. PLoS genetics 5 (4):e1000461. doi:10.1371/journal.pgen.1000461

Symer DE, Connelly C, Szak ST, Caputo EM, Cost GJ, Parmigiani G, Boeke JD (2002) Human l1 retrotransposition is associated with genetic instability in vivo. Cell 110 (3):327-338

Szak ST, Pickeral OK, Landsman D, Boeke JD (2003) Identifying related L1 retrotransposons by analyzing 3' transduced sequences. Genome biology 4 (5):R30

Szak ST, Pickeral OK, Makalowski W, Boguski MS, Landsman D, Boeke JD (2002) Molecular archeology of L1 insertions in the human genome. Genome biology 3 (10):research0052

Takahashi H, Fujiwara H (2002) Transplantation of target site specificity by swapping the endonuclease domains of two LINEs. The EMBO journal 21 (3):408-417

Talis AL, Huibregtse JM, Howley PM (1998) The role of E6AP in the regulation of p53 protein levels in human papillomavirus (HPV)-positive and HPV-negative cells. The Journal of biological chemistry 273 (11):6439-6445

Tchenio T, Casella JF, Heidmann T (2000) Members of the SRY family regulate the human LINE retrotransposons. Nucleic acids research 28 (2):411-415

Teneng I, Stribinskis V, Ramos KS (2007) Context-specific regulation of LINE-1. Genes to cells : devoted to molecular & cellular mechanisms 12 (10):1101-1110. doi:10.1111/j.1365-2443.2007.01117.x

Thacker J (1999) A surfeit of RAD51-like genes? Trends in genetics : TIG 15 (5):166-168

Thayer RE, Singer MF, Fanning TG (1993) Undermethylation of specific LINE-1 sequences in human cells producing a LINE-1-encoded protein. Gene 133 (2):273-277

Tsujishita Y, Guo S, Stolz LE, York JD, Hurley JH (2001) Specificity determinants in phosphoinositide dephosphorylation: crystal structure of an archetypal inositol polyphosphate 5-phosphatase. Cell 105 (3):379-389

Tuck SP, Crawford L (1989) Overexpression of normal human p53 in established fibroblasts leads to their tumorigenic conversion. Oncogene research 4 (2):81-96

Turelli P, Mangeat B, Jost S, Vianin S, Trono D (2004) Inhibition of hepatitis B virus replication by APOBEC3G. Science 303 (5665):1829. doi:10.1126/science.1092066

Turner G, Barbulescu M, Su M, Jensen-Seaman MI, Kidd KK, Lenz J (2001) Insertional polymorphisms of full-length endogenous retroviruses in humans. Current biology : CB 11 (19):1531-1535

Ullu E, Tschudi C (1984) Alu sequences are processed 7SL RNA genes. Nature 312 (5990):171-172

Upton KR, Gerhardt DJ, Jesuadian JS, Richardson SR, Sanchez-Luque FJ, Bodea GO, Ewing AD, Salvador-Palomeque C, van der Knaap MS, Brennan PM, Vanderver A, Faulkner GJ (2015) Ubiquitous L1 mosaicism in hippocampal neurons. Cell 161 (2):228-239. doi:10.1016/j.cell.2015.03.026

Urnov FD, Miller JC, Lee YL, Beausejour CM, Rock JM, Augustus S, Jamieson AC, Porteus MH, Gregory PD, Holmes MC (2005) Highly efficient endogenous human gene correction using designed zinc-finger nucleases. Nature 435 (7042):646-651. doi:10.1038/nature03556

Van Arsdell SW, Weiner AM (1984) Pseudogenes for human U2 small nuclear RNA do not have a fixed site of 3' truncation. Nucleic acids research 12 (3):1463-1471

Van Komen S, Petukhova G, Sigurdsson S, Stratton S, Sung P (2000) Superhelicity-driven homologous DNA pairing by yeast recombination factors Rad51 and Rad54. Molecular cell 6 (3):563-572

Vanin EF (1985) Processed pseudogenes: characteristics and evolution. Annual review of genetics 19:253-272. doi:10.1146/annurev.ge.19.120185.001345

Verkaik NS, Esveldt-van Lange RE, van Heemst D, Bruggenwirth HT, Hoeijmakers JH, Zdzienicka MZ, van Gent DC (2002) Different types of V(D)J recombination and end-joining defects in DNA double-strand break repair mutant mammalian cells. European journal of immunology 32 (3):701-709. doi:10.1002/1521-4141(200203)32:3<701::AID-IMMU701>3.0.CO;2-T

Voigt K, Gogol-Doring A, Miskey C, Chen W, Cathomen T, Izsvak Z, Ivics Z (2012) Retargeting sleeping beauty transposon insertions by engineered zinc finger DNA-

binding domains. Molecular therapy : the journal of the American Society of Gene Therapy 20 (10):1852-1862. doi:10.1038/mt.2012.126

Volpers C, Kochanek S (2004) Adenoviral vectors for gene transfer and therapy. The journal of gene medicine 6 Suppl 1:S164-171. doi:10.1002/jgm.496

Walisko O, Schorn A, Rolfs F, Devaraj A, Miskey C, Izsvak Z, Ivics Z (2008) Transcriptional activities of the Sleeping Beauty transposon and shielding its genetic cargo with insulators. Molecular therapy : the journal of the American Society of Gene Therapy 16 (2):359-369. doi:10.1038/sj.mt.6300366

Walker JR, Corpina RA, Goldberg J (2001) Structure of the Ku heterodimer bound to DNA and its implications for double-strand break repair. Nature 412 (6847):607-614. doi:10.1038/35088000

Wang H, Rosidi B, Perrault R, Wang M, Zhang L, Windhofer F, Iliakis G (2005) DNA ligase III as a candidate component of backup pathways of nonhomologous end joining. Cancer research 65 (10):4020-4030. doi:10.1158/0008-5472.CAN-04-3055

Wang H, Zeng ZC, Bui TA, Sonoda E, Takata M, Takeda S, Iliakis G (2001) Efficient rejoining of radiation-induced DNA double-strand breaks in vertebrate cells deficient in genes of the RAD52 epistasis group. Oncogene 20 (18):2212-2224. doi:10.1038/sj.onc.1204350

Wang M, Wu W, Wu W, Rosidi B, Zhang L, Wang H, Iliakis G (2006) PARP-1 and Ku compete for repair of DNA double strand breaks by distinct NHEJ pathways. Nucleic acids research 34 (21):6170-6182. doi:10.1093/nar/gkl840

Waterhouse PM, Graham MW, Wang MB (1998) Virus resistance and gene silencing in plants can be induced by simultaneous expression of sense and antisense RNA. Proceedings of the National Academy of Sciences of the United States of America 95 (23):13959-13964

Wei W, Gilbert N, Ooi SL, Lawler JF, Ostertag EM, Kazazian HH, Boeke JD, Moran JV (2001) Human L1 retrotransposition: cis preference versus trans complementation. Molecular and cellular biology 21 (4):1429-1439. doi:10.1128/MCB.21.4.1429-1439.2001

Wei W, Morrish TA, Alisch RS, Moran JV (2000) A transient assay reveals that cultured human cells can accommodate multiple LINE-1 retrotransposition events. Anal Biochem 284 (2):435-438. doi:10.1006/abio.2000.4675

Weichenrieder O, Repanas K, Perrakis A (2004) Crystal structure of the targeting endonuclease of the human LINE-1 retrotransposon. Structure 12 (6):975-986. doi:10.1016/j.str.2004.04.011

Weichenrieder O, Wild K, Strub K, Cusack S (2000) Structure and assembly of the Alu domain of the mammalian signal recognition particle. Nature 408 (6809):167-173.

Weinberg RA, Penman S (1968) Small molecular weight monodisperse nuclear RNA. Journal of molecular biology 38 (3):289-304

Weitzel JM, Buhrmester H, Stratling WH (1997) Chicken MAR-binding protein ARBP is homologous to rat methyl-CpG-binding protein MeCP2. Molecular and cellular biology 17 (9):5656-5666

West SC (2003) Cross-links between Fanconi anaemia and BRCA2. DNA repair 2 (2):231-234

Weterings E, Chen DJ (2007) DNA-dependent protein kinase in nonhomologous end joining: a lock with multiple keys? The Journal of cell biology 179 (2):183-186.

Weterings E, Verkaik NS, Bruggenwirth HT, Hoeijmakers JH, van Gent DC (2003) The role of DNA dependent protein kinase in synapsis of DNA ends. Nucleic acids research 31 (24):7238-7246

Wianny F, Zernicka-Goetz M (2000) Specific interference with gene function by double-stranded RNA in early mouse development. Nature cell biology 2 (2):70-75.

Wiegand HL, Doehle BP, Bogerd HP, Cullen BR (2004) A second human antiretroviral factor, APOBEC3F, is suppressed by the HIV-1 and HIV-2 Vif proteins. The EMBO journal 23 (12):2451-2458. doi:10.1038/sj.emboj.7600246

Wienert B, Funnell AP, Norton LJ, Pearson RC, Wilkinson-White LE, Lester K, Vadolas J, Porteus MH, Matthews JM, Quinlan KG, Crossley M (2015) Editing the genome to introduce a beneficial naturally occurring mutation associated with increased fetal globin. Nat Commun 6:7085. doi:10.1038/ncomms8085

Wilber A, Frandsen JL, Geurts JL, Largaespada DA, Hackett PB, McIvor RS (2006) RNA as a source of transposase for Sleeping Beauty-mediated gene insertion and expression in somatic cells and tissues. Molecular therapy : the journal of the American Society of Gene Therapy 13 (3):625-630. doi:10.1016/j.ymthe.2005.10.014

Williams DA (2008) Sleeping beauty vector system moves toward human trials in the United States. Molecular therapy : the journal of the American Society of Gene Therapy 16 (9):1515-1516. doi:10.1038/mt.2008.169

Wold MS (1997) Replication protein A: a heterotrimeric, single-stranded DNA-binding protein required for eukaryotic DNA metabolism. Annual review of biochemistry 66:61-92. doi:10.1146/annurev.biochem.66.1.61

Wolner B, Peterson CL (2005) ATP-dependent and ATP-independent roles for the Rad54 chromatin remodeling enzyme during recombinational repair of a DNA double strand break. The Journal of biological chemistry 280 (11):10855-10860.

Wong MJ, Goldberger G, Isenman DE, Minta JO (1995) Processing of human factor I in COS-1 cells co-transfected with factor I and paired basic amino acid cleaving enzyme (PACE) cDNA. Molecular immunology 32 (5):379-387

Woodcock DM, Lawler CB, Linsenmeyer ME, Doherty JP, Warren WD (1997) Asymmetric methylation in the hypermethylated CpG promoter region of the human L1 retrotransposon. The Journal of biological chemistry 272 (12):7810-7816

Wu L, Bachrati CZ, Ou J, Xu C, Yin J, Chang M, Wang W, Li L, Brown GW, Hickson ID (2006) BLAP75/RMI1 promotes the BLM-dependent dissolution of homologous recombination intermediates. Proceedings of the National Academy of Sciences of the United States of America 103 (11):4068-4073. doi:10.1073/pnas.0508295103

Wu L, Hickson ID (2003) The Bloom's syndrome helicase suppresses crossing over during homologous recombination. Nature 426 (6968):870-874. doi:10.1038/nature02253

Wu W, Wang M, Wu W, Singh SK, Mussfeldt T, Iliakis G (2008) Repair of radiation induced DNA double strand breaks by backup NHEJ is enhanced in G2. DNA repair 7 (2):329-338. doi:10.1016/j.dnarep.2007.11.008

Xie Y, Rosser JM, Thompson TL, Boeke JD, An W (2011) Characterization of L1 retrotransposition with high-throughput dual-luciferase assays. Nucleic acids research 39 (3):e16. doi:10.1093/nar/gkq1076

Xiong Y, Eickbush TH (1990) Origin and evolution of retroelements based upon their reverse transcriptase sequences. The EMBO journal 9 (10):3353-3362

Yan CT, Boboila C, Souza EK, Franco S, Hickernell TR, Murphy M, Gumaste S, Geyer M, Zarrin AA, Manis JP, Rajewsky K, Alt FW (2007) IgH class switching and translocations use a robust non-classical end-joining pathway. Nature 449 (7161):478-482. doi:10.1038/nature06020

Yang D, Lu H, Erickson JW (2000) Evidence that processed small dsRNAs may mediate sequence-specific mRNA degradation during RNAi in Drosophila embryos. Current biology : CB 10 (19):1191-1200

Yang H, Jeffrey PD, Miller J, Kinnucan E, Sun Y, Thoma NH, Zheng N, Chen PL, Lee WH, Pavletich NP (2002) BRCA2 function in DNA binding and recombination from a BRCA2-DSS1-ssDNA structure. Science 297 (5588):1837-1848.

Yang H, Li Q, Fan J, Holloman WK, Pavletich NP (2005) The BRCA2 homologue Brh2 nucleates RAD51 filament formation at a dsDNA-ssDNA junction. Nature 433 (7026):653-657. doi:10.1038/nature03234

Yang J, Malik HS, Eickbush TH (1999) Identification of the endonuclease domain encoded by R2 and other site-specific, non-long terminal repeat retrotransposable elements. Proceedings of the National Academy of Sciences of the United States of America 96 (14):7847-7852

Yang N, Zhang L, Zhang Y, Kazazian HH, Jr. (2003) An important role for RUNX3 in human L1 transcription and retrotransposition. Nucleic acids research 31 (16):4929-4940

Yant SR, Meuse L, Chiu W, Ivics Z, Izsvak Z, Kay MA (2000) Somatic integration and long-term transgene expression in normal and haemophilic mice using a DNA transposon system. Nature genetics 25 (1):35-41. doi:10.1038/75568

Yates JL, Warren N, Sugden B (1985) Stable replication of plasmids derived from Epstein-Barr virus in various mammalian cells. Nature 313 (6005):812-815

Yu F, Thiesen J, Stratling WH (2000) Histone deacetylase-independent transcriptional repression by methyl-CpG-binding protein 2. Nucleic acids research 28 (10):2201-2206

Yu F, Zingler N, Schumann G, Stratling WH (2001) Methyl-CpG-binding protein 2 represses LINE-1 expression and retrotransposition but not Alu transcription. Nucleic acids research 29 (21):4493-4501

Yu JY, DeRuiter SL, Turner DL (2002) RNA interference by expression of short-interfering RNAs and hairpin RNAs in mammalian cells. Proceedings of the National Academy of Sciences of the United States of America 99 (9):6047-6052.

Yuan J, Wang J, Crain K, Fearns C, Kim KA, Hua KL, Gregory PD, Holmes MC, Torbett BE (2012) Zinc-finger nuclease editing of human cxcr4 promotes HIV-1 CD4(+) T cell resistance and enrichment. Molecular therapy : the journal of the American Society of Gene Therapy 20 (4):849-859. doi:10.1038/mt.2011.310

Zamore PD, Tuschl T, Sharp PA, Bartel DP (2000) RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. Cell 101 (1):25-33. doi:10.1016/S0092-8674(00)80620-0

Zayed H, Izsvak Z, Walisko O, Ivics Z (2004) Development of hyperactive sleeping beauty transposon vectors by mutational analysis. Molecular therapy : the journal of the American Society of Gene Therapy 9 (2):292-304. doi:10.1016/j.ymthe.2003.11.024

Zhang H, Yang B, Pomerantz RJ, Zhang C, Arunachalam SC, Gao L (2003) The cytidine deaminase CEM15 induces hypermutation in newly synthesized HIV-1 DNA. Nature 424 (6944):94-98. doi:10.1038/nature01707

Zheng YH, Irwin D, Kurosu T, Tokunaga K, Sata T, Peterlin BM (2004) Human APOBEC3F is another host factor that blocks human immunodeficiency virus type 1 replication. Journal of virology 78 (11):6073-6076. doi:10.1128/JVI.78.11.6073-6076.2004

Zingler N, Weichenrieder O, Schumann GG (2005a) APE-type non-LTR retrotransposons: determinants involved in target site recognition. Cytogenetic and genome research 110 (1-4):250-268. doi:10.1159/000084959

Zingler N, Willhoeft U, Brose HP, Schoder V, Jahns T, Hanschmann KM, Morrish TA, Lower J, Schumann GG (2005b) Analysis of 5' junctions of human LINE-1 and Alu retrotransposons suggests an alternative model for 5'-end attachment requiring microhomology-mediated end-joining. Genome research 15 (6):780-789. doi:10.1101/gr.3421505

LILIANA ELISABETH LAYER

Date of birth: December 8, 1977

Place of birth: Frankenthal (Pfalz), Germany

RESEARCH AND WORK EXPERIENCE

Since Sept. 2013	CEO of Nova Space Biotechnology GmbH, Zurich, Switzerland
Since August 2013	Research associate of Prof. Dr. Dr. Oliver Ullrich at the University of Zurich, Switzerland, in the department of Cellular Biology working on immunology and gravitational biology.
August 2012- July 2013	Research and development at PromoCell GmbH in Heidelberg, Germany, in the field of primary human cells and cell type specific media.
December 2009- July 2012	Technical support and product specialist at PromoCell GmbH in Heidelberg, in the field of primary human cells and cell analysis.
November 2007- November 2009	Research associate in the department of Prof. Dr. P. H. Seeburg, Molecular Neurobiology at the Max-Planck Institute for Medical Research in Heidelberg, in the group of Dr. R. Sprengel.
November 2003- October 2007	Ph.D. student in the department of Human Retroelements at the Paul-Ehrlich Institute in Langen, Germany, in the group of Prof. Dr. Gerald G. Schumann.
May 2003- August 2003	Research associate in the department of Molecular Neurobiology at the Max-Planck Institute for Medical Research in Heidelberg.

EDUCATION

October 1997- April 2003	Studies of Biology at the Ruprecht-Karls University in Heidelberg. Subjects: Molecular Biology, Cellular Biology, Biochemistry. Diploma thesis in the department of Molecular Neurobiology at the Max-Planck Institute for Medical Research in Heidelberg.
Summer 1997	"Abitur" degree at the Integrierte Gesamtschule Ernst Bloch in Oggersheim, Germany. Major subjects: Biology, English, French.
August 1994- August 1995	Attendance of 11th grade of Sunset Senior High School in Miami, Florida, USA.

PUBLICATIONS ASSOCIATED WITH THIS THESIS

Patent:

Liliana E. Layer & Gerald G. Schumann (2008) Controlled activation of non-LTR retrotransposons in mammals. EP 2205750 B1

Peer-reviewed Paper:

Repanas K, Zingler N, Layer LE, Schumann GG, Perrakis A, Weichenrieder O (2007) Determinants for DNA target structure selectivity of the human LINE-1 retrotransposon endonuclease. Nucleic acids research 35 (14):4914-4926. doi:10.1093/nar/gkm516

Poster Presentations:

Liliana E. Layer, Nora Zingler, Kostas Repanas, Anastassis Perrakis, Oliver Weichenrieder and Gerald G. Schumann (2005) Structure-guided Mutational Analyses of the Human LINE-1 Endonuclease. Frühjahrsakademie der Gesellschaft für Genetik, Wittenberg

Liliana E. Layer and Gerald G. Schumann (2007) Identification of Host-encoded Double-strand Break (DSB) Repair Factors Involved in Human LINE-1 Retrotransposition. FASEB Summer Research Conference, Tucson (AZ), USA

ERKLÄRUNG

Ich erkläre hiermit, daß ich mich bisher keiner Doktorprüfung im Mathematisch-Naturwissenschaftlichen Bereich unterzogen habe.

Frankfurt am Main, den

VERSICHERUNG

Ich erkläre hiermit, daß ich die vorgelegte Dissertation über

Regulation of Human LINE-1 Activity and Integration Specificity by L1-EN Mutations, DSB Repair Proteins and the Application of the Tet/Dox-inducible Expression System

selbständig angefertigt und mich anderer Hilfsmittel als der in ihr angegebenen nicht bedient habe, insbesondere, daß alle Entlehnungen aus anderen Schriften mit Angabe der betreffenden Schrift gekennzeichnet sind.
Ich versichere, die Grundsätze der guten wissenschaftlichen Praxis beachtet, und nicht die Hilfe einer kommerziellen Promotionsvermittlung in Anspruch genommen zu haben.

Frankfurt am Main, den