Alma Mater Studiorum – Università di Bologna

DOTTORATO DI RICERCA IN

BIODIVERSITA' ED EVOLUZIONE

Ciclo XXVIII

Settore Concorsuale di afferenza: 05/B1 - Zoologia e Antropologia

Settore Scientifico disciplinare: BIO/05 – Zoologia

SEX AND REPETITIVE SEQUENCE DYNAMICS IN *Bacillus* STICK INSECTS (PHASMIDA, BACILLIDAE)

Presentata da: Dott.ssa LIVIA BONANDIN

Coordinatore Dottorato
Prof.ssa Barbara Mantovani

Relatore
Prof.ssa Barbara Mantovani

Correlatore
Dr. Andrea Luchetti

Esame finale anno 2016

This thesis is for my family and for Davide,

who have never left me alone during this experience!

INDEX

ABBREVIATIONS

APE: apurinic/apyrimidinic endonuclease

cDNA: complementary DNA

DAPI: 4',6-Diamidino-2-Phenylindole Dihydrochloride.

DIRS: *Dictyostelium* intermediate repeat sequence

DNA: deoxyribonucleic acid

EN: endonuclease

ETS: external transcribed spacer

FISH: fluorescence in situ hybridization

GTM: genomic turnover mechanism

IGS: intergenic spacer

ITS: internal transcribed spacer

Kb: kilobases

LINE: long interspersed nuclear element

LTR: long terminal repeat

MITE: miniature inverted repeat

mRNA: messenger RNA

non-LTR: non long terminal repeat

OH: hydroxyl group

ORF: open reading frame

Pol: polymerase

rDNA: ribosomal DNA

RLE: restriction enzyme-like endonuclease

RNA: ribonucleic acid

RNP: ribonucleoprotein

rRNA: ribosomal ribonucleic acid

RT: reverse transcriptase

SINE: short interspersed nuclear element

tDNA: transfer DNA

TE: transposable element

TIR: terminal inverted repeat

TPRT: target primed reverse transcription

TSD: target site duplication

UTR: untranslated region

ZF: zinc finger motif

CHAPTER 1

INTRODUCTION

1.1 The Order Phasmida

The genus *Bacillus* Berthold, 1827 is included in the Order Phasmida, which comprises terrestrial species distributed over tropical and subtropical regions, well-known for their exceptional morphological and behavioural mimicry. These phytophagous insects in fact well reproduce the shape of branches and leaves and the movement caused by the wind on them; for this reason they are also known as stick and leaf insects, respectively.

In particular, stick insects are of medium or large size and are characterized by an orthopteroid body structure with the first abdominal segment fused with the metathorax in the median segment.

Their head has moniliform antennae composed up to 100 segments, anterolateral small compound eyes, eyespots (only in males of some winged species) and a masticatory apparatus.

The prothorax is less developed than the mesothorax and the metathorax, especially in the winged forms. Most of them are apterous and the winged ones are brachypterous. Legs are long, slender and adapted to walking. Aside from the mimicry, these insects adopted the autotomy of limbs as an anti-predatory strategy; the lost legs can be regenerated during the moulting.

Their walking habits, together with the parthenogenetic reproduction of many taxa, affects population structure because they may lead to the origin of a deme from a single couple of individuals or even from a single female (Scali and Mantovani, 1989).

The abdomen has 10 segments with two terminal cerci. Female genitalia are covered by the *operculum*, while male genitalia are covered by the *lamina subgenitalis* or *poculum* (Bragg, 1997).

Reproductive strategies include bisexuality and/or unisexuality; the egg development can undergo diapause induced by the photoperiod. The development is hemimetabolous; with respect to adults, the nymphs are characterized by a lower number of antennal segments, an incomplete development of genitalia, and if present at the adult stage, by the absence of eyespots and wings. Sexual dimorphism can be observed with females usually larger than males. Supernumerary moulting can occur also when the animal is at the adult stage.

As far as we know from the literature, until now the genus *Bacillus* seems to be the only one of the Order analysed in several aspects: field behaviour, reproductive biology, body/egg/sperm morphology, chromosome typing, electrophoretic multilocus analysis, molecular phylogeny on mitochondrial genes, satellite DNA characterization and vitellogenin comparisons (Scali and Mantovani, 1989; Scali *et al*., 2003).

1.1.1 The genus *Bacillus* (Phasmida, Bacillidae): morphology, ecology, taxonomy and distribution

The genus *Bacillus* includes apterous stick insects long up to 10 cm. All species are nocturnal. Some taxa are polyphagous and can feed on bramble (*Rubus spp*.), blackthorn (*Prunus spp*.), myrtle (*Myrtus communis*), buckthorn (*Rhamnus alaternus*) and lentisk (*Pistacia lentiscus*; Scali and Mantovani, 1989). Some species are instead monophagous, as the lentisk-feeder *B*. *atticus*. It appears that lentisk and myrtle could have been the original feeding plants, bramble becoming the feeding niche for some species only recently (Mantovani *et al*., 1991; Mantovani and Scali, 1993).

The genus *Bacillus* has a holomediterranean distribution mainly in the coastal regions (Figure 1).

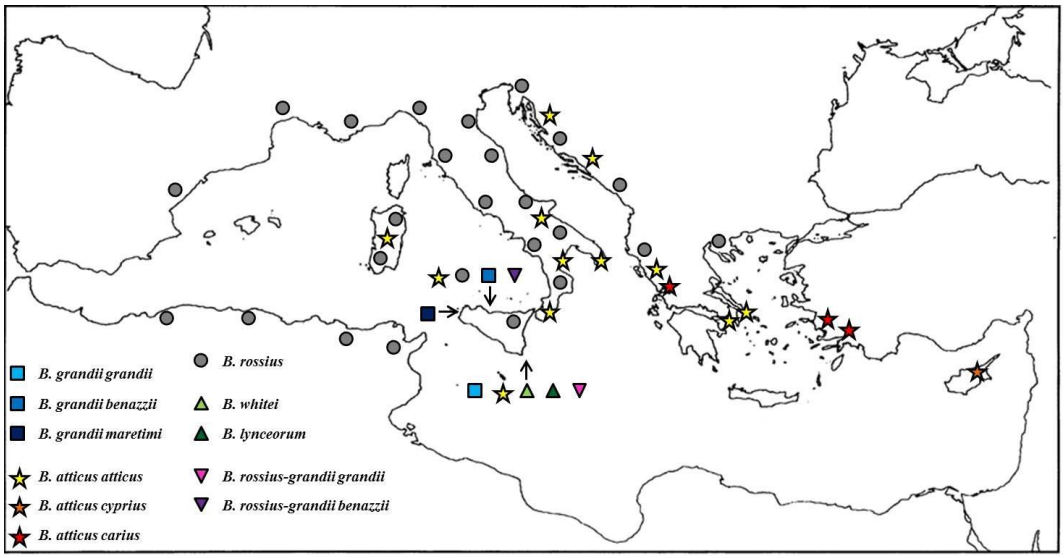


Figure 1: *Bacillus* stick insect distribution in the Mediterranean Basin (modified from Mantovani *et al*., 1999).

It comprises the strictly gonochoric *B*. *grandii*, the obligatory parthenogenetic *B*. *atticus* and the facultative parthenogenetic *B*. *rossius*.

The strictly gonochoric (i.e. with only bisexual populations) *B*. *grandii* Nascetti & Bullini, 1981 (Figure 2; 2n= 33/34, X0/XX) is endemic to the Sicilian area and includes three subspecies: *B*. *grandii grandii* Nascetti & Bullini, 1981; *B*. *grandii benazzii* Scali, 1991 and *B*. *grandii maretimi* Scali & Mantovani, 1990. *B*. *grandii grandii* is present with few relict populations in the Iblean area; *B*. *grandii benazzii* is located in North-Western Sicily and on Levanzo Island (Egadi Archipelago); *B*. *grandii maretimi* occurs on Marettimo Island (Egadi Archipelago; Figure 1; Mantovani, 1998).

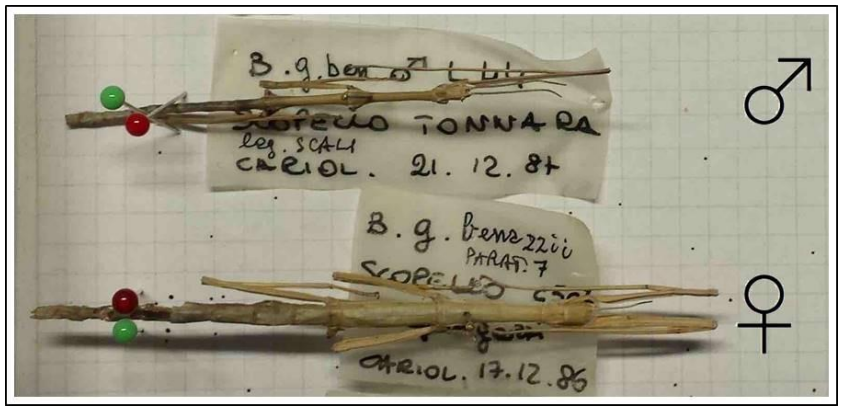


Figure 2: A male and a female specimens of the strictly gonochoric *B*. *grandii benazzii*. Sexual dimorphism can be observed. Photo taken at the Museum of Zoology, Alma Mater Studiorum-University of Bologna.

The obligate parthenogenetic (i.e. only all-females populations) *B*. *atticus* (Figure 3) Brunner, 1882 is distributed in the Central-Eastern region of the Mediterranean area with three karyological/allozymic subspecies: the diploids *B*. *atticus atticus* Brunner, 1882 and *B*. *atticus cyprius* Uvarov, 1936 (2n= 34 XX and 32 XX, respectively) and the diploid or triploid *B*. *atticus carius* Mantovani & Scali, 1985 (2n= 34 XX and 3n= 48-51 XXX). *B*. *atticus atticus* is widespread all through *B*. *atticus* range, being also the only Italian race. *B*. *atticus cyprius* is endemic to Cyprus and *B*. *atticus carius* is present with few populations in Turkey and Greece (Figure 1; Mantovani and Scali, 1993; Mantovani *et al*., 1999).

B. *atticus* parthenogenesis takes place with an automictic mechanism.

Figure 3: A specimen of the obligate parthenogenetic *B*. *atticus atticus* on *Pistacia lentiscus*. Photo taken at the MoZoo lab, BiGeA Department.

The facultative parthenogenetic (i.e. either bisexual populations or all-females ones) *B*. *rossius* Rossi, 1790 (Figure 4; 2n= 35/36 X0/XX) ranges in the Western Mediterranean area with its eight subspecies: *B*. *rossius tripolitanus* A Haan, 1842; *B*. *rossius tripolitanus* B Haan, 1842; *B*. *rossius lobipes* Lucas, 1849; *B*. *rossius montalentii* Bullini, 1982; *B*. *rossius medeae* Nascetti & Bullini, 1983; *B*. *rossius catalauniae* Bullini, 1982; *B*. *rossius rossius* Rossi, 1790 and *B*. *rossius redtenbacheri* Padewieth, 1899 (Figure 5).

Figure 4: A specimen of the facultative parthenogenetic *B. rossius rossius* on *Rubus spp*. Photo taken at the MoZoo lab, BiGeA Department.

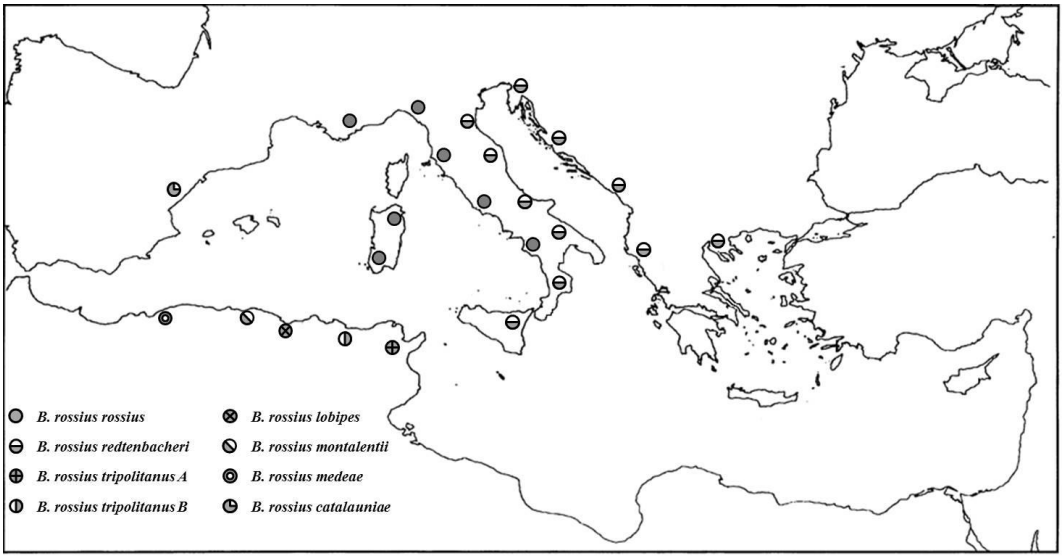


Figure 5: Distribution of *B. rossius* taxa in the Mediterranean basin (modified from Scali and Mantovani, 1989).

In Italy *B. rossius rossius* and *B. rossius redtenbacheri* occur. In particular, *B. rossius rossius* is distributed over the Northern Tyrrhenian coasts and in Sardinia, with the exception of the South-Eastern part of the region. In the latter, as in the other Italian area, *B. rossius redtenbacheri* is widespread (Mantovani *et al*., 1991). *B. rossius redtenbacheri* can be found also in Slovenia, Croatia, Bosnia, Albania and Greece. All the other races

instead, except the Spanish *B. rossius catalauniae*, are present in North-Africa. In Italy, a geographical variability can be observed about the reproductive modalities of the species: in fact, while Northern demes of either *B. rossius rossius* and *B. rossius redtenbacheri* consist mainly of all-females populations, in Southern Tyrrhenian, Adriatic and Ionian regions both bisexual and all-females populations have been observed (Scali, 2009). The North-African subspecies are all characterized by only bisexual populations (Mantovani *et al*., 1999). Bisexual demes are also present in the Western Mediterranean basin, as in Southern France and in the Spanish Catalaunia (Bullini, 1966; Tinti, 1993). Instead, the Eastern part is inhabited only by all-females populations.

In Sicily, 3 parental species gave origin to several hybrids (Figure 6).

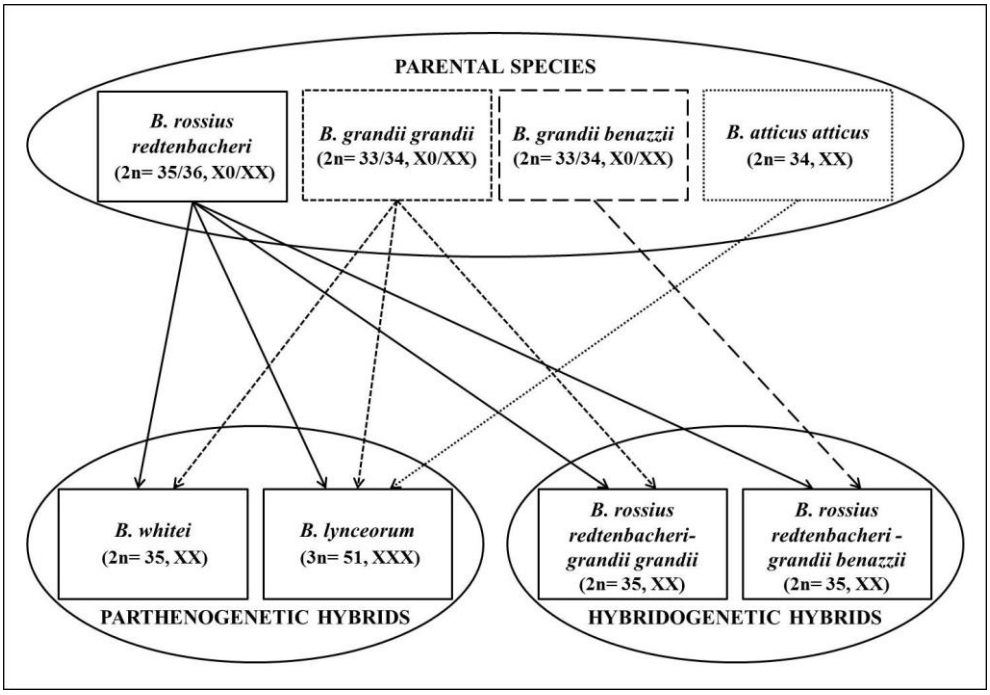


Figure 6: Reticulate evolution of the genus *Bacillus* showing the complex pattern of interactions among taxa and the origin of hybrid lineages (modified from Mantovani *et al*., 1999).

In particular, *B. rossius redtenbacheri*, *B. grandii grandii* and *B. atticus atticus* are the ancestors of two parthenogenons: the diploid *B. whitei* (Figure 7) Nascetti & Bullini, 1981 (*B. rossius redtenbacheri*/*B. grandii grandii*, 2n= 35 XX) and the triploid *B. lynceorum*

(Figure 8) Bullini, Nascetti & Bianchi Bullini, 1984 (*B*. *rossius redtenbacheri*/*B*. *grandii grandii*/*B*. *atticus atticus*, 3n= 51 XXX). Both reproduce by apomictic parthenogenesis.

Figure 7: A specimen of the facultative parthenogenetic *B*. *whitei* on *Rubus spp*. Photo taken at the MoZoo lab, BiGeA Department.

Figure 8: A specimen of the facultative parthenogenetic *B*. *lynceorum* on *Rubus spp*. Photo taken at the MoZoo lab, BiGeA Department.

B. *rossius* and *B*. *grandii* are also the ancestors of two hybridogenetic strains (*sensu* Schultz, 1961): *B*. *rossius redtenbacheri-grandii grandii* and *B*. *rossius redtenbacheri-*

grandii benazzii (Scali, 2009). Molecular mitochondrial analyses demonstrated that the female parental species for all hybrids is *B. rossius* (Mantovani *et al*., 2001).

1.1.2 The genus *Bacillus* (Phasmida, Bacillidae): reproductive biology

The reproductive biology of the genus *Bacillus* has been widely investigated with several approaches: field and laboratory observations, karyotype structures, egg meiotic process, phylogenetic relationships. The collected data revealed that *Bacillus* taxa, besides canonical gonochorism, can reproduce by sexual non-canonical strategies: these include geographic facultative and obligate thelytokous parthenogenesis, hybridogenesis and androgenesis (Scali, 2009). Their mechanism will be described in the next paragraphs.

The wide variety of reproductive modes, showed by the species of the genus *Bacillus*, defines a complex pattern of interactions called "reticulate evolution" (Figure 6; Scali, 2009).

Moreover, due to the variety of reproductive strategies and events of recombination that may occur in parthenogenesis, the definition "metasexual" has been suggested for the *Bacillus* species (Scali *et al*., 2003).

1.1.2.1 The facultative automictic parthenogenesis in *B. rossius*

Parthenogenesis is defined "facultative" when it occurs in a species together with bisexuality. It may characterize a part of the life cycle (cyclic parthenogenesis) or may happen in some populations of the species range where males are absent (geographic parthenogenesis). Through cyclic parthenogenesis both males and females can be produced, while geographic parthenogenesis is thelytokous, i.e. produces only female descendants. Facultative parthenogenesis may also be a way of sex determination as in the haplodiploid Hymenoptera where females are bisexually produced while males derive from unfertilized eggs through parthenogenesis (Tram and Sullivan, 2000).

B. *rossius* shows a facultative geographic parthenogenesis that takes place through one of the possible automictic mechanism: a canonical meiosis occurs forming an haploid blastula; only some of the thousands of cells, originated from egg segmentation, restore diploidy by anaphase restitution and contribute to the embryo development. Derived females therefore are genetically homozygous (Figure 9; Scali *et al*., 2003).

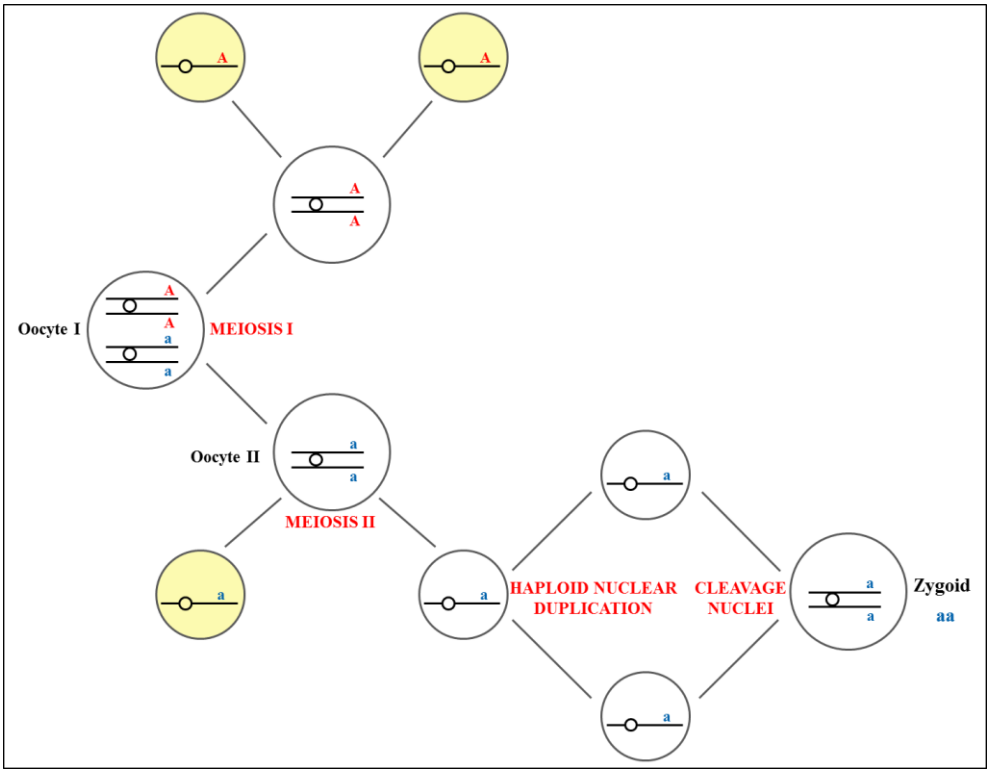


Figure 9: Mechanism of the facultative automictic parthenogenesis in *B*. *rossius* (modified from Scali *et al*., 2003).

Shifts from bisexual to all-females populations can easily take place with the production of polyclonal and homozygous progeny from maternal females heterozygous at some loci. If the reproduction of these females goes on by parthenogenesis, the clones will be genetically fixed (Scali *et al*., 2003 and reference therein).

1.1.2.2 The obligatory automictic parthenogenesis in *B. atticus*

When parthenogenesis represents the only way of reproduction of a species, we are dealing with obligatory parthenogenesis. The automictic mechanism of the obligatory parthenogenetic *B. atticus* follows more or less the same pattern in either diploid females or triploid ones (Figure 10).

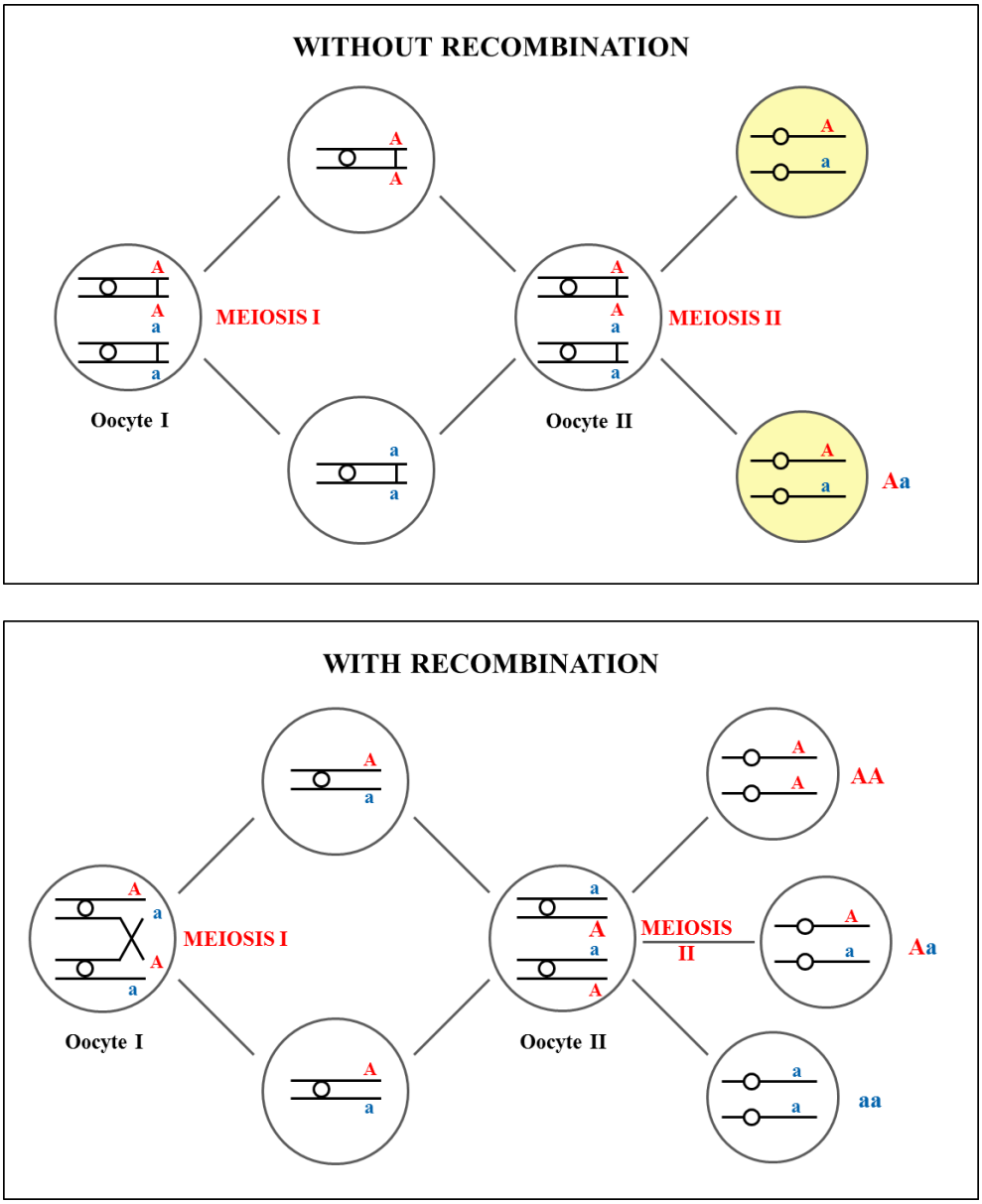


Figure 10: Mechanism of the obligatory automictic parthenogenesis in *B. atticus* (modified from Scali *et al*., 2003).

In particular, a standard first meiotic division occurs with the homologous chromosome segregation. The two haploid nuclei produced then fuse, giving origin to a diploid nucleus. The latter starts the second meiotic division that leads to a degenerative polocyte and a nucleus from which the embryo will develop.

In the first meiotic division of triploid females, asynaptic, heterologous or multivalent chromosomal associations can be observed together with canonical bivalents. The deriving unbalanced segregation is amended by the subsequent fusion of the two derived nuclei with the development of an unreduced zygoid nucleus (Scali *et al*., 2003).

The parthenogenetic mechanism of either diploid and triploid demes leads to the clonal maintenance of chromosomal rearrangements and of fixed heterozygosities at some loci. Further possible recombination events occurring during the first division lead to the production of homozygous genotypes from the heterozygous ones (Scali *et al*., 2003).

1.1.2.3 Hybrid reproductive mechanisms

The two parthenogenetic hybrids *B. whitei* and *B. lynceorum* are characterized by the same apomictic mechanism. The DNA extra-doubling occurring in prophase I gives origin to tetrachromatidic chromosomes forming autobivalent pairs. By mean of the two next divisions three polar bodies and an unreduced pronucleus are formed.

Owing to the DNA extra-synthesis, the produced thelytokous offspring presents most of the maternal fixed heterozygosities (Scali *et al*., 2003). However, some instances of low variability in the progeny can be observed: this could derive from recombination events during prophase I and it is probably involved in the origin of polyclonal populations (Mantovani *et al*., 1992; Scali *et al*., 1995).

In the two *Bacillus* hybridogenetic taxa, gametes are produced through the paternal nuclear genome elimination and the hemiclonal transmission of the maternal one. The hybrid

condition is restored through egg fertilization by sperms of the paternal parental species (Scali *et al*., 2003). The all-female *B*. *rossius redtenbacheri-grandii grandii* and *B*. *rossius redtenbacheri-grandii benazzii* strains show the hemiclonal transmission of the *B*. *rossius* maternal haploset, the paternal *B*. *grandii* one being eliminated. The maternal set is subsequently doubled, leading to a chromosomally balanced meiosis in the egg. Fertilization by syntopic *B*. *grandii* males will restore the hybrid condition (Scali *et al*., 2003).

Hybridogenetic females, thanks to the egg physiological polyspermy, can also reproduce by androgenesis. In this non-canonical strategy, observed in the animal kingdom for the first time in the genus *Bacillus* (Mantovani and Scali, 1992), the complete maternal genome is discarded; the fusion of two male pronuclei restores diploidy. This leads to a progeny of both sexes, with a nuclear *B*. *grandii* genome coupled with a *B*. *rossius* maternal mitochondrial genome (Scali *et al*., 2003 and reference therein). Lab-obtained androgenetic progeny is viable and fertile (Scali *et al*., 2003).

Some data on B. *whitei* suggest that also parthenogenetic taxa may show some instances of androgenesis (Mantovani and Scali, 1992; Scali, 2009).

1.2 Eukaryotic genomes and repetitive DNA

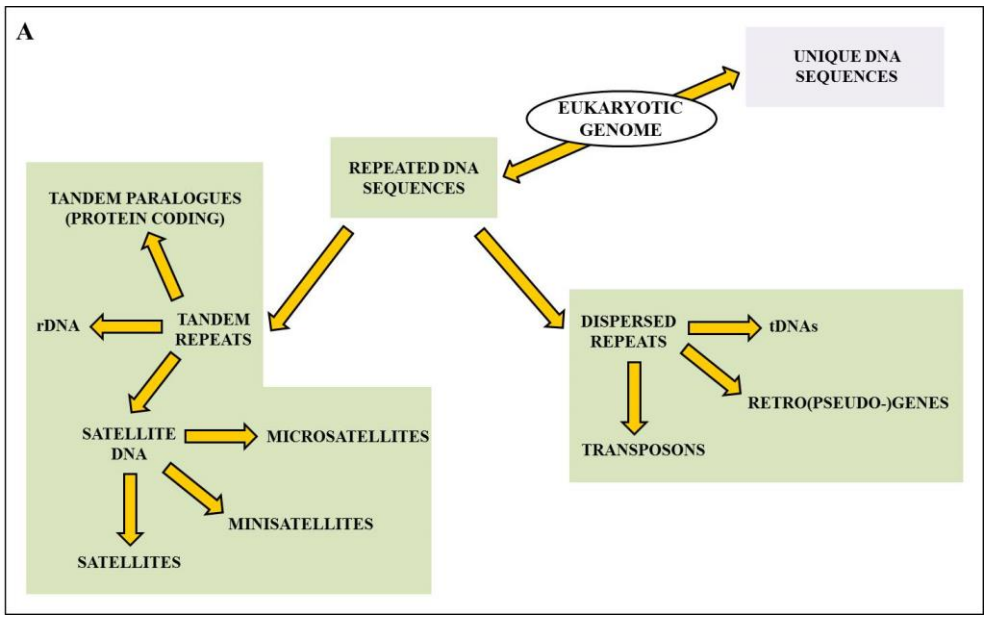
Starting from denaturation-renaturation experiments up to genome sequencing it appeared clear that a large portion of eukaryotic genomes is made up of repetitive DNA sequences (Richard *et al*., 2008).

The genomic content in repetitive elements may explain the "C-value paradox" (Richard *et al*., 2008). The "C-value" represents the DNA amount contained in the haploid nucleus of a cell. The definition of "C-value paradox" was coined because it was highlighted that the DNA content could be not directly proportional to the species complexity. For example, it

was observed that amphibians and fishes are characterized by up to 20 times more DNA per nucleus than mammals, wrongly supposed richer in gene number, or even more shocking that the unicellular *Amoeba dubia* presents a C-value 200 times higher than that seen in *Homo sapiens* (Richard *et al*., 2008).

In recent years, it has been understood that the C value depends on the content of repetitive sequences and not on the number of coding genes. In fact, the latter seems to be of the same order of magnitude in all eukaryotes: it ranges from roughly 6000 in the unicellular *Saccharomyces cerevisiae* to approximately 25.000 in the genome of *Homo sapiens*, which is 200 times bigger (Richard *et al*., 2008).

A possible very general classification of repetitive elements can be done according to their organization. In fact, they can be either "tandem repeats" or "dispersed repeats" (Figure 11A-B). Tandem repeats have a head-to-tail organization, while the dispersed ones are distributed in different sites of the genome.



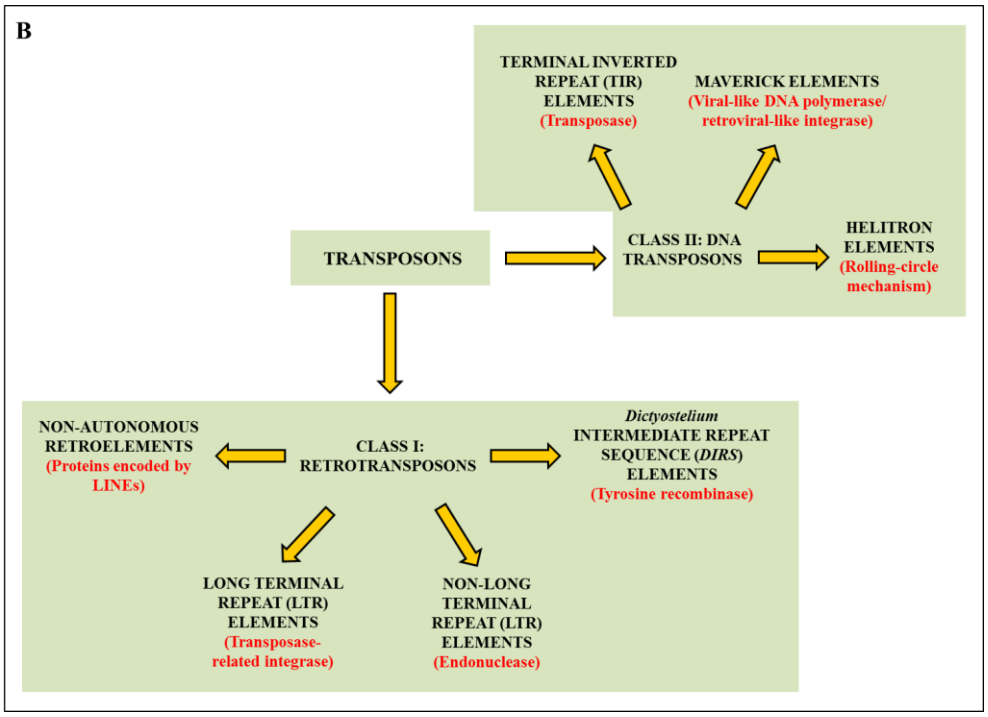


Figure 11: Repetitive DNA sequences in eukaryotic genomes. The classification in the two main families (tandem repeats and dispersed repeats) is shown, together with their subfamilies (A; modified from Richard *et al*., 2008). The classification of transposons based on the enzymes involved in their transposition mechanism, as reported by Chénais *et al*. (2012), is also represented (B).

At first, repeated elements were either indicated as "junk DNA" or "selfish DNA" mainly on the apparent lack of significance (Eddy, 2012; Dodsworth *et al*., 2015). At present we know that repetitive sequences range in their biological relevance for the host genome from structure (f.e. satellite DNAs in centromere) to variability (mobile elements) (Plohl *et al*., 2008, Chénais *et al*., 2012).

Given that my thesis will deal with the R2 non-long terminal repeat (non-LTR) retrotransposon specifically inserting into the rDNA 28S gene, I will detail on these two kinds of repetitive sequences.

1.2.1 Transposable elements

Transposable elements (TE) are mobile repetitive DNA sequences discovered for the first time in *Zea mays* by Barbara McClintock (McClintock, 1950). After this first characterization in the '40s, transposons were described in several eukaryotes (Richard *et al*., 2008; Chénais *et al*., 2012 and reference therein). Their occurrence in plant genomes is quite variable: in fact, their content range from ~65% in the rice to ~85% in the maize (Chénais *et al*., 2012 and reference therein). In the genome of animals, the percentage of TEs seems lower in invertebrate than in vertebrate, with values ranging from 12% in *Caenorhabditis elegans* and 15-22% in *Drosophila melanogaster* to 45% in *Homo sapiens* and to 77% in *Pelophylax esculentus* (Chénais *et al*., 2012 and reference therein).

The ability to move in different sites of the host genome is given by specific proteins, encoded by elements themselves, that mediate the transposition event; though, elements without coding capacity are also well represented. Therefore, based on the presence or absence of coding sequences, TEs are classified in autonomous and non-autonomous elements. To transpose, the latter rely on proteins coded by autonomous elements (Richard *et al*., 2008; Chénais *et al*., 2012). A schematic TE classification is given in Figure 11B.

1.2.2 Non-LTR retrotransposons

Non-LTR elements are class I transposons having length of 4-7 kb, characterized by the absence of LTRs. Usually they present target site duplication (TSD) at both ends, that are produced upon insertion, and a poly (A) tail at the 3' end (Fujiwara, 2015).

They move between different chromosomal sites by a "copy and paste" mechanism (Figure 12): new copies of DNA are synthesized from the messenger RNA (mRNA) by the reverse transcriptase (RT) encoded by the retrotransposons. The new copies of DNA may insert either randomly or specifically within the genome (Fujiwara, 2015).

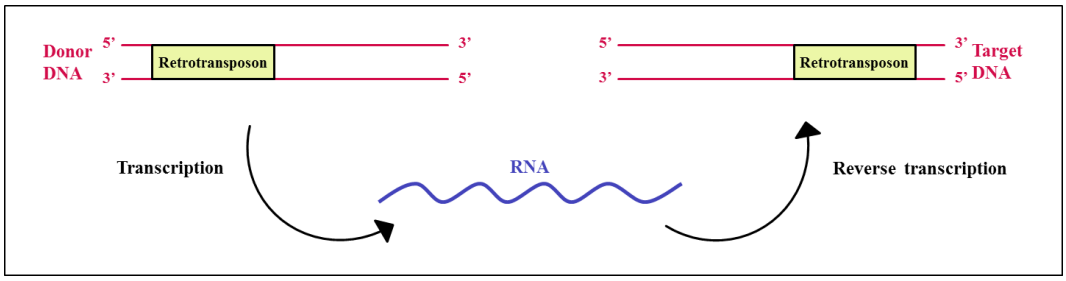


Figure 12: Replicative "copy and paste" mechanism of class I transposons.

Non-LTRs include the autonomous long interspersed elements (LINEs) and the non-autonomous short interspersed elements (SINEs) (Beauregard *et al*., 2008). The latter are unable to encode the enzymes that catalyze the transposition process and use the proteins produced by an autonomous partner LINE (Kajikawa and Okada, 2002; Dewannieux and Heidmann, 2005; Roy-Engel, 2012).

According to structural and phylogenetic features, non-LTRs are classified in two groups (Malik *et al*., 1999; Kapitonov *et al*., 2009). The first is characterized by a single open reading frame (ORF) with an N-terminal encoding DNA-binding motif, an RT and a C-terminal restriction enzyme-like endonuclease (RLE) domains (Eickbush and Jamburuthugoda, 2008; Fujiwara, 2015). This group is further divided in five clades (Fujiwara, 2015), each containing elements that insert in specific target sites (Figure 16A). The most studied element of this class is represented by R2 (Eickbush, 2002).

The second group of non-LTRs have two ORFs (ORF 1 and ORF2). ORF 1 is involved in the RNA binding, while ORF 2 presents RT and endonuclease (EN) domains (Eickbush and Jamburuthugoda, 2008 and reference therein). The EN domain of this group shows homology to apurinic/apyrimidinic endonuclease (APE). Contrary to RLE-encoding non-LTRs, only two (Tx1 and R1) out more than 20 clades include site-specific elements (Figure 16A-B; Fujiwara, 2015). The mammalian-wide LINE-1 (L1) is, probably, the best known retrotransposon of the group (Moran and Gilbert, 2002).

1.2.3 The R series and the R2 element

Retrotransposons of the R series insert specifically in rDNA genes. So far only two families have been detected in the 18S gene (R7 and R8) and seven in the 28S gene (R1, R2, R4, R5, R6, R9 and RT; Figure 13; Kojima *et al*., 2006; Eickbush and Eickbush, 2007; Gladyshev and Arkhipova, 2009).

R2 specifically inserts into the target 28S rDNA sequence 5'-TTAAGG↓TAGC-3'. To date, it has only been detected in the animal kingdom: in the diploblastic phylum Cnidaria and in the four triploblastic phyla Platyhelmintes, Arthropoda, Echinodermata and Chordata (Kojima *et al*., 2006 and reference therein).

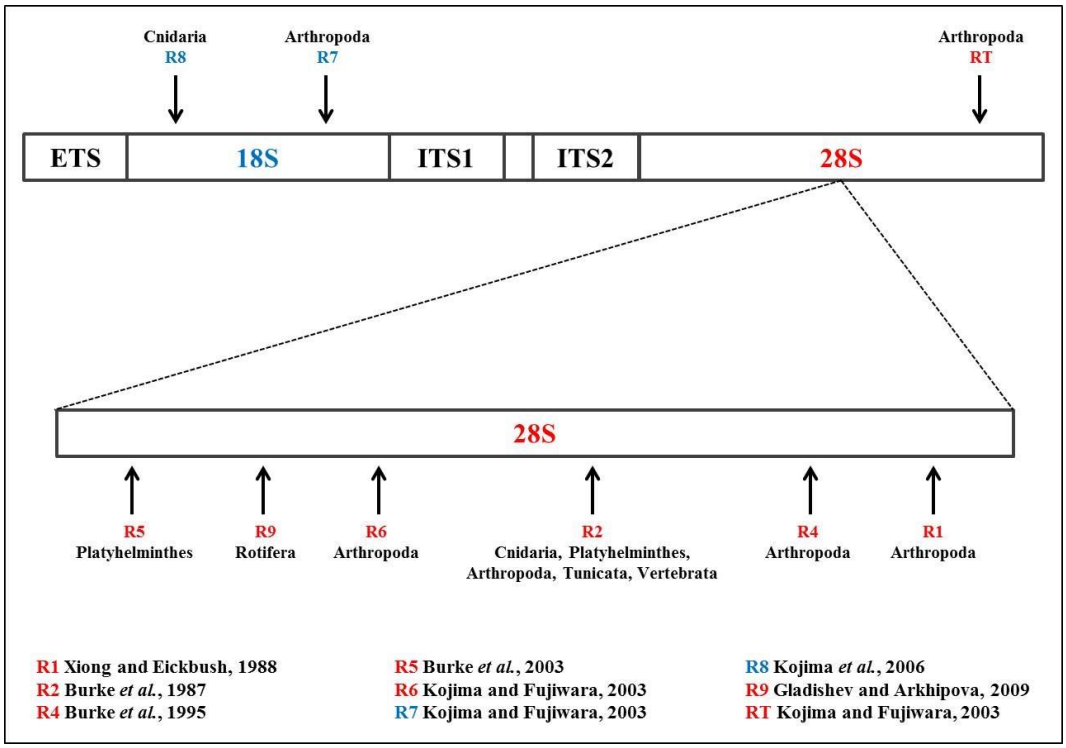


Figure 13: Non-LTR retrotransposons of the R series specifically inserting rDNA genes. For each family the transposition target in the 18S or 28S gene and the phyla in which they have been detected are indicated. References are reported at the bottom (modified from Eickbush and Eickbush, 2007).

1.2.3.1 R2 structure

R2 is characterized by the presence of a single ORF flanked by two untranslated regions (UTR) of variable length (Figure 14). The ORF includes the N-terminal, the RT and the C-terminal domains. The latter contains a zinc finger (ZF) motif (CCHC) and encodes the RLE, while the central RT domain the reverse transcriptase. The DNA binding N-terminal domain can include one (CCHH), two (CCHH + CCHH or CCHC + CCHH), or three (CCHH + CCHC + CCHH) ZF motifs. R2 phylogenesis is based on the number of N-terminal ZF motifs. Four clades have been defined: R2-A, R2-B, R2-C and R2-D. R2 elements belonging to R2-A, R2-C and R2-D clades show three (CCHH + CCHC + CCHH), two (CCHH + CCHH) and one (CCHH) ZF motifs, respectively (Kojima and Fujiwara, 2005; Luchetti and Mantovani, 2013). Only recently it has been demonstrated that the elements of the R2-B clade present two ZF motifs with a different combination than the R2-C clade: CCHC + CCHH (Luchetti and Mantovani, 2013).

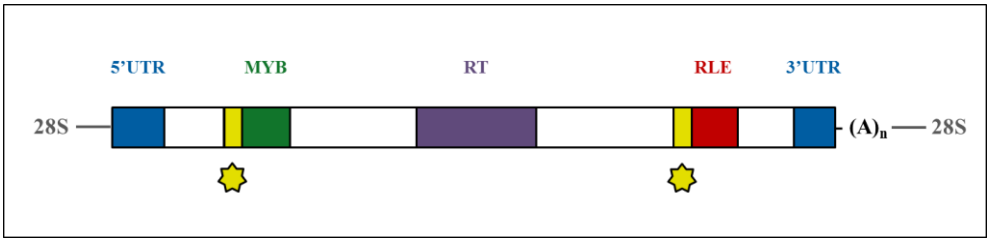


Figure 14: R2 structure. The blue boxes indicate the two flanking untranslated regions (UTRs). The single ORF is represented by the white boxes with the specific domains highlighted: the C-terminal domain (red box) coding the restriction enzyme-like endonuclease (RLE), the central domain (violet box) coding the reverse transcriptase (RT) and the N-terminal one (green box) binding DNA sequences by ZF motifs.
Yellow stars evidence the presence of ZF motifs contained in the C and N-terminal domains.

1.2.3.2 R2 retrotranscription mechanism

The R2 retrotransposition process is a copy and paste mechanism named "target primed reverse transcription" (TPRT) (Figure 15; Eickbush, 2002; Eickbush and Eickbush, 2010).

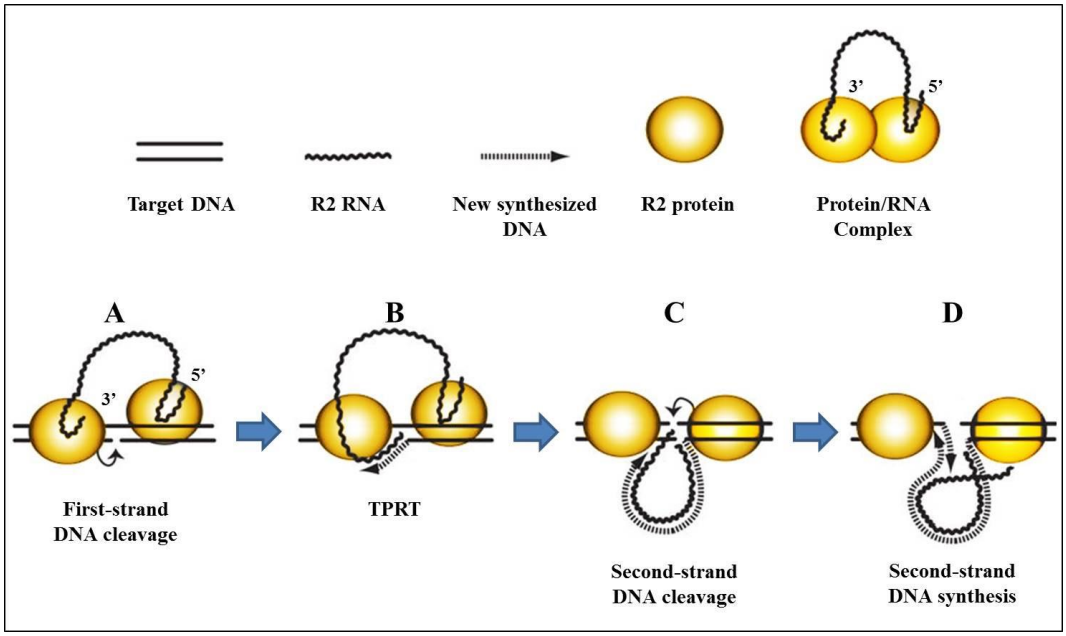


Figure 15: R2 TPRT mechanism (modified from Eickbush and Jamburuthugoda, 2008).

In the nucleus, the RNA polymerase (pol) I catalyzes the cotranscription of R2 and of the 28S gene. The R2 5' UTR acts a self-cleavage of the R2-28S cotranscript (Eickbush and Eickbush, 2010). The R2 transcript moves to the cytoplasm where it is translated. The resulting R2 protein includes a RT and an EN domain. R2 protein subunits bind either the 3' or the 5' end of the R2 transcript giving origin to a "ribonucleoprotein" (RNP) complex (Eickbush and Jamburuthugoda, 2008). The RNP complex is reintroduced into the nucleus. The R2 protein subunit that binds the 3' end of the R2 transcript will bind upstream the insertion site on the target DNA and cleave the lower strand through the endonuclease encoded from the C-terminal domain. The 3' OH, that derives from this nick, is used as primer for the reverse transcription of the RNA in the complementary DNA (cDNA). The R2 protein subunit binding the 5' end of the R2 transcript will bind downstream the

insertion site on the target DNA and cut the upper DNA strand. This leads to the synthesis of the second DNA strand, during which the disjunction between the first DNA strand and the R2 RNA occurs (Eickbush and Jamburuthugoda, 2008).

If during the TPRT the cellular degradation of the RNA transcript occurs or the reverse transcriptase fails to reach the 5' end, variable length deletions at the 5' end will be produced instead of full-length R2 elements (Christensen *et al*., 2006). However, according to the recurrent observation of inserted R2 copies in the genome with 5' end truncations, even if the synthesis of the cDNA is stopped before completing, the insertion can be completed (Burke *et al*., 1993).

5' truncated copies are used to determine R2 activity (Pérez-Gonzáles and Eickbush, 2001). In particular, the detection of new variants indicates that transposition events occurred and that the TE is active. On the contrary, the loss of pre-existing variants means that turnover mechanisms are actively eliminating the inserted units (Ghesini *et al*., 2011). Several studies, conducted on different organisms (f.e. *Drosophila* spp., *Blatella germanica*, *Triops cancriformis*, *Reticulitermes* spp., *Kalotermes* spp., and *Bacillus* spp.), supported the reliability of this approach to evaluate R2 dynamics (Pérez-González and Eickbush, 2001; Zhang and Eickbush, 2005; Kagramanova *et al*., 2007; Mingazzini *et al*., 2010; Ghesini *et al*., 2011). Nevertheless, the method showed two limits represented by the inability to discriminate a new insertion from a pre-existing one of the same length and the loss of a copy from a multicopy variant (Pérez-Gonzáles *et al*., 2003).

1.2.4 rDNA and TE evolutionary dynamics

Transposable elements whose integration occurs in specific sites often target multicopy genes, such as rDNA and telomeric sequences. Indeed, the high copy number is one of the main features that a chromosomal locus should have to become a potential specific retrotransposition target. TEs spreading and persistence within the genome, in this

instance, can be maintained as it is ensured that damages to single copy genes are avoided (Figure 16A).

This symbiotic strategy (*sensu* Fujiwara, 2015) diverges from the parasitic-like random pattern followed by most TEs (Figure 16B; Fujiwara, 2015 and reference therein).

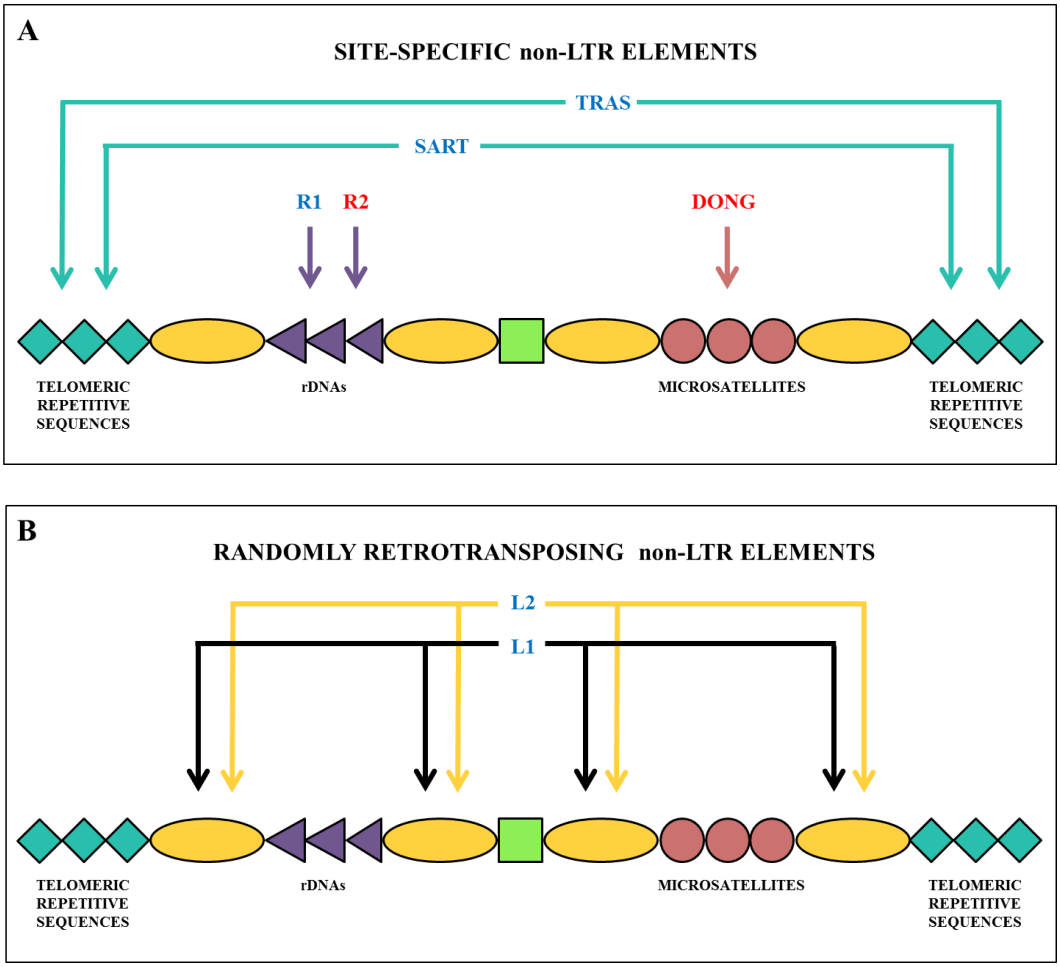


Figure 16: The integration preference of non-LTR retrotransposons in target DNA is shown. (A) restriction enzyme-like endonuclease (RLE)-encoding non LTRs (indicated in red) and few clades of apurinic/apyrimidinic endonuclease (APE)-encoding non LTRs (represented in blue) are characterized by specific insertion sites in the host genome, that are generally located in multicopy genes; elements specifically inserting in high copy number genes are defined as having a symbiotic strategy. (B) most of APE-encoding non-LTRs inserts in random chromosomal localizations with a parasitic-like trend (modified from Fujiwara *et al*., 2015).

The genes encoding rRNA are particularly prone to insertions of these non-LTRs. These tandemly repeated sequences are highly conserved genes among eukaryotes: in all lineages

the rDNA unit contains the 18S, 5.8S and 28S rRNA genes which codes for the small (18S) and large (5.8S and 28S) ribosome subunit RNAs (Figure 17).

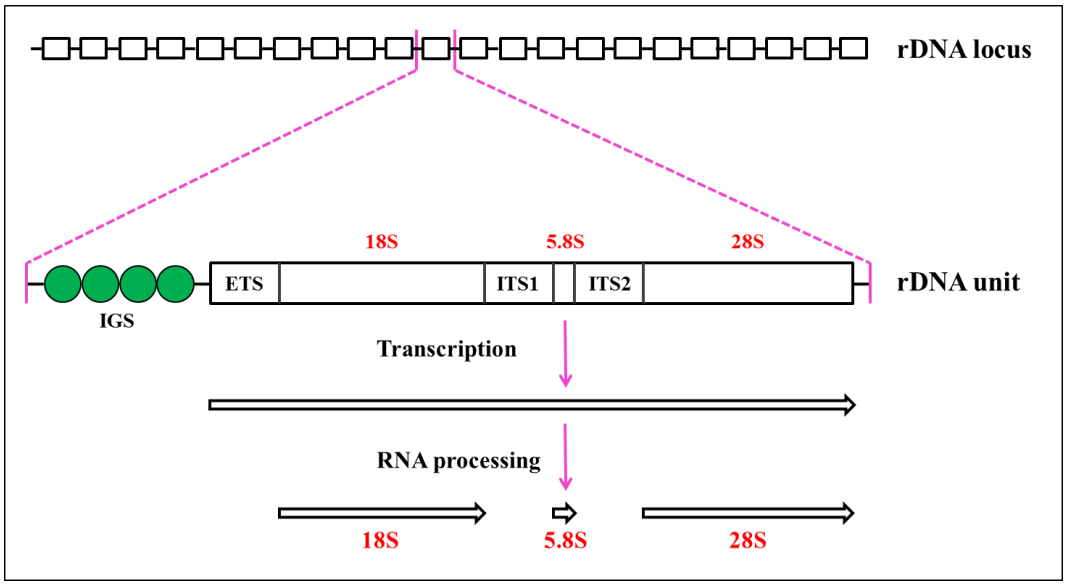


Figure 17: Eukaryotic rRNA genes. The organization of rDNA in tandemly repeated units and the structure of a single rDNA unit are described. Green circles represent the intergenic spacers (IGS) located between the different transcription units; the ETS and the ITS are indicated, as well as the three coding regions: 18S, 5.8S and 28S. With dashed arrows are shown the transcribed region and the mature rRNA, respectively (modified from Eickbush and Eickbush, 2007).

Every rDNA unit is separated from the others by intergenic spacers (IGS) that are generally made up of internally repeated sequences. In the rDNA unit, upstream the 18S gene there is an external transcribed spacer (ETS), while the three rRNA genes are separated by internal transcribed spacers (ITS) (Eickbush and Eickbush, 2007).

Owing to the abundance of ribosomes requested for protein synthesis in every process of development and growth, eukaryotes always host several copies of this transcription unit: rDNA units are, therefore, organized in tandem arrays, on one or few chromosomes, called rDNA locus(i). The latter will produce the nucleolus, which is visible during the transcription and maturation of rRNA molecules (Eickbush and Eickbush, 2007).

It has been observed that tandem repeats evolution show a peculiar pattern of sequence variability distribution, that is the non-independent evolution of each single unit. This evolutionary dynamics, has been explained by three main hypotheses: the concerted evolution, the birth-and-death and the magnification-and-fixation models (Dover, 1982; Nei and Rooney, 2005; Eickbush and Eickbush, 2007; Mukha *et al*., 2011).

In the '80s, Dover described the concerted evolution model trying to explain why rDNA units seem to vary their sequences in an orchestrated manner. He suggested that the concerted evolution of repetitive DNA monomers was the consequence of a two-level process, called "molecular drive", that leads to the homogenization of the mutations through the members of a repetitive family (genome level) and to their fixation in a group of reproductively linked organisms (population level; Dover, 1982, 1986).

In particular, homogenization can be the result of events of nonreciprocal transfer as: unequal crossing-over, gene conversion, rolling circle replication and reinsertion, and transposon-mediated exchange (Plohl *et al*., 2008 and reference therein). These mechanisms act with the highest effect on adjacent monomer subsets of a repeated sequence, while they show a progressively weak influence on arrays of the same chromosome, homologous and heterologous chromosomes (Figure 18A-B; Dover, 1986). Therefore, as predicted by theoretical models, monomers located at array ends, owing to the low force of the homogenization on them, will show less similarity than the centrally distributed ones (Plohl *et al*., 2008 and reference therein).

Instead, fixation depends on population characteristics and the main feature involved is represented by the bisexual reproduction. Indeed, the variant monomer spread among individuals is due to the random chromosomal meiotic segregation and to amphimixis (Figure 18C).

Thus, based on homogenization and fixation mechanisms, molecular drive is considered a stochastic event that allows the gradual but fast accumulation of mutations. As a result of

concerted evolution the homogeneity within lineages will be higher than between them (Plohl *et al*., 2008 and reference therein).

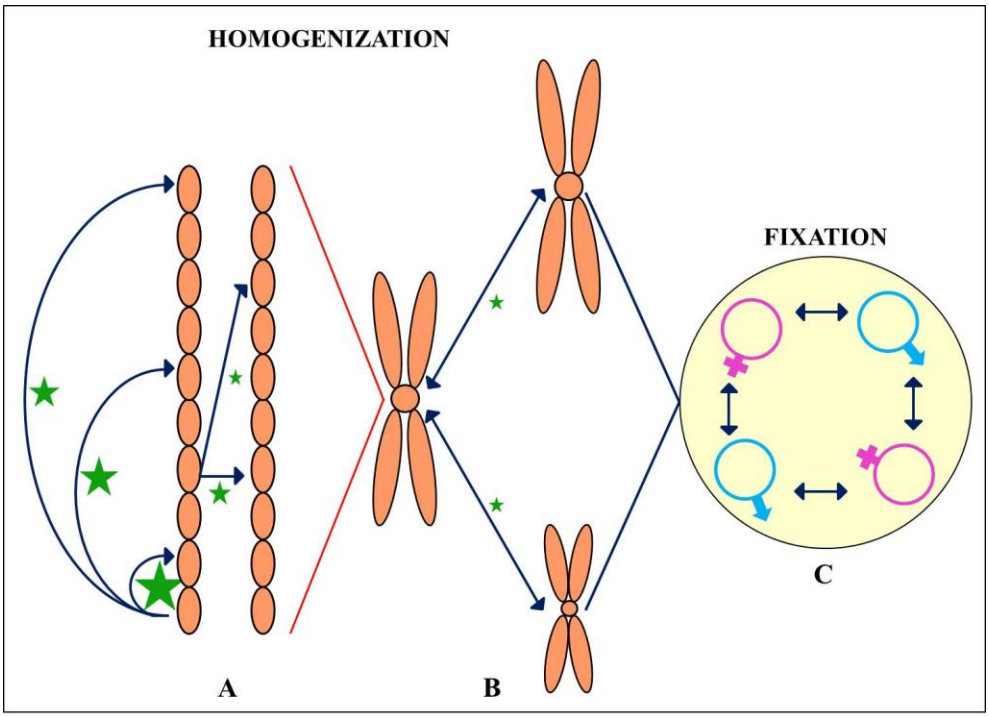


Figure 18: The concerted evolution. Molecular drive is a two-level process with homogenization of a mutation within the genome and its fixation through the population. (A) Homogenization effect on close monomers of a repeated sequence and on monomers distributed in sister chromatids. (B) Homogenization efficiency on monomers located in homologues and non-homologues chromosomes. Homogenization effectiveness is proportional to the size of green stars. (C) Fixation in bisexual organisms: reproduction isolation allows spreading in the population (modified from Plohl *et al*., 2008).

The birth-and-death (Nei and Rooney, 2005) model describes that tandem repetitive sequences undergo several duplications that give origin to new copies. Only some of these duplicated copies are conserved in the genome, the others will be inactivated or deleted. Consequently, the homogeneity of the repeated array could be more or less amplified (Nei and Rooney, 2005).

In the two steps magnification-and-fixation model, a quick amplification of repetitive sequences occurs (magnification) followed by their spreading through the array (fixation).

This mechanism, which seems to involve variants with adaptive importance, is led by events of gene conversion and of purifying selection (Mukha *et al*., 2011).

Mobile elements inserting in rDNA genes give origin to "variants" that can be used to analyse the concerted evolution of these loci (Eickbush and Eickbush, 2007). In particular, the loss of variants indicates that genomic turnover mechanisms (GTM) are actively eliminating the inserted units (Ghesini *et al*., 2011).

It has been also proved that high rates of R2 retrotranscription activity lead to the deletion of loci located upstream the new insertions and so to the elimination of possible pre-existing R2 variants (Zhang *et al*., 2008; Zhou and Eickbush, 2009). Inter-chromosomal and sister-chromatid unequal crossing overs, together with the subsequent selection of longer variants, re-establish the size of the rDNA array (Ghesini *et al*., 2011). Thus, the TE dynamics appears to be related to the distribution of R2 inserted units in the rDNA locus and it is not dependent on the length of the rDNA locus or on the abundance of uninserted units (Eickbush *et al*., 2008; Zhou and Eickbush, 2009).

However, the replacement of inserted loci with uninserted ones, by means of concerted evolution, is an unstable condition, since novel sites can become the target of new insertions and thus R2 remains active (Eickbush and Eickbush, 2007).

As suggested by Eickbush *et al*. in (2008), high contents of functional R2 insertions in transcriptionally active rDNA copies may be compartmentalized in restricted heterochromatinizable rDNA regions through the rDNA rearrangements resulting from the GTM. The silencing of densely inserted rDNA arrays may derive also by the nucleolar dominance occurring after the amphymixis with a partner bearing an R2-silenced rDNA array.

1.2.5 TEs and reproductive biology

Given their ability to reduce host fitness and because of their independent activity, TEs can be in general terms considered either deleterious mutations (under the Muller's ratchet theory) or "genomic parasites" (following the Red Queen hypothesis; Capy *et al*., 2000; Dolgin and Charlesworth, 2006; Lee and Langley, 2010).

In bisexual organisms, in the light of the Muller's ratchet theory, the host genome tries to limit the negative effects of their proliferation through events of recombination and of independent assortment of chromosomes (Felsenstein, 1974). On the other hand, on the basis of the Red Queen hypothesis TEs increase their copy number in host genomes faster also thanks to the "copy and paste" replicative mechanism (see section 1.2.3.2; Hamilton, 1975; Dolgin and Charlesworth, 2006). However, bisexual reproduction can both control and favor TEs spread. In fact, the above cited mechanisms set up by the host organism produce high variance and allow natural selection to more easily counteract their increasing load because a TE present in a heterozygous state is transmitted only to half of the progeny (Arkhipova, 2005; Dolgin and Charlesworth, 2006; Loewe and Lamatsch, 2008; Mingazzini, 2011; Ph.D. thesis). In the same way, TEs can disperse because the fusion of two haploid genomes helps the TEs spreading in transposition targets of "virgin" genomes (Hickey, 1982; Arkhipova, 2005). The mating between individuals with and without TE copies, implies the possibility for the TE to insert in the half-genome without variants (Zhang and Eickbush, 2005). TE can therefore be fixed in a population.

The scenario in the case of unisexual reproduction is rather different. It is worth noting that unisexuals are low or no-recombining organisms (parthenogenetic genomes); therefore following the Muller's ratchet rationale, they should accumulate TE insertions without the possibility of eliminating them eventually leading to lineage extinction by overloading the genome of deleterious insertions. It is, thus, assumed that TEs are not present in unisexuals, because of genome purging (Loewe and Lamatsch, 2008), unless the

occurrence of horizontal transmission events that allow them to spread (Kordis and Gubensek, 1999; Lampe *et al*., 2003; Sanchez-Gracia *et al*., 2005; Keeling and Palmer, 2008). Therefore, the survival of unisexuals bearing TEs seems to be linked to mechanisms able to keep their load low (Bull *et al*., 1991; Wright and Finnegan, 2001; Eickbush *et al*., 2008).

Theoretical studies evidenced a causal link between TEs load and the reproductive biology of the host organism (Wright and Finnegan, 2001; Nuzhdin and Petrov, 2003). They showed how the reproduction can play an important role in the fate of genomic parasites, as retrotransposons can be considered, and in the genome structuring. In particular, they proposed that the ability of TEs to increase their transposition rate was selected in bisexual organisms, despite their deleterious effects, because of the ability of bisexual genomes of counteract their load. On the contrary, they suggested that the presence of TEs in unisexual organisms could be advantageous in the short-term, but lethal in the long-term. The early beneficial effect of TEs in unisexual genomes seems to be involved in the production of genetic variability in otherwise clonal lineages.

Data on TE content in parthenogenetic lineages derive from analyses on the crustacean *Daphnia pulex* and from Rotifera (Valizadeh and Crease, 2008; Gladyshev and Arkhipova, 2010a and 2010b; Schaack *et al*., 2010).

Daphnia pulex is a cyclical parthenogenetic aquatic crustacean, whose populations reproduce clonally, unless the occurrence of stressful environmental conditions. In this case they shift to sexual reproduction; however, some populations lost this ability and became obligatory unisexuals (Schaack *et al*., 2010). Six TE families of the class II were assayed (Tc1-1, Tc1-2, Helitron 1, Helitron 2, a hAT homologue and the piggyBac element Pokey): cyclically parthenogenetic populations showed higher TE loads with respect to obligate unisexual ones (Valizadeh and Crease, 2008; Schaack *et al*., 2010). Rotifers of the Class Bdelloidea are characterized by an ancient unisexuality. The genome

of these microscopic freshwater invertebrates contains in telomeric regions transposons either of class I and II, and also retrovirus-like-retrotransposons (Gladyshev and Arkhipova, 2010a and 2010b).

1.3 Aims of the thesis

Main aim of my Ph.D. thesis was to determine the dynamics of repeated DNA sequences in *Bacillus* stick insects. The choice of this genus was made on the basis of its peculiar reproductive biology, ranging from bisexuality to facultative/obligatory parthenogenesis. Therefore, *Bacillus* was used as model group to test if and how the reproductive modality may affect the accumulation and the activity of repetitive sequences, with a remarkable attention to TEs to test also the threat of extinction from Muller's ratchet theory in its species. In particular, the role of the reproductive strategy has been deeply analysed for the most studied non-LTR retrotransposon R2.

To compare R2 load in different reproductive milieaux and given previous data at the species/population levels on the facultative parthenogen *B. rossius* (bisexuals *B. rossius rossius* population from Anzio and *B. rossius redtenbacheri* one from Patti; all-females *B. rossius redtenbacheri* population from Curcuraci) I focused on the strictly gonochoric *B. grandii maretimi* (Marettimo population) one side, and on the obligatory parthenogenetic *B. atticus atticus* (Necropoli Camarina and Scoglitti populations), the other. R2 dynamics was analysed in the progeny of facultative parthenogenetic isolates of *B. rossius redtenbacheri* (Castanea delle Furie, Curcuraci and Massa San Nicola populations) and in the male offspring of crosses involving parental females either from all-females populations (*B. rossius redtenbacheri* Curcuraci ♀ X *B. rossius rossius* Anzio ♂) and from bisexual ones (*B. rossius rossius* Anzio ♀ X *B. rossius rossius* Anzio ♂ and *B. rossius rossius* Capalbio ♀ X *B. rossius rossius* Capalbio ♂). The activity was also calculated in the descendants of *B. atticus atticus* obligatory parthenogenetic isolates (Necropoli

Camarina and Scoglitti populations). Moreover, in the all-females *B. rossius redtenbacheri* population from Castanea delle Furie, the R2 dynamics relative to a field collected female was studied up to the 5th generation.

I widened the study of repeated elements in *Bacillus*, through the contribution to the analyses of the TE genomic content in *B. grandii*, *B. rossius* and *B. atticus* for which both low coverage libraries and Southern blot experiments were conducted. The identification of tandem repeats and of TEs was performed by specific software together with a de novo analysis by self-comparison of each library, to find other interspersed repeats.

CHAPTER 2

RESULTS

The majority of the data obtained during my Ph.D. research have been published in three papers on ISI journals. I will here briefly introduce main aspects.

2.1 Evolutionary dynamics of R2 retroelement

Bonandin, L., Scavariello, C., Luchetti, A. and Mantovani, B. (2014) Evolutionary dynamics of R2 retroelement and insertion inheritance in the genome of bisexual and parthenogenetic *Bacillus rossius* populations (Insecta, Phasmida). Insect Molecular Biology 23(6): 808-820.

This part of my work was devoted to characterize and evaluate the dynamics of the non-LTR retrotransposon R2 in the facultative parthenogenetic *B. rossius* in the light of the Muller's ratchet expectations. The isolation and sequencing of the R2 complete elements from bisexual populations of *B. rossius rossius* and *B. rossius redtenbacheri* (see Scavariello, 2016; Ph.D. thesis) allowed me to analyse R2 presence and activity through the insertion display method, in selected parental individuals and in samples of their descendants; in particular I checked the offspring of *B. rossius redtenbacheri* parthenogenetic isolates and the male progeny of crosses involving parental females either from *B. rossius redtenbacheri* all-females populations and from *B. rossius rossius* bisexual ones. On the whole, 10 offspring individuals from two females for each of three parthenogenetic populations and 10-20 specimens of the progeny of five crosses were analysed. Insertions analyses indicate that R2 is actively transposing and the turnover

mechanisms are effective even in parthenogenetic genomes, contrary to Muller's ratchet expectations.

Insect Molecular Biology

Insect Molecular Biology (2014) **23**(6), 808–820 doi: 10.1111/imb.12126

Evolutionary dynamics of R2 retroelement and insertion inheritance in the genome of bisexual and parthenogenetic *Bacillus rossius* populations (Insecta Phasmida)

L. Bonandin*, C. Scavariello*, A. Luchetti and B. Mantovani

Dipartimento di Scienze Biologiche, Geologiche e Ambientali, Università di Bologna, Bologna, Italy

Abstract

Theoretical and empirical studies have shown differential management of transposable elements in organisms with different reproductive strategies. To investigate this issue, we analysed the R2 retroelement structure and variability in parthenogenetic and bisexual populations of *Bacillus rossius* stick insects, as well as insertions inheritance in the offspring of parthenogenetic isolates and of crosses. The *B. rossius* genome hosts a functional (R2Br^{fun}) and a degenerate (R2Br^{deg}) element, their presence correlating with neither reproductive strategies nor population distribution. The median-joining network method indicated that R2Br^{fun} duplicates through a multiple source model, while R2Br^{deg} is apparently still duplicating via a master gene model. Offspring analyses showed that unisexual and bisexual offspring have a similar number of R2Br-occupied sites. Multiple or recent shifts from gonochoric to parthenogenetic reproduction may explain the observed data. Moreover, insertion frequency spectra show that higher-frequency insertions in unisexual offspring significantly outnumber those in bisexual offspring. This suggests that unisexual offspring eliminate insertions with lower efficiency. A comparison with simulated insertion frequencies shows that inherited insertions in unisexual and bisexual offspring are significantly different from the expectation. On the whole, different mechanisms of R2 elimination in unisexual vs bisexual offspring and a complex interplay between recombination effectiveness, natural selection and time can explain the observed data.

Keywords: *Bacillus rossius* stick insects, bisexuality, parthenogenesis, R2 non-LTR element, retrotransposon insertion frequency, ribosomal DNA (rDNA).

Introduction

All living organisms contain in their genome a significant fraction of transposable elements (TEs), i.e. interspersed repeats able to move independently from one genomic location to another. They can be classified as class I TEs, moving via an RNA intermediate through 'copy-out/copy-in' (non-LTR elements) or 'copy out/paste in' (LTR elements) mechanisms, and class II elements, moving via a DNA intermediate by 'cut and paste' processes. TEs affect the host genome through a number of interactions that can be either beneficial or deleterious: they can modify gene expression and gene structure or promote recombination and genomic rearrangements (Kazazian, 2004).

Although they can be occasionally beneficial, TEs act as selfish elements and can replicate to high copy numbers affecting the host fitness (Doolittle & Sapienza, 1980). Mechanisms limiting TEs' proliferation, however, have evolved to allow host survival: cellular endogenous mechanisms, for example, silence TEs by methylation or through small RNA interference (O'Donnell & Burns, 2010). Nevertheless, much of the TEs population dynamics depend on the evolutionary dynamics of host organisms. In particular, theoretical studies linked differential TE loads with different reproductive strategies (Wright & Finnegan, 2001; Nuzhdin & Petrov, 2003 and reference therein). The idea that TEs can disperse in a population under bisexual reproduction was first put forward by Hickey (1982). Furthermore, according to Muller's ratchet

First published online 18 August 2014.

Correspondence: Dr Andrea Luchetti, Dip. BiGeA, via Selmi 3, 40126 Bologna, Italy. Tel.: + 39 051 2094173; e-mail: andrea.luchetti@unibo.it

*Equally contributing Authors.

theory, only bisexually reproducing organisms should be able to control and limit the deleterious proliferation of TEs thanks to mutation, recombination and the independent assortment of chromosomes. This would give natural selection more opportunities to operate (Arkhipova, 2005; Dolgin & Charlesworth, 2006; Loewe & Lamatsch, 2008). For the same reasons, in selfing or non-recombining genomes, TEs should either accumulate until host lineage extinction or be eliminated by genome purging (Wright & Finnegan, 2001; Nuzhdin & Petrov, 2003 and references therein). Actually, the few studies that have specifically addressed this issue showed contrasting patterns: while comparisons between selfing and outcrossing relatives in *Arabidopsis* and *Caenorhabditis* showed that TEs persist longer in selfing lineages, in obligate parthenogenetic lineages of *Daphnia pulex* the opposite result emerged (Valizadeh & Crease, 2008; Schaack *et al*., 2010). Moreover, in the wasp *Leptopilina claviceps,* only DNA transposons were found to have a higher copy number in the parthenogenetic lineage (Kraaijeveld *et al*., 2012).

Early-branching class I TEs show the particular feature of insertion into specific sequences, often tandem repeats; this strategy has been interpreted as an adaptation in order to bring little damage to the host genome, thus escaping genome purging (Malik *et al*., 1999; Malik & Eickbush, 2000). Among these TEs, R2 elements are the best studied retrotransposons so far and serve as a model to understand the biology of transposition of the whole non-LTR element subclass (Eickbush, 2002).

R2 is a non-LTR retroelement that inserts into the ribosomal DNA locus. R2 has a single open reading frame (ORF) flanked by two untranslated sequences of variable length. The ORF comprises the central reverse transcriptase (RT) domain, the DNA-binding motifs at the N-terminus and the restriction enzyme-like endonuclease (RLE) domain at the C-terminus. The C-terminal end of the R2 protein includes a cysteine-histidine (zinc finger, ZF) motif (CCHC), while the N terminal domain can contain one (CCHH), two (CCHH + CCHH or CCHC + CCHH), or three (CCHH + CCHC + CCHH) ZF motifs (Kojima & Fujiwara, 2005; Luchetti & Mantovani, 2013). R2 multiplies its copy number through a 'copy–out/copy-in' mechanism called target-primed reverse transcription (Eickbush, 2002; Eickbush & Eickbush, 2007). Target-primed reverse transcription frequently generates variable length deletions at the 5′ end of R2 and accounts for the numerous 5′ truncated copies in some insects (Burke *et al*., 1993). These 5′ truncations probably derive from the cellular degradation of the RNA transcript or from the inability of the reverse transcriptase to reach the 5′ end of the transcript (Christensen *et al*., 2006).

R2 is found in a variety of metazoan genomes and it is vertically inherited (Kojima & Fujiwara, 2005; Luchetti & Mantovani, 2013). The wide occurrence of R2 from Cnidaria to vertebrates is possibly attributable to the element niche provided by the ribosomal DNA locus (rDNA): in particular, R2 specifically inserts into the 28S rRNA gene sequence. This very specific target sequence makes the study of R2 activity quite interesting because of the interconnected TE-rDNA genomic dynamics.

rDNA shows a pattern of nucleotide variability known as concerted evolution, occurring when tandemly repeated sequences are more similar to each other within than between reproductive units (populations, subspecies, species, etc). This is achieved by a dual process: at the genome level, by molecular drive, through the homogenization of paralogous sequences by unequal DNA exchanges (mainly unequal crossing over, gene conversion, replication slippage; also known as genomic turnover mechanisms, GTM) and, at the reproductive unit level, by fixation of specific mutation profiles by means of bisexual reproduction (Eickbush & Eickbush, 2007; Plohl *et al*., 2008). The interplay between R2 dynamics and rDNA molecular drive is suggested by the deletion of rDNA units upon R2 insertion and rDNA unit replacement through compensatory GTM (Zhang *et al*., 2008; Mingazzini *et al*., 2011). This can be advantageous as 28S-inserted copies can be eliminated and replaced with new, functional 28S copies but, in this way, new sites for further R2 insertions are also provided (Eickbush & Eickbush, 2007; Zhang *et al*., 2008).

The short-term R2 inheritance and the interplay with rDNA have been well investigated in *Drosophila* spp. (Pérez-Gonzalez & Eickbush, 2001; Zhang & Eickbush, 2005; Zhang *et al*., 2008; Zhou & Eickbush, 2009; Zhou *et al*., 2013) and a few comparable studies address this issue in other organisms (Kagramanova *et al*., 2010; Ghesini *et al*., 2011; Mingazzini *et al*., 2011). The characterization of R2 elements in the genus *Bacillus* (Insecta Phasmida) should provide, however, the opportunity for analysis linking R2 dynamics to reproductive strategies. The genus is in fact a well-known example of reticulate evolution, and it includes the two highly differentiated species *Bacillus rossius* (bisexual with facultative parthenogenetic populations) and *Bacillus grandii* (strictly bisexual) as well as the obligatory parthenogenetic taxon *Bacillus atticus* and their related diploid and triploid non-Mendelian hybrids: the hybridogenetic lineages *B. rossius-grandii* and the obligatory parthenogenetic diploid *Bacillus whitei* and triploid *Bacillus lynceorum*. *B. rossius* is distributed over most of the Mediterranean basin with eight zymoraces, while *B. grandii* subspecies (*B. grandii grandii, B. grandii benazzii,* and *B. grandii maretimi*) and the hybrids are endemic to the Sicilian area; *B. atticus*, finally, occurs in the eastern part of the basin (Scali *et al*., 2003 and references therein; Fig. S1).

In the genus *Bacillus*, parthenogenesis occurs through a variety of mechanisms, ranging from automictic to apomictic processes. In *B. rossius*, in particular, females of parthenogenetic populations produce through meiosis haploid eggs; these start the development and then diploidize when reaching some thousands of cells, the mechanism generating an all-female offspring homozygous at all loci (Scali *et al*., 2003 and references therein).

Parthenogenesis may have major effects on concerted evolution: at variance of bisexual species, obligatory parthenogenetic taxa of the genus *Bacillus* show little or no fixation at all of either pericentromeric tandem repeats (Cesari *et al*., 2003; Luchetti *et al*., 2003) or rDNA intergenic spacer tandem repeats (Ricci *et al*., 2008). Generally speaking, it is to be recalled that a concerted evolution pattern in rDNA repeats can partly be fostered by selective pressures acting on both coding (18S, 5.8S and 28S genes) and regulatory (tandem repeats) regions (Ricci *et al*., 2008; Ambrose & Crease, 2011).

In the present study we report the characterization of R2 elements in the genome of *B. rossius* stick insects, along with inheritance studies in bisexual and parthenogenetic populations.

Results

R2 sequence characterization

The R2 element isolated from the *B. rossius rossius* population from Capalbio (roCAP) is 3492 base pairs (bp) long, excluding the 3′ terminal poly-(A) tail of about 23 nucleotides, with an ORF of 3165 bp encoding a protein of 1054 amino acids. The amino acid sequence shows an RT domain, an RLE domain at its C-terminal end and a single ZF motif (CCHH type) at the N-terminal end (Fig. 1). These observations are consistent with preliminary data from Mingazzini (2011). By contrast, the R2 element isolated from the *B. rossius redtenbacheri* sample from Patti (rePAT) is 3517 bp long and exhibits 14 frameshift mutations and one internal stop codon with respect to the roCAP R2 ORF (Fig. 1). The RT and the RLE domains as well as a single ZF motif at the N-terminal end are recognizable in the degenerate rePAT element. On the whole, the two R2 sequences are 9.2% divergent, the 5′ untranslated region (UTR) being more variable (16.4%) than either the ORF or the 3′ UTR (8.8% and 8.5%, respectively). For clarity, the R2 elements with a functional ORF and those with the degenerate one are hereafter referred to as R2Br^{fun} and R2Br^{deg}, respectively. The distribution of functional and degenerate R2 variants was studied in a population survey performed through PCR amplification and subsequent sequencing of the 3′ half of the element. Only the 1042 bp of the ORF, including the C-2480 frameshift mutation (Fig. 1, marked by the asterisk) were considered because, as a result of the presence of long poly-(A) tails, the sequencing of the 3′ UTR was not of sufficient quality. The three sampled populations of *B. rossius rossius* (roFOL, roCAP and roANZ) and the *B. rossius redtenbacheri* reBER sample show 100% functional elements; conversely, *B. rossius redtenbacheri* Sicilian samples (rePAT, reCDF, reMSN and reCUR) show 100% degenerated R2 sequences, while the remaining *B. rossius redtenbacheri* populations (reTDS, reVIR, reCOM and reGAB) and the *B. rossius tripolitanus* A population have both variants (Table 1, Fig. 2).

R2Br^{fun} sequence diversity within population ranges from 0.0012 ± 0.0007 to 0.0527 ± 0.0047. R2Br^{deg} variation is far more limited, being comprised between 0.0024 ± 0.0014 and 0.0162 ± 0.0057 (Table 1). On the

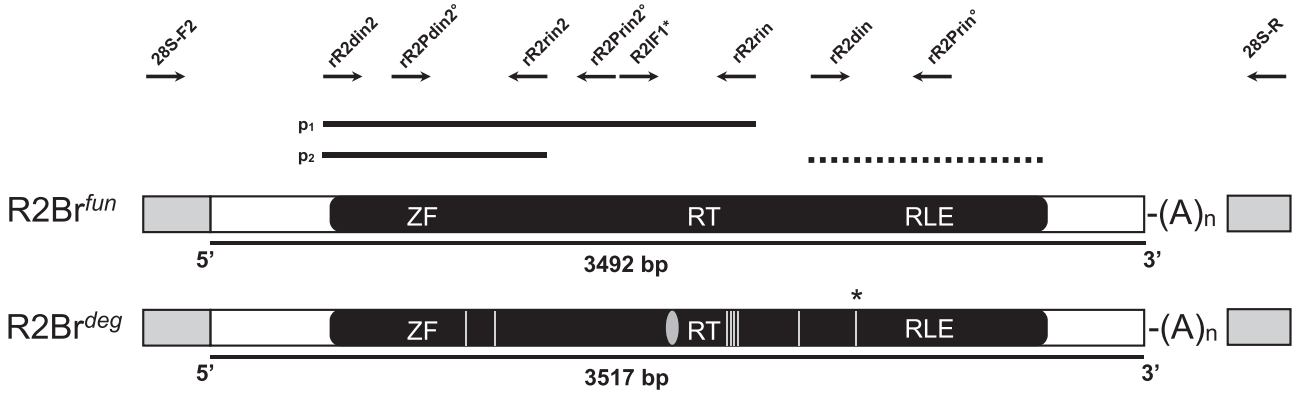


Figure 1. Schematic representation of R2Br functional (R2Br^{fun}) and degenerate (R2Br^{deg}) elements. Grey boxes indicate flanking 28S gene sequences; black boxes indicate the open reading frame (ORF) with the zinc finger (ZF), reverse transcriptase (RT) and restriction enzyme-like endonuclease (RLE) domains. In the R2Br^{deg} ORF, vertical grey lines indicate the frameshift mutations, the asterisk marking the C-2480 frameshift mutation used to distinguish the two R2 variants in the population survey; the oval represents the stop codon. In the panel above, the primers used for sequencing and transposon display are reported: the primer from Kojima & Fujiwara (2005) is marked with *, while those indicated with ° are primers specific for the degenerate variant. Thick bars p1 and p2 mark the regions used as probes in the transposon display blotting; the dashed bar indicates the region sequenced in the population survey.

Table 1. *Bacillus rossius* sampling with acronyms and reproductive strategy

Subspecies/collecting site	Acronym	RS	R2Br^{fun} (p-D ± SE)	R2Br^{deg} (p-D ± SE)
Bacillus rossius rossius				
Follonica	roFOL	P	7 (0.0075 ± 0.0025)	0
Capalbio	roCAP	G	10 (0.0032 ± 0.0011)	0
Anzio	roANZ	G	10 (0.0075 ± 0.0019)	0
B. rossius redtenbacheri				
Patti	rePAT	G	0	9 (0.0036 ± 0.0012)
Castanea delle Furie	reCDF	P	0	6 (0.0034 ± 0.0014)
Massa San Nicola	reMSN	P	0	6 (0.0054 ± 0.0018)
Curcuraci	reCUR	P	0	5 (0.0024 ± 0.0014)
Torino di Sangro	reTDS	G	12 (0.0028 ± 0.0011)	2 (0.0162 ± 0.0057)
Villa Rosa	reVIR	P	7 (0.0012 ± 0.0007)	1 (n/a)
Bertinoro	reBER	G	6 (0.0202 ± 0.0035)	0
Comacchio	reCOM	P	4 (0.0061 ± 0.0022)	5 (0.0038 ± 0.0016)
Gabonjin	reGAB	P	1 (n/a)	9 (0.0036 ± 0.0012)
B. rossius tripolitanus A				
Korbous	trKOR	G	8 (0.0527 ± 0.0047)	1 (n/a)

The number of functional (R2Br^{fun}) and degenerated (R2Br^{deg}) elements are given with their sequence diversity (p-D = p-distance) ± SE per population in parentheses. n/a, not applicable; RS, reproductive strategy; G, gonochoric; P, parthenogenetic.

whole, sequence variability between the functional and degenerate dataset was not significant (Student's *t*-test = 1.128, $P = 0.291$). Within the two datasets, even the comparison of R2Br variability between parthenogenetic and gonochoric populations was not significant ($t = 1.287$, $P = 0.263$; $t = 0.978$, $P = 0.505$).

The phylogenetic analysis included the two full-length elements and all 107 R2 3′ end sequences. These sequences split into two well-defined, main clusters: one includes only R2Br^{deg} elements, the other contains only R2Br^{fun} elements (Fig. 3A). Three sequences, though, fall outside the two main clades: one R2Br^{deg} from the reTDS population (reTDS-f1) and two R2Br^{fun} from the reGAB (reGAB-10) and trKOR (trKOR-9) samples (all indicated by arrowheads in Fig. 3A). The peculiar position of these sequences was explained by gene conversion events detected between functional and degenerate sequences: two functional elements showed extensive tracts exchanged with degenerate ones (309 bp in the reGAB-10 and 434 bp in trKOR-9), while only one R2Br^{deg} showed a converted tract (124 bp, reTDS-f1).

The median-joining network also confirmed the clear separation between functional and degenerate sequences, with the only exception of the trKOR-9

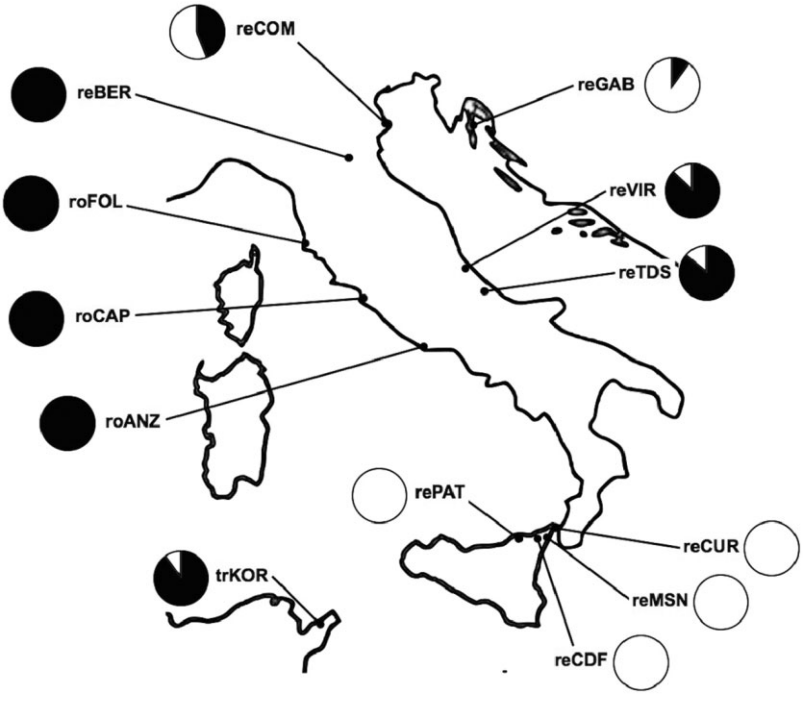


Figure 2. Geographic distribution of sampling localities. Pie charts indicate the proportion of R2Br functional (black) and degenerate (white) elements scored by sequencing in each sample. Acronyms are as in Table 1.

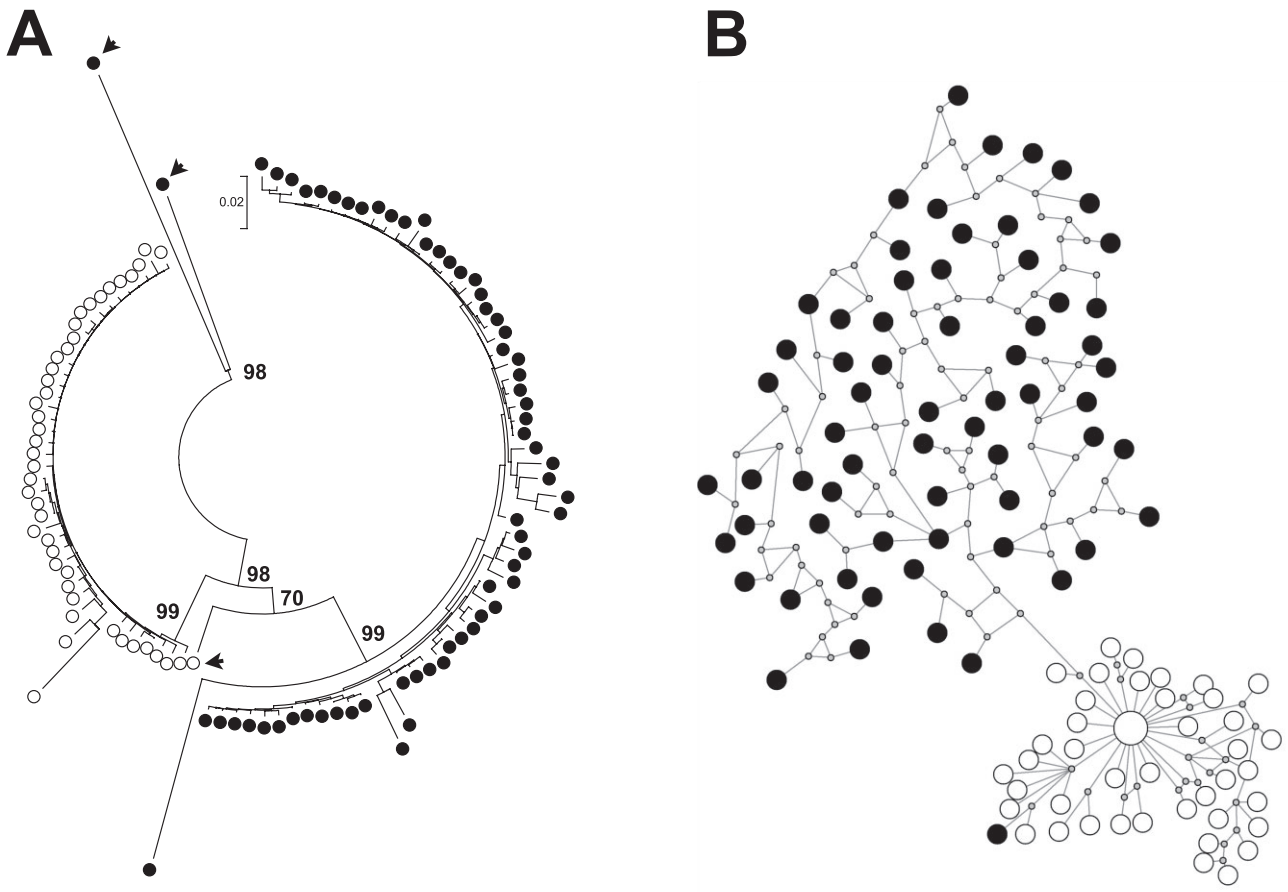


Figure 3. Phylogenetic analyses of R2 fragments sequenced in the population survey. (A) Minimum Evolution tree; numbers at nodes represent bootstrap values ≥70% obtained after 500 replicates. Arrowheads indicate sequences subject to gene conversion events. (B) Median-joining network; circles magnitude is proportional to the sequence variant frequency; the small, grey dots represent median-joining vectors. For both analyses, black-filled circles represent R2Br^{fun} sequences, while the white-filled ones indicate the R2Br^{deg} elements.

sequence. This sequence lacked the C-2480 frameshift mutation – and was therefore classified as functional – but otherwise showed a close relationship with R2Br^{deg} elements (Fig. 3B), owing to the above reported gene conversion event. The sub-networks corresponding to the two element types showed a quite different topology. The degenerate sequences exhibited a star-like pattern with a most common sequence, the ancestral one, from which all others originated. It is worth noting that the consensus R2Br^{deg} sequence is identical to this ancestral element. On the other hand, the sub-network built by functional sequences has a reticulated pattern, with no evident ancestry of a specific element (Fig. 3B).

Tajima's *D* values on R2Br sequences were significantly negative for both functional and degenerate datasets: $D = -2.685$, $P < 0.001$, and $D = -2.683$, $P < 0.001$, respectively. This indicates that both functional and degenerate elements are either under purifying selection or that they are both experiencing a relatively recent copy number expansion.

R2 insertions display

R2 activity was determined through the insertion display, the bands of variable length representing the 5′ end deletions occurring upon insertions (Pérez-Gonzalez & Eickbush, 2001). The detection of new 5′ truncated copies indicates insertion activity and, therefore, that the element is actively transposing (with a greater or lesser activity depending on the number of insertions). While this approach has some limitations, being unable to detect new insertions if the new truncated variant is of the same length as pre-existing bands or if the new insertion involves a full-length element (i.e. no. 5′ deletions occur), it has proven to be highly informative about the element's activity (Pérez-Gonzalez & Eickbush, 2001; Zhang & Eickbush, 2005; Zhang *et al.*, 2008; Zhou & Eickbush, 2009; Mingazzini *et al.*, 2011).

In parental parthenogenetic females, the number of sites occupied by R2 (S_p ♀) was in the range of 4–15 (Table 2). In their unisexual offspring, the total number of sites

Table 2. Parameters calculated on the basis of the R2 insertion display.

Parental individual(s)	S_p ♀	S_p ♂	S_p^{share}	S_p^{tot}	N_o	S_o^{tot}	S_o^{new}	ñ
reCDF-♀3	11			11	10♀	16	5	12.10
reCDF-♀4	14			14	10♀	18	4	12.80
reMSN-♀1	15			15	10♀	21	6	14.60
reMSN-♀2	4			4	10♀	4	1	3.60
reCUR-♀1	13			13	10♀	13	0	9.10
reCUR-♀2	15			15	10♀	17	2	12.10
reCUR-♀5 × roANZ-♂15	2	5	4	11	20♂	11	0	7.25
reCUR-♀6 × roANZ-♂9	0	0	4	4	20♂	4	0	3.70
roANZ-♀21 × roANZ-♂8	11	2	6	19	20♂	19	1	11.05
roCAP-♀1 × roCAP-♂4	2	8	5	15	10♂	19	4	9.50
roCAP-♀2 × roCAP-♂3	4	3	5	12	10♂	20	8	9.10

The number of sites occupied by R2 detected in parental individual(s), either in the single specimen (S_p♀, S_p♂) or shared between the two mates (S_p^{share}), and their total number ($S_p^{tot} = S_p♀ + S_p♂ + S_p^{share}$). The number of analyzed offspring individuals (N_o) and the total number of sites occupied by R2 in the offspring (S_o^{tot}) and the number of new sites (i.e. not occurring in parents) occupied by R2 (S_o^{new}); the estimated R2 copy number per haploid genome is indicated by ñ.

occupied by R2 (S_o^{tot}) was in the range of 4–21, with 1–6 new insertions (S_o^{new}) detected in all isolates with the exception of reCUR-♀1 progeny (Table 2). Interestingly, new sites occupied by R2 have been found also in those isolates that, from the sequence survey, showed only the presence of R2Br^{deg} (i.e. reCUR, reMSN, reCDF; Table 2), suggesting that these genomes could harbour active R2 copies.

In crossing experiments, the number of sites occupied by R2 was in the range of 2–11 within the female (S_p ♀) or the male (S_p ♂) parent, while 4–6 additional occupied sites are shared between the two mates (S_p^{share}; Table 2). The bisexual offspring showed S_o^{tot} in the range of 4–20; no new insertions were detected in the offspring of the two crosses involving former-unisexual females (reCUR-♀5 × roANZ-♂15 and reCUR-♀6 × roANZ-♂9), while 1–8 S_o^{new} were found in the descendants of the remaining three crosses (Table 2).

The R2 copy number per haploid genome (ñ) scored in the offspring varied widely: 3.60–14.60 in the unisexual dataset and 3.70–11.05 in the bisexual one (Table 2). On the whole, unisexual offspring showed a higher number of sites occupied by R2 (S_o^{tot}) than the bisexual offspring, 89 vs 73, although this difference was not significant ($t_{one\text{-}tailed} = 0.059$, $P = 0.477$). Moreover, the number of new inserted sites in the offspring (S_o^{new}) appeared higher in unisexuals than in bisexuals, but the difference was again not significant ($t_{one\text{-}tailed} = 0.220$, $P = 0.416$). The same also applies to the average R2 copy number per haploid genome in the unisexual–bisexual comparison ($t_{one\text{-}tailed} = 1.277$, $P = 0.117$).

By contrast, insertion frequency spectra, i.e. the frequencies' distribution of R2 insertions in the sample, were significantly different between unisexual and bisexual offspring (two-sample Kolmogorov–Smirnov test: $D = 0.351$, $P < 0.0001$; Fig. 4): in particular, high-frequency insertions (i.e. insertions occurring with a frequency >0.5) observed in unisexual offspring outnumber those found in bisexual offspring and most of the insertions remains close to 1.0 frequency. Conversely,

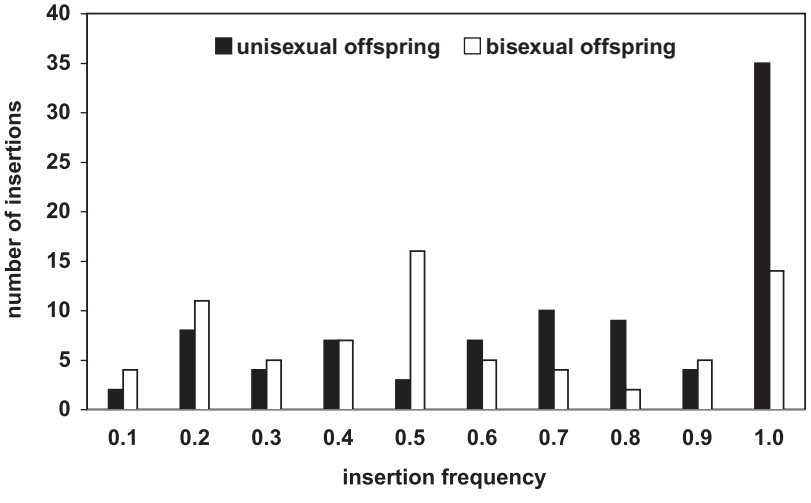


Figure 4. R2 insertion frequency spectra relative to unisexual and bisexual offspring. Each bin represents a class of insertion frequency in the offspring (i.e. the fraction of individuals carrying a certain R2 insertion over the total number of individuals) of either unisexuals (black) or bisexuals (white). On the *y*-axis are reported the number of R2 insertions recorded for each class.

low-frequency insertions (≤0.5) are more represented in bisexuals.

To obtain information about the contribution of recombination and selection in the elimination of R2 insertions, we derived expected data for unisexual and bisexual offspring: every deviation from strict insertion inheritance (in parthenogenetic offspring) or from Mendelian ratios (in crosses offspring) can be attributed to recombination or selection.

Based on the strict inheritance expected from the mechanism of parthenogenesis producing an all-homozygous progeny, all insertions of parthenogenetic mothers should have been present in the offspring at frequency = 1.0. However, data from unisexual progeny indicate quite a different scenario: excluding the new insertions occurring only in the progeny (S_o^{new}), maternally inherited insertion frequencies ranged from 0.1 to 1.0. Moreover, no offspring of any isolated female showed all mother's insertions occurring with frequency = 1.0 (not shown).

Expected insertion frequencies in crosses are, however, more difficult to establish with certainty as no information about the insertion heterozygosity can be derived (except for the two formerly unisexual females from the reCUR population that are homozygous). We thus simulated three different conditions: i) both parents are completely homozygous (all.ho); ii) both parents, with the exception of the reCUR females, are completely heterozygous (all.he); or iii) one parent is homozygous and the other one is heterozygous (one.ho). In all three simulations, bisexual offspring showed significant deviation from the Mendelian expectations of insertion frequencies, with the only exception of roCAP-♀1 × roCAP-♂4 descendants which fit the all.he condition (Table S1).

Observed R2 copy number per haploid genome is significant reduced from the expectation in unisexual offspring ($t_{paired} = 5.258$, $P < 0.01$; Fig. 5A). The same holds for bisexual offspring under the all.ho simulation ($t_{paired} = 3.630$, $P < 0.05$; Fig. 5B) but they did not show any significant difference from the expected inherited copy number under the all.he ($t_{paired} = 2.491$, $P = 0.067$) or one.ho simulation ($t_{paired} = 2.691$, $P = 0.055$), the R2 copy number resulting only slightly lower (Fig. 5B).

Discussion

R2Br^{fun} and R2Br^{deg} diversity and evolution

The present study first demonstrates the presence in the genome of *B. rossius* of two variants of the same R2 element: a functional one (R2Br^{fun}) and a degenerate one (R2Br^{deg}). The population sequence survey shows a peculiar distribution of the two variants in the Italian subspecies. On one hand, all *B. r. rossius* samples appear to host only R2Br^{fun}, and on the other hand, *B. r. redtenbacheri*

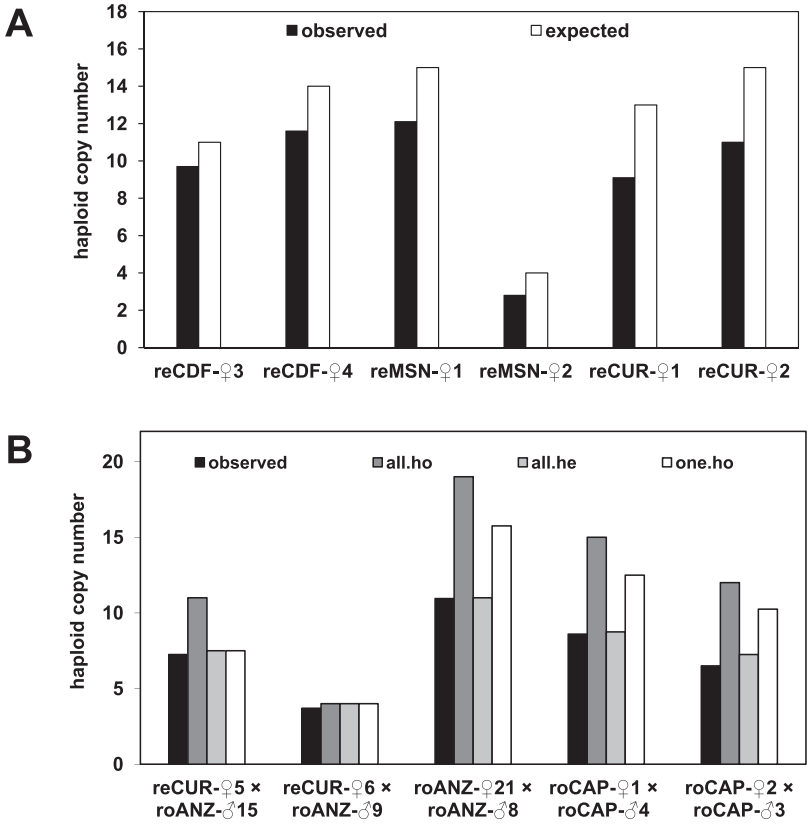


Figure 5. Observed and expected inherited R2 copy number in unisexual (A) and bisexual (B) offspring. Each bin represents parthenogenetic (A) or crosses (B) progeny observed (black) or expected (white and grey) R2 copy number per haploid genome, calculated as the sum of expected insertion frequencies based on simulation data. New insertions (i.e. those occurring only in the progeny but not in the parental individuals) are excluded. For unisexual offspring, the expectation is based on the fact that, given the parthenogenesis mechanism, progeny should have all the insertions of the mother. For bisexual offspring, as it is not possible to know if parental individuals have homozygous or heterozygous condition for the scored insertions, three simulations have been carried out: i) parental individuals are both homozygous (all.ho); ii) parental individuals, except reCUR females, are both heterozygous (all.he); iii) only one of the two parental individuals is homozygous while the other is heterozygous (one.ho). Acronyms are as in Table 1.

populations show widely different combinations: in all Sicilian populations only R2Br^{deg} was found, while peninsular populations contained both variants, in different percentages, or even only the R2Br^{fun} variant (reBER). The R2 insertion display, however, indicates that in the parthenogenetic *B. r. redtenbacheri* Sicilian populations reCUR, reMSN and reCDF new insertions occur, suggesting that some R2Br^{fun} copies could exist in their genome.

On the whole, no link appears between R2Br variant distribution and subspecies or reproductive strategies. Taking into account the evolutionary history of *B. rossius*, the presence of R2Br^{deg} in *B. r. tripolitanus* A indicates that it was already present in the ancestral *B. rossius* genome. When Europe and North Africa separated, leading to allopatric *B. rossius* lineage break (>5 Myr ago; Mantovani *et al*., 2001), R2Br^{deg} has been conserved in the extant genomes along with R2Br^{fun}.

Interestingly, the analysis of mutation distribution among sequenced R2Br ORF 3′ ends showed a non-random pattern. Tajima's *D* values calculated on both R2Br^{fun} and R2Br^{deg} datasets were significantly negative, suggesting that both variants are under purifying selection or in a state of a relatively recent copy number expansion. As purifying selection can be reasonably ruled out when dealing with mobile genetic elements, and particularly for the degenerate elements, Tajima's *D*s would indicate a recent burst of duplication. While this is perfectly conceivable for R2Br^{fun} and reflects its activity, the same cannot hold for R2Br^{deg} elements that do not code for a functional protein. Moreover, the median-joining network indicates that the replication of the two variants follows different patterns. Functional elements appear to duplicate through a multiple source model, reflected by the reticulated network, while degenerate elements seem to replicate via a master gene model, producing a star-like network. This implies that the R2Br^{deg} element differentiated once in the ancestral *B. rossius* genome and then gave origin to all extant R2Br^{deg} copies.

On the whole, therefore, data indicate that both R2Br variants can replicate, but their recent history shows the signature of different evolutionary dynamics.

While the sequence analyses provide evidence that R2Br^{deg} is undergoing duplication, the question remains how. At present we can only speculate about two possible explanations: i) R2Br^{deg}, after having lost its ability to code for a functional protein, became a non-autonomous element exploiting the retrotransposition machinery of an R2Br^{fun} not identified in our survey or ii) it duplicates when the host 28S sequence itself duplicates thanks to unequal crossing over or other GTMs. Both possibilities are likely to explain the observed data. In the *Drosophila* genome, non-autonomous elements derived from R2, or from a fusion between R2 and R1, (SIDE elements) were found to insert within the 28S by parasitizing the enzymatic machinery of parental elements (Eickbush & Eickbush, 2012). Moreover, again in the *Drosophila* genome, some R2 insertions have been found in multiple copies as a result of rDNA unequal crossing over, even if at a very low frequency (Zhou *et al*., 2013). Future studies addressing this specific issue will shed light on the mechanisms allowing R2Br^{deg} element duplication and, therefore, its survival.

R2Br insertion display and host reproductive strategies

The study of R2Br activity through the insertion display method, even if analysing a single generation, highlighted some interesting features. First of all, unisexual and bisexual offspring showed a similar number of R2-occupied sites (S_o^{tot}); however, unisexual offspring exhibited more high-frequency insertions than bisexual offspring, indicating that each insertion is less likely eliminated in the unisexual than in a bisexual condition. This pattern is partly consistent with theoretical predictions and with previous analyses of TE insertions in outcrossing and selfing lineages, where outcrossers showed more insertions at low frequency than the selfing relatives in both *Arabidopsis* spp. (Lockton & Gaut, 2010 and references therein) and *Caenorhabditis* spp. (Dolgin *et al*., 2008). These data contrast, however, with the dynamics of r*Pokey*, a rDNA-targeting element in the *D. pulex* genome. In this instance, obligate unisexuals have a significantly lower number of insertions with respect to cyclically unisexual isolates, but their frequency spectra are very similar (Valizadeh & Crease, 2008). This has been explained by the polyphyletic, recent origin of obligate parthenogenetic isolates that inherited the insertion profile of their cyclical ancestors and, then, possibly underwent clonal selection (Valizadeh & Crease, 2008). *B. rossius* bisexual populations can rapidly shift to unisexuality through male loss, as it has been directly observed in the wild (Scali *et al*., 2003), while the chance of sex chromosomes meiotic missegregation accounts for the appearance of males in laboratory-reared parthenogenetic lines (Mantovani B., pers. obs.). Multiple or recent shifts may explain why unisexuals and bisexuals have the same number of R2-occupied sites, but differences remain with *D. pulex*. One possibility is that the crustacean may undergo heavier environmental stresses than stick insects and this would cause either stronger clonal selection or, simply, stronger selection favouring genomes with a lower r*Pokey* load. This might also explain differences in relative frequency spectra, even though the type of parthenogenesis can play a role.

D. pulex eggs are diploid by abortive meiosis (Hiruta *et al*., 2010), therefore the heterozygosity of the mother is maintained in the offspring even if loss of heterozygous mutations has been documented by ameiotic recombination (Omilian *et al*., 2006). In contrast, *B. rossius* parthe-

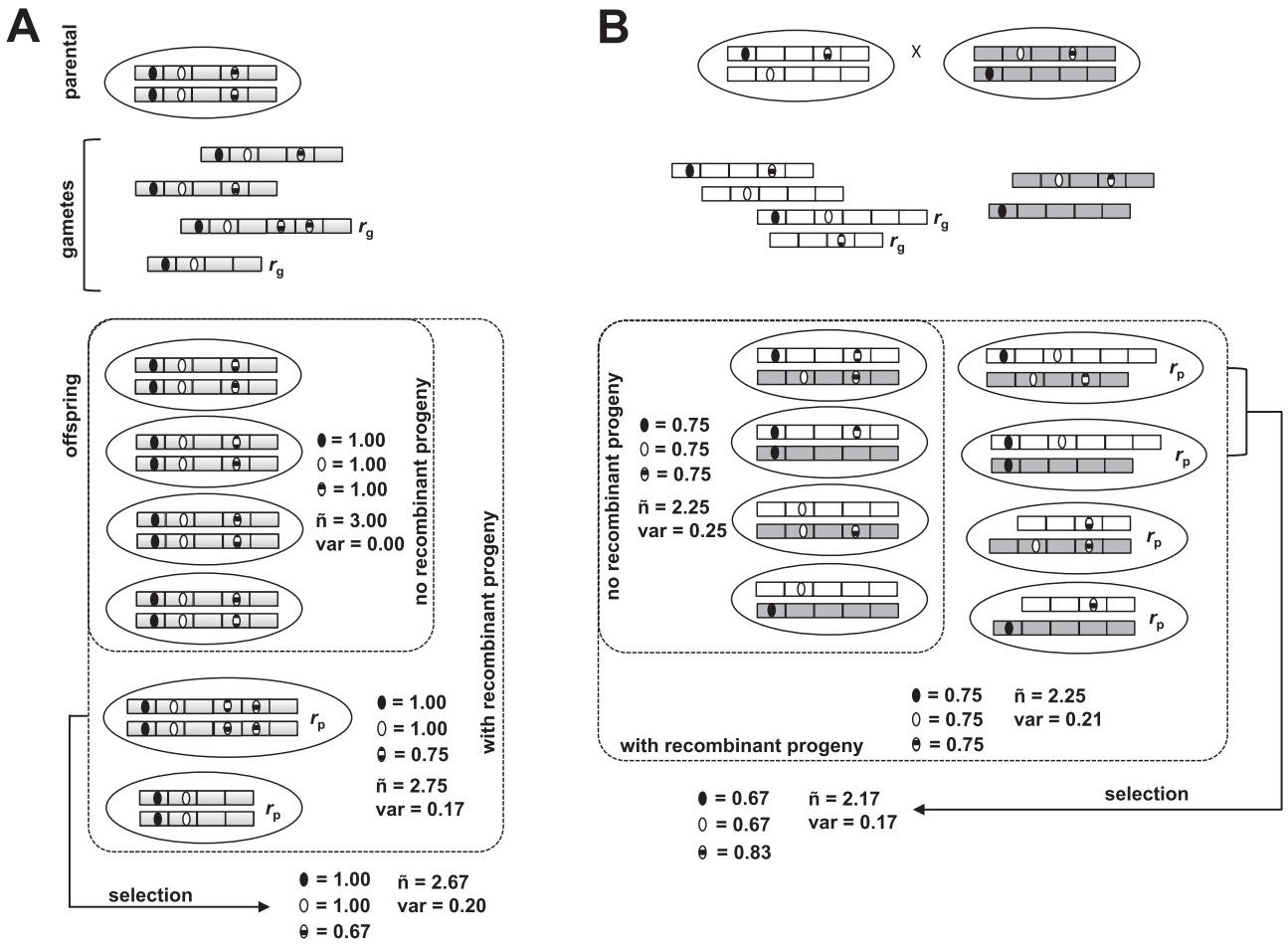


Figure 6. Scenarios of R2 insertion inheritance under unisexuality (A) and bisexuality, all.he simulation (B). Each box represents a 28S rRNA and dots are R2 insertions (different colours = different insertions); for each parental individual, gametes (r_g: recombinant gamete) and offspring (r_p: recombinant progeny) are reported. For three possible situations (progeny without recombinant; progeny with recombinant; progeny after selection) R2 insertion frequencies, copy number (ñ) and its variance (var) have been calculated. In this model, selection has been considered as acting on offspring with longer rDNA array. In the bisexual post-selection scenario, insertion frequencies are significantly different from those calculated under the two other scenarios ($G = 8.1$; d.f. = 3; $p < 0.05$).

nogenetic development is based on eggs deriving from a normal meiosis, therefore they start haploid and then diploidize. This mechanism generates a fully homozygous offspring that should inherit the 100% of the mother's insertions.

The comparison between presently obtained data and simulated data, however, clearly showed that observed frequencies of inherited insertions are significantly different from expected ones. In Fig. 6 and Fig. S2, we show possible scenarios that may explain the results of our simulations. Deviations from an insertion frequency of 1.0 in the unisexual progeny may only happen if unequal recombination (probably the most common GTM), possibly coupled with selection on particular genotypes, has eliminated some of the insertions (Fig. 6A). In our dataset, such recombination events should have occurred several times, as multiple insertion loss can be observed in all parthenogenetic offspring; however, from a parental homozygous condition, the probabilities of generating insertion-purged gametes are lower than if starting from a heterozygous condition. In fact, the insertion frequency distribution of unisexual is skewed toward high frequencies.

In crosses' progeny, with only one exception, Mendelian ratios are rejected in all simulated expectations and, quite interestingly, the observed inherited R2 copy numbers per haploid genome are slightly lower but not significantly different from the expected ones in the all.he and the one.ho simulations. If parents are all homozygous for R2 insertions (all.ho), the scenario does not change much from the unisexual condition, with the offspring expected to inherit all insertions at frequency =1.0. If parents are heterozygous (all.he simulation), recombination events, such as unequal crossing over, would simply reshuffle the rDNA, relocating the insertions within the array; thus, when eventually checked using transposon display, the observed frequencies and copy number would not be

changed in a single generation (Fig. 6B). Other unequal DNA exchanges such as gene conversions, actually, would probably lower the copy number in a single generation. On the other hand, selection on particular genotypes would modify the frequencies, although the copy number decrease is decidedly less than in unisexual offspring (Fig 6B), in line with our comparison between observed and simulated data. If both parents produce recombinant gametes, insertion frequency and copy number variation behave similarly (Fig. S2A). Generally speaking, however, whether or not recombination and/or selection occur, the bisexual progeny will show a greater copy number variance than unisexual offspring (Fig. 6B; Fig. S2A).

Under the one.ho simulated scenario, however, the management of R2 insertion seems to change substantially: unequal recombination is effective in one generation only if occurring in the homozygous parent, regardless of whether the heterozygous parent produces recombinants or not. Moreover, selection appears to have a weak effect on both deviation of insertion frequencies and copy number reduction (Fig. S2B–D). This is probably attributable to a buffering effect of the homozygous parent that will redistribute all its insertions to the offspring. It is likely that, in this instance, the process of R2 insertion elimination would be slower and, therefore, more evident in the subsequent generations.

Following the above explanation (Fig. 6; Fig. S2), the observed pattern, compared with simulated data, suggests two different mechanisms of R2 elimination in unisexual vs bisexual offspring. In the case of unisexual offspring, unequal crossing over generates two new genotypes, one with an insertion loss and one with insertion duplication. The genotype subject to the loss would give rise to a 'purged' clonal line, counteracting, theoretically speaking, the Muller's ratchet effect. The genotype with the duplication may be eliminated or not, depending on the effects on the fitness. It is to be noted, however, that such duplications would not be detected using the insertion display method, as duplicated insertions will unavoidably result in the same PCR amplified band. The scenario depicted in the one.ho simulation (Fig. S2B–D) supports this pattern, unequal crossing over having major effects when occurring in the homozygous parent.

In all-heterozygous bisexual offspring, the R2 insertion elimination would follow a more complicated route. Recombination does not seem to affect the retrotransposon load in a single generation. More importantly, outcrossing generates higher R2 copy number variance in the progeny, resulting in a higher variance of inserted/uninserted 28S rRNAs. This would give natural selection more options to operate in the following generations (Mingazzini *et al*., 2011); in fact, in this instance, selection has a greater effect than recombination.

Generally speaking, based on theoretical predictions such as Muller's ratchet theory, the retention of high frequency TE insertions in unisexual, selfing or asexual lineages is probably attributable to the absence (or very low efficiency) of truly effective recombination mechanisms. The data in the present study suggest that, whether recombination is effective or not, the time available (in terms of number of generations) for insertion elimination may also play a role. In situations where organisms are completely/highly homozygous or have very high selfing rates, more generations may be necessary to achieve elimination efficiencies similar to those of more heterozygous organisms. In line with this, variance in TE copy number should play a greater role in the fate of the whole TE load. Moreover, the rate of TE elimination in homozygous organisms is further contrasted by the TE insertional activity itself, continuously providing new TE copies.

Another interesting result of our analysis is that unisexual and bisexual offspring showed the same number of new R2 insertions (i.e. those occurring in the offspring only); this would indicate that R2Br has the same activity despite host's reproductive strategies, even if insertions within unisexual genomes persist longer than in the bisexual ones. It appears, therefore, as if no mechanisms are acting on R2 transposition rate in these parthenogenetic genomes. Suppression of transposition is a common phenomenon, mainly performed by micro-RNAs or methylation, aiming to keep at bay TEs activity in order to avoid potentially deleterious replicative burst (O'Donnell & Burns, 2010). This is particularly important in a homozygous genome such as that of parthenogenetic stick insects, because natural selection would have fewer opportunities to eliminate deleterious insertions. However, things can be slightly different for rDNA-targeting TEs: the specific insertion into repeated genes (as rDNA) undergoing GTMs will ensure that there would always be gene copies free from TE insertions, still allowing cellular functionality. Site-specificity might therefore be an adaptive strategy in order to provoke little damage to the host and then to escape purging (Malik & Eickbush, 2000). In fruit flies, however, R2 transcription and, therefore, transposition are largely controlled by heterochromatinization of densely inserted rDNA arrays followed by possible nucleolar dominance (Eickbush *et al*., 2008); this can be achieved by compartmentalizing functional R2s in restricted rDNA regions that can be transcriptionally silenced. Starting from a condition in which functional R2s are located within transcriptionally active rDNA copies, silencing can be obtained by either i) rearranging the array by GTMs (mainly unequal crossing over) so that R2s are moved to heterochromatinizable regions, or ii) the establishment of nucleolar dominance after mating with a partner bearing an R2-silenced rDNA array (Eickbush

et al., 2008). In the case of stick insect parthenogenetic populations, the latter does not appear possible, so that recombination would be a major factor maintaining R2 at bay. Based on the rate of R2 elimination evidenced by transposon display data, unequal recombination occurs at an appreciable rate as it can be deduced in isolates' offspring. This can be explained if we consider GTMs as part of a mechanism pushing R2 load toward regions of the rDNA array that can be silenced.

Experimental procedures

Sampling and DNA isolation

Specimens were field-collected from 13 localities (Table 1, Fig. 2) and either reared in plastic boxes to obtain offspring or frozen at –80 °C until molecular analysis. Total DNA was extracted from single stick insect legs or from the whole body of first instar larvae with the standard phenol/chloroform protocol.

R2 isolation and sequencing

R2 presence was first checked by a PCR assay, using a primer anchored on the 28S rRNA gene, downstream the R2 insertion site, and a collection of degenerate primers, complementary to the element ORF region (Fig. 1; Kojima & Fujiwara, 2005). PCR amplifications were performed in a 50-μl reaction mix using the GoTaq Flexi (Promega, Madison, WI, USA), following the manufacturer's instructions. Thermal cycling was as follows: initial denaturation step at 95 °C for 5 min, 35 cycles of denaturation at 95 °C for 30 s, annealing at 48 °C for 30 s and extension at 72 °C for 3 min, and a final extension at 72 °C for 7 min. PCR products were run on a 1.5% agarose electrophoresis gel and bands of ~1800 bp were eluted using the Wizard SV Gel and PCR Clean-Up System kit (Promega). Fragments were inserted into a pGEM-T Easy vector (Promega) and used to transform *Escherichia coli* DH5α cells. Recombinant colonies were PCR-amplified with T7/SP6 primers and sequenced at Macrogen Inc. – Europe Lab (Amsterdam, The Netherlands).

This approach led to the recovery of R2 3′ ends from which R2 complete sequences were obtained through primer walking coupling a 28S-anchored primer (28S-F2), annealing upstream of the insertion site, with specifically designed primers complementary to R2 internal regions (Table S2; Fig. 1). Full-length R2 elements were therefore isolated and characterized from one female specimen each of gonochoric populations of *B. rossius rossius* (Capalbio, Tuscany) and *B. rossius redtenbacheri* (Patti, Sicily; Table 1; Fig. 2). Partial R2 sequences at the 3′ end were then isolated from one individual for each of the 13 localities reported in Table 1 and Fig. 2 through PCR amplification using the primer pair rR2din/28S-R (Table S2; Fig. 1). PCR reaction conditions, cloning and sequencing were as described above.

R2 sequence analyses

Sequences were edited and assembled using MEGA v. 5.2 (Tamura *et al*., 2011) and ORFs were searched with the ORF Finder tool (http://www.ncbi.nlm.nih.gov/gorf/gorf.html). Sequence alignments through the CLUSTALW algorithm, p-distance (p-D) ± SE and Minimum Evolution tree (with nodal support after 500 bootstrap replicates) were carried out using MEGA v. 5.2. Median-joining network (Bandelt *et al*., 1999) was calculated with Network v. 4.6 (Fluxus Engineering, Clare, Suffolk, UK). Tajima's *D* test for selection and the detection of gene conversion events were performed with DnaSP v. 5.1 (Librado & Rozas, 2009).

Estimation of R2 activity

The analysis was performed on selected parental individuals and in a sample (10–20 individuals) of their offspring (Table 2). In particular, we analysed the thelytokous offspring of two females for each of three parthenogenetic *B. rossius redtenbacheri* populations (Castanea delle Furie, Massa San Nicola and Curcuraci) and the male progeny of crosses involving as parental females either parthenogenetic specimens of *B. rossius redtenbacheri* from Curcuraci (reCUR ♀ X roANZ ♂, two crosses) or *B. rossius rossius* gonochoric females from Anzio (roANZ ♀ X roANZ ♂, one cross) or Capalbio (roCAP ♀ X roCAP ♂, two crosses). Males were chosen in the latter instances because they certainly represent descendants of bisexuality, while female progeny may still be of parthenogenetic origin. For each progeny, both early-hatched individuals and late-hatched individuals were chosen. This is of particular importance for cross descendants to be sure that they represent distinct meiotic products of the parental male.

The estimate of R2 activity was assayed through the insertion display technique: given that the R2 5′ end is subject to deletions of various lengths upon insertion, a single 5′ end deletion corresponds to a single insertion event. Through this technique, the various deletions occurring in a single individual can be visualized and, therefore, the number of insertions can be estimated. For R2 5′ end insertion display, we used a strategy based on PCR amplification, as both the element and the flanking region sequences are known. A 28S-anchored primer (28S-F2), annealing upstream of the insertion site, was alternately coupled with two primers, rR2rin2 and rR2rin, specifically designed to anneal 1241 bp and 1862 bp from the 5′ end, respectively (Table S2; Fig. 1). The two PCR reactions were performed with the GoTaq Flexi kit (Promega) following the manufacturer's instructions and the PCR conditions: initial denaturation at 95 °C for 2 min, 35 cycles of denaturation at 94 °C for 30 s, annealing at 48 °C for 30 s and extension at 72 °C for 1 min 30 s or 2 min (with the primers rR2rin2 and rR2rin, respectively) and a final extension at 72 °C for 5 min. PCR products were electrophoresed on a 1.5% agarose gel and the complete elements and every 5′ end deleted R2 resulted, thus, in a PCR amplification band.

The gel was then Southern blotted onto a positively charged nylon membrane. Membranes were hybridized with PCR-amplified R2-specific probes; the amplicon 28S-F2>rR2rin2 was hybridized with the probe rR2din2>rR2rin2, and the amplicon 28S-F2>rR2rin with the probe rR2din2>rR2rin, and the resulting signal was detected using the AlkPhos Direct Labelling and Detection System kit (GE Healthcare, Piscataway, NJ, USA). For each insertion display, two replicates were carried out to be sure of band scoring correctness. Data from the two PCR reactions were then combined to obtain a single 5′ end deletion profile per assayed individual and the number of bands visualized on the gel was calculated per individual and per offspring. These numbers are referred to as the number of sites occupied by R2 (S) and indicate the number of inserted 28S scored with this analysis. The following variables were calculated: the number of sites occupied

by R2 in parental individual(s) (S_p), the total number of sites occupied by R2 (S_o^{tot}) and the number of new sites occupied by R2 (new insertions; S_o^{new}) in the offspring.

Moreover, the R2 copy number per haploid genome ($\tilde{n}$) has been calculated in the offspring of both isolates and crosses as

$$\tilde{n} = \sum^{s} \frac{b_S}{N_O}$$

where b_S is the number of bands displayed at the S_{th} site occupied by R2 and N_O is the number of assayed offspring individuals (Charlesworth & Charlesworth, 1983). Insertion frequency spectra (i.e. the distributions of frequencies with which R2 insertions occur in the sample) were calculated, partitioning data on the basis of reproductive strategies: unisexual vs bisexual.

Expected frequencies in the progeny were calculated in the unisexual dataset taking into account that, because of their parthenogenesis mechanism, parental individuals are homozygous and so are their offspring (Scali *et al*., 2003): therefore, leaving aside new insertions, offspring should strictly inherit the same insertions as the mother. It is, thus, expected that the 100% of the mother's insertions should appear in the progeny (frequency = 1.0).

Expected frequencies in crosses' offspring, however, cannot be calculated so simply, as we have no information about the insertions' heterozygosity in parental individuals (except the two reCUR females that, being from parthenogenetic population, are homozygous); therefore, we drew three different simulations of Mendelian expectation for each R2 site: i) parental individuals are both homozygous (all.ho); ii) parental individuals, except reCUR females, are both heterozygous (all.he); and iii) only one of the two parental individuals is homozygous while the other is heterozygous (one.ho). In other words, we simulated simple Mendelian crosses between individuals considering their insertion pattern, revealed by the insertion display analysis, as being made by insertions all homozygous (all.ho), all heterozygous (all.he) or homozygous in one individual and heterozygous in the other one (one.ho). We, thus, obtained the genotype frequencies in the simulated progeny (from which we derived the expected insertion frequencies), assuming no recombination and no selection, and compared these data with those observed in the offspring insertion display experiments. The G-test was applied to check if observed insertion frequencies fit the expectation. Moreover, expected inherited R2 copy numbers per haploid genome based on expected insertion frequencies have been calculated and compared with the observed ones, computed as above, but excluding new insertions (i.e. S_o^{new}).

Data availability

Sequences have been submitted to GenBank under the following accession numbers: **KJ958674** ($R2Br^{fun}$), **KJ958675** ($R2Br^{deg}$) and **KJ958565-KJ958673** (partial R2 elements for population survey).

Acknowledgements

This work was supported by RFO and Canziani funding to B.M. We are grateful to Danna Eickbush for the critical reading of the manuscript and for fruitful discussions and to two anonymous referees which helped greatly in improving the manuscript.

References

Ambrose, C.D. and Crease, T.J. (2011) Evolution of the nuclear ribosomal DNA intergenic spacer in four species of the *Daphnia pulex* complex. *BMC Genet* **12**: 13.

Arkhipova, I.R. (2005) Mobile genetic elements and sexual reproduction. *Cytogenet Genome Res* **110**: 372–382.

Bandelt, H.J., Forster, P. and Röhl, A. (1999) Median-joining networks for inferring intraspecific phylogenies. *Mol Biol Evol* **16**: 37–48.

Burke, W.D., Eickbush, D.G., Xiong, Y., Jakubczak, J. and Eickbush, T.H. (1993) Sequence relationship of retrotransposable elements R1 and R2 within and between divergent insect species. *Mol Biol Evol* **10**: 163–185.

Cesari, M., Luchetti, A., Passamonti, M., Scali, V. and Mantovani, B. (2003) Polymerase chain reaction amplification of the Bag320 satellite family reveals the ancestral library and past gene conversion events in *Bacillus rossius* (Insecta Phasmatodea). *Gene* **312**: 289–295.

Charlesworth, B. and Charlesworth, D. (1983) The population dynamics of transposable elements. *Genet Res* **42**: 1–27.

Christensen, S.M., Ye, J. and Eickbush, T.H. (2006) RNA from the 5′ end of the R2 retrotransposon controls R2 protein binding to and cleavage of its DNA target site. *Proc Natl Acad Sci USA* **103**: 17602–17607.

Dolgin, E.S. and Charlesworth, B. (2006) The fate of transposable elements in asexual populations. *Genetics* **174**: 817–827.

Dolgin, E.S., Charlesworth, B. and Cutter, A.D. (2008) Population frequencies of transposable elements in selfing and outcrossing *Caenorhabditis* nematodes. *Genet Res* **90**: 317–329.

Doolittle, W.F. and Sapienza, C. (1980) Selfish genes, the phenotype paradigm and genome evolution. *Nature* **284**: 601–603.

Eickbush, D.G. and Eickbush, T.H. (2012) R2 and R2/R1 hybrid non-autonomous retrotransposons derived by internal deletions of full-length elements. *Mobile DNA* **3**: 10.

Eickbush, D.G., Ye, J., Zhang, X., Burke, W.D. and Eickbush, T.H. (2008) Epigenetic regulation of retrotransposon within the nucleolus of *Drosophila*. *Mol Cell Biol* **28**: 6452–6461.

Eickbush, T.H. (2002) R2 and related site-specific non-LTR retrotransposons. In *Mobile DNA II* (Craig, N., Craigie, R., Gellert, M. and Lambowit, A., eds), pp. 813–835. American Society of Microbiology Press, Washington, DC.

Eickbush, T.H. and Eickbush, D.G. (2007) Finely orchestrated movements: evolution of the ribosomal RNA genes. *Genetics* **175**: 477–485.

Ghesini, S., Luchetti, A., Marini, M. and Mantovani, M. (2011) The non-LTR retrotransposon R2 in termites (Insecta, Isoptera): characterization and dynamics. *J Mol Evol* **72**: 296–305.

Hickey, D.A. (1982) Selfish DNA: a sexually-transmitted nuclear parasite. *Genetics* **101**: 519–531.

Hiruta, C., Nishida, C. and Tochinai, S. (2010) Abortive meiosis in the oogenesis of parthenogenetic *Daphnia pulex*. *Chromosome Res* **18**: 833–840.

Kagramanova, A.S., Korolev, A.L. and Mukha, D.M.V. (2010) Analysis of the inheritance patterns of 5′-truncated copies of the german cockroach R2 retroposons in individual crosses. *Russian J Genet* **46**: 1290–1294.

Kazazian, H.H. (2004) Mobile elements: drivers of genome evolution. *Science* **303**: 1626–1632.

Kojima, K.K. and Fujiwara, H. (2005) Long-term inheritance of the 28S rDNA-specific retrotransposon R2. *Mol Biol Evol* **22**: 2157–2165.

Kraaijeveld, K., Zwanenburg, B., Hubert, B., Vieira, C., de Pater, S., van Alphen, J.J.M. *et al.* (2012) Transposon proliferation in an asexual parasitoid. *Mol Ecol* **21**: 3898–3906.

Librado, P. and Rozas, J. (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25**: 1451–1452.

Lockton, S. and Gaut, B.S. (2010) The evolution of transposable elements in natural populations of self-fertilizing *Arabidopsis thaliana* and its outcrossing relative *Arabidopsis lyrata*. *BMC Evol Biol* **10**: 10.

Loewe, L. and Lamatsch, D.K. (2008) Quantifying the threat of extinction from Muller's ratchet in the diploid Amazon molly (*Poecilia formosa*). *BMC Evol Biol* **19**: 88.

Luchetti, A. and Mantovani, B. (2013) Non-LTR R2 element evolutionary patterns: phylogenetic incongruences, rapid radiation and the maintenance of multiple lineages. *PLoS ONE* **8**: e57076.

Luchetti, A., Cesari, M., Carrara, G., Cavicchi, S., Passamonti, M., Scali, V. *et al.* (2003) Unisexuality and molecular drive: Bag320 sequence diversity in *Bacillus* taxa (Insecta Phasmatodea). *J Mol Evol* **56**: 587–596.

Malik, H.S. and Eickbush, T.H. (2000) NeSL-1, an ancient lineage of site-specific non-LTR retrotransposons from *Caenorhabditis elegans*. *Genetics* **154**: 193–203.

Malik, H.S., Burke, W.D. and Eickbush, T.H. (1999) The age and evolution of non-LTR retrotransposable elements. *Mol Biol Evol* **16**: 793–805.

Mantovani, B., Passamonti, M. and Scali, V. (2001) The mitochondrial cytochrome oxidase II gene in *Bacillus* stick insects: ancestry of hybrids, androgenesis, and phylogenetic relationships. *Mol Phylogenet Evol* **19**: 157–163.

Mingazzini, V. (2011) Transposable elements: structure and dynamic of the autonomous retrotransposon R2 in Arthropoda with non-canonical reproduction. PhD thesis, University of Bologna.

Mingazzini, V., Luchetti, A. and Mantovani, B. (2011) R2 dynamics in *Triops cancriformis* (Bosc, 1801) (Crustacea, Branchiopoda, Notostraca): turnover rate and 28S concerted evolution. *Heredity* **106**: 567–575.

Nuzhdin, S.V. and Petrov, D.A. (2003) Transposable elements in clonal lineages: lethal hangover from sex. *Biol J Linn Soc* **79**: 33–41.

O'Donnell, K.A. and Burns, K.H. (2010) Mobilizing diversity: transposable element insertions in genetic variation and disease. *Mobile DNA* **1**: 21.

Omilian, A.R., Cristescu, M.E.A., Dudycha, J.L. and Lynch, M. (2006) Ameiotic recombination in asexual lineages of *Daphnia*. *Proc Natl Acad Sci USA* **103**: 18638–18643.

Pérez-Gonzalez, C.E. and Eickbush, T.H. (2001) Dynamics of R1 and R2 elements in the rDNA locus of *Drosophila simulans*. *Genetics* **158**: 1557–1567.

Plohl, M., Luchetti, A., Meštrovic′, N. and Mantovani, B. (2008) Satellite DNAs between selfishness and functionality: structure, genomics and evolution of tandem repeats in centromeric (hetero)chromatin. *Gene* **409**: 72–82.

Ricci, A., Scali, V. and Passamonti, M. (2008) The IGS-ETS in *Bacillus* (Insecta Phasmida): molecular characterization and the relevance of sex in ribosomal DNA evolution. *BMC Evol Biol* **8**: 278.

Scali, V., Passamonti, M., Marescalchi, O. and Mantovani, B. (2003) Linkage between sexual and asexual lineages: genome evolution in *Bacillus* stick insects. *Biol J Linn Soc* **79**: 137–150.

Schaack, S., Pritham, E.J., Wolf, A. and Lynch, M. (2010) DNA transposon dynamics in populations of *Daphnia pulex* with and without sex. *Proc R Soc Lond B* **277**: 2381–2387.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. and Kumar, S. (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance and maximum parsimony methods. *Mol Biol Evol* **28**: 2731–2739.

Valizadeh, P. and Crease, T.J. (2008) The association between breeding system and transposable element dynamics in *Daphnia pulex*. *J Mol Evol* **66**: 643–654.

Wright, S. and Finnegan, D. (2001) Genome evolution: sex and transposable elements. *Curr Biol* **11**: R296–R299.

Zhang, X. and Eickbush, T.H. (2005) Characterization of active R2 retrotransposition in the rDNA locus of *Drosophila simulans*. *Genetics* **170**: 195–205.

Zhang, X., Zhou, J. and Eickbush, T.H. (2008) Rapid R2 retrotransposition leads to the loss of previously inserted copies via large deletions of the rDNA locus. *Mol Biol Evol* **25**: 229–237.

Zhou, J. and Eickbush, T.H. (2009) The pattern of R2 retrotransposon activity in natural populations of *Drosophila simulans* reflects the dynamic nature of the rDNA locus. *Plos Genet* **5**: e1000386.

Zhou, J., Eickbush, M.T. and Eickbush, T.H. (2013) A population genetic model for the maintenance of R2 retrotransposons in rRNA gene loci. *PLoS Genet* **9**: e1003179.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Figure S1. Biogeographic distribution of *Bacillus* stick insect taxa.

Figure S2. (A) Scenario inferred for the all.he simulation with unequal recombination in both parental individuals. Post-selection insertion frequencies are significantly different from those calculated for no recombination and post-recombination ($G = 14.5$; d.f. = 3; $P < 0.01$). (B) Scenario inferred for the one.ho simulation with unequal recombination in the heterozygous parental. (C) Scenario inferred for the one.ho simulation with unequal recombination in the homozygous parental. Post-recombination insertion frequencies are significantly different from those calculated for no recombination ($G = 29.8$; d.f. = 3; $P < 0.001$). Post-selection insertion frequencies are not significantly different from those calculated for post recombination ($G = 0.97$; d.f. = 3; $P = 0.241$). (D) Scenario inferred for the one.ho simulation with unequal recombination in both parental individuals. Post-recombination insertion frequencies are significantly different from those calculated for no recombination ($G = 20.4$; d.f. = 3; $P < 0.001$). Post-selection insertion frequencies are not significantly different from those calculated for post recombination ($G = 0.07$; d.f. = 3; $P = 0.105$).

Table S1. Goodness-of-fit G test of observed R2 insertion frequencies to Mendelian expectation under the three scenarios all.ho, all.he and one.ho for the bisexual offspring (see text).

Table S2. Primers used for R2 PCR amplification and sequencing.

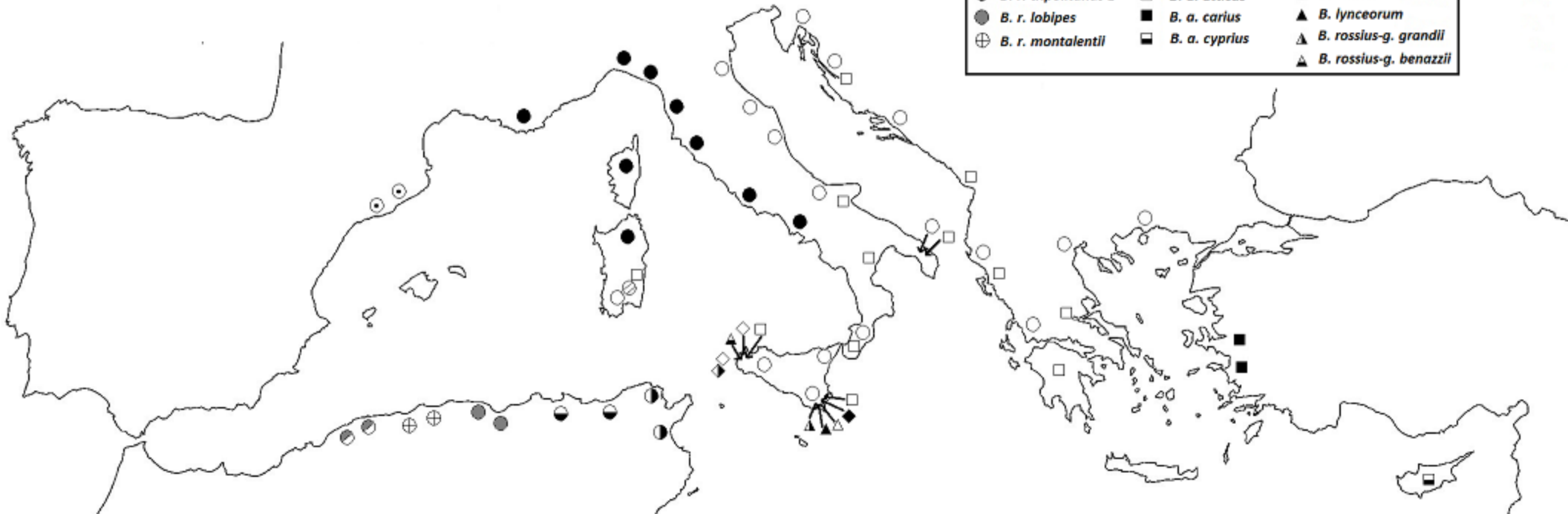

B. r. rossius
B. r. redtenbacheri
B. r. tripolitanus A
B. r. tripolitanus B
B. r. lobipes
B. r. montalentii
B. r. medeae
B. r. cataluniae
B. r. ssp.
B. a. atticus
B. a. carius
B. a. cyprius
B. g. grandii
B. g. benazzii
B. g. maretimi
B. whitei
B. lynceorum
B. rossius-g. grandii
B. rossius-g. benazzii

Figure S2A. Scenario inferred for the all.he simulation with unequal recombination in both parentals. Post-selection insertion frequencies are significantly different from those calculated for no recombination and post-recombination (G = 14.5; d.f. = 3; p < 0.01).

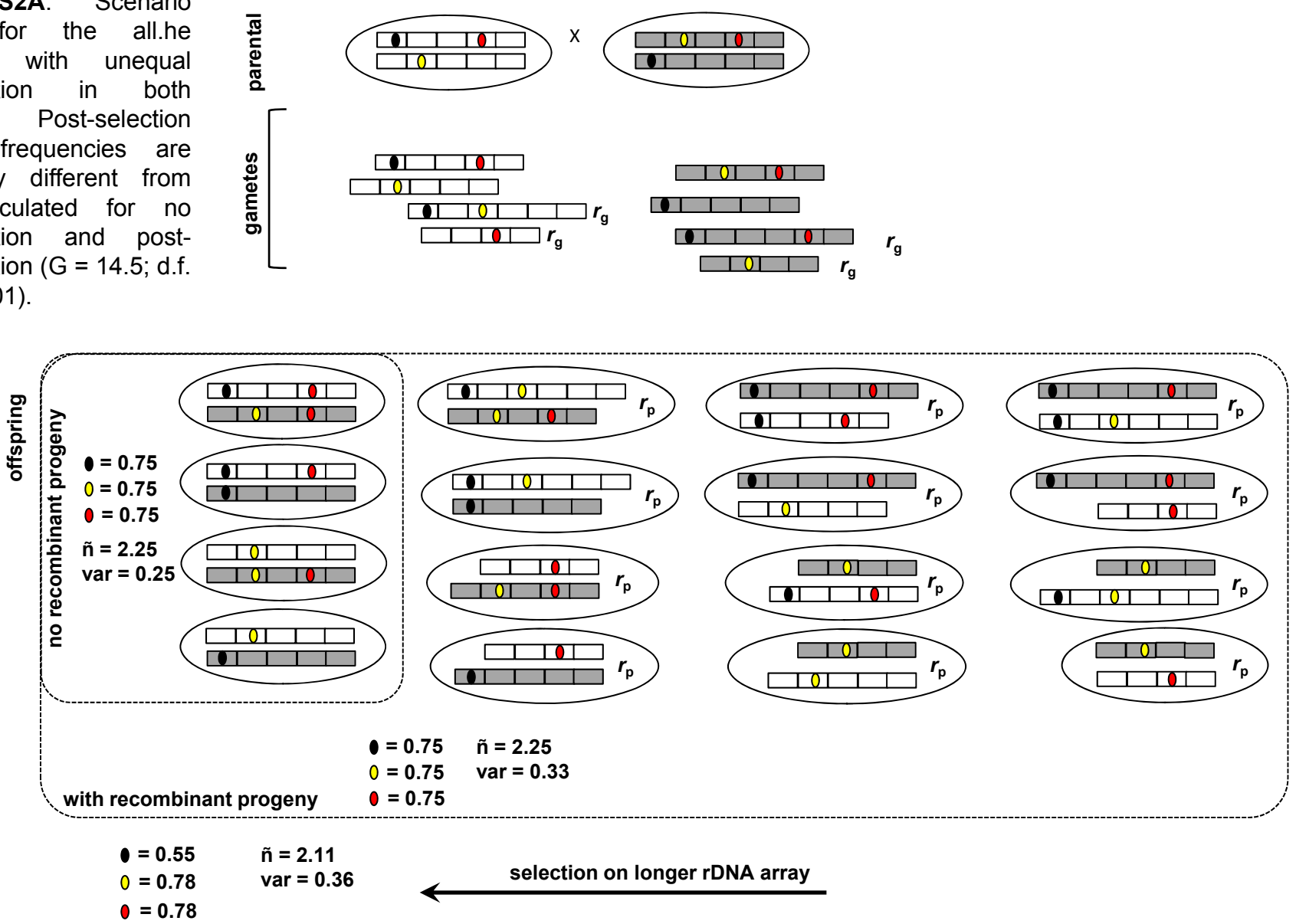


Figure S2B. Scenario inferred for the one.ho simulation with unequal recombination in the heterozygous parental.

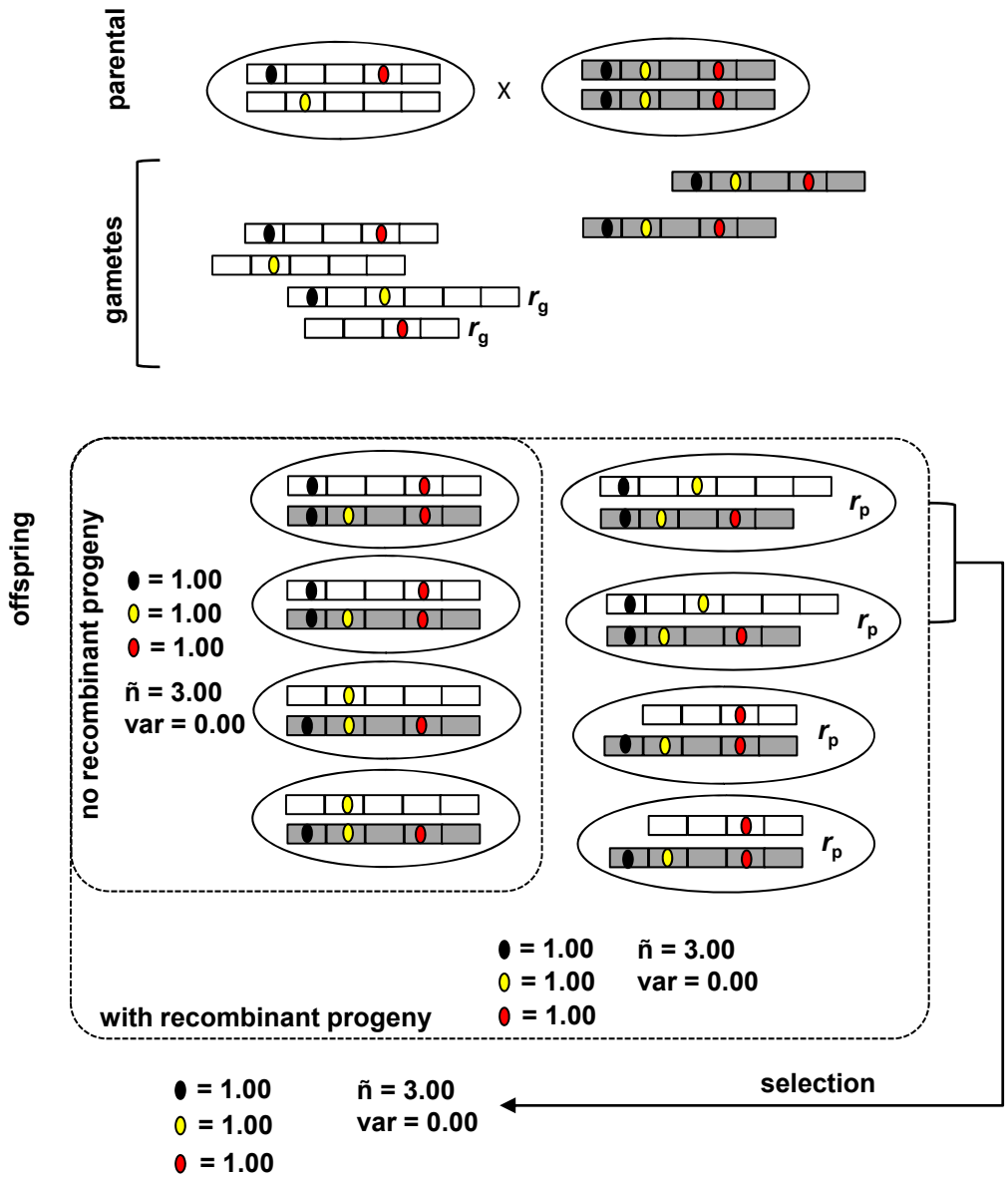


Figure S2C. Scenario inferred for the one.ho simulation with unequal recombination in the homozygous parental. Post-recombination insertion frequencies are significantly different from those calculated for no recombination (G = 29.8; d.f. = 3; p < 0.001). Post-selection insertion frequencies are not significantly different from those calculated for post recombination (G = 0.97; d.f. = 3; p = 0.241).

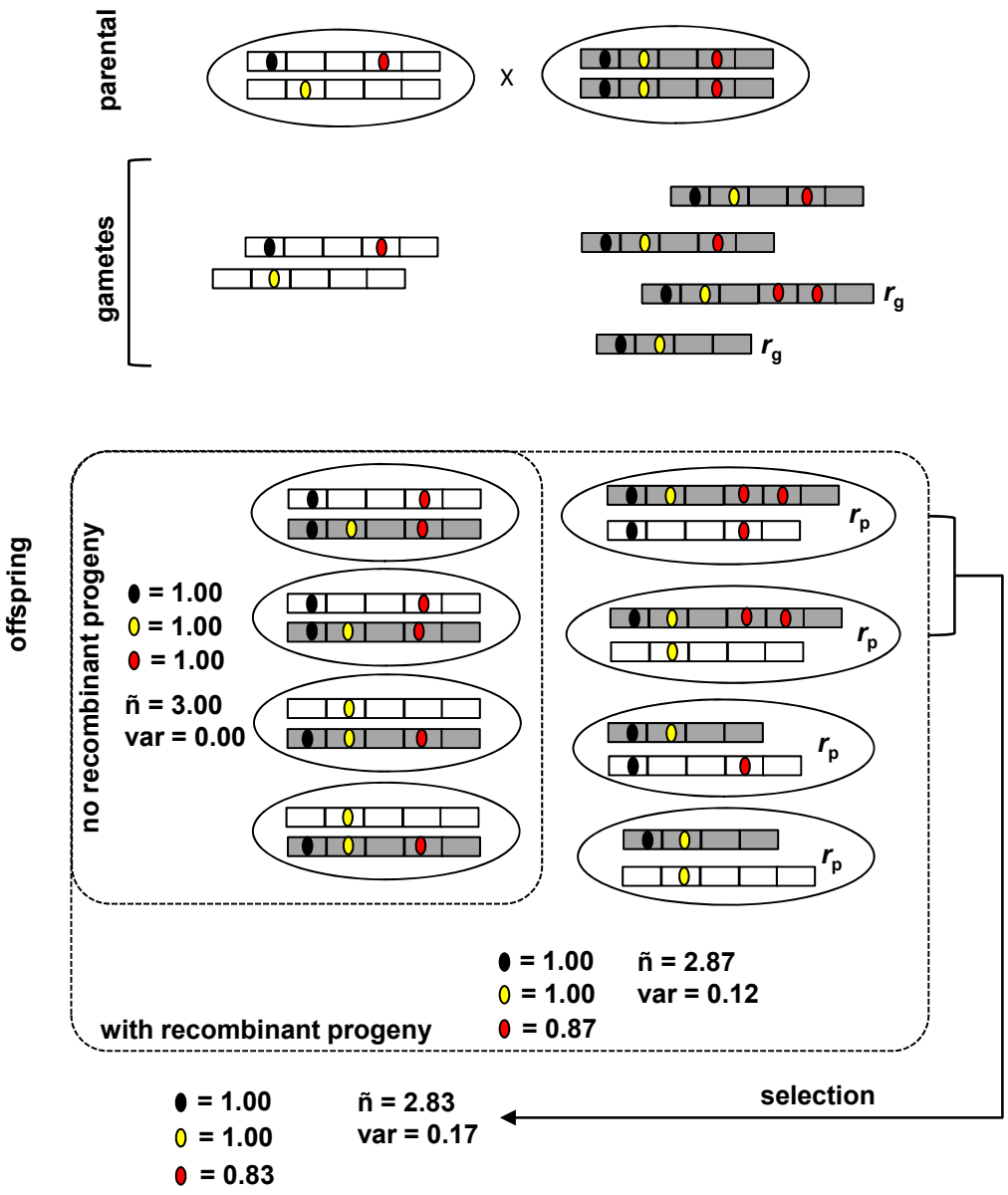
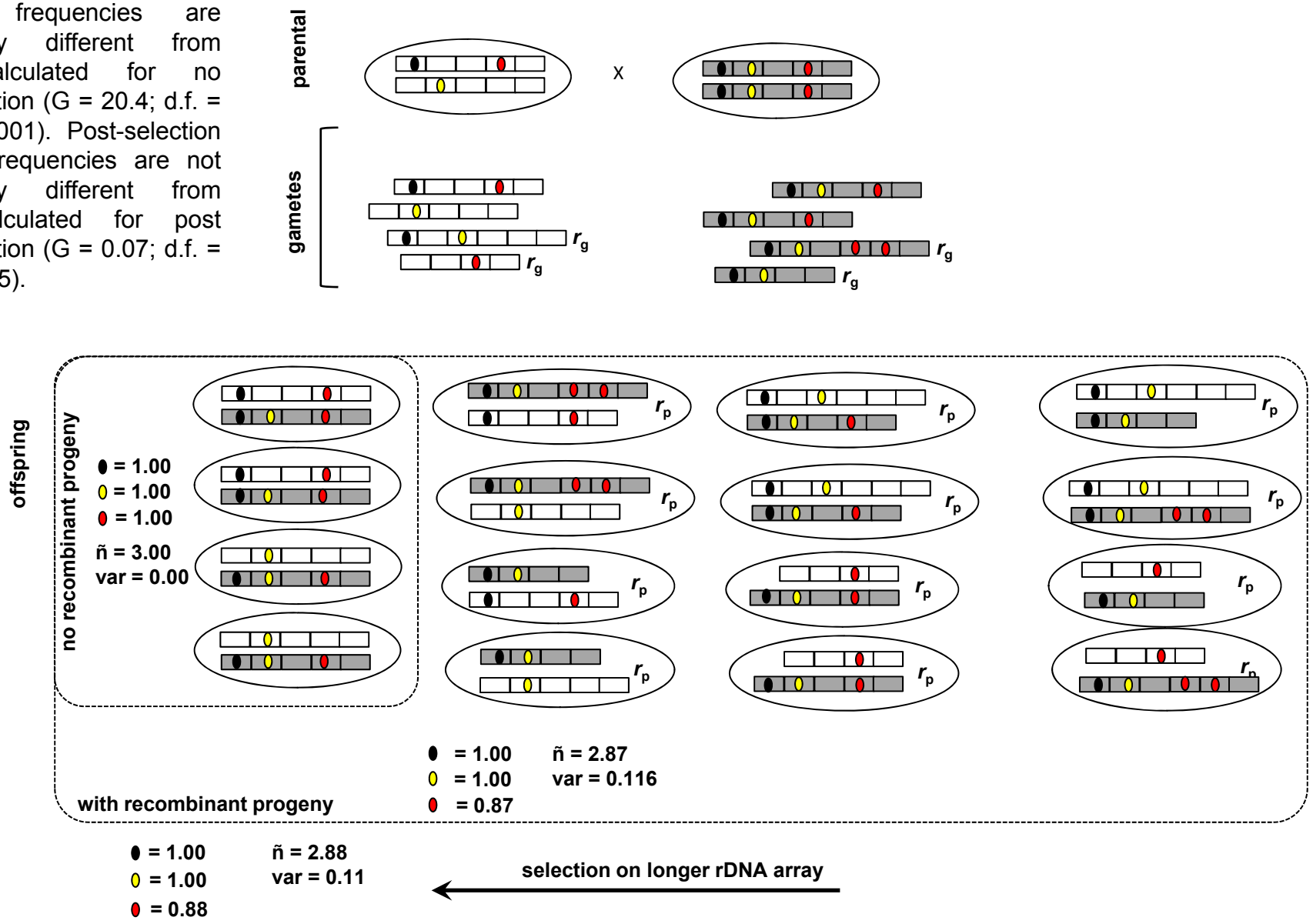


Figure S2D. Scenario inferred for the one.ho simulation with unequal recombination in both parentals. Post-recombination insertion frequencies are significantly different from those calculated for no recombination (G = 20.4; d.f. = 3; p < 0.001). Post-selection insertion frequencies are not significantly different from those calculated for post recombination (G = 0.07; d.f. = 3; p = 0.105).



Supporting Discussion of Fig. S2B-D. One.ho simulation scenarios description. If unequal recombination occurs in the heterozygous parental, nothing changes even after selection: in the progeny all insertion will be detected at 100% frequency (Suppl. Fig. S2B). However, the offspring is now heterozygous for at least one insertion, therefore it is likely that changes in frequencies and copy number will occur in the following generations. On the other hand, insertion frequencies deviate and copy number decreases just after recombination occurs, weakly changing after selection, when the homozygous partner undergoes unequal recombination (Suppl. Fig. S2C). Finally, if both parental in the one.ho simulation produce recombinant, frequencies and copy number decrease occur but there is weak effect of selection (Suppl. Fig. S2D).

Table S1. Goodness-of-fit G test of observed R2 insertion frequencies to Mendelian expectation under the three scenarios all.ho, all.he and one.ho for the bisexual offspring (see text).

Crosses	Expected frequency sets	G (d.f.)	p
reCUR-♀5 × roANZ-♂15	all.ho	1640.9 (11)	< 0.0001
	all.he	385.4 (11)	< 0.0001
	one.ho	385.4 (11)	< 0.0001
reCUR-♀6 × roANZ-♂9	all.ho	214.7 (4)	< 0.0001
	all.he	214.7 (4)	< 0.0001
	one.ho	214.7 (4)	< 0.0001
roANZ-♀21 × roANZ-♂8	all.ho	4283.1 (19)	< 0.0001
	all.he	115.9 (19)	< 0.0001
	one.ho	1202 (19)	< 0.0001
roCAP-♀1 × roCAP-♂4	all.ho	1468.3 (15)	< 0.0001
	all.he	16.53 (15)	0.063
	one.ho	451.84 (15)	< 0.0001
roCAP-♀2 × roCAP-♂3	all.ho	1931.8 (12)	< 0.0001
	all.he	48.95 (12)	< 0.0001
	one.ho	307.00 (12)	< 0.0001

Table S2. Primers used for R2 PCR amplification and sequencing.

Primer name	Sequence 5' -> 3'	Reference
R2IF1	AAGCARGGNGAYCCNCTNTC	Kojima and Fujiwara (2005)
28S-F2	GAATCCGACTGTCTAATTAAAACAAAG	Mingazzini et al. (2011)
28S-R	TCCATTGCTGCGCGTCACTAATTAGATGAC	this study
rR2rin	GACTGTCCAACAATAGGAGGGAAT	this study
rR2rin2	CACCAGGAGATTAGTTTGGTTTCT	this study
rR2din2	GCATGTCCAAGGATAAAGTCTAAAA	this study
rR2din	AACGACTATCAGTTCGTTGAATAGG	this study
rR2Prin	GACTGTCCAACAATAGGAGGGAAT	this study
rR2Pdin	GCATGTCCAAGGATAAAGTCTAAAA	this study
rR2Prin2	CACCAGGAGATTAGTTTGGTTTCT	this study

2.2 Obligatory parthenogenesis and TE load

Bonandin, L., Scavariello, C., Mingazzini, V., Luchetti, A. and Mantovani, B. (2016) Obligatory parthenogenesis and TE load: *Bacillus* stick insects and the R2 non-LTR retrotransposon. Accepted in Insect Science, doi: 10.1111/1744-7917.12322.

The results presented in this paper widen the analyses about R2 activity in two more *Bacillus* species: the strictly gonochoric *B. grandii maretimi* and the obligate parthenogen *B. atticus atticus*. Through the insertion display approach, I analysed R2 dynamics at the population level in both taxa, and in the offspring of three *B. atticus atticus* females. These data were obviously compared with previous one on the facultative parthenogen *B. rossius*. On the whole, main hints are:

- bisexual/facultative parthenogenetic populations have higher R2 load than obligatory parthenogenetic ones at the population level, with *B. atticus atticus* showing R2 presence at a very low content; the occurrence of truncated variants indicates that in the obligatory unisexual *B. atticus atticus* R2 is active but mechanisms of molecular turnover are effective;
- in *B. atticus atticus* offspring the R2 load is even further reduced; this indicates that, at variance of the facultative parthenogenetic *B. rossius,* the R2 activity is held at a lower rate in *B. atticus*. This appears of particular interest for R2 and the taxon evolutionary survival.

Moreover, models of R2 parental-offspring descendants under different reproductive strategies (bisexuality, facultative parthenogenesis and obligatory parthenogenesis) suggest that recombination hold a major role than selection, in counteracting the R2 deleterious proliferation. However, also silencing mechanisms and even clonal selection should be considered.

I also studied R2 dynamics up to the 5th generation of an isolated female of the facultative parthenogenetic *B. rossius redtenbacheri* population from Castanea delle Furie. On the whole, 27 individuals were analysed. R2 accumulation through the different generations was always low (unpublished data).

From: <zhaoyx@ioz.ac.cn<mailto:zhaoyx@ioz.ac.cn>>
Subject: Insect Science - Decision on Manuscript ID INS-2015-09-366.R1 [email ref: DL-RW-1-a]
Date: 4 January 2016 at 03:16:04 GMT+1
To: <andrea.luchetti@unibo.it<mailto:andrea.luchetti@unibo.it>>

03-Jan-2016

Dear Dr. Luchetti:

It is a pleasure to accept your manuscript entitled "Obligatory parthenogenesis and TE load: Bacillus stick insects and the R2 non-LTR retrotransposon" in its current form for publication in Insect Science. The comments of the reviewer(s) who reviewed your manuscript are included at the foot of this letter.

Please note although the manuscript is accepted the files will now be checked to ensure that everything is ready for publication, and you may be contacted if final versions of files for publication are required.

Thank you for your fine contribution. On behalf of the Editors of Insect Science, we look forward to your continued contributions to the Journal.

Sincerely,
Dr. Yun Xian Zhao
Executive editor, Insect Science
zhaoyx@ioz.ac.cn<mailto:zhaoyx@ioz.ac.cn>

Obligatory parthenogenesis and TE load: *Bacillus* stick insects and the R2 non-LTR retrotransposon

Livia Bonandin, Claudia Scavariello, Valentina Mingazzini, Andrea Luchetti*, and Barbara Mantovani

Dipartimento di Scienze Biologiche, Geologiche e Ambientali, Università di Bologna, Bologna, Italy

Running title: *Bacillus* stick insects and R2 load

Keywords: insertion activity, Muller's ratchet, Phasmida, recombination, retrotransposon, unisexuality.

***Correspondence:** *Andrea Luchetti, Dipartimento di Scienze Biologiche, Geologiche e Ambientali, via Selmi 3, 40126, Università di Bologna, Bologna, Italy*; phone: +39-051-2094173; fax: +39-051-2094286; e-mail: andrea.luchetti@unibo.it

Abstract

Transposable elements (TEs) are selfish genetic elements whose self-replication is contrasted by the host genome. In this context, host reproductive strategies are predicted to impact on both TEs load and activity. The presence and insertion distribution of the non-LTR retrotransposon R2 was here studied in populations of the strictly bisexual *Bacillus grandii maretimi* and of the obligatory parthenogenetic *B. atticus atticus*. Furthermore, data were also obtained from the offspring of selected *B. a. atticus* females. At the population level, the gonochoric *B. g. maretimi* showed a significantly higher R2 load than the obligatory parthenogenetic *B. a. atticus.* The comparison with bisexual and unisexual *B. rossius* populations showed that their values were higher than those recorded for *B. a. atticus* and similar, or even higher, than those of *B. g. maretimi*. Consistently, an R2 load reduction is scored in *B. a. atticus* offspring even if with a great variance. On the whole, data here produced indicate that in the obligatory unisexual *B. a. atticus* R2 is active and that mechanisms of molecular turnover are effective. Furthermore, progeny analyses show that, at variance of the facultative parthenogenetic *B. rossius,* the R2 activity is held at a lower rate. Modelling parental-offspring inheritance, suggests that in *B. a. atticus* recombination plays a major role in eliminating insertions rather than selection, as previously suggested for unisexual *B. rossius* progeny, even if in both cases a high variance is observed. In addition to this, mechanisms of R2 silencing or chances of clonal selection cannot be ruled out.

Introduction

Transposable elements (TEs) evolutionary dynamics reflects the interplay between their ability to replicate to high copy number, affecting the host fitness, and the effectiveness of the mechanisms set up by the host genome to limit their proliferation. The latter aspect is further linked to the evolutionary dynamics of host organisms, with special regard to their reproductive strategies. In fact, theoretical and empirical studies evidenced a causal link between differential TE loads and reproductive strategies, with the bisexuality allowing the control of TEs deleterious proliferation through recombination and the independent assortment of chromosomes which give to natural selection more opportunity to operate. On the contrary, selfing or non-recombining genomes either accumulate TEs until host lineage extinction or proceed with TEs elimination by genome purging (Wright & Finnegan, 2001; Nuzhdin & Petrov, 2003 and references therein). Only few studies have empirically considered this topic with different results (Valizadeh & Crease, 2008; Schaack *et al.*, 2010; Kraaijeveld *et al.*, 2012).

Transposable elements (TEs) of the R series are non-LTR elements known to specifically insert in the 28S (R1, R2, R4, R5, R6, R9, and RT families) or into the 18S (R7 and R8 families) ribosomal genes (Kojima *et al.*, 2006; Gladyshev & Arkhipova, 2009). R2 is the most studied non-LTR TE and it was used as a model to decipher the non-LTR mechanism of mobilisation and re-integration, that appears also mediated by recombination (Eickbush, 2002; Fujimoto *et al.*, 2004; Mukha *et al.*, 2013).

R2 inserts at the 5'-TTAAGG/TAGCCA-3' sequence of the 28S rRNA gene through a 'copy-out/copy-in' mechanism known as target primed reverse transcription complemented by recombination. During its re-integration, the first strand synthesis may be incomplete giving rise to 5' end truncated copies: these length variations can be used to screen the element activity (Pérez-Gonzalez & Eickbush, 2001). R2 occurs across the

Animal kingdom, from the diploblastic phylum Cnidaria to Vertebrates (Kojima & Fujiwara, 2005; Luchetti & Mantovani 2013). Divergence *vs* age evaluations revealed the absence of horizontal transfer and placed the origin of R2 between 600 and 850 Myr ago in the Pre-Cambrian Era (Kojima & Fujiwara, 2005).

R2 dynamics is also affected by the peculiar evolution of its niche represented by the rRNA tandemly repeated genes. It is known that multiple copies from the same genome are more similar to each other than between genomes. Three models have been proposed to explain this pattern: the concerted evolution, the birth-and-death and the magnification-and-fixation models (Dover, 1982; Nei & Rooney, 2005; Mukha *et al.*, 2011). In the first model, the molecular drive (a dual process that takes place through homogenization and fixation of sequence variants) leads to a variability pattern known as concerted evolution (Plohl *et al.*, 2008). Homogenization occurs at the genomic level with a variety of genomic turnover mechanisms, such as unequal crossing-over, gene conversion or replication slippage: these allow the spread of new units within the same genome. Fixation promotes the diffusion of sequence variants at the population level, through the bisexual reproduction. In the birth-and-death model, tandem repeat variants are subject to multiple duplications followed by differential deletion of some copies, eventually resulting in more (or less) homogeneous tandem repeat array (Nei & Rooney, 2005). Finally, in the magnification-and-fixation model two steps are hypothesized: a sudden amplification of certain repeat variants with possible adaptive significance (magnification), followed by their spreading along the array (fixation) driven by directional gene conversion and purifying selection (Mukha *et al.*, 2011).

The genus *Bacillus* Peletier de Saint Fargeau & Serville, 1828 (Insecta Phasmida) comprises the western Mediterranean facultative parthenogenetic *B. rossius* (Rossi, 1790) (2*n* = 35/36, XO-XX), with eight races; the strictly bisexual *B. grandii* Nascetti & Bullini, 1981 (2*n* = 33/34, XO-XX), endemic to small areas in the Sicilian region and formally split

into the three subspecies *B. grandii grandii* Nascetti & Bullini, 1981*, B. grandii benazzii* Scali 1991*,* and *B. grandii maretimi* Scali & Mantovani, 1990; the eastern Mediterranean *B. atticus* Brunner, an all-female obligate parthenogenetic taxon that includes a complex of three different karyological (2*n* = 32; 2*n* = 34; 3*n* = 48–51) and allozyme races. These species are the ancestors of the Sicilian parthenogenetic hybrids *B. whitei* (= *rossius/grandii*) Nascetti & Bullini, 1982 and *B. lynceorum* (= *rossius/grandii/atticus*) Bullini, Nascetti & Bianchi Bullini, 1983. In Sicily, *B. rossius* and *B. grandii* have also hybridized to produce two different hybridogenetic strains (*sensu* Schultz, 1961), which pass to their progeny the maternal *rossius* haploset, renewing the paternal genome at each generation through crosses with syntopic *B. grandii* males.

Parthenogenesis in *Bacillus* takes place with different mechanisms (Scali *et al*., 2003). In particular, facultative parthenogenesis in *B. rossius* comprises a normal meiosis; diploidization occurs in some cells of the blastula through anaphase restitution, this automictic mechanism leading to a female offspring homozygous at all loci in one generation. The meiotic divisions in the eggs of the obligatory parthenogenetic diploid *B. atticus atticus* (2*n* = 34) takes place through a first reduction division, the two nuclei then fusing to restore a diploid egg nucleus at the onset of prophase II. The second meiotic division produces a degenerating polocyte and a quickly dividing, diploid nucleus from which the parthenogenetic embryo develops. This automictic type of parthenogenesis explains the clonal maintenance of chromosomal rearrangements and of fixed heterozygosity at some loci, but it allows the production also of recombinants if an appropriate cross-over occurs during the first division. On the other hand, in the hybrids an apomictic mechanism occurs.

The mobile DNA content of *B. rossius, B. grandii* and *B. atticus* was first analyzed through the survey of random genomic libraries (Ricci *et al*., 2013). This lead to the scoring of 19 families of autonomous transposable elements (two LTR and six non-LTR

retrotransposons; 11 DNA transposons) and of the first MITE family observed in polyneopteran genomes. Contrary to expectations, the obligatory parthenogenetic *B. atticus* does not show a lower transposable element content compared with the other two species analyzed (22.9% vs 23.3% in *B. rossius* and 18% in *B. grandii*). It is impossible to state whether *B. atticus* selected for less harmful TEs or the taxon is not unisexual since enough time to allow an efficient elimination of the TE load. It is worth noting that the origin of *B. atticus* dates back to 15.37 ± 2.65 Myr ago, but the hybrid origin of *B. lynceorum* suggests that it was bisexual until 1 Myr ago (Mantovani *et al*., 2001).

As far as R2 is concerned, we showed that the *B. rossius* genome hosts a functional ($R2Br^{fun}$) and a closely related, degenerate ($R2Br^{deg}$) element (Bonandin *et al*., 2014). Both elements actively replicate, but while $R2Br^{fun}$ showed a canonical pattern of a functional R2, $R2Br^{deg}$ copies are exclusively full-length, i.e. they have no associated 5' end deletions, and lost the ability to fold into the ribozyme necessary for R2 RNA maturation, suggesting they replicate via 28S rDNA unequal recombination (Martoni *et al*., 2015).

Unisexual and bisexual *B. rossius* offspring have a similar number of $R2Br^{fun}$-occupied sites, that could be explained by multiple or recent shifts from gonochoric to parthenogenetic reproduction. On the other hand, higher-frequency insertions in unisexual offspring significantly outnumber those in bisexual offspring, suggesting that unisexuals can eliminate insertions with lower efficiency (Bonandin *et al*., 2014).

In the present paper, we survey R2 insertion patterns in field collected populations and lab reared offspring of the obligatory bisexual *B. g. maretimi* and the obligatory parthenogenetic *B. a. atticus* to screen the R2 inheritance and to compare results with those obtained from bisexual and unisexual populations of the facultative parthenogenetic *B. rossius*.

Materials and Methods

Sampling and DNA isolation

Stick insects samples of the strictly gonochoric *Bacillus g. maretimi* and of the obligatory parthenogenetic *Bacillus a. atticus* were field-collected in the Sicilian area and either immediately frozen at -80 °C or reared in the laboratory to obtain offspring for molecular analysis. Total DNA was extracted from single stick insect legs or from the whole body of first instar larvae with the standard phenol/chloroform protocol.

R2s from *B. g. maretimi* and *B. a. atticus* have been sequenced following the strategy described in Bonandin *et al.* (2014) and their detailed description will be given in a future report (Scavariello, unpublished data).

R2 activity was determined in field collected females and males of *B. g. maretimi* (Marettimo, Egadi Islands, Sicily) and in females of two *B. a. atticus* populations (Necropoli Camarina and Scoglitti, near Ragusa, southern Sicily) (Suppl. Table S1). The analysis was not performed on *B. g. grandii and B. g. benazzi* owing to the lack of sufficient specimens. Results were compared with data from Mingazzini (2011).

For comparison with the facultative parthenogenetic *B. rossius* (Bonandin *et al*., 2014), R2 activity analyses were performed also in selected parental females of the obligate parthenogenetic *B. a. atticus* and in a sample (10 individuals) of their offspring (Suppl. Table S1). In particular, the thelytokous progeny of three females field-collected in Necropoli Camarina (NC ♀2) and in Scoglitti (SCO ♀1, SCO ♀2; Suppl. Table S1) were analyzed. Both early- and late-hatched individuals from each progeny were analyzed.

Insertion analyses

R2 activity was determined, as described in Pérez-Gonzalez & Eickbush (2001) through the insertion display of the 5' truncated copies generated during the target-primed reverse transcription mechanism adopted by R2 for its transposition.

Insertion analysis of *B. g. maretimi* and *B. a. atticus* specimens was performed using a 28S-anchored primer (28S-F2 or 28SFBGG) annealing upstream of the element insertion site, coupled with a specific primer (R2Ins2_R or BGGR2ARIN) annealing inside R2, 1448 bp or 2906 bp from the 5' end (Suppl. Table S2). PCR amplifications with the GoTaq Flexi Kit (Promega) were performed following the manufacturer's instructions using the following conditions: initial denaturation at 95 °C for 2 min, 35 cycles of denaturation at 94 °C for 30 s, annealing at 54°C (*B. g. maretimi)* or 48°C *(B. a. atticus)* for 30 s and extension at 72°C for 2 (*B. g. maretimi)* or 3 *(B. a. atticus)* min, and a final extension at 72°C for 5 min.

PCR products were electrophoresed on 1.5% gels; these were subsequently blotted on positively charged nylon membranes and membranes were hybridized with a PCR amplified R2-specific DNA probe obtained with the primers R2Ins2_F>R2Ins2_R for *B. g. maretimi* and BGGBDIN3>BGGR2ARIN for *B. a. atticus* (Suppl. Table S2).

Resulting signals were detected with the AlkPhos Direct Labelling and Detection System kit (GE Healthcare, Piscataway, NJ, USA).

Each band scored in the hybridization analysis (see for example Suppl. Fig. S1) represents an R2 insertion, thus the number of the band equals the number of rDNA sites occupied by R2 (S). We, then, calculated S in field collected samples (S_p^{tot}); this parameter was calculated separately for females and males in the strictly gonochoric *B. g. maretimi* to test if the two sexes could show significant differences. In *B. a. atticus* offspring analyses, S was calculated per parental females ($S_p♀$) and per offspring, calculating both the total number of R2 occupied sites in the progeny (S_o^{tot}) and the total number of new sites occupied by R2 (i.e. not occurring in the parental female; S_o^{new}) (Bonandin *et al.*, 2014).

The estimated R2 copy number per haploid genome was also calculated as $\tilde{n} = \sum^{S} b_S/N$, where b_S is the number of bands at the S^{th} site occupied by R2 and N is the number of assayed individuals (Charlesworth & Charlesworth, 1983).

Present results on *B. g. maretimi* and *B. a. atticus* field collected specimens have been compared with insertion analyses data on *B. rossius rossius* from Anzio (Rome; bisexual population), *B. rossius redtenbacheri* from Patti (near Messina in northeastern Sicily; bisexual population) and Curcuraci (Messina; parthenogenetic population) derived from Mingazzini (2011), while offspring results have been compared with data on the element activity in the offspring of parthenogenetic *B. r. redtenbacheri* individuals and of crosses involving as parental females either parthenogenetic specimens or *B. r. rossius* gonochoric individuals derived from Bonandin *et al*. (2014).

Results

Insertion analysis in field collected specimens

In the *B. g. maretimi* population, the number of R2 occupied sites in females ranged from six to 11, while in males they ranged from three to nine. A total of 22 sites occupied by R2 (S_p^{tot}) were scored in females, while 19 sites were detected in the sampled males (Table 1). Fifteen insertions were shared between females and males.

The estimated haploid copy number (ñ) in the collected females was 9.14 and in the sampled males 6.50 (Table 1).

Frequency distributions of R2 insertions between males and females of *B. g. maretimi* were not significantly different (two-sample Kolmogorov-Smirnov test: D_{KS}=0.37, P=0.095). As a comparison, the insertion frequency distributions between the sampled females and males of the bisexual *B. rossius* populations were not significantly different for both Anzio and Patti samples (D_{KS}=0.28, P=0.119 and D_{KS}=0.34, P=0.234, respectively).

In the two populations of the obligatory parthenogenetic *B. a. atticus* samples (Necropoli Camarina and Scoglitti) the number of R2 occupied sites in each individual ranged from two to four and from one to three in the two populations, respectively. The total number of sites occupied by R2 was five in each analyzed populations (Table 1), with

three insertions shared between them. The estimated R2 ñ in the population of Necropoli Camarina was higher (2.88) than that scored in the population of Scoglitti (1.64; Table 1). Frequency distributions of R2 insertions between the *B. a. atticus* populations of Necropoli Camarina and Scoglitti were not significantly different (D_{KS}=0.60, P=0.209).

When comparing the two taxa, S_p^{tot} is significantly higher in *B. g. maretimi* with respect to *B. a. atticus* ($t_{\text{one-tailed}}$= 6.01 P= 0.005; Suppl. Fig. 1). The same holds for the average R2 copy number per haploid genome (ñ; $t_{\text{one-tailed}}$= 2.90, P= 0.031). On the other hand, the R2 insertion frequency distributions were not significantly different (D_{KS}=0.28, P=0.680) and showed a higher number of low frequency insertions (Fig. 1A).

Regarding the facultative parthenogenetic *B. rossius*, the insertion frequency distribution in the two *B. r. redtenbacheri* populations (the bisexual Patti and the parthenogenetic Curcuraci) was not significantly different (D_{KS}=0.30, P=0.406; Fig. 1B). The same applies in the comparison between the two *B. r. redtenbacheri* populations (Patti and Curcuraci) and the *B. r. rossius* sample from Anzio (D_{KS}=0.35, P=0.067 and D_{KS}=0.26, P=0.439, respectively; Fig. 1B). As in *B. a. atticus* and in *B. g. maretimi*, low frequency insertion outnumber the high frequency ones in all samples, with the exception of the *B. r. redtenbacheri* bisexual population from Patti (Fig. 1B).

R2 insertion frequency spectra were also compared for all available samples (data not shown). All comparisons were not significantly different, the only exception being given by the one between *B. g. maretimi* and *B. r. redtenbacheri* Patti (D_{KS}=0.49, P=0.006).

R2 insertions display in parental B. a. atticus *females and their offspring*

In the NC ♀2 female, three sites were occupied by R2 (S_p♀); in its offspring a maternal truncated variant was completely deleted, but two new insertions were evident (S_o^{new}), leading to a total of four sites occupied by R2 (S_o^{tot}) and ñ = 2.3 (Table 2). The two parental females SCO ♀1 and SCO ♀2 showed two and three sites occupied by R2,

respectively. In their progeny three and 11 S_o^{tot} were detected with one and eight S_o^{new} scored. The ñ was equal to 2.1 and 2.8, respectively (Table 2). At variance of NC ♀2 and SCO ♀1, the progeny of SCO ♀2 showed a lot more S_o^{new}, clearly, all having very low frequency. The comparison between S_o^{tot} and S_p^{tot} did not give significant difference ($t_{one\text{-}tailed}$= 0.301, P= 0.389) and the same holds for the average R2 copy number per haploid genome ($t_{one\text{-}tailed}$= 0.256, P= 0.405). Owing to the very low number of insertions observed, the study of insertions frequency distributions for each single *B. a. atticus* female's progeny risks to be meaningless: it will, therefore, not further considered.

In order to check for possible differences between obligatory and facultative parthenogenesis, present data on *B. a. atticus* offspring were compared with the unisexual offspring of *B. rossius* analyzed in Bonandin *et al.* (2014). Obligatory unisexual offspring has significantly lower values for S_o^{tot} and ñ ($t_{one\text{-}tailed}$= 2.21, P= 0.029 and $t_{one\text{-}tailed}$= 3.40, P= 0.0046, respectively) and showed a decidedly different insertion frequencies distribution than that of the facultative unisexual progeny (D_{KS}= 0.49, P=0.0009; Fig. 2)

Discussion

Theoretical studies link the reproductive strategy adopted by the host with the transposable element load (Wright & Finnegan, 2001; Nuzhdin & Petrov, 2003). In bisexually reproducing organisms TEs can disperse (Hickey, 1982) but, according to the Muller's ratchet theory, events of recombination and independent assortment of chromosomes limit their deleterious proliferation, as natural selection has more opportunity to operate (Arkhipova, 2005; Dolgin & Charlesworth, 2006; Loewe & Lamatsch, 2008). In asexual, unisexual and low recombining organisms, TEs are expected to accumulate even up to the host lineage extinction (Loewe & Lamatsch, 2008).

To test the impact of reproductive strategies on R2 accumulation in host genomes, we analyzed its activity in individuals from two populations of the obligatory

parthenogenetic *B. a. atticus* subspecies and from the population of the strictly gonochoric *B. g. maretimi* subspecies. These data add and complete previous analyses on the facultative parthenogenetic *B. rossius* (Bonandin *et al.*, 2014)

The strictly gonochoric *B. g. maretimi*, despite the lower number of processed samples, showed at population level significantly more sites occupied by R2 (S_p^{tot}) than the obligatory parthenogenetic *B. a. atticus.* This is also reflected by the estimated R2 copy number per haploid genome of *B. g. maretimi*, which is more than three times that scored in *B. a. atticus*. Values for bisexual and unisexual populations of *B. rossius* were higher than those recorded for *B. atticus*, being similar or even higher than *B. g. maretimi* ones. Thus, it appears as if obligatory parthenogenesis in *B. atticus* allows the maintenance of a very low number of insertions. Consistently, insertion frequency spectra show, in *B. a. atticus*, a prevalence of low frequency insertion as in *B. g. maretimi* and bisexual *B. rossius*.

On the whole, data produced here indicate that, in the obligatory unisexual *B. a. atticus*, R2 is active and that mechanisms of molecular turnover are effective; on the other hand, the R2 transposition rate is lower than in the facultative parthenogenetic *B. rossius*.

A comparable situation is observed for class II elements assayed in *Daphnia pulex* (including two families of Tc1-like elements, two Helitron families, a hAt homologue and the piggyBac element *Pokey*; Valizadeh & Crease, 2008 and Schaack *et al.*, 2010). *D. pulex* is a cyclical parthenogenetic aquatic microcrustacean, with most populations reproducing clonally during the growing season and, as a consequence of environmental stresses, could shift to sex. However, some populations seem to have lost this ability, thus becoming obligatory unisexuals (Schaack *et al.*, 2010). As for *B. atticus* with respect to the facultative parthenogenetic *B. rossius*, obligate *D. pulex* unisexuals had lower TE loads than cyclical parthenogenetic, but they have comparable distribution of insertion frequency spectra (Valizadeh & Crease, 2008; Schaack *et al.*, 2010).

Data observed in *D. pulex* have been explained as the likely result of clonal selection: obligate unisexual lineages originated multiple times and it is possible that only those with lower TE load survived and are presently observable (Valizadeh & Crease, 2008). Therefore, those insertions embodied in the unisexuals' genome occur at low frequency. At variance of facultative parthenogenetic, where the possible switch from bisexuality to unisexuality and vice versa could maintain a higher number of insertions, in obligate unisexuals the reduction of R2 copy number is a required process, since the recovery from unisexuality to bisexuality is not possible. Mechanisms able to keep R2 load low may explain the survival of an obligate parthenogenetic as *B. a. atticus*, notwithstanding Muller's ratchet predictions. In the light of this theory, in fact, unisexuals should either be free of deleterious mutations or reach extinction owing to their supposed inability to get rid of them.

Modelling the parental offspring inheritance as in Bonandin *et al*. (2014), a major role in eliminating insertions appears to be hold by recombination rather than selection even with *B. atticus* parthenogenetic mechanism. However, both recombination and selection produce a noticeably high variance (Fig. 3): this, on the long term, would allow natural selection to more easily counteract an increase of R2 load.

The extent of the R2 load reduction in *B. atticus*, as well as the great variance observed in the assayed offspring isolates, is striking. In particular, the transposition burst observed in the SCO ♀2 offspring may support also the possibility of clonal selection. This could be verified through the analyses of the reproductive output of *B. atticus* isolates with increased R2 load. Due to the relatively long generation time of these insects (8-10 months), this will be reported in a future study.

The very limited number of R2 insertions in *B. a. atticus* may also suggest that not only recombination would contribute to hold R2 at bay. It could be likely that some silencing mechanism is acting on those few R2 copies to further reduce the

retrotransposition activity. Eickbush *et al*. (2008) suggested a model in which R2 can be silenced by heterochromatinization of R2 densely populated rDNA regions, followed by nucleolar dominance. In this view, silencing few copies is even more simple, leading to an overall very low activity. It cannot be excluded that in obligatory unisexual genomes less virulent TEs have been selected, in order to guarantee the survival of both host and TE (Bull *et al*., 1991; Wright & Finnegan 2001).

Data presented here highlighted very interesting differences not only between bisexual and unisexual reproductive strategies but also between the very mechanisms with which unisexuality is achieved. Together with *B. rossius* data, present analyses clearly indicate that bisexual and unisexual taxa behave differently, bisexual reproduction ensuring a continuous TE turnover (insertions/deletions). This is in line with theoretical studies that predict a higher representativeness of TEs (Hickey, 1982) but with a high elimination rate (Muller's ratchet). In unisexual taxa, as a general trend, the frequency of insertions are low at the population level, irrespective of the mechanism of parthenogenesis and the standing copy number. As every female is part of a vertical line whose members do not cross with others in the population, every insertion within the line is not diffused in the population with the net effect that their frequency is always relatively low. However, within each line (here represented by unisexual isolates) facultative unisexuals as *B. rossius* seem more prone to R2 accumulation than obligatory unisexuals as *B. a. atticus*. Parthenogenesis in *B. rossius* in fact leads to homozygous offspring while in *B. atticus* the genetic asset of the mother is maintained, that is if the mother is heterozygous then its offspring will be heterozygous (of course, excluding potential recombinants; Fig. 3A). If it is true that recombination plays a major role in eliminating the insertions in both mechanisms of parthenogenesis (Bonandin *et al*., 2014 - Fig. 6; present study - Fig. 3B), in *B. atticus* the variance produced (with or without selection on inserted

rDNAs) is higher than that predicted in *B. rossius*. This means that natural selection would have more opportunity to operate to keep the unisexual genome purged by TE load.

Acknowledgments

This work was supported by RFO and Canziani funding to B.M. and A.L. We wish to thank two anonymous reviewers whose comments and suggestions helped to improve the paper.

References

Arkhipova, I.R. (2005) Mobile genetic elements and sexual reproduction. *Cytogenetic and Genome Research*, 110, 372–382.

Bonandin, L., Scavariello, C., Luchetti, A. and Mantovani, B. (2014) Evolutionary dynamics of R2 retroelement and insertion inheritance in the genome of bisexual and parthenogenetic *Bacillus rossius* populations (Insecta, Phasmida). *Insect Molecular Biology*, 23(6), 808-820.

Bull, J.J., Molineux, I.J. and Rice, W.R. (1991) Selection of benevolence in a host-parasite system. *Evolution*, 45(4), 875–882.

Charlesworth, B. and Charlesworth, D. (1983) The population dynamics of transposable elements. *Genetic Research*, 42, 1–27.

Dolgin, E.S. and Charlesworth, B. (2006) The fate of transposable elements in asexual populations. *Genetics,* 174, 817–827.

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

Eickbush, D.G., Ye, J., Zhang, X., Burke, W.D. and Eickbush, T.H. (2008) Epigenetic regulation of retrotransposon within the nucleolus of *Drosophila*. *Molecular and Cellular Biology* 28, 6452–6461.

Eickbush, T.H. (2002) R2 and related site-specific non-LTR retrotransposons. *Mobile DNA II* (eds. N. Craig, R. Craigie, M. Gellert & A. Lambowitz), pp. 813–835. American Society of Microbiology Press, Washington, DC.

Fujimoto, H., Hirukawa, Y., Tani, H., Matsuura, Y., Hashido, K., Tsuchida, K., Takada, N., Kobayashi, M. and Maekawa, H. (2004) Integration of the 5’ end of the retrotransposon, R2Bm, can be complemented by homologous recombination. *Nucleic Acids Research* 32:1555–1565

Gladyshev, E.A. and Arkhipova, I.R. (2009) Rotifer rDNA-specific R9 retrotransposable elements generate an exceptionally long target site duplication upon insertion. *Gene,* 448(2), 145-150.

Hickey, D.A. (1982) Selfish DNA: a sexually-transmitted nuclear parasite. *Genetics* 101, 519–531.

Kojima, K.K. and Fujiwara, H. (2005) Long-term inheritance of the 28S rDNA-specific retrotransposon R2. *Molecular Biology and Evolution,* 22, 2157–2165

Kraaijeveld, K., Zwanenburg, B., Hubert, B., Vieira, C., de Pater, S., van Alphen, J.J.M., Den Dunnen J.T. and De Knijff, P. (2012) Transposon proliferation in an asexual parasitoid. *Molecular Ecology,* 21 (16), 3898–3906.

Kojima, K.K, Kuma, K.,Toh, H. and Fujiwara, H. (2006) Identification of rDNA-Specific Non-LTR Retrotransposons in Cnidaria *Molecular Biology and Evolution,*23(10),1984–1993.

Loewe, L. and Lamatsch, D.K. (2008) Quantifying the threat of extinction from Muller's ratchet in the diploid Amazon molly *(Poecilia formosa*). *BMC Evolutionary Biology*, 19, 88.

Luchetti, A. and Mantovani, B. (2013) Non-LTR R2 element evolutionary patterns: phylogenetic incongruences, rapid radiation and the maintenance of multiple lineages. *PLoS ONE,* 8: e57076.

Mantovani, B., Passamonti, M. and Scali, V. (2001) The mitochondrial cytochrome oxidase II gene in *Bacillus* stick insects: ancestry of hybrids, androgenesis, and phylogenetic relationships. *Molecular Phylogenetics and Evolution,* 19, 157–163.

Martoni, F., Eickbush, D.G., Scavariello, C., Luchetti, A. and Mantovani, B. (2015) Dead element replicating: Degenerate R2 element replication and rDNA genomic turnover in the *Bacillus rossius* stick insect (Insecta: Phasmida). *PLoS ONE,* 10, e0121831.

Mingazzini, V. (2011) Transposable elements: structure and dynamic of the autonomous retrotransposon R2 in Arthropoda with non-canonical reproduction. PhD thesis, University of Bologna.

Mukha, D.V., Mysina, V., Mavropulo, V. and Schal, C. (2011) Structure and molecular evolution of the ribosomal DNA external transcribed spacer in the cockroach genus *Blattella*. *Genome* 54:222-234.

Mukha, D.V., Pasyukova, E.G., Kapelinskaya, T.V. and Kagramanova, A.S. (2013) Endonuclease domain of the *Drosophila melanogaster* R2 non-LTR retrotransposon and related retroelements: a new model for transposition. *Frontiers in Genetics* 4:63.

Nei, M. and Rooney, A.P. (2005) Concerted and birth-and-death evolution of multigene families. Annual Review of Genetics 39:121–152.

Nuzhdin, S.V. and Petrov, D.A. (2003) Transposable elements in clonal lineages: lethal hangover from sex. *Biological Journal of the Linnean Society*, 79, 33–41.

Pérez-Gonzalez, C.E. and Eickbush, T.H. (2001) Dynamics of R1 and R2 elements in the rDNA locus of *Drosophila simulans*. *Genetics,* 158, 1557–1567.

Plohl, M., Luchetti, A., Meštrović, N. and Mantovani, B. (2008) Satellite DNAs between selfishness and functionality: structure, genomics and evolution of tandem repeats in centromeric (hetero)chromatin. *Gene,* 409, 72–82.

Ricci, M., Luchetti, A., Bonandin, L. and Mantovani, B. (2013) Random DNA libraries from three species of the stick-insect genus *Bacillus* (Insecta: Phasmida): repetitive DNA characterization and first observation of polyneopteran MITEs. Genome, 56(12), 729-735.

Scali, V., Passamonti, M., Marescalchi, O. and Mantovani, B. (2003) Linkage between sexual and asexual lineages: genome evolution in *Bacillus* stick insects. *Biological Journal of the Linnean Society,* 79, 137–150.

Schaack, S., Pritham, E.J., Wolf, A. and Lynch, M. (2010) DNA transposon dynamics in populations of *Daphnia pulex* with and without sex. *Proceedings Royal Society B*, 277, 2381–2387.

Schultz, R. J. (1961) Reproductive mechanism of unisexual and bisexual strains of the viviparous fish *Poeciliopsis*. *Evolution,* 15, 302-325.

Valizadeh, P. and Crease, T.J. (2008) The association between breeding system and transposable element dynamics in *Daphnia pulex*. *Journal of Molecular Evolution*, 66, 643–654.

Wright, S. and Finnegan, D. (2001) Genome evolution: sex and transposable elements. *Current Biology*, 11, R296–R299.

Figure legends

Figure 1. R2 insertion frequency spectra relative to the presently analyzed taxa/populations (A) together with *B. rossius* samples from Mingazzini *et al*. (2011; B). Each bin represents a class of insertion frequency. On the *y*-axis the number of R2 insertions recorded for each class is given.

Figure 2. R2 insertion frequency spectra of the obligatory parthenogenetic *B. a. atticus* offspring in comparison with the facultative unisexual offspring of *B. r. redtenbacheri* analyzed in Bonandin *et al.* (2014). Each bin represents a class of insertion frequency. On the *y*-axis the number of R2 insertions recorded for each class is given.

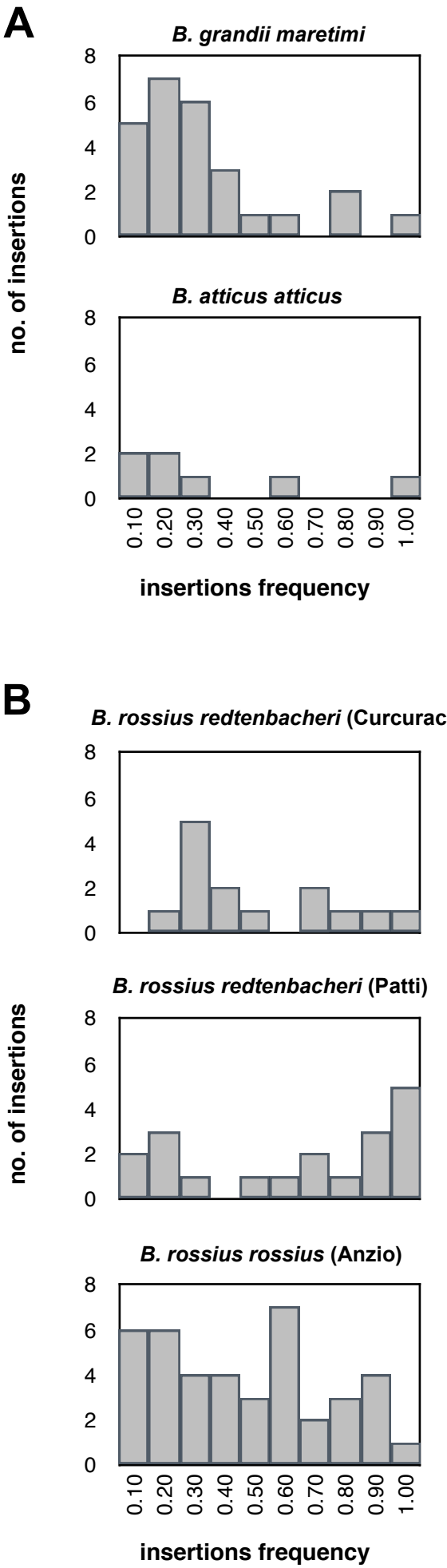
Figure 3. A) Scheme of *B. atticus* parthenogenetic mechanisms (modified from Scali *et al*., 2003). B) Scenario of R2 insertions inheritance under *B. atticus* unisexuality. Each box represents a 28S rRNA and dots are R2 insertions (different colours = different insertions); r_p = recombinant progeny. For three possible situations (progeny without recombinant; progeny with recombinant; progeny after selection) R2 insertion frequencies, copy number (ñ) and its variance (var) have been calculated. In this model, selection has been considered as acting on offspring with longer rDNA array.

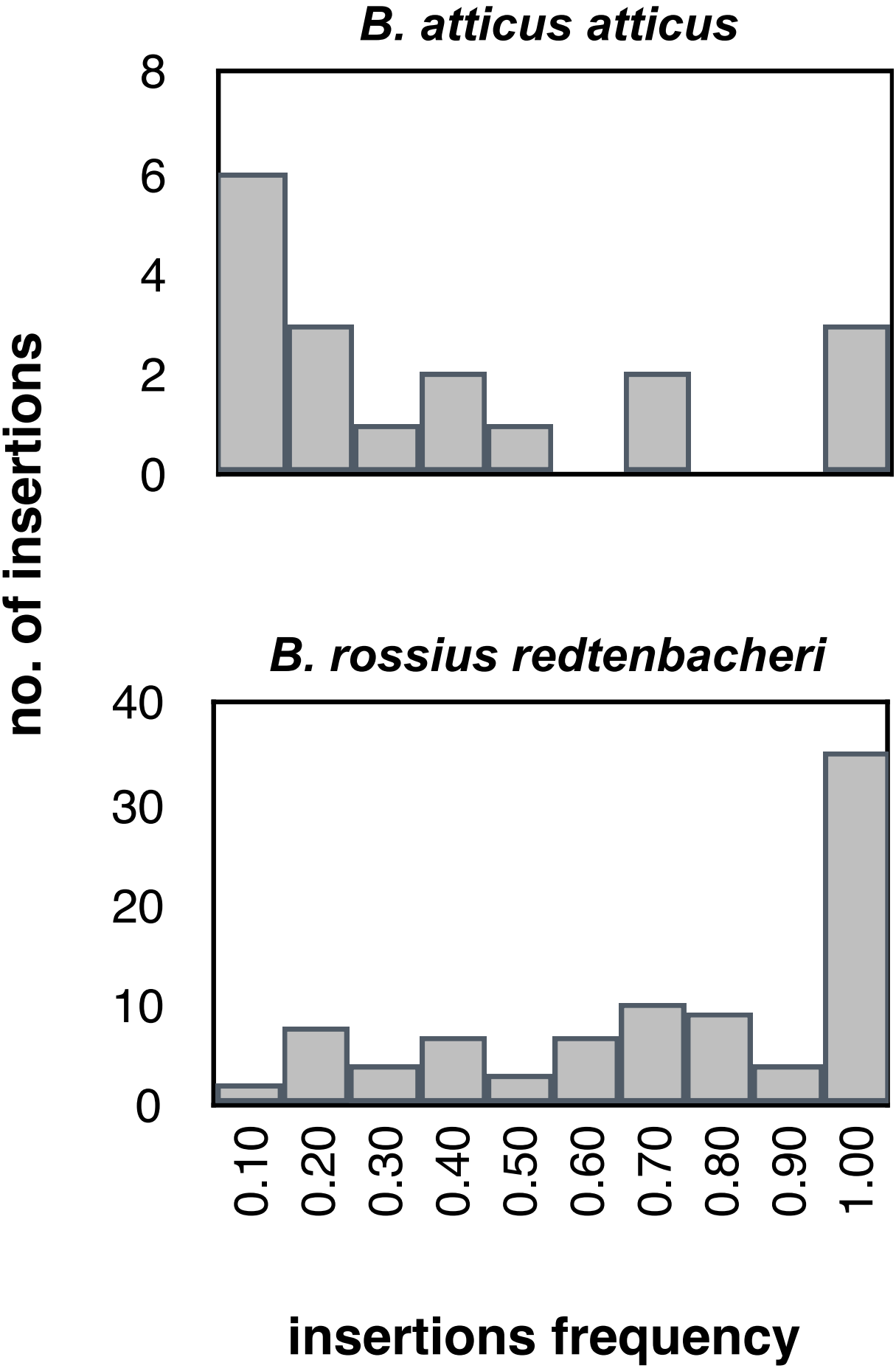
Table 1. Parameters calculated on the basis of the R2 insertion display in the analyzed taxa at the population level (N: number of individuals analyzed; S_p♀/♂: number of sites occupied by R2 detected in females/males; S_p^{tot}: total number of sites occupied by R2 in the population; ñ: estimated R2 copy number per haploid genome; *data from Mingazzini, 2011).

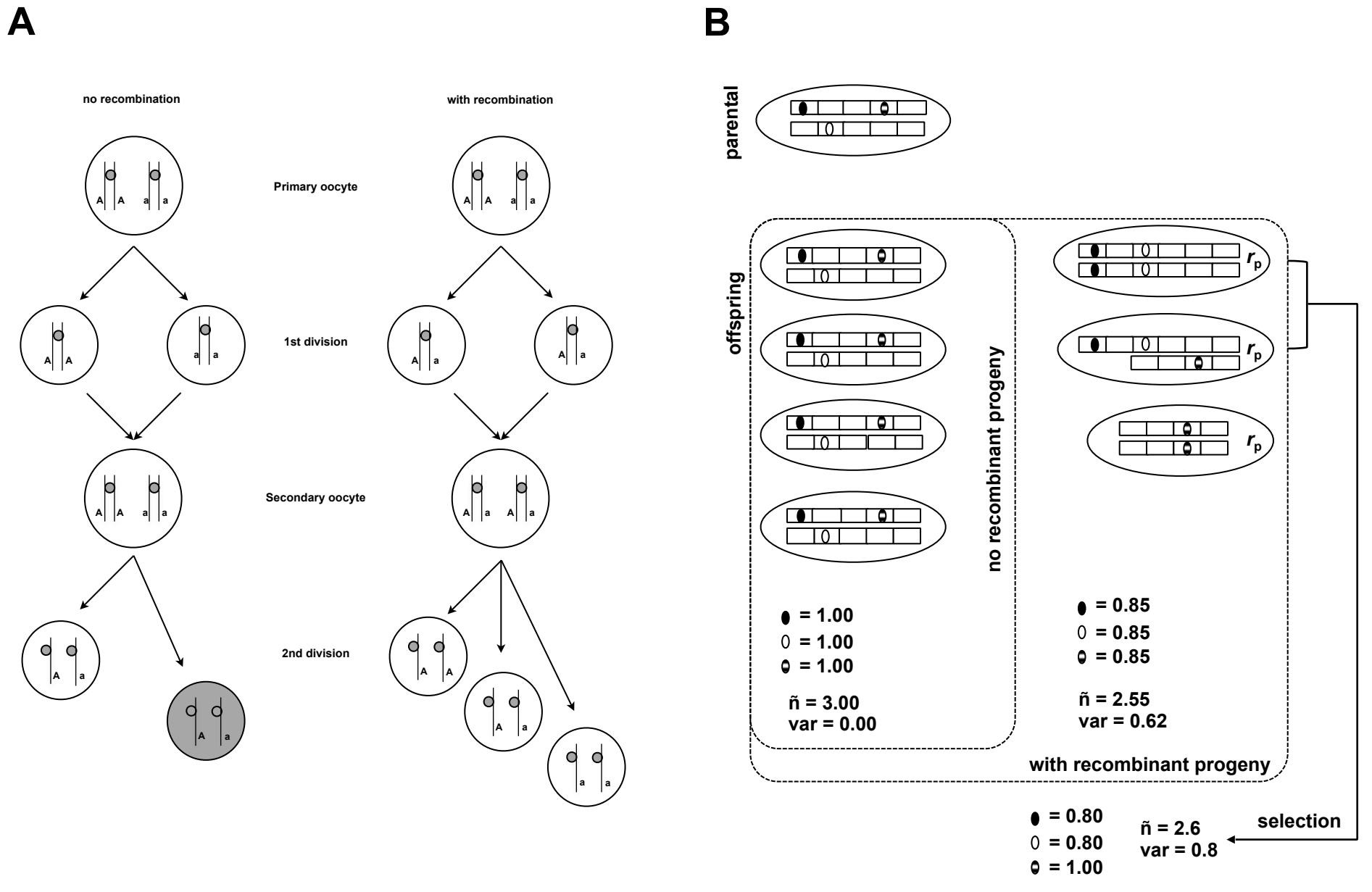
Taxon	**Population**	**N**	**S_p♀**	**S_p♂**	**S_p^{tot}**	**ñ♀**	**ñ♂**	**ñ^{tot}**
B. g. maretimi	Marettimo	7♀♀/8♂♂	22	19	26	9.14	6.5	7.73
B. atticus	Necropoli Camarina	8♀♀	5	-	5	2.88	-	2.88
B. atticus	Scoglitti	11♀♀	5	-	5	1.64	-	1.64
*B. r. rossius**	Anzio	10♀♀/11♂♂	35	33	40	15.50	18	16.81
*B. r. redtenbacheri**	Patti	10♀♀/8♂♂	17	16	19	12.40	9.75	11.22
*B. r. redtenbacheri**	Curcuraci	19♀♀	14	-	14	6.58	-	6.58

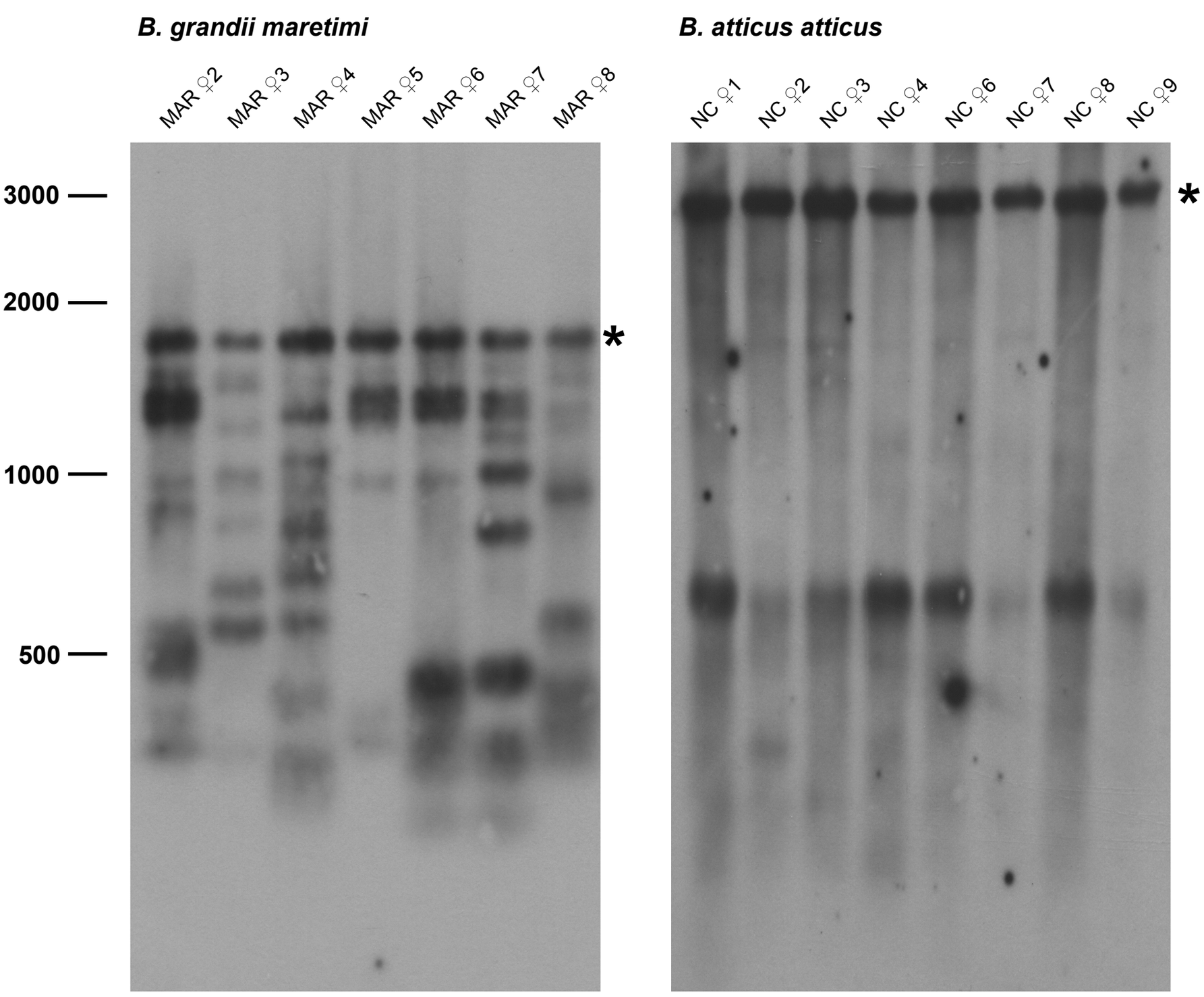
Table 2. Parameters calculated on the basis of the R2 insertion display in three *B. a. atticus* females and in 10 individuals each of their offspring (S_p♀: number of sites occupied by R2 detected in parental females; S_o^{new}: number of new sites occupied by R2 in the offspring; S_o^{tot}: the total number of sites occupied by R2 in the progeny; ñ: estimated R2 copy number per haploid genome).

Parental Individual	**S_p♀**	**S_o^{new}**	**S_o^{tot}**	**ñ**
NC ♀2	3	2	4	2.3
SCO ♀1	2	1	3	2.1
SCO ♀2	3	8	11	2.8


A
B. grandii maretimi
B. atticus atticus
no. of insertions
0
2
4
6
8
0.10
0.20
0.30
0.40
0.50
0.60
0.70
0.80
0.90
1.00
insertions frequency
B
B. rossius redtenbacheri (Curcuraci)
B. rossius redtenbacheri (Patti)
B. rossius rossius (Anzio)
no. of insertions
insertions frequency


B. atticus atticus
8
6
4
2
0
B. rossius redtenbacheri
40
30
20
10
0
0.10
0.20
0.30
0.40
0.50
0.60
0.70
0.80
0.90
1.00
no. of insertions
insertions frequency


A
no recombination
with recombination
Primary oocyte
1st division
Secondary oocyte
2nd division
B
parental
offspring
no recombinant progeny
with recombinant progeny
r_p
● = 1.00
○ = 1.00
⊖ = 1.00
ñ = 3.00
var = 0.00
● = 0.85
○ = 0.85
⊖ = 0.85
ñ = 2.55
var = 0.62
● = 0.80
○ = 0.80
⊖ = 1.00
ñ = 2.6
var = 0.8
selection



Suppl. Figure S1. Southern blot analysis for R2 insertion display carried out on *B. g. maretimi* Marettimo and *B. a. atticus* Necropoli Camarina populations. For *B. g. maretimi* only females are shown for a comparison with the all-female *B. a. atticus*. On the left, the molecular weight is given in base pairs; above each lane, the acronym of the assayed individual is reported as in Suppl. Table S1. Asterisks indicate the full length element.

Suppl. Table S1. Taxon, Sicilian sampling localities and acronyms of field collected and offspring specimens analyzed.

Taxon	Sampling locality	Field collected specimens	Offspring	Acronym
Bacillus atticus atticus				
	Necropoli Camarina			
		♀ 1		NC ♀1
		♀ 2		NC ♀2
			♀ 2-3	*Baa*NC♀2-F3
			♀ 2-4	*Baa*NC♀2-F4
			♀ 2-5	*Baa*NC♀2-F5
			♀ 2-6	*Baa*NC♀2-F6
			♀ 2-7	*Baa*NC♀2-F7
			♀ 2-28	*Baa*NC♀2-F28
			♀ 2-29	*Baa*NC♀2-F29
			♀ 2-30	*Baa*NC♀2-F30
			♀ 2-31	*Baa*NC♀2-F31
			♀ 2-32	*Baa*NC♀2-F32
		♀ 3		NC ♀3
		♀ 4		NC ♀4
		♀ 6		NC ♀6
		♀ 7		NC ♀7
		♀ 8		NC ♀8
		♀ 9		NC ♀9
	Scoglitti			
		♀ 1		SCO ♀1
			♀ 1-30	*Baa*SCO♀1-F30
			♀ 1-31	*Baa*SCO♀1-F31
			♀ 1-32	*Baa*SCO♀1-F32
			♀ 1-33	*Baa*SCO♀1-F33
			♀ 1-34	*Baa*SCO♀1-F34
			♀ 1-51	*Baa*SCO♀1-F51
			♀ 1-52	*Baa*SCO♀1-F52
			♀ 1-53	*Baa*SCO♀1-F53
			♀ 1-54	*Baa*SCO♀1-F54
			♀ 1-55	*Baa*SCO♀1-F55

		♀ 2	SCO ♀2
		♀ 2-1	*Baa*SCO♀2-F1
		♀ 2-2	*Baa*SCO♀2-F2
		♀ 2-3	*Baa*SCO♀2-F3
		♀ 2-4	*Baa*SCO♀2-F4
		♀ 2-5	*Baa*SCO♀2-F5
		♀ 2-31	*Baa*SCO♀2-F31
		♀ 2-32	*Baa*SCO♀2-F32
		♀ 2-33	*Baa*SCO♀2-F33
		♀ 2-34	*Baa*SCO♀2-F34
		♀ 2-35	*Baa*SCO♀2-F35
		♀ 3	SCO ♀3
		♀ 4	SCO ♀4
		♀ 9	SCO ♀9
		♀ 12	SCO ♀12
		♀ 13	SCO ♀13
		♀ 15	SCO ♀15
		♀ 19	SCO ♀19
		♀ 20	SCO ♀20
		♀ 23	SCO ♀23
Bacillus grandii maretimi			
	Marettimo		
		♀ 2	MAR ♀2
		♀ 3	MAR ♀3
		♀ 4	MAR ♀4
		♀ 5	MAR ♀5
		♀ 6	MAR ♀6
		♀ 7	MAR ♀7
		♀ 8	MAR ♀8
		♂ 4	MAR ♂4
		♂ 5	MAR ♂5
		♂ 6	MAR ♂6
		♂ 7	MAR ♂7
		♂ 8	MAR ♂8
		♂ 9	MAR ♂9

	♂ 10	MAR ♂10
	♂ 11	MAR ♂11

Suppl. Table S2. Primers used for R2 PCR amplifications for this study (F, forward primer; R, reverse primer)

Primer name	Sequence 5' -> 3'	Reference	Taxon	F/R
28S-F2	GTCAAAGTGAAGAAATTCAACGAAG	Mingazzini et al. (2011)	*Bgm*	F
R2Ins2_F	GGACAAGCGCACAGAGTCAG	this study	*Bgm*	F
R2Ins2_R	GGACATCCTGTTGGCGATTAC	this study	*Bgm*	R
28SFBgg	GAATCCGACTGTCTAATTAAAACAAAG	this study	*Baa*	F
BggBDIN3	TATGATATTCAACATGGTTATTG	Mingazzini (PhD thesis, 2011)	*Baa*	F
BggR2ARIN	CGATCTAGTAGGTACTTCACACATGCA	this study	*Baa*	R

2.3 Repetitive DNA content in *Bacillus*

Ricci, M., Luchetti, A., Bonandin, L. and Mantovani, B. (2013) Random DNA libraries from three species of the stick-insect genus *Bacillus* (Insecta: Phasmida): repetitive DNA characterization and first observation of polyneopteran MITEs. Genome 56(12): 729-735.

Given my interest in the relationship between TEs and reproductive biology, during my thesis I have also been involved in the research project of another Ph.D. student (see Ricci, 2014; Ph.D. thesis), whose results are presented in this paper. My contribution to this work was to verify the presence of some TE families in the genomes of *B. grandii* (strictly gonochoric), *B. rossius* (facultative parthenogenetic) and *B. atticus* (obligatory parthenogenetic) through Southern blot experiments. The data that I obtained completed the results on the libraries of the three considered species. On the whole, the genomic content of TEs was of 18%, 23.3% and 22.9% in *B. grandii*, *B. rossius* and *B. atticus*, respectively, while the tandem repeat abundance represented the 0.64% of the genome in *B. grandii*, the 1.32% in *B. rossius* and the 1.86% in *B. atticus*. Moreover, it is worth noting that in all three genomes a miniature inverted repeat (MITE) family has been observed for the first time in polyneopteran insects. These results highlighted that the occurrence of TEs revealed in the obligatory unisexual *B. atticus* was not significantly lower than that scored in *B. grandii* and *B. rossius*, counteracting Muller's ratchet expectations.

NRC Research Press

ARTICLE

Random DNA libraries from three species of the stick insect genus *Bacillus* (Insecta: Phasmida): repetitive DNA characterization and first observation of polyneopteran MITEs

Marco Ricci, Andrea Luchetti, Livia Bonandin, and Barbara Mantovani

Abstract: The repetitive DNA content of the stick insect species *Bacillus rossius* (facultative parthenogenetic), *Bacillus grandii* (gonochoric), and *Bacillus atticus* (obligate parthenogenetic) was analyzed through the survey of random genomic libraries roughly corresponding to 0.006% of the genome. By repeat masking, 19 families of transposable elements were identified (two LTR and six non-LTR retrotransposons; 11 DNA transposons). Moreover, a de novo analysis revealed, among the three libraries, the first MITE family observed in polyneopteran genomes. On the whole, transposable element abundance represented 23.3% of the genome in *B. rossius*, 22.9% in *B. atticus*, and 18% in *B. grandii*. Tandem repeat content in the three libraries is much lower: 1.32%, 0.64%, and 1.86% in *B. rossius*, *B. grandii*, and *B. atticus*, respectively. Microsatellites are the most abundant in all species. Minisatellites were only found in *B. rossius* and *B. atticus*, and five monomers belonging to the *Bag320* satellite family were detected in *B. atticus*. Assuming the survey provides adequate representation of the relative genome, the obligate parthenogenetic species (*B. atticus*), compared with the other two species analyzed, does not show a lower transposable element content, as expected from some theoretical and empirical studies.

Key words: genomic sequence survey, miniature inverted repeats (MITEs), stick insects, tandem repeats, transposable elements.

Résumé : Les auteurs ont analysé le contenu en ADN répété chez les phasmes *Bacillus rossius* (à parthénogenèse facultative), *Bacillus grandii* (gonochorique) et *Bacillus atticus* (à parthénogenèse obligatoire) en explorant aléatoirement des banques génomiques correspondant à environ 0,006 % du génome. En masquant les séquences répétées, 19 familles d'éléments transposables ont été identifiées (deux familles de rétrotransposons à LTR et six familles sans LTR; 11 familles de transposons à ADN). De plus, une analyse de novo a permis de découvrir, au sein des trois banques, la première famille d'éléments MITE au sein des polynéoptères. Globalement, l'abondance des éléments transposables représentait 23,3 % du génome chez le *B. rossius*, 22,9 % chez le *B. atticus* et 18 % chez le *B. grandii*. Les séquences répétées en tandem étaient beaucoup moins abondantes au sein des trois banques : 1,32 %, 0,64 % et 1,86 % respectivement chez le *B. rossius*, le *B. grandii* et le *B. atticus*. Les microsatellites étaient les plus abondants chez les trois espèces. Des minisatellites n'ont été observés que chez le *B. rossius* et le *B. atticus*, tandis que cinq monomères de la famille de satellites *Bag320* ont été détectés de manière unique chez le *B. atticus*. En supposant que cet échantillonnage du génome soit suffisamment représentatif de chaque génome, il s'avère que le génome de l'espèce à parthénogenèse obligatoire (*B. atticus*) ne présente pas un contenu en éléments transposables inférieur à celui des deux autres espèces, tel que le prédisaient certaines études théoriques et empiriques. [Traduit par la Rédaction]

Mots-clés : relevé de séquences génomiques, séquences répétées inversées miniatures (MITE), phasmes, répétitions en tandem, éléments transposables.

Introduction

A significant fraction of eukaryotic genomes harbours DNA sequences repeated either in tandem (head-to-tail arranged) or interspersed (Richard et al. 2008).

Tandem repeats are made by monomeric units, whose length is comprised between two and hundreds of base pairs (bp), organized in arrays ranging from few to thousands of units. They can be categorized into the following three main classes: microsatellites, minisatellites, and satellite DNAs (satDNAs). Although the three classes cannot be discriminated on the sole basis of unit length (Charlesworth et al. 1994; Richard et al. 2008), their approximate monomer length ranges can be considered as 2–10, 11–100, and >100 bp, respectively.

Interspersed repeats are, mainly, transposable elements (TEs), i.e., sequences able to move from one genomic location to another (Makałowski et al. 2012). There are two main classes of TEs: class I elements, moving via an RNA intermediate (retrotransposons), and class II elements, moving via a DNA intermediate (transposons). Within the two TE classes, autonomous elements, which are able to encode the proteins necessary for their transposition, and nonautonomous elements, which parasitize the transposition machinery of an autonomous partner, can be further distinguished (Makałowski et al. 2012). Among the latter, short interspersed elements (SINEs) are the most diverse and represented retrotransposons; their sequence is composed of (*i*) an RNA-related head, (*ii*) an anonymous body, and (*iii*) a simple sequence repeat tail. To date, the absence of SINEs has been reported only in species of *Drosophila* (Kramerov and Vassetzky 2011). Miniature inverted repeats (MITEs) are nonautonomous DNA transposons; they usually derive from autonomous elements through deletion of the internal protein-coding sequence and are characterized

Received 25 June 2013. Accepted 29 October 2013.

Corresponding Editor: T.E. Bureau.

M. Ricci, A. Luchetti, L. Bonandin, and B. Mantovani. Dipartimento di Scienze Biologiche, Geologiche e Ambientali (BiGeA), Università di Bologna, via Selmi 3, 40126 Bologna, Italy.

Corresponding author: Marco Ricci (e-mail: marco.ricci19@unibo.it).

Published at www.nrcresearchpress.com/gen on 7 November 2013.

by terminal or subterminal inverted repeats. MITEs are mainly found in plants, but they are also well represented in metazoan genomes (Feschotte et al. 2002).

Far from being just "junk DNA", repetitive DNA sequences are known to be biologically relevant for the host genome. SatDNAs, for example, are involved in structural functions, being the major component of centromeres in almost all eukaryotes (Plohl et al. 2008). On the other hand, TE genomic dynamics impact the host in different ways. For example, in many instances, TE insertions have been found to modify gene structures, gene expression profiles, and to promote recombination, thus introducing genetic diversity and even adaptive changes (Kazazian 2004). Besides positive interactions, repeated DNAs may also impact negatively on the host genome: for example, contraction or expansion of mini- and microsatellite arrays are directly involved in cancer development and other human pathologies (Galindo et al. 2011; El-Murr et al. 2012), while TE insertions can be deleterious for their involvement in gene disruption, negative alteration of gene expression, and ectopic recombination (Kazazian 2004).

The very existence of repeated DNAs has been long considered paradoxical because their accumulation should not be easily tolerated by the host genome (Charlesworth et al. 1994). In principle, TE accumulation or large expansions of tandem repeat arrays are contrasted by recombination that helps in eliminating deleterious alleles, for example, a TE insertion or a too large tandem array. Organisms with high genetic diversity and bisexual reproduction, therefore, would eliminate more efficiently an overload of repeated sequences. This relationship has been well depicted in the evolutionary hypothesis known as Muller's ratchet (Brookfield and Badge 1997; Wright and Schoen 1999): nonrecombining genomes would accumulate deleterious mutations that will drive them to extinction. However, the observation of unisexual and asexual taxa persisting over evolutionary time requires that mechanisms exist that are able to avoid the deleterious mutation load: the absence of TEs (or a very low TE activity) can be one such mechanism (Arkhipova and Meselson 2000, 2005; Sullender and Crease 2001). It is worth noting, though, that exceptions occur in that unisexual taxa show the same TE load than bisexual taxa (Kraaijeveld et al. 2012). Moreover, generally speaking, less virulent or even favourable parasites (here including TEs) can be selected in an unisexual or asexual lineage as a result of the strictly vertical transmission under these reproductive circumstances: this would allow both host and parasite to survive (Fine 1975; Bull et al. 1991; Wright and Finnegan 2001).

The *Bacillus* stick insect species complex is restricted to the Mediterranean area and shows a number of different reproductive biology issues. The genus comprises three so-called parental species: *Bacillus rossius*, with bisexual and parthenogenetic populations; the strictly bisexual *Bacillus grandii*; and the obligate parthenogenetic *Bacillus atticus*. Interspecific hybridization between or among the parental species produced both unisexual taxa and hybridogenetic lineages (Scali et al. 2003).

To go through the evolution of repeated DNAs, we undertook a genomic sequence survey by randomly cloning genomic fragments of the three parental species *B. rossius*, *B. grandii*, and *B. atticus*, obtaining low coverage DNA libraries. Low coverage sequencing, albeit giving partial genomic information, may provide a quick snapshot of the genome content, especially regarding repetitive DNA. For example, to de novo isolate SINE elements, the random sequencing of a relatively small portion of the genome is a recommended strategy (Nishihara and Okada 2008). In other instances, low coverage genomic surveys (even <0.1×) yield enough data for a good picture of the repeat content (Rasmussen and Noor 2009; Leese et al. 2012).

Here, we present the first results based on repeat masking and de novo characterization of repeated DNAs in genomes of *Bacillus*.

Table 1. Transposable element (TE) families found in the three analyzed genomes.

TE class/ superfamily	TE family	*Bacillus rossius*	*Bacillus grandii*	*Bacillus atticus*
Class I				
LTR		11	5	9
	Bel	2	2	*
	Gypsy	9	3	9
non-LTR		2	2	2
	Jockey	*	1	*
	L2B	n.f.	n.f.	1
	Nimb	1	*	*
	Outcast	*	1	*
	Penelope	1	*	*
	RTE	n.f.	*	1
Class II		16	18	20
	Academ	1	*	*
	Chapaev	*	*	1
	Harbinger	2	*	*
	hAT	5	1	2
	Helitron	1	3	3
	Kolobok	*	2	*
	Mariner	5	6	3
	P	*	1	*
	PiggyBac	*	1	*
	Polinton	2	3	10
	Sola	n.f.	1	1
Total		29	25	31

Note: The number of clones showing significant homology with listed TE families is given. Asterisks indicate TE presence verified through Southern Blot analysis. n.f., not found.

Data presented here will constitute the starting point for further analysis, aiming to clarify the relationships between repetitive DNA sequences and the reproductive biology of the host species.

Materials and methods

Samples and genomic DNA isolation

A *B. rossius* female (Patti, Sicily, gonochoric population), a *B. grandii* male (Ponte Manghisi, Sicily), and a *B. atticus* female (Scoglitti, Sicily) were utilized for the analyses. Gut-deprived specimens were maintained at –80 °C until the DNA isolation was performed through a standard phenol–chloroform procedure.

DNA library construction

For each library, 2 μg of genomic DNA was partially digested with *Eco*RI restriction enzyme (Invitrogen, Carlsbad, Calif., USA) for 2 h at 37 °C. Fragments were ligated to *Eco*RI-adapters (5′-CTCGTAGACTGCGTACC-3′; 5′-AATTGGTACGCAGTCTAC-3′) and then amplified with adaptor-specific primers (5′-GACTGC GTACCAATTCN-3′). After a 1% agarose gel electrophoresis, fragments between 800–1200 bp were recovered by gel extraction and cloned into a pGem-T Easy Vector (Promega, Madison, Wis., USA) used to transform *E. coli* DH5α competent cells (Invitrogen, Carlsbad, Calif., USA). Recombinant colonies were screened by PCR amplification with T7/SP6 primers, under standard PCR conditions. In total, 196 clones per species were sequenced at Macrogen Inc. (Korea).

Southern blot analysis

For each of the 14 TE families missing in at least one of the three libraries, probes were obtained by PCR amplification using specifically designed primers (supplementary data, Table S1)[1]. The PCR program was as follow: initial denaturation at 95 °C for 2 min; 35 cycle of denaturation at 95 °C for 30 s, annealing at 48 °C for 30 s, elongation at 72 °C for 30 s; and a final elongation step at

[1]Supplementary data are available with the article through the journal Web site at http://nrcresearchpress.com/doi/suppl/10.1139/gen-2013-0107.

Table 2. Nucleotide identity of transposable element (TE) families.

TE class	TE family	Species	Avg. identity (%)	*N*	Total matches (bp)
Intraspecific comparisons					
Class I	*Bel*	*Bacillus grandii*	99	2	885
	Gypsy	*Bacillus rossius*	96	7	3340
		Bacillus grandii	97	2	902
		Bacillus atticus	97	3	3063
Class II	*Helitron*	*Bacillus grandii*	66	2	596
		Bacillus atticus	67	2	202
	Mariner	*Bacillus rossius*	99	2	821
		Bacillus grandii	92	5	3630
	Polinton	*Bacillus grandii*	98	2	735
		Bacillus atticus	85	8	3180
Interspecific comparisons					
Class I	*Gypsy*	*Bacillus rossius-grandii*	97	5	4374
		Bacillus rossius-atticus	96	4	2642
		Bacillus grandii-atticus	96	3	1812
Class II	*Helitron*	*Bacillus rossius-grandii*	69	2	326
	Mariner	*Bacillus rossius-atticus*	94	2	867
	Polinton	*Bacillus rossius-atticus*	68	4	797
		Bacillus grandii-atticus	91	6	3054

Note: *N*, number of clones compared.

72 °C for 4 min. PCR reactions were performed with the GoTaq amplification kit (Promega, Madison, Wis., USA), using 30 ng of genomic DNA. Twenty microlitres of each amplification product was separated on a 1.5% agarose gel and Southern blotted onto a positively charged nylon membrane. Hybridization was performed using the AlkPhos labelling and detection kit (GE Healthcare, Pittsburg, Pa., USA) following the manufacturer's protocol. Stringency washes allowed up to 10% of nucleotidic divergence between probes and target DNA.

Sequence analysis

TE identification was done by repeat masking on Repbase Update database with CENSOR web tool (Kohany et al. 2006). We took into account only the hits with nucleotidic score >500 or amino acidic score >300, or those presenting simultaneously amino acidic score >200 and positives >0.5. For family identification, we considered accurate only the alignments with amino acidic score >200 and positives >0.5. Sequence identity for each TE family has been calculated as Σ IA × (LA / LSA), where IA is the identity of the alignment (max identity = 1), LA is the length of the alignment (bp), and LSA is the length of the sum of the alignments (bp).

A de novo search of interspersed repeats was done by self-comparison of each library; an e-value $\leq 10^{-5}$ was set to define significant high scoring segment pairs. Copy number of de novo identified interspersed repeats was calculated following the formula: (No. of occurrences in the library × genome size) / library size.

Differences in relative abundance of scored TEs in the three libraries were tested with repeated measure analysis of variance (ANOVA), followed by post-hoc paired *t* test with Holm correction.

Neighbor-joining tree, using uncorrected *p*-distance, was calculated with MEGA v.5 (Tamura et al. 2011); nodal support was obtained after 500 bootstrap replicates.

Tandem repeat searches were performed by Phobos v.3.3.11 (Mayer 2010), allowing extend exact search, repeat unit size from 2 to 500 bp, and at least four consecutive units.

Sequences were deposited in GenBank under accession numbers KF256266–KF256815.

Results and discussion

For each species, 196 random genomic fragments ~1000 bp long were obtained; on the whole, 144 250, 144 896, and 173 946 bp have been sequenced for *B. rossius*, *B. grandii*, and *B. atticus*, respectively. The genome sizes of these three stick insect species are 2.12–1.90, 2.55–2.11, and 2.25 Gbp, respectively; therefore, the sequencing corresponds to less than 0.006%–0.007% genomic coverage. The average GC content calculated within sequenced libraries is 40.3%, ranging from 39.3% in *B. atticus* to 42.2% in *B. rossius*.

Transposable elements

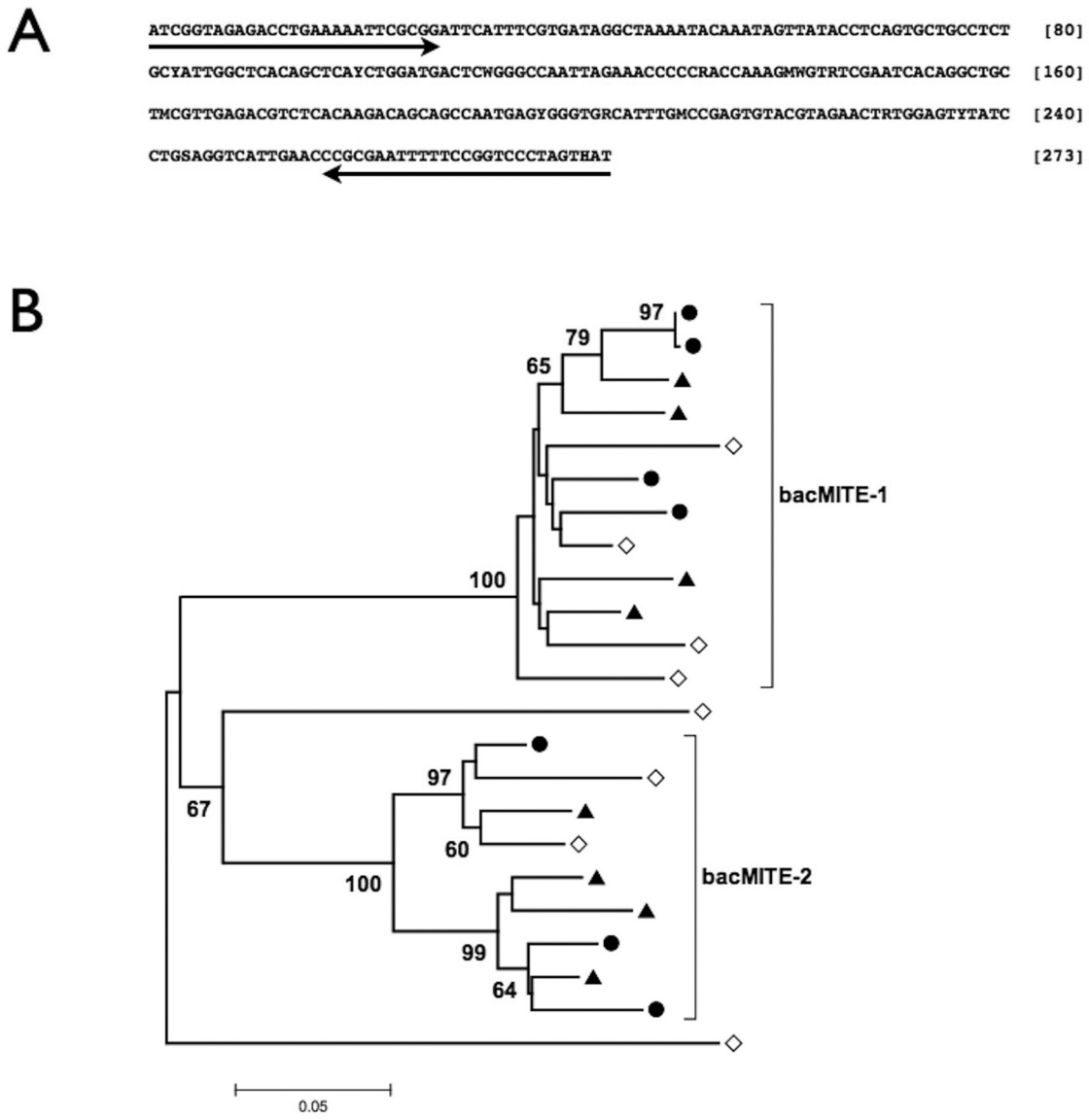
Repeat masking on the three libraries revealed 85 clones containing either class I or class II TEs.

Class I mobile elements belong to the following eight families: *Bel*, *Gypsy*, *Jokey*, *L2B*, *Nimb*, *Outcast*, *Penelope*, and *RTE*. *Bel* and *Gypsy* are LTR retroelements, while the other six are non-LTR retrotransposons (Table 1). Eleven families of class II elements have been identified as follows: *Academ*, *Chapaev*, *Harbinger*, *hAT*, *Helitron*, *Kolobok*, *Mariner*, *P*, *piggyBac*, *Polinton*, and *Sola* (Table 1).

All families have been identified in the three libraries either by clone sequencing or through Southern blot analysis, with the exceptions of *L2B*, *RTE*, and *Sola* in *B. rossius* and *L2B* in *B. grandii* (Table 1). On the whole, the parthenogenetic *B. atticus* shows the presence of all families and shares the majority of them with *B. grandii*, as it could be expected on the basis of phylogenetic relationships (Mantovani et al. 2001). The absence of some families in the gonochoric genomes of *B. rossius* and *B. grandii* may witness the greater ability of bisexuals in dealing with TEs; moreover, considering the time elapsed since the species splitting (23–17 Myr ago; Mantovani et al. 2001), these families could have had enough time for accumulating nucleotide divergence over 10%, and thus becoming undetectable under the Southern blot conditions used in this analysis.

For each library, TE families detected in at least two clones with overlapping regions were analyzed to evaluate the level of similarity (Table 2). These comparisons involved the LTR elements *BEL* and *Gypsy*, and the DNA families *Helitron*, *Mariner*, and *Polinton*. Identity values indicate a substantial intraspecific conservation of sequences, possibly being copies of the same element. The only exception is given by the *Helitron* elements in *B. grandii* and *B. atticus*, with identity values falling to 66% and 67%, respectively. Interspecific comparisons involve the LTR *Gypsy* and the class II elements *Helitron*, *Mariner*, and *Polinton* (Table 2). *Gypsy* and *Mariner* appear well conserved across species, as well as *Polinton* between *B. grandii* and *B. atticus*. On the other hand, a low degree of identity is found for *Helitron* in the *B. rossius* versus *B. grandii* comparison and *Polinton* in the *B. rossius* versus *B. atticus* comparison. On the whole, *Helitron* appears the less conserved element both within

Fig. 1. (A) The 50% majority rule consensus sequences of bacMITE elements. Arrows mark the terminal inverted repeats (TIRs). (B) Neighbor-joining tree based on uncorrected *p*-distance between bacMITE sequences; bootstrap values are calculated after 500 replicates. The two clusters corresponding to the two subfamilies have been indicated. Empty diamonds, *Bacillus rossius* clones; filled circles, *Bacillus grandii* clones; and filled triangles, *Bacillus atticus* clones.



and between genomes: further investigations will clarify if this could be due to higher element diversity.

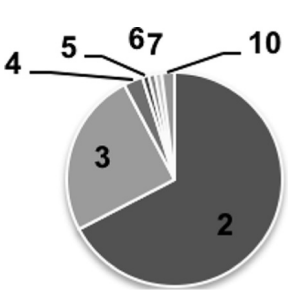
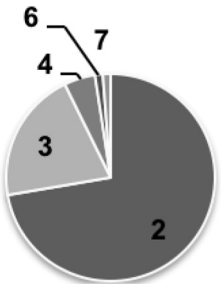
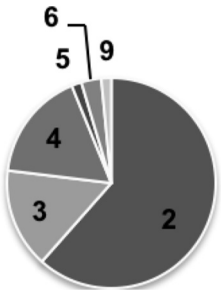
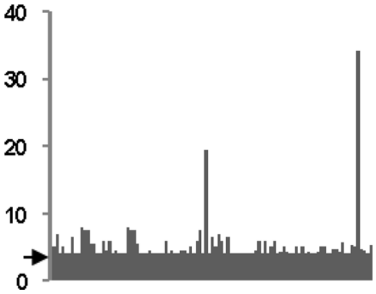
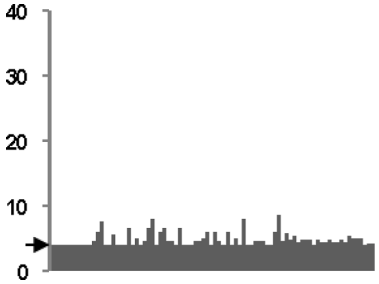
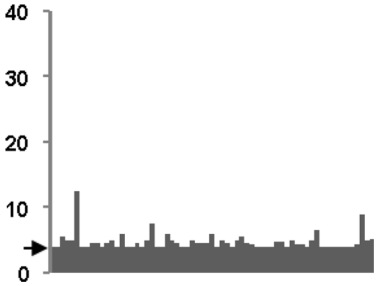
De novo identification of interspersed repeats led to the characterization of 23 homologous nucleotidic stretches 124–277 bp long, distributed in 21 clones from the three genomes, and sharing 77.3% sequence similarity. Aligned sequences gave a consensus length of 273 bp with terminal inverted repeats (TIRs; Fig. 1A). The short length and TIRs are common features of MITEs, small nonautonomous DNA transposon (Feschotte et al. 2002): we, therefore, named the retrieved sequences as bacMITEs. In the phylogenetic analysis, sequences are distributed in two main clusters having 100% nodal support; only two sequences fall outside the two clusters and may represent highly diverging or recombinant elements (Fig. 1B). Therefore, two possible MITE subfamilies can be identified (bacMITE-1 and bacMITE-2), showing within-subfamily sequence identity ranging from 81.5% to 91.2%. The relative copy numbers are 5.6×10^4, 4.8×10^4, 5.6×10^4, for bacMITE-1, and 5.6×10^4, 6.4×10^4, 5.6×10^4, for bacMITE–2, in *B. rossius*, *B. grandii*, and *B. atticus*, respectively. No flanking target site duplications (TSD) have been identified, but this could be due to the low number of identified full-length elements (three bacMITE-1 and one bacMITE-2). To our knowledge, bacMITEs are the first MITE elements ever found within a polyneopteran genome.

It is worth noting that no SINEs have been identified in this survey. Of course, this could be due to the limited sequencing

Table 3. Occurrences of transposable element (TE), relative percentages of TE classes among positive clones, and class II/class I ratio.

	Bacillus rossius	*Bacillus grandii*	*Bacillus atticus*
Total TEs/library	23.3%	18.0%	22.9%
Class I			
LTR	27.5%	11.8%	20.5%
non-LTR	10.0%	14.7%	6.8%
Class II	62.5%	73.5%	72.7%
Class II/class I	1.667	2.778	2.667

Table 4. Tandem repeat abundance (bp) detected in the three libraries.

	Bacillus rossius	*Bacillus grandii*	*Bacillus atticus*
Microsatellite	1.00% (103)	0.64% (83)	0.54% (65)
Unit length distribution (bp)			
Array length distribution (bp)			
Minisatellite	0.32% (2)	0.00% (0)	0.36% (3)
Unit length/copy number	15/11 45/7 —	n.f. — —	16/8 19/6 30/10
Satellite	0.00% (0)	0.00% (0)	0.96% (2)
Total	1.32% (105)	0.64% (23)	1.86% (70)

Note: The number of loci is given in parentheses. For microsatellite loci, array distributions are also shown (*x* axis, unit length; *y* axis, copy number; arrow indicates the four copy threshold used as the minimum tandem array length during Phobos v.3.3.11 search; each bin represents a single locus). n.f., not found.

relative to the genome size of *Bacillus* spp., but it should be noted that, in other instances, low coverage sequencing gave at least some SINE sequences. For example, in the termite *Reticulitermes lucifugus*, 25 SINE sequences belonging to four distinct SINE families have been found by random sequencing ~130 000 bp (approximately 0.012% of the termite genome; Luchetti and Mantovani 2011). In the hyrax genome, the random sequencing of 63 000 bp (~0.002% of the whole genome) revealed 26 AfroSINE elements (Nishihara and Okada, 2008). It is, thus, possible to hypothesize that the genome of *Bacillus* lacks SINEs, as observed in species of *Drosophila* (Kramerov and Vassetzky 2011), or that they are very poorly represented. Further sequencing or focused experiments on SINE search will allow more definitive conclusions.

As a general picture, LTR, non-LTR, and DNA transposons have quite different representativeness ($p = 0.002$), though this abundance distribution is not significantly different among the surveyed libraries ($p = 0.521$).

DNA transposons are generally prevailing in all libraries (class II versus LTR, $p = 0.00060$; class II versus non-LTR, $p = 0.00032$), their amount ranging from 1.666- to 2.778-fold higher than that of retrotransposons. This variability is not unexpected, taking into account that the class II/class I ratio may vary from ~100% class II to ~100% class I elements. For example, in two closely related insect species, the Culicidae mosquitoes *Anopheles gambiae* and *Aedes aegypti*, this ratio ranges from ~0.7 to ~2.3, respectively (reviewed in Feschotte and Pritham 2007). Within class I elements, LTR families outnumber non-LTR families (Table 3), but this difference is not significant ($p = 0.17433$).

Tandem repeats

On the whole, less than 2% of the sequenced genomic fragments in each species of *Bacillus* is constituted by tandem repeats.

As expected, microsatellite loci (2–10 bp) are more represented than minisatellite or satellite DNA. Di- and trinucleotide loci are the most abundant, with few instances of longer units (Table 4). However, most of the retrieved loci retains the minimum length imposed during the repeat search or remains below the 10 copies; only two loci in *B. rossius* and one locus in *B. atticus* showed longer arrays. In the former species, a dinucleotide array showed 19 repeats and a pentanucleotide locus has up to 34 repeat units. In the latter taxon, a dinucleotide array is made by 12 repeats (Table 4).

Minisatellites (11–100 bp) occurred only in *B. rossius* and *B. atticus* libraries, with repeat unit length ranging from 16 to 45 bp and copy number comprised between 6 and 10 (Table 4). Interestingly, a 45-mer array found in *B. rossius* has significant homologies with clones from both *B. grandii* and *B. atticus* libraries. Sequence analysis indicates that one clone from *B. grandii* (gra_af4) and five from *B. atticus* (att_ab7, att_ah5, att_ah7, att_ah11, and att_bc8) have from two to three tandemly arranged 45-mer repeat units, with an overall repeat units sequence identity of 62%. Therefore, the 5′ and the 3′ array flanking regions were compared to check if the same genomic locus harbours this minisatellite in all the three genomes: while the 3′ end flanking region did not show any homology among clones, the 5′ flanking region was significantly conserved even between species (74.1% pairwise identity; 84.4% of identical sites), with the exception of the clone att_ah11. The analysis of the consensus sequence generated from the alignment of 5′ flanking regions did not give any similarity with any known sequence in public databases. Tandem repeat arrays flanked by

homologous sequences are, actually, commonly found: in fact, they can originate within the TE sequence (Mogil et al. 2012; Sharma et al. 2013) or, in most cases, they are generated at retrotransposon tails when they reintegrate into new genomic locations (Lopez-Giraldez et al. 2006; Meglécz et al. 2007; Coates et al. 2009, 2011; Luchetti and Mantovani 2009, 2011). However, retrotransposons usually generate microsatellite loci, while in this case a 45-bp unit made the tandem array. At present, it is not possible to further explain such occurrence, especially because the homologous flanking sequence is not similar to any known TE sequence. However, it would be interesting to check whether this minisatellite locus has been generated upon the insertion of a retrotransposon or by some recombinative mechanism during the evolution of the genome in species of the genus *Bacillus*.

Satellite DNA sequences (>100 bp) have been retrieved only in the *B. atticus* library, where five *Bag320* monomers (Mantovani et al. 1997; Cesari et al. 2003; Luchetti et al. 2003) have been found in two clones (att_ac9 and att_ba1). With respect to previously isolated *Bag320* sequences, they show a sequence similarity with *B. atticus* specific monomers ranging from 94.1% to 96.0%. No *Bag320* sequences have been found in the genomes of *B. rossius* or *B. grandii*. For the former species, this was quite expected, as it is known that this satellite DNA occurs at a very low copy number in this genome and was only isolated by PCR amplification (Cesari et al. 2003). Its absence in the *B. grandii* library is, however, more surprising because this genome contains the highest copy number of the satellite family (15%–20% in *B. grandii* versus 2%–5% in *B. atticus*; Mantovani et al. 1997). As a general consideration, the use of *Eco*RI restriction enzyme for library production may have biased the genomic sampling of *Bag320* sequences, as they do not contain its cutting site in the considered species; therefore, its sampling from *B. atticus* genomes could be considered as mere chance.

On the whole, the sequenced libraries represent a small fraction of the whole genome of *B. rossius*, *B. grandii*, and *B. atticus*, with less than 1% genomic coverage. Yet, the survey allowed the retrieval of a number of repetitive DNAs, either interspersed or not. Most of the main TE families are represented, and a MITE family, the first ever discovered in polyneopteran insects, has been de novo characterized. Moreover, mini- and microsatellite loci were found even if characterized by short arrays. Unfortunately, it is not possible to make strong comparisons on the representativeness of repetitive DNAs within the sequenced libraries, as no polyneopteran genomes have been sequenced so far and the only polyneopteran DNA library available has been built from the termite species *R. lucifugus*. In this library (covering 0.012% of the genome), four SINE and one putative MITE families were de novo characterized (Luchetti and Mantovani 2011; A. Luchetti, unpublished data); moreover, 11 minisatellite and 298 microsatellite loci were isolated in a single termite species. Therefore, despite the smaller genome (~1 versus >2 Gbp), termites appear to have a more repetitive genomic landscape than stick insects. As a final remark, it is interesting to point out the relative TE content showed by unisexual (*B. atticus*) as opposed to gonochoric taxa (*B. rossius* and *B. grandii*). Both theoretical and empirical studies evidenced that parthenogenetic, i.e., low recombining, genomes should avoid TE accumulation to escape the effects of Muller's ratchet; otherwise, TE load would raise without the possibility of elimination eventually leading to the host species extinction (Arkhipova and Meselson 2000, 2005; Sullender and Crease 2001; Dolgin and Charlesworth 2006). On the other hand, in the parasitoid wasp *Leptopilina clavipes*, unisexual and bisexual lineages showed no difference in the overall TE genomic coverage (Kraaijeveld et al 2012). In line with this, *B. atticus* genome appears to have (at least) the same TE content of the bisexual species. As argued by Wright and Finnegan (2001), TE prevalence studies in obligate unisexual genomes are difficult to interpret mainly because they are derived from bisexual ancestors. Thus, at the moment, it is impossible to state whether *B. atticus* and *L. clavipes* TEs have been selected for less harmful elements or they are still in the "elimination phase", i.e., the two taxa are not unisexual since enough time to allow an efficient clearance of TE load. On the whole, although based on small datasets and considering the different reproductive strategies, the variation of ploidy, and the presence of hybrid taxa that characterize the *Bacillus* complex, the present survey provides interesting preliminary data to undertake further analysis in species of the genus *Bacillus*.

Acknowledgements

This work was funded by Canziani and RFO funds to B.M. M.R. produced the three libraries and carried out repeat masking; A.L. performed de novo TE searches and tandem repeat characterization; L.B. performed Southern blot analysis; B.M. supervised the project; and M.R., A.L., and B.M. wrote the article. Authors wish to thank the two reviewers whose criticisms and suggestions helped to substantially improve the paper.

References

Arkhipova, I., and Meselson, M. 2000. Transposable elements in sexual and ancient asexual taxa. Proc. Natl. Acad. Sci. U.S.A. **97**(26): 14473–14477. doi:10.1073/pnas.97.26.14473. PMID:11121049.

Arkhipova, I., and Meselson, M. 2005. Deleterious transposable elements and the extinction of asexuals. Bioessays, **27**(1): 76–85. doi:10.1002/bies.20159. PMID:15612027.

Brookfield, J.F.Y., and Badge, R.M. 1997. Population genetics models of transposable elements. Genetica, **100**(1–3): 281–294. doi:10.1023/A:1018310418744. PMID:9440281.

Bull, J.J., Molineux, I.J., and Rice, W.R. 1991. Selection of benevolence in a host-parasite system. Evolution, **45**(4): 875–882. doi:10.2307/2409695.

Cesari, M., Luchetti, A., Passamonti, M., Scali, V., and Mantovani, B. 2003. Polymerase chain reaction amplification of the *Bag320* satellite family reveals the ancestral library and past gene conversion events in *Bacillus rossius* (Insecta Phasmatodea). Gene, **312**: 289–295. doi:10.1016/S0378-1119(03)00625-5. PMID:12909366.

Charlesworth, B., Sniegowski, P., and Stephan, W. 1994. The evolutionary dynamics of repetitive DNA in eukaryotes. Nature, **371**(6494): 215–220. doi:10.1038/371215a0. PMID:8078581.

Coates, B.S., Sumerford, D.V., Hellmich, R.L., and Lewis, L.C. 2009. Repetitive genome elements in a European corn borer, *Ostrinia nubilalis*, bacterial artificial chromosome library were indicated by bacterial artificial chromosome end sequencing and development of sequence tag site markers: implications for lepidopteran genomic research. Genome, **52**(1): 57–67. doi:10.1139/G08-104. PMID:19132072.

Coates, B.S., Kroemer, J.A., Sumerford, D.V., and Hellmich, R.L. 2011. A novel class of miniature inverted repeat transposable elements (MITEs) that contain hitchhiking $(GTCY)_n$ microsatellites. Insect. Mol. Biol. **20**(1): 15–27. doi:10.1111/j.1365-2583.2010.01046.x. PMID:20977507.

Dolgin, E.S., and Charlesworth, B. 2006. The fate of transposable elements in asexual populations. Genetics, **174**(2): 817–827. doi:10.1534/genetics.106.060434. PMID:16888330.

El-Murr, N., Abidi, Z., Wanherdrick, K., Svrcek, M., Gaub, M.P., Fléjou, J.F., et al. 2012. *MiRNA* genes constitute new targets for microsatellite instability in colorectal cancer. PLoS ONE, **7**(2): e31862. doi:10.1371/journal.pone.0031862. PMID:22348132.

Feschotte, C., and Pritham, E. 2007. DNA transposons and the evolution of eukaryotic genomes. Annu. Rev. Genet. **41**: 331–368. doi:10.1146/annurev.genet.40.110405.090448. PMID:18076328.

Feschotte, C., Zhang, X., and Wessler, S.R. 2002. Miniature inverted-repeat transposable elements (MITEs) and their relationship with established DNA transposons. *In* Mobile DNA II. *Edited by* N. Craig, R. Craigie, M. Gellert, and A. Lambowitz. A.S.M. Press, Washington, D.C. pp. 1147–1158.

Fine, P.E.M. 1975. Vectors and vertical transmission: an epidemiologic perspective. Ann. N.Y. Acad. Sci. **266**: 173–194. doi:10.1111/j.1749-6632.1975.tb35099.x. PMID:829470.

Galindo, C.L., McIver, L.J., Tae, H., McCormick, J.F., Skinner, M.A., Hoeschele, I., et al. 2011. Sporadic breast cancer patients' germline DNA exhibit an AT rich microsatellite signature genes chromosomes. Cancer, **50**(4): 275–283. PMID:21319262.

Kazazian, H.H. 2004. Mobile elements: drivers of genome evolution. Science, **303**(5664): 1626–1632. doi:10.1126/science.1089670. PMID:15016989.

Kohany, O., Gentles, A.J., Hankus, L., and Jurka, J. 2006. Annotation, submission and screening of repetitive elements in Repbase: Repbase Submitter and Censor. BMC Bioinformatics, **25**(7): 474. PMID:17064419.

Kraaijeveld, K., Zwanenburg, B., Hubert, B., Vieira, C., De Pater, S., Van Alphen, J.J.M., et al. 2012. Transposon proliferation in an asexual parasi-

toid. Mol. Ecol. **21**(16): 3898–3906. doi:10.1111/j.1365-294X.2012.5582.x. PMID: 22548357.

Kramerov, D.A., and Vassetzky, N.S. 2011. Origin and evolution of SINEs in eukaryotic genomes. Heredity, **107**(6): 487–495. doi:10.1038/hdy.2011.43. PMID: 21673742.

Leese, F., Brand, P., Rozenberg, A., Mayer, C., Agrawal, S., Dambach, J., et al. 2012. Exploring Pandora's box: potential and pitfalls of low coverage genome surveys for evolutionary biology. PLoS ONE, **7**(11): e49202. doi:10.1371/journal.pone.0049202. PMID:23185309.

López-Giráldez, F., Andrés, O., Domingo-Roura, X., and Bosch, M. 2006. Analyses of carnivore microsatellites and their intimate association with tRNA-derived SINEs. BMC Genomics, **7**: 269. doi:10.1186/1471-2164-7-269.

Luchetti, A., and Mantovani, B. 2009. *Talua* SINE biology in the genome of the *Reticulitermes* subterranean termites (Isoptera, Rhinotermitidae). J. Mol. Evol. **69**(6): 589–600. doi:10.1007/s00239-009-9285-7. PMID:19904483.

Luchetti, A., and Mantovani, B. 2011. Molecular characterization, genomic distribution and evolutionary dynamics of Short INterspersed Elements in the termite genome. Mol. Genet. Genomics, **285**(2): 175–184. doi:10.1007/s00438-010-0595-7.

Luchetti, A., Cesari, M., Carrara, G., Cavicchi, S., Passamonti, M., Scali, V., and Mantovani, B. 2003. Unisexuality and molecular drive: *Bag320* sequence diversity in *Bacillus* taxa (Insecta Phasmatodea). J. Mol. Evol. **56**(5): 587–596. doi:10.1007/s00239-002-2427-9. PMID:12698295.

Makałowski, W., Pande, A., Gotea, V., and Makałowska, I. 2012. Transposable elements and their identification. Methods Mol. Biol. **855**: 337–359. doi:10.1007/978-1-61779-582-4_12. PMID:22407715.

Mantovani, B., Tinti, F., Bachmann, L., and Scali, V. 1997. The *Bag320* satellite DNA family in *Bacillus* stick insects (Phasmatodea): different rates of molecular evolution of highly repetitive DNA in bisexual and parthenogenetic taxa. Mol. Biol. Evol. **14**(12): 1197–1205. doi:10.1093/oxfordjournals.molbev.a025729. PMID:9402731.

Mantovani, B., Passamonti, M., and Scali, V. 2001. The mitochondrial cytochrome oxidase II gene in *Bacillus* stick insects: ancestry of hybrids, androgenesis, and phylogenetic relationships. Mol. Phylogenet. Evol. **19**(1): 157–163. doi:10.1006/mpev.2000.0850. PMID:11286500.

Mayer, C. 2010. Phobos 3.3.11. Available from http://worldwidewords.rub.de/spezoo/cm/cm_phobos.htm [accessed 22 April 2013].

Meglécz, E., Anderson, S.J., Bourguet, D., Butcher, R., Caldas, A., Cassel-Lundhagen, A., et al. 2007. Microsatellite flanking region similarities among different loci within insect species. Insect. Mol. Biol. **16**(2): 175–185. doi:10.1111/j.1365-2583.2006.00713.x. PMID:17298557.

Mogil, L.S., Slowikowski, K., and Laten, H.M. 2012. Computational and experimental analyses of retrotransposon-associated minisatellite DNAs in the soybean genome. BMC Bioinformatics, **13**(Suppl. 2): S13. doi:10.1186/1471-2105-13-S2-S13. PMID:22536864.

Nishihara, H., and Okada, N. 2008. Retroposons: genetic footprints on the evolutionary paths of life. Methods Mol. Biol. **422**: 201–225. doi:10.1007/978-1-59745-581-7_13. PMID:18629669.

Plohl, M., Luchetti, A., Meštrović, N., and Mantovani, B. 2008. Satellite DNAs between selfishness and functionality: structure, genomics and evolution of tandem repeats in centromeric (hetero)chromatin. Gene, **409**(1–2): 72–82. doi:10.1016/j.gene.2007.11.013. PMID:18182173.

Rasmussen, D.A., and Noor, M.A.F. 2009. What can you do with 0.1× genome coverage? A case study based on a genome survey of the scuttle fly *Megaselia scalaris* (Phoridae). BMC Genomics, **10**: 382. doi:10.1186/1471-2164-10-382.

Richard, G.-F., Kerrest, A., and Dujon, B. 2008. Comparative genomics and molecular dynamics of DNA repeats in eukaryotes. Microbiol. Mol. Biol. Rev. **72**(4): 686–727. doi:10.1128/MMBR.00011-08. PMID:19052325.

Scali, V., Passamonti, M., Marescalchi, O., and Mantovani, B. 2003. Linkage between sexual and asexual lineages: genome evolution in *Bacillus* stick insects. Biol. J. Linn. Soc. **79**: 137–150. doi:10.1046/j.1095-8312.2003.00172.x.

Sharma, A., Wolfgruber, T.K., and Presting, G.G. 2013. Tandem repeats derived from centromeric retrotransposons. BMC Genomics, **14**: 142. doi:10.1186/1471-2164-14-142. PMID:23452340.

Sullender, B.W., and Crease, T.J. 2001. The behavior of a *Daphnia pulex* transposable element in cyclically and obligately parthenogenetic populations. J. Mol. Evol. **53**(1): 63–69. PMID:11683324.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. 2011. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance and maximum parsimony methods. Mol. Biol. Evol. **28**: 2731–2739. doi:10.1093/molbev/msr121. PMID:21546353.

Wright, S., and Finnegan, D. 2001. Genome evolution: sex and the transposable element. Curr. Biol. **11**: R296–R299. doi:10.1016/S0960-9822(01)00168-3. PMID: 11369217.

Wright, S.I., and Schoen, D.J. 1999. Transposon dynamics and the breeding system. Genetica, **107**(1–3): 139–148. doi:10.1023/A:1003953126700. PMID: 10952207.

Table S1. List of primers designed for the families of TE subjected to Southern blot analysis. For each clone's name the abbreviations ROS, GRA and ATT indicate the species belonging.

Primer name	**Clone name**	**Probe size bps**	**sequence 5'>3'**	**TE family**
FSOLA	gra_ad9	250	CAC AAG ATA GAG TAA GCT CTG ACG	SOLA
RSOLA	gra_ad9	250	CCC TTT GCT TTG ATT TGT CA	SOLA
FBEL	gra_ad1	398	AAT TGG CAG CAA CCG ATT C	BEL
RBEL	gra_ad1	398	CCG ATC CAT CAG TAA TTT CCA	BEL
FP	gra_be3	393	AAT TCG TCC AGT CCC ATT TG	P
RP	gra_be3	393	ATT GCC ATT TCC TGA ACA CC	P
FPiggy bac	gra_ba9	398	ATA AGG TGA AAC CGC GAA TG	Piggy bac
RPiggy bac	gra_ba9	398	GCC GGG AGT GGG TAA TTT	Piggy bac
FKolobok	gra_bf3	488	TGT GAT GGC TTC AGT CCA AA	Kolobok
RKolobok	gra_bf3	488	TCT CTT GTT TTT CGA GTT GCT G	Kolobok
FOutcast	gra_be8	495	GGT GGT CGA ATC GAA AAC AT	Outcast
ROutcast	gra_be8	495	TGT TCC TTG AAT CAA AAG ACA AAA	Outcast
FPenelope	ros_ad8	369	TAG CTT CCA GGG AAA ACT GC	Penelope
RPenelope	ros_ad8	369	TGG CTG ATA TAT TGC TGT GAG G	Penelope
FNimb	ros_bd12	498	GAT GAA AGC GAC ATC GAT CC	Nimb

RNimb	ros_bd12	498	AAC AAC CAG CCA CGA ACA G	Nimb
FAcadem	ros_bd3	482	AAC TAT TGG CAT GGT GAC TGC	Academ
RAcadem	ros_bd3	482	CCA TTC AAA GAC TCC CTT GG	Academ
FHarbinger	ros_ae6	387	ACA CAC CAA TCC GAA AAT CC	Harbinger
RHarbinger	ros_ae6	387	GCG ATG CCG GAA AAT AAC T	Harbinger
FRTE	att_bb12	499	GAA ATG GAT CGG GCA TAC AC	RTE
RRTE	att_bb12	499	TGG CTT AAT CCC TGT TGT CC	RTE
FL2B	att_bd6	400	TTT CAC AAA GGT CGG GTG AC	L2B
RL2B	att_bd6	400	GGA ACA ATT GAG CGT GTG TC	L2B
FChapaev	att_be7	399	TCC TCG ACA GTG GAT CTC TAC A	Chapaev
RChapaev	att_be7	399	GGG CCC TAC GAA GCA GTT TA	Chapaev
NOGR F	gra_bc9	558	ATG GTC GCG AGT AGG ACAC	Jokey
NOGR R	gra_bc9	558	GCC GTG TGC ATT GAA AAT TAG	Jokey

These results of the thesis were also presented at several Congresses. Please read the abstracts in Appendix.

CHAPTER 3

DISCUSSION

My Ph.D. thesis concerns the relationship between sex and repetitive sequence dynamics. Eukaryotic genomes are widely made up of repeated elements, which play an important biological role in their structure and variability (Plohl *et al*., 2008; Richard *et al*., 2008; Chénais *et al*., 2012). Several studies have proposed a link between the reproductive biology of the host organism and the content of some repetitive sequences (Wright and Finnengan, 2001; Nuzhdin and Petrov, 2003; Schaack *et al*., 2010; Mingazzini, 2011; Ph.D. thesis; Ricci *et al*., 2013). Also my research project focused on the study of the impact of reproductive modality on the presence and activity of repeated elements. The analyses were conducted in *Bacillus* stick insects (Phasmida, Bacillidae) because of the complex pattern of reproductive strategies that characterizes the genus. In fact, in addition to canonical gonochoric species, the latter includes geographic facultative and obligate thelytokous parthenogenetic ones, as well as hybridogenetic and androgenetic taxa (see Chapter 1).

Therefore, the main aim of my thesis was to determine if and how sex can affect the load and the dynamics of repetitive sequences in *Bacillus* genomes, in particular transposable elements. These mobile repeated elements have been detected from bacteria to eukaryotes with a species-specific abundance (Chénais *et al*., 2012). According to the Muller's ratchet theory, TEs, being deleterious mutations, should be dealt only by bisexual genomes because they can counteract their proliferation through the recombination and the independent assortment of chromosomes (Felsenstein, 1974). On the contrary, unisexual genomes should not harbor TEs, apart from horizontal transmission events, because in this

instance their copies should accumulate leading to the host lineage extinction or be eliminated by genome purging (Kordis and Gubensek, 1999; Lampe *et al*., 2003; Sanchez-Gracia *et al*., 2005; Loewe and Lamatsch, 2008; Keeling and Palmer, 2008). In my thesis, the threat of extinction predicted by the Muller's ratchet theory was tested in parthenogenetic taxa of the genus *Bacillus*, with most analyses focused to evaluate the dynamics of the non-LTR element R2. This transposon represents the most studied TE, which inserts specifically in the 28S rDNA of many invertebrate and vertebrate phyla (Eickbush, 2002; Fujiwara, 2015).

In particular, the role played by sex in R2 activity was estimated at the species/population levels, through the insertion display method described by Pérez-Gonzáles and Eickbush (2001), in the strictly gonochoric *B*. *grandii maretimi* (Marettimo population) and in the obligate parthenogenetic *B*. *atticus atticus* (Necropoli Camarina and Scoglitti populations). These data completed a previous one on the facultative parthenogenetic *B*. *rossius* obtained either in bisexual populations (*B*. *rossius rossius* population from Anzio and *B*. *rossius redtenbacheri* one from Patti) or in the all-females *B*. *rossius redtenbacheri* population from Curcuraci (Mingazzini, 2011; Ph.D. thesis).

R2 dynamics was determined also in the offspring of facultative parthenogenetic isolates of *B*. *rossius redtenbacheri* (Castanea delle Furie, Massa San Nicola and Curcuraci) and in the male progeny of crosses involving parental females from bisexual populations (*B*. *rossius rossius* Anzio ♀ X *B*. *rossius rossius* Anzio ♂ and *B*. *rossius rossius* Capalbio ♀ X *B*. *rossius rossius* Capalbio ♂) and from all-females ones (*B*. *rossius redtenbacheri* Curcuraci ♀ X *B*. *rossius rossius* Anzio ♂).

R2 analyses were performed in the descendants of *B*. *atticus atticus* obligatory parthenogenetic isolates from the populations of Necropoli Camarina and Scoglitti, too. To compare R2 load through more generations, its activity was analysed up to the 5th

generation of a *B. rossius redtenbacheri* parthenogenetic isolate collected in Castanea delle Furie.

On the whole, R2 activity analyses, conducted in natural populations and in isolates' offspring, revealed that R2 is present and actively transposing in all *Bacillus* genomes, disregarding reproductive strategies. If in *B. rossius* parthenogenetic populations this can be ascribed to the facultative way of parthenogenesis and the possibility to switch between bi- and unisexuality, it was on the contrary quite unexpected in *B. atticus atticus*. Indeed, insertions were detected in field collected females of the considered populations of *B. atticus atticus* in a lower number (=5/population) than those found in facultative parthenogenetic *B. rossius* (=24/population) and gonochoric *B. grandii maretimi* one (=26). The greater occurrence of low frequency insertions highlighted in all samples, including *B. atticus atticus*, do not support the Muller's ratchet theory in the genus *Bacillus*, too. In fact, only in the bisexual *B. rossius rossius* population from Patti high frequency insertions were more represented. In *B. atticus atticus* progeny the R2 accumulation was even lower with respect to the facultative parthenogenetic *B. rossius* offspring (on average 6 *vs* 15) thus suggesting also that the mechanisms of genetic turnover are active. The insertion frequency spectra among the species were significantly different, but it is worth noting that this result could be due to the low R2 abundance in *B. atticus atticus*. Instead, intraspecific comparisons in *B. rossius* insertion patterns showed that R2 content was slightly higher in unisexual descendants than in bisexual ones, even if this difference was not significant. On the contrary, the insertion frequency spectra of unisexuals and bisexuals were significantly different: low frequency insertions were more represented in the bisexual offspring, while most of high frequency ones were detected in the unisexual offspring, as expected by theoretical predictions and as observed in previous analyses in *Arabidopsis* spp. and *Caenorhabditis* spp. (Lockton and Gaut, 2010 and reference therein, Dolgin *et al*., 2008).

The comparison of R2 dynamics in more generations, from G_0 to G_5, performed on the isolated female of the facultative parthenogenetic *B. rossius redtenbacheri* population from Castanea delle Furie, underlined a pattern in which the TE load is kept low. In particular, the number of R2 insertions revealed in this female offspring was lower than that scored in the other two selected females offspring from the same population (see Chapter 2). This suggests the need to widen R2 dynamics analyses given the possible intraspecies/population variation due to the origin age of the parthenogenetic lineage.

On the whole, data obtained in this thesis evidenced that the number of R2 inserted sites tends to be higher in the facultative parthenogen *B. rossius* than in the obligate parthenogen *B. atticus atticus*. The different trend could be due to the parthenogenetic mechanism; in fact, parthenogenesis in *B. rossius* can give origin to homozygous progeny, while in *B. atticus* the maternal genetic asset is maintained, unless the occurrence of events of recombination. Moreover, models of R2 parental-offspring inheritance under different reproductive strategies (bisexuality, facultative/obligatory parthenogenesis) demonstrated that recombination seems to play the major role in counteracting R2 proliferation but at the same time these simulations explained that also selection contributes to keep R2 at bay. In *B. atticus* and bisexual populations of *B. rossius* a greater variance of R2 insertions is expected. Thus, natural selection would have more opportunity to operate in *B. atticus*, keeping its genome purged by TE insertions.

It is finally to be noted that *B. rossius* parthenogenetic populations owing to the facultative origin of it are of undefined age, while *B. atticus* is parthenogenetic from about 1 Mya (Mantovani *et al*., 2001).

The results of my thesis confirm that obligate unisexuals can face TE presence (Valizadeh and Crease, 2008; Gladyshev and Arkhipova, 2010a and 2010b; Schaack *et al*., 2010). Their survival may be allowed by mechanisms able to keep the TE load low, that in addition to the above cited recombination and selection, could be constituted of silencing

mechanisms (Eickbush *et al*., 2008). As explained by Eickbush *et al*. (2008), R2 densely populated rDNA regions can be silenced by heterochromatinization. Therefore, if the number of R2 inserted sites is low, their silencing will be easily achieved and the R2 activity considerably reduced.

The selection of less virulent TEs can also be hypothesized to avoid host damage, and therefore TEs survival (Bull *et al*., 1991; Wright and Finnegan, 2001).

Insertion analyses conducted on several generations of *B*. *atticus* isolates could allow to test also the role of clonal selection in the control of R2 proliferation. Because of the relatively long generation time that characterizes the genus *Bacillus* (up to 10 months) this hypothesis has not yet been verified.

To conclude, data produced in my thesis indicate that R2 is present and active in all the analysed *Bacillus* taxa, irrespective of their reproductive strategy. Therefore, R2 is actively transposing even in the obligatory parthenogenetic *B*. *atticus*, counteracting Muller's ratchet predictions.

Moreover, results here presented show that sex affects the dynamics of R2 in *Bacillus* stick insects, owing to the different values of R2 load scored.

Main above reported was also demonstrated by the paper by Ricci *et al*. (2013), to which I contributed by Southern blot analyses to confirm the presence of 14 TE families. In particular, to check for a link between reproductive biology and the genomic content of repetitive sequences in *Bacillus* stick insects, three low coverage libraries (~ 0.006% of the genome) were prepared for the strictly gonochoric *B*. *grandii*, the facultative parthenogenetic *B*. *rossius* and the obligatory parthenogenetic *B*. *atticus* (Ricci *et al*., 2013). The results indicated that the genome content of tandem repeats was of 0.64% in *B*. *grandii*, 1.32% in *B*. *rossius* and 1.86% in *B*. *atticus*. As expected, microsatellites were the most represented in the analysed species; minisatellites were found only in the

parthenogenetic *B*. *rossius* and *B*. *atticus*, and satellite DNA sequences were detected only in the obligate parthenogenetic *B*. *atticus* with five *Bag320* monomers (Ricci *et al*., 2013). On the whole, data here obtained highlighted that the load of tandem elements is lower than that of TEs in the three considered genomes. In fact, the occurrence of TEs is of 18% in *B*. *grandii*, 23.3% in *B*. *rossius* and 22.9% in *B*. *atticus*. In particular, in the three examined taxa 19 TE families were identified (two LTRs, six non-LTRs and 11 DNA transposons). Moreover, through a *de novo* analysis, two MITE subfamilies (bacMITE-1 and bacMITE-2) were observed in the three libraries for the first time in polyneopteran genomes (Ricci *et al*., 2013). Even the obligate unisexual *B*. *atticus* is characterized by the presence of TEs and furthermore, their content not lower than that revealed in *B*. *rossius* and *B*. *grandii*, again counteracting Muller's ratchet expectations (see Chapter 2).

CHAPTER 4

OTHER RESEARCH ACTIVITIES CARRIED OUT DURING THE Ph.D. PROGRAMME

I widened the analyses of the *Bacillus* repeated component to the characterization of telomeric DNA sequences.

Telomeres are nucleoprotein structures that cap the ends of eukaryotic chromosomes (Bernadotte *et al*., 2016).

Telomere complex discovery dates back to the late 1930's with the cytogenetic studies conducted in flies and corn by Herman Muller and Barbara McClintock, respectively (Zakian, 2012).

Since their observations, the capping function of telomeres was described as an important protection from the deleterious genomic instability (Blackburn, 2001 and reference therein). In fact, their presence plays an important role for the chromosome positioning and segregation, as well as for their stability and integrity (Plohl *et al*., 2002). They are especially involved in the replication of the ends of linear DNA molecules, fundamental for eukaryotic chromosomes integrity (Blackburn, 2001; Plohl *et al*., 2002). DNA polymerase, in fact, is not able to complete the synthesis of the tips of linear DNA molecules, so that telomeres should become shorter with each cell division, leading to the loss of genetic information and to the consequent cellular "senescence" (Blackburn, 2001, Bernadotte *et al*., 2016). Telomere shortening can be used as cellular senescence indicator and as mechanism for counting cell divisions (Bernadotte *et al*., 2016). In most eukaryotes, during the initial phases of development, the enzyme telomerase counteracts the loss of

genetic information catalysing the addition of tandemly repeated, often short G-rich, telomeric sequences (Blackburn, 2001 and reference therein).

Structurally, telomeres are made up of these tandem repeats (Blackburn, 2001). Their length on the whole amounts to 4000-15000 nucleotides and their composition is species-specific. At the 3' end the antiparallel strand is missing; to avoid chromosomal fusion and degradation, the single strand tail forms a loop structure named "T-loop" and binds proteins involved in the DNA capping. Besides the tandem repeats of a short sequence unit, satellite DNA sequences are also present at the telomere. The latter constitute the internal telomere-associated DNA and form the subtelomeric regions (Plohl *et al*., 2002). Moreover mobile elements are known to occur embedded in telomeres (Levin and Moran, 2011).

Although telomeres and telomerase functions are the same throughout the eukaryotes, telomeric sequences vary between taxa and, being conserved at high taxonomic levels, are considered taxon-specific (Plohl *et al*., 2002). Vertebrates are characterized by the TTAGGG repeat motif (Blackburn, 2001; Plohl *et al*., 2002). Instead, telomeric repeats in invertebrates show heterogeneity in DNA sequence and length (Plohl *et al*., 2002 and reference therein). It is worth noting that the studies conducted so far in invertebrates are focused on Platyhelminthes, Nematoda, Mollusca, and Arthropoda (Müller *et al*., 1991; Teschke *et al*., 1991; Joffe *et al*., 1998, Plohl *et al*., 2002; Frydrychová *et al*., 2004; Mora *et al*., 2015).

In particular, the telomeric sequence $(TTAGG)_n$ has been retrieved from most insects (Sahara *et al*., 1999; Frydrychová *et al*., 2004). The pentanucleotide unit TTAGG probably represents the phylogenetically ancestral telomere motif in insects that has been replaced in several groups during evolution (Sahara *et al*., 1999; Kuznetsova *et al*., 2012). For example in the Tenebrionidae family of the order Coleoptera the DNA tandem repeat sequence within the telomere structure is $(TCAGG)_n$ (Mravinac *et al*., 2011). Most dipteran families

lacking telomerase maintain the chromosome length using arrays of satellites and retrotransposons (Nielsen and Edström, 1993; Frydrychová *et al*, 2004, Mason *et al*., 2008). The distribution of the canonical TTAGG within the class Insecta showed that it is conserved in the Apterygota but absent in the "primitive" Pterygota, the Palaeoptera. It is quite widespread in the lower Neoptera with the exception of Dermaptera, and in Paraneoptera in Heteroptera. In the Endopterygota orders it is heterogeneously distributed (Figure 19, Frydrychová *et al*., 2004; Korandová *et al*., 2014). A study of the distribution of TTAGG-specific telomerase activity in insects was conducted with the aim of analysing either the functionality of the $(TTAGG)_n$ telomere sequence or the TTAGG-telomerase system. Among the surveyed species, telomerase activity was detected in Neoptera (Isoptera and Blataria), in Paraneoptera (Sternorrhyncha) and in Endopterygota (Coleoptera, Trichoptera, Lepidoptera and Hymenoptera; Figure 19; Korandová *et al*., 2014). The $(TTAGG)_n$ sequence distribution and the telomerase activity one are consistent (Korandová *et al*., 2014).

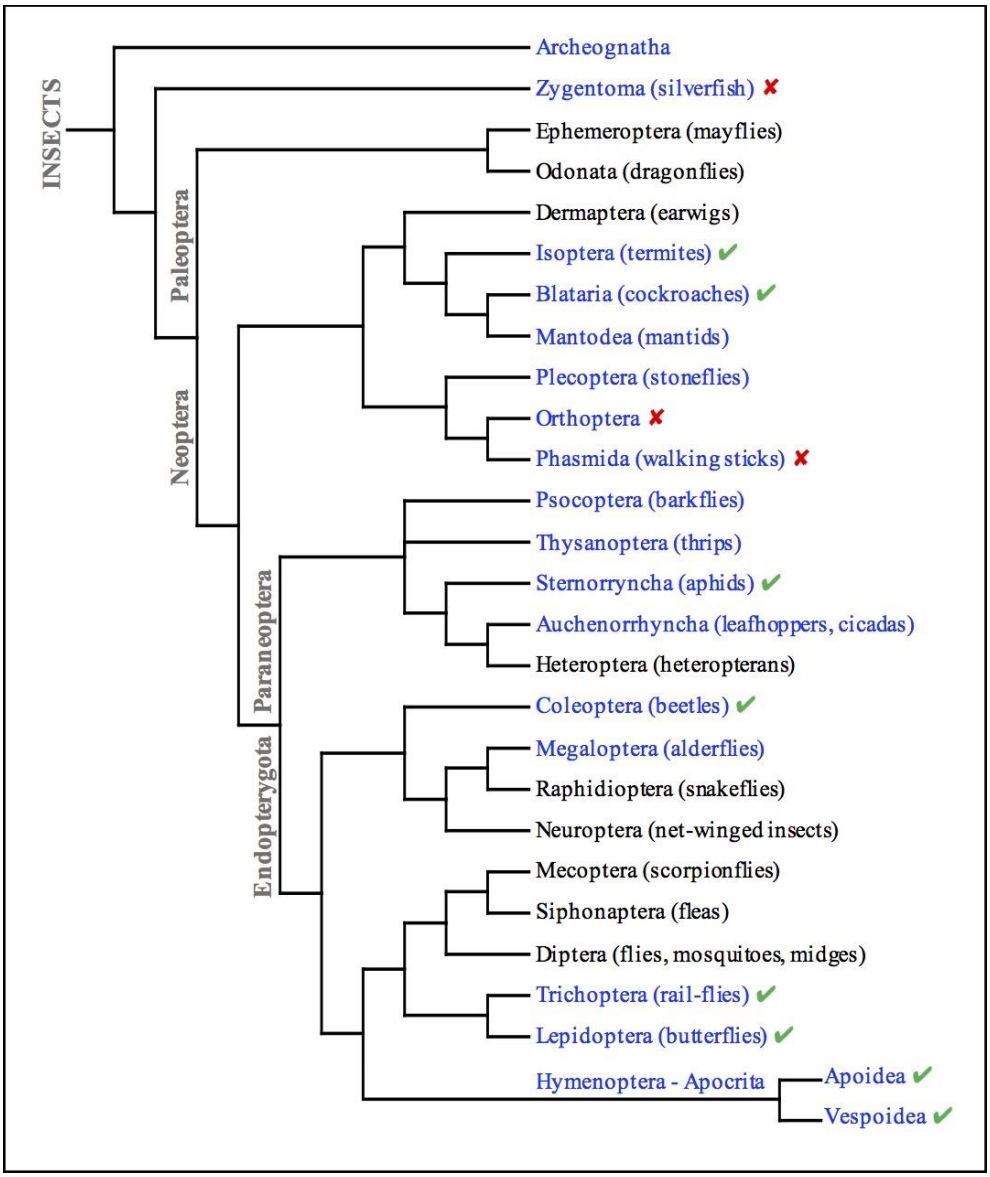


Figure 19: The distribution of $(TTAGG)_n$ telomere sequence and of TTAGG-specific telomerase activity in Insecta are reported. In the insect orders represented in blue the TTAGG repeat motif was found, in the ones evidenced in black it was not retrieved. The orders in which the TTAGG-specific telomerase activity was detected are indicated with (✓), the opposite with (**X**) (modified from Korandová *et al*., 2014).

Also the genus *Bacillus*, that constitutes a good example of reticulate evolution (see Chapter 1), could become an important model group for studies on karyotype structure and evolution.

Aim of this project is to reveal the composition of telomeres in the genus *Bacillus* that is characterized by different reproductive strategies to test if the reproduction can affect it.

This project was possible thanks to the collaboration going on from several years with the Laboratory of Heterochromatin Structure and Function, headed by dr. sc. Miroslav Plohl, of the Department of Molecular Biology-Ruđer Bošković Institute (Zagreb, Croatia). In particular, dr. sc. Brankica Mravinac planned and performed most of the experiments.

Thanks to this project I learned many techniques concerning the characterization of telomeric DNA sequences.

Preliminary analyses showed the presence also in the telomeres of *B. rossius*, *B. atticus* and *B. whitei,* by Southern blot hybridization, of the insect canonical telomeric TTAGG motif while *Bal31* DNA end-degradation assay proved that TTAGG repeats constitute the terminal regions of *Bacillus* chromosomes (Mravinac *et al*., 2013).

I confirmed the above data in the *B. lynceorum* population from Floridia, in south-eastern Sicily (data not shown).

Fluorescence In Situ Hybridization (FISH) experiments on gonads of adult specimens of *B. lynceorum* were performed to validate the chromosomal localization of the pentanucleotide (TTAGG) unit, as explained in Mravinac *et al*. (2013).

Hybridizations produced a signal at the ends of chromosomes (Figure 20) as in the previous analysed species *B. rossius*, *B. atticus* and *B. whitei* (Mravinac *et al*., 2013).

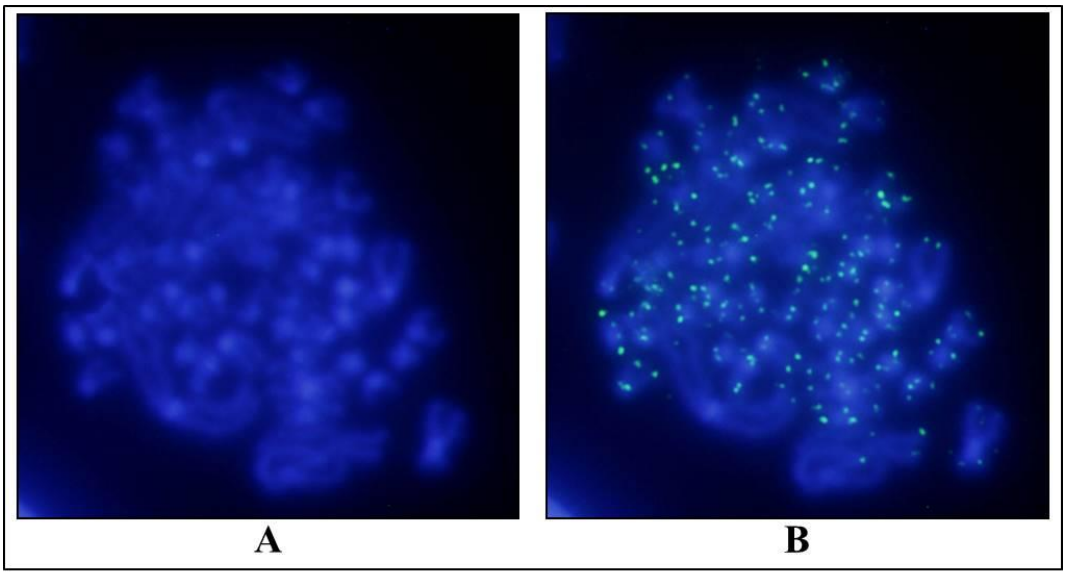


Figure 20: Chromosome localization of TTAGG repeats in *B*. *lynceorum*. Metaphase chromosome plates are counterstained with 4',6-Diamidino-2'-Phenylindole Dihydrochloride (DAPI; blue signal) and are reported either before (A) or after the hybridization with $(TTAGG)_n$ biotin labelled probe (B). The position of $(TTAGG)_n$ clusters is indicated by green spots deriving from the signal immunodetection with fluorescein avidin D (photo by Brankica Mravinac and Livia Bonandin).

It is to be noted that, as far as *B*. *rossius* and the obligatory parthenogenetic hybrid *B*. *whitei* are concerned, a very strong signal was observed in the telomeric region of certain chromosomes. These "mega-telomeres" (Mravinac *et al*., 2013) represented the first characterization of TTAGG mega-telomeres not only within the class Insecta but also in invertebrates (Mravinac *et al*., 2013). As known from the literature, there is the evidence of mega-telomeres only for vertebrate taxa as in the mouse and in the chicken (O'Hare and Delany, 2009).

Thus, to conclude, data here produced and the ones obtained by Mravinac *et al*. (2013) demonstrated that the telomeres of all analysed species of the stick insect genus *Bacillus* are made up of TTAGG motif, which constitutes the canonical telomeric repeat sequence within the class Insecta. Therefore, in the genus *Bacillus*, the presence of the pentanucleotide unit TTAGG is not correlated with the reproductive biology.

The composition of the telomeres in the species *B*. *rossius* and *B*. *whitei* seems characterized by longer arrays of $(TTAGG)_n$ sequences, representing TTAGG mega-telomeres (Mravinac *et al*., 2013). Their presence in the above mentioned species may be explained through the fact that *B*. *whitei* is the interspecific hybrid between *B*. *rossius redtenbacheri* and *B*. *grandii grandii* (Mravinac *et al*., 2013). My results on *B*. *lynceorum* (i.e. the absence of mega-telomeres), though, rise some questions: *B*. *lynceorum* is in fact the triploid hybrid deriving from *B*. *rossius redtenbacheri* X *B*. *grandii grandii* X *B*. *atticus atticus*. It remains therefore to explain the absence of TTAGG mega-telomeres in this taxon.

These results of the thesis were also presented at the 19th International Chromosome Conference, Bologna, 2nd-6th September 2013. Please read the abstract in Appendix.

CHAPTER 5
LITERATURE CITED

- Arkhipova, I.R. (2005) Mobile genetic elements and sexual reproduction. *Cytogenet Genome Res* 110: 372-82.

- Beauregard, A., Curcio, M.J. and Belfort, M. (2008) The take and give between retrotransposable elements and their hosts. *Annu Rev Genet* 42: 587-617.

- Bernadotte, A., Mikhelson, V.M and Spivak, I.M. (2016) Markers of cellular senescence. Telomere shortening as a marker of cellular senescence. *Aging* 8(1):3-11.

- Blackburn, E.H. (2001) Switching and signalling at the telomere. *Cell* 106(6):661-73.

- Bonandin, L., Scavariello, C., Luchetti, A. and Mantovani, B. (2014) Evolutionary dynamics of R2 retroelement and insertion inheritance in the genome of bisexual and parthenogenetic *Bacillus rossius* populations (Insecta, Phasmida). *Insect Molecular Biology* 23(6): 808-820.

- Bonandin, L., Scavariello, C., Mingazzini, V., Luchetti, A. and Mantovani, B. (2016) Obligatory parthenogenesis and TE load: *Bacillus* stick insects and the R2 non-LTR retrotransposon. *Insect Science* doi: 10.1111/1744-7917.12322.

- Bragg, P.E. (1997) A glossary of terms used to describe phasmids. *Phasmid Studies* 6(1):24-33.

- Bull, J.J., Molineux, I.J. and Rice, W.R. (1991) Selection of benevolence in a host- parasite system. *Evolution* 45(4): 875-882.

- Bullini, L. (1966) Spanandria e partenogenesi geografica in *Bacillus rossius* (Rossi). *Atti Acc Naz Lincei*, *Rend Sci Fis Mat Nat* 45: 926-932.

- Burke, W.D., Calalang, C.C. and Eickbush, T.H. (1987) The site-specific ribosomal insertion element type II of *Bombyx mori* (R2Bm) contains the coding sequence for a reverse transcriptase-like enzyme. *Mol Cell Biol* 7(6): 2221-30.

- Burke, W.D., Eickbush, D.G., Xiong, Y., Jakubczak, J. and Eickbush, T.H. (1993) Sequence relationship of retrotransposable elements R1 and R2 within and between divergent insect species. *Mol Biol Evol* 10(1): 163-185.

- Burke, W.D., Muller, F. and Eickbush, T.H. (1995) R4, a non- LTR retrotransposon specific to the large subunit rRNA genes of nematodes. *Nucleic Acids Res* 23: 4628-34.

- Burke, W.D., Singh, D. and Eickbush, T.H. (2003) R5 retrotransposons insert into a family of infrequently transcribed 28S rRNA genes of planaria. *Mol Biol Evol* 20(8): 1260-70.

- Capy, P., Gasperi, G., Bièmont, C. and Bazin, C. (2000) Stress and transposable elements: co-evolution or useful parasites? *Heredity* 85: 101-106.

- Chénais, B., Caruso, A., Hiard, S. AND Casse, N. (2012) The impact of transposable elements on eukaryotic genomes: from genome size increase to genetic adaptation to stressful environments. *Gene* 509(1): 7-15.

- Christensen, S.M., Ye, J. and Eickbush, T.H. (2006) RNA from the 5' end of the R2 retrotransposon controls R2 protein binding to and cleavage of its DNA target site. *Proc Natl Acad Sci USA* 103: 17602-17607.

- Dewannieux, M. and Heidmann, T. (2005) LINEs, SINEs and processed pseudogenes: parasitic strategies for genome modeling. *Cytogenet Genome Res* 110: 35-48.

- Dodsworth, S., Leitch, A.R. and Leitch, I.J. (2015) Genome size diversity in angiosperms and its influence on gene space. *Curr Opin Genet Dev* 35: 73-8.

- Dolgin, E.S. and Charlesworth, B. (2006) The fate of transposable elements in asexual populations. *Genetics* 174(2): 817-27.

- Dolgin, E.S., Charlesworth, B. and Cutter, A.D. (2008) Population frequencies of transposable elements in selfing and outcrossing *Caenorhabditis nematodes*. *Genet Res* 90: 317–329.

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

- Dover, G.A. (1986) Molecular drive in multigene families: how biological novelties arise, spread and are assimilated. *Trends Genet* 2: 159-165.

- Eddy, S.R. (2012) The C-value paradox, junk DNA and ENCODE. *Curr Biology* 22(21): R898-899.

- Eickbush, T.H. (2002) R2 and related site-specific non-LTR retrotransposons. *Mobile DNA II* (Craig, N., Craigie, R., Gellert, M., Lambowitz, A., eds), pp. 813-835. American Society of Microbiology Press, Washington, DC.

- Eickbush, T.H. and Eickbush, D. (2007) Finely orchestrated movements: evolution of the ribosomal RNA genes. *Genetics* 175: 477-485.

- Eickbush, T.H. and Jamburuthugoda, V.K. (2008) The diversity of retrotransposons and the properties of their reverse transcriptases. *Virus Res* 134: 221-234.

- Eickbush, D.G., Ye, J., Zhang, X., Burke, W.D. and Eickbush, T.H. (2008) Epigenetic regulation of retrotransposons within the nucleolus of *Drosophila*. *Mol Cell Biol* 28: 6452-6461.

- Eickbush, D.G. and Eickbush, T.H. (2010) R2 retrotransposons encode a self-cleaving ribozyme for processing from an rRNA cotranscript. *Mol Cell Biol* 30(13): 3142-3150.

- Felsenstein, J. (1974) The evolutionary advantage of recombination. *Genetics* 78: 737-756.

- Frydrychová, R., Grossmann, P., trubač, P., Vítková, M. and Marec, F. (2004) Phylogenetic distribution of TTAGG telomeric repeats in insects. *Genome* 47: 163-178.

- Fujiwara, H. (2015) Site specific non - LTR retrotransposons. *Microbiology spectrum* 3(2):MDNA3-0001-2014.

- Ghesini, S., Luchetti, A., Marini, M. and Mantovani, M. (2011) The Non-LTR retrotransposon R2 in termites (Insecta, Isoptera): Characterization and dynamics. *J Mol Evol* 72: 296-305.

- Gladyshev, E.A. and Arkhipova, I.R. (2009) Rotifer rDNA-specific R9 retrotransposable elements generate an exceptionally target site duplication upon insertion. *Gene* 448: 145-50.

- Gladyshev, E.A. and Arkhipova, I.R. (2010a) Subtelomeric non-LTR retrotransposon *Hebe* in the bdelloid rotifer *Adineta vaga* is subject to inactivation by deletions but not 5' truncations. *Mobile DNA* 1: 12.

- Gladyshev, E.A. and Arkhipova, I.R. (2010b) Genome structure of bdelloid rotifers: shaped by asexuality or desiccation? *Journal of Heredity* 2010:101(Supplement 1):S85–S93.

- Hamilton, W.D. (1975) Gamblers since life began: barnacles, aphids, elms (a review). *Quart Rev Biol* 50: 175-180.

- Hickey, D.A. (1982) Selfish DNA: a sexually-transmitted nuclear parasite. *Genetics* 101: 519-531.

- Joffe, B.I., Solovei, I.V. and Macgregor, H.C. (1998) Ordered arrangement and rearrangement of chromosomes during spermatogenesis in two species of planarians (Plathelminthes). *Chromosoma* 107(3): 173-183.

- Kagramanova, A.S. Kapelinskaya, T.V., Korolev, A.L. and Mukha, D.V. (2006) R1 and R2 retrotransposons of German cockroach *Blatella germanica*: A comparative study of 5'-truncated copies integrated into the genome. *Mol Biol* 41(4): 546-553.

- Kajikawa, M. and Okada, N. (2002) LINEs mobilize SINEs in the eel through a shared 3′sequence. *Cell* 111: 433-444.

- Kapitonov, V.V., Tempel, S. and Jurka, J. (2009) Simple and fast classification of non-LTR retrotransposons based on phylogeny of their RT domain protein sequences. *Gene* 448: 207-213.

- Keeling, P.J. and Palmer, J.D. (2008) Horizontal gene transfer in eukaryotic evolution. *Nat Rev Genet* 9(8): 605-18.

- Kojima, K.K. and Fujiwara, H. (2003) Evolution of target specificity in R1 clade non-LTR retrotransposons. *Mol Biol Evol* 20(3): 351-61.

- Kojima, K.K. and Fujiwara, H. (2005) Long-term inheritance of the 28S rDNA-specific retrotransposon R2. *Mol Biol Evol* 22: 2157-2165.

- Kojima, K.K., Kuma, K., Toh, H. and Fujiwara, H. (2006). Identification of rDNA-specific non-LTR retrotransposons in Cnidaria. *Mol Biol Evol* 23: 1984-93.

- Korandová, M., Krůček, T., Vrbová, K., Frydrychová, R.C. (2014) Distribution of TTAGG-specific telomerase activity in insects. *Chromosome Res* 22: 495-503.

- Kordis, D. and Gubensek, F. (1999) Horizontal transfer of non-LTR retrotransposons in vertebrates. *Genetica* 107(1-3): 121-8.

- Kuznetsova, V.G., Grozeva, S.M. and Anokhin, B.A. (2012) The first finding of (TTAGG)n telomeric repeat in chromosomes of true bugs (Heteroptera, Belostomatidae). *Comparative Cytogenet* 6(4): 341-346.

- Lampe, D.J., Witherspoon, D.J., Soto-Adames, F.N. and Robertson, H.M. (2003) Recent horizontal transfer of mellifera subfamily mariner transposons into insect lineages representing four different orders shows that selection acts only during horizontal transfer. *Mol Biol Evol* 20(4): 554-62.

- Lee, Y.C. and Langley, C.H. (2010) Transposable elements in natural populations of *Drosophila melanogaster*. *Philos Trans R Soc Lond B Biol Sci* 365(1544): 1219-28.

- Levin, H.L. and Moran, J.V. (2011) Dynamic interactions between transposable elements and their hosts. *Nat Rev Genet* 12(9): 615-627.

- Lockton, S. and Gaut, B.S. (2010) The evolution of transposable elements in natural populations of self-fertilizing *Arabidopsis thaliana* and its outcrossing relative *Arabidopsis lyrata*. *BMC Evol Biol* 10: 10.

- Loewe, L. and Lamatsch, D.K. (2008) Quantifying the threat of extinction from Muller's ratchet in the diploid Amazon molly (*Poecilia formosa*). *BMC Evol Biol* 19(8): 88.

- Luchetti, A. and Mantovani, B. (2013) Non-LTR R2 element evolutionary patterns: phylogenetic incongruences, rapid radiation and the maintenance of multiple lineages. *PLoS ONE* 8: e57076.

- Malik, H.S., Burke, W.D. and Eickbush, T.H. (1999) The age and evolution of non-LTR retrotransposable elements. *Mol Biol Evol* 16: 793-805.

- Mantovani, B., Scali, V. and Tinti, F. (1991) Allozyme analysis and phyletic relationship of two new stick-insects from north-west Sicily: *Bacillus grandi benazzii* and *B. rossius-grandii benazzii* (Insecta, Phasmatodea). *J Evol Bio* 4: 279-290.

- Mantovani, B. and Scali, V. (1992) Hybridogenesis and androgenesis in the stick insect *Bacillus rossius–grandii benazzii* (Insecta, Phasmatodea). *Evolution* 46: 783-796.

- Mantovani, B., Scali, V. and Tinti, F. (1992) New morphological and allozymic characterization of *Bacillus whitei* and *B. lynceorum* hybrid complexes (Insecta, Phasmatodea) *Biol Zen* Bl 111: 87-103.

- Mantovani, B. and Scali, V. (1993) Genetic structure and phyletic relationships of eastern Mediterranean *Bacillus atticus* Brunner (Insecta Phasmatodea): A biochemical study. *Biochemical Genetics* 31: 343-362.

- Mantovani, B. (1998) Satellite sequence turnover in parthenogenetic systems: the apomictic triploid hybrid *Bacillus lynceorum* (Insecta, Phasmatodea). *Mol Biol Evol* 15: 1288-1297.

- Mantovani, B., Passamonti, M. and Scali, V. (1999) Genomic evolution in parental and hybrid taxa of the genus *Bacillus* (Insecta, Phasmatodea). *It J Zool* 66: 265-72.

- Mantovani, B., Passamonti, M. and Scali, V. (2001) The mitochondrial cytocrome oxidase II gene in *Bacillus* stick insect: ancestry of hybrids, androgenesis, and phylogenetic relationships. *Mol Phylogenet Evol* 19(1): 157-63.

- Mason, J.M., Frydrychova, R.C. and Biessmann, H. (2008) *Drosophila* telomeres: an exception providing new insights. *Bioessays* 30: 25–37.

- McClintock, B. (1950) The origin and behavior of mutable loci in maize. *Proc Natl Acad Sci USA* 36: 344-355.

- Mingazzini, V. Luchetti, A. and Mantovani, B. (2010) R2 dynamics in *Triops cancriformis* (Bosc, 1801) (Crustacea, Branchiopoda, Notostraca): turnover rate and 28S concerted evolution. *Heredity* 106: 567-575.

- Mingazzini, V. (2011) Transposable elements: structure and dynamic of the autonomous retrotransposon R2 in Arthropoda with non-canonical reproduction. Ph.D. thesis, University of Bologna.

- Mora, P., Vela, J., Sanllorente, O., Palomeque, T. and Lorite, P. (2015) Molecular cytogenetic studies in the ladybird beetle *Henosepilachna argus* Geoffroy, 1762 (Coleoptera, Coccinellidae, Epilachninae). *Comparative Cytogenet* 9(3): 423-434.

- Moran, J.V. and Gilbert, N. (2002) Mammalian LINE-1 retrotransposons and related elements. *Mobile DNA II* (Craig, N., Craigie, R., Gellert, M., Lambowitz, A., eds), pp. 836-869. American Society of Microbiology Press, Washington, DC.

- Mravinac, B., Městrović, Čavrak, V.V. and Plohl, M. (2011) TCAGG, an alternative telomeric sequence in insects. *Chromosoma* 120: 367-376.

- Mravinac, B., Bonandin, L., Scavariello, C., Luchetti, A., Plohl, M. and Mantovani, B. (2013) TTAGG mega-telomeres in *Bacillus* stick-insects (Insecta: Phasmida). Book of abstracts 19th International Chromosome Conference, Bologna, 2nd-6th September 2013, pp. 34-34.

- Mukha, D.V., Mysina, V., Mavropulo, V. and Schal, C. (2011) Structure and molecular evolution of the ribosomal DNA external transcribed spacer in the cockroach genus *Blattella*. *Genome* 54: 222-234.

- Müller, F., Wicky, C., Spicher A. and Tobler, H. (1991) New telomere formation after developmentally regulated chromosomal breakage during the process of chromatin diminution in *Ascaris lumbricoides*. *Cell* 67: 815-822.

- Nei, M. and Rooney, A.P. (2005) Concerted and birth-and-death evolution of multigene families. *Annu Rev Genet* 39: 121-152.

- Nielsen, L. and Edström, J.E. (1993) Complex telomere-associated repeat units in members of the genus *Chironus* evolve from sequences similar to simple telomeric repeats. *Mol Cell Biol* 13(3): 1583-1589.

- Nuzhdin, S.V. and Petrov, D.A. (2003) Transposable elements in clonal lineages: lethal hangover from sex. *Biol J Linn Soc* 79: 33-41.

- O'Hare, T.H. and Delany, M.E. (2009) Genetic variation exists for telomeric array organization within and among the genomes of normal, immortalized and transformed chicken systems. *Chromosome Res* 17: 947-964.

- Pérez-González, C.E. and Eickbush, T.H. (2001) Dynamics of R1 and R2 elements in the rDNA locus of *Drosophila simulans*. *Genetics* 158: 1557-1567.

- Pérez-González, C.E., Burke, W.D. and Eickbush, T.H. (2003) R1 and R2 retrotransposition and deletion in the rDNA loci on the X and Y chromosomes of *Drosophila melanogaster*. *Genetics* 165: 675-685.

- Plohl, M., Prats, E., Martínez-Lage, A., Ana González-Tizón, A., Méndez, J. and Cornudella, L. (2002) Telomeric Localization of the Vertebrate-type Hexamer Repeat, (TTAGGG)n, in the Wedgeshell Clam *Donax trunculus* and Other Marine Invertebrate Genomes. *J Biol Chem* 277(22): 19839-19846.

- Plohl, M., Luchetti, A., Meštrovic′, N. and Mantovani, B. (2008) Satellite DNAs between selfishness and functionality: structure, genomics and evolution of tandem repeats in centromeric (hetero)chromatin. *Gene* 409: 72-82.

- Ricci, M., Luchetti, A., Bonandin, L. and Mantovani, B. (2013) Random DNA libraries from three species of the stick-insect genus *Bacillus* (Insecta: Phasmida): repetitive DNA characterization and first observation of polyneopteran MITEs. *Genome* 56(12): 729-735.

- Ricci, M. (2014) Transposable elements dynamics in taxa with different reproductive strategies or speciation rate. Ph.D. thesis, University of Bologna.

- Richard, G.F., Kerrest, A. and Dujon, B. (2008) Comparative genomics and molecular dynamics of DNA repeats in eukaryotes. *Microbiol Mol Biol Rev* 72: 686-727.

- Roy-Engel, A.M. (2012) A tale of an A-tail: the lifeline of SINE. *Mob Genet Elements* 2: 282-286.

- Sahara, K., Marec, F. and Traut, W. (1999). TTAGG telomeric repeats in chromosomes of some insects and other arthropods. *Chromosome Res* 7(6): 449-60.

- Sánchez-Gracia, A., Maside, X. and Charlesworth, B. (2005) High rate of horizontal transfer of transposable elements in *Drosophila*. *Trends Genet* 21(4): 200-3.

- Scali, V. and Mantovani, B. (1989) Updating of systematics and speciation mechanisms of *Bacillus* (Insecta, Phasmatodea). *Boll Zool* 56: 87-98.

- Scali, V., Tinti, F., Mantovani, B. and Marescalchi, O. (1995) Mate recognition and gamete cytology features allow hybrid species production and evolution in *Bacillus* stick insects. *Boll Zool* 62: 59-70.

- Scali, V., Passamonti, M., Marescalchi, O. and Mantovani, B. (2003) Linkage between sexual and asexual lineages: genome evolution in *Bacillus* sticks insects. *Biol J Linn Soc* 79: 137-50.

- Scali, V. (2009) *Lost sex: the evolutionary biology of parthenogenesis* (Schön, I., Martens, K., van Dijk P., eds), pp. 317-345. Metasexual stick insects: model pathways to losing sex and bringing it back. Springer, Heidelberg.

- Scavariello, C. (2016) Transposable elements in arthropods genomes with non-canonical reproductive strategies. Ph.D. thesis, University of Bologna.

- Schaack, S., Pritham, E.J., Wolf, A. and Lynch, M. (2010) DNA transposon dynamics in populations of *Daphnia pulex* with and without sex. *Proceedings Royal Society B* 277: 2381-2387.

- Schultz, R.J. (1961) Reproductive mechanism of unisexual and bisexual strains of the viviparous fish *Poeciliopsis*. *Evolution*, 15: 302-325.

- Teschke, C., Solleder, G. and Motitz, K.B. (1991) The highly variable pentameric repeats of the AT-rich germline limited DNA in *Parascaris univalens* are the telomeric repeats of somatic chromosomes. *Oxford University Press* 19(10): 2677-2684.

- Tinti, F. (1993) Morphological, genetic and chromosomal characterization of Corsican and Spanish *Bacillus rossius* (Insecta, Phasmatodea). *Vie Milieu* 43: 109-117.

- Tram, U. and Sullivan, W. (2000) Reciprocal inheritance of centrosomes in the parthenogenetic Hymenopteran *Nasonia vitripennis*. *Curr Biol* 10: 1413-1419.

- Valizadeh, P. and Crease, T.J. (2008) The association between breeding system and transposable element dynamics in *Daphnia pulex*. *J Mol Evol* 66: 643-654.

- Wright, S. and Finnegan, D. (2001) Genome evolution: sex and transposable elements. *Curr Biol*, 11: R296-R299.

- Xiong, Y. and Eickbush, T. H. (1988) Similarity of reverse transcriptase-like sequences of viruses, transposable elements, and mitochondrial introns. *Mol Biol Evol* 5(6): 675-90.

- Zakian V.A. (2012) Telomeres: The beginnings and the ends of eukaryotic chromosomes. *Exp Cell Res* 318(12): 1456-1460.

- Zhang, X. and Eickbush, T.H. (2005) Characterization of active R2 retrotransposition in the rDNA locus of *Drosophila simulans*. *Genetics* 170: 195-205.

- Zhang, X. Zhou, J. and Eickbush, T.H. (2008) Rapid R2 retrotransposition leads to the loss of previously inserted copies via large deletions of the rDNA locus. *Mol Biol Evol* 25: 229-237.

- Zhou, J. and Eickbush, T.H. (2009) The pattern of R2 retrotransposition activity in natural populations of *Drosophila simulans* reflects the dynamic nature of the rDNA locus. *PLoS Genetics* 5: 1-11.

APPENDIX

This appendix includes the abstracts of the Congress attendances in which the results of this thesis were presented.

1. Scavariello, C., Bonandin, L., Luchetti, A. and Mantovani, B. (2013) Non-LTR retrotransposon R2 molecular characterization and activity in *Bacillus rossius* (Phasmida, Bacillidae). Book of abstracts 5th Congress of the Italian Society for Evolutionary Biology, Trento, 28th-31st August 2013, pp. 57-58;

2. Mravinac, B., Bonandin, L., Scavariello, C., Luchetti, A., Plohl, M. and Mantovani, B. (2013) TTAGG mega-telomeres in *Bacillus* stick-insects (Insecta: Phasmida). Book of abstracts 19th International Chromosome Conference, Bologna, 2nd-6th September 2013, pp. 34-34;

3. Bonandin, L., Scavariello, C., Luchetti, A. and Mantovani, B. (2013) Caratterizzazione molecolare e dinamica del retrotrasposone non-LTR R2 nella specie *Bacillus atticus* (Phasmida, Bacillidae). Atti 74° Congresso Nazionale dell'Unione Zoologica Italiana, Modena, 30th September-3rd October 2013, pp. 122-122;

4. Bonandin, L., Luchetti, A. and Mantovani, B. (2015) Reproductive biology *versus* transposable elements load: the role of host reproductive strategy in the study of R2 dynamics in *Bacillus* stick insects (Phasmida, Bacillidae). Book of abstracts 6th Congress of the Italian Society for Evolutionary Biology, Bologna, 31st August-3rd September 2015, pp. 4-4.

Please also read the full abstracts attached below to have a more detailed view of the projects.

NEW PhDs ON THE BLOCKS

The symposium benefits from the collaboration with FIRS>T, the PhD programme of FEM. This symposium is reserved for young scientists who would like to talk about their research, their results (no matter at which stage) and their passion towards evolutionary biology in a friendly atmosphere.

Chairs: Lino Ometto and Alessandro Gretter (Fondazione Edmund Mach)

Genetic variability in the promoter of miR397 in *Picea abies*

Luisa Bresadola[1], Mingai Li[1], Silvio Schueler[2], Caroline Heiri[3], Lorenzo Bonosi[1], Claudio Varotto[1]
[1]Research and Innovation Centre, Fondazione Edmund Mach, Via Mach 1, 38010 San Michele all'Adige (TN), Italy.
[2]Bundesforschungszentrum für Wald, Hauptstrasse, 7 A-1140 Vienna, Austria
[3]Swiss Federal Institute for Forest, Snow and Landscape Research, Zuercherstrasse, 111 8903 Birmensdorf, Switzerland
Contact: luisa.bresadola@fmach.it

Norway spruce (*Picea abies* Karst.) is a tree species that belongs to conifers, a taxon that is extremely important both from an ecological and an economical point of view. Being perennial, this species has often to face suboptimal environmental conditions and to adapt to them: microRNAs are a fundamental class of regulatory molecules, often involved in stress responses and therefore, potentially very important for plant adaptive processes. The focus of this study is on miR397: in *Arabidopsis thaliana* miR397 was shown to be involved in the regulation of copper homeostasis and of the transcription of laccases. These enzymes operate during lignin biosynthesis and therefore their regulation is really important in woody plants like Norway spruce, in order to react to mechanical stress and to resist to the attack of pathogens. In order to understand miR397 regulatory mechanisms, its promoter was isolated in this species and putative regulatory elements were identified. This region, together with the microRNA stem-loop region, was sequenced in seeds produced by individuals originating from different alpine populations in Italy, Austria and Switzerland. In the mature miR397, the most important part for microRNA regulatory function, no polymorphism was found in the analyzed samples. This result suggests that purifying selection is probably acting on this sequence in order to preserve microRNA functionality. As regards the promoter region, several single nucleotide polymorphisms (SNPs) and some insertions/deletions were identified. Some of them are located in the putative regulatory elements, therefore they are good candidates to test if they influence the microRNA expression level and if they have consequences on phenotype that can be relevant in the process of adaptation. This will provide deeper insights into the adaptive role of microRNAs in Conifers.

,

Non-LTR retrotrasposon R2 molecular characterization and activity in *Bacillus rossius* (Phasmida, Bacillidae)

Claudia Scavariello[1], Livia Bonandin[1], Andrea Luchetti[1], Barbara Mantovani[1]
[1]Università degli Studi di Bologna, Via Selmi, 3 Bologna, Italy.
Contact: claudia.scavariello2@unibo.it

The non-LTR retrotransposon R2 is one of the most analyzed transposable elements (TEs), its presence being recovered from diploblastic organisms to lower vertebrates. It inserts in the sequence 5'-TTAA↓GGTAGC-3' of the 28S ribosomal gene, thus affecting the production of functional rRNAs. The evolutionary relationship between retrotransposon activity and reproductive biology of the host species is still debated: while some studies suggest that genomes with limited effective recombination (unisexuals and asexuals) accumulate TEs with a low capacity to eliminate them, gonochoric organisms better manage their proliferation (Muller's ratchet). In order to go through this issue, we are studying R2 distribution and dynamics in the facultative parthenogenetic stick-insect *Bacillus rossius*. In Italy, *B. rossius rossius*, spreading along the Western peninsular coasts and in North-Western Sardinia, and *B. rossius redtenbacheri*, distributed along the peninsular eastern coasts, in Sicily and in South-Eastern Sardinia, occur with gonochoric and unisexual (parthenogenetic) populations. The R2 complete sequence was PCR amplified and sequenced from gonochoric populations of *B. r. rossius* from Capalbio (Tuscany) and of *B. r. redtenbacheri* from Patti (Sicily). The R2 activity was studied through the 5' end deletions analysis in selected parental individuals and in a sample (10-20 individuals) of their offspring. In particular we analyzed the progeny of two females each of the parthenogenetic *B. r. redtenbacheri* populations from

Curcuraci, Massa San Nicola and Castanea delle Furie (Sicily) and the male progeny of crosses involving parental females either from parthenogenetic populations (*B. r. redtenbacheri* Curcuraci ♀ X *B. r. rossius* Anzio ♂; two crosses) or from *B. r. rossius* gonochoric ones (Anzio ♀ X Anzio ♂, one cross; Capalbio ♀ X Capalbio ♂, two crosses). The R2 complete sequence in the Capalbio population is 3,515 bp long and has an A+T content equal to 47%. The ORF is 3,165 bp long and encodes for 1,054 amino acids. The R2 complete sequence in the Patti population is 3,717 bp long and exhibits an A+T content equal to 49%. The ORF is 3,176 bp long, but it is degenerated having several stop codons and a frameshift mutation. Therefore, for a selection of parthenogenetic and gonochoric populations of both subspecies, a R2 sequence survey has been performed to verify the ORF structure. From this analysis, it emerged that *B. r. rossius* populations have R2 elements with an intact ORF, *B. r. redtenbacheri* Sicilian populations have elements with a degenerated ORF and *B. r. redtenbacheri* peninsular populations have both kinds of elements without any relationship to the reproductive strategy. The insertional activity analyses showed new insertions in the thelytokous offspring of parthenogenetic females and in the amphimictic progeny of crosses (1-6 and 1-8, respectively), indicating that the element is actively transposing. However, no new insertions have been observed in the offspring of crosses between parthenogenetic females and gonochoric males. The deletion of truncated variants was also detected: one event in the progeny of a parthenogenetic female (Massa San Nicola) and one event in the progeny of the cross Anzio ♀ X Anzio ♂. Contrary to Muller's ratchet expectations, this pattern of insertion indicates that genomic turnover mechanisms seem active even in parthenogenetic genomes.

Effects of cryptic genetic variation on innovability in *Escherichia coli*

Emanuele Rigato[1], Giulio Bertoloni[2], Giuseppe Fusco[1]
[1]Department of Biology, University of Padova, Via U. Bassi 58/B, Italy.
[2]Histology, Microbiology and Medical Biotechnologies Department, University of Padova, Via A. Gabelli 63, 35121 Padova
Contact: leistus9@yahoo.it

Widespread phenotype resistance to genetic mutation ('phenotype robustness') allows accumulation of neutral genetic variation. This variation, called "cryptic genetic variation" (CGV), has no effects on phenotypes in a particular genetic or environmental context, but can become phenotypically expressed as a consequence of genetic mutations or environmental changes. Thus CGV might have a key role in the kinetic of adaptation. Since evolutionary adaptation by natural selection requires phenotypic variation, the fraction of variation emerging from CGV could enhance innovability (sensu Wagner). This effect has been demonstrated with theoretical studies, using computational models, and with experimental studies on ribozymes. However, this has not been investigated yet in more complex evolving systems, as in whole organisms, where this phenomenon depends on a number of little known parameters of the genotype→phenotype map, such as the amount of epistasis, pleiotropy and neutrality. The aim of the present study is to test for the role of CGV in the adaptative performances of a whole organismal system, the prokaryote *Escherichia coli*, in exploiting different carbon sources (glycerol and lactate). The experimental design consists in generating populations with different amounts of CGV by subjecting the same genotype (BW30270) to EMS and UV light and subsequently exposing the clones to stabilizing selection in the context of their native carbon source (glucose). These populations with identical phenotypes, but different amount of CGV, are then tested for hundreds of generations for Darwinian adaptation in metabolizing new carbon sources. The kinetic of adaptation is measured and interpreted with appropriate models. Here, some preliminary results of the ongoing experiments are presented.

Towards a phylogeny of the geophilomorph centipede genus Stenotaenia

Laura Del Latte[1], Francesca Bortolin[1], Lucio Bonato[1], Giuseppe Fusco[1]
1Department of Biology, University of Padova, Via Ugo Bassi 58b, 35131 Padova, Italy.
Contact: laura.del.latte@gmail.com

In the context of a research investigating possible correlations between segmental modularity and developmental stability in geophilomorph centipedes, this study aims at reconstructing the phylogenetic relationships within the genus Stenotaenia. This will provide a basis for a forthcoming comparative analysis aimed at testing the existence of a trade-off between the number of body segments and the precision of their phenotypic expression. Stenotaenia is manly spread in the central-eastern part of the Mediterranean region. It exhibits high interspecific variability in the number of trunk segments and adult body size, while other morphological characters show very little variation. Presently, about five different morphospecies can be recognized, but the taxonomy of the genus is not well established. Many species are sampled only rarely and collection specimens suitable for DNA extraction are infrequent. We managed to sample a dozen populations from throughout the distribution range and representative of different morphospecies. Ongoing phylogenetic

TTAGG MEGA-TELOMERES IN *Bacillus* STICK INSECTS (INSECTA: PHASMIDA)

MRAVINAC B.[1], BONANDIN L.[2], SCAVARIELLO C.[2], LUCHETTI A.[2], PLOHL M.[1], MANTOVANI B.[2]
[1]*Divison of Molecular Biology, Ruđer Bošković Institute, Zagreb, Croatia;*
[2]*Dipartimento di Scienze Biologiche, Geologiche e Ambientali, Università di Bologna, Bologna, Italy*

The species of the genus *Bacillus*, with their complex genealogy and different reproductive strategies, represent a good model group to study karyotype evolution as well as chromosome structure itself. In order to reveal the composition of *Bacillus* telomeres, we analyzed the facultative partenogenetic species *Bacillus rossius* (with gonochoric and all-female populations) and the obligatory unisexual species *B. atticus* and *B. whitei*. By Southern blot hybridization we evidenced the presence of the insect canonical telomeric TTAGG motif in the genomes of all tested species. Bal31 DNA end-degradation assay proved that TTAGG repeats constitute the terminal regions of *Bacillus* chromosomes. In addition to the chromosome end positions, Bal31 trimming approach also disclosed distally located TTAGG arrays in *B. rossius* and *B. whitei*. The chromosomal distribution of TTAGG repeats was analyzed by fluorescence *in situ* hybridization (FISH), which authenticated the position of TTAGG sequence at the end of all chromosomes in the three species. Interestingly, FISH analysis showed that the telomeres of certain chromosomes of *B. rossius* and *B. whitei* are composed of extremely long TTAGG arrays, providing the first evidence of mega-telomeres not only in insects, but also in invertebrates.
As the species *B. whitei* is the interspecific hybrid between *B. rossius* and *B. grandii grandii*, TTAGG mega-telomeres might represent a chromosomal trait inherited through chromosome complement assembly. In our ongoing studies, we are addressing the evolutionary origin of *Bacillus* mega-telomeres, their persistence in different species and populations, and the mechanisms of their maintenance.
This work has been supported by grants from the Ministry of Science, Education and Sports of the Republic Croatia and Canziani funds.

CARATTERIZZAZIONE MOLECOLARE E DINAMICA DEL RETROTRASPOSONE NON-LTR R2 NELLA SPECIE *Bacillus atticus* (PHASMIDA, BACILLIDAE)

LIVIA BONANDIN, CLAUDIA SCAVARIELLO, ANDREA LUCHETTI, BARBARA MANTOVANI
Dipartimento di Scienze Biologiche, Geologiche e Ambientali (BiGeA), Università degli Studi di Bologna

La famiglia di elementi mobili ad inserzione specifica attualmente più studiata è rappresentata dal retrotrasposone autonomo non-LTR R2, le cui copie s'inseriscono nella sequenza 5'-TTAA↓GGTAGC-3' del gene ribosomale 28S, compromettendone la produzione di copie funzionali.
Al fine di studiare la dinamica evolutiva e la capacità di trasposizione in un genoma a trasmissione unisessuata, è stata condotta la caratterizzazione molecolare dell'elemento R2 nella specie partenogenetica obbligata di insetto stecco *Bacillus atticus*.
La sequenza completa di R2 è stata ottenuta mediante la tecnica del *primer walking*. Il numero di copie dei geni ribosomali e dell'elemento R2 è stato valutato con metodiche standard di *dot blot*, mentre l'attività di trasposizione è stata studiata sui profili di inserzione attraverso l'analisi delle varianti delete al 5' attraverso *Southern blot*. Lo studio è stato condotto su tre femmine parentali e sulle loro discendenze (10 individui ciascuna) di due popolazioni siciliane (Scoglitti e Necropoli Camarina).
L'elemento R2 in *B. atticus* ha una lunghezza di 3507 bp; la regione 5'UTR (171 bp) precede una ORF di 3177 bp, codificante una proteina di 1058 aminoacidi; la regione 3'UTR inclusa (159 bp) termina con una coda di poli-(A) caratteristica dei retrotrasposoni. La comparsa di nuove inserzioni di varianti delete al 5' (da 2 a 7) nelle discendenze rispetto alle due femmine parentali di Scoglitti e di 3 nuove inserzioni nella progenie della femmina parentale di Necropoli Camarina, mostra che l'elemento R2 è attivo. In quest'ultima discendenza sono stati anche osservati 2 eventi di delezione completa ad indicare che, contrariamente a quanto atteso secondo l'ipotesi *Muller's ratchet* i meccanismi di *turn-over* genomico sembrano attivi anche nei genomi partenogenetici.
I dati sono parzialmente in accordo con quelli precedentemente ottenuti su popolazioni unisessuate della specie partenogenetica facoltativa *B. rossius*; di particolare importanza sarà il confronto con i dati in via di ottenimento sulla specie gonocorica affine, *B. grandii*.

Reproductive biology *versus* transposable elements load: the role of host reproductive strategy in the study of R2 dynamics in *Bacillus* stick insects (Phasmida, Bacillidae)

Bonandin L., Luchetti A., Mantovani B.

Dipartimento di Scienze Biologiche, Geologiche ed Ambientali (BiGeA), Università di Bologna, Italy.

Transposable elements (TEs) are known to promote genome evolution, even if they are often defined as “genomic parasites”. According to the Red Queen hypothesis, conflicts between parasite TEs and the host genome are at equilibrium through strict competitive efforts. Host reproductive strategy is significantly involved in this dynamics. The R2 non-LTR retrotransposon activity was estimated, through insertions display analyses, in populations of the strictly gonochoric *Bacillus grandii maretimi,* the facultative parthenogenetic *B. rossius* and the obligatory parthenogenetic *B. atticus*. R2 activity was also evaluated in the progeny of parthenogenetic isolates of *B. rossius* and *B. atticus*, and in the male progeny of crosses between gonochoric individuals of *B. rossius*. Gonochoric populations showed higher R2 loads than the parthenogenetic ones, the lowest value being scored in *B. atticus*. In all samples low frequency insertions are the majority, with the only exception of a bisexual *B. rossius* population. R2 load was similar in unisexual and bisexual *B. rossius* offspring but lower in *B. atticus* progeny. Moreover, in *B. rossius* unisexual offspring high-frequency insertions were the most represented. Data on facultative unisexuals evidenced a low R2 elimination rate, with recombination having a major role, although a bisexual-like insertion profile emerge at the population level. In this regard, natural selection seems to predominate in bisexuals. Stick-insects obligatory parthenogenesis seems to allow only a very low R2 load, with possible burst of retrotransposition buffered at the population level. Our data agree in indicating TE dynamics deeply linked with host reproductive strategies.