

Untersuchungen zu positionsspezifisch integrierenden Retrotransposons in sozialen Amöben

Dissertation

zur Erlangung des akademischen Grades
„doctor rerum naturalium“ (Dr. rer. nat.)



vorgelegt dem Rat der Biologisch-Pharmazeutischen Fakultät
der Friedrich-Schiller-Universität Jena

von Diplom-Biologe Thomas Spaller
geboren am 5.4.1984 in Zwickau

1. Gutachter: Prof. Dr. Thomas Winckler (Friedrich-Schiller-Universität Jena)

2. Gutachter: PD Dr. Matthias Platzer (Fritz-Lipmann-Institut Jena)

3. Gutachter: Prof. Dr. Gerald G. Schumann (Paul-Ehrlich-Institut Langen)

Tag der Disputation: 26.05.2015

Inhalt

Inhaltsverzeichnis.....	i
Abbildungsverzeichnis.....	v
Tabellenverzeichnis.....	vii
Abkürzungsverzeichnis.....	viii
1 Einleitung.....	1
1.1 Die sozialen Amöben.....	1
1.1.2 Phylogenie innerhalb der Dictyostelia.....	3
1.1.3 Die Genome der sozialen Amöben.....	5
1.2 Mobile genetische Elemente.....	7
1.3 Retrotransposons.....	10
1.3.1 LTR-Retrotransposons.....	11
1.3.1.1 Die Retrotransposition der LTR-Retrotransposons.....	12
1.3.2 Non-LTR-Retrotransposons.....	13
1.3.2.1 Der Retrotranspositionsmechanismus der Non-LTR-Retrotransposons.....	15
1.4 DNA-Transposons.....	16
1.5 Einfluss mobiler genetischer Elemente auf Wirtsgenome.....	17
1.5.1 Wechselwirkungen zwischen mobilen genetischen Elementen und Wirtszellen.....	18
1.6 Spezifisch integrierende mobile genetische Elemente.....	20
1.7 Mobile genetische Elemente im Genom von <i>Dictyostelium discoideum</i>	21
1.7.1 Non-LTR-Retrotransposons.....	22
1.7.2 LTR-Retrotransposons.....	24
1.7.3 DNA-Transposons.....	26
1.8 RNA-Polymerase-III-transkribierte Gene.....	26
1.8.1 tRNA-Gene im Genom von <i>D. discoideum</i>	27
1.8.2 Modelle zur Erklärung der spezifischen Integration von Retrotransposons an tRNA-Genen.....	27
1.9 Genetisch markierte artifizielle TRE5-Retrotransposons.....	28
1.10 Zielsetzung der Arbeit.....	30
2 Material und Methoden.....	31
2.1 Materialien.....	31
2.1.1 Laborgeräte.....	31
2.1.2 Verbrauchsmaterialien.....	32
2.1.3 Verwendete Kits.....	32
2.1.4 Enzyme.....	33
2.1.5 DNA-Größenstandards.....	33
2.1.6 Antibiotika.....	33

2.1.7 Chemikalien.....	34
2.1.8 Medien.....	35
2.1.9 Puffer und Lösungen.....	35
2.1.10 Mikroorganismen.....	36
2.2 Verwendete Programme	37
2.3 Datenbanken.....	37
2.4 Methoden.....	38
2.4.1 Molekularbiologische Methoden.....	38
2.4.1.2 Isolation genomischer DNA.....	38
2.4.1.3 Extraktion verdaubarer genomischer DNA.....	38
2.4.1.4 Southern Blot mit biotinylierten Sonden.....	39
2.4.1.5 Inverse PCR	40
2.4.1.6 Extraktion von RNA und cDNA-Synthese.....	40
2.4.1.7 cRT-PCR	41
2.4.2 Zellkultur	41
2.4.2.1 Kultivierung von <i>Dictyostelium discoideum</i>	41
2.4.2.2 Kultivierung von Wildtyp-Stämmen	41
2.4.2.3 Herstellung von Zellpellets	42
2.4.2.4 Herstellung von Gefrierkulturen	42
2.4.2.5 CaCl ₂ -Transformation von <i>D. discoideum</i> -Zellen.....	42
2.4.2.6 Elektrotransformation von <i>D. discoideum</i> -Zellen.....	43
2.4.2.7 Retrotranspositionstest	43
2.5 In-silico-Analysen	44
2.5.1 Detektion neuer Retrotransposon-Spezies	44
2.5.2 Erstellung phylogenetischer Stammbäume.....	45
3 Ergebnisse.....	46
3.1 Identifizierung von Retrotransposons aus den Genomen der sozialen Amöben	47
3.1.1 Die Retrotransposons im Genom von <i>Dictyostelium purpureum</i>	47
3.1.1.1 tRNA-Gen-assoziierte LTR-Retrotransposons.....	47
3.1.1.2 Non-LTR-Retrotransposons aus <i>D. purpureum</i>	50
3.1.1.3 Zusammenfassende Bewertung.....	52
3.1.2 Die Retrotransposons im Genom von <i>Polysphondylium pallidum</i>	54
3.1.2.1 tRNA-Gen-spezifische LTR-Retrotransposons.....	54
3.1.2.2 Weitere LTR-Retrotransposons im Genom von <i>P. pallidum</i>	55
3.1.2.3 Die Non-LTR-Elemente von <i>P. pallidum</i>	56
3.1.2.4 Ein weiteres Non-LTR-Retrotransposon im Genom von <i>P. pallidum</i>	57
3.1.3 Die Retrotransposons im Genom von <i>Dictyostelium fasciculatum</i>	59
3.1.3.1 Die LTR-Retrotransposons von <i>D. fasciculatum</i>	59
3.1.3.2 Non-LTR-Retrotransposons	60
3.1.4 Phylogenie der Retrotransposons aus sozialen Amöben.....	63
3.1.4.1 Phylogenie der LTR-Retrotransposons	63
3.1.4.2 Phylogenie der Non-LTR-Retrotransposons	64
3.2 Experimenteller Teil	66
3.2.1 Genetisch markierte LTR-Retrotransposons.....	66

3.2.1.1 Das Skipper_ <i>msbri</i> -Element	67
3.2.1.2 Das DGLT-A_ <i>msbri</i> -Element.....	68
3.2.1.3 Austausch der RNase-H-Domäne in genetisch markierten Skipper-Elementen	68
3.2.1.4 Ein <i>msbri</i> -markiertes LTR-Element aus dem Genom von <i>D. fasciculatum</i>	69
3.2.2 Genomweite Sequenzierung von De-novo-Integrationen eines Non-LTR-Retrotransposons	72
3.2.2.1 Das Codon-adaptierte nicht-autonome TRE5-A_ <i>msbri</i> -Element	72
3.2.2.2 Überlegungen zur genomweiten Kartierung von De-novo-Integrationsereignissen	73
3.2.2.3 Bestimmung der Kopienzahl gesprungener Elemente	74
3.2.2.4 Überlegungen zu LAM-PCR-basierten Methoden.....	75
3.2.2.5 Etablierung der LAM-PCR zur Anreicherung von Integrationsereignissen	76
3.2.2.6 Charakterisierung der Vollständigkeit mobilisierter Retrotransposons	77
3.2.2.7 Sequenzierung tRNA-Gen-spezifischer PCR-Produkte zur Identifizierung neuer Integrationsereignisse	78
3.2.2.8 TRE5-A.1 ^{bsr} integriert oberhalb von tRNA-Genen	78
3.2.2.9 „Next-generation“-Sequenzierung tRNA-Gen-spezifischer PCR-Produkte	80
3.2.2.10 Sequenzierung von De-novo-Integrationen mit Hilfe der invers PCR-Methode	84
3.2.2.11 Mögliche Strategien zur Erzeugung von doppelsträngigen LAM-PCR-Produkten	86
3.2.2.12 Ligation eines ssDNA-Linkers an ein definiertes PCR-Produkt.....	86
3.2.2.13 Erzeugung und Sequenzierung einer Linker-basierten Bank von LAM-PCR-Produkten	88
3.2.3 TRE5-A-Integrationen in das extrachromosomale rDNA-Palindrom	89
3.2.3.1 Nachweis von Integrationen des artifiziellen TRE-5-A-Retrotransposons im Palindrom.....	91
3.2.3.2 cRT-PCR zur Charakterisierung von Palindrom-Transkripten.....	92
3.2.3.3 Das r5S-Gen als Integrationsort artifizieller Retrotransposons.....	94
3.2.3.4 Integrationen endogener Retrotransposons in Palindrom-Loci.....	95
3.2.4 Erzeugung eines genetisch markierten autonomen TRE5-A.1-Retrotransposons	96
3.2.4.1 Das autonome TRE5-A.1_ <i>msbri</i> -Retrotransposon.....	96
3.2.4.2 Untersuchungen zum <i>Cis</i> -Effekt der autonomen TRE5-A.1 ^{bsr} -Elemente	98
4 Diskussion.....	99
4.1 Retrotransposons in den Genomen der sozialen Amöben	99
4.2 Variabilität der Integrationsabstände zu tRNA-Genen	102
4.3 Ein neu entdecktes DGLT-A-Element aus <i>D. fasciculatum</i> : Verbindung zwischen Skipper- und DGLT-A-Familie	103
4.4 Sequenzierung Linker-ligierter LAM-PCR-Produkte	104
4.5 Genomweite Analyse Pol-III-assoziiierter Integrationsereignisse	106
4.6 Das artifizielle Retrotransposon TRE5-A.2 ^{bsf} als Werkzeug zur Identifizierung unbekannter RNA-Polymerase-III-Transkriptionskomplexe.....	108
4.7 Integrationen von TRE5-Retrotransposons in DDT-S-Elemente	108
4.8 Hinweise auf die Replikation des extrachromosomalen Palindroms	109
4.9 RNA-Polymerase-III-Transkriptionskomplexe: Einfluss auf globale Chromatinstrukturen.....	110
4.10 Genetisch markierte, autonome Retrotransposons	111
4.11 Ausblick.....	113

5 Zusammenfassung.....	115
6 Summary.....	117
Literaturverzeichnis.....	118
Anhang	127
Liste der wissenschaftlichen Veröffentlichungen.....	137
Lebenslauf.....	138
Ehrenwörtliche Erklärung	139
Danksagung	140

Abbildungsverzeichnis

Abb. 1: Phylogenetische Stellung der sozialen Amöben	1
Abb. 2: Lebenszyklus der sozialen Amöbe <i>Dictyostelium discoideum</i>	3
Abb. 3: Phylogenie der sozialen Amöben.....	4
Abb. 4: Anteile und Verteilung mobiler genetischer Elemente an Genomen verschiedener Spezies	8
Abb. 5: Charakteristische Lebenszyklen der drei Klassen mobiler genetischer Elemente	9
Abb. 6: Struktureller Aufbau, Klassifizierung und Phylogenie der Retrotransposons	10
Abb. 7: Modell der reversen Transkription der LTR-Retrotransposons	13
Abb. 8: Struktureller Aufbau der LINE- und SINE-Elemente.....	14
Abb. 9: Mechanismus der Retrotransposition bei Non-LTR-Retrotransposons	15
Abb. 10: Allgemeiner Aufbau eines klassischen DNA-Transposons	17
Abb. 11: Genomdynamik und TE-Aktivität	19
Abb. 12: Mobile genetische Elemente im Genom von <i>D. discoideum</i>	21
Abb. 13: Aufbau der verschiedenen Non-LTR-Retrotransposons aus <i>D. discoideum</i>	24
Abb. 14: Struktur der LTR-Retrotransposons von <i>D. discoideum</i>	26
Abb. 15: Modell der spezifischen Integration von TRE5-Elementen an tRNA-Gene.....	28
Abb. 16: Schema des Retrotranspositions-Assays.....	29
Abb. 17: Evolution und Phylogenie der sozialen Amöben	46
Abb. 18: Die Struktur der identifizierten LTR-Retrotransposons von <i>D. purpureum</i>	49
Abb. 19: Die Non-LTR-Retrotransposons aus <i>D. purpureum</i>	52
Abb. 20: Die LTR-Retrotransposons von <i>P. pallidum</i>	55
Abb. 21: Die Non-LTR-Retrotransposons der sozialen Amöbe <i>P. pallidum</i>	57
Abb. 22: Die LTR-Retrotransposons aus dem Genom von <i>D. fasciculatum</i>	60
Abb. 23: Die Non-LTR-Retrotransposons von <i>D. fasciculatum</i>	61
Abb. 24: Phylogenie der neu identifizierten LTR-Retrotransposons sozialer Amöben	63
Abb. 25: Stammbaum der Non-LTR-Retrotransposons sozialer Amöben.....	65
Abb. 26: Das genetisch markierte Skipper-Retrotransposon	67
Abb. 27: Das DGLT-A-Element aus <i>D. fasciculatum</i>	70
Abb. 28: Das Codon-adaptierte, nicht autonome TRE5-A ^{bsr} -Element	73
Abb. 29: Southern-Blot zur Bestimmung der Kopienzahl von Integrationsereignissen.....	75
Abb. 30: Effizienz der Kopplung von LAM-PCR-Produkten an Streptavidin-Kügelchen.....	77

Abb. 31: Charakterisierung der Länge von integrierten TRE5-A ^{bsr} -Retrotransposon-Kopien	78
Abb. 32: tRNA-Gen-spezifische PCR-Produkte nach der LAM-PCR	79
Abb. 33: Abstände der integrierten TRE5-A ^{bsr} -Retrotransposonkopien zu tRNA-Genen.....	81
Abb. 34: Unterscheidbarkeit verschiedener Integrationsereignisse am Locus Glu-UUC-2	81
Abb. 35: Invers-PCR-Strategie zur Amplifikation von TRE5-A ^{bsr} Integrationen.....	85
Abb. 36: Test der Ligation eines ssDNA-Linkers an ein LAM-PCR-Produkt	87
Abb. 37: A- und B-Boxen auf dem extrachromosomalen Palindrom.....	90
Abb. 38: PCR zur Identifizierung von Integrationen im Palindrom	91
Abb. 39: Beispiele für Integrationen artifizierender Retrotransposons an A- und B-Boxen im Palindrom.....	92
Abb. 40: Nachweis und Charakterisierung der Transkripte von Palindrom-Loci.....	93
Abb. 41: Integration eines TRE5-A ^{bsr} -Retrotransposons am r5S-Gen.....	94
Abb. 42: Integrationen endogener TRE5-A.1-Retrotransposons in Palindrom-Loci	95
Abb. 43: Das autonome Retrotransposons TRE5-A.1_ <i>msbri</i>	97
Abb. 44: Vergleich der Retrotranspositionsfrequenzen von autonomen und nicht-autonomen TRE5-A ^{bsr} -Retrotransposons	98
Abb. 45: Integrationsabstände der neu gefundenen Retrotransposons zu tRNA-Genen	102
Abb. 46: Vergleich der Sequenzier-Ergebnisse von tRNA-Gen-spezifischer und inverser PCR	107
Abb. 47: Extra-transkriptionelle Einflüsse von TFIIC	111

Tabellenverzeichnis

Tab. 1: Charakteristische Merkmale der Genome unterschiedlicher Arten sozialer Amöben.	6
Tab. 2: Beispiele für spezifisch integrierende mobile genetische Elemente.	20
Tab. 3: Die Retrotransposons im Genom von <i>Dictyostelium purpureum</i>	53
Tab. 4: Charakterisitika der Retrotransposons von <i>Polysphondylium pallidum</i>	58
Tab. 5: Retrotransposons im Genom von <i>Dictyostelium fasciculatum</i>	62
Tab. 6: Ergebnisse der Sequenzierung von tRNA-Gen-spezifischen LAM-PCR-Produkten.	82
Tab. 7: Verwendete Primer für <i>Nested</i> -PCRs zum Nachweis von Integrationen endogener TRE5-Retrotransposons im Palindrom.	95
Tab. 8: Genomcharakteristika, Anzahl der tRNA-Gene und Anteil von Retrotransposons verschiedener sozialer Amöben.	99
Tab. 9: Detektierbare Bestandteile der RNA-Interferenz-Maschinerie der untersuchten Spezies.	101

Abkürzungsverzeichnis

Abb.	Abbildung
AS	Aminosäure
Bp	Basenpaar
Brf	<i>B-related-factor</i>
BSA	bovines Serumalbumin
BSR	Blasticidin
cAMP	<i>cyclic</i> AMP (zyklisches Adenosin-Monophosphat)
cDNA	<i>complementary</i> -DNA
CDS	<i>coding DNA sequence</i> (codierende DNA-Sequenz)
ddCTP	2',3'-Didesoxycytidin-5'-Triphosphat
Dd	<i>Dictyostelium discoideum</i>
Df	<i>Dictyostelium fasciculatum</i>
DIRS	<i>Dictyostelium intermediate repeat sequence</i>
DNA	Desoxyribonukleinsäure
Dp	<i>Dictyostelium purpureum</i>
G418	Geneticin
HRP	<i>horseradish peroxidase</i> (Meerrettichperoxidase)
IN	Integrase
kb	Kilobasenpaar (1 kb = 1000 bp)
kDa	Kilo-Dalton (1 Da = $\frac{1}{12}$ der Masse eines Atoms des Kohlenstoff-Isotops ^{12}C)
LTR	<i>long terminal repeat</i> (terminale Sequenzwiederholung)
Mb	Megabasenpaar (1 Mb = 1000 kb)
min	Minute
mRNA	<i>messenger</i> -RNA (Boten-RNA)
n.d.	nicht definiert
ORF	<i>open reading frame</i> (Offener Leserahmen)
PCR	<i>polymerase chain reaction</i> (Polymerase-Kettenreaktion)
Pol	Polymerase
Pp	<i>Polysphondylium pallidum</i>
rDNA	ribosomale Desoxyribonukleinsäure

RNA	Ribonukleinsäure
RNase-H	Ribonuklease H
rpm	rounds per minute (Umdrehungen pro Minute)
Rrp	<i>RNA-dependent RNA-polymerase</i> (RNA-abhängige RNA-Polymerase)
RT	Reverse Transkriptase
sek	Sekunden
Tab.	Tabelle
TBP	TATA-bindendes Protein
TFIIIA/-B/-C	Transkriptionsfaktor-III-A/-B/-C
TPRT	target primed reverse transcription
TRE	tRNA <i>gene-targeted retroelement</i>
tRNA	Transfer-RNA
TSD	target site duplication (Zielsequenzverdopplung)
Ty	<i>transposon of yeast</i>
vgl.	vergleiche
z. B.	zum Beispiel

1 Einleitung

1.1 Die sozialen Amöben

Die sozialen Amöben (Dictyostelia), oder zellulären Schleimpilze, gehören gemeinsam mit den Myxogastria und den Protostelia zur Gruppe der Mycetozoa, welche durch ihren Lebenszyklus charakterisiert sind (Fiore-Donno *et al.* 2010). Dieser Lebenszyklus teilt sich in zwei Phasen: Zum einen findet man die Mycetozoa als amöboide Zellen, zum anderen haben sie die Möglichkeit zur Differenzierung zu Sporentragenden Fruchtkörpern. Mycetozoa bilden eine eigene Gruppe innerhalb der Amoebozoa, einer Supergruppe, zu denen Pseudopodien-bildende amöboide Eukaryoten gehören (Cavalier-Smith 1998; Adl *et al.* 2012) und die eine Schwestergruppe der Pilze und Tiere darstellt. Die Erstbeschreibung von *Dictyostelium discoideum* (Raper 1935) erfolgte einige Jahrzehnte nach der Entdeckung der sozialen Amöben. Als erste Spezies wurde 1869 *Dictyostelium mucoroides* isoliert (Brefeld 1869). Lange Zeit war die genaue phylogenetische Zugehörigkeit der Dictyostelia unklar, da sie aufgrund morphologischer Charakteristika sowohl Merkmale von Pilzen als auch von Protozoen aufweisen (Raper und Rahn 1984). Neuere phylogenetische Untersuchungen zeigten jedoch eindeutig, dass *D. discoideum* der Supergruppe der Amoebozoa angehört, die kurz vor der Supergruppe der Opisthokonta (Tiere und Pflanzen) abzweigt ist, etwa zur selben Zeit oder kurz nach den höheren Pflanzen (Abbildung 1) (Glöckner und Noegel 2013; Schilde und Schaap 2013).

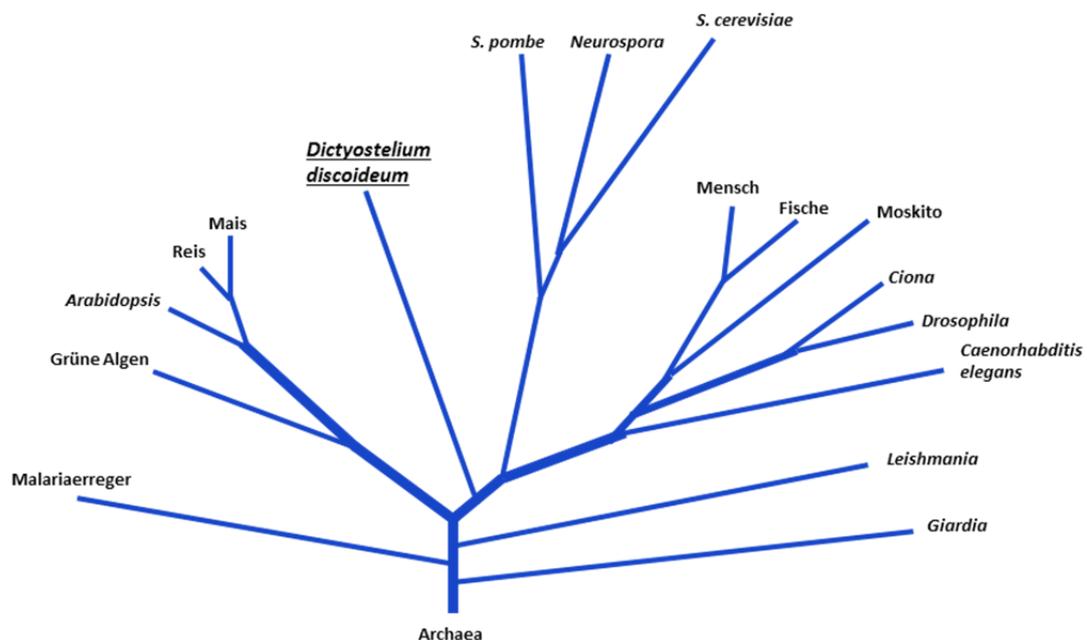


Abb. 1: Phylogenetische Stellung der sozialen Amöben.

Die sozialen Amöben sind phylogenetisch zwischen den Pflanzen und den Pilzen positioniert. Der Stammbaum beruht auf einem Vergleich von 5279 orthologen Proteinclustern der 17 Eukaryoten, welcher auf Daten von 159 Proteinclustern aus 6 Archaeen-Proteomen wurzelt (Eichinger *et al.* 2005).

Der Lebenszyklus der sozialen Amöben, am Beispiel von *Dictyostelium discoideum*, ist gekennzeichnet durch eine Unterteilung in eine vegetative, einzellige Phase und die Ausbildung multizellulärer Fruchtkörper bei Nahrungsmangel (Abbildung 2). Die Bezeichnung „soziale Amöben“ stammt in diesem Zusammenhang von dem Umstand, dass die Zellen einer Population während der Aggregation und Ausbildung der Fruchtkörper kooperieren, um einen vielzelligen Fruchtkörper zu bilden, welcher für die Gesamtheit der Population von großem Nutzen für die Verteilung von Sporen ist (Kessin *et al.* 1996; Kuzdzal-Fick *et al.* 2007; Li und Purugganan 2011).

Als Nahrung dienen den im Waldboden lebenden amöboiden Zellen Bakterien- und Hefezellen, welche durch Phagozytose aufgenommen werden. Die Fortbewegung der 9-13 µm großen Zellen geschieht durch die Ausbildung von Pseudopodien, welche durch das Cytoskelett gebildet werden (Bonner *et al.* 1969; Ashworth und Watts 1970).

Solange ein ausreichendes Nahrungsangebot vorhanden ist, durchlaufen die vegetativen Zellen mitotische Zellteilungen, die ca. drei bis vier Stunden dauern (Bonner 1959). Sobald ein Nahrungsmangel auftritt, beginnen die Zellen zu polarisieren und pulsierend extrazelluläre Lockstoffe zu sezernieren, welche von den umgebenden Zellen chemotaktisch wahrgenommen werden. Diese Zellen beginnen ebenfalls den Lockstoff pulsierend auszuschütten, wodurch sich das Signal weiter verstärkt (Goldbeter 2006). Die Art des sezernierten Lockstoffs variiert innerhalb der Dictyostelia (Asghar *et al.* 2012). *Dictyostelium discoideum* verwendet, wie auch weitere Arten der sozialen Amöben, zyklisches Adenosinmonophosphat (cAMP) (Burdine und Clarke 1995). Die Sezernierung und Perzeption des cAMPs durch *D. discoideum* bewirkt ein Wandern (engl.: „streaming“) der Zellen zum Aggregationszentrum und zur Bildung eines Pseudoplasmodiums im Aggregationszentrum, welches aus bis zu 100.000 Einzelzellen bestehen kann. Die Einzelzellen behalten dabei ihre Integrität, weshalb es sich nicht um ein echtes Plasmodium handelt. Das Pseudoplasmodium kann nun in ein motiles „slug“-Stadium übergehen. Geschützt durch eine extrazelluläre Matrix nimmt der „slug“ Umweltreize wie Temperatur, Licht, Feuchtigkeit, sowie chemische Gradienten wahr und reagiert darauf (Bonner 1952; Freeze und Loomis 1977). Der „slug“ dient dazu, als Zwischenschritt des Lebenszyklus der sozialen Amöben, eine geeignete Umgebung für die Entwicklung der Fruchtkörper zu finden. Bereits in dieser Phase der Entwicklung findet eine Differenzierung in unterschiedliche Zelltypen statt, wobei sich circa 20 % der Zellen zu Prästielzellen und rund 80 % zu Präsporenzellen entwickeln (Bonner 1952).

Die Kulmination, die Bildung des Fruchtkörpers (Sorokarps), beginnt in Abhängigkeit von verschiedenen Umweltfaktoren (Kirsten *et al.* 2005) mit der Ausbildung des sogenannten „mexican hat“ (Raper und Fennell 1952). Diese Phase ist charakterisiert durch die Abflachung des Pseudoplasmodiums und der Migration der Prästielzellen zum Zentrum. Anschließend beginnen die Prästielzellen mit der Bildung des Stiels (Dormann *et al.* 1996). Durch die Migration der Prästielzellen zum sich bildenden Stiel und die Auflagerung von Zellulose-Wänden wächst dieser und wird gleichzeitig stabilisiert. Die Prästielzellen vakuolisieren und sterben während dieses Prozesses (Raper

und Fennell 1952; Uchikawa *et al.* 2011). Schon während der Bildung des Stiels beginnen die Präsporenzellen an der Spitze mit der Differenzierung zu Sporen, welche sich durch die Bildung einer extrazellulären Schutzschicht aus Cellulose und Proteinen auszeichnet (Dormann *et al.* 1996; West 2003). Nach der Bildung des Stiels und der Sporen ist die Entwicklung des Fruchtkörpers abgeschlossen. Die Sporen können anschließend, zum Beispiel durch Tiere oder Wasser, verbreitet werden und sich unter günstigen Umweltbedingungen erneut zu solitären Amöben entwickeln (Cotter *et al.* 1992; Li und Purugganan 2011). Des Weiteren ist bekannt, dass auch sexuelle Zyklen durch die Bildung von Makrozysten möglich sind (Abbildung 2) (Kessin 2001).

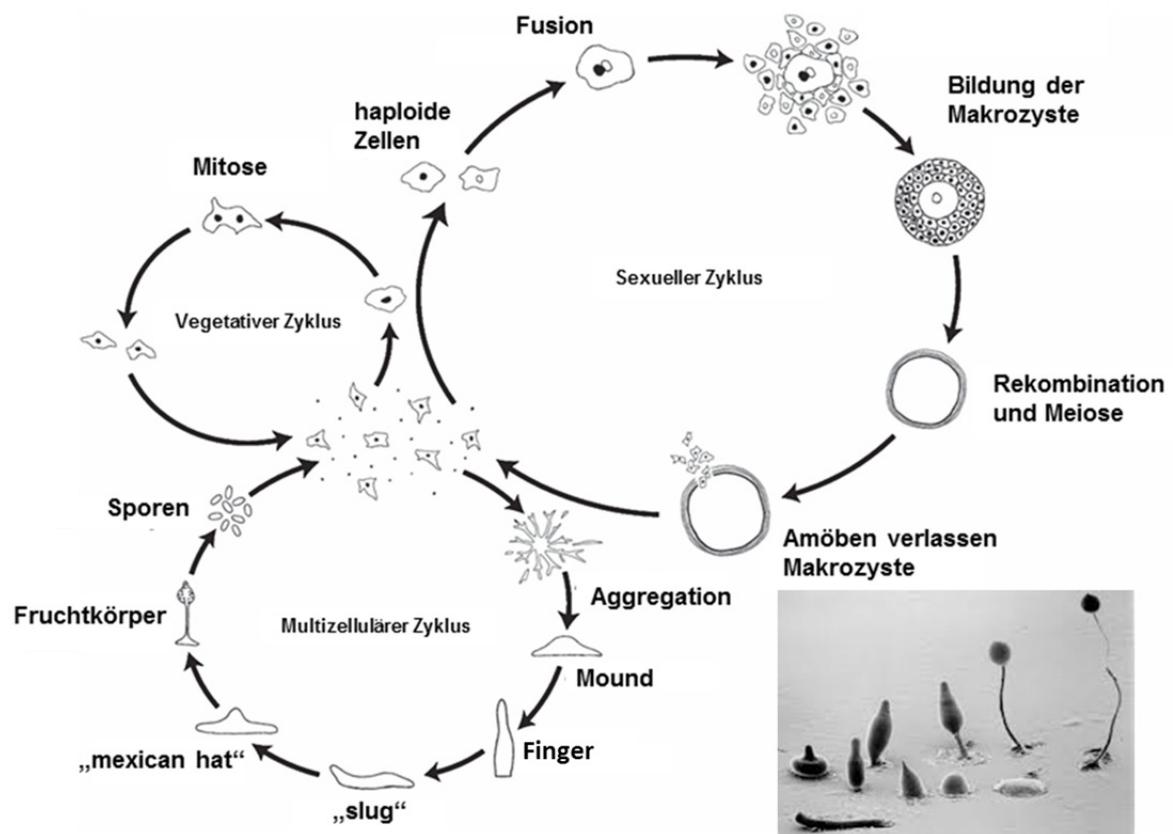


Abb.2: Lebenszyklus der sozialen Amöbe *Dictyostelium discoideum*.

Modifiziert nach Romeralo *et al.* 2012. (Mikroskopische Aufnahme: Mark Grimson & Lary Blanton, Electron Microscopy Laboratory, Department of Biological Sciences, Texas Tech University).

1.1.2 Phylogenie innerhalb der Dictyostelia

Zunächst wurden die Spezies der Dictyostelia anhand morphologischer Merkmale klassifiziert, Grundlage hierfür waren unterschiedliche Verzweigungstypen von Fruchtkörpern sowie die Morphologie der Sporen (Raper und Rahn 1984) (Abbildung 3B). Durch die Sequenzierung der 18S-rDNA von über 100 Arten der Dictyostelia konnte ein Stammbaum errechnet werden, der zeigt,

dass eine Einteilung der sozialen Amöben in 4 Gruppen möglich ist (Schaap *et al.* 2006) (Abbildung 3A). Der bekannteste Vertreter, *Dictyostelium discoideum* (*Dd*), gruppiert sich in der phylogenetisch jüngsten Gruppe 4 ein. Die weiteren Arten mit bisher komplett sequenzierten Genomen sind: *Dictyostelium purpureum* (Gruppe 4; *Dp*), *Polysphondylium pallidum* (Gruppe 2; *Pp*) und *Dictyostelium fasciculatum* (Gruppe 1; *Df*) (Eichinger *et al.* 2005; Heidel *et al.* 2011; Sugang *et al.* 2011).

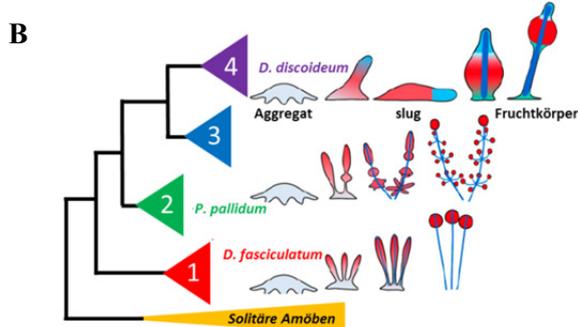
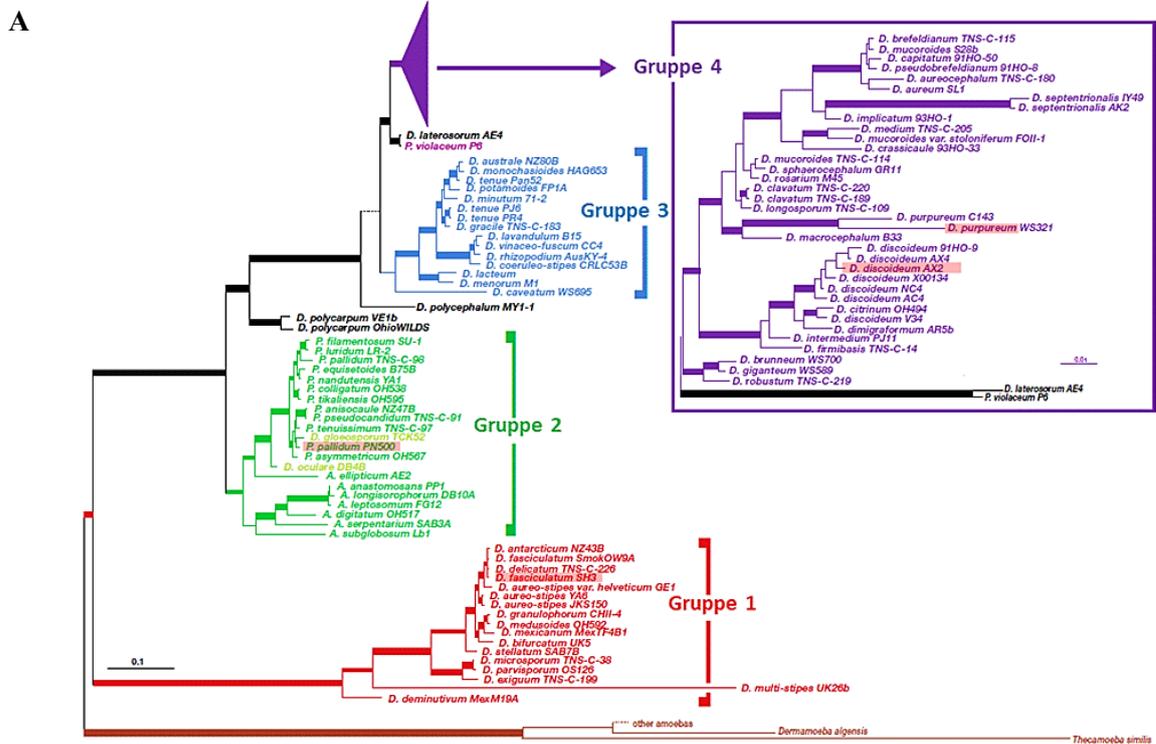


Abb. 3: Phylogenie der sozialen Amöben.
(A) Stammbaum der sozialen Amöben, basierend auf rDNA-Sequenzen (Schaap *et al.* 2006). Die sozialen Amöben können danach in 4 Gruppen eingeteilt werden. Spezies mit komplett sequenzierten Genomen sind rot markiert.
(B) Beispiele unterschiedlicher Fruchtkörpermorphologien in den jeweiligen Gruppen der sozialen Amöben. Modifiziert nach Heidel *et al.* 2011.

Durch phylogenetische Untersuchungen konnte des Weiteren gezeigt werden, dass der gemeinsame letzte Vorfahre dieser Arten vor mindestens 600 bis 1000 Mio. Jahren existierte. Damit sind die Dictyostelia evolutionär älter als die Gefäßpflanzen (Tracheophyta) und die vielzelligen Tiere (Metazoa) (Rubinstein *et al.* 2010; Heidel *et al.* 2011).

1.1.3 Die Genome der sozialen Amöben

Durch die Sequenzierung kompletter Genome verschiedener Arten der Dictyostelia (vgl. 1.1.2.) war es möglich, Rückschlüsse auf die individuelle Evolution der einzelnen Genome zu ziehen (Tabelle 1). Betrachtet man die Genome global, dann fällt auf, dass trotz der großen evolutionären Abstände zwischen den vier Arten nur marginale Größenunterschiede der Genome vorliegen. Sie variieren unbedeutend zwischen 33 und 35 Mb (Eichinger *et al.* 2005; Heidel *et al.* 2011; Sugang *et al.* 2011). Weitaus größer sind die Unterschiede des A/T-Gehalts der Genome, welcher zwischen 66,2 % bei *D. fasciculatum* und bis zu 77,6 % bei *D. discoideum* schwankt.

Ein weiteres Merkmal von *D. discoideum*, *D. fasciculatum* und *P. pallidum* ist das Vorhandensein von extrachromosomalen Palindromen, welche aus zwei gespiegelten Armen bestehen. Auf diesen liegen Kopien der Gene für die ribosomalen RNAs (Eichinger *et al.* 2005; Heidel *et al.* 2011). Die Palindrome kommen im Zellkern in großer Kopienzahl vor, der Mechanismus ihrer Replikation und Verteilung während der Zellteilung ist allerdings ungeklärt (Sugang *et al.* 2003). Eine Besonderheit von *D. discoideum* zeigte sich bei der Untersuchung der Telomerbereiche. Diese bestehen aus Bruchstücken von palindromischen rDNA-Einheiten (Eichinger *et al.* 2005), was im starken Kontrast zu den Telomeren von *D. fasciculatum* und *P. pallidum* steht. Diese bestehen aus normalen eukaryotischen Telomermotiven, die sie für die Stabilisierung ihrer Chromosomenenden benötigen. Unterschiede gibt es auch bei der Struktur der Centromerregionen. Während die Centromere von *Dd* und *Df* aus Clustern von DIRS-Elementen bestehen, einem Retrotransposon, konnten solche Strukturen bei *Pp* nicht nachgewiesen werden (Heidel *et al.* 2011).

Das gesamte codierende Potential der einzelnen Genome schwankt nur gering, berechnet wurden zwischen 12.000 bis 13.000 codierende Sequenzen (CDS) für die jeweiligen Arten (Heidel *et al.* 2011; Sugang *et al.* 2011). Damit besitzen die sozialen Amöben rund doppelt so viele Gene wie *Saccharomyces cerevisiae* und halb so viele Gene wie der Mensch (Goffeau *et al.* 1996; Lander *et al.* 2001). Die Gene haben eine durchschnittliche Länge von 1,6 kb und besitzen nur wenige Introns. Entsprechend der geringen Größe der Genome und der vorhergesagten Anzahl von Genen ergibt sich bei allen Arten eine hohe Gendichte mit geringen intergenischen Abständen. Im Durchschnitt liegen zwischen zwei benachbarten Genen lediglich 1 kb nicht-codierender Sequenz (Heidel *et al.* 2011).

Des Weiteren zeigte sich, dass vor allem Gene, die für basale Zellfunktionen codieren, stark konserviert sind und in allen Arten Orthologe nachweisbar sind. Das betrifft vor allem Gene, die am primären Zellmetabolismus, der Bildung des Zytoskeletts und intrazellulären Signalprozessen beteiligt sind. Vermutet wird des Weiteren, dass viele Gene nicht als Orthologe identifiziert werden können, da der evolutionäre Abstand zwischen den Arten so groß ist, dass die Akkumulation von Mutationen eine gemeinsame Zuordnung verhindert (Heidel *et al.* 2011).

Durch die Analyse von Syntenie-Regionen konnte auch gezeigt werden, dass es im Lauf der Zeit viele Genomumstrukturierungen gegeben haben muss. So konnten nur 1079 benachbarte Gene gefunden werden, die im Vergleich zwischen *D. discoideum*, *D. fasciculatum* und *P. pallidum* ihre Nachbarschaft

behalten haben, was weniger als 10 % der gesamten Protein-codierenden Sequenzen entspricht (Heidel *et al.* 2011). Dies spricht für eine starke Dynamik während der individuellen Genomevolution.

Eine weitere Besonderheit des Genoms von *D. discoideum* ist der große Anteil von mobilen genetischen Elementen. Während rund 10 % der genomischen DNA von *Dd* auf mobile genetische Elemente zurückzuführen ist, sind die Anteile dieser Elemente in den Genomen von *Df*, *Dp* und *Pp* um ein vielfaches geringer (Heidel *et al.* 2011). Bekannt ist, dass repetitive DNA einen großen Einfluss auf die Dynamik der Genomentwicklung hat. Ein Hinweis auf die dadurch entstehende Genomplastizität ergibt sich beim Blick auf die Anzahl der tRNA-Gene in den jeweiligen Genomen. Es zeigte sich, dass sich die Anzahl der tRNA-Gene offensichtlich von Gruppe 1 zu Gruppe 4 verdoppelt hat (Abbildung 3). Zu vermuten ist, dass die Expansion der mobilen genetischen Elemente im Genom von *D. discoideum* mit verantwortlich für diese Unterschiede ist (Eichinger *et al.* 2005; Heidel *et al.* 2011).

Tab. 1: Charakteristische Merkmale der Genome unterschiedlicher Arten sozialer Amöben.
(Eichinger *et al.* 2005; Heidel *et al.* 2011; Sucgang *et al.* 2011).

	<i>D. discoideum</i>	<i>D. purpureum</i>	<i>P. pallidum</i>	<i>D. fasciculatum</i>
Genomgröße [Mb]	35	33	33	31
Anzahl Chromosomen	6	n.d.	7	6
globaler A/T-Gehalt [%]	77,6	75,5	68	66,2
A/T-Gehalt in CDS	72,6	n.d.	63,8	63,2
Größe Palindromarm [kb]	45	n.d.	15	28
Anzahl vorhergesagte CDS	13.433	12.410	12.373	12.173
Durchschnittliche Genlänge [bp]	1579	1760	1552	1672
Gendichte [CDS pro Mb]	396	376	375	392
durchschnittlicher Genabstand [bp]	~1000	~900	~1100	~900
Anzahl tRNA-Gene	418	353	273	198

Aus diesen Daten lässt sich ableiten, dass die Genome der sozialen Amöben insgesamt durchaus vergleichbar sind, dass jedoch eine große Genomplastizität existiert, welche, bedingt durch die großen evolutionären Distanzen, allerdings nicht erstaunt. *D. discoideum*, als am besten untersuchte Art, scheint jedoch im Vergleich mit Spezies der Gruppe 1 (*Dictyostelium fasciculatum*) und der Gruppe 2 (*Polysphondylium pallidum*) einige Charakteristika aufzuweisen, die nach jetzigem Wissensstand als uncharakteristisch für die Gesamtheit der sozialen Amöben einzuordnen sind.

1.2 Mobile genetische Elemente

Als Barbara McClintock in den 1940er Jahren des letzten Jahrhunderts reversible Verfärbungen von Maiskörnern beschrieb und zum ersten Mal die Existenz „mobiler“ DNA postulierte (McClintock 1948), war noch nicht absehbar, welchen Einfluss diese Elemente (TE: transposable Elemente) auf verschiedenste genetische Prozesse in vielen Organismen haben.

Mit der Bezeichnung „*selfish*“ oder „*junk*“ DNA zeigt sich der Stellenwert, den man mobilen genetischen Elementen eine lange Zeit gab (Orgel und Crick 1980). Zunächst wurden diese nur als genomische Parasiten betrachtet, die dem Wirt keine Vorteile verschaffen. Erst im Lauf der Zeit und vor allem mit dem Einsatz moderner molekularbiologischer Methoden, allen voran den neuen Hochdurchsatz-Sequenziermethoden („*Next-generation sequencing*“), die es ermöglichten die Biologie dieser Elemente besser zu verstehen, wurde immer deutlicher, dass transposable DNA einen großen Einfluss auf die Evolution, Speziation und Organisation von Genomen hat (Kazazian 2004; Cordaux und Batzer 2009; Chenais *et al.* 2012; Sun *et al.* 2012). Zudem konnten in letzter Zeit auch einige humane Krankheiten mit Integrationen von mobilen genetischen Elementen in Gene in Verbindung gebracht werden (Hancks und Kazazian 2012; Kaer und Speek 2013).

Durch die vollständige Sequenzierung der Genome verschiedenster Spezies wurde schnell offensichtlich, dass mobile genetische Elemente teilweise beträchtliche Anteile an der gesamten genomischen DNA einnehmen können. So besteht z. B. die genomische DNA des Mais zu ca. 85 % (Schnable *et al.* 2009), das Genom der Maus zu ca. 37,5 % (Chinwalla *et al.* 2002) und das menschliche Genom zu 46 % bis 65 % (Lander *et al.* 2001; de Koning *et al.* 2011) aus Sequenzen, die mobilen genetischen Elementen zugeordnet werden können (Abbildung 4). Dabei scheint es auch eine Abhängigkeit der Genomgrößen von dem darin enthaltenen Anteil an mobilen genetischen Elementen zu geben, denn vor allem extrem aufgeblähte Genome beinhalten sehr oft viele dieser Elemente (Sun *et al.* 2012).

Anhand dieser Beispiele lässt sich bereits erkennen, dass die Beziehungen zwischen den Wirtszellen und den mobilen genetischen Elementen, allein schon durch deren große Anteile an vielen Genomen, sehr eng sind. So gibt es auf Seiten der Wirtszelle Mechanismen, um die ungehemmte Amplifikation dieser Elemente zu unterdrücken, wie zum Beispiel die Eliminierung der RNA-Zwischenstufen durch die RNA-Interferenz, oder die epigenetische, transkriptionelle Inaktivierung der genomischen Kopien durch Methylierung (Buchon und Vauray 2006; Suzuki *et al.* 2007; Crichton *et al.* 2014). Möglich sind auch „*repeat-induced point mutations*“ (RIP), bei der repetitive Sequenzen durch Punktmutationen so verändert werden, dass keine erfolgreiche Mobilisierung mehr möglich ist (Attard *et al.* 2005). Ebenso versuchen einige Transposons genomische „Nischen“ zu besetzen (Chalker und Sandmeyer 1992; Marschalek *et al.* 1992), oder auch unterschiedliche entwicklungsabhängige Amplifikationsstrategien zu nutzen, um sich der Wirtsabwehr zu entziehen (Cecco *et al.* 2013).

Heute bilden mobile genetische Elemente eine eigene Klasse innerhalb der repetitiven DNA-Elemente. Zu diesen repetitiven Elementen zählen des Weiteren die Satelliten-DNA, zu denen die „*short tandem*

repeats“ sowie die Mini- und Mikrosatelliten zählen und ebenso Solo-LTR-Sequenzen, welche Relikte früherer Integrationen von Retroviren bzw. Retrotransposons darstellen (Charlesworth *et al.* 1994). Generell können die mobilen genetischen Elemente in zwei große Gruppen unterteilt werden. Die Retrotransposons (Klasse I) und die DNA-Transposons (Klasse II) (Wicker *et al.* 2007). Zugrunde liegt dieser Unterteilung ein unterschiedlicher Transpositions-Mechanismus (Abbildung 5). Während DNA-Transposons mittels eines sogenannten „*cut and paste*“-Mechanismus ohne Amplifikationsschritt direkt mobilisiert werden, nutzen Retrotransposons einen „*copy and paste*“-Mechanismus. Dabei wird über eine RNA-Zwischenstufe und einer anschließenden reversen Transkription der RNA in doppelsträngige DNA ein neues Element generiert.

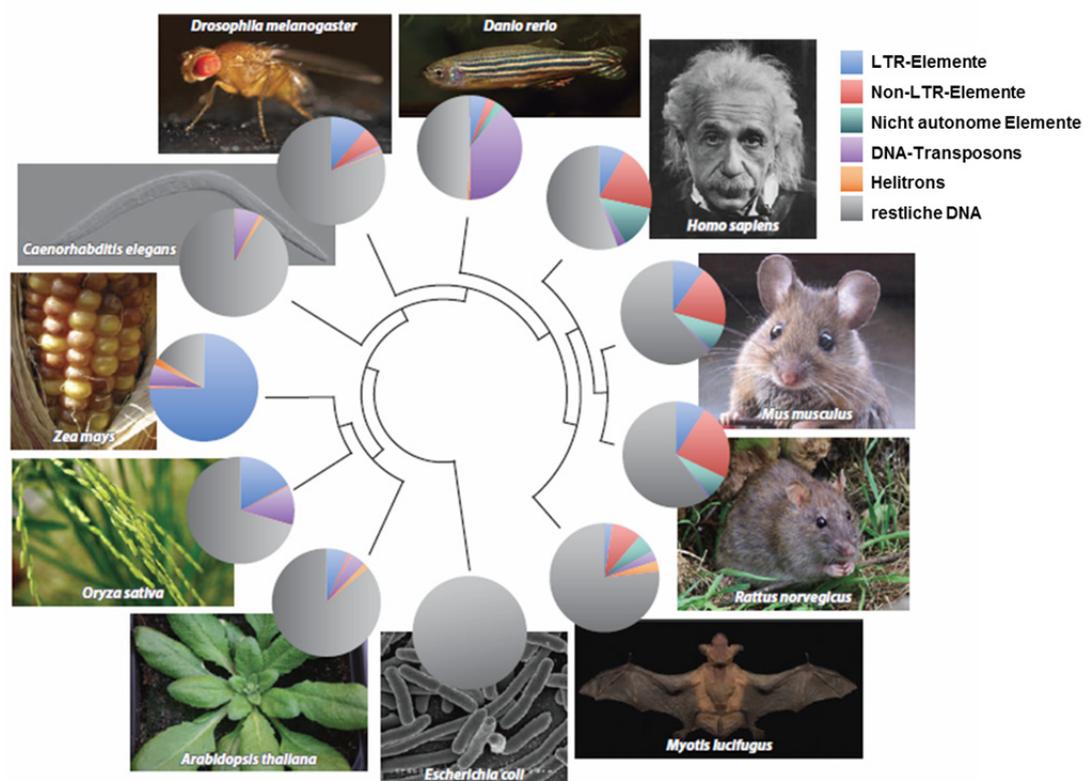


Abb. 4: Anteile und Verteilung mobiler genetischer Elemente an Genomen verschiedener Spezies.
(Huang *et al.* 2012)

Somit unterscheiden sich beide Gruppen transposabler Elemente grundlegend darin, dass die Kopienzahl bei der Mobilisierung von DNA-Transposons konstant bleibt, wohingegen Retrotransposons mit jedem Sprungereignis amplifiziert werden (Abbildung 5). Eine Amplifikation der DNA-Transposons ist daher nur durch Ereignisse möglich, die auf Wirtszell-basierten Reparaturmechanismen nach der Exzision des Retrotransposons, oder Sprungereignissen während der Zellteilung beruhen (Craig *et al.* 2002). Die Verhältnisse von DNA-Transposons und Retrotransposons schwanken teilweise erheblich zwischen verschiedenen, auch nahe verwandten, Spezies (Feschotte und Pritham 2007).

Die Gruppe der Retrotransposons wird in die LTR- (*long terminal repeat*) und Non-LTR-Retrotransposons unterteilt (Wicker *et al.* 2007). Die Non-LTR-Elemente werden Synonym auch als LINE (engl. „*long interspersed nuclear elements*“) bezeichnet. Eine weitere Unterteilung der Non-LTR-Retrotransposons erfolgt in autonome und nicht-autonome Elemente. Die nicht-autonomen Elemente besitzen keine proteincodierenden offenen Leserahmen (ORFs), um eigenständig zu transponieren. Sie sind daher auf die Hilfe der Proteine autonomer Elemente angewiesen (*Transmobilisierung*). Das häufigste humane, nicht-autonome Non-LTR-Retrotransposon ist das Alu-Element, welches ca. 10 % des gesamten Genoms ausmacht (Lander *et al.* 2001; Batzer und Deininger 2002). Das bekannteste humane autonome Retrotransposon, das 6 kb große L1-Element, kommt in etwa 500.000 Kopien im menschlichen Genom vor, was ungefähr 17 % der gesamten genomischen DNA entspricht (Singer *et al.* 1993; Lander *et al.* 2001). Allerdings sind offensichtlich nur rund 80 bis 100 Kopien aktiv, da die restlichen Elemente 5' deletiert und somit inaktiviert sind (Sassaman *et al.* 1997; Brouha *et al.* 2003). Sowohl DNA-Transposons, als auch LTR-Retrotransposons sowie die Retroviren zeigen im Bereich der Integrase-/Transposasedomäne starke Ähnlichkeiten (Capy *et al.* 1997). Das gemeinsame DD(35)E-Motiv hat als katalytische Triade eine essentielle Funktion, die evolutionär stark konserviert ist. Bereits der Austausch einer Aminosäure innerhalb des Motivs inaktiviert das entsprechende mobile Element (Doak *et al.* 1994; Vos *et al.* 1996; Lohe *et al.* 1997).

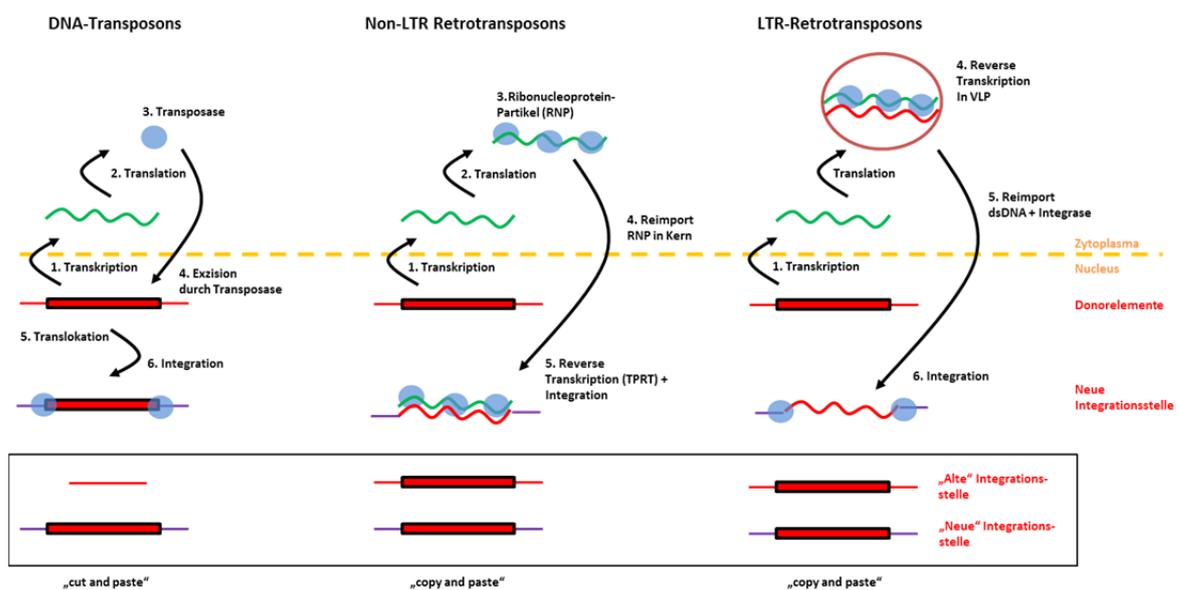


Abb. 5: Charakteristische „Lebenszyklen“ der drei Klassen mobiler genetischer Elemente.

Während die DNA-Transposons direkt mobilisiert werden und keine (direkte) Amplifizierung erfahren wird bei den Retrotransposons über eine RNA-Zwischenstufe ein neues Element erzeugt. Dabei unterscheiden sich beide Klassen von Retrotransposons durch den verwendeten Mechanismus zur Erzeugung der cDNA (vgl. 1.3.1.1. und 1.3.2.1). RNA ist in Grün dargestellt. TPRT: „*Target-primed reverse transcription*“; VLP: „*Virus-like particle*“

1.3 Retrotransposons

Zur Gruppe der Retrotransposons zählen alle mobilen genetischen Elemente, die im Zuge ihrer Mobilisierung eine RNA als Zwischenstufe verwenden (Wicker *et al.* 2007). Charakteristisches Merkmal dieser Gruppe ist daher die Nutzung einer RNA-abhängigen DNA-Polymerase, der Reversen Transkriptase (RT). Anhand konservierter Domänen der RT und strukturellen Unterschieden können mehrere Gruppen von Retrotransposons unterschieden werden (Wicker *et al.* 2007; Eickbush und Jamburuthugoda 2008).

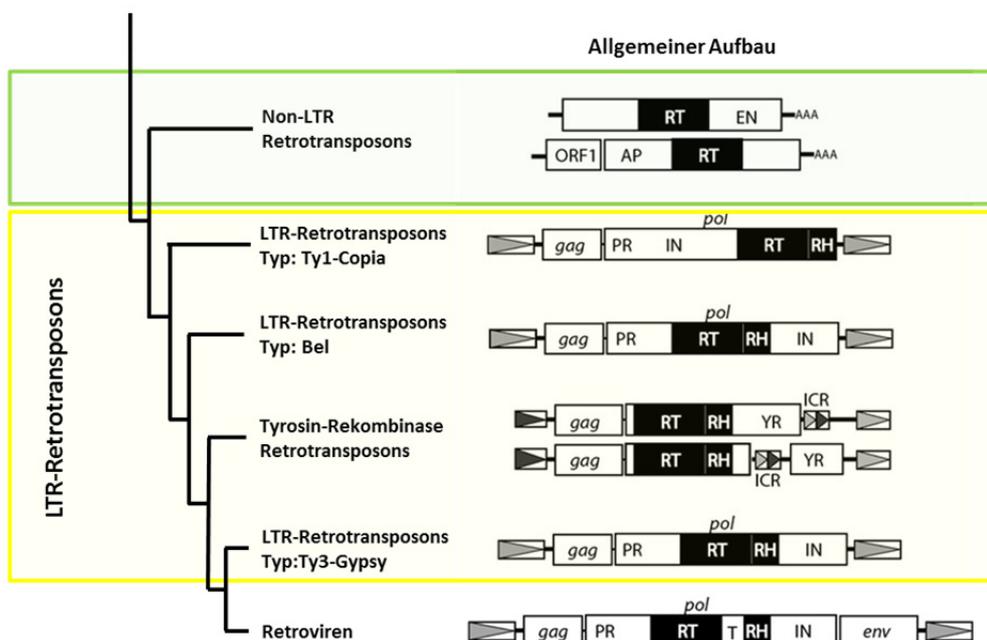


Abb. 6: Struktureller Aufbau, Klassifizierung und Phylogenie der Retrotransposons.

Der Stammbaum wurde auf Basis der Ähnlichkeiten der jeweiligen Reversen Transkriptase-Domänen errechnet. Markiert sind die beiden großen Klassen der Retrotransposons. Abkürzungen: RT: Reverse Transkriptase; EN: Endonuklease; AP: Apurinische Endonuklease; EN: Integrase; Uri: mobile Gruppe I-Intron ähnliche Endonuklease; YR: Tyrosin Rekombinase; ICR: Interne komplementäre Wiederholung; gag: Gruppenspezifisches Antigen; env: Hüllprotein; RH: RNase-H-Domäne. Modifiziert nach Eickbush und Jamburuthugoda 2008.

Es wird davon ausgegangen, dass alle Retrotransposons einen gemeinsamen Vorgänger haben, was mit der starken Konservierung der RT-Domäne begründet wird (Xiong und Eickbush 1990). Ein ebenfalls verbreitetes Merkmal der Retrotransposons ist das Vorhandensein von terminalen Sequenzwiederholungen (LTR, engl.: „*long terminal repeats*“). Lediglich die Gruppe der Non-LTR-Retrotransposons verzichtet auf diese Sequenzduplikationen, hier findet sich ein völlig anderer Mechanismus der Transposition als bei den LTR-Retrotransposons (Luan *et al.* 1993). Die Retrotranspositionsstrategie der LTR-Retrotransposons basiert dahingegen vollständig auf der modularen Struktur der LTRs. Beide Retrotransposon-Gruppen benötigen aber in funktioneller

Hinsicht weitgehend identische Proteine für eine erfolgreiche Mobilisierung. Dazu zählt zum einen die Reverse Transkriptase, welche die RNA-Zwischenstufe in eine doppelsträngige DNA umschreibt. Des Weiteren wird eine Endonuklease benötigt, um die Ziel-DNA zu schneiden, sowie eine Integraseaktivität, um am Ende eine neue Retrotransposonkopie zu integrieren.

Die Retroviren zählen im engeren Sinn nicht zu den endogenen mobilen genetischen Elementen, da sie durch den Erwerb eines „*envelope*“-Gens (*env*, dt.: „Hülle“), welches sie dazu befähigt die Wirtszellen zu verlassen und andere Zellen zu infizieren, einen primär nicht auf die Wirtszelle beschränkten Mobilisierungsmechanismus verwenden (Malik 2000). Allerdings wurde nachgewiesen, dass die Retroviren aus den LTR-Elementen hervorgegangen sind. Dadurch gleichen sie auch strukturell den endogenen Retrotransposons und werden somit phylogenetisch zur Gruppe der Retrotransposons gezählt (Xiong und Eickbush 1990). Zudem können Retroviren durch den Verlust der für die extrazellulären Verbreitung notwendigen Gene wieder zu obligat intrazellulären mobilen genetischen Elementen werden (Wicker *et al.* 2007).

1.3.1 LTR-Retrotransposons

Charakteristisches Merkmal dieser Gruppe mobiler genetischer Elemente ist das Vorhandensein von terminalen Sequenzwiederholungen, den LTRs. Die Gesamtgröße der Elemente variiert von wenigen 100 bp bis zu mehreren 1000 bp, wobei die meisten Elemente zwischen 4 und 10 kb groß sind. Die LTRs der Retroelemente werden durch kurze, terminale, invertierte Wiederholungen begrenzt, die meist mit dem Motiv 5'-TG..CA-3' assoziiert sind (Boeke und Corces 1989). Die LTRs sind aus drei Teilen zusammengesetzt, bestehend aus den Regionen U3-R-U5. Die U3-Region besitzt dabei sowohl Enhancer- als auch Promotoraktivität (Arkhipova *et al.* 1986; Falzon und Kuff 1988; Boeke und Corces 1989). Ein häufiges Merkmal von Retrotransposonintegrationen sind flankierende, meist wenige Basenpaar lange, Zielsequenzduplikationen (TSD, engl.: „*Target site duplication*“), welche auch Möglichkeiten zur automatischen Erkennung von LTR-Sequenzen bieten (Benachou *et al.* 2009). Die TSD entsteht während der Integration der neuen Retrotransposon-Kopie in die genomische Ziel-DNA durch versetzte Schnitte der genomischen DNA und dem Auffüllen der Lücken nach erfolgreicher Integration der Retrotransposon-Kopie durch zelleigene Reparaturmechanismen (Boeke und Corces 1989). Nahezu alle autonomen Retrotransposons besitzen zwei ORFs, welche für das Gruppenspezifische Antigen (*gag*; engl.: „*Group specific antigen*“) und ein Polyprotein (*pol*) codieren (Havecker *et al.* 2004). Diese können in einem gemeinsamen oder in getrennten ORFs vorliegen. Interessanterweise werden auch bei getrennten ORFs beide Gene als gemeinsame RNA transkribiert. Die Translation geschieht durch Frameshifts des Ribosoms, also ein Wechsel während der Translation in einen anderen Leserahmen, oder durch das Überlesen des Stoppcodons des *gag*-Gens (Wilhelm und Wilhelm 2001; Gao 2003). Das *gag*-Gen codiert für die Proteine welche für die Bildung der cytosolischen „*Virus-like-particles*“ (VLP; dt.: „Virus-ähnliche Partikel“) benötigt werden. In den VLPs findet die reverse Transkription und somit die Bildung der doppelsträngigen DNA aus der mRNA des Retroelements statt (siehe Abbildung 5) (Eichinger und Boeke 1988). Die für die reverse

Transkription und anschließende Integration benötigten Proteine sind im *pol*-Gen codiert, welche aus dem Polyprotein durch die Protease herausgeschnitten werden. Dazu zählen Domänen für die Protease (PR), Endonuklease (EN), die Reverse Transkriptase (RT), die Ribonuklease-H (RH oder RNase-H) sowie die Integraseaktivität (IN) (Eickbush und Jamburuthugoda 2008). Beträchtliche Anteile von LTR-Retrotransposons wurden vor allem in Pflanzengenomen gefunden (SanMiguel *et al.* 1996; Bennetzen *et al.* 2005), sie kommen aber auch in geringerem Maß in Säugetierzellen vor, dort vor allem in Form von sogenannten endogenen Retroviren und Solo-LTRs. Rund 8 % des menschlichen Genoms können auf LTR-Retroelemente zurückgeführt werden (Lander *et al.* 2001).

1.3.1.1 Die Retrotransposition der LTR-Retrotransposons

Der Mechanismus der Retrotransposition wurde vor allem an dem LTR-Retrotransposon Ty1 aus *Saccharomyces cerevisiae* untersucht (Lauermaun und Boeke 1997), kann aber auf die meisten LTR-Retrotransposons übertragen werden.

Die Retrotransposition der LTR-Elemente beginnt mit der Transkription der genomischen Ausgangskopie durch die RNA-Polymerase-II und der Erzeugung einer RNA, welche am Start der 5'-R-Region beginnt und am Ende der 3'-R-Region terminiert (Abbildung 7A). Diese RNA wird in zwei Kopien, gemeinsam mit den benötigten Proteinen, in die Virus-ähnlichen Partikel gepackt. Anschließend bindet eine zelleigene tRNA an die Primerbindestelle (PBS) des Transkripts und die Reverse Transkriptase synthetisiert, mit Hilfe der tRNA als Primer, ein kurzes Stück cDNA. Diese besteht aus der 5'-U5- und R-Region (Abbildung 7B), welche als „*Minus strand strong stop*“-DNA (-sssDNA) bezeichnet wird (Lauermaun und Boeke 1997; Wilhelm und Wilhelm 2001). Die dazu komplementäre RNA wird währenddessen durch die RNase-H-Aktivität der RT degradiert (Abbildung 7C)(Champoux 1993). Nach dieser ersten Elongation transloziert die cDNA und bindet mit der R-Region an den komplementären Bereich der 3'-R-Region an die RNA-Matrize (Abbildung 7D). Dabei kann entweder das gleiche RNA-Molekül verwendet werden, oder es findet eine intermolekulare Translokation der -sssDNA zu einem kompletten RNA-Strang statt. Davon hängt ab, ob die -sssDNA im nächsten Schritt die 5'-U5-R-Region enthält oder nicht. Bei einer intramolekularen Translokation, in der diese Region nicht synthetisiert werden kann (Verdau durch RNase-H in Schritt C), wird im letzten Schritt die zur PBS komplementäre Region am 3'Ende der -sssDNA verwendet (Lauermaun und Boeke 1997). Nun wird die -sssDNA durch die RT bis zum Ende der 5'-U5-Region verlängert (Abbildung 7E). Gleichzeitig wird die komplementäre RNA durch die RNase-H-Aktivität der RT abgebaut, wobei ein kurzer Bereich der RNA in der Poly-Purin-Region (PPT, engl.: „*Poly-purin tract*“) verbleibt und dieser im nächsten Schritt als Primer für die Erzeugung der „*Plus strand strong stop*“-DNA (+sssDNA) durch die RT dient (Abbildung 7F/G) (Heyman *et al.* 2003). Der PPT-Primer wird anschließend durch die RNase-H-Aktivität der RT degradiert. Nun folgt eine Translokation der +sssDNA zur R-U5-Region des 5'-LTRs (Abbildung 7H). Als letzter Schritt der Bildung einer doppelsträngigen DNA aus den einzelsträngigen RNA-Matrizen werden beide Stränge vervollständigt

(Abbildung 7I) (Lauermaun und Boeke 1997; Wilhelm und Wilhelm 2001). Somit kann die neu synthetisierte Kopie in eine neue Stelle der Wirts-DNA integriert werden.

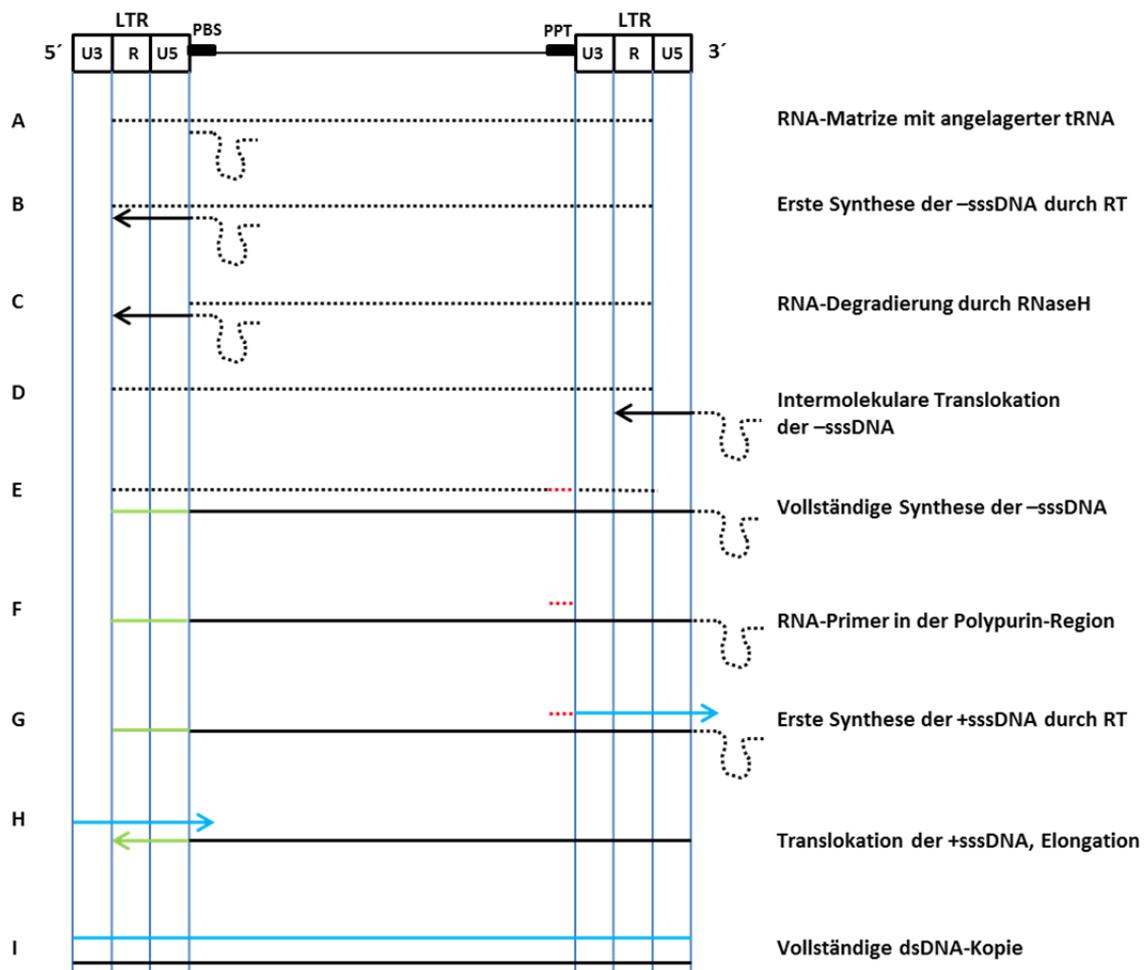


Abb.7: Modell der reversen Transkription der LTR-Retrotransposons.

Punktierte Linien stellen RNA, durchgezogene Linien DNA dar; In Grün ist der Teil der -sssDNA markiert, der nur bei der intermolekularen Translokation der -sssDNA vorhanden ist. Modifiziert nach Lauermaun und Boeke 1997.

1.3.2 Non- LTR-Retrotransposons

Die Familie der Non-LTR-Retrotransposons, oder LINEs (engl.: „*long interspersed nuclear elements*“), wurde, verglichen mit den DNA-Transposons und den LTR-Elementen, erst spät entdeckt (Weiner *et al.* 1986). Das ist vor allem darauf zurückzuführen, dass in dieser Gruppe keine redundanten Regionen, wie LTRs oder invertierte terminale Wiederholungen (ITR, vgl. DNA-Transposons) existieren. Auffallend groß ist auch die Anzahl deletierter Kopien von Non-LTR-Elementen. Diese entstehen durch vorzeitige Abbrüche während der reversen Transkription. Aus diesem Grund sind ein großer Teil der neu integrierenden Kopien inaktiviert (Sassaman *et al.* 1997).

Eines der am besten untersuchten Non-LTR-Retrotransposons ist das menschliche L1-Element, welches mit rund 500.000 Kopien ungefähr 17 % der gesamten genomischen DNA ausmacht (Lander *et al.* 2001). LINE-Elemente bestehen im Allgemeinen aus einer 5'-untranslatierten Region (5'-UTR), welche eine interne Promotoraktivität besitzt (Swergold 1990). Darauf folgt ein erster offener Leserahmen (ORF1), der für ein RNA-bindendes Protein mit essentieller Chaperonaktivität codiert (Martin *et al.* 2000; Martin *et al.* 2005). An dieses erste Gen schließt der ORF2 an, der für ein Multidomänenprotein mit einer Endonuklease- (Feng *et al.* 1996), Reversen Transkriptase- (Mathias *et al.* 1991) und in einigen Fällen einer eigenen RNase-H-Domäne codiert (Wicker *et al.* 2007) (Abbildung 8). Des Weiteren lässt sich am Ende des zweiten ORFs in der Regel eine Cystein-reiche Domäne ohne bekannte Funktion nachweisen (Fanning und Singer 1987). Das 3'-Ende der Non-LTR-Retrotransposons wird von einer weiteren untranslatierten Region gebildet (3'-UTR), die in einem homopolymeren poly(A)-Anhang endet (Dombroski *et al.* 1991).

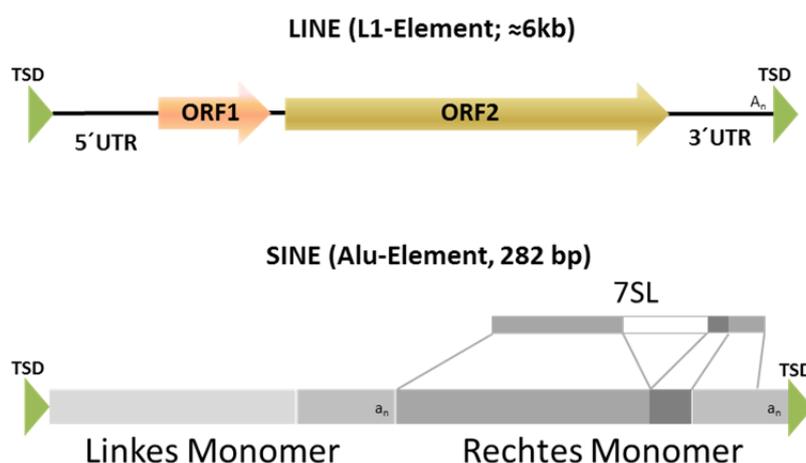


Abb. 8: Struktureller Aufbau der LINE- und SINE-Elemente.

Repräsentativ für die autonomen Non-LTR-Elemente (LINE), ist das L1-Element dargestellt. Als Vergleich wird darunter das Alu-Element gezeigt, welches zu den nicht-autonomen Non-LTR-Elementen (SINE) zählt. Beide Retrotransposons sind im menschlichen Genom zu finden (Lander *et al.* 2001).

Im Gegensatz zu den autonomen Elementen besitzen die nicht-autonomen, oder SINE (engl.: „*Short interspersed nuclear elements*“) genannten, Non-LTR-Retrotransposons keine codierenden Kapazitäten (Abbildung 8). Aus diesem Grund sind sie auf die *Trans*-Mobilisierung durch Proteine autonomer Non-LTR-Elemente angewiesen (Ostertag und Kazazian Jr 2001; Raiz *et al.* 2012). Im menschlichen Genom sind die nicht autonomen Non-LTR-Retrotransposons wahrscheinlich durch die reverse Transkription von zellulären RNA-Polymerase-III-Transkripten durch die RT-Proteine autonomer Elemente entstanden. Nachweisbar sind Homologien mit 7SL-, 5S- und tRNA-Genen (Ohshima und Okada 2005), zudem besitzen sie interne RNA-Polymerase-III-Promotoren. Die Integrationen nicht-autonomer Elemente haben starke Ähnlichkeiten mit denen der autonomen LINE1-

Elemente, was ebenfalls für die *Trans*-Mobilisierung durch Proteine der LINE-Elemente spricht (Ostertag und Kazazian Jr 2001). Es ist auch bekannt, dass inaktivierte Kopien autonomer Elemente durch *Trans*-Mobilisierung erneut transponieren können, jedoch geschieht das mit einer sehr geringen Effizienz. Das zeigt, dass offenbar eine starke *Cis*-Präferenz vorliegt, also die Bevorzugung der eigenen Transkripte durch die codierten Proteine autonomer Elemente (Wei *et al.* 2001).

1.3.2.1 Der Retrotranspositionsmechanismus der Non-LTR-Retrotransposons

Der Mechanismus zur Mobilisierung beginnt, wie auch bei den LTR-Retrotransposons, mit der Transkription einer kompletten genomischen Kopie durch die zelleigene RNA-Polymerase-II (Abbildung 9A) (Mizrokhi *et al.* 1988). Anschließend wird diese RNA in das Cytoplasma transportiert, wo die Translation des bicistronischen Transkripts stattfindet. Für beide daraus entstehenden Proteine (ORF1 und ORF2) wird angenommen, dass sie eine starke *Cis*-Präferenz für ihr eigenes Transkript besitzen (Wei *et al.* 2001; Martin *et al.* 2005). Dadurch entsteht ein Ribonuklein-Proteinkomplex (RNP), wobei das ORF1-Protein die RNA für den Transport in den Zellkern stabilisiert (Hohjoh und Singer 1996; Doucet *et al.* 2010). Nachdem der RNP-Komplex die Ziel-DNA erreicht erfolgt die reverse Transkription. Non-LTR-Retroelemente nutzen dafür den TPRT-Mechanismus („*Target primed reverse transcription*“; Abbildung 9B) (Hohjoh und Singer 1996).

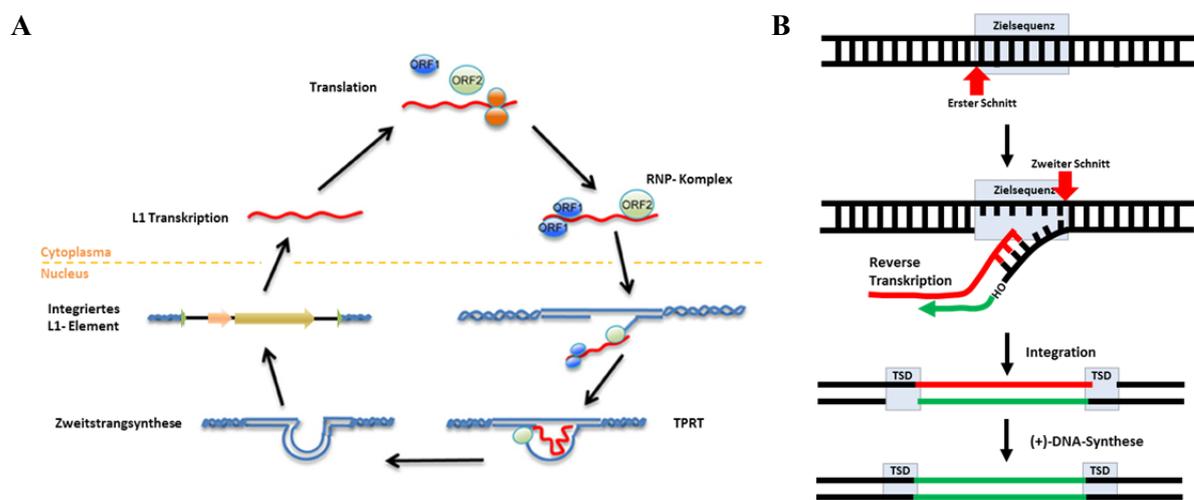


Abb. 9: Mechanismus der Retrotransposition bei Non-LTR-Retrotransposons.

(A) „Lebenszyklus der Non-LTR-Retrotransposons. Modifiziert nach Arias *et al.* 2012. (B) Der „*Targed primed reverse transcription*“-Mechanismus. Durch die Aktivität der Endonuklease entsteht das für die reverse Transkription benötigte 3'-OH-Ende. Durch auffüllen der versetzten Schnitte entstehen die charakteristischen Zielsequenzduplikationen. Rot: RNA; Grün: DNA. Modifiziert nach Ostertag *et al.* 2001.

Dazu schneidet zunächst die Endonuklease den Minus-Strang der Ziel-DNA. Anschließend lagert sich die RNA an kurze komplementäre Bereiche des DNA-Einzelstrangs an. Durch die an diesem Strang

frei gewordene 3'-Hydroxylgruppe, die als Primer fungiert, kann nun die reverse Transkriptase die cDNA des Retrotransposons erzeugen (Luan *et al.* 1993). Zur Integration der cDNA wird anschließend durch die Endonuklease der Plusstrang etwas oberhalb des ersten Schnittes geschnitten. Dadurch entsteht ein versetzter Schnitt, der nach Integration der cDNA und der Synthese des komplementären Strangs, durch die Auffüllung der entstandenen Überhänge, zu den charakteristischen Zielsequenzduplikationen (TSD) führt. Dies wird wahrscheinlich durch die zelleigenen DNA-Reparaturmechanismen realisiert (Luan *et al.* 1993; Gilbert *et al.* 2002). Es ist aber nach wie vor nicht im Detail bekannt, wie der komplementäre DNA-Strang erzeugt und die neue Kopie integriert wird. Möglicherweise dient der frei werdende 3'-Überhang der (+)-Strang Ziel-DNA als Primer für die Synthese des zweiten Strangs (Luan *et al.* 1993; Gilbert *et al.* 2002).

1.4 DNA-Transposons

Im Gegensatz zu Retrotransposons werden DNA-Transposons ohne RNA-Zwischenstufe mobilisiert. Daher benötigen sie ein geringeres Repertoire an Proteinen, um aktiv zu transponieren, als die LTR-Retroelemente. DNA-Transposons besitzen invertierte terminale Sequenzwiederholungen (ITR, engl.: „*inverted terminal repeat*“) und verfügen über mindestens einen ORF, der für eine Transposase codiert, die sowohl eine Endonuklease- als auch Integraseaktivität besitzt (Abbildung 10). Die Transposase, welche durch eine in der 5'-ITR codierte Promotoraktivität transkribiert wird, erkennt die ITRs, schneidet das DNA-Transposon mit Hilfe der Endonukleaseaktivität aus der Donor-DNA und integriert es in einen neuen Locus des Wirtsgenoms (Vos *et al.* 1996). Die Integrationsstellen sind in den meisten Fällen ebenfalls von TSD flankiert (Wicker *et al.* 2007).

Die DNA-Transposons werden in zwei Subklassen eingeteilt. Unterschieden wird dabei die Anzahl der geschnittenen Stränge während der Transposition. Die erste Subklasse besteht aus mehreren Familien, die sich als klassische „*cut and paste*“-DNA-Transposons charakterisieren lassen (Wicker *et al.* 2007). Anders als bei den Elementen der Subklasse II wird während der Exzision des DNA-Transposons der Doppelstrang der Donor-DNA geschnitten. Charakteristisch für die jeweiligen Unterklassen sind TSD mit bestimmten Längen und die individuellen Größen der ITRs. Die bekanntesten Vertreter dieser Klasse sind die Tc1/Mariner-Elemente. Elemente dieser Klasse sind zwischen 1,3 bis 2,4 kb lang, codieren für eine Transposase und besitzen ITRs. Tc1/Mariner-Elemente integrieren in TA-Dinukleotide (Miskey *et al.* 2005). Die zweite Subklasse der DNA-Transposons besteht aus Elementen, die während ihrer Mobilisierung nur einen Strang schneiden und einen völlig anderen Aufbau wie Elemente der Subklasse I haben (Wicker *et al.* 2007). Dazugehörige DNA-Transposons, wie zum Beispiel die Helitrons, replizieren über „*rolling circle*“-Mechanismen ohne Zielsequenzduplikationen zu erzeugen (Kapitonov und Jurka 2001).



Abb. 10: Allgemeiner Aufbau eines klassischen DNA-Transposons.
(Abbildung nicht maßstäblich)

1.5 Einfluss mobiler genetischer Elemente auf Wirtsgenome

Allein durch die zum Teil beträchtlichen Anteile der mobilen genetischen Elemente an der gesamten genomischen DNA der Wirte ergibt sich ein starker Einfluss auf die Genomplastizität. Ein potentieller Fitnessverlust des Wirts hätte im Lauf der Evolution unweigerlich zur Extinktion der genetischen Elemente geführt. Daher liegt der Schluss nahe, dass die Beherbergung von mobilen genetischen Elementen eher mit Vorteilen für den Wirt assoziiert sind.

Durch ihre genomische Redundanz dienen repetitive Sequenzen als potentielle Loci für homologe nicht-allelische Rekombinationsereignisse (Schwartz *et al.* 1998; Hill *et al.* 2000; Vitte und Panaud 2003). Diese Ereignisse sorgen somit für Genom-Umstrukturierungen, welche in der Wirtszelle neue genetische Möglichkeiten generieren können. Sehr häufig hinterlassen LTR-Retrotransposons durch homologe nicht-allelische Rekombinationen am ursprünglichen Integrationsort Solo-LTRs. Da die 3'-UTR-Region von LTRs sowohl Enhancer als auch Promotoraktivität besitzt, können die Solo-LTRs damit Einfluss auf die Expression flankierender Gene haben (Biémont und Vieira 2006). Doch nicht nur Solo-LTRs haben das Potential, flankierende Gene zu beeinflussen, auch von autonomen mobilen Elementen sind solche Einflüsse bekannt. Weiterhin konnte gezeigt werden, dass häufig im Zuge der Transposition 3'- oder 5'-flankierende Bereiche transloziert werden, die an einem anderen Locus auch neue Funktionen übernehmen können (Goodier *et al.* 2000). Möglich sind auch Integrationen in Introns, die neue Spleißvarianten erzeugen und somit alternative Transkripte generieren (Sorek *et al.* 2002; Britten 2006). Zudem konnte gezeigt werden, dass die 5'-UTR der humanen L1-Elemente eine „Antisense“-Promotoraktivität besitzt, welche ausreicht, um 5'-flankierende Gene zu transkribieren (Speck 2001).

Neben einem indirekten Einfluss mobiler genetischer Elemente auf die flankierenden genomischen Regionen sind auch viele Beispiele bekannt, in denen neue Gene direkt aus TEs hervorgegangen sind (Brandt *et al.* 2005; Volf 2006; Oliver und Greene 2009). Das prominenteste Beispiel ist die Telomerase, welche in eukaryotischen Zellen die linearen Chromosomenenden (Telomere) verlängert (Eickbush 1997). Telomere bestehen aus kurzen, tandemartig wiederholten Sequenzen, die sich bei jeder Zellteilung verkürzen. Die Telomerase kann mit Hilfe einer komplementären RNA-Matrize die Telomere nach jeder Zellteilung verlängern und somit das Hayflick-Limit aufheben (Hayflick 1965). Das ist vor allem für sich endlos teilende Zellen entscheidend, wie Stammzellen oder vegetativ einzellige Eukaryoten. Es konnte gezeigt werden, dass eine enge Verwandtschaft mit der reversen Transkriptase der Non-LTR-Retrotransposons besteht. Es scheint deshalb offensichtlich, dass hier eine

Domestikation einer Reversen Transkriptase aus dieser Klasse der mobilen genetischen Elemente stattgefunden hat (Eickbush 1997). Ein weiterer, jedoch im Vergleich zu den meisten Eukaryoten völlig anderer Mechanismus der Telomer-Erhaltung findet sich bei der Fruchtfliege *Drosophila melanogaster*. Hier werden die Chromosomenenden vollständig von Non-LTR-Retrotransposons erzeugt und stabilisiert (Pardue und DeBaryshe 2003). Ein weiteres Beispiel für die Domestikation von mobilen genetischen Elementen sind die RAG1- und RAG2-Gene des Menschen, deren Produkte direkt an der V(D)J-Rekombination der Antikörpergene beteiligt sind (Agrawal *et al.* 1998). Eine weitere Möglichkeit der TE-abhängigen Regulation von Wirtsgenen besteht in der Erzeugung von miRNAs durch mobile genetische Elemente, die innerhalb der RNA-Interferenz-Maschinerie ihre Wirkung entfalten (Piriyaopongsa *et al.* 2007; Ahn *et al.* 2013). Anhand dieser Beispiele lässt sich der starke Einfluss der mobilen genetischen Elemente auf die individuelle Genomevolution erkennen.

Neben den positiven Effekten können mobile genetische Elemente allerdings auch negative Auswirkungen auf die Genome ihrer Wirte haben. Ein offensichtliches Problem ist die Integration in Gene und die damit verbundene Mutagenese. Wenn es sich dabei um essentielle Gene handelt, kann dies direkte drastische Folgen für den Wirt haben. Aber auch die Akkumulation TE-induzierter rezessiver Mutationen kann die Fitness einer Spezies reduzieren. Ungefähr 0,27 % aller Mutationsbedingten Erkrankungen des Menschen lassen sich auf Integrationen mobiler genetischer Elemente zurückführen (Callinan und Batzer 2006; Kaer und Speek 2013). Da diese Ereignisse mit einer negativen Selektion verknüpft sind, erstaunt es nicht, dass mobile genetische Elemente in Regionen des menschlichen Genoms mit einer hohen Gendichte unterrepräsentiert sind (Medstrand *et al.* 2002). Ein weiteres Phänomen ist die Deletion von flankierenden Bereichen der Ziel-DNA durch Non-LTR-Retrotransposons (Han *et al.* 2005; Callinan und Batzer 2006). Ebenso können auch die intra- und interchromosomalen nicht-allelischen Rekombinationen nachteilige Effekte für den Wirt haben (Moran 1999; Sen *et al.* 2006).

1.5.1 Wechselwirkungen zwischen mobilen genetischen Elementen und Wirtszellen

Die Möglichkeiten der Regulation durch den Wirt und die für die Genomevolution funktionellen Eigenschaften der repetitiven Elemente legen die Vermutung nahe, dass es sich teilweise um eine Art Domestikation der TEs durch die Wirtszellen handelt (Miller *et al.* 1997). Denkbar ist ein Gleichgewichtszustand, der beeinflusst wird durch die geregelte Aktivität mobiler genetischer Elemente, aus denen sich unter Umständen Vorteile für den Wirt ergeben, und der daraus resultierenden Genomstabilität (Abbildung 11) (Oliver und Greene 2009).

Da es seit jeher eine Koevolution zwischen den mobilen genetischen Elementen und den Wirtszellen gab, sind im Lauf der Zeit auch verschiedene Mechanismen entstanden, die den Wirtszellen die Möglichkeit geben die TEs zu kontrollieren. Ebenso versuchen die mobilen Elemente auf vielfältige Weise stabil weitervererbt zu werden. Vor allem die RNA-Interferenz gilt als ein basaler Wirtsabwehrmechanismus gegenüber TEs, der sich offensichtlich auch als Antwort auf

intrachromosomale „Nukleinsäure-Parasiten“ entwickelt hat (Buchon und Vaury 2006). Die Regulation der mobilen genetischen Elemente basiert dabei auf der Erzeugung kurzer, einzelsträngiger RNA-Oligonukleotide aus dsRNA, der Bindung dieser an die Ziel-mRNAs und deren anschließenden Degradierung durch den RISC (engl.: „*RNA induced silencing complex*“), an dem verschiedene Proteine beteiligt sind (Hannon 2002).

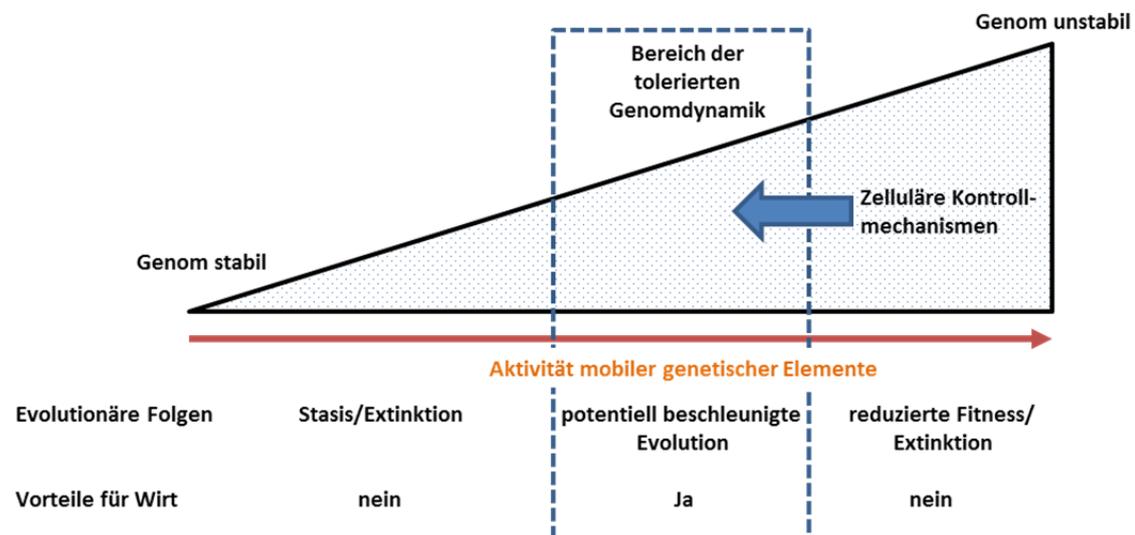


Abb. 11: Genomdynamik und TE-Aktivität.

Evolutionäre Vorteile für Wirte von mobilen genetischen Elementen in Abhängigkeit von der Aktivität der transposablen Elemente. Modifiziert nach Oliver und Greene 2009.

Des Weiteren wird vermutet, dass auch die Methylierung von genomischer DNA primär zur Inaktivierung von mobilen genetischen Elementen diene und sich erst später zu einem Werkzeug der Genregulation entwickelte (Barlow 1993; Yoder *et al.* 1997). Außer in sich hauptsächlich mitotisch teilenden Einzellern, wie zum Beispiel den sozialen Amöben, sind mobile genetische Elemente des Weiteren darauf angewiesen in pluripotenten- oder Keimbahnzellen aktiv zu sein, da somatische Retrotranspositionseignisse nicht weiter vererbt werden. Es konnte gezeigt werden, dass TEs gerade in diesen Zelltypen aktiviert werden (Ostertag *et al.* 2002; van den Hurk *et al.* 2007; Ribet *et al.* 2008). Viele mobile genetische Elemente befinden sich in Heterochromatinbereichen, die die meiste Zeit des Zellzyklus transkriptionell inaktiviert sind. Wahrscheinlich nutzen sie in den Phasen der Keimbahnentwicklung die globale Hypomethylierung der DNA um transkriptionell reaktiviert zu werden (Popp *et al.* 2010). Neben der Methylierung von Cytosin-Nukleotiden konnte gezeigt werden, dass weitere Modifikationen, wie zum Beispiel die Trimethylierung des Lysinrests 9 von Histon 3, zu einer Inaktivierung von mobilen genetischen Elementen führen können (Reichmann *et al.* 2012). Die Regulation durch den Wirt wird somit auf mehreren Ebenen realisiert und besteht aus einem engen Wechselspiel zwischen den mobilen genetischen Elementen und den Wirtseigenen Abwehrmechanismen.

1.6 Spezifisch integrierende mobile genetische Elemente

Da Integrationen in essentielle Gene oder in wichtige nicht-codierende regulatorische Bereiche der genomischen DNA zur Reduktion der Fitness der Wirtszelle führen können oder gar letal für die Wirtszellen sind, haben einige mobile Elemente zur Sicherung ihres eigenen „Überlebens“ eine Integrationspezifität entwickelt. Diese erlaubt es ihnen, ohne kritische genomische Regionen zu gefährden, mobilisiert zu werden. Bekannte Beispiele für mobile genetische Elemente mit Integrationspezifitäten sind die Ty-Elemente aus der Bäckerhefe *Saccharomyces cerevisiae* oder die R1- und R2-Elemente, welche in verschiedenen Insektenarten gefunden werden können (Eickbush und Eickbush 1995; Yieh *et al.* 2002).

Die Ty-Elemente sind LTR-Retrotransposons, welche innerhalb einer definierten Region oberhalb von RNA-Polymerase-III-transkribierten Genen integrieren (Ji *et al.* 1993; Boeke und Devine 1998). Dazu zählen alle tRNA-Gene sowie einige „*small nuclear RNA*“-Gene (Chalker und Sandmeyer 1992). Die Non-LTR-Elemente R1 und R2 integrieren dahingegen spezifisch in das 28S-rRNA-Gen in den Genomen von *Drosophila melanogaster* und *Bombyx mori* sowie weiteren Insektenarten (Burke *et al.* 1987). Darüber hinaus gibt es noch weitere spezifisch integrierende Retroviren und mobile genetische Elemente (Tabelle 2), wie die im Folgenden näher erläuterten TRE-Elemente aus dem Genom der sozialen Amöbe *Dictyostelium discoideum*.

Tab. 2: Beispiele für spezifisch integrierende mobile genetische Elemente.

Ziel	Name	Typ	Quelle
oberhalb von tRNA-Genen RNA-Pol-III-Transkriptionsstart	Ty1-4	LTR	(Yieh <i>et al.</i> 2002)
Heterochromatin der Telomere „ <i>silent mating loci</i> “	Ty5	LTR	(Brady <i>et al.</i> 2008)
RNA-Pol-II-Promotoren	Tf1	LTR	(Guo und Levin 2010)
oberhalb von tRNA-Genen RNA-Pol-III-Transkriptionseinheiten	TRE5	Non-LTR	(Marschalek <i>et al.</i> 1989)
unterhalb von tRNA-Genen RNA-Pol-III-Transkriptionseinheiten	TRE3	Non-LTR	(Winckler <i>et al.</i> 1998a)
28S-rRNA-Gen	R1/R2	Non-LTR	(Burke <i>et al.</i> 1987)
Telomere	TART/ HeT-A	Non-LTR	(Pardue und DeBaryshe 2003)
Promotoren aktiver Gene	MLV	Retrovirus	(Wu <i>et al.</i> 2003)

1.7 Mobile genetische Elemente im Genom von *Dictyostelium discoideum*

Das sehr kompakte Genom von *D. discoideum* beherbergt mit einem Anteil von rund 10 % an der gesamten genomischen DNA vergleichsweise viele mobile genetische Elemente. Es sind sowohl DNA-Transposons als auch autonome und nicht-autonomen Retrotransposons gefunden worden (Abbildung 12) (Glöckner *et al.* 2001).

Insbesondere die Gruppe der Non-LTR-Retrotransposons hat sich dabei offensichtlich an die hohe Gendichte angepasst, da sie nur in definierten Abständen zu tRNA-Genen in nicht-codierenden Bereichen gefunden werden können (Marschalek *et al.* 1989; Winckler *et al.* 2002). Innerhalb dieser Gruppe können zwei Familien unterschieden werden: Die TRE5-Elemente, welche oberhalb von tRNA-Genen integrieren, sowie die Gruppe der TRE3-Retrotransposons, die unterhalb von tRNA-Genen gefunden werden können (Winckler 1998b; Szafranski *et al.* 1999). DGLT-A, ein LTR-Retrotransposon, besitzt ebenfalls eine Integrationspezifität unterhalb von tRNA-Genen, allerdings integriert es in kürzeren Abständen als die TRE5-Elemente (Glöckner *et al.* 2001).

Die anderen Elemente, welche nicht mit tRNA-Genen assoziiert sind, bilden sehr häufig heterochromatische Cluster, was dafür spricht, dass sie entweder präferentiell ineinander integrieren, oder eine negative Selektion auf Integrationen außerhalb der Cluster besteht. Ungefähr 67 % aller nicht-spezifisch integrierenden mobilen genetischen Elemente werden von anderen Elementen flankiert (Winckler *et al.* 2005). Dahingegen werden 80 % der tRNA-Gen-assoziierten Elemente in euchromatischen Bereichen gefunden, obwohl sie nur 39 % aller mobilen genetischen Elemente im Genom von *D. discoideum* ausmachen (Winckler *et al.* 2011).

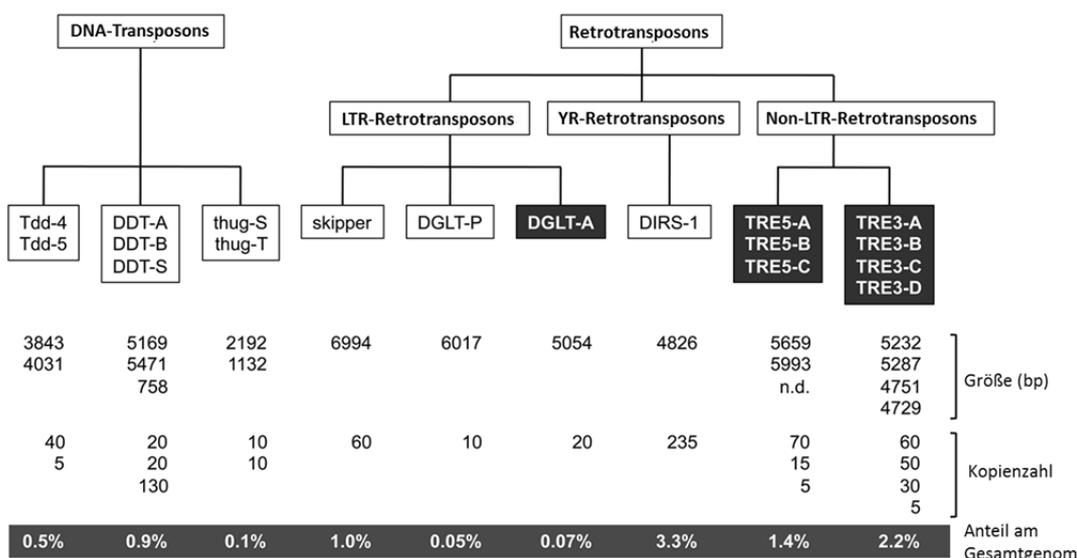


Abb. 12: Mobile genetische Elemente im Genom von *D. discoideum*.

Die tRNA-Gen-spezifisch integrierenden Retrotransposons sind in schwarzen Boxen hervorgehoben. Modifiziert nach Winckler *et al.* 2011.

1.7.1 Non-LTR-Retrotransposons

TRE5-Elemente

Die Gruppe der TRE5-Elemente zeichnet sich durch ihre obligate Integrationsspezifität unterhalb von tRNA-Genen aus. Es werden 3 Familien unterschieden: Die TRE5-A-, -B und -C-Elemente. Die Non-LTR-Retrotransposons der TRE5-A-Familie sind die am besten untersuchte Gruppe, welche auch als erste tRNA-Gen-assoziierte Elemente im Genom von *D. discoideum* entdeckt wurden (Marschalek *et al.* 1989; Winckler 1998b). Die TRE5-B- und TRE5-C-Elemente wurden erst im Zuge der kompletten Sequenzierung des Genoms detektiert (Glöckner *et al.* 2001) (Abbildung 13).

Ein vollständiges TRE5-A.1-Element ist 5657 bp groß und codiert für zwei ORFs, welche in unterschiedlichen, sich ungefähr 70 bp überlappenden; Leserahmen angeordnet sind (Marschalek *et al.* 1992). Interessanterweise fehlt dem zweiten ORF innerhalb der ersten 100 Aminosäuren ein Startcodon. Möglicherweise sorgt ein „ribosomaler Shift“ am Ende von ORF1 für eine korrekte Translationsinitiation von ORF2 (Marschalek *et al.* 1992; Winckler *et al.* 2002). Der erste ORF von TRE5-A codiert für ein 52 kDa großes Protein, welches, analog zu den ORF1-Proteinen des menschlichen L1-Elements, vermutlich an der Bildung der Ribonukleoprotein-Partikel beteiligt ist und eine *Cis*-Präferenz für die eigenen Transkripte besitzt (Hentschel *et al.* 2001). Allerdings konnten keine Homologien mit anderen Proteinen gefunden werden (Winckler *et al.* 2002). ORF2 codiert für ein 135 kDa großes Polyprotein, in dem Homologien mit Domänen anderer Non-LTR-Retrotransposons gefunden wurden. Die Übereinstimmungen zeigen, dass ORF2 für eine reverse Transkriptase, eine apurinisch/apyrimidinische Endonuklease und eine Cystein-reiche Region (HC-Motiv) codiert (Marschalek *et al.* 1992). Charakteristisches Merkmal aller TRE5-Elemente ist der modulare Aufbau der flankierenden untranslatierten Regionen. TRE5-A.1 beginnt mit einem 200 bp großen A-Modul, welches eine interne RNA-Polymerase-II-Promotoraktivität besitzt (Marschalek *et al.* 1992; Schumann *et al.* 1994). Die genomischen Kopien von TRE5-A.1 können mehrere tandemartig wiederholte A-Module besitzen (Winckler *et al.* 2005). Im sich anschließenden B-Modul befindet sich der Translationsstart des ORF1-Proteins. Das B-Modul wiederholt sich am Ende des zweiten offenen Leserahmens. Das 3'-Ende des Elements wird durch das sehr A/T-reiche C-Modul definiert, welches mit einem homopolymeren A-Stretch die TRE5-A.1-Elemente terminiert (Marschalek *et al.* 1992). Das C-Modul besitzt ebenfalls eine interne Promotoraktivität, die eine Minus-Strang-RNA erzeugt (Schumann *et al.* 1994).

Die 2,6 kb großen TRE5-A.2-Retrotransposons sind nicht-autonome Elemente, bei denen der komplette zweite ORF deletiert ist. Des Weiteren findet sich zwischen dem B- und dem C-Modul eine charakteristische 72 bp große Deletion (Marschalek *et al.* 1992). Die TRE5-A.2-Elemente können offensichtlich sehr erfolgreich durch die ORF2-Proteine der TRE5-A.1-Elemente *in trans* mobilisiert werden, da bei einem Test zur Untersuchung von neuen Integrationen endogener Elemente je zur Hälfte *De-novo*-Integrationen von TRE5-A.1- und TRE5-A.2-Elemente nachgewiesen wurden (Beck *et al.* 2002). Das nicht-autonome TRE5-A.3-Retrotransposon, welches aus je einem A-, einem B- und

einem C-Modul besteht, ist wahrscheinlich durch eine Rekombination innerhalb der TRE5-Elemente entstanden (Winckler *et al.* 2002).

Analog zu den TRE5-A-Elementen besitzen auch die TRE5-B- und -C-Elemente ein 5' gelegenes A- und B-Modul. Dem 5,5 kb großen TRE5-B-Element fehlen allerdings sowohl C-Modul als auch das benachbarte B-Modul (Winckler *et al.* 2002). Die A-Module von TRE5-A und TRE5-B stimmen zu rund 80 % in ihrer Nukleotidsequenz überein, was für eine starke Konservierung spricht und die offensichtlich essentielle Funktion der A-Module zeigt (Winckler *et al.* 2005). Ein Vergleich der Aminosäuresequenzen von ORF2 zeigt rund 40 % Homologie zwischen den TRE5-A und TRE5-B Elementen. Die TRE5-C Elemente konnten bislang nicht vollständig rekonstruiert werden, besitzen aber ein A- und B-Modul sowie ein terminal gelegenes C-Modul (Winckler *et al.* 2002; Winckler *et al.* 2005). Phylogenetische Untersuchungen zeigten, dass das TRE5-C Element näher mit TRE5-A verwandt ist als mit TRE5-B (Glöckner *et al.* 2001). Alle bisher gefundenen genomischen Kopien der drei unterschiedlichen TRE5-Elemente sind in einem Abstand von ungefähr 50 bp oberhalb von tRNA-Genen lokalisiert, was dafür spricht, dass der Mechanismus für die spezifische Integration der TRE5-Elemente konserviert ist. Eine Gemeinsamkeit der TRE5-Retrotransposons ist auch, dass der Großteil der integrierten Kopien 5' deletiert sind und somit durch das Fehlen eines intakten A-Moduls sofort nach der Integration inaktiviert wurden (Winckler *et al.* 2002).

Bei einem Blick auf die jeweiligen genomischen Anteile fällt auf, dass die TRE5-A-Familie offensichtlich erfolgreicher mobilisiert wird als die beiden anderen TRE5-Familien. So können im Genom von *D. discoideum* ungefähr 70 Loci mit TRE5-A-Fragmenten nachgewiesen werden, wohingegen nur jeweils 15 bzw. 5 Integrationen von TRE5-B und TRE5-C gefunden wurden (Glöckner *et al.* 2001). Die Größe der Zielsequenzduplikationen schwankt bei den TRE-Elementen zwischen 12 bis 16 bp (Hofmann *et al.* 1991; Szafranski *et al.* 1999).

TRE3

Die TRE3-Familie besteht aus vier unterschiedlichen Elementen, die Größen zwischen 4,7-5,5 kb besitzen und alle nach dem gleichen Prinzip aufgebaut sind (Glöckner *et al.* 2001; Winckler *et al.* 2002). Flankiert von zwei untranslatierten Regionen codieren die Elemente für zwei Gene, die neben dem ORF1-Protein im zweiten Gen für eine Endonuklease, eine Reverse Transkriptase und ein HC-Motiv codieren. Angeordnet sind beide ORFs in unterschiedlichen Leserahmen auf dem gleichen DNA-Strang, welche um wenige Basenpaare überlappen (Szafranski *et al.* 1999). Die TRE3-Elemente integrieren oberhalb von tRNA-Genen, wobei sie in durchschnittlichen Abständen von ca. 100 bzw. 150 bp gefunden werden können (Szafranski *et al.* 1999). Der größere Abstand resultiert möglicherweise aus der Bindung des RNA-Polymerase-III-Transkriptionsfaktors TFIIC an einer ca. 40 bp unterhalb von tRNA-Genen gelegenen externen B-Box (Bukerberger *et al.* 1991), wodurch sich auch hier ein durchschnittlicher Abstand der TRE3-Elemente von rund 100 bp zur externen-B-Box ergibt (Szafranski *et al.* 1999). Viele der genomischen Kopien sind, analog zu den TRE5-Elementen, 5' verkürzt und damit nicht aktiv (Winckler *et al.* 1998b). Zielsequenzduplikationen konnten nicht bei

allen TRE3-A- und TRE3-B-Elementen gefunden werden, wohingegen alle TRE3-C- und TRE3-D-Elemente TSD besitzen (Winckler *et al.* 2002).

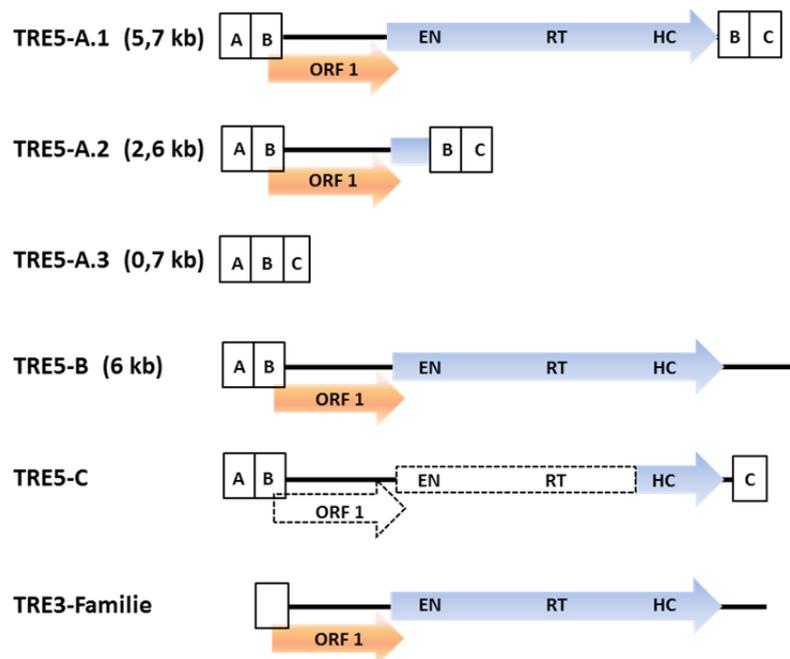


Abb. 13: Aufbau der verschiedenen Non-LTR-Retrotransposons aus *D. discoideum*.

EN: Endonuklease; RT: reverse Transkriptase; HC: Cystein-reiche Region. Modifiziert nach Winckler *et al.* 2005.

1.7.2 LTR-Retrotransposons

DGLT-A

Ein weiteres mit tRNA-Genen assoziiertes mobiles genetisches Element ist das 5 kb große LTR-Retrotransposon DGLT-A (engl.: „*Dictyostelium gypsy like Transposon Subtype A*“), das auf Basis von Sequenzvergleichen zur Familie der Ty3/Gypsy-Retrotransposons gezählt wird. DGLT-A hat 268 bp lange LTR-Sequenzen und besitzt einen ORF, der ein 1437 Aminosäuren großes Polyprotein codiert. Sowohl Protease-, Integrase-, Reverse Transkriptase- als auch RNase-H-Domänen sind darin nachweisbar (Glöckner *et al.* 2001; Winckler *et al.* 2002). DGLT-A-Elemente finden sich in Abständen von 10–30 bp unterhalb von tRNA-Genen (Marschalek *et al.* 1989; Hofmann *et al.* 1991). Es konnten nur einige wenige vollständige Elemente im Genom gefunden werden, die durch Mutationen offensichtlich inaktiviert wurden (Winckler *et al.* 2005). Des Weiteren lassen sich 15 Solo-LTRs im Genom von *D. discoideum* detektieren (Glöckner *et al.* 2001) (Abbildung 14).

Skipper

Skipper gehört zur Klasse der Ty3/Gypsy-ähnlichen LTR-Retrotransposons (Leng *et al.* 1998). Das 7 kb große Element kommt in ungefähr 50 kompletten Kopien und 10 Solo-LTRs im Genom von *D. discoideum* vor (Glöckner *et al.* 2001). Skipper besitzt zwei ORFs, die in unterschiedlichen Leserahmen angeordnet sind und für die zur Mobilisierung benötigten Proteine codieren. Das erste Gen codiert für das Gag-Protein und eine Proteasedomäne, welche durch ein Stoppcodon voneinander getrennt sind. Der zweite ORF codiert die für LTR-Retrotransposons typischen Reverse Transkriptase- und Integrasedomänen, welche als Polyprotein transkribiert werden (Leng *et al.* 1998). Im ORF2 ist C-terminal eine Chromodomäne detektierbar, die an der Erkennung von potentiellen Integrationsstellen beteiligt sein könnte (Malik und Eickbush 1999). Die LTRs von Skipper sind 390 bp lang und besitzen die für LTR-Retrotransposons typischen terminalen Dinukleotide 5'-TG.CA-3'. Die genomischen Kopien werden von 5 bp großen Zielsequenzduplikationen flankiert (Glöckner *et al.* 2002) (Abbildung 14).

DIRS-1

Das Retrotransposon DIRS-1 besitzt, im Gegensatz zu anderen LTR-Retrotransposons, nicht-identische, invertierte terminale Sequenzwiederholungen (ITR). Ein weiterer Unterschied besteht in der Verwendung einer Tyrosin-Rekombinase anstelle der sonst bei LTR-Retrotransposons typischen „DDE“-Integrasen. Das 4,8 kb große DIRS-1 Element codiert für drei Gene, die vom gleichen DNA-Strang in unterschiedlichen Leserahmen codiert werden. ORF1 entspricht vermutlich der für LTR-Retrotransposons typischen Gag-Domäne, wobei ORF2 und ORF3 die C-terminale Domäne von ORF1 überlappen. Für ORF2 konnte keine Funktion nachgewiesen werden, wohingegen der dritte ORF für eine Reverse Transkriptase und eine RNase-H codiert (Capello *et al.* 1985; Goodwin und Poulter 2001). Die 3'-ITR wird rechts von einer 27 bp langen Sequenz (RE) flankiert, sowie links von einer 88 bp langen internen komplementären Region (ICR). Die ersten 33 bp der ICR sind komplementär mit dem Start des 5'-ITRs, gefolgt von 55 bp die dem 3'-Ende des DIRS-1 Elements inklusive der RE-Region entsprechen (Cappello *et al.* 1985; Goodwin und Poulter 2001). Da diese Elemente auch in anderen DIRS-ähnlichen Retrotransposons gefunden wurden, wird vermutet, dass sie kritisch für die Replikation von DIRS-1-Elementen sind (Goodwin und Poulter 2001). Aufgrund dieser Merkmale wird ebenfalls angenommen, dass DIRS-Elemente durch die Bildung von doppelsträngigen, zirkulären Intermediaten einen anderen Retrotranspositionsmechanismus verwenden als LTR-Retrotransposons (Cappello *et al.* 1985; Duncan *et al.* 2002). Das Genom von *D. discoideum* enthält ungefähr 200 Fragmente des DIRS-1-Elements, welche als Cluster an den Enden der Chromosomen vorliegen. Daher wird vermutet, dass DIRS-1-Cluster an der Bildung der Centromerregionen der Chromosomen von *D. discoideum* beteiligt sind (Eichinger *et al.* 2005). Mit einem Anteil von über 3 % an der genomischen DNA ist es das häufigste mobile genetische Element im Genom von *D. discoideum* (Winckler *et al.* 2005) (Abbildung 14).

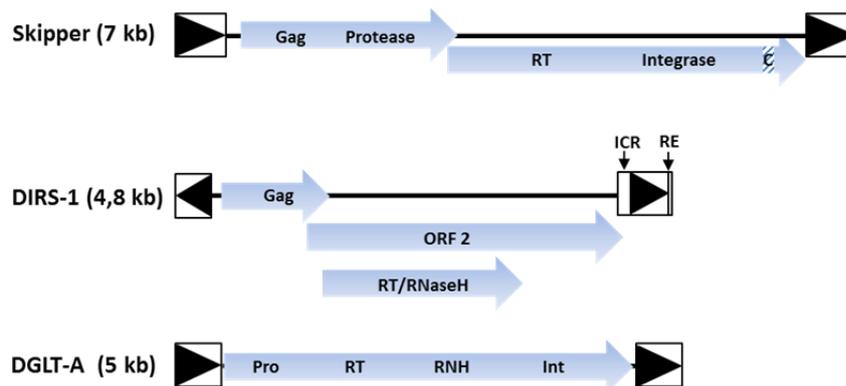


Abb. 14: Struktur der LTR-Retrotransposons von *D. discoideum*.

Pro: Protease; RT: Reverse Transkriptase; RNH: RNase-H; Int: Integrase; C: Chromodomäne; Gag: Gruppenspezifisches Antigen. Modifiziert nach Winckler *et al.* 2005

1.7.3 DNA-Transposons

Im Genom von *D. discoideum* wurden weitere mobile genetische Elemente identifiziert, die vermutlich zur Gruppe der DNA-Transposons gehören (Glöckner *et al.* 2001). Im Rahmen dieser Arbeit soll allerdings nicht näher darauf eingegangen werden.

1.8 RNA-Polymerase-III-transkribierte Gene

Alle bisher entdeckten spezifisch integrierenden Retrotransposons im Genom von *D. discoideum* konnten mit Genen in Verbindung gebracht werden, die von der RNA-Polymerase-III (Pol III) transkribiert werden. Essentiell für die Transkription der Pol-III-Gene ist die Bindung von Transkriptionsfaktoren, welche meist durch interne Promotorelemente rekrutiert werden. Auf dieser Grundlage lassen sich 3 Typen von Pol-III-transkribierten Genen unterscheiden (Schramm und Hernandez 2002):

TypI: 5S-rRNA-Gen, besitzt A- und C-Box als interne Promotorelemente

TypII: tRNA-Gene, welche sowohl interne A- als auch B-Boxen als Promotorelemente enthalten

TypIII: haben keine internen Promotorelemente, besitzen aber oberhalb gelegene Promotor- oder Enhancerelemente; charakteristisch ist eine TATA-Box

Zunächst bindet der Transkriptionsfaktor TFIIC an die A- und B-Box der tRNA-Gene, wodurch anschließend durch Protein-Protein Interaktionen der Transkriptionsinitiationsfaktor TFIIB oberhalb von TFIIC rekrutiert wird (Kassavetis *et al.* 1990). Dieser Komplex initiiert die Transkription durch die RNA-Polymerase-III (Abbildung 15). Die Assemblierung der Transkriptionsfaktoren vom Typ I beruht auf der Erkennung der A- und C-Box durch den Transkriptionsfaktor TFIIIA, welcher wiederum die Rekrutierung von TFIIC möglich macht. Anschließend rekrutiert der TFIIIA:TFIIC-

Komplex TFIIB, der die Transkription des ribosomalen 5S-Gens durch die Pol-III ermöglicht (Schramm und Hernandez 2002).

1.8.1 tRNA-Gene im Genom von *D. discoideum*

Im Genom von *D. discoideum* konnten bislang 418 tRNA-Gene identifiziert werden (Stand: Dezember 2014; Quelle: <http://www.dictybase.org>). Die Architektur der tRNA-Gene folgt dem allgemeinen Schema der Typ-II-transkribierten Pol-III-Gene mit internen A-Box- und B-Box-Promotorelementen. Eine Besonderheit sind die externen B-Box-Motive (exB-Box), die bei 66 % aller tRNA-Gene gefunden werden können. Die exB-Boxen liegen im Schnitt 41 bp unterhalb der internen B-Box (Abbildung 15) (Hofmann *et al.* 1991; Siol *et al.* 2006). Da TFIIC an B-Boxen bindet, wird vermutet, dass die Integration von TRE3-Elementen in größeren Abständen (~150 bp) zu den tRNA-Genen durch eine Interaktion mit TFIIC, welcher an die externen B-Boxen bindet, zustande kommt (Bukenberger *et al.* 1991; Bukenberger *et al.* 1994; Winckler *et al.* 1998a). Die B-Box-Motive der tRNA-Gene in *D. discoideum* sind stark konserviert und entsprechen den typischen eukaryotischen Motiven. In den tRNA-Genen von *D. discoideum* konnte für die internen B-Boxen das Konsensusmotiv GTCnnnnG⁵³TTCRAnYC abgeleitet werden (Siol *et al.* 2006). Den externen B-Boxen fehlt allerdings das 5'-gelegene GTC-Motiv der internen B-Boxen. Rund 54 % aller tRNA-Gene im Genom von *D. discoideum* sind mit Integrationen von TRE-Elementen assoziiert (Winckler *et al.* 2005).

1.8.2 Modelle zur Erklärung der spezifischen Integration von Retrotransposons an tRNA-Genen

Anhand von Experimenten mit LTR-Retrotransposons der Ty-Familie aus *S. cerevisiae* konnte gezeigt werden, dass Interaktionen der Pol-III-Transkriptionsfaktoren mit den Retrotransposons für eine spezifische Integration verantwortlich sind. Bei In-vitro-Versuchen mit Ty3, welches rund 700 bp oberhalb von Pol-III-Genen integriert, waren bereits die beiden TFIIB-Untereinheiten Brf1 und TBP1 ausreichend für eine erfolgreiche Integration der Retrotransposon-cDNA (Kirchner *et al.* 1995; Yieh *et al.* 2000). Ebenso ist bekannt, dass die spezifische Integration an den Transkriptionsstartpunkten von Pol III-Genen durch das LTR-Retrotransposon Ty3 auf Grund von Interaktionen des Präintegrationskomplexes mit den TFIIB-Untereinheiten TBP und Brf1 zustande kommt (Yieh *et al.* 2002). Für das *D. discoideum* TRE5-A.1-Element konnte in Bakterien-Zwei-Hybrid Versuchen gezeigt werden, dass das ORF1-Protein mit Untereinheiten von TFIIB interagiert, besonders die TBP-Untereinheit scheint dabei kritisch zu sein (Abbildung 15) (Chung *et al.* 2007). Interaktionen konnten auch zwischen dem ORF1-Protein von TRE5-B und TFIIB-Untereinheiten nachgewiesen werden (Chung *et al.* 2007). In weiteren Versuchen konnten Integrationen endogener Retrotransposons aus *D. discoideum* an isolierten B-Boxen und dem 5S-rRNA-Gen gefunden werden, was ebenfalls ein Hinweis auf Interaktionen mit den zuvor beschriebenen Transkriptionsfaktoren ist. Entscheidend scheint dabei die Interaktion mit TFIIB, da das 5S-rRNA-Gen keine B-Box besitzt und diese Beobachtung dafür spricht, dass keine Sequenzspezifische Erkennung einer B-Box durch den Prä-

Integrationskomplex von TRE5-A stattfindet (Siol *et al.* 2006). Aus diesen Ergebnissen lässt sich ein Modell für die spezifische Integration von Retrotransposons an tRNA-Genen oder anderen Pol-III-transkribierten Genen ableiten (Abbildung 15), welches nahelegt, dass eine Protein-Protein Interaktion zwischen dem ORF1-Protein des Retrotransposons und dem TFIIB-Transkriptionsfaktor die Integrationsstelle definiert (Siol *et al.* 2006; Chung *et al.* 2007).

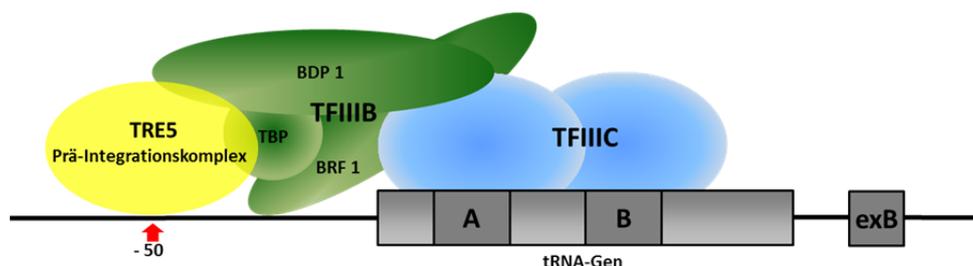


Abb. 15: Modell der spezifischen Integration von TRE5-Elementen an tRNA-Genen.

Durch Interaktion des Prä-Integrationskomplexes von TRE5 mit TFIIB-Untereinheiten kommt es zur Integration der TRE5-Elemente rund 50 bp oberhalb des tRNA-Gens. Modifiziert nach Siol *et al.* 2006.

1.9 Genetisch markierte artifizielle TRE5-Retrotransposons

Nachdem gezeigt wurde, dass das Genom von *D. discoideum* eine aktive Population von TRE5-Elementen beherbergt, wurde versucht, durch die Generierung von artifiziellen TRE5-Elementen die Strukturen zu definieren, welche für eine erfolgreiche Integration an tRNA-Genen essentiell sind. Basierend auf der Beobachtung, dass bei In-vivo-Versuchen zu einem erheblichen Anteil De-novo-Integrationen von nicht-autonomen TRE5-A.2-Elementen nachgewiesen wurden (Siol *et al.* 2006), konnte ein artifizielles, nicht-autonomes TRE5-Element erzeugt werden, welches in Zellkulturen aktiv ist (Siol *et al.* 2011). Das Element beinhaltet neben einem Promotor, der die Transkription des kompletten Elements ermöglicht, ebenfalls das ORF1-Gen von TRE5-A.1. An dieses Gen wurde eine Blasticidin-Kassette in umgekehrter Orientierung zum Element fusioniert, welche durch ein Intron, dass in gleicher Orientierung wie das Element liegt, unterbrochen wird. Die Kassette wird durch einen ebenfalls revers gelegenen Aktin-15-Promotor transkribiert. Terminiert wird das artifizielle TRE5-Element durch ein C-Modul (Abbildung 16) (Siol *et al.* 2011).

In Zellkulturen wird dieses Element *in trans* durch ORF2-Proteine endogener Elemente mobilisiert. Dabei wird aus der kompletten RNA des Elements während der Transkription das Intron herausgespleißt. Nach erfolgreicher reverser Transkription und Integration des nicht-autonomen Elementes entstehen so Kopien an neuen Integrationsorten, die durch das Fehlen des Introns eine funktionelle Blasticidin-Kassette besitzen und so die Zellen gegen das Antibiotikum resistent machen. Auf diese Weise kann auf Zellen selektioniert werden, die De-novo-Integrationen des artifiziellen Elements im Genom aufweisen. So konnte gezeigt werden, dass die künstlichen Retrotransposons die gleichen Integrationsorte präferieren wie die endogene TRE5-Population und auch die Abstände zu tRNA-Genen identisch sind (Siol *et al.* 2011). Zudem wurden Integrationen nachgewiesen, die sich an

isoliert gelegenen B-Boxen auf dem rDNA-tragenden Palindrom befanden. Des Weiteren konnte mit diesem System gezeigt werden, dass das A-Modul nicht essentiell für die Mobilisierung des Retrotransposons ist, da Elemente mit artifiziellen Promotoren, wie dem Aktin-6-Promotor, sogar noch aktiver waren als Elemente mit einem authentischen A-Modul. Das C-Modul scheint allerdings unentbehrlich für eine erfolgreiche Retrotransposition zu sein, denn Elemente ohne C-Modul waren inaktiv (Siol *et al.* 2011). Strukturanalysen des C-Moduls legen nahe, dass eine RNA-Sekundärstruktur wichtig für die Aktivität der TRE5-A-Elemente ist. Durch Mutationen veränderte C-Module, bei denen die vorhergesagte RNA-Sekundärstruktur zerstört wurde, zeigten dementsprechend eine stark reduzierte Aktivität oder waren komplett inaktiviert (Siol *et al.* 2011).

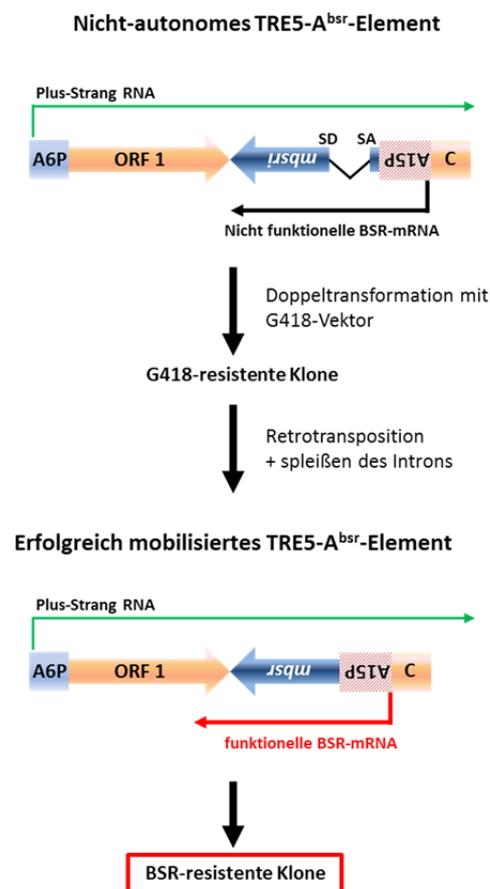


Abb. 16: Schema des Retrotranspositions-Assays.

Nicht-autonome, artifizielle TRE5-Elemente vermitteln nach erfolgreicher Retrotransposition Blasticidin-Resistenz. SD: Spleißdonor; SA: Spleißakzeptor; BSR: Blasticidin; A15P: Aktin-15-Promotor; A6P: Aktin-6-Promotor; C: C-Modul. Modifiziert nach Siol *et al.* 2011.

1.10 Zielsetzung der Arbeit

Mit der Sequenzierung der Genome verschiedener Arten der sozialen Amöben ergibt sich eine Möglichkeit, die evolutionäre Entwicklung der spezifisch integrierenden Retrotransposons näher zu analysieren. Dazu sollten in der vorliegenden Arbeit zunächst Genomanalysen durchgeführt werden, um einen Großteil der mobilen genetischen Elemente der neu sequenzierten Arten *D. fasciculatum*, *P. pallidum* und *D. purpureum* zu detektieren, diese zu annotieren und zu charakterisieren. Auf Basis dieser Daten sollten phylogenetische Zusammenhänge zwischen der Entwicklung der tRNA-Gen-abhängigen Integration mobiler genetischer Elemente und der jeweiligen Genomarchitektur aufgefunden werden. In Vorarbeiten wurde ein In-vivo-Retrotranspositionstest etabliert, mit dem das natürliche Integrationsverhalten von TRE5-A-Retrotransposons im Genom von *D. discoideum* untersucht werden kann. Als Erweiterung dieser Arbeiten sollen weitere Retrotransposons genetisch markiert und ihre authentische Integrationspräferenz charakterisiert werden.

Der genannte Retrotranspositions-Assay bietet die Möglichkeit, Zellen mit De-novo-Integrationen zu isolieren und anschließend die Integrationsstellen zu analysieren. Bislang beruhte die Sequenzierung von Integrationsstellen auf einer PCR-basierten Methode, die ausschließlich an tRNA-Genen integrierte Elemente nachweisen konnte. Daher war es nicht möglich Integrationen aufzuspüren, welche nicht die bekannten Integrationsorte umfassen („off target“-Integrationen). Deshalb soll in dieser Arbeit der Retrotranspositionstest erweitert werden, um genomweit und unabhängig von potentiellen Integrationsorten eine große Anzahl neuer Integrationen zu sequenzieren und zu analysieren. Auf Grund dieser Daten sollen Rückschlüsse auf die exakte Spezifität der TRE5-Elemente gezogen werden.

Es wurde in Vorarbeiten gezeigt, dass sich auf dem extrachromosomalen, palindromischen DNA-Element, das in *D. discoideum* die ribosomalen RNA-Gene trägt, isolierte B-Box-Motive befinden, die als Integrationsorte für artifizielle TRE5-Elemente in Frage kommen. Da bislang wenig Daten über Integrationen endogener TRE5-Elemente auf dem Palindrom verfügbar sind, soll im Rahmen dieser Arbeit systematisch untersucht werden, ob die B-Boxen als potentielle Integrationsorte auf dem Palindrom von TRE5-A erkannt und die Elemente dort integriert werden. In diesem Zusammenhang sollten die B-Boxen auf mögliche transkriptionelle Aktivitäten überprüft und charakterisiert werden, um eine Verbindung zur zuvor erläuterten Hypothese herzustellen, dass Transkriptionsfaktoren des RNA-Polymerase-III-Komplexes zur spezifischen Integration von TRE5-Retrotransposons benötigt werden.

2 Material und Methoden

2.1 Materialien

2.1.1 Laborgeräte

Name	Typ	Hersteller
Blockthermostate	Combi-Therm CH100	Peqlab
	9i-D1100	Peqlab
Elektroporator	Gene Pulser Xcell	Bio-Rad
Gel-Elektrophorese	Perfect blue Gelsystem Mini	Peqlab
Geldokumentation	Gel Jet Imager 2004	Intas
Lichtmikroskope	Binokular-Mikroskop SZX9	Olympus
	Mirkoskop CXX 41	Olympus
Magnetisches Rack	DynaMag™-2	Invitrogen
pH-Wert Messgerät	pH 1000L	VWR
Real-Time-PCR System	Mx3000P	Stratagene
Schüttelinkubatoren	Multitron 2 HAT-Infors	Infors
Schüttelwasserbad	SW22	Julabo
Schüttler	KS250 basic	IKA Labortechnik
Sterilwerkbank	MRF-B	Steag
Thermocycler	Primus 96 advanced	Peqlab
UV/VIS-Spektrometer	BioPhotometer	Eppendorf
Vakuumentrifuge	Eppendorf	Concentrator 5301
Zellzählsystem	Model DT	Schärfe System GmbH
Zentrifugen	Centrifuge 5415 D/5424	Eppendorf
	Centrifuge 5417 R	Eppendorf
	Centrifuge 5702	Eppendorf
	Centrifuge 5810 R	Eppendorf
	Rotina 420 R	Hettich
	Avanti J-E	Beckmann Coulter

2.1.2 Verbrauchsmaterialien

Name	Lieferant
CL-Xposure® Film	Thermo Scientific
Elektroporationsküvetten (1 mm, 2 mm, 4 mm)	PeqLab
Film-Entwickler GBX	Kodak
Film-Fixierer GBX	Kodak
Filterpapier 3 MM	Whatman
DNA-LoBind 1,5 ml Reaktionsgefäße	Eppendorf
Magnetische Streptavidin-Beads M-280	Invitrogen
Mikrotiterplatten (24 Wells)	Greiner
Nitrocellulose-Membran Hybond™-N ⁺	Amersham
Petrischalen	Greiner
Pipettenspitzen, unsteril	Greiner
Rotilabo® Spritzenfilter steril 0,22 µm	Roth
Vakuum-Filter-System	Renner GmbH

2.1.3 Verwendete Kits

Name	Hersteller/Lieferant
5'-DNA Adenylation Kit	New England Biolabs
Agarose Gel Extraction Kit	Jena Bioscience
Biotin PCR Labeling Core Kit	Jena Bioscience
ECL® Advance Western Blotting Detection-Kit	GE Healthcare
Gateway® LR Clonase® II Enzyme mix	Invitrogen
High Fidelity Polymerase Kit	Jena Bioscience
Omniscript Reverse Transcription Kit	Qiagen
PCR Extender System	5 Prime
pCR®8/GW/TOPO® TA Cloning Kit	Invitrogen
pGEM-T® Vector System	Promega
TOPO® XL PCR Cloning Kit	Invitrogen
QIAquick® Gel Extraction Kit	Qiagen
QIAquick® PCR Purification Kit	Qiagen
RNase-Free DNase-Set	Invitrogen
RNeasy Mini Kit	Qiagen

2.1.4 Enzyme

Name	Hersteller/Lieferant
Alkalische Phosphatase (CIP)	New England Biolabs
FastAP Alkalische Phosphatase	Fermentas/Thermo Scientific
Proteinase-K	Fermentas/Thermo Scientific
Restriktionsendonukleasen	NEB
RiboLock RNase Inhibitor	Thermo Scientific
Ribonuklease-A	New England Biolabs
T4-DNA-Ligase	Fermentas/Thermo Scientific
T4-Polynukleotidkinase	New England Biolabs
T4-RNA-Ligase	New England Biolabs
5'-App-DNA/RNA-Ligase	New England Biolabs
Streptavidin-HRP	Invitrogen

2.1.5 DNA-Größenstandards

Name	Hersteller/Lieferant
Biotinylated 2-Log DNA-Ladder	New Englands Biolabs
pUC19 BsiSI (HpaII)	Jena Bioscience
Lambda DNA/PstI	Jena Bioscience
Low-Range DNA-Ladder	Jena Bioscience
Mid-Range DNA-Ladder	Jena Bioscience

2.1.6 Anitibiotika

Name	Hersteller/Lieferant
Ampicillin	Roth
Blasticidin S	Invitrogen
Geneticin Disufat (G418)	Calbiochem/Invitrogen
Hygromicin	Roth
Kanamycin	Serva
Spectinomycin	Sigma

2.1.7 Chemikalien

Alle Chemikalien entsprechen dem Reinheitsgrad „Pro Analysis“.

Name	Lieferant
2-(N-morpholino)ethanesulfonic Säure (MES)	Serva
2-(4-(2-Hydroxyethyl)-1-piperazinyl)-ethansulfonsäure (HEPES)	Roth
5-Brom-4-chlor-3-indolyl- α -D-galactopyranosid (X- α -Gal)	Roth
50×Denhardt's-Lösung	Invitrogen
Agar-Agar	Roth
Agarose (peqGold Universal-Agarose)	PeqLab
Bovine Serum Albumin (BSA) Fraktion V	Roth
Chloroform/Isoamylalkohol (24:1)	Roth
D-Glucose-Monohydrat	Roth
Desoxyribonukleosidtriphosphate (dNTPs)	Roth
Ethylendiamintetraessigsäure (EDTA)	Roth
Ethidiumbromid	Roth
EvaGreen TM Fluorescent DNA Stain	Jena Bioscience
Formamid	Roth
Glycerin 86 %	Roth
Hefeextrakt	Serva
HL5-Medium	ForMedium
Isopropyl- β -D-thiogalactopyranosid (IPTG)	Roth
Milchpulver	Roth
Natriumhydroxid	Roth
Natriumlaurylsulfat (SDS)	Roth
Pepton	Roth
Polyethylenglykol (PEG) 8000	New England Biolabs
Phenol/Chloroform/Isoamylalkohol (25:24:1)	Roth
ROX Reference Dye	Jena Bioscience
Tris-Puffer	Roth
Triton X-100	Sigma
Trypton	Serva

2.1.8 Medien

Alle Angaben beziehen sich auf die Herstellung von 1 l Medium. Alle Medien wurden autoklaviert. Für Agar-Platten wurden dem Medium vor dem autoklavieren 18 g/l Agar zugegeben.

HL5-Medium:	22 g HL5-Fertigmedium (Hersteller: ForMedium); nach dem Autoklavieren 25 ml 40 % Glucose zugeben
LB-Medium:	10 g Trypton; 5 g Hefeextrakt; 10 g NaCl
MES-Medium:	10 g Bacton-Pepton; 5 g Hefeextrakt; 1,3 g MES; mit 10 M NaOH auf pH 7,1 einstellen; nach dem Autoklavieren 25 ml 40 % Glucose zugeben
Na-Medium:	0,5 g Bacto-Pepton; 0,5g Glucose; 1×Phosphatpuffer ad. 1l; auf pH 6,2 einstellen
Phosphat-Agar:	20 ml 50×Phosphatpuffer; auf pH von 6,5 einstellen
SOC-Medium:	5 g Trypton; 5 g Hefeextrakt; 1 g Glucose

2.1.9 Puffer und Lösungen

2× HBS (250 ml)	4 g NaCl; 0,18 g KCl; 0,05 g Na ₂ HPO ₄ ; 2,5 g HEPES; 0,5 g Glucose; auf pH 7 einstellen und steril filtrieren
20×SSC	3 M NaCl; 0,3 M Na-Citrat
50× Phosphatpuffer (1 l)	99,86 g KH ₂ PO ₄ ; 17,80 g Na ₂ HPO ₄ ; auf pH 5,7 einstellen und autoklavieren
Denaturierlösung	0,5 M NaOH; 1,5 M NaCl
Elektroporationspuffer (für <i>D. discoideum</i>)	10 mM KH ₂ PO ₄ ; 50 mM Saccharose; auf pH-Wert 6,1 einstellen und autoklavieren
Hybridisierlösung (50 ml)	25 ml Formamid; 12,5 ml 20×SSC; 5 ml 10 % SDS; 1 ml 50 × Dennhardts Lösung; 6,5 ml MQ-Wasser; 5 µg/ml Heringsperma-DNA
10×Low-TE Puffer:	100 mM Tris-HCl; 1 mM EDTA; auf pH 8 einstellen
Lysepuffer (gDNA)	0,32 M Saccharose; 10 mM Tris-HCl, pH 7,5; 5mM MgCl ₂ ; 1 % Triton X-100
Pufferlösung A (gDNA)	10 mM Tris-HCl; pH 7,5; 10 mM EDTA
Pufferlösung B (gDNA)	10 mM Tris-HCl; pH 7,5; 0,7 % SDS
Neutralisierlösung:	0,5 M Tris-HCl, pH 7,5; 1,5 M NaCl; 1 mM EDTA

Prä-Hybridisierungslsg. (50 ml)	25 ml Formamid; 12,5 ml 20×SSC; 5 ml 10 % SDS; 5,5 ml MQ-Wasser; 2 ml 50× Denhardt's Lösung; 50 µg/ml Heringssperma-DNA (vor Zugabe aufkochen)
Schocklösung (2ml)	1 ml 2×HBS; 0,4 ml MQ-Wasser; 0,6 ml 60 % (m/v) Glucose
Southern-Basispuffer (SB-Puffer)	0,1 M Tris-HCl, pH 7,5; 0,15 M NaCl

2.1.10 Mikroorganismen

Soziale Amöben

D. discoideum; Stamm AX2; axenisch wachsender Laborstamm

D. discoideum; RrpAC⁻ Mutante (Wiegand *et al.* 2013)

P. pallidum; Wildtyp-Stamm

D. fasciculatum; Wildtyp-Stamm

E. coli K12-Stämme

DH5α: F⁻ Φ80lacZΔM15 Δ(lacZYA-argF) U169 *recA1 endA1 hsdR17* (rk⁻, mk⁺) *phoA supE44 λ- thi-1 gyrA96 relA1*

Top 10: F⁻ *mcrA* Δ(*mrr-hsdRMS-mcrBC*) φ80lacZΔM15 Δ*lacX74 recA1 araD139* Δ(*ara-leu*) 7697 *galU galK rpsL* (Str^R) *endA1 nupG λ-*

JM109: *endA1, recA1, gyrA96, thi, hsdR17* (rk⁻, mk⁺), *relA1, supE44, λ-,* Δ(*lac-proAB*), [F', *traD36, proAB, lacIqZΔM15*], IDE3

SURE2: e14-(McrA-) Δ(*mcrCB-hsdSMR-mrr*)171 *endA1 gyrA96 thi-1 supE44 relA lac recB recJ sbcC umuC::Tn5* (Kanr) *uvrC* [F' *proAB lacIqZΔM15 Tn10* (Tetr) Amy Camr].

Für die Kultivierung von Wildtyp-Stämmen wurde zudem *Klebsiella planticola* verwendet (Bezugsquelle: *Dictyostelium*-Stock-Center).

2.2 Verwendete Programme

Name/Version	Beschreibung	Referenz/URL
Argo GenomeViewer	Genombrowser	http://www.broadinstitute.org/annotation/argo
Artemis	Genombrowser und Visualisierung	(Rutherford <i>et al.</i> 2000)
Blastall	Generierung und Nutzung lokaler Blast-Datenbanken	(Altschul <i>et al.</i> 1990)
ClustalX 2.0.12	Erzeugung von Alignments	(Larkin <i>et al.</i> 2007)
DNA-Plotter	Erstellung zirkulärer genomweiter Abbildungen	(Carver <i>et al.</i> 2009)
Emboss 1.5	Software-Sammlung, Bearbeitung von Sequenzdaten	(Rice <i>et al.</i> 2000)
GCUA	Analyse des Spezies-spezifischen Codon-Gebrauchs	(Fuhrmann <i>et al.</i> 2004)
LTRfinder	Automatische Detektion von LTR-Sequenzen	(Zhao und Wang 2007)
MEGA5	Alignment-Programm, Erzeugung von Stammbäumen	(Tamura <i>et al.</i> 2011)
Multalign	Alignment-Programm	http://multalin.toulouse.inra.fr/multalin/
pDRAW 1.1.121	Erzeugung von Vektorkarten, Bearbeitung von Sequenzen	http://www.acaclone.com/
Phylip	Phylogenie-Programm	http://evolution.genetics.washington.edu/phylip.html
Puzzle	Phylogenie-Programm	(Schmidt <i>et al.</i> 2002)
staden package	Assemblierung von Sequenzen, Generierung von Konsensussequenzen	http://staden.sourceforge.net/
Tandem Repeat Finder 4.07b	Detektion von repetitiven Sequenzen	(Benson 1999)
tRNAscan-SE	Automatische Detektion von tRNA-Genen	(Schattner <i>et al.</i> 2005)

2.3 Datenbanken

Name	Beschreibung	URL
dictybase	Genomdatenbanken für <i>D. discoideum</i> ; <i>D. fasciculatum</i> ; <i>P. pallidum</i> ; <i>D. purpureum</i>	http://www.dictybase.org
InterProScan	Proteindatenbank	http://www.ebi.ac.uk/interpro/
NCBI	Allgemeine Nukleotid- und Proteindatenbanken	http://blast.ncbi.nlm.nih.gov/Blast.cgi
Prosite	Proteindatenbank	http://prosite.expasy.org/
SACGB	Genombrowser für soziale Amöben	http://sacgb.fli-leibniz.de/cgi/index.pl

2.4 Methoden

2.4.1 Molekularbiologische Methoden

2.4.1.2 Isolation genomischer DNA

Für die Isolation von genomischer DNA, welche für PCR-Ansätze geeignet ist, wurden Pellets mit 2×10^7 Zellen verwendet. Die Pellets wurden in folgendem Puffer resuspendiert:

336 μ l ddH₂O
40 μ l 10x Taq-Polymerase-Puffer
20 μ l Triton X-100 (10 %)
4 μ l Proteinase K (10 mg/ml)

Es folgte eine Inkubation bei 56 °C für 60 min und eine anschließende Hitzeinaktivierung der Proteinase-K für 10 min bei 95 °C. Die Suspension wurde nun für 1 min bei 30.000 g zentrifugiert und der Überstand in ein neues 1,5 ml Reaktionsgefäß überführt. Der Verdau der RNA erfolgte durch Zugabe von 2 μ l RNase A-Lösung (1 mg/ml) und einer weiteren Inkubation für 30 min bei 37 °C. Es wurden 2–10 μ l der Lösung je PCR-Ansatz verwendet.

2.4.1.3 Extraktion verdaubarer genomischer DNA

Zur Isolation verdaubarer genomischer DNA wurden zunächst 5×10^8 Zellen bei 4 °C und 150 g in einem 50 ml Falcon-Röhrchen zentrifugiert. Das Pellet wurde anschließend in 50 ml kalten Lysepuffer resuspendiert. Nach Zugabe von 1 ml 10 % Triton-X und einer 10-minütigen Inkubation auf Eis folgte eine Zentrifugation bei 14.000 g und 4 °C für 10 Minuten. Der Überstand wurde verworfen und das Zellkernpellet in 5 ml Pufferlösung A (10 mM Tris-HCl, pH-Wert 7,5; 10 mM EDTA) aufgenommen. Nach Zugabe von 5,5 ml Pufferlösung B (10 mM Tris-HCl, pH-Wert 7,5; 0,7 % SDS) und 30 μ l RNaseA (1mg/ml) folgte eine Inkubation bei 37 °C für 30 Minuten. Anschließend wurden 40 μ l Proteinase K (20 mg/ml) zugegeben und die Lösung für eine weitere Stunde bei 55 °C inkubiert. Nun wurden 10 ml Phenol/Chloroform/Isoamylalkohol (Verhältnis: 25:24:1) zugegeben und vermischt. Nach einer Zentrifugation für 10 min bei 4 °C und 14.000 g wurden 8 ml der oberen wässrigen Phase vorsichtig abgenommen und in ein neues 50 ml Falcon-Röhrchen gegeben. Anschließend wurden 8 ml Chloroform/Isoamylalkohol (Verhältnis 24:1) zugegeben und die Lösung gemischt. Nach einer Zentrifugation für 10 min bei 4 °C und 14.000 g wurden wiederum 5 ml der oberen Phase abgenommen und in ein neues 50 ml Falcon überführt. Nun wurden 12,5 ml 100 % Ethanol und 500 μ l 5M NaCl zugegeben und die Lösung über Nacht bei -20 °C inkubiert um die genomische DNA aus der Lösung zu fällen. Es folgte eine Zentrifugation für 30 min bei 4 °C und 14.000 g. Das dadurch erhaltene DNA-Pellet wurde mit 10 ml 70 % Ethanol gewaschen und anschließend in 400 μ l MQ-Wasser aufgenommen. Die Lösung wurde nun in ein 1,5 ml Reaktionsgefäß überführt. Nach der Zugabe von 1 ml 100 % Ethanol und 40 μ l 5M NaCl folgte eine weitere Fällung bei -20 °C für 30 Minuten. Es folgte eine Zentrifugation für 20 min bei 4 °C und 30.000 g. Das DNA-Pellet wurde drei Mal mit 1 ml 70 % Ethanol gewaschen und anschließend bei 37 °C vorsichtig getrocknet. Zum

Schluss wurde die DNA in 400 µl Low-TE Puffer aufgenommen und die Konzentration durch eine photometrische Messung bei 260 nm ermittelt.

2.4.1.4 Southern-Blot mit biotinylierten Sonden

(A) Erzeugung biotinylierter Sonden für das Blastocidin-Gen

Verwendet wurde das „Biotin PCR Labeling Core Kit“ von Jena Bioscience nach Herstellerangaben. Die PCR-Reaktionen wurden in 5 parallelen Ansätzen mit den Primern BSR-TEST-01 und BSR-TEST-02 durchgeführt. Als Matrize diente ein Plasmid, welches eine Blastocidin-Kassette enthält. Die PCR-Produkte wurden vereinigt, über eine Säule gereinigt und anschließend in 150 µl MQ-Wasser eluiert.

(B) Restriktionsverdau genomischer DNA

Dazu wurden 15 µg verdaubarer genomischer DNA mit 10 µl des entsprechenden 10-fachen Enzym-Puffer und jeweils 20 Units des Restriktionsenzym vermischt und die Lösung auf 100 µl mit MQ-Wasser aufgefüllt. Nach einer Inkubation über Nacht bei 37 °C folgte eine Ethanol-Fällung. Das dadurch erhaltene DNA-Pellet wurde in 20 µl Low-TE Puffer resuspendiert.

(C) Blotten der verdauten DNA

Die verdaute genomische DNA wurde für 3 h bei 100 V in einem Agarosegel mit einer Konzentration von 1,2 % elektrophoretisch aufgetrennt. Als Größenstandard diente der „Biotinylated 2-Log DNA Ladder“. Das Gel wurde anschließend für mehrere Stunden in 200 ml frischer Denaturierlösung geschwenkt um einzelsträngige DNA zu erhalten. Nach einer Inkubation für 15 min in 200 ml Neutralisierlösung wurde die DNA über Nacht mittels Kapillar-Blot und 20×SSC auf eine Hybond™-N⁺ Nitrocellulose-Membran übertragen. Am folgenden Tag wurde die DNA durch eine Inkubation bei 80 °C für 2 Stunden an der Membran fixiert.

(D) Hybridisierung der biotinylierten Sonde

Die Membran wurde für mehrere Stunden in 50 ml Prä-Hybridisierlösung bei 42 °C geschwenkt. Anschließend wurde die Prä-Hybridisierlösung durch 50 ml Hybridisierlösung ersetzt, in welche die gereinigte Sonde gegeben wurde. Die Hybridisierlösung wurde vor der Verwendung für 5 min auf 100 °C erhitzt und anschließend auf rund 70 °C abgekühlt. Die Hybridisierung der Sonde erfolgte über Nacht bei 42 °C im Schüttel-Wasserbad.

(E) Waschen und Blocken der Membran

Nach der Hybridisierung wurde die Membran zunächst zweimal für 3 min in 2×SSC/0,1 % (v/v) SDS bei Raumtemperatur gewaschen. Es folgte ein zweiter Waschschrift für 3 min mit 0,25×SSC/0,1 % (v/v) SDS, der ebenfalls wiederholt wurde. Nun wurde die Membran zweimal für 15 min in 0,16×SSC/0,1 % (v/v) SDS bei 50 °C gewaschen und anschließend kurz in 2×SSC geschwenkt.

Zum Blocken unspezifischer Bindungen des Antikörpers wurde die Membran zunächst kurz mit SB-Puffer gewaschen und anschließend für 1 h bei RT in SB-Puffer mit 5 % (m/v) Milchpulver inkubiert. Die Membran wurde nun dreimal mit SB-Puffer gewaschen. Es folgte eine Inkubation für 1 h bei 65 °C in SB-Puffer mit 3 % (m/v) BSA. Anschließend wurde die Membran für 1 h bei RT in 50 ml SB-Puffer mit 10 µl Streptavidin-HRP inkubiert. Zum Schluss wurde die Membran drei Mal für 15 min mit SB-Puffer gewaschen.

(F) Detektion der Sonde

Die Detektion erfolgte mit dem „ECL[®] Advance Western-Blotting Detection-Kit“ nach Herstellerangaben und unterschiedlichen Belichtungszeiten.

2.4.1.5 Inverse PCR

Für die inverse PCR wurden zunächst 25 µg genomische DNA mit 40 Units des Restriktionsenzym Sau3AI für 4 Stunden bei 37 °C verdaut und anschließend für 20 min bei 65 °C Hitzeinaktiviert. Das Volumen wurde daraufhin auf 600 µl mit MQ-Wasser erhöht und die DNA mit einer Phenol-/Chloroform-Fällung gereinigt. Das DNA-Pellet wurde in 50 µl MQ-Wasser aufgenommen. Die Ligation erfolgte für 2 bis 3 Tage bei 22°C. Dazu wurden mehrere Ansätze mit unterschiedlichen Mengen der verdauten DNA in je 100 µl Gesamtvolumen unter Verwendung der T4-DNA-Ligase (Fermentas) nach Herstellerangaben ligiert. Die eingesetzten Mengen verdauter genomischer DNA variierten dabei von 0,1 ng/µl bis 1 ng/µl. Anschließend wurden alle Ligationsreaktionen vereinigt und erneut durch eine Phenol-/Chloroform-Extraktion gereinigt. Das Pellet wurde in 10 µl MQ-Wasser aufgenommen. Die ligierte DNA konnte nun für die Retrotransposon-spezifischen PCRs verwendet werden. Für die erste PCR wurden die Primer iPCR_ORF1opt-01 und -02 verwendet. Für eine folgende *Nested*-PCR wurden die Primer Rep-235 und iPCR-ORF1-opt-02 eingesetzt.

2.4.1.6 Extraktion von RNA und cDNA-Synthese

Für die Isolation von Gesamt-RNA aus *D. discoideum*-Zellen wurden 2×10^7 Zellen eingesetzt. Verwendet wurde dafür das *RNeasy Mini-Kit* (Qiagen) nach dem Protokoll für die Extraktion von RNA aus Zellkulturen. Zusätzlich erfolgte ein Verdau der genomischen DNA mit dem *RNase-Free DNase-Set* (Qiagen) nach Herstellerangaben. Anschließend wurde die RNA-Konzentration durch photometrische Messung bei 260 nm bestimmt.

Die cDNA-Synthese wurde mit dem *Omniscript Reverse Transcription Kit* (Qiagen) nach Herstellerangaben durchgeführt. Dabei wurden, soweit nicht anders angegeben, je Reaktion 500 ng RNA eingesetzt. Für unspezifische cDNA-Synthesen wurde eine Mischung (1:1) von Oligo-dT/Random-Hexamer-Primern eingesetzt. Für spezifische reverse Transkriptionsreaktionen wurden 300 pmol der jeweiligen Primer verwendet. Für die anschließende PCR wurden verschiedenen Mengen der cDNA eingesetzt.

2.4.1.7 cRT-PCR

Die cRT-PCR dient der Zirkularisierung von RNA und der anschließenden Erzeugung von cDNA um unbekannte 5' und 3'-Enden von Transkripten durch eine PCR zu amplifizieren und zu sequenzieren. Dazu wurden zunächst 15 µg RNA mit 2 Units FastAP Alkalische Phosphatase und 3 µl 10×FastAP-Puffer für 30 min in einem Gesamtvolumen von 30 µl inkubiert. Anschließend wurde die dephosphorylierte RNA durch eine EtOH-Fällung gereinigt und das Pellet in 20 µl MQ-Wasser aufgenommen. Die Phosphorylierung des 5' Endes der RNA erfolgte mit Hilfe der T4-Polynukleotid-Kinase (T4-PNK). Dazu wurden zu der dephosphorylierten RNA 3 µl 10×T4-PNK Puffer, 3 µl 10 mM ATP und 10 Units T4-PNK gegeben und auf 30 µl mit MQ-Wasser aufgefüllt. Es folgte eine Inkubation für 30 min bei 37 °C und eine Hitzeinaktivierung für 20 min bei 65 °C. Anschließend wurde die RNA erneut gefällt und in 20 µl MQ-Wasser aufgenommen. Im folgenden Schritt wurden 2 µg der 5'-phosphorylierten RNA mit 2 µl 10×T4-RNA-Ligase Puffer, 2 µl 50 % PEG-8000; 2 µl 10 mM ATP, 1 µl RNase-Inhibitor und 20 Units T4-RNA-Ligase vermischt und auf ein Gesamtvolumen von 20 µl mit MQ-Wasser aufgefüllt. Es folgte eine Inkubation für 30 min bei 37 °C und eine anschließende Hitzeinaktivierung für 15 min bei 65 °C. Nach einer erneuten Fällung der RNA wurden für die cDNA-Synthese je 500 ng der zirkularisierten RNA eingesetzt.

2.4.2 Zellkultur

2.4.2.1 Kultivierung von *Dictyostelium discoideum*

Anders als die Wildtyp-Stämme der sozialen Amöben wurden von *Dictyostelium discoideum* Laborstämme etabliert, die axenisch, das heißt ohne Bakterien, kultiviert werden können (Ashworth & Watts, 1970). Der in dieser Arbeit verwendete Stamm AX2 wurde dazu in Petrischalen bei 22 °C in HL5-Vollmedium kultiviert. Schüttelkulturen wurden bei 150 rpm und 22 °C bis zu einer Dichte von $4-6 \times 10^6$ Zellen pro ml inkubiert, um einen Übergang in die stationäre Phase zu vermeiden. Für Selektionen wurden dem HL5-Medium verschiedene Antibiotika zugegeben. Die Konzentrationen betragen 10 µg/ml G418, 5 µg/ml Blastidicin und 35 µg/ml Hygromycin.

2.4.2.2 Kultivierung von Wildtyp-Stämmen

Da die Wildtyp-Stämme *Polysphondylium pallidum* und *Dictyostelium fasciculatum* nicht axenisch wachsen können mussten für die Kultivierung zunächst frische Bakterienkulturen von *Klebsiella planticola* hergestellt werden. In 1 ml der Bakteriensuspension wurden die Sporen einer Gefrierkultur mit einer Impföse vermischt und auf NA-Agarplatten ausgestrichen. Die Bakterien dienten den Wildtyp-Stämmen als Nahrung. Sobald keine Bakterien mehr vorhanden waren begannen die Amöben mit der Fruchtkörperbildung. Für die Extraktion von genomischer DNA wurden Kulturen verwendet, die kurz vor Beginn der Differenzierung abgeerntet und pelletiert wurden. Für die Herstellung von Gefrier-Stocks und eine erneute Kulturpassage wurden die Zellen so lang kultiviert, bis Sporentragende Fruchtkörper gebildet wurden.

2.4.2.3 Herstellung von Zellpellets

Für die Isolierung von Gesamt-RNA und genomischer DNA für die PCR wurden Pellets mit 2×10^7 Zellen verwendet. Dazu wurde die entsprechende Zellanzahl bei 4 °C und 150 g für 5 min zentrifugiert. Der Überstand wurde verworfen, das Pellet in 20 ml kaltem 1×Phosphatpuffer resuspendiert und erneut zentrifugiert. Dieser Schritt wurde ein weiteres Mal wiederholt. Anschließend wurde das Pellet in 1 ml 1×Phosphatpuffer resuspendiert und für 1 min bei 30.000 g zentrifugiert. Die Pellets wurden bei -80 °C gelagert und konnten für die Isolierung von genomischer DNA und RNA verwendet werden.

2.4.2.4 Herstellung von Gefrierkulturen

Für die dauerhafte Lagerung von AX2-Kulturen und Wildtyp-Stämmen wurden Gefrierkulturen von Sporen hergestellt. Dazu wurden die Zellen zunächst in Schüttelkulturen bis zu einer Dichte von $4-6 \times 10^6$ Zellen pro ml kultiviert. Anschließend wurden davon 1×10^8 Zellen für 5 min bei 150 g und 4 °C zentrifugiert und der Überstand verworfen. Das Zellpellet wurde nun in 20 ml kaltem 1×Phosphatpuffer resuspendiert und erneut für 5 min mit 150 g bei 4 °C zentrifugiert um Medienreste zu entfernen. Die Zellen wurden in 5 ml 1×Phosphatpuffer aufgenommen und auf eine Phosphatplatte gegeben. Nach einer Inkubation für 30 min, während der sich die Zellen absetzen, wurde der Phosphatpuffer abgesaugt und die Platten kurz gerocknet. Aufgrund des Nahrungsmangels beginnen die Zellen nach einiger Zeit mit der Aggregation und der darauf folgenden Bildung von Fruchtkörpern, welche innerhalb von 1–2 Tagen abgeschlossen ist. Zur Herstellung der Gefrierkulturen wurden die Platten abgekratzt und die Sporen in 500 µl 86 % Glycerin resuspendiert. Die Stocks wurden bei -80 °C gelagert. Für die erneute Kultivierung wurde wenig Material der Stocks in HL5-Medium ohne Antibiotika resuspendiert. Nachdem die vegetativen Zellen erneut ausgebildet waren, wurden die Zellen bei Bedarf in Selektionsmedium weiter kultiviert.

2.4.2.5 CaCl₂-Transformation von *D. discoideum*-Zellen

Zur stabilen Integration von Plasmiden in das Genom wurden die Zellen mit einer CaCl₂-basierten Methode transformiert (Gaudet *et al.* 2007). Dazu wurden zunächst Zellen in Petrischalen bis zu einer Konfluenz von rund 90 % in HL5-Medium kultiviert. Am Tag der Transformation wurde das HL5-Medium zunächst durch 10 ml MES-Medium ersetzt und die Zellen für mindestens 45 min darin inkubiert. Für Doppel-Transformationen wurden 15 µg, für Einzeltransformationen 30 µg je Plasmid, mit 300 µl 2×HBS und dem entsprechendem Volumen MQ-Wasser in einem Gesamtvolumen von 600 µl vermischt. Zu dieser Lösung wurden anschließend 38 µl 2M CaCl₂ unter moderatem Schütteln tropfenweise zugegeben. Es folgte eine Inkubation für 25 min bei Raumtemperatur. Nun wurde das MES-Medium von den Zellen abgesaugt und die Plasmid-Lösung vorsichtig auf die Mitte der Platten getropft. Anschließend wurden die Platten vorsichtig geschwenkt, die Zellen für 30 min inkubiert und danach 10 ml MES-Medium zugegeben. Die Zellen wurden nun für mindestens 3 Stunden stehen

gelassen. Nun wurde das MES-Medium abgenommen und durch 2 ml Schocklösung ersetzt, die vorsichtig an eine zuvor markierte Stelle des Petrischalen-Rands gegeben und durch vorsichtiges Schwenken verteilt wurden. Nach einer Inkubation für 8 min wurde die Schocklösung abgesaugt und durch HL5-Medium ersetzt. Am folgenden Tag wurde das Medium durch das entsprechende Antibiotika-haltige HL5-Medium ersetzt. Das Selektionsmedium wurde alle 2–3 Tage ersetzt. Nach 7-10 Tagen wurden resistente Klone sichtbar, welche die entsprechenden Plasmide aufgenommen hatten. Die Klone konnten entweder als Einzelklone gepickt und in Mikrotiterplatten weiter kultiviert werden, oder sie wurden für die folgenden Experimente vermischt (Pool). Für die weitere Lagerung wurden Gefrierkulturen hergestellt.

2.4.2.6 Elektrotransformation von *D. discoideum*-Zellen

Eine weitere Möglichkeit der Transformation von *D. discoideum* mit Plasmid-DNA ist die Elektroporation der Zellen. Dazu wurden zunächst Schüttelkulturen bis zu einer Dichte von $4\text{--}6 \times 10^6$ Zellen pro ml kultiviert. Nun wurden je Transformation $3\text{--}5 \times 10^7$ Zellen bei $4\text{ }^\circ\text{C}$ und 200 g für 5 min zentrifugiert, der Überstand anschließend verworfen und das Zellpellet in 20 ml kaltem $1 \times$ Phosphatpuffer aufgenommen. Nach einer weiteren Zentrifugation bei $4\text{ }^\circ\text{C}$ und 200 g für 5 min wurde der Überstand erneut verworfen und das Pellet in 20 ml kaltem Elektroporationspuffer resuspendiert. Es folgte eine weitere Zentrifugation ($4\text{ }^\circ\text{C}$, 200 g, 5 min) wonach der Überstand verworfen und das Pellet in 300 μl Elektroporationspuffer aufgenommen wurde. Die Zellsuspension wurde in vorgekühlte 4 mm Elektroporationsküvetten überführt und jeweils 8 μg Plasmid-DNA zugegeben. Die Elektroporation erfolgte anschließend mit folgenden Parametern: 1000 V; 1 ms Pulslänge; 2 Pulse mit 5 sek Pulsintervall. Nach einer Inkubation für 10 min bei $22\text{ }^\circ\text{C}$ wurden die Zellen in 1,5 ml Reaktionsgefäße mit je 3 μl 0,1 M CaCl_2 und 0,1 M MgCl_2 überführt. Es folgte eine Inkubation bei $22\text{ }^\circ\text{C}$ und 150 rpm für 15 Minuten im Schüttelinkubator. Anschließend wurden die Zellen in 10 ml HL5-Medium ausplattiert und das Medium am nächsten Tag durch Selektionsmedium ersetzt. Die Selektion und die weitere Kultur erfolgten analog zu CaCl_2 -transformierten Zellen.

2.4.2.7 Retrotranspositionstest

Der Retrotranspositionstest dient der Mobilisierung und Quantifizierung artifizierender Retrotransposons. Zunächst musste dazu das artifizierende Retrotransposon durch eine Transformation stabil in das Genom der Zellen integriert werden. Dies erfolgte durch eine Co-Transformation mit einem G418-Resistenz vermittelnden Plasmid (pISAR). Bei der Co-Transformation werden beide Plasmide in das Genom integriert, wobei zu erwarten ist, dass erfolgreich transformierte Zellen mehrere Kopien beider Plasmide enthalten. Nach der Transformation wurden die Zellen auf G418-resistente Klone selektioniert und diese nach 7–10 Tagen gepoolt. Anschließend wurden von den Pools Schüttelkulturen hergestellt, welche bis zu einer Zelldichte von $4\text{--}6 \times 10^6$ Zellen pro ml kultiviert wurden. Von diesen Schüttelkulturen wurden nun $1\text{--}10 \times 10^6$ Zellen in Petrischalen mit 10 ml HL5-Medium gegeben und bei $22\text{ }^\circ\text{C}$ inkubiert bis sich die Zellen auf dem Boden abgesetzt hatten. Das

Medium wurde daraufhin abgesaugt und durch 10 ml HL5-Medium mit 5 µg/ml Blastocidin ersetzt. Die weitere Selektion erfolgte für 7–10 Tage, wobei alle 2–3 Tage das Medium vorsichtig an einer zuvor markierten Stelle abgesaugt und durch frisches Selektionsmedium ersetzt wurde. Zellen in denen die artifiziellen Retrotransposons mobilisiert wurden sind Blastocidin-resistent. Die Blastocidin-resistenten Klone wurden einzeln gepickt und weiter kultiviert oder wurden gepoolt. Für weitere Analysen wurden davon jeweils Zellpellets angefertigt.

Für den Vergleich der Retrotranspositionsfrequenzen verschiedener Elemente wurden mehrere unabhängig erzeugte G418-Pools verwendet, wobei darauf geachtet wurde, dass die Kulturzeiten ähnlich waren. Nach der Auszählung der Blastocidin-resistenten Klone je Platte wurde die Frequenz der Sprungereignisse, bezogen auf 1×10^7 Zellen, errechnet.

2.5 In-silico-Analysen

2.5.1 Detektion neuer Retrotransposon-Spezies

Für die Suche nach bisher unbekanntem Retrotransposons aus den Genomen von *P. pallidum*, *D. fasciculatum* und *D. purpureum* wurden BLAST-Suchen durchgeführt. Grundlage für die Detektion waren die Reverse-Transkriptase-Domänen (RT) der Retrotransposons von *D. discoideum*. Dazu wurden zunächst lokale Blast-Datenbanken generiert und anschließend ein genomweiter Vergleich der RT-Aminosäuresequenzen aus *D. discoideum* mit allen drei möglichen offenen Leserahmen in beiden Orientierungen durchgeführt (*tblastn*). Gewertet wurden dabei alle Treffer mit einem Wahrscheinlichkeitswert kleiner als 1×10^{-15} . Die Regionen der Treffer wurden anschließend um einige Kilobasenpaare ober- und unterhalb erweitert und aus den Genomdaten extrahiert. Diese Sequenzen wurden entweder in Teilstücke fragmentiert und unter Verwendung des „*staden package*“-Programms assembliert oder mittels anderer Alignment-Software (Clustal W, MEGA 5) analysiert. Parallel zur Assemblierung wurden die erhaltenen DNA-Sequenzen nach charakteristischen Merkmalen von Retrotransposons, wie LTR-Sequenzen, Zielsequenzduplikationen und offenen Leserahmen abgesucht. Durch einen Vergleich dieser Daten konnten neue Retroelemente gefunden und die potentiell vollständigen Sequenzen der Retrotransposons generiert werden. Anschließend wurden weitere BLAST-Suchen mit den Elementsequenzen gegen das Gesamtgenom durchgeführt (*blastn*) um alle Fragmente zu finden und die entsprechenden Loci zu analysieren.

2.5.2 Erstellung phylogenetischer Stammbäume

Für die Erstellung phylogenetischer Stammbäume wurden die Aminosäuresequenzen der Reversen-Transkriptase-Domänen der Retrotransposons verwendet. Dazu wurden zunächst Alignment erstellt um konservierte Motive zu detektieren. Anschließend wurden alle Bereiche mit geringer Konservierung aus dem Datensatz entfernt. Für die Erstellung der Phylogenetischen Stammbäume wurden jeweils verschiedene Stammbäume mit der „*Neighbor-joining*“- , der Parsimonie- und der „*Maximum likelihood*“-Methode, unter Zuhilfenahme verschiedener Programme errechnet. Dabei

wurden jeweils 100 Bootstrap-Replikate erzeugt. Der Stammbaum mit den höchsten Bootstrap-Werten, welcher auch eine vergleichbare Topologie im Vergleich mit anderen Methoden aufwies, wurde als repräsentativ für die jeweiligen phylogenetischen Beziehungen angesehen.

3 Ergebnisse

In der vorliegenden Arbeit wird die Hypothese untersucht, dass die Integrationsspezifität bestimmter mobiler genetischer Elemente in der Nähe von tRNA-Genen eine Anpassung an ein kompaktes Genom wie das von *D. discoideum* ist. Es wird vermutet, dass die Integration in der Nähe von tRNA-Genen Insertionsmutagenesen in proteincodierenden Genen vermeidet, so dass den mobilen Elementen eine begrenzte Ausbreitung möglich ist, ohne dass die Fitness des Wirtes übermäßig reduziert wird. Entsprechend dieser Hypothese müsste man erwarten, dass in den vergleichbar kompakten Genomen anderer sozialer Amöben ähnliche Anpassungen von transposablen Elementen aufzufinden sind.

Die Frage, wie die Integrationsspezifität an tRNA-Genen entstanden ist, kann eventuell durch die Analyse von Genomen anderer sozialer Amöben abgeleitet werden (z. B. die Frage, wie die unterschiedliche Spezifität der phylogenetisch eng verwandten TRE5- und TRE3-Elemente in *D. discoideum* entstanden ist). In *D. discoideum* kann die Integrationsspezifität zum Teil experimentell untersucht werden. Das erlaubt jedoch keine Aussagen über die Evolution dieser Spezifität, insbesondere darüber, ob die Spezifität schon vor der Ausbreitung solcher Elemente in sozialen Amöben bestand oder tatsächlich erst in den sozialen Amöben als Anpassung an deren kompakte Genome entstanden ist. Allerdings ist zu bedenken, dass die Analyse von Retrotransposons durch erhebliche Genomveränderungen selbst zwischen vermeintlich nah verwandten Spezies der sozialen Amöben erschwert wird, da die Evolution dieser Gruppe von Organismen eine Zeitspanne von mehreren hundert Millionen Jahren umfasst (Abbildung 17) (Glöckner und Noegel 2013).

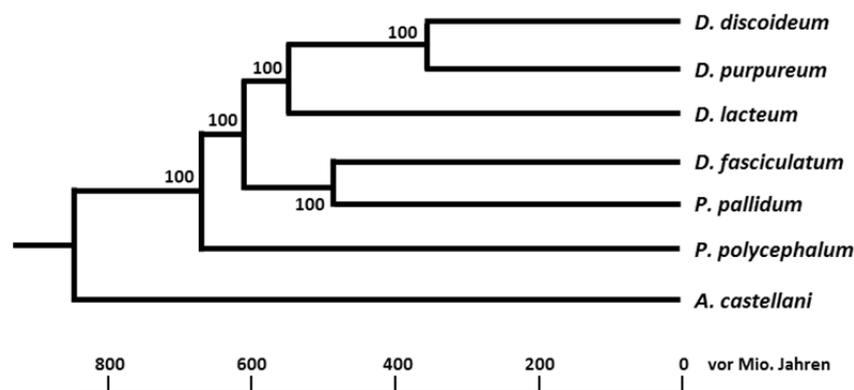


Abb. 17: Evolution und Phylogenie der sozialen Amöben.

Der letzte gemeinsame Vorfahre aller sozialen Amöben existierte vor rund 600 Mio. Jahren. Modifiziert nach Glöckner und Noegel 2013.

Im ersten Teil dieser Arbeit (Abschnitt 3.1) wird zunächst eine bioinformatische Auswertung verfügbarer Genome bezüglich identifizierbarer Retrotransposons durchgeführt. Danach werden im zweiten Teil (Abschnitt 3.2) experimentelle Untersuchungen zur genomweiten Ausbreitung artifiziell

in *D. discoideum*-Zellen eingebrachter Retrotransposons beschrieben, mit dem Ziel, durch die genomweite Auswertung von Neintegrationen Rückschlüsse auf Mechanismen der Zielerkennung zu ziehen. Die Aufklärung von Mechanismen der tRNA-Gen-Erkennung durch Retrotransposons in *D. discoideum* könnte nachfolgend auf mobile Elemente anderer Spezies ausgeweitet werden, um zu untersuchen, inwieweit es sich um allgemeine Prinzipien der Anpassung an kompakte Genome handelt.

3.1 Identifizierung von Retrotransposons aus den Genomen sozialer Amöben

Durch die Nutzung neuer Sequenzieretechnologien ist es in den letzten Jahren zu einem enormen Anstieg an verfügbaren Sequenzdaten ganzer Genome gekommen. Diese Entwicklung umfasste auch die Forschung an sozialen Amöben. Seit 2005 das Genom der sozialen Amöbe *D. discoideum* publiziert wurde (Eichinger *et al.* 2005), sind im Lauf der letzten Jahre die Genomdaten für *Polysphondylium pallidum*, *Dictyostelium fasciculatum* und *Dictyostelium purpureum* dazugekommen. Auf Grundlage dieser Daten war es daher interessant, auch in diesen Genomen nach Retrotransposons zu suchen, die sich an die Kompaktheit der Genome angepasst haben. Im Folgenden werden die Ergebnisse dieser Analysen näher erläutert. Die Konsensussequenzen der Elemente befinden sich auf einem Datenträger, welcher der vorliegenden Arbeit beiliegt.

3.1.1 Die Retrotransposons im Genom von *Dictyostelium purpureum*

Die genomischen Daten der sozialen Amöbe *D. purpureum* wurden erstmals im Jahr 2011 publiziert (Sucgang *et al.* 2011). Derzeit (Stand: August 2014) besteht das Genom aus 793 Supercontigs. Eine Zuordnung der annotierten Sequenzen zu einzelnen Chromosomen konnte daher nicht erfolgen.

Die Genomarchitektur ist vergleichbar mit *D. discoideum*. Sowohl Gengrößen, als auch die kurzen intergenischen Bereiche, zeigen die Charakteristika eines sehr kompakten Genoms. Die Anzahl der tRNA-Gene ist mit 353 Kopien nur unwesentlich geringer als bei *D. discoideum*. Basierend auf diesen Daten war es sehr wahrscheinlich, dass auch im Genom von *D. purpureum* mobile genetische Elemente zu finden sind, die sich, vergleichbar mit den tRNA-Gen-spezifischen Retrotransposons aus *D. discoideum*, an das kompakte Genom angepasst haben.

3.1.1.1 tRNA-Gen-assoziierte LTR-Retrotransposons

Dp_DGLT-A.1

Insgesamt konnten 15 Loci mit diesem Element detektiert werden. Ein vollständiges DGLT-A.1-Element hat eine Größe von 5436 bp und besitzt terminale LTR-Sequenzen mit einer Länge von 492 bp (Abbildung 18). Die Elemente wurden immer oberhalb von tRNA-Genen gefunden, wobei viele Elemente durch die Integration anderer Retrotransposons nicht mehr direkt neben einem tRNA-Gen liegen. Der durchschnittliche Abstand der sechs Elemente direkt oberhalb von tRNA-Genen beträgt 15 Basenpaare, was vergleichbar mit den Abständen des DGLT-A-Elements aus *D. discoideum*

ist (Marschalek *et al.* 1989; Hofmann *et al.* 1991). Auffällig ist, dass offenbar nur ein vollständiges Element mit intakten ORFs existiert, wohingegen die meisten anderen Fragmente Solo-LTR-Sequenzen darstellen. Es konnten keine Zielsequenzduplikationen gefunden werden. *Dp_DGLT-A.1* besitzt einen ORF mit 1413 Aminosäuren, der für alle zur Mobilisierung notwendigen Domänen codiert.

Dp_DGLT-A.2

Ein weiteres potentiell tRNA-Gen-spezifisches Retrotransposon ist das Element *Dp_DGLT-A.2*. Mit einer Größe von 6114 bp und einem einzigen ORF mit 1693 Aminosäuren ist es etwas größer als das *Dp-DGLT-A.1*-Element. Die LTR-Sequenzen haben eine Länge von 389 bp (Abbildung 18). Es konnte nur ein vollständiges Element annotiert werden, ein weiteres, nahezu komplettes Element wies eine Deletion auf, welche zu einer Unterbrechung des ORFs führt. Zu vermuten ist, dass auch dieses LTR-Retrotransposon 5' von tRNA-Genen spezifisch integriert, denn alle detektierten Fragmente wurden in unterschiedlichen Distanzen oberhalb von tRNA-Genen gefunden. Nur das komplette Element liegt direkt 15 bp oberhalb eines Serin-(TGA)-Gens. Bei größeren Abständen sind andere Retrotransposon-Spezies zwischen den tRNA-Genen und den *Dp_DGLT-A.2*-Fragmenten zu finden. Es konnten bei keinem der insgesamt neun Loci Zielsequenzduplikationen detektiert werden.

Dp_DGLT-A.3

Das LTR-Retrotransposon *Dp_DGLT-A.3* ist ebenfalls nur mit einer kompletten Kopie im Genom vorhanden. Das vollständige Fragment hat eine Größe von 5587 bp und besitzt einen ORF, welcher für 1384 Aminosäuren codiert. Die terminalen Sequenzwiederholungen haben eine Länge von 563 bp (Abbildung 18). Neben dem vollständigen Element konnten sieben weitere Loci gefunden werden, die mit *Dp_DGLT-A.3* assoziiert sind. Bei sieben Fragmenten konnten tRNA-Gene detektiert werden, wobei drei Kopien oberhalb von tRNA-Genen mit Abständen von 10 und 11 bp integriert sind. Die anderen Fragmente wurden durch neuere Integrationen anderer Retrotransposons von den benachbarten tRNA-Genen verdrängt. Bei vier *Dp_DGLT-A.3*-Elementen konnten TSD mit jeweils 4 bp Länge nachgewiesen werden.

Dp_Skipper-2

Ein weiteres tRNA-Gen-spezifisch integrierendes LTR-Retrotransposon ist das Element *Dp_Skipper-2*. Es konnten zahlreiche Fragmente unterhalb von tRNA-Genen identifiziert werden. Auffällig ist dabei die starke Fragmentierung, wobei die meisten Sequenzen durch Integrationen von Non-LTR-Retrotransposons der *Dp_TRE3*-Familie unterbrochen werden. Es konnten drei komplette Elemente detektiert werden, welche eine Länge von 5676 bp haben. Ein weiteres nahezu vollständiges Element hat eine Deletion von rund 200 bp. Flankiert von 315 bp großen LTR-Sequenzen codieren die vollständigen Skipper-2 Elemente für zwei offene Leserahmen, welche sich um 104 bp überlappen. Der erste ORF entspricht wahrscheinlich dem *GAG*-Gen, das zweite Gen mit 1219 Aminosäuren

codiert für eine Protease-, Integrase-, RNase-H- sowie eine Reverse Transkriptase-Domäne (Abbildung 18). Die vollständigen Elemente werden von 4 bp langen TSD flankiert, welche bei den zahlreichen kleineren Fragmenten, vermutlich durch die Integration weiterer mobiler genetischer Elemente, nicht gefunden werden konnten. Die Abstände zu den jeweiligen tRNA-Genen konnten aus diesem Grund auch nur für sieben Kopien ermittelt werden. Diese zeigen, dass die Retrotransposons zwischen 7 und 133 bp unterhalb von tRNA-Genen integriert sind.

Weitere potentielle tRNA-Gen-spezifische LTR-Retrotransposons

Neben den bislang erläuterten mobilen genetischen Elementen konnten im Genom von *D. pupureum* weitere repetitive Sequenzen nachgewiesen werden, welche offensichtlich mit tRNA-Genen assoziiert sind. Es war allerdings nicht möglich diese Elemente genauer zu charakterisieren. Möglicherweise handelt es sich dabei um LTR-Retrotransposons, die im Laufe der Zeit inaktiviert wurden und somit durch wiederholte Integrationen aktiver mobiler Elemente stark fragmentiert wurden. Dadurch ist es schwer ein komplettes Element zu rekonstruieren sowie phylogenetische Analysen durchzuführen.

Für das LTR-Retrotransposon-E konnten 20 Solo-LTR Sequenzen nachgewiesen werden. Des Weiteren war es möglich drei größere Fragmente mit einer maximalen Länge von 3447 bp zu identifizieren. Diese besitzen einen ORF, welcher für eine Zinkfingerdomäne vom CCHC-Typ, eine Protease- und den Teil einer Reversen Transkriptase-Domäne codiert. Die Abstände der sieben direkt mit tRNA-Genen assoziierten Kopien betragen durchschnittlich 55 bp. Nahezu alle identifizierten Fragmente befinden sich oberhalb von tRNA-Gen Loci.

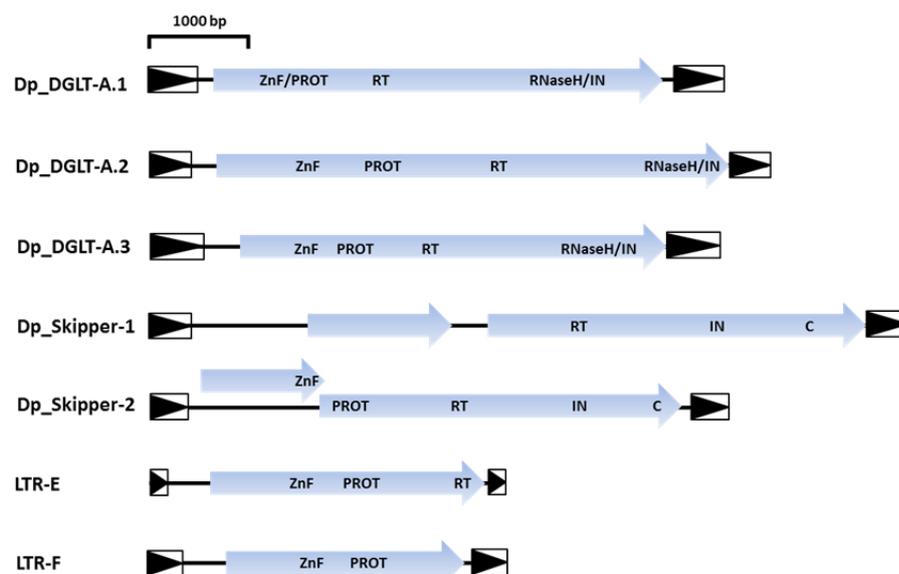


Abb. 18: Die Struktur der identifizierten LTR-Retrotransposons von *D. pupureum*.

Außer *Dp_Skipper-1* besitzen alle LTR-Retrotransposons eine tRNA-Gen-Spezifität (siehe Tabelle 3).

Das LTR-Retrotransposon-F, welches mit maximal 3440 bp ähnlich groß wie LTR-Element-E ist, wurde ebenfalls oberhalb von tRNA-Genen gefunden. Mit einem durchschnittlichen Abstand von 14 bp integrieren diese Retrotransposons ähnlich nah an tRNA-Gene wie die Vertreter der DGLT-A Familie. Neben 32 Solo-LTRs konnten jedoch nur drei größere Fragmente mit einem kurzen ORF gefunden werden. Dieser codiert lediglich für eine Zinkfinger- und eine Proteasedomäne.

Das bislang einzige auffindbare Retrotransposon ohne eine offensichtliche tRNA-Gen-Spezifität im Genom von *D. purpureum* ist das Element *Dp_Skipper-1*. Das vollständige LTR-Retrotransposon *Dp_Skipper-1* hat eine Länge von 7485 bp und besitzt LTR-Sequenzen mit einer Länge von 388 Basenpaaren (Abbildung 18).

3.1.1.2 Non-LTR-Retrotransposons aus *D. purpureum*

TRE3-Elemente

Im Genom von *D. purpureum* konnten drei unterschiedliche Non-LTR-Retrotransposons gefunden werden, welche unterhalb von tRNA-Genen integrieren. Ein Vergleich der Nukleotidsequenzen der einzelnen Retrotransposons zeigte keine signifikante Übereinstimmung. Ein Alignment der ORF2-Aminosäuresequenzen zeigt eine Übereinstimmung von 46 % zwischen *Dp_TRE3-A* und *Dp_TRE3-B*. Ein Alignment mit *Dp_TRE3-C* ist nicht möglich, da es sich vermutlich nur um Teile eines vollständigen Non-LTR-Retrotransposons handelt. Die Übereinstimmungen zwischen *Dp_TRE3-A* und *Dp_TRE3-B* finden sich nur in funktionell konservierten Bereichen, weshalb die drei Gruppen als eigenständige Familien betrachtet werden. Der Aufbau von *Dp_TRE3-A* und *Dp_TRE3-B* entspricht den TRE3-Elementen aus *D. discoideum*. Zwischen einer 5'-untranslatierten Region befinden sich zwei ORFs, welche nicht überlappen. Die Elemente enden 3' in einer weiteren UTR, welche durch einen Poly-A-Stretch terminiert wird.

Dp_TRE3-A

Betrachtet man die detektierbaren Kopien der TRE3-A-Familie fällt zunächst die starke Fragmentierung auf. Es finden sich zahlreiche kurze Sequenzen, welche nur wenige hundert Basenpaare groß sind. Es konnten nur zwei komplette und offensichtlich intakte Kopien im Genom von *D. purpureum* aufgefunden werden. Beide besitzen zwei ORFs, die für die zu erwartenden Domänen eines Non-LTR-Retrotransposons codieren. Das größte Element hat eine Länge von ≈ 5150 bp und besitzt einen 19 bp Poly-A-Anhang. Es konnte jedoch nur für drei kurze Fragmente der rund 60 Kopien eine potentielle Zielsequenzduplikation mit einer Größe von 6 bis 10 bp nachgewiesen werden. Daher ist es nicht möglich das 5'-Ende der TRE3-A-Elemente exakt zu definieren.

Viele Kopien befinden sich in größerer Distanz zu benachbarten tRNA-Genen, was auf die Integration anderer tRNA-Gen-assoziiierter Retrotransposon-Spezies zurückzuführen ist. Der mittlere Abstand der TRE3-A-Kopien direkt unterhalb von tRNA-Genen beträgt rund 98 bp. Dabei variiert der Bereich zwischen 34 und 161 bp.

Dp_TRE3-B

Mit rund 5210 bp ist das *Dp_TRE3-B* Non-LTR-Retrotransposon nur unwesentlich größer als die *Dp_TRE3-A*-Elemente. Es findet sich auch hier eine starke Fragmentierung. Neben insgesamt 20 Kopien ist im Genom nur ein vermutlich intaktes Element nachweisbar. Die *Dp_TRE3-A*-Kopien befinden sich in Abständen von 64 bis 185 bp unterhalb von tRNA-Genen, wobei der mittlere Abstand 108 bp beträgt. Wie schon bei den *Dp_TRE3-A*-Retrotransposons finden sich bei den *Dp_TRE3-B*-Elementen nahezu keine TSD, was eine exakte Abgrenzung der Fragmente verhindert.

Dp_TRE3-C

Bei *Dp_TRE3-C* handelt es sich vermutlich nur um den Teil eines autonomen Retrotransposons, oder um ein nicht-autonomes Element welches *in trans* durch andere Retrotransposons mobilisiert wird. Mit rund 1620 bp ist es wesentlich kürzer als die *Dp_TRE3-A*- und *Dp_TRE3-B*-Elemente. Es codiert nur für einen 936 bp großen ORF, in dem keine bislang bekannte Proteindomäne nachweisbar ist. Das Element endet in einem Poly-A-Schwanz, was dafür spricht, dass es sich bei *Dp_TRE3-C* um ein Non-LTR-Retrotransposon handelt. Es können rund 30 Kopien im Genom von *D. purpureum* detektiert werden, allerdings sind keine TSD nachweisbar. Die *Dp_TRE3-C*-Elemente integrieren offenbar in einem etwas größeren Abstand als *Dp_TRE3-A* und *Dp_TRE3-B* unterhalb von tRNA-Genen. Für die direkt mit tRNA-Genen assoziierten Elemente konnte ein durchschnittlicher Abstand von rund 136 bp ermittelt werden.

Dp_TRE5-ähnliche-Elemente

Neben den *Dp_TRE3*-Elementen konnte ein weiteres Non-LTR-Retrotransposon gefunden werden, welches eine tRNA-Gen-Spezifität besitzt. Das *Dp_TRE5-like* Element integriert in einem Abstand von durchschnittlich 4 bp oberhalb von tRNA-Genen und damit wesentlich näher als die bislang bekannten TRE5-Elemente. Ein intaktes Element hat eine Länge von rund 7440 bp und codiert für zwei offene Leserahmen, welche um 14 bp überlappen. Interessanterweise kann im ORF2, neben einer Endo-/Exonuklease- und einer Reversen Transkriptase-Domäne, auch eine RNase-H-Domäne gefunden werden. Die bislang im Genom von *D. discoideum* identifizierten TRE5-Elemente zeigen keine Homologien mit bekannten RNase-H-Domänen.

Eine weitere markante Auffälligkeit, im Vergleich zu den anderen annotierten Retrotransposon-Spezies im Genom von *D. purpureum*, ist das Vorhandensein von längeren Zielsequenzduplikationen bei einem Großteil der 28 auffindbaren Kopien. Nahezu alle Fragmente direkt oberhalb von tRNA-Genen besitzen TSD mit einer Länge von 15 bis 17 bp.

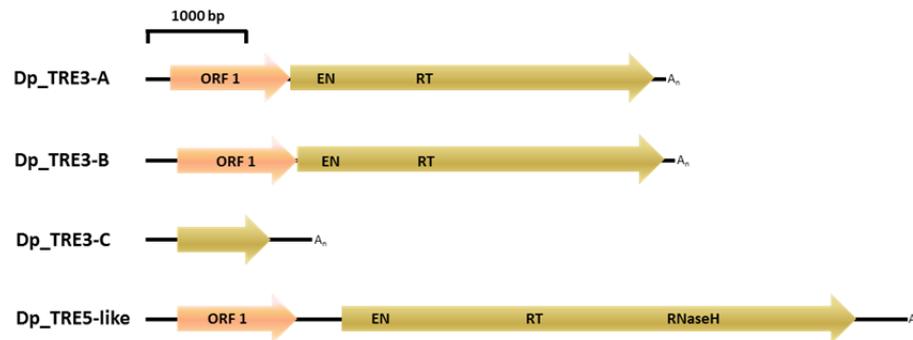


Abb. 19: Die Non-LTR-Retrotransposons aus *D. purpureum*.

3.1.1.3 Zusammenfassende Bewertung

Vor allem die Anzahl, Diversität und Dichte von tRNA-Gen-spezifischen LTR-Retrotransposons im Genom von *D. purpureum* steht im starken Kontrast zur Situation, die im Genom von *D. discoideum* vorliegt. In *D. discoideum* ist mit DGLT-A nur ein tRNA-Gen-spezifisches LTR-Retrotransposon bekannt (Glöckner *et al.* 2001; Winckler *et al.* 2002). In der vorliegenden Arbeit konnten dahingegen sechs unterschiedliche LTR-Elemente im Genom von *D. purpureum* identifiziert werden, die offensichtlich tRNA-Gen-abhängig integriert wurden. Mit einem Anteil von rund 1,5 % am Genom von *D. discoideum* ist die TRE5-Familie in diesem Organismus relativ stark vertreten. Dahingegen konnten in *D. purpureum* keine vergleichbaren TRE5-Elemente gefunden werden. Mit dem *Dp_TRE5-like*-Element existiert zwar ein Non-LTR-Retrotransposon das oberhalb von tRNA-Genen integriert, das Vorhandensein einer RNase-H-Domäne und der sehr kurze Abstand zu den tRNA-Genen spricht allerdings dafür, dass es sich dabei um eine andere Non-LTR-Retrotransposon-Spezies handelt. Daher ist es erstaunlich, dass es, vor allem in Hinblick auf die geringe phylogenetische Distanz zwischen beiden Arten, zu einer vollständigen Extinktion der TRE5-Familie in *D. purpureum* gekommen ist. Es ist jedoch nicht auszuschließen, dass noch rudimentäre Fragmente ehemaliger TRE5-Elemente im Genom von *D. purpureum* vorhanden sind, diese allerdings im Rahmen dieser Arbeit nicht gefunden wurden. Insgesamt liegt der genomische Anteil der Retrotransposons in *D. purpureum* mit rund 1 % deutlich unter dem Wert von *D. discoideum* mit 8 % (Glöckner *et al.* 2001).

Tab. 3: Die Retrotransposons im Genom von *Dictyostelium purpureum*.

Name	genomischer Anteil [kb]	genomischer Anteil in %	Kopienzahl	Größe [bp]	Anzahl Solo-LTRs	LTR-Größe [bp]	tRNA-Gen-spezifisch	tRNA-Gen-Abstände [bp]	Ø-Abstand zu tRNA-Gen [bp]	TSD	Größe TSD [bp]
DGLT-A.1	18	0,055	15	5436	9	492	5'	13-16	15	nein	-
DGLT-A.2	14,5	0,044	9	6114	5	389	5'	15	15	nein	-
DGLT-A.3	12,4	0,038	8	5589	6	563	5'	10-11	10	ja	4
Skipper-1	29	0,088	≈12	7485	6	388	nein	-	-	n. d.	-
Skipper-2	36	0,109	23	5676	5	315	3'	7-133	64	ja	4
LTR-E	21	0,064	30	3447	20	206	5'	19-116	55	nein	-
LTR-F	18	0,055	32	3440	32	354	5'	10-19	14	nein	-
TRE3-A	40,5	0,123	≈60	5150	-	-	3'	34-161	98	ja	6-10
TRE3-B	31	0,094	≈20	5210	-	-	3'	64-185	108	nein	-
TRE3-C	31,5	0,095	≈30	1620	-	-	3'	67-242	136	nein	-
TRE5-like	79,5	0,241	28	7438	-	-	5'	3-6	4	ja	15-17
Σ 1,006 %											
LTR-Retrotransposons						Non-LTR-Retrotransposons					

3.1.2 Die Retrotransposons im Genom von *Polysphondylium pallidum*

Mit 273 tRNA-Genen beherbergt das Genom der sozialen Amöbe *P. pallidum* rund 130 potentielle Integrationsorte weniger als *D. discoideum*. Der evolutionäre Abstand zwischen diesen beiden Arten ist wesentlich größer als zwischen *D. discoideum* und *D. purpureum* (Schaap *et al.* 2006; Heidel *et al.* 2011; Glöckner und Noegel 2013). Insgesamt sind die Genome von der Kompaktheit allerdings vergleichbar; sowohl in Hinblick auf die Gendichte, also die codierenden Einheiten pro Megabase, als auch in den Abständen der benachbarten Gene voneinander. (Heidel *et al.* 2011; vgl. Tab. 1). Somit war zu erwarten, dass es in *P. pallidum* spezialisierte mobile Elemente gibt, welche sich mit ihrer Integrationspezifität an die Gegebenheiten eines kompakten Genoms angepasst haben.

3.1.2.1 tRNA-Gen-spezifische LTR-Retrotransposons

Pp_Skipper-1

Das 5589 bp große LTR-Retrotransposon integriert unterhalb von tRNA-Genen. Es wird durch 226 bp große LTR-Sequenzen begrenzt und codiert für einen ORF. Charakteristisch für Mitglieder der Skipper-Familie ist das Vorhandensein einer Chromodomäne am Ende des ORFs. Daneben lassen sich auch eine Protease-, Reverse Transkriptase- sowie RNase-H-/Integrase-Domäne finden (Abbildung 20, Tabelle 4). Im Genom von *P. pallidum* lässt sich bislang nur ein vollständiges Element nachweisen, wobei insgesamt 14 Loci mit *Pp_Skipper-1* assoziiert sind. Viele der weiteren genomischen Kopien bestehen nur aus einem Teil der LTR-Sequenz. Für neun Kopien, welche direkt an tRNA-Genen gefunden wurden, konnte ein mittlerer Abstand von 77 bp ermittelt werden.

Pp_Skipper-2

Das Retrotransposon *Pp_Skipper-2* konnte nicht vollständig rekonstruiert werden. Bei keiner der Kopien sind flankierende LTR-Sequenzen oder TSD zu finden. Das größte Fragment im Genom von *P. pallidum* ist rund 3675 bp lang und codiert für zwei nicht überlappende offene Leserahmen. Im ersten Leserahmen finden sich Homologien mit einer Protease- sowie einer Reversen Transkriptase-Domäne (Abbildung 20, Tabelle 4). Ein Datenbankvergleich zeigt, dass die Aminosäuresequenz der RT Ähnlichkeiten mit Reversen Transkriptasen anderer LTR-Retrotransposons aufweist (vgl. Abschnitt 1.3). Des Weiteren sind im ORF2 eine RNase-H-/Integrase- und eine Chromodomäne nachweisbar. Das Vorhandensein einer Chromodomäne und die Homologie der Reversen Transkriptase-Domänen legt eine Verwandtschaft mit der Skipper-Familie nahe. Insgesamt zehn Elemente sind im Genom nachweisbar, wobei zwei Fragmente direkt unterhalb von tRNA-Genen in Abständen von 85 und 151 bp integriert sind. Vermutlich handelt es sich um ein inaktiviertes Element. Dies wäre eine Erklärung, warum eine starke Fragmentierung zu beobachten ist und viele der Kopien nicht mehr in der Nähe von tRNA-Genen gefunden werden können.

Pp_LTR-Retrotransposon-H

Von diesem Element konnte, neben zahlreichen kleineren Fragmenten, nur eine einzelne größere Sequenz gefunden werden. Dadurch war es nicht möglich die exakten Grenzen des Retrotransposons zu bestimmen. Insgesamt wurden zehn Fragmente identifiziert, wobei alle oberhalb von tRNA-Genen zu finden sind. Das größte Element ist rund 3650 bp lang und codiert für einen ORF in dem eine Protease, eine Reverse Transkriptase sowie ein Zinkfingermotiv vom Typ CCHC nachweisbar sind (Abbildung 20, Tabelle 4). Die RT zeigt signifikante Übereinstimmungen mit Sequenzen von Reversen Transkriptasen anderer LTR-Retrotransposons. Ein weiteres Indiz dafür, dass dieses Element zur Gruppe der LTR-Retrotransposons gehört, ist der Beginn der Sequenzen mit dem für LTR-Elemente charakteristischen TG-Dinukleotid. Allerdings konnten keine kompletten flankierenden LTR-Sequenzen nachgewiesen werden. Es ist davon auszugehen, dass es sich bei den gefundenen Fragmenten nur noch um Relikte eines ehemals aktiven Elements handelt. Dadurch kommt es zur Akkumulation von Mutationen, was die Detektion von kompletten LTR-Sequenzen zusätzlich erschwert. Drei der Elemente liegen mit Abständen von 25, 14 und 16 bp direkt oberhalb von tRNA-Genen, die anderen sind durch Integrationen anderer Retrotransposons von den tRNA-Genen verdrängt worden. Diese Integrationsspezifität lässt den Schluss zu, dass es bei dem Element um ein DGLT-A-ähnliches Retrotransposon handelt

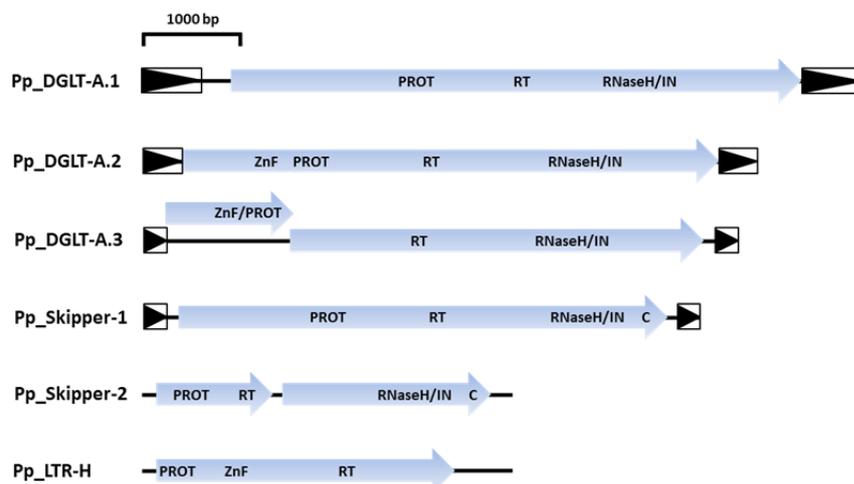


Abb. 20: Die LTR-Retrotransposons von *P. pallidum*.

3.1.2.2 Weitere LTR-Retrotransposons im Genom von *P. pallidum*

Pp_DGLTA-Familie

Im Genom von *P. pallidum* sind drei unterschiedliche Gruppen von LTR-Retrotransposons gefunden worden, welche aufgrund struktureller und phylogenetischer Merkmale eine Verwandtschaft mit Mitgliedern der DGLT-A-Familie zeigen (Abbildung 20; Tabelle 4). Bei keinem der Elemente konnten tRNA-Gene in der näheren Umgebung identifiziert werden, weshalb diese Elemente als nicht-tRNA-

Gen-spezifisch eingeordnet wurden. Die vollständigen Elemente dieser Gruppe werden durch LTR-Sequenzen flankiert und codieren, in ein oder zwei offenen Leserahmen, für alle zur autonomen Mobilisierung benötigten Proteine, weshalb davon auszugehen ist, dass es sich um aktive Populationen handelt.

3.1.2.3 Die Non-LTR-Elemente von *P. pallidum*

Insgesamt konnten im Genom von *P. pallidum* fünf unterschiedliche Elemente identifiziert werden, die der Klasse der Non-LTR-Retrotransposons zugeordnet werden konnten. Neben vier spezifisch integrierenden Elementen konnte des Weiteren ein offensichtlich nicht-tRNA-Gen-spezifisches Non-LTR-Retrotransposon nachgewiesen werden.

Pp_TRE3-Elemente

Unterhalb von tRNA-Genen konnten zwei unterschiedliche Non-LTR-Retrotransposons identifiziert werden. Beide Elemente besitzen je zwei offene Leserahmen, es sind jedoch nur sehr wenige konservierte Domänen nachweisbar. Beide Elemente enden in einem für Non-LTR-Elemente typischen Poly-A-Schwanz. Für das erste Element, das 4515 bp große *Pp_TRE3-1*, konnte lediglich eine Reverse Transkriptase-Domäne gefunden werden. Im ORF2 vom 4741 bp großen *Pp_TRE3-2* sind eine Exo-/Endonuklease- und eine RT-Domäne detektierbar. Für den ORF1 beider Elemente konnte keine konservierte Domäne detektiert werden, was für die ORF1-Proteine der Non-LTR-Retrotransposons, durch die allgemein geringe Konservierung, zu erwarten war.

Bei einem Blick auf das *Pp_TRE3-1*-Element fiel auf, dass beide ORFs durch 1,1 kb nicht-codierender Sequenz getrennt werden. Diese Lücke wurde allerdings bei allen nahezu vollständigen Elementen vorgefunden, was darauf schließen lässt, dass diese Retrotransposons auch mit dieser Sequenz mobilisiert wurden. Eine Vermutung wäre, dass es sich um eine *Trans*-Mobilisierung handelt, da wichtige Proteine durch eine Mutation des Start-Codons im ORF2 offensichtlich fehlen. Von den rund 30 genomischen *Pp_TRE3-1*-Kopien konnten zehn direkt vor tRNA-Genen gefunden werden, wobei ein mittlerer Abstand von 67 bp festgestellt wurde. Viele dieser Elemente konnten mit TSD von 7 bis 15 bp Länge in Verbindung gebracht werden.

In den *Pp_TRE3-2* Elementen beträgt der Abstand zwischen ORF1 und ORF2 nur 332 bp. Mit einem durchschnittlichen Abstand von 87 bp zu den oberhalb liegenden tRNA-Genen, scheint dieses Retrotransposon in etwas größerem Abstand zu integrieren als *Pp_TRE3-1*. Im Gegensatz zu *Pp_TRE3-1* konnten bei den rund 30 tRNA-Gen-assoziierten Elementen keine flankierenden Zielsequenzverdopplungen nachgewiesen werden.

Pp_TRE5

Das Non-LTR-Retrotransposon *Pp_TRE5* integriert im Mittel mit einem Abstand von 45 bp oberhalb von tRNA-Genen. Auffällig ist, dass die Hälfte der Elemente exakt 38 bp neben das tRNA-Gen gesprungen ist. Das 1169 bp große mobile genetische Element ist nicht-autonom, da es in einem

einzigem ORF lediglich für den Teil einer Reversen Transkriptase codiert (Abbildung 21, Tabelle 4). Am 3'-Ende wird es durch einen Poly-A-Schwanz terminiert, was die Zugehörigkeit zu den Non-LTR-Retrotransposons erkennen lässt. Von den 19 auffindbaren Elementen im Genom befinden sich zwölf direkt an tRNA-Genen. Diese Beobachtung könnte ein Indiz für relativ junge Sprungereignisse von *Pp_TRE5* sein.

TRE5-ähnliche Elemente

Ein weiteres Non-LTR-Retrotransposon mit einer tRNA-Gen-Spezifität, integriert vermutlich mit einem Abstand von 50 bp oberhalb von tRNA-Genen. Im Genom von *P. pallidum* konnte lediglich ein größeres Fragment identifiziert werden, weshalb eine weitere Charakterisierung nur schwer möglich war. Neben diesem Element mit einer Größe von rund 5550 bp konnten nur zwei weitere kleinere Fragmente in der Nähe von tRNA-Genen annotiert werden. Das TRE5-ähnliche Element scheint eine Vielzahl von Mutationen akkumuliert zu haben, was eine mögliche Erklärung für das Vorhandensein von vier potentiellen ORFs wäre. Davon lassen sich in drei offenen Leserahmen jeweils Domänen für eine Endo-/Exonuklease, eine RT vom Non-LTR-Typ sowie eine RNase-H detektieren. Das Element endet 3' offenbar mit einem Poly-A-Schwanz. Phylogenetische Analysen zeigten, dass offenbar keine Verwandtschaft mit den TRE5-Elementen aus *D. discoideum* vorliegt (siehe Abbildung 25 in Abschnitt 3.1.4.2).

3.1.2.4 Ein weiteres Non-LTR-Retrotransposon im Genom von *P. pallidum*

Neben den tRNA-Gen-spezifisch integrierenden Non-LTR-Elementen konnte ein weiteres Element identifiziert werden, das offensichtlich keine Integrationspezifität für tRNA-Gene besitzt. Das 3536 bp große Element *Pp_Non-LTR-1* ist mit zwölf Kopien im Genom von *P. pallidum* vertreten. In einem der drei ORFs ist eine RT-Domäne detektierbar, die Ähnlichkeiten mit Reversen Transkriptasen anderer Non-LTR-Retrotransposons aufweist. Es scheint allerdings keine Verwandtschaft mit einem anderen mobilen genetischen Element aus dem Genom von *P. pallidum* vorzuliegen.

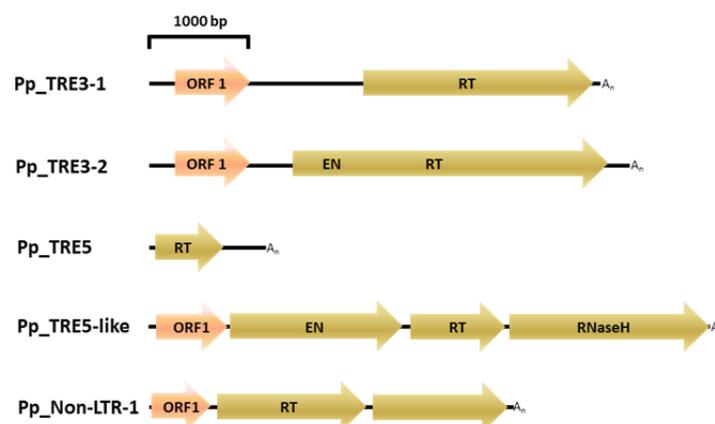


Abb. 21: Die Non-LTR-Retrotransposons der sozialen Amöbe *P. pallidum*.

Tab. 4: Charakteristika der Retrotransposons von *Polysphondylium pallidum*.

Name	genomischer Anteil [kb]	genomischer Anteil in %	Kopienzahl	Größe [bp]	Anzahl Solo-LTRs*	LTR-Größe [bp]	tRNA-Gen-spezifisch	tRNA-Gen-Abstände [bp]	Ø-Abstand zu tRNA-Gen [bp]	TSD	Größe TSD [bp]
Skipper-1	11	0,033	14	5589	10	226	3'	54-136	77	?	4-?
Skipper-2	18	0,055	9	3675	-	-	3'	85/151	118	nein	-
LTR-H	6,5	0,020	10	≈ 3650	-	-	5'	14/16/25	≈ 20	nein	-
DGLT-A.1	41	0,124	23	7295	13	601	nein	-	-	ja	5
DGLT-A.2	20,5	0,062	12	6160	6	393	nein	-	-	ja	3/5
DGLT-A.3	23	0,070	11	5942	3	212	nein	-	-	ja	5
TRE3-1	39	0,118	≈ 30	4515	-	-	3'	46-138	67	ja	7-15
TRE3-2	27	0,082	≈ 30	4741	-	-	3'	57-151	87	nein	-
TRE5	10	0,030	19	1169	-	-	5'	38-74	45	ja	8-11
TRE5-like	7,5	0,023	3	≈ 5500	-	-	5'	≈ 50	≈ 50	nein	-
Non-LTR-1	17	0,052	12	3536	-	-	nein	-	-	nein	-
Σ 0,669 %											
LTR-Retrotransposons						Non-LTR-Retrotransposons					

3.1.3 Die Retrotransposons im Genom von *Dictyostelium fasciculatum*

Das Genom von *D. fasciculatum* beherbergt mit 198 tRNA-Genen nur gut die Hälfte an potentiellen Integrationsorten von tRNA-Gen-spezifischen Retrotransposons verglichen mit *D. discoideum*. Die grundlegende Genomarchitektur gleicht der des Genoms von *D. discoideum* (Eichinger *et al.* 2005; Heidel *et al.* 2011). Daher sind auch hier tRNA-Gen-spezifische Retrotransposons zu erwarten. Im Folgenden werden die einzelnen Retrotransposons näher erläutert, die bei der Analyse des Genoms von *D. fasciculatum* identifiziert wurden.

3.1.3.1 Die LTR-Retrotransposons von *D. fasciculatum*

Df_DGLT-A

Das LTR-Retrotransposon *Df_DGLT-A* stellt eine Besonderheit in der Gesamtheit aller identifizierten LTR-Retrotransposons aus den Genomen der sozialen Amöben dar, denn es vereint typische Charakteristika von DGLT-A- und Skipper-Elementen. Das 5708 bp große Element codiert in zwei nicht-überlappenden Leserahmen für alle zur autonomen Mobilisierung benötigten Proteine (Abbildung 22). Neben einer für Skipper-Elemente typischen Chromodomäne findet sich eine tRNA-Primerbindestelle (PBS), welche Ähnlichkeiten mit den DGLT-A-Elementen aufweist (siehe Abbildung 24 in Abschnitt 3.1.4.1).

Alle bisher identifizierten Kopien liegen unterhalb von tRNA-Genen, mit einem mittleren Abstand von 58 bp. Insgesamt konnten, neben fünf Solo-LTR-Fragmenten, sieben potentiell vollständige und somit vermutlich aktive Elemente im Genom von *D. fasciculatum* identifiziert werden. Ungewöhnlich für LTR-Retrotransposons ist der Beginn der LTR-Sequenzen mit dem Dinukleotid 5'-TA..CA-3'. Offensichtlich werden die Elemente mit diesem Motiv mobilisiert, da alle detektierbaren genomischen Kopien damit assoziiert sind. Bislang sind nur wenige LTR-Retrotransposons mit einem ähnlichen Motiv bekannt (Yin *et al.* 2013). Bei mehreren Elementen konnten TSD mit einer Länge von 5 bis 7 bp charakterisiert werden.

Df_Skipper-1

Im Genom von *D. fasciculatum* konnte ein Retrotransposon gefunden werden, das einen ähnlichen Aufbau wie die Elemente der Skipper-Familie besitzt. So lässt sich am Ende des zweiten ORFs eine Chromodomäne detektieren, welche für diese Art von LTR-Retrotransposons typisch ist (Abbildung 22). Allerdings war es nicht möglich flankierende LTR-Sequenzen nachzuweisen. Das Element besitzt keine tRNA-Gen-Spezifität. Insgesamt sind nur fünf Loci mit Fragmenten von *Df_Skipper-1* auffindbar.

Df_DGLT-P

Für das Element *Df_DGLT-P* konnte keine vollständige Konsensussequenz generiert werden, die vermutliche Größe liegt bei rund 4,2 kb. Auf der Grundlage von Datenbankvergleichen konnten

Homologien mit dem DGLT-P-Retrotransposon aus dem Genom von *D. discoideum* nachgewiesen werden. Es liegt keine tRNA-Gen-Spezifität vor. Es codiert für mehrere kürzere ORFs, in denen allerdings keine funktionellen Domänen nachweisbar sind.

Df_DIRS

Die 20 genomischen Kopien dieses LTR-Retrotransposons zeigen keine tRNA-Gen-Spezifität. Es konnte eine Sequenz mit einer Länge von 2283 bp rekonstruiert werden, welche Homologien mit dem DIRS-Element aus *D. discoideum* aufweist. Allerdings konnten keine Hinweise darauf gefunden werden, dass *Df_DIRS* Cluster bildet, die, analog zu den Elementen aus *D. discoideum*, eventuell an der Bildung von Centromeren beteiligt sein können (Eichinger *et al.* 2005).

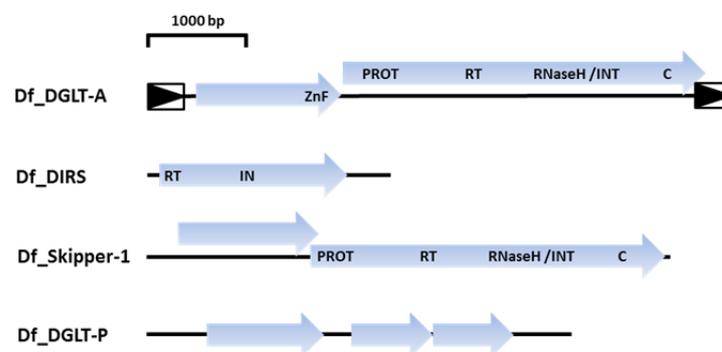


Abb. 22: Die LTR-Retrotransposons aus dem Genom von *D. fasciculatum*.

3.1.3.2 Non-LTR-Retrotransposons

TRE5-Elemente

Insgesamt konnten vier unterschiedliche Gruppen von Non-LTR-Retrotransposons oberhalb von tRNA-Genen nachgewiesen werden. Zum einen gibt es drei voneinander unterscheidbare Elemente, welche allerdings identische 3'-Enden besitzen und im Folgenden als *Df_TRE5-A*-Familie bezeichnet werden. Zum anderen konnte ein weiteres Non-LTR-Retrotransposon 5' von tRNA-Genen identifiziert werden, welches keine Sequenzhomologien mit den *Df_TRE5-A*-Elementen aufweist und darum *Df_TRE5-B* genannt wird. Alle *Df_TRE5*-Elemente enden 3' in einem für Non-LTR-Retrotransposons typischen Poly-A-Schwanz variabler Länge (Abbildung 23).

***Df_TRE5-A*-Familie**

Das größte Element der TRE5-A-Familie, das Non-LTR-Retrotransposon *Df_TRE5-A.1*, ist maximal 2587 bp lang. Es besitzt einen ORF mit 701 Aminosäuren, der lediglich für eine Reverse Transkriptase codiert. Die beiden anderen Elemente, *Df_TRE5-A.2* und *Df_TRE5-A.3*, haben keine größeren offenen Leserahmen. Im Durchschnitt integrieren alle drei Gruppen zwischen 55 und 59 bp oberhalb

von tRNA-Genen, wobei bei vielen Kopien auch Zielsequenzduplikationen nachgewiesen werden können. Bei den *Df_TRE5-A.2*-Elementen finden sich zwei unterschiedliche Untergruppen. Zum einen existieren Elemente, bei denen nur die letzten 429 bp identisch mit den *Df_TRE5-A.1*-Elementen sind. Zum anderen existieren vier Kopien, welche am 3'-Ende 830 bp Sequenzidentität mit *Df_TRE5-A.1* aufweisen. Die 5'-gelegene Sequenz aller *Df_TRE5-A.2*-Elemente, welche sich von den *Df_TRE5-A.1*-Retrotransposons unterscheidet, ist jedoch identisch. Vermutlich handelt es sich bei allen drei Gruppen entweder um bereits ausgestorbene Elemente, oder sie werden weiter *in trans* durch Proteine anderer Retrotransposons mobilisiert. Ein Argument für die bestehende Aktivität der *Df_TRE5-A*-Elemente ist die Anzahl von Zielsequenzverdopplungen (Tabelle 5), die bei einer inaktivierten Population relativ schnell durch die Akkumulation von Mutationen degeneriert und somit nicht länger nachweisbar wären.

Df_TRE5-B

Das Non-LTR-Retrotransposon *Df_TRE5-B* ist mit maximal 1584 bp ebenfalls sehr kurz und codiert in einem ORF nur für eine Reverse Transkriptase. Im Genom von *D. fasciculatum* konnte nur ein Fragment gefunden werden, das dieser Größe entspricht. Der mittlere Abstand zu benachbarten tRNA-Genen beträgt 57 bp. Zielsequenzverdopplungen konnten nur bei zwei der 26 genomischen Kopien festgestellt werden, was ein Indiz dafür sein kann, dass es sich bei den *Df_TRE5-B*-Retrotransposons um eine bereits inaktivierte Population von mobilen genetischen Elementen handelt.

Df_TRE3

Mit dem *Df_TRE3*-Element aus *D. fasciculatum* konnte lediglich eine Gruppe von Non-LTR-Retrotransposons nachgewiesen werden, die unterhalb von tRNA-Genen integrieren. Dabei schwankt der Bereich der Integrationen von minimal 29 bis über 100 bp, mit einem durchschnittlichen Abstand von 79 bp. Bei sieben von 14 Elementen konnten TSD nachgewiesen werden. Nur eine Kopie besitzt mit einer Länge von 1867 bp einen einzigen ORF, welcher für eine Reverse Transkriptase codiert.

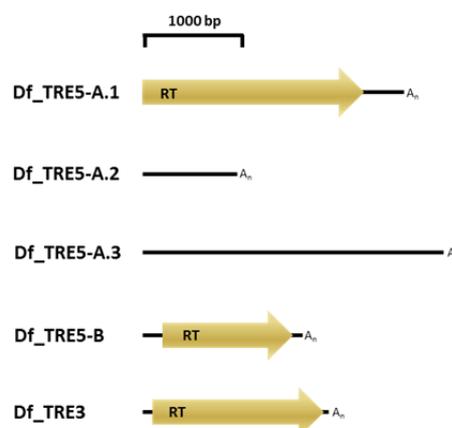


Abb. 23: Die Non-LTR-Retrotransposons von *D. fasciculatum*.

Tab. 5: Retrotransposons im Genom von *Dictyostelium fasciculatum*.

Name	genomischer Anteil [kb]	genomischer Anteil in %	Kopienzahl	Größe [bp]	Anzahl Solo-LTRs	LTR-Größe [bp]	tRNA-Gen-spezifisch	tRNA-Gen-Abstände [bp]	Ø Abstand zu tRNA-Gen [bp]	TSD	Größe TSD [bp]
DGLT-A	41	0,132	12	5708	5	312	3'	26-97	58	ja	5-7
DIRS	9,5	0,031	7	≈2300	-	-	nein	-	-	nein	-
Skipper	17	0,055	5	≈5120	-	-	nein	-	-	nein	-
DGLT-P	25	0,081	7	4-6,5 kb	-	-	nein	-	-	nein	-
TRE5-A.1	46	0,148	55	2587	-	-	5'	44-88	59	ja	4-16
TRE5-A.2	17,5	0,056	20	890/1275	-	-	5'	31-98	59	ja	7-17
TRE5-A.3	8,5	0,027	7	2941	-	-	5'	39-67	55	ja	7-10
TRE5-B	15	0,048	26	1534	-	-	5'	25-90	57	?	6/11
TRE3	8	0,026	16	1867	-	-	3'	29-154	79	ja	9-14
LTR-Retrotransposons											
Non-LTR-Retrotransposons											

3.1.4 Phylogenie der Retrotransposons aus sozialen Amöben

3.1.4.1 Phylogenie der LTR-Retrotransposons

Durch die Identifizierung der unterschiedlichsten Retrotransposons aus den Genomen der sozialen Amöben wurde die Grundlage für die Durchführung phylogenetischer Analysen geschaffen. Dies diente zum einen zur Überprüfung der Zuordnung zu den jeweiligen Familien, welche zunächst auf Grund von strukturellen Merkmalen erfolgte, und zum anderen um Rückschlüsse auf die Entwicklung der tRNA-Gen-abhängigen Integrationspezifität einzelner mobiler genetischer Elemente zu ziehen. Für die Erstellung eines Stammbaums der LTR-Retrotransposons wurden die Aminosäuresequenzen der RNase-H und der Reversen Transkriptase-Domänen verwendet. Somit konnten nur die Elemente zur Berechnung des Stammbaums genutzt werden, bei denen die entsprechenden Domänen im Rahmen der Genomanalyse identifiziert werden konnten.

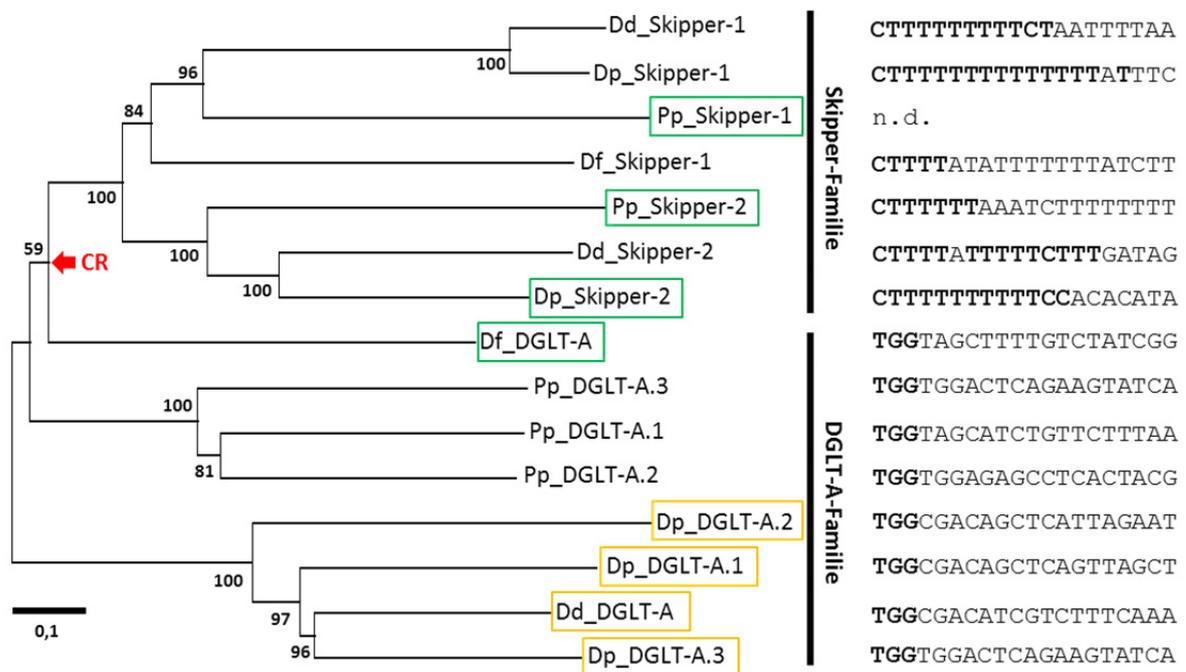


Abb. 24: Phylogenie der neu identifizierten LTR-Retrotransposons sozialer Amöben.

Ungewurzelter Stammbaum, berechnet mit Hilfe der „*maximum-likelihood*“-Methode auf der Grundlage von Aminosäuresequenzen der RT- und RNase-H-Domänen mit Hilfe des Programms MEGA5 (Tamura *et al.* 2011). Die Bootstrap-Werte an den Verzweigungen beziehen sich auf 100 Replikate. Die LTR-Retrotransposons der sozialen Amöben lassen sich entweder der DGLT-A- oder der Skipper-Familie zuordnen. Grüne Boxen markieren LTR-Retrotransposons, welche unterhalb von tRNA-Genen integrieren. Gelbe Boxen zeigen Elemente, die oberhalb von tRNA-Genen gefunden wurden. Der rote Pfeil bezeichnet die Position, ab der alle Elemente eine Chromodomäne besitzen. Rechts neben dem Stammbaum sind die Sequenzen der Polypyrimidin-Regionen der Skipper-Familie und die Primerbindestellen der DGLT-A-Elemente dargestellt.

Wie in Abbildung 24 zu erkennen ist, können alle neu identifizierten LTR-Retrotransposons, neben ihren strukturellen Eigenschaften, auch phylogenetisch in zwei Familien eingeteilt werden. Der Stammbaum ist im Einklang mit strukturellen Unterschieden der beiden Familien, was neben dem Besitz einer Chromodomäne vor allem die Unterschiede beim Aufbau der Polypyrimidin-Regionen und der Primerbindestellen betrifft.

Eine Besonderheit stellt jedoch das DGLT-A-Element aus *D. fasciculatum* dar. Strukturell gleicht es der DGLT-A-Familie, vor allem durch das Vorhandensein einer Polypyrimidin-Region (Abbildung 24). Allerdings besitzt es ebenfalls eine Chromodomäne, welche ein Kennzeichen der Skipper-Retrotransposons ist. Zudem integriert *Df_DGLT-A* unterhalb von tRNA-Genen, wohingegen alle anderen spezifisch integrierenden Mitglieder der DGLT-A-Familie oberhalb von tRNA-Genen zu finden sind. Es gruppiert sich im Stammbaum auch nicht mit den anderen DGLT-A-LTR-Retrotransposons, sondern scheint eine Schwestergruppe aller Skipper-Elemente zu sein.

Es kann daher zwei mögliche Erklärungen für die Struktur der DGLT-A- und Skipper-Elemente im Vergleich mit dem LTR-Retrotransposon *Df_DGLT-A* geben: Zum einen ist es möglich, dass im letzten gemeinsamen Vorläufer der sozialen Amöben nur LTR-Retrotransposons existierten, die Chromodomänen besaßen und diese im Lauf der Evolution von den DGLT-A-Elementen verloren wurde. Eine weitere Möglichkeit besteht in der unabhängigen Akquirierung einer Chromodomäne durch das DGLT-A-Element aus *D. fasciculatum*. Um diese Frage klären zu können, müssten ähnliche Elemente aus Spezies in relativ naher Verwandtschaft, aber außerhalb der Gruppe der Dictyosteliden, identifiziert werden. Bisher sind solche Elemente jedoch nicht bekannt.

3.1.4.2 Phylogenie der Non-LTR-Retrotransposons

Für die Erstellung eines Stammbaums der Non-LTR-Retrotransposons eigneten sich nur die Aminosäuresequenzen der Reversen Transkriptase-Domänen. Die weiteren Domänen zeigen eine sehr schwache Konservierung und sind damit ungeeignet für phylogenetische Untersuchungen.

Alle Elemente der TRE5- und TRE3-Familie ordnen sich im Stammbaum der Non-LTR-Retrotransposons in die Gruppe der LINE-1-Retrotransposons ein. Sie bilden zwei getrennte Taxa, die sich an der jeweiligen Integrationsspezifität orientiert (Abbildung 25). Die Zugehörigkeit der TRE-Elemente zur Gruppe der LINE-1-Elemente wurde bereits 1999 durch Malik *et al.* gezeigt (Malik *et al.* 1999), welche auch in dem hier dargestellten Stammbaum mit entsprechend hohen Bootstrap-Werten gut wiedergegeben wird. Dahingegen zählen die beiden TRE5-ähnlichen Elemente aus *D. purpureum* und *P. pallidum*, welche beide eine zusätzliche RNase-H-Domäne besitzen, offensichtlich nicht zur Gruppe der orthodoxen TRE-Elemente. Der Bootstrap-Wert von 70 an der Verzweigung von LINE-Elementen und anderen Klassen von Non-LTR-Retrotransposons, bei denen sich auch diese beiden Elemente befinden, scheint die Annahme zu unterstützen, dass es sich um eine weitere Gruppe von mobilen genetischen Elementen handelt, die sich an ein sehr kompaktes Genom

angepasst haben. Zudem scheinen die beiden TRE5-ähnlichen Elemente eng miteinander verwandt zu sein.

Das einzige bislang bekannte nicht tRNA-Gen-spezifische Non-LTR-Retrotransposon der sozialen Amöben (*Pp*_Non-LTR-1) scheint phylogenetisch eher mit diesen beiden mobilen genetischen Elementen verwandt zu sein als mit den TRE-Elementen.

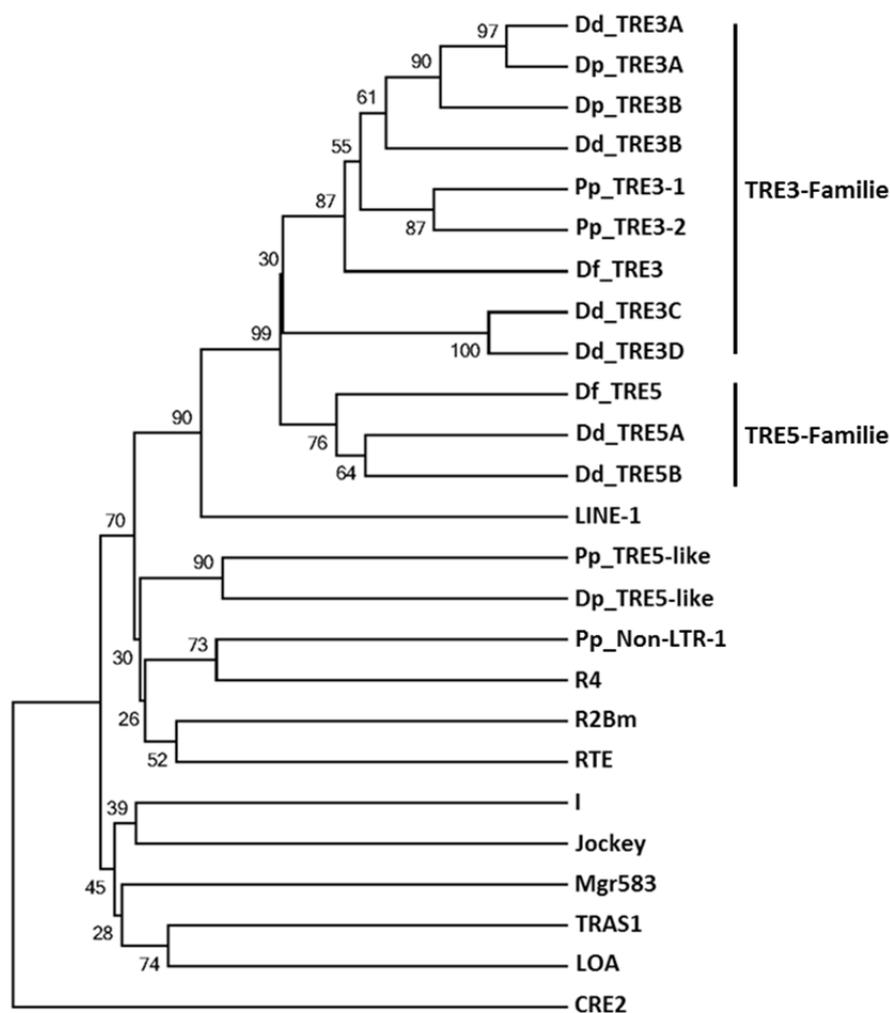


Abb. 25: Stammbaum der Non-LTR-Retrotransposons sozialer Amöben.

Basierend auf Sequenzen der Reversen Transkriptase. Erstellt mit der „Neighbor-joining“-Methode mit Hilfe des Programms MEGA (Tamura *et al.* 2011). Als Wurzel wurde das Element CRE2 verwendet. Die Bootstrap-Werte beziehen sich auf 100 Replikate. Die TRE-Elemente sind eng verwandt mit dem humanen LINE-Element, wohingegen die drei anderen Non-LTR-Retrotransposons *Pp*_TRE5-like, *Dp*_TRE5-like und *Pp*_Non-LTR-1 Verwandtschaften mit anderen Klassen mobiler genetischer Elemente zeigen. NCBI-Accession-Nummern: CRE2 [GI:1362544]; R2Bm [GI:903695]; R4 [GI:2146863]; LINE-1 [GI:74705367]; RTE [GI:1587247]; Mgr583 [GI:2454622]; TRAS1 [GI:1549144]; LOA [GI:229358279]; I [GI:51950578]; Jockey [GI:134083].

3.2 Experimenteller Teil

3.2.1 Genetisch markierte LTR-Retrotransposons

Die experimentelle Forschung an spezifisch integrierenden mobilen genetischen Elementen der sozialen Amöben beschränkt sich bislang auf die Non-LTR-Retrotransposons aus dem Genom von *D. discoideum*. Mit der Generierung eines genetisch markierten, nicht-autonomen TRE5-A.1-Elements aus *Dictyostelium discoideum* wurde eine Möglichkeit geschaffen, *in vivo* die Retrotransposition eines Non-LTR-Retrotransposons zu beobachten und zu analysieren (Siol *et al.* 2011). Die Funktionsweise dieses Systems basiert auf einer invertierten Resistenzkassette, welche durch ein Intron unterbrochen wird und von einem eigenen, ebenfalls invertierten Promotor transkribiert wird (Sassaman *et al.* 1997). Das Intron ist in Bezug auf das artifizielle Retrotransposon vorwärts orientiert. Bei einer Mobilisierung wird aus dem RNA-Intermediat das Intron gespleißt und nach der erfolgreichen reversen Transkription und der anschließenden Integration des Elements entsteht eine funktionelle Resistenzkassette ohne Intron. Eine Selektion auf resistente Zellen ergibt in der Zellkultur einen Pool von Klonen, die neue Sprungereignisse beherbergen und anschließend analysiert werden können (vgl. Abschnitt 1.9).

Mit dem DGLT-A-Element existiert auch in *D. discoideum* eine Population von LTR-Retrotransposons, welche eine tRNA-Gen-abhängige Integrationspezifität besitzen (Winckler 1998b; Glöckner *et al.* 2001). Bislang sind allerdings alle Versuche gescheitert, eine Aktivität eines genetisch markierten DGLT-A-Retrotransposons in einer Zellkultur nachzuweisen (Dissertation Jana Schiefner, 2014). Eine Ursache dafür könnte in der Tatsache bestehen, dass bislang keine Kopie eines vollständigen, potentiell aktiven, endogenen DGLT-A-Elements im Genom von *D. discoideum* bekannt ist. Daher mussten die Sequenzen auf der Grundlage hypothetischer Konsensussequenzen nachträglich repariert werden, um einen vollständigen ORF zu erhalten. Somit ist es möglich, dass dadurch kein funktionsfähiger ORF generiert werden konnte und das Element aus diesem Grund keine Aktivität in einer Zellkultur gezeigt hat (vgl. 3.2.1.2).

Anders als bei DGLT-A existieren von dem LTR-Retrotransposon Skipper mehrere vollständige genomische Kopien im Genom von *D. discoideum*, welche potentiell aktiv sein können (Leng *et al.* 1998). Da LTR-Retrotransposons einen wesentlich komplexeren Mechanismus zur Erzeugung neuer genomischer Kopien verwenden als Non-LTR-Retrotransposons, bestand vor allem die Frage nach der möglichen Position für die Integration einer Blasticidin-Kassette. Ziel war es daher, zunächst ein Skipper-Element zu isolieren und anschließend mit einem genetischen Marker auszustatten, der es erlaubt, die Retrotransposition dieses Elements in einer Zellkultur nachzuweisen. Sollte ein retrotranspositions kompetentes Skipper generierbar sein und somit ein "*Proof-of-concept*" liefern, sollten auf dieser Grundlage anschließend weitere Experimente mit DGLT-A folgen, die sich mit der tRNA-Gen-spezifischen Integration von LTR-Retrotransposons auseinandersetzen sollten.

3.2.1.1 Das Skipper_*msbri*-Element

Für die Isolation eines Skipper-Retrotransposons wurde ein Element auf Chromosom 1 ausgewählt (dictybase Gene ID: DDB_G0270954). In einem ersten Schritt wurde an genomischer DNA von AX2 eine PCR mit den Primern Skipper-01 und Skipper-02_AgeI durchgeführt, um das 5'-gelegene LTR mit beiden ORFs zu amplifizieren. Dieses Fragment wurde in den Vektor pCR/XL/TOPO kloniert. Im nächsten Schritt wurde die *msbri*-Kassette als AgeI/PstI-Fragment aus dem TRE5-A^{bsr}-Retrotransposon unter Verwendung der internen PstI-Schnittstelle des pCR/XL/TOPO-Vektors in das zuvor erzeugte Plasmid kloniert. Der 3'-LTR wurde mit den Primern Skipper-03 und Skipper-04 von genomischer DNA amplifiziert und in den Vektor pGEM-T kloniert. Zum Schluss wurde der 3'-LTR als PstI-Fragment in den pCR/XL/TOPO-Vektor kloniert. Dadurch wurde ein vollständiges Skipper-Retrotransposon mit einer invertierten BSR-Kassette generiert (Abbildung 26).

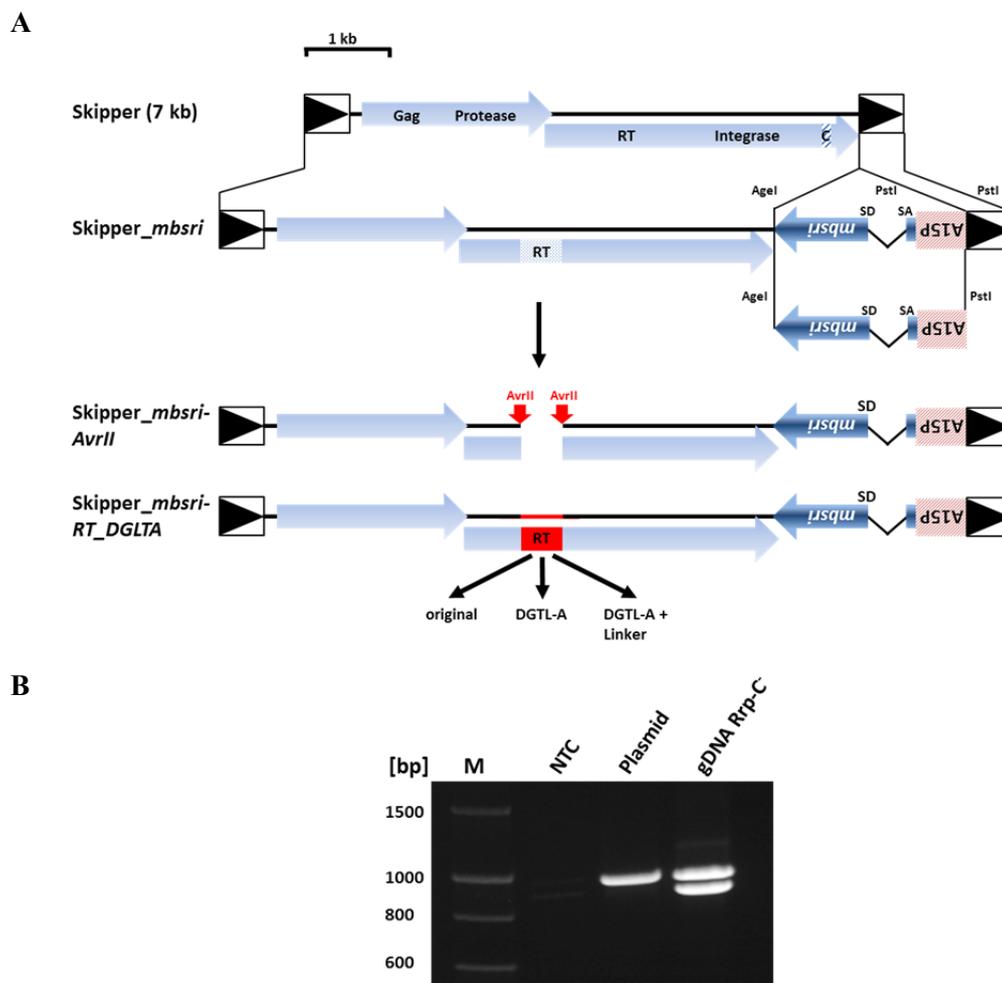


Abb. 26: Das genetisch markierte Skipper-Retrotransposon.

(A) Erzeugung genetisch markierter Skipper-Retrotransposons und der Austausch der RNase-H-Domäne in einem Skipper_*msbri*-Element. durch Einfügen von AvrII-Restriktionsschnittstellen. Es wurden anschließend drei unterschiedliche Elemente erzeugt (vgl. 3.2.1.3). (B) PCR zur Überprüfung der BSR-resistenten Zellen auf Sprungereignisse des genetisch markierten Skipper-Retrotransposons. Die Sequenz der unteren Bande der PCR-Produkte von genomischer DNA aus der Rrp-C-Mutante zeigte, dass das Intron aus dem genetisch markierten Skipper-Element gespleißt wurde (Daten nicht gezeigt). Verwendete Primer: Q-Skipper-05 und BSR-19.

Zunächst wurde dieses Retrotransposon in AX2-Zellen transformiert und auf BSR-resistente Klone selektioniert. Allerdings konnten keine Blastocidin-resistenten Klone beobachtet werden. Eine mögliche Ursache dafür besteht in der Möglichkeit, dass zelluläre Abwehrmechanismen die Retrotransposition des Skipper-Retrotransposons verhindern (vgl. Abschnitt 1.5.1).

In der Arbeitsgruppe von Prof. Dr. Christian Hamann (Jacobs University Bremen) wurden zuvor „RNA-Interferenz“-Mutanten generiert, welche durch Knock-Outs verschiedener RNA-abhängiger RNA-Polymerasen (Rrp; engl.: „*RNA-dependent RNA-polymerase*“) erzeugt wurden (Wiegand *et al.* 2013). Das genetisch markierte Skipper-Element wurde daher in diese Zelllinien transformiert und anschließend auf Sprungereignisse selektioniert. Dabei konnte gezeigt werden, dass das genetisch markierte Skipper-Retrotransposon aktiv in Mutanten mobilisiert wird, welche einen „Knock-Out“ der Rrp-C im Genom beherbergen (siehe Abbildung 26 B). Daher wurden alle weiteren Experimente in einem Stamm mit einem „Doppel-Knock-Out“ der Gene für die Rrp-C und Rrp-A durchgeführt.

3.2.1.2 Das DGLT-A_ *mbsri*-Element

Das LTR-Retrotransposon DGLT-A aus *D. discoideum* wurde mit der *mbsri*-Kassette genetisch markiert, um in einer Zellkultur anschließend auf neue Sprungereignisse selektionieren zu können. Die Retrotranspositionskassette wurde dabei an der gleichen Stelle integriert, zwischen dem Ende des ORFs und dem Beginn des 3'-LTRs, wie bei dem aktiv mobilisierten Skipper-*mbsri*-Element. Allerdings zeigte das DGLT-A-Retrotransposon weder in einer AX2-Zellkultur oder in einer der verfügbaren Rrp-Mutanten eine aktive Retrotransposition. Auch eine Klonierung des Elements in einen Expressionsvektor, der durch einen zusätzlichen Promotor die Transkription des Elements verbessern könnte, blieb ohne Erfolg. Es ist daher unklar, warum eine Aktivierung des DGLT-A-Retrotransposons bislang erfolglos blieb, jedoch scheint die Position der Retrotranspositionskassette als Ursache ausgeschlossen werden zu können.

3.2.1.3 Austausch der RNase-H-Domäne im genetisch markierten Skipper-Element

Nachdem gezeigt wurde, dass das *mbsri*-markierte Skipper-Retrotransposon in einer Zellkultur mobilisiert wird, sollte anschließend die RNase-H-Domäne des Elements durch die RNase-H-Domäne des DGLT-A-Retrotransposons ersetzt werden. In der Dissertation von Jana Schiefner (2014) konnte zuvor in Hefe-Zwei-Hybrid-Experimenten gezeigt werden, dass eine Interaktion der RNase-H-Domäne von DGLT-A mit der Untereinheit TFC4 des Transkriptionsfaktors TFIIC wahrscheinlich an der Erkennung von tRNA-Genen als Integrationsziele beteiligt ist. Da kein aktives, genetisch markiertes DGLT-A zur In-vivo-Untersuchung dieser Interaktion verfügbar war, sollte daher versucht werden, diese Experimente im Hintergrund des Skipper-Elements durchzuführen.

Zur Erzeugung chimärer Skipper-Elemente mussten in einem ersten Schritt zwei PCR-Fragmente erzeugt werden, welche an den äußeren Grenzen der natürlichen RNH-Domäne von Skipper

Schnittstellen für das Restriktionsenzym AvrII besitzen (Abbildung 26 A). Für die Amplifikation der Fragmente wurden die Primer Skipper-01/Skipper-AvrII_NotI und Skipper-AvrII_for/Skipper-NotI_rev verwendet. Nachdem beide Fragmente in den pCR8™/GW/TOPO®-Vektor kloniert wurden (Skipper_mbsri_AvrII), konnte in die AvrII-Schnittstelle die RNase-H-Domäne von DGLT-A eingefügt werden, die als AvrII-Fragment zuvor amplifiziert und subkloniert wurde (Primer DGLT-A_AvrII-01/-02). Neben der originalen RNase-H-Domäne von DGLT-A wurde durch Gensynthese ein weiteres Fragment erzeugt, welches neben der RNase-H-Domäne von DGLT-A eine Linkersequenz der Integrasedomäne von Skipper besitzt (im Folgenden als chimäre RNase-H-Domäne genannt). Dieses Fragment wurde erzeugt, da es Hinweise auf eine zusätzliche Beteiligung dieser Sequenz an einer spezifischen Integration eines DGLT-A-Retrotransposons gab (Dissertation Jana Schiefner, 2014). Zur Kontrolle, ob die AvrII-Schnittstellen unter Umständen eine erfolgreiche Retrotransposition verhindert, wurde auch die originale RNase-H-Domäne von Skipper als AvrII-Produkt amplifiziert und in das Skipper_mbsri_AvrII-Element kloniert (Abbildung 26 A; Primer: DGLTA-AvrII-03/-06). Die Retrotransposons wurden anschließend auf ihre Aktivität in einer RrpAC⁻-Zellkultur getestet.

Bei den Retrotransposons mit der originalen RNase-H-Domäne von Skipper sowie dem Element mit der chimären RT-Domäne von DGLT-A und der Integrasedomäne von Skipper konnten BSR-resistente Klone beobachtet werden. Eine anschließende PCR an genomischer DNA zeigte, dass das Intron gespleißt wurde und somit von einer aktiven Retrotransposition dieses Element ausgegangen werden kann (Daten nicht gezeigt). Warum es nicht möglich war, BSR-resistente Klone des Skipper-Retrotransposons, welches ausschließlich die RNase-H-Domäne von DGLT-A besitzt, zu erhalten, konnte nicht geklärt werden.

Eine folgende LAM-PCR (vgl. Abschnitt 3.2.2) an genomischer DNA von BSR-resistenten Klonen des Skipper-Elementes mit der chimären RNase-H-Domäne erbrachte keinen Hinweis auf eine tRNA-Gen-spezifische Integration (Daten nicht gezeigt). Das deutet darauf hin, dass neben der RNase-H-Domäne eventuell noch andere Proteine des Retrotransposons an einer tRNA-Gen-spezifischen Integration beteiligt sind.

3.2.1.4 Ein *mbsri*-markiertes LTR-Element aus dem Genom von *D. fasciculatum*

Nach der Identifizierung eines DGLT-A-Elements im Genom von *D. fasciculatum*, welches sowohl Merkmale von DGLT-A- als auch Skipper-Elementen aufweist (vgl. Abschnitt 3.1.1.3.) und tRNA-Gen-spezifisch integriert, schien es lohnenswert dieses LTR-Retrotransposon zu isolieren und ebenfalls mit einem genetischen Marker auszustatten. Da die Transformation von *D. fasciculatum* nicht etabliert war und die Zellen zudem eine natürliche Blasticidin-Resistenz aufweisen, war das Ziel der Arbeiten zunächst, Integrationen des *Df_DGLT-A_mbsri*-Retrotransposons an tRNA-Genen im Genom von *D. discoideum* nachzuweisen. Dies würde zudem ein Beweis für die Konservierung der Integrationsmechanismen zwischen phylogenetisch weit entfernten Spezies darstellen.

Zur Isolierung des LTR-Retrotransposons wurde zunächst aus einer Zellkultur von *D. fasciculatum* genomische DNA gewonnen. Durch die Genomanalyse wurden sieben potentiell intakte Kopien des Elements identifiziert, welche als Matrize für eine Amplifikation in Frage kamen. In einem ersten Schritt wurde eines dieser Elemente mit den flankierenden Primern Dfas-DGLTA-12 und Dfas-DGLTA-15 amplifiziert und in den Vektor pGEM-T kloniert (pGEMT-DGLT-A_Dfas). Dieses Element wurde anschließend in RrpAC⁻-Zellen transformiert und mit einer RT-PCR auf eine transkriptionelle Aktivität überprüft (Abbildung 27). So konnte gezeigt werden, dass das Retrotransposon in RrpAC⁻-Zellen über die komplette Länge exprimiert wurde. Des Weiteren wurde versucht, mit Southern-Blots mögliche Sprungereignisse nachzuweisen (Daten nicht gezeigt). Allerdings scheiterten diese Versuche vermutlich daran, weil es sich dabei um „somatische“ Ereignisse handelte, die nur sehr selten vorkamen und daher innerhalb der Gesamtpopulation extrem schwach repräsentiert waren. Durch die Verwendung eines genetisch nicht markierten Elements war es auch nicht möglich auf diese Ereignisse zu selektionieren. Nachweisbar wären solche Ereignisse nur, wenn eine neue Retrotransposon-Kopie in der gesamten Population fixiert wird, was im Falle eines neutralen Ereignisses, innerhalb eines überschaubaren Zeitraumes, nur bedingt möglich erscheint.

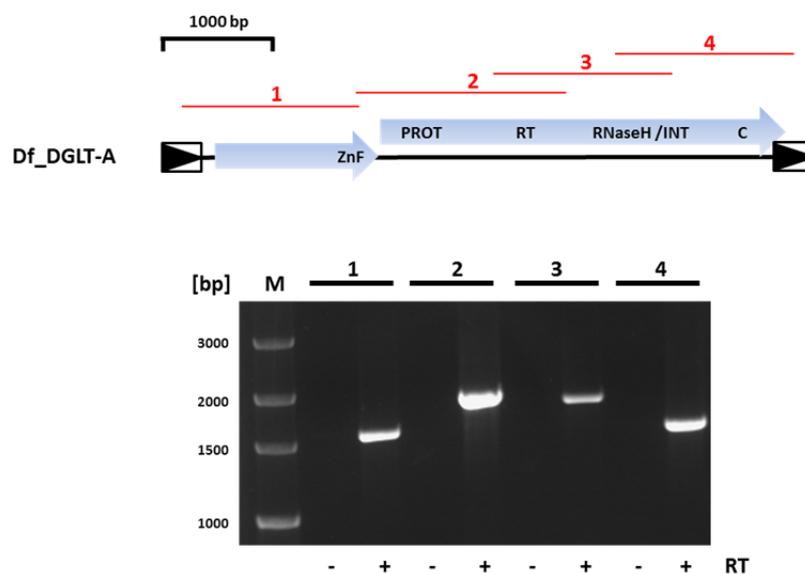


Abb. 27: Das DGLT-A-Element aus *D. fasciculatum*.

Nachweis der transkriptionellen Aktivität des Retrotransposons *Df_DGLT-A* in einer Zellkultur von *D. discoideum* (Rrp-AC⁻). PCR an unspezifisch erzeugter cDNA. Verwendete Primer: Produkt 1: Dfas-DGLTA-27/-25; 2: Dfas-DGLTA-19/-22; 3: Dfas-DGLTA-21/-20; 4: Dfas-DGLTA-4/-32.

Eine sensitivere Suche nach Integrationen des *Df*_DGLT-A-Elementes wäre möglich, wenn das Element mit dem *mbsri*-Selektionsmarker ausgestattet wäre. Für die Erzeugung eines Blasticidin-markierten Elements wurde der 5'-LTR und beide ORFs als *Sma*I/*Eag*I-Fragment mit den Primern DGLTA-39 und DGLTA-40 von dem Plasmid pGEMT-DGLT-A_Dfas amplifiziert und in den pCR-XL-TOPO-Vektor kloniert (pCR-XL-F1_*Df*_DGLTA). Dieses Fragment wurde anschließend als *Sma*I/*Eag*I-Fragment in den Vektor pBT1 kloniert (pBT-F1_*Df*_DGLT-A). Die Blasticidin-Kassette und der 3'-LTR wurden als „*overlap-extension*“-PCR-Produkt mit flankierenden *Eag*I-Schnittstellen amplifiziert und mit Hilfe des Restriktionsenzym *Eag*I in den Vektor pBT-F1_*Df*_DGLT-A kloniert (Primer: *mbsri*-01/-02 und *Df*-DGLTA-35/-36). Dadurch entstand ein *mbsri*-markiertes DGLT-A-Element aus *D. fasciculatum* (pBT1-F1_*Df*_DGLT-A_*mbsri*). Die Retrotranspositionskassette befand sich dabei an der gleichen Position wie in dem aktiv mobilisierten Skipper-Retrotransposon.

Nachdem dieses Element sowohl in eine AX2-, als auch in eine RrpAC⁻-Zellkultur transformiert wurde, konnten nach einer anschließenden Selektion auf Blasticidin-resistente Klone allerdings keine Zellen beobachtet werden, die Sprungereignisse akkumuliert hatten. Da bekannt ist, dass mobile genetische Elemente teilweise auch entwicklungs- oder stressabhängig aktiviert werden, wurde ebenfalls versucht, das markierte *Df*_DGLT-A-Element auf diese Weise zu mobilisieren. Allerdings konnten auch nach einem Entwicklungszyklus oder Kälte- und Hitzeschocks keine Blasticidin-resistenten Klone beobachtet werden. Somit war es nicht möglich, das genetisch markierte *Df*_DGLT-A-Retrotransposon in einer *Dictyostelium*-Zellkultur zu mobilisieren.

3.2.2 Genomweite Sequenzierung von De-novo-Integrationen eines Non-LTR-Retrotransposons

Mit der Erzeugung eines nicht-autonomen TRE5-A.1-Retrotransposons war es erstmals möglich De-novo-Integrationen eines artifiziellen mobilen genetischen Elements direkt im Genom von *Dictyostelium discoideum* nachzuweisen und zu analysieren (Siol *et al.* 2011). Bereits zuvor konnten mit der Etablierung der „TRE-Falle“ De-novo-Integrationen endogener TRE-Elemente an tRNA-Genen nachgewiesen werden (Beck *et al.* 2002). Bei beiden Studien beruht die Identifizierung der Integrationsereignisse auf PCR-basierten Methoden, bei denen nur relativ wenige, einzelne Ereignisse mit einem verhältnismäßig großen Aufwand analysiert werden können. Daher war ein Ziel dieser Arbeit, ein Protokoll zu entwickeln, das es möglich macht, eine größtmögliche Anzahl unabhängiger Integrationsereignisse eines tRNA-Gen-spezifischen Retrotransposons im Rahmen eines einzigen Experiments zu sequenzieren und damit analysieren zu können. Ein wesentlicher Aspekt war dabei, die Möglichkeit zu schaffen, auch Integrationsereignisse zu detektieren die womöglich nicht an orthodoxen Integrationsorten, also den tRNA-Genen, stattfinden.

Eine weitere Frage ergibt sich bei einem Blick auf das Referenzgenom. Hier finden sich bei insgesamt mehr als 400 tRNA-Genen im gesamten Genom von *D. discoideum*, lediglich rund 100 unterschiedliche Integrationsorte von TRE5-A-Retrotransposons (Daten nicht gezeigt). Daher sollte im Zuge dieser Untersuchungen ebenfalls geklärt werden, ob grundsätzlich alle Gene, die von der RNA-Polymerase-III transkribiert werden, als Integrationsorte von TRE5-A-Elementen erkannt werden.

3.2.2.1 Das Codon-adaptierte nicht-autonome TRE5-A_ *msbri*-Element

Eine Voraussetzung für die Entwicklung eines Systems zur genomweiten Kartierung unabhängiger Integrationsereignisse bestand darin, ein artifizielles TRE5-A-Retrotransposon zu verwenden, das von der endogenen TRE5-Population unterschieden werden kann. Das bereits vorhandene, genetisch markierte, nicht-autonome Retrotransposon TRE5-A_ *msbri* (A6P-ORF1-*msbri*-C; Siol *et al.* 2011) sollte daher zunächst so modifiziert werden, dass es diese Voraussetzungen erfüllt.

Ein Ansatz stellt hier die Veränderung der ORF1-codierenden DNA-Sequenzen dar. Da der genetische Code degeneriert ist, besteht die Möglichkeit die ORF1-Sequenz so zu verändern, dass sie sich vom Original unterscheidet, aber die gleiche Aminosäuresequenz codiert. Die originale ORF1-Sequenz der endogenen TRE5-Elemente zeigt zudem einen verhältnismäßig schlechten Codon-Gebrauch. Die veränderte und an den optimalen Codon-Gebrauch von *D. discoideum* angepasste ORF1-DNA-Sequenz von TRE5-A weicht zu 19 % von der originalen Sequenz ab und zeigt an den meisten Positionen einen verbesserten Codon-Gebrauch (Abbildung 28 A). Das veränderte ORF1-Gen wurde synthetisch erzeugt (Vektor pJ1.2-ORF1_opt_Dd) und anschließend im Vektor A6P-ORF1-*msbri*-C als HindIII-/AgeI-Fragment mit der ursprünglichen ORF1-Sequenz ausgetauscht. Das resultierende, nicht-autonome Retrotransposon A6P-ORF1_Dd/adapt-*msbri*-C (im Folgenden als TRE5-A^{bsr} bezeichnet) wurde anschließend in einer AX2-Zellkultur auf seine Aktivität überprüft. Dazu wurden die Blasticidin-resistenten Klone vereinigt und daraus genomische DNA isoliert. An dieser genomischen

DNA wurde eine Kontroll-PCR mit den Primern A15P-seq und BSR-01 durchgeführt (Abbildung 28 B). Die daraus resultierende Doppelbande spricht für eine Mobilisierung des Retrotransposons. Die obere Bande entspricht dabei den transformierten, nicht mobilisierten genomischen Kopien („Masterelemente“), welche das Intron besitzen. Die untere Bande repräsentiert die erfolgreich transponierten Elemente ohne Intron. Die Sequenzierung des kleineren PCR-Produkts zeigte, dass das Intron gespleißt wurde (Daten nicht gezeigt).

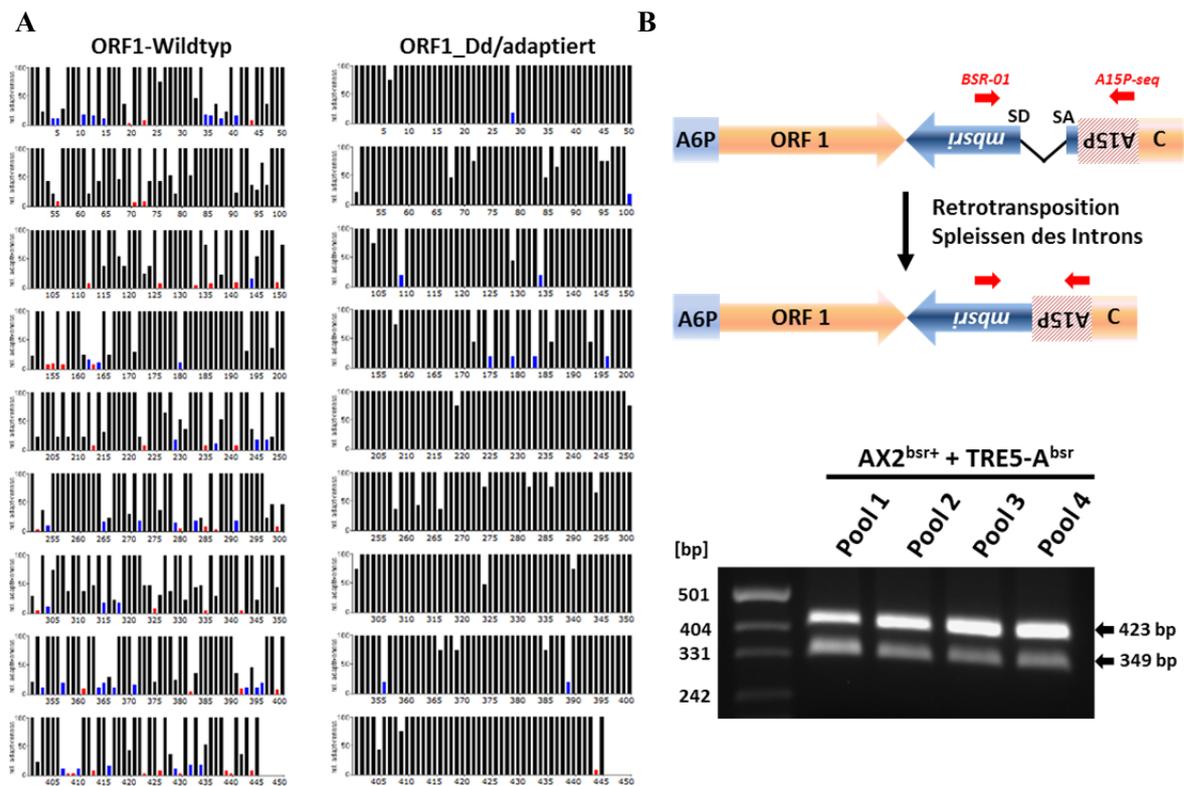


Abb. 28: Das Codon-adaptierte, nicht autonome TRE5-A^{bsr}-Element.

(A) Anpassung der ORF1-Sequenz des Retrotransposons TRE5-A. Dargestellt sind die Spezies-spezifischen Nutzungsfrequenzen der jeweiligen Codons. Links ist die Originalsequenz, rechts die angepasste ORF1-Sequenz abgebildet. Die Abbildungen wurden mit dem GCUA-Online-Tool (Fuhrmann *et al.* 2004), unter Berücksichtigung der von *D. discoideum* bevorzugten Triplets, erstellt. Die authentische ORF1-Sequenz besitzt eine schlechte „Codon-Usage“ im Vergleich mit der adaptierten Sequenz. (B) PCR zur Überprüfung der Aktivität des Codon-adaptierten TRE5-A^{bsr}-Elements. Die untere Bande mit einer Größe von 349 bp repräsentiert die gesprungenen Kopien des artifiziiellen Retrotransposons. Die PCR wurde an genomischer DNA von Pools BSR-resistenter Zellen mit den Primern A15P-seq und BSR-01 durchgeführt.

3.2.2.2 Überlegungen zur genomweiten Kartierung von De-novo-Integrationsereignissen

In den letzten Jahren wurden mit der Entwicklung der modernen Hochdurchsatz-Sequenzierungsmethoden zahlreiche Protokolle verfügbar, welche zur Identifizierung und Analyse von Integrationen mobiler genetischer Elemente verwendet werden können (Schmidt *et al.* 2007;

Paruzynski *et al.* 2010; Qi *et al.* 2012). Viele dieser Protokolle beziehen sich auf die Detektion der Integrationsstellen von LTR-Retrotransposons. Der Vorteil bei der Arbeit mit LTR-Retrotransposons ist dabei die Konstanz der 5'- und 3'-Enden, bedingt durch den von LTR-Retrotransposons verwendeten Retrotranspositionsmechanismus (vgl. Abschnitt 1.3.1.1). Möchte man die Integrationsstellen von Non-LTR-Retrotransposons bestimmen, stößt man relativ schnell auf das Problem der vorzeitigen Abbrüche der Reversen Transkription, was variable 5'-Enden der Elemente erzeugt (George *et al.* 1996; Sassaman *et al.* 1997). Daher hat man es unter Umständen, bezogen auf das 5'-Ende, mit einer sehr inhomogenen Population von Elementen zu tun. Die Alternative, über das konstante 3'-Ende Integrationsstellen zu bestimmen, hat allerdings auch einige Nachteile. Zum einen besitzen die Non-LTR-Elemente Poly-A-Regionen mit variabler Länge, die bei einer Sequenzierung zu Problemen führen können. Des Weiteren ist speziell bei den TRE-Elementen die Wahrscheinlichkeit groß, dass an einer orthodoxen Integrationsstelle bereits weitere genomische Kopien der endogenen TRE-Familien integriert sind. Da neu integrierende Elemente jeweils vor bereits vorhandene Elemente springen würden, wäre eine Kartierung des angesprungenen Locus, bedingt durch die Redundanz der Elemente, vermutlich schwierig. Da viele der tRNA-Gene in *D. discoideum* mit TRE-Retrotransposons assoziiert sind, ist die Wahrscheinlichkeit einer solchen, nicht-auswertbaren Integration als relativ hoch einzustufen. Aus diesen Gründen wurde die Strategie, für die Etablierung einer Methode zur genomweiten Kartierung von De-novo-Integrationen, auf die 5'-Enden eines in ORF1 Codon-adaptierten, genetisch markierten TRE5-A_{mbstri}-Elements ausgelegt. Als Ausgangsmaterial diente die genomische DNA eines Pools von Zellen, welcher rund 15.000 Blasticidin-resistente Klone repräsentierte und damit eine größtmögliche Abdeckung aller potentiellen Integrationsorte gewährleisten sollte.

3.2.2.3 Bestimmung der Kopienzahl gesprungener Elemente

Zunächst sollte ermittelt werden, wie viele Integrationsereignisse innerhalb einer einzelnen Zelle akkumulieren. Dazu wurde von individuellen Klonen genomische DNA gewonnen, welche für ein Southern-Blot-Experiment zur Bestimmung der Kopienzahl der mobilisierten TRE5-A^{bsr}-Retrotransposons verwendet wurde. Die genomische DNA wurde mit den Restriktionsenzymen AgeI und NsiI verdaut, wodurch die Retrotranspositionskassette und das C-Modul mit einem Teil des Vektors aus den Masterkopien geschnitten wurden. Die Detektion der artifiziellen Retrotransposons erfolgte mit einer biotinylierten Sonde, welche durch eine PCR am *mbstri*-Gen gewonnen wurde (Primer: BSR-TEST-01/BSR-TEST-02). In Abbildung 29 sind die Masterelemente nach AgeI-/NsiI-Verdau im Genom der jeweiligen Klone als Fragmente von rund 1,9 kb Größe zu erkennen, was der zu erwartenden Größe der Fragmente bei einem Verdau des im Genom verankerten Expressionsvektors entspricht. Des Weiteren ließen sich schwächere Banden erkennen, die je nach Klon unterschiedlich groß waren. Diese Fragmente entsprechen den gesprungenen Kopien, da die Restriktionsschnittstellen in den flankierenden Regionen der Integrationen in unterschiedlichen Distanzen vorkommen. Somit kann davon ausgegangen werden, dass jeder Klon eine bis drei Integrationen des artifiziellen

Retrotransposons im Genom beherbergt. Der in den folgenden Experimenten verwendete Pool von rund 15.000 Blastocidin-resistenten Klonen repräsentiert daher eine größere Anzahl von unabhängigen Integrationsereignissen, als es die reine Anzahl eingesetzter Klone vermuten lässt.

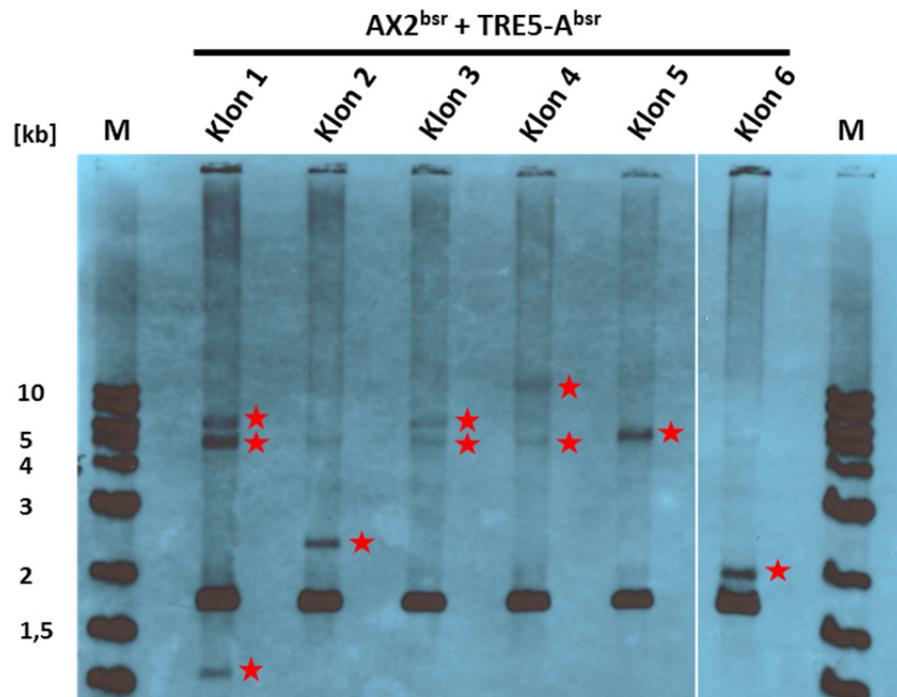


Abb. 29: Southern Blot zur Bestimmung der Kopienzahl von Integrationsereignissen.

Genomische DNA von BSR-resistenten Einzelklonen wurde mit den Restriktionsenzymen NsiI und AgeI verdaut. Die Detektion der genomischen Kopien erfolgte mit einer BSR-spezifischen, biotinylierten Sonde. Als Marker diente der „2 log-biotinylated Ladder“. Die Sterne zeigen die jeweiligen Banden, welche unterschiedliche Integrationsereignisse des Codon-adaptierten TRE5-A_ *mbsri*-Elements darstellen.

3.2.2.4 Überlegungen zu LAM-PCR-basierten Methoden

Eine Methode zur Anreicherung von unabhängigen Integrationsereignissen aus einem Pool von BSR-resistenten Zellen ist die Methode der LAM-PCR („*linear amplification mediated-PCR*“; Schmidt *et al.* 2007). Die LAM-PCR basiert auf der Erzeugung von einzelsträngigen PCR-Produkten. Anders als bei der konventionellen PCR, bei der zwei Primer eingesetzt werden, wird dabei allerdings nur ein Primer verwendet. Dadurch ist eine exponentielle Vervielfältigung der DNA nicht möglich, dementsprechend werden mehr Elongationszyklen als bei der konventionellen PCR benötigt. Als Primer wird ein 5'-biotinyliertes Oligonukleotid verwendet. Durch die Biotinylierung der LAM-PCR-Produkte können diese in einem weiteren Schritt an magnetische Streptavidin-Kügelchen gebunden werden und anschließend die unerwünschte genomische DNA abgereichert werden. Somit sollte am Ende nur die gekoppelte, durch die LAM-PCR generierte, DNA vorliegen. Ein Nachteil bereits etablierter Methoden zur Analyse von Integrationsereignissen ist die Voraussetzung, dass in den

flankierenden genomischen Regionen geeignete Restriktionsschnittstellen vorhanden sein müssen, um entsprechende doppelsträngige Linker zu ligieren oder eine inverse PCR durchzuführen (Schmidt *et al.* 2007; Qi *et al.* 2012). Daher sollte im Folgenden eine Methode etabliert werden, die ausgehend von einzelsträngigen LAM-PCR-Produkten, die Erzeugung von doppelsträngigen PCR-Produkten ermöglicht, ohne auf Restriktionsschnittstellen in den flankierenden genomischen Regionen angewiesen zu sein. Ziel war die Sequenzierung der PCR-Produkte mit „Next-generation“-Sequenzierungsmethoden, um eine sehr große Menge an unabhängigen Integrationsereignissen innerhalb eines Experiments analysieren zu können.

3.2.2.5 Etablierung der LAM-PCR zur Anreicherung von Integrationsereignissen

Zunächst sollte versucht werden, die LAM-PCR in Verbindung mit dem Codon-adaptierten TRE5-A^{bsf}-Retrotransposon zu etablieren. So sollte gezeigt werden, dass eine effektive Reduzierung unerwünschter genomischer DNA (inklusive endogener TRE5-A-Elemente) nach der Kopplung von 5'-biotinylierten LAM-PCR-Produkten an Streptavidin-Kügelchen erreicht werden kann.

Für die Amplifizierung wurden zwei lineare PCRs mit je 50 Zyklen und einer Elongationszeit von 45 sek durchgeführt, wobei nach der ersten PCR erneut 0,5 µl High-Fidelity-Polymerase zugegeben wurde. Als Primer wurde das 5'-biotinylierte Oligonukleotid Rep-234_{bio} verwendet, das spezifisch mit der Codon-adaptierten ORF1-Sequenz hybridisiert. Als Negativ-Kontrolle wurde die LAM-PCR ohne diesen Primer durchgeführt.

Nach einer Kopplung an Streptavidin-Kügelchen über Nacht und einem anschließenden Waschschrift mit 100 µl Wasser, welcher dreimal wiederholt wurde, konnte die gekoppelte DNA für die folgenden Test-PCRs verwendet werden. Als Kontrollen wurden Primer für das Gen der Glycerinaldehyd-3-Phosphat-Dehydrogenase (Q-GAPDH-01/Q-GAPDH-02), sowie für multiple Kopien des DIRS-1-Retrotransposons (Q-DIRS-01/Q-DIRS-02) verwendet. Für die Unterscheidung der endogenen TRE5-A-Retrotransposons von den artifiziellen Elementen wurden Primerkombinationen verwendet, die in Vorversuchen gezeigt hatten, dass sie spezifisch die unterschiedlichen ORF1-Sequenzen amplifizieren (Rep-196_454A/Rep-197 für das Codon-adaptierte artifizielle Retrotransposon und Rep-196_WT/Rep-197_WT für die endogene TRE5-A Population; Daten nicht gezeigt). Wie in Abbildung 30 zu erkennen ist, kann nach der Kopplung der LAM-PCR-Produkte lediglich das Codon-adaptierte TRE5-A^{bsf}-Retrotransposon nachgewiesen werden. Somit findet eine Abreicherung unerwünschter genomischer DNA und damit auch endogener TRE5-A-Elemente, durch Kopplung der LAM-PCR-Produkte an Streptavidin-Kügelchen statt. Die LAM-PCR ist daher eine geeignete Methode, um zunächst die DNA zu erhalten, welche mit der artifiziellen Retrotransposon-Sequenz assoziiert ist.

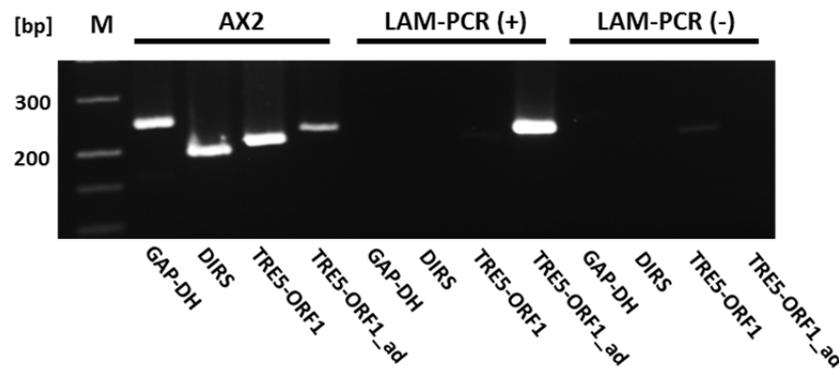


Abb. 30: Effizienz der Kopplung von LAM-PCR-Produkten an Streptavidin-Kügelchen.

Zu erkennen ist eine vollständige Eliminierung der unerwünschten genomischen DNA. Als Negativ-Kontrolle wurde die LAM-PCR ohne den biotinylierten Primer durchgeführt (rechts).

3.2.2.6 Charakterisierung der Vollständigkeit mobilisierter Retrotransposons

Ein Problem der „Next-generation“-Methoden sind die relativ kurzen Längen der erzeugten Sequenzen von wenigen hundert Basenpaaren (Mardis 2013). Daher schien es wichtig, die Methode dahingehend zu optimieren, dass die finalen PCR-Produkte ebenfalls möglichst klein sind. Dazu wurde zunächst untersucht inwieweit vorzeitige Abbrüche während der Reversen Transkription zu verkürzten Kopien des mobilisierten TRE5-A^{bsr}-Elements führen. Für die LAM-PCR wurde ein Primer verwendet (BSR-enrich-01), der innerhalb der Blastocidin-Kassette überlappend über der Position des gepleißten Introns hybridisiert und somit bevorzugt die durch Retrotransposition integrierten Kopien amplifizieren sollte. Durch die Lage des Primers sollten auch potentiell verkürzte Kopien nach erfolgreicher Integration mit amplifiziert werden. Anschließend folgten PCRs mit verschiedenen tRNA-Gen- und einem Palindrom-spezifischen Primer, sowie einem reversen Retrotransposon-spezifischen Primer, der hinter dem Ende der codierenden ORF1-Sequenz in der Blastocidin-Kassette bindet (Primer BSR-09). Wie in Abbildung 31 zu erkennen ist, sind ausschließlich einheitlich große Fragmente amplifiziert worden, die der Größe eines vollständigen Retrotransposons entsprechen. Sollte es zur Verkürzung der Kopien durch einen vorzeitigen Abbruch der reversen Transkription kommen, sollten neben diesen Fragmenten ebenfalls verkürzte PCR-Produkte zu beobachten sein. Dieses Ergebnis deutet darauf hin, dass vorzeitige Abbrüche während der reversen Transkription offensichtlich sehr selten stattfinden. Somit konnte im weiteren Verlauf bei der LAM-PCR ein Primer verwendet werden, der nahe des 5`-Endes des Retrotransposons hybridisiert und dennoch mit einer hohen Wahrscheinlichkeit alle Integrationsereignisse amplifiziert. Für die folgenden Experimente wurde daher der Primer Rep-234_bio für die Erzeugung von LAM-PCR-Produkten verwendet, der 220 bp unterhalb des Start-Codons vom Codon-adaptierten ORF1-Gen hybridisiert.

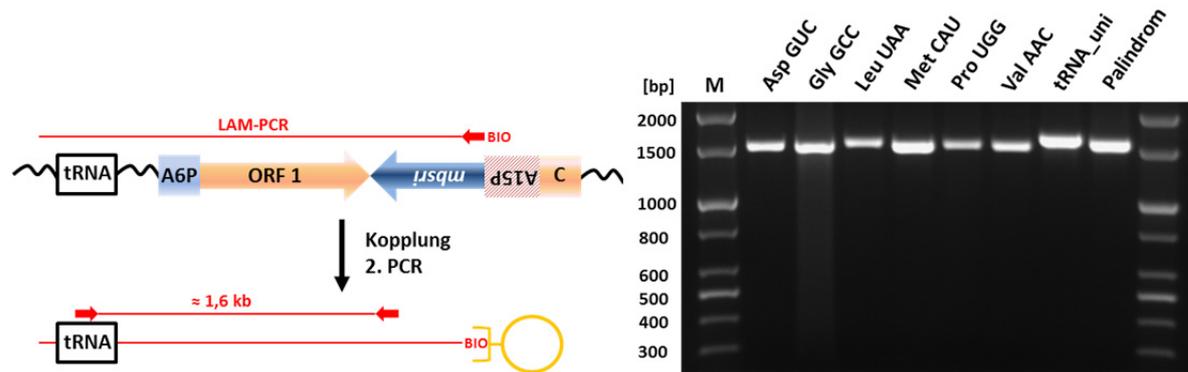


Abb. 31: Charakterisierung der Länge von integrierten TRE5-A^{bsr}-Retrotransposon-Kopien.

(A) Schema zur Erzeugung von LAM-PCR-Produkten. Vollständige Kopien des TRE5-A^{bsr}-Retrotransposons haben eine Länge von rund 1,6 kb. (B) PCR-Produkte, welche mit unterschiedlichen tRNA-Gen-Primern erzeugt werden, zeigen das es zu keiner Verkürzung der Elemente während der Retrotransposition kommt. Für die LAM-PCR wurde das Oligonukleotid BSR-enrich-biotin verwendet. Als Blasticidin-spezifischer Primer wurde BSR-09 eingesetzt. Der Primer tRNA_uni ist degeneriert und amplifiziert mehrere unterschiedliche Loci, welche mit B-Box-Motiven assoziiert sind. Für die Amplifikation von Integrationen in das Palindrom wurde der Primer Palindrom-01 verwendet.

3.2.2.7 Sequenzierung tRNA-Gen-spezifischer PCR-Produkte zur Identifizierung neuer Integrationsereignisse

Derzeit sind im Genom von *Dictyostelium discoideum* 418 tRNA-Gene annotiert (Stand: Dezember 2014; Quelle: <http://www.dictybase.org>). Eine Analyse der flankierenden Regionen der jeweiligen Loci ergab, dass sich bei 102 tRNA-Genen Vertreter der endogenen TRE5-A-Population nachweisen lassen, was rund 25 % aller tRNA-Gene entspricht. Diese Loci lassen sich 27 unterschiedlichen tRNA-Gen Familien zuordnen, wobei im gesamten Genom von *D. discoideum* insgesamt 42 tRNA-Gen-Familien vorzufinden sind.

Ausgehend von der Hypothese, dass TRE5-A-Retrotransposons zur Erkennung und spezifischen Integration an tRNA-Genen mit Untereinheiten von Transkriptionsfaktoren der RNA-Polymerase-III interagieren (Siol *et al.* 2006; Chung *et al.* 2007), sollten alle aktiv transkribierten tRNA-Gene im Genom von *D. discoideum* als potentielle Integrationsorte von TRE5-A erkannt werden. Dementsprechend sollten bei einer Analyse von mehreren hundert De-novo-Integrationen des artifiziellen TRE5-A-Retrotransposons nahezu alle tRNA-Gene mit Kopien des Retrotransposons assoziiert sein, setzt man voraus, das alle tRNA-Gene durch die RNA-Pol-III exprimiert werden.

3.2.2.8 TRE5-A^{bsr} integriert oberhalb von tRNA-Genen

Ausgehend von genomischer DNA, die rund 15.000 Blasticidin-resistenten Klonen entsprach, wurden zunächst LAM-PCR-Produkte des artifiziellen, Codon-adaptierten TRE5-A-Retrotransposons mit dem Primer Rep-234_Bio erzeugt. Nachdem die unerwünschte genomische DNA abgereichert wurde,

wurden anschließend PCRs mit Primern durchgeführt, welche selektiv unterschiedliche tRNA-Gen-Familien amplifizieren (Abbildung 32 A). Um die Identität der PCR-Produkte zu verifizieren, wurden diese in den Vektor pGEM-T kloniert und anschließend konventionell sequenziert. Auf diese Weise war es möglich, Integrationen des artifiziellen TRE5-A-Retrotransposons an verschiedenen tRNA-Gen-Loci zu identifizieren und zu analysieren. Ein Beispiel hierfür ist der Locus Lys-CUU-2 auf Chromosom 1. An diesem tRNA-Gen konnten sechs unterschiedliche Integrationsereignisse nachgewiesen werden (Abbildung 32 B). Alle Kopien des Retrotransposons besitzen am 5'-Ende einen Teil des Aktin-6-Promotors mit einer Länge von rund 42 bp. Der Beginn dieser 5'-UTR entspricht dem publizierten Transkriptionsstart des authentischen Aktin-6-Gens (McKeown und Firtel 1981). Alle sechs Integrationen an diesem Locus ereigneten sich innerhalb eines Bereichs von 44 bis 48 bp oberhalb des tRNA-Gens. Die unterschiedlichen Integrationen lassen sich dabei nur anhand kleiner Unterschiede des 5'-Endes der Retrotransposons und der Inkorporation von Extranukleotiden an der Integrationsstelle unterscheiden. Die Präzision und der Abstand der integrierten Kopien sprechen für eine authentische Retrotransposition des artifiziellen, Codon-adaptierten Retrotransposons im Vergleich mit den endogenen TRE5-A-Elementen.

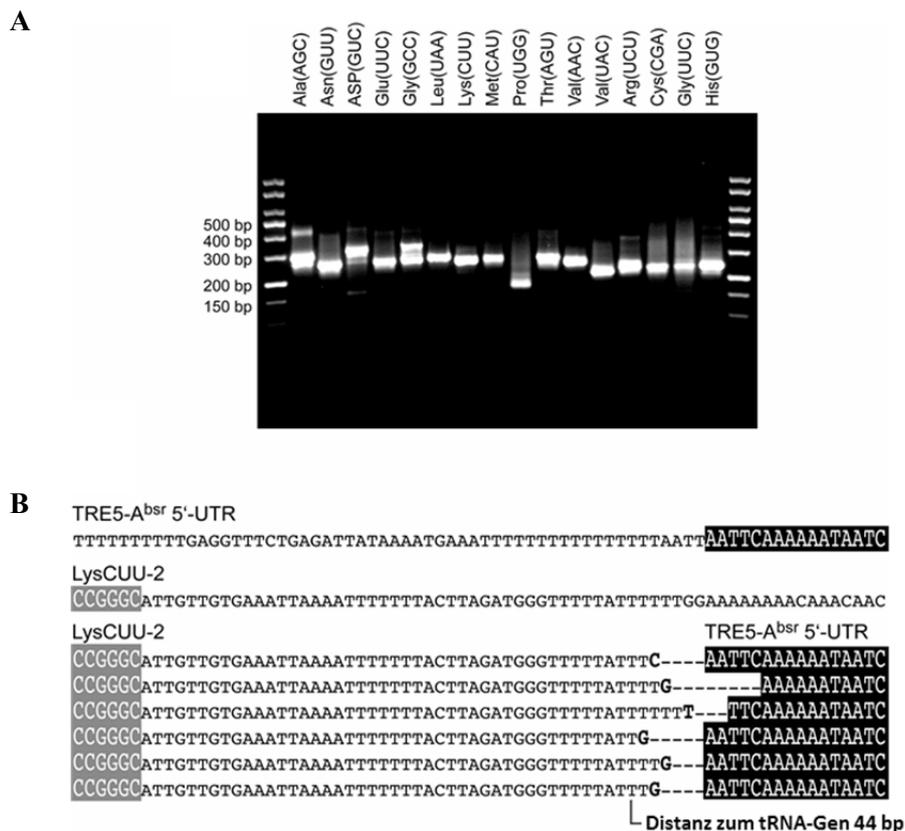


Abb. 32: tRNA-Gen-spezifische PCR-Produkte nach der LAM-PCR.

(A) Verschiedene tRNA-Gen-PCR-Produkte, erzeugt mit tRNA-Gen-Familien-spezifischen Primern. Als TRE5-A^{bsr}-spezifischer Primer wurde Rep-232_454A verwendet. (B) Dargestellt sind sechs individuelle Integrationen des TRE5-A^{bsr}-Elements oberhalb des Locus Lys-CUU-2. Fett markiert sind Extranukleotide an der Integrationsstelle.

3.2.2.9 „Next-generation“-Sequenzierung tRNA-Gen-spezifischer PCR-Produkte

Nachdem gezeigt wurde, dass es sich bei den PCR-Produkten um diverse De-novo-Integrationen des artifiziellen TRE5-A-Retrotransposons handelt, sollte in einem weiteren Experiment eine „Next-generation“-Sequenzierung verschiedener PCR-Produkte erfolgen. Die verwendeten Primer repräsentieren dabei 29 tRNA-Gen Familien, welche, bedingt durch die unterschiedlichen Kopienzahlen (von 1 bis 23; Tabelle 6) insgesamt 355 unterschiedliche tRNA-Gen-Loci abdecken. Dadurch wurden rund 85 % aller tRNA-Gene im Genom von *D. discoideum* erfasst. Als spezifischer Primer für das Codon-adaptierte TRE5-A-Retrotransposon wurde das Oligonukleotid Rep-235 verwendet. Alle PCR-Produkte wurden gelelektrophoretisch aufgetrennt und durch Gelextraktion gereinigt. Die Produktgrößen variierten zwischen 150–200 bp (Daten nicht gezeigt). Die PCR-Produkte wurden für die folgende Sequenzierung vereinigt.

Die Sequenzierung der PCR-Produkte erfolgte mit Hilfe des *MiSeq*-Sequenzierers von Illumina in der Arbeitsgruppe von PD Dr. Matthias Platzer am Fritz-Lipmann-Institut Jena. Die „paired-end“-Sequenzierung erzeugte 414.722 Sequenzpaare, wovon 380.745 miteinander assembliert werden konnten (Pseudoreads). Für eine erste Filterung der Daten wurden anschließend alle Paare nach dem 5'-Ende des artifiziellen Retrotransposons und einer tRNA-Gen-Sequenz durchsucht. Dieses Merkmal konnte bei 236.410 Pseudoreads festgestellt werden, was 62,1 % aller assemblierten Sequenzpaare entsprach. Die entsprechenden Sequenzen wurden anschließend auf das Referenzgenom „gemappt“, wobei alle Treffer mit einer minimalen Übereinstimmung von mindestens 40 bp berücksichtigt wurden. Wurden nur einmalige (unique) Treffer zugelassen, so erhielt man von diesen 236.410 Reads insgesamt 206.652 Sequenzen (87,4 %), die eine signifikante Übereinstimmung mit Loci aus dem Referenzgenom aufwiesen. Auf Grundlage dieser Daten wurden anschließend die jeweiligen angesprungenen Positionen mit Hilfe des IGV-Genombrowsers bestimmt (siehe Anhang, Tabelle C).

Durch die Kartierung der De-novo-Integrationen wurden bei insgesamt 170 tRNA-Genen neue Kopien des artifiziellen Retrotransposons festgestellt (siehe Tabelle 6), was rund 40 % der tRNA-Gen-Loci im Genom von *D. discoideum* entspricht. Lediglich 47 dieser Loci sind im Referenzgenom mit endogenen TRE5-A-Retrotransposons assoziiert. Somit konnten 123 tRNA-Gene als „neue“ Integrationsorte von TRE5-A definiert werden. Die Kartierung der Retrotranspositionereignisse zeigt, dass sich die 170 Integrationsorte auf 36 unterschiedliche tRNA-Gen-Familien verteilen, was 95 % aller möglichen tRNA-Gen-Familien im Genom von *D. discoideum* entspricht. Da die sequenzierte Bank eigentlich auf einem Pool von 29 Genfamilien basierte, haben bei der exponentiellen PCR offenbar Kreuzhybridisierungen von Primern mit sieben weiteren Genfamilien stattgefunden. Eine Analyse der jeweiligen Abstände zum 5'-Ende der tRNA-Gene unterstreicht, dass sich das artifizielle Retrotransposon analog zur endogenen TRE5-A-Population verhält. Die meisten der Kopien sind in Abständen zwischen 40 bis 50 bp oberhalb von tRNA-Genen integriert worden. Ein Maximum ist bei einem Abstand von 47 bp zu erkennen (Abbildung 33), wobei 84,5 % der Integrationen in einem

Bereich von 47 ± 3 bp identifiziert werden konnten. Allerdings gibt es auch einige Retrotransposons, die außerhalb dieses Bereichs integriert wurden. Die Abstände variieren dabei von 29 bis 68 bp.

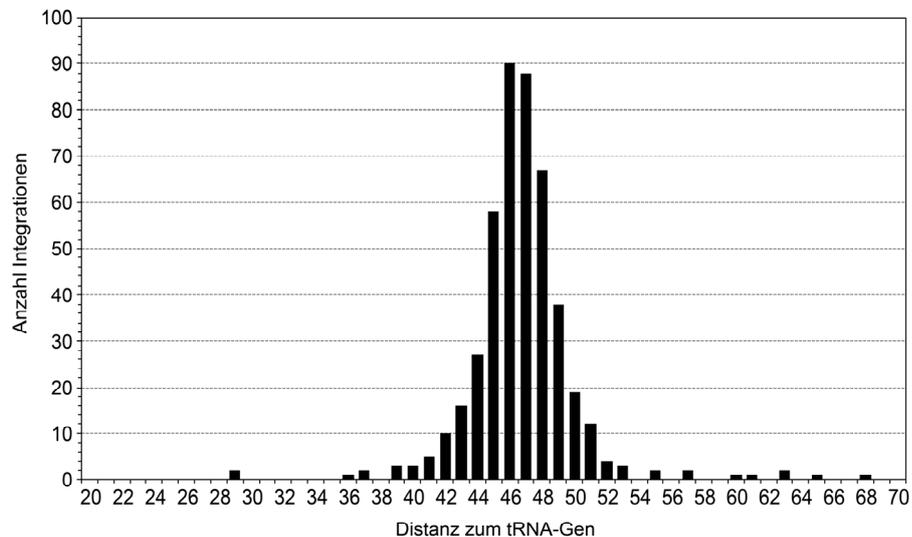


Abb. 33: Abstände der integrierten TRE5-A^{bsr}-Retrotransposonkopien zu tRNA-Genen.

Dargestellt sind die Abstände zur +1-Position der tRNA-Gene, basierend auf 458 individuellen Integrationsereignissen an 170 tRNA-Genen, die aufgrund unterschiedlicher Abstände unterscheidbar sind. Die Verteilung entspricht dem zuvor beschriebenen beobachteten Muster der endogenen TRE5-Population (Hofmann *et al.* 1991; Beck *et al.* 2002).

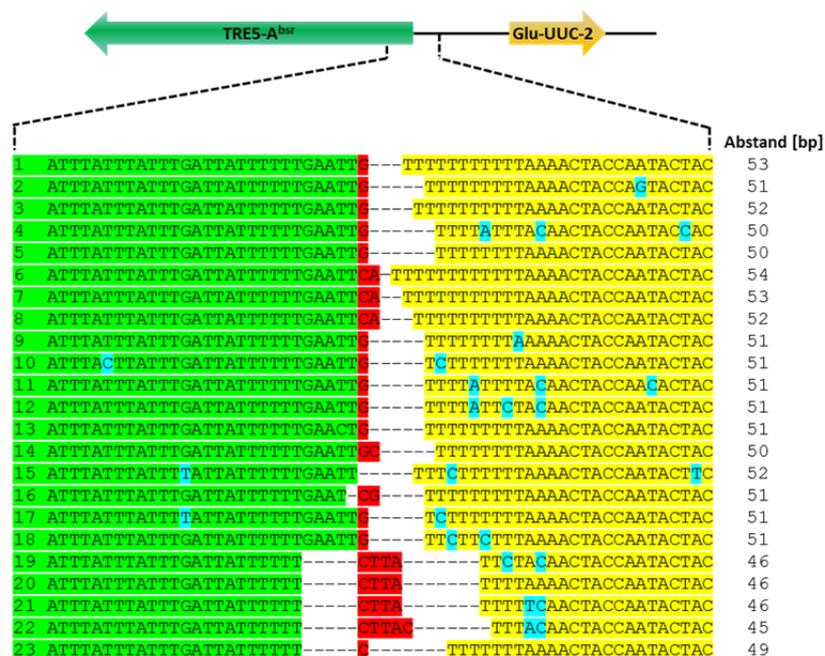


Abb. 34: Unterscheidbarkeit verschiedener Integrationsereignisse am Locus Glu-UUC-2.

In Rot sind Extranukleotide markiert, die weder Template- noch Retrotransposon-spezifisch sind. Blau dargestellt sind Sequenzunterschiede.

Um einen Eindruck von der gesamten Anzahl sequenzierter Integrationsereignisse zu bekommen, wurden exemplarisch die „mappbaren“ Reads eines tRNA-Gen Locus näher analysiert. Verwendet wurden hierfür die Sequenziererergebnisse des Locus Glu-UUC-2 von Chromosom 1. Dazu wurden von 4800 Reads zunächst alle Sequenzen 50 bp ober- und unterhalb der Integrationsstelle entfernt. Anschließend wurde ein Alignment der verbliebenen Sequenzen erzeugt, welches einen Eindruck bezüglich der Diversität in diesem Bereich gab (Abbildung 34). Die einzelnen Varianten ließen sich somit mindestens 23 unterscheidbaren Integrationsereignissen allein an dem Glu-UUC-2 tRNA-Gen zuordnen. Legt man die Annahme zugrunde, dass an allen 170 getroffenen tRNA-Genen eine ähnliche Anzahl individueller Integrationsereignisse sequenziert wurden und nimmt man weiter an, dass mit hoher Wahrscheinlichkeit viele Integrationen an exakt derselben Stelle passierten und nicht unterschieden werden konnten, dann kann man davon ausgehen, dass insgesamt der überwiegende Teil der ca. 15.000 Integrationen der analysierten Klon-Sammlung in dem tRNA-Gen-spezifischen PCR-Pool vorhanden war.

Tab. 6: Ergebnisse der Sequenzierung von tRNA-Gen-spezifischen LAM-PCR-Produkten.

Gegenüberstellung der bekannten Integrationen des endogenen TRE5-A-Elements und der neu identifizierten Integrationsereignisse durch die Sequenzierung tRNA-Gen-spezifischer LAM-PCR-Produkte des TRE5-A^{bsr}-Retrotransposons.

Gen-familie	genomische Kopien	In Referenzgenom mit TRE5-A-Integration assoziiert	TRE5-A ^{bsr} -Integration in LAM-PCR-Bank
AlaAGC	15	AlaAGC-8, AlaAGC-14, AlaAGC-15	AlaAGC-1, AlaAGC-2, AlaAGC-3, AlaAGC-5, AlaAGC-6, AlaAGC-8, AlaAGC-10, AlaAGC-11, AlaAGC-14
AlaUGC	8	AlaUGC-4, AlaUGC-7	AlaUGC-1, AlaUGC-4, AlaUGC-6, AlaUGC-7, AlaUGC-8
ArgACG	7	ArgACG-1, ArgACG-7	ArgACG-1, ArgACG-2, ArgACG-3, ArgACG-6, ArgACG-7
ArgCCU	1	-	-
ArgUCG	1	ArgUCG-1	-
ArgUCU	12	ArgUCU-3, ArgUCU-12	ArgUCU-3, ArgUCU-9, ArgUCU-12
AsnGUU	19	AsnGUU-3, AsnGUU-7, AsnGUU-8	AsnGUU-4, AsnGUU-5
AspGUC	22	AspGUC-10, AspGUC-13, AspGUC-18, AspGUC-19, AspGUC-22	AspGUC-2, AspGUC-4, AspGUC-5, AspGUC-6, AspGUC-8, AspGUC-10, AspGUC-14, AspGUC-15, AspGUC-18, AspGUC-21, AspGUC-22
CysGCA	8	-	CysGCA-3, CysGCA-4, CysGCA-6, CysGCA-7
GlnCUG	1	-	-
GlnUUG	14	GlnUUG-3, GlnUUG-4, GlnUUG-6, GlnUUG-8, GlnUUG-11	GlnUUG-4
GluCUC	3	-	GluCUC-2, GluCUC-3
GluUUC	20	GluUUC-3, GluUUC-7, GluUUC-10, GluUUC-12, GluUUC-16	GluUUC-2, GluUUC-3, GluUUC-4, GluUUC-6, GluUUC-8, GluUUC-9, GluUUC-12, GluUUC-15, GluUUC-16, GluUUC-17, GluUUC-18, GluUUC-20
GlyGCC	18	GlyGCC-3, GlyGCC-16	GlyGCC-6, GlyGCC-14, GlyGCC-16, GlyGCC-18
GlyUCC	5	GlyUCC-2	GlyUCC-1, GlyUCC-2, GlyUCC-3, GlyUCC-4
HisGUG	10	HisGUG-6, HisGUG-7, HisGUG-9	HisGUG-5, HisGUG-6, HisGUG-8, HisGUG-9, HisGUG-10

Gen-familie	genomische Kopien	In Referenzgenom mit TRE5-A-Integration assoziiert	TRE5-A ^{bsf} -Integration in LAM-PCR-Bank
IleAAU	17	IleAAU-1, IleAAU-3, IleAAU-10, IleAAU-11	IleAAU-4, IleAAU-5, IleAAU-6, IleAAU-9, IleAAU-11, IleAAU-15, IleAAU-16
IleGAU	2	–	–
IleUAU	4	–	IleUAU-1, IleUAU-2
LeuAAG	11	LeuAAG-3, LeuAAG-9	LeuAAG-6, LeuAAG-7, LeuAAG-10
LeuCAA	4	LeuCAA-1	LeuCAA-2
LeuCAG	1	–	–
LeuUAA	18	LeuUAA-1, LeuUAA-2, LeuUAA-3, LeuUAA-6, LeuUAA-11	LeuUAA-1, LeuUAA-2, LeuUAA-6, LeuUAA-7, LeuUAA-11, LeuUAA-12, LeuUAA-13, LeuUAA-15, LeuUAA-16, LeuUAA-17, LeuUAA-18
LeuUAG	3	–	–
LysCUU	10	LysCUU-2	LysCUU-2, LysCUU-5
LysUUU	23	LysUUU-4, LysUUU-6, LysUUU-7, LysUUU-15, LysUUU-22	LysUUU-2, LysUUU-3, LysUUU-4, LysUUU-5, LysUUU-6, LysUUU-10, LysUUU-12, LysUUU-13, LysUUU-14, LysUUU-15, LysUUU-16, LysUUU-17, LysUUU-18, LysUUU-19, LysUUU-22, LysUUU-23
MetCAU	17	MetCAU-12, MetCAU-15, MetCAU-16	MetCAU-7
PheGAA	16	PheGAA-6, PheGAA-11	PheGAA-11, PheGAA-12
ProAGG	1	–	–
ProUGG	16	ProUGG-3, ProUGG-8	ProUGG-2, ProUGG-5, ProUGG-7, ProUGG-14
SerAGA	9	SerAGA-5, SerAGA-9	SerAGA-1, SerAGA-2, SerAGA-5, SerAGA-6, SerAGA-7, SerAGA-8, SerAGA-9
SerCGA	1	–	SerCGA-1
SerGCU	12	SerGCU-1, SerGCU-2, SerGCU-5, SerGCU-8, SerGCU-11	SerGCU-2, SerGCU-3, SerGCU-5, SerGCU-6, SerGCU-7, SerGCU-9, SerGCU-10, SerGCU-11, SerGCU-12
SerUGA	15	SerUGA-14	SerUGA-1, SerUGA-2, SerUGA-8, SerUGA-11, SerUGA-13, SerUGA-15
ThrAGU	18	ThrAGU-4, ThrAGU-8, ThrAGU-9, ThrAGU-11	ThrAGU-1, ThrAGU-2, ThrAGU-5, ThrAGU-6, ThrAGU-7, ThrAGU-8, ThrAGU-9, ThrAGU-12
ThrCGU	1	ThrCGU-1	ThrCGU-1
ThrUGU	6	ThrUGU-1	ThrUGU-2, ThrUGU-5, ThrUGU-6
TrpCCA	7	TrpCCA-1, TrpCCA-6, TrpCCA-7	TrpCCA-1, TrpCCA-2, TrpCCA-3, TrpCCA-4, TrpCCA-6, TrpCCA-7
TyrGUA	13	TyrGUA-9, TyrGUA-11	TyrGUA-11
ValAAC	21	ValAAC-14, ValAAC-16	ValAAC-7, ValAAC-8, ValAAC-12, ValAAC-14, ValAAC-15, ValAAC-19, ValAAC-20
ValCAC	1	–	ValCAC-1
ValUAC	6	ValUAC-2, ValUAC-6, ValUAC-20	ValUAC-2, ValUAC-3, ValUAC-6, ValUAC-7

3.2.2.10 Sequenzierung von De-novo-Integrationen mit Hilfe der invers-PCR-Methode

Mit der Methode der inversen PCR (iPCR) besteht eine weitere Möglichkeit zur Amplifikation von Integrationsereignissen. Ein Vorteil dieser Methode besteht darin, Integrationsereignisse zu analysieren, welche unabhängig von vorher definierten spezifischen Integrationszielen, wie beispielsweise tRNA-Genen, amplifiziert werden können. Eine Limitation besteht allerdings in der Voraussetzung, dass an den potentiellen Integrationsorten entsprechende Restriktionsschnittstellen vorhanden sein müssen. Bei der hier angewendeten Methode wurde daher das Restriktionsenzym Sau3AI verwendet. Das Enzym erkennt und schneidet das DNA-Motiv [GATC]. Eine Analyse der genomischen Sequenzen von *D. discoideum* zeigt, dass diese Sequenz im Durchschnitt alle 500 bp im Genom zu finden ist (Daten nicht gezeigt). Entsprechende Ligationsprodukte sollten daher eine Größe von wenigen hundert Basenpaaren nicht überschreiten und somit geeignet für eine folgende Amplifikation und Sequenzierung sein. Zudem sollte das von Sau3AI erkannte Sequenzmotiv hinreichend oft im Genom vorhanden sein, um möglichst viele potentielle Integrationsstellen des TRE5-A^{bsr}-Retrotransposons abzudecken.

Um eine Amplifikation der endogenen TRE5-A-Retrotransposons zu verhindern, wurden bei der PCR ausschließlich Primer verwendet, die das Codon-adaptierte TRE5-A^{bsr}-Element spezifisch amplifizieren. Für die iPCR wurden zunächst 25 µg genomische DNA verdaut und anschließend durch eine Phenol-/Chloroform-Extraktion gereinigt. Für die Ligationen wurden unterschiedliche Konzentrationen der gefällten genomischen DNA verwendet (von 0,1 ng/µl bis 1 ng/µl). Nach einer weiteren Extraktion wurde das Pellet in Wasser aufgenommen und die DNA für die folgenden PCRs verwendet (siehe Abschnitt 2.4.1.5.; Abbildung 35 A).

Nachdem einzelne iPCR-Produkte durch konventionelle Klonierung und Sequenzierung untersucht wurden und gezeigt werden konnte, dass es sich um iPCR-Produkte von De-novo-Integrationen des *mbsri*-markierten TRE5-A-Retrotransposons handelt (Daten nicht gezeigt), wurde anschließend der gesamte iPCR-Fragment-Pool mit dem „MiSeq“-Sequenzierer (Illumina) sequenziert. Auf diese Weise wurden 648.642 „paired-end“-Sequenzen erzeugt, von denen 528.370 (81,5 %) miteinander verbunden werden konnten. Von diesen Pseudoreads enthielten 250.515 die 5'-UTR des TRE5-A^{bsr}-Retrotransposons sowie eine flankierende *D. discoideum*-spezifische Sequenz (47,4 %). Von diesen Reads konnten anschließend 230.762 mit Positionen des Referenzgenoms in Verbindung gebracht werden (92,1 %) (siehe Anhang, Tabelle D).

Eine Analyse dieser Integrationsorte ergab, dass es sich um 86 unterschiedliche tRNA-Gen-Loci handelte, welche 31 Genfamilien zugeordnet werden konnten. Im Vergleich mit den tRNA-Gen-spezifischen LAM-PCR-Produkten waren in der iPCR-Bank 20 zuvor nicht gefundene Loci enthalten. Die Abstände zu den tRNA-Genen lagen im zu erwartenden Bereich von 40 bis 50 bp (Daten nicht gezeigt). Interessanterweise wurde in dieser Bank auch eine Integration sequenziert, die sich innerhalb des DNA-Transposons DDT-S ereignet hat. Dieser Befund war zunächst überraschend und wurde daher näher analysiert. Da TRE5-A der Erwartung nach nur Ziele ansteuert, die funktionelle B-Box-

Motive aufweisen, wurden die genomischen DDT-S-Sequenzen auf A- oder B-Box-Motive untersucht. Es zeigte sich, dass sich diese Motive in 35 DDT-S-Kopien nachweisen lassen (Abbildung 35 C), die sich auf alle sechs Chromosomen verteilen. DDT-S besitzt somit offensichtlich bisher nicht identifizierte Bindungsstellen für die Transkriptionsfaktoren TFIIIB/TFIIIC der RNA-Polymerase-III. Diese Theorie wird dadurch erhärtet, dass Integrationen des *msri*-markierten TRE5-A in Positionen 42 bis 51 bp oberhalb der tRNA-Gen-analogen +1-Position beobachtet wurden (Abbildung 35 C), die DDT-S-Elemente also analog einem tRNA-Gen als Integrationsziel erkannt wurden. Da die DDT-S-Sequenzen im Bereich der TRE5-A-Integrationsstelle identisch sind, konnte nicht bestimmt werden, welche konkreten DDT-S-Kopien im Genom von *D. discoideum* von TRE5-A angesteuert wurden.

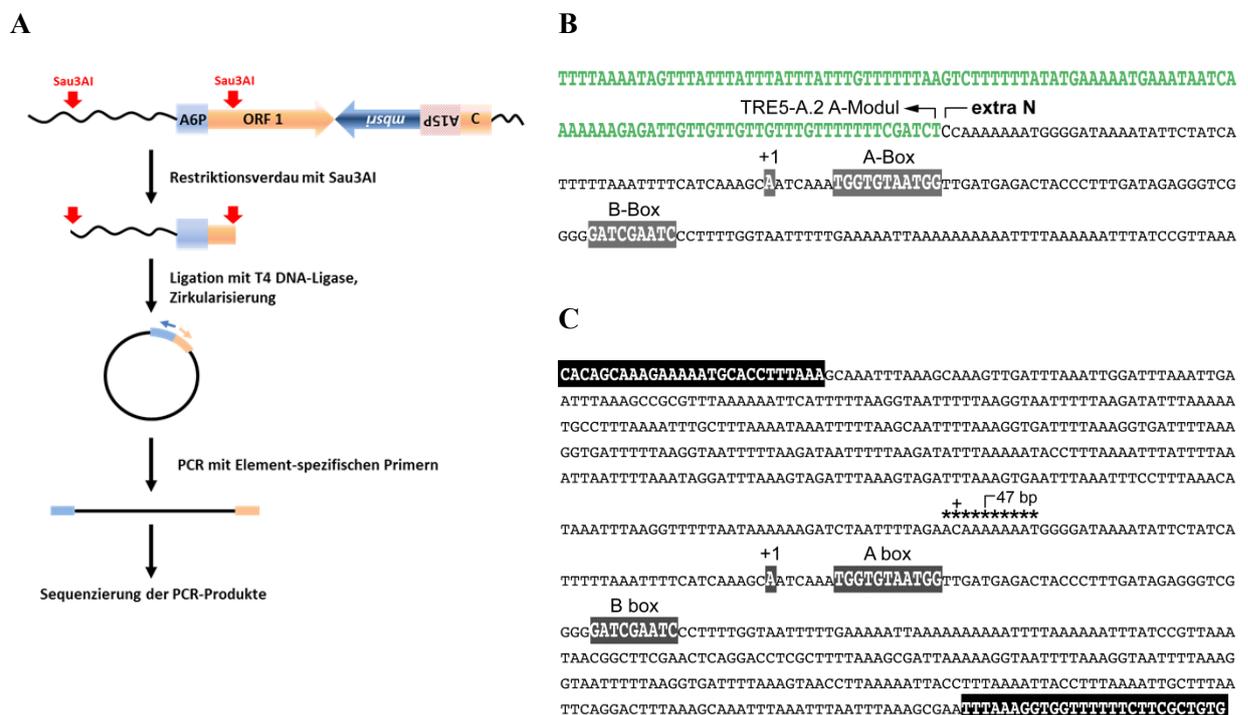


Abb. 35: Invers-PCR-Strategie zur Amplifikation von TRE5-A^{bsr} Integrationen.

(A) Schema der invers-PCR zur Identifizierung neuer Integrationsereignisse von TRE5-A^{bsr}. (B) Locus auf Chromosom 1 mit authentischer Integration eines TRE5-A.2-Retrotransposons (Grün) in ein DDT-S-Element (Schwarz). Der Abstand zur theoretischen +1-Position beträgt exakt 50 bp. (C) Unterscheidbare Integrationen des artifiziellen Retrotransposons in DDT-S-Transposons. Die Sterne (*) markieren unterschiedliche Integrationsereignisse des artifiziellen TRE5-A-Retrotransposons mit verschiedener Distanz zur theoretischen +1-Position in DDT-S. Mit (+) markiert ist die Integrationsstelle des endogenen TRE5-A.2-Elements aus (B). In schwarzen Boxen markiert sind die invertierten Sequenzwiederholungen von DDT-S.

Wenn davon ausgegangen wird, dass es sich bei den Integrationen des klonierten TRE5-A-Elementes in DDT-S um ein Abbild der authentischen Integrationspräferenz von TRE5-A handelt, sollte man erwarten, dass solche Integrationen auch im Genom von *D. discoideum* vorhanden sind.

Eine Untersuchung des Referenzgenoms nach Integrationen endogener TRE5-A-Retrotransposons in DDT-S Kopien zeigte tatsächlich einen Locus auf Chromosom 1, an dem ein TRE5-A.2-Element exakt 50 bp hinter die theoretische +1-Position, bezogen auf die Situation in tRNA-Genen, in ein DDT-S-DNA-Transposon gesprungen ist (Abbildung 35 B). Flankiert wird dieses Element von 16 bp großen Zielsequenzduplikationen, was ein Indiz für eine authentische Integration durch den für Non-LTR-Retrotransposons typischen TPRT-Mechanismus ist und somit Rekombinationsereignisse als Ursache ausgeschlossen werden können. DDT-S stellt also auch für die endogene TRE5-A-Population ein authentisches Ziel dar.

3.2.2.11 Mögliche Strategien zur Erzeugung von doppelsträngigen LAM-PCR-Produkten

LAM-PCR-Produkte liegen zunächst als DNA-Einzelstränge vor. Dies ist für die bereits beschriebene Suche nach tRNA-Gen-spezifischen Integrationen mittels konventioneller PCR unkritisch, da die unbekannt flankierenden Sequenzen durch tRNA-Gen-spezifische Primer amplifiziert wurden. Für die genomweite Sequenzierung der LAM-PCR-Produkte mit unbekannt flankierenden Regionen war es jedoch notwendig, die LAM-PCR-Produkte in Doppelstrang-DNA zu konvertieren. Dazu mussten die 3'-Enden der einzelsträngigen DNA-Moleküle so modifiziert werden, dass ein komplementärer Primer als Startpunkt für die Synthese des Zweitstrangs verwendet werden konnte. Es wurden daher mehrere Möglichkeiten für die Modifikation der 3'-Enden in Betracht gezogen:

- (A) Synthese von Oligo-dC- oder Oligo-dA-Sequenzen mit Hilfe einer Terminalen Transferase und anschließende Zweitstrangsynthese mit einem komplementären Primer.
- (B) Ligation eines einzelsträngigen DNA-Linkers mit Hilfe einer T4-RNA-Ligase, welche eine geringe Aktivität für die Ligation von DNA-Molekülen besitzt (Tessier *et al.* 1986; Schmidt *et al.* 2007).
- (C) Ligation eines einzelsträngigen DNA-Linkers durch eine spezifische DNA-Ligase, welche optimiert für die Ligation von einzelsträngigen DNA-Molekülen ist (Zhelkovsky und McReynolds 2012).

Experimente mit der Terminalen Transferase und der T4-RNA-Ligase erbrachten keine zufriedenstellenden Ergebnisse. Daher werden im Weiteren die Resultate der Versuche mit der spezifischen DNA-Ligase erläutert.

3.2.2.12 Ligation eines ssDNA-Linkers an ein definiertes PCR-Produkt

Zunächst sollte nachgewiesen werden, dass ein ssDNA-Linker effizient an eine einzelsträngige Ziel-DNA ligiert werden kann. Dazu wurde in einem Modell-Versuch ein definiertes PCR-Produkt mit den Primern Rep-201 und Rep-234_{bio} erzeugt. Als Matrize diente das Plasmid mit dem artifiziellen, Codon-adaptierten TRE5-A^{bsr}-Retrotransposon. Das 999 bp große, 5'-biotinylierte PCR-Produkt wurde anschließend über Säulen gereinigt und an magnetische Streptavidin-Kügelchen gebunden. Zur

Erzeugung von DNA-Einzelstrangmolekülen wurde die gebundene DNA für 15 min in 0,1 M NaOH inkubiert, der Überstand abgenommen und die gebundene DNA schließlich in Low-TE-Puffer aufgenommen. Die für die Ligations-Reaktion notwendige 5'-Adenylierung des ssDNA-Linkers (454B_Compl_5'P+3'ddCTP) erfolgte nach Herstellerangaben mit dem 5'-Adenylation-Kit von NEB. Um eine mehrfache, tandemartige Ligation des ssDNA-Linkers zu verhindern wurde das terminale Nukleotid am 3'-Ende des Linkers als Didesoxy-CTP synthetisiert.

Die Ligation des Linkers an die gebundene DNA erfolgte nach Herstellerangaben mit Hilfe der Thermostabilen 5'-AppDNA/RNA-Ligase (NEB). Während der einstündigen Inkubation bei 65 °C wurde die DNA mehrfach durchmischt, um ein Absinken der gebundenen DNA an den Boden des Reaktionsgefäßes zu unterbinden. Als Kontrolle wurden jeweils Reaktionen ohne Enzym und Linker durchgeführt. Des Weiteren sollte geprüft werden, ob eine vorherige Reinigung des Linkers nach der Adenylierungsreaktion zu unterschiedlichen Ergebnissen bei der Linker-Ligation führt. Daher wurde ein Teil der adenylierten Linker vor der Ligationsreaktion durch eine Ethanol-Fällung gereinigt.

Für die anschließenden Test-PCRs wurden zwei Primer-Kombinationen verwendet (Abbildung 36 A). Ein 295 bp großes PCR-Produkt mit den Primern Rep-235 und Q-A6P-01 diene als Ladungskontrolle. Der Nachweis einer erfolgreichen Ligation des Linkers an die gebundene DNA erfolgte mit Hilfe der Primer 454B_-2bp_Linker, welcher revers komplementär zur Sequenz des Linkers ist, und dem Primer Rep-235. Das daraus resultierende, 763 bp große PCR-Produkt sollte nur amplifiziert werden, wenn der adenylierte ssDNA-Linker an die gebundene DNA ligiert wurde.

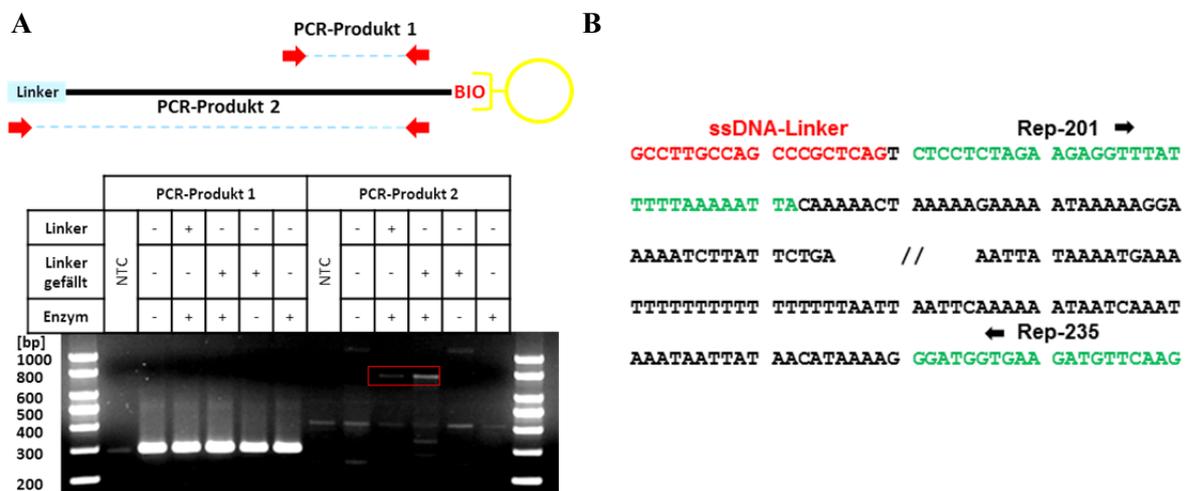


Abb. 36: Test der Ligation eines ssDNA-Linkers an ein LAM-PCR-Produkt.

(A) Ligation eines ssDNA-Linkers an gebundene, einzelsträngige DNA. In der roten Box ist das schwach amplifizierte PCR-Produkt markiert welches die Linker-Sequenz trägt. (B) Sequenz des PCR-Produkts aus (A). Der ssDNA-Linker wurde an die einzelsträngige DNA ligiert.

Wie in Abbildung 36 A zu erkennen ist, wurden PCR-Produkte nur bei den Reaktionen beobachtet, welche alle benötigten Komponenten enthielten (PCR-Produkt 2; Ansatz 2 und 3). Allerdings ist im Vergleich mit den PCR-Produkten der Ladekontrolle nur eine sehr schwache Bande in der entsprechenden Größe zu erkennen. Das deutet darauf hin, dass die Ligationsreaktion nur sehr ineffizient abläuft. Das PCR-Produkt wurde anschließend zum Nachweis der erfolgreichen Ligation des Linkers sequenziert. Es zeigte sich, dass es sich tatsächlich um das zu erwartende Ligationsprodukt gehandelt hat (Abbildung 36 B). Ebenfalls zu erkennen ist, dass die Ligation mit dem gefällten Linker, im Vergleich mit dem nicht gereinigten Linker, besser zu funktionieren schien. Daher wurden für die Erzeugung der Bank von Integrationsereignissen der gefällte, adenylierte Linker verwendet. Es wurde in den folgenden Experimenten davon ausgegangen, dass die relativ geringe Ligationseffizienz kein Problem darstellt, da zum einen nur eine relativ geringe Menge DNA für die Sequenzierung benötigt wird und zum anderen eine zweite PCR die Ausbeute an relevanten PCR-Produkten wesentlich erhöhen sollte.

3.2.2.13 Erzeugung und Sequenzierung einer Linker-basierten Bank von LAM-PCR-Produkten

Nachdem gezeigt wurde, dass die Ligation eines ssDNA-Linkers an LAM-PCR-Produkte möglich ist, sollte anschließend eine Bank von Integrationsereignissen auf diese Weise erzeugt werden. Zunächst wurde die LAM-PCR in mehreren parallelen Ansätzen unter Verwendung des Primer Rep-234_bio durchgeführt und die DNA anschließend über Säulen gereinigt. Die Kopplung der DNA an magnetische Streptavidin-Kügelchen erfolgte bei Raumtemperatur über Nacht. Um sicherzustellen, dass ein Maximum der DNA als einzelsträngiges Substrat für die Linker-Ligation zur Verfügung stand, wurde die gebundene DNA im Folgenden für 30 min in 0,1 M NaOH inkubiert. Die Adenylierung des Linkers erfolgte nach Herstellerangaben. Vor der Ligation wurde der Linker durch eine Ethanol-Fällung gereinigt und anschließend in Low-TE-Puffer resuspendiert. Die Ligation erfolgte nach Herstellerangaben, abweichend davon jedoch für 2 h bei 65 °C unter gelegentlichem Resuspendieren der gebundenen DNA. Nachdem die DNA drei Mal mit ddH₂O gewaschen wurde folgte die erste PCR. Verwendet wurden dazu die Primer Rep-232_454A, welcher spezifisch mit dem Codon-adaptierten Retrotransposon hybridisiert, und 454B_Nru_Bio. Mit Hilfe des Primers 454B_Nru_Bio sollten ausschließlich PCR-Produkte von der gebundenen DNA erzeugt werden, welche nach einer erfolgreichen Ligation am 3'-Ende die Linkersequenz besaßen.

Die 5'-biotinylierten PCR-Produkte wurden anschließend erneut an magnetische Streptavidin-Kügelchen gebunden. Um eine Eliminierung der Masterelemente zu erreichen, wurde die DNA anschließend mit dem Restriktionsenzym NcoI verdaut. Die Restriktionsschnittstelle für NcoI liegt in einem Teil des Aktin-6-Promotors, der bei den mobilisierten Kopien des artifiziellen Retrotransposons nicht mehr vorhanden war und somit nur bei den Masterelementen zu finden war. Dadurch wurde verhindert, dass die Masterelemente bei der sich anschließenden zweiten PCR amplifiziert wurden. Für die zweite PCR wurden der Retrotransposon-spezifische Primer Rep-235, sowie der Primer 454B-

Linker-2bp verwendet, welcher spezifisch mit der Linkersequenz hybridisiert. Die amplifizierte DNA wurde zum Schluss über eine Säule gereinigt und in Low-TE-Puffer aufgenommen.

Um die DNA zu testen, wurde sie einerseits direkt kloniert, um die Fragmente auf ihre Identität zu prüfen, und andererseits für Test-PCRs mit tRNA-Gen-spezifischen Primern eingesetzt. Sowohl die tRNA-Gen-spezifischen PCR-Produkte, wie auch die direkte Klonierung von PCR-Produkten zeigten, dass zum einen die Linker-Ligation funktionierte. Zum anderen konnte nachgewiesen werden, dass verschiedene Integrationsereignisse in dem Pool der finalen PCR-Produkte vorhanden waren (Daten nicht gezeigt).

Die „Paired-end“-Sequenzierung des PCR-Pools erfolgte mit dem *MiSeq*-Sequenzierautomaten in der Arbeitsgruppe von PD Dr. Matthias Platzer. Insgesamt wurden dabei 4.713.736 Sequenzpaare erzeugt, von denen 3.368.171 miteinander verbunden werden konnten. Von diesen Sequenzen enthielten 2.522.187 Paare die entsprechende Retrotransposon- und flankierende tRNA-Gen-Sequenzen (74,9 %). Beim anschließenden Mapping konnten davon 2.409.187 Sequenzen einer Position im Referenzgenom zugeordnet werden (95,5 %), wobei jedoch insgesamt lediglich 31 unterschiedliche Loci, bei Zulassen unikalere Treffer, kartierbar waren (siehe Anhang, Tabelle E). Von diesen Positionen konnten 29 mit unterschiedlichen tRNA-Genen in Verbindung gebracht werden. Die beiden verbleibenden Treffer wurden den Aktin-6- und Aktin-20-Genen zugeordnet, was auf eine Kontamination der Bank mit Masterkopien hindeutete.

Bei einem Blick auf die erzeugten Sequenzen fiel auf, dass es offensichtlich im Zuge des angewendeten Protokolls zu vielen Kreuzhybridisierungen zwischen unterschiedlichen tRNA-Genen gekommen war. Vermehrt fanden sich einzelne Sequenzen mit mehreren Linkersequenzen an unterschiedlichen Positionen, die nur mit Fehlern während der PCR erklärbar waren. Bedingt durch die Redundanz der tRNA-Gene und einer teilweise großen Homologie zwischen verschiedenen Familien bestand offenbar die Möglichkeit, dass durch fehlerhaftes Annealing Hybride aus zwei nicht-identischen DNA-Molekülen entstanden. Durch die multiplen Zyklen während der PCR wurde die „fehlerhafte Komplexität“ der Bank scheinbar schrittweise verstärkt, was im Endergebnis zu nicht auswertbaren Sequenzen führte. Ein Unterschied mit den beiden zuvor sequenzierten Banken lag im Vorhandensein kompletter tRNA-Gene, was unter Umständen unerwünschte Hybridisierungen verstärkt hat. Somit muss leider konstatiert werden, dass der Versuch, ein genomweites Mapping von TRE5-A-Integrationen durch eine LAM-PCR-basierte Methode durchzuführen, zunächst gescheitert war.

3.2.3 TRE5-A-Integrationen in das extrachromosomale rDNA-Palindrom

Im Zellkern von *D. discoideum* existieren neben der genomischen DNA auch rund 100 Kopien eines extrachromosomalen DNA-Moleküls. Die 88 kb großen Moleküle liegen in zwei 44 kb großen Armen gespiegelt vor, weshalb sie als Palindrom bezeichnet werden. Die Palindrome beherbergen die Kopien der ribosomalen RNA-Gene (Sucgang *et al.* 2003; Abbildung 37 B). Nach wie vor nicht verstanden ist allerdings die Replikation der Palindrome während der Zellteilung. Im Zuge der Untersuchung von

De-novo-Integrationen des nicht-autonomen TRE5-A-Elements (A6P-ORF1-*msri*-C) wurden erstmals Integrationsereignisse des TRE5-A-Retrotransposons auf dem Palindrom beschrieben (Siol *et al.* 2011). Bei einer Analyse der DNA-Sequenz des Palindroms konnten keine tRNA-Gen-ähnlichen Sequenzen gefunden werden. Die Integrationen ereigneten sich allerdings an einer Position in der Nähe eines B-Box-Motivs, welches ebenfalls in vielen tRNA-Gen-Familien gefunden werden kann und als Bindungsstelle für den Transkriptionsfaktor TFIIC dient (Siol *et al.* 2011; Kassavetis *et al.* 1989). Es konnte des Weiteren gezeigt werden, dass neben der mit der TRE5-A-Integration assoziierten B-Box weitere, ähnliche B-Box-Motive innerhalb des Palindroms existieren (Siol *et al.* 2011; Abbildung 37 A).

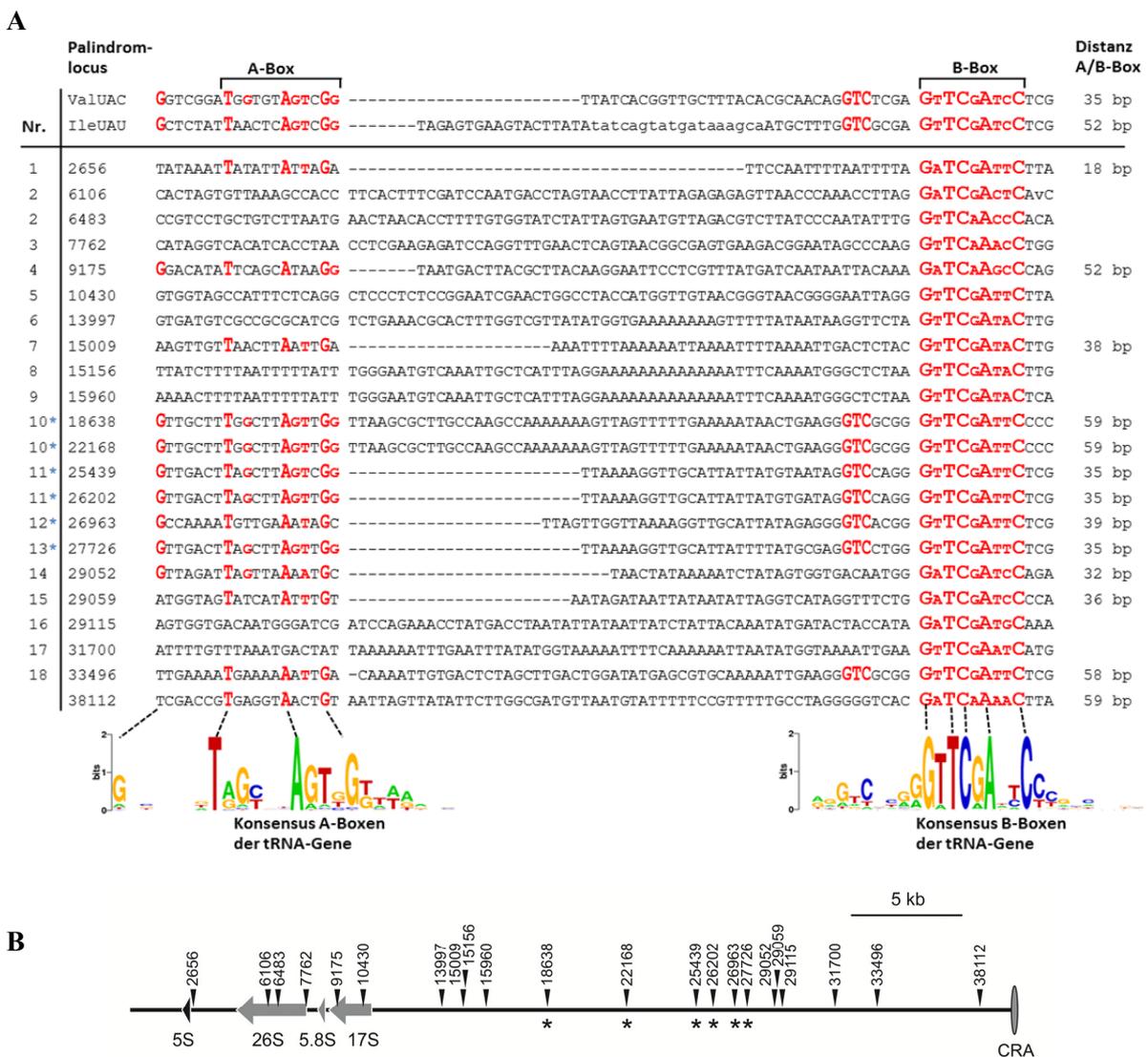


Abb. 37: A- und B-Boxen auf dem extrachromosomalen Palindrom.

(A) Alignment der detektierbaren A- und B-Box-Motive. Mehrere Loci sind identisch und können nicht unterschieden werden, daher wurden hier gleiche Nummern vergeben. Darunter angegeben sind die Konsensussequenzen der A- und B-Boxen der rund 400 tRNA-Gene. (B) Schema eines Arms des extrachromosomalen Palindroms. Dargestellt sind die Positionen der B-Box-Motive aus (A) und die ribosomalen Gene (unten). Mit (*) markiert sind die Loci, die tRNA-Gen-ähnlichen A- und B-Box-Motiven entsprechen. CRA: „central region of asymmetry“.

Ziel dieser Arbeit war es daher, zunächst zu untersuchen, ob auch Integrationen des artifiziellen Retrotransposons an weiteren B-Box-Motiven im Palindrom erfolgen können. Aufbauend auf der Hypothese, dass TRE5-A-Integrationsstellen aktive RNA-Polymerase-III-Transkriptionseinheiten darstellen, sollten zudem die Regionen der B-Boxen auf mögliche transkriptionelle Aktivität untersucht und die Transkripte näher charakterisiert werden. Ein weiteres Experiment sollte nachweisen, dass auch endogene TRE5-Elemente die B-Box-Motive als Integrationsort erkennen und somit im Palindrom integriert werden. Mit dem r5S-Gen existiert ein von der RNA-Polymerase-III transkribiertes Gen innerhalb des Palindroms. Bislang konnten noch keine Integrationen im natürlichen Kontext des r5S-Gens nachgewiesen werden. Daher sollte versucht werden sowohl Integrationen der artifiziellen, wie auch der endogenen TRE5-A-Elemente nachzuweisen.

3.2.3.1 Nachweis von Integrationen des artifiziellen TRE5-A-Retrotransposons im Palindrom

Mit Hilfe der LAM-PCR-Methode wurde zunächst eine Anreicherung von Sequenzen mit Integrationsereignissen von TRE5-A^{bsr} generiert. Dazu wurde der Primer Rep-234_{bio} verwendet, der spezifisch die Codon-adaptierten Elemente amplifiziert. Anschließend wurde eine Serie von PCRs an der angereicherten DNA durchgeführt. Dazu wurden Palindrom-spezifische Primer, welche in reverser Orientierung in den Regionen der jeweiligen B-Boxen hybridisieren, sowie ein Retrotransposon-spezifischer Primer (Rep-196_454A) verwendet.



Abb. 38: PCR zur Identifizierung von Integrationen im Palindrom.

Nach der Anreicherung von Integrationsereignissen durch eine LAM-PCR wurde eine PCR mit verschiedenen Palindrom-Primern und einem Retrotransposon-spezifischen Primer durchgeführt. Es werden ausschließlich Produkte von orthodoxen Integrationsorten amplifiziert. Die Nummern entsprechen den in Abbildung 37 dargestellten Loci auf dem Palindrom. Das PCR-Produkt von Locus 13 wurde sehr schwach amplifiziert, konnte aber durch die anschließende Sequenzierung verifiziert werden.

In Abbildung 38 ist zu sehen, dass Integrationen an den Orten nachweisbar waren, die tRNA-Gen-ähnliche A- und B-Boxen besaßen (Palindrom-Loci 10, 11, 12 und 13; siehe Abbildung 37). Zur Verifizierung wurden die PCR-Produkte kloniert und sequenziert. Es zeigte sich, dass es sich bei den

amplifizierten Fragmenten tatsächlich um Integrationsereignisse des artifiziiellen Retrotransposons handelte (Abbildung 39). Dabei integrieren die Retrotransposons exakt in der Distanz zur theoretischen +1-Position, die auch bei Integrationen an tRNA-Genen beobachtet werden können (Siol *et al.* 2011). Die Abstände variieren zwischen 40 und 49 bp in Bezug zur +1-Position.

Palindrom Locus 10; Abstand 46 bp; 9 Extranukleotide

```

CTGAAATTTT TTGTTGACAT TTATGTTGTT TAATATTTGT AGTATTATTT
GTAATTTTCA TTTTAAGCTT GAACATCTTC ACCATCCATT TTATATTATA
TTTATTTATC TGATTATTTT TTGAATAAAT CTATTCCTATT ATTTTTCAAA
AAAATGGAAT TTTAATTTTT TTGCTAAAAA GTTGCTTTGG CTTAGTTGGT
TAAGCGCTTG CCAAGCCAAA AAAAGTTAGT TTTTGAAAAA TAACTGAAGG
GTCACGGGTT CGATTCCTCCG GAATTTCTAA AAAAATGAAA AAAGTGGACC

```

Palindrom Locus 11; Abstand 48 bp, 1 Extranukleotid

```

ATGTTGTTA ATATTTGTAG TATTATTTGT AATTTTCATT TTAAGCTTGA
ACATCTTAC CATCCATTTT ATATTATATT TATTATTTG ATTATTTTTT
GAATTTTGA CTATACCCTC CGAAAAAAGT TACCATGTGA AATTTGGCCA
AAATGTTGAC CTAGCTTAGT CGGTTAAAAG GTTGCATTAT TATGTAATAG
GTCACGGGTT CGATTCCTCGG TTTATTTTTTA GACATCTGCT GACTAATCTA

```

Palindrom Locus 12; Abstand 40 bp, 3 Extranukleotide

```

TTATGTTGTT TAATATTTGT AGTATTATTT GTAATTTTCA TTTTAAGCTT
GAACATCTTC ACCATCCATT TTATATTATA TTTATTTATT TGATTATTTT
TTGAATGAT TTTGACTATA CCTCTGAAAA AAATTA AAAA GTAAAATTTG
GCCAAAATGT TGAATAGCT TAGTTGGTTA AAAGTTGCA TTATAGAGGG
GTCACGGGTT CGATTCCTCGA TTTTAACTTT TGACATATGA TGACTAATCC

```

Palindrom Locus 13; Abstand 49 bp, keine Extranukleotide

```

TGTCCTCTA ATATTTGTAG TATTATTTGT AATTTTCATT TTAAGCTTGA
ACATCTTCC CATCCATTTT ATATTATATT TATTATTTG ATTATTTTTT
GAATTTTGG ACTATCATTT CCGAAAAAAG TGAAAAAAA ATGGTGGCCA
AAATGTTTAC TTAACCTTAGT TGGTAAAAG GTTGCATTTT TTTTGGCCAG
GTCCTGGGTT CCATTCCTCAT GAACTTTCTT TACCCCTCGT TAACAAATCC

```

Abb. 39: Beispiele für Integrationen artifiziieller Retrotransposons an A- und B-Boxen im Palindrom.

Die Sequenzen stammen von den PCR-Produkten aus Abbildung 38. Grün markiert ist die Retrotransposonsequenz, gelb sind Extranukleotide dargestellt. Grau hervorgehoben ist die theoretische Position +1, analog zum Beginn von tRNA-Genen. Blau markiert sind die A-Boxen, in Rot sind die B-Box-Motive zu erkennen.

3.2.3.2 cRT-PCR zur Charakterisierung von Palindrom-Transkripten

Nachdem erfolgreich unabhängige Integrationen an mehreren Loci auf dem Palindrom nachgewiesen wurden, sollte geprüft werden, ob diese transkriptionell aktiv sind. Dazu wurden zunächst konventionelle RT-PCRs durchgeführt (Abbildung 40 A). Ein Problem schien dabei die vermutlich geringe Transkriptionsaktivität der Palindrom-Loci zu sein. So war es nicht möglich, für alle angesprochenen Integrationsstellen des Palindroms durch konventionelle RT-PCRs Transkripte nachzuweisen. In einem weiteren Experiment wurde daher versucht, mögliche Transkripte mit einer cRT-PCR zu amplifizieren. Die cRT-PCR basiert dabei auf einer intramolekularen Zirkularisierung der gesamten zellulären RNA-Population und einer Amplifikation der anschließend generierten cDNA durch eine *Nested*-PCR (Abbildung 40 B). Für die Synthese der cDNA wird dabei ein Transkript-

spezifischer Primer verwendet. Die entsprechenden PCR-Produkte wurden per Gelelektrophorese gereinigt, in den Vektor pGEM-T kloniert und sequenziert. Für alle untersuchten Palindrom-Loci war es möglich, cRT-PCR-Produkte nachzuweisen (Abbildung 40 C).

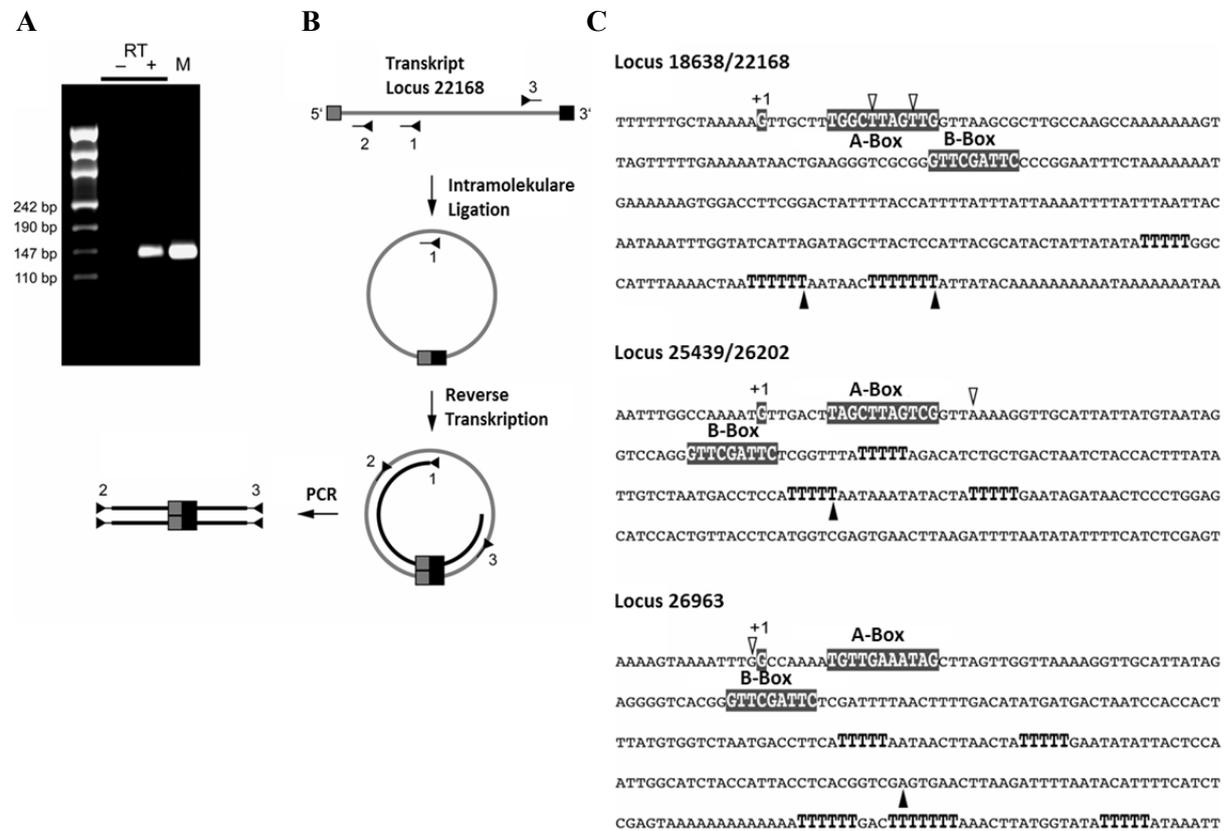


Abb. 40: Nachweis und Charakterisierung der Transkripte von Palindrom-Loci

(A) Palindrom Locus 18638 wird transkribiert. RT-PCR mit den Primern Pal-RT_04 und Pal-RT_05 an cDNA, welche spezifisch mit dem Primer Pal-RT_04 erzeugt wurde. (B) Schema der cRT-PCR zur Charakterisierung der Palindrom-Transkripte. (C) Mit cRT-PCR detektierte Transkripte der drei Palindrom-Loci 18638/22168; 25439/26202 und 26963. Weiße Pfeile markieren den Transkriptstart, schwarze Pfeile zeigen das Ende der identifizierten Transkripte an.

Die detektierten Transkripte scheinen sehr schwach exprimiert zu werden, so dass erst mit Hilfe einer *Nested-PCR* eine ausreichende Amplifikation möglich war. Der Nachweis mittels Northern-Blot scheiterte zuvor ebenfalls. Für den Locus 10 (Position 18638/22168) konnten zwei alternative Transkriptionsstartpunkte bestimmt werden, welche beide innerhalb der A-Box liegen. Die Transkripte endeten an kurzen Poly-T-Sequenzen, welche typische RNA-Polymerase-III-Terminatoren darstellen (Arimbasseri *et al.* 2013). Das Transkript an Locus 18638/22168 hat somit eine Länge von ungefähr 230 bp. Das Transkript von Locus 11 (Position 25439/26202) endet ebenfalls an einer kurzen Poly-T-Sequenz. Mit einem Transkriptionsstart kurz oberhalb der A-Box ergibt sich eine Transkriptlänge von 107 bp. Deutlich von diesen beiden Loci unterscheidet sich das beobachtete Transkript von Locus 12 (Position 26963). Das Transkript beginnt bereits ein Nukleotid vor der theoretischen +1-Position und

endet auch nicht nach 196 bp an einer Poly-T-Sequenz. Mögliche Ursachen für die Verkürzung der 5'-Enden der Transkripte der Loci 18638/22168 und 25439/26202 können 5'-Degradierungen während der RNA-Aufbereitung oder 5'-Prozessierung sein, welche auch bei tRNAs beobachtet werden kann (Castaño *et al.* 1986; Li und Deutscher 1996).

Diese Ergebnisse belegen, dass es sich bei den von TRE5-A^{bsr} angesprungenen Orten auf dem Palindrom um Loci handelt, die vermutlich aktiv von der RNA-Polymerase-III erkannt und transkribiert werden. Daraus ergibt sich allerdings auch die Frage nach der Funktion dieser Transkripte, welche im Rahmen dieser Arbeit allerdings nicht untersucht werden konnte.

3.2.3.3 Das r5S-Gen als Integrationsort artifiziieller Retrotransposons

Mit der „TRE-Falle“ wurde bereits gezeigt, dass endogene TRE5-Retrotransposons grundsätzlich das r5S-Gen, welches ebenfalls von der RNA-Polymerase-III transkribiert wird, als Ziel erkennen und dort integriert werden können (Siol *et al.* 2006). Allerdings war es unklar, ob diese Beobachtung auch auf die Situation an dem authentischen r5S-Gen-Locus auf dem rDNA-Palindrom übertragbar ist, da das r5S-Gen in der TRE-Falle in eine artifiziielle Umgebung kloniert wurde.

Mit Hilfe der zuvor beschriebenen Methoden konnten Integrationen des artifiziiellen TRE5-A^{bsr}-Retrotransposons am r5S-Locus nachgewiesen werden. Zur Verifizierung wurde ein spezifisches PCR-Produkt mit den Primern rDNA-03 und Rep-235 an LAM-PCR-Produkten erzeugt und anschließend in den Vektor pGEM-T kloniert. Dadurch konnte eine Integration eines artifiziiellen TRE5-A^{bsr}-Retrotransposon in einem Abstand von 38 bp zum r5S-Gen nachgewiesen werden (Abbildung 41). Somit wird der authentische r5S-Gen-Locus im Genom von *D. discoideum* von artifiziiellen TRE5-A^{bsr}-Retrotransposons erkannt.

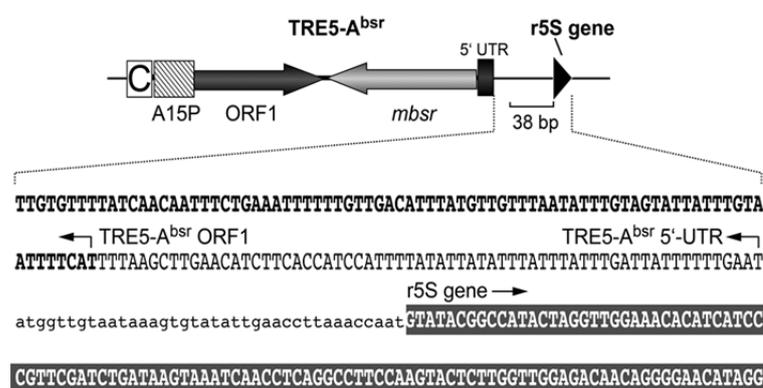


Abb. 41: Integration eines TRE5-A^{bsr}-Retrotransposons am r5S-Gen.

Das artifiziielle Retrotransposon ist in einem Abstand von 38 bp an das ribosomale 5S-Gen auf dem extrachromosomalen Palindrom gesprungen.

3.2.3.4 Integrationen endogener Retrotransposons in Palindrom-Loci

Aus den Daten der Genomsequenzierung von *D. discoideum* ergeben sich keine Hinweise auf Integrationen eines endogenen TRE5-Retrotransposons in das rDNA-Palindrom (Eichinger *et al.* 2005). Dies steht in Kontrast zu den beobachteten Integrationsereignissen der artifiziellen Elemente. Daher sollte untersucht werden, ob Integration endogener TRE5-A-Retrotransposons in das Palindrom stattfinden.

Da eine Amplifikation mit Hilfe einer einzelnen PCR zunächst keine Ergebnisse erbrachte, wurde anschließend versucht durch *Nested-PCR*-Experimente eine Amplifikation dieser Ereignisse zu erreichen. Verwendet wurde dafür genomische DNA von untransformierten AX2-Zellen. Die PCRs wurden mit den in Tabelle 7 gelisteten Primern durchgeführt. Für die zweite PCR wurde jeweils 1 µl der ersten PCR eingesetzt. Die nach der zweiten PCR erzeugten Produkte wurden anschließend kloniert und sequenziert.

Tab. 7: Verwendete Primer für *Nested-PCRs* zum Nachweis von Integrationen endogener TRE5-Retrotransposons im Palindrom.

	1.PCR	2.PCR
r5S-DNA	5S-rDNA-04 / Rep-190	5S-rDNA-03 / Rep-109Q
Pal-18639/22168	Pal-06 / Rep-190	Pal-18638_rev / Rep-157Q

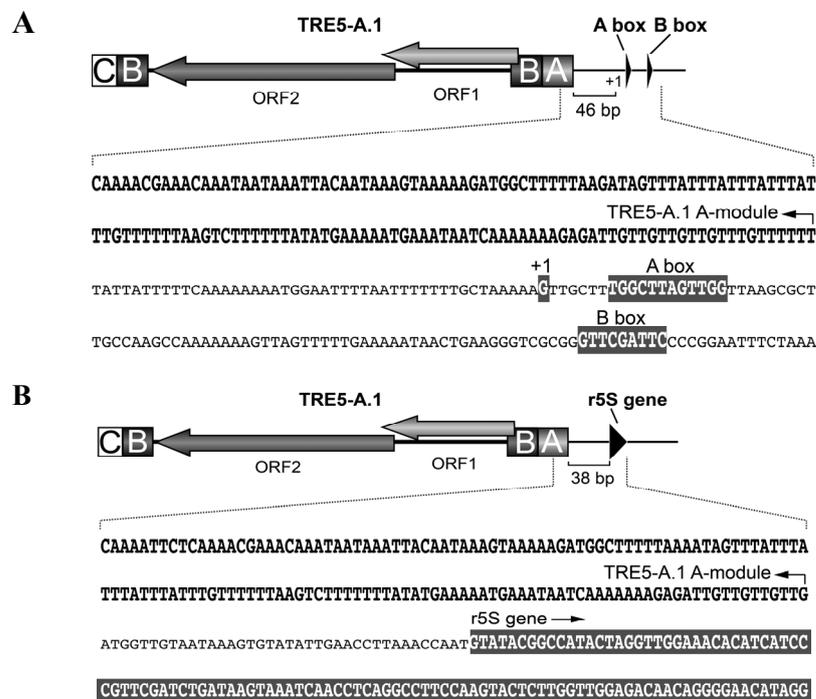


Abb. 42: Integrationen endogener TRE5-A.1-Retrotransposons in Palindrom-Loci.

(A) Integration in den Palindrom-Locus 18638/22168. Der Abstand zur theoretischen Position +1 beträgt exakt 46 bp. (B) Integration eines TRE5-A.1-Retrotransposons am r5S-Gen. Das mobile genetische Element ist in einem Abstand von 38 bp unterhalb des r5S-Gens integriert.

Auf diese Weise war es möglich, Integrationen von TRE5-A.1-Retrotransposons sowohl in der Nähe des r5S-Gens, als auch im Locus 18638/22168 des Palindroms nachzuweisen (Abbildung 42). Auffällig ist dabei, dass sowohl die endogenen als auch die artifiziellen TRE5-Retrotransposons exakt 38 bp vor dem r5S-Gen integrieren. Der Locus 18639/22168 wurde mit einem authentischen Abstand von 46 bp von dem endogenen TRE5-A.1 angesprungen.

Diese Ergebnisse zeigen, dass auch die Non-LTR-Retrotransposons der endogenen TRE5-Population die von der RNA-Polymerase-III-transkribierten Loci auf dem Palindrom als Ziele erkennen und dort in einem definierten Abstand integriert werden.

3.2.4 Erzeugung eines genetisch markierten autonomen TRE5-A.1-Retrotransposons

Die Möglichkeiten des bisher beschriebenen nicht-autonomen TRE5-A_*mbstri*-Elementes sind begrenzt, wenn es um die Charakterisierung des ORF2-Proteins von TRE5-A geht, dass die TRE5-A^{bsr}-Elemente *in trans* mobilisiert und von endogenen Elementen beigesteuert wird. Eine Untersuchung der Funktionen des ORF2-Proteins, wie die Reverse Transkriptase oder der Endonuklease, ist mit dem nicht-autonomen TRE5-A^{bsr}-Element nicht möglich. Daher schien es lohnenswert, genetisch markierte TRE5-A-Retrotransposons zu erzeugen, welche alle Voraussetzungen für eine autonome Mobilisierung erfüllen. Ausgehend von der bereits etablierten Blasticidin-Kassette mit Intron (*mbstri*), sollte daher ein autonomes TRE5-A.1-Element generiert und auf seine Aktivität in Zellkulturen getestet werden. Für die Unterscheidung der Elemente wurde die Nomenklatur verändert, so dass das nicht-autonome, genetisch markierte TRE5-Element als TRE5-A.2^{bsr} (zuvor als TRE5-A^{bsr} bezeichnet) und das autonome Element im Folgenden als TRE5-A.1^{bsr} bezeichnet wird.

3.2.4.1 Das autonome TRE5-A.1_*mbstri*-Retrotransposon

Die Erzeugung eines autonomen TRE5-A.1-Elementes (A6P-ORF1-ORF2-*mbstri*-C) soll vor allem zur Untersuchung der Domänen des zweiten offenen Leserahmens dienen. Als Ausgangspunkt wurde das bereits etablierte, nicht-autonome Blasticidin-markierte TRE5-A.2-Element verwendet (A6P-ORF1-*mbstri*-C; Siol *et al.* 2011). Des Weiteren war ein Plasmid verfügbar (PB3), in das bereits ein TRE5-A.1-Element mit kompletten ORFs kloniert worden war (Marschalek *et al.* 1992). Das Schema der einzelnen Klonierungen und die Struktur des kompletten TRE5-A.1^{bsr}-Elements sind in Abbildung 43 A dargestellt. Zunächst wurde der Aktin-6-Promotor als AgeI-Fragment von Plasmid 2.0128 (A6P-ORF1-*mbstri*-C) mit den Primern Rep-226_SalI und Rep-227_SalI amplifiziert und subkloniert. Das Fragment ORF1-ORF2 wurde als AgeI-/SalI-Produkt von Plasmid PB3 mit den Primern Rep-228_AgeI/Rep-229_SalI per PCR amplifiziert und in den pCR/XL/TOPO-Vektor (Invitrogen) via TA-Klonierung ligiert. Das hintere Ende, mit der invertierten BSR-Kassette, dem Intron und dem C-Modul, wurde wiederum von Plasmid 2.0128 als AgeI-/AgeI-Fragment amplifiziert und subkloniert. Verwendet wurden hierfür die Primer Rep-205 und Rep-230_AgeI. Zur Generierung des vollständigen A6P-ORF1-ORF2-*mbstri*-C-Retrotransposons wurden der A6-Promotor und die Resistenzkassette mit

dem C-Modul in den pCR/XL/TOPO-Vektor mit den beiden ORF2 des TRE5-A.1-Retrotransposons kloniert. Das dadurch generierte, vollständige TRE5-A.1^{bsr}-Retrotransposon wurde anschließend in AX2-Zellen Co-transformiert und die G418-resistenten Zellen anschließend mit Blasticidin selektioniert. Die entstandenen BSR-resistenten Klone wurden vereinigt und daraus genomische DNA gewonnen. Zur Verifizierung, dass es sich um Klone handelte, die Sprungereignisse von dem autonomen TRE5-A.1^{bsr}-Element akkumuliert hatten, wurde aus diesen Kulturen genomische DNA gewonnen und anschließend eine PCR mit dem ORF2-spezifischen Primer Rep-43 und A15P-seq durchgeführt (Abbildung 43 B). Parallel dazu wurde auch genomische DNA von einzelnen Klonen gewonnen und diese PCR durchgeführt. Die resultierende Doppelbande sprach für eine erfolgreiche Retrotransposition des autonomen TRE5-A.1^{bsr}-Elements. Die untere der beiden Banden, welche die gesprungenen Kopien repräsentiert, wurde anschließend gereinigt und sequenziert. Die Sequenz zeigte, dass das Intron aus der Blasticidin-Kassette gespleißt wurde (Daten nicht gezeigt). Somit konnte gezeigt werden, dass das autonome TRE5-A.1^{bsr}-Retrotransposon aktiv in AX2-Zellkulturen mobilisiert worden war.

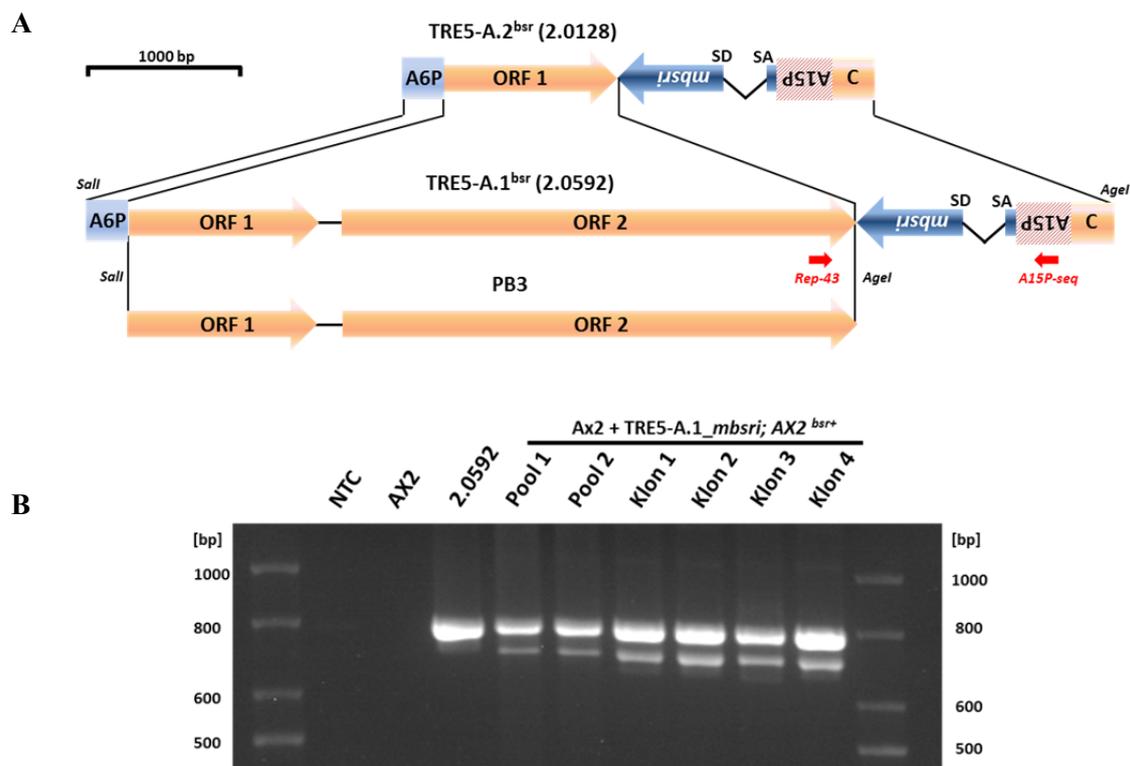


Abb 43: Das autonome Retrotransposons TRE5-A.1_{msbri}

(A) Strategie zur Erzeugung eines *msbri*-markierten autonomen TRE5-A.1-Retrotransposons. Die roten Pfeile geben die Positionen der Primer für die PCR zur Überprüfung der Retrotransposition an. (B) Test der Aktivität des TRE5-A.1_{msbri}-Retrotransposons in einer AX2-Zellkultur. Durchgeführt wurde eine PCR an genomischer DNA von BSR-resistenten Zellen mit den Primern Rep-43 und A15P-seq (siehe A). Zu erwarten sind Produkte von 807 bp (mit Intron) und 733 bp (ohne Intron). Als Negativkontrolle diente genomische DNA von AX2.

3.2.4.2 Untersuchungen zum *Cis*-Effekt der autonomen TRE5-A.1^{bsr}-Elemente

Nachdem gezeigt wurde, dass das autonome TRE5-A.1^{bsr}-Element aktiv in AX2-Zellkulturen mobilisiert wird, sollte gemessen werden, inwieweit ORF2 einen *Cis*-Effekt generiert. Ein *Cis*-Effekt liegt vor, wenn die Proteine des Retrotransposons direkt nach der Translation bevorzugt die eigene RNA binden und somit die Effektivität der Retrotransposition erhöht wird (Esnault *et al.* 2000; Wei *et al.* 2001).

Dazu sollten in parallelen Experimenten die Retrotranspositionsfrequenzen des nicht-autonomen TRE5-A.2^{bsr}- und des autonomen TRE5-A.1^{bsr}-Retrotransposons verglichen werden (Abbildung 44). Bei einem messbaren *Cis*-Effekt sollten deutlich mehr BSR-resistente Klone des autonomen TRE5-Retrotransposons zu beobachten sein, als bei dem nicht-autonomen Element. Voraussetzung für eine genaue Messbarkeit ist dabei eine synchrone Durchführung der Experimente. Die CaCl₂-Transformationen fanden daher am gleichen Tag statt und es wurde während der gesamten Kulturzeit auf die Synchronität der Kulturen geachtet.

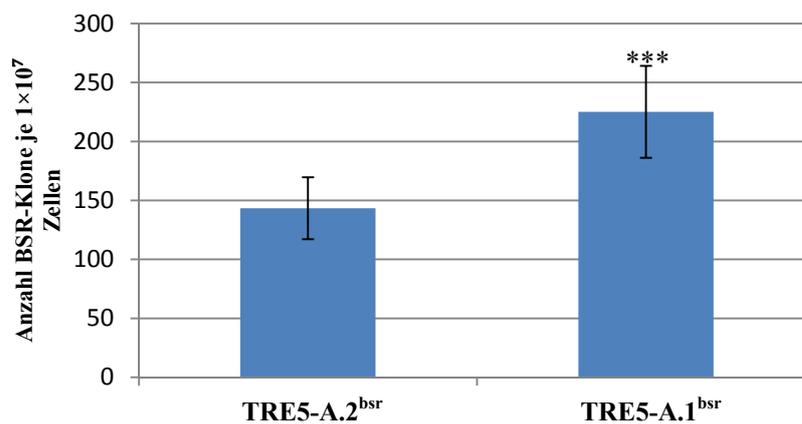


Abb. 44: Vergleich der Retrotranspositionsfrequenzen von autonomen und nicht-autonomen TRE5-A^{bsr}-Retrotransposons.

Autonome TRE5-Elemente werden häufiger erfolgreich mobilisiert, was vermutlich auf einem *Cis*-Effekt beruht. Die Daten basieren auf je drei unabhängigen Transformationen. Für die BSR-Selektionen wurden pro Transformation je drei Petrischalen mit 1×10^7 G418-resistenten Zellen verwendet. Die Selektion auf Sprungereignisse erfolgte für sieben Tage mit Blasticidin-haltigem Medium. Anschließend wurde die Anzahl der Klone je Petrischale bestimmt. n=9; *** p < 0.0001; Student's t-Test.

Die Retrotranspositionsfrequenz des nicht-autonomen TRE5-A.2^{bsr}-Retrotransposons lag bei 63 % des autonomen TRE5-A.1^{bsr}-Retrotransposons. Somit kann von einem *Cis*-Effekt ausgegangen werden, der es den autonomen, *mbsri*-markierten TRE5-A.1^{bsr}-Elementen ermöglicht, mit einer größeren Häufigkeit erfolgreich an eine neue Stelle im Genom integriert zu werden. Es konnte daher zum ersten Mal gezeigt werden, dass auch die Proteine des autonomen TRE5-A.1-Retrotransposons bevorzugt ihre eigene RNA binden und sich somit die Retrotranspositionsrate, im Vergleich mit Elementen welche *in trans* mobilisiert werden, erhöht.

4 Diskussion

4.1 Retrotransposons in den Genomen der sozialen Amöben

In dieser Arbeit wurden zum ersten Mal detailliert die Genome der sozialen Amöben *D. fasciculatum*, *D. purpureum* und *P. pallidum* nach Retrotransposon-Spezies untersucht. Ein besonderes Augenmerk lag dabei auf der Detektion von tRNA-Gen-spezifischen mobilen genetischen Elementen. Es konnten in allen untersuchten Genomen Retrotransposons mit einer tRNA-Gen-Spezifität nachgewiesen werden, was dafür spricht, dass bereits der letzte gemeinsame Vorläufer aller sozialen Amöben tRNA-Gen-spezifisch integrierende mobile genetische Elemente besaß. Große Unterschiede gibt es bei der Anzahl an Retrotransposons im Vergleich mit der bislang am besten untersuchten Art *D. discoideum*. Bei keiner der drei in dieser Arbeit untersuchten Spezies konnte eine ähnlich hohe Dichte von Retrotransposons wie im Genom von *D. discoideum* gefunden werden. Mit rund 1 % der Genomgröße beherbergt *D. purpureum* im Vergleich mit *D. discoideum* gerade einmal ein Achtel an mobilen genetischen Elemente, die den Retrotransposons zugeordnet werden können. Mit nur 0,6 % Anteil am Gesamtgenom sind die Retrotransposons in *D. fasciculatum* und *P. pallidum* sogar noch schwächer vertreten. Somit stellt sich die Frage nach den genetischen Faktoren, welche es den mobilen Elementen in den drei neu untersuchten Arten entweder erschweren im Genom weiter zu expandieren, oder im Genom von *D. discoideum* eine Expansion von Retrotransposons ermöglichen.

Tab. 8: Genomcharakteristika, Anzahl der tRNA-Gene und Anteil von Retrotransposons verschiedener sozialer Amöben. (*ohne DIRS)

	<i>D. discoideum</i>	<i>D. purpureum</i>	<i>P. pallidum</i>	<i>D. fasciculatum</i>
Genomgröße [Mb]	35	33	33	31
Anzahl tRNA-Gene	416	353	273	198
Gendichte [CDS pro Mb]	396	376	375	392
Genomanteil Retrotransposons [%]	8,058	1,006	0,668	0,605
Genomanteil tRNA-Gen-spezifische Elemente [%]	3,743	0,918	0,361	0,437
Anteil tRNA-Gen-spezifische Elemente [%]	43,6 (78,0)*	91,4	54,0	72,2

Schaut man zunächst auf die individuelle Genomarchitektur, so gibt es keine gravierenden Unterschiede, die als offensichtliche Erklärung in Frage kommen. Sowohl die Kompaktheit des Genoms, als auch die Genomgrößen sind bei allen vier Arten sehr ähnlich (Tabelle 8). Betrachtet man die Anzahl der tRNA-Gene, so fällt auf, dass im Laufe der Evolution sozialer Amöben eine Amplifikation stattgefunden hat. So beherbergt das Genom von *D. fasciculatum*, die zuerst divergierte

Spezies im Vergleich mit *D. discoideum*, nur etwa die Hälfte an tRNA-Genen (Eichinger *et al.* 2005; Heidel *et al.* 2011). Es gibt Hinweise darauf, dass die Anzahl der tRNA-Gene im Genom von *D. discoideum* durch Duplikationen entstanden sind, denn zahlreiche identische tRNA-Gene liegen als Tandems oder Triplets in sehr geringen Abständen zueinander im Genom vor (Eichinger *et al.* 2005). Somit könnte eine ausgedehnte Amplifikation von tRNA-Genen ein Phänomen der Gruppe 4 der sozialen Amöben sein, da in *D. purpureum* eine ähnlich hohe Anzahl von tRNA-Genen wie in *D. discoideum* detektiert wurde (Tabelle 8) (Sucgang *et al.* 2011). Es stellt sich daher die Frage, ob die Expansion der tRNA-Gen-spezifischen Elemente im Genom von *D. discoideum* mit einer Erhöhung der Anzahl von tRNA-Genen in Verbindung steht. Dieses Argument scheint allerdings bei einem Blick auf die Verhältnisse in *D. purpureum* nicht haltbar, denn in diesem Genom finden sich nur unwesentlich weniger tRNA-Gene als bei *D. discoideum*. Im Gegensatz dazu konnte hier allerdings nur 1 % des Genoms mit Retrotransposons in Verbindung gebracht werden, was dafür spricht, dass nicht die Anzahl von tRNA-Genen als Faktor für die Expansion von spezifisch integrierenden mobilen Elementen in Frage kommt. Interessant ist in diesem Zusammenhang der relative Anteil tRNA-Gen-spezifisch-integrierender Elemente an der Gesamtzahl vorhandener Retrotransposons. Mit einem Anteil von 91,4 % sind nahezu alle Retrotransposons im Genom von *D. purpureum* in der Nähe von tRNA-Genen zu finden (Tabelle 8). Für *D. discoideum* liegt dieser Wert bei 78 %, wenn das DIRS-Element nicht in dieser Berechnung berücksichtigt wird. In *D. discoideum* wurde das DIRS-Element stark amplifiziert und hat eine Wirtsfunktion übernommen, indem es maßgeblich zur Bildung der Centromere beiträgt (Eichinger *et al.* 2005). Dies ist in der Form in den anderen hier untersuchten Spezies nicht der Fall, z. B. ist das DIRS-Element in *D. purpureum* komplett abwesend.

Eine ähnliche Häufigkeit tRNA-Gen-spezifisch integrierender Retrotransposons lässt sich mit einem Anteil von 72,2 % aller Retrotransposons für *D. fasciculatum* berechnen. Im Genom von *P. pallidum* sind dagegen nur rund die Hälfte (54 %) aller Retrotransposons in der Umgebung von tRNA-Genen zu finden, bei einem ähnlichen genomischen Gesamtanteil von Retrotransposons wie im Genom von *D. fasciculatum*. Die soziale Amöbe *P. pallidum* besitzt jedoch mit 273 tRNA-Genen insgesamt 75 tRNA-Gene mehr als *D. fasciculatum*. Damit ist diese Beobachtung ein weiteres Argument dafür, dass die Anzahl von tRNA-Genen in einem Genom nicht ausschlaggebend für die Expansion von tRNA-Gen-spezifischen Retrotransposons ist. Somit muss es andere Faktoren geben, die für die beobachteten Unterschiede in Frage kommen.

Eine weitere mögliche Erklärung für die starke Expansion tRNA-Gen-spezifischer Retrotransposons könnte in Änderungen der Wirtszell-basierten Abwehr gegen mobile genetischen Elemente zu suchen sein. Womöglich haben es die Retrotransposons in *D. discoideum* mit einer, im Vergleich mit den drei anderen Arten, schwächeren Wirtsabwehr zu tun? Betrachtet man zentrale Bestandteile der RNA-Interferenz-Maschinerie, so sind zunächst keine gravierenden Unterschiede festzustellen (Tabelle 9). Mit 5 unterschiedlichen Argonaut-Genen scheint *D. discoideum* lediglich innerhalb eines Teils der RNAi-Maschinerie über ein breiteres Spektrum zur post-transkriptionellen Regulation zu verfügen

(Hutvagner und Simard 2008). Bezogen auf diese Beobachtung könnte man vermuten, dass in *D. discoideum* eine größere Palette an Werkzeugen zur Bekämpfung von mobilen genetischen Elementen existiert. Im Umkehrschluss müsste dann aber eher eine Reduktion der TRE-Elemente zu beobachten sein. Eine weitere Möglichkeit besteht darin, dass es erst vor einer evolutionär „kurzen Zeit“ zu einer starken Amplifikation der Retrotransposons in *D. discoideum* gekommen ist. Interessant ist in diesem Zusammenhang, dass das LTR-Retrotransposon Skipper stark durch posttranskriptionelle Regulation unterdrückt wird und offensichtlich in Wildtyp-Zellen kaum erfolgreich mobilisiert wird (Abschnitt 3.2.1.1). Im Gegensatz dazu scheinen die TRE5-Elemente wesentlich schwächer durch RNAi- oder andere Regulationsmechanismen beeinflusst zu sein und werden in *D. discoideum*-Zellen permanent mobilisiert.

Tab. 9: Detektierbare Bestandteile der RNA-Interferenz-Maschinerie der untersuchten Spezies.

	<i>D. discoideum</i>	<i>D. purpureum</i>	<i>P. pallidum</i>	<i>D. fasciculatum</i>
Dicer	2	1	2	2
RdRP	3	2	2	2
AGOs	5	2	2	3

RdRP: RNA-abhängige RNA-Polymerase; AGO: Argonaut-ähnliche Proteine

Die Kontrolle der mobilen genetischen Elemente kann auf mehreren Ebenen stattfinden, von denen RNA-Interferenz nur eine ist (vgl. Abschnitt 1.5.1). Eine weitere Möglichkeit der Zelle eine ungehemmte Akkumulation von mobilen genetischen Elementen zu verhindern, besteht in der Hemmung der Transkription von genomischen Kopien durch deren Methylierung. Für die Retrotransposons Skipper und DIRS wurden bereits Experimente durchgeführt, die auf eine Abhängigkeit der Aktivität von Retrotransposons von der zelleigenen Methylierungs-Maschinerie hindeuten. Es konnte dabei gezeigt werden, dass der Knock-Out eines an der Methylierung von genomischer DNA beteiligten Gens (*dnmA*) zu einem Anstieg der Retrotransposition von Skipper-Retrotransposons führt (Kuhlmann *et al.* 2005). Denkbar ist daher auch, dass vor allem TRE5-Elemente in *D. discoideum* weniger stark methyliert sind und daher verhältnismäßig oft mobilisiert werden. Derzeit gibt es noch keine Erkenntnisse über den globalen Methylierungsstatus der TRE-Elemente im Genom von *D. discoideum*.

Ein weiterer experimenteller Beweis für solche Unterschiede könnte darin bestehen, eines der Non-LTR-Retrotransposons der drei anderen Arten mit einer Retrotranspositionskassette auszustatten. Somit könnten Retrotranspositionsfrequenzen mit denen der bereits etablierten artifiziellen TRE5-Elementen in *D. discoideum* verglichen werden. Experimente in Mutanten, welche zum Beispiel Defekte in Bestandteilen der Methylierungs- oder RNAi-Maschinerie besitzen, könnten ebenfalls Aufschluss auf vermeintliche Unterschiede zwischen den Arten geben. Ein weiterer Weg führt in Zukunft sicher über die Sequenzierung und Assemblierung von Genomen weiterer Vertreter der sozialen Amöben, insbesondere der mit *D. discoideum* am nächsten verwandten Vertreter aus Gruppe 4. Dadurch wäre möglicherweise erkennbar, ob es sich bei den in *D. discoideum* beschriebenen

Verhältnissen um eine Ausnahme handelt -was die in dieser Arbeit dargestellten Ergebnisse nahelegen- oder ob sich ähnlich große Anteile mobiler genetischer Elemente auch in anderen Genomen auffinden lassen.

4.2 Variabilität der Integrationsabstände zu tRNA-Genen

Ein wesentlicher Fokus dieser Arbeit lag darin, die Evolution der tRNA-Gen-spezifischen Integration von Retroelementen besser verstehen zu können. Ein Vergleich der Durchschnittswerte des Abstandes zu tRNA-Genen der jeweiligen Retrotransposon-Familien lässt dabei einige Rückschlüsse zu:

- (1) Es gibt zwei relativ eng begrenzte Integrationsfenster oberhalb von tRNA-Genen
- (2) Unterhalb von tRNA-Genen gibt es eine breitere Streuung
- (3) Integrationen 3' von tRNA-Genen sind erst ab rund 50 bp möglich
- (4) DGLT-A-Elemente integrieren sehr eng oberhalb von tRNA-Genen
- (5) Spezifische Skipper-Retrotransposons integrieren unterhalb von tRNA-Genen
- (6) Alle LTR-Elemente unterhalb von tRNA-Genen besitzen eine Chromodomäne

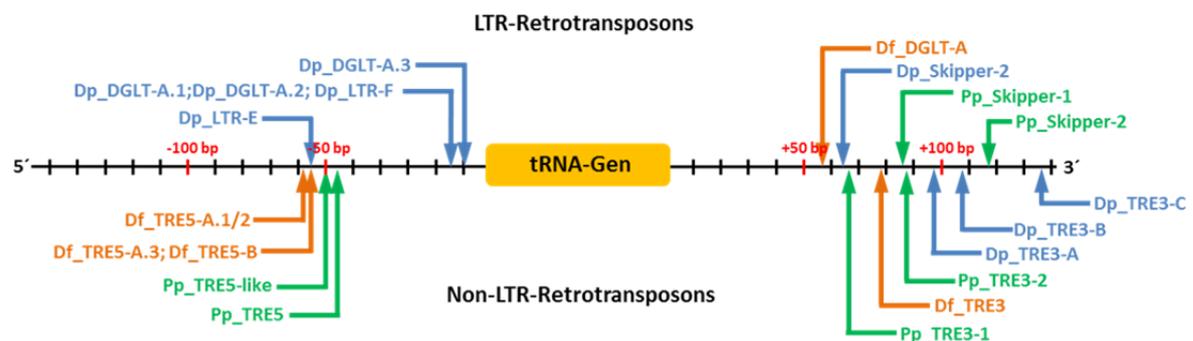


Abb. 45: Integrationsabstände der neu gefundenen Retrotransposons zu tRNA-Genen.

Dargestellt sind die Durchschnittswerte aller bekannten Integrationsereignisse der jeweiligen Retrotransposons. Oberhalb sind die Positionen der tRNA-Gen-spezifischen LTR-Retrotransposons dargestellt, unterhalb die Abstände der Non-LTR-Elemente.

Die dargestellten Integrationspräferenzen (Abbildung 45) lassen darauf schließen, dass die Wechselwirkung zwischen den Transkriptionsfaktoren TFIIB/TFIIC und den oberhalb integrierenden Retrotransposons stark konserviert ist. Das lässt sich daran erkennen, dass selbst in der evolutionär am weitesten von *D. discoideum* entfernten Spezies, *D. fasciculatum*, noch TRE-Elemente nachweisbar sind. Diese werden, analog zu den TRE5-Elementen aus *D. discoideum*, in rund 50 bp Abstand oberhalb von tRNA-Genen integriert. Derartige Beobachtungen gelten ebenfalls für die TRE5-Elemente aus *P. pallidum*. Andere Verhältnisse lassen sich unterhalb von tRNA-Genen beobachten. Im Bereich 3' von tRNA-Genen integrieren mobile genetische Elemente in einem relativ großen Fenster,

welches, beginnend bei etwa 50 bp, bis weit über 100 bp hinaus ragt. Das lässt sich zum einen mit dem Vorhandensein von Extra-B-Boxen erklären, die rund 40 bp unterhalb von ca. 66 % der tRNA-Gene in *D. discoideum* vorkommen (Hofmann *et al.* 1991; Marschalek und Dingermann 1991; Siol *et al.* 2006). Zudem ist bekannt, dass tRNA-Gene und deren unmittelbare Umgebung beispielsweise in *S. cerevisiae* grundsätzlich frei von Nucleosomen sind. Dabei ist die Grenze zu den flankierenden Nucleosomen oberhalb von tRNA-Genen starrer als unterhalb (Kumar und Bhargava 2013). Des Weiteren wurde für die Hefe gezeigt, dass der TFIIB:TFIIC-Komplex eine gewisse Flexibilität aufweist (Nagarajavel *et al.* 2013). Bezogen auf die Situation in *D. discoideum* wären das mögliche Erklärungen für die breitere Streuung der Integrationspräferenzen unterhalb und das relativ kleine Fenster oberhalb von tRNA-Genen. Da die Ermittlung des durchschnittlichen Abstands maßgeblich durch die Anzahl von individuellen Integrationsereignissen beeinflusst wird und für einige Elemente nur wenige Daten diesbezüglich vorliegen, kann die in Abbildung 45 gezeigte Auswertung jedoch nur zu hypothetischen Überlegungen anregen. Erst mit der Analyse der Genomdaten von weiteren sozialen Amöben kann sich das hier vorgestellte Bild vervollständigen.

4.3 Ein neu entdecktes DGLT-A-Element aus *D. fasciculatum*: Verbindung zwischen Skipper und DGLT-A-Familie?

Die in dieser Arbeit durchgeführten Genomanalysen deuten erstmals auf eine Verteilung der LTR-Retrotransposons auf zwei Familien hin: DGLT-A und Skipper. Ein Großteil der neu identifizierten LTR-Retrotransposons lässt sich aufgrund von strukturellen und phylogenetischen Merkmalen (vgl. Abbildung 24) einer dieser beiden Gruppen zuordnen. Einen Sonderfall scheint das DGLT-A-Element aus *D. fasciculatum* darzustellen. Als mutmaßliches Mitglied der DGLT-A-Familie integriert dieses Retrotransposon unterhalb von tRNA-Genen, wohingegen alle anderen bisher detektierten DGLT-A-Elemente stets oberhalb von tRNA-Genen gefunden wurden. Womöglich beruht dieser Unterschied auf der Tatsache, dass *Df_DGLT-A* eine zusätzliche Chromodomäne besitzt, die sonst nur bei LTR-Retrotransposons der Skipper-Familie zu finden ist (Leng *et al.* 1998; Malik und Eickbush 1999). Denn auch die anderen neu identifizierten tRNA-Gen-spezifisch integrierenden Elemente der Skipper-Familie, wie *Dp_Skipper-2* und die beiden Skipper-Retrotransposons aus *P. pallidum*, integrieren obligat hinter tRNA-Genen. Daher stellt sich die Frage, ob eine Interaktion der Chromodomäne mit Transkriptionsfaktoren der RNA-Polymerase-III für die Integration der Skipper-Retrotransposons oberhalb von tRNA-Genen mit verantwortlich ist. Dem widerspricht zunächst die Erkenntnis, dass es auch eine Reihe von nicht tRNA-Gen-spezifischen Skipper-Elementen gibt (Winckler 2002), wofür drei Erklärungen in Frage kommen:

- (1) Die nicht-spezifisch integrierenden Skipper-Elemente haben ihre Spezifität sekundär verloren
- (2) Es handelt sich um inaktivierte Elemente, die im Lauf der Zeit von ihren Zielen getrennt wurden
- (3) Die Chromodomäne ist nicht an der Zielerkennung unterhalb von tRNA-Genen beteiligt

Mit der Voraussetzung, dass ein aktives, genetisch markiertes *Df_DGLT-A*-Element verfügbar ist, könnte ein Experiment, bei der die Chromodomäne entweder manipuliert oder ersetzt wird, Aufschluss darauf geben, ob diese strukturelle Eigenheit des *Df_DGLT-A*-Elements für die Integration unterhalb von tRNA-Genen verantwortlich ist.

4.4 Sequenzierung Linker-ligierter LAM-PCR-Produkte

Mit der fortschreitenden Entwicklung der modernen Sequenziermethoden folgten in den letzten Jahren auch zahlreiche Protokolle zur Charakterisierung von neuen Integrationsereignissen verschiedenster mobiler genetischer Elemente (Badge *et al.* 2003; Gabriel *et al.* 2009; Ewing und Kazazian 2010; Roggo *et al.* 2013). Vor allem für das humane Non-LTR-Retrotransposon L1 sind verschiedene Möglichkeiten vorhanden, den exakten Integrationsort einer neuen Kopie im Genom experimentell zu bestimmen. Zu Beginn dieser Arbeit stellte sich die Frage nach einem relativ einfachen Protokoll, welches das Potential besitzen sollte, möglichst viele unabhängige neue Integrationsereignisse des artifiziellen TRE5-A.2^{bsr}-Elements innerhalb eines Experiments analysieren zu können. Daher sollen im Folgenden kurz die Grundlagen der bis dato publizierten Methoden erläutert werden. Zunächst gibt es zwei mögliche Ansätze: Entweder, ausgehend von gesamtgenomischer DNA, direkt selektiv Sequenzen mobiler genetischer mit flankierenden Regionen zu amplifizieren (Badge *et al.* 2003; Qi *et al.* 2012), oder zunächst eine Anreicherung der betreffenden Sequenzen durch eine LAM-PCR zu erreichen (Gabriel *et al.* 2009; Schmidt *et al.* 2007). Beide Methoden erfordern, als entscheidenden Schritt, die Ligation eines für alle erzeugten Fragmente identischen Linkers. Dabei erfolgt die Ligation des Linkers, ausgehend von doppelsträngiger DNA, auf der Grundlage von im Genom vorhandenen Restriktionsschnittstellen (Badge *et al.* 2003; Qi *et al.* 2012). Für die einzelsträngigen LAM-PCR-Produkte bestehen folgende Möglichkeiten zur Ligation eines ssDNA-Linkers, beziehungsweise zur Erzeugung eines Doppelstrangs:

- (1) Ligation eines ssDNA-Linkers mit Hilfe einer Ligase: T4-RNA-Ligase (Paruzynski *et al.* 2010) bzw. ssDNA-Ligase (Zhelkovsky und McReynolds 2012)
- (2) Erzeugung des Doppelstrangs durch degenerierte Hexanukleotide und der Klenow-Polymerase, anschließender Restriktionsverdau und Ligation eines dsDNA-Linkers (Schmidt *et al.* 2007)
- (3) Hemispezifische PCR mit degenerierten Primern (Ewing *et al.* 2010)

Alle beschriebenen Methoden zielen darauf ab, eine bekannte Linker-Sequenz am Ende der unbekannt, flankierenden Region zu generieren, um anschließend mit einem Element- und einem Linker-spezifischen Primer PCR-Produkte zu erhalten, die den Übergang zwischen dem mobilen genetischen Element und der flankierenden genomischen Region beinhalten. Der Vorteil der (neueren) LAM-PCR-basierten Methoden liegt darin, dass keine Restriktionsschnittstellen im Genom benötigt werden. Hierbei muss jedoch ein wesentlich größerer Aufwand betrieben werden, um am Ende

doppelsträngige DNA zu erhalten. Andererseits sind Methoden, wie die inverse PCR und die Ligation von doppelsträngigen Linkern an verdaute DNA mit Hilfe der T4-DNA-Ligase wesentlich etablierter. Den größten Erfolg für die Ziele in dieser Arbeit versprach eine LAM-PCR-basierte Methode, bei der ein einzelsträngiger Linker an das freie 3'-Ende ligiert wird und anschließend die Fragmente mittels PCR amplifiziert werden können. Vor allem durch die Verfügbarkeit einer speziellen Einzelstrang-DNA-Ligase (Zhelkovsky und McReynolds 2012) sollte die Ligation des Linkers, im Gegensatz zu Methoden mit einer T4-RNA-Ligase (Troutt *et al.* 1992; Paruzynski *et al.* 2010), weitaus effizienter sein, was sich allerdings im Nachhinein nicht bestätigte. Zur Detektion neuer Integrationsereignisse von TRE5-A-Elementen wurden im Rahmen dieser Arbeit aber auch weitere, zuvor vorgestellte Protokolle angewendet und modifiziert. Es zeigte sich jedoch, dass eine Anpassung schwierig zu sein scheint, was folgende Ursachen haben kann:

- (1) Der sehr hohe A/T-Gehalt des Genoms von *D. discoideum* (Eichinger *et al.* 2005), insbesondere in den nicht-codierenden Bereichen, also auch zwischen tRNA-Genen und TRE5-A-Kopien
- (2) Die Redundanz der tRNA-Gene kann zu Kreuzhybridisierungen während der PCR führen
- (3) Einzelsträngige LAM-PCR-Produkte von tRNA-Genen können Sekundärstrukturen bilden

Während der Etablierung eines LAM-PCR-Protokolls zur Anreicherung von TRE5-A.^{2^{bsr}}-Integrationen konnte gezeigt werden, dass der für die Zweitstrangsynthese und Amplifikation benötigte Linker an einzelsträngige LAM-PCR-Produkte ligiert wurde. Die Klonierung und anschließende Analyse einzelner PCR-Produkte suggerierte eine erfolgreiche Präparation der Bank für die anschließende Hochdurchsatz-Sequenzierung. Allerdings zeigte sich, dass es offensichtlich zu Problemen während der Amplifizierung oder Sequenzierung kam und somit nur wenige Integrationsereignisse mit Hilfe dieses Protokolls analysiert werden konnten. Im Folgenden sollen daher mögliche Fehler diskutiert werden. Zwei Auffälligkeiten wurden bei einem Großteil der Reads festgestellt: Zum einen die mehrfache Sequenz des Linkers in einem Read und zum anderen die schlechte Qualität der Sequenzen in den flankierenden genomischen Bereichen der tRNA-Gene. Beide Beobachtungen deuten auf ungewollte Rekombinationsereignisse zwischen verschiedenen tRNA-Genen hin. Ein wesentlicher Unterschied dieser Bank, im Vergleich mit den beiden anderen sequenzierten Banken (siehe Abschnitt 4.5.), liegt im Vorhandensein kompletter tRNA-Gene. Außerdem trägt jedes PCR-Produkt, mit dem Linker am 3'-Ende, eine identische Sequenz. Beide Charakteristika können dafür verantwortlich sein, dass es zu Rekombinationen zwischen redundanten Sequenzen, also gleichen tRNA-Gen-Familien oder den Linker-Sequenzen kommt. Dadurch können Mischsequenzen entstehen, die in einer folgenden Sequenzierung nur unzureichende und fehlerhafte Informationen generieren. Zur Unterbindung dieser Probleme wäre es denkbar beispielsweise kürzere LAM-PCR-Produkte zu generieren, die keine kompletten tRNA-Gene mehr beinhalten. Das würde wahrscheinlich die Kreuzhybridisierungen in diesen Bereichen verhindern oder reduzieren. Zum anderen wäre es möglich, mehrere unterschiedliche ssDNA-Linker für die Ligation an LAM-PCR-

Produkten zu verwenden, um bei einer anschließenden PCR auch an den 3'-Enden eine größere Diversität zu erhalten. Die vorliegenden Ergebnisse deuten jedoch darauf hin, dass dieses Protokoll in Kombination mit der hier verwendeten Sequenzier-Methode, zur Identifizierung der Integrationsorte von tRNA-Gen-spezifischen Retrotransposons, als ungeeignet eingestuft werden muss.

4.5 Genomweite Analyse Pol-III-assoziierter Integrationsereignisse

In dieser Arbeit wurden Assays etabliert, mit denen genomweit neue Integrationsereignisse nachweisbar wurden. Mit der Etablierung der LAM-PCR zur Anreicherung dieser Ereignisse wurde eine Möglichkeit geschaffen, auch relativ selten angesprungene Integrationsorte, wie zum Beispiel am ribosomalen 5S-Gen, für die Amplifikation und Analyse mittels PCR zugänglich zu machen. Auf diese Weise konnten zum ersten Mal simultan viele De-novo-Retrotranspositionsergebnisse des TRE5-A.2^{bsr}-Retrotransposons innerhalb eines Experiments erfasst werden (Abbildung 45).

Es wurde gezeigt, dass sich das artifizielle Retrotransposon TRE5-A.2^{bsr} analog zu den endogenen TRE5-Elementen verhält und in einem definierten Bereich von 44-50 bp oberhalb von tRNA-Genen integriert wird. Charakteristische Merkmale des von den Non-LTR-Retrotransposons verwendeten TPRT-Retrotranspositionsmechanismus (Luan *et al.* 1993), wie 5'-Deletionen des integrierten Elements oder Inkorporation von Extranukleotiden an der Integrationsstelle (Sassaman *et al.* 1997) zeigen ebenfalls, dass sich das klonierte, genetisch markierte TRE5-Retrotransposon wie die Elemente der endogenen TRE5-Population verhält und wahrscheinlich nach dem TPRT-Mechanismus amplifiziert. Die Beobachtung, dass das TRE5-A.1^{bsr}-Element nahezu alle in diesem Experiment abgedeckten tRNA-Gene als Ziele erkennt und anspringt, zeigt, dass offenbar die weitaus meisten tRNA-Gene im Genom von *D. discoideum* grundsätzlich als Integrationsorte zur Verfügung stehen.

Mit der zur LAM-PCR komplementären Methode, der inversen PCR, wurden diese Ergebnisse bestätigt (Abbildung 46). Die kombinierten Ergebnisse der beiden Methoden geben einen guten Einblick in die Dynamik der TRE5-Population. Mit der Beobachtung, dass wahrscheinlich alle tRNA-Gene als Integrationsorte von artifiziellen TRE5-Elementen erkannt werden und die Integrationspräferenz, sowie die vorgefundenen Muster nach einer Integration artifizieller TRE5-Retrotransposons, den endogenen TRE5-Elementen entsprechen, zeigt, dass diese Elemente als experimentell zugängliches Modell für die endogenen TRE5-Elemente verwendet werden können. Das zeigt, dass die Situation im Referenzgenom eher den aktuellen Ist-Zustand widerspiegelt, als ein exaktes Abbild der gesamten Population darzustellen. Betrachtet man dahingegen die Gesamtpopulation, so ist wahrscheinlich mit einer recht heterogenen Besetzung der tRNA-Gene mit spezifisch integrierten Retrotransposons zu rechnen. Zu vermuten ist, dass sich die maximal mögliche „Beladung“ unterschiedlicher tRNA-Gene innerhalb einer Zelle aus der daraus resultierenden Genomstabilität ableitet (Oliver & Greene, 2009), und dass das der limitierende Faktor ist für die Amplifikation der TRE-Elemente ist. Eine Analyse der Dynamik von Integrationsereignissen ist daher nur durch die experimentelle Untersuchung von *De-novo*-Integrationen möglich, was in dieser Arbeit für die TRE5-A.2^{bsr}-Elemente von *D. discoideum* erstmals versucht wurde.

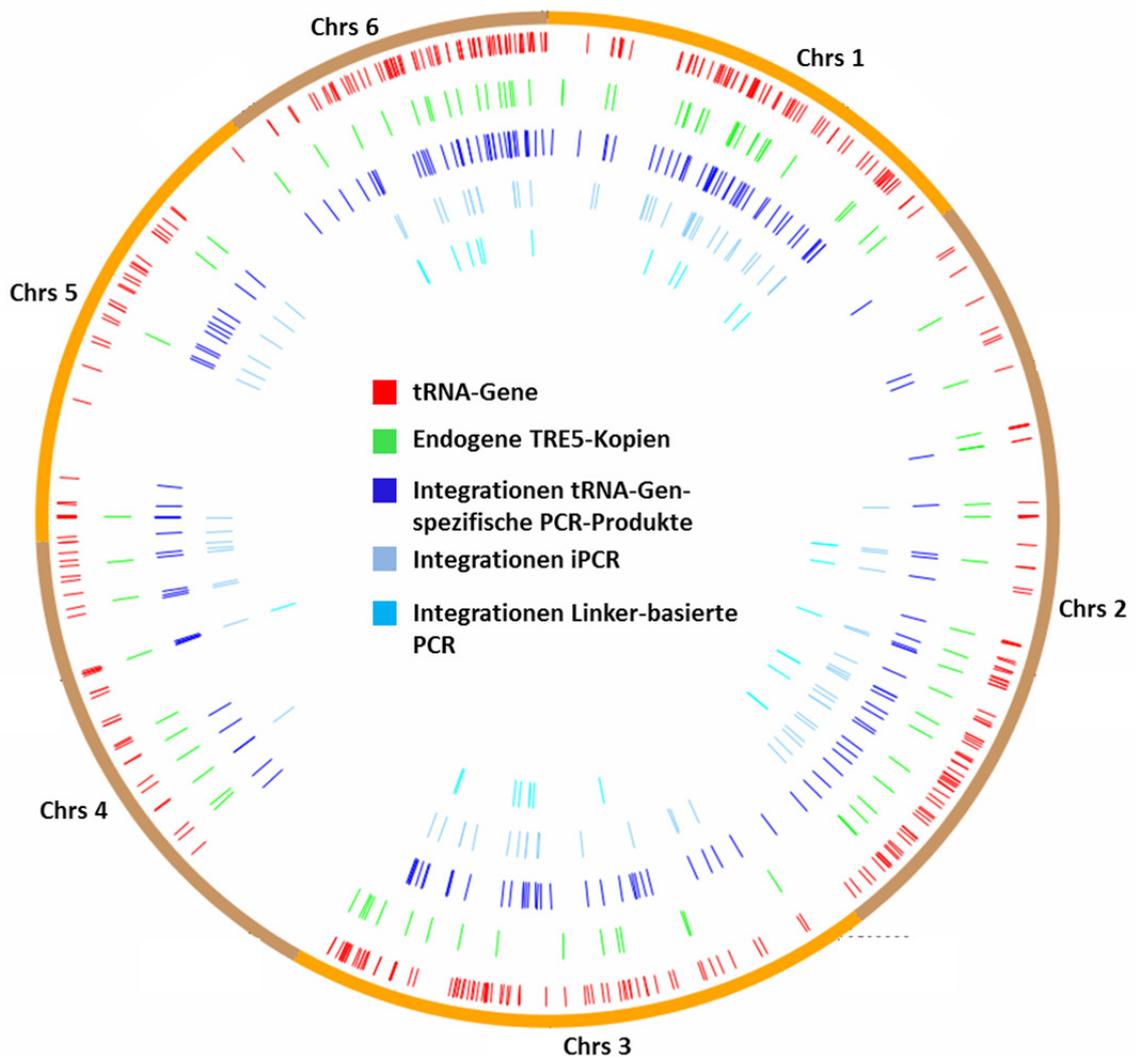


Abb 46: Vergleich der Sequenzier-Ergebnisse von tRNA-Gen-spezifischer und inverser PCR.

Dargestellt sind zum einen die tRNA-Gene (Rot) sowie die detektierbaren endogenen TRE5-A-Kopien (Grün) im Referenzgenom von *D. discoideum*. Dunkelblau markiert sind alle Positionen, die mit Hilfe der tRNA-Gen-spezifischen PCR detektiert wurden. Hellblau dargestellt sind die Loci, die durch die inverse PCR mit Integrationen des artifiziiellen Retrotransposons in Verbindung gebracht werden konnten. Die Ergebnisse der Linker-basierten Bank sind in Cyan dargestellt. Die Abbildung wurde mit Hilfe des Artemis-Genombrowsers und des DNAPlotter-Tools erzeugt (Rutherford *et al.* 2000; Carver *et al.* 2009).

4.6 Das artifizielle Retrotransposon TRE5-A.2^{bsr} als Werkzeug zur Identifizierung unbekannter RNA-Polymerase-III-Transkriptionskomplexe

Die Ergebnisse dieser Arbeit untermauern die Hypothese, dass TRE5-A nur solche Positionen als Integrationsort erkennt, die von TFIIC und TFIIB besetzt sind (Siol *et al.* 2006; Chung *et al.* 2007). Da TFIIB durch TFIIC rekrutiert wird und somit alle Voraussetzungen für die Assemblierung der RNA-Pol-III gegeben sind, ist davon auszugehen, dass alle von TRE5-A als Integrationsort erkannten Loci aktive Transkriptionseinheiten darstellen (Geiduschek und Kassavetis 1992). Mit der Identifizierung von Integrationen des TRE5-A.2^{bsr}-Elements an allen tRNA-Gen-ähnlichen A-/B-Box-Motiven auf dem extrachromosomalen Palindrom zeigt sich die Möglichkeit, mit Hilfe genetisch markierter TRE5-A-Elemente TFIIC/TFIIB-Bindestellen zu finden, welche nicht mit tRNA-Genen assoziiert sind. Der Nachweis von Transkripten, welche ausgehend von den jeweiligen Palindrom-Loci gebildet werden und die Integration sowohl endogener, als auch artifizieller TRE5-Elemente an diesen Positionen, kann als Hinweis auf eine Beteiligung des kompletten RNA-Polymerase-III-Transkriptionskomplexes gedeutet werden. Mit der Identifizierung von Integrationen innerhalb des DDT-S-Transposons und des Palindroms zeigt sich die potentielle Möglichkeit, durch die Detektion dieser Ereignisse die Bindestellen von RNA-Pol-III-Transkriptionsfaktoren zu bestimmen. Weiterführende Experimente sollten allerdings klären, ob diese Loci wirklich von diesen Faktoren erkannt und gebunden werden. In Frage kommt dabei zum Beispiel die Chromatin-Immunopräzipitation (Collas 2010), bei der Proteinkomplexe an ihren Zielsequenzen fixiert und anschließend die Positionen identifiziert werden können. Dies sollte vor allem zur Verifizierung der neu identifizierten Integrationsstellen von TRE5-A im Palindrom und in DDT-S-Elementen dienen, da die hier gezeigten Ergebnisse nur als ein Hinweis auf eine Beteiligung von RNA-Polymerase-III-assoziierten Transkriptionsfaktoren gelten können. Zudem sollte geklärt werden, inwieweit an bestimmten Positionen nur einzelne RNA-Pol-III-Faktoren binden und dann keine Transkription, wie auf dem Palindrom, nachweisbar ist. Mit TFIIC-assoziierte Loci, bei denen es keine Hinweise auf eine Assemblierung weiterer RNA-Pol-III-Transkriptionsfaktoren gibt, wurden zum Beispiel für die Hefe *Schizosaccharomyces pombe* beschrieben (Moqtaderi und Struhl 2004). Eine solche Beobachtung wäre ein Indiz für nicht-transkriptionelle Funktionen der RNA-Pol-III-Komplexe, welche in Abschnitt 4.9. näher erläutert werden sollen. Die hier mit artifiziellen TRE5-A.2^{bsr}-Elemente erzielten Ergebnisse haben insofern zwei prinzipiell neue mögliche Pol-III-Transkriptionseinheiten detektiert: B-Box-Loci auf dem extrachromosomalen Palindrom und DDT-S-Elemente.

4.7 Integrationen von TRE5-Retrotransposons in DDT-S-Elemente

DDT-S-Elemente liegen im Genom, zumeist in Clustern mit anderen mobilen genetischen Elementen, in rund 35 Kopien vor (Glöckner *et al.* 2001). Sie besitzen keine codierende Kapazität, allerdings haben ihre invertierten Sequenzwiederholungen Ähnlichkeiten mit denen von DDT-A und DDT-B (Glöckner *et al.* 2001), was vermuten lässt, das DDT-S als nicht-autonome DNA-Transposons über einen „cut and paste“-Mechanismus *in trans* mobilisiert wird.

Im Zuge der Sequenzierung von neuen Integrationsereignissen des artifiziiellen TRE5-A.2^{bsr}-Elements wurden auch Integrationen in DDT-S-Elemente identifiziert, was nach unserer Hypothese auf funktionelle RNA-Polymerase-III-Komplexe in DDT-S hinweist. Dabei ist noch unklar, ob die DDT-S-Elemente auch von der RNA-Polymerase-III transkribiert werden. Dennoch scheint unsere Beobachtung somit in Kontrast zur bisherigen Annahme zu stehen, dass die DNA-Transposons ausschließlich von der RNA-Polymerase-II erkannt werden (Sims III *et al.* 2004). Vor allem der exakte Abstand der TRE5-A-Integrationen innerhalb von DDT-S in Bezug zur theoretischen +1 Position vor der A-Box (vgl. Abschnitt 3.2.8.), bezogen auf die Situation in tRNA-Genen, zeigt allerdings, wie konserviert die Interaktion zwischen den TFIIC- und TFIIB-Untereinheiten und den Proteinen des Prä-Integrationskomplexes von TRE5-A in DDT-S sein muss. Dies legt nahe, dass Teile von DDT-S in der Tat durch Pol-III transkribiert werden. Welche Bedeutung diese Transkription für die Biologie des Elementes hat, bleibt allerdings unklar. Eine attraktive Hypothese wäre, dass die DDT-S-assoziierten Bindestellen, analog zur Situation der B-Boxen auf dem extrachromosomalen Palindrom, für Pol-III-Transkriptionsfaktoren Ankerpunkte für die subnukleäre Chromatin-Organisation schaffen (siehe Abschnitt 4.9). Da DDT-S ein mobiles Element darstellt, wären diese Ankerpunkte ebenfalls mobil und würden der Zelle eventuell evolutionäre Vorteile durch eine erhöhte Flexibilität der Organisation der Genexpression unter bestimmten Selektionsbedingungen verschaffen.

4.8 Hinweise auf die Replikation des extrachromosomalen Palindroms

Das Palindrom macht mit rund 100 Kopien pro Zelle ca. 25 % des Gesamtgenoms von *D. discoideum* aus (Sucgang *et al.* 2003). Bislang ist allerdings noch nicht genau verstanden wie diese Moleküle repliziert werden. Eine Hypothese postuliert eine Masterkopie auf Chromosom 4 (Sucgang *et al.* 2003). In dem Paper wurde die mögliche Position der Masterkopie anhand genetischer Marker an einer Stelle postuliert, die sich nach der Assemblierung von Chromosom 4 im Rahmen des *Dictyostelium*-Genomprojektes (Eichinger *et al.* 2005) als falsch herausgestellt hat. Gegen eine Masterkopie als Quelle der Amplifikation des extrachromosomalen Palindroms sprechen auch die hier dargestellten Ergebnisse. Mit dem Nachweis von Integrationen sowohl oberhalb von tRNA-Gen-ähnlichen A- und B-Box-Motiven, als auch dem 5S-rDNA-Gen stellt sich die Frage, warum es im Falle des Vorhandenseins einer Masterkopie nicht zu einer Akkumulation von integrierten Retrotransposons in dieser „Masterkopie“ des Palindroms gekommen sein soll? Da davon auszugehen ist, dass auch eine vermeintliche Masterkopie als Ziel von spezifisch integrierenden Retrotransposons erkannt werden müsste, wäre zu vermuten, dass alle daraus resultierenden „freien“ Palindrome ebenfalls Zeichen dieser Integrationen aufweisen würden. Die genomischen Sequenzdaten geben allerdings keine Hinweise darauf. Die Ergebnisse der vorliegenden Arbeit legen vielmehr eine relativ rasche Elimination von Integrationsereignissen aus der Population von Palindrom-Molekülen nahe. Das spricht für einen schnellen Abbau der Einzelmoleküle. Geht man davon aus, dass nur selten Integrationen in die jeweiligen Palindrom-Loci stattfinden, gäbe es zu jedem Zeitpunkt innerhalb der Gesamtpopulation nur einen kleinen Teil von Molekülen, die diese neuen Merkmale tragen. Somit

wäre eine Fixierung nur sehr unwahrscheinlich und ein Großteil der Palindrom-Moleküle würde nach wie vor keine Anzeichen einer Integration von Retrotransposons aufweisen. Die verbleibenden Palindrom-Moleküle, welche nach wie vor Integrationsereignisse akkumuliert hätten, würden im Zuge einer genomweiten Sequenzierung wahrscheinlich nicht auffallen, da die Transposon-Sequenzen bei der Assemblierung der Sequenzdaten eliminiert werden. Die vorgestellte Hypothese würde sich allerdings nur stützen lassen, wenn man postuliert, dass die Palindrome als extrachromosomale genetische Einheiten unabhängig von einer genomischen Masterkopie repliziert werden. Die vorliegenden Ergebnisse stützen diese Vermutung, wobei offen bleibt, wie der Mechanismus der Replikation aussehen könnte.

4.9 RNA-Polymerase-III-Transkriptionskomplexe: Einfluss auf globale Chromatinstrukturen

Neben der offensichtlichen Funktion der RNA-Polymerase-III-Transkriptionskomplexe, der Synthese der RNA von Pol-III transkribierten Genen, gibt es Hinweise auf weitere, sogenannte extratranskriptionelle Funktionen von RNA-Pol-III-Transkriptionskomplexen. Vor allem der Transkriptionsfaktor TFIIC scheint dabei eine besondere Rolle zu spielen. Bekannte nicht-transkriptionelle Einflüsse des RNA-Pol-III-Transkriptionskomplexes, oder Teilen davon, umfassen unter anderem (Donze 2012; Pascali und Teichmann 2013) (Abbildung 47):

- Inhibition und Beeinflussung der Transkription von nahegelegenen RNA-Polymerase II-Genen
- Veränderung der Positionierung von Nucleosomen
- Markierung von Positionen für Pausen während der DNA-Replikation
- Einfluss auf die Genomorganisation

Daraus ergeben sich eine ganze Reihe unterschiedlicher Funktionen, welche zusammengenommen zu einer globalen Bedeutung des RNA-Pol-III-Transkriptionskomplexes für die Chromatin-Organisation führen. Dabei scheint sich abzuzeichnen, dass vor allem die Organisation der genomischen DNA innerhalb des Zellkerns, sowie der Pol-I- und Pol-III-Transkription im oder nahe des Nukleolus maßgeblich durch RNA-Pol-III-Komplexe beeinflusst zu sein scheint (Pascali und Teichmann 2013). So konnte für Hefezellen gezeigt werden, dass die tRNA-Gene vornehmlich in der Nähe der tandemartig im Nucleolus arrangierten ribosomalen 5S-Gene lokalisiert sind, unabhängig von deren chromosomaler Position (Wang *et al.* 2005). Bekannt ist des Weiteren, dass die Anzahl von Bindestellen, an denen nur der Transkriptionsfaktor TFIIC nachweisbar ist (Extra-TFIIC-Loci), die Anzahl der Loci übertrifft, die mit der RNA-Polymerase-III in humanen Zellen und in Hefezellen assoziiert sind (Simms *et al.* 2008; Moqtaderi *et al.* 2010). Dabei scheint es eine Korrelation von spezifischen Histon-Modifikationen mit dem Vorhandensein der Extra-TFIIC-Loci zu geben (Moqtaderi *et al.* 2010).

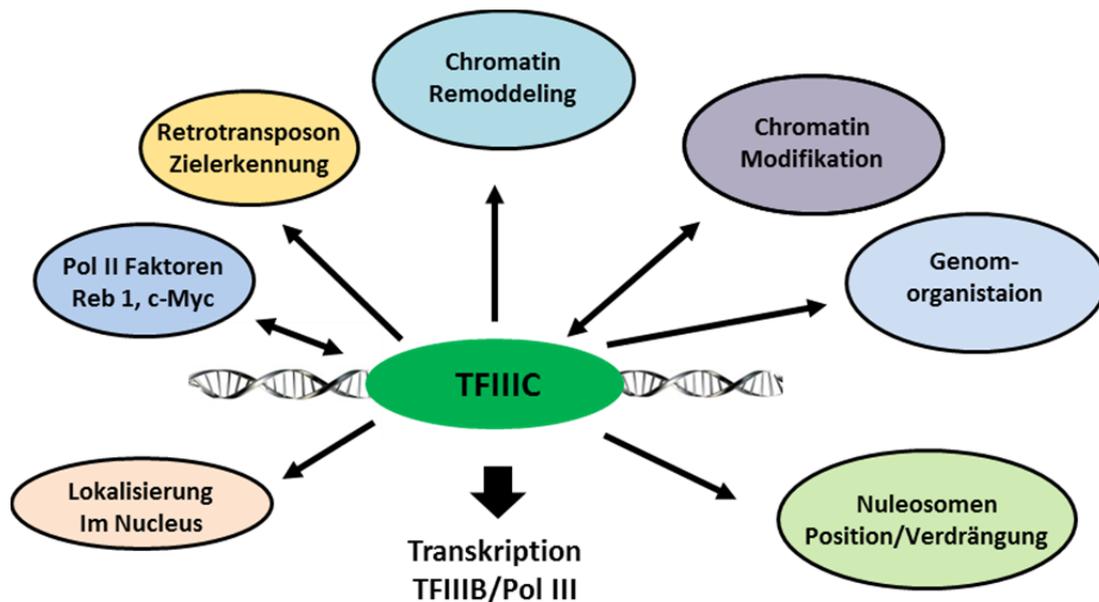


Abb 47: Extra-transkriptionelle Einflüsse von TFIIC.

Neben der Assemblierung des Transkriptionskomplexes der RNA-Polymerase-III besitzt der Transkriptionsfaktor TFIIC weitere Eigenschaften, die genomweite Einflüsse haben können. Modifiziert nach Donze 2012.

Somit stellt sich die Frage nach der Bedeutung ähnlicher Elemente im Genom von *D. discoideum*. Mit den hier dargestellten Ergebnissen zeigt sich die Möglichkeit, ein Retrotransposon zu verwenden, um bislang unbekannte Pol-III-Transkriptionskomplexe, beziehungsweise einzelne damit assoziierte Transkriptionsfaktoren zu detektieren. Vor allem die Loci auf den extrachromosomalen Palindromen, welche mit TRE5-Retrotransposons assoziiert sind, lassen die Vermutung zu, dass diese mit TFIIC assoziierten Positionen eventuell einen Einfluss auf die gesamte Genom-Architektur besitzen. Wie zuvor beschrieben, ist es daher denkbar, dass die extrachromosomal organisierten rDNA-Gene, analog zur Situation in Hefezellen (Wang *et al.* 2005), möglicherweise subnukleäre Kompartimente bilden, die eine Ko-Expression von RNA-Pol-I- und Pol-III-transkribierten ribosomalen Genen, unter der Einbeziehung von tRNA-Genen ermöglichen.

4.10 Genetisch markierte, autonome Retrotransposons

Mit der Erzeugung eines genetisch markierten, autonomen TRE5-A.1-Retrotransposons wurde die Grundlage für weiterführende Experimente geschaffen. Anders als das bislang verwendete artifizielle TRE5-A.2^{bsr}-Element besitzt das erzeugte autonome Element den kompletten ORF2 der endogenen TRE5-A.1-Elemente, der für die an der Retrotransposition direkt beteiligten Proteine codiert. Durch den in dieser Arbeit erbrachten Nachweis einer *Cis*-Präferenz des ORF2-Proteins in TRE5-A.1^{bsr}-Elementen können direkt Effekte dieser Proteine gemessen werden. Dies ermöglicht die Untersuchung von benötigten Strukturen bei der Mobilisierung von TRE5-A.1, zum Beispiel durch die Veränderung von Domänen des ORF2-Proteins.

Für die weitere Entwicklung der Erforschung von spezifisch integrierenden Retrotransposons in den Genomen sozialer Amöben ist es von großem Interesse mit aktiv mobilisierten, genetisch markierten LTR-Retrotransposons zu arbeiten. Bislang scheiterten alle Bemühungen ein solches Element auf Grundlage des DGLT-A LTR-Retrotransposons aus *D. discoideum* zu generieren. Daher sollte in dieser Arbeit, ausgehend von den Skipper-Elementen aus *D. discoideum*, welche in kompletten Kopien im Genom vorhanden sind, geklärt werden, ob es überhaupt möglich ist ein LTR-Retrotransposon genetisch zu markieren und dabei die Funktionsfähigkeit aufrecht zu erhalten.

Es ist gelungen, ein aktives, genetisch markiertes Skipper-Element zu erzeugen. Es konnte gezeigt werden, dass die Mobilisierung nur in RNAi-Mutanten stattfinden kann. Damit ist ein Werkzeug geschaffen worden, das es ermöglicht, in RNAi-Mutanten nicht nur Veränderungen der Transkription entsprechender Elemente nachzuweisen, sondern auch deren Mobilisierung direkt zu beweisen. Offensichtlich werden die LTR-Retrotransposons effektiv in den Wildtyp-Zellen inhibiert, so dass hier keine De-novo-Retrotranspositionseignisse nachweisbar sind. Frühere Experimente zeigten bereits, dass Änderungen des globalen Methylierungsstatus zu einer Verstärkung der Expression von endogenen Skipper-Retrotransposons führen (Kuhlmann *et al.* 2005). Es ist daher nicht erstaunlich, dass das genetisch markierte Skipper-Element in einer Rrp-AC⁻-Mutante auch tatsächlich aktiv mobilisiert wird. Somit findet in Wildtyp-Zellen offensichtlich eine sehr effektive post-transkriptionelle Regulation von mobilen genetischen Elementen statt (Wiegand *et al.* 2013), welche es verhindert neue Integrationsereignisse von genetisch markierten LTR-Retrotransposons nachzuweisen. Mit der Beobachtung, dass das genetisch markierte Skipper-Retrotransposon aktiv mobilisiert wird, konnte die Frage geklärt werden, an welcher Position die Retrotranspositionskassette in einem LTR-Retrotransposon integriert werden muss, ohne die Aktivität des Elements zu inhibieren.

In weiterführenden Experimenten müsste versucht werden, weitere geeignete genetische Marker zu integrieren, um die artifiziellen LTR-Retrotransposons von den Elementen der endogenen Population unterscheiden zu können. Denn anders als die Codon-adaptierten Non-LTR-Elemente der TRE5-Familie, bei denen eine Unterscheidung aufgrund der modifizierten Nukleotidsequenz möglich ist, besitzen die aktiv mobilisierten, genetisch markierten Skipper-Elemente bislang die authentische Sequenz der endogenen Skipper-Population. Für ein genomweites Mapping von Skipper-Integrationen wäre eine Unterscheidung dieser Elemente von den endogenen Elementen somit nur anhand der Retrotranspositionskassette möglich. Diese liegt allerdings rund 300 bp von der flankierenden genomischen Region entfernt, was ein zusätzliches Hindernis bei der Sequenzierung darstellt.

Eine mögliche Modifikation kann daher zum Beispiel die Erzeugung einer neuen Restriktionsschnittstelle innerhalb der LTR-Sequenzen sein. Für Ty1-Elemente konnte bereits gezeigt werden, dass eine Modifikation einer einzigen LTR-Sequenz möglich ist, ohne die Aktivität der Elemente zu inhibieren (Baller *et al.* 2012; Mularoni *et al.* 2012). Die neu erzeugte Restriktionsschnittstelle könnte anschließend in einem mobilisierten LTR-Retrotransposon für eine inverse PCR verwendet werden. Somit sollten nur PCR-Produkte entstehen, welche durch die Ligation

der neu erzeugten Restriktionsschnittstellen innerhalb des LTR-Retrotransposons mit flankierenden genomischen Schnittstellen ermöglicht wurden. Somit wäre eine Unterscheidung der artifiziellen Elemente von den genomischen Kopien der endogenen Population möglich.

4.11 Ausblick

Besonders die Sequenzierung und Assemblierung weiterer Genome sozialer Amöben sollte neue Erkenntnisse zur Evolution der tRNA-Gen-spezifischen Retrotransposons führen. Wichtig wäre dabei auch die Untersuchung von nahe mit den sozialen Amöben verwandten Spezies, wie zum Beispiel *Protostelium*, um zu klären, ob bereits in diesen Organismen TRE-Elemente nachweisbar sind. In diesem Zusammenhang könnte eventuell geklärt werden, ob die Entstehung der unterschiedlichen Integrationspräferenzen ober- und unterhalb der tRNA-Gene bereits außerhalb der sozialen Amöben entstanden ist. Interessant sind in diesem Zusammenhang allerdings auch Untersuchungen der Genome weiterer Vertreter innerhalb der Gruppe 4. Hier sollten Genomanalysen die Frage beantworten, ob es sich bei den in *D. discoideum* beschriebenen Verhältnissen um eine Ausnahme handelt, oder ob weitere Spezies existieren, die ähnlich hohe Anteile mobiler genetischer Elemente in ihren Genomen beherbergen.

Ein wesentlicher Fokus sollte des Weiteren auf der Etablierung weiterer genetisch markierter Retrotransposons liegen, die in einer Zellkultur eine nachweisbare Aktivität besitzen. Dabei scheint vor allem die Auswahl einer geeigneten *Dictyostelium*-Mutante entscheidend zu sein. Da eine mögliche Suppression der Retrotransposons durch die Zelle auf verschiedenen Ebenen stattfindet, wäre es denkbar, die zunächst nicht aktiven, artifiziellen Retrotransposons in unterschiedlichen Mutanten, welche Defekte in der (post-) transkriptionellen Regulationsmaschinerie aufweisen, zu untersuchen und auf eine mögliche Aktivität zu überprüfen. Mit den bereits etablierten, aktiven LTR-Retrotransposons sollten in Zukunft Experimente durchgeführt werden, welche auf Strukturelle Voraussetzungen für tRNA-Gen-spezifische Integrationen abzielen. Denkbar sind dabei Experimente, die zum Ziel haben durch strukturelle Modifikationen entweder die tRNA-Gen-Spezifität zu inhibieren, oder aber ein unspezifisch integrierendes Element mit einer solchen Fähigkeit auszustatten. Wichtige Erkenntnisse sind auch von der Arbeit mit dem neu generierten, autonomen und genetisch markierten TRE5-A.1-Retrotransposon zu erwarten. Zum einen werden Studien zur Bindung von RNA an das C-Modul und anderen, hypothetischen Strukturen von Interesse sein. Zum anderen wäre es interessant, Knock-Out Mutanten zu erzeugen, welche auf eine Elimination der tRNA-Gen-Spezifität bei einer gleichzeitigen Aufrechterhaltung der Mobilität des TRE5-A.1-Elements abzielen. Diese Experimente würden entscheidende Hinweise auf die beteiligten Strukturen bei der Interaktion zwischen dem Prä-Integrationskomplex des Retrotransposons und den Pol-III-Transkriptionsfaktoren geben.

Mit Blick auf das extrachromosomale Palindrom wäre es vorstellbar, mit Hilfe der dort integrierenden Retrotransposons den bislang nicht verstandenen Mechanismus der Replikation dieser speziellen DNA-Einheiten nachzuvollziehen. Dazu könnte zum Beispiel versucht werden, ein integriertes

Retrotransposon nachzuverfolgen, um so den „*Turn-Over*“ zu bestimmen mit dem die einzelnen Palindrome gebildet werden. Wichtig sind in diesem Zusammenhang auch Experimente zur Lokalisation der Palindrome innerhalb des Nucleus während unterschiedlicher Zellzyklus-Phasen. Geeignet wären für diese Untersuchungen zum Beispiel „Fluoreszenz-in-situ-Hybridisierungs“-Experimente (FISH), unter Verwendung von Palindrom-spezifischen Sonden. Denkbar sind in dieser Hinsicht auch FISH-Experimente mit einer TRE5-spezifischen Sonde, welche zeigen sollten, wie viele Palindrom-Moleküle zu einem bestimmten Zeitpunkt innerhalb einer Zelle mit Retrotransposon-Kopien besetzt sind. Die Aufklärung einer vermeintlichen Funktion der neu beschriebenen Palindrom-spezifischen Transkripte stellt eine weitere Fragestellung dar, die verbunden werden sollte mit Experimenten, welche beweisen, dass tatsächlich die RNA-Polymerase-III für die Transkription dieser Loci verantwortlich ist. Eine solche Fragestellung könnte zum Beispiel durch die Methode der Chromatin-Immunopräzipitation beantwortet werden.

5. Zusammenfassung

Die Genome der sozialen Amöben sind äußerst kompakt organisiert. Daher kann vermutet werden, dass mobile genetische Elemente in diesen Genomen besondere Anpassungen benötigen, um Insertionsmutagenesen zu vermeiden. Bekannt war bisher, dass Retrotransposons in der Spezies *Dictyostelium discoideum* die nicht-codierenden Bereiche in der Nähe von tRNA-Genen als "sichere" Integrationsorte aufsuchen. In der vorliegenden Arbeit wurden erstmals die Genome der sozialen Amöben *D. purpureum*, *Polysphondylium pallidum* und *D. fasciculatum* nach mobilen genetischen Elementen durchsucht. In allen analysierten Genomen konnten tRNA-Gen-spezifisch integrierende Retrotransposons nachgewiesen werden. Die gefundenen Anteile der Elemente am Gesamtgenom liegen dabei weit unter den Werten, welche für *D. discoideum* beschrieben wurden. Daher kann ausgeschlossen werden, dass die Expansion tRNA-Gen-spezifischer Retrotransposons mit der Amplifikation von tRNA-Genen in den jeweiligen Genomen einhergeht. Vielmehr scheint ein Gleichgewicht zwischen der Aktivität der mobilen genetischen Elemente und der damit verbundenen Genominstabilität der limitierende Faktor für die Expansion der Retrotransposons in den jeweiligen Genomen zu sein.

Mit der Identifizierung von LTR-Retrotransposons in allen untersuchten Genomen und einer anschließenden strukturellen und phylogenetischen Analyse dieser Elemente konnte gezeigt werden, dass innerhalb der sozialen Amöben ein Großteil der Elemente in zwei Familien eingeteilt werden kann: die DGLT-A- und die Skipper-Familie. Mit der Identifizierung eines LTR-Retrotransposons im Genom von *D. fasciculatum*, welches strukturelle Merkmale beider LTR-Retrotransposon-Familien trägt, könnte ein Vorläufer beider Element-Typen gefunden worden sein.

In Vorarbeiten wurde ein TRE5-A-Retrotransposon mit einem Selektionsmarker ausgestattet, der es erlaubt, die Amplifikation solcher Elemente *in vivo* zu verfolgen. Mit der Generierung eines aktiven, Codon-adaptierten TRE5-A-Retrotransposons war es erstmals möglich, artifizielle, mit dieser Selektionskassette versehene Elemente von der endogenen TRE5-A-Population zu unterscheiden. In der vorliegenden Arbeit wurden Assays etabliert, die es ermöglichen, genomweit De-Novo-Integrationen des artifiziellen Retrotransposons zu analysieren. So konnte gezeigt werden, dass ein großer Teil der genomischen tRNA-Gen-Kopien von TRE5-A-Elementen als Integrationsziel erkannt wird. Des Weiteren wurden Integrationen des artifiziellen Retrotransposons auf dem extrachromosomalen Palindrom untersucht. Es zeigte sich, dass tRNA-Gen-ähnliche A- und B-Box-Motive als Ziel von TRE5-A erkannt werden und diese Loci aktiv transkribiert werden. Dies ist ein Hinweis auf eine Beteiligung des RNA-Polymerase-III-Transkriptionskomplexes, der an der Erkennung von Integrationszielen durch TRE5-A beteiligt ist. Mit der Generierung eines autonomen, genetisch markierten TRE5-A.1-Elements konnte gezeigt werden, dass eine *Cis*-Präferenz der ORF2-Proteine von TRE5-A für die eigenen Transkripte vorliegt und so die Retrotransposition der autonomen mobilen genetischen Elemente unterstützt.

Unerwartet wurden Integrationen des artifiziellen TRE5-A-Retrotransposons in das repetitive Element DDT-S identifiziert. In diesem Zusammenhang wurden tRNA-Gen-ähnliche A- und B-Box-Motive innerhalb der genomischen DDT-S-Sequenzen identifiziert, welche nahelegen, dass diese Transposons vom RNA-Polymerase-III-Transkriptionskomplex erkannt werden und die tRNA-Gen-ähnliche Topologie der beteiligten Transkriptionsfaktoren auf DDT-S-Elementen für die charakteristischen TRE5-A-Integrationen verantwortlich sind.

In der vorliegenden Arbeit wurden autonome, genetisch markierte LTR-Retrotransposons erzeugt, die aktiv in *Dictyostelium*-Zellen mobilisiert werden. Damit konnte geklärt werden, an welcher Position solcher Elemente ein Retrotranspositionsmarker eingebaut werden kann, ohne die Aktivität der Retrotransposons zu inhibieren. Allerdings scheint die zellulären RNA-Interferenz-Maschinerie die mobilen genetischen Elemente so stark zu inhibieren, dass eine Mobilisierung nur in Zellen mit defekter Abwehr möglich ist. Möglicherweise ist das der Grund warum es nicht möglich war ein genetisch markiertes, tRNA-Gen-spezifisches LTR-Retrotransposon aus *D. fasciculatum* in einer *D. discoideum*-Zellkultur zu mobilisieren. Es kann allerdings nicht ausgeschlossen werden, dass die Mobilisierung der inaktiven Elemente in bislang nicht berücksichtigten Mutanten möglich ist.

6 Summary

Social amoebae have very compact genomes. Therefore, it is likely that tRNA-gene-specific integration in the genome of the social amoeba *Dictyostelium discoideum* is an adaptation to avoid insertional mutagenesis. In this study, a survey of mobile genetic elements was performed in the recently sequenced species *D. purpureum*, *Polysphondylium pallidum* and *D. fasciculatum* with emphasis on tRNA gene-targeting retrotransposons. It was assumed that analysis of several compact genomes could reveal specific integration as a common means to avoid instability of compact genomes caused by retrotransposon activity. Considering phylogenetic relations between different species, the detection of tRNA gene-specific elements in all species indicates even the last common ancestor of social amoebae possessed such elements within their genomes. Structural and phylogenetic analysis of the newly identified LTR-retrotransposons indicates the presence of two different families of elements within the social amoebae: the DGLT-A and the Skipper family. Interestingly, an element from *D. fasciculatum* shows features of both families; it could represent a precursor of the Skipper and DGLT-A retrotransposons.

In previous work a TRE5-A retrotransposon was generated that contained a selection marker to detect new retrotransposition events. Here, the generation of a codon-adapted, tagged TRE5-A element enabled the discrimination of these tagged retrotransposons from the endogenous TRE5-A population. This was a prerequisite to establish protocols for the genome-wide search for de novo TRE5-A integrations in the *D. discoideum* genome. It was found that most tRNA genes of *D. discoideum* are principle targets for TRE5-A integration. Furthermore, it was possible to generate an autonomously amplifying, genetically tagged and full-length TRE5-A.1-retrotransposon. A comparison of retrotransposition frequencies indicates a *cis*-preference of the TRE5-A ORF2-protein for its own RNA.

New TRE5-A integrations were observed to occur into the extrachromosomal rDNA palindrome that carries the ribosomal RNA genes. It was shown that, distinct A- and B-box-motives on the palindrome, which have strong similarities with the features of a tRNA-gene, are recognized by the artificial TRE5-A-retrotransposons. Furthermore, TRE5-A integrations upstream of the ribosomal 5S gene were observed. The authentic distances of integrated TRE5-A copies to the targeted B box motifs and the 5S gene indicates an involvement of RNA-polymerase-III-transcription units in integration site recognition. It was shown that the targeted B box loci are actively transcribed, which argues for an involvement of the RNA polymerase III complex in target site selection by TRE5-A.

In this work genetically tagged LTR retrotransposons were developed and characterized. It could be shown that a selection marker for retrotransposition events can be introduced into the elements without disruption of transposon activity. Interestingly, it seemed that such tagged retrotransposons were only mobilized in *D. discoideum* mutants with a defective RNA interference machinery.

Literaturverzeichnis

- Adl, S. M., A. G. Simpson, C. E. Lane, *et al.* (2012). The revised classification of eukaryotes. *J Eukaryot Microbiol* **59**(5): 429-493.
- Agrawal, A., Q. M. Eastman and D. G. Schatz (1998). Transposition mediated by RAG1 and RAG2 and its implications for the evolution of the immune system. *Nature* **394**(6695): 744-751.
- Ahn, K., J. A. Gim, H. S. Ha, *et al.* (2013). The novel MER transposon-derived miRNAs in human genome. *Gene* **512**(2): 422-428.
- Altschul, S. F., W. Gish, W. Miller, *et al.* (1990). Basic local alignment search tool. *Journal of Molecular Biology* **215**(3): 403-410.
- Arias, J. F., T. Koyama, M. Kinomoto, *et al.* (2012). Retroelements versus APOBEC3 family members: No great escape from the magnificent seven. *Front Microbiol* **3**: 275.
- Arimbasseri, A. G., K. Rijal and R. J. Maraia (2013). Transcription termination by the eukaryotic RNA polymerase III. *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms* **1829**(3): 318-330.
- Arkhipova, I. R., A. M. Mazo, V. A. Cherkasova, *et al.* (1986). The steps of reverse transcription of *Drosophila* mobile dispersed genetic elements and U3-R-U5 structure of their LTRs. *Cell* **44**(4): 555-563.
- Asghar, A., M. Groth, O. Siol, *et al.* (2012). Developmental gene regulation by an ancient intercellular communication system in social amoebae. *Protist* **163**(1): 25-37.
- Ashworth, J. M. and D. J. Watts (1970). Metabolism of the cellular slime mould *Dictyostelium discoideum* grown in axenic culture. *Biochem J* **119**(2): 175-182.
- Attard, A., L. Gout, S. Ross, *et al.* (2005). Truncated and RIP-degenerated copies of the LTR retrotransposon Pholy are clustered in a pericentromeric region of the *Leptosphaeria maculans* genome. *Fungal Genet Biol* **42**(1): 30-41.
- Badge, R. M., R. S. Alisch and J. V. Moran (2003). ATLAS: a system to selectively identify human-specific L1 insertions. *The American Journal of Human Genetics* **72**(4): 823-838.
- Baller, J. A., J. Gao, R. Stamenova, *et al.* (2012). A nucleosomal surface defines an integration hotspot for the *Saccharomyces cerevisiae* Ty1 retrotransposon. *Genome Research* **22**(4): 704-713.
- Barlow, D. P. (1993). Methylation and imprinting: from host defense to gene regulation? *Science* **260**: 309-310.
- Batzer, M. A. and P. L. Deininger (2002). Alu repeats and human genomic diversity. *Nat Rev Genet* **3**(5): 370-379.
- Beck, P., T. Dingermann and T. Winckler (2002). Transfer RNA Gene-targeted Retrotransposition of *Dictyostelium* TRE5-A into a Chromosomal UMP Synthase Gene Trap. *Journal of Molecular Biology* **318**(2): 273-285.
- Benachou, F., P. Jern, M. Oja, *et al.* (2009). Evolutionary conservation of orthoretroviral long terminal repeats (LTRs) and ab initio detection of single LTRs in genomic data. *PLoS One* **4**(4): e5179.
- Bennetzen, J. L., J. Ma and K. M. Devos (2005). Mechanisms of recent genome size variation in flowering plants. *Ann Bot* **95**(1): 127-132.
- Benson, G. (1999). Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Res* **27**(2): 573.
- Biémont, C. and C. Vieira (2006). Genetics: junk DNA as an evolutionary force. *Nature* **443**(7111): 521-524.
- Boeke, J. and V. G. Corces (1989). Transcription and reverse transcription of retrotransposons. *Annual Reviews in Microbiology* **43**(1): 403-434.
- Boeke, J. D. and S. E. Devine (1998). Yeast retrotransposons: finding a nice quiet neighborhood. *Cell* **93**(7): 1087-1089.
- Bonner, J. T. (1952). The pattern of differentiation in amoeboid slime molds. *American Naturalist*: 79-89.
- Bonner, J. T. (1959). The cellular slime molds. *The cellular slime molds*.
- Bonner, J. T., D. S. Barkley, E. M. Hall, *et al.* (1969). Acrasin, Acrasinase, and the sensitivity to acrasin in *Dictyostelium discoideum*. *Dev Biol* **20**(1): 72-87.

- Brady, T. L., C. L. Schmidt and D. F. Voytas** (2008). Targeting integration of the *Saccharomyces Ty5* retrotransposon. *Chromosomal Mutagenesis*, Springer: 153-163.
- Brandt, J., S. Schrauth, A. M. Veith, et al.** (2005). Transposable elements as a source of genetic innovation: expression and evolution of a family of retrotransposon-derived neogenes in mammals. *Gene* **345**(1): 101-111.
- Brefeld, O.** (1869). *Dictyostelium mucoroides*: ein neuer Organismus aus der Verwandtschaft der Myxomyceten, C. Winter.
- Britten, R.** (2006). Transposable elements have contributed to thousands of human proteins. *Proc Natl Acad Sci U S A* **103**(6): 1798-1803.
- Brouha, B., J. Schustak, R. M. Badge, et al.** (2003). Hot L1s account for the bulk of retrotransposition in the human population. *Proc Natl Acad Sci U S A* **100**(9): 5280-5285.
- Buchon, N. and C. Vauray** (2006). RNAi: a defensive RNA-silencing against viruses and transposable elements. *Heredity (Edinb)* **96**(2): 195-202.
- Bukenberger, M., T. Dingermann, W. Meissner, et al.** (1994). Isolation of transcription factor III C from *Dictyostelium discoideum*. *European Journal of Biochemistry* **220**(3): 839-846.
- Bukenberger, M., R. Marschalek, W. Salzburger, et al.** (1991). Nuclear factors which bind to *Dictyostelium discoideum* transfer RNA genes. *Current Genetics* **20**(1-2): 129-135.
- Burdine, V. and M. Clarke** (1995). Genetic and physiologic modulation of the prestarvation response in *Dictyostelium discoideum*. *Molecular biology of the cell* **6**(3): 311-325.
- Burke, W., C. Calalang and T. Eickbush** (1987). The site-specific ribosomal insertion element type II of *Bombyx mori* (R2Bm) contains the coding sequence for a reverse transcriptase-like enzyme. *Mol Cell Biol* **7**(6): 2221-2230.
- Callinan, P. and M. Batzer** (2006). Retrotransposable elements and human disease.
- Cappello, J., K. Handelsman and H. F. Lodish** (1985). Sequence of *Dictyostelium* DIRS-1: an apparent retrotransposon with inverted terminal repeats and an internal circle junction sequence. *Cell* **43**(1): 105-115.
- Capy, P., T. Langin, D. Higuete, et al.** (1997). Do the integrases of LTR-retrotransposons and class II element transposases have a common ancestor? *Genetica* **100**(1-3): 63-72.
- Carver, T., N. Thomson, A. Bleasby, et al.** (2009). DNAPlotter: circular and linear interactive genome visualization. *Bioinformatics* **25**(1): 119-120.
- Castaño, J. G., R. Ornberg, J. G. Koster, et al.** (1986). Eukaryotic pre-tRNA 5' processing nuclease: copurification with a complex cylindrical particle. *Cell* **46**(3): 377-387.
- Cavalier-Smith, T.** (1998). A revised six-kingdom system of life. *Biol Rev Camb Philos Soc* **73**(3): 203-266.
- Cecco, M., S. W. Criscione, E. J. Peckham, et al.** (2013). Genomes of replicatively senescent cells undergo global epigenetic changes leading to gene silencing and activation of transposable elements. *Aging Cell* **12**(2): 247-256.
- Chalker, D. L. and S. B. Sandmeyer** (1992). Ty3 integrates within the region of RNA polymerase III transcription initiation. *Genes & Development* **6**(1): 117-128.
- Champoux, J. J.** (1993). 6 Roles of Ribonuclease H in Reverse Transcription. *Cold Spring Harbor Monograph Archive* **23**: 103-117.
- Charlesworth, B., P. Sniegowski and W. Stephan** (1994). The evolutionary dynamics of repetitive DNA in eukaryotes.
- Chenais, B., A. Caruso, S. Hiard, et al.** (2012). The impact of transposable elements on eukaryotic genomes: from genome size increase to genetic adaptation to stressful environments. *Gene* **509**(1): 7-15.
- Chinwalla, A. T., L. L. Cook, K. D. Delehaunty, et al.** (2002). Initial sequencing and comparative analysis of the mouse genome. *Nature* **420**(6915): 520-562.
- Chung, T., O. Siol, T. Dingermann, et al.** (2007). Protein interactions involved in tRNA gene-specific integration of *Dictyostelium discoideum* non-long terminal repeat retrotransposon TRE5-A. *Mol Cell Biol* **27**(24): 8492-8501.
- Collas, P.** (2010). The current state of chromatin immunoprecipitation. *Molecular biotechnology* **45**(1): 87-100.
- Cordaux, R. and M. A. Batzer** (2009). The impact of retrotransposons on human genome evolution. *Nat Rev Genet* **10**(10): 691-703.

- Cotter, D. A., T. W. Sands, K. J. Viridy, *et al.* (1992). Patterning of development in Dictyostelium discoideum: factors regulating growth, differentiation, spore dormancy, and germination. *Biochemistry and Cell Biology* **70**(10-11): 892-919.
- Craig, N. L., R. Craigie, M. Gellert, *et al.* (2002). Mobile dna ii, ASM Press Washington, DC.
- Crichton, J. H., D. S. Dunican, M. Maclennan, *et al.* (2014). Defending the genome from the enemy within: mechanisms of retrotransposon suppression in the mouse germline. *Cell Mol Life Sci* **71**(9): 1581-1605.
- de Koning, A. J., W. Gu, T. A. Castoe, *et al.* (2011). Repetitive elements may comprise over two-thirds of the human genome. *PLoS Genet* **7**(12): e1002384.
- Doak, T. G., F. P. Doerder, C. L. Jahn, *et al.* (1994). A proposed superfamily of transposase genes: transposon-like elements in ciliated protozoa and a common "D35E" motif. *Proceedings of the National Academy of Sciences* **91**(3): 942-946.
- Dombroski, B. A., S. L. Mathias, E. Nanthakumar, *et al.* (1991). Isolation of an active human transposable element. *Science* **254**(5039): 1805-1808.
- Donze, D. (2012). Extra-transcriptional functions of RNA Polymerase III complexes: TFIIC as a potential global chromatin bookmark. *Gene* **493**(2): 169-175.
- Dormann, D., F. Siegert and C. J. Weijer (1996). Analysis of cell movement during the culmination phase of Dictyostelium development. *Development* **122**(3): 761-769.
- Doucet, A. J., A. E. Hulme, E. Sahinovic, *et al.* (2010). Characterization of LINE-1 ribonucleoprotein particles. *PLoS Genet* **6**(10).
- Duncan, L., K. Bouckaert, F. Yeh, *et al.* (2002). kangaroo, a mobile element from Volvox carteri, is a member of a newly recognized third class of retrotransposons. *Genetics* **162**(4): 1617-1630.
- Eichinger, D. J. and J. D. Boeke (1988). The DNA intermediate in yeast Ty1 element transposition copurifies with virus-like particles: cell-free Ty1 transposition. *Cell* **54**(7): 955-966.
- Eichinger, L., J. A. Pachebat, G. Glöckner, *et al.* (2005). The genome of the social amoeba Dictyostelium discoideum. *Nature* **435**(7038): 43-57.
- Eickbush, D. G. and T. H. Eickbush (1995). Vertical transmission of the retrotransposable elements R1 and R2 during the evolution of the Drosophila melanogaster species subgroup. *Genetics* **139**(2): 671-684.
- Eickbush, T. H. (1997). Telomerase and retrotransposons: which came first? *Science* **277**(5328): 911-912.
- Eickbush, T. H. and V. K. Jamburuthugoda (2008). The diversity of retrotransposons and the properties of their reverse transcriptases. *Virus Res* **134**(1-2): 221-234.
- Esnault, C., J. Maestre and T. Heidmann (2000). Human LINE retrotransposons generate processed pseudogenes. *Nat Genet* **24**(4): 363-367.
- Ewing, A. D. and H. H. Kazazian (2010). High-throughput sequencing reveals extensive variation in human-specific L1 content in individual human genomes. *Genome Research* **20**(9): 1262-1270.
- Falzon, M. and E. Kuff (1988). Multiple protein-binding sites in an intracisternal A particle long terminal repeat. *J Virol* **62**(11): 4070-4077.
- Fanning, T. G. and M. F. Singer (1987). LINE-1: a mammalian transposable element. *Biochim Biophys Acta* **910**(3): 203-212.
- Feng, Q., J. V. Moran, H. H. Kazazian Jr, *et al.* (1996). Human L1 retrotransposon encodes a conserved endonuclease required for retrotransposition. *Cell* **87**(5): 905-916.
- Feschotte, C. and E. J. Pritham (2007). DNA transposons and the evolution of eukaryotic genomes. *Annu Rev Genet* **41**: 331.
- Fiore-Donno, A. M., S. I. Nikolaev, M. Nelson, *et al.* (2010). Deep phylogeny and evolution of slime moulds (mycetozoa). *Protist* **161**(1): 55-70.
- Freeze, H. and W. F. Loomis (1977). Isolation and characterization of a component of the surface sheath of Dictyostelium discoideum. *J Biol Chem* **252**(3): 820-824.
- Fuhrmann, M., A. Hausherr, L. Ferbitz, *et al.* (2004). Monitoring dynamic expression of nuclear genes in Chlamydomonas reinhardtii by using a synthetic luciferase reporter gene. *Plant molecular biology* **55**(6): 869-881.
- Gabriel, R., R. Eckenberg, A. Paruzynski, *et al.* (2009). Comprehensive genomic access to vector integration in clinical gene therapy. *Nature medicine* **15**(12): 1431-1436.

- Gao, X.** (2003). Translational recoding signals between gag and pol in diverse LTR retrotransposons. *Rna* **9**(12): 1422-1430.
- Gaudet, P., K. E. Pilcher, P. Fey, et al.** (2007). Transformation of Dictyostelium discoideum with plasmid DNA. *Nature protocols* **2**(6): 1317-1324.
- Geiduschek, E. P. and G. A. Kassavetis** (1992). RNA Polymerase III Transcription Complexes. *Cold Spring Harbor Monograph Archive* **22**: 247-280.
- George, J. A., W. D. Burke and T. H. Eickbush** (1996). Analysis of the 5' junctions of R2 insertions with the 28S gene: implications for non-LTR retrotransposition. *Genetics* **142**(3): 853-863.
- Gilbert, N., S. Lutz-Prigge and J. V. Moran** (2002). Genomic deletions created upon LINE-1 retrotransposition. *Cell* **110**(3): 315-325.
- Glöckner, G. and A. A. Noegel** (2013). Comparative genomics in the Amoebozoa clade. *Biological Reviews* **88**(1): 215-225.
- Glöckner, G., K. Szafranski, T. Winckler, et al.** (2001). The complex repeats of Dictyostelium discoideum. *Genome Res* **11**(4): 585-594.
- Goffeau, A., B. Barrell, H. Bussey, et al.** (1996). Life with 6000 genes. *Science* **274**(5287): 546-567.
- Goldbeter, A.** (2006). Oscillations and waves of cyclic AMP in Dictyostelium: a prototype for spatio-temporal organization and pulsatile intercellular communication. *Bull Math Biol* **68**(5): 1095-1109.
- Goodier, J. L., E. M. Ostertag and H. H. Kazazian Jr** (2000). Transduction of 3' -flanking sequences is common in L1 retrotransposition. *Hum Mol Genet* **9**(4): 653-657.
- Goodwin, T. J. and R. T. Poulter** (2001). The DIRS1 group of retrotransposons. *Mol Biol Evol* **18**(11): 2067-2082.
- Guo, Y. and H. L. Levin** (2010). High-throughput sequencing of retrotransposon integration provides a saturated profile of target activity in Schizosaccharomyces pombe. *Genome Res* **20**(2): 239-248.
- Han, K., S. K. Sen, J. Wang, et al.** (2005). Genomic rearrangements by LINE-1 insertion-mediated deletion in the human and chimpanzee lineages. *Nucleic Acids Res* **33**(13): 4040-4052.
- Hancks, D. C. and H. H. Kazazian, Jr.** (2012). Active human retrotransposons: variation and disease. *Curr Opin Genet Dev* **22**(3): 191-203.
- Hannon, G. J.** (2002). RNA interference. *Nature* **418**(6894): 244-251.
- Havecker, E. R., X. Gao and D. F. Voytas** (2004). The diversity of LTR retrotransposons. *GENOME BIOLOGY* **5**(6): 225-225.
- Hayflick, L.** (1965). The Limited in Vitro Lifetime of Human Diploid Cell Strains. *Exp Cell Res* **37**: 614-636.
- Heidel, A. J., H. M. Lawal, M. Felder, et al.** (2011). Phylogeny-wide analysis of social amoeba genomes highlights ancient origins for complex intercellular communication. *Genome Res* **21**(11): 1882-1891.
- Hentschel, U., I. Zündorf, T. Dingermann, et al.** (2001). On the Problem of Establishing the Subcellular Localization of Dictyostelium Retrotransposon TRE5-A Proteins by Biochemical Analysis of Nuclear Extracts. *Anal Biochem* **296**(1): 83-91.
- Heyman, T., M. Wilhelm and F. X. Wilhelm** (2003). The Central PPT of the Yeast Retrotransposon Ty1 is not Essential for Transposition. *Journal of Molecular Biology* **331**(2): 315-320.
- Hill, A. S., N. J. Foot, T. L. Chaplin, et al.** (2000). The most frequent constitutional translocation in humans, the t(11; 22)(q23; q11) is due to a highly specific Alu-mediated recombination. *Hum Mol Genet* **9**(10): 1525-1532.
- Hofmann, J., G. Schumann, G. Borschet, et al.** (1991). Transfer RNA genes from Dictyostelium discoideum are frequently associated with repetitive elements and contain consensus boxes in their 5' and 3' -flanking regions. *Journal of Molecular Biology* **222**(3): 537-552.
- Hohjoh, H. and M. F. Singer** (1996). Cytoplasmic ribonucleoprotein complexes containing human LINE-1 protein and RNA. *The EMBO journal* **15**(3): 630.
- Huang, C. R. L., K. H. Burns and J. D. Boeke** (2012). Active transposition in genomes. *Annu Rev Genet* **46**: 651.
- Hutvagner, G. and M. J. Simard** (2008). Argonaute proteins: key players in RNA silencing. *Nature reviews Molecular cell biology* **9**(1): 22-32.

- Ji, H., H. Moore and M. A. Blomberg** (1993). Hotspots for unselected Tyl transposition events on yeast chromosome I11 are near tRNA genes and LTR sequences. *Cell* **73**: 1007-1018.
- Kaer, K. and M. Speek** (2013). Retroelements in human disease. *Gene* **518**(2): 231-241.
- Kapitonov, V. V. and J. Jurka** (2001). Rolling-circle transposons in eukaryotes. *Proc Natl Acad Sci U S A* **98**(15): 8714-8719.
- Kassavetis, G. A., B. R. Braun, L. H. Nguyen, et al.** (1990). S. cerevisiae TFIIB is the transcription initiation factor proper of RNA polymerase III, while TFIIA and TFIIC are assembly factors. *Cell* **60**(2): 235-245.
- Kazazian, H. H., Jr.** (2004). Mobile elements: drivers of genome evolution. *Science* **303**(5664): 1626-1632.
- Kessin, R. H.** (2001). Dictyostelium: evolution, cell biology, and the development of multicellularity, Cambridge University Press.
- Kessin, R. H., G. G. Gundersen, V. Zaydfudim, et al.** (1996). How cellular slime molds evade nematodes. *Proceedings of the National Academy of Sciences* **93**(10): 4857-4861.
- Kirchner, J., C. M. Connolly and S. B. Sandmeyer** (1995). Requirement of RNA polymerase III transcription factors for in vitro position-specific integration of a retroviruslike element. *Science* **267**(5203): 1488-1491.
- Kirsten, J. H., Y. Xiong, A. J. Dunbar, et al.** (2005). Ammonium transporter C of Dictyostelium discoideum is required for correct prestalk gene expression and for regulating the choice between slug migration and culmination. *Dev Biol* **287**(1): 146-156.
- Kuhlmann, M., B. E. Borisova, M. Kaller, et al.** (2005). Silencing of retrotransposons in Dictyostelium by DNA methylation and RNAi. *Nucleic Acids Res* **33**(19): 6405-6417.
- Kumar, Y. and P. Bhargava** (2013). A unique nucleosome arrangement, maintained actively by chromatin remodelers facilitates transcription of yeast tRNA genes. *BMC genomics* **14**(1): 402.
- Kuzdzal-Fick, J. J., K. R. Foster, D. C. Queller, et al.** (2007). Exploiting new terrain: an advantage to sociality in the slime mold Dictyostelium discoideum. *Behavioral Ecology* **18**(2): 433-437.
- Lander, E. S., L. M. Linton, B. Birren, et al.** (2001). Initial sequencing and analysis of the human genome. *Nature* **409**(6822): 860-921.
- Larkin, M. A., G. Blackshields, N. Brown, et al.** (2007). Clustal W and Clustal X version 2.0. *Bioinformatics* **23**(21): 2947-2948.
- Lauermann, V. and J. D. Boeke** (1997). Plus strand strong - stop DNA transfer in yeast Ty retrotransposons. *The EMBO journal* **16**(21): 6603-6612.
- Leng, P., D. H. Klatt, T. L. Steck, et al.** (1998). Skipper, an LTR retrotransposon of Dictyostelium. *Nucleic Acids Res* **26**(8): 2008-2015.
- Li, S. I. and M. D. Purugganan** (2011). The cooperative amoeba: Dictyostelium as a model for social evolution. *Trends Genet* **27**(2): 48-54.
- Li, Z. and M. P. Deutscher** (1996). Maturation pathways for E. coli tRNA precursors: a random multienzyme process in vivo. *Cell* **86**(3): 503-512.
- Lohe, A. R., D. De Aguiar and D. L. Hartl** (1997). Mutations in the mariner transposase: the D, D (35) E consensus sequence is nonfunctional. *Proceedings of the National Academy of Sciences* **94**(4): 1293-1297.
- Luan, D. D., M. H. Korman, J. L. Jakubczak, et al.** (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.
- Malik, H. S.** (2000). Poised for Contagion: Evolutionary Origins of the Infectious Abilities of Invertebrate Retroviruses. *Genome Research* **10**(9): 1307-1318.
- Malik, H. S. and T. H. Eickbush** (1999). Modular evolution of the integrase domain in the Ty3/Gypsy class of LTR retrotransposons. *J Virol* **73**(6): 5186-5190.
- Mardis, E. R.** (2013). Next-generation sequencing platforms. *Annual review of analytical chemistry* **6**: 287-303.
- Marschalek, R., T. Brechner, E. Amon-Bohm, et al.** (1989). Transfer RNA genes: landmarks for integration of mobile genetic elements in Dictyostelium discoideum. *Science* **244**(4911): 1493-1496.

- Marschalek, R., J. Hofmann, G. Schumann, *et al.* (1992). Two distinct subforms of the retrotransposable DRE element in NC4 strains of *Dictyostelium discoideum*. *Nucleic Acids Res* **20**(23): 6247-6252.
- Marschalek, R., J. Hofmann, G. Schumann, *et al.* (1992). Structure of DRE, a retrotransposable element which integrates with position specificity upstream of *Dictyostelium discoideum* tRNA genes. *Mol Cell Biol* **12**(1): 229-239.
- Martin, S. L., M. Cruceanu, D. Branciforte, *et al.* (2005). LINE-1 retrotransposition requires the nucleic acid chaperone activity of the ORF1 protein. *J Mol Biol* **348**(3): 549-561.
- Martin, S. L., J. Li and J. A. Weisz (2000). Deletion analysis defines distinct functional domains for protein-protein and nucleic acid interactions in the ORF1 protein of mouse LINE-1. *Journal of Molecular Biology* **304**(1): 11-20.
- Mathias, S. L., A. F. Scott, H. H. Kazazian, *et al.* (1991). Reverse transcriptase encoded by a human transposable element. *Science* **254**(5039): 1808-1810.
- McClintock, B. (1948). Mutable loci in maize. *Carnegie Inst. Wash. Year Book* **47**(155).
- McKeown, M. and R. A. Firtel (1981). Differential expression and 5' end mapping of actin genes in *Dictyostelium*. *Cell* **24**(3): 799-807.
- Medstrand, P., L. N. van de Lagemaat and D. L. Mager (2002). Retroelement distributions in the human genome: variations associated with age and proximity to genes. *Genome Res* **12**(10): 1483-1495.
- Miller, W. J., J. F. McDonald and W. Pinsker (1997). Molecular domestication of mobile elements. *Evolution and Impact of Transposable Elements*, Springer: 261-270.
- Miskey, C., Z. Izsvák, K. Kawakami, *et al.* (2005). DNA transposons in vertebrate functional genomics. *Cellular and molecular life sciences* **62**(6): 629-641.
- Mizrokhi, L. J., S. G. Georgieva and Y. V. Ilyin (1988). jockey, a mobile drosophila element similar to mammalian LINES, is transcribed from the internal promoter by RNA polymerase II. *Cell* **54**(5): 685-691.
- Moqtaderi, Z. and K. Struhl (2004). Genome-wide occupancy profile of the RNA polymerase III machinery in *Saccharomyces cerevisiae* reveals loci with incomplete transcription complexes. *Mol Cell Biol* **24**(10): 4118-4127.
- Moqtaderi, Z., J. Wang, D. Raha, *et al.* (2010). Genomic binding profiles of functionally distinct RNA polymerase III transcription complexes in human cells. *Nature structural & molecular biology* **17**(5): 635-640.
- Moran, J. V. (1999). Exon Shuffling by L1 Retrotransposition. *Science* **283**(5407): 1530-1534.
- Mularoni, L., Y. Zhou, T. Bowen, *et al.* (2012). Retrotransposon Ty1 integration targets specifically positioned asymmetric nucleosomal DNA segments in tRNA hotspots. *Genome Research* **22**(4): 693-703.
- Nagarajavel, V., J. R. Iben, B. H. Howard, *et al.* (2013). Global 'bootprinting' reveals the elastic architecture of the yeast TFIIB-TFIIC transcription complex in vivo. *Nucleic Acids Res* **41**(17): 8135-8143.
- Ohshima, K. and N. Okada (2005). SINEs and LINES: symbionts of eukaryotic genomes with a common tail. *Cytogenet Genome Res* **110**(1-4): 475-490.
- Oliver, K. R. and W. K. Greene (2009). Transposable elements: powerful facilitators of evolution. *Bioessays* **31**(7): 703-714.
- Orgel, L. E. and F. H. Crick (1980). Selfish DNA: the ultimate parasite. *Nature* **284**: 604-607.
- Ostertag, E. M., R. J. DeBerardinis, J. L. Goodier, *et al.* (2002). A mouse model of human L1 retrotransposition. *Nat Genet* **32**(4): 655-660.
- Ostertag, E. M. and H. H. Kazazian Jr (2001). Biology of mammalian L1 retrotransposons. *Annu Rev Genet* **35**(1): 501-538.
- Pardue, M. L. and P. G. DeBaryshe (2003). Retrotransposons provide an evolutionarily robust non-telomerase mechanism to maintain telomeres. *Annu Rev Genet* **37**: 485-511.
- Paruzynski, A., A. Arens, R. Gabriel, *et al.* (2010). Genome-wide high-throughput integrome analyses by nrLAM-PCR and next-generation sequencing. *Nature protocols* **5**(8): 1379-1395.
- Pascali, C. and M. Teichmann (2013). RNA Polymerase III Transcription-Regulated by Chromatin Structure and Regulator of Nuclear Chromatin Organization. *Epigenetics: Development and Disease*, Springer: 261-287.

- Piriyapongsa, J., L. Marino-Ramirez and I. K. Jordan** (2007). Origin and evolution of human microRNAs from transposable elements. *Genetics* **176**(2): 1323-1337.
- Popp, C., W. Dean, S. Feng, et al.** (2010). Genome-wide erasure of DNA methylation in mouse primordial germ cells is affected by AID deficiency. *Nature* **463**(7284): 1101-1105.
- Qi, X., K. Daily, K. Nguyen, et al.** (2012). Retrotransposon profiling of RNA polymerase III initiation sites. *Genome Research* **22**(4): 681-692.
- Raiz, J., A. Damert, S. Chira, et al.** (2012). The non-autonomous retrotransposon SVA is trans-mobilized by the human LINE-1 protein machinery. *Nucleic Acids Res* **40**(4): 1666-1683.
- Raper, K. B.** (1935). Dictyostelium discoideum, a new species of slime mold from decaying forest leaves. *J. Agr. Res* **50**: 135-147.
- Raper, K. B. and D. I. Fennell** (1952). Stalk formation in Dictyostelium. *Bulletin of the Torrey Botanical Club*: 25-51.
- Raper, K. B. and A. W. Rahn** (1984). The dictyostelids, Princeton University Press Princeton.
- Reichmann, J., J. H. Crichton, M. J. Madej, et al.** (2012). Microarray analysis of LTR retrotransposon silencing identifies Hda1 as a regulator of retrotransposon expression in mouse embryonic stem cells. *PLoS Comput Biol* **8**(4): e1002486.
- Ribet, D., S. Louvet-Vallee, F. Harper, et al.** (2008). Murine endogenous retrovirus MuERV-L is the progenitor of the "orphan" epsilon viruslike particles of the early mouse embryo. *J Virol* **82**(3): 1622-1625.
- Rice, P., I. Longden and A. Bleasby** (2000). EMBOSS: the European molecular biology open software suite. *Trends in genetics* **16**(6): 276-277.
- Roggo, C., E. Coronado, S. K. Moreno - Forero, et al.** (2013). Genome - wide transposon insertion scanning of environmental survival functions in the polycyclic aromatic hydrocarbon degrading bacterium *Sphingomonas wittichii* RW1. *Environmental microbiology* **15**(10): 2681-2695.
- Romeralo, M., R. Escalante and S. L. Baldauf** (2012). Evolution and diversity of dictyostelid social amoebae. *Protist* **163**(3): 327-343.
- Rubinstein, C. V., P. Gerrienne, G. S. de la Puente, et al.** (2010). Early Middle Ordovician evidence for land plants in Argentina (eastern Gondwana). *New Phytol* **188**(2): 365-369.
- Rutherford, K., J. Parkhill, J. Crook, et al.** (2000). Artemis: sequence visualization and annotation. *Bioinformatics* **16**(10): 944-945.
- SanMiguel, P., A. Tikhonov, Y.-K. Jin, et al.** (1996). Nested retrotransposons in the intergenic regions of the maize genome. *Science* **274**(5288): 765-768.
- Sassaman, D. M., B. A. Dombroski, J. V. Moran, et al.** (1997). Many human L1 elements are capable of retrotransposition. *Nat Genet* **16**(1): 37-43.
- Schaap, P., T. Winckler, M. Nelson, et al.** (2006). Molecular phylogeny and evolution of morphology in the social amoebas. *Science* **314**(5799): 661-663.
- Schattner, P., A. N. Brooks and T. M. Lowe** (2005). The tRNAscan-SE, snoscan and snoGPS web servers for the detection of tRNAs and snoRNAs. *Nucleic Acids Res* **33**(suppl 2): W686-W689.
- Schilde, C. and P. Schaap** (2013). The Amoebozoa. *Dictyostelium discoideum Protocols*, Springer: 1-15.
- Schmidt, H. A., K. Strimmer, M. Vingron, et al.** (2002). TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing. *Bioinformatics* **18**(3): 502-504.
- Schmidt, M., K. Schwarzwaelder, C. Bartholomae, et al.** (2007). High-resolution insertion-site analysis by linear amplification-mediated PCR (LAM-PCR). *Nat Methods* **4**(12): 1051-1057.
- Schnable, P. S., D. Ware, R. S. Fulton, et al.** (2009). The B73 maize genome: complexity, diversity, and dynamics. *Science* **326**(5956): 1112-1115.
- Schramm, L. and N. Hernandez** (2002). Recruitment of RNA polymerase III to its target promoters. *Genes Dev* **16**(20): 2593-2620.
- Schumann, G., I. Zündorf, J. Hofmann, et al.** (1994). Internally located and oppositely oriented polymerase II promoters direct convergent transcription of a LINE-like retroelement, the Dictyostelium repetitive element, from Dictyostelium discoideum. *Mol Cell Biol* **14**(5): 3074-3084.

- Schwartz, A., D. C. Chan, L. G. Brown, *et al.* (1998). Reconstructing hominid Y evolution: X-homologous block, created by X–Y transposition, was disrupted by Yp inversion through LINE—LINE recombination. *Hum Mol Genet* **7**(1): 1-11.
- Sen, S. K., K. Han, J. Wang, *et al.* (2006). Human Genomic Deletions Mediated by Recombination between Alu-Elements. *The American Journal of Human Genetics* **79**(1): 41-53.
- Simms, T. A., S. L. Dugas, J. C. Gremillion, *et al.* (2008). TFIIC binding sites function as both heterochromatin barriers and chromatin insulators in *Saccharomyces cerevisiae*. *Eukaryotic cell* **7**(12): 2078-2086.
- Sims III, R. J., S. S. Mandal and D. Reinberg (2004). Recent highlights of RNA-polymerase-II-mediated transcription. *Current opinion in cell biology* **16**(3): 263-271.
- Singer, M. F., V. Krek, J. P. McMillan, *et al.* (1993). LINE-1: a human transposable element. *Gene* **135**(1-2): 183-188.
- Siol, O., M. Boutiliss, T. Chung, *et al.* (2006). Role of RNA polymerase III transcription factors in the selection of integration sites by the dictyostelium non-long terminal repeat retrotransposon TRE5-A. *Mol Cell Biol* **26**(22): 8242-8251.
- Siol, O., T. Spaller, J. Schiefner, *et al.* (2011). Genetically tagged TRE5-A retrotransposons reveal high amplification rates and authentic target site preference in the *Dictyostelium discoideum* genome. *Nucleic Acids Res* **39**(15): 6608-6619.
- Sorek, R., G. Ast and D. Graur (2002). Alu-containing exons are alternatively spliced. *Genome Res* **12**(7): 1060-1067.
- Speek, M. (2001). Antisense promoter of human L1 retrotransposon drives transcription of adjacent cellular genes. *Mol Cell Biol* **21**(6): 1973-1985.
- Sucgang, R., A. Kuo, X. Tian, *et al.* (2011). Comparative genomics of the social amoebae *Dictyostelium discoideum* and *Dictyostelium purpureum*. *Genome Biol* **12**(2): R20.
- Sun, C., D. B. Shepard, R. A. Chong, *et al.* (2012). LTR retrotransposons contribute to genomic gigantism in plethodontid salamanders. *Genome biology and evolution* **4**(2): 168-183.
- Suzuki, S., R. Ono, T. Narita, *et al.* (2007). Retrotransposon silencing by DNA methylation can drive mammalian genomic imprinting. *PLoS Genet* **3**(4): e55.
- Swergold, G. D. (1990). Identification, characterization, and cell specificity of a human LINE-1 promoter. *Mol Cell Biol* **10**(12): 6718-6729.
- Szafranski, K., G. Glöckner, T. Dinger, *et al.* (1999). Non-LTR retrotransposons with unique integration preferences downstream of *Dictyostelium discoideum* tRNA genes. *Molecular and General Genetics MGG* **262**(4-5): 772-780.
- Tamura, K., D. Peterson, N. Peterson, *et al.* (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* **28**(10): 2731-2739.
- Tessier, D. C., R. Brousseau and T. Vernet (1986). Ligation of single-stranded oligodeoxyribonucleotides by T4 RNA ligase. *Anal Biochem* **158**(1): 171-178.
- Troutt, A. B., M. G. McHeyzer-Williams, B. Pulendran, *et al.* (1992). Ligation-anchored PCR: a simple amplification technique with single-sided specificity. *Proceedings of the National Academy of Sciences* **89**(20): 9823-9825.
- Uchikawa, T., A. Yamamoto and K. Inouye (2011). Origin and function of the stalk-cell vacuole in *Dictyostelium*. *Dev Biol* **352**(1): 48-57.
- van den Hurk, J. A., I. C. Meij, M. C. Seleme, *et al.* (2007). L1 retrotransposition can occur early in human embryonic development. *Hum Mol Genet* **16**(13): 1587-1592.
- Vitte, C. and O. Panaud (2003). Formation of solo-LTRs through unequal homologous recombination counterbalances amplifications of LTR retrotransposons in rice *Oryza sativa* L. *Mol Biol Evol* **20**(4): 528-540.
- Volff, J. N. (2006). Turning junk into gold: domestication of transposable elements and the creation of new genes in eukaryotes. *Bioessays* **28**(9): 913-922.
- Vos, J. C., I. De Baere and R. H. Plasterk (1996). Transposase is the only nematode protein required for in vitro transposition of Tc1. *Genes & Development* **10**(6): 755-761.
- Wang, L., R. A. Haeusler, P. D. Good, *et al.* (2005). Silencing near tRNA genes requires nucleolar localization. *Journal of Biological Chemistry* **280**(10): 8637-8639.
- Wei, W., N. Gilbert, S. L. Ooi, *et al.* (2001). Human L1 retrotransposition: cis preference versus trans complementation. *Mol Cell Biol* **21**(4): 1429-1439.

- Weiner, A. M., P. L. Deininger and A. Efstratiadis** (1986). Nonviral retroposons: genes, pseudogenes, and transposable elements generated by the reverse flow of genetic information. *Annu Rev Biochem* **55**: 631-661.
- West, C. M.** (2003). Comparative analysis of spore coat Formation, structure, and function in Dictyostelium. *International review of cytology* **222**: 237-293.
- Wicker, T., F. Sabot, A. Hua-Van, et al.** (2007). A unified classification system for eukaryotic transposable elements. *Nature Reviews Genetics* **8**(12): 973-982.
- Wiegand, S., D. Meier, C. Seehafer, et al.** (2013). The Dictyostelium discoideum RNA-dependent RNA polymerase RrpC silences the centromeric retrotransposon DIRS-1 post-transcriptionally and is required for the spreading of RNA silencing signals. *Nucleic Acids Res*: gkt1337.
- Wilhelm, M. and F.-X. Wilhelm** (2001). Reverse transcription of retroviruses and LTR retrotransposons. *Cellular and Molecular Life Sciences CMLS* **58**(9): 1246-1262.
- Winckler, T.** (1998b). Retrotransposable elements in the Dictyostelium discoideum genome. *Cellular and Molecular Life Sciences CMLS* **54**(5): 383-393.
- Winckler, T., T. Dingermann and G. Glöckner** (2002). Dictyostelium mobile elements: strategies to amplify in a compact genome. *Cellular and Molecular Life Sciences CMLS* **59**(12): 2097-2111.
- Winckler, T., J. Schiefner, T. Spaller, et al.** (2011). Dictyostelium transfer RNA gene-targeting retrotransposons: Studying mobile element-host interactions in a compact genome. *Mob Genet Elements* **1**(2): 145-150.
- Winckler, T., K. Szafranski and G. Glöckner** (2005). Transfer RNA gene-targeted integration: an adaptation of retrotransposable elements to survive in the compact Dictyostelium discoideum genome. *Cytogenet Genome Res* **110**(1-4): 288-298.
- Winckler, T., C. Tschepke, E. De Hostos, et al.** (1998a). Tdd-3, a tRNA gene-associated poly (A) retrotransposon from Dictyostelium discoideum. *Molecular and General Genetics MGG* **257**(6): 655-661.
- Wu, X., Y. Li, B. Crise, et al.** (2003). Transcription start regions in the human genome are favored targets for MLV integration. *Science* **300**(5626): 1749-1751.
- Xiong, Y. and T. H. Eickbush** (1990). Origin and evolution of retroelements based upon their reverse transcriptase sequences. *The EMBO journal* **9**(10): 3353.
- Yieh, L., H. Hatzis, G. Kassavetis, et al.** (2002). Mutational analysis of the transcription factor IIIB-DNA target of Ty3 retroelement integration. *J Biol Chem* **277**(29): 25920-25928.
- Yieh, L., G. Kassavetis, E. P. Geiduschek, et al.** (2000). The Brf and TATA-binding protein subunits of the RNA polymerase III transcription factor IIIB mediate position-specific integration of the gypsy-like element, Ty3. *J Biol Chem* **275**(38): 29800-29807.
- Yin, H., J. Liu, Y. Xu, et al.** (2013). TARE1, a Mutated Copia-Like LTR Retrotransposon Followed by Recent Massive Amplification in Tomato. *PLoS One* **8**(7): e68587.
- Yoder, J. A., C. P. Walsh and T. H. Bestor** (1997). Cytosine methylation and the ecology of intragenomic parasites. *Trends in genetics* **13**(8): 335-340.
- Zhao, X. and H. Wang** (2007). LTR-FINDER: An efficient tool for the prediction of full-length LTR retrotransposons (2007). *Nucleic Acids Res* **35**: W265-W268.
- Zhelkovsky, A. M. and L. A. McReynolds** (2012). Structure-function analysis of Methanobacterium thermoautotrophicum RNA ligase—engineering a thermostable ATP independent enzyme. *BMC molecular biology* **13**(1): 24.

Anhang

A Verwendete Primer

Name	Sequenz 5' -> 3'
Transkriptionsnachweis <i>Df_DGLT-A</i>	
Dfas-DGLTA-27	GCGACGAGCTTGGTGGATC
Dfas-DGLTA-25	GCAAGTGACAGGTGGCACC
Dfas-DGLTA-19	GCTCCACACCCATGTCAAC
Dfas-DGLTA-21	GGACGGCACTCTCCGTATG
Dfas-DGLTA-22	CAGCGTCTTCAATTCGACTG
Dfas-DGLTA-32	GGCACCGCCATAATGC
Dfas-DGLTA-20	GTGATCTCATCTTCATTCG
Dfas-DGLTA-4	CTATCCTATTACTATCTATATCG
Klonierung Skipper-Elemente	
Skipper-01	GATATTTCTTGGTGATTAGTTG
Skipper-02_AgeI	<i>accggt</i> CTATAAACTCTTGGGTGTAAC
Skipper-03	<i>ctgcag</i> TCCACCTCTGAGCCATAGC
Skipper-04	CAACTCATCTCCAAACCCAAC
DGLT-A_AvrII-01	<i>cctagg</i> TAAAGTTTTCCACAAGATACAGATG
DGLT-A_AvrII-02	<i>cctagg</i> ATCACAATCAATTGGCTGTTGATCTC
Skipper-AvrII-NotI	<i>gcgccgc</i> cctaggGATTGGCATGTCTGTTGATGC
Skipper-AvrII_for	<i>cctagg</i> ATTAAGGAAGAACAATCAATTCTC
Skipper-NotI_rev	<i>gcgccgc</i> CAACTCATCTCCAAACCCAAC
DGLTA-AvrII-03	CAGACATGCCAATC <i>cctagg</i> GATAATAATAATCT
DGLTA-AvrII-06	TTGTTCTTCCTTAAT <i>cctagg</i> TAAAAACGATTCATC
Klonierung <i>Df_DGLT-A</i>-Elemente	
Dfas-DGLTA-12	CCCGCAAGGGAGATAC
Dfas-DGLTA-15	CCTAGTTTGTGAGGTGTGG
DGLTA-39	AATA <i>cggcgg</i> GTCTAATTGGAAGCTGGAGTATTG
DGLTA-40	GG <i>cccg</i> gTTTTTGGCCTTATCATCCTTTGTTAGG
mbsrI-01	CAATAAG <i>accggt</i> TTTATTAATTTCCGGG
mbsrI-02	GAATTCGG <i>ctgcag</i> GTCGACGGATCCTC
Dfas-DGLTA-35	G <i>Agg</i> atccGTCGACCTGCAGCCGAATTCGGGGAAGGACTTTAAAGTAGATAATAAT
Dfas-DGLTA-36	CCTTGATAGCAAATAGAGGGGGCCTG
Nachweis Sprungereignisse TRE5-A.2^{bsr}	
A15P-seq	AAATGGTTTAAAAAACTTGGGTTGG
BSR-01	CTGCCGAAATGATTTCTCC
Sonde Southern Blot	
BSR-TEST-01	GATAATAAACATCATGTGGGAGCG
BSR-TEST-02	GTTCTTCAATCGTAGTTTGTGAC
Etablierung LAM-PCR	
Q-GAPDH-01	GGTTGTCCCAATTGGTATTAATGG
Q-GAPDH-02	CCGTGGGTTGAATCATATTTGAAC
Q-DIRS-01	GAATTAACATCGTTAGATGATACAAACTC
Q-DIRS-02	CAATACCTGACTTTGAGAGTGTGATAG
Rep-196_454A	GCCTCCCTCGCGCCATCAGATAAATAACTGTTGAAATTGAATC
Rep-197	CAACATAAATGTCAACAAAAAATTTT
196_WT	ATAGATTACTGTACTAATACTGTCT
Rep-197_WT	CAACACAAATGCCAACAAAAATAAG
BSR-enrich-01	Biotin-TAACATTTCTCAACAAGATCTAGAATTAGT
Rep-232_454A	CGCTCCCTCGCGCCATCAGAAATTTTTGTTGACATTTATGTTG
Rep-234_bio	Biotin-CAAAATAAATAACTGTTGAAATTG
Rep-235	C TTGAACATCTTCACCATCC

Name	Sequenz 5' -> 3'
reverse Primer tRNA-Gene	
BSR-09 AlaAGC ArgUCU AsnGUU AspGUC CysGCA GlnUUG GluUUC GlyGCC GlyUUC GlyGCC GlyUCC HisGUG IleAAU IleUAU LeuUAA LeuUAG LysCUU LysUUU MetCAU PheGAA ProUGG SerAGA SerGUC SerUGA ThrAGU ThrCGU ThrUGU TrpCCA TyrGUA ValAAC ValCAC ValUAC tRNA_uni	CTCATTCCACTCAAATATACCCG CCCTTTACCTCTCGCATGCTAAGC CATTAATAGCTTAGAAGGCTACTG GGACCTAGCGGTTAACAGCCGCTC CCTTTCGCGTGACAGGCGAAAATC GACCTAACGATTTGCAGTCGCTTG CAGATTACAGGATTCAAAAGTCTTG CGAGGTACCAGTGTGAAAGACTAG GTATGCTGGGAATCGAACCC CGGTTCAACTGCTTGAAGGCAGC GTATGCTGGGAATCGAACCC CGGTTCAACTGCTTGAAGGCAGC GGTCAACAGAGCCACAATCTGATG CGACCGTGGCGTTATTAGCACCAC CTCGCGACCAAAGCATTGC GAACCGAATGGATGAGATCTTAAG GAACCGAATGGATGAGATCTTAAG GCTCAGTCGGTAGAGCGC CCGACCACAAGGTTAAAAGCCTTG CTAGGACCTCGAGGTTATGG CGACCTTAAGATCTTCAGTCTCAC CGAGAATCGAACTCAGGACG CGAAGGGATCTGATTTCTAGTC GAGAACGCCTTAGCAGGGC GCGCCTCCTAGGAGATTTGATTTT GACCTCTTCTTACTAGGGAAGCG GTATAATTTTATACGAGAGAAATGTCTTACC GACCCTCTGTTATCGCTTTAGCAC CTGGACCAACTGCTAACTTCTAAC CAACCGATTGATCTCAATTATTAC ACGACCTAATCCGTGTTAAGGATT GACCTAACCCGTGTGAAGGGC GAGACCTGTTGCGTGTAAGCAAC CTGCAGRRRNTCGAACYNNGAC
PCR für TRE5-A.2^{bsr}-Integrationen an B-Box-Loci auf dem Palindrom	
Pal-2656-rev Pal-6106-rev Pal-6483-rev Pal-7762-rev Pal-9175-rev Pal-10430-rev Pal-13997-rev Pal-15009-rev Pal-15156-rev Pal-18638-rev Pal-25439-rev Pal-26963-rev Pal-27726-rev Pal-29052-rev Pal-29059-rev Pal-29115-rev Pal-31700-rev	GAATTTAGAGATTA AAAACTTTTTGGAG CTTTAGTGCAGATCTTGATGGTAGTAG GGGTATCATTTTTAATTCACCTAATTGG CACCTTCTACGACTCGAAAATCCAAAC TGTGATGCCCTTAGATACCTTGGGCCG GGTATGATATTTATGTCACTACTTC CCCCTTTTTTGTACCCTTTGTAGGTC TTCCTAAATGAGCAATTTGACATTCCC AACAACTTTTTTAAGGTATAGTTACTC AAATGGTAAAATAGTCCGAAGGTCCAC GTGGATGCTCCAGGGAGTTATCTATTC GTAGATGCCAATTGGAGTAATATATTC TGAGGTAACAGTAGATTCTCCATTGAG ATTAGGTCATAGGTTTCTGGATCGATC CAACGGGGAATGGAACGTCTTGAAG TCTAAGTTAGATTTTTACTAGTAATTC TACCATTAGATAGCCCTATTTCTCTAG

Name	Sequenz 5' -> 3'
Pal-33496-rev	GGAGTTGTCTATCCAACAAGGGTAATC
PCR für TRE5-A.2^{bsr}-Integrationen am r5S-Gen	
5S-rDNA-03 Rep-235	CCAACCAGAGTACTTGGAAGGC CTTGAACATCTTCACCATCC
PCR für Integrationen endogener TRE5-A-Elemente am r5S-Gen	
5S-rDNA-04 Rep-190 5S-rDNA-03 Rep-109Q	GTATACAGCACCCCTATGTTCCCTG CCATAGGTTGGGCTGAAACT CCAACCAGAGTACTTGGAAGGC GTTAGATTGTCTAGTTCAATGATAGTGTC
PCR für Integrationen endogener TRE5-A-Elemente an palindromische B-Box-Loci	
Palindrom-06 Rep-190 Pal-18638-rev Rep-157Q	GTAATGGAGTAAGCTATCTAATGATAC CCATAGGTTGGGCTGAAACT AAATGGTAAAATAGTCCGAAGGTCCAC CTTCCGTTGTTAGTCTCTTTGTTTCGCC
cRT-PCR Palindrom-Locus 18638	
cRT-PCR_Pal-18638-01 cRT-PCR_Pal-18638-02 cRT-PCR_Pal-18638-03 cRT-PCR_Pal-18638-04 cRT-PCR_Pal-18638-05	AATGGTAAAATAGTCCGAAGGTCCAC CGAACCCGCGACCCCTCAG GTATCATTAGATAGCTTACTCCATTAC GTTGCTTTGGCTTAGTTGGTTAAGCG GCATACTATTATATATTTTT GGCC
cRT-PCR Palindrom-Locus 25439	
cRT-PCR_Pal-25439-01 cRT-PCR_Pal-25439-02 cRT-PCR_Pal-25439-03 cRT-PCR_Pal-25439-04 cRT-PCR_Pal-25439-05	CTATTACATAATAATGCAACC GAGAATCGAACCCTGGAC CTTTATATTGTCTAATGACCTCC GATGTCTAAAAATAAACCGAG TGCTGACTAATCTACCAC
cRT-PCR Palindrom- Locus 26963	
cRT-PCR_Pal-26963-01 cRT-PCR_Pal-26963-02 cRT-PCR_Pal-26963-03 cRT-PCR_Pal-26963-04 cRT-PCR_Pal-26963-05	GACCCCTCTATAATGCAACC AAAATCGAGAATCGAACCCG CAATTGGCATCTACCATTACC GGTGGATTAGTCATCATATGTC CTTTATGTGGTCTAATGACCTTC
RT-PCR Palindrom-Locus 18638/22168	
Pal-RT_04 Pal-RT_05	GTTGCTTTGGCTTAGTTGGTTAAGCG AATGGTAAAATAGTCCGAAGGTCCAC
iPCR	
iPCR_ORF1opt-01 iPCR_ORF1opt-02	CAAAATAAATAACTGTTGAAATTG AATTAAAAGTGAAATTAAGGTGAG
Linker-Ligation	
454B_Compl_5'P+3'ddCTP Rep-201 Q-A6P-01 454B_-2bp_Linker 454B_Nru_Bio Rep-196_454A	Pho-CTGAGCGGGCTGGCAAGGCGTCATCGGAT-ddCTP CTCCTCTAGAAGAGGTTTATTTTTAAAAATTAC GTAATAAAACAAAGAATTGATATTCTGATGCC CCGATGACGCCTTGCCAGCC Biotin-GAGAGAGAGAtcgcgaGCCTTGCCAGCCCGCTCAG GCCTCCCTCGGCCATCAGATAAATAACTGTTGAAATGAATC
Klonierung TRE5-A.2^{bsr}	
Rep-205 Rep-230_AgeI Rep-226_Sall Rep-227_Sall Rep-228_AgeI Rep-229_Sall	CACAATAAGACCGGTTTTATTAATTTCCG accggtGGGGATCCCCAATTCTTTTAAAAATG gtcgacCGCATGCTCCTCTAGAAG gtcgacAATTTGCTTATTTTTGTTGGC accggtGTTATTTATAGAAAGTACTGAGTG gtcgacAAAACCTCAATTAATAAAAAACAATTC

B Erzeugte Plasmide

In der folgenden Tabelle sind die in dieser Arbeit erzeugten Plasmide gelistet. Die Nummern beziehen sich auf die intern verwendete Datenbank, in der alle Plasmiddaten gespeichert werden.

Plasmid Name	Nummer	Inserts; Art der Klonierung
pJ1.2-ORF1_opt_Dd	2.0363	Codon-adaptierter ORF1 von TRE5-A.1; synthetisches Plasmid
pGEMT-A6P-Orf1_Ddopt-BSRi-C_Age	2.0365	A6P-ORF1_Dd/adapt-msbri-C; als HindIII/AgeI-Fragment aus 2.0363 in 2.0128 kloniert
pGEMT-DGLT-A_Dfas	2.0373	komplettes <i>Df_DGLT-A</i> -Element von <i>D. fasciculatum</i> , PCR mit Primern DGLT-A-12/-15 an gDNA
TOPO-XL-F1 (Df_DGLT-A_bsri)	2.0501	<i>Df_DGLT-A</i> ohne 3'LTR
pGEMT-F3 (Df_DGLT-A_bsri)	2.0502	<i>msbri</i> -5'LTR von <i>Df_DGLT-A</i> ; Ligation von zwei PCR-Produkten
pBT1-F1_LTR3	2.0524	Fragment 1 vom <i>Df_DGLT-A</i> -Element; als EagI/Sma-Fragment aus 2.0501 in pBT1 kloniert
pBT-Dfas-DGLT-2_mBSRi	2.0533	komplettes DGLT-A aus <i>D. fasciculatum</i> mit <i>msbri</i> -Kassette vor 3'LTR; EagI-Fragment aus 20502 in 20524 kloniert
TOPOXL-Skipper-F1	2.0544	großes Fragment Skipper aus genomischer DNA <i>Dictyostelium discoideum</i> ; Primer: Skipper-01 und Skipper-02_Age
pGEMT-Skipper-F3	2.0546	kleines Fragment Skipper 3'LTR; PCR an gDNA mit Primern Skipper-03_PstI + Skipper-04 an gDNA-Ddis, kloniert in pGEMT
TOPOXL-A6P-ORF1-BSRi-C_Spe/Xho	2.0557	A6P-ORF1-BSRi-C_Spe/Xho; PCR-Produkt mit Primern Rep-224 und Rep-225 von 2.0128 in TOPOXL kloniert
pGEMT-F1_ORF2E	2.0588	SalI-A6P-SalI; PCR mit Primern Rep-226_SalI + Rep-227_SalI an Plasmid 2.0128; Produkt in pGEMT kloniert
TOPOXL-ORF2E_F2-SalI	2.0589	SalI-ORF1-ORF2-AgeI TRE5A1; PCR mit Primern Rep-228_AgeI + Rep-229_SalI an Plasmid pB3 (1.0821), in TOPOXL-Vektor kloniert
TOPOXL-F4_ORF2E	2.0590	SalI-A6P-ORF1-ORF2-AgeI; SalI-Fragment aus 20588 in SalI-Schnittstelle von 2.0589 kloniert

Plasmid Name	Nummer	Inserts; Art der Klonierung
pGEMT-F3_ORF2E	2.0591	AgeI-BSRi-A6P-C-AgeI; PCR mit Primern Rep-205 und Rep-230_AgeI an Plasmid 2.0128; Produkt in pGEMT kloniert
TOPOXL-ORF2E	2.0592	A6P-ORF1-ORF2-BSRi-C; BSRi-C als AgeI-Fragment aus 2.0591 in AgeI-Schnittstelle von 2.0590 kloniert
TOPOXL-Skipper_ <i>mbstri</i>	2.0702	BSR-Kassette aus 2.0128 als AgeI/PstI-Fragment in 20544 kloniert, anschließend 3'LTR als PstI-Fragment in Vektor kloniert
pCR8/GW/TOPO-Skipper_F1_AvrII-NotI	2.0755	5' Fragment Skipper (F1) für Klonierung des Skipper-RNaseH_DGLTA Elements; PCR mit Primern Skipper-01 + Skipper-AvrII-NruI an Plasmid 2.0702, Produkt in pCR8/GW/TOPO ligiert
pGEMT-Skipper_F2_AvrII-NotI	2.0756	3' Fragment Skipper (F2) für Klonierung des Skipper-RNaseH_DGLTA Elements; PCR mit Primern Skipper-AvrII_for + Skipper-NruI_rev an Plasmid 2.0702, Produkt in pGEMT ligiert
pCR8/GW/TOPO-RNaseH_DGLTA_AvrII	2.0757	RNase-H von DGLTA als AvrII-Fragment; PCR mit Primern DGLT_A_AvrII-01 + DGLT_A_AvrII-02 an 2.0570, Produkt in pCR8/GW/TOPO
pCR8/GW/TOPO-Skipper-F3_AvrII	2.0758	Skipper ohne RNaseH-Domäne, dafür mit AvrII-Schnittstelle um RNaseH-Domäne auszutauschen; AvrII/NotI-Fragment aus 20756 in AvrII/Not-Sites von 2.0755 kloniert
pGEMT-RNaseH_Skipper_AvrII	2.0759	Rnase-H von Skipper als AvrII-Fragment; PCR mit Primern DGLTA-AvrII-03 + DGLTA-AvrII-06, Produkt in pGEMT kloniert
pCR8/GW-Skipper_RNaseH_DGLTA	2.0770	Skipper-BSRi mit RNaseH-Domäne von DGLTA als AvrII Fragment; Rnase-H-Domäne aus 2.0757 als AvrII-Fragment in 2.0758 kloniert
pCR8/GW-Skipper_RNaseH_Skipper	2.0771	Skipper-BSRi mit RNaseH-Domäne von Skipper als AvrII Fragment; Rnase-H-Domäne aus 2.0759 als AvrII-Fragment in 2.0758 kloniert
pUC57-Skipper_chim	2.0776	Chimäre aus RNaseH DGLTA + Linker Skipper + In-NTD.1 von DGLTA als AvrII-Fragment; Synthetisches Plasmid
pCR8/GW-Skipper_RNH_Skipper_chim	2.0793	Chimäres Skipper-Element mit BSR-Kassette und RNaseH DGLTA + Linker Skipper + In-NTD.1 von DGLTA als AvrII-Fragment von 2.0776 in 2.0758 kloniert

Mappbare Integrationen von TRE5-A^{bsr} der Hochdurchsatz-Sequenzierungs-Experimente

Die DDB-Nummern der Chromosomen entsprechen der Nomenklatur von www.dictybase.org.

C Ergebnisse der Sequenzierung tRNA-Gen-spezifischer LAM-PCR-Produkte

Chromosom	Chromosomale Position		tRNA-Gen
DDB0220052	33256	34248	ThrAGU-1, ThrAGU-2
DDB0220052	49183	49385	–
DDB0220052	49923	50331	–
DDB0232428	62538	62816	–
DDB0232428	68060	68478	–
DDB0232428	448295	448694	LysUUU-2
DDB0232428	818479	818877	SerUGA-1
DDB0232428	835509	835907	ProUGG-2
DDB0232428	945847	946047	SerUGA-2
DDB0232428	1511239	1511637	TrpCCA-4
DDB0232428	1618045	1618732	ProUGG-5
DDB0232428	1760194	1760627	ArgAGC-2
DDB0232428	1863182	1863592	ThrAGU-5
DDB0232428	1995314	1995715	ThrAGU-6
DDB0232428	1996992	1997402	ThrAGU-7
DDB0232428	2023066	2023464	ArgACG-1
DDB0232428	2265405	2265803	AspGUC-6
DDB0232428	2366670	2367090	GluUUC-2
DDB0232428	2368151	2368578	GluUUC-6
DDB0232428	2398096	2398494	LysUUU-3
DDB0232428	2404490	2404889	ThrAGU-8
DDB0232428	2411750	2412164	TrpCCA-3
DDB0232428	2424672	2425074	SerGCU-2
DDB0232428	2503328	2503726	LysUUU-4
DDB0232428	2558421	2558621	SerAGA-1
DDB0232428	2687032	2687432	GluUUC-3
DDB0232428	2696843	2697245	LysCUU-2
DDB0232428	2817298	2817699	TrpCCA-1
DDB0232428	2855126	2855524	GlnUUG-4
DDB0232428	2877505	2878502	LysUUU-5
DDB0232428	2912820	2913219	TrpCCA-2
DDB0232428	3008285	3008683	GluUUC-4
DDB0232428	3222572	3222971	ThrAGU-9
DDB0232428	3256654	3257081	IleAAU-5
DDB0232428	3431048	3431469	AlaAGC-1
DDB0232428	3579526	3579924	AsnGUU-5
DDB0232428	3644036	3644458	ThrAGU-12
DDB0232428	3865599	3866000	AspGUC-2
DDB0232428	3927833	3928301	ValCAC-1
DDB0232428	4158702	4159135	LysUUU-6
DDB0232428	4163434	4163843	IleAAU-4
DDB0232428	4185478	4185876	AlaUGC-1
DDB0232428	4249671	4250069	AspGUC-5
DDB0232428	4250427	4250825	AspGUC-4
DDB0232429	346897	347295	ValAAC-12
DDB0232429	1492264	1492662	IleAAU-6
DDB0232429	1609857	1610257	–
DDB0232429	2661337	2661735	LeuUAA-1
DDB0232429	3370004	3370407	LeuUAA-6
DDB0232429	4055304	4056074	–
DDB0232429	4103111	4103511	LeuUAA-2
DDB0232429	4399624	4400035	SerAGA-2
DDB0232429	4994845	4995248	TrpCCA-6
DDB0232429	5289855	5290210	HisGUG-5
DDB0232429	5416018	5417050	MetCAU-7
DDB0232429	5452139	5452537	GluUUC-9
DDB0232429	5454520	5454918	GluUUC-8
DDB0232429	5820851	5821249	TrpCCA-7
DDB0232429	6126981	6127381	ValAAC-7
DDB0232429	6132078	6132479	AspGUC-8

Chromosom	Chromosomale Position		tRNA-Gen
DDB0232429	6187067	6187467	ArgUCU-3
DDB0232429	6386730	6387128	ArgACG-3
DDB0232429	6482041	6482439	LeuCAA-2
DDB0232429	6680417	6680819	ThrUGU-2
DDB0232429	6724005	6724403	ArgACG-6
DDB0232429	7004045	7004446	ProUGG-7
DDB0232429	7122453	7122851	AlaAGC-2
DDB0232429	7237087	7237485	SerAGA-6
DDB0232429	7435415	7435824	ValAAC-8
DDB0232429	7675861	7676259	AspGUC-10
DDB0232429	7878584	7878983	SerAGA-5
DDB0232429	8141240	8141639	AlaAGC-3
DDB0232430	548862	549217	HisGUG-6
DDB0232430	1112640	1113039	LysUUU-13
DDB0232430	1431019	1431452	LysUUU-12
DDB0232430	1573952	1574354	CysGCA-6
DDB0232430	1807751	1808156	ValUAC-2
DDB0232430	2450050	2450448	CysGCA-3
DDB0232430	2539672	2540071	SerAGA-8
DDB0232430	2632237	2632635	CysGCA-4
DDB0232430	2675375	2675814	GluUUC-12
DDB0232430	2700409	2700807	IleAAU-9
DDB0232430	2875993	2876397	ArgACG-7
DDB0232430	3143360	3143765	SerAGA-7
DDB0232430	3365899	3366942	GlnUUG-11
DDB0232430	3876377	3876831	GlyGCC-6
DDB0232430	4014532	4014732	AlaAGC-5
DDB0232430	4067716	4068115	AlaAGC-6
DDB0232430	4090026	4090499	AlaUGC-6
DDB0232430	4194284	4194682	AlaUGC-4
DDB0232430	4240543	4240941	GlyUCC-1
DDB0232430	4266440	4266935	LysCUU-5
DDB0232430	4267826	4268370	LysUUU-10
DDB0232430	4457786	4458185	LeuUAA-7
DDB0232430	4567710	4568209	ArgUCU-9
DDB0232430	5070497	5070901	ThrCGU-1
DDB0232430	5325746	5326195	GlyUCC-4
DDB0232430	5333460	5333865	SerGCU-9
DDB0232430	5339236	5339650	SerGCU-7
DDB0232430	5679848	5680253	SerCGA-1
DDB0232430	5706343	5707231	GlyUCC-3, GlyUCC-2
DDB0232430	5781462	5781860	SerGCU-6
DDB0232430	5867761	5868159	ValAAC-14
DDB0232430	5869879	5870284	ValAAC-15
DDB0232430	5898210	5898611	SerGCU-3
DDB0232430	5906656	5907056	SerGCU-5
DDB0232431	1588992	1589390	SerUGA-8
DDB0232431	1802240	1802638	LysUUU-15
DDB0232431	2220059	2220478	SerGCU-11
DDB0232431	2227384	2227784	SerGCU-10
DDB0232431	2563293	2563693	PheGAA-11
DDB0232431	2845576	2845974	AspGUC-18
DDB0232431	3919695	3920093	AspGUC-15
DDB0232431	3942159	3942557	AspGUC-14
DDB0232431	3947589	3947995	AlaAGC-8
DDB0232431	3970617	3971017	LysUUU-14
DDB0232431	4598249	4598449	PheGAA-12
DDB0232431	4641038	4641436	-
DDB0232431	4682686	4683085	AlaUGC-7
DDB0232431	5168027	5168426	LeuUAA-11
DDB0232431	5226743	5227141	LeuUAA-12
DDB0232432	40637	41043	LysUUU-16
DDB0232432	254601	254999	AlaAGC-10
DDB0232432	278485	278883	IleAAU-11
DDB0232432	431961	432359	LeuAAG-7
DDB0232432	712605	713006	ThrUGU-6
DDB0232432	2550015	2550413	LeuAAG-6

Chromosom	Chromosomale Position		tRNA-Gen
DDB0232432	2593478	2593876	ArgUCU-12
DDB0232432	2727144	2727542	LysUUU-17
DDB0232432	2768390	2768790	LysUUU-18
DDB0232432	2970286	2970684	ProUGG-14
DDB0232432	3061119	3061517	ValUAC-3
DDB0232432	3113003	3113401	LysUUU-19
DDB0232432	3208055	3208460	ValAAC-19
DDB0232432	3336403	3336803	ThrUGU-5
DDB0232432	3761573	3761974	ValUAC-6
DDB0232432	3986939	3987362	AlaUGC-8
DDB0232433	7992	8460	–
DDB0232433	323171	323574	GluUUC-15
DDB0232433	585890	586293	ValAAC-20
DDB0232433	858910	859312	GlyGCC-14
DDB0232433	1044661	1045096	LeuUAA-18
DDB0232433	1096171	1096617	–
DDB0232433	1097075	1097509	SerUGA-15
DDB0232433	1132978	1133376	SerUGA-13
DDB0232433	1703136	1703540	LysUUU-23
DDB0232433	1775953	1776376	LeuUAA-13
DDB0232433	1837421	1838188	GlyGCC-18, HisGUG-10
DDB0232433	1870301	1870736	LeuUAA-17
DDB0232433	1912637	1913038	GlyGCC-16
DDB0232433	2120259	2120661	GluUUC-16
DDB0232433	2258353	2258774	CysGCA-7
DDB0232433	2340407	2340813	SerAGA-9
DDB0232433	2347599	2347997	SerGCU-12
DDB0232433	2364871	2365269	GluCUC-2
DDB0232433	2478390	2478788	GluCUC-3
DDB0232433	2479550	2479948	AlaAGC-14
DDB0232433	2624377	2624776	LeuUAA-16
DDB0232433	2729935	2730333	GluUUC-20
DDB0232433	2768158	2768562	ValUAC-7
DDB0232433	2777239	2777639	AspGUC-21
DDB0232433	2836618	2837017	LeuAAG-10
DDB0232433	2840179	2840576	HisGUG-8
DDB0232433	2935908	2936306	IleAAU-16
DDB0232433	3008483	3008881	AspGUC-22
DDB0232433	3052901	3053301	GluUUC-17
DDB0232433	3065099	3065303	–
DDB0232433	3068337	3068735	TyrGUA-11
DDB0232433	3132092	3132491	AlaAGC-11
DDB0232433	3137229	3137666	LysUUU-22
DDB0232433	3173380	3173821	HisGUG-9
DDB0232433	3291507	3291905	GluUUC-18
DDB0232433	3299689	3300156	IleUAU-2
DDB0232433	3312452	3312854	IleUAU-1
DDB0232433	3325307	3325705	SerUGA-11
DDB0232433	3431836	3432236	LeuUAA-15
DDB0232433	3534697	3535098	IleAAU-15

D Auswertung der iPCR-Bank

Chromosom	Chromosomale Position		tRNA-Gen
DDB0232428	448293	448695	LysUUU-2
DDB0232428	749910	750307	AsnGUU-8
DDB0232428	818481	818879	SerUGA-1
DDB0232428	1615335	1615774	ProUGG-6
DDB0232428	1618283	1618725	ProUGG-5
DDB0232428	1687977	1688375	ThrAGU-4
DDB0232428	1759507	1760626	LysCUU-1, ArgACG-2
DDB0232428	1996992	1997403	ThrAGU-7
DDB0232428	2369953	2370352	GlnUGG-5
DDB0232428	2398096	2398494	LysUUU-3
DDB0232428	2404485	2404894	ThrAGU-8
DDB0232428	2411760	2412160	TrpCCA-3
DDB0232428	2424625	2425073	SerGCU-2

Chromosom	Chromosomale Position		tRNA-Gen
DDB0232428	2503326	2503730	LysUUU-4
DDB0232428	2558421	2558819	SerAGA-1
DDB0232428	2855080	2855526	GlnUUG-4
DDB0232428	3008285	3008683	GluUUC-4
DDB0232428	3222572	3222975	ThrAGU-9
DDB0232428	3256659	3257082	IleAAU-5
DDB0232428	3579526	3579927	AsnGUU-5
DDB0232428	3644027	3644461	ThrAGU-12
DDB0232428	3927873	3928301	ValCAC-1
DDB0232428	4163430	4163850	IleAAU-4
DDB0232428	4185468	4185877	AlaUGC-1
DDB0232429	3370004	3370407	LeuUAA-6
DDB0232429	4055327	4056082	-
DDB0232429	4103099	4103513	LeuUAA-2
DDB0232429	4399612	4400161	SerAGA-2
DDB0232429	5416636	5417053	MetCAU-7
DDB0232429	5454486	5454923	GluUUC-8
DDB0232429	5934237	5934447	AsnGUU-9
DDB0232429	6132080	6132482	AspGUC-8
DDB0232429	6187067	6187465	ArgUCU-3
DDB0232429	6226882	6227282	ArgUCU-6
DDB0232429	6571102	6571504	SerAGA-4
DDB0232429	6647006	6647404	ArgUCU-5
DDB0232429	6766725	6767126	PheGAA-5
DDB0232429	6771382	6771780	ProUGG-9
DDB0232429	7122453	7122856	AlaAGC-2
DDB0232429	7187737	7188135	PheGAA-6
DDB0232429	7435426	7435827	ValAAC-8
DDB0232429	7524204	7524602	ProUGG-8
DDB0232429	7714989	7715391	AspGUC-9
DDB0232429	7876977	7877208	-
DDB0232430	1431015	1431459	LysUUU-12
DDB0232430	1773465	1773872	ArgUCG-1
DDB0232430	1807751	1808149	ValUAC-2
DDB0232430	2539673	2540071	SerAGA-8
DDB0232430	3365897	3366632	GlnUUG-11
DDB0232430	4067716	4068114	AlaAGC-6
DDB0232430	4090079	4090499	AlaUGC-6
DDB0232430	4267924	4268367	LysUUU-10
DDB0232430	4270141	4270587	LysUUU-9
DDB0232430	4400198	4400631	GlyGCC-9
DDB0232430	4567788	4568212	ArgUCU-9
DDB0232430	5126000	5126398	ArgUCU-8
DDB0232430	5333433	5333866	SerGCU-9
DDB0232430	5339235	5339650	SerGCU-7
DDB0232430	5679854	5680252	SerCGA-1
DDB0232430	5869884	5870284	ValAAC-15
DDB0232431	2220040	2220483	SerGCU-11
DDB0232431	3945422	3946009	SerUGA-7
DDB0232431	4598050	4598448	PheGAA-12
DDB0232431	4641022	4641439	-
DDB0232431	5168025	5168431	LeuUAA-11
DDB0232431	5226743	5227141	LeuUAA-12
DDB0232431	5312700	5313098	TyrGUA-8
DDB0232432	40638	41036	LysUUU-16
DDB0232432	254601	254999	AlaAGC-10
DDB0232432	2550015	2550413	LeuAAG-6
DDB0232432	2727143	2727542	LysUUU-17
DDB0232432	2956433	2956840	ProUGG-13
DDB0232432	3336349	3336806	ThrUGU-5
DDB0232432	3693331	3693729	ArgUCU-11
DDB0232432	3986941	3987366	AlaUGC-8
DDB0232433	1097061	1097509	SerUGA-15
DDB0232433	1132978	1133376	SerUGA-13
DDB0232433	1775956	1776358	LeuUAA-13
DDB0232433	1870336	1870735	LeuUAA-17
DDB0232433	2258353	2258772	CysGCA-7

Chromosom	Chromosomale Position		tRNA-Gen
DDB0232433	2340407	2340814	SerAGA-9
DDB0232433	2347563	2348003	SerGCU-12
DDB0232433	2478386	2478788	GluCUC-3
DDB0232433	2479535	2479953	AlaAGC-14
DDB0232433	3052901	3053300	GluUUC-17
DDB0232433	3068337	3068735	TyrGUA-11
DDB0232433	3132088	3132506	AlaAGC-11
DDB0232433	3325308	3325706	SerUGA-11

E Auswertung der Linker-ligierten LAM-PCR-Bank

Chromosom	Chromosomale Position		tRNA-Gen
DDB0232428	1759524	1760628	LysCUU-1, ArgACG-2
DDB0232428	1997000	1997398	ThrAGU-7
DDB0232428	2560333	2560536	LysCUU-3
DDB0232428	2696822	2697500	LysCUU-2
DDB0232428	3927897	3928295	ValCAC-1
DDB0232428	4163398	4163853	IleAAU-4
DDB0232429	4048676	4049104	–
DDB0232429	4055611	4056080	–
DDB0232429	4399636	4400037	SerAGA-2
DDB0232429	5400604	5401012	LysCUU-4
DDB0232429	6364299	6364792	GlnUUG-8
DDB0232429	6386963	6387172	ArgACG-3
DDB0232429	6723964	6724172	ArgACG-6
DDB0232429	7430048	7430462	AlaUGC-2
DDB0232429	7434376	7434780	AlaUGC-3
DDB0232430	2875957	2876171	ArgACG-7
DDB0232430	4194316	4194730	AlaUGC-4
DDB0232430	4267918	4268367	LysUUU-10
DDB0232430	4270147	4270381	LysUUU-9
DDB0232430	4457784	4458192	LeuUAA-7
DDB0232430	4567788	4568206	ArgUCU-9
DDB0232430	5652699	5653106	LysCUU-6
DDB0232430	5679792	5680257	SerCGA-1
DDB0232431	3947563	3948000	AlaAGC-8
DDB0232433	1097066	1097505	SerUGA-15
DDB0232433	1132972	1133376	SerUGA-13
DDB0232433	1794609	1794818	LysUUU-20
DDB0232433	2082285	2082682	LeuAAG-8
DDB0232433	2258355	2258765	CysGCA-7
DDB0232433	2340412	2340810	SerAGA-9
DDB0232433	3299695	3300153	IleUAU-2

Liste der wissenschaftlichen Veröffentlichungen

Siol, O.; **Spaller, T.**; Schiefner, J.; Winckler, T. (2011). Genetically tagged TRE5-A retrotransposons reveal high amplification rates and authentic target site preference in the *Dictyostelium discoideum* genome. *Nucleic Acids Res* **39**(15): 6608-6619.

Winckler, T., Schiefner, J.; **Spaller, T.**; Siol, O.; (2011). *Dictyostelium* transfer RNA gene-targeting retrotransposons: Studying mobile element-host interactions in a compact genome. *Mob Genet Elements* **1**(2): 145-150.

Schmith, A.; Groth, M.; Ratka, J.; Gatz, S.; **Spaller, T.**; Siol, O.; Glöckner, G.; Winckler, T. (2013). Conserved Gene Regulatory Function of the Carboxy-Terminal Domain of Dictyostelid C-Module-Binding Factor. *Eukaryotic cell* **12**(3): 460-468.

Spaller, T.; Schiefner, J.; Glöckner, G.; Winckler, T. (2012). How to watch retrotransposons jumping around in the *Dictyostelium* genome. Posterpräsentation Annual International *Dictyostelium* Conference 2012, Madrid (Spanien).

Lebenslauf

Persönliche Daten

Name: Spaller
Vorname: Thomas
Geboren am: 05.04.1984
Geboren in: Zwickau
Familienstand: verheiratet, 1 Kind
Anschrift: Hospitalgasse 2
09350 Lichtenstein

Schulische Ausbildung

1990-1994 Sigmund-Jähn-Grundschule Zwickau
1994-2002 Georgengymnasium Zwickau, Abschluss: Abitur (Note: 2,7)

Zivildienst

08/2002-05/2003 Jakobusheim Mülsen St. Jakob, Tätigkeit in der Pflege

Berufliche Ausbildung

Studium

10/2003-05/2009 Friedrich-Schiller-Universität Jena
Studiengang: Biologie (Diplom)
Abschluss: Diplom (Note: 1,3)

Beruflicher Werdegang

06/2009-09/2009 Wissenschaftliche Hilfskraft am Fritz-Lipmann-Institut Jena
seit 01/10/09 Wissenschaftlicher Mitarbeiter am Lehrstuhl für pharmazeutische Biologie, FSU Jena. Angestrebter Abschluss: Promotion. Thema: „Untersuchungen zu positionsspezifisch integrierenden Retrotransposons in sozialen Amöben“. Betreuer: Prof. Dr. Thomas Winckler.

Thomas Spaller

Jena, den 20.1.2015

Ehrenwörtliche Erklärung

Hiermit erkläre ich, gemäß der Promotionsordnung der Biologisch-Pharmazeutischen Fakultät der FSU Jena vom 04. Dezember 2012, dass die vorliegende Arbeit selbständig verfasst und nur die angegebenen Quellen und Hilfsmittel verwendet wurden.

Die geltende Promotionsordnung der Biologisch-Pharmazeutischen Fakultät der Friedrich-Schiller-Universität Jena ist mir bekannt.

Die Hilfe eines Promotionsberaters wurde von mir nicht in Anspruch genommen.

Niemand hat von mir unmittelbar oder mittelbar geldwerte Leistungen für Arbeiten erhalten, die im Zusammenhang mit der vorgelegten Dissertation stehen.

Die vorliegende Arbeit wurde in gleicher oder ähnlicher Form noch bei keiner anderen Prüfungsbehörde oder Hochschule eingereicht.

Thomas Spaller

Jena, den 20.1.2015

Danksagung

Mein besonderer Dank gilt Herrn Prof. Dr. Thomas Winckler für die Bereitstellung des Themas und die zahlreichen, sehr lehrreichen Gespräche rund um die alltäglichen Laborprobleme, oder auch die „großen Fragen“ des bearbeiteten Themas.

Des Weiteren möchte ich Herrn Dr. Gernot Glöckner danken. Die Anregungen in den Gesprächsrunden waren stets hilfreich für den Fortgang der Arbeit. Die Bereitstellung von, zum Beginn der Arbeit, unpublizierten Genomdaten bildete eine wesentliche Basis für die vorliegende Arbeit.

Ein großer Dank geht an meine „Bürokollegen“: Dr. Friedemann Gaube, Robert Penzis und Eva Kling. Euch möchte ich für die allzeit gute Laune und die extrem tollen Geschenke danken, mit denen ich meinen Arbeitsplatz verzieren durfte! Auch für die Tipps während der Erstellung dieser Arbeit bin ich euch dankbar (jetzt kenne ich die Bedeutung von geschützten Leerzeichen)!

Allen anderen Kollegen am Lehrstuhl für pharmazeutische Biologie danke ich für das wunderbare Arbeitsklima und die Hilfe bei allen möglichen Fragen während der täglichen Arbeit.

Ein großer Dank geht an Herrn PD Dr. Matthias Platzer für die Möglichkeit der Hochdurchsatz-Sequenzierung in seiner Arbeitsgruppe und die Begutachtung der vorliegenden Arbeit.

In diesem Zusammenhang möchte ich auch Herrn Dr. Marco Groth für seine Hilfe bei der Sequenzierung und der anschließenden Auswertung der Daten danken.

Herrn Prof. Dr. Gerald Schumann danke ich für seine Bereitschaft zur Begutachtung der Arbeit.

Ein großer Dank geht an meine Eltern, die mich überhaupt erst dazu gebracht haben Biologie zu studieren. Vor allem ihre Unterstützung, auch wenn es mal nicht so lief, hat mich immer wieder motiviert.

Abschließend gilt mein Dank vor allem meiner Familie. Meiner Frau Jessica danke ich für ihre Geduld und ihre Mühe, mir den Rücken frei zu halten, ohne die diese Arbeit nicht möglich gewesen wäre! Sie hat es geschafft, mich auch in schwierigen Phasen zu unterstützen und mit viel Verständnis so manche schlechte Laune („wenn mal wieder was nicht geklappt hat“) akzeptiert. Ein großer Dank gilt auch meinen beiden Kindern Emma und Ben.