The University of South Bohemia
Faculty of Science

Isolation and characterization of highly repetitive fraction of codling moth, *Cydia pomonella*

Master Thesis

Bc. Pavlína Věchtová

supervisor: RNDr. Magda Vítková, PhD.

České Budějovice
2011

Věchtová, P., 2011: Isolation and characterization of highly repetitive fraction of codling moth, *Cydia pomonella.* Master Thesis, in English – 57 pp., Faculty of Science, The University of South Bohemia, České Budějovice, Czech Republic.

Annotation:
Repetitive DNA comprises substantial part of the eukaryotic genome. "Junk DNA", as it was originally understood at the beginning of its discovery has attracted a lot of attention lately due to many studies proving its functional perspectives. Analysis of its dynamics, characteristics and distribution has been widely studied in organisms with monocentric chromosomes. Holokinetic system, however, was left behind in these efforts and whole image of repetitive DNA distribution and dynamics in this system remains to be elucidated.
In this thesis various approaches were used to isolate and characterise repetitive DNA in the genome of the codling moth, *Cydia pomonella*. Satellite DNA CPSAT-1 was successfully isolated, characterised with Dot blot and Southern blot and mapped with FISH in the genome of *C. pomonella*. 17 microsatellite probes were used to localize microsatellite arrays in the genome of *C. pomonella*. Method of microsatellite FISH revealed distribution of all tested microsatellites in *C. pomonella* complement.

This thesis is part of the project funded by grant KJB501410901 of the Grant Agency of the Academy of Sciences of the Czech Republic.

Prohlašuji, že svoji diplomovou práci jsem vypracoval/a samostatně pouze s použitím pramenů a literatury uvedených v seznamu citované literatury.
Prohlašuji, že v souladu s § 47b zákona č. 111/1998 Sb. v platném znění souhlasím se zveřejněním své diplomové, a to v nezkrácené podobě elektronickou cestou ve veřejně přístupné části databáze STAG provozované Jihočeskou univerzitou v Českých Budějovicích na jejích internetových stránkách, a to se zachováním mého autorského práva k odevzdanému textu této kvalifikační práce. Souhlasím dále s tím, aby toutéž elektronickou cestou byly v souladu s uvedeným ustanovením zákona č. 111/1998 Sb. zveřejněny posudky školitele a oponentů práce i záznam o průběhu a výsledku obhajoby kvalifikační práce. Rovněž souhlasím s porovnáním textu mé kvalifikační práce s databází kvalifikačních prací Theses.cz provozovanou Národním registrem vysokoškolských kvalifikačních prací a systémem na odhalování plagiátů.

V Českých Budějovicích, 29. 4. 2011 .....

Acknowledgement

Hereby I would like to thank to my family for support and patience during my master studies. My great words of thanks belong to Magda Vítková for her perfect work as a tutor. I appreciate her attentive supervising and her enthusiasm for the whole project, which really meant a great deal for me and helped me in getting through the whole problematics. Her support and patience especially during final works on the thesis were priceless. I am very grateful to Eduard Kejnovský from Laboratory of Plant Developmental Genetics in Brno for providing me selflessly with microsatellite probes. I also thank to Franta Marec, the head of the Laboratory of Molecular Cytogenetics for accepting me to his lab family and providing me with place and resources to pursue my project. He was also willing to give me immediate and selfless expert consultations whenever I needed. I want to thank to all members of the laboratory of Molecular Cytogenetics for their advices and creating nice atmosphere in the lab. My special thank belongs to Marťa Daliková who advised me and virtually led me through the substantial part of my thesis concerning isolation of satellite DNA. For critical reading I would like to thank to both Franta Marec and Radka Frydrychová-Čapková. They both were very helpful not only in revising the content of the thesis but also in correcting English grammar and stylistics. I thank to Magda Vítková, Petr Nguyen, Valeriya Zabelina and Michal Zapletal for providing me with insect used in this work.

Contents

1. Introduction ... 1
   1.1. Repetitive DNA ... 1
   1.2. Transposable elements (TEs) ... 2
   1.3. Tandem repeats ... 3
      1.3.1. Satellites ... 3
      1.3.2. Minisatellites ... 4
      1.3.3. Microsatellites ... 5
   1.4. Negative effect of repetitive DNA presence and activity on the host genome ... 7
   1.5. Repetitive DNA in human service ... 9
   1.6. Repetitive DNA in lepidopteran genome ... 10
   1.7. Holokinetic chromosomes ... 11
   1.8. *Cydia pomonella* ... 12
2. Aims of the thesis ... 13
3. Material and Methods ... 14
   3.1. Insects ... 14
   3.2. Genomic self-priming PCR (GSP-PCR) ... 15
   3.3. Chromosome preparations ... 16
   3.4. DNA extraction ... 16
   3.5. Visualization of tandem repeats via restriction analysis ... 17
   3.6. Cloning and sequencing ... 18
   3.7. Probes for hybridization techniques ... 19
   3.8. Southern hybridization ... 19
   3.9. Dot-blot hybridization ... 20
   3.10. Fluorescence in situ hybridization (FISH) ... 20
      3.10.1. FISH with biotin/digoxigenin labeled probe ... 20
      3.10.2. FISH with microsatellite probes ... 22
4. Results ... 23
   4.1. GSP-PCR ... 23
   4.2. Isolation and characterisation of satellite DNA ... 25
      4.2.1. Isolation of satellite DNA ... 25
      4.2.2. Sequencing and characterization of satellite DNA with Southern blot ... 27
      4.2.3. Distribution of CPSAT-1 in *C. pomonella* chromosomes ... 28
      4.2.4. Dot blot analysis ... 28
   4.3. Microsatellite FISH ... 30
5. Discussion ... 35
   5.1. GSP-PCR ... 35
   5.2. Distribution of satellite DNA in the genome of *C. pomonella* ... 35
   5.3. Microsatellite distribution in the genome of *C. pomonella* ... 36
6. Summary ... 40
7. References ... 41

1. Introduction

1.1. Repetitive DNA

Genomes of prokaryotic organisms consist predominantly of low-copy number DNA sequences. Genome sizes of different prokaryotic species vary by less than an order of magnitude (Kingsbury, 1969), and there is almost perfect correlation between genome size and the total number of genes. In sharp contrast to prokaryotes, the eukaryotic genomes that are generally much larger than their prokaryotic counterparts show enormous interspecific differences in the size. Also genome sizes of very similar (and therefore similarly complex) organisms may be very different. Among all organisms, amoebae (*Polychaos dubium*) has the largest known genome counting up to 670 000 Mbp (McGrath and Katz, 2004), while the smallest known non-viral genome belongs to bacteria Carsonella rudi with its 160 kb large genome (Than, K., 2006). In eukaryotes, the differences in the genome size are given by large-scale fluctuations in the amount of "junk", noncoding, repetitive DNA. The genome size variation that does not correlate with the complexity of organism is termed as C-value paradox (Thomas 1971). Identification of function and origin of repetitive DNA has always been the primary mission of many geneticists since we have known about the existence of the repetitive sequences. The C_ot analysis (i.e. DNA reassociation kinetics) performed in many organisms demonstrated that repetitive fraction may represent even more than half of the genome (Britten and Kohne, 1968). In the plant genome, the share of repetitive sequences can reach up to 95% (Flavell et al., 1994). Trumpet lily, *Lilium longiflorum,* for instance, has the largest known plant genome counting up to 90 000 Mbp where most of the DNA content is caused by high repetitive DNA content (Joseph et al., 1990) and other plants demonstrate similar tendency.

Repetitive DNA was originally considered as a purely selfish and parasitic DNA by most geneticists (Orgel and Crick, 1980). Large portions of eukaryotic genome are composed of the non-coding DNA which expands in parasitic-like manner. Theory of selfish DNA was formulated based on these observations and it explained the relationship between host and repetitive DNA quite sufficiently. However, accumulating evidence doubting this theory as the only possible explanation of this relationship were emerging with a development of molecular techniques and sequencing projects. What is more, hypotheses describing evolution of this relationship were formulated. At the beginning of the repetitive DNA expansion this relationship could have been clearly selfish and in the course of evolution this relationship could develop in symbiotic manner (Brosius, 1999a;

Kidwell and Lisch, 2000; Kidwell and Lisch, 2001) Character and distribution of repetitive DNA create two distinctive groups of repeats - transposable elements and tandem repeats.

1.2. Transposable elements (TEs)

Transposones are semiautonomous elements that can move themselves to new positions within the host genome. Transposable elements are divided into two classes according to their mechanism of transposition, which can be either "cut and paste" (class of DNA transposons) or "copy and paste" (class of retrotransposons) (Berg & Howe, 1989). Character of their activity radically influences specific genome size and constitution of the host genome. It also provides organisms with a potential source of variability, thereby possibility to evolve from the long-term point of view (Kidwell and Lisch, 2001).

Transposones and transposone-derived elements (TE) appear to be abundant part of eukaryotic genomes. Their sequences comprise up to 44% of the human genome (Smit, 1996), 39% in mice (Waterston, et al., 2002), cca 50% in maize (SanMiguel et al., 1996), cca 35% in the silkworm *Bombyx mori* (Futahashi et al., 2008), cca 10% in *Drosophila melanogaster* (Pimpinelli et al., 1995), and 52% in the recently sequenced genome of opossum (Gentles et al., 2007). The data indicates that transposons make up a large fraction of the C-value of eukaryotic cell.

Activity of mobile elements in the genome is the main source of genetic variation giving space to evolutionary processes. In the human genome, however, the activity of DNA transposons was lost based on accumulation of mutations during evolution, and out of mobile elements, only retrotransposons are truly active in humans nowadays (International Human Genome Sequence Consortium, 2001). There is a whole range of phenotypic changes caused by TE's activity. It is noteworthy that TE's activity creates much broader spectrum of mutations than any other mutator (Kidwell and Lisch, 2001). One of the common features of TE's activity is production of double-strand breakage of DNA strand causing chromosomal rearrangements such as translocations or inversions, thereby contributing to faster genome evolution (Kidwell and Lisch, 2001). Mobile element activity causes insertional mutation by transduction of structurally or regulatory important sequences or by leaving 5' or 3' flanking sequence after imprecise excision. Such activity results in production of chimeric mRNA or changes in expression pattern of the respective gene. TEs can greatly contribute to gene evolution by their ability to duplicate certain parts of chromosomes by means of ectopic recombination or imprecise excision. The

stated behaviour causes mostly negative phenotype manifesting as various diseases mentioned later in this work.

"Molecular domestication" is another phenomenon observed in connection with TE's activity. The definition says that molecular domestication involves "the transition of a genomic parasite to a stable integrated gene useful to the host over evolutionary time" (Miller et al., 1997). There are plenty of conspicuous examples of molecular domestication discovered so far. The most well-known of all is probably the case of *D. melanogaster* and its system of maintaining the telomeres integrity (Traverse & Pardue, 1988) or mechanism of VDJ recombination employed in production of human immune system variability (Lewis and Wu, 1997).

In order to survive parasitic behaviour of TEs it was necessary for the genome to develop a system down-regulating or repressing TE's activity. Only recently it has been proposed that a cell might have developed a silencing mechanism based on process involving small RNA molecules, RNA interference, which has always been considered to be developed by the cell primarily as a tool for regulation of gene expression (Agrawal et al., 2003; Barlow 1993; Yoder et al., 1997). DNA methylation is a neat alternative of a gene regulation based on regulatory proteins interfering with initiation of transcription (Jones and Takai, 2001). Besides gene regulation, proteins involved in DNA methylation participate on gene silencing as well (Klose and Bird, 1999) and the origin and evolution of this mechanism seems to be the same as the one of RNAi (Jordan and Miller, 2008). TEs also developed mechanisms to downregulate their own activity in order to increase fitness of their host. The main mechanism of selfregulation lies in less frequent transposition (Robertson and Engels, 1989) as well as in selective insertion into transcriptionally less active locations of the genome or into pre-existing elements (San Miguel et al., 1996).

1.3. Tandem repeats

1.3.1. Satellites

Satellite DNA is a long, tandemly organized sequence motif with a repeat length of one to several thousand base pairs. The sequence of satellite DNA is rich in adenine and thymine, which gives it different density character in density gradient thereby satellite DNA forms there an extra, satellite band (Beridze, 1986). The iteration of the satellite units reaches up to several thousands of repeats. Distribution of satellite blocks is in concordance with localization of heterochromatic blocks (John, 1988) that can be easily visualized by C-band technique (Sumner, 1972; Jorge and Yasmineh; 1971, Charlesworth

et al., 1994). Most satellite DNA is localized to telomeric or centromeric regions of monocentric chromosomes (Charlesworth et al., 1994). While sequence of the satellite unit is well conserved across related taxa, the length of the repeat can be highly variable even in closely related species (Beritze, 1986).

According to recent findings, satellite DNA proves distinct conserved structural features proposing the involvement of some repeats in specific chromatin structures. Satellite DNA is proven to be an important structural unit in centromere (Dimitri et al., 2005). Henikoff et al. (2001) suggested a possible role of satellite unit in nucleosome formation when observing suspicious concordance of satellite unit distribution with distribution of nucleosomes and similar length of the repeated unit between organisms.

Moreover, other functions such are maintenance and spreading of silent chromatin, dosage compensation, programmed DNA elimination or RNAi mediated heterochromatin assembly or even post-transcriptional gene regulation through RNA-induced silencing complex (RISC) have been suggested for satellite DNA as well (reviewed in Palomeque and Lorite, 2008).

Satellite DNA was proven to be functional not only thanks to its structural features but also through its transcripts. Several studies managed to detect satellite DNA transcripts in vertebrates, invertebrates and plants. Transcripts demonstrate tissue specific and temporal expression, which may attribute them regulatory role. Their possible function still remains speculation (reviewed in Ugarkovic, 2005).

Satellite sequence units within a genome must necessarily face the similar mutational processes due to the similar character of their sequence. Various satellite DNAs present within one organism are subjected to gene conversion and unequal crossing-over being considered as a main tool for satellite unit diversification (Smith, 1976). The fact may contribute to speciation of the species through the evolution of the basal chromosome structures that satellite DNA forms thanks to its specific structural features recruited by the cell (Ugarkovic and Plohl, 2002). Satellite DNA sequence conservation is another indicator of selective constraint put on the sequence due to its involvement in establishment of the centromere (Csink and Henikoff, 1998).

1.3.2. Minisatellites

Minisatellites are tandemly repeated motives, when a length of repeated unit stretches from 9 up to 100 nucleotides. Their array length extends from 0.5 to 30 kbp (Jeffreys et al., 1985; Nakamura et al., 1987). Studies of minisatellite genomic distribution revealed predominantly terminal subtelomeric location. Unlike satellite DNA, minisatellites spread in

euchromatic regions of the chromosomes. Their occurrence is often reported to be upstream or downstream of genes, sometimes even within introns (reviewed in Ramel, 1997). Number of minisatellite repeat units varies largely within individuals. This phenomenon called variable number of tandem repeats (VNTR) is widely used for "DNA fingerprint analyses" being used in forensic medicine, for linkage studies in genetic analyses or for paternity determination (Ramel, 1997). Mechanisms causing this kind of variability within individuals proceed mainly by means of replication slippage, intramolecular recombination or gene conversion (Armour and Jeffreys, 1992).

1.3.3. Microsatellites

Microsatellites or simple sequence repeats (SSR) are 1-6 nucleotide tandem repeat units iterating mostly 10 to 100 times (Queller et al, 1993). Out of all the tandem repeats microsatellites are the most spread within a genome and show high level of length polymorphism (Li et al., 2002). Mechanisms of their amplification throughout the genome predetermine them to cause random dispersive pattern. Their origin is attributed predominantly to polymerase slippage during replication stage causing extension of pre-existing microsatellite.

The other minor mechanism is spreading of microsatellite repeats during retrotransposition as a part of the sequence of some transposable elements. It is supported by numerous studies proving frequent coincidence of transposable elements and microsatellites (Hoekstra et al., 1997; Ramsay et al., 1999; Akagi et al., 2001; Fagerberg et al., 2001; Temnykh et al., 2001; Wilder and Hollocher, 2001; Johnson et al., 2006).The fact that microsatellites cluster into families based on the similarities in the sequence of their flanking regions in some taxa accounts for this phenomenon as well (Meglécz et al., 2004). TEs therefore serve as a molecular transport for microsatellite distributing it through the genome (Li et al., 2002). Nadir et al. (1997) also suggest one possible scenario of microsatellite origin through 3'-extension of retrotranscripts, a mechanism similar to mRNA polyadenylation.

The distribution of microsatellites within the genome was proved to be nonrandom. All types of SSRs in selected organisms gather up predominantly in non-coding regions of the genome most probably due to selection against frame-shift mutations deleterious when occurring in coding regions of the genome (Metzgar et al., 2000).

SSR used to be considered as an evolutionarily neutral DNA. Only recently did lots of studies point out on the possible functional perspectives of SSRs. The occurrence of microsatellite tracts is the most often connected with causing negative phenotypes like

human trinucleotide expansion disorders. Like TEs, SSRs can interfere with various regulatory processes (Handa et al., 2003; Jasinka et al., 2003) including induction of methylation resulting in transcriptional silencing (Coffee et al., 1999).

When listing changes caused by SSRs there are several mutations causing beneficial phenotype to the host organism. In bacteria, for instance, SSRs have a key role in the generation and maintenance of the high level of phenotypic diversity required for successful inhabitation of new niches. Namely, the human pathogen *Haemophilus influenzae* parasiting on human red cells show variable expression pattern in the genes responsible for translation of hemoglobin-binding proteins. All three genes differ right in the length of tetrameric repeats (Ren et al., 1999) included in their sequences causing phase variation in expression (Morton et al., 1999). The same phase variation has been observed at *Neisseria meningitides* (Bayliss, Field & Moxton, 2001) and *Escherichia coli* (Torrez-Cruz and van der Woude, 2003).

Statistical linkage of SSR and variable phenotypic traits can be used for quantitative trait loci (QTL) studies (Darvasi and Pisante Shalom, 2002). It has been proposed by King et al. (1997) that SSRs may quantitatively modify phenotype of host organism based on their distribution. What is more the SSRs have been proven to be part of various control or coding sequences, therefore affecting function of the respective gene by exposing the gene by its presence to higher mutation rate and regulating its expression in quantitative manner (Li et al., 2002). Babich et al. (1999) when analyzing the distribution pattern of Alu sequences in the human genome found out that SSR containing high affinity binding sites for thyroid hormone, retinoic acid, and oestrogen receptors are often located in promoter regions of genes activated in response to hormonal signals. The presence of SSR within particular genes in the species across very diverse taxonomic groups implies that those may actively participate on important cellular regulatory processes. For example, the HLA-DRB1 gene contains a $(GT)_n(GA)_n$ SSR within intron 2 in almost all vertebrates (Schwaiger and Epplen, 1995). This SSR segment is able to bind the transcriptional factor CTCF (Arnold et al., 2000). Effects caused by SSRs mutations often appear to be tissue or locus specific (Cleary et al., 2002). For instance, when being present within flanking region of the certain gene, SSR can play a role as a trans-factor or cis-element altering the gene expression. Specific example of such an effect present Cleary and Pearson (2003) when they proved that $(CTG)_n$ expansion participating on the Myotonic dystrophy has flanking binding sites for Zn finger insulatory protein CTCF. In human, promoter of Growth inhibitory Factor/Metallothionen III contains a (CTG)25 repressor motif that prevents gene expression in non-neural tissue (Imagawa et al.,

1995). The same goes for human CD20 whose promoter contains a $(CCAT)_n$ repressor sequence. Loss or expansion of this motif causes over-expression thereby lymphoproliferative disease (Croager et al., 2000).

1.4. Negative effect of repetitive DNA presence and activity on the host genome

Seldom do transposon insertions cause a disease since most of such transposons have been selected against in the course of evolution. Out of 65 diseases caused by TE insertions recorded so far the most of these events have been attributed to L1, Alu and SVA elements as these are the most active and, therefore, the most numerous elements present in human genome nowadays. Frequently affected alleles are coagulation factor VIII or IX, BTK gene and dystrophin causing coagulopathies (Kazazian et al., 1988; Ostertag and Kazazian, 2001), immunodeficiency syndromes (Rohrer et al., 1999) and muscular dystrophies and cardiomyopathies, respectively. Majority of TE insertions remain unrevealed since their effect is either lethal to their host or they cause silent mutations. They may also be concealed by recessive character of the mutation. The latter case usually manifests in hemizygotic constitution of X-linked alleles in male carriers (Narita et al., 1993; Yoshida et al., 1998). Autosomal TE insertion induced mutations are not that prevalent since a loss of their function is concealed by the other unimpaired allele present in the diploid genome. If autosomal mutation manifests, it predominantly mimics phenotype of autosomal dominant disease. Examples include impaired NF1 tumor suppressor leading to clinical neurofibromatosis (Wallace et al., 1991) and often affected fibroblast growth factor receptor 2 (FGFR2) causing Apert syndrome (Oldridge et al., 1999; Bochukova et al., 1999).

Minisatellite stretches were also proven to cause several human genetic diseases. Noteworthy is minisatellite tract that is located 1000 bp downstream of the polyadenylation signal of the Ha-ras protooncogene locus. Some alleles of this locus are 3 times more coincidental in cancer patients than it is in case of common alleles (Capon et al., 1983; Kasperczyk et al, 1990; Krontiris, et al., 1993; Krontiris, 1995). Somewhat similar case was discovered when studying minisatellite tract of insulin gene. One allelic form with the shortest minisatellite tract present nearby the gene is the most often associated with doubled risk of type I diabetes mellitus (IDDM) incidence (Krontiris, 1995).

Microsatellite tracts are dynamic entities of the genome. Their length changes under the influence of many cell mechanisms such as DNA polymerase slippage during replication or ectopic recombination. Most of these processes cause microsatellite

expansion. Stability of microsatellite tracts within genome depends by large on performance of mismatch repair system of the host cell. The lowered performance or impaired function of the system was proven to cause increased microsatellite instability at *Saccharomyces* (Strand et al., 1993). The comparable case was recorded concerning the human colon cancer. A gene involved in HNPCC (Hereditory Non-Polyposis Colon Cancer) is associated with instability of AC microsatellite tract. The impaired gene shares homology with mismatch repair gene *MutS* in *E. coli* and *MSH2* in yeast (Fishel et al., 1993; Leach et al., 1993,). There are also other forms of cancer associated with microsatellite instability, e.g. lung or Barrett's esophageal cancer (Meltzer et al., 1994). Microsatellites can influence a gene function even directly where an expression of a functional gene is affected by microsatellite instability within or outside the coding sequence. Those diseases are characterized by the presence of unusually long trinucleotide microsatellite tracts within genes encoding proteins ensuring certain neurological functions (Bates and Lehrach, 1994; Ashley and Warren, 1995). Such an affected gene causes malfunction, improper localization or impaired structure of the translated protein. There are several well known cases of such microsatellite expansion causing human diseases. Expansion of CGG and CTG trinucleotide causes Fragile X syndrome and Myotonic dystrophy, respectively (Green, 1993). Huntington disease is caused by expansion of CAG microsatellite tract in the huntingtin gene (Andrew et al., 1994). Spinobulbar muscular atrophy is caused by polyglutamine expansion in the gene coding for androgene receptor (Chamberlain et al., 1994). Another disease directly associated to polyglutamine array extension is Spinocerebellar ataxia. There have been more than 25 types of spinocerebellar ataxia types recognized so far depending on the gene in which microsatellite expansion occurred (Bird, 1998). Recent studies have pointed out that other diseases like schizophrenia or bipolar disorder may be linked to trinucleotide tract expansion as well (O'Donovan et al., 1995). Extent and time of manifestation of the disease correlate with the extension of the microsatellite array. The fact that only trinucleotide repeats expansion causes diseases may be related to the phenomenon of frameshift mutation. Microsatellite expansions where other than tri or hexaplets are involved result in complete suppression of gene function causing immediate dead of its carrier.

1.5. Repetitive DNA in human service

With a development of current molecular techniques it is possible to construct synthetic retrotransposones and pursue directed mutagenesis (Han and Boeke, 2004; An et al., 2006).

Ever since the discovery of the DNA transposone P element in *D. melanogaster* in late 80s' scientists revealed a possibility to use the P element as a tool for genetic manipulations. Nowadays it is used by large for fly transformation and insertional mutagenesis.

Ivics et al. (1997) managed to reconstruct an active ancestral vertebrate's transposable element Sleeping Beauty (SB) from teleost fish. Numerous studies carried out recently give a promise of successful application of SB construct in gene therapy (Yant et al., 2000; Montini et al., 2002; Oritz-Urda et al., 2003).

The *piggyBac* vector is a system developed for insect transgenesis. It is a DNA transposon first isolated from the cabbage looper moth, *Trichoplusia ni* (Cary et al., 1989).

Any transgenesis would not be plausible without position effect correlation. Position effect during transgenesis is successfully suppressed with a *gypsy* insulator, a part of the *gypsy* retroposon discovered when observing unusual character of mutations caused by *gypsy* retroposon insertions in *D. melanogaster* (Holdridge and Dorsett, 1991).

It is noteworthy that the sequence motif of alpha satellite DNA in human is specific for every chromosome (Jorgensen, 1997). Chromosome specific alpha-satellite DNA probes are used in clinical cytogenetics and biomedical science for detection of aneuploidies, non-disjunctions, chromosomal abnormalities and rearrangements involving centromere, studying the parent of origin effect in cancer cytogenetics and genomic imprinting. Alpha satellite variants of particular chromosome among different individuals allow tracking back the pedigree and kinship of tested individuals, in a similar way as a common polymorphic marker (O'Keefe et al., 1996).

Variability of microsatellite and minisatellite repeats (VNTR) is used for DNA profiling in forensic studies, in paternity test or other kinship determination (Jeffreys et al., 1985; Litt and Luty, 1989; Tautz, 1989; Weber and May, 1989). What is more, the extensive length polymorphism of microsatellite loci even within closely related species enables using them as efficient molecular markers in population genetic studies using RFLP or AFLP (Goldstein and Schlotterer, 1999). Isolation of microsatellite loci is possible owing to their unique flanking regions. This characteristic is used for the construction of high density linkage maps (Beckman and Soller, 1990; Morgante and Olivieri, 1993).

1.6. Repetitive DNA in lepidopteran genome

The very first report on the repetitive DNA analysis in lepidopteran genome dates back to 1976. Efstratiadis in cooperation with Britten made a study based on the Cot analysis of the genome of Chinese oak silkmoth, *Antheraea pernyi*. They concluded that a large fraction of the genome is composed of interspersed sequences and repetitive DNA (Efstratiadis et al, 1976).

Analysis of a repetitive DNA depends by large on the knowledge of genome sequence of an analyzed organism. The only representative of Lepidoptera with known genome sequence is the silkworm *B. mori*. Futahashi et al. (2008) conducted screening of the *B. mori* genome focusing on the characterization and genome distribution of transposable elements. Transposable elements, represented mainly by SINEs and non-LTR transposons, comprises up to 35% of the silkworm genome, which represents the second highest portion presented among all insect species analyzed so far. The detailed analysis of *B. mori* W chromosome revealed that its prevalent part is composed of nested structures of various transposable elements (Abe et al., 2005).

First evidence on the composition of the lepidopteran genome in terms of repetitive DNA came from Prasad et al. (2004) who carried out an analysis of *B. mori* genome focused on frequency and distribution of microsatellites and their conservation within lepidopteran species. They found that 0.31 % of the *B. mori* genome is composed of microsatellite sequences when microsatellites with high AT content and short repeat motif where the most abundant.

Franck et al. (2004) isolated 24 microsatellite markers from codling moth *Cydia pomonella* and recommended them for further use in population genetic studies when almost in the same time Meglecz et al. (2004) discovered that butterflies` flanking regions of microsatellite loci demonstrate very high similarity. The fact makes use of microsatellite loci as a population genetic marker impossible for Lepidopera. Van't Hof, et al. (2007) conducted survey of microsatellite loci of lepidopteran representative *Bicyclus anynana* and based on his observation he suggested several mechanisms as a possible explanation for this phenomenon.

B. mori also shows the highest proportion of microsatellite families out of 10 analyzed insect species (Meglécz et al., 2007). These findings suggest that the presence of microsatellite families is at least partially the result of the association between microsatellites and interspersed repetitive elements (Meglécz et al., 2007). Curious consequences result from these findings. As microsatellite loci serve as markers in

population genetics studies it is therefore difficult to isolate microsatellite markers at certain species and enable their further use or characterization and distribution in the genome as in the case of lepidopterans mentioned above.

Mandrioli et al. (2003) managed to isolate the first, and so far the only satellite DNA in Lepidoptera. A satellite, designated MBSAT1, was isolated from the cell line of the cabbage moth, *Mamestra brassicae.* The MBSAT1 was localized into heterochromatic regions of sex chromosomes by FISH.

To the best of my knowledge information presented above are the only conclusions made about repetitive DNA in lepidopterans. To widen knowledge about distribution and character of repetitive DNA in Lepidoptera, the group of insect with holokinetic chromosomes, I decided to conduct my research on a well-known lepidopteran pest, codling moth, *Cydia pomonella.*

1.7. Holokinetic chromosomes

In most organisms the kinetics of the chromosomes during the course of their segregation is ensured by kinetic apparatus localized in centromere (Mola and Papeschi, 2006). However some organisms show different organization where the kinetic activity is more or less evenly distributed along chromosomes (termed as holokinetic or holocentric chromosomes). Holokinetic or holocentric chromosomes are terms referring to chromosomes without a primary restriction, the centromere, which implies particular structure and behaviour that those chromosomes assume during a cell cycle especially during mitosis or meiosis (Mola and Papeschi, 2006).

Holokinetic chromosomes have been discovered within wide range of organisms even though their occurrence is rather scarce. They have been found in representatives of both plants and animals, and also in unicellular organisms such as *Rhizaria*. In insects, they are present in numerous orders including Lepidoptera and their sister order Trichoptera (for details see review of Mola and Papeschi, 2006).

Presence of the holokinetic chromosomes lends the host organism specific features and provides the species with several advantages in comparison to the systems with monocentric chromosomes. The structure of kinetochore apparatus determines character of heritability of chromosome fragments emerging in response to exposure to various clastogenic agents or irradiation. The fragments of holokinetic chromosomes were proven to be attached to the mitotic spindle during chromosome segregation and they can carry through the cell cycle. Unlike monocentric chromosomes, holokinetic chromosomes do not create dicentric fragments that block mitosis and are therefore fatal for the cell (Hughes-

Schrader and Ris, 1941; Ris, 1942; Hughes-Schrader and Schrader, 1961; Nordenskiold, 1963). The fact has the large consequences on the evolution of karyotype in holokinetic systems. Holokinetic chromosomes also show different way of evolution of chromosome structure not only in terms of distribution of heterochromatin compared to monocentric chromosomes (Mola & Papeschi, 2006).

1.8. *Cydia pomonella*

The presented master thesis deals with a model organism, the codling moth *Cydia pomonella* (Tortricidae; Lepidoptera). The karyotype of *C. pomonella* comprises 56 chromosomes of the holokinetic type and almost uniform size. Fuková et al. (2005) sorted out the chromosomes into 5 size classes. It is still impossible to identify individual chromosomes due to the absence of banding techniques at Lepidoptera. The sex chromosomes (WZ/ZZ; female/male) are the only chromosomes recognizable of the *C. pomonella* karyotype. The Z chromosome is the biggest element of the complement. The W chromosome, which is only slightly smaller, is composed of heterochromatin and it forms a heterochromatic body, the so-called sex chromatin, in interphase nuclei of somatic cells.

2. Aims of the thesis

Codling moth is the major pest of pome and walnut orchards of the moderate climate worldwide. A systematic research of this species is therefore of the highest importance (Fuková et al., 2005). Numerous studies focusing on genome structure and evolution has been carried out so far. Only scarcely do these studies concern organisms possessing holokinetic chromosomes. A complete image of the chromosome structure requires the understanding of the repetitive DNA distribution (Ferreira and Martins, 2008). While the location of repetitive DNA, particularly satellite DNA in monocentric chromosomes has its typical pericentric and subtelomeric location, the situation in the holocentric chromosomes must be different due to the special kinetochore structure (Charlesworth et al., 1994; Mola and Papeschi, 2006). The whole genome sequencing (WGS) project which would facilitate an effort to make an overall picture of the genome organization in holokinetic systems has been carried out only in *C. elegans* worm (*C. elegans* sequencing consortium, 1998) and silkworm so far (Mita et al., 2004). Such knowledge therefore requires further more detailed survey.

Until WGS project is underway a further research depends by large on the physical markers mapped by in situ hybridization to the genome. Since previous studies reported repetitive elements demonstrating selective, sometime even chromosome specific distribution pattern in monocentric chromosomes (Willard et al., 1987), it would be interesting to find out whether organisms with holokinetic chromosomes demonstrate the same tendency (considering that most repetitive DNA gather around centromere and subtelomeric regions).

Speaking of *C. pomonella* only limited knowledge of its genome has been gathered so far (Fuková et al., 2005, 2007, 2009; Makee et al., 2008). Since codling moth does not represent a model organism of the Lepidoptera genetics, the whole genome sequence is not available for this species. It is therefore necessary to use another approach for genome analysis. The aim of this thesis was to isolate and characterise different types of repetitive DNA in the genome of *C. pomonella*. The work consisted in partial steps as followed. (1) To test various approaches allowing isolation of repetitive DNA from the genome of the codling moth. (2) To reveal the character of isolated DNA with Southern blot, to obtain the sequence of the respective repetitive element and localize it in *C. pomonella* complement with FISH. (3) To verify its conservativeness in different lepidopteran species with cross-species Dot blot. (4) To map microsatellite probes to the genome of *C. pomonella* by means of microsatellite FISH

3. Material and Methods

3.1. Insects

Cydia pomonella (codling moth; Tortricidae) instar of strain Krym-61 available in Laboratory of Molecular Cytogenetics, Institute of Entomology, Biology Centre, ASCR, České Budějovice, Czech Republic. Details about the strain are discussed in work of Fuková et al. (2005).

Grapholita molesta (Oriental fruit moth; Tortricidae) is available in Laboratory of Molecular Cytogenetics, Institute of Entomology. The strain was provided by Prof. Silvia Dorn (Institute of Plant, Animal and Agroecosystem Sciences, Applied Entomology, ETH Zurich, Switzerland) Details of the strain listed in work of Notter-Hausmann, C. and Dorn, S. (2010)

Lobesia botrana (European grapevine moth; Tortricidae) is available in Laboratory of Molecular Cytogenetics, Institute of Entomology. The strain was provided by Prof. Annette Reineke (Research Center Geisenheiminstitute, of Biology, Department of Phytomedicine, Entomology, Geisenheim, Germany).

Ephestia kuehniella (flour moth; Pyralidae), strain WT-C was available in Laboratory of Molecular Cytogenetics, Institute of Entomology. Details of the strain listed in work Marec (1990).

Bombyx mori (silkworm; Bombycidae), strain P29, was provided by Valeriya Zabelina, Biological Center, AV ČR, České Budějovice, Czech Republic.

Xestia c- nigrum (Spotted cutworm; Noctuidae) was generously provided by Michal Zapletal, Institute of Entomology, The collection was made in moth collector in České Budějovice, Czech Republic.

Phthorimaea operculella (Potato tuber moth; Gelechiidae) was available in a breeding stock of Laboratory of Molecular Cytogenetics, Institute of Entomology. The strain was provided by Dr. Hayat Makee (Department of Molecular Biology and Biotechnology, Atomic Energy Commission of Syria, Damascus, Syria). Details of the strain are listed in Makee, H. and G. Saour (1999).

Pieris brassicae (Large white; Pieridae) was provided by my colleague from Laboratory of Molecular Cytogenetics, Petr Nguyen. Specimens were collected in natural population in Ohrazení near České Budějovice, Czech Republic. Details of the specimens listed in work Nguyen et al. (2010).

Polyommatus bellargus (Adonis blue; Lyceanidae) was provided by my colleague from Laboratory of Molecular Cytogenetics, Petr Nguyen. Specimens were collected in natural population in Ohrazení near České Budějovice, Czech Republic. Details of the specimens listed in work Nguyen et al. (2010).

Inachis io (European peacock; Nymphalidae) was provided by my tutor Magda Vítková from Laboratory of Molecular Cytogenetics. Specimens were collected in natural population near Veselí nad Lužnicí, Czech Republic.

Lasiommata megera (Wall brown, Nymphalidae) was provided by my tutor Magda Vítková from Laboratory of Molecular Cytogenetics. Specimens were collected in natural population near České Budějovice, Czech Republic.

3.2. Genomic self-priming PCR (GSP-PCR)

The GSP-PCR was carried out according to the protocol listed in the work of Buntjer and Lenstra (1999). 50 ng of genomic DNA (gDNA) was denaturated at 100°C for 15 min and mixed it with 50 ng of intact gDNA. The DNA was used in 25 µl PCR mix (200 µM DNTP mix, 1x Ex Taq buffer, 4 mM $MgCl_2$, 1.25 U of Taq polymerase (TaKaRa Bio Inc, Otsu, Japan) and the PCR reaction was performed with following PCR profile: predenaturation at 95°C for 5 min, 30 cycles of denaturation at 92°C for 1 min, annealing at 50°C for 2 min, elongation at 72°C for 2 min and 1 cycle of postelongation at 72°C for 7 min. 5 µl of the product was used for reamplification. The products were reamplified once and twice, respectively, and subsequently separated electrophoretically on 1% agarose gel at 5V/cm. The product amplified once was also digested with selected restriction endonucleases (*Afa* I, *Alu* I, *Hinf* I, *Eco*R I, *Pst* I, *Rsa* I, *Sau*3A I, *Taq* I) and the products were separated electrophoretically on 1% agarose gel. Alternatively, gDNA was digested with selected restriction endonucleases, products were mixed with the same amount of intact gDNA and used for GSP-PCR. The products were digested with same restriction endonucelases and the products were separated electrophoretically on 1% agarose gel. The gel was stained

with Gel Red™ Nucleic Acid Gel Stain (Biotium, Inc., California, USA) and the outcome was documented under UV light in White/UV Transilluminator with rainbow TV zoom lens (UVP, LLC, Japan). The scheme of the experiment is listed below for the better understanding of the process (see Fig.1.).

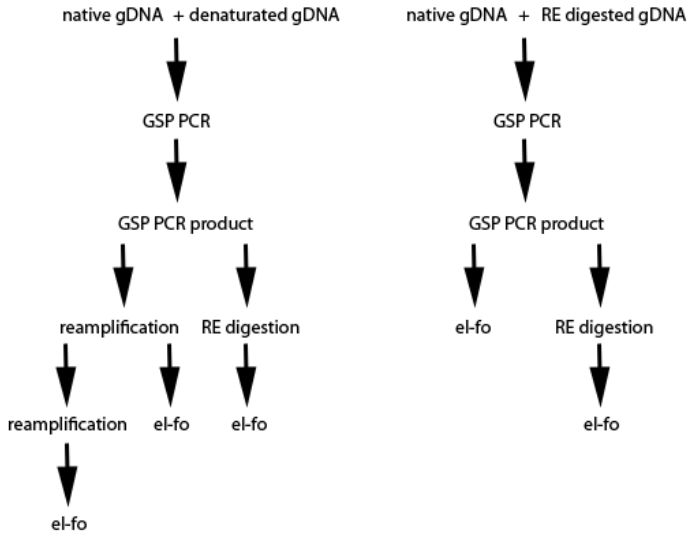

Fig.1. A scheme of GSP-PCR. RE = restriction endonuclease, el-fo = visualization with agarose-gel electrotrophoresis.

3.3. Chromosome preparations

Both mitotic and meiotic chromosome preparations were obtained from 5th instar female larvae of *Cydia pomonella.* Mitotic preparations were made of imaginal wing discs. Preparations were made according to a slightly modified procedure of Sahara et al. (1999). Ovaria were dissected in Ringer physiological solution according to Glaser (1917) (0.9% Sodium Chloride, 0.042% Potassium Chloride, 0.025% Calcium Chloride, 0.02% Sodium Hydrogen Carbonate) treated in hypotonic solution (0,075M KCl), and transferred into Carnoy fixative (Ethanol-Chloroform-Acetic acid, 6:3:1) where they were left for 15 min. After fixation the ovaria were transferred on the slide into a drop of 60% acetic acid and macerated with tungsten needles. Macerated material was spread onto histological plate heated to 45°C. For more efficient spreading hot and cold was changed when switching histological plate with frozen plate.

3.4. DNA extraction

DNA extraction was performed using standard phenol-chloroform extraction. Adult females of *Cydia pomonella, Grapholita molesta, Lobesia botrana, Ephestia kuehniella,*

Bombyx mori, Xestia c-nigrum and Phthorimaea operculella were collected and their abdomens were removed. The remaining tissue was either homogenized in liquid nitrogen and incubated in extraction buffer (0.5 g of tissue/10 ml of extraction buffer: 100 mM NaCl, 50 mM EDTA, 100 µg/ml Proteinase K, 0.5% Sarkosyl, 10 mM Tris-Cl, pH 8) overnight while shaking gently at 37°C in case of *C. pomonella* or were homogenized in 500 µl of extraction buffer in 1.5 µl eppendorf tube with little pestle and were incubated overnight while shaking gently at 37°C in case of remaining specimens. The next day the RNase A was added (final concentration 10 µg/ml) and the mixture was incubated another one hour at 37°C while gently shaking. Then phenol (pH 8) of the amount corresponding to the sample volume was added to each sample and incubated at room temperature (RT) for 30 min while gently shaking. Sample was centrifuged for 15 min at 5000 g. Upper phase containing DNA was transferred into a clean tube. This step was repeated once or twice depending on protein contamination. Then, same volume of phenol/chloroform/isoamylalcohol mixture (ratio 25:24:1) was added, incubated and centrifuged as during phenol treatment. The last extraction step was performed with chloroform/isoamylalcohol mixture (ratio 24:1). DNA was precipitated by addition of 3M Na-acetate (1/10 of sample volume) and isopropanol (7/10 of sample volume). Precipitated DNA was briefly washed in 70% ethanol and dissolved in TE buffer. Quality of DNA was verified in use of restriction analysis (1 µg of DNA was digested with *Hind* III at 37°C for one hour) and following electrophoretic separation in 1% agarose gel in TAE. Concentration of DNA was determined on fluorometer DyNA Quant 200 (Amersham Biosciences, Buckinghamshire, UK). Genomic DNA was used for restricion analysis of tandem repeats, GSP-PCR, Southern hybridization and Dot blot hybridization.

3.5. Visualization of tandem repeats via restriction analysis

Isolated genomic DNA (gDNA) was digested with 27 different restriction endonucleases listed in the table below with number of bands it produced, 10-20 µl of digestion reaction contained 5-10 µg of gDNA, 1x restriction buffer, 1 µl of respecive restriction endonuclease (number of units ranged from 10 to 15 depending on the enzyme). Reaction mix was incubated overnight at 37°C while gently shaking. Restriction products were separated electrophoretically on 1.5% agarose gel in TBE at 5V/cm, labelled with Gel Red Red™ Nucleic Acid Gel Stain for 10 min, and checked under the UV light.

3.6. Cloning and sequencing

Desirable bands were cut out under UV light and isolated DNA with Wizard SV Gel and PCR-up System (Promega Corporation, Madison, WI, USA). Blunted ends of the product were incubated with Mung Bean nuclease at 37°C for 60 min and dATPs were added to the blunt ends with rTaq polymerase at 72°C for 30-60 min. Adjusted DNA was cloned by use of pGEM-T easy vector system (Promega Corporation, Madison, WI, USA). Alternatively the pUC19 cloning system (Invitrogen, Carlsbad, CA, USA) was used. Vector pUC19 was digested with respective restriction endonuclease, 50 ng of the vector was used in ligation reaction along with 11-25 ng of the digested gDNA, 1x buffer and T4 DNA ligase (TaKaRa Bio Inc, Otsu, Japan). Ligation reaction was incubated overnight at 16°C. The product of ligation reaction was used for heat-shock transformation of chemically competent cells DH5α. Heat shock of 42°C was applied for 90 s. Cells were quickly cooled down on ice and let subsequently regenerate in LB medium with 2% glucose at 37°C for 45 to 60 min while shaking. Transformed cells were further transferred on Petri dishes with LB medium containing agar (LB medium, 2% agar, 100 μg/ml ampicilin, 350 μM isopropyl β-D-1-thiogalactopyranosid, 35 μg/ml X-gal). Positive white colonies were tested for the presence of insert by use of PCR with universal M13-24 primer (5´-CGC CAG GGT TTT CCC AGT CAC GAC-3´) and M 13-26 primer (5´-CAG GAA ACA GCT ATG AC-3´). 12.5 μl reaction mix contained 1x Ex-Taq buffer, dNTP mix (200 μM of each nucleotide), 6 μM of each primer, 0.5 U of Ex Taq DNA polymerase, HS (TaKaRa). PCR profile was as followed. Predenaturation at 94°C for 3 min, denaturation at 94°C for 30 s, annealing at 57°C for 30 s, elongation at 72°C for 60 s (steps 2-4 were repeated 30 times), postelongation at 72°C for 7 min. Final products were separated electrophoretically on 1% agarose gel in TAE. Clones with inserts of various sizes were used for further sequencing. Sequencing reactions were carried out with universal primers M 13-24 and M 13-26 listed above and with use of BigDye Terminator in 3.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, USA) in the Laboratory of genomics, Biology centre ASCR. Acquired sequences were adjusted in program MEGA4 (Tamura et al., 2007) available for free at www.megasoftwarenet.com. Partial vector sequence was removed in on-line program BLAST, subfunction VecScreen available in NCBI database and the identity of the insert sequences was verified in NCBI database in subfunction BLASTN (http://www.ncbi.nlm.nih.gov/).

3.7. Probes for hybridization techniques

DIG labelled probe for Southern hybridization and Dot blot hybridization was generated by means of PCR from pUC19 plasmids containing inserts. The PCR reaction and conditions were exactly same as used to test the colonies for insert presence, except the dNTP mix, which contained DIG-11-dUTP (Roche Diagnostics, Basel, Switzerland) (dNTP mix: 1 mM dGTP, dCTP, dATP, 0.65 mM dTTP, 0.35 mM DIG-dUTP).

Unlabelled telomeric probe $(TTAGG)_n$ was generated by non-template PCR with $(TTAGG)_4$ and $(CCTAA)_4$ primers custom-made by Generi Biotech (Hradec Králové, Czech Republic). 25 µl PCR reaction contained 1x Takara ExTaq buffer, 200 µmol/L dNTP mix, 0.5 µmol/L of each primer and 1.25 units of Ex Taq DNA polymerase, HS (TaKaRa). PCR conditions were: Predenaturation at 94°C for 90 s, denaturation at 94°C for 45 s, annealing at 52°C for 30 s, elongation at 72°C for 60 s (steps 2-4 were repeated 30 times), postelongation at 72°C for 10 min. PCR product was consequently labelled either with digoxigenin (a probe designated for FISH with biotin labelled satellite probe) or with SpectrumGreen (a probe designated for FISH with Cy3 labelled microsatellite probes).
DIG-11-dUTP labelling was performed using High Prime DNA Labeling Kit (Roche Diagnostics, Mannheim, Germany) following the manufacturer protocol. Latter probe was labeled with SpectrumGreen-dUTP (Vysis Inc., Downers Grove, USA) by means of Nick Translation Mix (Roche Diagnostics). We used our own dNTP mixture containing 50 µM dCTP, dGTP, dATP, 14 µM dTTP), with 35 µM SpectrumGreen-dUTP per 300 ng of unlabelled telomeric probe. The reaction run overnight at 16°C and then was stopped by adding 1 µl of 0.5 M EDTA. Length of the probes was checked on 1% agarose gel in TAE. Optimal length of the labelled probe was 500-1000 bp.

Cy3-directly labelled microsatellite probes (VBC-Biotech, Vienna, Austria) were a courtesy of doc. RNDr. Eduard Kejnovský, CSc. (Institute of Biophysics ASCR, Brno, Czech Republic).

3.8. Southern hybridization

Approximately 6 µg of *Cydia pomonella* genomic DNA per sample was digested with 15 U of *Xba* I endonuclease for one hour at 37°C. Then, another 15 U of *Xba* I were added to the reaction and incubated overnight at 37°C. Next day the DNA fragments were separated on a 1% agarose gel in 1x TBE buffer (5V/cm for 2-4 hours). The gel was stained with Gel Red™ Nucleic Acid Gel Stain and checked under UV light in order to confirm presence of sufficient amount of DNA and its proper fragmentation. The gel was consequently treated for 10 min with 0,25 M HCl at RT, 2 x 15 min with denaturation

solution (0.5 M NaOH, 1.5 M NaCl) at RT, and 2 x 15 min with neutralization solution (0.5 M Tris-HCl, 3M NaCl, pH 7.5). Finally, DNA was blotted overnight onto a Hybond-N+ nylon membrane (Amersham Biosciences, Buckinghamshire, UK) by capillary transfer in 20x SSC. Then the membrane was briefly washed in 2x SSC and crosslinked in the UV crosslinker (Amersham Life Science, Little Chalfont, UK) at 1200 mJ. Hybridization of genomic DNAs with probes was performed at 42°C for 20 hours in 6.5 ml of a hybridization solution prepared from DIG Easy Hyb Granules (Roche Diagnostics), containing 50 ng of denaturated respective probe. Washes of 2 x 5 min in 0.1% SDS/2x SSC at RT and a stringent wash in 0.1% SDS/0.2x SSC for 15 min at 68°C were followed by the procedure for signal detection at RT: 5 min in washing solution (0.3% Tween 20 in TBS (TBS: 250mM Tris-HCl, 2M NaCl, pH 7,5)), 30 min in blocking solution (5% fat free dry milk in TBS; Difco skim milk, Becton, Dickinson and Company, Sparks, MD, USA), 30 min in 100 mL blocking solution containing 7.5 U Anti-Digoxigenin-AP (Roche Diagnostics), 2 x 15 min in washing solution and 5 min in pH-adjusting solution (0.1 M Tris-HCl, 0.1 M NaCl, pH 9.5). Chemiluminescent signals were obtained by incubation of the membrane in CDP-*Star* ready-to-use substrate (Roche Diagnostics) for 10-30 min at RT and images were recorded with a LAS-3000 Lumi-Imager (Fuji Photo Film Europe GmbH, Düsseldorf, Germany). Exposure times ranged from 1 to 10 min.

3.9. Dot-blot hybridization

500 ng of genomic DNAs of *Cydia pomonella, Grapholita molesta, Lobesia botrana, Ephestia kuehniella, Bombyx mori, Xestia c- nigrum,* and *Phthorimaea operculella* was denaturated at 95°C for 5 minutes, chilled on ice and dripped on the Hybond-N+ nylon membrane (Amersham Biosciences, Buckinghamshire, UK). The membrane with DNA samples was cross-linked in UV crosslinker at 1200 mJ, soaked in 2x SSC and hybridized with 50 ng of denaturated DIG-labelled satellite probe in 6.5 mL of DIG Easy Hyb solution at 42°C or 52°C overnight. Washing and detection conditions were the same as in Southern hybridization.

3.10. Fluorescence in situ hybridization (FISH)

3.10.1. FISH with biotin/digoxigenin labeled probe

Slides were taken out of the freezer (-20°C), dehydrated in ethanol series 70%-80%-100% and let dry for 10 min. In order to decrease background signal caused by excessive cytoplasm chromosome slides were pretreated with solution of RNAseA and Proteinase K. Chromosome preparations were baked at 60°C for four hours to avoid washing the

material off the slide during following Proteinase K pretreatment. 100 µl of RNAse A solution (200 µg/ml of RNAse A in 2x SSC) were dripped on the preparation, covered with a cover slip (24x50 mm) and let incubate in a moisty chamber at 37°C for one hour. Further, the pretreatment in PBS (phosphate buffered saline) containing Proteinase K (1 µg/ml) was applied. The preparations were incubated in the Proteinase K solution in a Coplin jar at 37°C for 5 min in a rocking bath. Preparations were washed twice in PBS at 37°C for 5 min in a rocking bath. Preparations were blocked in 60 ml of 5x Denhardt's reagens (100x stock solution: 2% BSA, 2% Ficol, 2% Polyvinylpyrrolidone in 3x SSPE buffer) at 37°C in a rocking bath. After pretreatment the preparations were immediately denaturated. 100 µl of the 70% formamide in 2x SSC was dripped on the preparation and covered with a 24x50 mm cover slip. The denaturation proceeded at 68°C for 3 min 30 s. The cover glass was shook off and preparations were put into 70% cold ethanol for 1 min. Dehydration was finished with the ethanol series of 80% and 100% ethanol and preparation was let dry.

In the mean time the hybridization mix was prepared. Following amounts are intended for one slide. The probe or probes (20 ng of satellite probe, 40 ng of telomeric probe) were mixed with 25 µg of sonicated DNA from salmon sperm (Sigma-Aldrich, St. Louis, USA). 1/10 volume of 3 M Na-acetate and 2.5x volume of 100% cold ethanol were added to the hybridization mix and it was let precipitate for 30-60 min at -80°C. Hybridization mix was centrifugated at 13 000 rpm for 20 min. A supernatant was removed and the pellet was washed by addition of 200 µl of 70% cold ethanol. Then it was centrifugated at 13 000 rpm for 5 min. A supernatant was removed and pellet was let dry in order to get rid of excessive ethanol. A pellet was resuspended in 10 µl of deionised formamide at 37°C for 30 min. 10 µl of 20% dextran sulfate in 4x SSC was added and the mixture was denaturated at 90°C for 5 min. The mixture was cooled down on ice immediately and let stand there for at least 3 min. 20 µl of the hybridization mix was applied on denaturated and perfectly dry preparations. Preparations were covered with cover slip 24x32 mm and sealed with rubber cement (Marabuwerke, Germany). The preparations were placed into a moisty chamber misted with 2x SSC and let hybridize at 37°C overnight. The next day the cover glass was shook off and preparations were washed 3 times for 5 min in 50% formamide in 2x SSC at 46°C in rocking bath, further 5 times for 3 min in 2x SSC at 46°C in rocking bath, 3 times for 5 min in 0.1x SSC at 62°C in a rocking bath and 5 min in 4x SSC in 0.1% Tween 20 at RT in a darkness while shaking. Probe detection consisted of three sets of steps, each set started with blocking with 500 µl of 2.5% BSA in 2x SSC for 20 min at RT in darkness, then incubation with

antibodies, and finally washing three times in 4x SSC with 0.1% Tween 20 at 37°C for 5 min in a rocking bath. The antibodies and respective conditions were as follow. In the first set, 100 µl of mix of Streptavidin-Cy3 (Jackson ImmunoResearch Laboratories Inc., Pennsylvania, USA) dissolved 1:1000 in blocking solution (2.5% BSA) and Anti-DIG mouse monoclonal antibody (Roche, Switzerland) dissolved 1:25 in blocking solution] were applied on each slide. The incubation lasted one hour at 37°C in darkness. In the second set, 50 µl per preparation of antibody mix containing biotinilated Antistreptavidin (Vector Lab Inc., California, USA) dissolved 1:25 in a blocking solution and Anti mouse-IG antibody-DIG (Roche, Switzerland) dissolved 1:25 in a blocking solution was applied. Incubation conditions were the same as above. In the last set, the mixture of antibodies contained Streptavidin-Cy3 (Jackson ImmunoResearch Laboratories Inc., Pennsylvania, USA) dissolved 1:1000 in a blocking solution and anti-DIG antibody-fluorescein (Roche, Switzerland), dissolved 1:25 in a blocking solution. 100 µl of the mixture was applied on each slide. The incubation carried on 20 min at 37°C in darkness.

The DAPI staining was performed immediately. The preparations were washed in 1% Triton X in PBS at RT for 5 min. The preparations were incubated in 1% Triton X in PBS with DAPI (50 ng/ml) at RT for 15 min. After staining the preparations were washed in 1% PhotoFlo (Eastman Kodak Company, Rochester, USA) in PBS at RT for 2-5 min and in 1% PhotoFlo in miliQ H_2O for 1 min. An excessive fluid was drawn off, 20 µl of DAPCO (1,4-diazabicyclo[2.2.2]octane, Sigma-Aldrich, St Louis, Missouri, USA ; 20 mM Tris-HCl pH 8, 90% glycerol) was applied and preparation was covered with a 24x32 mm cover slip, an excessive liquid was forced out and the preparation was sealed with a nail polish. The hybridization signals were documented on the preparations with fluorescent microscope Zeiss Axioplan 2 (Carl Zeiss, Jena, Germany). The pictures were taken with F-view CCD camera and program AnalySIS 3.2 (Soft Imaging System, Münster, Germany). The Adobe Photoshop CS 8.0. was used for final arrangement of the pictures.

3.10.2. FISH with microsatellite probes

The preparations were treated with the very same pretreatment as it was used in case of FISH with biotin/DIG labeled probe in chap. 3.10.1. However, the pretreatment with Denhardt's reagens was skipped and the denaturation of preparations described in chap. 3.10.1. was performed immediately after post Proteinase K washes.

The hybridization mix for one preparation was prepared as followed. 300 ng of labeled telomeric probe and 300-500 ng of microsatellite probe were mixed. Precipitation, resuspendation and denaturation of the hybridization mix are described in chap. 3.10.1.

20 µl of denaturated hybridization mix was applied on the denaturated preparation, it was covered with 24x32 mm cover slip, sealed with rubber cement and let incubate at 37°C for 3 days in moisty chamber with 2x SSC. The third day the cover slip was shook off and the preparation were washed twice in 2x SSC at room temperature in rocking bath, twice in 1x SSC at room temperature in rocking bath and briefly in 1x PBS at room temperature. The preparations were stained with DAPI as described in chap. 3.10.1. Alternatively, the preparations were directly mounted in DAPI in DAPCO (0.5 mg/ml). The preparation was covered with 24x32 mm cover slip, an excessive liquid was pressed out and the preparation was sealed with nail polish.

4. Results

4.1. GSP-PCR

GSP-PCR is a method, which amplifies unknown repetitive DNA using fragmented gDNA as primers and intact gDNA as a template. The desired output is a band or bands, which contain highly repetitive sequences. I fragmented gDNA by denaturation and by digestion with one of 8 selected restriction endonucleases *Afa* I, *Alu* I, *Eco*R I, *Hin*f I, *Pst* I, *Rsa* I, *Sau*3A I *Taq* I. All Restriction endonucleases gave the same negative results therefore I present only the result of three of them and products of *Alu* I, *Pst* I, *Rsa* I, *Sau*3A I and *Taq* therefore are not shown. The fragmented DNA was used as a primer in GSP-PCR. In order to enhance the possible yield of repetitive DNA, I used the PCR product either for reamplification in case of PCR with denaturated gDNA or for digestion by respective restriction endonuclease in case of both types of GSP-PCR products (a scheme of the experiments is in Fig.1. in chap. 3.2.). Unfortunately, all the combinations gave nothing but a smear, therefore I decided to stop experiments with GSP-PCR and focus on other approaches of isolation of repetitive DNA. Outcomes of GSP-PCR are presented at the Figs.2 and 3.

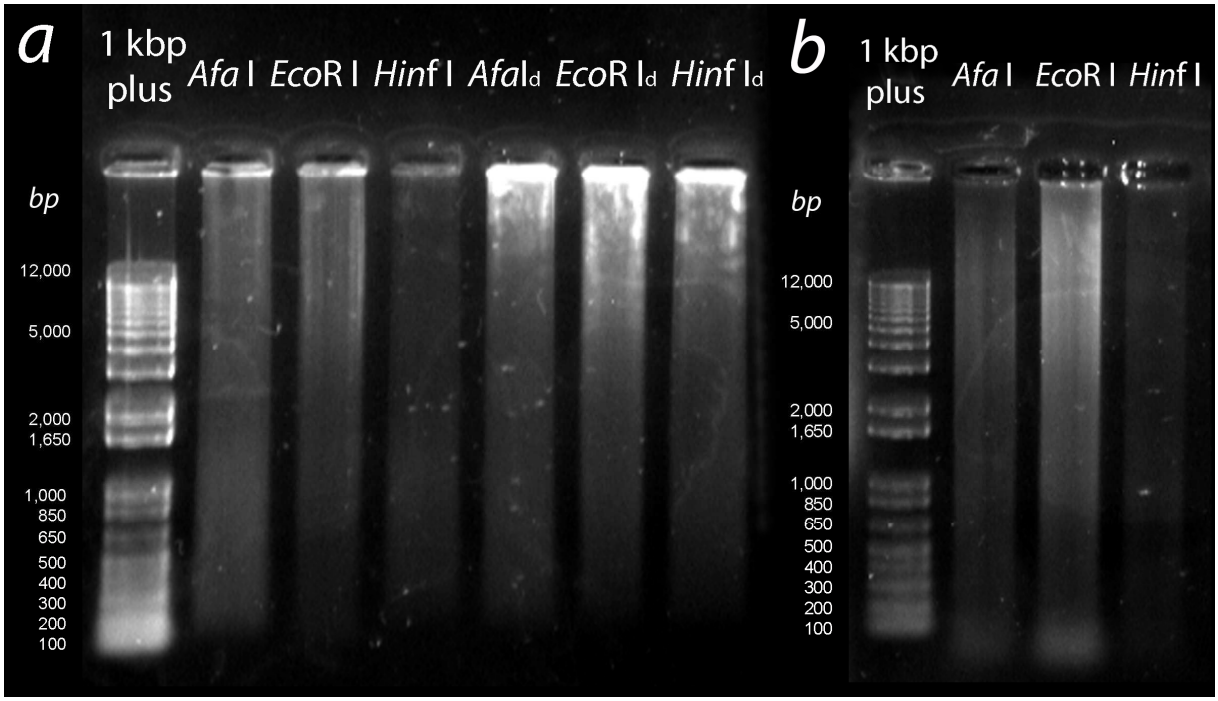

Fig.2. GSP-PCR with gDNA digested with selected restriction endonucleases. Fig. 2a presents two sets of products. Samples without lower "d" present products of GSP-PCR using DNA digested with respective restriction endonucleases as primers. Samples with lower "d" present products of GSP-PCR with denaturated gDNA as primers. Those products were subsequently digested with restriction endonucelases. Fig. 2b present GSP-PCR product with use of gDNA digested with selected restriction endonucleases as a primer.

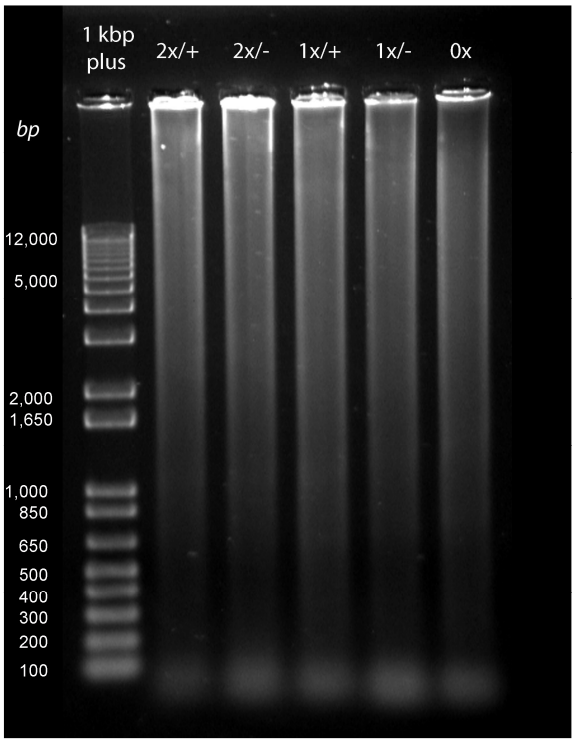

Fig.3. GSP-PCR with denaturated gDNA with subsequent reamplification once and twice. "+" symbol behind the slash presents GSP-PCR product with addition of denaturated gDNA as a primer and "-" symbol behind the slash presents sample without addition of the denaturated gDNA.

4.2. Isolation and characterisation of satellite DNA

4.2.1. Isolation of satellite DNA

I digested 5 µl of gDNA of *C. pomonella* with selected restriction endonucleases and separated it electrophoretically on 1.5% agarose gel. Restriction analysis provided various restriction products. Out of the 27 tested restriction endonucleases only 14 of them provided visible restriction bands (see Table 1). For the following analysis I used restriction products of enzymes *Cla* I, *Eco*R I, *Kpn* I, *Not* I, *Nsp* I, *Pst* I, *Taq* I and *Xba* I. Remaining restriction endonucleases provided long products, usually about several thousand base pairs long, therefore it would be uneasy to clone them and submit them to the further analysis. Products *Pst* I, *Xba* I and *Not* I restriction endonucelases each created 4 restriction bands ranging from 100 bp to 500 bp. Enzyme *Cla* I produced 3 bands of 350 bp, 250 bp and 100 bp. Restriction endonucleases *Eco*R I produced 3 bands of 2,400; 2,000; 1,100 bp and *Nsp* I and *Taq* I produced one band of 12,000 and 3,000 bp. I successfully cloned 100 bp and 150 bp restriction bands of *Xba* I, 400 bp band of *Kpn* I enzyme and 500 bp, 350 bp and 150 bp bands of *Pst* I enzyme (Fig. 4).

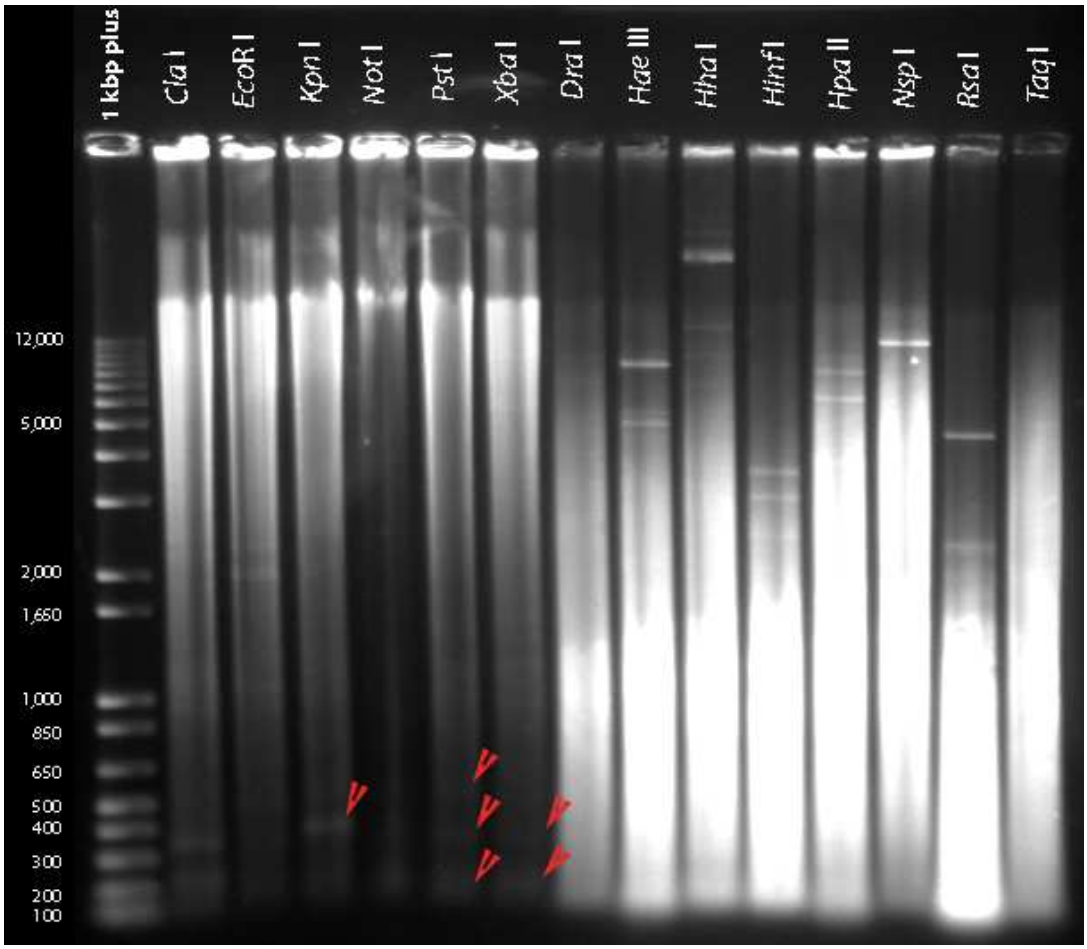

Fig.4. Result of the restriction analysis by agarose gel electrophoresis. Analysis presents only products of those restriction endonucleases producing restriction bands. Red arrows point at the successfully cloned bands.

Tab.1. Summary of products obtained in restriction analysis

Restriction endonucleas	Number of bands	Length of the bands (bp)	Isolated bands
Alu I	0		
Apa I	0		
*Bam*H I	0		
Bgl I	0		
Cla I	3	350; 250; 100	
Dra I	2	9,500; 7,000	
*Eco*R I	3	2,400; 2,000; 1,100	
*Eco*R V	0		
Hae III	4	10,000; 9,000; 5,500; 5,000	
Hha I	4	3 bands larger than 12,000; 9,500	
Hind III	0		
Hinf I	3	3,500; 3,000; 2,500	
Hpa II	3	9,500; 8,000; 6,000	
Kpn I	2	400; 180	400
Mbp II	0		
Nde I	0		
Nha I	0		
Not I	4	500; 350; 150; 100	
Nru I	0		
Nsp I	1	12,000	
Pst I	4	500; 350; 150; 100	500; 350; 150
Rsa I	2	4,500; 2,200	
*Sau*3A I	0		
Sme I	0		
Ssp I	0		
Taq I	1	3,000	
Xba I	4	500; 350; 150; 100	150; 100
Xho I	0		

4.2.2. Sequencing and characterization of satellite DNA with Southern blot

I sequenced the obtained clones and submitted them to the identification in NCBI. The 100 bp band of *Xba* I enzyme (*Xba* I 1-1) turned out to be analogous to microsatellite sequence (NCBI: *Cydia pomonella* clone CP5.173 microsatellite sequence, Accesion gi 88866987 DO394030.1, E value 8e-32). The remaining clones appeared to be of unknown sequences.

The presented thesis focuses on the complete analysis of isolated satellite DNA and limited time did not allow finishing analysis of the remaining clones. Thereby I submitted only selected clones to Southern blot to find out character of the isolated repetitive sequences. Only clone *Xba* I 1-1 proved to be a satellite repeat (Fig. 5), since it expressed a typical ladder-like pattern of bands of decreasing intensity towards longer fragments. Since it is the first satellite sequence isolated from *C. pomonella*, I named it CPSAT-1. Clone *Xba* I 2-2 and clone *Kpn* I 1-1 both seem to be an abundant dispersed repetition of an unknown mobile element. My conclusion are supported by Southern blot and FISH (data not shown).

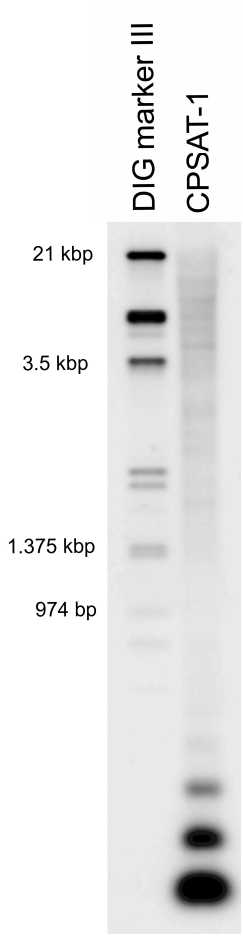

Fig.5. Southern hybridization of CPSAT-1 with *Xba* I digested *C. pomonella* genomic DNA. Applied size marker was DIG Marker III (Roche Diagnostics) Exposition time was 1 minute.

4.2.3. Distribution of CPSAT-1 in *C. pomonella* chromosomes

In order to find out the distribution of CPSAT-1 I performed FISH with mitotic preparations. Satellite created rather random dispersive pattern. Signals are basically scattered throughout all the chromosomes except for one medium size pair of chromosomes where signals gather into bigger little block (Fig. 6).

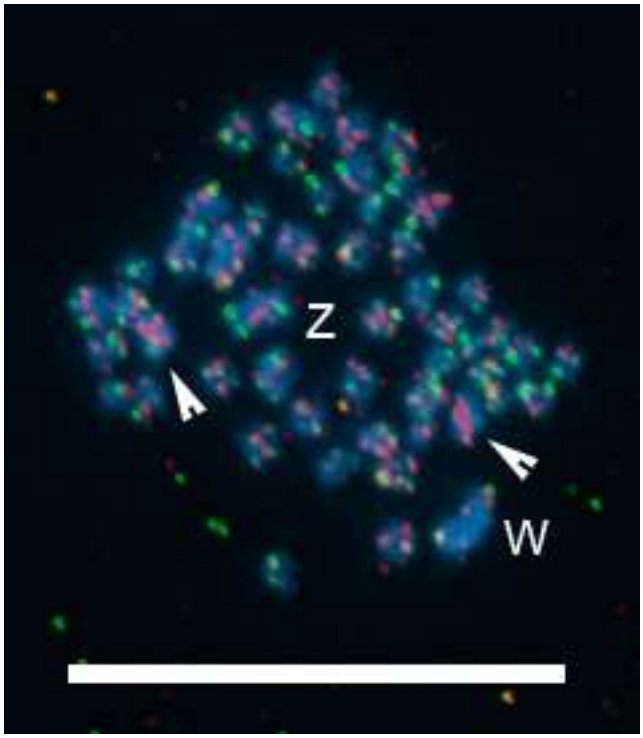

Fig.6. FISH with CPSAT-1 on *C. pomonella* female mitotic chromosomes. Hybridization signal of the CPSAT-1 is red, telomeric probe is green. Chromosomes are counterstained with DAPI. The arrowheads point at chromosomes with a distinct satellite block. Scale bar = 20 μm.

4.2.4. Dot blot analysis

I submitted the CPSAT-1 to dot-blot analysis with gDNA of various lepidopteran species in order to find out possible presence of this sequence in other lepidopteran species, as well as its abundance and conservativeness. The species were selected to represent taxa closely related as well as distant to *C. pomonella*. Hybridization was carried out at two different temperatures, one at 42°C considered as a standard temperature and more stringent temperature of 52°C in order to determine quantity or rate of divergence of CPSAT-1 in other species. *C. pomonella* created the strongest positive hybridization signal, but it was present in some of the tested species as well. The other species beginning with *Bombyx mori* along with *Phthorimaea operculella, Xestia* c- *nigrum, Grapholita molesta,* and *Lobesia botrana* demonstrated weaker signal. Signal weakened in all species except

in C. pomonella when more stringent temperature of 52°C was applied during hybridization. *Ephestia kuehniella* demonstrated the least visible signal at 42°C that completely disappeared when the hybridization was performed at more stringent temperature of 52°C (Fig. 7). The other lepidopteran species used in this analysis, namely *Pieris brassicae*, *Polyommatus bellargus*, *Inachis io*, and *Lasiommata megera* did not show any hybridization signal at any of the two hybridization temperatures (data not shown).

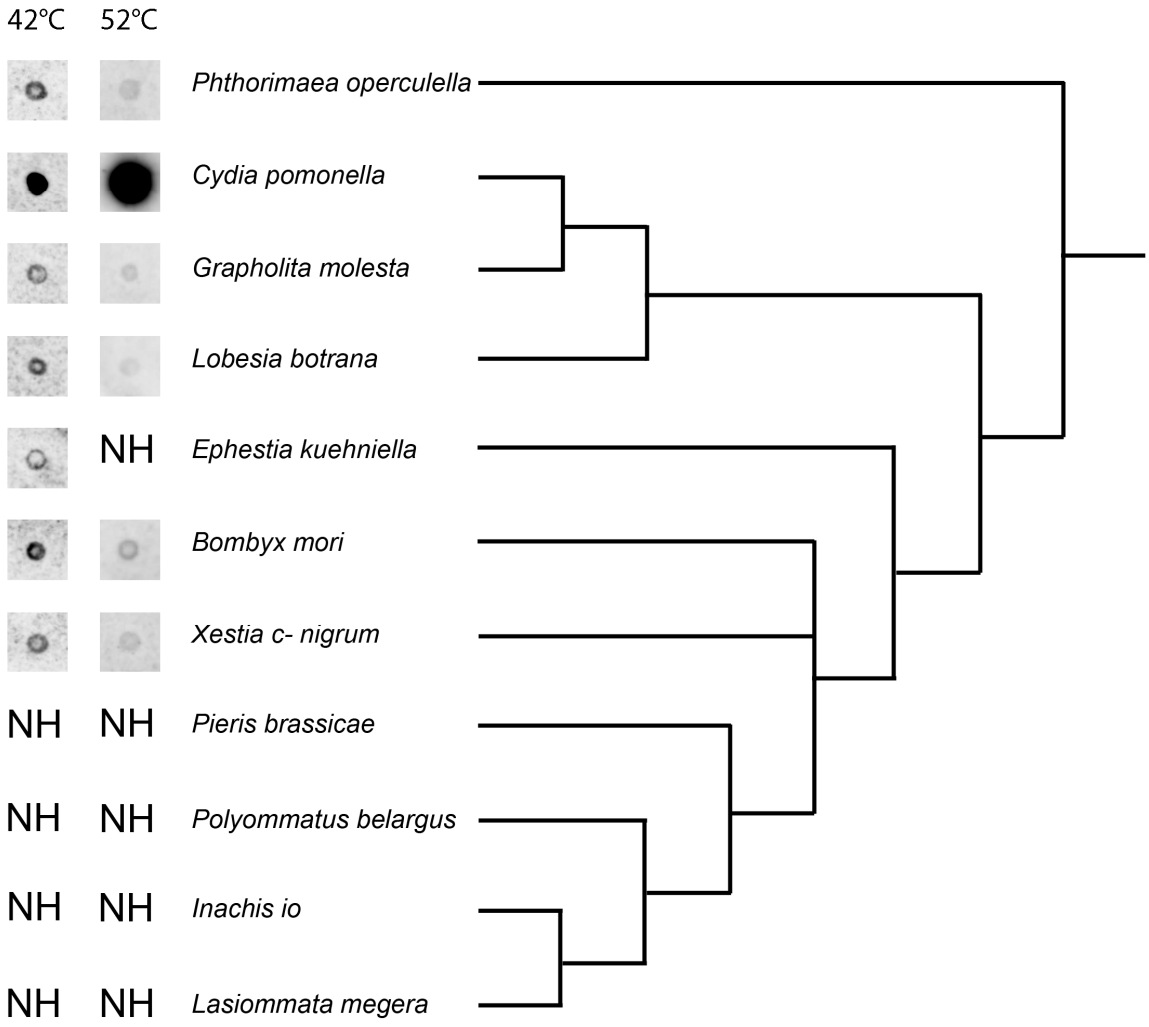

Fig.7. Dot blot with CPSAT-1 probe hybridized on DNA of *C.pomonella*, *G. molesta*, *L. botrana* (all Tortricidae), *E. kuehiella* (Pyralidae), *B. mori* (Bombycidae), *X. c-nigrum* (Noctuidae), *P. brassicae* (Pieridae), *P. belargus* (Lycaenidae), *I. io* (Nymphalidae), *L. megera* (Nymphalidae), and *P. operculella* (Gelechiidae). Hybridization was carried out at 42°C and 52°C, respectively. Picture also demonstrates phylogenetic relationship of the species (Komai, 1999; Mutanen, 2010; Horak, 2006, Kim et al., 2010, Kristensen et al.,2007). NH = no hybridization.

4.3. Microsatellite FISH

For analysis of microsatellite distribution I used FISH with 16 possible combinations of mono- di- and trinucleotides probes $(A)_{30}$, $(C)_{30}$, $(CA)_{15}$, $(GA)_{15}$, $(GC)_{15}$, $(TA)_{15}$, $(CAA)_{10}$, $(CAC)_{10}$, $(CAG)_{10}$, $(CAT)_{10}$, $(CGG)_{10}$, $(GAA)_{10}$, $(GAC)_{10}$, $(GAG)_{10}$, $(TAA)_{10}$, $(TAC)_{10}$, a courtesy of doc. RNDr. Eduard Kejnovský, CSc., according to Kubát et al., 2008. Additionally, I tested distribution of GATA tetranucleotide, since it has been shown before that it may be engaged in sex chromosome evolution of some species (e.g. Subramanian et al, 2002). I analyzed the distribution on mitotic preparations. As a control probe I used telomeric probe labeled with Spectrum Green (not shown) in order to evaluate the quality of hybridization procedure and determine an orientation of mitotic chromosomes. Most of the microsatellites were more or less evenly distributed along chromosomes, however, there were exceptions. Also, different signal intensity was recorded in comparison of individual microsatellites. Distribution of respective microsatellites and their pattern on sex chromosomes are summarized in Tab. 2. In order to get reliable data on the microsatellite distribution pattern, I carried out FISH with each probe at least twice. Unfortunately, lots of preparations provided high background signals despite using pretreatment against high background, therefore it was necessary to perform FISH several times.

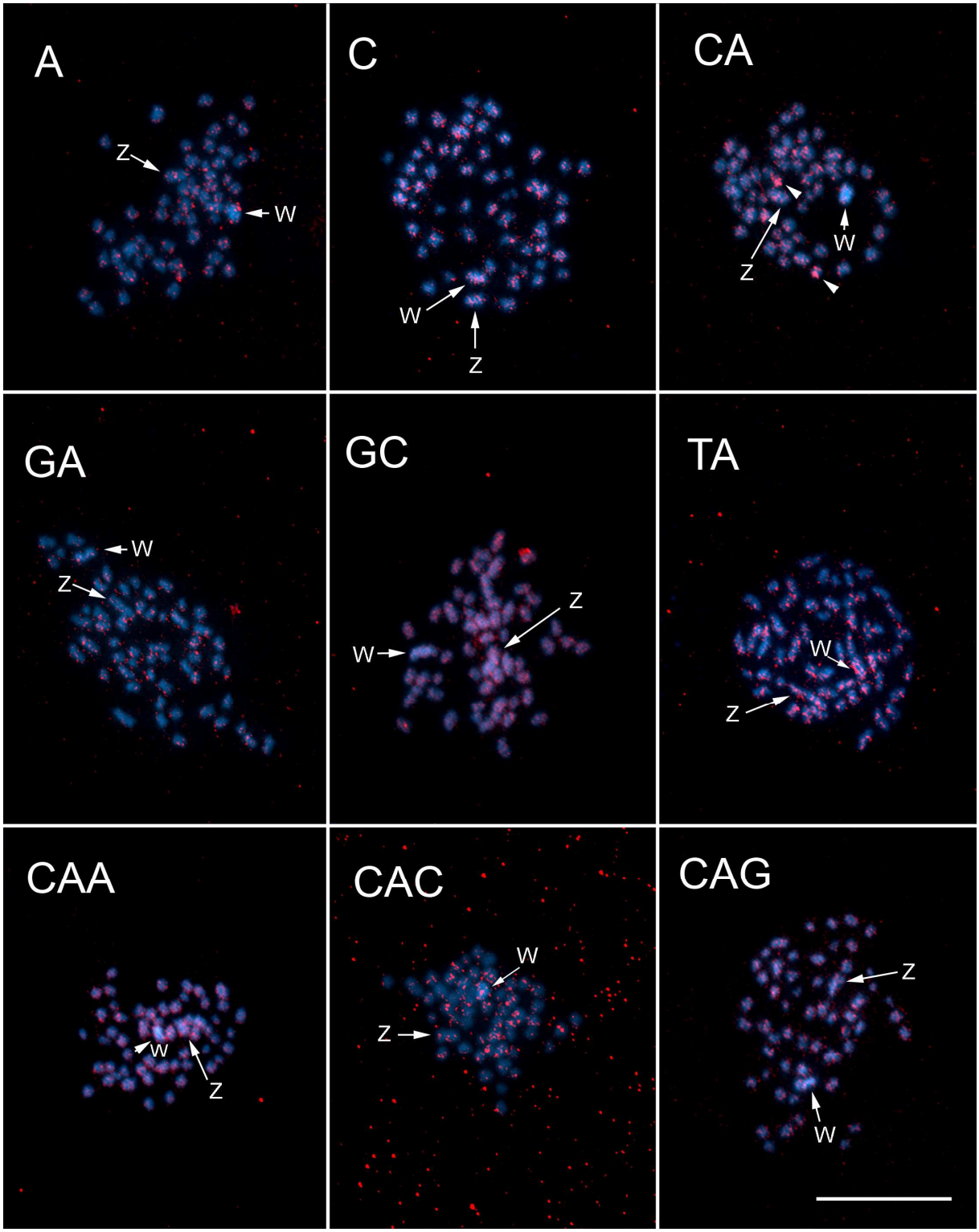

Fig.8a. FISH with microsatellite probes listed in the picture. Female *C. pomonella* chromosomes are in mitotic phase. Microsatellite probe is labelled with red fluorochrome, chromosomes are counterstained with DAPI. Scale bar = 20 μm

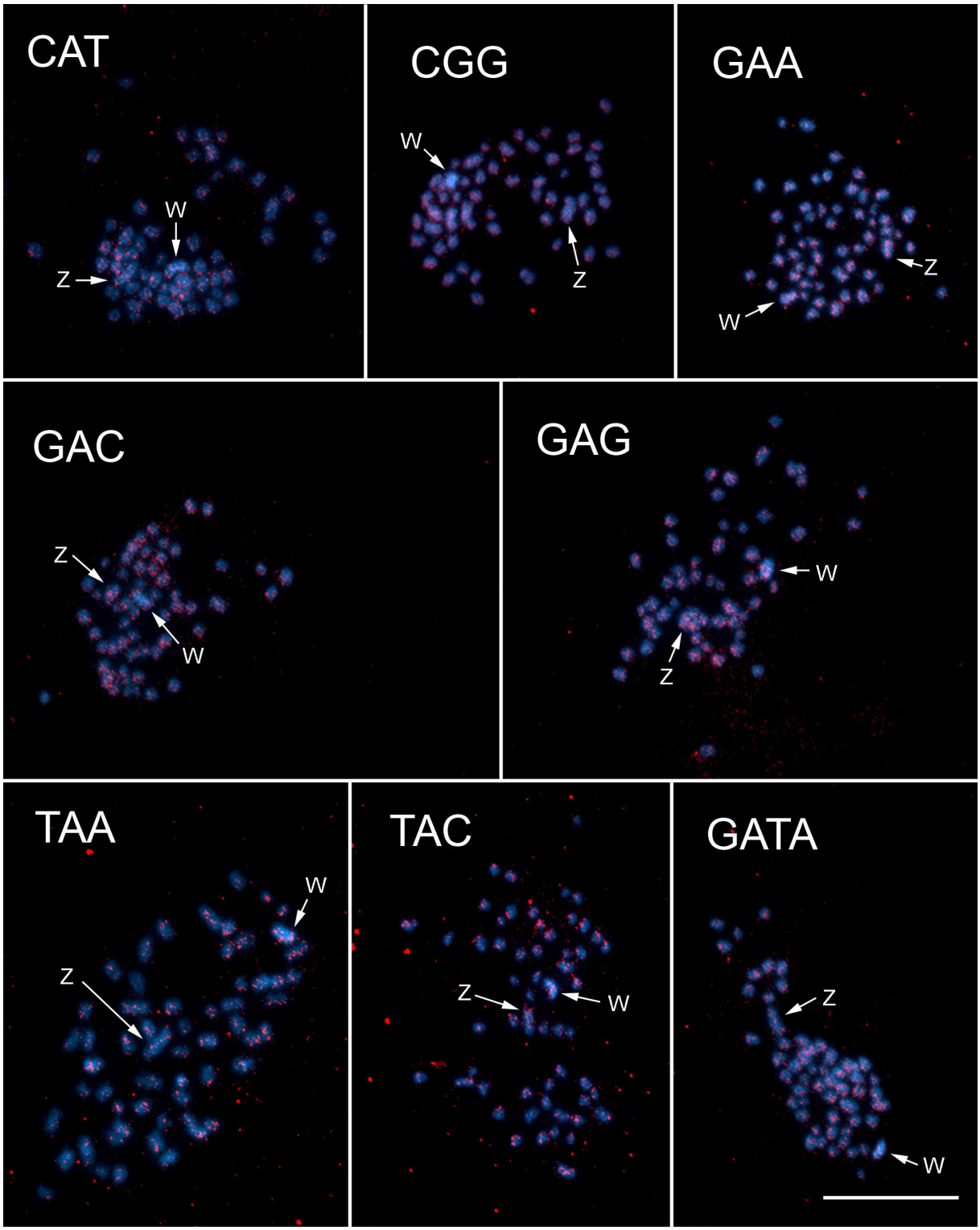

Fig. 8b FISH with microsatellite probes listed in the picture. Female *C. pomonella* chromosomes are in mitotic phase. Microsatellite probe is labelled with red fluorochrome, chromosomes are counterstained with DAPI. Scale bar = 20 μm

Tab.2. Evaluation of distribution pattern and signal intensity of the individual microsatellites

microsatellite probe	W/Z distribution	autosomal distribution, signal intensity, character of the signal
A	no difference compared to autosomes	Microsatellite is intensively distributed on several pairs of medium size chromosomes, where it rather forms discrete signals predominantly gathered in subtelomeric regions. It has the strongest signal of all microsatellite probes analysed.
C	no difference compared to autosomes	Microsatellite creates signal of very distinct little blocks, this pattern is especially visible at several pairs of chromosomes.
CA	signal seems avoid W	Microsatellite is preferentially distributed on one pair of the smallest chromosomes, signal creates evenly distributed little
GA	no difference compared to autosomes	Microsatellite creates signal of very tiny little blocks scattered throughout all chromosomes.
GC	no difference compared to autosomes	Microsatellite creates distinct little blocks on several chromosomes but it demonstrate rather dispersive pattern on the others.
TA	no difference compared to autosomes	Hybridization signal is very strong, almost comparable to the one formed by microsatellite A, microsatellite creates very distinct little blocks, this tendency is especially visible on several pairs of large chromosomes, where it forms conspicuous distinct little blocks.
CAA	no difference compared to autosomes	Microsatellite creates very distinct little blocks of the comparable size, signal is evenly distributed along all chromosomes.
CAC	no difference compared to autosomes	Microsatellite signal creates tiny little blocks evenly distributed along chromosomes.
CAG	no difference compared to autosomes	Microsatellite creates rather small little blocks unevenly distributed along individual chromosomes, signals often group into bigger little blocks
CAT	no difference compared to autosomes	Signal is of a dispersive character, rarely does it form little blocks.
CGG	no difference compared to autosomes	Signal creates very distinct little blocks distributed on all chromosomes.
GAA	no difference compared to autosomes	Signal creates discrete, not very distinct little blocks.
GAC	no difference compared to autosomes	Signal creates discrete, not very distinct little blocks.

GAG	no difference compared to autosomes	Signal is of a dispersive character, the intensity is lower at 4 pairs of the smallest chromosomes.
TAA	Signal has specific pattern on the W chromosome chromosome.	Microsatellite creates dispersive pattern evenly distributed along all chromosomes.
TAC	no difference compared to autosomes	Microsatellite forms one distinct little block or grouping of more little signals in subtelomeric part of two little and one medium size chromosome pairs.
GATA	Signal seems to have polar distribution on W chromosome, where it is preferentially distributed on one half of the chromosome.	Microsatellite creates dispersive signal more or less evenly distributed along all chromosomes.

5. Discussion

Ever since its discovery, the repetitive DNA attracted lots of attention. Its unusually high content and uneven distribution within eukaryotic genome implied its significant role in genome structure and evolution. Nowadays, as research of repetitive DNA advanced we can make some general conclusions about its characteristic and behaviour within the genome. Tandem arrays comprise substantial part of the eukaryotic genome. Their character and distribution is, for the most part, well recognized speaking of monocentric chromosomes (Palomeque and Lorite, 2008). However, the character and distribution in holokinetic system remain elusive due to the absence of heterochromatic hotspot represented by centromere in case of monocentric chromosomes.

In this thesis I aimed to isolate and characterise satellite DNA in the genome of codling moth. For this purpose I chose different approaches successfully used in previous publications.

5.1. GSP-PCR

In this study I attempted to selectively amplify and isolate satellite DNA in the genome of codling moth by means of GSP-PCR method. My research, however, did not bring any positive results in contrast to studies conducted previously. Buntjer and Lenstra (1998), who developed the method, successfully amplified satellite DNA in all tested organisms, namely horse, chicken, ostrich, dolphin, and cattle. Owing to their research a new satellite in ostrich was identified and isolated. The same conclusions reached Macas et al. (2000) who isolated two new families of tandem repeats in *Vicia* using this method.

The method may not be sensitive to short tendem repeat stretches. My findings may be explained by holokinetic character of lepidopteran chromosomes that in case of *C. pomonella* seem to lack bigger blocks of tandem arrays. I base these conclusions on dispersive distribution of satellite DNA in the codling moth genome as I proved in this work.

5.2. Distribution of satellite DNA in the genome of *C. pomonella*

Satellite sequence CPSAT-1 I managed to isolate is the second satellite DNA successfully isolated from the lepidopteran genome so far. The only Lepidopteran satellite sequence, MBSAT1, was isolated from the cell line of the cabbage moth *Mamestra brassicae* (Mandrioli et al., 2003). Distribution of this satellite in the *M. brassicae* cell line was limited to the heterochromatic regions of sex chromosomes. However, research of the

respective satellite sequence on the chromosomes coming from the wild population of *M. brassicae* near České Budějovice did not detect any positive signal (Magda Vítková, personal communication).

Satellite DNA in the genome of *C. pomonella* presented in this thesis is more or less evenly distributed through the genome and does not cluster or create any specific pattern. The only exception is the medium size chromosome pair where hybridization signals aggregate to a bigger block. The finding is in contrast to general fact that satellite DNA as a major part of heterochromatin often clusters in regions with very low recombination which is often represented by heteromorphic chromosomes (Charlesworth et al., 1994; Stephan and Cho, 1994).

Weaker hybridization signals in Dot blot at higher temperature observed in all species except *C. pomonella* evidences for lower sequence similarity rather than smaller amount of the satellite in the respective genomes. Interestingly, the most divergent sequence occurs in *E. kuehniella*, while less divergent sequence was found in phylogenetically more distant *P. operculella*. Lack of the hybridization signals in *Pieris brassicae*, *Polymmatus bellargus*, *Inachis io*, and *Lasiommata megera* implies that genomes of respective species devoid the satellite DNA entirely. The satellite sequence might have disappeared randomly during evolution or its sequence lost most of its original homology in extent unrecognizable by the probe used during hybridization.

It is common knowledge that monocentric chromosomes are characteristic for the presence of heterochromatin blocks located predominantly in centromere, subtelomeric, and telomeric regions, in NOR and in sex chromosomes (Ray and Venketeswaran, 1978). Some particular satellite DNA sequences are also known as a functional part of the centromeres in eukaryotes. Particularly satellite DNA binding kinetochore proteins in human is alpha-satellite DNA (Willard et al., 1987). Holokinetic chromosomes have dispersed centromere therefore the motif that interacts with kinetochore is dispersed in the similar way. However, considering random and non-regular distribution of satellite DNA within *C. pomonella* genome I conclude that the isolated satellite DNA CPSAT-1 does not have a function similar to that of human alpha-satellite DNA.

5.3. Microsatellite distribution in the genome of *C. pomonella*

Besides satellite DNA, microsatellite DNA is another member of tandemly organised repetitive DNA class. Character of the sequence and genomic distribution along with behaviour within the genome differ considerably between satellite and microsatellite repeats. Microsatellites are an object of extensive research due to characteristics it

presents. Character of distribution and variability in the length of microsatellite array predetermine them to be employed as a useful tool in various genetic applications, such are construction of genetic maps, various linkage analysis, population genetic studies etcetera (Dietrich et al., 1994; Dib et al., 1996; Schlotterer and Pemberton, 1998; Goldstein and Schlotterer, 1999). Research focusing on their further analysis is therefore of a major interest.

Studies dealing with microsatellite characterisation and distribution mostly focus on analysis of the information obtained from genome sequence project and BAC sequences of chosen organisms and physical distribution of microsatellites in the complement. Research focused on the distribution of microsatellite arrays in the holokinetic system is still scarce, limited only to survey of the genome sequence of *B. mori* (Prasad et al., 2004) and studies focused on analysis of microsatellites as markers for population genetic studies (Meglécz, 2004, 2007; Franck et al., 2005). Analysis of its physical distribution within holokinetic system may elucidate their function in the genome and help in understanding of the evolutionary dynamics of the holokinetic chromosomes. It may also shed lights on the structure of *C. pomonella* genome and push us forward in the research of lepidopteran cytogenetics. In the presented thesis I analyzed physical distribution of microsatellite in the *C. pomonella* complement.

There are several papers focusing on the microsatellite distribution in the genome of plants and animals. Surprisingly, their findings are contradictory. For instance, Charlesworth et al. (1994) reported microsatellites avoiding heterochromatic regions, which was supported later by research on sugar beet (Schmidt and Heslop-Harris, 1996). On the other hand, microsatellite (GAA) in barley prefers heterochromatin regions, and so do microsatellites in wolf fish (Cioffi et al. 2010).

Mapping of the microsatellite distribution in this study revealed several differences in hybridization of particular microsatellites. The overall picture, however, gives an impression of more or less uniform density of most microsatellites in the complement of *C. pomonella* compared to findings made in organisms with monocentric chromosomes (e.g. Kubát et al., 2008; Cioffi et al. 2010; Schmidt and Heslop-Harrison, 1996). Character of microsatellite distribution is rather dispersive, signals are more or less evenly scattered through the chromosomes and most microsatellites give non-specific pattern. My conclusions neither accept nor reject findings presented above since the distribution of most microsatellites in *C. pomonella* genome was more or less evenly dispersed along all chromosomes. This includes the sex chromosomes regardless the fact that W chromosome in case of *C. pomonella* is completely heterochromatic (Traut and Marec,

1996). Such finding opposes a theory that microsatellites may have function in sex chromosome evolution (e.g. Subramanian et al., 2003). Several other studies reported accumulation or specific distribution of microsatellites in sex chromosomes. Kubát et al. (2008) conducted research on microsatellite distribution in *Silene latifolia* and discovered accumulation of microsatellite DNA on Y chromosome and they support their findings by models predicting accumulation on non-recombining regions of the genome. Their findings also agree with observations of several above mentioned studies that microsatellites tend to aggregate in euchromatic part of chromosomes since *S. latifolia* Y chromosome is for the most part created by euchromatin with only small pericentric and subtelomeric heterochromatic blocks (Kubát et al., 2008).

Several authors observed chromosome specific pattern of tetranucleotide GATA in their survey of microsatellite distribution. Analysis of GATA microsatellite repeat in human genome proved predominant occurrence of this repeat on the sex chromosomes. Also distribution on the sex chromosomes was proven to be nonrandom. Y specific region almost devoided of GATA while the region homologous with X chromosome was rich in this repeat. Based on these observations, authors even suggested possible role of this repeat in higher order chromatin organization (Subramanian et al., 2003). Their observations drove me to include GATA tetreanucleotide in this study. My findings, however, did not indicate any preferential aggregation of GATA on sex chromosomes and rejected therefore the idea proposed by Subramanian in reference to *C. pomonella*.

Random distribution of microsatellites in *C. pomonella* genome can be caused by the presence and activity of transposable elements. Transposable elements often contain microsatellites and extension of the retrotranscripts has been suggested as well (Nadir et al., 1996). On the other hand, Abe et al. (2005) in the sequence analysis of W chromosome showed aggregation of transposable elements on the W chromosome and so suggested Fuková et al. (2007) in *C. pomonella* W chromosome. My findings, however, do not indicate any microsatellite aggregation on any of the sex chromosomes. The possible explanation is the absence of crossing-over in lepidopteran females and therefore absence of ectopic recombination, which is one of the mechanisms involved in tandem repeats elongation.

Despite evenly dense distribution and dispersed pattern of microsatellite repeats observed in this research several conclusions can be made about specificity of hybridization in several microsatellite repeats. Based on the intensity of hybridization signal an AT rich microsatellites were the most abundant ones, particularly A, and AT microsatellite repeats. These findings are in congruence with findings in the genome of

Arabidopsis thaliana (Cardle et al., 2000), *Saccharomyces cerevisiae* (Young et al., 2000), and *B. mori* (Prasad et al., 2004) as the only representative of the organism with holokinetic system where the characterisation of the microsatellites has been conducted. Genome analyses of the rice (Temnykh et al., 2001), maize (Chin et al., 1996), and human (Jurka and Pethiyagoda, 1995) contradicts these observations when showing abundance of GC-rich microsatellites.

All the findings mentioned above evidence for chromosome specific distribution of microsatellite DNA, which seems to be general feature of eukaryotic organisms (Schmidt and Heslop-Harrison, 1996). These findings also contribute to the idea that monocentric chromosomes show specific pattern compared to rather non-selective scattered pattern observed in this study, which may be caused right by the holokinetic state of the chromosomes.

The uniform microsatellite distribution can be attributed to holokinetic state of *C. pomonella* chromosomes presenting different evolutionary dynamics compared to chromosomes monocentric where the distribution was proven rather selective. Survey conducted by d'Alencon et al. (2010) proved high macrosynteny conservation and simultaneous high rate of local genome rearrangements in lepidopterans. The observation is unforeseen considering holokinetic state of its chromosomes proven to be more prone to rearrangements compared to chromosomes monocentric (Mola and Papeschi, 2006).

Another explanation of differential microsatellite distribution between monocentric and holokinetic chromosomes may refer to general characteristic of Eukaryotes often showing largely conserved synteny of coding sequences sometimes among distantly related taxonomic groups (Paterson et al., 1995; Shields, 1993; Moore, 1995). Repetitive DNA, however, is not under strong selective pressure and its evolutionary dynamics differs widely among closely related species (Schmidtt and Harrison, 1996). This study proves that evolutionary dynamics of the genome can be very unexpected and lots of further research needs to be done until some conclusions about this problematic will be made.

6. Summary

This thesis deals with the characterisation of repetitive DNA in the genome of codling moth, *Cydia pomonella*. Several approaches were used in order to isolate, characterize and localize different types of tandemly repeated DNA.

For the isolation of satellite DNA I used GSP-PCR, which failed to yield any satellite DNA, and restriction analysis of genomic DNA (gDNA). Restriction analysis managed to detect products in 14 out of 27 tested restriction endonucelases (RE). Products of three of them were successfully isolated and sequenced and subsequent Southern blot revealed tandem organisation of the product created by *Xba* I RE. The satellite was named CPSAT-1 since it is the first satellite DNA isolated from *C. pomonella*. CPSAT-1 was further submitted to Dot blot hybridization with gDNAs of various lepidopteran species. The analysis revealed that a sequence or sequences similar to CPSAT-1 occur in representatives of rather distant families except Papilionoidea, where it is completely missing

The physical distribution of the CPSAT-1 within *C. pomonella* genome was accomplished by means of FISH, which showed evenly dispersed pattern along all chromosomes. The one exception presented a pair of medium size chromosomes where a larger block of satellite grouping was observed.

The last approach was mapping of microsatellite distribution on *C. pomonella* chromosomes by means of FISH with microsatellite probes. Analysis revealed differences in distribution and intensity of several microsatellite probes, although most microsatellites did not show any specific pattern. Most of the microsatellites neither avoided nor preferred heterochromatic W chromosome including microsatellite GATA, which is believed to participate on sex chromosome evolution.

7. References

Abe, H., Mita, K., Yasukochi, Y., Oshiki, T., and Shimada, T., 2005a. Retrotransposable elements on the W chromosome of the silkworm, *Bombyx mori*. Cytogenet Genome Res 110: 144-151.

Agrawal, N., Dasaradhi, P.V.N., Mohmmed, A., Malhotra, P., Bhatnagar, R.K., and Mukherjee, S.K., 2003. RNA interference: biology, mechanism, and applications. Microbiol Mol Biol Rev 67: 657-685.

Akagi, H., Yokozeki, Y., Inagaki, A., Mori, K. and Fujimura, T., 2001. Micron, a microsatellite-targeting transposable element in the rice genome. Mol Genet Genomics 266: 471–480.

d'Alencon, E., Sezutsu, H., Legeai, F., Permal, E., Bernard-Samain, S., Gimenez, S., Gagneur, C., Cousserans, F., Shimomura, M., Brun-Barale, A., Flutre, T., Couloux, A., East, P., Gordon, K., Mita, K., Quesneville, H., Fournier, P., Feyereisen, R., 2010. Extensive synteny conservation of holocentric chromosomes in Lepidoptera despite high rates of local genome rearrangements. PNAS 107 (17): 7680-7685.

An, W., Han, J. S., Wheelan, S. J., Davis, E. S., Coombes. C. E., Ye, P., Triplett, C., Boeke, J. D., 2006. Active retrotransposition by a synthetic L1 element in mice. Proc Natl Acad Sci USA 103: 18662-18667.

Andrew, S. E., Goldberg, Y. P., Theilmann, J., Zeisler, J. R., Hayden, M. R., 1994. ACCG repeat polymorphism adjacent to the CAG repeat in the Huntington disease gene: implication for diagnostic accuracy and predictive testing. Hum Mol Genet 3:65-67.

Armour, J. A. L. and Jefreys, A., J., 1992. Biology and application of human minisatellite loci. Curr Opin Genet Dev 2: 850-856.

Arnold, R., Maueler, W., Bassili, G., Lutz, M., Burke, L., Epplen, T. J. & Renkawitz, R., 2000. The insulator protein CTCF represses transcription on binding to the (gt)22(ga)15 microsatellite in intron 2 of the HLA-DRB1*0401 gene. Gene 253: 209–214.

Ashley, C.T., and S,T. Warren. 1995. Trinucleotide repeat expansion and human disease. Annu Rev Genet 29: 703-728.

Babich, V., Aksenov, N., Alexeenko, V., Oei, S. L., Buchlow, G. & Tomilin, N., 1999. Association of some potential hormone response elements in human genes with the *Alu* family repeats. Gene 239: 341–349.

Barlow, D.P., 1993. Methylation and imprinting: from host defense to gene regulation? Science 260: 309–310.

Bates, G., Lehrach, H., 1994. Trinucleotide repeat expansions and human genetic disease. BioEssays 16: 277–284.

Bayliss, C. D., Field, D. & Moxon, E. R., 2001. The simple sequence contingency loci of *Haemophilus influenzae* and *Neisseria meningitidis*. J Clin Invest 107: 657–662.

Beckman, J. S. and Soller, M., 1990. Towards unified approach to the genetic mapping of eukaryotes based on sequence-tagged microsatellite sites. Bio/Technology 8: 930-932.

Berg, D. E. and Howe, M. M., 1989. Mobile DNA. American Society of Microbiology Publication, Washington, D.C.

Beritze T., 1986. Satellite DNA. Springer-Verlag. Berlin/Heidelberg, Germany.

Bird, T. D. 1994. Hereditary ataxia overview. GeneReviews. University of Washington, Seattle.

Bochukova, E. G., Roscioli, T., Hedges, D. J., Taylor, I. B., Johnson, D., David, D.J., Deininger, P. L., Wilkie, A. O.,2009. Rare mutations of FGFR2 causing Apert syndrome: identification of the first partial gene deletion, and an *Alu* element insertion from a new subfamily. Hum Mutation 30:204-211.

Britten, R.J. and Kohne, D.E., 1968. Repeated sequences in DNA. Science 161: 529-540.

Brosius, J., 1999a. Vertebrate genomes were forged by massive bombardments with transposable elements and retrotransposones. Genetica 107: 209-238.

Buntjer, J. B., Lenstra, J. A., 1998. Self-amplification of satellite DNA in vitro. Genome 41: 429-434.

C. elegans Sequencing Consortium. 1998. Genome sequence of the nematode *C. elegans*: a platform for investigating biology. Science 282(5396): 2012-2018.

Capon, D. J., Chen, E. Y., Levinson, A. D., Seeburg, P. H., Goedder, D. V., 1983. Complete nucleotide sequence of the T24fhuman bladder carcinoma oncogene and its normal homologue. Nature 302: 33-37.

Cardle, L., L. Ramsay, D. Milbourne, M. Macaulay, D. Marshall et al., 2000. Computational and experimental characterization of physically clustered simple sequence repeats in plants. Genetics 156**:** 847–854.

Cary, L. C., Goebel, M., Corsaro, B. G., Wang, H. G., Rosen, E., Fraser, M. J., 1989. Transposon mutagenesis of baculoviruses: analysis of *Trichoplusia ni* transposon

IFP2 insertions within the FP-locus of nuclear polyhedrosis viruses. Virology 172(1):156-169.

Chamberlain, N. L., Driver, E. D., Mjesfeld, R. L, 1994. The length and location of CAG trinucleotide repeat in the androgen receptor N-terminal domain affect transactivation function. Nucleic Acids Res 22:3181-3186.

Charlesworth, B., Sniegowsky, P., Stephan, W., 1994. The evolutionary dynamics of repetitive DNA in Eukaryotes. Nature 371: 215-220.

Chin, E. C., M. L. Senior, H. Shu and J. S. Smith, 1996 Maize simple repetitive DNA sequences: abundance and allelic variation. Genome 39**:** 866–873.

Cioffi, M. B., Martins, C., Bertollo, L. A. C., 2009. Comparative chromosome mapping of repetitive sequences. Implications for genomic evolution in the fish, *Hoplias malabaricus.* BMC Genetics 10: 34.

Cleary, J. D., Nichol, K., Wang, Y.-H. & Pearson, C. E., 2002. Evidence of cis-acting factors in replication-mediated trinucleotide repeat instability in primate cells. Nature Genet 31: 37–46.

Cleary, J.D. & Pearson, C. E., 2003. The contribution of cis-elements to disease-associated repeat instability : clinical and experimental evidence. Cytogenet Genome Res 100: 25–55.

Coffee, B., Zhang, F., Warren, S.T. & Reines, D., 1999. Acetylated histones are associated with FMR1 in normal but not fragile X-syndrome cells. Nature Genet 22: 98–101.

Croager, E. J., Gout, A.M. & Abraham, L. J., 2000. Involvement of Sp1 and microsatellite repressor sequences in the transcriptional control of the human CD30 gene. Am J Pathol 156: 1723–1731.

Csink, A. K. and Henikoff, S., 1998. Spmething from nothing: the evolution and utility of satellite repeats. Tig May 14(5): 200-204.

Darvasi, A. & Pisente-Shalom, A., 2002. Complexities in the genetic dissection of quantitative trait loci. Trends Genet 18: 489–491.

Dib, C., Faure, S., Fizames, C., Samson, D., Drouot, N., Vignal, A., Millasseau, P., Marc, S., Hazan, J., Seboun, E., Lathrop, M., Gyapay, G., Morissette, J., Weissenbach, J., 1996. A comprehensive genetic map of the human genome based on 5,264 microsatellites. Nature 380:152–154

Dietrich, W. F., Miller, J. C., Steen, R. G., Merchant, M., Darmon, D., Nahf, R., Gross, A., Joyse, D. C., Wessel, M., Dredge, R. D., Marquis, A., Stein, L. D., Goodman, N., Page, D. C., Lander, E. S., 1994. A genetic map of the mouse with 4,006 simple sequence length polymorphism. Nat Genet 7: 220-245.

Dimitri, P., Corradini, N., Rossi, F. and Vernì, F., 2005. The paradox of functional heterochromatin. BioEssays 27: 29-41.

Dimitri, P., Vernì, F., Mei, E., Rossi, F. and Corradini, N., 2005. Transposable elements as artisans of the heterochromatic genome. Cytogenetics and Genome Research 110: 165-172.

Efstratiadis, A., Crain, W. R., Britten, R. J., Davidson, E. H., Kafatos, F. C., 1976. DNA sequence organisation in the lepidopteran *Antheraea pernyi*. Proc Natl Acas Aci 73 (7): 2289-2296.

Fagerberg, A. J., Fulton, R. E. and Black, W. C., 2001. Microsatellite loci are not abundant in all arthropod genomes: analyses in the hard tick, *Ixodes scapularis* and the yellow fever mosquito, *Aedes aegypti*. Insect Mol Biol 10: 225–236.

Ferreira, I. A. and Martins, C., 2008. Physical chromosome mapping of repetitive DNA sequences in Nila tilapia *Orechromis niloticus*: Evidence for a differential distribution of repetitive elements in the sex chromosomes. Micron 39: 411-418.

Fishel, R., Lescoe, M. K., Rao, M. R., Copeland, N. G., Jenkins, N. A., Garber, J., Kane, M., Kolodner, R, 1993. The human mutator gene homolog MSH2 and its association with hereditary nonpolyposis colon cancer. Cell 75:1027-1038.

Flavell, R. B., Bennett, M. D., Smith, J. B., Smith, D. B., 1994. Genome size and the proportion of repeated nucleotide sequence DNA in plants. biochem. Genet. 12: 257-269.

Franck, P., Guérin, F., Loiseau, A., Sauphanor, B., 2005. Isolation and characterisation of microsatellite loci in the codling moth, *Cydia pomonella* L. (Lepidoptera, Tortricidae). Mol Ecol Notes 5: 99-102.

Fuková, I., Nguyen, P., Marec, F., 2005. Codling moth cytogenetics: karyotype, chromosomal location of rDNA, and molecular differentiation of sex chromosomes. Genome 48 (6): 1083-1092.

Fuková, I., Traut, W., Vitkova, M., Nguyen, P., Kubickova, S., Marec, F., 2007. Probing the W chromosome of the codling moth, *Cydia pomonella*, with sequences from microdissected sex chromatin. Chromosoma 116: 135-145.

Fukova I., Neven L.G., Barcenas N.M., Gund N.A., Dalikova M., Marec F., 2009. Rapid assessment of the sex of codling moth *Cydia pomonella* (Linnaeus) (Lepidoptera: Tortricidae) eggs and larvae. J. Appl. Entomol. 133: 249-261.

Gentles, A. J., Wakefield, M. J., Kohany, O., Gu, W., Batzer, M. A., Pollock, D. D., Jurka, J., 2007. Evolutionary dynamics of transposable elements in the short-tailed opossum *Monodelphis domestica*. Genome Res 17: 992-1004.

Glaser, R. W., 1917. 'Ringer' solutions and some notes on the physiological basis of their ionic composition. Comp Biochem Physiol 2: 241-289.

Goldstein, D. B. & Schlötterer, C. 1999. Microsatellites: Evolution and Applications. Oxford University Press, New York, 368pp.

Green, H., 1993. Human genetic diseases due to codon reiteration: relationship to an evolutionary mechanism. Cell 74: 955-956.

Han, J. S., Boeke, J. D., 2004. A highly active synthetic mammalian retrotransposon. Nature 429: 314-318

Handa, V., Saha, T. & Usdin, K., 2003. The fragile X syndrome repeats form RNA hairpins that do not activate the interferoninducible protein kinase, PKR, but are cut by Dicer. Nucleic Acids Res 31: 6243–6248.

Henikoff, S., Ahmad, K., Malik, H. S., 2001. The centromere paradox: stable inheritance with rapidly evolving DNA. Science 293: 1098-1102.

Hoekstra, R., CriadoFornelio, A., Fakkeldij, J., Bergman, J. and Roos, M. H., 1997. Microsatellites of the parasitic nematode *Haemonchus contortus*: polymorphism and linkage with a direct repeat. Mol Biochem Parasitol 89: 97–107.

Holdridge, CH. and Dorsett, D., 1991. Repression od *hsp70* heat shock gene transcription by the suppressor of Hairy-wing protein *Drosophila melanogaster*. Mol Cell Biol 11: 1894-1900.

Horak, M. 2006. Olethreutinae moths of Australia (Lepidoptera: Tortricidae). Monographs on Australian Lepidoptera 10. 522pp.

Hughes-Schrader, S., Ris, H., 1941. The diffuse spindle attachment of coccids, verified by the mitotic behavior of induced chromosome fragments. J Exp Zool: 87(3): 429-456.

Hughes-Schrader, S., Scgrader, F., 1961. The holokinetic kinetochore of the Hemiptera. Chromosoma 12: 327-350.

Imagawa, M., Ishikawa, Y., Shimano, H., Osada, S. & Nishihara, T., 1995. CTG triplet repeat in mouse growth inhibitory factor/metallothionein III gene promoter represses the transcriptional activity of the heterologous promoters. J Biol Chem 270: 20898–20900.

International Human Genome Sequencing Consortium, 2001. Initial sequencing and analysis of the human genome. Nature 409: 860–921.

Ivics, Z., Hackett, P. B., Plasterk, R. H., Izsvák, Z., 1997. Molecular reconstruction of *Sleeping Beauty*, a Tc1-like transposon from fish, and its transposition in human cells. Cell 91: 501-510.

Jasinka, A., Michelewski, G., De Mezer, M., Sobczak, K. & Kozlowski, P., Napierala, M. & Krzyzosiak, W. J., 2003. Structures of trinucleotide repeats in human transcripts and their functional implications. Nucleic Acids Res 31: 5463– 5468.

John, B. 1988. The biology of heterochromatin, In: V. R S., editor. Heterochromatin: Molecular and structural aspect. Cambridge University Press. pp. 1-128.

Johnson, P.C.D., Webster, L.M.I., Adam, A., Buckland, R., Dawson, D.A. and Keller, L.F., 2006. Abundant variation in microsatellites of the parasitic nematode *Trichostrongylus tenuis* and linkage to a tandem repeat. Mol Biochem Parasitol 148: 210–218.

Jones, P.A. and Takai, D**.,** 2001. The role of DNA methylation in mammalian epigenetics. Science 293: 1068–1070.

Jordan, I.K. and W.M. Miller, 2008. Genome defense against transposable elements and the origins of regulatory RNA. in Genome Dynamics and Stability, (Lankenau, D.H. and S. Scharz eds.) Springer, Berlin.

Jorge J. Yunis, J. J. and Walid G. Yasmineh, W. G., 1971. Heterochromatin, Satellite DNA, and Cell Function. Science 174(4015): 1200-1209.

Jorgensen, 1997. Alphoid repetitive DNA in human chromosomes. Dan Med Bull 44(5): 522-534.

Joseph, J. L., Sentry, J. W., Smyth, D. R., 1990. Interspecies distribution of abundant DNA sequences in*Lilium* . J Mol Evol 30(2): 146-154.

Jeffreys, A. J., Wilson, V., Thein, S. L., 1985. Hypervariable minisatellite regions in human DNA. Nature 314: 67-73.

Jurka, J., and C. Pethiyagoda, 1995. Simple repetitive DNA sequences from primates: compilation and analysis. J. Mol. Evol 40**:** 120–126.

Kasperczyk, A., DiMartino, N. A., Krontiris, T. G., 1990. Minisatellite allele diversification: the origin of rare alleles of the HRASJ locus. Am J Hum Genet 47: 854-859.

Kazazian, H. H. Jr, Wong, C., Youssoufian, H., Scott, A. F., Phillips, D. G., Antonarakis, S.E., 1988. Haemophilia A resulting from de novo insertion of L1 sequences represents a novel mechanism for mutation in man. Nature 332: 164-166.

Kidwell, M. G., Lisch, D. R., 2000. Transposable element as a source of variation in animals and plants. Proc Natl Acad Sci USA 94: 7704-7711.

Kidwell, M. G., Lisch, D., 2001. Transposable elements, parasitic DNA, and genome evolution. Evolution 55: 1-24.

King, D. G., Soller, M. & Kashi, Y., 1997. Evolutionary turning knobs. Endeavour 21: 36–40.

Kingsbury, D. T., 1969. Estimate of the genome size of various microorganisms. J Bacteriol 98(3): 1400–1401.

Klose, R. J. and Bird, A. P., 1999. Genomic DNA methylation: the mark and its mediators. Trends Bioch Sci 31(2): 89-97.

Komai, F., 1999. A taxonomic review of the genus Grapholita and allied genera (Lepidoptera: Tortricidae) in the Palearctic region. Entomologica Scandinavica, Supplement 55: 1–226 .

Kristensen, N. P., Scoble, M. J., Karsholt, O., 2007. Lepidoptera phylogeny and systematics: the state of inventorying moth and butterfly diversity. Zootaxa 1668: 699-747.

Krontiris, T. G., Derlin, B., Karp, D. D., Robert, N. J., Risch, N., 1993. An association between the risk of cancer and mutations in the HRASJ minisatellite locus. N Engl J Med 329: 517-523.

Krontiris, T. G., 1995. Minisatellites and human disease. Science 269: 1682-1683.

Kubát, Z. Hobza, R., Vyskot, B.,Kejnovský, E., 2008. Microsatellite accumulation on the Y chromosome in *Silene latifolia*. Genome 51: 350-356.

Leach, F. S., Nicolaides, N. C., Papadopoulos, N., Liu, B., Jen, J., Parsons, R., Peltomaki, P., Sistonen, P., Aaltonen, L. A., Lystrom-Lahti, M. et al., 1993. Mutations of a MutS homolog in hereditay nonpolyposis colorectal cancer. Cell 75:1215-1225.

Lewis, S. M. and Wu, G. E., 1997. The origins of W(D)J recombination. Cell 88(2): 159-162.

Li, Y., Korol, A. B., Fahima, T., belies, A. and Nevo, E., 2002. Microsatellites: genomic distribution, putative functions and, mutational mechanisms: a review. Mol Ecol 11: 2453-2465.

Litt, M., Luty, J. A.,1989. A hypervariable microsatellite revealed by *in vitro* amplification of a dinucleotide repeat within the cardiac muscle actin gene. Am J Hum Genet 44: 397-401.

Macas, J., Požárková, D., Navrátilová., Nouzová, M., Neumann, P., 2000. Two new families of tandem repeats isolated from genus *Vicia* using genomic self-priming PCR. Mol Gen Genet 263: 741-751.

Makee, H., Tafesh, N. & Harrown, H., 2008. Mapping of pachytene bivalents of female codling moth *Cydia pomonella* (Linnaeus) (Lep., Tortricidae). J Appl Entomol 134: 182-190.

Makee, H., Saour, G., 1999. Non-recovery of fertility in partially sterile male *Phthorimaea operculella* (Lepidoptera:Gelechiidae). J. Econ. Entomol. 92: 516–520.

Mandrioli, M., Mnicardi, G. C., Marec, F., 2003. Cytogenetic and molecular characterization of the MBSAT1 satellite DNA in holokinetic chromosomes of the cabbage moth, *Mamestra brassicae* (Lepidoptera). Chromosome Res. 11: 51-56.

Marec F., 1990. Genetic control of pest Lepidoptera: Induction of sex-linked recessive lethal mutations in *Ephestia kuehniella* (Pyralidae). Acta Entomol Bohemoslov 87: 445-458.

McGrath, C. L. and Katz, L. A., 2004. Genome diversity in microbial eukaryotes. Trends Ecol Evol 19: 32-38.

Meglécz, E., Petenian, F., Danchin, E., d'Acier A. R., Rasplus, J. and Faure, E., 2004., High similarity between flanking regions of different microsatellites detected within each of two species of Lepidoptera: *Parnassius apollo* and *Euphydryas aurinia.* Mol Ecol 13: 1693-1700.

Meglécz, E., Anderson, S. J., Bourguet, D., Butcher, R., Caldas, A., Cassel-Lundhagen, A., Coeur d'Acier, A., Dawson, D. A., Faure, N., Fauvelot, C., Franck, P., Harper, G., Keyghobadi, N., Kluetsch, C., Muthulakshmi, M., Nagaraju, J., Patt, A., Petenian, F., Silvain, J. F. and Wilcock, H. R. (2007). Microsatellite flanking region similarities among different loci within insect species. Insect Molecular Biology 16(2): 175–185.

Meltzer, S. J., Yin, J., Manin, B., Rhyu, M. G., Cottrell, J., Hudson, E., Redd, J. L., Krasna, M. J., Abraham, J. M., Reid, B. J., 1994. Microsatellite instability occurs frequently and in both diploid and aneuploid cell populations of Barrett's- associated esophageal adenocarcinoma. Cancer Res 54: 3379-3382.

Metzgar D., Bytof J., Wills C., 2000. Selection against frameshift mutations limits microsatellite expansion in coding DNA. Genome Res 10: 72–80.

Mita, K., Kasahara, M., Sasaki, S., Nagayasu, Y., Yamada, T., Kanamori, H., Namiki, N., Kitagawa, M., Yamashita, H., Yasukochi, Y., et al., 2004. The genome sequence of Silkworm, *Bombyx mori*. DNA Res 11: 27-35.

Miller, W. J., Donald, J. F., Pinsker, W., 1997. Molecular domestication of mobile elements. Genetica 100: 261-270.

Mola, L. M. & Papeschi, A. G., 2006. Holokinetic chromosomes at a glance. J Basic Appl Genet 17(1): 17-33.

Montini, E., Held, P. K., Noll, M., Morcinek, N., Al-Dhalimy, M., Finegold, M., Yant, S. R., Kay, M. A., Grompe, M., 2002. In vivo correction of murine tyrosinemia type I by DNA mediated transposition. Mol Ther 6: 759-769.

Moore, G., Devos, K. M., Wang, Z., Gale, M. D., 1995. Cereal genome evolution, grasses line up and form a circle. Curr Biol 5: 737-739.

Morgante,M., and A. M. Olivieri. 1993. PCR-amplified microsatellites as markers in plant genetics. Plant J 3: 175-182.

Mutanen, M., Wahlberg, N., Kaila, L., 2010. Comprehensive gene and taxon coverage elucidates radiation patterns in moths and butterflies. Proc R Soc Lond 277(1695): 2839-2848.

Nadir, E., Margalit, H., Gallili, T. and Ben-Sasson, S. A., 1996. Microsatellite spreading in the human genome: evolutionary mechanisms and structural implications. Proc Natl Acad Sci U. S. A. 93: 6470-6475.

Nakabachi, A., Yamashita, A., Toh, H., Ishikawa, H., Dunbar, H., Moran, N., Hattori, M. 2006. "The 160-kilobase genome of the bacterial endosymbiont *Carsonella*.". Science 314(5797): 267.

Nakamura, Y., Leppert, M., O'Connell, P., Wolff, R., Holm, T., Culver, M., Martin, C., Fujimoto, E., Hoff, M., Kumlin, E., et al., 1987. Variable number of tandem repeats (VNTR) markers for human gene mapping. Science 235: 1616-1622.

Narita, N., Nishio, H., Kitoh, Y., Ishikawa, Y., Ishikawa, Y., Minami, R., Nakamura, H., Matsuo, M., 1993. Insertion of a 5' truncated L1 element into the 3' end of exon 44 of the dystrophin gene resulted in skipping of the exon during splicing in a case of Duchenne muscular dystrophy. J Clin Invest 91: 1862-1867.

Nguyen P., Sahara K., Yoshido A., Marec F. (2010) Evolutionary dynamics of rDNA clusters on chromosomes of moths and butterflies (Lepidoptera). Genetica 138: 343–354.

Nordenskiold, H., 1963. A Study of meiosis in Luzula purpurea. Hereditas 48: 503-519.

Notter-Hausman, C., Dorn, S., 2010. Relationship between behavior and physiology in an invasive pest species: oviposition site selection and temperature-dependent development of the oriental fruit moth (Lepidoptera: Tortricidae). Environ Entomol 39: 561-569.

O'Donovan, M. C., Guy, C., Craddock, N., Murphy, K. C., Cardno, G., Jones ,L. A., Owen, M. J., McGuffin, P., 1995. Expanded CAG repeats in schizophrenia and bipolar disorders. Nature Genet 10: 380-381.

O'Keefe, CH. L., Warburton, P. E. and Matera, G., 1996. Oligonucleotide probes for alpha satellite DNA variants can distinguish homologous chromosomes by FISH. Hum Mol Gen 5(11): 1793-1799

Oldridge, M., Zackai, E. H., McDonald-McGinn, D. M., Iseki, S., Morriss-Kay, G. M., Twigg, S. R., Johnson, D., Wall, S. A., Jiang, W., Theda, C., et al.,1999. De novo aluelement insertions in FGFR2 identify a distinct pathological basis for Apert syndrome. Am J Hum Genetics 64: 446-461.

Orgel, L. E., Crick F. H. C. & Sapienza, C., 1980. Selfish DNA. Nature 288: 645-646.

Ortiz-Urda, S, Lin, Q., Yant, S. R., Keene, D., Kay, M. A., Khavari, P. A., 2003. Sustainable correction of junctional epidermolysis bullosa via transposon-mediated nonviral gene transfer. Gene Ther 10: 1099-1104.

Osanai-Futahashi, M., Suetsuga, Y., Mita, K., Fujiwara, H., 2008. Genome.wide screening and characterization of trasnposable elements and their distribution analysis in the silkworm, *Bombyx mori*. Insect Biochem Mol Biol 38: 1046-1057.

Ostertag, E. M., Kazazian, H. H. Jr., 2001. Biology of mammalian L1 retrotransposons. Ann Rev Genetics 35: 501-538.

Palomeque, T. and Lorite, P., 2008. Satellite DNA in insects: a review. Heredity 100: 564-573.

Paterson, A. H., Lin, Y. -R., Li, Z., Schertz, K. F., Doebley, J. F., Pinson, S. R. M., Liu, S. -C., Stansel, J. W., Irvine, J. E., 1995. Convergent domestication of cereal crops by independent mutations at corresponding genetic loci. Science 269: 1714-1718.

Pimpinelli S, Berloco M, Fanti L et al., 1995. Transposable elements are structural components of Drosophila melanogaster heterochromatin. Proc Natl Acad Sci USA 92: 3804-3808.

Prasad, M. D., and J. Nagaraju, 2004. A comparative phylogenetic analysis of full-length *mariner* elements isolated from the Indian tasar silkmoth, *Antheraea mylitta* (Lepidoptera: Saturniidae). J Biosci 28**:** 443–453.

Queller, D. C., Strassman,J. E. & Hughes, C. R., 1993. "Microsatellites and Kinship". Trends Ecol Evol 8: 285–288.

Ramel, C., 1997. Mini- and Microsatellites. Environ Health Perspect 105(4): 781-789.

Ramsay, L., Macaulay, M., Cardle, L., Morgante, M., degli Ivanissevich, S., Maestri, E., et al., 1999. Intimate association of microsatellite repeats with retrotransposons and other dispersed repetitive elements in barley. Plant J 17: 415–425.

Ray, J. H. and Venketeswaran, S., 1978. Constitutive heterochromatin distribution in monocentric and polycentric chromosomes. Chromosoma 66: 341 -350.

Ren, Z., Jin, H., Whitby, P. W., Morton, D. J. & Stull, T.L., 1999. Role of CCAA nucleotide repeats in regulation of hemoglobin and hemoglobin-haptoglobin binding protein genes of *Haemophilus influenzae*. J Bacteriol 181: 5865–5870.

Ris, H., 1942. A cytological and experimental analysis of the meiotic behavior of the univalent X-chromosomes in the bearberry aphid *Tamalia* (=Phyllaphis) *coweni* (Ckll.). J Exp Zool 90: 267-326.

Robertson, H. M. and Engels, W. R., 1989. Modified P elements that mimic the P cytotype in *Drosophila melanogaster*. Genetics 123**:** 815-823.

Rohrer, J., Minegishi, Y., Richter, D., Eguiguren, J., Conley, M. E., 1999. Unusual mutations in Btk: an insertion, a duplication, an inversion and four large deletions. Clin Immun Orlando 90: 28-37.

Sahara, K., Marec, F., Traut, W., 1999. TTAGG telomeric repeat in chromosomes of some insect and other arthropods. Chromosome Res 7: 449-460.

SanMiguel, P., Tikhonov, A., Jin, Y. K., Motchoulskaia, N., Zakharov, D., Melake-Berhan A., Springer, P. S., Edwards, K. J., Lee, M., Avramova, Z. and Jeffrey L. Bennetzen 1996. Nested retrotransposons in the intergenic regions of the maize genome. Science 274: 765–768.

Schlotterer, C., Pemberton, J., 1998. The use of microsatellites for genetic analysis of natural populations – a critical review. In: DeSalle, R., Schierwater, B. (eds). Molecular approaches to ecology and evolution. Birkhauser, Basel, pp 71-86.

Schmidt, T. and Heslop-Harrison, J. S., 1996. The physical and genomic organisation of microsatellites in sugar beet. Proc Natl Acad Sci 93: 8761-8765.

Schwaiger, F. & Epplen, J., 1995. Exonic MHC-DRB polymorphisms and intronic simple repeat sequences: Janus' faces of DNA sequence evolution. Immunological Reviews 143: 199–224.

Shapiro, J.A., 1999a. Natural genetic engineering, adaptive mutation and bacterial evolution, pp. 259–275, in Microbial Ecology and Infectious Disease, edited by E. Rosenberg, ASM Press, Washington.

Shapiro, J.A., 1999b. Genome system architecture and natural genetic engineering in evolution. Annal NY Acad Sci 870: 23–35.

Shields, R. 1993. Pastoral synteny. Nature 365: 297-298.

Smit, A.F.A., 1996. The origin of interspersed repeats in the humangenome. Curr Opin Genet Dev 6: 743–748.

Smith, P. G., 1976. Evolution of repeated sequences by unequal crossover. Science 191: 528-535.

Stephan, W., Cho, S., 1994. Possible role of natural selection the formation of tandem-repetitive noncoding DNA. Genetics 136: 333-341.

Strand, M., Prolla, T. A., Liskay, R. M., Petes, T. D, 1993. Destabilization of tracts of simple repetitive DNA in yeast by mutations affecting DNA mismatch repair. Nature 365: 274-276.

Subramanian, S., Mishra, R. K., Singh, L., 2003. Genome-wide analysis of Bmk sequence (GATA repeats): predominant association with sex chromosomes and potentioal role in higher order chromatin organisation and function. Bioinformatics 19(6): 681-685.

Sumner, A.T., 1972. A simple technique for demonstrating centromeric heterochromatin. Exp. Cell Res. 75: 304–306.

Tamura, K., Dudley, J., Nei, M., Kumar, S. 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol Biol Evol 24: 1596-1599.

Tautz, D., 1989. Hypervariability of simple sequences as a general source for polymorphic DNA markers. Nucleic Acids Res 17: 6463-6471.

Temnykh, S., DeClerck, G., Lukashova, A., Lipovich, L., Cartinhour, S. and McCouch, S., 2001. Computational and experimental analysis of microsatellites in rice (*Oryza sativa* L.): Frequency, length variation, transposon associations, and genetic marker potential. Genome Res 11: 1441–1452.

Thomas, C. A. Jr, 1971. "The genetic organization of chromosomes". Annual Review of Genetics 5: 237–256.

Torrez-Cruz, J. & Van der Woude, M. W., 2003. Slippedstrand mispairing can function as a phase variation mechanism in *Escherichia coli*. J Bacteriol 185: 6990–6994.

Traut, W., and Marec, F. 1996. Sex chromatin in Lepidoptera. Q Rev Biol 71: 239–256.

Traverse, K.L., Pardue, M.L., 1988. A spontaneously opened ring chromosome of *Drosophila melanogaster* has acquired He-T DNA sequences at both new telomeres. Proc Natl Acad Sci USA 85: 8116--8120.

Ugarkovic, D., 2005. Functional elements residing within satellite DNAs. EMBO reports 6: 1035-1039.

Ugarkovic, D. and Plohl, M., 2002. Variation in satellite profiles – causes and effects. The EMBO Journal 21(22): 5955-5959.

Van't Hof, A. E., Brakefield, P. M., Saccheri, I. J., Zwaan, B. J., 2007. Evolutionary dynamics of multilocus microsatellite arrangements in the genome of the butterfly *Bicyclus anynana*, with implications for other Lepidoptera. Heredity: 1-9.

Wallace, M. R., Andersen, L. B., Saulino, A. M., Gregory, P. E., Glover, T. W., Collins, F. S., 1991. A de novo *Alu* insertion results in neurofibromatosis type 1. Nature 353: 864-866.

Waterston, R. H., Lindblad-Toh, K., Birney, E., Rogers, J., Abril, J. F., Agarwal, P., Agarwala, R., Ainscough, R., Alexandersson, M., An, P., et al., 2002. Initial sequencing and comparative analysis of the mouse genome. Nature 420: 520-562.

Weber, J. L., May, P. E., 1989. Abundant class of human DNA polymorphism which can be typed using the polymerase chain reaction. Am J Hum Genet 44: 388-396.

Wilder, J. and Hollocher, H., 2001. Mobile elements and the genesis of microsatellites in dipterans. Mol Biol Evol 18: 384–392.

Willard, H. F. and Waye, J. S., 1987. Hierarchical order in chromosome-specific human alpha satellite DNA. Trends Genet 3: 192-198.

Yant, S.R., Meuse, L., Chiu, W., Ivics, Z., Izsvák, Z., Kay, M. A., 2000. Somatic integration and long-term transgene expression in normal and haemophilic mice using a DNA transposon system. Nat Genet 25: 35-41.

Yoder,J.A., Walsh,C.P., and Bestor,T.H., 1997. "Cytosine methylation and the ecology of intragenomic parasites." Trends Genet 13: 335–340.

Yoshida, K., Nakamura, A., Yazaki, M., Ikeda, S., Takeda, S., 1998. Insertional mutation by transposable element, L1, in the DMD gene results in X-linked dilated cardiomyopathy. Hum Molec Genetics 7:1129-1132.

Young, E. T., J. S. Sloan and K. Van Riper, 2000. Trinucleotide repeats are clustered in regulatory regions in *Saccharomyces cerevisiae.* Genetic