

Genomanalyse

von

Prodiamesa olivacea

Dissertation

zur Erlangung des Grades

Doktor der Naturwissenschaften

(Dr. rer. nat.)

am Fachbereich Biologie der

Johannes Gutenberg-Universität in Mainz

Sarah Brunck

geb. 08.08.1987 in Mainz

Mainz, 2016

Dekan:

1. Berichterstatter:

2. Berichterstatter:

Tag der mündlichen Prüfung:

Inhaltsverzeichnis

Inhaltsverzeichnis.....	iii
1 Einleitung	1
1.1 Die Familie der Chironomiden.....	1
1.1.1 Die Gattung <i>Chironomus</i>	3
1.1.2 Die Gattung <i>Prodiamesa</i>	6
1.2 Die Struktur von Insekten-Genomen am Beispiel der Chironomiden	9
1.2.1 Hochrepetitive DNA-Sequenzen.....	11
1.2.2 Mittelrepetitive DNA-Sequenzen bzw. Gen-Familien	13
1.2.3 Gene und genregulatorische Sequenzen.....	17
1.3 Zielsetzung	20
2 Material und Methoden.....	22
2.1 Versuchstiere.....	22
2.2 Speziesbestimmung <i>P. olivacea</i>	22
2.3 Geschlechtsbestimmung bei Chironomiden-Larven.....	23
Molekularbiologische Methoden	24
2.4 Isolation von Nukleinsäuren	24
2.4.1 Isolation von genomischer DNA aus Chironomiden-Larven	24
2.4.2 Isolation von Einzellarven- DNA.....	24
2.4.3 Isolation von Plasmid- DNA.....	25
2.4.4 Präparation von Gesamt-RNA (Guanidinthiocyanat (GTC)- Methode) (Chomczynski und Sacchi 1987).....	25
2.4.5 Präparation von Gesamt-RNA aus Einzellarven (GTC-Methode)..... (Chomczynski und Sacchi 1987).....	26

2.5 Konzentrationsbestimmung von Nukleinsäuren	26
2.5.1 NanoDrop®.....	26
2.5.2 QuBit™.....	26
2.5.3 Agilent 2100 Bioanalyzer.....	27
2.6 Aufreinigung von Nukleinsäuren.....	27
2.6.1 Phenol-Chloroform-Extraktion.....	27
2.6.2 Fällern von Nukleinsäuren	27
2.6.3 Entfernen von Verunreinigungen	28
2.7 Elektrophoretische Auftrennung von DNA	28
2.7.1 Testgele	29
2.7.2 Präparative bzw. analytische Gele.....	29
2.8 Wiedergewinnung von DNA-Fragmenten aus Agarosegelen.....	29
2.8.1 Elektroelution.....	29
2.8.2 Wiedergewinnung mittels Silica-Membran.....	30
2.9 Klonierung von DNA.....	30
2.9.1 Restriktion genomischer DNA	30
2.9.2 Ligation	30
2.9.3 Transformation mittels Elektroporation	31
2.10 DNA-Markierungstechniken.....	31
2.10.1 Random Primed Oligo Labeling.....	31
2.10.2 Nick Translation	32
2.11 Hybridisierungstechniken.....	32
2.11.1 Fluoreszenz- <i>in-situ</i> -Hybridisierung (FISH)	32
2.11.2 Southern-Hybridisierung bzw. „Pirrotta“	34
2.11.3 Koloniefilter-Hybridisierung (Grunstein und Hogness 1975)	36
2.12 Polymerase Kettenreaktion (PCR) (Mullis et al. 1986)	37

2.12.1 Standard-PCR.....	37
2.12.2 cDNA-Synthese.....	38
2.12.3 RT-PCR.....	38
2.13 Sequenzierung	38
2.13.1 Sanger Sequenzierung	38
2.13.2 Illumina <i>Next-Generation Sequencing</i>	39
2.14 Bioinformatische Methoden.....	40
2.14.1 Konvertierung der Rohdaten.....	40
2.14.2 Importieren der Rohdaten.....	40
2.14.3 Trimmen der Rohdaten	40
2.14.4 <i>De novo</i> -Assembly der getrimmten Daten	40
2.14.5 Kartierung von genomischen bzw. transkriptomischen Daten.....	41
2.14.6 Blast	41
2.15 Puffer und Lösungen	42
3 Ergebnisse.....	44
3.1 Isolierung von Nukleinsäuren aus <i>P. olivacea</i> -Larven/-isolierten Gehirnen.....	44
3.1.1 Isolierung von genomischer DNA aus unterschiedlichen Geweben.....	44
3.1.2 Präparation von Gesamt-RNA aus <i>P. olivacea</i> -Larven	45
3.2 NGS-generierte Daten.....	45
3.3 Analyse von hochrepetitiver DNA aus <i>P. olivacea</i>	47
3.3.1 Analyse von hochrepetitiver DNA mittels Restriktionsverdau genomischer DNA	47
3.3.2 Analyse von hochrepetitiver DNA durch Southern-Analyse	48
3.3.3 Chromosomale Lokalisation von hochrepetitiver DNA mittels FISH.....	49
3.3.4 Sequenzanalyse von hochrepetitiver DNA aus <i>P. olivacea</i>	50
3.4 Die Globin-Gene von <i>P. olivacea</i>	52
3.4.1 Chromosomale Lokalisation und genomische Organisation.....	52

3.4.2 Genstruktur	54
3.5 Die Histon-Gene von <i>P. olivacea</i>	58
3.5.1 Die Organisation.....	58
3.5.2 Die Lokalisation	58
3.5.3 Die Nukleotid-Sequenzen	59
3.5.4 <i>Orphon</i> -Histon-Gene	60
3.6 Die Identifikation von homologen Genen aus anderen Dipteren in <i>P. olivacea</i>	63
3.6.1 Identifikation, Charakterisierung und Lokalisation von Genen aus der Geschlechtsbestimmungskaskade von <i>D. melanogaster</i> in <i>P. olivacea</i>	63
3.6.2 Identifikation von homologen Genen aus der geschlechtsbestimmende Region von <i>C. thummi</i>	77
3.7 Analyse des larvalen <i>P. olivacea</i> -Transkriptoms.....	87
3.8 Chromosomenkartierung bei <i>C. thummi/piger</i> und <i>P. olivacea</i>	89
4 Diskussion	93
4.1 Das <i>P. olivacea</i> -Genom.....	93
4.1.1 Sat-DNA des <i>P. olivacea</i> -Genoms.....	93
4.1.2 Genfamilien aus <i>P. olivacea</i>	95
4.1.3 <i>Low-copy</i> -DNA-Sequenzen.....	102
4.2 Die Chromosomenevolution der Chironomiden.....	112
5 Zusammenfassung	115
6 Abkürzungsverzeichnis.....	117
7 Abbildungsverzeichnis	121
8 Tabellenverzeichnis.....	125
9 Literaturverzeichnis	126
10 Anhang.....	144
11 Curriculum Vitae.....	152

12 Danksagung	154
13 Eidesstattliche Erklärung	155

1 Einleitung

1.1 Die Familie der Chironomiden

Die zirka (ca.) 15.000 bekannten Arten (Armitage et al. 2012), welche zu der Familie der Zuckmücken gehören, sind weltweit verbreitet. Sie gehören der Ordnung Diptera, Unterordnung Nematocera und der Teilordnung Culicomorpha an. Die Zuckmücken sind holometabole Insekten, welche ihre Eier in einer gallertigen Masse, dem Gelege, ablegen. Die Larven, die bei der Art *Chironomus riparius* (*C. riparius*) nach ca. drei Tagen schlüpfen, durchlaufen meist vier abgrenzbare larvale Stadien. Den längsten Entwicklungsabschnitt, den der Larve, verbringen sie im Sediment und ernähren sich von Algen und Detritus, wobei auch Spezies mit anderen Ernährungsformen zum Beispiel (z. B.) räuberische existieren (Armitage et al. 2012). Die Ansprüche an abiotische und biotische Umweltfaktoren unterscheiden zwischen den zahlreichen Arten stark. Viele *Chironomus*-Arten besitzen als Larven den Blutfarbstoff Hämoglobin, wodurch eine rote Färbung der Larven zustande kommt. Das Hämoglobin der Chironomiden ist in der Hämolymphe der Larven frei gelöst und ermöglicht das Überleben in zum Teil (z. T.) sehr sauerstoffarmen Substrat (Burmester und Hankeln 2007), (Hankeln et al. 1988), (Nowak et al. 2006), (Weber et al. 1985). Neben sauerstoffarmen Habitaten kommen Chironomiden auch in kalten oder stark verunreinigten Habitaten vor. Aufgrund solcher Eigenschaften beziehungsweise (bzw.) Anpassungen werden Chironomiden als Bioindikatoren zur Bestimmung der Gewässergüte eingesetzt (Oliver 1971), (Pinder 1986), (Taenzler et al. 2007). Die Art *Belgica antarctica* (*B. antarctica*, Abbildung 1, A rot), deren Genom vor kurzem mittels *Next Generation Sequencing* (NGS) sequenziert wurde, besiedelt z.B. Habitate in der Antarktis und ist ein Beispiel für die Anpassung an extrem kalte Umweltbedingungen (Kelley et al. 2014).

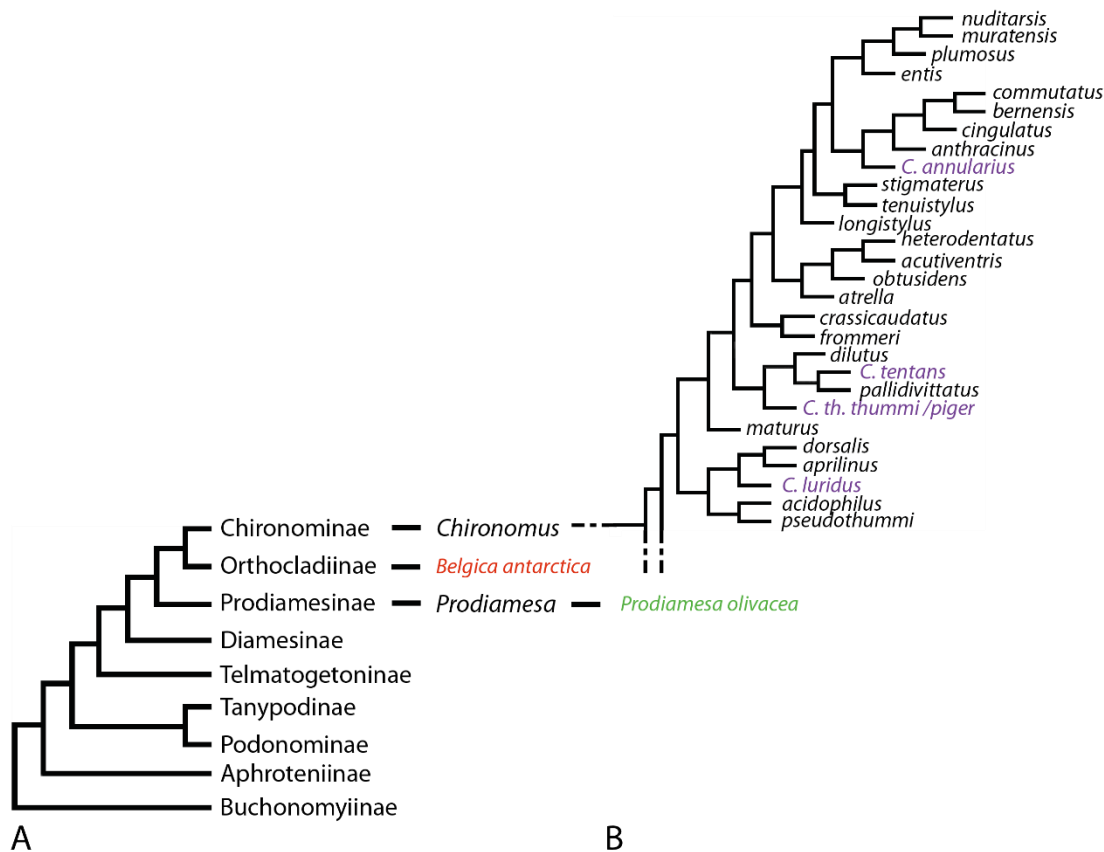


Abbildung 1 **Stammbaum der Familie der Zuckmücken (Chironomidae) und der Gattung Chironomus.** A zeigt den Stammbaum der Familie der Zuckmücken mit neun Unterfamilien. Dazu zählen unter anderem Chironominae und Prodiamesinae. Zur Unterfamilie Prodiamesinae gehört die Gattung Prodiamesa. Eine der für diese Gattung bekannten elf Arten, *Prodiamesa olivacea* (*P. olivacea*; grün), ist Untersuchungsobjekt dieser Arbeit. Die Gattung *Chironomus* gehört zur Unterfamilie Chironominae und besitzt eine Vielzahl an Arten. B zeigt einen kleinen Ausschnitt des Stammbaums der Gattung *Chironomus* mit den in dieser Arbeit vergleichend untersuchten Arten (violett). A verändert nach (Cranston et al. 2012); B verändert nach (Martin 2010)

Es gibt Arten, deren Larven sogenannte Wohnraumröhrchen ausbilden. Die Wohnraumröhrchen werden von den Larven mittels des in sogenannten Speicheldrüsen der Larven produzierten Spinnsekrets gesponnen und von außen zugeführten Sedimentteilen gebaut. Die Larven verpuppen sich nach vier Larvenstadien und innerhalb von wenigen Tagen vollenden sie ihre Metamorphose zu Imagines. Mit Hilfe von „Schwimmfächern“ gelangen die Tiere an die Wasseroberfläche und leben meist nur wenige Tage als fliegende Mücke, deren einziges Ziel die Verpaarung und Eiablage ist. Meist nehmen die Imagines nicht einmal mehr Nahrung auf und sterben kurz nach der Eiablage. Die Imago weist im Gegensatz zur Larve einen morphologisch distinkten Geschlechts-Dimorphismus auf. Männliche Imagines besitzen im Vergleich zu weiblichen auffällig große, wirtelig behaarte Antennen. In der Wissenschaft wurde das Interesse für die Zuckmücken durch die Entdeckung der Balbiani-Ringe in Riesenchromosomen-Präparaten aus Speicheldrüsenzellen der

Chironomiden schon früh geweckt (Balbiani 1881), wobei allerdings zu diesem Zeitpunkt nicht bekannt war, dass es sich bei den beschriebenen „Kernfäden“ um polytäre Chromosomen handelte. Die gut sichtbaren Chromosomen dieser Zellen sind aufgrund ihres hohen Polytäniegrads $2^{14}n$ (entspricht $16,384n$) und höher für zytologische Analysen gut geeignet. So können solche Chromosomenpräparate z. B. für Fluoreszenz-*in-situ*-Hybridisierungen (FISH) zur Lokalisation von Genen und anderen DNA-Sequenzen als Grundlage dienen. Die Bedeutung der Chironomiden und deren Speicheldrüsenchromosomen für die genetische Wissenschaft wird unterstrichen durch die kürzlich durchgeführte Entschlüsselung des Genoms der *Chironomus*-Art *C. tentans* (Abbildung 1, B violett) (Kutsenko et al. 2014).

1.1.1 Die Gattung *Chironomus*

Die Gattung *Chironomus* gehört zu der Unterfamilie Chironominae. Viele wissenschaftliche Arbeiten beschäftigen sich mit den Arten der Gattung *Chironomus*. Es gibt weitaus mehr Information zu genetischen Themen als zur Gattung *Prodiamesa* (Abbildung 1, A/B), die das bevorzugte Untersuchungsobjekt in dieser Arbeit ist. Für manche Fragestellungen war es daher notwendig, die Ergebnisse aus *Prodiamesa olivacea* (*P. olivacea*) mit Untersuchungen und Ergebnissen der Arten *C. thummi* bzw. *C. piger*, *C. luridus*, *C. tentans* und/oder *C. annularius* zu ergänzen (Abbildung 1, B violett). Die Art *C. thummi* (Kieffer 1911) wurde zuvor von (Meigen 1804) als *C. riparius* beschrieben, in dieser Arbeit wird sie weiterhin als *C. thummi* bezeichnet, da auch viele Einträge in Datenbanken unter dem Speziesname „*thummi*“ existieren.

Die Chromosomenanzahl der meisten Chironomiden variiert zwischen $n = 3$ oder $n = 4$ homologen Paaren. Häufig ist im Laufe der Evolution zu unterschiedlichen Chromosomenarm-Kombinationen gekommen, die vermutlich Folge von reziproken Translokationen sind (Keyl und Keyl 1959), (Keyl 1961), (Keyl 1962). In vielen Spezies treten keine typischen Heterochromosomen auf (Bauer und Beermann 1952), (Beermann 1955). In manchen Fällen kann anhand von zytologischen oder genetischen Merkmalen ein Proto-Y identifiziert werden. Das Proto-Y der Gattung *Chironomus* wurde über männchenspezifische Inversionen oder mittels männchenspezifischer heterochromatischer Banden identifiziert. Umfangreiche Untersuchungen bei vielen Spezies zeigten, dass der männchenbestimmende

Faktor (M) auf unterschiedlichen Chromosomenarmen lokalisiert sein kann, jedoch nie auf Chromosomenarm E (Acton 1957), (Acton 1959), (Atchley und Martin 1971), (Beermann 1955), (Hägele 1985), (Keyl 1961), (Keyl 1962), (Martin und Wülker 1971), (Martin 1974), (Martin et al. 1978), (Martin et al. 1980), (Martin und Lee 1984), (Rempel et al. 1962), (Rosin und Fischer 1972), (Thompson und Bowen 1972), (Wulker und Am Klotzli 1973). Teils kann der Locus sogar innerhalb einer Art auf unterschiedlichen Chromosomen lokalisiert sein. In Tabelle 1 sind die Chromosomenarme, welche die *sex determining region* (SDR) tragen aufgelistet.

Tabelle 1 Die chromosomale Lage der SDR in *C. annularius*, *C. luridus*, *C. tentans*, *C. pallidivittatus* und der Schwesterspezies *C. thummi* und *C. piger*. Die chromosomale Lage der SDR und ihr spezifisches Erkennungsmerkmal für die entsprechende Art mit der Angabe des Autors, welcher dies zuerst beschrieben hat. Tabelle verändert nach Stiehl, unveröffentlicht.

Arm	Spezies	Erkennungsmerkmal	Autor
F	<i>C. annularius</i>	komplexe Inversion	(Beermann 1955) (Keyl 1962)
F	<i>C. luridus</i>	einfache Inversion	(Acton 1957)
B ^a C ^b F ^c	<i>C. tentans</i> (EU)	^a komplexe Inversion ^b komplexe perizentrische I. mit Arm F ^c komplexe Inversion	^a ; ^b ; ^c (Beermann 1955) ^a (Rosin und Fischer 1972)
F	<i>C. piger</i>	Spezieshybriden	(Hägele 1985)
B ^a F ^c	<i>C. pallidivittatus</i>	^a komplexe Inversion ^c komplexe Inversion	^a ; ^c (Beermann 1955) ^a (Rosin und Fischer 1972)
F	<i>C. thummi</i>	Spezieshybriden, heterochromatische Bande	(Hägele 1985)

Die geschlechtsspezifischen Inversionen, welche in *C. tentans* und *C. pallidivittatus* zuerst von Beermann bzw. Rosin und Fischer beschrieben wurden, treten nur im Männchen und dort stets heterozygot auf (Beermann 1955), (Rosin und Fischer 1972). Daher kann angenommen werden, dass das Prinzip der männlichen Heterogamete bei der Gattung *Chironomus* zutrifft. Es wird angenommen, dass ein dominanter männchenbestimmender Faktor M für die Ausprägung des männlichen Geschlechts verantwortlich ist.

Mit den Schwesternarten *C. thummi* und *C. piger* wurden Kreuzungsversuche durchgeführt, welche zur Identifikation von M auf Chromosomenarm F führten (Hägele 1985). Die Rückkreuzung männlicher F1-Individuen mit weiblichen *C. piger* zeigte dabei in dem Chromosomen III eine Asynapsis in der Region D₃d-g, Arm F. Innerhalb dieser Region zeigt sich in *C. thummi*-Männchen eine heterochromatische Bande an Position D₃e₁ (Hägele 1985). Obwohl diese Bande in männlichen *C. piger*-Individuen nicht zu finden ist, kann davon

ausgegangen werden, dass der Geschlechtsbestimmer M dort ebenfalls lokalisiert ist, da die Heterochromatinbildung nicht zwangsläufig im funktionellen Zusammenhang mit der Determination des Geschlechtes steht. Die Lage von M in *C. thummi* bzw. *piger* scheint zu einem der beiden identifizierten Loci aus *C. tentans* und *C. annularius* (Tabelle 1) homolog zu sein. Der homologe Locus auf Chromosomenarm F der besagten Arten trägt M, wobei es sich möglicherweise um dasselbe Gen handelt (Kraemer und Schmidt 1993).

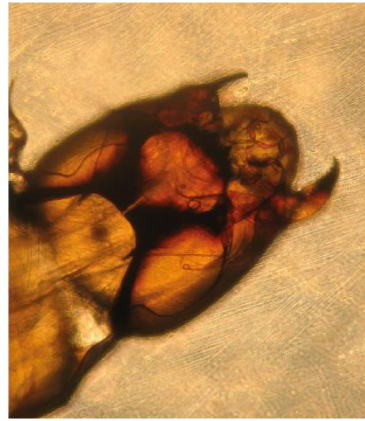
Die Sequenz der SDR aus *C. thummi/piger* wurde mittels DNA-Bibliotheken, *chromosome walk* und NGS-Sequenzierungen ermittelt und analysiert (Kraemer und Schmidt 1993), (Blumensatt 2009), (Funke 2011), Stiehl, unveröffentlicht. Die 200 kb-lange Sequenz aus *C. thummi* umfasst 15 Gene (Abbildung 35, B) und die 120 kb-lange Sequenz aus *C. piger* 10 dieser Gene. Die Zuordnung dieser Referenzsequenzen zu Proto-X oder Proto-Y war bisher nicht eindeutig möglich. Jedoch konnte über genomische Southern Analysen von Stiehl ein geschlechtsspezifischer Unterschied für das Gen *fs(1)K10-like* identifiziert werden (Stiehl 2009). Das Proto-Y der *Chironomus*-Spezies *piger*, *thummi* und *luridus* trägt eine Duplikation des Gens ((Stiehl 2009), (Stiehl, unveröffentlicht), (C. Kraemer, persönliche Mitteilung)).

1.1.2 Die Gattung *Prodiamesa*

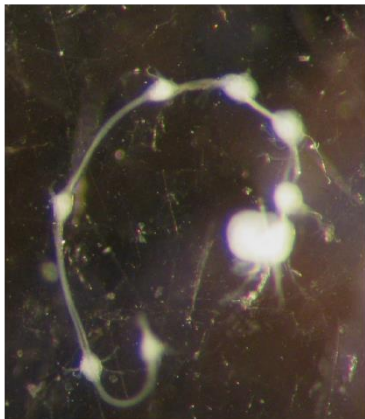
P. olivacea Meigen 1818 gehört zu der Nematoceren-Familie Chironomidae. Die elf Arten der Gattung *Prodiamesa* (Kieffer 1906, Abbildung 1, A), welche im Gegensatz zu vielen Chironomiden eine gelblich-weiße Farbe der Larve zeigen, sind morphologisch gut bestimmbar. Ein rot-brauner Kopf (Leathers 1922) der Larve von *P. olivacea* zeigt eine charakteristische schwarze Zeichnung auf der Ventralseite und ist neben den Tasthaaren der Ventromentalplatte typisch für diese Spezies (Cambridge Univ Press 2007) (Abbildung 2, B). Diese Merkmale dienen der eindeutigen Bestimmung von *P. olivacea*-Larven. Larven von *P. olivacea* kommen bevorzugt in sandigen Fließgewässern von Ende April bis Mai, Juli und August bis September vor (Leathers 1922), (Schmid 1992). Mitglieder der Unterfamilie Prodiamesinae besitzen einen haploiden Karyotyp mit nur drei Chromosomen (Petrova und Zhirov 2014). Von den drei mitotischen identifizierbaren Chromosomenpaaren sind zwei Chromosomenpaare metazentrisch und ein Chromosomenpaar submetazentrisch organisiert (Michailova 1977), (Zacharias 1979). Das Genom von *P. olivacea* enthält 0,25 pg DNA (2C DNA). Dies entspricht einer Genomgröße von $1,3 \times 10^8$ bp (Zacharias 1979). Ein Chromozentrum, welches in polytänen Speicheldrüsen-Zellkernen (Speicheldrüsen in Abbildung 2, D) sichtbar ist, verbindet die zwei beiden Chromosomen (Abbildung 2, E). Das dritte und kleinste Chromosom ist nicht Teil des Chromozentrums (Abbildung 2, E). Es zeichnet sich durch eine auffällige konzentrische Bänderung und einen verdichteten heterochromatischen „Knopf“ aus. Es umfasst 20,8 % der zytoplasmatisch messbaren „total-Feulgen-DNA“. Während der larvalen Polytänisierung nimmt sein Anteil jedoch auf 6,8 % ab. Dieses Ergebnis wird von Zacharias mit der Unterreplikation des rechten Arms des dritten Chromosoms erklärt (Zacharias 1979). Die Unterreplikation ist im Weibchen im Zellkern von Zellen aus Speicheldrüsen, Malpighi-Gefäßen und aus dem Darm zu beobachten. In den entsprechenden Zellkernen aus männlichen Tieren ist nur ein Gehalt von 3 % statt der 6 % im Weibchen festzustellen. Dies wird als selektive Replikation des Y-Chromosoms bei *P. olivacea* beschrieben (Zacharias 1979).



A



B



C



D



E

Abbildung 2 *P. olivacea*. A Ganzkörper-Aufnahme einer *P. olivacea*-Larve B Nahaufnahme eines *P. olivacea*-Larven-Kopfs, welcher die schwarze Zeichnung und die Tastaare auf der Ventromentalplatte zeigt C Frei-präpariertes Nervensystem einer *P. olivacea*-Larve: Oberschlundganglion und Unterschlundganglion (Gehirn) und z.T. ventraler Nervenstrang mit segmentalen Ganglien (vergrößert) D Nahaufnahme einer Speicheldrüse mit Zellen E Mikroskopische Aufnahme von polytären Riesenchromosomen aus einer *P. olivacea*-Larve

Zacharias postulierte, dass es sich bei dem Geschlechtsbestimmungsmechanismus in *P. olivacea* um männliche Heterogametie handeln muss, da im Speicheldrüsenzellkern das dritte Chromosom im Weibchen als zwei Chromosomen vorliegen, im Männchen jedoch immer nur ein Chromosom zu erkennen ist (Zacharias 1981).

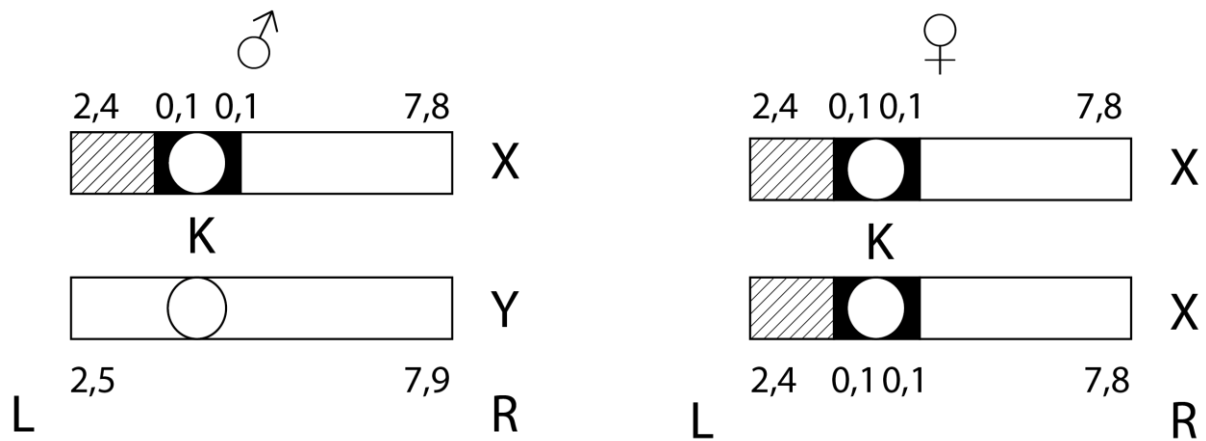


Abbildung 3 **Hypothetische Zusammensetzung des dritten Chromosomenpaares in weiblichen und männlichen diploiden *P. olivacea*-Individuen.** Links ist das männliche Chromosomenpaar (XY) abgebildet, wobei das Y-Chromosom vollständig unterrepliziert ist, wogegen das X-Chromosom einen euchromatischen linken Arm (schraffierte Fläche) aufzeigt. Rechts ist das weibliche Chromosomenpaar (XX) abgebildet. Ein Zellkern weist somit das Doppelte an Euchromatin (schraffierte Fläche) im Vergleich zum männlichen diploiden Zellkern auf. Schwarz dargestellt ist das polytänisierte β -Heterochromatin, welches das nicht polytänisierte α -Heterochromatin umschließt. K: Kinetochore L:Links R:Rechts X: X-Chromosom Y: Y-Chromosom Abbildung verändert nach (Zacharias 1981)

Aufgrund der unterschiedlichen Polytänisierung der Chromosomenarme zu verschiedenen Zeitpunkten vermutet Zacharias, dass sich die Gene zur Geschlechtsbestimmung auf dem Chromosomenarm 3L befinden, da die Chromosomenarme 3R in beiden Geschlechtern unterrepliziert vorliegen (Zacharias 1981). Der homologe Chromosomenarm könnte vielmehr Gene tragen, welche in der Zellteilung benötigt werden, da diese in nicht-zellteilenden Geweben (z. B. Speicheldrüsen) entbehrlich sind. Weiter geht Zacharias nicht davon aus, dass sich auf dem Chromosomenarm 3R Sat-DNA befindet, da eine entsprechende CsCL-Dichtezentrifugation mit Gehirn-Material aus 600 männlich und weiblichen gemischten Larven durchgeführt wurde ohne, dass eine typische Satellitenbande zu entdecken war (Zacharias 1981).

1.2 Die Struktur von Insekten-Genomen am Beispiel der Chironomiden

Die Gesamtheit der genetischen Information eines Lebewesens wird als Genom bezeichnet. Dabei unterscheiden sich Organismen mit Zellkernen (Eukaryoten) von denen ohne Zellkern. Seit langer Zeit ist bekannt, dass eukaryotischer Genome mehr oder weniger große Anteile an repetitiver DNA besitzen (Kuhn, G C S 2015) (Abbildung 4, blaue Flächen). Der kodierende Teil des Genoms enthält die Gene und genregulatorische Sequenzen. Zu denen mit Genen gekoppelten Sequenzen zählen die Intron-Sequenzen, Promotor-Bereiche, untranslatierte Regionen (UTR) und Pseudogene (Abbildung 4, grüne Flächen). Der nicht-kodierende Teil enthält repetitive DNA-Sequenzen (wie Makro-, Mini-, Mikro-Satelliten-DNA, Genfamilien und Retrotransposons) und singuläre DNA-Sequenzen wie regulatorische Bereiche und miRNAs (Abbildung 4, blaue Flächen).

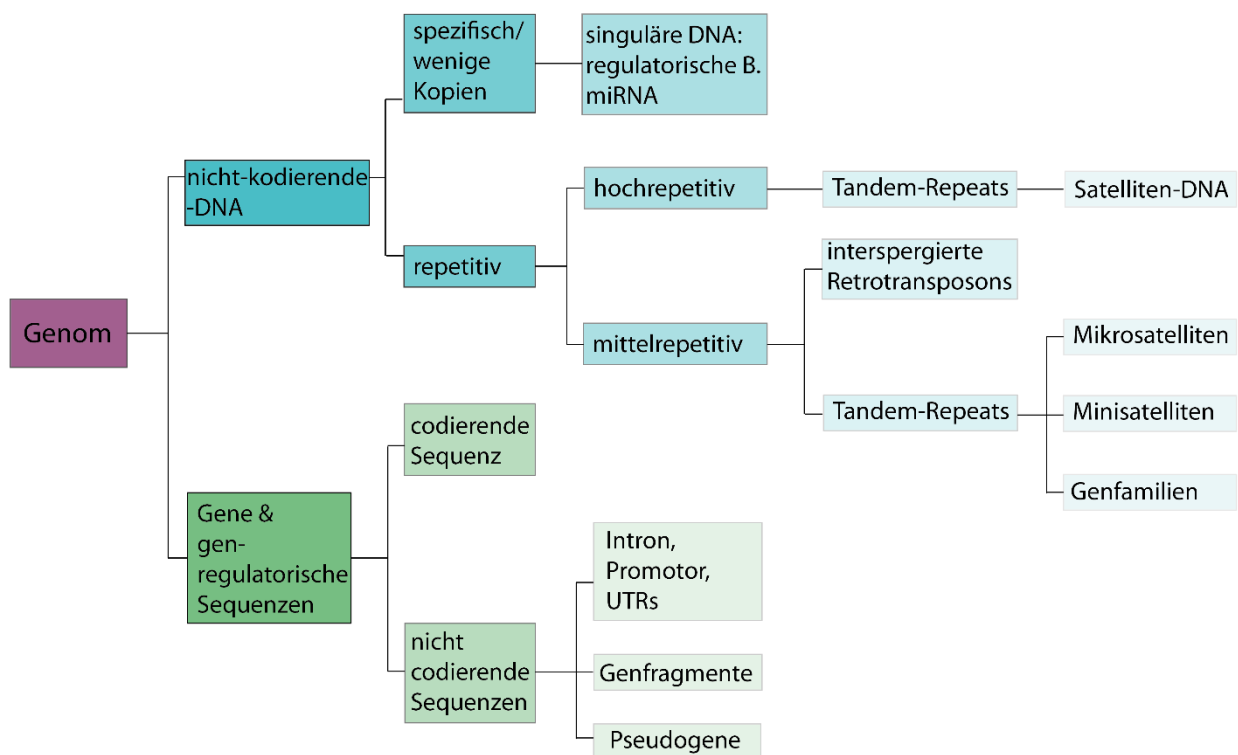


Abbildung 4 Zusammensetzung eines eukaryotischen Genoms. Bereiche (B.), untranslatierte Region (UTR)

Abbildung verändert nach (Watson 2011), (Klug et al. 2007) und (Fletcher und Hickey 2013)

Repetitive DNA-Sequenzen lassen sich in hoch- und mittelrepetitive Sequenzen unterteilen. Zu hochrepetitiven Sequenzen gehört die Satelliten-DNA (Sat-DNA), die häufig in der Umgebung der Zentromere zu finden ist. Zu den mittelrepetitiven Sequenzen zählen

interspergierte Retrotranspons (LINE, SINE) und tandem-repetitive Sequenzen. Die tandem-repetitiven Sequenzen bilden häufig Sequenzfamilien oder Mikro- bzw. Minisatelliten.

Die Zusammensetzung von Insekten-Genomen

Aufgrund der in den letzten Jahren erworbenen Sequenzierungstechniken der nächsten Generation (NGS) wurden viele Genome sequenziert darunter auch etliche Genome von Insekten. Einen Auszug aus den bisher sequenzierten Insekten-Genomen zeigt Tabelle 2.

Tabelle 2 Genomgrößen und Anzahl der Gene unterschiedlicher Insekten.

Art	Ordnung	Genomgröße	Gen-anzahl	TE [%]	Autor
<i>Aedes aegypti</i>	Diptera	1376 Mb	15.419	47	(Nene et al. 2007)
<i>Anopheles gambiae</i>	Diptera	278 Mb	13.683	16	(Holt et al. 2002)
<i>Apis mellifera</i> (Honigbiene)	Hymenoptera	236 Mb	10.157		(Weinstock et al. 2006)
<i>Belgica antarctica</i>	Diptera	99 Mb	13.517	0,12	(Kelley et al. 2014)
<i>Bombyx mori</i>	Lepidoptera	432 Mb	14.623		(Abe et al. 2005a)
<i>Chironomus tentans</i>	Diptera	200 Mb	15.120	10-15	(Kutsenko et al. 2014)
<i>Culex quinquefasciatus</i>	Diptera	540 Mb	18.883	29	(Arensburger et al. 2010)
<i>Drosophila melanogaster</i>	Diptera	130 Mb	13.600	20	(Adams et al. 2000)
<i>Prodiamesa olivacea</i>	Diptera	113 Mb	?	?	(Zacharias 1979)

Es zeigt sich, dass die Genomgröße innerhalb der Klasse Insecta sehr unterschiedlich ist. Die 2014 sequenzierte Chironomiden-Art *B. antarctica* hat mit 99 Mb das bisher kleinste bekannte Insekten-Genom (Kelley et al. 2014). Dagegen hat die Gelbfiebermücke *Aedes aegypti* (*A. aegypti*) mit 1376 Mb ein für Insekten recht großes Genom (Nene et al. 2007). Häufig korreliert die Anzahl der Gene jedoch nicht mit der Genomgröße. Während diese Zahl wenig schwankt, ist die Menge an repetitiven DNA-Sequenzen sehr unterschiedlich (Kidwell 2002). Bei Arten mit großen Genomen ist repetitive-DNA in größerer Menge vorhanden. In der Gelbfiebermücke *A. aegypti* machen repetitive, transposable Elemente (TE) fast 50 % des Gesamt-Genoms aus, wogegen sie in *Anopheles gambiae* (*A. gambiae*) bei einer Genomgröße von

278 Mb nur 16 % darstellen (Nene et al. 2007), (Holt et al. 2002). Die Mücke *B. antarctica* hat sogar nur einen Anteil von 0,12 % TE in ihrem Genom (Kelley et al. 2014). Eine ähnliche Korrelation findet sich hinsichtlich der Länge der Introns: diese ist bei Arten mit kleinem Genom und weniger TE ebenfalls kleiner (Kelley et al. 2014), (Kutsenko et al. 2014), (Mount et al. 1992).

Relativ unterschiedliche Genomgrößen sind auch für die Schwesternarten *C. thummi* und *C. piger* beschrieben (Keyl 1965). Der Unterschied beruht ebenfalls vor allem auf hochrepetitiven DNA-Sequenzen (Schmidt 1981).

1.2.1 Hochrepetitive DNA-Sequenzen

Sat-DNA wurde zuerst von Kit als DNA beschrieben, welche durch ihre ungewöhnliche Basenzusammensetzung in der CsCl-Dichte-zentrifugation sich als "Satelliten"-Bande von der Haupt-DNA-Bande trennen lässt (Kit 1961). Sat-DNA kann einen erheblichen Anteil an der genomischen DNA von eukaryotischen Genomen (bei Insekten z. T. > 50 %, (Davis und Wyatt 1989)) bilden. Dabei ist sie häufig die Hauptkomponente des (perizentromerischen) Heterochromatins (Charlesworth et al. 1994). Viele Monomer-Einheiten (*repeat*-Einheiten), welche *head-to-tail* angeordnet sind, bilden lange tandem-repetitive-Cluster (Palomeque und Lorite 2008). Aufgrund dieser molekularen Organisation kann Sat-DNA durch geeignete Restriktionsenzyme in monomere und multimerer Repeat-Einheiten verdaut und als Leiternmuster sichtbar gemacht werden. Eine Studie der Insekten-Sat-DNAs deutete eine evolutionäre Konservierung gewisser Merkmale trotz ihrer Sequenz-Heterogenität an. Dabei wurden Merkmale wie Gesamtlänge, Monomer-Länge, Motive, besondere Regionen und sekundär bzw. tertiär Strukturen untersucht (Palomeque und Lorite 2008).

Im Genom von *C. thummi/piger* konnte eine perizentromere Sat-DNA nachgewiesen werden (Schmidt 1981). Diese Sat-DNA hat eine Monomer-Repeatlänge von ~ 120 bp. Typischerweise besitzt ein Monomer eine Schnittstelle für das Restriktionsenzym ClaI daher resultiert die Bezeichnung "Cla-Elemente" (Schmidt 1984), (Abbildung 5). Die Cla-Elemente im Genom von *C. thummi* und *C. piger* bilden einen erheblichen Teil des Unterschieds in der Genomgröße, die ca. 30 % zwischen Schwesternarten *C. thummi* und *C. piger* beträgt (Keyl 1965). Die Menge und Verteilung der Cla-Elemente in den Chromosomen von *C. thummi* und *C. piger* sind extrem unterschiedlich. In *C. thummi* gibt es ca. 70.000 Kopien an Cla-Elementen gegenüber nur

ca. 10.000 in *C. piger*. Während im Genom von *C. thummi* Cla-Elemente in mehr als 100 chromosomalen Loci auf allen Chromosomen sowohl im Eu- wie im Heterochromatin vorkommen, sind in *C. piger* die Cla-Elemente weitgehend auf die Zentromerregion beschränkt (Schmidt 1981), (Schmidt 1984).

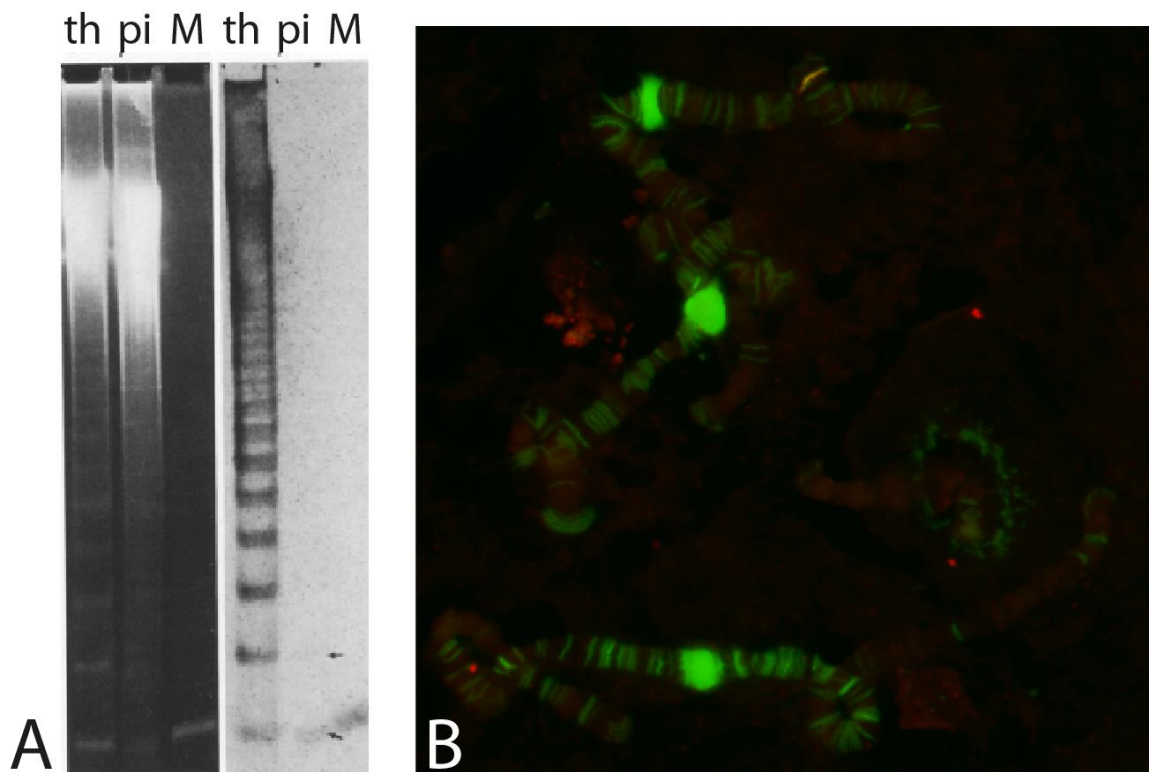


Abbildung 5 *Sat-DNA* aus *C. thummi/piger*. A zeigt das charakteristische Leiternmuster, welches nach der gelelektrophoretischen Auftrennung von restringierter genomischer DNA (th= thummi, pi= piger) mit ClaI bzw. isolierter Monomer (M)-DNA entsteht. B zeigt das Ergebnis einer FISH mit einer Cla-Element-Sonde (Trimer) (grüne Signale, FITC) an *C. thummi* Speicheldrüsenchromosomen.

A verändert nach (Schaefer und Schmidt 1981)

Die unterschiedliche Anzahl der Cla-Elemente in den Schwesternarten wird zudem als erster Schritt der Arten-Trennung gesehen (Israelewski 1983). Die Cla-Elemente wurden ebenfalls im *non-transcribed-spacer* (NTS) der rDNA- Gene von *C. thummi* identifiziert. In *C. piger* fehlen die Cla-Elemente an dieser Stelle. Dieser Zustand unterstützt die Theorie, dass es sich bei den Cla-Elementen um TE handelt (Schmidt und Godwin 1983).

1.2.2 Mittelrepetitive DNA-Sequenzen bzw. Gen-Familien

Zu der Klasse der mittelrepetitiven DNA-Sequenzen zählen vor allem die Mitglieder von Genfamilien.

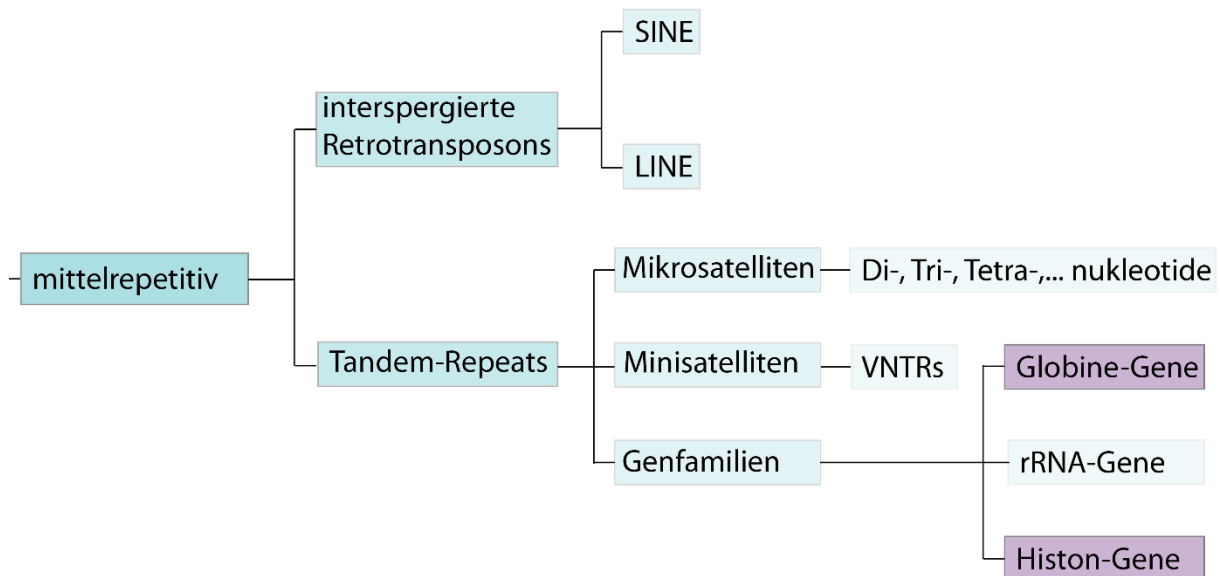


Abbildung 6 *Mittelrepetitive DNA-Sequenzen eines eukaryotischen Genoms.* Die in dieser Arbeit untersuchten Genfamilien sind lila unterlegt. Abbildung verändert nach (Klug et al. 2007)

Genfamilien entstehen durch Duplikationsereignisse von zunächst "single copy"-Genen. Die Vervielfachung von Genen kann mehrere Ursachen haben. Entweder wird das gesamte Genom oder zumindest einzelne Chromosomen dupliziert (Ohno 2013) oder Teilabschnitte von Chromosomen. Letzteres kommt oft durch inäquales (nicht homologes) Crossing-over oder Retrotransposition zustande. Meist findet inäquales Crossing-over zwischen DNA-Abschnitten mit hoher Sequenzidentität statt (Friedl 2001). Dies ist beispielsweise der Fall in der menschlichen Beta-Globin-Gen-Gruppe, dort findet diese Art von Rekombinationsereignis bevorzugt in einer 50 bp-langen identischen Sequenz statt (Metzenberg et al. 1991). Das Ergebnis einer Genduplikation oder Vervielfältigung kann unterschiedliche Konsequenzen haben. Zum einem hat es oft eine erhöhte Menge des Transkriptprodukts bzw. Proteins zur Folge (tRNA, rRNA und Histon-Gene) (Elgin und Weintraub 1975). Zum anderen können die duplizierten Gene sich unabhängig voneinander entwickeln. Dies kann durch Mutationen zum Funktionsverlust bzw. in der Entstehung von Pseudogenen führen. Dieser Mechanismus ist für die Immunglobulin-Gene (Huang et al. 1980), die 5 S rRNA-Gene (Jacq et al. 1977) oder die Globin-Gene (Nishioka et al. 1980) bekannt.

Die Hämoglobine der Chironomiden

Die Genfamilie der Hämoglobin-Gene entwickelte sich wahrscheinlich durch Genduplikationen und deren anschließende Diversifizierung. Es handelt sich dabei um eine sehr alte Genfamilie, da sie in allen fünf Reichen zumindest sporadisch auftritt (Friedl 2001). In Invertebraten zeichnen sich die Hämoglobine durch strukturelle und funktionelle Vielfalt aus (Dewilde 1999). Moens et al. zeigten, dass lediglich die Aminosäuren (AS)-Reste CD1 Phenylalanin (Phe) und F8 Histidin (His), welche zur Molekül-typischen Faltung wichtig sind, konserviert vorliegen (Moens et al. 1996).

Chironomiden gehören zu den wenigen Insektenpezies, die Hämoglobin als Sauerstoff-Transporter in ihrer Hämolymphe besitzen. Viele Chironomidenspezies leben als Larven in z. T. sehr sauerstoffarmen Substrat wie Schlamm aus eutrophierten Gewässern. So ist es nicht verwunderlich, dass Hämoglobin in Larven das Leben in einer sauerstoffarmen Umgebung ermöglicht. Rollet beschrieb 1861 zuerst den roten Blutfarbstoff der Zuckmücken-Larven (zitiert von (Weber 1980)), der von Lankaster 1867 als Hämoglobin bezeichnet wurde. Das Hämoglobin tritt in den vier Larvenstadien, Ovarien von Imagos, sowie in Eiern und Embryonen auf (Shilova K.S.), (Bergtrom et al. 1976), (Trewitt und Bergtrom 1981), (Trewitt et al. 1986). Das Hämoglobin besteht jedoch nicht nur aus einem Polypeptid, sondern ist eine Mischung aus mehreren Protein-Varianten. Dabei handelt es sich um Sequenz-Varianten die auf verschiedene Gene und nicht auf Allele zurückzuführen sind (Thompson und English 1966). Braunitzer et al. gelang es elektrophoretisch, zwölf verschiedene Hämoglobin-Komponenten von *C. thummi* zu trennen (Braun et al. 1968), (Goodman et al. 1983) und (Goodman et al. 1988). In der Hämolymphe kommen sowohl Monomeren als auch Dimeren-Hämoglobin-Proteine vor. Neben frei in der Hämolymphe gelösten Hämoglobine gibt es auch intrazellulär vorliegende Hämoglobine (Dewilde 1999). Die sezernierten primären Translationsprodukte tragen am N-Terminus ein 16 AS-langes, hydrophobes Signalpeptid, welches die Sekretion des Hämoglobins aus den Hämoglobin synthetisierenden "Fettzellen" ermöglicht (Baldwin et al. 1981), (Antoine und Niessing 1984). Es gibt zahlreiche Studien über invariante oder semi-invariante AS-Positionen in Hämoglobinen verschiedener Arten, welche Erkenntnisse über Funktion, Sauerstoffbindung, Dimer-Bindestelle oder interhelikale Strukturen liefern konnten (Goodman et al. 1983), (Bashford et al. 1987), (Moens et al. 1996), (Ptitsyn und Ting 1999), (Friedl 2001). Auch in anderen Dipteren wie in der zur Unterfamilie

Brachycera zugehörigen Gattung *Drosophila* und den zwei Mosquito-Arten *A. gambiae* und *A. aegypti* konnten Globin-Gene identifiziert werden (Burmester und Hankeln 1999), (Burmester et al. 2007). In den Chironomiden-Art *C. thummi* konnten zwei getrennte Genloci auf den Chromosomenarmen D und E nachgewiesen werden (Hankeln et al. 1988), (Schmidt et al. 1988). Die Hämoglobin-Gene sind in Clustern angeordnet, welche je nach Chromosomenarm Lage entsprechend Cluster D bzw. E bezeichnet werden. Die genomische Analyse zeigte eine Vielfalt an Hämoglobin-Genen, welche die zuvor beschriebene Protein-Vielfalt noch übersteigt (Friedl 2001).

Viele der untersuchten Globin-Gene aus Pflanzen, Vertebraten und Invertebraten besitzen meist mehrere Introns (Hardison 1996), (Moens et al. 1996), (Vinogradov et al. 1992). In Vertebraten ist die Verteilung der Introns einheitlich. Es gibt drei Exons welche durch zwei Introns unterbrochen sind. Die Position der Introns liegt in den Helices B und G der abgeleiteten Proteine (Abbildung 7) (Hardison 1996), (Blanchetot et al. 1983), (Blanchetot et al. 1986), (Weller et al. 1984). In Pflanzen gibt es ein zusätzliches Intron in der Helix E, welches aufgrund seiner Lage als „zentrales“ Intron bezeichnet wird (Gö 1981), (Wiborg et al. 1982), (Jensen et al. 1981). Die Intronverteilung der Hämoglobin-Gene der Invertebraten ist im Vergleich zu den Hämoglobin-Genen stark variabel. Analysen in Chironomiden zeigten, dass viele Hämoglobin-Gene keine Introns besitzen (Antoine und Niessing 1984), (Antoine et al. 1987), (Trewitt et al. 1987), (Hankeln et al. 1988), (Hankeln et al. 1997), (Rozynek et al. 1991), (Rozynek et al. 1989). Die drei Gene 2 β , 9 und 7A weisen jedoch in der Helix E ein Intron auf (Kao et al. 1994), (Gruhl et al. 1997), (Hankeln et al. 1997). Dies stellt eine Besonderheit dar, da Orthologe anderer Chironomiden-Gene kein Intron besitzen (Hankeln et al. 1991), (Hankeln et al. 1997). Aufgrund des sporadischen Auftretens und der nicht konservierten Lage der Introns wird vermutet, dass es sich um ein Integrationsereignis handelt (Hankeln et al. 1997), (Friedl 2001). In anderen Dipteren-Arten wie *Drosophila melanogaster* (*D. melanogaster*), *A. gambiae* und *A. aegypti* wurden ebenfalls Introns in Globin-Genen nachgewiesen (Burmester et al. 2006), (Burmester et al. 2007). Dabei weisen alle drei *globin-1*-Gene alle die Position G7.0 (Abbildung 7). Die Intron-Position B12.2 konnte in allen *globin-2*-Genen und in *globin-1* aus *A. aegypti* und *A. gambiae* nachgewiesen werden. Die Position D7.0 tritt nur in *D. melanogaster globin-1* auf aber nicht in den Globin-Genen der Arten *A. aegypti* oder *A. gambiae*. Das *globin-1*-Gen aus *A. aegypti* trägt als einzige der drei Arten an Position E18.0 ein Intron.

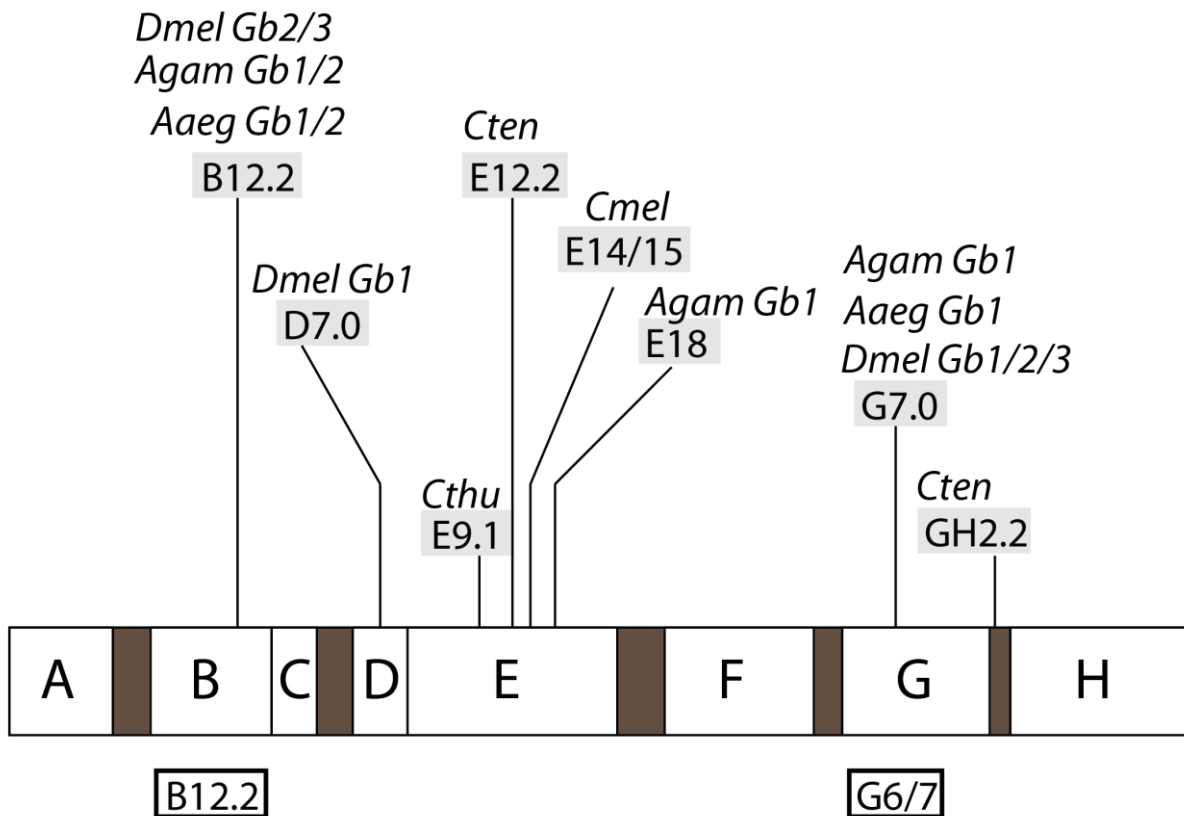


Abbildung 7 **Intronpositionen der Hämoglobin-Gene verschiedener Dipteren-Arten.** Die Intronpositionen der Arten *A. aegypti*, *A. gambiae*, *C. melanotus*, *C. tentans*, *C. thummi* und *D. melanogaster* sind abgebildet. Die Positionen beziehen sich auf die Helices der entsprechenden Proteine. Intron B12.2 beispielsweise ist in Helix B, Kodon 12 zwischen Nukleotid 2 und 3 lokalisiert. Die Positionen B12.2 und G6/7 sind konserviert in Vertebraten, Invertebraten und Pflanzen. Abb. verändert nach (Friedl 2001)

Die Erhebung zusätzlicher Daten in anderen Spezies wie z. B. *P. olivacea* wäre hilfreich zur Absicherung dieser Hypothese.

Die Histon-Gene

Die Histon-Gene sind ubiquitär in Eukaryoten vorhanden, ihre Produkte bilden eine Familie von fünf kleinen basischen DNA-bindenden Proteinen. Die hochkonservierten Sequenzen der Histon-Gene sind meist tandem-repetitiv in Gengruppen organisiert. Jede enthält je eine Kopie von jedem der fünf Histon-Gene (H1, H2A, H2B, H3, H4) (Isenberg 1979), (Maxson et al. 1983), (Hentschel und Birnstiel 1981), (Wells und McBride 1989). Histone sind Multidomänenproteine (Böhm und Crane-Robinson 1984), die DNA binden und mit anderen Histonen interagieren können. Hinsichtlich der Arthropoden wurden die Histon-Gene zuerst in *D. melanogaster* und *D. hydei* analysiert (Lifton et al. 1978), (Matsuo und Yamazaki 1989), (Kremer und Hennig 1990). Die Lokalisation der Histon-Gene zeigte für *D. melanogaster* einen Locus (Pardue et al. 1977), (Lifton et al. 1978), (Felger und Pinsker 1987). In anderen *Drosophila*-Arten wurden bis zu vier Histon-Loci gefunden (Steinemann et al. 1984). Die Histon-Cluster einiger Chironomiden-Arten sind bereits bekannt und lokalisiert auf Chromosom Arm D (Hankeln und Schmidt 1991), (Hankeln et al. 1993). In *C. melanesiens* und in *C. pallidivittatus* wurden sechs Histon-Loci identifiziert. In den in dieser Arbeit untersuchten Arten *C. thummi*, *C. piger*, *C. luridus*, *C. tentans* und *C. annularius* wurden zuvor fünf Histon-Loci nachgewiesen. Im Gegensatz zu *D. melanogaster* zeigt *C. thummi* verschiedene Varianten von H1-Genen (Hoyer-Fender und Grossbach 1988). Die Histon-Gene von *P. olivacea* wurden ebenfalls von Hankeln et al. durch FISH lokalisiert (Hankeln et al. 1993). Über die Sequenz und Organisation ist jedoch bisher nichts bekannt. Daher soll in der vorliegenden Arbeit die Struktur, Sequenz und die genomische Organisation der Histon-Gene von *P. olivacea* aufgeklärt werden.

1.2.3 Gene und genregulatorische Sequenzen

Das Genom besteht neben den hoch- bzw. mittelrepetitiven DNA-Sequenzen noch aus *single*- bzw. *low-copy*-DNA-Sequenzen (Abbildung 4). Dabei handelt es sich nicht nur um *single-copy*-Gene sondern auch u. a. um regulatorische Bereiche, miRNA und Pseudogene. Vertreter für *single (low) copy*-Gene sind z. B. die Gene aus der Geschlechtsbestimmungskaskade von *D. melanogaster*.

In Eukaryoten ist die Entwicklung in unterschiedliche Geschlechter weit verbreitet. Trotz dieses allgemeinen Prinzip sind die Mechanismen, welche die Geschlechtsbestimmung und

Geschlechterentwicklung verursachen, auch in nahverwandten Arten sehr verschieden (Bull und others 1983), (Graves, Jennifer A Marshall 2008). Dabei gibt es lediglich über eine kleine Anzahl an Taxa umfangreiches Wissen. Die Diversität in den Klassen der Säuger und Vögel ist eher invariabel, wogegen die der Fische und Insekten sehr verschiedene Mechanismen aufweisen (Devlin und Nagahama 2002), (Sánchez R. Lucas und others 2008). Die Geschlechtsbestimmung in Insekten ist charakterisiert durch eine Genkaskade, deren weiter unten stehende Gene besser konserviert zu sein scheinen. Am Anfang der Kaskade gibt es ein weniger stark konserviertes Gen *transformer (tra)*. *Tra* ist nachgewiesen in den Ordnungen Hymenoptera, Coleoptera, Trichoptera und Diptera, in anderen Ordnungen wie Lepidoptera und Strepsiptera jedoch nicht (Geuverink und Beukeboom 2014). Das terminale Gen der Geschlechtsbestimmungskaskade ist in allen bisher untersuchten Insekten das Gen *doublesex (dsx)* (Sánchez R. Lucas und others 2008), (Gempe und Beye 2011).

dsx

Im Jahr 1965 wurde eine rezessive Mutation in *D. melanogaster* beschrieben, die in Individuen mit genetisch festgelegtem Geschlecht zur Entwicklung von zwitterigen Individuen führte. Diese Mutation bzw. das dazugehörige Gen wurde als *dsx* bezeichnet (Hildreth 1965). Die molekulare Analyse des Gens offenbarte seine bifunktionelle Eigenschaft und Rolle in der somatischen Geschlechtsbestimmung. Diese wird durch geschlechtsspezifisch-alternatives Spleißen der *dsx*-Transkripte erreicht, da dabei geschlechtsspezifische Proteine entstehen (Burtis und Baker 1989), (Baker und Wolfner 1988). Etwa zur gleichen Zeit wurde in *Caenorhabditis elegans (C. elegans)* das Gen *male abnormal-3 (mab-3)* identifiziert, welches für Männchen-spezifische Funktionen benötigt wird (Shen und Hodgkin 1988). Beide Gene enthalten eine DNA-bindende Domäne, die nach *dsx/mab-3* als DM-Domäne benannt wurde. Daraufhin wurden weitere Proteine aus Säugern und anderen Vertebraten, welche die DM-Domänen besitzen, identifiziert. Alle Invertebraten- und Vertebraten-Genome enthalten mehrere DM-Domänen-Proteine, von denen nicht alle an der Geschlechtsbestimmung beteiligt sind (zusammengefasst in (Hong et al. 2007)). Dennoch scheint diese Familie von *dsx/mab-3-related genes (Dmrt genes)* in allen untersuchten Fällen an der Geschlechtsbestimmung beteiligt zu sein. Das *dsx*-Gen aus *D. melanogaster* und seine Orthologen besitzen zwei Oligomerisierungsdomänen (Abbildung 8), eine geschlechtsunabhängige DM-Domäne (in allen *Dmrt*-Genen) und eine geschlechtsspezifische

OD2-Domäne (nur in *dsx* und Orthologen) (An et al. 1996), (Erdman et al. 1996). Die DM-Domäne setzt sich aus einer DNA-bindenden Domäne (DBD) und einer Oligomerisierungsdomäne (OD1) zusammen. Die OD2 Domäne hat einen geschlechtsunabhängigen N-Terminus und einen geschlechtsspezifischen C-Terminus, resultierend aus dem alternativen Spleißen der *dsx*-Transkripte.

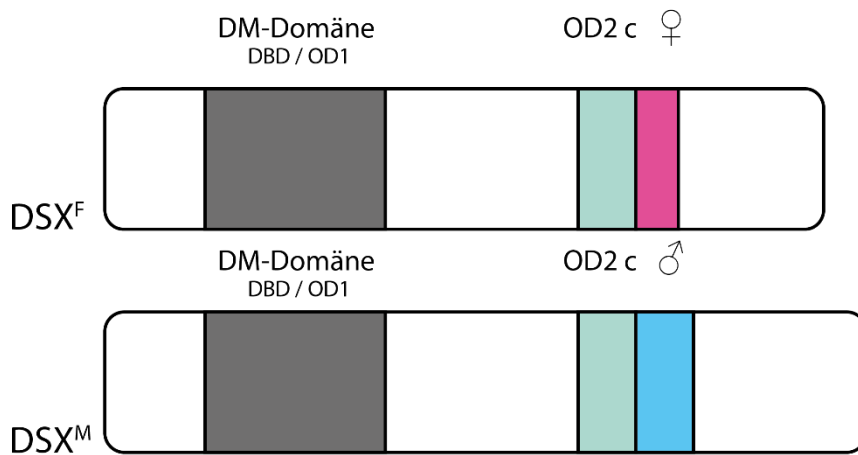


Abbildung 8 Schema des DSX^F bzw. ^M Proteins aus *D. melanogaster* mit seinen funktionellen Proteindomänen. Die DM-Domäne (grau) und die OD2 (common: hellgrün; Weibchen-spezi.: rosa, Männchen-spezi.: blau) sind eingezeichnet.

Somit scheint DSX über die Klasse der Insekten hinaus eine wichtige Rolle in der Geschlechtsbestimmung von *C. elegans* (*mab-3*) oder in Vertebraten (wie *Dmrt1* im Mensch) (Shen und Hodgkin 1988), (Raymond et al. 1998) zu spielen. Trotz der rasch evolvierenden Geschlechtsbestimmungsmechanismen in selbst nah-verwandten Arten sind die DM-Faktoren stark konserviert. Dies wurde zuerst von Wilkins beobachtet und als *bottom-up*-Hypothese bezeichnet, da die Gene am Ende der Kaskade konservierter zu sein scheinen als das primäre Signal oder daran nachgeschaltete Gene (Wilkins 1995).

Es gibt Ausnahmen, in denen DSX durch einen anderen Faktor ersetzt wird. Im Zentralnervensystem übernimmt seine Rolle entweder der Transkriptionsfaktor (TF) FRU, welcher von dem Gen *fruitless* (*fru*) kodiert wird (Ryner et al. 1996) oder der TF DSF, welcher von dem Gen *dissatisfaction* (*dsf*) kodiert wird (Finley et al. 1997).

1.3 Zielsetzung

Die Gattung *Prodiamesa* ist phylogenetisch nahe verwandt mit der Gattung *Chironomus* (Abbildung 1). Die Gattung *Chironomus* ist seit über 130 Jahren für ihre in Speicheldrüsen vorkommenden Riesenchromosomen bekannt (Balbiani 1881). Seither gibt es eine Vielzahl von wissenschaftlichen Arbeiten die sich mit den *Chironomus*-Spezies befassen. Das (genetische) Wissen über die Art *P. olivacea* beschränkt sich lediglich auf eine geringe Anzahl an wissenschaftlichen Studien (vgl. (Zacharias 1979), (Zacharias 1981), (Michailova 1977), (Petrova und Zhironov 2014)). Dabei weisen die Larven durch ihre weiß-gelbliche Farbe und durch das Vorkommen eines morphologisch identifizierbaren Y-Chromosoms eine Besonderheit zu den *Chironomus*-Arten *C. thummi* und *C. piger* auf (Leathers 1922), (Zacharias 1981).

Im Rahmen dieser Arbeit soll sowohl Genom- als auch Transkriptomanalyse von *P. olivacea* mittels *next generation sequencing* erfolgen. Die Analyse ausgewählter Abschnitte des *P. olivacea*-Genoms steht dabei im Vordergrund. Neben hochrepetitiven sollen mittelrepetitiven Sequenzen wie Genfamilien (Histon-Gene, Hämoglobin-Gene) mittels bereits bekannter Sequenzen aus der Gattung *Chironomus* per BLAST identifiziert und detailliert untersucht werden. Darüber hinaus soll eine Lokalisation der Gene bzw. repetitiven Elemente per FISH an Speicheldrüsenchromosomen von *P. olivacea* erfolgen. Von besonderem Interesse sind dabei die Hämoglobin-Gene, da *P. olivacea*-Larven im Gegensatz zu *Chironomus* augenscheinlich (weiß-gelbliche Farbe) keine signifikanten Mengen an Hb besitzen. Mittels bioinformatischer Analyse soll versucht werden, Erkenntnisse über die Funktionalität möglicher Hämoglobin-Gene bei *P. olivacea* zu gewinnen. Ein weiteres interessantes Thema ist die Geschlechtsbestimmung bei Chironomiden. Deshalb sollen die bei anderen Insekten in die Geschlechtsbestimmung involvierten Gene identifiziert und bioinformatisch analysiert werden. Eine mögliche Rolle für die Geschlechtsbestimmung bei *P. olivacea* soll über Sequenzhomologie, Expressionsstärken und mögliche geschlechts-spezifische Expression überprüft werden. Als Vertreter für *single-* bzw. *low-copy*-Sequenzen sollen homologe Gene aus der Geschlechtsbestimmungskaskade von *D. melanogaster* und homologe Gene aus der SDR von *C. thummi* per BLAST identifiziert werden. Diese homologen Gene aus *P. olivacea* sollen ebenfalls charakterisiert und mittels FISH an Speicheldrüsenchromosomen lokalisiert

werden. Die Ergebnisse des Larven-Transkriptoms sollen für die Bestimmung der Genstruktur, wie auch Bestimmung von Spleißvarianten z.B. bei *P. olivacea*-Männchen und -Weibchen aufzeigen. Die *in-situ*-lokalisierten Gene bzw. Sequenzen soll die Grundlage einer Chromosomenkarte von *P. olivacea* sein und somit eine Identifizierung der Chromosomenarme auf der Basis des Referenzkaryotyps von *C. thummi* erfolgen.

2 Material und Methoden

2.1 Versuchstiere

Für die vorliegenden Analysen wurden Zuckmückenlarven der Arten *P. olivacea*, *C. thummi* und *C. piger* (Chironomidae, Diptera) verwendet. Die Gelege und Larven der Spezies *C. thummi* und *C. piger* stammten aus der permanenten Laborzucht des Instituts für Molekulargenetik, gentechnologische Sicherheitsforschung und Beratung (IMSB) der Johannes Gutenberg-Universität Mainz. Die Gelege und Larven der Spezies *P. olivacea* wurden im Zeitraum von März 2013 bis April 2015 in Osthofen aus dem Seebach (N49° 42' 36.133" E8° 20' 21.574" & N49° 42' 25.744" E8° 20' 10.993") entnommen. Die Aufzucht der Larven bis zum L4- Stadium erfolgte in belüfteten Boxen bei 11°C. Als Futter diente Brennesselpulver und Zellstoff. Die Spezies wurde anhand ihres Chromosomensatzes und/oder per Art-spezifischer PCR bestimmt.

2.2 Speziesbestimmung *P. olivacea*

Die Spezies wurde anhand des Karyotyp $n=3$ und Morphologie der Chromosomen (Zacharias 1979) und bzw. oder per Art-spezifischer Primer (Cranston et al. 2012), *GenBank Acession no.* GU356718.1 bestimmt. Eine morphologische Bestimmung der Larven erfolgte über die verlängerte Ventromentalplatte mit charakteristischer Zeichnung und den darunterliegenden Tasthaaren (siehe Abbildung 9)



Abbildung 9 *Kopfkapsel einer P. olivacea -Larve*. Frontalansicht der Ventromentalplatte mit Zeichnung (schwarz) und Tasthaaren (anliegend, deshalb schlecht zu erkennen).

2.3 Geschlechtsbestimmung bei Chironomiden-Larven

Die Geschlechtsbestimmung im Larvenstadium kann anhand der Morphologie ihrer Genitalimaginalscheiben erfolgen (Wülker und Götz 1968). Die weiblichen Larven besitzen 2 Strukturen im achten bzw. neunten Segment, wohingegen die Männchen nur eine Genitalimaginalscheibe im neunten Segment besitzen (Abbildung 10). Dabei wurden die Genitalimaginalscheiben im L4- Stadium identifiziert. Die Larven wurden dabei auf einem Objektträger mit der Ventralseite nach oben und gestreckten Nachschiebern am Abdomen fixiert. Zur Bestimmung diente ein Mikroskop der Fa. Zeiss (Axiostar plus; Göttingen, Deutschland). Die Entwicklung der Genitalimaginalscheiben ist im L4-Stadium in 9 Phasen unterteilt. Die eindeutige Zuordnung kann ab Phase 3 erfolgen. Abbildung 10-A zeigt die weibliche *P. olivacea*-Genitalimaginalscheibe in Segment 9 in Phase 5 Abbildung 10-B zeigt die männliche *P. olivacea*-Genitalimaginalscheibe in Segment 9 in Phase 7.

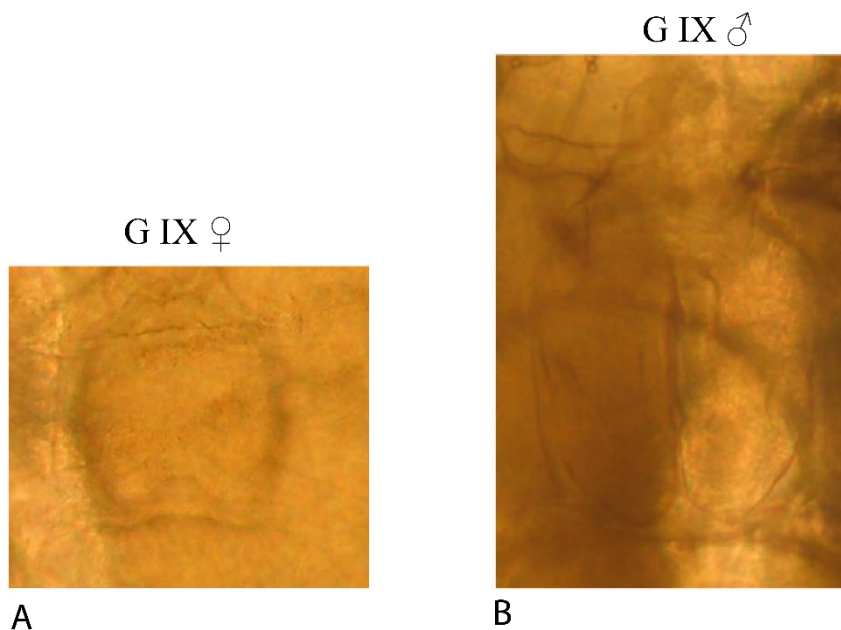


Abbildung 10 *Genitalimaginalscheiben von P. olivacea im Vergleich zwischen Weibchen und Männchen.* In A ist die weibliche *P. olivacea*-Genitalimaginalscheibe im 9ten Segment zu abgebildet. Die Abbildung der zweiten weiblichen Genitalimaginalscheibe im Segment 8 kann aufgrund des nicht ausreichenden Kontrasts nicht abgebildet werden. In B ist die Genitalimaginalscheibe eines *P. olivacea*-Männchen im Segment 9 dargestellt.

Die Morphologie der *P. olivacea*-Genitalimaginalscheiben weicht zu der für *Chironomus* beschriebenen Morphologie (Wülker und Götz 1968) leicht ab, kann jedoch eindeutig bestimmt werden.

Molekularbiologische Methoden

2.4 Isolation von Nukleinsäuren

2.4.1 Isolation von genomischer DNA aus Chironomiden-Larven

Die Isolation genomischer DNA aus Chironomiden-Larven erfolgte nach einem verändertem Protokoll von (Schmidt et al. 1980). Dafür wurden etwa 0,3 bis 2 g Larven (*C. thummi*, *C. piger* und *P. olivacea*, ♂ und ♀ getrennt) mit 4 ml Homogenisierungspuffer (+ 1 % (v/v) Triton-X-100) in einem Glas-Teflon-Homogenisator auf Eis mechanisch aufgeschlossen. Das Homogenat wurde durch eine Gaze (250 µm) filtriert und die Gaze nach Bedarf mit Homogenisierungspuffer nachgespült. Das Filtrat wurde in einem 50 ml Reaktionsgefäß aufgefangen und 15 min bei 3500 g und 4°C zentrifugiert (Kühlzentrifuge 3K30, Fa. Sigma, Osterode am Harz, Deutschland). Das Kernpellet wurde in 5 ml Homogenisierungspuffer resuspendiert und abermals zentrifugiert. Der Waschschrift wurde 2 bis 3-mal wiederholt. Das Pellet wurde in 2 ml Homogenisierungspuffer resuspendiert. Die Suspension wurde mit $\frac{1}{10}$ Vol. SDS (10 %), $\frac{1}{10}$ 10x Dialysepuffer und einer Spatelspitze Proteinase K 30 bis 60 min bei 60°C inkubiert. Zum Entfernen der Proteine aus der Lösung erfolgten mehrere PCI- und CI-Extraktionen. Zu der DNA-haltigen Lösung wurde $\frac{1}{10}$ Vol. 10x Dialysepuffer zugegeben und mit dem 2,5-fachen Vol. absolutem eiskaltem Ethanol überschichtet. Die DNA, welche an der Grenzschicht ausfällt wurde mittels Glasstab aufgewickelt, in 70%igem EtOH gewaschen und luftgetrocknet. Die DNA wurde in 500 µl HPLC-H₂O abgewickelt. Die genomische DNA wurde bei 4°C gelagert.

2.4.2 Isolation von Einzellarven- DNA

Eine Larve mittels Pistill in einem 1,5 ml Eppendorf-Gefäß in 50 µl Homogenisierungspuffer aufnehmen und zusammen mit 5 µl 10 % Triton-X-100 zerkleinern. Dies dann 15 min bei 4 °C und 5000 Upm zentrifugieren. Den Überstand abziehen und verwerfen. Das Pellet in 50 µl Homogenisierungspuffer resuspendieren. Diesen Schritt bis zu dreimal wiederholen. Eine Spatelspitze Proteinase K in 110 µl 10 x Dialysepuffer lösen. Je Ansatz 5 µl Proteinase K -Dialysepuffer-Gemisch und 5 µl 10 % SDS dazugeben. Die Ansätze 30 min bei 60 °C inkubieren. Nach Zugabe von 6 µl gesättigtem Tris (pH 8,5), 16 µl Natriumperchlorat

und 82 µl PCI den Ansatz 10x invertieren und 5 min bei RT und 5000 Upm zentrifugieren. Die PCI-Extraktion mit der oberen, wässrigen Phase zweimal wiederholen, dabei abgeschnittene Piptettenspitzen verwenden. Die wässrige Phase nach der letzten PCI mit gleichem Vol. CI versetzen 10x invertieren, 5 min bei RT und 5000 Upm zentrifugieren. Zur oberen, wässrigen Phase 1/10 Vol. 10x Dialysepuffer und 2 Vol. eiskalter Ethanol (abs.) geben. Die Ansätze mehrmals invertieren und mind. 30 min bei -20°C inkubieren. Die Ansätze mind. 30 min bei 4 °C und 14000 Upm zentrifugieren. Den Überstand abnehmen und verwerfen. Das Pellet mit 70 % Ethanol (RT) mittels 10 min Zentrifugation bei RT und 14000 Upm waschen. Den Überstand abziehen und das Pellet trocknen lassen. Nach dem Trocknen das Pellet in 30 µl HPLC-H₂O lösen.

2.4.3 Isolation von Plasmid- DNA

Die Plasmid-Präparation (pUC-Klone aus pUC-Bibliothek oder pGEM T-easy-Klone aus der Klonierung von PCR-Produkten) erfolgte aus ü. N.- Kulturen mittels des peqGOLD Plasmid Miniprep Kit I (Peqlab, Erlangen, Deutschland). Abweichend zum Skript wurde die isolierte Plasmid-DNA in 30 µl HPLC-H₂O eluiert.

2.4.4 Präparation von Gesamt-RNA (Guanidinthiocyanat (GTC)- Methode) (Chomczynski und Sacchi 1987)

Je 0,1 g männliche und weibliche *P. olivacea*-Larven wurden in einem Keramik-Mörser in flüssigem Stickstoff gehalten gemörsert und anschließend in 1000 µl GTC-Arbeitslösung aufgenommen. 1/10 Vol. 2 M Natriumacetat hinzugeben und invertieren. Zur Lösung ein Volumen PCI geben, invertieren und 5 min auf Eis stehen lassen. Die Lösung dann 15 min bei 4 °C bei 3500 Upm zentrifugieren. Diesen Schritt je nach Bedarf mehrmals wiederholen, sowie 2 anschließende CI-Extraktionen. Es erfolgte eine Standardfällung mit 100 %igem EtOH. Das Pellet wurde in Nuklease-freiem H₂O gelöst. Zur Lagerung wurde die RNA bei -80 °C aufbewahrt.

2.4.5 Präparation von Gesamt-RNA aus Einzellarven (GTC-Methode)

(Chomczynski und Sacchi 1987)

Zur RNA-Präparation aus Einzellarven wurden Einweg-Pistill der Fa. Argos passend für ein 1,5 ml Eppendorfggefäß verwendet. Diese wurden zunächst bei -20°C vorgekühlt. Die 1,5 ml Eppendorfggefäße mit jeweils einer Larve wurden mittels flüssigen Stickstoff gekühlt und die Larven je mit einem neuen Einweg-Pistill auf Eis gemörsert. Nach dem die Larven vollständig zerkleinert waren, wurde der Pistill mit $300\ \mu\text{l}$ GTC-Arbeitslösung abgespült. Zur GTC-Larven-Lösung wurde $\frac{1}{10}$ Vol. 2 M Natriumacetat pipettiert. Danach erfolgten 3 PCI-Extraktionen und 2 CI-Extraktionen wie unter 2.4.4 Präparation von Gesamt-RNA (Guanidinthiocyanat (GTC)- Methode) beschrieben. Der wässrige Überstand wurde mit 2 Vol. eiskaltem EtOH abs. versetzt. Nach Invertieren der Lösung wurde diese mindestens 30 min bei -20°C gefällt. Die Lösung wurde dann bei 4°C mindestens 30 min zentrifugiert. Der Überstand wurde abgezogen und verworfen, das Pellet mit $100\ \mu\text{l}$ 70%igem EtOH gewaschen in dem es erneut bei 4°C 15 min zentrifugiert wird. Der Überstand wird wieder verworfen und das Pellet auf Eis luftgetrocknet. Das getrocknete Pellet wird in $50\ \mu\text{l}$ Nuklease-freiem H_2O gelöst. Zur Lagerung wurde die RNA bei -80°C aufbewahrt.

2.5 Konzentrationsbestimmung von Nukleinsäuren

2.5.1 NanoDrop®

Die Konzentration der DNA wurde mit dem NanoDrop® ND-1000 (PeqLab, Erlangen, Deutschland) photometrisch bestimmt. Hierfür wurde $1\ \mu\text{l}$ DNA-Lösung eingesetzt.

2.5.2 QuBit™

Die fluorometrische Konzentrationsbestimmung mittels Qubit™ (Thermo Fisher Scientific, Waltham, MA, USA) erfolgte über den Qubit™ dsDNA HS Assay. Dieser eignet sich um die Konzentration von DNA-Lösungen mit einer Konzentration von $10\ \text{pg}/\mu\text{l}$ bis $100\ \text{ng}/\mu\text{l}$ zu messen. Das Assay erfolgte nach Herstellerangaben.

2.5.3 Agilent 2100 Bioanalyzer

Der Agilent 2100 Bioanalyzer (Agilent, Santa Clara, Kalifornien, USA) ist eine weitere Methode der fluorometrischen Konzentrationsbestimmung unter anderem von nukleinsäurehaltigen Lösungen. Für die Messungen von RNA-Lösungen wurde der Agilent RNA 6000 Nano gesamt RNA Assay verwendet. Dieser bestimmt in einem Bereich von 25 bis 250 ng/ μ l zuverlässig die Konzentration einer RNA-Lösung.

2.6 Aufreinigung von Nukleinsäuren

2.6.1 Phenol-Chloroform-Extraktion

Eine Extraktion mit Phenol, Chloroform und Isoamylalkohol (PCI; 25:24:1, v:v:v) erfolgte um Proteinverunreinigungen aus Nukleinsäurelösungen zu entfernen. Dafür wurde 1 Vol. PCI zur Nukleinsäurelösung gegeben und die dadurch entstandenen Phasen durch 5-minütiges Invertieren durchmischt. Die Phasentrennung wurde mittels 15-minütiger Zentrifugation bei RT und 3500 g (15-bzw. 50-ml Falcon) bzw. 16000 g (1,5 ml-Eppendorf-Reaktionsgefäß) beschleunigt. Anschließend wurde die obere nukleinsäurehaltige Phase in ein neues Reaktionsgefäß überführt und eine erneute PCI-Extraktion durchgeführt. Je nach Grad der Verunreinigung wurden 2 bis 5 PCI-Extraktionen durchgeführt. Auf die letzte PCI-Extraktion folgte, zum Entfernen der Phenolreste, eine Extraktion mit Chloroform und Isoamylalkohol (CI; 24:1, v:v).

2.6.2 Fällern von Nukleinsäuren

Zur Erhöhung der DNA-Konzentration einer Lösung oder zum Entfernen von Salzen aus der Lösung, wurde die DNA mit $\frac{1}{10}$ Vol. 10x Dialysepuffer und 2 bis 2,5 Vol. eiskaltem EtOH abs. bei -20 °C mind. 30 min gefällt. Das Pelletieren der DNA erfolgte durch zentrifugieren (4 °C, 30 min, 16000 x g). Das Pellet wurde mit 70 %igem EtOH gewaschen. Das Pellet wurde luftgetrocknet und in HPLC-H₂O gelöst.

2.6.3 Entfernen von Verunreinigungen

RNase-Verdau

Die isolierte DNA wurde, um RNA-Kontaminationen zu entfernen, mit 1 µl RNase A für 30 min bei 37 °C inkubiert. Die Lagerung der DNA erfolgte bei 4 °C.

DNase-Verdau

Die isolierte RNA wurde, um DNA-Kontaminationen zu entfernen in einem 100 µl Reaktionsansatz mit DNase I verdaut. In den Ansatz wurden 10 µl 10 x DNase I-Puffer und 1 µl DNase I (NEB, Frankfurt am Main, Deutschland) gegeben und 20 bis 30 min bei 37 °C inkubiert. Zur Inaktivierung der DNase I wurde zum Ansatz 1 µl EDTA (0,5 M) gegeben und die Lösung 10 min bei 75 °C inkubiert. Die Lagerung der RNA erfolgte bei – 80 °C.

2.7 Elektrophoretische Auftrennung von DNA

Je nach Anwendung wurden unterschiedliche Gelelektrophorese-Systeme zur Auftrennung von Nukleinsäuren verwendet. Je nach Größe der Nukleinsäure-Fragmente wurden Agarosegele mit verschiedenen Prozentigkeiten zwischen 0,8 % bis 2 % (1 x TBE, w:v) verwendet. Als Größenmarker wurden die xx gezeigten Molekulargewichtsstandards (Fermentas, St. Leon-Rot, Deutschland) eingesetzt. Die Proben wurden um ihre Dichte zu erhöhen mit einem Probenpuffer versetzt. Gleichzeitig diente der Probenpuffer als Referenz für die Laufgeschwindigkeit der Proben während der gelelektrophoretischen Auftrennung. Den Proben wurde 1/6 Vol. Bromphenolblau-DNA-Ladepuffer (BPB) zugegeben. Die Gele wurden je nach Dicke in einer Ethidiumbromid(EtBr)-Färbelösung (5 µg/ml in 1x E-Puffer) inkubiert und dann kurz gewässert. Über die Anregung des in die DNA interkalierenden Farbstoffs EtBr kann die DNA bei 312 nm auf dem Transilluminator (Bachofer, Reutlingen, Deutschland) sichtbar gemacht werden. Die Dokumentation erfolgte über das Carestream Gel Logic 112 System (Carestream Health, Rochester, USA). Die Gelbilder wurden farblich invertiert.

2.7.1 Testgele

Zur Analyse von DNA-Lösungen wurden diese auf vertikalen GENterphorese™ (GENterprise, Mainz, Deutschland) gelelektrophoretisch aufgetrennt. Die Auftrennung in den 7 cm x 6,5 cm x 0,1 cm großen Gelen erfolgte in einem TBE-Puffersystem bei 100 bis 150 mA für 25 bis 45 min. Zu den Proben wurde für die oben genannten Marker 2 µl aufgetragen. Die kleinen Gele wurden für ca. 3 bis 5 min in der EtBr-Färbelösung inkubiert.

2.7.2 Präparative bzw. analytische Gele

Um DNA-Fragmente sauber aufzutrennen, erfolgte eine vertikale Gelelektrophorese in großen Gelkammern (18 cm x 13 cm x 0,4 cm). Die Auftrennung wurde in einem TBE-Puffersystem bei 30 mA für 3 bis 4 Stunden durchgeführt. Für die große vertikale Gelelektrophorese wurden 5 µl Molekulargewichtsstandard eingesetzt. Das Färben des Gels in der EtBr-Färbelösung erfolgte für 5 bis 10 min.

2.8 Wiedergewinnung von DNA-Fragmenten aus Agarosegelen

2.8.1 Elektroelution

Für die Wiedergewinnung größerer DNA-Fragmente wurde die Elektroelution angewendet. Die gewünschten DNA-Fragmente wurden aus dem Agarosegel ausgeschnitten und je ein 0,5 cm breites Gelstück mit 600 µl 0,5 x TBE-Puffer in einen Dialyseschlauch (Fa. Medicell International, London, UK) gegeben. Der Dialyseschlauch wurde mittels Kunststoff-Klammer verschlossen und in eine Elektrophoresekammer, welche mit 0,5 x TBE befüllt wurde, gebettet. Die Elektroelution erfolgte bei 100 mA für 1 bis 2 Stunden unter Eiskühlung. Die Umpolung der elektrischen Schaltung erfolgte 5 Minuten und eine Umpolung der ursprünglichen Richtung für 1 Minute. Die DNA-Lösung wurde aus dem Dialyseschlauch in ein Eppendorf-Gefäß überführt und der Schlauch mit 300 µl 0,5 x TBE-Puffer nachgespült und mit der DNA-Lösung vereint.

2.8.2 Wiedergewinnung mittels Silica-Membran

Die Wiedergewinnung kleinerer DNA-Fragmente bzw. PCR-Produkte aus Agarosegelen erfolgte mit dem GenElute™ Gel Extraction Kit (Sigma Aldrich, St.Louis, USA). Die Wiedergewinnung wurde nach den Angaben des Herstellers durchgeführt.

2.9 Klonierung von DNA

2.9.1 Restriktion genomischer DNA

Restriktion genomischer DNA

Als Grundlage der Southern Blots (2.11.2 Southern-Hybridisierung bzw. „Pirrotta“) wurde genomische DNA von *P. olivacea* (ungesext, Weibchen oder Männchen) mit unterschiedlichen Restriktionsenzymen geschnitten. Dabei wurden ~ 5 µg DNA mit 2 µl des jeweiligen Enzyms in einem Gesamtvolumen von 150 µl für 2 Stunden bei 37 °C restringiert. Der Ansatz wurde wie unter 2.6.2 Fällen von Nukleinsäuren beschrieben gefällt und wie oben beschrieben erneut verdaut und gefällt. Über ein Agarosetestgel wurde die vollständige Restriktion der DNA überprüft, indem $\frac{1}{10}$ Vol. des Restriktionsansatzes aufgetragen wurde.

Restriktion Plasmid-DNA

Zur Überprüfung potenziell positiver rekombinanter Plasmide (pGem™T Easy), wurden diese mit EcoRI restringiert. Potenziell rekombinante Plasmide auf pUC-Basis wurden in einem Doppelverdau mit EcoRI und HindIII geschnitten. Alle Enzyme wurden von der Fa. Thermo Fisher Scientific (Waltham, MA, USA) bezogen. Es wurden ~ 300 ng Plasmid-DNA in einem Volumen von 20 µl mit 1 µl Enzym für 2 Stunden bei 37 °C restringiert.

2.9.2 Ligation

Die PCR-Produkte (~ 100 ng) wurden in einem 10 µl-Ansatz in 5 ng pGem®-T Easy Vektor (Promega, Madison, USA) bei 18 °C ü. N. ligiert (5 U T4 DNA Ligase, Fermentas, St. Leon-Rot, Deutschland). Der Ligationsansatz wurde gefällt (2.6.2 Fällen von Nukleinsäuren) und in 10 µl HPLC-H₂O gelöst. 1-2 µl der Ligation wurde in elektrokompotent *E. coli*- Zellen (DH10B, Laborstamm) transformiert.

2.9.3 Transformation mittels Elektroporation

Die elektrokompetenten *E. coli*- Zellen (40 µl) werden mit 1-2 µl des rekombinanten Plasmids versetzt und bei RT 5 Minuten inkubiert. Die Suspension wurde in eine vorgekühlte Küvette (Gene Pulser® Cuvette, Spaltbreite: 0,1 cm) der Fa. Bio-Rad Laboratories (Hercules, USA) luftblasenfrei pipettiert. Die Transformation erfolgte im Elektroporator Micro Pulser™ (Bio-Rad Laboratories, Hercules, USA) bei 1,8 kV für 3 bis 5 ms. Die Zellsuspension wurde sofort in 1 ml vorgewärmtes L-Medium aufgenommen und für 45 bis 60 min bei 37 °C unter Schütteln in einem 15-ml-Reaktionsgefäß inkubiert. Auf vorgewärmte Agarplatten wurden 100 µl des Transformationsansatzes ausplattiert und für 16 bis 20 Stunden bei 37 °C bebrütet.

2.10 DNA-Markierungstechniken

2.10.1 Random Primed Oligo Labeling

Die einzelsträngige DNA-Matrize bindet Oligonukleotide zufälliger Sequenz (Feinberg und Vogelstein 1983), welche als Primer für das Klenow-Enzym (Fragment der DNA-Polymerase I aus *E. coli*) dienen. Die Elongation findet unter Einbau markierter Nukleotide statt. Diese Methode wurde hier für die Markierung mit Digoxigenin sowohl in der radioaktiven Markierung von [$\alpha^{32}\text{P}$]dATP als auch in der Fluoreszenzmarkierung eingesetzt.

Fluoreszenzmarkierung mit Digoxigenin

Die Markierung von DNA-Sonden mit Digoxigenin (DIG)-markiertem dUTP, welche für die Fluoreszenz-*in-situ*-Hybridisierung eingesetzt wurden, wurden mit dem DIG DNA Labeling Kit der Fa. Roche Diagnostics GmbH (Mannheim, Deutschland) durchgeführt. Die Menge der zu markierenden DNA betrug ca. 100 ng und wurde auf ein Volumen von 15 µl mit HPLC-H₂O aufgefüllt. Die DNA wurde 10 min bei 96 °C denaturiert und auf Eis abgekühlt. Die DNA-Lösung wurde mit 2 µl Hexanukleotid-Mix, 2 µl dNTP-Mix sowie 1 µl Klenow-Enzym versetzt. Die Markierungsreaktion erfolgte bei RT ü.N. Nach der Markierung erfolgte die Zugabe von 30 µl HPLC-H₂O, 50 µl 10 x SSC, 1 µl 1 % SDS sowie 4 µl heterologe *E. coli*-DNA („Carrier-DNA 50 µg/ml).

Radioaktive Markierung mit [$\alpha^{32}\text{P}$]-dATP

Die DNA (~ 200 ng) welche als Sonden dienen sollte wurden in 11 μl aufgenommen und 10 Minuten bei 96 °C denaturiert. Zur auf Eis inkubierten DNA wurden je 1 μl dTTP, dGTP und dCTP (je 0,5 mM), sowie 2 μl Hexanukleotidmix aus dem Random Primed DNA Labeling Kit der Fa. Roche Diagnostics GmbH (Mannheim, Deutschland) pipettiert. Im Isotopenlabor erfolgte die Zugabe von 5 μl [$\alpha^{32}\text{P}$]-dATP (10 $\mu\text{Ci}/\mu\text{l}$) und 1 μl Klenow-Enzym (2 U/ μl). Die Markierungsreaktion erfolgte für 2,5 Stunden bei RT. Die gefällte DNA wurde in 100 μl HPLC-H₂O gelöst.

2.10.2 Nick Translation

Die Markierung der DNA-Sonde mit Biotin für die FISH erfolgte über das BioNick™ Labeling System der Fa. Invitrogen (Karlsruhe, Deutschland). Zur Markierung von 1 μg DNA wurde diese auf 40 μl mit HPLC-H₂O aufgefüllt und 5 μl dNTP-Mix und 5 μl Enzym-Mix zugesetzt. Nach 2,5 h Inkubation bei RT wurden zur Markierungsreaktion 5 μl Stopp-Puffer gegeben, sowie 50 μl 10 x SSC, 1 μl 1 % SDS und 4 μl heterologe *E. coli*-DNA („Carrier-DNA 50 $\mu\text{g}/\text{ml}$).

2.11 Hybridisierungstechniken

2.11.1 Fluoreszenz-*in-situ*-Hybridisierung (FISH)

Chromosomenpräparation

Die FISH wurde an polytären Riesenchromosomen aus Speichdrüsen von Chironomiden-Larven welche im 4.ten Larval-Stadium sind durchgeführt. Die Chironomiden-Larve wurde dekaptiert und die Speicheldrüsen treten aus. Diese wurden separiert und für 5 bis 20 Minuten in 50 %iger Essigsäure fixiert. Die fixierten Speicheldrüsen wurden mittels Deckgläschen gequetscht und mindestens 20 Minuten auf Kohlendäureschnee mit dem Objekt nach unten inkubiert. Das Deckgläschen wurde mittels Skalpell vom Objektträger entfernt. Die Objektträger wurden in 100 %igem Isopropanol bei – 20 °C gelagert.

Denaturierung der Chromosomenpräparate

Zur Vorbereitung der Präparate wurden diese je 5 Minuten in 70 %, 50 % und 30 % Isopropanol sowie in 0,1 x SSC überführt. Die Stabilisierung der Chromosomen erfolgte dann über eine Inkubation der Präparate in 2 x SSC für 30 Minuten bei 70 °C. Die Präparate wurden nach dem Abkühlen auf RT abermals in 0,1 x SSC überführt und für 1 Minute in 0,05 N NaOH denaturiert. Die Neutralisierung erfolgte in eiskaltem 0,1 x SSC und wurden die Alkoholreihe aufwärts geführt bis zu 100 % Isopropanol und erneut bei -20 °C gelagert oder luft-getrocknet zur Hybridisierung verwendet.

Hybridisierung der Präparate

Die luft-getrockneten Präparate wurden für Einzel-Hybridisierungen mit 8 µl, bei Doppel-Hybridisierungen mit je 4 µl einer Fluoreszenz-markierten Sonde bedeckt.

Die Sonden wurden zuvor bei 96 °C für 10 Minuten denaturiert und auf Eis abgekühlt.

Die Präparate mit den Sonden wurden mittels Deckgläschen und Flüssiggummi abgedichtet und für mindestens 6 Stunden in einer feuchten Kammer bei 55 °C hybridisiert. Nach der Hybridisierung wurden die Präparate 5 Minuten in 2 x SSC und 10 Minuten in 1 x PBS gewaschen.

Nachweis der fluoreszenz-markierten DNA

Der Nachweis der fluoreszenz-markierten DNA erfolgte über Fluoresceinisothiocyanat (FITC)-markierter Anti-DIG-Antikörper (1:30 verdünnt mit PBS/BSA) bzw. Rhodaminisothiocyanat (TRITC)-markiertes Avidin (1:75 verdünnt mit PBS/BSA) oder vice versa. Je nach Hybridisierungsart wurden 8 µl eines Nachweis-Reagenz oder je 4 µl zweier Nachweis-Reagenzien dazugegeben.

Die Präparate wurden erneut mit Deckgläschen und Flüssiggummi abgedichtet und die Nachweisreaktion erfolgte für 3 Stunden bei 37 °C in einer feuchten Kammer. Die Präparate wurden wieder für 10 Minuten in 1 x PBS gewaschen. Zum Schutz vor zu schnellem Ausbleichen wurde die Präparate in 8 µl Glyzerin-Para-Phenylendiamin-Mix (*Anti-fading-Mix*) eingebettet und mit Lack versiegelt.

Die Auswertung wurde mit Hilfe des Fluoreszenzmikroskop Olympus BX61 der Fa. Olympus (Hamburg, Deutschland) durchgeführt. Die Dokumentation erfolgte mit der Software CellF und der Kamera Color View III der Fa. Soft Imaging System GmbH (Münster, Deutschland).

2.11.2 Southern-Hybridisierung bzw. „Pirrotta“

Southern-Blot (Southern 1975)

Die DNA-DNA-Hybridisierung (Southern 1975) dient der Detektion von restringierter DNA welche zuvor gelelektrophoretisch aufgetrennt und durch Diffusionsprozesse auf eine Nitrozellulosemembran transferiert wurde. Es wurden je nach Anwendung uni- bzw. bidirektionale Blots durchgeführt. Das Gel wurde mit EtBr gefärbt und die Markerbanden mit einem Spatel durchstoßen.

Die Denaturierung der DNA im Gel erfolgte durch 15 minütige Inkubation des Gels in Denaturierungslösung (0,5 M NaOH/1,5 M NaCl). Danach wird das Gel für 25 Minuten in Neutralisierungslösung (3 M NaCl/0,5 M Tris pH 7,0) neutralisiert. Das Gel wurde zwischen 2 bzw. auf eine Nitrozellulosemembran luftblasenfrei gelegt. Zum Verstärken der Diffusion wurde der bidirektionale Blot mit Filterpapier und Handtuchpapier umgeben und beim unidirektionalen Blot eine Salzbrücke mit 20 x SSC erzeugt. Nach dem ü.N. erfolgtem Transfer wurden die durch Anregung mit UV-Licht sichtbar gemachten Markerbanden mit einem Fettstift markiert. Die DNA wurde auf der Membran fixiert dazu wurde die Membran 2 Stunden bei 80 °C inkubiert. Die Präinkubation der Blots fand bei 60 °C für 3 bis 4 Stunden in 2 x PM/6 x SSC/1 % SDS (Denhardt 1966) statt. Den, die wie in

Radioaktive Markierung mit [$\alpha^{32}\text{P}$]-dATP beschrieben, markierten Sonden wurden 100 μl Carrier-DNA (*E. coli*, 50 $\mu\text{g/ml}$) zugegeben. Die zuvor bei 96 °C denaturierten Sonden wurde in 20 ml Hybridisierungslösung aufgenommen und mit der Nitrozellulosemembran in Glas-Hybridisierungsröhrchen überführt. Die Hybridisierung erfolgte für 18 bis 20 Stunden bei 60 °C. Die Nitrozellulosemembran wurde mehrmals mit 2 x SSC bei 60 °C gewaschen. Die getrocknete Nitrozellulosemembran wurde auf Filterpapier geklebt und in Folie verpackt. Die Exposition eines für Röntgenstrahlen empfindlichen Films der Fa. Fuji Medical X-Ray (Japan) erfolgte in einer Expositionskassette mit Verstärkerfolie bei -80 °C.

„Pirrotta“-Blot

Der „Pirrotta“-Blot dient zur Identifikation von Sonden mit repetitiven Elementen, er wird auch als reverser Southern Blot bezeichnet. Im Gegensatz zum Southern Blot wurde als Sonde radioaktiv markierte genomische DNA zur Hybridisierung der Nitrozellulosemembran eingesetzt. Der Rest der Durchführung entspricht der des oben beschriebenen Southern- Blots.

2.11.3 Koloniefilter-Hybridisierung (Grunstein und Hogness 1975)

Von einer Agarplatte mit potenziell rekombinanten Bakterienkolonien wurde mittels Nitrozelluloserundfilter ein Abklatsch angefertigt. Dieser Rundfilter wurde für 7 Minuten in 0,5 M NaOH inkubiert und getrocknet. Danach für je 5 Minuten in 1,5 M Tris (pH 7) und in 0,5 M Tris/1,5 M NaCl (pH 7,4) inkubiert. Die getrockneten Filter wurden in eine Proteinase K- Lösung (1 mg/ml in 1 x SSC) überführt und kurz durch 100 % EtOH gezogen. Nach 5-minütiger Inkubation in 0,3 M NaCl wurden Koloniereste durch festes Drücken zwischen Filterpapieren entfernt. Die Fixierung der DNA auf dem Filter erfolgte bei 80 °C für 2 Stunden. Die Vorbereitung für die Hybridisierung der Filter erfolgte wie unter 2.11.2 Southern-Hybridisierung bzw. „Pirrotta“ beschrieben. Die Hybridisierung mit 900 µl Hybridisierungslösung mit Sonde fand bei 60 °C ü. N. in Petrischalen statt. Die Nitrozelluloserundfilter wurden wie in 2.11.2 Southern-Hybridisierung bzw. „Pirrotta“ beschrieben mehrmals bei 60 °C mit 2 x SSC gewaschen und anschließend getrocknet. Die Exposition erfolgte wie in 2.11.2 Southern-Hybridisierung bzw. „Pirrotta“ beschrieben.

Die Bestimmung von möglichen Positiv-Klonen erfolgte über den Vergleich der Autoradiogramme mit den bewachsenen Agarplatten. Zur Bestätigung der positiven Signale wurden die entsprechenden Kolonien auf einem gerasterten Nitrozelluloserundfilter ausgestrichen. Von dem Filter wurde ein Replikat angefertigt, von welchem die positiven Klone gepickt und vermehrt werden konnten.

Die Inkubation der Filter erfolgte bei 37 °C ü. N. und die Behandlung erfolgte wie oben beschrieben.

2.12 Polymerase Kettenreaktion (PCR) (Mullis et al. 1986)

2.12.1 Standard-PCR

Die Amplifikation spezifischer DNA-Abschnitte wurde mittels Polymerase-Ketten-Reaktion erzielt. Standardmäßig diente dazu die GoTaq®DNA Polymerase der Fa. Promega (Madison, USA) sowie der Thermocycler T3 (Biometra, Göttingen, Deutschland). Die sequenz-spezifisch kreierten Oligonukleotide wurden von der Fa. Sigma Aldrich (St. Louis, USA) bezogen. Eine Liste der entsprechenden Sequenzen befindet sich im elektronischen Anhang.

Tabelle 3 Standard PCR-Ansatz mit der GoTaq®DNA-Polymerase.

5 x GoTaq Puffer	10 µl
MgCl ₂	4 µl
dNTPS	1 µl
<i>forward</i> -Primer	2 µl
<i>reverse</i> -Primer	2 µl
Matrizen-DNA (~ 20 ng/µl)	1 µl
HPLC-H ₂ O	29,8 µl
GoTaq® DNA Polymerase	0,2 µl

In Tabelle 3 ist eine Standard-PCR und in Tabelle 2 ein Standard-PCR-Programm aufgelistet. Wobei gewisse Parameter, wie Annealing-Temperatur oder Elongations-Zeit variieren konnten.

Tabelle 4 PCR-Programm eines GoTaq®-PCR-Ansatzes.

Initiale Denaturierung	94 °C	5 min	
Denaturierung	94 °C	30 sek	40 x
Annealing	50- 62 °C	30 sek	
Elongation	72 °C	1 min/kb	
Finale Elongation	72 °C	5 min	

2.12.2 cDNA-Synthese

Die cDNA-Synthese erfolgte mittels SuperScriptIII™ RT der Fa. Invitrogen (Karlsruhe, Deutschland). Das Endvolumen der Erststrangsynthese umfasste 20 µl je Ansatz, wobei sich dieser aus 2-5 µg total -RNA und 1 µl mM dNTP Mix, aufgefüllt auf 14 µl, zusammensetzte.

2.12.3 RT-PCR

Die RT-PCR funktioniert wie eine Standard-PCR (2.12.1 Standard-PCR) jedoch diente als Template-DNA die unter 2.12.2 hergestellte cDNA. Diese wurde je nach Anwendung unverdünnt, 1:10 oder 1:100 verdünnt eingesetzt.

2.13 Sequenzierung

2.13.1 Sanger Sequenzierung

Exonuklease I/alkalische Phosphatase-Behandlung

Die Direkt-Sequenzierung von PCR-Produkten erfordert eine Exonuklease I/alkalische Phosphatase-Behandlung. Hierzu wurden je PCR-Ansatz 10 U Exonuklease I und 0,9 U alkalische Phosphatase der Fa. Thermo Scientific (Waltham, MA, USA) hinzugegeben und für 25 Minuten bei 37 °C inkubiert. Die Inaktivierung der Enzyme erfolgte bei 72 °C für 15 Minuten.

Sanger-Sequenzierung

Die Sequenz von PCR-Produkten (direkt/kloniert) wurde mittels Sequenzierung nach Sanger von der Fa. Starseq GmbH (Mainz, Deutschland) mittels einer Kapillarelektrophorese ermittelt. Hierfür wurden bei einer Direkt-Sequenzierung 50 bis 200 ng aufgereinigte DNA eingesetzt. Für die Sequenzierung von pGem-Klonen waren je 400 bis 700 ng erforderlich. Die für die Sequenzierung benötigten Primer waren entweder sequenz-spezifische Primer (10 pmol/ μ l) für PCR-Produkte oder Standard-Primer wie T7 bzw. Sp6 (10 pmol/ μ l) für pGem-Klone.

Sequenzauswertung

Die aus der Sequenzierung der Fa. Starseq resultierenden Sequenzen wurde über das jeweilige Chromatogramm mittels dem Programm Finch TV der Fa. Geospiza Inc. (Seattle, USA) einer optischen Kontrolle unterzogen. Hatten sie eine entsprechende Qualität wurden die Sequenzen mit dem Lasergene8® Programmopaket (insbesondere EditSeq und SeqMan) der Fa. DNASTAR Inc (Madison, USA) editiert. Zum Vergleich der Sequenzen mit bekannten Referenzsequenzen wurden diese z. B. mittels SeqMan assembliert.

Der BLAST-Algorithmus () wurde zum Abgleich mit der auf NCBI hinterlegten Datenbank nach ähnlichen Sequenzen durchsucht.

2.13.2 Illumina *Next-Generation Sequencing*

Das Genom und Transkriptom von *P. olivacea* (ungesext, ♀, ♂) wurde auf dem Illumina HighSeq 2000 des Nukleinsäure-Analyse-Zentrums (IMSB, Mainz, Deutschland) im Rahmen dieser Arbeit sequenziert. Dabei handelt es sich ausschließlich um sogenannte *paired-end*-Bibliotheken aus männlichen bzw. weiblichen oder ungesexten L4-Larven der Spezies *P. olivacea*. Mit allen Bibliotheken der Spezies *P. olivacea* wurden 100 bp *paired-end*-Läufe durchgeführt. Die NGS-Daten der anderen Arten (*C. thummi*, *C. piger*, *C. luridus*, *C. annularius*) sind im Rahmen von FI-Studentenpraktika entstanden.

2.14 Bioinformatische Methoden

2.14.1 Konvertierung der Rohdaten

Die mit Hilfe des Illumina HiSeq™2000 generierten Sequenzdaten wurde im qseq-Format gespeichert. Die Sequenzen mit den gleichen Adaptersequenzen wurden mit Hilfe des Skripts `qseq_pe_tag_sort_list.pl` (Benjamin Rieger, IMSB, Mainz, Deutschland) sortiert. Die qseq-Dateien wurden mit dem Skript `qseq2fastq.pl` umgewandelt. Neben der Umwandlung fand eine erste Filterung der Sequenzen nach dem Illumina spezifischen Chastitiy-Filter statt.

2.14.2 Importieren der Rohdaten

Die Sequenzanalyse erfolgte hauptsächlich über das Programm CLC Genomics Workbench (CLC bio, Aarhus, Dänemark), daher wurde die Daten mittels der Funktion „Import-> Illumina...“ importiert.

2.14.3 Trimmen der Rohdaten

Die Rohdaten wurden nach verschiedenen Parametern getrimmt, dies erfolgte nach dem `fastq_qual_trim_script` (Benjamin Rieger, IMSB) oder mittels der CLC Genomics Workbench (CLC bio, Aarhus, Dänemark). Am 5'Ende wurden sechs Nukleotide und am 3'Ende wurden fünf Nukleotide abgeschnitten. Teilsequenzen mit einem Phred-Score kleiner 20 wurden entfernt.

2.14.4 *De novo*-Assembly der getrimmten Daten

Die *De novo*-Assemblies der getrimmten Daten erfolgte über die CLC Genomics Workbench (lokal bzw. Server-Plugin) oder über die Assembly Cell (ebenfalls CLC bio, Aarhus, Dänemark). Die daraus resultierenden Assemblierungen waren die Basis der bioinformatischen Sequenzanalyse. Der allgemein zugrundeliegende Algorithmus beruht auf de Bruijn Graphen (Miller et al. 2010). Die Sequenzen wurde in kurze Subsequenzen k-mers zerlegt, die Größe der k-mers variierte dabei zwischen 15 und 64 bp. Die *bubble size* betrug in jedem Assembly 300 nt. Die Mindestlänge eines Contigs betrug 200 bp.

2.14.5 Kartierung von genomischen bzw. transkriptomischen Daten

Die getrimmten Daten (DNA oder RNA) wurden mittels CLC Genomics Workbench gegen die jeweiligen Referenzsequenzen mittels der Funktion „NGS Core Tools -> Map Reads to Reference“ bei DNA-reads oder mittels „Large Gap Read Mapping“ (LGRM) bzw. „RNA-Seq“ bei RNA-reads kartiert. Die Stringenz der Parameter *length fraction* (LF) und *similarity fraction* (SIM) wurde je nach Anwendung angepasst.

2.14.6 Blast

Der BLAST-Algorithmus („**B**asic **L**ocal **A**lignment **S**earch **T**ool“, (Altschul et al. 1990) ist ein heuristischer Algorithmus welcher kurze Übereinstimmungen von 2 Sequenzen feststellt.

2.15 Puffer und Lösungen

Agarosegel (1%)	1 % (w/v) Agarose in 1 x TBE
Agarplatten (500 ml)	7 g Agar-Agar 25 mg IPTG (in 500 µl DMF) 50 mg X-Gal (in 500 µl DMF) 5 ml Ampicillin-Stammlösung
Ampicillin-Stammlösung	100 mg Ampicillin in 10 ml HPLC-H ₂ O
“Antifading-Mix”	0,01 g para-Phenylendiamin 6,6 g Glycerin in 1 x PBS
Bromphenolblau- DNA-Ladepuffer (BPB)	4 M Harnstoff 0,1 M Na ₂ EDTA 50 % (w/v) Saccharose 0,1 % (w/v) BPB
Chloroform/Isoamylalkohol (CI) (24:1)	98 % Chloroform 2 % Isoamylalkohol
DEPC-Wasser	1 ml Diethylpyrocarbonat in 1 L VE-Wasser
Denaturierungspuffer	50 mmol/l NaOH 1,5 mol/l NaCl
Dialysepuffer (10x)	3 M NaCl 0,25 M Tris-HCl 1 M Na ₂ EDTA
Elektrophorese-Puffer (10 x) (E-Puffer)	0,36 M Tris 0,3 M NaOH 0,1 M Na ₂ EDTA
EtBr-Färbelösung	0,1 % EtBr-Stammlösung (10 mg/ml) In 1x E-Puffer
Guanidinthiocyanat (GTC)- Stammlösung	4,4 ml 0,75 M Na-Citrat pH 7,0 6,6 ml 10 % N-Lauroylsarkosin 62,5 g Guanidinthiocyanat in 73,3 ml DEPC-Wasser
Guanidinthiocyanat (GTC)- Arbeitslösung	12,5 ml GTC-Stammlösung 87,5 µl β-Mercaptoethanol

Homogenisierungspuffer (500 ml)	58 g Saccharose 10 ml 0,1 M Na ₂ EDTA, pH 7,4 2,5 ml 0,1 M Na ₂ EGTA, pH 7,4 50 ml Puffer A
Hybridisierungspuffer	1 x PM (in 3xSSC) 0,5 % SDS
LB-Medium	0,5 % (w/v) Natriumchlorid 0,5 % (w/v) Hefe-Extrakt 1 % (w/v) Trypton
MOPS-Puffer (1 x)	0,2 M MOPS, pH 7,0 20 mM NaAcetat 10 mM Na ₂ EDTA pH 8,0
Phosphate-Buffered-Saline (PBS) (10 x)	1,37 M Natriumchlorid 30 mM Kaliumchlorid 10 mM Kaliumdihydrogenphosphate
Phenol/Chloroform/ Isoamylalkohol (PCI) (25:24:1)	50 % 50 mM Phenol 48 % Chloroform 2 % Isoamylalkohol
Präinkubationsmedium (PM)	0,02 % (w/v) Ficoll 400 0,02 % (w/v) Polyvinylpyrrolidon 0,02 % (w/v) BSA in 3 x SSC
Standard-Salz-Citrat-Lösung (SSC) (10 x)	1,5 M NaCl 0,15 M NaCitrat
TBE-Puffer (1x)	90 mM Tris 90 mM Borsäure 25 mM Na ₂ EDTA
50T/50E-Puffer, pH 8,0	50 mM Tris 50 mM Na ₂ EDTA

3 Ergebnisse

3.1 Isolierung von Nukleinsäuren aus *P. olivacea*-Larven/-isolierten Gehirnen

3.1.1 Isolierung von genomischer DNA aus unterschiedlichen Geweben

Als Ausgangsmaterial wurde hochmolekulare DNA isoliert und eingesetzt. In Abbildung 11 sind die unterschiedlichen genomischen DNA-Präparationen, die aus *P. olivacea* -Larven bzw. isolierter Gehirne isoliert wurden, gelelektrophoretisch aufgetrennt. Als Ausgangsmaterial dienten ganze Larven (überwiegend polytänes Gewebe, Abbildung 11 A, B & C) oder isolierte Gehirne (überwiegend diploides Gewebe, Abbildung 11 D). Die genomischen DNA-Präparationen, welche in Abbildung 11 A dargestellt sind, stammen aus älteren Präparationen aus ganzen Larven (Schmidt, persönliche Mitteilung) Abbildung 11 B, C und D sind eigene DNA-Präparationen.

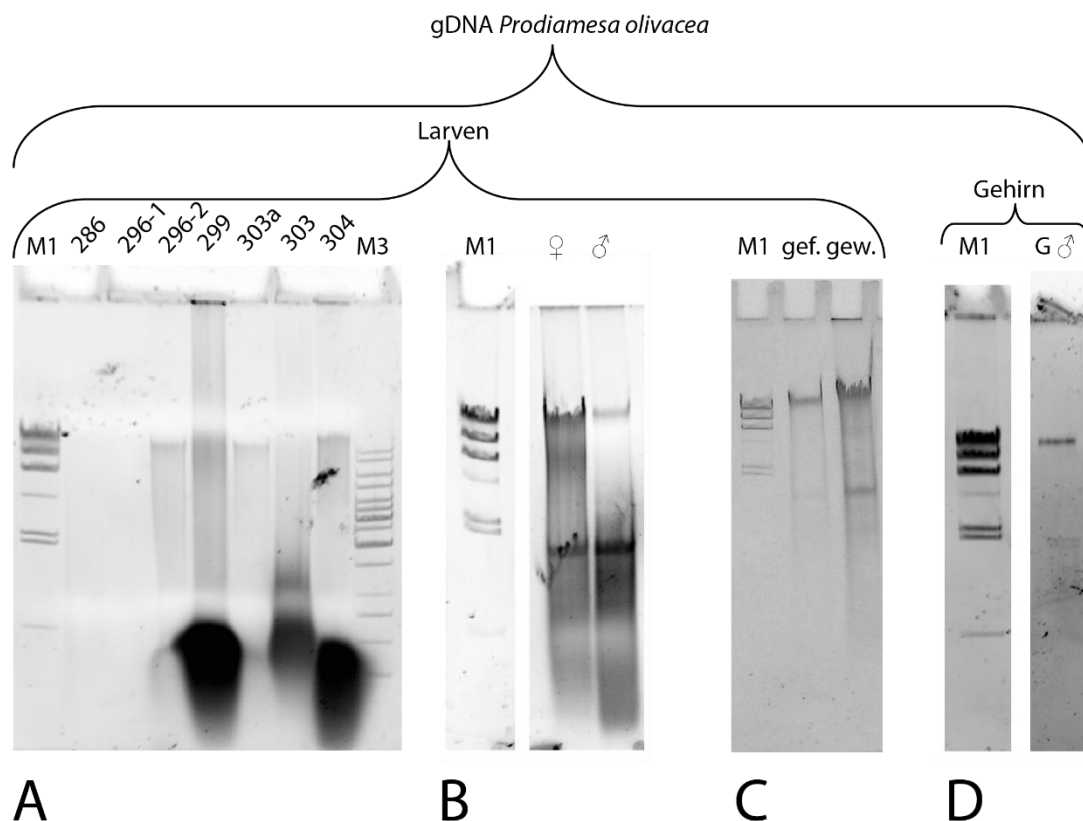


Abbildung 11 Gelelektrophoretische Auftrennung genomischer *P. olivacea*-DNA. A: Genomische DNA aus einem Gemisch von weiblichen & männlichen-*P. olivacea*-Larven B: Genomische DNA aus weiblichen oder männlichen *P. olivacea*-Larven C: Genomische DNA aus einem Gemisch von weiblichen & männlichen-*P. olivacea*-Larven, linke Spur auf Glasstab gewickelt (gew.) rechte Spur Ethanol-gefällt (gef.) D: Genomische DNA aus Gehirnen von männlichen *P. olivacea*-Larven **Marker 1 (M1)**: Lambda DNA/HindIII Marker **Marker 3 (M3)**: GeneRuler™ 1 kb DNA Ladder

3.1.2 Präparation von Gesamt-RNA aus *P. olivacea*-Larven

Nach der RNA-Präparation aus ganzen Larven (♀ und ♂ getrennt) erfolgte eine Qualitätskontrolle und Konzentrationsbestimmung mittels NanoDrop® und Agilent 2100 Bioanalyzer. Die Konzentrationsbestimmung mit Hilfe der NanoDrop-Messung ergab eine Konzentration von 349,89 ng/μl der ♂-RNA-Lösung und 1382,01 ng/μl der ♀-RNA-Lösung. Diese Werte decken sich in etwa mit den Werten, welche via Agilent 2100 Bioanalyzer generiert wurden (Abbildung 12). Die in Abbildung 12 dargestellten Elektropherogramme (links:♂, rechts:♀) zeigen das Profil der elektrophoretischen Auftrennung der RNA mittels Agilent 2100 Bioanalyzer. Die Elektropherogramme zeigen das für Dipteren typische Profil mit nur einem rRNA-Peak korrespondierend zum 18SrRNA-Peak. Der 28S-rRNA-Peak fehlt, da bei Dipteren die 28S-rRNA in zwei Hälften unter denaturierenden Bedingungen zerfällt. Eine RIN-Berechnung ist daher nicht möglich. Im Elektropherogramm ist so gut wie kein RNA-Abbau zu erkennen (Abbildung 12).

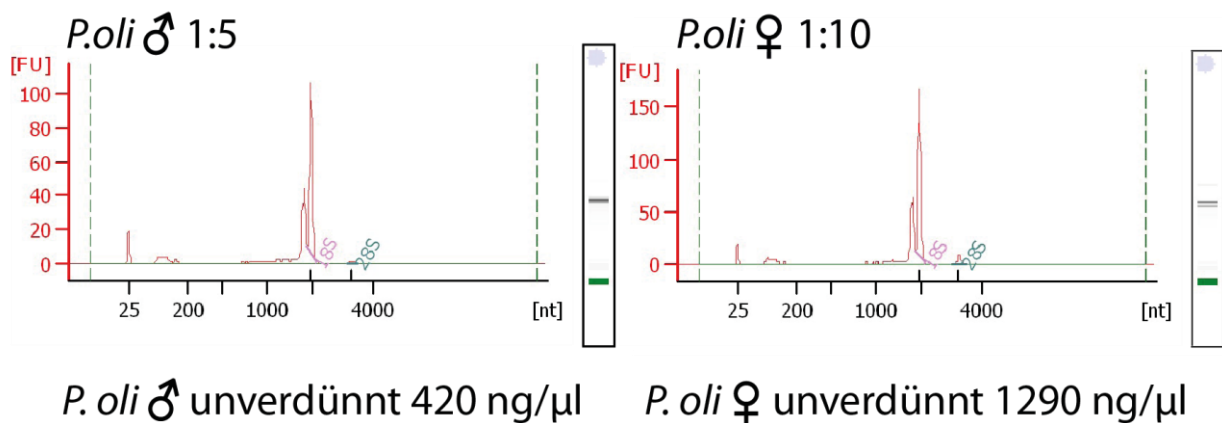


Abbildung 12 Konzentrations- und Qualitätsbestimmung der isolierten Gesamt-RNA aus weiblichen bzw. männlichen *P. olivacea*-Larven. Elektropherogramm der Bioanalyzer-Messung der aus *P. olivacea*-Larven isolierten Gesamt-RNA. Die Ergebnisse der NanoDrop-Messung ergaben Konzentrationen von ~350 ng/μl (♂) bzw. 1382 ng/μl (♀) der jeweiligen RNA-Lösung. Der 260/280-Ratio ergab für die ♂-RNA-Lösung einen Wert von 1,94 und für die ♀-RNA-Lösung 1,68.

Weitere Qualitätsinformationen können dem 260/280-Ratio entnommen werden (Abbildung 12, A), dieser beträgt 1,94 (♂) bzw. 1,68 (♀).

3.2 NGS-generierte Daten

Die erhobenen Daten der verschiedenen durchgeführten Sequenzierungen sind in Tabelle 5 zusammengefasst. Bei allen Sequenzierungen wurden 100-nt-paired-end-libraries eingesetzt welche aus genomischer DNA oder aus Gesamt-RNA aus *P. olivacea* basieren. Es wurde bei

den Datensätzen #468 bis 552 0,7 bis 1 g Larven als Ausgangsmaterial für die genomische DNA-Präparation gewählt. Das Ausgangsmaterial für den Datensatz 729 war genomische DNA aus ca. 90 männlichen *P. olivacea*-Larvengehirnen. Die Datensätze #539 bis #552 und #729 basierten auf männlichen oder weiblichen *P. olivacea*-Larven, wogegen der Datensatz #468 männlichen und weiblichen-*P. olivacea*-Larvenmaterial basiert. Für die Datensätze wurde je ein durchschnittlicher *paired-end*-Abstand der Sequenz-*reads* ermittelt, welcher zwischen 100 bis 600 bp liegt. Der kleinste Datensatz ist #729 mit 91.681.756 Sequenzen. Der durchschnittliche AT-Gehalt der genomischen DNA-Datensätze liegt zwischen 66 bis 68 %. Die Assemblierungen wurden mit kmers von 15 bis 64 durchgeführt und jeweils die qualitativ Besten¹ in Tabelle 5 vermerkt.

Tabelle 5 Übersicht der Sequenz-Datensätze aus *P. olivacea*, welche mittels Illumina HiSeq 2000 generiert wurden.

NGS-Daten					Sequenzen		Assembly		
Material	#	♂/♀	Typ	Library-Typ	Paired-end (PE)Distanz	Rohdaten	getrimmte Daten	kmer-size	AT-Gehalt
<i>P. olivacea</i> -Larven	552	♂	DNA	100 nt, PE	100-470 bp	229.346.242	222.265.470	20,35,52,62	66 %
	540	♂	RNA	100 nt, PE	100-370 bp	213.577.186	209.644.437	18,20,25	
<i>P. olivacea</i> -Gehirn	551	♀	DNA	100 nt, PE	120-550 bp	385.536.188	378.794.318	20,35,59,61	67 %
	539	♀	RNA	100 nt, PE	100-370 bp	245.740.342	242.606.549	18,20,25	
	468	♂/♀	DNA	100 nt, PE	130-570 bp	306.091.484	296.848.286	63,64	67 %
<i>P. olivacea</i> -Gehirn	729	♂	DNA	100 nt, PE	200-600 bp	91.681.756	88.914.889	24	68 %

Zudem wurde ein "Master-Assembly" angefertigt (aus allen DNA-Datensätzen der Tabelle 5; Rapp & Rieger; IMSB Mainz), welches automatisiert über *spliced reads* annotiert wurde (Rapp, IMSB Mainz). Die 45.388 Contigs aus dem "Master-Assembly" decken im günstigsten Fall 70 % des Genoms ab (Tabelle 6). Die automatisierte Annotation ergab eine Zahl von 14.000 potenziellen Genen. Darüber hinaus wurden 21.700 mögliche Transkripte mit Hilfe der *spliced reads* aus den Datensätzen 539/540 annotiert.

Tabelle 6 Übersicht des „Master“-Assemblies aus *P. olivacea* generiert aus den Datensätzen 468, 551, 552 und 729

Art	ASM	Bp	Contigzahl	Contigs mit gemappten reads	potenzielle Gene	potenzielle Transkripte
<i>P. olivacea</i>	„master“	74 Mb	45.388	36.535	~14.000	~21.700

¹ Dabei wurden 4 Parameter zugrunde gelegt: Gesamt Nukleotidanzahl, N50-Wert, *average contig size* und geringe Zahl an *ambiguous bases* (N).

3.3 Analyse von hochrepetitiver DNA aus *P. olivacea*

3.3.1 Analyse von hochrepetitiver DNA mittels Restriktionsverdau genomischer DNA

Die genomische DNA aus *P. olivacea*-Larven (σ^7 und φ gemischt, siehe Abbildung 11 C) wurde mittels Verdau mit Restriktionsendonukleasen des Typs II analysiert. Dabei wurden 10 unterschiedliche Restriktionsenzyme AluI, ClaI, HincII, HinfI, TaqI, EcoRV, VspI, SspI, RsaI und MboI eingesetzt. Acht der zehn verwendeten Enzyme (AluI, ClaI, HincII, HinfI, TaqI, EcoRV, SspI und RsaI) zeigten nach gelelektrophoretischer Auftrennung kein prominentes Bandenmuster (Beispiele siehe Abbildung 13 A SspI und RsaI).

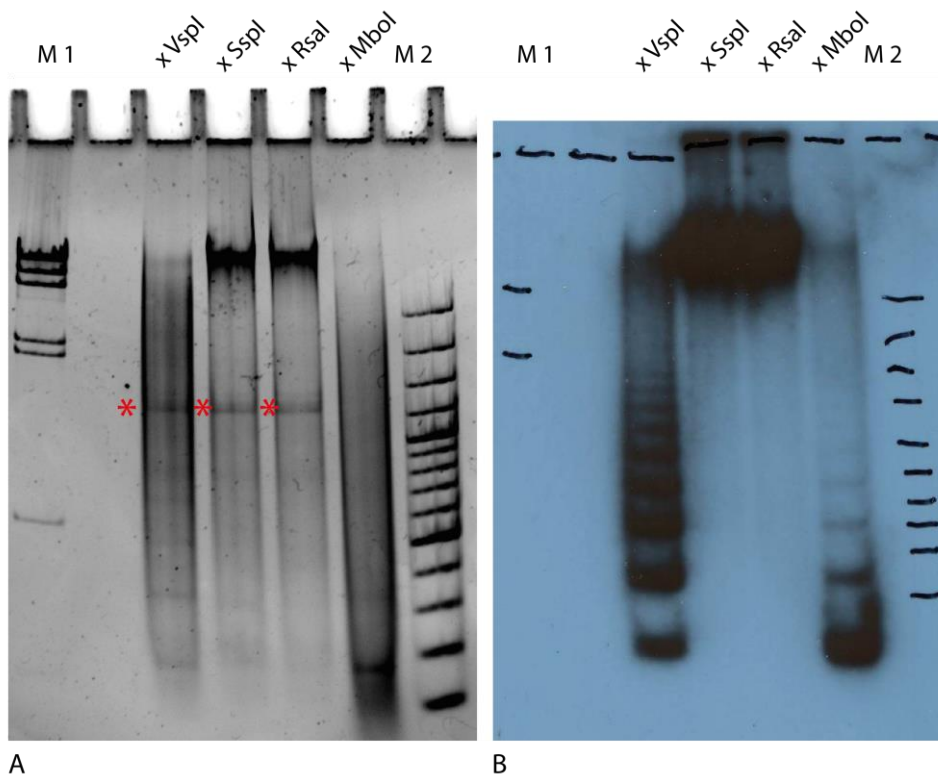


Abbildung 13 **Southern Blot** genomischer *P. olivacea*-DNA. A: Gelelektrophoretische Auftrennung restringierter gDNA aus *P. olivacea* mit VspI, SspI, RsaI und MboI. Rote Sternchen markiert vermutliche RNA-Bande. Als Molekulargewichtsstandard diente der Lambda DNA/HindIII Marker (M1) und der GeneRuler 100 bp plus DNA Ladder (M2) B: Autoradiogramm der Sonde VspI. In den Restriktionsansätzen welche mit VspI und MboI restringiert wurden ist ein schwaches Leitermuster zu sehen. Der Restriktionsansatz mit VspI zeigt dabei mehr Signale als der Ansatz mit MboI. Die 2 Restriktionsansätze mit SspI oder RsaI zeigen ebenfalls ein Signal im dem Bereich, wo sich im Gel die genomische DNA bewegt hat.

Die gelelektrophoretische Auftrennung von VspI- bzw. MboI- verdauter Spur zeigt ein schwaches Leiterbandenmuster auf einem kräftigen Hintergrundschmier (Abbildung 13 A). Die Fragmentgröße des jeweils kleinsten Fragments der Bandenmuster ist gleich und lässt sich anhand des Molekulargewichtsstandard auf ~180 bp schätzen. Die mit VspI-geschnittene

DNA zeigt drei weitere deutliche Banden, welche ungefähr das Doppelte bzw. das Dreifache Molekulargewicht des kleinsten Fragments zeigen. Alle Restriktionsansätze (außer der Ansatz mit MboI) zeigen eine weitere deutliche Bande bei ~1200 bp, die kein Vielfaches des kleinsten Fragments ist und nicht aus dem Restriktionsverdau resultiert (vergleiche Abbildung 11 C), dabei handelt es sich vermutlich um die noch vorhandene RNA.

3.3.2 Analyse von hochrepetitiver DNA durch Southern-Analyse

Die 180 bp-Fraktion der Restriktionsansatzes MboI bzw. VspI wurde aus dem Gel wiedergewonnen, dann radioaktiv markiert und als Sonde zur Southern-Hybridisierung eingesetzt. Die Abbildung 13 B zeigt das Autoradiogramm, welches aus der Hybridisierung der Sonde VspI-180 bp-Fraktion resultierte. Das Autoradiogramm für die Ansätze gDNA x SspI und gDNA x RsaI zeigt kein für tandem-repetitive DNA typisches Bandenmuster. Für die Restriktion mit VspI und MboI zeigt sich, ähnlich wie nach der elektrophoretischen Auftrennung (Abbildung 13 A) zuvor, für tandem-repetitive DNA typisches Bandenmuster, wobei bei der Verdauung mit MboI Oligomere bis zum Pentamer, mit VspI bis zum Oktamer erkennbar sind. Beide gelelektrophoretisch aufgetrennten Restriktionsansätze zeigen nach der Hybridisierung mit der spezifischen Sonde (jeweils die kleinste Fraktion) ein ähnliches Ergebnis. Daher ist nur eines der beiden Autoradiographieergebnisse abgebildet (Sonde VspI).

3.3.3 Chromosomale Lokalisation von hochrepetitiver DNA mittels FISH

Zur chromosomalen Lokalisation wurde die tandem-repetitive 180 bp-Fraktion-VspI an Polytänchromosomen aus Speicheldrüsen der Art *P. olivacea* hybridisiert.

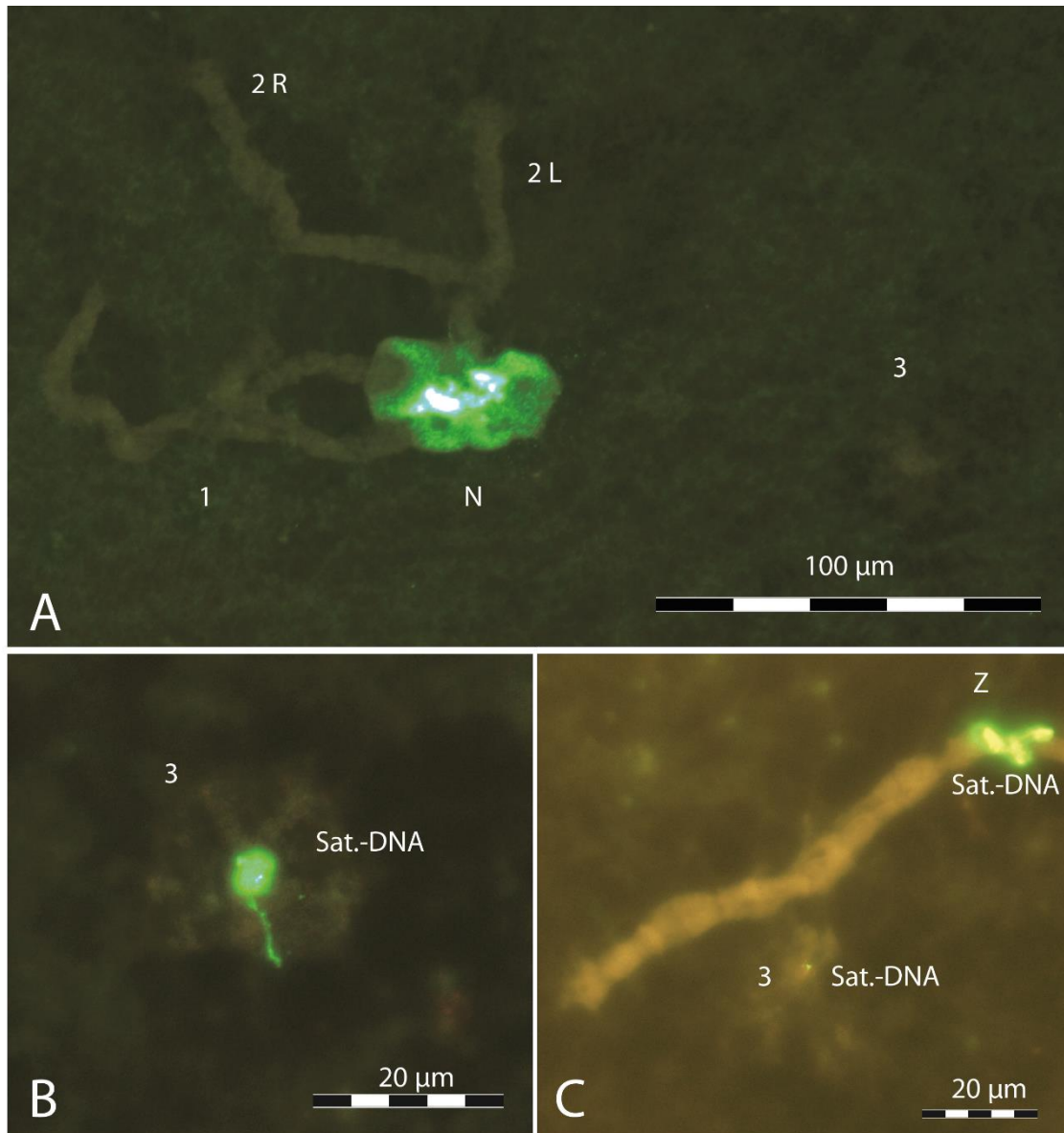


Abbildung 14 Ergebnis der FISH an polytänen Speicheldrüsenchromosomen aus *P. olivacea* mit der Sonde VspI-180 bp-Fraktion. A: Starke Hybridisierungssignale (grüne Signale, FITC) erfolgten am Zentromer der zwei großen Chromosomen (I und II). B: Die Hybridisierung (grünes Signal, FITC) erfolgte am kleinsten, unterreplizierten Chromosom (Chromosom III) C: Die Hybridisierung (grünes Signal, FITC) erfolgte am Zentromer des Chromosom I oder II und am dritten Chromosom, am heterochromatischen „Knopf“ (beschrieben von (Zacharias 1979)). N: Nukleolus 1:Chromosom I 2:Chromosom II 3: Chromosom III

Die Hybridisierungen (Abbildung 14, A) zeigen Signale an den Zentromeren der zwei großen Chromosomen, Chromosom I und II. Auch eine Hybridisierung an Chromosom III (X, Y-Chromosom) konnte bei stärkerer Vergrößerung beobachtet werden (Abbildung 14, B/C). Dort

zeigt die Sonde ein Signal am heterochromatischen „Knopf“ (beschrieben durch (Zacharias 1979)). Abbildung 14-B zeigt zudem die konzentrisch umschließenden Banden um den heterochromatischen „Knopf“(HK). Aufgrund der Lokalisation am Zentromer eines jeden Chromosoms und die zuvor durch Restriktionsendonukleasen identifizierte hohe tandemartige Repetitivität dieser DNA-Fraktion, lässt sich schließen, dass es sich hierbei um klassische Sat-DNA handelt.

3.3.4 Sequenzanalyse von hochrepetitiver DNA aus *P. olivacea*

Die Sat-DNA aus *P. olivacea* hat Wiederholungseinheiten von ~168 bp Länge und weist einen GC-Gehalt von 36 % auf. Eine Monomer-Einheit enthält je eine Schnittstelle der Restriktionsenzyme MboI (/GATC) bzw. AsnI (AT/TAAT), welche 63 bp voneinander entfernt liegen. Über die Schnittstellen-Sequenz ATTAAT konnte bioinformatisch mit Hilfe von k-mer-Analysen (Rieger, IMSB, Mainz) Sequenzvarianten des Monomers ermittelt und experimentell verifiziert werden.

Varianten der Sat-DNA-Repeats aus *P. olivacea*

Für die Sequenzierung der zuvor aus genomischer DNA isolierten 168 bp-Fraktion wurden zunächst Repeats in *E. coli* kloniert und anschließend sequenziert. Dabei zeigten sich mehrere Varianten der Satelliten-Repeats, die sowohl in der Länge der Monomer-Einheiten als auch in der Nukleotidsequenz voneinander abweichen.

Darüber hinaus wurden aus einer genomischer *P. olivacea*-DNA Genbank mittels Koloniefilterhybridisierung mit einer radioaktiv-markierten Sonde aus der 168 bp-Fraktion-VspI zwei ~1 kb-Klone identifiziert, die im Dotplot-Sequenz-Vergleich mit der Monomer-Sequenz eine große Übereinstimmung zeigen. Bei Klon 1 konnten mindestens sechs Satelliten-Monomer-Einheiten identifiziert werden. Diese weisen eine Übereinstimmung zwischen 95,8 bis 99,4 % auf (Abbildung 15). Klon 2 zeigt eine Übereinstimmung von 95,1 % für eine Monomer-Einheit, die restlichen ~850 bp ergeben keinen signifikanten Sequenzübereinstimmungen. Die beiden Klone wurden an Polytäanchromosomen hybridisiert und zeigen ein ähnliches Ergebnis wie isolierte Sat-DNA (vgl. Abbildung 14).

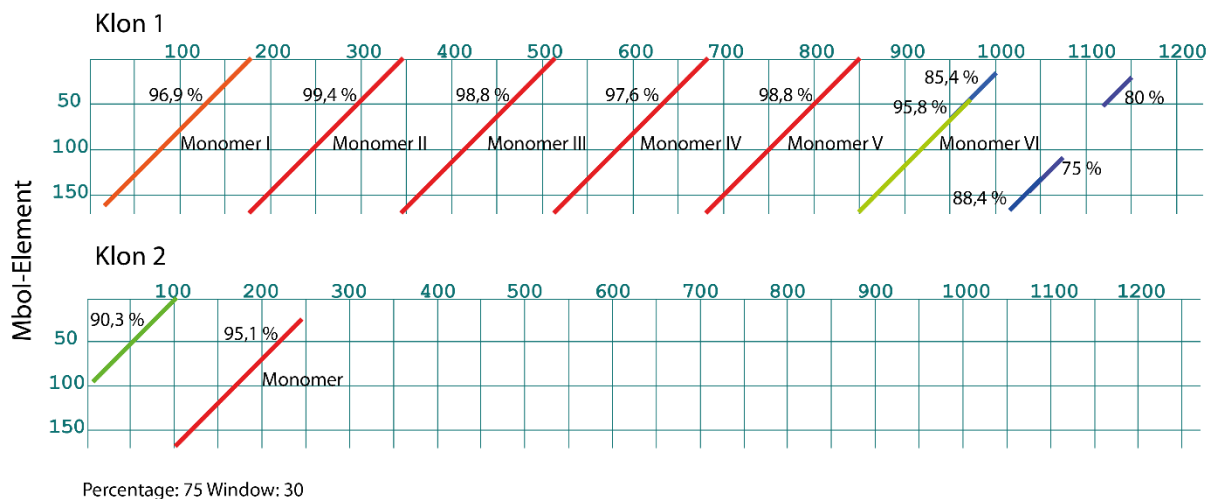


Abbildung 15 Dotplot der Mbol-I-Monomer-Sequenz mit den Klon-Sequenzen aus dem Genbank-Screening.

Kopienanzahl der Monomer-Einheiten

Über bioinformatische Analysen konnte eine ungefähre Anzahl der Monomer-Repeats bestimmt werden. Dazu wurde über die genomischen Sequenzdaten die Abdeckung von einem *single-copy*-Gen mit der Abdeckung einer Repeat-Einheit verglichen (Tabelle 7).

Tabelle 7 Bestimmung der Kopienzahl der Monomer-Einheit der Sat-DNA.
Mapping-Parameter: SIM: 0,98 LF: 0,95

Datensatz	Name	Sequenz-Typ	Referenzlänge	Total read count	Ø Abdeckung	reads total	Kopie-Zahl	Kopienanzahl / mio. reads
male Gehirn	bcn92	single copy	612 bp	350	61	99.402.046	1	
male Gehirn	Mbol-Element	hoch-repetitiv	168 bp	1.411.847	854.391	99.402.046	~14.000	140

Das 612 bp lange Gen *bcn92-like* ist durchschnittlich 61-fach abgedeckt. Eine Repeat-Einheit der Sat-DNA ist durchschnittlich 854.391-fach abgedeckt. Daraus resultiert für eine Monomer-Einheit des Satelliten eine Kopien-Zahl von ~14.000 Stück.

Über die CLC-Funktion *duplicate reads removal* wurden keine weiteren hochrepetitiven DNA-Sequenzen im *P. olivacea*-Genom identifiziert. Es konnten nach Assemblieren der *duplicate reads* fünf Contigs (Contig 1, 2, 3, 4, 5) mit einer hohen durchschnittlichen Abdeckung (474 bis 18,8) generiert werden. Einer der Contigs stimmt mit der bereits bekannten Sat-DNA überein (Contig 5). Dabei zeigte sich für die Sat-DNA-Sequenz die höchste durchschnittliche Abdeckung von 474. Drei der Contigs (1, 2, 4) zeigten nach einem BLASTX einen bakteriellen

Ursprung. Dagegen erzielte der Contig 3 keinen BLAST-Treffer. Die auf der Sequenz von Contig 3 basierte Sonde erzielte in einer FISH kein Signal.

3.4 Die Globin-Gene von *P. olivacea*

Über die Assemblierungen (vgl. 3.2) konnten drei Contigs mit BLAST-Treffern zu Dipteren – Globin -Sequenzen identifiziert werden. Contig_118 ist mit 284.392 bp der längste Contig und zeigt zwei Treffer zu Globin-Genen und. Contig_115 umfasst 94.875 bp und 7 mögliche Globin-Gene. Der dritte Contig, Contig_117, enthält innerhalb der 31.176 bp- langen Sequenz ein Globin-Gen.

3.4.1 Chromosomale Lokalisation und genomische Organisation

Zwei der drei Contigs wurden auf demselben Chromosomenarm lokalisiert. Die Hybridisierung an polytären Speicheldrüsenchromosomen aus *P. olivacea* fand an Chromosom I, Arm D' statt (Abbildung 16; A/C). Die Sonde des Contigs Globin 118 hybridisierte im Vergleich zur Sonde des Contigs_115 mehr mittig innerhalb des Chromosomenarms D' (Abbildung 16, A). Die Sonde basierend auf der Sequenz von Contig_115 hybridisierte telomernah (Abbildung 16, C).

Das Globin-Gen, welches auf Contig_117 identifiziert wurde, wurde mittels FISH an Speicheldrüsenchromosomen auf Chromosom II, Arm L lokalisiert (Abbildung 16-B).

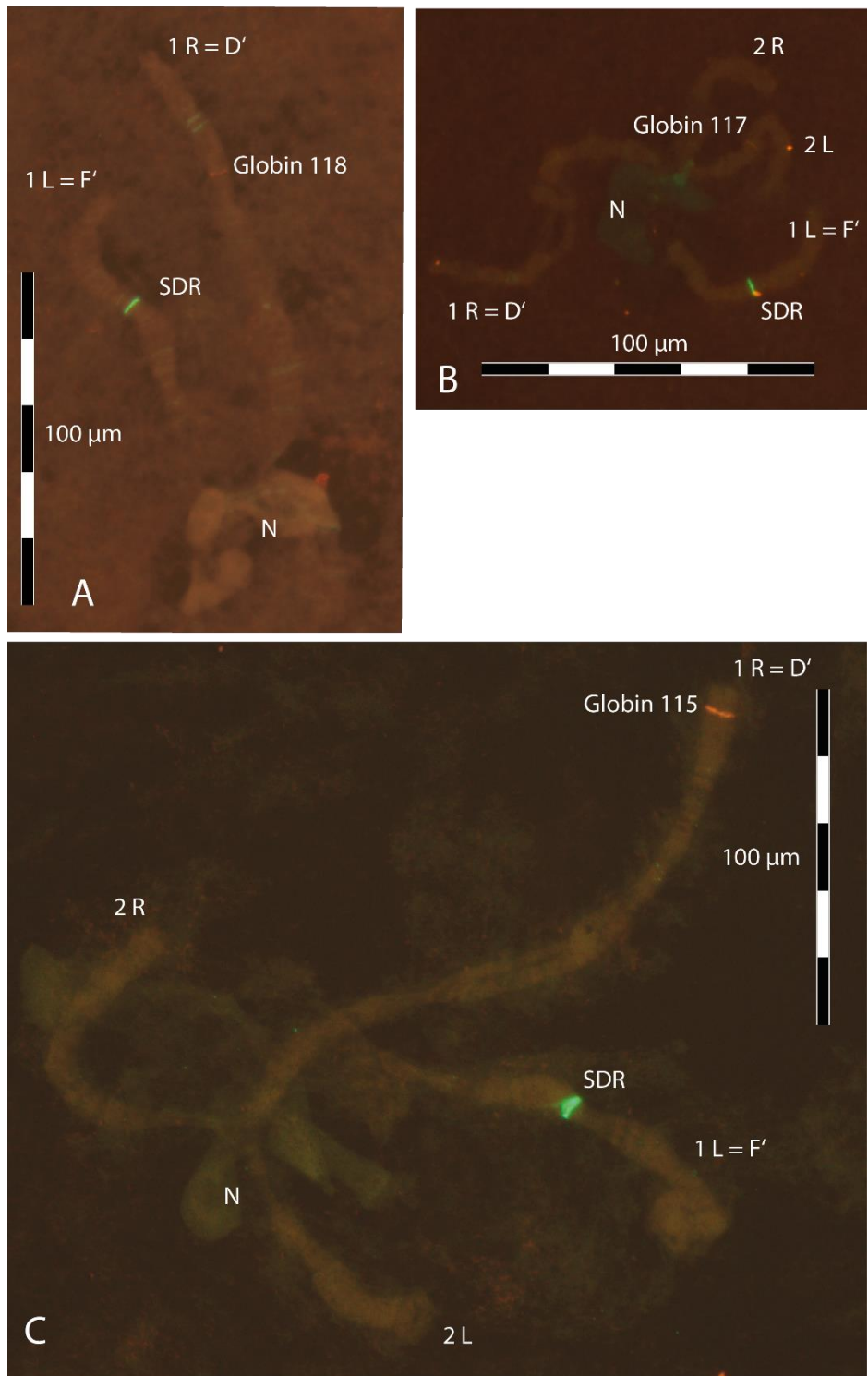


Abbildung 16 Mikroskopische Aufnahme der FISH an Polytäanchrosomen aus *P. olivacea* mit den Sonden basierend auf Contig_115, Contig_118 und Contig_117. A: Die Hybridisierung der Sonde Globin_118 (rotes Signal, TRITC) erfolgte an Chromosom I Arm D'. Als Referenzsonde diente die SDR (grüne Signale, FITC), welche neben dem Hauptsignal noch weitere, schwache Signale aufweist. B: Die Hybridisierung von Contig_Gb_117 (rotes Signal, TRITC) erfolgte an Chromosom II Arm F' in der Mitte des Chromosomarms. C: Die Hybridisierung von GB_115 (rotes Signal, TRITC) erfolgte an Chromosom I Arm D' telomernah. N: Nukleolus

3.4.2 Genstruktur

Die Globin-Gene sind auf beide DNA-Stränge verteilt, zwei (D, G) der sieben Globin-Gene von Contig_115 liegen auf dem Plusstrang (Abbildung 17; Tabelle 8) und die anderen fünf (A bis E) auf dem Minusstrang. Die Gene A bis E weisen je drei Introns und vier Exons auf. Die Gengröße von A bis E liegt zwischen 860 und 1328 bp. Die vermutliche Transkriptlänge beträgt 1152 bp für Globin A, 995 bp für Globin B, 756 bp für Globin C, 658 bp für Globin D und 802 bp für Globin E. Für alle fünf Gene wurden potenzielle ORFs identifiziert, welche zwischen 486 und 537 bp lang sind. Die daraus resultierende AS-Sequenz ist 179 AS für Globin A, 169 AS für Globin B, 164 AS für Globin C, 169 AS für Globin D und 162 AS für Globin E lang. Die Introngrößen der Globin-Gene A bis F variieren für das erste Intron zwischen 57 und 71 bp, für das zweite Intron zwischen 54 und 67 bp und für das dritte Intron zwischen 56 und 79 bp. Außer dem Globin-Gen G weist jedes Gen nur einen möglichen ORF auf. Dieses besitzt zwei mögliche ORFs, welche mit 90 bzw. 147 bp im Vergleich zu den Genen A bis F relativ kurz sind. Eine Analyse mittels BLASTX ergab für diese zwei kurzen Sequenzen keinen signifikanten Treffer. Die Gene A bis F besitzen einen (mind.) 435 bp langen ORF. Die BLASTX-Suche mit den ORF-Sequenzen der Gene A bis F zeigten signifikante Treffer zu Dipteren-Globinen.

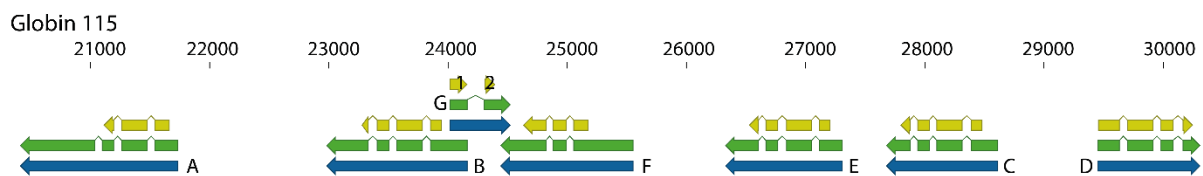


Abbildung 17 Die chromosomale Organisation des Globin Cluster auf dem Contig 115 von *P. olivacea*. Die Globin-Gene A bis G sind je nach Orientierung als blaue Pfeile annotiert. Die dazugehörigen mRNAs sind als grüne Pfeile und die daraus resultierenden ORFs als gelbe Pfeile eingezeichnet. Abgebildet ist nur der Bereich des Globin Clusters, welcher die Region 20 kb bis 31 kb umfasst.

Der Contig 118 trägt zwei potenzielle Globin-Gene, welche 1496 bzw. 980 bp groß sind (Abbildung 18; Tabelle 8). Beide Gene besitzen laut bioinformatischer Analyse (bezieht sich auf die Expression in der Larve) jeweils nur ein mögliches Transkript. Für das Globin-Gen 118 -1 wurde ein 537 bp-großer ORF identifiziert. Dieser entspricht einer AS-Sequenz von 179 AS. Ein Datenbankabgleich ergab einen signifikanten Treffer zu Dipteren-Globinen. Die Genstruktur weicht von der der Globin-Gene A bis F ab. Das Gen 118-1 hat ebenfalls drei Introns, aber das dritte Intron ist mit 525 bp jedoch deutlich größer als das dritte Intron der

Gene A bis F. Das zweite Gen, Globin-Gen 118 -2, besitzt zwei mögliche ORFs, welche mit 219 bzw. 270 bp verhältnismäßig kurz sind. Sie resultieren wie die ORFs des Gens G in kürzeren AS-Sequenzen und führen somit zu einem verkürzten Polypeptid.

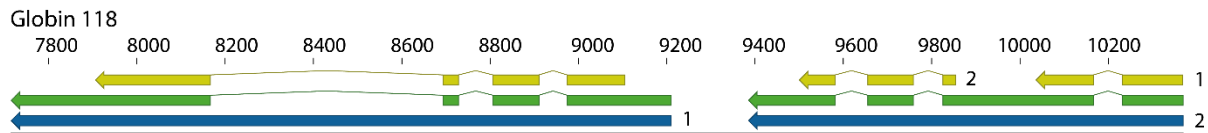


Abbildung 18 Die chromosomale Organisation des Contigs_118 mit den Globin-Genen 118_1 und 118_2. Beide Gene liegen auf dem Minusstrang des Contigs_118. Die Gene sind als blaue Pfeile annotiert. Die dazugehörigen mRNAs und ORFs sind als grüne und gelbe Pfeile annotiert. Abgebildet ist die Region 7,8 kb bis 10,5 kb des Contigs 118.

Der dritte Contig trägt ein mögliches 1359 bp langes Globin-Gen (Abbildung 19, Tabelle 8). Die Analyse der Genstruktur ergab zwei potenzielle Transkripte. Bei Transkript 1 wird das zentral liegende Intron gespleißt, wogegen bei Transkript 2 das 62 bp große Intron nicht gespleißt wird. Es existiert jedoch für beide Transkript-Varianten nur ein möglicher 480 bp-großer ORF, da das Intron im 3'-UTR liegt. Die daraus resultierende AS-Sequenz ergab in der BLAST-Suche einen signifikanten Treffer zu *globin 1* aus *A. darlingi*.

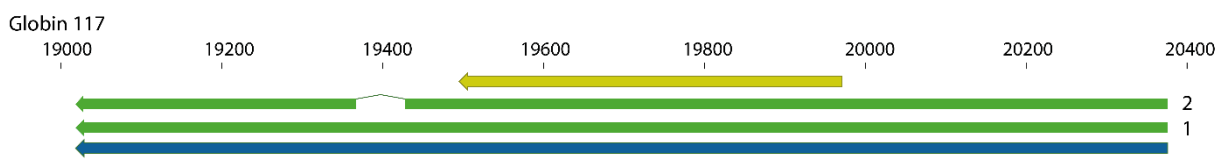


Abbildung 19 Die Genstruktur des Globin Gens 117. Die Annotation des Gens besteht aus dem Gen (blauer Pfeil), den zwei möglichen Transkripten 1 und 2 (mRNA, grün) und dem ORF (gelb). Es ist die Region 19 kb bis 20,5 kb des Contigs_117 abgebildet.

Zusammengefasst wird für acht der zehn identifizierten Globin- Gene aus *P. olivacea* je eine mRNA synthetisiert, welche als Matrize für ein funktionelles Polypeptid fungieren könnte.

Tabelle 8 Übersicht der in *P. olivacea* identifizierten Globin-Gene

Contig	Gesamt [bp]	Globin -Gene	Strang	Gen [bp]	Anzahl Transkripte pro Gen	Transkript [bp]	ORF [bp]	AS	Intron [bp]
Contig_117	31.176	117	-	1.359	2	1.344	477	159	(1) 62
						1.406			
Contig_118	284.392	118-1	-	1.496	1	833	537	179	(1) 62 (2)76 (3) 525
		118-2		980	1	785	270 219		(1) 64 (2) 65 (3) 72
Contig_115	94.875				5				
		A	-	1.328	1	1.152	537	179	(1) 60 (2) 61 (3) 56
		B	-	1.185	1	995	507	169	(1) 61 (2) 61 (3) 68
		C	-	940	1	756	492	164	(1) 71 (2) 55 (3) 70
		D	+	860	1	658	507	169	(1) 66 (2) 63 (3) 79
		E	-	983	1	802	486	162	(1) 60 (2) 67 (3) 56
		F	-	1.116	1	1.005	438	146	(1) 57 (2) 54
		G	+	512	1	379	147 90		(1) 133

Entsprechend der abgeleiteten AS-Sequenz konnte für sieben der zehn Globin-Gene aus *P. olivacea* eine Einteilung der AS-Sequenz in die Helices A bis H erfolgen (Abbildung 20, dicke violette Pfeile). Die Helices sind auf Nukleotidebene z. T. durch Intronsequenzen unterbrochen (Abbildung 20, dünne violette Striche zwischen dicken violetten Pfeilen).

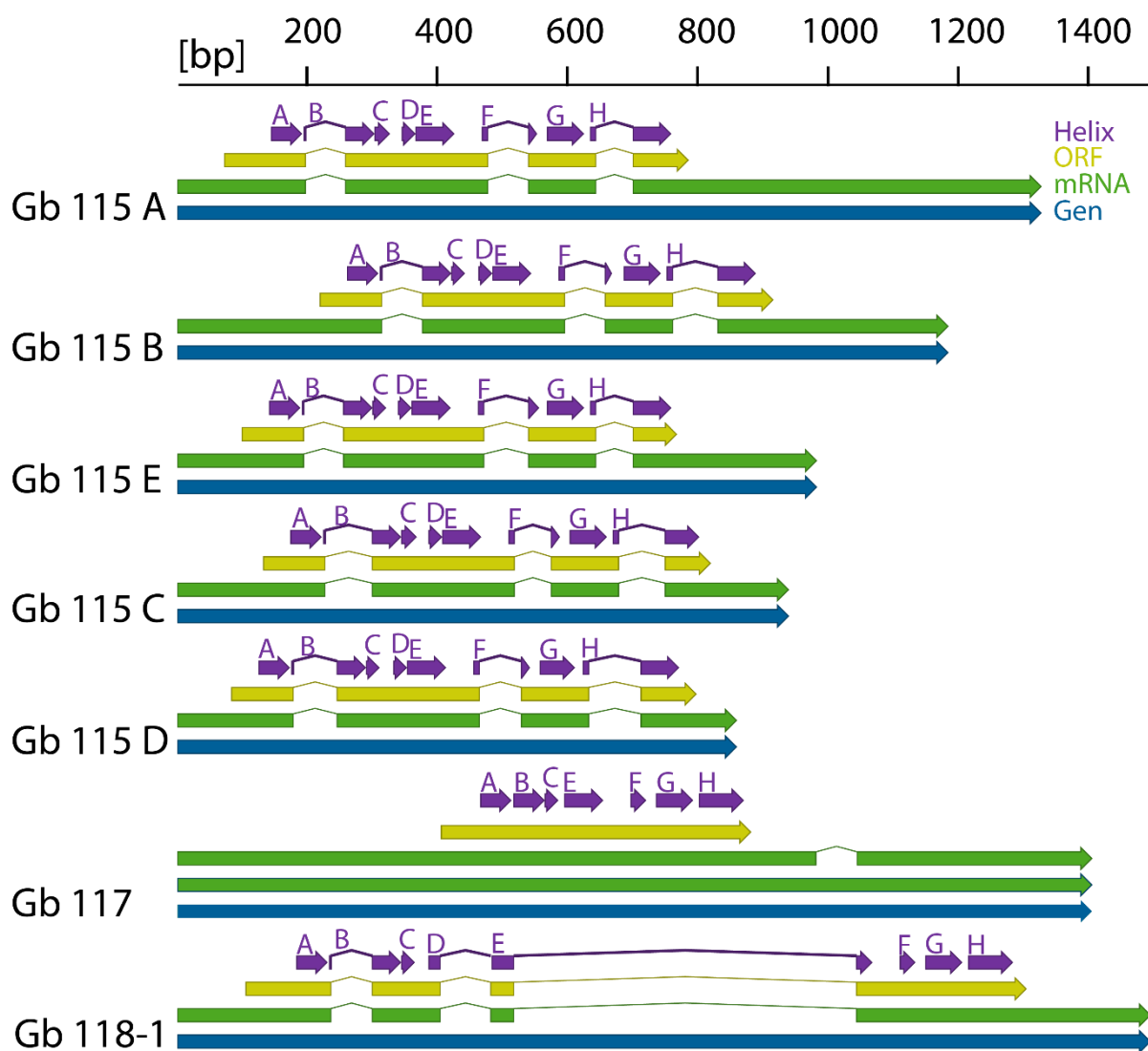


Abbildung 20 Auswahl an Globin-Genen aus *P. olivacea* inklusive Annotationen der Gene, mRNAs, Introns, ORFs und Helices der abgeleiteten Proteine. Abgebildet sind die sieben Globin-Genen aus *P. olivacea* bei denen eine Identifikation der Helices A bis H möglich war. Die Gen-Annotationen sind durch einen blauen Pfeil, die mRNA-Annotationen durch grüne Pfeile, Exons stellen die Kästen der mRNA-Annotation dar, die Introns sind durch Striche zwischen den Exons der mRNA dargestellt, die möglichen ORFs durch gelbe Pfeile und die Protein-Helices durch violette Pfeile und entsprechend nach A bis H beschriftet.

Für die Globin-Genen *P. olivacea* 115 A bis D und 118-1 konnten alle Helices A bis H identifiziert werden. Für *P. olivacea* Globin 117 konnten alle Helices außer Helix D identifiziert werden. Für *P. olivacea* Globin 115 F konnten alle Helices B bis H und eine nur unvollständige Helix A identifiziert werden (nicht in Abbildung 20 dargestellt).

3.5 Die Histon-Gene von *P. olivacea*

3.5.1 Die Organisation

Die Histon-Gene von *P. olivacea* sind wie folgt angeordnet: <H3 H4> <H2A H2B> H1> (Abbildung 21). Für die Histon-Gene H3, H4, H2A und H2B wurde eine TATA-Box identifiziert. Zudem zeigte die bioinformatische Analyse für alle fünf Gene eine 16 bp lange palindromische Sequenz (*conserved 3' hairpin loop, c3'h*; (Busslinger et al. 1979)), welche 45 bis 90 bp hinter dem Stopp-Kodon liegt (Tabelle 9, Abbildung 21).

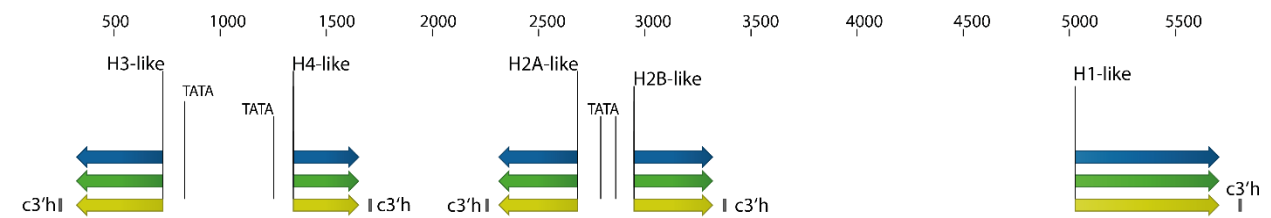


Abbildung 21 **Organisation der Histon-Gene H1, H2A, H2B, H3 und H4 aus *P. olivacea*.** Die Gene sind blau, die mRNA grün und der open reading frame (ORF) gelb annotiert. Die 16 bp lange palindromische Sequenz, auch bezeichnet als *conserved 3' hairpin (c3'h)*, ist durch einen schwarzen Strich markiert. Der Promotor der Gene H2A, H2B, H3 und H4 ist mit TATA markiert.

3.5.2 Die Lokalisation

Die Histon-Gene in *P. olivacea* wurden von Hankeln et al. mit einer Sonde, welche auf der Sequenz aus *C. thummi* basiert, lokalisiert (Hankeln et al. 1993). Daher wurde über die bioinformatisch ermittelte *P. olivacea*-Histon-Sequenz eine Sequenz-spezifische Sonde an Polytäanchrosomen aus *P. olivacea* hybridisiert.

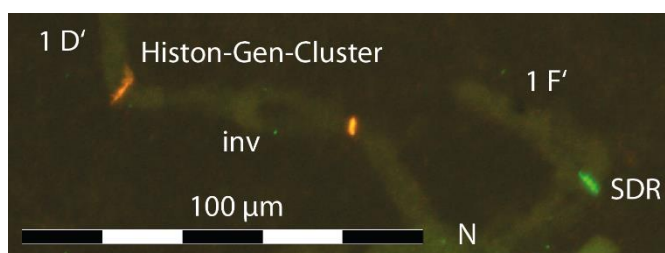


Abbildung 22 **Mikroskopische Aufnahme von Polytäanchrosomen aus *P. olivacea*, in der das Ergebnis der FISH mit einer *P. olivacea*-Sequenz-spezifischen Histon-Sonde dargestellt ist.** Als Referenz-Sonde (grünes Signal, FITC) diente eine Sonde aus der vermutlichen SDR. Die Hybridisierung der Histon-Sonde (rote Signale, TRITC) erfolgte an zwei Loci an Chromosom I Arm R. Zwischen den Histon-Gen liegt (wenn vorhanden) die variable Inversion, welche von (Zacharias 1979) beschrieben wurde.

Abbildung 22 zeigt zwei Loci auf Chromosom I Arm D', für die Histon-Gen-Cluster, wie von Hankeln et al. zuvor beschrieben wurde (Hankeln et al. 1993).

3.5.3 Die Nukleotid-Sequenzen

Die Nukleotid-Sequenz des *P. olivacea* eines Histon-Gen-Clusters hat einen durchschnittlichen AT-Gehalt von 68 %. Die einzelnen Histon- Gene weisen einen geringeren durchschnittlichen AT-Gehalt von 56 % (H1 und H4) oder 57 % (H2a, H2B und H3) auf, die Intergenregionen einen höheren durchschnittlichen AT-Gehalt von 70 bis zu 77 % auf (Abbildung 23).

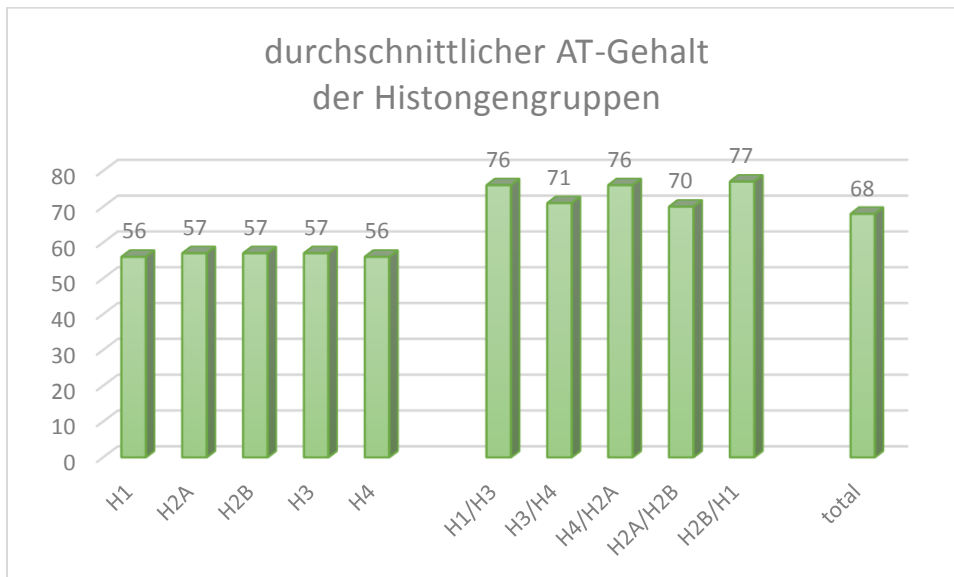


Abbildung 23 Analyse der Histon-Gensequenzen und Intergen-Sequenzen aus *P. olivacea* hinsichtlich ihres AT-Gehalts. Die Histon-Gene zeigen einen geringen AT-Gehalt von 56 oder 57 %, wogegen die Intergen-Regionen im Vergleich einen erhöhten AT-Gehalt von 70 bis 77 % aufweisen.

Die Nukleotidsequenz bzw. die abgeleitete AS-Sequenz des *P. olivacea* Histon-Gens H1 wurde mit der Sequenz des H1-Gens bzw. dessen abgeleitete AS-Sequenz aus *C. thummi* verglichen (ClustalW2). Dabei zeigt sich eine Sequenzidentität auf DNA-Ebene von 71% und eine von 64 % zwischen den AS-Sequenzen der beiden Arten (Tabelle 9). Die Nukleotidaustausche finden mit 89 Austauschen überwiegend an Kodon-Position 3 statt, Nukleotidaustausche an Kodon-Position 1 bzw. 2 dagegen nur 52- bzw. 49-mal (Anhang, Abbildung 52). Die Lücken im Nukleotid-Alignment (Poli H1 Position 379, 427 und 466) zeigen 27 deletierte oder inserierte Nukleotide (Anhang, Abbildung 52) und 12 inserierte oder deletierte Nukleotide ab Position 550; 592 & 700 (Abbildung 52, Cthu H1).

Die Nukleotidaustausche (Anhang, Abbildung 52, türkis unterlegt) resultieren in 74 AS-Austauschen. Wobei 27 durch AS mit schwach ähnlichen Eigenschaften ausgetauscht werden und 22 durch AS mit stark ähnlichen Eigenschaften. Die restlichen 25 AS werden durch AS, welche unterschiedliche chemische Eigenschaften aufweisen, ersetzt (Abbildung 47).

Die Lücken im Nukleotid-Sequenz-Alignment (Anhang, Abbildung 52) ergeben insgesamt 13 deletierte oder inserierte AS, wovon 9 AS in *P. olivacea* im Vergleich zu *C. thummi* fehlen und nur 4 AS in *C. thummi* im Vergleich zum abgeleiteten H1-Protein aus *P. olivacea* (Abbildung 47).

Die Vergleiche der Nukleotid-Sequenzen der Histon-Gene H2A, H3 und H4 aus *P. olivacea* und *C. thummi* zeigen eine Sequenzidentität von durchschnittlich 90 %. Der Nukleotid-Sequenzvergleich des Histon-Gens H2B zeigt dagegen eine geringere Identität von 83 % auf (Tabelle 9).

Die AS-Sequenzen der Histon-Gene H3 und H4 aus *P. olivacea* und *C. thummi* sind identisch. Die Übereinstimmung der AS-Sequenz der Histon-Gene H2A und H2B liegt bei 97 bzw. 96 % (Tabelle 9).

Tabelle 9 Vergleich der Histon-Gene,-Proteine und der regulatorischen Sequenzen aus *C. thummi* und *P. olivacea*.

Histon	TATA-Box	Sequenzidentität		conserved 3'hairpin
		DNA	AS	
H1	-	71 %	64 %	+
H2A	+	90 %	97 %	+
H2B	+	83 %	96 %	+
H3	+	90 %	100 %	+
H4	+	90 %	100 %	+

-: nicht identifiziert +: identifiziert

Die Alignments der Nukleotid-Sequenzen bzw. der AS-Sequenzen der Histone-Gene bzw. abgeleiteten Proteine H2A, H2B, H3 und H4 aus *P. olivacea* und *C. thummi* sind dem Anhang (Abbildung 53, Abbildung 54, Abbildung 55 und Abbildung 56) zu entnehmen.

3.5.4 Orphon-Histon-Gene

Die BLAST-Analysen der Sequenz von *P. olivacea* haben zwei weitere, sogenannte verwaiste H3-Gene (*orphon*-Loci) identifiziert. In *D. melanogaster* sind ebenfalls verwaiste H3-Gene bekannt (Childs et al. 1981) und (Ahmad und Henikoff 2002a). In *C. thummi* konnten für alle Cluster-Gene auch *orphon*-Gene identifiziert werden (Hankeln et al. 1993). Die zwei *orphon* H3-Gene besitzen abweichend von ein bzw. drei Introns. Für die beiden H3-Varianten wurde je ein ORF bestimmt. Im Alignment der drei AS-Sequenzen zeigt sich, dass die *orphon* H3-

Variante 1 (H3V1) mit der H3-Cluster-Variante zu 97 % identisch ist. Dagegen zeigt die *orphan* H3-Variante 2 (H3V2) mit nur 67 % Sequenz-Identität eine deutlich geringere Ähnlichkeit zu den beiden anderen Varianten (Abbildung 24). Neben der AS-Sequenz unterscheiden sich die H3-Varianten auch in ihrer Länge. Die H3-Cluster-Variante und die H3V1 bestehen aus 136 AS, wogegen die H3V2 mit insgesamt 158 AS länger ist (Abbildung 24).

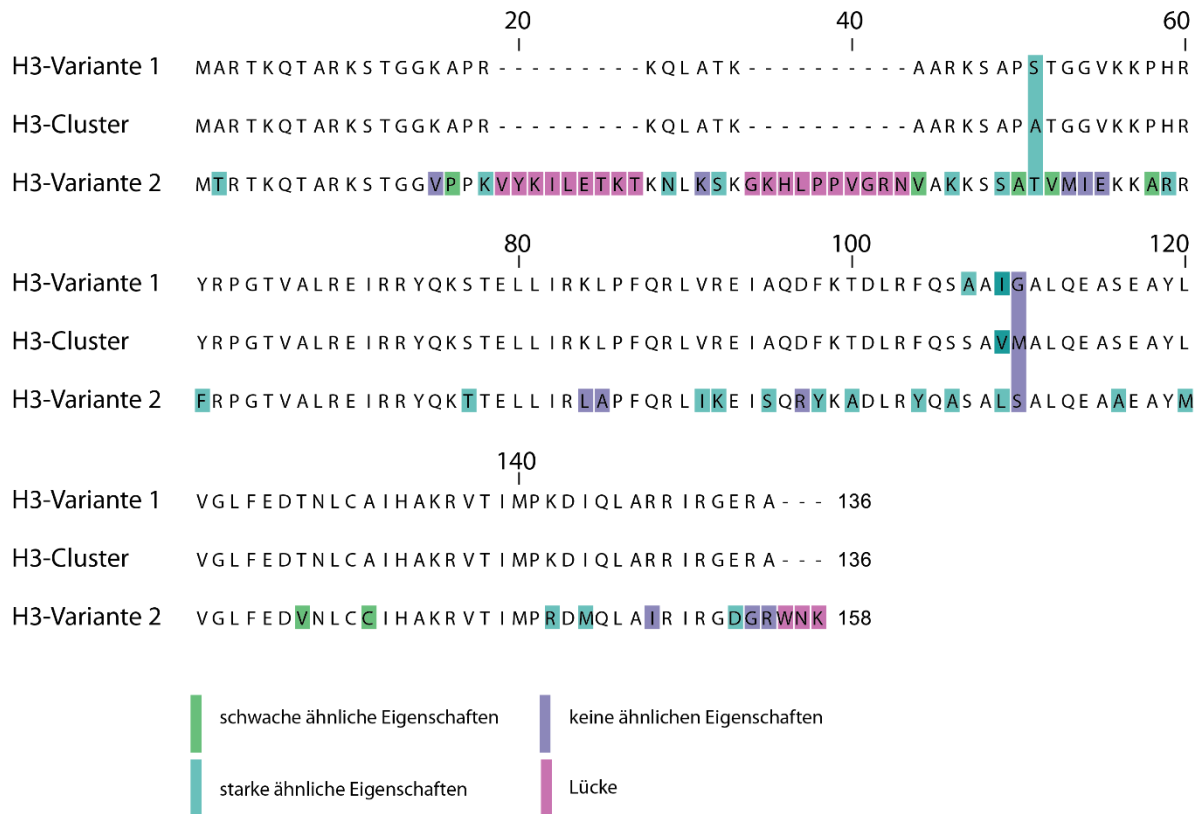


Abbildung 24 MSA der drei H3-Varianten aus *P. olivacea*. AS, welche gegen AS mit starken ähnlichen Eigenschaften ausgetauscht wurden sind türkis unterlegt. Austausch durch AS, welche schwache ähnliche Eigenschaften zeigen, sind grün unterlegt. AS-Austausche von AS gegen AS mit keinen ähnlichen Eigenschaften sind lila unterlegt. Lücken im Alignment sind rosa markiert.

Die H3V1 zeigt eine Sequenzidentität von 100 % zu der Histon 3-Variante H3.3 aus *D. melanogaster*. Das Gen *H3V2* wurde mittels FISH lokalisiert (Abbildung 25). Die mittels PCR erzeugte Sonde hybridisiert zentromernah an Chromosom I, Arm R bzw. D'. Das Gen ist somit auf demselben Chromosomenarm wie die zwei Histon-Gen-Cluster lokalisiert. Die H3V2-Sonde hybridisierte zusätzlich an den zwei Loci der Histon-Gen-Cluster (Abbildung 25, markiert durch weiße Pfeile).

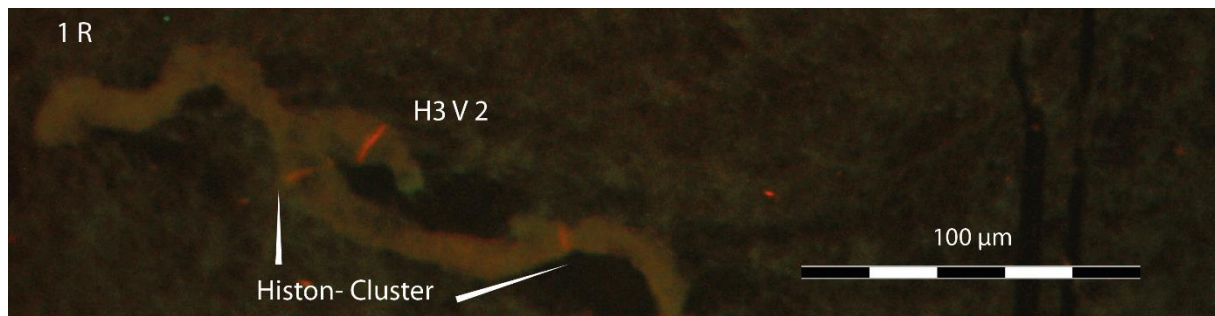


Abbildung 25 Mikroskopische Aufnahme der FISH an Polytäanchromosomen aus *P. olivacea* mit der Sonde H3V2. Die Sonde H3V2 hybridisiert (rote Signale, TRITC) an Chromosom I Arm R bzw. D' an drei Loci. Zwei der Hybridisierungen entsprechen den Histon-Cluster Loci (markiert durch weiße Pfeile). Eine dritte Hybridisierung ist zentromernah erfolgt.

Über BLAST konnte in *C. thummi* ein homologes Gen zur H3V2 aus *P. olivacea* identifiziert werden. Das identifizierte homologe H3-Gen aus *C. thummi* besitzt zwei Introns. Die AS-Sequenzidentität zur *P. olivacea*-H3V2 ist mit 69 % höher als die Übereinstimmung zwischen den *P. olivacea*-H3-Varianten untereinander. Beide Proteine sind mit 158 AS (H3V2 aus *P. olivacea*) bzw. 155 AS (H3V2 aus *C. thummi*) fast gleich groß.

In *P. olivacea* gibt es laut verschiedener Assemblierungen der Datensätze über BLAST-Suchen, ähnlich wie in *D. melanogaster*, keine Hinweise auf H1-Varianten.

3.6 Die Identifikation von homologen Genen aus anderen Dipteren in *P. olivacea*

3.6.1 Identifikation, Charakterisierung und Lokalisation von Genen aus der

Geschlechtsbestimmungskaskade von *D. melanogaster* in *P. olivacea*

Das Gen *Sxl*

Identifikation und Charakterisierung der Genstruktur von *Sxl*

Mittels des TBLASTX-Tool wurde in den assemblierten Sequenzen des *P. olivacea* Datensatzs 729 (Tabelle 5) ein Contig mit einem zu *Sxl* homologen Gen identifiziert. Dieser Contig, contig_550_asm_729_def, ist 181.692 bp lang. Das Gen *Sxl* ist 17.321 bp lang (Abbildung 26). Das *Sxl*-Gen enthält maximal fünf Introns. Die potenziell am stärksten transkribierte mRNA des *Sxl*-Gens ist mit 3.719 bp deutlich kürzer als das dazugehörige Gen. Eine ORF-Suche ergab drei mögliche offene Leserahmen, wobei zwei mit 156 bzw. 222 bp relativ kurz sind und abgeleitete AS-Sequenzen keine konservierten Domänen identifiziert werden konnten. Der dritte und längste potenzielle offene Leserahmen ist mit 924 bp deutlich länger als die anderen beiden. Eine SMART-Domänen-Suche ergab für das potenziell 308 AS-lange Polypeptid zwei konservierte RNA recognition motifs (RRM), welche mit einem *e-value* von $1,53 \cdot 10^{-19}$ bzw. $8,77 \cdot 10^{-20}$ signifikant sind. Diese zwei RRM sind auch im SXL-Protein aus *D. melanogaster* (Handa et al. 1999) und in der AS-Sequenz von *C. thummi* (persönliche Mitteilung Christiane Kraemer) zu finden.

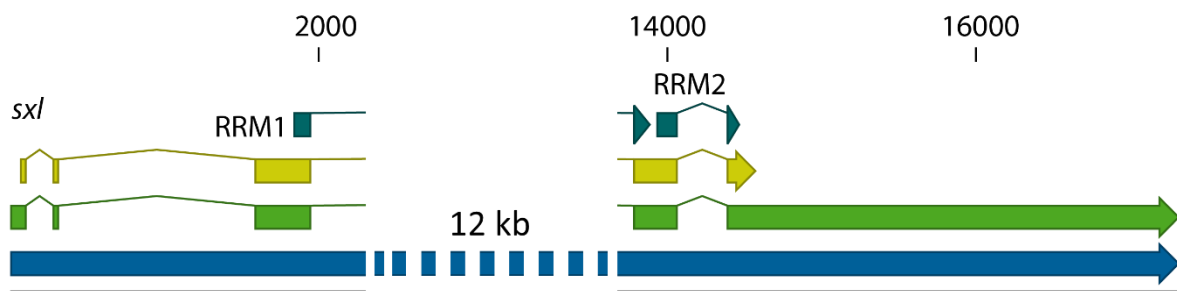


Abbildung 26 Die Genstruktur des *Sxl*-Gens aus *P. olivacea*. Das Gen (blauer Pfeil) hat eine mögliche Länge von 17 kb. Dargestellt ist das am stärksten exprimierte Transkript (grün) der fünf potenziellen Transkriptvarianten. Weiter wurden die zwei identifizierten RNA recognition motifs (RRM) annotiert (dunkelgrün). Das dritte Intron ist verkürzt dargestellt (zwischen 2 kb und 14 kb).

Ein Vergleich der SXL-Sequenzen aus *C. piger* und *P. olivacea* ergab eine Sequenzidentität von 82 %. Abweichungen der AS-Sequenzen beschränken sich dabei hauptsächlich auf den N - Terminus des Polypeptids. Das in *D. melanogaster* alternativ-gespleißte Exon 3 konnte per BLAST nicht identifiziert werden.

Die chromosomale Lokalisation und die genomische Struktur von *Sxl*

Über PCR wurden zwei Abschnitte des *Sxl*-Gens amplifiziert und gepoolt. Die gepoolte DNA diente als Sonde zur FISH an polytären Speicheldrüsen-Chromosomen aus *P. olivacea*. Die Hybridisierung erfolgte an Chromosomen getrennt, welche sowohl aus männlichen als auch aus weiblichen Larven stammten. Dabei konnten zwischen "männlichen" und "weiblichen" Chromosomensätzen keine Unterschiede bei Hybridisierungsergebnissen festgestellt werden. In Abbildung 27 ist exemplarisch eine Hybridisierung an einem Präparat aus einer Larve, deren Geschlecht nicht bestimmt wurde, gezeigt. Die Hybridisierung von *Sxl* ergab nur 1 Signal an Chromosom II Arm B', nach der Lage von *Sxl* in *C. thummi* bzw. *piger* auf Chromosom I Arm B (persönliche Mitteilung Christiane Kraemer, (Frank 2001)).

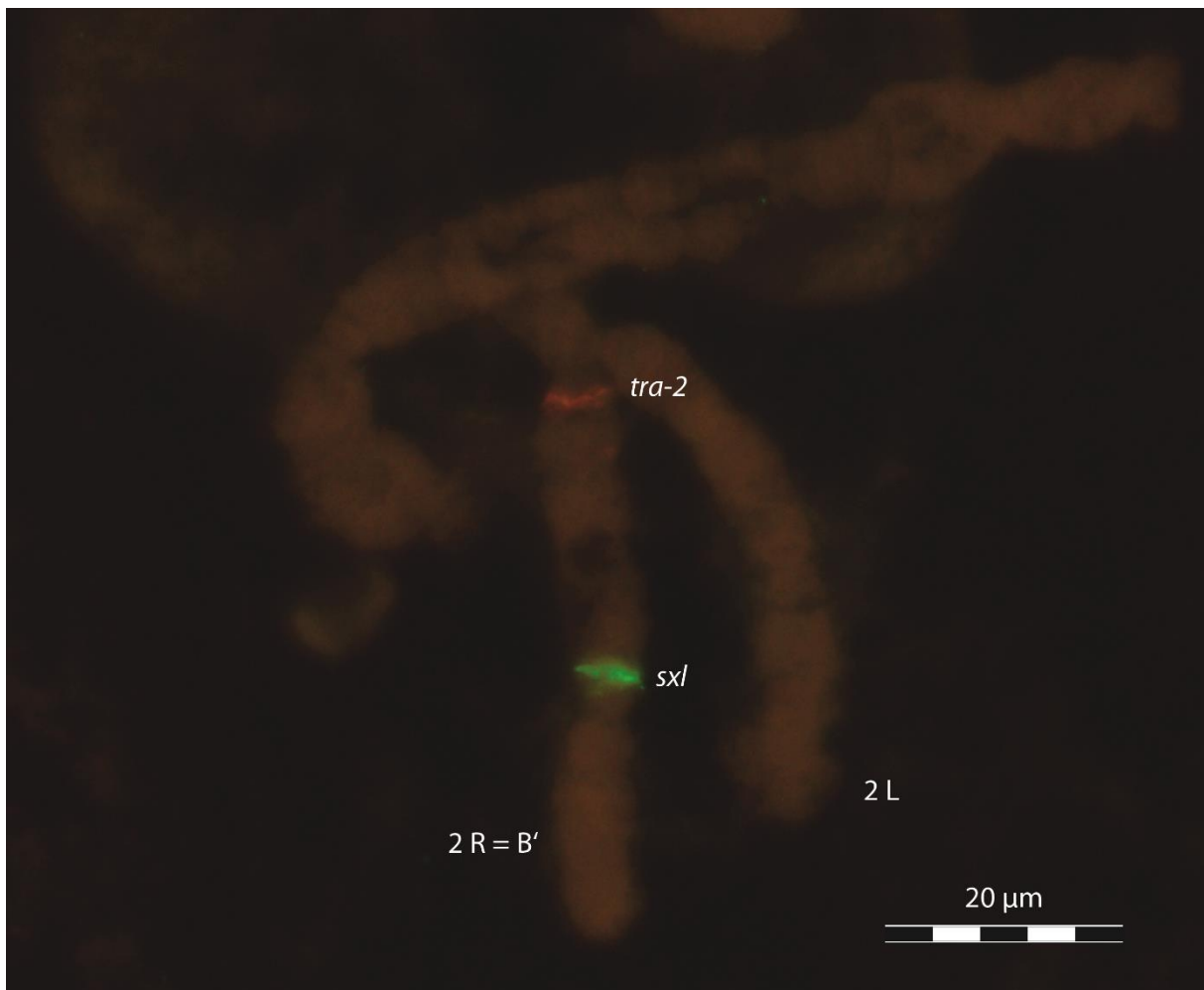


Abbildung 27 Mikroskopische Aufnahme des FISH-Ergebnis mit der Sonde *Sxl* an Polytänchromosomen aus *P. olivacea*. Die Hybridisierung der Sonde *Sxl* (grünes Signal, FITC) erfolgte an Chromosom II Arm R bzw. B'. Die Referenzsonde *tra-2* (rotes Signal, TRITC) hybridisierte ebenfalls an Chromosomarm 2R bzw. B'.

Transkription von *Sxl*

Eine RNA-Seq Analyse zeigt, dass *Sxl* in *P. olivacea* transkribiert und alternativ gespleißt wird. Exon 3 kann in vier Varianten vorkommen (369 bp, 360 bp, 327 bp oder 318 bp). Dies tritt sowohl im Weibchen-als auch im Männchen-Transkriptom auf. Dabei zeigt sich eine tendenziell 3 bis 9% stärkere Expression im Weibchen für das 318 bp-Exon 3 und im Männchen eine tendenziell 3 % stärkere Expression für das 360 bp-Exon 3. Exon 5 wird ebenfalls alternativ gespleißt. Entweder stellt es mit 2938 bp das letzte Exon der *Sxl*-mRNA dar oder es wird nach 161 bp durch ein weiteres Intron unterbrochen und enthält ein zusätzliches Exon 6 mit 1613 bp. Generell zeigt *Sxl* im Männchen eine leicht erhöhte Expression ($1 \times \sigma^7$ entspricht $0,8 \times \varphi$).

Das Gen *tra-2*

Identifikation und Analyse der Genstruktur von *tra-2*

Wie für das Gen *Sxl* wurde auch für *tra-2* eine TBLASTX-Suche durchgeführt. Daraus resultierte ein Contig aus dem Datensatz #468 (Tabelle 5). Der contig_283_asm_468_k19_b300 ist 106.137 bp lang. Das zum *D. melanogaster* homologe *tra-2*-Gen ist 1.489 bp lang (Abbildung 28, blauer Pfeil). Über RNA-Daten wurde eine potenzielle mRNA mit 1.225 bp bestimmt (Abbildung 28, grüner Pfeil). Die Gensequenz zeigt drei Introns, deren Spleißstellen verifiziert wurden. Die Suche nach einem offenen Leserahmen ergab drei mögliche Leserahmen. Dabei handelt es sich um zwei kurze und einen langen Leserahmen. Der längste Leserahmen mit 822 bp (Abbildung 28, gelber Balken) zeigt ein RRM (*e-value* von $8,77 \times 10^{-20}$) (Abbildung 28, hellblauer Pfeil).

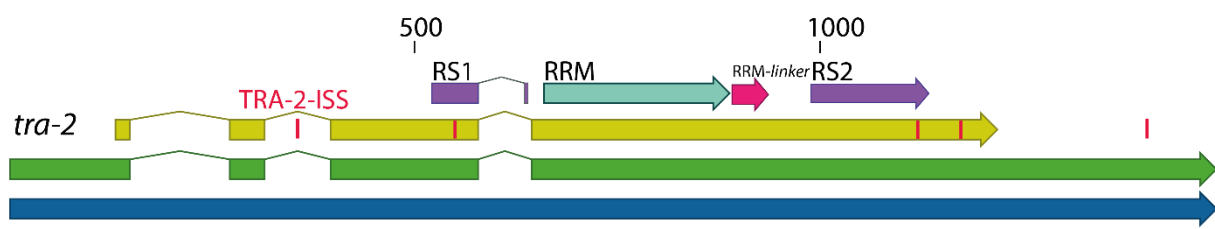


Abbildung 28 Die Genstruktur des *tra-2*-Gens aus *P. olivacea*. Das 1,5 kb lange *tra-2*-Gen (dunkelblau) ist mit dem am stärksten exprimierten Transkript (aus vier möglichen Transkripten) bestehend aus vier Exons (grün) und dem daraus resultierenden 822 bp-ORF (gelb) abgebildet. Das identifizierte RNA recognition motif (RRM, türkis) ist in Exon 4 lokalisiert, darauf folgt der RRM-linker (pink). Upstream bzw. downstream des RRM sind die RS1-bzw. die RS2-Domäne zu finden (violett). Die identifizierten TRA-2-ISS-Erkennungssequenzen sind in Rot eingezeichnet.

Dieser schließt sich eine 16 AS lange RRM-linker Region an (Abbildung 28, pinker Pfeil). Vor und hinter dem RRM erfolgt eine Abfolge von Arginin (R) bzw. Serin (S)-Dipeptiden, welche dementsprechend als RS1 und RS2 bezeichnet werden (Abbildung 28, dunkelviolette Pfeile). Das in *D. melanogaster* beschriebene ISS (*intronic splicing silencer*)-Motiv (Qi et al. 2006) wurde mittels BLASTN 5-mal innerhalb der Politra-2-Sequenz identifiziert. Ein TRA-2-ISS-Motiv ist im Intron 2 lokalisiert. Die übrigen vier Motive sind in Exon 3 oder 4 lokalisiert (Abbildung 28, rote vertikale Striche).

Lokalisation von *tra-2*

Basierend auf der Gensequenz von *tra-2* wurde mittels PCR ein Genabschnitt amplifiziert und als Sonde zur FISH an Polytänochrosomen aus den Speicheldrüsen von *P. olivacea* verwendet. Mit Hilfe der FISH (Abbildung 27 und Abbildung 29) kann das *tra-2*-Gen auf Chromosom II Arm B' lokalisiert werden. Es ist somit auf demselben Chromosomenarm lokalisiert wie *Sxl*. Im Gegensatz zu *Sxl* liegt es allerdings zentromernah.

Transkription von *tra-2*

Die RNA-Seq-Analyse für *tra-2* ergab vier mögliche Transkripte, die alternativ gespleißt werden. Im larvalen Transkriptom machen drei der vier Transkripte mit 11 % (σ) bzw. 18% (φ) nur einen geringen Anteil aus. Das dominante Transkript mit 1225 bp (Abbildung 28, grüner Pfeil) kodiert als einziges für den vollständigen 822 bp-ORF (Abbildung 28, gelber Pfeil) und die daraus resultierenden 274 AS. Die drei schwach vertretenen Transkripte besitzen nur einen verkürzten ORF (471 bp) und somit resultieren daraus je ein verkürztes Polypeptid mit einer verkürzten RRM-Domäne (vollständige RRM-Domäne in Abbildung 28, türkisfarbener Pfeil). In allen Transkripten wird immer Intron 3 gespleißt, jedoch Intron 1 oder 2 können alternativ gespleißt werden (Abbildung 48).

Das Gen *tra* und das Gen *l(3)73Ah*

In einigen Dipteren konnte das Gen *tra* über das Nachbar-Gen *l(3)73Ah* ermittelt werden (Pane et al. 2002), (Lagos et al. 2007) und (Ruiz et al. 2007). Daher wurde mittels BLAST das Gen *l(3)73Ah* im Assembly des Datensatzes #551 (Tabelle 5) in contig_300_asm_551_k61_b300 identifiziert. Der Contig umfasst 20 potenzielle Gene, wobei sich ein RRM-haltiges Gen 21 kb von *l(3)73Ah* entfernt bestimmen ließ. Dieses Gen zeigt zudem Ähnlichkeit zu SR-reichen Proteinen. Der signifikanteste Treffer der BLASTX-Suche zeigt einen e-value von 8×10^{-87} zu AGAP003369 -PA-like Protein aus *Anopheles sinensis*. Ein Alignment mit bekannten *tra*-Gensequenzen aus *D. melanogaster*, *Bactrocera oleae* (*B. oleae*), *Anastrepha obliqua* (*A. obliqua*) und *Ceratitis capitata* (*C. capitata*) (McKeown et al. 1987), (Lagos et al. 2007), (Ruiz et al. 2007) und (Pane et al. 2002) zeigt keine große Sequenzähnlichkeit.

Lokalisation von *l(3)73Ah*

Die Lokalisation des Gens *l(3)73Ah* erfolgte mittels FISH an polytären Speicheldrüsenchromosomen aus *P. olivacea*. Dabei zeigte sich eine Hybridisierung auf Chromosom II Arm B', auf dem zuvor schon die Gene *tra-2* und *Sxl* identifiziert wurden (Abbildung 29).

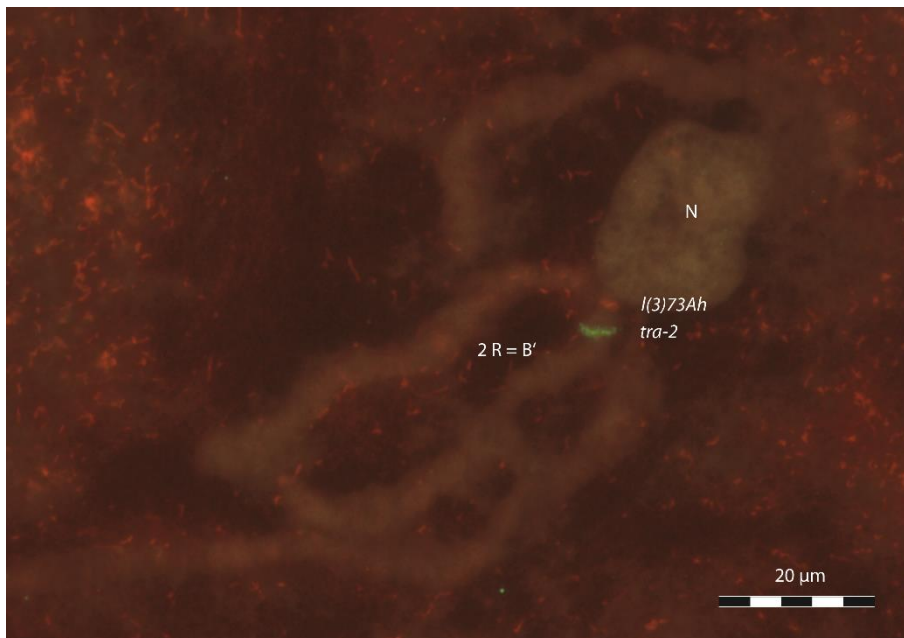


Abbildung 29 Mikroskopische Aufnahme der FISH an Polytänchromosomen aus Speicheldrüsen von *P. olivacea* mit den Sonden *tra-2* und *l(3)73Ah*. Die Hybridisierung der Sonde *tra-2* (grünes Signal, FITC) erfolgte an Chromosom II Arm R bzw. B'. Die Sonde *l(3)73Ah* hybridisierte (rotes Signal) an demselben Chromosomarm, jedoch proximal von *tra-2*.

Das Gen *dsx*

Identifikation und Charakterisierung der Genstruktur von *dsx*

Mittels TBLASTX-Funktion wurde ein zu *dsx* aus *D. melanogaster* homologes Gen, *Polidsx1*, auf contig_259 (aus ASM 729_kdef_b300) identifiziert (Abbildung 30).

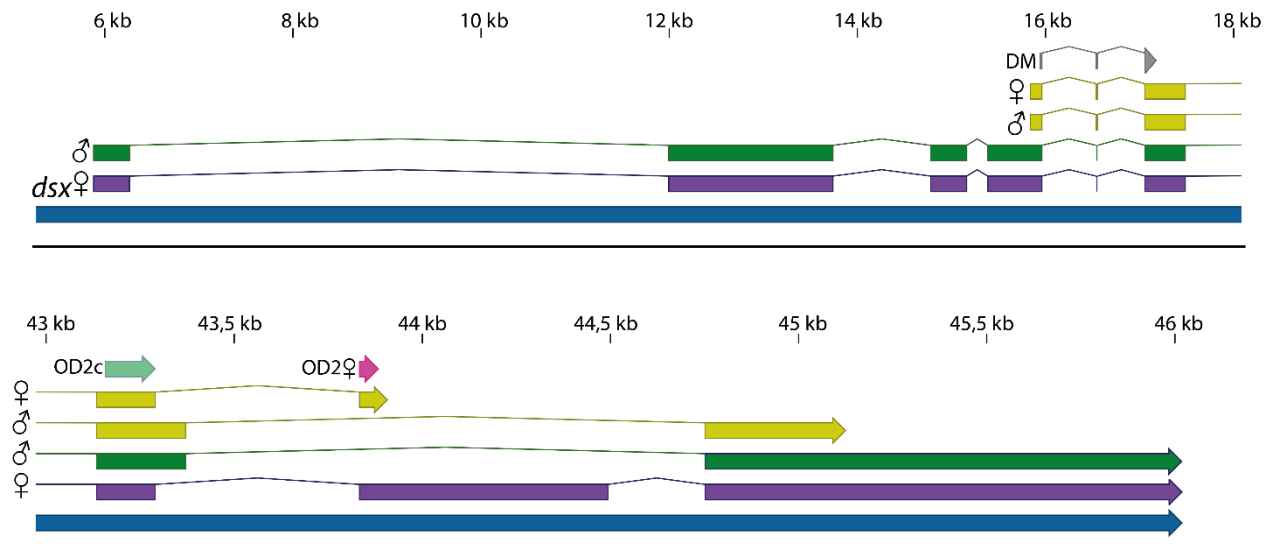


Abbildung 30 Die Genstruktur von *dsx1* (*Polidsx1*) aus *P. olivacea*. Die Abbildung zeigt die Region 6 kb bis 18 kb und die Region 43 kb bis 46 kb des *dsx*-Gens. Der obere Teil der Abbildung zeigt das geschlechtsunabhängig gespleißte 5'-Ende des Gens und der untere Teil das geschlechtsspezifisch gespleißte 3'-Ende des *dsx*-Gens. Die Annotation umfasst das Gen (blau), Weibchen-Transkript (lila), Männchen-Transkript (grün) und die jeweiligen ORFs (gelb). Zudem sind die Domänen, DM (DBD und OD1, in grau) und OD2 common (türkis) bzw. Weibchen-spezifisch (pink) annotiert.

Polidsx1 zeigt im Alignment der AS-Sequenz große Übereinstimmungen zur DM (DBD, OD1)-Domäne und zur OD2 common bzw. OD2 female. Zudem wurde ein weiteres Gen (*Polidsx2*), welches ebenfalls Sequenzähnlichkeit zu *dsx* aufweist, identifiziert. *Polidsx1* umfasst 40.144 bp auf contig_259 (Position 6680 bis 46823). Das Gen umfasst 9 mögliche Exons. Durch eine ORF-Suche konnten zwei Haupttranskripte von 5,1 kb (*Polidsx1T1*) bzw. 5,7 kb (*Polidsx1T2*) identifiziert werden. Die möglicherweise daraus abgeleiteten Proteine, PoliDSX1T1 bzw. 2 bestehen aus 403 bzw. 281 AS. Eine SMART-Domänen-Suche ergab für PoliDSX1T1 und T2 Treffer zur DM-Domäne (e -value $6,14 \times e^{-26}$) und zu DSX_Dimer (e -value $1,4 \times e^{-7}$). *Polidsx2* umfasst 2,6 kb mit bis zu fünf möglichen Exons. Die beiden Genkopien liegen 15,8 kb voneinander entfernt und sind in entgegengesetzter Orientierung angeordnet.

Lokalisation von *dsx* mittels FISH

Ein *Polidsx1*-Klon wurde mittels FISH an polytären Speicheldrüsenchromosomen aus *P. olivacea* zur Lokalisation von *Polidsx1* und *Polidsx2* verwendet, da *Polidsx2* 15,8 kb von *Polidsx1* getrennt liegt und somit nicht separat lokalisiert werden kann. Die Hybridisierung erfolgte an Chromosom I Arm F' (Abbildung 31).

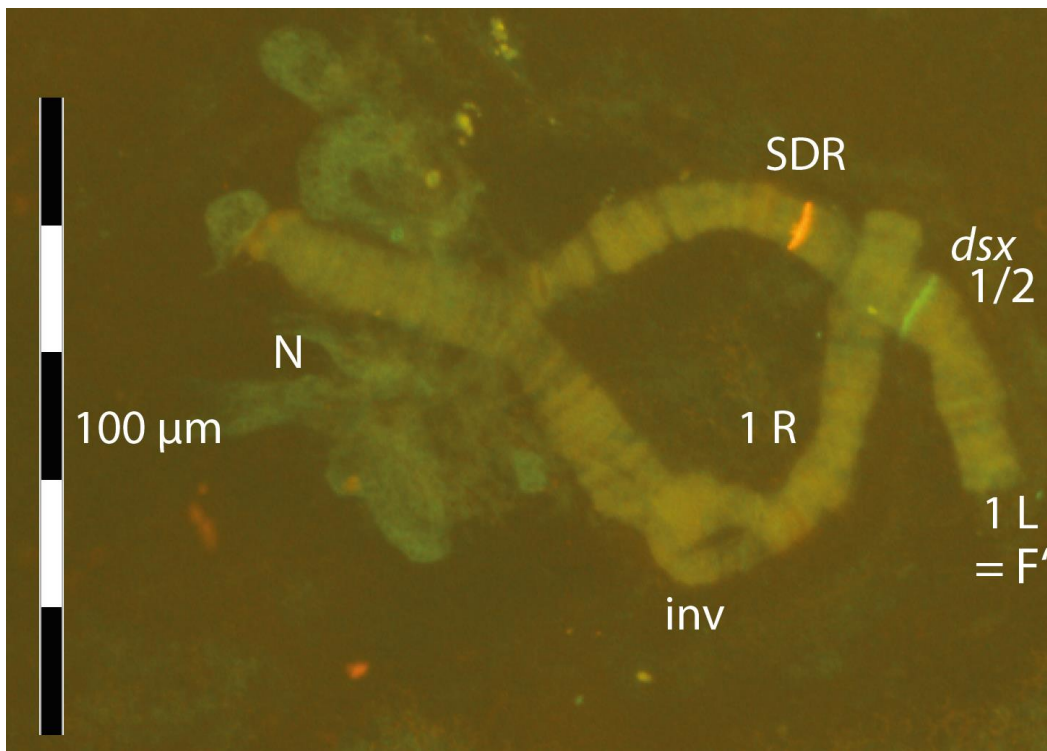


Abbildung 31 Mikroskopische Aufnahme des FISH Ergebnis mit der Sonde *Polidsx1* an Polytänchromosomen aus *P. olivacea*. Die Sonde *Polidsx1* hybridisierte (grünes Signal, FITC) an Chromosom I Arm L bzw. F'. Die Referenzsonde potenzielle SDR hybridisierte (rotes Signal, TRITC) ebenfalls am selben Chromosomenarm, jedoch proximal von der *Polidsx1*-Sonde.

Geschlechtsspezifische Prozessierung des 3'Ende von *Polidsx1*

Ein Vergleich der genomischen Sequenz von *Polidsx1* mit RNA -Daten aus weiblichen bzw. männlichen *P. olivacea*-Larven zeigt, dass das Gen geschlechtsspezifisch alternativ-gespleißt wird. Der C-Terminus des abgeleiteten Proteins unterscheidet sich in Männchen und Weibchen stark.

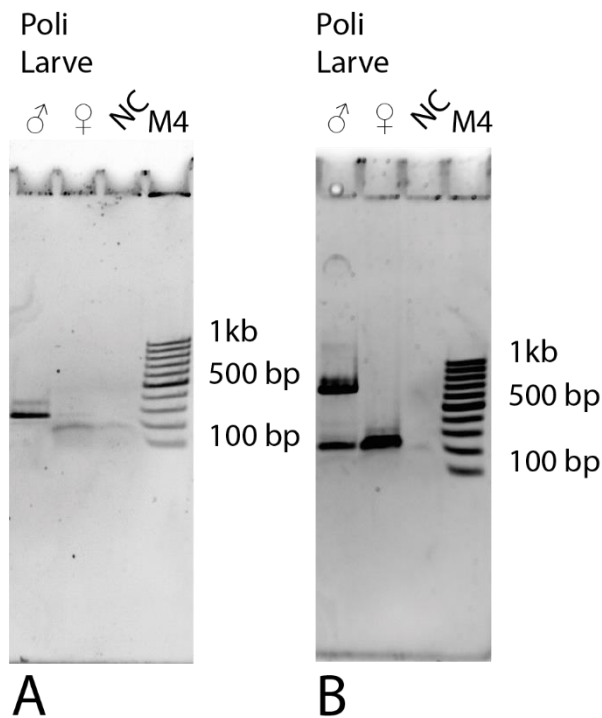


Abbildung 32 Gelelektrophoretische Auftrennung der RT-PCR, in welcher das Männchen-spezifische 3'Ende (A) bzw. das Weibchen-spezifische 3'Ende (B) amplifiziert wurde. A Männchen-spezifische RT-PCR zur Amplifikation des 3'-Endes. B Weibchen-spezifische RT-PCR zur Amplifikation des 3'-Endes. Als Template diente isolierte Gesamt-RNA aus weiblichen bzw. männlichen *P. olivacea*-Larven, welche zuvor zu cDNA umgeschrieben wurde. NC: Negativ-Kontrolle Als Molekulargewichtsstandard (M4) diente der GeneRuler 100 p DNA Ladder (Fa. Life Technologies)

Das 5,702 kb lange Weibchen-spezifische Transkript umfasst Exon 1-4 (5'-UTR), Exon 4-7 (geschlechtsunabhängig kodierender Bereich) und Weibchen-spezifisch kodierendes Exon 8 und Exon 9 (3'-UTR). Der daraus resultierende ORF ist 861 bp lang und kodiert für 287 AS. Die Translation wird durch ein zweifaches-Stopp-Kodon (Amber-Opal, Position: 3072, Abbildung 33) beendet. Das Männchen-spezifische Transkript umfasst 5,118 kb, wobei der C-Terminus aus dem gemeinsamen Exon 7 (kodierend) und Exon 9 (kodierend und 3'-UTR) besteht. Der identifizierte ORF ist 1224 bp und kodiert für ein potenziell 408 AS-Protein. Auf den Männchen-spezifischen ORF folgt ein Stopp-Kodon (Amber) an Position 3435. Die bioinformatischen Daten wurden experimentell mittels RT-PCR verifiziert. Über eine RT-PCR mit Männchen-spezifischen Primern wurde ein Amplifikat von 230 bp auf männliche cDNA generiert (Abbildung 32, A). Dieses Primerpaar könnte auch im Weibchen-Transkript binden, allerdings wäre das resultierende Produkt mit 814 bp recht groß und somit schwierig zu amplifizieren. Auf cDNA aus weiblichen *P. olivacea*-Larven wurde ein PCR-Produkt mit 175 bp erwartet, welches Weibchen-spezifisch ist (Abbildung 32, B). Im Männchen sollte laut NGS-Daten für dieses Primerpaar kein Produkt amplifiziert werden. In Abbildung 32, B sind

jedoch zwei PCR-Produkte (~700 bp & ~200 bp), welche aufgrund des Männchen-Template amplifiziert wurden, zu sehen. Die Sequenzierung der beiden Amplifikate ergab eine Übereinstimmung der 700 bp-Bande zur genomischen Sequenz, wogegen die 200 bp-Bande mit der Sequenz des Weibchen-Transkripts übereinstimmt.

Spleißmechanismus in *Polidsx1*

Die Spleiß-Akzeptor- und Donor-Stellen der in *Polidsx1* 3' gelegenen Exons sind in Tabelle 10 aufgelistet. Für die 3'Spleiß-Akzeptoren aller annotierten *P. olivacea*-Gene (Tabelle 6) wurde bioinformatisch ein Durchschnitt von 4,02 Purine bzw. 7,98 Pyrimidinen berechnet.

Tabelle 10 3'Spleiß-Akzeptorstellen der 3'Region des *Polidsx1*-Gens

Gen	Art der Spleißstelle	Exon	Position im Gen [bp]	Länge	Sequenz [((Y) _n) ₁₂ - N-YAG]	# der Y im (Y) _n
<i>Polidsx1</i>	3'Akzeptor-Spleißstelle	Exon 7.1	34124..34139	16 nt	GTTTTCAATAAT-T-TAG	7/12
	3'Akzeptor-Spleißstelle	Exon 7.2	34142..34157	16 nt	TATTTTAAAAAT-T-CAG	6/12
	3'Akzeptor-Spleißstelle	♀-Exon 8	34823..34838	16 nt	TTTTTAAACTTC-A-TAG	9/12
	3'Akzeptor-Spleißstelle	♂/common Exon 9	35741..35756	16 nt	GAAATTCTATCT-A-CAG	7/12

Ein Vergleich des Durchschnittswerts bzw. der Konsensus-Sequenz zeigen in den 3'Spleißakzeptoren der 3'Region des *Polidsx1*-Gens keine signifikanten Abweichungen wie sie in *D. melanogaster* (Lynch und Maniatis 1995) beschrieben wurden (Tabelle 10).

Der Vergleich mit bekannten cis-regulatorischen Sequenzen innerhalb des Weibchen-spezifischen Exons ergab drei mögliche *purin rich elements* (PRE) und sechs TRA/TRA-2-Motive (Abbildung 33,). Dabei liegen alle identifizierten Sequenzen, außer ein PRE, im 3'-UTR des Weibchen-Transkripts.

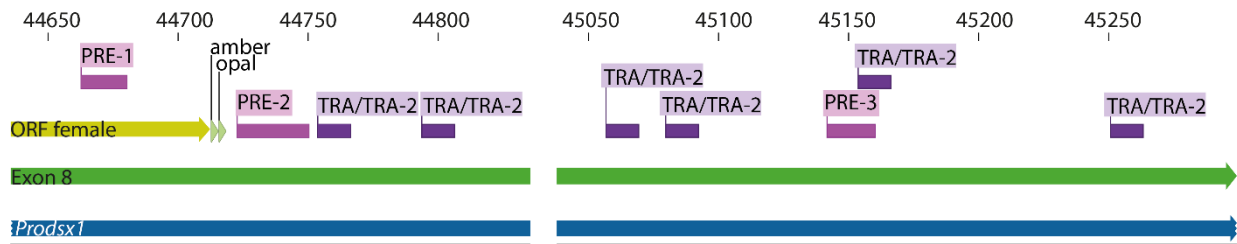


Abbildung 33 **Potenzielle cis-regulatorische Sequenzen im Weibchen-spezifischen Exon des Polidsx1-Gens.** Dargestellt ist der ORF (gelb), die mRNA (grün) und das Gen (blau). Zudem sind die potenziell cis regulatorischen Sequenzen in lila eingezeichnet (purin rich element: PRE und TRA/TRA-2). Das zweifaches-Stopp-Kodon (amber, opal) ist in hellgrün eingezeichnet.

Fünf der sechs TRA/TRA-2-Motive zeigen eine Übereinstimmung von über 69 % zu den Motiven aus *D. melanogaster*. Nur eins der sechs identifizierten Motive zeigt eine geringe Übereinstimmung von 61 %. Dabei wurde mittels bioinformatischen Analysen die Sequenz der sechs identifizierten Motive in allen in *P. olivacea* identifizierten Intron-Sequenzen innerhalb eines Fensters von 664 bp identifiziert. Kein weiteres Fenster zeigte eine signifikante Übereinstimmung zu den TRA/TRA-2-Motiven.

Tabelle 11 In *Polidsx1* identifizierte cis-regulatorische Elemente im 664 bp-Weibchen-spezifischen Exon

Gen	Element	Position im 664 bp-Exon	Länge	Sequenz	Identität	Art mit signifikantester Übereinstimmung
<i>Polidsx1</i>	TRA/TRA-2_1	119-131	13 nt	ACAAAAATCAACT	76,9	<i>Aaeg</i>
	TRA/TRA-2_2	159-171	13 nt	AGTCAAATCAACA	69,2	<i>Dmel, Aaeg</i>
	TRA/TRA-2_3	422-434	13 nt	TCAACAATCAACA	100	<i>Dmel</i>
	TRA/TRA-2_4	445-457	13 nt	GCTTCAATCTCTC	76,9	<i>Cqui</i>
	TRA/TRA-2_5	519-531	13 nt	AATTCAATCAACA	84,6	<i>Dmel, Aaeg</i>
	TRA/TRA-2_6	616-628	13 nt	TCGGCGTTCAAAA	69,2	<i>Agam</i>

Die identifizierten PREs weisen über eine Länge von 18 bis 28 Nukleotiden nur 4 oder 5 Pyrimidine auf (Tabelle 12, Abbildung 33).

Tabelle 12 Weitere in *Polidsx1* identifizierte cis-regulatorische Elemente im 664 bp-Weibchen-spezifischen Exon

Gen	Element	Position im 664 bp-Exon	Länge	Sequenz	Y/N
<i>Polidsx1</i>	PRE-1	28-45	18 nt	GAAAAAATAATATAAGCA	4/18
	PRE-2	88-115	28 nt	GAAAATATTAATAACGAAGAAAAAAAAA	5/28
	PRE-3	507-525	19 nt	AAAGGAAGCAAAAATTCAA	4/19

Zwei der PREs und zwei TRA/TRA-2-Motive wurden in der 5'Region des Weibchen-spezifischen Exons 8 lokalisiert. Das dritte PRE und die vier übrigen TRA/TRA-2-Motive wurden im 3'Bereich des Weibchen-spezifischen Exons 8 identifiziert (Abbildung 33).

Weitere cis regulatorische Elemente (wie RBP1b, NvTRA/TRA-2 oder TRA/TRA-2-ISS) aus *D. melanogaster* oder *Nasonia vitripennis* zeigten keine signifikanten Übereinstimmungen innerhalb Exon 8.

Geschlechtsunabhängiges-alternatives Spleißen von *Polidsx1*

Weiter ergab der Vergleich von RNA-Daten aus *P. olivacea*-Larven alternative Spleiß-Varianten, welche geschlechtsunabhängig erfolgen. Nur Exon 5 und 9 werden nicht alternativ gespleißt. Exon 1 zeigt sechs Varianten, ebenso wie Exon 7. Exon 3 zeigt mit mindestens acht Varianten die größte Varietät. Exon 2 und 6 können in fünf unterschiedlichen Varianten auftreten, wogegen Exon 4 nur in vier möglichen Variationen gespleißt werden kann. Die Vielzahl an Varianten wurde experimentell mittels RT-PCR oder mittels bioinformatischer Auswertung von RNA-Seqs identifiziert. Nicht alle Varianten sind experimentell und bioinformatisch und in beiden Geschlechtern abgedeckt. Die genauen Daten der Exon-Varianten sind im elektronischen Anhang zu finden.

Geschlechtsspezifisches-alternatives Spleißen von *Polidsx2*

Die *dsx*-Genkopie *Polidsx2* wird ebenfalls geschlechtsspezifisch gespleißt. Im Weibchen gibt es drei mögliche, im Männchen zwei mögliche Transkripte. Transkript 1(2 kb) gibt es in beiden Geschlechtern, die restlichen Transkripte sind geschlechtsspezifisch. Die Suche nach einem ORF zeigt für alle Transkripte jedoch nur einen in allen Transkripten identischen ORF von 633 bp.

Zielgene der Geschlechtsdifferenzierung von *dsx*

Um die Funktion von *dsx* in der Geschlechtsdifferenzierung zu verstehen, wurden in vielen Insekten *dsx*-Zielgene untersucht. Von Verhulst und van de Zande wurden alle bis dahin identifizierten Gene zusammengefasst (Verhulst und van de Zande 2015). Nach der Identifizierung der *dsx*-Zielgene aus anderen Insekten in *P. olivacea* wurde eine Expressionsanalyse durchgeführt. Lediglich das Gen *SP1* (*sex specific storage protein-1*, auch *hexamerin-2β*) (Suzuki et al. 2003) konnte in *P. olivacea* als im larval Transkriptom geschlechtsspezifisch exprimiert identifiziert werden. Im Weibchen-Larven-Transkriptom findet eine 5,11-fach höhere Expression statt (Normalisiert auf *reads per kilobase per million*, *rpkm*).

Weitere Gene der Geschlechtsbestimmungskaskade aus *D. melanogaster*

Über *Sxl*, *tra-2* und *dsx* hinaus konnten weitere Gene aus der Geschlechtsbestimmungskaskade von *D. melanogaster* identifiziert werden. Es handelt sich dabei um *emc*, *gro*, *runt*, *da*, *fru*, *fl(2)d*, *intersex (ix)*, *vir*, *snf* und *dis* (Tabelle 13). Für die Gene *sis A/B/C*, *her*, *dpn* und *tra* wurde kein homologes Gen in *P. olivacea* identifiziert. Eine Übersicht der genannten Gene erfolgt in Tabelle 13.

Tabelle 13 Gene welche in *D. melanogaster* in der Geschlechtsbestimmung beteiligt sind

Gen	Gen-Symbol	Art	Zelltyp	Funktion	homologes Gen in <i>P. olivacea</i>
<i>runt</i>	-	Numerator	Soma		✓
<i>sisterless A B C</i>	<i>sis A B C</i>	Numerator	Soma	Aktiviert ♀-p _e SXL	
<i>deadpan</i>	<i>dpn</i>	Denominator	Soma	TF, neg. Reg. von <i>Sxl</i>	
<i>extra macrochaetae</i>	<i>emc</i>	maternal	Soma	negative Reg. von <i>Sxl</i>	✓
<i>groucho</i>	<i>gro</i>	maternal	Soma	negative Reg. von <i>Sxl</i>	✓
<i>daughterless</i>	<i>da</i>	maternal	Soma	positive Reg. von <i>Sxl</i>	✓
<i>hermaphrodite</i>	<i>her</i>	maternal	Soma	positive Reg. von <i>Sxl</i>	
<i>female lethal (2) d</i>	<i>fl(2)d</i>	SR-Protein	Soma	Autoreg. von p _m <i>Sxl</i>	✓
<i>sans fille</i>	<i>snf</i>	RRM	Soma	Autoreg. von p _m <i>Sxl</i>	✓
<i>virilizer</i>	<i>vir</i>	NS-binding	Soma	Autoreg. von p _m <i>Sxl</i> , blockiert <i>splice site</i> Exon 3	✓
<i>intersex</i>	<i>ix</i>	T Coaktivator	Soma	Aktivität von DSX ^F	✓
<i>fruitless</i>	<i>fru</i>	TF	ZNS	Sexualverhalten	✓
<i>dissatisfaction</i>	<i>dis</i>	TF	ZNS	Sexualverhalten	✓
<i>Sex-lethal</i>	<i>Sxl</i>	RRM	S, Z, K	Schlüsselgen	✓
<i>transformer</i>	<i>tra</i>	SR	S, Z	Spleißfaktor	
<i>transformer-2</i>	<i>tra-2</i>	SR	S, K	Spleißfaktor	✓
<i>doublesex</i>	<i>dsx</i>	TF	Soma	Anschaalten Differenzierungsgene /Doppelschalter	✓

Das Gen *fru* zeigt ähnlich wie *Polidsx1* viele Transkript-Varianten im Larven-Transkriptom von Männchen und Weibchen. Dabei zeigt sich eine signifikante differenzielle Genexpression zwischen den beiden Geschlechtern (Tabelle 14).

Tabelle 14 Expressionsanalyse der homologen Gene aus der Geschlechtsbestimmungskaskade von *D. melanogaster* in *P. olivacea*. Die nicht differenziell exprimierten Gene sind grün unterlegt, die im Weibchen hoch exprimierten Transkripte sind rosa unterlegt, die im Männchen blau.

Gen	Gen-Symbol	Expression (normalisiert)		fold change	Transkript [Bp]
		female	male		
<i>runt</i>		44.425	47.503	~1	2904
<i>extra macrochaetae</i>	<i>emc</i>	262.993	204.626	1,29	1851
<i>groucho</i>	<i>gro</i>	117.938	117.459	1	3528
<i>daughterless1</i>	<i>da1</i>	88.225	60.966	1,45	3881
<i>daughterless2</i>	<i>da2</i>	171.065	208.631	0,8	3950
<i>female lethal (2) d</i>	<i>fl(2)d</i>	27.789	32.123	0,86	899
<i>sans fille</i>	<i>snf</i>	45	72	0,63	798
<i>virilizer</i>	<i>vir</i>	144.045	164.408	0,88	5243
<i>intersex</i>	<i>ix</i>	13.077	16.700	0,78	674
<i>fruitless1</i>	<i>fru1</i>	17.260	13.335	1,29	4707
<i>fruitless2</i>	<i>fru2</i>	721	3.408	0,21	4827
<i>dissatisfaction</i>	<i>dis</i>	240	681	0,35	1745

Im Männchen wird *Polifru1* höher exprimiert, im Weibchen dagegen *Polifru2*. Die Transkripte unterscheiden sich im 5'-UTR, jedoch nicht im ORF. Zudem wird das Gen *da* alternativ gespleißt. Die zwei Transkript-Varianten zeigen eine differenzielle Expression. *Da1* wird im Weibchen höher exprimiert, wogegen *da2* eine erhöhte Expression im Männchen zeigt. Das Gen *emc* wird im Weibchen leicht höher exprimiert. Die Gene *fl2d*, *snf*, *vir*, *ix* und *dis* werden dagegen im Männchen-Transkriptom höher exprimiert.

Die übrigen oben genannten Gene zeigten keine geschlechtsspezifischen Unterschiede hinsichtlich ihrer Expression.

3.6.2 Identifikation von homologen Genen aus der geschlechtsbestimmende Region von *C. thummi*

Mittels TBLASTX gelang es, 10 Gene aus der SDR von *C. thummi* in *P. olivacea* auf einem Contig zu identifizieren. Über FISH (Abbildung 34) konnte gezeigt werden, dass diese Gene in Bezug auf *C. thummi* in invertierter Abfolge vorliegen. Der Locus der potenziellen SDR konnte ebenfalls über FISH bestimmt werden auf Chromosom I Arm L (Abbildung 31; Abbildung 34). Dieser wird daher weiter als F' bezeichnet, da die SDR in anderen (bzw. in der Mehrzahl der) Chironomiden-Arten auf Arm F identifiziert wurde (Martin et al. 1980).

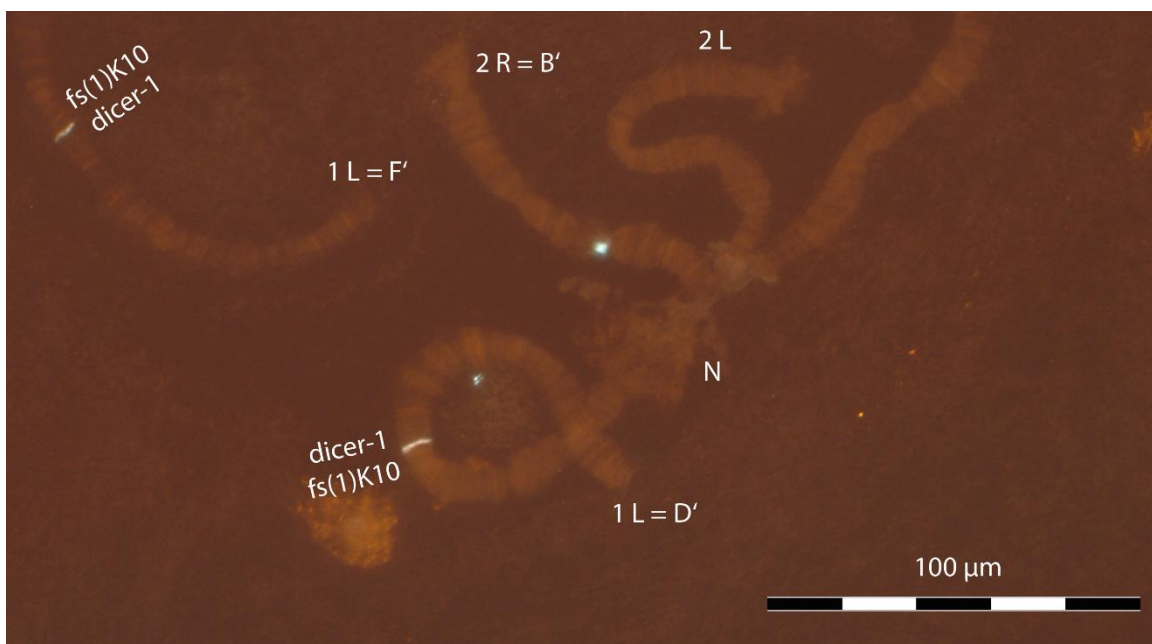


Abbildung 34 Mikroskopische Aufnahme nach der FISH an Speicheldrüsenchromosomen aus *P. olivacea* mit den Sonden *fs(1)K10-like* und *dicer-1*. Die Hybridisierung der Sonde *fs(1)K10-like* (rotes Signal, TRITC) erfolgte an Chromosomenarm 1L bzw. F' proximal im Vergleich zu der Hybridisierung der Sonde *dicer-1* (grünes Signal, FITC). Abgebildet ist ein Ausschnitt des Präparats, welcher zweimal den Chromosomenarm 1L zeigt. N: Nukleolus

Vergleich der chromosomalen Organisation der SDR-Gene und benachbarter Gene

Die homologe Region zwischen der SDR aus *C. thummi* und der potenziellen SDR aus *P. olivacea* erstreckt sich von *mesoderm induction early response 1-like* (*mi-er-1-like*) bis zu *Hermansky-Pudlak-syndrom-like* (*hps-like*, Abbildung 35). Die jeweiligen Bruchpunkte der homologen Region wurden mittels PCR analysiert und zusätzlich mit genomischer NGS-Daten via Mapping verifiziert (Daten nicht gezeigt). Die homologe Region ist in Abbildung 35 dargestellt (violett unterlegt). Sie umfasst in *C. thummi*-Weibchen knapp 70 kb und in *P. olivacea* 46,5 kb. In *P. olivacea* liegt zudem das Gen *6PGD* innerhalb der homologen Region,

so dass keine vollständige Syntäne zu *C. thummi* existiert. Daher umfasst der tatsächlich syntäne Bereich nur 44 kb.

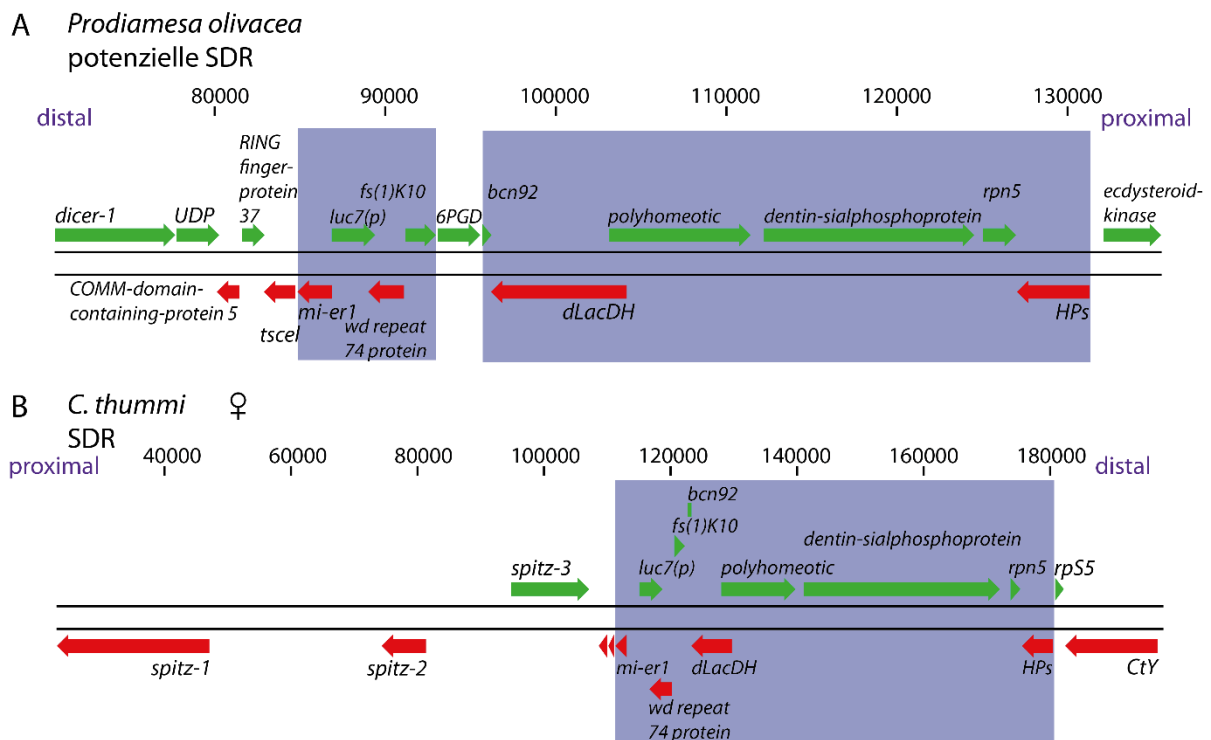


Abbildung 35 Ausschnitt der SDR aus *C. thummi* und der potenziellen SDR aus *P. olivacea* welcher die syntäne Region (blau) umfasst. A zeigt die potenzielle SDR aus *P. olivacea* mit den identifizierten Genen (grün = Plusstrang, rot = Minusstrang) von 70,6 bis 135,5 kb der potenzielle SDR (64,9 kb). B zeigt einen Ausschnitt der SDR aus *C. thummi* Weibchen von 22,9 bis 197,9 kb (174,9 kb), die identifizierten Gene sind hier für den Plus- und den Minusstrang annotiert. Die syntänen Regionen sind blau unterlegt.

Die Gene, welche an den jeweiligen Bruchpunkten des syntänen Bereichs liegen, wurden mittels BLASTX in der Art, in der sie dort nicht auftreten, identifiziert.

Identifikation SDR-angrenzender Gene aus *C. thummi* bzw. *piger* in *P. olivacea*

Das in *C. thummi/piger* distal auf Chromosom III, Arm F gelegene Gen *CtY* bzw. *CpY* wurde in *P. olivacea*-Contigs identifiziert und mittels FISH an polytären Speichdrüsenchromosomen aus *P. olivacea*-Larven mit einer spezifischen *CpY*-Sonde lokalisiert (Abbildung 36). Das gleiche wurde für das proximal gelegene Gen *spitz-like* durchgeführt (Abbildung 36). Beide Gene konnten auf Chromosom I, Arm F' wie zuvor die potenzielle SDR lokalisiert werden. Die beiden Gene liegen jedoch proximal bzw. zentromernah im Vergleich zur potenziellen SDR. Zudem liegen beide Gene relativ nah aneinander, wobei *spitz-like* proximal von *CpY* liegt.

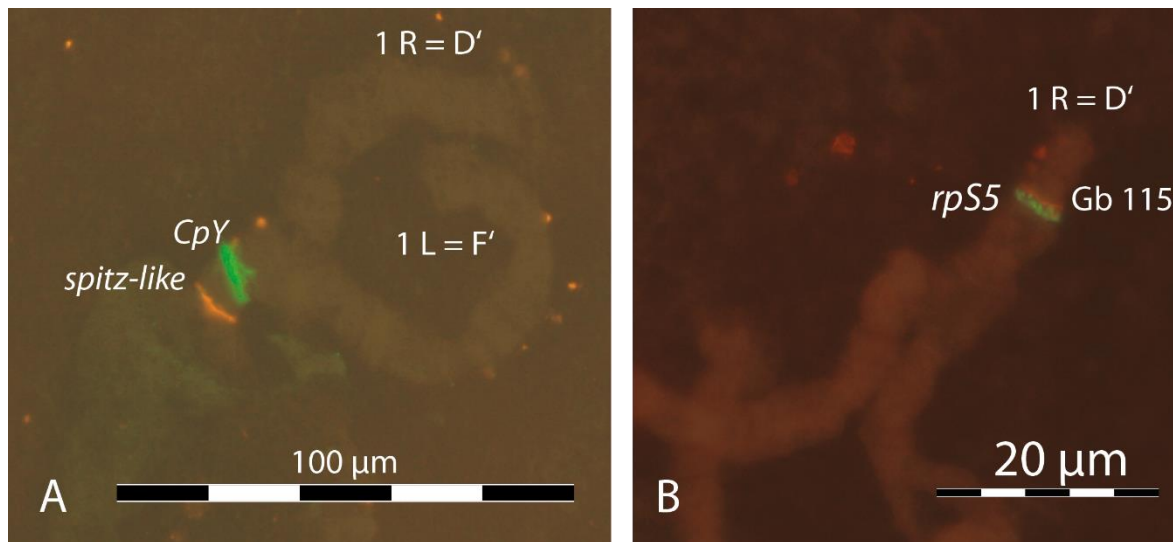


Abbildung 36 Mikroskopische Aufnahme des Ergebnisses einer FISH an Polytäanchromosomen aus *P. olivacea* mit den Sonden CpY und *spitz-like* (A) und den Sonden *rpS5-like* und Gb_115 (B). Die Hybridisierung der Sonden CpY und *spitz* erfolgte an Chromosom I Arm L bzw. F' wie zuvor die potenzielle SDR. Die Hybridisierung der Sonde CpY (grünes Signal, FITC) erfolgte distal im Vergleich zu der Hybridisierung der Sonde *spitz-like* (rotes Signal, TRITC), welche zentromernah hybridisierte. Die Hybridisierung der Sonde *rpS5* (grünes Signal, FITC) erfolgte an Chromosom I Arm R proximal von der bereits lokalisierten Sonde Gb_115 (rotes Signal, TRITC).

Der Contig mit dem BLAST-Treffer für *spitz-like* umfasst in 41 kb mindestens vier Gene, darunter zwei *spitz-like*-Kopien. Der Contig mit dem BLAST-Treffer für CpY ist dagegen mit 225 kb deutlich länger und umfasst mind. 29 Gene.

Die Annotation der Gene und die darauffolgende Expressions-Analyse der potenzielle SDR zeigten keine Unterschiede zwischen Männchen und Weibchen.

Weiter wurde das Gen *rpS5-like*, welches in der SDR von *C. thummi* aber nicht in der potenziellen SDR von *P. olivacea* liegt, identifiziert. Der 114,8 kb-Contig umfasst 21 Gene inklusive *rpS5-like*. Eine FISH mit einer spezifischen Sonde an polytänen Speicheldrüsenchromosomen aus *P. olivacea* zeigte eine Hybridisierung an Chromosom I, Arm D' (Abbildung 36). Als Referenzsonde diente Gb_115 (Abbildung 16, C), diese hybridisiert distal von der Sonde *rpS5-like*.

Das Gen *rpS5-like* wird durch zwei Intron-Sequenzen unterbrochen und zeigt in der Expressions-Analyse keine Unterschiede zwischen *P. olivacea*-Weibchen-Larven und Männchen-Larven.

Identifikation SDR-benachbarter Gene aus *P. olivacea* in *C. thummi/piger*

Weiter wurde das Gen *6PGD* in *C. thummi/piger* identifiziert und chromosomal lokalisiert. Der längste Contig umfasst 637,6 kb und stammt aus *C. thummi*. In *C. piger* wurde die Sequenz des Gens ebenfalls ermittelt. Eine FISH an Speicheldrüsenchromosomen aus *C. thummi/piger* zeigte eine Hybridisierung der Sonde *6PGD* an Chromosom IV, Arm G (Abbildung 37, A).

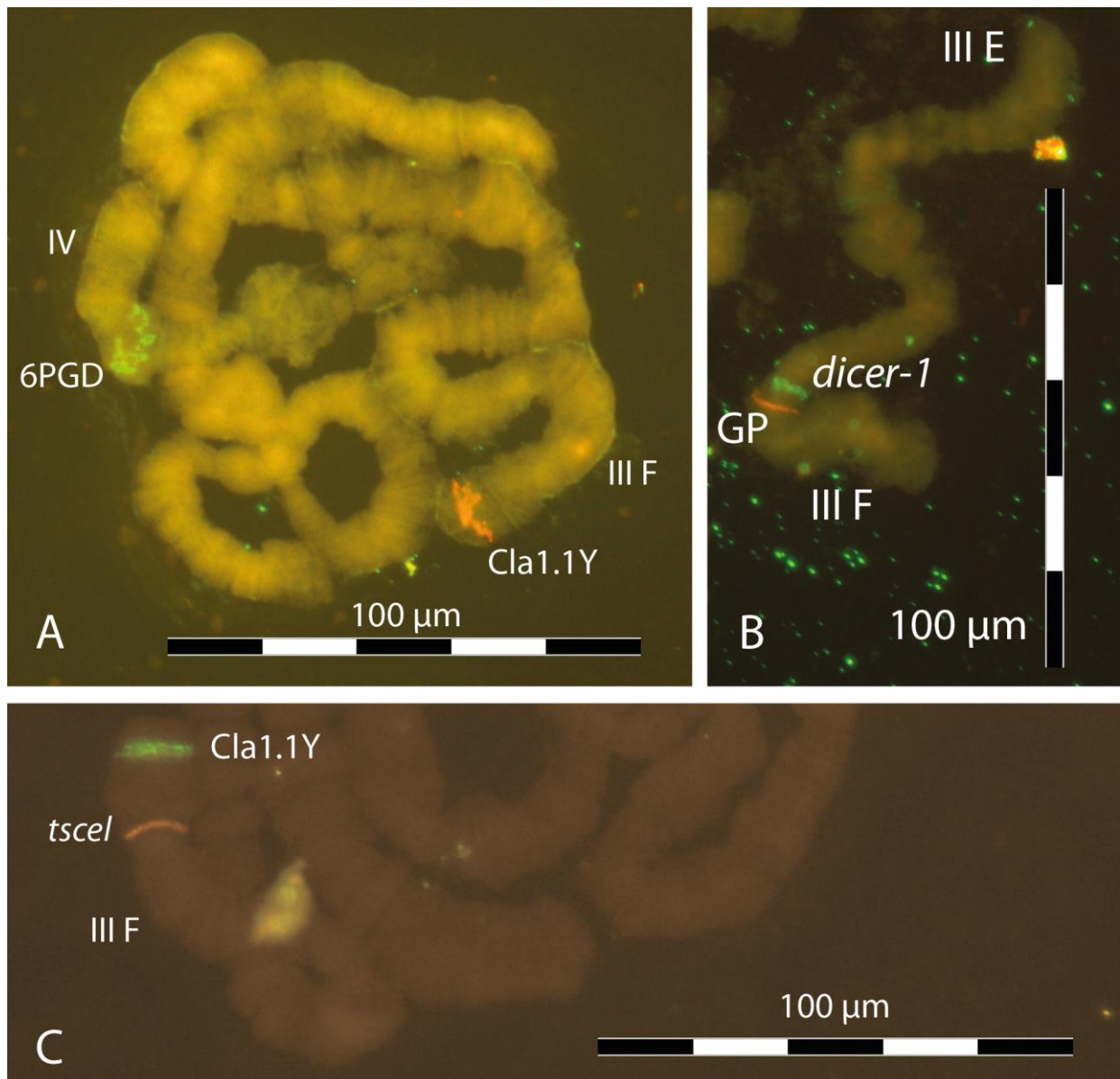


Abbildung 37 Mikroskopische Aufnahmen des Ergebnisses einer FISH an *C. thummi* Speicheldrüsenchromosomen mit den Sonden *6PGD* (A), *Cla1.1Y* (A, C), *dicer-1* (B) Gencluster-*piger* (GP) (B) und *tscel* (C). A: Die Hybridisierung der Sonde *6PGD* erfolgte am vierten und kleinsten Chromosom (grünes Signal, FITC). Der Klon *Cla1.1Y* diente als Referenzsonde und hybridisierte, wie bereits bekannt, an Chromosom III Arm F (rotes Signal, TRITC) (Kraemer und Schmidt 1993). B: Die Hybridisierung der Referenzsonde *dicer-1* erfolgte an Chromosom III Arm F (rotes Signal). Die Sonde GP hybridisierte distal von *dicer-1* auf dem selben Chromosomarm C: Die Sonde *tscel* hybridisiert (rotes Signal, TRITC) an Chromosomarm III F, ebenso wie die Referenzsonde *Cla1.1Y* (grünes Signal, FITC)..

Zudem wurden die in *P. olivacea* distal an die potenzielle SDR angrenzenden Gene *dicer-1*, *UDP* und *tsc1* in *C. thummi/piger* identifiziert und lokalisiert. Die Gene liegen nicht auf einem Contig wie in *P. olivacea* (Abbildung 35, A), sondern auf drei Contigs verteilt. Die Lokalisation via FISH zeigte jedoch, dass alle drei Gene auf einem Chromosomenarm liegen. Das Hybridisierungsergebnis an Chromosomenarm III F Region D2 (nach (King 1975)) für *tsc1* ist in Abbildung 37 C abgebildet.

Die in *P. olivacea* proximal der potenziellen SDR gelegenen Gene *ecdysteroid kinase (ek)*, *venom dipeptidyl peptidase 4-like (vdp4)*, *conserved oligomeric Golgi complex subunit 3-like (coGcs3)*, *casein kinase I isoform gamma-3-like (ckIg-3)*, *leucine rich repeat -like (LLR) 2* und *major facilitator superfamily domain containing protein 6-like (mfsdcp6)* finden sich in *C. piger* auf einem 25.638 bp-Contig (contig_16311_Cpi_female), jedoch in anderer Reihenfolge, wieder (Abbildung 38).

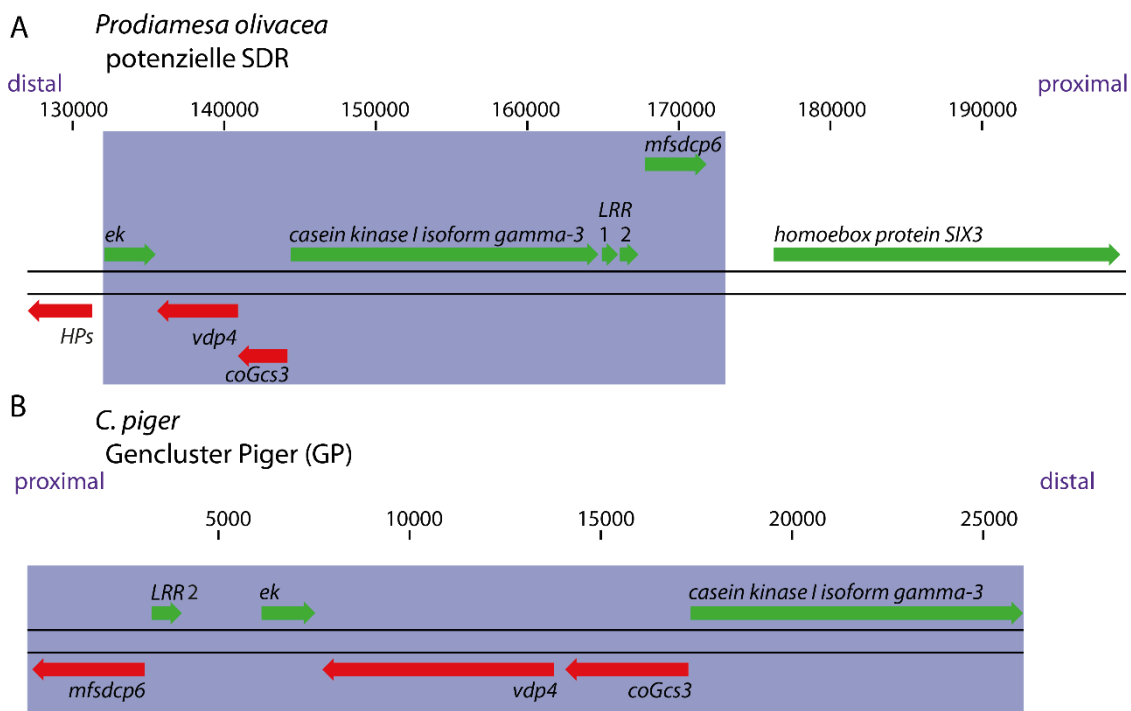


Abbildung 38 Der proximale Abschnitt der potenziellen *P. olivacea*-SDR im Vergleich zum Gencluster-piger Contig aus *C. piger*. Der oben abgebildete Ausschnitt zeigt den proximalen Abschnitt der potenziellen *P. olivacea*-SDR (A), welcher nicht mit der SDR aus *C. thummi/piger* übereinstimmt. Ein Contig aus einem *C. piger*-Assembly (contig_16311_asm_841_kdef25_b300) zeigte im BLASTX Alignment Übereinstimmungen zur proximalen *P. olivacea*-SDR. In B ist der 25 kb lange Contig mit den identifizierten Genen dargestellt (Gene auf dem Plusstrang in grün, Gene auf dem Minusstrang in rot). Die Gene sind im Vergleich zu *P. olivacea* in veränderter Abfolge angeordnet.

Dieser Abschnitt mit konservierten Genen in anderer Abfolge wird weiter als Gencluster-piger (GP) bezeichnet. Das Gen *LRR 1* aus *P. olivacea*, welches zwischen *ckIig-3* und *LLR-2* liegt, fehlt an dieser Stelle bei *C. thummi/piger*. Zudem ist das Gen *ckIig-3-like* nur teilweise durch den

Tabelle 15 Sequenzidentität der SDR-Gene und ihrer Produkte zwischen den zwei Arten *C. thummi* (♀) und *P. olivacea*

Gen	Sequenzidentität (Ebene)	Sequenzidentität (%)	Gen	Sequenzidentität (Ebene)	Sequenzidentität (%)
<i>mi-er-1</i>	genomisch	65,75	<i>dLacDH</i>	genomisch	60,96
	mRNA	66,57		mRNA	71,54
	ORF	70,32		ORF	72,34
	AS	57,81		AS	71,34
<i>luc7(p)</i>	genomisch	65,2	<i>polyhomeotic</i>	genomisch	68,86
	mRNA	63,31		mRNA	70,82
	ORF	69,82		ORF	71,74
	AS	62,78		AS	63,53
<i>wd74</i>	genomisch	68,4	<i>dentin-sialphospho-protein</i>	genomisch	77,6
	mRNA	69,95		mRNA	69,53
	ORF	67,56		ORF	72,26
	AS	70,4		AS	60,78
<i>fs(1)K10</i>	genomisch	62	<i>rpn5</i>	genomisch	75,64
	mRNA	61,74		mRNA	77,84
	ORF	69,31		ORF	78,64
	AS	44,83		AS	86,34
<i>bcn92</i>	genomisch	66,18	<i>HPs</i>	genomisch	67,24
	mRNA	64,48		mRNA	67,84
	ORF	70,5		ORF	71,77
	AS	64,37		AS	61,29

Überprüfung einer möglichen Männchen-spezifischen Duplikation des Gens *fs(1)K10-like* in *P. olivacea* mittels

a) genomischer Southern Analyse

In den *Chironomus*-Arten *thummi*, *piger* und *luridus* liegt im Männchen eine Männchen-spezifische Duplikation des in der SDR liegenden Gens *fs(1)K10-like* vor (Stiehl, unveröffentlicht). In *P. olivacea* wurde die potenzielle SDR inklusive des Gens *fs(1)K10-like* identifiziert. Der entsprechende Contig liefert jedoch keinen Hinweis auf eine Duplikation von *fs(1)K10-like*. Zur Überprüfung wurde ein Southern Blot bzw. Klonen aus Genbanken zuvor verifiziert wurde, wurde ein Southern Blot (Abbildung 40) mit genomischer, restringierter Weibchen- bzw. Männchen-DNA aus *P. olivacea*-Larven durchgeführt. Die gelelektrophoretische Auftrennung (Abbildung 40, B) der HincII-restringierten gDNA aus Weibchen bzw. Männchen zeigt einen Schmier, der im Bereich von 1 kb bis 100 bp in der Weibchen-Probe intensiver als in Männchen-Probe fluoresziert. Die mit SspI-verdaute DNA zeigt ebenfalls einen Schmier, zudem ist ein Rest hochmolekularer DNA zu sehen. Alle vier Spuren zeigen zwischen 1,2 und 1,5 kb eine prägnante Bande. Diese resultiert aus der Isolationsmethode der gDNA und es handelt sich dabei um RNA. Die Restriktion der gDNA mit HincII sollte entsprechend der Referenzsequenz und mit der in Abbildung 40 eingezeichneten Sonde zwei Signale (1,6 kb und 3 kb) im Autoradiogramm hervorrufen. Im Ergebnis sind die erwarteten zwei Signale sowohl in der Weibchen-Probe als auch in der Männchen-Probe zu erkennen (Abbildung 40, C hellgrüne Pfeile). Darüber hinaus zeigen sich ein prominentes Signal bei 1,2 kb und drei weniger prägnante bei ca. 5 kb, 800 bp und 500 bp (Abbildung 40 C, weiße Pfeile). Zur Aufklärung dieser Bandenvielfalt wurde ein weiterer Southern Blot durchgeführt. Der Southern Blot von dem HincII-restringierten Template der Sonde *fs(1)K10-like* (Abbildung 40, D und E) zeigt keine Variation der Bandengröße und-zahl. Die gelelektrophoretische Auftrennung des restringierten bzw. unrestringierten PCR-Produkts (Template der Sonde) von *fs(1)K10-like* (Abbildung 40, D) zeigt eine schwache Bande bei 2 kb (Abbildung 40, E grauer Pfeil) und zwei starke Banden bei 1,2 kb und 800 bp (Abbildung 40, E violette Pfeile).

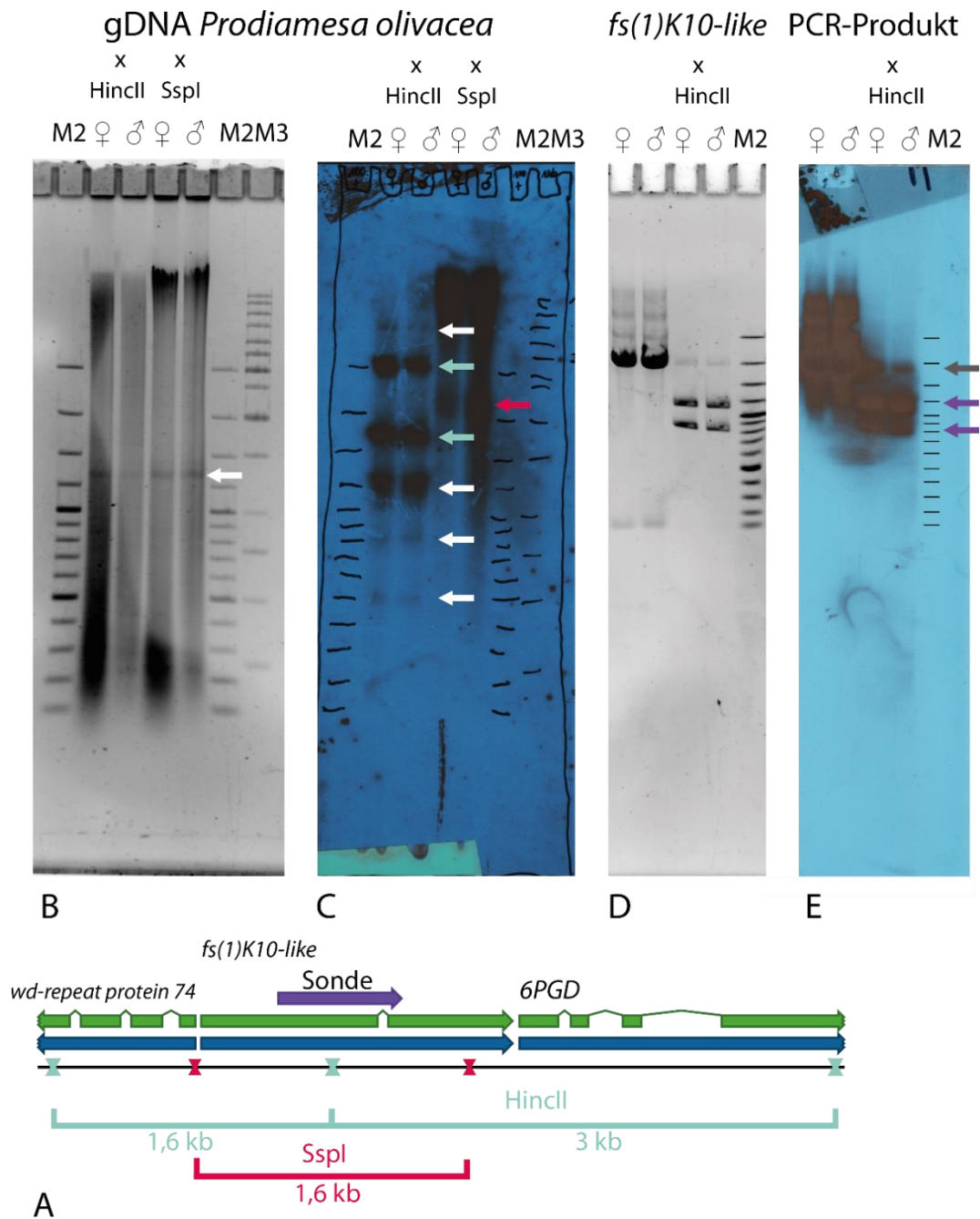


Abbildung 40 Southern Blot mit restringierter genomischer *P. olivacea* DNA aus weiblichen und männlichen Larven hybridisiert mit der Sonde *fs(1)K10-like*. A zeigt die genomische *P. olivacea*-Referenzsequenz mit den Genen *wd-repeat protein 74-like*, *fs(1)K10-like* und *6PGD* (blau = Gensequenz, grün = mRNA-Sequenz) und die Lage der Sonde *fs(1)K10-like* (violett). Die aus der Restriktion mit *HincII* (hellgrün) bzw. *SspI* (pink) resultierenden Fragmente sind mit Größen eingezeichnet. B zeigt die gelelektrophoretische Auftrennung der restringierten (*HincII*, *SspI*) *P. olivacea*-DNA aus Weibchen bzw. Männchen. In allen vier Spuren ist eine prominente Bande bei 1,3 kb zu sehen (weißer Pfeil), dies ist RNA, welche aus der Isolationsmethode resultiert. Als Molekulargewichtsstandard wurden der GeneRuler 100 bp plus DNA Ladder (M2) und der GeneRuler™ 1 kb DNA Ladder (M3) aufgetragen. C zeigt das Autoradiogramm nach der Hybridisierung mit der Sonde *fs(1)K10-like*. Die weißen Pfeile markieren die Signale. D zeigt die gelelektrophoretische Auftrennung des PCR-Produkts von *fs(1)K10-like* (unrestringiert bzw. restringiert). Als Molekulargewichtsstandard diente M2. E zeigt das aus der Hybridisierung von D resultierende Autoradiogramm.

Diese beiden Banden (Abbildung 40 E, violette Pfeile) entsprechen den Erwartungen, da innerhalb des Gens eine Schnittstelle für HincII liegt. Die dritte Bande (Abbildung 40 E, grauer Pfeil) entspricht dem unrestringiertem PCR-Produkt. Zur Kontrolle wurde das unrestringierte PCR-Produkt von *fs(1)K10-like* in den Spuren 1 (♀) und 2 (♂) aufgetragen (Abbildung 40, D).

Die Restriktion mit SspI zeigte ebenfalls in beiden Geschlechtern das gleiche Signalmuster, der vermutlich unvollständige Verdau der gDNA führt zur Überlagerung des eigentlichen Signals bei 1,6 kb, welches aber (vor allem in der Weibchen-Probe) gut zu erkennen ist (Abbildung 40, C pinker Pfeil).

b) NGS-basierter Analyse

Zur Überprüfung einer möglichen Duplikation von *fs(1)K10-like* im Genom von *P. olivacea* (- Männchen) wurde per CLC Genomics Workbench ein Mapping mit genomischen Daten 729 (Tabelle 5) durchgeführt. Als *single-copy*-Referenz diente (wie unter 3.3.4) das Gen *bcn92*. Das Mapping (LF: 0,96 SIM: 0,98) ergab für das 1837 bp lange Gen *fs(1)K10-like* eine durchschnittliche Abdeckung von 66. Dies entspricht einer Kopienzahl von 1 im Genom (Tabelle 16).

Tabelle 16 Überprüfen der Kopienzahl von *fs(1)K10-like* mittels Abdeckung genomischen DNA-reads im Mapping

Datensatz	Name	Sequenz-Typ	Referenz-länge	Total read count	Ø Abdeckung	reads total	Kopie-Zahl
male Gehirn	<i>bcn92</i>	<i>single copy</i>	612 bp	350	61	99.402.046	1
male Gehirn	<i>fs(1)K10-like</i>	<i>single copy</i>	1837 bp	1136	66	99.402.046	1

3.7 Analyse des larvalen *P. olivacea*-Transkriptoms

Das zuvor generierte und annotierte Master-Assembly (Tabelle 6, Benjamin Rieger und Steffen Rapp, IMSB, Mainz) aus *P. olivacea* diente als Grundlage der RNA-Seq-Analyse. Die Auswertung beschränkte sich auf Gene bzw. Transkripte, die einen *fold change* (*fc*) > 4 und eine *range* von >100 aufwiesen.

Tabelle 17 zeigt eine Auswahl der differenziell exprimierten Gene des *P. olivacea*-Larven-Transkriptoms. Die Auswahl beschränkt sich auf im Männchen höher exprimierte Gene. Ein Vergleich mit Literaturdatenbanken gibt einen Hinweis über die mögliche Funktion der Gene. 91 der 784 im Männchen hoch exprimierten Gene zeigen keine Expression im Weibchen. Dagegen zeigen nur 83 Gene im Weibchen eine erhöhte Expression.

Tabelle 17 Übersicht der Männchen-spezifisch exprimierten Gene des Larven-Transkriptoms aus *P. olivacea*.
n.v.: nicht verfügbar

Position	Gen	Funktion/Prozess	Expression		fold change
			♀	♂	
10	<i>sperm_flagellar_protein_1-like</i>	Spermatogenese	0	36	n.v.
11	<i>msRNA 98 Cb</i>	Spermatogenese	0	371	n.v.
17	<i>osiris</i>	Pathogenabwehr	0	11	n.v..
48	<i>radial_spoke_head_10_homolog_B-like</i>	Spermatogenese	0	4	n.v.
96	<i>testis-specific_serine/threonine-protein_kinase_3-like</i>	Spermatogenese	0	118	n.v.
97	<i>male specific RNA 98 Cb-2</i>	Spermatogenese	0	117	n.v..
186	<i>exuperantia-1</i>	Spermatogenese	1,96	557	0,003
212	<i>axonemal_dynein_intermediate_chain</i>	Spermatogenese	0,36	83	0,004

Die Auswertung der Analyse zeigte u.a. einen signifikanten Expressionsunterschied für ein Gen, *male specific RNA 98Cb*, welches in weiblichen *P. olivacea*-Larven (*e-value*: 0) nicht, in männlichen *P. olivacea*-Larven dagegen stark (*e-value*: 204,5) exprimiert wird.

Eine RT-PCR mit sequenzspezifischem Primer bestätigte dieses Ergebnis. Das zu erwartende RT-PCR-Produkt sollte eine Länge von 411 bp haben (Abbildung 41, A; B).

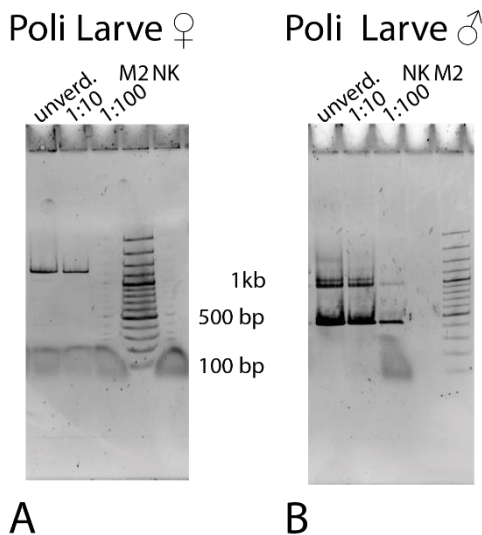


Abbildung 41 Gelelektrophoretische Auftrennung der RT-PCR-Produkte für das Primerpaar *msRNA98Cb* auf weibliche (A) bzw. männliche (B) cDNA aus *P. olivacea*-Larven. A zeigt für die weiblichen Proben mit unverdünnter cDNA und 1:10 verdünnter cDNA als Template eine starke Bande bei 1,3 kb. In der Probe mit der 1:100 verdünnten cDNA als Template ist kein Produkt zu sehen. B zeigt in allen drei Proben (cDNA unverdünnt; 1:10; 1:100) zwei Banden, eine bei 1,3 kb und eine bei ~400 bp. Die 400 bp-Bande ist in jeder Probe die prägnanteste. In den Proben mit unverdünnter oder 1:10 verdünnter cDNA sind noch weitere leichte Banden bzw. ein Schmier zusehen. Dabei handelt es sich vermutlich um PCR-Artefakte bzw. Einzelstränge. Diese treten in der dritten Probe (1:100 verdünnt) nicht auf, und die Sequenzierung der dritten Probe (Template cDNA 1:100) liefert ein eindeutiges und signifikantes Ergebnis.

Die gelelektrophoretische Auftrennung des RT-PCR-Produkts, bei dem weibliche cDNA als Template-DNA eingesetzt wurde, zeigt eine Bande bei 1,3 kb (Abbildung 41, A). Diese entspricht der genomischen Organisation des Gens. Die Auftrennung des auf männlicher cDNA basierenden RT-PCR-Produkts zeigt zwei Banden mit den Größen 1,3 kb und ~400 bp. Das entspricht dem zu erwartenden RT-PCR-Produkt (411 bp) aus der mRNA und dem genomischen Abschnitt von 1,3 kb. Dies wurde mittels Sanger-Sequenzierung der jeweiligen Amplifikate verifiziert.

In *C. thummi/piger* wurde mittels BLASTX ebenfalls ein orthologes Gen identifiziert. Dieses zeigt in einer Expressionsanalyse ebenfalls eine Expression für männliche *C. thummi*- bzw. *piger*-Larven. Darüber hinaus wurde in *P. olivacea* zusätzlich ein paraloges Gen (*male specific RNA 98Cb-2*) zu *male specific RNA98Cb-1* identifiziert. Das in *C. thummi/piger* identifizierte *male specific RNA 98Cb* zeigt eine höhere Sequenzidentität zu *PolimsRNA98Cb-2*, daher wird es entsprechend benannt.

Per FISH wurde das Gen in *C. thummi/piger* und die beiden paraloge Gene in *P. olivacea* chromosomal lokalisiert (Abbildung 43: Schematische Abbildung, Anhang: Abbildung 57, Abbildung 58 und Abbildung 59).

3.8 Chromosomenkartierung bei *C. thummi/piger* und *P. olivacea*

Die Lokalisation von 20 Genen bzw. Contigs (Abbildung 14, Abbildung 22 und (Hankeln et al. 1993), Abbildung 25, Abbildung 16, Abbildung 27, Abbildung 29, Abbildung 31, Abbildung 34, Abbildung 36) im *P. olivacea*-Genom ist in Abbildung 42 schematisch (nicht maßstabsgetreu) dargestellt.

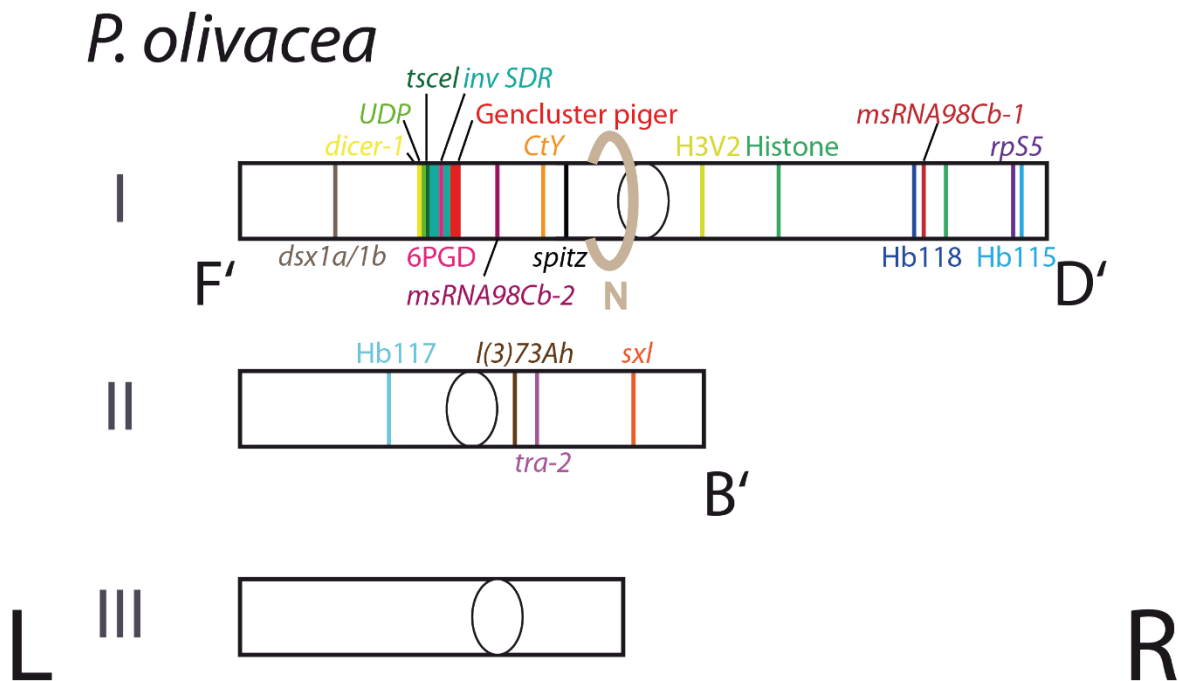


Abbildung 42 **Chromosomenkartierung bei *P. olivacea***. Abgebildet eine Chromosomenkarte der Lage von 20 Genen oder Contigs in *P. olivacea* Chromosomen durchnummeriert I-IV und die Chromosomenarme mit B', D', F².

Die in dieser Arbeit in *C. thummi/piger* sechs neu lokalisierten Gene bzw. Contigs wie *tscel* (Abbildung 37), *dicer-1* (Abbildung 37), *UDP*, *msRNA98Cb-2*, *6PGD* (Abbildung 37) oder das GP (Abbildung 37) in Abbildung 43 eingetragen. Die bekannten Loci von *C. thummi* 12 Genen bzw. Contigs wie *Sxl* (Frank 2001), *l(3)73Ah* (Frank 2001), *tra-2* (Gerbracht 2014), *dsx1a* (Frank 2001), *dsx1b* (Hammann 2013) bzw. Histon- und Hämoglobin-Cluster B und D (Schmidt et al. 1988), (Hankeln und Schmidt 1991) und die SDR (Kraemer und Schmidt 1993) wurden ebenfalls in das Schema aus Abbildung 43 eingetragen. Die Loci zwischen *C. thummi* und *C. piger* sind weitgehend an homologen Positionen, so dass *C. thummi* auch für *C. piger*. Die Lokalisation erfolgte per FISH an polytären Speicheldrüsenchromosomen der Arten

² definiert für *P. olivacea* in dieser Arbeit

P. olivacea, *C. thummi* und *C. piger*. Die Bestimmung dieser Loci erlauben Rückschlüsse auf die Chromosomenumbauten zwischen den Arten *C. thummi/piger* und *P. olivacea*.

C. thummi

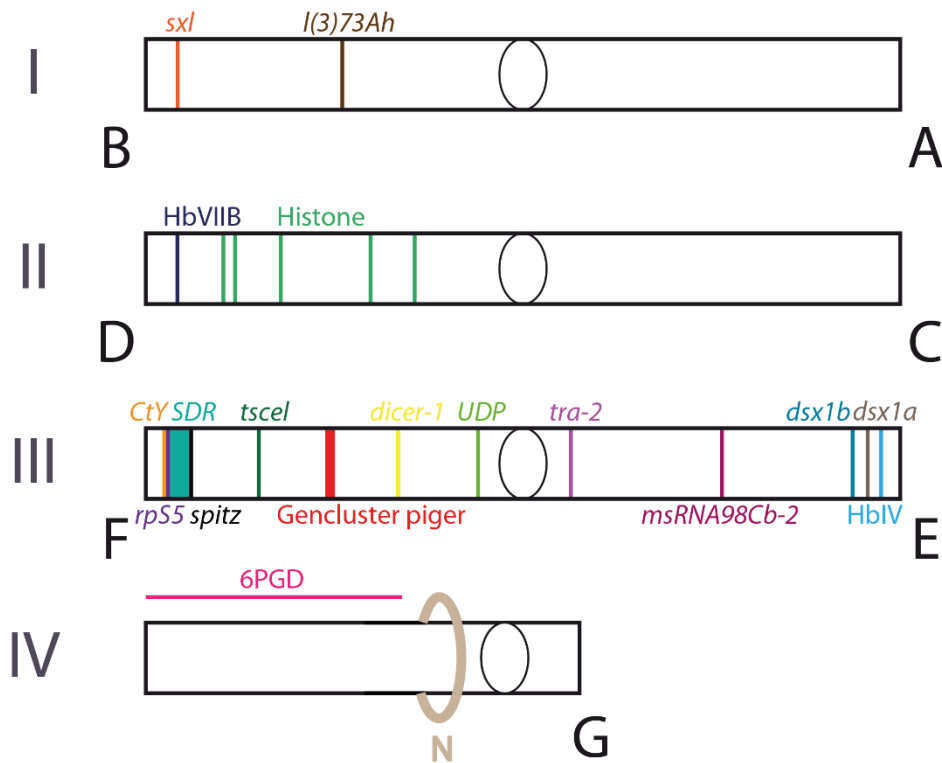


Abbildung 43 **Chromosomenkartierung bei *C. thummi/piger***. Abgebildet eine Chromosomenkarte der Lage von 18 Genen aus *C. thummi/piger* Chromosomen durchnummeriert I-IV und die Chromosomenarme mit A bis G³.

In *C. thummi/piger* liegen die Gene *Sxl* und *l(3)73Ah* auf dem Chromosom I Arm B, in *P. olivacea* liegen die Gene ebenfalls auf einem Chromosomenarm (Chromosom II Arm R). Auch die Orientierung der Gene *Sxl* (distal) und *l(3)73Ah* (proximal) ist in beiden Arten gleich. Daher wird für den Chromosomenarm II R die Bezeichnung B' eingeführt. es gibt jedoch einen Unterschied zwischen den beiden Arten: In *P. olivacea* liegt auf dem Chromosomen II Arm B' noch ein weiteres der identifizierten Gene, *tra-2*. *Tra-2* wurde in *C. thummi* auf Chromosom III Arm E nahe dem Zentromer lokalisiert. In *C. thummi* wurden auf Chromosom III Arm E noch weitere Gene identifiziert (*msRNA98Cb-2*, Hb-Cluster E und *dsx*). Die Gene *dsx* und *msRNA98Cb-2* wurden in *P. olivacea* auf Chromosom I Arm L lokalisiert. Auf diesem Arm wurden zudem *dicer-1*, *UDP*, *tscel*, potenzielle SDR, *6PGD*, Gen-Cluster, *CtY* und *spitz-like*

³ definiert für *C. thummi/piger* von Keyl 1962

identifiziert. Da die SDR in den meisten untersuchten *Chironomus*-Arten auf Arm F lokalisiert ist und auch die anderen oben genannten Gene für *Chironomus* auf Arm F liegen, wird der Chromosomenarm I L in *P. olivacea* als F' bezeichnet. Das Gen *6PGD*, welches in *P. olivacea* in der Region der potenziellen SDR vorkommt, liegt in *C. thummi/piger* auf Chromosom IV Arm G. Die exakte Position war nicht auszumachen, da die Chromosomenstruktur durch den Nukleolus stark beeinträchtigt wird. Das Gen *rpS5* aus der SDR von *C. thummi/piger* wurde zwar auf demselben Chromosom wie die potenzielle SDR, jedoch auf dem anderen Arm (Chr I Arm R) in *P. olivacea* lokalisiert. Auf dem Chromosomenarm I Arm D' wurden zudem die Histon-Gen-Cluster, *orphon-H3V2*, *msRNA98Cb-1* und die Poli Gb Gen-Cluster 118 und 115 lokalisiert. In *C. thummi* sind die Histon-Gen-Cluster auf Chromosom II Arm D lokalisiert. Zudem ist auf diesem Chromosomenarm ein weiteres Hb-Cluster zu finden. Daher wird der Arm in *P. olivacea* als D' bezeichnet. Auf dem Chromosom II Arm L konnte nur Poli Gb117 lokalisiert werden. Da das Gen keine signifikante Ähnlichkeit zu aus *C. thummi* oder *C. piger* bekannten Hämoglobin-Genen aufweist, ist keine Aussage über die Zuordnung zu den *C. thummi/piger*-Chromosomenarmen nach (Keyl und Keyl 1959) bzw. (Keyl 1962) möglich. Auf Chromosom III konnte in *P. olivacea* keine Hybridisierung bzw. kein Fluoreszenz-Signal außer dem Signal der Sat-DNA am Zentromer beobachtet werden.

Tabelle 18 Lokalisation von DNA-Sequenzen (single copy-Gene oder repetitive Elemente) per FISH in *C. thummi/piger* und *P. olivacea*.

Spezies	Gen	Chromosom	Arm	Locus	Autor	Spezies	Gen	Chromosom	Arm	Locus	Autor												
<i>P. olivacea</i>			a	b		<i>C. thummi/piger</i>			a'	b	c												
	<i>Sxl</i>	2	R	B'			<i>Sxl</i>	I	L	B	A1c/d	(Frank 2001)											
	<i>tra-2</i>	2	R	B'	diese Arbeit		<i>tra-2</i>	III	R	E		(Gerbracht 2014)											
	<i>l(3)73Ah</i>	2	R	B'	diese Arbeit		<i>l(3)73Ah</i>	I	L	B	B3d-g	(Frank 2001)											
	<i>dsx1 und 2</i>	1	L	F'	diese Arbeit		<i>dsx1a</i>	III	R	E	A1b/c	(Frank 2001)											
	<i>msRN98Cb-1</i>	1	R	D'	diese Arbeit		<i>dsx 1b</i>	III	R	E		(Hammann 2013)											
	<i>msRN98Cb-2</i>	1	L	F'	diese Arbeit		<i>msRN98Cb-2</i>	III	R	E		diese Arbeit											
	Globin 115	1	R	D'	diese Arbeit		HB VII	II	L	D	F2b3	(Schmidt et al. 1988)											
	Globin 118	1	R	D'	diese Arbeit		HB IV	III	R	E	A1b2	(Schmidt et al. 1988)											
	Globin 117	2	L		diese Arbeit																		
	Histon-Gen-Cluster	1	R	D'	(Hankeln et al. 1993)		Histon-Gen-Cluster	II	L	D	5 Loci: C4f, C6e, D3c, E1k, E2c	(Hankeln und Schmidt 1991)											
	H3V2	1	R	D'	diese Arbeit																		
	Sat-DNA	1, 2, 3			Zentromer	diese Arbeit	Sat-DNA	1, 2, 3, 4		*	Zentromer	(Schmidt 1984)											
	<i>CtY</i>	1	L	F'	diese Arbeit		<i>CtY</i> bzw. <i>CpY</i>	III	L	F	D3ef	(Kraemer, Schmidt 1993; Kraemer 1994)											
	potenzielle SDR	1	L	F'	diese Arbeit		SDR	III	L	F	D3e	(Kraemer, Schmidt 1993)											
	<i>rpS5-like</i>	1	R	F'	diese Arbeit		<i>rpS5-like</i>	III	L	F	D3e	(Kraemer, Schmidt 1993)											
	<i>spitz-like</i>	1	L	F'	diese Arbeit		<i>spitz-like</i>	III	L	F	D3e	(Kraemer, Schmidt 1993)											
	Gencluster-piger	1	L	F'	diese Arbeit		Gencluster-piger	III	L	F	C3e-f	diese Arbeit											
	<i>6PGD</i>	1	L	F'	diese Arbeit		<i>6PGD</i>	IV		G		diese Arbeit											
	<i>dicer-1</i>	1	L	F'	diese Arbeit		<i>dicer-1</i>	III	L	F	C3c	diese Arbeit											
	<i>UDP</i>	1	L	F'	diese Arbeit		<i>UDP</i>	III		F	B3h-j	diese Arbeit											
	<i>tscel</i>	1	L	F'	diese Arbeit		<i>tscel</i>	III		F	D2l-m	diese Arbeit											
			a: nach (Zacharias 1979)						a': nach (Keyl und Keyl 1959)			b: nach (Keyl 1962)						* in <i>C. thummi</i> auch perizentrisch			c: nach (Hägele 1970)		

4 Diskussion

4.1 Das *P. olivacea*-Genom

Eine erste Analyse der *P. olivacea*-Chromosomen wurde 1979 von Zacharias durchgeführt (Zacharias 1979). Dabei gelang es anhand von Größe und Bandenstruktur, die drei Chromosomen bzw. die sechs Chromosomenarme zu benennen. Die durchgeführten Analysen beschränkten sich dabei vor allem auf zytologische Merkmale und DNA-Gehalt verschiedener Zelltypen. Eine geschätzte Genomgröße von 0,25 pg DNA für 2 C und ein geschätzter GC-Gehalt von 33 % für polytänes Gewebe sowie 35 % für diploides Gewebe war zum Beginn der vorliegenden Arbeit bekannt (Zacharias 1981). Eine NGS-Analyse genomischer DNA von *P. olivacea* (Tabelle 5) bestätigen einen GC-Gehalt zwischen 32 und 34 %. Dieser Wert mit dem GC-Gehalt stimmt gut mit dem der Gattung *Chironomus* überein und weicht erheblich von dem GC-Gehalt der Genome anderer Dipteren z. B. *D. melanogaster* (42 %) ab (Steinemann 1978). Warum Chironomiden einen vergleichsweise hohen AT-Gehalt haben ist unklar.

4.1.1 Sat-DNA des *P. olivacea*-Genoms

Über den Anteil an hochrepetitiven DNA-Sequenzen des *P. olivacea*-Genoms konnte bisher keine Aussage getroffen werden. Eine CsCl-Dichtezentrifugationsanalyse von *P. olivacea* konnte keine typische Sat-DNA nachweisen (Zacharias 1981). Dies trifft ebenfalls für die Art *C. thummi* zu (Steinemann 1978). Bei *C. melanotus* und *Glyptotendipes barbipes* konnten typische Sat-DNAs über CsCl-Dichtezentrifugationsanalysen nachgewiesen (Steinemann 1978), (Schmidt 1980). In *P. olivacea* besitzt die DNA-Sequenz des Zentromer-Satelliten einen GC-Gehalt der dem der genomische Fraktion entspricht (Tabelle 5, 3.3.4). Daher ist es nicht möglich, den Zentromer-Satelliten unter diesen Bedingungen mittels CsCl-Dichtezentrifugation abzutrennen. Wenn davon ausgegangen wird, dass es sich mit einem 113 Mb großen Genom um ein eher kleines Insekten-Genom handelt, dann sollte der Anteil an repetitiven Elementen im Vergleich eher gering ausfallen. Die in dieser Arbeit identifizierte Sat-DNA beschränkt sich nach den Ergebnissen der FISH (3.3.3, Abbildung 14) auf die Zentromere der drei Chromosomen. Dies entspricht dem Zustand von *C. piger*. Dort befinden sich die Cla-Elemente ebenfalls nur im Zentromer und nicht wie in *C. thummi* auch in der

Peripherie. Wie von Schaefer und Schmidt über C₀t-Analysen bestimmt wurde, enthält das Genom der Schwesternart *piger* nur 3 % hochrepetitive DNA, wogegen *thummi* 12 % hochrepetitive DNA enthält (Schaefer und Schmidt 1981). Weiter wurde eine Kopien-Zahl von ~71.000 für *thummi* und ~13.000 für *piger* geschätzt (Schaefer und Schmidt 1981). Für *P. olivacea* konnte eine Kopien-Zahl von ~14.000 Sat-Repeats geschätzt werden (Tabelle 7). Dies Kopien-Zahl von Cla-Elementen bei *C. piger*, die bei dieser Spezies typischerweise ebenfalls überwiegend in den Zentromerregionen lokalisiert sind.

Weitere Sequenzen von hochrepetitiver DNA wurden nicht identifiziert (vgl. 3.3.4), die zuvor identifizierten Sequenzen weisen einen bakteriellen Ursprung auf. Dies lässt sich mit einer Verunreinigung durch Bakterien-DNA bei der Sequenzierung bzw. beim Erstellen der sequenzierten DNA-Bank erklären. Somit lässt sich keine konkrete Aussage über den Umfang der hochrepetitiven Fraktion treffen. Wenn jedoch davon ausgegangen wird, dass das *C. piger*-Genom 30 % kleiner ist als das Genom von *C. thummi* (Keyl 1965) und sich für *C. piger* eine Genomgröße von 140 Mb schätzen lässt. Ist eine Schätzung des Anteils von 3 % hochrepetitiver DNA-Sequenzen nur sehr grob möglich. Da das Master-Assembly im günstigsten Fall ~84 % des *P. olivacea*-Genoms abdeckt (Tabelle 6), würde das bedeuten, dass noch 13 % des *P. olivacea*-Genoms bisher nicht assembliert werden konnten. Dabei kann es sich natürlich auch um mittelrepetitive Sequenzen handeln, deren Anteil ähnlich dem der hochrepetitiven DNA schwer zu messen ist.

4.1.2 Genfamilien aus *P. olivacea*

Die Analyse der Genfamilien mit multiplen Kopien pro *P. olivacea*-Genom beschränkte sich vor allem auf die Globin- und Histon-Gene.

Die Hämoglobin-Gene

Aus der Hämoglobin-Familie konnten zehn Hämoglobin-Gene an drei verschiedenen Chromosomenpositionen in *P. olivacea* identifiziert werden (3.4). In *C. thummi* sind dagegen zwei Cluster mit über 30 Genen bekannt (Friedl 2001), (Schmidt et al. 1988), (Hankeln et al. 1991), (Rozynek et al. 1991). Die Existenz von Hämoglobin-Genen in *P. olivacea* ist insofern bemerkenswert als die Larven der Gattung *Prodiamesa* nicht die für Hämoglobin-Anwesenheit typische Rotfärbung zeigen. Die Farbe der *P. olivacea*-Larven ist gelblich-weiß. Die Hämoglobin-Gene von *P. olivacea* enthalten keine typischen Sequenzen die für Signalpeptide kodieren. In Abbildung 45 ist ein multiples AS-Sequenzalignment (MSA) der Polypeptide CthuHbI, DmelGb1 und der *P. olivacea* Globine 115 A bis F des Globin-Clusters von Contig_115 zu sehen. CthuHbI besitzt als einziges der zu vergleichenden Proteine ein Signalpeptid (Abbildung 45, rosa unterlegt), das die Sekretion in die Hämolymphe ermöglicht. Bei den in *P. olivacea* identifizierten Hämoglobinen handelt es sich eventuell um intrazelluläre Globine.

Alle in *P. olivacea* identifizierten Gene enthalten Intronsequenzen (Abbildung 17, Abbildung 19, Abbildung 19, Abbildung 44). Die Hämoglobin-Gene der Chironomiden galten, im Gegensatz den denen der Vertebraten oder Pflanzen, lange als intronlose Vertreter. Die anfängliche These, dass es zum Verlust der Introns in den Globin-

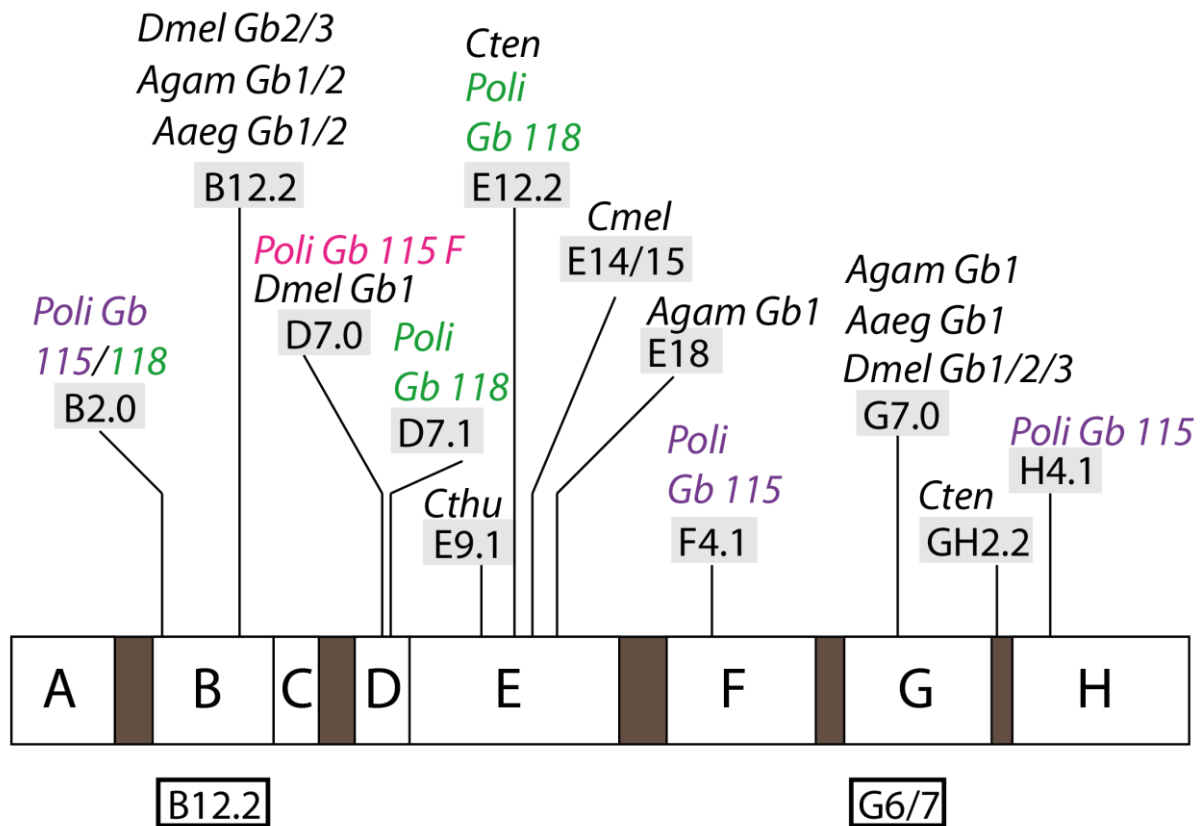


Abbildung 44 **Intronpositionen der Hämoglobin-Gene von Dipteren-Arten.** Die Intronpositionen der Dipteren-Spezies *P. olivacea* (Poli, lila, grün und pink), *D. melanogaster* (Dmel) (Burmester et al. 2006), *Anopheles gambiae* (Agam), *Aedes aegypti* (Aaeg) (Burmester et al. 2007) und *C. thummi*(Cthu), *C. tentans* (Cten), *C. melanotus* (Cmel) sind dargestellt. Die Positionen sind bezogen auf die Helices (A-H) der Globin-Proteine. Intron B12.2 beispielsweise ist in Helix B, Kodon 12 zwischen Nukleotid 2 und 3 lokalisiert. Die Positionen in Helix B und G (dicke schwarz umrandete Kästen) sind konserviert in Vertebraten, Invertebraten und Pflanzen. Abb. verändert nach (Friedl 2001)

Genen von *Chironomus* kam, wurde von Friedl eher ausgeschlossen (Friedl 2001). Die Klassifizierung der Intronsequenzen erfolgt über ihre Lage in der entsprechenden Helix des Proteins. Die Intron Grenzen der Poli Gb 115 A bis E sind an den Positionen B2.0, F4.1 und H4.1 zu finden (Abbildung 44, lila) für Poli Gb118 an den Positionen B2.0, D7.1 und E12.2 (Abbildung 44, grün). In Poli Gb117 liegt das Intron außerhalb der kodierenden Region und kann somit keiner Helix zugeordnet werden. Diese Positionen wurden für keine anderen Globin-Gene zuvor beschrieben. Ein Intron in Helix H wurde ebenfalls noch nie in einer anderen Art beschrieben. Innerhalb der Art *P. olivacea* sind die oben genannten Intronpositionen nur teilweise konserviert (markiert für Poli Gb115 in MSA in Abbildung 45, violetter Strich). Dies gilt auch in Bezug auf die (Hämo-)Globin-Gene aus der Arten *D. melanogaster* und *C. thummi* Eine Ausnahme bildet dabei Poli Gb115F (Abbildung 44, pink), welches die zweite Intronposition (D7.0) mit *D. melanogaster* Glob1 teilt. Zudem teilt das Poli Gb118 die Position E12.2 mit Genen aus *C. tentans*. Die Annahme von Friedl, dass sich bei

weiteren untersuchten Arten noch eine Vielzahl an Intronpositionen finden lassen würde, wird durch diesen Befund noch bekräftigt (Friedl 2001). Ob die möglicherweise intrazellulären *P. olivacea*-Globine auch eine sauerstoffbindende Funktion besitzen, muss noch geklärt werden. Über die bekannte Röntgenstruktur der monomeren Hämoglobin-Komponente III aus *C. thummi* (Steigemann und Weber 1979) kann die AS-Sequenz anderer Hämoglobine bewertet werden. Dabei spielen Reste mit Hämkontakt, Sauerstoffbindung, Dimerbindestellen oder AS-Reste mit interhelikalen Kontakten eine Rolle. Verschiedene Autoren haben sich mit diesen sogenannten invarianten bzw. semi-invarianten AS-Positionen beschäftigt (Goodman et al. 1983), (Bashford et al. 1987), (Moens et al. 1996), (Ptitsyn und Ting 1999).

Der Vergleich der *P. olivacea*-Globine mit *C. thummi*-Hämoglobin I-Protein und *D. melanogaster*-Globin 1-Protein liefert Hinweise für eine mögliche Funktion der zuvor identifizierten *P. olivacea*-Globine. Die Hämoglobin-spezifischen Helices A bis H wurden über die Referenz-Proteine CttHbI und DmelGb1 für *P. olivacea*-Gb 115 A bis F identifiziert (Abbildung 45, grau unterlegt). Die für die Häm-Bindung beschriebenen neun wichtigsten AS-Reste sind, wie im MSA zu sehen ist, konserviert und entsprechen somit einer der „erlaubten“ AS für die entsprechende Position (Abbildung 45, türkis unterlegt). Weitere AS-Reste, welche für Häm- oder interhelikale Kontakte verantwortlich sind, sind nicht explizit markiert, sondern einer Tabelle aus dem elektronischen Anhang zu entnehmen. Diese sind größtenteils ebenfalls konserviert. Die zwei His-Reste E7 und F8 sind in Abbildung 45 violett unterlegt. Alle identifizierten *P. olivacea* Gb 115 Proteine tragen an der Position F8 den konservierten His-Rest. An der Position E7 zeigen die möglichen Proteine Poli Gb115 D und E einen anderen AS-Rest. Das Poli Gb117 zeigt dagegen an Position F8 keinen His-Rest. In dem Poli Gb118 tragen beide Positionen die konservierten His-Reste. Somit kommt eine Hämbindung durch die *P. olivacea* Globine 115 A, B, C, F und Poli Gb118 in Frage. Über die Funktion der Poli Gb115 D und E kann keine ausreichende Aussage getroffen werden. Der AS-Austausch an Position E7 könnte zu einer funktionslosen Genkopie geführt haben. Das Gleiche gilt für die Genkopie von Poli Gb 117.

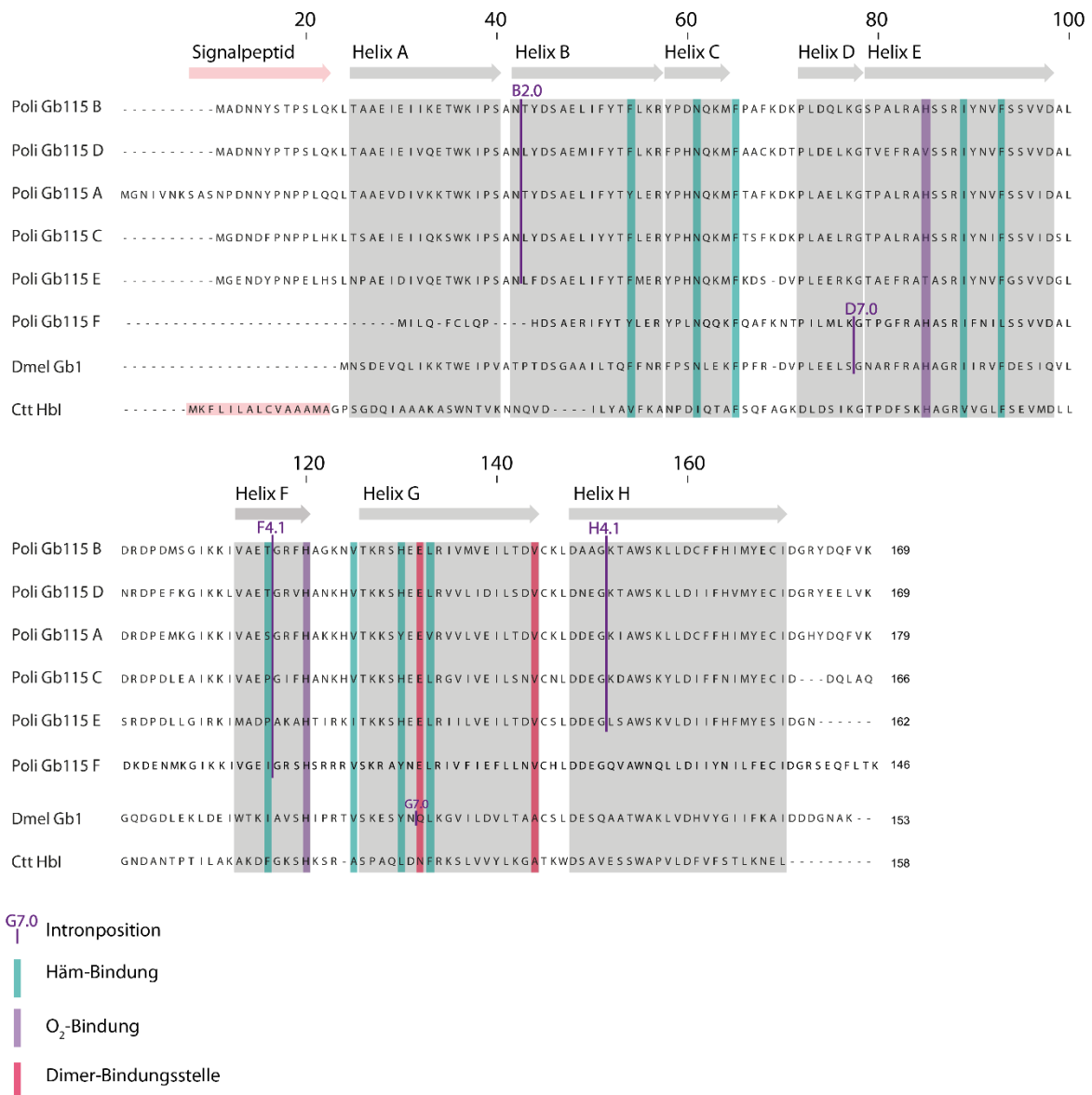


Abbildung 45 MSA der abgeleiteten Proteine des Globin-Gen-Clusters 115 aus *P. olivacea* mit den *D. melanogaster* Globin-1 und *C. thummi* HbI. Die Helices A bis H sind im MSA grau unterlegt. Das Signalpeptid von CthuHbI ist rosa unterlegt. Weiter sind funktionell bzw. strukturell wichtige AS-Reste hervorgehoben. Wie zum Beispiel die Positionen (bzw. AS-Reste), welche bei der Ausbildung von Häm-Kontakten beteiligt sind (türkis unterlegt) oder AS-Reste, welche für die Sauerstoffbindung verantwortlich sind (lila unterlegt aber auch Positionen, welche an der Dimerisierung beteiligt sind (pinkrot unterlegt). Intronengrenzen der dazugehörigen Gensequenzen sind im AS-Sequenzalignment violett unter Positionsangabe in der Helix markiert.

Die Dimer-Bindestellen (Abbildung 45, rot unterlegt) sind innerhalb der *P. olivacea* Globine des Clusters 115 konserviert (G7 = E und G19 = V). An Position G7 ist in allen Poli Gb115 Proteinen eine Glutaminsäure in der Helix G vorhanden. An der Position G19 weisen die Poli Gb115 weder die AS Lysin, His noch Asparagin auf. Die MSA für die darüber hinaus identifizierten *P. olivacea* Globine (Globine 117 und 118) sind aufgrund ihrer Redundanz dem Anhang zu entnehmen (Abbildung 60 und Abbildung 61). Es liegt daher die Vermutung nahe,

dass es sich bei den in *P. olivacea* identifizierten Globinen um Monomere handelt (vgl. (Friedl 2001)).

Eine *RaptorX structure prediction* für die *P. olivacea*-Globine sagt eine ähnliche Struktur wie die von DmelGb1 vorher (Abbildung 46, B).

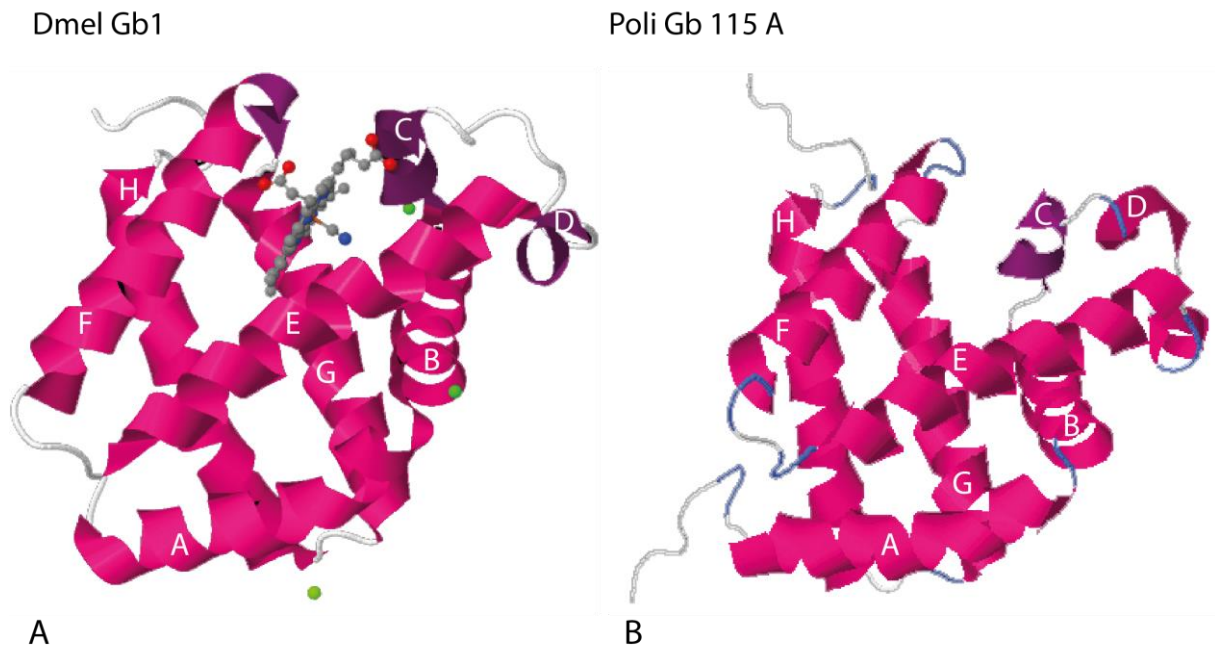


Abbildung 46 Vorhersage der 3D-Struktur von *P. olivacea* Gb 115 A mittels *RaptorX structure prediction-Tool*. A zeigt die Referenzstruktur von DmelGb1 mit den Helices A bis H und einem gebundenen Häm. B zeigt die für Poli Gb115 A vorhergesagte 3D-Struktur, die Helices sind nach der Referenz DmelGb1 bezeichnet.

Dabei verlaufen die Helices A, E und F im Vordergrund und die Helices G und H im Hintergrund antiparallel zu einander. Die Helices C und D weisen in Poli Gb115A einen geringeren Abstand und somit auch einen anderen Winkel zu einander auf. Die vorhergesagte Helix D neigt sich dadurch ins Innere des Peptids. Auffällig ist jedoch der Raum, zwischen den Helices, welcher es ermöglicht eine Häm-Gruppe zu binden.

Zusammengefasst zeigen die Globin-Gene von *P. olivacea* bzw. deren abgeleitete Proteine Übereinstimmungen in Genstruktur und konservierten AS-Resten. Allerdings finden sich auch viele Unterschiede: fünf der zehn Gene zeigen die konservierten AS-Reste denen mögliche Funktion bei der Sauerstoffbindung zugeschrieben werden. Schwierig ist eine Aussage darüber wie die Verwandtschaftsbeziehungen zu den bekannten Chironomiden-Hämoglobin-Genen sind. Wahrscheinlich ist eine phylogenetische Beziehung zu Globin 1 aus *D. melanogaster* insbesondere wegen der großen Sequenzübereinstimmung.

Die Histon-Gene

Bei der Familie der Histon-Gene handelt es sich um hoch konservierte evolutionär sehr alte Gene, welche in allen Eukaryoten vorkommen (Isenberg 1979). Die Organisation der Histon-Gene bei *P. olivacea* entspricht der bei anderen Dipteren-Arten in Clustern. Diese Genfamilie ist wie erwartet auch in *P. olivacea* hochkonserviert (Tabelle 9, Abbildung 23). Die Sequenzidentität zwischen *P. olivacea* und *C. thummi* nimmt von 100 % für H3/H4 über ~96 % für H2A/H2B bis zu 64 % für H1 ab (Tabelle 9). Dieses Verhältnis stimmt mit den bekannten Histon-Sequenzen aus anderen Arten überein (Isenberg 1979).

In Abbildung 47 ist das MSA von *C. thummi*-H1-Protein und *P. olivacea*-H1-Protein dargestellt. Die AS-Austausche nehmen von 27 AS-Austauschen für Austausch mit AS welche schwächer ähnlichen Eigenschaften besitzen über 25 AS-Austausche mit unterschiedlichen Eigenschaften auf 22 Austausch für AS mit stärker ähnlichen Eigenschaften ab. Zudem gibt es 13 Indels.

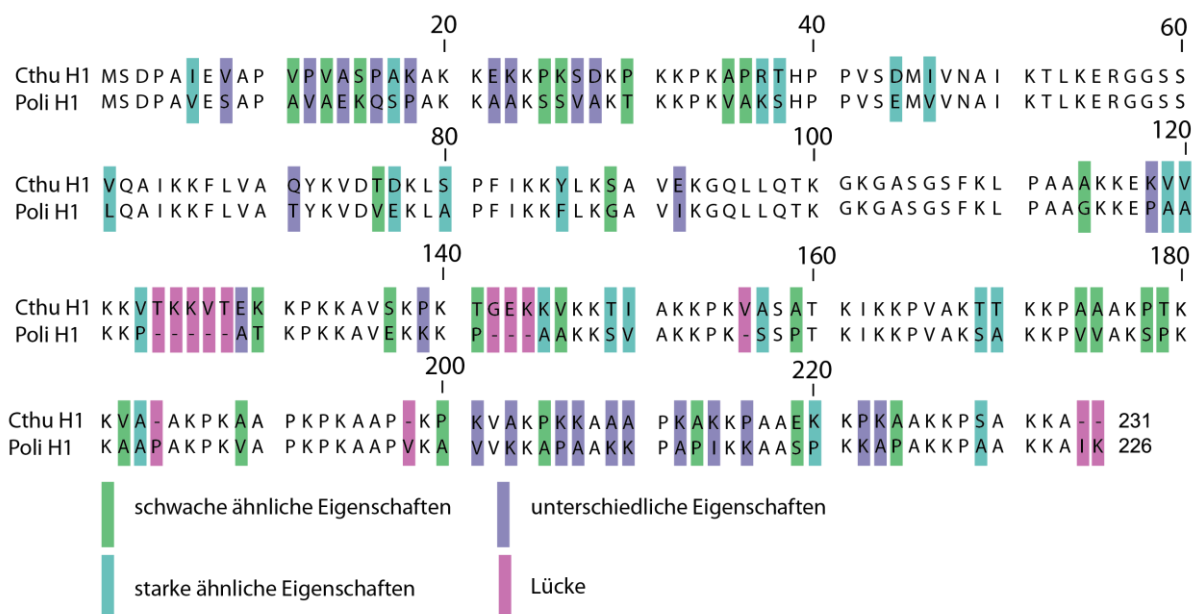


Abbildung 47 Alignment der AS-Sequenzen von H1-Proteinen aus *C. thummi* (Cthu H1) und *P. olivacea* (Poli H1). Grün hinterlegte AS sind AS-Austausche gegen AS mit schwachen ähnlichen Eigenschaften. Türkis unterlegte AS zeigen AS-Austausche gegen AS mit starken ähnlichen Eigenschaften. AS-Austausche gegen AS ohne ähnliche Eigenschaften sind lila unterlegt. AS-Indels sind durch rosa Markierungen dargestellt.

Dabei zeigt das MSA der H1-Proteine das am stärksten unterschiedliche Histon-Protein.

Der Vergleich mit einer dritten Dipteren-Art *D. melanogaster* zeigt, dass das H1-Gen in Bezug auf die Nematoceren-Arten, *C. thummi* und *P. olivacea*, bei *D. melanogaster* invertiert vorliegt. Die Genabfolge von H3/H4 und H2A/H2B im Cluster ist in allen drei Arten identisch (Tabelle

19). Die Länge der jeweiligen Cluster beträgt bei 4,8 bzw. 5 kb in *D. melanogaster* bis 6,2 kb in *C. thummi* (Matsuo und Yamazaki 1989), (Hankeln und Schmidt 1991). Die Zahl der Histon-Gen-Loci ist bei *C. thummi* mit fünf Hybridisierungssignalen auf Chromosom II, Arm D⁴ die höchste (Hankeln und Schmidt 1991). Für *D. melanogaster* ist nur ein Locus für die Histone-Gene bekannt (Pardue et al. 1977). In *P. olivacea* sind die Histon-Cluster in zwei Loci auf Chromosom I, Arm R⁵ aufgeteilt ((Hankeln et al. 1993), Abbildung 22). Daher wird Chromosom-Arm I R als D' bezeichnet. Die Histon-Gene der drei Arten besitzen TAA als Stopp-Kodon (Hankeln und Schmidt 1991), (Matsuo und Yamazaki 1989). Die zuvor in *C. thummi* und *D. melanogaster* identifizierten konservierten 3' *hairpin loops* (c3'h) wurden auch in *P. olivacea* gefunden. Sie sind in *P. olivacea* 45 bis 90 bp hinter dem Stopp-Kodon lokalisiert und stimmen zu 75 % (H2A) bzw. 87,5 % (H1, H3, H4, H2B) mit der Sequenz von *D. melanogaster* überein. Zu *C. thummi* zeigen sie eine 100 % Übereinstimmung.

Tabelle 19 Vergleich der Histon-Cluster aus drei Dipteren-Spezies

	Cthu	Poli	Dmel
Anordnung Gene	<H3 H4> <H2A H2B> H1>	<H3 H4> <H2A H2B> H1>	<H3 H4> <H2A H2B> <H1
Länge des "Cluster"	6262 bp	5921 bp	4,8 oder 5 kb
Southern Analyse	SacI: 6,6 6,2 & 5,6 kb		
FISH Lokalisation	5 Loci: Chr II Arm D ¹	2 Loci: Chr I Arm R ² (D')	1 Locus: Chr 2 Arm L
Stopp-Kodon	TAA	TAA	TAA
con. 3' <i>hairpin</i>	+	+	+
Abstand zu TAA:			
-H1	59 bp	90 bp	40 bis 100 bp
-H2A	32 bp	46 bp	
-H2B	33 bp	46 bp	
-H3	43 bp	66 bp	
-H4	37 bp	45 bp	
<i>orphon</i> Gene	+; alle	+; 2-mal H3	+;3-mal H3 (H3.3A/B; Cid)
H1-Varianten	+	-	-

Unterschiede finden sich dagegen in sogenannten *Orphon*-Histon-Genen, welche außerhalb der Cluster liegen. Für *C. thummi* sind für alle fünf Histone-Gene *Orphon*-Gene bekannt. In *P. olivacea* konnten, ebenso wie in *D. melanogaster*, nur für H3-*Orphon*-Gene identifiziert

⁴ definiert für *C. thummi* von Keyl 1962

⁵ definiert für *P. olivacea* von Zacharias 1979

werden (vgl. 3.4.3). Diese zeigen die Besonderheit, ein oder mehrere Introns zu besitzen. Dies kommt bei den Cluster-Histon-Genen nicht vor. Histon-Sequenz-Varianten sind schon seit über 30 Jahren bekannt, doch über den Mechanismus der Abgrenzung und ihre Funktion ist wenig bekannt (Ahmad und Henikoff 2002a). Auch Histon-Gene, deren mRNA durch Intron-Sequenzen unterbrochen sind, wurden 1982 von Engel et al. beschrieben (Engel et al. 1982). Dabei wurde eine H3-Variante im Huhn mit zwei Intron-Sequenzen identifiziert. Im Huhn wird die intronhaltige H3-Variante in allen Geweben transkribiert, die intronlose H3-Variante dagegen nur im Embryo (Engel et al. 1982). Auch für *P. olivacea* und *C. thummi/piger* konnten zwei bzw. eine H3-Variante mit Intron-Sequenzen identifiziert werden (3.5.4). Es könnte sich bei den in *P. olivacea* bzw. *C. thummi/piger* identifizierten H3-Varianten dementsprechend um Entwicklungsstadien-/gewebsspezifische-Varianten handeln. Vermutlich handelt es sich bei den H3-Varianten *H3V2* aus *P. olivacea* und *C. thummi/piger* wirklich um homologe Genkopien. Die sich nicht nur aus deren intronhaltiger Genstruktur schließen, sondern auch aus der Struktur der chromosomalen Region. In beiden Arten grenzt das Gen *H3V2* an das gleiche Gen (*hematopoietic signal peptide-containing membrane domain-containing 1*, Daten nicht gezeigt). Das Gen *H3V1* aus *P. olivacea* zeigt im Gegensatz zur vorher beschriebenen H3-Variante eine 100 %ige Sequenzidentität zur *D. melanogaster* H3-Variante H3.3. H3.3 hat in *D. melanogaster* eine Funktion in der Transkriptionsregulation bei aktivem Chromatin (Ahmad und Henikoff 2002b). Ob die *H3V1*-Variante in *P. olivacea* die gleiche Funktion ausübt, muss noch überprüft werden.

4.1.3 Low-copy-DNA-Sequenzen

Die Assemblies der *P. olivacea*-DNA-Daten ergaben 14.000 potenzielle intronlose Gene. Die Anzahl scheint möglich, da auch andere Insekten ähnliche Zahlen aufweisen (vgl. Tabelle 2). Es muss aber daraufhin gewiesen werden, dass es durch die große Contiganzahl (Tabelle 6) dazu kommen kann, dass Gene auf zwei Contigs (oder sogar mehr) „verteilt“ liegen. Dies spiegelt dann aber nicht die tatsächliche Genomsituation wider, sondern erhöht künstlich die Anzahl der Gesamtgene. Umgekehrt werden nur Gene welche im Larven-Transkriptom enthalten sind erfasst. Dies ist durch die automatisierte Annotations-Methode bedingt, welche sich auf die *spliced reads* der RNA-Datensätze stützt. Es müssten also weitere Daten erhoben werden, um das Genom von *P. olivacea* umfassend zu charakterisieren.

Homologe Gene aus anderen Dipteren in *P. olivacea*

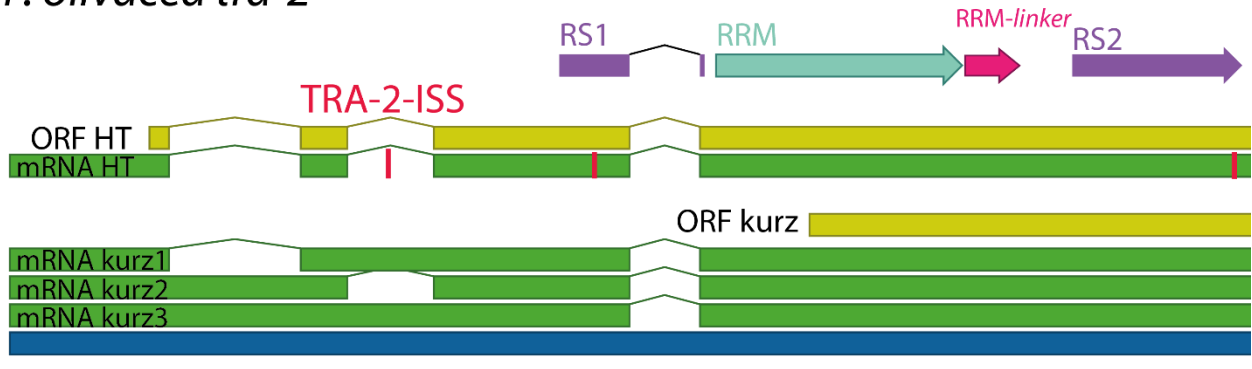
Die Gene der Geschlechtsbestimmungskaskade aus *D. melanogaster*

Über die Geschlechtsbestimmung von *D. melanogaster* ist im Vergleich zu anderen Dipteren bis heute am meisten bekannt (zusammengefasst in (Schütt und Nöthiger 2000)). Die dabei am besten beschriebenen Gene sind *Sxl*, *tra* und *dsx*. Die Gene *Sxl* und *dsx* konnten für die Arten *C. thummi/piger* (C. Kraemer, persönliche Mitteilung, (Frank 2001), (Hammann 2013) und die Art *P. olivacea* identifiziert werden (vgl. 3.6.1). Lediglich für *dsx* scheint eine Mitwirkung in der Geschlechtsbestimmung bisher gesichert. Zudem wurden weitere Gene aus der *D. melanogaster* Geschlechtsbestimmungskaskade in *P. olivacea* identifiziert wie unter anderem *tra-2*, *fru*, *da* und *dis* (Tabelle 13).

Sxl konnte auch in anderen Dipteren-Arten wie *Musca domestica* (Dübendorfer et al. 1992), *C. capitata* (Saccone et al. 1998), *Chrysomya rufifacies* (Müller-Holtkamp 1995) und *Megaselia scalaris* (Mainx 1966), (Traut 1994) identifiziert werden. In den untersuchten Arten gibt es jedoch kein geschlechtsspezifisch gespleißtes *Sxl*-Transkript. Lediglich für *Chrysomya rufifacies* wird von Müller-Holtkamp eine Beteiligung an der Geschlechtsbestimmung diskutiert (Müller-Holtkamp 1995). Für die nahverwandten Arten *C. thummi/piger* wurde eine Rolle von *Sxl* in der Geschlechtsbestimmung durch Frank ausgeschlossen, da das zu *D. melanogaster* homologe Exon 3 fehlt (Frank 2001). Dies unterstützt die Vermutung, dass *Sxl* in *P. olivacea* ebenfalls keine Rolle in der Geschlechtsbestimmung spielt, da Exon 3 auch im *P. olivacea*-Homologen fehlt.

Tra-2 konnte bereits früher in *C. thummi/piger* identifiziert werden (Gerbracht 2014). Im Gegensatz zu *P. olivacea* konnten allerdings keine alternativen Spleißvarianten detektiert werden (*C. thummi/piger*: Daten nicht gezeigt; *P. olivacea*: 3.6.1 und Abbildung 48).

P. olivacea tra-2



D. melanogaster tra-2

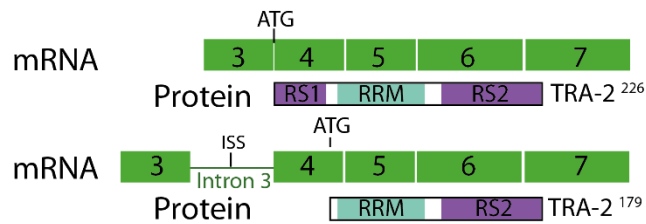


Abbildung 48 Ausschnitt der Genstruktur von *tra-2* aus *P. olivacea* mit den möglichen Transkript-Varianten und zwei der drei in *D. melanogaster* beschriebenen *tra-2*-Isoformen und der abgeleiteten Proteine. Für *P. olivacea* sind die möglichen Transkripte (grün) und deren ORF (gelb) dargestellt. Die Domänen (RS1, RS2, RRM, RRM-linker) der resultierenden Proteine und die cis regulatorischen TRA-2-ISS-Motive (rot) sind eingezeichnet. Für *D. melanogaster* sind 2 Transkripte ebenfalls mit den abgeleiteten Proteindomänen eingezeichnet. Zudem ist das Start-Kodon (ATG) des jeweiligen Transkripts markiert.

In *D. melanogaster* wurde zuerst eine Autoregulation von *tra-2* beschrieben (Mattox et al. 1996). Dabei wird in der männlichen Keimbahn die Expression von *tra-2* mittels einer Transkript-Variante (TRA-2²²⁶) reprimiert, welche das Spleißen von Intron M1 (Intron 3) unterdrückt (Abbildung 48). Dies erfolgt durch den *intronic splicing silencer* (ISS), welcher im besagten Intron vorkommt. Ein solches Motiv wurde 5-mal innerhalb der *P. olivacea tra-2* Sequenz identifiziert (*P. olivacea*: Abbildung 25; *C. thummi/piger*: 6-mal, Daten nicht gezeigt). Die drei in *P. olivacea* unterrepräsentierten Transkripte zeigen alle alternative Varianten des Intron 2, in Bezug auf das Haupt-Transkript, auf. Dieses enthält wie das Intron 3 des *tra-2*-Gens aus *D. melanogaster* ein TRA-2-ISS Motiv, welches dort bei der negativen Autoregulation von *tra-2* in der männlichen Keimbahn beteiligt ist. Beide Varianten sowohl aus *D. melanogaster* als auch aus *P. olivacea* resultieren in einem verkürzten Protein, wobei in *D. melanogaster* lediglich die RS1-Domäne (N-Terminus) fehlt, dagegen in *P. olivacea* zusätzlich zur RS1-Domäne ein Teil des RRM. Diese Gemeinsamkeiten könnten darauf hinweisen, dass es in *P. olivacea* ähnlich wie in *D. melanogaster* eine Autoregulation von *tra-2* gibt. Womöglich findet diese ebenfalls in der Spermatogenese statt. Dagegen spricht jedoch das Vorkommen der Isoformen im Weibchen (vgl. 3.6.1). Zudem beziehen sich die Daten auf RNA-Daten der gesamten Larve.

Um genauere Aussagen treffen zu können, wäre es nötig, Keimbahn-spezifische Expressionsanalysen zu machen.

In anderen Dipteren-Spezies wie *C. capitata* und *B. oleae* wurde das Geschlechts-bestimmende Gen *tra* über die Nähe des Gens *l(3)73Ah* identifiziert (Pane et al. 2002), (Lagos et al. 2007). Dies wurde ebenfalls in *C. thummi/piger* versucht, allerdings ohne den gewünschten Erfolg (persönliche Mitteilung, C. Kraemer). In *P. olivacea* lieferte die Suche nach *l3(73)Ah* ein 20 kb entferntes SR-reiches Protein. SR-Proteine stellen oft Spleißfaktoren mit einem hohen Anteil an S/R dar. Da es sich bei dem von *tra* kodierten Protein um einen solchen Spleißfaktor handelt, wurde die aus dem in *P. olivacea* identifizierten Gen abgeleitete AS-Sequenz mit den AS-Sequenzen von bekannten TRA-Proteinen aus anderen Dipteren verglichen (vgl. 3.6.1). Hier ließ sich jedoch keine Sequenzidentität feststellen. Da *tra* innerhalb der Dipteren schlecht konserviert ist, kann eine Beteiligung an der Geschlechtsbestimmung nicht ausgeschlossen werden. Die Expression, welche sich in männlichen *P. olivacea*-Larven nicht von der aus weiblichen *P. olivacea*-Larven unterscheidet, unterstützt jedoch die Hypothese, dass es sich bei dem 21 kb von *l3(73)Ah* entfernten Gen nicht um *tra* handelt. Auch liegt das Gen bei den oben angesprochenen Arten nur wenige hundert Basen entfernt von *tra* lokalisiert (*B. oleae* und *A. obliqua*; (Lagos et al. 2007), (Ruiz et al. 2007)) bzw. überlappt sogar mit dem *tra*-Gen (*D. melanogaster* und *C. capitata*; (Irminger-Finger und Nöthiger 1995), (Pane et al. 2002)).

Besser konserviert ist dagegen das in der Kaskade *downstream* vorkommende Gen *dsx*. Dieses zeigt über die Klasse der Insekten hinaus eine Beteiligung an der Geschlechtsbestimmung (Raymond et al. 1998), (Shen und Hodgkin 1988). Für *C. thummi/piger* (Frank 2001), (Hammann 2013) und *P. olivacea* wurde ein *dsx*-orthologes sowie ein -paraloges Gen identifiziert. Beide orthologen Gene weisen geschlechtsspezifisches Spleißen des C-Terminus auf (*C. thummi/piger*: (Hammann 2013); *P. olivacea*: 3.6.1, Abbildung 32), wobei die RT-PCR in *P. olivacea* mit Weibchen-spezifischen Primern im Männchen zwei verschiedene Amplifikate aufzeigt (vgl. Abbildung 32, B). Die 700 bp Bande umfasst die genomische Region und erklärt sich durch einen unvollständigen DNase-Verdau der zur cDNA-Synthese eingesetzten RNA. Die Bande mit ~200 bp stimmt mit der Sequenz des Weibchen-spezifischen Transkripts überein. Eine basale Transkription des Weibchen-spezifischen Transkripts wurde zuvor in der Honigbiene, dem Seidenspinner und der Erzwespe beschrieben (Cho et al. 2007), (Suzuki et al. 2001), (Verhulst 2011). Es wird postuliert, dass es sich bei der weiblichen Isoform um den

Grundzustand handelt (Cho et al. 2007). In *D. melanogaster* ist es dagegen umgekehrt, dort ist die männliche Isoform der Grundzustand. Das heißt, in *P. olivacea* muss entweder das Ausschalten eines Weibchen-spezifischen Faktors oder das Einschalten eines Männchen-spezifischen Faktors wie M erfolgen, um die Männchen-spezifische *dsx*-Isoform zu exprimieren.

Der Spleißmechanismus, welcher zu den geschlechtsspezifischen Isoformen führt, ist für die Art *D. melanogaster* bekannt (Lynch und Maniatis 1995), (Inoue et al. 1992). Dabei handelt es sich um eine schwache 3'Spleißakzeptorstelle des Introns vor dem Weibchen-spezifischen Exon 4 und das Vorkommen von TRA/TRA-2-Motive und PREs im 3'-UTR des Weibchen-spezifischen Transkripts. Für die Arten *A. aegypti*, *A. gambiae* und *C. quinquefasciatus* wurden unterschiedliche Hypothesen aufgestellt, die es noch zu überprüfen gilt (Salvemini et al. 2011), (Scali et al. 2005), (Price et al. 2015). Für *P. olivacea* kann aufgrund vorheriger Analysen (Abbildung 33, Tabelle 10, Tabelle 11 und Tabelle 12) angenommen werden, dass es sich bei der 3'Akzeptorspleißstelle vor dem Weibchen-spezifischen Exon 8 nicht um eine schwache Spleißstelle handelt. Weiter wurde die Mehrzahl der regulatorischen Elemente (TRA/TRA-2 und PREs, Abbildung 33) zwar im 3'-UTR des Weibchen-spezifischen Transkripts ermittelt, jedoch näher an der 5'Donorspleißstelle des Introns nach dem Weibchen-spezifischen Exons 8. Dies ist vergleichbar mit dem Zustand, der für das *dsx*-Transkript in *A. gambiae* beschrieben wurde (Scali et al. 2005). Möglicherweise liegt also ein ähnlicher Spleißmechanismus zwischen den beiden Arten vor.

Von den im Weibchen zusätzlich auftretenden Isoformen der 5'Region von Exon 7 (Abbildung 49,1.1) wird jedoch angenommen, dass es sich bei der kurzen Variante 7.2 um eine schwächere Spleißstelle als die der Variante 7.1 handelt (Tabelle 10). Dies wird auch von den NGS-Daten unterstützt.

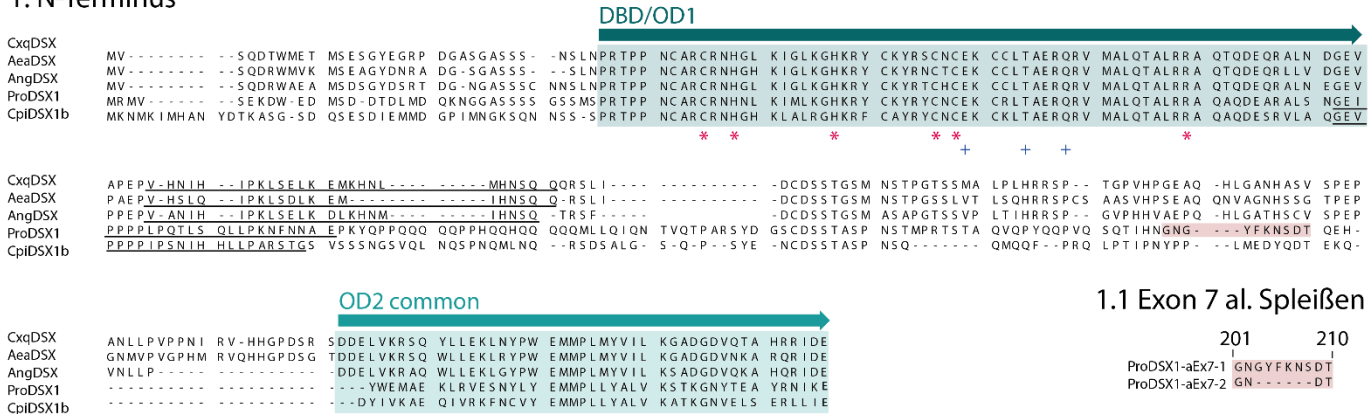
In der Familie der Culicidae sind bereits DSX-Proteine der Arten *A. aegypti*, *A. gambiae* und *C. quinquefasciatus* beschrieben (Salvemini et al. 2011), (Scali et al. 2005), (Price et al. 2015). Der Vergleich der Nematoceren DSX-Sequenzen zeigt, dass die DNA-Bindung durch die sechs verantwortlichen AS-Reste für PoliDSX1 und PoliDSX2 gewährleistet ist (* in Abbildung 49, 1). Diese sechs Reste liegen in *P. olivacea* konserviert vor. Zudem sind die drei Reste, welche für die DM-Domäne in DSX charakteristisch sind, ebenfalls konserviert (+ Abbildung 49, 1). In

dem MSA fällt weiter auf, dass das Moskito-spezifische Spleiß-Ereignis (Price et al. 2015) (variables *in-frame* Intron; Abbildung 49, 1, unterstrichen im MSA) nach der DM-Domäne ebenfalls in *P. olivacea* (69 bp) und in *C. thummi/piger* (63 bp) auftritt. In *A. aegypti* ist es wie in *C. thummi/piger* 63 bp und in *A. gambiae* 72 bp (Salvemini et al. 2011) lang und in *C. quinquefasciatus* mit 75 bp (Price et al. 2015) am längsten.

Der Vergleich der OD2-Domänen bestätigt vorherige Analysen, dass die Weibchen-spezifische OD2-Domäne zwischen verschiedenen Arten stärker konserviert ist als die OD2-Domäne der DSX^M-Proteine. Für den Männchen-spezifischen Teil wird eine schnellere Veränderung im Vergleich zur Weibchen-Variante angenommen (Price et al. 2015).

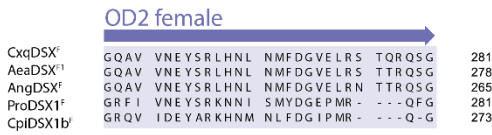
Eine Besonderheit konnte in *P. olivacea* ermittelt werden, hier zeigen sich Transkripte im Männchen, die sich am 3'Ende unterscheiden (Abbildung 49, blau unterlegt). Ein vorzeitiger Translationsstopp führt zu zum Teil verkürzten Proteinen. Zuvor wurden mehrere Isoformen nur für Weibchen-Transkripte in anderen Insekten wie *Apis mellifera* oder *A. aegypti* beschrieben (Cho et al. 2007), (Salvemini et al. 2011).

1. N-Terminus



* AS verantwortlich für DNA-Bindung
+ spezifische Aminosäuren für DSX-DM-Domäne

2. weibchen-spezifische Region



3. männchen-spezifische Region

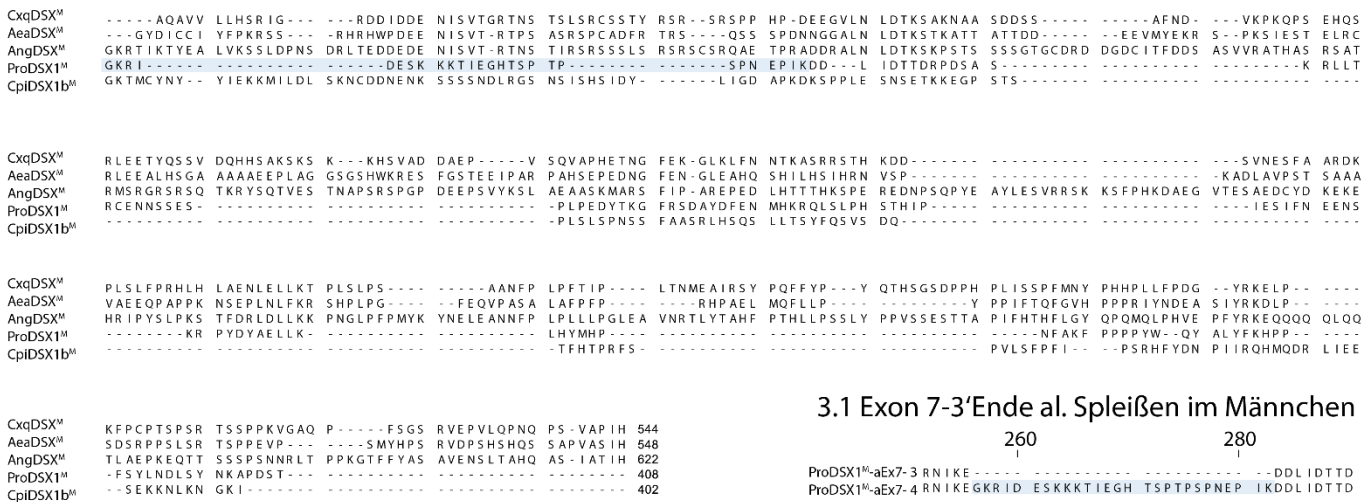


Abbildung 49 Nematoceren DSX-MSA.

1. N-Terminus der DSX-Proteine: Dunkelgrün bzw. Türkis unterlegt sind die gemeinsamen, in beiden Geschlechtern vorkommenden Domänen DM-Domäne (DBD/OD1) bzw. OD2common. AS-Reste der DBD, welche für die DNA-Bindung zuständig sind, sind mit einem * gekennzeichnet. AS-Reste, welche spezifisch für DSX-DM-Domänen sind, sind mit einem + gekennzeichnet. Das Nematoceren-spezifische in-frame Intron ist schwarz unterstrichen.

1.1 Alternatives Spleißen des 5'Ende von Exon 7 ist rosa unterlegt.

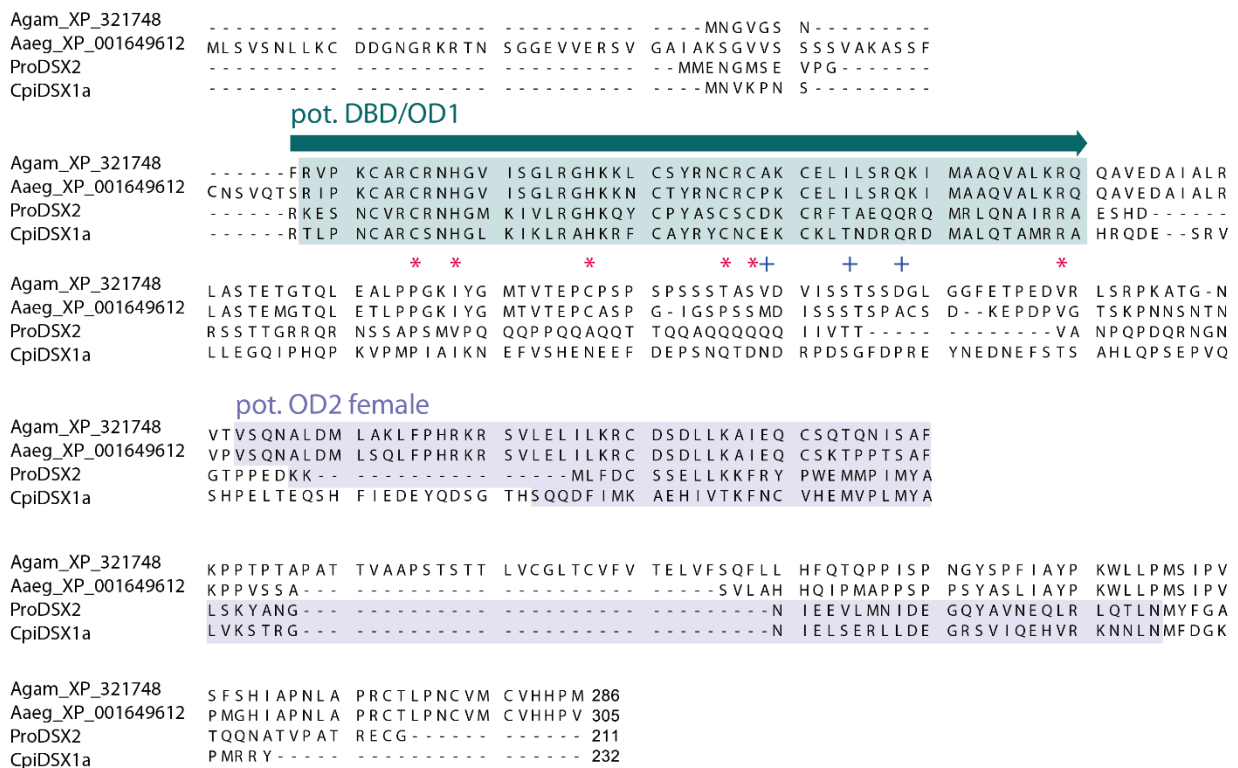
2. Weibchen-spezifische OD2 Alignment (violett unterlegt)

3. Männchen-spezifisches MSA.

3.1 Alternatives 3'Ende im Männchen von Exon 7 ist grau unterlegt.

In *C. thummi/piger* wurden zuvor zwei Gene, welche mit *dsx* Sequenzübereinstimmungen aufweisen, identifiziert. *Dsx1b* zeigt signifikante Übereinstimmung und wurde als *Dmeldsx*-Orthologes identifiziert. Das Gen *dsx1a* ist paralog zur Genkopie *dsx1b* in *Chironomus* (untersucht in den Arten *C. piger*, *C. thummi*, *C. luridus* und *C. annularius*, Daten nicht gezeigt), es gibt vor allem signifikante Sequenzübereinstimmungen der AS-Sequenzen, welche der DM-Domäne entsprechen. So sind in beiden resultierenden Proteinen der Genkopien die konservierten AS-Reste für die DNA-Bindung und die DSX-DM spezifischen-Domäne konserviert (Abbildung 49 und Abbildung 50).

dsx- Duplikation in Nematocera-Spezies



* Aminosäure verantwortlich für DNA-Bindung
+ spezifische Aminosäure für DSX-DM-Domäne

Abbildung 50 MSA der Nematoceren DSX-Paralogen. Dargestellt ist ein MSA mit den AS-Sequenzen der dsx-Kopien aus vier Nematoceren-Spezies *A. gambiae*, *A. aegypti*, *C. piger* und *P. olivacea*. Dunkelgrün unterlegt ist die gemeinsame, in beiden Geschlechtern vorkommende, potenzielle DSX-DM-Domäne (DBD/OD1). AS-Reste der DBD, welche für die DNA-Bindung zuständig sind, sind mit einem * gekennzeichnet. Positionen mit spezifischen AS-Resten für DSX-DM-Domänen, sind mit einem + gekennzeichnet. Die potenzielle Weibchen-spezifische Region OD2 ist im Alignment violett.

Das daraus resultierende Protein umfasst 211 AS. PoliDSX2 zeigt im MSA ebenfalls die konservierten sechs AS-Reste. Von den drei DSX-DM-Domänen-spezifischen AS-Resten (Oliveira et al. 2009) sind jedoch nur zwei (T, Q) konserviert. Anstelle der konservierten

Glutaminsäure erfolgte ein Austausch zur Asparaginsäure (Abbildung 50). Eine SMART-Domänen-Suche ergab signifikante Treffer für die DM-Domäne ($e\text{-value } 1,51 \times 10^{-16}$) und für die DSX-Dimer-Domäne ($e\text{-value } 1,18 \times 10^{-13}$). Auch über die Familie der Zuckmücken hinaus wurden paraloge Gene von *dsx* gefunden (in *A. aegypti*; XP_001649612; (Salvemini et al. 2011)). Im Rahmen dieser Arbeit wurde ein paraloges Gen noch per Datenbanksuche für *A. gambiae* identifiziert (XP_321748). Das Gen umfasst 1,7 kb und ist wie *Agdsx* auf Chromosom 2 R lokalisiert. Ein MSA (Abbildung 50) der AS-Sequenzen, welche aus den paralogenen *dsx*-Gensequenzen abgeleitet sind, zeigt, dass die sechs AS-Reste der DM-Domäne konserviert sind (Abbildung 50,*). Von den drei AS-Resten (+), welche als DSX-DM Domänen-spezifisch identifiziert wurden (Oliveira et al. 2009), ist in *A. aegypti* bzw. *A. gambiae* jedoch nur das Glutamin (Q) konserviert. Der N-Terminus von AeaDSXp ist im Vergleich zu den anderen drei Proteinen ungefähr um 40 AS verlängert. In PoliDSX2 fand eine Insertion/Deletion von 13 AS statt. Die Region, welche der putativen OD2 (Weibchen-spezifisch)-Domäne entspricht, ist zwischen *Cpidsx1a* und *Polidsx2* besser konserviert als in den paralogenen Genen aus *A. aegypti* bzw. *A. gambiae*. Weiter zeigen die AS-Sequenzen aus *A. aegypti*, *P. olivacea* und *C. thummi/piger* eine Insertion/Deletion (28 AS in *A. aegypti*, 32 in *P. olivacea* und *C. thummi/piger*) am C-Terminus der putativen OD2-Domäne. Am C-Terminus von PoliDSX2 und *CpiDSX1a* hat ebenfalls eine Insertion/Deletion (12 AS in PoliDSX2; 21 AS in *CpiDSX1a*) stattgefunden.

Ob die paraloge Genkopie, welche über die Zuckmücken hinaus identifiziert werden konnte, eine Rolle in der Geschlechtsbestimmung spielt, ist nicht eindeutig. Die resultierenden Proteine zeigen eine höhere Ähnlichkeit zur weiblichen Isoform, jedoch kein geschlechtsspezifisches Spleißen, welches in unterschiedlichen Proteinen resultiert. Eine Funktion als Transkriptionsfaktor sollte jedoch nicht ausgeschlossen werden, da die für die DNA-Bindung verantwortlichen sechs AS-Reste konserviert vorliegen. Zudem ließ sich eine stärkere Diversifizierung der Gensequenzen bzw.-strukturen gegenüber seinem Vorläufer-Gen ausmachen. Die geschlechtsunabhängig-alternativ-gespleißten *dsx*-Transkripte in *P. olivacea* unterscheiden sich im 5'-UTR. Die Vielzahl der *dsx*-Transkripte kann dadurch erklärt werden, dass als Grundlage die gesamte Larve eingesetzt wurde. Wenn es sich also bei den Transkript-Varianten um gewebs- oder stadienspezifische Transkripte handelt, konnten sie auf diese Art nicht identifiziert werden. Eine gewebs- bzw. stadienspezifische Funktion der

Transkripte mit unterschiedlichem 5'-UTR kann nicht ausgeschlossen werden, da zuvor schon Analysen zu UTR-Sequenzen gemacht wurden. Diese zeigen beispielsweise, dass der UTR Einfluss auf Transkriptstabilität und Translationseffizienz hat (Hughes 2006).

Eine Vielzahl von Autoren hat sich bereits mit den Zielgenen von *DSX^M* bzw. *DSX^F* beschäftigt, welche für die Geschlechtsdifferenzierung verantwortlich sind. Einige wurden von Verhulst und van de Zande zusammengefasst (Verhulst und van de Zande 2015). In *P. olivacea* führten ähnliche Analysen zu einem im Weibchen hochregulierten Speicherprotein *hexamerin-2* (vgl. 3.6.1). Da sich die von Verhulst und van de Zande untersuchten Gene meist auf andere, hier nicht untersuchte, Entwicklungsstadien (Puppe, Imago) beziehen, können hier keine weiteren Überlegungen angestellt werden. Dazu müssten weitere Daten erhoben werden (Verhulst und van de Zande 2015).

Die Gene aus der SDR von *C. thummi*

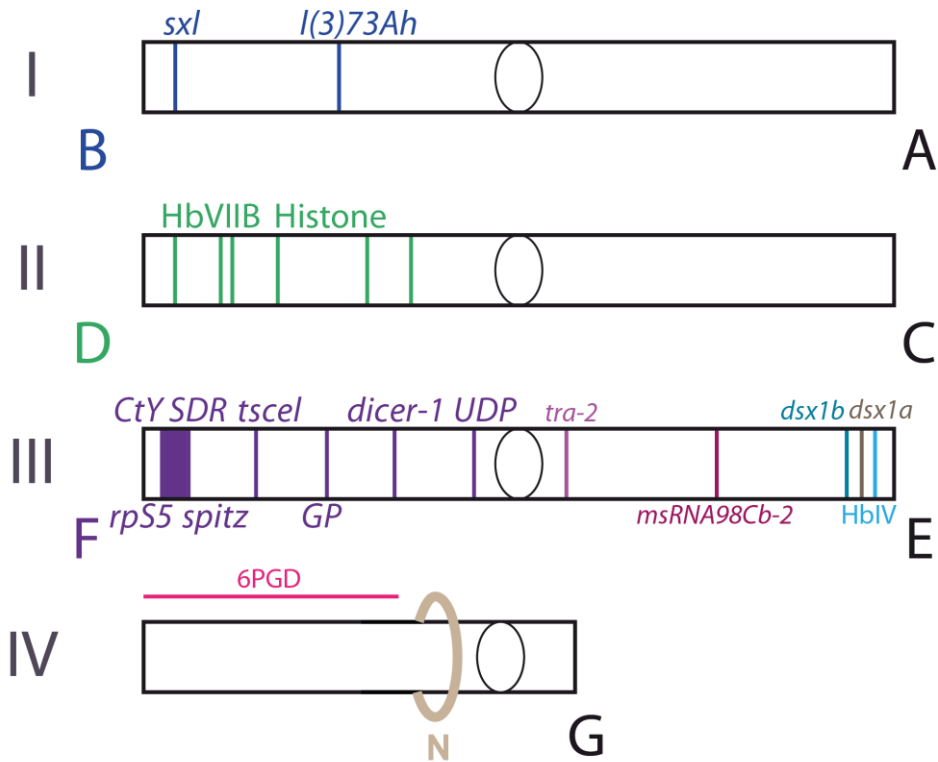
Die Rolle von *dsx* in der Geschlechtsbestimmung der Zuckmücken-Arten *P. olivacea* und *C. thummi/piger* scheint ein verbindendes Element zu sein (vgl. 3.6.1 und 4.2.1). Ansonsten weisen die zwei Arten z. T. große Unterschiede auf. Auf zytologischer Ebene zeigt *P. olivacea* laut Zacharias ein distinktes Y-Chromosom, wogegen *C. thummi/piger* ein Proto-Y-Chromosom aufweist (Zacharias 1981). Über die Identifikation der SDR aus *C. thummi/piger* in *P. olivacea* konnten eine Vielzahl von Chromosomenumbauten beobachtet werden (3.8, Abbildung 43 und Tabelle 18), ein „Kern“ von 10 syntänen Genen konnte ermittelt werden. Unter diesen 10 Genen befindet sich auch der stärkste Kandidat für M, *fs(1)K10-like*. Dieser liegt in männlichen *C. piger/thummi*-Individuen dupliziert vor (Stiehl 2009). Eine Duplikation des Gens in *P. olivacea*, am Locus der potenziellen SDR ist jedoch mit großer Wahrscheinlichkeit auszuschließen. In einer genomischen Southern Analyse dieser Region (Abbildung 40) traten zwar zur Referenz unterschiedliche Ergebnisse auf, diese lassen sich jedoch auf das ebenfalls in der Region untersuchte Gen 6 *PGD* zurückführen. Auffällig ist jedoch, dass eben dieses Gen *downstream* von *fs(1)K10-like* liegt. Dort fand das Duplikationsereignis in *C. thummi/piger/luridus* statt. Die Annahme, dass es sich bei dieser Region um einen Hotspot für Mutationsereignisse handelt, wird durch einen kurzen, dennoch signifikanten Treffer zu einem TE (*jockey*-Element, LINE) homologen noch verstärkt werden.

In *Bombyx mori* konnte gezeigt werden, dass das W-Chromosom viele TE enthält (Abe et al. 1998), (Abe et al. 2005b), (Abe et al. 2005a). Somit könnte es sich bei den Sequenzen auf Arm 3R um TE oder andere repetitive Sequenzen handeln. Ein Zusammenhang zwischen der Amplifikation von Mikrosatelliten und der Diversifizierung von Geschlechtschromosomen wurde vor kurzer Zeit von Matsubara et al. untersucht (Matsubara et al. 2015).

4.2 Die Chromosomenevolution der Chironomiden

Chromosomenumbauten der Gattung *Chironomus* sind schon intensiv in der Vergangenheit untersucht und für phylogenetische Analysen benutzt worden (Beermann 1962), (Keyl 1961), (Keyl 1962). Es gelang für drei der insgesamt sieben Chromosomenarme eine Einteilung nach Querscheibenabfolge. So gelang die Aufstellung eines Schemas zur phylogenetischen Beziehung für viele Arten innerhalb der Gattung *Chironomus*. Für Chironomiden außerhalb der Gattung *Chironomus* ist deutlich weniger bekannt. Chironomiden der Gattung *Prodiamesa* zeigen einen Karyotyp von $n=3$, während die meisten bekannten *Chironomus*-Arten einen Karyotyp von $n=4$ aufweisen. Dies macht den Vergleich des Querscheibenmusters schwieriger. Mit Hilfe von FISH-Ergebnissen (zusammengefasst in Abbildung 43) ist es jedoch möglich, homologe Sequenzen auf Chromosomenarmen bzw. -abschnitten in *P. olivacea* und *C. thummi/piger* zu lokalisieren und damit auch Chromosomen(arme) zu benennen. Auf diese Weise gelang es in der Arbeit die Arme B', D' und F' identifizieren. Für die Arme A und C aus *Chironomus* konnte keine entsprechende Gensonde generiert und somit keine entsprechende DNA-Sequenz in *P. olivacea* lokalisiert werden. Ein Gen des Chromosomenarms G aus *C. thummi/piger* (6PGD) wurde in *P. olivacea* auf Arm F' identifiziert. Da Arm G in einigen anderen *Chironomus*-Arten durch Translokation mit dem Arm E fusioniert ist ((Keyl 1962), Abbildung 51), liegt es nahe, dass eine Fusion von Arm G an F' in *P. olivacea* der Grund für einen Karyotyp von $n=3$ sein könnte.

C. thummi



P. olivacea

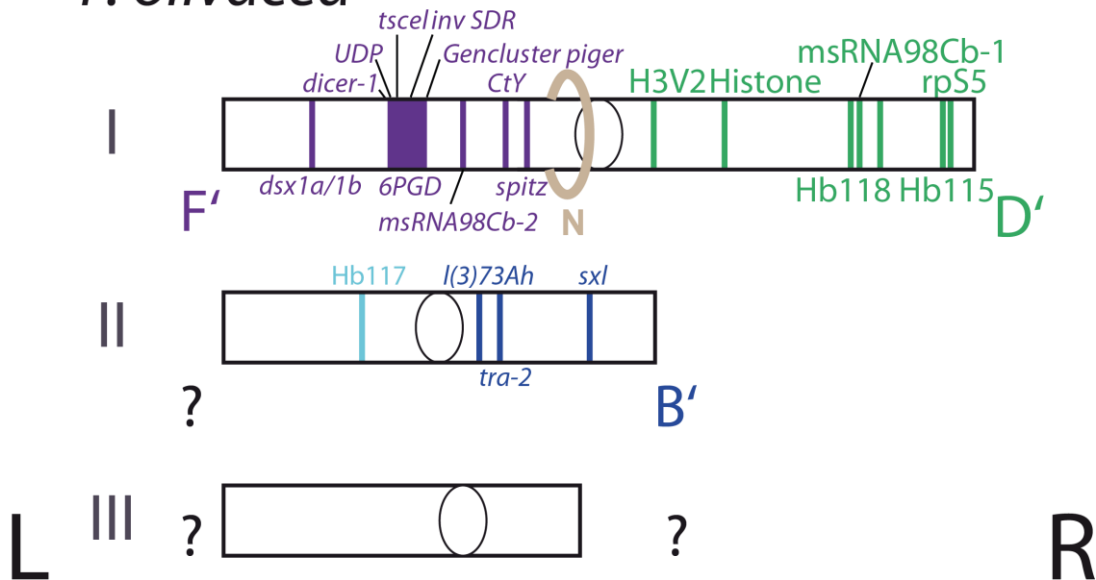


Abbildung 51 Schema der Chromosomen aus *P. olivacea* und *C. thummi*. Zudem wurden alle zuvor in *C. thummi/piger* auf Arm E identifizierten Gene in *P. olivacea* auf unterschiedlichen Chromosomenarmen lokalisiert (Abbildung 43 und Tabelle 18). Somit konnte keiner der in *P. olivacea* verbleibenden Arme eindeutig Arm E zugeordnet werden. Vielmehr liegt es nahe, dass Veränderungen in der Gattung *Chironomus* oder nach der Abspaltung der Gattung

Prodiamesa in *P. olivacea* stattgefunden haben muss. Um eine dieser Vermutungen zu bestätigen, müssten jedoch weitere Analysen mit einer sogenannten „Outgroup“ durchgeführt werde. Dazu bietet sich eine Art der Gattung *Orthocladinae* an (Abbildung 1), da diese phylogenetisch gesehen zwischen den beiden untersuchten Gattungen steht.

Die in *P. olivacea* identifizierten Globin-Gene ließen bisher keinen Rückschluss auf die evolutionären Vorgänge zwischen der Gattung *Chironomus* und *Prodiamesa* zu, da diese auf Sequenzebene weder dem in *C. thummi* identifizierten Cluster D noch E zugeordnet werden konnten. Die Lokalisation von Poli Gb115 und Poli Gb118 erfolgte allerdings am Chromosomenarm D' (Abbildung 16, Abbildung 43 und Tabelle 18). Deshalb könnten die Gene des Clusters 115 und 118 möglicherweise Homologe der Globin-Gene von Cluster D aus *C. thummi* darstellen. Die Anzahl der Kopien in den Clustern zwischen den Arten unterscheidet sich jedoch erheblich. Die Globin-Gene unterliegen offenbar starken Veränderungen in relativ kurzer Zeit. Dabei ist auch zu berücksichtigen, dass *P. olivacea* möglicherweise andere Funktionen der Globin-Proteine nutzt aus *Chironomus*.

5 Zusammenfassung

Die NGS-Sequenzierungen von *P. olivacea*-DNA bzw. RNA ermöglichte eine partielle Analyse des ~113 Mb (Zacharias 1979) großem *P. olivacea*-Genoms. Das hier sequenzierte Genom deckt davon ~74 Mb ab und umfasst ~14.000 Gene (vgl. Tabelle 6). Der Anteil an hochrepetitiver Sat-DNA kann grob auf 3 % geschätzt werden. Die Sat-DNA umfasst eine 168 bp-lange Sequenz mit 64 % AT-Gehalt, welche überwiegend oder auch vollständig mit ~14.000 Kopien in den Zentromerregionen der Chromosomen lokalisiert ist (vgl., 3.3, Abbildung 14, Tabelle 7). Es konnten 10 intronhaltige Hämoglobin-Gene an drei chromosomalen Loci identifiziert werden (vgl. Abbildung 16). Nach Analyse der abgeleiteten AS-Sequenz, der Einteilung der Helices A bis H und der möglichen 3D-Struktur wird angenommen, dass fünf der zehn abgeleiteten Hämoglobinproteine eine Rolle im Sauerstofftransport/-speicherung spielen. Allerdings trägt keins der Polypeptide am N-Terminus das 16-AS-lange Signalpeptid, welches die Sekretion des Hämoglobins aus den Hämoglobin-synthetisierenden "Fettzellen" ermöglichen könnte (vgl. Abbildung 45, Abbildung 46). Die AS-Sequenz-Analyse ergab weiter, dass es sich bei den *P. olivacea*-Hämoglobinen um Monomere handeln muss. Vier der sechs in *P. olivacea*-Hämoglobinen vorkommenden Intronpositionen (vgl. Abbildung 44) stimmen nicht mit den aus anderen Dipteren zuvor identifizierten Positionen überein und unterstützen somit die Hypothese von Intronintegrationsereignissen (Hankeln et al. 1997). Die identifizierten Histon-Cluster-Gene zeigen wie erwartet hochkonservierte Organisation und Sequenzen (vgl. Abbildung 21). Daneben konnten zwei Orphon-H3-Varianten identifiziert werden (vgl. 3.5.4), jedoch keine Orphon-H1-Variante wie in *C. thummi*. Die Orphon-H3-Variante 1 zeigt große Übereinstimmung zur in *D. melanogaster* identifizierten H3.3-Variante, wogegen für die H3-Variante 2 eine homologe Variante in *C. thummi* identifiziert werden konnte. Die Analyse von homologen Genen aus der Geschlechtsbestimmungskaskade aus *D. melanogaster* ergab für 13 von 17 Genen (inklu. *Sxl* und *tra-2*) homologe Gene in *P. olivacea* (vgl. 3.6.1). Dabei ist nur für das Gen *dsx* eine Rolle in der Geschlechtsbestimmung wahrscheinlich. Dies wird durch die larvalen Transkriptom-Daten von *P. olivacea*-Weibchen bzw. -Männchen unterstützt (vgl. Abbildung 32). Darüber hinaus konnte ein orthologes *dsx*-Gen in *P. olivacea* identifiziert werden. Dies konnte zuvor für die Moskito-Arten *A. aegypti* (Salvemini et al. 2011), *A. gambiae* (Scali et al. 2005) und *C. quinquefasciatus* (Price et al. 2015) sowie in *Chironomus (thummi, piger, luridus, annularius)* ((Frank 2001), diese Arbeit) erfolgen. Eine weitere Besonderheit stellt ein

in-frame-Intron des *dsx*-Homologen dar, welches zuvor als Moskito-spezifisch identifiziert wurde (Price et al. 2015). Dies ist ebenfalls in *P. olivacea* und *C. thummi/piger* zu finden. Nach der Analyse der Spleiß-Akzeptor und -Donorstellen ist ein Spleißmechanismus von *dsx* in *P. olivacea* wie in *A. gambiae* am wahrscheinlichsten (vgl. 3.6.1 *dsx*, (Scali et al. 2005)). Für die homologen Gene aus der SDR von *C. thummi* konnten 10 Homologe identifiziert werden. Eine Southern-Analyse ergab für das Gen *fs(1)K10-like* keine Duplikation in *P. olivacea*-Männchen (vgl. 3.6.2) wie es in *C. thummi/piger/luridus*-Männchen der Fall ist ((Stiehl 2009), Stiehl unveröffentlicht, C. Kraemer persönliche Mitteilung). Somit ist eine Beteiligung der homologen Gene aus der SDR von *C. thummi* in *P. olivacea* an der Geschlechtsbestimmung eher unwahrscheinlich. Über die Lokalisation von 20 Genen bzw. Contigs in *P. olivacea* und 18 Genen/Contigs in *C. thummi/piger* konnte eine Kartierung der *P. olivacea*-Chromosomenarme 1L = F', 1R = D' und 2R = B' erreicht werden (vgl. 3.8).

6 Abkürzungsverzeichnis

<i>A. aegypti</i>	<i>Aedes aegypti</i>
<i>A. darlingi</i>	<i>Anopheles darlingi</i>
<i>A. gambiae</i>	<i>Anopheles gambiae</i>
<i>A. obliqua</i>	<i>Anastrepha obliqua</i>
AS	Aminosäure(n)
asm	Assembly
B.	Bereiche
<i>B. antarctica</i>	<i>Belgica antarctica</i>
<i>B. oleae</i>	<i>Bactrocera oleae</i>
bp	Basenpaare
BLAST	<u>B</u> asic <u>L</u> ocal <u>A</u> lignment <u>S</u> earch <u>T</u> ool
bzw.	beziehungsweise
C.	<i>Chironomus</i>
<i>C. capitata</i>	<i>Ceratitis capitata</i>
<i>C. elegans</i>	<i>Caenorhabditis elegans</i>
ca.	zirka
<i>c3'h</i>	<i>conserved 3' hairpin</i>
°C	Grad Celsius
cDNA	<i>complementary DNA</i>
CI	Chloroform-Isoamylalkohol
D.	<i>Drosophila</i>
<i>da/DA</i>	<i>daughterless</i> (Gen), DAUGHTERLESS (Protein)
DBD	DNA bindende Domäne
def	default
<i>dis</i>	<i>dissatisfaction</i>
DM	<i>dsx/mab-3</i> -Domäne
<i>dmrt</i>	<i>dsx/mab-3-related</i>
DNA	<i>deoxyribonucleic acid</i> bzw. Desoxyribonukleinsäure
<i>dpn</i>	<i>deadpan</i> (Gen)

<i>dsx</i> /DSX	<i>doublesex</i> (Gen)/DOUBLESEX (Protein)
E	East bzw. Ost
<i>emc</i>	<i>extra machrochatae</i> (Gen)
EtOH	Ethanol
Fa.	Firma
FISH	Fluoreszenz- <i>in-situ</i> -Hybridisierung
<i>fl(2)d</i>	<i>female lethal (2) d</i> (Gen)
<i>fru</i> /FRU	<i>fruitless</i> (Gen)/FRUITLESS (Protein)
g	Gramm
x g	mittlere Erdbeschleunigung
gDNA	genomische DNA
gef.	gefällt
gew.	gewickelt
<i>gro</i>	<i>groucho</i> (Gen)
GTC	Guanidinthiocyanat
H1, 2A, 2B, 3, 4	Histon 1, 2A, 2B, 3 ,4
<i>her</i>	<i>hermaphrodite</i> (Gen)
His	Histidin
HLPC	<i>high performance liquid chromatography</i>
H ₂ O	Wasser
H3V1/2	Histon3-Variante 1/2
HT	Haupt-Transkript
IMSB	Institut für Molekulargenetik, gentechnologische Sicherheitsforschung und Beratung
Indel	Insertion-Deletion
<i>ix</i>	<i>intersex</i>
kb	Kilo-Basenpaare
L4	Larvenstadium 4
LF	<i>length fraction</i>
LGRM	Large Gap Read Mapping
M	Monomer

M1,M2, M3, M4	Marker 1, Marker 2, Marker 3, Marker 4
mA	Milliampere
<i>mab-3</i>	<i>male abnormal-3</i>
Mb	Megabase
min	Minuten
mind.	mindestens
miRNA	micro RNA
ml	Milliliter
mRNA	<i>messenger</i> RNA
MSA	Multiples Sequenzalignment
µl	Mikroliter
N	Nord
ng	Nanogramm
NGS	<i>Next Generation Sequencing</i>
nt	Nukleotid
OD1/2	Oligodimerisierungsdomäne 1/2
ORF	<i>open reading frame</i> (offener Leserahmen)
<i>P.</i>	<i>Prodiamesa</i>
PCI	Phenol-Chloroform-Isoamylalkohol
PCR	<i>polymerase chain reaction</i> bzw. Polymerasekettenreaktion
pg	Pikogramm
Phe	Phenylalanin
pi	piger
pUC	<i>plasmid <u>U</u>niversity of <u>C</u>alifornia</i>
RRM	<i>RNA recognition motif</i>
RNA	<i>ribonucleic acid</i> bzw. Ribonukleinsäure
RS1/2	R-S-reiche Domäne 1/2
RT	Raumtemperatur
Sat-DNA	Satelliten-DNA
SDR	<i>sex determing region</i> bzw. geschlechtsbestimmende Region

SDS	<i>sodium dodecyl sulfate</i> bzw. Natriumlaurylsulfat
<i>snf</i>	<i>sans fille</i> (Gen)
SIM	<i>similarity fraction</i>
<i>sis A/B/C</i>	<i>sisterless A/B/C</i> (Gene)
<i>Sxl/SXL</i>	<i>Sex lethal</i> (Gen)/SEX LETHAL (Protein)
TBE	Tris-Borat-EDTA
TE	transposable Elemente
TF	Transkriptionsfaktor
th	thummi
<i>tra/TRA</i>	<i>transformer</i> (Gen)/TRANSFORMER (Protein)
<i>tra-2/TRA-2</i>	<i>transformer-2</i> (Gen)/TRANSFORMER 2 (Protein)
ü. N.	über Nacht
Upm	Umdrehung pro Minute
UTR	untranslatierte Region
vgl.	vergleiche
<i>vir</i>	<i>virilizer</i> (Gen)
Vol.	Volumen
z. B.	zum Beispiel
z. T.	zum Teil

7 Abbildungsverzeichnis

Abbildung 1 Stammbaum der Familie der Zuckmücken (Chironomidae) und der Gattung <i>Chironomus</i>	2
Abbildung 2 <i>P. olivacea</i>	7
Abbildung 3 Hypothetische Zusammensetzung des dritten Chromosomenpaares in weiblichen und männlichen diploiden <i>P. olivacea</i> -Individuen.....	8
Abbildung 4 Zusammensetzung eines eukaryotischen Genoms.	9
Abbildung 5 Sat-DNA aus <i>C. thummi/piger</i>	12
Abbildung 6 Mittelrepetitive DNA-Sequenzen eines eukaryotischen Genoms.	13
Abbildung 7 Intronpositionen der Hämoglobin-Gene verschiedener Dipteren-Arten.....	16
Abbildung 8 Schema des DSX ^F bzw. ^M Proteins aus <i>D. melanogaster</i> mit seinen funktionellen Proteindomänen.	19
Abbildung 9 Kopfkapsel einer <i>P. olivacea</i> -Larve.	22
Abbildung 10 Genitalimaginalscheiben von <i>P. olivacea</i> im Vergleich zwischen Weibchen und Männchen.	23
Abbildung 11 Gelelektrophoretische Auftrennung genomischer <i>P. olivacea</i> -DNA.	44
Abbildung 12 Konzentrations- und Qualitätsbestimmung der isolierten Gesamt-RNA aus weiblichen bzw. männlichen <i>P. olivacea</i> -Larven.	45
Abbildung 13 Southern Blot genomischer <i>P. olivacea</i> -DNA.	47
Abbildung 14 Ergebnis der FISH an polytären Speicheldrüsenchromosomen aus <i>P. olivacea</i> mit der Sonde VspI-180 bp-Fraktion.	49
Abbildung 15 Dotplot der MboI-Monomer-Sequenz mit den Klon-Sequenzen aus dem Genbank-Screening.	51
Abbildung 16 Mikroskopische Aufnahme der FISH an Polytäanchromosomen aus <i>P. olivacea</i> mit den Sonden basierend auf Contig_115, Contig_118 und Contig_117.....	53
Abbildung 17 Die chromosomale Organisation des Globin Cluster auf dem Contig 115 von <i>P. olivacea</i>	54
Abbildung 18 Die chromosomale Organisation des Contigs_118 mit den Globin-Genen 118_1 und 118_2.....	55
Abbildung 19 Die Genstruktur des Globin Gens 117.....	55

Abbildung 20 Auswahl an Globin-Genen aus <i>P. olivacea</i> inklusive Annotationen der Gene, mRNAs, Introns, ORFs und Helices der abgeleiteten Proteine.	57
Abbildung 21 Organisation der Histon-Gene H1, H2A, H2B, H3 und H4 aus <i>P. olivacea</i> ..	58
Abbildung 22 Mikroskopische Aufnahme von Polytäanchromosomen aus <i>P. olivacea</i> , in der das Ergebnis der FISH mit einer <i>P. olivacea</i> -Sequenz-spezifischen Histon-Sonde dargestellt ist.	58
Abbildung 23 Analyse der Histon-Gensequenzen und Intergen-Sequenzen aus <i>P. olivacea</i> hinsichtlich ihres AT-Gehalts.	59
Abbildung 24 MSA der drei H3-Varianten aus <i>P. olivacea</i>	61
Abbildung 25 Mikroskopische Aufnahme der FISH an Polytäanchromosomen aus <i>P. olivacea</i> mit der Sonde H3V2.	62
Abbildung 26 Die Genstruktur des <i>Sxl</i> -Gens aus <i>P. olivacea</i>	63
Abbildung 27 Mikroskopische Aufnahme des FISH-Ergebnis mit der Sonde <i>Sxl</i> an Polytäanchromosomen aus <i>P. olivacea</i>	64
Abbildung 28 Die Genstruktur des <i>tra-2</i> -Gens aus <i>P. olivacea</i>	65
Abbildung 29 Mikroskopische Aufnahme der FISH an Polytäanchromosomen aus Speicheldrüsen von <i>P. olivacea</i> mit den Sonden <i>tra-2</i> und <i>l(3)73Ah</i>	67
Abbildung 30 Die Genstruktur von <i>dsx1</i> (<i>Polidsx1</i>) aus <i>P. olivacea</i>	68
Abbildung 31 Mikroskopische Aufnahme des FISH Ergebnis mit der Sonde <i>Polidsx1</i> an Polytäanchromosomen aus <i>P. olivacea</i>	69
Abbildung 32 Gelelektrophoretische Auftrennung der RT-PCR, in welcher das Männchen-spezifische 3'Ende (A) bzw. das Weibchen-spezifische 3'Ende (B) amplifiziert wurde.	70
Abbildung 33 Potenzielle cis-regulatorische Sequenzen im Weibchen-spezifischen Exon des <i>Polidsx1</i> -Gens.	72
Abbildung 34 Mikroskopische Aufnahme nach der FISH an Speicheldrüsenchromosomen aus <i>P. olivacea</i> mit den Sonden <i>fs(1)K10-like</i> und <i>dicer-1</i>	77
Abbildung 35 Ausschnitt der SDR aus <i>C. thummi</i> und der potenziellen SDR aus <i>P. olivacea</i> welcher die syntäne Region (blau) umfasst.	78
Abbildung 36 Mikroskopische Aufnahme des Ergebnisses einer FISH an Polytäanchromosomen aus <i>P. olivacea</i> mit den Sonden CpY und <i>spitz-like</i> (A) und den Sonden <i>rpS5-like</i> und Gb_115 (B).	79

Abbildung 37 Mikroskopische Aufnahmen des Ergebnisses einer FISH an <i>C. thummi</i> Speicheldrüsenchromosomen mit den Sonden 6PGD (A), Cla1.1Y (A, C), <i>dicer-1</i> (B) Gencluster-piger (GP) (B) und <i>tsc1</i> (C).	80
Abbildung 38 Der proximale Abschnitt der potenziellen <i>P. olivacea</i> -SDR im Vergleich zum Gencluster-piger Contig aus <i>C. piger</i>	81
Abbildung 39 AS-Sequenzalignment von BCN92 aus <i>C. thummi</i> und <i>P. olivacea</i> mit Genstruktur von <i>bcn92</i>	82
Abbildung 40 Southern Blot mit restringierter genomischer <i>P. olivacea</i> -DNA aus weiblichen und männlichen Larven hybridisiert mit der Sonde <i>fs(1)K10-like</i>	85
Abbildung 41 Gelelektrophoretische Auftrennung der RT-PCR-Produkte für das Primerpaar msRNA98Cb auf weibliche (A) bzw. männliche (B) cDNA aus <i>P. olivacea</i> -Larven.	88
Abbildung 42 Chromosomenkartierung bei <i>P. olivacea</i>	89
Abbildung 43 Chromosomenkartierung bei <i>C. thummi/piger</i>	90
Abbildung 44 Intronpositionen der Hämoglobin-Gene von Dipteren-Arten.	96
Abbildung 45 MSA der abgeleiteten Proteine des Globin-Gen-Clusters 115 aus <i>P. olivacea</i> mit den <i>D. melanogaster</i> Globin-1 und <i>C. thummi</i> HbI.	98
Abbildung 46 Vorhersage der 3D-Struktur von <i>P. olivacea</i> Gb 115 A mittels RaptorX structure prediction- Tool.	99
Abbildung 47 Alignment der AS-Sequenzen von H1-Proteinen aus <i>C. thummi</i> (Cthu H1) und <i>P. olivacea</i> (Poli H1).	100
Abbildung 48 Ausschnitt der Genstruktur von <i>tra-2</i> aus <i>P. olivacea</i> mit den möglichen Transkript-Varianten und zwei der drei in <i>D. melanogaster</i> beschriebenen <i>tra-2</i> -Isoformen und der abgeleiteten Proteine.	104
Abbildung 49 Nematoceren DSX-MSA.	108
Abbildung 50 MSA der Nematoceren DSX-Paralogen.	109
Abbildung 51 Schema der Chromosomen aus <i>P. olivacea</i> und <i>C. thummi</i>	113
Abbildung 52 Cthu und Poli H1 Nukleotid-Alignment.	144
Abbildung 53 Cthu und Poli H2A Nukleotid-Alignment.	145
Abbildung 54 Cthu und Poli H2B Nukleotid-Alignment.	145
Abbildung 55 Cthu und Poli H3 Nukleotid-Alignment.	145
Abbildung 56 Cthu und Poli H4 Nukleotid-Alignment.	145

Abbildung 57 Mikroskopische Aufnahme von Polytänchromosomen aus <i>C. thummi</i> in der das Ergebnis der FISH mit den Sonden Cla1.1Y (TRITC, rot) und msRNACb (FITC, grün) abgebildet ist.....	147
Abbildung 58 Mikroskopische Aufnahme von Polytänchromosomen aus <i>P. olivacea</i> in der das Ergebnis der FISH mit den Sonden SDR (TRITC, rot) und msRNACb-1 (FITC, grün) abgebildet ist.....	148
Abbildung 59 Mikroskopische Aufnahme von Polytänchromosomen aus <i>P. olivacea</i> in der das Ergebnis der FISH mit den Sonden SDR (FITC, grün) und msRNACb-2 (TRITC, rot) abgebildet ist.....	149
Abbildung 60 MSA des Globin-Clusters 118 aus <i>P. olivacea</i> und den Referenzsequenzen <i>D. melanogaster</i> Globin-1 und <i>C. thummi</i> HbI.....	150
Abbildung 61 MSA des Globin-Clusters 117 aus <i>P. olivacea</i> und den Referenzsequenzen <i>D. melanogaster</i> Globin-1 und <i>C. thummi</i> HbI.....	151

8 Tabellenverzeichnis

Tabelle 1 Die chromosomale Lage der SDR in <i>C. annularius</i> , <i>C. luridus</i> , <i>C. tentans</i> , <i>C. pallidivittatus</i> und der Schwesternspezies <i>C. thummi</i> und <i>C. piger</i>	4
Tabelle 2 Genomgrößen und Anzahl der Gene unterschiedlicher Insekten.....	10
Tabelle 3 Standard PCR-Ansatz mit der GoTaq®DNA-Polymerase.....	37
Tabelle 4 PCR-Programm eines GoTaq®-PCR-Ansatzes.....	38
Tabelle 5 Übersicht der Sequenz-Datensätze aus <i>P. olivacea</i> , welche mittels Illumina HiSeq 2000 generiert wurden.....	46
Tabelle 6 Übersicht des „Master“-Assemblies aus <i>P. olivacea</i> generiert aus den Datensätzen 468, 551, 552 und 729.....	46
Tabelle 7 Bestimmung der Kopienzahl der Monomer-Einheit der Sat-DNA.....	51
Tabelle 8 Übersicht der in <i>P. olivacea</i> identifizierten Globin-Gene.....	56
Tabelle 9 Vergleich der Histon-Gene,-Proteine und der regulatorischen Sequenzen aus <i>C. thummi</i> und <i>P. olivacea</i>	60
Tabelle 10 3'Spleiß-Akzeptorstellen der 3'Region des <i>Polidsx1</i> -Gens.....	71
Tabelle 11 In <i>Polidsx1</i> identifizierte cis-regulatorische Elemente im 664 bp-Weibchenspezifischen Exon.....	72
Tabelle 12 Weitere in <i>Polidsx1</i> identifizierte cis-regulatorische Elemente im 664 bp-Weibchenspezifischen Exon.....	73
Tabelle 13 Gene welche in <i>D. melanogaster</i> in der Geschlechtsbestimmung beteiligt sind.....	75
Tabelle 14 Expressionsanalyse der homologen Gene aus der Geschlechtsbestimmungskaskade von <i>D. melanogaster</i> in <i>P. olivacea</i>	76
Tabelle 15 Sequenzidentität der SDR-Gene und ihrer Produkte zwischen den zwei Arten <i>C. thummi</i> (♀) und <i>P. olivacea</i>	83
Tabelle 16 Überprüfen der Kopienzahl von <i>fs(1)K10-like</i> mittels Abdeckung genomischen DNA-reads im Mapping.....	86
Tabelle 17 Übersicht der Männchen-spezifisch exprimierten Gene des Larven-Transkriptoms aus <i>P. olivacea</i>	87
Tabelle 18 Lokalisation von DNA-Sequenzen (single copy-Gene oder repetitive Elemente) per FISH in <i>C. thummi/piger</i> und <i>P. olivacea</i>	92
Tabelle 19 Vergleich der Histon-Cluster aus drei Dipteren-Spezies.....	101

9 Literaturverzeichnis

Abe, H.; Mita, K.; Yasukochi, Y.; Oshiki, T.; Shimada, T. (2005a): Retrotransposable elements on the W chromosome of the silkworm, *Bombyx mori*. In: *Cytogenetic and genome research* 110 (1-4), S. 144–151.

Abe, H.; Seki, M.; Ohbayashi, F.; Tanaka, N.; Yamashita, J.; Fujii, T. et al. (2005b): Partial deletions of the W chromosome due to reciprocal translocation in the silkworm *Bombyx mori*. In: *Insect molecular biology* 14 (4), S. 339–352.

Abe, Hiroaki; Kanehara, Mariko; Terada, Tomoko; Ohbayashi, Fumi; Shimada, Toru; Kawai, Shinya et al. (1998): Identification of novel random amplified polymorphic DNAs (RAPDs) on the W chromosome of the domesticated silkworm, *Bombyx mori*, and the wild silkworm, *B. mandarina*, and their retrotransposable element-related nucleotide sequences. In: *Genes Genet. Syst.* 73 (4), S. 243–254. DOI: 10.1266/ggs.73.243.

Acton, A. B. (1957): Sex-chromosome inversions in *Chironomus*. In: *American Naturalist*, S. 57–59.

Acton, A. B. (1959): A study of the differences between widely separated populations of *Chironomus* (= *Tendipes*) tentans (Diptera). In: *Proceedings of the Royal Society of London B: Biological Sciences* 151 (943), S. 277–296.

Adams, M. D.; Celniker, S. E.; Holt, R. A.; Evans, C. A.; Gocayne, J. D.; Amanatides, P. G. et al. (2000): The genome sequence of *Drosophila melanogaster*. In: *Science (New York, N.Y.)* 287 (5461), S. 2185–2195.

Ahmad, Kami; Henikoff, Steven (2002a): Histone H3 variants specify modes of chromatin assembly. In: *Proceedings of the National Academy of Sciences* 99 (suppl 4), S. 16477–16484.

Ahmad, Kami; Henikoff, Steven (2002b): The histone variant H3.3 marks active chromatin by replication-independent nucleosome assembly. In: *Molecular cell* 9 (6), S. 1191–1200.

Altschul, Stephen F.; Gish, Warren; Miller, Webb; Myers, Eugene W.; Lipman, David J. (1990): Basic local alignment search tool. In: *Journal of molecular biology* 215 (3), S. 403–410. DOI: 10.1016/S0022-2836(05)80360-2.

An, Wenqian; Cho, Sayeon; Ishii, Haruhiko; Wensink, Pieter C. (1996): Sex-specific and non-sex-specific oligomerization domains in both of the doublesex transcription factors from *Drosophila melanogaster*. In: *Molecular and cellular biology* 16 (6), S. 3106–3111.

Antoine, Marianne; Erbil, Cebraïl; Münch, Eva; Schnell, Sigrid; Niessing, Jürgen (1987): Genomic organization and primary structure of five homologous pairs of intron-less genes encoding secretory globins from the insect *Chironomus thummi thummi*. In: *Gene* 56 (1), S. 41–51.

Antoine, Marianne; Niessing, Jürgen (1984): Intron-less globin genes in the insect *Chironomus thummi thummi*.

Arensburger, Peter; Megy, Karine; Waterhouse, Robert M.; Abrudan, Jenica; Amedeo, Paolo; Antelo, Beatriz et al. (2010): Sequencing of *Culex quinquefasciatus* establishes a platform for mosquito comparative genomics. In: *Science (New York, N.Y.)* 330 (6000), S. 86–88. DOI: 10.1126/science.1191864.

Armitage, Patrick D.; Pinder, L. C.; Cranston, Peter (2012): *The Chironomidae: biology and ecology of non-biting midges*: Springer Science & Business Media.

Atchley, William R.; Martin, Jon (1971): A morphometric analysis of differential sexual dimorphism in larvae of *Chironomus* (Diptera). In: *The Canadian Entomologist* 103 (03), S. 319–327.

Baker, B. S.; Wolfner, M. F. (1988): A molecular analysis of doublesex, a bifunctional gene that controls both male and female sexual differentiation in *Drosophila melanogaster*. In: *Genes & development* 2 (4), S. 477–489.

Balbani, E. G. (1881): Sur la structure du noyau des cellules salivaires chez les larves de *Chironomus*. In: *Zool. Anz* 4 (637-641), S. 662–667.

Baldwin, Albert S.; Bergtrom, Gerald; Haralson, Michael A. (1981): Cell-free synthesis of *Chironomus thummi* (Diptera) globins. In: *Archives of biochemistry and biophysics* 211 (1), S. 390–395.

Bashford, Donald; Chothia, Cyrus; Lesk, Arthur M. (1987): Determinants of a protein fold. In: *Journal of molecular biology* 196 (1), S. 199–216. DOI: 10.1016/0022-2836(87)90521-3.

- Bauer, Hans; Beermann, Wolfgang (1952): Der Chromosomencyclus der Orthoclaadien (Nematocera, Diptera). In: *Zeitschrift für Naturforschung B* 7 (9-10), S. 557–563.
- Beermann, W. (1955): Geschlechtsbestimmung und Evolution der genetischen Y-Chromosomen bei Chironomus. In: *Biol. Zbl* 74, S. 525–544.
- Beermann, W. (1962): Riesenchromosomen. Protoplasmatologia Handbuch der Protoplasmaforschung. Band IVd: Springer-Verlag, Vienna.
- Bergtrom, Gerald; Laufer, Hans; Rogers, Ralph (1976): Fat body: a site of hemoglobin synthesis in *Chironomus thummi* (diptera). In: *The Journal of cell biology* 69 (2), S. 264–274.
- Blanchetot, A.; Wilson, V.; Wood, D.; Jeffreys, A. J. (1983): The seal myoglobin gene: an unusually long globin gene.
- Blanchetot, Alain; Price, Melanie; Jeffreys, Alec J. (1986): The mouse myoglobin gene. In: *Eur J Biochem* 159 (3), S. 469–474.
- Blumensatt, Marcel (2009): Klonierung und Analyse der geschlechtsbestimmenden Region von *Chironomus piger* durch in vitro Amplifikation. Diplomarbeit. Johannes Gutenberg-Universität, Mainz. IMSB.
- Böhm, Lothar; Crane-Robinson, Colyn (1984): Proteases as structural probes for chromatin: the domain structure of histones. In: *Bioscience reports* 4 (5), S. 365–386.
- Braun, Volkmar; Crichton, Robert R.; BRAUNITZER, Gerhard (1968): Hämoglobine, XV. Über monomere und dimerie Insektenhämoglobine (*Chironomus thummi*). In: *Hoppe-Seyler's Zeitschrift für physiologische Chemie* 349 (1), S. 197–210.
- Bull, James J.; others (1983): Evolution of sex determining mechanisms: The Benjamin/Cummings Publishing Company, Inc.
- Burmester, Thorsten; Hankeln, Thomas (1999): A globin gene of *Drosophila melanogaster*. In: *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology* 124, S. S93.
- Burmester, Thorsten; Hankeln, Thomas (2007): The respiratory proteins of insects. In: *Journal of insect physiology* 53 (4), S. 285–294. DOI: 10.1016/j.jinsphys.2006.12.006.

Burmester, Thorsten; Klawitter, Sabine; Hankeln, Thomas (2007): Characterization of two globin genes from the malaria mosquito *Anopheles gambiae*: divergent origin of nematoceran haemoglobins. In: *Insect molecular biology* 16 (2), S. 133–142.

Burmester, Thorsten; Storf, Joachim; Hasenjäger, Anne; Klawitter, Sabine; Hankeln, Thomas (2006): The hemoglobin genes of *Drosophila*. In: *The FEBS journal* 273 (3), S. 468–480. DOI: 10.1111/j.1742-4658.2005.05073.x.

Burtis, Kenneth C.; Baker, Bruce S. (1989): *Drosophila* doublesex gene controls somatic sexual differentiation by producing alternatively spliced mRNAs encoding related sex-specific polypeptides. In: *Cell* 56 (6), S. 997–1010.

Busslinger, Meinrad; Portmann, Rudolf; Birnstiel, Max L. (1979): A regulatory sequence near the 3' end of sea urchin histone genes. In: *Nucleic acids research* 6 (9), S. 2997–3008.

Cambridge Univ Press (2007): *Annales de Limnologie-International Journal of Limnology*.

Charlesworth, B.; Sniegowski, P.; Stephan, W. (1994): The evolutionary dynamics of repetitive DNA in eukaryotes. In: *Nature* 371 (6494), S. 215–220. DOI: 10.1038/371215a0.

Childs, Geoffrey; Maxson, Rob; Cohn, Ronald H.; Kedes, Larry (1981): Orphans: Dispersed genetic elements derived from tandem repetitive genes of eucaryotes. In: *Cell* 23 (3), S. 651–663. DOI: 10.1016/0092-8674(81)90428-1.

Cho, S.; Huang, Z. Y.; Zhang, J. (2007): Sex-Specific Splicing of the Honeybee doublesex Gene Reveals 300 Million Years of Evolution at the Bottom of the Insect Sex-Determination Pathway. In: *Genetics* 177 (3), S. 1733–1741. DOI: 10.1534/genetics.107.078980.

Chomczynski, Piotr; Sacchi, Nicoletta (1987): Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. In: *Analytical Biochemistry* 162 (1), S. 156–159. DOI: 10.1016/0003-2697(87)90021-2.

Cranston, Peter S.; Hardy, Nate B.; Morse, Geoffrey E. (2012): A dated molecular phylogeny for the Chironomidae (Diptera). In: *Systematic Entomology* 37 (1), S. 172–188. DOI: 10.1111/j.1365-3113.2011.00603.x.

Davis, Claytus A.; Wyatt, Gerard R. (1989): Distribution and sequence homogeneity of an abundant satellite DNA in the beetle, *Tenebrio Molitor*. In: *Nucleic acids research* 17 (14), S. 5579–5586. DOI: 10.1093/nar/17.14.5579.

Denhardt, David T. (1966): A membrane-filter technique for the detection of complementary DNA. In: *Biochemical and Biophysical Research Communications* 23 (5), S. 641–646. DOI: 10.1016/0006-291X(66)90447-5.

Devlin, Robert H.; Nagahama, Yoshitaka (2002): Sex determination and sex differentiation in fish: an overview of genetic, physiological, and environmental influences. In: *Aquaculture* 208 (3), S. 191–364.

Dewilde, S. (1999): Nonvertebrate globins and globin genes: structure, function and evolution: Universiteit Antwerpen. Faculteit Wetenschappen (Proefschriften UA-WET). Online verfügbar unter <https://books.google.de/books?id=IM32rQEACAAJ>.

Dübendorfer, A.; Hilfiker-Kleiner, D.; Nöthiger, R. (1992): Sex determination mechanisms in dipteran insects: the case of *Musca domestica*. In: *Dev. Biol* 3, S. 349–356.

Elgin, Sarah C. R.; Weintraub, Harold (1975): Chromosomal proteins and chromatin structure. In: *Annual review of biochemistry* 44 (1), S. 725–774.

Engel, James Douglas; Sugarman, Barry J.; Dodgson, Jerry B. (1982): A chicken histone H3 gene contains intervening sequences. In: *Nature* 297 (5865), S. 434–436. DOI: 10.1038/297434a0.

Erdman, Scott E.; Chen, Hui-Ju; Burtis, Kenneth C. (1996): Functional and genetic characterization of the oligomerization and DNA binding properties of the *Drosophila* doublesex proteins. In: *Genetics* 144 (4), S. 1639–1652.

Feinberg, Andrew P.; Vogelstein, Bert (1983): A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. In: *Analytical Biochemistry* 132 (1), S. 6–13. DOI: 10.1016/0003-2697(83)90418-9.

Felger, Ingrid; Pinsker, W. (1987): Histone gene transposition in the phylogeny of the *Drosophila obscura* group. In: *Journal of Zoological Systematics and Evolutionary Research* 25 (2), S. 127–140. DOI: 10.1111/j.1439-0469.1987.tb00596.x.

Finley, K. D.; Taylor, B. J.; Milstein, M.; McKeown, M. (1997): dissatisfaction, a gene involved in sex-specific behavior and neural development of *Drosophila melanogaster*. In: *Proceedings of the National Academy of Sciences of the United States of America* 94 (3), S. 913–918.

Fletcher, Hugh; Hickey, Ivor (2013): Genetik. Für Biologen, Biochemiker, Pharmazeuten und Mediziner. 1. Aufl. Weinheim: Wiley-VCH (Verdammt clever!).

Frank, Tanja (2001): Nachweis und Lokalisation von möglicherweise geschlechtsbestimmenden Genen in *Chironomus thummi* und *Chironomus piger*. Diplomarbeit. Johannes Gutenberg-Universität, Mainz. IMSB.

Friedl, Heike (2001): Die molekulare Evolution der Hämoglobin-Genfamilie in *Chironomus tentans*. Dissertation. Johannes Gutenberg-Universität, Mainz. IMSB.

Funke, Elisabeth (2011): Charakterisierung und Analyse der geschlechtsbestimmenden Region von *Chironomus riparius*. Disstertation. Johannes Gutenberg-Universität, Mainz. IMSB.

Gō, Mitiko (1981): Correlation of DNA exonic regions with protein structural units in haemoglobin.

Gempe, Tanja; Beye, Martin (2011): Function and evolution of sex determination mechanisms, genes and pathways in insects. In: *Bioessays* 33 (1), S. 52–60.

Gerbracht, Jennifer (2014): Molekulare Analyse des *transformer-2*-Gens in *Chironomus*. Bachelor-Arbeit. Johannes Gutenberg-Universität, Mainz. IMSB.

Geuverink, E.; Beukeboom, L. W. (2014): Phylogenetic distribution and evolutionary dynamics of the sex determination genes doublesex and transformer in insects. In: *Sexual development : genetics, molecular biology, evolution, endocrinology, embryology, and pathology of sex determination and differentiation* 8 (1-3), S. 38–49. DOI: 10.1159/000357056.

Goodman, Morris; Braunitzer, Gerhard; Kleinschmidt, Traute; Aschauer, Heinz (1983): The Analysis of a Protein-Polymorphism. Evolution of Monomeric and Homodimeric Haemoglobins (Erythrocrucorins) of *Chironomus thummi thummi* (Insecta, Diptera). In: *Hoppe-Seyler's Zeitschrift für physiologische Chemie* 364 (1), S. 205–218. DOI: 10.1515/bchm2.1983.364.1.205.

Goodman, Morris; Pedwaydon, Janet; Czelusniak, John; Suzuki, Tomohiko; Gotoh, Toshio; Moens, Luc et al. (1988): An evolutionary tree for invertebrate globin sequences. In: *Journal of molecular evolution* 27 (3), S. 236–249.

- Graves, Jennifer A Marshall (2008): Weird animal genomes and the evolution of vertebrate sex and sex chromosomes. In: *Annu. Rev. Genet.* 42, S. 565–586.
- Gruhl, Mary; Kao, Wen-Yen; Bergtrom, Gerald (1997): Evolution of orthologous intronless and intron-bearing globin genes in two insect species. In: *Journal of molecular evolution* 45 (5), S. 499–508.
- Grunstein, M.; Hogness, D. S. (1975): Colony hybridization: a method for the isolation of cloned DNAs that contain a specific gene. In: *Proceedings of the National Academy of Sciences of the United States of America* 72 (10), S. 3961–3965.
- Hägele, K. (1970): DNS-Replikationsmuster der Speicheldrüsen-Chromosomen von Chironomiden. In: *Chromosoma* 31 (1), S. 91–138.
- Hägele, Klaus (1985): Identification of a polytene chromosome band containing a male sex determiner of *Chironomus thummi thummi*. In: *Chromosoma* 91 (3-4), S. 167–171.
- Hammann, Jens (2013): Struktur und Expression des dsx-Gens bei Chironomus. Johannes Gutenberg-Universität, Mainz. IMSB.
- Handa, N.; Nureki, O.; Kurimoto, K.; Kim, I.; Sakamoto, H.; Shimura, Y. et al. (1999): Structural basis for recognition of the tra mRNA precursor by the Sex-lethal protein. In: *Nature* 398 (6728), S. 579–585. DOI: 10.1038/19242.
- Hankeln, T.; Keyl, H. G.; Ross, R.; Schmidt, E. R. (1993): Evolution of histone gene loci in chironomid midges. In: *Genome / National Research Council Canada = Génome / Conseil national de recherches Canada* 36 (5), S. 852–862.
- Hankeln, T.; Luther, C.; Rozynek, P.; Schmidt, E. R. (1991): Complete nucleotide sequence of a hemoglobin gene cluster from the midge *Chironomus thummi piger*. In: *Structure and Function of Invertebrate Oxygen Carriers*: Springer, S. 287–296.
- Hankeln, T.; Schmidt, E. R. (1991): The organization, localization and nucleotide sequence of the histone genes of the midge *Chironomus thummi*. In: *Chromosoma* 101 (1), S. 25–31.
- Hankeln, Thomas; Friedl, Heike; Ebersberger, Ingo; Martin, Jon; Schmidt, Erwin R. (1997): A variable intron distribution in globin genes of *Chironomus*: evidence for recent intron gain. In: *Gene* 205 (1), S. 151–160.

Hankeln, Thomas; Rozynek, Peter; Schmidt, Erwin R. (1988): The nucleotide sequence and in situ localization of a gene for a dimeric haemoglobin from the midge *Chironomus thummi piger*. In: *Gene* 64 (2), S. 297–304.

Hardison, Ross C. (1996): A brief history of hemoglobins: plant, animal, protist, and bacteria. In: *Proceedings of the National Academy of Sciences* 93 (12), S. 5675–5679.

Hentschel, Christopher C.; Birnstiel, Max L. (1981): The organization and expression of histone gene families. In: *Cell* 25 (2), S. 301–313. DOI: 10.1016/0092-8674(81)90048-9.

Hildreth, Philip E. (1965): Doublesex, a Recessive Gene That Transforms Both Males and Females of *Drosophila* into Intersexes. In: *Genetics* 51 (4), S. 659.

Holt, Robert A.; Subramanian, G. Mani; Halpern, Aaron; Sutton, Granger G.; Charlab, Rosane; Nusskern, Deborah R. et al. (2002): The genome sequence of the malaria mosquito *Anopheles gambiae*. In: *Science (New York, N.Y.)* 298 (5591), S. 129–149. DOI: 10.1126/science.1076181.

Hong, Chang-Soo; Park, Byung-Yong; Saint-Jeannet, Jean-Pierre (2007): The function of *Dmrt* genes in vertebrate development: it is not just about sex. In: *Developmental biology* 310 (1), S. 1–9.

Hoyer-Fender, S.; Grossbach, U. (1988): Histone H1 heterogeneity in the midge, *Chironomus thummi*. Structural comparison of the H1 variants in an organism where their intrachromosomal localization is possible. In: *European journal of biochemistry / FEBS* 176 (1), S. 139–152.

Huang, H.; Crews, S.; Hood, L. (1980): An immunoglobulin VH pseudogene. In: *Journal of molecular and applied genetics* 1 (2), S. 93–101.

Hughes, Thomas A. (2006): Regulation of gene expression by alternative untranslated regions. In: *Trends Genet.* 22 (3), S. 119–122. DOI: 10.1016/j.tig.2006.01.001.

Inoue, Kunio; Hoshijima, Kazuyuki; Higuchi, Ikuko; Sakamoto, Hiroshi; Shimura, Yoshiro (1992): Binding of the *Drosophila* transformer and transformer-2 proteins to the regulatory elements of doublesex primary transcript for sex-specific RNA processing. In: *Proceedings of the National Academy of Sciences* 89 (17), S. 8092–8096.

- Irminger-Finger, Irmgard; Nöthiger, Rolf (1995): The *Drosophila melanogaster* gene lethal (3) 73Ah encodes a ring finger protein homologous to the oncoproteins MEL-18 and BMI-1. In: *Gene* 163 (2), S. 203–208.
- Isenberg, I. (1979): Histones. In: *Annual review of biochemistry* 48, S. 159–191. DOI: 10.1146/annurev.bi.48.070179.001111.
- Israelewski, Norbert (1983): Structure and function of an AT-rich, interspersed repetitive sequence from *Chironomus thummi*. Solenoidal DNA, 142 bp palindrome-frame and homologies with the sequence for site-specific recombination of bacterial transposons. In: *Nucleic acids research* 11 (20), S. 6985–6996. DOI: 10.1093/nar/11.20.6985.
- Jacq, C.; Miller, JR; Brownlee, G. G. (1977): A pseudogene structure in 5S DNA of *Xenopus laevis*. In: *Cell* 12 (1), S. 109–120.
- Jensen, Erik Østergaard; Paludan, Kirsten; Hyldig-Nielsen, J. J.; Jørgensen, Poul; Marcker, K. A. (1981): The structure of a chromosomal leghaemoglobin gene from soybean.
- Kao, Wen-Yen; Trewitt, Patrick M.; Bergtrom, Gerald (1994): Intron-containing globin genes in the insect *Chironomus thummi*. In: *Journal of molecular evolution* 38 (3), S. 241–249.
- Kelley, Joanna L.; Peyton, Justin T.; Fiston-Lavier, Anna-Sophie; Teets, Nicholas M.; Yee, Muh-Ching; Johnston, J. Spencer et al. (2014): Compact genome of the Antarctic midge is likely an adaptation to an extreme environment. In: *Nat Comms* 5. DOI: 10.1038/ncomms5611.
- Keyl, H. G.; Keyl, Ilse (1959): Die cytologische Diagnostik der Chironomiden. I. Bestimmungstabelle für die Gattung *Chironomus* auf Grund der Speicheldrüsen-Chromosomen. In: *Arch. Hydrobiol* 56 (1/2), S. 43–57.
- Keyl, H-G (1961): Chromosomenevolution bei *Chironomus*. In: *Chromosoma* 12 (1), S. 26–47.
- Keyl, H-G (1962): Chromosomenevolution bei *Chironomus*. In: *Chromosoma* 13 (4), S. 464–514.
- Keyl, H-G (1965): A demonstrable local and geometric increase in the chromosomal DNA of *Chironomus*. In: *Experientia* 21 (4), S. 191–193.
- Kidwell, Margaret G. (2002): Transposable elements and the evolution of genome size in eukaryotes. In: *Genetica* 115 (1), S. 49–63.

- Kieffer, J. J. (1911): Nouvelles descriptions de chironomides obtenus d'éclosion. In: *Bulletin De La Société D'histoire Naturelle De Metz* 27, S. 1–60.
- King, R. C. (1975): Handbook of Genetics: Invertebrates of genetic interest: Plenum Press (Handbook of Genetics). Online verfügbar unter <https://books.google.de/books?id=2yO8MAEACAAJ>.
- Kit, Saul (1961): Equilibrium sedimentation in density gradients of DNA preparations from animal tissues. In: *Journal of molecular biology* 3 (6), S. 711 - IN2.
- Klug, William S.; Cummings, Michael R.; Spencer, Charlotte A.; Thomm, Michael (2007): Genetik. 8., aktualisierte Aufl. München: Pearson Studium (Biologie).
- Kraemer, C.: Molekulare Analyse der geschlechtsbestimmenden Region von Chironomiden. Ph D thesis. Johannes Gutenberg-University Mainz, Germany, Mainz. IMSB.
- Kraemer, Christiane; Schmidt, Erwin R. (1993): The sex determining region of *Chironomus thummi* is associated with highly repetitive DNA and transposable elements. In: *Chromosoma* 102 (8), S. 553–562.
- Kremer, Hannie; Hennig, Wolfgang (1990): Isolation and characterization of a *Drosophila hydei* histone DNA repeat unit. In: *Nucleic acids research* 18 (6), S. 1573–1586. DOI: 10.1093/nar/18.6.1573.
- Kuhn, G C S (2015): 'Satellite DNA transcripts have diverse biological roles in *Drosophila*'. In: *Heredity* 115 (1), S. 1–2. DOI: 10.1038/hdy.2015.12.
- Kutsenko, Alexey; Svensson, Thomas; Nystedt, Björn; Lundeberg, Joakim; Björk, Petra; Sonnhammer, Erik et al. (2014): The *Chironomus tentans* genome sequence and the organization of the Balbiani ring genes. In: *BMC Genomics* 15 (1), S. 819. DOI: 10.1186/1471-2164-15-819.
- Lagos, D.; Koukidou, M.; Savakis, C.; Komitopoulou, K. (2007): The transformer gene in *Bactrocera oleae*: the genetic switch that determines its sex fate. In: *Insect molecular biology* 16 (2), S. 221–230.
- Leathers, Adelbert Llewellyn (1922): Ecological study of aquatic midges and some related insects with special reference to feeding habits: Govt. Print. Off.

- Lifton, R. P.; Goldberg, M. L.; Karp, R. W.; Hogness, D. S. (1978): The organization of the histone genes in *Drosophila melanogaster*: functional and evolutionary implications. In: *Cold Spring Harbor symposia on quantitative biology* 42 Pt 2, S. 1047–1051.
- Lynch, Kristen W.; Maniatis, Tom (1995): Synergistic interactions between two distinct elements of a regulated splicing enhancer. In: *Genes & development* 9 (3), S. 284–293.
- Mainx, F. (1966): Die Geschlechtsbestimmung bei *Megaselia scalaris* Loew (Phoridae). In: *Zeitschrift für Vererbungslehre* 98 (1), S. 49–60.
- Martin, J. (2010): A phylogenetic analysis of the change in genomic location of the sex determining region in the genus *Chironomus* (Diptera: Chironomidae). In: *VOGiS Herald* 14, S. 62–69.
- Martin, J.; Lee, B. T.O.; Conner, E. (1978): Apparent incipient speciation in the midge *Chironomus oppositus* Walker (Diptera: Chironomidae). In: *Australian Journal of Zoology* 26 (2), S. 323–329.
- Martin, Jon (1974): A review of the genus *Chironomus* (Diptera, Chironomidae). In: *Chromosoma* 45 (1), S. 91–98.
- Martin, Jon; Kuvangkadilok, Chaliow; Peart, Dianne H.; Lee, Barry T O (1980): Multiple sex determining regions in a group of related *Chironomus* species (Diptera:Chironomidae). In: *Heredity* 44 (3), S. 367–382. DOI: 10.1038/hdy.1980.34.
- Martin, Jon; Lee, B. T.O. (1984): A phylogenetic study of sex determiner location in a group of Australasian *Chironomus* species (Diptera, Chironomidae). In: *Chromosoma* 90 (3), S. 190–197.
- Martin, Jon; Wülker, W. (1971): Inversion polymorphism in *Chironomus staegeri* Lundbeck. In: *Canadian Journal of Genetics and Cytology* 13 (2), S. 306–321.
- Matsubara, Kazumi; O'Meally, Denis; Azad, Bhumika; Georges, Arthur; Sarre, Stephen D.; Graves, Jennifer A Marshall et al. (2015): Amplification of microsatellite repeat motifs is associated with the evolutionary differentiation and heterochromatinization of sex chromosomes in *Sauropsida*. In: *Chromosoma*. DOI: 10.1007/s00412-015-0531-z.
- Matsuo, Y.; Yamazaki, T. (1989): tRNA derived insertion element in histone gene repeating unit of *Drosophila melanogaster*. In: *Nucleic acids research* 17 (1), S. 225–238.

- Mattox, William; McGuffin, M. Elaine; Baker, Bruce S. (1996): A negative feedback mechanism revealed by functional analysis of the alternative isoforms of the *Drosophila* splicing regulator transformer-2. In: *Genetics* 143 (1), S. 303.
- Maxson, R.; Cohn, R.; Kedes, L.; Mohun, T. (1983): Expression and organization of histone genes. In: *Annu. Rev. Genet.* 17, S. 239–277. DOI: 10.1146/annurev.ge.17.120183.001323.
- McKeown, Michael; Belote, John M.; Baker, Bruce S. (1987): A molecular analysis of transformer, a gene in *Drosophila melanogaster* that controls female sexual differentiation. In: *Cell* 48 (3), S. 489–499.
- Meigen, Johann Wilhelm (1804): Klassifikation und beschreibung der europäischen zweiflügligen insekten:(Diptera Linn.): Karl Reichard (1).
- Metzenberg, A. B.; Wurzer, G.; Huisman, T. H.; Smithies, O. (1991): Homology requirements for unequal crossing over in humans. In: *Genetics* 128 (1), S. 143–161.
- Michailova, P. (1977): Karyotaxonomische Charakteristik der *Prodiamesa olivacea* Meigen und *Prodiamesa bureschi* sp. n.(Diptera, Chironomidae). In: *Zoologische Beiträge*.
- Miller, Jason R.; Koren, Sergey; Sutton, Granger (2010): Assembly algorithms for next-generation sequencing data. In: *Genomics* 95 (6), S. 315–327. DOI: 10.1016/j.ygeno.2010.03.001.
- Moens, L.; Vanfleteren, J.; Van de Peer, Y; Peeters, K.; Kapp, O.; Czeluzniak, J. et al. (1996): Globins in nonvertebrate species: dispersal by horizontal gene transfer and evolution of the structure-function relationships. In: *Molecular Biology and Evolution* 13 (2), S. 324–333.
- Mount, Stephen M.; Burks, Christian; Herts, Gerald; Stormo, Gary D.; White, Owen; Fields, Chris (1992): Splicing signals in *Drosophila*: intron size, information content, and consensus sequences. In: *Nucleic acids research* 20 (16), S. 4255–4262.
- Müller-Holtkamp, Felix (1995): The Sex-lethal gene homologue in *Chrysomya rufifacies* is highly conserved in sequence and exon-intron organization. In: *Journal of molecular evolution* 41 (4), S. 467–477.
- Mullis, K.; Faloona, F.; Scharf, S.; Saiki, R.; Horn, G.; Erlich, H. (1986): Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. In: *Cold Spring Harbor symposia on quantitative biology* 51 Pt 1, S. 263–273.

- Nene, Vishvanath; Wortman, Jennifer R.; Lawson, Daniel; Haas, Brian; Kodira, Chinnappa; Tu, Zhijian Jake et al. (2007): Genome sequence of *Aedes aegypti*, a major arbovirus vector. In: *Science (New York, N.Y.)* 316 (5832), S. 1718–1723. DOI: 10.1126/science.1138878.
- Nishioka, Y.; Leder, A.; Leder, P. (1980): Unusual alpha-globin-like gene that has cleanly lost both globin intervening sequences. In: *Proceedings of the National Academy of Sciences* 77 (5), S. 2806–2809.
- Nowak, Carsten; Hankeln, Thomas; Schmidt, Erwin R.; Schwenk, Klaus (2006): Development and localization of microsatellite markers for the sibling species *Chironomus riparius* and *Chironomus piger* (Diptera: Chironomidae). In: *Molecular Ecology Notes* 6 (3), S. 915–917.
- Ohno, Susumu (2013): Evolution by gene duplication: Springer Science & Business Media.
- Oliveira, Deodoro CSG; Werren, John H.; Verhulst, Eveline C.; Giebel, Jonathan D.; Kamping, Albert; Beukeboom, Leo W.; van de Zande, Louis (2009): Identification and characterization of the doublesex gene of *Nasonia*. In: *Insect molecular biology* 18 (3), S. 315–324.
- Oliver (1971): Life history of the Chironomidae. In: *Annual review of entomology* 16 (1), S. 211–230.
- Palomeque, T.; Lorite, P. (2008): Satellite DNA in insects: a review. In: *Heredity* 100 (6), S. 564–573.
- Pane, Attilio; Salvemini, Marco; Delli Bovi, Pasquale; Polito, Catello; Saccone, Giuseppe (2002): The transformer gene in *Ceratitis capitata* provides a genetic basis for selecting and remembering the sexual fate. In: *Development (Cambridge, England)* 129 (15), S. 3715–3725.
- Pardue, M. L.; Kedes, L. H.; Weinberg, E. S.; Birnstiel, M. L. (1977): Localization of sequences coding for histone messenger RNA in the chromosomes of *Drosophila melanogaster*. In: *Chromosoma* 63 (2), S. 135–151.
- Petrova, N. A.; Zhirov, S. V. (2014): Characteristics of the Karyotypes of three subfamilies of chironomids (Diptera, Chironomidae: Tanypodinae, Diamesinae, Prodiamesinae) of the world fauna. In: *Entmol. Rev.* 94 (2), S. 157–165. DOI: 10.1134/S001387381402002X.
- Pinder, L. C.V. (1986): Biology of freshwater Chironomidae. In: *Annual review of entomology* 31 (1), S. 1–23.

- Price, Dana C.; Egizi, Andrea; Fonseca, Dina M. (2015): Characterization of the doublesex gene within the *Culex pipiens* complex suggests regulatory plasticity at the base of the mosquito sex determination cascade. In: *BMC evolutionary biology* 15 (1), S. 108. DOI: 10.1186/s12862-015-0386-1.
- Ptitsyn, O. B.; Ting, K. L. (1999): Non-functional conserved residues in globins and their possible role as a folding nucleus. In: *Journal of molecular biology* 291 (3), S. 671–682. DOI: 10.1006/jmbi.1999.2920.
- Qi, J.; Su, S.; Mattox, W. (2006): The doublesex Splicing Enhancer Components Tra2 and Rbp1 Also Repress Splicing through an Intronic Silencer. In: *Molecular and cellular biology* 27 (2), S. 699–708. DOI: 10.1128/MCB.01572-06.
- Raymond, Christopher S.; Shamu, Caroline E.; Shen, Michael M.; Seifert, Kelly J.; Hirsch, Betsy; Hodgkin, Jonathan; Zarkower, David (1998): Evidence for evolutionary conservation of sex-determining genes. In: *Nature* 391 (6668), S. 691–695.
- Rempel, J. G.; Naylor, J. M.; Rothfels, K.; Ottonen, B. (1962): The sex chromosome constitution of chironomid intersexes parasitized by nematodes. In: *Canadian Journal of Genetics and Cytology* 4 (1), S. 92–96.
- Rosin, S.; Fischer, J. (1972): Polymorphismus des Realisators für männliches Geschlecht bei *Chironomus*. In: *Revue suisse de zoologie*.
- Rozynek, P.; Broecker, M.; Hankeln, T.; Schmidt, E. R. (1991): The primary structure of several hemoglobin genes from the genome of *Chironomus tentans*. In: *Structure and function of invertebrate oxygen carriers*: Springer, S. 297–303.
- Rozynek, Peter; Hankeln, Thomas; SCHMIDT, Erwin Robert (1989): Structure of a hemoglobin gene cluster and nucleotide sequence of three hemoglobin genes from the midge *Chironomus thummi piger* (Diptera, Insecta). In: *Biological chemistry Hoppe-Seyler* 370 (1), S. 533–542.
- Ruiz, María; Perondini, Andre L. P.; Selivon, Denise; Polito, Catello; Saccone, Giuseppe; Sanchez, Lucas (2007): The gene transformer of *Anastrepha* fruit flies (Diptera, tephritidae) and its evolution in insects. In: *PLoS One* 2 (11), S. e1239.

- Ryner, Lisa C.; Goodwin, Stephen F.; Castrillon, Diego H.; Anand, Anuranjan; Vilella, Adriana; Baker, Bruce S. et al. (1996): Control of Male Sexual Behavior and Sexual Orientation in *Drosophila* by the fruitless Gene. In: *Cell* 87 (6), S. 1079–1089. DOI: 10.1016/S0092-8674(00)81802-4.
- Saccone, G.; Peluso, I.; Artiaco, D.; Giordano, E.; Bopp, D.; Polito, L. C. (1998): The *Ceratitis capitata* homologue of the *Drosophila* sex-determining gene *sex-lethal* is structurally conserved, but not sex-specifically regulated. In: *Development (Cambridge, England)* 125 (8), S. 1495–1500.
- Salvemini, Marco; Mauro, Umberto; Lombardo, Fabrizio; Milano, Andreina; Zazzaro, Vincenzo; Arcà, Bruno et al. (2011): Genomic organization and splicing evolution of the *doublesex* gene, a *Drosophila* regulator of sexual differentiation, in the dengue and yellow fever mosquito *Aedes aegypti*. In: *BMC evolutionary biology* 11, S. 41. DOI: 10.1186/1471-2148-11-41.
- Sánchez R. Lucas; others (2008): Sex-determining mechanisms in insects. In: *International journal of developmental biology* 52 (7), S. 837–856.
- Scali, Christina; Catteruccia, Flaminia; Li, Qiuxiang; Crisanti, Andrea (2005): Identification of sex-specific transcripts of the *Anopheles gambiae* *doublesex* gene. In: *The Journal of experimental biology* 208 (Pt 19), S. 3701–3709. DOI: 10.1242/jeb.01819.
- Schaefer, Jutta; Schmidt, Erwin R. (1981): Different repetition frequencies of a 120 base-pair DNA-element and its arrangement in *Chironomus thummi thummi* and *Chironomus thummi piger*. In: *Chromosoma* 84 (1), S. 61–66.
- Schmid, P. E. (1992): Population dynamics and resource utilization by larval Chironomidae (Diptera) in a backwater area of the River Danube. In: *Freshwater Biology* (28), S. 111–127.
- Schmidt, E. R. (1980): Two AT-rich satellite DNAs in the chironomid *Glyptotendipes barbipes* (Staeger): isolation and localization in polytene chromosomes of *G. barbipes* and *Chironomus thummi*. In: *Chromosoma* 79 (3), S. 315–328.
- 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. In: *Journal of molecular biology* 178 (1), S. 1–15.

Schmidt, E. R.; Godwin, E. A. (1983): The nucleotide sequence of an unusual non-transcribed spacer and its ancestor in the rDNA in *Chironomus thummi*. In: *The EMBO journal* 2 (7), S. 1177.

Schmidt, E. R.; Keyl, H.-G.; Hankeln, T. (1988): In situ localization of two haemoglobin gene clusters in the chromosomes of 13 species of *Chironomus*. In: *Chromosoma* 96 (5), S. 353–359. DOI: 10.1007/BF00330701.

Schmidt, E. R.; Vistorin, Gerda; Keyl, H.-G. (1980): An AT-rich DNA component in the genomes of *Chironomus thummi thummi* and *Chironomus thummi piger*. In: *Chromosoma* 76 (1), S. 35–45. DOI: 10.1007/BF00292224.

Schmidt, Erwin R. (1981): The development of a 120 basepair repetitive DNA sequence in *Chironomus thummi* is correlated to the duplication of defined chromosomal segments. In: *FEBS letters* 129 (1), S. 21–24.

Schütt, C.; Nöthiger, R. (2000): Structure, function and evolution of sex-determining systems in Dipteran insects. In: *Development (Cambridge, England)* 127 (4), S. 667–677.

Shen, Michael M.; Hodgkin, Jonathan (1988): *mab-3*, a gene required for sex-specific yolk protein expression and a male-specific lineage in *C. elegans*. In: *Cell* 54 (7), S. 1019–1031.

Shilova K.S.: Biologie of insect pupae from fish ponds. In: *Trudy biol. Sta. Borok* 1958 (3), S. 250–258.

Southern, E. M. (1975): Detection of specific sequences among DNA fragments separated by gel electrophoresis. In: *Journal of molecular biology* 98 (3), S. 503–517. DOI: 10.1016/S0022-2836(75)80083-0.

Steigemann, Wolfgang; Weber, Ernst (1979): Structure of erythrocrucorin in different ligand states refined at 1.4 Å resolution. In: *Journal of molecular biology* 127 (3), S. 309–338. DOI: 10.1016/0022-2836(79)90332-2.

Steinemann, M. (1978): Co-replication of satellite DNA of *Chironomus melanotus* with mainband DNA during polytenization. In: *Chromosoma* 66 (2), S. 127–139.

Steinemann, M.; Pinsker, W.; Sperlich, D. (1984): Chromosome homologies within the *Drosophila obscura* group probed by in situ hybridization. In: *Chromosoma* 91 (1), S. 46–53.

- Stiehl, Regina (2009): Klonierung und Analyse der geschlechtsbestimmenden Region bei *Chironomus piger* und *Chironomus thummi*. Diplomarbeit. Johannes Gutenberg-Universität, Mainz. IMSB.
- Suzuki, Masataka G.; Funaguma, Shunsuke; Kanda, Toshio; Tamura, Toshiki; Shimada, Toru (2003): Analysis of the biological functions of a doublesex homologue in *Bombyx mori*. In: *Development Genes and Evolution* 213 (7), S. 345–354. DOI: 10.1007/s00427-003-0334-8.
- Suzuki, Masataka G.; Ohbayashi, Fumi; Mita, Kazuei; Shimada, Toru (2001): The mechanism of sex-specific splicing at the doublesex gene is different between *Drosophila melanogaster* and *Bombyx mori*. In: *Insect Biochemistry and Molecular Biology* 31 (12), S. 1201–1211. DOI: 10.1016/S0965-1748(01)00067-4.
- Taenzler, Verena; Bruns, Eric; Dorgerloh, Michael; Pfeifle, Verena; Weltje, Lennart (2007): Chironomids: suitable test organisms for risk assessment investigations on the potential endocrine disrupting properties of pesticides. In: *Ecotoxicology* 16 (1), S. 221–230.
- Thompson, Peter E.; Bowen, Jean S. (1972): Interactions of differentiated primary sex factors in *Chironomus tentans*. In: *Genetics* 70 (3), S. 491–493.
- Thompson, Peter E.; English, Darrel S. (1966): Multiplicity of hemoglobins in the genus *Chironomus* (Tendipes). In: *Science* 152 (3718), S. 75–76.
- Traut, Walther (1994): Sex determination in the fly *Megaselia scalaris*, a model system for primary steps of sex chromosome evolution. In: *Genetics* 136 (3), S. 1097–1104.
- Trewitt, Patrick M.; Bergtrom, Gerald (1981): Immunological characterization of the haemoglobins of *Chironomus thummi* (Diptera). In: *Insect Biochemistry* 11 (6), S. 635–644.
- Trewitt, Patrick M.; Boyer, David R.; Bergtrom, Gerald (1986): Characterization of maternal haemoglobins in the eggs and embryos of *Chironomus thummi*. In: *Journal of insect physiology* 32 (11), S. 963–969.
- Trewitt, Patrick M.; Saffarini, Daad A.; Bergtrom, Gerald (1987): Nucleotide sequence of the intronless gene expressing a member of the globin VIIB subfamily from *Chironomus thummi* (Diptera). In: *Nucleic acids research* 15 (13), S. 5494.
- Verhulst, E. C.; van de Zande, L. (2015): Double nexus--Doublesex is the connecting element in sex determination. In: *Briefings in Functional Genomics*. DOI: 10.1093/bfpg/elv005.

- Verhulst, Eveline Catharina (2011): Genetic basis of sex determination in the haplodiploid wasp 'Nasonia vitripennis'. [S.l.: s.n.].
- Vinogradov, Serge N.; Walz, Daniel A.; Pohajdak, Bill (1992): Organization of non-vertebrate globin genes. In: *Comparative Biochemistry and Physiology Part B: Comparative Biochemistry* 103 (4), S. 759–773.
- Watson, James D. (2011): Watson Molekularbiologie. 6., aktualisierte Aufl., [1. dt. Ausg.]. München: Pearson Studium.
- Weber, Roy E. (1980): Functions of invertebrate hemoglobins with special reference to adaptations to environmental hypoxia. In: *American Zoologist* 20 (1), S. 79–101.
- Weber, Roy E.; Braunitzer, Gerhard; Kleinschmidt, Traute (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. In: *Comparative Biochemistry and Physiology Part B: Comparative Biochemistry* 80 (4), S. 747–753.
- Weinstock, George M.; Robinson, Gene E.; Gibbs, Richard A.; Worley, Kim C.; Evans, Jay D.; Maleszka, Ryszard et al. (2006): Insights into social insects from the genome of the honeybee *Apis mellifera*. In: *Nature* 443 (7114), S. 931–949. DOI: 10.1038/nature05260.
- Weller, P.; Jeffreys, A. J.; Wilson, V.; Blanchetot, A. (1984): Organization of the human myoglobin gene. In: *The EMBO journal* 3 (2), S. 439.
- Wells, Dan; McBride, Celeste (1989): A comprehensive compilation and alignment of histones and histone genes. In: *Nucleic acids research* 17 (Suppl), S. r311.
- Wiborg, Ove; Hyldig-Nielsen, Jens Jørgen; Jensen, Erik Ø.; Paludan, Kirsten; Marcker, Kjeld A. (1982): The nucleotide sequences of two leghemoglobin genes from soybean. In: *Nucleic acids research* 10 (11), S. 3487–3494.
- Wilkins, Adam S. (1995): Moving up the hierarchy: a hypothesis on the evolution of a genetic sex determination pathway. In: *Bioessays* 17 (1), S. 71–77.
- Wulker, W.; Am Klotzli (1973): Revision der Gattung *Chironomus* Meig. IV. Arten des *Lacunarius*-(*commutatus*-) Komplexes. In: *Archiv für Hydrobiologie*.

Wülker, Wolfgang; Götz, Peter (1968): Die Verwendung der Imaginalscheiben zur Bestimmung des Entwicklungszustandes von Chironomus-Larven (Dipt.). In: *Zeitschrift für Morphologie der Tiere* 62 (4), S. 363–388.

Zacharias, H. (1981): Sex-linked difference in DNA content of a polytene chromosome in *Prodiamesa* (Chironomidae). In: *Chromosoma* 82 (5), S. 657–672.

Zacharias, Helmut (1979): Underreplication of a polytene chromosome arm in the chironomid *Prodiamesa olivacea*. In: *Chromosoma* 72 (1), S. 23–51. DOI: 10.1007/BF00286428.

10 Anhang

2 Material und Methoden

2.12 Polymerase Kettenreaktion (PCR) (Mullis et al. 1986)

zu 2.12.1 Standard-PCR

-Primerliste: siehe Excel-Datei im elektronischen Anhang:

Primerübersicht_Doktorarbeit_Sarah_Brunck_Prodiamesa.xlsx

3 Ergebnisse

3.5 Die Histon-Gene der Zuckmücke *P. olivacea*

zu 3.5.3 Der Vergleich der Histone mit anderen Dipteren-Arten

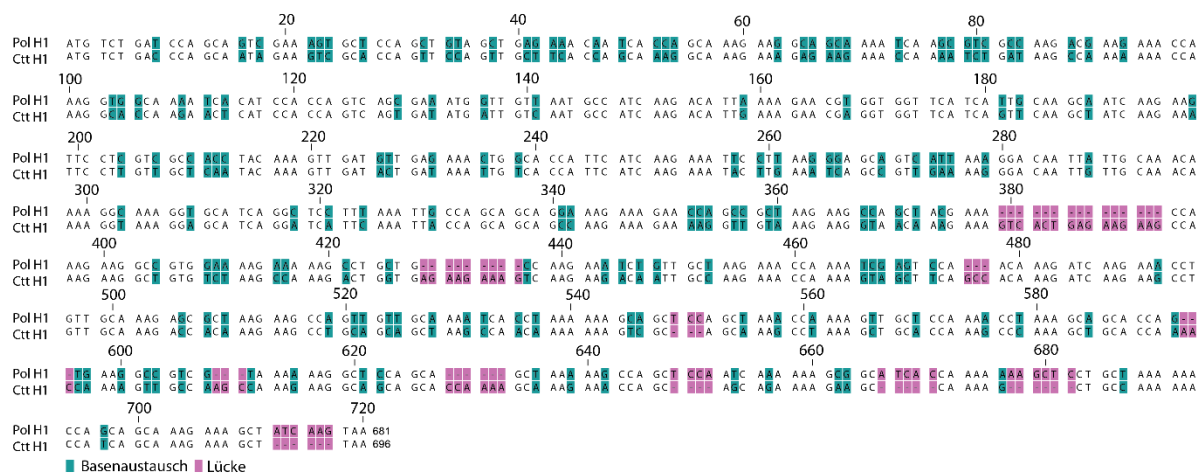


Abbildung 52 *Cthu* und *Poli* H1 Nukleotid-Alignment

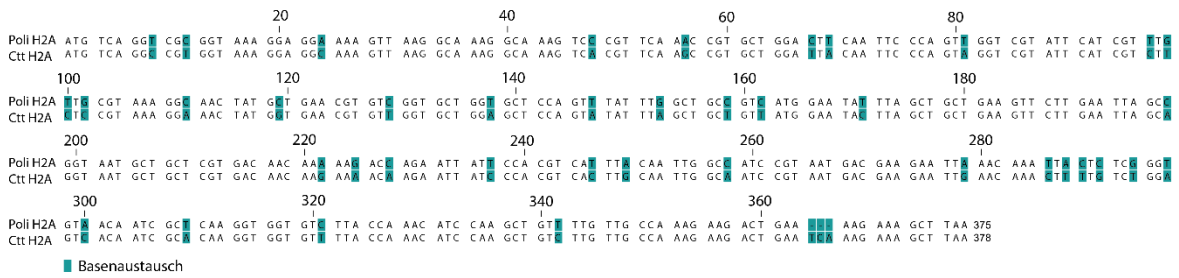


Abbildung 53 Cthu und Poli H2A Nukleotid-Alignment

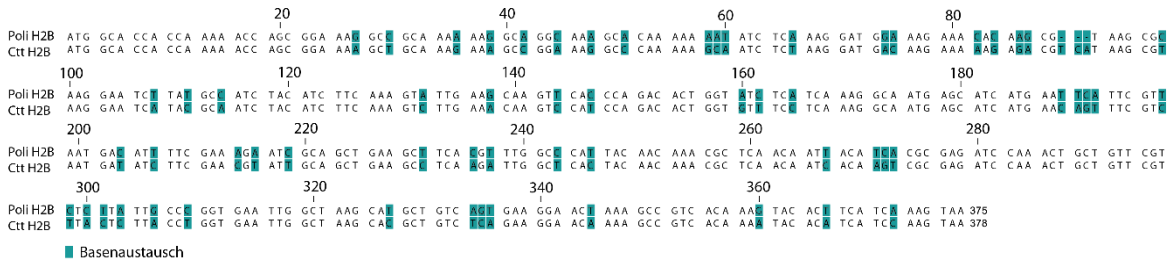


Abbildung 54 Cthu und Poli H2B Nukleotid-Alignment

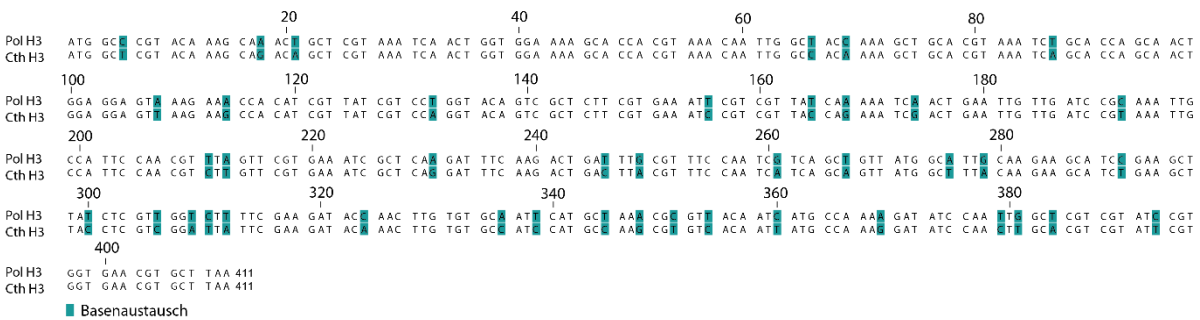


Abbildung 55 Cthu und Poli H3 Nukleotid-Alignment

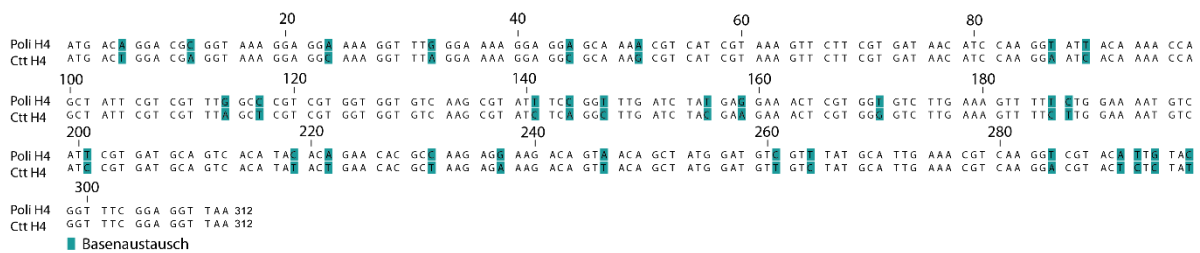


Abbildung 56 Cthu und Poli H4 Nukleotid-Alignment

3.6 Die Identifikation von homologen Genen aus anderen Dipteren in *P. olivacea*

3.6.1 Identifikation, Charakterisierung und Lokalisation von Genen aus der Geschlechtsbestimmungskaskade von *D. melanogaster* in *P. olivacea*

dsx

-geschlechtsunspezifische Exon-Varianten:

Excel-Datei im elektronischen Anhang: Exon Varianten in *Cthu dsx1b* und *Poli dsx1*.xlsx

3.7 Analyse des larvalen *P. olivacea*-Transkriptoms

msRNA98Cb-1 bzw.-2

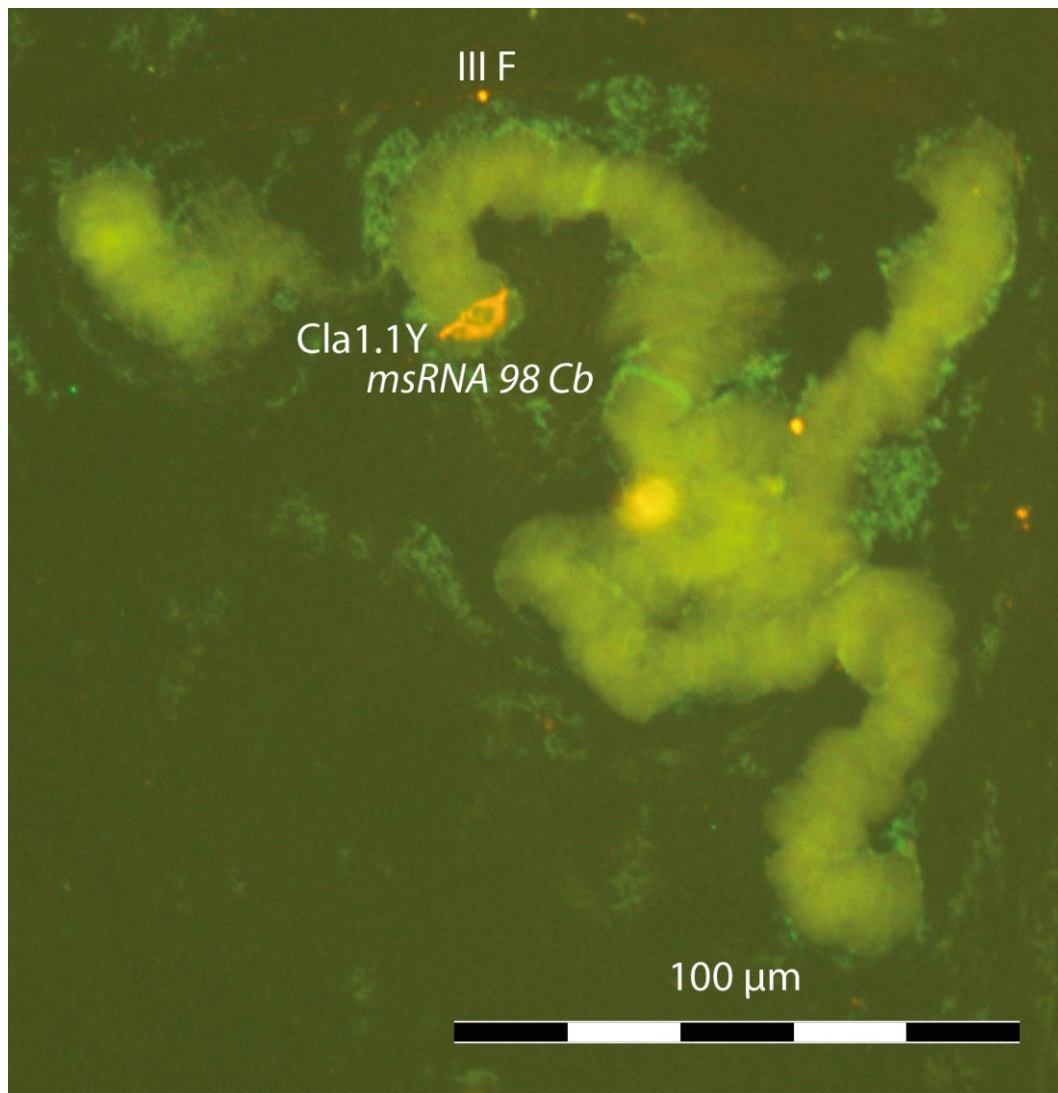


Abbildung 57 Mikroskopische Aufnahme von Polytrichrosomen aus *C. thummi* in der das Ergebnis der FISH mit den Sonden *Cla1.1Y* (TRITC, rot) und *msRNACb* (FITC, grün) abgebildet ist.

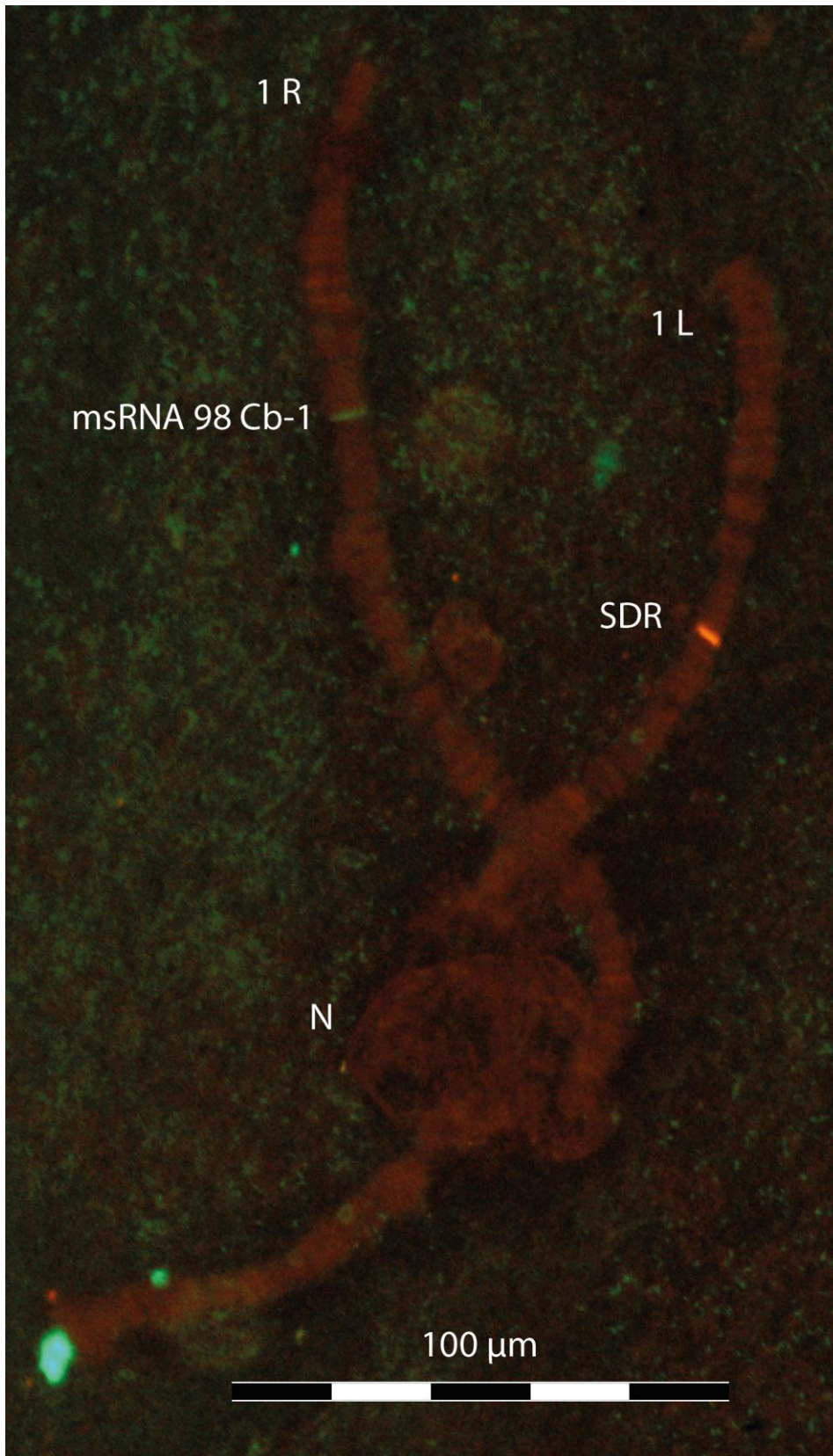


Abbildung 58 Mikroskopische Aufnahme von Polytänchromosomen aus *P. olivacea* in der das Ergebnis der FISH mit den Sonden SDR (TRITC, rot) und msRNACb-1 (FITC, grün) abgebildet ist.

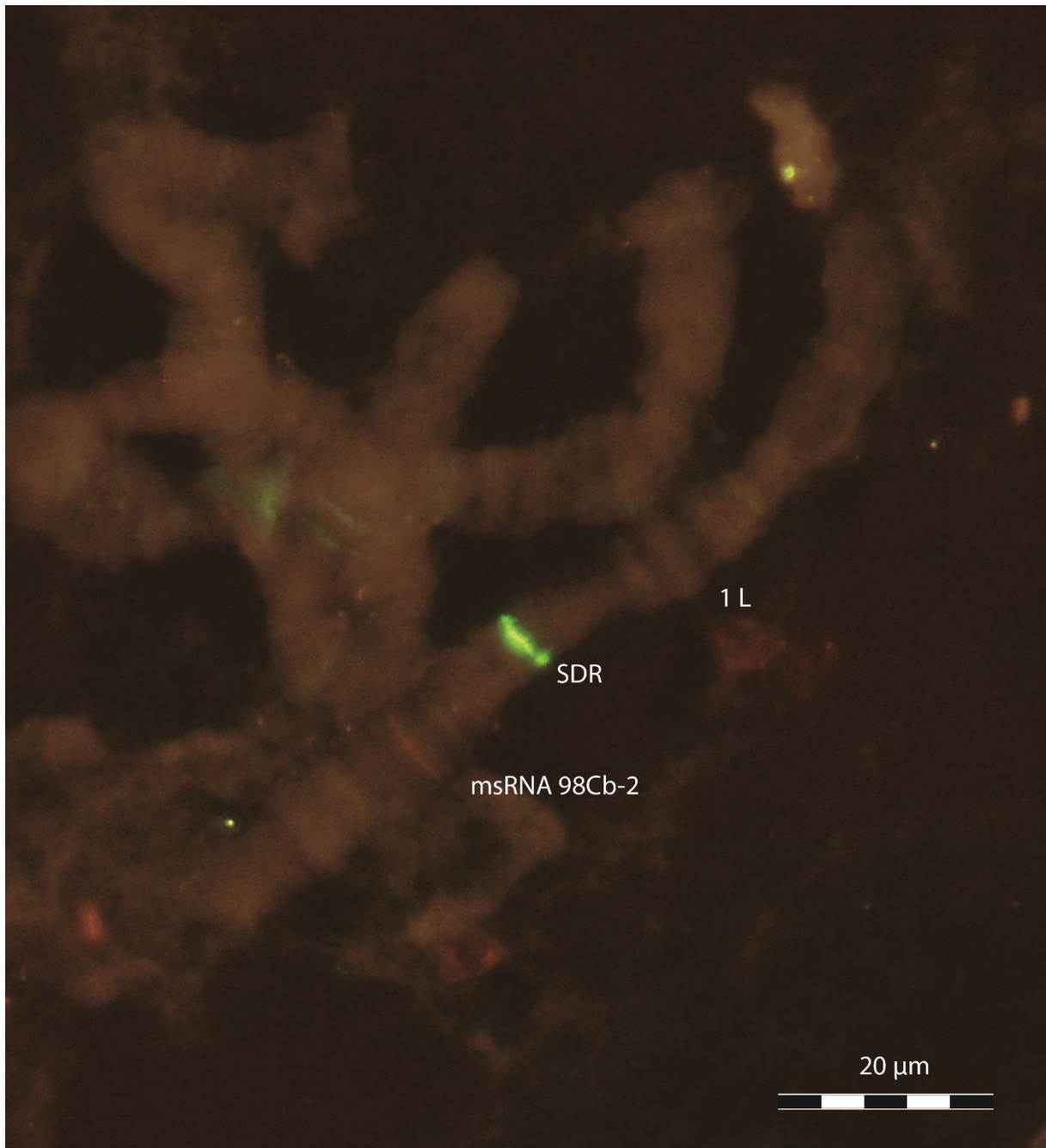


Abbildung 59 Mikroskopische Aufnahme von Polytänchromosomen aus *P. olivacea* in der das Ergebnis der FISH mit den Sonden SDR (FITC, grün) und msRNACb-2 (TRITC, rot) abgebildet ist.

4 Diskussion

4.1.2 Genfamilien aus *P. olivacea*

Die Hämoglobine-Gene

-Excel-Datei: Globin_all_invariant_semiinvariante_Reste.xlsx

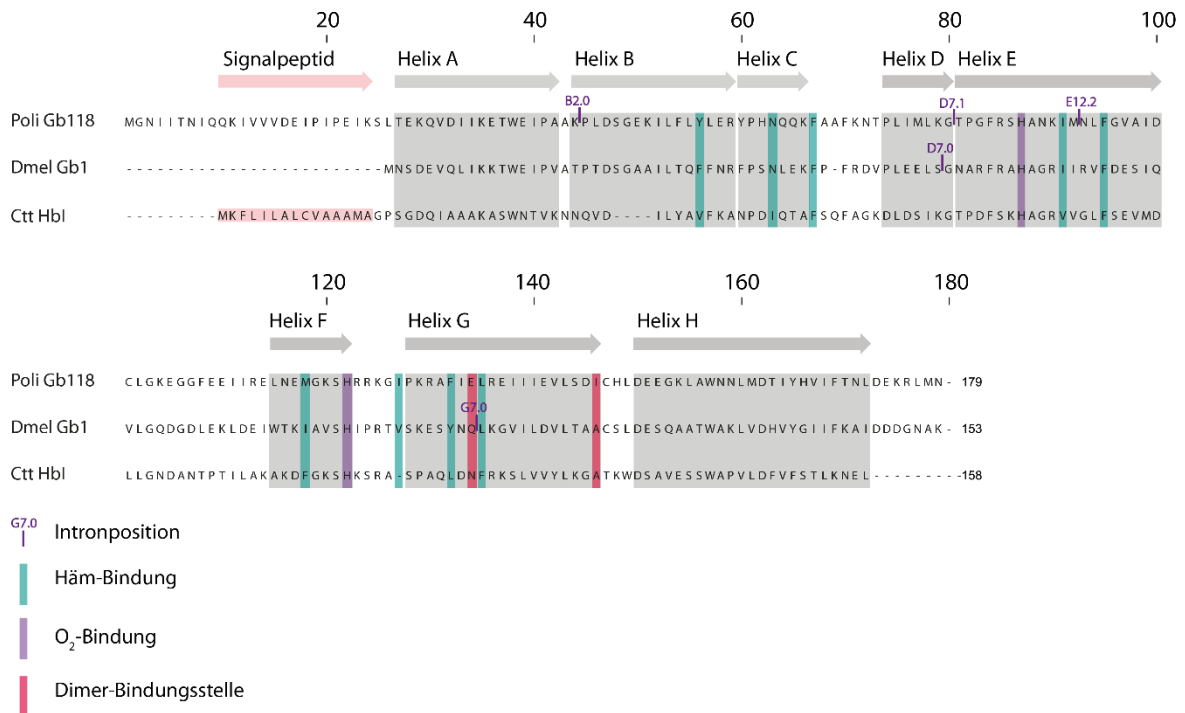


Abbildung 60 MSA des Globin-Clusters 118 aus *P. olivacea* und den Referenzsequenzen *D. melanogaster* Globin-1 und *C. thummi* HbI. Die Helices A bis H sind im MSA grau unterlegt. Das Signalpeptid von CthuHbI ist rosa unterlegt. Weiter sind funktionell bzw. strukturell wichtige AS-Reste hervorgehoben. Wie zum Beispiel Positionen (bzw. AS-Reste) welche bei der Ausbildung von Häm-Kontakten beteiligt sind (türkis unterlegt) oder AS-Reste welche für die Sauerstoffbindung verantwortlich sind (lila unterlegt). Aber auch Positionen welche an der Dimerisierung beteiligt sind (pinkrot unterlegt). Introngrenzen der dazugehörigen Gensequenzen sind im AS-sequenzalignment violett unter Positionsangabe in der Helix markiert.

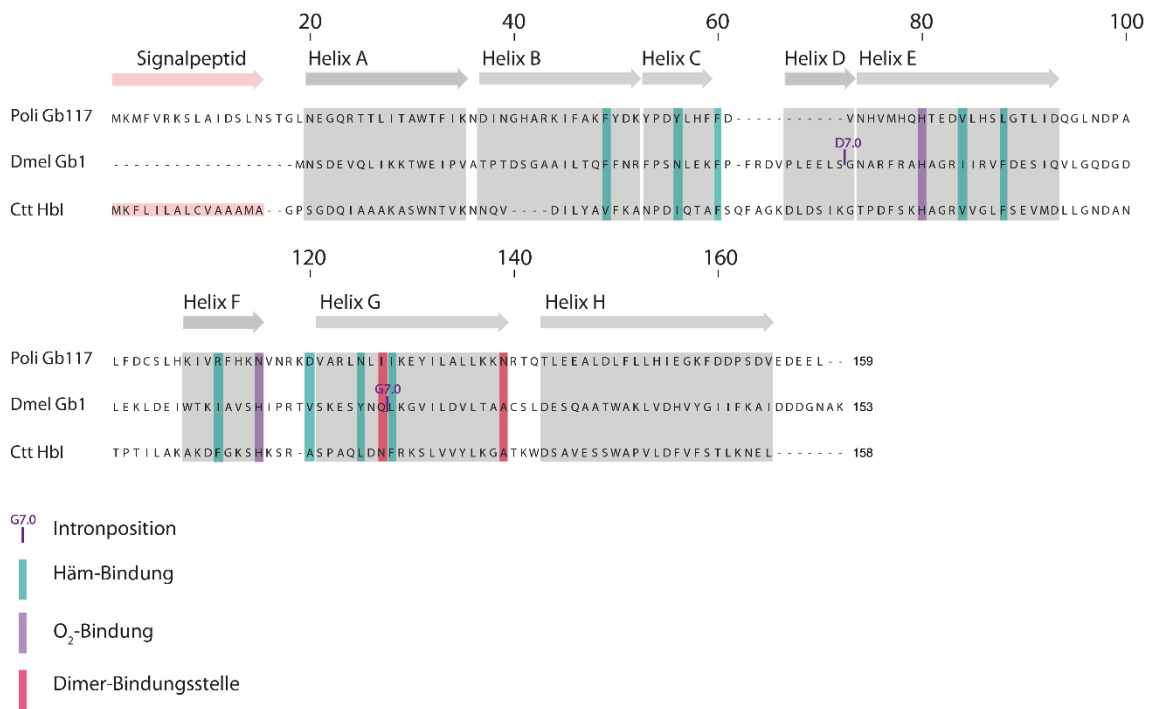


Abbildung 61 MSA des Globin-Clusters 117 aus *P. olivacea* und den Referenzsequenzen *D. melanogaster* Globin-1 und *C. thummi* HbI. Die Helices A bis H sind im MSA grau unterlegt. Das Signalpeptid von *CthHbI* ist rosa unterlegt. Weiter sind funktionell bzw. strukturell wichtige AS-Reste hervorgehoben. Wie zum Beispiel Positionen (bzw. AS-Reste) welche bei der Ausbildung von Häm-Kontakten beteiligt sind (türkis unterlegt) oder AS-Reste welche für die Sauerstoffbindung verantwortlich sind (lila unterlegt). Aber auch Positionen welche an der Dimerisierung beteiligt sind (pinkrot unterlegt). Intronengrenzen der dazugehörigen Gensequenzen sind im AS-Sequenzalignment violett unter Positionsangabe in der Helix markiert.

11 Curriculum Vitae

12 Danksagung

13 Eidesstattliche Erklärung

Versicherung gemäß § 11, Abs. 3d der Promotionsordnung

1. Ich habe die als Dissertation vorgelegte Arbeit selbst angefertigt und alle benutzten Hilfsmittel (Literatur, Apparaturen und Material) in der Arbeit angegeben.
2. Ich habe und hatte die als vorgelegte Dissertation/Arbeit nicht als Prüfungsarbeit für eine staatliche oder andere wissenschaftliche Prüfung eingereicht.
3. Ich hatte weder die als Dissertation vorgelegte Arbeit noch Teile einer Abhandlung davon bei einer anderen Fakultät bzw. einem anderen Fachbereich als Dissertation eingereicht.

Mainz, den 17. Februar 2016
