University of South Bohemia in České Budějovice

Faculty of Science

Testing the synteny of genes in *Yponomeuta evonymella*

Master thesis

Bc. et Bc. Veronika Běhálková

Supervisor: RNDr. Petr Nguyen, Ph.D.

Supervisor-specialist: Mgr. Irena Provazníková

České Budějovice 2019

Běhálková, V., 2019: Testing the synteny of genes in *Yponomeuta evonymella*. Mgr. Thesis, in English – 34 pp., Faculty of Science, University of South Bohemia, České Budějovice, Czech Republic.

Annotation

The aim of this thesis was to create a physical map of the bird-cherry ermine, *Yponomeuta evonymella*, chromosome Z_1 and chromosome 15 using BAC-FISH mapping. The synteny of genes was compared to the closest outgroup of *Yponomeuta* spp., *Plutella xylostella*, in order to determine whether the *P. xylostella* genome sequence could be used as reference in reference-assissted genome assembly of *Yponomeuta* spp.

Declaration

I hereby declare that I worked on this Master thesis independently and used only the sources listed in the bibliography.

I hereby declare that, in accordance with Article 47b of Act No. 111/1998 in the valid wording, I agree to the publication of this Master thesis in full in electronic form on the publicly accessible web page of the STAG database operated by University of South Bohemia in České Budějovice, the copyright of this thesis text being retained.
I further agree to the electronic publication of the supervisor's and the opponent's assessments of this thesis as well as the electronic publication of the record of the proceedings and the result of this thesis defense in accordance with the above mentioned Act No. 111/1998.
I also agree to the comparison of the text of my thesis with the Theses.cz database operated by the National Registry of University Theses and with a plagiarism detection system.

České Budějovice, 11th December, 2019

…….

Bc. et Bc. Veronika Běhálková

Acknowledgements

I would like to express my sincere gratitude to my supervisor, RNDr. Petr Nguyen, Ph.D., who kindly accepted me into his laboratory, readily shared his expert knowledge, provided excellent and timely guidance and lessened the strain of the hard work involved with his kind and positive approach and his good sense of humor.

I would also like to express my sincerest gratitude to my supervisor-specialist, Mgr. Irena Provazníková, who taught me the research methods employed in this study with great kindness and patience, answered all of my inquisitive questions, and also conducted the research preceding this study. I would like to thank Irenka for always being there for me when I needed her and for her always being such a ray of sunshine in our laboratory.

In addition, I would like to thank Martina, Sander, Anička, Lafi, Atsuo and Magda, who kindly helped me whenever I asked them, and created the friendliest of atmospheres in our laboratory.

Last, but not least, I would like to thank my family, my boyfriend and my friends who kindly supported me throughout my master's degree program.

The present research was funded by the Czech Science Foundation grant reg. no. 17-17211S.

Table of contents

1 INTRODUCTION

1.1 The role of chromosomal rearrangements in speciation

Chromosomal rearrangements, while sometimes deleterious to an afflicted organism, also represent an evolutionary mechanism that contributes to the complex processes of adaptation and speciation on the molecular level. Changes in the number and/or structure of chromosomes affect the rate of recombination by altering linkage relationships (Lukhtanov et al. 2018). However, the specifics of accumulation and fixation of novel rearrangements remain unresolved (Vershinina and Lukhtanov 2017).

One of the rearrangements that plays an important role in local adaptation and diversification by maintaining polymorphism in complex traits are chromosomal inversions. Inverted sequences are protected from recombination and consequently, act as an effective carrier of locally adapted sets of alleles that can be retained in the face of gene flow (Wellenreuther and Bernatchez 2018). Several studies in *Drosophila* and some butterflies have shown that inversions are more abundantly distributed in sympatric, rather than allopatric sister species (Brown et al. 2004; Kandul, Lukhtanov, and Pierce 2007; Noor et al. 2001), thus supporting their role in establishing genetic isolation between an incipient species and its source population.

Recently, Lukhtanov et al. (2018) hypothesized that complex multivalents segregation in heterokaryotypes may be facilitated by means of inverted meiosis in organisms with holocentric chromosomes such as moths and butterfies. The numerous and diverse Lepidoptera thus represent an ideal model system for studying chromosomal rearrangements and their role in speciation.

1.2 The karyotype of Lepidoptera

Lepidoptera (moths and butterflies) represent one of the largest groups of animals with nearly 160,000 described species (Zhang 2011) and also the most diverse group of animals with the WZ/ZZ (female/male) sex determination system (Bachtrog et al. 2014). Both physical and genetic mapping as well as genome studies across different lepidopteran taxa confirmed the ancestral haploid chromosome number to be n = 31 (2n = 62) (Van't Hof et al. 2013; Ahola et al. 2014). It was shown that synteny, i.e. the chromosomal localization of orthologous genes between species, remains well conserved in Lepidoptera for more than 140 My (Ahola et al.

2014) and the overall stability of the lepidopteran genome is thus very high. However, there are many known examples of lepidopteran species harboring a different haploid chromosome number as the result of rearrangements. For instance, the blue butterfly *Polyommatus atlantica* (Lycaenidae) represents with its n = 224–226 (Lukhtanov 2015) the upper limit of the lepidopteran chromosome number distribution, which is the broadest of all insects (Blackmon, Ross, and Bachtrog 2017).

Such seemingly contradictory observations might be resolved, if we look at the cytogenetic characteristics of the lepidopteran karyotype. First of all, lepidopteran chromosomes are numerous and small which reduces the probability of interchromosomal interactions (Van't Hof et al. 2013). Secondly, chromosomes of moths and butterflies are holokinetic, i.e. they lack primary constriction (centromere) and kinetochores cover a substantial part of their surface, so that microtubules attach along the whole length of the chromosomes during cell division (Bureš, Zedek, and Marková 2013). Lepidopteran karyotype is therefore fairly resistant to the deleterious consequences of chromosomal rearrangements such as gene loss and unbalanced gametes because even a smaller fragment of the original chromosome can be passed on without substantial damage.

1.3 Neo-sex chromosomes in Lepidoptera

Considerable attention has been drawn to the relatively high frequency of the fusions of sex chromosomes with autosomes. In several lepidopteran species, the haploid chromosome number had diverged from the ancestral n = 31 by just one rearrangement event – the formation of a neo-sex chromosome. In the codling moth (*Cydia pomonella*, Tortricidae), chromosome Z was shown to form a long fused element with an autosome corresponding to Chr15 in the *Bombyx mori* reference genome (Nguyen et al. 2013). Sex chromosomes in the monarch butterfly (*Danaus plexippus*, Nymphalidae) and its congeners have been reported to form neo-Z and neo-W (Mongue et al. 2017). The monarch Z-autosome fusion involved the autosome homeologous to *B. mori* Chr16. Recently, Provazníková and Nguyen (unpublished data) described in detail the neo-W chromosome in the bird-cherry ermine (*Yponomeuta evonymella*, Yponomeutidae), the fused autosome corresponding to Chr2 in the *B. mori* reference genome.

What makes lepidopteran sex chromosomes so prone to fusions? Chromosome W is known to be non-recombining due to achiasmatic female meiosis and largely heterochromatic and abundant with repetitive elements (Marec and Traut 1993; Sahara, Yoshido, and Traut 2012). Although recombination occurs between Z chromosomes in males, its overall frequency

is considerably reduced compared to autosomes, since Z is present in a hemizygous state in females (Sahara, Yoshido, and Traut 2012). Consequently, Z undergoes faster evolution and can accumulate repetitive sequences, which was confirmed by Bellott et al. (2010) in chicken. Repetitive sequences may, in turn, facilitate ectopic recombination and consequently, formation of rearrangements (Pennell et al. 2015).

Regarding the autosomes involved in rearrangements, Ahola et al. (2014) reported that shorter chromosomes with high frequency of transposable elements show higher rearrangement rates than long ones in Lepidoptera. In addition, the autosomes involved in fusions with sex chromosomes do not seem to be random. It is hypothesized (Nguyen, personal communication) that they carry gene clusters expressed in ovaries as shown in the *B. mori* reference genome (Suetsugu et al. 2013). This is in accordance with sexually antagonistic selection driving sex chromosome turnover (Charlesworth D and Charlesworth B 1980).

1.4 Ermine moths of the genus *Yponomeuta*

The small ermine moth genus *Yponomeuta* (Yponomeutidae) includes 76 species with a wide distribution across the Palaearctic and an occasional distribution in other parts of the world (Turner et al. 2010). The genus belongs to the superfamily Yponomeutoidea which represents one of the earliest diverging ditrysian lineages (Mutanen, Wahlberg, and Kaila 2010). Most of the *Yponomeuta* species feed on trees and shrubs of the spindle family Celastraceae. However, a number of European species are specialist monophagous feeders on Rosaceae and Salicaceae. Celastraceae, abundant in Asia, have been identified as the ancestral host plant association of the *Yponomeuta* ermine moths, while the Rosaceae and Salicaceae associations are considered a result of secondary host shifts that probably occurred in the ancestor of the European '*Y. cagnagella-irrorella* clade' (Turner et al. 2010). Since the early 1970s, the genus has been a model system for multidisciplinary studies in the evolution of phytophagous insects and their host plant associations (Menken, Herrebout, and Wiebes 1992).

Specialized phytophagous insects employ highly sensitive and specific chemoreceptors to recognize plant compounds that signal the suitability of host plants (Roessingh et al. 1999). Electrophysiological experiments revealed that the *Euonymus*-feeding *Yponomeuta* species are chemosensitive to the sugar alcohol dulcitol predominant in Celastraceae, whereas sorbitol, the predominant sugar alcohol in Rosaceae, acts as a phagostimulant to the Rosaceae-feeders (Menken, Herrebout, and Wiebes 1992). Besides, *Prunus* (Rosaceae)-feeders reacted positively to dulcitol as well, which was presently explained by low amounts of dulcitol contained in

Prunus spp. It was therefore suggested that these sugar alcohols have played a key role in the shift from the ancestral hosts of the family Celastraceae to Rosaceae and that the low amount of dulcitol present in some Rosaceae might have facilitated the host shift (Menken and Roessingh 1998).

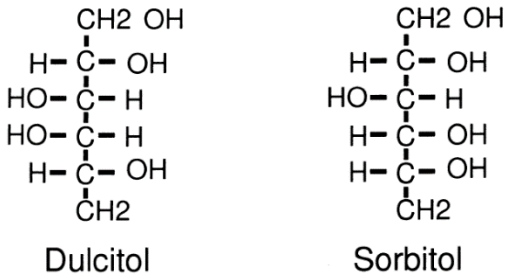

Figure 1: Dulcitol and sorbitol are stereo-isomers and differ only in the orientation of a single hydroxyl group (Roessingh et al. 1999).

Roessingh et al. (1999) further studied the larval perception of the two sugar alcohols in *Y. evonymella* feeding exclusively on *Prunus padus* (Rosaceae) and in inter-specific hybrids of *Y. cagnagella* feeding on the *Euonymus europaeus* (Celastraceae) and *Y. padella* which is oligophagous on a number of Rosaceae species. Roessingh reported that the sugar alcohol sensitivity is localized in a single cell and argues that a relatively simple genetic modification of the sugar receptor proteins might have enabled the host-shift from Celastraceae to Rosaceae.

Additional research is needed to clarify this evolutionary novelty. Yet gustatory receptor genes were successfully identified and located in specific chromosomes in the model species, *B. mori* (Wanner and Robertson 2008; Guo et al. 2017). Putative sugar receptor genes *BmGr4*, *BmGr5* and *BmGr6* were reported to form a cluster on Chr15 in *B. mori* (Guo et al. 2017).

1.5 *Yponomeuta evonymella*

Yponomeuta evonymella (bird-cherry ermine) represents a model organism for studying sympatric speciation, insect-host associations, host races formation, and perception of plant phagostimulants (Menken, Herrebout, and Wiebes 1992; Roessingh et al. 1999; Turner et al. 2010). It is a day active moth that feeds exclusively on the bird cherry (*Prunus padus*) and its distribution ranges from Europe to northern and eastern Asia. Bird-cherry ermine is univoltine, the larvae are gregarious and produce extensive common nests. The larvae may defoliate entire trees during severe infestation. Therefore, *Y. evonymella* is considered a minor pest (Menken, Herrebout, and Wiebes 1992).

As mentioned above, the karyotype of *Y. evonymella* differs from the ancestral chromosome number n = 31 by one sex chromosome-autosome fusion (Nilsson, Löfstedt, and Dävring 1988; Provazníková and Nguyen unpublished). During cell division, a WZ_1Z_2 trivalent forms in females whereas two bivalents Z_1Z_1 and Z_2Z_2 form in males (Figure 2). As mentioned above, the autosome participating in this fusion event was identified as a homeolog of *B. mori* Chr2 (Provazníková and Nguyen unpublished).

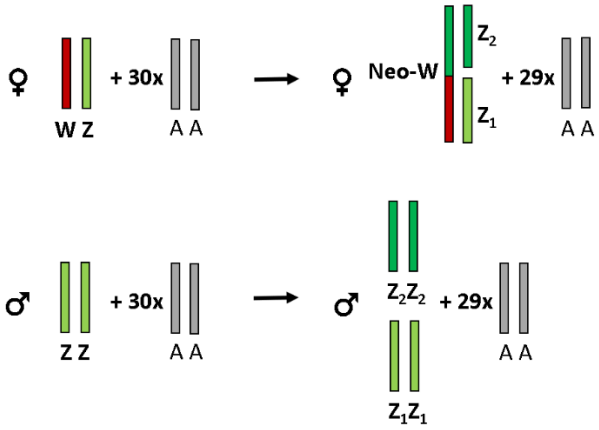

Figure 2: Schematic formation of neo-W in *Y. evonymella* by W-autosome fusion. Ancestral W element shown in red.

1.6 Current shortcomings in genome assembly

Each year, the number of assembled genomes increases owing to advances in DNA sequencing technology. However, building the genome architecture from individual DNA sequences up to the level of highly dynamic chromosome organization remains a complex and arduous task. Historical approaches to genome assembly included chromosome-based large insert clones such as bacterial artificial chromosomes (BACs) or yeast artificial chromosomes (YACs) resulting in a chromosome-based assembly. On the other hand, most recently assembled genomes feature short contigs and often lack even a basic physical map or chromosome number (Deakin et al. 2019). Cytogeneticists are concerned that without precise interrelationship between taxonomy, cytogenetics, and genomics, we might end up with numerous draft genome assemblies whose contigs and scaffolds bear no relationship to proper linkage maps or the chromosomes with which they are associated (Traut et al. 2017).

At present, researchers focus on assembling genomes of non-model species, e.g. one of the goals of our laboratory is to acquire the *Yponomeuta cagnagella* genome sequence. Since

many genomes of model organisms have been successfully assembled and reassembled reaching chromosome level quality, a novel approach to genome assemblies employing a reference genome was proposed: reference-assisted chromosome assembly (Tamazian et al. 2016). This type of genome assembly approach includes arranging contigs and scaffolds into putative chromosomes using information from the reference genome of a closely related species, thus reducing the overall number of DNA sequence fragments from thousands to hundreds or even dozens, and altogether simplifying the whole process of genomic analysis (Tamazian et al. 2016). However, such approach should be accompanied by cytogenetic analysis of synteny of genes between the species of interest and the reference, which would support the suitability of the reference genome. Conserved synteny of genes is often assumed and anticipated but not tested. Even closely related species might have undergone major changes in karyotypes such as various chromosomal rearrangements. These might cause errors in reference-assisted assembly and therefore, cytogenetic data should be employed to clarify potential misinterpretations (Deakin et al. 2019). One example of such unconfirmed mapping included transcripts for boa and pygmy rattlesnake mapped to protein-coding genes of the lizard *Anolis carolinensis* which represented the closest reptile relative with a sequenced and annotated genome (Vicoso et al. 2013). While Vicoso et al. (2013) reported various level of differentiation of W sex chromosomes in three snake lineages with sex chromosomes being homomorphic in boas, Gamble et al. (2017) discovered that boas and phytons have XY sex chromosome systems of independent origin.

Identification of individual lepidopteran chromosomes is not easily accomplished due to their lack of discernible features. Lepidopteran chromosomes contain hardly any heterochromatin (apart from the sex chromosome W). Furthermore, traditional methods such as banding do not work in Lepidoptera (De Prins and Saitoh 2003). A new and significant road in mapping and analysis was opened by the introduction of fluorescence *in situ* hybridization (FISH) and its modifications. Labeled probes were employed in order to localize a target sequence identifying the chromosome of interest (Goldsmith and Marec 2010).

Effective experimental studies of lepidopteran karyotypes have been enabled by FISH mapping employing the above-mentioned BAC libraries (BAC-FISH). A genomic BAC library contains 100–200 kbp long DNA inserts that represent a genome of a given organism. BAC DNA can be labeled and used as a probe for *in situ* hybridizations. Fluorescence *in situ* hybridization of BAC probes represents an efficient tool for determining the chromosomal location of specific genes (gene mapping) or detecting rearranged regions in chromosomes (Yasukochi et al. 2009; Nguyen et al. 2013; Van't Hof et al. 2013). BAC-FISH is a useful

technique in phylogenetic and comparative studies which test hypotheses on the evolutionary history of chromosomes, most particularly sex chromosomes (Janes et al. 2011).

2 AIMS

The ultimate goal of our laboratory is to study the role of chromosomal rearrangements such as inversions in the adaptation and speciation of *Yponomeuta* species. The aim of the present thesis was to compare synteny of genes between *Yponomeuta* spp. (Yponomeutidae) and their close relative, a diamondback moth *Plutella xylostella* (Plutellidae), and to determine whether the available genome sequence of *P. xylostella* can be used for reference-assisted genome assembly in *Yponomeuta* spp. We took advantage of available BAC library of *Yponomeuta evonymella* to create physical maps of selected chromosomes of *Y. evonymella*, namely $ChrZ_1$ and Chr15, using BAC-FISH.

3 MATERIAL AND METHODS

3.1 Experimental individuals

The *Yponomeuta evonymella* larvae used in this study were collected from their food plant, the bird cherry *Prunus padus*, in the village of Vrábče, Czech Republic, on May, 26th, 2019.

3.2 Preparation of *Y. evonymella* chromosome spreads

Meiotic chromosome spreads were prepared from gonads of IV. instar larvae. Testes and ovaries were dissected from the larvae in cold saline for *Ephestia kuehniella* (0.9 % NaCl, 0.042 % KCl, 0.025 % $CaCl_2$, 0.02 % $NaHCO_3$; Glaser 1917 cited in Lockwood 1961) and soaked in hypotonic solution (75 mM KCl) for 10 min. Afterward, the tissue was fixed in freshly prepared Carnoy's solution (ethanol:chloroform:acetic acid, 6:3:1) for 15 min. The tissue was macerated in 10 µl of 60 % acetic acid on a clean microscope slide until the material was homogenized. The slide was placed on a plate preheated to 45 °C and the cell suspension was evenly spread. Finally, the chromosome spreads were dehydrated in the series of 70 %, 80 % and 100 % ethanol, 1 min each, and stored at –20 °C for further use. The remains of the dissected larvae were frozen in liquid nitrogen and stored at –20 °C for later extraction of genomic DNA.

3.3 Isolation of *Y. evonymella* genomic DNA

Genomic DNA was isolated from the IV. instar larvae using the NucleoSpin® DNA Insect purification kit (Macherey-Nagel, GmbH & Co. KG, Düren, Germany). The manufacturer's protocol was slightly modified. At the beginning, the tissue was homogenized with a pestle in the mixture of 100 µl of the Elution Buffer BE, 40 µl of the Buffer MG and 10 µl of Proteinase K provided by the manufacturer. The samples were briefly centrifuged and incubated at RT for 20 min. 600 µl of Buffer MG was added and from this point, the manufacturer's protocol was followed. In the end, DNA was eluted twice. The concentration of the yielded DNA was measured using Qubit 3.0 (Thermo Fisher Scientific, Waltham, MA, USA).

3.4 Identification of chromosomal markers

In order to create a physical map of the *Yponomeuta evonymella* sex chromosome Z_1 ($ChrZ_1$) and the chromosome 15 (Chr15), nine and seven markers were chosen, respectively, based on

the known orthologs from the reference genome of the closely related diamondback moth, *Plutella xylostella*. The chosen orthologs are characterized in Table 1. For $ChrZ_1$, the markers matched those mapped previously by Petr Nguyen in *P. xylostella* (Fig. 3, Nguyen P, unpublished data) apart from *Henna,* which was identified as Z-linked in *P. xylostella* by Dalíková et al. (2017).

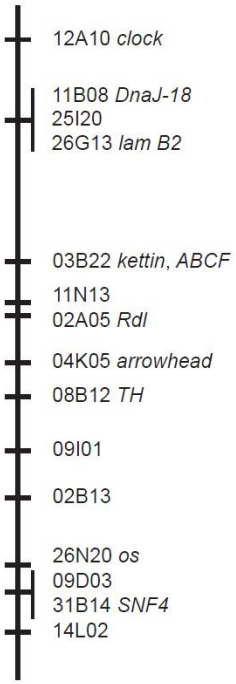

Figure 3: Comparative map of *P. xylostella* $ChrZ_1$ (Nguyen P, unpublished data).

The markers for Chr15 apart from *acetylcholinesterase I* were chosen to match the set of the *P. xylostella* BAC clones bearing single-copy genes identified by Yasukochi et al. (2011). *Acetylcholinesterase I* is localized in Chr15 in *Bombyx mori* but interestingly, was identified as Z-linked in *Cydia pomonella* as the result of the fusion between $ChrZ_1$ and Ch15 establishing neo-Z (Nguyen et al. 2013).

Table 1: Overview of the chosen chromosomal markers.

Name of gene ortholog	Abbreviation	Chromosomal location	*Bombyx mori* ortholog	Localization in *B. mori*
*henna/ tryptophan phenylalanine 4-monooxygenase**	*henna*	Z_1	BMgn003866	chr1:21842454-21845665
*SNF4/AMP-activated protein kinase gamma subunit**	*SNF4*		BMgn012310	chr1:15595590-15678086
fushi tarazu	*ftz*		BMgn000716	chr1:11668015-11753033
tyrosin hydroxylase	*TH*		BMgn000563	chr1:8795363-8803219
arrowhead	*arh*		BMgn003888	chr1:21849863-21852235
*gamma-amminobutyric-acid A receptor, beta subunit (Resistant to dieldrin)**	*rdl*		BMgn000568	chr1:8060590-8089612
surfeit 1	*surf*		BMgn000722	chr1:12008314-12011576
DnaJ	*DnaJ*		BMgn000608	chr1:5808189-5812185
clock	*clock*		BMgn000498	chr1:12215106-12256266
ribosomal protein S23	*RpS23*	15	BMgn007645	chr15:1859286-1862652
ribosomal protein L7A	*RpL7A*		BMgn007661	chr15:2912132-2916752
ribosomal protein L5	*RpL5*		BMgn007879	chr15:5494651-5498849
ribosomal protein S5	*RpS5*		BMgn007710	chr15:7586843-7587692
ribosomal protein L8	*RpL8*		BMgn007743	chr15:9127595-9130339
ribosomal protein L10A	*RpL10A*		BMgn003337	chr15:14491197-14493497
acetylcholinesterase 1	*Ace1*		BMgn003320	chr15:15498774-15629164

*Name of ortholog adapted from flybase.org (25th November 2019).

3.5 *Y. evonymella* BAC library screening

3.5.1 *Y. evonymella* BAC library

The *Y. evonymella* genomic library of bacterial artificial chromosomes (BACs) was constructed by AC Amplicon Express (Pullman, WA, USA). High molecular weight gDNA of *Y. evonymella* males was partially digested by *Hind* III and cloned into the pCC1BAC (Epicentre, Madison, WI, USA) vector, which was then transformed into the DH10B *Escherichia coli* cells. The library consists of 20736 clones with the average insert of about 125 Kbp.

To identify BAC clones bearing genes of interest, i.e. a particular plate, row, and column of the BAC library, the library was screened by means of PCR according to the manufacturer's

guidebook. For the screening procedure, individual BAC clones were merged into 18 sets (or superpools) by the manufacturer, each containing 1,152 BAC clones. Round I PCR (screening of superpools) is performed on all of the superpools to identify which superpool contains BAC clone(s) with the sequence of interest (there may be more than one superpool identified). Each superpool corresponds to an another BAC clones subset consisting of 21 matrixpools. The matrixpools are plate, row and column pools combined, so that the sought sequence of interest could be identified in one of 3 plates which were previously identified by the positive superpool, i.e. one of the 3x16 rows and one of the 3x24 columns. Round II PCR screening of the respective matrixpools serves as the final indicator of the plate, row and column of the BAC library where the clone bearing the gene of interest is located.

3.5.2 Primer design

The *Y. evonymella* trascriptome (Nguyen, unpublished data) was searched for orthologs of selected *P. xylostella* genes using BLAST (https://blast.ncbi.nlm.nih.gov/Blast.cgi; Altschul et al. 1990). Exon boundaries were further identified by alignment of identified transcripts with the genome sequence of closely related *Yponomeuta cagnagella* (Nguyen, unpublished data). Primers were designed in a chosen exon of each gene. The primers used for screening are shown in Table 2.

Table 2: Primers used for screening and identified BAC clones.

Marker	Forward primer sequence (5'→3')	Reverse primer sequence (5'→3')	Annealing t [°C]	Product size [bp]	BAC clone
henna	AACCTCAGCCA CATCGAGTC	GCTCTGAGCCATA CGACAGG	60	680	26J01
SNF4	CGTCACAGGAG GCCGAGAAG	CCCCCTGGAGTCT CTAAAGAGT	60	122	15N20
ftz	GTGCTAGATCA CCTCCACCAG	TCGAATTTGAGTT CCGCCAGT	62	164	16I13
TH	GTGGAGTTCGG TTTGTGCAAG	GGCGTCTTCGAA ACTTTCAGC	60	201	43E06
arh	GGGCTACCACA AGAGTAAGGC	CCGGTATGCTGGT GTTTCTTC	60	207	39G04
rdl	GTTACTCGGCAC TACTCCGTC	CACGACATCGGA AACATGCAG	58	118	-
	CTCGTTTGTGTC CAACTGCTG	CATAGCTGATGTG GGTGTCCA	58	259	-
surf	TCGAACAGATG TCAGCCCAC	ACAAAATAGCGG TGCCACATG	62	190	20D10
DnaJ	TGGTTGCTTGGT GGCATTTTC	CGTGGTCAGGGC TCCAGTTAG	62	163	54G24
clock	CGATTGGAAGC CGACGTTTC	CGGATTATGTCCC AGCAGTGA	58	145	-
	TGAAACGAGGT GATCCGACAG	CTATGCTTGTCCA CTGCTGGA	58	289	-
RpS23	ACTCCAGGAAT GTCACCAACG	CAGCCCAACTCTG CTATCCG	62	173	14I05
RpL7A	ATATCCAGCCA ACACGTGACC	GGTCTGGGTGAA CTGGTTGAT	58	128	-
RpL5	ACCCAAATACC GCCTGATTGT	TGGGGAGTTCAT GTGAGTAAGC	62	116	18H08
RpS5	CCTAACTGGCG AAAACCCTCT	CCGCAATGGTCTT GATGTTCC	58	206	-
RpL8	CAACCATGCCC CTGTTACTTG	GGAAATGTAATG CCAGTAGGTGC	58	205	09N16
RpL10A	TTCTGTGTCTGT CTGTGGCTG	GCTTAGTACAATC TCTGGGGAGG	60	172	13C11
Ace1	TCCTTACGTGAC TGATGTGGC	GGGCTAATTCGTT AACGGCAC	60	146	47O13

3.5.3 Confirmation of the primers on the *Y. evonymella* gDNA

As the first step of the BAC library screening, primers were tested and their annealing temperature was optimized on a temperature gradient. Three annealing temperatures, 58, 60 and 62 °C, were tested. PCR reaction of the final volume of 10 µl consisted of 15–20 ng of *Y. evonymella* gDNA, 0.25 U of One *Taq*® Quick-Load® DNA Polymerase (NEB, Ipswich, MA, USA), 0.2 mM dNTPs, 1x One *Taq*® Quick-Load® Reaction Buffer (NEB), and 3 µM each forward and reverse primer. PCR reactions were run in a thermal cycler set to the following temperature profile: initial denaturation at 95 °C for 3 min; 30 cycles consisting of denaturation at 95 °C for 30 s, annealing for 30 s, and elongation at 68 °C for 30 s; and final elongation at

68 °C for 3 min. The amplicons were separated on a 1.5 % agarose/1x TAE gel and visualized by ethidium bromide staining at a final concentration of 0.5 µg/ml.

3.5.4 Screening of superpools (Round I PCR)

Secondly, 18 superpools provided by the manufacturer of the BAC library were each used as template in PCR reaction using the validated primers. One superpools screening reaction of the final volume of 10 µl consisted of 1 µl of the provided superpool DNA, 0.25 U of One *Taq*® Quick-Load® DNA Polymerase (NEB), 0.2 mM dNTPs, 1x One *Taq*® Quick-Load® Reaction Buffer (NEB), and 3 µM each forward and reverse primer. 15–20 ng of *Y. evonymella* gDNA was used for positive control. The superpools screening reactions were run in a thermal cycler with the following temperature profile: initial denaturation at 94 °C for 5 min; 35 cycles of denaturation at 94 °C for 30 s, annealing for 30 s and elongation at 68 °C for 30 s; and final elongation at 68 °C for 3 min. The amplicons were separated by gel electrophoresis and visualized as described in the previous chapter. The marker's sequence was detected only in the superpool or superpools containing the sought BAC clone. Three repeats of Round I PCR were performed for each marker on the three superpools sets provided by the manufacturer. Only the superpool(s) with the positive hit was further investigated.

3.5.5 Screening of matrixpools (Round II PCR)

Thirdly, the matrixpools belonging to the identified positive superpool were screened. One matrixpools screening reaction of the final volume of 10 µl contained the same compounds as the superpools screening reaction except for the template that consisted of 1 µl each of the provided 21 matrixpools. 15–20 ng of *Y. evonymella* gDNA was used for positive control and the matrixpools screening reactions were run in a thermal cycler set to the temperature profile described in the previous chapter. The amplicons were separated by gel electrophoresis, stained and visualized as described in Chapter 3.5.3. The presence of the marker's sequence in a particular bacterial artificial chromosome and its location in a particular plate, row and column of the *E. coli* BAC library was determined.

3.5.6 Clone test

Lastly, *E. coli* cells containing the identified BAC were streaked onto LB agar plates containing chloramphenicol (12.5 µg/ml) and the plates were incubated overnight at 37 °C. To confirm the presence of the gene of interest, colony PCR was performed using single *E. coli* colonies as template. One clone test reaction of the final volume of 10 µl consisted of the same compounds as one superpools screening reaction, and the clone test reactions were run in a thermal cycle set to the same temperature profile (Chapter 3.5.4). The amplicons were visualized in the same

way as described in Chapter 3.5.3. The presence of the gene of interest in the BAC clone identified by screening was thus confirmed.

3.6 BAC DNA isolation

The positive *E. coli* clones bearing the genes of interest were cultured in 200 ml of LB medium with chloramphenicol (12.5 µg/ml) for ~15 hours at 37 °C at 200 rpm. The cells were harvested by centrifugation and BAC DNA was isolated using NucleoBond® Xtra Midi plasmid DNA purification kit (Macherey-Nagel) according to the manufacturer's protocol. The isolated BAC DNA was dissolved in sterile water at 4 °C overnight, the concentration was measured using Qubit 3.0 (Thermo Fisher Scientific) and a small sample was checked on a 1 % agarose gel in 1x TAE.

3.7 Preparation of glycerol stocks

The confirmed *E. coli* clones were stocked for later use. *E. coli* cells were cultured in 4 ml of LB medium with chloramphenicol (12.5 µg/ml) and incubated for ~15 hours at 37 °C at 200 rpm. 500 µl of the cell culture was mixed with 500 µl of 25 % glycerol, inverted gently a few times and immediately frozen in liquid nitrogen. The glycerol stocks were stored in –80 °C.

3.8 Preparation of *Y. evonymella* competitor DNA

Competitor DNA was isolated from the frozen larvae using the CTAB (hexadecyltrimethylammonium bromide) method which was adapted from Winnepenninckx et al. (1993). The larval tissue was disrupted with a pestle in 800 µl of the CTAB buffer consisting of 2 % CTAB (Sigma-Aldrich, St. Louis, MO, USA), 100 mM Tris-HCl (pH 8), 40 mM EDTA, 1.4 M NaCl, 0.2 % β-mercaptoethanol and 0.1 mg/ml of Proteinase K (Macherey-Nagel) and shaken at 400 rpm at 60 °C overnight. Next day, the same volume of chloroform was added, the samples were mixed gently by inverting for 2 min and centrifuged at 14 000 x g for 10 min at 4 °C. The separated DNA in the aqueous phase was transferred to a clean tube and the chloroform step was repeated to increase the purity of the isolated DNA. Afterward, 62.5 µg/ml of RNase A (Sigma-Aldrich) was added and the samples were incubated for 30 min at 37 °C. Two thirds of the recovered volume of isopropanol was added, the samples were mixed gently by inverting and left to precipitate at 4 °C overnight. Next day, the DNA was pelleted by centrifugation at 14 000 x g for 15 min at 4 °C. The pellets were washed twice with cold 70 % ethanol (max speed, 15 min, 4 °C). The pellets were left to air dry and resuspended in sterile

H_2O overnight. DNA concentration was measured by Qubit 3.0 and a small sample was checked on a 1 % agarose gel in 1x TAE. The isolated DNA was repeatedly amplified with ~50 ng of template using Illustra GenomiPhi HY DNA Amplification Kit (GE Healthcare, Buckinghamshire, UK) according to the manufacturer's protocol. Quantification of DNA was performed as described above. Prior to use in BAC FISH, competitor DNA was fragmented by boiling at 99 °C for 20 min.

3.9 Labeling of probes by nick-translation

Isolated BAC DNA was labeled using nick-translation. The protocol for the nick-translation reaction was based on Kato et al. (2006) and slightly modified. BAC DNA was labeled by additional incorporation of fluorescent aminoallyl-dUTP-Cy3 (Jena Bioscience, Jena, Germany). The modified final volume of 20 µl of the nick-translation reaction contained 1 µg of BAC DNA, 20 U of DNA Polymerase I (Thermo Fisher Scientific), 20 µM aminoallyl-dUTP-Cy3 (Jena Bioscience), 50 µM dATP, dCTP, dGTP and 10 µM dTTP, 0.00025 U of DNase I (RNase-free, Thermo Fisher Scientific), 10 mM β-mercaptoethanol, and 1x nick-translation buffer (0.5 M Tris-HCl pH 7.5, 50 mM $MgCl_2$, and 0.05 % BSA). The reaction mixture was incubated in a thermal cycler at 15 °C for 4.5 hours. Immediately afterward, the enzymes were inactivated by 1x loading buffer consisting of 25 mM EDTA (pH 8), 0.6 mM bromphenol blue, and 5 % glycerol. To verify the length of the labeled fragments which should range between 300 and 500 bp for optimal performance in hybridization, a small sample of the nick-translation reaction was run on a 1.5 % agarose gel in 1x TAE. The labeled probes were kept in dark at –20 °C.

3.10 BAC-FISH

Fluorescence *in situ* hybridization of BAC probes was conducted according to Nguyen et al. (2013).

3.10.1 Preparation of the hybridization mix

The hybridization mix for one slide was prepared by combining 500 ng of probe with 2.5 µg of *Y. evonymella* competitor DNA (see Chapter 3.8) and 25 µg of salmon sperm acting as a carrier and unspecific competitor. For precipitation, 1/10 of the hybridization mix volume of 3M sodium acetate and 2.5x of the hybridization mix volume of cold 100 % ethanol were added. The mixture was left to precipitate in –80 °C for 1 hour. The precipitated DNA was centrifuged at 15 000 rpm at 4 °C for 20 min. The supernatant was discarded and DNA was washed with

500 µl of cold 70 % ethanol and centrifuged at 15 000 rpm at 4 °C for 10 min. The supernatant was carefully removed in stages and the DNA was left to partially air dry. The DNA was dissolved in 50 % deionized formamide and the mixture was incubated in 37 °C for 30 min, occasionally vortexed and spun down. 10 % dextran sulphate in 2x SSC was added to the dissolved probe and the mixture was vortexed and briefly centrifuged. Afterward, the probe was denatured at 90 °C for 5 min and immediately chilled on ice for 4 min. After vortexing and short centrifuging, the denatured probe was ready for application on a slide.

3.10.2 Preparation of slides

The slides were removed from a freezer, dehydrated by immersing in the series of 70 %, 80 % and 100 % ethanol, 1 min each, and air dried. The chromosome spreads were pretreated by RNase (200 µg/ml) and incubated at 37 °C for 1 hour. Afterward, the slides were washed three times 5 min in 2x SSC at room temperature in water bath with shaking. Next, the slides were incubated in 5x Denhardt's solution (0.1 % polyvinylpyrolidol, 0.1 % BSA, 0.1 % Ficoll 400) at 37 °C for 30 min in water bath with shaking. The chromosome spreads were denatured in 70 % formamide at 68 °C for 3.5 min and immediately dehydrated in 70 % ethanol precooled to –20 °C for 1 min and then in room temperature 80 % and 100 % ethanol, 1 min each. Once the preparations air dried, hybridization mix was applied on them and sealed under the cover slip with rubber cement glue. The preparations were placed into a wet chamber to hybridize at 37 °C for 3 days.

3.10.3 Washing

After 3 days, the rubber cement glue and cover slip were removed and the slides were washed in 1 % Triton X-100 in 0.1x SSC at 62 °C for 5 min in water bath with shaking. Then, the slides were washed in 1 % Kodak PhotoFlo at RT for 2 min. The slides were left to partially air dry for a couple of minutes and the chromosomes were mounted with DAPI (Sigma-Aldrich) in DABCO (Sigma-Aldrich) at the concentration of 0.5 µg/ml. The cover slip was sealed with nail polish and the slides were stored at 4 °C in dark for at least 12 hours before documentation.

3.11 Documentation and image processing

The hybridization was documented using Zeiss Axioplan 2 microscope (Carl Zeiss, Jena, Germany) equipped with respective filters and the Olympus DP73 camera (Olympus Europa Holding, Hamburg, Germany) using the cellSens Standard software. Black and white photographs were pseudocolored and superimposed in Adobe Photoshop CS4, version 11.0.

3.12 Reprobing

For the reprobing process, the hybridization mix was prepared and denatured as described in Chapter 3.10.1. Nail polish and cover slip were removed from the slides and the preparations were washed twice in 2x SSC at RT for 2 min in water bath with shaking. Afterward, the preparations were postfixed in freshly prepared 4 % paraformaldehyde in 2x SSC at RT for 10 min in water bath with shaking. The slides were washed three times in 2x SSC at RT for 5 min. Afterward, the chromosome spreads were denatured by incubation in 50 % formamide 1 % Triton X-100 in 0.1x SSC at 70 °C for 10 min in water bath with shaking. The slides were immediately placed into –20 °C 70 % ethanol for 2 min and further dehydrated in the ethanol series (80 %, 100 %), 1 min each. The slides were air dried and thus prepared for the application of another hybridization mix. After 3 days, the slides were washed as described in Chapter 3.10.3 and visualized as described in Chapter 3.11.

4 RESULTS

4.1 Identification of BAC clones bearing the markers' sequences

The *Y. evonymella* BAC library was screened for 16 markers in total (Table 1). Gene markers for $ChrZ_1$ included *henna*, *arh*, *ftz*, *surf*, *rdl*, *TH*, *SNF4*, *DnaJ* and *clock*. Gene markers for Chr15 included the *RpS23*, *RpL7A*, *RpL5*, *RpS5*, *RpL8*, *RpL10A* and *Ace 1*. Consequently, 16 pairs of primers (one for each gene marker) were designed for the screening (Table 2). In 12 out of 16 cases, the BAC clone bearing the marker's sequence was identified (Table 2). An illustrative example of screening of superpools and matrixpools for *ribosomal protein L8* can be seen in Figure 4 and Figure 5, respectively. In the cases of *rdl* and *clock*, the presence of the screened marker was detected in all of the 18 superpools as can be seen in Figure 6. Therefore, an additional series of superpool screening was conducted using new primers designed in a different exon of the gene. However, many false positive results were again detected and I did not continue with the screening. In the cases of *RpL7A* and *RpS5*, a number of positive superpools were identified but the presence of the marker's sequence was detected in all of the screened matrixpools, so these markers were dismissed as well. Eventually, BAC clones were identified for 7 markers in $ChrZ_1$, namely *henna*, *arh*, *ftz*, *surf*, *TH*, *SNF4* and *DnaJ*, and 5 markers in Chr15, namely *RpS23*, *RpL5*, *RpL8*, *RpL10A* and *Ace 1*.

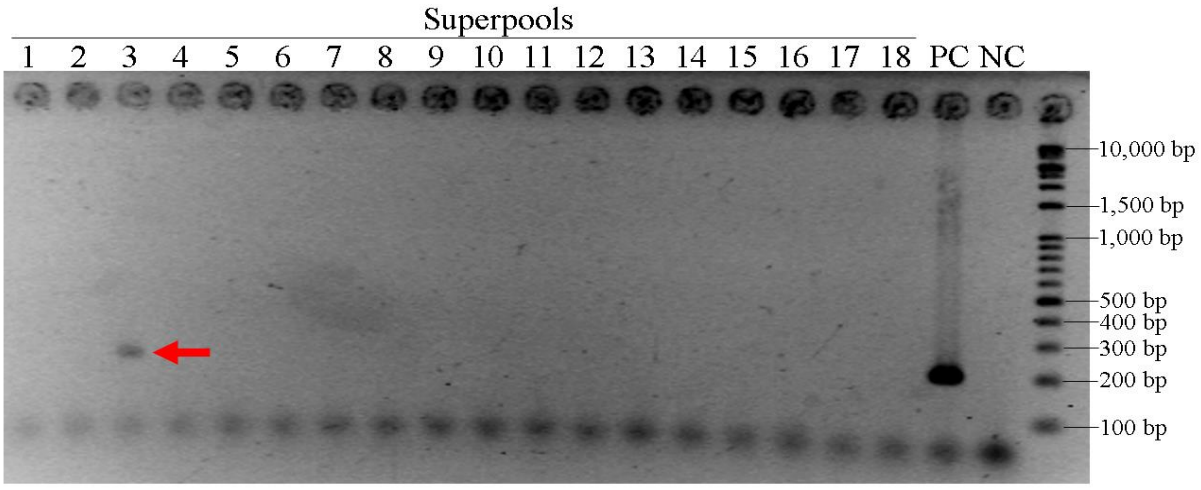

Figure 4: Superpools screening for *ribosomal protein L8*. The clone bearing the *RpL8* sequence is present in superpool 3. *Y. evonymella* genomic DNA was used for positive control (PC). Primer dimers can be seen in the lower part of the gel. NC – negative control.

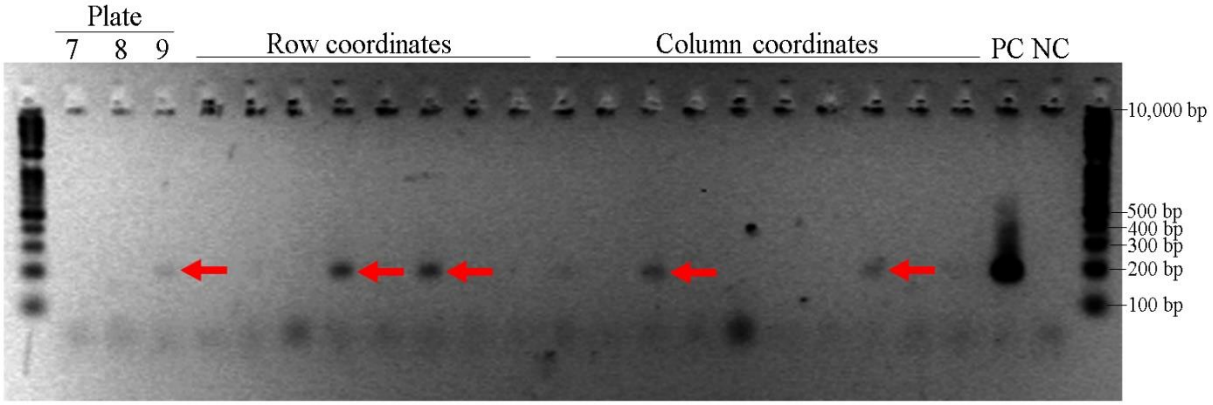

Figure 5: Screening of the 21 matrixpools belonging to superpool 3. The *RpL8* sequence is present in the third plate corresponding to superpool 3, i.e. plate 9, in row N and column 16 of the BAC library as evaluated according to the manufacturer's guidebook. *Y. evonymella* genomic DNA was used for positive control (PC). NC – negative control.

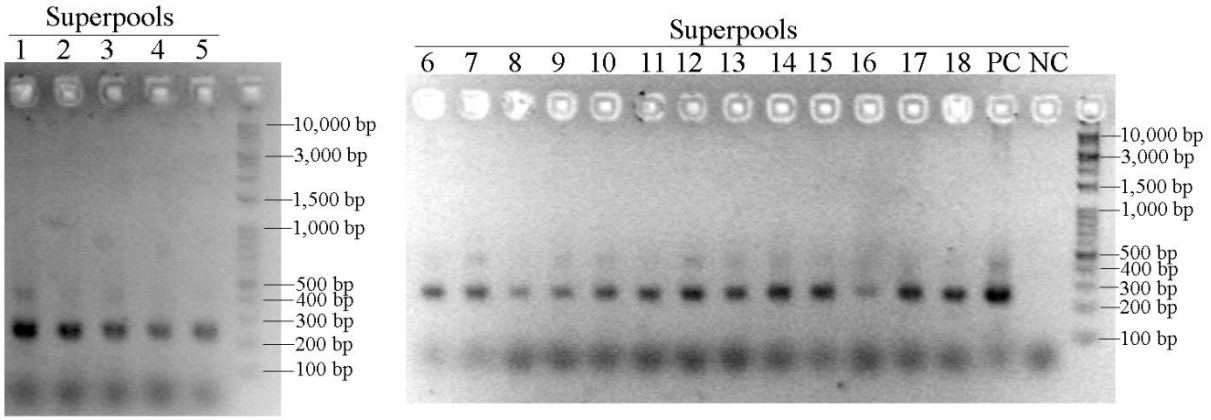

Figure 6: Superpools screening for the *resistant to dieldrin* gene. The marker's sequence seems to be present in all superpools. *Y. evonymella* genomic DNA was used for positive control (PC). NC – negative control.

4.2 Optimization of the BAC probes labeling

BAC DNA was labeled by the nick-translation reaction. Its incubation time was optimized to produce 300–500 bp long fragments suitable for effective fluorescence *in situ* hybridization. In our laboratory, final inactivation of DNA polymerase I and DNase I by incubation at 70 °C for 10 min was firmly established. However in the conducted labeling reactions, the activity of DNase I was not hindered by the increased temperature and BAC DNA was digested into excessively short fragments as can be seen in Figure 7a. The time of the labeling was gradually shortened from 4.5 hours down to 2 hours in the next several attempts but BAC DNA was

always excessively cleaved. Therefore, final inactivation of the enzymes by EDTA was attempted. For practical reasons, loading buffer (25 mM EDTA, pH 8; 0.6 mM bromphenol blue and 5 % glycerol) was added to the labeling reaction in the end instead of the temperature-based step. This modification of the nick-translation protocol produced much longer BAC DNA fragments as can be seen in Figure 7b. The fluorescent signal of BAC DNA probes prepared in this way and hybridized according to the BAC-FISH protocol was successfully detected (see Figure 8).

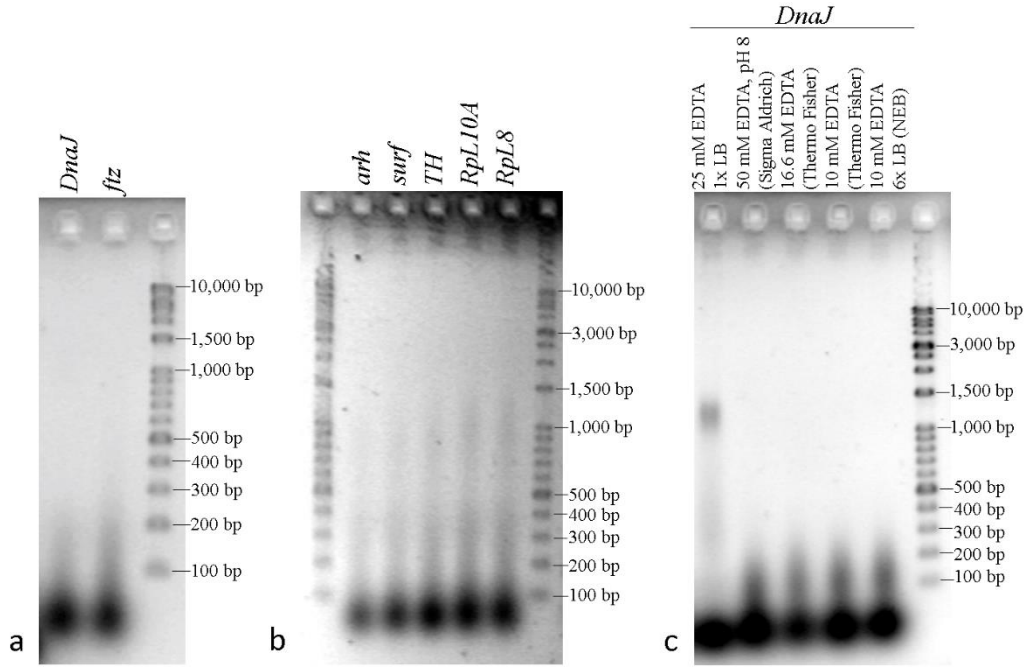

Figure 7: **(a)** Temperature-based inactivation of DNase I at the end of the labeling reaction. BAC DNA is excessively digested. **(b)** DNase I inactivation by 1x loading buffer (25 mM EDTA, pH 8; 0.6 mM bromphenol blue and 5 % glycerol). Labeled BAC DNA fragments varied in length up to ~1000 bp. **(c)** Inactivation of DNase I by EDTA of different origin and different concentrations. Only previously confirmed inactivation by 1x loading buffer resulted in the desired length of BAC DNA fragments. All samples were labeled 4.5 hours.

Further investigation into the DNase I inactivation issue showed that interestingly, neither 50 mM EDTA, pH 8 (Sigma Aldrich) nor 10 mM EDTA supplied with DNase I (Thermo Fisher Scientific) nor 10 mM EDTA in 6x loading buffer (NEB) is able to inactivate DNase I in the labeling reaction (Figure 7c). However, I did not look into this problem any further due to time reasons.

4.3 The reprobing approach

In our laboratory, the use of two fluorescent dyes for labeling has been well established. These are aminoallyl-dUTP-Cy3 and fluorescein-12-dUTP. Despite all efforts, I was unable to optimize the fluorescein-12-dUTP labeling of BAC DNA and achieve sufficient incorporation of fluorescently labeled nucleotides, i.e. sufficient level of signal intensity in BAC-FISH, due to technical issues. Therefore, BAC probes were labeled only by aminoallyl-dUTP-Cy3 and chromosome spreads were reprobed two times after the first hybridization.

4.4 Experimental design of the mapping of ChrZ$_1$ and Chr15

Experimental design of the mapping of ChrZ$_1$ and Chr15 is presented in Table 3. Seven ChrZ$_1$ markers were hybridized to male Z$_1$ pachytene bivalents. In the first BAC-FISH, the markers *DnaJ*, *ftz* and *SNF4* were hybridized to the chromosome spreads ♂1, ♂2 and ♂3, respectively. In the second BAC-FISH, the previously confirmed (Provazníková I, unpublished data) terminal marker, *henna*, was hybridized to all male spreads, representing an anchor for all Z$_1$ hybridizations. In the third BAC-FISH, the markers *arh*, *surf* and *TH* were hybridized to ♂1, ♂2 and ♂3, respectively. Thus, 3 gene markers were successively hybridized to each of the male chromosome spreads, *henna* being the common anchor marker.

Five gene markers were hybridized to pachytene bivalents of Chr15 present in female chromosome spreads. In the first BAC-FISH, the markers *RpL5* and *Ace1* were hybridized to the chromosome spreads ♀1 and ♀2, respectively. In the second BAC-FISH, the anchor marker *RpS23* was hybridized to both ♀1 and ♀2. In the third BAC-FISH, the markers *RpL10A* and *RpL8* were hybridized to ♀1 and ♀2, respectively. Thus, 3 gene markers were successively hybridized to each of the female chromosome spreads, *RpS23* being the common anchor marker.

Table 3: ChrZ$_1$ and Chr15 mapping design.

Chromosome	Chromosome spread	BAC FISH	1st reprobing	2nd reprobing
Z1	♂1	*DnaJ*	*henna*	*arh*
	♂2	*ftz*	*henna*	*surf*
	♂3	*SNF4*	*henna*	*TH*
15	♀1	*RpL5*	*RpS23*	*RpL10A*
	♀2	*Ace1*	*RpS23*	*RpL8*

4.5 Mapping of ChrZ$_1$ and Chr15

The position of the anchor marker, *henna*, was localized in 24 pachytene Z_1 bivalents found in all three male chromosome preparations. The positions of *DnaJ* and *arh* were localized in 10 pachytene bivalents of Z_1 present in chromosome spread ♂1. The positions of *ftz* and *surf* were localized in 8 pachytene bivalents of Z_1 in chromosome spread ♂2 and *SNF4* and *TH* were localized in 6 pachytene bivalents of Z_1 in chromosome spread ♂3.

The position of the anchor marker, *RpS23*, was localized in 25 pachytene bivalents of Chr15 found in both female preparations. The positions of *RpL5* and *RpL10A* were localized in 12 pachytene bivalents of Chr15 found in chromosome spread ♀1, and in 13 pachytene bivalents of 15 in chromosome spread ♀2, *Ace1* and *RpL8* were localized. Representative examples of the BAC-FISH mapping of YeChrZ$_1$ (*Y. evonymella* ChrZ$_1$) and YeChr15 (*Y. evonymella* Chr15) are shown in Figure 8.

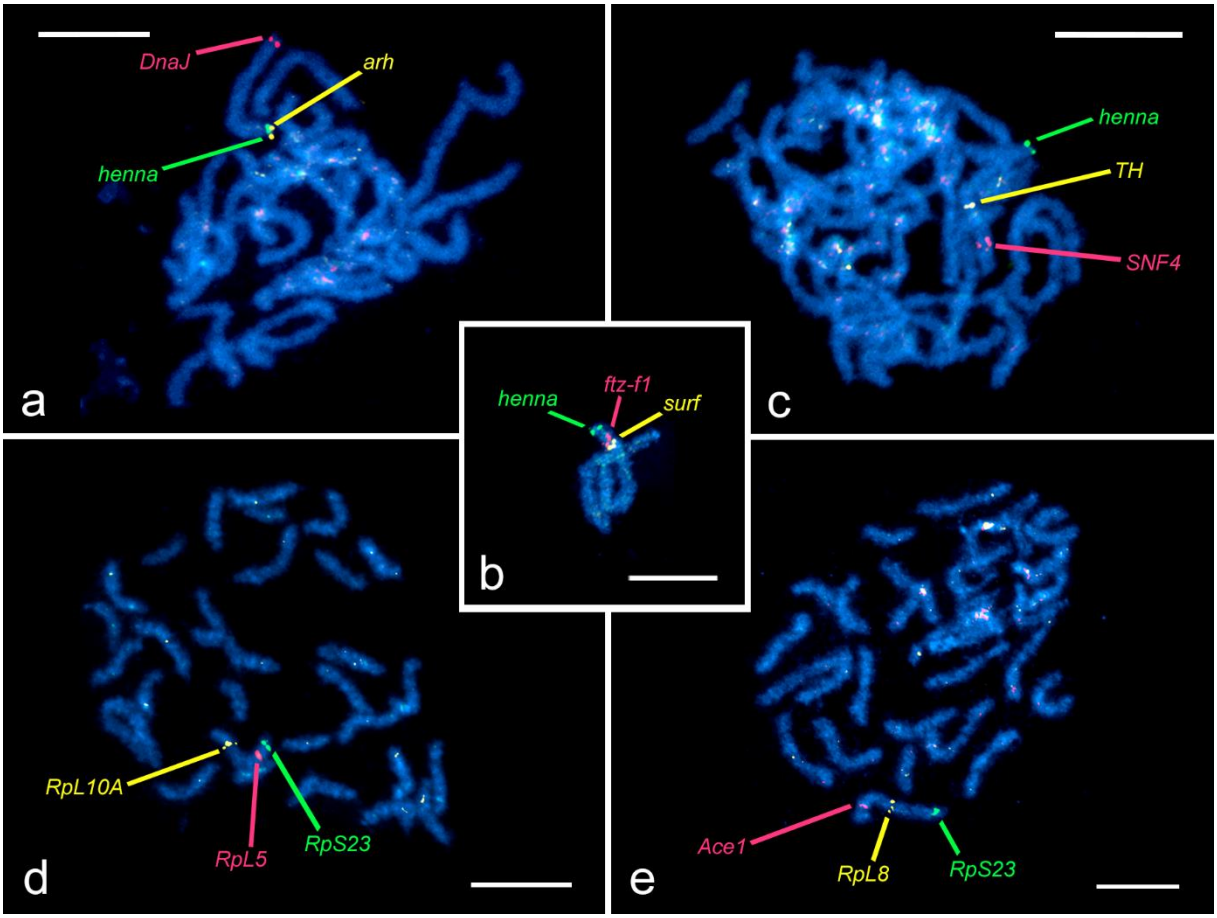

Figure 8: BAC-FISH mapping of genes on chromosome preparations of *Y. evonymella*. Chromosomes were stained with DAPI (blue). Hybridization signals of BAC probes (green, red, yellow) indicate the physical postitions of the markers. **(a-c)** Three runs of BAC-FISH localized the positions of *henna, arh,*

ftz, *surf*, *TH*, *SNF4* and *DnaJ* in YeChrZ$_1$. **(d, e)** Three runs of BAC-FISH localized the positions of *RpS23*, *RpL5*, *RpL8*, *RpL10A* and *Ace 1* in YeChr15. Scale bar = 10 μm.

4.6 Measuring and calculation of relative distances

Using the ImageJ software, the distance between a marker and the anchor-marked end of the chromosome was measured, and related to the whole length of the chromosome. Each pachytene bivalent was measured three times to increase the accuracy of the calculation. The average relative positions of the markers can be seen in Table 4 and Table 5. In ChrZ$_1$, the markers are located in the following order: *henna*, *arh*, *ftz*, s*urf*, *TH*, *SNF4* and *DnaJ*. In Chr15, the markers are located as follows: *RpS23*, *RpL5*, *RpL8*, *RpL10A* and *Ace 1*.

Table 4: Average positions of the markers in ChrZ$_1$. N represents the number of measurements.

Marker	***henna***	***arh***	***ftz***	***surf***	***TH***	***SNF4***	***DnaJ***
N	24	10	8	8	6	6	10
Relative position (mean)	0.039115	0.048356	0.140189	0.163113	0.595837	0.802514	0.949359
± SE	0.015002	0.021855	0.021197	0.021191	0.037464	0.036484	0.014508

Table 5: Average positions of the markers in Chr15.

Marker	***RpS23***	***RpL5***	***RpL8***	***RpL10A***	***Ace1***
N	25	12	13	12	13
Relative position (mean)	0.117286	0.298542	0.485748	0.776554	0.821699
± SE	0.02578	0.023282	0.033233	0.018747	0.030922

Mean positions of the markers were statistically compared using unpaired two-tailed t-test with Holm-Sidak correction. Most of the markers' positions differed at 0.01% level of significance. The positions of *henna* and *arh* proved different at 5% level of significance and the positions of *ftz* and *surf* tested different at 0.1% level of significance (Table 6 and Table 7).

Table 6: Statistical significance of distances between positions of the markers in $ChrZ_1$. The values represent the mean distance of the two markers and SE of the mean distance.

x	***henna***	***arh***	***ftz***	***surf***	***TH***	***SNF4***	***DnaJ***
henna	x	P= 0.0155215	P<0.0001	P<0.0001	P<0.0001	P<0.0001	P<0.0001
arh	0.0092 ± 0.0038	x	P<0.0001	P<0.0001	P<0.0001	P<0.0001	P<0.0001
ftz	0.1011 ± 0.0039	0.0918 ± 0.0059	x	P= 0.00049811	P<0.0001	P<0.0001	P<0.0001
surf	0.1240 ± 0.0039	0.1148 ± 0.0059	0.0229 ± 0.0061	x	P<0.0001	P<0.0001	P<0.0001
TH	0.5567 ± 0.0056	0.5475 ± 0.0085	0.4556 ± 0.0091	0.4327 ± 0.0091	x	P<0.0001	P<0.0001
SNF4	0.7634 ± 0.0055	0.7542 ± 0.0084	0.6623 ± 0.0090	0.6394 ± 0.0090	0.2067 ± 0.0123	x	P<0.0001
DnaJ	0.9102 ± 0.0032	0.9010 ± 0.0048	0.8092 ± 0.0049	0.7862 ± 0.0049	0.3535 ± 0.0076	0.1468 ± 0.0075	x

Table 7: Statistical significance of distances between positions of the markers in Chr15. The values represent the mean distance of the two markers and SE of the mean distance..

x	***RpS23***	***RpL5***	***RpL8***	***RpL10A***	***Ace1***
RpS23	x	*P*<0.0001	*P*<0.0001	*P*<0.0001	*P*<0.0001
RpL5	0.1813 ± 0.0051	x	*P*<0.0001	*P*<0.0001	*P*<0.0001
RpL8	0.3685 ± 0.0056	0.1872 ± 0.0067	x	*P*<0.0001	*P*<0.0001
RpL10A	0.6593 ± 0.0048	0.4780 ± 0.0050	0.2908 ± 0.0063	x	*P*<0.0001
Ace1	0.7044 ± 0.0055	0.5232 ± 0.0064	0.3360 ± 0.0073	0.0451 ± 0.0060	x

The average relative positions of the markers were depicted in two gene-based chromosome maps (Figure 9a for $YeChrZ_1$ and Figure 9b for YeChr15). In addition, the comparison was made between *Y. evonymella* and *P. xylostella* and *B. mori*.

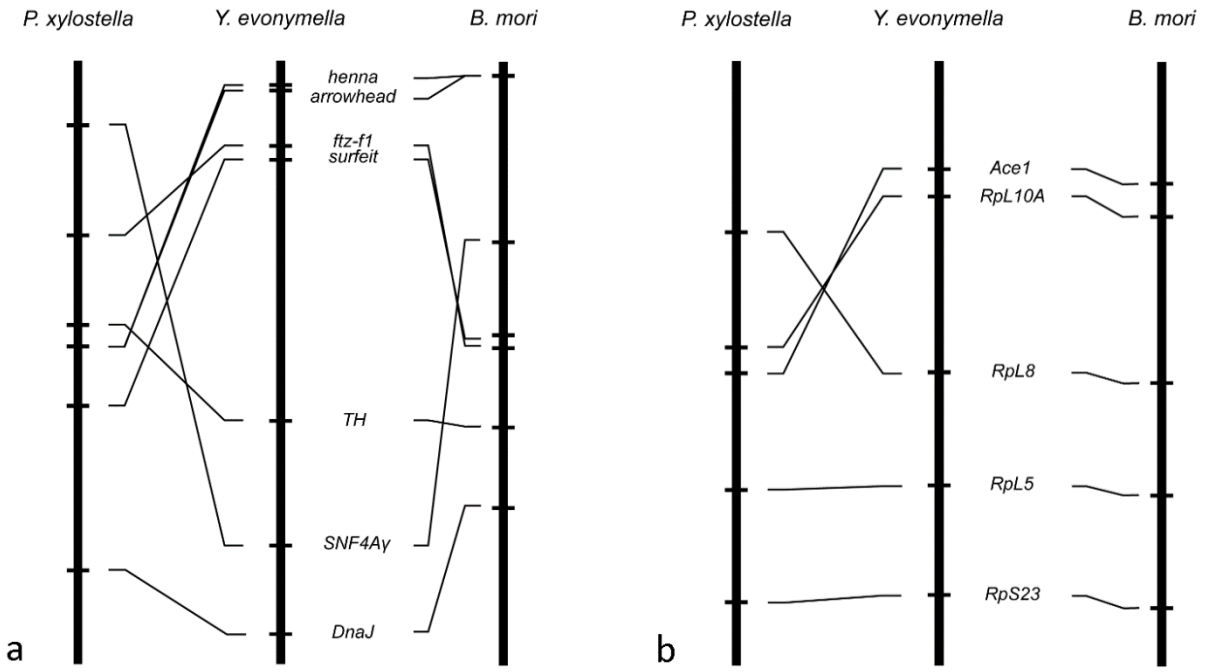

Figure 9: Comparison of synteny of mapped markers between *Yponomeuta evonymella* (this study) and *Plutella xylostella* and *Bombyx mori* (Whiteford and Darby, unpublished data and KAIKObase, respectively). **(a)** Chromosome Z_1 of Yponomeuta evonymella corresponding to ancestral Z chromosome shared with both *P. xylostella* and *B. mori* and **(b)** chromosome 15.

5 DISCUSSION

The ultimate goal of our laboratory is to acquire the *Y. cagnagella* genome sequence. In order to evaluate whether the newest and yet unpublished *Plutella xylostella* genome sequence could be used as the suitable reference in reference-assissted genome assembly in *Yponomeuta* spp., Unfortunately, the cross-hybridization experiments, i.e. fluorescence *in situ* hybridizations of available *Y. evonymella* BACs on chromosome preparations of *Y. cagnagella*, conducted by Provazníková (unpublished data) were unsuccessful. Therefore, we decided to test the synteny of genes on *Y. evonymella* chromosome preparations instead.

Sixteen *P. xylostella* gene orthologs were chosen as markers (Table 1). Primers were designed in the respective exons identified in the available *Y. evonymella* trascriptome and used for *Y. evonymella* BAC library screening (Table 2). BAC clones were identified for seven $ChrZ_1$ markers and five Chr15 markers (Table 2). Twelve markers in total were mapped to *Y. evonymella* male and female pachytene nuclei using fluorescence *in situ* hybridization of BAC probes (Figure 8). The resulting physical maps were compared to *P. xylostella* and *B. mori* ChrZ and Chr15 (Figure 9).

Lepidopteran sex chromosomes are prone to rearrangements and $YeChrZ_1$ proved to be no exception. All markers for $YeChrZ_1$ mapped to a single bivalent in male pachytene nuclei (Figure 8a-c) thus confirming conserved synteny of the Z-linked genes between *Y. evonymella* and both reference genomes of *B. mori* and *P. xylostella* (Figure 9a). However, the gene order was not conserved between *Y. evonymella* and *B. mori*, which suggests that many rearrangements took place since their split more than 120 My ago (Kawahara et al. 2019). The same was revealed by comparison between *Y. evonymella* and *P. xylostella*. Although *P. xylostella* represents the sister family to the family Yponomeutidae including the genus *Yponomeuta*, the two lineages have diverged more than 80 My ago (Kawahara et al. 2019). Given the complexity of changes in their gene order, it is not possible to reconstruct the chain of events differentiating the Z chromosomes of the compared species. However, it is reasonable to assume that multiple inversions were involved (cf Van't Hof et al. 2013).

The lepidopteran genome is characterized by two seemingly opposing attributes. The overall synteny of genes is well conserved in Lepidoptera conferring a high stability to the genome architecture as a whole. However, the lepidopteran karyotype shows in some cases a high frequency of rearrangements, which corresponds to the cytogenetic characteristics of the lepidopteran chromosomes such as holokinetic organization (Hill et al. 2019). Consequently, the lepidopteran chromosomes can be considered rather prone to balanced rearrangements and

simultaneously, rather resistant to otherwise deleterious rearrangements. These factors act together with the features typical for sex chromosomes in the lepidopteran Z. The rate of recombination in Z is considerably lowered due to its hemizygous state in females, which enables accumulation of repetitive sequences and consequently, ectopic recombination resulting in rearrangements and faster evolution. Van't Hof at al. (2013) constructed a linkage map for the peppered moth (*Biston betularia*) and compared the synteny and order of its genes in autosomes as well as in the chromosome Z to orthologs in *B. mori*. The order of the Z-linked genes suggested extensive intrachromosomal rearrangements compared to overall stable autosomes. However, the content of genes in the Z chromosome remained largely conserved. These observations correspond to the results of this study presented in Figure 9. It can be seen that the order of Z-linked genes between *Y. evonymella*, *P. xylostella* and *B. mori* is by no means colinear, whereas only one conspicuous rearrangement between YeChr15 and PxChr15 (*P. xylostella* Chr15) was detected. The extensively rearranged order of the mapped Z-linked genes presented in Figure 9a points to the special role sex chromosomes and their rearrangements play in evolution (Van't Hof et al. 2013). Especially the Z chromosome inversions could be important in the formation of new species as they retain favorable haplotypes and at the same time, genetic incompatibilities may be accumulated in them (Faria and Navarro 2010; Van't Hof et al. 2013).

The evolution of genes in chromosome Z is subject to the so-called Faster-Z effect (Mank et al. 2010). Novel mutations of in Z-linked genes are directly exposed to selection in the lepidopteran heterozygous females, since they possess only one Z chromosome. Therefore, the rate of adaptive changes in Z-linked genes can be accelerated in comparison with autosome-linked genes. However, it was proposed that the strength of the Faster-Z evolution depends, among other factors, on the effective population size (N_e) of chromosome Z. In a population characterized by equal numbers of breeding males and females and random variation in offspring, the N_e of Z chromosome is equal to ¾ of the N_e of an autosome. However, if the reproductive success of males is lower compared to the reproductive success of females (in a hypothetical example, if almost every female manages to reproduce but only 1 out of 10 males is reproductively successful), the N_e of chromosome Z drops under the ¾ of the N_e of an autosome and the Faster-Z effect in such a population is supposedly very strong due to genetic drift (Mank et al. 2010). The high frequency of rearrangements observed in the *Y. evonymella* Z chromosome (Figure 9a) thus might have been influenced also by a strong Faster-Z effect, since chromosome rearrangements can be fixed by genetic drift.

On the contrary, no or little changes were observed between markers for Chr15 between *Y. evonymella* and *B. mori* and *P. xylostella*, respectively (Figure 9b). The results support synteny of all markers examined as these mapped to a single bivalent in *Y. evonymella*. While gene order and surprisingly even relative position of loci were fully conserved between *Y. evonymella* and *B. mori*, the order of three markers, namely *Ace1*, *RpL10A*, and *RpL8*, differed between *Y. evonymella* and *P. xylostella*. The most parsimonious explanation for the observed change is a single inversion encompassing the corresponding region.

Synteny of genes in Chr15 was investigated both for the purpose of comparison of chromosome Z to an autosome and also because putative sugar receptors have been localized in Chr15 in *B. mori* (Roessingh et al. 1999). Since *B. mori* represents an outgroup to *Y. evonymella* and *P. xylostella*, the gene order shared between YeCh15 and *B. mori* Chr15 suggests that it is the ancestral state. The rearrangement between YeChr15 and PxChr15, most likely an inversion, probably occurred independently in *P. xylostella* and does not relate to the host shift of the ancestor of *Y. evonymella*. However, the physical map of YeChr15 might contribute to the identification of clones bearing the putative sugar receptor genes and their sequencing and analysis using e.g. chromosome walking.

The *P. xylostella*-assissted genome assembly in *Yponomeuta* species should be approached with caution because of the altered architecture shown in this study. The results of BAC-FISH mapping confirm rearrangements of both Z chromosome and the representative autosome, Chr15, between the two species. In spite of the close taxonomic relationship between the bird-cherry ermine and the diamondback moth, their lineages are separated by considerable evolutionary distance (>80 My) and PxChrZ cannot be used as reference in chromosome level assembly of $YeChrZ_1$. The use of *P. xylostella* autosomes as reference might represent less risk. However, detailed cytogenetic analysis should back up the bioinformatics.

In conclusion, the results of this study indicate that synteny of genes in reference-assissted genome assemblies encompassing relatively closely related species cannot be taken for granted and that the ‘chromosomics’ approach (Deakin et al. 2019), i.e. the conjunction of genome sequencing, cytogenetics, and cell biology, should be implemented in the overall endeavor to assembly as many genomes as possible (G10KCOS 2009).

6 CONCLUSION

The aim of this thesis was to create a physical map of the bird-cherry ermine (*Yponomeuta evonymella*) chromosome Z_1 (YeChrZ_1) and chromosome 15 (YeChr15) using BAC-FISH mapping and to compare synteny of genes with the *Plutella xylostella* and *Bombyx mori* genome sequences. Nine markers, i.e. single-copy genes, namely *henna*, *arh*, *ftz*, *surf*, *rdl*, *TH*, *SNF4*, *DnaJ* and *clock* were chosen according to previous mapping in *P. xylostella* for YeChrZ_1, and seven markers, namely *RpS23*, *RpL7A*, *RpL5*, *RpS5*, *RpL8*, *RpL10A* and *Ace 1*, were chosen for YeCh15. The comparative map for YeChrZ_1 showed that the gene order was not conserved between the three species of interest and that multiple but undefined rearrangements most probably occurred. The comparative map for YeCh15 showed conserved synteny of genes between the three species. However, one rearrangement was detected between *P. xylostella* and *Y. evonymella* encompassing *Ace1*, *RpL10A* and *RpL8*, most probably an inversion. These observations are in accordance with the results of the synteny analysis between the peppered moth (*Biston betularia*) and *B. mori* conducted by Van't Hof et. al (2013) and support the hypothesis that lepidopteran sex chromosomes undergo faster evolution and can accumulate more rearrangements than lepidopteran autosomes. The ultimate goal of our laboratory is to acquire *Yponomeuta* genome sequence. The synteny of genes between *Y. evonymella* and *P. xylostella* was tested in order to determine whether the *P. xylostella* yet unpublished genome could be used as a reference in reference-assissted genome assembly of *Yponomeuta*. The gene order comparison made in this study suggests that PxChrZ should not be used because of the high rearrangement frequency between PxChrZ and YeChrZ_1. As for chromosome 15, i.e. autosomes, synteny of genes could be conserved sufficiently but cytogenetic data should accompany the bioinformatics because of possible rearrangements such as the one identified in this study.

7 REFERENCES

Ahola, Virpi et al. 2014. "The Glanville Fritillary Genome Retains an Ancient Karyotype and Reveals Selective Chromosomal Fusions in Lepidoptera." *Nature Communications* 5: 4737.

Altschul, Stephen F. et al. 1990. "Basic Local Alignment Search Tool." *Journal of Molecular Biology* 215(3): 403–10.

Bachtrog, Doris et al. 2014. "Sex Determination: Why So Many Ways of Doing It?" *PLoS Biology* 12(7): e1001899.

Bellott, Daniel W. et al. 2010. "Convergent Evolution of Chicken Z and Human X Chromosomes by Expansion and Gene Acquisition." *Nature* 466(7306): 612–16.

Blackmon, Heath, Laura Ross, and Doris Bachtrog. 2017. "Sex Determination, Sex Chromosomes, and Karyotype Evolution in Insects." *Journal of Heredity* 108(1): 78–93.

Brown, Kirsten M, Lisa M Burk, Loren M Henagan, and Mohamed A F Noor. 2004. "A Test of the Chromosomal Rearrangement Model of Speciation in *Drosophila pseudoobscura*." *Evolution* 58(8): 1856–60.

Bureš, Petr, František Zedek, and Michaela Marková. 2013. "Holocentric Chromosomes." In *Plant Genome Diversity Volume 2*, eds. Ilia J. Leitch, Johann Greilhuber, Johathan Wendel, and Jaroslav Doležel. Springer Verlag, 187–208.

Charlesworth, D., and B. Charlesworth. 1980. "Sex Differences in Fitness and Selection for Centric Fusions between Sex-Chromosomes and Autosomes." *Genetical Research* 35(2): 205–14.

Dalíková, Martina et al. 2017. "New Insights into the Evolution of the W Chromosome in Lepidoptera." *The Journal of Heredity* 108(7): 709–19.

Deakin, Janine E. et al. 2019. "Chromosomics: Bridging the Gap between Genomes and Chromosomes." *Genes* 10(8): E627.

Faria, Rui, and Arcadi Navarro. 2010. "Chromosomal Speciation Revisited: Rearranging Theory with Pieces of Evidence." *Trends in Ecology & Evolution* 25(11): 660–69. http://dx.doi.org/10.1016/j.tree.2010.07.008.

G10KCOS. 2009. "Genome 10K: A Proposal to Obtain Whole-Genome Sequence for 10 000 Vertebrate Species." *Journal of Heredity* 100(6): 659–74.

Gamble, Tony et al. 2017. "The Discovery of XY Sex Chromosomes in a *Boa* and *Python*." *Current Biology* 27(14): 2148–53. http://dx.doi.org/10.1016/j.cub.2017.06.010.

Goldsmith, Marian R., and František Marec. 2010. *Molecular Biology and Genetics of the*

Lepidoptera. 1st ed. Boca Raton, FL, USA: CRC Press.

Guo, Huizhen et al. 2017. "Expression Map of a Complete Set of Gustatory Receptor Genes in Chemosensory Organs of *Bombyx mori*." *Insect Biochemistry and Molecular Biology* 82: 74–82.

Hill, Jason et al. 2019. "Unprecedented Reorganization of Holocentric Chromosomes Provides Insights into the Enigma of Lepidopteran Chromosome Evolution." *Science Advances* 5(6): eaau3648.

Janes, Daniel E. et al. 2011. "Sex Chromosome Evolution in Amniotes: Applications for Bacterial Artificial Chromosome Libraries." *Journal of Biomedicine and Biotechnology* 2011: 132975.

Kandul, Nikolai P, Vladimir A Lukhtanov, and Naomi E Pierce. 2007. "Karyotypic Diversity and Speciation in *Agrodiaetus* Butterflies." *Evolution* 61(3): 546–59.

Kato, A., P. S. Albert, J. M. Vega, and J. A. Birchler. 2006. "Sensitive Fluorescence *in situ* Hybridization Signal Detection in Maize Using Directly Labeled Probes Produced by High Concentration DNA Polymerase Nick Translation." *Biotechnic and Histochemistry* 81(2–3): 71–78.

Kawahara, Akito Y. et al. 2019. "Phylogenomics Reveals the Evolutionary Timing and Pattern of Butterflies and Moths." *PNAS* 116(45): 22657–63.

Lockwood, A. P. M. 1961. "'Ringer' Solutions and Some Notes on the Physiological Basis of Their Ionic Composition." *Comparative Biochemistry and Physiology* 2: 241–89.

Lukhtanov, Vladimir A. 2015. "The Blue Butterfly *Polyommatus* (*Plebicula*) *atlanticus* (Lepidoptera, Lycaenidae) Holds the Record of the Highest Number of Chromosomes in the Non-Polyploid Eukaryotic Organisms." *Comparative Cytogenetics* 9(4): 683–90.

———. 2018. "Versatility of Multivalent Orientation, Inverted Meiosis, and Rescued Fitness in Holocentric Chromosomal Hybrids." *PNAS* 115(41): E9610–19.

Mank, Judith E., Beatriz Vicoso, Sofia Berlin, and Brian Charlesworth. 2010. "Effective Population Size and the Faster-X Effect: Empirical Results and Their Interpretation." *Evolution* 64(3): 663–74.

Marec, František, and Walther Traut. 1993. "Synaptonemal Complexes in Female and Male Meiotic Prophase of *Ephestia kuehniella* (Lepidoptera)." *Heredity* 71: 394–404.

Menken, S. B. J., W. M. Herrebout, and J. T. Wiebes. 1992. "Small Ermine Moths (*Yponomeuta*): Their Host Relations and Evolution." *Annual Review of Entomology* 37: 41–66.

Menken, S. B. J., and P. Roessingh. 1998. "Evolution of Insect-Plant Associations: Sensory

Perception and Receptor Modifications Direct Food Specialization and Host Shifts in Phytophagous Insects." In *Endless Forms: Species and Speciation*, eds. D. Howard and S. H. Berlocher. Oxford University Press, 145–56.

Mongue, Andrew J., Petr Nguyen, Anna Voleníková, and James R. Walters. 2017. "Neo-Sex Chromosomes in the Monarch Butterfly, *Danaus plexippus*." *G3* 7(10): 3281–94.

Mutanen, Marko, Niklas Wahlberg, and Lauri Kaila. 2010. "Comprehensive Gene and Taxon Coverage Elucidates Radiation Patterns in Moths and Butterflies." *Proceedings. Biological Sciences.* 277(1695): 2839–48.

Nguyen, Petr et al. 2013. "Neo-Sex Chromosomes and Adaptive Potential in Tortricid Pests." *PNAS* 110(17): 6931–36.

Nilsson, Nils-Otto, Christer Löfstedt, and Lars Dävring. 1988. "Unusual Sex Chromosome Inheritance in Six Species of Small Ermine Moths (*Yponomeuta*, Yponomeutidae, Lepidoptera)." *Hereditas* 108: 259–65.

Noor, Mohamed A F, Katherine L Grams, Lisa A Bertucci, and Jane Reiland. 2001. "Chromosomal Inversions and the Reproductive Isolation of Species." *PNAS* 98(21): 12084–88.

Pennell, Matthew W. et al. 2015. "Y Fuse? Sex Chromosome Fusions in Fishes and Reptiles." *PLOS Genetics* 11(5): e1005237.

De Prins, Jurate, and Kazuo Saitoh. 2003. "Karyology and Sex Determination." In *Handbook of Zoology*, eds. Niels P. Kristensen and Maximilian Fischer. De Gruyter, 449–68.

Roessingh, P., K. H. Hora, J. J. A. van Loon, and S. B. J. Menken. 1999. "Evolution of Gustatory Sensitivity in *Yponomeuta* Caterpillars: Sensitivity to the Stereo-Isomers Dulcitol and Sorbitol Is Localised in a Single Sensory Cell." *Journal of Comparative Physiology A* 184(1): 119–26.

Sahara, Ken, Atsuo Yoshido, and Walther Traut. 2012. "Sex Chromosome Evolution in Moths and Butterflies." *Chromosome Research* 20(1): 83–94.

Suetsugu, Yoshitaka et al. 2013. "Large Scale Full-Length cDNA Sequencing Reveals a Unique Genomic Landscape in a Lepidopteran Model Insect, *Bombyx mori*." *G3* 3(9): 1481–92.

Tamazian, Gaik et al. 2016. "Chromosomer: A Reference-Based Genome Arrangement Tool for Producing Draft Chromosome Sequences." *GigaScience* 5(1): 38. http://dx.doi.org/10.1186/s13742-016-0141-6.

Traut, Walther et al. 2017. "Karyotypes versus Genomes: The Nymphalid Butterflies *Melitaea cinxia*, *Danaus plexippus*, and *D. chrysippus*." *Cytogenetic and Genome*

Research 153(1): 46–53.

Turner, Hubert, Niek Lieshout, Wil E. Van Ginkel, and S. B. J. Menken. 2010. "Molecular Phylogeny of the Small Ermine Moth Genus *Yponomeuta* (Lepidoptera , Yponomeutidae) in the Palaearctic." *PLoS One* 5(3): e9933.

Van't Hof, A. E. et al. 2013. "Linkage Map of the Peppered Moth, *Biston betularia* (Lepidoptera, Geometridae): A Model of Industrial Melanism." *Heredity* 110(3): 283–95.

Vershinina, Alisa O., and Vladimir A. Lukhtanov. 2017. "Evolutionary Mechanisms of Runaway Chromosome Number Change in *Agrodiaetus* Butterflies." *Scientific Reports* 7: 8199. http://dx.doi.org/10.1038/s41598-017-08525-6.

Vicoso, Beatriz et al. 2013. "Comparative Sex Chromosome Genomics in Snakes: Differentiation, Evolutionary Strata, and Lack of Global Dosage Compensation." *PLOS Biology* 11(8): e1001643.

Wanner, K. W., and H. M. Robertson. 2008. "The Gustatory Receptor Family in the Silkworm Moth *Bombyx mori* Is Characterized by a Large Expansion of a Single Lineage of Putative Bitter Receptors." *Insect Molecular Biology* 17(6): 621–29.

Wellenreuther, Maren, and Louis Bernatchez. 2018. "Eco-Evolutionary Genomics of Chromosomal Inversions." *Trends in Ecology & Evolution* 33(6): 427–40. http://dx.doi.org/10.1016/j.tree.2018.04.002.

Winnepenninckx, Birgitta, Thierry Backeljau, and Rupert De Wachter. 1993. "Extraction of High Molecular Weight DNA from Molluscs." *Trends in Genetics* 9(12): 407 (Technical Tips).

Yasukochi, Yuji et al. 2009. "Extensive Conserved Synteny of Genes between the Karyotypes of *Manduca sexta* and *Bombyx mori* Revealed by BAC-FISH Mapping." *PLoS ONE* 4(10): e7465.

———. 2011. "Isolation of BAC Clones Containing Conserved Genes from Libraries of Three Distantly Related Moths: A Useful Resource for Comparative Genomics of Lepidoptera." *Journal of Biomedicine and Biotechnology* epub: Article ID 165894.

Zhang, Zhi-Qiang. 2011. "Animal Biodiversity: An Outline of Higher-Level Classification and Survey of Taxonomic Richness." *Zootaxa* 3148: 1–237.