

**Vergleichende Analyse
der Gene und der Genomstruktur
der geschlechtsbestimmenden
Chromosomenregion
von *Chironomus* und
*Camptochironomus***

Dissertation

zur Erlangung des Grades
Doktor der Naturwissenschaften
(Dr. rer. nat.)

am Fachbereich Biologie der
Johannes Gutenberg-Universität in Mainz



Thomas Herold

Geboren am 08. November 1981 in Hanau

Im Oktober 2012

Dekan:

1. Berichtstatter:

2. Berichtstatter:

Tag der mündlichen Prüfung: 06.11.2012

Inhaltsverzeichnis

1	Einleitung.....	1
1.1	<i>Chironomus</i>	1
1.2	Mechanismen der Geschlechtsbestimmung bei Insekten und ihr einheitliches Grundprinzip	3
1.2.1	Heterogametie	3
1.2.2	Haplo-/ Diploidie (Arrhenotokie)	4
1.2.3	Maternale Monogenie	5
1.2.4	Umweltbedingte Geschlechtsdetermination.....	6
1.3	Geschlechtsbestimmende Gene	6
1.3.1	Geschlechtsbestimmende Gene in Säugetieren.....	6
1.3.2	Geschlechtsbestimmende Gene bei <i>Drosophila melanogaster</i>	8
1.3.3	Geschlechtsbestimmende Gene bei <i>Musca domestica</i> und <i>Ceratitis capitata</i>	11
1.4	Evolution der Geschlechtschromosomen	14
1.5	Geschlechtsdetermination in <i>Chironomus</i>	17
1.6	Zielsetzung der Arbeit	21
2	Material und Methoden	22
2.1	Zellkultur	22
2.2	Versuchstiere	22
2.3	Bestimmung des Geschlechts von Chironomiden	22
2.4	Molekularbiologische Standardmethoden.....	23
2.4.1	Phenol-Chloroform-Isoamylalkohol (PCI)-Extraktion	23
2.4.2	Fällung von DNA (Ethanol-fällung)	23
2.4.3	Konzentrationsbestimmung (DNA)	24
2.4.4	Agarose-Gelelektrophorese	24
2.4.5	Wiedergewinnung von DNA aus Agarosegelen.....	25
2.4.6	Restriktionen	25
2.5	DNA-Präparation.....	25
2.5.1	Isolierung genomischer DNA aus Zellkultur.....	25
2.5.2	Isolierung von Plasmid-DNA	26
2.5.2.1	Plasmid-Präparation nach dem Whitehead-Protokoll (Whitehead Institute/ MIT)	26
2.5.2.2	Plasmid-Präparation im 96er Mikrotiter-Plattenformat	27
2.6	Klonierung.....	28
2.6.1	Verwendeter Bakterienstamm	28
2.6.2	Herstellung elektrokompenter Zellen	28
2.6.3	Verwendete Vektoren.....	29

2.6.4	Ligation	29
2.6.5	Transformation	29
2.7	Erstellen einer BAC-Bibliothek aus <i>C. tentans</i>	29
2.7.1	Integrat-DNA	29
2.7.2	pBeloBAC11-Vektor Vorbereitung	30
2.8	Erstellen einer Shotgun-Bibliothek.....	30
2.9	DNA-Aufreinigung	31
2.9.1	DNA-Aufreinigung mittels CsCl-Dichtegradientenzentrifugation	31
2.9.2	Aufreinigung von PCR-Produkten	31
2.10	Primerdesign.....	32
2.11	Polymerase-Kettenreaktion (PCR)	32
2.12	DNA-Sequenzierung	34
2.13	Bioinformatische Sequenzauswertung	34
2.14	DNA-Markierungstechniken.....	34
2.15	Hybridisierungstechniken	35
2.15.1	Fluoreszenz in-situ-Hybridisierung (FISH).....	35
2.15.2	Southern-Hybridisierung.....	35
2.15.3	Koloniefilter-Hybridisierung	36
2.16	RNA-Präparation (GTC-Methode)	37
2.17	DNase-Verdau	37
2.18	Fällung von RNA	37
2.19	Poly-A ⁺ -Aufreinigung	38
2.20	Qualitäts- und Konzentrationsbestimmung von RNA per BioAnalyzer	38
2.21	cDNA-Synthese (Reverse Transkription).....	39
2.22	„Rapid Amplification of cDNA ends“ (RACE)	39
2.23	Illumina HiSeq 2000	39
3	Ergebnisse.....	40
3.1	Klonierung der SDR von <i>C. tentans</i>	40
3.1.1	Screening der BAC-Bibliotheken von <i>C. tentans</i>	40
3.1.2	Repetitive Elemente	44
3.1.3	Lokalisation der BAC-Klone per FISH	45
3.1.4	Überprüfung der Organisation der SDR von <i>C. tentans</i> mit Hilfe von Southern-Analysen.....	48
3.2	Analyse der Struktur der in der SDR befindlichen Gene per RT-PCR und RACE sowie der Hochdurchsatz-RNA-Sequenzierung.....	55
3.2.1	Analyse der Promotorbereiche der Gene der SDR.....	78

4	Diskussion	85
4.1	Vergleich der geschlechtsbestimmenden Region (SDR) von <i>C. tentans</i> und <i>C. thummi</i> sowie <i>C. thummi</i> ♀ und <i>C. thummi</i> ♂	85
4.2	Die Gene der SDR	93
4.2.1	Analysen der RACE-Ergebnisse	93
4.2.1.1	Gene mit verschiedenen 5'UTR-Varianten	93
4.2.1.2	„Scheinbar“ multiple 3'UTR-Transkriptlängen durch Fehlpriming des Oligo-d(T)-Primers	95
4.2.2	Analyse und Vergleich der Promotorbereiche der Gene der SDR aus <i>C. thummi</i>	97
4.2.3	Gene und ihre Produkte innerhalb der SDR von <i>Chironomus</i> sowie ihre potentielle Beteiligung an der Geschlechtsbestimmung.....	100
4.2.4	Gene und ihre Produkte der SDR mit NLS-Signal	102
4.2.5	Das Gen <i>luc7(p)</i> -like und seine mögliche <i>tra</i> -ähnliche Funktion	103
4.2.6	Potentielle Kandidaten-Gene für <i>M</i>	110
4.2.6.1	<i>fs(1)K10</i> -like	110
4.2.6.2	<i>polyhomeotic</i> -like	112
4.2.6.3	<i>rpn5</i> -like	113
4.3	Zusammenfassung und Ausblick.....	115
5	Anhang	117
5.1	Standardlösungen und Materialien	117
5.2	Bezugsquellen	120
5.3	Primer	122
5.3.1	Analyse der SDR bei <i>C. tentans</i>	122
5.3.2	RT-PCR und RACE Primer	124
5.3.3	Primer Gene Racer Kit/ dUTP-Kit/ Oligo-dT	127
6	Literaturverzeichnis	129
7	Eidesstattliche Erklärung	150

Tabellenverzeichnis

Tabelle 2.1: PCR-Ansätze	33
Tabelle 2.2: Programm einer PCR/ RT-PCR.....	33
Tabelle 2.3: Programm einer „Touchdown“-PCR (links) sowie einer Long-range-PCR (rechts).....	33
Tabelle 3.1: Übersicht über die zum Screenen der BAC-Bibliotheken verwendeten Sonden sowie die Lage der im Contig SDR von <i>C. tentans</i> liegenden Klone.....	42
Tabelle 3.2: Lage der Gene sowie der repetitiven Elemente im Contig SDR von <i>C. tentans</i>	45
Tabelle 3.3: Sequenzidentitäten der verschiedenen Gene und ihrer Produkte der SDR zwischen den beiden Arten <i>C. tentans</i> und <i>C. thummi</i>	77
Tabelle 3.4: Analysierte Promotorelemente	78
Tabelle 4.1: Lage repetitiver und transposabler Elemente im Contig SDR von <i>C. thummi</i>	86
Tabelle 4.2: Überprüfung der 3'-RACE-Produkte auf PolyA-Signale.....	96
Tabelle 5.1: Primer SDR-Analyse <i>C. tentans</i>	122
Tabelle 5.2: Genspezifische Primer RT-PCR.....	124
Tabelle 5.3: Primer 5' RACE.....	125
Tabelle 5.4: Primer 3' RACE.....	126
Tabelle 5.5: Primer Gene Racer Kit (Invitrogen)	127
Tabelle 5.6: Oligo-dT Primer.....	128
Tabelle 5.7: Primer dUTP Kit.....	128

Abbildungsverzeichnis

Abbildung 1.1: Kreuzungsschema der komplementären Geschlechtsdetermination ...	5
Abbildung 1.2: Mechanismus der Geschlechtsdetermination in <i>D. melanogaster</i>	11
Abbildung 1.3: Mechanismus der Geschlechtsdetermination bei <i>M. domestica</i>	13
Abbildung 1.4: Lage der BAC-Klone aus „Chromosomenwalk“ in <i>C. thummi</i>	20
Abbildung 3.1: Übersicht über die Lage der Gene der SDR von <i>C. thummi</i>	41
Abbildung 3.2: Übersicht über den Contig SDR von <i>C. tentans</i> inklusive der darin liegenden Gene sowie repetitiven Elemente	43
Abbildung 3.3: „In situ“-Hybridisierung der verschiedenen BAC-Klone aus <i>C. tentans</i> an Polytäanchrosomen der Spezies <i>C. thummi</i>	47
Abbildung 3.4: Autoradiogramm des genomischen Southern-Blots von <i>C. tentans</i> nach Hybridisierung mit der cDNA-Sonde <i>mi-er1-like</i>	49
Abbildung 3.5: Autoradiogramm des genomischen Southern-Blots von <i>C. tentans</i> nach Hybridisierung mit der cDNA-Sonde <i>luc7(p)-like</i>	50
Abbildung 3.6: Autoradiogramm des genomischen Southern-Blots von <i>C. tentans</i> nach Hybridisierung mit der cDNA-Sonde <i>fs(1)K10-like</i>	51
Abbildung 3.7: Proximaler Teil des Contigs SDR von <i>C. tentans</i> inklusive der dort lokalisierten BAC-Klone sowie der für die Southern-Analysen verwendeten Sonden ...	53
Abbildung 3.8: Distaler Teil des Contigs SDR von <i>C. tentans</i> inklusive der dort lokalisierten BAC-Klone sowie der für die Southern-Analysen verwendeten Sonden ...	54
Abbildung 3.9: Gelbild der RT-PCR mit dem Primerpaar <i>polyhomeotic_3for/3rev</i>	56
Abbildung 3.10: Lage und Struktur des Gens <i>mi-er1-like</i> im Contig SDR von <i>C. thummi</i>	57
Abbildung 3.11: Domänen im Protein MI-ER1-like nach Suche mit InterProScan	58
Abbildung 3.12: Lage und Orientierung des Gens <i>luc7(p)-like</i> im Contig SDR von <i>C. thummi</i>	59
Abbildung 3.13: Domänen im Protein Luc7(p)-like nach Suche mit InterProScan	59
Abbildung 3.14: Lage und Struktur des Gens <i>wd74-like</i> im Contig SDR von <i>C. thummi</i>	60
Abbildung 3.15: Domänen im Protein wd74-like nach Suche mit InterProScan	61
Abbildung 3.16: Lage und Struktur des Gens <i>fs(1)K10-like</i> im Contig SDR von <i>C. thummi</i>	62
Abbildung 3.17: Protein-Alignment der drei <i>fs(1)K10</i> -Varianten (X, Y1 und Y2)	63
Abbildung 3.18: Struktur der männchenspezifischen Duplikation des Gens <i>fs(1)K10-like</i> (Y ₁ ; Y ₂) in <i>C. thummi</i>	64
Abbildung 3.19: Lage und Struktur des Gens <i>bcn92-like</i> im Contig SDR von <i>C. thummi</i>	65
Abbildung 3.20: Domänen im Protein bcn92-like nach Suche mit InterProScan	66
Abbildung 3.21: Lage und Struktur des Gens <i>dLacDH-like</i> im Contig SDR von <i>C. thummi</i>	67
Abbildung 3.22: Domänen im Protein dLacDH-like nach Suche mit InterProScan	68
Abbildung 3.23: Lage und Struktur des Gens <i>polyhomeotic-like</i> im Contig SDR von <i>C. thummi</i>	69

Abbildung 3.24: Domänen im Protein polyhomeotic-like nach Suche mit InterProScan	70
Abbildung 3.25: Lage und Struktur des Gens CG15207-like im Contig SDR von <i>C. thummi</i>	71
Abbildung 3.26: Lage und Struktur des Gens CG11203-like im Contig SDR von <i>C. thummi</i>	72
Abbildung 3.27: Position sowie Chromatogramme der allelischen Varianten im Gen CG11203-like in <i>C. tentans</i>	73
Abbildung 3.28: Lage und Struktur des Gens <i>rpn5</i> -like im Contig SDR von <i>C. thummi</i>	74
Abbildung 3.29: Domänen im Protein <i>rpn5</i> -like nach Suche mit InterProScan.....	75
Abbildung 3.30: Lage und Struktur des Gens <i>hermansky pudlak</i> -like im Contig SDR von <i>C. thummi</i>	76
Abbildung 3.31: Übersicht über die Promotorbereiche der Gene der SDR.....	84
Abbildung 4.1: Vergleich der repetitiven und transposablen Elemente zwischen <i>C. thummi</i> und <i>C. tentans</i> sowie deren Lage im Contig SDR.....	87
Abbildung 4.2: Vergleiche zweier Abschnitte mit repetitiven bzw. transposablen Elementen des Contigs SDR zwischen <i>C. thummi</i> und <i>C. tentans</i> per DotPlot.....	88
Abbildung 4.3: Überschneidung der transkribierten Bereiche von <i>luc7(p)</i> -like und <i>wd74</i> -like sowie Lage des Indels 116 im 3'UTR-Bereich der beiden Gene.	92
Abbildung 4.4: RS-Domäne des abgeleiteten Polypeptids von <i>luc7(p)</i> -like aus <i>Chironomus</i> sowie Sequenz-Vergleich zu einem konservierten TRA-Abschnitt aus <i>Drosophila</i>	108
Abbildung 4.5: Sequenzalignment des Protein-BLAST Treffers mit Homologie zu fs(1)K10 aus <i>D. melanogaster</i> nach Suche mit der fs(1)K10-Sequenz aus <i>C. thummi</i>	111

Abkürzungsverzeichnis

A	Absorption [nm]
Abb.	Abbildung
abs.	absolute
Amp	Ampicillin
ANT-C	Antennepedia-Komplex
AP	alkalische Phosphatase
AS	Aminosäure(n)
ATP	2'-Adenosintriphosphat
BAC	Bacterial Artificial Chromosome
Bio	Biotin
BLAST	„Basic local alignment search tool“
Bp	Basenpaare
BRE	„TFIIB recognition element“
BSA	Rinderserumalbumin (Bovines Serumalbumin)
BX-C	Bithorax-Komplex
bzw.	beziehungsweise
CBC	Cap-Bindekomplex
CBF	„CCAAT-box binding factor“
CCE	CCC-Element (transposables Element)
Ci	Curie Einheit der Radioaktivität (1 Curie = $3,7 \cdot 10^{10}$ Becquerels)
CI	Chloroform/Isoamylalkohol (24:1; v/v)
CO ₂	Kohlenstoffdioxid
<i>C. piger</i>	<i>Chironomus piger</i>
CsCl	Cäsiumchlorid
<i>csd</i> / CSD	Gen/ Protein „ <i>complementary sex determiner</i> “
<i>C. tentans</i>	<i>Chironomus tentans</i>
<i>C. thummi</i>	<i>Chironomus thummi</i>
CtY	Gen aus <i>C. thummi</i> ; Y-chromosomal entdeckt
DIG	Digoxigenin
<i>dLacDH</i>	Gen d-Laktatdehydrogenase
DMF	Dimethylformamid
DNA	„Desoxyribonucleic acid“ (Desoxyribonukleinsäure)
DNase	Desoxyribonuklease
dNTP	2'-Desoxynucleosidtriphosphat
DPE	„downstream core promoter element“

DR	„direct repeat“
dsDNA	doppelsträngige DNA
dsx/ DSX	Gen/ Protein „ <i>doublesex</i> “
<i>E. coli</i>	<i>Escherichia coli</i>
EDTA	Ethylendiamintetraacetat
ELM2	Proteindomäne (Homologie zu EGL-27 und MTA1)
E-Puffer	Tris-Phosphat-Elektrophoresepuffer
ESE	„exonic splicing enhancer“
EtOH	Ethanol
et al.	et altera
EtBr	Ethidiumbromid
Fa.	Firma
FCS	„fetal calf serum“ (fötales Kälberserum)
FISH	Fluoreszenz „in situ“-Hybridisierung
FITC	Fluoresceinisothiocyanat
for	forward-(Primer)
g	Einheit der Erdbeschleunigung (gravity)
°C	Grad Celsius
GFP	grün fluoreszierendes Protein
ggf.	gegebenenfalls
<i>grk</i>	Gen „ <i>gurken</i> “
GTC	Guanidinthiocyanat
HMG	Proteindomäne („High-mobility group“)
HPLC	„High-Performance Liquid Chromatography“
IMSB	Institut für Molekulargenetik, gentechnologische Sicherheitsforschung und Beratung
Indel	Insertion/ Deletion
IPTG	isopropyl-1-thio-β-d-galactopyranosid
IR	„inverted repeat“
kb	Kilo-Basenpaare (10 ³ Basenpaare)
kV	Kilovolt = 10 ³ V
L	Liter
LB-Medium	Luria Broth; Bakterienmedium
λ	lambda
M	molar
mA	Milli-Ampere
MCS	„multiple cloning site“ (multiple Klonierungsschnittstelle)

<i>mi-er1</i>	Gen „mesoderm induction early response prot. 1“
min	Minute
ml	Milliliter
µl	Mikroliter
mg	Milligramm
µg	Mikrogramm
mM	Millimolar
mRNA	messenger-RNA
ms	Millisekunden
MSY	„male specific region on the Y chromosome“
MTP	Mikrotiterplatte
NaOH	Natriumhydroxid
NCBI	„National Center for Biotechnology Information“
NLS	Kern-Lokalisierungs-Signal
nm	Nanometer
nt	Nukleotid(e)
OD	Optische Dichte
o. g.	oben genannt
ORF	offener Leserahmen („open reading frame“)
PBS	„phosphate buffered saline“
PcG	„polycomb“-Gruppe
PCI	Phenol/Chloroform/Isoamylalkohol (25:24:1; v/v/v)
PCI	Proteindomäne „proteasome, COP9 signalosome, initiation factor 3“
PCR	Polymerasekettenreaktion („polymerase chain reaction“)
Pen/ Strep	Penicillin/ Streptomycin
pH	negativer, dekadischer Logarithmus der Wasserstoffionenkonzentration
PM	Präinkubationsmedium
pg	Picogramm
pmol	Picomol
RACE	rapid amplification of cDNA ends
Read	Einzelsequenz (Hochdurchsatzsequenzierung)
rev	reverse-(Primer)
RNA	Ribonukleinsäure
RNase	Ribonuklease
<i>rpn/ RPN</i>	Gen/ Protein „ <i>ribonucleoprotein</i> “

RRM	Proteindomäne „RNA recognition motif“
RS	Proteindomäne (mit RS- und SR-Dimeren)
RT	Raumtemperatur
RT-PCR	Reverse Transkriptase PCR (PCR auf cDNA)
SAM	Proteindomäne („sterile alpha motif“)
SANT	Proteindomäne
SDR	geschlechts-determinierende Region
SDS	Sodiumdodecylsulfat
sec.	Sekunde
SINE	„short interspersed nuclear element“
snRNP	„small nuclear Ribonucleoprotein particle“
SPT2	Proteindomäne
<i>sry</i> / SRY	Gen/ Protein „ <i>sex determining region Y</i> “
SSC	Standard-Salz-Citrat („Standard-Saline-Citrat“)
Std.	Stunde
s. o.	siehe oben
s. u.	siehe unten
<i>sxl</i> / Sxl	Gen/ Protein „ <i>sex-lethal</i> “
<i>Taq</i> -Polymerase	<i>Thermus aquaticus</i> DNA-Polymerase
TBE	Tris-Borat-EDTA
TE	Tris-EDTA
TECth1	Transposables Element <i>C. thummi</i> 1
TFB1	„Thummi-Foldback“-Element (transpos. Element)
<i>tra</i> , <i>tra2</i> / TRA, TRA2	Gen/ Protein „ <i>transformer</i> , <i>transformer2</i> “
Tris	Tris(hydroxymethyl)aminomethan
TRITC	Tetramethylrhodaminisothiocyanat
U	Unit (Einheit)
Upm	Umdrehungen pro Minute
UTR	untranslatierter Bereich („untranslated region“)
UV	Ultraviolett
UZ	Ultrazentrifuge
VE-Wasser	voll-entsalztes Wasser
vgl.	vergleiche
Vol.	Volumen
v/v	Volumen/Volumen (Volumenprozent)
w/v	Masse/Volumen (Gewichtsprozent)
X-Gal	5-Brom-4-chlor-3-indolyl- α -D-galactopyranosid
z. B.	zum Beispiel

1 Einleitung

Die Existenz von weiblichen und männlichen Individuen ist bei Eukaryoten der Normalfall. Bei höheren Eukaryoten erfolgt die Fortpflanzung in den meisten Fällen nach dem Prinzip der geschlechtlichen Vermehrung. Hierbei unterscheidet man zwischen Monözie und Diözie. Bei der Monözie werden sowohl männliche wie auch weibliche Gameten von einem Individuum gebildet. Dies kommt insbesondere bei Pflanzen sowie marinen Invertebraten vor. Hingegen werden bei der Diözie die Gameten von unterschiedlichen Individuen derselben Art gebildet. Diözie findet man vor allem in höheren Eukaryoten wie Arthropoden und Vertebraten (Maynard Smith 1978).

Die Ausprägung des Geschlechts wird in den meisten getrenntgeschlechtlichen Arten durch Geschlechtschromosomen und den darauf liegenden geschlechtsdeterminierenden Faktoren bestimmt. In einigen Spezies wirken äußere Einflüsse, wie z.B. die Temperatur auf die Determination des Geschlechts mit ein. Die Mechanismen der Geschlechtsbestimmung sind hierbei wiederum recht vielfältig.

Im Folgenden soll zunächst die Familie der Chironomidae (Zuckmücken) mit den beiden hier untersuchten Vertretern *C. tentans* und *C. thummi* vorgestellt werden. Des Weiteren soll ein Überblick über die verschiedenen Mechanismen der Geschlechtsdetermination gegeben werden. Im Vordergrund werden hierbei die verschiedenen Geschlechtsbestimmungs-Mechanismen in der Klasse der Insekten stehen. Außerdem soll zum Abschluss ein Einblick in das Thema der Evolution der Geschlechtschromosomen gegeben werden.

1.1 *Chironomus*

Chironomiden gehören innerhalb der Klasse der Insekten zur Ordnung der Diptera (Zweiflügler), der Unterordnung der Nematocera (Mücken) und der Familie der Chironomidae (Zuckmücken). In Deutschland kommen etwa 700 der rund 10000-15000 weltweit bekannten Chironomiden-Arten vor (Armitage et al. 1995).

Innerhalb der Chironomiden gibt es verschiedene Gruppen, welche sich durch reziproke Translokationen von verschiedenen Chromosomenarmen voneinander unterscheiden lassen. So spricht man bei der untersuchten Art *C. tentans* von Camptochironomiden, wohingegen die beiden Arten *C. thummi* und *C. piger* (Schwestern-Arten) zu den Chironomiden zählen. Diese beiden Gruppen unterscheiden sich durch eine reziproke Translokation der beiden Chromosomenarme CD und EF hin zu DE und CF voneinander (Keyl 1962). Die untersuchte Art *C. thummi* (Kieffer 1911), welche auch als *C. riparius* bezeichnet wird, wurde erstmals von Meigen (1804) beschrieben. Credland (1973) zeigte, dass die beiden durch Kieffer und Meigen beschriebenen Arten identisch sind.

Bei Chironomiden handelt es sich um die größte und am weitesten verbreitete Gruppe von im Süßwasser lebenden holometabolen Insekten (Pinder 1986). Diese durchlaufen während ihres Lebenszyklus eine vollständige Metamorphose mit den vier Stadien Ei,

Larve, Puppe und Imago. Zur Paarung bilden die Männchen vieler Arten sogenannte Hochzeitsschwärme (Kon 1987). Durch diese werden Weibchen angelockt, und es kommt zur Kopulation. Die Übertragung des Spermas erfolgt hierbei in Form einer Spermatophore (Downes 1969). Die Eiablage erfolgt im Anschluss entweder durch das Abwerfen der Eiballen über der Wasseroberfläche oder durch das Anheften an im Wasser befindliches Substrat. Nach rund drei Tagen schlüpfen aus den Eiern die Larven. Diese durchlaufen meist vier distinkte Larvenstadien. Das Larvenstadium stellt in der Regel den längsten Abschnitt im Lebenszyklus einer Chironomide dar (Pinder 1986; Armitage et al. 1995). Die Larven ernähren sich von Algen und Detritus, welche sie aus dem Wasser filtrieren. Im Allgemeinen leben die Larven der Chironomiden im Sediment oder an dessen Oberfläche. Mit Hilfe von Substrat bauen sie sich ihre Wohnröhren, welche durch das in den Speicheldrüsen (Spinndrüsen) der Larven erzeugte Spinnsekret zusammengehalten werden. In den Spinndrüsen besitzen die Larven gut ausgebildete polytäre Riesenchromosomen (Polytäniegrad bis $2^{14}n = 16384n$). An diesen wurden durch Balbiani (1881) erstmals die so genannten Balbiani-Ringe entdeckt und beschrieben. Das auf das Larvenstadium folgende Puppenstadium ist mit rund drei Tagen das kürzeste Stadium im Leben der Chironomiden. Allerdings geht es mit tiefgreifenden inneren wie äußeren Veränderungen der Organisation des Körpers einher. Am Ende des Puppenstadiums treiben die Puppen an die Wasseroberfläche. Hier schlüpfen innerhalb kürzester Zeit die adulten Chironomiden. Die Geschlechter der adulten Tiere lassen sich leicht anhand ihres Geschlechts-Dimorphismus unterscheiden. Männliche Tiere besitzen wirtelig behaarte Antennen, wohingegen die Antennen der Weibchen unbehaart sind. Je nach Lebensdauer ernähren sich die Imagines entweder von Pollen und Nektar oder nehmen gar keine Nahrung auf (Schlee 1977). Bei den meisten Arten der Chironomiden dauert der gesamte Lebenszyklus nur einige Wochen. Diese Arten durchlaufen häufig mehrere Lebenszyklen pro Jahr und werden daher als bivoltin (zwei Lebenszyklen) oder multivoltin bezeichnet.

Mit Hilfe des für Insekten untypischen Blutfarbstoffs Hämoglobin ist es Chironomiden möglich, Sauerstoff effektiv aus dem Wasser aufzunehmen und somit selbst in extrem sauerstoffarmen Gewässern zu überleben (Weber et al. 1985).

Auf Grund der unterschiedlichen Bedürfnisse der verschiedenen Chironomiden-Arten bezüglich Sauerstoffgehalt, pH-Wert, Salzgehalt sowie Temperatur des Wassers eignen sich die Larven der unterschiedlichen Arten gut als Bioindikatoren für Umweltbelastungen im Wasser (Oliver 1971; Pinder 1986; Taenzler et al. 2007).

In der Medizin gelten Chironomiden als Auslöser von allergischen Reaktionen sowie Verursacher bronchialen Asthmas (Kay et al. 1978; Baur et al. 1980).

1.2 Mechanismen der Geschlechtsbestimmung bei Insekten und ihr einheitliches Grundprinzip

In Insekten ist eine große Fülle von unterschiedlichen Mechanismen der Geschlechtsbestimmung beschrieben (Bull 1983). Trotz ihrer Fülle lassen sich diese in drei Hauptkategorien einordnen. Hierbei ist das primäre geschlechtsdeterminierende Signal entweder zygotischen, maternalen oder externen, umweltbedingten Ursprungs (Sánchez 2008).

Bereits 1985 wurde die Hypothese aufgestellt, dass bei allen Insekten ein gemeinsames Prinzip der Geschlechtsbestimmung vorhanden ist (Nöthiger und Steinmann-Zwicky 1985). Nach dieser Hypothese leiten sich alle Mechanismen der Geschlechtsdetermination von einem einfachen Prinzip der männlichen Heterogamete ab. Die Hypothese beruht auf der Beobachtung, dass in nah verwandten Spezies oder sogar in verschiedenen Stämmen innerhalb derselben Spezies unterschiedliche Mechanismen der Geschlechtsbestimmung auftreten können. Der Hypothese nach entstehen diese vom Grundprinzip abweichenden Mechanismen der Geschlechtsdetermination durch allelische Variationen oder Mutationen innerhalb der Kontrollgene. Nach der Hypothese liegt den verschiedenen Mechanismen der Geschlechtsbestimmung bei Insekten ein einfaches Modell mit den drei Faktoren M , F und F' zugrunde. Hierbei führen aktive Produkte von F zu einer weiblichen Differenzierung. Allerdings wird F durch das dominante Allel M reprimiert. Bei Individuen, die dieses dominante Allel M tragen, erfolgt eine männliche Differenzierung. In einigen Fällen kann für die Expression aktiver F Produkte und somit der weiblichen Differenzierung zusätzlich das maternale Genprodukt F' erforderlich sein.

Im Folgenden sollen einige Mechanismen der Geschlechtsdetermination kurz vorgestellt werden.

1.2.1 Heterogamete

Bei Heterogamete entscheidet die An- oder Abwesenheit eines dominanten geschlechtsdeterminierenden Faktors über die Differenzierung des Geschlechts.

Bei der weit verbreiteten Form der männlichen Heterogamete erzeugen Männchen verschiedene Gameten, nämlich solche mit einem X- oder mit einem Y-Chromosom. Weibchen hingegen erzeugen nur einheitliche Gameten mit jeweils einem X-Chromosom. Hieraus resultieren Zygoten die entweder $XX=\text{♀}$ oder $XY=\text{♂}$ sind. Der dominante geschlechtsdeterminierende Faktor M ist hierbei auf dem genetischen Y-Chromosom lokalisiert. Bei weiblichen Individuen fehlt dieser dominante Faktor M . In Folge dessen ist bei ihnen der weibliche Faktor F aktiv. Das zahlreiche Vorkommen der männlichen Heterogamete wurde durch Bull (1983) dokumentiert. Männliche Heterogamete kommt in Vertebraten bei Säugern, Reptilien, Amphibien und Fischen vor. In Säugetieren stellt die männliche Heterogamete hierbei den typischen Mechanismus der Geschlechtsdetermination dar. Bei den Invertebraten ist männliche Heterogamete in Arachniden und Crustaceen vertreten. Des Weiteren kommt sie in

vielen Insekten-Arten vor. Hierzu zählen die meisten Arten der Chironomiden (Bauer und Beermann 1952; Beermann 1955; Acton 1957; Keyl 1961; Keyl 1962; Rosin und Fischer 1972; Martin 1980; Hägele 1985), die meisten Stämme von *Musca domestica* (Stevens 1908; Dübendorfer et al. 2002), *Ceratitis capitata* (Willhoeft und Franz 1996), *Megaselia scalaris* (Mainx 1959; Mainx 1964; Mainx 1966) sowie *Calliphora erythrocephala* (Ullerich 1963).

Eine Sonderform der männlichen Heterogametie ist in *Drosophila melanogaster* zu finden. Die Entscheidung über die Bestimmung des Geschlechts erfolgt in diesem Fall nicht einfach durch die An- oder Abwesenheit des dominanten geschlechtsbestimmenden Faktors *M*. Vielmehr ist das Verhältnis von Autosomen zu X-Chromosomen entscheidend für die Determination des Geschlechts (Bridges 1921). Man spricht hierbei auch kurz vom X:A-Verhältnis. In Individuen mit einem X:A-Verhältnis von 1 (XX:AA) kommt es zur weiblichen Differenzierung, bei einem X:A-Verhältnis von 0,5 (X:AA) zur männlichen Differenzierung (Cline und Meyer 1996; Schütt und Nöthiger 2000). Detailliert wird hierauf in Kapitel 1.3.2 eingegangen.

Bei weiblicher Heterogametie bilden die Weibchen zwei verschiedene Arten von Gameten, während Männchen keine verschiedenen Gameten erzeugen. Zur besseren Abgrenzung von der männlichen Heterogametie verwendet man bei weiblicher Heterogametie die Bezeichnungen W und Z für die Geschlechtschromosomen. Hierbei liegt in weiblichen Individuen die Chromosomenkonstellation WZ und in männlichen Individuen die Konstellation ZZ vor. Nach der Hypothese von Nöthiger und Steinmann-Zwicky (1985) liegt bei weiblicher Heterogametie der dominante geschlechtsbestimmende Faktor F^D auf dem W-Chromosom. Das dominante Allel F^D wirkt epistatisch über *M* und bestimmt die Differenzierung zum weiblichen Geschlecht. Bei Vertebraten kommt weibliche Heterogametie vor allem bei Vögeln und Fischen vor (Stevens 1997). Bei Insekten kommt sie außerdem unter den Köcherfliegen (Bull 1983), den Schmetterlingen wie z.B. *Bombyx mori* (Traut und Mosbacher 1968; Bull 1983; Fujii und Shimada 2007) und einigen Stämmen von *M. domestica* (Franco et al. 1982; Hilfiker-Kleiner et al. 1993; Schmidt et al. 1997) vor.

1.2.2 Haplo-/ Diploidie (Arrhenotokie)

Haplo-/ Diploidie kommt in der Ordnung der Hautflügler (Hymenoptera) vor. Aus unbefruchteten Eiern entwickeln sich männliche Individuen, die haploid sind, wohingegen sich befruchtete Eier zu diploiden, weiblichen Individuen entwickeln (Whiting und Whiting 1925). Mit Hilfe von Kreuzungsexperimenten an der parasitären Wespe *Bracon hebetor* konnte gezeigt werden, dass hier eine komplementäre Geschlechtsdetermination vorliegt (Whiting 1933; Whiting 1943). Weibliche Differenzierung findet immer dann statt, wenn verschiedene Allele eines geschlechtsbestimmenden Faktors *F* (Single-Locus) oder mehrerer geschlechtsbestimmender Faktoren *F* (Multi-Locus) heterozygot vorliegen. Zur männlichen Entwicklung kommt es im Gegensatz hierzu, wenn die Tiere hemi- oder homozygot für den oder die entsprechenden geschlechtsbestimmenden Faktor(en)

sind (Crozier 1971). Zur Verdeutlichung siehe hierzu auch Abbildung 1.1. Der Einfachheit halber soll hier nur auf die Form des „Single-Locus“ eingegangen werden. Nach der Hypothese von Nöthiger und Steinmann-Zwicky (1985) hat der geschlechtsdeterminierende Faktor *M* im System der Arrhenotokie seine Funktion verloren.

Parental	♀ aa'	x	♂ a
F1	aa' ♀ (diploid; heterozygot)		a ♂ (unbefruchtet; haploid)
Rückkreuzung von Sohn (a) und Mutter (aa')	aa' ♀ (diploid; heterozygot)		aa ♂ (diploid; homozygot)

Abbildung 1.1: Kreuzungsschema der komplementären Geschlechtsdetermination

Zur weiblichen Differenzierung kommt es durch das heterozygote Vorliegen von Allelen eines oder mehrerer geschlechtsbestimmender Faktoren. Liegen diese hingegen hemi- oder homozygot vor, kommt es zur männlichen Differenzierung.

Nachgewiesen wurde der geschlechtsdeterminierende Mechanismus der Haplo-/Diploidie an der Honigbiene *Apis mellifera* (Drescher und Rothebbuhler 1964) sowie der Ameise *Solenopsis invicta* (Ross und Fletcher 1985). Der als *csd* (*complementary sex determiner*) bezeichnete Faktor stellt das primäre Signal in der Determination des Geschlechts in Hymenopteren dar (Cook 1993). In *Apis* existieren mindestens 15 Allele eines einzelnen *csd*-Gens. Dieses kodiert für ein Arginin-Serin-reiches (SR-) Protein, welches strukturelle Ähnlichkeit zu TRA-Proteinen aus Dipteren zeigt. Das Gen *csd* wird in *Apis mellifera* in beiden Geschlechtern transkribiert. Es besitzt eine hochvariable Region, welche zu verschiedenen Isoformen des Genprodukts führt. Es wird angenommen, dass für die weibliche Differenzierung ein funktioneller Komplex aus zwei unterschiedlichen Isoformen des Csd gebildet werden muss. In der Differenzierung zu männlichen Individuen wird nur eine Csd-Isoform gebildet. Hierdurch kann kein funktioneller Komplex gebildet werden, und es erfolgt die männchenspezifische Entwicklung (Beye et al. 2003; Beye 2004).

1.2.3 Maternale Monogenie

Bei maternaler Monogenie wird das Geschlecht der Zygote ausschließlich durch die An- oder Abwesenheit des maternalen Faktors *F'* in der Oozyte bestimmt. Hierbei findet man zwei Arten von weiblichen Individuen. Thelygene Weibchen sind heterozygot (*F'f*) für den Faktor *F'*. Die Nachkommen von diesen Weibchen sind immer weiblich, wobei man thelygene und arrhenogene Weibchen im Verhältnis 1:1 findet. Die Nachkommen von arrhenogenen Weibchen sind männlich. Bei arrhenogenen Weibchen liegt der Faktor *F'* homozygot rezessiv (*ff*) vor. Maternale Monogenie kommt bei Vertretern der Gattung *Sciara* (Metz und Schmuck 1929; Sánchez 2008) sowie *Chrysomya rufifacies* und *Chrysomya albiceps* vor (Ullerich 1963; Ullerich 1973; Ullerich 1975; Ullerich 1976; Ullerich 1984).

1.2.4 Umweltbedingte Geschlechtsdetermination

Bei der umweltbedingten Geschlechtsdetermination wurden verschiedene externe Umweltfaktoren wie Temperatur, pH-Wert und Hormone als primäres Signal identifiziert. Die Geschlechtsdetermination als Folge von bestimmten externen Gegebenheiten kommt bei vielen Reptilienarten sowie bei einigen Fischen vor. Im Buntbarsch *Cichlasoma citrinellum* ist die Größe des jungen Fisches ausschlaggebend für dessen Geschlecht im adulten Stadium (Francis und Barlow 1993). Hierbei entwickeln sich aus kleineren Fischen weibliche Individuen, wohingegen sich größere Fische zu männlichen Tieren entwickeln. Bei Vertretern der Schildkröten (Bull und Vogt 1979; Bull und Vogt 1981), Echsen (Charnier 1966; Bull 1987) sowie bei Vertretern der Krokodile (Janzen und Paukstis 1991) entscheidet die Inkubations-Temperatur der Eier während der Embryogenese über die geschlechtliche Differenzierung der Tiere. Hierbei kann sich eine niedrigere oder höhere Temperatur in verschiedenen Arten auch unterschiedlich auf die Entscheidung zwischen weiblichem und männlichem Geschlecht auswirken. Innerhalb der Insekten ist eine umweltbedingte Geschlechtsdetermination beispielsweise bei *Sciara ocellaris* (Nigro et al. 2007) sowie *Heteropeza pygmaea* (Went und Camenzind 1980) zu finden. Bei *Sciara ocellaris* führt eine erhöhte Temperatur (24-28°C) während der Embryogenese zu einer erhöhten weiblichen Nachkommenschaft. Eine erniedrigte Temperatur (12°C) hingegen führt zu mehr männlichen Nachkommen. Bei *Heteropeza pygmaea* entscheidet der Ernährungszustand der Mutter und deren Hämolymphe über die Entscheidung zwischen männlichen und weiblichen Nachkommen.

1.3 Geschlechtsbestimmende Gene

1.3.1 Geschlechtsbestimmende Gene in Säugetieren

In Säugetieren findet man überwiegend den Mechanismus der männlichen Heterogamete als geschlechtsdeterminierendes Prinzip. Hierbei entscheidet die Anwesenheit eines dominanten Faktors *M* auf dem Y-Chromosom über die Differenzierung des Organismus zu männlichen Individuen. Daher wird *M* auch als *TDF* (Testis-determinierender-Faktor) bezeichnet. Die beiden Heterosomen X und Y unterscheiden sich in der Regel morphologisch klar in ihrer Struktur und Größe. Das Y-Chromosom ist hierbei oftmals kleiner als das X-Chromosom. Große Bereiche des Y-Chromosoms bestehen aus konstitutivem Heterochromatin (Pathak und Stock 1974). Als typisches Beispiel hierfür kann der Mensch dienen.

Funktionell kann man das humane Y-Chromosom in zwei Bereiche unterteilen. Der Bereich, in welchem der *TDF* lokalisiert ist, unterscheidet sich stark vom X-Chromosom. Er zeigt keine Homologien zwischen den beiden Heterosomen. Innerhalb dieses Bereichs des Y-Chromosoms findet in der Regel auch keine Rekombination mit dem X-Chromosom statt. Der zweite Bereich des Y-Chromosoms besteht aus zwei Regionen, die während der Meiose mit homologen Bereichen des

X-Chromosoms paaren können. Sie werden deshalb auch als pseudoautosomale Regionen (PAR) bezeichnet. Für eine funktionierende Segregation der Heterosomen während der Meiose sind diese beiden pseudoautosomalen Regionen auf dem Y-Chromosom essentiell. Die beiden pseudoautosomalen Regionen befinden sich an den distalen Enden der beiden Arme des Y-Chromosoms (Ellis und Goodfellow 1989; Freije et al. 1992). Durch Untersuchungen von XX-Männern, welche translozierte Y-chromosomale Abschnitte aufwiesen, sowie von XY-Frauen mit Deletionen Y-chromosomaler Abschnitte konnte eine Eingrenzung des *TDF* auf eine Region in direkter Nachbarschaft zur pseudoautosomalen Region auf dem kurzen Arm des Y-Chromosoms erfolgen (Affara et al. 1986; Müller et al. 1986; Palmer et al. 1989).

Innerhalb dieser Region mit einer Größe von 35 kb konnte das *TDF*-Gen gefunden werden (Sinclair et al. 1990). Dieses liegt etwa 8 kb von der pseudosomalen Region entfernt und wurde als *SRY* (sex determining region Y-chromosome gene) bezeichnet (Goodfellow und Lovell-Badge 1993; Bogan und Page 1994). Durch die Ergebnisse der Untersuchungen von XY-Frauen, welche ein Y-Chromosom mit einer Punktmutation im *SRY*-Gen besitzen, die zu einer Verschiebung des ORFs führt, wurden die Vermutung erhärtet, dass es sich bei dem *SRY*-Gen um *TDF* handelt (Berta et al. 1990). Ein weiterer Hinweis, dass es sich bei *SRY* um *TDF* handelt, stammt aus Keimbahn-Transformationsexperimenten an Mäusen. In Mäusen wird das homologe Gen zu *SRY* mit *Sry* bezeichnet, der *TDF* als *Tdy*. Im Rahmen der Experimente wurde weiblichen Mäusen ein 14 kb großes Fragment aus der *Tdy*-Region injiziert, das das Gen *Sry* enthält. Weitere längere ORFs, die auf die Anwesenheit eines weiteren Gens hinweisen, fehlen innerhalb dieses Fragments (Gubbay et al. 1992). Die transgenen XX-Mäuse, die mit dem 14 kb Fragment transformiert worden waren, entwickelten Testes (Koopmann et al. 1991). Hieraus konnte geschlussfolgert werden, dass das 14 kb Fragment alle Informationen enthält, die zur Entwicklung der Testes nötig sind. Da innerhalb des Fragments keine weiteren Gene lokalisiert werden konnten, ist bewiesen, dass es sich bei *Sry* um das Gen für *Tdy* handelt.

Genauere Untersuchungen des *SRY*-Gens zeigten, dass es sich hierbei um ein kleines Gen ohne Introns handelt (Behlke et al. 1993). In den Testes liegen Transkripte mit 900 bp (Clepet et al. 1993) bzw. 1,1 kb (Su und Lau 1993) vor. Die unterschiedlichen Größen der Transkripte sind Folge von multiplen Transkriptionsstartpunkten (Clepet et al. 1993). Die aus dem ORF abgeleitete AS-Sequenz besitzt eine Länge von 204 AS. Sie zeigt eine HMG-(high-mobility group) Box mit einer Größe von 79 AS (Jantzen et al. 1990; Landsman und Bustin 1993; Grosschedl et al. 1994). *SRY* wird zusammen mit anderen *SRY*-ähnlichen Proteinen mit HMG-Box zur Familie der SOX-(*SRY*-ähnliche HMG-Box) Gene gezählt. Proteine mit einer HMG-Box haben eine DNA-bindende Funktion (Harley et al. 1992; Thomas 2001; Thomas und Travers 2001; Stros et al. 2007). Aus diesem Grund nimmt man an, dass es sich bei *SRY* um einen Transkriptionsfaktor handelt (Münsterberg und Lovell-Badge 1991; Haaq et al. 1993). Als ein möglicher Interaktionspartner des *SRY*-Genprodukts gilt KRAB (Krüppel-associated box only protein). Nach diesem Modell rekrutiert *SRY* den KRAB-KAP1 (KRAB associating protein 1)-Komplex, welcher Einfluss auf das Chromatin nimmt (Oh und Lau 2006). *SRY*-Homologe konnten außer in der Maus auch in Rindern

(Daneau et al. 1995), Marsupialiern (Foster et al. 1992; Foster und Graves 1994) und Primaten (Whitfield et al. 1993) isoliert werden. Innerhalb der Klasse der Säugetiere ist das Gen *SRY* im Taxon der Theria (Zitzentiere) hochkonserviert. Zu diesem Taxon zählen sowohl die Unterklasse der Marsupialia (Beuteltiere) wie auch die der Placentalia (Plazentatiere). In Vertretern der Monotremata (Kloakentiere) wie Platypus (Schnabeltier) oder Echidna (Ameisenigel) hingegen fehlt ein *SRY*-homologes Gen (Wallis et al. 2008).

1.3.2 Geschlechtsbestimmende Gene bei *Drosophila melanogaster*

Der in der Genetik häufig verwendete Modellorganismus *D. melanogaster* gehört zu der Familie der Taufliegen (Drosophilidae). In *D. melanogaster* liegt eine Sonderform der männlichen Heterogametie vor. Das Y-Chromosom besitzt in diesem Fall keine geschlechtsdeterminierende Funktion. Allerdings sind auf dem Y-Chromosom Gene lokalisiert, welche für die männliche Fertilität von Wichtigkeit sind. Die Determination des Geschlechts erfolgt in *D. melanogaster* über das Verhältnis von X-Chromosomen zu Autosomen (kurz X:A-Verhältnis) (Bridges 1921). Dieses Verhältnis stellt das primäre geschlechtsdeterminierende Signal dar.

Ziel dieses komplexen Mechanismus ist die Regulation des Schlüsselgens *Sex-lethal* (*Sxl*). Dieses steht an der Spitze einer Genregulationskaskade, welche über die Geschlechtsdetermination entscheidet.

Die Regulation des Gens *Sxl* erfolgt über die Aktivierung oder Repression des frühen Promotors von *Sxl*. Dieser wird bei einem X:A-Verhältnis von 1 aktiviert. Bei einem X:A-Verhältnis von 0,5 hingegen wird der frühe Promotor nicht aktiviert und *Sxl* bleibt zunächst inaktiv. Diese Entscheidung findet in einer frühen Phase der Embryogenese (vor dem Blastoderm-Stadium) zellautonom in jeder Zelle statt (Sánchez und Nöthiger 1983; Cline 1984; Gergen 1987; Bopp et al. 1991; Keyes et al. 1992). Die Regulation der Transkription von *Sxl* und somit die Umsetzung des X:A-Verhältnisses erfolgt über die Genklassen der maternalen Genprodukte, Numeratoren und Denominatoren. Alle drei Genklassen kodieren hierbei für Transkriptionsfaktoren.

Die Numeratorgene sind auf dem X-Chromosom lokalisiert. Bisher wurden vier Gene dieser Klasse identifiziert (Schütt und Nöthiger 2000). Es handelt sich hierbei um die Gene *sisterless-A* (*sisA*) (Cline 1988; Erickson und Cline 1993; Walker et al. 2000), *sisterless-B* (*sisB*), auch *scute* genannt (Cline 1988; Torres und Sanchez 1989; Erickson und Cline 1991; Parkhurst et al. 1993; Deshpande et al. 1995), *sisterless-C* (*sisC*), auch *outstretched* genannt (Sefton et al. 2000) sowie *runt* (*run*) (Duffy und Gergen 1991; Kramer et al. 1999). Das bislang einzige bekannte Gen aus der Gruppe der Denominatoren ist *deadpan* (*dpn*) (Parkhurst und Ish-Horowicz 1992; Younger-Shepherd et al. 1992; Sánchez et al. 1994). Im Gegensatz zu den Numeratoren liegt *dpn* auf einem Autosom. Das Genprodukt von *dpn* wirkt reprimierend auf die Transkription von *Sxl* (Schütt und Nöthiger 2000). Für die Numeratorgene *sisB* und *runt* sowie das Denominatorgen *dpn* konnte gezeigt werden, dass die jeweiligen Genprodukte an spezifischen Stellen des *Sxl*-Promotors binden können, um diesen so

aktivieren (*sisB*, *runt*) bzw. reprimieren (*dpn*) zu können (Hoshijima et al. 1995; Kramer et al. 1999).

Zusätzlich zu den Numerator- und Denominatorgenen sind maternale Genprodukte an der Regulation von *Sxl* beteiligt. Bisher konnten vier maternale Gene identifiziert werden, deren Produkte die Transkription von *Sxl* entweder positiv oder negativ regulieren. Positiv auf die Transkription wirken hierbei die Genprodukte der Gene *daughterless* (*da*) (Cronmiller und Cline 1987; Gergen 1987; Cronmiller et al. 1988) sowie *hermaphrodite* (*her*) (Pultz und Baker 1995). Die Genprodukte der beiden Gene *groucho* (*gro*) (Paroush et al. 1994; Aronson et al. 1997) und *extra macrochaetae* (*emc*) (Younger-Shepherd et al. 1992; Barbash und Cline 1995) wirken hingegen als negative regulatorische Faktoren auf die Transkription von *Sxl*. In Schütt und Nöthiger (2000) ist ein vereinfachtes Modell der Interaktion zwischen den Numeratoren (*sisA*, *sisB*), dem Denominator (*dpn*) sowie den maternalen Faktoren (*da*) und deren Wirkung auf den frühen *Sxl*-Promotor aufgezeigt.

Bei weiblichen Individuen wird der frühe Promotor (P_E) des Gens *Sxl* aktiviert. Die hieraus resultierenden frühen Transkripte werden weibchenspezifisch gespleißt. SXL wirkt hierbei selbst als Spleiß-Faktor. Die Genprodukte der frühen Transkripte bewirken, dass auch die Transkripte des späten Promotors (P_M) von *Sxl* weibchenspezifisch gespleißt werden (Estes et al. 1995). Hierbei wird das dritte Exon aus dem Transkript entfernt, welches ein Stopcodon enthält. Über einen „Feedback“-Mechanismus regulieren die Genprodukte der Transkripte des späten Promotors ebenfalls ein weibchenspezifisches Spleißen der späten Transkripte von *Sxl* (Cline 1984).

In männlichen Individuen hingegen wird der frühe Promotor des Gens *Sxl* nicht aktiviert. Nach kurzer Zeit wird der in beiden Geschlechtern genutzte späte Promotor des Gens *Sxl* aktiv. Die hieraus resultierenden Transkripte werden männchenspezifisch gespleißt, da kein *Sxl*-Genprodukt des frühen Promotors vorliegt, wie dies in weiblichen Individuen der Fall ist (Keyes et al. 1992). Hierdurch enthalten männliche *Sxl*-Transkripte das dritte Exon, in welchem sich ein Stopcodon befindet. Die erzeugten SXL-Proteine sind somit verkürzt und nicht funktionsfähig (Bopp et al. 1991; Horabin und Schedl 1993).

SXL bewirkt nicht nur ein geschlechtsspezifisches Spleißen seiner eigenen Transkripte, sondern auch das differentielle Spleißen der Transkripte des in der Kaskade nachgeschalteten Gens *Transformer* (*tra*). SXL besitzt hierfür zwei RNA-Erkennungsmotive (RRM), welche spezifisch an RNA binden können (Bell et al. 1991; Samuels et al. 1998). Zusätzlich werden für das Spleißen der *tra*-Transkripte allerdings noch weitere Gene benötigt. Für das weibchenspezifische Spleißen von *Sxl* und *tra* wurden *sans fille* (*snf*) auch als „*splicing necessary factor*“ bezeichnet (Albrecht und Salz 1993; Flickinger und Salz 1994; Cline et al. 1999), *female lethal (2)d* (*fl(2)d*) (Granadino et al. 1990; Granadino et al. 1996) sowie *virilizer* (*vir*) (Hilfiker und Nöthiger 1991; Hilfiker et al. 1995) als beteiligte Gene nachgewiesen. Bei dem Genprodukt von *snf* handelt es sich um einen Spleiß-Faktor mit Homologien zu den U1A- und U1B-Proteinen aus Säugetieren. Die Funktion der Genprodukte von *fl(2)d* und *vir* sind bisher unbekannt (Schütt und Nöthiger 2000).

Das in der Genkaskade auf *Sxl* nachfolgende Gen *transformer (tra)* wird durch SXL ebenfalls geschlechtsspezifisch differentiell gespleißt (Boggs et al. 1987; Nagoshi et al. 1988; Sosnowski et al. 1989; Inoue et al. 1990). Hierbei wird aus dem weiblichen Transkript das zweite Exon entfernt, welches genau wie das dritte Exon von *Sxl* ein Stopcodon beinhaltet. In männlichen Individuen wird das zweite Exon hingegen nicht herausgespleißt. Dies führt in männlichen Individuen zu einem früheren Translationsstopp und somit zu verkürzten Tra-Genprodukten, welche wie auch die verkürzten SXL-Genprodukte nicht funktionsfähig sind.

Das Gen *doublesex (dsx)* ist in *D. melanogaster* das letztes Gen der geschlechtsdeterminierenden Kaskade. Im Gegensatz zu den Genen *Sxl* und *tra* enthalten die Transkripte von *dsx* in Individuen beider Geschlechter lange ORFs, welche in beiden Geschlechtern zu funktionell aktiven DSX-Proteinen führen (Burtis und Baker 1989). TRA bewirkt zusammen mit den Genprodukten des konstitutiv exprimierten Gens *tra2* ein geschlechtsspezifisches Spleißen sowie eine weibchenspezifische Polyadenylierung des Gens *dsx* (Amrein et al. 1988; Hodgkin 1989; Amrein et al. 1990; Hedley und Maniatis 1991; Hoshijima et al. 1991; Inoue et al. 1992). Tra2 fungiert hierbei als ein positiver Regulator der Prozessierung von *dsx*-Transkripten. In männlichen Individuen erfolgt eine Verknüpfung der Exons 1-3 mit den Exons 5+6. In weiblichen Individuen hingegen werden die Exons 1-3 mit Exon 4 verknüpft. Die aus den unterschiedlich gespleißten Primärtranskripten entstehenden Proteine DSXM und DSXF sind im N-terminalen Bereich identisch, unterscheiden sich jedoch im C-terminalen Bereich. Hierbei ist das DSXM-Protein mit 152 AS im Gegensatz zum DSXF-Protein mit 30 AS im C-terminalen Bereich um 122 AS länger (Burtis und Baker 1989). Bei beiden DSX-Proteinen handelt es sich um Transkriptionsfaktoren. Diese regulieren die Aktivität der nachgeschalteten Differenzierungsgene. Der in DSXM und DSXF identische N-terminale Proteinanteil enthält eine DNA-Bindedomäne, über welche DSX mit den Zielgenen interagieren kann. Der C-terminale Teil hingegen ist für die unterschiedliche Regulation der Differenzierungsgene verantwortlich. DSXM bewirkt hierbei eine Aktivierung von männlichen Differenzierungsgenen und reprimiert weibliche Differenzierungsgene (Burtis und Baker 1989; Jursnich und Burtis 1993). DSXF hingegen wirkt zusammen mit dem Genprodukt des Gens *intersex (ix)* reprimierend auf die männlichen Differenzierungsgene. Das Genprodukt IX fungiert hierbei als ein Co-Faktor bei der DNA-Bindung von DSXF. Eine Interaktion mit DSXM konnte nicht gezeigt werden (Chase und Baker 1995; Garrett-Engele et al. 2002). Eine grafische Übersicht über die Geschlechtsdetermination in *D. melanogaster* ist in Abbildung 1.2 dargestellt.

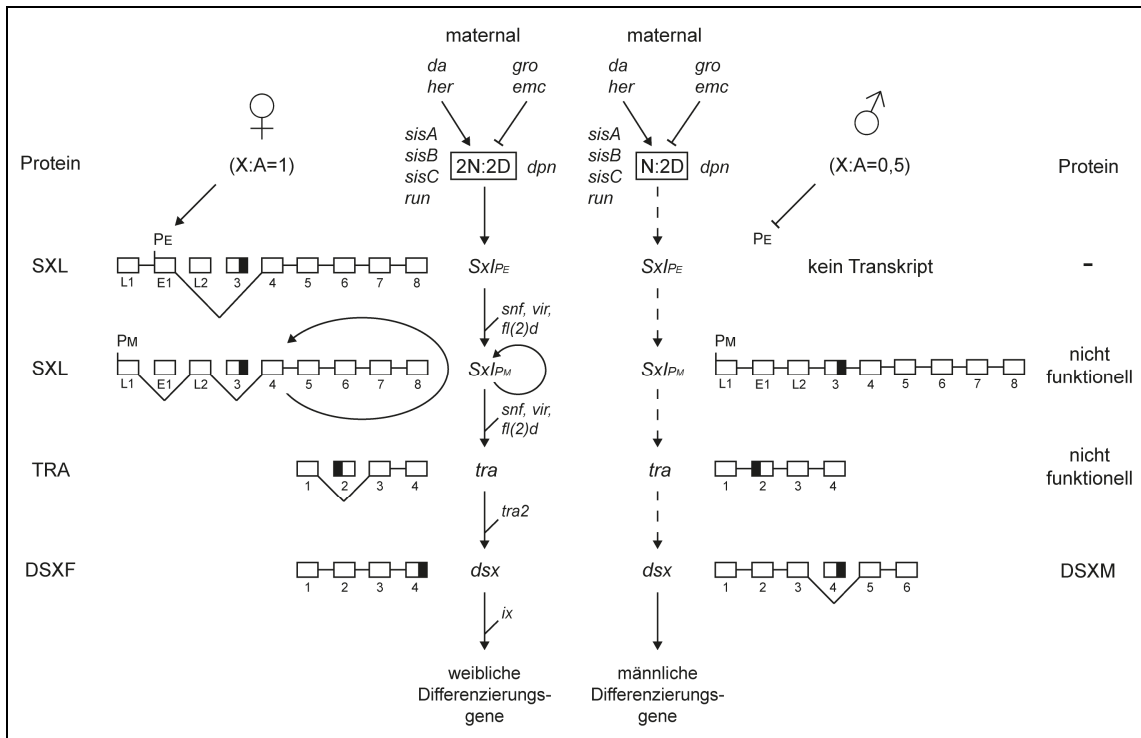


Abbildung 1.2: Mechanismus der Geschlechtsdetermination in *D. melanogaster*

Die Abbildung zeigt schematisch den Ablauf der geschlechtsbestimmenden Genkaskade bei *D. melanogaster*. Hierbei sind in der linken Hälfte die regulatorischen Prozesse im Weibchen, auf der rechten Seite die Prozesse im Männchen dargestellt. Das Verhältnis von X-chromosomalen Numeratoren (N) zu autosomalen Denominatoren (D) stellt hierbei das primäre geschlechtsdeterminierende Signal dar. Die Kästchen der Gene stellen Exons dar. Durch Dreiecke verbundene Exons sind das Ergebnis von differentiellen Spleiß-Prozessen. Bei schwarzen Bereichen innerhalb der Exons handelt es sich um Stopcodons. Werden diese Exons nicht durch Spleißen entfernt, können hieraus wie im ♂ verkürzte Transkripte mit nicht funktionsfähigen Genprodukten resultieren.

1.3.3 Geschlechtsbestimmende Gene bei *Musca domestica* und *Ceratitis capitata*

Die beiden Arten *Musca domestica* und *Ceratitis capitata* gehören innerhalb der Ordnung der Dipteren (Zweiflügler) zu der Unterordnung der Brachycera (Fliegen). In beiden Arten erfolgt die Determination des Geschlechts auf eine weitaus einfachere Art, als dies in *D. melanogaster* (Kapitel 1.3.2) der Fall ist (Dübendorfer et al. 2002).

In den sogenannten Standard-Stämmen von *M. domestica* trägt das Y-Chromosom einen dominanten Faktor *M*, welcher zu männchenspezifischer Differenzierung führt (Perje 1948). Genetisch betrachtet ist die Anwesenheit dieses Faktors *M* auf dem Y-Chromosom hierbei der einzige Unterschied zwischen den beiden Heterosomen (Dübendorfer et al. 2002). X- und Y-Chromosom sind beide überwiegend heterochromatisch. Sie können jedoch auf morphologischer Ebene unterschieden werden, da sich die repetitiven Einheiten der heterochromatischen Bereiche zwischen X- und Y-Chromosom unterscheiden (Hediger et al. 1998). Die euchromatischen

Bereiche sind hingegen auf beiden Heterosomen gleich. Diese hohe Übereinstimmung zwischen den beiden Geschlechtschromosomen scheint eine Dosiskompensation überflüssig zu machen (Dübendorfer et al. 2002).

In weiteren Stämmen von *M. domestica* wurde der dominante Faktor *M* auf allen fünf Autosomen (Hiroyoshi 1964; Wagoner 1969; Rubini et al. 1972) sowie dem X-Chromosom lokalisiert (Rubini und Palenzona 1967; Denholm et al. 1983). In einigen gefundenen Stämmen wird das Geschlecht nicht durch den dominanten Faktor *M*, sondern durch einen weibchenbestimmenden, dominanten Faktor F^D bestimmt. In diesen Stämmen sind beide Geschlechter homozygot für den autosomalen Faktor *M* (Rubini et al. 1972; McDonald et al. 1978). Die weiblichen Individuen dieser Stämme sind heterozygot für den dominanten Faktor F^D . Dies führt auch bei Anwesenheit des Faktors *M* zu einer weiblichen Differenzierung (Dübendorfer et al. 2002). In weiteren Stämmen mit einer „loss-of-function“-Mutation von *F* (F^{man}) kommt es bei homozygoter Ausprägung zu einer männlichen Entwicklung der Individuen, obwohl diese chromosomal gesehen weiblich sind (Schmidt et al. 1997). In beiden Stämmen wurde der Faktor *F* auf Chromosom 4 lokalisiert.

In allen aufgeführten Stämmen von *M. domestica* fungiert der Faktor *F* als Schlüsselgen der Geschlechtsdifferenzierung. Inzwischen konnte gezeigt werden, dass es sich bei dem für die Differenzierung entscheidenden Faktor *F* in *M. domestica* um das Gen *transformer* (*tra*) handelt (Hediger et al. 2010). Eine kontinuierliche Expression von weibchenspezifisch gespleißtem *tra* wird für die weibliche Differenzierung benötigt. Dies wird über eine „Feedback“-Schleife erreicht. Beim weibchenspezifischen Spleißen wird das zweite Exon, welches ein Stopcodon enthält, aus dem *tra*-Transkript entfernt. Am weibchenspezifischen Spleißen von *tra* sowie dem in der Kaskade nachfolgenden *dsx* sind sowohl TRA wie auch TRA2 beteiligt. Hierbei fungiert TRA2 als Co-Faktor von TRA und ist für die weibliche Differenzierung nötig (Burghardt et al. 2005). Die Initiation der „Feedback“-Schleife erfolgt über maternale Genprodukte von *tra* und *tra2*. Der männchenspezifische Faktor *M* wirkt zygotisch reprimierend auf die Ausbildung der „Feedback“-Schleife. Hierdurch kann kein funktionelles TRA gebildet werden. Dies führt im Folgenden zu einem männchenspezifischen Spleißen von *dsx* und somit zu einer männlichen Differenzierung. In Abwesenheit des Faktors *M* wird *F* (*tra*) aktiv und es kommt zur weibchenspezifischen Entwicklung (Dübendorfer et al. 1992).

Das in *D. melanogaster* als Schlüsselgen fungierende *Sex-lethal* (*Sxl*) hingegen wird in *M. domestica* in beiden Geschlechtern gleichermaßen exprimiert und ist offensichtlich nicht an der Geschlechtsdetermination beteiligt (Meise et al. 1998). Somit unterscheiden sich die Genkaskaden der Geschlechtsbestimmung von *D. melanogaster* und *M. domestica* in ihren jeweiligen Schlüsselgenen (Hediger et al. 2004). Betrachtet man hingegen den terminalen Regulator der Genkaskade *doublesex* (*dsx*), so ist festzustellen, dass es sich hierbei um das höchst-konservierte Gen der Geschlechtsdifferenzierung innerhalb der Dipteren handelt (Permpoon und Thanaphum 2010). Dies unterstützt die Hypothese, dass sich die Kaskade der geschlechtsdeterminierenden Gene „vom Grund zur Spitze“ entwickelt hat (Wilkins

1995). Eine grafische Übersicht über die Geschlechtsdetermination in *M. domestica* ist in Abbildung 1.3 gegeben.

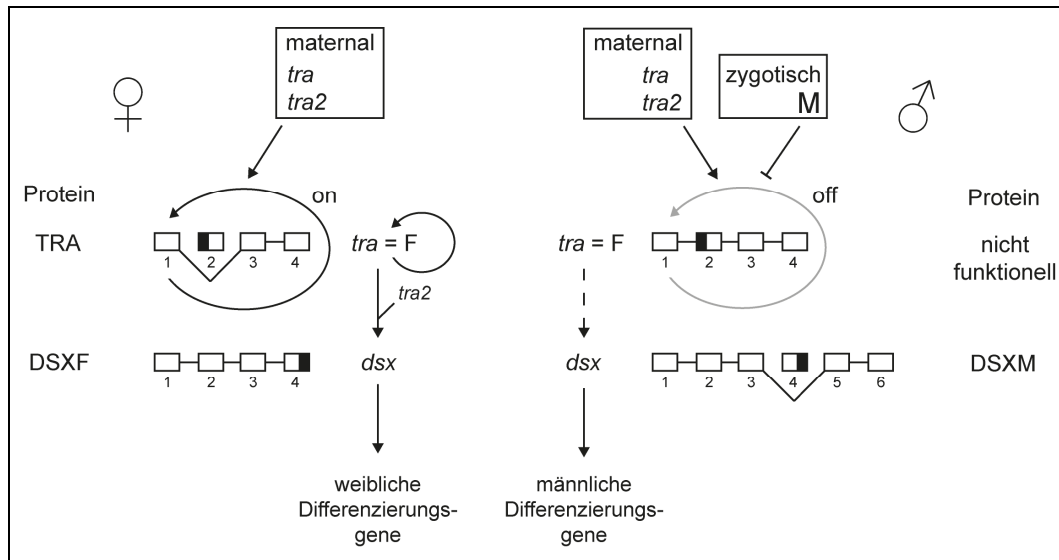


Abbildung 1.3: Mechanismus der Geschlechtsdetermination bei *M. domestica*

Die Abbildung zeigt schematisch den Ablauf der geschlechtsbestimmenden Genkaskade bei *M. domestica*. Hierbei sind in der linken Hälfte die regulatorischen Prozesse im Weibchen, auf der rechten Seite die Prozesse im Männchen dargestellt. Das Schlüsselgen dieser Kaskade ist *transformer* (*tra*). Hierbei handelt es sich um den viel beschriebenen Faktor *F*. Die Kästchen der Gene stellen Exons dar. Durch Dreiecke verbundene Exons sind das Ergebnis von Spleiß-Prozessen. Bei schwarzen Bereichen innerhalb der Exons handelt es sich um Stopcodons. Werden diese Exons nicht durch Spleißen entfernt, können hieraus wie im ♂ bei *tra* verkürzte Transkripte mit nicht funktionsfähigen Genprodukten resultieren.

Auch in *Ceratitis capitata* konnte *transformer* (*tra*) als Schlüsselgen der geschlechtsdeterminierenden Genkaskade nachgewiesen werden (Pane et al. 2002). Das Gen *tra* weicht hierbei in seiner Sequenz stark von dem homologen Gen aus *D. melanogaster* ab. Die Funktion des Genprodukts ist jedoch so stark konserviert, dass es sogar das Genprodukt in *D. melanogaster* ersetzen kann (Pane et al. 2005). Die Genprodukte von *tra* und *tra2* sind einerseits für die Aufrechterhaltung der „Feedback“-Schleife notwendig, welche zum weibchenspezifischen Spleißen von *tra* führt (Salvemini et al. 2009). Außerdem kommt es durch eine Interaktion von TRA und TRA2 zu einem weibchenspezifischen Spleißen des in der Kaskade nachfolgenden Gens *dsx* (Saccone et al. 2000; Saccone et al. 2002). Durch die Anwesenheit eines dominanten männchenbestimmenden Faktors *M*, welcher auf dem Y-Chromosom lokalisiert ist, kommt es zu einer männchenspezifischen Differenzierung (Willhoef und Franz 1996). In Anwesenheit von *M* wird die Ausbildung der autoregulatorischen Schleife des Gens *tra* verhindert. Im Folgenden wird *tra* nicht weibchenspezifisch gespleißt. Es kommt zu einem männchenspezifischen Transkript mit verkürztem Leserahmen. Das hieraus resultierende Genprodukt ist funktionell inaktiv. Das in der Genkaskade nachfolgende Gen *dsx* wird männchenspezifisch gespleißt, und es kommt zur männlichen Differenzierung.

Das in *D. melanogaster* als Schlüsselgen fungierende *Sxl* weist in *C. capitata* keine geschlechtsspezifischen Transkripte auf und ist somit vermutlich nicht an der Determination des Geschlechts beteiligt (Saccone et al. 1998).

1.4 Evolution der Geschlechtschromosomen

Ein Vergleich der Geschlechtschromosomen von *Musca domestica* und *Homo sapiens* zeigt zwei unterschiedliche Stadien in der Evolution der Entstehung differenzierter Heterosomen auf. Auf genetischer Ebene lassen sich die beiden Geschlechtschromosomen von Standardstämmen von *M. domestica* durch den dominanten Faktor *M* unterscheiden (Dübendorfer et al. 2002). Ausserdem lassen sich die beiden Chromosomen mit Hilfe morphologischer Analysen anhand ihrer unterschiedlich stark ausgeprägten heterochromatischen Bereiche unterscheiden (Hediger et al. 1998). In *H. sapiens* hingegen ist die Differenzierung der Geschlechtschromosomen voll ausgeprägt. Hier lassen sich die beiden Heterosomen bereits bei kurzer Betrachtung anhand ihrer Größe unterscheiden.

Die Ausprägung der männlichen bzw. weiblichen Heterogametie erfolgte in unterschiedlichen Taxa mehrfach und getrennt voneinander. Hierbei kommt es häufig zum Auftreten morphologisch differenzierter Geschlechtschromosomen.

Nach der Hypothese von Muller (1918) entwickeln sich Heterosomen aus zwei ursprünglich homologen Autosomen. Nach diesem Modell entsteht hierbei zuerst ein kleiner, für die Geschlechtsdetermination wichtiger Bereich auf den Autosomen. Dieser Bereich trägt zwei Allele eines Gens, welches für die Determination des Geschlechts ausschlaggebend ist. Hierdurch entstehen im Fall männlicher Heterogametie die sogenannten Proto-X- und Proto-Y-Chromosomen. Der Grund für eine unterdrückte Rekombination liegt demnach in der Erhaltung der männchen- oder weibchendeterminierenden Funktion. Die Hypothese von Muller wurde durch die Beobachtungen von unterschiedlichen Differenzierungsstadien der Geschlechtschromosomen in nah verwandten Arten eines Taxons unterstützt. Die Stadien reichen hierbei von undifferenzierten bis zu stark differenzierten Heterosomen (Ohno 1967; Singh et al. 1980; Galetti et al. 1981). Auch die schnelle Degeneration ehemals autosomaler Gene nach Translokation eines Autosoms auf das Y-Chromosom in *D. miranda* (Steinemann 1982; Steinemann und Steinemann 1990; 1992; 1993) unterstützt die Hypothese von Muller. Inzwischen erklären mehrere Modelle die Entstehung differenzierter Heterosomen als Folge dauerhafter Heterozygotie des geschlechtsdeterminierenden Gens (Muller 1964; Felsenstein 1974; Charlesworth 1978; Lucchesi 1978).

Die Heterosomen lassen sich hierbei durch unterdrückte Rekombination, eine Degeneration und Heterochromatinisierung des Y- oder W-Chromosoms, die Anreicherung von repetitiver DNA oder transponierbarer Elemente auf dem Y- oder W-Chromosom sowie eine meist nötige Dosiskompensation voneinander unterscheiden.

Die Unterdrückung der Rekombination zwischen den beiden Geschlechts-Chromosomen ist vermutlich der erste Schritt in der Entstehung von differenzierten Heterosomen. Hierbei kann die Unterdrückung der Rekombination entweder das gesamte Geschlechtschromosom betreffen oder nur einen Teilbereich. So kann in *H. sapiens* Rekombination zwischen dem X- und Y-Chromosom innerhalb der pseudo-autosomalen Regionen (PAR) des Y-Chromosoms erfolgen (Ellis und Goodfellow 1989). Die für die Determination des Geschlechts relevante Region bleibt hingegen immer von der Rekombination ausgeschlossen. Eine Unterdrückung der Rekombination lässt sich sowohl genetisch wie auch indirekt cytologisch nachweisen. Genetisch ist sie durch eine strikte X-gekoppelte Vererbung nachweisbar. Beim Menschen existieren bekannte X-gekoppelte Erbkrankheiten. So werden z.B. die beiden Varianten A und B der Bluterkrankheit (Hämophilie) X-chromosomal vererbt. Auch bei der Rot-Grün-Sehschwäche oder -Blindheit (Dyschromatopsie) handelt es sich um eine X-chromosomal vererbte Krankheit. Cytologisch liefert das Fehlen von Chiasmata zwischen den gesamten Heterosomen oder innerhalb des für die Geschlechts-determination relevanten Bereichs in der Meiose einen indirekten Nachweis für unterdrückte Rekombination (Bull 1983).

Die Anzahl funktionsfähiger Gene differiert zwischen den beiden Heterosomen stark. Während beim Menschen zahlreiche X-chromosomale Gene identifiziert werden konnten (NCBI: 1672 Gene; Stand April 2012), ist die Anzahl der Y-chromosomalen Gene wesentlich geringer. Zwar konnten inzwischen auch hier mit Hilfe der Hochdurchsatz-Sequenzierung 429 Gene (NCBI; Stand April 2012) identifiziert werden. Hierbei handelt es sich jedoch in vielen Fällen um Pseudogene sowie um nicht-proteinkodierende Transkripte. Als funktionsfähige Gene konnten neben dem für die Determination des Geschlechts relevanten Gen *SRY* (Sinclair et al. 1990) noch weitere Gene wie z.B. das *GBY* (gonado blastoma locus on the Y chromosome) und das *YRRM* (Y chromosome RNA recognition motif) (Tsuchiya et al. 1995), das *TSPY* (human Y-chromosome-derived cDNA) (Arnemann et al. 1991) sowie das *ZFY* (Zink-Finger-Y) (Page et al. 1987) gefunden werden. In *D. melanogaster* wurde die Zahl der X-chromosomalen Gene auf Grund von Untersuchungen der Letalität nach Bestrahlung auf 800-1000 Gene geschätzt (Lefevre 1974). Neueste Angaben geben für das X-Chromosom von *D. melanogaster* 2274 Gene an (NCBI; Stand April 2012). Für das Y-Chromosom sind bisher nur 12 Gene beschrieben (Koerich et al. 2008). Schätzungen gehen von insgesamt etwa 20 Genen auf dem Y-Chromosom von *D. melanogaster* aus.

Die weitaus geringere Anzahl an funktionsfähigen Genen auf dem Y-Chromosom sowie die gleichzeitig weitaus höhere Anzahl an Pseudogenen lassen sich mit Hilfe der unterdrückten Rekombination erklären. Hierdurch kommt es auf dem Proto-Y-Chromosom zu einer Anreicherung von Mutationen. Gene, welche einen Vorteil für die Selektion bieten, können zu dem eigentlichen geschlechtsdeterminierenden Gen hinzustoßen. Hierdurch vergrößert sich der Bereich des Proto-Y-Chromosoms, auf dem die Rekombination unterdrückt wird.

Einhergehend mit der Anreicherung von Mutationen in Folge der unterdrückten Rekombination kommt es außerdem zu weiteren weitreichenden Veränderungen auf

dem Proto-Y-Chromosom. Hierzu gehören Deletionen, Duplikationen sowie die Anreicherung mit transposablen Elementen. Dies führt zu einer Degeneration und Heterochromatinisierung des Proto-Y-Chromosoms, bis letztlich zu einer reduzierten Größe des differenzierten Y-Chromosoms (Clark und Wall 1996; Ayling und Griffin 2002). In *D. miranda* (Dobzhansky 1935) konnte der Prozess der Degeneration des Proto-Y-Chromosoms durch die Translokation eines Autosoms auf ein Y-Chromosom deutlich gemacht werden. Durch die erfolgte Translokation entstand ein neo-Y-Chromosom (Macknight 1939). Ein Vergleich einiger Gene des neo-Y-Chromosoms mit dem Teil des homologen Autosoms konnte zeigen, dass sich die ehemals aktiven Gene des Autosoms zu degenerierten Pseudogenen entwickelt hatten (Steinemann 1982). Weiterhin konnten auf dem neo-Y-Chromosom Deletionen, Duplikationen und transposable Elemente nachgewiesen werden (Steinemann und Steinemann 1990; 1992; 1993; Steinemann et al. 1993; Lucchesi 1994; Steinemann und Steinemann 1998). Dies spricht für eine zentrale Rolle von transposablen Elementen in der Entstehung von Heterosomen.

Ein weiteres Merkmal differenzierter Y-Chromosomen ist ein hoher Gehalt an repetitiver DNA. Bei einem großen Anteil der repetitiven DNA handelt es sich um Satelliten-DNA. Diese ist auf dem X-Chromosom in der Regel nur im Zentromerbereich zu finden (Kunkel und Smith 1982; Steinemann 1982). Hierbei handelt es sich häufig um sogenannte „simple quarternary repeats“. Diese kommen z.B. in Säugern, Fischen und Fliegen vor (Eppelen et al. 1982; Levinson et al. 1985; Miklos et al. 1989). In *D. melanogaster* bestehen etwa 80 % des Y-Chromosoms aus solchen kurzen repetitiven Einheiten (Bonaccorsi und Lohe 1991).

Eine morphologische Unterscheidung noch nicht ausdifferenzierter Heterosomen kann mit Hilfe von Färbetechniken erfolgen. So unterscheiden sich viele Chromosomen evolutiv früher Differenzierungsstadien in unterschiedlich stark ausgeprägten heterochromatischen Bereichen. Diese lassen sich z.B. mit Hilfe der C-Bandierungstechnik spezifisch anfärben. In *M. domestica* lassen sich so die beiden nicht ausdifferenzierten Heterosomen voneinander unterscheiden (Hediger et al. 1998). In anderen Arten wie z.B. *C. thummi* stellen spezifische, heterochromatische Banden im heterogametischen Geschlecht das einzige Unterscheidungsmerkmal zwischen den noch nicht ausdifferenzierten Heterosomen dar (siehe Kapitel 1.5).

Durch die Degeneration des Y-Chromosoms und das damit verbundene Fehlen vieler X-chromosomaler Gene im männlichen Y-Chromosom kommt es zu einer Veränderung der Gendosis zwischen dem heterogametischen und dem homogametischen Geschlecht. So besitzt das homogametische Geschlecht die doppelte Gendosis X-chromosomaler Gene wie das heterogametische Geschlecht. Um die unterschiedliche Gendosis zwischen den beiden Geschlechtern auszugleichen, hat sich der Mechanismus der Dosiskompensation entwickelt. Die Dosiskompensation erfolgt hierbei in verschiedenen Spezies auf unterschiedliche Weise. In *H. sapiens* kommt es zur Inaktivierung eines der beiden X-Chromosomen, so dass in beiden Geschlechtern nur ein Gen aktiv ist (Lyon 1961). Das heterochromatische, inaktivierte X-Chromosom ist in cytologischen Präparaten als Barr-Körperchen zu erkennen. In *D. melanogaster* erfolgt dagegen die Dosiskompensation nicht durch

Inaktivierung eines X-Chromosoms des Weibchens, sondern durch eine Überexpression X-chromosomaler Gene in männlichen Individuen (Kuroda et al. 1991). In *C. elegans* wiederum erfolgt die Dosiskompensation durch Herunterregulierung der Genexpression auf beiden X-Chromosomen bei Hermaphroditen (Kelly et al. 2002).

1.5 Geschlechtsdetermination in *Chironomus*

Die meisten Arten von *Chironomus* weisen keine heteromorphen Geschlechts-Chromosomen auf (Bauer und Beermann 1952; Beermann 1955). Das Prinzip männlicher Heterogamie konnte jedoch als Mechanismus der Geschlechtsdifferenzierung nachgewiesen werden (Beermann 1955). Die Entscheidung über die Differenzierung des Geschlechts wird hierbei durch den dominanten Faktor *M* getroffen. Ist dieser vorhanden, reprimiert er vermutlich den Faktor *F*, und es kommt zur männchenspezifischen Differenzierung (Bauer und Beermann 1952; Beermann 1955; Keyl 1961; Keyl 1962). Das Chromosom, welches den dominanten Männchen-Bestimmer *M* trägt, wird als Proto-Y-Chromosom bezeichnet. Dieses ist genetisch gesehen das Y-Chromosom (Beermann 1955), lässt sich aber morphologisch meist nicht vom Proto-X-Chromosom unterscheiden.

Durch Untersuchungen männchenspezifischer Inversionen in den Arten *C. annularius*, *C. tentans* und *C. pallidivittatus* konnte die Lage des Faktors *M* in diesen Arten ermittelt werden (Beermann 1955). In *C. annularius* konnte der eng mit der männchenspezifischen Inversion gekoppelte Faktor *M* in einer Region nahe dem Telomer des Chromosomenarms F (Chromosom 3L) lokalisiert werden. In den beiden Arten der Camptochironomiden *C. tentans* und *C. pallidivittatus* konnten zwei unterschiedliche geschlechtsbestimmende Loci identifiziert werden. Hierbei ist der geschlechtsdeterminierende Faktor *M* entweder nahe dem Telomer auf Arm F (Chromosom 1R) oder nahe dem Telomer auf Arm B (Chromosom 2L) lokalisierbar (Beermann 1955). Auch in den Arten *C. nuditarsis* (Rosin und Fischer 1972), *C. obtusidens* (Keyl 1961; Keyl 1962) und *C. plumosus* (Acton 1957; Keyl 1962) wurden männchenspezifische Inversionen beschrieben. In den meisten Arten von *Chironomus* kommen hingegen keine männchenspezifischen Inversionen vor. In einigen Arten wie *C. australis* (Martin 1980), *C. obtusidens* (Keyl 1961; Keyl 1962; Martin 1980), einigen Stämmen von *C. nuditarsis* (Fischer und Tichy 1980) sowie in *C. thummi* (Hägele 1985) konnte eine hemizygot, heterochromatische Bande identifiziert werden. Da diese nur in Männchen zu finden ist, muss sie eng mit dem geschlechtsdeterminierenden Faktor gekoppelt sein. Innerhalb der untersuchten *Chironomus*-Arten konnte die Lage des Faktors *M* auf allen Chromosomenarmen außer auf Chromosomenarm E nachgewiesen werden (Martin 1980; Martin 2010).

Um in den beiden Schwester-Arten *C. thummi* und *C. piger* (Keyl und Strenzke 1956), welche keine geschlechtsgekoppelte Inversion zeigen, die Lage des Faktors *M* zu bestimmen, führte Hägele (1985; 1986) Rückkreuzungen durch. Hierbei wurden FI-Hybrid-Männchen der Kreuzung aus *C. thummi* und *C. piger* mit den Weibchen der jeweiligen Elternart verpaart. Im Folgenden wurde das Paarungsverhalten der

Polytänchromosomen untersucht. Hierbei zeigen die Nachkommen der Rückkreuzungen geschlechtsspezifische Unterschiede. Die Unterschiede in der Paarung treten im Bereich D₃d-f auf Arm F (Chromosom 3) auf. Sie äußern sich in einer Asynapsis der Chromosomen der Hybriden in dieser Region bei männlichen Tieren. Die chromosomale Lage der Region (Arm F; nahe dem Telomer) stimmt hierbei mit einer der beiden bekannten Lagen (Arm F; nahe dem Telomer/ Arm B; nahe dem Telomer) der geschlechtsdeterminierenden Region der schwedischen Population von *C. tentans* überein (Beermann 1955; Hägele 1985). Dies deutet darauf hin, dass in beiden Arten dasselbe Gen für die Determination des Geschlechts verantwortlich sein könnte.

Durch Anfärbung der Polytänchromosomen von *C. thummi* mit Hilfe der RB- (Rot-Blau) und Quinacrin-Färbemethoden konnte eine hemizygoten männchenspezifische Bande in Region D₃e₁ lokalisiert werden. Mit Hilfe der RB-Färbung konnte gezeigt werden, dass es sich bei dieser hemizygoten Bande um einen heterochromatischen Bereich handelt. Weiterhin konnte durch die Anfärbung mittels Quinacrin nachgewiesen werden, dass diese Bande einen hohen Anteil AT-reicher Sequenzen enthalten muss (Hägele 1985). In der Schwesternart *C. piger* ist in diesem Bereich weder in Weibchen noch in Männchen eine hemizygoten Bande zu finden. Da sich der Faktor *M* in den beiden Spezies in homologen Loci befindet (Hägele 1985; Hägele 1986), kann es sich bei der hemizygoten, heterochromatischen Bande per se nicht um den dominanten Faktor *M* handeln.

Zwischen den beiden Arten *C. thummi* und *C. piger* gibt es generelle Unterschiede in der Menge und der Verteilung der repetitiven DNA sowie der heterochromatischen Bereiche. So konnte mittels cytologischer Untersuchungen in *C. thummi* ein etwa 30 % größeres Genom als in der Schwesternart *C. piger* nachgewiesen werden (Keyl 1965). Etwa 20 % des höheren DNA-Gehalts sind auf tandem-repetitive *Cla*-Elemente zurückzuführen. Diese stellen eine Familie hochrepetitiver DNA-Sequenzen mit einer Monomerlänge von etwa 120 bp dar. Sie weisen eine Restriktionsschnittstelle für das Restriktionsenzym *Cla*I auf, wodurch sie ihre Bezeichnung erhielten (Schmidt 1981). *Cla*-Elemente besitzen mehrere palindromische Sequenzen und einen AT-Gehalt von mehr als 80 % (Schmidt 1984). In *C. thummi* sind *Cla*-Elemente in der Zentromerregion sowie an mehr als 200 weiteren Loci auf allen Chromosomen über das Genom verteilt zu finden. Innerhalb dieser Loci liegen die *Cla*-Elemente als tandem-repetitive Cluster vor. Die Gesamtzahl der *Cla*-Elemente beläuft sich hierbei auf etwa 70000 Kopien (Schmidt 1984; Hankeln und Schmidt 1987; Hankeln et al. 1994). Selbst in den nicht-transkribierten Spacern (NTS) der rDNA-Gene von *C. thummi* sind *Cla*-Elemente zu finden (Israelewski und Schmidt 1982; Schmidt et al. 1982). In *C. piger* ist das Vorkommen der *Cla*-Elemente hingegen hauptsächlich auf die Zentromerregion beschränkt. Im Vergleich zu *C. thummi* ist ihre Kopienzahl in *C. piger* etwa um den Faktor 5,5 verringert (Schaefer und Schmidt 1981; Schmidt 1981; Schmidt und Godwin 1983). Durch das erhöhte Vorkommen von *Cla*-Elementen sowie das damit verbundene größere Genom unterscheidet sich *C. thummi* klar von der Schwesternart *C. piger*. Da sich *C. thummi* in seiner Genomgröße außerdem auch von anderen

Chironomus-Arten unterscheidet, handelt es sich bei *C. thummi* vermutlich um die phylogenetisch abgeleitete Art (Keyl 1962; Keyl 1965).

Durch „in situ“-Hybridisierungen mit *Cla*-Elementen an Polytänychromosomen von *C. thummi* konnte ein männchenspezifisches hemizygoten Hybridisierungssignal im Bereich D_{3e_1} auf Arm F (Chromosom 3) lokalisiert werden (Kraemer und Schmidt 1993). Die Lage entspricht hierbei dem Bereich der von Hägele (1985) beschriebenen hemizygoten, heterochromatischen Bande. In einigen Fällen taucht zusätzlich zu dem starken, hemizygoten Hybridisierungssignal in Region D_{3e_1} ein weiteres männchenspezifisches Hybridisierungssignal auf. Dieses ist schwächer und im Bereich D_{3f} des Arms F (Chromosom 3) lokalisierbar. In *C. thummi* ♀ sind in dieser Region keine Hybridisierungssignale zu finden. Somit muss das hemizygoten *Cla*-Element-Cluster ebenfalls eng mit dem dominanten Faktor *M* gekoppelt sein. In der Schwesternart *C. piger* sind weder in Männchen noch in Weibchen Hybridisierungssignale der *Cla*-Elemente in dieser Region zu verzeichnen (Kraemer und Schmidt 1993). Da sich der geschlechtsdeterminierende Faktor *M* in den beiden Spezies in homologen Loci befindet (Hägele 1985; Hägele 1986), ist auch eine direkte Beteiligung der *Cla*-Elemente an der Determination des Geschlechts unwahrscheinlich.

Die Diskrepanz im Auftreten der hemizygoten, heterochromatischen Bande sowie des hemizygoten *Cla*-Element Clusters zwischen den beiden Schwesternarten *C. thummi* und *C. piger* bedeutet, dass es innerhalb des kurzen Zeitraums der evolutiven Trennung der beiden Arten zu einer Akkumulation repetitiver Sequenzen wie den *Cla*-Elementen sowie zu einer Heterochromatinisierung innerhalb der geschlechtsbestimmenden Region gekommen sein muss. Dies muss als Folge der unterdrückten Rekombination innerhalb dieser Region angesehen werden (Kraemer und Schmidt 1993). Somit bietet das Proto-Y-Chromosom bei *C. thummi* die Möglichkeit, ein frühes Stadium im evolutiven Prozess der Entstehung von Geschlechtschromosomen zu analysieren (Kraemer 1994).

Mit Hilfe von *Cla*-Elementen als Hybridisierungssonde konnten innerhalb einer λ -Phagen-Bibliothek mit DNA aus *C. thummi* ♂ zwei Klone gefunden werden, welche das hemizygoten *Cla*-Element Cluster in Region D_{3e_1} zu beiden Seiten hin flankieren. Der Klon λ -*Cla*1.8Z liegt hierbei in proximaler Lage zum hemizygoten *Cla*-Element Cluster. Der Klon λ -*Cla*1.1Y konnte hingegen auf der distalen Seite des Clusters lokalisiert werden. Ausgehend von den beiden Klonen λ -*Cla*1.8Z und λ -*Cla*1.1Y wurden „Chromosomenwalks“ mit Hilfe genomischer λ - und BAC-Bibliotheken gestartet. Hierdurch konnte der Bereich zwischen den Positionen beider Klone vollständig kloniert werden. Insgesamt wurden bisher etwa 200 kb kloniert und sequenziert. Die meisten Sequenzinformationen stammen hierbei aus den Analysen bereits analysierter BAC-Klone (Kraemer 1994; Kunz 2001; Knies 2006; Berwanger 2007; Funke 2011). Eine Übersicht über die Lage der BAC-Klone innerhalb des sequenzierten Bereichs ist in Abbildung 1.4 gezeigt. Innerhalb dieser Region konnten bislang 20 Gene identifiziert werden. Da die analysierte Region höchstwahrscheinlich den männchenbestimmenden Faktor *M* enthält und sich die gesamten Sequenzinformationen aus verschiedenen Analysen zusammensetzen, wird der untersuchte Bereich im Folgenden als Contig SDR (sex determining region) oder geschlechtsbestimmende Region bezeichnet.

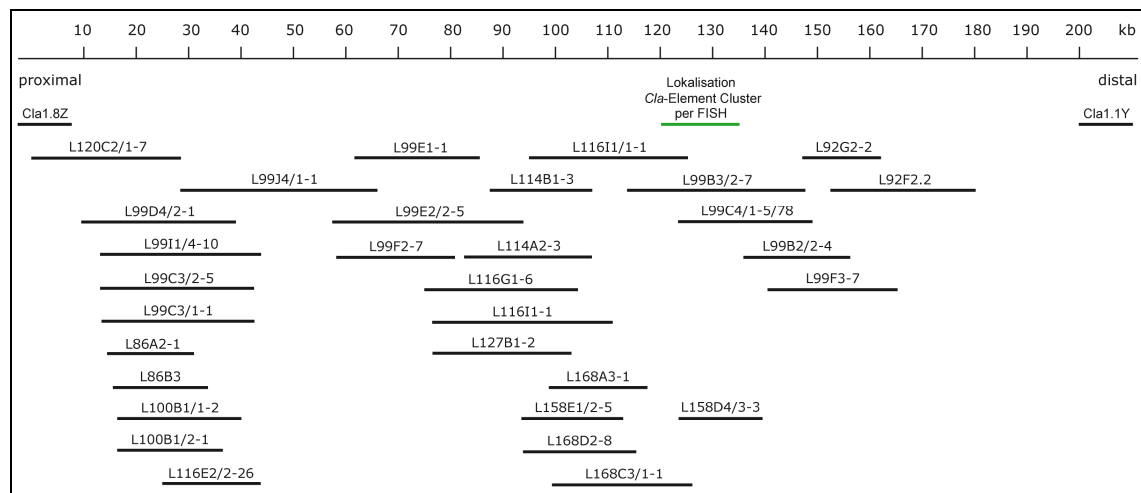


Abbildung 1.4: Lage der BAC-Klone aus „Chromosomenwalk“ in *C. thummi*

Gezeigt ist eine Übersicht über die in *C. thummi* existierenden BAC-Klone zwischen den beiden Lambda-Klonen λ -Cla1.8Z und λ -Cla1.1Y. Weiterhin dargestellt ist die per FISH analysierte Position des hemizygoten *Cla*-Element Clusters (grüner Balken) innerhalb des sequenzierten Abschnitts.

Bislang konnte nicht geklärt werden, ob es sich bei den vorliegenden Sequenzen des Contigs SDR um X- oder Y-chromosomale Abschnitte handelt. Da allerdings das in dieser Region in *C. thummi* ♂ per FISH lokalisierbare hemizygoten *Cla*-Element Cluster bisher nicht in der Sequenz zu finden ist, scheint es sich bei den vorliegenden Sequenzen des Contigs eher um X-chromosomale Sequenzen zu handeln. Hierfür spricht auch ein kurzer Sequenzabschnitt im Bereich des Gens *dLacDH*-like, welcher vermutlich das proximale Ende des *Cla*-Element Clusters darstellt. Dieser Bereich konnte per PCR mit einer Kombination aus Primern der geschlechtsbestimmenden Region sowie *Cla*-spezifischen Primern lokalisiert werden (Stiehl, pers. Mitteilung). Weiterhin konnte mit Hilfe von genomischen Southern-Analysen sowie des „Genome Walk“-Verfahrens eine männchenspezifische Duplikation im Bereich der Gene *wd74*-like, *fs(1)K10*-like sowie *dLacDH*-like nachgewiesen werden (Stiehl 2009, Kraemer und Stiehl pers. Mitteilung). Hierbei liegt im Fall des Gens *fs(1)K10*-like eine Duplikation des gesamten Genbereichs vor.

1.6 Zielsetzung der Arbeit

Im Rahmen dieser Arbeit soll der Bereich der geschlechtsbestimmenden Region aus *C. tentans* kloniert, sequenziert und analysiert werden. Bei der Camptochironomiden-Art *C. tentans* liegt innerhalb der SDR kein *Cla*-Element Cluster vor, welches die Klonierung bestimmter Sequenz-Bereiche erschwert. Die gefundenen BAC-Klone der SDR sollen möglichst vollständig sequenziert werden. Die so gewonnenen Sequenzinformationen sollen zu einer zusammenhängenden Sequenz des „Contigs SDR“ von *C. tentans* führen. Dieser sollte auf bereits aus *C. thummi* bekannte Gene der SDR sowie repetitive Elemente hin untersucht werden. Hierbei sollte insbesondere darauf geachtet werden, ob in der SDR der beiden Arten dieselben Gene zu finden sind und ob diese in beiden Arten in identischer Reihenfolge vorliegen. Durch Vergleiche von genomischen Southern-Hybridisierungen von *C. tentans* soll untersucht werden, ob der mit Hilfe der Sequenzen der BAC-Klone erstellte Contig SDR mit der genomischen Organisation der SDR in *C. tentans* übereinstimmt.

Da im Rahmen von vorangegangenen Arbeiten zwar inzwischen viele Gene innerhalb der SDR von *Chironomus* identifiziert werden konnten, der dominante männchenbestimmende Faktor *M* jedoch noch nicht identifiziert werden konnte, sollten die Gene sowie Genprodukte der SDR im Rahmen dieser Arbeit näher untersucht werden. Hierfür sollten mit Hilfe der RT-PCR sowie des RACE-Verfahrens möglichst Vollängen-Transkripte der Gene der SDR erhalten werden. Hierbei sollte vor allem ein Vergleich zwischen larvalen Transkripten männlicher und weiblicher L4-Larven aus *C. thummi* erfolgen. Unterschiede im Spleißmuster zwischen den beiden Geschlechtern können hierbei einen Hinweis auf eine mögliche differentielle Struktur, Regulation oder unterschiedliche Stabilität der Transkripte geben.

Weiterhin sollten die Genprodukte auf das Vorkommen von konservierten Domänen hin untersucht werden. Da es sich bei dem Faktor *M* mit aller Wahrscheinlichkeit um einen Transkriptionsfaktor handelt, ist hierbei von besonderem Interesse, ob sich auch Domänen identifizieren lassen, welche in anderen Transkriptionsfaktoren oder sogar homologen Proteinen der Geschlechtsdifferenzierung anderer Arten zu finden sind.

2 Material und Methoden

2.1 Zellkultur

Die epitheliale Zellkultur von *Chironomus tentans* (Wyss 1982) des Instituts für Molekulargenetik, gentechnologische Sicherheitsforschung und Beratung (IMSB) der Johannes Gutenberg Universität Mainz wurde ursprünglich von Prof. Dr. M. Spindler-Barth (Universität Ulm) zur Verfügung gestellt.

Die *Chironomus tentans*-Zellen wurden zu Beginn der Arbeit in „culture medium ZO buffered with PIPES“ (Wyss 1982) der Fa. Cell culture technologies (Lausanne, Schweiz), später in Promo Cell Schneider´s Drosophila Medium der Fa. Promocell (Heidelberg, Deutschland) unter Zugabe von L-Glutamin (Endkonzentration 13 mM) kultiviert. Pro 500 ml wurden den Medien je 50 ml FCS, 5 ml einer Pepton-Stammlösung (10 %), 5 ml Penicillin/Streptomycin (100 x konzentriert) sowie 50 µl Insulin zugegeben. Zur besseren Überwachung des pH-Wertes wurden 500 ml Medium mit 1 ml 0,5 % Phenolrotlösung (Farbindikator) versetzt.

Das Medium der Fa. Cell culture technologies wurde zusätzlich zu den Herstellerangaben mit PIPES (Endkonzentration 50 mM) gepuffert, um einen konstanten pH-Wert unter 7,0 zu gewährleisten. Überschreitet das Medium diesen pH-Wert, so ist das Zellwachstum gehemmt.

Die Kultivierung der Zellen erfolgte bei 25°C in 200 ml T-Flaschen der Fa. Greiner (Frickhausen, Deutschland). Wöchentlich wurden die Zellen 1:4 aufgeteilt und mit frischem Nährmedium versetzt.

2.2 Versuchstiere

Die verwendeten Versuchstiere der Spezies *Chironomus thummi* sowie *Chironomus piger* stammen aus der permanenten Laborzucht des Instituts für Molekulargenetik, gentechnologische Sicherheitsforschung und Beratung (IMSB) der Johannes Gutenberg Universität Mainz. In den vergangenen Jahren wurden neue Freilandgelege in Zucht gebracht, um somit den „Genpool“ zu erweitern. Die Larven wurden für die Versuche entweder bis zum 4. Larvalstadium (L4) oder bis zum Imaginalstadium aufgezogen.

2.3 Bestimmung des Geschlechts von Chironomiden

Zur Analyse geschlechtsspezifischer Unterschiede wurden sowohl Larven wie auch Adulttiere nach ihrem Geschlecht getrennt aufgearbeitet und die anschließenden Analysen separiert durchgeführt. Im Laufe der Arbeit wurden sowohl Larven wie auch Adulttiere der beiden Spezies *C. thummi* und *C. piger* aufgearbeitet.

Die Unterscheidung des Geschlechts im L4-Larvalstadium konnte nur anhand der Morphologie der Genital-Imaginalscheiben (Wülker 1968) durchgeführt werden. Hierzu

mussten die Segmente 8+9 der L4-Larven betrachtet werden. Weibliche Tiere besitzen in beiden Segmenten weibchenspezifische Strukturen, wohingegen männliche Tiere lediglich im 9. Segment männchenspezifische Strukturen aufzeigen. Die Larven wurden für dieses Verfahren ohne Nahrung für 1 Tag in abgestandenes Wasser umgesetzt. Dies diente der Entleerung des Darms, wodurch eine bessere Betrachtung der Imaginalscheiben ermöglicht wurde. Zur Betrachtung der Ventralseite der Larven mussten diese in gestreckter Position (Nachschieber am Abdomen gespreizt) auf der Dorsalseite liegend fixiert werden. Die Betrachtung der Larven erfolgte unter dem Mikroskop „Axiostar plus“ der Fa. Zeiss (Göttingen, Deutschland). Larven, bei denen das Geschlecht eindeutig bestimmt werden konnte, wurden zuerst in einem Reaktionsgefäß auf Kohlendäureschnee schockgefrostet und später bei -80°C gelagert.

Die Unterscheidung des Geschlechts von Adulttieren erfolgte unter einem „MS5-Binokular“ der Fa. Leica (Solms, Deutschland). Hierzu wurden die Adulttiere mit Hilfe von CO_2 betäubt und anschließend unter dem Binokular betrachtet und sortiert. Die Bestimmung des Geschlechts erfolgte hierbei mit Hilfe eines Geschlechts-Dimorphismus der Antennen. Bei männlichen Tieren sind diese wirtelig behaart, wohingegen weibliche Tiere diese Behaarung nicht aufweisen. Bis zur weiteren Verwendung wurden die Tiere bei -80°C gelagert.

2.4 Molekularbiologische Standardmethoden

2.4.1 Phenol-Chloroform-Isoamylalkohol (PCI)-Extraktion

Zur Reinigung einer DNA-Lösung von Proteinen wurde die Lösung mit 1 Vol. PCI (Mischungsverhältnis 25:24:1, v/v/v) versetzt, das Reaktionsgefäß mehrfach invertiert und die Lösung im Anschluss für 10 min bei $4500 \times g$ (für Falcons), beziehungsweise $18600 \times g$ (für 2 ml Reaktionsgefäße) bei Raumtemperatur zentrifugiert. Im Anschluss wurde die obere, DNA-haltige Phase in ein frisches Reaktionsgefäß überführt. Je nach Grad der Verunreinigung wurde der PCI-Schritt mehrfach wiederholt. Um Phenolreste aus der DNA-haltigen Lösung zu entfernen, wurde im Anschluss stets eine Chloroform-Isoamylalkohol (CI)-Extraktion (Mischungsverhältnis 24:1, v/v) durchgeführt.

2.4.2 Fällung von DNA (Ethanol fällung)

Zur Entfernung von Salzen und zur Aufkonzentrierung der DNA erfolgte die Zugabe von 1/10 Vol. Dialysepuffer (10 x konzentriert) sowie 2,5 Vol. Ethanol (abs., kalt). Die Lösung wurde invertiert und für mindestens 60 min bei -20°C inkubiert. Im Anschluss erfolgte eine Zentrifugation für mindestens 45 min bei 4°C und $16100 \times g$. Der Überstand wurde vorsichtig abgezogen, das Pellet mit 70 % Ethanol überschichtet und die Lösung erneut für 5-10 min bei RT und $18600 \times g$ zentrifugiert. Der Überstand wurde verworfen und das Pellet in einer Vakuumzentrifuge „Concentrator 5301“ der

Fa. Eppendorf (Hamburg, Deutschland) getrocknet. Je nach Versuch wurden die Pellets in 5-500 µl frischem HPLC-Wasser gelöst.

2.4.3 Konzentrationsbestimmung (DNA)

Die Quantifizierung von DNA erfolgte einerseits durch den Vergleich mit definierten Banden der Molekulargewichtsmarker auf einem Agarosegel sowie mit Hilfe einer photometrischen Bestimmung. Diese erfolgte am NanoDrop® ND-1000 der Fa. PeqLab (Erlangen, Deutschland). Hierbei wurden die Absorptionen der Lösung bei 260 nm und bei 280 nm (Michelson 1958) gemessen. Auf Grund der Abweichung des Messwertes des NanoDrop durch Messung von RNA (280 nm) in einer DNA-haltigen Lösung war ein Vergleich mit einem Gel immer von Vorteil.

2.4.4 Agarose-Gelelektrophorese

Zur Größenauftrennung von DNA-Fragmenten aus den unterschiedlichen Versuchen (PCR, RT-PCR, Plasmidisolierung/ -restriktion, genomische DNA-Restriktion) wurden Agarosegele (0,8-2,5 % Agarose in 1x TBE oder 1 x E-Puffer) verwendet.

Für Testgele von Plasmid-Restriktionen, PCR- und RT-PCR-Proben wurden hierfür die GENTERPHORESE™-Gelkammersystem (7 cm x 6 cm x 0,1 cm) der Fa. Genterprise (Mainz, Deutschland) sowie das horizontale Gelkammersystem Gibco BRL Sunrise™ 96 der Fa. Life Technologies (Carlsbad, USA) verwendet. Die Auftrennung erfolgte bei 80-130 mA für 30-60 min.

Die Auftrennung von Restriktionen genomischer DNA erfolgte über Maxi-Gelkassetten (18 cm x 13 cm x 0,4 cm). Hierbei erfolgte die Auftrennung entweder bei 25-30 mA für 3-4 Std. in 1 x TBE, oder bei 14-16 mA für 16-18 Std. in 1 x E-Puffer.

Für präparative Gele wurde Agarose vom Typ II der Fa Sigma-Aldrich (St. Louis, USA) verwendet.

Die Proben wurden vor der Auftrennung mit 1/6 Vol. DNA-Ladepuffer versetzt. Als Molekulargewichtsmarker dienten je nach Anwendung die beiden Marker „GeneRuler™ 100 bp Plus DNA Ladder“ sowie „Lambda DNA x HindIII“ der Fa. Fermentas (St. Leon-Roth, Deutschland). Des Weiteren wurde in einigen Fällen der Molekulargewichtsmarker „1 kb Ladder“ der Fa. New England Biolabs (Ipswich, UK) verwendet. Die Färbung der Gele nach erfolgter Elektrophorese wurde in einer Ethidiumbromid-Färbelösung (5 µl/ml Ethidiumbromid in 1 x E-Puffer) durchgeführt. Im Anschluss wurden die Gele auf einem UV-Transilluminator der Fa. Bachofer (Reutlingen, Deutschland) bei einer Wellenlänge von 312 nm betrachtet und mit Hilfe des „Carestream GelLogic 112 Systems“ der Fa. Carestream Health (Rochester, USA) dokumentiert.

2.4.5 Wiedergewinnung von DNA aus Agarosegelen

Die Wiedergewinnung von PCR-Produkten aus Agarosegelen erfolgte mit Hilfe des Nucleo Spin® Extract II Kits der Fa. Machery-Nagel (Düren, Deutschland) nach den Angaben des Herstellers.

Handelte es sich um die Wiedergewinnung restringierter genomischer DNA, so erfolgte die Wiedergewinnung durch Elektroelution. Hierbei wurde das „blind“ aus dem Gel ausgeschnittene Gelstück zusammen mit 0,5 x TBE in einen Dialyseschlauch der Fa. Medicell International (London, UK) überführt und für 2 Std. bei 100 mA eluiert. Die Ausrichtung des Gelstückes im luftblasenfreien Dialyseschlauch erfolgte hierbei quer zum Stromfluss, wobei das Gelstück auf der Kathodenseite des Dialyseschlauchs lag. Nach erfolgter Elution wurde die DNA-haltige 0,5 x TBE-Lösung aus dem Dialyseschlauch entnommen. Optional wurde das Reaktionsgefäß zum Entfernen von Agaroseresten für 10 min auf Eis gestellt und im Anschluss für 5 min bei 4°C und 16100 x g zentrifugiert. Der Überstand wurde in diesem Fall in ein frisches Reaktionsgefäß überführt. Es folgten eine PCI/CI-Extraktion und eine Ethanol-fällung der DNA, bevor diese in einem geringen Volumen (20-30 µl) frischen HPLC-Wassers gelöst wurde.

2.4.6 Restriktionen

Die durchgeführten Restriktionen dienten der Überprüfung der Integratgrößen von DNA-Fragmenten in verschiedenen Vektoren, der Erstellung von Southern-Blots sowie der Erstellung genomischer BAC-Bibliotheken.

Die Dauer der Restriktion richtete sich hierbei nach dem jeweiligen Verfahren. So wurden die unvollständigen Restriktionen (*Sau3A1*) genomischer DNA aus *C. tentans* zur Erstellung von BAC-Bibliotheken in einem großen Volumen (3-5 ml) mit wenig Enzym (0,1 µl/ml Ansatz) und einer kurzen Dauer (1-4 min) durchgeführt. Restriktionen zur Überprüfung von Integratgrößen hingegen wurden in einem Volumen von 20 µl mit 1 µl Enzym für die Dauer von 3-4 Std. durchgeführt. Die Restriktionen genomischer DNA zur Erstellung von genomischen Southern-Blots erfolgten in Volumina von 200-500 µl über eine Zeit von 16-18 Std. mit 2 µl Enzym und der weiteren Zugabe von 0,5-1 µl Enzym nach mindestens 2 Stunden.

Alle verwendeten Enzyme stammen von den Firmen Roche (Mannheim, Deutschland), Fermentas (St. Leon-Roth, Deutschland) und New England Biolabs (Ipswich, UK).

2.5 DNA-Präparation

2.5.1 Isolierung genomischer DNA aus Zellkultur

Die 40 ml Kulturlösung einer dichtgewachsenen Zellkulturflasche wurden für 15 min bei RT und 300 x g abzentrifugiert. Die Pellets aus mehreren Flaschen wurden in 5-10 ml

1 x PBS resuspendiert, zusammengeführt und erneut für 15 min bei 300 x g zentrifugiert. Der Überstand wurde verworfen und die Pellets entweder direkt weiterbearbeitet oder bei -20°C/-80°C gelagert.

Die weitere Bearbeitung der Zellkulturpellets erfolgte in L-Puffer. Diese wurden in 5 ml L-Puffer resuspendiert und erneut für 15 min bei 300 x g abzentrifugiert. Der Überstand wurde verworfen, die Pellets in 1,6 ml L-Puffer resuspendiert und die Lösung durch Zugabe von SDS-Lösung (Endkonzentration 1-2 %) und 0,008 g Proteinase K der Fa. Merck (Darmstadt, Deutschland) komplementiert. Der Proteinase K-Verdau erfolgte für 60 min bei 50°C. Im Anschluss wurde die Lösung in einen Dialyseschlauch der Fa. Medicell International (London, UK) gefüllt und über Nacht gegen autoklaviertes VE-Wasser dialysiert. Die so erhaltene DNA-Lösung konnte direkt zur Erstellung von BAC-Bibliotheken eingesetzt werden oder nach PCI/CI-Extraktion bei 4°C gelagert werden.

2.5.2 Isolierung von Plasmid-DNA

Allen hier aufgeführten Plasmid-Präparationen liegt die Methode der alkalischen Lyse (Birnboim und Doly 1979) zugrunde. Die Vektorpräparation der Vektoren pBeloBAC11 sowie pUC18 erfolgte mit Hilfe des Whitehead-Protokolls. Die Präparationen der Plasmid-DNA von BAC-Klonen wurde ebenfalls mit Hilfe des Whitehead-Protokolls durchgeführt. Alternativ wurde hierfür auch das „PureYield™ Plasmid Maxiprep System“ der Fa. Promega (Fitchburg, USA) verwendet. Die Plasmid-Präparation wurde hierbei nach den Angaben des Herstellers durchgeführt. Sämtliche Klone, welche aus den Ligationen von DNA-Fragmenten in den Vektor pUC18 oder den Vektor pGem®T Easy der Fa. Promega (Fitchburg, USA) hervorgegangen waren, wurden mit Hilfe des „peqGold Plasmid Miniprep Kit I“ der Fa. Peqlab (Erlangen, Deutschland) oder des „PureYield™ Plasmid Miniprep Systems“ der Fa. Promega (Fitchburg, USA) nach den Angaben der Hersteller aufgearbeitet. Klone, welche aus Shotgun-Bibliotheken hervorgegangen waren, wurden mit Hilfe der Plasmid-Präparation im 96er-Format aufgearbeitet.

2.5.2.1 Plasmid-Präparation nach dem Whitehead-Protokoll (Whitehead Institute/ MIT)

Für eine Vorkultur wurden 5 ml LB-Medium mit 1/100 Vol. Antibiotikum-Stammlösung versetzt und zusammen mit der entsprechenden Bakterienkolonie für mindestens 6 Std. bei 37°C inkubiert. Das Wachstum der Bakterienkultur (500 ml; 1/100 Vol. Antibiotikum-Stammlösung) erfolgte im Anschluss für 14-18 Std. bei 37°C. Die Kultur wurde für 10 min bei 4°C und 5000 x g abzentrifugiert. Der Überstand wurde verworfen und das Bakterien-Pellet in 10 ml Lösung I resuspendiert. Es folgte eine Inkubation von 5 min bei RT. Nach Zugabe von 20 ml Lösung II und vorsichtigem Invertieren der Lösung erfolgte eine weitere Inkubation von 10 min bei RT. Die Neutralisierung der

Lösung wurde durch die Zugabe von 15 ml Lösung III erreicht, welcher ein erneutes vorsichtiges Invertieren und eine Inkubation von 10 min auf Eis folgten. Im Anschluss wurde die Lösung für 20 min bei 4°C und 21500 x g zentrifugiert. Unter Umständen wurde der Schritt so lange wiederholt, bis ein klarer Überstand erreicht war. Die im Überstand befindlichen Plasmide wurden durch Zugabe von 1 Vol. Isopropanol gefällt und im Anschluss für 20 min bei RT und 5000 x g pelletiert. Nach Verwerfen des Überstandes wurden die Plasmide in 6 ml TE-Puffer resuspendiert. Nach Zugabe von 3 ml 7,5 M Kaliumacetat erfolgte eine Inkubation von mindestens 45 min bei -80°C. Bei der anschließenden Zentrifugation (15 min, RT, 5000 x g) wurde das in Lösung befindliche SDS als Kalium-SDS pelletiert. Die im Überstand befindlichen Plasmide wurden in ein neues Falcon überführt, durch Zugabe von 2 Vol. Ethanol (abs.) und Inkubation von mindestens 20 min bei -20°C gefällt und bei der anschließenden Zentrifugation (15 min, 4°C, 5000 x g) pelletiert. Der Überstand wurde verworfen und das Pellet in 700 µl 50T/50E-Puffer resuspendiert. Nach Zugabe von 1 µl Rnase A Lösung (E.Z.N.A. Plasmid Miniprep Kit I) der Fa. Peqlab (Erlangen, Deutschland) und anschließender Inkubation für 45 min bei 37°C erfolgten drei PCI- sowie eine CI-Extraktion. Die in der wässrigen Phase befindlichen Plasmide wurden mit Hilfe von 1 Vol. Isopropanol gefällt und durch Zentrifugation (15 min, RT, 14000 x g) pelletiert. Nach anschließendem Waschen in 70 % Ethanol wurden die Plasmide getrocknet und in 100-300 µl HPLC-Wasser gelöst.

2.5.2.2 Plasmid-Präparation im 96er Mikrotiter-Plattenformat

Die Plasmid-Präparation im Rahmen der Shotgun-Bibliotheken erfolgte auf Grund der hohen Anzahl an aufzuarbeitenden Plasmiden im 96er Plattenformat. Für die Präparation wurden „Sephadex™ G-50 Fine MTP“ vorbereitet. Die Aufreinigung über Sephadex™ G-50 beruht hierbei auf dem Prinzip der Gelfiltration (Porath und Flodin 1959). Für die Vorbereitung der „Sephadex™ G-50 Fine MTP“ wurden 96er Multiscreen™-HV Filterplatten der Fa. Millipore (Billerica, USA) mit 30 mg Sephadex™ G50-Fine der Fa. GE Healthcare (München, Deutschland) pro Position befüllt. Im Anschluss wurden diese mit Hilfe von je 300 µl HPLC-Wasser zum Quellen gebracht und für mindestens 3 Std. bei 4°C gelagert. Nach Zentrifugation für 3 min bei 4°C und 870 x g erfolgte die erneute Zugabe von 150 µl HPLC-Wasser und die Lagerung bei 4°C.

Die Anzucht der Bakterienkulturen erfolgte in 96er-„Deep-Well“-Platten der Fa. ABgene (Epsom, UK). Pro Klon wurden hierfür 1,2 ml 2 x LB-Medium mit 1/100 Vol. Ampicillin-Stammlösung vorgelegt. Das Wachstum der Kolonien erfolgte für 24 Std. bei 37°C. Zum Pelletieren der Bakterien erfolgte eine Zentrifugation für 10 min bei 4°C und 2500 x g. Der Überstand wurde verworfen und die Bakterien-Pellets in 335 µl Lösung I resuspendiert. Nach Inkubation für 2 min auf Eis erfolgte durch die Zugabe von 665 µl Lösung II und vorsichtiges Invertieren die Lyse der Zellen. Im Anschluss wurden die Platten erneut für 10 min auf Eis inkubiert, bevor die Zugabe von 500 µl Lösung III zur Neutralisierung erfolgte. Nach vorsichtigem Invertieren wurden die Platten für 15 min

bei 4°C und 2500 x g zentrifugiert. Im Anschluss wurde je 1 ml Überstand in eine frische 96er-„Deep-Well“-Platte überführt, in welcher sich bereits je 1 ml Isopropanol pro Position befand. Die Plasmide wurden durch Zentrifugation (30 min, RT, 2500 x g) pelletiert. Der Überstand wurde verworfen und die Plasmide durch Zugabe von je 50 µl HPLC-Wasser und Inkubation für mindestens 1 Std. gelöst. Die gelösten Plasmide wurden auf die vorbereiteten „Sephadex™ G-50 Fine MTP“ aufgetragen und durch Zentrifugation (5 min, RT, 870 x g) aufgereinigt. Das Eluat wurde in 96er „Multiple Well Platten“ der Fa. Sarstedt (Newton, USA) aufgefangen. Die Lagerung der Platten erfolgte bei -20°C.

2.6 Klonierung

2.6.1 Verwendeter Bakterienstamm

Für alle durchgeführten Experimente wurde der *E. coli* - Stamm DH10B genutzt. Dieser wurde erstmals 1990 erstellt (Grant et al. 1990) und besitzt folgenden Genotyp:

F- mcrA Δ(mrr-hsdRMS- mcrBC) φ80lacZΔM15 ΔlacX74 recA1 endA1 araD139 Δ(ara,leu)7697 galU galKλ- rpsL nupG

2.6.2 Herstellung elektrokompenter Zellen

Für die Herstellung elektrokompenter Zellen wurden 50 ml LB-Medium als Vorkultur mit einer Einzelkolonie des *E.coli* Stammes DH10B angeimpft und für mindestens 14 Std. bei 37°C inkubiert. Je 475 ml frisches LB-Medium wurden mit 25 ml der Vorkultur versetzt und bis zu einer optischen Dichte (OD) von 0,71-0,72 bei 37°C inkubiert. Die OD wurde hierfür mit einem BioPhotometer der Fa. Eppendorf (Hamburg, Deutschland) bestimmt. Die Zellsuspension wurde im Anschluss für mindestens 40 min auf Eis inkubiert und im Folgenden pelletiert (15 min, 4°C, 1300 x g). Der Überstand wurde verworfen, die Zellen in 160 ml frischem HPLC-Wasser resuspendiert und erneut zentrifugiert (20 min, 4°C, 1300 x g). Der Überstand wurde verworfen und die Zellen in 80 ml frischem HPLC-Wasser gelöst. Nach einer weiteren Zentrifugation (20 min, 4°C, 1300 x g) und Verwerfen des Überstandes erfolgte die Resuspension in 40 ml 10 % Glycerin. Im Anschluss wurden die Zellen ein weiteres Mal pelletiert (20 min, 4°C, 1300 x g) und der Überstand vorsichtig abgeschüttet. Das Bakterien-Pellet wurde im noch vorhandenen Glycerin resuspendiert und in 40 µl Aliquots aufgeteilt. Die Lagerung der Zellen erfolgte bei -80°C.

Die Überprüfung der Zellen erfolgte mit einem pUC18-Standard (10 pg/µl). Die Transformationsrate ist hierbei definiert als die Zahl der Zellen, die durch 1 µg Plasmid-DNA transformiert werden.

2.6.3 Verwendete Vektoren

Für die durchgeführte Arbeit wurden drei verschiedene Vektoren verwendet. Alle verwendeten Vektoren besitzen den Vorteil der „Blau-Weiß“-Selektion, welcher mit Hilfe der α -Komplementation des LacZ-Gens (Langley et al. 1975) erreicht wird.

Für die Erstellung genomischer BAC-Bibliotheken wurde der Plasmid-Vektor pBeloBAC11 (Mori et al. 1986; Shizuya et al. 1992; Kim et al. 1996) verwendet. Dieser besitzt die Fähigkeit, DNA-Fragmente bis zu einer Größe von 300 kb aufnehmen.

Zur Subklonierung der aufgearbeiteten BAC-Klone mittels Shotgun-Bibliotheken wurde der Vektor pUC18/19 (Yanisch-Perron et al. 1985) verwendet.

Zur Klonierung von PCR-Produkten wurde der Plasmid-Vektor pGem®T-Easy der Fa. Promega (Fitchburg, USA) genutzt. Dieser weist den Vorteil von 3' T-Überhängen auf, durch die eine besondere Effizienz der Klonierung von PCR-Produkten, welche 3' A-Überhänge besitzen, gewährleistet wird.

2.6.4 Ligation

Die unterschiedlichen Ligationen wurden in einem Gesamtvolumen von 10-50 μ l angesetzt. Die eingesetzte Vektormenge betrug je nach Ligation 5-50 ng. Als Ligase wurden je 1-2,5 U der T4-Ligase der Fa. Roche Diagnostics GmbH (Mannheim, Deutschland) verwendet. Nach erfolgter Ligation (ü.N. bei RT) wurde der Ligations-Ansatz gefällt und in 5-10 μ l HPLC-Wasser gelöst.

2.6.5 Transformation

Die Transformation der elektrokompenten Zellen erfolgte nach der Methode der Elektroporation (Dower et al. 1988). Hierfür wurde 1 μ l des rekombinanten Vektors mit einem Aliquot der elektrokompenten Zellen vermischt und in eine vorgekühlte Küvette (Gene Pulser® Cuvette, Spaltbreite 0,1 cm) der Fa. Bio-Rad (München, Deutschland) überführt. Nach erfolgter Transformation bei 1,8 kV im Micro Pulser™ der Fa. Bio-Rad erfolgte die Aufnahme in 1 ml vorgewärmtes LB-Medium sowie die Inkubation für 45-60 min bei 37°C. Nach erfolgtem Ausplattieren des Transformationsansatzes auf entsprechenden Agar-Platten erfolgte die Inkubation der Platten bei 37°C über Nacht.

2.7 Erstellen einer BAC-Bibliothek aus *C. tentans*

2.7.1 Integrat-DNA

Die Aufarbeitung der genomischen DNA aus Zellkultur-Pellets erfolgte wie in Kapitel 2.5.1 beschrieben. Als Ausgangsmenge an DNA für eine BAC-Bibliothek wurde hierbei die DNA-Menge aus 1-2 Zellkulturpellets (etwa 6 x 200 ml T-Flaschen) verwendet.

Nach erfolgter Aufarbeitung wurde die genomische DNA mit dem Restriktionsenzym *Sau3AI* (0,1 µl / 1 ml Ansatz) der Fa. Roche (Mannheim, Deutschland) für 1-4 min bei 37°C restringiert. Hierdurch erfolgte ein gewünschter, unvollständiger Verdau der genomischen DNA. Zum Abstoppen der Restriktion wurde dem Ansatz 1/10 Vol. 0,5 M EDTA zugefügt. Des Weiteren erfolgte eine Inaktivierung für 20 min. bei 60°C sowie eine PCI/CI-Extraktion. Im Anschluss wurde die restringierte DNA gefällt. Nach erfolgter Aufnahme des DNA-Pellets in frisches HPLC-Wasser erfolgte die Auftrennung über ein Agarosegel (0,8 %, 1x E-Puffer, Maxi-Gelkassetten) bei 14-16 mA für 16-18 Stunden. Die Fragmente der aufgetrennten DNA mit einer Größe von mehr als 20 kb wurden „blind“ aus dem Gel ausgeschnitten und per Elektroelution wiedergewonnen. Nach einer Fällung wurde die DNA erneut in HPLC-Wasser aufgenommen und ein zweites Mal einer Größenselektionierung über ein Agarosegel (0,8 %, 1x E-Puffer, Maxi-Gelkassetten) unterzogen. Im Anschluss an die zweite Wiedergewinnung mit anschließender Fällung stand die DNA zur Ligation bereit.

2.7.2 pBeloBAC11-Vektor Vorbereitung

Der Plasmid-Vektor pBeloBAC11 wurde per Whitehead-Protokoll (siehe Kapitel 2.5.2.1) aus einer Einzelkolonie aufgearbeitet. Die Aufreinigung der Vektor-DNA erfolgte danach über die CsCl-Dichtegradientenzentrifugation (siehe Kapitel 2.9.1). Etwa 5 µg des Vektors wurden mit dem Restriktionsenzym BamHI (20 U) der Fa. Roche (Mannheim, Deutschland) für 4 Std. restringiert. Die Restriktion wurde auf einem Agarosegel (1 %, GENterphoresis™-Gelkammersystem) getestet. Hierzu wurde 1/10 Vol. des Restriktionsansatzes (10 µl) verwendet. Im Anschluss wurde der restringierte Vektor mit Hilfe der Antarctic-Phosphatase der Fa. New England Biolabs (Ipswich, UK) dephosphoryliert. Die Dephosphorylierung erfolgte hierbei für 15 min bei 37°C. Der so behandelte pBeloBAC11-Vektor wurde aliquotiert und bei -20°C bis zur Erstellung von BAC-Bibliotheken gelagert.

2.8 Erstellen einer Shotgun-Bibliothek

Als Ausgangsmenge an DNA wurden 30-50 µg UZ-aufgereinigte (Kapitel 2.9.1) Plasmid-DNA (BAC-Klon) eingesetzt. Die DNA-Lösung wurde mit 10 mM Tris (pH 7,5) auf ein Volumen von 2 ml aufgefüllt und für 25 sec. bei 1 bar in einem „Nebulizer“ der Fa. Life Technologies (Carlsbad, USA) mechanisch geschert. Im Anschluss erfolgte eine Ethanol-fällung. Die Pellets wurden in einem Gesamtvolumen von 16 µl HPLC-Wasser gelöst. Zur Überprüfung der gescherten Plasmid-DNA wurden 1,5 µl der Probe auf ein Agarosegel (1 %, GENterphoresis™-Gelkammersystem) aufgetragen.

In Anschluss erfolgte das Auffüllen der gescherten DNA-Enden. Hierzu wurden zu 14,5 µl DNA 5 U Klenow-Enzym sowie 3 U T4-DNA-Polymerase der Fa. Roche (Mannheim, Deutschland) zugegeben (30 µl Ansatz) und für 30 min bei RT inkubiert. Es folgte die Auftrennung auf einem präparativen Agarosegel (1 %).

Die DNA-Fractionen mit Größen von 1-2 kb und 2-4,5 kb wurden „blind“ aus dem Gel ausgeschnitten und per Elektroelution (hier in 1 x TBE, 1 Std., 100 mA) aus dem Gel wiedergewonnen. Nach einer PCI/CI-Extraktion wurde die DNA gefällt und anschließend in 10 µl HPLC-Wasser gelöst. Zur Überprüfung der wiedergewonnenen DNA-Fractionen wurden diese auf einem Agarosegel (1 %, GENterphorese™-Gelkammersystem) getestet.

Die Ligation der DNA in den Vektor pUC18 (30 ng, *Sma*I restringiert, dephosphoryliert) erfolgte mit 1,5 U T4-DNA-Ligase der Fa. Roche (Mannheim, Deutschland) bei 4°C über Nacht. Die Ligation wurde gefällt und anschließend in 5 µl HPLC-Wasser gelöst. Für die Transformation wurde je 1 µl des gefällten Ligationsansatzes eingesetzt.

2.9 DNA-Aufreinigung

2.9.1 DNA-Aufreinigung mittels CsCl-Dichtegradientenzentrifugation

Für die Durchführung der CsCl-Dichtegradientenzentrifugation (Chirgwin et al. 1979) wurde die aufzureinigende Plasmid-DNA in 1x TE-Puffer mit CsCl (1 g/ ml) aufgenommen. Die eingesetzte DNA-Menge betrug hierbei 30-50 µg. Das Volumen wurde mit 1,6 ml (als Mindestvolumen für die Zentrifugenröhrchen) gewählt. Die DNA-haltige Lösung wurde mit 64 µl EtBr-Lösung (5 mg/ ml) überschichtet und im Anschluss durch schnelles Invertieren vermischt. Die Einstellung des richtigen Brechungsindex (1,391-1,395 bei 20°C) erfolgte mit Hilfe eines Refraktometers der Fa. Carl Zeiss (Oberkochen, Deutschland) durch Zugabe von gesättigter CsCl-Lösung. Ein Gesamt-Füllvolumen von 1,8 ml pro Zentrifugenröhrchen dürfte hierbei nicht überschritten werden. Die Zentrifugation erfolgte für 24 Std. bei 20°C und 40000 Upm in der Ultrazentrifuge (UZ) „Sorval® Discovery™ M120 SE“ der Fa. Thermo Fisher Scientific (Waltham, USA). Nach erfolgter Zentrifugation wurde die Bande der gewünschten supercoiled DNA unter UV-Licht abgezogen. Im Anschluss erfolgte eine Extraktion durch Zugabe von 1 Vol. Isoamylalkohol zur Entfernung des Ethidiumbromids. Die Lösung wurde mehrfach invertiert und im Folgenden zentrifugiert (5 min, RT, 18500 x g). Die obere organische Phase wurde verworfen und die Extraktion mit Isoamylalkohol so lange durchgeführt, bis keine Färbung mehr in der organischen Phase zu verzeichnen war. Durch Dialyse (12-16 Std., 4°C) gegen 1 x Dialysepuffer wurde das CsCl aus der Lösung entfernt. Im Anschluss wurde die Plasmid-Lösung gefällt und in HPLC-Wasser aufgenommen. Zur quantitativen Überprüfung wurden 1-2 µl der DNA auf einem Agarosegel (1 %, GENterphorese™-Gelkammersystem) getestet.

2.9.2 Aufreinigung von PCR-Produkten

Die Aufreinigung von PCR-Produkten erfolgte entweder mit Hilfe des NucleoSpin® Extract II Kits der Fa. Macherey-Nagel (Düren, Deutschland) nach den Angaben des

Herstellers oder durch enzymatischen Verdau. Hierfür wurden dem PCR-Ansatz 10 U Exonuklease I und 0,9 U alkalische Phosphatase (AP) der Fa. Fermentas (St. Leon-Roth, Deutschland) zugegeben. Die Inkubation erfolgte für 25 min bei 37°C, gefolgt von einer Inaktivierung der Enzyme für 15 min bei 72°C.

2.10 Primerdesign

Je nach Anwendung wurden unterschiedliche Primer konstruiert. Primer für Standard-PCR sowie RT-PCR wurden mit einer Länge von 18-22 Nukleotiden bestellt. Für RACE-Experimente wurden hingegen Primer mit einer Länge von bis zu 28 Nukleotiden bestellt. Alle verwendeten Primer wurden so konstruiert, dass sie einen GC-Gehalt von 40-60 % aufweisen und keine stabileren Haarnadel-Strukturen ausbilden. Weiterhin wurde darauf geachtet, dass die Primer möglichst keine stabileren Dimere (Rychlik 1995) mit sich selbst (Homodimere) oder dem zweiten PCR-Primer (Heterodimere) bilden. Die Konstruktion der Primer erfolgte mit Hilfe der online verfügbaren Programme „Oligonucleotide Properties Calculator“ (Kibbe 2007) sowie „IDT SciTools OligoAnalyzer 3.1“ (Owczarzy et al. 2008).

Alle Primer wurden bei den Firmen Invitrogen (Carlsbad, USA) oder Biomers (Ulm, Deutschland) bestellt und auf 10 pmol/ µl (10 µM) eingestellt.

2.11 Polymerase-Kettenreaktion (PCR)

Im Laufe der vorliegenden Arbeit wurden unterschiedliche PCR (Mullis et al. 1992*) durchgeführt (PCR, „Touchdown“-PCR, „Long-range“-PCR, RT-PCR). In allen Fällen wurde die PCR in einem Reaktions-Volumen von 50 µl durchgeführt. Mit Ausnahme der „Long-range“-PCR wurden alle PCR mit der GoTaq®-Polymerase der Fa. Promega (Fitchburg, USA) durchgeführt und in den Thermocyclern „Tprofessional“ Fa. Biometra (Göttingen, Deutschland) sowie „Primus96“ der Fa. Peqlab (Erlangen, Deutschland) durchgeführt. Die genauen Zusammenstellungen der PCR-Ansätze sind in Tabelle 2.1 aufgeführt. Des Weiteren sind in Tabelle 2.2 und Tabelle 2.3 die durchgeführten Programme aufgelistet.

Tabelle 2.1: PCR-Ansätze

Komponenten	PCR/ RT-PCR/ "Touchdown"-PCR	"Long-range"-PCR
genomische DNA/ cDNA	10-30 ng genomische DNA/ cDNA	20-200 ng genomische DNA
dNTP-Mix (10 mM each)	1 µl	2 µl
Primer I (10 µM)	1 µl	1 µl
Primer II (10µM)	1 µl	1 µl
MgCl ₂ (25 mM)	4 µl	-
PCR-Puffer GoTaq (5x)	10 µl	-
PCR-Puffer LongAmp (5x)	-	10 µl
GoTaq-Polymerase (5U/µl)	0,25 U	-
Enzym-Mix	-	5 U
HPLC-Wasser	x µl	x µl
Gesamtvolumen	50 µl	50 µl

Tabelle 2.2: Programm einer PCR/ RT-PCR

Schritt	Temperatur	Dauer	Zyklen
1. Initiale Denaturierung	94°C	4 min	1
2. Denaturierung	94°C	30 sec	35
3. Annealing	55-65 °C	30 sec	
4. Elongation	72°C	30-120 sec	
5. Finale Elongation	72°C	10 min	1
	16°C	∞	1

Tabelle 2.3: Programm einer „Touchdown“-PCR (links) sowie einer Long-range-PCR (rechts)

Schritt	Temperatur	Dauer	Zyklen	Temperatur	Dauer	Zyklen
1. Initiale Denaturierung	94°C	4 min	1	94°C	4 min	1
2. Denaturierung	94°C	30 sec	10	94°C	30 sec	10
3. Annealing	60-70 °C	30 sec	(-1°C pro Zyklus)	52-65 °C	30 sec	
4. Elongation	72°C	30-180 sec		65°C	8 min	
5. Denaturierung	94°C	30 sec	35	94°C	30 sec	25
6. Annealing	55-65 °C	30 sec		52-65 °C	30 sec	
7. Elongation	72°C	30-180 sec		65°C	8 min	
8. Finale Elongation	72°C	10 min	1	65°C	10 min	1
	16°C	∞	1	16°C	∞	1

2.12 DNA-Sequenzierung

Die durchgeführten DNA-Sequenzierungen beruhen auf dem Prinzip der Kettenabbruchsynthese (Sanger et al. 1977; Prober et al. 1987*; Lee et al. 1992) und wurden von der Fa. Genterprise Genomics (Mainz, Deutschland) durchgeführt. Bei allen Sequenzierungen wurde hierbei ein U-Mix Ansatz aus maximal 6 µl DNA und Wasser sowie 1 µl Sequenzier-Primer (10 µM) vorbereitet. Als Primer dienten die Standardprimer Sp6 und T7 (BAC-Klone und pGem®T Easy-Klone), M13for und M13rev (pUC18/ 19-Klone) sowie produktspezifische Primer für die Sequenzierung von PCR-Fragmenten.

2.13 Bioinformatische Sequenzauswertung

Die von der Fa. Genterprise Genomics (Mainz, Deutschland) erhaltenen Sequenzdaten wurden auf unterschiedlichen Wegen näher analysiert. Zu einer ersten Beurteilung der Daten wurden diese mit dem Programm „FintchTV v.1.4.0“ der Fa. Geospiza (Seattle, USA) betrachtet und mit Hilfe des Internettools VecScreen (NCBI) gegebenenfalls auf Vektoranteile hin untersucht. Weitere Sequenz-Informationen konnten mit Hilfe von BLAST-Analysen (Altschul et al. 1990) erlangt werden. Der größte Teil der Analysen erfolgte mit Hilfe der Programme „DNASTAR Lasergene®“- Software Paket der Fa. Lasergene (Madison, USA) sowie „CLC Main Workbench 5“ der Fa. CLC bio (Katrinebjerg, Dänemark). Für die Analysen von Intron/Exon-Strukturen und offenen Leserahmen (ORF) wurden weiterhin der ORF-Finder sowie Spidey (beide NCBI), das Programm GENSCAN (Burge und Karlin 1997) und das Programm AUGUSTUS (Stanke und Morgenstern 2005) genutzt. Zur Suche nach in Genen vorkommenden Domänen wurde das Programm InterProScan (Quevillon et al. 2005) verwendet.

2.14 DNA-Markierungstechniken

Im Laufe der Arbeit erfolgten sowohl radioaktive wie auch nicht-radioaktive Markierungen von DNA. Die Markierung der DNA erfolgte hierbei mit Hilfe des „Random Primed Oligo Labeling Kits“ (radioaktiv) oder des „DIG DNA Labeling Kits“ (nicht radioaktiv) der Fa. Roche (Mannheim, Deutschland) nach der Methode von (Feinberg und Vogelstein 1983) sowie dem „Bio Nick™ Labeling System“ der Fa. Invitrogen (Carlsbad, USA) nach der Methode von (Macgregor und Mizuno 1976). Für alle Markierungen wurde nach den Angaben des Herstellers verfahren. Die hierbei eingesetzten DNA-Mengen lagen bei der radioaktiven Markierung, wie auch bei den nicht-radioaktiven Markierungen bei 200 ng- 1 µg. Für die radioaktive Markierung wurden zwischen 30 µCi $\alpha^{32}\text{PdATP}$ (Southern-Hybridisierung von BAC-Klonen) und 70 µCi $\alpha^{32}\text{PdATP}$ (Koloniefilter-Hybridisierung) eingesetzt.

2.15 Hybridisierungstechniken

2.15.1 Fluoreszenz in-situ-Hybridisierung (FISH)

Die Lokalisation der positiven BAC-Klone aus *C. tentans* erfolgte per FISH (Langer-Safer et al. 1982) an den Polytänochrosomen von *C. thummi*. Die entsprechenden BAC-Klone wurden hierfür mit Hilfe des „Bio Nick™ Labeling System“, die für eine Doppelhybridisierung verwendeten synthetischen Cla-Elemente mit Hilfe des „DIG DNA Labeling Kits“ markiert.

Nach der Präparation der Speicheldrüsen aus *C. thummi*-Larven (L4-Stadium) wurden diese für 5 min in Essigsäure (50 %) inkubiert, gequetscht und für mindestens 10 min auf Kohlendäureschnee gelagert. Im Folgenden wurde das Deckglas abgesprengt und die Objektträger in einer Kuvette mit Isopropanol (100 %) bei -20°C für mindestens 16 Std. gelagert. Die Präparate wurden im Folgenden den weiteren Behandlungen unterzogen:

Alkoholreihe Isopropanol (70 %, 50 %, 30 %, RT, je 5 min), 0,1 x SSC (RT, 5 min), 2 x SSC (70°C, 30 min), 2 x SSC (RT, 10 min), 0,05 N NaOH (RT, 1 min), 0,1 x SSC kalt (RT, 5 min), Alkoholreihe Isopropanol (30 %, 50 %, 70 %, RT, je 5 min). Die Lagerung der stabilisierten und denaturierten Chromosomenpräparate erfolgte bis zur weiteren Verwendung in Isopropanol (100 %) bei -20°C.

Vor dem Aufbringen der markierten Sonden wurden die Präparate an der Luft getrocknet. Es folgte das Aufbringen der Sonden (je 5 µl bei Doppelhybridisierung, 8 µl bei Einzelhybridisierung) und die Inkubation in einer feuchten Kammer (52°C, 14-18 Std.). Die Präparate wurden in 2 x SSC (5 min) und 1 x PBS (10 min) gewaschen. Im Anschluss wurden die entsprechenden Antikörper auf die Präparate aufgegeben. Es folgte eine Inkubation bei 37°C für mindestens 2,5 Stunden. Nach der Entfernung überschüssigen Antikörpers von den Präparaten (1 x PBS, 10 min) wurden diese mit 8 µl Glycerin-Para-Phenylendiamin-Mix eingebettet und mit Lack versiegelt.

Die Dokumentation der Präparate erfolgte am Fluoreszenzmikroskop „Olympus BX61“ der Fa. Olympus (Hamburg, Deutschland) mit Hilfe der Software „Cell F“ sowie einer „Color View III“-Kamera der Fa. Soft Imaging System (Münster, Deutschland).

2.15.2 Southern-Hybridisierung

Zum Transfer von DNA-Fragmenten auf eine Nitrocellulosemembran wurde die Methode nach (Southern 1975) angewendet. Hierbei wurden sowohl das unidirektionale wie auch das bidirektionale Bloten durchgeführt. Nach erfolgter Auftrennung der DNA-Fragmente durch Gelelektrophorese wurden die Markerbanden im Gel durchstochen, das Gel denaturiert (20 min, Denaturierungslösung), neutralisiert (20 min, Neutralisierungslösung), in 2 x SSC gewässert und geblotet (RT, 12-16 Std.). Im Anschluss konnten die Markerbanden und die Konturen des Gels auf der Nitrocellulosemembran mit einem Fettstift markiert werden. Die Fixierung der DNA an die Nitrocellulosemembran erfolgte durch Inkubation für 2 Std. bei 80°C.

Zur Vorbereitung der Hybridisierung wurde die Nitrocellulosemembran für 2-4 Std. bei 60°C in 2 x PM in 6 x SSC (Denhardt 1966) mit SDS (Endkonzentration 1 %) präinkubiert. Die Hybridisierung erfolgte im Anschluss in 1 x PM in 3 x SSC in einer Hybridisierungsröhre (60°C, 12-16 Std.). Durch mehrfaches Waschen der Nitrocellulosemembran mit 2 x SSC bei 60°C wurden überschüssige Sondenmoleküle entfernt. Die trockene Nitrocellulosemembran wurde in einer mit einer Verstärkerfolie ausgestatteten Expositionskassette zusammen mit einem Röntgenfilm der Fa. Fuji Medical X-Ray (Tokyo, Japan) für die Dauer von einigen Minuten bis zu mehreren Tagen bei -80°C exponiert.

2.15.3 Koloniefilter-Hybridisierung

Um BAC-Klone mit spezifischen Integraten aus einer BAC-Bibliothek herauszufiltern, wurde die Methode der Koloniefilter-Hybridisierung angewendet (Grunstein und Hogness 1975). Hierfür wurde ein „Abklatsch“ der auf Agarplatten befindlichen Bakterien-Kolonien einer BAC-Bibliothek auf Nitrocellulose-Rundfilter erzeugt.

Hierzu wurden diese vorsichtig auf die Agarplatten gelegt, asymmetrisch markiert und wieder vorsichtig von den Platten abgezogen.

Zur Bestätigung von vermeintlich positiven BAC-Klonen einer BAC-Bibliothek wurden diese in identischer Weise auf zwei gerasterte Nitrocellulose-Rundfilter übertragen, welche zum Wachstum der Bakterien auf entsprechenden Agarplatten inkubiert wurden (37°C, 12-16 Std.). Einer der beiden Filter wurde anschließend für die Koloniefilter-Hybridisierung vorbereitet.

In beiden Fällen wurden die Nitrocellulose-Filter dann auf Filterpapier mit 0,5 M NaOH inkubiert (7 min) und getrocknet. Im Anschluss folgten die Inkubationen auf Filterpapier mit 1,5 M Tris (pH 7) sowie einer 0,5 M Tris/ 1,5 M NaCl-Lösung (pH 7,4) für je 5 Minuten. Nach erneutem Trocknen der Filter wurden diese für 30 min in eine Proteinase K-Lösung (1 mg/ml in 1 x SSC) gelegt und im Anschluss kurz durch EtOH (100 %) gezogen. Die trockenen Filter wurden in 0,3 M NaCl inkubiert (5 min) und anschließend zum Entfernen der Bakterienreste zwischen zwei Lagen Filterpapier gepresst. Die fertig behandelten Nitrocellulose-Filter wurden im Folgenden für 2 Std. bei 80°C inkubiert. Hierdurch erfolgte das Fixieren der DNA an die Nitrocellulose.

Zur Vorbereitung der Hybridisierung wurden die Nitrocellulose-Filter für 2-4 Std. bei 60°C in 2 x PM in 6 x SSC (Denhardt 1966) mit SDS (Endkonzentration 1 %) präinkubiert. Die Hybridisierung erfolgte danach in 1 x PM in 3 x SSC. Hierfür wurden 15-20 Filter in Petrischalen gelegt, mit Sonde überschichtet und bei 60°C für 12-16 Std. inkubiert. Durch zwei- bis dreimaliges Waschen der Nitrocellulosemembran mit 2 x SSC bei 60°C wurden die überschüssigen Sondenmoleküle von den Nitrocellulose-Filtern entfernt. Die getrockneten Filter wurden in einer mit einer Verstärkerfolie ausgestatteten Expositionskassette zusammen mit einem Röntgenfilm der Fa. Fuji Medical X-Ray (Tokyo, Japan) für 5-24 Std. bei -80°C exponiert.

2.16 RNA-Präparation (GTC-Methode)

Die Isolation von Total-RNA aus Gewebe beruht auf der GTC-Methode (Chomczynski und Sacchi 1987). Hierbei erfolgt die RNA-Isolation mit Hilfe von Guanidinthiocyanat (GTC). Für die Isolation wurden 0,5-1 g L4-Larven, bzw. 0,35-0,45 g Adulttiere als Ausgangsmaterial eingesetzt. Diese wurden unter flüssigem Stickstoff solange gemörsert, bis keine festen Gewebestandteile mehr sichtbar waren. Die pulverisierte Probe wurde in ein 15 ml Falcon überführt und in 4 ml Arbeitslösung resuspendiert. Nach der Zugabe von 1/10 Vol. 2 M NaAcetat (pH 4) und 1 Vol. PCI erfolgte eine PCI-Extraktion. Hierfür wurde die Lösung vorsichtig mehrmals invertiert und im Anschluss zentrifugiert (5 min, RT, 5000 x g). Der wässrige Überstand wurde in ein frisches Falcon überführt und die PCI-Extraktion solange wiederholt, bis keine Interphase mehr sichtbar war. Nach erfolgreicher PCI-Extraktion wurde noch eine CI-Extraktion angeschlossen, bevor die RNA mit Ethanol gefällt wurde. Die RNA-Fällung wurde wie in Kapitel 2.18 beschrieben durchgeführt. Die gefällte RNA wurde in 200-400 µl RNase-freiem Wasser gelöst und im Anschluss direkt einem DNase-Verdau (Kapitel 2.17) unterzogen, bevor sie bis zur weiteren Verwendung bei -80°C gelagert wurde.

2.17 DNase-Verdau

Zur Entfernung von DNA-Resten aus der unter Punkt 2.16 gewonnenen Total-RNA wurde der Ansatz mit DNase I verdaut. Hierzu erfolgte die Zugabe von 1/10 Vol. 10 x DNase I-Puffer sowie 2U DNase I der Fa. New England Biolabs (Ipswich, UK). Der Ansatz wurde für 30 min bei 37°C inkubiert. Im Anschluss erfolgte durch Zugabe von 1/10 Vol. 50 mM EDTA und Inkubation für 15 min bei 65°C die Inaktivierung der Dnase I. Die Lagerung der RNA erfolgte bis zur weiteren Verwendung bei -80°C.

2.18 Fällung von RNA

Wie die Fällung von DNA dient auch die Fällung von RNA zur Aufkonzentrierung der RNA sowie der Entfernung von Salzen. Die Fällung von RNA erfolgte im Gegensatz zur Fällung von DNA im sauren Milieu. Zur Fällung wurde das Volumen mit RNase-freiem Wasser auf wenigstens 90 µl erhöht. Dann erfolgte die Zugabe von 1/10 Vol. 3 M NaAcetat (pH 5,2) sowie 2,5 Vol. kaltem EtOH (abs.). Der Ansatz wurde durch Invertieren gemischt und für mindestens 45 min bei -20°C oder alternativ für mindestens 10 min auf Kohlendäureschnee inkubiert. Es folgten eine Zentrifugation (45-60 min, 4°C, 16100 x g), das Abziehen und Verwerfen des Überstandes und das Waschen des Pellets in 500 µl 70 % EtOH. Nach einer weiteren Zentrifugation (10 min, RT, 18600 x g) und Verwerfen des Überstandes wurde das Pellet in der Vakuumentrifuge „Eppendorf Concentrator 5301“ der Fa. Eppendorf (Hamburg,

Deutschland) getrocknet. Im Anschluss wurde die RNA in 20 µl RNase-freiem Wasser gelöst und bis zur weiteren Bearbeitung bei -80°C gelagert.

2.19 Poly-A⁺-Aufreinigung

Aus der Dnase-verdauten Total-RNA (Kapitel 2.17) wurde mit Hilfe eines modifizierten Protokolls des „NucleoTrap® mRNA Mini Kits“ der Fa. Macherey-Nagel (Düren, Deutschland) mRNA angereichert. Bei diesem Protokoll wurden ebenfalls die oligo-(dT) Latex-Beads genutzt. Allerdings wurde bei der Bindung der mRNA an die Latex-Beads auf den NucleoTrap®-Mikrofilter verzichtet. Die Total-RNA wurde mit RNase-freiem Wasser auf 250 µl aufgefüllt und mit 1 Vol. RM0-Puffer vermischt. Die Zugabe der Latex-Beads erfolgte in Abhängigkeit von der eingesetzten Menge an Total-RNA (15 µl oligo (dT) Latex-Beads/ 100 µg Total-RNA). Im Anschluss wurde die Lösung für 5 min bei 68°C und weitere 10 min bei RT inkubiert. Die Latex-Beads wurden durch Zentrifugation (15 sec, 2000 x g und 5 min, 11000 x g) pelletiert, der Überstand abgezogen und das Pellet gewaschen (1 x in 600 µl RM2-Puffer, 2 x in 500 µl RM3-Puffer). Nach jedem Resuspendieren des Pellets erfolgte eine Zentrifugation (15 sec, 2000 x g und 5 min, 11000 x g) und das Abziehen des Überstandes. Nach dem zweiten Waschschritt mit RM3-Puffer wurde das getrocknete Pellet in RNase-freiem Wasser (30 µl Wasser/ 15 µl eingesetzten oligo-(dT) Latex-Beads) gelöst und für 7 min bei 68°C inkubiert. Im Anschluss erfolgte eine Zentrifugation (4 min, RT, 11000 x g). Der Überstand enthielt nun die aufgereinigte mRNA und wurde in ein neues Reaktionsgefäß überführt. Die Qualität und die Konzentration der mRNA wurde mit Hilfe des Agilent 2100-Bioanalyzers (Kapitel 2.20) bestimmt. Die Lagerung der mRNA erfolgte bis zur weiteren Bearbeitung bei -80°C.

2.20 Qualitäts- und Konzentrationsbestimmung von RNA per BioAnalyzer

Zur Überprüfung der Qualität sowie der Konzentration der aufgearbeiteten Total-RNA oder mRNA wurde diese mit Hilfe des Agilent 2100-Bioanalyzers der Fa. Agilent (Santa Clara, USA) untersucht. Hierbei wurden die Total-RNA mit Hilfe des RNA 6000 Nano Kits analysiert. Für die Analyse der mRNA wurde sowohl das RNA 6000 Pico Kit wie auch das RNA 6000 Nano Kit genutzt. Alle Analysen wurden nach den Angaben des Herstellers durchgeführt. Hierbei musste beachtet werden, dass die RIN-Werte (Fleige und Pfaffl 2006; Schroeder et al. 2006) der Messungen in diesem Fall keine hohe Aussagekraft über die Qualität der RNA haben, da die 28S rRNA bei Dipteren zerfällt und somit kein realistisches Verhältnis von 28S rRNA zu 18S rRNA berechnet werden kann.

Des Weiteren wurde zur Bestimmung der Konzentration der Total-RNA der NanoDrop ND-1000 der Fa. PeqLab (Erlangen, Deutschland) eingesetzt.

2.21 cDNA-Synthese (Reverse Transkription)

Die cDNA-Synthese erfolgte unter Verwendung der „SuperScript™ III Reverse Transcriptase“ der Fa. Invitrogen (Carlsbad, USA) nach den Angaben des Herstellers. Hierbei wurden im Laufe der Arbeit sowohl Total-RNA (1-5 µg) wie auch mRNA (>100 ng) als Ausgangsmaterial in die cDNA-Synthese eingesetzt. Die cDNA-Synthese erfolgte in den meisten Fällen mit einem oligo-(dT)-Primer und ohne die Zugabe von „RNaseOUT™“.

2.22 „Rapid Amplification of cDNA ends“ (RACE)

Um vollständige Transkripte verschiedener Gene der SDR zu erhalten, wurden zwei verschiedene Protokolle der RACE verwendet. Beide Protokolle dienen hierbei sowohl zur Analyse der 5'UTR- wie auch der 3'UTR-Bereiche der verschiedenen Gene. Zum einen wurde hierfür das Protokoll „Targeted rapid amplification of cDNA ends (T-RACE)“ nach (Bower und Johnston 2010) leicht modifiziert durchgeführt. Hierbei wurde für die cDNA-Synthese die „SuperScript™ III-Reverse Transcriptase“ der Fa. Invitrogen (Carlsbad, USA) verwendet. Für die Durchführung der RT-PCR wurde ausserdem die GoTaq®-Polymerase der Fa. Promega (Fitchburg, USA) eingesetzt. Weiterhin wurde das „GeneRacer® Kit with SuperScript® III RT and TOPO TA Cloning® Kit for Sequencing“ der Fa. Invitrogen (Carlsbad, USA) verwendet. Bei der Synthese der cDNA wurde hierbei nach den Angaben des Herstellers verfahren. Die Lagerung der cDNA erfolgte bei -20°C.

2.23 Illumina HiSeq 2000

Die cDNAs der Larven von *C. thummi* ♀ und *C. thummi* ♂ wurden auf dem Illumina HiSeq 2000 der Fa. Illumina (San Diego, USA) sequenziert. Hierfür wurde RNA aus je 0,5-1 g Larven gewonnen. Die Präparation wurde nach GTC-Protokoll (siehe Kapitel 2.16) durchgeführt. Die Erstellung der cDNA-Bibliothek erfolgte nach dem Protokoll des TruSeq RNA-Kit der Fa. Illumina (San Diego, USA). Hierbei wurden 1,5 µg Total-RNA als Ausgangsmenge eingesetzt. Die Total-RNA wurde im Laufe des Protokolls einer zweifachen Poly-A⁺-Aufreinigung unterzogen. Pro Probe wurde für den Lauf eine Spur der Flowcell genutzt. Hierbei erfolgte ein 51 bp „single end“ bzw. ein 100 bp „paired end“ Lauf. Die Daten wurden im Anschluss mit Hilfe der „CLC Main Workbench 5“ der Fa. CLC bio (Katrinebjerg, Dänemark) getrimmed und für weitere Analysen genutzt.

3 Ergebnisse

3.1 Klonierung der SDR von *C. tentans*

Zur Klonierung von Bereichen aus der SDR in *C. tentans* wurden BAC-Bibliotheken von *C. tentans* mit „single copy“-Sonden aus SDR-Bereichen von *C. thummi* mit Hilfe der Koloniefilterhybridisierung (Kapitel 2.15.3) durchsucht. Hybridisierende Klone wurden angezüchtet, die BAC-DNA gereinigt und durch Restriktionsverdau charakterisiert. Anschließend erfolgte eine Teilsequenzierung der Klonintegrate mittels Primern, die an Vektorsequenzen binden. Die so erhaltenen Randsequenzen konnten zur Lokalisation der BAC-Klone im Contig SDR genutzt werden. Da zu Beginn der Arbeit keine Referenzsequenzen von *C. tentans* vorlagen, erfolgte die Einordnung der Randsequenzen anfänglich durch Sequenzvergleich zum bereits bestehenden Contig der SDR von *C. thummi* (Kraemer 1994; Funke 2006; Knies 2006; Berwanger 2007; Herold 2008). Einzelne Klone wurden im Folgenden mit Hilfe der FISH (Kapitel 2.15.1) an Polytänchromosomen der Spezies *C. thummi* und gegebenenfalls durch DNA-Sequenzierung näher analysiert. Die Sequenzen der BAC-Klone sowie einige Sequenzen, welche mit Hilfe der „Primer Walking“-Strategie generiert werden konnten, wurden zu einem finalen Contig der SDR von *C. tentans* mit einer Größe von etwa 87 kb zusammengefügt. Dieser Contig wurde bioinformatisch auf die bereits aus *C. thummi* und *C. piger* bekannten Gene sowie auf repetitive Elemente hin untersucht.

3.1.1 Screening der BAC-Bibliotheken von *C. tentans*

Zu Beginn der Arbeit lagen keine Sequenzinformationen der SDR von *C. tentans* vor. Allerdings bestand bereits ein Contig der SDR aus *C. thummi* mit einer Größe von etwa 200 kb. In Abbildung 3.1 ist eine Übersicht über den distalen Bereich dieser Region aus *C. thummi* sowie der darin enthaltenen Gene dargestellt.

Zum Screenen der BAC-Bibliotheken von *C. tentans* wurden Sonden aus unterschiedlichen Genen im Bereich von 100-200 kb der SDR von *C. thummi* verwendet. In diesem Abschnitt konnten durch bioinformatische Analysen bisher insgesamt 16 Gene vorhergesagt werden (Knies 2006; Berwanger 2007; Mack 2009; Funke 2011).

Um störende Signale, die durch repetitive Sequenzen hervorgerufen werden, auf den Autoradiogrammen der Koloniefilter möglichst zu vermeiden, wurden die BAC-Bibliotheken mehrmals nacheinander mit relativ kleinen „single copy“-Sonden aus Exonbereichen verschiedener Gene aus der SDR von *C. thummi* durchsucht. Alle in der Arbeit gefundenen und näher untersuchten BAC-Klone stammen aus den beiden BAC-Bibliotheken L13 und L16, die aus Zellkulturzellen von *C. tentans* hergestellt wurden. Hierbei stammen alle BAC-Klone der ersten drei Screenings aus der BAC-Bibliothek L13. Lediglich das vierte Screening wurde mit der BAC-Bibliothek L16 durchgeführt.

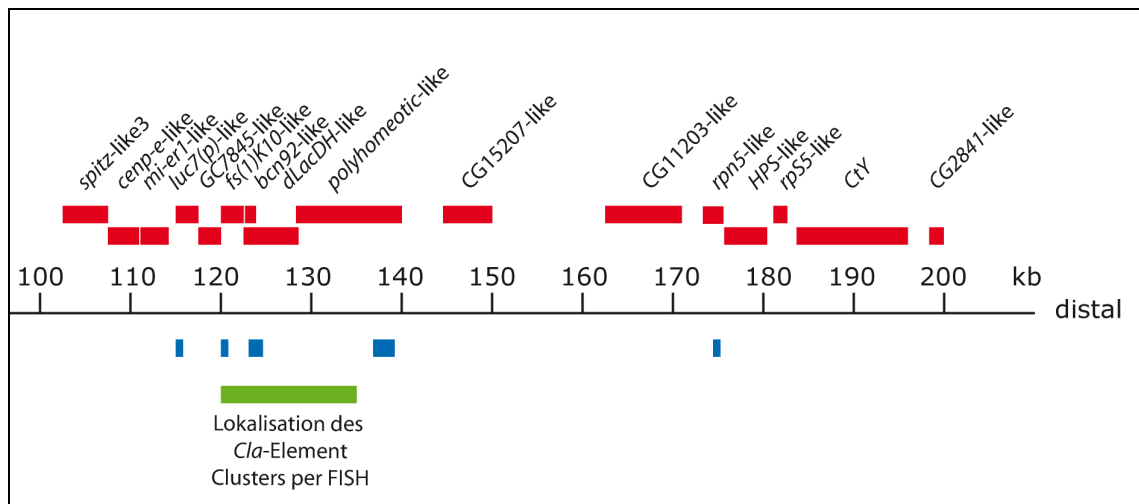


Abbildung 3.1: Übersicht über die Lage der Gene der SDR von *C. thummi*

Dargestellt ist der Contig SDR von *C. thummi* im Bereich von 100-200 kb. In Rot dargestellt sind die in diesem Bereich identifizierten Gene. Die obere Reihe entspricht den auf dem „Watson“-Strang liegenden Genen. Die untere Reihe repräsentiert hingegen die auf dem „Crick“-Strang befindlichen Gene. In Blau gezeigt sind die zum „Screening“ der BAC-Bibliotheken von *C. tentans* genutzten Sonden (Genbereiche: *luc7(p)*-like, *fs(1)K10*-like, *dLacDH*-like, *polyhomeotic*-like sowie *rpn5*-like). In Grün dargestellt ist der Bereich, auf den das *Cla*-Element Cluster in *C. thummi* per FISH eingegrenzt werden konnte (Herold 2008).

In Tabelle 3.1 sind die durch die Screenings gefundenen BAC-Klone inklusive ihrer Lage im Contig SDR von *C. tentans* angegeben. Weiterhin zeigt die Tabelle, mit welchen Sonden die entsprechenden BAC-Klone gefunden werden konnten. Die weiterführenden Analysen, mit deren Hilfe die Sequenz des Contigs SDR von *C. tentans* erstellt werden konnte, sind in der letzten Spalte der Tabelle angegeben. Ein großer Bereich des Contigs SDR stammt hierbei aus Sequenzen der vier BAC-Klone L13B4/1-1, L13A2/1-5, L13G3/2-3 sowie L13C2/1-1. Der Klon L13H4/1-1 wurde nicht näher analysiert, da die Lage des Klons in weiten Teilen mit den beiden analysierten BAC-Klonen L13A2/1-5 und L13G3/2-3 übereinstimmt. Der BAC-Klon L13G3/2-3 weist eine Größe von 45,2 kb auf. Allerdings bricht die Sequenzhomologie zum Contig SDR von *C. thummi* sowie zu den in diesem Bereich liegenden BAC-Klonen L13A2/1-5, L16D3-1 und L16F3-1 aus *C. tentans* an Position 44316 bp des Contigs SDR von *C. tentans* (Position 25934 bp im Klon) ab. Da die Position des Abbruchs der Sequenzhomologie zwischen dem BAC-Klon L13G3/2-3 und dem Contig SDR von *C. thummi* exakt mit dem distalen Ende des BAC-Klons L13A2/1-5 übereinstimmt und sich an dieser Position ebenfalls eine Restriktionsschnittstelle des Enzyms *Sau3AI* befindet, handelt es sich bei dem Klon L13G3/2-3 möglicherweise um ein „In vitro“-Ligat. Diese Vermutung wird durch die Daten der FISH (Kapitel 3.1.3) untermauert.

Insgesamt wurden Sequenzinformationen von 18 kb des Contigs SDR (Bereich 44-62 kb) mit Hilfe der „Primer Walking“-Strategie gewonnen. Das Verfahren wurde hier auf die BAC-Klone L16D3-1, L16F3-1 sowie L13J21 angewendet. Der 4,5 kb lange Bereich von 53-57,5 kb ist der einzige Abschnitt im gesamten Contig, der nicht durch

einen BAC-Klon abgedeckt wird. Dieser Bereich wurde mit Hilfe einer PCR mit genomischer DNA von *C. tentans* erschlossen. Die PCR wurde hierfür mit Primern aus dem distalen Randbereich des Klons L16D3-1 sowie dem proximalen Randbereich des Klons L13J21 durchgeführt. Das so erhaltene Produkt wurde kloniert und ebenfalls per „Primer Walking“-Strategie sequenziert.

Tabelle 3.1: Übersicht über die zum Screenen der BAC-Bibliotheken verwendeten Sonden sowie die Lage der im Contig SDR von *C. tentans* liegenden Klone

Die Tabelle zeigt die Lage der Klone im Contig SDR von *C. tentans*. Weiterhin gezeigt sind die zum Screening der BAC-Bibliotheken verwendeten Sonden sowie die im Folgenden angewendeten Verfahren zur Gewinnung der Sequenzinformation der Klone.

Position (bp)	Klon	verwendete Sonde(n)	durchgeführte Analysen
1-20287	L13B4/1-1	<i>luc7(p)</i> -like	"Shotgun"-Bibliothek
18382-44316	L13G3/2-3	<i>luc7(p)</i> -like, <i>fs(1)K10</i> -like, <i>dLacDH</i> -like, <i>polyhomeotic</i> -like	"Shotgun"-Bibliothek
21123-44316	L13A2/1-5	<i>luc7(p)</i> -like, <i>fs(1)K10</i> -like, <i>dLacDH</i> -like, <i>polyhomeotic</i> -like	"Shotgun"-Bibliothek
22119-45838	L13H4/1-1	<i>luc7(p)</i> -like, <i>fs(1)K10</i> -like, <i>dLacDH</i> -like, <i>polyhomeotic</i> -like	keine
27764-53260	L16D3-1	<i>polyhomeotic</i> -like	"Primerwalk" (44121-53260 bp)
29822-53117	L16F3-1	<i>polyhomeotic</i> -like	"Primerwalk" (44121-53117 bp)
53187-57482	pGem	keine (genomische PCR)	"Primerwalk" (53187-57482 bp)
57411-80438	L13J21	<i>rpn5</i> -like	"Primerwalk" (57411-62413 bp)
60178-86665	L13C2/1-1	<i>rpn5</i> -like	"Shotgun"-Bibliothek

Eine grafische Übersicht über den gesamten Contig SDR von *C. tentans* mit einer Gesamtlänge von 86665 bp ist in Abbildung 3.2 gezeigt.

Mit Hilfe von BLAST-Analysen und einem Vergleich der Sequenzdaten zwischen den Spezies *C. tentans* und *C. thummi* konnten 13 Gene dieses Bereichs bestätigt werden, welche bereits aus dem Contig SDR von *C. thummi* bekannt waren. Diese liegen im Contig SDR von *C. tentans* in derselben Anordnung, wie im Contig SDR von *C. thummi* vor. Weiterhin wurden einige repetitive Elemente im Contig SDR von *C. tentans* lokalisiert. In Abbildung 3.2 sind die lokalisierten Gene in zwei Reihen untereinander angeordnet. Hierbei entspricht die obere Reihe den auf dem „Watson“-Strang liegenden Genen, während die untere Reihe die auf dem „Crick“-Strang lokalisierten Gene repräsentiert.

Weiterhin ist in der Abbildung der Bereich dargestellt, auf den im Contig SDR von *C. thummi* das hemizygoten *Cla*-Element Cluster per FISH eingegrenzt werden konnte (Herold 2008). Dieser Bereich dient als Orientierungspunkt bei der Überprüfung der Lage der BAC-Klone aus *C. tentans* im Contig SDR von *C. thummi* mit Hilfe der FISH (Kapitel 3.1.2). Eine Auflistung der genauen Lage der Gene sowie der repetitiven Elemente im Contig SDR von *C. tentans* ist in Tabelle 3.2 gezeigt.

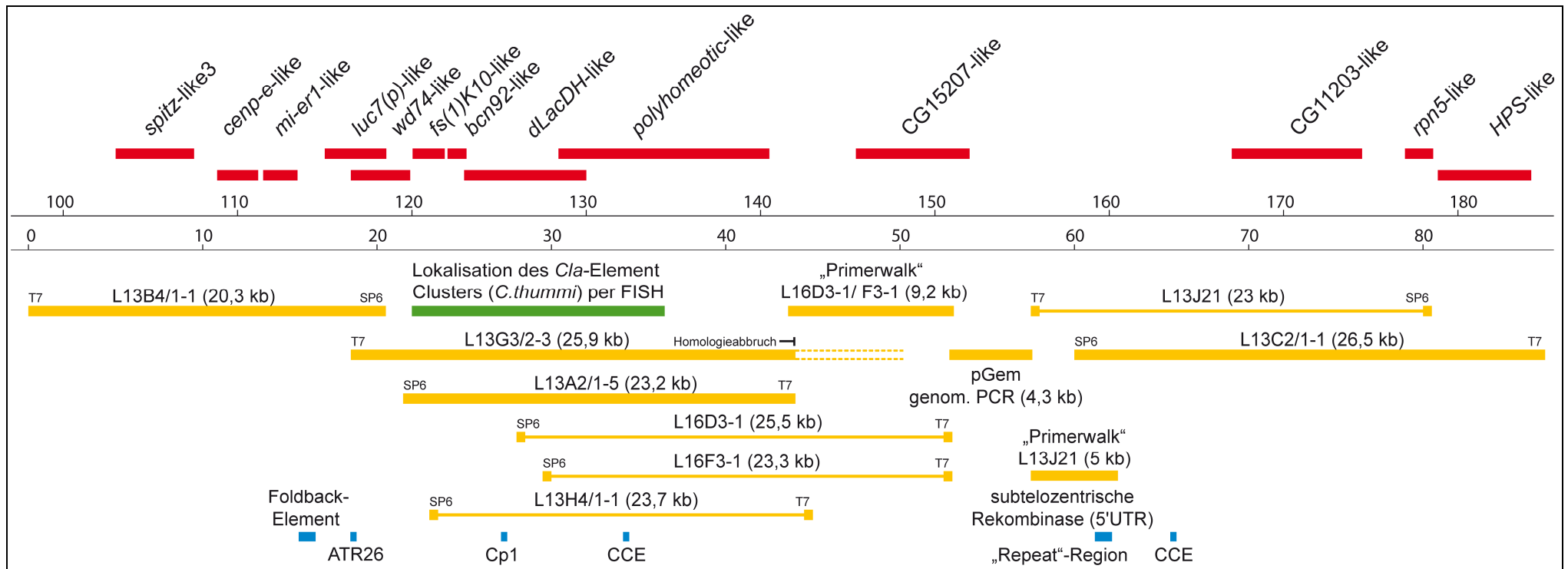


Abbildung 3.2: Übersicht über den Contig SDR von *C. tentans* inklusive der darin liegenden Gene sowie repetitiven Elemente

Dargestellt ist der gesamte Contig SDR von *C. tentans* mit einer Größe von 86665 bp. Die Skalierungen geben die Position der BAC-Klone im Contig SDR von *C. tentans* (untere Achse) sowie im Contig SDR von *C. thummi* (obere Achse) an. Unterhalb der Skalierungsachsen sind die in den Screenings der BAC-Bibliotheken gefundenen BAC-Klone dargestellt. Die per DNA-Sequenzierung vollständig analysierten BAC-Klone sowie die Bereiche, in denen das Verfahren der „Primer Walking“-Strategie angewendet wurde, sind als durchgehende gelbe Balken dargestellt. BAC-Klone, mit denen ein „Primer Walk“ durchgeführt wurde oder welche nicht näher analysiert wurden, sind als dünnere gelbe Linie gezeigt. Die Größen der entsprechenden BAC-Klone oder Abschnitte sind in Klammern hinter den Namen angegeben. Oberhalb der Skalierungsachsen sind die im Contig SDR identifizierten Gene dargestellt (rot). Diese konnten per BLAST-Analysen und Sequenzvergleich zum Contig SDR von *C. thummi* identifiziert werden. Die beiden Reihen geben hierbei die auf dem „Watson“-Strang liegenden Gene (obere Reihe) sowie die auf dem „Crick“-Strang liegenden Gene (untere Reihe) an. In Grün dargestellt ist der Bereich des Contigs SDR, auf den das männchenspezifische hemizygoten *Cla*-Element Cluster in der SDR von *C. thummi* per FISH eingegrenzt werden konnte (Herold 2008).

3.1.2 Repetitive Elemente

Thummi Foldback-Element (TFB1):

Hierbei handelt es sich um ein transposables Element, welches erstmals in der Intergenregion der Histone H1-H3 von *C. thummi* beschrieben wurde, jedoch auch im restlichen Genom zu finden ist (Hankeln und Schmidt 1990). Es zeichnet sich durch lange „Inverted repeats“ (IR) aus, welche von „Target site“-Duplikationen mit einer Länge von 9-10 Nukleotiden flankiert werden.

Im Bereich von 15811-16668 bp konnte eine Indel-Sequenz mit einer Länge von 858 bp identifiziert werden. Diese zeigt 20 bp lange IRs, welche von 4 bp langen „Direct repeats“ (DRs) flankiert werden. Hierbei könnte es sich um ein Foldback-Element handeln.

ATR26:

ATR26 ist ein tandem-repetitives Element, welches in *C. annularius* im Bereich des „intergenic spacer“ von rDNA-Genen gefunden wurde (GenBank: AJ279434.1; Koepf, Hankeln, Schmidt; Organization and evolution of the rDNA genes in *Chironomus*; unpubliziert).

Cp1:

Bei Cp1 handelt es sich um ein transposables Element der Gruppe der „short interspersed nuclear elements“ (SINEs). Cp1 wurde in *C. tentans* entdeckt (He et al. 1995) und weist eine Größe von 3694 bp auf. Es kommt an vielen Stellen des Genoms vor und ist meist von 10-14 bp langen „Target site“-Duplikationen umgeben.

Bei dem im Contig SDR von *C. tentans* gefundenen Bereich handelt es sich um einen 128 bp langen Abschnitt mit guter Homologie zu Cp1.

CCC-Element (CCE):

Die CCC-Elemente gehören ebenfalls zu der Klasse der transposablen Elemente und wurden erstmal in *C. thummi* beschrieben (Kraemer und Schmidt 1993). Das von C. Kraemer gefundene CCE besitzt eine Größe von 394 bp. Es zeichnet sich durch zwei 62-64 bp lange interne G/C-reiche DRs aus, welche dem Element seinen Namen geben. Des Weiteren besitzt das Element zwei terminale IRs mit einer Größe von 27 bp. Es enthält keinen ORF, so dass keine Transposase kodiert werden kann.

Die beiden in *C. tentans* gefundenen Bereiche weisen mit 59 bp und 66 bp kurze Abschnitte mit Homologien zum ersten IR und Teilen des dahinter liegenden DR des CCE aus *C. thummi* auf.

Subtelozentrische Rekombinase:

Der BLAST-Treffer des Bereichs 61334-62175 bp der SDR aus *C. tentans* zeigt starke Homologien zu einem 5463 bp langen DNA-Abschnitt des Telomerbereichs des telozentrischen vierten Chromosoms von *C. pallidivittatus* (Rosen et al. 2002). Dieser weist 155 bp lange zentromerspezifische „Repeats“ sowie einen ORF mit Ähnlichkeit zur Transposase auf. Die Homologie der Sequenz aus *C. tentans* besteht zur

„Repeat“-Region des DNA-Abschnitts aus *C. pallidivittatus* und liegt weit vor dem ORF der Transposase.

Tabelle 3.2: Lage der Gene sowie der repetitiven Elemente im Contig SDR von *C. tentans*

Die Tabelle zeigt die exakte Lage der einzelnen Gene sowie der repetitiven Elemente im Contig SDR von *C. tentans*. Sie ergänzt somit Abbildung 3.2.

Position (bp)	Gen	repetitives Element
6243-9251	<i>spitz-like 3</i>	
10907-13511	<i>cenp-e-like</i>	
13587-15665	<i>mi-er1-like</i>	
15811-16668		Foldback-Element
16896-20404	<i>luc7(p)-like</i>	
18707-18862		ATR26; tandem-repetitiv aus „intergenic spacer“ (IGS) der rRNA-Gene; <i>C. annularius</i>
18473-21700	<i>wd74-like</i>	
21880-23798	<i>fs(1)K10-like</i>	
23854-24837	<i>bcn92-like</i>	
27144-27271		Cp1 (SINE); <i>C. tentans</i>
24622-32331	<i>dLacDH-like</i>	
30485-42550	<i>polyhomeotic-like</i>	
34430-34488		CCC-Element; <i>C. thummi</i>
47563-54016	CG15207-like	
61334-62175		Subtelozentrische Rekombinase („Repeat“-Region vor ORF); <i>C. pallidivittatus</i>
65578-65643		CCC-Element; <i>C. thummi</i>
69330-76700	CG11203-like	
78929-80756	<i>rpn5-like</i>	
80785-86361	<i>hermanky pudlak-like</i>	
86262-86665	<i>rpS5-like</i> (Teilbereich)	

3.1.3 Lokalisation der BAC-Klone per FISH

Die durch die Screenings der BAC-Bibliotheken L13 und L16 von *C. tentans* gefundenen BAC-Klone wurden mit Hilfe des Verfahrens der FISH an Polytänchromosomen der Spezies *C. thummi* lokalisiert (*C. tentans* Larven standen leider nicht zur Verfügung). Hierdurch konnte die Lage der BAC-Klone im Contig SDR sowie die relative Lage der BAC-Klone untereinander überprüft und bestätigt werden. Zur besseren Lokalisation der BAC-Klone wurden die Hybridisierungen zusammen mit synthetisch hergestellten *Cla*-Elementen an Polytänchromosomen aus männlichen Larven der Spezies *C. thummi* durchgeführt. Diese zeigen in der SDR eine hemizygoten Bande in ♂. Weitere Signale der repetitiven *Cla*-Elemente sind auf allen Chromosomen

verteilt zu finden. Besonders starke Signale sind hierbei an den Zentromeren der Chromosomen und am Nukleolus erkennbar. Durch die Doppelhybridisierung mit *Cla*-Elementen kann die Lage der BAC-Klone in Relation zu der Position des hemizygoten *Cla*-Element Clusters bestimmt werden (vgl. hierzu Abbildung 3.2). Zusätzlich kann mittels FISH festgestellt werden, ob ein Klon repetitive Anteile enthält. In diesem Fall würden Anteile des Klons an verschiedenen Stellen der Polytänchromosomen hybridisieren. Die Möglichkeit eines „in vitro“-Ligats (hier bei Klon L13G3/2-3 vermutet; siehe Kapitel 3.1.1) kann mit Hilfe der FISH ebenfalls überprüft werden. Sollte es sich bei dem Integrat eines Klons um ein „in vitro“-Ligat handeln, so sind im Allgemeinen mindestens zwei Hybridisierungssignale an unterschiedlichen Positionen der Polytänchromosomen zu erwarten.

Alle verwendeten BAC-Klone wurden für die durchgeführten Analysen mit Biotin markiert und mittels Avidin-Rhodamin nachgewiesen. Die synthetischen *Cla*-Elemente wurden generell mit Digoxigenin markiert und mit Hilfe von Anti-Digoxigenin-Fluorescein detektiert. In Abbildung 3.3 ist die Lokalisation der verschiedenen BAC-Klone der SDR von *C. tentans* aufgezeigt. Hierbei hybridisieren alle untersuchten BAC-Klone der Spezies *C. tentans* innerhalb der SDR von *C. thummi*. Die Anordnung der Bilder (A-F) entspricht hierbei der Anordnung der Klone innerhalb der SDR von proximal nach distal. So hybridisiert der BAC-Klon L13B4/1-1 in proximaler Position zum hemizygoten *Cla*-Element Cluster innerhalb der SDR von *C. thummi* (Abbildung 3.3 A). Die Hybridisierungssignale der beiden BAC-Klone L13A2/1-5 (Abbildung 3.3 B) und L16D3-1 (Abbildung 3.3 E) zeigen Überlagerungen der Hybridisierungssignale mit dem Signal des hemizygoten *Cla*-Element Clusters. Der BAC-Klon L13C2/1-1 zeigt ein Hybridisierungssignal mit distaler Lokalisation zum hemizygoten *Cla*-Element Cluster in *C. thummi* (Abbildung 3.3 F). Die Ergebnisse der FISH bestätigen somit in allen Fällen die Lage der BAC-Klone im Contig SDR aus den Ergebnissen der Sequenzierungen (Abbildung 3.2).

Das Hybridisierungssignal des BAC-Klons L13G3/2-3 (Abbildung 3.3 C) überlagert sich ebenfalls mit dem Hybridisierungssignal des hemizygoten *Cla*-Element Clusters. Auch dies bestätigt die Lage des Klons im Contig SDR von *C. tentans*. Zusätzlich zum Hybridisierungssignal des BAC-Klons in der SDR von *C. thummi* taucht ein zweites Hybridisierungssignal auf einem anderen Chromosom auf (Abbildung 3.3 D). Zusammen mit den Ergebnissen der Sequenzierung des Klons (bestehend *Sau3A*-Restriktionsschnittstelle; Homologieabbruch zur bestehenden Sequenz des Contig SDR; Kapitel 3.1.1) kann gefolgert werden, dass es sich bei dem BAC-Klon L13G3/2-3 vermutlich um ein „In vitro“-Ligat handelt.

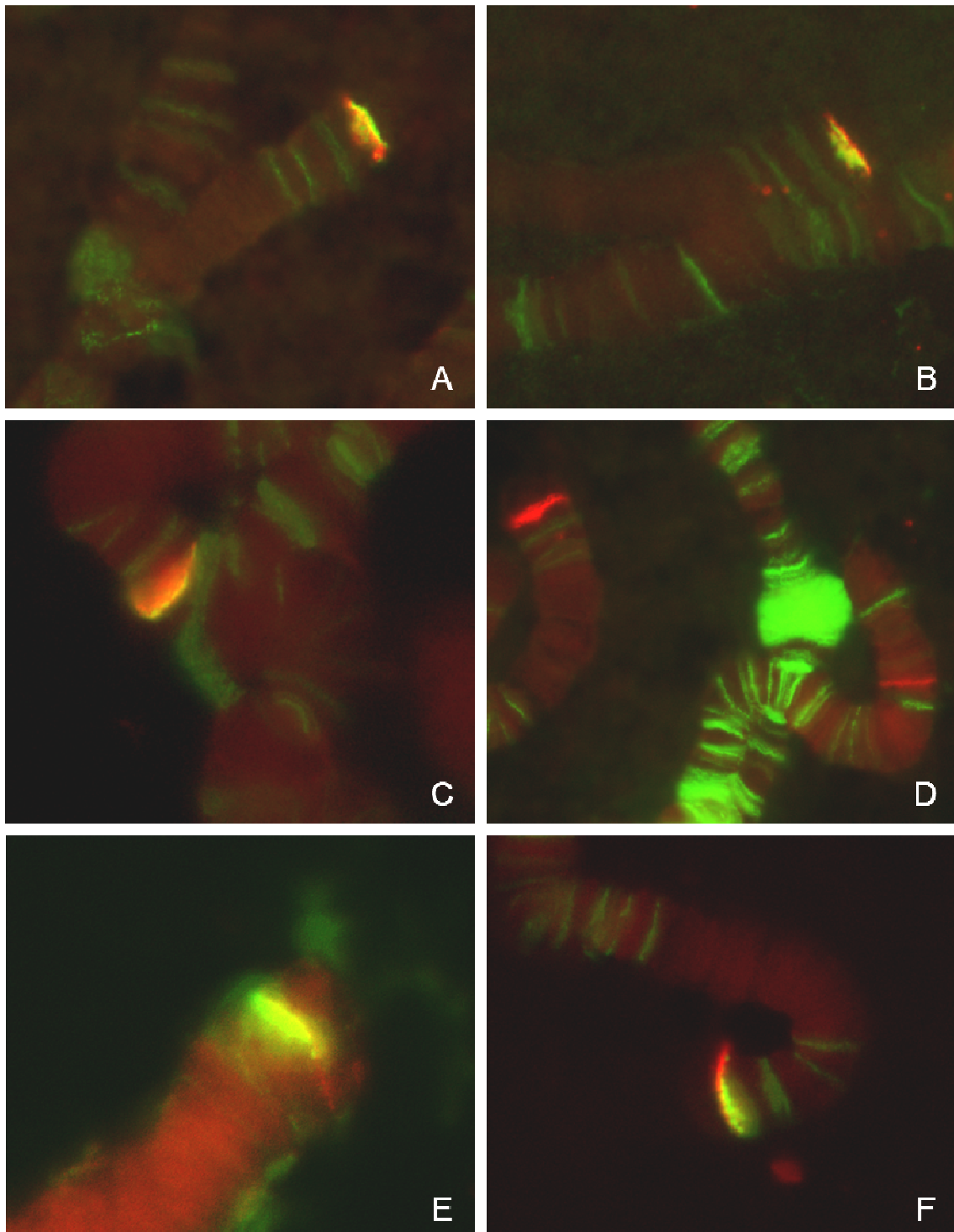


Abbildung 3.3: „In situ“-Hybridisierung der verschiedenen BAC-Klone aus *C. tentans* an Polytänchromosomen der Spezies *C. thummi*

Die Abbildung zeigt die Hybridisierungen der BAC-Klone aus *C. tentans* (rote Signale) in relativer Lage zum hemizygoten *Cla*-Element Cluster der SDR (grüne Signale von repetitiven *Cla*-Elementen) an Polytänchromosomen von *C. thummi* Männchen. **A:** BAC-Klon L13B4/1-1 mit proximalem Hybridisierungssignal (rot) zum hemizygoten *Cla*-Element Cluster (grün). **B:** BAC-Klon L13A2/1-5 mit überlagerndem Hybridisierungssignal (rot) zum hemizygoten *Cla*-Element Cluster (grün). **C:** BAC-Klon L13G3/2-3 mit überlagerndem Hybridisierungssignal (rot) zum hemizygoten *Cla*-Element Cluster (grün). **D:** Hybridisierung des BAC-Klons L13G3/2-3

an Polytänchromosomen von *C. thummi* Weibchen. Es sind zwei Hybridisierungssignale (rot) an zwei unterschiedlichen Chromosomen zu erkennen. Das Hybridisierungssignal im telomernahen Bereich des Chromosoms entspricht hierbei der Hybridisierung des Klon in der SDR (Region D3e1, Arm F). Das zweite Hybridisierungssignal liegt auf einem anderen Chromosom und deutet somit darauf hin, dass es sich bei dem BAC-Klon L13G3/2-3 um ein „in vitro“-Ligat handeln könnte. **E:** BAC-Klon L16D3-1 mit leicht überlagerndem Hybridisierungssignal (rot) zum hemizygoten *Cla*-Element Cluster (grün). **F:** BAC-Klon L13C2/1-1 mit distalem Hybridisierungssignal (rot) zum hemizygoten *Cla*-Element Cluster (grün).

3.1.4 Überprüfung der Organisation der SDR von *C. tentans* mit Hilfe von Southern-Analysen

Um zu überprüfen, ob die mit Hilfe der BAC-Klone sowie des „Primer Walks“ bestimmte Sequenz der SDR in *C. tentans* tatsächlich der genomischen Organisation entspricht, wurden genomische Southern-Analysen durchgeführt. Hierzu wurde zunächst mit Hilfe des Programms SeqBuilder aus dem „DNASTAR Lasergene®“- Software Paket eine Restriktionskarte der gesamten Region mit verschiedenen Enzymen erstellt. Hierbei wurde darauf geachtet, dass Restriktionsenzyme ausgewählt wurden, deren vorhergesagten Restriktionsfragmente in der SDR eine Größe von 1-10 kb besitzen, da diese gut elektrophoretisch analysiert werden können. Zur Überprüfung, ob die ermittelte Sequenz des Contigs SDR mit der genomischen Organisation dieser Region übereinstimmt, wurden genomische Southern-Blots von *C. tentans* erstellt. Im Anschluss wurden diese mit ausgewählten Sonden aus der SDR hybridisiert.

Zuvor waren bereits Blots restringierter BAC-Klone aus *C. tentans* mit den gleichen Sonden hybridisiert worden. Hierdurch konnte überprüft werden, ob die mittels Seqbuilder errechneten Fragmente mit den Signalgrößen auf den Autoradiogrammen nach Hybridisierung mit den entsprechenden Sonden übereinstimmen.

Nach Hybridisierung des genomischen Southern-Blots von *C. tentans* mit der cDNA-Sonde *mi-er1-like* (Position 13866-15050 bp im Contig SDR von *C. tentans*) ist zu erkennen, dass die auf dem Autoradiogramm zu verzeichnenden Signale nicht mit den erwarteten Fragmentgrößen übereinstimmen. Auf dem Autoradiogramm des genomischen Southern-Blots (Abbildung 3.4) ist zusätzlich zu den erwarteten Signalen (Tabelle rechts neben dem Autoradiogramm) in den Spuren der Restriktionsverdau *EcoRI*, *HindIII* und *NsiI* je ein weiteres starkes Hybridisierungssignal zu verzeichnen. Die zusätzlichen Hybridisierungssignale weisen hierbei eine Größe von etwa 5,2-5,6 kb auf. Auch in der Spur des Restriktionsverdau mit dem Enzym *AccI* ist nicht auszuschließen, dass es sich bei der zu erkennenden Bande um eine Doppelbande handeln könnte.

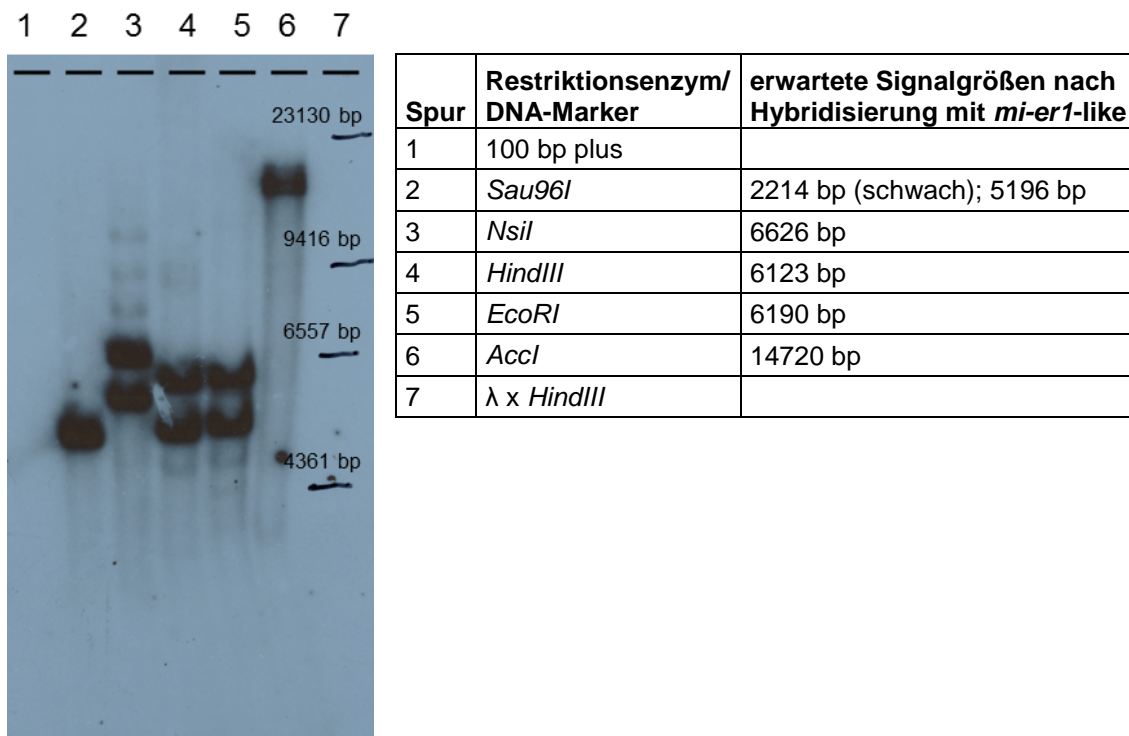


Abbildung 3.4: Autoradiogramm des genomischen Southern-Blots von *C. tentans* nach Hybridisierung mit der cDNA-Sonde *mi-er1-like*

Die Abbildung zeigt die Hybridisierungssignale nach Hybridisierung restringierter genomischer DNA von *C. tentans* mit der cDNA-Sonde *mi-er1-like* (Position 13866-15050 bp im Contig SDR von *C. tentans*). Die erwarteten Signalgrößen der verschiedenen Restriktionsenzyme nach Hybridisierung mit der Sonde *mi-er1-like* sind in der Tabelle rechts neben dem Autoradiogramm angegeben. Zu erkennen ist, dass in mehreren Spuren (*EcoRI*, *HindIII*, *NsiI*) zusätzlich zu den aus der SDR-Sequenz abgeleiteten Fragmenten je ein weiteres starkes Hybridisierungssignal zu verzeichnen ist. Die Größen der zusätzlichen Banden liegen bei etwa 5,2-5,6 kb. Expositionszeit ~110 Std.

Zur Überprüfung der genomischen Organisation der Region erfolgte ebenfalls eine Hybridisierung eines genomischen Southern-Blots von *C. tentans* mit der verwendeten cDNA-Sonde *luc7(p)-like* (Position 16992-17991 bp im Contig SDR von *C. tentans*). Hierbei ist zu beachten, dass in den Spuren der Restriktionsverdau mit *AccI*, *EcoRI*, *HindIII* sowie *NsiI* dieselben Hybridisierungssignale zu erwarten sind wie nach der Hybridisierung mit der Sonde *mi-er1-like* (siehe auch Abbildung 3.7).

Vergleicht man die in Abbildung 3.5 gezeigten Hybridisierungssignale des Autoradiogrammes mit den erwarteten Fragmentgrößen (Tabelle rechts neben dem Autoradiogramm) nach Hybridisierung mit der cDNA-Sonde *luc7(p)-like*, so zeigt sich ein ähnliches Muster wie nach Hybridisierung mit der cDNA-Sonde *mi-er1-like* (Abbildung 3.4). In den Spuren der Restriktionsverdau mit *EcoRI*, *HindIII* und *NsiI* ist auf dem Autoradiogramm neben den erwarteten Signalen je ein weiteres starkes Signal zu verzeichnen. Die Hybridisierungsmuster (Abbildung 3.4 und Abbildung 3.5) scheinen identisch zu sein. Die Größen der zusätzlichen Banden liegen, genau wie nach der Hybridisierung mit der cDNA-Sonde *mi-er1-like*, auch in diesem Fall im Bereich von 5,2-5,6 kb.

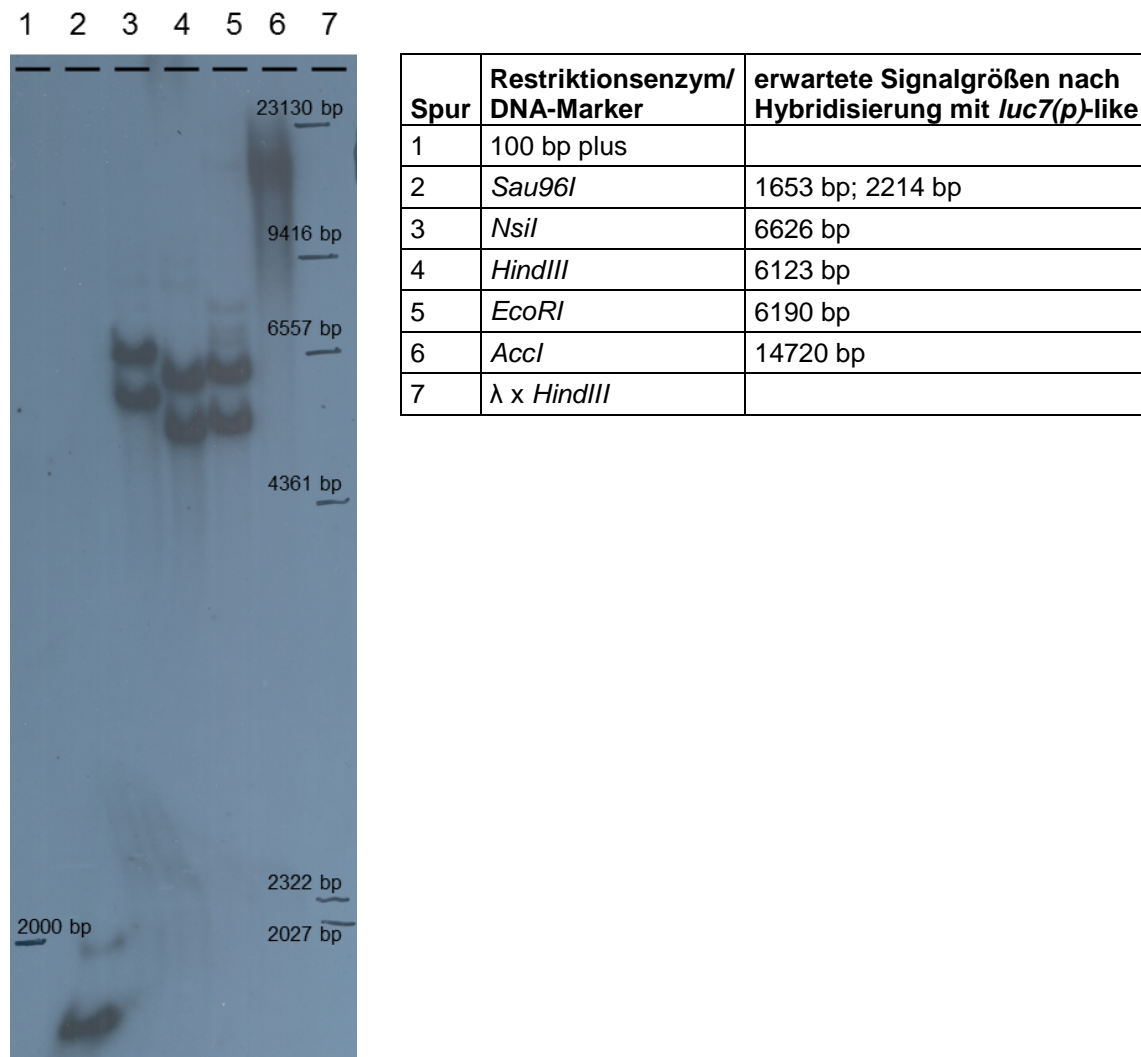


Abbildung 3.5: Autoradiogramm des genomischen Southern-Blots von *C. tentans* nach Hybridisierung mit der cDNA-Sonde *luc7(p)*-like

Die Abbildung zeigt die Hybridisierungssignale nach Hybridisierung restringierter genomischer DNA von *C. tentans* mit der cDNA-Sonde *luc7(p)*-like (Position 16992-17991 bp im Contig SDR von *C. tentans*). Die erwarteten Signalgrößen der verschiedenen Restriktionsenzyme nach Hybridisierung mit der Sonde *mi-er1*-like sind in der Tabelle rechts neben dem Autoradiogramm angegeben. Sie stimmen bis auf die Fragmente der *Sau96I*-Restriktion mit den Größen nach Hybridisierung mit der Sonde *mi-er1*-like (siehe Abbildung 3.4) überein. In den Spuren der Restriktionsverdau *EcoRI*, *HindIII* sowie *NsiI* ist zusätzlich zu den aus der SDR-Sequenz abgeleiteten Fragmenten je ein weiteres starkes Hybridisierungssignal zu verzeichnen. Die Größen der zusätzlichen Banden liegen bei etwa 5,2-5,6 kb. Expositionszeit ~110 Std.

Um zu analysieren, worauf die zusätzlichen Hybridisierungssignale auf den Autoradiogrammen der genomischen Southern-Blots zurückzuführen sind, wurden verschiedenen PCRs auf genomische DNA von *C. tentans* durchgeführt. Auf diese Weise ist es gelungen im Bereich 15811-16668 bp des Contigs SDR von *C. tentans* eine Indel-Sequenz mit einer Länge von 858 bp zu identifizieren. Diese findet sich im BAC-Klon L13B4/1-1. Mit genomischer DNA als Template sind per PCR sowohl eine Variante mit Insertion wie auch eine Variante ohne Insertion nachweisbar. Bei den

Bereichen 15815-15834 bp sowie 16649-16668 bp, den Randbereichen des Indels, handelt es sich um eine 20 bp lange IR-Struktur. Flankiert wird diese durch eine 4 bp lange DR-Struktur (CTAG; 15811-15814 bp sowie 16669-16672 bp). Die DR-Struktur an den Rändern des Indels spricht für eine „Target site“-Duplikation die bei vielen transposablen Ereignissen auftritt. Innerhalb des Indels konnte kein ORF identifiziert werden, der für eine mögliche Transposase kodiert. Das Vorkommen von „Direct repeats“ (DRs) und „Inverted repeats“ (IRs) sowie das Fehlen eines Transposasesgens weisen darauf hin, dass es sich bei dem untersuchten Indel-Bereich um ein Foldback-Element handeln könnte (Potter et al. 1980).

Der analysierte Indel-Bereich erklärt die zusätzlichen Hybridisierungssignale auf den genomischen Southern-Blots nach Hybridisierung mit den Sonden *mi-er1*-like (siehe Abbildung 3.4) sowie *luc7(p)*-like (siehe Abbildung 3.5).

Bei allen weiteren durchgeführten Southern-Hybridisierungen zeigen sich auf den Autoradiogrammen keine Abweichungen von den erwarteten Größen der Hybridisierungssignale.

In Abbildung 3.6 ist beispielhaft das Autoradiogramm der Hybridisierung eines genomischen Southern-Blots von *C. tentans* mit der Sonde *fs(1)K10*-like dargestellt.

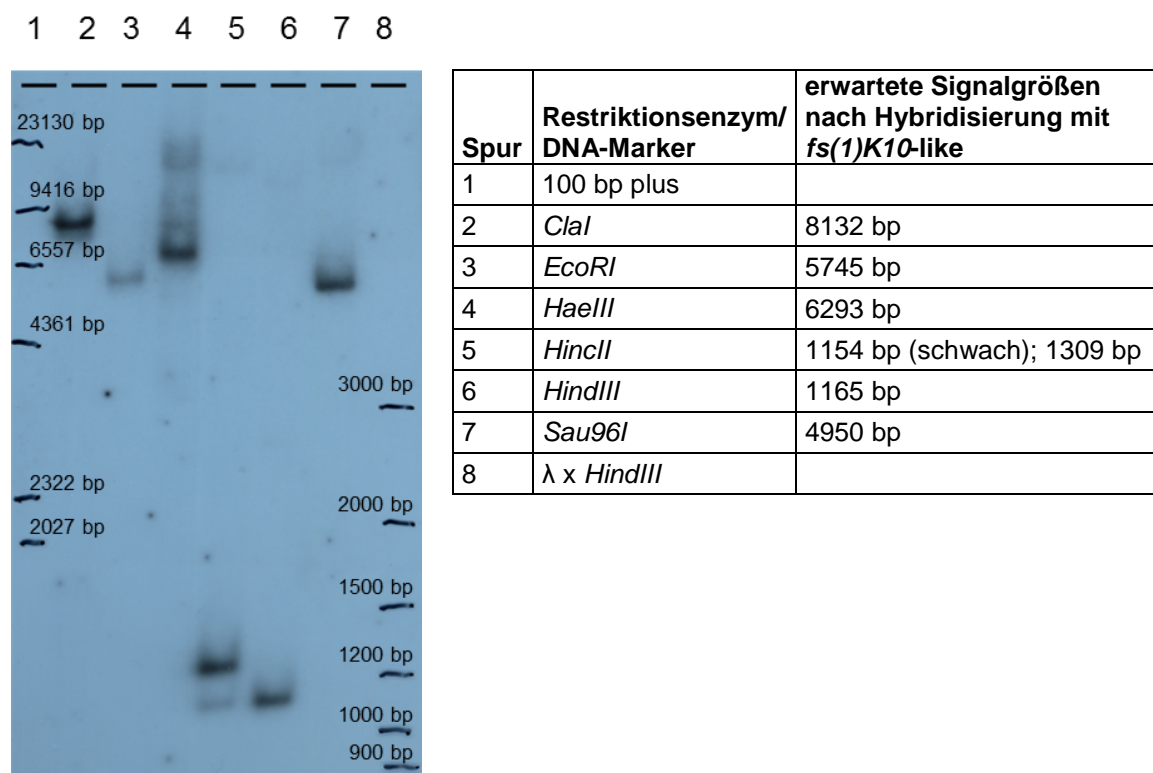


Abbildung 3.6: Autoradiogramm des genomischen Southern-Blots von *C. tentans* nach Hybridisierung mit der cDNA-Sonde *fs(1)K10*-like

Die Abbildung zeigt die auf dem Autoradiogramm erhaltenen Signale nach erfolgter Hybridisierung restringierter genomischer DNA von *C. tentans* mit der cDNA-Sonde *fs(1)K10*-like (Position 22495-22879 bp im Contig SDR von *C. tentans*). Die erwarteten Signalgrößen der verschiedenen Restriktionsenzyme nach Hybridisierung mit der Sonde *fs(1)K10*-like sind in der Tabelle rechts neben dem Autoradiogramm angegeben. Hierbei ist zu erkennen, dass in allen Spuren der Restriktionsverdau die erwarteten Signale zu verzeichnen sind. Expositionszeit ~120 Std.

Die auf dem Autoradiogramm sichtbaren Hybridisierungssignale stimmen mit den erwarteten Fragmentgrößen (Tabelle rechts neben dem Autoradiogramm) nach Hybridisierung mit der cDNA-Sonde *fs(1)K10-like* überein. In *C. tentans* konnte somit mit Hilfe der genomischen Southern-Analysen kein Hinweis auf eine mögliche Duplikation im Bereich des Gens *fs(1)K10-like* gefunden werden, wie diese in *C. thummi* ♂ und *C. piger* ♂ zu finden ist (Stiehl 2009).

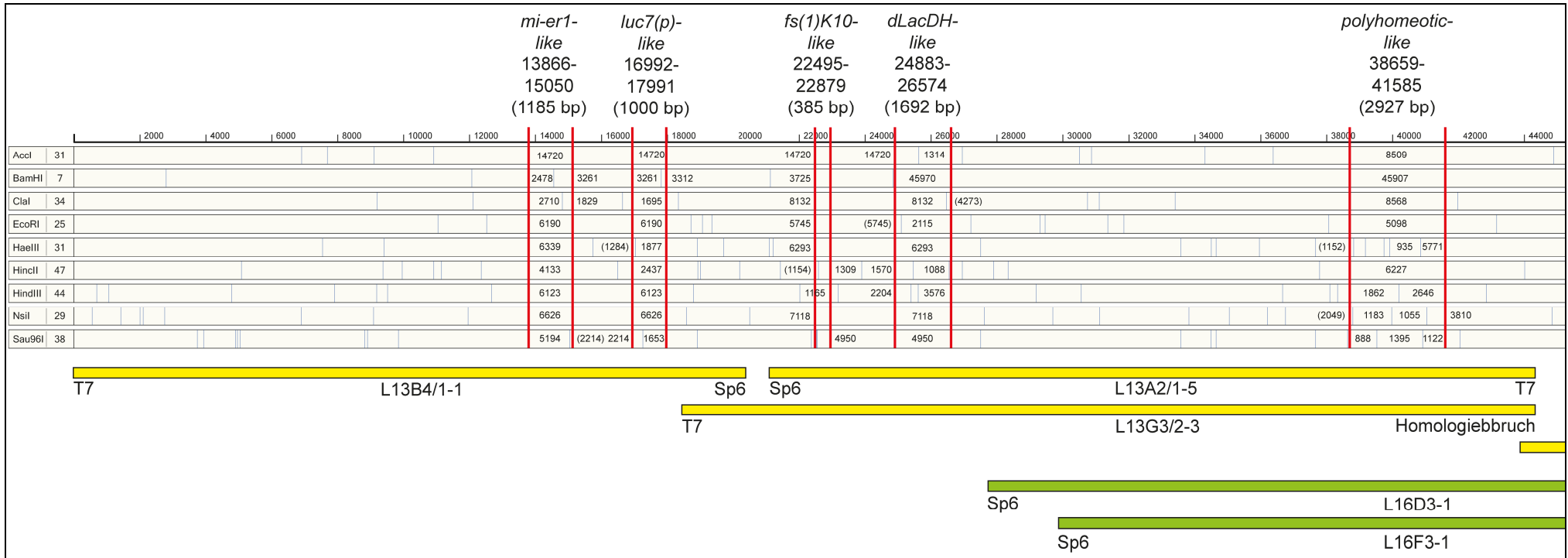


Abbildung 3.7: Proximaler Teil des Contigs SDR von *C. tentans* inklusive der dort lokalisierten BAC-Klone sowie der für die Southern-Analysen verwendeten Sonden

Die Abbildung zeigt den proximalen Teil der SDR von *C. tentans* inklusive der darin liegenden BAC-Klone. Die per „Shotgun“-Verfahren analysierten Klone sind gelb dargestellt. Klone, mit denen ein „Primerwalk“ durchgeführt wurde, sind in grün gezeigt. Die für die Southern-Analysen verwendeten Sonden sind über der Restriktionskarte inklusive ihrer Position im Contig SDR von *C. tentans* sowie ihrer Länge angegeben. Die Bereiche der Sonden sind durch rote Balken zu beiden Seiten hin begrenzt. Die auf den Autoradiogrammen der genomischen Southern-Blots erwarteten Signalgrößen sind im Bereich der Sonden in den jeweiligen Spuren der Restriktionsenzyme angegeben.

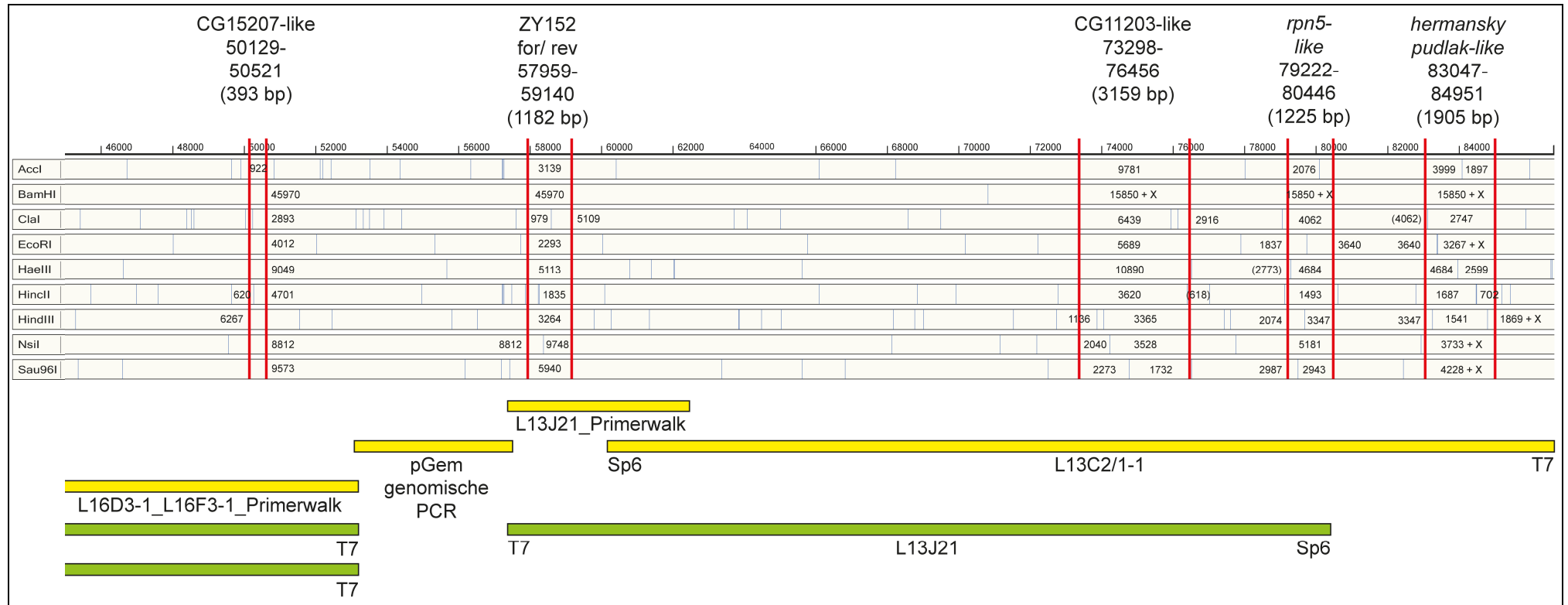


Abbildung 3.8: Distaler Teil des Contigs SDR von *C. tentans* inklusive der dort lokalisierten BAC-Klone sowie der für die Southern-Analysen verwendeten Sonden

Die Abbildung zeigt den distalen Teil der SDR von *C. tentans* inklusive der darin liegenden BAC-Klone sowie des pGem-Klons des genomischen PCR-Fragments. Der per „Shotgun“-Verfahren analysierte BAC-Klon L13C2/1-1 sowie die per „Primerwalk“ erschlossenen Bereiche sind gelb dargestellt. Klone, mit denen ein „Primerwalk“ durchgeführt wurde, sind in grün gezeigt. Die für die Southern-Analysen verwendeten Sonden sind über der Restriktionskarte inklusive ihrer Position im Contig SDR von *C. tentans* sowie ihrer Länge angegeben. Die Bereiche der Sonden sind durch rote Balken zu beiden Seiten hin begrenzt. Die auf den Autoradiogrammen der genomischen Southern-Blots erwarteten Signalgrößen sind im Bereich der Sonden in den jeweiligen Spuren der Restriktionsenzyme angegeben.

3.2 Analyse der Struktur der in der SDR befindlichen Gene per RT-PCR und RACE sowie der Hochdurchsatz-RNA-Sequenzierung

Innerhalb der SDR von *C. tentans* konnten im Laufe der Arbeit mit Hilfe von Genvorhersageprogrammen sowie durch den Vergleich zu den bekannten Genen aus *C. thummi* (Blumensatt 2009; Mack 2009; Funke 2011) insgesamt 14 Gene identifiziert werden. Bei dem am Weitesten proximal liegenden Gen handelt es sich um das Gen *spitz-like3*. Das am Weitesten distal liegende Gen ist *rpS5-like*. Von diesem Gen liegt im Contig SDR von *C. tentans* nur der proximale Teil als Sequenz vor. Die restliche Sequenz des Gens ist in *C. tentans* unbekannt, da der Contig innerhalb des Gens *rpS5-like* endet.

Im Contig SDR von *C. thummi* sind im Bereich von 100-200 kb bislang 16 Gene per Genvorhersage identifiziert worden. Bei insgesamt 11 dieser Gene wurden im Laufe dieser Arbeit die vorhergesagten Genstrukturen per RT-PCR sowie RACE überprüft. Die Überprüfung erfolgte mit Primerpaaren, deren RT-PCR-Produkte die vorhergesagten Introns überspannen. Somit konnten mit Hilfe der RT-PCR die vorhergesagten Intron-Exon-Grenzen bestätigt oder widerlegt werden. Die Durchführung der RT-PCR erfolgte mit cDNA der Spezies *C. thummi*, *C. piger* und *C. tentans* als Template. Das Hauptaugenmerk wurde hierbei auf die Spezies *C. thummi* gelegt, da für die cDNA-Synthese aus *C. tentans* nur Zellkulturen zur Verfügung standen, die keine Unterscheidung zwischen männchen- und weibchenspezifischen Produkten erlaubt. Weiterhin kann bei der cDNA aus *C. tentans*-Zellkultur keine Unterscheidung zwischen larvalem und adultem Gewebe erfolgen. Somit boten sich bei *C. thummi* umfassendere Untersuchungsmöglichkeiten an. Zum besseren Vergleich von möglichen Unterschieden in den Transkripten der verschiedenen Gene erfolgte zusätzlich die Durchführung der RT-PCR an cDNAs der Spezies *C. piger* sowie der aus Zellkultur stammenden cDNA von *C. tentans*.

Bei der detaillierteren Untersuchung der 5'UTR- und 3'UTR-Bereiche der Gene per RACE wurde die Untersuchung auf die cDNA aus larvalem Gewebe (L4-Stadium) von männlichen und weiblichen Tieren der Spezies *C. thummi* beschränkt.

Die Untersuchung der Intron-Exon-Strukturen der Gene im Bereich der SDR per RT-PCR in den Spezies *C. thummi* und *C. piger* wies mit den verwendeten Primerpaaren keine Unterschiede in der Größe der PCR-Produkte auf. In nahezu allen Genen zeigen sich bei den verwendeten cDNAs die erwarteten Produktgrößen. Es konnten keine Unterschiede zwischen den beiden Spezies *C. thummi* und *C. piger* sowie den beiden untersuchten Stadien (larval und adult) festgestellt werden. In Abbildung 3.9 ist exemplarisch eine gelelektrophoretische Auftrennung der RT-PCR-Produkte gezeigt. Hierbei handelt es sich um das untersuchte Gen *polyhomeotic-like*. Die erwartete Produktgröße liegt bei 850 bp. Wie auf dem Gelbild zu erkennen ist, entspricht die erwartete Größe von 850 bp den Größen der Banden auf dem Gel.

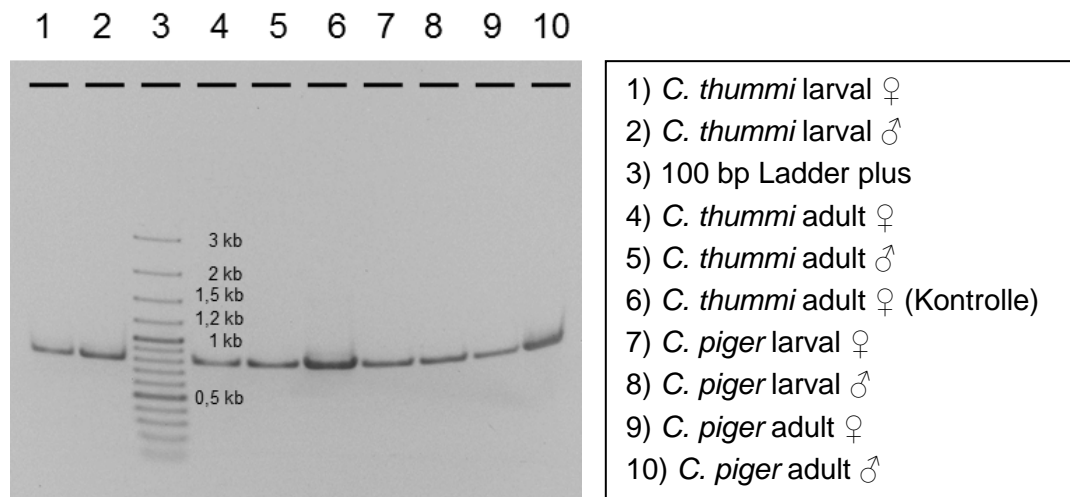


Abbildung 3.9: Gelbild der RT-PCR mit dem Primerpaar *polyhomeotic_3for/3rev*

Die Abbildung zeigt in allen Spuren des Gels die erwarteten Fragmente nach RT-PCR mit einer Größe von 850 bp. Auf dem Gel sind keine Unterschiede in den Produkten der RT-PCR zwischen den Spezies *C. thummi* und *C. piger* oder den verwendeten Stadien (larval und adult) zu erkennen.

Im Laufe der Arbeit ist es gelungen, von den meisten der in der SDR liegenden Gene aus *C. thummi* Sequenzdaten des gesamten transkribierten Bereichs inklusive der 5'UTR- und 3'UTR-Bereiche zu erhalten. In einigen Fällen ist es nicht gelungen, in beiden cDNA-Proben (larvaler RNA aus *C. thummi* ♀ und *C. thummi* ♂) ein Produkt nachzuweisen. Am wenigsten erfolgreich waren hierbei die Versuche, die beiden UTR-Bereiche der Gene CG15207-like und CG11203-like zu erlangen. Wie sich mit Hilfe der Hochdurchsatzsequenzierung zeigte, liegt dies vermutlich an einem sehr niedrigen Expressionslevel der beiden Gene im analysierten Stadium.

Im Folgenden soll nun detailliert auf die einzelnen Gene eingegangen werden. Die Reihenfolge, in welcher die Gene behandelt werden, entspricht hierbei der Anordnung der Gene im Contig SDR (von proximal nach distal). Ein Vergleich der Sequenzidentitäten zwischen den beiden Arten *C. thummi* und *C. tentans* auf genomischer Ebene, auf Ebene der prozessierten RNA, im proteinkodierenden Bereich der Gene sowie auf Polypeptidebene ist in Tabelle 3.3 am Ende des Kapitels aufgelistet. Die Daten stammen aus einem Sequenzvergleich der entsprechenden Bereiche aus *C. thummi* zum genomischen Contig von *C. tentans*. Der Sequenzvergleich wurde mit Hilfe des Programms MegAlign aus dem „DNASTAR Lasergene®“- Software Paket der Fa. Lasergene (Madison, USA) durchgeführt. Große Bereiche der Transkripte aus *C. tentans* sind zusätzlich mit Hilfe der RT-PCR überprüft worden. Die Sequenzdaten der RT-PCR bestätigen hierbei die im Sequenzvergleich angegebenen Intron-Exon-Strukturen der Gene und untermauern somit die angegebenen Sequenzidentitäten der einzelnen Gene.

mi-er1-like:

Die genomische Lage sowie die Struktur des Transkripts *mi-er1-like* im Contig SDR von *C. thummi* ist in Abbildung 3.10 gezeigt. Der transkribierte Bereich des Gens *mi-er1-like* erstreckt sich im Contig SDR von *C. thummi* auf dem „Crick“-Strang über einen genomischen Bereich von ≥ 1974 bp (111177-113150 bp). Die Länge des prozessierten Transkripts beträgt in *C. thummi* ♀ und *C. thummi* ♂ ≥ 1813 bp. Der transkribierte Bereich erstreckt sich über drei Exons, welche durch zwei Introns mit Längen von 100 bp und 61 bp voneinander getrennt werden. Der proteinkodierende Teil des Gens erstreckt sich über die Exons 2 und 3. Er weist eine Länge von 1479 Basen auf. Dies entspricht einer Proteinlänge von 493 AS.

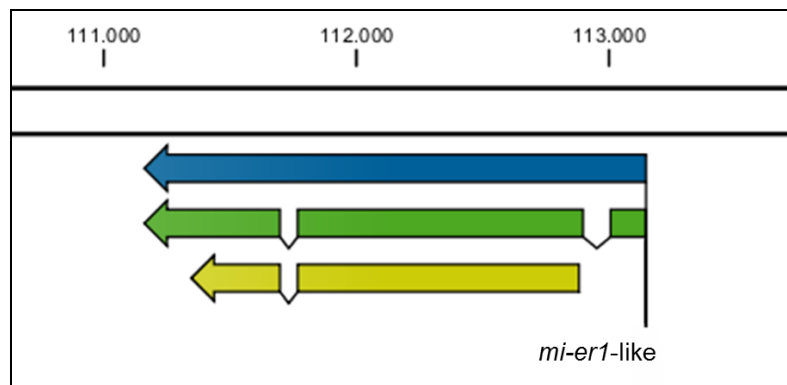


Abbildung 3.10: Lage und Struktur des Gens *mi-er1-like* im Contig SDR von *C. thummi*

Dargestellt sind der genomische Abschnitt des transkribierten Bereichs (blau), das prozessierte Transkript des Gens (grün) inklusive der beiden Introns sowie der proteinkodierende Teil (gelb). Die genauen Größen der einzelnen Bereiche sind dem Text zu entnehmen.

Die Überprüfung der abgeleiteten AS-Sequenz auf konservierte Domänen mit den beiden Programmen SMART und InterProScan lieferte zwei signifikante Treffer. Im Bereich von 344-393 AS besteht ein Treffer zu einer SANT-Domäne. Im Bereich von 238-337 AS liegt eine ELM2-Domäne. In Abbildung 3.11 sind die Domänen mit Hilfe des Programms InterProScan dargestellt. Weiterhin konnte mit Hilfe des Programms MotifScan ein mögliches Kern-Lokalisierungs-Signal (NLS) im Bereich von 121-138 AS gefunden werden. Der Bereich der ersten 100 AS weist zusätzlich eine hohe Anzahl der sauren AS Asparagin- und Glutaminsäure (D=15 und E=20) auf.

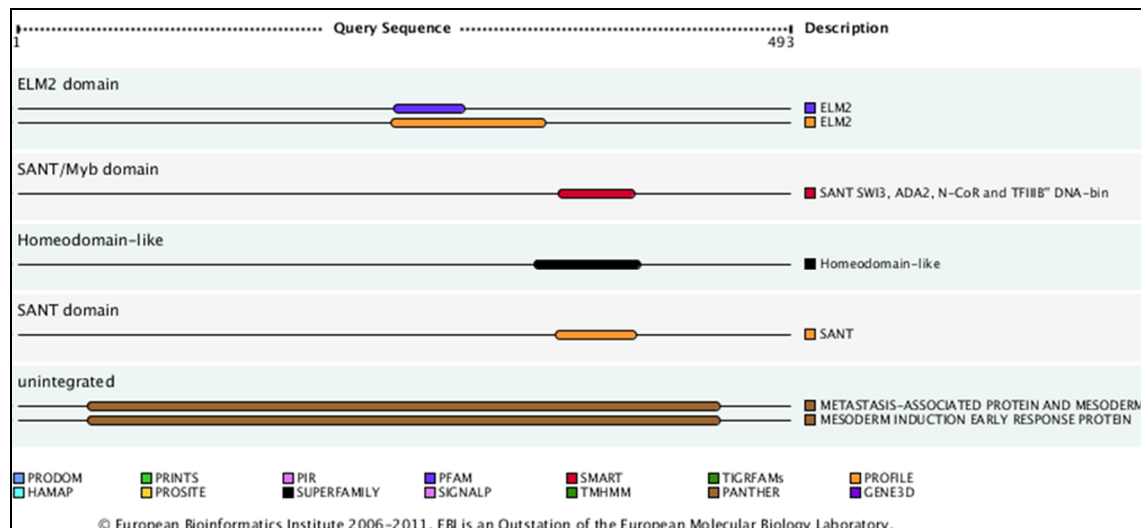


Abbildung 3.11: Domänen im Protein MI-ER1-like nach Suche mit InterProScan

Gezeigt sind die beiden Domänen im Protein MI-ER1-like. Die unterschiedlichen Farben der Balken zeigen an, in welcher Datenbank die entsprechende Domäne gefunden wurde. Im Protein MI-ER1-like konnten die beiden Domänen ELM2 sowie SANT in den Datenbanken PFAM, PROFILE, SMART, SUPERFAMILY und PANTHER lokalisiert werden.

luc7(p)-like:

Der transkribierte Bereich des Gens *luc7(p)-like* erstreckt sich im Contig SDR von *C. thummi* auf dem „Watson“-Strang über einen genomischen Bereich von ≥ 3839 bp (114983-118821 bp). Per RT-PCR konnten für das Gen fünf Exons nachgewiesen werden, welche durch Introns mit Längen von 117 bp, 57 bp, 64 bp und 1808 bp unterbrochen werden. Das prozessierte Transkript weist eine Gesamtlänge von mindestens 1793 bp auf. Der proteinkodierende Bereich des Gens *luc7(p)-like* erstreckt sich über die Exons 1-4 und besitzt eine Länge von 957 Basen, was einer Länge von 319 AS entspricht.

Zusätzlich konnte mit Hilfe der Hochdurchsatzsequenzierung in *C. thummi* ♀ ein zusätzliches, offensichtlich differentiell gespleißtes Intron identifiziert werden. Das Intron weist die typischen Grenzen (GT..AG) auf. In zwei „Reads“ fehlt ein Bereich von 65 bp (115690-115754 bp) in Exon 2, die restlichen „Reads“ stimmen zu 100 % mit dem vorhandenen Contig SDR von *C. thummi* überein. Durch das differentiell gespleißte Intron verschiebt sich der Leserahmen, und der proteinkodierende Teil des Gens wird durch ein früheres Stopcodon verkürzt. Die Länge des proteinkodierenden Bereichs liegt im verkürzten Transkript bei 555 Basen, was einer Polypeptidlänge von 185 AS entspricht. Dem verkürzten Polypeptid fehlt somit der gesamte C-terminale Abschnitt.

In Abbildung 3.12 sind genomische Lage sowie die Struktur des Transkripts *luc7(p)-like* im Contig SDR von *C. thummi* gezeigt.

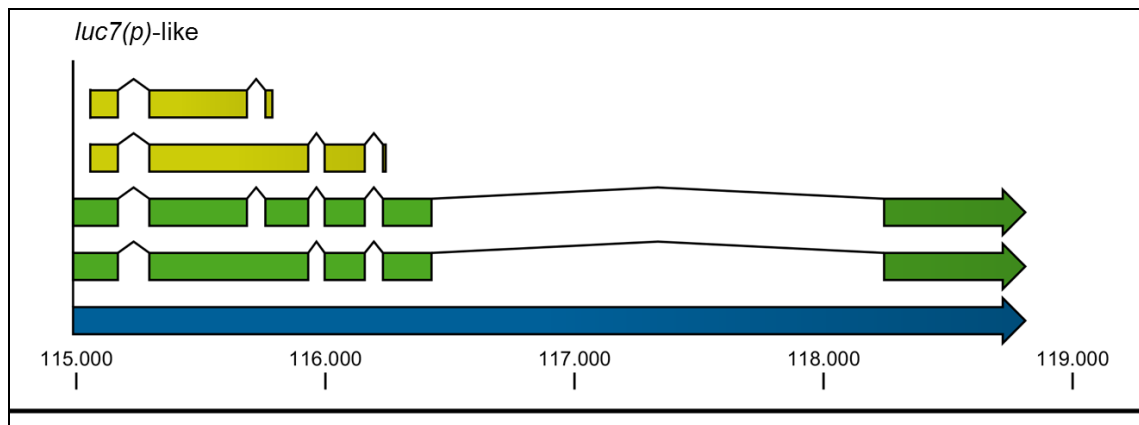


Abbildung 3.12: Lage und Orientierung des Gens *luc7(p)*-like im Contig SDR von *C. thummi*

Dargestellt sind der genomische Bereich der prä-mRNA (blau), die prozessierten Transkripte des Gens (grün) inklusive der vier Intronen sowie der proteinkodierende Teil (gelb). Die genauen Größen der einzelnen Bereiche sind dem Text zu entnehmen. Des Weiteren wird in *C. thummi* ♀ durch das differenziell gespleißte Intron im eigentlichen Exon 2 der proteinkodierende Teil des Gens durch ein früheres Stopcodon verkürzt.

Die Suche nach konservierten Domänen mit den beiden Programmen SMART und InterProScan ergab einen Treffer im Bereich von 2-254 AS zur Luc7-Proteinfamilie. Der Treffer des Programms InterProScan ist in Abbildung 3.13 gezeigt.

Weiterhin konnte im C-terminalen Bereich 7x das AS-Motiv RS und 7x das Motiv SR gefunden werden. Hierbei weisen die letzten 80 AS einen hohen Anteil (49 AS) an Arginin und Serin auf. Das in der Literatur ebenfalls für den C-terminalen Bereich beschriebene AS-Motiv RE konnte in *C. thummi* nicht bestätigt werden. Ebenfalls nur mit schwacher Homologie bestätigt werden konnten die in der Literatur angegebenen Zink-Finger-Motive im N-terminalen Bereich des abgeleiteten Polypeptids. Sie liegen in den Bereichen der AS 31-48 sowie 189-216. Ein Kern-Lokalisierungs-Signal (NLS) konnte mit Hilfe der Programme MotifScan und ProSite im Bereich der AS 260-274 lokalisiert werden.

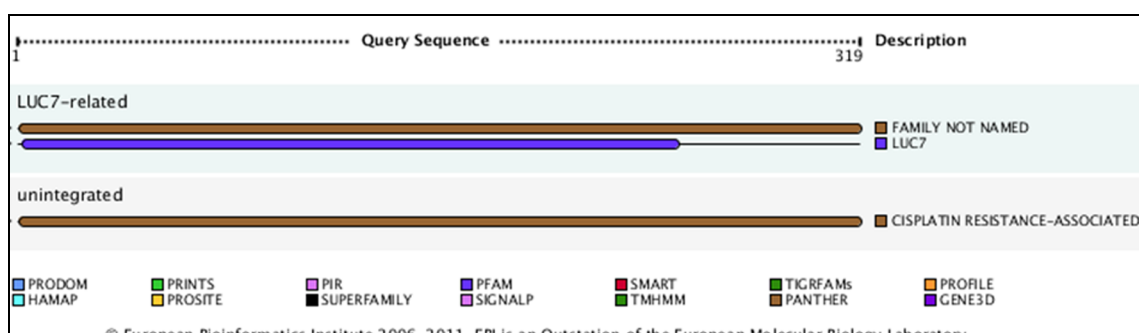


Abbildung 3.13: Domänen im Protein *Luc7(p)*-like nach Suche mit InterProScan

Die unterschiedlichen Farben der Balken zeigen, in welchen Datenbanken die entsprechenden Domänen gefunden wurden. In diesem Fall erfolgten die Treffer in den Datenbanken PFAM und PANTHER. Bei dem Treffer handelt es sich um einen Treffer zur Luc7-Proteinfamilie.

wd74-like:

Der transkribierte Bereich des Gens *wd74-like* ist im Contig SDR von *C. thummi* im Bereich von 116547-120316 bp auf dem „Crick“-Strang lokalisiert und erstreckt sich über einen genomischen Bereich von ≥ 3770 bp. Das prozessierte Transkript des Gens besitzt in *C. thummi* eine Gesamtlänge von ≥ 2185 bp und erstreckt sich über vier Exons. Die Introns des Gens sind 61 bp, 177 bp sowie 552 bp lang. Der proteinkodierende Teil des Gens reicht von Exon 1 bis zu Exon 4 und zeigt keine Unterschiede zwischen *C. thummi* ♀ und *C. thummi* ♂. Er besitzt eine Länge von 1122 Basen, was einer Länge von 374 AS entspricht. Der 3'UTR-Bereich des Gens *wd74-like* überschneidet sich mit dem 3'UTR-Bereich des Gens *luc7(p)-like*, welches auf dem „Watson“-Strang bis zur Position 118894 bp reicht. In *C. thummi* ♂ konnte im 3'UTR-Bereich des Gens *wd74-like* mit Hilfe von RACE-Produkten ein genomischer Bereich nachgewiesen werden, welcher durch die RACE-Produkte überspannt wird. Hierbei handelt es sich um den 795 bp langen Abschnitt (116822-117616 bp) des Contigs SDR von *C. thummi*. An den Rändern des durch die RACE-Produkte überspannten Bereichs zeigen sich keine typischen Intron-Exon-Strukturen. Analysen des Bereichs zeigen terminale 232 bp lange IRs (116822-117040 bp/ 117378-117607 bp) sowie hierzu flankierende 9 bp lange DRs (116813-116821 bp/ 117608-117616 bp). Ein Vergleich mit vorhandenen Sequenzdaten zeigt, dass es sich bei dem analysierten Bereich um die Indel-Sequenz von I116 handelt (Funke 2011). In Abbildung 3.14 sind die genomische Lage sowie die Struktur des Transkripts *wd74-like* im Contig SDR von *C. thummi* gezeigt.

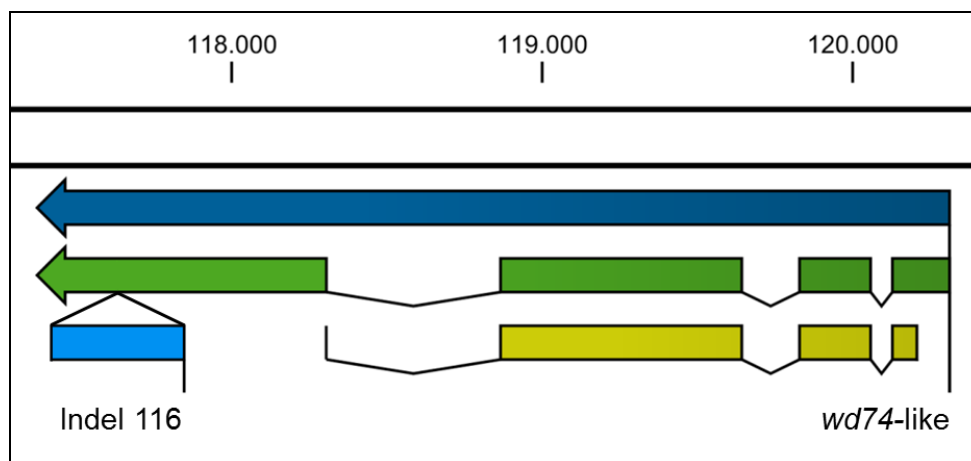


Abbildung 3.14: Lage und Struktur des Gens *wd74-like* im Contig SDR von *C. thummi*. Das Gen liegt auf dem „Crick“-Strang. Dargestellt sind der genomische Abschnitt des Transkripts (blau), das prozessierte Transkript des Gens (grün) inklusive der Introns sowie der proteinkodierende Teil (gelb). Weiterhin gezeigt ist der Indel-Bereich 116 im 3'UTR-Bereich des Gens (hellblau). Die genauen Größen der einzelnen Bereiche sind dem Text zu entnehmen.

Die Domänensuche mit den beiden Programmen SMART und InterProScan ergab einen Treffer. Über den proteinkodierenden Bereich sind fünf WD40-Repeats verteilt. Diese sind in den Bereichen der AS 41-83, 87-124, 126-164, 167-216 sowie 262-301

lokalisiert. In Abbildung 3.15 sind die Treffer der Domänensuche mit Hilfe des Programms InterProScan aufgezeigt. Im Falle des Proteins WD74 wurden die Domänen in den Datenbanken PFAM, SMART, PROFILE, SUPERFAMILY und PANTHER gefunden. Innerhalb der abgeleiteten AS-Sequenz konnte kein Kern-Lokalisierungs-Signal (NLS) gefunden werden.

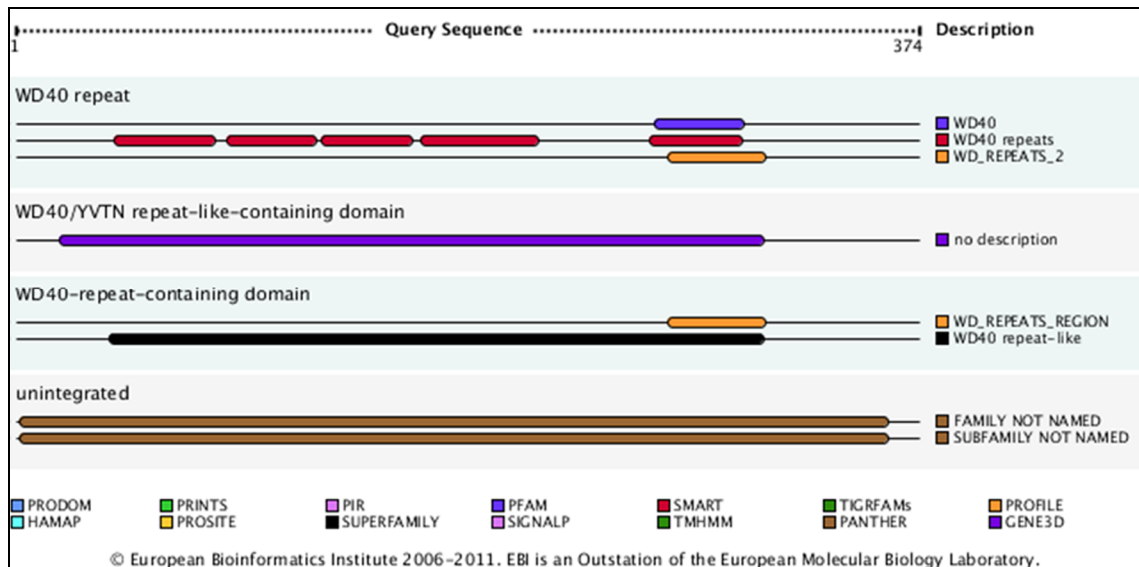


Abbildung 3.15: Domänen im Protein wd74-like nach Suche mit InterProScan

Die unterschiedlichen Farben der Balken zeigen an, in welcher Datenbank die entsprechende Domäne gefunden wurde. In diesem Fall wurden die Treffer zur Familie der WD40-Repeat-Proteine in den Datenbanken PFAM, SMART, PROFILE, SUPERFAMILY und PANTHER gefunden.

fs(1)K10-like:

Der transkribierte Bereich des Gens *fs(1)K10-like* erstreckt sich im Contig SDR von *C. thummi* auf dem „Watson“-Strang über den Bereich von 120500-122358 bp. In *C. thummi* ♂ liegt eine Duplikation des Gens *fs(1)K10-like* vor (Stiehl 2009, Kraemer, persönliche Mitteilung). Auch die Ergebnisse der RT-PCR sowie die Daten der Hochdurchsatz-Sequenzierung zeigen, dass in *C. thummi* ♂ alle drei Transkripte (X, Y1 und Y2) exprimiert werden.

Der genomische Bereich, den die X-chromosomale Variante in *C. thummi* abdeckt, ist ≥ 1859 bp lang. Das Gen besteht aus zwei Exons. Das prozessierte Transkript besitzt in *C. thummi* ♀ eine Größe von ≥ 1793 bp. Die proteinkodierende Sequenz besitzt eine Länge von 1131 Basen, was einer Polypeptidlänge von 377 AS entspricht. Der proteinkodierende Bereich erstreckt sich über beide Exons. Diese werden durch ein Intron mit 66 bp Länge voneinander getrennt. Bei dem in Abbildung 3.16 dargestellten Transkript von *fs(1)K10-like* handelt es sich um die X-chromosomale Variante.

Eine Überprüfung des Proteins *fs(1)K10-like* auf konservierte Domänen mit den Programmen SMART und InterProScan erbrachte keine Ergebnisse. Allerdings konnte

mit dem Programm MotifScan eine Reihe von Regionen nachgewiesen werden, die einen erhöhten Anteil an bestimmten AS aufweisen. Im Bereich von 11-106 AS sowie von 185-220 AS sind zwei Asparagin-reiche Regionen zu verzeichnen. Weiterhin sind eine Glutamin-reiche Region (21-138 AS), eine Prolin-reiche Region (108-178 AS) und eine Methionin-reiche Region (118-177 AS) zu finden. Zwei weitere Regionen mit einer Anreicherung an Lysin (202-319 AS) sowie Alanin (282-308 AS) werden von den e-Value Werten nur schwächer unterstützt. Weiterhin konnte ein Kern-Lokalisierungs-Signal (NLS) im Bereich von 209-225 AS identifiziert werden.

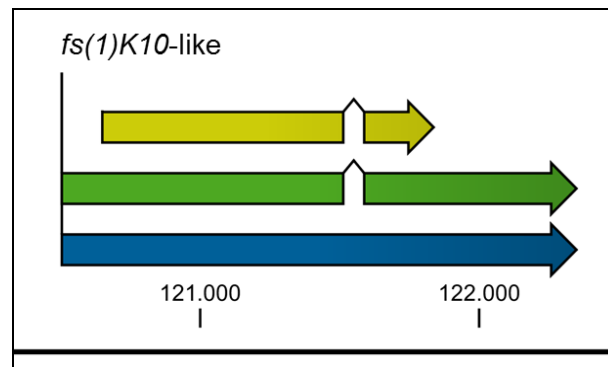


Abbildung 3.16: Lage und Struktur des Gens *fs(1)K10-like* im Contig SDR von *C. thummi*

Das Gen ist auf dem „Watson“-Strang lokalisiert. Dargestellt sind der genomische Abschnitt der prä-mRNA (blau), das prozessierte Transkript des Gens (grün) inklusive des Introns sowie der proteinkodierende Teil (gelb). Die genauen Größen der einzelnen Bereiche sind dem Text zu entnehmen.

Die in *C. thummi* ♂ aufgetretene Duplikation des Genbereichs *fs(1)K10* wird hier mit dem Zusatz Y1 für die erste Y-chromosomale Kopie des Gens, bzw. Y2 für die zweite Y-chromosomale Kopie des Gens bezeichnet. Die Struktur des duplizierten transkribierten Bereichs ist in Abbildung 3.18 gezeigt.

Das Transkript der Y1-Variante erstreckt sich hierbei über einen genomischen Bereich von ≥ 1751 bp. Das Transkript besteht aus zwei Exons. Die Länge des prozessierten Y1-Transkripts beträgt ≥ 1685 bp. Das Intron der Y1-Variante besitzt eine Länge von 64 bp. Der proteinkodierende Bereich weist eine Länge von 1116 Basen auf. Dies entspricht einer Länge von 372 AS. Hiermit besitzt die Y1-Variante in etwa dieselbe Länge wie die X-chromosomale Variante (siehe Abbildung 3.16).

Der transkribierte Bereich der Y2-Variante deckt einen genomischen Bereich von ≥ 1881 bp ab. Das prozessierte Transkript besteht ebenfalls aus zwei Exons und besitzt eine Länge von ≥ 1817 bp. Die beiden Exons werden durch ein 64 bp langes Intron voneinander getrennt. Der proteinkodierende Bereich der Y2-Variante ist mit 1032 Basen kürzer als die Y1-Variante und als die X-chromosomale Variante. Die abgeleitete AS-Sequenz besitzt eine Länge von 344 AS.

Ein Vergleich der beiden Y-chromosomalen Varianten des *fs(1)K10*-Proteins mit der X-chromosomalen Variante zeigt auf AS-Ebene mit 90,2 % eine höhere

Sequenzidentität zwischen dem Y1-Allel und dem X-Allel als zwischen dem Y2-Allel und dem X-Allel. Hier beträgt die Sequenzidentität nur 76,4 %. Ein Vergleich der beiden Y-chromosomalen Varianten zeigt eine Sequenzidentität von 76,6 % auf. Ein Protein-Alignment ist in Abbildung 3.17 gezeigt.

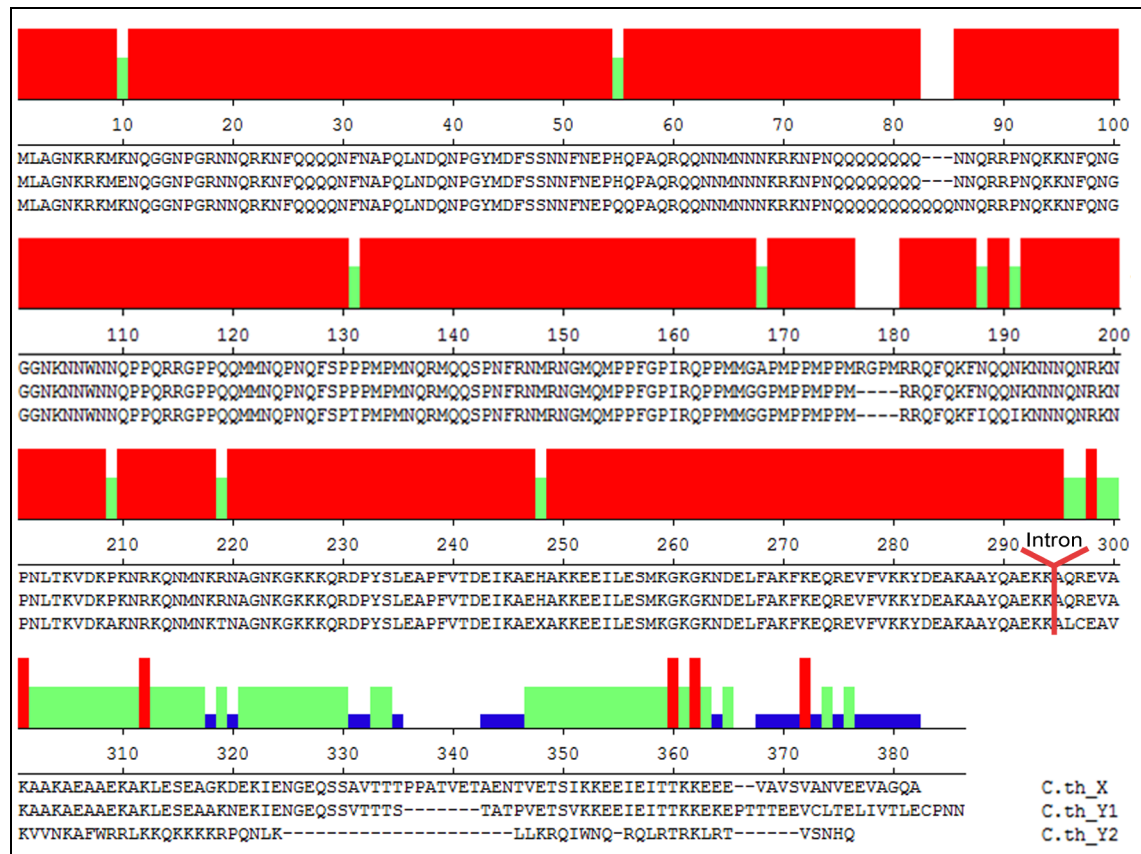


Abbildung 3.17: Protein-Alignment der drei fs(1)K10-Varianten (X, Y1 und Y2)

Die Abbildung zeigt ein Alignment der drei abgeleiteten AS-Sequenzen (X, Y1 und Y2) von *fs(1)K10-like* aus *C. thummi* ♂. Deutlich ist zu erkennen, dass die Übereinstimmung der drei Varianten innerhalb des N-terminalen Bereichs (Exon 1) deutlich höher ist als innerhalb des C-terminalen Abschnitts (Exon 2). Grafisch getrennt werden die beiden Exonbereiche durch die mit einem roten Einschub markierte Position des Introns im Transkript. Vor allem innerhalb des C-terminalen Bereichs ist die größere Homologie zwischen der X-chromosomalen sowie der Y-chromosomalen Variante Y1 zu erkennen. Ausgedrückt wird dies auch durch die prozentualen Übereinstimmungen (siehe Text).

Eine Überprüfung der beiden Y-chromosomalen Varianten des abgeleiteten Polypeptids fs(1)K10-like auf konservierte Domänen mit dem Programm InterProScan erbrachte, wie schon bei der Überprüfung der X-chromosomalen AS-Sequenz, keine signifikanten Ergebnisse. Die Analysen der Y-chromosomalen Varianten auf Regionen mit einem erhöhten Anteil einer bestimmten AS mit Hilfe des Programms MotifScan erbrachte ähnliche Ergebnisse wie die Untersuchung der X-chromosomalen Variante. Die Y1-Variante des Proteins fs(1)K10-like zeigte hierbei Bereiche mit einer erhöhten Anzahl der AS Asparagin (11-216 AS), Glutamin (21-138 AS), Prolin (108-172 AS) sowie Methionin (118-173 AS) auf. Weiterhin konnten zwei mittels e-Value etwas

schwächer unterstützte Bereiche mit einem erhöhten Anteil an Lysin (198-315 AS) sowie Alanin (278-311 AS) gefunden werden. Im C-terminalen Bereich von 272-317 AS konnte in der Y1-Variante mit Hilfe des Programms SMART eine „coiled-coil“ Struktur gefunden werden. Des Weiteren konnte, wie auch in der X-chromosomalen Variante, ein Kern-Lokalisierungs-Signal (NLS) im Bereich von 205-221 AS gefunden werden. Die Y2-Variante zeigt leichte Abweichungen gegenüber den Varianten Y1 und der X-chromosomalen Variante. Hierbei variieren die AS-Anreicherungen für bestimmte Regionen. Die Y2-Variante zeigt Bereiche mit einem erhöhten Anteil von Asparagin (11-109 AS), Glutamin (21-141 AS), Prolin (111-175 AS) und Methionin (121-176 AS). Ein weiterer Bereich mit einem erhöhten Anteil von Asparagin (189-218 AS) wird in der Y2-Variante nur schwächer unterstützt, ein Bereich mit einem erhöhten Anteil an Lysin (201-323 AS) hingegen stärker. Eine Anreicherung an Alanin konnte in der Y2-Variante nicht nachgewiesen werden. Ein Kern-Lokalisierungs-Signal (NLS) konnte auch in der Y2-Variante nachgewiesen werden (208-224 AS).

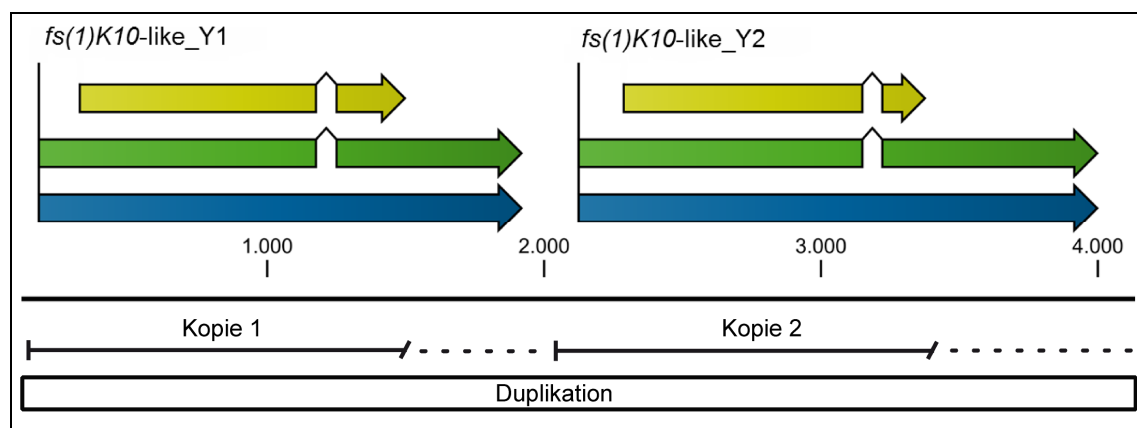


Abbildung 3.18: Struktur der männchenspezifischen Duplikation des Gens *fs(1)K10-like* (Y_1 ; Y_2) in *C. thummi*

Die Abbildung zeigt die Duplikation des Gens *fs(1)K10-like* in *C. thummi* ♂. Diese wurde mit Hilfe genomischer Southern-Analysen entdeckt (Stiehl 2009) und die genomische Sequenz per „Genome Walk“ und RACE erschlossen (Krämer; pers. Mitteilung). Die beiden Transkripte wurden im Laufe dieser Arbeit per RT-PCR sowie mit Hilfe der NGS-Daten bestätigt. Dargestellt sind die genomischen Abschnitte der prä-mRNA (blau), die prozessierten Transkripte des duplizierten Gens (grün) inklusive des Introns sowie der proteinkodierende Teil (gelb). Die genauen Größen der einzelnen Bereiche sind dem Text zu entnehmen. Die mit Kopie 1 und Kopie 2 gekennzeichneten Bereiche (unterer Bereich der Abb.) stellen die Sequenzbereiche dar, für die mittels Alignment eine Sequenzhomologie von 80 % oder höher gezeigt werden konnte. Im Bereich der gestrichelten Linien (jeweils 3'UTR) hingegen ist keine signifikante Sequenzhomologie zwischen den beiden Kopien zu erkennen.

***bcn92-like*:**

Das Gen *bcn92-like* liegt im Contig SDR von *C. thummi* in direkter distaler Nachbarschaft zum Gen *fs(1)K10-like*. Der transkribierte Bereich erstreckt sich über einen genomischen Abschnitt von ≥ 828 bp (122575-123402 bp). Lokalisiert ist das Gen auf dem „Watson“-Strang. Das prozessierte Transkript besitzt eine Gesamtlänge von ≥ 490 bp und zeigt keine Unterschiede zwischen den beiden Geschlechtern. Der proteinkodierende Bereich des Gens erstreckt sich über alle drei Exons des Gens und weist eine Länge von 261 Basen auf. Dies entspricht einer Länge des Polypeptids von 87 AS. Getrennt werden die drei Exons durch Introns der Längen 270 bp und 68 bp. In Abbildung 3.19 sind die genomische Lage sowie die Struktur des Transkripts *bcn92-like* im Contig SDR von *C. thummi* dargestellt.

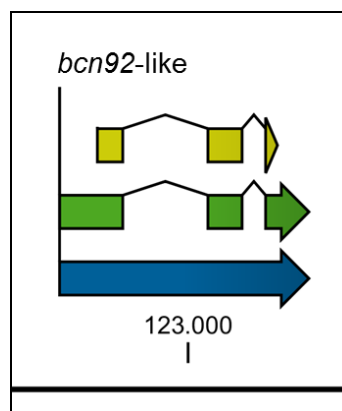


Abbildung 3.19: Lage und Struktur des Gens *bcn92-like* im Contig SDR von *C. thummi*

Das Gen ist auf dem „Watson“-Strang lokalisiert. Dargestellt sind der genomische Abschnitt der prä-mRNA (blau), das prozessierte Transkript des Gens (grün) inklusive der Introns sowie der proteinkodierende Teil (gelb). Die genauen Größen der einzelnen Bereiche sind dem Text zu entnehmen.

Die Überprüfung des abgeleiteten Polypeptids von *bcn92-like* auf konservierte Domänen mit Hilfe der beiden Programme SMART und InterProScan zeigte im Bereich der AS 9-71 einen Treffer zu einem Proteinkomplex mit LYR-Domäne. In der AS-Sequenz von *C. thummi* konnte an Position 14-16 AS das Tripeptid LYK gefunden werden. Hierbei ist demnach die stark basische AS Arginin durch die ebenfalls stark basische AS Lysin ersetzt. In Abbildung 3.20 ist der Treffer des Programms InterProScan gezeigt. Die Domänen wurden hierbei in den Datenbanken PFAM und PANTHER gefunden. Weiterhin wurde am N-Terminus des Proteins durch das Programm SIGNALP ein potentiell Signalpeptid gefunden. Ein Kern-Lokalisierungs-Signal konnte für das abgeleitete Polypeptid nicht gefunden werden.

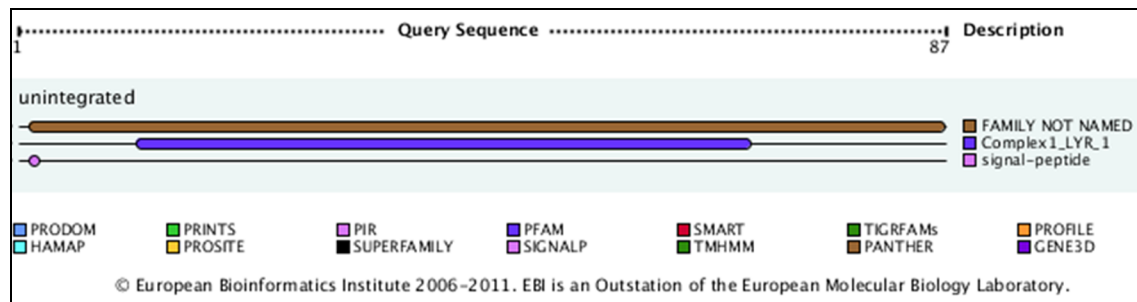


Abbildung 3.20: Domänen im Protein bcn92-like nach Suche mit InterProScan

Die unterschiedlichen Farben der Balken zeigen, in welcher Datenbank die entsprechende Domäne gefunden wurde. In diesem Fall wurden der Treffer zu einem Proteinkomplex mit LYR-Domäne in den Datenbanken PFAM und PANTHER gefunden. Weiterhin wurde in der Datenbank SIGNALP ein potentielles Signalpeptid am N-Terminus angegeben.

***d*-Laktat-Dehydrogenase-like:**

Das Gen *dLacDH* gehört zu einer Familie der Oxidoreduktasen, welche eine konservierte FAD-Bindedomäne aufweisen (Fraaije et al. 1998). Das Gen *dLacDH* besitzt eine Funktion im Energiestoffwechsel der Zelle. Es katalysiert die Oxidation von D-Laktat zu Pyruvat (Lodi et al. 1994).

Der transkribierte Bereich des Gens *dLacDH*-like im Contig SDR von *C. thummi* erstreckt sich über einen genomischen Abschnitt von ≥ 6586 bp (123196-129781 bp) auf dem „Crick“-Strang. Im 3'UTR-Bereich überlappt der transkribierte Bereich mit dem 3'UTR-Bereich des auf dem „Watson“-Strang liegenden Transkripts von *bcn92*-like. Sowohl in *C. thummi* ♀ wie auch in *C. thummi* ♂ konnten per RT-PCR zwei unterschiedliche Transkripte gefunden werden. Der Unterschied betrifft hierbei das zweite Exon des Gens und somit auch den proteinkodierenden Teil des Gens *dLacDH*-like. Das längere, prozessierte Transkript des Gens besitzt eine Länge von ≥ 1831 bp. Der proteinkodierende Bereich ist in diesem Fall 1497 Basen lang, was einer Polypeptidlänge von 499 AS entspricht. Beim kürzeren, prozessierten Transkript fehlen durch die Verwendung einer anderen 5'-Spleißstelle des 2. Exons 854 bp. Der daraus resultierende proteinkodierende Bereich des Gens ist auf eine Länge von 627 Basen (Bereich 123391-124238 bp) verkürzt, was einer Polypeptidlänge von 209 AS entspricht. Somit fehlt im kürzeren Transkript der gesamte N-terminale Bereich des Polypeptids. Getrennt werden die vier Exons durch drei Introns mit Längen von 4354 bp (bzw. 5208 bp im alternativen, kürzeren Transkript), 70 bp und 151 bp. In Abbildung 3.21 ist die genomische Lage des Transkripts *dLacDH*-like sowie dessen Struktur gezeigt.

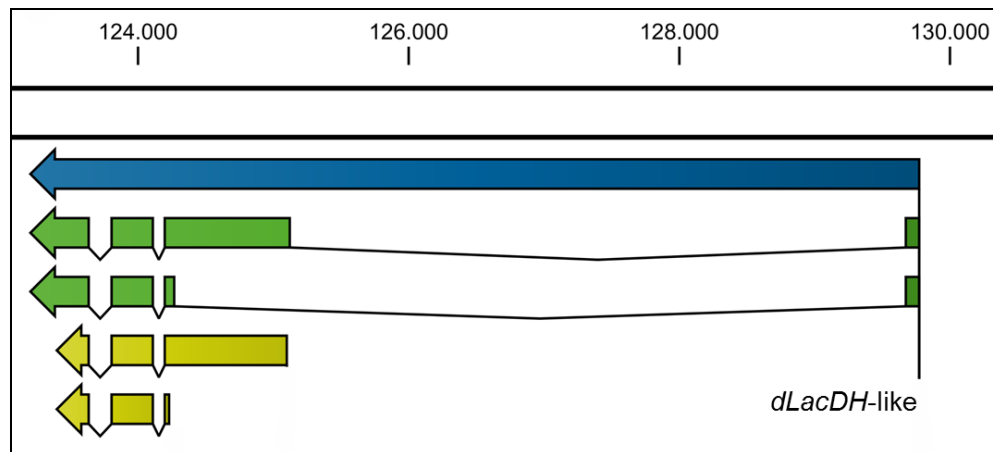


Abbildung 3.21: Lage und Struktur des Gens *dLacDH-like* im Contig SDR von *C. thummi*

Das Gen liegt auf dem „Crick“-Strang. Dargestellt sind der genomische Bereich der prä-mRNA (blau), die beiden prozessierten Transkripte des Gens (grün) inklusive der drei Introns sowie die proteinkodierenden Bereiche (gelb). Die genauen Größen der einzelnen Bereiche sind dem Text zu entnehmen. Zu erkennen ist, dass beim verkürzten, prozessierten Transkript der gesamte N-terminale Bereich des abgeleiteten Polypeptids fehlt.

Die Suche nach konservierten Domänen erfolgte mit Hilfe der Programme SMART und InterProScan und erbrachte zwei Treffer. Im N-terminalen Bereich (79-218 AS) ist eine FAD-Bindedomäne-4 zu verzeichnen. Viele Enzyme nutzen FAD als Co-Faktor. Die meisten Enzyme sind hierbei sauerstoffabhängige Oxidoreduktasen. Der zweite Treffer zeigt eine FAD gekoppelte Oxidase im C-terminalen Bereich der abgeleiteten AS-Sequenz (254-495 AS). Die beiden Domänen wurden in den Datenbanken PFAM, SUPERFAMILY, PROFILE, GENE3D und PANTHER gefunden und sind in Abbildung 3.22 gezeigt.

In der N-terminal verkürzten Variante des Polypeptids mit lediglich 209 AS statt 499 AS, ergab die Suche nach konservierten Domänen einen Treffer. Im Bereich von 1-205 AS wird ein Treffer zur FAD gekoppelten Oxidase gezeigt. Der Treffer wurde in den Datenbanken PFAM, SUPERFAMILY, GENE3D und PANTHER gefunden. Der Anteil des Proteins, welcher die N-terminale FAD-Bindedomäne-4 trägt, ist in der verkürzten Variante des abgeleiteten Polypeptids nicht vorhanden. Die Domänen der verkürzten Variante sind hier nicht getrennt dargestellt. Ein Kern-Lokalisierungs-Signal konnte in beiden Varianten nicht gefunden werden.

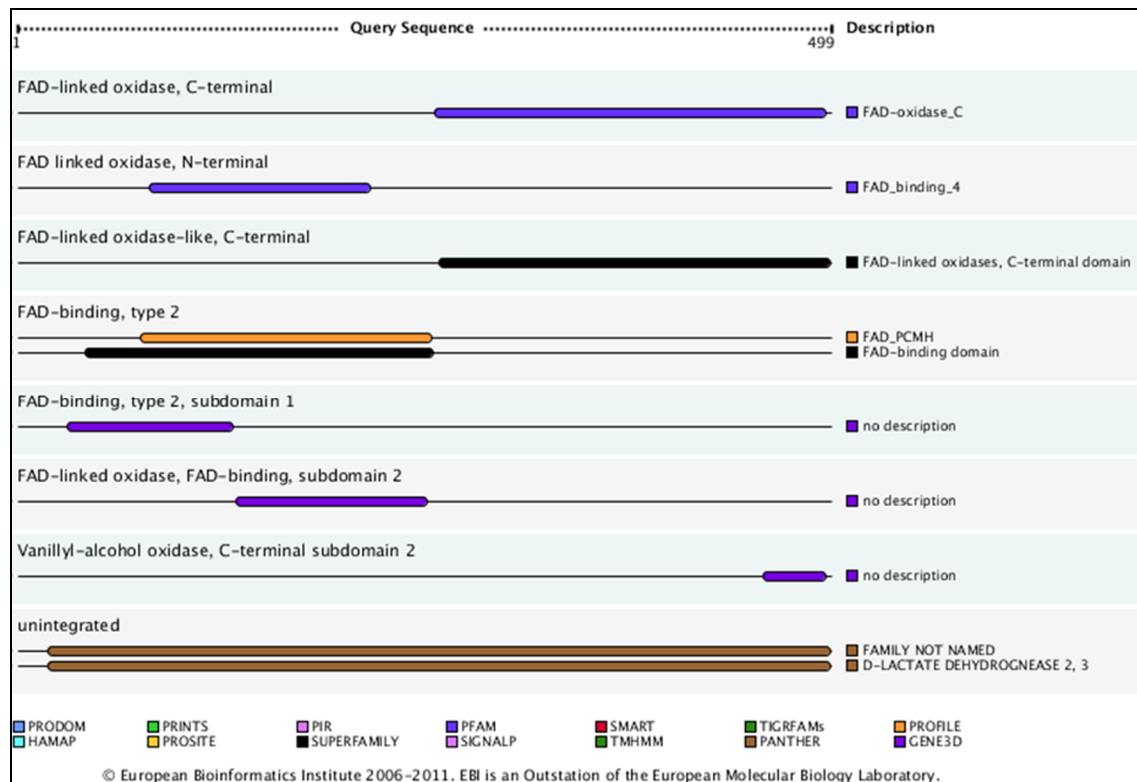


Abbildung 3.22: Domänen im Protein dLacDH-like nach Suche mit InterProScan
Gezeigt sind die Domänen im Protein dLacDH-like (lange Transkriptvariante). Die unterschiedlichen Farben der Balken zeigen, in welchen Datenbanken die entsprechende Domäne gefunden wurde. Im diesem Fall wurden die FAD-Bindedomäne-4 und die FAD-gekoppelte Oxidase in den Datenbanken PFAM, SUPERFAMILY, PROFILE, GENE3D und PANTHER gefunden.

polyhomeotic-like:

Der transkribierte Bereich des Gens *polyhomeotic-like* erstreckt sich im Contig SDR von *C. thummi* auf dem „Watson“-Strang über einen genomischen Abschnitt von ≥ 11924 bp (127928-139851 bp). Die Transkripte setzen sich aus mindestens 6 Exons zusammen, wobei zwei alternativ gespleißte erste Exons (Exon 1a sowie Exon 1b) vorliegen. Die prozessierten Transkripte weisen Längen von ≥ 4226 bp (Exon 1a) sowie ≥ 4405 bp (Exon 1b) auf. Mit Hilfe der RT-PCR sowie der Daten der beiden Hochdurchsatz-Sequenzierungen konnten die beiden ersten Exons (1a und 1b) in beiden Geschlechtern bestätigt werden. Allerdings liegen bislang nur 5'-RACE-Produkte aus *C. thummi* ♂ vor.

Der proteinkodierende Bereich des Gens *polyhomeotic-like* erstreckt sich im Transkript mit Exon 1a (weiter proximal lokalisiert) über die letzten vier Exons und besitzt eine Länge von 3183 Basen. Dies entspricht einer Polypeptidlänge von 1061 AS. Im Transkript mit Exon 1b (weiter distal lokalisiert) ist ein früherer Translationsstart möglich. Die Länge des proteinkodierenden Abschnitts beläuft sich hier auf 3483 Basen, was einer Proteinlänge von 1161 AS entspricht. In Abbildung 3.23 sind

die genomische Lage des Transkripts *polyhomeotic-like* im Contig SDR von *C. thummi* sowie die unterschiedlichen Transkriptvarianten gezeigt.

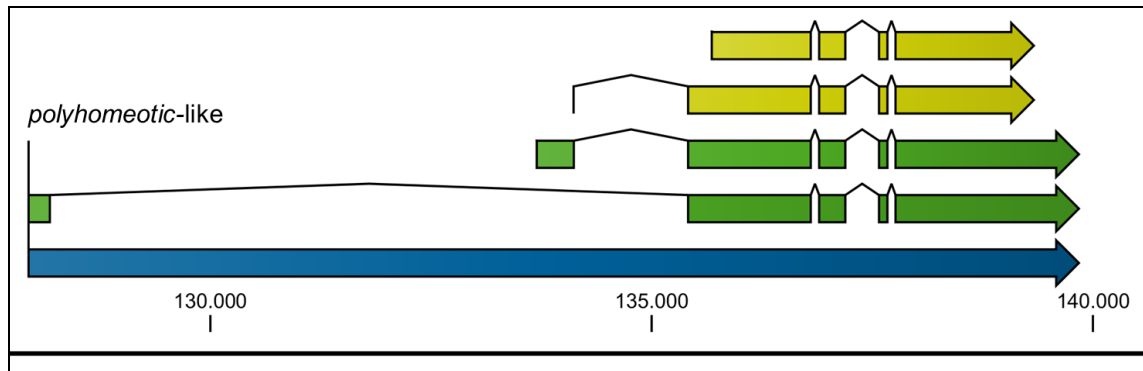


Abbildung 3.23: Lage und Struktur des Gens *polyhomeotic-like* im Contig SDR von *C. thummi*

Das Gen ist auf dem „Watson“-Strang lokalisiert. Dargestellt sind der genomische Abschnitt der prä-mRNA (blau), die prozessierten Transkripte des Gens (grün) inklusive der Introns sowie der proteinkodierende Teil (gelb). Die genauen Größen der einzelnen Bereiche sind dem Text zu entnehmen. Zu erkennen ist, dass zwei differentiell gespleißte Exons (Exon 1a und Exon 1b) verwendet werden.

Die Suche nach konservierten Domänen im Protein *polyhomeotic-like* mit Hilfe der beiden Programme SMART und InterProScan lieferte zwei Treffer. Im Bereich von 1077-1144 AS des Transkripts mit Exon 1b ist eine SAM-Domäne lokalisiert. Der zweite Treffer ist im Bereich 992-1025 AS lokalisiert. Hierbei handelt es sich um ein Zink-Finger-Motiv. Die Treffer wurden in den Datenbanken SMART, PROFILE, PFAM, GENE3D, SUPERFAMILY und PANTHER gefunden. In Abbildung 3.24 sind die einzelnen Treffer der Domänen- und Motivsuche abgebildet. Zusätzlich zu den beiden gefundenen Domänen konnten in der abgeleiteten AS-Sequenz fünf Glutamin-reiche Regionen identifiziert werden. Drei Regionen enthalten hierbei einen besonders hohen Anteil der AS Glutamin. Hierbei handelt es sich um die Bereiche 109-126 AS, 382-401 AS und 620-649 AS. Die anderen beiden Regionen sind ebenfalls reich an Glutamin, enthalten hiervon aber weitaus weniger als die ersten drei genannten Regionen. Es handelt sich hierbei um die Bereiche 164-186 AS sowie 335-353 AS. Alle fünf Glutamin-reichen Regionen liegen innerhalb des Bereichs (105-702 AS), welcher mit Hilfe des Programms MotifScan als Glutamin-reich angegeben wird. Bei der Serin/Threonin-reichen Region handelt es sich um die Region 847-886 AS. Diese weist mit 17 von 39 AS einen recht hohen Anteil der AS Serin und Threonin auf. Auch diese Region liegt innerhalb des durch MotifScan vorhergesagten Bereichs (778-887 AS). Ein potentiell Kern-Lokalisierungs-Signal konnte mit Hilfe des Programms MotifScan lediglich innerhalb des abgeleiteten Polypeptids mit Exon 1b im Bereich von 10-26 AS identifiziert werden.

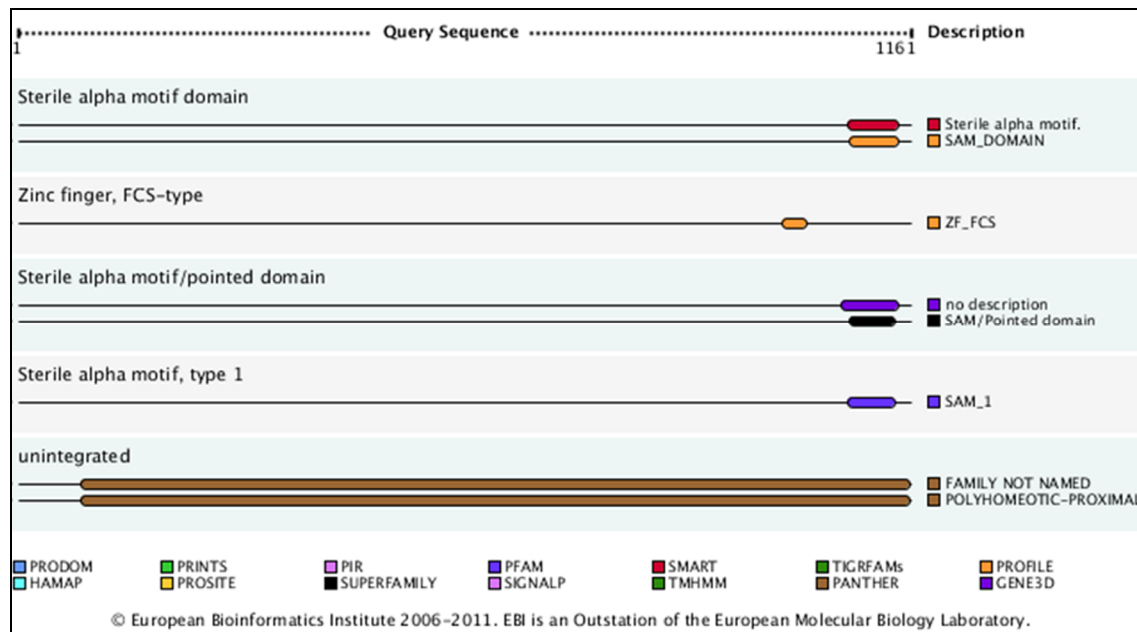


Abbildung 3.24: Domänen im Protein polyhomeotic-like nach Suche mit InterProScan Gezeigt sind die beiden Domänen im Protein polyhomeotic-like (lange Transkriptvariante). Die unterschiedlichen Farben der Balken zeigen, in welcher Datenbank die entsprechende Domäne gefunden wurde. Hier wurden die SAM-Domäne und das Zink-Finger-Motiv in den Datenbanken SMART, PROFILE, PFAM, GENE3D, SUPERFAMILY und PANTHER gefunden.

CG15207-like:

Das Gen CG15207 ist in *D. melanogaster* als proteinkodierend beschrieben. Es liegt auf dem X-Chromosom in der Region 10A1-2 (Babenko et al. 2009). Die Funktion des Gens ist bisher unbekannt.

Der transkribierte Bereich des Gens CG15207-like erstreckt sich im Contig SDR von *C. thummi* auf dem „Watson“-Strang über einen genomischen Bereich von ≥ 6063 bp (144127-150189 bp). Die prozessierte Transkriptlänge beträgt in *C. thummi* ♂ ≥ 3018 bp, wobei der 3'UTR-Bereich per RACE nicht komplett amplifiziert werden konnte. In *C. thummi* ♀ konnten weder 5'UTR- noch 3'UTR-Bereich analysiert werden. Das Gen besitzt mindestens vier Exons, die durch Introns mit Längen von 92 bp, 669 bp und 2284 bp unterbrochen werden. Der proteinkodierende Bereich des Gens erstreckt sich über den Bereich der Exons 2-4 und besitzt eine Länge von 2892 Basen. Dies entspricht einer Polypeptidlänge von 964 AS. Der proteinkodierende Bereich endet kurz vor dem bisher bekannten distalen Ende des Transkripts. In Abbildung 3.25 ist die genomische Lage des transkribierten Bereichs von CG15207-like im Contig SDR von *C. thummi* aufgezeigt.

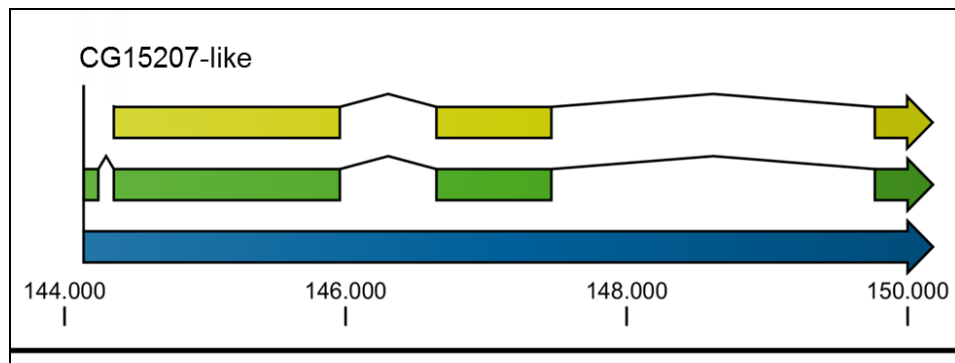


Abbildung 3.25: Lage und Struktur des Gens CG15207-like im Contig SDR von *C. thummi*

Das Gen ist auf dem „Watson“-Strang lokalisiert. Dargestellt sind der genomische Abschnitt der prä-mRNA (blau), das prozessierte Transkript des Gens (grün) inklusive der Introns sowie der proteinkodierende Teil (gelb). Die genauen Größen der einzelnen Bereiche sind dem Text zu entnehmen. Zu erkennen ist, dass es sich bei dem Transkript nicht um ein vollständiges Transkript handelt, da der 3'UTR-Bereich fehlt.

Die Untersuchung der abgeleiteten AS-Sequenz auf konservierte Domänen mit Hilfe der beiden Programme SMART und InterProScan lieferte keine Treffer zu konservierten Domänen. Allerdings zeigte das Programm SMART zwei Bereiche in der Sequenz auf, die interne „Repeats“ aufweisen. Hierbei handelt es sich um die Bereiche von 205-280 AS sowie 377-447 AS. Diese wurden mittels Analyse mit dem Programm Prospero gefunden.

Eine weitere Untersuchung des abgeleiteten Polypeptids mit MotifScan zeigte vier Regionen mit Anhäufungen von bestimmten AS. Im Bereich von 865-935 AS ist eine Glutamin-reiche Region zu finden. Schwach unterstützt werden hingegen die Regionen 261-313 AS (Serin-reich), 725-730 AS (Asparagin-reich) sowie 791-909 AS (Cystein-reich). Ein Kern-Lokalisierungs-Signal konnte innerhalb der AS-Sequenz nicht gefunden werden.

CG11203-like:

Das Gen CG11203 ist in *D. melanogaster* ebenfalls als proteinkodierend beschrieben und liegt auf dem X-Chromosom in distaler Nachbarschaft zu dem Gen CG15207-like (Huntley und Golding 2004; Babenko et al. 2009). Die Funktion des Gens ist ebenfalls bisher unbekannt.

Der transkribierte Bereich des Gens CG11203-like erstreckt sich im Contig SDR von *C. thummi* auf dem „Watson“-Strang über einen genomischen Bereich von ≥ 8893 bp (162183-171075 bp). Es konnte ein prozessiertes Transkript mit einer Länge von mindestens 2956 bp in *C. thummi* ♂ identifiziert werden. Allerdings konnte der 5'UTR-Bereich nicht komplett per RACE amplifiziert werden. In *C. thummi* ♀ konnten weder 5'UTR- noch 3'UTR-Bereich analysiert werden. Das Gen besteht aus mindestens sieben Exons. Diese sind durch Introns mit Längen von 5386 bp, 141 bp,

101 bp, 173 bp, 68 bp und 68 bp voneinander getrennt. Der proteinkodierende Bereich des Gens weist eine Länge von 2743 Basen auf. Ein mögliches Startcodon befindet sich im Contig von *C. thummi*, 14 bp vor dem bisher sequenzierten Bereich. Die daraus resultierende Länge des proteinkodierenden Abschnitts läge somit bei 2757 Basen, was einer Polypeptidlänge von 919 AS entsprechen würde. Die Lage und Struktur des Transkripts CG11203-like sind in Abbildung 3.26 gezeigt.

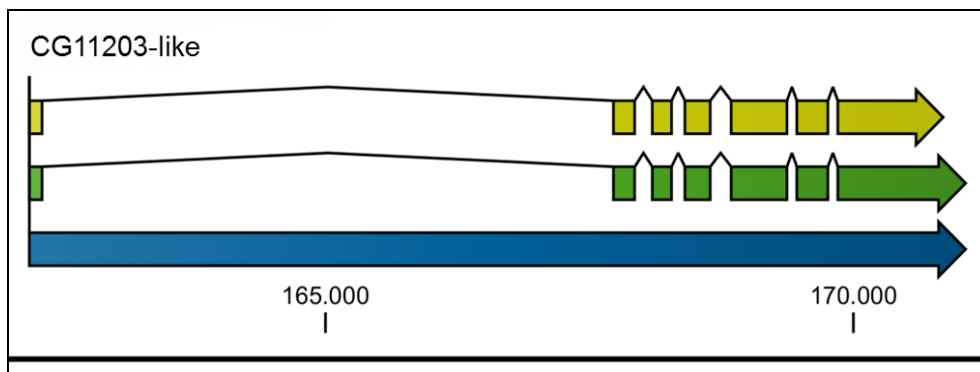


Abbildung 3.26: Lage und Struktur des Gens CG11203-like im Contig SDR von *C. thummi*

Das Gen ist auf dem „Watson“-Strang lokalisiert. Dargestellt sind der genomische Abschnitt der prä-mRNA (blau), das prozessierte Transkript des Gens (grün) inklusive der Introns sowie der proteinkodierende Teil (gelb). Die genauen Größen der einzelnen Bereiche sind dem Text zu entnehmen. Zu erkennen ist, dass es sich bei dem analysierten Bereich nicht um die Volllänge des Transkripts handelt, da der 5'UTR-Bereich fehlt.

Die Suche nach konservierten Domänen im Protein CG11203-like mit Hilfe der Programme SMART und InterProScan lieferte einen Treffer. Dieser wurde in der Datenbank PANTHER gefunden und erstreckt sich über den Bereich 251-763 AS. Nach der Klassifizierung der Datenbank gehört das Protein in die Familie der am Cytoskelett beteiligten Actine.

Eine weitere Untersuchung mittels MotifScan zeigte vier Regionen mit Anhäufungen von bestimmten AS. Zwei der vier Treffer auf Regionen werden hierbei als schwach unterstützt eingestuft, die anderen beiden Treffer sind gesichert. Im Bereich von 341-358 AS ist eine Serin-reiche Region zu finden. Weiterhin wird im Bereich 563-720 AS eine Glutamin-reiche Region angezeigt. Schwach unterstützt werden hingegen die Regionen 237-257 AS (ebenfalls Serin-reich) sowie 581-597 AS (Asparagin-reich). Des Weiteren besteht im Bereich 361-374 AS ein potentieller Treffer zu einem Kern-Lokalisierungs-Signal (NLS).

Im Contig *C. tentans* ist an Position 75195-75197 bp (Exon 6) ein Stopcodon (TAA) im Gen zu finden. Dieser Bereich des Contigs resultiert aus den Sequenzierungen der Subklone des BAC-Klons L13C2/1-1. Der Bereich wurde aus diesem Grund per PCR auf genomische DNA sowie cDNA von *C. tentans* erneut überprüft. Hierbei konnte festgestellt werden, dass es sich bei diesem Basen-Triplett scheinbar um eine allelische Variante handelt. Die Sequenzierungen der PCR-Produkte mit genomischer

DNA sowie cDNA als Template haben gezeigt, dass sowohl das Triplet TAA, wie auch GGA im sequenzierten Bereich vorkommen. Das Triplet GGA kodiert für die Aminosäure Glycin und erhält den ORF des Gens. In Abbildung 3.27 sind die Chromatogramme der unterschiedlichen Sequenzierungen des Bereichs gezeigt.

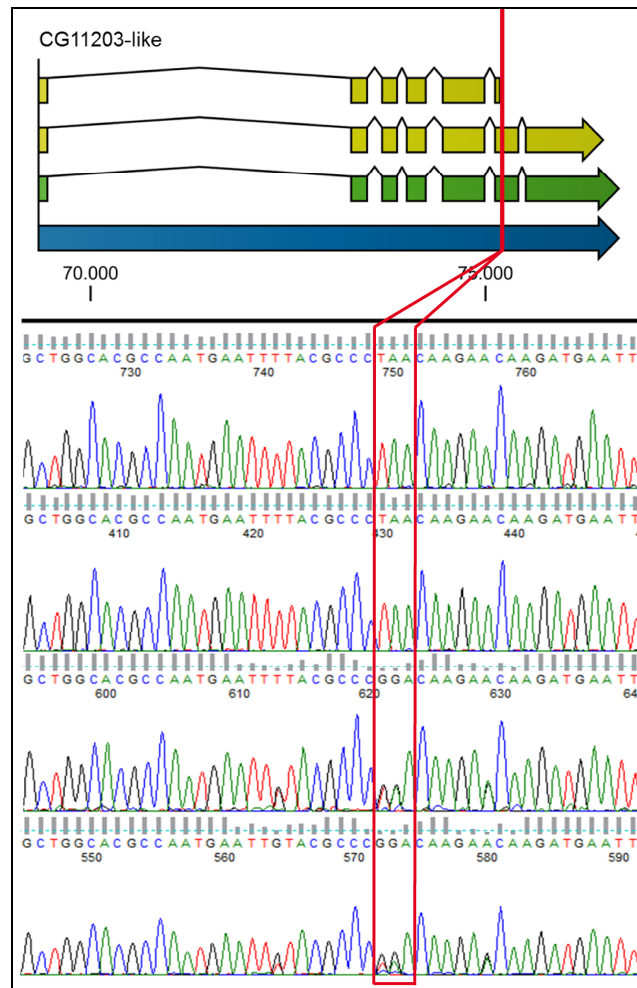


Abbildung 3.27: Position sowie Chromatogramme der allelischen Varianten im Gen CG11203-like in *C. tentans*

Die Abbildung zeigt die Position der allelischen Varianz im Exon 6 des Gens CG11203-like in *C. tentans* sowie ein Chromatogramm-Alignment von vier Sequenzen. Die oberen beiden Sequenzen stammen aus der Sequenzierung des BAC-Klons L13C2/1-1, die dritte und vierte Sequenz aus Sequenzierungen von PCR-Produkten auf genomische DNA bzw. cDNA aus *C. tentans*. In den Sequenzen der PCRs mit genomische DNA sowie cDNA als Template ist deutlich zu erkennen, dass bei den ersten beiden Basen des Triplets heterozygote Signale vorliegen.

***rpn5-like*:**

Der transkribierte Bereich des Gens *rpn5-like* erstreckt sich im Contig SDR von *C. thummi* auf dem „Watson“-Strang über einen genomischen Bereich von ≥ 1834 bp (173622-175455 bp). Im Rahmen der vorliegenden Arbeit ist es gelungen, mit Hilfe der RT-PCR sowie RACE ein vollständiges Transkript des Gens *rpn5-like* zu erhalten. Das prozessierte Transkript besitzt eine Länge von ≥ 1577 bp und zeigt keine Unterschiede zwischen *C. thummi* ♀ und *C. thummi* ♂. Das Transkript besteht aus vier Exons, die durch Introns mit Längen von 133 bp, 64 bp sowie 60 bp voneinander getrennt sind. Der proteinkodierende Bereich erstreckt sich über den Bereich der Exons 2-4 und besitzt eine Länge von 1362 Basen. Dies entspricht einer Polypeptidlänge von 454 AS. In Abbildung 3.28 ist die genomische Lage des Transkripts im Contig SDR von *C. thummi* sowie dessen Struktur dargestellt.

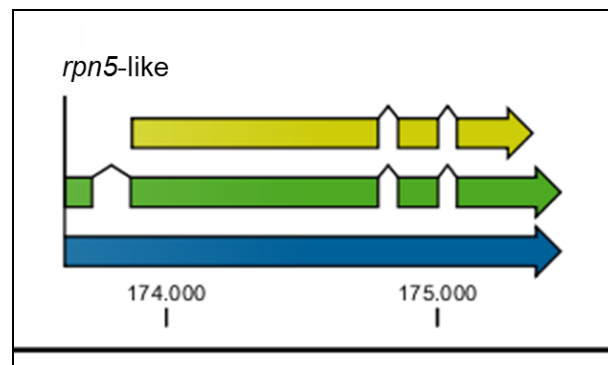


Abbildung 3.28: Lage und Struktur des Gens *rpn5-like* im Contig SDR von *C. thummi*. Das Gen ist auf dem „Watson“-Strang lokalisiert. Dargestellt sind der genomische Bereich der prä-mRNA (blau), das prozessierte Transkript des Gens (grün) inklusive der Introns sowie der proteinkodierende Teil (gelb). Die genauen Größen der einzelnen Bereiche sind dem Text zu entnehmen.

Die Analyse des Proteins auf konservierte Domänen mit den beiden Programmen SMART und InterProScan erbrachte einen Treffer. Im Bereich von 346-432 AS zeigt das Protein einen Treffer zu einer PCI-, oder auch PINT-Domäne. Abbildung 3.29 zeigt die Treffer der Domänensuche mit dem Programm InterProScan. Die Treffer wurden in den Datenbanken PFAM, SMART, GENE3D, PANTHER und SUPERFAMILY gefunden. Das Programm MotifScan gibt die PCI-Domäne für den Bereich von 240-414 AS an. Weiterhin zeigt MotifScan ein mögliches Kern-Lokalisierungs-Signal im Bereich von 82-99 AS.

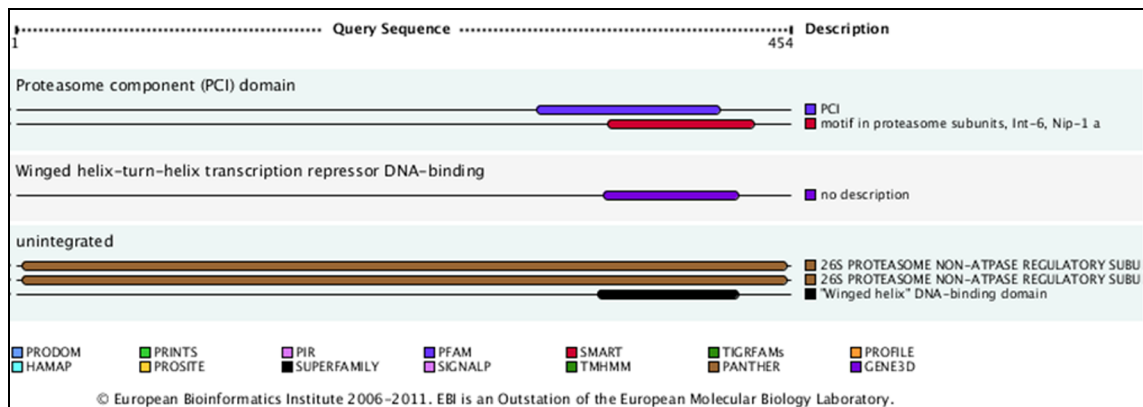


Abbildung 3.29: Domänen im Protein *rpn5-like* nach Suche mit InterProScan

Die unterschiedlichen Farben der Balken zeigen, in welchen Datenbanken die entsprechende Domäne gefunden wurde. Hier wurden die PCI-Domäne (auch PINT-Domäne genannt) in den Datenbanken PFAM, SMART, GENE3D, PANTHER und SUPERFAMILY gefunden.

hermansky pudlak-like:

Der transkribierte Bereich des Gens *hermansky pudlak-like* erstreckt sich im Contig SDR von *C. thummi* auf dem „Crick“-Strang über einen genomischen Bereich von ≥ 5398 bp (175487-180884 bp). Mit Hilfe von RT-PCR und RACE konnten in *C. thummi* ♂ zwei 5'-RACE-Produkte gefunden werden, die aus der Verwendung alternativer erster Exons resultieren (Exon 1a sowie Exon 1b). Das Produkt mit Exon 1b (weiter proximal; näher an Exon 2) aus *C. thummi* ♂ gleicht hierbei dem gefundenen Produkt aus *C. thummi* ♀. Das Gen besteht aus fünf Exons. Diese werden durch Introns mit Längen von 453 bp (Exon 1a) bzw. 236 bp (Exon 1b), 60 bp und 62 bp unterbrochen. Die per RT-PCR und RACE analysierten Produkte weisen in *C. thummi* ♂ je nach Verwendung des ersten Exons Transkriptlängen von ≥ 4427 bp (Exon 1a) sowie ≥ 4319 bp (Exon 1b) auf. In *C. thummi* ♀ konnte bisher nur die Transkriptvariante unter Verwendung von Exon 1b mit einer Länge von ≥ 4319 bp bestätigt werden. Der proteinkodierende Bereich erstreckt sich über die Exons 2-4 und besitzt eine Länge von 3678 Basen. Dies entspricht einer Polypeptidlänge von 1226 AS.

Im 3'UTR-Bereich des Gens *hermansky pudlak-like* konnte mit Hilfe von RT-PCR und RACE-Produkten ein 396 bp langer genomischer Bereich (175662-176057 bp) nachgewiesen werden, welcher durch die PCR-Produkte überspannt wird. An den Rändern des durch die PCR-Produkte überspannten Bereichs, zeigen sich keine typischen Intron-Exon-Strukturen. Analysen des Bereichs zeigen 121 bp lange IRs (175718-175839 bp/ 175848-175969 bp) sowie an den Rändern des Abschnitts liegende 8 bp lange DRs (175654-175661 bp/ 176050-176057 bp).

In Abbildung 3.30 ist die genomische Lage des Transkripts *hermansky pudlak-like* im Contig SDR von *C. thummi* sowie dessen Struktur gezeigt.

Die Suche mit den Programmen SMART und InterProScan auf konservierte Domänen innerhalb des abgeleiteten Polypeptids von *hermansky pudlak-like* erbrachte keine

Treffer. Auch eine Suche mit MotifScan lieferte keine Ergebnisse. Die in der Literatur im Protein HPS3 gefundene Clathrin-Bindedomäne konnte somit für *C. thummi* nicht bestätigt werden. Ein Kern-Lokalisierungs-Signal konnte innerhalb der AS-Sequenz nicht identifiziert werden.

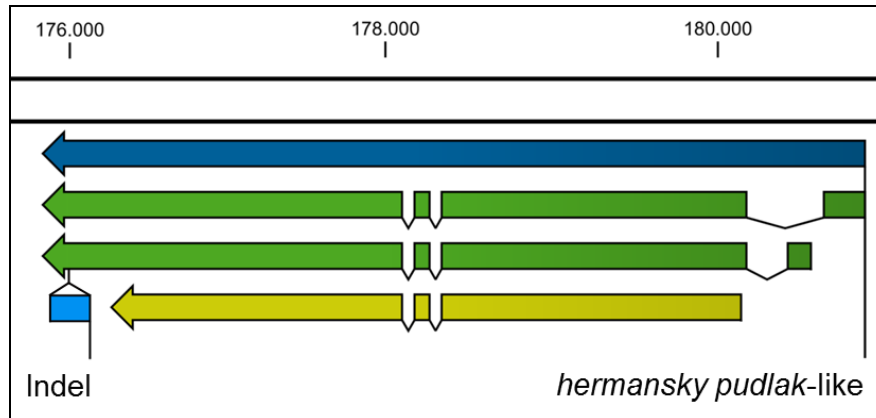


Abbildung 3.30: Lage und Struktur des Gens *hermansky pudlak-like* im Contig SDR von *C. thummi*

Das Gen ist auf dem „Crick“-Strang lokalisiert. Dargestellt sind der genomische Bereich der prä-mRNA (blau), die prozessierten Transkripte des Gens (grün) inklusive der Introns sowie der proteinkodierende Teil (gelb). Des Weiteren ist der Indel-Bereich im 3'UTR-Bereich des Gens (hellblau) gezeigt. Die genauen Größen der einzelnen Bereiche sind dem Text zu entnehmen.

Tabelle 3.3: Sequenzidentitäten der verschiedenen Gene und ihrer Produkte der SDR zwischen den beiden Arten *C. tentans* und *C. thummi*

Die Tabelle zeigt die Sequenzidentitäten der einzelnen Gene der SDR sowie deren Produkte zwischen den beiden Arten *C. tentans* und *C. thummi*. Verglichen wurden hierbei die Übereinstimmungen der genomischen Bereiche, der transkribierten Bereiche (cDNA), der proteinkodierenden Bereiche sowie die Sequenzidentität auf Polypeptidebene (AS).

Gen	Sequenzidentität (Ebene)	Sequenzidentität (%)	Gen	Sequenzidentität (Ebene)	Sequenzidentität (%)
<i>mi-er1-like</i>	genomisch	82,6 %	<i>polyhomeotic-like</i>	genomisch	81,5 %
	cDNA	85,4 %		cDNA	87,5 %
	proteinkodierend	85,8 %		proteinkodierend	90,3 %
	AS	87,6 %		AS	93,8 %
<i>luc7(p)-like</i>	genomisch	62,9 %	CG15207-like	genomisch	83,7 %
	cDNA	79,9 %		cDNA	88,8 %
	proteinkodierend	90,1 %		proteinkodierend	88,8 %
	AS	97,2 %		AS	92,0 %
<i>wd74-like</i>	genomisch	63,6 %	CG11203-like	genomisch	61,0 %
	cDNA	82,6 %		cDNA	92,3 %
	proteinkodierend	86,6 %		proteinkodierend	92,7 %
	AS	91,9 %		AS	96,9 %
<i>fs(1)K10-like</i>	genomisch	78,5 %	<i>rpn5-like</i>	genomisch	88,7 %
	cDNA	78,7 %		cDNA	92,0 %
	proteinkodierend	86,8 %		proteinkodierend	94,1 %
	AS	90,4 %		AS	96,5 %
<i>bcn92-like</i>	genomisch	64,1 %	<i>hermansky pudlak-like</i>	genomisch	76,1 %
	cDNA	82,2 %		cDNA	80,6-88,8 %
	proteinkodierend	88,9 %		proteinkodierend	89,7 %
	AS	90,8 %		AS	93,0 %
<i>dLacDH-like</i>	genomisch	65,5 %	Mittelwert aller Gene	genomisch	73,5 %
	cDNA	84,4 %		cDNA	85,4 %
	proteinkodierend	89,2 %		proteinkodierend	89,4 %
	AS	95,8 %		AS	93,3 %

3.2.1 Analyse der Promotorbereiche der Gene der SDR

Mit Hilfe der in Kapitel 3.2 durchgeführten RT-PCR und RACE-Experimente konnten von den meisten Genen der SDR von *C. thummi* Vollängen-Transkripte gewonnen werden. Die Transkripte der verschiedenen Gene sind in den entsprechenden Abbildungen in Kapitel 3.2 dargestellt. Durch Sequenzvergleiche der erhaltenen Transkripte aus *C. thummi* zu den genomischen Contigs aus *C. tentans* und *C. piger* konnte die Lage sowie die Intron-Exon-Struktur der einzelnen Gene in *C. tentans* und *C. piger* vorhergesagt werden.

Zur Analyse der Promotorbereiche der einzelnen Gene wurden etwa 500 bp vor dem bekannten 5'UTR-Bereich detailliert analysiert. Dieser Bereich wurde zwischen den drei Spezies *C. thummi*, *C. piger* und *C. tentans* mit Hilfe des Programms MegAlign aus dem „DNASTAR Lasergene®“- Software Paket der Fa. Lasergene (Madison, USA) verglichen. Im Folgenden wurde der Bereich nach möglichen Transkriptionsstartpunkten, Kernpromotorelementen sowie proximalen Promotorelementen durchsucht. Die Suche konzentrierte sich hierbei einerseits auf Motive, die aus früheren Studien an Chironomiden bekannt sind (Kraemer 1994). Weiterhin wurden degenerierte Sequenzmotive der verschiedenen Promotorelemente genutzt (Juven-Gershon et al. 2008). Die zur Suche verwendeten Sequenzmotive sind in Tabelle 3.4 angegeben.

Tabelle 3.4: Analyisierte Promotorelemente

In der Tabelle sind die analysierten Promotorelemente inklusive der zur Suche verwendeten Sequenzmotive sowie der relativen Position der Elemente zum Transkriptionsstartpunkt angegeben

Promotorelement	Abkürzung	Sequenzmotiv	Position
Start der Transkription; Initiatorregion	INR	TCAKTY (<i>D. melanogaster</i>)	-2...+4
Start der Transkription; Initiatorregion	INR	YYANWYY (<i>H. sapiens</i>)	-2...+4
TATA-Box	TATA-Box	TATA, TATAAT, TATAAA	-31...-26
„TFIIB recognition element“ (up)	BREup	SSRCGCC	-37...-32
„TFIIB recognition element“ (down)	BREdown	RTDKKKK	-25...-19
„Motif ten element“	MTE	CSARCSSAAC	+18...+27
„Downstream Core Promotor Element“	DPE	RGWYVT	+28...+33
CAAT-Box	CAAT-Box	CCAAT, CAAT	-50...-200
GC-Box	GC-Box	GGGCGG, GCGC, CCCGC, GGGC	-50...-200

Hierbei beschränkt sich der Kernpromotorbereich auf den Bereich von -50 bis +50 bp, bezogen auf den Transkriptionsstartpunkt. Hierbei ist die TATA-Box im Bereich von etwa -31 bis -26 bp lokalisiert. Allerdings gibt es viele Gene ohne eine TATA-Box im Promotorbereich (Juven-Gershon et al. 2008). In direkter Nachbarschaft zur TATA-Box liegen die „TFIIB recognition element“ (BRE)-Elemente. In relativer Nähe zur Initiatorregion, distal des Transkriptionsstartpunkts liegt das „Motif Ten Element“ (MTE; +18...+27 bp) sowie das „Downstream Core Promotor Element“ (DPE; +28...+32 bp). Das Initiatormotiv (INR) liegt direkt am Startpunkt der

Transkription bei -2 bis +4 bp. Der proximale Promotorbereich ist zwischen den Basen -50 und -200 lokalisiert. Er kann die CAAT-Box, die GC-Box sowie CpG-Inseln beinhalten. Auf die CpG-Inseln wird im Rahmen dieser Arbeit nicht weiter eingegangen.

In Abbildung 3.31 sind die Promotorbereiche der Gene der SDR grafisch dargestellt. Angegeben ist hierbei der Bereich von etwa -190 bis +35 bp, bezogen auf den möglichen Transkriptionsstartpunkt der Gene. Die Sequenzen der drei miteinander verglichenen Arten *C. thummi*, *C. piger* sowie *C. tentans* sind jeweils als Sequenz-Alignment untereinander abgebildet. Hierbei ist bei *C. thummi* der durch die 5'-RACE erhaltene Bereich in roten Buchstaben dargestellt. Bei den Genen des distalen Bereichs der SDR fehlen die Sequenzinformationen aus *C. piger*. Aus diesem Grund zeigen die Alignments nur den Sequenzvergleich zwischen den beiden Arten *C. thummi* und *C. tentans*.

Die möglichen Initiatorregionen inklusive des Transkriptionsstartpunktes sind durch rote Kästen gekennzeichnet. In einigen Fällen ist zusätzlich eine mögliche Initiatorregion mit dem Sequenzmotiv aus *Homo sapiens* in olivgrün angegeben, wenn diese in entsprechender Entfernung zur TATA-Box oder BRE liegt und sich nicht mit einer existierenden Initiatorregion mit dem Sequenzmotiv aus *D. melanogaster* überlagert. Der Bereich der TATA-Boxen ist dunkelblau gekennzeichnet. Das Element BRE_{down} ist in violett dargestellt. In keinem der analysierten Gene der SDR wurde ein Element der Sorte BRE_{up} gefunden. In distaler Lage zur Initiatorregion ist das Element DPE lokalisiert. Dieses ist in hellblau dargestellt. Das Element MTE wurde ebenfalls in keinem der analysierten Gene der SDR gefunden.

Im proximalen Promotorbereich sind die CAAT-Boxen in gelb dargestellt. Des Weiteren sind die GC-Boxen in grün gezeigt.

Innerhalb der Promotorbereiche sind Sequenzmotive, die in mindestens einer Spezies vorkommen, jedoch in den anderen Arten an entsprechender Stelle Basenaustausche zeigen, in grau dargestellt.

Bei den Analysen wurde nicht ausschließlich nach den in Tabelle 3.4 angegebenen Sequenzmotiven gesucht. Es wurde auch auf verkürzte Sequenzmotive und Abwandlungen innerhalb der Sequenzen der Motive geschaut. Hierbei wurde ausgehend von den Ergebnissen der Sequenzierung der 5'RACE nach dem ersten stromaufwärts liegenden Motiv für einen möglichen Transkriptionsstart gesucht. Im Umkreis dieser möglichen Initiatorregion wurde im Folgenden nach allen weiteren Elementen gesucht.

4 Diskussion

4.1 Vergleich der geschlechtsbestimmenden Region (SDR) von *C. tentans* und *C. thummi* sowie *C. thummi* ♀ und *C. thummi* ♂

Die Geschlechtschromosomen von *C. tentans* stellen wie in den meisten Arten von *Chironomus* (Bauer und Beermann 1952; Beermann 1955) keine heteromorphen Heterosomen dar. Das Chromosom, welches den dominanten männchenbestimmenden Faktor *M* trägt wird als Proto-Y-Chromosom bezeichnet. Die Lage des Faktors *M* konnte in den *Camptochironomus*-Arten *C. tentans* und *C. pallidivittatus* auf eine Region nahe dem Telomer von Chromosomenarm F oder Arm B (Beermann 1955) sowie in der *Chironomus*-Art *C. thummi* (Hägele 1985) auf eine Region nahe dem Telomer von Chromosomenarm F lokalisiert werden. Die chromosomale Lage der geschlechtsbestimmenden Region auf Arm F stimmt hierbei zwischen der schwedischen Population von *C. tentans* und *C. thummi* überein (Beermann 1955; Hägele 1985). Bisher konnte die Lage des dominanten Faktors *M* in vielen verschiedenen Arten von *Chironomus* nachgewiesen werden. Hierbei ist *M* auf allen Chromosomenarmen außer auf Arm E zu finden. Eine Übersicht über die Lage des Faktors *M* in 24 verschiedenen Arten ist in Martin (2010) gegeben.

Nach der Hypothese von Muller (1918) entstehen heteromorphe Geschlechtschromosomen aus einem Autosomenpaar. Eines der beiden Autosomen trägt hierbei den dominanten geschlechtsdeterminierenden Faktor. Im Folgenden kommt es zu einer Unterdrückung der Rekombination zwischen den beiden Chromosomen. Hierdurch können sich repetitive und transposable Elemente innerhalb der geschlechtsbestimmenden Region ansammeln, da diese auf Grund der fehlenden Rekombination nicht mehr entfernt werden können (Steinemann 1982; Steinemann und Steinemann 1992; 1998). Im Laufe der Zeit erfolgt somit eine Heterochromatinisierung des Proto-Y-Chromosoms (Charlesworth 1991).

In *C. thummi* konnte in der Region D_{3e1} des Arms F (Chromosom 3L) eine hemizygot heterochromatische Bande nachgewiesen werden, welche männchenspezifisch ist (Hägele 1985). Ebenfalls wurde in dieser Region ein männchenspezifisches hemizygot *Cla*-Element Cluster nachgewiesen (Kraemer und Schmidt 1993). Sowohl in *C. thummi* ♀ wie auch in beiden Geschlechtern der Schwesternart *C. piger* und in *C. tentans* fehlen das hemizygot *Cla*-Element Cluster und die hemizygot, heterochromatische Bande. Diese Unterschiede im Proto-Y-Chromosom zwischen den Arten bieten die Möglichkeit, ein frühes Stadium im evolutiven Prozess der Entstehung von Heterosomen zu untersuchen (Kraemer 1994). Hierbei handelt es sich bei *C. thummi* um die phylogenetisch jüngere Art (Keyl 1962; Keyl 1965).

In *C. tentans* konnten im Laufe dieser Arbeit sechs Bereiche mit repetitiven bzw. transposablen Elementen innerhalb der untersuchten Region der SDR mit einer Größe von rund 87 kb identifiziert werden. Hierbei handelt es sich um ein ATR26-tandemrepetitives Element aus dem intergenischen Bereich der rRNA-Gene,

zwei Bereiche mit Homologien zu den transposablen CCC-Elementen (Kraemer und Schmidt 1993), einen Bereich des transposablen Elements Cp1 (He et al. 1995) aus der Gruppe der SINEs sowie einen Bereich mit Homologie zum 5'UTR-Bereich einer subtelozentrischen Rekombinase (Rosen et al. 2002). Die genauen Positionen der Elemente im Contig SDR sind in Tabelle 3.2 aufgelistet.

Innerhalb der analysierten 200 kb des Contigs SDR von *C. thummi* wurden insgesamt zehn repetitive bzw. transposable Elemente lokalisiert. Hierbei handelt es sich überwiegend um transposable CCC-Elemente (Kraemer und Schmidt 1993) und TFB1 (Thummi Foldback)-Elemente (Hankeln und Schmidt 1990). Weiterhin sind ein *Hinf(lur)*-Element (Roß 1993) sowie ein TECth1-Element (Wobus et al. 1990) zu finden. Im proximalen Teil des Contigs SDR von *C. thummi* ist zusätzlich ein weiterer Treffer zum 7B Globin-Gen Cluster zu verzeichnen. Die in *C. thummi* gefundene Sequenz befindet sich hierbei im nicht-kodierenden Bereich des Globin-Gen Clusters. Die Lage der repetitiven bzw. transposablen Elemente ist in Tabelle 4.1 aufgelistet. Sechs der zehn gefundenen Elemente befinden sich hierbei innerhalb der zweiten 100 kb des Contigs SDR. In *C. tentans* konnten im analysierten, homologen Bereich der SDR mit einer Größe von 87 kb insgesamt ebenfalls sechs repetitive bzw. transposable Elemente identifiziert werden (vgl. hierzu Tabelle 3.2).

Tabelle 4.1: Lage repetitiver und transposabler Elemente im Contig SDR von *C. thummi*

In der Tabelle sind die im Contig SDR von *C. thummi* lokalisierten repetitiven und transposablen Elemente sowie deren Lage im Contig aufgelistet.

Position (bp)	repetitives Element
2962-4312	TFB1-Element
19136-19244	CCC-Element
27449-27583	7B Globin-Gen Cluster (nicht-kodierender Bereich)
47701-47847	CCC-Element
113430-113951	TFB1 mit CCE-Insertion
114528-114714	<i>Hinf(lur)</i> -Element
161414-161686	TFB1-Element
164492-164673	TECth1-Element
166645-166795	CCC-Element

Alle in *C. tentans* und *C. thummi* identifizierten repetitiven oder transposablen Elemente liegen in Bereichen zwischen den Genen oder innerhalb größerer Introns, welche sich häufig innerhalb der UTR-Bereiche der Gene befinden.

Ein Vergleich der im Contig SDR von *C. thummi* und *C. tentans* identifizierten repetitiven und transposablen Elemente sowie deren Lage innerhalb des Contigs ist in Abbildung 4.1 gezeigt.

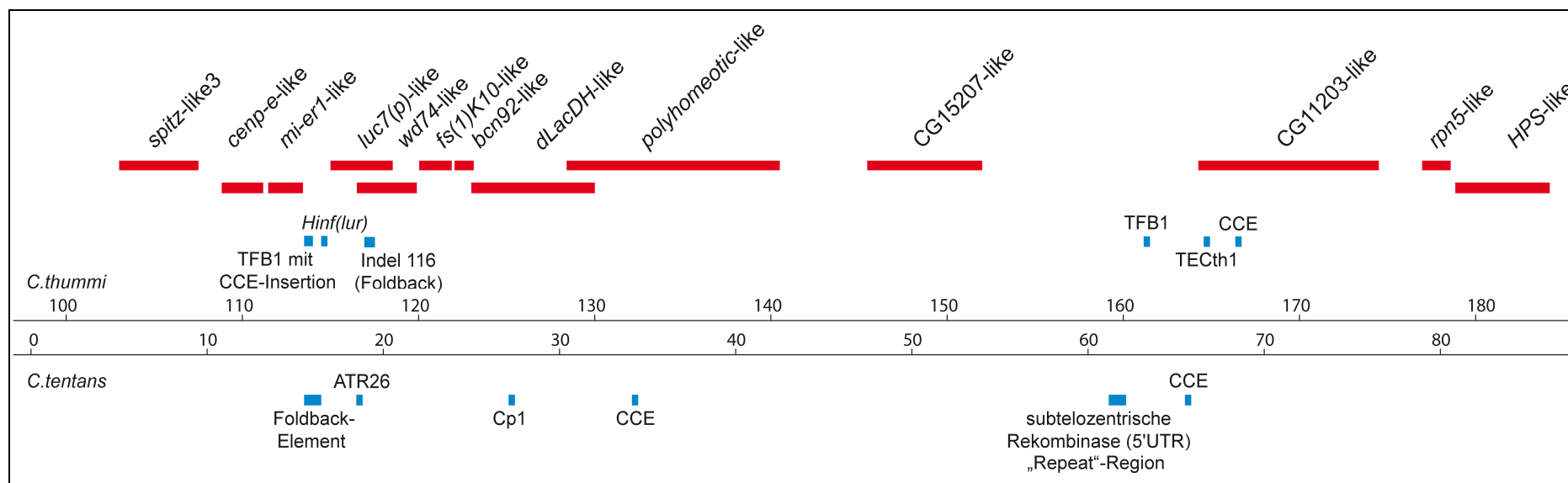


Abbildung 4.1: Vergleich der repetitiven und transposablen Elemente zwischen *C. thummi* und *C. tentans* sowie deren Lage im Contig SDR

Die Abbildung zeigt die im Contig SDR von *C. thummi* und *C. tentans* identifizierten repetitiven bzw. transposablen Elemente. Verglichen wird der sequenzierte und analysierte Bereich aus *C. tentans* mit dem entsprechenden distalen Bereich der SDR von *C. thummi*. Gezeigt sind die repetitiven Elemente innerhalb der SDR der beiden *Chironomus*-Arten (blau) sowie die Lage der analysierten Gene (rot). In der oberen Reihe sind hierbei die auf dem „Watson“-Strang liegenden Gene lokalisiert. In der unteren Reihe sind die auf dem „Crick“-Strang liegenden Gene gezeigt. Die oberen Skalierungsachsen geben die Positionen im Contig SDR von *C. thummi* (obere Achse) sowie von *C. tentans* (untere Achse) an. Alle gezeigten Gene kommen sowohl in *C. thummi* als auch in *C. tentans* vor.

Hierbei fällt auf, dass die gefundenen repetitiven sowie transposablen Elemente in fast allen Fällen unterschiedliche Positionen in den Chromosomenabschnitten beider Spezies besetzen. In keinem Fall sind identische repetitive Elemente in beiden Arten an der gleichen Position zu finden. Ein Abschnitt, welcher in beiden Arten repetitive bzw. transposable Elemente aufweist, ist der Bereich zwischen den transkribierten Bereichen der Gene *mi-er1-like* und *luc7(p)-like*. Innerhalb dieses Bereichs konnte in *C. tentans* eine Indel-Sequenz identifiziert werden, bei der es sich auf Grund von IR- und DR-Strukturen vermutlich um ein transposables Foldback-Element handelt. In *C. thummi* befindet sich innerhalb dieses Abschnitts ein TFB1 mit CCE-Insertion sowie ein *Hinf(lur)*-Element. In Abbildung 4.2 sind zwei Bereiche aus *C. tentans* und *C. thummi* miteinander per DotPlot verglichen, welche in jeweils einer der beiden *Chironomus*-Arten ein repetitives bzw. transposables Element zeigen. Hierbei handelt es sich um den Bereich des Indels 116 (Funke 2011) aus *C. thummi* mit Lage im 3'UTR-Bereich des Gens *wd74-like* sowie einen Abschnitt aus *C. tentans* mit einem BLAST-Treffer zu einem tandemrepetitiven Cluster (5'UTR) der subtelozentrischen Rekombinase aus *C. pallidivittatus*.

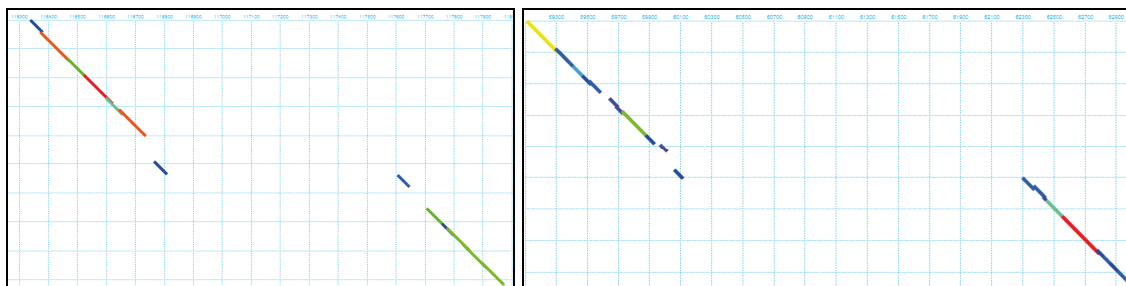


Abbildung 4.2: Vergleiche zweier Abschnitte mit repetitiven bzw. transposablen Elementen des Contigs SDR zwischen *C. thummi* und *C. tentans* per DotPlot

Gezeigt ist ein Vergleich der Bereiche des Indels 116 aus dem 3'UTR-Bereich des Gens *wd74-like* (links) sowie des tandemrepetitiven Bereichs (5'UTR) der subtelozentrischen Rekombinase (rechts) per DotPlot. Die Sequenz von *C. thummi* ist in der linken Abbildung auf der X-Achse abgebildet, die Sequenz von *C. tentans* auf der Y-Achse. Der Bereich der Lücke entspricht dem Indel 116 (116822-117616 bp im Contig *C. thummi*). Hierbei handelt es sich um ein transposables Element aus der Gruppe der Foldback-Elemente. In der rechten Abbildung sind die Achsen vertauscht (X-Achse= *C. tentans*; Y-Achse= *C. thummi*). Der Bereich der Lücke entspricht dem Bereich aus *C. tentans*, für welchen die BLAST-Analysen einen Treffer zu einem tandemrepetitiven Bereich der subtelozentrischen Rekombinase (5'UTR) aus *C. pallidivittatus* zeigen.

Zusätzlich zu den repetitiven und transposablen Elementen wurden in *C. thummi* mit Hilfe von „Skimming“-Analysen Indels gefunden, welche von „Direct repeats“ flankiert werden (Funke 2011). Da diese Repeats „target site“-Duplikationen von transposablen Elementen sein können, ist dies ein weiterer Hinweis auf mögliche Transpositionereignisse innerhalb der geschlechtsdeterminierenden Region. Diese Transpositionereignisse wurden von Steinemann und Steinemann (1992) als eine treibende Kraft während der voranschreitenden Degeneration des Y-Chromosoms beschrieben.

Innerhalb der ersten proximal gelegenen 50 kb des Contigs SDR von *C. thummi* sind durch vorangegangene Arbeiten sechs größere „Inverted repeat“ (IR)-Strukturen identifiziert worden. Eine weitere IR-Struktur liegt im Bereich des Gens CG11203-like innerhalb des ersten Introns. Die meisten der IR-Strukturen flankieren demnach den Bereich der SDR, in welchem sich die meisten Gene befinden. In *H. sapiens* wurden innerhalb der Y-spezifischen Region des Y-Chromosoms (MSY) sowohl acht palindromische sowie weitere nicht-palindromische IR-Strukturen lokalisiert. Diese decken zusammen etwa 25 % der MSY-Region ab und schließen in vielen Fällen Genbereiche mit ein (Skaletsky et al. 2003). Ob auch innerhalb der SDR von *C. tentans* IR-Strukturen wie in den ersten 50 kb des Contigs SDR von *C. thummi* vorhanden sind, kann nicht gesagt werden, da sich dieser Bereich außerhalb der bisher analysierten Region von *C. tentans* befindet. Innerhalb des bislang analysierten Bereichs konnten hingegen keine IR-Strukturen lokalisiert werden. Eine Übersicht mit den Positionen der IR-Strukturen bei *C. thummi* ist im elektronischen Anhang sowie in der Arbeit von Funke (2011) zu finden.

Das Vorkommen von repetitiven und transposablen Elementen in *C. tentans* und *C. thummi*, die in *C. thummi* vorhandenen „Direct repeat“-Strukturen an den Rändern der analysierten Indels, welche auf Transpositionereignisse hindeuten, sowie die IR-Strukturen innerhalb der proximalen 50 kb des Contigs SDR von *C. thummi* sind Indizien für eine Anhäufung von Transposons und damit für eine fortschreitende Degeneration des Proto-Y-Chromosoms von *Chironomus*.

In *C. thummi* befinden sich innerhalb der zweiten 100 kb des Contigs SDR 16 der insgesamt 20 identifizierten Gene. Im homologen Bereich der SDR von *C. tentans* mit einer Größe von 87 kb konnten die ersten 14 dieser 16 Gene ebenfalls gefunden werden. Im distalen Bereich um das Gen *CtY* liegen in *C. tentans* bisher keine Sequenzen vor. Ein Vergleich der beiden Contigs SDR von *C. tentans* und *C. thummi* zeigt, dass die Abfolge der Gene innerhalb der Contigs identisch ist. Dies untermauert die Ergebnisse der FISH (Kapitel 3.1.3), mit deren Hilfe die Lage der BAC-Klone aus *C. tentans* innerhalb des Contigs von *C. thummi* und in relativer Position zum hemizygoten *Cl*-Element Cluster untersucht wurde. Auch die Ergebnisse der bisher sequenzierten Bereiche der SDR von *C. piger* (Blumensatt 2009 sowie Stiehl persönliche Mitteilung) bestätigen die Abfolge der Gene innerhalb des Contigs SDR. Innerhalb des Contigs SDR ist eine vergleichsweise hohe Dichte an Genen zu verzeichnen. Betrachtet man den distalen Bereich des Contig SDR von *C. thummi* mit 16 Genen, so enthält dieser im Vergleich zum proximalen Bereich bereits die 4-fache Anzahl an Genen. Hierbei liegen die identifizierten Genbereiche, vor allem in den Regionen von 115-140 kb und 173,5-196,5 kb in *C. thummi* sowie 16,5-42,5 kb und 78,5-86,5 kb in *C. tentans* eng aneinander. Rein rechnerisch ergibt sich innerhalb des distalen Bereichs eine Gendichte von 1 Gen pro 6 kb. Hierbei ist die durchschnittliche Länge eines Gens nicht berücksichtigt worden. Auch für *C. tentans* ergibt sich ein errechneter Wert von 1 Gen pro 6 kb-Abschnitt des analysierten Contigs SDR. Die Übereinstimmung in den Werten der Gendichte zwischen den Arten *C. tentans* und *C. thummi* verwundert nicht weiter, da bereits gezeigt wurde, dass die Abfolge der Gene und somit die Struktur des Contigs SDR große Homologien zwischen den beiden

Arten zeigt. Die in der Literatur angegebene Gendichte von *D. melanogaster* liegt bei etwa 1 Gen pro 12-12,5 kb, die Gendichte von *H. sapiens* sogar bei 1 Gen pro 125-160 kb (Munk 2010; Watson et al. 2011). Diese Werte sind gemittelte Werte und können nur Tendenzen angeben. Allerdings macht der Vergleich deutlich, dass sich innerhalb des analysierten Bereichs der SDR von *C. tentans* sowie des homologen Bereichs aus *C. thummi* eine erhöhte Anzahl von Genen befindet. Zwischen den Genbereichen des „Watson“- und des „Crick“-Stranges kommt es in einigen Fällen sogar zu Überlappungen. Diese Überlappungen betreffen in allen Fällen nur die UTR-Bereiche der Gene. Zu einer Überlappung der kodierenden Bereiche zwischen den beiden Strängen kommt es in keinem Fall. Ob ein Zusammenhang zwischen der erhöhten Gendichte und der unterdrückten Rekombination innerhalb des geschlechtsdeterminierenden Bereichs besteht, kann nicht mit Sicherheit gesagt werden.

Als ein weiterer Schritt innerhalb des Prozesses der Degeneration des Proto-Y-Chromosoms wird neben dem vermehrten Vorkommen von transposablen Elementen (Steinmann und Steinmann 1993; Steinmann et al. 1993) das Vorkommen von Deletionen und Duplikationen sowie eine Anreicherung repetitiver Satelliten-DNA innerhalb der geschlechtsdeterminierenden Region beschrieben (Steinmann und Steinmann 1998). Neben den identifizierten Indel-Strukturen mit flankierenden DRs in *C. thummi*, welche auf Transpositionereignisse hindeuten (Funke 2011) konnte innerhalb der SDR von *C. thummi* ein duplizierter Bereich nachgewiesen werden. Die Duplikation konnte im Bereich der Gene *wd74-like*, *fs(1)K10-like*, *dLacDH-like* lokalisiert werden und kommt nur in *C. thummi* ♂ sowie *C. piger* ♂ vor. Erste Hinweise auf diese männchenspezifische Duplikation lieferte ein genomischer Southern-Blot nach Hybridisierung mit einer Sonde aus dem Bereich des Gens *fs(1)K10-like* (Stiehl 2009). Die DNA männlicher Tiere zeigte neben den erwarteten Banden zusätzlich noch weitere männchenspezifische Banden. Eine weitere Hybridisierung unter Verwendung einer *dLacDH*-Sonde zeigte als Ergebnis ebenfalls zusätzliche, männchenspezifische Banden. Mit Hilfe einer Inversen-PCR sowie der „Genome Walk“-Strategie konnte eine männchenspezifische Sequenz des Contigs SDR von *C. thummi* und *C. piger* erhalten werden (Stiehl 2009, Stiehl persönliche Mitteilung, Kraemer persönliche Mitteilung). Innerhalb dieser Sequenz konnte auch das proximale Ende eines *Cla*-Element Clusters nachgewiesen werden. Dieses konnte zuvor nur mit Hilfe der FISH auf einen Bereich von etwa 20 kb zwischen 115-135 kb des Contigs SDR eingegrenzt werden (Herold 2008). Hierzu erfolgten Doppelhybridisierungen von bekannten BAC-Klonen der SDR von *C. thummi* zusammen mit repetitiven *Cla*-Elementen an Polytänchromosomen männlicher *C. thummi*-Larven. Innerhalb der geschlechtsdeterminierenden Region von *C. thummi* ♂ sind *Cla*-Elemente Proto-Y-spezifisch in Form eines hemizygoten Clusters vorhanden (Kraemer und Schmidt 1993).

Als Teil der Proto-Y-spezifischen Sequenz konnte eine Duplikation des Gens *fs(1)K10-like* sowohl genomische als auch als cDNA nachgewiesen werden (Kraemer, persönliche Mitteilung). Mit Hilfe von RT-PCR konnte im Rahmen dieser Arbeit ebenfalls bestätigt werden, dass es sich bei beiden Genbereichen der Duplikation um

transkribierte Bereiche handelt (siehe Kapitel 3.2). Beide Transkripte weisen durchgehende funktionelle ORFs auf. Hierbei besitzt die Y1-Variante mit 372 AS im Gegensatz zur Y2-Variante mit 344 AS den längeren ORF. Die beiden abgeleiteten AS-Sequenzen zeigen im N-terminalen Bereich nur wenige Unterschiede. Im C-terminalen Bereich hingegen unterscheiden sich die beiden Varianten klar ab AS-Position 291.

In *C. piger* konnte die in *C. thummi* beschriebene Duplikation in diesem Bereich ebenfalls mit Hilfe genomischer Southern-Blots (Stiehl 2009) und durch Sequenzierung nachgewiesen werden. Hierzu wurde wie bei *C. thummi* die „Genome Walk“-Strategie angewendet (Kraemer, persönliche Mitteilung). Beide Varianten (Y1 und Y2) der Duplikation werden transkribiert, was mit Hilfe von RT-PCR und RACE bestätigt werden konnte (Stiehl, persönliche Mitteilung).

In *C. tentans* konnte hingegen weder mit Hilfe von PCR mit genomischer DNA als Template unter Verwendung männchenspezifischer Primer aus *C. thummi*, noch mit genomischen Southern-Blots im Bereich der Gene *fs(1)K10*-like (siehe Kapitel 3.1.4) und *dLacDH*-like eine Duplikation nachgewiesen werden. Allerdings konnten auf genomischen Southern-Blots nach Hybridisierung mit Sonden aus den benachbarten Genbereichen *mi-er1*-like und *luc(7)p*-like neben den erwarteten noch weitere Banden mit einer Größe von 5,2-5,6 kb in den Spuren der Restriktionsverdau mit *EcoRI*, *HindIII* sowie *NsiI* festgestellt werden (siehe Kapitel 3.1.4). Mit Hilfe von PCRs mit genomischer DNA als Template konnte ein 858 bp großer Indel-Bereich (15811-16668 bp des Contigs SDR von *C. tentans*) identifiziert werden. Dieser zeigt im Randbereich 20 bp lange DRs sowie 4 bp lange IRs. Dies könnte darauf hindeuten, dass es sich bei diesem analysierten Indel um ein transposables Foldback-Element handelt. Ob die Indel-Sequenz geschlechtsspezifisch auftritt, war für *C. tentans* nicht zu klären, da nur DNA aus Zellkulturzellen zur Verfügung stand (Geschlecht nicht bekannt; wahrscheinlich ♀:♂ 50:50).

In *C. thummi* konnten in beiden Geschlechtern auf Autoradiogrammen genomischer Southern-Blots nach Hybridisierung mit einer Sonde aus dem Bereich des Gens *luc7(p)*-like zusätzliche Signale verzeichnet werden. Die Intensität der zusätzlichen Signale stimmt, genau wie in *C. tentans* mit der Intensität der erwarteten Signale überein. Auch in ihrer Größe ähneln die Signale aus *C. thummi* denen von *C. tentans*.

In *C. piger* sind im Bereich des Gens *luc7(p)*-like in beiden Geschlechtern keine zusätzlichen Hybridisierungssignale zu verzeichnen (Stiehl, persönliche Mitteilung). Aus dem Bereich des Gens *mi-er1*-like liegen bislang weder aus *C. thummi* noch aus *C. piger* Autoradiogramme vor.

Auf Grund der in *C. thummi* in beiden Geschlechtern nachgewiesenen zusätzlichen Hybridisierungssignale sowie der in *C. thummi* und *C. tentans* vergleichbaren Intensitäten der zusätzlichen Signale zu den per Southern-Analyse erwarteten Signale kann demnach davon ausgegangen werden, dass es sich bei dem Indel um einen normalen, geschlechtsunspezifischen Polymorphismus handelt.

In *C. thummi* liegen im Bereich zwischen den Genen *mi-er1*-like und *luc7(p)*-like, in welchem sich in *C. tentans* der Indel-Bereich befindet, die transposablen Elemente CCE, TFB1 und *Hinf(lur)* (vgl. Abbildung 4.1). Da transposable Elemente in *C. piger* im

Vergleich zu *C. thummi* in deutlich verminderter Häufigkeit vorkommen, kann dies unter Umständen auch die fehlenden zusätzlichen Hybridisierungssignale im Bereich des Gens *luc7(p)*-like in *C. piger* erklären.

Im Bereich der Gene *mi-er1*-like und *luc7(p)*-like des Contigs SDR von *C. thummi* liegen zwei der in Funke (2011) beschriebenen Indels, welche in den flankierenden Bereichen „Direct repeats“ aufweisen (I114 sowie I116). Im Bereich des Indels 114 (114525 bp) ist im Contig SDR von *C. thummi* ein *Hinf*(*lur*)-Element lokalisiert. Hierbei handelt es sich bei der Klongruppe, aus welcher sich die Sequenz des Contigs zusammensetzt, in diesem Fall um eine Deletion. Die Sequenz der Insertion ist somit im Contig nicht repräsentiert. Die zweite Indel-Sequenz (I116) dieser Region ist im Bereich von 116822-117616 bp des Contigs SDR von *C. thummi* zu finden. Hierbei handelt es sich um eine Insertion, bei welcher die Sequenz der Insertion in der Konsensussequenz des Contigs SDR vertreten ist. Der Indel-Bereich 116 liegt somit innerhalb des 3'UTR-Bereichs des auf dem „Crick“-Strang liegenden Gens *wd74*-like (siehe auch Kapitel 3.2). Innerhalb des auf dem „Watson“-Strang liegenden Gens *luc7(p)*-like, welches im 3'UTR Bereich mit dem Gen *wd74*-like überlappt, befindet sich der Indel-Bereich 116 im vierten Intron (siehe Abbildung 4.3). Eine direkte Auswirkung auf das Transkript oder den kodierenden Bereich von *luc7(p)*-like ist somit unwahrscheinlich.

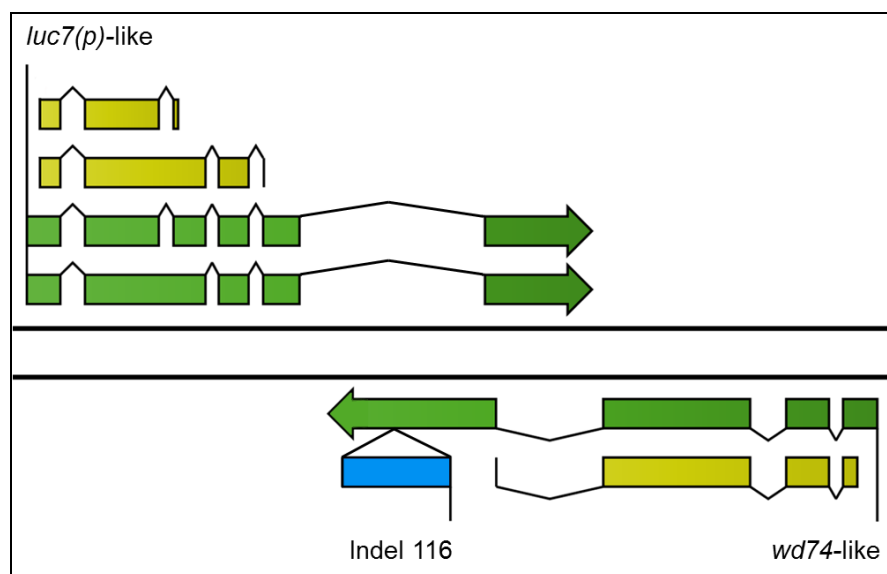


Abbildung 4.3: Überschneidung der transkribierten Bereiche von *luc7(p)*-like und *wd74*-like sowie Lage des Indels 116 im 3'UTR-Bereich der beiden Gene.

Die Abbildung zeigt die Überlappung der transkribierten Bereiche (grün) der beiden Gene *luc7(p)*-like („Watson“-Strang) und *wd74*-like („Crick“-Strang). Der proteinkodierende Bereich der beiden Gene (gelb) überlappt hierbei nicht. Die Lage des Indels 116 ist in beiden Genen auf den 3'UTR-Bereich lokalisierbar, wobei der Indel-Bereich im Fall des Gens *luc7(p)*-like sogar innerhalb des Intronbereichs liegt.

Die Sequenz des Indels 116 weist lange terminale „Inverted repeat“-Strukturen auf (Funke 2011). Zusammen mit den flankierenden kurzen „Direct repeats“ sowie dem

Fehlen des Transposasegens deutet dies auf ein transposables Element der „Foldback“-Gruppe hin (Potter et al. 1980; Truett et al. 1981; Potter 1982). In *C. thummi* sowie *C. piger* ist aus dieser Gruppe das TFB1-Element beschrieben. Dieses ist an weitreichenden genomischen Veränderungen beteiligt, welche während der unterschiedlichen Evolution der Schwesternarten *C. thummi* und *C. piger* stattgefunden haben (Hankeln und Schmidt 1990). Hierbei sind TFB1-homologe Sequenzen in *C. thummi* über das ganze Genom verteilt. In *C. piger* hingegen sind diese hauptsächlich in den zentromernahen Regionen zu finden und in ihrer Gesamtheit viel schwächer vertreten als in *C. thummi*.

Alle dargestellten Ergebnisse stimmen mit den in *D. miranda* gezeigten Ereignissen während der fortschreitenden Degeneration des Proto-Y-Chromosoms überein (Steinemann und Steinemann 1990; 1992; Steinemann et al. 1993; Steinemann und Steinemann 1998). So konnten sowohl in *C. tentans* wie auch in *C. thummi* Bereiche mit Homologien zu transposablen Elementen gefunden werden. Weiterhin wurden in *C. thummi* „Inverted repeat“-Strukturen gefunden, welche sich im proximalen Bereich des Contigs SDR über rund 50 kb erstrecken. Ebenfalls in *C. thummi* konnte das proximale Ende des *Cla*-Element Clusters innerhalb des Contigs SDR lokalisiert werden (Stiehl 2009). Hierbei handelt es sich sehr wahrscheinlich um das in *C. thummi* ♂ mittels FISH als hemizygot erkennbare Cluster tandemrepetitiver DNA, welches eng mit dem dominanten Faktor *M* gekoppelt ist. Dieses ist in *C. tentans* sowie in *C. piger* in dieser Form nicht nachweisbar. Die in den beiden Arten *C. thummi* und *C. piger* gefundenen Duplikationen des Bereichs um das Gen *fs(1)K10*-like deuten auf eine Umorganisation des Proto-Y-Chromosoms hin. Auch die in *C. tentans* gefundenen zusätzlichen Hybridisierungssignale auf den Autoradiogrammen der genomischen Southern-Blots deuten auf eine Duplikation innerhalb des Contigs SDR von *C. tentans* im Bereich der Gene *mi-er1*-like sowie *luc7(p)*-like hin.

4.2 Die Gene der SDR

4.2.1 Analysen der RACE-Ergebnisse

4.2.1.1 Gene mit verschiedenen 5'UTR-Varianten

Innerhalb des analysierten Bereichs der SDR von *C. thummi* wurden im Rahmen der Arbeit 12 Gene per RACE-Verfahren näher untersucht. In drei Genen konnten mindestens zwei unterschiedliche 5'UTR-Bereiche identifiziert werden. Hierbei handelt es sich um die Gene *bcn92*-like, *polyhomeotic*-like sowie *hermanky pudlak*-like.

Da Sequenzen innerhalb der Promotor- und 5'UTR-Bereich für die Regulation der Transkription und Translation eine Rolle spielen (Lahiri et al. 2003; Szymanowska et al. 2004) könnten die unterschiedlichen 5'UTR-Varianten zu Veränderungen der Expressionstärke des Gens oder zu einer veränderten Translation führen.

Das Gen *bcn92*-like weist sowohl in *C. thummi* ♀ wie auch in *C. thummi* ♂ zwei unterschiedliche 5'UTR-Varianten auf. Diese unterscheiden sich in ihrer Länge um

72 bp. Der proteinkodierende Bereich des Gens ist durch die zwei unterschiedlichen Varianten nicht betroffen. Bei der längeren 5'UTR-Variante kann die gefundene CAAT-Box auf Grund ihrer relativen Nähe zum vermuteten Transkriptions-Startpunkt (~35 bp) als regulatorisches Element ausgeschlossen werden.

Im Gen *hermansky pudlak-like* konnten ebenfalls zwei 5'UTR-Varianten gefunden werden. Hierbei liegt die kürzere Variante sowohl in *C. thummi* ♀ wie auch in *C. thummi* ♂ vor. Die längere Variante konnte nur in *C. thummi* ♂ per RACE gefunden werden. Ob diese Variante allerdings tatsächlich existiert, bleibt fraglich, da RT-PCRs zur Bestätigung dieser Variante bislang auch in *C. thummi* ♂ keine Produkte lieferten. Auch die Daten der Hochdurchsatz-Sequenzierung sprechen eher gegen diese längere Transkriptvariante, da in einem Bereich von ~85 bp nach der Intron-Exon-Grenze praktisch keine „Reads“ zu finden sind. Ein Vergleich der RACE-Sequenz mit dem Contig SDR aus *C. thummi* liefert allerdings keine Anzeichen auf ein Fehlpriming eines Primers. Des Weiteren sind die Intron-Exon-Grenzen innerhalb der Sequenz klar definierbar. Der proteinkodierende Teil des Gens ist wie bei *bcn92-like* auch hier nicht von den unterschiedlichen 5'UTR-Varianten betroffen.

Das Gen *polyhomeotic-like* zeigt ebenfalls zwei 5'UTR-Varianten (Exon 1a und 1b). Diese konnten bisher nur in *C. thummi* ♂ vollständig analysiert werden. Allerdings konnte per Intron-überspannender RT-PCRs sowie mit Hilfe der Daten der Hochdurchsatz-Sequenzierung gezeigt werden, dass diese beiden Varianten auch in *C. thummi* ♀ vorliegen müssen. Die beiden Varianten werden durch Introns von 7206 bp (Exon 1a) bzw. 1269 bp (Exon 1b) vom zweiten Exon getrennt. Der proteinkodierende Teil des Gens liegt in der Variante mit Exon 1a mit einer Länge von 3183 Basen vor. Im Transkript mit Exon 1b kann es zu einer Verlängerung des proteinkodierenden Abschnitts kommen. Hierbei wurde im Exon 1b ein zusätzliches Startcodon gefunden. Die hieraus resultierende Sequenzlänge des proteinkodierenden Bereichs beträgt 3483 Basen und somit 100 AS mehr als im Transkript mit Exon 1a. Dies ist insbesondere deshalb interessant, da im Bereich der AS 10-26 ein potentiell NLS erkannt werden konnte. Innerhalb der kürzeren AS-Sequenz wurde hingegen kein NLS identifiziert. Aus welchem Grund bei *polyhomeotic-like* die beiden 5'UTR-Varianten vorliegen und ob diese differentiell reguliert werden, oder sogar unterschiedliche Funktionen aufweisen, kann zum jetzigen Zeitpunkt jedoch nicht gesagt werden.

4.2.1.2 „Scheinbar“ multiple 3'UTR-Transkriptlängen durch Fehlpriming des Oligo-d(T)-Primers

Mit Hilfe der RT-PCR und RACE ist es im Rahmen dieser Arbeit gelungen, von fast allen Genen der SDR von *C. thummi* Vollängen-Transkripte zu erhalten (siehe Kapitel 3.2). Hierbei fällt auf, dass in mehreren Fällen unterschiedlich lange PCR-Produkte im 3'UTR-Bereich der Gene amplifiziert wurden. Dies ist bei den Genen *luc7(p)-like*, *wd74-like*, *fs(1)K10-like*, *bcn92-like*, *polyhomeotic-like* sowie *hermansky pudlak-like* der Fall. Da diese Anzahl an PCR-Produkten auch das Resultat von Fehlpriming des Oligo-d(T)-Primers während der cDNA-Synthese sein kann, wurden die erhaltenen 3'UTR-Enden noch einmal näher untersucht. Ein Fehlpriming des Oligo-d(T)-Primers ist in *Chironomus* auf Grund des extrem hohen AT-Gehalts durchaus wahrscheinlich. Der AT-Gehalt des Contigs SDR von *C. thummi* mit einer Größe von etwa 200 kb und 20 Genbereichen beträgt etwa 69 %. Der AT-Gehalt des untersuchten Contigs SDR von *C. tentans* mit 87 kb liegt knapp darunter bei etwa 68 %. Der AT-Gehalt der 3'UTR-Bereiche der analysierten Gene in *C. thummi* liegt mit 71-85 % (gemittelt: 77 %) nochmals über den Durchschnittswerten des gesamten Contigs. Für die Überprüfung der 3'-Enden wurden diese erneut per RT-PCR mit einer Primerkombination überprüft. Hierbei liegt ein Primer in kurzer Distanz vor dem 3'-Ende des Transkripts, der zweite Primer liegt vor dem letzten Intron des Transkripts. Auf Grund der Größe des Transkripts sowie dessen Sequenz kann somit ausgeschlossen werden, dass es sich um ein genomisches PCR-Produkt handelt, welches durch DNA-Kontamination innerhalb der zur cDNA-Synthese verwendeten RNA entstanden ist. Allerdings kann mit dieser Methode immer nur das längste Transkript überprüft werden. Kann dieses bestätigt werden, können alle kürzeren Transkripte durch Fehlpriming innerhalb der Sequenz entstanden sein. Aus diesem Grund wurden die 3'-Enden der RACE-Produkte neben der Überprüfung per RT-PCR auch auf die Anzahl der Adenin-Basen innerhalb der 20 bp hinter dem Ende des PCR-Produkts hin untersucht. Liegt in diesem Bereich ein erhöhter Anteil an Adenin vor und ist weiterhin kein Poly-Adenylierungssignal (AATAAA oder leicht abweichend (Retelska et al. 2006)) zu finden, so kann mit hoher Sicherheit von einem Fehlpriming des Oligo-d(T)-Primers innerhalb der Sequenz ausgegangen werden. In Tabelle 4.2 sind die gefundenen Poly-Adenylierungs-Signale der im Rahmen der Arbeit analysierten Gene der SDR von *C. thummi* zusammengefasst. Im oberen Teil der Tabelle sind hierbei die sechs Gene aufgelistet, bei denen innerhalb des 3'UTR-Bereichs mehrere RACE-Produkte vorlagen. Bei den verkürzten Produkten konnte jedoch in fast allen Fällen kein PolyA-Signal identifiziert werden, so dass von einem Fehlpriming des Oligo-d(T)-Primers innerhalb des 3'UTR-Bereichs ausgegangen werden muss. Angegeben sind in der Tabelle somit lediglich die längsten analysierten RACE-Produkte inklusive des Abstandes des PolyA-Signals zum Ende des Transkripts. Weiterhin sind die Positionen des PolyA-Signals sowie der vermeintlich letzten Base des Transkripts im Contig SDR von *C. thummi* angegeben. Im unteren Teil der Tabelle sind die vier Gene gelistet, bei denen nach durchgeführter RACE nur ein Produkt vorlag.

Tabelle 4.2: Überprüfung der 3'-RACE-Produkte auf PolyA-Signale

Die Tabelle gibt einen Überblick über den Abstand des PolyA-Signals zum durch RACE lokalisierten vermutlichen Transkriptende. Weiterhin angegeben ist die genaue Position des PolyA-Signals im Contig SDR von *C. thummi* sowie die Position des per RACE erhaltenen vermutlichen Ende des Transkripts.

Gen	Pos. PolyA-Signal vor Transkriptende	Pos. PolyA-Signal im Contig SDR	Ende Transkript in Contig SDR
<i>luc7(p)</i> -like	(-33..-27 bp)	118589-118794 bp	118821 bp
<i>wd74</i> -like	(-32..-26 bp)	116573-116578 bp	116547 bp
<i>fs(1)K10</i> -like	(-25..-19 bp)	122334-122339 bp	122358 bp
<i>bcn92</i> -like	(-26..-20 bp)	123377-123382 bp	123402 bp
<i>polyhomeotic</i> -like	(-24..-18 bp)	139828-139833 bp	139851 bp
<i>hermansky pudlak</i> -like	(-46..-40 bp)	175527-175532 bp	175487 bp
<i>mi-er1</i> -like	(-45..-39 bp)	111216-111221 bp	111177 bp
<i>dlacDH</i> -like	(-27..-21 bp)	123217-123222 bp	123196 bp
CG11203-like	(-19..-13 bp); AATACA	171057-171062 bp	171075 bp
<i>rpn5</i> -like	(-26..-20 bp)	175430-175435 bp	175455 bp

In allen analysierten Transkripten konnte ein PolyA-Signal innerhalb der Sequenz ermittelt werden. Bei den beiden Genen *mi-er1*-like sowie *hermansky pudlak*-like liegt das lokalisierte PolyA-Signal leicht ausserhalb des Bereichs (-38..-24 bp), in welchem es für gewöhnlich zu finden ist (Retelska et al. 2006). In allen anderen Sequenzen ist das PolyA-Signal (AATAAA) innerhalb dieses Bereichs zu finden. Im Fall des Gens CG11203-like konnte im Bereich (-19..-13 bp) ein alternatives PolyA-Signal (AATACA) gefunden werden. Dieses kommt in *Drosophila* mit einer Frequenz von etwa 2,9 % ebenfalls vor. Das am häufigsten vorkommende PolyA-Signal (AATAAA) besitzt in *Drosophila* eine Frequenz von etwa 47,8 % (Retelska et al. 2006).

4.2.2 Analyse und Vergleich der Promotorbereiche der Gene der SDR aus *C. thummi*

Nach erfolgter Durchführung der RACE-Versuche lagen von 10 der 11 mittels RACE näher untersuchten Gene von *C. thummi* 5'UTR-Daten vor. Einzige Ausnahme ist das Gen CG11203-like, von dem weder in *C. thummi* ♀, noch in *C. thummi* ♂ Daten aus dem 5'UTR-Bereich gewonnen werden konnten.

Bei den meisten Genen der SDR konnten in beiden Geschlechtern Daten aus dem 5'UTR-Bereich erhalten werden. Lediglich im Gen CG15207-like sowie den beiden *polyhomeotic*-like Transkripten (Exon 1a und 1b) liegen keine 5'UTR-Daten aus *C. thummi* ♀ vor. Bei *fs(1)K10*-like liegen zwar Daten aus dem 5'UTR-Bereich vor. Hierbei kann allerdings auf Grund der Duplikation des Gens in *C. thummi* ♂ nicht mit Sicherheit gesagt werden, ob es sich bei der Sequenz um eine X-chromosomale oder eine Y-chromosomale handelt. Lediglich die Y₂-Variante des Gens kann nach der Sequenzanalyse eindeutig ausgeschlossen werden.

Mit Hilfe der vorliegenden Daten wurden Sequenzvergleiche zwischen *C. thummi*, *C. tentans* und *C. piger* vorgenommen. Hierbei wurden genomische Bereiche analysiert, welche sich im Bereich von -500 bp bis +100 bp zum bisherigen Ende der 5'-RACE Produkte befinden. Die Analysen konzentrierten sich hierbei vor allem auf den Kernpromotorbereich (-50 bp bis +50 bp um den Transkriptions-Startpunkt) sowie den proximalen Promotorbereich (-200 bp bis -50 bp stromaufwärts vom Transkriptionsstartpunkt). Die in diesem Bereich gesuchten Elemente sowie deren in der Literatur angegebenen Positionen (Juven-Gershon et al. 2008) sind in Tabelle 3.4 aufgeführt. Zu Beginn wurde nach möglichen Initiatorregionen (INR) der Gene gesucht. Hierzu wurde der genomische Bereich der Gene näher betrachtet, in welchem die Sequenzen der 5'-RACE-Produkte beginnen. Ausgehend von diesen Daten wurde im Folgenden nach Motiven wie der TATA-Box sowie den BREs („TFIIB recognition element“) oberhalb und unterhalb der TATA-Box gesucht. Ein Vergleich der Sequenzen der drei Arten *C. thummi*, *C. piger* und *C. tentans* sowie der Abstände von INR, TATA-Box, BRE führten in den meisten Fällen zur Identifizierung eines vermutlichen Startpunktes der Transkription. In einigen Fällen konnte jedoch nicht nur ein potentieller Startpunkt angegeben werden. In diesen Fällen wurden oft mehrere mögliche Sequenzmotive gefunden, welche als INR in Frage kommen. In einigen Fällen besteht hierbei eine Diskrepanz zwischen der Sequenz, welche am nächsten zum vorhandenen Sequenzende der 5'-RACE liegt und einer weiter aufwärts liegenden Sequenz, welche in ihrem Abstand besser mit weiteren Kern-Promotor-Elementen übereinstimmt. Das für die Suche verwendete Sequenzmotiv der INR aus *Drosophila* (TCAKTY) (Ohler et al. 2002; FitzGerald et al. 2006) konnte lediglich in den Promotorbereichen von *polyhomeotic*-like (Exon 1a) sowie *polyhomeotic*-like (Exon 1b) aus *C. thummi* ♂ vollständig identifiziert werden. Alle weiteren INR-Sequenzen liegen in einer, im Vergleich zur Sequenz aus *Drosophila*, abgewandelten Form vor. Hierbei konnte im Kernbereich der INR-Sequenzen das Basentriplett CAG (7x), CAT (8x), CAC (4x) sowie CAA (2x) identifiziert werden. Bei den beiden gefundenen CAA-Sequenzen fällt auf, dass auf das zweite Adenin jeweils ein Guanin folgt. Das

Thymin vor dem Tripletts konnte in 12 von 21 Fällen gefunden werden, das Thymin nach dem Tripletts in 10 von 21 Fällen. Somit kann als Konsensussequenz für die INR der SDR von *C. thummi* die Basenfolge (T)CAN(T) angegeben werden. Als Base für den Transkriptionsstart konnte, wie allgemein üblich, auch hier die Purin-Base Adenin identifiziert werden.

In vielen Kernpromotorbereichen der Gene der SDR konnten TATA-Boxen oder TATA-Box-ähnliche Sequenzen identifiziert werden. Ausgehend von der ersten Position des Motivs liegen die gefundenen Abschnitte im Bereich von -22 bis -48 bp in Bezug auf den vermuteten Startpunkt der Transkription (+1). Die Bereiche der gefundenen TATA-Boxen oder TATA-Box-ähnlichen Sequenzen liegen somit teilweise leicht über dem in der Literatur angegebenen Wert von 26 bis 34 Nukleotiden (Breathnach und Chambon 1981). Auch in Juven-Gershon (2008) wird die Position der ersten Base der TATA-Box auf die Positionen -30 oder -31 Nukleotide vor dem Startpunkt der Transkription angegeben. Die teilweise von den Literaturwerten nach oben abweichenden Abstände der TATA-Box wurden für *Chironomus* bereits in der Arbeit von Kraemer (1994) für das ebenfalls in der SDR liegende Gen CtY beschrieben. Hierbei liegt die TATA-Box 38 Nukleotide vom Transkriptionsstartpunkt entfernt. Weiterhin konnten in Histon- und Hämoglobingenen von *Chironomus* ebenfalls leicht erhöhte Abstände von 31-35 Nukleotiden verzeichnet werden (Rozynek et al. 1989; Hankeln und Schmidt 1991). Ob zwischen dem Abstand der TATA-Box und dem Expressionslevel der Gene ein genereller Zusammenhang besteht, ist unklar. Allerdings liegt z.B. die identifizierte TATA-Box des schwach exprimierten Gens CG15207-like mit -24 bp recht nah am vermuteten Startpunkt der Transkription. Die in der Literatur oft angegebene Konsensussequenz der TATA-Box (TATAWAWR) konnte in keinem Promotorbereich der Gene der SDR gefunden werden. Ob es sich bei allen im Promotorbereich gefundenen TATA-Box-ähnlichen Sequenzen tatsächlich um genutzte TATA-Boxen handelt, ist ebenfalls unklar. Speziell im distalen Bereich, in welchem von *C. piger* keine Sequenzen mehr vorliegen, konnte eine TATA-Box-ähnliche Sequenz häufig nur in *C. thummi* identifiziert werden. Yang et al. (2007) konnten zeigen, dass in *H. sapiens* lediglich 24 % der Gene eine TATA-Box-ähnlichen Sequenzen im Kernpromotorbereich aufweisen. Des Weiteren weisen nur 10 % dieser Sequenzen hierbei tatsächlich das vollständige Motiv der Konsensussequenz auf. Demnach ist es durchaus vorstellbar, dass nicht alle gefundenen TATA-Box-ähnlichen Strukturen auch tatsächlich als TATA-Box fungieren müssen.

In einigen Kernpromotorbereichen der Gene der SDR von *C. thummi* konnten BREs („TFIIB recognition element“) identifiziert werden. Das BRE^u (upstream) wurde ursprünglich in einem Teil der Gene mit TATA-Box als direkt vor der TATA-Box liegendes Element identifiziert (Lagrange et al. 1998). Später wurden auch Elemente gefunden, welche direkt nach der TATA-Box lokalisierbar sind. Diese wurden im Folgenden als BRE^d (downstream BRE) bezeichnet (Deng und Roberts 2005). Die Suche nach BRE^u-Elementen innerhalb der Promotorbereiche erbrachte keinen Treffer. Das Sequenzmotiv des BRE^d konnte hingegen in 7 von 21 Fällen identifiziert werden. Doch lediglich in den Promotorbereichen von *mi-er1*-like, *bcn92*-like sowie CG11203-like konnte die Sequenz im direkten Anschluß an die Sequenz der

TATA-Box erkannt werden. In wie weit ein Abstand von einigen Nukleotiden Abstand zur TATA-Box die Funktion des BRE beeinflusst, ist unklar.

Ein weiteres Kernpromotor-Element, das DPE („downstream core promoter element“) ist in seiner Sequenz von *Drosophila* bis zu *H. sapiens* hochkonserviert. Es handelt sich hierbei um eine Erkennungssequenz des TFIIID, welche für eine basale Aktivität der Transkription wichtig ist (Burke und Kadonaga 1996). Die Konsensussequenz des DPE (RGWYVT) mit einer Position von +28 bis +33 relativ zum Startpunkt der Transkription (Kutach und Kadonaga 2000) konnte in *mi-er1-like*, *fs(1)K10-like* sowie *polyhomeotic-like* gefunden werden. Allerdings fehlt in allen Sequenzen bis auf *mi-er1-like* ($L_{\text{♀}}$, $L_{\text{♂}_2}$) das distale Thymin. Weiterhin schwanken die Abstände zum Startpunkt der Transkription von +10 zu +41 Nukleotide. Lediglich in *mi-er1-like* ($L_{\text{♀}}$, $L_{\text{♂}_2}$) sowie *polyhomeotic-like* (Exon 1b) konnte der in der Literatur angegebene Abstand von +27 bzw. +28 Nukleotiden gemessen werden.

Innerhalb des proximalen Promotorbereichs (-50 bis -200 Nukleotide) wurde nach den Motiven der GC-Box und der CAAT-Box gesucht. In vielen Promotorbereichen wurde hierbei die Kern-Sequenz der CAAT-Box (CAAT) gefunden. In den Genen *mi-er1-like* *wd74-like*, *polyhomeotic-like* (Exon 1b) sowie *rpn5-like* konnten keine CAAT-Boxen gefunden werden. Die vollständige Sequenz der CAAT-Box (CCAAT) konnte in den Genen *fs(1)K10-like*, *dLacDH-like* sowie *polyhomeotic-like* (Exon 1a ♂_1) gefunden werden. Weiterhin konnte die Sequenz in den Promotorbereichen von *C. tentans* in den Genen *luc7(p)-like* sowie *hermansky pudlak-like* lokalisiert werden. In den Genen *dLacDH-like*, *fs(1)K10-like* sowie *CG15207-like* konnten zwei CAAT-Boxen im Bereich bis -200 Nukleotide gefunden werden. Die gefundenen CAAT-Boxen sind alle im Bereich von -57 bis -195 Nukleotide lokalisierbar und stellen eine Bindestelle für den generellen Transkriptionsfaktor NF1 („nuclear factor1“) bzw. CBF („CCAAT-Box binding factor“) dar (Jones et al. 1987).

Die GC-Box mit einer Konsensus-Sequenz GGGCGG stellt eine Bindestelle für den Transkriptionsfaktor Sp1 dar. Diese Konsensus-Sequenz konnte in keinem proximalen Promotorbereich der Gene der SDR ausgemacht werden. Bereits für das Gen *CtY* konnten innerhalb des Promotorbereichs zwar Sequenzen mit vier bis fünf GC-Nukleotiden gefunden werden. Aber auch hier konnte keine vollständige Konsensus-Sequenz identifiziert werden. Kraemer (1994) stellte die Hypothese auf, dass dies unter Umständen mit dem hohen AT-Gehalt des Genoms von *Chironomus* zu tun haben könnte und die verkürzten Sequenzen trotzdem für eine Bindung von Sp1 ausreichen. Auch in den Histongenen, welche in der Regel gut konservierte GC-Boxen aufweisen, sind in *Chironomus* keine derartigen konservierten Sequenzen innerhalb der Promotorbereiche zu finden (Ross 1993). Innerhalb der Promotorbereiche der Gene der SDR konnten GC-Sequenzen mit vier bis sechs Nukleotiden lediglich in den Genen *mi-er1-like* sowie *polyhomeotic-like* (Exon 1a ♂_2) gefunden werden. Alle anderen Gene zeigen keine GC-reichen Sequenzen von mindestens vier Nukleotiden innerhalb des Bereichs von -50 bis -200 Nukleotiden.

4.2.3 Gene und ihre Produkte innerhalb der SDR von *Chironomus* sowie ihre potentielle Beteiligung an der Geschlechtsbestimmung

Im Rahmen der vorliegenden Arbeit wurden 11 Gene mit Lage innerhalb der SDR von *C. thummi* näher untersucht (siehe Kapitel 3.2). Hierbei wurde überprüft, welche Funktion für die Gene beschrieben ist und ob das jeweilige Gen an der Geschlechtsdetermination bei *Chironomus* beteiligt sein kann. Hierfür erfolgte ein Vergleich der Domänenstrukturen der abgeleiteten AS-Sequenzen mit den beschriebenen Strukturen der Proteine der geschlechtsbestimmenden Kaskaden anderer Organismen. Dabei wurde auch auf Bereiche innerhalb der Sequenzen der Polypeptide geachtet, welche Anreicherungen von bestimmten AS aufzeigen. Ein Vergleich der mit Hilfe der Programme MotifScan, InterProScan und SMART gefundenen Domänen und Regionen mit Anreicherungen bestimmter AS ergab lediglich bei der abgeleiteten AS-Sequenz des Gens *luc7(p)*-like eine Übereinstimmung zu den Proteinen mit Beteiligung an der Geschlechtsdetermination (z.B. Sxl, CSD, TRA, TRA2, DSX, SRY) anderer untersuchter Organismen (*D. melanogaster*, *M. domestica*, *C. capitata*, *A. mellifera*, *H. sapiens*, *M. musculus*). Hierbei konnten Homologien zu Proteinen mit RS-Domäne festgestellt werden. Auf das Gen *luc7(p)*-like sowie die mögliche Funktion des Proteins innerhalb der Geschlechtsdetermination wird in Kapitel 4.2.5 eingegangen.

Für weitere Gene des Contigs SDR werden in der Literatur Funktionen innerhalb von Entwicklungsprozessen beschrieben. Hierbei handelt es sich um die Proteine von *fs(1)K10*-like, *polyhomeotic*-like sowie *rpn5*-like. Ob es sich bei diesen Genen um potentiellen Kandidaten für den Faktor *M* handelt, wird zusammen mit einer detaillierten Beschreibung der drei Gene in den Kapiteln 4.2.6.1 bis 4.2.6.3 diskutiert.

Für einen Teil der in der SDR von *C. thummi* sowie *C. tentans* identifizierten Gene sowie deren Produkte werden Funktionen innerhalb allgemeiner Zellprozesse beschrieben. So sind z.B. für *mi-er1*-like sowie das nicht näher untersuchte *spitz*-like (Rutledge et al. 1992) Beteiligungen an Prozessen der Zellentwicklung beschrieben. Für *bcn92*-like sowie *dLacDH*-like werden Beteiligungen an Oxidations- und Reduktionsprozessen beschrieben. So katalysiert das Protein *dLacDH*-like die Oxidation von D-Laktat zu Pyruvat (Fraaije et al. 1998) und spielt somit eine zentrale Rolle im Energiestoffwechsel. Die beiden Gene *CG15207*-like und *CG11203*-like werden als proteinkodierend beschrieben, jedoch sind die Funktionen der beiden Gene bzw. deren Produkte bisher unbekannt. Im Folgenden soll auf einige Gene sowie deren Produkte näher eingegangen werden. Eine Beteiligung dieser Gene an der Geschlechtsdetermination scheint auf Grund der beschriebenen oder nicht beschriebenen Funktionen jedoch eher unwahrscheinlich.

***mi-er1*-like:**

Bei *mi-er1* (mesoderm induction early response 1) handelt es sich um ein Gen, dessen Expressionslevel während der Differenzierung von embryonalen Zellen zu mesodermalem Gewebe erhöht ist. Der Anstieg der Expression wird hierbei durch FGF

(fibroblast growth factor) vermittelt (Paterno et al. 1998). Das Genprodukt von *mi-er1* ist ausschließlich im Zellkern zu finden (Paterno et al. 1997). MI-ER1 fungiert hierbei als ein Repressor der Transkription von einer Reihe von Genen (Thorne et al. 2005). Für das Protein MI-ER1 sind eine SANT-Domäne sowie eine vorgelagerte ELM2-Domäne beschrieben (Ding et al. 2003). Die Region um die SANT-Domäne ist hierbei die am stärksten konservierte Region im Protein (Thorne et al. 2005). Die SANT-Domäne wird sowohl als DNA-bindend wie auch an Protein-Protein-Interaktionen beteiligt beschrieben (Paterno et al. 1997). Die ELM2-Domäne zeigt eine N-terminale Lage zu anderen DNA-Bindedomänen (Solari et al. 1999) und ist an der Regulation der Transkription beteiligt (Ding et al. 2003). Weiterhin wurden durch computergestützte Analysen zwei mögliche Kern-Lokalisierungs-Signale (NLS), vier saure AS-Regionen am N-Terminus sowie eine Prolin-reiche Region am C-Terminus gefunden. Hierbei fungiert die N-terminale saure Region als ein wirksamer Transaktivator, weshalb eine Funktion als Transkriptionsfaktor für MI-ER1 angenommen wird (Paterno et al. 1997).

wd74-like:

Die WD-Repeat enthaltenden Proteine gehören zu einer großen Proteinfamilie. Diese besteht aus mehr als 30 Subfamilien (Yu et al. 2000). Vertreter dieser Familien sind in allen Eukaryoten, jedoch nicht in Prokaryoten zu finden. Eine WD-Repeat-Einheit umfasst etwa 40 AS. Typischerweise beginnen diese mit einem Glycin-Histidin-Paar und enden mit einem Tryptophan-Asparaginsäure (WD)-Paar (Neer et al. 1994). Die Funktionen von WD-Repeat haltigen Proteinen sind sehr weitreichend. Sie reichen von der Regulation der Zelldifferenzierung über transmembrane Signaltransduktion und Vesikelverschmelzung bis hin zur Kontrolle der Zusammenlagerung des Initiationskomplexes bei der Transkription (Neer et al. 1994; Smith et al. 1999). Hierbei haben Proteine mit ähnlichen Oberflächen häufig auch gleiche Interaktionspartner sowie gleiche Funktionen.

bcn92-like:

Das Gen *bcn92* wurde 1995 in *Drosophila* als ein Gen mit 3 Exons und einem möglichen Protein mit einer Länge von 92 AS beschrieben (Cirera et al. 1995). Die im Gen enthaltene LYR-Domäne, kommt in vielen Proteinen des Eisen-Schwefel (Fe-S) Metabolismus vor (Ghezzi et al. 2009). Benannt ist das Motiv nach dem hochkonservierten Tripeptid (LYR) in der Nähe des N-Terminus der Proteine. Möglicherweise hat *bcn92* die Funktion einer Oxidoreduktase (Beisswanger et al. 2006).

In den abgeleiteten AS-Sequenzen von *C. thummi* und *C. tentans* konnte an den Positionen 14-16 AS jeweils das Tripeptid LYK gefunden werden. Hierbei ist die stark basische AS Arginin durch die ebenfalls stark basische AS Lysin ersetzt. Es handelt sich somit um einen konservativen AS-Austausch.

hermansky pudlak-like:

Das im Contig SDR gefundene Gen *hermansky pudlak-like* zeigt als BLAST-Treffer eine Übereinstimmung zu dem Gen *hermansky-pudlak 3*. Als Hermansky-Pudlak Syndrom (HPS) wird eine Gruppe von genetischen Störungen beschrieben, die durch Defekte im endosomal-lysosomalen System hervorgerufen werden. Erstmals wurde diese Krankheit 1959 von den beiden Pathologen F.Hermansky und P.Pudlak beschrieben (Huizing et al. 2000; Dell'Angelica 2004). Das heterogene Krankheitsbild wird durch Abnormalitäten in der Synthese und/ oder dem Transport von Lysosomenverwandten Organellen (LROs) ausgelöst (Gautam et al. 2006). Die Proteine HPS3, HPS5 und HPS6 sind Komponenten des Protein-Komplexes BLOC-2 (biogenesis of lysosome-related organelles complex-2). Dieser reguliert den Vesikel-Transport zu den LROs sowie die Synthese der LROs. Die Expression der Gene *hps3*, *hps5* und *hps6* erfolgt nur in höheren Eukaryoten. Es existieren keine homologen Gene in Hefen (Huizing et al. 2002; Gautam et al. 2004). In *Drosophila* ist ein zu *hps3* homologes Gen bekannt. Hierbei handelt es sich um das Gen CG14562. Das Genprodukt ist bisher nicht näher charakterisiert (Dell'Angelica 2004; Falcon-Perez et al. 2007). Das Protein HPS3 zeigt eine mögliche Clathrin-Bindedomäne (LLDFE) (Boissy et al. 2005; Helip-Wooley et al. 2005).

4.2.4 Gene und ihre Produkte der SDR mit NLS-Signal

Im Rahmen dieser Arbeit wurde die SDR von *C. thummi* sowie *C. tentans* auf Gene hin untersucht, die als potentielle Kandidaten für den dominanten männchenbestimmenden Faktor *M* in Frage kommen können. Da dieser Faktor ein Transkriptionsfaktor sein könnte und somit im Zellkern wirken muss, wurden die aus der Nukleotidsequenz abgeleiteten Polypeptide auf Kern-Lokalisierungs-Signale (NLS) hin untersucht. Das NLS wird benötigt, da die Aufnahme von Proteinen in den Zellkern äußerst selektiv ist (Dingwall und Laskey 1986; Garcia-Bustos et al. 1991). Obwohl die Sequenz des NLS recht einfach ist, konnten die bisher analysierten Signale auf keinen einheitlichen Nenner gebracht werden. Obwohl bereits einige Muster von Konsensus-Sequenzen vorgeschlagen wurden (Gomez-Marquez und Segade 1988), ist wohl die Definition des „bipartite nuclear localization signal“ (Dingwall und Laskey 1991) die beste Möglichkeit ein NLS zu finden. Hiernach beginnt das Signal mit zwei benachbarten basischen AS (Arg oder Lys). Es folgt ein „Spacer“ von 10 AS, bevor das Signal innerhalb der letzten fünf AS wiederum drei basische AS aufweist.

Die Analyse der Polypeptide der SDR aus *Chironomus* auf mögliche Kern-Lokalisierungs-Signale erfolgte mit den beiden Programmen MotifScan und ProSite. Hierbei wurden mit ProSite NLS innerhalb der abgeleiteten AS-Sequenzen von *luc7(p)*-like sowie *fs(1)K10*-like identifiziert. Diese wurden ebenfalls mit dem Programm MotifScan bestätigt. Zusätzlich konnten mit MotifScan innerhalb von vier weiteren AS-Sequenzen potentielle NLS lokalisiert werden. Die Unterstützungswerte dieser Treffer liegen hierbei knapp unter den Werten der als signifikant angegebenen Treffer

der Genprodukte von *luc7(p)*-like und *fs(1)K10*-like. Bei den zusätzlichen Treffern handelt es sich um die Produkte der Gene *mi-er1*-like, *polyhomeotic*-like (mit Exon 1b), CG11203-like sowie *rpn5*-like. Für alle weiteren abgeleiteten AS-Sequenzen konnten keine NLS gefunden werden.

Bei den abgeleiteten AS-Sequenzen von *mi-er1*-like, *luc7(p)*-like, *polyhomeotic*-like und *rpn5*-like handelt es sich um Proteine, deren Homologe in der Literatur als potentielle Transkriptionsfaktoren, Spleißfaktoren oder als im Zellkern lokalisierbar beschrieben wurden. Die gefundenen Domänen der beschriebenen Proteine (ELM2, SANT, RS, Zink-Finger, SAM und PCI) sprechen für Protein-DNA- oder Protein-Protein-Interaktionen. Selbst unter der Annahme, dass es sich bei den nicht signifikanten Treffern des Programms MotifScan um keine realen NLS handelt, besteht somit für die Proteine *mi-er1*-like, *polyhomeotic*-like und *rpn5*-like die Möglichkeit einer Aufnahme in den Zellkern als Bestandteil eines Proteinkomplexes, da diese Proteine Domänen zur Protein-Protein-Interaktion aufweisen. Die Möglichkeit der Aufnahme von Proteinen in den Zellkern als Teil eines Proteinkomplexes wurde für RPN5 bereits diskutiert (Yen et al. 2003).

Für das *Drosophila*-Homolog des abgeleiteten Polypeptids von *fs(1)K10* wird ebenfalls ein Motiv mit DNA-bindender Funktion beschrieben (siehe auch 4.2.6.1). Es wird daher angenommen, dass es sich bei dem Protein um einen möglichen TF handelt (Prost et al. 1988; Suzuki et al. 1994). Allerdings konnte bei *fs(1)K10*-like aus *C. thummi* dieses DNA-bindende Motiv nicht gefunden werden. Für das Genprodukt von CG11203-like konnten bisher keine beschriebene Funktion von homologen Proteinen gefunden werden. Des Weiteren konnten in den abgeleiteten AS-Sequenzen von *fs(1)K10*-like und CG11203-like keine signifikanten Treffer zu konservierten Domänen identifiziert werden. In welcher Weise diese Produkte demnach im Zellkern wirken, bedarf weiterer Untersuchungen.

Mit Hilfe der Suche nach abgeleiteten Polypeptiden mit Kern-Lokalisierungs-Signalen konnten die potentiellen Kandidaten für den dominanten Faktor *M*, unter der Annahme seiner Funktion als TF, auf die im Zellkern lokalisierbaren Proteine begrenzt werden.

4.2.5 Das Gen *luc7(p)*-like und seine mögliche *tra*-ähnliche Funktion

Ein besonderes Interesse weckt das Gen *luc7(p)*-like. Das Gen *luc7(p)* (lethal unless CBC is produced) wurde erstmals als eine letale Mutation in einem Hefestamm identifiziert, welchem der nukleare Cap-Bindekomplex fehlt. *Luc7(p)* ist eine Komponente des U1-snRNPs, welches essentiell für das vegetative Wachstum bei Hefen ist. Das Fehlen von *Luc7(p)* führt zu einer Störung der Interaktion von U1-snRNP und CBC (Cap Bindekomplex). Hierdurch wird die 5'-Spleißstellen-Erkennung gestört und die Effizienz des Spleißens gesenkt (Fortes et al. 1999). Für den N-terminalen Bereich des Proteins wird eine Zink-Finger-Domäne beschrieben, welche im 5'-Exon der prä-mRNA in der Nähe des Cap bindet. Weiterhin wird im C-terminalen Bereich eine Region mit den AS-Repeats Arginin-Serin (RS) und Arginin-

Glutaminsäure (RE) beschrieben. Diese sind charakteristisch für Spleiß-Faktoren (Fortes et al. 1999; Puig et al. 2007).

Die Untersuchung der abgeleiteten Polypeptide der SDR von *C. thummi* auf Anreicherungen der beiden AS Arginin (R) und Serin (S) zeigt, dass innerhalb des C-terminalen Bereichs des Genprodukts von *luc7(p)*-like ein Bereich zu finden ist, welcher eine erhöhte Anzahl dieser beiden Aminosäuren aufweist. Bei 49 der 79 C-terminalen AS des abgeleiteten Polypeptids von *luc7(p)*-like (ab AS-Position 241) handelt es sich um die AS Arginin oder Serin. Dies entspricht einem Prozentsatz von etwa 62 % der C-terminalen AS. Eine Anreicherung des AS-Repeats (RE) konnte hingegen nicht festgestellt werden.

Innerhalb des Contigs SDR von *C. thummi* konnte lediglich ein weiteres abgeleitetes Polypeptid gefunden werden, welches ebenfalls eine leicht erhöhte Anzahl an SR- oder RS-Dimeren aufweist. Hierbei handelt es sich um das Produkt des Gens CG11203-like. In diesem sind hier 8 RS- oder SR-Dimere lokalisierbar. Diese sind aber im Gegensatz zum Protein *luc7(p)*-like über die gesamte Länge des abgeleiteten Polypeptids mit einer Größe von 914 AS verteilt. Eine Anreicherung des AS-Repeats (RE) konnte genau wie bei *luc7(p)*-like auch bei CG11203-like nicht festgestellt werden. Auf Grund der zerstreuten Lage der Dimere (RS, SR) innerhalb des abgeleiteten Polypeptids sowie der, im Vergleich zum Protein *luc7(p)*-like (319 AS), erheblichen Größe scheint es sich allerdings hierbei nicht um eine signifikante Anreicherung zu handeln.

Das Protein *luc7(p)*-like ist hingegen auf Grund der Anreicherung an Arginin und Serin im C-terminalen Bereich sehr interessant. Eine Anreicherung der beiden AS Arginin und Serin liegt auch in den beiden Proteinen TRA und TRA2 vor. Diese werden der Familie der SR-Proteine zugeordnet (Amrein et al. 1988; Long und Caceres 2009). Die SR-Proteine TRA und TRA2 spielen in Arten wie *D. melanogaster* (Amrein et al. 1988; Sciabica und Hertel 2006), *M. domestica* (Hediger et al. 2010) und *C. capitata* (Pane et al. 2002; Salvemini et al. 2009) eine entscheidende Rolle in der Kaskade der Geschlechtsdifferenzierung. Auch in *A. mellifera* zeigt das Produkt des geschlechtsdeterminierenden Gens *csd* Homologien zu TRA auf und besitzt im C-terminalen Bereich eine RS-reiche Region (Beye et al. 2003; Beye 2004).

SR-Proteine wirken sowohl als generelle Spleiß-Faktoren wie auch als wichtige Regulatoren von alternativem Spleißen (Long und Caceres 2009). Hierbei kommt es durch die Auswahl von verschiedenen 5'-Spleißstellen zu einem alternativen Spleißen verschiedenster prä-mRNAs (Zahler et al. 1992; Fu 1993; Shen und Green 2004). Das Spleißen unterschiedlicher prä-mRNAs erfordert vermutlich auch verschiedene Komplexe von SR-Proteinen (Fu 1993). Generell sind SR-Proteine hierbei für den ersten Schritt bei der Bildung des Spliceosoms wichtig. Dieser besteht in der Interaktion des U1-snRNP-Komplexes mit der 5'-Spleißseite (Fu 1993; Zahler und Roth 1995). Weiterhin sind SR-Proteine für die Rekrutierung von allgemeinen Spleiß-Faktoren während der Bildung des Spliceosoms verantwortlich (Wu und Maniatis 1993; Kohtz et al. 1994). Durch die Interaktion von SR-Proteinen mit RNA-Sequenzelementen, den ESEs („exonic splicing enhancers“), kommt es zu einer Aktivierung von schwachen Spleißstellen. Dies geschieht durch die Rekrutierung eines Komplexes verschiedenster Spleiß-Enhancer. Die Bindestellen der Spleiß-Enhancer

werden durch die ESEs bereitgestellt (Graveley 2000; Graveley et al. 2001; Lam und Hertel 2002; Lam et al. 2003). Bei *D. melanogaster* führt die Interaktion der SR-Proteine TRA, TRA2 sowie RBP1 mit dem weibchenspezifischen Spleiß-Enhancer des Gens *dsx* zu einem weibchenspezifischen Spleißen des Transkripts (Lynch und Maniatis 1996).

Zusätzlich zur C-terminalen RS-Domäne weisen SR-Proteine häufig ein bis zwei RRM (RNA recognition motif)-Domänen mit einer Länge von 80-90 AS im N-terminalen Bereich auf. Diese sorgen für die nötige Spezifität der Proteine bei der RNA-Bindung (Dreyfuss et al. 1988; Mattaj 1989). Für die Bindung der RNA ist jedoch nicht zwingend eine RRM-Domäne nötig. Alternativ kann auch eine direkte Bindung an die prä-mRNA erfolgen. Hierbei interagiert die RS-Domäne mit der 5'-Spleißstelle sowie der Verzweigungsstelle der prä-mRNA (Shen und Green 2004; Shen et al. 2004; Hertel und Graveley 2005). Generell können SR-Proteine in zwei Gruppen aufgeteilt werden. Die erste Gruppe besitzt am N-Terminus eine RRM-Domäne (Hastings und Krainer 2001). Diese scheint innerhalb der SR-Proteine konserviert zu sein (Dauwalder et al. 1996; Blencowe et al. 1999). Die zweite Gruppe besitzt keine RRM-Domäne. Diese Gruppe wird als SR-ähnliche Proteine bezeichnet (Blencowe et al. 1999).

In *H. sapiens* konnte mit *LUC7L* (*Luc7*-like; humanes *Luc7p*) ein homologes Gen zu *luc7(p)*-like beschrieben werden. Innerhalb der Sequenz des Proteins konnten eine RS-Domäne sowie zwei Zink-Finger-Domänen identifiziert werden (Graveley 2000). Innerhalb der abgeleiteten AS-Sequenz von *luc7(p)*-like aus *C. thummi* konnten zwei Bereiche identifiziert werden, welche schwache Homologien zu Zink-Finger-Motiven zeigen. Diese liegen in den Bereichen 31-48 AS sowie 189-216 AS und somit in ähnlichen Regionen wie beim humanen *Luc7p*.

Sowohl die RS-Domäne im C-terminalen Bereich des abgeleiteten Polypeptids von *luc7(p)*-like wie auch die in der Literatur beschriebene Beteiligung an der Erkennung der 5'-Spleißstelle sind hierbei Anzeichen für eine Verwandtschaft zu SR-Proteinen. Die Untersuchung auf konservierte Domänen wie auch der Sequenzvergleich zu verschiedenen Proteinen mit RRM-Domäne wie z.B. TRA2 und Sxl zeigt, dass innerhalb der abgeleiteten AS-Sequenz von *luc7(p)*-like keine derartige Domäne zu finden ist. Allerdings zeigen auch die Proteine CSD aus *A. mellifera* sowie die TRA-Proteine der verschiedenen untersuchten *Drosophila*-Arten keine RRM-Domäne. Die Identifikation von homologen Genen zu *transformer* gestaltet sich auf Grund der schlecht konservierten Sequenz als schwierig. Die hohe Variabilität der Nukleotidsequenz ist Folge eines schnell evolvierenden *transformer*-Gens (McAllister und McVean 2000). So beträgt z.B. die Sequenzidentität zwischen TRA aus *D. melanogaster* und dem TRA-Orthologen CSD aus *A. mellifera* nur 16 % (Beye et al. 2003). Die Übereinstimmung auf AS-Ebene zwischen TRA aus *Drosophila* und *C. capitata* liegt bei nur 18 % (Pane et al. 2005). Selbst zwischen verschiedenen Stämmen von *Drosophila* ist eine hohe Sequenzvariabilität feststellbar (O'Neil und Belote 1992). Ein Vergleich der abgeleiteten AS-Sequenz von *luc7(p)*-like zu TRA aus *D. melanogaster* zeigt eine Übereinstimmung von etwa 22 %. Beschränkt man den Sequenzvergleich auf die RS-Domäne (241-319 AS), erhöht sich die Übereinstimmung auf etwa 34 %. Ein Sequenzvergleich der RS-Domäne von CSD aus *A. mellifera* mit

der RS-Domäne der abgeleiteten AS-Sequenz von *luc7(p)*-like aus *C. thummi* zeigt eine Übereinstimmung von etwa 28 %.

Für die Analyse des Proteins CSD aus *A. mellifera* (Beye et al. 2003) auf Homologien zu TRA wurde eine kurze 11 AS lange, konservierte Sequenz aus TRA-Proteinen von *C. Capitata* sowie mehrerer *Drosophila*-Arten genutzt (Pane et al. 2002). In CSD ist keine Übereinstimmung mit dieser konservierten Sequenz zu finden. Mit Hilfe dieser kurzen konservierten Sequenz wurde auch in dieser Arbeit ein Vergleich zu den Produkten der Gene der SDR durchgeführt. Hierbei wurde lediglich in der abgeleiteten AS-Sequenz von dem Gen *luc7(p)*-like eine signifikante Übereinstimmung zu der konservierten Sequenz aus TRA gefunden. Diese Sequenzhomologie ist im Bereich von 252-265 AS des abgeleiteten Polypeptids erkennbar und liegt somit innerhalb des Bereichs der RS-Domäne (241-319 AS). Dies entspricht der Lage des konservierten AS-Abschnitts in *Ceratitis* und *Drosophila*. Hierbei liegt dieser Bereich in *C. capitata*, *D. erecta*, *D. melanogaster* und *D. simulans* ebenfalls innerhalb der RS-Domäne. Lediglich in den beiden Arten *D. hydei* und *D. virilis* ist der konservierte AS-Abschnitt kurz vor der RS-Domäne lokalisierbar. Eine grafische Übersicht über die Sequenzhomologie der RS-Domäne von *luc7(p)*-like-Proteinen innerhalb der untersuchten *Chironomus*-Arten ist in Abbildung 4.4 gegeben. Ebenfalls in dieser Abbildung dargestellt ist die Übereinstimmung der konservierten TRA-Sequenz aus *Drosophila* mit der Sequenz aus *Chironomus*. Die Übereinstimmung der konservierten Sequenz von *Chironomus* mit der konservierten Sequenz von *Drosophila* beträgt in diesem Bereich 57,1 %. Auch in dem TRA-Homologen aus *M. domestica* konnte diese konservierte Sequenz gefunden werden. Die Übereinstimmung mit der *Drosophila*-Sequenz beträgt 78,6 %. In den beiden Mücken-Arten *Anopheles gambiae* und *Aedes aegypti* liegen auf Grund der niedrigen Konserviertheit der TRA-Homologen bisher keine Kandidaten für TRA und somit auch keine Sequenzinformationen vor. Aus diesem Grund wurde per BLAST-Analyse nach homologen Proteinen zum analysierten *Luc7(p)*-like Protein aus *C. thummi* gesucht. Hierbei wurden je zwei Proteine gefunden, welche im N-terminalen Bereich starke Homologien zum Protein *Luc7(p)*-like aus *C. thummi* aufweisen. Bei *A. gambiae* handelt es sich um die Proteine AGAP010454-PA (XP_311495.4) sowie AGAP008314-PA (XP_317153.4), bei *A. aegypti* um die Proteine AST3 (alternative splice type 3; XP_001660952.1) und pTSR1 (XP_001659775.1). Auch diese Proteine wurden auf den in *Drosophila* und *Ceratitis* vorliegenden konservierten AS-Abschnitt sowie auf RS- und SR-Repeats innerhalb dieser Region hin untersucht. Alle vier Proteine weisen im C-terminalen Bereich ein erhöhtes Vorkommen an RS- und SR-Repeats auf. Des Weiteren konnte in allen vier Proteinen ein AS-Abschnitt mit Homologie zum konservierten Bereich aus *Drosophila* und *Ceratitis* gefunden werden. Die Sequenzübereinstimmung liegt im Falle des Proteins pTSR1 (XP_001659775.1) aus *A. aegypti* bei 50 %. In den anderen drei Proteinen liegt die Homologie, genau wie in *C. thummi*, bei 57,1 %. Die Sequenz-Alignments der gefundenen Bereiche mit Homologie zum konservierten AS-Abschnitt aus *Drosophila* und *Ceratitis* sind ebenfalls in Abbildung 4.4 zu finden.

Die Existenz einer RS-Domäne innerhalb des C-terminalen Bereichs der abgeleiteten AS-Sequenz von *luc7(p)*-like (241-319 AS), die gezeigte Sequenzhomologie zu dem

kurzen konservierten Abschnitt der TRA-Sequenz aus *Drosophila* (252-265 AS) sowie die für Luc7p in der Literatur beschriebene Beteiligung an der 5'-Spleißstellen-Erkennung sind Anzeichen für eine mögliche TRA-ähnliche Funktion von luc7(p)-like. Auf Grund der vorhandenen RS-Domäne sowie der fehlenden RRM-Domäne kann das abgeleitete Polypeptid von luc7(p)-like zu der Gruppe der SR-ähnlichen Proteine zugeordnet werden. Zu dieser Gruppe gehört auch das humane Luc7p-Protein (Graveley 2000).

Handelt es sich bei luc7(p)-like tatsächlich um ein Protein mit TRA-ähnlicher Funktion, so müsste mindestens ein alternatives Spleiß-Produkt zu finden sein. Mit Hilfe der Hochdurchsatz-Sequenzierung konnte im Laufe dieser Arbeit ein verkürztes Produkt des Gens *luc7(p)*-like gefunden werden (siehe Kapitel 3.2). Dieses konnte bisher nur in *C. thummi* ♀ sowie *C. piger* ♀ gefunden werden und liegt in den Daten der Hochdurchsatz-Sequenzierung des Transkriptoms von L4-Larven stark unterrepräsentiert vor. Hierbei konnte im eigentlichen Exon 2 ein zusätzliches, offensichtlich differentiell gespleißtes Intron identifiziert werden (siehe Abbildung 3.12 sowie Abbildung 4.3). Hierdurch verschiebt sich der Leserahmen, und es kommt zu einem verkürzten Transkript mit lediglich 555 bp. Das hieraus resultierende abgeleitete Polypeptid besitzt eine Länge von 185 AS. Dem Polypeptid fehlt somit der gesamte C-terminale Bereich, in welchem sich die RS-Domäne befindet (ab Position 241 AS im eigentlichen Polypeptid).

In *C. piger* konnte ebenfalls in Exon 2 des Gens *luc7(p)*-like ein zusätzliches Intron gefunden werden. Dieses ist zwar 15 bp größer als das Intron in *C. thummi*. Vergleicht man jedoch die 5'-Introngrenze des zusätzlichen Introns sowie die Position des Stopcodons zwischen den beiden Arten *C. thummi* und *C. piger*, so ist festzustellen, dass diese an den selben Positionen liegen.

Auf Grund der geringen Menge an Sequenzdaten, welche das alternativ gespleißte Intron bestätigen, liegt die Vermutung nahe, dass das alternative Produkt im L4-Stadium nur schwach vertreten vorliegt. Um zu überprüfen, ob das Produkt eventuell während der frühen Embryonalentwicklung stärker exprimiert wird, bedarf es der zusätzlichen Untersuchung früherer Entwicklungsstadien. Sollte sich das alternativ gespleißte Produkt auch hier nur in *C. thummi* ♀ bzw. *C. piger* ♀ bestätigen, so wäre dies sehr interessant, da in diesem Fall im ♀ eine differentielle Regulation erfolgen würde, die vermutlich zu einem nicht funktionsfähigen Spleißfaktor führt.

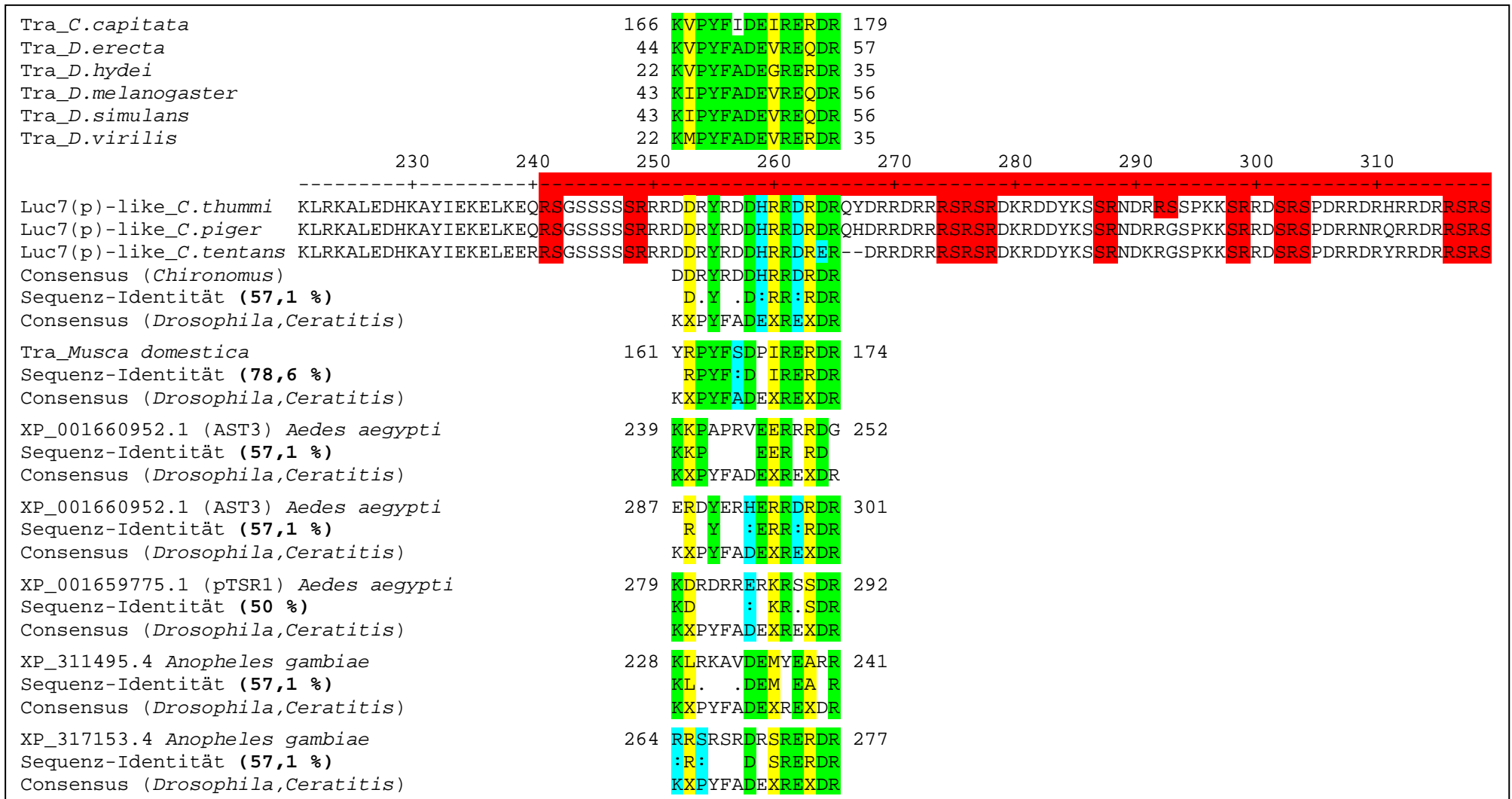


Abbildung 4.4: RS-Domäne des abgeleiteten Polypeptids von *luc7(p)*-like aus *Chironomus* sowie Sequenz-Vergleich zu einem konservierten TRA-Abschnitt aus *Drosophila* (Erläuterungen siehe nächste Seite)

In der Abbildung ist der C-terminale Bereich des abgeleiteten Polypeptids von *luc7(p)*-like aus *Chironomus* gezeigt. Im Bereich der AS 241-319 (79 AS, roter Balken) konnten 49 Mal die AS Arginin (R) oder Serin (S) identifiziert werden. Die Sequenz-Motive RS oder SR innerhalb dieser Region (ebenfalls in rot dargestellt) weisen eine starke Konservierung zwischen den untersuchten *Chironomus*-Arten auf. Der Sequenzvergleich eines konservierten TRA-Abschnitts aus *Drosophila* und *Ceratitis* (Sequenz-Alignment im oberen Bildteil) mit der AS-Sequenz des abgeleiteten Proteins von *luc7(p)*-like aus *Chironomus* zeigt eine Sequenz-Homologie von 57,1 %. Hierbei liegt der gefundene homologe AS-Abschnitt genau wie in den meisten Fällen bei *Drosophila* innerhalb des Bereichs der RS-Domäne. Konservierte AS-Positionen sind innerhalb der Sequenzvergleiche in grün gehalten; variable AS-Positionen hingegen in gelb. Konservative AS-Austausche sind mit (:) gekennzeichnet und hellblau hinterlegt. Weniger konservierte AS-Austausche sind mit (.) gekennzeichnet. Auch im TRA-Protein von *M. domestica* sowie den per BLAST-Analysen gefundenen *Luc7(p)*-Homologen aus *A. aegypti* und *A. gambiae* konnte ein Bereich identifiziert werden, welcher Homologien zum konservierten AS-Abschnitt aus *Drosophila* aufweist. Die Übereinstimmungen zu *Drosophila* liegen bei 78,6 % (Tra *M. domestica*), 50 % (pTSR1 *A. aegypti*) sowie 57,1 % für die drei weiteren Proteine aus *A. aegypti* und *A. gambiae*. In allen analysierten Proteinen der drei Arten *M. domestica*, *A. aegypti* sowie *A. gambiae* befindet sich der AS-Abschnitt mit Homologien zum konservierten Bereich des TRA-Proteins aus *Drosophila* wie in den verschiedenen *Drosophila*-Arten entweder vor oder am Beginn der RS-Domäne. Auf Grund der niedrigen Sequenzidentitäten zwischen den RS-Domänen der einzelnen TRA-Proteine (*Drosophila*-Stämme) bzw. CSD (*A. mellifera*) zu der abgeleiteten AS-Sequenz aus *Chironomus* sowie der fehlenden Übereinstimmung des CSD-Proteins (*A. mellifera*) zur konservierten AS-Sequenz aus *Drosophila* und *C. capitata*, sind diese Alignments in der Abbildung nicht gezeigt.

4.2.6 Potentielle Kandidaten-Gene für *M*

Innerhalb des Contigs SDR konnten drei Gene ausgemacht werden, für welche in anderen Organismen eine Funktion innerhalb geschlechtsspezifischer Entwicklungsprozesse nachgewiesen werden konnte. Hierbei handelt es sich um die Gene *fs(1)K10-like*, *polyhomeotic-like* sowie *rpn5-like*.

4.2.6.1 *fs(1)K10-like*

In *Drosophila* erfolgt die Expression des für die Ausbildung der dorso-ventralen Achse wichtigen Gens *fs(1)K10* vor allem in Stadien der frühen Oogenese. Die Transkription des Gens erfolgt hierbei in den Nährzellen. Im Folgenden wird die mRNA über zytoplasmatische Brücken zum posterioren Pol der Oocyte transportiert. Für das Protein *fs(1)K10* konnte eine Anreicherung innerhalb des Zellkerns der Oocyte nachgewiesen werden (Haenlin et al. 1987; Prost et al. 1988; Serano und Cohen 1995). Liegt eine homozygote Mutation des X-chromosomalen Gens *fs(1)K10* vor, so führt dies durch einen maternalen Effekt zu Sterilität in Weibchen der Folgegeneration. Hierbei wird die Lokalisation von *gurken* (*grk*)-mRNA im dorsal-anterioren Bereich der Oocyte verhindert (Neuman-Silberberg und Schüpbach 1993). Die hieraus resultierenden Eier zeigen eine abnormale, zylindrische Form. In nahezu allen Fällen entwickeln sich aus diesen Eiern keine Larven (Wieschaus et al. 1978).

Innerhalb des Proteins *fs(1)K10* konnten acht „repeats“ des Oktamers SPNQQQHP (oder leicht abweichend) nachgewiesen werden. Es wird vermutet, dass es sich bei dem Protein um einen Transkriptionsfaktor handelt (Prost et al. 1988; Suzuki et al. 1994). Die Interaktion mit der DNA erfolgt hierbei über die Oktamere, welche jeweils einen A/T-reichen Bereich von 4 Basenpaaren überspannen.

Sowohl in *C. thummi* wie auch in *C. tentans* konnte innerhalb der SDR ein Gen identifiziert werden, dessen abgeleitete AS-Sequenz nach BLAST-Analysen Homologien zu *fs(1)K10* aus *D. melanogaster* aufzeigt. Die Übereinstimmung zwischen den beiden AS-Sequenzen aus *D. melanogaster* und *C. thummi* von 39 % beschränkt sich hierbei auf den 92 AS langen Bereich von AS 230-322. In Abbildung 4.5 ist das Ergebnis des Protein-BLASTs mit der abgeleiteten AS-Sequenz des Gens *fs(1)K10-like* aus *C. thummi* aufgezeigt. Weitere BLAST-Treffer zeigen in diesem Bereich Homologien zum *fs(1)K10*-Protein aus *Culex quinquefasciatus* (46 % im Bereich 215-299 AS), dem Protein AGAP013462-PA aus *A. gambiae* (45 % im Bereich 227-300 AS) sowie AND_16864 aus *Anopheles darlingi* (29 % im Bereich 31-373 AS). Bei Letzterem handelt es sich um ein hypothetisches Protein, welches laut Datenbankeintrag durch Homologie zu anderen Proteinen identifiziert wurde.

Zwar handelt es sich bei den BLAST-Treffern zu den K10-Proteinen aus *D. melanogaster* sowie *C. quinquefasciatus* um die Treffer mit der größten Homologie zum *fs(1)K10*-Protein aus *C. thummi*, jedoch erstrecken sich diese Treffer lediglich über einen Abschnitt von rund 25 % der Länge des Proteins. Auch auf Grund der mäßigen Homologie von 39-46 % in diesem Bereich ist es fraglich, ob es sich bei dem

innerhalb des Contigs SDR von *C. thummi* gefundenen Protein tatsächlich um ein fs(1)K10-Orthologes handelt.

GENE ID: 47781 fs(1)K10 female sterile (1) K10 [Drosophila melanogaster]	
(Over 100 PubMed links)	
Score = 67.8 bits (164), Expect = 1e-09, Method: Compositional matrix adjust. Identities = 39/99 (39%), Positives = 62/99 (63%), Gaps = 8/99 (8%)	
Query	230 YSLEAPFVTDEIKAEHAKKEEILESMKGGKNDLFAKFKEQREVFKKYDEAKAAYQAE 289
	Y +E P+VTDEI++EH KK +I +KG K+DELFA++K QR+ FV Y+ A+ Y +
Sbjct	366 YPIEKPWVTDEIRSEHDKKVDIENRLKGH-KDELFAQYKGRDKFVSLYEAREEYLKQ 424
Query	290 KKAQREVAKAAKAE-----AAEKAKLESEAGKDEKIEN 322
	+ A + AK AK++ +++ A ++ + KD I N
Sbjct	425 EAATVK-AKDAKSDKDKNAISSQSAAPKAGSAKDATIPN 462

Abbildung 4.5: Sequenzalignment des Protein-BLAST Treffers mit Homologie zu fs(1)K10 aus *D. melanogaster* nach Suche mit der fs(1)K10-Sequenz aus *C. thummi*. Das Sequenzalignment zeigt eine Übereinstimmung der beiden fs(1)K10-Sequenzen aus *D. melanogaster* und *C. thummi* von 39 % über einen Bereich von 92 AS. Die entspricht einem Abschnitt von etwa 25 % der Proteingröße des fs(1)K10-Proteins aus *C. thummi* (377 AS). Die obere Sequenz im Alignment entspricht hierbei der Sequenz aus *C. thummi*, während die untere Sequenz einem Abschnitt des fs(1)K10-Proteins aus *D. melanogaster* entspricht.

In *C. thummi* sowie der Schwesternart *C. piger* konnte mit Hilfe genomischer Southern-Analysen eine männchenspezifische Duplikation im Bereich des Gens *fs(1)K10-like* nachgewiesen werden (Stiehl 2009). Die Sequenz dieser Duplikation konnte mit Hilfe der „inversen PCR“ sowie der Strategie des „Genome Walks“ erhalten werden (Stiehl 2009, Kraemer und Stiehl persönliche Mitteilung). Hierbei unterscheiden sich die AS-Sequenzen der beiden Y-chromosomalen Varianten sowohl untereinander wie auch von der X-chromosomalen Variante im C-terminalen Bereich (siehe Kapitel 3.2). Im N-terminalen Bereich gleichen sich die drei AS-Sequenzen hingegen fast vollständig.

Bei einem Vergleich der drei in *C. thummi* gefundenen X- und Y-chromosomalen Produkte des Gens *fs(1)K10-like* mit dem fs(1)K10-Protein aus *D. melanogaster* konnte kein „Oktamer-Repeat“ der Sequenz SPNQQQHP gefunden werden. Auch andere „Repeats“ konnten innerhalb der Sequenz nicht identifiziert werden.

Weitere Untersuchungen der abgeleiteten AS-Sequenzen aus *C. thummi* auf konservierte Domänen mit den Programmen InterProScan, SMART und MotifScan lieferten keine signifikanten Treffer. Zwei schwache Treffer mit Abschnitten des X-chromosomalen Produkts im Bereich von 216-309 AS (SPT2-Domäne) sowie 222-295 (HMG-Domäne) zeigen Übereinstimmungen zu Domänen, die in Transkriptionsfaktoren vertreten sind (Harley et al. 2003; Novoseler et al. 2005). Diese Treffer sind auch in den Y-chromosomalen Varianten zu finden. Inwiefern diese schwachen Übereinstimmungen eine Aussage über die Funktion des Genprodukts von *fs(1)K10-like* geben können, ist fraglich.

In allen drei abgeleiteten AS-Sequenzen konnten signifikante Anreicherungen der Aminosäuren Asparagin, Glutamin, Methionin und Prolin gefunden werden. Hinzu kommen schwache Übereinstimmungen mit Anreicherungen der AS Alanin, Lysin und Threonin. Diese Übereinstimmungen sind nicht in allen AS-Sequenzen vertreten. Ob

diese AS-Anreicherungen eine Rolle in der Funktion des Genprodukts spielen und inwiefern sie die Funktion des Proteins beeinflussen, ist nicht bekannt.

Eine Untersuchung der Daten aus der Hochdurchsatz-Sequenzierung des Transkriptoms von *C. thummi* zeigt, dass auch ein Mapping mit niedriger Stringenz keine „Reads“ mit Übereinstimmung zum *fs(1)K10*-Gen von *Drosophila* liefert.

Auf Grund der durchgeführten Analysen ist davon auszugehen, dass es sich bei dem in der SDR lokalisierten Gen nicht um das *fs(1)K10*-Orthologe von *Drosophila* handelt. Welche Funktion das kodierte Protein von diesem Gen ausübt, ist unbekannt. Somit sind das Gen und das davon kodierte Protein weiterhin sehr interessant und bedürfen weiterer Analysen.

4.2.6.2 polyhomeotic-like

Das innerhalb der SDR von *C. thummi* liegende Gen *polyhomeotic-like* zeigt Homologien zu dem zur Polycomb (PcG)-Gruppe gehörigen Gen *polyhomeotic* aus *Drosophila*. Die Genprodukte der PcG-Gruppe sind für die korrekte und differentielle Expression der beiden homeotischen Genkomplexe ANT-C (Antennapedia-Komplex) und BX-C (Bithorax-Komplex) sowie die Aufrechterhaltung der segmentalen Identität nötig (Deatrick et al. 1991; DeCamillis et al. 1992; Kim et al. 2005). Weiterhin ist *polyhomeotic* essentiell für die epidermale Zellentwicklung der Ventralseite (Dura et al. 1987). Der *polyhomeotic*-Locus besteht aus zwei genetisch unabhängigen Einheiten mit tandem-repetitiver Anordnung. Diese sind durch eine Duplikation entstanden. Die Funktionen der beiden Einheiten sind hierbei weitgehend redundant (Dura et al. 1987; Deatrick et al. 1991). Allerdings führen Umordnungen der Struktur einer Einheit zu einer veränderten Funktion des Locus (Dura et al. 1987). Eine Überexpression von *polyhomeotic* führt in *Drosophila* zu einer reduzierten Fertilität der Weibchen. Hierbei kommt es zu negativen Veränderungen der Follikel-Struktur. Ebenfalls in der Struktur veränderte Follikel entstehen bei einer fehlenden Expression von *polyhomeotic* (Narbonne et al. 2004). Hierbei wird die Funktion des Proteins Polyhomeotic für die Proliferation und Differenzierung der somatischen Zellen des Keimstocks benötigt.

Das Protein Polyhomeotic von *Drosophila* zeigt eine SAM (sterile alpha motif)-Domäne sowie eine Zink-Finger-Domäne auf. Weiterhin konnten mehrere Glutamin-reiche und ein Serin/Threonin-reicher Bereich gefunden werden (DeCamillis et al. 1992; Franke et al. 1992). Ein Vergleich der in der Literatur beschriebenen Domänen mit den Ergebnissen der Domänensuche der abgeleiteten Polypeptidsequenz von *polyhomeotic-like* von *C. thummi* zeigt, dass auch hier die entsprechenden Domänen und Bereiche mit Anreicherungen der AS Glutamin bzw. Serin/Threonin gefunden werden konnten (siehe Kapitel 3.2). Ein Vergleich der abgeleiteten *polyhomeotic* AS-Sequenz von *Chironomus* und *Drosophila* zeigt hierbei lediglich eine Sequenzhomologie von etwa 28 %. Vergleicht man jedoch nur den Bereich der SAM-Domäne erhöht sich dieser Wert auf 73 %. Die in Polyhomeotic vorkommende Zink-Finger-Domäne weist auf eine Protein-DNA-Interaktion hin, was typisch für einen Transkriptionsfaktor ist. Über die gefundene SAM-Domäne kann eine Interaktion mit

anderen Proteinen oder auch mit RNA erfolgen. Hierbei ist in vielen Fällen eine Interaktion mit Proteinen typisch, welche ebenfalls eine SAM- oder eine SH2-Domäne besitzen. Proteine mit einer konservierten SAM-Domäne sind in verschiedene Entwicklungsprozesse involviert und weisen diverse Funktionen auf. Diese reichen von der Rezeptor-Thyrosin-Kinase bis hin zu Transkriptionsfaktoren (Schultz et al. 1997; Kim et al. 2005). In Hefen sind verschiedene Proteine mit SAM-Domänen in sexuelle Differenzierungsprozesse involviert. Mutationen in den entsprechenden Genen führen hierbei zu Sterilität (Rhodes et al. 1990; Okazaki et al. 1991; Ponting 1995).

Für Polyhomeotic ist eine Interaktion innerhalb eines Proteinkomplexes mit SAM-Proteinen der PcG-Gruppe beschrieben (Kim et al. 2005). Als negativer Regulator wirkt Polyhomeotic hierbei auf die homeotischen Gene sowie weitere Loci und den eigenen Locus ein (Fauvarque et al. 1995).

Die in der AS-Sequenz von *polyhomeotic-like* von *C. thummi* gefundene Zink-Finger-Domäne spricht für eine Funktion als Transkriptionsfaktor. Über die ebenfalls bestätigte SAM-Domäne kann außerdem eine Interaktion mit anderen Proteinen erfolgen. Inwieweit sich die beiden bisher nur in Männchen gefundenen und vom Weibchen abweichenden 5'UTR-Bereiche auf die Regulation der Genexpression auswirken, ist auf Grundlage der vorliegenden Daten nicht zu entscheiden. Der kodierende Bereich zeigt zwischen Männchen und Weibchen keine Unterschiede.

Die Funktion von *polyhomeotic-like* könnte mit der in *Drosophila* beschriebenen Funktion übereinstimmen. Allerdings ist auf Grund der nicht besonders hohen Sequenzhomologie von 28 % auch eine andere Funktion als die eines Transkriptionsfaktors denkbar. Somit könnte das Gen *polyhomeotic-like* prinzipiell einen möglichen Kandidaten für den dominanten Faktor *M* darstellen.

4.2.6.3 *rpn5-like*

Das innerhalb des Contigs SDR von *C. thummi* recht weit distal liegende Gen *rpn5-like* kodiert für ein hoch-konserviertes Protein. Hierbei handelt es sich bei RPN5 um eine Untereinheit des Deckels des 26S-Proteasoms mit dosisabhängiger regulatorischer Funktion. Der Zusammenbau des Proteasoms erfolgt an der inneren Membran des Zellkerns. Dies konnte durch Versuche mit GFP-markiertem RPN5 in *S. pombe* gezeigt werden. Da RPN5 ein offensichtliches Kern-Lokalisierungssignal (NLS) fehlt, ist davon auszugehen, dass das Protein den Zellkern als ein Komplex zusammen mit anderen Proteinen erreicht (Yen et al. 2003). Das Protein RPN5 ist für den korrekten Aufbau sowie die Funktion des Proteasoms wichtig (Yen et al. 2003; Isono et al. 2007). Die Funktion des 26S-Proteasoms liegt im Abbau von Proteinen des Zytoplasmas sowie des Zellkerns von Eukaryoten. Diese liegen in den meisten Fällen bereits im Vorfeld ubiquitiniert vor. Bei dem 26S-Proteasom handelt es sich um einen ATP-abhängigem Proteasekomplex (Voges et al. 1999; Book et al. 2009). Das Protein ist in der Bäckerhefe (*Saccharomyces cerevisiae*) essentiell, obwohl die genaue Funktion von RPN5 bislang nicht geklärt werden konnte (Yang et al. 2004).

In *Arabidopsis* konnten mit *rpn5a* und *rpn5b* zwei Isoformen von *rpn5* gefunden werden. Die Funktionen der entsprechenden Proteine sind hierbei in großen Teilen redundant. Allerdings konnten nach Überexpression der *rpn5*-Gene abnormale Phänotypen beobachtet werden. Dies spricht für zusätzliche spezifische Funktionen der beiden Isoformen. Pflanzen, denen beide Isoformen fehlen, sind nicht lebensfähig. Fehlt lediglich das Protein RPN5b, so hat dies keine phänotypischen Auswirkungen. Fehlt dagegen das Protein RPN5a, bleibt die Pollenproduktion aus. Außerdem ist die Entwicklung von Embryo und Keimling gestört. Das Fehlen von RPN5a äußert sich in einer herabgesetzten Aktivität und Funktionalität des 26S-Proteasoms, die sich in einer Anreicherung an ubiquitinierten Proteinen sowie einer Hochregulation anderer Proteasomen-Einheiten zeigt (Book et al. 2009). In *S. cerevisiae* und *S. pombe* treten in *rpn5*-Mutanten Fehler in der Zellteilung auf. Dies könnte die ausbleibende Pollenproduktion sowie die veränderten Zellteilungsraten in *Arabidopsis* erklären. Zusammengenommen sind dies Hinweise auf die Bedeutung des 26S-Proteasoms sowie den Abbau ubiquitinylierter Proteine an den zahlreichen Kontrollpunkten innerhalb eines Zellzyklus. Offensichtlich spielt RPN5 eine entscheidende Rolle in der Mitose und Meiose (Yen et al. 2003; Isono et al. 2007; Book et al. 2009).

Die hohe Konserviertheit der Funktion des Proteins RPN5 lässt sich durch „Rescue“-Versuche zeigen. So kann das Protein RPN5 aus *Arabidopsis* die Funktion des homologen Proteins der Hefe ersetzen. Und auch das homologe Protein des *p55*-Gens aus *H. sapiens* (Finley et al. 1998) ersetzt die Funktion des Homologen in der Hefe (Yen et al. 2003; Yang et al. 2004). Ein Sequenzvergleich der RPN5-Proteine aus *C. thummi*, *D. melanogaster* und *A. thaliana* zeigt eine Übereinstimmung von etwa 62 % zwischen *C. thummi* und *D. melanogaster* sowie etwa 34 % und 40 % zwischen *C. thummi* und *A. thaliana* (RPN5a, RPN5b).

Im RPN5-Protein ist im C-terminalen Bereich eine PCI-Domäne lokalisierbar. Diese Domäne konnte bisher nur in eukaryotischen Proteinen gefunden werden. Die PCI-Domäne, für die eine α -helikale Struktur vorhergesagt wurde, ist verantwortlich für die Protein-Protein-Interaktion. Proteine mit dieser Domäne sind oft Bestandteil von Multiprotein-Komplexen (Glickman 2000; Glickman und Ciechanover 2002; Fleige und Pfaffl 2006).

Das viel diskutierte potentielle Kern-Lokalisierungs-Signal konnte nach Analyse der RPN5-Proteine aus *Arabidopsis thaliana* (RPN5a, RPN5b) sowie aus *D. melanogaster* und *C. thummi* nur in RPN5a von *A. thaliana* bestätigt werden. In allen anderen Proteinen wird ein NLS durch bioinformatische Analysen nur sehr vage vorhergesagt. Da für RPN5 auch ein Zellkernimport in einem Komplex mit anderen Molekülen diskutiert wird (Yen et al. 2003), ist es fraglich, ob das Protein RPN5 überhaupt auf eine NLS-Domäne angewiesen ist.

Auf Grund der beschriebenen Konserviertheit von RPN5 ist auch in *C. thummi* eine entsprechende Funktion innerhalb des 26S-Proteasoms recht wahrscheinlich. Somit wäre RPN5 an Prozessen innerhalb der Mitose und Meiose beteiligt. Mit Hilfe der PCI-Domäne wäre jedoch auch eine Interaktion mit weiteren Proteinen innerhalb von anderen Prozessen wie z.B. geschlechtsbestimmenden Schritten möglich. Da RPN5 keine DNA-Bindedomänen zeigt, wäre jedoch für eine entwicklungsbestimmende

Funktion eine Interaktion mit weiteren Proteinen nötig. Ob für den Faktor *M* eine direkte Interaktion mit der DNA nötig ist und ob diese mögliche Interaktion nicht auch im Rahmen eines Proteinkomplexes erfolgen kann, ist nicht bekannt.

4.3 Zusammenfassung und Ausblick

Im Rahmen der vorliegenden Arbeit wurde ein Bereich aus der geschlechtsbestimmenden Chromosomenregion, der „Contig SDR“, von *C. tentans* mit einer Größe von ~87 kb untersucht. Zur Erstellung des Contigs wurden 8 BAC-Klone aus *C. tentans* isoliert, teilweise subkloniert und sequenziert. Innerhalb des Contigs SDR konnten insgesamt 13 Gene sowie ein Teilbereich des Gens *rpS5*-like im distalen Bereich des Contigs SDR identifiziert werden. Hierbei handelt es sich um dieselben Gene, welche schon im Contig SDR von *C. thummi* identifiziert werden konnten. Ein Vergleich der beiden Contigs zeigt, dass die Abfolge der Gene zwischen den beiden Arten *C. thummi* und *C. tentans* identisch ist. Weiterhin konnten im Contig SDR von *C. tentans* sechs Bereiche lokalisiert werden, in denen repetitive oder transposable Elemente zu finden sind (siehe Tabelle 3.2).

Ein Vergleich der larvalen Transkripte (L4-Stadium) von 11 Genen des Contigs SDR aus *C. thummi* ♂ und *C. thummi* ♀ per RT-PCR und Hochdurchsatz-Sequenzierung zeigte mit Ausnahme der Gene *luc7(p)*-like sowie *fs(1)K10*-like bislang keine weiteren geschlechtsspezifischen Unterschiede. Im Gen *luc7(p)*-like konnte in *C. thummi* ♀ und *C. piger* ♀ ein alternativ gespleißtes Intron im eigentlichen Exon 2 identifiziert werden. Bei *fs(1)K10*-like konnte in *C. thummi* ♂ eine Duplikation des Gens nachgewiesen werden.

Weiterhin wurden mit Hilfe des RACE-Verfahrens die 5'UTR- und 3'UTR-Bereiche der Transkripte analysiert. Hierdurch konnten differentielle Spleißprodukte identifiziert werden, welche jedoch nicht geschlechtsspezifisch auftreten.

Die bioinformatische Bearbeitung der von den Genen der SDR kodierten Proteine auf konservierte Domänen zeigt vier Proteine, die möglicherweise als Transkriptions- oder Spleißfaktoren wirken können. Hierbei handelt es sich um die Proteine der Gene *mi-er1*-like, *luc7(p)*-like, *polyhomeotic*-like sowie *rpn5*-like. Für das auf Grund der männchenspezifischen Duplikation in *C. thummi* in den Fokus geratene Genprodukt von *fs(1)K10*-like konnten keine Domänen vorhergesagt werden. Somit kann nicht gesagt werden, ob das Protein im Rahmen der Geschlechtsdetermination eine Funktion haben könnte.

Die Sequenzidentität der abgeleiteten AS-Sequenz liegt für alle Gene außer *mi-er1*-like (87,6 %) zwischen den beiden Arten *C. tentans* und *C. thummi* bei 93,3 %.

In Zukunft sollte weiterhin versucht werden, eine möglichst vollständige Sequenz des Proto-Y-Chromosoms des Contigs SDR von *C. thummi* zu erhalten. Dies kann über Screenings von männchenspezifischen BAC-Bibliotheken mit Proto-Y-Chromosom-spezifischen Sonden erreicht werden. Darüber hinaus sollte versucht werden, auch das distale Ende des männchenspezifischen *Cla*-Element Clusters zu bestimmen.

Für eine detailliertere Analyse der Gene der SDR sollte weiterhin versucht werden, alternative Transkripte mit möglichen geschlechtsspezifischen Unterschieden zu erhalten. Hierfür sollten zusätzlich auch frühere Entwicklungsstadien als Larven (z.B. Embryonen) untersucht werden. Dazu wäre eine Hochdurchsatz-Sequenzierung von embryonaler RNA von *C. thummi* gut geeignet.

Mit Genen der SDR, deren Proteine in der Literatur als im Zellkern lokalisiert oder als möglicher Transkriptionsfaktor beschrieben werden, sollten Versuche zum „gene silencing“ durchgeführt werden. Als Kandidaten für diese Versuche bieten sich die fünf Gene *mi-er1-like*, *luc7(p)-like*, *fs(1)K10-like*, *polyhomeotic-like* sowie *rpn5-like* an. Diese weisen alle ein potentiell Kern-Lokalisierungs-Signal auf. Weiterhin werden die entsprechenden Proteine in der Literatur als Protein-DNA- oder Protein-Protein-interagierend beschrieben. Die für diese Interaktionen bekannten Domänen sind in den abgeleiteten AS-Sequenzen der genannten Gene zumindest laut bioinformatischer Analyse mit Ausnahme von *fs(1)K10-like* vorhanden.

5 Anhang

5.1 Standardlösungen und Materialien

Agarosegel (1 %)	1 % (w/v) Agarose in 1 x TBE
Agar-Platten (500 ml) (mit Antibiotikum und "Blau-Weiß"-Selektion)	15 g Agar-Agar 5 ml Antibiotikum-Stammlösung 0,2 mmol IPTG (dioxanfrei) 0,5 % (w/v) X-Gal in Dimethylformamid in 500 ml 1x LB-Medium
Ampicillin-Stammlösung (100x)	10 mg/ml Ampicillin in HPLC-Wasser
„Antifadingmittel“ (10 ml)	0,01 g para-Phenylendiamin 6,6 g Glycerin in 1 x PBS
Bromphenolblau DNA-Ladepuffer	4 M Harnstoff 0,1 M EDTA- Dinatriumsalz 50 % (w/v) Saccharose 0,1 % (w/v) Bromphenolblau in HPLC-Wasser
DEPC-Wasser	1 ml Diethylpyrocarbonat in 1 Liter VE-Wasser
Chloramphenicol-Stammlösung (100x)	1,5 mg/ml Chloramphenicol in HPLC-Wasser
Denaturierungspuffer (Southern-Blot)	50 mmol/l NaOH 1,5 mol/l NaCl
Dialyse-Puffer (10x)	3 M Natriumchlorid 0,25 mM Tris-HCl 1 M EDTA-Dinatriumsalz in VE-Wasser
Elektrophoresepuffer (10x) (E-Puffer)	0,36 M Tris 0,3 M NaOH 0,1 M EDTA-Dinatriumsalz in VE-Wasser
Ethidiumbromid-Färbelösung	0,001 % (v/v) EtBr-Stammlösung in 1x E-Puffer
Ethidiumbromid-Stammlösung	0,5 % (w/v) Ethidiumbromid in 1x E-Puffer

Glycerintöpfchen (pro Töpfchen):	500 µl 2 x L-Medium 500 µl Glycerin
GTC-Stammlösung	4,4 ml 0,75 M NaCitrat pH 7,0 6,6 ml 10 % N-Lauroylsarkosin 62,5 g Guanidin-Thiocyanat In 73,3 ml DEPC-Wasser
GTC-Arbeitslösung	12,5 ml Stammlösung 87,5 µl β-Mercaptoethanol
Hybridisierungspuffer	1x PM (in 3 x SSC) 0,5 % SDS
LB-Medium	0,005 % (w/v) Natriumchlorid 0,005 % (w/v) Hefe-Extrakt 0,01 % (w/v) Trypton in VE-Wasser
Lösung I (Whitehead)	10 mM EDTA-Dinatriumsalz in VE-Wasser; pH 8
Lösung I (96er Plasmidpräparation) (Whitehead)	+7,5 mg RNase A/ 100 ml Lösung I
Lösung II (Whitehead)	0,2 M Natriumhydroxid 1 % (w/v) Sodiumdodecylsulfat (SDS) in VE-Wasser
Lösung II (96er Plasmidpräparation)	0,2 M Natriumhydroxid 2 % (w/v) Sodiumdodecylsulfat (SDS) in VE-Wasser
Lösung III (Whitehead/ 96er Plasmidpräparation)	1,875 M Kaliumacetat in VE-Wasser; pH 5,2
Neutralisierungspuffer (Southern-Blot)	0,5 mol/l Tris/ HCl, pH 7,4 3 mol/l NaCl
PBS/BSA	10 mM MgCl ₂ 10 mg/ml BSA in 1x PBS
Penicillin/ Streptomycin-Stammlösung (100x)	10.000 U Penicillin+10 mg/ml Streptomycin in HPLC-Wasser oder 0,9 % NaCl (steril-filtriert)

Phenol/ Chloroform/ Isoamylalkohol (25:24:1)	50 % 50 mM Phenol 48 % Chloroform 2 % Isoamylalkohol
<u>Phosphate Buffered Saline</u> (10x) (PBS)	1,37 M Natriumchlorid 30 mM Kaliumchlorid 10 mM Kaliumdihydrogenphosphat 0,1 M Dinatriumhydrogenphosphat- Dihydrat in VE-Wasser
Präinkubationsmedium (PM) (nach Denhardt, 1966)	0,02 % (w/v) Ficoll 400 0,02 % (w/v) Polyvinylpyrrolidon 0,02 % (w/v) BSA in 3 x SSC
<u>Standard-Saline-Citrate-Puffer</u> (20 x SSC)	3 M NaCl 0,3 M Natriumcitrat
<u>Tris-Borat-Elektrophorese-Puffer</u> (1x TBE)	90 mM Tris 90 mM Borsäure 25 mM EDTA-Dinatriumsalz in VE-Wasser
TE-Puffer (Whitehead)	10 mM Tris, pH 8,0 50 mM EDTA-Dinatriumsalz, pH 8,0
50T/ 50E-Puffer	50 mM Tris 50 mM EDTA-Dinatriumsalz, pH 8,0
Tris-Phosphat-Elektrophoresepuffer (E-Puffer)	0,36 M Tris 0,3 M Natriumdihydrogenphosphat 0,1 M EDTA-Dinatriumsalz
Zellkulturmedium (500 ml) komplementieren	50 ml FCS 5 ml Pepton- Stammlösung (10 %) 5 ml Pen/ Strep-Stammlösung 50 µl Insulin 1 ml 0,5 % Phenolrotlösung (Farbindikator) 33 ml L-Glutamin (bei „Promo Cell Schneider’s Drosophila Medium“)

5.2 Bezugsquellen

Agarose	Roti®GAROSE Fa. Roth (Karlsruhe, Deutschland)
	Agarose Typ II Fa. Sigma-Aldrich (St. Louis, USA)
$\alpha^{32}\text{P}$ dATP	Fa. Analytic Hartmann (Braunschweig, Deutschland)
Ampicillin	Fa. Roche (Mannheim, Deutschland)
Antarctic Phosphatase	Fa. New England Biolabs (Ipswich, UK)
Chloramphenicol	Fa. Sigma-Aldrich (St. Louis, USA)
Dialyseschläuche	Fa. Medicell International (London, UK)
DNA-Molekulargewichtsmarker	„100 bp Plus DNA Ladder“ Fa. Fermentas (St. Leon-Roth, Deutschland) „Lambda DNA x HindIII“ Fa. Fermentas (St. Leon-Roth, Deutschland) „1 kb Ladder“ Fa. New England Biolabs (Ipswich, UK)
DNase I	Fa. New England Biolabs (Ipswich, UK)
dNTP's/ dUTP	Fa. Fermentas (St. Leon-Roth, Deutschland)
Elektroporationsküvetten	Fa. Bio-Rad (München, Deutschland) Fa. Biozym (Oldendorf, Deutschland)
Exonuklease I	Fa. Fermentas (St. Leon-Roth, Deutschland)
Fast-AP	Fa. Fermentas (St. Leon-Roth, Deutschland)
FCS	Fa. PAA (Cölbe, Deutschland)
Insulin	Fa. Sigma-Aldrich (St. Louis, USA)
IPTG	Fa. Fermentas (St. Leon-Roth, Deutschland)
Labeling-Kit	„Random Primed Oligo Labeling Kit“ Fa. Roche (Mannheim, Deutschland) „DIG DNA Labeling Kit“ Fa. Roche (Mannheim, Deutschland) „Bio Nick™ Labeling System“ Fa. Invitrogen (Carlsbad, USA)
L-Glutamin	Fa. Promocell (Heidelberg, Deutschland)
Nebulizer	Fa. Life Technologies (Carlsbad, USA)
Nitrozellulosefilter	Fa. Whatman (Dassel, Deutschland)
Nucleo Spin® Extract II Kit	Fa. Machery-Nagel (Düren, Deutschland)
PBS	Fa. PAA (Cölbe, Deutschland)
Penicillin/ Streptomycin	Fa. Sigma-Aldrich (St. Louis, USA)
Pepton	Fa. Sigma-Aldrich (St. Louis, USA)
Phenolrot	Fa. Sigma-Aldrich (St. Louis, USA)
PIPES	Fa. Sigma-Aldrich (St. Louis, USA)
Plasmidpräparation	„peqGold Plasmid Miniprep Kit I“ Fa. Peqlab (Erlangen, Deutschland) „PureYield™ Plasmid Maxiprep System“ Fa. Promega (Fitchburg, USA)

	„PureYield™ Plasmid Miniprep System”
	Fa. Promega (Fitchburg, USA)
Poly-A ⁺ Kit	„NucleoTrap® mRNA Mini Kit”
	Fa. Macherey-Nagel (Düren, Deutschland)
Primer	Fa. Invitrogen (Carlsbad, USA)
	Fa. Biomers (Ulm, Deutschland)
Proteinase K	Fa. Merck (Darmstadt, Deutschland)
3MM-Papier	Fa. Whatman (Dassel, Deutschland)
RACE-Kit	„GeneRacer® Kit with SuperScript® III RT and TOPO TA Cloning® kit for Sequencing“
	Fa. Invitrogen (Carlsbad, USA)
Restriktionsendonukleasen	Fa. Fermentas (St. Leon-Roth, Deutschland)
	Fa. Roche (Mannheim, Deutschland)
	Fa. New England Biolabs (Ipswich, UK)
Reverse Transkriptase	„SuperScript™ III Reverse Transcriptase”
	Fa. Invitrogen (Carlsbad, USA)
RNA 6000 Nano Kit (Bioanalyzer)	Fa. Agilent (Santa Clara, USA)
RNA 6000 Pico Kit (Bioanalyzer)	Fa. Agilent (Santa Clara, USA)
RNase A-Lösung	Fa. Peqlab (Erlangen, Deutschland)
Röntgenfilm	Fa. Fuji (Tokyo, Japan)
Röntgenfilm-Entwickler/ -Fixierer	Fa. Adefo-Chemie (Dietzenbach, Deutschland)
Sephadex™ G50-Fine	Fa. GE Healthcare (München, Deutschland)
Taq-Polymerase	„GoTaq®-Polymerase”
	Fa. Promega (Fitchburg, USA)
	“LongAmp® Taq DNA-Polymerase”
	Fa. New England Biolabs (Ipswich, UK)
T-Flaschen (Zellkultur)	Fa. Greiner GmbH (Frickenhausen, Deutschland)
T4 DNA-Polymerase	Fa. Roche (Mannheim, Deutschland)
T4 DNA-Ligase	Fa. Roche (Mannheim, Deutschland)
	Fa. Fermentas (St. Leon-Roth, Deutschland)
Uracil DNA Glykosylase (UDG)	Fa. New England Biolabs (Ipswich, UK)
Vektor	„pGem®T Easy“, Fa. Promega (Fitchburg, USA)
Verstärkerfolie	Fa. Dupont; Dr. Goos-Suprema (Heidelberg, Deutschland)
X-Gal	Fa. Fermentas (St. Leon-Roth, Deutschland)
Zellkulturmedium	„culture medium ZO buffered with PIPES ”
	Fa. Cell culture technologies (Lausanne, Schweiz)
	„Promo Cell Schneider´s Drosophila Medium“
	Fa. Promocell (Heidelberg, Deutschland)
96er „Multiple Well Platten“	Fa. Sarstedt (Newton, USA)

5.3 Primer

5.3.1 Analyse der SDR bei *C. tentans*

Tabelle 5.1: Primer SDR-Analyse *C. tentans*

L13A1_1-4T7_P1	GTGGCTCTTGTGATTCAAAC
L13A1_1-4T7_P2	TGATGCATCAAATTGACAGC
L13A1_1-4T7_P3	CTCCCTGATCTACTGGTC
L13A2_1-5_1r	CTTTACACGATACTCTCCAC
L13C2_1-1_2350bp	CACCAACAACCATCGTCAG
L13G3_2-3forAbbruch	ACATGTGGTCGCTGTTTCG
L13G3_2-3revAbbruch	GACCGCACTATGAACTGC
L13G3_2-3revAbbruch_2	CATCCACACGCATAGACC
L13H4_1-1T7_P1	CCTTTCTTTAATCGTGCTG
L13H4_1-1T7_P1rev	GTCTCTTCATCATCGGTC
L13H4_1-1T7_P2	GAGTATCCTGGGATGCAC
L13H4_1-1T7_P2rev	ACGTCCATTCAGCGGCAG
L13H4_1-1T7_P3	CGGTTTCGTTCCGAAATGAC
L13H4_1-1T7_P3rev	CGAGTCATCATCATCATCC
L13H4_1-1T7_P4	GCGGTGCTTATGATGAGG
L13H4_1-1T7_P4rev	GGATGTAGAACAATTGCTGC
L13H4_1-1T7_P5	GCAGCAATTGTTCTACATCC
L13J21_T7_710bp	CGCTTTGCTGATCATTTCGC
L13J21_T7_for	AATCCTTGGCGTCAACGC
L13J21_T7_rev	GCGAATGATCAGCAAAGCG
L13J21T7_P1	CGCCAAGGATTTTACTATG
L16A1-3Sp6_P1	GACCAGTAGATCAGGGAG
L16D3-1T7_P1	CTTTCTGTGACATGAGC
L16F3-1T7_P1	GACGCACTCTTGGGATAG
L16F3-1T7_P1rev	CATAGCAAACAATAGAACGG
L16F3-1T7_P1rev2	CCATAACTAAGCCGACCATC
L16F3-1T7_P2	GTCAGGAAGAAGAAAGGAAC
L16F3-1T7_P2rev	ACTAGCCATGCAATATTGC
L16F3-1T7_P3	CCACATGAGAAGCTATAGC
L16F3-1T7_P3rev	GTACGTGCGGAGATCATG
L16F3-1T7_P4	GGTGTGCGTCCATCATCG
L16F3-1T7_P4rev	GAATCATCTACTGAAATCGG
L16F3-1T7_P5	CAGTAGATGATTCTGGCG
pGem_L13J21T7_P1rev	GTCGTCCTTGACCTCTCC
pGem_Sp6	CAGCCACAGCAAATACAGC
pGem_T7	CGCATTCCGAGCATTGTG
pGem_THE133	TGGCGTAGCAAAGGGTG
pGem_THE135	CTCCCTTTTCTGGACTAGC
pGem_THE158for	CAACAGGAGAAAGTGAGG
pGem_THE158rev	GTCATCGTGGCCATATCG

pGem_THE159for	CTCTCATCGTCGTCATCATC
pGem_THE159rev	CAATGATGATGATGACGACG
Primer_THE100_721bp	CCCACCATCTAGTCCATG
Primer_THE99_668bp	TGTTCAATTTGGCGGACGG
Primer1_L13A2_1-5_A11P2f	GACGAAGCTAAGGCTGC
Primer1_L13C2_1-1_224b	CGTGAAATCGGATAGAGG
Primer1+Sp6	CGTGACAAGAATAAGAGTC
Primer2_L13A2_1-5_A11P2f	CATACTCATTGGAAGCTCC
Primer2_L13C2_1-1_896bp	CAGGGTAGATCTACGTGC
Primer2_L13C2_1-1_903b	CTGGTTTACAGGGTAGATC
Primer2+Sp6	CTTATGTGGCAGTTATCAGC
tent01_L13B4_525_rev	GATCGGCATGTGTTCAAGAC
tent02_L13B4_1593_rev	TATCCATTCTTAACTGCCAC
tent03_L13B4_7752_for	AAAGGTCCGAGGTAGAGG
tent04_L13B4_10489_rev	GGTTGAAAGTTGAATGTTGG
tent05_L13B4_14907_for	ACCGAGTTGCTGATAATGC
tent06_L13B4_16152_for	AAGAATAAGGCAATGCTGAG
tent07_L13B4_16600_for	ATGAGGATTGGTAGTTCTGC
tent08_L13B4_17449_for	TCCATTGTCAAAGCAGGCTG
tent09_L13B4_18931_rev	GTAATTGACGAGAAGTTGAC
tent10_L13B4_18653_rev	CCACCGATAATCTCAAGCC
tent11_L13G3_19800_for	AACAAATCGCCCTTAACAGC
tent12_L13G3_20890_rev	GGACATAATAGAAACAGGAG
tent13_L13G3_22207_for	TTTAATGAACCACAGCAACC
tent14_L13G3_27187_for	ACCCAAACTAAACAACCCC
tent15_L13A2_28768_for	TACTACACTCTTGAAATGCC
tent16_L13A2_30569_rev	ACATTAAACTCCAGTGCCAG
tent17_L13A2_36312_for	ACATTATGATCGTTTCTCTCC
tent18_L16F3_46049_rev	TCAATCAATCTGCTTCTTCC
tent19_L16D3_46856_rev	TACCCAATTCACACATACAC
tent20_L16D3_49640_rev	TTTCACTTCTCCATTCC
tent21_L16D3_49127_rev	AAACTTTCCCATCATCTGAC
tent22_pGem5_53443_rev	GTATATGGTAGGAAAGGCAC
tent23_pGem5_55352_rev	TCTATCACAGGTCACGCC
tent24_L13J21_59754_rev	GGACGAATGTAATCCTTCC
tent25_L13J21_59376_rev	CAATGTTTGTGGAGATGGG
tent26_L13J21_58792_rev	ACCATTCCTGATAAATTGC
tent27_L13J21_58298_rev	GTTTCTGTGTGGTGATTGTG
tent28_L13J21_58075_rev	GTGTATGATTGTGGACTCTC
tent29_L13C2_67957_for	TGGATATGGATAAGAGTTCCG
tent30_L13C2_69317_rev	CATGATTGCGAAGGTTAAGG
tent31_L13C2_70675_for	TTAATCTGAAAGGTCATGGG
tent32_L13C2_71746_for	TTCCCATCACCCATCATCC
tent33_L13C2_83342_rev	GTGTCGGCAGCAAATTCG
tent34_L13C2_85994_rev	GGTGTGTATCATCCATGCC

tent35_L13C2_86618_for	GAGATTTTCATGCTAGTTTCG
tent36_L13C2_72565_rev	TGCAAGAATCAATAGACTCC
tent37_pGem5_55700_for	GTGTTGTGATCGGAATGGC
tent38_pGem5_56220_for	GGGATTGAATTTGGAAATGTG
tent39_L13J21_59600_for	CCACACGTTTGATTCTGAC
tent40_L13C2_61870_for	CCTTACAAATGTTTCATGGAC
tent41_L13B4_360_rev	TATGCTTCTCGAACACACG
tent42_L13B4_240_rev	TTGTGATGATCTGCTGCTG
tent43_L13B4_8090_for	CTATATGGCAGTAATCAAACC
tent44_L13B4_10643_rev	ACCTAATACAGAGAAAGTTGC
3250bp_proximal	CGTTTGTGACTGTAATTTCC
5300bp_proximal	CTCTTAATTCTCCTTCGACG
5650bp_proximal	CAACTCTTTATTATCGCACTC
5750bp_distal	TCTCACATCATAAGGAATGG
6300bp_distal	CTCTAATCCTTTTCGTAATCG
C64_left	CTTGACGATGTGACTCGC
C64_right	CCCATTCTGCTTCTGTCTG
C65_left	CGGCATTCTCGTCGATC
C66_left	GATATCCGTTAGGTGTGG
C66_right	GCCTAAGTTCCACTAAATGTG
C67_left	CATAATCAGCATCGAGAAAG
C67_right	GGGGCGGATTTCAAAAATC
C68_left	ATCGCACTTCCTTAATATGG
C68_right	GAAAGAGGAAGGTGTTACTG
C69_left	TTGCGTTAGTATTTGCACTG
C69_right	GAATTCTATGCGCAAGCG
C70_right	ACTGGAAGAGAGAGTGAG
C71_left	AGAATATTCGATGCAGAGTC
C71_right	CCTGCAGCTACATTAATAATCC
Ct1fL13B4_1-1Sp6	TATTGCCTCGTCACTGAC
Ct1r2L13A2_2-5Sp6	GAGGTGATTCTGTTATTCG
Ct1rL13A2_2-5Sp6	CCTGGAGTCAAGAGTAAG

5.3.2 RT-PCR und RACE Primer

Tabelle 5.2: Genspezifische Primer RT-PCR

dLacDH_1for	GCAACTTAAAATCATCAATCG
dLacDH_1rev	TCTGGACTCTTTGAATAGTG
dLacDH_2for	GAGGTGCTTCTAATGTGCG
dLacDH_test31kb_exon_for	CCTTAAACCGTGTGTGAG
dLacDH_test31kb_exon_rev	CGATTGATGATTTTAAGTTGC
fs1K10_1for	ATGTTGGCTGGAAATAAGAG
fs1K10_1rev	AACTTCTTCCTCCTTCTTCG
hermanskypudlak_1a_for	ATAATCCATCCATCGTTACG

hermanskypudlak_1rev	CATTTAGGAAGCACTTTGAG
hermanskypudlak_2for	TACAATAAGTCAGCAAGAGG
hermanskypudlak_2rev	GCAAATCGACTAAACAACTC
hermanskypudlak_3for	GTCTAATCTTCTATGAAATGGC
hermanskypudlak_3rev	TATGATGACTTCCAATGTGC
hermanskypudlak_4for	AGAATGAAAGAAAGCACAGC
hermanskypudlak_4rev	CTTCAAACACATCAAGAC
hypProtein71kb_for	GGATTTAATAAGAGACCGTC
hypProtein71kb_rev	TGTTATTGCTTTCTCCATCC
K10_male2_760for	GACAAAAGTCGACAAAAGCG
K10fem_1250rev	CAACATTAGCTACAGAAACAGC
K10fem_650for	CACCTATGCCTCCAATGCG
K10male1_1670rev	GATCAGCATATGTTTGAAGGC
K10male2_2050rev	CTTTTTTCATCATTTTCATATTCACAC
luc7_1for	TGAATTGATGGGAAGGAACC
luc7_1rev	GAAGACTTGTAATCATCACG
luc7_2for	GAAGGGATAGAGATAGACAG
luc7_2rev	GACATTGTAAGGAAAACAAACG
mier1_1for	ATGGATATGGAAGTGGACC
mier1_1rev	TGCTGCTCGTTTCATTTTCG
mier1_GSP_14900for	GCCGTGATATGAAAACGATGAAGGTG
polyh_test30kb_exon_for	CAGTATTTGAAGTTTAGTCCG
polyh_test36kb_exon_for	CCACGAGGTTTCTTTACG
polyh_testupstream_exon_rev	GATGTGGACTTTGTGGTG
polyhomeotic_1for	CAACACAACAGCAACAAGG
polyhomeotic_1rev	TGTTGATTCGTGGCTTGAG
polyhomeotic_2b_rev	GAGTTTGTGGTGATGAGAG
polyhomeotic_2for	TTCCAACAGAGTCCAACG
polyhomeotic_2rev	CTGCTGAAGTTGTAATTGAG
polyhomeotic_3for	CCACAATTACAACAGCAAGG
polyhomeotic_3rev	GACAATTTTAGGGCTTTTAC
polyhomeotic_last_rev	TATTGTGGCGATTGATCTCC
rpn5_1for	TGGAAGTAGATTATAGCAGC
rpn5_1rev	GTTTCATCGAGATCCAAAAGC
wd74_1for	GGAACACAGACAGGAAGC
wd74_1rev	CTGTTTACTTGATTCCTTACG

Tabelle 5.3: Primer 5' RACE

46_52kb_5UTR_46860	GTCAGTCTTCCCAGTATCCACATTTGC
46_52kb_5UTR_472700	CAACTGATGCTACTGACGATGACG
69_75kb_5UTR_69870	CCCGTTATAGCCTGAGCAGTTCC
69_75kb_5UTR_70200	GGGCTTTCCCTCCTCATCCGTAGC
bcn92_5UTR_24720	CCAATCGACGACTTTGATGAGTAGC
bcn92_5UTR_25100	GGTTCCATCTGTTATTGACTTGTTCTC

bcn92_5UTR_GSPfor	TAATTCAACATACATTTGGCTC
dLacDH_5UTR_intern_430	CGCCTGATTGACATCTGACAATCCCAG
dLacDH_5UTR_nested_330	CCACCAACCAAGCCAGTATTTCCACC
fs1K10fem_5UTR_22700	CGCTGATTATTTTCGTCTGGATTTC
hermansky_5UTR_intern_370	CACGAAGTGGTGCCTGAGACTGTGC
hermansky_5UTR_nested_270	CATAATTGCTAATGCGTCCTTCGACGG
hp_5UTR_kurz_GSPfor	CATGAAACAGTGATATAGTGTG
hp_5UTR_lang_GSPfor	TGACAGATACCTTGATGTCTG
hypoth71_5UTR_intern_530	CATGATAACTATGCCGAGCATGTGCCG
hypoth71_5UTR_nested_430	GCTCCATGATCCGCTATGTTGATGCTG
K10fem_5UTR_22700	CGCTGATTATTTTCGTCTGGATTTC
K10fem_5UTR_GSPfor	CTAGGTTAGTGAGACTATTCC
K10fem_5UTR_intern_500	GCATTGGAGGCATAGGTGGCATTGG
K10fem_5UTR_nested_320	GCGTCTCTGAGGTGGCTGATTATTCC
luc7_5UTR_intern_550	GCATACCTCCATTTCTTCTGCTCCG
luc7_5UTR_nested_420	CACGCCAGCTTCTTCTGCTTCACG
mier1_5UTR_14250	GCTGTTGGATCGAGTTTCAATTCTCCGTC
mier1_5UTR_14440	CTTCACTCTTTTCTGGAATTCTTGTGG
mier1_5UTR_intern_350	GCGTCTCTTCTCGATGATGACGATGC
mier1_5UTR_nested_220	GTTCTTTGCGGCCATTGCTTCTTCG
polyhom_5UTR_37600	GGACGTTCTTATTTGGGGTTCTCG
polyhom_5UTR_intern_600	GAAACTTTGAGCTCCTGACGCTGTCCG
polyhom_5UTR_nested_550	CCCGCCGTGCTAAGCATTGTGG
rpn5_5UTR_intern_170	CTGATCCGATACGAGCTTGCTTCTC
rpn5_5UTR_nested_30	CACGATTCTGCCTCCCTCATATGTAGG
rpn5_GSP_5UTR_75640	CTAAAAACATTACACAAGTCTCAC
wd74_5UTR_20950	TCCTCGTCACTGACATCGACAGTCTG
wd74_5UTR_21500	GTTGTTCTTTCTATCCTTGCCACCGAC
wd74_5UTR_intern_320	GCGACAATTGTGTCTTGACTCCATGCC
wd74_5UTR_nested_280	CCTTCGGGAATCGATAGGTCTGTCTC

Tabelle 5.4: Primer 3' RACE

46_52kb_3UTR_48940	GTGTCAATTTGTGAAGACGACGAGG
46_52kb_3UTR_49440	GGATACACAGATGCTGATGATGGAC
69_75kb_3UTR_71980	CAGCCACAACAACCACAACAAAGTG
69_75kb_3UTR_72420	CTGGAGAAGGAGGGGAGGAGG
bcn92_3UTR_24720	GCTACTCATCAAAGTCGTCGATTGG
bcn92_3UTR_25100	GAGAACAAGTCAATAACAGATGGAACC
dLacDH_3UTR_intern	GGCAAGTACAACAATGTGC
dLacDH_3UTR_nested	CTGGTTATGGTCATGTGG
hermansky_3UTR_intern	TGCCATTAACCAAGCCG
hermansky_3UTR_nested	GTTTGATAACGATTGGAGC
hypo71kb_3UTR_intern	CAACAAACAATTAGAGTTCCAG
hypo71kb_3UTR_nested	AGTCAGGATGGAGAAAGC

K10female_3UTR_intern	GAAGCTCCATTTGTTACCG
K10female_3UTR_nested	ACACCACCAGCAACAGTC
luc7_3UTR_intern	CCGCAGTCGTTCATAGAC
luc7_3UTR_nested	CCTAGTTTATGGATCTGCG
mier1_3UTR_intern	TACGTGATGATGAGCAGG
mier1_3UTR_nested	TGAATCCTTGTGTAAGTAC
polyhom3UTR_intern	GCTCATTAAACTGTGCTCG
polyhom3UTR_nested	TGATGAAATGGAGCGTCTC
rpn5_3UTR_intern	GAGTTTGGAGGAAATACCC
rpn5_3UTR_nested	GGAATTATTGACTCTAACTG
wd74_3UTR_intern1	CATTAGATCGTTATGTCCG
wd74_3UTR_intern2	CGAGATATTAGTGAAAGAAGC
wd74_3UTR_nested	CATAAATACCTCGGAAGTGG
bcn92_RACEa1_rev	CTGCCTGATATAGATGACC
bcn92_RACEa3_rev	CTTCATCATAAAATGCTTTATTTG
dLacDH_RACE_rev	GGTCATCTATATCAGGCAG
fs1K10_RACEa1_rev	GCGACCTCTTCAACATTAGC
fs1K10_RACEa2_rev	AAATCGTTAGAAAATCCTTTTCG
fs1K10_RACEa3_rev	AAAAGAAAATAAGAGGGTCC
fs1K10_RACEa4_rev	GAATGCTCTAAAACCTTACACG
hp_female_RACE_rev	TGGTTATGGCGTCATAAAGC
hp_male_RACE_rev	TCACGTTACACACAATTTACG
hypo2(69_75)_RACEa1_rev	GCTTCACCTTCAGGCACG
hypo2(69_75)_RACEa2_rev	AGACAGATATAGTAAGTCTTGC
luc7_RACEa4_rev	ATCGCACTTCCTTAATATGG
mier1_RACE_rev	TCCTGTTCCGTTTGACGATC
poly_RACE_a1_rev	ATACCAATTAAGAGTTATATAGC
poly_RACE_a2_rev	TACACTACAAATATCATCCCCG
poly_RACE_for	GGACGAAACAGACTCCCCG
poly_test454_for	GGATATAATAATTGAGAGGTGC
poly5UTR_454_intern	CGTAAAGAAACCTCGTGGACGTTCC
poly5UTR_454_nested	TCTGCTCTGCCTGCTCATTCAAATGC
rpn5_RACE_rev	TCTTTCTGGATGAGATGCG
wd74_RACEa2_rev	TTTCGTAAAAATGATATGGTCG
wd74_RACEa3_rev	TCTGAGTCAGAAAGTCCG

5.3.3 Primer Gene Racer Kit/ dUTP-Kit/ Oligo-dT

Tabelle 5.5: Primer Gene Racer Kit (Invitrogen)

GeneRacer_3UTR	GCTGTCAACGATACGCTACGTAACG
GeneRacer_3UTRnested	CGCTACGTAACGGCATGACAGTG
GeneRacer_5UTR	CGACTGGAGCACGAGGACACTGA
GeneRacer_5UTRnested	GGACACTGACATGGACTGAAGGAGTA

Tabelle 5.6: Oligo-dT Primer

anchoredOligodT_lang	GATCCTAGACTGTGACC(T) ₂₄ VN
anchoredOligodT_mitPrimer	GATCCTAGACTGTGACC(T) ₁₅ VN
nested_zu_anchoredOligodT	GGATCCTAGACTGTGACC

Tabelle 5.7: Primer dUTP Kit

3'T_RACEPrimer	CTCGAGTTCGCGCCGGATC
5'T_RACEPrimer	ATATGCACTGCCGCGTCTGA
AdapterA_dUTP	AUCUCGAGUUCGCGCCGGAUCC(T) ₂₅ VN
AdapterB_dUTP	AUAUGCACUGCCGCGUCUGA(G) ₈
PrimerA_dUTP	CUCGAGUUCGCGCCGGAUC
PrimerB_dUTP	AUAUGCACUGCCGCGUCUGA

6 Literaturverzeichnis

- Acton, A. B. (1957).** "Sex-Chromosome Inversions in Chironomus." *The American Naturalist* 91: 57-59.
- Affara, N. A., et al. (1986).** "Variable transfer of Y-specific sequences in XX males." *Nucleic Acids Res* 14(13): 5375-87.
- Albrecht, E. B. und H. K. Salz (1993).** "The *Drosophila* sex determination gene *snf* is utilized for the establishment of the female-specific splicing pattern of *Sex-lethal*." *Genetics* 134(3): 801-7.
- Altschul, S. F., et al. (1990).** "Basic local alignment search tool." *J Mol Biol* 215(3): 403-10.
- Amrein, H., et al. (1988).** "The sex-determining gene *tra-2* of *Drosophila* encodes a putative RNA binding protein." *Cell* 55(6): 1025-35.
- Amrein, H., et al. (1990).** "Alternatively spliced transcripts of the sex-determining gene *tra-2* of *Drosophila* encode functional proteins of different size." *EMBO J* 9(11): 3619-29.
- Armitage, P., et al. (1995).** "The Chironomidae. The biology and ecology of non-biting midges." Chapman & Hall, London.
- Arnemann, J., et al. (1991).** "Cloning and sequence analysis of a human Y-chromosome-derived, testicular cDNA, TSPY." *Genomics* 11(1): 108-14.
- Aronson, B. D., et al. (1997).** "Groucho-dependent and -independent repression activities of Runt domain proteins." *Mol Cell Biol* 17(9): 5581-7.
- Ayling, L. J. und D. K. Griffin (2002).** "The evolution of sex chromosomes." *Cytogenet Genome Res* 99(1-4): 125-40.
- Babenko, V. N., et al. (2009).** "Characteristics of molecular-genetic organization of intercalary heterochromatin band 10A1-2 in X chromosome of *Drosophila melanogaster*." *Dokl Biochem Biophys* 424: 27-30.
- Balbani, E. (1881).** "Sur la structure du noyau des cellules salivaires chez les larves de *Chironomus*." *Zool. Anz.* 4: 637-641.
- Barbash, D. A. und T. W. Cline (1995).** "Genetic and molecular analysis of the autosomal component of the primary sex determination signal of *Drosophila melanogaster*." *Genetics* 141(4): 1451-71.
- Bauer, H. und W. Beermann (1952).** "Der Chromosomenzyklus der Orthocladiniinen (Nematocera, Diptera)." *Z. Naturforsch.* 7b: 557-563.
- Baur, X., et al. (1980).** "Detection of potent insect antigens for humans: hemoglobins (erythrocrucorins) of chironomids." *Naturwissenschaften* 67: 365-366.
- Beermann, W. (1955).** "[Cytological analysis of a *Camptochironomus* hybrid species. I. Results of crossing and development of the karyotype.]" *Chromosoma* 7(2-3): 198-259.
- Beermann, W. (1955).** "Geschlechtsbestimmung und Evolution der genetischen Y-Chromosomen bei *Chironomus*." *Biol. Zbl.* 74: 525-544.

- Behlke, M. A., et al. (1993).** "Evidence that the SRY protein is encoded by a single exon on the human Y chromosome." *Genomics* 17(3): 736-9.
- Beisswanger, S., et al. (2006).** "Evidence for a selective sweep in the wapl region of *Drosophila melanogaster*." *Genetics* 172(1): 265-74.
- Bell, L. R., et al. (1991).** "Positive autoregulation of sex-lethal by alternative splicing maintains the female determined state in *Drosophila*." *Cell* 65(2): 229-39.
- Berta, P., et al. (1990).** "Genetic evidence equating SRY and the testis-determining factor." *Nature* 348(6300): 448-50.
- Berwanger, C. (2007).** "Sequenzbestimmung und bioinformatische Analyse der geschlechtsbestimmenden Chromosomenregion der Chironomiden." Diplomarbeit, Fachhochschule Bingen.
- Beye, M. (2004).** "The dice of fate: the *csd* gene and how its allelic composition regulates sexual development in the honey bee, *Apis mellifera*." *Bioessays* 26(10): 1131-9.
- Beye, M., et al. (2003).** "The gene *csd* is the primary signal for sexual development in the honeybee and encodes an SR-type protein." *Cell* 114(4): 419-29.
- Birnboim, H. C. und J. Doly (1979).** "A rapid alkaline extraction procedure for screening recombinant plasmid DNA." *Nucleic Acids Res* 7(6): 1513-23.
- Blencowe, B. J., et al. (1999).** "SR-related proteins and the processing of messenger RNA precursors." *Biochem Cell Biol* 77(4): 277-91.
- Blumensatt, M. (2009).** "Klonierung und Analyse der geschlechtsbestimmenden Region von *Chironomus piger* durch in vitro Amplifikation." Diplomarbeit, Johannes-Gutenberg-Universität, Mainz.
- Bogan, J. S. und D. C. Page (1994).** "Ovary? Testis?--A mammalian dilemma." *Cell* 76(4): 603-7.
- Boggs, R. T., et al. (1987).** "Regulation of sexual differentiation in *D. melanogaster* via alternative splicing of RNA from the transformer gene." *Cell* 50(5): 739-47.
- Boissy, R. E., et al. (2005).** "Melanocyte-specific proteins are aberrantly trafficked in melanocytes of Hermansky-Pudlak syndrome-type 3." *Am J Pathol* 166(1): 231-40.
- Bonaccorsi, S. und A. Lohe (1991).** "Fine mapping of satellite DNA sequences along the Y chromosome of *Drosophila melanogaster*: relationships between satellite sequences and fertility factors." *Genetics* 129(1): 177-89.
- Book, A. J., et al. (2009).** "The RPN5 subunit of the 26s proteasome is essential for gametogenesis, sporophyte development, and complex assembly in *Arabidopsis*." *Plant Cell* 21(2): 460-78.
- Bopp, D., et al. (1991).** "Developmental distribution of female-specific Sex-lethal proteins in *Drosophila melanogaster*." *Genes Dev* 5(3): 403-15.
- Bower, N. I. und I. A. Johnston (2010).** "Targeted rapid amplification of cDNA ends (T-RACE)--an improved RACE reaction through degradation of non-target sequences." *Nucleic Acids Res* 38(21): e194.
- Breathnach, R. und P. Chambon (1981).** "Organization and expression of eucaryotic split genes coding for proteins." *Annu Rev Biochem* 50: 349-83.

- Bridges, C. B. (1921).** "Triploid Intersexes in *Drosophila Melanogaster*." *Science* 54(1394): 252-4.
- Bull, J. (1983).** "Evolution of sex determining mechanisms." The Benjamin/Cummings Publishing Company, Inc., Menlo Park, California, USA.
- Bull, J. (1987).** "Temperature-sensitive periods of sex determination in a lizard: similarities with turtles and crocodilians." *J. Exp. Zool.* 241: 143-148.
- Bull, J. J. und R. C. Vogt (1979).** "Temperature-dependent sex determination in turtles." *Science* 206(4423): 1186-8.
- Bull, J. J. und R. C. Vogt (1981).** "Temperature-sensitive periods of sex determination in Emydid turtles." *J Exp Zool* 218(3): 435-40.
- Burge, C. und S. Karlin (1997).** "Prediction of complete gene structures in human genomic DNA." *J Mol Biol* 268(1): 78-94.
- Burghardt, G., et al. (2005).** "The transformer2 gene in *Musca domestica* is required for selecting and maintaining the female pathway of development." *Dev Genes Evol* 215(4): 165-76.
- Burke, T. W. und J. T. Kadonaga (1996).** "*Drosophila* TFIID binds to a conserved downstream basal promoter element that is present in many TATA-box-deficient promoters." *Genes Dev* 10(6): 711-24.
- Burtis, K. C. und B. S. Baker (1989).** "*Drosophila* doublesex gene controls somatic sexual differentiation by producing alternatively spliced mRNAs encoding related sex-specific polypeptides." *Cell* 56(6): 997-1010.
- Charlesworth, B. (1978).** "Model for evolution of Y chromosomes and dosage compensation." *Proc Natl Acad Sci U S A* 75(11): 5618-22.
- Charlesworth, B. (1991).** "The evolution of sex chromosomes." *Science* 251(4997): 1030-3.
- Charnier, M. (1966).** "[Action of temperature on the sex ratio in the *Agama agama* (*Agamidae*, *Lacertilia*) embryo]." *C R Seances Soc Biol Fil* 160(3): 620-2.
- Chase, B. A. und B. S. Baker (1995).** "A genetic analysis of intersex, a gene regulating sexual differentiation in *Drosophila melanogaster* females." *Genetics* 139(4): 1649-61.
- Chirgwin, J. M., et al. (1979).** "Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease." *Biochemistry* 18(24): 5294-9.
- Chomczynski, P. und N. Sacchi (1987).** "Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction." *Anal Biochem* 162(1): 156-9.
- Cirera, S., et al. (1995).** "Molecular characterization of the breakpoints of an inversion fixed between *Drosophila melanogaster* and *D. subobscura*." *Genetics* 139(1): 321-6.
- Clark, M. und W. Wall (1996).** "Chromosomes, the complex code." Chapman and Hall, London.
- Clepet, C., et al. (1993).** "The human SRY transcript." *Hum Mol Genet* 2(12): 2007-12.
- Cline, T. W. (1984).** "Autoregulatory functioning of a *Drosophila* gene product that establishes and maintains the sexually determined state." *Genetics* 107(2): 231-77.

- Cline, T. W. (1984).** "Autoregulatory functioning of a *Drosophila* gene product that establishes and maintains the sexually determined state." *Genetics* 107(2): 231-77.
- Cline, T. W. (1988).** "Evidence that *sisterless-a* and *sisterless-b* are two of several discrete "numerator elements" of the X/A sex determination signal in *Drosophila* that switch *Sxl* between two alternative stable expression states." *Genetics* 119(4): 829-62.
- Cline, T. W. und B. J. Meyer (1996).** "Vive la difference: males vs females in flies vs worms." *Annu Rev Genet* 30: 637-702.
- Cline, T. W., et al. (1999).** "Functioning of the *Drosophila* integral U1/U2 protein *Snf* independent of U1 and U2 small nuclear ribonucleoprotein particles is revealed by *snf(+)* gene dose effects." *Proc Natl Acad Sci U S A* 96(25): 14451-8.
- Cook, J. (1993).** "Sex determination in the hymenoptera: a review of models and evidence." *Heredity* 71: 421-435.
- Credland, P. (1973).** "The taxonomic status of *Chironomus riparius* Meigen and *Chironomus thummi* Kieffer (Diptera: Chironomidae)." *J. nat. Hist.* 7: 209-216.
- Cronmiller, C. und T. W. Cline (1987).** "The *Drosophila* sex determination gene *daughterless* has different functions in the germ line versus the soma." *Cell* 48(3): 479-87.
- Cronmiller, C., et al. (1988).** "Molecular characterization of *daughterless*, a *Drosophila* sex determination gene with multiple roles in development." *Genes Dev* 2(12A): 1666-76.
- Crozier, R. (1971).** "Heterozygosity and sex determination in haplo-diploidy." *Amer. Natur.* 105: 399-412.
- Daneau, I., et al. (1995).** "Bovine SRY gene locus: cloning and testicular expression." *Biol Reprod* 52(3): 591-9.
- Dauwalder, B., et al. (1996).** "A human homologue of the *Drosophila* sex determination factor *transformer-2* has conserved splicing regulatory functions." *Proc Natl Acad Sci U S A* 93(17): 9004-9.
- Deatrick, J., et al. (1991).** "The complex genetic locus *polyhomeotic* in *Drosophila melanogaster* potentially encodes two homologous zinc-finger proteins." *Gene* 105(2): 185-95.
- DeCamillis, M., et al. (1992).** "The *polyhomeotic* gene of *Drosophila* encodes a chromatin protein that shares polytene chromosome-binding sites with *Polycomb*." *Genes Dev* 6(2): 223-32.
- Dell'Angelica, E. C. (2004).** "The building BLOC(k)s of lysosomes and related organelles." *Curr Opin Cell Biol* 16(4): 458-64.
- Deng, W. und S. G. Roberts (2005).** "A core promoter element downstream of the TATA box that is recognized by TFIIB." *Genes Dev* 19(20): 2418-23.
- Denhardt, D. T. (1966).** "A membrane-filter technique for the detection of complementary DNA." *Biochem Biophys Res Commun* 23(5): 641-6.
- Denholm, I., et al. (1983).** "Identification of a male determinant on the X chromosome of housefly (*Musca domestica*) populations in South-East." *England. Genet. Res.* 42: 311-322.

- Deshpande, G., et al. (1995).** "scute (sis-b) function in Drosophila sex determination." *Mol Cell Biol* 15(8): 4430-40.
- Ding, Z., et al. (2003).** "Human MI-ER1 alpha and beta function as transcriptional repressors by recruitment of histone deacetylase 1 to their conserved ELM2 domain." *Mol Cell Biol* 23(1): 250-8.
- Dingwall, C. und R. A. Laskey (1986).** "Protein import into the cell nucleus." *Annu Rev Cell Biol* 2: 367-90.
- Dingwall, C. und R. A. Laskey (1991).** "Nuclear targeting sequences--a consensus?" *Trends Biochem Sci* 16(12): 478-81.
- Dobzhansky, T. (1935).** "Drosophila Miranda, a New Species." *Genetics* 20(4): 377-91.
- Dower, W. J., et al. (1988).** "High efficiency transformation of E. coli by high voltage electroporation." *Nucleic Acids Res* 16(13): 6127-45.
- Downes, J. A. (1969).** "The swarming and mating flight of Diptera." *Annual Review of Entomology* 14: 271-298.
- Drescher, W. und W. C. Rothebbuhler (1964).** "sex detremination in the honey bee." *J. Heredity* 5: 91-96.
- Dreyfuss, G., et al. (1988).** "Heterogeneous nuclear ribonucleoprotein particles and the pathway of mRNA formation." *Trends Biochem Sci* 13(3): 86-91.
- Dübendorfer, A., et al. (2002).** "Musca domestica, a window on the evolution of sex-determining mechanisms in insects." *Int J Dev Biol* 46(1): 75-9.
- Dübendorfer, A., et al. (1992).** "Sex determination mechanisms in dipteran insects: the case of Musca domestica." *Sem. Dev. Biol.* 3: 349-356.
- Duffy, J. B. und J. P. Gergen (1991).** "The Drosophila segmentation gene runt acts as a position-specific numerator element necessary for the uniform expression of the sex-determining gene Sex-lethal." *Genes Dev* 5(12A): 2176-87.
- Dura, J. M., et al. (1987).** "A complex genetic locus, polyhomeotic, is required for segmental specification and epidermal development in D. melanogaster." *Cell* 51(5): 829-39.
- Ellis, N. und P. N. Goodfellow (1989).** "The mammalian pseudoautosomal region." *Trends Genet* 5(12): 406-10.
- Epplen, J. T., et al. (1982).** "Base sequence of a cloned snake W-chromosome DNA fragment and identification of a male-specific putative mRNA in the mouse." *Proc Natl Acad Sci U S A* 79(12): 3798-802.
- Erickson, J. W. und T. W. Cline (1991).** "Molecular nature of the Drosophila sex determination signal and its link to neurogenesis." *Science* 251(4997): 1071-4.
- Erickson, J. W. und T. W. Cline (1993).** "A bZIP protein, sisterless-a, collaborates with bHLH transcription factors early in Drosophila development to determine sex." *Genes Dev* 7(9): 1688-702.
- Estes, P. A., et al. (1995).** "Multiple response elements in the Sex-lethal early promoter ensure its female-specific expression pattern." *Mol Cell Biol* 15(2): 904-17.

- Falcon-Perez, J. M., et al. (2007).** "The *Drosophila* pigmentation gene pink (p) encodes a homologue of human Hermansky-Pudlak syndrome 5 (HPS5)." *Traffic* 8(2): 154-68.
- Fauvarque, M. O., et al. (1995).** "Regulation of polyhomeotic transcription may involve local changes in chromatin activity in *Drosophila*." *Mech Dev* 52(2-3): 343-55.
- Feinberg, A. P. und B. Vogelstein (1983).** "A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity." *Anal Biochem* 132(1): 6-13.
- Felsenstein, J. (1974).** "The evolutionary advantage of recombination." *Genetics* 78(2): 737-56.
- Finley, D., et al. (1998).** "Unified nomenclature for subunits of the *Saccharomyces cerevisiae* proteasome regulatory particle." *Trends Biochem Sci* 23(7): 244-5.
- Fischer, J. und H. Tichy (1980).** "Über eine Heterochromatin-Mutation aus einer Wildpopulation von *Chironomus Nuditarsis*
I. Zur Funktion des veränderten Genom-Abschnitts." *Genetica* 54: 41-43.
- FitzGerald, P. C., et al. (2006).** "Comparative genomics of *Drosophila* and human core promoters." *Genome Biol* 7(7): R53.
- Fleige, S. und M. W. Pfaffl (2006).** "RNA integrity and the effect on the real-time qRT-PCR performance." *Mol Aspects Med* 27(2-3): 126-39.
- Flickinger, T. W. und H. K. Salz (1994).** "The *Drosophila* sex determination gene *snf* encodes a nuclear protein with sequence and functional similarity to the mammalian U1A snRNP protein." *Genes Dev* 8(8): 914-25.
- Fortes, P., et al. (1999).** "Luc7p, a novel yeast U1 snRNP protein with a role in 5' splice site recognition." *Genes Dev* 13(18): 2425-38.
- Foster, J. W., et al. (1992).** "Evolution of sex determination and the Y chromosome: SRY-related sequences in marsupials." *Nature* 359(6395): 531-3.
- Foster, J. W. und J. A. Graves (1994).** "An SRY-related sequence on the marsupial X chromosome: implications for the evolution of the mammalian testis-determining gene." *Proc Natl Acad Sci U S A* 91(5): 1927-31.
- Fraaije, M. W., et al. (1998).** "A novel oxidoreductase family sharing a conserved FAD-binding domain." *Trends Biochem Sci* 23(6): 206-7.
- Francis, R. C. und G. W. Barlow (1993).** "Social Control of Primary Sex Differentiation in the Midas Cichlid." *National Academy of Sciences* 90: 10673-10675.
- Franco, M. G., et al. (1982).** "Sex-determinants and their distribution in various populations of *Musca domestica* L. of Western Europe." *Genet Res* 40(3): 279-93.
- Franke, A., et al. (1992).** "Polycomb and polyhomeotic are constituents of a multimeric protein complex in chromatin of *Drosophila melanogaster*." *EMBO J* 11(8): 2941-50.
- Freije, D., et al. (1992).** "Identification of a second pseudoautosomal region near the Xq and Yq telomeres." *Science* 258(5089): 1784-7.
- Fu, X. D. (1993).** "Specific commitment of different pre-mRNAs to splicing by single SR proteins." *Nature* 365(6441): 82-5.

- Fujii, T. und T. Shimada (2007).** "Sex determination in the silkworm, *Bombyx mori*: a female determinant on the W chromosome and the sex-determining gene cascade." *Semin Cell Dev Biol* 18(3): 379-88.
- Funke, E. (2006).** "Molekulare Analyse der geschlechtsbestimmenden Region von *Chironomus thummi*." Diplomarbeit, Johannes-Gutenberg-Universität, Mainz.
- Funke, E. (2011).** "Charakterisierung und Analyse der geschlechtsbestimmenden Region von *Chironomus riparius*." Dissertation, Johannes-Gutenberg-Universität, Mainz.
- Galetti, P. M., Jr., et al. (1981).** "Heteromorphic sex chromosomes in three species of the genus *Leporinus* (Pisces, Anostomidae)." *Cytogenet Cell Genet* 29(3): 138-42.
- Garcia-Bustos, J., et al. (1991).** "Nuclear protein localization." *Biochim Biophys Acta* 1071(1): 83-101.
- Garrett-Engle, C. M., et al. (2002).** "intersex, a gene required for female sexual development in *Drosophila*, is expressed in both sexes and functions together with doublesex to regulate terminal differentiation." *Development* 129(20): 4661-75.
- Gautam, R., et al. (2004).** "The Hermansky-Pudlak syndrome 3 (cocoa) protein is a component of the biogenesis of lysosome-related organelles complex-2 (BLOC-2)." *J Biol Chem* 279(13): 12935-42.
- Gautam, R., et al. (2006).** "Interaction of Hermansky-Pudlak Syndrome genes in the regulation of lysosome-related organelles." *Traffic* 7(7): 779-92.
- Gergen, J. P. (1987).** "Dosage Compensation in *Drosophila*: Evidence That daughterless and Sex-lethal Control X Chromosome Activity at the Blastoderm Stage of Embryogenesis." *Genetics* 117(3): 477-85.
- Ghezzi, D., et al. (2009).** "SDHAF1, encoding a LYR complex-II specific assembly factor, is mutated in SDH-defective infantile leukoencephalopathy." *Nat Genet* 41(6): 654-6.
- Glickman, M. H. (2000).** "Getting in and out of the proteasome." *Semin Cell Dev Biol* 11(3): 149-58.
- Glickman, M. H. und A. Ciechanover (2002).** "The ubiquitin-proteasome proteolytic pathway: destruction for the sake of construction." *Physiol Rev* 82(2): 373-428.
- Gomez-Marquez, J. und F. Segade (1988).** "Prothymosin alpha is a nuclear protein." *FEBS Lett* 226(2): 217-9.
- Goodfellow, P. N. und R. Lovell-Badge (1993).** "SRY and sex determination in mammals." *Annu Rev Genet* 27: 71-92.
- Granadino, B., et al. (1990).** "The *Drosophila melanogaster* fl(2)d gene is needed for the female-specific splicing of Sex-lethal RNA." *EMBO J* 9(8): 2597-602.
- Granadino, B., et al. (1996).** "The gene fl(2)d is needed for the sex-specific splicing of transformer pre-mRNA but not for double-sex pre-mRNA in *Drosophila melanogaster*." *Mol Gen Genet* 253(1-2): 26-31.
- Grant, S. G., et al. (1990).** "Differential plasmid rescue from transgenic mouse DNAs into *Escherichia coli* methylation-restriction mutants." *Proc Natl Acad Sci U S A* 87(12): 4645-9.

- Graveley, B. R. (2000).** "Sorting out the complexity of SR protein functions." *RNA* 6(9): 1197-211.
- Graveley, B. R., et al. (2001).** "The role of U2AF35 and U2AF65 in enhancer-dependent splicing." *RNA* 7(6): 806-18.
- Grosschedl, R., et al. (1994).** "HMG domain proteins: architectural elements in the assembly of nucleoprotein structures." *Trends Genet* 10(3): 94-100.
- Grunstein, M. und D. S. Hogness (1975).** "Colony hybridization: a method for the isolation of cloned DNAs that contain a specific gene." *Proc Natl Acad Sci U S A* 72(10): 3961-5.
- Gubbay, J., et al. (1992).** "Inverted repeat structure of the Sry locus in mice." *Proc Natl Acad Sci U S A* 89(17): 7953-7.
- Haaq, C. M., et al. (1993).** "SRY Recognizes Conserved DNA Sites in Sex-Specific Promoters." *National Academy of Sciences* 90: 1097-1101.
- Haenlin, M., et al. (1987).** "Oocyte-specific transcription of fs(1)K10: a Drosophila gene affecting dorsal-ventral developmental polarity." *EMBO J* 6(3): 801-7.
- Hägele, K. (1985).** "Identification of a polytene chromosome band containing a male sex determiner of *Chironomus thummi thummi*." *Chromosoma* 91(3-4): 167-71.
- Hägele, K. (1986).** "Localization of a male sex determining chromosome region in *Chironomus thummi piger*." *Genetica* 70: 187-190.
- Hankeln, T., et al. (1994).** "Transposition of minisatellite-like DNA in *Chironomus* midges." *Genome* 37(4): 542-9.
- Hankeln, T. und E. R. Schmidt (1987).** "Cotransposition of a highly repetitive DNA element with flanking sequences in the genome of the midge *Chironomus thummi*." *J Mol Evol* 26(4): 311-9.
- Hankeln, T. und E. R. Schmidt (1990).** "New foldback transposable element TFB1 found in histone genes of the midge *Chironomus thummi*." *J Mol Biol* 215(4): 477-82.
- Hankeln, T. und E. R. Schmidt (1991).** "The organization, localization and nucleotide sequence of the histone genes of the midge *Chironomus thummi*." *Chromosoma* 101(1): 25-31.
- Harley, V. R., et al. (2003).** "The molecular action and regulation of the testis-determining factors, SRY (sex-determining region on the Y chromosome) and SOX9 [SRY-related high-mobility group (HMG) box 9]." *Endocr Rev* 24(4): 466-87.
- Harley, V. R., et al. (1992).** "DNA binding activity of recombinant SRY from normal males and XY females." *Science* 255(5043): 453-6.
- Hastings, M. L. und A. R. Krainer (2001).** "Pre-mRNA splicing in the new millennium." *Curr Opin Cell Biol* 13(3): 302-9.
- He, H., et al. (1995).** "Polymorphic SINEs in chironomids with DNA derived from the R2 insertion site." *J Mol Biol* 245(1): 34-42.
- Hediger, M., et al. (2004).** "Sex determination in *Drosophila melanogaster* and *Musca domestica* converges at the level of the terminal regulator doublesex." *Dev Genes Evol* 214(1): 29-42.

- Hediger, M., et al. (2010).** "Molecular characterization of the key switch F provides a basis for understanding the rapid divergence of the sex-determining pathway in the housefly." *Genetics* 184(1): 155-70.
- Hediger, M., et al. (1998).** "Distribution of heterochromatin on the mitotic chromosomes of *Musca domestica* L. in relation to the activity of male-determining factors." *Chromosoma* 107(4): 267-71.
- Hedley, M. L. und T. Maniatis (1991).** "Sex-specific splicing and polyadenylation of *dsx* pre-mRNA requires a sequence that binds specifically to *tra-2* protein in vitro." *Cell* 65(4): 579-86.
- Helip-Wooley, A., et al. (2005).** "Association of the Hermansky-Pudlak syndrome type-3 protein with clathrin." *BMC Cell Biol* 6: 33.
- Herold, T. (2008).** "Molekulare Charakterisierung der geschlechtsdifferenzierenden Chromosomenregion bei Chironomiden." Diplomarbeit, Johannes-Gutenberg-Universität, Mainz.
- Hertel, K. J. und B. R. Graveley (2005).** "RS domains contact the pre-mRNA throughout spliceosome assembly." *Trends Biochem Sci* 30(3): 115-8.
- Hilfiker-Kleiner, D., et al. (1993).** "Developmental analysis of two sex-determining genes, M and F, in the housefly, *Musca domestica*." *Genetics* 134(4): 1187-94.
- Hilfiker, A., et al. (1995).** "The gene *virilizer* is required for female-specific splicing controlled by *Sxl*, the master gene for sexual development in *Drosophila*." *Development* 121(12): 4017-26.
- Hilfiker, A. und R. Nöthiger (1991).** "The temperature-sensitive mutation *vir* (*virilizer*) identifies a new gene involved in sex determination of *Drosophila*." *Roux's Arch. Dev. Biol.* 200: 240-248.
- Hiroyoshi, T. (1964).** "Sex-Limited Inheritance and Abnormal Sex Ratio in Strains of the Housefly." *Genetics* 50: 373-85.
- Hodgkin, J. (1989).** "Drosophila Sex Determination: A Cascade of Regulated Splicing." *Cell* 56: 905-906.
- Horabin, J. I. und P. Schedl (1993).** "Regulated splicing of the *Drosophila* sex-lethal male exon involves a blockage mechanism." *Mol Cell Biol* 13(3): 1408-14.
- Hoshijima, K., et al. (1991).** "Control of doublesex alternative splicing by transformer and transformer-2 in *Drosophila*." *Science* 252(5007): 833-6.
- Hoshijima, K., et al. (1995).** "Transcriptional regulation of the Sex-lethal gene by helix-loop-helix proteins." *Nucleic Acids Res* 23(17): 3441-8.
- Huizing, M., et al. (2000).** "Hermansky-Pudlak syndrome and related disorders of organelle formation." *Traffic* 1(11): 823-35.
- Huizing, M., et al. (2002).** "Hermansky-Pudlak syndrome: vesicle formation from yeast to man." *Pigment Cell Res* 15(6): 405-19.
- Huntley, M. A. und G. B. Golding (2004).** "Neurological proteins are not enriched for repetitive sequences." *Genetics* 166(3): 1141-54.
- Inoue, K., et al. (1992).** "Binding of the *Drosophila* transformer and transformer-2 proteins to the regulatory elements of doublesex primary transcript for sex-specific RNA processing." *Proc Natl Acad Sci U S A* 89(17): 8092-6.

- Inoue, K., et al. (1990).** "Binding of the *Drosophila* sex-lethal gene product to the alternative splice site of transformer primary transcript." *Nature* 344(6265): 461-3.
- Isono, E., et al. (2007).** "The assembly pathway of the 19S regulatory particle of the yeast 26S proteasome." *Mol Biol Cell* 18(2): 569-80.
- Israelewski, N. und E. R. Schmidt (1982).** "Spacer size heterogeneity in ribosomal DNA of *Chironomus thummi* is due to a 120 bp repeat homologous to a predominantly centromeric repeated sequence." *Nucleic Acids Res* 10(23): 7689-700.
- Jantzen, H. M., et al. (1990).** "Nucleolar transcription factor hUBF contains a DNA-binding motif with homology to HMG proteins." *Nature* 344(6269): 830-6.
- Janzen, F. J. und G. L. Paukstis (1991).** "Environmental sex determination in reptiles: ecology, evolution, and experimental design." *Q Rev Biol* 66(2): 149-79.
- Jones, K. A., et al. (1987).** "A cellular DNA-binding protein that activates eukaryotic transcription and DNA replication." *Cell* 48(1): 79-89.
- Jurnsich, V. A. und K. C. Burtis (1993).** "A positive role in differentiation for the male doublesex protein of *Drosophila*." *Dev Biol* 155(1): 235-49.
- Juven-Gershon, T., et al. (2008).** "The RNA polymerase II core promoter - the gateway to transcription." *Curr Opin Cell Biol* 20(3): 253-9.
- Kay, A. B., et al. (1978).** "Widespread IgE-mediated hypersensitivity in Northern Sudan to the chironomid *Cladotanytarsus lewisi* ('green nimitti')." *Clin Exp Immunol* 34: 106-110.
- Kelly, W. G., et al. (2002).** "X-chromosome silencing in the germline of *C. elegans*." *Development* 129(2): 479-92.
- Keyes, L. N., et al. (1992).** "The primary sex determination signal of *Drosophila* acts at the level of transcription." *Cell* 68(5): 933-43.
- Keyl, H.-G. und K. Strenzke (1956).** "Taxonomie und Cytologie von zwei Subspezies der Art *Chironomus thummi*." *Zeitschrift für Naturforschung* 11b: 12.
- Keyl, H. G. (1961).** "[Chromosome evolution in *Chironomus*. I. Structural variants of salivary gland chromosomes]." *Chromosoma* 12: 26-47.
- Keyl, H. G. (1962).** "Chromosomenevolution bei *Chironomus* 2." *Chromosoma* 13: 464-513.
- Keyl, H. G. (1965).** "[Duplications of subunits of chromosome DNA during the evolution of *Chironomus thummi*]." *Chromosoma* 17(2): 139-80.
- Kibbe, W. A. (2007).** "OligoCalc: an online oligonucleotide properties calculator." *Nucleic Acids Res* 35(Web Server issue): W43-6.
- Kieffer, J.-J. (1911).** "Nouvelles descriptions de Chironomides obtenus d'éclosion." *Bull. Soc. Hist. nat. Metz* 27: 1-60.
- Kim, C. A., et al. (2005).** "Structural organization of a Sex-comb-on-midleg/polyhomeotic copolymer." *J Biol Chem* 280(30): 27769-75.
- Kim, U. J., et al. (1996).** "Construction and characterization of a human bacterial artificial chromosome library." *Genomics* 34(2): 213-8.

- Knies, D. (2006).** "Sequenzierung, in situ Lokalisation und genomische Analyse eines ca. 55656 Bp langen Chromosomenabschnitts aus der geschlechtsbestimmenden Region von *C.thummi*. ." Diplomarbeit, Johannes-Gutenberg-Universität, Mainz.
- Koerich, L. B., et al. (2008).** "Low conservation of gene content in the Drosophila Y chromosome." *Nature* 456(7224): 949-51.
- Kohtz, J. D., et al. (1994).** "Protein-protein interactions and 5'-splice-site recognition in mammalian mRNA precursors." *Nature* 368(6467): 119-24.
- Kon, M. (1987).** "The mating system of chironomid midges (Diptera: Chironomidae): a review." *Mem Fac Sci Kyoto Univ Ser B*: 129-134.
- Koopmann, P., et al. (1991).** "Male development of chromosomally female mice transgenic for Sry." *Nature* 342: 940-942.
- Kraemer, C. (1994).** "Molekulare Analyse der geschlechtsbestimmenden Region von Chironomiden." Dissertation, Johannes-Gutenberg-Universität Mainz.
- Kraemer, C. (1994).** "Molekulare Analyse der geschlechtsbestimmenden Region von Chironomiden." Dissertation, Johannes-Gutenberg-Universität, Mainz.
- Kraemer, C. und E. R. Schmidt (1993).** "The sex determining region of *Chironomus thummi* is associated with highly repetitive DNA and transposable elements." *Chromosoma* 102(8): 553-62.
- Kramer, S. G., et al. (1999).** "Direct activation of Sex-lethal transcription by the Drosophila runt protein." *Development* 126(1): 191-200.
- Kunkel, L. M. und K. D. Smith (1982).** "Evolution of human Y-chromosome DNA." *Chromosoma* 86(2): 209-28.
- Kunz, S. (2001).** "Struktur der geschlechtsbestimmenden Region von *Chironomus thummi* (Diptera, Insecta) und Vergleich der Gene *CtY* und *CpY*." Dissertation, Johannes-Gutenberg-Universität Mainz.
- Kuroda, M. I., et al. (1991).** "The maleless protein associates with the X chromosome to regulate dosage compensation in Drosophila." *Cell* 66(5): 935-47.
- Kutach, A. K. und J. T. Kadonaga (2000).** "The downstream promoter element DPE appears to be as widely used as the TATA box in Drosophila core promoters." *Mol Cell Biol* 20(13): 4754-64.
- Lagrange, T., et al. (1998).** "New core promoter element in RNA polymerase II-dependent transcription: sequence-specific DNA binding by transcription factor IIB." *Genes Dev* 12(1): 34-44.
- Lahiri, D. K., et al. (2003).** "Role of cytokines in the gene expression of amyloid beta-protein precursor: identification of a 5'-UTR-binding nuclear factor and its implications in Alzheimer's disease." *J Alzheimers Dis* 5(2): 81-90.
- Lam, B. J., et al. (2003).** "Enhancer-dependent 5'-splice site control of fruitless pre-mRNA splicing." *J Biol Chem* 278(25): 22740-7.
- Lam, B. J. und K. J. Hertel (2002).** "A general role for splicing enhancers in exon definition." *RNA* 8(10): 1233-41.
- Landsman, D. und M. Bustin (1993).** "A signature for the HMG-1 box DNA-binding proteins." *Bioessays* 15(8): 539-46.

- Langer-Safer, P. R., et al. (1982).** "Immunological method for mapping genes on *Drosophila* polytene chromosomes." *Proc Natl Acad Sci U S A* 79(14): 4381-5.
- Langley, K. E., et al. (1975).** "Molecular basis of beta-galactosidase alpha-complementation." *Proc Natl Acad Sci U S A* 72(4): 1254-7.
- Lee, L. G., et al. (1992).** "DNA sequencing with dye-labeled terminators and T7 DNA polymerase: effect of dyes and dNTPs on incorporation of dye-terminators and probability analysis of termination fragments." *Nucleic Acids Res* 20(10): 2471-83.
- Lefevre, G., Jr. (1974).** "The relationship between genes and polytene chromosome bands." *Annu Rev Genet* 8: 51-62.
- Levinson, G., et al. (1985).** "Cross-hybridizing snake satellite, *Drosophila*, and mouse DNA sequences may have arisen independently." *Mol Biol Evol* 2(6): 494-504.
- Lodi, T., et al. (1994).** "Carbon catabolite repression in *Kluyveromyces lactis*: isolation and characterization of the KIDLD gene encoding the mitochondrial enzyme D-lactate ferricytochrome c oxidoreductase." *Mol Gen Genet* 244(6): 622-9.
- Long, J. C. und J. F. Caceres (2009).** "The SR protein family of splicing factors: master regulators of gene expression." *Biochem J* 417(1): 15-27.
- Lucchesi, J. C. (1978).** "Gene dosage compensation and the evolution of sex chromosomes." *Science* 202(4369): 711-6.
- Lucchesi, J. C. (1994).** "The evolution of heteromorphic sex chromosomes." *Bioessays* 16(2): 81-3.
- Lynch, K. W. und T. Maniatis (1996).** "Assembly of specific SR protein complexes on distinct regulatory elements of the *Drosophila* doublesex splicing enhancer." *Genes Dev* 10(16): 2089-101.
- Lyon, M. F. (1961).** "Gene action in the X-chromosome of the mouse (*Mus musculus* L.)." *Nature* 190: 372-3.
- Macgregor, H. C. und S. Mizuno (1976).** "In situ hybridization of "nick-translated" 3H-ribosomal DNA to chromosomes from salamanders." *Chromosoma* 54(1): 15-25.
- Mack, T. S. (2009).** "Genexpressionsanalyse der geschlechtsdeterminierenden Region von *Chironomus thummi*." Diplomarbeit, Johannes-Gutenberg-Universität, Mainz.
- Macknight, R. H. (1939).** "The Sex-Determining Mechanism of *Drosophila* *Miranda*." *Genetics* 24(2): 180-201.
- Mainx, F. (1959).** "[The sex conditions of the Phoridae *Megaselia scalaris* and the problem of an alternating sex determination]." *Z Vererbungsl* 90: 251-6.
- Mainx, F. (1964).** "The genetics of *Megaselia scalaris* Loew (Phoridae): a new type of sex determination in Diptera." *Am. Nat.* 98: 415-430.
- Mainx, F. (1966).** "[Sex determination in *Megaselia scalaris* Loew (Phoridae)]." *Z Vererbungsl* 98(1): 49-60.
- Martin, J. (1980).** "Multiple sex determining regions in a group of related *Chironomus* species (Diptera:Chironomidae)." *heredity* 44.
- Martin, J. (2010).** "A phylogenetic analysis of the change in genomic location of the sex determining region in the genus *Chironomus* (Diptera: Chironomidae)." *VOGiS Herald* 14: 62-69.

- Mattaj, I. W. (1989).** "A binding consensus: RNA-protein interactions in splicing, snRNPs, and sex." *Cell* 57(1): 1-3.
- Maynard Smith, J. (1978).** "The evolution of sex." Cambridge University Press, Cambridge.
- McAllister, B. F. und G. A. McVean (2000).** "Neutral evolution of the sex-determining gene transformer in *Drosophila*." *Genetics* 154(4): 1711-20.
- McDonald, I., et al. (1978).** "housefly genetics: isolation of a female determining factor on chromosome 4." *H. Ann. Entomol. Soc. Am.* 71: 692-694.
- Meigen, J. W. (1804).** "Klassifikation und Beschreibung der europäischen zweiflügligen Insekten (Diptera, Linn.)." Reichard, Braunschweig.
- Meise, M., et al. (1998).** "Sex-lethal, the master sex-determining gene in *Drosophila*, is not sex-specifically regulated in *Musca domestica*." *Development* 125(8): 1487-94.
- Metz, C. W. und M. L. Schmuck (1929).** "Unisexual Progenies and the Sex Chromosome Mechanism in *Sciara*." *Proc Natl Acad Sci U S A* 15(12): 863-6.
- Michelson, A. M. (1958).** "Hyperchromicity and nucleic acids." *Nature* 182(4648): 1502-3.
- Miklos, G. L., et al. (1989).** "Occurrence of the (GATA)_n sequences in vertebrate and invertebrate genomes." *Chromosoma* 98(3): 194-200.
- Mori, H., et al. (1986).** "Structure and function of the F plasmid genes essential for partitioning." *J Mol Biol* 192(1): 1-15.
- Muller, H. J. (1918).** "Genetic Variability, Twin Hybrids and Constant Hybrids, in a Case of Balanced Lethal Factors." *Genetics* 3(5): 422-99.
- Muller, H. J. (1964).** "The Relation of Recombination to Mutational Advance." *Mutat Res* 106: 2-9.
- Müller, U., et al. (1986).** "Deletion mapping of the testis determining locus with DNA probes in 46,XX males and in 46,XY and 46,X,dic(Y) females." *Nucleic Acids Res* 14(16): 6489-505.
- Mullis, K., et al. (1992*).** "Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. 1986." *Biotechnology* 24: 17-27.
- Munk, K. (2010).** "Genetik." Thieme Verlag.
- Münsterberg, A. und R. Lovell-Badge (1991).** "Expression of the mouse anti-müllerian hormone gene suggests a role in both male and female sexual differentiation." *Development* 113(2): 613-24.
- Nagoshi, R. N., et al. (1988).** "The control of alternative splicing at genes regulating sexual differentiation in *D. melanogaster*." *Cell* 53(2): 229-36.
- Narbonne, K., et al. (2004).** "polyhomeotic is required for somatic cell proliferation and differentiation during ovarian follicle formation in *Drosophila*." *Development* 131(6): 1389-400.
- Neer, E. J., et al. (1994).** "The ancient regulatory-protein family of WD-repeat proteins." *Nature* 371(6495): 297-300.

- Neuman-Silberberg, F. S. und T. Schüpbach (1993).** "The *Drosophila* dorsoventral patterning gene *gurken* produces a dorsally localized RNA and encodes a TGF alpha-like protein." *Cell* 75(1): 165-74.
- Nigro, R., et al. (2007).** "Temperature and the progeny sexratio in *Sciara ocellaris* (Diptera, Sciaridae)." *Genet. Mol. Biol.* 30: 152-158.
- Nöthiger, R. und M. Steinmann-Zwicky (1985).** "A single principle for sex determination in insects." *Cold Spring Harb Symp Quant Biol* 50: 615-21.
- Novoseler, M., et al. (2005).** "Functional domains of the yeast chromatin protein Sin1p/Spt2p can bind four-way junction and crossing DNA structures." *J Biol Chem* 280(7): 5169-77.
- O'Neil, M. T. und J. M. Belote (1992).** "Interspecific comparison of the transformer gene of *Drosophila* reveals an unusually high degree of evolutionary divergence." *Genetics* 131(1): 113-28.
- Oh, H. J. und Y. F. Lau (2006).** "KRAB: a partner for SRY action on chromatin." *Mol Cell Endocrinol* 247(1-2): 47-52.
- Ohler, U., et al. (2002).** "Computational analysis of core promoters in the *Drosophila* genome." *Genome Biol* 3(12): RESEARCH0087.
- Ohno, S. (1967).** "Sex chromosomes and sex linked genes." Springer, Berlin.
- Okazaki, N., et al. (1991).** "The *ste4+* gene, essential for sexual differentiation of *Schizosaccharomyces pombe*, encodes a protein with a leucine zipper motif." *Nucleic Acids Res* 19(25): 7043-7.
- Oliver, D. (1971).** "Life history of Chironomidae." *Annu. Rev. Entomol.* 16: 211-230.
- Owczarzy, R., et al. (2008).** "IDT SciTools: a suite for analysis and design of nucleic acid oligomers." *Nucleic Acids Res* 36(Web Server issue): W163-9.
- Page, D. C., et al. (1987).** "The sex-determining region of the human Y chromosome encodes a finger protein." *Cell* 51(6): 1091-104.
- Palmer, M. S., et al. (1989).** "Genetic evidence that ZFY is not the testis-determining factor." *Nature* 342(6252): 937-9.
- Pane, A., et al. (2005).** "Evolutionary conservation of *Ceratitis capitata* transformer gene function." *Genetics* 171(2): 615-24.
- Pane, A., et al. (2002).** "The transformer gene in *Ceratitis capitata* provides a genetic basis for selecting and remembering the sexual fate." *Development* 129(15): 3715-25.
- Parkhurst, S. M. und D. Ish-Horowicz (1992).** "Common denominators for sex." *Curr Biol* 2(12): 629-31.
- Parkhurst, S. M., et al. (1993).** "achaete-scute feminizing activities and *Drosophila* sex determination." *Development* 117(2): 737-49.
- Paroush, Z., et al. (1994).** "Groucho is required for *Drosophila* neurogenesis, segmentation, and sex determination and interacts directly with hairy-related bHLH proteins." *Cell* 79(5): 805-15.
- Paterno, G. D., et al. (1997).** "cDNA cloning of a novel, developmentally regulated immediate early gene activated by fibroblast growth factor and encoding a nuclear protein." *J Biol Chem* 272(41): 25591-5.

- Paterno, G. D., et al. (1998).** "Molecular cloning of human er1 cDNA and its differential expression in breast tumours and tumour-derived cell lines." *Gene* 222(1): 77-82.
- Pathak, S. und A. D. Stock (1974).** "The X chromosomes of mammals: karyological homology as revealed by banding techniques." *Genetics* 78(2): 703-14.
- Perje, A. M. (1948).** "Studies on the spermatogenesis in *Musca domestica*." *Hereditas* 34: 209-232.
- Permpoon, R. und S. Thanaphum (2010).** "Isolation and characterization of oligomerization domain I and II coding regions of doublesex genes in agricultural fruit flies (Diptera: Tephritidae)." *Eur. J. Entomol.* 107: 121–126.
- Pinder (1986).** "Biology of Freshwater Chironomidae." *Annual Reviews of Entomology* 31: 1-23.
- Ponting, C. P. (1995).** "SAM: a novel motif in yeast sterile and *Drosophila* polyhomeotic proteins." *Protein Sci* 4(9): 1928-30.
- Porath, J. und P. Flodin (1959).** "Gel filtration: a method for desalting and group separation." *Nature* 183(4676): 1657-9.
- Potter, S., et al. (1980).** "Eucaryotic transposable genetic elements with inverted terminal repeats." *Cell* 20(3): 639-47.
- Potter, S. S. (1982).** "DNA sequence of a foldback transposable element in *Drosophila*." *Nature* 297(5863): 201-4.
- Prober, J. M., et al. (1987*).** "A system for rapid DNA sequencing with fluorescent chain-terminating dideoxynucleotides." *Science* 238(4825): 336-41.
- Prost, E., et al. (1988).** "Role of the oocyte nucleus in determination of the dorsoventral polarity of *Drosophila* as revealed by molecular analysis of the K10 gene." *Genes Dev* 2(7): 891-900.
- Puig, O., et al. (2007).** "The U1 snRNP-associated factor Luc7p affects 5' splice site selection in yeast and human." *Nucleic Acids Res* 35(17): 5874-85.
- Pultz, M. A. und B. S. Baker (1995).** "The dual role of hermaphrodite in the *Drosophila* sex determination regulatory hierarchy." *Development* 121(1): 99-111.
- Quevillon, E., et al. (2005).** "InterProScan: protein domains identifier." *Nucleic Acids Res* 33(Web Server issue): W116-20.
- Retelska, D., et al. (2006).** "Similarities and differences of polyadenylation signals in human and fly." *BMC Genomics* 7: 176.
- Rhodes, N., et al. (1990).** "STE11 is a protein kinase required for cell-type-specific transcription and signal transduction in yeast." *Genes Dev* 4(11): 1862-74.
- Rosen, M., et al. (2002).** "Telomere terminating with centromere-specific repeats is closely associated with a transposon derived gene in *Chironomus pallidivittatus*." *Chromosoma* 110(8): 532-41.
- Rosin, S. und J. Fischer (1972).** "[Polymorphism of male realization in *Chironomus*]." *Rev Suisse Zool: Suppl.* 119-41.
- Ross, K. G. und D. J. C. Fletcher (1985).** "Genetic Origin of Male Diploidy in the Fire Ant, *Solenopsis invicta* (Hymenoptera: Formicidae), and its Evolutionary Significance." *Evolution* 39: 888-903.

- Roß, R. (1993).** "Struktur der Histongene in Korrelation zum Heterochromatingehalt verschiedener Chironomiden-Spezies." Dissertation, Johannes-Gutenberg-Universität, Mainz.
- Ross, R. (1993).** "Struktur der Histongene in Korrelation zum Heterochromatingehalt verschiedener Chironomiden-Spezies." Dissertation, Johannes-Gutenberg-Universität Mainz.
- Rozynek, P., et al. (1989).** "Structure of a hemoglobin gene cluster and nucleotide sequence of three hemoglobin genes from the midge *Chironomus thummi piger* (Diptera, Insecta)." *Biol Chem Hoppe Seyler* 370(6): 533-42.
- Rubini, P. G., et al. (1972).** "Polymorphisms for heterochromosomes and autosomal sex-determinants in *Musca domestica*." *L. Atti del IX Congresso Nazionale Italiano di Entomologia*: 341-352.
- Rubini, P. G. und D. Palenzona (1967).** "Response to selection for high number of heterochromosomes in *Musca domestica*." *L. Gen. Agr.* 21: 101-110.
- Rutledge, B. J., et al. (1992).** "The *Drosophila spitz* gene encodes a putative EGF-like growth factor involved in dorsal-ventral axis formation and neurogenesis." *Genes Dev* 6(8): 1503-17.
- Rychlik, W. (1995).** "Selection of primers for polymerase chain reaction." *Mol Biotechnol* 3(2): 129-34.
- Saccone, G., et al. (2002).** "Sex determination in flies, fruitflies and butterflies." *Genetica* 116(1): 15-23.
- Saccone, G., et al. (2000).** "Sex determination in medfly: a molecular approach. Area-wide control of fruitflies and other pest insects." ed. K.-H. Tan: 491-496 Penang: Penerbit USM.
- Saccone, G., et al. (1998).** "The *Ceratitis capitata* homologue of the *Drosophila* sex-determining gene *sex-lethal* is structurally conserved, but not sex-specifically regulated." *Development* 125(8): 1495-500.
- Salvemini, M., et al. (2009).** "*Ceratitis capitata* transformer-2 gene is required to establish and maintain the autoregulation of *Cctra*, the master gene for female sex determination." *Int J Dev Biol* 53(1): 109-20.
- Samuels, M., et al. (1998).** "Activities of the *Sex-lethal* protein in RNA binding and protein:protein interactions." *Nucleic Acids Res* 26(11): 2625-37.
- Sánchez, L. (2008).** "Sex-determining mechanisms in insects." *Int. J. Dev. Biol.* 52: 837-856.
- Sánchez, L., et al. (1994).** "Sex determination in *Drosophila melanogaster*: X-linked genes involved in the initial step of *sex-lethal* activation." *Dev Genet* 15(3): 251-64.
- Sánchez, L. und R. Nöthiger (1983).** "Sex determination and dosage compensation in *Drosophila melanogaster*: production of male clones in XX females." *EMBO J* 2(4): 485-91.
- Sanger, F., et al. (1977).** "DNA sequencing with chain-terminating inhibitors." *Proc Natl Acad Sci U S A* 74(12): 5463-7.

- Schaefer, J. und E. R. Schmidt (1981).** "Different repetition frequencies of a 120 base-pair DNA-element and its arrangement in *Chironomus thummi thummi* and *Chironomus thummi piger*." *Chromosoma* 84(1): 61-6.
- Schlee, D. (1977).** "Florale und extraflorale Nektarien sowie Insektenkot als Nahrungsquelle für Chironomidae-Imagines (und andere Dipteren)." *Stuttgarter Beitr. Naturk. Ser. A* 300: 1-16.
- Schmidt, E. R. (1981).** "The development of A 120 basepair repetitive DNA sequence in *chironomus thummi* is correlated to the duplication of defined chromosomal segments." *FEBS Letters* 129: 21-24.
- Schmidt, E. R. (1984).** "Clustered and interspersed repetitive DNA sequence family of *Chironomus*. The nucleotide sequence of the Cla-elements and of various flanking sequences." *J Mol Biol* 178(1): 1-15.
- Schmidt, E. R. und E. A. Godwin (1983).** "The nucleotide sequence of an unusual non-transcribed spacer and its ancestor in the rDNA in *Chironomus thummi*." *EMBO J* 2(7): 1177-83.
- Schmidt, E. R., et al. (1982).** "Cloning and analysis of ribosomal DNA of *Chironomus thummi piger* and *Chironomus thummi thummi*. The nontranscribed spacer of Ch. th. thummi contains a highly repetitive DNA sequence." *Chromosoma* 87(4): 389-407.
- Schmidt, R., et al. (1997).** "The mutation masculinizer (man) defines a sex-determining gene with maternal and zygotic functions in *Musca domestica* L." *Genetics* 145(1): 173-83.
- Schroeder, A., et al. (2006).** "The RIN: an RNA integrity number for assigning integrity values to RNA measurements." *BMC Mol Biol* 7: 3.
- Schultz, J., et al. (1997).** "SAM as a protein interaction domain involved in developmental regulation." *Protein Sci* 6(1): 249-53.
- Schütt, C. und R. Nöthiger (2000).** "Structure, function and evolution of sex-determining systems in Dipteran insects." *Development* 127(4): 667-77.
- Sciabica, K. S. und K. J. Hertel (2006).** "The splicing regulators Tra and Tra2 are unusually potent activators of pre-mRNA splicing." *Nucleic Acids Res* 34(22): 6612-20.
- Sefton, L., et al. (2000).** "An extracellular activator of the *Drosophila* JAK/STAT pathway is a sex-determination signal element." *Nature* 405(6789): 970-3.
- Serano, T. L. und R. S. Cohen (1995).** "Gratuitous mRNA localization in the *Drosophila* oocyte." *Development* 121(9): 3013-21.
- Shen, H. und M. R. Green (2004).** "A pathway of sequential arginine-serine-rich domain-splicing signal interactions during mammalian spliceosome assembly." *Mol Cell* 16(3): 363-73.
- Shen, H., et al. (2004).** "Arginine-serine-rich domains bound at splicing enhancers contact the branchpoint to promote prespliceosome assembly." *Mol Cell* 13(3): 367-76.
- Shizuya, H., et al. (1992).** "Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in *Escherichia coli* using an F-factor-based vector." *Proc Natl Acad Sci U S A* 89(18): 8794-7.
- Sinclair, A. H., et al. (1990).** "A gene from the human sex-determining region encodes a protein with homology to a conserved DNA-binding motif." *Nature* 346(6281): 240-4.

- Singh, L., et al. (1980).** "Sex chromosome associated satellite DNA: evolution and conservation." *Chromosoma* 79(2): 137-57.
- Skaletsky, H., et al. (2003).** "The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes." *Nature* 423(6942): 825-37.
- Smith, T. F., et al. (1999).** "The WD repeat: a common architecture for diverse functions." *Trends Biochem Sci* 24(5): 181-5.
- Solari, F., et al. (1999).** "The *Caenorhabditis elegans* genes *egl-27* and *egr-1* are similar to MTA1, a member of a chromatin regulatory complex, and are redundantly required for embryonic patterning." *Development* 126(11): 2483-94.
- Sosnowski, B. A., et al. (1989).** "Sex-specific alternative splicing of RNA from the transformer gene results from sequence-dependent splice site blockage." *Cell* 58(3): 449-59.
- Southern, E. M. (1975).** "Detection of specific sequences among DNA fragments separated by gel electrophoresis." *J Mol Biol* 98(3): 503-17.
- Stanke, M. und B. Morgenstern (2005).** "AUGUSTUS: a web server for gene prediction in eukaryotes that allows user-defined constraints." *Nucleic Acids Res* 33(Web Server issue): W465-7.
- Steinemann, M. (1982).** "Multiple sex chromosomes in *Drosophila miranda*: a system to study the degeneration of a chromosome." *Chromosoma* 86(1): 59-76.
- Steinemann, M. und S. Steinemann (1990).** "Evolutionary changes in the organization of the major LCP gene cluster during sex chromosomal differentiation in the sibling species *Drosophila perMmills*, *D. pseudoobscura* and *D. miranda*." *Chromosoma* 99: 424-431.
- Steinemann, M. und S. Steinemann (1992).** "Degenerating Y chromosome of *Drosophila miranda*: a trap for retrotransposons." *Proc Natl Acad Sci U S A* 89(16): 7591-5.
- Steinemann, M. und S. Steinemann (1993).** "A duplication including the Y allele of *Lcp2* and the TRIM retrotransposon at the *Lcp* locus on the degenerating neo-Y chromosome of *Drosophila miranda*: molecular structure and mechanisms by which it may have arisen." *Genetics* 134(2): 497-505.
- Steinemann, M. und S. Steinemann (1998).** "Enigma of Y chromosome degeneration: neo-Y and neo-X chromosomes of *Drosophila miranda* a model for sex chromosome evolution." *Genetica* 102-103(1-6): 409-20.
- Steinemann, M., et al. (1993).** "How Y chromosomes become genetically inert." *Proc Natl Acad Sci U S A* 90(12): 5737-41.
- Stevens, L. (1997).** "Sex chromosomes and sex determining mechanisms in birds." *Sci Prog* 80 (Pt 3): 197-216.
- Stevens, N. M. (1908).** "A study of the germ cells of certain diptera with reference to the heterochromosomes and the phenomena of synapsis." *J. Exp. Zool.* 5: 359-374.
- Stiehl, R. (2009).** "Klonierung und Analyse der geschlechtsbestimmenden Region bei *Chironomus piger* und *Chironomus thummi*." Diplomarbeit, Johannes-Gutenberg-Universität, Mainz.

- Stros, M., et al. (2007).** "The HMG-box: a versatile protein domain occurring in a wide variety of DNA-binding proteins." *Cell Mol Life Sci* 64(19-20): 2590-606.
- Su, H. und Y. F. Lau (1993).** "Identification of the transcriptional unit, structural organization, and promoter sequence of the human sex-determining region Y (SRY) gene, using a reverse genetic approach." *Am J Hum Genet* 52(1): 24-38.
- Suzuki, M., et al. (1994).** "Solution structure of the DNA binding octapeptide repeat of the K10 gene product." *Protein Eng* 7(4): 461-70.
- Szymanowska, M., et al. (2004).** "Transcription factor binding to variable nucleotide sequences in 5'-flanking regions of bovine casein genes." *International Dairy Journal* 14(2): 103-115.
- Taenzler, V., et al. (2007).** "Chironomids: suitable test organisms for risk assessment investigations on the potential endocrine disrupting properties of pesticides." *Ecotoxicology* 16(1): 221-30.
- Thomas, J. O. (2001).** "HMG1 and 2: architectural DNA-binding proteins." *Biochem Soc Trans* 29(Pt 4): 395-401.
- Thomas, J. O. und A. A. Travers (2001).** "HMG1 and 2, and related 'architectural' DNA-binding proteins." *Trends Biochem Sci* 26(3): 167-74.
- Thorne, L. B., et al. (2005).** "Cloning and characterization of the mouse ortholog of *mir1*." *DNA Seq* 16(3): 237-40.
- Torres, M. und L. Sanchez (1989).** "The scute (T4) gene acts as a numerator element of the X:a signal that determines the state of activity of sex-lethal in *Drosophila*." *EMBO J* 8(10): 3079-86.
- Traut, W. und G. C. Mosbacher (1968).** "[Sex chromatin in Lepidoptera]." *Chromosoma* 25(3): 343-56.
- Truett, M. A., et al. (1981).** "Unusual structure of the FB family of transposable elements in *Drosophila*." *Cell* 24(3): 753-63.
- Tsuchiya, K., et al. (1995).** "Gonadoblastoma: molecular definition of the susceptibility region on the Y chromosome." *Am J Hum Genet* 57(6): 1400-7.
- Ullerich, F. H. (1963).** "Geschlechtschromosomen und Geschlechtsbestimmung bei einigen Calliphorinen (Calliphoridae, Diptera)." *Chromosoma* 14: 45-110.
- Ullerich, F. H. (1973).** "Die genetische Grundlage der Monogenie bei der Schmeigfliege *Chrysomya rufifacies* (Calliphoridae, Diptera)." *Molec. gen. Genet.* 125: 157-172.
- Ullerich, F. H. (1975).** "[Identification of the genetic sex chromosomes in the monogenic blowfly *Chrysomya rufifacies* (Calliphoridae, Diptera)]." *Chromosoma* 50(4): 393-419.
- Ullerich, F. H. (1976).** "Chromosomenverhältnisse, konstitutives Heterochromatin und Geschlechtsbestimmung bei einigen Arten der Gattung *Chrysomya* (Calliphoridae, Diptera)" *Chromosoma* 58: 113-136.
- Ullerich, F. H. (1984).** "Analysis of Sex Determination in the Monogenic Blowfly *Chrysomya rufifacies* by Pole Cell Transplantation." *Mol Gen Genet* 193: 479-487.
- Voges, D., et al. (1999).** "The 26S proteasome: a molecular machine designed for controlled proteolysis." *Annu Rev Biochem* 68: 1015-68.

- Wagoner, D. E. (1969).** "Presence of male determining factors found on three autosomes in the house fly, *Musca domestica*." *Nature* 223(5202): 187-8.
- Walker, J. J., et al. (2000).** "The *Drosophila melanogaster* sex determination gene *sisA* is required in yolk nuclei for midgut formation." *Genetics* 155(1): 191-202.
- Wallis, M. C., et al. (2008).** "Sex determination in mammals--before and after the evolution of SRY." *Cell Mol Life Sci* 65(20): 3182-95.
- Watson, J. D., et al. (2011).** "Watson Molekularbiologie, 6. aktualisierte Auflage." Pearson Studium.
- Weber, R. E., et al. (1985).** "Functional multiplicity and structural correlations in the hemoglobin system of larvae of *Chironomus thummi thummi* (Insecta, Diptera): Hb components CTT I, CTT II beta, CTT III, CTT IV, CTT VI, CTT VIIB, CTT IX and CTT X." *Comp Biochem Physiol B* 80(4): 747-53.
- Went, D. F. und R. Camenzind (1980).** "Sex determination in the dipteran insect *Heteropeza pygmaea*." *Genetica* 52/53: 373.
- Whitfield, L. S., et al. (1993).** "Rapid sequence evolution of the mammalian sex-determining gene SRY." *Nature* 364(6439): 713-5.
- Whiting, P. W. (1933).** "Selective Fertilization and Sexdetermination in Hymenoptera." *Science* 78(2032): 537-8.
- Whiting, P. W. (1943).** "Multiple Alleles in Complementary Sex Determination of *Habrobracon*." *Genetics* 28(5): 365-82.
- Whiting, P. W. und A. R. Whiting (1925).** "Diploid Males from Fertilized Eggs in Hymenoptera." *Science* 62(1611): 437.
- Wieschaus, E., et al. (1978).** "*fs(1)K10*, a germline-dependent female sterile mutation causing abnormal chorion morphology in *Drosophila melanogaster*." *Roux's Arch. Dev. Biol.* 184: 75-82.
- Wilkins, A. S. (1995).** "Moving up the hierarchy: a hypothesis on the evolution of a genetic sex determination pathway." *Bioessays* 17(1): 71-7.
- Willhoeft, U. und G. Franz (1996).** "Identification of the sex-determining region of the *Ceratitis capitata* Y chromosome by deletion mapping." *Genetics* 144(2): 737-45.
- Wobus, U., et al. (1990).** "A new transposable element in *Chironomus thummi*." *Mol Gen Genet* 222(2-3): 311-6.
- Wu, J. Y. und T. Maniatis (1993).** "Specific interactions between proteins implicated in splice site selection and regulated alternative splicing." *Cell* 75(6): 1061-70.
- Wülker, W. (1968).** "Die Verwendung der Imaginalscheiben zur Bestimmung des Entwicklungszustandes von *Chironomus*-Larven (Dipt.)." *Z. Morphol. Tiere*.
- Wyss, C. (1982).** "*Chironomus tentans* epithelial cell lines sensitive to ecdysteroids, juvenile hormone, insulin and heat shock." *Exp Cell Res* 139(2): 309-19.
- Wyss, C. (1982).** "Ecdysterone, insulin and fly extract needed for the proliferation of normal *Drosophila* cells in defined medium." *Exp Cell Res* 139(2): 297-307.
- Yang, C., et al. (2007).** "Prevalence of the initiator over the TATA box in human and yeast genes and identification of DNA motifs enriched in human TATA-less core promoters." *Gene* 389(1): 52-65.

- Yang, P., et al. (2004).** "Purification of the Arabidopsis 26 S proteasome: biochemical and molecular analyses revealed the presence of multiple isoforms." *J Biol Chem* 279(8): 6401-13.
- Yanisch-Perron, C., et al. (1985).** "Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors." *Gene* 33(1): 103-19.
- Yen, H. C., et al. (2003).** "Rpn5 is a conserved proteasome subunit and required for proper proteasome localization and assembly." *J Biol Chem* 278(33): 30669-76.
- Yen, H. C., et al. (2003).** "Schizosaccharomyces pombe Int6 and Ras homologs regulate cell division and mitotic fidelity via the proteasome." *Cell* 112(2): 207-17.
- Younger-Shepherd, S., et al. (1992).** "deadpan, an essential pan-neural gene encoding an HLH protein, acts as a denominator in Drosophila sex determination." *Cell* 70: 911-922.
- Yu, L., et al. (2000).** "Thirty-plus functional families from a single motif." *Protein Sci* 9(12): 2470-6.
- Zahler, A. M., et al. (1992).** "SR proteins: a conserved family of pre-mRNA splicing factors." *Genes Dev* 6(5): 837-47.
- Zahler, A. M. und M. B. Roth (1995).** "Distinct functions of SR proteins in recruitment of U1 small nuclear ribonucleoprotein to alternative 5' splice sites." *Proc Natl Acad Sci U S A* 92(7): 2642-6.

7 Eidesstattliche Erklärung

Hiermit erkläre ich, die vorliegende Dissertation selbstständig und ohne fremde Hilfe angefertigt zu haben. Weiterhin versichere ich, wörtlich übernommene Ausführungen anderer Autoren oder dem Sinne nach anderen Veröffentlichungen entnommene Stellen entsprechend gekennzeichnet und die Quellen zitiert zu haben.

.....
(Ort, Datum)

.....
(Thomas Herold)