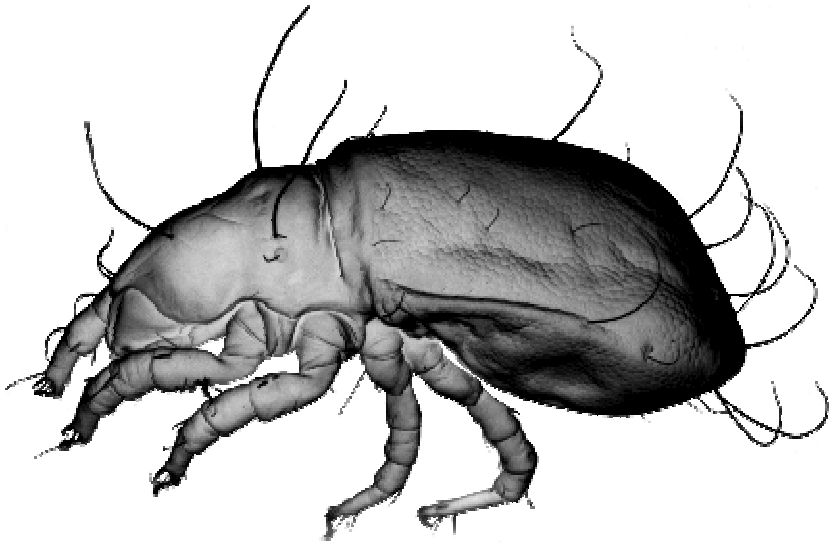
Genetic diversity and evolutionary age of parthenogenetic oribatid mites (Acari: Oribatida)



Mucronothrus nasalis

Dem Fachbereich Biologie der Technischen Universität Darmstadt
zur Erlangung des akademischen Grades
Doctor rerum naturalium
genehmigte Dissertation
von

Michael Heethoff
aus Wiesbaden

Berichterstatter: Prof. Dr. Stefan Scheu
Mitberichterstatter: Prof. Dr. Ulrich Göringer, Prof. Dr. Roy A. Norton
Tag der Einreichung: 22. Oktober 2003
Tag der mündlichen Prüfung: 16. Dezember 2003

Darmstadt 2004

D17

“All truths are easy to understand once they are discovered; the point is to discover them.”

Galileo Galilei

“Wissenschaft ist „auf der Suche sein“, nicht „gefunden haben“.”

Mark Maraun

Die vorliegende Arbeit wurde unter Leitung von Prof. Stefan Scheu am Institut für Zoologie der Technischen Universität Darmstadt durchgeführt und unter dem Geschäftszeichen SCHE 376/9 von der Deutschen Forschungsgemeinschaft (DFG) gefördert.

Lebenslauf

Persönliche Daten	Michael Heethoff Geboren 10. Mai 1973 in Wiesbaden Ledig	
Studium	**1994 – 1996**	Grundstudium der Biologie an der TU Darmstadt
	1996 – 1999	Hauptstudium der Zoologie an der TU Darmstadt mit den Nebenfächern Ökologie, Tierphysiologie und Organische Chemie (Gesamtnote: Sehr gut) Thema der Diplomarbeit „Genetische Variabilität bei der parthenogenetischen Hornmilbe *Platynothrus peltifer*“
	1999	Auszeichnung mit dem Förderpreis für „Hervorragende wissenschaftliche Leistung“ der Freunde der TU Darmstadt
	2000 - 2003	Promotion an der TU Darmstadt mit dem Themenschwerpunkt „Genetische Diversität pathenogenetischer Bodenorganismen“ Gefördert von der Deutschen Forschungsgemeinschaft.
	2003	Disputation mit Auszeichnung
	Seit 2002	Kollegiat im Graduiertenkolleg 340 „Communication in biological systems – From the molecule to the organism in its environment“
	Seit 2003	Gewählter studentischer Sprecher des GRK 340
Tätigkeiten	**1992 – 1996**	Werkstudent bei der Kühlschmierstoff-Firma Houghton Deutschland mit dem Arbeitsgebiet Qualitätssicherung, Produktentwicklung und elektronische Datenverarbeitung
	1996 – 2003	Betreuung folgender Hochschul-Praktika ♦ Grundpraktikum zur Anatomie der Tiere ♦ Grundpraktikum zur Ökologie und Kenntnis einheimischer Fauna ♦ Grundpraktikum Tierphysiologie „Osmotische Resistenz von Säugererythrozyten“ ♦ Hauptpraktikum Ökologie „Limnologie“ ♦ Hauptpraktikum Ökologie „Marine Lebensräume“ ♦ Hauptpraktikum Ökologie „Terrestrische Lebensräume“ ♦ Hauptpraktikum Zoologie „Prinzipien der Phylogenie“ mit Schwerpunkt auf molekularer Phylogenie ♦ Großpraktikum Zoologie „Anatomie und Physiologie der Arthropoda“ ♦ Betreuung verschiedener Forschungspraktika, Examens- und Diplomarbeiten zu Thema „Molekulare Evolution“
Schulbildung	**1979 – 1983**	Grundschule Bad Camberg-Erbach
	1983 – 1985	Förderstufe Bad Camberg
	1985 – 1989	Gymniasaler Zweig der Taunusschule Bad Camberg
	1989 – 1992	Oberstufe des Pestalozzi-Gymnasiums Idstein
	1992	Abitur

Zivildienst	**1993**	Jugendherberge Limburg / Lahn
	1994	Mobiler Pflegedienst der Caritas Bad Camberg
Zusatzqualifikationen	**2003**	„Argumentation und Verhandlung“. Anwendung des Harvard-Konzeptes. Fortbildungsprogramm der TU Darmstadt
	2003	„Das Wissen der Forschung – verständlich für Laien“. Seminar für Öffentlichkeits- und Pressearbeit der DFG.
Sprachkenntnisse	**Englisch**	Sehr gute Kenntnisse in Wort und Schrift
	Französisch	Grundkenntnisse
EDV-Kenntnisse	**MS-Office**	Sehr gute Kenntnisse in Word, Excel und Powerpoint
	Phylogenie	Sehr gute Kenntnisse im Umgang mit verschiedensten Software-Paketen wie z.B. PAUP, MEGA, DAMBE etc.
	Netzwerk	Gute Kenntnisse zu Netzwerkaufbau und –administration
	Sonstige	Sehr gute Kenntnisse in der Bildbearbeitung, Videobearbeitung, HTML-Erstellung und Systempflege

Eidesstattliche Erklärung

Ich erkläre hiermit an Eides statt, dass ich die vorliegende Dissertation selbständig und nur mit den angegebenen Hilfsmitteln angefertigt habe. Ich habe noch keinen Promotionsversuch unternommen.

Darmstadt, den 22.10.2003

(Michael Heethoff)

Teile der vorliegenden Arbeit sowie anderer Projekte während der Promotionszeit wurden bisher wie folgt publiziert bzw. auf Konferenzen präsentiert:

Heethoff, M., Maraun, M., Norton, R. A., Scheu, S. (2003). No need for sex? Evidence for ancient asexual lineages of oribatid mites. *J. Evol. Biol.*, submitted.

Heethoff, M., Etzold, K., Scheu, S. (2003). Mitochondrial COII sequences indicate that the parthenogenetic earthworm *Octolasion tyrtaeum* (Savigny 1826) constitutes of two lineages differing in body size and genotype. Pedobiologia 47, in press.

Maraun, M., **Heethoff, M.**, Scheu, S., Norton, R. A., Weigmann, G., Thomas, R. H. (2003). Radiation in sexual and parthenogenetic oribatid mites (Oribatida, Acari) as indicated by genetic divergence of closely related species. Exp. Appl. Acarol. 29: 265-277.

Heethoff, M., Laumann, M., Domes, K., Maraun, M., Norton, R. A., Scheu, S. (2003). No need for sex? Evolution without recombination in oribatid mites. ESEBIX: 13.8P. (Poster)

Maraun, M., **Heethoff, M.**, Schneider, K., Scheu, S., Weigmann, G., Cianciolo, J., Thomas, R. H., Norton, R. A. (2003). Molecular phylogeny of oribatid mites (Oribatida, Acari): Evidence for multiple radiations of parthenogenetic lineages. Exp. Appl. Acarol., in review.

Maraun, M., **Heethoff, M.**, Schneider, K., Scheu, S., Thomas, R. H., Norton, R. A. (2003). Radiation in ancient asexuals: Oribatid mites (Acari). ESEBIX: 16.4.

Heethoff, M., Maraun, M., Scheu, S. (2002). Ancient parthenogenetic species: Methods for detection and a candidate from oribatid mites. Zoology 105: 25.

Maraun, M., **Heethoff, M.**, Scheu, S. (2002). Parthenogenetic radiation and the re-evolution of sex. Evidence from molecular phylogeny of oribatid mites. Zoology 105: 93.

Heethoff, M. (2002). Ist Parthenogenese eine Sackgasse der Evolution? Berichte der AG Bodenmesofauna 18, 16-18.

Heethoff, M., Maraun, M., Scheu, S. (2000). Asexuals in soil-living „Oribatida“: ancient or recent? (Poster)

Heethoff, M., Maraun, M., Scheu, S. (2000). Genetic variability in ribosomal ITS 1-sequences of the parthenogenetic oribatid mite *Platynothrus peltifer* (C.L. Koch, 1839) (Acari: Oribatida). Ber. Nat.-med. Verein Innsbruck 87: 339-354.

Contents

Summary

Theories on the evolution and maintenance of sex are challenged by the existence of ancient asexual lineages like bdelloid rotifers and darwinulid ostracods. Ancient asexual origin of several speciose taxa has also been proposed for oribatid mites (Acari).

In the present work, I analysed genetic divergences in different parthenogenetic and sexual lineages of oribatid mites to estimate the age of these lineages and to evaluate if the evolution of oribatid mites was without recombination.

Age estimations were based on genetic distances of the mitochondrial gene for cytochrome oxidase I (COI) between clonal lineages of different parthenogenetic oribatid mite species. Species from two parthenogenetic taxa (*Platynothrus peltifer*, *P. yamasakii*, *P. sibiricus, Heminothrus thori* and *Camisia horrida* (Camisiidae) and *Mucronothrus nasalis* (Trhypochthoniidae)) were sampled in different sites in Europe, North America, South America, Australia and Asia. Corrected distances reached up to 157% on DNA level and 10% on amino acid level. Using different calibrations for a molecular clock, an age of up to 240 million years since the separation of the lineages was estimated. With this age, oribatid mites contain the oldest living parthenogenetic metazoan taxa.

Further analyses investigated if the high evolutionary age of *Platynothrus* and other oribatid mite taxa (*Mucronothrus nasalis*, *Tectocepheus velatus*,

Atropacarus striculus, *Steganacarus magnus* and *Metabelba pulverulenta*) was attained without gene recombination. To address this question, the two alleles of the *hsp82* gene (coding for the 82 KD heat shock protein) were analysed for their genetic divergence. Due to gene recombination, sexual species in general have maximum divergences of 4% for this gene; higher amounts of divergence indicate absence of gene recombination as expected for parthenogenetic organisms. Genetic distances were up to 70% on DNA level and 15% on amino acid level indicating ancient lack of recombination. The absence of recombination in *hsp82* was estimated to have happened 350 million years ago. The amount of genetic divergences between the alleles was in the same range in sexual and asexual oribatid mite species.

High allelic divergence between *hsp82* alleles in sexual and asexual species may be explained by the special mechanism of inheritance in oribatid mites: chromosomes are holokinetic and inverted meiosis has been inferred. If inverted meiosis occurs together with chiasma terminalisation gene recombination is suppressed in both sexual and parthenogenetic species. Together with terminal fusion in parthenogenetic species this mechanism resembles mitotic cloning.

Inverted meiosis is a simple form of meiosis and it is discussed if this mechanism might be ancestral to “normal” meiosis. A scenario is proposed on how “normal” meiosis and monocentric chromosomes evolved from inverted meiosis and holokinetic chromosomes.

Particular conditions may favour the frequent independent evolution and persistence of parthenogenetic clusters in oribatid mites. The absence of gene recombination in sexual and asexual oribatid mite species for at least 350 million years and the evolution of perhaps 100,000 oribatid mite species without gene recombination is an evolutionary mystery contradicting theories on sex and recombination and is unique in the animal kingdom.

Chapter One

1. Introduction

1.1. Oribatid mites

Oribatid mites (Acari, Oribatida) are a speciose group of mainly soil living invertebrates with about 10,000 described species (Schatz 2002), a conservatively estimated total number of 50,000 species (Travé et al. 1996) or maybe even 100,000 (Schatz 2002). The first indisputable fossil records of oribatid mites are from Devonian sediments deposited at least 380 million years ago (Shear et al. 1984; Norton et al. 1988) but the origin of the group presumably is older, about 400-440 million years (Lindquist 1984). About 16% of the oribatid mite species show a cosmopolitan distribution; these species presumably predated the breakup of Pangea which was about 200 million years ago and kept their distinct morphology (Hammer and Wallwork 1979).

Oribatid mites are important decomposers (Lussenhop 1992) in forest ecosystems, fallows, fields and meadows with densities up to 500,000 per square meter in acidic soils of northern boreal forests (Maraun and Scheu 2000). A strong co-evolution between oribatid mites and fungi was hypothesised (Wallwork 1983; Bernini 1986) and there is trophic niche differentiation concerning feeding preferences between different oribatid mite species (Scheu and Falca 2000; Maraun et al. 2003a).

The reproduction rate of oribatid mites is low compared to other soil microarthropods (Travé et al. 1996; Maraun 1997), due to their longevity oribatid mites are presumably generally iteropar (Mitchell 1977).

Oribatid mites are divided in six groups: Palaeosomata, Enarthronota, Parhyposomata, Circumdehiscentiae (=Brachypylina), Mixonomata and Nothroidea (=Desmonomata) (Grandjean 1969). The first four groups are assumed to be monophyletic while Mixonomata and Desmonomata are probably paraphyletic (Norton et al. 1993). A state-of-the-art phylogeny of oribatid mite groups is shown in Figure 1.1 (Maraun et al. submitted).

The monophyly of oribatid mites is questioned in general due to the possible origin of the Astigmata within oribatid mites (OConnor 1984; Norton 1994). OConnor (1984) postulated that the Astigmata originated within Desmonomata and Norton (1994) assumed that they are a sister taxon of the Trhypochthonoidea, an exclusively asexual group within the Desmonomata. Recent molecular studies do not support these hypotheses, but they sustain the idea that the Astigmata evolved within oribatid mites (Maraun et al. submitted) and that oribatid mites therefore are a paraphyletic group.

Parthenogenesis in oribatid mites was first reported in 1941 (Grandjean 1941). It is a widespread phenomenon in these mites and there is morphological and molecular evidence for radiations of several speciose asexual monophyletic taxa (Norton and Palmer 1991; Palmer and Norton 1992; Maraun et al. 2003b, Maraun et al. submitted). Important groups of Desmonomata with asexual radiations are the monophyletic Trhypochthonoidea comprising 68 asexual species, the Camisiidae (92 spp.), Nanhermanniidae (56 spp.), Malaconothridae (104 spp.) and the genus *Nothrus* (54 spp.). All these taxa lack closely related sexual species and are presumably ancient (Norton and Palmer 1991). In most asexual oribatid mite species, sterile males are produced at low frequencies (Grandjean 1941). This is a widespread phenomenon among obligate parthenogens (Lynch 1984). Due to the sterlility of males (Taberly 1988) and fixed heterozygosity in asexual oribatid mites (Palmer and Norton 1992) sexual reproduction presumably is absent despite the sporadic presence of males.

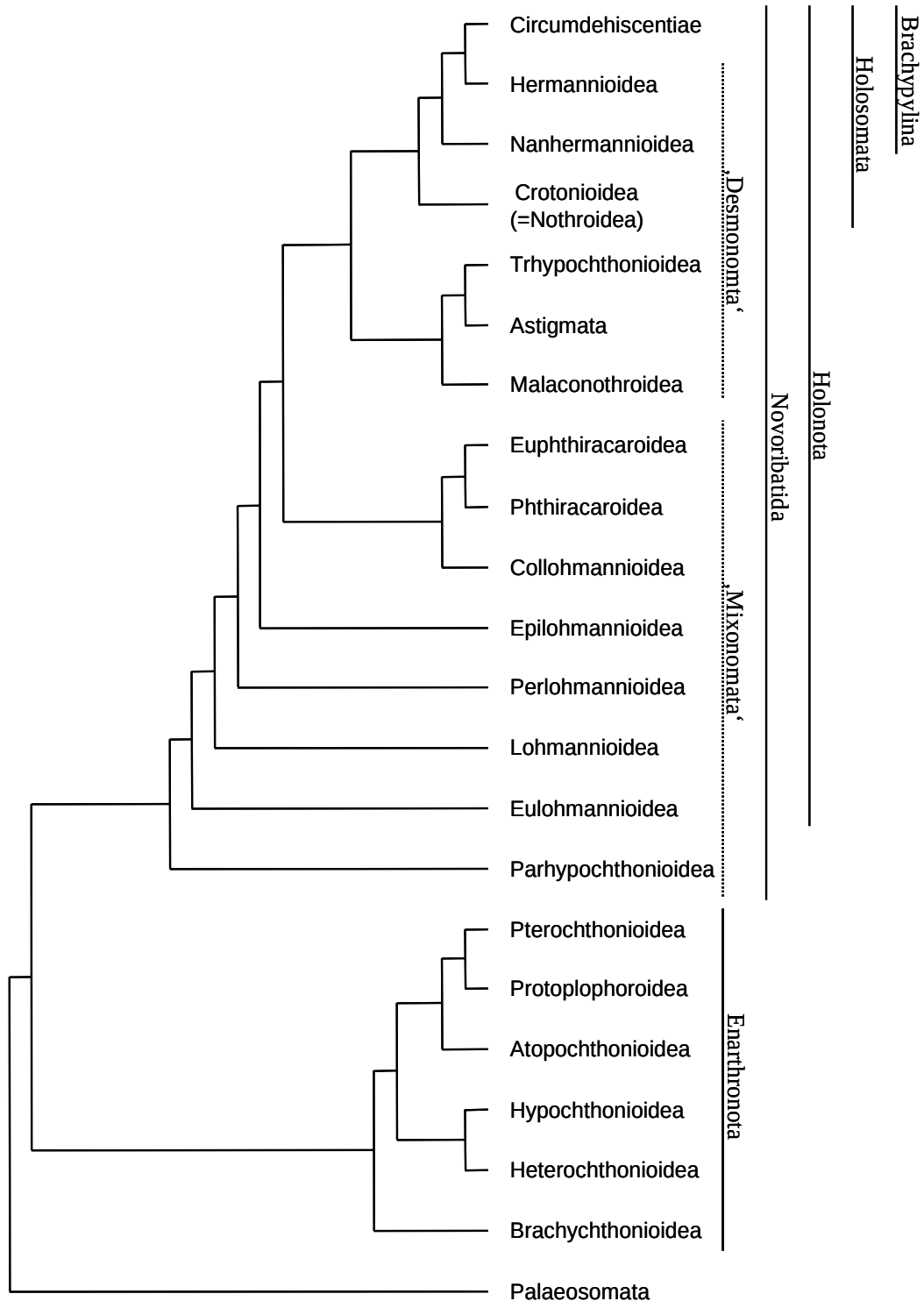


Figure 1.1
Phylogenetic relationships of the main oribatid mite groups based primarily on the hypothesis by Haumann (1991), with modifications according to Weigmann (1996) and including the Astigmata as suggested by Norton (1994) (Maraun et al. submitted).
Palaeosomata, Enarthronota, Circumdehiscentiae and Parhyposomata are monophyletic, Mixonomata and Desmonomata are paraphyletic groups.

1.2. The aim of this work

The ancient absence of sexual reproduction challenges evolutionary theories on sex. Asexual lineages are assumed to be short-lived and doomed to extinction on evolutionary small timescales (Maynard Smith 1978). Just a few groups of organisms are known to contradict theses theories and reproduce asexually for millions of years (Normark et al. 2003; Milius 2003). These organisms may serve as model systems in many questions of evolutionary biology.

Oribatid mites contain several speciose taxa comprising many asexual species without sexual relatives (Norton and Palmer 1991). The group is hardly investigated by evolutionary biologists and there are few molecular studies regarding the asexual status of oribatid mites.

Different questions are addressed in this work: (i) How old are asexual lineages in different oribatid mite taxa? (ii) Is recombination absent? (iii) Is the evolutionary success of asexual oribatid mites based on a special genetic mechanism?

After an introduction about theories on sex and parthenogenesis at the end of this chapter, the first question will be addressed in chapter two. The worldwide genetic divergences of a part of the mitochondrial gene for cytochrome oxidase I (COI) on both, DNA and protein levels in different species of Camisiidae and Trhypochthoniidae were analysed. Using different methods for the calibration of a molecular clock for COI, divergences indicate an evolutionary age of up to 240 million years for the studied oribatid mite species. The second question on the absence of recombination will be addressed in chapter three, where intraindividual allelic comparisons of a highly conserved gene (82 kD heat shock protein: *hsp82*) will be used to show that recombination in oribatid mites presumably has been ceased for about 350 million years. Chapter four deals with the genetic mechanism of inheritance in oribatid mites. The evolution of meiosis will be discussed in general with a focus on evolutionary consequences of holokinetic chromosomes and inverted meiosis known from oribatid mites.

1.3. Sex and parthenogenesis

Sexual reproduction is the predominant form of reproduction in eukaryotes and both probably evolved together (Cavalier-Smith 2002). Eukaryotes probably originated about 2.0-3.5 billion years ago (Miyamoto and Fitch 1996); recent calculations date the origin of eukaryotes and sex back to 2.5 billion years (Gu 1997). Eukaryotes without sexual reproduction therefore have abandoned sex at some time in their evolution due to hybridisation, cytological dysfunction or bacterial infection (Lynch 1984; Hurst et al. 1993).

The first records of parthenogenesis were described from aphids at the end of the 17th century (Leeuwenhoek 1695). Parthenogenetic organisms were construed as abnormal and incomplete individuals (Steenstrup 1842) until the late 19th century when cytological studies by August Weismann clarified mechanisms of parthenogenesis (Lynch 1984). Parthenogenetic species can be found in almost all groups of organisms (about 2000 parthenogenetic species are known, Milius 2003); even within vertebrates, some 50 species reproduce without sex (Lynch 1984). Almost all organisms are assumed to have a certain potential for asexual reproduction (White 1973).

Besides the widespread occurrence of parthenogenesis it is a reproductive strategy with a patchy taxonomic distribution, presumably due to long-term costs that offset the short-term numerical advantages (Maynard Smith 1978; Bell 1982; Kondrashov 1993; Butlin et al. 1999; Butlin 2002). The short-term advantages are twofold due to the absence of males and it's been called the "queen of problems in evolutionary biology" to explain why most eukaryotic organisms reproduce sexually (Bell 1982). By 1993, at least 20 theories had been proposed to explain the widespread occurrence of sexual reproduction (Kondrashov 1993). Most theories fall into one of two categories: either sex increases the rate of adaptive evolution (short-term advantages) by generating new gene combinations, or it prevents the accumulation of deleterious mutations (long-term advantages) (Butlin 2002; Figure 1.2). None of the theories is broadly

accepted. Therefore, a pluralist approach to sex and recombination was suggested (West et al. 1999) combining aspects of different theories.

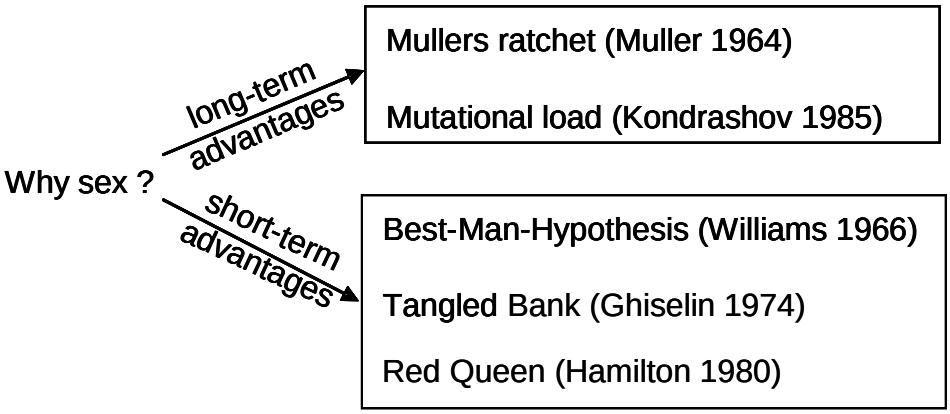


Figure 1.2
Schematical overview of theories on the evolution and maintenance of sexual reproduction.
Mullers ratchet: favourable genotypes and mutation free genotypes may be lost in the long-term by genetic drift (Muller 1964). Mutational load: Due to the absence of recombination, asexual organisms accumulate deleterious mutations (Kondrashov 1985).
Best-Man-Hypothesis: Sex provides genetic variability for progeny in a changing environment (Williams 1966). Tangled Bank: Sex provides genetic variability in capricious environments (Ghiselin 1974; Bell 1982). Red Queen: Sex provides genetic variability to compete with predators or parasites (van Valen 1973; Hamilton 1980).

Whatever theories state, parthenogentic lineages are assumed to be short-lived and evolutionary dead ends (Muller 1964; Maynard Smith 1978; Kondrashov 1988). Few exceptions contradict this dogma of evolutionary biology and were called “evolutionary scandals” (Maynard Smith 1978).

Known ancient asexual groups are the bdelloid rotifers (363 species, Mark Welch and Meselson 2000), the darwinulid ostracods (26 species, Martens et al. 2003) and probably several speciose groups of asexual oribatid mites (Norton and Palmer 1991; Maraun et al. 2003b; Heethoff et al. submitted). In general, there is little empirical data on the age of parthenogenetic lineages (Judson and Normark 1996; Schön et al. 1996,

1998; Sandoval et al. 1998; Butlin et al. 1999; Mark Welch and Meselson 2000; Butlin 2002; Normark et al. 2003) but bdelloid rotifers may have abandoned sex for 80 million years (Butlin 2002), darwinulid ostracods for 200 million years (Martens et al. 2003) and some oribatid mites for about 240 million years (Hammer and Wallwork 1979).

Parthenogenetic reproduction is defined by the development of an individual from an unfertilised egg. Parthenogenesis is a collective term for different genetic mechanisms. Some hymenopterans and mesostigmatid mites undergo arrhenotoky where unfertilised eggs develop into haploid males and fertilised eggs develop into diploid females (Stearns 1987). Another form of parthenogenesis is thelytoky where bisexual reproduction is completely absent. Apomictic thelytoky describes the development of a female from an unfertilised egg without reduction of chromosome number and meiosis (Hughes 1989) and is realised by many plants and animals (Suomalainen et al. 1987). Automictic thelytoky includes some kind of meiosis in oogenesis and is more common in animals than in plants (Suomalainen et al. 1987). In automictic thelytoky, there is some potential for intrachromosomal recombination but recombination is usually suppressed (Lynch 1984).

In parthenogenetic oribatid mites oogenesis is automictic, diploidy is restored by terminal fusion and inverted meiosis has been inferred (Wrensch et al. 1994; see also chapter four). This set of features genetically mimics apomixis and rules out recombination (Palmer and Norton 1992; Wrensch et al. 1994).

Chapter Two

2. Age of parthenogenetic oribatid mite lineages

2.1. Introduction

Mitochondria own a separate genome which is almost independent from the nuclear genome of the cell. Mitochondrial genomes show large size variations, ranging from 14 kb to 450 kb (Lewin 1991). Despite this variation, the gene content of the mitochondrial genome is highly conserved; size differences are due to variations in noncoding regions (Harrison 1989) or multiple repeats of some genes (Stanton et al. 1994). Metazoan mitochondria comprise in general 36 or 37 genes: 2 for ribosomal RNA, 22 for tRNAs, 13 for subunits of multimere proteins (e.g. subunits I-III of the cytochrome oxidases) and a control region of variable size (Boore and Brown 1995; Harrison 1989).

Mitochondria are uniparentally inherited, maternal in animals, paternal in some plants, and do not recombine (Horak et al. 1974; Giles et al. 1980; but see Rokas et al. 2003). This fact and the presence of different evolutionary rates in different regions of the mitochondrial genome make it a powerful tool for phylogenetic and evolutionary investigations (Harrison 1989; Wolstenholme 1992; Simon et al. 1994; Lunt et al. 1996). As mitochondria are clonally inherited, evolutionary rates are not expected to be influenced by reproductive mode (Schön et al. 1998) and should therefore be the same in both sexual and asexual species.

Especially the genes coding for the cytochrome oxidase subunits I and II (COI and COII) are commonly used regions to address phylogenetic and evolutionary questions (e.g. Simon et al. 1994; Lunt et al. 1996; Salomone et al. 2002; Heethoff et al. 2003). The COI gene is an auspicious candidate for calculations of divergence times (e.g. Knowlton 1993; Sandoval et al. 1998; Andersen et al. 2000; Salomone et al. 2002). Evolutionary rates of COI differ slightly in different taxonomic groups, but on average they seem to evolve in a clock-like manner in arthropods (DeSalle et al. 1987; Brower 1994) with a divergence rate of 2-2.3% corresponding to an evolutionary rate of 1-1.15% per million years. This rate was also used for COI divergence time estimations for oribatid mites (Salomone et al. 2002).

Cytochrome c oxidase is an indispensable enzyme which is found in all organisms which perform aerobic respiration (Adkins et al. 1996). It is the terminal and, possibly, the rate limiting (Poyton et al. 1988) component of the mitochondrial respiratory chain and is embedded in the inner membrane of the organelle (Schmidt et al. 1997). In its active form in mammals it is a dimeric enzyme composed of two monomers, each of which contains 3 subunits (I-III) encoded by the mitochondrial genome and 10 subunits encoded by the nuclear genome (Cooper et al. 1991). The essential role played by COI-III is indicated by the fact that these are the only subunits with homologs in both eukaryotes and prokaryotes. Only a few eukaryotes lack mitochondria and are strictly anaerobic (Müller 1988). Cytochrome c oxidase performs a four-electron reduction of oxygen to water in conjugation with the transfer of protons into the intermembrane space during the final stage of electron transfer. The electrochemical gradient formed by this process is an intermediate step in the conversion of redox energy to ATP (Castresana et al. 1994; Taanman and Williams 2001).

The insect COI protein comprises 511 amino acids which are arranged in 25 regions: besides the NH2- and COOH-termini there are 12 transmembrane domains, 6 external and 5 internal loops (Clary and Wolstenholme 1985; Lunt et al. 1996). The COOH-terminal is the most

variable region of the protein, amino acids in the reaction centres are highly conserved but they do not dominate the entire COI molecule, allowing scope for considerable variability in some regions (Liu and Beckenbach 1992; Lunt et al. 1996). There is no significant rate variability between transmembrane domains, internal or external loops but there are differences in the mean variability of different regions of the same structural class with the external loop E4, the internal loop I1 and the transmembrane domains M2, M6 and M8 being the most conserved regions (Lunt et al. 1996).

The extent of DNA sequence variation between parthenogenetic lineages reflects time since the split of these lineages (Avise 1994). By analysing worldwide divergences of the COI gene and its corresponding protein in monophyletic asexual taxa and different calibrations of a molecular clock I aimed to date the split of the different oribatid mite lineages and duration of asexual reproduction.

2.2. Materials and methods

Sampling of species

Five species of Camisiidae representing the three principal genera, including 61 specimens of *Platynothrus peltifer* from 18 sites in North America, Europe and Asia, and 11 specimens from other Camisiidae (*Platynothrus yamasakii*, *P. sibiricus, Camisia horrida* and *Heminothrus thori*), and also 23 specimens of *Mucronothrus nasalis* (Trhypochthoniidae) from 6 sites in North and South America, Europe, Asia and Australia were analysed (Figure 2.1, Table 2.1). All analysed species reproduce exclusively via parthenogenesis.

Platynothrus peltifer is distributed in soils of the whole Palaearctic region (Karppinen 1958; Dalenius 1960) with broad niche adaptations (Siepel 1990). *P. peltifer* feeds on decomposing litter, wood or fungi (Hartenstein 1962) with fungi being probably the most important food resource (Luxton 1972; Maraun 1997) and was classified like most other oribatid mites as sapro-mycophagous (Vera-Ziegler et al. 1990). More recent studies using

stable isotopes (^{15}N) indicate that *P. peltifer* is primary saprophagous (Scheu and Falca 2000).

Reproduction is by automictic thelytoky (Taberly 1987) with 1-4 eggs laid per clutch (Grandjean 1950) which are laid once a year between March and September (Harding 1971). Therefore, *P. peltifer* has a K-style reproductive biology despite its parthenogenetic reproduction.

Mucronothrus nasalis lives in a very distinctive habitat: it is always found in cold springs or icy melt-water at high altitudes (Hammer and Wallwork 1979). Due to its wide distribution and its restriction to freshwater habitats it has been assumed that the distribution of extant populations is due to continental drift (Hammer and Wallwork 1979); a conclusion which was raised by Hammer (1965) and subsequently supported by morphological studies (Travé 1971, 1973). Reproduction seems to be distributed over the whole year with May and June being the months with the lowest reproduction rate (Norton et al. 1988). As a trhypochthonoid mite, *M. nasalis* is assumed to reproduce also by automictic thelytoky (Taberly 1987). The feeding behaviour of *M. nasalis* is not well studied; diatoms, filamentous algae, organic particles and fungal hyphae were detected by gut content analyses (Norton et al. 1988).

Figure 2.1
Origin of analysed species. Grey dots: Camisiidae, black dots: *Mucronothrus nasalis.*

Table 2.1
Origin of the analysed species; n=numbers of analysed specimens

Species	Origin	n	Abbreviation
Platynothrus peltifer	USA, Washington D. C.	5	PPUW
	USA, New York, Tully	3	PPUH
	Norway, Bergen	2	PPNB
	Germany, Schwedt	1	PPDB
	Germany, Solling	5	PPDS
	Germany, Darmstadt	4	PPKW
	Belgium, Rockroi	2	PPBA
	Belgium, Calestienne	2	PPBC
	Belgium, Ottignies	2	PPBO
	Austria, Graz	6	PPOG
	Italy, Trento	3	PPIC
	Italy, Siena	3	PPIS
	Italy, Elba	4	PPIE
	Italy, Tirol	1	PPIO
	Italy, Monte Bodone	1	PPIM
	Kashmir, Srinagar	4	PPK
	Japan, Yatsugatake	8	PPJY
	Japan, Fuji Yoshida	5	PPJF
Platynothrus sibiricus	Costa Rica	4	PSCR
Platynothrus yamasakii	China, Beijing	3	PYCH
Heminothrus thori	Germany, Schwedt	2	HTDB
Camisia horrida	Costa Rica	2	CH
Mucronothrus nasalis	Canada, Toronto	6	MNCT
	Canada, British Columbia	4	MNCB
	USA, Colorado	1	MNUC
	Chile, Cap Horn	4	MNCH
	Norway, Finse	6	MNNF
	Papua New Guinea	2	MNNG

Molecular techniques

Oribatid mite specimens were preserved in 70% v/v ethanol until preparation. Due to the small size of oribatid mites (< 1 mm) it is difficult to obtain enough DNA for analysis from single specimens. Therefore, different principles of DNA extraction from single specimens were

conducted. Highest amounts of DNA were obtained with the DNeasy Tissue Kit (Qiagen). The principle of this technique is the use of a silica-gel-membrane for efficient purification of total cellular DNA without organic extraction or ethanol precipitation. Specimens were frozen on liquid nitrogen and squeezed with a mortar in an Eppendorff tube. The homogenised cells were lysed with a buffer containing detergent and proteinase K for digestion of cellular proteins and histones (Lottspeich and Zorbas 1998). DNA was bound to a silica-gel-membrane in presence of chaotropic reagents (Vogelstein and Gillespie 1979), washed and eluted in 30 µl of water. Of these, 5 µl with unknown concentration were used for PCR amplification (Saiki et al. 1985) with the HotStarTaq Master Mix Kit (Qiagen) and primers COIarch1 (5'GGTCAACAAATCATAAAGAYA TYG3') and COIarch2 (5'TAAACTTCAGGGTGACCAAAAAATCA3') (Thomas, personal communication). The total reaction volume of 50 µl contained 1.5 mM $MgCl_2$, 200 µM of each dNTP, 200 pmol of each primer and 2.5 units of Taq polymerase. PCR was specific; conditions were 15 min at 95°C for polymerase activation, 30 sec at 94°C for denaturation, 60 sec at 51°C for primer annealing and 60 sec at 72°C for elongation. Thirty-six cycles were performed followed by a terminal elongation (10 min) at 72°C. Products were purified on a 1% w/v agarose gel and stained with ethidium bromide; bands were excised, purified using chaotropic reagents, cloned with the Perfectly Blunt Cloning Kit (Novagen) and transfected in Nova Blue Singles™ competent cells (Novagen) by heat shock. Positive clones were selected by blue/white screening. Plasmids were purified by alkaline lysis. Inserts were sequenced in both directions by SRD GmbH (Oberursel, Germany) on an ABI capillary sequencer. All sequences are available at Genbank (AN: AY279416-AY279511).

Data analysis

All phylogenetic methods make assumptions, whether explicit or implicit, about the evolutionary process of DNA substitutions (Felsenstein 1988). For example, an assumption common to many phylogenetic methods is that bifurcating trees are appropriate representations of species

phylogenies (Huelsenbeck and Crandall 1997). Due to the nature of nucleotide sequences (only four character states are possible for every site and these four characters can change in all directions) the occurrence of convergent evolution between old and strongly diverged sequences is likely (Avise 1994; Page and Holmes 1998). As time goes by, the number of differences between two sequences becomes less and less of an accurate estimator of the actual number of substitutions that occurred since two sequences diverged from their common ancestor meaning that the distance of two sequences becomes saturated (Page and Holmes 1998). Classical phylogenetic methods using nonparametric maximum parsimony approaches to reconstruct evolutionary history therefore often are misleading. Given that observed distances may underestimate the actual amount of evolutionary change, many parametrical models were developed to convert the observed distances into measures of actual evolutionary distances (e.g. Jukes and Cantor 1969; Felsenstein 1981; Kimura 1980; Hasegawa et al. 1985; Rodriguez et al. 1990; Yang et al. 1994).

In addition, all evolutionary models can be modified by parametric assumptions on different substitution probabilities for different nucleotide positions in a gene (like third codon positions) underlying a Γ distribution with the shape parameter α which specifies the range of rate variation among sites (Yang 1996). Small values of α result in an L-shaped distribution with extreme variation in rates (most sites are invariable but few have very high rates of substitution). Conversely, the larger α the smaller is the range of rates (for $\alpha > 1$ the distribution gets bell-shaped; Yang 1996).

The measurement of genetic distances and phylogenetic analyses of nucleotide sequences are strongly dependant on the choice of the evolutionary model and the estimated parameters (Goldman 1993). A likelihood ratio test uses log likelihood scores to establish the model of DNA evolution that best fits the data with a minimum number of estimated parameters by comparison of more than 50 models (Posada and Crandall 1998, Figure 2.2).

A likelihood ratio test statistic is

$$\delta = 2\log\frac{\max[L_0(NullModel/Data)]}{\max[L_1(AlternativeModel/Data)]}$$

where L_0 is the likelihood under the null hypothesis (simple model) and L_1 is the likelihood under the alternative hypothesis (more complex model) (Posada and Crandall 1998).

Likelihood ratio tests were conducted with ModelTest 3.06 (Posada and Crandall 1998), genetic distances were calculated in PAUP* 4b10 (Swofford 1999) based on the evolutionary model and the parameters established by ModelTest 3.06 (Table 2.2).

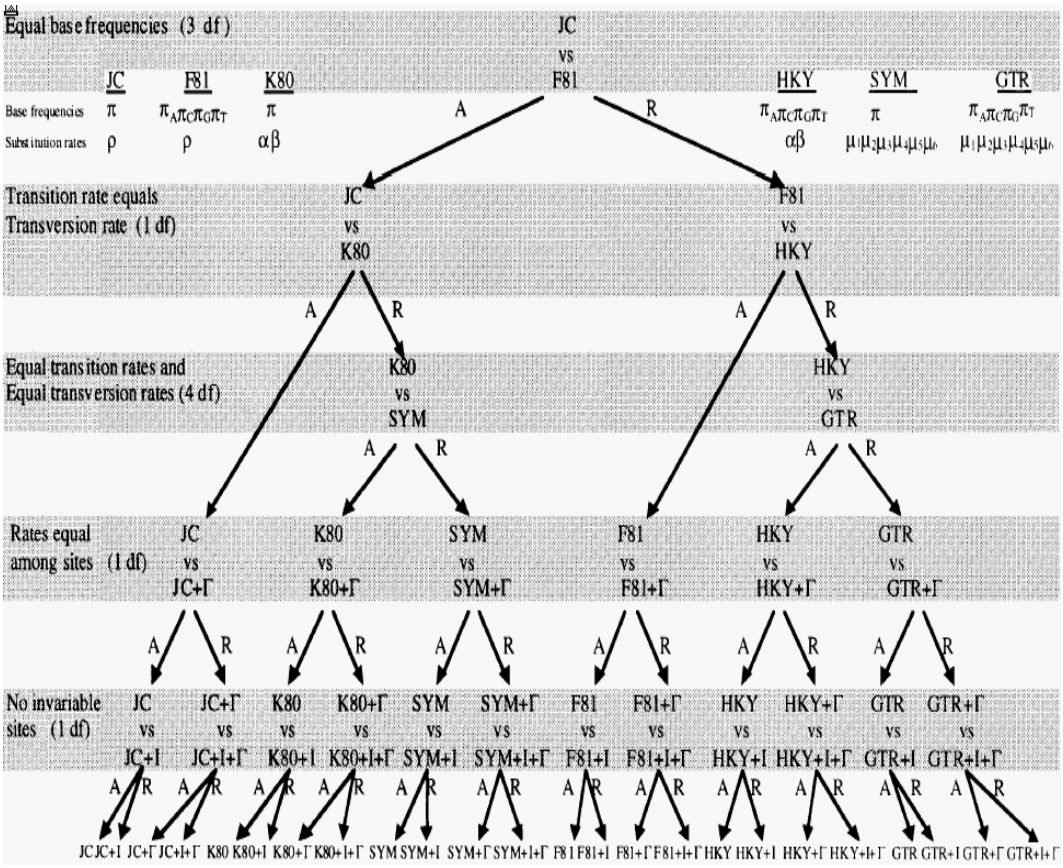


Figure 2.2
Hierachical hypothesis testing in ModelTest (Posada and Crandall 1998). At each level the null hypothesis (upper model) is either accepted (A) or rejected (R). Γ: gamma distribution; I: proportion of invariable sites; df: degrees of freedom; π: frequency of nucleotides; α: transition rate; β: transversion rate.

The admissibility to use genetic distances for time estimations depends on the clock-like evolution of the analysed genes. A simple way to test the accuracy of the molecular clock is to estimate the difference in number of substitutions between two closely related taxa in comparison with a third, more distantly related outgroup species by a relative rate test (Page and Holmes 1998). This test does not require any knowledge of the divergence times of the taxa in question. Tajimas relative rate test (Tajima 1993; Figure 2.3) is implemented in MEGA2.1 (Kumar et al. 2001); *Oribatula tibialis* (Oribatida: Poronota) was used as outgroup for pair-wise comparison of all Camisiidae and *M. nasalis*.

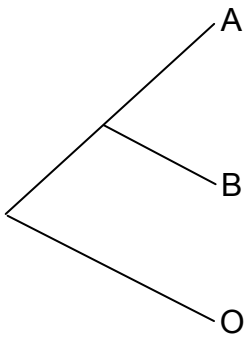


Figure 2.3

Consider three sequences A, B and O and let O be the outgroup. Let n*abo* be the observed number of sites (n) where sequences A, B and O have nucleotides *a*, *b* and *o*. Under the molecular clock hypothesis, E(n*abo*)=E(n*bao*) irrespective of the evolutionary model and whether or not the substitution rate varies with site. If this hypothesis is rejected, then the molecular clock hypothesis can be rejected for this set of sequences (Tajima 1993).

Genes coding for proteins are usually under negative selection when synonymous substitution (ss) outnumber those which are non-synonymous (ns) (ss>ns; Page and Holmes 1998). The rate of ss and ns can therefore be used to identify mitochondrial pseudogenes in the nuclear genome that have no function and are under neutral selection (ss=ns). A measurement of the difference between ss and ns is therefore important to verify that origin of sequences is from mitochondrial genes rather than from nuclear pseudogenes. Different algorithms have been developed to measure ss and ns (Li et al. 1985; Pamilo et al. 1993; Li 1993; Nei and Kumar 2000); all of them have some critical aspects in estimating the number of potential and realised synonymous and non-

synonymous substitutions (Nei and Kumar 2000). Therefore, all above mentioned methods to estimate ss and ns were used.

A valuable nonparametric measurement of genetic distances in coding regions is the percentage distance estimation in fourfold degenerate sites (D4; Mark Welsh and Meselson 2000). D4 sites do not effect the protein sequence and are assumed to have a slower saturation and less sensitivity to transition-transversion bias than other sites (Li 1993). Identification of D4 sites and translation of DNA into protein were conducted in MEGA2.1 (Kumar et al. 2001).

2.3. Results and discussion

A fragment of 600 bp of the COI gene was analysed corresponding to the positions 61-660 of the *Drosophila yakuba* COI gene and 21-220 of the *D. yakuba* protein (Clary and Wolstenholme 1985) (Appendix A). Sequences were verified by comparison with known sequences in GenBank using the BLAST search algorithm (Altschul et al. 1997). Nucleotide sequences were translated into amino acids using the invertebrate mitochondrial genetic code (Clary and Wolstenholme 1985). Sequences were aligned by hand; the alignment was free of gaps and unambiguous. The analysed part of the protein comprised 200 amino acids with two complete and one partial external loops, two complete internal loops and four complete and one partial transmembrane domains (Figure 2.7). With the internal loop I1 and the transmembrane domain M2 two of the most conserved regions of the COI protein were included in the analyses (Lunt et al. 1996).

Phylogenetic analyses did not indicate correlation of highly diverged sequences of *Platynothrus peltifer* and *Mucronothrus nasalis* with their geographical origin for either species indicating that COI lineages separated before separation of the populations (at least 200 milion years for *M. nasalis*; Hammer and Wallwork 1979; Norton et al. 1988b). In addition, genotypes of the different Camisiidae species were not associated with particular species indicating that COI lineages separated

before the radiation of Camisiidae into extant genera or species and confirms the hypothesis of asexual radiation of Camisiidae (Figures 2.4, 2.5). Therefore, all species of Camisiidae were merged for further calculations (Table 2.2).

Due to the persistence of different lineages and the apparent absence of lineage sorting it is unlikely that sexual reproduction occurred within or between areas.

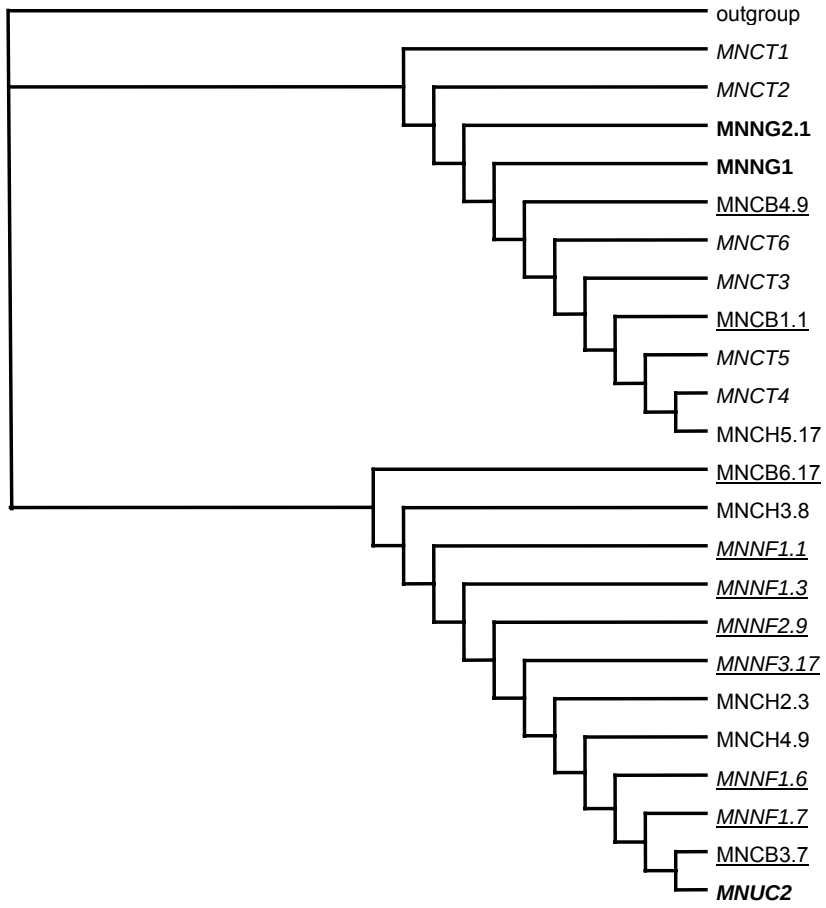


Figure 2.4
Phylogenetic analysis of *M. nasalis* nucleotide sequences. The tree is a Neighbour Joining tree calculated in PAUP*. Distances were calculated based on the evolutionary model estimated by ModelTest with suggested parameters (Table 2.2). The 14 haplotypes are distributed over at least 2 lineages. Note that lineages and geographic origin are poorly correlated (geographic origin indicated by different accentuations). Abbreviations see Table 1.

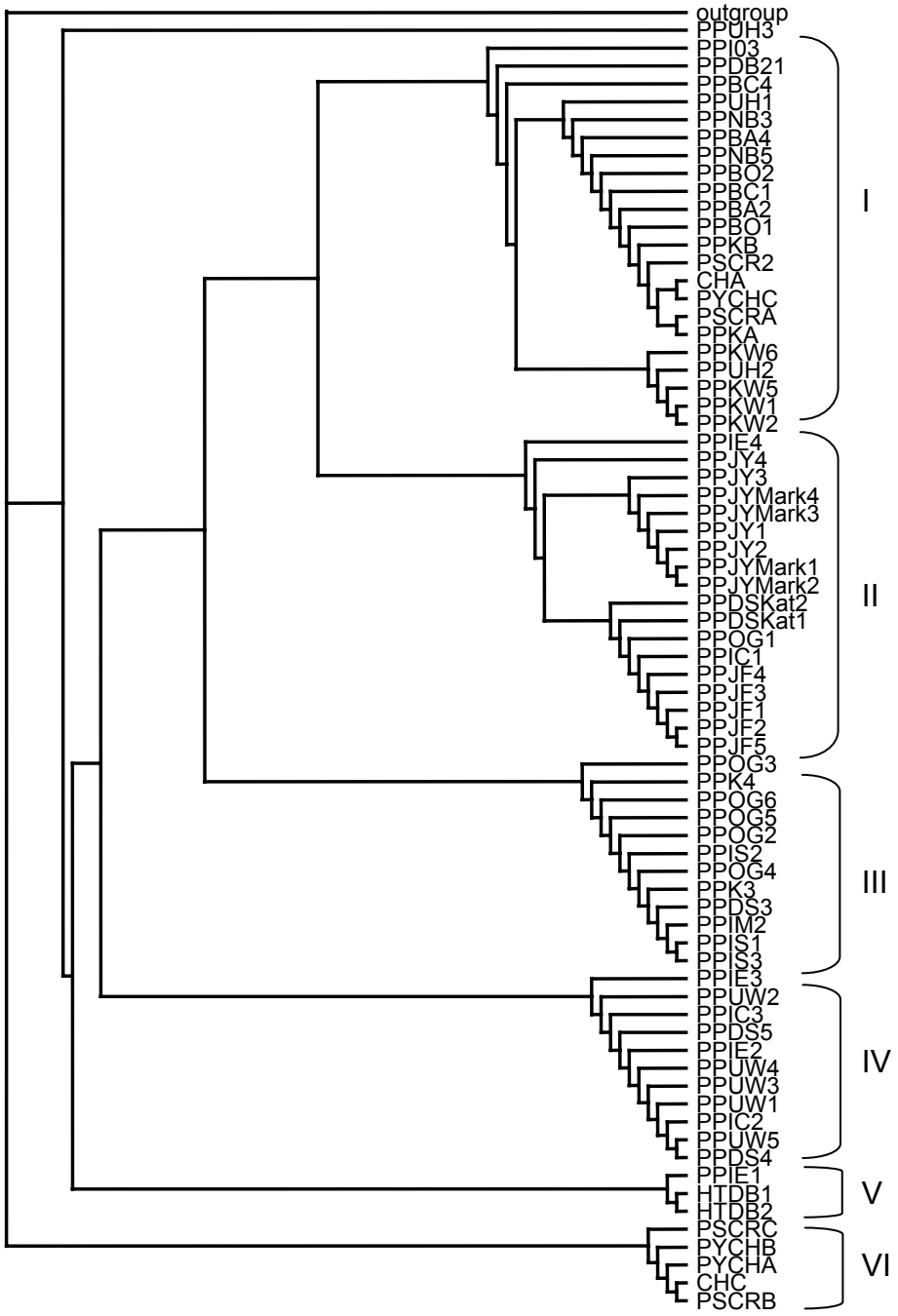


Figure 2.5
Phylogenetic analysis of Camisiidae nucleotide sequences. The tree is a Neighbour Joining tree calculated in PAUP*. Distances were calculated based on the evolutionary model estimated by ModelTest with suggested parameters (Table 2.2). The 42 haplotypes are distributed over at least 6 lineages (I-VI). There is no correlation between lineages and geographical origin of specimens nor a correlation between species (*P. peltifer, P. yamasakii, P. sibiricus, H. thori, C. horrida*) and genotype. For abbreviations see Table 1.

All sequences were free of stop codons, and highly conserved positions were equal in all sequences. For Camisiidae, 261 nucleotide positions were variable and formed 42 different haplotypes. The *M. nasalis* sequences consisted of 14 haplotypes with 249 polymorphic sites. The preponderant type of substitutions was synonymous (ss-sn averaged 1.1 for *M. nasalis* and 1.2 for Camisiidae). Therefore, it is unlikely that sequences represent nuclear pseudogenes.

Table 2.2

Genetic variability for the 600 bp cytochrome oxidase I region of Camisiidae (72 specimens from 22 sites) and *Mucronothrus nasalis* (23 specimens from 6 sites).

	Camisiidae	*M. nasalis*
Polymorphic sites for DNA	261	249
Number of haplotypes	42	14
A frequency	0.2993	0.2893
C frequency	0.2435	0.2489
G frequency	0.1354	0.1352
T frequency	0.3218	0.3266
Likelihood ratio test (model choice)	HKY+I+Γ	
-lnL (log likelihood)	4151	2988
Gamma-shape (Γ)	1.0076	1.1272
Invariant sites (I)	0.4858	0.5246
Transition/transversion ratio	4.6592	5.5250

It is interesting to note that the A+T content (57% and 58% for Camisiidae and *M. nasalis*, respectively), highly biased in the COI gene in other studied Chelicerata (Avise et al. 1994; Navajas et al. 1996; Salomone et al. 1996), tends to be much less biased in Camisiidae and *M. nasalis* being the only known arthropods with A+T contents in the COI gene less than 60%. Several hypotheses have been put forward to explain biases in nucleotide contents (Moriyama and Gojobori 1992; Jermiin and Crozier 1994; Wirth et al. 1999; Xia 2000) but a consistent explanation for this high variability in nucleotide frequencies has not yet been provided. In order to

determine the mechanisms governing the base composition in Camisiidae and *M. nasalis* more detailed analyses on the pattern and direction of mutation in the mitochondrial genome of these taxa are necessary.

About 15% of all sites in the analysed COI gene were fourfold degenerate (D4); in Camisiidae and *M. nasalis*, 100% and 99% of these were variable, respectively, suggesting high saturation of sequences. The maximum pairwise percentage distance of D4 sites was 76% for Camisiidae and 73% for *M. nasalis*. Using this nonparametric distance estimation with a maximum conservative evolutionary rate of 5% (Tautz et al. 2003) an age of more than 7 million years of asexual reproduction was estimated for Camisiidae and *M. nasalis*. This estimation can be refined by parametrical approaches.

Therefore, maximum pair wise corrected distances between lineages were calculated. Due to high saturation, distances were corrected using an evolutionary model selected by a likelihood ratio test (Posada and Crandall 1998). Distances were calculated for Camisiidae and separately for *M. nasalis* using the estimated model with suggested parameters (HKY+I+ Γ (Hasegawa et al. 1985); Table 2.2). For Camisiidae the genetic distance averaged 61% with a maximum of 136%; for *M. nasalis* it averaged 67% with a maximum of 157%. Distances of more than 100% are possible if backmutations are taken into account. Maximum distances apply to the deepest branching in the group and thus to its age (Avise 1994).

The intraspecific variability for Camisiidae and *M. nasalis* is more than 15-fold higher than in sexual populations of the oribatid mite species *Steganacarus magnus* and the insect species *Lymantria dispar* and even about 30-fold higher than in the putative ancient asexual ostracod *D. Stevensoni* (Figure 2.6). These findings indicate an ancient split of the different COI lineages and the absence of lineage sorting.

Relative rate tests (Tajima 1993; Figure 2.3) indicate that evolution of COI sequences was clockwise for Camisiidae and *M. nasalis* (no rate heterogeneity with $X^2<3.85$ and $p>0.05$). Therefore, the time elapsed since separation of lineages can be estimated.

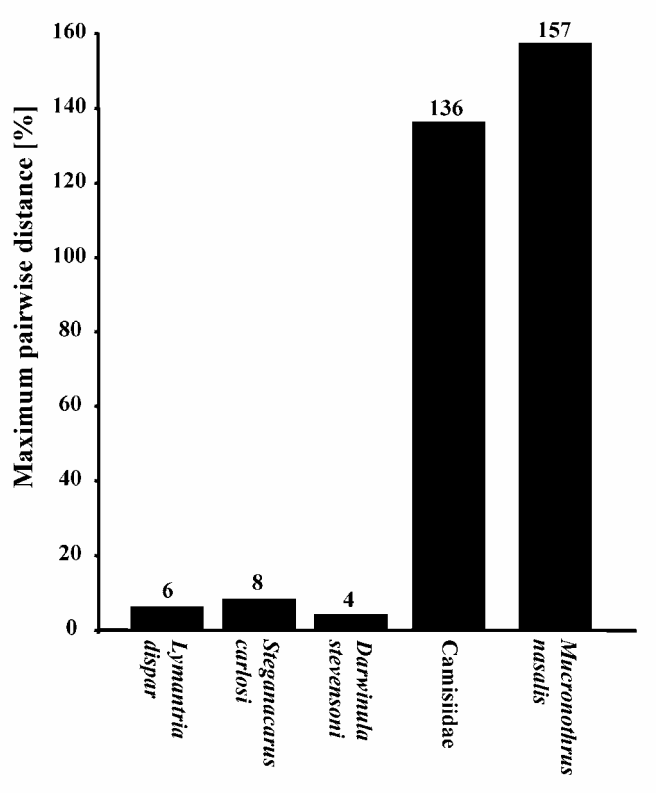


Figure 2.6
Divergence of the COI gene for some invertebrate species. *Lymantria dispar* (Hexapoda: Lepidoptera; Bogdanowicz et al. 2000) ; *Steganacarus carlosi* (Acari: Oribatida; Salomone et al. 2002) ; *Darwinula stevensoni* (Crustacea: Ostracoda; Schön et al. 1998).

The COI gene of oribatid mites was determined to evolve at a divergence rate of 2.15% per million years (Salomone et al. 2002) suggesting an age of 63 and 73 million years for Camisiidae and *M. nasalis*, respectively. Since sequences were highly saturated, backmutations likely occur and the calculated ages are minimum estimates.

To further improve time estimations we used COI protein sequences of Camisiidae and *M. nasalis*. The structure of the COI protein in oribatid mites is similar to that in insects (Lunt et al. 1996). The evolutionary rate of the protein was estimated using those of 6 Neoptera insect species (2 Caelifera and 4 Diptera; Table 2.3). The last common ancestor of these insect species lived in the Carboniferous about 285-360 million years ago (Gullan and Cranston 1994).

Table 2.3
Origin of the data concerning the COI protein of insect species.

Species	Accession number	Reference
Chorthippus parallelus	AF229010	Lunt et al. 1996
Locusta migratoria	X80245	Flook et al. 1995
Anopheles gambiae	L20934	Beard et al. 1993
Anopheles quadrimaculatus	L04272	Cockburn et al. 1990
Drosophila yakuba	X03240	Clary and Wolstenholme 1985
Protophormia terraenovae	L14946	Wells and Sperling 2000

Of the 200 analysed amino acids in Neoptera, 39 were variable and the maximum pair wise distance was 15% between *Locusta migratoria* and *Anopheles quadrimaculatus*. This distance is according to an age of 285-360 million years indicating an evolutionary rate of 0.04-0.05% of divergence per million years. Within Camisiidae and also in *M. nasalis* 34 amino acids were variable (Figure 2.7). The maximum pairwise distance was 7.5% and 10% for Camisiidae and *M. nasalis*, respectively, indicating an age of at least 143-180 million years for Camisiidae and 190-240 million years for *M. nasalis*. Therefore, the hypothesis that *M. nasalis* predated the breakup of Pangaea 200 million years ago (Hammer and Wallwork 1979; Norton et al. 1988b) is strongly supported, and the Camisiidae probably also existed at that time. Due to the absence of closely related sexual species, frequent evolution of parthenogenetic lineages from sexual ancestors is unlikely (Norton and Palmer 1991; Palmer and Norton 1992).

Mucronothrus nasalis is considered a single species with little morphological variability (Norton et al. 1996) but the strong variability in the COI gene, its disjunct biogeographical distribution (Hammer and Wallwork 1979) and its clonal population structure (Palmer and Norton 1992) suggests that it is an ancient asexual species.

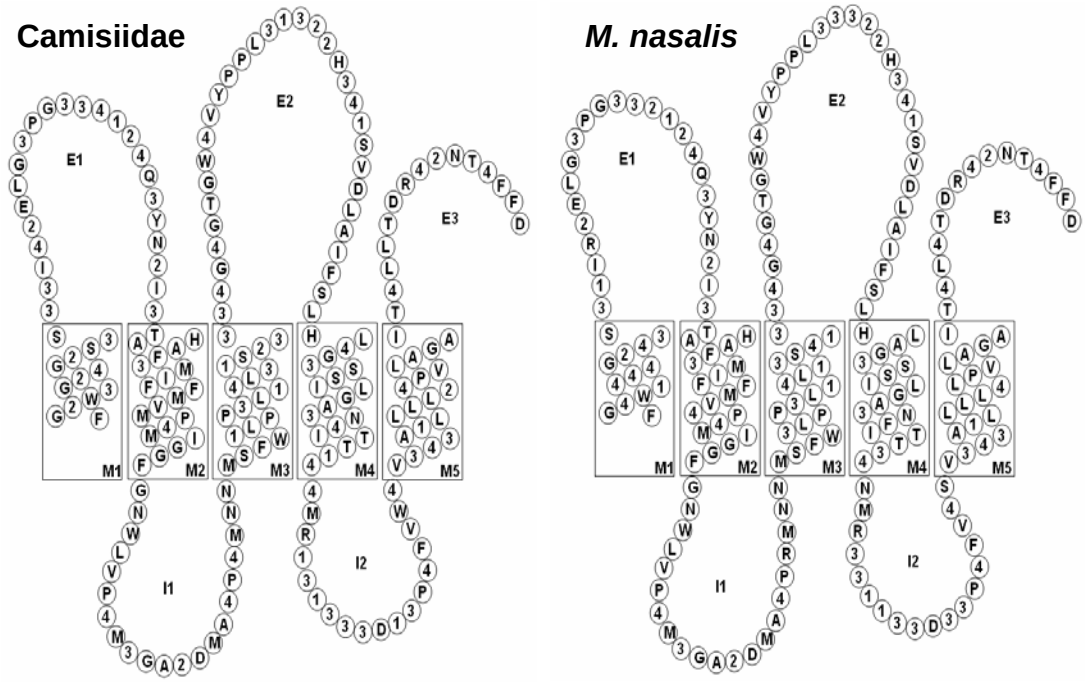


Figure 2.7: Partial structure of the cytochrome oxidase I protein
The analysed part of the COI protein corresponds to positions 21-220 of the insect COI. Conserved amino acids for the neopteran and the oribatid mite protein are indicated in one-letter-code. Positions marked with “1” are variable between and within insects and oribatid mites, “2” are variable between insects and oribatid mites, “3” are variable within insects and “4” are variable only within oribatid mites. M1-M5: transmembrane domains; E1-E3: external loops; I1-I2: internal loops

Similar in Camisiidae, the variability in COI, the absence of sexual reproduction (Norton and Palmer 1991) and the clonal structure of the examined species indicate ancient asexual origin. Oribatid mites therefore represent more than a third “evolutionary scandal”; they contain several ancient and speciose parthenogenetic groups representing a number of evolutionary scandals. With 240 million years they probably contain the oldest living asexual metazoan taxa. Presumably, a combination of factors - including habitat stability and an automictic genetic system with inverted meiosis - is responsible for the independent evolution, longevity, and radiation of multiple lineages of parthenogenetic oribatid mites.

Appendix A

Nucleotide sequence alignment of the 72 Camisiidae specimens

```
            ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                5         15         25         35         45         55         65
CHA         TTTGGAACCT GGGCTGGATT GATAGGATCT TCTTTAAGTG CTTTGATTCG TCTTGAGTTA GGTCAACCCG
CHC         TTTCGGTACC TGGGTGGCcT AATGGGATCT TCTCTAAGAG CATTAATTCG ATTGGAATTA GGACAACCAG
HTDB1       TTTGGTACAT GAGCTGGATT AATAGGGTCA TCATTAAGGG CTTTAATTCG ACTTGAGTTA GGGCAGCCCG
HTDB2       TTTGGTACAT GAGCTGGATT AATAGGGTCA TCATTAAGGG CTTTAATTCG ACTTGAGTTA GGGCAGCCCG
PSCR2       TTTGGAACCT GGGCTGGATT GATAGGATCT TCTTTAAGTG CTTTGATTCG TCTTGAGTTA GGTCAACCCG
PSCRA       TTTGGAACCT GGGCTGGATT GATAGGATCT TCTTTAAGTG CTTTGATTCG TCTTGAGTTA GGTCAACCCG
PSCRB       TtcGGTACCT GGGCTGGCCT AATGGGATCT TCTCTAAGAG CATTAATTCG ATTGGAATTA GGACAACCAG
PSCRC       TTCGGTACTT GAGCGGGACT AATGGGCTCT TCACTGAGAG CCCTGATTCG TTTAGAATTA GGACAACCTG
PYCHA       TTCGGtACCT GGGCTGGCCT AATGGGATCT TCTCTAAGAG CATTAATTCG ATTGGAATTA GGACAACCAG
PYCHB       TTCGGTACCT GGGCTGGCCT AATGGGATCT TCTCTAAGAG CATTAATTCG ATTGGAATTA GGACAACCAG
PYCHC       TTTGGAACCT GGGCTGGATT GATAGGATCT TCTTTAAGtG CTTTGATTCG TCTTGAGTTA GGTCAACCCG
PPK3        TTTGGTACAT GAGCTGGACT AATAGGCTCC TCCCTAAGTG CCTTGATTCG TCTTGAATTA GGACAGCCCG
PPK4        TTTGGtaCAT GAGCTGGACT AATAGGCTCC TCCCTAAGTG CTTTGATTCG TCTTGAATTA GGACAGCCCG
PPKA        TTTGGAACCT GGGCTGGATT GATAGGATCT TCTTTAAGtG CTTTGATTCG TCTTGAGTTA GGtCAACCCG
PPKB        TTTGGAACCT GGGCTGGATT GATAGGATCT TCTTTAAGTG CTTTGATTCG TCTTGAGTTA GGTCAACCCG
PPUW1       TTTGGAACTT GAGCGGGATT GATAGGTTCC TCTTTAAGGG CGCTGATCCG GTTAGAACTA GGGCAGCCAG
PPUW2       TTTGGAACTT GAGCGGGATT GATAGGTTCC TCTTTAAGGG CGCTGATCCG GTTAGAACTA GGGCAGCCAG
PPUW3       TTTGGAACTT GAGCGGGATT GATAGGTTCC TCTTTAAGGG CGCTGATCCG GTTAGAACTA GGGCAGCCAG
PPUW4       TTTGGAACTT GAGCGGGATT GATAGGTTCC TCTTTAAGGG CGCTGATCCG GTTAGAACTA GGGCAGCCAG
PPUW5       TTTGGAACTT GAGCGGGATT GATAGGTTCC TCTTTAAGGG CGCTGATCCG GTTAGAACTA GGGCAGCCAG
PPUH1       TTTGGAACCT GGGCTGGATT GATAGGATCT TCTTTAAGTG CTTTGATTCG TCTTGAGTTA GGTCAACCCG
PPUH2       TTTGGAACCT GGGCTGGATT GATAGGATCT TCTTTAAGTG CTTTGATTCG TCTTGAGTTG GGTCAACCCG
PPUH3       TTCGGCACAT GGGCTGGGCT GATGGGCTCG TCTTTAAGAG CTCTAATTCG GCTAGAACTA GGTCAACCCG
PPJYMark1   TTTGGGACTT GAGCTGGGTT AATGGGATCT TCTCTAAGTG CTTTAATTCG CCTCGAGTTA GGACAACCTG
PPJYMark2   TTTGGGACTT GAGCTGGGTT AATGGGATCT TCTCTAAGTG CTTTAATTCG CCTCGAGTTA GGACAACCTG
PPJYMark3   TTTGGGACTT GAGCTGGGTT AATGGGATCT TCTCTAAGTG CTTTAATTCG CCTCGAGTTA GGACAACCTG
PPJYMark4   TTTGGGACTT GAGCTGGGTT AATGGGATCT TCTCTAAGTG CTTTAATTCG CCTCGAGTTA GGACAACCTG
PPJY1       TTTGGGACTT GAGCTGGGTT AATGGGATCT TCTCTAAGTG CTTTAATTCG CCTCGAGTTA GGACAACCTG
PPJY2       TTTGGGACTT GAGCTGGGTT AATGGGATCT TCTCTAAGTG CTTTAATTCG CCTCGAGTTA GGACAACCTG
PPJY3       TTTGGGACTT GAGCTGGGTT AATGGGATCT TCTCTAAGTG CTTTAATTCG CCTCGAGTTA GGACAACCTG
PPJY4       TTTGGGACTT GAGCTGGGTT AATGGGATCT TCTCTAAGTG CTTTAATTCG CCTCGAGTTA GGACAACCTG
PPJF1       TTTGGGACTT GAGCTGGGTT AATGGGATCT TCTCTAAGTG CTTTAATTCG CCTCGAGTTA GGACAACCTG
PPJF2       TTTGGGACTT GAGCTGGGTT AATGGGATCT TCTCTAAGTG CTTTAATTCG CCTCGAGTTA GGACAACCTG
PPJF3       TTTGGGACTT GAGCTGGGTT AATGGGATCT TCTCTAAGTG CTTTAATTCG CCTCGAGTTA GGACAACCTG
PPJF4       TTTGGGACTT GAGCTGGGTT AATGGGATCT TCTCTAAGTG CTTTAATTCG CCTCGAGTTA GGACAACCTG
PPJF5       TTTGGGACTT GAGCTGGGTT AATGGGATCT TCTCTAAGTG CTTTAATTCG CCTCGAGTTA GGACAACCTG
PPIE1       TTTGGTACAT GAGCTGGATT AATAGGGTCA TCATTAAGGG CTTTAATTCG ACTTGAGTTA GGGCAGCCCG
PPIE2       TTTGGAACTT GAGCGGGATT GATAGGTTCC TCTTTAAGGG CGCTGATCCG GTTAGAACTA GGGCAGCCAG
PPIE3       TTTGGAACTT GAGCGGGATT GATAGGTTCC TCTTTAAGGG CGCTGATCCG GTTAGAACTA GGGCAGCCAG
PPIE4       TTTGGGACTT GAGCTGGGTT AATGGGATCT TCTCTAAGTG CTTTAATTCG CCTCGAGTTA GGACAACCTG
PPIC1       TTTGGGACTT GAGCTGGGTT AATGGGATCT TCTCTAAGTG CTTTAATTCG CCTCGAGTTA GGACAACCTG
PPIC2       TTTGGAACTT GAGCGGGATT GATAGGTTCC TCTTTAAGGG CGCTGATCCG GTTAGAACTA GGGCAGCCAG
PPIC3       TTTGGAACTT GAGCGGGATT GATAGGCTCC TCTTTAAGGG CGCTGATCCG GTTAGAACTA GGGCAGCCAG
PPIS1       TTTGGTACAT GAGCTGGACT AATAGGCTCC TCCCTAAGTG CCTTGATTCG TCTTGAATTA GGACAGCCCG
PPIS2       TTTGGTACAT GAGCTGGACT AATAGGCTCC TCCCTAAGTG CCTTGATTCG TCTTGAATTA GGACAGCCCG
PPIS3       TTTGGTACAT GAGCTGGACT AATAGGCTCC TCCCTAAGTG CCTTGATTCG TCTTGAATTA GGACAGCCCG
PPIM2       TTTGGTACAT GAGCTGGACT AATAGGCTCC TCCCTAAGTG CCTTGATTCG TCTTGAATTA GGACAGCCCG
PPI03       TTTGGAACTT GAGcTGGGCT AATAGGATCT TCTTTAAGTG CCCTGATCCG TCTCGAGTTG GGACAGCCCG
PPOG1       TTTGGGACTT GAGCTGGGTT AATGGGATCT TCTCTAAGTG CTTTAATTCG CCTCGAGTTA GGACAACCTG
PPOG2       TTTGGTACAT GAGCTGGACT AATAGGCTCC TCCCTAAGTG CCTTGATTCG TCTTGAATTA GGACAGCCCG
PPOG3       TTTGGTACAT GAGCTGGACT AATAGGCTCC TCCCTAAGTG CCTTGATTCG TCTTGAATTA GGACAACCCG
PPOG4       TTTGGTACAT GAGCTGGACT AATAGGCTCC TCCCTAAGTG CCTTGATTCG TCTTGAATTA GGACAGCCCG
PPOG5       TTTGGTACAT GAGCTGGACT AATAGGCTCC TCCCTAAGTG CCTTGATTCG TCTTGAATTA GGACAGCCCG
PPOG6       TTTGGTACAT GAGCTGGACT AATAGGCTCC TCCCTAAGTG CCTTGATTCG TCTTGAATTA GGACAGCCCG
PPNB3       TTTGGAACCT GGGCTGGATT GATAGGATCT TCTTTAAGTG CTTTGATTCG TCTTGAGTTA GGTCAACCCG
PPNB5       TTTGGGACCT GGGCTGGATT GATAGGATCT TCTTTAAGTG CTTTGATTTG TCTTGAGTTA GGTCAACCCG
PPBC1       TTTGGAACCT GGGCTGGATT GGTAGGATCT TCTTTAAGTG CTTTGATTCG TCTTGAGTTA GGTCAACCCG
PPBC4       TTTGGAACCT GGGCTGGATT AATAGGATCT TCTTTAAGTG CTTTGATTCG TCTTGAGTTA GGTCAACCCG
PPBA2       TTTGGAACCT GGGCTGGATT GATAGGATCT TCTTTAAGTG CTTTGATTCG TCTTGAGTTA GGTCAACCCG
PPBA4       TTTGGAACCT GGGCTGGATT GATAGGATCT TCTTTAAGTG CTTTGATTCG TCTTGAGTTA GGTCAACCCG
PPBO1       TTTGGAACCT GGGCTGGATT GATAGGATCT TCTTTAAGTG CTTTGATTCG TCTTGAGTTA GGTCAACCCG
PPBO2       TTTGGAACCT GGGCTGGATT GATAGGATCT TCTTTAAGTG CTTTGATTCG TCTTGAGTTA GGTCAACCCG
PPDB21      TTTGGAACCT GGGCTGGATT AATAGGATCT TCCTTAAGTG CCTTGATTCG TCTTGAGTTG GGTCAACCTG
PPDSKat1    TTTGGGACTT GAGCTGGGTT AATGGGATCT TCTCTAAGTG CTTTAATTCG CCTCGAGTTA GGACAACCTG
PPDSKat2    TTTGGGACTT GAGCTGGGTT AATGGGATCT TCTCTAAGTG CTTTAATTCG CCTCGAGTTA GGACAACCTG
PPDS3       TTTGGTACAT GAGCTGGACT AATAGGCTCC TCCCTAAGtG CCTTGATTCG TCTTGAATTA GGACAGCCCG
PPDS4       TTTGGAACTT GAGCGGGATT GATAGGTTCC TCTTTAAGGG CGCTGATCCG GTTAGAACTA GGGCAGCCAG
PPDS5       TTTGGAACTT GAGCGGGATT GATAGGTTCC TCTTTAAGGG CGCTGATCCG GTTAGAACTA GGGCAGCCAG
PPKW1       TTTGGAACCT GGGCTGGATT GATAGGATCT TCTTTAAGTG CTTTGATTCG TCTTGAGTTG GGTCAACCCG
PPKW2       TTTGGAACCT GGGCTGGATT GATAGGATCT TCTTTAAGTG CTTTGATTCG TCTTGAGTTG GGTCAACCCG
PPKW5       TTTGGAACCT GGGCTGGATT GATAGGATCT TCTTTAAGTG CTTTGATTCG TCTTGAGTTG GGTCAACCCG
PPKW6       TTTGGAACCT GGGCTGGATT GATAGGATCT TCTTTAAGTG CTTTGATTCG TCTTGAGTTG GGTCAACCCG
```

```
                ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                    75         85         95        105        115        125        135
CHA             GTTCACTACT AGACAACGAC CAAATTTATA ACACCATTGT TACAGCCCAC GCCTTTGTAA TAATTTTTTT
CHC             GCTCTCTTCT CGAAAATGAC CAAATTTACA ACACAATTGT CACTGCACAC GCCTTCGTAA TAATCTTCTT
HTDB1           GATCACTTTT AGATAACGAC CAAATCTATA ACACCATCGT TACAGCACAC GCTTTCGTAA TAATTTTTTT
HTDB2           GATCACTTTT AgATAACGAC CAAATCTATA ACACCATCGT TACAGCACAC GCTTTCGTAA TAATTTTTTT
PSCR2           GTTCACTACT AGACAACGAC CAAATTTATA ACACCATTGT TACAGCCCAC GCCTTTGTAA TAATTTTTTT
PSCRA           GTTCACTACT AGACAACGGC CAAATTTATA ACACCATTGT TACAGCCCAC GCCTTTGTAA TAATTTTTTT
PSCRB           GCTCTCTTCT CGAAAATGAC CAAATTTACA ACACAATTGT CACTGCACAC GCCTTCGTAA TAATCTTCTT
PSCRC           GTTCCCTCCT TGAAAATGAT CAAATCTATA ACACTATTGT CACTGCCCAT GCCTTCGtAA TAATTTTTTT
PYCHA           GCTCTCTTCT CGAAAATGAC CAAATTTACA ACACAATTGT CACTGCACAC GCCTTCGTAA TAATCTTCTT
PYCHB           GCTCTCTTCT CGAAAATGAC CAAATTTACA ACACAATTGT CACTGCACAC GCCTTCGTAA TAATCTTCTT
PYCHC           GTTCACTACT AGACAACGAC CAAATTTATA ACACCATTGT TACAGCCCAC GCCTTTGTAA TAATTTTTTT
PPK3            GTTCACTTCT AGATAACGAC CAAATCTACA ACACTATTGT CACAGCTCAC GCCTTCGTAA TGATTTTTTT
PPK4            GTTCACTTCT AGATAACGAC CAAATCTACA ACACTATTGT CACAGCTCAC GCCTTCGTAA TGATTTTTTT
PPKA            GTTCACTACT AGACAACGAC CAAATTTATA ACACCATTGT TACAGCCCAC GCCTTTGTAA TAATTTTTTT
PPKB            GTTCACTACT AGACAACGAC CAAATTTATA ACACCATTGT TACAGCCCAC GCCTTTGTAA TAATTTTTTT
PPUW1           GATCATTACT AGACAATGAC CAAATTTATA ACACAATTGT TACGGCTCAC GCTTTCGTTA TGATTTTCTT
PPUW2           GATCATTACT AGACAATGAC CAAATTTATA ACACAATTGT TACAGCTCAC GCTTTCGTTA TGATTTTCTT
PPUW3           GATCATTACT AGACAATGAC CAAATTTATA ACACAATTGT TACGGCTCAC GCTTTCGTTA TGATTTTCTT
PPUW4           GATCATTACT AGACAATGAC CAAATTTATA ACACAATTGT TACGGCTCAC GCTTTCGTTA TGATTTTCTT
PPUW5           GATCATTACT AGACAATGAC CAAATTTATA ACACAATTGT TACGGCTCAC GCTTTCGTTA TGATTTTCTT
PPUH1           GTTCACTACT AGACAACGAC CAAATTTATA ACACCATTGT TACAGCCCAC GCCTTTGTAA TAATTTTTTT
PPUH2           GTTCACTACT AGACAACGAC CAAATTTATA ACACCATTGT TACAGCCCAC GCCTTTGTAA TAATTTTTTT
PPUH3           GATCCCTTAT AGATAATGAC CAAATCTACA ACACTATTGT TACTGCTCAC GCTTTCGTTA TGATCTTCTT
PPJYMark1       GATCTCTCTT AGATAACGAT CAAATCTACA ATACTATTGT TACAGCTCAC GCCTTCGTAA TAATTTTCTT
PPJYMark2       GATCTCTCTT AGATAACGAT CAAATCTACA ATACTATTGT TACAGCTCAC GCCTTCGTAA TAATTTTCTT
PPJYMark3       GGTCTCTCCT AGATAACGAT CAAATCTACA ATACTATTGT TACAGCCCAC GCCTTCGTAA TAATTTTCTT
PPJYMark4       GGTCTCTCCT AGATAACGAT CAAATCTACA ATACTATTGT TACAGCCCAC GCCTTCGTAA TAATTTTCTT
PPJY1           GATCTCTCCT AGATAACGAT CAAATCTACA ATACTATTGT TACAGCCCAC GCCTTCGTAA TAATTTTCTT
PPJY2           GATCTCTCTT AGATAACGAT CAAATCTACA ATACTATTGT TACAGCTCAC GCCTTCGTAA TAATTTTCTT
PPJY3           GGTCTCTCCT AGATAACGAT CAAATCTACA ATACTATTGT TACAGCCCAC GCCTTCGTAA TAATTTTCTT
PPJY4           GGTCTCTCCT AGATAACGAT CAAATCTACA ATACTATTGT TACAGCCCAC GCCTTCGTAA TAATTTTCTT
PPJF1           GGTCTCTCCT AGATAACGAT CAAATCTACA ATACTATTGT TACAGCCCAC GCCTTCGTAA TAATTTTCTT
PPJF2           GGTCTCTCCT AGATAACGAT CAAATCTACA ATACTATTGT TACAGCCCAC GCCTTCGTAA TAATTTTCTT
PPJF3           GGTCTCTCCT AGATAACGAT CAAATCTACA ATACTATTGT TACAGCCCAC GCCTTCGTAA TAATTTTCTT
PPJF4           GGTCTCTCCT AGATAACGAT CAAATCTACA ATACTATTGT TACAGCCCAC GCCTTCGTAA TAATTTTCTT
PPJF5           GGTCTCTCCT AGATAACGAT CAAATCTACA ATACTATTGT TACAGCCCAC GCCTTCGTAA TAATTTTCTT
PPIE1           GATCACTTTT AGATAACGAC CAAATCTATA ACACCATCGT TACAGCACAC GCTTTCGTAA TAATTTTTTT
PPIE2           GATCATTACT AGACAATGAC CAAATTTATA ACACAATTGT TACGGCTCAC GCTTTCGTTA TGATTTTCTT
PPIE3           GATCATTACT AGACAATGAC CAAATTTATA ACACAATTGT TACGGCTCAC GCTTTCGTTA TGATTTTCTT
PPIE4           GGTCTCTCCT AGATAACGAT CAAATCTACA ATACTATTGT TACAGCCCAC GCCTTCGTAA TAATTTTCTT
PPIC1           GGTCTCTCCT AGATAACGAT CAAATCTACA ATACTATTGT TACAGCCCAC GCCTTCGTAA TAATTTTCTT
PPIC2           GATCATTACT AGACAATGAC CAAATTTATA ACACAATTGT TACGGCTCAC GCTTTCGTTA TGATTTTCTT
PPIC3           GATCATTACT AGACAATGAC CAAATTTATA ACACAATTGT TACGGCTCAC GCTTTCGTTA TGATTTTCTT
PPIS1           GTTCACTTCT AGATAACGAC CAAATCTACA ACACTATTGT CACAGCTCAC GCTTTCGTAA TGATTTTTTT
PPIS2           GTTCACTTCT AGATAACGAC CAAATCTACA ACACTATTGT CACAGCTCAC GCCTTCGTAA TGATTTTTTT
PPIS3           GTTCACTTCT AGATAACGAC CAAATCTACA ACACTATTGT CACAGCTCAC GCTTTCGTAA TGATTTTTTT
PPIM2           GTTCACTTCT AGATAACGAC CAAATCTACA ACACTATTGT CACAGCTCAC GCTTTCGTAA TGATTTTTTT
PPI03           GTTCACTTTT GGACAATGAC CAAATTTATA ACACCATTGT TACGGCCCAC GCCTTCGTAA TAATTTTCTT
PPOG1           GGTCTCTCCT AGATAACGAT CAAATCTACA ATACTATTGT TACAGCCCAC GCCTTCGTAA TAATTTTCTT
PPOG2           GTTCACTTCT AGATAACGAC CAAATCTACA ACACTATTGT CACAGCTCAC GCCTTCGTAA TGATTTTTTT
PPOG3           GTTCACTTCT AGATAACGAC CAAATCTACA ACACTATTGT CACAGCTCAC GCCTTCGTAA TGATTTTTTT
PPOG4           GTTCACTTCT AGATAACGAC CAAATCTACA ACACTATTGT CACAGCTCAC GCCTTCGTAA TGATTTTTTT
PPOG5           GTTCACTTCT AGATAACGAC CAAATCTACA ACACTATTGT CACAGCTCAC GCCTTCGTAA TGATTTTTTT
PPOG6           GTTCACTTCT AGATAACGAC CAAATCTACA ACACTATTGT CACAGCTCAC GCCTTCGTAA TGATTTTTTT
PPNB3           GTTCACTACT AGACAACGAC CAAATTTATA ACACCATTGT TACAGCCCAC GCCTTTGTAA TAATTTTTTT
PPNB5           GTTCACTACT AGACAACGAC CAAATTTATA ACACCATTGT TACAGCCCAC GCCTTTGTAA TAATTTTTTT
PPBC1           GTTCACTACT AGACAACGAC CAAATTTATA ACACCATTGT TACAGCCCAC GCCTTTGTAA TAATTTTTTT
PPBC4           GTTCACTACT AGACAACGAC CAAATTTATA ACACCATTGT TACAGCCCAC GCCTTTGTAA TAATTTTTTT
PPBA2           GTTCACTACT AGACAACGAC CAAATTTATA ACACCATTGT TACAGCCCAC GCCTTTGTAA TAATTTTTTT
PPBA4           GTTCACTACT AGACAACGAC CAAATTTATA ACACCATTGT TACAGCCCAC GCCTTTGTAA TAATTTTTTT
PPBO1           GTTCACTACT AGACAACGAC CAAATTTATA ACACCATTGT TACAGCCCAC GCCTTTGTAA TAATTTTTTT
PPBO2           GTTCACTACT AGACAACGAC CAAATTTATA ACACCATTGT TACAGCCCAC GCCTTTGTAA TAATTTTTTT
PPDB21          GTTCACTACT AGACAACGAC CAAATTTATA ACACCATTGT TACAGCCCAC GCTTTTGTAA TAATTTTTTT
PPDSKat1        GGTCTCTCCT AGATAACGAT CAAATCTACA ATACTATTGT TACAGCCCAC GCCTTCGTAA TAATTTTCTT
PPDSKat2        GGTCTCTCCT AGATAACGAT CAAATCTACA ATACTATTGT TACAGCCCAC GCCTTCGTAA TAATTTTCTT
PPDS3           GTTCACTTCT AGATAACGAC CAAATCTACA ACACTATTGT CACAGcTCAC GCTTTCGTAA TGATTTTTTT
PPDS4           GATCATTACT AGACAATGAC CAAATTTATA ACACAATTGT TACGGCTCAC GCTTTCGTTA TGATTTTCTT
PPDS5           GATCATTACT AGACAATGAC CAAATTTATA ACACAATTGT TACGGCTCAC GCTTTCGTTA TGATTTTCTT
PPKW1           GTTCACTACT AGACAACGAC CAAATTTATA ACACCATTGT TACAGCCCAC GCCTTTGTAA TAATTTTTTT
PPKW2           GTTCACTACT AGACAACGAC CAAATTTATA ACACCATTGT TACAGCCCAC GCCTTTGTAA TAATTTTTTT
PPKW5           GTTCACTACT AGACAACGAC CAAATTTATA ACACCATTGT TACAGCCCAC GCCTTTGTAA TAATTTTTTT
PPKW6           GTTCACTACT AGACAACGAC CAAATTTATA ACACCATTGT TACAGCCCAC GCCTTTGTAA TAATTTTTTT
```

```
            ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
               145        155        165        175        185        195        205
CHA         CATAGTTATG CCTGGAATAA TTGGTGGGTT TGGAAATTGA TTAGTTCCGT TAATAATTGG AGCCCAAGAT
CHC         TATAGTAATG CCTGTAATAA TTGGAGGGTT TGGAAACTGG CTAGTACCCA TGATAATCGG AGCCCAAGAC
HTDB1       CATAGTTATA CCGGTAATAA TTGGAGGATT TGGAAATTGA TTAGTTCCCC TAATAATTGG CGCACAAGAT
HTDB2       CATAGTTATA CCGGTAATAA TTGGAGGATT TGGAAATTGA TTAGTTCCCC TAATAATTGG CGCACAAGAT
PSCR2       CATAGTTATG CCTGGAATAA TTGGTGGGTT TGGAAATTGA TTAGTTCCGT TAATAATTGG AGCCCAAGAT
PSCRA       CATAGTTATG CCCGGAATAA TTGGTGGGTT TGGAAATTGA TTAGTTCCGT TAATAATTGG AGCCCAAGAT
PSCRB       TATAGTAATG CCTGTAATAA TTGGAGGGTT TGGAAACTGG CTAGTACCCA TGATAATCGG AGCCCAAGAC
PSCRC       TATAGTAATG CCTGTTATAA TCGGAGGTTT CGGGAATTGA CTAGTTCCAA TAATAATtGG TGCTCAAGAT
PYCHA       TATAGTAATG CCTGtAATAA TTGGAGGGTT TGGAAACTGG CTAGTACCCA TGATAATCGG AGCCCAAGAC
PYCHB       TATAGTAATG CCTGTAATAA TTGGAGGGTT TGGAAACTGG CTAGTACCCA TGATAATCGG AGCCCAAGAC
PYCHC       CATAGTTATG CCTGGAATAA TTGGTGGGTT TGGAAATTGA TTAGTTCCGT TAATAATTGG AGCCCAAGAT
PPK3        CATAGTAATA CCAGTAATAA TTGGGGGGTT TGGAAACTGG CTAGTACCGC TGATAATTGG GGCACAAGAT
PPK4        CATAGTAATA CCAGTAATAA TTGGGGGGTT TGGAAACTGG CTAGTACCGC TGATAATTGG GGCACAAGAT
PPKA        CATAGTTATG CCTGGAATAA TTGGTGGGTT TGGAAATTGA TTAGTTCCGT TAATAATTGG AGCCCAAGAT
PPKB        CATAGTTATG CCTGGAATAA TTGGtGGGTT TGGAAATTGA TTAGTTCCGT TAATAATTGG AGCCCAAGAT
PPUW1       TATAGTCATG CCTGTTATAA TTGGAGGGTT CGGAAACTGA CTGGTACCTC TGATGATCGG AGCTCAAGAT
PPUW2       TATAGTCATG CCTGTTATAA TTGGAGGGTT CGGAAACTGA CTGGTACCTC TGATGATCGG AGCTCAAGAT
PPUW3       TATAGTCATG CCTGTTATAA TTGGAGGGTT CGGAAACTGA CTGGTACCTC TGATGATCGG AGCTCAAGAT
PPUW4       TATAGTCATG CCTGTTATAA TTGGAGGGTT CGGAAACTGA CTGGTACCTC TGATGATCGG AGCTCAAGAT
PPUW5       TATAGTCATG CCTGTTATAA TTGGAGGGTT CGGAAACTGA CTGGTACCTC TGATGATCGG AGCTCAAGAT
PPUH1       CATAGTTATG CCTGTAATAA TTGGTGGGTT TGGGAATTGA TTAGTTCCGT TAATAATTGG AGCCCAAGAT
PPUH2       CATAGTTATG CCTGTAATAA TTGGTGGGTT TGGAAATTGA TTAGTTCCGT TAATAATTGG AGCCCAAGAT
PPUH3       TATAGTCATG CCCGTAATAA TTGGAGGATT TGGAAACTGA CTTGTCCCCC TAATAATTGG AGCCCAAGAC
PPJYMark1   CATGGTAATG CCAGTGATAA TCGGTGGATT TGGGAATTGA CTAGTTCCAC TGATAATTGG AGCTCAAGAC
PPJYMark2   CATGGTAATG CCAGTGATAA TCGGTGGATT TGGGAATTGA CTAGTTCCAC TGATAATTGG AGCTCAAGAC
PPJYMark3   CATGGTAATG CCAGTGATAA TCGGTGGATT TGGGAATTGG CTAGTTCCAC TGATAATTGG AGCTCAAGAC
PPJYMark4   CATGGTAATG CCAGTGATAA TCGGTGGATT TGGGAATTGG CTAGTTCCAC TGATAATTGG AGCTCAAGAC
PPJY1       CATGGTAATG CCAGTGATAA TCGGAGGATT TGGGAATTGG CTAGTTCCAC TGATAATTGG AGCTCAAGAC
PPJY2       CATGGTAATG CCAGTGATAA TCGGTGGATT TGGGAATTGA CTAGTTCCAC TGATAATTGG AGCTCAAGAC
PPJY3       CATGGTAATG CCAGTGATAA TCGGTGGATT TGGGAATTGG CTAGTTCCAC TGATAATTGG AGCTCAAGAC
PPJY4       CATGGTAATG CCAGTGATAA TCGGTGGATT TGGGAATTGG CTAGTTCCAC TGATAATTGG AGCTCAAGAC
PPJF1       CATGGTAATG CCAGTGATAA TCGGTGGATT TGGGAATTGG CTAGTTCCAC TGATAATTGG AGCTCAAGAC
PPJF2       CATGGTAATG CCAGTGATAA TCGGTGGATT TGGGAATTGG CTAGTTCCAC TGATAATTGG AGCTCAAGAC
PPJF3       CATGGTAATG CCAGTGATAA TCGGTGGATT TGGGAATTGG CTAGTTCCAC TGATAATTGG AGCTCAAGAC
PPJF4       CATGGTAATG CCAGTGATAA TCGGTGGATT TGGGAATTGG CTAGTTCCAC TGATAATTGG AGCTCAAGAC
PPJF5       CATGGTAATG CCAGTGATAA TCGGTGGATT TGGGAATTGG CTAGTTCCAC TGATAATTGG AGCTCAAGAC
PPIE1       CATAGTTATA CCGGTAATAA TTGGAGGATT TGGAAATTGA TTAGTTCCCC TAATAATTGG CGCACAAGAT
PPIE2       TATAGTCATG CCTGTTATAA TTGGAGGGTT CGGAAACTGA CTGGTACCTC TGATGATCGG AGCTCAAGAT
PPIE3       TATAGTCATG CCTGTTATAA TTGGAGGGTT CGGAAACTGA CTGGTACCTC TGATGATCGG AGCTCAAGAT
PPIE4       CATGGTAATG CCAGTGATAA TCGGTGGATT TGGGAATTGG CTAGTTCCAC TGATAATTGG AGCTCAAGAC
PPIC1       CATGGTAATG CCAGTGATAA TCGGTGGATT TGGGAATTGG CTAGTTCCAC TGATAATTGG AGCTCAAGAC
PPIC2       TATAGTCATG CCTGTTATAA TTGGAGGGTT CGGAAACTGA CTGGTACCTC TGATGATCGG AGCTCAAGAT
PPIC3       TATAGTCATG CCTGTTATAA TTGGAGGGTT CGGAAACTGA CTGGTACCTC TGATGATCGG AGCTCAAGAT
PPIS1       CATAGTAATA CCAGTAATAA TTGGGGGGTT TGGAAACTGG CTAGTACCGC TGATAATTGG GGCACAAGAT
PPIS2       CATAGTAATA CCGGTAATAA TTGGGGGGTT TGGAAACTGG CTAGTACCGC TGATAATTGG GGCACAAGAT
PPIS3       CATAGTAATA CCAGTAATAA TTGGGGGGTT TGGAAACTGG CTAGTACCGC TGATAATTGG GGCACAAGAT
PPIM2       CATAGTAATA CCAGTAATAA TTGGGGGGTT TGGAAACTGG CTAGTACCGC TGATAATTGG GGCACAAGAT
PPI03       CATGGTAATG CCTGTAATAA TTGGTGGATT CGGGAATTGA TTAGTTCCAT TAATAATCGG AGCTCAAGAC
PPOG1       CATGGTAATG CCAGTGATAA TCGGTGGATT TGGGAATTGG CTAGTTCCAC TGATAATTGG AGCTCAAGAC
PPOG2       CATAGTAATA CCAGTAATAA TTGGGGGGTT TGGAAACTGG CTAGTACCGC TGATAATTGG GGCACAAGAT
PPOG3       CATAGTAATA CCAGTAATAA TTGGGGGGTT TGGAAACTGG CTAGTACCGC TGATAATTGG GGCACAAGAT
PPOG4       CATAGTAATA CCGGTAATAA TTGGGGGGTT TGGAAACTGG CTAGTACCGC TGATAATTGG GGCACAAGAT
PPOG5       CATAGTAATA CCAGTAATAA TTGGGGGGTT TGGAAACTGG CTAGTACCGC TGATAATTGG GGCACAAGAT
PPOG6       CATAGTAATA CCAGTAATAA TTGGGGGGTT TGGAAACTGG CTAGTACCGC TGATAATTGG GGCACAAGAT
PPNB3       CATAGTTATG CCTGTAATAA TTGGTGGGTT TGGAAATTGA TTAGTTCCGT TAATAATTGG AGCCCAAGAT
PPNB5       CATAGTTATG CCTGTAATAA TTGGTGGGTT TGGAAATTGA TTAGTTCCGT TAATAATTGG AGCCCAAGAT
PPBC1       CATAGTTATG CCTGTAATAA TTGGTGGGTT TGGAAATTGA TTAGTTCCGT TAATAATTGG AGCCCAAGAT
PPBC4       CATAGTTATG CCTGTAATAA TTGGTGGGTT TGGAAATTGA TTAGTTCCGT TAATAATTGG AGCCCAAGAT
PPBA2       CATAGTTATG CCTGTAATAA TTGGTGGGTT TGGAAATTGA TTAGTTCCGT TAATAATTGG AGCCCAAGAT
PPBA4       CATAGTTATG CCTGTAATAA TTGGTGGGTT TGGAAATTGA TTAGTTCCGT TAATAATTGG AGCCCAAGAT
PPBO1       CATAGTTATG CCTGGAATAA TTGGTGGGTT TGGAAATTGA TTAGTTCCGT TAATAATTGG AGCCCAAGAT
PPBO2       CATAGTTATG CCTGTAATAA TTGGTGGGTT TGGAAATTGA TTAGTTCCGT TAATAATTGG AGCCCAAGAT
PPDB21      CATAGTCATG CCTGTAATAA TTGGTGGGTT TGGAAATTGA TTAGTTCCGT TAATAATTGG AGCTCAAGAT
PPDSKat1    CATGGTAATG CCAGTGATAA TCGGTGGATT TGGGAATTGG CTAGTTCCAC TGATAATTGG AGCTCAAGAC
PPDSKat2    CATGGTAATG CCAGTGATAA TCGGTGGATT TGGGAATTGG CTAGTTCCAC TGATAATTGG AGCTCAAGAC
PPDS3       CATAGTAATA CCAGTAATAA TTGGGGGGTT TGGAAACTGG CTAGTACCGC TGATAATTGG GGCACAAGAT
PPDS4       TATAGTCATG CCTGTTATAA TTGGAGGGTT CGGAAACTGA CTGGTACCTC TGATGATCGG AGCTCAAGAT
PPDS5       TATAGTCATG CCTGTTATAA TTGGAGGGTT CGGAAACTGA CTGGTACCTC TGATGATCGG AGCTCAAGAT
PPKW1       CATAGTTATG CCTGTAATAA TTGGTGGGTT TGGAAATTGA TTAGTTCCGT TAATAATTGG AGCCCAAGAT
PPKW2       CATAGTTATG CCTGTAATAA TTGGTGGGTT TGGAAATTGA TTAGTTCCGT TAATAATTGG AGCCCAAGAT
PPKW5       CATAGTTATG CCTGTAATAA TTGGTGGGTT TGGAAATTGA TTAGTTCCGT TAATAATTGG AGCCCAAGAT
PPKW6       CATAGTTATG CCTGTAATAA TTGGTGGGTT TGGAAATTGA TTAGTTCCGT TAATAATTGG AGCCCAAGAT
```

```
            ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
              215        225        235        245        255        265        275
CHA         ATAGCTTTCC CTCGTATAAA TAATATGAGA TTCTGACTTC TTCCCCCATC ACTGTGTCTC CTTTCAGCAT
CHC         ATGGCATTCC CaCGAATGAA TAACATGAGA TTCTGATTAC TACCACCCTC TCTCACTCTT TTGTCAGCAT
HTDB1       ATAGCCTTCC CACGAATAAA CAACATGAGA TTCTGACTTC TTCCGCCCTC TTTATGCCTT CTTTCGGCTT
HTDB2       ATAGCCTTCC CACGAATAAA CAACATGAGA TTCTGACTTC TTCCGCCCTC TTTATGCCTT CTTTCGGCTT
PSCR2       ATAGCTTTCC CTCGTATAAA TAATATGAGA TTCTGACTTC TTCCCCCATC ACTGTGTCTC CTTTCAGCAT
PSCRA       ATAGCTTTCC CTCGTATAAA TAATATGAGA TTCTGACTTC TTCCCCCATC ACTGTGTCTC CTTTCAGCAT
PSCRB       ATGGCATTCC CACAAATGAA TAACATGAGA TTCTGATTAC TACCACCCTC TcTCACTCTT TTGTCAGCAT
PSCRC       ATGGCATTCC CTCGAATAAA TAACATGAGA TTTTGACTAC TGCCTCCCTC TTTAACACTC CTATCAGCCT
PYCHA       ATGGCATTCC CACGAATGAA TAACATGAGA TTCTGATTAC TACCACCCTc TCTCACTCTT TTGTCAGCAT
PYCHB       ATGGCATTCC CACGAATGAA TAACATGAGA TTCTGATTAC TACCACCCTC TCTCACTCTT TTGTCAGCAT
PYCHC       ATAGCTTTCC CTCGTATAAA TAATATGAGA TTcTGACTTC TTCCCCCATC ACTGTGTCTC CTTTCAGCAT
PPK3        ATGGCCTTCC CCCGCATAAA TAATATAAGA TTCTGACTCC TTCCTCCTTC TCTATGTCTT CTTTCCGCAT
PPK4        ATGGCCCTCC CCCGCATAAA TAATATAAGA TTCTGACTCC TTCCTCCTTC TCTATGTCTT CTTTCCGCAT
PPKA        ATAGCTTTCC CTCGTATAAA TAATATGAGA TTCTGACTTC TTCCCCCATC ACTGTGTCTC CTTTCAGCAT
PPKB        ATAGCTTTCC CTCGTATAAA TAATATGAGA TTCTGACTTC TTCCCCCATC ACTGtGTCTC CTTTCAGCAT
PPUW1       ATGGCTTTTC CCCGTATGAA CAACATGAGA TTCTGGCTTC TCCCACCATC CTTGTGTCTT CTTTCAGCCT
PPUW2       ATGGCTTTTC CCCGTATGAA CAACATGAGA TTCTGGCTTC TCCCACCATC CTTGTGTCTT CTTTCAGCCT
PPUW3       ATGGCTTTTC CCCGTATGAA CAACATGAGA TTCTGGCTTC TCCCACCATC CTTGTGTCTT CTTTCAGCCT
PPUW4       ATGGCTTTTC CCCGTATGAA CAACATGAGA TTCTGGCTTC TCCCACCATC CTTGTGTCTT CTTTCAGCCT
PPUW5       ATGGCTTTTC CCCGTATGAA CAACATGAGA TTCTGGCTTC TCCCACCATC CTTGTGTCTT CTTTCAGCCT
PPUH1       ATAGCTTTCC CTCGTATAAA TAATATGAGA TTCTGACTTC TTCCCCCATC ACTGTGTCTC CTTTCAGCAT
PPUH2       ATAGCTTTCC CTCGTATAAA TAATATGAGA TTCTGACTTC TTCCCCCATC ACTGTGTCTC CTTTCAGCAT
PPUH3       ATAGCATTCC CACGAATAAA CAACATAAGG TTTTGGCTTC TACCTCCGTC CTTATGTCTT CTATCAGCAT
PPJYMark1   ATGGCCTTCC CTCGTATGAA CAACATAAGA TTCTGGCTCC TTCCTCCATC ACTGTGCCTC CTCTCAGCTT
PPJYMark2   ATGGCCTTCC CTCGTATGAA CAACATAAGA TTCTGGCTCC TTCCTCCATC ACTGTGCCTC CTCTCAGCTT
PPJYMark3   ATGGCCTTCC CTCGTATGAA CAACATAAGA TTCTGGCTCC TTCCTCCATC ACTGTGCCTC CTCTCAGCTT
PPJYMark4   ATGGCCTTCC CTCGTATGAA CAACATAAGA TTCTGGCTCC TTCCTCCATC ACTGTGCCTC CTCTCAGCTT
PPJY1       ATGGCCTTCC CTCGTATGAA CAACATAAGA TTCTGGCTCC TTCCTCCATC ACTGTGCCTC CTCTCAGCTT
PPJY2       ATGGCCTTCC CTCGTATGAA CAACATAAGA TTCTGGCTCC TTCCTCCATC ACTGTGCCTC CTCTCAGCTT
PPJY3       ATGGCCTTCC CTCGTATGAA CAACATAAGA TTCTGGCTCC TTCCTCCATC ACTGTGCCTC CTCTCAGCTT
PPJY4       ATGGCCTTCC CTCGTATGAA CAACATAAGA TTCTGGCTCC TTCCTCCATC ACTGTGCCTC CTCTCAGCTT
PPJF1       ATGGCCTTCC CTCGTATGAA CAACATAAGA TTCTGGCTCC TTCCTCCATC ACTGTGCCTC CTCTCAGCTT
PPJF2       ATGGCCTTCC CTCGTATGAA CAACATAAGA TTCTGGCTCC TTCCTCCATC ACTGTGCCTC CTCTCAGCTT
PPJF3       ATGGCCTTCC CTCGTATGAA CAACATAAGA TTCTGGCTCC TTCCTCCATC ACTGTGCCTC CTCTCAGCTT
PPJF4       ATGGCCTTCC CTCGTATGAA CAACATAAGA TTCTGGCTCC TTCCTCCATC ACTGTGCCTC CTCTCAGCTT
PPJF5       ATGGCCTTCC CTCGTATGAA CAACATAAGA TTCTGGCTCC TTCCTCCATC ACTGTGCCTC CTCTCAGCTT
PPIE1       ATAGCCTTCC CACGAATAAA CAACATGAGA TTCTGACTTC TTCCGCCCTC TTTATGCCTT CTTTCGGCTT
PPIE2       ATGGCTTTTC CCCGTATGAA CAACATGAGA TTCTGGCTTC TCCCACCATC CTTGTGTCTT CTTTCAGCCT
PPIE3       ATGGCTTTTC CCCGTATGAA CAACATGAGA TTCTGGCTTC TCCCACCATC CTTGTGTCTT CTTTCGGCCT
PPIE4       ATGGCCTTCC CTCGTATGAA CAACATAAGA TTCTGGCTCC TTCCTCCATC ACTGTGCCTC CTCTCAGCTT
PPIC1       ATGGCCTTCC CTCGTATGAA CAACATAAGA TTCTGGCTCC TTCCTCCATC ACTGTGCCTC CTCTCAGCTT
PPIC2       ATGGCTTTTC CCCGTATGAA CAACATGAGA TTCTGGCTTC TCCCACCATC CTTGTGTCTT CTTTCAGCCT
PPIC3       ATGGCTTTTC CCCGTATGAA CAACATGAGA TTCTGGCTTC TCCCACCATC CTTGTGTCTT CTTTCAGCCT
PPIS1       ATGGCCTTCC CCCGCATAAA TAATATAAGA TTCTGACTCC TTCCTCCTTC TCTATGTCTT CTTTCCTCAT
PPIS2       ATGGCCTTCC CCCGCATAAA TAATATAAGA TTCTGACCCC TTCCTCCTTC TCTATGTCTT CTTTCCGCAT
PPIS3       ATGGCCTTCC CCCGCATAAA TAATATAAGA TTCTGACTCC TTCCTCCTTC TCTATGTCTT CTTTCCTCAT
PPIM2       ATGGCCTTCC CCCGCATAAA TAATATAAGA TTCTGACTCC TTCCTCCTTC TCTATGTCTT CTTTCCTCAT
PPI03       ATGGCCTTCC CTCGAATAAA CAACATAAGA TTCTGGCTCC TCCCTCCGTC ACTATGCCTC CTCTCAGCAT
PPOG1       ATGGCCTTCC CTCGTATGAA CAACATAAGA TTCTGGCTCC TTCCTCCATC ACTGTGCCTC CTCTCAGCTT
PPOG2       ATGGCCTTCC CCCGCATAAA TAATATAAGA TTCTGACTCC TTCCTCCTTC TCTATGTCTT CTTTCCGCAT
PPOG3       ATGGCCTTCC CCCGCATAAA TAATATAAGA TTCTGACTCC TTCCTCCTTC TCTATGTCTT CTTTCCGCAT
PPOG4       ATGGCCTTCC CCCGCATAAA TAATATAAGA TTCTGACTCC TTCCTCCTTC TCTATGTCTT CTTTCCGCAT
PPOG5       ATGGCCTTCC CCCGCATAAA TAATATAAGA TTCTGACTCC TTCCTCCTTC TCTATGTCTT CTTTCCGCAT
PPOG6       ATGGCCTTCC CCCGCATAAA TAATATAAGA TTCTGACTCC TTCCTCCTTC TCTATGTCTT CTTTCCGCAT
PPNB3       ATAGCTTTCC CTCGTATAAA TAATATGAGA TTCTGACTTC TTCCCCCATC ACTGTGTCTC CTTTCAGCAT
PPNB5       ATAGCTTTCC CTCGTATAAA TAATATGAGA TTCTGACTTC TTCCCCCATC ACTGTGTCTC CTTTCAGCAT
PPBC1       ATAGCTTTCC CTCGTATAAA TAATATGAGA TTCTGACTTC TTCCCCCATC ACTGTGTCCC CTTTCAGCAT
PPBC4       ATAGCTTTCC CTCGTATAAA TAATATGAGA TTCTGACTTC TTCCCCCATC ACTGTGTCTC CTTTCAGCAT
PPBA2       ATAGCTTTCC CTCGTATAAA TAATATGAGA TTCTGACTTC TTCCCCCATC ACTGTGTCTC CTTTCAGCAT
PPBA4       ATAGCTTTCC CTCGTATAAA TAATATGAGA TTCTGACTTC TTCCCCCATC ACTGTGTCTC CTTTCAGCAT
PPBO1       ATAGCTTTCC CTCGTATAAA TAATATGAGA TTCTGACTTC TTCCCCCATC ACTGTGTCTC CTTTCAGCAT
PPBO2       ATAGCTTTCC CTCGTATAAA TAATATGAGA TTCTGACTTC TTCCCCCATC ACTGTGTCTC CTTTCAGCAT
PPDB21      ATGGCTTTCC CTCGTATAAA TAATATGAGA TTCTGGCTCC TTCCTCCATC CCTATGTCTC CTTTCAGCAT
PPDSKat1    ATGGCCTTCC CTCGTATGAA CAACATAAGA TTCTGGCTCC TTCCTCCATC ACTGTGCCTC CTCTCAGCTT
PPDSKat2    ATGGCCTTCC CTCGTATGAA CAACATAAGA TTCTGGCTCC TTCCTCCATC ACTGTGCCTC CTCTCAGCTT
PPDS3       ATGGCCTTcC CCCGCATAAA TAATATAAGA TTCTGACTCC TTCCTCCTTC TCTATGTcTT cTTTCCTCAT
PPDS4       ATGGCTTTTC CCCGTATGAA CAACATGAGA TTCTGGCTTC TCCCACCATC CTTGTGTCTT CTTTCAGCCT
PPDS5       ATGGCTTTTC CCCGTATGAA CAACATGAGA TTCTGGCTTC TCCCACCATC CTTGTGTCTT CTTTCAGCCT
PPKW1       ATAGCTTTCC CTCGTATAAA TAATATGAGA TTCTGACTTC TTCCCCCATC ACTGTGTCTC CTTTCAGCAT
PPKW2       ATAGCTTTCC CTCGTATAAA TAATATGAGA TTCTGACTTC TTCCCCCATC ACTGTGTCTC CTTTCAGCAT
PPKW5       ATAGCTTTCC CTCGTATAAA TAATATGAGA TTCTGACTTC TTCCCCCATC ACTGTGTCTC CTTTCAGCAT
PPKW6       ATAGCTTTCC CTCGTATAAA TAATATGAGA TTCTGACTTC TTCCCCCATC ACTGTGTCTC CTTTCAGCAT
```

```
                ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                   285        295        305        315        325        335        345
CHA             CAGCCTTTGC CGGGCTAGGA GTGGGAACTG GGTGAACTGT CTACCCCCCT CTAGCCGGGA ACCTCTTCCA
CHC             CTGCCTTTGC AGGAATAGGA GCAGGAACAG GATGAACAGT ATATCCTCCC TTAGCCGGAA ATTTATTTCA
HTDB1           CAGCCTTTGC TGGTCTAGGT GTAGGAACCG GATGAACAGT GTATCCTCCA TTGGCTGGAA ACCTGTTTCA
HTDB2           CAGCCTTTGC TGGTCTAGGT GTAGGAACCG GATGAACAGT GTATCCTCCA TTGGCTGGAA ACCTGTTTCA
PSCR2           CAGCCTTTGC CGGGCTAGGA GTGGGAACTG GGTGAACTGT CTACCCCCCT CTAGCCGGGA ACCTCTTCCA
PSCRA           CAGCCTTTGC CGGGCTAGGA GTGGGAACTG GGTGAACTGT CTACCCCCCT CTAGCCGGGA ACCTCTTCCA
PSCRB           CTGCCTTTGC AGGAATAGGA GCAGGAACAG GATGAACAGT ATATCCTCCC TTAGCCGGAA ATTTATTTCA
PSCRC           CAGCTTTTGC AGGAATAGGG GCTGGAACAG GATGAACGGT ATACCCGCCA TTGGCTGGAA ATCTCTTCCA
PYCHA           CTGCCTTTGC AGGAATAGGA GCAGGAACAG GATGAACAGT ATATCCTCCC TTAGCCGGAA ATTTATTTCA
PYCHB           CTGCCTTTGC AGGAATAGGA GCAGGAACAG GATGAACAGT ATATCCTCCC TTAGCCGGAA ATTTATTTCA
PYCHC           CAGCCTTTGC CGGGCTAGGA GTGGGAACTG GGTGAACTGT CTACCCCCCT CTAGCCGGGA ACCTCTTCCA
PPK3            CTGCTTTCGC CGGATTAGGT GTAGGGACCG GGTGAACCGT GTACCCCCCA CTAGCTGGAA ATCTTTTCCA
PPK4            CTGCTTTCGC CGGATTAGGT GTAGGGACCG GGTGAACCGT GTACCCCCCA CTAGCTGGAA ATCTTTTCCA
PPKA            CAGCCTTTGC CGGGCTAGGA GTGGGAACTG GGTGAGCTGT CTACCCCCCT CTAGCCGGGA ACCTCTTCCA
PPKB            CAGCCTTTGC CGGGCTAGGA GTGGGAACTG GGTGAACTGT CTACCCCCCT CTAGCCGGGA ACCTCTTCCA
PPUW1           CTGCTTTCGC AGGGTTAGGT GTAGGGACAG GATGGACAGT ATACCCCCCA CTCGCTGGGA ATTTATTTCA
PPUW2           CTGCTTTCGC AGGGTTAGGT GTAGGGACAG GATGGACAGT ATACCCCCCA CTCGCTGGGA ATTTATTTCA
PPUW3           CTGCTTTCGC AGGGTTAGGT GTAGGGACAG GATGGACAGT ATACCCCCCA CTCGCTGGGA ATTTATTTCA
PPUW4           CTGCTTTCGC AGGGTTAGGT GTAGGGACAG GATGGACAGT ATACCCCCCA CTCGCTGGGA ATTTATTTCA
PPUW5           CTGCTTTCGC AGGGTTAGGT GTAGGGACAG GATGGACAGT ATACCCCCCA CTCGCTGGGA ATTTATTTCA
PPUH1           CAGCCTTTGC CGGGCTAGGA GTGGGAACTG GGTGAACTGT CTACCCCCCT CTAGCCGGGA ACCTCTTCCA
PPUH2           CAGCCTTTGC CGGGCTAGGA GTGGGAACTG GGTGAACTGT CTACCCCCCT CTAGCCGGGA ACCTCTTCCA
PPUH3           CTGCCTTCGC AGGACTAGGG GTAGGAACAG GATGAACCGT ATATCCACCT CTAGCAGGAA ACTTATTCCA
PPJYMark1       CTGCTTTTGC TGGATTAGGG GTGGGAACCG GATGAACTGT GTATCCTCCC CTAGCAGGAA ATCTTTTCCA
PPJYMark2       CTGCTTTTGC TGGATTAGGG GTGGGAACCG GATGAACTGT GTATCCTCCC CTAGCAGGAA ATCTTTTCCA
PPJYMark3       CTGCTTTTGC TGGATTAGGA GTGGGAACCG GATGAACTGT ATATCCTCCC CTAGCAGGAA ATCTTTTCCA
PPJYMark4       CTGCTTTTGC TGGATTAGGA GTGGGAACCG GATGAACTGT ATATCCTCCC CTAGCAGGAA ATCTTTTCCA
PPJY1           CTGCTTTTGC TGGATTAGGA GTGGGAACCG GATGAACTGT ATATCCTCCC CTAGCAGGAA ATCTTTTCCA
PPJY2           CTGCTTTTGC TGGATTAGGG GTGGGAACCG GATGAACTGT GTATCCTCCC CTAGCAGGAA ATCTTTTcCA
PPJY3           CTGCTTTTGC TGGATTAGGA GTGGGAACCG GATGAACTGT ATATCCTCCC CTAGCAGGAA ATCTTTTCCA
PPJY4           CTGCTTTTGC TGGATTAGGA GTGGGAACCG GATGAACTGT ATATCCTCCC CTAGCAGGAA ATCTTTTCCA
PPJF1           CTGCTTTTGC TGGATTAGGA GTGGGAACCG GATGAACTGT ATATCCTCCC CTAGCAGGAA ATCTTTTCCA
PPJF2           CTGCTTTTGC TGGATTAGGA GTGGGAACCG GATGAACTGT ATATCCTCCC CTAGCAGAAA ATCTTTTCCA
PPJF3           CTGCTTTTGC TGGATTAGGA GTGGGAACCG GATGAACTGT ATATCCTCCC CTAGCAGGAA ATCTTTTCCA
PPJF4           CTGCTTTTGC TGGATTAGGA GTGGGAACCG GATGAACTGT ATATCCTCCC CTAGCAGGAA ATCTTTTCCA
PPJF5           CTGCTTTTGC TGGATTAGGA GTGGGAACCG GATGAACTGT ATATCCTCCC CTAGCAGGAA ATCTTTTCCA
PPIE1           CAGCCTTTGC TGGTCTAGGT GTAGGAACCG GATGAACAGT GTATCCTCCA TTGGCTGGAA ACCTGTTTCA
PPIE2           CTGCTTTCGC AGGGTTAGGT GTAGGGACAG GATGGACAGT ATACCCCCCA CTCGCTGGGA ATTTATTTCA
PPIE3           CTGCTTTCGC AGGGTTAGGT GTAGGGACAG GATGGACAGT ATACCCCCCA CTCGCTGGGA ATTTATTTCA
PPIE4           CTGCTTTTGC TGGATTAGGA GTGGGAACCG GATGAACTGT ATATCCTCCC CTAGCAGGAA ATCTTTTCCA
PPIC1           CTGCTTTTGC TGGATTAGGA GTGGGAACCG GATGAACTGT ATATCCTCCC CTAGCAGGAA ATCTTTTCCA
PPIC2           CTGCTTTCGC AGGGTTAGGT GTAGGGACAG GATGGACAGT ATACCCCCCA CTCGCTGGGA ATTTATTTCA
PPIC3           CTGCTTTCGC AGGGTTAGGT GTAGGGACAG GATGGACAGT ATACCCCCCA CTCGCTGGGA ATTTATTTCA
PPIS1           CTGCTTTCGC CGGATTAGGT GTAGGGACCG GGTGAACCGT GTACCCCCCA CTAGCTGGAA ATCTTTTCCA
PPIS2           CTGCTTTCGC CGGATTAGGT GTAGGGACCG GGTGAACCGT GTACCCCCCA CTAGCTGGAA ATCTTTTCCA
PPIS3           CTGCTTTCGC CGGATTAGGT GTAGGGACCG GGTGAACCGT GTACCCCCCA CTAGCTGGAA ATCTTTTCCA
PPIM2           CTGCTTTCGC CGGATTAGGT GTAGGGACCG GGTGAACCGT GTACCCCCCA CTAGCTGGAA ATCTTTTCCA
PPI03           CTGCTTTTGC CGGATTAGGA GTGGGGACTG GCTGAACTGT CTACCCTCCC TTGGCTGGAA ATCTTTTCCA
PPOG1           CTGCTTTTGC TGGATTAGGA GTGGGAACCG GATGAACTGT ATATCCTCCC CTAGCAGGAA ATCTTTTCCA
PPOG2           CTGCTTTCGC CGGATTAGGT GTAGGGACCG GGTGAACCGT GTACCCCCCA CTAGCTGGAA ATCTTTTCCA
PPOG3           CTGCTTTCGC CGGATTAGGT GTAGGAACCG GGTGAACCGT GTACCCCCCA CTAGCTGGAA ATCTTTTCCA
PPOG4           CTGCTTTCGC CGGATTAGGT GTAGGGACCG GGTGAACCGT GTACCCCCCA CTAGCTGGAA ATCTTTTCCA
PPOG5           CTGCTTTCGC CGGATTAGGT GTAGGGACCG GGTGAACCGT GTACCCCCCA CTAGCTGGAA ATCTTTTCCA
PPOG6           CTGCTTTCGC CGGATTAGGT GTAGGGACCG GGTGAACCGT GTACCCCCCA CTAGCTGGAA ATCTTTTCCA
PPNB3           CAGCCTTTGC CGGGCTAGGA GTGGGAACTG GGTGAACTGT CTACCCCCCT CTAGCCGGGA ACCTCTTCCA
PPNB5           CAGCCTTTGC CGGGCTAGGA GTGGGAACTG GGTGAACTGT CTACCCCCCT CTAGCCGGGA ACCTCTTCCA
PPBC1           CAGCCTTTGC CGGGCTAGGA GTGGGAACTG GGTGAACTGT CTACCCCCCT CTAGCCGGGA ACCTCTTCCA
PPBC4           CAGCCTTTGC CGGGCTAGGA GTGGGAACTG GGTGAACTGT CTACCCCCCT CTAGCCGGGA ACCTCTTCCA
PPBA2           CAGCCTTTGC CGGGCTAGGA GTGGGAACTG GGTGAACTGT CTACCCCCCT CTAGCCGGGA ACCTCTTCCA
PPBA4           CAGCCTTTGC CGGGCTAGGA GTGGGAACTG GGTGAACTGT CTACCCCCCT CTAGCCGGGA ACCTCTTCCA
PPBO1           CAGCCTTTGC CGGGCTAGGA GTGGGAACTG GGTGAACTGT CTACCCCCCT CTAGCCGGGA ACCTCTTCCA
PPBO2           CAGCCTTTGC CGGGCTAGGA GTGGGAACTG GGTGAACTGT CTACCCCCCT CTAGCCGGGA ACCTCTTCCA
PPDB21          CAGCCTTTGC TGGGCTAGGA GTGGGAACTG GGTGAACTGT CTACCCCCCT CTAGCCGGAA ACCTCTTCCA
PPDSKat1        CTGCTTTTGC TGGATTAGGA GTGGGAACCG GATGAACTGT ATATCCTCCC CTAGCAGGAA ATCTTTTCCA
PPDSKat2        CTGCTTTTGC TGGATTAGGA GTGGGAACCG GATGAACTGT ATATCCTCCC CTAGCAGGAA ATCTTTTCCA
PPDS3           CTGCTTTCGC CGGATTAgGT GTAGGGACCG GGTGAACCGT GTACCCCCCA CTAGCTGGAA ATCTTTTTCA
PPDS4           CTGCTTTCGC AGGGTTAGGT GTAGGGACAG GATGGACAGT ATACCCCCCA CTCGCTGGGA ATTTATTTCA
PPDS5           CTGCTTTCGC AGGGTTAGGT GTAGGGACAG GATGGACAGT ATACCCCCCA CTCGCTGGGA ATTTATTTCA
PPKW1           CAGCCTTTGC CGGGCTAGGA GTGGGAACTG GGTGAACTGT CTACCCCCCT CTAGCCGGGA ACCTCTTCCA
PPKW2           CAGCCTTTGC CGGGCTAGGA GTGGGAACTG GGTGAACTGT CTACCCCCCT CTAGCCGGGA ACCTCTTCCA
PPKW5           CAGCCTTTGC CGGGCTAGGA GTGGGAACTG GGTGAACTGT CTACCCCCCT CTAGCCGGGA ACCTCTTCCA
PPKW6           CAGCCTTTGC CGGGCTAGGA GTGGGAACTG GGTGAACTGT CTACCCCCCT CTAGCCGGGA ACCTCTTCCA
```

```
                ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                   355        365        375        385        395        405        415
CHA             CTCAGGACTT TCCGTAGATC TCGCAATCTT TAGCCTTCAC CTAGCTGGTG CGTCCTCCAT TCTTGGAGCT
CHC             TTCCAGAATC TCAGTAGACC TAGCTATTTT TAGACTACAT TTAGCTGGGG CCTCCTCCAT CTTAGGGGCA
HTDB1           CTCTGGATTC TCTGTAGATC TAGCGATTTT CAGGtTACAC TTAGCAGGAG CTTCATCAAT CTTAGGAGCT
HTDB2           CTCTGGATTC TCTGTAGATC TAGCGATTTT CAGGTTACAC TTAGCAGGAG CTTCATCAAT CTTAGGAGCT
PSCR2           CTCAGGATTT TCCGTAGATC TCGCAATCTT TAGCCTTCAC CTAGCTGGTG CGTCCTCCAT TCTTGGAGCT
PSCRA           CTCAGGATTT TCCGTAGATC TCGCAATCTT TAGCCTTCAC CTAGCTGGTG CGTCCTCCAT TCTTGGAGCT
PSCRB           TTCCAGAATC TCAGTAGACC TAGCTATTTT TAGACTACAT TTAGCTGGGG CCTCCTCCAT CTTAGGGGCA
PSCRC           CTCTAGAATT TCAGTAGATT TGGCAATTTT CAGATTACAC CTAGCTGGTG CCTCTTCCAT TTTAGGGGCA
PYCHA           TTCCAGAATC TCAGTAGACC TAGCTATTTT TAGACTACAT TTAGCTGGGG CCTCCTCCAT CTTAGGGGCA
PYCHB           TTCCAGAATC TCAGTAGACC TAGCTATTTT TAGACTACAT TTAGCTGGGG CCTCCTCCAT CTTAGGGGCA
PYCHC           CTCAGGACTT TCCGTAGATC TCGCAATCTT TAGCCTTCAC CTAGCTGGTG CGTCCTCCAT TCTTGGAGCT
PPK3            CTCAGGATTT TCAGTAGACC TTGCAATTTT CAGGCTCCAT CTTGCAGGAG CTTCTTCTAT CTTGGGGGCC
PPK4            CTCAGGATTT TCAGTAGACC TTGCAATTTT CAGGCTCCAT CTTGCAGGAG CTTCTTCTAT CTTGGGGGCC
PPKA            CTCAGGATTT TCCGTAGATC TCGCAATCTT TAGCCTTCAC CTAGCTGGTG cGTCCTCCAT TCTTGGAGCT
PPKB            CTCAGGATTT TCCGTAGATC TCGCAATCTT TAGCCTTCAC CTAGCTGGTG CGTCCTCCAT TCTTGGAGCT
PPUW1           CTCCGGTTTC TCTGTAGATT TAGCAATCTT TAGCCTTCAC CTAGCGGGTG CCTCTTCCAT TTTAGGTGCC
PPUW2           CTCCGGTTTC TCTGTAGATT TAGCAATCTT TAGCCTTCAC CTAGCGGGTG CCTCTTCCAT TTTAGGTGCC
PPUW3           CTCCGGTTTC TCTGTAGATT TAGCAATCTT TAGCCTTCAC CTAGCGGGTG CCTCTTCCAT TTTAGGTGCC
PPUW4           CTCCGGTTTC TCTGTAGATT TAGCAATCTT TAGCCTTCAC CTAGCGGGTG CCTCTTCCAT TTTAGGTGCC
PPUW5           CTCCGGTTTC TCTGTAGATT TAGCAATCTT TAGCCTTCAC CTAACGGGTG CCTCTTCCAT TTTAGGTGCC
PPUH1           CTCAGGATTT TCCGTAGATC TCGCAATCTT TAGCCTTCAC CTAGCTGGTG CGTCCTCCAT TCTTGGAGCT
PPUH2           CTCAGGATTT TCCGTAGATC TCGCAATCTT TAGCCTTCAC CTAGCTGGTG CGTCCTCCAT TCTTGGAGCT
PPUH3           CTCTGGATTC TCAGTTGATT TGGCAATTTT CAGCCTCCAC TTAGCCGGAG CCTCTTCCAT CTTAGGGGCC
PPJYMark1       CTCAGGATTT TCTGTAGATC TTGCAATTTT TAGCCTCCAC TTGGCAGGTG CCTCCTCTAT TCTGGGGGCC
PPJYMark2       CTCAGGATTT TCTGTAGATC TTGCAATTTT TAGCCTCCAC TTGGCAGGTG CCTCCTCTAT TCTGGGGGCC
PPJYMark3       CTCAGGATTT TCTGTAGATC TTGCAATTTT TAGCCTCCAC TTGGCAGGTG CCTCCTCTAT TCTGGGGGCC
PPJYMark4       CTCAGGATTT TCTGTAGATC TTGCAATTTT TAGCCTCCAC TTGGCAGGTG CCTCCTCTAT TCTGGGGGCC
PPJY1           CTCAGGATTT TCTGTAGATC TTGCAATTTT TAGCCTCCAC TTGGCAGGTG CCTCCTCTAT TCTGGGGGCC
PPJY2           CTCAGGATTT TCTGTAGATC TtGCAATTTT TAGCCTCCAC TTGGCAGGTG CCTCCTCTAT TCTGGGGGCC
PPJY3           CTCAGGATTT TCTGTAGATC TTGCAATTTT TAGCCTCCAC TTGGCAGGTG CCTCCTCTAT TCTGGGGGCC
PPJY4           CTCAGGATTT TCTGTAGATC TTGCAATTTT TAGCCTCCAC TTGGCAGGTG CCTCCTCTAT TCTGGGGGCC
PPJF1           CTCAGGATTT TCTGTAGATC TTGCAATTTT TAGACTCCAC TTGGCAGGTG CCTCCTCTAT TCTGGGGGCC
PPJF2           CTCAGGATTT TCTGTAGATC TTGCAATTTT TAGACTCCAC TTGGCAGGTG CCTCCTCTAT TCTGGGGGCC
PPJF3           CTCAGGATTT TCTGTAGATC TTGCAATTTT TAGACTCCAC TTGGCAGGTG CCTCCTCTAT TCTGGGGGCC
PPJF4           CTCAGGATTT TCTGTAGATC TTGCAATTTT TAGACTCCAC TTGGCAGGTG CCTCCTCTAT TCTGGGGGCC
PPJF5           CTCAGGATTT TCTGTAGATC TTGCAATTTT TAGACTCCAC TTGGCAGGTG CCTCCTCTAT TCTGGGGGCC
PPIE1           CTCTGGATTC TCTGTAGATC TAGCGATTTT CAGGTTACAC TTAGCAGGAG CTTCATCAAT CTTAGGAGCT
PPIE2           CTCCGGTTTC TCTGTAGATT TAGCAATCTT TAGCCTTCAC CTAGCGGGTG CCTCTTCCAT TTTAGGTGCC
PPIE3           CTCCGGTTTC TCTGTAGATT TAGCAATCTT TAGCCTTCAC CTAGCGGGTG CCTCTTCCAT TTTAGGTGCC
PPIE4           CTCAGGATTT TCTGTAGATC TTGCAATTTT TAGCCTCCAC TTGGCAGGTG CCTCCTCTAT TCTGGGGGCC
PPIC1           CTCAGGATTT TCTGTAGATC TTGCAATTTT TAGACTCCAC TTGGCAGGTG CCTCCTCTAT TCTGGGGGCC
PPIC2           CTCCGGTTTC TCTGTAGATT TAGCAATCTT CAGCCTTCAC CTAGCGGGTG CCTCTTCCAT TTTAGGTGCC
PPIC3           CTCCGGTTTC TCTGTAGATT TAGCAATCTT TAGCCTTCAC CTAGCGGGTG CCTCTTCCAT TTTAGGTGCC
PPIS1           CTCAGGATTT TCAGTAGACC TTGCAATTTT CAGGCTCCAT CTTGCAGGAG CTTCTTCTAT CTTGGGGGCC
PPIS2           CTCAGGATTT TCAGTAGACC TTGCAATTTT CAGGCTCCAT CTTGCAGGAG CTTCTTCTAT CTTGGGGGCC
PPIS3           CTCAGGATTT TCAGTAGACC TTGCAATTTT CAGGCTCCAT CTTGCAGGAG CTTCTTCTAT CTTGGGGGCC
PPIM2           CTCAGGATTT TCAGTAGACC TTGCAATTTT CAGGCTCCAT CTTGCAGGAG CTTCTTCTAT CTTGGGGGCC
PPI03           CTCAGGATTT TCCGTAGATC TTGCAATCTT TAGCCTTCAT CTAGCAGGTG CCTCCTCCAT TCTGGGGGCC
PPOG1           CTCAGGATTT TCTGTAGATC TTGCAATTTT TAGACTCCAC TTGGCAGGTG CCTCCTCTAT TCTGGGGGCC
PPOG2           CTCAGGATTT TCAGTAGACC TTGCAATTTT CAGGCTCCAT CTTGCAGGAG CTTCTTCTAT CTTGGGGGCC
PPOG3           CTCAGGATTT TCAGTAGACC TTGCAATTTT CAGGCTCCAT CTTGCAGGAG CTTCTTCTAT CTTGGGGGCC
PPOG4           CTCAGGATTT TCAGTAGACC TTGCAATTTT CAGGCTCCAT CTTGCAGGAG CTTCTTCTAT CTTGGGGGCC
PPOG5           CTCAGGATTT TCAGTAGACC TTGCAATTTT CAGGCTCCAT CTTGCAGGAG CTTCTTCTAT CTTGGGGGCC
PPOG6           CTCAGGATTT TCAGTAGACC TTGCAATTTT CAGGCTCCAT CTTGCAGGAG CTTCTTCTAT CTTGGGGGCC
PPNB3           CTCAGGATTT TCCGTAGATC TCGCAATCTT TAGCCTTCAC CTAGCTGGTG CGTCCTCCAT TCTTGGAGCT
PPNB5           CTCAGGATTT TCCGTAGATC TCGCAATCTT TAGCCTTCAC CTAGCTGGTG CGTCCTCCAT TCTTGGAGCT
PPBC1           CTCAGGATTT TCCGTAGATC TCGCAATCTT TAGCCTTCAC CTAGCTGGTG CGTCCTCCAT TCTTGGAGCT
PPBC4           CTCAGGATTT TCCGTAGATC TCGCAATCTT TAGCCTTCAC CTAGCTGGTG CGTCCTCCAT TCTTGGAGCT
PPBA2           CTCAGGATTT TCCGTAGATC TCGCAATCTT TAGCCTTCAC CTAGCTGGTG CGTCCTCCAT TCTTGGAGCT
PPBA4           CTCAGGATTT TCCGTAGATC TCGCAATCTT TAGCCTTCAC CTAGCTGGTG CGTCCTCCAT TCTTGGAGCT
PPBO1           CTCAGGATTT TCCGTAGATC TCGCAATCTT TAGCCTTCAC CTAGCTGGTG CGTCCTCCAT TCTTGGAGCT
PPBO2           CTCAGGATTT TCCGTAGATC TCGCAATCTT TAGCCTTCAC CTAGCTGGTG CGTCCTCCAT TCTTGGAGCT
PPDB21          CTCAGGATTT TCCGTAGATC TTGCAATCTT TAGTCTCCAC TTAGCAGGTG CGTCCTCCAT TCTTGGAGCT
PPDSKat1        CTCAGGATTT TCTGTAGATC TTGCAATTTT TAGACTCCAC TTGGCAGGTG CCTCCTCTAT TCTGGGGGCC
PPDSKat2        CTCAGGATTT TCTGTAGATC TTGCAATTTT TAGACTCCAC TTGGCAGGTG CCTCCTCTAT TCTGGGGGCC
PPDS3           CTCAGGATTT TCAGTaGACC TTGcAATTTT CAgGCTcCAT CTTGCAgGAG CTTcTTcTAT cTTGGGGGCC
PPDS4           CTCCGGTTTC TCTGTAGATT TAGCAATCTT TAGCCTTCAC CTAGCGGGTG CCTCTTCCAT TTTAGGTGCC
PPDS5           CTCCGGTTTC TCTGTAGATT TAGCAATCTT TAGCCTTCAC CTAGCGGGTG CCTCTTCCAT TTTAGGTGCC
PPKW1           CTCAGGATTT TCCGTAGATC TCGCAATCTT TAGCCTTCAC CTAGCTGGTG CGTCCTCTAT TCTTGGAGCT
PPKW2           CTCAGGATTT TCCGTAGATC TCGCAATCTT TAGCCTTCAC CTAGCTGGTG CGTCCTCTAT TCTTGGAGCT
PPKW5           CTCAGGATTT TCCGTAGATC TCGCAATCTT TAGCCTTCAC CTAGCTGGTG CGTCCTCTAT TCTTGGAGCT
PPKW6           CTCAGGATTT TCCGTAGATC TCGCAATCTT TAGCCTTCAC CTAGCTGGTG CGTCCTCCAT TCTTGGAGCT
```

```
                ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                   425        435        445        455        465        475        485
CHA             ATCAATTTCA TCACCACAAT CTTAAATATA CGATCATCCA CAATAAGCTT GGACTCCATC CCACTATTCG
CHC             ATCAACTTTA TCACAACAAT CATAAACATA CGCTCTTCAT CAATAAGGTT GGATTCAATC CCCCTTTTCG
HTDB1           ATTAACTTCA TTACTACCAT TCTcAACATA CGGtCCTCCa CAATAAGACT AgACTCAATT CCCTTATTTG
HTDB2           ATTAACTTCA TTACTaCCAT TCTcAACATA CGGtCCtccA CAATAAgACT AGACTCAATT CCCTTATTTG
PSCR2           ATCAATTTCA TCACCACAAT CTTAAATATA CGATCATCCA CAATAAGCTT GGACTCCATC CCACTATTCG
PSCRA           ATCAATTTCA TCACCACAAT CTTAAATATA CGATCATCCA CAATAAGCTT GGACTCCATC CCACTATTCG
PSCRB           ATCAACTTTA TCACAACAAT CATAAACATA CGCTCTTCAT CAATAAGGTT GGATTCAATC CCCCTTTTCG
PSCRC           ATTAACTTTA TCACTACAAT CATAAATATa CGGTCTTCTT CCATGAGGCT TGATTCAATT CCCCTTTTTG
PYCHA           ATCAACTTTA TCACAACAAT CATAAACATA CGCTCTTCAT CAATAAGGtT GGATTCAATC CCCCTTTTCG
PYCHB           ATCAACTTTA TCACAACAAT CATAAACATA CGCTCTTCAT CAATAAGGTT GGATTCAATC CCCCTTTTCG
PYCHC           ATCAATTTCA TCACCACAAT CTTAAATATA CGATCATCCA CAATAAGCTT GGACTCCATC CCACTATTCG
PPK3            ATTAATTTCA TCACTACAAT TCTAAACATA CGATCTTCCA CAATAAGCCT AGACTTCATC CCATTATTCG
PPK4            ATTAATTTCA TCACTACAAT TCTAAACATA CGATCTTcCA CAATAAGCCT AgACTCCATC CCATTATTCG
PPKA            ATCAATTTCA TCACCACAAT CTTAAATATA CGATCATCCA CAATAAGCTT GGACTCCATC CCACTATTCG
PPKB            ATCAATTTCA TCACCACAAT CTTAAATATA CGATCATCCA CAATAAGCTT GGACTCCATC CCACTATTCG
PPUW1           ATCAACTTCA TCACTACTAT TTTAAATATA CGTTCCTCTA CCATAAGGCT AGATTCAATT CCTTTATTTG
PPUW2           ATCAACTTCA TCACTACTAT TTTAAATATA CGTTCCTCTA CCATAAGGCT AGATTCAATT CCTTTATTTG
PPUW3           ATCAACTTCA TCACTACTAT TTTAAATATA CGTTCCTCTA CCATAAGGCT AGATTCAATT CCTTTATTTG
PPUW4           ATCAACTTCA TCACTACTAT TTTAAATATA CGTTCCTCTA CCATAAGGCT AGATTCAATT CCTTTATTTG
PPUW5           ATCAACTTCA TCACTACTAT TTTAAATATA CGTTCCTCTA CCATAAGGCT AGATTCAATT CCTTTATTTG
PPUH1           ATCAATTTCA TCACCACAAT CTTAAATATA CGATCATCCA CAATAAGCTT GGACTCCATC CCACCATTCG
PPUH2           ATCAATTTCA TCACCACAAT CTTAAATATA CGATCATCCA CAATAAGCTT GGACTCCATC CCACTATTCG
PPUH3           ATTAACTTCA TCACCACCAT TCTAAATATA CGATCCTCTA CTATAAGCCT AGATTCAATC CCCCTATTCG
PPJYMark1       ATCAATTTCA TCACCACAAT CTTAAACATA CGGTCATCCA CAATGAGATT AGATTCCATC CCATTATTCG
PPJYMark2       ATCAATTTCA TCACCACAAT CTTAAACATA CGGTCATCCA CAATGAGATT AGATTCCATC CCATTATTCG
PPJYMark3       ATCAATTTCA TCACCACAAT CTTAAACATA CGGTCATCCA CAATGAGATT AGATTCCATC CCATTATTCG
PPJYMark4       ATCAATTTCA TCACCACAAT CTTAAACATA CGGTCATCCA CAATGAGATT AGATTCCATC CCATTATTCG
PPJY1           ATCAATTTCA TCACCACAAT CTTAAACATA CGGTCATCCA CAATGAGATT AGATTCCATC CCATTATTCG
PPJY2           ATCAATTTCA TCACCACAAT CTTAAACATA CGGTCATCCA CAATGAGATT AGATTCCATC CCATTATTCG
PPJY3           ATCAATTTCA TCACCACAAT CTTAAACATA CGGTCATCCA CAATGAGATT AGATTCCATC CCATTATTCG
PPJY4           ATCAATTTCA TCACCACAAT CTTAAACATA CGGTCATCCA CAATGAGATT AGATTCCATC CCATTATTCG
PPJF1           ATCAATTTCA TCACCACAAT CTTAAACATA CGGTCATCCA CAATGAGATT AGACTCCATC CCATTATTCG
PPJF2           ATCAATTTCA TCACCACAAT CTTAAACATA CGGTCATCCA CAATGAGATT AGACTCCATC CCATTATTCG
PPJF3           ATCAATTTCA TCACCACAAT CTTAAACATA CGGTCATCCA CAATGAGATT AGACTCCATC CCATTATTCG
PPJF4           ATCAATTTCA TCACCACAAT CTTAAACATA CGGTCATCCA CAATGAGATT AGACTCCATC CCATTATTCG
PPJF5           ATCAATTTCA TCACCACAAT CTTAAGCATA CGGTCATCCA CAATGAGATT AGACTCCATC CCATTATTCG
PPIE1           ATTAACTTCA TTACTACCAT TCTCAACATA CGGTCCTCCA CAATAAGACT AGACTCAATT CCCTTATTTG
PPIE2           ATCAACTTCA TCACTACTAT TTTAAATATA CGTTCCTCTA CCATAAGGCT AGATTCAATT CCTTTATTTG
PPIE3           ATCAACTTCA TCACCACAAT CTTAAATATA CGATCATCCA CAATAAGCTT GGACTCCATC CCACTATTCG
PPIE4           ATCAATTTCA TCACCACAAT CTTAAACATA CGGTCATCCA CAATGAGATT AGATTCCATC CCATTATTCG
PPIC1           ATCAATTTCA TCACCACAAT CTTAAACATA CGGTCATCCA CAATGAGATT AGACTCCATC CCATTATTCG
PPIC2           ATCAACTTCA TCACTACTAT TTTAAATATA CGTTCCTCTA CCATAAGGCT AGATTCAATT CCTTTATTTG
PPIC3           ATCAACTTCA TCACTACTAT TTTAAATATA CGTTCCTCTA CCATAAGGCT AGATTCAATT CCTTTATTTG
PPIS1           ATTAATTTCA TCACTACAAT TCTAAACATA CGATCTTCCA CAATAAGCCT AGACTCCATC CCATTATTCG
PPIS2           ATTAATTTCA TCACTACAAT TCTAAACATA CGATCTTCCA CAATAAGCCT AGACTCCATC CCACTATTCG
PPIS3           ATTAATTTCA TCACTACAAT TCTAAACATA CGATCTTCCA CAATAAGCCT AGACTCCATC CCATTATTCG
PPIM2           ATTAATTTCA TCACTACAAT TCTAAACATA CGATCTTCCA CAATAAGCCT AGACTCCATC CCATTATTCG
PPI03           ATCAATTTCA TCACCACAAT CTTAAATATA CGGTCATCCA CAATAAGTTT AGACTCCATC CCACTATTCG
PPOG1           ATCAATTTCA TCACCACAAT CTTAAACATA CGGTCATCCA CAATGAGATT AGACTCCATC CCATTATTCG
PPOG2           ATTAATTTCA TCACTACAAT TCTAAACATA CGATCTTCCA CAATAAGCCT AGACTCCATC CCATTATTCG
PPOG3           ATTAATTTCA TCACTACAAT TCTAAACATA CGATCTTCCA CAATAAGCCT AGACTCCATC CCATTATTCG
PPOG4           ATTAATTTCA TCACTACAAT TCTAAACATA CGATCTTCCA CAATAAGCCT AGACTCCATC CCATTATTCG
PPOG5           ATTAATTTCA TCACTACAAT TCTAAACATA CGATCTTCCA CAATAAGCCT AGACTCCATC CCATTATTCG
PPOG6           ATTAATTTCA TCACTACAAT TCTAAACATA CGATCTTCCA CAATAAGCCT AGACTCCATC CCATTATTCG
PPNB3           ATCAATTTCA TCACCACAAT CTTAAATATA CGATCATCCA CAATAAGCTT GGACTCCATC CCACTATTCG
PPNB5           ATCAATTTCA TCACCACAAC CTTAAATATA CGATCATCCA CAATAAGCTT GGACTCCATC CCACTATTCG
PPBC1           ATCAATTTCA TCACCACAAT CTTAAATATA CGATCATCCA CAATAAGCTT GGACTCCATC CCACTATTCG
PPBC4           ATCAATTTCA TCACCACAAT CTTAAATATA CGATCATCCA CAATAAGCTT GGACTCCATC CCACTATTCG
PPBA2           ATCAATTTCA TCACCACAAT CTTAAATATA CGACCATCCA CAATAAGCTT GGACTCCATC CCACTATTCG
PPBA4           ATCAATTTCA TCACCACAAT CTTAAATATA CGATCATCCA CAATAAGCTT GGACTCCATC CCACTATTCG
PPBO1           ATCAATTTCA TCACCACAAT CTTAAATATA CGATCATCCA CAATAAGCTT GGACTCCATC CCACTATTCG
PPBO2           ATCAATTCCA TCACCACAAT CTTAAATATA CGATCATCCA CAATAAGCTT GGACTCCATC CCACTATTCG
PPDB21          ATCAATTTCA TCACCACAAT CTTAAACATA CGATCATCCA CAATAAGCTT AGACTCCATC CCACTATTCG
PPDSKat1        ATCAATTTCA TCACCACAAT CTTAAACATA CGGTCATCCA CAATGAGATT AGACTCCATC CCATTATTCG
PPDSKat2        ATCAATTTCA TCACCACAAT CTTAAACATA CGGTCATCCA CAATGAGATT AGACTCCATC CCATTATTCG
PPDS3           ATTAATTTCA TCACTACAAT TTTAAACATA CGATCTTCCA CAATAAGCCT AGACTCCATC CCATTATTCG
PPDS4           ATCAACTTCA TCACTACTAT TTTAAATATA CGTTCCTCTA CCATAAGGCT AGATTCAATT CCTTTATTTG
PPDS5           ATCAACTTCA TCACTACTAT TTTAAATATA CGTTCCTCTA CCATAAGGCT AGATTCAATT CCTTTATTTG
PPKW1           ATCAATTTCA TCACCACAAT CTTAAATATA CGATCATCCA CAATAAGCTT GGACTCCATC CCACTATTCG
PPKW2           ATCAATTTCA TCACCACAAT CTTAAATATA CGATCATCCA CAATAAGCTT GGACTCCATC CCACTATTCG
PPKW5           ATCAATTTCA TCACCACAAT CTTAAATATA CGATCATCCA CAATAAGCTT GGACTCCATC CCACTATTCG
PPKW6           ATCAATTTCA TCACCACAAT CTTAAATATA CGATCATCCA CAATAAGCTT GGACTCCATC CCACTATTCG
```

```
                ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                   495        505        515        525        535        545        555
CHA             TTTGATCAGT CCTAATTACC GCAGTATTGT TATTACTGGC TTTACCAGTG CTGGCTGGAG CCATCACAAC
CHC             TCTGATCTGT ACTAATCACA GCCGTCCTTT TACTTCTAGC CCTCCCAGTT TTGGCAGGGG CTATCACAAT
HTDB1           tTTGAtCAGT TCTAATCACC GCAATTCTCC TACTTCTAGC ACTTCCCGTA CTAGCGGGAG CTATCACTAT
HTDB2           tTTGATCAGt TCTAATCACC GCAATTcTcC TACTTcTAGC ACTTCCCGTA CTAGCGGGAG CTATCACTAT
PSCR2           TTTGATCAGT CCTAATTACC GCAGTATTGT TATTACTGGC TTTACCAGTG CTGGCTGGAG CCATCACAAC
PSCRA           TTTGATCAGT CCTAATTACC GCAGTATTGT TATTACTGGC TTTACCAGTG CTGGCTGGAG CCATCACAAC
PSCRB           TCTGATCTGT ACTAATCACA GCCGTCCTTT TACTTCTAGC CCTCCCAGTT TTGGCAGGGG CTATCACAAT
PSCRC           TATGATCTGT TCTAATTaCA GCAATCCTAT TACTCTTAGC TCTTCCAGTA TTGGCAGGGG CCATTACTAT
PYCHA           TCTGATCTGt ACTAATCACA GCCGtCCTTT TACTTCTAGC CCTcCCAGTT TTGGCAGGGG CTATCACAAT
PYCHB           TCTGATCTGT ACTAATCACA GCCGTCCTTT TACTTCTAGC CCTCCCAGTT TTGGCAGGGG CTATCACAAT
PYCHC           TTTGATCAGT CCTAATTACC GCAGTATTGT TATTACTGGC TTTACCAGTG CTGGCTGGAG CCATCACAAC
PPK3            TATGATCCGT TTTAATCACT GCCGTCTTAT TACTACTGGC GCTACCCGTG TTAGCAGGAG CTATTACCAT
PPK4            TATGATCCGT TTTAATCACT GCCGTCTTAT TACTACTGGC GCTACCCGTG TTAGCAGGAG CTATTACCAT
PPKA            TTTGACCAGT CCTAATTACC GCAGTATTGT TATTACTGGC TTTACCAGTG CTGGCTGGAG CCATCACAAC
PPKB            tTTGATCAGT CCTAATTACC GCAGTATTGT TATTACTGGC TTTACCAGTG CTGGCTGGAG CCATCACAAC
PPUW1           TGTGATCAGT TTTGATTACT GCTGTATTGC TTCTTTTAGC GCTACCCGTT TTAGCAGGAG CTATCACCAT
PPUW2           TGTGATCAGT TTTGATTACT GCTGTATTGC TTCTTTTAGC GCTACCCGTT TTAGCAGGAG CTATCACCAT
PPUW3           TGTGATCAGT TTTGATTACT GCTGTATTGC TTCTTTTAGC GCTACCCGTT TTAGCAGGAG CTATCACCAT
PPUW4           TGTGATCAGT TTTGATTACT GCTGTATTGC TTCTTTTAGC GCTACCCGTT TTAGCAGGAG CTATCACCAT
PPUW5           TGTGATCAGT TTTGATTACT GCTGTATTGC TTCTTTTAGC GCTACCCGTT TTAGCAGGAG CTATCACCAT
PPUH1           TTTGATCAGT CCTAATTACC GCAGTATTGT TATTACTGGC TTTACCAGTG CTGGCTGGAG CCATCACAAT
PPUH2           TTTGATCAGT CCTAATTACC GCAGTATTGT TATTACTGGC TTTACCAGTG CTGGCTGGAG CCATCACAAT
PPUH3           TATGATCTGT TCTGATTACC GCAGTACTTC TTCTATTAGC CCTCCCAGTG CTGGCTGGAG CTATCACTAT
PPJYMark1       TTTGATCTGT CTTAATCACT GCCGTTCTAT TACTTCTGGC CCTGCCAGTA TTAGCTGGAG CCATCACGAT
PPJYMark2       TTTGATCTGT CTTAATCACT GCCGTTCTAT TACTTCTGGC CCTGCCAGTA TTAGCTGGAG CCATCACGAT
PPJYMark3       TTTGATCTGT CTTAATCACT GCCGTTCTAT TACTTCTGGC CCTACCAGTA TTAGCTGGAG CCATCACGAT
PPJYMark4       TTTGATCTGT CTTAATCACT GCCGTTCTAT TACTTCTGGC CCTACCAGTA TTAGCTGGAG CCATCACGAT
PPJY1           TTTGATCTGT CTTAATCACT GCCGTTCTAT TACTTCTGGC CCCACCAGTA TTAGCTGGAG CCATCACGAT
PPJY2           TTTGATCTGT cTTAATCACT GCCGTTCTAT TACTTCTGGC CCTGCCAGTA TTAGCTGGAG CCATCACGAT
PPJY3           TTTGATCTGT CTTAATCACT GCCGTTCTAT TACTTCTGGC CCTACCAGTA TTAGCTGGAG CCATCACGAT
PPJY4           TTTGATCTGT CTTAATCACT GCCGTTCTAT TACTTCTGGC CCTACCAGTA TTAGCTGGAG CCATCACGAT
PPJF1           TTTGATCTGT CTTAATCACT GCCGTTCTAT TACTTCTGGC CCTACCAGTA TTAGCTGGAG CCATCACGAT
PPJF2           TTTGATCTGT CTTAACCACT GCCGTTCTAT TACTTCTGGC CCTACCAGTA TTAGCTGGAG CCATCACGAT
PPJF3           TTTGATCTGT CTTAATCACT GCCGTTCTAT TACTTCTGGC CCTACCAGTA TTAGCTGGAG CCATCACGAT
PPJF4           TTTGATCTGT CTTAATCACT GCCGTTCTAT TACTTCTGGC CCTACCAGTA TTAGCTGGAG CCATCACGAT
PPJF5           TTTGATCTGT CTTAATCACT GCCGTTCTAT TACTTCTGGC CCTACCAGTA TTAGCTGGAG CCATCACGAT
PPIE1           TTTGATCAGT TCTAATCACC GCAATTCTCC TACTTCTAGC ACTTCCCGTA CTAGCGGGAG CTATCACTAT
PPIE2           TGTGATCAGT TTTGATTACT GCTGTATTGC TTCTTTTAGC GCTACCCGTT TTAGCAGGAG CTATCACCAT
PPIE3           TTTGATCAGT CCTAATTACC GCAGTATTGT TATTACTGGC TTTACCAGTG CTGGCTGGAG CCATCACAAT
PPIE4           TTTGATCTGT CTTAATCACT GCCGTTCTAT TACTTCTGGC CCTACCAGTA TTAGCTGGAG CCATCACGAT
PPIC1           TTTGATCTGT CTTAATCACT GCCGTTCTAT TACTTCTGGC CCTACCAGTA TTAGCTGGAG CCATCACGAT
PPIC2           TGTGATCAGT TTTGATTACT GCTGTATTGC TTCTTTTAGC GCTACCCGTT TTAGCAGGAG CTATCACCAT
PPIC3           TGTGATCAGT TTTGATTACT GCTGTATTGC TTCTTTTAGC GCTACCCGTT TTAGCAGGAG CTATCACCAT
PPIS1           TATGATCCGT TTTAATCACT GCCGTCTTAT TACTACTGGC GCTACCCGTA TTAGCAGGAG CTATTACCAT
PPIS2           TATGATCCGT TTTAATCACT GCCGTCTTAT TATTACTGGC GCTACCCGTG TTAGCAGGAG CTATTACCAT
PPIS3           TATGATCCGT TTTAATCACT GCCGTCTTAT TACTACTGGC GCTACCCGTA TTAGCAGGAG CTATTACCAT
PPIM2           TATGATCCGT TTTAATCACT GCCGTCTTAT TACTACTGGC GCTACCCGTA TTAGCAGGAG CTATTACCAT
PPI03           TTTGATCAGT CTTAATCACC GCCGTCTTGT TACTACTAGC ATTACCAGTA TTAGCTGGAG CCATCACAAT
PPOG1           TTTGATCTGT CTTAATCACT GCCGTTCTAT TACTTCTGGC CCTACCAGTA TTAGCTGGAG CCATCACGAT
PPOG2           TATGATCCGT TTTAATCACT GCCGTCTTAT TACTACTGGC GCTACCCGTG TTAGCAGGAG CTATTACCAT
PPOG3           TATGATCCGT TTTAATCACT GCCGTCTTAT TACTACTGGC GCTACCCGTG TTAGCAGGAG CTATTACCAT
PPOG4           TATGATCCGT TTTAATCACT GCCGTCTTAT TACTACTGGC GCTACCCGTG TTAGCAGGAG CTATTACCAT
PPOG5           TATGATCCGT TTTAATCACT GCCGTCTTAT TACTACTGGC GCTACCCGTG TTAGCAGGAG CTATTACCAT
PPOG6           TATGATCCGT TTTAATCACT GCCGTCTTAT TACTACTGGC GCTACCCGTG TTAGCAGGAG CTATTACCAT
PPNB3           TTTGATCAGT CCTAATTACC GCAGTATTGT TATTACTGGC TTTACCAGTG CTGGCTGGAG CCATCACAAT
PPNB5           TTTGATCAGT CCTAATTACC GCAGTATTGT TATTACTGGC TTTACCAGTG CTGGCTGGAG CCATCACAAT
PPBC1           TTTGATCAGT CCTAATTACC GCAGTATTGT TATTACTGGC TTTACCAGTG CTGGCTGGAG CCATCACAAT
PPBC4           TTTGATCAGT CCTAATTACC GCAGTATTGT TATTACTGGC TTTACCAGTG CTGGCTGGAG CCATCACAAT
PPBA2           TTTGATCAGT CCTAATTACC GCAGTATTGT TATTACTGGC TTTACCAGTG CTGGCTGGAG CCATCACAAT
PPBA4           TTTGATCAGT CCTAATTACC GCAGTATTGT TATTACTGGC TTTACCAGTG CTGGCTGGAG CCATCACAAT
PPBO1           TTTGATCAGT CCTAATTACC GCAGTATTGT TATTACTGGC TTTACCAGTG CTGGCTGGAG CCATCACAAC
PPBO2           TTTGATCAGT CCTAATTACC GCAGTATTGT TATTACTGGC TTTACCAGTG CTGGCTGGAG CCATCACAAT
PPDB21          TTTGATCAGT TCTAATTACC GCAGTCTTGT TATTACTGGC TTTACCGGTG TTGGCTGGAG CCATTACAAT
PPDSKat1        TTTGATCTGT CTTAATCACT GCCGTTCTAT TACTTCTGGC CCTACCAGTA TTAGCTGGAG CCATCACGAT
PPDSKat2        TTTGATCTGT CTTAATCACT GCCGTTCTAT TACTTCTGGC CCTACCAGTA TTAGCTGGAG CCATCACGAT
PPDS3           TATGATCCGT TTTAATCACT GCCGTTTTAT TACTACTGGC GCTACCCGTA TTAGCAGGAG CTATTACCAT
PPDS4           TGTGATCAGT TTTGATTACT GCTGTATTGC TTCTTTTAGC GCTACCCGTT TTAGCAGGAG CTATCACCAT
PPDS5           TGTGATCAGT TTTGATTACT GCTGTATTGC TTCTTTTAGC GCTACCCGTT TTAGCAGGAG CTATCACCAT
PPKW1           TTTGATCAGT CCTAATTACC GCAGTATTGT TATTACTGGC TTTACCAGTG CTGGCTGGAG CCATCACAAT
PPKW2           TTTGATCAGT CCTAATTACC GCAGTATTGT TATTACTGGC TTTACCAGTG CTGGCTGGAG CCATCACAAT
PPKW5           TTTGATCAGT CCTAATTACC GCAGTATTGT TATTACTGGC TTTACCAGTG CTGGCTGGAG CCATCACAAT
PPKW6           TTTGATCAGT CCTAATTACC GCAGTATTGT TATTACTGGC TTTACCAGTG CTGGCTGGAG CCATCACAAT
```

```
            ....|....| ....|....| ....|....| ....|....|
               565        575        585        595
CHA         ACTCTTAACA GACCGAAACT TCAATACAAC ATTCTTCGAC
CHC         GCTTTTAACT GACCGAAACT TTAACACATC ATTTTTTGAT
HTDB1       ACTCTTAACG GATCGAAACT TTAACACTAC ATTTTTTGAT
HTDB2       ACTCTTAACG GATCGAAACT TTAACACTAC ATTTTTTGAT
PSCR2       ACTCTTAACA GACCGAAACT TCAATACAAC ATTCTTCGAC
PSCRA       ACTCTTAACA GACCGAAACT TCAATACAAC ATTCTTCGAC
PSCRB       GCTTTTAACT GACCGAAACT TTAACACATC ATTTTTTGAT
PSCRC       ACTTTTAACA GATCGTAATT TTAATACCTC ATTCTTTGAT
PYCHA       GCTTTTAACT GaCCGAAACT TTAACACATC ATTTTTTGAT
PYCHB       GCTTTTAACT GACCGAAACT TTAACACATC ATTTTTTGAT
PYCHC       ACTCTTAACA GACCGAAACT TCAATACAAC ATTcTTCGAC
PPK3        ACTTTTGACA GACCGTAATT TCAACACAAC GTTTTTTGAT
PPK4        ACTTTTGACA GACCGTAATT TCAACACAAC GTTTTTTGAT
PPKA        ACTcTTAACA GACCGAAACT TCAATACAAC ATTCTTCGAC
PPKB        ACTcTTAACA GACCGaAACT TCAATACAAC ATTcTTCGAC
PPUW1       ACTTCTAACG GACCGAAACT TCAATACAAC ATTCTTTGAC
PPUW2       ACTTCTAACG GACCGAAACT TCAATACAAC ATTCTTTGAC
PPUW3       ACTTCTAACG GACCGAAACT TCAATACAAC ATTCTTTGAC
PPUW4       ACTTCTAACG GACCGAAACT TCAATACAAC ATTCTTTGAC
PPUW5       ACTTCTAACG GACCGAAACT TCAATACAAC ATTCTTTGAC
PPUH1       ACTCTTAACA GACCGAAACT TCAATACAAC ATTCTTCGAC
PPUH2       ACTCTTAACA GACCGAAACT TCAATACAAC ATTCTTCGAC
PPUH3       ACTTTTAACG GACCGAAACT TCAACACAAC ATTCTTTGAT
PPJYMark1   ACTCTTAACA GACCGGAATT TTAATACAAC ATTTTTTGAC
PPJYMark2   ACTCTTAACA GACCGGAATT TTAATACAAC ATTTTTTGAC
PPJYMark3   ACTCTTAACA GACCGAAATT TTAATACAAC ATTTTTTGAC
PPJYMark4   ACTCTTAACA GACCGAAATT TTAATACAAC ATTTTTTGAC
PPJY1       ACTCTTAACA GACCGAAATT TTAATACAAC ATTTTTTGAC
PPJY2       ACTCTTAACA GACCGGAATT TTAATACAAC ATTTTTTGAC
PPJY3       ACTCTTAACA GACCGAAATT TTAATACAAC ATTTTTTGAC
PPJY4       ACTCTTAACA GACCGAAATT TTAATACAAC ATTTTTTGAC
PPJF1       ACTCTTAACA GACCGAAATT TTAATACAAC ATTTTTTGAC
PPJF2       ACTCTTAACA GACCGAAATT TTAATACAAC ATTTTTTGAC
PPJF3       ACTCTTAACA GACCGAAATT TTAATACAAC ATTTTTTGAC
PPJF4       ACTCTTAACA GACCGAAATT TTAATACAAC ATTTTTTGAC
PPJF5       ACTCTTAACA GACCGAAATT TTAATACAAC ATTTTTTGAC
PPIE1       ACTCTTAACG GATCGAAACT TTAACACTAC ATTTTTTGAT
PPIE2       ACTTCTAACG GACCGAAACT TCAATACAAC ATTCTTTGAC
PPIE3       ACTCTTAACA GACCGAAGCT TCAATACAAC ATTCTTCGAC
PPIE4       ACTCTTAACA GACCGAAATT TTAATACAAC ATTTTTTGAC
PPIC1       ACTCTTAACA GACCGAAATT TTAATACAAC ATTTTTTGAC
PPIC2       ACTTCTAACG GACCGAAACT TCAATACAAC ATTCTTTGAC
PPIC3       ACTTCTAACG GACCGAAACT TCAATACAAC ATTCTTTGAC
PPIS1       ACTTTTGACA GACCGTAATT TCAACACAAC GTTTTTTGAT
PPIS2       ACTTTTGACA GACCGTAATT TCAACACAAC GTTTTTTGAT
PPIS3       ACTTTTGACA GACCGTAATT TCAACACAAC GTTTTTTGAT
PPIM2       ACTTTTGACA GACCGTAATT TCAACACAAC GTTTTTTGAT
PPI03       ACTCCTAACG GACCGAAACT TCAATACAAC ATTTTTTGAT
PPOG1       ACTCTTAACA GACCGAAATT TTAATACAAC ATTTTTTGAC
PPOG2       ACTTTTGACA GACCGTAATT TCAACACAAC GTTTTTTGAT
PPOG3       ACTTTTGACA GACCGTAATT TCAACACAAC GTTTTTTGAT
PPOG4       ACTTTTGACA GACCGTAATT TCAACACAAC GTTTTTTGAT
PPOG5       ACTTTTGACA GACCGTAATT TCAACACAAC GTTTTTTGAT
PPOG6       ACTTTTGACA GACCGTAATT TCAACACAAC GTTTTTTGAT
PPNB3       ACTCTTAACA GACCGAAACT TCAATACAAC ATTCTTCGAC
PPNB5       ACTCTTAACA GACCGAAACT TCAATACAAC ATTCTTCGAC
PPBC1       ACTCTTAACA GACCGAAACT TCAATACAAC ATTCTTCGAC
PPBC4       ACTCTTAACA GACCGAAACT TCAATACAAC ATTCTTCGAC
PPBA2       ACTCTTAACA GACCGAAACT TCAATACAAC ATTCTTCGAC
PPBA4       ACTCTTAACA GACCGAAACT TCAATACAAC ATTCTTCGAC
PPBO1       ACTCTTAACA GACCGAAACT TCAATACAAC ATTCTTCGAC
PPBO2       ACTCTTAACA GACCGAAACT TCAATACAAC ATTCTTCGAC
PPDB21      ACTCTTAACA GACCGAAACT TCAATACAAC ATTCTTCGAC
PPDSKat1    ACTCTTAACA GACCGAAATT TTAATACAAC ATTTTTTGAC
PPDSKat2    ACTCTTAACA GACCGAAATT TTAATACAAC ATTTTTTGAC
PPDS3       ACTTTTGACA GACCGTAATT TCAACACAAC GTTTTTTGAT
PPDS4       ACTTCTAACG GACCGAAGCT TCAATACAAC ATTCTTTGAC
PPDS5       ACTTCTAACG GACCGAAACT TCAATACAAC ATTCTTTGAC
PPKW1       ACTATTAACA GACCGAAACT TCAATACAAC ATTCTTCGAC
PPKW2       ACTATTAACA GACCGAAACT TCAATACAAC ATTCTTCGAC
PPKW5       ACTATTAACA GACCGAAACT TCAATACAAC ATTCTTCGAC
PPKW6       ACTCTTAACA GACCGAAACT TCAATACAAC ATTCTTCGAC
```

Nucleotide sequence alignment of the 23 *Mucronothrus nasalis* specimens

```
                ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                    5         15         25         35         45         55         65
MNCT1           TTTGGAACTT GAGCGGGATT GATAGGTTCC TCTTTAAGGG CGCTGATCCG GTTAGAACTA GGGCAGCCAG
MNCT2           TTTGGTACAT GAGCTGGACT AATAGGCTCC TCCCTAAGTG CTTTGATTCG TCTTGAATTA GGACAGCCCG
MNCT3           TTTGgAACCT GGGcTGGATT GATAGGATCT TCTTTAAGTG CTTTGATTCG TCTTGAGTTA GGTCAACCCG
MNCT4           TTTGGAACCT GGGCTGGATT GATAGGATCT CCTTTAAGTG CTTTGATTCG TCTTGAGTTA GGTCAACCCG
MNCT5           TTTGGAACCT GGGCTGGATT GATAGGATCT TCTTTAAGTG CTTTGATTCG TCTTGAGTTA GGTCAACCCG
MNCT6           TTTGGAACCT GGGCTGGATT GATAGGATCT TCTTTAAGTG CTTTGATTCG TCTTGAGTTA GGTCAACCCG
MNNF1.1         TTCGGTACTT GAGCGGGACT AATGGGCTcT TCACTGAGAG CCCTGATTCG TTTAGAATTA GGACAACCTG
MNNF1.3         TTCGGTACTT GAGCGGGACT AATGGGCTCT TCACTGAGAG CCCTGATTCG TTTAGAATTA GGACAACCTG
MNNF1.6         TTCGGTACTT GAGCGGGACT AATGGGCTCT TCACTGAGAG CCCTGATTCG TTTAGAATTA GGACAACCTG
MNNF1.7_        TTCGGTACTT GAGCGGGACT AATGGGcTCT TCACTGAGAG CCCTGATTCG TTTAGAATTA GGACAACCTG
MNNF2.9         TTCGGTACTT GAGCGGGACT AATGGGCTCT TCACTGAGAG CCCTGATTCG TTTAGAATTA GGACAACCTG
MNNF3.17_       TTCGGTACTT GAGCGGGACT AATGGGcTCT TCACTgAGAG CCCTGATTCG TTTAGAATTA GGACAACCTG
MNCB1.1         TTTGGAACCT GGGCTGGATT GATAGGATCT TCTTTAAGTG CTTTGATTCG TCTTGAGTTA GGTCAACCCG
MNCB3.7         TTCGGTACTT GAGCGGGACT AATGGGCTCT TCACTGAGAG CCCTGATTCG TTTAGAATTA GGACAACCTG
MNCB4.9         TTTGGAACCT GGGCTGGATT GATAGGATCT TCTTTAAGTG CTTTGATTCG TCTTGAGTTA GGTCAACCCG
MNCB6.17        TTCGGTGCTT GGGCCGGCAT ACTCGGCTCC TCCTTAAGAG CCATCATTCG ATTAGAATTG GGTCAACCAG
MNCH2.3_        TTCGGTACTT GAGCGGGACT AATGGGCTCT TCACTGAGAG CCCTGATTCG TTTAGAATTA GGACAACCTG
MNCH3.8         TTCGGTACCT GGGATGGCCT AATGGGATCT TCTCTAAGAG CATTAATTCG ATTGGAATTA GGACAACCAG
MNCH4.9         TTCGGTACTT GAGCGGGACT AATGGGCTCT TCACTGAGAG CCCTGATTCG TTTAGAATTA GGACAACCTG
MNCH5.17        TTTGGAACCT GGGCTGGATT GATAGGATCT TCTTTAAGTG CTTTGATTCG TCTTGAGTTA GGTCAACCCG
MNNG1           TTTGGAACCT GGGCTGGATT GATAGGATCT TCTTTAAGTG CTTTGATTCG TCTTGAGTTA GGTCAACCCG
MNNG2.1         TTTGGAACCT GGGCTGGATT GATAGGATCT TCTTTAAGTG CTTTGATTCG TCTTGAGTTA GGTCAACCCG
MNUC2           TTCGGTACTT GAGCGGAACT AATGGGCTCT TCACTGAGAG CCCTGATTCG TTTAGAATTA GGACAACCTG

                ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                   75         85         95        105        115        125        135
MNCT1           GATCATTACT AGACAATGAC CAAATTTATA ACACAATTGT TACGGCTCAC GCTTTCGTTA TGATTTTCTT
MNCT2           GTTCACTTCT AGATAACGAC CAAATCTACA ACACTATTGT CACAGCTCAC GCCTTCGTAA TGATTTTTTT
MNCT3           GTTCACTACT AGACAACGAC CAAATTTATA ACACCATTGT TACAGCCCAC GCCTTTGTAA TAATTTTTTT
MNCT4           GTTCACTACT AGACAACGAC CAAATTTATA ACACCATTGT TACAGCCCAC GCCTTTGtAA TAATTTTTTT
MNCT5           GTTCACTACT AGACAACGAC CAAATTTATA ACACCATTGT TACAGCCCAC GCCTTTGTAA TAATTTTTTT
MNCT6           GTTCACTACT AGACAACGAC CAAATTTATA ACACCATTGT TACAGCCCAC GCCTTTGTAA TAATTTTTTT
MNNF1.1         GTTCCCTCCT TGAAAATGAT CAAATCTATA ACACTATTGT CACTGCCCAT GCCTTCGTAA TAATTTTTTT
MNNF1.3         GTTCCCTCCT TGAAAATGAT CAAATCTATA ACACTATTGT CACTGCCCAT GCCTTCGTAA TAATTTTTTT
MNNF1.6         GTTCCCTCCT TGAAAATGAT CAAATCTATA ACACTATTGT CACTGCCCAT GCCTTCGTAA TAATTTTTTT
MNNF1.7_        GTTCCcTCCT TGAAAATGAT CAAATCTATA ACACTATTGT CACTGCCCAT GCCTTCGTAA TAATTTTTTT
MNNF2.9         GTTCCCTCCT TGAAAATGAT CAAATCTATA ACACTATTGT CACTGCCCAT GCCTTCGTAA TAATTTTTTT
MNNF3.17_       GTTCCCTCCT TGAAAATGAT CAAATCTATA ACACTATTGT CACTGCCCAT GCCTTCGTAA TAATTTTTTT
MNCB1.1         GTTCACTACT AGACAACGAC CAAATTTATA ACACCATTGT TACAGCCCAC GCCTTTGTAA TAATTTTTTT
MNCB3.7         GTTCCCTCCT TGAAAATGAT CAAATCTATA ACACTATTGT CACTGCCCAT GCCTTCGTAA TAATTTTTTT
MNCB4.9         GTTCACTACT AGACAACGAC CAAATTTATA ACACCATTGT TACAGCCCAC GCCTTTGTAA TAATTTTTTT
MNCB6.17        GCTCCCTACT AGAGAATGAC CAAATTTACA ACACCATCGT CACCGCCCAC GCATTTGTCA TGATTTTCTT
MNCH2.3_        GTTCCCTCCT TGAAAATGAT CAAATCTATA ACACTATTGT CACTGCCCAT GCCTTCGTAA TAATTTTTTT
MNCH3.8         GCTCTCTTCT CGAAAATGAC CAAATTTACA ACACAATTGT CACTGCACAC GCCTTCGTAA TAATCTTCTT
MNCH4.9         GTTCCCTCCT TGAAAATGAT CAAATCTATA ACACTATTGT CACTGCCCAT GCCTTCGTAA TAATTTTTTT
MNCH5.17        GTTCACTACT AGACAACGGC CAAATTTATA ACACCATTGT TACAGCCCAC GCCTTTGTAA TAATTTTTTT
MNNG1           GTTCACTACT AGACAACGAC CAAATTTATA ACACCATTGT TACAGCCCAC GCCTTTGTAA TAATTTTTTT
MNNG2.1         GTTCACTACT AGACAACGAC CAAATTTATA ACACCATTGT TACAGCCCAC GCCTTTGTAA TAATTTTTTT
MNUC2           GTTCCCTCCT TGAAAATGAT CAAATCTATA ACACTATTGT CACTGCCCAT GCCTTCGTAA TAATTTTTTT

                ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                  145        155        165        175        185        195        205
MNCT1           TATAGTCATG CCTGTTATAA TTGGAGGGTT CGGAAACTGA CTGGTACCTC TGATGATCGG AGCTCAAGAT
MNCT2           CATAGTAATA CCAGTAATAA TTGGGGGGTT TGGAAACTGG CTAGTACCGC TGATAATTGG GGCACAAGAT
MNCT3           CATAGTTATG CCTGGAATAA TTGGTGGGTT TGGAAATTGA TTAGTTCCGT TAATAATTGG AGCCCAAGAT
MNCT4           CATAGTTATG CCTGGAATAA TTGGTGGGTT TGGAAATTGA TTAGTTCCGT TAATAATTGg AGCCCAAGAT
MNCT5           CATAGTTATG CCTGGAATAA TTGGtGGGTT TGGAAATTGA TTAGTTCCGT TAATAATTGG AGCCCAAGAT
MNCT6           CATAGTTATG CCTGGAATAA TTGGTGGGTT TGGAAATTGA TTAGTTCCGT TAATAATTGG AGCCCAAGAT
MNNF1.1         TATAGTAATG CCTGTTATAA TCGGAGGTTT CGGGAATTGA CTAGTTCCAA TAATAATTGG TGCTCAAGAT
MNNF1.3         TATAGTAATG CCTGTTATAA TCGGAGGTTT CGGGAATTGA CTAGTTCCAA TAATAATTGG TGCTCAAGAT
MNNF1.6         TATAGTAATG CCTGTTATAA TCGGAGGTTT CGGGAATTGA CTAGTTCCAA TAATAATTGG TGCTCAAGAT
MNNF1.7_        TATAGTAGTG CCTGTTATAA TCGGAGGTTT CGGGAATTGA CTAGTTCCAA TAATAATTGG TGCTCAAGAT
MNNF2.9         TATAGTAATG CCTGTTATAA TCGGAGGTTT CGGGAATTGA CTAGTTCCAA TAATAATTGG TGCTCAAGAT
MNNF3.17_       TATAGTAATG CCTGTTATAA TCGGAGGTTT CGGGAATTGA CTAGTTCCAA TAATAATTGG TGCTCAAGAT
MNCB1.1         CATAGTTATG CCTGGAATAA TTGGTGGGTT TGGAAATTGA TTAGTTCCGT TAATAATTGG AGCCCAAGAT
MNCB3.7         TATAGTAATG CCTGTTATAA TCGGAGGTTT CGGGAATTGA CTAGTTCCAA TAATAATTGG TGCTCAAGAT
MNCB4.9         CATAGTTATG CCTGGAATAA TTGGTGGGTT TGGAAATTGA TTAGTTCCGT TAATAATTGG AGCCCAAGAT
MNCB6.17        TATAGTAATG CCTGTAATAA TTGGAGGATT CGGGAATTGA CTGGTTCCAT TAATAATCGG AGCACAAGAC
MNCH2.3_        TATAGTAATG CCTGTTATAA TCGGAGGTTT CGGGAATTGA CTAGTTCCAA TAATAATTGG TGCTCAAGAT
MNCH3.8         TATAGTAATG CCTGTAATAA TTGGAGGGTT TGGAAACTGG CTAGTACCCA TGATAATCGG AGCCCAAGAC
MNCH4.9         TATAGTAATG CCTGTTATAA TCGGAGGTTT CGGGAATTGA CTAGTTCCAA TAATAATTGG TGCTCAAGAT
MNCH5.17        CATAGTTATG CCCGGAATAA TTGGTGGGTT TGGAAATTGA TTAGTTCCGT TAATAATTGG AGCCCAAGAT
MNNG1           CATAGTTATG CCTGGAATAA TTGGtGGGTT TGGAAATTGA TTAGTTCCGT TAATAATTGG AGCCCAAGAT
MNNG2.1         CATAGTTATG CCTGGAATAA TTGGTGGGTT TGGAAATTGA TTAGTTCCGT TAATAATTGG AGCCCAAGAT
MNUC2           TATAGTAATG CCTGTTATAA TCGGAGGTTT CGGGAATTGA CTAGTTCCAA TAATAATTGG TGCTCAAGAT
```

```
                ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                   215        225        235        245        255        265        275
MNCT1           ATGGCTTTTC CCCGTATGAA CAACATGAGA TTCTGGCTTC TCCCACCATC CTTGTGTCTT CTTTCAGCCT
MNCT2           ATGGCCCTCC CCCGCATAAA TAATATAAGA TTCTGACTCC TTCCTCCTTC TCTATGTCTT CTTTCCGCAT
MNCT3           ATAGCTTTCC CTCGTATAAA TAATATGAGA TTCTGACTTC TTCCCCCATC ACTGTGTCTC CTTTCAGCAT
MNCT4           ATAGCTTTCC CTCGTATAAA TAATATGAGA TTCTGACTTC TTCCCCCATC ACTGTGTCTC CTTTCAGCAT
MNCT5           ATAGCTTTCC CTCGTATAAA TAATATGAGA TTCTGACTTC TTCCCCCATC ACTGTGTcTC CTTTCAGCAT
MNCT6           ATAGCTTTCC CTCGTATAAA TAATATGAGA TTCTGACTTC TTCCCCCATC ACTGTGTCTC CTTTCAGCAT
MNNF1.1         ATGGCATTCC CTCGAATAAA TAACATGAGA TTTTGACTAC TGCCTCCCTC TTTAACACTC CTATCAGCCT
MNNF1.3         ATGGCATTCC CTCGAATAAA TAACATGAGA TTTTGACTAC TGCCTCCCTC TTTAACACTC CTATCAGCCT
MNNF1.6         ATGGCATTCC CTCGAATAAA TAACATGAGA TTTTGACTAC TGCCTCCCTC TTTAACACTC CTATCAGCCT
MNNF1.7_        ATGGCATTCC CTCGAATAAA TAACATGAGA TTTTGACTAC TGCCTCCCTC TTTAACACTC CTATCAGCCT
MNNF2.9         ATGGCATTCC CTCGAATAAA TAACATGAGA TTTTGACTAC TGCCTCCCTC TTTAACACTC CTATCAGCCT
MNNF3.17_       ATGGCATTCC CTCGAATAAA TAACATGAGA TTTTGACTAC TGCCTCCCTC TTTAACACTC CTATCAGCCT
MNCB1.1         ATAGCTTTCC CTCGTATAAA TAATATGAGA TTCTGACTTC TTCCCCCATC ACTGTGTCTC CTTTCAGCAT
MNCB3.7         ATGGCATTCC CTCGAATAAA TAACATGAGA TTTTGACTAC TGCCTCCCTC TTTAACACTC CTATCAGCCT
MNCB4.9         ATAGCTTTCC CTCGTATAAA TAATATGAGA TTCTGACTTC TTCCCCCATC ACTGTGTCTC CTTTCAGCAT
MNCB6.17        ATAGCCTTTC CACGTATAAA CAATATAAGG TTTTGACTTC TACCGCCTTC TTTAACCTTC CTAACGGCAT
MNCH2.3_        ATGGCATTCC CTCGAATAAA TAACATGAGA TTTTGACTAC TGCCTCCCTC TTTAACACTC CTATCAGCCT
MNCH3.8         ATGGCATTCC CACGAATGAA TAACATGAGA TTCTGATTAC TACCACCCTC TCTCACTCTT TTGTCAGCAT
MNCH4.9         ATGGCATTCC CTCGAATAAA TAACATGAGA TTTTGACTAC TGCCTCCCTC TTTAACACTC CTATCAGCCT
MNCH5.17        ATAGCTTTCC CTCGTATAAA TAATATGAGA TTCTGACTTC TTCCCCCATC ACTGTGTCTC CTTTCAgCAT
MNNG1           ATAGCTTTCC CTCGTATAAA TAATATGAGA TTCTGACTTC TTCCCCCATC ACTGTGTCTC CTTTCAGCAT
MNNG2.1         ATAGCTTTCC CTCGTATAAA TAATATGAGA TTCTGACTTC TTCCCCCATC ACTGTGTCTC CTTTCAGCAT
MNUC2           ATGGCATTCC CTCGAATAAA TAACATGAGA TTTTGACTAC TGCCTCCCTC TTTAACACTC CTATCAGCCT

                ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                   285        295        305        315        325        335        345
MNCT1           CTGCTTTCGC AGGGTTAGGT GTAGGGACAG GATGGACAGT ATACCCCCCA CTCGCTGGGA ATTTATTTCA
MNCT2           CTGCTTTCGC CGGATTAGGT GTAGGGACCG GGTGAACCGT GTACCCCCCA CTAGCTGGAA ATCTTTTCCA
MNCT3           CAGCCTTTGC CGGGCTAGGA GTGGGAACTG GGTGAACTGT CTACCCCCCT CTAGCCGGGA ACCTCTTCCA
MNCT4           CAGCCCTTGC CGGGCTAGGA GTGGGAACTG GGTGAACTGT CTACCCCCCT CTAGCCGGGA ACCTCTTCCA
MNCT5           CAGCCTTCGC CGGGCTAGGA GTGGGAACTG GGTGAACTGT CTACCCCCCT CTAgCCGGGA ACCTCTTcCA
MNCT6           CAGCCTTTGC CGGGCTAGGA GTGGGAACTG GGTGAACTGT CTACCCCCCT CTAGCCGGGA ACCTCTTCCA
MNNF1.1         CAGCTTTTGC AGGAATAGGG GCTGGAACAG GATGAACGGT ATACCCGCCA TTGGCTGGAA ATCTCTTCCA
MNNF1.3         CAGCTTTTGC AGGAATAGGG GCTGGAACAG GATGAACGGT ATACCCGCCA TTGGCTGGAA ATCTCTTCCA
MNNF1.6         CAGCTTTTGC AGGAATAGGG GCTGGAACAG GATGAACGGT ATACCCGCCA TTGGCTGGAA ATCTCTTCCA
MNNF1.7_        CAGCTTTTGC AGGAATAGGG GCTGGAACAG GATGAACGGT ATACCCGCCA TTGGCTGGAA ATCTCTTCCA
MNNF2.9         CAGCTTTTGC AGGAATAGGG GCTGGAACAG GATGAACGGT ATACCCGCCA TTGGCTGGAA ATCTCTTCCA
MNNF3.17_       CAGCTTTTGC AGGAATAGGG GCTGGAACAG GATGAACGGT ATACCCGCCA TTGGCTGGAA ATCTCTTCCA
MNCB1.1         CAGCCTTTGC CGGGCTAGGA GTGGGAACTG GGTGAGCTGT CTACCCCCCT CTAGCCGGGA ACCTCTTCCA
MNCB3.7         CAGCTTTTGC AGGAATAGGG GCTGGAACAG GATGAACGGT ATACCCGCCA TTGGCTGGAA ATCTCTTCCA
MNCB4.9         CAGCCTTTGC CGGGCTAGGA GTGGGAACTG GGTGAACTGT CTACCCCCCT CTAGCCGGGA ACCTCTTCCA
MNCB6.17        CTGGGTTTGC TGGAATAGGA GCAGGAACGG GATGAACAGT ATATCCACCA CTAGCAGGGA ACTTGTTTCA
MNCH2.3_        CAGCTTTTGC AGGAATAGGG GCTGGAACAG GATGAACGGT ATACCCGCCA TTGGCTGGAA ATCTCTTCCA
MNCH3.8         CTGCCTTTGC AGGAATAGGA GCAGGAACAG GATGAACAGT ATATCCTCCC TTAGCCGGAA ATTTATTTCA
MNCH4.9         CAGCTTTTGC AGGAATAGGG GCTGGAACAG GATGAACGGT ATACCCGCCA TTGGCTGGAA ATCTCTTCCA
MNCH5.17        CAGCCTTTGC CGGGCTAGGA GTGGGAACTG GGTGAACTGt CTACCCCCCT CTAgCCGGGA ACCTCTTCCA
MNNG1           CAGCCTTTGC CGGGCTAGGA GTGGGAACTG GGTGAACTGT CTACCCCCCT CTAGCCGGGA ACCTCTTCCA
MNNG2.1         CAGCCTTTGC CGGGCTAGGA GTGGGAACTG GGTGAACTGT CTACCCCCCT CTAGCCGGGA ACCTCTTCCA
MNUC2           CAGCTTTTGC AGGAATAGGG GCTGGAACAG GATGAACGGT ATACCCGCCA TTGGCTGGAA ATCTCTTCCA

                ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                   355        365        375        385        395        405        415
MNCT1           CTCCGGTTTC TCTGTAGATT TAGCAATCTT TAGCCTTCAC CTAGCGGGTG CCTCTTCCAT TTTAGGTGCC
MNCT2           CTCAGGATTT TCAGTAGACC TTGCAATTTT CAGGCTCCAT CTTGCAGGAG CTTCTTCTAT CTTGGGGGCC
MNCT3           CTCAGGATTT TCCGTAGATC TCGCAATCTT TAGCCTTCAC CTAGCTGGTG CGTCCTCCAT TCTTGGAGCT
MNCT4           CTCAGGATTT TCCGTAGATC TCGCAATCTT TAGCCTTCAC CTAGcTGGTG CGTCCTCCAT TCTTGGAGCT
MNCT5           CTCAGGATTT TCCGTAGATC TCGCAATCTT TAGCCTTCAC CTAGCTGGtG CGTcCTCCAT TCTTGGAGCT
MNCT6           CTCAGGATTT TCCGTAGATC TCGCAATCTT TAGCCTTCAC CTAGCTGGTG CGTCCTCCAT TCTTGGAGCT
MNNF1.1         CTCTAGAATT TCAGTAGATT TGGCAATTTT CAGATTACAC CTAGCTGGTG CCTCTTCCAT TTTAGGGGCA
MNNF1.3         CTCTAGAATT TCAGTAGATT TGGCAATTTT CAGATTACAC CTAGCTGGTG CCTCTTCCAT TTTAGGGGCA
MNNF1.6         CTCTAGAATT TCAGTAGATT TGGCAATTTT CAGATTACAC CTAGCTGGTG CCTCTTCCAT TTTAGGGGCA
MNNF1.7_        CTCTAGAATT TCAGTAGATT TGGCAATTTT CAGATTACAC CTAGCTGGTG CCTCTTCCAT TTTAGGGGCA
MNNF2.9         CTCTAGAATT TCAGTAGATT TGGCAATTTT CAGATTACAC CTAGCTGGTG CCTCTTCCAT TTTAGGGGCA
MNNF3.17_       CTCTAGAATT TCAGTAGATT TGGCAATTTT CAGATTACAC CTAGCTGGTG CCTCTTCCAT TTTAGGGGCA
MNCB1.1         CTCAGGATTT TCCGTAGATC TCGCAATCTT TAGCCTTCAC CTAGCTGGTG CGTCCTCCAT TCTTGGAGCT
MNCB3.7         CTCTAGAATT TCAGTAGATT TGGCAATTTT CAGATTACAC CTAGCTGGTG CCTCTTCCAT TTTAGGGGCA
MNCB4.9         CTCAGGATTT TCCGTAGATC TCGCAATCTT TAGCCTTCAC CTAGCTGGTG CGTCCTCCAT TCTTGGAGCT
MNCB6.17        TTCAGGGATC TCAGTAGACC TAGCAATCTT CAGGTTACAC TTAGCAGGTG CATCCTCCAT TCTGGGAGCA
MNCH2.3_        CTCTAGAATT TCAGTAGATT TGGCAATTTT CAGATTACAC CTAGCTGGTG CCTCTTCCAT TTTAGGGGCA
MNCH3.8         TTCCAGAATC TCAGTAGACC TAGCTATTTT TAGACTACAT TTAGCTGGGG CCTCCTCCAT CTTAGGGGCA
MNCH4.9         CTCTAGAATT TCAGTAGATT TGGCAATTTT CAGATTACAC CTAGCTGGTG CCTCTTCCAT TTTAGGGGCA
MNCH5.17        CTCAGGATTT TCCGTAGATC TCGCAATCTT TAGCCTTCAC CTAGCTGGTG CGTCCTCCAT TCTTGGAGCT
MNNG1           CTCAGGATTT TCCGTAGATC TCGCAATCTT TAGCCTTCAC CTAGCTGGTG CGTCCTCCAT TCTTGGAGCT
MNNG2.1         CTCAGGATTT TCCGTAGATC TCGCAATCTT TAGCCTTCAC CTAGCTGGTG CGTCCTCCAT TCTTGGAGCT
MNUC2           CTCTAGAATT TCAGTAGATT TGGCAATTTT CAGATTACAC CTAGCTGGTG CCTCTTCCAT TTTAGGGGCA
```

```
                ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                   425        435        445        455        465        475        485
MNCT1           ATCAACTTCA TCACTACTAT TTTAAATATA CGTTCCTCTA CCATAAGGCT AGATTCAATT CCTTTATTTG
MNCT2           ATTAATTTCA TCACTACAAT TCTAAACATA CGATCTTCCA CAATAAGCCT AGACTCCATC CCATTATTCG
MNCT3           ATCAATTTCA TCACCACAAT CTTAAATATA CGATCATCCA CAATAAGCTT GGACTCCATC CCACTATTCG
MNCT4           ATCAATTTCA TCACCACAAT CTTAAATATA CGATCATCCA CAATAAGCTT GGACTCCATC CCACTATTCG
MNCT5           ATCAATTTCA TCACCACAAT CTTAAATATA CGATCATCCA CAATAAGCTT GGACTCCATC CCACTATTCG
MNCT6           ATCAATTTCA TCACCACAAT CTTAAATATA CGATCATCCA CAATAAGCTT GGACTCCATC CCACTATTCG
MNNF1.1         ATTAACTTTA TCACTACAAT CATAAATATA CGGTCTTcTT CCATGAGGCT TGATTCAATT CCCCTTTTTG
MNNF1.3         ATTAACTTTA TCACTACAAT CATAAATATA CGGtCTTCTT CCATGAGGCT TGATTCAATT CCCCTTTTTG
MNNF1.6         ATTAACTTTA TCACTACAAT CATAAATATA CGGTCTTCTT CCATGAGGCT TGATTCAATT CCCCTTTTTG
MNNF1.7_        ATTAACTTTA TCACTACAAT CATAAATATA CGGTCTTCTT CCATGAGGCT TGATTCAATT CCCCTTTTTG
MNNF2.9         ATTAACTTTA TCACTACAAT CATAAATATA CGGTCTTCTT CCATGAGGCT TGATTCAATT CCCCTTTTTG
MNNF3.17_       ATTAACTTTA TCACTACAAT CATAAATATA CGGTCTTCTT CCATGAGGCT TGATTCAATT CCCCTTTTTG
MNCB1.1         ATCAATTTCA TCACCACAAT CTTAAATATA CGATCATCCA CAATAAGCTT GGACTCCATC CCACTATTCG
MNCB3.7         ATTAACTTTA TCACTACAAT CATAAATATA CGGTCTTCTT CCATGAGGCT TGATTCAATT CCCCTTTTTG
MNCB4.9         ATCAATTTCA TCACCACAAT CTTAAATATA CGATCATCCA CAATAAGCTT GGACTCCATC CCACTATTCG
MNCB6.17        ATTAATTTCA TTACCACGAT CATAAACATA CGTTCCTCAA CCAGAAGCCT GGACTCAATC CCACTATTTG
MNCH2.3_        ATTAACTTTA TCACTACAAT CATAAATATA CGGTCTTCTT CCATGAGGCT TGATTCAATT CCCCTTTTTG
MNCH3.8         ATCAACTTTA TCACAACAAT CATAAACATA CGCTCTTCAT CAATAAGGtT GGATTCAATC CCCCTTTTCG
MNCH4.9         ATTAACTTTA TCACTACAAT CATAAATATA CGGTCTTCTT CCATGAGGCT TGATTCAATT CCCCTTTTTG
MNCH5.17        ATCAATTTCA TCACCACAAT CTTAAATATA CGATCATCCA CAATAAGCTT GGACTCCATC CCACTATTCG
MNNG1           ATCAATTTCA TCACCACAAT CTTAAATATA CGATCATCCA CAATAAGCTT GGACTCCATC CCACTATTCG
MNNG2.1         ATCAATTTCA TCACCACAAT CTTAAATATA CGATCATCCA CAATAAGCTT GGACTCCATC CCACTATTCG
MNUC2           ATTAACTTTA TCACTACAAT CATAAATATA CGGTCTTCTT CCATGAGGCT TGATTCAATT CCCCCTTTTG

                ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                   495        505        515        525        535        545        555
MNCT1           TGTGATCAGT TTTGATTACT GCTGTATTGC TTCTTTTAGC GCTACCCGTT TTAGCAGGAG CTATCACCAT
MNCT2           TATGATCCGT TTTAATCACT GCCGTCTTAT TACTACTGGC GCTACCCGTG TTAGCAGGAG CTATTACCAT
MNCT3           TTTGATCAGT CCTAATTACC GCAGTATTGT TATTACTGGC TTTACCAGTG CTGGCTGGAG CCATCACAAC
MNCT4           TTTGATCAGT CCTAATTACC GCAGTATTGT TATTACTGGC TTTACCAGTG CTGGCTGGAG CCATCACAAC
MNCT5           TTCGATCAGT CCTAATTACC GCAGTATTGT TATTACTGGC TTTACCAGTG CTGGCTGGAG CCATCACAAC
MNCT6           TTTGATCAGT CCTAATTACC GCAGTATTGT TATTACTGGC TTTACCAGTG CTGGCTGGAG CCATCACAAC
MNNF1.1         TATGATCTGT TCTAATTACA GCAATCCTAT TACTCTTAGC TCTTCCAGTA TTGGCAGGGG CCATTACTAT
MNNF1.3         tATGATCTGt TCTAATTACA GCAATCCTAT TACTCTTAGC TCTTCCAGTA TTGGCAGGGG CCATTACTAT
MNNF1.6         TATGATCTGT TCTAATTACA GCAATCCTAT TACTCTTAGC TCTTCCAGTA TTGGCGGGGG CCATTACTAT
MNNF1.7_        TATGATCTGT TCTAATTACA GCAATCCTAT TACTCTTAGC TCTTCCAGTA TTGGCAGGGG CCATTACTAT
MNNF2.9         TATGATCTGt TCTAATTACA GCAATCCTAT TACTCTTAGC TCTTCCAGTA TTGGCAGGGG CCATTACTAT
MNNF3.17_       TATGATCTGT TCTAATTACA GCAATCCTAT TACTCTTAGC TCTTCCAGTA TTGGCAGGGG CCATTACTAT
MNCB1.1         TTTGATCAGT CCTAATTACC GCAGTATTGt TATTACTGGC TTTACCAGTG CTGGCTGGAg CCATCACAAC
MNCB3.7         TATGATCTGT TCTAACTACA GCAATCCTAT TACTCTTAGC TCTTCCAGTA TTGGCAGGGG CCATTaCTAT
MNCB4.9         TTTGATCAGT CCTAATTACC GCAGTATTGT TATTACTGGC TTTACCAGTG CTGGCTGGAG CCATCACAAC
MNCB6.17        TATGATCAGT ACTAATTACT GCAGTGCTAT TACTACTGtC CTTACCAGTA TTAGCAGGGG CAATTaCTAT
MNCH2.3_        TATGATCTGT TCTAATTACA GCAATCCTAT TACTCTTAGC TCTTCCAGTA TTGGCAGGGG CCATTACTAT
MNCH3.8         TCTGATCTGT aCTAATCACA GCCGtCCTTT TACTTCTAGC CCTcCCAGTT TTGGCAGGGG CTATCACAAT
MNCH4.9         TATGATCTGt TCTAATTACA GCAATCCTAT TACTCTTAGC TCTTcCAGTA TTGGCAgGGG CCATTACTAT
MNCH5.17        TTTGATCAGT CCTAATTACC GCAGTATTGt TATTACTGGC TTTACCAGTG CTGGCTGGAG CCATCACAAC
MNNG1           TTTGATCAGT CCTAATTACC GCAGTATTGT TATTACTGGC TTTACCAGTG CTGGCTGGAG CCATCACAAC
MNNG2.1         TTTGATCAGT CCTAATTACC GCAGTATTGT TATTACTGGC TTTACCAGTG CTGGCTGGAG CCATCACAAC
MNUC2           TATGATCTGT TCTAATTACA GCAATCCTAT TACTCTTAGC TCTTCCAGTA TTGGCAGGGG CCATTACTAT

                ....|....| ....|....| ....|....| ....|....|
                   565        575        585        595
MNCT1           ACTTCTAACG GACCGAAGCT TCAATACAAC ATTCTTTGAC
MNCT2           ACTTTTGACA GACCGTAATT TCAACACAAC GTTTTTTGAT
MNCT3           ACTCTTAACA GACCGAAACT TCAATACAAC ATTCTTCGAC
MNCT4           ACTCTTAACA GACCGAAACT TCAATACAAC ATTCTTCGAC
MNCT5           ACTCTTAACA GACCGAAACT TCAATACAAC ATTCTTCGAC
MNCT6           ACTCTTAACA GACCGAAACT TCAATACAAC ATTCTTCGAC
MNNF1.1         ACTTTTAACA GATCGTAATT TTAATACCTC ATTCTTTGAT
MNNF1.3         ACTTTTAACA GATCGtAATT TTAATACCTC ATTCTTTGAT
MNNF1.6         ACTTTTAACA GATCGTAATT TTAATaCCTC ATTCTTTGAT
MNNF1.7_        ACTTTTAACA GATCGTAATT TTAATACCTC ATTCTTTGAT
MNNF2.9         ACTTTTAACA GATCGTAATT TTAATACCTC ATTCTTTGAT
MNNF3.17_       ACTTTTAACA GATCGTAATT TTAATACCTC ATTCTTTGAT
MNCB1.1         ACTCTTAACA GACCGAAACT TCAATACAAC ATTCTTCGAC
MNCB3.7         aCTTTCAACA GATCGTAATT TTAATACCTC ATTCTTTGAT
MNCB4.9         ACTCTTAACA GACCGAAACT TCAATACAAC ATTCTTCGAC
MNCB6.17        ACTATTAACA GACCGTAACT TCAACACTTc ATTCTTTGAC
MNCH2.3_        ACTTTTAACA GATCGTAATT TTAATACCTC ATTCTTTGAT
MNCH3.8         GCTTTTAACT GACCGAAACT TTAACACATC ATTTTTTGAT
MNCH4.9         ACTTTTAACA GATCGTAATT TTAATACCTC ATTCTTTGAT
MNCH5.17        ACTCTTAACA GACCGAAACT TCAATACAAC ATTCTTcGAC
MNNG1           ACTCTTAACA GACCGAAACT TCAATACAAC ATTCTTCGAC
MNNG2.1         ACTCTTAACA GACCGAAACT TCAATACAAC ATTCTTCGAC
MNUC2           ACTTTTAACA GATCGTAATT TTAATACCTC ATTCTTTGAT
```

Chapter Three

3. Absence of recombination in oribatid mites

3.1. Introduction

Recombination is assumed to be one of the main advantages of sexual reproduction (Kondrashov 1993; West et al. 1999) bringing together favourable mutations (Fisher 1930; Müller 1932) and being an important mechanism in DNA repair (Bernstein and Bernstein 1991; Michod 1993; Michod 1998). Recombination is also responsible for a non-biased distribution of genes after meiosis and to avoid segregation distorters from acting (Ridley 2000). Therefore, the absence of recombination is assumed to be an evolutionary dead end (Maynard Smith 1978; see also chapter 1.3).

Just one exception is known to be evolutionary successful in the long-term without recombination: Bdelloid rotifers (Mark Welsh 1999; Mark Welsh and Meselson 2000).

In sexually reproducing species, recombination and segregation allow random drift to drive selectively neutral alleles toward fixation or extinction,

limiting the divergence between allelic sequences (Kimura and Crow 1964; Kreitman 1983; Mark Welsh and Meselson 2000). Reported species averages for synonymous site diversity in a wide range of animal species are between 0.1 an 4% (Li and Sadler 1991; Palumbi and Metz 1991; Avise 1994; Moriyama and Powell 1996; Wang et al. 1997; Mark Welsh and Meselson 2000).

Alleles are alternative forms of a single gene, differing from each other in sequence but coding for the same polypeptide or RNA. They usually occupy the same locus on particular homologous chromosomes and are referred to as homozygous if the sequence is identical or heterozygous if it is not (Birky 1996).

Asexual species lack the ability of recombination and segregation. Descendants of formerly allelic sequences within individual genomes therefore are supposed to be highly divergent if not lost by deletion, mitotic crossing over, gene conversion or reduction of ploidy (Birky 1996; Mark Welsh and Meselson 2000). Chromosomes may also acquire different rearrangements resulting in heteromorphy (White 1973).

Allele sequence divergence confounds phylogenetic analysis, causing gene trees to depart drastically from species trees and making it difficult or impossible to recover the correct tree topology (Birky 1996). A remarkable consequence of asexual reproduction is that the two alleles in an individual may differ from each other more than each does from an allele in a related species or clone of the same asexual clade (Birky 1996). If asexual radiation occurred, a phylogeny of the two alleles should consist of two clades for each allele containing the separate species (Figure 3.1).

Therefore, the analysis of the two alleles and their divergence may lead to conclusions about presence or absence of gene recombination during the radiation of asexual lineages and look different in sexual and asexual species (Birky 1996; Figure 3.1).

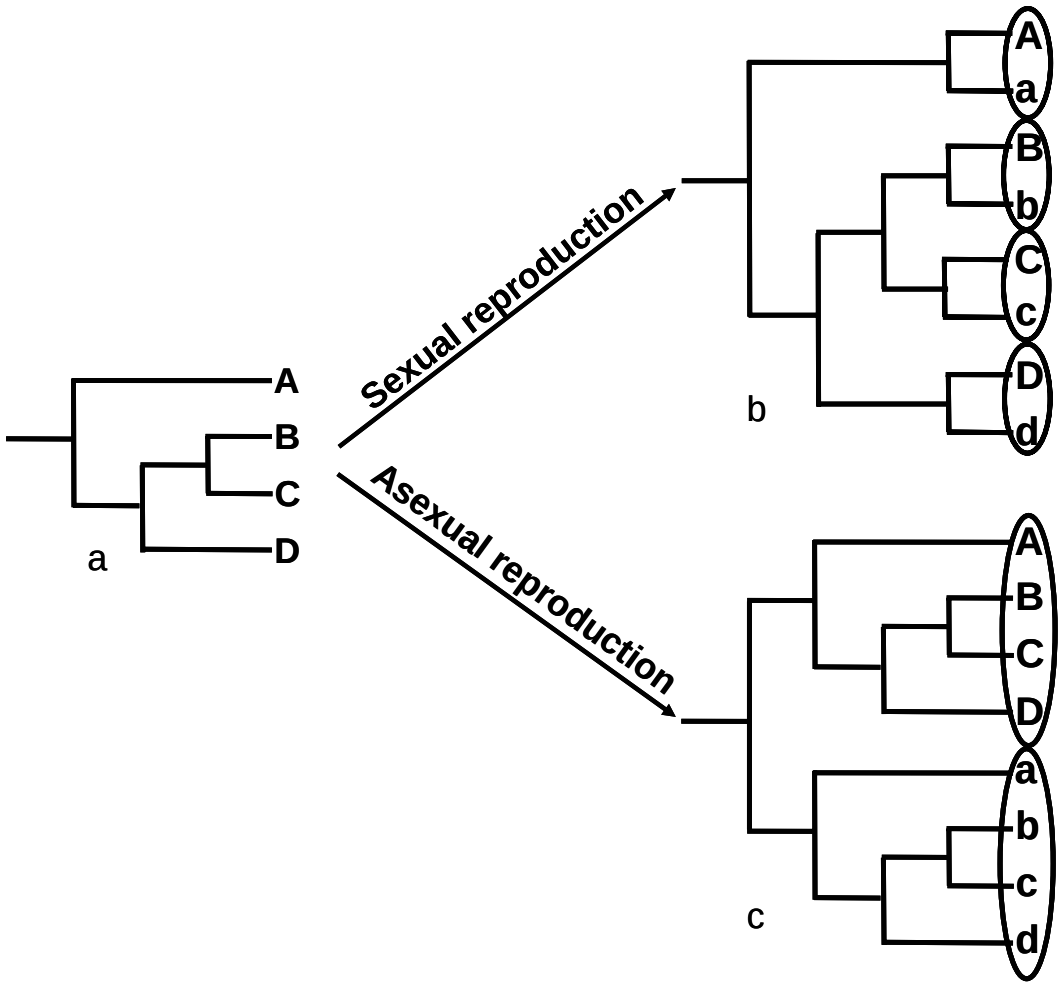


Figure 3.1

Phylogeny of four species A, B, C and D which are related like (a). Phylogenetic analyses using both alleles of each individual (indicated by large and small letters) lead to different tree topologies according to reproductive mode. With sexual reproduction and recombination, allele pairs within an individual are more closely related to each other than to any allele in another individual (b). With asexual reproduction and lack of recombination, each allele within an individual is more closely related to its "sister" in another individual than to the second, homologous allele within the same individual (c) (Birky 1996).

This analysis is most appropriate for single copy genes which are not accompanied by a paralog (Mark Welch and Meselson 2000). An auspicious candidate is the *hsp82* gene for which BLAST searches (Altschul et al. 1997) revealed no species in which the gene has a paralog

since the origin of metazoans (Mark Welsh and Meselson 2000). In eukaryotic cells, specific *hsp90* homologs exist in the cytosol (Cy) and the endoplasmatic reticulum (ER) (Gupta 1995); homologs of HSP90 have also been found in prokaryotes (referred to as HptG) (Bardwell and Craig 1987). The Cy and the ER *hsp90* homologs constitute paralogous families which diverged from each other at a very early stage in the evolution of eukaryotic cells and another Cy *hsp90* gene differentiation has happened in the vertebrate clade (Gupta 1995). How these two forms differ in their physiological function is not known (Lindquist and Craig 1988).

The *hsp82* gene is coding for an 82kD protein of the HSP90[1] family, proteins which are highly conserved in animals, fungi and prokaryotes (Picard 2002). HSP90 corresponds to about 1% of the cytosolic proteins in unstressed cells and is therefore one of the most abundant proteins (Lai et al. 1984; Nollen and Morimoto 2002). Due to the high concentration in unstressed cells and its only few-fold induction after stress, the name “heat shock protein” is somewhat of a misnomer for HSP90 (Picard 2002). HSP90 contributes to the folding, maintenance of structural integrity and proper regulation of a subset of cytosolic proteins involved e.g. in cell cycle control and signal transduction (Wiech et al. 1992; Miyata and Yahara 1992).

By analysing allelic divergence of the *hsp82* gene it was aimed to investigate if the assumed 180-240 million years since the split of the mitochondrial COI lineages (see chapter 2) were really without intrachromosomal recombination in the nuclear genome. To address this question, species with different reproductive modes were analysed. Three species from exclusively parthenogenetic taxa, one sexual and one asexual species from a group where both reproductive modes exist, and one sexual species from an exclusively sexual taxon were included. This set of species enables comparisons of allelic divergence between species and taxa with different modes of reproduction.

[1] The HSP90 family refers to different names in the literature: HSP80, 82, 83, 84, 85, 86 and 90 (Gupta 1995). Here, HSP82 and HSP90 are used as referring to the same protein in oribatid mites.

3.2. Materials and methods

Sampling of species

A partial sequence of the *hsp82* gene was analysed in different sexual and asexual oribatid mite species (Table 3.1). Between 2 and 20 sequences were obtained from each specimen. Analysed specimens for which less than five sequences were obtained and all of these had the same sequence were excluded from the analysis and are not listed.

Table 3.1
Analysed oribatid mite species. For details on the geographic origins see chapter 2.2. P: parthenogenetic species; S: sexual species.

Species	Taxon	Origin	Abbreviation	Reproduction
Platynothrus yamasakii	Camisiidae	China	PYCH	P
Platynothrus peltifer	Camisiidae	Germany Germany Kashmir	PPKWX PPKWY PPKA	P
Mucronothrus nasalis	Trhypochthoniidae	New Guinea Canada	MNNG MNCB	P
Tectocepheus velatus	*Tectocepheus*	Germany	TVDO	P
Atropacarus striculus	Mixonomata	Germany	ASA	P
Steganacarus magnus	Mixonomata	Germany	SM	S
Metabelba pulverulenta	Belbidae	Germany	MPA	S

This set of species was chosen to span a wide range of taxa with different patterns of reproductive biology. *Platynothrus* is a member of Camisiidae. Camisiidae form one of the speciose (92 species) taxa of oribatid mites which exclusively reproduces by parthenogenesis and presumably

radiated without sexual reproduction. Two species of *Platynothrus* were analysed and from *P. peltifer*, two different geographical origins were considered. More details on the biology of the analysed species from Camisiidae are given in chapter two. *M. nasalis* belongs to the Trhypochthoniidae. This also is an exclusively asexual taxon comprising 68 species. The biology of *M. nasalis* is also described in chapter two. *T. velatus* and the whole genus *Tectocepheus* are presumably also parthenogenetic (Nübel-Reidelbach 1994). There are nearly forty species-group names under this genus but maybe just fifteen are valid (Norton, personal communication). Phylogenetically, *Tectocepheus* were usually assigned to the Apheredermata; molecular analysis and recent morphological studies more likely assign them to the Poronota (Maraun et al. 2003b). *Tectocepheus velatus* is of cosmopolitan distribution, has a wide ecological niche and is amongst the most abundant oribatid mite species in soil (Murphy and Jalil 1964). Like the Camisiidae and Trhypochthoniidae they are assumed to have evolved by asexual radiation. The Mixonomata comprise sexual and asexual species. I analysed the sexual species *Steganacarus magnus* and the parthenogen *Atropacarus striculus*. As the Mixonomata are a group in which both reproductive modes exist (while sexual reproduction is more common; Norton and Palmer 1991) it is an important group for comparisons with exclusively parthenogenetic and exclusively sexual taxa. Although the Mixonomata comprise sexual and asexual species, cyclical or geographical parthenogenesis are absent, as like as in all other oribatid mites. As a representative of an exclusively sexually reproducing taxon *Metabelba pulverulenta* (Damaeidae) was sampled.

This assembly of species enables comparisons among taxa with parthenogenetic reproduction, taxa with mixed reproduction and taxa with sexual reproduction.

Molecular techniques

Most sequences were obtained by cloning (see chapter 2.2). I designed the denatured primer pair hspeu1 (5'TCNATGATHGGNG ARTTYGGTGTNGGTTTYTA3') and hspeu2 (5'YTTNACNGCTCA RTGRTCYTCCCARTCRTT3') and I used primers designed by Schön and Martens (2003) (hsp1.2: 5'TGCTCTAGAGCACARTTYGGTGTNGGT TTYTA3'; hsp8.x: 5'ACGTTCTAGARTGRTCYTCCCARTCRTTNGT3') for amplification via PCR. The total reaction volume of 50 µl contained 1.5 mM $MgCl_2$, 200 µM of each dNTP, 200 pmol of each primer and 2.5 units of Taq polymerase. The two-step PCR conditions were (i) 15 min at 95°C for polymerase activation, 50 sec at 94°C for denaturation, 50 sec at 50°C for primer annealing and 120 sec at 72°C for elongation for 10 cycles and (ii) 50 sec at 94°C for denaturation, 50 sec at 55°C for primer annealing and 120 sec at 72°C for elongation for 36 cycles followed by a terminal 10 min elongation at 72°C.

In addition, TTGE (Temporal Temperature Gradient Electrophoresis; Zoller et al. 1998) analyses were conducted for some samples to confirm the identification of the two alleles; the differences in length of the two *hsp82* alleles were too small to be separated by agarose gel electrophoresis. TTGE was performed with the DCode™ universal mutation detection system (BioRad, Munich) which allows to differentiate between two sequences of equal length which differ in at least one nucleotide position. The technique is an advancement of the commonly used DGGE (Denaturing Gradient Gel Electrophoresis; Fischer and Lerman 1979) and TGGE (Temperature Gradient Gel Electrophoresis; Rosenbaum and Riesner 1987). The replacement of spatial temperature gradients by temporal temperature gradients allows the use of simpler electrophoretic systems with a higher reproducibility (Yoshino et al. 1991; Wiese et al. 1995; Borresen-Dale et al. 1997). TTGE uses a single constant concentration of urea and formamid for the denaturation of doublestranded DNA while gradually increasing the temperature during the run (Borresen-Dale et al. 1997). TTGE gels contained 6% polyacrylamide/bis (37.5:1) and 7M urea. Urea decreases the melting temperature of double stranded

DNA by 2°C/M in the gel and allows TTGE to be used at moderate temperatures (Lottespeich and Zorbas 1998; Zoller et al. 1998). The temperature range for electrophoresis of *hsp82* was estimated using WinMelt (BioRad), a program for calculation of melting temperatures of sequences for TTGE analysis (Figure 3.2). The program is based on an algorithm to estimate the temperature where half of each basepair is bound and half is melted (Lerman and Silverstein 1987). Despite the fact that every basepair has an individual melting temperature it was shown that domains of up to 300 basepairs are correlated with their melting properties (Lerman and Silverstein 1987) and usually, sequences contain more than one of such melting domains (Myers et al. 1985). The temperature range for different runs was corrected according to the urea concentration of the gel.

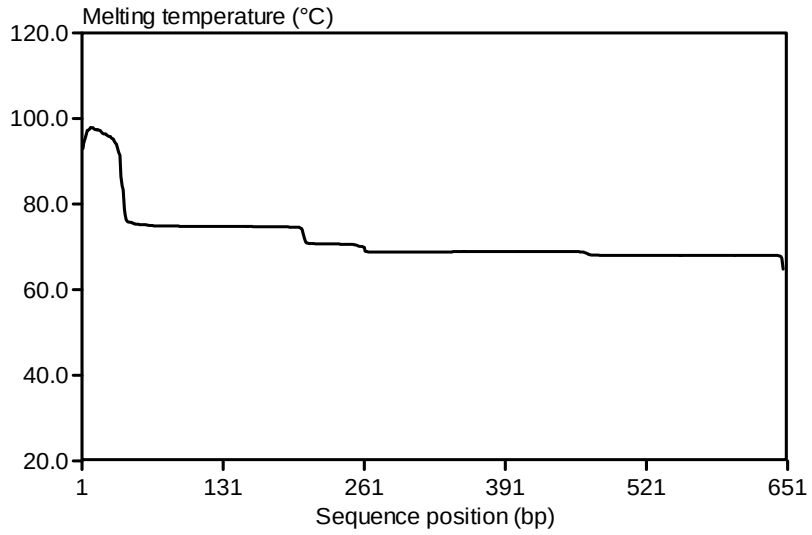


Figure 3.2
Melting profile of *hsp82* as estimated by the program WinMelt.

To avoid total dissociation of double stranded DNA, one primer (hsp1.2) was modified by a GC-clamp constituting of 40 GC nucleotides (Borresen et al. 1991).

According to 7 M urea concentration, the temperature range was set to 48-66°C (corresponding to 62-80°C in a non-denaturing gel) with a ramp between 1.5 and 5°C/h. The highest temperature does not melt the GC-clamp domain ensuring that DNA remains double stranded during the run.

Data analysis

Affiliations of obtained sequences to the *hsp82* gene were ensured via BLAST search (Altschul et al. 1997) and comparisons of signal sequences in the analysed region of the protein (Gupta 1995). Sequences were compared with the paralogous sequences from HSP90 known to be acting in the endoplasmatic reticulum to ensure that all analysed sequences belong to the cytoplasmatic *hsp82* copy (Table 3.2). Contaminations from other organisms, such as fungi, nematodes or bacteria representing parasites or gut content were excluded by a deep branch phylogenetic analysis of the HSP90 protein. The Genbank database was screened for available sequences of HSP90 from different taxa (Table 3.3).

Alignments of amino acid sequences were performed using ClustalX (Higgins and Sharp 1988, 1989; Higgins et al. 1992; Thompson et al. 1994; Thompson et al. 1997) with the BLOSUM 30 scoring matrix for weighing of different substitutions (BLOck SUbstitution Matrix; Henikoff and Henikoff 1992; Figure 3.2), a gap opening penalty of 10 and a gap extension penalty of 0.1. The BLOSUM 30 matrix assumes that two sequences are closely related when their amino acid sequences match in more than 30% of positions and clusters these sequences.

Table 3.2
Sequences and accession numbers used for testing of ER copies

Taxon	Species	Location	AN
Viridiplantae	*Catharantus roseus*	ER	L14594
Vertebrata	*Canis familiaris*	ER	U01153
	Sus scrofa	ER	X76301

Table 3.3
Sequences and accession numbers used for testing of contaminations

Taxon	Species	AN
Prokaryota	*Campylobacter jejuni jejuni*	NC_002163
	Borellia burgdorferi	NC_001318
	Mycobacterium leprae	NC_002677
	Pseudomonas putida	NC_002947
	Escherichia coli	NC_002655
	Salmonella typhimurium	NC_003197
	Shigella flexneri	NC_004741
"Protozoa"	*Tetrahymena thermophila*	AF151114
	Babesia bovis	AF136649
	Theileria parva	M57386
	Eimeria tenella	AF042329
	Crypthecodinium cohnii	AF421541
	Leishmania infantum	X87770
	Trypanosoma brucei	X14176
Viridiplantae	*Ipomoea nil*	M99431
	Arabidopsis thaliana	NM_124983
	Zea mays	S59780
Fungi	*Saccharomyces cerevisiae*	NC_001148
	Schizosaccharomyces pombe	NC_003424
	Candida tropicalis	AF251005
	Ajellomyces capsulatus	M55629
Nematoda	*Brugia pahangi*	AJ005784
	Meloidogyne javanica	AF201338
	Heterodera glycines	AF461150
Rotifera	*Oncicola* sp.	AF375826
	Habrotrocha constricta	AF249999
	Philodina roseola	AF250004
Hexapoda	*Spodoptera frugiperda*	AF254880
	Bombyx mori	AB060275
	Anopheles albimanus	L47285
	Drosophila melanogaster	U57473
	Sarcophaga crassipalpis	AF261773
Pisces	*Salmo salar*	AF135117
	Oncorhynchus tshawytscha	U89945
	Danio rerio	NM_131310
Mammalia	*Sus scrofa*	U94395
	Equus caballus	AB043677
	Rattus norvegicus	AJ428213

For comparison of the oribatid mite sequences, nucleotide sequences were translated into amino acids using the standard genetic code in BioEdit 5.0.9 (Hall 1999). Amino acid sequences were aligned to avoid codons from being disrupted during the alignment of nucleic acid sequences. Aligned amino acid sequences were translated back to nucleotide sequences and the originally used codons.

Allelic distances were estimated by calculating uncorrected percentage distances of fourfold degenerate sites (see chapter 2.2.) and percentage distance of DNA and amino acid sequences.

	A	R	N	D	C	Q	E	G	H	I	L	K	M	F	P	S	T	W	Y	V	B	Z	X	*
A	4	-1	0	0	-3	1	0	0	-2	0	-1	0	1	-2	-1	1	1	-5	-4	1	0	0	0	-7
R	-1	8	-2	-1	-2	3	-1	-2	-1	-3	-2	1	0	-1	-1	-1	-3	0	0	-1	-2	0	-1	-7
N	0	-2	8	1	-1	-1	-1	0	-1	0	-2	0	0	-1	-3	0	1	-7	-4	-2	4	-1	0	-7
D	0	-1	1	9	-3	-1	1	-1	-2	-4	-1	0	-3	-5	-1	0	-1	-4	-1	-2	5	0	-1	-7
C	-3	-2	-1	-3	17	-2	1	-4	-5	-2	0	-3	-2	-3	-3	-2	-2	-2	-6	-2	-2	0	-2	-7
Q	1	3	-1	-1	-2	8	2	-2	0	-2	-2	0	-1	-3	0	-1	0	-1	-1	-3	-1	4	0	-7
E	0	-1	-1	1	1	2	6	-2	0	-3	-1	2	-1	-4	1	0	-2	-1	-2	-3	0	5	-1	-7
G	0	-2	0	-1	-4	-2	-2	8	-3	-1	-2	-1	-2	-3	-1	0	-2	1	-3	-3	0	-2	-1	-7
H	-2	-1	-1	-2	-5	0	0	-3	14	-2	-1	-2	2	-3	1	-1	-2	-5	0	-3	-2	0	-1	-7
I	0	-3	0	-4	-2	-2	-3	-1	-2	6	2	-2	1	0	-3	-1	0	-3	-1	4	-2	-3	0	-7
L	-1	-2	-2	-1	0	-2	-1	-2	-1	2	4	-2	2	2	-3	-2	0	-2	3	1	-1	-1	0	-7
K	0	1	0	0	-3	0	2	-1	-2	-2	-2	4	2	-1	1	0	-1	-2	-1	-2	0	1	0	-7
M	1	0	0	-3	-2	-1	-1	-2	2	1	2	2	6	-2	-4	-2	0	-3	-1	0	-2	-1	0	-7
F	-2	-1	-1	-5	-3	-3	-4	-3	-3	0	2	-1	-2	10	-4	-1	-2	1	3	1	-3	-4	-1	-7
P	-1	-1	-3	-1	-3	0	1	-1	1	-3	-3	1	-4	-4	11	-1	0	-3	-2	-4	-2	0	-1	-7
S	1	-1	0	0	-2	-1	0	0	-1	-1	-2	0	-2	-1	-1	4	2	-3	-2	-1	0	-1	0	-7
T	1	-3	1	-1	-2	0	-2	-2	-2	0	0	-1	0	-2	0	2	5	-5	-1	1	0	-1	0	-7
W	-5	0	-7	-4	-2	-1	-1	1	-5	-3	-2	-2	-3	1	-3	-3	-5	20	5	-3	-5	-1	-2	-7
Y	-4	0	-4	-1	-6	-1	-2	-3	0	-1	3	-1	-1	3	-2	-2	-1	5	9	1	-3	-2	-1	-7
V	1	-1	-2	-2	-2	-3	-3	-3	-3	4	1	-2	0	1	-4	-1	1	-3	1	5	-2	-3	0	-7
B	0	-2	4	5	-2	-1	0	0	-2	-2	-1	0	-2	-3	-2	0	0	-5	-3	-2	5	0	-1	-7
Z	0	0	-1	0	0	4	5	-2	0	-3	-1	1	-1	-4	0	-1	-1	-1	-2	-3	0	4	0	-7
X	0	-1	0	-1	-2	0	-1	-1	-1	0	0	0	0	-1	-1	0	0	-2	-1	0	-1	0	-1	-7
*	-7	-7	-7	-7	-7	-7	-7	-7	-7	-7	-7	-7	-7	-7	-7	-7	-7	-7	-7	-7	-7	-7	-7	1

Figure 3.2
BLOSUM 30 matrix with scores and penalties for different substitution types. Scores are between -7 and 20. Positive values indicate common substitutions and are shaded grey. Negative values are uncommon substitutions.

Likelihood ratio tests were conducted (see chapter 2.2) to estimate an applicable evolutionary model for the nucleotide sequence dataset. Phylogenetic analyses of the nucleotide sequences were performed using maximum likelihood algorithms in PAUP* (Swofford 1999) with the estimated parameters. Maximum likelihood approaches chose between different trees by calculating which tree (H) has the highest likelihood (L) to be the outcome of the particular dataset (D; L=P(D/H)). The probability of given substitutions is calculated for each site of the dataset by involving discrete models of sequence evolution and is summed up for the whole dataset. The tree with the highest likelihood value is assumed to be the best tree for the observed data (Page and Holmes 1998).

Phylogenetic analyses of amino acid sequences were conducted using Neighbour Joining (NJ; Saitou and Nei 1987) in PAUP* (Swofford 1999).

The robustness of all phylogenetic trees was tested by bootstrap analyses. All trees were tested with 10,000 bootstrap replicates and NJ as tree building algorithm.

3.3. Results

Two allelic sequences were obtained from five parthenogenetic and two sexual species (Table 3.1). Lengths of alleles were between 534 and 549 base pairs corresponding to 178 and 183 amino acids, respectively. All sequences were free of introns, senseless and stop codons and were verified by BLAST search (Altschul et al. 1997) and signal sequence comparisons (Gupta 1995).

The proportion of variable sites for the entire sequences was between 0% for the sexual species *S. magnus* and 26% for the parthenogen *P. peltifer*. The average distance of the two alleles was 18% over all species examined. D4 distances were at a maximum in the sexual species *M. pulverulenta* and smallest in *S. magnus*. Over all species, D4 distances averaged 54%. Protein distances between the two alleles were on

average 11% with a maximum value of 15% for *M. nasalis* and *P. peltifer* (Figure 3.3, Table 3.4). Three different sequences were found in two specimens (*A. striculus* and *M. nasalis*).

Table 3.4
Allelic divergences of analysed specimens

	Number of alleles	DNA length [bp]	**DNA variable [%]**	D4 sites [bp]	**D4 variable [%]**	Protein length [aa]	**Protein variable [%]**
Platynothrus yamasakii PYCH	2	537 546	**20**	46	**61**	179 182	**14**
Platynothrus peltifer PPKWX	2	534 534	**26**	47	**70**	178 178	**15**
Platynothrus peltifer PPKWY	2	534 546	**19**	49	**55**	178 182	**13**
Platynothrus peltifer PPKA	2	537 546	**19**	46	**59**	179 182	**11**
Mucronothrus nasalis MNNG	2	546 549	**17**	48	**63**	182 183	**9**
Mucronothrus nasalis MNCB	3	540 546 546	**19**	46	**50**	180 182 182	**15**
Atropacarus striculus ASA	3	534 543 546	**21**	48	**69**	178 181 182	**12**
Tectocepheus velatus TVDO	2	540 546	**16**	46	**41**	180 182	**12**
Steganacarus magnus SM	2	534 534	**0**	53	**0**	178 178	**0**
Metabelba pulverulenta MPA	2	534 546	**24**	48	**73**	178 182	**10**

Four specimens representing two species of the parthenogenetic genus *Platynothrus* were analysed. Within *Platynothrus*, D4 divergences

averaged 61% and amino acid divergence averaged 13%. Of the two specimens of the parthenogen *M. nasalis*, one had two different sequences and one had three; divergence between alleles were on average 57% regarding D4 sites and 12% on amino acid level. *T. velatus* differed in 16% and 12% on DNA and amino acid level, respectively, with a divergence of 41% in neutrally evolving sites. The two sexual species showed different patterns: while for *S. magnus* just one sequence was obtained, *M. pulverulenta* had two sequences differing at 73% of D4 sites.

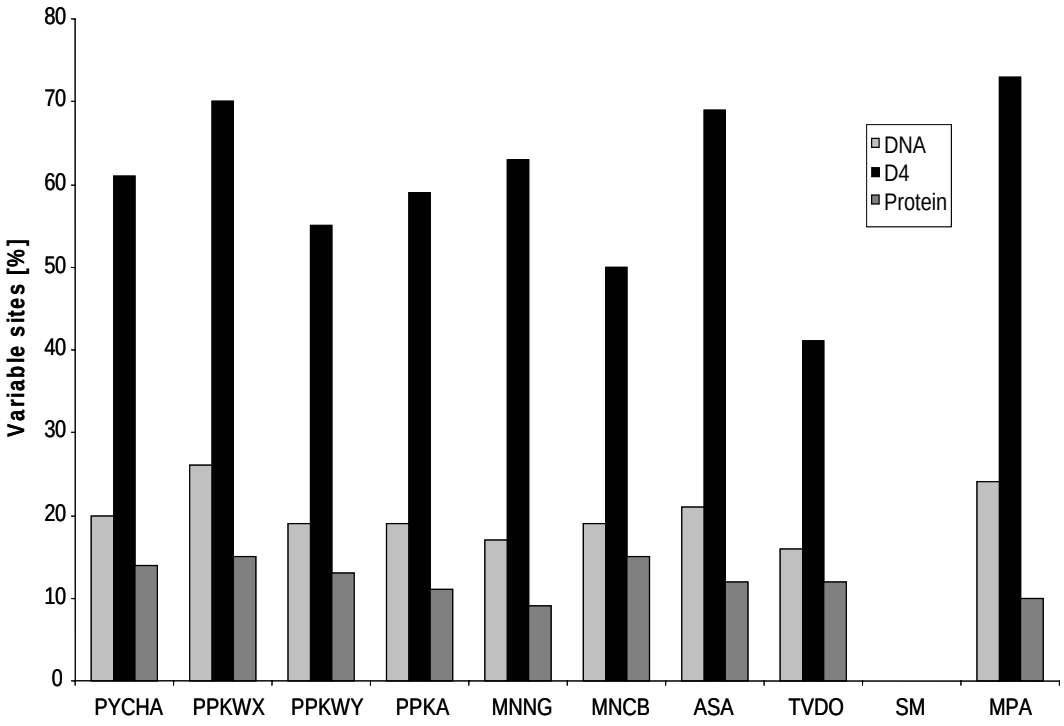


Figure 3.3
Proportion of variable sites as an indicator of allelic distances for analysed specimens. For *S. magnus* (SM) just one sequence was found. For abbreviations see Table 3.1.

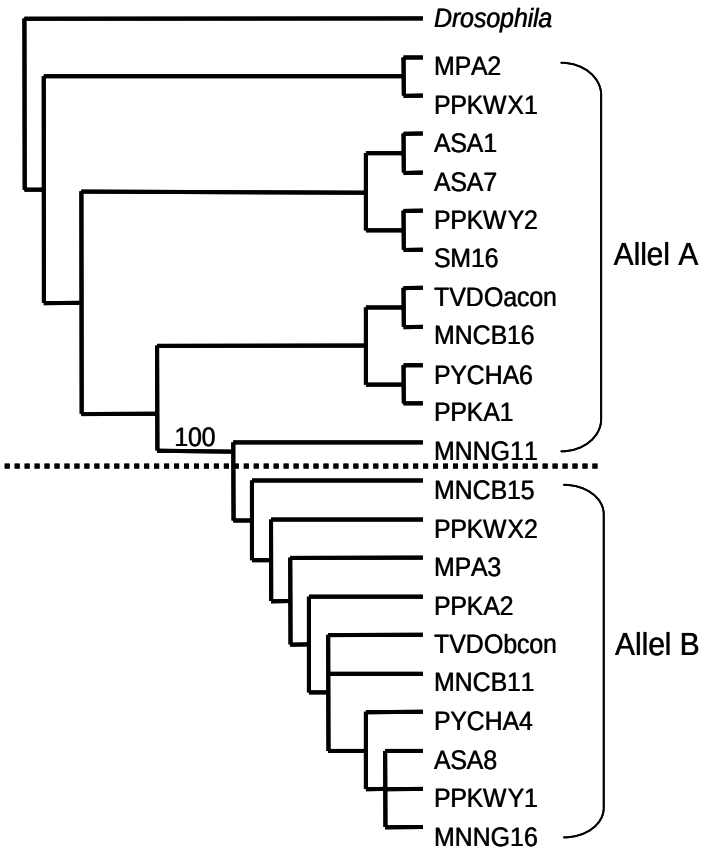


Figure 3.4
Phylogenetic analysis of nucleotide sequences. ML algorithm based on the TrN+G model (Tamura and Nei 1993; Yang 1996).). Base frequencies and substitution parameters as estimated by a likelihood ratio test: A (0.3552), C (0.1949), G (0.2963), T (0.1536); [A-G] (2.7885), [C-T] (6.865); Gamma: 0.4055.
The dashed line indicates the possible split of the two alleles. Numbers on branches indicate bootstrap values from 10,000 replicates (not all shown).

Phylogenetic analysis of all allelic sequences leads to a tree topology where alleles between individuals are more closely related to each other than those within individuals (Figure 3.4). The bootstrap value of 100 strongly supports a robust split of the tree in two parts, each containing one of the suspected alleles (allele A and B) from all specimens. All sequences from allele B are very uniform while allele A sequences are more diverse.

While in the sexual species *S. magnus* only one allele was found representing the allelic cluster A, the second sexual species, *M.*

pulverulenta, had two alleles with the one being represented in the allelic cluster A and the other in B. The allelic sequence divergence of *M. pulverulenta* was in the same range as in the analysed parthenogenetic species (Figure 3.3).

In the two specimens where three allelic sequences were obtained, two of these were represented in one of the allelic clusters (A for *A. striculus* and B for *M. nasalis*) and the third in the other allelic cluster. While the three sequences just differ slightly in *A. striculus* (indicated also by phylogenetic position of ASA1 and ASA7), the differences in the three sequences of *M. nasalis* were more pronounced (MNCB11, MNCB15 and MNCB16 are widely spread over the phylogenetic tree).

3.4. Discussion

The absence of recombination during the evolution and radiation of asexual species clusters leads to a specific phylogenetic topolology if both alleles are used for phylogenetic analyses (Birky 1996, Figure 3.1). Phylogenetic analyses of the obtained *hsp82* alleles had the expected topology under the assumption that intrachromosomal recombination was absent (Figure 3.4). If both sequences represent alleles then recombination was absent during the whole period of evolution of the analysed oribatid mite species. The last common ancestor of analysed species lived in the Carboniferous about 350 million years ago (Norton, personal communication) as indicated by fossil records. The high allelic divergences between alleles, up to 70% in neutrally evolving sites, 26% on DNA level and 15% in the amino acid sequence, are profound indicators that this in fact was the case.

To confirm this assumption, several alternative conclusions have to be considered: (i) one of the sequences may correspond to the *hsp90* paraloge from the endoplasmatic reticulum, (ii) one of the sequences may represent contaminations from gut content or parasites or (iii) the two

sequences may not represent alleles but paralogous genes due to a gene duplication.

To confirm that all obtained sequences belong to the cytoplasmatic copy of oribatid mites rather than to any kind of contamination from the ER or other organisms, a deep branch phylogeny was conducted with amino acid sequences of the ER paralog and cytoplasmatic sequences from a wide range of taxa from prokaryotes to vertebrates (Table 3.2 and 3.3; Figure 3.5).

All sequences from oribatid mites formed a distinct cluster in this phylogenetic analysis (Figure 3.5); this grouping is highly supported by a bootstrap value of 100. A high bootstrap value was also found for Arthropoda represented by oribatid mites and five insect species. As there are no sequences of possible contaminations related to oribatid mites, contaminations from gut content (fungi or plants) or parasites (nematodes or prokaryotes) are unlikely. The paralog from the endoplasmatic reticulum is at the basis of the phylogenetic tree and therefore, it is unlikely that it is represented in the obtained oribatid mite sequences. Hence, all sequences likely represent the cytosolic *hsp82* gene from the oribatid mite genome.

As assumed by Gupta (1995), HSP90 is a suitable genetic marker for deep branch phylogeny. My phylogenetic analysis supports this findings as all important groups are supported by high bootstrap values.

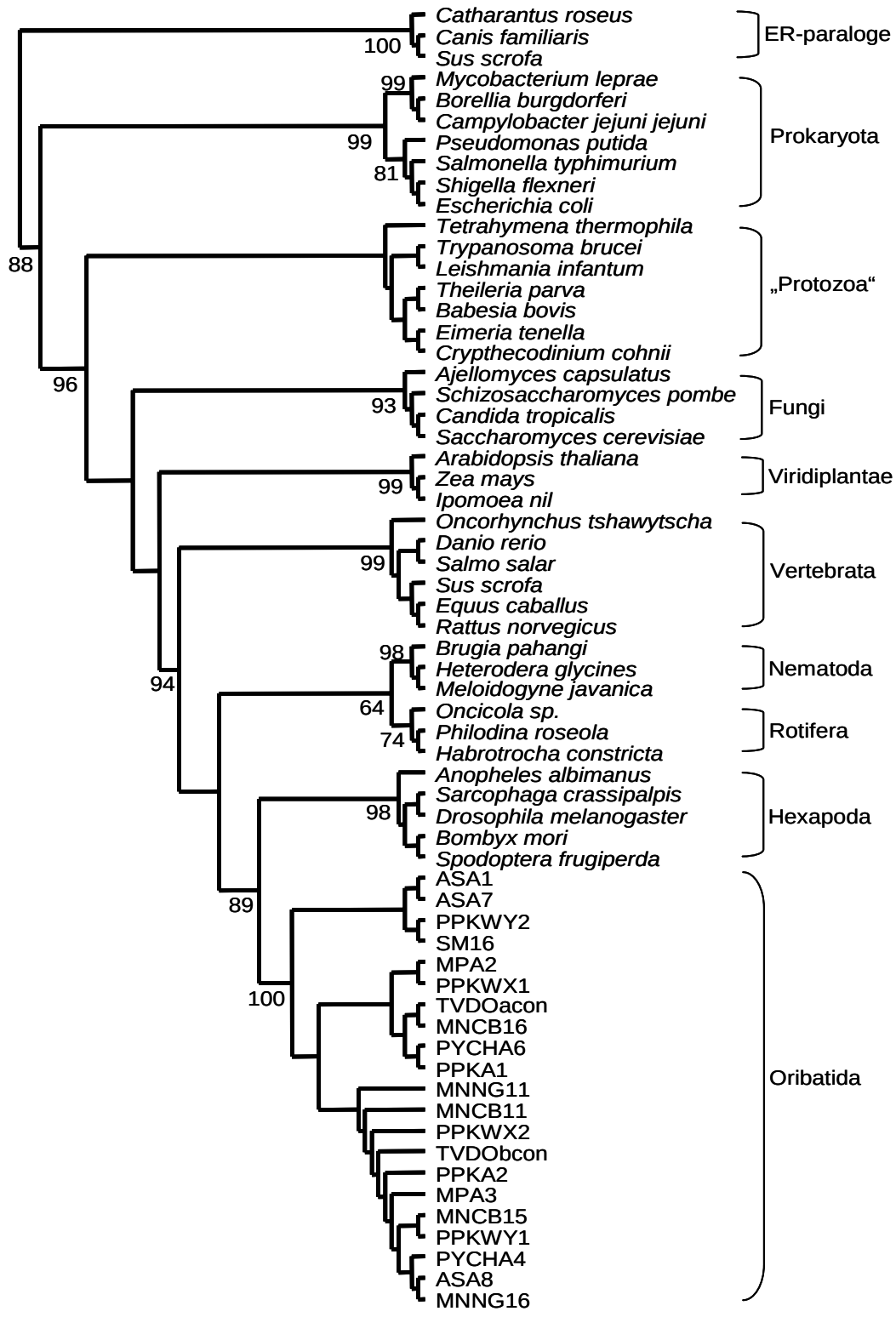


Figure 3.5
Phylogenetic analyses of HSP90 from a wide range of different taxa and from the eukaryotic ER paralogue. Alignment was conducted with ClustalX (BLOSUM30 matrix, gap open penalty: 10, gap extension penalty: 0.1). Values represent bootstrap analyses with 10,000 replicates. For abbreviations see Table 3.4.

An interesting finding of this phylogeny is that nematodes and arthropods are not closely related as assumed by Aguinaldo et al. (1997) and do not represent a monophyletic group “Ecdysozoa”, a taxon which was intensively debated during the last years (Aguinaldo et al. 1997; Giribet and Wheeler 1999). Although bootstrap values are not strongly supporting that nematods and rotifers are sister taxa a close relationship of nematodes and arthropods excluding rotifers is not supported at all. This is not in the focus of this thesis but further analyses of HSP90 from other groups of the “Ecdysozoa”, Annelida and Mollusca might be an interesting field of research to evaluate relationships among protostome animal taxa.

Due to the topology of the phylogenetic analyses, contaminations were shown to be unlikely represented in the obtained sequences attributive to oribatid mites in this study. However, it has to be demonstrated unequivocally that the sequences are not paraloges which originate from gene duplication within oribatid mites or chelicerates.

In general, gene duplications have been assumed to have an important influence on biological evolution even before the structure of the DNA was clarified by Watson and Crick (1953) (Bridges 1936; Stephens 1951). As more genomic information became available in the 1990s it became clear that a substantial part of prokaryotic and eukaryotic genomes originated from gene duplications (Zhang 2003) with up to 65% of all genes in *Arabidopsis thaliana* (Kaul et al. 2000) and 41% in *Drosophila melanogaster* (Rubin et al. 2000). Most likely the contribution of duplicated genes to the eukaryote genome is even higher because many duplicated genes diverged strongly thereby masking their origin (Zhang 2003). Gene duplications were estimated to arise and be fixed at an approximate rate of one gene per 100 million years (Lynch and Conery 2000) but many of the duplicated genes later become nonfunctional pseudogenes which might be deleted from the genome (Wang et al. 2001) or acquire a new function (Li 1983). Gene duplications can result from unequal cross over or retrotransposition. Unequal cross over usually generates tandem gene duplications with linked genes in a chromosome which might evolve together by concerted evolution (Dover 1982; Hurst and Smith 1998).

Retrotransposition occurs when mRNA is retrotranscribed to complementary DNA (cDNA) and then inserted into the genome. Due to this process, there are several molecular features of retrotransposition such as the loss of introns and regulatory sequences which are not represented in the mRNA (Long 2001). In addition, genes duplicated by retrotransposition are usually unlinked to the original gene because the insertion of cDNA into the genome is more or less random (Zhang 2003).
When a gene duplication event occurs, the duplicate genes have redundant functions. Many deleterious mutations may then be harmless, because even if one gene suffers a mutation, the redundant gene copy can provide a backup function (Wagner 2002). Therefore, one of the duplicates should experience relaxed selective constraints that result in elevated rates of evolution indicated by non-synonymous substitutions having the same rate as synonymous substitutions (Wagner 2002). Paralogs which originated by retrotransposition are therefore usually easy to detect and have a likely fate of pseudogenisation, the process by which a functional gene becomes a pseudogene (Zhang 2003). This pseudogenisation usually occurs in the first few million years after duplication if the duplicated gene is not under selection (Lynch and Conery 2000). Two genes with identical functions are unlikely to be stably maintained in the genome (Nowak et al. 1997) unless they differ in some aspects of their functions which can occur by subfunctionalisation like the division of gene expression after duplication (Force et al. 1999). One of the most important outcomes of gene duplication is the origin of a novel function (neofunctionalisation) although it seems improbable that entirely new functions could emerge in a duplicated gene (Zhang 2003). Generally, gene duplication probably is an important mechanism to make biological systems robust against genetic turbulence (Kitami and Nadeau 2002; Gu et al. 2003).

Can both copies of the *hsp82* gene found in oribatid mites be ascribed to alleles rather than to duplicated genes? Absence of introns in one of the two sequence clusters would be an indicator of retrotranspositional gene

duplication. As both sequence clusters were free of introns, the presence/absence of introns cannot be used as an indicator for gene duplication of one of the two copies. The predominant type of substitutions was synonymous in both alleles (ss-sn=0.95; Kumar 2001; chapter 2.2) indicating that both copies are under negative selection and therefore are likely functional. Also, all sequences are highly conserved in a signal sequence region of the HSP90 family (Gupta 1995) indicating affiliation to HSP90. Together with the ancient split of analysed oribatid mite species (350 million years) this makes pseudogenisation of one of the copies unlikely although allele A exhibits a higher degree of divergence than allele B (Figure 3.3). Expression studies of the two sequence clusters are lacking, therefore, it cannot be concluded if there is a subfunctionalisation between the two genes. Analyses of transcribed mRNA might give information, if both copies represent active genes. In fact, after RNA isolation and specific cDNA synthesis using the same primers as for DNA, both copies were represented in the mRNA pool of oribatid mites (Domes, Scheu, Heethoff unpublished; Figure 3.5). This proofs the functionality of both sequence clusters. Neofunctionalisation is unlikely due to the high amount of homology in signal sequences and structure of the protein.

A problem in distinguishing between alleles and paralogs remains due to two reasons: (i) both sequence clusters were present in one of the two sexual species (*M. pulverulenta*) and (ii) three sequences were found in *A. striculus* and *M. nasalis*.

(i) Finding two diverged copies in some sexual oribatid mite species was not surprising due to their genetic mechanism, holokinetic chromosomes and inverted meiosis (Wrensch et al. 1994), which prevents intrachromosomal recombination (see chapter 4). This is in accordance with finding just one copy in the other sexual species, *S. magnus*. Due to interchromosomal recombination a Hardy-Weinberg (Hardy 1908) distribution of alleles is expected among sexual oribatid mite species. Future studies have to evaluate if this is in fact the case. In recent studies (Domes, Scheu, Heethoff unpublished) both sequence clusters were found to be present also in *S. magnus* with allelic distances similar to those

reported here. Finding homozygous and heterozygous individuals in sexual oribatid mite species indicate linkage of the sequence clusters and that they represent alleles.

(ii) The presence of more than two copies in some specimens might be due to an ancient hybrid polyploid origin of the species which is unlikely since all investigated oribatid mites are diploid (Norton et al. 1993). Another possible explanation is the occurrence of somatic mutations, as postulated for ostracods (Schon and Martens 2003). This might explain the slight differences in the third sequence found in *A. striculus* but is unlikely for the three sequences found in *M. nasalis* because the differences here are more pronounced. One copy in *M. nasalis* (MNCB16) is identical to one of the alleles found in *T. velatus* (TVDOacon). A cross contamination is probably the most parsimonious explanation for this sequence.

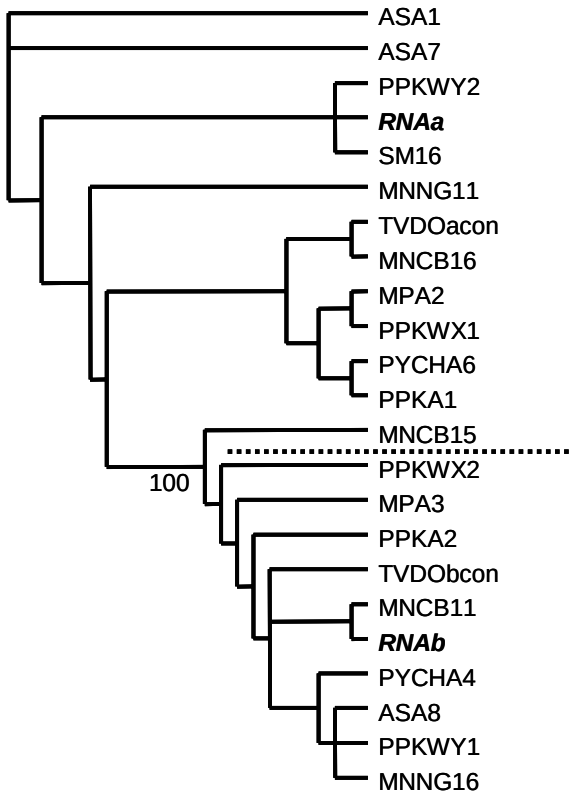


Figure 3.5
Phylogenetic analyses of *hsp82* mRNA. ML algorithm based on the TrN+G model (Tamura and Nei 1993).
The dashed line indicates the possible split of the two alleles. Numbers on branches indicate bootstrap values. cDNA sequences are represented in both sequence clusters (RNAa and RNAb).

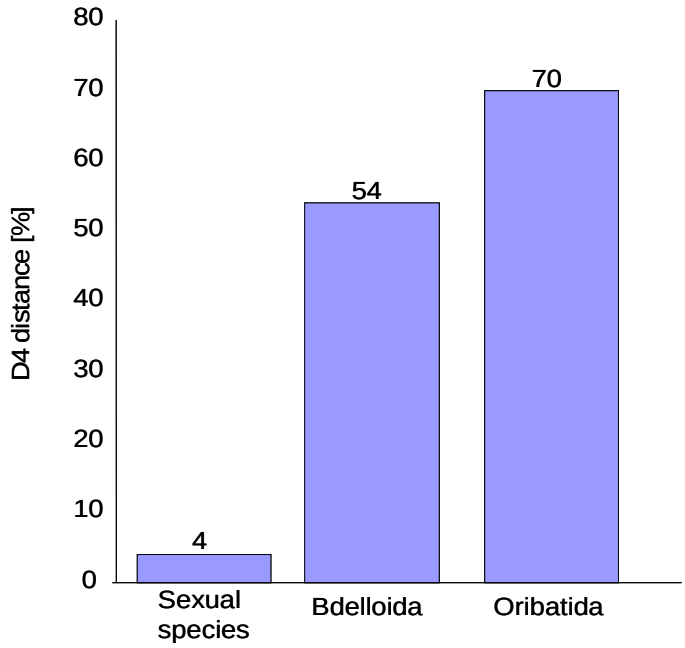


Figure 3.6
Camparisons of allelic divergences. Sexual species were analysed over a wide range of taxa, values for sexual species and bdelloid rotifers are from Mark Welsh and Meselson (2000).

In conclusion there is strong evidence that both sequence clusters represent in fact alleles of the *hsp82* gene although some further investigations are needed to unequivocally proof this assumption. The degree of divergence between both alleles is the highest known for allelic sequences (Figure 3.6). If both sequences represent alleles, this is an indication that the “Meselson effect” (Mark Welsh and Meselson 2000) acts in sexual and asexual oribatid mites. The absence of recombination during the evolution of oribatid mites, a taxon presumably comprising 100,000 species (Schatz 2002), contradicts evolutionary theory on the importance of sex and recombination in the long-term (Kondrashov 1993, West et al. 1999). Under certain genetic circumstances, such as inverted meiosis, evolution in the long-term presumably also is possible without recombination.

Appendix B

The following pages contain the sequence alignments of the alleles for different specimens. Shown are the two or three alleles found within specimens. Not all sequenced samples (>200) are shown, just one candidate from each allele within single specimens. The first block for each specimen constitutes of the DNA sequences, the second of amino acid sequences.

Platynothrus yamasakii (PYCH)

```
DNA          ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                 5         15         25         35         45         55         65
PYCHA.4      CAGTTTGGTG TGGGTTTTTA TTCCGCGTAT TTGATCGCCG ATAGAGTGAC CGTCACTTCA AAGCATAACG
PYCHA.6      CAGTTTGGTG TCGGTTTCTA TTCTGCGTAT TTGATCGCAG ATAAGGTCAC CGTTTACTCC AAACACAACG

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                 75         85         95        105        115        125        135
PYCHA.4      ACGACGAACA GTACGTTTGG GAGTCTTCTG CTGGCGGGTC ATTCACTATC AGAACTGACG CCACGGGTGA
PYCHA.6      ACGACGAACA GTACGTCTGG GGGTCGTCTG CTGGCGGGTC GTTCACTATA AAGCCAGACA ATACCGGA--

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                145        155        165        175        185        195        205
PYCHA.4      GCCTTTAGGT CGCGGAACTA AGATTGTATT GCACTTGAAA GAAGACCAGT TGGAATACGC GGAGGAAAAA
PYCHA.6      -CCTCTCGGC AGAGGCACCA AAATTATTCT TCACCTCAAA GAGGATCAGT TGGAGTACTC CGAAGAGAAA

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                215        225        235        245        255        265        275
PYCHA.4      CGCATTAAAG AAATTGTTAA AAAACACTCC CGATTCATCG GTTATCCCAT TAAACTATTG GTTCAGAAAG
PYCHA.6      CGTATAAAAG ATATCGTGAA AAAACACTCC CAATTCATCG GATATCCAAT CAAACTATTG GTCCAAAAGG

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                285        295        305        315        325        335        345
PYCHA.4      AGAGAGAAAA AGAAGTCTCT GATGACGAAG AAGACAAAGA GGAGGAGAAG ACTGAAGAAA AGAGTGAGGA
PYCHA.6      AGAGGGAAAA AGAGGTCTCA GACGACGAGG AAGAGAAAGA AGAGGACAAG AAAGAGGACG AAGAGAAGAA

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                355        365        375        385        395        405        415
PYCHA.4      AAATAAGACC GAAGAAGAGA AGATTGATGA AGACGAACCT AAAGTTGAAG ACGTCGAGGA CTCCGAAGAT
PYCHA.6      AGAGGACAAA ------GAGG GCGGAGATGA AGACGAACCC AAAGTAGAAG ACGTCGAAGA CTCTGACGAA

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                425        435        445        455        465        475        485
PYCHA.4      AAGAAAGACA AGAAAAAGAA GAAAAAAATT AAGGAAAAGT ATGTCGAAGA CGAAGAGCTG AATAAAACTA
PYCHA.6      AAGAAGGATA AGAAGAAGAA GAAGAAGATT AAGGAAAAGT ACGTCGAAGA CGAAGAGCTG AACAAAACGA

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|.
                495        505        515        525        535        545
PYCHA.4      AACCCATTTG GATGAGAAAT CCCGATGACA TCACTCAAGA AGAATACGGA GAATTC
PYCHA.6      AACCAATTTG GATGAGAAAT CCGGATGACA TCACTCAAGA AGAATACGGA GAATTC



Protein      ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                 5         15         25         35         45         55         65
PYCHA.4      QFGVGFYSAY LIADRVTVTS KHNDDEQYVW ESSAGGSFTI RTDATGEPLG RGTKIVLHLK EDQLEYAEEK
PYCHA.6      QFGVGFYSAY LIADKVTVYS KHNDDEQYVW GSSAGGSFTI KPDNTG-PLG RGTKIILHLK EDQLEYSEEK

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                 75         85         95        105        115        125        135
PYCHA.4      RIKEIVKKHS RFIGYPIKLL VQKEREKEVS DDEEDKEEEK TEEKSEENKT EEEKIDEDEP KVEDVEDSED
PYCHA.6      RIKDIVKKHS QFIGYPIKLL VQKEREKEVS DDEEEKEEDK KEDEEKKEDK --EGGDEDEP KVEDVEDSDE

             ....|....| ....|....| ....|....| ....|....| ..
                145        155        165        175
PYCHA.4      KKDKKKKKKI KEKYVEDEEL NKTKPIWMRN PDDITQEEYG EF
PYCHA.6      KKDKKKKKKI KEKYVEDEEL NKTKPIWMRN PDDITQEEYG EF
```

Platynothrus peltifer (PPKWX)

```
DNA          ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                 5          15         25         35         45         55         65
PPKWX1       CAGTTTGGTG TTGGTTTTTA CAGCGCATAC CTGATCGCCG ATAAGGTTGT GGTGACCTCT AAGCACAACG
PPKWX2       CAATTTGGTG TTGGTTTCTA TTCCGCGTAT TTGATCGCCG ATAGAGTGAC CGTCACTTCA AGGCATAACG

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                75         85         95        105        115        125        135
PPKWX1       ACGACGAGCA GTACGTGTGG GAGTCGTCGG CCGGCGGCTC GTTCACCATC CGGGCCGACA AC---ACCGA
PPKWX2       ACGACGAACA GTACGTTTGG GAGTCTTCTG CTGGCGGGTC ATTCACTATC AGAACTGACG CCACGGGCGA

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
               145        155        165        175        185        195        205
PPKWX1       GCCGTTGGGC AGAGGCACGA AGATTGTGCT GCaCCTGAAG GAAGACCAGT TGGAGTACGC GGAGGAGAAG
PPKWX2       GCCTTTAGGT CGCGGAACTA AGATTGTATT GCACTTGAAA GAAgACCAGT TGGAATACgC GGAGGAAAAA

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
               215        225        235        245        255        265        275
PPKWX1       CGCATCAGAG AGATCGTGAA GAAGCACTCG CAGTTCATCG GATACCCAAT CAAACTACTC GTGCAGAAGG
PPKWX2       CGCATTAAAG AAATTGTTAA AAAACACTCC CAATTCATCG GTTATCCCAT TAAACTATTG GTTCAGAAAG

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
               285        295        305        315        325        335        345
PPKWX1       AACGCGAGAA GGAGGTGTCC GACGACGAGG AGGAAGAGGC GAAG------ GACGAGAAGA AAGACGAGGA
PPKWX2       AGAGAGAAAA AGAAGTCTCT GATGACGAAG AAGACAAAGA GGAGGAGAAG ACTGAAGAAA AGAGTGAGGA

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
               355        365        375        385        395        405        415
PPKWX1       GAAGAAGGAG TCCGAA---G GCGGCGATGA GGACGAGCCG AAGGTCGAGG ACGCGGAGGA CTCGGACGAG
PPKWX2       AAATAAGACC GAAGAAGAGA AGATTGATGA AGACGAACCT AAAGTTGAAG ACGTCGAGGA CTCCGAAGAT

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
               425        435        445        455        465        475        485
PPKWX1       AAGAAAGACA AAAAGAAAAA GAAGAAAATA AAGGAGACGT ACGTCGAGGA CGAGGAGCTG AATAAGACTA
PPKWX2       AAGAAAGACG AGAAAAAGAA GAAAAAAATT AAGGAAAAGT ATGTCGAAGA CGGAGAGCTG AATAGAACTA

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|.
               495        505        515        525        535        545
PPKWX1       AGCCGTTATG GATGCGCAAC CCCGACGACA TCACTCAGGA AGAGTACGGC GAGTTT
PPKWX2       AACCCATTTG GATGAGAAAT CCCGATGACA TC-------- ----AACGGA GAATTC



Protein      ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                 5          15         25         35         45         55         65
PPKWX1       QFGVGFYSAY LIADKVVVTS KHNDDEQYVW ESSAGGSFTI RADN-TEPLG RGTKIVLHLK EDQLEYAEEK
PPKWX2       QFGVGFYSAY LIADRVTVTS RHNDDEQYVW ESSAGGSFTI RTDATGEPLG RGTKIVLHLK EDQLEYAEEK

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                75         85         95        105        115        125        135
PPKWX1       RIREIVKKHS QFIGYPIKLL VQKEREKEVS DDEEEEAK-- DEKKDEEKKE SE-GGDEDEP KVEDAEDSDE
PPKWX2       RIKEIVKKHS QFIGYPIKLL VQKEREKEVS DDEEDKEEEK TEEKSEENKT EEEKIDEDEP KVEDVEDSED

             ....|....| ....|....| ....|....| ....|....| ..
               145        155        165        175
PPKWX1       KKDKKKKKKI KETYVEDEEL NKTKPLWMRN PDDITQEEYG EF
PPKWX2       KKDEKKKKKI KEKYVEDGEL NRTKPIWMRN PDDI----NG EF
```

Platynothrus peltifer (PPKWY)

```
DNA          ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                 5         15         25         35         45         55         65
PPKWY.2      CAGTTTGGTG TAGGTTTCTA TTCCGCGTAT TTGATCGCCG ATAGAGTGAC CGTCACTTCA AAGCATAACG
PPKWY.14     CAGTTCGGTG TGGGTTTCTA TTCTGCGTAT TTGATTGCAG ATCGAGTGGT GGTTCACTCG AAGCACAACG

                 ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                     75         85         95        105        115        125        135
PPKWY.2      ACGACGAACA GTACGTTTGG GAGTCTTCTG CTGGCGGGTC ATTCACTATC AGAACTGACG CCACGGGTGA
PPKWY.14     ACCACGAGCA GTACGTCTGG GAGTCGGCAG CCGGCGGTTC GTTCACTATT CGTGTGGAC- --ACTGGCGA

                 ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                    145        155        165        175        185        195        205
PPKWY.2      GCCTTTAGGT CGCGGAACTA AGATTGTACT GCACTTGAAA GAAGACCAGT TGGAATACGC GGAGGAAAAA
PPKWY.14     GTCTTTGGGT CGCGGAACTA AGATTGTGCT CCATTTGAAA GAGGATCAGT TGGATTACAC TGAGGAGAGA

                 ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                    215        225        235        245        255        265        275
PPKWY.2      CGCATTAAAG AAATTGTTAA AAAACACTCC CGATTCATCG GTTATCCCAT TAAACTATTG GTTCAGAAAG
PPKWY.14     CGCATCAAAG ATATCGTTAA AAAGCACTCG CAGTTCATCG GGTATCCCAT CAAACTCGTG GTTCAAAAGG

                 ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                    285        295        305        315        325        335        345
PPKWY.2      AGAGAGAAAA AGAAGTCTCT GATGACGAAC AAGACAAAGA GGAGGAGAAG ACTGAAGAAA AGAGTGAGGA
PPKWY.14     AAAGAGAGAA AGAGATCTCT GATGACGAAG AAGAGAAGGA AGAGGAGAAA AAAGATGAAA CCGAGGAAAA

                 ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                    355        365        375        385        395        405        415
PPKWY.2      AAATAAGACC GAAGAAGAGA AGATTGATGA AGACGAACCT AAAGTTGAAG ACGTCGAGGA CTCCGAAGAT
PPKWY.14     GGAGAAAACC GAAGAGAAT- --------GA AGACCAACCG AAAGTCGAGG ACGTGGAGGA CTCGGAAGAC

                 ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                    425        435        445        455        465        475        485
PPKWY.2      AAGAAAGACA AGAAAAAGAA GAAAAAAATT AAGGAAAAGT ATGTCGAAGA CGAAGAGCTG AATAAAACTA
PPKWY.14     AAGAAAGACA AAAAGAAAAA GAAGAAAATA AAGGAAAAGT ATGTGGAGGA CGAAGAATTG AACAAAACGA

                 ....|....| ....|....| ....|....| ....|....| ....|....| ....|.
                    495        505        515        525        535        545
PPKWY.2      AACCCATTTG GATGAGAAAT CCCGATGACA TCACTCAAGA AGAATACGGA GAATTC
PPKWY.14     AACCAATTTG GATGCGAAAT CCCGATGACA TCACTCAAGA AGAGTACGGC GAGTTC



Protein      ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                 5         15         25         35         45         55         65
PPKWY.2      QFGVGFYSAY LIADRVTVTS KHNDDEQYVW ESSAGGSFTI RTDATGEPLG RGTKIVLHLK EDQLEYAEEK
PPKWY.14     QFGVGFYSAY LIADRVVVHS KHNDHEQYVW ESAAGGSFTI RVD-TGESLG RGTKIVLHLK EDQLDYTEER

                 ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                     75         85         95        105        115        125        135
PPKWY.2      RIKEIVKKHS RFIGYPIKLL VQKEREKEVS DDEQDKEEEK TEEKSEENKT EEEKIDEDEP KVEDVEDSED
PPKWY.14     RIKDIVKKHS QFIGYPIKLV VQKEREKEIS DDEEEKEEEK KDETEEKEKT EEN---EDQP KVEDVEDSED

                 ....|....| ....|....| ....|....| ....|....| ..
                    145        155        165        175
PPKWY.2      KKDKKKKKKI KEKYVEDEEL NKTKPIWMRN PDDITQEEYG EF
PPKWY.14     KKDKKKKKKI KEKYVEDEEL NKTKPIWMRN PDDITQEEYG EF
```

Platynothrus peltifer (PPKA)

```
DNA          ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                 5         15         25         35         45         55         65
PPKA2        CAGTTCGGTG TGGGTTTTTA TTCCGCGTAT TTGATCGCCG ATAGAGTGAC CGTCACTTCA AAGCATAACG
PPKA.1       CAGTTCGGTG TTGGTTTCTA TTCTGCGTAT TTGATCGCAG ATAAGGTCAC CGTTTACTCC AAACACAACG

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                75         85         95        105        115        125        135
PPKA2        ACGACGAACA GTACGTTTGG GAGTCTTCTG CTGGCGGGTC ATTCACTATC AGAACTGACG CCACGGGTGA
PPKA.1       ACGACGAACA GTACGTCTGG GAGTCGTCTG CTGGCGGGTC GTTCACTATA AAGCCAGACA ATACC---GA

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
               145        155        165        175        185        195        205
PPKA2        GCCTTTAGGT CGCGGAACTA AGATTGTATT GCACTTGAAA GAAGACCAGT TGGAATACGC GGAGGAAAAA
PPKA.1       ACCTCTCGGC AGAGGCACCA AAATTATTCT TCACCTCAAA GAGGATCAGT TGGAGTACTC CGAAGAGAAA

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
               215        225        235        245        255        265        275
PPKA2        CGCATTAAAG AAATTGTTAA AAAACACTCC CAATTCATCG GTTATCCCAT TAAACTATTG GTTCAGAAAG
PPKA.1       CGTATAAAAG ATATCGTGAA AAAACACTCC CAATTCATCG GATATCCAAT CAAACTATTG GTCCAAAAGG

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
               285        295        305        315        325        335        345
PPKA2        AGAGAGAAAA AGAAGTCTCT GATGACGAAG AAGACAAAGA GGAAGAGAAG ACTGAAGAAA AGAGTGAGGA
PPKA.1       AGAGGGAAAA AGAGGTCTCA GACGACGAGG AAGAGAAAGA AGAGGACAAG AAAGAG---- --GACGAAGA

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
               355        365        375        385        395        405        415
PPKA2        AAATAAGACC GAAGAAGAGA AGATTGATGA AGACGAACCT AAAGTTGAAG ACGTCGAGGA CTCCGAAGAT
PPKA.1       GAAGAAAGAG GACAAAGAGG GCGGAGATGA AGACGAACCC AAAGTAGAAG ACGTCGAAGA CTCTGACGAA

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
               425        435        445        455        465        475        485
PPKA2        AAGAAAGACA AGAAAAAGAA GAAAAAAATT AAGGAAAAGT ATGTCGAAGA CGAAGAGCTG AATAAAACTA
PPKA.1       AAGAAGGATA AGAAGAAGAA GAAGAAGATT AAGGAAAAGT ACGTCGAAGA CGAAGAGCTG AACAAAACGA

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|.
               495        505        515        525        535        545
PPKA2        AACCCATTTG GATGAGAAAT CCCGATGACA TCACTCAAGA AGAATACGGA GAATTC
PPKA.1       AACCAATTTG GATGAGAAAT CCGGATGACA TCACTCAAGA AGAATACGGA GAATTC



Protein      ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                 5         15         25         35         45         55         65
PPKA2        QFGVGFYSAY LIADRVTVTS KHNDDEQYVW ESSAGGSFTI RTDATGEPLG RGTKIVLHLK EDQLEYAEEK
PPKA.1       QFGVGFYSAY LIADKVTVYS KHNDDEQYVW ESSAGGSFTI KPDNT-EPLG RGTKIILHLK EDQLEYSEEK

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                75         85         95        105        115        125        135
PPKA2        RIKEIVKKHS QFIGYPIKLL VQKEREKEVS DDEEDKEEEK TEEKSEENKT EEEKIDEDEP KVEDVEDSED
PPKA.1       RIKDIVKKHS QFIGYPIKLL VQKEREKEVS DDEEEKEEDK KE--DEEKKE DKEGGDEDEP KVEDVEDSDE

             ....|....| ....|....| ....|....| ....|....| ..
               145        155        165        175
PPKA2        KKDKKKKKKI KEKYVEDEEL NKTKPIWMRN PDDITQEEYG EF
PPKA.1       KKDKKKKKKI KEKYVEDEEL NKTKPIWMRN PDDITQEEYG EF
```

Mucronothrus nasalis (MNNG)

```
DNA          ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                  5         15         25         35         45         55         65
MNNG1.1      CAGTTCGGTG TTGGTTTCTA TTCGGCGTAT TTGATCGCAG ACAGAGTGAC CGTCACTTCT AAGCATAACG
MNNG1.6      CAGTTTGGTG TCGGTTTTTA TTCCGCGTAT TTGATCGCCG ATAGAGTGAC CGTCACTTCA AAGCATAACG

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                 75         85         95        105        115        125        135
MNNG1.1      ACGACGAACA GTATGTCTGG GAGTCATCGG CCGGCGGCTC CTTCACTATC AGGACAGATA ATAGC---GA
MNNG1.6      ACGACGAACA GTACGTTTGG GAGTCTTCTG CTGGCGGGTC ATTCACTATC AGAACTGACG CCACGGGTGA

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                145        155        165        175        185        195        205
MNNG1.1      ACCATTAGGT CGAGGCACTA AAATTGTCTT ACTTCTCAAA GAAGACCAAT TAGAATACGC AGAAGAAAAA
MNNG1.6      GCCTTTAGGT CGCGGAACTA AGATTGTACT GCACTTGAAA GAAGACCAGT TGGAATACGC GGAGGAAAAA

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                215        225        235        245        255        265        275
MNNG1.1      CGTATTAAAG AGATCGTGAA AAAACACTCG CAATTCATCG GATATCCGAT CAAACTTGTC GTTCAAAAGG
MNNG1.6      CGCATTAAAG AAATTGTTAA AAAACACTCC CGATTCATCG GTTATCCCAT TAAACTATTG GTTCAGAAAG

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                285        295        305        315        325        335        345
MNNG1.1      AAAGAGAAAA AGAGATCTCA GACGACGAGG AAGACAAAGA AGAAGCCAAA GAAGATAAAG AAGACAAGAT
MNNG1.6      AGAGAGAAAA AGAAGTCTCT GATGACGAAG AAGACAAAGA GGAGGAGAAA ACT------G AAGAAAAGAG

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                355        365        375        385        395        405        415
MNNG1.1      CGAAGATGAA GACAAGACCG AAGAGAAGAA AGAAGAGGGC GACGAGCCTA AGGTCGAAGA CGTCGAGGAC
MNNG1.6      TGAGGAAAAT AAGACCGAAG AAGAGAAGAT TGATGAA--- GACGAACCTA AAGTTGAAGA CGTCGAGGAC

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                425        435        445        455        465        475        485
MNNG1.1      TCTGAAGATA AGAAAGACAA GAAAAAGAAG AAA---ATTA AGGAAAAGTA TGTCGAAGAC GAAGAACTAA
MNNG1.6      TCCGAAGATA AGAAAGACAA GAAAAAGAAG AAAAAAATTA AGGAAAAGTA TGTCGAAGAC GAAGAGCTGA

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|
                495        505        515        525        535        545        555
MNNG1.1      ACAAAACAAA ACCAATTTGG ATGAGAAATC CAGACGATAT CACTCAAGAA GAATACGGAG AATTC
MNNG1.6      ATAAAACTAA ACCCATTTGG ATGAGAAATC CCGATGACAT CACTCAAGAA GAATACGGAG AATTC



Protein      ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                  5         15         25         35         45         55         65
MNNG1.1      QFGVGFYSAY LIADRVTVTS KHNDDEQYVW ESSAGGSFTI RTDNS-EPLG RGTKIVLLLK EDQLEYAEEK
MNNG1.6      QFGVGFYSAY LIADRVTVTS KHNDDEQYVW ESSAGGSFTI RTDATGEPLG RGTKIVLHLK EDQLEYAEEK

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                 75         85         95        105        115        125        135
MNNG1.1      RIKEIVKKHS QFIGYPIKLV VQKEREKEIS DDEEDKEEAK EDKEDKIEDE DKTEEKKEEG DEPKVEDVED
MNNG1.6      RIKEIVKKHS RFIGYPIKLL VQKEREKEVS DDEEDKEEEK T--EEKSEEN KTEEEKIDE- DEPKVEDVED

             ....|....| ....|....| ....|....| ....|....| ....|
                145        155        165        175        185
MNNG1.1      SEDKKDKKKK K-IKEKYVED EELNKTKPIW MRNPDDITQE EYGEF
MNNG1.6      SEDKKDKKKK KKIKEKYVED EELNKTKPIW MRNPDDITQE EYGEF
```

Mucronothrus nasalis (MNCB)

```
DNA          ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                 5         15         25         35         45         55         65
MNCB1.1      CAGTTTGGTG TGGGTTTCTA TTCCGCGTAT TTGATCGCCG ATAGAGTGAC CGTCACTTCA AAGCATAATG
MNCB1.5      CAGTTCGGTG TTGGTTTCTA TTCCGCGTAT TTGATCGCCG ATAGAGTGAC CGTCACTTCA AAGCATAATG
MNCB1.6      CAGTTCGGTG TGGGTTTCTA TTCCGCGTAT TTGATCGCCG ATAAGGTGAC GGTTCACTCA AAACACAACG

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                 75         85         95        105        115        125        135
MNCB1.1      ACGACGAACA GTACGTTTGG GAGTCTTCTG CTGGCGGGTC ACTCACTATC AGAACTGACG CCACGGGTGA
MNCB1.5      ACGACGAACA GTGCGTTTGG GAGTCTTCTG CTGGCGGGTC ATTCACTATC AGAACTGACG CCACGGGCGA
MNCB1.6      ACGACGAACA GTACGTCTGG GAGTCATCTG CCGGCGGATC CTTCACTATA AAGCCAGACG TCGAGGGCGA

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                145        155        165        175        185        195        205
MNCB1.1      GCCTTTAGGT CGCGGAACTA AGATTGTATT GCACTTGAAA GAAGACCAGT TGGAATACGC GGAGGAAAAA
MNCB1.5      GCCTTTAGGT CGCGGAACTG AGATTGTATT GCACTTGAAA GAAGACCAGT TGGAATACGC GGAGGAAAAA
MNCB1.6      ACCTCTCGGC AGAGGCACCA AAATCATCCT TCACTTGAAA GAGGATCAGT TGGGGTACTC CGAAGAGAAG

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                215        225        235        245        255        265        275
MNCB1.1      CGCATTAAAG ATATTGTTAA AAAACACTCC CAATTCATCG GTTATCCCAT TAAACTCTTG GTTCAGAAAG
MNCB1.5      CGCATTAAAG AAATTGTTAA AAAACACTCC CAATTCATCG GTTATCCCAT TAAACTATTG GTTCAGAAAG
MNCB1.6      CGCATCAAAG ATATCGTGAA AAAACACTCA CAATTCATCG GATATCCCAT CAAACTATTG GTTCAAAAGG

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                285        295        305        315        325        335        345
MNCB1.1      AGAGAGAAAA AGAAGTCTCT GATGACGGAG AAGACAAAGA GGAGGAGAAG ACTGAAGAAA AGAGTGAGGA
MNCB1.5      AGAGAGAAAA AGAAGTCTCT GTTGACGAAC AAGACAAAGA GGAGGAGAAG ACTGAAGAAA AGAGTGAGGA
MNCB1.6      AAAGAGAAAA AGAGGTCTCC GATGACGAGG AGGACAAAGA AGAAGATAAG AAAGAAGAC- -----GACGA

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                355        365        375        385        395        405        415
MNCB1.1      AAATAAGACC GAAGAAGAGA AGATTGATGA AGACGAACCT AAAGTTGAAG ACGTCGAGGA CTCCGAAGAT
MNCB1.5      AAATAAGACC GAAGAAGAGA AGATTGATGA AGACGAACCT AAAGTTGAAG ACGTCGAGGA CTCCGAAGAT
MNCB1.6      AAAGAAAGAA GAGAAAGAGG GCGGAGACGA AGACGAACCC AAAGTTGAAG ACGTCGAGGA TTCAGAGGAA

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                425        435        445        455        465        475        485
MNCB1.1      AAGAAAGACA AGAAAGAGAA GAAAAAAATT AAGGAAAAGT ATGTCGAAGA CGAAGAGCTC AATAAAACTA
MNCB1.5      AAGAAAGACA AGAAGAAGAA GAAAAAAATT CAGGAAAAGT ATGTCGAAGA CGAAGAGCTG AATAAAACTA
MNCB1.6      AAGAAAGACA AAAAGAAGAA GAAGAAAATT AAGGAAAAGT ATGTCGAAGA CGAAGAACTT AACAAAACTA

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|.
                495        505        515        525        535        545
MNCB1.1      AACCCATTTG GATGAGAAAT CCCGATGACA TCACTCAAGA AGAATACGGC GAATTC
MNCB1.5      AACCCATTTG GATGAGAAAT CCCGATGACA TCACTCAAGA AGAATACGGA GAATTC
MNCB1.6      AACCCATTTG GATGAGAAAT CCCGATGACA TCACTCAAGA AGAATACGGA GAATTC




Protein      ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                 5         15         25         35         45         55         65
MNCB1.1      QFGVGFYSAY LIADRVTVTS KHNDDEQYVW ESSAGGSLTI RTDATGEPLG RGTKIVLHLK EDQLEYAEEK
MNCB1.5      QFGVGFYSAY LIADRVTVTS KHNDDEQCVW ESSAGGSFTI RTDATGEPLG RGTEIVLHLK EDQLEYAEEK
MNCB1.6      QFGVGFYSAY LIADKVTVHS KHNDDEQYVW ESSAGGSFTI KPDVEGEPLG RGTKIILHLK EDQLGYSEEK

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                 75         85         95        105        115        125        135
MNCB1.1      RIKDIVKKHS QFIGYPIKLL VQKEREKEVS DDGEDKEEEK TEEKSEENKT EEEKIDEDEP KVEDVEDSED
MNCB1.5      RIKEIVKKHS QFIGYPIKLL VQKEREKEVS VDEQDKEEEK TEEKSEENKT EEEKIDEDEP KVEDVEDSED
MNCB1.6      RIKDIVKKHS QFIGYPIKLL VQKEREKEVS DDEEDKEEDK KED--DEKKE EKEGGDEDEP KVEDVEDSEE

             ....|....| ....|....| ....|....| ....|....| ..
                145        155        165        175
MNCB1.1      KKDKKEKKKI KEKYVEDEEL NKTKPIWMRN PDDITQEEYG EF
MNCB1.5      KKDKKKKKKI QEKYVEDEEL NKTKPIWMRN PDDITQEEYG EF
MNCB1.6      KKDKKKKKKI KEKYVEDEEL NKTKPIWMRN PDDITQEEYG EF
```

Atropacarus striculus (ASA)

```
DNA          ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                 5         15         25         35         45         55         65
ASA1         CAGTTCGGTG TGGGTTTTTA TTCGGCATAT TTGGTTGCGG ACAGAGTTGT GGTTCACTCG AAACACAACG
ASA7         CAGTTTGGTG TCGGTTTCTA TTCGGCATAT TTGGTTGCGG ACAGAGTTGT GGTTCACTCG AAACACAACG
ASA8         CAGTTTGGTG TGGGTTTTTA TTCCGCGTAT TTGATCGCCG ATAGAGTGAC CGTCACTTCA AAGCATAACG

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                 75         85         95        105        115        125        135
ASA1         ACGACGAGCA GTACGTGTGG GAGTCTTCGG CCGGCGGTTC GTTCACCATT CGCGTCGAT- --AGCGGAGA
ASA7         ACGACGAGCA GTACGTGTGG GAGTCTTCGG CCGGCGGTTC GTTCACCATT CGCGTCGAT- --AGCGGAGA
ASA8         ACGACGAACA GTACGTTTGG GAGTCTTCTG CTGGCGGGTC ATTCACTATC AGAACTGACG CCACGGGTGA

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                145        155        165        175        185        195        205
ASA1         ATCTTTGGGT CGCGGAACCA AAATAATCCT TTTTTTGAAA GAAGATCAGT TGGATTACAC TGAGGAAAGA
ASA7         ATCTTTGGGT CGCGGAACCA AAATAATCCT TTTTTTGAAA GAAGATCAGT TGGATTACAC TGAGGAAAGA
ASA8         GCCTTTAGGT CGCGGAACTA AGATTGTACT GCACTTGAAA GAAGACCAGT TGGAATACGC GGAGGAAAAA

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                215        225        235        245        255        265        275
ASA1         CGTATCAAAG ATATCGTTAA AAAGCATTCG CAATTCATTG GATATCCGAT TAAGCTTTTG GTACAAAAGG
ASA7         CGTATCAAAG ATATCGTTAA AAAGCATTCG CAATTCATTG GATATCCGAT TAAGCTTTTG GTACAAAAGG
ASA8         CGCATTAAAG AAATTGTTAA AAAACACTCC CGATTCATCG GTTATCCCAT TAAACTATTG GTTCAGAAAG

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                285        295        305        315        325        335        345
ASA1         AGAGAGAAAA AGAGATCTCT GATGACGAGG AAGAGAAGGA AGAAGAG--- ---------A AAAAGGATGA
ASA7         AGAGAGAAAA AGAGATCTCT GATGACGAGG AAGAGAAGGA A---GAGAAG GAAGAAGAGA AAAAGGATGA
ASA8         AGAGAGAAAA AGAAGTCTCT GATGACGAAG AAGACAAAGA GGAGGAGAAG ACTGAAGAAA AGAGTGAGGA

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                355        365        375        385        395        405        415
ASA1         A------ACT GAGGAGAAAG AGAAGACCGA AGAGAATGAA GACGAGCCGA AAGTAGAGGA CGTGGAAGAC
ASA7         A------ACT GAGGAGAAAG AGAAGACCGA AGAGAATGAA GACGAGCCGA AAGTAGAGGA CGTGGAAGAC
ASA8         AAATAAGACC GAAGAA---- -----GAGAA GATTGATGAA GACGAACCTA AAGTTGAAGA CGTCGAGGAC

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                425        435        445        455        465        475        485
ASA1         TCTGAGGATA AGAAAGACAA AAAGAAAAAG AAGAAAATAA AGGAAAAGTA TGTCGAAGAC GAAGAACTGA
ASA7         TCTGAGGATA AGAAAGACAA AAAGAAAAAG AAGAAAATAA AGGAAAAGTA TGTCGAAGAC GAAGAACTGA
ASA8         TCCGAAGATA AGAAAGACAA GAAAAAGAAG AAAAAAATTA AGGAAAAGTA TGTCGAAGAC GAAGAGCTGA

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|
                495        505        515        525        535        545        555
ASA1         ACAAAACGAA ACCGATTTGG ATGCGAAACC CCGATGACAT CACTCAAGAA GAGTATGGAG AATTC
ASA7         ACAAAACGAA ACCGATTTGG ATGCGAAACC CCGATGACAT CACTCAAGAA GAGTATGGAG AATTC
ASA8         ATAAAACTAA ACCCATTTGG ATGAGAAATC CCGATGACAT CACTCAAGAA GAATACGGAG AATTC



Protein      ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                 5         15         25         35         45         55         65
ASA1         QFGVGFYSAY LVADRVVVHS KHNDDEQYVW ESSAGGSFTI RVD-SGESLG RGTKIILFLK EDQLDYTEER
ASA7         QFGVGFYSAY LVADRVVVHS KHNDDEQYVW ESSAGGSFTI RVD-SGESLG RGTKIILFLK EDQLDYTEER
ASA8         QFGVGFYSAY LIADRVTVTS KHNDDEQYVW ESSAGGSFTI RTDATGEPLG RGTKIVLHLK EDQLEYAEEK

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                 75         85         95        105        115        125        135
ASA1         RIKDIVKKHS QFIGYPIKLL VQKEREKEIS DDEEEKEEE- ---KKDE--T EEKEKTEENE DEPKVEDVED
ASA7         RIKDIVKKHS QFIGYPIKLL VQKEREKEIS DDEEEKE-EK EEEKKDE--T EEKEKTEENE DEPKVEDVED
ASA8         RIKEIVKKHS RFIGYPIKLL VQKEREKEVS DDEEDKEEEK TEEKSEENKT EE---EKIDE DEPKVEDVED

             ....|....| ....|....| ....|....| ....|....| ....|
                145        155        165        175        185
ASA1         SEDKKDKKKK KKIKEKYVED EELNKTKPIW MRNPDDITQE EYGEF
ASA7         SEDKKDKKKK KKIKEKYVED EELNKTKPIW MRNPDDITQE EYGEF
ASA8         SEDKKDKKKK KKIKEKYVED EELNKTKPIW MRNPDDITQE EYGEF
```

Tectocepheus velatus (TVDO)

```
DNA          ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                 5         15         25         35         45         55         65
TVDOa        CAGTTCGGTG TGGGTTTCTA TTCCGCGTAT TTGATCGCCG ATAAGGTGAC GGTTCACTCA AAACACAACG
TVDOb        CAGTTTGGTG TGGGTTTTTA TTCCGCGTAT TTGATCGCCG ATAGAGTGAC CGTCACTTCA AAGCATAACG

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                 75         85         95        105        115        125        135
TVDOa        ACGACGAACA GTACGTCTGG GAGTCATCTG CCGGCGGGTC GTTCACTATA AAGCCAGACG TCGAGGGCGA
TVDOb        ACGACGAACA GTACGTTTGG GAGTCTTCTG CTGGCGGGTC ATTCACTATC AGAACTGACG CCACGGGTGA

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                145        155        165        175        185        195        205
TVDOa        ACCTCTCGGC AGAGGCACCA AAATCATCCT TCACTTGAAA GAGGATCAGT TGGAGTACTC CGAAGAGAAA
TVDOb        GCCTTTAGGT CGCGGAACTA AGATTGTATT GCACTTGAAA GAAGACCAGT TGGAATACGC GGAGGAAAAA

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                215        225        235        245        255        265        275
TVDOa        CGCATCAAAG ATGTCGTGAA AAGACACTCA CAATTCATCG GATATCCCAT CAAACTATTG GTTCAAAAGG
TVDOb        CGCATTAAAG AAATTGTTAA AAAACACTCC CAATTCATCG GTTATCCCAT TAAACTATTG GTTCAGAAAG

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                285        295        305        315        325        335        345
TVDOa        AAAGAGAAAA AGAGGTCTCC GATGACGAGG AGGACAAAGA AGAAGATAAG AAAGAAGCC- -----GCCGA
TVDOb        AGAGAGAAAA AGAAGTCTCT GATGACGAAG AAGACAAAGA GGAGGAGAAG ACTGAAGAAA AGAGTGAGGA

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                355        365        375        385        395        405        415
TVDOa        AAAGAAAGAA GAGAAAGAGG GCGGAGACGA AGACGAACCC AAAGTTGAAG ACGTCGAGGA TTCAGAGGAA
TVDOb        AAATAAGACC GAAGAAGAGA AGATTGATGA AGACGAACCT AAAGTTGAAG ACGTCGAGGA CTCCGAAGAT

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                425        435        445        455        465        475        485
TVDOa        AAGAAAGACA AAAAGAAAAA GAAGAAAATT AAGGAAAAGT ATGTCGAAGA CGAAGAACTT AACAAAACTA
TVDOb        AAGAAAGACA AGAAAAAGAA GAAAAAAATT AAGGAAAAGT ATGTCGAAGA CGAAGAGCTG AATAAAACTA

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|.
                495        505        515        525        535        545
TVDOa        AACCCATTTG GATGAGAAAT CCCGATGACA TCACTCAAGA AGAATACGGA GAATTC
TVDOb        AACCCATTTG GATGAGAAAT CCCGATGACA TCACTCAAGA AGAATACGGA GAATTC



Protein      ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                 5         15         25         35         45         55         65
TVDOa        QFGVGFYSAY LIADKVTVHS KHNDDEQYVW ESSAGGSFTI KPDVEGEPLG RGTKIILHLK EDQLEYSEEK
TVDOb        QFGVGFYSAY LIADRVTVTS KHNDDEQYVW ESSAGGSFTI RTDATGEPLG RGTKIVLHLK EDQLEYAEEK

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                 75         85         95        105        115        125        135
TVDOa        RIKDVVKRHS QFIGYPIKLL VQKEREKEVS DDEEDKEEDK KEA--AEKKE EKEGGDEDEP KVEDVEDSEE
TVDOb        RIKEIVKKHS QFIGYPIKLL VQKEREKEVS DDEEDKEEEK TEEKSEENKT EEEKIDEDEP KVEDVEDSED

             ....|....| ....|....| ....|....| ....|....| ..
                145        155        165        175
TVDOa        KKDKKKKKKI KEKYVEDEEL NKTKPIWMRN PDDITQEEYG EF
TVDOb        KKDKKKKKKI KEKYVEDEEL NKTKPIWMRN PDDITQEEYG EF
```

Steganacarus magnus (SM)

```
DNA          ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                 5         15         25         35         45         55         65
SM1.6        CAATTCGGTG TAGGTTTTTA TTCTGCGTAT TTGATTGCAG ATCGAGTGGT GGTTCACTCG AAGCACAACG
SM1.2        CAATTCGGTG TAGGTTTTTA TTCTGCGTAT TTGATTGCAG ATCGAGTGGT GGTTCACTCG AAGCACAACG

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                 75         85         95        105        115        125        135
SM1.6        ACGACGAGCA GTACGTCTGG GAGTCGGCAG CCGGCGGTTC GTTCACTATT CGTGTGGACA CTGGCGAGTC
SM1.2        ACGACGAGCA GTACGTCTGG GAGTCGGCAG CCGGCGGTTC GTTCACTATT CGTGTGGACA CTGGCGAGTC

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                145        155        165        175        185        195        205
SM1.6        TTTGGGTCGC GGAACTAAGA TAGTGCTCCA TTTGAAAGAG GATCAGTTGG ATTACACTGA GGAGAGACGC
SM1.2        TTTGGGTCGC GGAACTAAGA TAGTGCTCCA TTTGAAAGAG GATCAGTTGG ATTACACTGA GGAGAGACGC

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                215        225        235        245        255        265        275
SM1.6        ATCAAAGATA TCGTTAAAAA GCACTCGCAG TTCATCGGGT ATCCCATCAA ACTCGTGGTT CAAAAGGAAA
SM1.2        ATCAAAGATA TCGTTAAAAA GCACTCGCAG TTCATCGGGT ATCCCATCAA ACTCGTGGTT CAAAAGGAAA

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                285        295        305        315        325        335        345
SM1.6        GAGAGAAAGA GATCTCTGAT GACGAAGAAG AGAAGGAAGA GGAGAAAAAA GATGAAACCG AGGAAAAGGA
SM1.2        GAGAGAAAGA GATCTCTGAT GACGAAGAAG AGAAGGAAGA GGAGAAAAAA GATGAAACCG AGGAAAAGGA

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                355        365        375        385        395        405        415
SM1.6        GAAAACCGAA GAGAATGAAG ACGAACCGAA AGTCGAGGAC GTGGAGGACT CGGAAGACAA GAAAGACAAA
SM1.2        GAAAACCGAA GAGAATGAAG ACGAACCGAA AGTCGAGGAC GTGGAGGACT CGGAAGACAA GAAAGACAAA

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                425        435        445        455        465        475        485
SM1.6        AAGAAAAAGA AGAAAATAAA GGAAAAGTAT GTGGAGGACG AAGAATTGAA CAAAACGAAA CCAATTTGGA
SM1.2        AAGAAAAAGA AGAAAATAAA GGAAAAGTAT GTGGAGGACG AAGAATTGAA CAAAACGAAA CCAATTTGGA

             ....|....| ....|....| ....|....| ....|....| ....
                495        505        515        525
SM1.6        TGCGAAATCC CGATGACATC ACTCAAGAAG AGTACGGCGA GTTC
SM1.2        TGCGAAATCC CGATGACATC ACTCAAGAAG AGTACGGCGA GTTC



Protein      ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                 5         15         25         35         45         55         65
SM16         QFGVGFYSAY LIADRVVVHS KHNDDEQYVW ESAAGGSFTI RVDTGESLGR GTKIVLHLKE DQLDYTEERR
SM1.6        QFGVGFYSAY LIADRVVVHS KHNDDEQYVW ESAAGGSFTI RVDTGESLGR GTKIVLHLKE DQLDYTEERR

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                 75         85         95        105        115        125        135
SM16         IKDIVKKHSQ FIGYPIKLVV QKEREKEISD DEEEKEEEKK DETEEKEKTE ENEDEPKVED VEDSEDKKDK
SM1.6        IKDIVKKHSQ FIGYPIKLVV QKEREKEISD DEEEKEEEKK DETEEKEKTE ENEDEPKVED VEDSEDKKDK

             ....|....| ....|....| ....|....| ....|...
                145        155        165        175
SM16         KKKKKIKEKY VEDEELNKTK PIWMRNPDDI TQEEYGEF
SM1.6        KKKKKIKEKY VEDEELNKTK PIWMRNPDDI TQEEYGEF
```

Metabelba pulverulenta (MPA)

```
DNA          ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                 5          15         25         35         45         55         65
MPA2         GAGTTCGGTG TTGGTTTTTA CAGCGCATAC CTGATCGCCG ATAAGGTTGT GGTGACCTCT AAGCACAACG
MPA3         GAGTTCGGTG TTGGTTTTTA TTCCGCGTAT TTGATCGCCG ATAGAGTGAC CGTCACTTCA AAGCATAACG

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                 75         85         95         105        115        125        135
MPA2         ACGACGAGCA GTACGTGTGG GAGTCGTCGG CCGGCGGCTC GTTCACCATC CGGGCCGACA ACACC---GA
MPA3         ACGACGAACA GTACGTTTGG GAGTCTTCTG CTGGCGGGTC ATTCACTATC AGAACTGACG CCACGGGTGA

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                145        155        165        175        185        195        205
MPA2         GCCGTTGGGC AGAGGCACGA AGATTGTGCT GCACCTGAAG GAAGACCAGT TGGAGTACGC GGAGGAGAAG
MPA3         GCCTTTAGGT CGCGGAACTA AGATTGTATT GCACTTGAAA GAAGACCAAT TGGAATACGC GGAGGAAAAA

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                215        225        235        245        255        265        275
MPA2         CGCATCAAAG AGATCGTGAA GAAGCACTCG CAGTTCATCG GATACCCAAT CAAACTACTC GTGCAGAAGG
MPA3         CGCATTAAAG AAATTGTTAA GAAACACTCC CAATTCATCG GTTATCCCAT TAAACTATTG GTTCAGAAAG

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                285        295        305        315        325        335        345
MPA2         AACGCGAGAA GGAGGTGTCC GACGACGAGG AGGAAGAGGC G------AAG GACGAGAAGA AAGACGAGGA
MPA3         AGAGAGAAAA AGAAGTCTCT GATGACGAAG AAGACAAAGA GGAGGAGAAG ACTGAAGAAA AGAGTGAGGA

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                355        365        375        385        395        405        415
MPA2         GAAGAAG--- GAGTCCGAAG GCGGCGATGA GGACGAGCCG AAGGTCGAGG ACGTGGAGGA CTCGGACGAG
MPA3         AAATAAGACC GAAGAAGAGA AGATTGATGA AGACGAACCT AAAGTTGAAG ACGTCGAGGA CTCCGAAGAT

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                425        435        445        455        465        475        485
MPA2         AAGAAAGACA AAAAGAAAAA GAAGAAAATA AAGGAGACGT ACGTCGAGGA CGAGGAGCTG AATAAGACTA
MPA3         AAGAAAGACA AGAAAAAGAA GAAAAAAATT AAGGAAAAGT ATGTCGAAGA CGAAGAGCTG AATAAAACTA

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|.
                495        505        515        525        535        545
MPA2         AGCCGTTATG GATGCGCAAC CCCGACGACA TCACTCAGGA AGAGTACGGC GAGTTC
MPA3         AACCCATTTG GATGAGAAAT CCCGATGACA CCACTCAAGA AGAATACGGA GAATTC
```

```
Protein      ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                 5          15         25         35         45         55         65
MPA2         EFGVGFYSAY LIADKVVVTS KHNDDEQYVW ESSAGGSFTI RADNT-EPLG RGTKIVLHLK EDQLEYAEEK
MPA3         EFGVGFYSAY LIADRVTVTS KHNDDEQYVW ESSAGGSFTI RTDATGEPLG RGTKIVLHLK EDQLEYAEEK

             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                 75         85         95         105        115        125        135
MPA2         RIKEIVKKHS QFIGYPIKLL VQKEREKEVS DDEEEEA--K DEKKDEEKK- ESEGGDEDEP KVEDVEDSDE
MPA3         RIKEIVKKHS QFIGYPIKLL VQKEREKEVS DDEEDKEEEK TEEKSEENKT EEEKIDEDEP KVEDVEDSED

             ....|....| ....|....| ....|....| ....|....| ..
                145        155        165        175
MPA2         KKDKKKKKKI KETYVEDEEL NKTKPLWMRN PDDITQEEYG EF
MPA3         KKDKKKKKKI KEKYVEDEEL NKTKPIWMRN PDDTTQEEYG EF
```

Alignment of amino acid sequences of all alleles from oribatid mites

```
                ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                    5         15         25         35         45         55         65
PYCHA4          QFGVGFYSAY LIADRVTVTS KHNDDEQYVW ESSAGGSFTI RTDATGEPLG RGTKIVLHLK EDQLEYAEEK
PYCHA6          QFGVGFYSAY LIADKVTVYS KHNDDEQYVW GSSAGGSFTI KPD-NTGPLG RGTKIILHLK EDQLEYSEEK
PPKWX1          QFGVGFYSAY LIADKVVVTS KHNDDEQYVW ESSAGGSFTI RAD-NTEPLG RGTKIVLHLK EDQLEYAEEK
PPKWX2          QFGVGFYSAY LIADRVTVTS RHNDDEQYVW ESSAGGSFTI RTDATGEPLG RGTKIVLHLK EDQLEYAEEK
PPKWY1          QFGVGFYSAY LIADRVTVTS KHNDDEQYVW ESSAGGSFTI RTDATGEPLG RGTKIVLHLK EDQLEYAEEK
PPKWY2          QFGVGFYSAY LIADRVVVHS KHNDHEQYVW ESAAGGSFTI RVD-TGESLG RGTKIVLHLK EDQLDYTEER
PPKA2           QFGVGFYSAY LIADRVTVTS KHNDDEQYVW ESSAGGSFTI RTDATGEPLG RGTKIVLHLK EDQLEYAEEK
PPKA1           QFGVGFYSAY LIADKVTVYS KHNDDEQYVW ESSAGGSFTI KPD-NTEPLG RGTKIILHLK EDQLEYSEEK
MNNG11          QFGVGFYSAY LIADRVTVTS KHNDDEQYVW ESSAGGSFTI RTD-NSEPLG RGTKIVLLLK EDQLEYAEEK
MNNG16          QFGVGFYSAY LIADRVTVTS KHNDDEQYVW ESSAGGSFTI RTDATGEPLG RGTKIVLHLK EDQLEYAEEK
MNCB11          QFGVGFYSAY LIADRVTVTS KHNDDEQYVW ESSAGGSLTI RTDATGEPLG RGTKIVLHLK EDQLEYAEEK
MNCB15          QFGVGFYSAY LIADRVTVTS KHNDDEQCVW ESSAGGSFTI RTDATGEPLG RGTEIVLHLK EDQLEYAEEK
MNCB16          QFGVGFYSAY LIADKVTVHS KHNDDEQYVW ESSAGGSFTI KPDVEGEPLG RGTKIILHLK EDQLGYSEEK
ASA1            QFGVGFYSAY LVADRVVVHS KHNDDEQYVW ESSAGGSFTI RVD-SGESLG RGTKIILFLK EDQLDYTEER
ASA7            QFGVGFYSAY LVADRVVVHS KHNDDEQYVW ESSAGGSFTI RVD-SGESLG RGTKIILFLK EDQLDYTEER
ASA8            QFGVGFYSAY LIADRVTVTS KHNDDEQYVW ESSAGGSFTI RTDATGEPLG RGTKIVLHLK EDQLEYAEEK
TVDOacon        QFGVGFYSAY LIADKVTVHS KHNDDEQYVW ESSAGGSFTI KPDVEGEPLG RGTKIILHLK EDQLEYSEEK
TVDObcon        QFGVGFYSAY LIADRVTVTS KHNDDEQYVW ESSAGGSFTI RTDATGEPLG RGTKIVLHLK EDQLEYAEEK
SM16            QFGVGFYSAY LIADRVVVHS KHNDDEQYVW ESAAGGSFTI RVD-TGESLG RGTKIVLHLK EDQLDYTEER
MPA2            EFGVGFYSAY LIADKVVVTS KHNDDEQYVW ESSAGGSFTI RAD-NTEPLG RGTKIVLHLK EDQLEYAEEK
MPA3            EFGVGFYSAY LIADRVTVTS KHNDDEQYVW ESSAGGSFTI RTDATGEPLG RGTKIVLHLK EDQLEYAEEK

                ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                    75         85         95        105        115        125        135
PYCHA4          RIKEIVKKHS RFIGYPIKLL VQKEREKEVS DDEEDKEE-- -EKTEEKSEE NKTEEEKIDE DEPKVEDVED
PYCHA6          RIKDIVKKHS QFIGYPIKLL VQKEREKEVS DDEEEKEE-- ---DKKEDEE KKEDKEGGDE DEPKVEDVED
PPKWX1          RIREIVKKHS QFIGYPIKLL VQKEREKEVS DDEEEEAK-- ---DEKKDEE KKES-EGGDE DEPKVEDAED
PPKWX2          RIKEIVKKHS QFIGYPIKLL VQKEREKEVS DDEEDKEE-- -EKTEEKSEE NKTEEEKIDE DEPKVEDVED
PPKWY1          RIKEIVKKHS RFIGYPIKLL VQKEREKEVS DDEQDKEE-- -EKTEEKSEE NKTEEEKIDE DEPKVEDVED
PPKWY2          RIKDIVKKHS QFIGYPIKLV VQKEREKEIS DDEEEKEE-- -EKKDETEEK EKTEE---NE DQPKVEDVED
PPKA2           RIKEIVKKHS QFIGYPIKLL VQKEREKEVS DDEEDKEE-- -EKTEEKSEE NKTEEEKIDE DEPKVEDVED
PPKA1           RIKDIVKKHS QFIGYPIKLL VQKEREKEVS DDEEEKEE-- ---DKKEDEE KKEDKEGGDE DEPKVEDVED
MNNG11          RIKEIVKKHS QFIGYPIKLV VQKEREKEIS DDEEDKEEAK EDKEDKIEDE DKTEEKKEEG DEPKVEDVED
MNNG16          RIKEIVKKHS RFIGYPIKLL VQKEREKEVS DDEEDKEE-- -EKTEEKSEE NKTEEEKIDE DEPKVEDVED
MNCB11          RIKDIVKKHS QFIGYPIKLL VQKEREKEVS DDGEDKEE-- -EKTEEKSEE NKTEEEKIDE DEPKVEDVED
MNCB15          RIKEIVKKHS QFIGYPIKLL VQKEREKEVS VDEQDKEE-- -EKTEEKSEE NKTEEEKIDE DEPKVEDVED
MNCB16          RIKDIVKKHS QFIGYPIKLL VQKEREKEVS DDEEDKEE-- ---DKKEDDE KKEEKEGGDE DEPKVEDVED
ASA1            RIKDIVKKHS QFIGYPIKLL VQKEREKEIS DDEEEKEE-- -EKKDETEEK EKTEE---NE DEPKVEDVED
ASA7            RIKDIVKKHS QFIGYPIKLL VQKEREKEIS DDEEEKEEKE EEKKDETEEK EKTEE---NE DEPKVEDVED
ASA8            RIKEIVKKHS RFIGYPIKLL VQKEREKEVS DDEEDKEE-- -EKTEEKSEE NKTEEEKIDE DEPKVEDVED
TVDOacon        RIKDVVKRHS QFIGYPIKLL VQKEREKEVS DDEEDKEE-- ---DKKEAAE KKEEKEGGDE DEPKVEDVED
TVDObcon        RIKEIVKKHS QFIGYPIKLL VQKEREKEVS DDEEDKEE-- -EKTEEKSEE NKTEEEKIDE DEPKVEDVED
SM16            RIKDIVKKHS QFIGYPIKLV VQKEREKEIS DDEEEKEE-- -EKKDETEEK EKTEE---NE DEPKVEDVED
MPA2            RIKEIVKKHS QFIGYPIKLL VQKEREKEVS DDEEEEAK-- ---DEKKDEE KKES-EGGDE DEPKVEDVED
MPA3            RIKEIVKKHS QFIGYPIKLL VQKEREKEVS DDEEDKEE-- -EKTEEKSEE NKTEEEKIDE DEPKVEDVED

                ....|....| ....|....| ....|....| ....|....| ....|
                   145        155        165        175        185
PYCHA4          SEDKKDKKKK KKIKEKYVED EELNKTKPIW MRNPDDITQE EYGEF
PYCHA6          SDEKKDKKKK KKIKEKYVED EELNKTKPIW MRNPDDITQE EYGEF
PPKWX1          SDEKKDKKKK KKIKETYVED EELNKTKPLW MRNPDDITQE EYGEF
PPKWX2          SEDKKDEKKK KKIKEKYVED GELNRTKPIW MRNPDDIN-- --GEF
PPKWY1          SEDKKDKKKK KKIKEKYVED EELNKTKPIW MRNPDDITQE EYGEF
PPKWY2          SEDKKDKKKK KKIKEKYVED EELNKTKPIW MRNPDDITQE EYGEF
PPKA2           SEDKKDKKKK KKIKEKYVED EELNKTKPIW MRNPDDITQE EYGEF
PPKA1           SDEKKDKKKK KKIKEKYVED EELNKTKPIW MRNPDDITQE EYGEF
MNNG11          SEDKKD-KKK KKIKEKYVED EELNKTKPIW MRNPDDITQE EYGEF
MNNG16          SEDKKDKKKK KKIKEKYVED EELNKTKPIW MRNPDDITQE EYGEF
MNCB11          SEDKKDKKEK KKIKEKYVED EELNKTKPIW MRNPDDITQE EYGEF
MNCB15          SEDKKDKKKK KKIQEKYVED EELNKTKPIW MRNPDDITQE EYGEF
MNCB16          SEEKKDKKKK KKIKEKYVED EELNKTKPIW MRNPDDITQE EYGEF
ASA1            SEDKKDKKKK KKIKEKYVED EELNKTKPIW MRNPDDITQE EYGEF
ASA7            SEDKKDKKKK KKIKEKYVED EELNKTKPIW MRNPDDITQE EYGEF
ASA8            SEDKKDKKKK KKIKEKYVED EELNKTKPIW MRNPDDITQE EYGEF
TVDOacon        SEEKKDKKKK KKIKEKYVED EELNKTKPIW MRNPDDITQE EYGEF
TVDObcon        SEDKKDKKKK KKIKEKYVED EELNKTKPIW MRNPDDITQE EYGEF
SM16            SEDKKDKKKK KKIKEKYVED EELNKTKPIW MRNPDDITQE EYGEF
MPA2            SDEKKDKKKK KKIKETYVED EELNKTKPLW MRNPDDITQE EYGEF
MPA3            SEDKKDKKKK KKIKEKYVED EELNKTKPIW MRNPDDTTQE EYGEF
```

Alignment of DNA sequences for all oribatid mite alleles

```
                ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                     5         15         25         35         45         55
ASA1            CAGTTCGGTG TGGGTTTTTA TTCGGCATAT TTGGTTGCGG ACAGAGTTGT GGTTCACTCG
ASA7            CAGTTTGGTG TCGGTTTCTA TTCGGCATAT TTGGTTGCGG ACAGAGTTGT GGTTCACTCG
ASA8            CAGTTTGGTG TGGGTTTTTA TTCCGCGTAT TTGATCGCCG ATAGAGTGAC CGTCACTTCA
TVDOacon        CAGTTCGGTG TGGGTTTcTA TtCCGCGTAT TTGATCGCCG ATAAGGTGAC GGTTCACTCA
TVDObcon        CAGTTTGGTG TGGGTTTTTA TTCCGCGTAT TTGATCGCCG ATAGAGTGAC CGTCACTTCA
MNCB11          CAGTTTGGTG TGGGTTTCTA TTCCGCGTAT TTGATCGCCG ATAGAGTGAC CGTCACTTCA
MNCB15          CAGTTCGGTG TTGGTTTCTA TTCCGCGTAT TtGATCGCCG ATAGAGTGAC CGTCACTTCA
MNCB16          CAGTTCGGTG TGGGTTTCTA TTCCGCGTAT TTGATCGCCG ATAAGGTGAC GGTTCACTCA
MNNG11          CAGTTCGGTG TTGGTTTCTA TTCGGCGTAT TTGATCGCAG ACAGAGTGAC CGTCACTTCT
MNNG16          CAGTTTGGTG TCGGTTTTTA TTCCGCGTAT TTGATCGCCG ATAGAGTGAC CGTCACTTCA
MPA2            GAGTTCGGTG TTGGTTTTTA CAGCGCATAC CTGATCGCCG ATAAGGTTGT GGTGACCTCT
MPA3            GAGTTCGGTG TTGGTTTTTA TTCCGCGTAT TTGATCGCCG ATAGAGTGAC CGTCACTTCA
PYCHA4          CAGTTTGGTG tGGGTTTTTA TTCCGCGtAT TTGATCGCCG ATAGAGTGAC CGTCACTTCA
PYCHA6          CAGTTTGGTG TCGGTTTCTA TTCTGCGTAT TTGATCGCAG ATAAGGTCAC CGTTTACTCC
PPKWX1          CAGTTTGGTG TTGGTTTTTA CAGCGCATAC CTGATCGCCG ATAAGGTTGT GGTGACCTCT
PPKWX2          CAATTTGGTG TTGGTTTCTA TTCCGCGTAT TTGATCGCCG ATAGAGTGAC CGTCACTTCA
PPKWY1          CAGTTTGGTG TAGGTTTCTA TTCCGCGTAT TTGATCGCCG ATAGAGTGAC CGTCACTTCA
PPKWY2          CAGTTCGGTG TGGGTTTCTA TTCTGCGTAT TTGATTGCAG ATCGAGTGGT GGTTCACTCG
PPKA2           CAGTTCGGTG TGGGTTTTTA TTCCGCGTAT TTGaTCGCCG ATAGAGTGAC CGTCACTTCA
PPKA1           CAGTTCGGTG TTGGTTTCTA TTCTGcGtAT TTGATCGCAG ATAAGGTCaC CGTTTACTCC
SM16            CAATTCGGTG TAGGTTTTTA TTCTGCGTAT TTGATTGCAG ATCGAGTGGT GGTTCACTCG

                ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                     65         75         85         95        105        115
ASA1            AAACaCAACG ACGACGAGCA GTACGTGTGG GAGTCTTCGG CCGGCGGTTC GTTCACCATT
ASA7            AAACACAACG ACGACGAGCA GTACGTGTGG GAGTCTTCGG CCGGCGGTTC GTTCACCATT
ASA8            AAGCATAACG ACGACGAACA GTACGTTTGG GAGTCTTCTG CTGGCGGGTC ATTCACTATC
TVDOacon        AAACACAACG ACGACGAACA GTACGTcTGG GAGTCATCTG CCGGCGGGTC GTTCACTATA
TVDObcon        AAGCATAACG ACGACGAACA GTACGTTTGG GAGTCTTCTG CTGGCGGGTC ATTCACTATC
MNCB11          AAGCATAATG ACGACGAACA GTACGTTTGG GAGTCTTCTG CTGGCGGGTC ACTCACTATC
MNCB15          AAGCATAATG ACGACgAACA GTgCGTTTGG GAGTCTTCTG CTGGCGGGTC ATTCACTATC
MNCB16          AAACACAACG ACGACGAACA GTACGTcTGG GAGTCATCTG CCGGCGGATC CTTCACTATA
MNNG11          AAGCATAACG ACGACGAACA GTATGTCTGG GAGTCATCGG CCGGCGGCTC CTTCACTATC
MNNG16          AAGCATAACG ACGACGAACA GTACGTTTGG GAGTCTTCTG CTGGCGGGTC ATTCACTATC
MPA2            AAGCACAACG ACGACGAGCA GTACGTGTGG GAGTCGTCGG CCGGCGGCTC GTTCACCATC
MPA3            AAGCATAACG ACGACGAACA GTACGTTTGG GAGTCTTCTG CTGGCGGGTC ATTCACTATC
PYCHA4          AAGCATAACG ACGACGAACA GTACGTTTGG GAGTCTTCTG CTGGCGGGTC ATTCACTATC
PYCHA6          AAACACAACG ACGACGAACA GTACGTCTGG GGGTCGTCTG CTGGCGGGTC GTTCACTATA
PPKWX1          AAGCACAACG ACGACGAGCA GTACGTGTGG GAGTCGTCGG CCGGCGGCTC GTTCACCATC
PPKWX2          AGGCATAACG ACGACGAACA GTACGTTTGG GAGTCTTCTG CTGGCGGGTC ATTCACTATC
PPKWY1          AAGCATAACG ACGACGAACA GTACGTTTGG GAGTCTTCTG CTGGCGGGTC ATTCACTATC
PPKWY2          AAGCACAACG ACCACGAGCA GTACGTCTGG GAGTCGGCAG CCGGCGGTTC GTTCACTATT
PPKA2           AAGCATAACG ACGACGAACA GTACGTTTGG GAGTCTTCTG CTGGCGGGTC ATTCACTATC
PPKA1           AAACACAACG ACGACGAACA GTACGTcTGG GAGTCGTCTG CTGGCGGGTC GTTCACTATA
SM16            AAGCACAACG ACGACGAGCA GTACGTcTGG GAGTCGGCAG CCGGCGGTTC GTTCACTATT

                ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                   125        135        145        155        165        175
ASA1            CGCGTCGAT- --AGCGGAGA ATCTTTGGGT CGCGGAACCA AAATAATCCT TTTTTTGAAA
ASA7            CGCGTCGAT- --AGCGGAGA ATCTTTGGGT CGCGGAACCA AAATAATCCT TTTTTTGAAA
ASA8            AGAACTGACG CCACGGGTGA GCCTTTAGGT CGCGGAACTA AGATTGTACT GCACTTGAAA
TVDOacon        AAGCCAGACG TCGAGGGCGA ACCTCTCGGC AGAGGCACCA AAATCATCCT TCACTTGAAA
TVDObcon        AGAACTGACG CCACGGGTGA GCCTTTAGGT CGCGGAACTA AGATTGTATT GCACTTGAAA
MNCB11          AGAACTGACG CCACGGGTGA GCCTTTAGGT CGCGGAACTA AGATTGTATT GCACTTGAAA
MNCB15          AgAACTGACG CCACGGGCGA GCCTTTAGGT CGCGGAACTG AGATTGTATT GCACTTGAAA
MNCB16          AAGCCAGACG TCGAGGGCGA ACCTCTCGGC AGAGGCACCA AAATCATCCT TCACTTGAAA
MNNG11          AGGACAGAT- --AATAGCGA ACCATTAGGT CGAGGCACTA AAATTGTCTT ACTTCTCAAA
MNNG16          AGAACTGACG CCACGGGTGA GCCTTTAGGT CGCGGAACTA AGATTGTACT GCACTTGAAA
MPA2            CGGGCCGAC- --AACACCGA GCCGTTGGGC AGAGGCACGA AGATTGTGCT GCACCTGAAG
MPA3            AGAACTGACG CCACGGGTGA GCCTTTAGGT CGCGGAACTA AGATTGTATT GCACTTGAAA
PYCHA4          AGAACTGACG CCACGGGTGA GCCTTTAGGT CGCGGAACTA AGATTGTATT GCACTTGAAA
PYCHA6          AAGCCAGAC- --AATACCGG ACCTCTCGGC AGAGGCACCA AAATTATTCT TCACCTCAAA
PPKWX1          CGGGCCGAC- --AACACCGA GCCGTTGGGC AGAGGCACGA AGATTGTGCT GCaCCTGAAG
PPKWX2          AGAACTGACG CCACGGGCGA GCCTTTAGGT CGCGGAACTA AGATTGTATT GCACTTGAAA
PPKWY1          AGAACTGACG CCACGGGTGA GCCTTTAGGT CGCGGAACTA AGATTGTACT GCACTTGAAA
PPKWY2          CGTGTGGAC- --ACTGGCGA GTCTTTGGGT CGCGGAACTA AGATTGTGCT CCATTTGAAA
PPKA2           AGAACTGACG CCACGGGTGA GCCTTTAGGT CGCGGAACTA AGATTGTATT GCACTTGAAA
PPKA1           AAGCCAGAC- --AATACCGA ACCTCTCGGC AGAGGCACCA AAATTATTcT TCACCTCAAA
SM16            CGTGTGGAC- --ACTGGCGA GTCTTTGGGT CGCGGAACTA AgATAGTGCT CCATTTGAAA
```

```
                ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                   185        195        205        215        225        235
ASA1            GAAGATCAGT TGGATTACAC TGAGGAAAGA CGTATCAAAG ATATCGTTAA AAAGCATTCG
ASA7            GAAGATCAGT TGGATTACAC TGAGGAAAGA CGTATCAAAG ATATCGTTAA AAAGCATTCG
ASA8            GAAGACCAGT TGGAATACGC GGAGGAAAAA CGCATTAAAG AAATTGTTAA AAAACACTCC
TVDOacon        GAGGATCAGT TGGAGTACTC CGAAGAGAAA CGCATCAAAG ATGTCGTGAA AAGACACTCA
TVDObcon        GAAGACCAGT TGGAATACGC GGAGGAAAAA CGCATTAAAG AAATTGTTAA AAAACACTCC
MNCB11          GAAGACCAGT TGGAATACGC GGAGGAAAAA CGCATTAAAG ATATTGTTAA AAAACACTCC
MNCB15          GAAGACCAGT TGGAATACgC GGAGGAAAAA CGCATTAAAG AAATTGTTAA AAAACACTCC
MNCB16          GAGGATCAGT TGGGGTACTC CGAAGAGAAG CGCATCAAAG ATATCGtGAA AAAACACTCA
MNNG11          GAAGACCAAT TAGAATACGC AGAAGAAAAA CGTATTAAAG AGATCGTGAA AAAACACTCG
MNNG16          GAAGACCAGT TGGAATACGC GGAGGAAAAA CGCATTAAAG AAATTGTTAA AAAACACTCC
MPA2            GAAGACCAgT TGGAGTACGC GGAGGAgAAG CGCATCAAAG AGATCGTGAA GAAGCACTCG
MPA3            GAAGACCAAT TGGAATACGC GGAGGAAAAA CGCATTAAAG AAATTGTTAA GAAACACTCC
PYCHA4          GAAGACCAGT TGGAATACGC GGAGGAAAAA CGCATTAAAG AAATTGTTAA AAAACACTCC
PYCHA6          GAGGATCAGT TGGAGTACTC CGAAGAGAAA CGTATAAAAG ATATCGTGAA AAAACACTCC
PPKWX1          GAAGACCAGT TGGAGTACGC GGAGGAGAAG CGCATCAGAG AGATCGTGAA GAAGCACTCG
PPKWX2          GAAgACCAGT TGGAATACgC GGAGGAAAAA CGCATTAAAG AAATTGTTAA AAAACACTCC
PPKWY1          GAAGACCAGT TGGAATACGC GGAGGAAAAA CGCATTAAAG AAATTGTTAA AAAACACTCC
PPKWY2          GAGGATCAGT TGGATTACAC TGAGGAGAGA CGCATCAAAG ATATCGTTAA AAAGCACTCG
PPKA2           GAAGACCAGT TGGAATACGC GGAGGAAAAA CGCATTAAAG AAATTGTTAA AAAACACTCC
PPKA1           GAGGATCAGT TGGAGTACTC CGAAGAGAAA CGTATAAAAG ATATCGTGAA AAAACACTCC
SM16            GAGGATCAGT TGGATTACAC TGAGGAGAGA CgCATCAAAG ATATCGTTAA AAAGCACTCG

                ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                   245        255        265        275        285        295
ASA1            CAATTCATTG GATATCCGAT TAAGCTTTTG GTACAAAAGG AGAGAGAAAA AGAGATCTCT
ASA7            CAATTCATTG GATATCCGAT TAAGCTTTTG GTaCAAAAGG AGAGAGAAAA AGAGATCTCT
ASA8            CGATTCATCG GTTATCCCAT TAAACTATTG GTTCAGAAAG AGAGAGAAAA AGAAGTCTCT
TVDOacon        CAATTCATCG GATATCCCAT CAAACTATTG GTTCAAAAGG AAAGAGAAAA AGAGGTCTCC
TVDObcon        CAATTCATCG GTTATCCCAT TAAACTATTG GTTCAGAAAG AGAGAGAAAA AGAAGTCTCT
MNCB11          CAATTCATCG GTTATCCCAT TAAACTCTTG GTTCAGAAAG AGAGAGAAAA AGAAGTCTCT
MNCB15          CAATTCATCG GTTATCCCAT TAAACTATTG GTTCAGAAAG AGAGAGAAAA AGAAgTCTcT
MNCB16          CAATTCATCG GATATCCCAT CAAACTATTG GTTCAAAAGG AAAGAGAAAA AGAGGTCTCC
MNNG11          CAATTCATCG GATATCCGAT CAAACTTGTC GTTCAAAAGG AAAGAGAAAA AGAGATCTCA
MNNG16          CGATTCATCG GTTATCCCAT TAAACTATTG GTTCAGAAAG AGAGAGAAAA AGAAGTCTCT
MPA2            CAGTTCATCG GATACCCAAT CAAACTACTc GTGCAgAAGG AACgCGAGAA GGAGGTGTCC
MPA3            CAATTCATCG GTTATCCCAT TAAACTATTG GTTCAGAAAG AGAGAGAAAA AGAAGTCTCT
PYCHA4          CGATTCATCG GTTATCCCAT TAAACTATTG GTTCAGAAAG AGAGAGAAAA AGAAGTCTCT
PYCHA6          CAATTCATCG GATATCCAAT CAAACTATTG GTCCAAAAGG AGAGGGAAAA AGAGGTCTCA
PPKWX1          CAGTTCATCG GATACCCAAT CAAACTACTC GTGCAGAAGG AACGCGAGAA GGAGGTGTCC
PPKWX2          CAATTCATCG GTTATCCCAT TAAACTATTG GTTCAGAAAG AGAGAGAAAA AGAAGTCTCT
PPKWY1          CGATTCATCG GTTATCCCAT TAAACTATTG GTTCAGAAAG AGAGAGAAAA AGAAGTCTCT
PPKWY2          CAGTTCATCG GGTATCCCAT CAAACTCGTG GTTCAAAAGG AAAGAGAGAA AGAGATCTCT
PPKA2           CAATTCATCG GTTATCCCAT TAAACTATTG GTTCAGAAAG AGAGAGAAAA AGAAGTcTCT
PPKA1           CAATTCATCG GATATCCAAT CAAACTATTG GTCCAAAAGG AGAGGGAAAA AGAGGTCTCA
SM16            CAGTTCATCG GGTATCCCAT CAAACTCGTG GTTCAAAAGG AAAGAGAGAA AGAGATCTcT

                ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                   305        315        325        335        345        355
ASA1            GATGACGAGG AAGAGAAGGA AGAA------ ---GAGAAAA AGGATGAAAC TGAGGAGAAA
ASA7            GATGACGAGG AAGAGAAGGA AGAGAAGGAA GAAGAGAAAA AGGATGAAAC TGAGGAGAAA
ASA8            GATGACGAAG AAGACAAAGA GGAG------ ---GAGAAGA CTGAAGAAAA GAGTGAGGAA
TVDOacon        GATGACGAGG AGGACAAAGA AGAA------ ---------G ATAAGAAAGA AGcCGCCGAA
TVDObcon        GATGACGAAG AAGACAAAGA GGAG------ ---GAGAAGA CTGAAGAAAA GAGTGAGGAA
MNCB11          GATGACGGAG AAGACAAAGA GGAG------ ---GAGAAGA CTGAAGAAAA GAGTGAGGAA
MNCB15          GTTGACGAAC AAGACAAAGA GGAG------ ---GAGAAGA CTGAAGAAAA GAGTGAGGAA
MNCB16          GATGACGAGG AGGACAAAGA AGAA------ ---------G ATAAGAAAGA AGACgACgAA
MNNG11          GACGACGAGG AAGACAAAGA AGAAGCCAAA GAAGATAAAG AAGACAAGAT CGAAGATGAA
MNNG16          GATGACGAAG AAGACAAAGA GGAG------ ---GAGAAAA CTGAAGAAAA GAGTGAGGAA
MPA2            GACgACgAGG AGGAAGAGGC GAAG------ ---------G ACGAGAAGAA AGACgAGGAG
MPA3            GATGACGAAG AAGACAAAGA GGAG------ ---GAGAAGA CTGAAGAAAA GAGTGAGGAA
PYCHA4          GATGACGAAG AAGACAAAGA GGAG------ ---GAGAAGA CTGAAGAAAA GAGTGAGGAA
PYCHA6          GACGACGAGG AAGAGAAAGA AGAG------ ---------G ACAAGAAAGA GGACGAAGAG
PPKWX1          GACGACGAGG AGGAAGAGGC GAAG------ ---------G ACGAGAAGAA AGACGAGGAG
PPKWX2          GATGACGAAG AAGACAAAGA GGAG------ ---GAGAAGA CTGAAGAAAA GAGTGAGGAA
PPKWY1          GATGACGAAC AAGACAAAGA GGAG------ ---GAGAAGA CTGAAGAAAA GAGTGAGGAA
PPKWY2          GATGACGAAg AAgAGAAGGA AGAG------ ---GAGAAAA AAGATGAAAC CGAGGAAAAG
PPKA2           GATGACGAAG AAGACAAAGA GGAA------ ---GAGAAGA CTGAAGAAAA GAGTGAGGAA
PPKA1           GACGACGAGG AAGAGAAAGA AGAG------ ---------G ACAAGAAAGA GGACGAAGAG
SM16            GATGACgAAG AAGAGAAGGA AGAG------ ---GAGAAAA AAGATGAAAC CGAGGAAAAG
```

```
                ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                   365        375        385        395        405        415
ASA1            GAGAAGACCG AAGAG----- ----AATGAA GACGAGCCGA AAGTAGAGGA CGTGGAAGAC
ASA7            GAGAAGACCG AAGAG----- ----AATGAA GACGAGCCGA AAGTAGAGGA CGTGGAAGAC
ASA8            AATAAGACCG AAGAAGAGAA GATTGATGAA GACGAACCTA AAGTTGAAGA CGTCGAGGAC
TVDOacon        AAGAAAGAAG AGAAAGAGGG CGGAGACGAA GACGAACCCA AAGTTGAAGA CGTCGAGGAT
TVDObcon        AATAAGACCG AAGAAGAGAA GATTGATGAA GACGAACCTA AAGTTGAAGA CGTCGAGGAC
MNCB11          AATAAGACCG AAGAAGAGAA GATTGATGAA GACGAACCTA AAGTTGAAGA CGTCGAGGAC
MNCB15          AATAAGACCG AAGAAGAGAA GATTGATGAA GACgAACCTA AAGTTGAAGA CGTCGAGGAC
MNCB16          AAGAAAGAAG AGAAAGAGGG CGGAGACGAA GACGAACCCA AAGTTGAAGA CGTCgAGGAT
MNNG11          GACAAGACCG AAGAGAAGAA AGAAGAGGGC GACGAGCCTA AGGTCGAAGA CGTCGAGGAC
MNNG16          AATAAGACCG AAGAAGAGAA GATTGATGAA GACGAACCTA AAGTTGAAGA CGTCGAGGAC
MPA2            AAGAAGGAGT cC---GAAGG CGGCGATgAG GACGAGCCGA AGGTcgAGGA CGTGGAGGAC
MPA3            AATAAGACCG AAGAAGAGAA GATTGATGAA GACGAACCTA AAGTTGAAGA CGTCGAGGAC
PYCHA4          AATAAGACCG AAGAAGAGAA GATTGATGAA GACGAACCTA AAGTTGAAGA CGTCGAGGAC
PYCHA6          AAGAAAGAGG ACAAAGAGGG CGGAGATGAA GACGAACCCA AAGTAGAAGA CGTCGAAGAC
PPKWX1          AAGAAGGAGT CC---GAAGG CGGCGATGAG GACGAGCCGA AGGTCGAGGA CGCGGAGGAC
PPKWX2          AATAAGACCG AAGAAGAGAA GATTGATGAA GACGAACCTA AAGTTGAAGA CGTCGAGGAC
PPKWY1          AATAAGACCG AAGAAGAGAA GATTGATGAA GACGAACCTA AAGTTGAAGA CGTCGAGGAC
PPKWY2          GAGAAAACCG AAGAG----- ----AATGAA GACCAACCGA AAGTCGAGGA CGTGGAGGAC
PPKA2           AATAAGACCG AAGAAGAGAA GATTGATGAA GACGAACCTA AAGTTGAAGA CGTCGAGGAC
PPKA1           AAGAAAGAGG ACAAAGAGGG CGGAGATGAA GACGAACCCA AAGTAGAAGA CGTCGAAGAC
SM16            GAGAAAACCG AAGAG----- ----AATGAA GACgAACCGA AAGTCGAGGA CGTGGAGGAC

                ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                   425        435        445        455        465        475
ASA1            TCTGAGGATA AGAAAGaCaA AAAGAAAAAG AAGAAAATAA AGGAAAAGTA TGTCGAAGAC
ASA7            TCTGAGGATA AGAAAGACAA AAAGAAAAAG AAGAAAATaA AGGAAAAGTA TGTCGAAGAC
ASA8            TCCGAAGATA AGAAAGACAA GAAAAAGAAG AAAAAAATTA AGGAAAAGTA TGTCGAAGAC
TVDOacon        TCAGAGGAAA AGAAAGACAA AAAGAAAAAG AAGAAAATTA AGGAAAAGTA TGTCGAAGAC
TVDObcon        TCCGAAGATA AGAAAGACAA GAAAAAGAAG AAAAAAATTA AGGAAAAGTA TGTCGAAGAC
MNCB11          TCCGAAGATA AGAAAGACAA GAAAGAGAAG AAAAAAATTA AGGAAAAGTA TGTCGAAGAC
MNCB15          TCCGAAGATa AGAAAGACaa GAAGAAGAAG AAAAAAATTC AGGAAAAGTa TGTCGAAGAC
MNCB16          TCAGAGGAAA AGAAAGACAA AAAGAAGAAG AAGAAAATTA AGGAAAAGTA TGTCGAAGAC
MNNG11          TCTGAAGATA AGAAAGAC-- -AAGAAAAAG AAGAAAATTA AGGAAAAGTA TGTCGAAGAC
MNNG16          TCCGAAGATA AGAAAGACAA GAAAAAGAAG AAAAAAATTA AGGAAAAGTA TGTCGAAGAC
MPA2            TCGGACGAGA AGAAAGACAA AAAGAAAAAG AAGAAAATaA AGGAGACGTa CGTcGAGGAC
MPA3            TCCGAAGATA AGAAAGACAA GAAAAAGAAG AAAAAAATTA AGGAAAAGTA TGTCGAAGAC
PYCHA4          TCCGAAGATA AGAAAGACAA GAAAAAGAAG AAAAAAATTA AGGAAAAGTA TGTCGAAGAC
PYCHA6          TCTGACGAAA AGAAGGATAA GAAGAAGAAG AAgAAGATTA AGGAAAAGTA CGTCGAAGAC
PPKWX1          TCGGACGAGA AGAAAGACAA AAAGAAAAAG AAGAAAATAA AGGAGACGTA CGTCGAGGAC
PPKWX2          TCCGAAGATA AGAAAGACGA GAAAAAGAAG AAAAAAATTA AGGAAAAGTA TGTCGAAGAC
PPKWY1          TCCGAAGATA AGAAAGACAA GAAAAAGAAG AAAAAAATTA AGGAAAAGTA TGTCGAAGAC
PPKWY2          TCGGAAGACA AGAAAGACAA AAAGAAAAAG AAGAAAATAA AGGAAAAGTA TGTGGAGGAC
PPKA2           TCCGAAGATA AGAAAGACAA GAAAAAGAAG AAAAAAATTA AGGAAAAGTa TGTCgAAGAC
PPKA1           TCTGACGAAA AGAAGGATAA GAAGAAGAAG AAGAAGATTA AGGAAAAGTA CGTCGAAGAC
SM16            TCGGAAGACA AGAAAGACAA AAAGAAAAAG AAGAAAATAA AGGAAAAGTa TGTGGAGGAC

                ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                   485        495        505        515        525        535
ASA1            GAAGAACTGA ACAAAACGAA ACCGATTTGG ATGCGAAACC CCGATGACAT CACTCAAGAA
ASA7            GAAGAACTGA ACAAAACGAA ACCGATTTGG ATGCGAAACC CCGATGACAT CACTCAAGAA
ASA8            GAAGAGCTGA ATAAAACTAA ACCCATTTGG ATGAGAAATC CCGATGACAT CACTCAAGAA
TVDOacon        GAAGAACTTA ACAAAACTAA ACCCATTTGG ATGAGAAATC CCGATGACAT CACTCAAGAA
TVDObcon        GAAGAGCTGA ATAAAACTAA ACCCATTTGG ATGAGAAATC CCGATGACAT CACTCAAGAA
MNCB11          GAAGAGCTCA ATAAAACTAA ACCCATTTGG ATGAGAAATC CCGATGACAT CACTCAAGAA
MNCB15          GAAGAGCTGA ATAAAACTAA ACCCATTTGG ATGAGAAATC CCGATGACaT CACTCAAGAA
MNCB16          GAAGAACTTA ACAAAACTAA ACCCATTTGG ATGAGAAATC CCGATGACAT CACTCAAGAA
MNNG11          GAAGAACTAA ACAAAACAAA ACCAATTTGG ATGAGAAATC CAGACGATAT CACTCAAGAA
MNNG16          GAAGAGCTGA ATAAAACTAA ACCCATTTGG ATGAGAAATC CCGATGACAT CACTCAAGAA
MPA2            gAGGAGCTGA ATAAGACTAA GCCGTTATGG ATGCGCAACC CCGACGACAT cACTcAGGAA
MPA3            GAAGAGCTGA ATAAAACTAA ACCCATTTGG ATGAGAAATC CCGATGACAC CACTCAAGAA
PYCHA4          GAAGAGCTGA ATAAAACTAA ACCCATTTGG ATGAGAAATC CCGATGACAT CACTCAAGAA
PYCHA6          GAAGAGCTGA ACAAAACGAA ACCAATTTGG ATGAGAAATC CGGATGACAT CACTCAAGAA
PPKWX1          GAGGAGCTGA ATAAGACTAA GCCGTTATGG ATGCGCAACC CCGACGACAT CACTCAGGAA
PPKWX2          GGAGAGCTGA ATAGAACTAA ACCCATTTGG ATGAGAAATC CCGATGACAT CAAC------
PPKWY1          GAAGAGCTGA ATAAAACTAA ACCCATTTGG ATGAGAAATC CCGATGACAT CACTCAAGAA
PPKWY2          GAAgAATTGA ACAAAACGAA ACCAATTTGG ATGCGAAATC CCGATGACAT CACTCAAGAA
PPKA2           GAAGAGCTGA ATAAAACTAA aCCCATTTGG ATGAGAAATC CCGATGACAT CACTCAAGAA
PPKA1           GAAGAGCTGA ACAAAACGAA ACCAATTTGG ATGAGAAATC CGGATGACAT CACTCAAGAA
SM16            GAAGAATTGA ACAAAACGAA ACCAATTTGG ATGCGAAATC CCGATGACAT CACTCAAGAA
```

```
              ....|....| ....|
                 545        555
ASA1          gAGTATGGAG AATTC
ASA7          GAGTATGGAG AATTC
ASA8          GAATACGGAG AATTC
TVDOacon      GAATACGGAG AATTC
TVDObcon      GAATACGGAG AATTC
MNCB11        GAATACGGCG AATTC
MNCB15        GAATACGGAG AATTC
MNCB16        GAATACGGAG AATTC
MNNG11        GAATACGGAG AATTC
MNNG16        GAATACGGAG AaTTC
MPA2          GAGTaCgGcG AGTTC
MPA3          GAATACGGAG AATTC
PYCHA4        GAATACGGAG AATTC
PYCHA6        GAATACGGAG AATTC
PPKWX1        GAGTACGGCG AGTTT
PPKWX2        ------GGAG AATTC
PPKWY1        GAATACGGAG AATTC
PPKWY2        GAGTACGGCG AGTTC
PPKA2         GAATACGGAG AATTC
PPKA1         GAATACGGAG AATTC
SM16          GAGTACGGCG AGTTC
```

Alignment of amino acid sequences used for testing of contaminations. For AN, see Table 3.2 and 3.3.

```
            ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                5         15         25         35         45         55         65
M57385      FGVGFYSAYL VADKVTVVSK NNADD---QY VWESTASGHF TVKKDDSHEP --LKRGTRLI LHLK-----E
AF136649    FGVGFYSAYL VADKVTVVSK NNNDD---QY VWESNASGHF TVTKDESEDQ --LKRGTRLI LHLK-----D
AF042329    FGVGFYSAYL VADSVTVVSK HNDDE---QY VWESAAGGSF TVQKDDKYEP --LGRGTRII LHLK-----E
AF421541    FGVGFYSGYL VADKIRVVSK HNDDE---QY VWESGAGGSF TVQKDTEMVH GEIKRGTKII CYLK-----E
X14167      FGVGFYSAYL VADRVTVVSK NNEDD---AY TWESSAGGTF TVTSTP---D CDLKRGTRIV LHLK-----E
X87770      FGVGFYSAYL VADRVTVTSK NNSDE---SY VWESSACGTF TITSTP---E SDMKRGTRIT LHLK-----E
AF151114    FGVGFYSAYL VAEKVEVISK SNDDES--QW RWESSAGGTF TVVNDDENPE -KLTRGTKII LHMK-----N
PPKWX2      FGVGFYSAYL IADRVTVTSR HNDDE---QY VWESSAGGSF TIRTDATGEP --LGRGTKIV LHLK-----E
TVDObcon    FGVGFYSAYL IADRVTVTSK HNDDE---QY VWESSAGGSF TIRTDATGEP --LGRGTKIV LHLK-----E
PPKA2       FGVGFYSAYL IADRVTVTSK HNDDE---QY VWESSAGGSF TIRTDATGEP --LGRGTKIV LHLK-----E
MPA3        FGVGFYSAYL IADRVTVTSK HNDDE---QY VWESSAGGSF TIRTDATGEP --LGRGTKIV LHLK-----E
ASA8        FGVGFYSAYL IADRVTVTSK HNDDE---QY VWESSAGGSF TIRTDATGEP --LGRGTKIV LHLK-----E
MNNG16      FGVGFYSAYL IADRVTVTSK HNDDE---QY VWESSAGGSF TIRTDATGEP --LGRGTKIV LHLK-----E
PYCHA4      FGVGFYSAYL IADRVTVTSK HNDDE---QY VWESSAGGSF TIRTDATGEP --LGRGTKIV LHLK-----E
MNCB15      FGVGFYSAYL IADRVTVTSK HNDDE---QC VWESSAGGSF TIRTDATGEP --LGRGTEIV LHLK-----E
PPKWY1      FGVGFYSAYL IADRVTVTSK HNDDE---QY VWESSAGGSF TIRTDATGEP --LGRGTKIV LHLK-----E
MNCB11      FGVGFYSAYL IADRVTVTSK HNDDE---QY VWESSAGGSL TIRTDATGEP --LGRGTKIV LHLK-----E
MNNG11      FGVGFYSAYL IADRVTVTSK HNDDE---QY VWESSAGGSF TIRTDNS-EP --LGRGTKIV LLLK-----E
TVDOacon    FGVGFYSAYL IADKVTVHSK HNDDE---QY VWESSAGGSF TIKPDVEGEP --LGRGTKII LHLK-----E
MNCB16      FGVGFYSAYL IADKVTVHSK HNDDE---QY VWESSAGGSF TIKPDVEGEP --LGRGTKII LHLK-----E
PYCHA6      FGVGFYSAYL IADKVTVYSK HNDDE---QY VWGSSAGGSF TIKPDN-TGP --LGRGTKII LHLK-----E
PPKA1       FGVGFYSAYL IADKVTVYSK HNDDE---QY VWESSAGGSF TIKPDN-TEP --LGRGTKII LHLK-----E
MPA2        FGVGFYSAYL IADKVVVTSK HNDDE---QY VWESSAGGSF TIRADN-TEP --LGRGTKIV LHLK-----E
PPKWX1      FGVGFYSAYL IADKVVVTSK HNDDE---QY VWESSAGGSF TIRADN-TEP --LGRGTKIV LHLK-----E
ASA1        FGVGFYSAYL VADRVVVHSK HNDDE---QY VWESSAGGSF TIRVDS-GES --LGRGTKII LFLK-----E
ASA7        FGVGFYSAYL VADRVVVHSK HNDDE---QY VWESSAGGSF TIRVDS-GES --LGRGTKII LFLK-----E
PPKWY2      FGVGFYSAYL IADRVVVHSK HNDHE---QY VWESAAGGSF TIRVDT-GES --LGRGTKIV LHLK-----E
SM16        FGVGFYSAYL IADRVVVHSK HNDDE---QY VWESAAGGSF TIRVDT-GES --LGRGTKIV LHLK-----E
AF261773    FGVGFYSAYL VADKVTVTSK HNDDE---QY IWESSAGGSF TVKPDS-SEP --LGRGTKIV LYIK-----E
U57473      FGVGFYSAYL VADKVTVTSK NNDDE---QY VWESSAGGSF TVRADN-SEP --LGRGTKIV LYIK-----E
AB060275    FGVGFYSSYL VADRVTVHSK HNDDE---QY VWESSAGGSF TVRPDS-GEP --LGRGTKIV LHVK-----E
AF254880    FGVGFYSCYL VADRVTVHSK HNDDE---QY MWESSAGGSF TVRPDP-GEP --LGRGTKIV LHIK-----E
L47285      FGVGFYSAYL VADKVVVTSK NNDDE---QY VWESSAGGSF TVRADS-GEP --LGRGTKIV LHIK-----E
AB043677    FGVGFYSAYL VAEKVTVITK HNDDE---QY AWESSAGGSF TVRTDT-GEP --MGRGTKVI LHLK-----E
AJ428213    FGVGFYSAYL VAEKVTVITK HNDDE---QY AWESSAGGSF TVRTDT-GEP --MGRGTKVI LHLK-----E
U94395      FGVGFYSAYL VAEKVTVITK HNDDE---QY AWESSAGGSF TVRTDT-GEP --MGRGTKVI LHLK-----E
NM_131310   FGVGFYSAYL VAEKVTVITK HNDDE---QY AWESSAGGSF TVKVDH-GEP --IGRGTKVI LHLK-----E
AF135117    FGVGFYSAYL VAERVTVITK HNDDE---QY IWESSAGGSF TVKVDT-GEP --MLRGTKVI LHMK-----E
U89945      FGVGFYSAYL VAERVTVITK HNDDE---QY IWESAAGGSF TVKVDT-GES --IGRGTRVI LHMK-----E
AF461150    FGVGFYSAFL VADRVTVTSK HNDDD---CY QWESSAGGSF IIRNCADPE- --VTRGTKIV LHLK-----E
AF201338    FGVGFYSAFL VADRVTVTSK HNDDD---CH QWESSAGGSF IIRNCVDPE- --MTRGTKIT LYLK-----E
AJ005784    FGAGFYSAFL VADKVVVASK HNDDD---CY QWESSAGGSF IIRQVNDPE- --LTRGTKIT LYIK-----E
AF250004    FGVGFYSSYL VSDKVIVTSK HNDDE---QY VWESSAGGSF TIKRDTTGEP --LGRGTKIV MYMK-----E
AF249999    FGVGFYSCYL VADKVVVTSK NNDDE---QY IWESSAGGSF TIKRDTTGEP --IGRGTKIV MYLK-----E
AF375826    FGVGFYSCYL VADKVTVTSK HNDDE---QY IWESSAGGNF SVSIDKHGER --LGRGTKVT LYMK-----E
AF251005    FGVGFYSLFL VADHVQVVSK HNDDE---QY IWESNAGGKF TVTLDETNER --LGRGTMLR LFLK-----E
NC_001148   FGVGFYSLFL VADRVQVISK SNDDE---QY IWESNAGGSF TVTLDEVNER --IGRGTILR LFLK-----D
NC_003424   FGVGFYSAYL VADKVQVVSK HNDDE---QY IWESSAGGSF TVTLDTDGPR --LLRGTEIR LFMK-----E
M55629      FGVGFYSAYL VADKVTVISK SNDDE---QY IWESNAGGTF KVTQDDDGRA --IGRGTKMI LHLK-----D
S59780      FGVGFYSAYL VADRVMVTTK HNDDE---QY VWESQAGGSF TVTHDTTGEQ --LGRGTKIT LFLK-----D
M99431      FGVGFYSAYL VAEKVIVTTK HNDDE---QY IWESQAGGSF TVTRDVDGEQ --LGRGTKIT LFLK-----E
NM_124983   FGVGFYSAYL VADKVVVTTK HNDDE---QY VWESQAGGSF TVTRDTSGEA --LGRGTKMV LYLK-----E
U01153DogE  FGVGFYSAFL VADKVIVTSK HNNDT---QH IWESDSN-EF SVIADPRGNT --LGRGTTIT LVLK-----E
X76301PigE  FGVGFYSAFL VADKVIVTSK HNNDT---QH IWESDSN-EF SVIADPRGNT --LGRGTTIT LVLK-----E
L14594Plan  FGVGFYSVYL VPDYVEVISK HNDDK---QY IWESKADGAF AISEDVWNEP --LGRGTEIR LHLR-----D
NC_004741   FGVGFYSAFI VADKVTVRTR AAGEKPENGV FWESAGEGEY TVADITKEDR -----GTEIT LHLR-----E
NC_002655   FGVGFYSAFI VADKVTVRTR AAGEKPENGV FWESAGEGEY TVADITKEDR -----GTEIT LHLR-----E
NC_003197   FGVGFYSAFI VADKVTVRTR AAGDKPENGV FWESAGEGEY TVADITKNDR -----GTEIT LHLR-----E
NC_002947   FGVGFYSAFI VADKVDVYSR RAGQPAAEGV HWSSKGEGEF EVATIDKPQR -----GTRIV LHLK-----K
NC001318    FGVGFYSAFI VSEKVEVTSK KALES--DAY IWSSDGKTGY EIEKAKKEES -----GTEIK LYLN-----K
NC_002163   FGVGFYSAFM VASKIEVLSK KALDD--KAY LWSSDAN-GY EIDDANKEEQ -----GTSIT LYLK-----D
NC_002677   FGIGFYSSFM VANKVELLTR KAGET--AAT RWSSDGEATY TIESVDEAPQ -----GTSVT LHLKPEDFED
```

```
            ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                75         85         95        105        115        125        135
M57385      DQTEYLEERR LKELVKKHSE FISFPISLSV EKTQETEVTD DEAELDEDKK PEEEK---PK DDKVEDVTDE
AF136649    DQSEYLEERR LKELVKKHSE FISFPIRLSV EKTTETEVTD DEAEPTEAES KPEEK---IT DVTEEEEEKE
AF042329    DQGEYLEERR LKDLVKKHSE FISFPIELAV EKTHEREVTE SEDEEEKKAD EKAEE---KE GEEKKEGEEK
AF421541    DQSEFLEERR LKDLVKKHSE FIGFPIELYV EKSKEKEVTD SEEDEEDEKK KEEGA---EG DEPKIEEVDE
X14167      DQQEYLEERR LKDLIKKHSE FIGYDIELMV ENTTEKEVTD EDEDEEAAKK AEEGE---EP KVEEVKDGVD
X87770      DQMEYLEPRR LKELIKKHSE FIGYDIELMV EKTTEKEVTD EDE--EDTKK ADEDE---EP KVEEVRE--G
AF151114    DNLEFLEERR IKDLIKKHSE FIAFPIELQV EKTEEKEETD EEDEEKEKED KEKTD---EP EIKEETEKKD
PPKWX2      DQLEYAEEKR IKEIVKKHSQ FIGYPIKLLV QKEREKEVSD DEEDKEEEKT EEKS---EEN KTEEEKIDED
TVDObcon    DQLEYAEEKR IKEIVKKHSQ FIGYPIKLLV QKEREKEVSD DEEDKEEEKT EEKS---EEN KTEEEKIDED
PPKA2       DQLEYAEEKR IKEIVKKHSQ FIGYPIKLLV QKEREKEVSD DEEDKEEEKT EEKS---EEN KTEEEKIDED
MPA3        DQLEYAEEKR IKEIVKKHSQ FIGYPIKLLV QKEREKEVSD DEEDKEEEKT EEKS---EEN KTEEEKIDED
ASA8        DQLEYAEEKR IKEIVKKHSR FIGYPIKLLV QKEREKEVSD DEEDKEEEKT EEKS---EEN KTEEEKIDED
MNNG16      DQLEYAEEKR IKEIVKKHSR FIGYPIKLLV QKEREKEVSD DEEDKEEEKT EEKS---EEN KTEEEKIDED
PYCHA4      DQLEYAEEKR IKEIVKKHSR FIGYPIKLLV QKEREKEVSD DEEDKEEEKT EEKS---EEN KTEEEKIDED
MNCB15      DQLEYAEEKR IKEIVKKHSQ FIGYPIKLLV QKEREKEVSV DEQDKEEEKT EEKS---EEN KTEEEKIDED
PPKWY1      DQLEYAEEKR IKEIVKKHSR FIGYPIKLLV QKEREKEVSD DEQDKEEEKT EEKS---EEN KTEEEKIDED
MNCB11      DQLEYAEEKR IKDIVKKHSQ FIGYPIKLLV QKEREKEVSD DGEDKEEEKT EEKS---EEN KTEEEKIDED
MNNG11      DQLEYAEEKR IKEIVKKHSQ FIGYPIKLVV QKEREKEISD DEEDKEEAKE DKEDKIEDED KTEEKKEEGD
TVDOacon    DQLEYSEEKR IKDVVKRHSQ FIGYPIKLLV QKEREKEVSD DEEDKEEDKK EAAE-----K KEEKEGGDED
MNCB16      DQLGYSEEKR IKDIVKKHSQ FIGYPIKLLV QKEREKEVSD DEEDKEEDKK EDDE-----K KEEKEGGDED
PYCHA6      DQLEYSEEKR IKDIVKKHSQ FIGYPIKLLV QKEREKEVSD DEEEKEEDKK EDEE-----K KEDKEGGDED
PPKA1       DQLEYSEEKR IKDIVKKHSQ FIGYPIKLLV QKEREKEVSD DEEEKEEDKK EDEE-----K KEDKEGGDED
MPA2        DQLEYAEEKR IKEIVKKHSQ FIGYPIKLLV QKEREKEVSD DEEEEAKDEK KDEE-----K KES-EGGDED
PPKWX1      DQLEYAEEKR IREIVKKHSQ FIGYPIKLLV QKEREKEVSD DEEEEAKDEK KDEE-----K KES-EGGDED
ASA1        DQLDYTEERR IKDIVKKHSQ FIGYPIKLLV QKEREKEISD DEEEKEE--- EKKDE---TE EKEKTEENED
ASA7        DQLDYTEERR IKDIVKKHSQ FIGYPIKLLV QKEREKEISD DEEEKEEKEE EKKDE---TE EKEKTEENED
PPKWY2      DQLDYTEERR IKDIVKKHSQ FIGYPIKLVV QKEREKEISD DEEEKEE--- EKKDE---TE EKEKTEENED
SM16        DQLDYTEERR IKDIVKKHSQ FIGYPIKLVV QKEREKEISD DEEEKEE--- EKKDE---TE EKEKTEENED
AF261773    DQTEYLEESK IKEIVNKHSQ FIGYPIKLLV QKERDQEVSD DEAEEEKKE- ---------- --MDTDEPKI
U57473      DQTDYLEESK IKEIVNKHSQ FIGYPIKLLV EKEREKEVSD DEADDEKKEG DEKE------ --METDEPKI
AB060275    DLAEFMEEHK IKEIVKKHSQ FIGYPIKLMV EKEREKELSD DEAEEEKKEE E--------- ----DEKPKI
AF254880    DLTEYLEEHK IKEIVKKHSQ FIGYPIKLMV EKEREKELSD DEAEEEKKED EK-------- ---EDDKPKI
L47285      DQLEYLEESK IKQIVNKHSQ FIGYPIKLLV EKEREKEVSD DEAEDDKKEE KK-------- ----EEDKKD
AB043677    DQTEYLEERR IKEIVKKHSQ FIGYPITLFV EKERDKEVSD DEAEEKEDKE EEKEK---EE KESDDKPEIE
AJ428213    DQTEYLEERR IKEIVKKHSQ FIGYPITLFV EKERDKEVSD DEAEEKEEKE EEKEK---EE KESDDKPEIE
U94395      DQTEYLEERR IKEIVKKHSQ FIGYPITLFV EKERDKEVSD DEAEEKEDKE EEKEK---EE KESEDKPEIE
NM_131310   DQTEYIEEKR VKEVVKKHSQ FIGYPITLYV EKERDKEISD DEA---EEEK AEKEE---KE EEGEDKPKIE
AF135117    DQTEYVEEKR VKEVVKKHSQ FIGYPITLFV EKEREKEISD D-----EEEK AEEEK---EE KEAEDKPKIE
U89945      DQFEYCEEKR VKEVVKKHSQ FIGYPITLFV EKSREKEVDL EEG---EKDE EADKD---SA AEDQDKPKIE
AF461150    DQTDYLEERR VREVVKKHPQ FIGYPIKLLV EKERDKEISD DEAEEEKK-- ----E-DEAK EEEKKPEDDV
AF201338    DQTDYLEERR IREVVKKHSQ FIGYPIKLLV EKERDKEISD DEAEDEKKDV KKEEE-KEEE KEIKKEEGED
AJ005784    DQTDYLEERR IKEIVKKHSQ FIGYPIKLTV EKERDKEVSD DEAEEEKK-- ---------- DEDKEKKEGE
AF250004    DQTEYLEERR LKEVVKKHSQ FIGYPIKLLV EKERDKEISD DEAEDEKK-- KTETK-DEDD TKKDAAKVEE
AF249999    DQTEYLEEKR IKEVIKKHSQ FIGYPIKLLV EKERDKEISD DEAEDEKD-- KKEK--DEDE TKKDEAKVEE
AF375826    DQLEFLEERK IKEIVKKHSQ FIGYPIKLVM QKEREKEVSD DEEEKKEEEM EVEKK----- -DEDKEKVDV
AF251005    DQLEYLEEKR IIEVVKKHSE FVAYPIQLVV TKEVEKEVPE EETLAEDEDK ----A----- ---TGEDDKK
NC_001148   DQLEYLEEKR IKEVIKRHSE FVAYPIQLVV TKEVEKEVPI PEEEKKDEEK KDEEK----- ---KDEDDKK
NC_003424   DQLQYLEEKT IKDTVKKHSE FISYPIQLVV TREVEKEVPE EEETEEVKNE ---------- -----EDDKA
M55629      EQTEYLNESK IKEVVKKQSE FIFYPIYLHV LKENEKEVPD EDAEEVKEDG ---------- ------DDKA
S59780      DQLEYLEERR LKDLVKKHSE FISYPIYLWT EKTTEKEISD DEEEEDNKKE EEGD------ ----------
M99431      DQLEYLEERR IKDLVKKHSE FISYPIYLWT EKTTEKEISD DEDDEP-KKE EEGD------ ----------
NM_124983   DQMEYIEERR LKDLVKKHSE FISYPISLWI EKTIEKEISD DEEEEE-KKD EEGK------ ----------
U01153DogE  EASDYLELDT IKNLVKKYSQ FINFPIYVWS SKTETVEEPM EEEEAAKEEK EDSD------ ----------
X76301PigE  EASDYLELDT IKNLVKKYSQ FINFPIYVWS SKTETVEEPM EEEEAAKEEK EESD------ ----------
L14594Plan  EAQEYLDEFK LKELVKRYSE FINFPIYLWA SKEVEVEVPA EEDDSSDDED NKSES----- -------SSS
NC_004741   GEDEFLDDWR VRSIISKYSD HIALPVEIEK R--------- ---------- ---------- ----------
NC_002655   GEDEFLDDWR VRSIISKYSD HIALPVEIEK R--------- ---------- ---------- ----------
NC_003197   GEDEFLDDWR VRSIISKYSD HIALPVEIEK R--------- ---------- ---------- ----------
NC_002947   DEQEFADGWR LRNVVKKYSD HIALPIQLPK EQA------- ---------- ---------- ----------
NC001318    EGLEYANKWK IQEIIKKYSN HINYPIYIKY SEP-IMK--- ---------- ---------- ----------
NC_002163   D--EFANAYK IESIIEKYSN HIQFPIFMEK EEFTPAK--- ---------- ---------- ----------
NC_002677   ELHDYTSEWK IRELVKKYSD FIAWPIRMEV ERRAPAT--- ---------- ---------- ----------
```

```
            ....|....| ....|....| ....|....| ....|....| ....|....| ....
               145        155        165        175        185
M57385      KVTDVTDEEE KKEEKKKKKR KVTNVTREWE MLNKQKPIWM RLPSEVTNEE YAAF
AF136649    KEAEKDGEE- ---KTEKKKR KVTNVTREWE MLNKQKPIWM RLPTEVTNEE YASF
AF042329    KEGEEEKKE- ----KTGKTK KVQEVTREWE QLNKQKPLWM RKPEEVTEEE YASF
AF421541    EKEKEEKKK- -------KTK KVKEVSHEWE QLNKNKPLWM RKSEDVTNEE YASF
X14167      ADAKK----- ------KKTK KVKEVKQEFV VQNKHKPLWT RDPKDVTKEE YASF
X87770      DEGEK----- ------KKTK KVKEVTKEYE VQNKHKPLWT RDPKDVTKEE YAAF
AF151114    KKKKK----- ---------- -VKVVHTEFE EQNKNKPLWM RKPEEITKEE YVNF
PPKWX2      EPKVEDVEDS EDKKDEKKKK KIKEKYVEDG ELNRTKPIWM RNPDDIN--- -GEF
TVDObcon    EPKVEDVEDS EDKKDKKKKK KIKEKYVEDE ELNKTKPIWM RNPDDITQEE YGEF
PPKA2       EPKVEDVEDS EDKKDKKKKK KIKEKYVEDE ELNKTKPIWM RNPDDITQEE YGEF
MPA3        EPKVEDVEDS EDKKDKKKKK KIKEKYVEDE ELNKTKPIWM RNPDDTTQEE YGEF
ASA8        EPKVEDVEDS EDKKDKKKKK KIKEKYVEDE ELNKTKPIWM RNPDDITQEE YGEF
MNNG16      EPKVEDVEDS EDKKDKKKKK KIKEKYVEDE ELNKTKPIWM RNPDDITQEE YGEF
PYCHA4      EPKVEDVEDS EDKKDKKKKK KIKEKYVEDE ELNKTKPIWM RNPDDITQEE YGEF
MNCB15      EPKVEDVEDS EDKKDKKKKK KIQEKYVEDE ELNKTKPIWM RNPDDITQEE YGEF
PPKWY1      EPKVEDVEDS EDKKDKKKKK KIKEKYVEDE ELNKTKPIWM RNPDDITQEE YGEF
MNCB11      EPKVEDVEDS EDKKDKKEKK KIKEKYVEDE ELNKTKPIWM RNPDDITQEE YGEF
MNNG11      EPKVEDVEDS EDKKD-KKKK KIKEKYVEDE ELNKTKPIWM RNPDDITQEE YGEF
TVDOacon    EPKVEDVEDS EEKKDKKKKK KIKEKYVEDE ELNKTKPIWM RNPDDITQEE YGEF
MNCB16      EPKVEDVEDS EEKKDKKKKK KIKEKYVEDE ELNKTKPIWM RNPDDITQEE YGEF
PYCHA6      EPKVEDVEDS DEKKDKKKKK KIKEKYVEDE ELNKTKPIWM RNPDDITQEE YGEF
PPKA1       EPKVEDVEDS DEKKDKKKKK KIKEKYVEDE ELNKTKPIWM RNPDDITQEE YGEF
MPA2        EPKVEDVEDS DEKKDKKKKK KIKETYVEDE ELNKTKPLWM RNPDDITQEE YGEF
PPKWX1      EPKVEDAEDS DEKKDKKKKK KIKETYVEDE ELNKTKPLWM RNPDDITQEE YGEF
ASA1        EPKVEDVEDS EDKKDKKKKK KIKEKYVEDE ELNKTKPIWM RNPDDITQEE YGEF
ASA7        EPKVEDVEDS EDKKDKKKKK KIKEKYVEDE ELNKTKPIWM RNPDDITQEE YGEF
PPKWY2      QPKVEDVEDS EDKKDKKKKK KIKEKYVEDE ELNKTKPIWM RNPDDITQEE YGEF
SM16        EPKVEDVEDS EDKKDKKKKK KIKEKYVEDE ELNKTKPIWM RNPDDITQEE YGEF
AF261773    EDVGEDEDAD KKDKDGKKKK TIKVAYTEDE ELNKTKPIWT RNPDDITQAE YGDF
U57473      EDVGEDEDAD KKDKDAKKKK TIKEKYTEDE ELNKTKPIWT RNPDDISQEE YGEF
AB060275    EDVGEDEDED KKDTK-KKKK TIKEKYTEDE ELNKTKPIWT RNADDITQDE YGDF
AF254880    EDVGEDDEED KKDKK-KKK- TIKEKYTEDE ELNKTKPIWT RNADDITQEE YGDF
L47285      DEPKLEDAED DDDKKDKKKK TVKVKYTEDE ELNKTKPIWT RNADDISQEE YGEF
AB043677    DVGSDEEEEE KKDGDKKKKK KIKEKYIDQE ELNKTKPIWT RNPDDITNEE YGEF
AJ428213    DVGSDEEEEE KKDGDKKKKK KIKEKYIDQE ELNKTKPIWT RNPDDITNEE YGEF
U94395      DVGSDEEEEE KKDGDKKKKK KIKEKYIDQE ELNKTKPIWT RNPDDITNEE YGEF
NM_131310   DVGSDDEED- TKDKDKKKKK KIKEKYIDQE ELNKTKPIWT RNPDDISNEE YGEF
AF135117    DVGSDDEED- SKDKDKKKTK KIKEKYIDQE ELNKTKPIWT RNPDDITMEE YGEF
U89945      DVGSDEDED- TKDSKNKRKK KVKEKYIDAE ELNKTKPIWT RNPDDITNEE YGEF
AF461150    SDDEAEKKKE EGDKKKKKTK KIKEKYTEDE ELNKTKPIWT RNPDDISNEE YAEF
AF201338    KEGEDEDKDK KDGEKKKKTK KIKEKYTEDE ELNKTKPIWT RNPDDITNEE YAEF
AJ005784    IEDVGEDEEE DKKDKDKKKK KIKEKYHEDE ELNKTKPIWT RNPDDISNEE YAEF
AF250004    VEDDDDDDDK KNDKDKKKKK KIKEKYIDEE ELNKQKPIWT RNPEDISTEE YAEF
AF249999    VEDDDDDD-K KKDTDKKKKK KIKEKYTDEE ELNKQKPIWT RNPEDISTEE YAEF
AF375826    EEVSDESDDE NKDKEKKKKK KIKEKYIDEE ELNKTKPIWT RNPEDIKHEE YAEF
AF251005    PKLEEVK-DE EEETKEKKTK KIKEEVTETE ELNKTKPLWT RNPSDITQEE YNAF
NC_001148   PKLEEVD-EE EE--KKPKTK KVKEEVQEIE ELNKTKPLWT RNPSDITQEE YNAF
NC_003424   PKIEEVD-DE SEK-KEKKTK KVKETTTETE ELNKTKPIWT RNPSEVTKEE YASF
M55629      PKVEEVDEDE EDKTKAKKTK KIKENKIEEE ELNKTKPIWT RNPADITQEE YASF
S59780      -VEEVDDEDK DTKDKSKKKK KVKEVSHEWV QINKQKPIWL RKPEEITRDE YASF
M99431      -IEEVDE--- DKEKEGKKKK KIKEVSHEWQ LINKQKPIWL RKPEEITKEE YASF
NM_124983   -VEEVDE--- EKEKEEKKKK KIKEVSHEWD LVNKQKPIWM RKPEEINKEE YAAF
U01153DogE  ---DEAAVEE EEEEKKPKTK KVEKTVWDWE LMNDIKPIWQ RPSKEVEDDE YKAF
X76301PigE  ---DEAAVEE EEEEKKPKTK KVEKTVWDWE LMNDIKPIWQ RPSKEVEDDE YKAF
L14594Plan  EEGEEEETEK EEDEKKPKTK KVKETTYEWE LLNDMKAIWL RNPKDVTDDE YTKF
NC_004741   ---------- --------EE KDGETVISWA KINKAQALWT RNKSEITDEE YKEF
NC_002655   ---------- --------EE KDGETVISWE KINKAQALWT RNKSEITDEE YKEF
NC_003197   ---------- --------EE KDGETVISWE KINKAQALWT RNKSEIKDDE YNEF
NC_002947   ---------- -------ATE GEEQPAEEWE TVNRASALWT RSRTEVKDEE YQEF
NC001318    ---------- ---DG----K QEG-IEEKEE KLNETTALWT KNKSEIKAEE YNEF
NC_002163   ---------- ---EG----E EEGKTELKIS QINKANALWR MQKSSLKAED YERF
NC_002677   ---------- ---SDGEGAD GEEQVTIETQ TINSMKALWT KSKDEVSEDE YKEF
```

Chapter Four

An understanding of the nature and mechanisms of meiosis is basic to the analysis of many biological problems for while its causes are rooted to cell biology its consequences lie ultimately in evolution.

Bernard John (1990)

4. Genetic basis for inheritance in oribatid mites

4.1. Origin and structure of chromosomes

It has long been customary to divide organisms into two groups: the prokaryotes, whose DNA is located in the cytoplasm, and the eukaryotes, whose DNA is separated from the cytoplasm by a nuclear membrane. Eukaryotic chromosomes are linear, change from a diffuse to a highly condensed form during mitosis, chromosomal DNA has specialised ends (telomeres) and is usually packed into nucleosomes by histones. In contrast, prokaryotic chromosomes have been thought to lack these properties. These differences have been related to the presence or absence of a nuclear membrane (Bendich and Drlica 2000). However, recent studies suggest that the distinction between eukaryotic and

prokaryotic chromosome features are less straightforward, e.g. histone-based DNA packing has been found in euryarcheotes (Sandman et al. 1998; Li et al. 1999). In addition, some prokaryotic chromosomes undergo large movements similar to those found in eukaryotic mitosis (Margolin 1998; Sharpe 1999). There are other exceptions to commonly held views on chromosome multiplicity, ploidy, linearity, heterochromatinisation, partitioning, and histone-based packing that chromosomal properties do not correlate well with the presence or absence of a nuclear membrane (Bendich and Drlica 2000).

Evidence for a dichotomy between eukaryotes and prokaryotes comes from nucleosome-based packing (although it is absent in some eukaryotes; Sala-Rovira et al. 1991; Schreiner et al. 1995). In this dichotomy, some of the euryarcheotes are grouped with eukaryotes and are distinguished from non-nucleosomal cell types (Bendich and Drlica 2000). This distinction also applies to the presence of histones in nucleosomal organisms (Sandman et al. 1998; Li et al. 1999). The presence of nucleosomal packing probably predates the evolution of a nuclear membrane and therefore represents a plesiomorphic character for eukaryotic organisms (Bendich and Drlica 2000). As indicated by chromosome structure, eukaryotic origin presumably was by symbiosis involving a euryarcheote as one of the symbionts (Lopez-Garcia et al. 1999). Some prokaryotes contain a nucleoid bounded by a single (Fuerst et al. 1998) or a double bilayer membrane (Fuerst et al. 1991); however, it remains to be demonstrated unequivocally if these membranes represent homologies between prokaryotes and eukaryotes or convergent evolution (Bendich and Drlica 2000).

The major DNA component of the eukaryotic cell is located in the nucleus and associated with basic histone proteins to form a molecular complex referred to as chromatin. This chromatin is organised into a series of subunits (nucleosomes), each of which contain a combination of 200 base pairs of DNA together with nine histone molecules (Matthews 1981). Collectively the nucleosomes are organised around a protein scaffold to form a system of two or more individual threads, the chromosomes, which

are highly diffused within the nucleus (John 1990). During cell divisions, the chromosomes are packed and undergo a movement along a spindle of microtubules. This movement is an active process involving kinetochores which usually are associated with the centromere region of the chromosome (Rieder 1982). According to the location of the centromere, chromosomes can be defined as metacentric, acrocentric, telocentric or, if no centromere structure exists, as acentric (=holocentric, holokinetic) (Figure 4.1). Acentric chromosomes have no localised centromere and the microtubules attach over the whole length of the chromosome during its movement, leading to the term holocentric. As not the centromere per se but the kinetochore is the functional part in the movement of the chromosome, holocentric chromosomes were termed holokinetic (John 1990; Helle et al. 1984; Wrensch et al. 1994; Dernburg 2001).

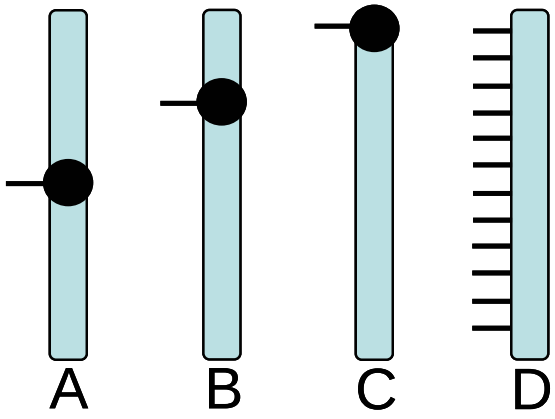


Figure 4.1
Different types of chromosomes. Black dots indicate centromeres: A. metacentric, B. acrocentric, C. telocentric, D. acentric (holocentric, holokinetic). Lines indicate attachment zones of the microtubules to the chromosome.

Holokinetic systems are fundamentally different from monocentric ones (Evans and Pond 1964; John 1990; Wrensch et al. 1994). It is still not clear which type is ancestral to the other but it has been suggested that

monocentric chromosomes evolved from holokinetic ones (Wrensch et al. 1994) and that centromeres derived from telomeric regions of holokinetic chromosomes (Holmquist and Dancis 1980; Stern and Hotta 1987). Telomeres are important features of linear chromosomes. They consist of repeated small G-rich sequences and form a small loop at the end. DNA-polymerases act in one direction and they need a double stranded part of DNA to start polymerisation. Due to this end-replication problem, telomeric DNA is shortened as the cell divides (Watson 1972). In most eukaryotes, this shortening of telomeric DNA is compensated by the activity of an enzyme called telomere terminal transferase (telomerase) that synthezises telomeric DNA de novo (Blackburn 1992) but is strictly regulated to be inactive in many somatic cells (Harley et al. 1990). Certain proteins bind tightly to the folded back structure at the telomeric ends; these are thought to protect the ends of linear chromosomes from end-to-end fusion and exonucleolytic erosion (Blackburn 1991; Lodish et al. 1995; Ishikawa and Naito 1999). Although linear chromosomes are some burden for the cell, the linearity of chromosomes and the presence of telomeres seems to be an important precondition for sexual reproduction and meiosis (de Lange 1998; Ishikawa and Naito 1999), since the movement and pairing of homologous chromosomes is based on the attachment of telomeres at the nuclear membrane (de Lange 1998).

In the following, the origin of mitosis and meiosis will be discussed based on the behaviour of monocentric and holokinetic chromosomes in eukaryotic cells.

4.2. Mitosis in monocentric and holokinetic systems

Conventionally, the life cycle of a eukaryote incorporates two principal phases of activity: development and growth on the one hand, and maturation and reproduction on the other (John 1990). In multicellular eukaryotes both phases of activity involve cell divisions. The reproduction of a cell is impossible without the duplication of the molecules out of which that cell is composed (John 1990). In the simplest and most common category of cell division the molecules are first multiplied and then separated into two identical daughter cells. This reproduction of the cell is termed mitosis (Greek, mitos: threads) based on the occurrence of chromosome threads within the nucleus at the outset of mitosis. Chromosomes are highly packed by higher order coiling of the primary chain of nucleosomes.

In several unicellular eukaryotes mitosis serves as a form of reproduction, referred to as asexual. Mitosis is also the basis of vegetative propagation in multicellular eukaryotes (Suomalainen 1987; Hughes 1989; John 1990).

Focussing on the chromosomes, mitosis involves two phases. During the phase of replication (synthetic or S phase) each chromosome is copied within the intact nuclear membrane and then consists of two identical sister threads (chromatids). Histone synthesis also occurs during S phase and is loosely coupled with DNA synthesis. Like DNA synthesis itself, the products of histone replication also segregate semi-conservatively (Matthews 1981). The second phase is the mitotic or M phase, when the sister chromatids of each condensed chromosome are separated accurately and equally into two nuclei (John 1990). This usually occurs in conjunction with the division of the cell itself. The separation of sister chromatids is due to their active movement along the spindle to opposite poles of the cell. In this movement the first substantial differences between monocentric and holokinetic chromosomes become obvious (Figure 4.2). Monocentric chromosomes are attached to a single spindle in the centromere region while holokinetic chromosomes have a number of active kinetochore regions. Therefore, chromosomal breaks are lost in

monocentric systems but can be restored in holokinetic systems. The mitotic phase can be subdivided in different stages (prophase, metaphase, anaphase and telophase). These stages are not distinct but they can be defined by chromosomal morphology (Griffiths et al. 1996).

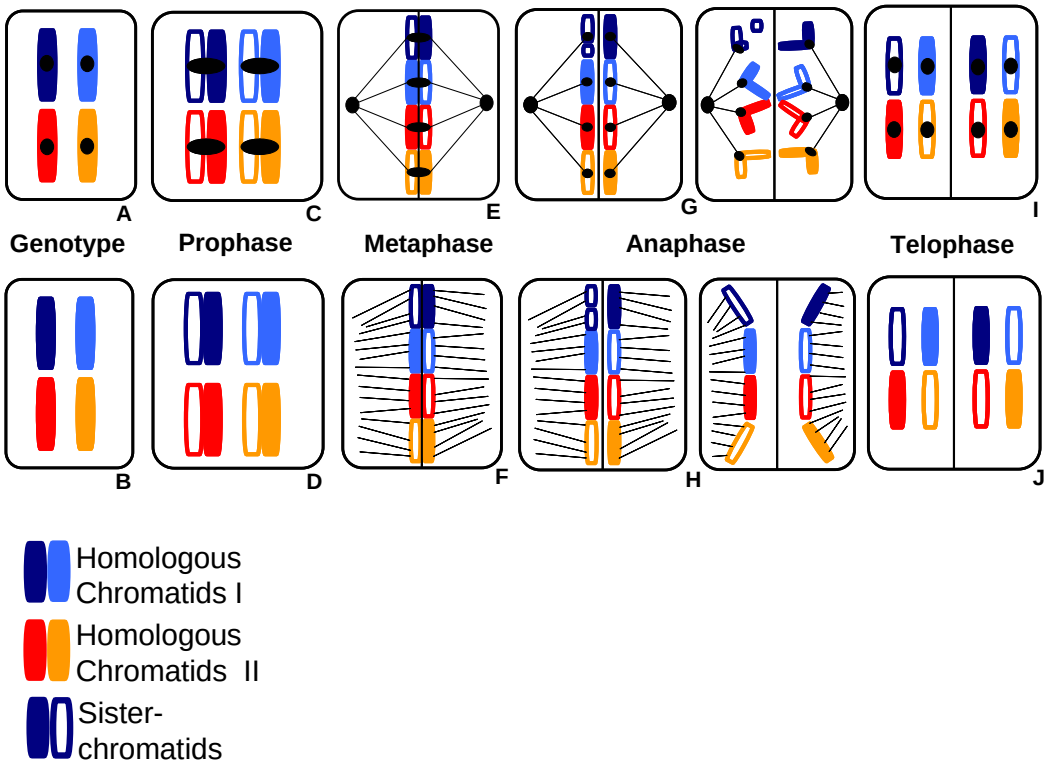


Figure 4.2

Schematic overview of the behaviour of monocentric (upper row) and holokinetic (lower row) chromosomes during mitosis. (A,B) Genotype of the diploid mother cell with 2n=4, black dots indicate centromeres in the upper row. (C,D) After replication of chromosomes and completion of the S phase, monocentric sister chromatids are attached at their centromere regions only, holokinetic sister chromatids are aligned over their whole length. (E,F) After orientation of chromosomes at the equatorial plate, monocentric chromosomes attach to the spindle by a single attachment located at the kinetochore in the centromere region. In holokinetic systems, spindles attach at various regions dispersed over the whole length of the chromosome. Holokinetic systems usually lack centrioles. (G, H) In anaphase, the divided sister chromatids move toward the poles. (I, J) The end of the telophase is also the end of the M phase. Independent from the presence of a centromere there are now two identical daughter cells in both monocentric and holokinetic systems. Nuclear membranes are not shown (after White 1973; John 1990; Wrensch et al. 1994; Lodish et al. 1995; Griffiths et al. 1996; Dernburg 2001).

4.3. Meiosis in monocentric and holokinetic systems

Meiosis is a basic process of most sexual eukaryotes, as it forms the basis of sexual reproduction (Solari 2002). In comparison to mitosis, where one cell is divided into two identical cells after a duplication of the chromosomes, meiosis consists of an additional division leading to four haploid instead of two diploid cells. Despite sex is present in the vast majority of eukaryotes, the origin of meiosis is still unknown. Presumably, meiosis evolved by adding an additional step to a mitotic cell division (Wrensch et al. 1994). Gessler and Xu (1999) also suggested that mitosis is ancestral to meiosis. While meiosis is inevitably linked with sexual reproduction and mitosis with asexual reproduction, a certain type of parthenogenesis includes meiosis (automoxis: Bell 1982; Suomalainen et al. 1987; Hughes 1989). In many automictic organisms the genome of the egg and one of the polar nuclei fuse although there are a number of alternative possibilities to restore diploidy (described in chapter 4.4). Often recombination is inhibited in automicts; in this case automixis may resemble apomixis (reproduction by mitosis; Lynch 1984).

Meiosis is a collective term for different types and mechanisms (John 1990). In “normal” meiosis the first of the two meiotic divisions is reductional and the second is equational. In “inverted” meiosis, the first division is equational and the second is reductional. The two types of meioses are also defined in respect to the orientation of bivalents: axial for normal meiosis and equatorial for inverted meiosis.

It remains unclear if “normal” meiosis as documented in textbooks or inverted meiosis is ancestral. On the following pages, I will analyse meiotic features of holokinetic and monocentric chromosomes in normal and inverted meiosis and try to answers this question.

Meiotic mechanisms in monocentric and holokinetic systems

As documented above the behaviour of monocentric chromosomes is different from holokinetic chromosomes in mitosis. The differences are even more pronounced in meiosis as holokinetic chromosomes have the

potential to undergo an inverted meiosis, which is strikingly varying from normal meiosis (Figures 4.3 and 4.4).

In normal meiosis, the first meiotic division (meiosis I) in reductional because homologous chromosomes (identical chromosomes with different parental origin) are separated and the two daughter cells are haploid after meiosis I. Sister chromatids are attached in their centromere region only, kinetic activity for the movement of the chromosome is restricted to kinetochores in the centromere region. In prophase I chiasmata occur leading to cross over and non-sister exchange of homologous chromosomes (=gene recombination, intrachromosomal recombination). Sister chromatids are separated in the second meiotic division leading to four haploid cells: one egg and three polar bodies in females or four sperm in males. Meiosis II is equational in normal meiosis because the ploidy number is unchanged after the second division. Due to the occurrence of gene recombination, the chromosomes in the four haploid cells differ from the parental chromosomes and from each other (John 1990).

In inverted meiosis, the first division is equational because sister chromatids are separated and homologous chromosomes remain paired, leaving the ploidy number unchanged. The pairing of homologous chromosomes is by a single chiasma in the telomere region which does not lead to gene recombination in any other chromosomal regions (Wrensch et al. 1994). Due to the absence of gene recombination and the equational kind, meiosis I in inverted systems resembles mitosis. Homologous chromosomes are separated in meiosis II in inverted systems leading to four haploid cells (reductional division).

The chromosomal composition after meiosis is different in monocentric and holokinetic systems. Both systems produce four haploid cells by meiosis where the assembly of the homologous paternal and maternal chromosomes is composed by chance (referred to as interchromosomal recombination). In monocentric systems, chromosomes underwent additional intrachromosomal recombination.

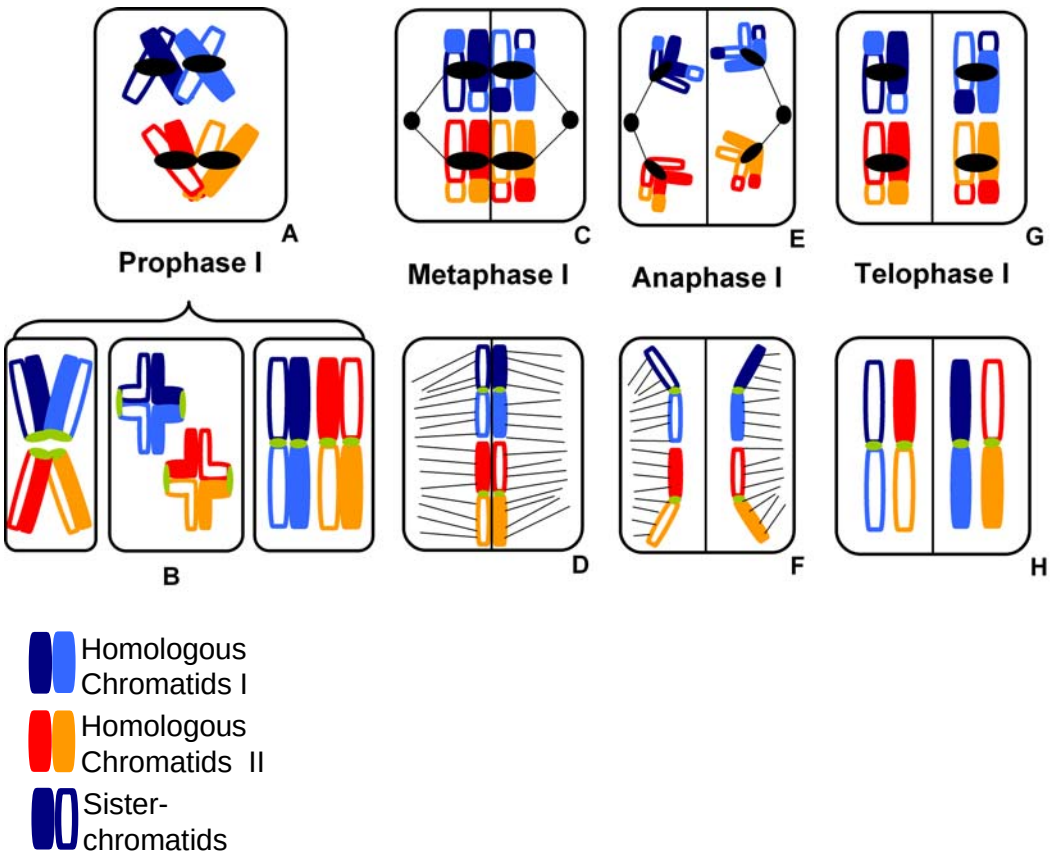


Figure 4.3

Schematic overview of the first meiotic division in monocentric (upper row) and holokinetic systems (lower row). (A) After replication of the DNA sister chromatids are attached in their centromere regions. Homologous chromosomes are aligned. Chiasmata may occur in several regions leading to intrachromosomal recombination. (B) In holokinetic systems replicated chromosomes are associated in their telomeric regions by a terminalised chiasma. By the end of prophase I, homologous chromatids are connected in telomeric regions; sister chromatids have lost their connection. (C, E) In metaphase I of monocentric systems, sister chromatids remain paired and homologous chromosomes become separated by anaphase I. (D, F) In holokinetic systems, homologous chromatids remain paired in the first division while sister chromatids are separated by anaphase I. (G) In monocentric systems, the first meiotic division is reductional as the ploidy number after the initial replication is reduced. (H) The first meiotic division in inverted meiosis leads to two diploid cells and is therefore equational (after White 1973; John 1990; Wrensch et al. 1994; Lodish et al. 1995; Griffiths et al. 1996).

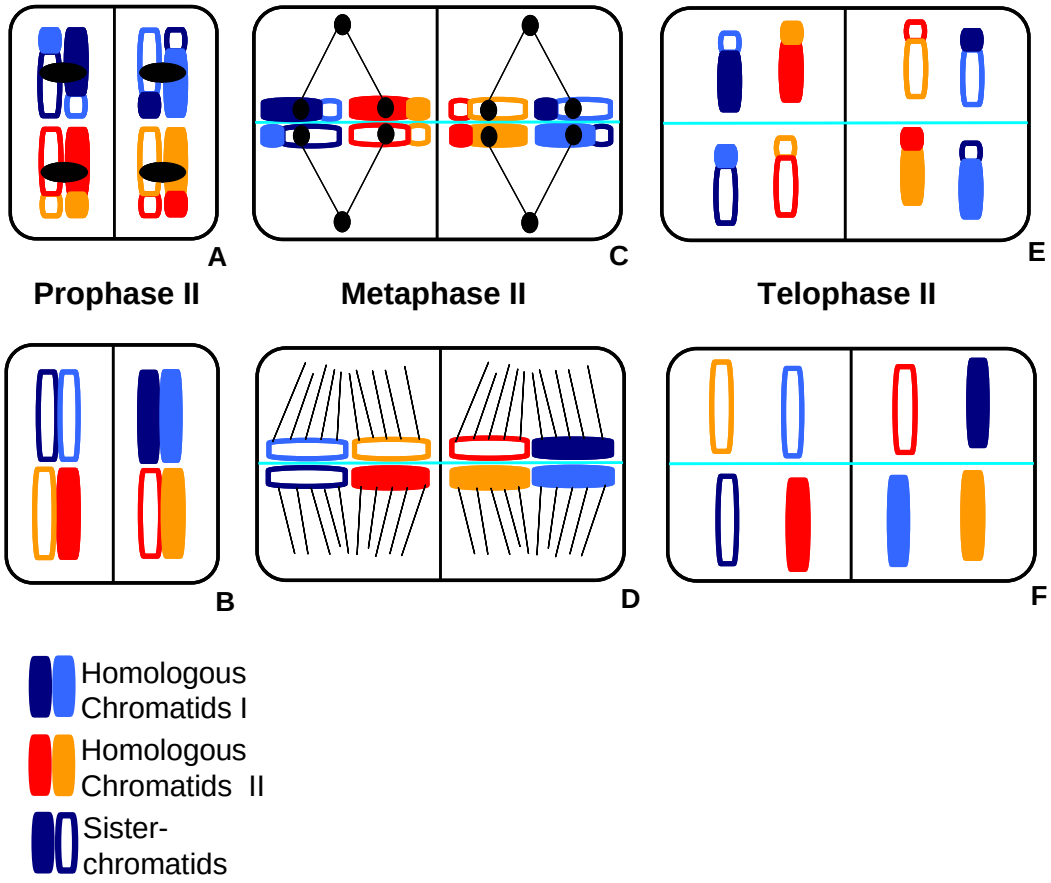


Figure 4.4

Schematic overview of the second meiotic division in monocentric (upper row) and holokinetic systems (lower row). (A) After meiosis I in monocentric systems the two haploid cells undergo a second division without further DNA replication. (C, E) This cell division in meiosis II is equational in monocentric systems; sister chromatids are separated, as in mitosis and result in four haploid cells. (B) Holokinetic systems with inverted meiosis constitute two diploid cells after meiosis I containing one of each sister chromatids. (D, E) The second meiotic division is reductional. Here, homologous chromatids are separated resulting in four haploid cells. (E, F) As chiasmata are terminalised in holokinetic chromosomes there is no recombination of functional genes (after White 1973; John 1990; Wrensch et al. 1994; Lodish et al. 1995; Griffiths et al. 1996).

The success of sexual reproduction and meiosis is due to the ability to recombine the genomes of individuals in a way that makes selection more efficient (Rice and Chippindale 2001; see also chapter 1.3). Meiotic recombination is also responsible for a non-biased distribution of genes after meiosis and for preventing segregation distorters from acting (Ridley 2000).

As meiosis acts only in the germ line and produces haploid cells, a mechanism to restore diploidy in somatic cells of the offspring is necessary.

The origin of meiosis

To investigate the origin of meiosis organisms that are basal in the tree of eukaryotes have to be studied. It is remarkable that while sexual reproduction and meiosis are present in virtually all eukaryotes they are absent in archea and bacteria (Solari 2002). However, components of sexual reproduction and meiosis are present in archea and/or bacteria, such as molecular mechanisms of gene recombination, genome packing, separation of parental and replicated genomes and mechanisms for the association of genomes with membranes (Solari 2002).

One of the most primitive component of sexual reproduction which is present in bacteria is the molecular apparatus for gene recombination. The first recognised recombinase, RecA, was isolated from the bacterium *E. coli* (Clark and Margulis 1965) and it has homologs in all studied eukaryotes. While RecA is not essential for sporadic bacterial gene recombination (instead, it is more important for the DNA synthesis when blocked by DNA damage; Courcelle et al. 2001) the homologs in eukaryotes are necessary for intrachromosomal recombination. In a similar way the bacterial proteins MutLS and MutH are essential in prokaryotes for the correction of mismatch errors in DNA after replication while the homologs in eukaryotes evolved an additional function: during meiosis, they are essential for the resolution of crossovers and mismatch repair at heteroduplex DNA in cross over regions during the prophase (Kirkpatrick 1999). However, there are large differences between

prokaryotes and eukaryotes: prokaryotes lack a nuclear envelope, nuclear pores, microtubules and a cytoskeleton similar to that of eukaryotes (Drlica and Bendich 2000; Bendich and Drlica 2000). These differences are related to the absence of a true mitotic process among prokaryotes (Lewis 2001) which is an essential precondition for meiosis because it ensures the equal distribution of genomes in daughter cells (Solari 2002).

Meiosis in extant organisms requires the activity of hundreds of genes, many of which are also needed for mitosis, and meiosis could have hardly evolved without a functional mitotic process for genome separation. Therefore, mitosis presumably evolved from prokaryotic cell division and meiosis evolved from mitosis (Wrensch et al. 1994; Gessler and Xu 1999).

Ancestrality of inverted meiosis to normal meiosis

As meiosis is assumed to have derived from mitosis and the first meiotic division with inverted meiosis resembles mitosis, this mechanism probably represents the ancestral kind of meiosis. As inverted meiosis is only possible with holokinetic chromosomes these probably also represent the ancestral type of chromosomes (Halkka 1959; Rhoades 1961; Suomalainen and Halkka 1963; John 1990; Wrensch et al. 1994).

While holokinetic chromosomes are a precondition for inverted meiosis, holokinetic systems do not necessarily undergo an inverted meiosis with chiasma terminalisation (Nokkala and Nokkala 1997). Also, kinetic activity of holokinetic chromosomes is not always diffused over the whole length of the chromatids in meiosis but may become restricted to one of the telomeric regions (Perez et al. 2000). Probably both the absence of chiasma terminalisation and the restriction of kinetic activity to the telomeric region are linked because chiasma proximity inhibits kinetic activity (Camacho et al. 1985). This activation of localised kinetochores is another way, besides inverted meiosis, for holokinetic chromosomes to segregate correctly (Loidl, personal communication) and is realised in *Caenorhabditis elegans* and other nematodes (Albertson and Thomson 1993). The occurrence of holokinetic chromosomes is widespread in fungi, plants and animals (for an overview, see Wrensch et al. 1994),

predominant in many arthropod groups and presumably universal in nematodes, mites and many insect orders, such as Hemiptera, Homoptera and Lepidoptera (Pazi 1997; Vanzela et al. 1998; Mandrioli et al. 1999; Perez et al. 2000; Rebagliati et al. 2001; Nokkala et al. 2002). Inverted meiosis is little studied but was inferred to occur in plants (Pazy 1997) and mites (Wrensch et al. 1994).

The evolution of a synaptonemal complex allows almost all eukaryotes to keep maternal and paternal chromosomes paired tight during the prophase. Meiotic recombination is highly regulated and the number of crossovers and illegitime recombination is lower than in protists (Solari 2002). One of the main adaptive advantages of the synaptonemal complex presumably is that it controls recombination by chiasmatic interference (inhibition of another recombinational event in the immediacy of a previous one; Pigozzi and Solari 1999; Gorlov and Gorlova 2001). A synaptonemal complex is also present in holokinetic systems with inverted meiosis. It breaks down together with the nuclear membrane when the telomeric association of sister chromatids switches to an association of homologous chromatids (Figure 4.3; Wrensch et al. 1994).

Virtually all eukaryotes have linear chromosomes with conserved telomeric structures and it has been assumed that these are necessary for regular meiosis with recombination (Ishikawa and Naito 1999). Indeed, the presence of circular chromosomes poses a number of problems for meiosis: if two circular homologous chromosomes undergo one or any other uneven number of crossovers, they will produce dicentric chromosomes which should break up at meiotic division. Furthermore, the circularity of chromosomes raises difficulties for their pairing with homologous regions (Solari 2002). The evolution of linear chromosomes with differentiated telomeres therefore presumably preceded the origin of meiosis as the association of telomeres with the inner side of the nuclear envelope is also needed for the initiation of meiosis. Again, this indicates that holokinetic chromosomes are ancestral to monocentric chromosomes because they have the same functionality of telomeres but lack a complex centromere which is not important for the initiation of meiosis.

Evolution of normal meiosis from inverted meiosis

Possibly “normal” meiosis and monocentric chromosomes evolved from inverted meiosis and holokinetic chromosomes (Figure 4.5). In this scenario, holokinetic chromosomes are ancestral and kinetic activity is distributed over the whole length of the chromosome. Due to the switch of telomeric association between sister chromatids and homologous chromatids with a teminalised chiasma in prophase I, meiosis is inverted and the first meiotic division is equational. If chiasmata are not completely terminalised, sister chromatids remain paired after prophase I and kinetic activity may become restricted to telomeres as there is no inhibiting chiasma in the telomeric region (Keyl 1957). Now, the first meiotic division is reductional. Here, telomeric regions are responsible for kinetic movement of the chromosome and for the regular functioning of telomeres. At this stage, an additional centromere function may form in the telomeric region (Holmquist and Dancis 1980). Centromeres are chromosomal regions of repetitive DNA sequences and are associated with kinetochores (Griffiths et al. 1996). Telomeres also contain repetitive DNA sequences; together with the restricted kinetic activity to telomeric regions, the evolution of a centromere under these circumstances is a conceivable step. As the centromere moves from the telomere to a more central point of the chromosome leading to acrocentric or metacentric chromosome morphology, telomeres retain their original function and the kinetic activity becomes restricted to the centromere region. This is the situation which is commonly depicted as “normal” meiosis in textbooks. If in fact “normal” meiosis evolved from “inverted” meiosis the terms are quite misleading as they swap evolutionary reality (Wrensch et al. 1994).

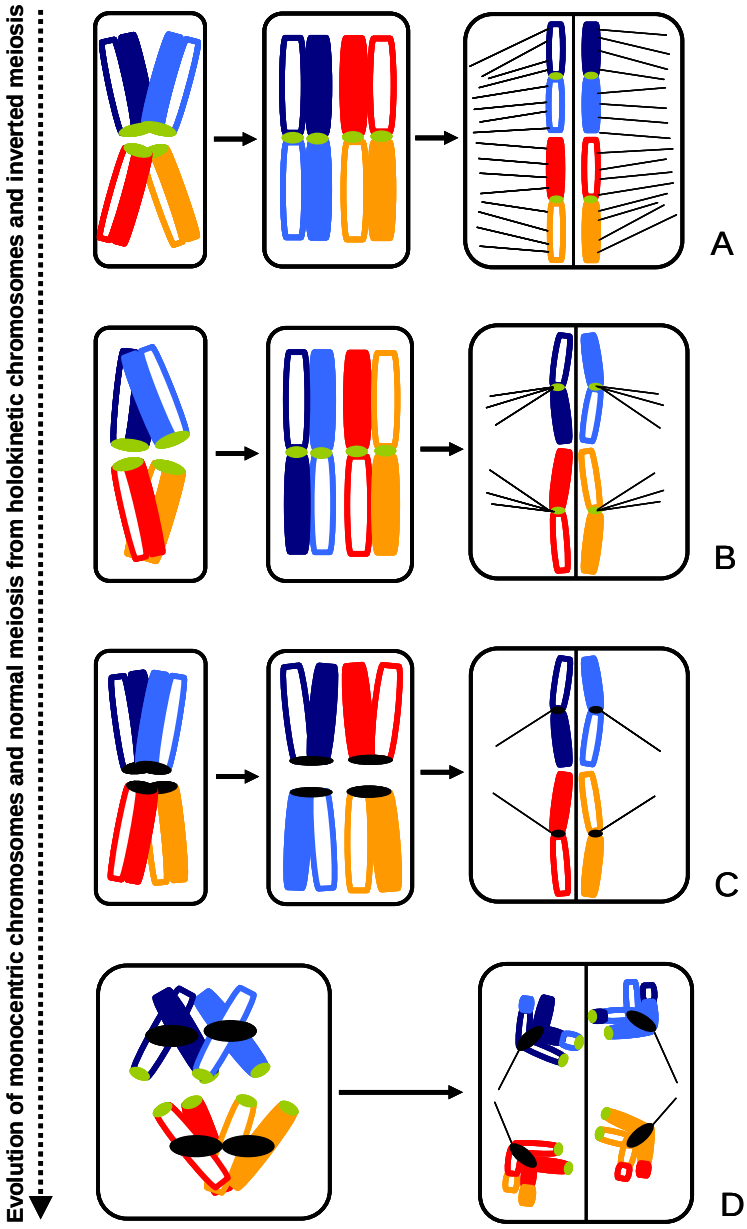


Figure 4.5

First meiotic division. Proposed scenario for the evolution of monocentric chromosomes and normal meiosis (D) from holokinetic chromosomes and inverted meiosis (A). Grey dots indicate telomeres and are shown on one end of the chromosome only, black dots indicate centromeres. (A) Chromosomal activity in holokinetic systems with a terminalised chiasma and inverted meiosis (first meiotic division is equational leading to two diploid daughter cells; see Figure 4.3). (B) The chiasma is not terminalised, therefore, kinetic activity could be restricted to telomeric regions. Sister chromatids remain paired leading to a reductional first division. (C) The telomeres acquire an additional function as centromeres. Kinetic activity is restricted to the centromere region. (D) The centromere moves down the chromosome leading to acrocentric or metacentric morphology. Except in (A) in all steps intrachromosomal recombination may occur.

4.4. Meiosis in parthenogenetic organisms

The former paragraphs dealt with mitosis and meiosis without considering whether the two homologous genomes are from different parents (sexual reproduction) or from the same individual (parthenogenesis). While apomixis (mitotic reproduction) is the most common form of parthenogenetic reproduction in unicellular eukaryotes and plants, automixis (parthenogenetic reproduction including meiosis) is common in many metazoan taxa (Suomalainen et al. 1987; Hughes 1989). There are different ways for automictic organisms to restore diploidy.

One mechanism is gamete duplication where the haploid egg is duplicated and these two products fuse to give rise to complete homozygous diploidy (Figure 4.6). This is known from some species of crustaceans and insects (Suomalainen et al. 1987).

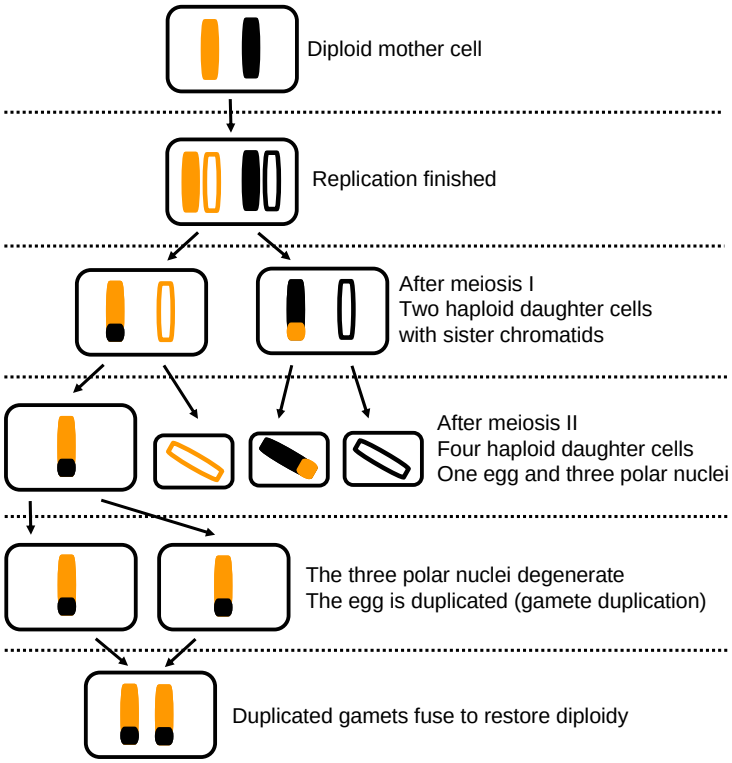


Figure 4.6

Automixis by gamete duplication, in the context of normal meiosis. The result is a diploid embryo which is completely homozygous, also in cross over regions.

Another mechanism to restore diploidy is central fusion. With central fusion, the two central polar nuclei fuse. Depending on the segregation of chromosomes in meiosis II, this may restores the heterozygous state of the mother except in cross over regions or also in cross over regions (Figure 4.7). Central fusion automixis is known from some insect species (Suomalainen et al. 1987).

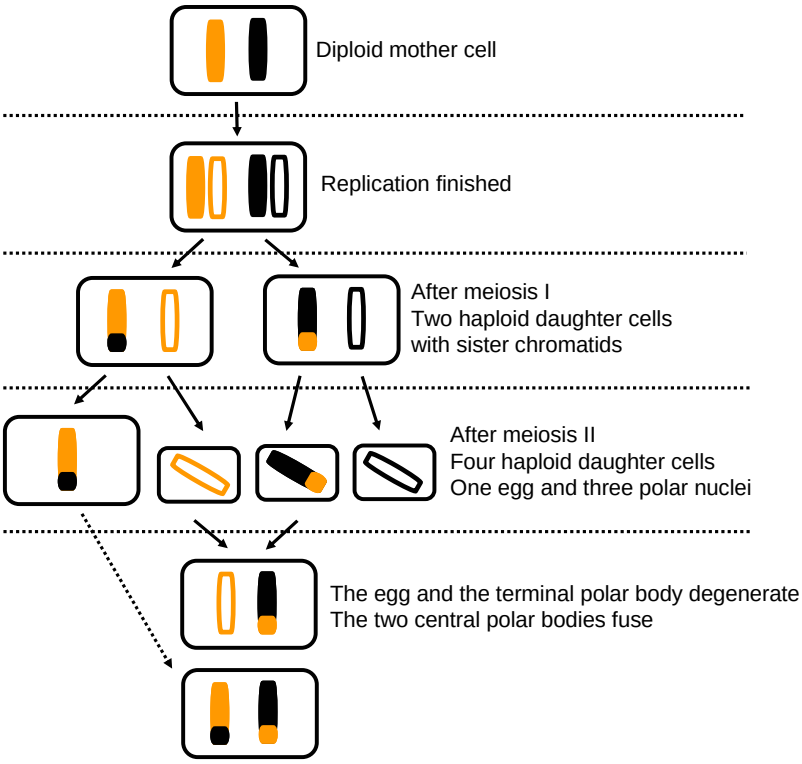


Figure 4.7

Automixis by central fusion, in the context of normal meiosis. The egg and the terminal polar body degenerate and the two central polar nuclei fuse. The result is a diploid embryo that is heterozygous except in cross over regions or a complete heterozygous embryo, depending on segregation of chromosomes in meiosis II.

A third mechanism is terminal fusion (Figure 4.8.), in which the egg fuses with the second polar nucleus. Terminal fusion is realised in some species of nematodes, lumbricids, crustaceans and insects (Suomalainen et al. 1987), as well as some mites. In monocentric systems with normal meiosis this leads to complete homozygosity except in cross over regions, since sister chromatids fuse to give rise to a diploid embryo.

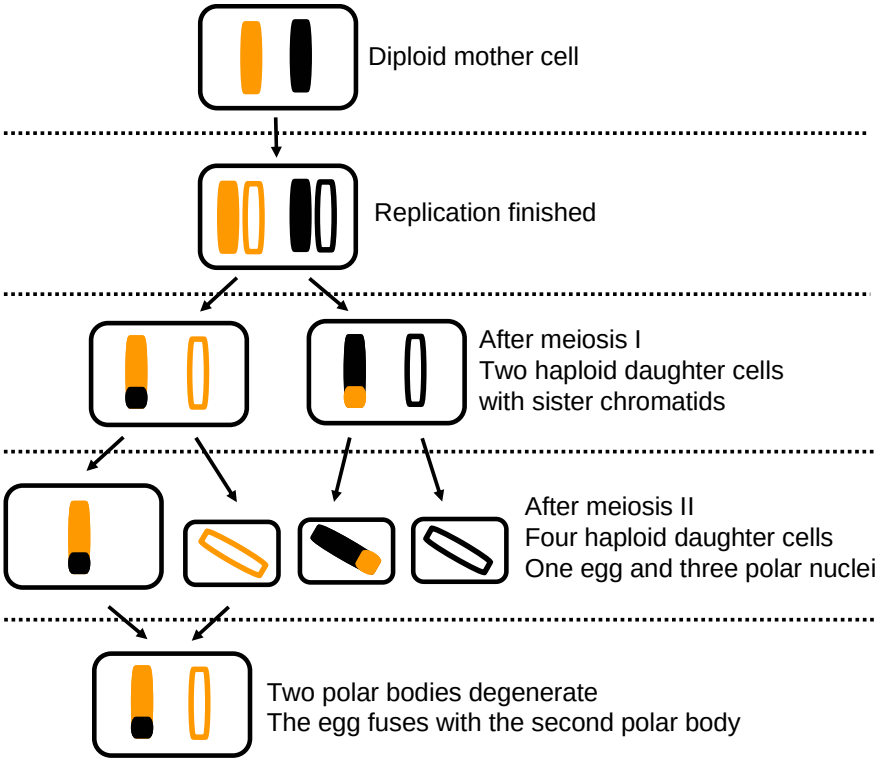


Figure 4.8

Automixis by terminal fusion, in the context of normal meiosis. Because replicated sister chromatids fuse, the diploid embryo is homozygous except in cross over regions.

In holokinetic systems with inverted meiosis, terminal fusion leads to completely different results than with normal meiosis (Figure 4.9, 4.10; Figure 4.8). If chiasmata are terminalised and there is no intrachromosomal recombination the result of terminal fusion is an exact restoration of the mothers genotype. Therefore, the combination of terminalised chiasmata, inverted meiosis and terminal fusion mimics apomixis as the outcomes of mitosis and of terminal fusion are identical. This type of parthenogenesis is known from some insects like the Homoptera (Suomalainen et al. 1987) and was also inferred from data for oribatid mites (Wrensch et al. 1994).

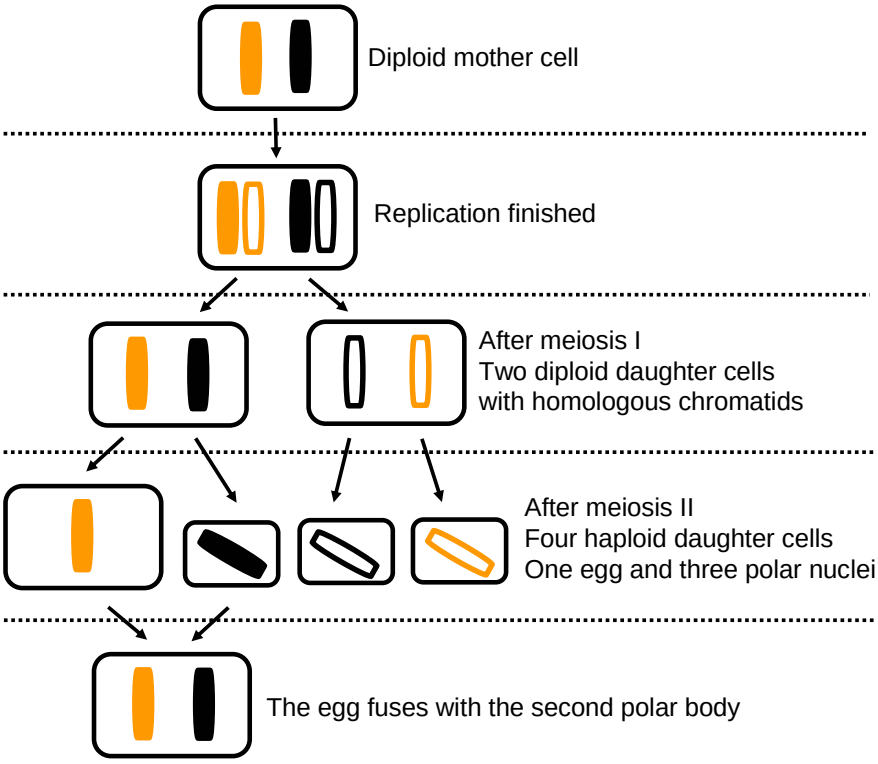


Figure 4.9

Automixis by terminal fusion with chiasma terminalisation, in the context of inverted meiosis. As recombination is absent except in the telomeric regions and no exchange of functional genes occurs the fusion of the egg nucleus with the second polar nucleus restores the genotype of the mother and therefore is equivalent to apomixis.

If chiasmata are not terminalised, the genetic constitution of the embryo depends on the segregation of chromosomes in meiosis I (Figure 4.10). Without terminalisation of chiasma, holokinetic chromosomes may also segregate in a normal meiosis instead of inverted meiosis (Nokkala and Nokkala 1997).

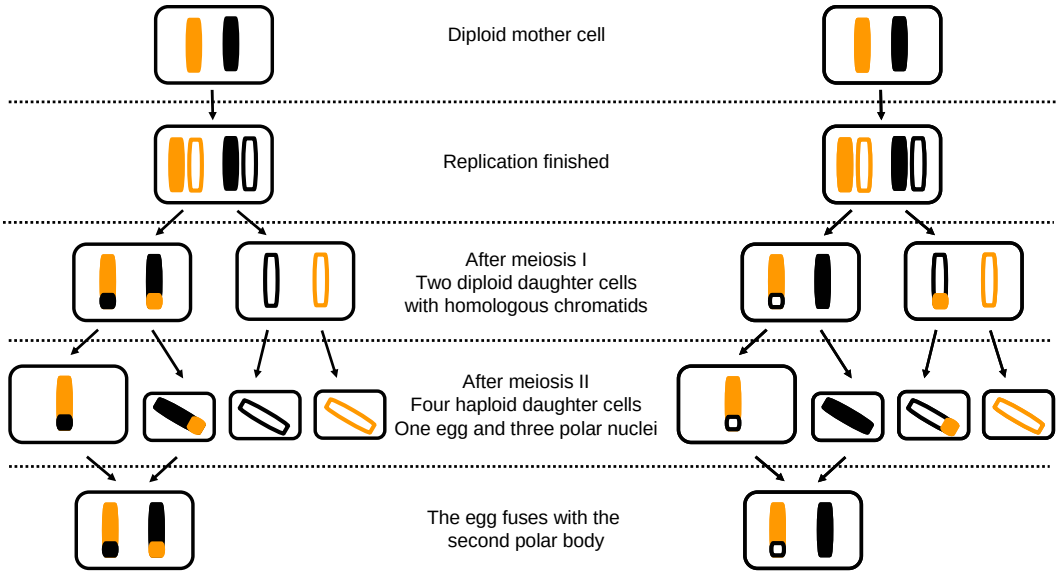


Figure 4.10

Automixis by terminal fusion without chiasma terminalisation, in the context of inverted meiosis. Gene recombination may occur, the genetic constitution of the embryo depends on segregation of chromosomes in meiosis I and may be complete heterozygous (left) or heterozygous except in cross over regions (right).

In some species of trematodes, crustaceans and insects the egg fuses with a polar body of the first meiotic division. The result is a heterozygous embryo except in cross over regions. From the chromosomal view, this mechanism resembles central fusion (Figure 4.7).

Another common mechanism to restore diploidy in automictic parthenogenesis is termed premeiotic doubling (Figure 4.11). Here, a second chromosome replication is added after the first has finished leading to a tetraploid cell. This cell undergoes a normal meiosis leading to a diploid egg and three diploid polar bodies which degenerate. Depending on the occurrence of cross over, the embryo is heterozygous except in these recombinant regions. Premeiotic doubling is a widespread mechanism and can be found in turbellaria, lumbricids, insects, tardigrads and also in vertebrates (Suomalainen et al. 1987).

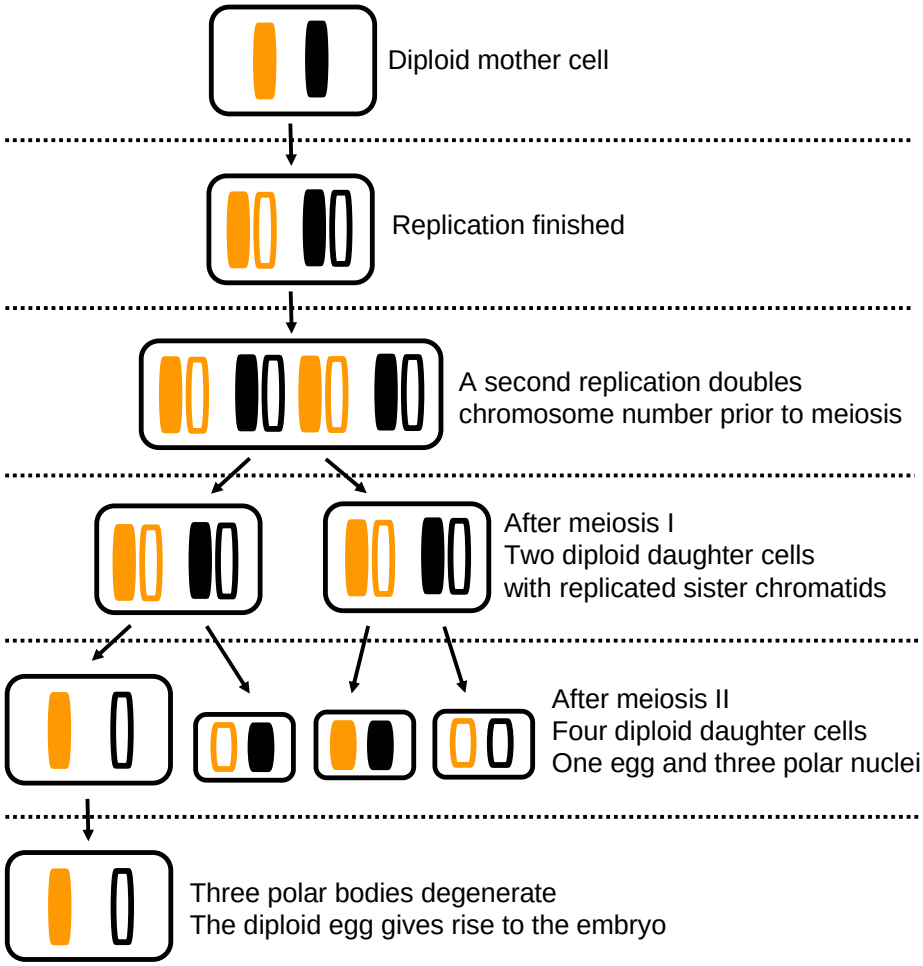


Figure 4.11

Premeiotic doubling. Before entering meiosis, a second chromosome replication occurs. The diploid embryo therefore develops without the fusion of any meiotic products.

These are not all of the diploidy restoring mechanisms exhibited by automictic organisms, but the other forms are so incompletely known that one can not ascribe them to any type of automictic parthenogenesis (Suomalainen et al. 1987). As the eggs of the automictic parthenogenetic species have undergone meiosis and, except in terminal fusion with inverted meiosis, cross over and segregation have taken place, they differ considerably from apomicts in both genetic and evolutionary aspects. However, many automictic organisms have evolved mechanisms to suppress recombination, and these can be referred to as functionally apomictic organisms (Lynch 1984).

4.5. Holokinetic chromosomes and inverted meiosis in oribatid mites

Taberly (1987) performed the single available cytological study of meiosis during oogenesis in oribatid mites, in which the parthenogenetic species *Platynothrus peltifer* and *Trhypochthonius tectorum* were examined. Although he did not use the terminology, both species restore diploidy by terminal fusion automixis. Taberly discussed both possibilities, normal meiosis and inverted meiosis, with the latter discussed under the term “post-réduction”. As documented in figures 4.8 and 4.9, the occurrence of normal or inverted meiosis could be tested by analysing if the embryos have fixed hetero- or homozygosity. If heterozygosity is fixed, the only possible explanation for terminal fusion automixis is inverted meiosis, since terminal fusion after normal meiosis leads to fixed homozygosity. This issue was first investigated by Palmer and Norton (1992) using isozyme techniques. They found fixed heterozygosity, absence of complete homozygosity and absence of recombination in nine parthenogenetic oribatid mite species. The findings were discussed on the basis of Taberly’s work and the authors wondered how fixed heterozygosity could occur in organisms with terminal fusion. In light of conflicting data, they assumed that apomixis or central fusion automixis should be more common in parthenogenetic oribatid mites, even if neither mechanism had been discovered. This point of view changed when Dana Wrensch and colleagues raised the idea of inverted meiosis in oribatid mites (Wrensch et al. 1994). Fixed heterozygosity and terminal fusion were brought together with the occurrence of holokinetic chromosomes. With this new idea, both Taberly’s work and that work of Palmer and Norton could be explained.

Investigations of the present study on the allelic divergence of *hsp82* (chapter 3.3.) also indicate fixed heterozygosity. However, this has to be analysed in more detail on the population level to prove if intrachromosomal recombination and sex are in fact absent in putative parthenogenetic oribatid mite species. The absence of recombination in the whole genome can be tested by analysing linkage disequilibrium of

different genes. Absence of sexual reproduction can be tested by analysing the allelic frequency distribution of a single locus like *hsp82*. For this locus, results of this study (chaper 3.3) suggest that intrachromosomal recombination was absent in both the sexual oribatid mite species *Steganacarus magnus* and *Metabelba pulverulenta* and the parthenogenetic species *Platynothrus peltifer*, *P. yamasakii*, *Mucronothrus nasalis*, *Atropacarus striculus* and *Tectocepheus velatus* since the evolution and radiation of oribatid mites about 350 million years ago. Possibly, however, *hsp82* is located near the telomeric region where recombination is inhibited and other more distant genes recombine. However, as stated above, a general lack of recombination was assumed by Palmer and Norton (1992). The findings strongly support the suggestion that, independent of the mode of reproduction, oribatid mites undergo inverted meiosis with chiasma terminalisation.

Without recombination and mixis, the whole genome serves as a single unit for selection and may evolve as a general purpose genotype (Lynch 1984). This is also true for apomictic organisms. As inverted meiosis with terminal fusion and without recombination mimics apomixis, parthenogenetic oribatid mites may be examples for general purpose genotypes. It remains unsolved why this type of automixis is maintained when it is identical to apomixis. Wrensch et al. (1994) stated "… Fusion of these segregants restores the maternal genotype insofar as there is no crossingover. By minimizing chiasmata, such a system would maximize conservation of the maternal genotype while providing the benefits of DNA repair… The terminal chiasma in holokinetic chromosomes does not result in the recombination of functional genes. …". I do not see scope for a different DNA repair in meiosis than in mitosis except in cross over regions. If chiasmata are terminalised and recombination is restricted to telomeres, this might be true for the telomeric regions only and is unlikely an evolutionary advantage of inverted meiosis over apomixis. However, the absence of intrachromosomal recombination in sexual species affects evolutionary theories as this is assumed to be one of the most important mechanisms to maintain sexual reproduction (Kondrashov 1993; West et

al. 1999). If recombination is really absent in sexual oribatid mite species this lack should facilitate the evolution of parthenogenetic lineages as the advantage of sexuality is reduced. In fact, about 10% of all oribatid mite species reproduce by parthenogenesis (Norton and Palmer 1991), whereas only about 1% of other eukaryotic taxa have abandoned sex (Norton et al. 1993). In oribatid mites, some parthenogenetic lineages are taxonomically scattered as expected from theory, but others are clustered without sexual relatives (Norton and Palmer 1991). The evolution of the high proportion of parthenogens in oribatid mites might be explained by their special genetic mechanism of meiosis. However, the maintenance of parthenogenesis in the long-term presumably has to be explained by ecological rather than genetic features. There is no strong correlation between habitat stability and reproductive mode in oribatid mites. Parthenogenetic oribatid mite species frequently occur in both stable habitats, such euedaphic soil horizons, and habitats of high temporal and spatial variability, such as early successional stages (temperate ecosystems as compared to the tropics) (Norton and Palmer 1991). Thelytoky is also predominant in freshwater habitats, probably due to a selection pressure against the ancestral spermatophore mating system (Norton et al. 1993).

There is no evidence that the reproductive mode is related to the strength of biotic interactions, such as predator/prey or host/parasite interactions (red queen; see chapter 1.3) in oribatid mites as it is predicted for typical parthenogens (Cianciolo and Norton unpublished). Although few studies exist, oribatid mites may serve as prey for a variety of vertebrate and invertebrate predators (e.g. Schuster 1966; Kupfer and Maraun 2003), as hosts for different parasitic fungi and protozoans (van der Geest et al. 2000), and as intermediate hosts for tapeworms (Anoplocephalidae, Denegri 1993; Trowe 1997). However, strong top-down control of oribatid mites is unlikely (Maraun 1997). Generally, selective forces driving most eukaryotic organisms to reproduce sexually seem to be reduced in soil, a stable and predictable habitat (Norton et al. 1993).

Sex ratios may be influenced by different mechanisms and it is known for haplodiploid systems that females can control the sex ratio of the offspring by providing sperm or not to the eggs (Hamilton 1993). In diplodiploid systems sex determination is usually by sex chromosomes and the ratio is assumed to be 1:1 (Fisher 1930). Sex determination may also be influenced by ecological factors, such as temperature (Ewert et al. 1994) or by hormonal or pheromonal determination (White 1973). Oribatid mites lack sex chromosomes despite their diplodiploidy and the sex determination is unknown (Wrensch et al. 1994). However, sexual oribatid mite species have sex ratios of 1:1 whereas males in parthenogenetic species are sporadic and sterile (Palmer and Norton 1992; Taberly 1988). A genetic mechanism based on inverted meiosis with terminal fusion automixis which mimics apomixis can not explain the occurrence of sporadic males. An ecological determination is also unlikely because sexual and parthenogenetic oribatid mite species coexist under the same ecological circumstances. Probably females can influence the sex ratio of the offspring; however, the mechanism for the sporadic production of sterile males remains to be uncovered.

The proportion of parthenogenesis in soil is very high although a review on this topic is lacking: in addition to several oribatid mite species many species of earthworms, nematodes and collembolans reproduce by parthenogenesis and these represent the most widespread and abundant organisms in soil (e.g. Bell 1982; Suomalainen 1987; Hughes 1989; Hopkin 1997; Sims and Gerard 1999; Reul 1999). It remains to be demonstrated which characteristics of the soil habitat favour the frequent occurrence of parthenogenetic species and higher taxa.

References

Adkins, R. M., Honeycutt, R. L., Disotell, T. R. (1996). Evolution of Eutherian Cytochrome c Oxidase Subunit II: Heterogeneous Rates of Protein Evolution and Altered Interaction with Cytochrome c. Mol. Biol. Evol. 13: 1393-1404.

Aguinaldo, A. M. A., Turbeville, J. M., Linford, L. S., Rivera, M. C., Garey, J. R., Raff, R. A., Lake, J. A. (1997). Evidence for a clade of nematodes, arthropods and other moulting animals. Nature 387: 489-492.

Albertson, D. G., Thomson, J. N. (1993). Segregation of holocentric chromosomes at meiosis in the nematode *Caenorhabditis elegans*. Chrom. Res. 1: 15-26.

Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., Lipman, J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25: 3389-3402.

Andersen, N. M., Cheng, L., Damgaard, J., Sperling, F. A. H. (2000). Mitochondrial DNA sequence variation and phylogeography of oceanic insects (Hemiptera: Gerridae: *Halobates* spp.). Marine Biology 136: 421-430.

Avise, J. C. (1994). Molecular markers, natural history and evolution. Chapman and Hall, London.

Bahler, J., Wyler, T., Loidl, J., Kohli, J. (1993). Unusual nuclear structures in meiotic prophase of fission yeast: a cytological analysis. J. Cell Biol. 124: 241-256.

Bardwell, J. C. A., Craig, E. A. (1987). Eukaryotic Mr 83,000 heat shock protein has a homologue in *Escherichia coli*. Proc. Natl. Acad. Sci. USA 84: 5177-5181.

Beard, C. B., Hamm, D. M., Collins, F. H. (1993). The mitochondrial genome of the mosquito *Anopheles gambiae*: DNA sequence, genome organization, and comparisons with mitochondrial sequences of other insects. Insect Mol. Biol. 2: 103-124.

Bell, G. (1982) The Masterpiece of Nature. Univ. Calif. Press, California.

Bendich, A. J., Drlica, K. (2000). Prokaryotic and eukaryotic chromosomes: what's the difference? BioEssays 22: 481-486.

Bernini, F. (1986) Current ideas on the phylogeny and the adaptive radiations of Acarida. Boll. Zool. 53 : 279-313.

Bernstein, C., Bernstein, H. (1991) Ageing, sex, and DNA repair. Academic Press.

Birky, C. W. (1996). Heterozygosity, heteromorphy, and phylogenetic trees in asexual eukaryotes. Genetics 144: 427-437.

Blackburn, E. H. (1991). Structure and function of telomeres. Nature 350: 569-573.

Blackburn, E. H. (1992). Telomerase. Annu. Rev. Biochem. 61: 113-129.

Blackman, R. K., Meselson, M. (1986). Interspecific nucleotide sequence comparisons used to identify regulatory and structural features of *Drosophila* hsp82 gene. J. Mol. Biol. 188:499-515.

Bogdanowicz, S. M., Schaefer, P. W., Harrison, R. G. (2000). Mitochondrial DNA variation among worldwide populations of gypsy moths, *Lymantria dispar*. Mol. Phyl. Evol. 15: 487-495.

Boore, J. L., Brown, W. M. (1995). Complete sequence of the mitochondrial DNA of the annelid worm *Lumbricus terrestris*. Genetics 141: 305-319.

Borresen A. L., Hovig, E., Smith-Sorensen, S., Malkin, D., Lystad, S., Andersen, T. I., Nesland, J. M., Isselbacher, K. J., Friend, S. H. (1991). Constant Denaturant Gel Electrophoresis as a rapid screening technique for p53 mutations. Proc. Natl. Acad. Sci. USA 88: 8405-8409.

Borresen-Dale, A. L., Lystad, S. Langerod, A. (1997). Temporal temperature gradient gel electrophoresis on the DCode™ system – A comparison to DGGE and CDGE in mutation screening. US/EG Bulletin 2133, BioRad.

Bridges, C. B. (1936). The bar “gene” a duplication. Science 83: 210-211.

Brower, A. V. Z. (1994). Rapid morphological radiation and convergence among races of the butterfly *Heliconius erato* inferred from patterns of mitochondrial DNA evolution. Proc. Natl. Acad. Sci. USA 91: 6491-6495.

Butlin, R., Schön, I., Martens, K. (1999). Origin, age and diversity of clones. J. Evol. Biol. 12 : 1020-1022.

Butlin, R. (2002). The costs and benefits of sex : new insights from old asexual lineages. Nat. Rev. Genet. 3: 311-317.

Cala, S. E., Jones, L. R. (1994). Grp94 resides within cardiac sarcoplasmic reticulum vesicles and is phosphorylated by casein kinase II. J. Biol. Chem. 269: 5926-5931.

Camacho, J. P. M., Belda, J., Cabrero, J. (1985). Meiotic behaviour of the holocentric chromosomes of *Nezara viridula* (Insecta, Heteroptera) analysed by C-banding and Silver impregnation. Can. J. Genet. Cytol. 27: 491-497.

Castresana, J., Lübben, M., Saraste, M., Higgins, D. G. (1994). Evolution of cytochrome oxidase, an enzyme older than atmospheric oxygen. EMBO Journal 13: 2516-2525.

Cavalier-Smith, T. (2002). Origins of the machinery of recombination and sex. Heredity 88: 125-141.

Clark, A. J., Margulis, A. D. (1965). Isolation and characterization of recombination in deficient mutants of *E. coli* K-12. Proc. Natl. Acad. Sci. USA 53: 451-459.

Clary, D. O., Wolstenholme, D. R. (1985) The mitochondrial DNA molecule of *Drosophila yakuba*: nucleotide sequence, gene organization, and genetic code. J. Mol. Evol. 22: 252-271.

Cockburn, A. F., Mitchell, S. E., Seawright, J. A. (1990). Cloning of the mitochondrial genome of *Anopheles quadrimaculatus*. Arch. Insect Biochem. Physiol. 14: 31-36.

Cooper, C. E., Nicholls, P., Freedman, J. A. (1991). Cytochrome c oxidase: structure, function and membrane topology of the polypeptide subunits. Biochem. Cell Biol. 69: 586-607.

Courcelle, J., Ganesan, A. K., Hanawalt, P. C. (2001). Therefore, what are recombination proteins there for? BioEssays 23: 463-470.

Dalenius, P. (1960). Studies on the Oribatei (Acari) of the Torneträsk territory in Swedish Lapland. A list of habitats, and the composition of their oribatid fauna. Oikos 2: 80-124.

Dechert, U., Weber, P., Konig, B., Ortwein, C., Nilson, I., Linxweiler, W., Wollny, E., Gassen, H. G. (1994). A protein kinase isolated from porcine brain microvessels is similar to a class of heat-shock proteins. Eur. J. Biochem. 225: 805-809.

de Lange, T. (1998). Ending up with the right partner. Nature 392: 753-754.

Denegri, G. M. (1993). Review of oribatid mites as intermediate hosts of tapeworms of the Anoplocephalidae. Exp. Appl. Acarol. 17: 567-580.

Dernburg, A. F. (2001). Here, there and everywhere: Kinetochore function on holocentric chromosomes. J. Cell Biol. 153: F33-F38.

DeSalle, R., Freedman, T., Prager, E. M., Wilson, A. C. (1987). Tempo and mode of sequence evolution in mitochondrial DNA of Hawaiian *Drosophila*. J. Mol. Evol. 26 : 157-164.

Dover, G. A. (1982). Molecular drive – a cohesive mode of species evolution. Nature 299: 111-117.

Drlica, K., Bendich, A. J. (2000). Chromosome, Bacterial. Enzyclopedia of Microbiology, Vol. 1, 2nd ed, Academic press, New York, pp. 808-821.

Egel, R. (2000). Fission yeast on the brink of meiosis. BioEssays 22: 854-860.

Evans, H. J., Pond, V. (1964). The influence of the centromere on chromosome fragment frequency under chronic irradiation. Port. Acta Biol. (A) 8: 125-146.

Ewert, M. A., Jackson, D. R., Nelson, C. E. (1994). Patterns of the temperature-dependant sex determination in turtles. J. Exp. Zool. 270: 3-15.

Felsenstein, J. (1981). Evolutionary trees from DNA sequences : A maximum likelihood approach. J. Mol. Evol. 17: 368-376.

Felsenstein, J. (1983). Parsimony in systematics: biological and statistical issues. Ann. Rev. Ecology Sys. 14: 313-333.

Felsenstein, J. (1988). Phylogenies from molecular sequences: inference and reliability. Annu. Rev. Genet. 22: 521-565.

Fischer, S., Lerman, L. (1979). Two-dimensional electrophoretic separation of restriction enzyme fragments of DNA. Methods in Enzymol. 68: 183-191.

Fisher, R. A. (1930). The genetical theory of sexual selection. Clarendon Press.

Flook, P.K., Rowell, C. H. F., Gellissen, G. (1995. The sequence, organisation and evolution of the *Locusta migratoria* mitochondrial genome. J. Mol. Evol. 41: 928-941.

Force, A., Lynch, M., Pickett, F. B., Amores, A., Yan, Y. L., Postlethwait, J. (1997). Preservation of duplicate genes by complementary, degenerative mutations. Genetics 151: 1531-1545.

Fuerst, J. A., Webb, R. I. (1991). Membrane-bounded nucleoid in the eubacterium *Gemmata obscuriglobus*. Proc. Natl. Acad. Sci. USA 88: 8184-8188.

Fuerst, J. A., Webb, R. I., Garson, M. J., Hardy, L., Reiswig, H. M. (1998). Membrane-bounded nucleoids in microbial symbionts of marine sponges. FEMS Microbiol. Lett. 166: 29-34.

Gessler, D. D. G., Xu, S. (1999). On the evolution of recombination and meiosis. Genet. Res. 73: 119-131.

Ghiselin, M. T. (1974). The economy of nature and the evolution of sex. Univ. Calif. Press, California.

Giles, R. E., Blanc, H., Cann, H. M., Wallace, D. C. (1980). Maternal inheritance of human mitochondrial DNA. Proc. Natl. Acad. Sci. USA 77: 6715-6719.

Giribet, G., Wheeler, C. (1999). The position of arthropods in the animal kingdom: Ecdysozoa, islands, trees, and the “parsimony ratchet”. Mol. Phyl. Evol. 133: 619-623.

Goldman, N. (1993). Statistical tests of models of DNA substitution. J. Mol. Evol. 36, 182-198.

Gorlov, I. P., Gorlova, Y. O. (2001). Cost-benefit analysis of recombination and its application for understanding of chiasma interference. J. Theoret. Biol. 213 : 1-8.

Grandjean, F. (1941). Statistique sexuelle et parthénogenèse chez les Oribates (Acariens). C. R. Seanc. Acad. Sci. 212: 463-467.

Grandjean, F. (1950). Observations éthologiques sur *Camisia segnis* (Herm.) et *Platynothrus peltifer* (Koch). Bull. Mus. Hist. Nat. Paris, 2^{e} sér. 22: 224-231.

Grandjean, F. (1969). Considérations sur le classement des Oribates. Leur division en 6 groupes majeurs. Acarologia 11: 127-153.

Griffiths, A. F. J., Miller, J. H., Suzuki, D. T., Lewontin, R. C., Gelbart, W. M. (1996). An introduction to genetic analysis (6th edition). W. H. Freeman and company, USA.

Gu, X. (1997). The age of the common ancestor of eukaryotes and prokaryotes: Statistical inferences. Mol. Biol. Evol. 14: 861-866.

Gu, Z., Steinmetz, L. M., Gu, X., Scharfe, C., Davis, R. W., Li, W. H. (2003). Role of duplicate genes in genetic robustness against null mutations. Nature 421: 63-66.

Gullan, P. J., Cranston, P. S. (1994). The Insects: An Outline of Entomololgy. Chapman and Hall, London.

Gupta, R. S. (1995). Phylogenetic analysis of the 90 kD heat shock family of protein sequences and an examination of the relationship among animals, plants, and fungi species. Mol. Biol. Evol. 12: 1063-1073.

Halkka, O. (1959). Chromosome studies on the Hemiptera, Homoptera, Auchenorrhyncha. Ann. Acad. Sci. Fenn. (Ser. A IV) 43: 1-72.

Hall, T. A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl. Acids Symp. Ser. 41: 95-98.

Hamilton, W. D. (1980). Sex versus non-sex versus parasites. Oikos 35: 282-290.

Hamilton, W. D. (1993). Inbreeding in Egypt and in this book: A childish perspective. In: Thornhill, N. W. (ed.). The natural history of inbreeding and outbreeding. Univ. Chicago Press, Chicago, pp. 429-450.

Hammer, M. (1965). Are low temperatures a species-preserving factor? Illustrated by the oribatid mite *Mucronothrus nasalis* (Willm.). Acta Univ. Lund. 2: 1-10.

Hammer, M., Wallwork, J. A. (1979). A review of the world distribution of oribatid mites (Acari: Cryptostigmata) in relation to continental drift. Biol. Skr. Dan. Vid. Selsk. 22: 1-31.

Harding, D. J. L. (1973). The phenology of Platynothrus peltifer (C. L. K.) In: Daniel, M., Rosicky, B. (eds) Proc. 3rd Congr. Acarol., Prague, pp 79-83.

Hardy, G. H. (1908). Mendelian proportions in a mixed population. Science 28: 49-50.

Harley, C. B., Villeponteau, B. (1995). Telomeres and telomerase in ageing and in cancer. Curr. Opin. Genet. Dev. 5: 249-225.

Harrison, R. G. (1989). Animal mitochondrial DNA as a genetic marker in population and evolutionary biology. Trends Ecol. Evol. 4: 6-11.

Hasegawa, M., Kishino, H., Yano. T. A. (1985). Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. J. Mol. Evol. 22: 160-174.

Haumann, G. (1991). Zur Phylogenie primitiver Oribatiden (Acari: Oribatida). Verlag Techn. Univ. Graz.

Heath, I. B. (1980). Variant mitoses in lower eukaryotes: indicators for the evolution of mitosis? Int. Rev. Cytol. 64: 1-80.

Heethoff, M., Maraun, M., Scheu, S. (2000). Genetische Variabilität der ribosomalen ITS1-Sequenzen bei der parthenogenetischen Hornmilbe *Platynothrus peltifer* (C.L. KOCH, 1839). Ber. Nat.-med. Verein Innsbruck 87: 339-354.

Heethoff, M., Etzold, K., Scheu, S. (2003). Mitochondrial COII sequences indicate that the parthenogenetic earthworm *Octolasion tyrtaeum* (Savigny 1826) constitutes of two lineages differing in body size and genotype. Pedobiologia 47, in press.

Helle, W., Bolland, H. R., Jeurissen, S. H. M., van Seventer, G. A. (1984). Chromosome data on the Actinedida, Tarsonemida and Oribatida. In: Griffiths, D. A., Bowman, C. E. (eds.). Acarology VI, Vol 1. Ellis Horwood, Chinchester, pp. 449-454.

Henikoff, S., Henikoff, J. G. (1992). Amino acid substitution matrices from protein blocks. Proc. Natl. Acad. Sci. USA 89: 10915-10919.

Higgins, D.G., Sharp, P.M. (1988). CLUSTAL: a package for performing multiple sequence alignment on a microcomputer. Gene 73: 237-244.

Higgins, D.G., Sharp, P.M. (1989). Fast and sensitive multiple sequence alignments on a microcomputer. CABIOS 5: 151-153.

Higgins, D.G., Bleasby, A.J. and Fuchs, R. (1992). CLUSTAL V: improved software for multiple sequence alignment. CABIOS 8: 189-191.

Hillis, M., Moritz, C., Mable, B. K. (1996). Molecular systematics. Sinauer Associates, Canada.

Hoffmann, T., Hovemann, B. (1988). Heat shock proteins *hsp84* and *hsp86* of mice and men: two related genes encode formerly identidied tumor-specific transplantation antigens. Gene 74: 491-501.

Holmquist, G., Dancis, B. M. (1980). A general model of karyotype evolution. Genetica 52/53: 151-163.

Hopkin, S. P. (1997). Biology of the springtails (Insecta: Collembola). Oxford Univ. Press, Oxford.

Horak, I., Coon, H. G., Dawid, I. B. (1974). Interspecific recombination of mitochondrial DNA molecules in hybrid somatic cells. Proc. Natl. Acad. Sci. USA 71: 1828-1832.

Hoyle, D. C., Higgs, P. G. (2003). Factors affecting the errors in the estimation of evolutionary distances between sequences. Mol. Biol. Evol. 20: 1-9.

Huelsenbeck, J. P., Crandall, K. A. (1997). Phylogeny estimation and hypothesis testing using maximum likelihood. Annu. Rev. Ecol. Syst. 28: 437-466.

Hughes, R. N. (1989) A functional biology of clonal animals. Chapman and Hall, London.

Hurst, G. D. D., Hurst, L. D., Majerus, M. E. N. (1993). Altering sex ratios: The games microbes play. BioEssays 15: 695-697.

Hurst, L. D., Smith, N. G. C. (1998). The evolution of concerted evolution. Proc. R. Soc. Lond. Ser. B 265: 121-127.

Ishikawa, F., Naito, K. (1999). Why do we have linear chromosomes? A matter of Adam and Eve. Mut. Res. 434: 99-107.

Jermiin, L. S., Crozier, R. H. (1994). The cytochrome b region in the mitochondrial DNA of the ant *Tetraponera rufoniger*: sequence divergence in hymenoptera may be associated with nucleotide content. J. Mol. Evol. 38: 282-294.

John, B. (1990). Meiosis. Cambridge Univ. Press, UK.

Judson, O. P., Normark, B. B. (1996). Ancient asexual scandals. Trends Ecol. Evol. 11: 41-46.

Jukes, T. H., Cantor, C. R. (1969). Evolution of protein molecules. In: Munro, H. N. (ed) Mammalian Protein Metabolism III. Academic Press, New York, pp 21-132.

Karppinen, E. (1958). Über die Oribatiden (Acer.) der finnischen Waldböden. Ann. Zool. Soc. „Vanamo“ 19: 1-43.

Kaul, S., Koo, H. L., Jenkins, J., Rizzo, M., Rooney, T., Tallon, L. J., Feldblyum, T., Nierman, W., Benito, M. I., Lin, X. Y., Town, C. D., Venter, J. C., Fraser, C. M., Tabata, S., Nakamura, Y., Kaneko, T., Sato, S., Asamizu, E., Kato, T., Kotani, H., Sasamoto, S., Ecker, J. R., Theologis, A., Federspiel, N. A., Palm, C. J., Osborne, B. I., Shinn, P., Conway, A. B., Vysotskaia, V. S., Dewar, K., Conn, L., Lenz, C. A., Kim, C. J., Hansen, N. F., Liu, S. X., Buehler, E., Altafi, H., Sakano, H., Dunn, P., Lam, B. (2000). Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature 408: 796-815.

Keyl, H. G. (1957). Zur Karyologie der Hydrachnellen (Acarina). Chromosoma 8:719-729.

Kimura, M., Crow, J. F. (1964). The number of alleles that can be maintained in a finite population. Genetics 49: 725-38.

Kimura, M. (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J. Mol. Evol. 16: 111-120.

Kitami, T., Nadeau, J. H. (2002). Biochemical networking contributes more to genetic buffering in human and mouse metabolic pathways than does gene duplication. Nat. Genet. 32: 191-194.

Kirkpatrick, D. T. (1999). Roles of the mismatch repair and nucleotide excision repair proteins during meiosis. Cell. Mol. Life Sci. 55: 437-449.

Knowlton, N., Weigt, L. A., Solorzano, L. A., Mills, D. K., Bermingham, E. (1993). Divergence in Proteins, Mitochondrial DNA, and Reproductive Compability Across the Isthmus of Panama. Science 260: 1629-1632.

Kondrashov, A. S. (1985). Deletirious mutations as an evolutionary factor. II. Facultative apomixes and selfing. Genetics 111: 635-653.

Kondrashov, A. S. (1988). Deleterious mutations and the evolution of sexual reproduction. Nature 336: 435-440.

Kondrashov, A. S. (1993). Classification of hypotheses on the advantage of amphimixis. J. Hered. 84: 372-387.

Kreitman, M. (1983). Nucleotide polymorphism at the alcohol dehydrogenase locus of *Drosophila melanogaster*. Nature 304: 412-417.

Kumar, S., Tamura, K., Jakobsen, I. B., Nei, M. (2001). MEGA2.1 : Molecular Evolutionary Genetics Analysis software. Bioinformatics 17: 1244-1245.

Kupfer, A., Maraun, M. (2003). *Ichthyophis kohtaoensis*: Diet. Herpetol. Rev. 34: 226.

Lai, B. T., Chin, N. W., Stanek, A. E., Keh, W., Lanks, K. W. (1984). Quantitation and intracellular localization of the 85K heat shock protein by using monoclonal and polyclonal antibodies. Mol. Cell. Biol. 4: 2802-2810.

Lerman, L. S., Silverstein, K. (1987). Computational simulation of DNA melting and its application to denaturing gradient gel electrophoresis. Meth. Enzymol. 155: 482-501.

Lewis, P. J. (2001). Bacterial chromosome segregation. Microbiology 147: 519-526.

Li, J. Y., Arnold-Schulz-Gahmen, B., Kellenberger, E. (1999). Histones and histone-like DNA-binding proteins: correlations between structural differences, properties and functions. Microbiology 145: 1-2.

Li, W. H. (1983). Evolution of duplicated genes. In: Nei, M., Koehn, R. K. (eds) Evolution of genes and proteins. Sinauer Associates, Sunderland, MA.

Li, W. H., Wu, C. I., Luo, C. C. (1985). A new method for estimating synonymous and nonsynonymous rates of nucleotide substitution considering the relative likelihood of nucleotide and codon changes. Mol. Biol. Evol. 2: 150-174.

Li, W. H., Sadler, L. A. (1991). Low nucleotide diversity in man. Genetics 129: 513-523.

Li, W. H. (1993). Unbiased estimation of the rates of synonymous and nonsynonymous substitution. J. Mol. Evol. 36: 96-99.

Lindquist, E. E. (1984). Current theories on the evolution of major groups of Acari and on their relationships with other groups of Arachnida, with consequent implications for their classification. In: Griffith, D. A., Bowman, C. E. (eds) Acarology VI (1). Ellis Horwood Publ., Chichester, England, pp 28-62.

Lindquist, S., Craig, E. A. (1988). The heat shock proteins. Annu. Rev. Genet. 22: 631-677.

Liu, L., Beckenbach, A. T. (1992). Evolution of mitochondrial cytochrome oxidase II gene among 10 orders of insects. Mol. Phyl. Evol. 1: 41-52.

Lodish, H., Baltimore, D., Berk, A., Zipursky, S. L., Matsudaira, P., Darnell, J. (1995). Molecular Cell Biology (3rd ed.). Scientific American Books, Inc., USA.

Long, M. (2001). Evolution of novel genes. Curr. Opin. Genet. Dev. 11: 673-680.

Lopez-Garcia, P., Moreira, D. (1999). Metabolic symbiosis at the origin of eukaryotes. Trends Biochem. Sci. 24: 88-93.

Lottspeich, F., Zorbas, H. (1998). Bioanalytik. Spektrum Akademischer Verlag, Heidelberg.

Lunt, D. H., Zhang, D. X., Szymura, J. M., Hewitt, G. M. (1996). The insect cytochrome oxidase I gene: evolutionary patterns and conserved primers for phylogenetic studies. Ins. Mol. Biol. 5: 153-165.

Lussenhop, J. (1992). Mechanisms of microarthropod-microbial interactions in soil. In: Begon, M., Fitter, A. H. (eds) Ecological Research 23, Academic Press, London, pp 1-33.

Luxton, M. (1972). Studies on the oribatid mites of a Danish beech wood soil. I. Nutritional biology. Pedobiologia 12: 434-463.

Lynch, M. (1984). Destabilizing hybridization, general-purpose genotypes and geographic parthenogenesis. The Quat. Rev. Biol. 59: 257-290.

Lynch, M., Conery, J. S. (2000). The evolutionary fate and consequences of duplicate genes. Science 290: 1151-1155.

Mandrioli, M., Ganassi, S., Bizzaro, D., Manicardi, G. C. (1999). Cytogenetic analysis of the holocentric chromosomes of the aphid *Schizaphis graminum*. Hereditas 131: 185-190.

Maraun, M. (1997). Hornmilben (Oribatida) in Buchenwäldern: Nahrungsbiologie und Einfluß auf Stoff-Flüsse. Berichte des Forschungszentrums Waldökosysteme, Reihe A, Bd. 143.

Maraun, M., Scheu, S. (2000). The structure of oribatid mite communities (Acari, Oribatida): patterns, mechanisms and implications for future research. Ecography 23: 374-383.

Maraun, M., Martens, H., Migge, S., Theenhaus, A., Scheu, S. (2003a). Adding to ‚the enigma of soil animal diversity': fungal feeders and saprophagous soil invertebrates prefer similar food substrates. Eur. J. Soil Biol. 39: 85-95.

Maraun, M., Heethoff, M., Scheu, S., Norton, R. A., Weigmann, G., Thomas, R. H. (2003b). Radiation in sexual and parthenogenetic oribatid mites (Oribatida, Acari) as indicated by genetic divergence of closely related species. Exp. Appl. Acarol. 29: 265-277.

Maraun, M., Heethoff, M., Schneider, K. Scheu, S., Weigmann, G. Cianciolo, J., Thomas, R. H., Norton, R. A. (submitted). Molecular phylogeny of oribatid mites (Oribatida, Acari): evidence for multiple radiations of parthenogenetic lineages.

Margolin, W. (1998). A green light for the bacterial cytoskeleton. Trends Micriobiol. 6: 233-238.

Mark Welsh, D. B. (1999). Evidence for the evolution of bdelloid rotifers without sexual reproduction or genetic exchange. Thesis, Harvard University, Cambridge, Massachusetts.

Mark Welsh, D. B., Meselson, M. (2000). Evidence for the evolution of bdelloid rotifers without sexual reproduction or genetic exchange. Science 288: 1211-1215.

Martens, K., Rossetti, G., Horne, D. J. (2003). How ancient are ancient asexuals? Proc. R. Soc. Lond. B. 270: 723-729.

Matthews, H. R. (1981). Chromatin patterns and progress through the cell cycle. In: John, P. C. L. (ed.). The Cell Cycle. Cambridge Univ. Press, UK, pp. 223-246.

Maynard Smith, J. (1978). The evolution of sex. Cambridge University Press, Cambridge.

McGuire, J. A., Poellinger, L., Wikstrom, A. C. (1992). Clonino and regulation by glucocorticoid receptor ligands of a hsp90. J. Steroid Biochem. Mol. Biol. 42: 813-822.

Michod, R. E. (1993). Genetic error, sex, and diploidy. J. Hered. 84: 360-371.

Michod, R. E. (1998). Origin of sex for error repair. Theor. Pop. Biol. 53: 60-74.

Milius, S. (2003). Life without sex. So, how many million years has it been? Science News 163(26): 406.

Mitchell, M. J. (1977). Lifehistory strategies of oribatid mites. In: Dindal, D. L. (ed) Biology of oribatid mites. SUNY-CESF, Syracuse, pp 65-69.

Miyamoto, M. M., Fitch, W. M. (1996). Constraints on protein evolution and the age of the eubacteria/eukaryote split. Syst. Biol. 45: 568-575.

Miyata, Y., Yahara, I. (1992). The 90-kDa heat shock protein, HSP90, binds and protects casein kinase II from self-aggregation and enhances its kinase activity. J. Biol. Chem. 267: 7042-7047.

Moriyama, E. N., Gojobori, T. (1992). Rates of synonymous substitution and base composition of nuclear genes in *Drosophila*. Genetics 130: 855-864.

Moriyama, E. N., Powell, J. R. (1996). Intraspecific nuclear DNA variation in *Drosophila*. Mol. Biol. Evol. 13: 261-277.

Müller, H. J. (1932). Some genetic aspects of sex. Am. Natl. 66: 118-138.

Muller, H. J. (1964). The relation of recombination to mutational advance. Mut. Res. 1: 2-9.

Müller, M. (1988). Enery-metobolism of protozoa without mitochondria. Ann. Rev. Microbiol. 42: 465-488.

Murphy, P. W., Jalil, M. (1964). Some observations on the genus *Tectocepheus*. Acarologia 6: 187-197.

Myers, R. M., Lerman, L. S., Maniatis, T. (1985). A general method for saturation mutagenesis of cloned DNA fragments. Science 229: 242-247.

Navajas, M., Gutierrez, J., Lagnel, J., Boursot, P. (1996). Mitochondrial cytochrome oxidase I in tetranychid mites : a comparison between molecular phylogeny and changes of morphological and life history traits. Bull. Entomol. Res. 86: 407-417.

Nei, M., Kumar, S. (2000). Molecular Evolution and Phylogenetics. Oxford University Press, New York.

Nokkala, S., Nokkala, C. (1997). The absence of chiasma terminalization and inverted meiosis on males and females of *Myrmus miriformis* Fn. (Coritidae, Heteroptera). Heredity 78: 561-566.

Nokkala, S., Laukkanen, A., Nokkala, C. (2002). Mitotic and meiotic chromosomes in *Somatochlora metallica* (Cordulidae, Odonata). The absence of localized centromeres and inverted meiosis. Hereditas 136: 7-12.

Nollen, E. A., Morimoto, R. I. (2002). Chaperoning signalling pathways: molecular chaperones as stress-sensing ‘heat shock’ proteins. J. Cell. Sci. 115: 2809-2816.

Normark, B. B., Judson, O. P., Moran, N. A. (2003). Genomic signatures of ancient asexual lineages. Biol. J. Linn. Soc. 79: 69-84.

Norton, R. A., Bonamo, P. M., Grierson, J. D., Shear, W. A. (1988). Oribatid mite fossils from a terrestrial Devonian deposit near Gilboa, New York. J. Paleont. 62: 421-499.

Norton, R. A., Williams, D. D., Hogg, I. D., Palmer, S. C. (1988b). Biology of the oribatid mite *Mucronothrus nasalis* (Acari: Oribatida: Trypochthomiidae) from a small coldwater springbrook in eastern Canada. Can. J. Zool. 66: 622-629.

Norton, R. A., Palmer, S. C. (1991). The distribution, mechanisms and evolutionary significance of parthenogenesis in oribatid mites. In: Schuster, R., Murphy, W. (eds) The Acari: reproduction, development and life-history strategies. Chapman and Hall, London, pp 107-136.

Norton, R. A., Kethley, J. B., Johnston, D. E., OConnor, B. M. (1993). Phylogenetic perspectives on genetic systems and reproductive modes of mites. In: Wrensch, D. L., Ebbert, M. A. (eds) Evolution and diversity of sex ratio. Chapman and Hall, New York, pp 8-99.

Norton, R. A. (1994) Evolutionary aspects of oribatid mite life histories and consequences for the origin of the Astigmata. In: Houck, M. A. (ed) Mites. Ecological and evolutionary analyses of life-history pattern. Chapman and Hall, New York, pp 99-135.

Norton, R. A., Behan-Pelletier, V. M., Wang, H. F. (1996). The aquatic oribatid mite genus Mucronothrus in Canada and western U.S.A. (Acari: Trhypochtoniidae). Can. J. Zool. 74: 926-949.

Nowak, M. A., Boerlijst, M. C., Cooke, J., Maynard Smith, J. (1997). Evolution of genetic redundancy. Nature 388: 167-171.

Nübel-Reidelbach, E. (1994). Taxonomie und Systematik der Gattung Tectocepheus Berlese, 1895 (Acari, Oribatei). Andrias 12: 3-94.

OConnor, B. B. (1984) Phylogenetic relationships among higher taxa in the Acariformes, with particular reference to the Astigmata. In: Griffith, D. A., Bowman C. E. (eds) Acarology VI (1). Ellis Horwood Publ., Chichester , England, pp. 19-27.

Page, R. D. M., Holmes, E. C. (1998). Molecular Evolution – A Phylogenetic Approach. Blackwell Science, Malden.

Palmer, S. C., Norton, R. A. (1990). Further experimental proof of thelytokous parthenogenesis in oribatid mites (Acari: Oribatida: Desmonomata). Exp. Appl. Acarol. 8: 149-159.

Palmer, S. C., Norton, R. A. (1991). Taxonomic, geographic and seasonal distribution of thelytokous parthenogenesis in the Desmonomata (Acari: Oribatida). Exp. Appl. Acarol. 12: 67-81.

Palmer, S. C., Norton, R. A. (1992). Genetic diversity in thelytokous oribatid mites (Acari; Acariformes; Desmonomata). Biochem. Sys. Ecol. 20:219-231.

Palumbi, S. R., Metz, E. C. (1991). Strong reproductive isolation between closely related tropical sea urchins (genus *Echinometra*). Mol. Biol. Evol. 8: 227–239.

Pamilo, P., Bianchi, N. O. (1993). Evolution of the Zfx and Zfy, genes: Rates and interdependence between the genes. Mol. Biol. Evol. 10: 271-281.

Pazy, B. (1997). Supernumerary chromosomes and their behaviour in meiosis of the holocentric *Cuscuta babylonica* Choisy. Bol. J. Linn. Soc. 123: 173-176.

Perez, R., Rufas, J. S., Suja, J. A., Page, J., Panzera, F. (2000). Meiosis in holocentric chromosomes : orientation and segregation of an autosome and sex chromosomes in *Triatoma infestans* (Heteroptera). Chrom. Res. 8: 17-25.

Philippe, H., Germot, A., Moreira, D. (2000). The new phylogeny of eukaryotes. Curr. Op. Genet. Dev. 10: 596-601.

Picard, D. (2002). Heat-shock protein 90, a chaperone for folding and regulation. Cell. Mol. Life Sci. 59: 1640-1648.

Pigozzi, M. I., Solari, A. J. (1999). The ZW pairs of two paleognath birds from two orders show transitional stages of sex chromosome differentiation. Chrom. Res. 7: 541-551.

Posada, D., Crandall, K. A. (1998). Modeltest: testing the model of DNA substitution. Bioinformatics 14: 817-818.

Poyton, R. O., Trueblood, C. E., Wright, R. M., Farrell, L. E. (1988). Expression and function of cytochrome c oxidase subunit isologues. Modulators of cellular energy production. Cytochrome oxidase: structure, function and physiology. Ann. N. Y. Acad. Sci. 550: 289-307.

Rebagliati, P. J., Mola, L. M., Papeschi, A. G. (2001). Karyotype and meiotic behaviour of the holokinetic chromosomes of six Argentine species of Pentatomidae (Heteroptera). Caryologia 54: 339-347.

Reul, C. (1999). Sexuelle Fortpflanzungsstrategien im Tierreich – Das Beispiel Nematoda. Examensarbeit, Institut Biologie II, Aachen.

Rhoades, M. M. (1961). Meiosis. In: Brachet, J., Mirsky, A. E. (eds.). The cell: Biochemistry, physiology, morphology, Vol. III, Meiosis and mitosis, pp. 1-75, Academic Press, NY.

Rice, W. R., Chippindale, A. K. (2000). Sexual recombination and the power of natural selection. Science 294: 555-559.

Ridley, M. (2000). Mendels's Demon: Gene justice and the complexity of life. The Guernsey Press Co. Ltd., Guernsey, UK.

Rieder, C. L. (1982). The formation, structure and composition of the mammalian kinetochore fibre. Int. Rev. Cyt. 79: 1-58.

Rodriguez, F., Oliver, J. L., Marin, A., Medina, J. R. (1990). The general stochastic model of nucleotide substitution. J. Theor. Biol. 142: 485-501.

Rokas, A., Ladoukakis, E., Zouros, E. (2003). Animal mitochondrial DNA recombination revisited. Trends Ecol. Evol. 18: 411-417.

Rosenbaum, V., Riesner, D. (1987). Temperature-Gradient Gel-Electrophoresis – Thermodynamic analysis of nucleic-acids and proteins in purified form and in cellular-extracts. Biophys. Chem. 26: 235-246.

Rubin, G. M., Yandell, M. D., Wortman, J. R., Gabor Miklos, G. L., Nelson, C. R., Hariharan, I. K., Fortini, M. E., Li, P. W., Apweiler, R., Fleischmann, W., Cherry, J. M., Henikoff, S., Skupski, M. P., Misra, S., Ashburner, M., Birney, E., Boguski, M. S., Brody, T., Brokstein, P., Celniker, S. E., Chervitz, S. A., Coates, D., Cravchik, A., Gabrielian, A., Galle, R. F., Gelbart, W. M., George, R. A., Goldstein, L. S., Gong, F., Guan, P., Harris, N. L., Hay, B. A., Hoskins, R. A., Li, J., Li, Z., Hynes, R. O., Jones, S. J., Kuehl, P. M., Lemaitre, B., Littleton, J. T., Morrison, D. K., Mungall, C., O'Farrell, P. H., Pickeral, O. K., Shue, C., Vosshall, L. B., Zhang, J., Zhao, Q., Zheng, X. H., Lewis, S. (2000). Comparative genomics of the eukaryotes. Science 287: 2204-2215.

Saiki, R. K., Scharf, S., Faloona, F., Mullis, K. B., Horu, G. T., Erlich, H. A., Arnheim, N. (1985). Enzymatic amplification of ß-globin sequences and restriction site analysis for diagnosis of sickle cell anemia. Science 230: 1350-1354.

Saitou, N., Nei, M. (1987). The neighbour-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4: 406-425.

Sala-Rovira, M., Geraud, M. L., Caput, D., Jaques, F., Soyer-Gobillard, M. O., Fernet, G., Herzog, M. (1991). Molecular cloning and immunolocalization of two variants of the major basic nuclear protein (HCc) from the histone-less eukaryote *Crypthecodinium cohnii* (Pyrrophyta). Chromosoma 100: 510-518.

Salomone, N., Frati, F., Bernini, F. (1996). Investigation on the taxonomic status of *Steganacarus magnus* and *Steganacarus anomalus* (Acari: Oribatida) using mitochondrial DNA sequences. Exp. Appl. Acarol. 20: 607-615.

Salomone, N., Emerson, B. C., Hewitt, G. M., Bernini, F. (2002). Phylogenetic relationships among the Canary Island Steganacaridae (Acari, Oribatida) inferred from mitochondrial DNA sequence data. Mol. Ecol. 11: 79-89.

Sandman, K., Pereira, S. L., Reeve, J. N. (1998). Diversity of procaryotic chromosomal proteins and the origin of the nucleosome. Cell. Mol. Life Sci. 54: 1350-1364.

Sandoval, C., Carmean, D. A., Crespi, B. J. (1998). Molecular phylogenetics of sexual and parthenogenetic *Timema* walking-sticks. Proc. R. Soc. Lond. B. 265: 589-595.

Schatz, H. (2002). Die Oribatidenliteratur und die beschriebenen Oribatidenarten (1758-2001) – Eine Analyse. Abh. Ber. Naturkundemus. Görlitz 72: 37-45.

Schelling, D. W., Jones, G. (1996). Analysis of induction of hsc70, hsp82, and juvenile hormone esterase genes by heat shock in *Trichoplusia ni*. J. Ins. Phys. 42: 295-301.

Scheu, S., Falca, M. (2000) The soil food web of two beech forests (*Fagus sylvatica*) of contrasting humus type: stable isotope analysis of a macro- and a mesofauna-dominated community. Oekologia 123: 285-296.

Schmidt, T. R., Jaradat, S. A., Goodman, M., Lomax, M. I., Grossman, L. I. (1997). Molecular Evolution of Cytochrome c Oxidase: Rate Variation Among Subunits Via Isoforms. Mol. Biol. Evol. 14: 595-601.

Schreiner, M., Geisert, M., Oed, M., Arendes, J., Gungerich, U., Breter, H. J., Stuber, K., Weinblum, D. (1995). Phylogenetic relationship of the green alga *Nanochlorum eukaryotum* deduced from its chloroplast rRNA sequences. J. Mol. Evol. 40: 428-442.

Sims, R. W., Gerard, B. M. (1999). Earthworms. The Linn. Soc. Lond. And the Estuarine and Coastal Sciences Ass., London.

Solari, A. J. (1995). Mitosis and genome partition in trypanosomes. Biocell 19: 65-84.

Solari, A. J. (2002). Primitive forms of meiosis: The possible evolution of meiosis. Biocell 26: 1-13.

Schön, I., Martens, K., Rossi, V. (1996). Ancient asexuals: scandal or artefact? Trends. Ecol. Evol. 11: 296-297.

Schön, I., Butlin, R. K., Griffiths, H. I., Martens, K. (1998). Slow molecular evolution in an ancient asexual ostracod. Proc. R. Soc. Lond. B. 265: 235-242.

Schon, I., Martens, K. (2003). No slave to sex. Proc. R. Soc. Lond. B. 270: 827-833.

Schroder, G., Beck, M., Eichel, J., Vetter, H. P., Schroder, J. (1993). HSP90 homologue from Madagascar periwinkle (*Catharanthus roseus*): cDNA sequence, regulation of protein expression and location in the endoplasmatic reticulum. Plant Mol. Biol. 23: 583-594.

Schuster, R. (1966). Scymaeniden-Larven als Milbenräuber. Naturwisscenschaften 53: 439-440.

Shear, W. A., Bonamo, M., Grierson, J. D., Rolfe, W. D. I., Smith, E. L., Norton, R. A. (1984). Early land animals on North America: evidence from Devonian age arthropods from Gilboa, New York. Science 224: 492-494.

Siepel, H. (1990). Niche relationships between two panphytophageous soil mites, *Nothrus silvestris* Nicolet (Acari, Oribatida, Nothridae) and *Platynothrus peltifer* Koch (Acari, Oribatida, Camisiidae). Biol. Fertil. Soils. 9: 139-144.

Simon, C., Frati, F., Beckenbach, A. Crespi, B., Liu, H., Flook, P. (1994). Evolution, weighting and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. Ent. Soc. Am. 87: 651-701.

Sharpe, M. E., Errington, J. (1999). Upheaval in the bacterial nucleoid. An active chromosome segregation mechanism. Trends Genet. 15: 70-74.

Stanton, D. J., Daehler, L. L., Moritz, C. C., Brown, W. M. (1994). Sequences with the potential to form stem-and-loop structures are associated with coding-region duplications in animal mitochondrial DNA. Genetics 137: 233-241.

Stearns, S. C. (1987). The evolution of sex and its consequences. Birkhäuser, Basel.

Steenstrup, J. J. H. (1842). Ueber den Generationenwechsel. Verlag C. A. Reitzel, Copenhagen.

Stephens, S. G. (1951). Possible significance of duplication in evolution. Adv. Genet. 4: 247-265.

Stern, H., Hotta, Y. (1987). The biochemistry of meiosis. In: Moens, P. B. (ed) Meiosis. Academic Press, NY, pp. 303-333.

Suomalainen, E., Halkka, O. (1963). The mode of meiosis in the Psyllina. Chromosoma 14: 498-510.

Suomalainen, E., Saura, A., Lokki, J. (1987). Cytology and Evolution in Parthenogenesis. CRC Press, Florida.

Swofford, D. (1999). PAUP*: Phylogenetic Analysis Using Parsimony (*and other methods). Version 4. Sinauer Associates, Sunderland.

Taanman, J. W., Williams, S. L. (2001). Assembly of cytochrome c oxidase : what can we learn from patients with cytochrome c oxidase deficiency? Biochem. Soc. Trans. 29 : 446-451.

Taberly, G. (1987). Recherches sur la parthéogenèse thélytoque de deux espèces d'acariens oribatides : *Trypochthonius tectorum* (Berlese) et *Platynothrus peltifer* (Koch). III. Etude anatomique, histologique et cytologique des femelles parthénogenétiques. Acarologia 28: 389-403.

Taberly, G. (1988). Recherches sur la parthéogenèse thélytoque de deux espèces d'acariens oribatides : *Trypochthonius tectorum* (Berlese) et *Platynothrus peltifer* (Koch). IV. Observations sur les males atavique. Acarologia 29: 95-107.

Tajima, F. (1993). Simple methods for testing molecular clock hypothesis. Genetics 135: 599-607.

Tamura, K., Nei, M. (1993). Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Mol. Biol. Evol. 10: 512-526.

Tautz, D., Arctander, P., Minelli, A., Thomas, R. H., Vogler, A. P. (2003). A plea for DNA taxonomy. Trends Ecol. Evol. 18: 70-74.

Thompson, J.D., Higgins, D.G., Gibson, T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. Nucl. Acids Res. 22: 4673-4680.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., Higgins, D.G. (1997). The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucl.Acids Res. 24: 4876-4882.

Travé, J. (1971). The specific stability and the chaetotactic variations in *Mucronothrus nasalis* (Willm.) (Oribatei). Proc. 3rd. Int. Congr. Acarology, Prague, 1971, pp 303-309.

Travé, J. (1973). Les variations chaetotaxiques dans quelques populations de *Mucronothrus nasalis* (Willm.) (Oribatei). Acarologia 15, 521-533.

Travé, J., André, H. M., Taberly, G., Bernini, F. (1996). Les Acariens Oribates : Éditions AGAR et SIALF, Belgique.

Trowe, S. (1997). Morphometrische Differenzierung von Anoplocephalidae-Zystizerkoiden mit einem Beitrag zur Reproduktionsfähigkeit experimentell infizierter Oribatiden. Dissertation, Freie Universität Berlin, Journal-Nr. 2100.

Van der Geest, L. P. S., Elliot, S. L., Breeuwer, J. A. J., Beerling, E. A. M (2000). Diseases of mites. Exp. Appl. Acarol. 24: 497-560.

Van Valen, L. (1973). A new evolutionary law. Evol. Theory 1: 1-30.

Vanzela, A. L. L., Caudrado, A., Jouve, N., Luceno, M., Guerra, M. (1998). Multiple locations of the rDNA sites in holocentric chromosomes of *Rhynchospora* (Cyperaceae). Chrom. Res. 6: 345-349.

Vera-Ziegler, H., Backeljau, T., Wauthy, G. (1990). Alloenzyme variation in two forest populations of *Platynothrus peltifer* (Acari, Oribatida). Biochem. Syst. Ecol. 18: 57-62.

Vogelstein, B., Gillespie, D. (1979). Preparative and analytical purification of DNA from Agarose. Proc. Natl. Acad. Sci. USA 76: 615-619.

Wagner, A. (2002). Selection and gene duplication: a view from the genome. Genome Biology 3: reviews 1012.1-1012.3.

Wallwork, J. A. (1983). Oribatids in forest ecosystems. Ann. Rev. Entomol. 28: 109-130.

Wang, R. L., Wakeley, J., Hey, J. (1997). Gene flow and natural selection in the origin of *Drosophila pseudoobscura* and close relatives. Genetics 147: 1091-1106.

Wang, Y., Gu, J., Gu, X. (2001). Patterns of functional divergence after gene duplication in vertebrate gene families: altered functional constraints. CBGIST 2001: 181-185.

Watson, J. D., Crick, F. H. C. (1953). Molecular structure of nucleic acids. Nature 171: 737-738.

Watson, J. D. (1972). Origin of cocatemeric T7 DNA. Nat. New Biol. 239: 197-201.

Wiech, H., Buchner, J., Zimmermann, R., Jakob, U. (1992). HSP90 chaperones protein folding in vitro. Nature 358: 169-170.

Wirth, T., Le Guellec, R., Veuillee, M. (1999). Directional substitution and evolution of nucleotide contents in the cytochrome oxidase II gene in earwings (Dermaptera, Insects). Mol. Biol. Evol. 16: 1645-1653.

Weigmann, G. (1996). Hypostome morphology of Malaconothridae and phylogenetic conclusions on primitive Oribatida. In: Mitchell, R., Horn, D. J., Needham, G. J., Welbourn, W. C. (eds) Acarology IX (1), Ohio Biol. Survey, Columbus, Ohio, pp. 273-276.

Wells, J. D., Sperling, F. A. (2000). A DNA-based approach to the identification of insect species used for postmortem interval estimation and partial sequencing of the cytochrome oxydase b subunit gene I: a tool for the identification of European species of blow flies for postmortem interval estimation. J. Forensic Sci. 45: 1358-1359.

West, S., Lively, C. M., Read, A. F. (1999). A pluralist approach to sex and recombination. J. Evol. Biol. 12: 1003-1012.

White, M. J. D. (1973). Animal Cytology and Evolution. Cambridge University Press, Cambridge.

Wiese, U., Wulfert, M., Prusiner, S.B., Riesner, D. (1995). Scanning for mutations in the human prion protein open reading frame by temporal temperature gradient gel electrophoresis. Electrophoresis 16: 1851-1860.

Williams, G. C. (1966). Adaptation and natural selection. Princeton Univ. Press, Princeton.

Wolstenholme, D. R. (1992). Animal mitochondrial DNA: structure and evolution. Int. Rev. Cyt. 141: 173-216.

Wrensch, D. L., Kethley, J. B., Norton, R. A. (1994). Cytogenetics of holokinetic chromosomes and inverted meiosis: keys to the evolutionary success of mites, with generalizations on eukaryotes. In: Houck, M. A. (ed) Mites: Ecological and Evolutionary Analyses of Life-history Pattern. Chapman and Hall, New York, pp 282-343.

Xia, X. (2000). Data analysis in molecular biology and evolution. Kluwer Academic Publishers, Boston.

Yang, Z., Goldman, N., Friday, A. (1994). Comparison of models for nucleotide substitution used in maximum likelihood phylogenetic estimation. Mol. Biol. Evol. 11: 316-324.

Yang, Z. (1996). Among-site variation and its impact on phylogenetic analyses. Trends. Ecol. Evol. 11: 367-371.

Yoshino, K., Nishigaki, K., Husimi, Y. (1991). Temperature sweep gel electrophoresis: a simple method to detect point mutations. Nucl. Acids. Res. 19: 3153.

Zhang, J. (2003). Evolution by gene duplication: an update. Trends Ecol. Evol. 18: 292-298.

Zoller, P., Redila-Flores, T., Chu, D., Patel, A. (1998). Temporal Temperature Gradient Electrophoresis – A powerful technique to screen mutation. Tech. note 2808, BioRad.

Acknowledgements

I would like to thank Stefan Scheu as my thesis supervisor. He was always a source of motivation, knowledge and ideas, he provided me with scope for development to realise my scientific ideas and he always found time and patience to keep in touch with my work.
Next, I have to express my deepest gratitude to Roy A. Norton. He never got tired of providing all imaginable kind of help and he shared all the ups and downs of my work. If I would print out all my emails he answered, this thesis would at least be doubled…
I'd like to acknowledge Uli Göringer for providing a substantial proportion of our lab equipment. Without his support molecular work would've been hardly possible in our lab. He always found time for discussions or commenting manuscripts.
For being on my examination committee and spending time to review this thesis I'd like to acknowledge Gerd Thiel and Barbara Drossel. Both of them were available for discussions and questions under different circumstances.
Friedrich Zimmermann provided me with an oceanic amount of money to set up a molecular laboratory when he left university. Without this initiating support and his confidence in my person none of my work would have been possible.

Special thanks go to Mark Maraun for nice discussions on all fields of scientific and non-scientific areas, he was always being helpful in all parts of my work.
Michael Lauman and Katja Domes both provided some of their data to this thesis. Michael analysed the allelic divergence of *Tectocepheus velatus*, Katja established the TTGE and provided *hsp82* cDNA sequences. Both suffered from my disaffection during the period of writing this thesis – I promise to make good for this.
Thanks to Reinhart Schuster (Graz), Heinz Schatz (Innsbruck), Valerie Behan-Pelletier (Ottawa), Nobuhiro Kaneko (Yokohama), Torstein Solhoy (Bergen), Gerd Weigmann (Berlin), Philippe Lebrun (Louvain), Fabio

Bernini (Siena), Katja Schneider (Darmstadt) and Sonja Migge (Göttingen) for providing specimens for analysis and Richard Thomas (London) and Isa Schön (Brussels) for providing primer sequences.
Constructive criticisms on chapter two of this thesis were offered by Joel Peck (Sussex), Pekka Pamilo (Uppsala), Alex Kondrashov (Bethesda), Roger Butlin (Leeds), Diethard Tautz (Cologne), Andrew Weeks (Victoria), Isa Schön (Brussels) and Benjamin Normark (Massachusetts).

Sven and Stevie – thanks for a nice time in room 261 and at any other place! Knut (Morsche), Jens and Olaf (Kippchen) – thanks for providing at least on joke per day. Michael L. and Katja D. – thanks for a nice cooperative time in the lab. Katja S. – thanks for your enthusiasm in keeping the lab free of contaminants…
All members of the Scheu group are acknowledged for always running away when I started to talk about inverted meiosis or genes and parthenogenesis and for being much more friends than colleagues ;-). I loved the years in this group – there was not a single day where I didn't like to go to work! Thank you all for everything.

My parents always gave me all kind of support (although they couldn't explain to me what this curious inverted meiosis is good for). Thanks a lot for all the years...

Finally, I'd like to thank Alex for being by my side and making every day of the last seven years something very special. Thanks for bearing my moods and providing me with love and patience…

Erratum

After this thesis and the disputation were finished and accepted by the faculty we realised that some sequences in the dataset may represent cross contaminations. Therefore, some of the interpretations in this thesis may be misleading and remain to be confirmed.