



Università Politecnica delle Marche

Ph.D. School in Science

Curriculum Biomolecular Sciences

PhD Thesis:

**Molecular and cytogenetic characterisation  
of repetitive DNAs in squamate reptiles (Lacertidae)**

PhD candidate:

Dr. Paola Nisi Cerioni

Academic tutor:

Prof. Vincenzo Caputo Barucchi

Cycle XIV

2012-2015

# Index

## INTRODUCTION

<b>1. Tandem Repeats – Definitions and Characteristics</b> .....	1
1.2 Satellite DNA (satDNA) .....	3
1.3 Evolutionary dynamics of satellite DNA .....	4
1.4 Concerted evolution of satellite DNA .....	6
1.5 Functional potential of satellite DNAs .....	9
<b>2. Sex Determining Mechanisms in Vertebrates</b> .....	11
2.1 Distribution of sex determining mechanisms among vertebrate taxa .....	11
2.2 Origin sex chromosomes .....	13
2.3 Evolution of sex chromosomes in squamate reptiles .....	16
<b>3. The studied species</b> .....	19
3.1 The family Lacertidae (Reptilia, Squamata) .....	19
3.2 The main groupings of lacertids .....	20
3.3 Historical biogeography of Lacertidae .....	22
3.4 The Tribe Lacertini Opper, 1811 .....	23
3.5 Present assessments of relationships within the Lacertini .....	29
3.6 The genus Iberolacerta Arribas, 1997 .....	31
3.7 The genera Lacerta (Linneus, 1758) .....	33
<b>4. References</b> .....	37
<b>AIM OF RESEARCH</b> .....	54

<b>Chapter I.....</b>	<b>56</b>
<b>Isolation and Characterization of Two Satellite DNAs in Some Iberian Rock Lizards (Squamata, Lacertidae).</b>	
<b>Chapter II.....</b>	<b>71</b>
<b>Evolutionary dynamics of two satellite DNA families in rock lizards of the genus Iberolacerta (Squamata, Lacertidae): different histories but common traits.</b>	
<b>Chapter III: .....</b>	<b>93</b>
<b>Characterisation of a satellite DNA involved in the W chromosome differentiation in the genus <i>Lacerta</i> Linnaeus, 1758 (Reptilia, Lacertidae).</b>	
<b>CONCLUSIONS .....</b>	<b>120</b>

# ***INTRODUCTION***



# Introduction

## 1. Tandem Repeats – Definitions and Characteristics

Repetitive DNA sequences represent a substantial portion of the eukaryote genome (Dover, 1982; John and Miklos, 1988). In the 1960s, scientists identified these repetitive elements as the explanation for the negative correlation between an organism's phenotypic complexity and its genome size (Hartl, 2000). For instance, repeats constitute almost 46% of the entire human genome and prokaryotic genomes contain roughly 10% repetitive regions, a significant amount considering their small sizes (Van Belkum et al., 1998).

However, since no protein coding function could be primarily associated with repetitive DNAs, early hypothesis considered them as useless genomic elements accumulated as junk (Ohno, 1972), or alternatively, as sequences that represent genomic parasites proliferating for their own sake (Orgel and Crick, 1980). In many cases these sequences seem to be maintained solely by their ability to replicate within the genome (the 'selfish DNA' hypothesis) (Doolittle and Sapienza, 1980; Orgel and Crick, 1980). Far from conferring benefits, their behaviour can sometimes result in a fitness loss to the host (Mackay, 1986). Some human genetic diseases are known to be caused in this way, including mutations due to insertions of transposable element (Wallace et al., 1991 and Holmes et al., 1994), to chromosomal rearrangements induced by recombination between repeated sequences (Lakich et al., 1993), or to the amplification of microsatellite sequences (Kuhl et al., 1993). It has often been proposed that repetitive sequences are functionally important for the host organism or are maintained because their mutagenic activities contribute to long-term evolutionary potential of the population. But these may be consequences rather than causes of the presence of repeat sequences (Charlesworth et al., 1994).

Two categories of repetitive sequences exist: interspersed repeats and tandem repeats (TRs).

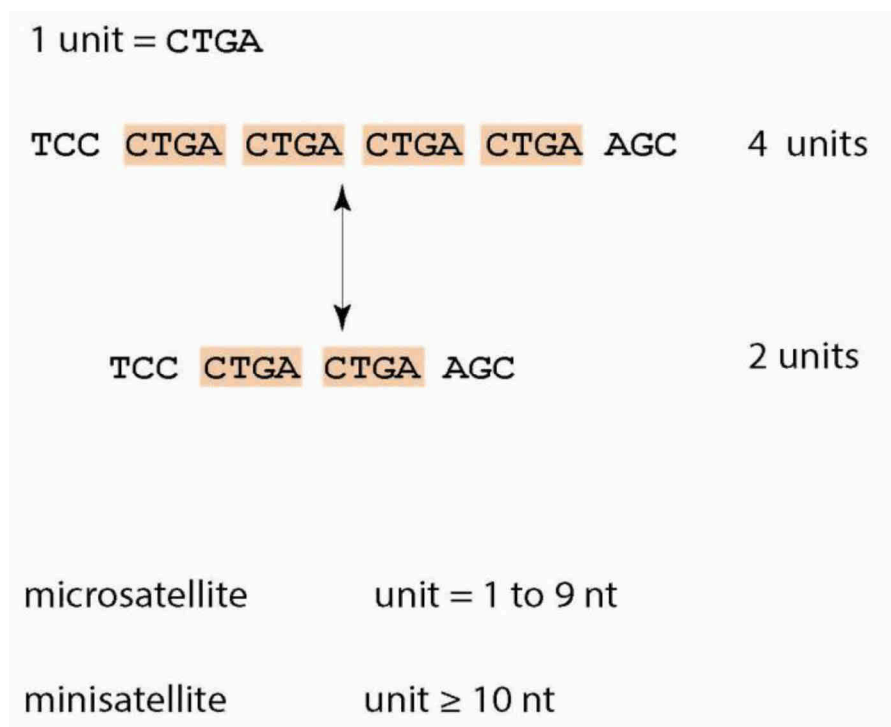
Interspersed repeats, the more predominant type of repeat, are remnants of transposons dispersed throughout the genome. Such elements are responsible for the diverse array of genome sizes amongst various species (Feschotte et al., 2009). On the other hand, TRs are repetitive DNA sequences, which exist directly adjacent, or in tandem, to one another (Figure 1). TRs are often referred to as satellite DNA because they were first identified as sequences constituting the second or "satellite" band that is detected after density-gradient centrifugation of DNA (Kit, 1961).

On the basis of unit length (unit= repeated sequence of DNA - see Figure 1), TRs are further divided into three subcategories-microsatellites, minisatellites and satellite.

Microsatellites, or simple sequence repeats (SSRs), are short TRs with unit length between one to five nucleotides, found in vertebrate, insect and plant genome, located in the euchromatin. Copy numbers are characteristically variable within a population, typically with mean array sizes of the order of 100 but with multiple array size classes distributed around the mean (Bruford, 1993; Di Renzo, 1994).

Minisatellites are TRs with unit length larger than ten nucleotides, generally involving mean array lengths of 0.5-30 Kb. They are found in euchromatic regions of the genome of vertebrates, fungi and plants, and are also highly variable in array size (Armour, 1992).

Satellite are TRs that can be similar in length (5-10bp) to micro- or minisatellites, or much larger (~100pb). They are typically organized as large [up to 100 megabases (Mb)] clusters in the heterochromatic regions of chromosomes, near centromeres and telomeres or on the Y chromosome (Cavalier-Smith, 1985; John and Miklos, 1988; Tyler-Smith and Willard, 1993; Lohe et al., 1993). They are apparently not as variable in array size within populations as micro- and minisatellites (Jabs et al., 1989; Wevrick et al., 1989).



**Figure 1.** Main definitions and characteristics of tandem repeats (TRs). TRs are unstable due to frequent changes in the number of repeat units. TRs with short unit length are also called microsatellites and those with long units are called minisatellites. nt= nucleotides (from Gemaley et al., 2012)

### ***1.1 Instability of Repeats***

Tandem repeats are evolutionarily pertinent due to their instability; they mutate at rates between  $10^{-3}$  and  $10^{-6}$  per cellular generation (i.e., 1 to 10 orders of magnitude greater than point mutations) (Verstrepen et al., 2005). Variation in the numbers and lengths of tandemly repeated units occurs on many different scales. Several genetic mechanisms can affect the number of repeating units in tandem arrays, usually occur from the addition or deletion of repeat units, rather than nucleotide substitutions. For instance, in a CTGA tract, most mutations occur by the addition or deletion of an entire CTGA unit as opposed to rare cases in which only a part of the repeat unit is altered (e.g., deletion of two nucleotides GA) (Figure 1) (Gemaley et al., 2012).

A set of major models have been proposed to explain TR expansions and contractions:

Strand-slippage replication (slipped-strand mispairing or DNA slippage) is a DNA replication error by which mispairing occurs between the template and nascent strands. As such, the template strand can loop out, causing contraction; the nascent strand can also loop out, leading to repeat expansion.

Recombination events, such as unequal crossing over and gene conversion may additionally lead to contractions and expansions of TR sequences (Verstrepen et al., 2005; Paques et al., 1998).

Rolling circle amplification (Felicciello et al., 2006) here, circular plasmids created by intrastrand exchange integrate (at some low level) into arrays by homologous recombination, either on the same chromosome or possibly on non-homologous chromosomes. If the circular plasmids contain the replication origins, rolling circle replication can greatly expand a short array sequence on a plasmid, allowing for rapid amplification (Hourcade et al., 1973; Flavell, 1982).

### ***1.2 Satellite DNA (satDNA)***

Satellite DNAs can be defined as highly reiterated noncoding DNA sequences, organized as long arrays of head-to-tail linked repeats located in the constitutive heterochromatin, the part of eukaryotic genomes that remain condensed throughout the cell cycle (Heitz, 1928).

The term “satellite DNA” is historical, because this kind of sequences was initially isolated from satellite bands in experiments with gradient centrifugation, due to the difference in A+T content from the rest of genomic DNA (Szybalski, 1968).

The basic repeating units, satDNA monomers, are often A+T rich and range in length from only few bp up to more than 1 kb, building up to 100 Mb long arrays. The preferential monomer length of 150–180 bp and 300–360 bp detected in many satellites in both plants and animals is often considered to mirror requirements of DNA length wrapped around one or two nucleosomes (Schmidt and Heslop-Harrison, 1998; Henikoff et al., 2001).

The copy number is substantially conserved within the populations, but the monomeric unit may show several variants regarding the nucleotide sequence (Charlesworth et al., 1994; Ugarkovic and Plohl, 2002).

Satellite DNA contribution to total genomic content varies significantly among species, exceeding sometimes 50% of total DNA (Elder and Turner, 1995; Schmidt and Heslop-Harrison, 1998), and consequently they are involved in the enormous variation of genome size in eukaryotes (Doolittle and Sapienza, 1980; Cavalier-Smith, 1985; Gregory et al., 2007).

Satellite sequences are the main constituent of centromeric and pericentromeric heterochromatin, two epigenetically determined regions responsible for correct pairing and disjunction of eukaryotic chromosomes in cell divisions (see for example Arney and Fisher, 2004; Hall et al., 2004; Bloom, 2007).

It is observed that principal DNA components underlying the majority of centromeres in plants and animals are satellite repeats, as corroborated by chromatin immunoprecipitation (ChIP) data (e.g. Nagaki et al., 2003; Zhong et al., 2002; Lee et al., 2005). Centromere is a multidomain locus necessary for poleward chromosomal segregation in mitosis and meiosis. Functional centromeres are usually embedded into large blocks of pericentromeric heterochromatin, but chromatin structure in centromeres is distinct from that in heterochromatin and in euchromatin (Sullivan and Karpen, 2004). While centromere structure and function is conserved through complex eukaryotes, DNA sequences in that region are paradoxically variable (Henikoff et al., 2001).

Satellite families in (peri)centromeric regions vary significantly in copy number, nucleotide sequence, organizational patterns, number and nature of inserted non-satellite DNA sequences (Plohl et al., 2008). Domains formed by single satellites are usually several hundreds kb, or even Mb long, such as in humans (Shiels et al., 1997; Mahtani and Willard, 1998; Schueler et al., 2001). Among individuals, array length of a single satellite can be highly polymorphic. For example, array length in alpha-satellite from human X varies almost 3 times (Mahtani and Willard, 1990).

It has been hypothesized that low abundance of satellite repeats in the centromere represents an early stage in the centromere evolution, characterized by progressive accumulation of satellite repeats in mature centromeres. (Wu et al., 2004; Nagaki et al., 2004).

### ***1.3 Evolutionary dynamics of satellite DNA.***

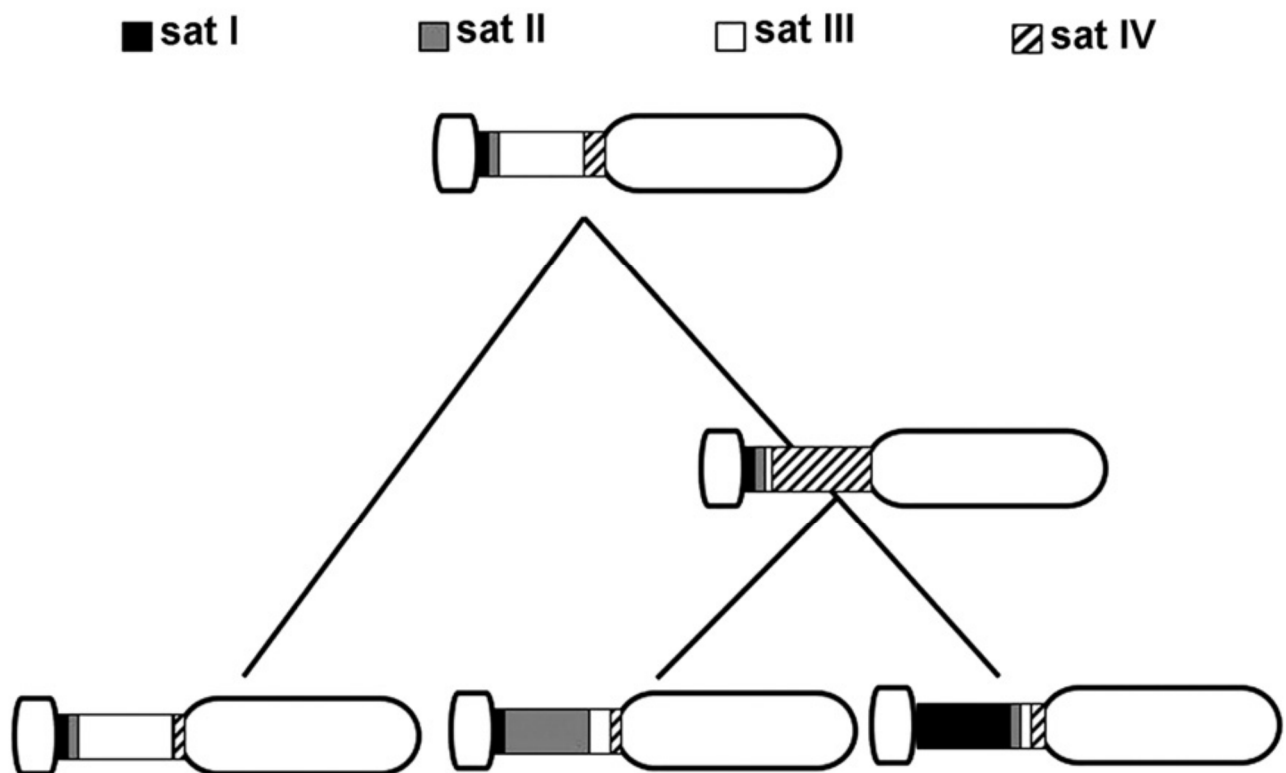
As discussed above, satDNAs in (peri)centromeric heterochromatin, represent rapidly evolving genomic components. Consequently, even among the most closely related species, they differ in

nucleotide sequence, copy number, and/or composition of satellite families (reviewed in Schmidt and Heslop-Harrison, 1998; Ugarković and Plohl, 2002).

Rapid evolution of satDNA sequences is possible owing to the accumulation of nucleotide divergences, usually with a high rate and in a gradual manner (Bachmann and Sperlich, 1993).

Gradual accumulation of mutations follows phylogeny at different hierarchical ranks. At the species level, centromeric satDNAs were informative in phylogenetic studies of the *Drosophila obscura* group (Bachmann and Sperlich, 1993), or in the study of the fish family Sparidae (Garrido-Ramos et al., 1999). Even within a genome, distinct forms of satellite DNAs can accumulate mutations at different rates, thus producing diversity of sequence patterns in (peri)centromeric areas. Interestingly, centromerically located higher-order units diverge more rapidly than pericentromerically located monomeric repeats (Rudd et al., 2006).

Accumulation of mutations in satellite families is not the only way to alter specific profiles of satellite repeats in short evolutionary periods. Since more than one satellite family exists in a genome, expansions and contractions of satellite arrays can efficiently change a landscape of DNA sequences in heterochromatin by replacing one dominant (major) satellite repeat with another one less represented (reviewed in Ugarković and Plohl, 2002) (fig.2). Unequal crossover is proposed to be the major mechanism responsible for dramatic fluctuations in the copy number of satellite DNAs (Smith, 1976).



**Figure 2.** The library model. In a genome, several satellite DNA families are coexisting on chromosomes, one family being often at high copy number as the major satellite. Different families can be preferentially amplified in the derived chromosomes, therefore changing the relative contribution of each family to the (peri)centromeric chromatin and leading to species-specific profiles of satellite repeats (from Plohl et al., 2008).

Satellite repeats may be the preferred form of DNA sequences in functional centromeres and their flanking regions just because of their unique characteristic to maintain sequence homogeneity over long stretches of DNA, and simultaneously to change rapidly in evolution.

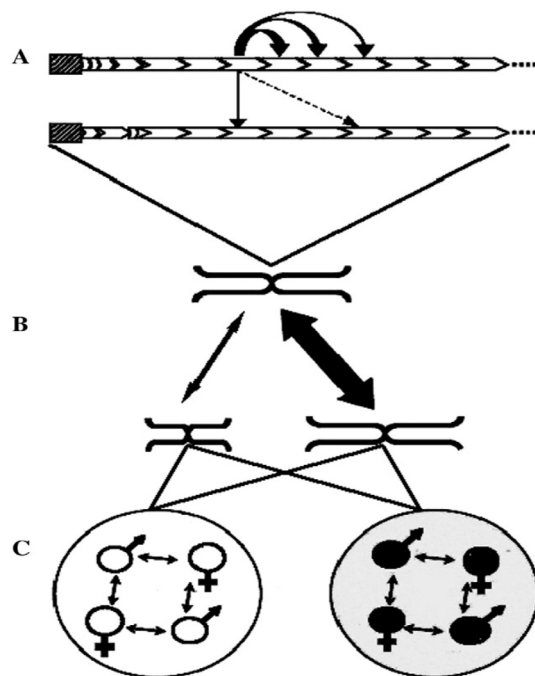
This characteristic is achieved by non-independent evolution of monomers. It is a consequence of molecular drive, a two-level process in which mutations are homogenized throughout members of a repetitive family, and concomitantly fixed within a group of reproductively linked organisms (Dover, 1982, 1986). The consequence is concerted evolution of monomers constituting a satellite DNA family.

#### ***1.4 Concerted evolution of satellite DNA***

Evolution of satDNA sequences is governed by principles of concerted evolution, in which mutations are homogenized throughout members of a repetitive family and fixed within a group of reproductively linked organisms in a stochastic process of molecular drive (Dover, 1986).

Sequence homogenization is due to diverse molecular mechanisms of nonreciprocal transfer, such as unequal crossover, gene conversion, rolling circle replication and reinsertion, and transposon-mediated exchange (Stephan, 1986; Dover, 2002; Glinka et al., 2006). While it is not clear which of the above reported mechanisms is preferentially involved in sequence homogenization, it is generally acknowledged that these mechanisms act more efficiently within localized subsets of satellite repeats, while efficiency drops progressively when changes are homogenized between arrays on the same chromosome, homologous and heterologous chromosomes (Fig. 3; Dover, 1986). Because of differences in rates of local and global sequence homogenization, adjacent monomers show a higher degree of sequence similarity than those retrieved at random, and can be often grouped into subsets or subfamilies, defined by diagnostic mutations (Willard and Wayne, 1987; Durfy and Willard, 1989; Schindelbauer and Schwarz, 2002; Hall et al., 2005; Roizes, 2006). Distinctive groups of monomer variants are usually chromosome-specific. As predicted by theoretical models (Smith, 1976; Stephan, 1989), monomers at array ends are more divergent than those located centrally due to the low efficiency of homogenization mechanisms (predominantly

unequal crossover) in bordering regions of the satellite array (Mashkova et al., 1998; McAllister and Werren, 1999; Bassi et al., 2000; Schueler et al., 2005). Adjacent monomer variants can be sometimes homogenized together and form a new, composite higher-order repeat (HOR) unit in which former monomers became subrepeats or subunits (Willard and Wayne, 1987; Warburton and Willard, 1990). Since a HOR is a homogenization unit, HORs generally show high level of sequence identity, while substantial sequence divergence is accumulated among constituent subunits (Willard and Wayne, 1987; Roizes, 2006 and references therein). While homogenization depends on mechanisms of genomic turnover, fixation results from random chromosomal assortment in sexual reproduction through meiosis and amphimixis, depending thus on population factors (Plohl et al., 2008).



**Figure 3.** Molecular drive and concerted evolution. A) Homogenization of mutated variant throughout the members of a repetitive family within an array and between sister chromatids. Arrow appearance is correlated with the homogenization efficiency. B) Efficiency of homogenization between homologous and non-homologous chromosomes. C) Variant spread among individuals (fixation) depends on bisexuality and population factors. Reproductive isolation leads to fixation of different repeat variants in genomes of different evolutionary units (white and black circles) (from Plohl et al., 2008).

These mechanisms induce high turnover of satellite sequences and rapid changes in copy number, nucleotide sequence and composition of satellites in the genome, resulting in high within-species homogeneity of satellite repeats and alterations in satellite profiles between species.

Some satellite sequences indeed change rapidly in evolution and accumulate mutations even at the population level, for example in the pupfish (Elder and Turner, 1994). On the other hand, some are widely distributed across species, providing evidence of nucleotide sequence conservation for long-term evolutionary periods. A study on *Pimelia* (Coleoptera, Tenebrionidae) suggested that a satDNA, highly abundant in all examined species, has persisted for more than 8 million years (Pons et al., 2002). Satellites in species from the *Drosophila virilis* group remained conserved for about 20 Myr (Heikkinen et al. 1995), and cetacean satellite DNA persisted for at least 40 Myr (Arnason et al. 1992). Low-copy number repeats indistinguishable in their nucleotide sequence from high-copy *Palorus ratzeburgii* satDNA have been detected in distant species *Pimelia elevata*, although these taxa diverged 60 Myr ago (Mravinac et al., 2002). Conservation or “freezing” of repeat families for long evolutionary periods can be explained as a consequence of mechanisms involved in concerted evolution, if a small bias in turnover mechanisms is anticipated (Dover and Flaveli, 1984). Such preference to maintain some putative “optimal” set of monomer variants may be due to functional constraints imposed on the nucleotide sequence (Mravinac et al., 2005). Slow rates of sequence change and of concerted evolution in some satellites were explained as specificity of slow general genomic evolution in sturgeons (Robles et al., 2004) and in whales (Arnason et al., 1992). Therefore, forces (i.e., gene conversion) leading to concerted evolution might be acting in these species, but that speciation by hybridization and/or polyploidization events played an important role in forming of their phylogeny. Thus, such a trend could be leading to the apparent failure of the concerted evolution process (Robles et al., 2004).

Furthermore, the final outcome of an extreme conservation of nucleotide sequence can be also predicted by the model, if “non-desirable” mutations are preferentially eliminated instead of being spread throughout a satellite family. The nucleotide sequence of some satellite families indeed remained “frozen” for long periods, even for tens of Myrs. Although the basis for favouring one sequence variant over another is usually not known, it might mirror constraints imposed on satellite sequences by some functional interactions. In that case, the evolution of at least some satellites seems to be driven by an interplay of selective constraints and stochastic events (by Plohl et al., 2008 and references therein).

Accumulation of mutations in satellite families is not the only way to alter specific profiles of satellite repeats in short evolutionary periods. In addition to sequence changes, satDNAs are permanently altered in copy number by expanding and contracting arrays of satellite monomers (Ugarković and Plohl, 2002; Plohl et al., 2012). Because usually more than one satellite family exists in a genome, fluctuations in their copy numbers can change very efficiently and rapidly any



profile of genomic satDNA. The library model of satDNA evolution explains the occurrence of species-specific satellite profiles as a result of differential amplifications and/or contractions within a collection, or library, of satellite sequences shares by related species (Fry and Salser, 1977; Mestrovic et al., 1998; Ugarković and Plohl, 2002). Not only distinct satDNAs, but also monomer variants or subfamilies from a single family can be distributed in genomes in the form of a library (Cesari et al., 2003).

### ***1.5 Functional potential of satellite DNAs***

The possible role of this fraction of the genome has been long discussed. Various evidences led scientists to formulate some hypotheses on the involvement of satellite DNAs in a series of functions ranging from centromere formation and function, heterochromatin assembly, regulation of gene expression and in epigenetic regulatory processes (reviewed in Ugarković, 2009; Pezer et al., 2012).

Some studies supported the role played by heterochromatic genomic compartments in a proper chromosomal behaviour in mitosis and meiosis (Csink and Henikoff, 1998). Indeed, satellite DNAs appear to be major constituents of functional centromeres, as shown in detail in *Drosophila melanogaster* (Sun et al., 1997) and in humans (Schueler et al., 2001). Satellite repeats should contain sequence motifs recognized by protein components. The best known is CENP-B box, the 17-bp long sequence motif found in its functional form in a subset of higher-order alpha satellite monomers. The motif binds the CEN-B protein, suggested to facilitate kinetochore formation (Ikeno et al., 1994; Masumoto et al., 2004; Schueler et al., 2005). Motifs resembling to CENP-B box were observed in diverse satellite families from various species, but their true functional significance is not known (Canapa et al., 2000; Lorite et al., 2004).

Transcriptional activity was not expected for repetitive DNA sequences residing in the transcriptionally suppressive heterochromatin environment. However, satellite transcripts have been so far in many animal and plant taxa indicating that satellite transcription might be a general phenomenon (for example, Varadaraj and Skinner, 1994; Lorite et al., 2002; Rudert et al., 1995; Pathak et al., 2006; Lee et al., 2006). Satellite DNAs never show any prominent open reading frame, and accordingly, transcript translation has never been demonstrated (Plohl et al., 2008). However, various functional roles have been hypothesized for these transcripts. The strand or tissue or stage specificity observed in some cases suggests the involvement of satellite transcripts in regulatory functions (Varadaraj and Skinner, 1994; Lorite et al., 2002; Pathak et al., 2006).

The most striking example is the sequence of human alpha satellite DNA, which is found highly conserved in chicken and zebrafish (Li and Kirby, 2003). This sequence is transcribed during early embryogenesis of both species and it is proposed that it may serve as a control element in gene regulation. It has recently been shown that transcripts of satellite DNAs and other repetitive sequences are functional in the form of small interfering RNAs which act as signals necessary for establishment and/or maintenance of heterochromatin in different eukaryotes (Volpe et al., 2002; Aravin et al., 2003).

Some satellite DNAs from insects, nematodes and amphibians produce hammerhead structures with a possible ribozymic activity (Rojas et al., 2000 and references therein). Transcripts of centromeric satellite in maize were shown to remain tightly bound within centromeric chromatin and contribute to initiation and stabilization of kinetochore chromatin structure (Topp et al., 2004).

Another proposed role of satellite DNA transcripts attracted particular attention in recent years. It was observed that satellite DNA transcripts are involved in the initiation of histone H3 methylation, a necessary prerequisite in heterochromatin formation and maintenance (Volpe et al., 2002; Martienssen, 2003). Transcripts from centromeric satellites are processed to produce small interfering RNAs (siRNA) that mobilize a number of proteins and specifically target their coding sequence. This sequence is then packed into the transcription-inhibiting heterochromatin structure (reviewed in Grewal and Elgin, 2007). This mechanism requires low-levels of transcription and may be universal, since siRNAs processed from centromeric satellite repeats were identified in several eukaryotic species (Lee et al., 2006 and references therein). However, the relationship between transcription of centromeric satellite repeats and centromeric silencing/centromere function is still unclear.

Furthermore, a potential indirect role of satellite DNAs in chromosomal repatterning should be mentioned. This role would be important in that karyotype rearrangements can be important in promoting reproductive isolation between populations, ultimately leading to speciation (Coghlan et al., 2005). For example, a link between satellite DNAs and chromosomal instability was studied in the genus *Ctenomys*, one of the most speciose and karyotypically diverse mammalian taxon. The high karyotypic variability was associated with amplifications and deletions of the major *Ctenomys* satellite DNA and with the number of species (Slamovits et al., 2001; Hartmann and Scherthan, 2004; Ellingsen et al., 2007). Satellite DNAs appear also involved in genome restructuring during development in different organisms. The process of chromatin diminution is known to occur during development in different organisms, such as in the nematode *Parascaris univalens* and *Ascaris suum*, in copepods and in a hagfish. The quantity of lost DNA ranges up to 94%, and is mainly

composed of satellite sequences (Stanley et al., 1984; Drouin, 2006, and references therein). The hypothesis that RNAi-related mechanisms are involved in chromatin diminution has been put forward (Drouin, 2006).

## **2 Sex Determining Mechanisms in Vertebrates**

Sexual reproduction is a nearly universal feature of eukaryotic organisms. Given its ubiquity and shared core features, sex is thought to have arisen once in the last common ancestor to all eukaryotes. Vertebrates have various sex determining mechanisms. These have been broadly classified as either environmental sex determination (ESD) or genotypic sex determination (GSD).

The term environmental sex determination (ESD) implies that the sex of an individual is determined by the environmental conditions experienced during early development. The key environmental factor can be pH (e.g., in fish; Römer and Beisenherz, 1996), or social conditions or relative juvenile size (e.g., in fish; Francis and Barlow, 1993; Holmgren and Mosegaard, 1996), and the temperature. This latter is the most common and most studied environmental parameter inducing sexual determination and it is generally referred to as TSD (Temperature-dependent Sex Determination) system.

In genotypic sex determination (GSD) the sex of a zygote is determined entirely by its genotype, and the sex of an individual is fixed at fertilisation. The most common type of GSD involves sex chromosomes. Two main systems of sex chromosomes can be found in vertebrates. One type is termed as XX/XY system and male (XY) is the heterogametic sex. The other type is indicated as ZZ/ZW system with the heterogametic sex represented by female (ZW). In polygenic sex determination, which is less common than the preceding one, sex is determined by a number of genes, each with minor effect, distributed throughout the chromosome complement (Liew et al., 2012)

### ***2.1 Distribution of sex determining mechanisms among vertebrate taxa***

ESD is common among reptiles, and occurs also in amphibians and fish. TSD is the form of ESD found in reptiles. In some lizards and in alligators, eggs incubated at low temperature (less than 27°C) give rise to 100% females, and eggs incubated at high temperatures (above 30°C) give rise to 100% males. In many turtles it is the other way round: 100% males at low temperatures and 100% females at high temperatures. In other turtles and in crocodiles, incubation at intermediate temperatures leads to 100% males, whereas both low and high temperatures lead to females only (reviewed in Bull,

1983). In all these cases there is only a very narrow temperature range in which both sexes are produced. This range, however, may vary within a species and is heritable (Bull et al., 1982, Janzen, 1992). In fishes, the sex ratios usually vary less extremely with temperature, but nevertheless temperature-dependent sex determination has been found in various species (e.g. Conover and Heins, 1987, reviewed in Francis 1992, Römer and Beisenherz, 1996, Goto et al., 2000). The molecular mechanism by which temperature triggers sex determination in these taxa is completely unknown, although an influence of temperature on hormones (estrogens) and their receptors has been proposed (Sarre et al., 2004).

In mammals, birds, amphibians, and many reptiles and fish, sex is determined genetically. GSD mechanisms range from those that depend on allelic variation at a single locus to those in which the sex-determining gene is borne on a pair of differentiated sex chromosomes (Ezaz et al., 2005). In all these systems, one sex (homogametic sex) produces gametes that are homozygous for the same sex allele or chromosome and the other sex (heterogametic sex) produces equal proportions of two kinds of gametes (one containing the X or Z and the other the Y or W) that specify equal numbers of males and females in the offspring. There is a fundamental distinction between species with male heterogamety (XY male, XX female, such as mammals), and female heterogamety (ZZ male, ZW female such as birds). However, the two systems share many parallels; for instance, the X and Z chromosomes are usually larger and contain many more active genes than the sex-specific Y or W. Environmental and genetic sex determination have traditionally been thought to constitute completely different triggers, but these systems sometimes interact, as in some reptile species, amphibians, and fish where we find have elements of both (Quinn et al., 2007). Genotypic and environmental sex determination have been viewed as two ends of a continuum of sex-determining mechanisms (Shine et al., 2002; Sarre et al., 2004). For example, XX/XY sex chromosomes were described in a scincid lizard (*Bassiana duperreyi*) that presumably possessed TSD mechanisms (Shine et al., 2002). The mechanism for long-term maintenance of these sex chromosomes in a population with ESD is not clear (Bull, 2008). The continuum between ESD and GSD can be explained by the existence of GSD with environmental effects (Valenzuela et al., 2003), where environmental conditions can influence the observed sex ratio of hatchlings, but sex of an individual is still determined by its genotype (e.g. by differential fertilization of gametes producing particular sex or sex-specific mortality of embryos), or by the thermal induction of sex revertants (i.e. production of individuals with the wrong gonad type for their genotypes). Sex revertants induced by extreme developmental temperatures are well known in many GSD vertebrate lineages (e.g. Witschi, 1929) and were recently documented in two species of lizards as well (Quinn et al., 2007 in the ZZ/ZW agamid *Pogona vitticeps*, and Radder et al., 2008 in the XX/XY skink *Bassiana*

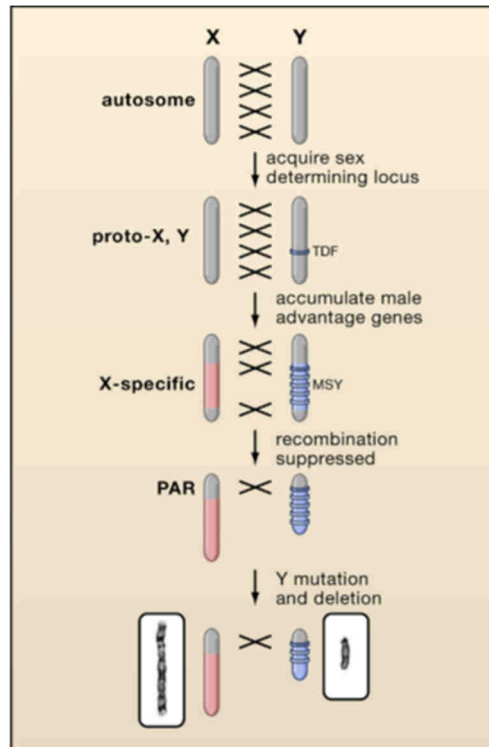
duperreyi). In both cases, the existence of thermally induced sex reverted individuals was interpreted by the authors as evidence for the co-occurrence of GSD and TSD or for a transitional state between TSD and GSD. Nevertheless, gonadally and phenotypically reverted individuals still possess sex chromosomes corresponding to their genotypic sex, and when crossed with non reverted mates, they produce progeny with a skewed sex ratio. We understand GSD and TSD as two dichotomous sex determining systems that do not differ in thermal dependency of sex ratios, but basically in the presence or absence of sex chromosomes (Pokorná et al., 2009).

## ***2.2 Origin of sex chromosomes***

Sex chromosomes are very atypical of the genome. Evidently the acquisition of a sex determining allele confers on a chromosome special properties and a special fate. Sex chromosomes are considered to be the most variable region of the genome. However, this is true only for the sex-specific element (the Y or W). The mammal X and the bird Z are extremely conserved (Graves, 2008).

Sex chromosomes differ from autosomes in that the two members of the sex chromosome pair typically vary in morphology, size, number, staining, and gene content. They are highly specialized and appear to have evolved independently many times in vertebrates (for review see Graves, 2008). Sex chromosomes are thought to evolve from autosomes (Ohno, 1967). However, vertebrate XY and ZW sex chromosomes are not homologous, suggesting independent evolution of sex chromosomes in different lineages from non-homologous ancestral autosomes (Fridolfsson et al., 1998; Nanda et al., 2000, 2002). The autosomal origin of sex chromosomes is also supported by the fact that sex chromosomes of one group of organisms are autosomal in another group (e.g., Matsubara et al., 2006; Pokorná et al., 2011), which would be consistent with an independent origin of sex determination in vertebrates (e.g. Graves, 2008).

The trajectory from autosomes to sex chromosomes may start with the emergence of a mutation that confers a sexual advantage. Additional sex-linked mutations in other genes then accumulated on the same homologue. Recombination between the primordial sex chromosomes was suppressed by chromosomal rearrangements such as inversions to preserve the block of sex-linked genes. The absence of recombination fostered the accumulation of mutations and repetitive sequences with subsequent 'heterochromatization' of the sex-specific chromosome. Deletions of heterochromatin account for the smaller sizes usually observed for the Y or W chromosomes compared with the X or Z chromosomes, respectively (by Modi et al., 2005) (Fig. 4).



**Figure 4.** Differentiation of an X and Y chromosome from an ancestral autosome.

This process is initiated when one partner acquires a sex-determining locus such as the testis-determining factor (TDF). Accumulation of male-specific alleles selects for repression of recombination (represented by crosses), creating an X-specific region on the X and a male-specific region on the Y (MSY). Exclusion from recombination leads to rapid degradation of the MSY leaving only a small pseudoautosomal region (PAR). Active genes are lost, leaving largely genes that have, or acquire, a male advantage. This model accounts for the differences in size and gene content of the human X (left) and Y (right) (Graves 2006).

The initiating mechanisms of recombination suppression are not yet clear. However, two hypotheses have been proposed to explain how restriction of recombination spreads along sex chromosomes. Stepwise model highlights the role of inversions in suppression of recombination, which is supported by evolutionary strata of different age, i.e. sex chromosome regions of different levels of divergence, observed in mammals (Kohn et al. 2004, Macha et al. 2012), birds (Nam and Ellegren 2008), snakes (Vicoso et al. 2013a), papaya (Wang et al. 2012), and white campion, *Silene latifolia* (Bergero et al. 2007, 2013, Hobza et al. 2007). However, it was shown that the multiple inversions on the Y chromosome are a consequence rather than a cause of suppressed recombination in *S. latifolia* (Bergero et al. 2008). Thus, other mechanisms such as heterochiasmy, i.e. sex-specific differences in recombination, are probably involved in a gradually proceeding cessation of recombination in the early stage of sex chromosome differentiation (Perrin 2009, Bachtrog 2013, Natri et al. 2013).

The suppression of recombination between the heterochromosome and its homologue would start the gradual erosion of the heterochromosome (Y or W) itself because genes that are not essential for males (in XY systems) or females (in ZW systems) show accelerated rates of mutation and deletion. Consequently, the heterochromosome becomes progressively gene-poor (see Handley et al., 2004) and in the extreme case the simplification process can lead to the complete loss of the heterochromosome (e.g., Just et al., 2007).

Concerning the heterochromatinization of heterochromosomes (W or Y), it is still debated whether repetitive sequences are the cause of the suppression of recombination or whether these elements accumulate on the Y or W chromosome as a consequence of the block of genetic exchange between sex-chromosomes. The observation of an evolving incipient Y chromosome in the fish families Gasterosteidae (Peichel et al., 2004) and Cichlidae (Griffin et al., 2002) supports the first hypothesis. Indeed, despite sex chromosomes are not clearly cytogenetically distinct, a molecular analysis revealed an accumulation of heterochromatin in the sex-determining region of the “proto-Y” resulting in a reduction in recombination. On the other hand, accumulation of heterochromatin could be a mere effect of cessation of recombination. In fact, the impossibility of recombination between the unpaired sex chromosomes (W or Y) opens the door to the “invasion” of various repetitive sequences, like transposons, microsatellites and tandem repeats, on sex chromosomes. In this respect, it has been hypothesized that the heterochromatinization may be a mechanism for the defense against invasive transposable elements (Kidwell, 2002; Steinemann & Steinemann, 2005). However, little is known about how this occurs or about how the absence of recombination affects the subsequent evolutionary fate of the repetitive sequences in the W or Y chromosome.

The repetitive DNA sequences or TEs accumulation and expansions on one hand and contractions on the other hand are stepwise or are occurring simultaneously (Kejnovsky et al. 2009). However, the lifetime of an old Y chromosome is often prolonged by the addition of segments transferred from autosomes (Graves, 2005). Acquisition of new genes from autosomes mediated by retrotransposition has been shown in humans (Lahn and Page, 1999), and a similar duplicative transfer has also been shown in the young Y chromosomes of *Silene latifolia* (Matsunaga et al., 2003). Kejnovsky et al. (2008) have recently discussed the potential for junk DNA accumulation to start at an early stage in the evolution of sex chromosomes. Both past cytogenetic analyses and recent genome projects have revealed that many animal Y chromosomes have more abundant heterochromatin derived from repetitive sequences compared with X chromosomes and autosomes. Accumulation of repetitive sequences induces abnormal recombination and chromosome breaks. Thus, junk DNA accumulation may well be a factor in the generation of differences in morphology and size observed between X and Y chromosomes; for example, in both the fruit fly, *Drosophila*

*melanogaster*, and in humans, the Y chromosome is drastically smaller than the X chromosome (Adams et al., 2000; Skaletsky et al., 2003). These findings in animal species show that the accumulation of junk DNA is an important step in promoting the morphogenesis of sex chromosomes.

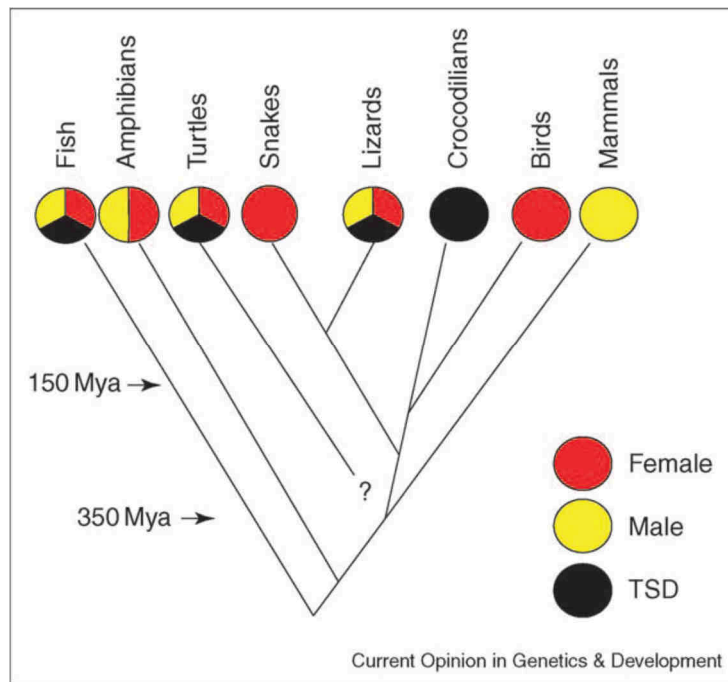
Junk DNA accumulation on Y chromosomes has been believed to be a symptom of Y-chromosome degeneration. Insertion of repetitive sequences into coding genes and regulatory regions induces alteration in the genes' functions and results in gene loss. However, there is no correlation between the insertion of the transposable element and gene dysfunction on the Y chromosome of *D. miranda* (Bachtrog et al., 2008). The contradiction between gene acquisition and accumulation of highly repetitive sequences on the *Drosophila* Y chromosomes, indicates that junk DNA accumulation is not always directly connected with Y chromosome degeneration (Matsunaga 2009).

Then the Y chromosome either remains as a genetic entity or can be lost entirely. A new autosomal pair can then be chosen to become a new pair of sex chromosomes and the cyclic process can continue. The persistence of the Y chromosome indicates that it can repeatedly arise *de novo*, for example, by the fusion between an autosome and an X chromosome followed by the fixation of the neo-X and the neo-Y chromosomes as was shown in grasshopper *Podisma pedestris* (Westerman and Hewitt, 1985; Veltsos et al., 2008).

### ***2.3 Evolution of sex chromosomes in squamate reptiles***

The study of reptilian genome is of great interest because reptiles occupy a pivotal position in the phylogeny of vertebrates — they are the direct ancestor to birds and mammals — and because they also possess several unique biological attributes that, if better understood, could contribute significantly to understanding basic evolutionary biology and the molecular mechanisms behind human health and disease (Modi and Crews, 2005). Reptiles exhibit some of the most extraordinary variability in sex chromosome structure and patterns of sex determination modes seen among vertebrates (Fig. 5) (Valenzuela and Lance, 2004). These modes include gonochorism (separate sexes) and parthenogenesis, oviparity, viviparity, and ovoviviparity, genotypic sex determination (GSD), with either male (XX/XY) or female (ZZ/ ZW) heterogamety, and temperature-dependent sex determination (TSD) (Ezaz et al., 2009a).

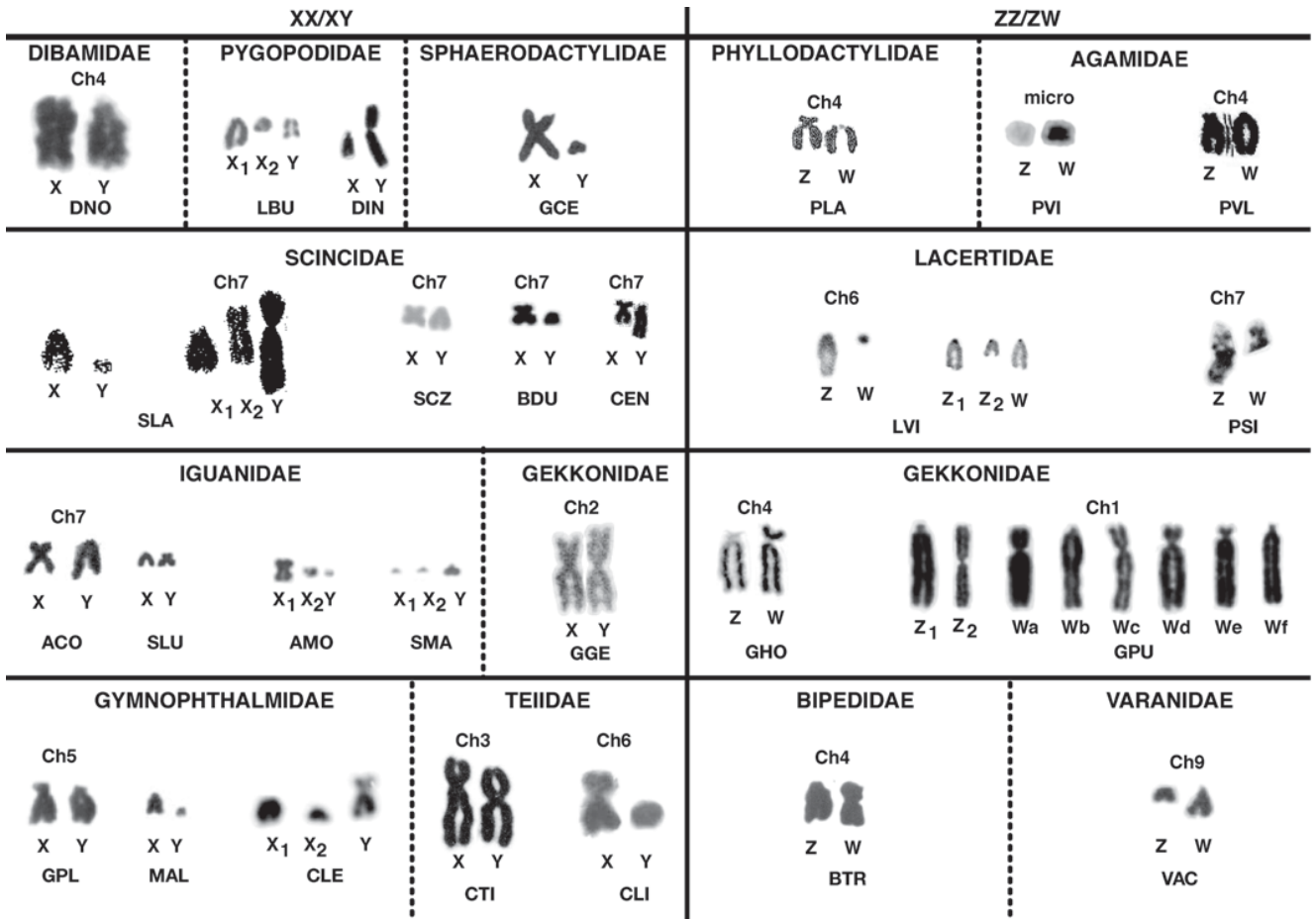




**Figure 5.** Vertebrate phylogeny illustrating sex determination modes in different taxa. ‘Female’ and ‘Male’ represent genetic sex determination with female and male heterogamety, respectively. TSD represents temperature-dependent sex determination. An unanswered question in contemporary reptilian phylogenomics regards the relationships of turtles to other reptiles (from Modi e Crews, 2005).

The variability seen among reptilian sex chromosomes suggests that sex chromosome and sex determination systems have evolved independently in different lineages (Modi e Crews, 2005).

Lizards (order Squamata, suborder Sauria) are particularly fascinating because the distribution of sex determining mechanisms shows no clear phylogenetic segregation. This implies that there have been multiple transitions between TSD and GSD, and between XY and ZW sex chromosome systems (Ezaz et al., 2009a). Lizards with GSD display remarkable diversity in sex chromosome differentiation, ranging from cryptic or homomorphic to highly differentiated. Much of this variation occurs within families, often among closely related species and even within the various races or populations of the same species. For example, the gekkonid lizard *Gehyra purpurascens* displays two Z chromosome and six W chromosome morphs, primarily as the result of centromeric inversions (Moritz, 1984) ( fig. 6 ). Variation in the morphology of sex chromosomes among closely-related taxa, or populations of one taxon, indicate that morphological evolution of sex chromosomes, and perhaps also sex-determining mechanisms in lizards may occur relatively easily in comparison to mammals and birds (Ezaz et al., 2009a).



**Figure 6.** A snapshot of morphological diversity of sex chromosomes in lizards. Ch: Chromosome. ACO: *Anolis conspersus* AMO: *Anolis monesis*, BDU: *Bassiana duperreyi*, BTR: *Bipes tridactylus*, CEN: *Claireascincus entrecasteauxii*, CLE: *Calyptommatus leiolepis*, CLI: *Cnemidophorus littoralis*, CTI: *Cnemidophorus tigris*, DIN: *Delma inornata*, DNO: *Dibamus novaeguineae*, GCE: *Gonatodes ceciliae*, GGE: *Gekko gekko*, GH0: *Gekko hokouensis*, GPL: *Gymnophthalmus pleei*, GPU: *Gehyra purpurascens*, LBU: *Lialis burtonis*, LVI: *Lacerta vivipera*, MAL: *Micrablepharus allicolus*, PLA: *Phyllodactylus lanei*, PSI: *Podarchis sicula*, PVI: *Pogona vitticeps*, PVL: *Phrynocephalus vlangalii*, SCZ: *Saproscincus czechurai*, SLA: *Scincella lateralis*, SLU: *Sceloporus lundelli*, SMA: *Sceloporus maculosus*, VAC: *Varanus acanthurus* (from Ezaz et al., 2009a).

Like birds, turtles and snakes, most lizards have a karyotype composed of macrochromosomes (ranging from  $2n = 10-46$ ) and microchromosomes (ranging from  $2n = 0-26$ ) (Olmo and Signorino, 2005). Microchromosomes have been found to be gene rich in birds with 2–3 times the number of genes contained in macrochromosomes (Smith et al., 2000) and to have higher recombination rates (Rodionov et al., 1992). In addition, microchromosomes are GC and CpG-rich and contain few repetitive elements (Hillier et al., 2004) and therefore, are likely to be important for generating genetic variation (Organ et al., 2008). Importantly, in some species of lizards, sex chromosomes

have been found to be microchromosomes (Gorman and Atkins, 1966; Gorman, 1973; Bull, 1983; Ezaz et al., 2005; Ezaz et al., 2009a). The patterns of differentiation of these sex microchromosomes are highly variable within and among groups (fig. 6), and have evolved primarily via the accumulation and amplification of heterochromatin (Ezaz et al., 2009a). Sex microchromosomes appear to be highly labile in at least one family of lizards, with agamids exhibiting a substantial array of forms among closely related species. In this group, the W chromosomes are highly to moderately heterochromatic, whereas the Z chromosomes are euchromatic and can be detected only by mapping sex chromosome specific DNA sequences (Ezaz et al., 2005, 2009a) or by mapping sex chromosome specific BAC clones (Ezaz et al., 2009a). These patterns of heterochromatic variability suggest various stages of sex chromosome differentiation within closely related species.

The accumulation of repetitive sequences during sex chromosome evolution in reptiles has been studied only in snakes (Jones and Singh, 1985; O'Meally et al., 2010). Pythons, considered basal in snake phylogeny, show homomorphic sex chromosomes, without accumulation of repetitive DNAs. On the contrary, in many advanced snakes like colubrids or elapids (Colubroidea) the heteromorphic W sex chromosome exhibit a strong accumulation of repeats (Jones and Singh, 1985; O'Meally et al., 2010). For example, the W chromosome in the elapid *Notechis scutatus* is composed almost entirely of repetitive sequences, including 18S rDNA and the banded krait minor-satellite (Bkm) repeats (Lee et al., 2007). The Bkm repeats consist of tandem arrays of 26 and 12 copies, respectively, of two tetranucleotides, GATA and GACA (Epplen et al., 1982). Bkm-related repeats are also accumulated on the heterogametic sex chromosomes in many vertebrates and also in plants (Jones and Singh, 1981; Parasnis et al., 1999), suggesting their possible role in the transcriptional activation of sex chromosome heterochromatin (Singh et al., 1976).

One interesting reptilian group in which sex chromosomes evolution involves heterochromatinization is represented by the lizards belonging to the Lacertidae family.

### **3. The study species**

#### **3.1 The family Lacertidae (Reptilia, Squamata)**

The family Lacertidae consist of about 42 genera including 321 species widespread in the Palaearctic region (Uetz and Hošek, 2015). Recent molecular analyses strongly support the monophyly of lacertids, and suggest that Lacertidae may be the sister-group of Amphisbaenia, the worm lizards (Townsend et al. 2004; Vidal and Hedges 2004), otherwise its nearest relatives are Teiioidea, a group of squamates currently exclusive to the American continent (Arnold et al., 2007).

Lacertids are defined as a clade by a number of mainly exclusive synapomorphies:

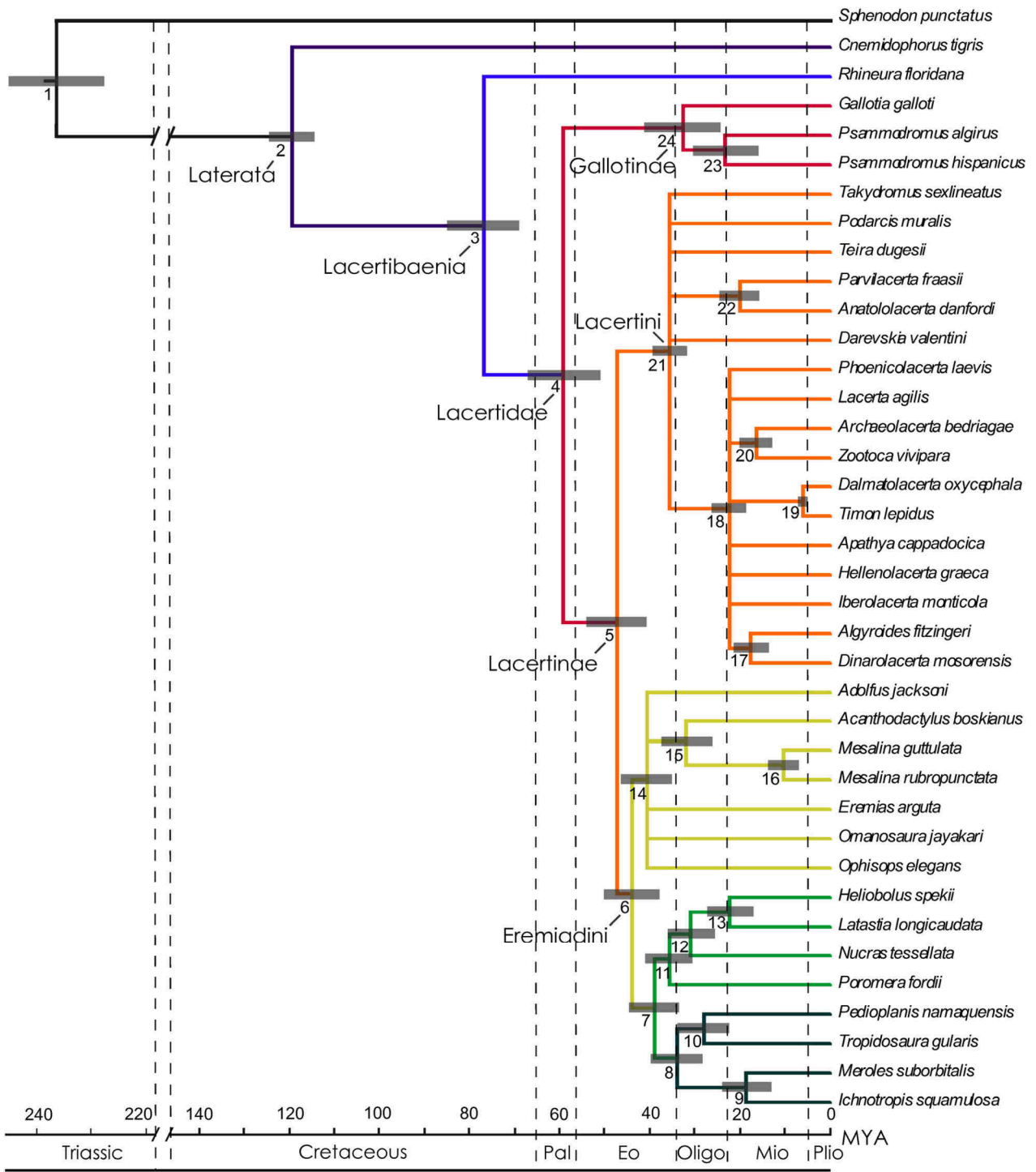
- lack of downgrowths on the parietal bone (Estes et al., 1988);
- supratemporal fenestra largely or wholly filled by postfrontal bone (a feature shared with Scincidae);
- usual presence of sexual variation in the number of presacral vertebrae;
- bodenaponeurosis divided into two lobes caudally, and a parasagittal vertical sheet connecting the quadrate aponeurosis to the temporal fascia (Rieppel, 1980);
- abdominal fat-bodies largely outside the peritoneum (Arnold, 1989a);
- either the lobes of the hemipenis invested by the retractor penis magnus muscle, or the lobes usually omplexly folded;
- the erect hemipenis supported by an elaborate cartilaginous supporting structure, termed armature (Arnold, 1973, 1986, 1989a).

To these features presence of a microornamentation on the epidermis of the hemipenial lobes consisting of individual cells that are typically hook-shaped spines or crown-shaped tubercles can probably be added (Klemmer 1957; Böhme 1971; Arnold 1973, 1986, 1989a). Additional putative synapomorphies of the Lacertidae, involving the scaling of the posterior dorsal surface of the head, such as widespread presence of an occipital scale, have also been put forward (Borsuk-Bialynicka et al. 1999).

### **3.2 The main groupings of lacertids**

In recent years this family has been the subject of several taxonomical studies, considering both molecular and morphological characters. The analysis of a relatively large morphological data set (84 characters, equivalent to 112 binary characters) allowed the recognition of a Palaearctic and Oriental group of relatively primitive forms, and a monophyletic group consisting of Afrotropical and advanced Saharan and Eurasian taxa (Arnold 1989a).

Afterwards a series of studies based on sequences of mitochondrial DNA genes (cytochrome b, 12S rRNA, 16S rRNA and Cytochrome Oxidase subunit I) (Harris et al. 1998a, Fu 2000) and two nuclear genes (RAG-1 and c-mos) (Mayer and Pavličev 2007) explored the relationships of lacertid lizards. All data sets strongly indicate that the family Lacertidae is phylogenetically arranged in two subfamilies, Gallotiinae and Lacertinae. The Gallotiinae, which includes two genera, *Gallotia* and *Psammmodromus*, is sister to the Lacertinae. This latter comprises two monophyletic tribes, the Eremiadini of Africa and arid southwest and central Asia, and the Lacertini of Europe, northwest Africa and southwest and east Asia (Kapli et al., 2011) ( Fig. 7).



**Figure 7.** 95% majority rule consensus tree for Lacertidae with divergences estimated under an Uncorrelated Lognormal relaxed molecular clock, based on a concatenated data set of 3 mitochondrial and 2 nuclear genes. Gray bars represent mean divergence dates  $\pm$  1 standard deviation. Nodes are numbered consecutively and correspond to node numbers in the Additional file 1. A geological time scale in millions of years is shown below. (from Hipsley et al., 2009).

In all Bayesian phylogenetic analysis carried out by Hipsley et al. (2009), the Amphisbaenia-Lacertidae split was dated to the Cretaceous, about 75 million years ago, while within Lacertidae, the split between Gallotinae and Lacertinae, is estimated to have occurred in the Paleocene (56–58 Mya), with the initial radiation of the African clade occurring in the mid-Eocene (44–46 Mya). Within the Eremiadini, the separation of the Saharo-Eurasian and Ethiopian clades occurred after their split from the Lacertini, about 40–43 Mya. The subfamily Gallotinae diverged into its component genera, *Gallotia* and *Psammodromus*, during the Oligocene, 29–32 Mya (Fig. 7). Within the Lacertidae, the majority of divergences occur in the mid- to late Eocene, after the Eremiadini split from their palearctic sister clade, giving rise to a large number of species (Hipsley et al., 2009 and references therein).

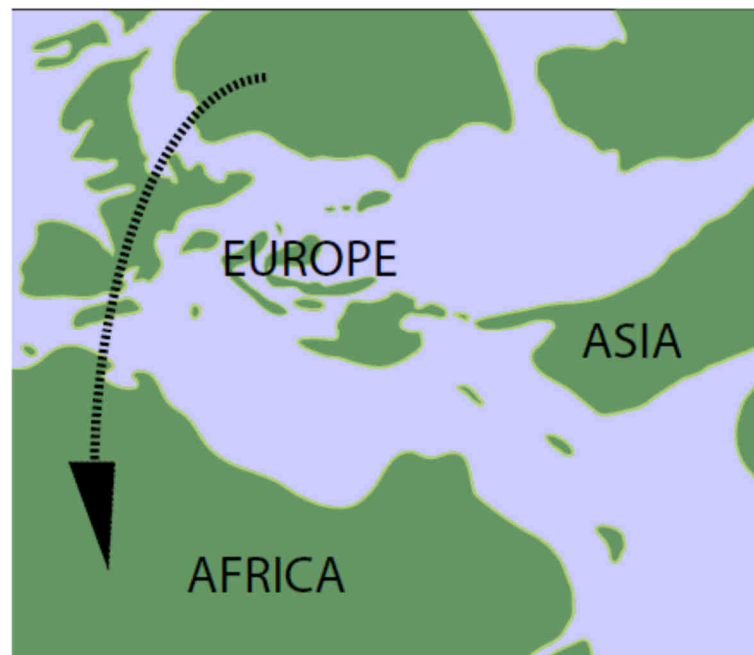
### 3.3 Historical biogeography of Lacertidae

Most authors agree that lacertids originated in Europe, as indicated by the mainly European distribution of the basal Gallotinae (Arnold et al., 2007). According to most reliable model (ULN), the majority of the lacertid radiation occurred in the mid-Eocene, 43–46 Mya. In that epoch, Europe was an archipelago composed of larger and smaller islands separated by shallow bodies of water (Hipsley et al., 2009 and references therein). The appearance of land bridges in the Eocene as well as increasing aridity are thought to have played an important role in terrestrial vertebrate migration, and evidence for faunal exchange between Europe and Africa can be seen in the fossil records of mammals and alligators (Hipsley et al., 2009 and references therein).

Three hypotheses about how this migration took place have been proposed:

1. A first hypothesis is that lacertids entered North Africa at its northwestern edge via a chain of islands and diversified as they moved towards the southern tip of the continent (Fig. 8). A primarily western migration for African lacertids is supported by modern biogeography, since the basal most taxa of both the European and African radiations are found along the western edges of the continents. The basal-most palearctic genus in our analysis (*Podarcis*; ULN, DM, CPP 50% consensus trees) occurs primarily in the western Mediterranean region and *Atlantolacerta andreanskyi*, which morphologically and genetically appears basal in the African radiation (Arnold et al., 2007) is restricted to the Atlas Mountains in northern Africa.
2. According to a second hypothesis, proposed from Mayer and Benyr 1994, and Arnold et al. 2007, the colonization of Africa by Lacertidae occurred in the Miocene over the land bridge connecting Arabia and Africa, which remained up until the early Pliocene. However, these dates for the initial radiation of African lacertids conflict with hypothesis Hipsley et al., 2009.

3. An alternative colonization scenario proposes that the African lineage split from the Lacertini in Europe prior to migrating to Africa, and then only later radiated into its component lineages after reaching the African continent. The discovery of a fossil lacertid in Europe with African-like trait would support this hypothesis. Indeed, the Baltic amber lizard *Succinilacerta* from mid-Eocene Poland was for some time assigned to the south African genus *Nucras*, suggesting that it resembles an African lacertid, at least superficially (Hipsley et al., 2009 and references therein).



**Figure 8.** Paleogeographic map of Europe and North Africa in the Late Eocene.

Arrow indicates possible lacertid migration route to Africa between southwestern Europe and northwestern Africa via small island chains (by Hipsley et al., 2009).

### ***3.4 The Tribe Lacertini Oppel, 1811***

***Morphological and chromosomal features*** (Arnold et al., 2007 and references therein):

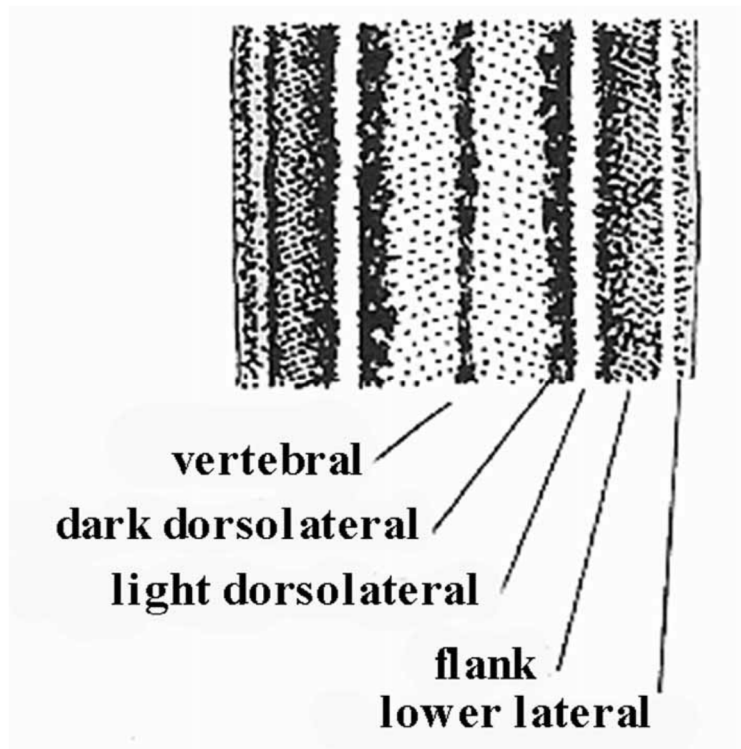
Tribe Lacertini exhibits a distinctive syndrome of morphological features, including flattened heads and bodies, fenestrated supraocular osteoderms in adults, often unkeeled dorsal body scales and little posterior overlap of the ventral ones, slender and fragile tails and, frequently, dorsal patterns in which longitudinal striping is reduced or absent.

*Size and shape*

- *Body size.* Adults of most Lacertini species are around 55–90 mm from snout to vent (exceptionally over 90 mm in *Podarcis*). A few forms are smaller including some *Algyroides* (*A. fitzingeri* not more than about 45 mm and *A. moreoticus* often under 50 mm). In contrast, *Lacerta* and *Timon* are generally much larger than other Lacertini, with respective adult sizes of about 70–175 mm and 100–210 mm from snout to vent or more. Adult males are usually larger than females, exceptions are represented by some *Darevskia*, *Lacerta*, *Iberolacerta*, *Takydromus* and *Zootoca*.
- *Body shape.* The head and body is fairly deep in most Lacertini, particularly in *Lacerta*, *Timon* and *Zootoca*. In contrast, a number of species that regularly use crevices as refuges are moderately to very depressed.
- *Dorsal body scales.* The body is usually covered above and on its flanks by small scales that are not as large as those on the tail. The number of scales in a transverse row across the mid-body is often 40–80 but figures may reach or even exceed 100 in some populations of the *Timon lepidus* group, or be as low as 25 in some *Zootoca*. Dorsal scales are frequently lightly keeled and more strongly so in forms like *Lacerta*, *Parvilacerta parva*, *Podarcis taurica*, *Timon princeps* and many *Zootoca* species. However, keeling is absent in many forms and, in ones that regularly use crevices as refuges, such as *Dalmatolacerta*, *Dinarolacerta* and *Iberolacerta horvathi*, the scales themselves are flattened. *Algyroides* and most *Takydromus* differ from all other Lacertini in having dorsal scales that are much bigger than those on the tail.
- *Femoral pores.* The femoral pores under the thigh are arranged in most Lacertini in a row numbering 7–31, beginning close to the midline of the body. Normally the row extends to the knee, but it is sometimes shortened distally, for example in *Darevskia derjugini* and some members of the *Lacerta trilineata* group. *Takydromus* is exceptional in often having just one or two femoral pores on each side or sometimes none at all; the maximum number found in this genus is five.
- *Toes.* The toes of Lacertini usually have one or two rows of unkeeled tubercular lamellae beneath, but the lamellae bear a single row of subdigital keels in *Apathya* and sometimes a faint double row in *Parvilacerta parva*.
- *Tail.* The tail scales are arranged in regular whorls, two to each vertebra. The whorls may be subequal in length but sometimes are alternatively markedly longer and shorter. The scales bordering the ventral mid-line of the tail are usually about the same width as neighbouring ones, but they are markedly expanded in *Dalmatolacerta* (Fig. 18) and to a lesser extent in some *Hellenolacerta*.



- *Colouring*. Lacertini exhibits considerable variation in their colouring and there are often extensive intraspecific differences, both between and within different populations.
- *Sexual dimorphism*. Some species show little difference between males and females in their dorsal colouring, but sexual dimorphism is sometimes well developed, like in *Podarcis*. It also occurs in varying extents in *Algyroides moreoticus*, some *Lacerta* populations (particularly well developed in *L. schreiberi* and *L. agilis*), and some populations of *Hellenolacerta*, *Darevskia*, *Iberolacerta* and *Teira*. Some Lacertini have broad dorsolateral stripes that are lighter than the flanks and mid-dorsum but sometimes consist of ground colour rather than being lighter than this; they occur in some *Anatololacerta*, *Apathya*, *Scelarcis*, *Takydromus* and *Teira*. Absence of striping in juveniles is uncommon but is found in *Archaeolacerta*, *Dalmatolacerta*, *Timon* and some *Dinarolacerta*, *Podarcis*, *Scelarcis* and many *Takydromus*. In these cases, animals are spotted or reticulated from hatching. Dorsal ground colour is often various shades of brown, buff or grey but it may be at least transiently bright green in many *Lacerta* and *Timon* and a wide though sporadic range of other forms. Melanism occurs sporadically in some taxa, but is especially common in *Podarcis* populations on small islands and in *Dalmatolacerta* at high altitudes. It arises in three different ways: general darkening of the ground colour (the commonest), increase in the number of dark markings, and the spread of those markings already present. These conditions were respectively named melanismus, abundismus and nigrismus, by Reinig (1937).

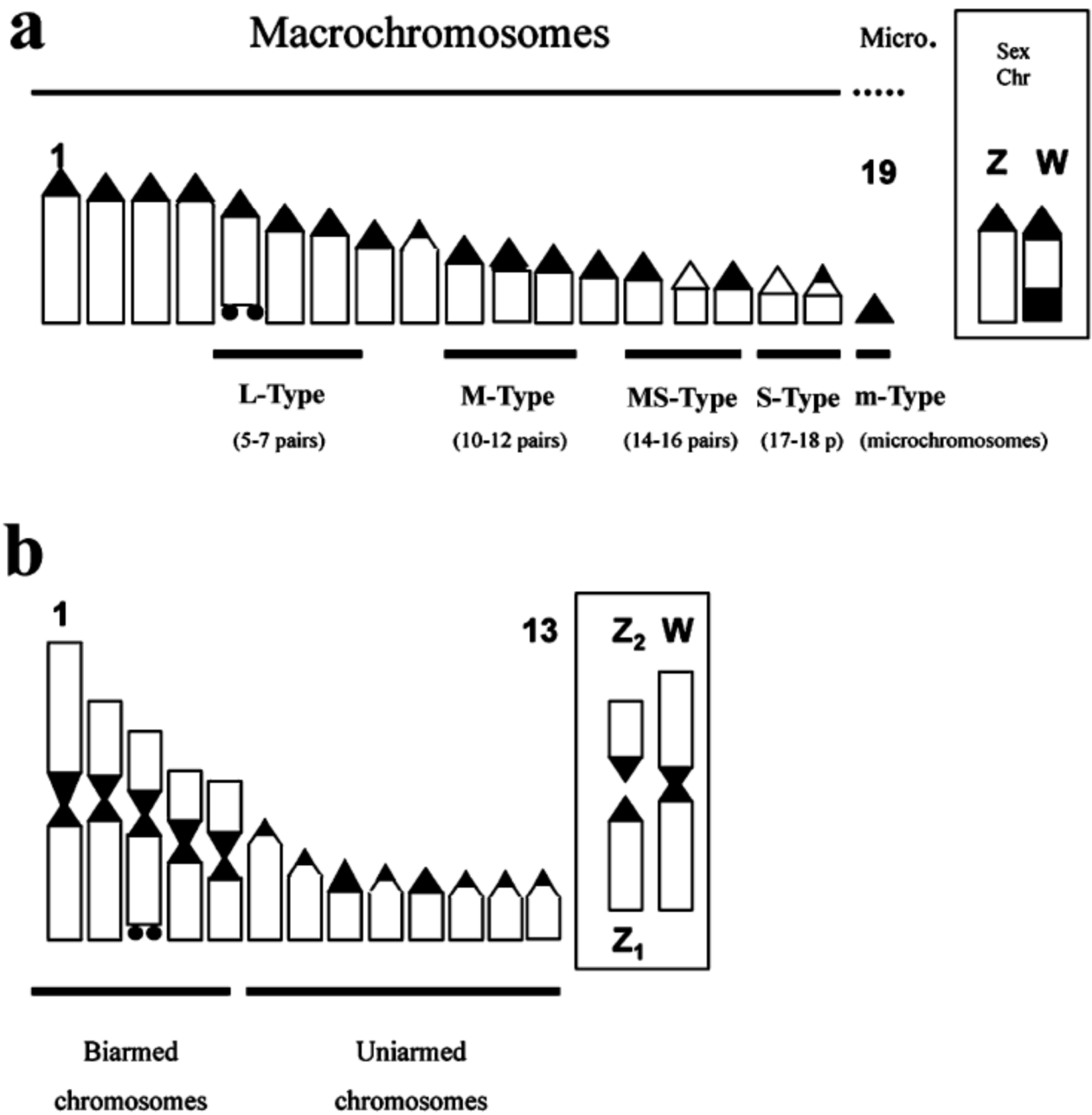


**Figure 9.** Common elements of patterning in lacertines: mid-back showing frequent positions of longitudinal stripes or rows of spots. From Arnold & Burton (1978).

- *Hemipenis*. Lobes having especially thin walls, which are complexly folded when the hemipenis is retracted.

*Chromosomes:*

The diploid number of chromosomes in Lacertini is usually 38, consisting of 36 single-armed macrochromosomes (otherwise known as uniarmed, acrocentric or subtelocentric) and two microchromosomes. The total number of chromosome arms is termed the Fundamental Number (FN), which is usually 38 in Lacertini (figure 10) (Arnold et al., 2007).



**Figure 10.** Diagrammatic representations of lacertine karyotypes (a single chromosome of each homologous pair is represented). (a) Common haploid condition with 18 single armed (that is uniarmed or acrocentric macrochromosomes) and one microchromosome, producing a diploid number of 38. The boxes in figure 10 illustrate the sex chromosomes, which are part of the usual complement of chromosomes illustrated to the left but revealed in cytological preparations by specific dyes; in this case they are of the ZW type, with ZZ males and ZW females. Chromosomes where the nucleolar organiser (NOR) may be situated are marked by horizontal bars and organisers are assigned to categories based on the size of these: if the NOR is situated in a macrochromosome this may be large (L-type), medium (M-type), medium-small (MS-type) or small (S-type); nucleolar organisers may also occur on the microchromosomes (m-type). In the present case the nucleolar organiser, indicated by two small black dots, is L-type. (b) Derived haploid condition with many double armed (that is biarmed, metacentric or submetacentric) chromosomes, based here on *Iberolacerta aurelioi*, but similar conditions occur in other Pyrenean *Iberolacerta* and in *Parvilacerta*. The five double-armed chromosomes

appear to have each been produced by fusion of two chromosomes (Robertsonian fusions). The sex chromosome system is  $Z_1Z_2W$ , in which males have  $Z_1Z_1Z_2Z_2$  and females  $Z_1Z_2W$ . Here the W chromosome is bi-armed as a result of Robertsonian fusion of two chromosomes, giving females one fewer chromosome than males (by Arnold et al. 2007).

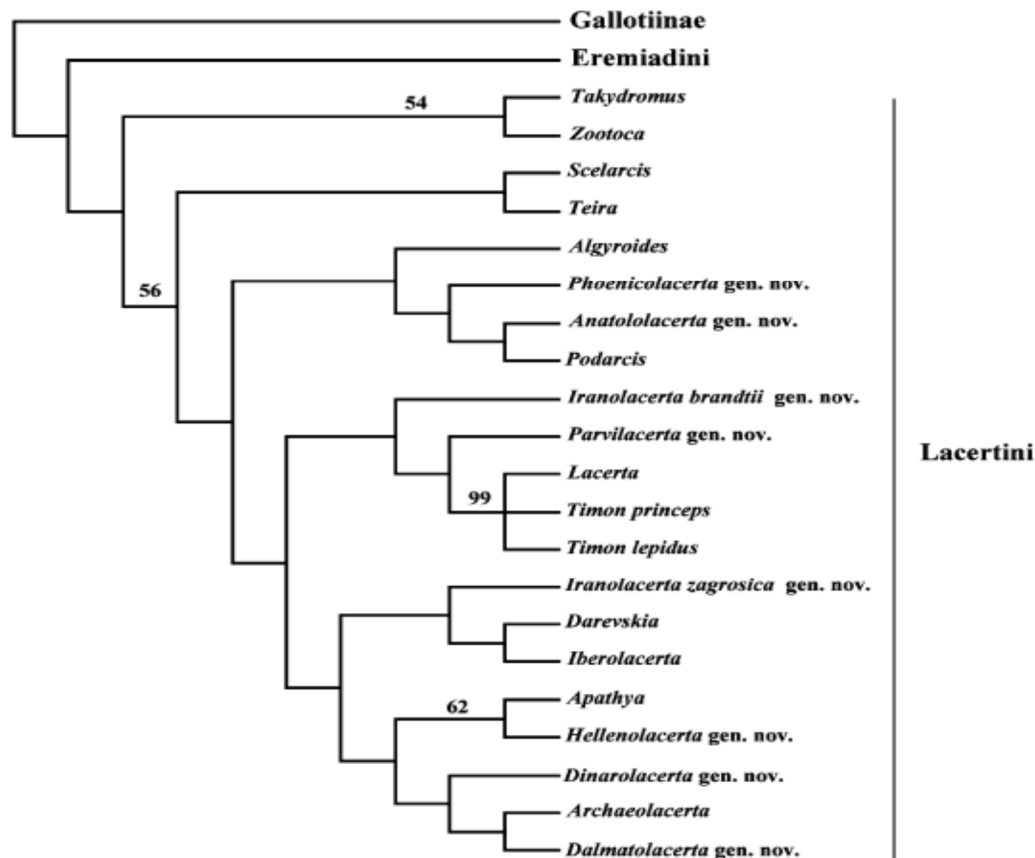
Conditions in other lacertids suggest this overall pattern is the primitive state in this tribe. *Iranolacerta brandtii* is distinctive in having a different arrangement of macrochromosomes, with 34 single-armed ones and a pair that are double-armed (metacentric). With the two microchromosomes, this may indicate a fundamental number of 40 ( $2n = 38$ ,  $FN = 40$ ), which is also found in some Gallotiinae (*Gallotia* -  $2n = 40$ ,  $FN = 40$ ; most *Psammodromus algirus* -  $2n = 38$ ,  $FN = 38$ , but others are reported as  $2n = 40$ ,  $FN = 40$ ). *Takydromus* has diploid numbers from 38 to 42, with 36 macrochromosomes and 2–6 microchromosomes. The diploid number is reduced to no more than 36 in *Zootoca* and *Iberolacerta* by loss of the microchromosomes. A similar degree of reduction in chromosome number occurs in *Timon* but here there has apparently been fusion (Robertsonian fusion) of two pairs of the single-armed macrochromosomes resulting in one pair of two-armed macrochromosomes. A greater degree of reduction is found in some Pyrenean *Iberolacerta*, resulting in diploid numbers of 26 or 24 in males and 26, 25 or 23 in females, the last being the lowest chromosome number known in lacertid lizards. In male *Iberolacerta* with reduced chromosome numbers, there are 16 or 12 single-armed macrochromosomes and 10 or 12 double-armed ones. Similarly, in *Parvilacerta* the diploid number is reduced to 24 by seven fusions so that there are eight single-armed macrochromosomes, 14 double armed ones and two microchromosomes. In *Darevskia*, hybrids between sexual males and parthenogenetic females may be triploid with  $3n = 57$  chromosomes. *Sex chromosomes.* Specific chromosomes among the total complement described above determine the sex of individual lizards. In the widespread ZW system, males have two Z chromosomes (ZZ in the diploid cells) and females one Z and one W chromosome (ZW in the diploid cells). In *Iberolacerta* with reduced chromosome numbers, the sex chromosome system is  $Z_1Z_2W$ , in which males have  $Z_1Z_1Z_2Z_2$  and females  $Z_1Z_2W$ . Populations of *Zootoca* exhibit a range of conditions, including the primitive ZW one in *Z. vivipara carniolica* and some Hungarian populations still considered as *Z. v. vivipara* (Odierna et al. 2004), and the  $Z_1Z_2W$  system across most of the vast distribution of the genus, from the egg-laying Iberian populations to the Pacific coast of Siberia and Sakhalin island (*Z. v. sachalinensis*). In the  $Z_1Z_2W$  system, females have a total chromosome number of 35 chromosomes rather than the 36 usual in males. The W chromosome of *Zootoca* is doublearmed in some populations but has single-armed by heterochromatinization and loss of chromosome fragments in others.

*Nucleolar organiser*. This is situated in chromosomes of different sizes in different species. It may occur

in large (L-type), medium (M-type), medium-small (MS-type) or small (S-type) macrochromosomes, or in a microchromosome (m-type). Details are given in Fig. 24. As, L-type nucleolar organisers are known in the Lacertini but are so far unrecorded from other lacertids, they are likely to represent a derived condition. In some individuals or populations of *Timon lepidus* group, a second nucleolar organiser may occur (Arnold et al., 2007 and references therein).

### ***3.5 Present assessments of relationships within the Lacertini***

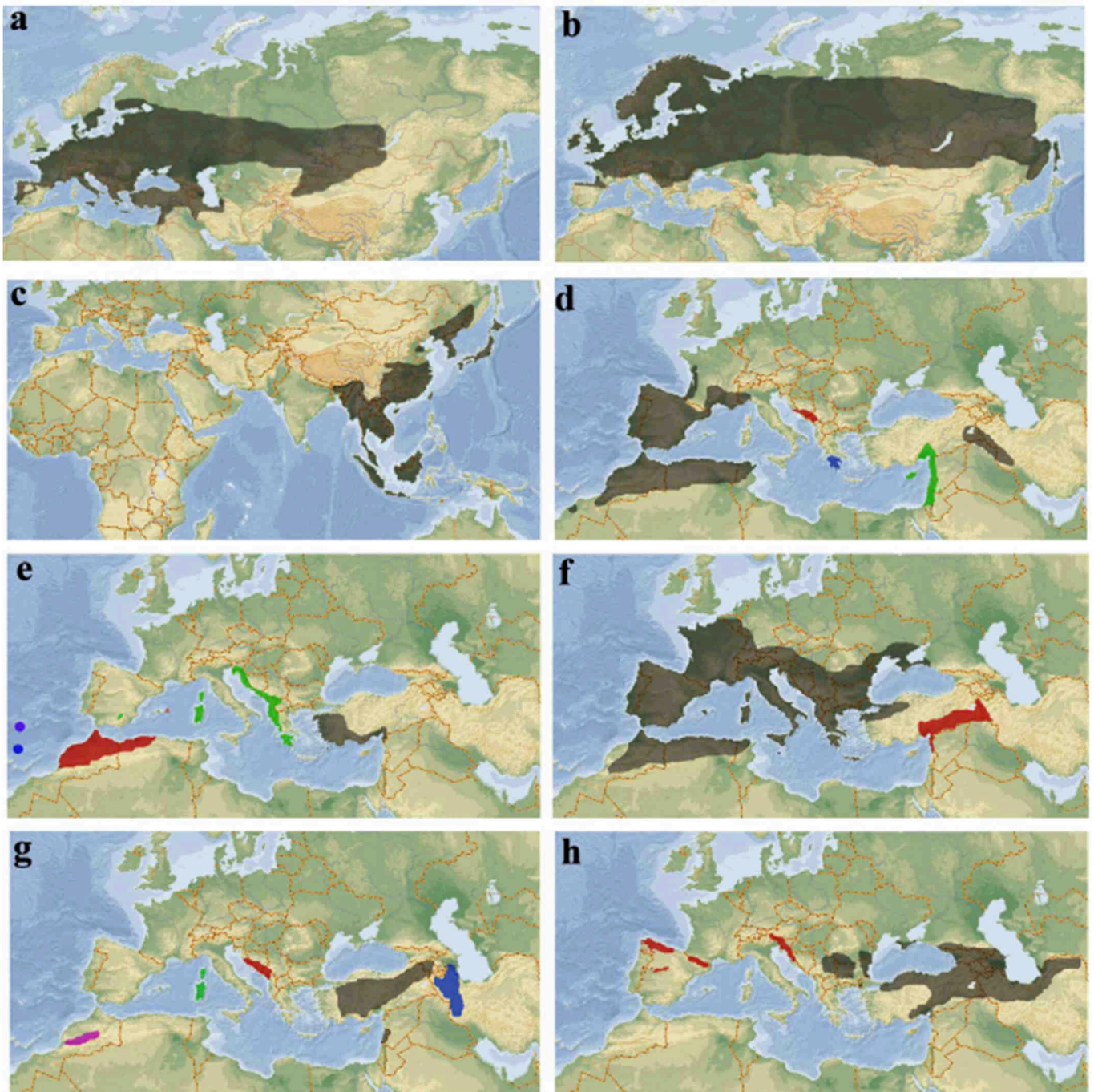
Taxonomic treatment of the species now assigned to Lacertini has varied considerably over time, but a number of assemblages are now commonly recognised. The fact that taxonomy was initially based on morphology and now it relies more on molecular data has inevitably led to confusion and instability in nomenclature. As a foundation for a more rational and comprehensive system of names for groups of Lacertini, we use new and old evidence for relationships within the group, based both on DNA sequences and morphology. The mitochondrial studies of Harris et al.(1998) and Fu (2000), and reanalyses of data presented by Arnold et al. (2007), support many of the recognised groupings within the Lacertini that have more than one species to corroborate the clade status of 19 groups (see Fig. 11) (Arnold et al. 2007).



**Figure 11.** Phylogeny of Lacertini based on a total of 64 morphological characters (58 parsimony-informative); equivalent to 83 binary characters.. Figures above the nodes indicate bootstrap support (from Arnold et al. 2007).

In all, 19 units were recognised which molecular clocks suggest separated 12–15 My ago (by Arnold et al., 2007).

If Europe is the source area for modern lacertids, there must have been several invasions of other regions. The molecular clock used here indicates that the Lacertini split into most of its component living genera 12–16 My ago, so they underwent quite rapid speciation at this time (but see Hipsley et al., 2009 for older divergence time estimates, around 43-46 mya). Most genera in the Lacertini have largely allopatric and often disjunct ranges, which may mean that initial spread of the group was followed or accompanied by multiple vicariance (see Fig. 12) (Arnold et al., 2007). A few units do not fit this pattern and have large ranges that overlap with several other taxa, although these too may have begun as vicariant isolates with small ranges and then spread (Arnold et al., 2007 and references therein).



**Figure 12.** Distribution maps of the 19 genera of Lacertini, and of *Atlantolacerta* (Eremiadini): (a) *Lacerta*; (b) *Zootoca*; (c) *Takydromus*; (d) *Timon* (black), *Dinarolacerta* gen. nov. (red), *Hellenolacerta* gen. nov. (blue), *Phoenicolacerta* gen. nov. (green); (e) *Teira* (blue), *Scelarcis* (red), *Algyroides* (green), *Anatololacerta* gen. nov. (black), (f) *Podarcis* (black), *Apathya* (red); (g) *Archaeolacerta* (green), *Dalmatolacerta* gen. nov. (red), *Parvilacerta* gen. nov. (black), *Iranolacerta* gen. nov. (blue); *Atlantolacerta* (purple); (h) *Iberolacerta* (red), *Darevskia* (black) (by Arnold et al. 2007).

### 3.6 The genus *Iberolacerta* Arribas, 1997

The genus *Iberolacerta* is a monophyletic group of rock lizards mainly distributed in highland areas of Western Europe (Fig. 13). This genus comprises 8 species, which can be subdivided into three



main units: (1) *I. horvathi*, occurring in the Eastern Alps and the north of the Dinaric Chains; (2) the subgenus Pyrenesaura, which includes the three species found in the Pyrenees Mountains, namely *I. aranica*, *I. aurelioi* and *I. bonnali*; and (3) the four species included in the ‘Iberian group’, i.e. *I. cyreni*, *I. martinezricai*, *I. galani*, and *I. monticola*, with disjunct distributions in central and northern mountain ranges of the Iberian Peninsula.

Previous cytogenetic surveys of the Iberolacerta species (Capula et al. 1989; Odierna et al. 1996; Arribas and Odierna 2005; Arribas et al. 2006; Rojo et al. 2014) showed a common diploid number of  $2n = 36$  and a similar karyotypic macrostructure, with all acrocentric chromosomes. Only the karyotypes of the three Pyrenean species differ from this formula, with reduced diploid numbers that range from  $2n = 24$  to  $26$  in males and from  $23$  to  $26$  in females and numerous biarmed chromosomes, which probably evolved from the ancestral acrocentric complement through a series of Robertsonian fusions (Odierna et al. 1996). Sex chromosomes of the ZW-type occur in *I. aranica*, *I. monticola*, *I. cyreni*, *I. martinezricai*, *I. galani* and *I. horvathi*, and of the  $Z_1Z_2W$ -type in *I. aurelioi* and *I. bonnali*, this latter system is correlated with lower chromosome numbers in females ( $25$  compared with  $26$  in *I. aurelioi* and  $23$  compared with  $24$  in *I. bonnali*). Nucleolar organizer are borne by large macrochromosome (L-type—*I. aranica*, *I. aurelioi*, *I. bonnali*, *I. monticola*, *I. galani*, *I. horvathi*) or by medium-sized chromosomes (M-type—*I. cyreni* and *I. martinezricai*) (Arnold et al., 2007 and references herein).

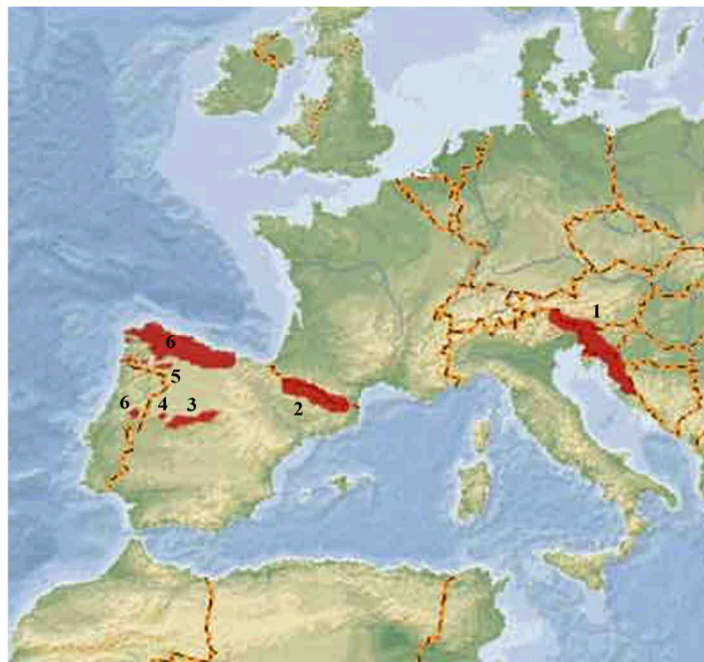


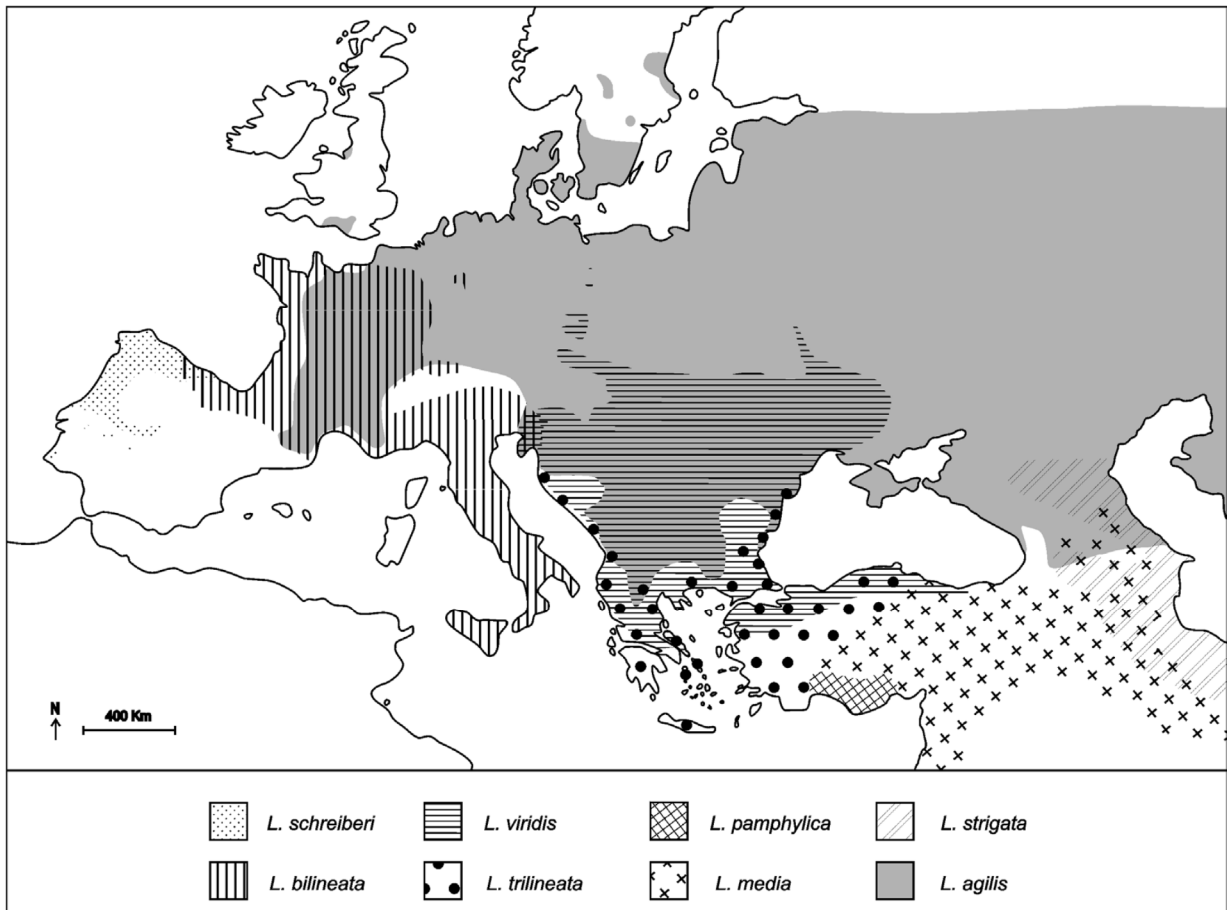
Figure 13. (Top) Distribution map for the genus *Iberolacerta*. 1: *I. horvathi*; 2: Pyrenean group; 3: *I. cyreni*; 4: *I. martinezricai*; 5: *I. galani*; 6: *I. monticola*. (Bottom) Adult male of *I. monticola* from the Natural Park of Fragas do Eume (A Coruña, Spain) (by Arnold et al. 2007).



The phylogeny of this genus has been under continual revision, but the evolutionary relationships among some taxa still remain unresolved (Mayer and Arribas, 2003; Carranza et al., 2004; Arribas et al., 2006). According to the most recently published phylogeny (Arribas et al., 2014), speciation within *Iberolacerta* started *ca.* 13.5 Million years ago (Mya; 95% confidence interval 11.6 – 15.6), with the split between the clades formed by *I. horvathi* and the Iberian group, on one side, and by the Pyrenean species, on the other. This event was rapidly followed, or almost simultaneous, to the divergence of *I. horvathi*, which took place approximately 11.5 Mya (9.6 – 13.7). Within the Iberian group, *I. cyreni* split earlier, from 7.3 to 8.5 Mya, while the speciation events within the clade formed by *I. martinezricai*, *I. galani* and *I. monticola* occurred considerably later, at the beginning of the Pleistocene, roughly 2.5 Mya (2.1 – 2.9). The three Pyrenean species probably originated in rapid succession *ca.* 3.8 Mya (2.7 – 4.9), although this phylogenetic analysis suggests that *I. bonnali* split first, shortly before the separation between *I. aranica* and *I. aurelioi*, 3.3 Mya (2.3 – 4.3).

### **3.7 The genus *Lacerta* (Linneus, 1758)**

Green lizards of the genus *Lacerta sensu stricto* inhabit a large area extending from Western Europe to Central Asia. With the exception of the more widespread *Lacerta agilis*, they are almost restricted to the southern European peninsulas well known for exceptional richness in biodiversity. They are essentially parapatric, with relatively small areas where two or rarely three species are found and each species substitutes the other in a clinal continuum throughout the Mediterranean belt (fig. 14). They are easily identifiable from other lizards by their relatively large size and often brilliant colour, but some are difficult to discriminate from each other phenotypically (Godinho et al., 2005 and references herein). This genus comprises 8 species: *Lacerta agilis* Linnaeus, 1758; *Lacerta bilineata* Daudin, 1802; *Lacerta media* Lantz and Cyren, 1920; *Lacerta pamphylica* Schmidtler, 1975; *Lacerta schreiberi* Bedriaga, 1878; *Lacerta strigata* Eichwald, 1831; *Lacerta trilineata* Bedriaga, 1886; *Lacerta viridis* (Laurenti, 1768).



**Figure 14.** Map showing the distribution of green lizard species in Europe and Western Asia (adapted from Gasc et al., 1998 and Nettmann, 2001).

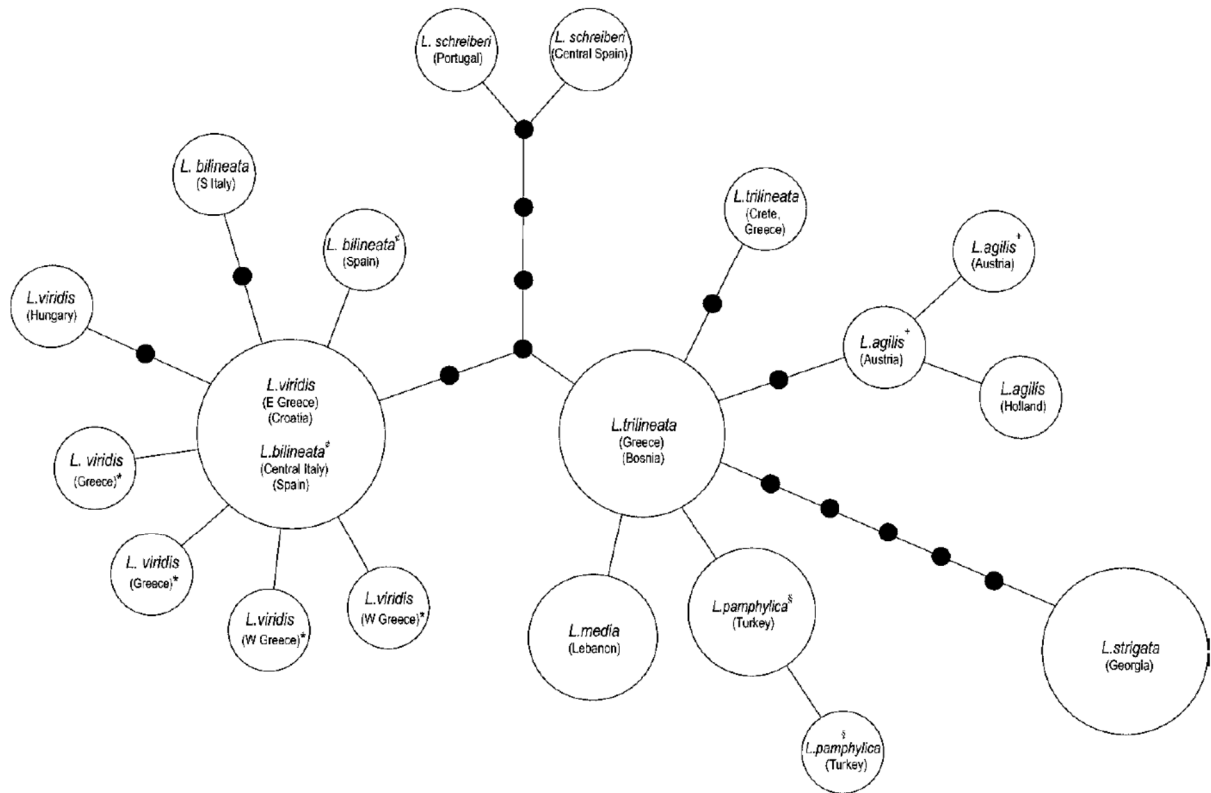
Based on morphological characters Arnold (1973) suggested that green lizards and the *Lacerta lepida* group (*L. lepida*, *L. pater* and *L. princeps*) formed a clade. Later, Rykena and Nettmann (1986) showed that the *L. lepida* group is strongly supported as a monophyletic clade by a chromosomal arrangement unique among lacertids (36 diploid with two biarmed chromosomes instead of the typical 38 diploid number). Subsequently, evidence from mtDNA sequences suggested that the *L. lepida* group is the sister clade to the green lizards (Godinho et al., 2005 and references herein).

No cladistic attempt has previously been made to reconstruct the phylogeny of the green lizards group. Using the micro-complement fixation technique, Lutz and Mayer (1985) indicated that *L. trilineata* and *L. viridis* are more closely related to each other than to *L. agilis*, but did not examine other species. Later, Rykena (1996) suggested a distant relationship of *L. schreiberi* to the other green lizards based on hybridization experiments. Similarly, Brückner et al. (2001) described a close association of *L. viridis* with *L. bilineata* relative to *L. trilineata* using mtDNA sequences and Mayer and Beyerlein (2001) described the paraphyly of *L. trilineata* cluster in relation to *L.*

*pamphylica*. However, no other green lizards were included in these studies. Therefore, at present, the phylogeny of the group is unresolved. Furthermore, it is unclear whether species and subspecies defined on the basis of morphological and ecological features are well supported by molecular data. To address this question we have used new and published sequences from three mtDNA (cytochrome b, 12S rRNA and 16S rRNA) and two nuclear ( $\beta$ -fibrinogen intron 7 ( $\beta$ -fibint7) and C-mos) genes to assess the phylogeny of all eight species of green lizards, including many of the distinct subspecies (fig. 15) (Godinho et al., 2005 and references herein).

The phylogenetic inferences derived from the different mtDNA partial genes are very similar and support the monophyly of *Lacerta sensu stricto*, *L. agilis*, *L. schreiberi* and *L. strigata* are all strongly supported as distinct genetic units based on the separate mtDNA data sets (bootstrap support between 95-100%) but their relationship relative to the other green lizards is not well supported by the mtDNA sequence data. The species status of *L. strigata* has been questioned by some authors due to its ecological and morphological similarities to *L. viridis* (Schmidtler, 1986b). However, hybridization experiments (Rykena, 1996, 2001) and the high genetic separation based on the DNA sequence data suggest it is a distinct species. The species status of *L. strigata* has been questioned by some authors due to its ecological and morphological similarities to *L. viridis* (Schmidtler, 1986b). However, hybridization experiments (Rykena, 1996, 2001) and the high genetic separation based on the DNA sequence data suggest it is a distinct species (Godinho et al., 2005 and references herein). The results obtained in these two species could indicate an ancient introgression (exchange of genetic material from one species to another, only partially insulated from the first through interspecific hybridization), confirmed by the fact that it is known the hybridisation between the two (Godinho et al., 2005).

Concordant with their overall morphological similarity, *L. trilineata*, *L. media* and *L. pamphylica* as a clade is supported by the 12S and 16S rRNA sequence data. Of these three species, *L. media* is the sister clade to the other two, thus constituting a group paraphyletic (Godinho et al., 2005).



**Figure 15.** Most parsimonious network of C-mos sequences. There were no homoplastic characters. Size of circles is relative to the frequency of haplotypes. Filled circles indicate presumed missing haplotypes. \* indicates estimated haplotypes. §, \* and + indicate two haplotypes from the same individual ( by Godinho et al., 2005).

Seven out of the eight species that form the green lizard group are almost exclusively restricted to the southern European peninsulas or the Caucasian mountains while a single one, *L. agilis* occupies Central Asia, nearly all of Europe, including parts of England and southern regions of Scandinavia, but is not present in the most southerly refugia. Kalyabina et al. (2001) produced a phylogeographic analysis of *L. agilis* using mtDNA and suggested that the three subspecies that represent the European genetic lineages of *L. agilis* come from a Balkan-Carpathian refugium, after a radiation from the Caucasian mountains in the Late Pliocene where the species presumably originated (Bischoff, 1988). According to this scenario, the colonization of Europe after the last glacial age, where a single species colonized the central and northern regions of the continent coming from one of the southern refugia (Godinho et al., 2005 and references herein).

## 4. References

- ADAMS M.D., CELNIKER S.E., HOLT R.A., EVANS C.A., GOCAYNE J.D., AMANATIDES P.G. ET AL. (2000). The genome sequence of *Drosophila melanogaster*. *Science*, 287: 2185–2195
- ARAVIN A.A., LAGOS-QUINTANA M., YALCIN A., ZAVOLAN M., MARKS D., SNYDER B., GAASTERLAND T., MEYER J., TUSCHL T. (2003). The small RNA profile during *Drosophila melanogaster* development. *Cell*. 5:337–350
- ARMOUR J. A. L., JEFFREYS A. J. (1992). Biology and applications of human minisatellite loci. *Current Opinion in Genetics e Development*, 2: 850-856.
- ARNASON U., GRETARSDOTTIR S., WIDEGREN B. (1992). Mysticete (baleen whale) relationship based upon the sequence of the common cetacean DNA satellite. *Mol. Biol. Evol.*, 9: 1018-1028.
- ARNOLD, E.N. (1973). Relationships of the Palaearctic lizards assigned to the genera *Lacerta*, *Algyroides* and *Psammodromus* (Reptilia, Lacertidae). *Bulletin of the British Museum, London (Zoology)*, 29: 289–366.
- ARNOLD, E.N. (1986). The hemipenis of lacertid lizards (Reptilia: Lacertidae): structure, variation and systematic implications. *Journal of Natural History*, 20: 1221–1257.
- ARNOLD E.N. (1989a). Towards a phylogeny and bigeography of the Lacertidae: relationships within an Old-world family of lizards derived from morphology. *Bulletin of the British Museum, London (Zoology)*, 55: 209–257.
- ARNOLD E.N., ARRIBAS O., CARRANZA S. (2007). Systematics of the Paleaertic and Oriental lizard tribe Lacertini (Squamata: Lacertidae: Lacertinae) with descriptions of eight new genera. *Zootaxa*, 1430: 1-86.
- ARNEY K.L., FISHER A.G. (2004). Epigenetic aspects of differentiation. *J. Cell Sci.*, 117: 4355-4363.
- ARRIBAS O., ODIERNA G. (2005). Karyological and osteological data supporting the specific status of *Iberolacerta (cyreni) martinezricai* (Arribas, 1996). *Amphibia-Reptilia* 25: 359- 367.
- ARRIBAS O., CARRANZA S., ODIERNA G. (2006). Description of a new endemic species of mountain lizard from Northwestern Spain: *Iberolacerta galani* sp. nov. (Squamata: Lacertidae). *Zootaxa* 1240: 1-55.
- BACHMANN L., SPERLICH D. (1993). Gradual evolution of a specific satellite DNA family in *Drosophila ambigua*, *D. tristis*, and *D. obscura*. *Mol. Biol Evol.*, 10: 647-659.

- BACHTROG D.** (2006). Expression profile of a degenerating neo-Y chromosome in *Drosophila*. *Curr. Biology*, 16: 1694–1699.
- BACHTROG D., HOM E., WONG K.M., MASIDE X., DE JONG P.** (2008). Genomic degradation of a young Y chromosome in *Drosophila miranda*. *Genome Biol.*, 9: R30.
- BACHTROG D.** (2013). Y-chromosome evolution: emerging insights into processes of Y-chromosome degeneration. *Nat Rev Genet.*, 14: 113-124.
- BASSI, C., ET AL.** (2000). Molecular structure and evolution of DNA sequences located at the alpha satellite boundary of chromosome 20. *Gene*, 256: 43–50.
- BERGERO R., FORREST A., KAMAU E., CHARLESWORTH D.** (2007). Evolutionary strata on the X chromosomes of the dioecious plant *Silene latifolia*: Evidence from new sex-linked genes. *Genetics*, 175: 1945-1954.
- BERGERO R., CHARLESWORTH D., FILATOV D.A., MOORE R.C.** (2008). Defining regions and rearrangements of the *Silene latifolia* Y chromosome. *Genetics*, 178: 2045-2053.
- BISCHOFF W.** (1988). Zur verbreitung und systematik der zauneidechse, *Lacerta agilis* Linnaeus, 1758. *Mertensiella*, 1: 11-30.
- BLOOM K.** (2007). Centromere dynamics. *Curr. Opin. Genet. Dev.*, 17: 151-156.
- BÖHME W.** (1971) Über das Stachelepithel am Hemipenis lacertider Eidechsen und seine systematische Bedeutung. *Zeitschrift für Zoologische Systematik und Evolutionsforschung*, 9: 187–223.
- BORSUK-BIALYNICKA M., LUBKA M., BÖHME W.** (1999). A lizard from Baltic amber (Eocene) and the ancestry of the crown group of lacertids. *Acta Palaeontologica Polonica*, 44: 349–382.
- BRÜCKNER M., DÜRING A.** (2001). PCR-RFLP: A fast and inexpensive biochemical method for detecting species boundaries in the *Lacerta viridis/bilineata* species complex. *Mertensiella*, 13: 40-44.
- BRUFORD M. W., WAYNE R. K.** (1993). Microsatellites and their application to population genetic studies. *Current Opinion in Genetics e Development*, 3: 939–943.
- BULL J.J., VOGT R.C., BULMER M.G.** (1982) Heritability of sex ratio in turtles with environmental sex determination. *Evolution*, 36, 333-341.
- BULL J.J.** (1983). *Evolution of sex determining mechanisms*. Menlo Park, CA: Benjamin/Cummings.

- CANAPA A., BARUCCA M., NISI CERIONI P., OLMO E.** (2000). A satellite DNA containing CENP-B box-like motifs is present in the antarctic scallop *Adamussium colbecki*. *Gene*, 247: 175-180.
- CAPULA M., LAPINI L., CAPANNA E.** (1989). The karyotype of *Lacerta horváthi* (Reptilia, Sauria, Lacertidae) *Genetica* 79: 11-16.
- CARRANZA, S. & WADE, E.** (2004) Taxonomic revision of Algero-Tunisian *Pleurodeles* (Caudata: Salamandridae) using molecular and morphological data. Revalidation of the taxon *Pleurodeles nebulosus* (Guichenot, 1850). *Zootaxa*, 488, 1–24.
- CARRANZA S., ARNOLD E.N., AMAT F.** (2004) DNA phylogeny of *Lacerta (Iberolacerta)* and other lacertine lizards (Reptilia: Lacertidae): did competition cause long-term mountain restriction? *Systematics and Biodiversity*, 2, 57–77.
- CAVALIER-SMITH T.** (1985). *The Evolution of Genome Size* (ed. Cavalier-Smith, T.) 105–184 (John Wiley, Chichester).
- CESARI M., LUCHETTI A., PASSAMONTI M., SCALI V., MANTOVANI B.** (2003). PCR 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.** (1991). The evolution of sex chromosomes. *Science*, 251: 1030–1033.
- CHARLESWORTH B., SNEGOWSKI P., STEPHAN W.** (1994). The evolutionary dynamics of repetitive DNA in eukaryotes. *Nature*, 371: 215–220.
- CONOVER D.O., HEINS S.W.** (1987) Adaptive variation in environmental and genetic sex determination in a fish. *Nature*, **326**, 496-498.
- CSINK A.K., HENIKOFF S.** (1998). Something from nothing: the evolution and utility of satellite repeats. *Trends Genet.*, 14: 200–204.
- DI RENZO A., PETERSON A. C., GARZA J. C., VALDES A. M., SLATKIN M., FREIMER N. B.** (1994). Mutational processes of simple-sequence repeat loci in human populations. *Proceedings of the National Academy of Sciences of the United States of America*, 91: 3166-3170.
- DOOLITTLE W.F., SAPIENZA C.** (1980). Selfish genes, the phenotype paradigm and genome evolution. *Nature*, 284: 601-603.
- DOVER G.A.** (1982). Molecular drive: a cohesive mode of species evolution. *Nature*, 299: 111-117.
- DOVER G.A., FLAVELL R.B.,** (1984). Molecular coevolution: DNA divergence and the maintenance of function. *Cell*, 38: 623-624.
- DOVER G.A.** (1986). Molecular drive in multigene families: how biological novelties arise, spread and are assimilated.. *Trends Genet.*, 2: 159-165.
- DOVER G.A.** (1987). DNA turnover and the molecular clock. *J. Mol. Evol.* 26: 47–58.

- DOVER G.A.** (2002). Molecular drive. *Trends Genet.*, 18: 587-589.
- DROUIN G.** (2006). Chromatin diminution in the copepod *Mesocyclops edax*: diminution of tandemly repeated DNA families from somatic cells. *Genome* 49, 657–665.
- DURFY S.J., WILLARD H.F.**, (1989). Patterns of intra- and interarray sequence variation in alpha-satellite from the human X chromosome: evidence for short-range homogenization of tandemly repeated DNA sequences. *Genomics*, 5: 810–821.
- ELDER JR J.F., TURNER B.J.** (1995). Concerted evolution of repetitive DNA sequences in eukaryotes. *Q. Rev. Biol.*, 70: 297–320.
- ELLINGSEN A., SLAMOVITS C.H., ROSSI M.S.**, (2007). Sequence evolution of the major satellite DNA of the genus *Ctenomys* (Octodontidae, Rodentia). *Gene*, 392: 283–289.
- EPPLEN J.T., MCCARREY J.R., SUTOU S., OHNO S.** (1982). Base sequence of a cloned snake W-chromosome DNA fragment and identification of a male-specific putative mRNA in the mouse. *Proc. Natl. Acad. Sci. USA . Genetics*, 79: 3798-3802.
- ESTES R., DE QUEIROZ K., GAUTHIER J.** (1988). Phylogenetic relationships within Squamata. In: *Phylogenetic relationships of the lizard families* (Estes, R. & Pregill, G. Eds), Stanford University Press: Stanford, California, pp. 119–281.
- EZAZ T, QUINN A.E., MIURA I., SARRE S.D., GEORGES A., MARSHALL GRAVES J.A.** (2005). The dragon lizard *Pogona vitticeps* has ZZ/ZW micro-sex chromosomes. *Chromosome Res.*, 13: 763–776.
- EZAZ T., SARRE S.D., O’MEALLY D., GRAVES J.A.M.**, (2009a). Sex Chromosome Evolution in Lizards: Independent Origins and Rapid Transitions. *Cytogenetic and Genome Research*, 127: 249-260.
- FESCHOTTE C., KESWANI U., RANGANATHAN N., GUIBOTSY M.L., LEVINE D.** (2009). Exploring repetitive DNA landscapes using REPCLASS, a tool that automates the classification of transposable elements in eukaryotic genomes. *Genome Biol. Evol.*, 1: 205-220.
- FLAVELL B.** (1982). Sequence amplification, deletion and rearrangement: major sources of variation during species divergence. *Genome Evolution*, 301-323.
- FRANCIS R.C.** (1992). Sexual lability in teleosts: developmental factors. *The Quarterly Review of Biology*, 67: 1-18.
- FRANCIS R.C., BARLOW G.W.** (1993) Social control of primary sex differentiation in the Midas cichlid. *Proceedings of the National Academy of Sciences*, 90: 10673-10675.



- FRIDOLFSSON A.K., CHENG H., COPELAND N.G., JENKINS N.A., LIU H.C., RAUDSEPP T., ET AL.** (1998). Evolution of the avian sex chromosomes from an ancestral pair of autosomes. *Proc Natl Acad Sci USA*, 95: 8147–8152
- FU J.** (2000). Toward the phylogeny of the family Lacertidae—Why 4708 base pairs of mtDNA sequences cannot draw the picture. *Biological Journal of the Linnean Society*, 71:203-217.
- GARRIDO-RAMOS M.A., DE LA HERRAN R., JAMILENA R., LOZANO R., RUIZ-REJON C., RUIZ-REJON M.,** (1999). Evolution of centromeric satellite DNA and its use in phylogenetic studies of the Sparidae family (Pisces, Perciformes). *Mol. Phylogenet. Evol.*, 12: 200–204.
- GEMAYEL R., CHO J., BOEYNAEMS S., VERSTREPEN K. J.,** (2012). Beyond Junk-Variable Tandem Repeats as Facilitators of Rapid Evolution of Regulatory and Coding Sequences. *Genes*, 3: 461-480.
- GRAVES J.** (2005). Recycling the Y chromosome. *Science*, 307: 50–51.
- GRAVES J.** (2008). Weird Animal Genomes and the Evolution of Vertebrate Sex and Sex Chromosomes. *Annu. Rev. Genet.*, 42: 565-586.
- GREGORY, T.R., ET AL.,** 2007. Eukaryotic genome size databases. *Nucleic Acids Res.* 35, D332–D338.
- GREWAL S.I., ELGIN S.C.** (2007). Transcription and RNA interference in the formation of heterochromatin. *Nature*, 447: 399–406.
- GLINKA S., DE LORENZO D., STEPHAN W.** (2006). Evidence of gene conversion associated with a selective sweep in *Drosophila melanogaster*. *Mol. Biol. Evol.*, 23: 1869–1878.
- GODINHO R., CRESPO E. G., FERRAND N. D., HARRIS J.** (2005). Phylogeny and evolution of the green lizards, *Lacerta* spp. (Squamata: Lacertidae) based on mitochondrial and nuclear DNA sequences. *Amphibia-Reptilia* , 26: 271-285.
- GORMAN G.C.** (1973). The Chromosomes of the Reptilia, a Cytotaxonomic Interpretation, in Chiarelli AB, Capanna, E (eds): *Cytotaxonomy and Vertebrate Evolution*, 349–424. (Academic Press, London, New York).
- GORMAN G.C., ATKINS L.** (1966). Chromosomal heteromorphism in some male lizards of the genus *Anolis* . *Am Nat.*, 100: 579–583.
- GOTO R., KAYABA T., ADACHI S., YAMAUCHI K.** (2000) . Effects of temperature on sex determination in marble sole *Limanda yokohamae*. *Fisheries Science*, 66, 400-402.
- HALL S.E., KETTLER G., PREUSS D.,** (2003). Centromere Satellites From Arabidopsis populations: maintenance of conserved and variable domains. *Genome Res.*, 13: 195–205.
- HALL A.E., KEITH K.C., HALL S.E., COPENHAVER G.P., PREUSS D.,** (2004). The rapidly evolving field of plant centromeres. *Curr. Opin. Plant Biol.*, 7: 108–114.

- HALL S.E., LUO S., HALL A.E., PREUSS D.** (2005). Differential rates of local and global homogenization in centromere satellites from Arabidopsis relatives. *Genetics*, 170: 1913–1927.
- HANDLEY L.J., CEPLITIS H., ELLEGREN H.** (2004). Evolutionary strata on the chicken Z chromosome: implications for sex chromosome evolution. *Genetics*, 167:367-376.
- HARRIS D.J., ARNOLD E.N., THOMAS R.H.** (1998). Relationships of lacertid lizards (Reptilia: Lacertidae) estimated from mitochondrial DNA sequences and morphology. *Proceeding of Royal Society London B.*, 265: 1939–1948.
- HARTL, D.L.** (2000). Molecular melodies in high and low C. *Nat. Rev. Genet.*, 1: 145-149.
- HEIKKINEN E., LAUNONEN V., MULLER E., BACHMANN L.,** (1995). The pvB370 BamHI satellite DNA family of the *Drosophila virilis* group and its evolutionary relation to mobile dispersed genetic pDv elements. *J. Mol. Evol.*, 41: 604–614.
- HEITZ E.** (1928). Das Heterochromatin der Moose. *I. Jahrb. Wiss. Bot.*, 69: 762–818.
- HILLIER L.W., MILLER W., BIRNEY E., WARREN W., HARDISON R.C., PONTING C.P., ET AL** (2004). Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. *Nature*, 432: 695–716.
- HIPSLEY C. A., HIMMELMANN L., METZLER D., MÜLLER J.**(2009). Integration of Bayesian molecular clock methods and fossil-based soft bounds reveals early Cenozoic origin of African lacertid lizards. *BMC Evolutionary Biology*, 9:151
- HOBZA R., KEJNOVSKY E., VYSKOT B., WIDMER A.** (2007). The role of chromosomal rearrangements in the evolution of *Silene latifolia* sex chromosomes. *Mol Genet Genomics*, 278: 633-638.
- HOLMES S. E., DOMBROWSKI B. A., KREBS C. M., BOEHM C. D. KAZAZIAN H. G.** (1994). A new retrotransposable human L1 element from the LRE2 locus on chromosome 1q produces a chimaeric insertion. *Nature Genet.*, 7: 143–148.
- HOLMGREN K., MOSEGAARD H.,** (1996). Implications of individual growth status on the future sex of the European eel. *Journal of Fish Biology*, 49: 910-925.
- HOUCADE D., DRESSLER D., WOLFSON J.** (1973). The amplification of ribosomal RNA genes involves a rolling circle intermediate. *Proc. Nat. Acad. Sci. USA* vol. 70, 10: 2926-2930.
- IKENO M., MASUMOTO H., OKAZAKI T.** (1994). Distribution of CENP-B boxes reflected in CREST centromere antigenic sites on long-range alpha-satellite DNA arrays of human chromosome 21. *Hum. Mol. Genet.*, 3: 1245–1257.
- JABLONKA E.** (2004). The evolution of the peculiarities of mammalian sex chromosomes: an epigenetic view. *Bio Essays*, 26: 1327–1332.

- JABS E.W., GOBLE C. A., CUTTING G. R.** (1989). Macromolecular organization of human centromeric regions reveals high-frequency, polymorphic macro DNA repeats. *Proceedings of the National Academy of Sciences of the United States of America*, 86: 202-206.
- JANZEN F.J.** (1992). Heritable variation for sex ratio under environmental sex determination in the common snapping turtle (*Chelydra serpentina*). *Genetics*, 131, 155-161.
- JOHN B., MIKLOS, G. L. G.** (1988). *The Eukaryote Genome in Development and Evolution* (Allen and Unwin, London), p 305.
- JONES K.W., SINGH L.** (1981). Conserved repeated DNA sequences in vertebrate sex chromosomes. *Human Genetics*, 58: 46-53.
- JONES K.W., SINGH L.** (1985) . Snakes and the evolution of sex chromosomes. *Trends in Genetics*, 1: 55-61.
- JUST W., BAUMSTARK A., SUSS A., GRAPHODATSKY A., RENS W., SCHAFFER N., BAKLOUSHINSKAYA I., HAMEISTER H., VOGEL W.** (2007) . *Ellobius lutescens*: sex determination and sex chromosome ex Dev., 1:211-221.
- KALYABINA S.A., MILTO K.D., ANANJEVA N.B., LEGAL L., JOGER U., WINK M.** (2001). Phylogeography and systematics of *Lacerta agilis* based on mitochondrial cytochrome b gene sequences: first results. *Russ. J. Herp.*, 8: 149-158.
- KAPLI P., POULAKAKIS N., LYBERASKIS P., MYLONAS M.** (2011). A re-analysis of the molecular phylogeny of Lacertidae with currently available data. *Basic and Applied Herpetology*, 25: 97-104.
- KRZYWINSKI J., HOBZA R., CERMAK T., KUBAT Z., VYSKOT B.** (2009). The role of repetitive DNA in structure and evolution of sex chromosomes in plants. *Heredity*, 102: 533–541
- KIT, S.** (1961). Equilibrium sedimentation in density gradients of DNA preparations from animal tissues. *J. Mol. Biol.*, 3: 711-716.
- KLEMMER, K.** (1957). Untersuchungen zur Osteologie und Taxionomie der europäischen Mauereidechsen. *Abhandlungen von der Senckenbergischen Naturforschenden Gesellschaft*, 496: 1–56.
- KOHN M., KEHRER-SAWATZKI H., VOGEL W., GRAVES J.A.M., HAMEISTER H.** (2004). Wide genome comparisons reveal the origins of the human X chromosome. *Trends Genet.*, 20: 598-603.

- KUHL D. P. A., CASKEY C. T.** (1993). Trinucleotide repeats and genome variation. Trinucleotide repeats and genome variation Trinucleotide repeats and genome variation Trinucleotide repeats and genome variation *Curr. Opin. Genet. Dev.*, 3: 404–407
- LAKICH D., KAZAZIAN H. H., ANTONARAKIS S. E. & GITSCHIER J.** (1993). Inversions disrupting the factor VIII gene are a common cause of severe haemophilia A. *Nature Genet.*, 5: 236–241
- LAHN BT, PAGE D.C.** (1999). Retroposition of autosomal mRNA yielded testis-specific gene family on human Y chromosome. *Nature Genet.*, 21: 429–43.
- LEE H.R., ZHANG W., LANGDON T., JIN W., YAN H., CHENG Z., JIANG J.** (2005). Chromatin immunoprecipitation cloning reveals rapid evolutionary patterns of centromeric DNA in *Oryza* species. *Proc. Natl. Acad. Sci. USA*, 102: 11793–11798.
- LEE H.R., NEUMANN P., MACAS J., JIANG J.** (2006). Transcription and evolutionary dynamics of the centromeric satellite repeat CentO in rice. *Mol. Biol. Evol.*, 23: 2505–2520.
- LI Y.X., KIRBY M.L.** (2003). Coordinated and conserved expression of alphoid repeat and alphoid repeat-tagged coding sequences. *Dev Dynamics*, 228:72–81.
- LIEW W.C., BARTFAI R., LIM Z., SREENIVASAN R., SIEGFRIED K.R., ET AL.** (2012) Polygenic Sex Determination System in Zebrafish. *PLoS ONE*, 7(4): e34397.  
doi:10.1371/journal.pone.0034397.
- LOHE A. R., HILLIKER A. J., ROBERTS P. A.** (1993). Mapping simple repeated DNA sequences in heterochromatin of *Drosophila melanogaster*. *Genetics*, 134:1149–1174.
- LORITE P., RENAULT S., ROULEUX-BONNIN F., BIGOT S., PERIQUET G., PALOMEQUE T.** (2002). Genomic organization and transcription of satellite DNA in the ant *Aphaenogaster subterranea* (Hymenoptera, Formicidae). *Genome*, 45: 609–616.
- LORITE P., CARRILLO J.A., TINAUT A., PALOMEQUE T.** (2004). Evolutionary dynamics of satellite DNA in species of the genus *Formica* (Hymenoptera, Formicidae). *Gene*, 332: 159–168
- LUCHETTI A., CESARI M., CARRARA G ET AL.** (2003) Unisexuality and molecular drive: Bag320 sequence diversity in *Bacillus taxa* (Insecta Phasmatodea). *J Mol Evol* 56: 587-589.
- LUTZ, D., MAYER, W.** (1985). Albumin Evolution and its Phylogenetic and Taxonomic implications in several Lacertid Lizards. *Amphibia-Reptilia* 7: 53-61.
- MACHA J., TEICHMANOVA R., SATER A.K., WELLS D.E., TLAPAKOVA T. ET AL.** (2012). Deep ancestry of mammalian X chromosome revealed by comparison with the basal tetrapod *Xenopus tropicalis*. *BMC Genomics*, 13:315.

- MACKAY T.F.C.** (1986). Transposable element-induced fitness mutations in *Drosophila melanogaster*. *Genet.Res.*, 48: 77-87.
- MANTOVANI, B.** (1998). Satellite sequence turnover in parthenogenetic systems: the apomictic triploid hybrid *Bacillus lynceorum* (Insecta, Phasmatodea). *Mol. Biol. Evol.*, 15:1288–1297.
- MARAIS GAB, NICOLAS M, BERGERO R, CHAMBRIER P, KEJNOVSKY E, MONEGER F ET AL.** (2008). Evidence for degeneration of the Y chromosome in the dioecious plant *Silene latifolia*. *Current Biology*, 18: 1–5.
- MARTIENSSEN R.A.** (2003). Maintenance of heterochromatin by RNA interference of tandem repeats. *Nat. Genet.*, 35: 213–214.
- MASUMOTO H., NAKANO M., OHZEKI J.** (2004). The role of CENP-B and alpha-satellite DNA: de novo assembly and epigenetic maintenance of human centromeres. *Chromosom. Res.*, 12: 543–556.
- MATSUNAGA S, ISONO E, KEJNOVSKY E, VYSKOT B, DOLEZEL J, KAWANO S ET AL.** (2003). Duplicative transfer of a MADS box gene to a plant Y chromosome. *Mol. Biol. Evol.*, 20: 1062–1069.
- MAHTANI M.M., WILLARD H.F.** (1990). Pulsed-field gel analysis of alpha-satellite DNA at the human X chromosome centromere: high-frequency polymorphisms and rray size estimate. *Genomics*, 7: 607–613.
- MAHTANI M.M., WILLARD H.F.** (1998). Physical and genetic mapping of the human X chromosome centromere: repression of recombination. *Genome Res.*, 8: 100–110.
- MASHKOVA, T., ET AL.** (1998). Unequal crossing-over is involved in human alpha satellite DNA rearrangements on border of the satellite domain. *FEBS Lett.*, 441: 451–457.
- MAYER W., ARRIBAS O.** (2003). Phylogenetic relationships of the European lacertid genera *Archaeolacerta* and *Iberolacerta* and their relationships to some other “*Archaeolacertae*” (sensu lato) from Near East, derived from mitochondrial DNA sequences. *J Zool Syst. Evol. Res.*, 41: 157-161.
- MAYER W., BEYERLEIN P.** (2001). Genetic differentiation of the *Lacerta viridis/bilineata* complex and of *Lacerta trilineata* in Greece: mitochondrial DNA sequences. *Mertensiella*, 13: 52-59.
- MAYER W., BENYR G.** (1994). Albumin-Evolution und Phylogenese in der Familie Lacertidae. *Annual Naturhistorisches Museum Wien*, 96B: 621-648.
- MAYER W., PAVLICEV M.** (2007). The phylogeny of the family Lacertidae (Reptilia) based on nuclear DNA sequences: convergent adaptations to arid habitats within the subfamily Eremiainae. *Mol Phylogenet Evol.*, 44: 1155-1163.

- MATSUBARA K., TARUI H., TORIBA M., YAMADA K., NISHIDA-UMEHARA C. ET AL.** (2006). Evidence for different origin of sex chromosomes in snakes, birds, and mammals and step-wise differentiation of snake sex chromosomes. *Proc. Natl. Acad. Sci. USA*, 103:18190–95.
- MATSUNAGA S., E. ISONO, E. KEJNOVSKY, B. VYSKOT, J. DOLEZEL ET AL.** (2003). Duplicative transfer of a MADS box gene to a plant Y chromosome. *Mol. Biol. Evol.*, 20: 1062–1069.
- MATSUNAGA S.** (2009). Junk DNA promotes sex chromosome evolution. *Heredity*, 102 : 525–526.
- MCALLISTER BF, WERREN JH** (1999) Evolution of Tandemly Repeated Sequences: What Happens at the End of an Array? *J. Mol. Evol.*, 48: 469–481.
- MEŠTROVIĆ N., CASTAGNONE-SERENO P., PLOHL M.** (2006a). High conservation of the differentially amplified MPA2 satellite DNA family in parthenogenetic root–knot nematodes. *Gene*, 376: 260–267.
- MEŠTROVIĆ N., CASTAGNONE-SERENO P., PLOHL M.** (2006b). Interplay of selective pressure and stochastic events directs evolution of the MEL172 satellite DNA library in root–knot nematodes. *Mol. Biol. Evol.*, 23: 2316–2325.
- MODI V.S., CREWS D.** (2005). Sex chromosomes and sex determination in reptiles. *Current Opinion in Genetics & Development*, 15:660–665.
- MORITZ C.** (1984). The evolution of a highly variable sex chromosome in *Gehyra purpurascens* (Gekkonidae). *Chromosoma*, 90: 111–119.
- MRAVINAC B., PLOHL M., MEŠTROVIĆ N., UGARKOVIĆ D.** (2002). Sequence of PRAT satellite DNA “frozen” in some Coleopteran species. *J. Mol. Evol.*, 54: 774–783.
- MRAVINAC B., UGARKOVIĆ Đ., FRANJEVIĆ D., PLOHL M.** (2005). Long inversely oriented subunits form a complex monomer of *Tribolium brevicornis* satellite DNA. *J. Mol. Evol.*, 60: 513–525.
- NAM K., ELLEGREN H.** (2008). The chicken (*Gallus gallus*) Z chromosome contains at least three nonlinear evolutionary strata. *Genetics*, 180: 1131-1136.
- NAGAKI K., TALBERT P.B., ZHONG C.X., DAWE R.K., HENIKOFF S., JIANG J.**(2003). Chromatin immunoprecipitation reveals that the 180-bp satelliterepeat is the key functional DNA element of *Arabidopsis thaliana* centromeres. *Genetics*, 163: 1221–1225.
- NANDA I., ZEND-AJUSCH E., SHAN Z., GRUTZNER F., SCHARTL M., ET AL.** (2000). Conserved synteny between the chicken Z sex chromosome and human chromosome 9 includes the male regulatory gene DMRT1 : A comparative (re)view on avian sex determination. *Cytogenet. Cell. Genet.*, 89: 67–78

- NANDA I., HAAF T., SCHARTL M., SCHMID M., BURT D.W.** (2002). Comparative mapping of Z-orthologous genes in vertebrates: implications for the evolution of avian sex chromosomes. *Cytogenet. Genome Res.*, 99: 178–184.
- NATRI H.M., SHIKANO T., MERILA J.** (2013). Progressive recombination suppression and differentiation in recently evolved neo-sex chromosomes. *Mol Biol Evol* 30: 1131–1144.
- ODIERNA G., APREA G., ARRIBAS O., CAPRIGLIONE T., CAPUTO V., OLMO E.** (1996). The karyology of the Iberian rock lizards. *Herpetologica*, 52: 542–550.
- OHNO S.** (1967) *Sex Chromosomes and Sex Linked Genes*. Springer-Verlag: New York.
- OHNO S.** (1972). So much 2junk2 DNA in our genome. *Brookhaven Symp. Biol.*, 23: 366–370.
- ORGEL L.E., CRICK F.H.** (1980). Selfish DNA: the ultimate parasite. *Nature*, 284: 604–607.
- OHTA T., DOVER G.A.** (1984). The cohesive population genetics of molecular drive. *Genetics*, 108: 501–521.
- OLMO E., SIGNORINO G.** (2005). Chromorep: a reptile chromosomes database. Internet references: <http://193.206.118.100/professori/chromorep.pdf>, 15.09.06.
- O'MEALLY D., PATEL HR, STIGLEC R, SARRE, SD, GEORDES A, GRAVES, JAM, EZAZ T.** (2010). Nonhomologous sex chromosomes of birds and snakes share repetitive sequences. *Chromosome Res.*, 18: 787–800.
- ORGAN C., JANES D.E.** (2008). Sex chromosome evolution in sauropsids. *Integr. Comp. Biol.*, 48: 512–519.
- ORGAN C.L., MORENO R.G., EDWARDS S.V.** (2008). Three tiers of genome evolution in reptiles. *Integr, Comp. Biol.*, 48: 494–504.
- PARASNIS A.S., RAMAKRISHNA W., CHOWDARI K.V., GUPTA V.S., RANJEKAR P. K.** (1999). Microsatellite (GATA) n reveals sex-specific differences in papaya. *Theoretical and applied genetics*, 99(6): 1047–1052.
- PAQUES F., LEUNG W.Y., HABER J.E.** (1998). Expansions and contractions in a tandem repeat induced by double-strand break repair. *Mol. Cell. Biol.*, 18(4): 2045–54
- PATHAK D., SRIVASTAVA J., PREMI S., TIWARI M., GARG L. C., KUMAR S., ALI S.** (2006). Chromosomal localization, copy number assessment, and transcriptional status of BamHI repeat fractions in water buffalo *Bubalus bubalis*. *DNA Cell. Biol.*, 25: 206–214.
- PERRIN N.** (2009). Sex reversal: a fountain of youth for sex chromosomes? *Evolution* 63: 3043–3049.
- PERRY J., ASHWORTH A.** (1999). Evolutionary rate of a gene affected by chromosomal position. *Curr Biol.*, 9: 987–989.

- PEZER Ž., BRAJKOVIĆ J., FELICIELLO I., UGARKOVIĆ D.** (2012). Satellite DNA-mediated effects on genome regulation. *Genome Dyn.*, 7: 153–169.
- GARRIDO-RAMOS, M.A., DE LA HERRAN, R., JAMILENA, R., LOZANO, R., RUIZ-REJON, C., RUIZ-REJON, M.** (1999). Evolution of centromeric satellite DNA and its use in phylogenetic studies of the Sparidae family (Pisces, Perciformes). *Mol. Phylogenet. Evol.*, 12: 200–204.
- PLOHL M., LUCHETTI A., MEŠTROVIĆ N., MANTOVANI B.** (2008). Satellite DNAs between selfishness and functionality: Structure, genomics and evolution of tandem repeats in centromeric (hetero)chromatin. *Gene*, 409: 72– 82.
- PLOHL M., MEŠTROVIĆ N., MRAVINAC B.** (2012) Satellite DNA evolution. In: Garrido-Ramos MA, ed. Repetitive DNA. Genome dynamics vol 7. Basel: Karger, 126–152.
- POKORNÁ M., KRATOCHVÍL L.** (2009). Phylogeny of sex-determining mechanisms in squamate reptiles: are sex chromosomes an evolutionary trap? *Zoological Journal of the Linnean Society*, 156: 168–183.
- POKORNÁ M., GIOVANNOTTI M., KRATOCHVÍL L., KASAI F, TRIFONOV V.A., O'BRIEN P.C., CAPUTO V., OLMO E.,FERGUSON-SMITH M.A., RENS W.** (2011). Strong conservation of the bird Z chromosome in reptilian genomes is revealed by comparative painting despite 275 million years divergence. *Chromosoma*, 120: 455-468.
- PONS J., PETITPIERRE E. AND JUAN C.** (2002). Evolutionary dynamics of satellite DNA family PIM357 in species of the genus *Pimelia* (Tenebrionidae, Coleoptera). *Mol. Biol. Evol.*, 19, 1329-1340.
- QUINN AE, GEORGES A, SARRE SD, GUARINO F, EZAZ T, GRAVES JAM.** 2007. Temperature sex reversal implies sex gene dosage in a reptile. *Science*, 316:411
- RADDER RS, QUINN AE, GEORGES A, SARRE SD, SHINE R.** (2008). Genetic evidence for co-occurrence of chromosomal and thermal sex-determining systems in a lizard. *Biology Letters*, 4: 176–178.
- REINIG W.F.** (1937). *Melanismus, Albinismus und Rufinismus. Ein Beitrag zum Problem der Entstehung und Bedeutung tierischer Färbungen.* Georg Thieme Verlag. Leipzig.: 122.
- RIEPEL, O.** (1980) The trigeminal jaw adductor musculature of *Tupinambis*, with comments on the phylogenetic relationships of the Teiidae (Reptilia, Lacertilia). *Zoological Journal of the Linnean Society*, 69: 1–29.
- RYKENA, S.** (1996). Experimental interspecific hybridization in the genus *Lacerta*. *Isr. J. Zoo.*, 42: 171-184.



- RYKENA, S.** (2001). Experimental hybridization in Green Lizards (*Lacerta s. str.*), a tool to study species boundaries. *Mertensiella*, 13: 78-88.
- ROBLES F., DE LA HERRA ´N R., LUDWIG A., RUIZ REJO ´N C., RUIZ REJO ´N M., GARRIDO-RAMOS M.A.** (2004). Evolution of ancient satellite DNAs in sturgeon genomes. *Gene*, 338: 133-142.
- RODIONOV A, CHELYSHEVA LA, SOLOVEI IV, MIAKOSHINA IUA.** (1992) : Chiasma distribution in the lampbrush chromosomes of the chicken *Gallus g allus domesticus* : Hot spots of recombination and their possible role in proper disjunction of homologous chromosomes at the first meiotic division. *Genetika*, 28: 151–160.
- ROIZES G.** (2006). Human centromeric alphoid domains are periodically homogenized so that they vary substantially between homologues. Mechanism and implications for centromere functioning. *Nucleic Acids Res.*, 34: 1912–1924.
- ROJAS A.A., VAZQUEZ-TELLO A., FERBEYRE G., VENANZETTI F., BACHMANN L., PAQUIN B., SBORDONI V., CEDERGREN R.** (2000). Hammerhead-mediated processing of satellite pDo500 family transcripts from Dolichopoda cave crickets. *Nucleic Acids Res.*, 28: 4037–4043.
- ROJO V., GIOVANNOTTI M., NAVEIRA H., NISI CERIONI P., GONZÁLEZ\_TIZÓN A.M., CAPUTO BARUCCHI V., GALÀN P., OLMO E., MARTÍNEZ\_LANGE A.** (2014). Karyological characterization of the endemic Iberian Rock lizard, *Iberolacerta monticola* (Squamata, Lacertidae): Insights into sex chromosome evolution. *Cytogenet Genome Res* 142: 28–39.
- RÖMER U., BEISENHERTZ W.** (1996). Environmental determination of sex in *Apistogramma* (Cichlidae) and two other freshwater fishes (Teleostei). *Journal of Fish Biology*, 48: 714-725.
- RUDERT F., BRONNER S., GARNIER J.M., DOLLÉ P.** (1995). Transcripts from opposite strands of gamma satellite DNA are differentially expressed during mouse development. *Mamm. Genome*, 6: 76–83.
- RUDD M.K., WRAY G.A., WILLARD H.F.** (2006). The evolutionary dynamics of alpha-satellite. *Genome Res.*, 16: 88–96.
- SARRE S.D., GEORGES A., QUINN A.** (2004). The ends of a continuum: genetic and temperature-dependent sex determination in reptiles. *BioEssays* 26:639–45
- SCHUELER M.G., HIGGINS A.W., RUDD M.K., GUSTASHAW K., WILLARD H.F.** (2001). Genomic and genetic definition of a functional human centromere. *Science*, 294: 109–115.
- SCHUELER M.G., DUNN J.M., BIRD C.P., ROSS M.T., VIGGIANO L., ROCCHI M., WILLARD H.F., GREEN E.D.** (2005). Progressive proximal expansion of the primate X chromosome centromere. *Proc. Natl. Acad. Sci. USA*, 102: 10563–10568.

- SCHINDELHAUER D., SCHWARZ T.** (2002). Evidence for a fast, intrachromosomal conversion mechanism from mapping of nucleotide variants within a homogeneous alpha-satellite DNA array. *Genome Res.*, 12: 1815–1826.
- SCHMIDT T., HESLOP-HARRISON J.S.** (1998). Genomes, genes and junk: the largescale organization of plant chromosomes. *Trends Plant Sci.*, 3: 195–199.
- SCHMIDTLER J.F.** (1986b). Orientalische Smaragdeidechsen: 1. Zur systematik und verbreitung von *Lacerta viridis* in der Türkei (Sauria: Lacertidae). *Salamandra* 22: 29-46.
- SCHUELER M.G., HIGGINS A.W., RUDD M.K., GUSTASHAW K., WILLARD H.F.** (2001). Genomic and genetic definition of a functional human centromere. *Science*, 294: 109–115.
- SCHUELER M.G., DUNN J.M., BIRD C.P., ROSS M.T., VIGGIANO L., ROCCHI M., WILLARD H.F., GREEN E.D.** (2005). Progressive proximal expansion of the primate X chromosome centromere. *Proc. Natl. Acad. Sci. USA*, 102: 10563–10568.
- SHIELS C., COUTELLE C., HUXLEY C.** (1997). Contiguous arrays of satellites 1, 3, and beta form a 1.5-Mb domain on chromosome 22p. *Genomics*, 44: 35–44.
- SHINE R., ELPHICK MJ, DONNELLAN S.** (2002). Co-occurrence of multiple, supposedly incompatible modes of sex determination in a lizard population. *Ecology Letters* 5: 486–489.
- SINGH N., AGARWAL A.K., LATA A., KOHLI R.P.** (1976). Evaluation of ‘adaptogenic’ properties of *Withania somnifera*. *Proc. Indian Pharmacological Society*, 17.
- SKALETSKY H., KURODA-KAWAGUCHI T., MINX P.J., CORDUM H.S., HILLIER L., BROWN L.G. ET AL.** (2003). The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. *Nature*, 423: 825–837.
- SLAMOVITS C.H., COOK J.A., LESSA E.P., ROSSI M.S.** (2001). Recurrent amplifications and deletions of satellite DNA accompanied chromosomal diversification in South American Tuco-tucos (Genus *Ctenomys*, Rodentia: Octodontidae): a phylogenetic approach. *Mol. Biol. Evol.*, 18: 1708–1719.
- SMITH G.P.** (1976). Evolution of repeated DNA sequences by unequal crossover. *Science*, 191: 528–535.
- SMITH C.A., CLIFFORD V., WESTERN P.S., WILCOX S.A., BELL K.S., SINCLAIR A.H.** (2000). Cloning and expression of a *DAX1* homologue in the chicken embryo. *J Mol Endocrinol*, 24: 23–32.
- STEPHAN W.** (1986). Recombination and the evolution of satellite DNA. *Genet Res* 47: 167–174.
- STEPHAN W.** (1989). Quantitative variation and chromosomal location of satellite DNAs. *Genet Res* 50: 41–52.

- STRACHAN T., WEBB D., DOVER G.A.** (1985). Transition stages of molecular drive in multiple-copy DNA families in *Drosophila*. *EMBO J.*, 4: 1701–1708.
- SUN X., WAHLSTROM J., KARPEN G.** (1997). Molecular structure of a functional *Drosophila* centromere. *Cell* 91: 1007–1019.
- SZYBALSKI W.** (1968). Use of cesium sulfate for equilibrium density gradient centrifugation. *Methods Enzymol.* 12B, 330–360.
- TYLER-SMITH C., WILLARD H. F.** (1993). Mammalian chromosome structure. *Curr. Opin. Genet. Dev.*, 3:390–397.
- TOPP C.N., ZHONG C.X., DAWE R.K.** (2004). Centromere-encoded RNAs are integral components of the maize kinetochore. *Proc. Natl. Acad. Sci. U. S. A.*, 101: 15986–15991.
- TOWNSEND, T.M., LARSON, A., LOUIS, E. & MACEY, J.R.** (2004) Molecular phylogenetics of Squamata: the position of snakes, amphisbaenians and dibamids, and the root of the squamate tree. *Systematic Biology* 53, 735–757.
- VALENZUELA N., LANCE L.** (2004). *Temperature-Dependent Sex Determination in Vertebrates*. Washington: Smithsonian Books.
- VIDAL, N., HEDGES, S.B.** (2004). Molecular evidence for a terrestrial origin of snakes. *Proceedings of the Royal Society of London, B (Supplement)*, 271, 226–229.
- UETZ P., HOŠEK J.** (2015). The reptile database. Available at <http://www.reptile-database.org>. Accessed 1 June 2015.
- UGARKOVIC D., PLOHL M.** (2002). Variation in satellite DNA profiles—causes and effects. *EMBO J.* 21, 5955– 5959.
- UGARKOVIĆ D.** (2009) Centromere-competent DNA: structure and evolution. *Prog Mol Subcell Biol* 48: 53–76.
- VALENZUELA N., ADAMS D.C., JANZEN F.J.** (2003). Pattern does not equal process: exactly when is sex environmentally determined? *The American Naturalist* 161: 676–683.
- VALENZUELA N, LANCE L** (2004). *Temperature-Dependent Sex Determination in Vertebrates*. Washington: Smithsonian Books.
- VAN BELKUM A., SCHERER S., VAN ALPHEN L., VERBRUGH H.** (1998). Short-sequence DNA repeats in prokaryotic genomes. *Microbiol. Mol. Biol. Rev.* 62: 275-293.
- VARADARAJ K., SKINNER D.M.** (1994). Cytoplasmic localization of transcripts of a complex G+C-rich crab satellite DNA. *Chromosoma (Berlin)*, 103: 423–431.

- VELTSOS P., KELLER I., NICHOLS R.A.** (2008). The inexorable spread of a newly arisen neo-Y chromosome. *PLoS Genetics* 4: 1–9.
- VERSHININ A.V., ALKHIMOVA E.G., HESLOP-HARRISON J.S.** (1996). Molecular diversification of tandemly organized DNA sequences and heterochromatic chromosome regions in some Triticeae species. *Chromosom. Res.* 4: 517–525.
- VERSTREPEN K.J., JANSEN A., LEWITTER F., FINK G.R.** (2005). Intragenic tandem repeats generate functional variability. *Nat Genet.*, 37(9): 986–990.
- VIDAL N., HEDGES S.B.** (2004). Molecular evidence for a terrestrial origin of snakes. *Proc. R. Soc. Lond. B., (Suppl.)* 271: S226–S229.
- VICOSO B, EMERSON JJ, ZEKTSER Y, MAHAJAN S, BACHTROG D** (2013a) Comparative sex chromosome genomics in snakes: Differentiation, evolutionary strata, and lack of global dosage compensation. *PLoS Biol.*, 11: e1001643.
- VOLPE T.A., KIDNER C., HALL I.M., TENG G., GREWAL S.I.S.** (2002). Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. *Science*, 297: 1833–1837.
- WANG J.P., NA J.K., YU Q.Y., GSCHWEND A.R., HAN J. ET AL.** (2012) Sequencing papaya X and Y-h chromosomes reveals molecular basis of incipient sex chromosome evolution. *Proc Natl Acad Sci U S A*, 109: 13710-13715.
- WEVRICK R., WILLARD H. F.** (1989). Long-range organization of tandem arrays of  $\alpha$  satellite DNA at the centromeres of human chromosomes: high-frequency array-length polymorphism and meiotic stability. *Proceedings of the National Academy of Sciences of the United States of America*, 86: 9394-9398.
- WALLACE M.R., ANDERSEN L. B. , SAULINO A. M. , GREGORY P. E. , GLOVER T. W., COLLINS F. S.** (1991). A de novo Alu insertion results in neurofibromatosis type 1. *Nature*, 353: 864-865.
- WARBURTON P.E., WILLARD H.F.** (1990). Genomic analysis of sequence variation in tandemly repeated DNA. Evidence for localized homogeneous sequence domains within arrays of alpha-satellite DNA. *J. Mol. Biol.*, 216: 3–16.
- WESTERMAN M, HEWITT G.M.** (1985). Chromosome banding in *Podisma pedestris*. *Heredity* 55: 157–161.
- WILLARD H.F., WAYE J.S.** (1987). Chromosome-specific subsets of human alpha satellite DNA: analysis of sequence divergence within and between chromosomal subsets and evidence for an ancestral pentameric repeat. *J Mol. Evol.*, 25: 207–214.

**WU, J., ET AL.** (2004). Composition and structure of the centromeric region of rice chromosome 8. *Plant Cell* 16, 967–976.

**WITSCHI E.** (1929). Studies on sex differentiation and sex determination in amphibians. II. Sex reversal in female tadpoles of *Rana sylvatica* following the application of high temperature.

*Journal of Experimental Zoology* 52: 267–291.

**ZLUVOVA J, JANOUSEK B, NEGRUTIU I, VYSKOT B** (2005). Comparison of the X and Y chromosome organization in *Silene latifolia*. *Genetics* 170: 1431

## *Aims of research*

## Aims of research

Considering the well-known usefulness of satDNAs in facing phylogenetic issues (i.e., Martinsen et al., 2009), the aim of the present paper was to isolate and characterise satDNAs in some lacertid species in order to:

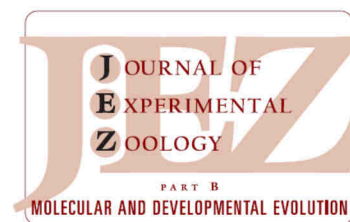
- I. increase the knowledge of this genomic elements in an important amniote group for which data on occurrence, genomic distribution, and evolutionary rates are limited to a handful of species;
- II. use the satDNAs isolated to verify the robustness of the proposed phylogenetic reconstruction for some *Iberolacerta* and *Lacerta* taxa on the light of independent molecular markers;
- III. to analyze the patterns of sequence variability, genomic organization, and chromosomal distribution of two satellite DNA families in all eight *Iberolacerta* species in order to understand the processes that determine the structure and evolutionary dynamics of these repetitive elements, and their possible role in chromosomal evolution;
- IV. evaluate the possible involvement of satDNA repeats, widespread in the lacertid genome, in the differentiation of W sex chromosome in the studied species.

# ***CHAPTER I***

**Isolation and Characterization of Two Satellite DNAs in some  
Iberian Rock Lizards (Squamata, Lacertidae).**



# Isolation and Characterization of Two Satellite DNAs in some Iberian Rock Lizards (Squamata, Lacertidae)



MASSIMO GIOVANNOTTI<sup>1</sup>, VERÓNICA ROJO<sup>2</sup>,  
PAOLA NISI CERIONI<sup>1</sup>, ANA GONZÁLEZ-TIZÓN<sup>2</sup>,  
ANDRÉS MARTÍNEZ-LAGE<sup>2</sup>, ANDREA SPLENDIANI<sup>1</sup>,  
HORACIO NAVEIRA<sup>2</sup>, PAOLO RUGGERI<sup>1</sup>,  
ÓSCAR ARRIBAS<sup>3</sup>, ETTORE OLMO<sup>1</sup>,  
AND VINCENZO CAPUTO BARUCCHI<sup>1,4\*</sup>

<sup>1</sup>Dipartimento di Scienze della Vita e dell'Ambiente, Università Politecnica delle Marche, Ancona, Italy

<sup>2</sup>Departamento de Biología Celular y Molecular, Universidade da Coruña, La Coruña, Spain

<sup>3</sup>Avda. Francisco Cambó 23, Barcelona, Spain

<sup>4</sup>Consiglio Nazionale delle Ricerche, Istituto di Scienze Marine Sezione Pesca Marittima, Ancona, Italy

## ABSTRACT

Satellite DNAs represent a large portion of all high eukaryotic genomes. They consist of numerous very similar repeated sequences, tandemly arranged in large clusters up to 100 million base pairs in length, usually located in the heterochromatic parts of chromosomes. The biological significance of satDNAs is still under discussion, but most of their proposed functions are related to heterochromatin and/or centromere formation and function. Because information about the structure of reptilian satDNA is far from exhaustive, we present a molecular and cytogenetic characterization of two satDNA families in four lacertid species. Two families of tandemly repeated DNAs, namely *TaqI* and *HindIII* satDNAs, have been cloned and sequenced from four species belonging to the genus *Iberolacerta*. These satDNAs are characterized by a monomer length of 171–188 and 170–172 bp, and by an AT content of 60.5% and 58.1%, respectively. FISH experiments with *TaqI* satDNA probe produced bright signals in pericentromeric regions of a subset of chromosomes whereas all the centromeres were marked by *HindIII* probe. The results obtained in this study suggest that chromosome location and abundance of satDNAs influence the evolution of these elements, with centromeric families evolving tenfold faster than interstitial/pericentromeric ones. Such different rates render different satellites useful for phylogenetic investigation at different taxonomic ranks. *J. Exp. Zool. (Mol. Dev. Evol.)* 9999B: 1–14, 2013. © 2013 Wiley Periodicals, Inc.

*J. Exp. Zool.*  
(*Mol. Dev. Evol.*)  
9999B:1–14, 2013

How to cite this article: Giovannotti M, Rojo V, Nisi Cerioni P, González-Tizón A, Martínez-Lage A, Splendiani A, Naveira H, Ruggeri P, Arribas Ó, Olmo E, Caputo Barucchi V. 2013. Isolation and characterization of two satellite DNAs in some Iberian rock lizards (Squamata, Lacertidae). *J. Exp. Zool. (Mol. Dev. Evol.)* 9999B:1–14.

Satellite DNAs (satDNAs) form a substantial part of eukaryotic genomes and consist of tandemly repeated DNA sequences typically arranged in large clusters of hundreds or thousands of copies usually located in the heterochromatic regions of chromosomes, mainly in the regions close to the centromeres and telomeres. The biological significance of satDNAs remains intriguing and challenging. The sequence conservation of some satellites over long evolutionary times, the presence of differentially expressed transcripts in several species and interactions with centromeric-specific proteins (e.g., the histone H3 variant CENH3) suggest a biological role for some satellites, although this is not fully understood (see Plohl et al., 2008; Plohl, 2010).

A satDNA family could arise in a phylogenetically short period by explosive amplification (Bachmann and Sperlich, '93) and afterwards its repeats could follow a gradual mode of sequence evolution during a long evolutionary time (Bachmann and Sperlich, '93). The processes by which satDNA families arise are not well known. A set of molecular-exchange mechanisms has been proposed to account for its origin by amplification of a tandem array of multi-copy sequences. These mechanisms include unequal crossing-over (Smith, '76), transposition (Miller et al., 2000), or extrachromosomal rolling-circle replication and reintegration of tandem arrays into the genome (Felicciello et al., 2006). A recently originated tandem array is initially homogeneous in sequence because of the multi-copy amplification of the same repeat. In the course of time, random mutations would accumulate and the repeats would diverge. However, the nonallelic repeats of a satDNA family do not evolve independently, but concertedly leading to near homogeneity for species-specific mutations (Bachmann and Sperlich, '93; Rudd et al., 2006). This phenomenon, known as concerted evolution, is achieved by a number of genomic mechanisms, mainly unequal crossing-over, biased gene conversion, slippage replication, and amplification by rolling-circle (Dover, '82; Walsh, '87; Charlesworth et al., '94). However, the rates of sequence change (homogenization and fixation) vary

for each satDNA family or even for the same satDNA family within different lineages. Levels of sequence variation among repeats would depend on factors such as mutation rate, inter- and intrachromosomal recombination rates, copy number, array size and structure, chromosomal distribution, chromosomal structure, population size, divergence time, and reproductive mode; it is also subject to random genetic drift and possibly natural selection (Strachan et al., '85; Stephan and Cho, '94; Luchetti et al., 2003; Navajas-Pérez et al., 2005; Dawe and Henikoff, 2006; Kuhn et al., 2007). The relative importance of each factor remains controversial.

In this context, very little information exists on satDNA array size, composition and long-range organization, especially in reptiles (see Giovannotti et al., 2009). An exception is represented by Lacertidae, a species rich family of squamate reptiles, widespread in the Palearctic region (Sindaco and Jeremcenko, 2008). This family comprises the subfamilies Gallotiinae and Lacertinae, with the latter comprising two monophyletic tribes, the Eremiadini of Africa and arid southwest and central Asia, and the Lacertini of Europe (Arnold et al., 2007). So far, five satDNA families have been described for the genome of the Lacertinae subfamily: the pLCS (190 bp in length) is shared by the genera *Algyroides*, *Teira*, *Lacerta*, and *Podarcis* (Capriglione et al., '89, '91; Capriglione, 2000); the pLHS (140 bp) is specific for *Podarcis* only (Capriglione et al., '94; Capriglione, 2000); the pGPS (185 bp) is present in the genome of *Podarcis* and in species belonging to the genera *Archaeolacerta*, *Algyroides*, *Lacerta*, and *Zootoca* (Capriglione et al., '98), so that its appearance would precede the divergence within the Lacertinae subfamily; the CLSat family is described for the Caucasian genus *Darevskia* (145–147 bp, Ciobanu et al., 2003; Grechko et al., 2006); the Agi160 is restricted to the genus *Lacerta* (138–184 bp, Ciobanu et al., 2004; Grechko et al., 2005). These satDNA families revealed several common features, such as the same range of monomer lengths (140–190 bp), AT content (tendency toward AT enrichment 50–65%) and homopolymeric ( $A_{3-4}$  and  $T_{3-4}$ ) stretches (Capriglione et al., '91; Ciobanu et al., 2001, 2004). All these features were also found in other nonreptilian satDNAs (see King and Cummings, '97).

The genus *Iberolacerta* (see Arribas, '99) has a disjunct range in mountain areas of western Europe: a portion comprises central Portugal, central and northern Spain and Pyrenees; another part embraces western Alps and northern Dinaric chain. Until recently the rock-lizard populations endemic to the Iberian Peninsula were considered to represent a single species, *Lacerta monticola* Boulenger, 1905 (see Salvador, '85), that has recently been split into the following taxa: *Iberolacerta aranica*, *I. aurelioi*, and *I. bonnali* restricted to the Pyrenees and *I. cyreni*, *I. galani*, *I. martinezricai*, and *I. monticola*, in the central-western parts of Iberian Peninsula (see Arribas et al., 2006). An additional species is represented by the east-Alpine and Dinaric species *I. horvathi*. This classification was based on (i) morphological (biometry, scalation), osteological, and karyological data; (ii) on the use of

Grant sponsor: Ministry of Education, University and Research (Italy) ("Ricerche di citogenetica molecolare sui sistemi di determinazione del sesso nei rettili squamati, sottordine Sauria"); grant number: PRIN2009/20093HYH97; grant sponsor: Ministerio de Ciencia y Tecnología, Spain; grant number: REN2003-02931/GLO; grant sponsor: Xunta de Galicia, Spain; grant numbers: PGIDIT03RFO10301PR, PGIDIT06RFO10301PR; grant sponsor: "FPU" fellowship; grant number: AP2008-02686; grant sponsor: Ministerio de Educación (Spain) (Beca de Formación de Profesorado Universitario).

Conflicts of interest: None.

\*Correspondence to: Vincenzo Caputo Barucchi, Dipartimento di Scienze della Vita e dell'Ambiente, Università Politecnica delle Marche, Ancona, Italy. E-mail: v.caputo@univpm.it

Received 8 April 2013; Revised 29 July 2013; Accepted 31 July 2013

DOI: 10.1002/jez.b.22530

Published online XX Month Year in Wiley Online Library

(wileyonlinelibrary.com).



molecular tools, namely nuclear (*c-mos*) and mitochondrial DNA (12S and cytochrome *b*), and (iii) on the construction of phylogenetic trees ranking the different allopatric populations based on the degree of genetic divergence, with *I. horvathi* as the most basal species (for a revision see Arribas et al., 2006). Another conceptual framework influencing the species subdivision of these largely allopatric lizards is the phylogenetic species concept, according to which species are segments of a phylogenetic lineage beyond nodes, irrespective of the degree of reproductive isolation (for a criticism see Mace, 2004). Considering the well-known usefulness of satDNAs in facing phylogenetic issues (i.e., Martinsen et al., 2009), the aim of the present paper was to isolate and characterise satDNA in some lacertid species in order to (i) increase the knowledge of this genomic elements in an important amniote group for which data on occurrence, genomic distribution, and evolutionary rates are limited to a handful of species; (ii) use the satDNAs isolated to verify the robustness of the proposed phylogenetic reconstruction for some *Iberolacerta* taxa on the light of independent molecular markers.

## MATERIALS AND METHODS

### Samples

Two males and two females of *Iberolacerta monticola* (from Fragas do Eume, A Capela, Galicia, Spain) and two males and two females of *I. galani* (from A Ponte, Pena Trevinca, A Veiga, Galicia, Spain) were used to make metaphase chromosomes and to extract genomic DNA. In addition, genomic DNA was extracted from seven ethanol preserved specimens of *I. cyreni* from three different Iberian locations (Navacerrada, Sierra de Guadarrama, Segovia-Madrid, Spain; Pico Zapatero, Sierra de la Paramera, Ávila; Puerto de Peña Negra, Sierra de Villafranca, Ávila, Spain) and one of *I. martinezricai* (Puerto El Portillo, Salamanca, Spain). Permissions for field work and experimental procedures were issued by the competent Spanish authorities: Xunta de Galicia (for *I. monticola* and *I. galani*) (permission number 79/2008) and Junta de Castilla y León (for *I. cyreni* and *I. martinezricai*) (permission numbers: 20051630007003/2005, 20061630024599/2006, 2007167004130/2007, 20081630020386/2008, 20092390004760/2009). Finally, genomic DNA of *Lacerta bilineata*, *Podarcis muralis*, *P. siculus*, and *Timon lepidus*, was extracted from ethanol preserved tissues of voucher specimens belonging to one of the authors (Vincenzo Caputo Barucchi).

### Isolation and Characterization of Satellite DNAs

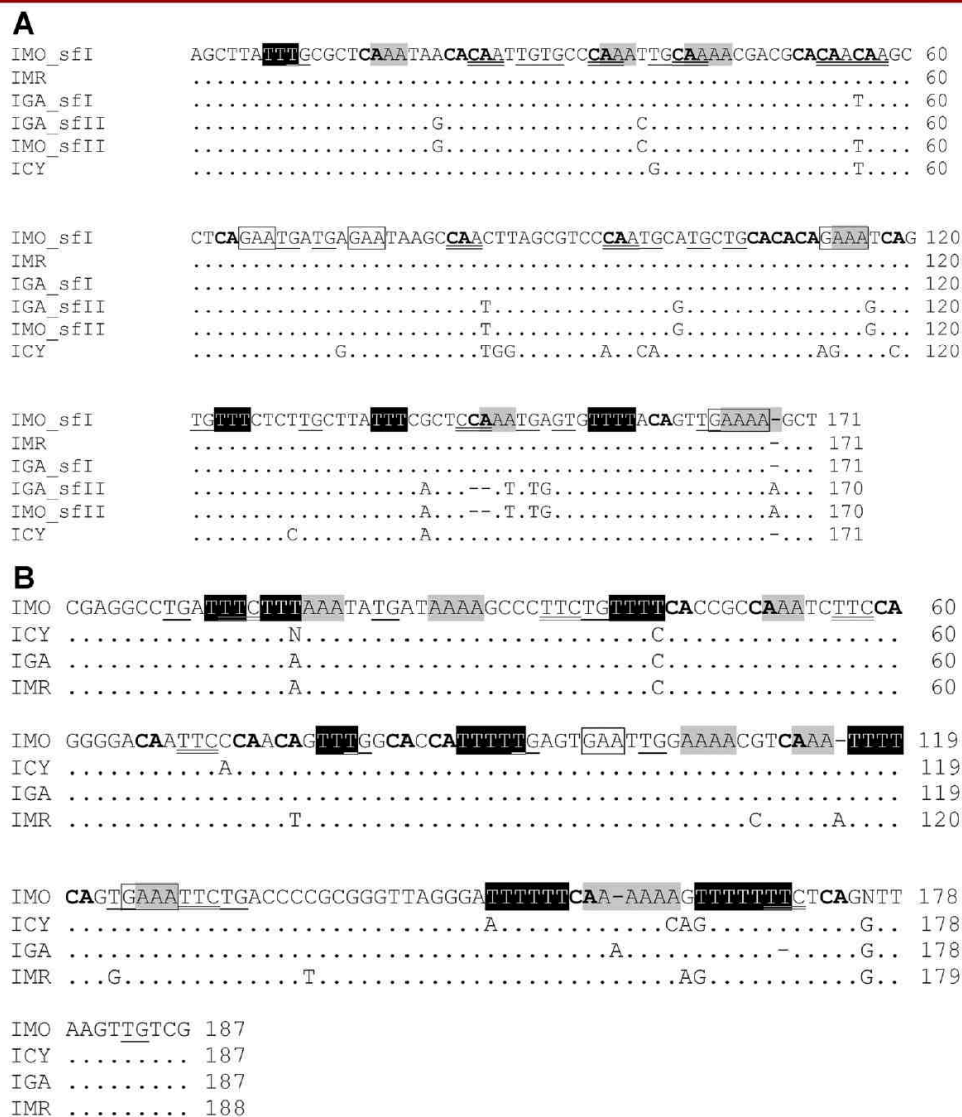
Genomic DNA was extracted from whole blood, using standard protocols with proteinase K digestion followed by phenol/chloroform extraction (see Sambrook et al., '89). Fifteen restriction endonucleases (*AluI*, *ApaI*, *Avall*, *BamHI*, *BclI*, *BglI*, *BglII*, *DraI*, *EcoRV*, *HindIII*, *MspI*, *RsaI*, *SmaI*, *TaqI*, *XbaI*) (Fermentas International, Inc., Burlington, ON, USA) were screened and about 8 µg of *I. monticola* and *I. galani* purified genomic DNA

were utilized for each digestion. Electrophoresis on 2% agarose gel of the digested DNA revealed a band of about 170 bp for *HindIII* and 190 bp for *TaqI*, corresponding to the monomeric unit of repetitive DNA (Fig. 1A), whereas no clear bands were produced by the remaining 13 endonucleases. The 170 and 190 bp fragments were excised from agarose gel, purified with Pure Link Quick Gel Extraction Kit (Invitrogen, Carlsbad, CA, USA) and cloned in the pCR®-blunt vector with Zero Blunt PCR Cloning Kit (Invitrogen) following the manufacturer's recommendations. Ten clones of each *I. monticola* satellite DNAs (*HindIII* and *TaqI* satDNAs henceforth) and 13 (*HindIII*) and 16 (*TaqI*) of *I. galani* satDNAs were sequenced on an ABI PRISM 3730XL (Applied Biosystems, Foster City, CA, USA) automatic sequencer.

Digoxigenin-labeled probes were produced by PCR amplification of single clones and used in Southern hybridization experiments to verify that the elements isolated were tandemly arranged, as expected for satDNAs. In these experiments, *HindIII* and *TaqI* digested genomic DNAs from *I. monticola* and other lizards (*I. cyreni*, *I. galani*, *I. martinezricai*, *Lacerta bilineata*, *Podarcis muralis*, *P. siculus*, *Timon lepidus*) were used in order to assess the presence of these repetitive elements in other genera of this family. The hybridization with the digoxigenin-labeled satDNA probes was performed at 50°C overnight with the Sure Blot CHEMI Hybridization and Detection Kit (EMD Millipore Co., Billerica, MA, USA) following the manufacturer's recommendations. The hybridization was detected with the same kit.

The genomic abundance of satDNAs was estimated by quantitative dot blot analysis. Dilutions of genomic DNA and clones containing *HindIII* and *TaqI* satDNAs used as a standard were blotted onto a nylon membrane with BIO-DOT® micro-filtration apparatus (Bio-Rad Laboratories, Hercules, CA, USA), following manufacturer's recommendations. In order to avoid errors due to the differences in the hybridization kinetics, sonicated salmon sperm DNA was used as a carrier and added to each sample up to a final amount of 0.5 µg DNA/sample (see Cafasso et al., 2003). Hybridization was performed overnight at 45°C. The same clones as those used as a standard were employed to produce digoxigenin-labeled probes. The detection protocol was carried out with the same protocol as the one used for Southern hybridization.

From the sequences of the monomers of *I. monticola* and *I. galani*, *HindIII* and *TaqI* satDNAs two pairs of primers (*HindIII*-F: 5'-TGAGTGTTTTACAGTTGAAAAGCT-3'; *HindIII*-R: 5'-CATTGTGTTATTTGAGCGCAA-3'; *TaqI*-F: 5'-ATTCTGACCCTGGGGGT-TAG-3'; *TaqI*-R: 5'-CATATTTAAAGAAATCAGGCCTCG-3') were designed and used for isolation of these satellites from the genomes of the other two *Iberolacerta* species. PCR products from the amplification of *Iberolacerta* genomic DNAs with above primers were run on 2% agarose gel, the band corresponding to the amplified monomers excised from the gel, purified with Pure Link Quick Gel Extraction Kit (Invitrogen) and cloned in the pCR®-blunt vector with Zero Blunt PCR Cloning Kit (Invitrogen)



**Figure 1.** Comparisons of consensus sequences of *Hind*III (a) and *Taq*I (b) between the four *Iberolacerta* species analysed. Repeated motifs are highlighted. sfi: *Hind*III satDNA subfamily I; sfII: *Hind*III satDNA subfamily II. Dots refer to nucleotide identity and dashes indicate indels.

following manufacturer's recommendations. Clones of *Hind*III and *Taq*I satDNAs were sequenced on an ABI PRISM 3730XL (Applied Biosystems) automatic sequencer. These sequences were then aligned in CLUSTAL W (Larkin et al., 2007), using default parameters. The visual inspection of sequence alignments was carried out to check for the presence of shared nucleotide changes, which could serve as diagnostic positions to define subsets (subfamilies) within each satDNA family. A GenBank search was performed in order to compare *Hind*III and *Taq*I satDNAs with other satDNAs in the database.

Maximum parsimony (MP), neighbor joining (NJ), maximum likelihood (ML), and Bayesian analyses (BA) were used to infer the phylogenetic relationships among sequences of each satDNA. MP consensus trees (50% majority rule) were constructed with PAUP\* version 4.0b10 (Swofford, 2002) using the heuristic search method with 1,000 random-addition-sequence replicates, tree-bisection-reconnection (TBR) branch swapping and holding 100 trees at each cycle of the stepwise-addition procedure. To increase the number of informative characters, gaps were coded as binary (presence/absence) characters.



NJ analyses were performed in MEGA version 5 (Tamura et al., 2011). The NJ trees were based on distances obtained by the maximum composite likelihood method, with pairwise deletion and 1,000 bootstrap replicates. ML analyses were conducted in MetaPIGA v.2.1.3 (<http://www.metapiga.org>) (Helaers and Milinkovitch, 2010) using the metapopulation genetic algorithm (metaGA) with probability consensus pruning among four populations of four individuals each. The best-fitting nucleotide substitution models [Jukes–Cantor (JC) for *HindIII* satDNA and Hasegawa–Kishino–Yano plus Gamma (HKY + G) for *TaqI* satDNA] were selected based on the Likelihood Ratio Test implemented in this software. Branch support values that approximate the posterior probability distribution of the corresponding branches were estimated by performing a minimum of 100 replicated metaGA searches that were stopped when the mean relative error (MRE) among 10 consecutive consensus trees remained below 5%. BA were carried out using the software MrBayes v.3.2.1 (Ronquist and Huelsenbeck, 2003). As in the MP analyses, gaps were coded as binary characters and included as a separate data partition in the matrix. A binary model (Iset coding = variable) was applied to the coded gaps, whereas the previously selected models of sequence evolution, JC and HKY + G, were applied to the DNA partitions of *HindIII* and *TaqI* satDNAs, respectively. The analyses included two separate concurrent Monte Carlo Markov Chain (MCMC) runs, each composed of four chains (one cold, three heated). Each Markov chain was started from a random tree and run for up to  $10^6$  generations, sampling every 500 generations. Stationarity was assessed using the software Tracer v.1.5 (Rambaut and Drummond, 2009). Samples obtained during the first 25% generations were discarded as burn-in, and the remaining data were used to generate a majority-rule consensus tree where the percentage of samples recovering any particular clade of the consensus tree represented the clade's posterior probability.

Intraspecific nucleotide diversity ( $\pi$ ) was estimated using DnaSP v. 5 (Librado and Rozas, 2009). Net average genetic distances between groups were calculated under the appropriate substitution model for each satDNA family (see above) with MEGA v. 5. Rates of *HindIII* and *TaqI* satDNAs evolution were determined according to the divergence times estimated for the four *Iberolacerta* species here investigated by Arribas et al. (2006).

The occurrence of genetic differentiation between the four species analyzed was assessed with the analysis of molecular variance (AMOVA) (Excoffier et al., '92) calculating  $\Phi$ -statistics. This test was performed at two hierarchical levels to test how satDNAs sequence variability was distributed within species and among species, for both *HindIII* and *TaqI* satDNAs. The test was based on pair wise genetic distances between clones and performed as implemented in ARLEQUIN 2.000 (Schneider et al., 2000), using 1,000 permutations.

The repeats of the analyzed species were compared using satDNA Analyzer version 1.2 (Navajas-Pérez et al., 2007). This

program allows the discrimination between shared and nonshared polymorphic sites. The program identifies polymorphic sites shared between two species when the same polymorphism is found in both species. When this occurs, we assume that these are ancestral sites that appeared before the split between the two species (Navajas-Pérez et al., 2005). By contrast, nonshared polymorphic sites are autapomorphies, representing different transitional stages in the process of intraspecific sequence homogenization and interspecific divergence. Under the assumption that concerted evolution is a time dependent process, the expected stages of transition during the spread of a variant repeat unit toward its fixation can be defined according to the model of Strachan et al. ('85). This is a method of partitioning the variation by analyzing the patterns of variation at each nucleotide site considered independently among all the repeats of a repetitive family when comparing a pair of species (Strachan et al., '85; Navajas-Pérez et al., 2007). This method examines the distribution of nucleotide sites among six stages (Classes I–VI) in the spread of variant repeats through the family and the species. Briefly, the Class I site represents complete homogeneity across all repeat units sampled from a pair of species, whereas Classes II, III, and IV represent intermediate stages in which one of the species shows a polymorphism. The frequency of the new nucleotide variant at the site considered is low in Class II and intermediate in Class III, while Class IV represents sites in which a mutation has replaced the progenitor base in most members of the repetitive family in the other species. Class V represents diagnostic sites in which a new variant is fully homogenized and fixed in all the members of one of the species while the other species retains the progenitor nucleotide. A Class VI site represents an additional step over the stage of Class V (new variants appear in some of the members of the repetitive family at a site fully divergent between the two species). The statistical significance ( $P$ -value) of the variation in the relative proportions of Strachan transitions stages among different interspecific comparisons was evaluated using chi-square heterogeneity tests that were performed in the interactive online calculator available at <http://www.quantpsy.org/chisq/chisq.htm> (Preacher, 2001).

#### Chromosome Analysis

For metaphase preparations, about 50  $\mu$ l of blood were taken from *I. monticola* and *I. galani* individuals with a sterile heparinized syringe and cultured in CO<sub>2</sub> incubators using the culture conditions indicated by Ezaz et al. (2005). Metaphase preparations were obtained by exposing cell cultures to 75 ng/ml of Demecolcine (Sigma-Aldrich Co., St Louis, MO, USA) for 4 hr before harvesting (Ezaz et al., 2005). Cells were hypotonized in KCl 0.75 M for 30 min at 37°C, prefixed by adding several drops of freshly prepared methanol:acetic acid fixative (3:1), then fixed through three changes of fixative. Suspensions of fixed cells were dropped onto microscope slides and air dried at room temperature.

Fluorescence in situ hybridization (FISH) experiments were performed on metaphase preparations using (i) a telomeric probe



(TTAGGG)<sub>n</sub> produced by PCR according to Ijdo et al. ('91), and (ii) the probes obtained by PCR amplification of *TaqI* and *HindIII* satDNA clones. Telomeric and *TaqI* probes were also used in two-color FISH experiments. The probes were labeled by PCR either with biotin-16-dUTP (Roche) or digoxigenin-11-dUTP (Roche Diagnostics GmbH, Mannheim, Germany). Slide pretreatment, denaturation, hybridization, post-hybridization washes, and detection were performed according to Schwarzacher and Heslop-Harrison (2000). The *HindIII* satDNA and telomeric probes were evidenced with fluorescein iso-thiocyanate (FITC) and tetramethyl rhodamine iso-thiocyanate (TRITC), respectively. Chromosomes were observed with a Nikon Eclipse 800 epifluorescence microscope and the images were captured and processed with a Leica CytoVision version 7.2 system.

In order to define the relationships between satDNAs and the constitutive heterochromatin, C-banding was performed on metaphase plates following Sumner ('72). The relations between AT-rich heterochromatic regions and satDNAs were determined by staining C-banded metaphases with 4',6-diamidino-2-phenylindole (DAPI) (Schweizer, '76).

## RESULTS

### Isolation and Characterization of Satellite DNAs

The digestion of *I. monticola* and *I. galani* genomic DNA with *HindIII* and *TaqI* restriction enzymes revealed bands corresponding to a monomer of a repetitive element of about 170 and 190 bp, respectively (not shown). PCR amplification using primers designed by aligning *I. monticola* and *I. galani* sequences of both satDNAs was successful in individuals representing the other two lineages of *Iberolacerta* recognized as distinct species (*I. martinezricai*, *I. cyreni*). The length of the 45 clones sequenced for *HindIII* ranged between 170 and 172 bp, whereas the length of the 42 clones sequenced for *TaqI* ranged between 171 and 188 bp (Table 1). Sequences of both satDNAs were deposited in GenBank

(*HindIII* accession numbers: from KF453637 to KF453681; *TaqI* accession numbers: from KF453682 to KF453723). When *HindIII* and *TaqI* satDNA sequences were subjected to a BLASTN search, no significant similarities with sequences deposited in databases were found.

Southern blot analysis revealed hybridization of both satDNA probes onto *Iberolacerta monticola* digested genomic DNA with a ladder-like pattern, indicating the tandem arrangement of repeating units which is typical of satDNAs. A strong hybridization signal was also produced on the other three *Iberolacerta* species with both *HindIII* and *TaqI* probes; this latter probe also produced a clear signal on the other lizards tested, whereas no signal appeared when *HindIII* probe was hybridized on representatives of the genera *Lacerta*, *Podarcis*, and *Timon* (not shown).

Quantitative dot blot analysis revealed that *HindIII* satDNA represents around 10% of *I. monticola* and *I. galani*, and 5% of *I. cyreni* and *I. martinezricai* genomes. *TaqI* satDNA represents 5% of *I. cyreni*, *I. galani*, and *I. monticola* genomes, and 2.5% in *I. martinezricai* (data not shown). The estimation of the number of repeats was not possible because the genome size of these lizards is not known.

The consensus sequences of the two satDNAs were very similar in the four *Iberolacerta* species, with an AT average content of 58.4% for *HindIII* and 60.3% for *TaqI*, indicating an enrichment in AT (Table 1). Both satellites repeats are characterized by the occurrence of short motifs such as A and T stretches and dinucleotides steps TG and CA, with more numerous and longer A (T) stretches in *TaqI* satDNA (Fig. 1), as expected from its higher AT content. Within *HindIII* satDNA, two monomer variants or subfamilies (I and II) were detected in *I. galani* and *I. monticola* (Fig. 1A). The consensus sequences of subfamily I in both species were virtually identical to the consensus of *I. martinezricai*, whereas subfamily II showed several (nine) randomly distributed diagnostic nucleotide substitutions, as well as three exclusive indels located in the terminal region of the monomer. Both

Table 1. Summary of repeat features and  $\pi$  values.

Species	<i>HindIII</i>				<i>TaqI</i>			
	<i>n</i>	%AT	Repeat length	$\pi$	<i>n</i>	%AT	Repeat length	Nucleotide diversity ( $\pi$ )
<i>I. cyreni</i>	11	57.0	171	0.0055 ± 0.0022	9	60.2	186–187	0.0384 ± 0.0058
<i>I. galani</i>	13	58.9	170–171	0.0358 ± 0.0033	16	60.1	186–187	0.0475 ± 0.0070
<i>I. galani</i> (sfl)	6	59.4	171	0.0175 ± 0.0031				
<i>I. galani</i> (sfl)	7	58.5	170	0.0101 ± 0.0020				
<i>I. monticola</i>	10	59.0	170–171	0.0187 ± 0.0035	10	60.8	171–188	0.0569 ± 0.0062
<i>I. monticola</i> (sfl)	9	59.0	171	0.0062 ± 0.0019				
<i>I. monticola</i> (sfl)	1	58.8	170	–				
<i>I. martinezricai</i>	10	58.7	171–172	0.0105 ± 0.0052	7	60.1	187–188	0.0428 ± 0.0114

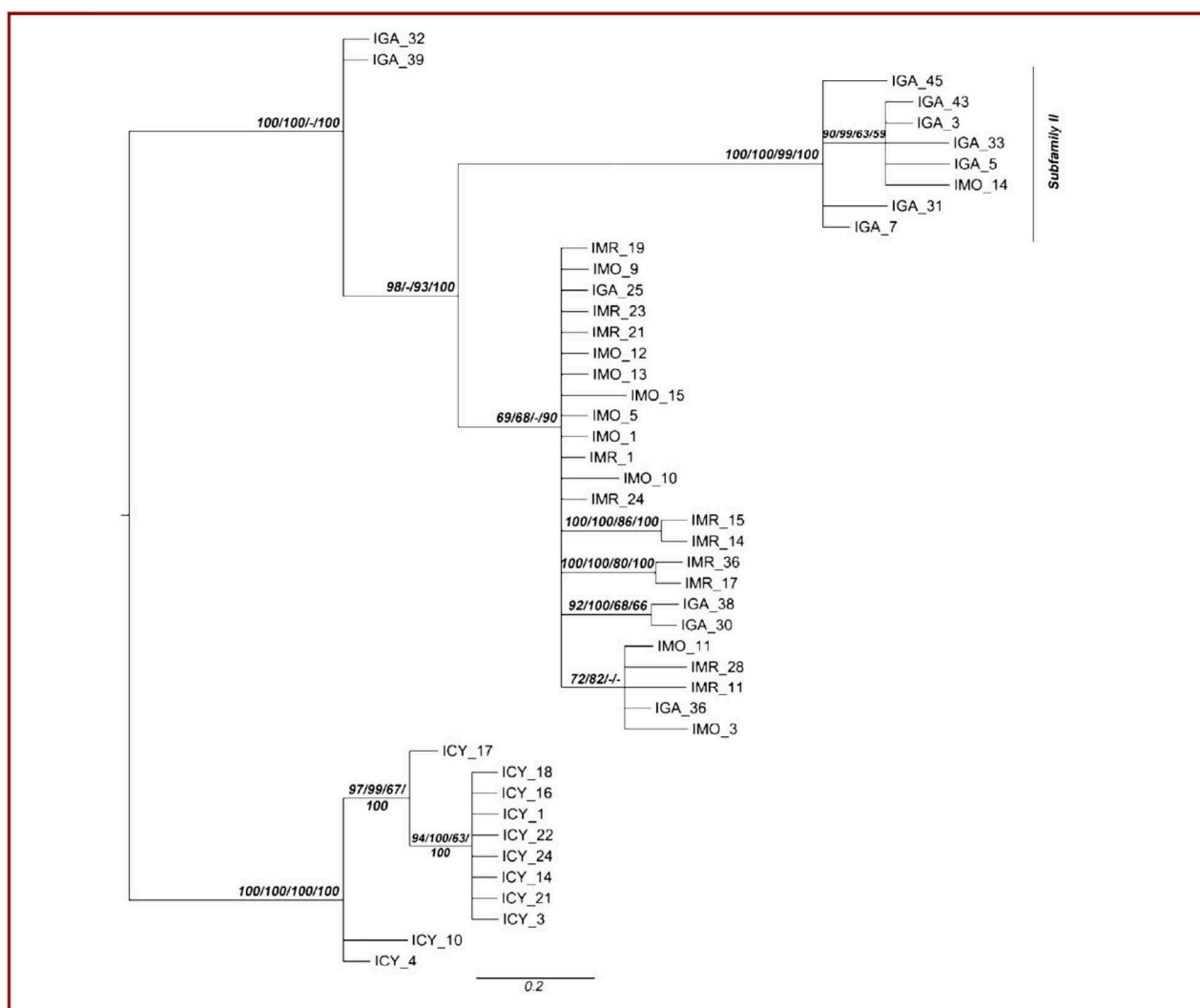
Number of monomeric repeats sequenced (*n*), nucleotide composition of repeats (AT), length of repeats (expressed in base pairs), and nucleotide diversity ( $\pi$ ) ± SE for both satDNAs for each *Iberolacerta* species investigated. sfl: *HindIII* satDNA subfamily I; sflI: *HindIII* satDNA subfamily II

monomer variants were present in similar proportions in the sequence data set of *I. galani*, but only one out of ten sequences in *I. monticola* belonged to subfamily II (Table 1).

The phylogenetic tree obtained from the Bayesian analysis of *Hind*III satDNA is shown in Figure 2. The four different phylogenetic analyses (NJ, MP, ML, and BA) yielded very similar topologies, with some minor incongruences. Two major clades were recovered with maximum support, one harbouring *I. cyreni* clones and the other the sequences of the remaining three *Iberolacerta* species. Within this second cluster, monomers of

subfamily II constitute a well-supported clade sister to that formed by sequences belonging to subfamily I, [with the exception of two clones from *I. galani* (IGA\_32 and IGA\_39) that share some private nucleotide substitutions]. Within subfamily I, relationships between most monomers were poorly resolved and they were not grouped according to the species of origin.

The Bayesian tree constructed using the sequences of *Taq*I satDNA was largely unresolved, regardless of the phylogenetic method employed, showing that this satellite cannot discriminate effectively the four *Iberolacerta* species here investigated (Fig. 3).



**Figure 2.** Bayesian phylogenetic tree depicting the the phylogenetic relationships between the 45 monomeric units of *Hind*III satDNA sequenced. Support values obtained by four different methods of analysis are shown at each node; from left to right: Bayes posterior probability (100×), metaGA branch support values (100×), NJ-bootstrap (%), and equally MP trees (%). A hyphen was inserted whenever a particular method did not support the Bayesian topology. Numbers after the species names are experimental number for clone identification. ICY: *Iberolacerta cyreni*; IGA: *Iberolacerta galani*; IMO: *Iberolacerta monticola*; IMR: *Iberolacerta martinezricai*.

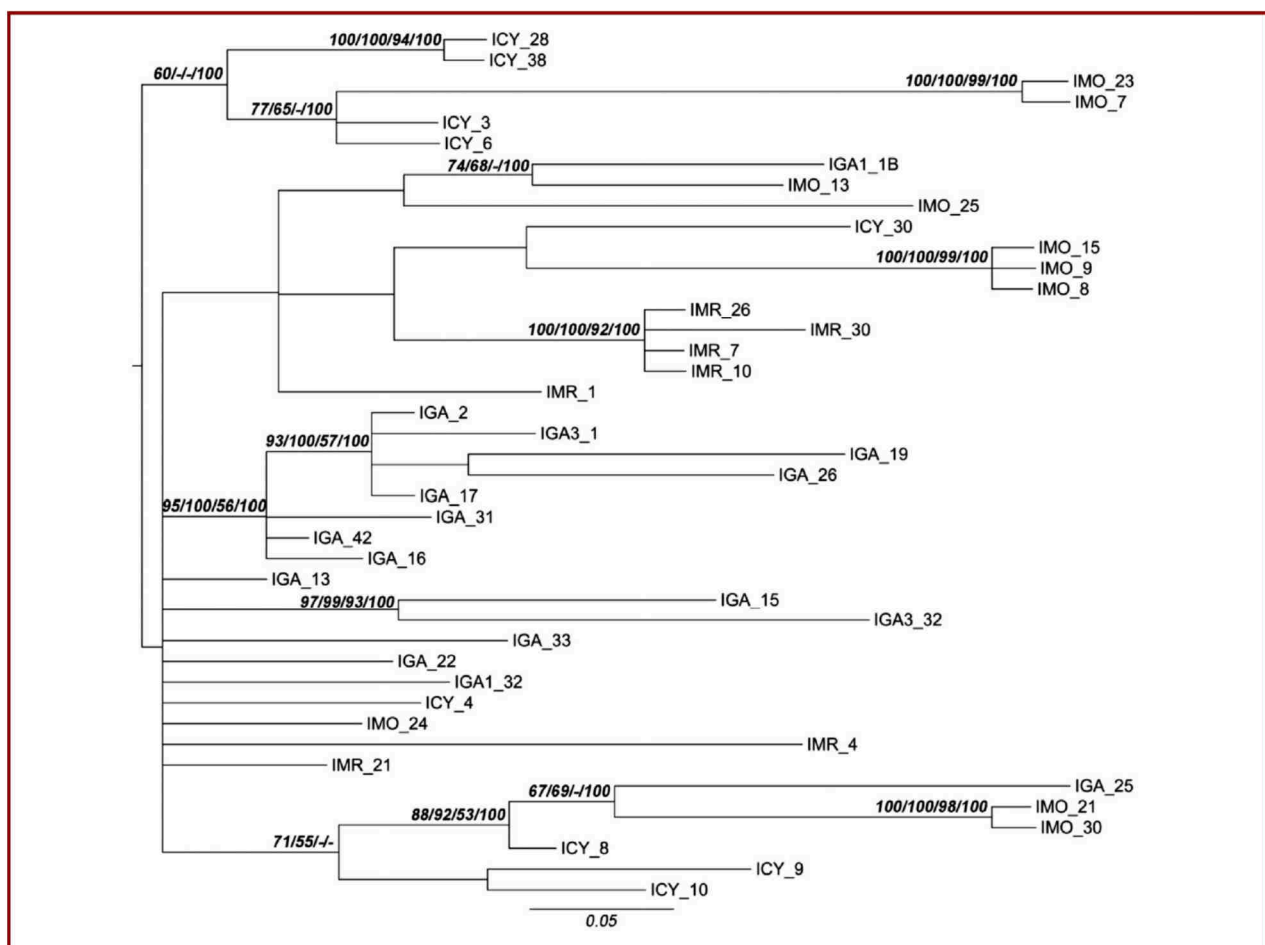
Even though several well-supported subclusters including conspecific monomers were recognized, the number of diagnostic mutations shared by these sequences was too low to be considered species-specific *TaqI* satDNA subfamilies (not shown).

The  $\pi$  values indicated that intraspecific sequence heterogeneity is higher for *TaqI* satDNA (from 3.84% in *I. cyreni* to 5.69% in *I. monticola*) than for *HindIII* satDNA (from 0.55% in *I. cyreni* to 3.58% in *I. galani*) (Table 1). Interspecific mean net distances are low and similar for both satellites when *I. cyreni* is excluded from the analysis of *HindIII* satDNA (from 0.04% between *I. monticola* subfamily I and *I. martinezricai* to 5.60% between *I. galani* subfamily II and *I. martinezricai* for *HindIII*, and from 0.90% between *I. galani* and *I. martinezricai* to 1.30% between *I. monticola* and *I. galani* for *TaqI* satDNA) (Tables 2 and 3). Pair wise

comparisons of *HindIII* satDNA involving *I. cyreni* and the other *Iberolacerta* analyzed, showed distance values substantially higher, between 8.40% and 13.90% (Table 2).

In addition, higher levels of sequence divergence were obtained in the comparisons between subfamilies I and II of *HindIII* satDNA in *I. galani* (4.5%) than in the comparisons between monomeric repeats belonging to subfamily I in different species (from 0.04% to 0.4%) (Table 2).

The evolutionary rate of these two satellites was then calculated based on sequence divergence between *I. cyreni* and the other three species, that were considered as a single taxonomic unit not being discriminated by either satellite. The values found are 1.2% for *HindIII* and 0.14% for *TaqI*, indicating an evolutionary rate almost 10-fold faster for the former.



**Figure 3.** Bayesian phylogenetic tree depicting the the phylogenetic relationships between the 42 monomeric units of *TaqI* satDNA sequenced. Support values obtained by four different methods of analysis are shown at each node; from left to right: Bayes posterior probability (100×), metaGA branch support values (100×), NJ-bootstrap (%) and equally MP trees (%). A hyphen was inserted whenever a particular method did not support the Bayesian topology. Numbers after the species names are experimental number for clone identification. ICY: *Iberolacerta cyreni*; IGA: *Iberolacerta galani*; IMO: *Iberolacerta monticola*; IMR: *Iberolacerta martinezricai*.



**Table 2.** Interspecific and intersubfamily comparative analysis of *HindIII* repeats

Species comparison	SP (%)	Strachan sites II–III (%)	Strachan sites IV–VI (%)	Genetic distance
<i>HindIII</i>				
<i>I. cyreni</i> versus <i>I. galani</i> (sfl)	0 (0%)	4 (2.3%)	15 (8.8%)	0.0838 ± 0.0232
<i>I. cyreni</i> versus <i>I. galani</i> (sflI)	0 (0%)	5 (2.9%)	21 (12.3%)	0.1388 ± 0.0326
<i>I. cyreni</i> versus <i>I. monticola</i> (sfl)	0 (0%)	5 (2.9%)	16 (9.4%)	0.1025 ± 0.0265
<i>I. cyreni</i> versus <i>I. martinezricai</i>	0 (0%)	5 (2.9%)	15 (8.8%)	0.0996 ± 0.0258
<i>I. galani</i> (sfl) versus <i>I. monticola</i> (sfl)	1 (0.59%)	7 (4.1%)	1 (0.59%)	0.0038 ± 0.0025
<i>I. galani</i> (sfl) versus <i>I. martinezricai</i>	2 (1.2%)	6 (3.5%)	1 (0.59%)	0.0034 ± 0.0026
<i>I. monticola</i> (sfl) versus <i>I. martinezricai</i>	2 (1.2%)	4 (2.3%)	0 (0%)	0.0004 ± 0.0005
<i>I. galani</i> (sflI) versus <i>I. monticola</i> (sfl)	0 (0%)	8 (4.7%)	9 (5.3%)	0.0545 ± 0.0190
<i>I. galani</i> (sflI) versus <i>I. martinezricai</i>	0 (0%)	8 (4.7%)	9 (5.3%)	0.0555 ± 0.0192
<i>I. galani</i> (sfl) versus <i>I. galani</i> (sflI)	0 (0%)	6 (3.5%)	10 (5.8%)	0.0447 ± 0.0160

The table reports number and percentage of shared polymorphic sites (SP); variable nucleotide sites classified according to Strachan et al. ('85); net genetic distances (Jukes–Cantor method) in pair wise comparisons of species. sfl: *HindIII* satDNA subfamily I; sflI: *HindIII* satDNA subfamily II.

The poor phylogenetic differentiation of these species based on the sequences of the satDNAs here isolated was confirmed by AMOVA analysis. When this test was performed on the *HindIII* sequences, most of the percentage of the molecular variation was distributed among species (69.60%;  $\Phi_{ST}$  0.69596,  $P < 0.0001$ ) whereas the percentage of variation within species was much lower, but still significant (30.40%;  $\Phi_{ST}$  0.69596,  $P < 0.001$ ) (Table 4). The variance among species became much lower (32.07%;  $\Phi_{ST}$  0.32072,  $P < 0.001$ ) and the one within populations became the preponderant variance component (67.93%;  $\Phi_{ST}$  0.32072,  $P < 0.001$ ) when the sequences of *I. cyreni* were excluded from the analysis (Table 4). This result can be explained by the fact that *I. cyreni* was recovered as a distinct cluster with a high support in the phylogeny based on *HindIII* sequences, whereas the other three cannot be discriminated by this molecular marker. The AMOVA test carried out on *TaqI* satDNA sequences produced results very similar to those obtained with *HindIII* sequences after excluding *I. cyreni*, with a preponderant variance component distributed

within species (82.69%;  $\Phi_{ST}$  0.17314,  $P < 0.001$ ), confirming that this satDNA cannot effectively discriminate between these *Iberolacerta* species (Table 4). These results emerged also by analyzing the pattern of variation at each nucleotide position considered independently among all *HindIII* repeats (Table 2). Indeed, when comparing *I. cyreni* with the other species, a high percentage of Strachan sites belonging to the categories IV, V, and VI were found (average = 9.9%), while 5.1% of sites per repeat were Strachan transition stages (II + III), and no shared polymorphic sites were observed. Conversely, for *TaqI* satDNA sites of the classes IV–VI were very few (average = 0.5%) in all the comparison, while 20.7% of the sites represented Strachan stages II–III and an average of 4.1% were polymorphic sites (Table 3). According to the chi-square heterogeneity test, these differences in the relative proportions of Strachan transition stages between *HindIII* and *TaqI* satDNAs are highly significant ( $P < 0.001$ ).

The relatively high degree of genetic differentiation detected in the analysis of sequence divergence between *HindIII* subfamily II

**Table 3.** Interspecific comparative analysis of *TaqI* repeats.

Species comparison	SP (%)	Strachan sites II–III (%)	Strachan sites IV–VI (%)	Genetic distance
<i>TaqI</i>				
<i>I. cyreni</i> versus <i>I. galani</i>	8 (4.3%)	51 (27.3%)	3 (1.6%)	0.0099 ± 0.0040
<i>I. cyreni</i> versus <i>I. monticola</i>	9 (4.8%)	25 (13.4%)	1 (0.5%)	0.0113 ± 0.0040
<i>I. cyreni</i> versus <i>I. martinezricai</i>	7 (3.7%)	25 (13.4%)	2 (1.1%)	0.0109 ± 0.0039
<i>I. galani</i> versus <i>I. monticola</i>	10 (5.3%)	43 (23%)	1 (0.5%)	0.0130 ± 0.0057
<i>I. galani</i> versus <i>I. martinezricai</i>	5 (2.7%)	57 (30.5%)	3 (1.6%)	0.0089 ± 0.0037
<i>I. monticola</i> versus <i>I. martinezricai</i>	7 (3.7%)	31 (16.6%)	2 (1.1%)	0.0114 ± 0.0040

The table reports number and percentage of shared polymorphic sites (SP); variable nucleotide sites classified according to Strachan et al. ('85); net genetic distances (maximum composite likelihood method) in pair wise comparisons of species.

Source of variation	Variance components	Percentage of variation
Among species	4.21275	69.60
	1.07719	32.07
	1.14218	17.31
Within species	1.84035	30.40
	2.28152	67.93
	5.45487	82.69

The test was carried on *Hind*III satDNA sequences including the four species selected for this study (first line of each hierarchical level), and removing *Iberolacerta cyreni* from the analysis (second line of each hierarchical level). The test on *Taq*I satDNA sequences included all four of the *Iberolacerta* investigated (third line of each hierarchical level).  $\Phi$ -statistics were highly significant in all comparisons ( $P < 0.001$ ).

and subfamilies I from *I. galani*, *I. monticola*, and *I. martinezricai* was also evident in the comparisons of Strachan transition stages among these groups (Table 2). No shared polymorphisms were found and the number of sites falling in classes IV and V (between 5% and 6%) was significantly larger ( $P < 0.001$ ) than the average frequency of these “differentiated sites” in the comparisons among subfamilies I in different species.

#### Chromosome Analysis

FISH experiments with *Hind*III satDNA probe on metaphase chromosomes of *I. galani* and *I. monticola* revealed that this repetitive element is widespread in the genome of these species, occurring at centromeres of all the 36 chromosomes of the diploid complement (Fig. 4A,B), with no differences between males and females. The occurrence of “bouquet” figures where chromosomes are linked together at the level of centromeres seems to indicate that this satDNA is involved in the interchromosome connection during mitosis (Fig. 4B). FISH with *Taq*I satDNA probe produced bright signals in interstitial position in a subset of 18 chromosomes in *I. galani* and 20 in *I. monticola*. No differences between males and females were detected with this probe either (Fig. 4C,D). Results of FISH experiments are consistent with the genomic abundance of *Hind*III and *Taq*I satDNAs as showed by quantitative dot blot analysis for these two species, with the former around twofold more abundant than the latter.

FISH with a telomeric probe (TTAGGG)<sub>n</sub> produced a fluorescent signal at telomeres of all the chromosomes. Besides telomeric signals, also interstitial telomeric sites (ITS) were marked in about five chromosome pairs. When a two-color FISH with both telomeric and *Taq*I satellite probes were performed, the fluorescent signals of ITS resulted distally located to the satellite ones (Fig. 4D).

C-banding, performed in order to assess the relationships between the isolated satellites and constitutive heterochromatin,

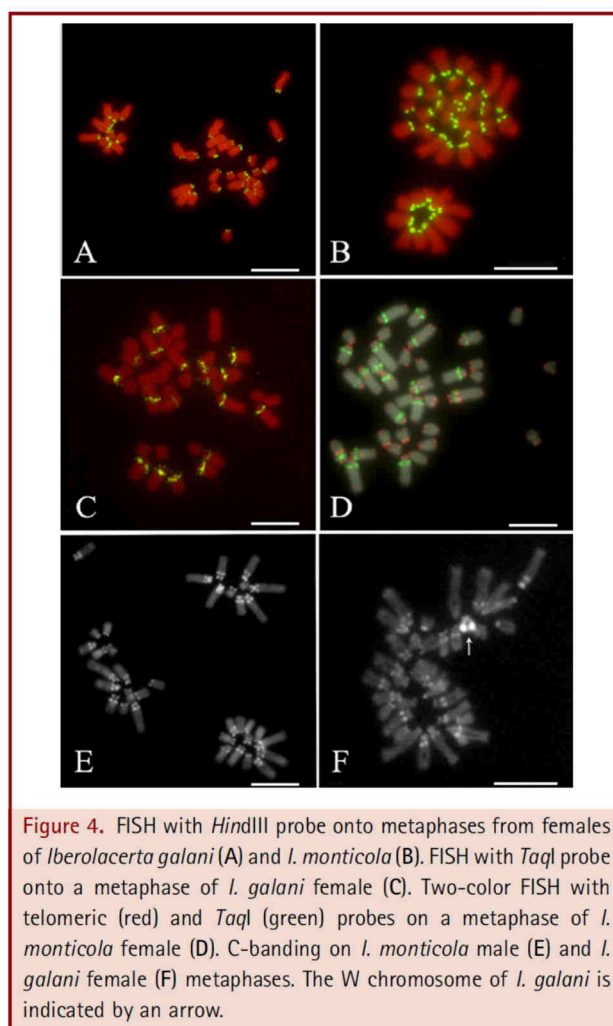


Figure 4. FISH with *Hind*III probe onto metaphases from females of *Iberolacerta galani* (A) and *I. monticola* (B). FISH with *Taq*I probe onto a metaphase of *I. galani* female (C). Two-color FISH with telomeric (red) and *Taq*I (green) probes on a metaphase of *I. monticola* female (D). C-banding on *I. monticola* male (E) and *I. galani* female (F) metaphases. The W chromosome of *I. galani* is indicated by an arrow.

revealed that in *Iberolacerta* the chromosomal distribution of *Hind*III satDNA overlaps the centromeric heterochromatic blocks, whereas *Taq*I probe colocalizes with pericentromeric heterochromatin (Fig. 4E,F).

#### DISCUSSION

Satellite DNAs represent rapidly evolving genomic elements, and therefore, even among most closely related species, they usually differ in nucleotide sequence, copy number, and/or composition of satellite families (Csink and Henikoff, '98). However, some satDNA families evolve more slowly than others and occur in several closely related species with different degrees of sequence similarity (Bachmann and Sperlich, '93; Mantovani et al., '97; Watabe et al., '97). Some satDNAs seem to be rather ancient and are widely distributed among higher taxa (Modi et al., 2004; Robles et al., 2004). Consequently, some satDNAs may be valuable



taxonomic identification tools while others might be useful for phylogenetic analyses at higher taxonomic levels. In the present study, we compared sequences of two different satDNA families (*Hind*III and *Taq*I) in four closely related lacertid species, allopatrically distributed in mountain areas of the Iberian Peninsula. These satDNAs seem to evolve at different rates in the studied lizards, with *Hind*III showing a 10-fold faster evolutionary rate than *Taq*I. Indeed, Southern blot analysis using *Iberolacerta* satellite probes revealed a clear hybridization pattern also in other lizard genera (namely, *Lacerta*, *Podarcis*, and *Timon*) only for *Taq*I repeats, whereas *Hind*III seems to be restricted to the genus *Iberolacerta*. However, a significant level of genetic divergence was detected only in comparisons involving *I. cyreni* when *Hind*III satDNA was considered. For this satDNA, analysis of turnover dynamics indicate the effectiveness of the molecular drive process, after species split, in the spreading of new sequence variants leading to intraspecific homogeneity (0.56% of sequence variation within *I. cyreni*) and interspecific divergence (around 9% of sequence divergence between *I. cyreni* and the other species), an evolutionary pattern known as concerted evolution (Dover, '82). The fact that the other species are scarcely differentiated at *Hind*III repeats can be interpreted in two alternative ways: (i) it may represent the outcome of the relatively recent (approximately 2 mya, Arribas et al., 2006) and rapid succession of speciation events within this group. In fact, previous molecular analyses based on nuclear and mitochondrial markers also failed to resolve the phylogenetic relationships or even track lineage splitting at this taxonomic level (Mayer and Arribas, 2003; Carranza et al., 2004; Crochet et al., 2004; Arribas et al., 2006; Arnold et al., 2007); (ii) the specific status for these three taxa might not have been reached yet. Indeed, estimation of divergence times among these three *Iberolacerta* species are similar to those recorded for different populations of the lizard *Podarcis muralis* that diverged genetically in separate refuges during glaciations, currently not showing evidence for reproductive isolation (Giovannotti et al., 2010).

The deep divergence observed between *I. cyreni* and the other *Iberolacerta* species here investigated with *Hind*III satellite is in good accordance with the molecular phylogenies published so far (Mayer and Arribas, 2003; Carranza et al., 2004; Crochet et al., 2004; Arribas et al., 2006; Arnold et al., 2007). This analysis showed that this species was the most diverged clade of the tree, with an estimated splitting time of about 7.5 million years. The relatively scarce representation of transitional stages (only 5% of the nucleotide positions) might suggest that the concerted evolution mechanisms have led to sequence differentiation between *I. cyreni* and the other species, probably due to the efficiency of the molecular-exchange homogenizing mechanisms among chromosomes.

The occurrence of two different types of monomeric variants or subfamilies was described for *Hind*III satDNA sequences. These subfamilies were defined according to a set of particular

nucleotide substitutions or indels, in two of the four species examined. However, given the almost simultaneous speciation processes between *I. monticola*, *I. galani*, and *I. martinezricai*, it seems unlikely that subfamily II constitutes a specific variant of *I. monticola* and *I. galani*. An interspecific analysis of the pattern of nucleotide change was not possible for subfamily II due to the lack of a representative number of sequences in *I. monticola* or *I. martinezricai*. Even so, our results show that both subfamilies are presumably evolving independently, as indicated by the substantially high percentage of transitions stages IV and V between the monomers of subfamily II (*I. galani*) and the sequences of subfamily I, either belonging to *I. galani*, *I. monticola*, or *I. martinezricai*. The coexistence and divergent evolution of satellite subfamilies in the genomes of these species could be in agreement with the Nijman and Lenstra model (2001), in which mutations inhibiting the interactions of repeat units in a satellite family would lead to sequence diversification and the independent amplification or contraction of concurrent sequence variants. Nevertheless, a more extensive survey of *Hind*III satDNA will be the subject of further studies, in order to assess the presence and abundance of both monomeric variants in other *Iberolacerta* species, as well as to elucidate the processes driving the evolution of this satellite family.

Conversely to *Hind*III sequences, the tandem arrays of *Taq*I show a low sequence change rate when comparing *I. cyreni* with the other *Iberolacerta*. In fact, we detected a low rate of sequence change (0.1% per Myr), a rate 10-fold lower than that estimated for *Hind*III sequences (about 1.2% per Myr) and only 1.1% of Strachan stages IV–VI compared to 18% of II–III stages. In addition, we also observed some shared polymorphic sites and a comparatively higher intraspecific heterogeneity, suggesting that most of the intraspecific variability in each species is ancestral, originated prior to the separation of these lineages; moreover, the high number of transitional stages of differentiation (Strachan stages II–III) suggest that after the allopatric isolation, processes of concerted evolution were less efficient than in the *Hind*III repeats. In addition, contrarily to *Hind*III, Southern hybridization with *Taq*I probe produced a clear signal also in other lacertid genera, like *Lacerta*, *Podarcis*, and *Timon*, also suggesting a strong conservation of this satellite DNA family.

Various factors were invoked to explain different evolutionary turnover rates between satDNA families, like interchromosomal and intrachromosomal recombination rates, copy number, array size and structure, chromosomal distribution, chromosomal structure, population size, divergence time and reproductive mode. Moreover, evolutionary conservation of satDNA repeats might be a likely indication of functional constraints and natural selection (see Plohl et al., 2008). Unfortunately, very few examples are found in the literature with both fast-evolving and slow-evolving satDNAs found within the same species. For instance, in the genus *Dolichopoda*, a comparison among three satDNA families showed a trend of sequence variability and copy number



being positively correlated, and a trend of sequence variability and length of repeats being negatively correlated (Martinsen et al., 2009). Like in *Dolichopoda*, it seems that also in the studied lizards an increase in copy number is linked to a trend of sequence homogenization. In fact, it was observed that *HindIII* repeats represent between 5% and 10% of the *Iberolacerta* genome, while *TaqI* satDNA between 2.5% and 5%. The different chromosome localization of the two satellites may also play a role in the different rate of sequence homogenization recorded for the two satDNA families. First of all, it should be noted that *HindIII* repeats are centromerically located on all the acrocentric chromosomes of *I. galani* and *I. monticola* karyotypes. In fact, it is reported that satellite DNAs at centromeres of acrocentric chromosomes show greater homology and a higher rate of homogenization than in noncentromeric locations or nonacrocentric chromosomes (Jantsch et al., '90; Bandyopadhyay et al., 2001). It has been hypothesized that homogenization occurs through physical association and crossing-over between nonhomologous chromosomes (Ohno et al., '61). Indeed, acrocentric chromosomes associate at the heterochromatic regions during meiotic prophase and somatic interphase (Schmid et al., '83; Tuck-Muller et al., '84; Kuznetsova et al., 2007) and we also observed typical "bouquet" figures, where chromosomes are linked together at the level of centromeres (Fig. 4B). This process may be the most important mechanisms for spontaneous chromosomal mutation, concerted evolution, and homogenization of satellite subfamilies of DNA among acrocentric chromosomes (Maeda and Smithies, '86).

Conversely, *TaqI* repeats are pericentromerically located on a lower number of chromosomes (10 pairs in *I. monticola* and 9 in *I. galani*). In this case, we could explain the low homogenization rate within single species in terms of primary rate of the homogenization process. That is, it is possible that the exchange between nonhomologous chromosomes having *TaqI* sequences is limited. The *TaqI* repeats are indeed restricted to a subset of chromosomes in these species and located in a pericentromeric position less prone to physical association: this could reduce interchromosomal exchange and homogenization, thus determining a lower rate of interspecific divergence and a higher degree of intraspecific repeat heterogeneity. Similar considerations were reported for satDNAs of *Rumex*, where repeats in nonrecombining Y chromosomes show low rates of concerted evolution and intraspecific variability increase with no interspecific divergence (Navajas-Pérez et al., 2009; see also Kuhn et al., 2008), and to explain the lower mutation rate of satDNAs in sturgeons as compared to sparids. In fact, the more symmetrical karyotypes of these latter fishes would represent no physical barrier to interchromosomal exchange (de la Herrán et al., 2001a,b). However, also these AT-rich pericentromeric repeats could represent chromosome sites favoring spontaneous rearrangements. Indeed, we observed that the majority of the *TaqI* repeats are flanked by interstitial telomeric sequences that would insert in these chromosome points during

the repair of double strand breaks (see Bolzán and Bianchi, 2006). These unstable sequences might explain the high rate of Robertsonian translocation observed in Pyrenean *Iberolacerta* (Odierna et al., '96).

In conclusion, our study suggests the effect of differential location and repeat copy number in the evolution of satDNAs, revealing features that could also improve the use of this genomic component as a molecular marker in phylogenetic analyses. Moreover, these results indicate that some molecular markers should be used cautiously in species identification when divergence times are shallow among the taxa compared.

#### ACKNOWLEDGMENTS

Special thanks are given to Pedro Galán for providing the samples of *Iberolacerta monticola* and *I. galani* used in this study. All authors declare that no conflict of interests exists for this manuscript.

#### LITERATURE CITED

- Arnold EN, Arribas O, Carranza S. 2007. Systematics of the palaeartic and oriental lizard tribe Lacertini (Squamata: Lacertidae: Lacertinae), with descriptions of eight new genera. *Zootaxa* 1430:1–86.
- Arribas O. 1999. Taxonomic revision of the Iberian "Archaeolacertae" 2: diagnosis, morphology, and geographic variation of "*Lacerta aurelioi*" Arribas, 1994 (Squamata: Sauria: Lacertidae). *Herpetozoa* 11:155–180.
- Arribas O, Carranza S, Odierna G. 2006. Description of a new endemic species of mountain lizard from Northwestern Spain: *Iberolacerta galani* sp. nov. (Squamata: Lacertidae). *Zootaxa* 1240:1–55.
- Bachmann L, Sperlich D. 1993. Gradual evolution of a specific satellite DNA family in *Drosophila ambigua*, *D. tristis*, and *D. obscura*. *Mol Biol Evol* 10:647–659.
- Bandyopadhyay R, McQuillan C, Page SL, Choo KHA, Shaffer LG. 2001. Identification and characterization of satellite III subfamilies to the acrocentric chromosomes. *Chromosome Res* 9:223–233.
- Bolzán AD, Bianchi MS. 2006. Telomeres, interstitial telomeric repeat sequences, and chromosomal aberration. *Mutat Res* 612:189–214.
- Cafasso D, Cozzolino S, Chinali G, De Luca P. 2003. An unusual satellite DNA from *Zamia paucijuga* (Cycadales) characterised by two different organisations of the repetitive unit in the plant genome. *Gene* 311:69–77.
- Capriglione T. 2000. Repetitive DNA as a tool to study the phylogeny of cold-blooded vertebrates. *Chromosomes Today* 13:183–194.
- Capriglione T, Olmo E, Odierna G, Smith DI, Miller OJ. 1989. Genome composition and tandemly repetitive sequence at some centromeres in the lizard *Podarcis s. sicula* Raf. *Genetica* 79:85–91.
- Capriglione T, Cardone A, Odierna G, Olmo E. 1991. Evolution of a centromeric satellite DNA and phylogeny of lacertid lizards. *Comp Biochem Physiol B* 100:641–645.
- Capriglione T, Cardone A, Odierna G, Olmo E. 1994. Further data on the occurrence and evolution of satellite DNA families in the lacertid genome. *Chromosome Res* 2:327–330.



- Capriglione T, De Santo MG, Odierna G, Olmo E. 1998. An alphoid-like satellite DNA sequence is present in the genome of lacertid lizards. *J Mol Evol* 46:240–244.
- Carranza S, Arnold EN, Amat F. 2004. DNA phylogeny of *Lacerta* (*Iberolacerta*) and other lacertine lizards (Reptilia: Lacertidae): did competition cause long-term mountain restriction? *Syst Biodivers* 2:57–77.
- Charlesworth B, Sniegowski P, Stephan W. 1994. The evolutionary dynamics of repetitive DNA in eukaryotes. *Nature* 371:215–220.
- Ciobanu D, Roudykh IA, Grechko VV, Kramerov DA. 2001. Molecular mechanisms of satellite DNA evolution in a group of closely related Caucasian rock lizards. In: Piruzian E, Volkova L, editors. International Symposium Molecular Mechanisms of Genetic Processes and Biotechnology, Moscow [Thesis]. p 26.
- Ciobanu D, Grechko VV, Darevsky IS. 2003. Molecular evolution of satellite DNA CLsat in lizards from the genus *Darevskia* (Sauria: Lacertidae): correlation with species diversity. *Russ J Genetics* 39:1527–1541.
- Ciobanu DG, Grechko VV, Darevsky IS, Kramerov DA. 2004. New satellite DNA in *Lacerta* s. str. lizards (Sauria: Lacertidae): evolutionary pathways and phylogenetic impact. *J Exp Zool* 302B:505–516.
- Crochet PA, Chaline O, Surget-Groba Y, Debain C, Cheylan M. 2004. Speciation in mountains: phylogeography and phylogeny of the rock lizards genus *Iberolacerta* (Reptilia: Lacertidae). *Mol Phylogenet Evol* 30:860–866.
- Csink AK, Henikoff S. 1998. Something from nothing: the evolution and utility of satellite repeats. *Trends Genet* 14:200–204.
- Dawe RK, Henikoff S. 2006. Centromeres put epigenetics in the drivers's seat. *Trends Biochem Sci* 31:662–669.
- de la Herrán R, Fontana F, Lanfredi M, et al. 2001a. Slow rates of evolution and sequence homogenization in an ancient satellite DNA family of sturgeons. *Mol Biol Evol* 18:432–436.
- de la Herrán R, Ruiz Rejon, Ruiz C, Rejon M, Garrido-Ramos MA. 2001b. The molecular phylogeny of the Sparidae (Pisces, Perciformes) based on two satellite DNA families. *Heredity* 87:691–697.
- Dover GA. 1982. Molecular drive: a cohesive mode of species evolution. *Nature* 299:111–117.
- Excoffier L, Smouse P, Quattro J. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* 131:479–491.
- Ezaz T, Quinn AE, Miura I, et al. 2005. The dragon lizard *Pogona vitticeps* has ZZ/ZW micro-sex chromosomes. *Chromosome Res* 13:763–776.
- Feliciello I, Picariello O, Chinali G. 2006. Intra-specific variability and unusual organization of the repetitive units in a satellite DNA from *Rana dalmatina*: molecular evidence of a new mechanism of DNA repair acting on satellite DNA. *Gene* 383:81–92.
- Giovannotti M, Nisi Cerioni P, Caputo V, Olmo E. 2009. Characterization of a GC-rich telomeric satellite DNA in *Eumeces schneideri* Daudin (Reptilia, Scincidae). *Cytogenet Genome Res* 125:272–278.
- Giovannotti M, Nisi Cerioni P, Caputo V. 2010. Mitochondrial DNA sequence analysis reveals multiple Pleistocene glacial refugia for *Podarcis muralis* (Laurenti, 1768) in the Italian Peninsula. *Ital J Zool* 77:277–288.
- Grechko VV, Ciobanu DG, Darevsky IS, Kramerov DA. 2005. Satellite DNA of lizards of the genus *Lacerta* s. str. (the group *L. agilis*), the family Lacertidae. *Dokl Biochem Biophys* 400:44–47.
- Grechko VV, Ciobanu DG, Darevsky IS, Kosushkin SA, Kramerov DA. 2006. Molecular evolution of satellite DNA repeats and speciation of lizards of the genus *Darevskia* (Sauria: Lacertidae). *Genome* 49:1297–1307.
- Helaers R, Milinkovitch MC. 2010. MetaPIGA v2.0: maximum likelihood large phylogeny estimation using the metapopulation genetic algorithm and other stochastic heuristics. *BMC Bioinformatics* 11:379.
- Ijdo JW, Wells RA, Baldini A, Reeders ST. 1991. Improved telomere detection using a telomere repeat probe (TTAGGG)<sub>n</sub> generated by PCR. *Nucleic Acids Res* 19:4780.
- Jantsch M, Hamilton B, Mayr B, Schweizer D. 1990. Meiotic chromosome behavior reflects levels of sequence divergence in *Sus scrofa domestica* satellite III DNA. *Chromosoma* 99:330–335.
- King LM, Cummings MP. 1997. Satellite DNA repeat sequence variation is low in three species of burying beetles in the genus *Nicrophorus* (Coleoptera: Silphidae). *Mol Biol Evol* 14:1088–1095.
- Kuhn GCS, Franco FF, Manfrin MH, Moreira-Filho O, Sene FM. 2007. Low rates of homogenization of the DBC-150 satellite DNA family restricted to a single pair of microchromosomes in species from the *Drosophila buzzatii* cluster. *Chromosome Res* 15:457–469.
- Kuhn GCS, Sene FM, Moreira-Filho O, Schwarzacher T, Heslop-Harrison JS. 2008. Sequence analysis, chromosomal distribution and long-range organization show that rapid turnover of new and old pBuM satellite DNA repeats leads to different patterns of variation in seven species of the *Drosophila buzzatii* cluster. *Chromosome Res* 16:307–324.
- Kuznetsova IS, Erukashvily NI, Noniashvili EM, et al. 2007. Evidence for the existence of satellite DNA-containing connection between metaphase chromosomes. *J Cell Biochem* 101:1046–1061.
- Larkin MA, Blackshields G, Brown NP, et al. 2007. ClustalW and ClustalX version 2.0. *Bioinformatics* 23:2947–2948.
- Librado P, Rozas J. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25:1451–1452.
- Luchetti A, Cesari M, Carrara G, et al. 2003. Unisexuality and molecular drive: Bag320 sequence diversity in *Bacillus* taxa (Insecta Phasmatodea). *J Mol Evol* 56:587–589.
- Mace GM. 2004. The role of taxonomy in species conservation. *Philos Trans R Soc Lond B* 359:711–719.
- Maeda N, Smithies O. 1986. The evolution of multigene families: human haptoglobin genes. *Annu Rev Genet* 20:81–108.
- Mantovani B, Tinti F, Bachmann L, 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 parthenogenic taxa. *Mol Biol Evol* 14:1197–1205.



- Martinsen L, Venanzetti F, Johnsen A, Sbordoni V, Bachman L. 2009. Molecular evolution of the *pDo500* satellite DNA family in *Dolichopoda* cave crickets (Rhaphidophoridae). *BMC Evol Biol* 9:1–14.
- Mayer W, Arribas O. 2003. Phylogenetic relationships of the European lacertid genera *Archaeolacerta* and *Iberolacerta* and their relationships to some other 'Archaeolacertae' (*sensu lato*) from Near East, derived from mitochondrial DNA sequences. *J Zool Syst Evol Res* 41:157–161.
- Miller WJ, Nagel A, Bachmann J, Bachmann L. 2000. Evolutionary dynamics of the SGM transposon family in the *Drosophila obscura* species group. *Mol Biol Evol* 17:1597–1609.
- Modi WS, Ivanov S, Gallagher DS. 2004. Concerted evolution and higher-order repeat structure of the 1.709 (satellite IV) family in bovids. *J Mol Evol* 58:460–465.
- Navajas-Pérez R, de la Herrán R, Jamilena M, et al. 2005. Reduced rates of sequence evolution of Y-linked satellite DNA in *Rumex* (Polygonaceae). *J Mol Evol* 60:391–399.
- Navajas-Pérez R, Ruiz Rejón M, Garrido-Ramos MA, Aznarte JL, Rubio-Escudero C. 2007. SatDNA Analyzer: a computing tool for satellite-DNA evolutionary analysis. *Bioinformatics* 23:767–768.
- Navajas-Pérez R, Quesada del Bosque ME, Garrido-Ramos MA. 2009. Effect of location, organization, and repeat-copy number in satellite-DNA evolution. *Mol Genet Genomics* 282:395–406.
- Nijman IJ, Lenstra JA. 2001. Mutation and recombination in cattle satellite DNA: a feedback model for the evolution of satellite repeats. *J Mol Evol* 52:361–371.
- Odierna G, Aprea G, Arribas O, et al. 1996. The karyology of the Iberian rock lizards. *Herpetologica* 52:542–550.
- Ohno S, Trujillo JM, Kaplan WD, Kinoshita R. 1961. Nucleolus-organizer in the causation of chromosomal abnormalities. *Lancet* 2:123–126.
- Plohl M. 2010. Those mysterious sequences of satellite DNAs. *Periodicum* 112:403–410.
- Plohl M, Luchetti A, Meštrović N, Mantovani B. 2008. Satellite DNAs between selfishness and functionality: structure, genomics and evolution of tandem repeats in centromeric (hetero) chromatin. *Gene* 409:72–82.
- Preacher KJ. 2001. Calculation for the chi-square test: an interactive calculation tool for chi-square tests of goodness of fit and independence [Computer software]. Available online at: <http://quantpsy.org> [Accessed February 13, 2013].
- Rambaut A, Drummond AJ. 2009. Tracer version 1.5 [computer program]. Available online at: <http://beast.bio.ed.ac.uk>.
- Robles F, de la Herrán R, Ludwig A, et al. 2004. Evolution of ancient satellite DNAs in sturgeon genomes. *Gene* 338:133–142.
- Ronquist F, Huelsenbeck JP. 2003. MRBAYES 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19:1572–1574.
- Rudd MK, Wray GA, Willard HF. 2006. The evolutionary dynamics of alpha-satellite. *Genome Res* 16:88–96.
- Salvador A. 1985. Guía de campo de los anfibios y reptiles de la Península Ibérica, islas Baleares y Canarias. León, Spain: Santiago García Ed.
- Sambrook J, Fritsch EF, Maniatis T. 1989. Molecular cloning: a laboratory manual. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
- Schmid M, Grunert D, Haaf T, Engel W. 1983. A direct demonstration of somatically paired heterochromatin of human chromosomes. *Cytogenet Cell Genet* 36:554–561.
- Schneider S, Roessler D, Excoffier L. 2000. Arlequin, version 2.000: a software for population genetic data analysis. Geneva, Switzerland: Genetics and Biometry Laboratory, University of Geneva.
- Schwarzacher T, Heslop-Harrison P. 2000. Practical in situ hybridization. New York, NY: Springer-Verlag.
- Schweizer D. 1976. Fluorescent chromosome banding in plants: applications, mechanisms and implications for chromosome structure. In: Davies DR, Hopwood AP, editors. The plant genome. Proc 4th John Innes Symposium. Norwich, UK: John Innes Charity. p 61–72.
- Sindaco R, Jeremcenko VK. 2008. The reptiles of the Western Palearctic. Latina, Italy: Edizioni Belvedere.
- Smith GP. 1976. Evolution of repeated DNA sequences by unequal crossover. *Science* 191:528–535.
- Stephan W, Cho S. 1994. Possible role of natural selection in the formation of tandem repetitive noncoding DNA. *Genetics* 136:333–341.
- Strachan T, Webb D, Dover G. 1985. Transition stages of molecular drive in multiple-copy DNA families in *Drosophila*. *EMBO J* 4:1701–1708.
- Sumner AT. 1972. A simple technique for demonstrating centromeric heterochromatin. *Exp Cell Res* 75:304–306.
- Swofford DL. 2002. PAUP\* Phylogenetic Analysis Using Parsimony (\*and other methods), Version 4.0b10a. Sunderland, MA: Sinauer Associates.
- Tamura K, Peterson D, Peterson N, et al. 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* 28:2731–2739.
- Tuck-Muller CM, Bordson BL, Varela M, Bennett JW. 1984. NOR associations with heterochromatin. *Cytogenet Cell Genet* 38:165–170.
- Walsh JB. 1987. Persistence of tandem arrays: implications for satellite and simple-sequence DNAs. *Genetics* 115:553–567.
- Watabe H, Bachmann L, Haring E, Sperlich D. 1997. Taxonomic and molecular studies on *Drosophila sinobscura* and *D. hubeiensis*, two sibling species of the *D. obscura* group. *J Zool Syst Evol Res* 35:81–94.

## ***CHAPTER II***

**Evolutionary dynamics of two satellite DNA families in  
Rock lizards of the genus *Iberolacerta* (Squamata,  
Lacertidae): different histories but common traits**

## Evolutionary dynamics of two satellite DNA families in rock lizards of the genus *Iberolacerta* (Squamata, Lacertidae): different histories but common traits

Verónica Rojo · Andrés Martínez-Lage · Massimo Giovannotti · Ana M. González-Tizón · Paola Nisi Cerioni · Vincenzo Caputo Barucchi · Pedro Galán · Ettore Olmo · Horacio Naveira

Published online: 18 September 2015  
© Springer Science+Business Media Dordrecht 2015

**Abstract** Satellite DNAs compose a large portion of all higher eukaryotic genomes. The turnover of these highly repetitive sequences is an important element in genome organization and evolution. However, information about the structure and dynamics of reptilian satellite DNA is still scarce. Two satellite DNA families, HindIII and TaqI, have been previously characterized in four species of the genus *Iberolacerta*. These families showed different

chromosomal locations, abundances, and evolutionary rates. Here, we extend the study of both satellite DNAs (satDNAs) to the remaining *Iberolacerta* species, with the aim to investigate the patterns of variability and factors influencing the evolution of these repetitive sequences. Our results revealed disparate patterns but also common traits in the evolutionary histories of these satellite families: (i) each satellite DNA is made up of a library of monomer variants or subfamilies shared by related species; (ii) species-specific profiles of satellite repeats are shaped by expansions and/or contractions of different variants from the library; (iii) different turnover rates, even among closely related species, result in great differences in overall sequence homogeneity and in concerted or non-concerted evolution patterns, which may not reflect the phylogenetic relationships among taxa. Contrasting turnover rates are possibly related to genomic constraints such as karyotype architecture and the interspersed organization of diverging repeat variants in satellite arrays. Moreover, rapid changes in copy number, especially in the centromeric HindIII satDNA, may have been associated with chromosomal rearrangements and even contributed to speciation within *Iberolacerta*.

Responsible Editors: Maria Assunta Biscotti, Pat Heslop-Harrison and Ettore Olmo.

**Electronic supplementary material** The online version of this article (doi:10.1007/s10577-015-9489-1) contains supplementary material, which is available to authorized users.

V. Rojo · A. Martínez-Lage · A. M. González-Tizón · H. Naveira (✉)  
Grupo de Investigación en Biología Evolutiva, Departamento de Biología Celular e Molecular, Universidade da Coruña, E-15071 A Coruña, Spain  
e-mail: horacio.naveira.fachal@udc.es

M. Giovannotti · P. N. Cerioni · V. C. Barucchi · E. Olmo  
Dipartimento di Scienze della Vita e dell'Ambiente, Università Politecnica delle Marche, via Brecce Bianche, 60131 Ancona, Italy

V. C. Barucchi  
Consiglio Nazionale delle Ricerche, Istituto di Scienze Marine  
Sezione Pesca Marittima, Largo Fiera della Pesca, 60125 Ancona, Italy

P. Galán  
Grupo de Investigación en Biología Evolutiva, Departamento de Biología Animal, Biología Vegetal e Ecología, Universidade da Coruña, E-15071 A Coruña, Spain

**Keywords** Concerted evolution · FISH · *Iberolacerta* · Library model · Satellite DNA · Squamate reptiles

### Abbreviations

Cy3      Cyanine 3  
dNTP      Deoxyribonucleotide triphosphate



FCA	Factorial correspondence analysis
FISH	Fluorescence in situ hybridization
FITC	Fluorescein iso-thiocyanate
Mya	Million years ago
$\pi$	Nucleotide diversity
satDNA	Satellite DNA

## Introduction

Satellite DNAs (satDNAs) represent one of the major classes of repetitive sequences in almost all eukaryotic genomes. They consist of tandemly repeated non-coding DNA sequences, typically arranged in large clusters of hundreds or thousands of copies usually located in the heterochromatic regions of chromosomes, close to the centromeres and telomeres (Charlesworth et al. 1994). Several satDNA families of independent origin are commonly found in the genome of a species or group of species, and they usually differ in nucleotide sequence, monomer length, and complexity, as well as in evolutionary history (Ugarković and Plohl 2002; Kuhn et al. 2008, 2010). The biological function of these sequences is not yet fully understood, although numerous reports point out the role of certain satellites in centromeric condensation, chromosome organization, or chromosome pairing (see Plohl et al. 2008). A growing field of research is also addressing the role of satDNA transcripts in the formation and maintenance of heterochromatin and even in regulation of gene expression (Ugarković 2009; Pezer et al. 2012). In addition, several examples support the hypothesis that the rapid evolution of satDNAs can act as a driver of population and species divergence (Ugarković and Plohl 2002; Felicicelo et al. 2015).

Despite their biological significance, satDNAs are still the least understood genomic component, underrepresented in outputs of most genome projects (Plohl et al. 2012). A common feature of many of them is that, even though monomers can be present in many thousand copies per genome, sequence divergence between repeats of the same family is often very low, usually less than 15 % (Plohl et al. 2008). The non-independent or concerted evolution of repeat units is postulated to be a consequence of a two-step process called molecular drive, consisting of the gradual spread of a sequence variant (1) through a genome (homogenization) and (2) through a species (fixation) (Dover 1982). Sequence

homogenization is due to diverse molecular mechanisms of nonreciprocal transfer, such as unequal crossing-over, gene conversion, rolling circle replication and reinsertion, and transposon-mediated exchange (Stephan 1986; Dover 2002), while fixation results from random chromosomal assortment in sexual reproduction, depending thus on population factors. This process results in rapid divergence of satellite sequences in reproductively isolated groups of organisms, and in this case, satDNAs can be used as phylogenetically informative markers (Plohl et al. 2012).

Accumulation of mutations in satellite families is not the only way to alter specific profiles of satellite repeats in short evolutionary periods. In addition to sequence changes, satDNAs are permanently altered in copy number by expanding and contracting arrays of satellite monomers (Ugarković and Plohl 2002; Plohl et al. 2012). Because usually more than one satellite family exists in a genome, fluctuations in their copy numbers can change very efficiently and rapidly any profile of genomic satDNA. The library model of satDNA evolution explains the occurrence of species-specific satellite profiles as a result of differential amplifications and/or contractions within a collection, or library, of satellite sequences shared by related species (Fry and Salser 1977; Meštrović et al. 1998; Ugarković and Plohl 2002). Not only distinct satDNAs but also monomer variants or subfamilies from a single family can be distributed in genomes in the form of a library (Cesari et al. 2003).

SatDNAs have been extensively studied in insects (Palomeque and Lorite 2008) and mammals (Erukashvily and Ponomartsev 2013), and less so in other taxa, although there are several exceptions. Squamata, by far the largest reptile order, is one of them (see, for example, Giovannotti et al. 2009, 2013; Chairasertsri et al. 2013). It includes the Lacertidae, a widespread species-rich group restricted to the Palearctic region, formed by two subfamilies, Gallotiinae and Lacertinae (Arnold et al. 2007; Sindaco and Jeremčenko 2008). So far, five satDNA families have been described in Lacertinae, with different taxonomic distributions. Three satellite families are genus-specific, namely, pLHS in *Podarcis* (Capriglione et al. 1994; Capriglione 2000), CLsat in *Darevskia* (Ciobanu et al. 2003; Grechko et al. 2006), and Agi160 in *Lacerta* (Ciobanu et al. 2004; Grechko et al. 2005). The other two families, on the contrary, are broadly distributed in Lacertinae: pLCS, shared by *Algyroides*, *Teira*, *Lacerta*, and *Podarcis*



(Capriglione et al. 1989, 1991; Capriglione 2000), and pGPS, present in *Podarcis*, *Archaeolacerta*, *Algyroides*, *Lacerta*, and *Zootoca* (Capriglione et al. 1998).

In a previous work (Giovannotti et al. 2014) we isolated two new satDNA families in the lacertid genus *Iberolacerta*, a monophyletic group of rock lizards mainly distributed in highland areas of Western Europe. This genus comprises eight species, which can be subdivided into three main units: (1) *I. horvathi*, occurring in the Eastern Alps and the north of the Dinaric Chains; (2) the subgenus *Pyrenesaura*, which includes the three species found in the Pyrenees, (*I. aranica*, *I. aurelioi*, and *I. bonnali*); and (3) the four species included in the "Iberian group" (*I. cyreni*, *I. martinezricai*, *I. galani*, and *I. monticola*), with disjunct distributions in central and northern mountain ranges of the Iberian Peninsula. Previous cytogenetic surveys of the *Iberolacerta* species (Capula et al. 1989; Odierna et al. 1996; Arribas and Odierna 2004; Arribas et al. 2006; Rojo et al. 2014) showed them to possess a diploid number of  $2n=36$ , and a similar karyotypic macrostructure, with all chromosomes acrocentric. Only the karyotypes of the three Pyrenean species differ from this formula, with reduced diploid numbers that range from  $2n=24$  to 26 in males and from  $2n=23$  to 26 in females, and many banded chromosomes that probably evolved from the ancestral acrocentric complement through a series of Robertsonian fusions (Odierna et al. 1996).

According to the most recently published phylogeny (Arribas et al. 2014), speciation within *Iberolacerta* started ca. 13.5 million years ago (Mya; 95 % credibility interval 11.6–15.6), with the split between the clades formed by *I. horvathi* and the Iberian group, on one side, and by the Pyrenean species, on the other. This event was most likely quickly followed by the separation of *I. horvathi*, which took place approximately 11.5 Mya (9.6–13.7). Within the Iberian group, *I. cyreni* split earlier (7.3–8.5 Mya), while the speciation events within the clade formed by *I. martinezricai*, *I. galani*, and *I. monticola* occurred considerably later, at the beginning of the Pleistocene, 2.1–2.9 Mya. The three Pyrenean species probably originated in rapid succession ca. 3.8 Mya (2.7–4.9), although this phylogenetic analysis suggests that *I. bonnali* split first, shortly before the separation between *I. aranica* and *I. aurelioi*, 3.3 Mya (2.3–4.3). Notwithstanding minor uncertainties still remaining, the mapping of satDNA differences on that species tree is likely to provide valuable information

about the time and mode of evolution of these repetitive sequences. In our previous work (Giovannotti et al. 2014), we analyzed two unrelated satDNA arrays in the Iberian clade of *Iberolacerta*: (1) the centromeric HindIII family, which comprises two subfamilies (I and II) and represents 5–10 % of the genome and (2) the TaqI family, which shows only interstitial loci and represents 2.5–5 % of the genome. The nucleotide sequences of the two families were presumably evolving at different rates, almost tenfold higher for centromeric than for interstitial repeats, after comparing *I. cyreni* vs. the other, relatively closer, species of the Iberian clade. In agreement with this conclusion, the HindIII family seems to be specific to the genus *Iberolacerta* (Capriglione et al. 1989, 1991, 1998; Capriglione 2000), whereas the TaqI satDNA has also been detected in representatives of three other genera of the subfamily Lacertinae (*Lacerta*, *Podarcis*, and *Timon*).

Here, we extend the study of both satDNAs to the remaining *Iberolacerta* species, and increase our dataset for HindIII satDNA, to further investigate the occurrence of two divergent subfamilies in the genomes of all these taxa. The results obtained offer a more complete portrait of the intra- and interspecific variability of these highly repetitive sequences and their genomic organization and chromosomal distribution, with the ultimate objective of contributing to assess the relative strength of the processes that determine their structure and mode of evolution.

## Material and methods

### Animals

Genomic DNA was isolated from a total of 20 specimens, representing all eight *Iberolacerta* species. The number of specimens per species and their geographical origin are given in Supplementary Table 1. In addition, one male and one female of *I. horvathi* and one female of *I. bonnali* were used to make metaphase chromosomes.

### DNA extraction, PCR, cloning and sequencing

Genomic DNA was extracted from ethanol preserved tissues using standard protocols with proteinase K digestion followed by phenol/chloroform extraction (see Sambrook et al. 1989). Two primer pairs designed in our previous work (HindIII-F: 5'-



TGAGTGTTTTACAGTTGAAAAGCT-3'; HindIII-R: 5'-CATTGTGTTATTTGAGCGCAA-3'; TaqI-F: 5'-ATTCTGACCCTGGGGGTTAG-3'; TaqI-R: 5'-CATATTTAAAGAAATCAGGCCTCG-3') were used for isolation of both satellite families from the genomes of *I. horvathi*, *I. bonnali*, *I. aranica*, and *I. aurelioi*. An additional primer pair was designed to specifically amplify HindIII-subfamily II in all eight *Iberolacerta* species (Hind\_sflI-F: 5'-CTCTTGCTTATTTTCGCTCCAATGA-3'; Hind\_sflI-R: 5'-ATTTCTGTGTGCAGCATGCAT TGG-3'). PCR reactions were performed in a final volume of 25  $\mu$ l containing ~25 ng of genomic DNA, 0.625 U of *Taq* DNA polymerase and 1 $\times$  PCR buffer (Roche Diagnostics), 5 nmol of each dNTP (Roche Diagnostics), and 20 pmol of each primer. The general reaction conditions were as follows: initial denaturation at 94 °C for 5 min; 35 cycles of denaturation at 94 °C for 30 s, annealing at the following temperatures (HindIII-F/HindIII-R, 55 °C; TaqI-F/TaqI-R, 47 °C; Hind\_sflI-F/Hind\_sflI-R, 58 °C) for 30 s, extension at 72 °C for 30–60 s, and a final extension at 72 °C for 7 min. The obtained PCR products were run on 1.5 % agarose gels; DNA in bands of interest was eluted using Pure Link Quick Gel Extraction Kit (Invitrogen) and cloned in the T&A cloning vector with T&A cloning kit (Yeastern Biotech) following manufacturer's recommendations. Positive clones were selected through PCR amplification using the M13 forward and M13 reverse primers. Bidirectional sequencing with the M13 primers was performed on an ABI PRISM 3730XL (Applied Biosystems) automatic sequencer.

#### Sequence analysis

The newly sequenced repeats were analyzed together with the previously reported sequences of the HindIII and TaqI satDNA families from *I. cyreni*, *I. monticola*, *I. galani*, and *I. martinezricai* (DDBJ/EMBL/GenBank accession numbers for HindIII: from KF453637 to KF453681; accession numbers for TaqI: from KF453682 to KF453723) (Giovannotti et al. 2014). Multiple sequence alignment was performed with MUSCLE (Edgar 2004), using default parameters, as implemented in Geneious version 8.0.5 (Kearse et al. 2012). After visual inspection of alignments, sequences

were classified into different sets according to shared nucleotide changes and indels.

Intraspecific nucleotide diversity ( $\pi$ ) was estimated using DnaSP v. 5 (Librado and Rozas 2009). Net average genetic distances between groups were calculated using the Maximum Composite Likelihood model (Tamura et al. 2004) in MEGA v. 6.0 (Tamura et al. 2013). Sequence variability among satellite repeats was further investigated by performing a factorial correspondence analysis (FCA), carried out with Genetix v. 4.05.2 (Belkhir et al. 2004). For this analysis, we constructed a matrix with all the sequences, where the nucleotide present at each diagnostic position was coded with a unique integer (100, 120, 140, or 160).

For the subsequent phylogenetic analysis, a consensus sequence was obtained for each sequence set by choosing the most frequent nucleotide at each position, except when a combination of dinucleotides of the three pairs CpG, CpA, and TpG was present at the same doublet position. In that case, the CpG dinucleotide was chosen as the consensus unless the T or A nucleotides were present in >70 % of the sequences. A phylogenetic network of the consensus sequences was constructed with TCS v. 1.21 (Clement et al. 2000) using the statistical parsimony algorithm under the 95 % parsimony criterion (Templeton et al. 1992).

#### Chromosome analysis

Metaphase chromosome spreads were prepared as described previously (Giovannotti et al. 2014). As for *I. horvathi*, individuals of this species were induced to autotomize their tail tips, the tissues were collected in the field following the protocol by Waters et al. (2008) and transferred to the laboratory for the establishment of primary cell cultures. For fluorescence in situ hybridization (FISH) experiments, we developed species-specific probes obtained by PCR amplification of HindIII and TaqI satDNA clones. The probes were labeled either with Cy3, using a PCR labeling kit (Jena Bioscience), or with FITC, using the Platinum Bright 495 labeling kit (KREATECH Biotechnology). Slide pretreatment, denaturation, hybridization, post-hybridization washes, and detection were performed according to Schwarzacher and Heslop-Harrison (2000). Images were captured using the epifluorescence microscopes (Nikon Microphot-FXA; Leica Leitz



DMRBE) equipped with monochrome cameras (Nikon DS-Qi1Mc; JAI CV-M4+CL). The NIS-Elements D 3.10 (Nikon Instruments) and Leica CytoVision version 7.2 (Leica Microsystems) softwares were used to process the images and reconstruct the karyotypes.

## Results

### Isolation and characterization of satellite DNAs

PCR amplification using primers specific for HindIII and TaqI satDNA was successful in all tested species and produced a ladder-like banding pattern, which is typical for satellite DNA. PCR products included complete monomers and multimers (from dimers up to hexamers), flanked by partial monomer sequences. Only clones with complete repeat units were sequenced and, for further analyses, multimers were separated into individual monomers. A total of 187 new sequences were obtained for HindIII, whereas 109 clones were sequenced for TaqI. Comparison of these new sequences with the HindIII and TaqI monomers isolated from *I. cyreni*, *I. monticola*, *I. galani*, and *I. martinezricai* in our previous study (Giovannotti et al. 2014) indicated that all of them belong to the same satDNA families. Altogether, our dataset comprises 232 HindIII and 151 TaqI monomers from all eight *Iberolacerta* species, which are likely to reflect the overall variability of the two satellite families in the genus.

Both HindIII and TaqI satDNAs are characterized by an AT bias (average AT content of 58.9 and 59.1 %, respectively) and by the occurrence of short repeat motifs such as A and T stretches, dinucleotide TG and CA, and trinucleotide CAA and TTC (Supplementary Figs. 1a, b). The size of HindIII repeats ranged between 169 and 172 bp, with the exception of two monomers with lengths of 151 bp (IAR\_99b) and 161 bp (ICY\_209c) (Table 1). TaqI repeats showed a broader range of length variation, from 155 to 191 bp (Table 1). Several indels varying in size from 1 to 31 bp are the causes of the repeat length variation in this satDNA family.

After alignment, monomers within each satDNA family were classified into subfamilies, according to the state of diagnostic positions, characterized by nucleotide substitutions or indels shared by at least 90 % of all the members grouped in the same subfamily. The subfamilies were designated with Roman numerals

following the nomenclature previously used in Giovannotti et al. (2014) for HindIII subfamilies I and II. Additional diagnostic positions further divided each subfamily into several sequence groups and subgroups, denoted by a Latin letter and a numeral, respectively, after the subfamily name (Table 2).

### Sequence variability within HindIII satDNA

Within HindIII satDNA, we found a total of 30 diagnostic positions, which identified three subfamilies—namely HI, HII, and HIII—and 27 sequence groups (Table 2a and Supplementary Fig. 1a). Their abundances ranged from 1.3 to 17 % (3–39 representatives) of the examined sequences. Figure 1a overlays data on the abundance and distribution of HindIII sequence groups onto a phylogenetic tree for *Iberolacerta* derived from mitochondrial markers (Arribas et al. 2014). As evidenced in this figure, sequence groups were not equally represented in the different species. The Pyrenean species (*I. aurelioi*, *I. aranica*, and *I. bonnali*) harbor a wide diversity of HindIII repeats, mainly belonging to subfamilies HI and HII. Only 12 monomers were retrieved from *I. horvathi*, and they are all members of subfamily HI. Similarly, subfamily HI is also the most abundant variant of the HindIII family in the Iberian species *I. martinezricai*, *I. monticola*, and *I. galani*. A strikingly different profile of HindIII repeats was found in *I. cyreni*, also an Iberian species, which is characterized by the presence of several private sequence groups belonging to subfamily HIII and one exclusive sequence group within subfamily HI.

The coexistence of more than one subfamily explains the higher nucleotide diversity values ( $\pi$ ) in species such as *I. bonnali* (4.91%) or *I. aurelioi* (3.96%), in comparison with the values obtained for those species in which all their HindIII repeats belonged to a single subfamily, i.e., *I. horvathi* (1.16%) and *I. martinezricai* (1.51%) (Table 1). Interestingly, despite their different abundances, mean  $\pi$  values for each subfamily were roughly similar (from 2.30 % in subfamily HII to 2.54 % in subfamily HIII).

The factorial correspondence analysis (FCA) based on diagnostic positions highlighted the differentiation among the three HindIII subfamilies, lending further support to our classification. Altogether, the three main axes of variation explain 96.53 % of the observed variation (Fig. 2a). The most informative is axis 1 (69.70 %), which identifies two main clusters,

**Table 1** Summary of repeat features of HindIII and TaqI satDNA

Species	HindIII				TaqI			
	Subfamily	n	Repeat length	Nucleotide diversity ( $\pi$ )	Subfamily	n	Repeat length	Nucleotide diversity ( $\pi$ )
<i>I. monticola</i>	All combined	34		0.0151±0.0018	All combined	10		0.0600±0.0089
	HI	30	171	0.0142±0.0023	TI	10	171–188	0.0600±0.0089
	HII	4	170	0.0177±0.0060				
<i>I. galani</i>	All combined	31		0.0331±0.0040	All combined	16		0.0489±0.0001
	HI	23	171	0.0148±0.0019	TI	16	186–188	0.0489±0.0001
	HII	8	169–170	0.0211±0.0082				
<i>I. martinæricai</i>	All combined	33		0.0151±0.0018	All combined	7		0.0541±0.0103
	HI	33	171–172	0.0151±0.0018	TI	7	187–188	0.0541±0.0103
<i>I. cyreni</i>	All combined	40		0.0356±0.0037	All combined	9		0.0406±0.0001
	HI	7		0.0180±0.0030	TI	9	186–187	0.0406±0.0001
	HIII	33	161–171	0.0240±0.0029				
<i>I. horvathi</i>	All combined	12		0.0116±0.0028	All combined	33		0.1218±0.0079
	HI	12	171	0.0116±0.0028	TI	31	167–191	0.1184±0.0083
					TII	2	189–191	0.0699±0.0349
<i>I. aurelioi</i>	All combined	25		0.0396±0.0034	All combined	20		0.0976±0.0086
	HI	14	171	0.0290±0.0048	TI	1	187	
	HII	11	170	0.0262±0.0026	TII	19	177–188	0.0908±0.0074
<i>I. aratica</i>	All combined	22		0.0355±0.0043	All combined	34		0.1209±0.0070
	HI	7	151–171	0.0265±0.0055	TI	14	175–190	0.1082±0.0126
	HII	15	170	0.0164±0.0028	TII	20	177–190	0.0960±0.0059
<i>I. bonnali</i>	All combined	35		0.0491±0.0050	All combined	22		0.1204±0.0096
	HI	17	171	0.0257±0.0027	TI	17	155–188	0.1060±0.0102
	HII	15	169–170	0.0230±0.0076	TII	5	177–190	0.0983±0.0156
	HIII	3	171	0.0195±0.0033				
All species combined	HI	143		0.0241±0.0015	TI	105		0.1342±0.0060
	HII	53		0.0230±0.0018	TII	46		0.0961±0.0044
	HIII	36		0.0254±0.0029				
	TOTAL	232		0.0539±0.0020	TOTAL	151		0.1567±0.0038

Number of monomeric repeats sequenced ( $n$ ), length of repeats (expressed in base pairs), and nucleotide diversities ( $\pi$ )±S.E. for both satDNAs for each *Iberolacerta* species investigated

corresponding to subfamily HIII repeats of *I. cyreni* and *I. bonnali* on one side, and to subfamilies HI and HII on the other. Axis 2, which accounts for 24.60 % of the observed variation, separates subfamilies HI and HII. Finally, axis 3, with 2.23 % of the observed variation, probably corresponds to sequence heterogeneity within each subfamily. The clustering of HindIII repeats revealed by the FCA matches the estimates of interspecies and inter-subfamilies net genetic distances, shown in Table 3a. Monomers of subfamily HIII are the

most divergent, with average genetic distances of 7.50 and 9.90 % from subfamily HI and HII, respectively. These values are substantially higher than the average distance between subfamilies HI and HII (around 4.0 %). When *I. cyreni* is excluded from the analysis, pairwise interspecies genetic distances within each subfamily are all very low and uncorrelated with relative divergence times between species, with average values of 1.0 % within subfamily HIII, 0.34 % within subfamily HII, and 0.33 % within subfamily HI. Net genetic distances



**Table 2** Nucleotide differences among the consensus sequences of the different groups of (a) HindIII subfamilies HI, HII, and HIII, and (b) TaqI subfamilies TI and TII. The second row refers to base positions relative to the alignment shown in Supplementary Fig. 1a (HindIII) and 1b (TaqI). The general consensus sequence of each satDNA was used as reference. Dots indicate identity with this reference sequence

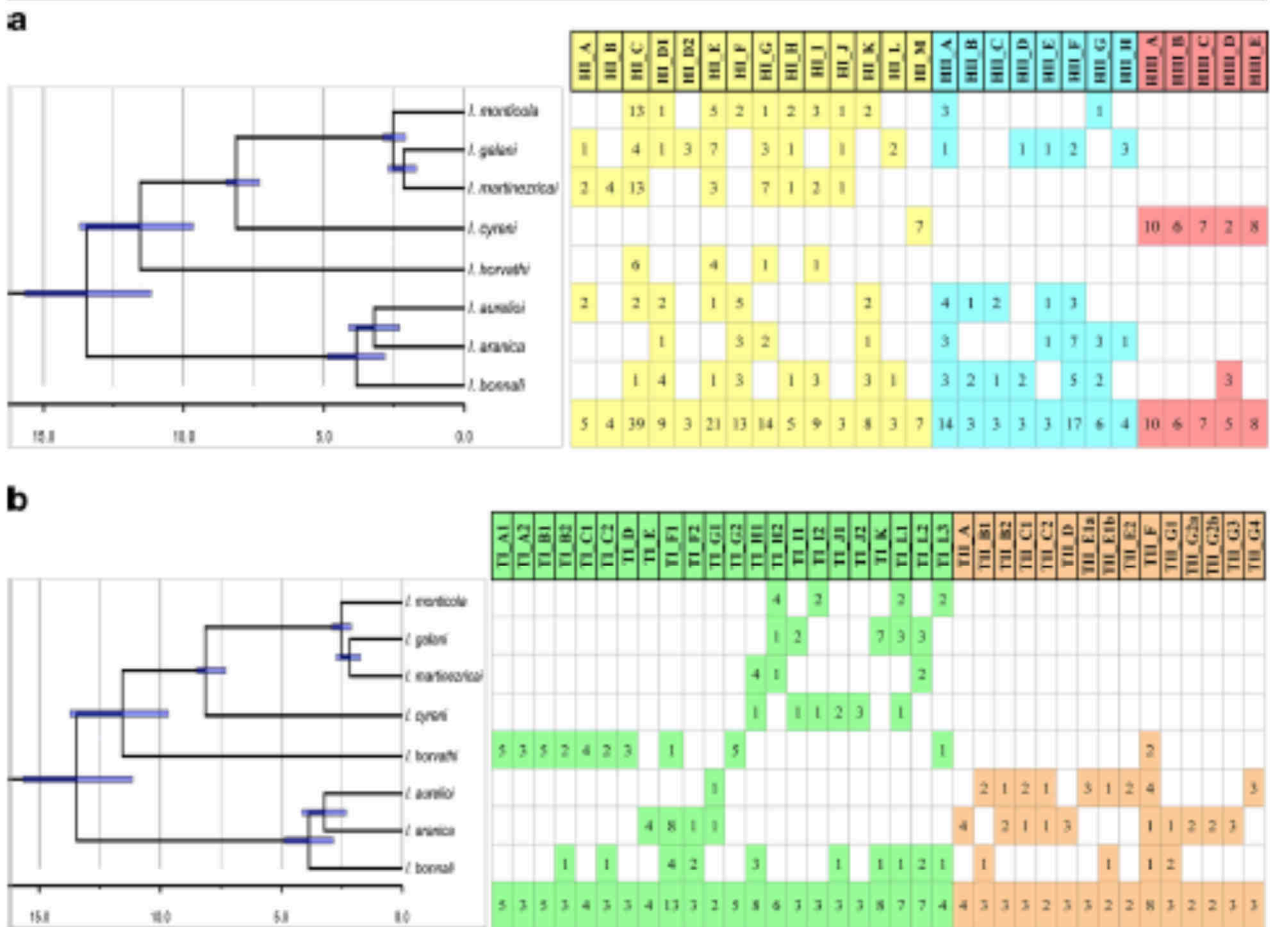
a)		1	2	3	4	5	6	7	8	10	11	12	13	14	15	17	18	19	20	21	22	23	24			
Positions		14	15	21	27	38	39	56	73	83	84	85	86	87	95	98	99	99	101							
Consensus		T	C	A	T	T	T	C	A	A	A	T	T	T	C	T	T	G	A	G	G	T	A			
HI_A		.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.			
HI_B		.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.			
HI_C		.	.	.	.	.	.	T	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.			
HI_D1		.	.	.	.	.	.	T	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.			
HI_D2		.	.	G	.	.	.	T	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.			
HI_E		.	.	.	.	.	.	T	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.			
HI_F		.	.	.	.	.	.	T	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.			
HI_G		.	.	.	C	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.			
HI_H		.	.	.	C	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.			
HI_I		.	G	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.			
HI_J		.	G	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.			
HI_K		.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.			
HI_L		.	.	.	C	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.			
HI_M		.	.	.	C	.	G	T	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.			
HII_A		.	.	.	C	.	G	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G			
HII_B		.	.	.	C	.	G	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.			
HII_C		.	.	.	.	C	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.			
HII_D		.	.	G	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.			
HII_E		A	.	G	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.			
HII_F		.	.	G	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.			
HII_G		.	.	G	.	C	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G			
HII_H		.	.	G	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G			
HII_I		.	.	.	C	.	G	T	.	G	.	.	.	G	.	A	.	.	.	A	.	.	.			
HII_J		.	.	.	C	.	G	T	.	G	.	.	.	G	.	.	.	.	A	.	.	.	.			
HII_K		.	.	.	C	.	G	T	.	G	.	.	.	G	.	.	.	.	A	.	.	.	.			
HII_L		.	.	.	C	.	G	T	.	G	.	.	.	G	.	.	.	.	A	.	.	.	.			
HII_M		.	.	.	C	.	G	T	.	G	.	.	.	G	.	.	.	.	A	.	.	.	.			
HII_N		.	.	.	C	.	G	T	.	G	.	.	.	G	.	.	.	.	A	.	.	.	.			
HII_O		.	.	.	C	.	G	T	.	G	.	.	.	G	.	.	.	.	A	.	.	.	.			
HII_P		.	.	.	C	.	G	T	.	G	.	.	.	G	.	.	.	.	A	.	.	.	.			
HII_Q		.	.	.	C	.	G	T	.	G	.	.	.	G	.	.	.	.	A	.	.	.	.			
HII_R		.	.	.	C	.	G	T	.	G	.	.	.	G	.	.	.	.	A	.	.	.	.			
HII_S		.	.	.	C	.	G	T	.	G	.	.	.	G	.	.	.	.	A	.	.	.	.			
HII_T		.	.	.	C	.	G	T	.	G	.	.	.	G	.	.	.	.	A	.	.	.	.			
TI		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	
Positions		1	5	6	9	14	16	17	18	19	20	21	22	31	33	43	54	56	58	59	60	61	66	71	72	
Consensus		C	G	C	G	C	T	A	A	C	C	T	A	A	C	C	T	T	C	C	A	G	G	C	C	G
TI_A1		.	.	.	.	.	.	C	.	G	G	.	.	.	.	T	.	C	.	.	G	C	.	T	.	.
TI_A2		.	.	.	.	.	.	C	.	G	G	.	.	.	.	T	.	C	.	.	G	C	.	T	.	.
TI_B1		.	.	.	.	.	.	C	.	G	G	.	.	.	.	T	.	C	.	.	G	C	.	T	.	.
TI_B2		.	.	.	.	.	.	C	.	G	G	.	.	.	.	T	.	C	.	.	G	C	.	T	.	.
TI_C1		G	.	T	.	.	.	C	.	G	G	.	G	.	.	.	.	C	.	.	G	C	.	T	.	.











**Fig. 1** Distribution and abundance of HindIII (a) and TaqI (b) subfamilies in *Iberolacerta* coupled to a Bayesian tree obtained from two mitochondrial loci (Cyt b, cytochrome b; CR, control region) (adapted from Arribas et al. 2014). Node bars indicate 95 %

credibility intervals (regions of highest posterior density) for the corresponding divergence time (in million years). Numbers in the table indicate the number of repeats of each subfamily retrieved from each species. Colors identify different subfamilies

between HI repeats involving *I. cyreni* are always considerably higher (from 2.0 % between *I. cyreni* and *I. aranica* to 3.40 % between *I. cyreni* and *I. horvathi*).

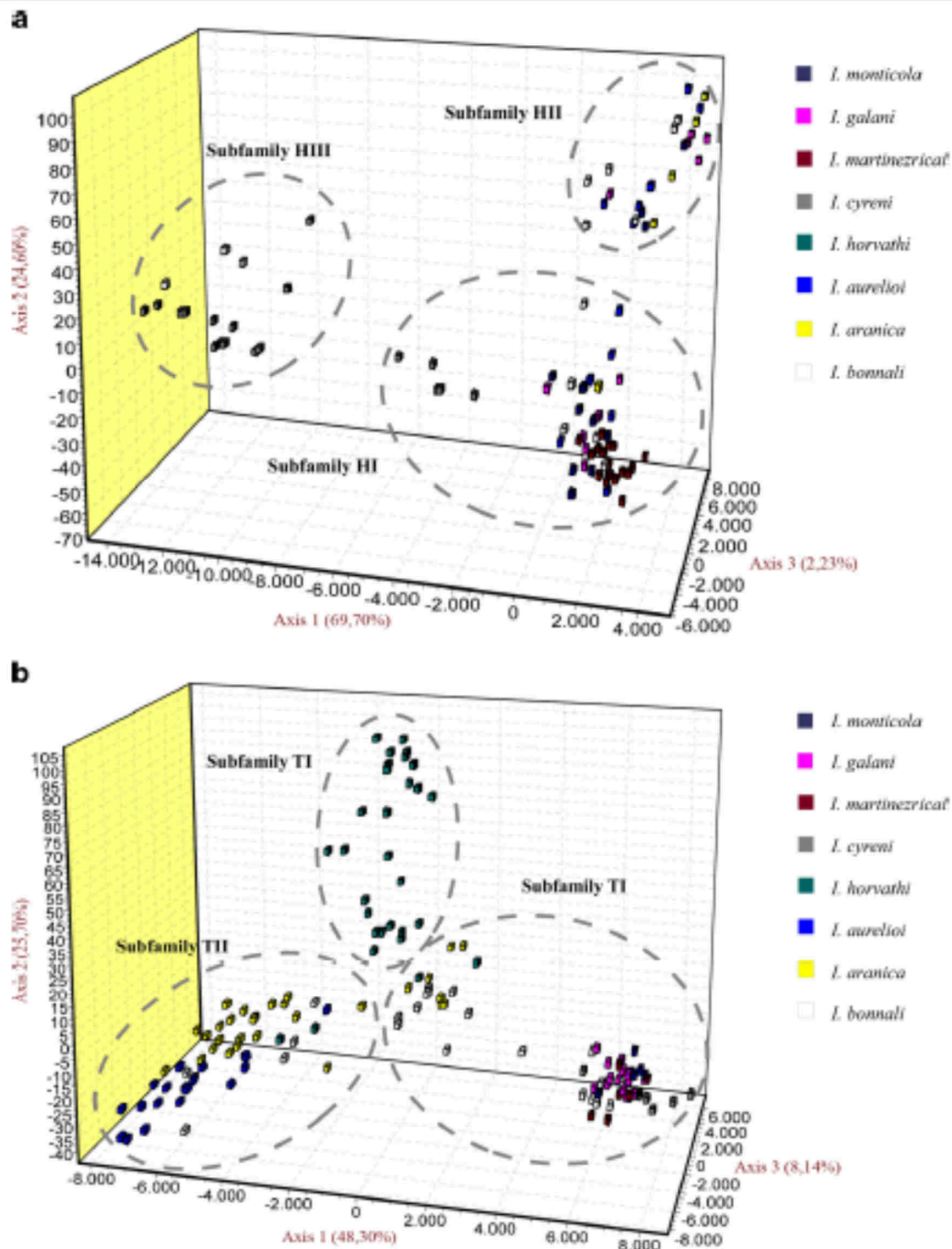
Sequence variability within TaqI satDNA

From the alignment of TaqI sequences, we identified a total of 50 diagnostic positions, which defined two main subfamilies—namely TI and TII—and 37 sequence groups, whose abundances ranged from 1.3 to 8.5 % (2–13 representatives) of the examined sequences (Table 2b and Supplementary Fig. 1b).

In general, the species of the Iberian clade were characterized by the presence of TaqI repeats belonging only to subfamily TI (Fig. 1b), with a substantial proportion of private sequence groups (four groups, comprising 15 out of 42 sequences).

Conversely, subfamily TII is essentially characteristic of the subgenus *Pyrenesaura*, although it has been residually observed also in *I. horvathi*. This subfamily appears to be the most abundant variant in the genomes of *I. aranica* and, above all, *I. aurelioi*, which show both species-specific and shared sequence groups. The sampled loci from *I. bonnali* and *I. horvathi* contain mostly T1 repeats. However, the clustering pattern of T1 repeats differs markedly between the two species: while all the monomers retrieved from *I. bonnali* were grouped together with monomers from other species, *I. horvathi* shows the highest proportion of species-specific repeats (25 out of 33), allocated to six private sequence groups.

As expected from the distribution of subfamilies TI and TII in the genomes of the *Iberolacerta* species, intraspecific nucleotide diversity values



**Fig. 2** Three-dimensional representation of a factorial correspondence analysis based on monomeric sequences of HindIII (a) and TaqI (b) satDNAs

are higher for *I. horvathi* and the Pyrenean species, which harbor both types of TaqI repeats in their genomes (Table 1). When each subfamily is analyzed separately,

$\pi$  values within subfamily TI are two- to threefold greater in these species than in the species of the Iberian clade (from 4.06 % in *I. cyreni* to 11.84 % in *I. horvathi*). High



**Table 3** Interspecific and inter-subfamily net genetic distances for HindIII (a) and TaqI (b) repeats. Standard error estimates are shown above the diagonal. Color codes represent the different

types of HindIII and TaqI subfamilies. Asterisks in b indicate those values obtained in comparisons involving IAU\_TI, represented by only one sequence

**a**

	IGA_HI	IMR_HI	IAU_HI	IHO_HI	IMO_HI	IBN_HI	IAR_HI	ICY_HI	IGA_HII	IBN_HII	IAU_HII	IAR_HII	IMO_HII	ICY_HII	IBN_HII
IGA_HI		0.001	0.003	0.001	0.000	0.002	0.003	0.011	0.017	0.016	0.016	0.017	0.016	0.023	0.026
IMR_HI	0.001		0.004	0.001	0.000	0.003	0.004	0.013	0.018	0.017	0.017	0.017	0.017	0.024	0.026
IAU_HI	0.006	0.008		0.005	0.003	0.000	0.001	0.011	0.015	0.013	0.013	0.014	0.012	0.023	0.025
IHO_HI	0.001	0.001	0.009		0.000	0.003	0.005	0.015	0.019	0.018	0.017	0.018	0.018	0.026	0.028
IMO_HI	0.000	0.001	0.005	0.000		0.002	0.003	0.012	0.017	0.016	0.016	0.017	0.016	0.024	0.026
IBN_HI	0.003	0.005	0.000	0.006	0.003		0.001	0.010	0.016	0.014	0.013	0.015	0.012	0.022	0.025
IAR_HI	0.005	0.006	0.001	0.008	0.004	0.001		0.010	0.016	0.014	0.013	0.015	0.013	0.021	0.024
ICY_HI	0.026	0.028	0.023	0.034	0.029	0.022	0.020		0.022	0.018	0.019	0.021	0.018	0.016	0.020
IGA_HII	0.047	0.048	0.039	0.050	0.046	0.042	0.041	0.065		0.001	0.004	0.001	0.007	0.032	0.030
IBN_HII	0.041	0.044	0.031	0.047	0.042	0.034	0.032	0.051	0.002		0.000	0.001	0.002	0.029	0.030
IAU_HII	0.038	0.042	0.026	0.044	0.038	0.029	0.028	0.051	0.007	0.000		0.002	0.001	0.029	0.030
IAR_HII	0.044	0.046	0.033	0.048	0.043	0.036	0.035	0.061	0.001	0.000	0.002		0.003	0.031	0.031
IMO_HII	0.039	0.043	0.024	0.045	0.038	0.027	0.026	0.049	0.012	0.004	0.000	0.010		0.028	0.030
ICY_HII	0.077	0.080	0.074	0.086	0.080	0.073	0.068	0.043	0.115	0.102	0.101	0.112	0.100		0.005
IBN_HII	0.088	0.090	0.086	0.097	0.091	0.085	0.080	0.063	0.104	0.103	0.105	0.106	0.108	0.011	

**b**

	IHO_TI	IBN_TI	IAR_TI	IAU_TI	IMR_TI	ICY_TI	IGA_TI	IMO_TI	IAR_TII	IAU_TII	IBN_TII	IHO_TII
IHO_TI		0.009	0.022	0.025	0.018	0.018	0.025	0.019	0.022	0.019	0.021	0.029
IBN_TI	0.033		0.004	0.027	0.006	0.006	0.004	0.008	0.020	0.021	0.015	0.019
IAR_TI	0.024	0.011		0.024	0.015	0.015	0.015	0.015	0.017	0.017	0.012	0.017
IAU_TI	0.100*	0.107*	0.084*		0.033	0.033	0.032	0.032	0.028	0.030	0.026	0.031
IMR_TI	0.066	0.014	0.050	0.152*		0.002	0.002	0.002	0.026	0.027	0.023	0.024
ICY_TI	0.070	0.016	0.051	0.154*	0.004		0.003	0.004	0.026	0.027	0.023	0.024
IGA_TI	0.064	0.014	0.049	0.147*	0.005	0.008		0.005	0.025	0.027	0.023	0.024
IMO_TI	0.061	0.016	0.047	0.146*	0.003	0.007	0.007		0.026	0.028	0.024	0.024
IAR_TII	0.062	0.075	0.056	0.122*	0.112	0.116	0.110	0.115		0.003	0.004	0.006
IAU_TII	0.066	0.075	0.056	0.128*	0.117	0.121	0.113	0.119	0.007		0.005	0.005
IBN_TII	0.044	0.054	0.036	0.104*	0.095	0.097	0.091	0.097	0.001	0.000		0.004
IHO_TII	0.055	0.057	0.046	0.120*	0.089	0.090	0.084	0.089	0.013	0.008	0.002	

$\pi$  values were also obtained for subfamily TII in those species with a large number of monomers examined (9.08 % in *I. aurelioi* and 9.60 % in *I. aranica*).

The factorial analysis of TaqI monomers identified a main axis of variation (axis 1 at Fig. 2b, explaining 48.30 % of the observed variation), corresponding to the separation between three groups of repeats: (1) subfamily TII (i.e., essentially *Pyrenesaura*); (2) a subset of subfamily TI, including all the monomers of Iberian species and a few monomers of *I. bonnali*; and (3) a subset of subfamily TI, made up of monomers from *I. horvathi*, *I. aranica*, and *I. bonnali*. Axis 2 in the FCA, which accounts for 25.70 % of the total variation, separates a fourth group of repeats, comprising the remaining TI monomers of *I. horvathi*. Net genetic distances between repeats from the different species (Table 3b) give additional support to the FCA results. Leaving aside the comparisons involving the single monomer of TI in *I. aurelioi*, larger distances between TI repeats correspond to pairs of the Iberian species with both *I. aranica* (4.70–5.10 %) and, above all,

*I. horvathi* (6.10–7.0 %). As for the TII repeats, all the pairwise comparisons, involving the subgenus *Pyrenesaura* and *I. horvathi*, produce rather low values (0.0–1.30 %).

Organization of consecutive monomeric units

The cloning and sequencing of multimeric products allowed us to characterize the organization of consecutive monomeric repeats. In both satDNA families, and in all the species analyzed, we observed that adjacent monomers in a satellite array usually belong to different sequence groups and even to different subfamilies (for a list of all HindIII and TaqI composite arrays sampled in the *Iberolacerta* species, see Supplementary Tables 2 and 3, respectively).

Phylogenetic analysis

The statistical parsimony network obtained for HindIII satDNA showed a high degree of reticulation among the

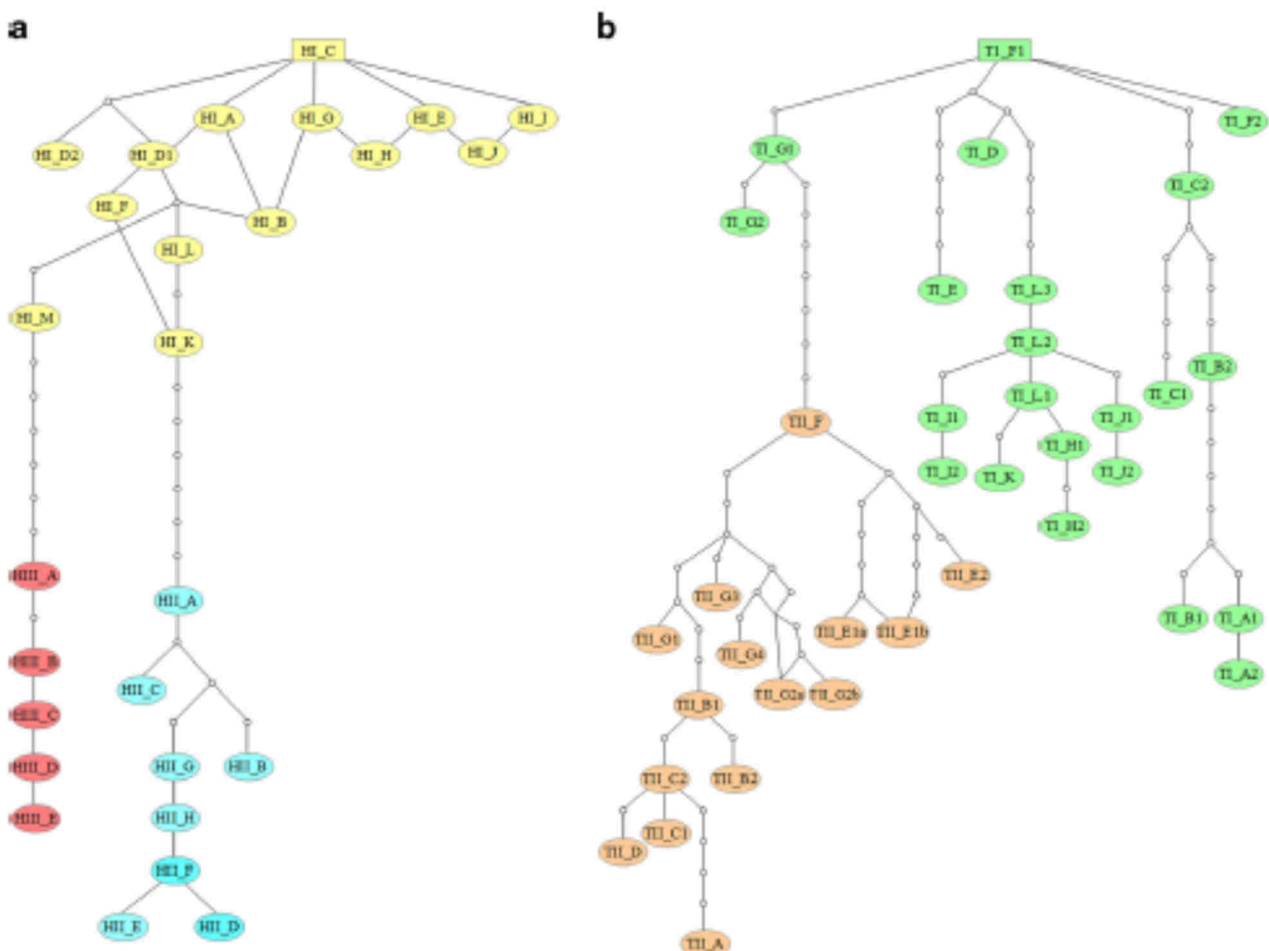
members of subfamily HI (Fig. 3a). This pattern suggests that rearrangements due to recombination events are an important force generating new monomers in this subfamily—the most widespread among *Iberolacerta* species—which occupies the central position of the parsimony network. Two sequence groups within this subfamily, HI\_K and HI\_M, branched into two separate lineages, corresponding to subfamilies HII and HIII, respectively. In contrast to subfamily HI, no evidence for recombination events has been found within subfamilies HII and HIII.

In the network of TaqI satDNA, all sequence groups converge on a group belonging to subfamily T1 (T1\_F1, Fig. 3b). The network shows a major separation of four clusters, connected to group T1\_F1 by a few mutational steps. Three of them (T1\_F2, T1\_C2, and T1\_G1, together with their related variants) include sequences only found in *I. horvathi* and in the subgenus

*Pyrenesaura*. All sequence groups belonging to subfamily TII occupy a peripheral position within cluster G1. The extensive diversification within subfamily TII has been promoted, in some cases, by recombination events that created new monomer variants (e.g., TII\_E1b or TII\_G2a). Within the fourth cluster, the prolific lineage T1\_L3 includes closely related sequence groups (separated by just one or two nucleotide changes), specific to the Iberian clade.

#### Chromosomal location of HindIII and TaqI satDNA families

FISH with HindIII satDNA probe on metaphase chromosomes of *I. monticola* and *I. galani* revealed that this repetitive element is present at centromeres of all the 36 chromosomes of the diploid complement (Fig. 4; Giovannotti et al. 2014). FISH on female metaphases



**Fig. 3** Statistical parsimony network constructed from the consensus sequences of the different sequence groups of a HindIII satDNA and b TaqI satDNA



of *I. bonnali*, carried out in this work, showed hybridization signals in the centromeric regions of all the 23 chromosomes of the karyotype, although with variable signal strength in different chromosome pairs (Fig. 4). Moreover, the overall intensity of HindIII signals in *I. bonnali* was noticeably lower than in *I. monticola* and *I. galani*. No hybridization signals were observed in the chromosomes of *I. horvathi*.

FISH with TaqI satDNA probe in *I. monticola* and *I. galani* produced bright signals in interstitial position in a subset of 20 and 18 chromosomes, respectively (Fig. 5). In *I. bonnali*, similarly intense signals were detected interstitially on both arms of 10 meta-/submetacentric chromosomes. In some metaphases, an additional faint signal could be observed in a medium-sized chromosome pair (Fig. 5). In *I. horvathi*, strong hybridization signals were also observed in interstitial position but just in six chromosomes. However, after increased exposure times, 10 additional chromosomes appeared weakly labeled (Fig. 5).

## Discussion

The turnover rate of a satDNA family is a complex feature that depends on many parameters, such as inter-chromosomal and intrachromosomal recombination rates, copy number and long-range organization of repeat units, genome location and distribution, putative functional interactions, reproductive mode, and population factors (Strachan et al. 1985; Dover 2002; Luchetti et al. 2003; Robles et al. 2004; Meštrović et al. 2006; Kuhn et al. 2008; Navajas-Pérez et al. 2009; Giovannotti et al. 2013). In consequence, sequence dynamics of satDNA families may differ not only among families but also, for a given family, among genomic regions (Kuhn et al. 2011), populations (Wei et al. 2014), species, or higher taxonomic groups (e.g., Macas et al. 2006; Kuhn et al. 2008; Martinsen et al. 2009; Plohl et al. 2010).

In agreement with Giovannotti et al. (2014), the results of the present work show that overall variability

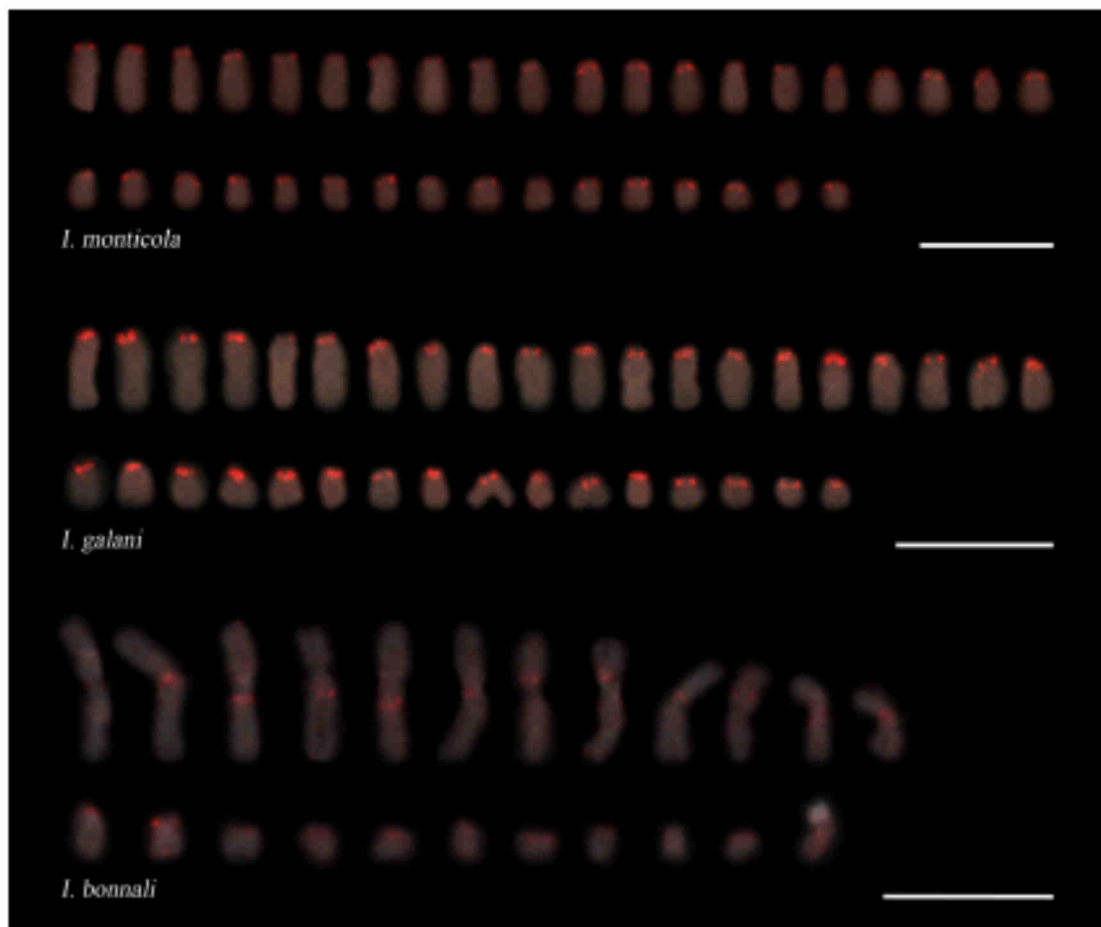
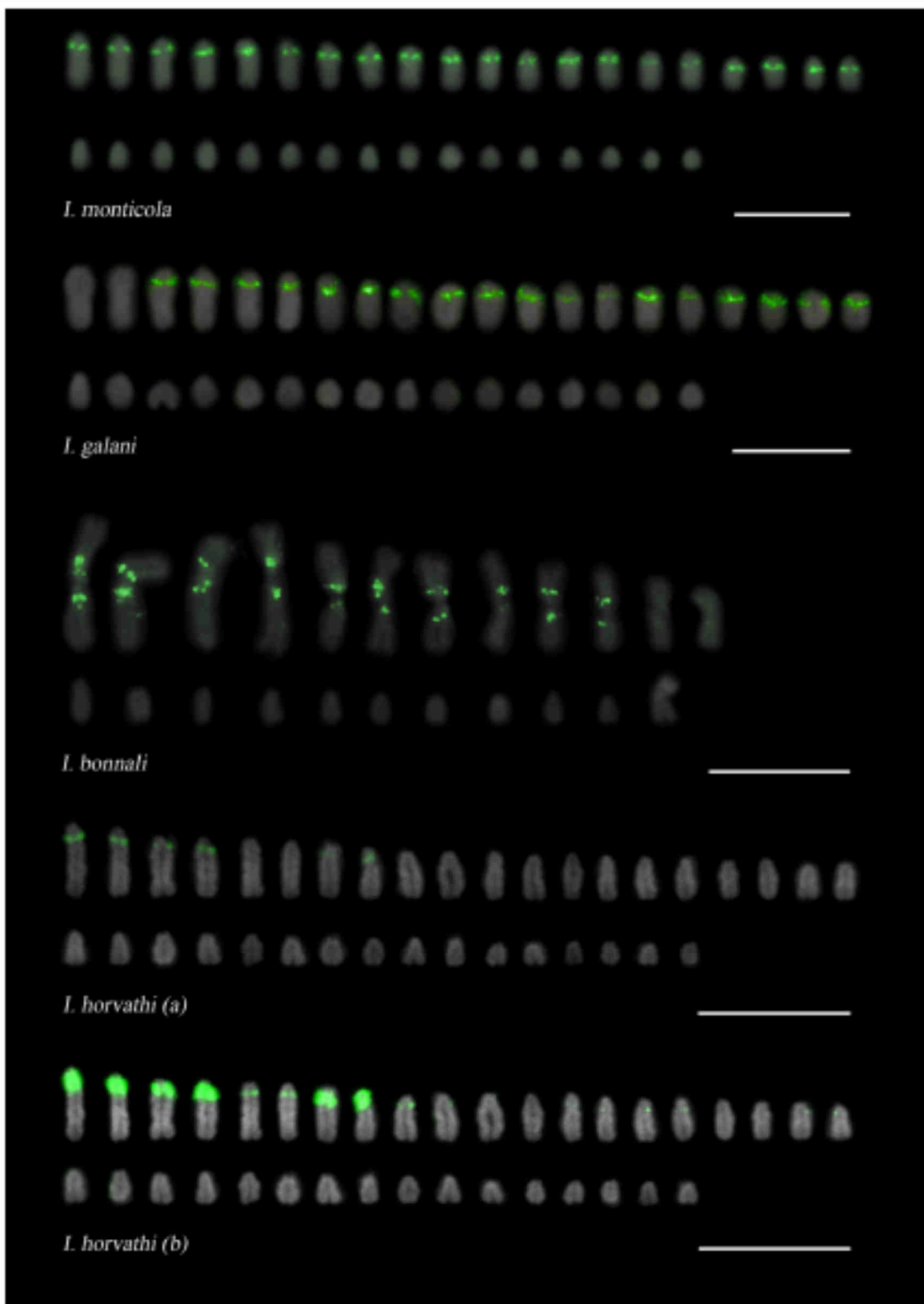


Fig. 4 Hybridization pattern of the HindIII probe in the karyotypes of *Iberolacerta monticola*, *I. galani* and *I. bonnali*. Scale bar=10  $\mu$ m



**Fig. 5** Hybridization pattern of the TaqI probe in the karyotypes of *Iberolacerta monticola*, *I. galani*, *I. bonnali*, and *I. horvathi*. FISH signals on *I. horvathi* chromosomes are shown at standard (a) and increased (b) exposure times. Scale bar=10  $\mu$ m



of TaqI repeats in the whole genus *Iberolacerta* is on average three times higher than the variability of HindIII repeats, which suggests a faster homogenization/fixation rate for the latter satDNA family. However, the detailed characterization of both satDNA families in all eight *Iberolacerta* species reveals that their evolutionary patterns are more complex than previously anticipated. The presence of HindIII\_H1 in all the species, and its central position in the phylogenetic network, suggests that this is the most ancestral variant of HindIII satDNA, from which subfamilies HII and HIII were derived. Interestingly, with the exception of *I. cyreni*, no intraspecific homogenization for any particular subfamily was detected in our study, and most different sequence groups of subfamilies HI and HII are widespread and shared by even distantly related species. Indeed, interspecific genetic distances within each subfamily are substantially lower than intraspecific genetic distances between repeats belonging to different subfamilies. On the contrary, *I. cyreni* shows a high proportion of private sequence groups belonging to subfamily HIII, and a well-differentiated subset of HI repeats, which explains the evidence of concerted evolution found for this species in our previous study. However, the finding of HIII repeats also in *I. bonnali* indicates that this subfamily is not exclusive of *I. cyreni*, but was already present in the common ancestral library of HindIII variants. Combining these data with the results of FISH experiments, the most parsimonious interpretation of HindIII satDNA evolution is that the diversification of HindIII repeats—which generated most of the extant variants—took place in the common ancestor of *Iberolacerta*, before species radiation, i.e., from 11.6 to 15.6 Mya (Arribas et al. 2014). In the ancestral species, HindIII satDNA might have been widely distributed in the centromeres of all chromosome pairs, with a subsequent decrease in copy number in *I. horvathi* and, at least, in the Pyrenean *I. bonnali*. In the latter species, and maybe also in the other two Pyrenean taxa, the reduced amounts of HindIII satDNA might obey to the possible involvement of this centromeric element in the Robertsonian fusions that originated the banded chromosomes characteristic of *Pyrenesaura* from the ancestral acrocentric karyotype, as has been suggested for other centromeric repeats in marsupials (Bulazel et al. 2007). Alternatively, HindIII could represent a minor satDNA family in the centromeres of the ancestral species, which was differentially amplified in the Iberian clade. In either case, the turnover of HindIII

repeats in the different lineages mainly involved the same pool of “old” repeat variants. Long-term conservation of ancestral repeats could be a consequence of selective constraints imposed on functional motifs or structural features of satellite monomers (see, for example, Meštrović et al. 2006; Plohl et al. 2012), involved in any of the roles ascribed to satDNAs (reviewed in Ugarković 2009). Thus, even if we did not find any evidence of function in HindIII satDNA, selection may have favored the maintenance of some repeat variants and/or limited the diversification of this repetitive element. Nevertheless, the loss of HindIII repeats in *I. horvathi* and *I. bonnali* (or, alternatively, the amplification in the Iberian species) suggests that even if functional, a satellite family may be replaced by another in a relatively short evolutionary time.

Actually, and in contrast to the highly conserved function of the centromeres, the rapid evolution and extensive changes in copy number of satDNAs is a general characteristic of centromeric regions (Henikoff et al. 2001). The detection of recombinant sequences within subfamily HI suggests that mechanisms such as unequal crossovers between sister chromatids and gene conversion may have been an important source of new sequence variants in HindIII satDNA (e.g. Smith 1976; Talbert and Henikoff 2010). Moreover, unequal crossover occurring between highly homogeneous arrays can induce copy number alterations of satDNA repeats, such as those observed in the *Iberolacerta* species (Stephan 1986). This fast evolution of centromeric satDNAs can be linked to reproductive isolation and speciation (Bachmann et al. 1989; Bachmann and Sperlich 1993). For example, divergence of centromeric satDNA in *Drosophila* species can inhibit chromosome segregation in hybrids and thus directly cause hybrid incompatibilities and postzygotic isolation (Ferree and Barbash 2009). Likewise, the high copy number polymorphisms and rapid shifts in centromere sequence composition could have contributed and even triggered species radiation within *Iberolacerta*.

The TaqI satDNA family appears to have a very different evolutionary history from the HindIII family, and to evolve much faster in the lineage that leads to *I. horvathi*. According to the parsimony network, TaqI\_T1, the most widespread subfamily among the analyzed species, would also be the most ancestral variant, from which subfamily TII was derived. Moreover, the phylogenetic distribution of the different sequence sets suggests that both subfamilies were



present in the common ancestor of *Iberolacerta*. Subsequently, subfamily TII spread in the Pyrenean species, whereas it was progressively lost in *I. horvathi* and maybe even completely removed from the genomes of the Iberian species. Altogether, TI repeats retrieved from *I. horvathi* show a general pattern of concerted evolution, with high interspecific distance values in all pairwise comparisons and a large subset of species-specific sequence groups. The allocation of these private groups (e.g., TI\_A2 or TI\_C1) in terminal clades of the statistical parsimony network indicates that they probably arose after the early separation of *I. horvathi* from the remaining species, about 11.5 Mya (9.6–13.7) (Arribas et al. 2014). The evolution of TaqI satDNA in *I. horvathi* was probably accompanied by a reduction in the abundance and chromosomal distribution, as inferred from the results of FISH experiments. TaqI satDNA also seems to evolve in concert in the Iberian clade but with a distinct pattern from that found in *I. horvathi*. In this case, the profile of TI repeats and the low levels of nucleotide diversity indicate that concerted evolution in the Iberian clade involved the preferential homogenization of a reduced subset of TaqI variants, all of which evolved from a single sequence lineage, TI\_L3. After cladogenesis, however, the rate at which TI repeats evolved within the Iberian clade is presumably low, since TaqI sequences are poorly differentiated between the four taxa and we found almost no species-specific sequence sets.

In contrast with *I. horvathi* and the Iberian species, the turnover of TaqI satDNA seems to be remarkably slow in the Pyrenean *I. bonnali*. TaqI repeats from this species belong mainly to “old” sequence sets of subfamily TI, and lack species-specific diagnostic positions, which indicates that most of the variability found in *I. bonnali* obeys to synapomorphisms, and that TaqI repeats have been evolving with a low rate of sequence change after speciation. Conversely, the evolution of TaqI satDNA in the other two Pyrenean species, *I. aranica* and *I. aurelioi*, is characterized by the amplification of subfamily TII. Phylogenetic studies suggest that the three species of the Pyrenean clade originated in rapid succession, though *I. bonnali* probably split first, roughly 3.8 Mya (2.7–4.9) (Arribas et al. 2006, 2014). According to this phylogenetic reconstruction, the amplification of subfamily TII in the genomes of *I. aranica* and *I. aurelioi* may have occurred in a short time, after the separation of *I. bonnali* and before the

divergence of both species, ca. 3.3 Mya (2.3–4.3). A rapid expansion of subfamily TII agrees well with the high levels of intraspecific nucleotide diversity and interspecific sequence conservation observed for this subfamily in both species.

The different turnover rates of TaqI repeats among the Pyrenean species, *I. horvathi* and the Iberian species, could be related to differences in their karyotypes. It is possible that interchromosomal exchange and homogenization between the asymmetric meta-/submetacentric chromosomes of the Pyrenean species is more limited than in the species with all acrocentric chromosomes, more homogeneous in shape and size. Similar considerations have been proposed to explain the lower evolutionary rate of satDNAs in sturgeons as compared to sparids (de la Herrán et al. 2001). Limited interchromosomal exchange would lead to a progressive compartmentalization of satellite repeats, followed by a reduction in their interactions and, eventually, by a lack of homogenization of different sequence variants. However, this hypothesis is at least partially contradicted by our analysis of consecutive monomeric units, which revealed that, in both HindIII and TaqI satDNA families, adjacent repeats are not necessarily more similar than are repeats selected at random and that members of different sequence groups or even subfamilies can be interspersed in the same array.

In fact, this pattern of composite repeats may be a key factor explaining the disparate turnover rates of each satDNA family in different species. In eukaryotes, homologous recombination within or between chromosomes can be inhibited by only one mutation per 200 bp (Nijman and Lenstra 2001 and references therein). Likewise, mutations in new monomer variants would inhibit the interactions of repeat units, leading to sequence diversification, divergent evolution, and the formation of satDNA subfamilies. Accordingly, our estimates of intraspecific genetic distances between repeats belonging to different subfamilies suggest that each subfamily within HindIII and TaqI satDNAs is evolving independently. In this context, the intermixing between subfamilies HI and HII within HindIII arrays in most of the species analyzed, and between TaqI subfamilies TI and TII in the Pyrenean taxa, would strongly reduce recombination and homogenization within each subfamily, resulting in the pattern of non-concerted evolution observed



in our study. Conversely, the amplification of subfamily HIII in *I. cyreni*, and the preponderance of subfamily TI in *I. horvathi* and the Iberian species, allows a more efficient homogenization of HindIII and TaqI repeats, respectively, which translates into the overall patterns of concerted evolution observed for these satDNA families in the species mentioned above.

Taken together, our results on the dynamics of HindIII and TaqI satDNAs in *Iberolacerta* are congruent with proposed models of satDNA evolution and life history, intended to explain the considerable fluctuations in copy number and variability of satDNAs shared by related species (Nijman and Lenstra 2001; Plohl et al. 2010). They also support the idea that the “library model” may be extended to monomer variants of the same satDNA family, which were already present in a common ancestor and are currently distributed in related species in variant copy numbers (Cesari et al. 2003). As observed in *Iberolacerta*, this particular evolutionary pattern may result in species-specific profiles of satDNAs which do not reflect the phylogenetic relationships among taxa.

In conclusion, an in-depth analysis of intragenomic variability of HindIII and TaqI satDNAs in *Iberolacerta* revealed two disparate evolutionary histories which, nevertheless, showed some common traits: (i) each satDNA family is made up of a library of monomer variants or subfamilies shared by related species; (ii) species-specific profiles of satellite repeats are shaped by expansions and/or contractions of different variants from the library; (iii) different turnover rates, even among closely related species, result in great differences in overall sequence homogeneity and in concerted or non-concerted evolution patterns. Contrasting turnover rates are possibly related to genomic constraints such as karyotype architecture and the interspersed organization of diverging repeat variants in satellite arrays and maybe also to functional interactions. On the whole, these satDNA families constitute highly dynamic systems, which may have a critical role on the evolution of genome and species. Further studies aimed at investigating the genome-wide variability and organization of reptilian satDNAs may not only be useful to test current hypothesis and identify mechanisms influencing the evolution of this genomic

component but also to improve its application as a molecular marker in phylogenetic studies.

**Acknowledgments** This work was supported by grants REN2003-02931/GLO (Ministerio de Ciencia y Tecnología, Spain), PGIDIT03RFO10301PR and PGIDIT06RFO10301PR (Xunta de Galicia, Spain) awarded to Horacio Naveira, GRC2014/050 awarded to Ana González, and by grant PRIN2009/20093HYH97 (Ministry of Education, University and Research, Italy) awarded to Vincenzo Caputo Barucchi. Verónica Rojo has been supported by a “FPU” fellowship from Ministerio de Educación, Cultura y Deporte (Spain).

**Ethical standards** Permissions for field work and experimental procedures were issued by the competent authorities: Xunta de Galicia (for *I. monticola* and *I. galani*), Junta de Castilla y León (for *I. cyreni* and *I. martinicensis*), Gobierno de Aragón (for *I. bonnali*), and Italian Environment Ministry (for *I. horvathi*). All institutional and national guidelines for the care and use of laboratory animals were followed.

## References

- Arnold EN, Arribas O, Carranza S (2007) Systematics of the palaearctic and oriental lizard tribe Lacertini (Squamata: Lacertidae: Lacertinae), with descriptions of eight new genera. *Zootaxa* 1430:1–86
- Arribas O, Odierna G (2004) Karyological and osteological data supporting the specific status of *Iberolacerta (cyreni) martinicensis* (Arribas, 1996). *Amphibia-Reptilia* 25: 359–367
- Arribas O, Carranza S, Odierna G (2006) Description of a new endemic species of mountain lizard from Northwestern Spain: *Iberolacerta galani* sp. nov. (Squamata: Lacertidae). *Zootaxa* 1240:1–55
- Arribas O, Galán P, Remón N, Naveira H (2014) A new mountain lizard from Montes de León (NW Iberian Peninsula): *Iberolacerta monticola astur* ssp. nov. (Squamata: Lacertidae). *Zootaxa* 3796:201–236
- Bachmann L, Sperlich D (1993) Gradual evolution of a specific satellite DNA family in *Drosophila ambigua*, *D. obscura*, and *D. obscura*. *Mol Biol Evol* 10:647–659
- Bachmann L, Raab M, Sperlich D (1989) Satellite DNA and speciation: a species specific satellite DNA of *Drosophila guanchei*. *J Zool Syst Evol Res* 27:84–93
- Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F (2004) GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations. Retrieved from: [www.genetix.univ-montp2.fr/genetix/genetix.htm](http://www.genetix.univ-montp2.fr/genetix/genetix.htm). 18/3/2015
- Bulazel KV, Ferreri GC, Eldridge MDB, O’Neill RJ (2007) Species-specific shifts in centromere sequence composition are coincident with breakpoint reuse in karyotypically divergent lineages. *Genome Biol* 8:R170
- Capriglione T (2000) Repetitive DNA as a tool to study the phylogeny of cold-blooded vertebrates. In: Olmo E, Redi CA (eds) *Chromosomes today*. Birkhäuser Verlag, Basel, pp 183–194



- Capriglione T, Olmo E, Odierna G, Smith DI, Miller OJ (1989) Genome composition and tandemly repetitive sequence at some centromeres in the lizard *Podarcis s. sicula* Raf. *Genetica* 79:85–91
- Capriglione T, Cardone A, Odierna G, Olmo E (1991) Evolution of a centromeric satellite DNA and phylogeny of lacertid lizards. *Comp Biochem Physiol B* 100:641–645
- Capriglione T, Cardone A, Odierna G, Olmo E (1994) Further data on the occurrence and evolution of satellite DNA families in the lacertid genome. *Chromosom Res* 2:327–330
- Capriglione T, de Santo MG, Odierna G, Olmo E (1998) An alphoid-like satellite DNA sequence is present in the genome of a lacertid lizard. *J Mol Evol* 46:240–244
- Capula M, Lapini L, Capanna E (1989) The karyotype of *Lacerta horvathi* (Reptilia, Sauria, Lacertidae). *Genetica* 79:11–16
- Cesari M, Luchetti A, Passamonti M, Sgali V, 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
- Chaiprasertsri N, Uno Y, Peyachokinagui S et al (2013) Highly species-specific centromeric repetitive DNA sequences in lizards: molecular cytogenetic characterization of a novel family of satellite DNA sequences isolated from the water monitor lizard (*Varanus salvator macromaculatus*, Platynota). *J Hered* 104:798–806
- Charlesworth B, Sniegowski P, Stephan W (1994) The evolutionary dynamics of repetitive DNA in eukaryotes. *Nature* 371:215–220
- Cioabanu D, Grechko VV, Darevsky IS (2003) Molecular evolution of satellite DNA CLsat in lizards from the Genus *Darevskia* (Sauria: Lacertidae): correlation with species diversity. *Russ J Genet* 39:1527–1541
- Cioabanu D, Grechko VV, Darevsky IS, Kramerov DA (2004) New satellite DNA in *Lacerta* s. str. lizards (Sauria: Lacertidae): evolutionary pathways and phylogenetic impact. *J Exp Zool B Mol Dev Evol* 302:505–516
- Clement M, Posada D, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. *Mol Ecol* 9:1657–1659
- de la Herrán R, Fontana F, Lanfredi M et al (2001) Slow rates of evolution and sequence homogenization in an ancient satellite DNA family of sturgeons. *Mol Biol Evol* 18:432–436
- Dover GA (1982) Molecular drive: a cohesive mode of species evolution. *Nature* 299:111–117
- Dover GA (2002) Molecular drive. *Trends Genet* 18:587–589
- Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32:1792–1797
- Enukashvily NI, Ponomartsev NV (2013) Mammalian satellite DNA: a speaking dumb. *Adv Protein Chem Struct Biol* 90:31–65
- Feliciello I, Akrup I, Brajković J, Zlatar I, Ugarković D (2015) Satellite DNA as a driver of population divergence in the red flour beetle *Tribolium castaneum*. *Genome Biol Evol* 7:228–239
- Ferree PM, Barbash DA (2009) Species-specific heterochromatin prevents mitotic chromosome segregation to cause hybrid lethality in *Drosophila*. *PLoS Biol* 7, e1000234
- Fry K, Salser W (1977) Nucleotide sequences of HS-A satellite DNA from kangaroo rat *Dipodomys ordii* and characterization of similar sequences in other rodents. *Cell* 12:1069–1084
- Giovannotti M, Nisi Cerioni P, Caputo V, Olmo E (2009) Characterisation of a GC-rich telomeric satellite DNA in *Eumeces schneideri* Daudin (Reptilia, Scincidae). *Cytogenet Genome Res* 125:272–278
- Giovannotti M, Nisi Cerioni P, Splendiani A, Ruggeri P, Olmo E, Caputo Barucchi V (2013) Slow evolving satellite DNAs: the case of a centromeric satellite in *Chalcides ocellatus* (Forskål, 1775) (Reptilia, Scincidae). *Amphibia-Reptilia* 34:401–411
- Giovannotti M, Rojo V, Nisi Cerioni P et al (2014) Isolation and characterization of two satellite DNAs in some Iberian rock lizards (Squamata, Lacertidae). *J Exp Zool B Mol Dev Evol* 322:13–26
- Grechko VV, Cioabanu DG, Darevsky IS, Kramerov DA (2005) Satellite DNA of lizards of the genus *Lacerta* s. str. (the group *L. agilis*), the family Lacertidae. *Dokl Biochem Biophys* 400:44–47
- Grechko VV, Cioabanu DG, Darevsky IS, Kosushkin SA, Kramerov DA (2006) Molecular evolution of satellite DNA repeats and speciation of lizards of the genus *Darevskia* (Sauria: Lacertidae). *Genome* 49:1297–1307
- Henikoff S, Ahmad K, Malik HS (2001) The centromere paradox: stable inheritance with rapidly evolving DNA. *Science* 293:1098–1102
- Kearse M, Moir R, Wilson A et al (2012) Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28:1647–1649
- Kuhn GCS, Sene FM, Moreira-Filho O, Schwarzacher T, Heslop-Harrison JS (2008) Sequence analysis, chromosomal distribution and long-range organization show that rapid turnover of new and old pBm satellite DNA repeats leads to different patterns of variation in seven species of the *Drosophila buzzatii* cluster. *Chromosom Res* 16:307–324
- Kuhn GCS, Schwarzacher T, Heslop-Harrison JS (2010) The non-regular orbit: three satellite DNAs in *Drosophila martensis* (*buzzatii* complex, *repleta* group) followed three different evolutionary pathways. *Mol Genet Genomics* 284:251–262
- Kuhn GCS, Kütler H, Moreira-Filho O, Heslop-Harrison JS (2011) The 1.688 repetitive DNA of *Drosophila*: concerted evolution at different genomic scales and association with genes. *Mol Biol Evol* msr173
- Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25:1451–1452
- Luchetti A, Cesari M, Carrara G et al (2003) Unisexuality and molecular drive: Bag320 sequence diversity in *Bacillus* taxa (Insecta Phasmatodea). *J Mol Evol* 56:587–589
- Macas J, Navrátilová A, Koblížková A (2006) Sequence homogenization and chromosomal localization of VicTR-B satellites differ between closely related *Vicia* species. *Chromosoma* 115:437–447
- Martinsen L, Venanzetti F, Johnsen A, Sbordoni V, Bachman L (2009) Molecular evolution of the *pDo.500* satellite DNA family in *Dolichopoda* cave crickets (Rhopidophoridae). *BMC Evol Biol* 9:1–14
- Meštrović N, Plohl M, Mravinac B, Ugarković D (1998) Evolution of satellite DNAs from the genus *palorus*—



- experimental evidence for the 'library' hypothesis. *Mol Biol Evol* 15:1062–1068
- Meštrović N, Castagnone-Sereno P, Plohl M (2006) Interplay of selective pressure and stochastic events directs evolution of the MEL172 satellite DNA library in root-knot nematodes. *Mol Biol Evol* 23:2316–2325
- Navajas-Pérez R, Quesada del Bosque ME, Garrido-Ramos MA (2009) Effect of location, organization, and repeat-copy number in satellite-DNA evolution. *Mol Genet Genomics* 282:395–406
- Nijman IJ, Lenstra JA (2001) Mutation and recombination in cattle satellite DNA: a feedback model for the evolution of satellite repeats. *J Mol Evol* 52:361–371
- Odierna G, Aprea G, Amibias O, Capriglione T, Caputo V, Olmo E (1996) The karyology of the Iberian rock lizards. *Herpetologica* 52:542–550
- Palomeque T, Lorite P (2008) Satellite DNA in insects: a review. *Heredity* 100:564–573
- Pezer Ž, Brajković J, Felicciello I, Ugarković D (2012) Satellite DNA-mediated effects on genome regulation. In: Garrido-Ramos MA (ed) *Repetitive DNA. Genome dynamics vol 7*. Karger, Basel, pp 153–169
- Plohl M, Luchetti A, Meštrović N, Mantovani B (2008) Satellite DNAs between selfishness and functionality: structure, genomics and evolution of tandem repeats in centromeric (hetero) chromatin. *Gene* 409:72–82
- Plohl M, Petrović V, Luchetti A et al (2010) Long-term conservation vs high sequence divergence: the case of an extraordinarily old satellite DNA in bivalve mollusks. *Heredity* 104:543–551
- Plohl M, Meštrović N, Mravinac B (2012) Satellite DNA evolution. In: Garrido-Ramos MA (ed) *Repetitive DNA, vol 7, Genome dynamics vol*. Karger, Basel, pp 126–152
- Robles F, de la Herrán R, Ludwig A, Rejón CR, Rejón MR, Garrido-Ramos MA (2004) Evolution of ancient satellite DNAs in sturgeon genomes. *Gene* 338:133–142
- Rojó V, Giovannotti M, Naveira H et al (2014) Karyological characterization of the endemic Iberian rock lizard, *Iberolacerta monticola* (Squamata, Lacertidae): insights into sex chromosome evolution. *Cytogenet Genome Res* 142:28–39
- Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular cloning: a laboratory manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York
- Schwartzacher T, Heslop-Harrison P (2000) *Practical in situ hybridization*. BIOS Scientific Publishers Ltd., Oxford
- Sindaco R, Jermolenko VK (2008) *The reptiles of the western paleartic, volume 1: annotated checklist and distributional atlas of the turtles, crocodiles, amphibians and lizards of Europe, North Africa, Middle East and Central Asia*. Latina: Edizioni Belvedere
- Smith GP (1976) Evolution of repeated DNA sequences by unequal crossover. *Science* 191:528–535
- Stephan W (1986) Recombination and the evolution of satellite DNA. *Genet Res* 47:167–174
- Strachan T, Webb D, Dover G (1985) Transition stages of molecular drive in multiple-copy DNA families in *Drosophila*. *EMBO J* 4:1701–1708
- Talbot PB, Henikoff S (2010) Centromeres convert but don't cross. *PLoS Biol* 8, e1000326
- Tamura K, Nei M, Kumar S (2004) Prospects for inferring very large phylogenies by using the neighbor-joining method. *Proc Natl Acad Sci U S A* 101:11030–11035
- Tamura K, Stecher G, Peterson D, Filipiński A, Kumar S (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol* 30:2725–2729
- Templeton AR, Crandall KA, Sing CF (1992) A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. *Genetics* 132:619–633
- Ugarković D (2009) Centromere-competent DNA: structure and evolution. *Prog Mol Subcell Biol* 48:53–76
- Ugarković D, Plohl M (2002) Variation in satellite DNA profiles—causes and effects. *EMBO J* 21:5955–5959
- Waters PD, Marshall Graves JA, Thompson K, Sankovic N, Ezzaz T (2008) Identification of cryptic sex chromosomes and isolation of X- and Y-borne genes. *Methods Mol Biol* 422:239–251
- Wei KH-C, Greiner JK, Barbash DA, Clark AG (2014) Correlated variation and population differentiation in satellite DNA abundance among lines of *Drosophila melanogaster*. *PNAS* 111:18793–18798

## ***CHAPTER III***

**Characterisation of a satellite DNA involved in the W  
chromosome differentiation in the genus *Lacerta* Linnaeus,  
1758 (Reptilia, Lacertidae).**

Paper in preparation.

## **Characterisation of a satellite DNA involved in the W chromosome differentiation in the genus *Lacerta* Linnaeus, 1758 (Reptilia, Lacertidae).**

### **Abstract**

Satellite DNAs represent a preponderant portion of eukaryotic genomes, and despite the ample literature on satDNAs of eukaryotes little is known about these repetitive elements in reptiles.

Satellite DNAs are tandemly arrayed, highly repetitive DNA sequences occurring in the eukaryotic genomes located in the constitutive heterochromatin (Ugarkovic & Plohl 2002). The repeats comprising a satellite-DNA family do not evolve independently of one another but rather follow concerted evolution (e.g., Plohl et al., 2008). That is, arrays of non-allelic homologous sequences, homogenized by transfer mechanisms such as unequal crossing-over and gene conversion, evolve as a unit (see Plohl et al., 2012). Here, was isolated and characterized the TaqI satDNAs, previously isolated in eight *Iberolacerta* species, from 4 species of the genus *Lacerta* and three of the genus *Timon* with the aim of gain further insights into its evolutionary dynamics, its occurrence among lacertids and to understand if it play any role in sex chromosome evolution in these seven species. The results here obtained highlighted the presence of this genomic element in the genome of all the species investigated, thus indicating that TaqI satDNA is evolutionary conserved among a wide variety of lacertids. In addition, this element was found as very abundant in the heterochromatin of the W-sex chromosome of the four *Lacerta* species investigated. TaqI satDNA occurrence on *Lacerta* heterochromosome suggests that it is involved in the differentiation of the W by heterochromatinization, and the fact that it is absent in the W of other lacertids investigated seems to confirm that repetitive DNA sequences would remain randomly trapped into the sex chromosomes, undergoing amplification as a consequence, not a cause, of the suppression of recombination.

**Key words: satDNA, *Lacerta*, *Timon*, W-specific repeats, FISH**



## Introduction

Squamates are an interesting vertebrate group to study the evolution of sexuality. Indeed, different sex determination mechanisms spanning from environmental cues like temperature (temperature-dependent sex determination, TSD) to genotypic sex determination (GSD) can be found in these reptiles. GSD is mainly based on the differentiation of sex chromosomes, with two reported types (ZZ male/ZW female and XX female /XY male) that differ in the heterogametic sex. Squamates with GSD display remarkable diversity in sex chromosome differentiation, ranging from homomorphic to highly differentiated XX/XY or ZZ/ZW systems (e.g., Organ & Janes, 2008; Ezaz et al., 2009; O'Meally et al., 2012). The first step to the evolution of sex chromosomes would imply the emergence of a locus with female fertility and male sterility and another locus with the opposite effects, leading to the constitution of a small sex-determining region on autosomes. To prevent the production of infertile individuals, selection favors the restriction of those loci to one sex by suppression of recombination, which may subsequently spread along most or all of the chromosome (e.g., Rice, 1996; Charlesworth et al., 2005). In the absence of recombination, retrotransposons invade sex chromosomes and tandem repeats are amplified in the non-recombining region (Charlesworth et al., 1994). Heterochromatin also accumulates, perhaps as a cellular defense against unchecked transposition (Steinemann & Steinemann, 2005).

Isolation of sex-chromosome (W or Y) specific repetitive sequences and their molecular and cytogenetic characterization would provide significant information on the process and mechanism of reptilian heterochromosomes evolution, that was extensively studied only snakes so far (e.g., Jones & Singh, 1985; O'Meally et al., 2010). Pythons, considered basal in snake phylogeny, show homomorphic sex chromosomes, without accumulation of repetitive DNAs. On the contrary, in many advanced snakes like colubrids or elapids (Colubroidea) the heteromorphic W sex chromosome exhibit a strong accumulation of repeats (Jones & Singh, 1985; O'Meally et al., 2010). For example, the W chromosome in the elapid *Notechis scutatus* is composed almost entirely of repetitive sequences, including 18S rDNA and the banded krait minor-satellite (Bkm) repeats (Lee et al., 2007). The Bkm repeats consist of tandem arrays of 26 and 12 copies, respectively, of two tetranucleotides, GATA and GACA (Epplen et al., 1982). Bkm-related repeats have also been isolated from the heteromorphic sex chromosomes of plants and many vertebrates (Jones & Singh, 1981; Parasnis et al., 1999), including birds (O'Meally et al., 2010), suggesting their possible role in the transcriptional activation of sex chromosome heterochromatin (Singh et al., 1976). However, the non-homologous relationship between bird and snake W-chromosomes might also indicate an

independent accumulation during sex chromosome differentiation (see Matsubara et al., 2006; O'Malley et al., 2010).

Another interesting reptilian group in which sex chromosomes differentiation involves heterochromatinization of the heterochromosome is represented by the lizards belonging to the family Lacertidae. This family consists of about 40 genera including 318 species widespread in the Palaearctic region (Uetz, 2014). According to Hipsley et al. (2009) modern lacertids arose shortly after the K/T transition and underwent quite rapid evolutionary diversification splitting into most of its component living genera, almost all possessing 36 acrocentrics plus 2 microchromosomes (Olmo & Signorino, 2005). Despite this substantial uniformity in the karyotypes and a widespread diffusion of a GSD system with ZW heterogamety, the W chromosome exhibits different morphologies, interpreted as different stages of evolution in an almost linear trend (e.g., Olmo et al., 1987). According to Olmo et al. (e.g., 1987, 1990) heterochromatin could have played a paramount role in the differentiation of sex chromosomes, starting with the storage of a specific highly repetitive DNA on either homolog accompanied by heterochromatinization of a homomorphic and heterochromatic W, with a mechanism similar to that of snakes. In fact, Capriglione et al. (1994) speculated that this early stage of W differentiation would have replication and spiralization cycles different from that of Z chromosome, thus hampering the recombination. The next step in this “linear model” would be a progressive deletion of the W, transforming it in a microchromosome (Olmo et al., 1987; Odierna et al., 1993). A deviation from that linear trend would be represented by some populations of *Zootoca vivipara* and the Pyrenean *Iberolacerta*, where original W would be fused with an autosome, giving rise to a  $Z_1Z_2W$  system (Odierna et al., 1996; Odierna et al., 2004). Despite this wide theoretical background, only for two lacertid species data on the composition of heterochromatin in the W chromosome is reported. One case concerns the eremiadine *Eremias velox*, in which Pokorná et al. (2011) observed an enrichment in some microsatellite sequences on either on the whole W chromosome or in its centromeric region, but no accumulation in the centromeres of the Z chromosomes and the autosomes. On the contrary, some microsatellite sequences are present on the Z chromosome and on autosomes, but they are ostensibly lacking on the W chromosome. The second one was recently described for the lacertine *Lacerta agilis*, where the micro-W chromosome is enriched in telomeric TTAGGG repeats (Srikulnath et al., 2014).

In order to better focus on the role played by heterochromatin in sex chromosome evolution in lacertid lizards, we cytogenetically analyzed by conventional and molecular cytogenetic techniques a number of species representative of the genus *Lacerta* Linnaeus, 1758, constituting a monophyletic clade and for which out-group relationships are well established (Harris, 1999;

Godinho et al., 2005; Arnold et al., 2007). In particular, we evaluated the possible involvement of the *TaqI* satellite DNA, widespread in the lacertid genome (Capriglione, 2000; Giovannotti et al., 2014), in the differentiation of W sex chromosome in the studied species.

## Materials and methods

### *Animals, DNA extraction and TaqI satDNA repeats isolation.*

Female and males individuals of *Lacerta bilineata*, *L. agilis*, *L. strigata*, *Timon lepidus*, *T. pater*, *T. tangitanus* and a female of *L. trilineata* were used in this study (Table I). Genomic DNA was extracted from whole blood, using standard protocols with proteinase K digestion followed by phenol/chloroform extraction (see Sambrook et al., 1989).

In order to PCR-amplify *TaqI* satDNA sequences in all the above species, degenerate primers were designed by aligning *Iberolacerta TaqI* satDNA sequences designed by Giovannotti et al. (2014): *TaqI*-F: 5'-AAATTCTGACCSYGSGGGTTAG-3'; *TaqI*-R: 5'-AAAATVGTGCCAACTGTTG-3'. PCR products were electrophoresed in 2% agarose gels, and the band corresponding to the amplified monomers was excised from the gel, purified with Pure Link Quick Gel Extraction Kit (Invitrogen) and cloned in the pCR<sup>®</sup>-blunt vector with Zero Blunt PCR Cloning Kit (Invitrogen) following manufacturer's recommendations. Clones of *TaqI* satDNA were sequenced on an ABI PRISM 3730XL (Applied Biosystems) automatic sequencer. These sequences were then aligned in CLUSTALW (Larkin et al. 2007), using default parameters. A GenBank search was performed in order to compare *TaqI* satDNA with other satDNAs in the database.

Digoxigenin labelled probes were produced by PCR amplification of single clones and used in Southern hybridisation experiments to verify that the elements isolated were tandemly arranged, as expected for satDNAs. In these experiments, *TaqI* digested genomic DNAs from *Lacerta* and *Timon* were used. The hybridisation with the digoxigenin-labelled satDNA probes was performed at 50°C overnight with the Sure Blot CHEMI Hybridisation and Detection Kit (Intergen) following the manufacturer's recommendations. The hybridisation was detected with the same kit.

Neighbor Joining (NJ), Maximum Parsimony (MP), Maximum Likelihood (ML) and Bayesian analyses (BA) were applied to infer the phylogenetic relationships among the sequences of *TaqI* satDNA from the species analysed. MP and NJ analyses were performed as implemented in MEGA version 5 (Tamura et al. 2011). NJ tree was constructed using genetic distances calculated according to Jukes-Cantor (JC) method, with the complete deletion option and 1000 bootstrap

replicates to test statistical support for nodes. Bootstrap values higher than 50% were considered as statistically significant. MP analysis was carried out as implemented in MEGA v. 5. Heuristic search was performed using the Tree-Bisection-Reconnection (TBR) method with the random addition of ten initial trees. One hundred trees were kept at each cycle of the addition procedure. The statistical support for nodes was tested with 100 bootstrap replicates, with bootstrap values higher than 50% regarded as significant. For ML and BA analyses, the best fit model of nucleotide substitution for TaqI satDNA repeats was selected among the 88 models available in jModeltest 2.1.3 (Darriba et al., 2012) using the Akaike Information Criterion (AIC). The most appropriate model was GTR + G. ML analysis was carried out using PhyML 3.0 (Guindon et al., 2010) using the NNI method, with the model parameters fitted to the data by likelihood maximization. Statistical reliability of the ML trees was assessed by 1000 bootstrap replicates (Felsenstein, 1985). Bootstrap values higher than 50% were regarded as statistically significant.

BA analysis was carried out with MrBayes v3.2 (Ronquist et al., 2012) using appropriate model of nucleotide substitution (GTR + G) selected with jModeltest 2.1.3. The BA analysis was run with four incrementally heated Markov chains for  $15 \times 10^6$  generations in two independent runs with samplings at intervals of 500 generations that produced 30000 trees. Once the stationarity had been reached, both in terms of likelihood scores and parameter estimation, the first  $7.5 \times 10^3$  trees (25% 'burn-in') were discarded in both runs and a majority-rule consensus tree was generated from the  $22.5 \times 10^3$  remaining (post burnin) trees. The posterior probability (pp) was calculated as the percentage of samples recovering any particular clade (Huelsenbeck and Ronquist, 2001), with  $pp \geq 0.95$  indicating a statistically significant support (Wilcox et al., 2002). Both ML and BI trees were displayed with FigTree v1.4.0 (<http://tree.bio.ed.ac.uk/software/figtree/>).

Average AT content of the monomeric units of TaqI satDNA was determined with MEGA version 5 (Tamura et al., 2011). Intraspecific nucleotide diversity ( $\pi$ ), haplotype numbers and haplotype diversity (h) were estimated for each species using DnaSP v. 5 (Librado and Rozas 2009).

The overall consensus sequence, consensus sequence of Timon, Lacerta autosomic repeats and Lacerta W-specific repeats were determined with the program EMBOSS that is available online at <http://emboss.bioinformatics.nl/>.

Net average genetic distances between groups were calculated under the appropriate substitution model with MEGA v. 5. Rate of TaqI satDNA evolution was determined according to the divergence times estimated for the four *Lacerta* species here investigated by Godinho et al. (2005).

The occurrence of genetic differentiation between monomeric repeats in the seven species analysed was also assessed with the analysis of molecular variance (AMOVA) (Excoffier et al. 1992) calculating  $\Phi$ -statistics. The test was carried out on the sequences from *Timon* (*T. lepidus*, *T. pater* et *T. tangitanus*) and *Lacerta* (*L. agilis*, *L. bilineata*, *L. stigata* et *L. trilineata*). In *Timon*, TaqI satDNA monomeric repeats were divided into two groups according to the sex of the individual from which they had been isolated. In *Lacerta*, four different groupings were made: i) one made on the basis of the sex of the individual from which monomeric repeats sequence had been isolated; ii) one according to the clades revored by the phylogenetic analyses [putative W-specific repeats (W) and putative autosome-specific repeats (A)]; iii) a third grouping was made by considering only the monomeric repeat sequences from females divided the W-specific repeats from the autosome-specific repeats (based on data from the phylogenetic analyses); iv) the fourth grouping considered a W-specific group of sequences and a group containing only sequences isolated from males. This test was performed at two hierarchical levels to test how satDNA sequence variability was distributed within a group of monomeric repeats and among groups of such repeats. The test was based on pair wise genetic distances between clones and performed as implemented in ARLEQUIN 2.000 (Schneider et al. 2000), using 1000 permutations.

The repeats of the analyzed species were compared using satDNA Analyzer version 1.2 (Navajas-Pérez et al. 2007). This program allows the discrimination between shared and non-shared polymorphic sites. The program identifies polymorphic sites shared between two species when the same polymorphism is found in both species. When this occurs, we assume that these are ancestral sites that appeared before the split between the two species (Navajas-Pérez et al. 2005). By contrast, nonshared polymorphic sites are autapomorphies, representing different transitional stages in the process of intra-specific sequence homogenization and inter-specific divergence. Under the assumption that concerted evolution is a time dependent process, the expected stages of transition during the spread of a variant repeat unit toward its fixation can be defined according to the model of Strachan et al. (1985). This is a method of partitioning the variation by analysing the patterns of variation at each nucleotide site considered independently among all the repeats of a repetitive family when comparing a pair of species (Strachan et al. 1985; Navajas-Pérez et al. 2007). This method examines the distribution of nucleotide sites among six stages (Classes I–VI) in the spread of variant repeats through the family and the species. Briefly, the Class I site represents complete homogeneity across all repeat units sampled from a pair of species, whereas Classes II, III, and IV represent intermediate stages in which one of the species shows a polymorphism. The frequency of the new nucleotide variant at the site considered is low in Class II and intermediate in Class III, while Class IV represents sites in which a mutation has replaced the progenitor base in most



members of the repetitive family in the other species. Class V represents diagnostic sites in which a new variant is fully homogenized and fixed in all the members of one of the species while the other species retains the progenitor nucleotide. A Class VI site represents an additional step over the stage of Class V (new variants appear in some of the members of the repetitive family at a site fully divergent between the two species). The statistical significance (P-value) of the variation in the relative proportions of Strachan transitions stages among different interspecific comparisons was evaluated using chi-square heterogeneity tests that were performed in the interactive online calculator available at <http://www.quantpsy.org/chisq/chisq.htm> (Preacher, 2001).

### Chromosome analysis

Metaphase chromosomes were prepared from females and males of *Lacerta bilineata* (LBI), *L. agilis* (LAG), *L. strigata* (LST), *Timon lepidus* (TLE), and a female of *L. trilineata* (LTR). Metaphases were prepared following the protocol reported by Rojo et al., 2014.

Fluorescence *in situ* hybridisation (FISH) experiments were performed on metaphase preparations using i) a telomeric probe (TTAGGG)<sub>n</sub> produced by PCR according to Ijdo et al. (1991), and ii) the probes obtained by PCR amplification of *TaqI* satDNA clones from each of the studied species. The probes were labelled by PCR either with biotin-16-dUTP (Roche) or digoxigenin-11-dUTP (Roche). Slide pretreatment, denaturation, hybridisation, post-hybridisation washes and detection were performed according to Schwarzacher and Heslop-Harrison (2000). The telomeric probes were evidenced with Fluorescein Iso-Thiocyanate (FITC) and Tetramethyl Rhodamine Iso-Thiocyanate (TRITC), respectively. Chromosomes were observed with a Leica Leitz DMRBE epifluorescence microscope and the images were captured and processed with a Leica CytoVision version 7.2 system.

In order to define the relationships between satDNAs and the constitutive heterochromatin, C-banding was performed on metaphase chromosomes following Sumner (1972). C-banded metaphases were mounted and stained with Vectashield mounting medium with 4',6-diamidino-2-phenylindole (DAPI; Vector Laboratories).

## Results

### *Isolation and analysis of satDNA sequences.*

Southern-blot hybridization of genomic DNAs digested with *TaqI* restriction enzyme revealed a typical ladder-like pattern consisting of multimeric units ranging from 170 to 190 bp.

These results suggest that the satDNA studied is tandemly arrayed in both genera here investigated. No hybridization differences were found between males and females, but the hybridisation signal was stronger in *Lacerta* than in *Timon* (Figure 1).

A total of 162 clones containing sequences of TaqI satDNA monomeric unit were sequenced for the three species of *Timon* and the four species of *Lacerta*. One hundred and thirteen haplotypes were detected (Table 1). This satDNA showed an average AT content of 59.2%, indicating an enrichment in AT, as typical for these genomic elements, with short A and T stretches ranging from 3 to 7 base pairs (Table 2).

Intraspecific  $\pi$  values (JC method), ranged from  $0.1526 \pm 0.0075$  to  $0.0810 \pm 0.0123$ . When, in *Lacerta*,  $\pi$  values were calculated separately for sequences of the two sub-clades W-specific and autosome-specific, (see below), the W-specific repeats resulted more heterogeneous than the autosomal ones (Table 1).

The phylogenetic tree obtained from the Bayesian analysis of TaqI satDNA is shown in Figure 2. The four different phylogenetic analyses applied (NJ, MP, ML, and BA) yielded very similar topologies, with some minor incongruences. Three major clades were recovered with maximum support, one harbouring *Timon* clones and the other two harbouring the sequences of the four *Lacerta* species investigated. One clade contained only sequences from females that were considered as W-specific TaqI satDNA repeats, and the other clade contained repeats from both males and females and that were therefore considered as repeats located in the autosomal arrays of TaqI satDNA.

AMOVA analysis was performed on TaqI sequences of *Lacerta* and *Timon*. The analysis produced results quite different in the two species. In *Timon* most of the molecular variation was distributed within groups of sequences (88.24%;  $\Phi_{ST}$  0.11756,  $P < 0.0001$ ) whereas the percentage of variation among groups of sequences (males vs females) was much lower, representing only 11.76% of the total variation ( $\Phi_{ST}$  0.11756,  $P < 0.0001$ ). The variance among groups of sequences became higher when the AMOVA was carried on *Lacerta* considering two groups of sequences. One group was always represented by the putative W-specific TaqI satDNA repeats, the other group always contained autosome-specific repeats (see Materials & Methods section for details). In the three tests on *Lacerta* the percentage of molecular variation among groups of sequences ranged from 25.72 to 49.54%, with  $\Phi_{ST}$  values always significant ( $P < 0.0001$ ), thus confirming the differentiation between autosomal and W-specific repeats (Table 3).

Strachan analysis of variable sites confirms what already highlighted by phylogenetic and AMOVA analyses. Indeed, diagnostic the occurrence of diagnostic sites (stage VI according to

Strachan et al.,1985) are detected when one of the two groups of sequences compared is represented by W-specific repeats (Table 4 and Figure 3).

### *Chromosome analysis*

The analysis of chromosomal location of TaqI satDNA by FISH with species-specific probes confirmed the results of Southern hybridization that produce a weaker signal on *Timon* as compared to *Lacerta*, indicating a smaller amount of TaqI repeats in the former genus. Indeed, in *Timon* the species-specific probe produced a fluorescent signal on nine autosomes in both males and females. In *Lacerta* the signal was detected on a higher number of autosomes (16 in *L. bilineata*, 20 in *L. agilis*, 24 in *L. strigata* and 16 in *L. trilineata*). In addition, the probe produced a bright signal also on the W sex-chromosomes of all the *Lacerta* species analysed here, whereas no signal was produced by TaqI satDNA probe onto *Timon* W chromosome (Figure 4 and 5). In addition, FISH experiments confirmed the occurrence in *Lacerta* species of two clearly distinguished groups of repeats, already highlighted by phylogenetic, AMOVA and Strachan sites analyses. In fact, when W-specific probe was hybridized onto female metaphases a very bright signal on the heterochromosome and a dimmer signal on autosomes (Figure 5) were produced. FISH with autosomes-specific probe onto female metaphases produced a dim signal on the W and a bright one on the autosomes. Finally, FISH experiments with a telomeric probe marked the telomeres of all the chromosomes of the complement with a certain degree of enrichment in telomeric repeats of the W chromosome in both *Timon* and *Lacerta* here investigated (Figure 6).

C-banded metaphases shown heterochromatin occurs in both pericentromeric and telomeric position. The pericentromeric heterochromatin is on a number of autosomes comparable to that of autosomes labelled by TaqI satDNA probe. In all species the W chromosome is extensively heterochromatic (Figure 7).

## **Discussion**

Satellite DNAs are tandemly arrayed, highly repetitive DNA sequences of the eukaryotic genomes located in the constitutive heterochromatin (Ugarkovic & Plohl 2002). The repeats belonging to a satellite-DNA family do not evolve independently of one another but rather follow concerted evolution (e.g., Plohl et al., 2008). That is, arrays of non-allelic homologous sequences, homogenized by transfer mechanisms such as unequal crossing-over and gene conversion, evolve as a unit (see Plohl et al., 2014). Factors affecting concerted evolution include rates of transfer between homologous and non-homologous chromosomes, arrangements of repeats, array sizes, and population structure. Bias in any of these factors can alter the rates of concerted evolution. Thus, for

instance, these rates will be reduced by mechanisms impeding chromosomal exchanges (i.e., recombination). Therefore, reduced rates of concerted evolution would be expected in satellite-DNA sequences in the non-recombining W or Y chromosomes, (e.g. Navajas-Perez et al., 2009) or depending on the location of satDNA clusters on the chromosomes (e.g. terminal position vs interstitial position; see Giovannotti et al., 2014).

A first remarkable result of this research is the finding that TaqI satDNA is a genomic element widely represented in lacertid lizards genome. In fact, in the paper by Giovannotti et al. (2014) this genomic element was isolated and characterized in the genome of 4 species of Iberian rock lizards (*Iberolacerta cyrenii*, *I. galani*, *I. martinezricai* and *I. monitcola*), and in the course of the present study it was successfully isolated from 4 species of the genus *Lacerta* and three of the genus *Timon*. These results highlight a strong conservation for this satellite that could be attributed to its slow evolutionary rate as determined for this element in the four Iberian rock lizards by Giovannotti et al. (2014). These authors highlighted an evolutionary rate that has a value half of the rate usually reported for these genomic elements (Bachmann and Sperlich 1993; De la Herrán et al. 2001). The slow evolutionary rate and poor phylogenetic resolution power of these repeats could be related to its genomic distribution and chromosomal location. In fact, TaqI repeats are in interstitial/pericentromeric position (less prone to physical association) on a subset of chromosomes (from 9 in *Timon* to 25 in *Lacerta strigata*), and it is possible that the exchange between non-homologous chromosomes bearing TaqI sequences is limited in these conditions. This could reduce interchromosomal exchange and homogenization, thus determining a lower rate of interspecific divergence and a higher degree of intraspecific repeat heterogeneity (as was shown from the comparison between HindIII and TaqI satDNAs in four *Iberolacerta* species by Giovannotti et al., 2014). These data on the phylogenetic distribution among lacertids indicate that the evolutionary history of this satellite DNA is at least 20 Million years old, as inferred from the dating of Lacertinae radiation (Hipsley et al., 2009). This slow evolutionary rate is also responsible for the low phylogenetic resolution power shown by this satellite that can resolve genera but not species (see Giovannotti et al., 2014 and the present study). The slow evolutionary rate of TaqI satDNA could be explained from a point of view different from that considering the chromosomal position of the array of monomers of this genomic element. In fact, if we look at the alignment of the consensus sequences of this satellite in the species here investigated it is possible to observe that most of the monomer sequence is conserved and that variable sites are confined to three regions of the monomeric unit. This observation could imply that the sequence is conserved for its majority by selective pressures. Similar considerations were made for a 70 Million years old centromeric satellite DNA that was described in two anurans families, Leptodactylidae and Hylodidae, and that

showed a conserved dominion that the authors hypothesized as the result of selective pressures, probably associated with centromeric chromatin structure (Vittorazzi et al., 2014). This hypothesis open the way to a new scenario in which the conservatism of this satellite would be the results of the constraints imposed by a function of this genomic element rather than the position along the chromosome. However, the satellite here studied is in pericentromeric/interstitial position and therefore a role in centromeric chromatin compaction is not likely for TaqI element.

The second interesting result of this investigation is represented by the fact that in the four representative of the genus *Lacerta* here investigated (*L. agilis*, *L. bilineata*, *L. strigata*, *L. trilineata*) the fluorescent probe of TaqI satDNA produced a very bright signal also on the W sex chromosome. Indeed, the fluorescent signal is absent on the W of the other species of lacertids so far tested by FISH with TaqI satDNA probe. In particular, this satellite occurs only on autosomes in the three species of *Timon* studied here (*T. lepidus*, *T. pater*, *T. tangitanus*) and in *Iberolacerta* (Giovannotti et al., 2014; Rojo et al., 2015). Therefore it seems that in these *Lacerta* species, TaqI satDNA is involved in the differentiation of the W sex chromosome by heterochromatinization.

Our study confirmed the enrichment in telomeric sequences of the W chromosome of *L. agilis*, as already demonstrated by Srikulnath et al. (2014), and the other three *Lacerta* species here investigated. Interestingly, we found an enrichment in telomeric sequences also on the W of the *Timon* species investigated. The occurrence of (TTAGGG)<sub>n</sub> on the W of *Lacerta* and *Timon* can be a confirmation of their close phylogenetic relationship (see Kapli et al. 2011), whereas the absence of TaqI satDNA repeats on the W of *Timon* and its occurrence on the W of *Lacerta* may indicate that this genomic element started to play a role in the heterochromatinization of the heterochromosome of this genus after the split from *Timon* lineage.

The absence of TaqI satDNA on the W of other species investigated [*Timon* (present study), and *Iberolacerta* (Giovannotti et al. 2014; Rojo et al., 2015)] is consistent with what already found in the heterochromosome of another lacertid (*Eremias velox*) (Pokorná et al., 2011), the heterochromatin of which is filled with microsatellite repeats. From these observations, it seems that is the contingency and not the characteristics of a particular sequence that determine which repetitive elements will accumulate on the sex chromosome of a certain lineage (see Pokorná et al., 2011 for similar considerations). Interestingly, *Lacerta* species are the only species among those so far investigated the sequences of which are clearly separated into two statistically supported clades in the phylogenetic analysis. This condition reflects the isolation of the repeats accumulated on the non-recombining W that would not undergo concerted evolution, consistently with the higher intraspecific heterogeneity recorded for W-specific repeats ( $\pi = 0,1195$ ) when compared to

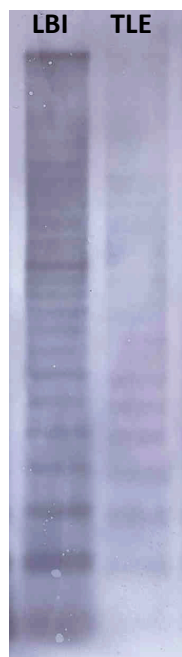


autosomal repeats ( $\pi= 0,0834$ ) (see Navajas-Perez et al., 2009 for similar considerations). This is confirmed by AMOVA analysis that gives a clear indication of the differentiation that is accumulating between autosomal and W-specific TaqI sequences in *Lacerta*. In three *Timon* species, where the sequences of this satDNA occur only on autosomes, the TaqI satDNA repeats form a single statistically supported cluster. The same indication comes from the analysis of Strachan transitional stages that show in *Lacerta* species a pattern similar to that found between genera due to the separation of TaqI satDNA in two distinct groups.

Finally, the enrichment in repeats represented by TaqI satDNA in the four *Lacerta* species investigated strongly support the hypothesis that the sex chromosomes are homologous within this group of lacertids. Indeed, it would be highly improbable that the same family of repeats accumulated on the W chromosomes of different species only as a result of a random process. On the contrary, it is very likely that these repeats were trapped onto the heterochromosome of the common ancestor of modern *Lacerta* species. The sharing of repetitive elements on the W among *Lacerta* species also support the monophyletism of these lizards. In addition, the telomeric repeats enrichment on the W chromosomes of both *Lacerta* and *Timon* seems to confirm the close phylogenetic relationship of these two genera already proposed by Kapli et al. (2011) based on nuclear and mitochondrial molecular markers.

In conclusion, tandemly repetitive DNA sequences would remain randomly trapped into the sex chromosomes, undergoing amplification as a consequence, not a cause, of the suppression of recombination. In addition, comparing female-specific (W linked) satDNA clones with that shared by both sexes (autosomic), we found that TaqI repeats located in the non-recombining W chromosome have lower rates of sequences evolution than the repeats of the autosomic chromosomes. These data imply that patterns of satDNA evolution depend on the chromosomal location of the repeats, such as the absence of recombination between sex chromosomes, that seriously influences the rate of satDNA sequence change.

## Figures and Tables



**Figure 1.** Hybridisation of TaqI satDNA doxigenin-labeled probe onto genomic DNAs of *Lacerta bilineata* (LBI) and *Timon lepidus* (TLE) species digested with TaqI restriction enzyme.

Species	Number of clones			% AT	Repeat Length	Nucleotide diversity ( $\pi$ )	Number of Haplotypes (H)	Haplotype diversity (Hd)
	Female	Male	Total					
<i>L. agilis</i>	12	10	22	58,7	186-189	0,1152±0,0056	15	0,944
<i>L. bilineata</i>	28	15	43	59,0	187-190	0,1432±0,0075	35	0,99
<i>L. trilineata</i>	26	-	26	59,2	186-190	0,1096±0,0048	18	0,951
<i>L. strigata</i>	19	11	30	60,2	173-187	0,1526±0,0075	18	0,956
<i>Lacerta A</i>	36	26	62	58,8	166-190	0,0834±0,0031	47	0,989
<i>Lacerta W</i>	59	-	59	59,8	173-189	0,1195±0,0047	38	0,964
<i>T. lepidus</i>	10	10	20	59,3	182-189	0,1124±0,0047	19	0,995
<i>T. pater</i>	2	8	10	59,3	182-188	0,0810±0,0123	3	0,733
<i>T. tangitanus</i>	4	7	11	58,8	187-189	0,1086±0,0124	5	0,818

**Table 1.** Summary of repeat features and  $\pi$  values. Number of monomeric repeats sequenced (n), average AT content of repeats (AT), length of repeats (expressed in base pairs), and nucleotide diversity ( $\pi$ )±SE for each species investigated, number of haplotypes (H) and haplotype diversity (Hd). A: autosome-specific repeats; W: W-specific repeats.

```

all      NCGAGGCCTGATTTCC TTTCC TTTGATNAAAAA CCCTTCTGTTT CACCGCCAAA TCTTC 60
TIM      -C.....T.....A..-TT.....C.C...C..... 58
W        NT.....C.....G..-AT.....T.A...G..... 59
A        -C.....C.....T..ATG.....T.A...C..... 59

all      CAGGGGAGTTNTGGCAACAGTTTGGCACCATTTTTGANGNAAATTGGAGAACGTCAGATT 120
TIM      .AGGG..G.TC.TG...C.G.....AG-TG..... 117
W        .AGGG..A.TG.GC...C.G.....NAGAA..... 119
A        .TTCA..G.CT.GG...A.T.....AG-GA..... 118

all      TTTGGTGAAATTCTGACCGCGNGGGGTTAGGGA TTTTTTCAAAAAA NNNTTTT CNGCA 180
TIM      ...A.....C.G..-.....A....ANNN...TCCG.. 176
W        ...G.....G..C..-.....N....-TNN...TCNG.. 177
A        ...G.....G..C..N.....A....AGTT...CNNN.. 178

all      GGTNAAGTTGTNGN 194
TIM      G..T..G...C.- 189
W        A..A..G...N.- 190
A        G..C..C....N.N 192

```

**Table 2.** Consensus sequences of TaqI satDNA repeats (all) obtained by aligning the consensus sequences of *Timon* (TIM), W-specific (W) and autosome-specific (A) repeats of *Lacerta* .

Source of variation	Groups	Variance components	Percentage of variation	$F_{ST}$
Among groups of sequences	TIM f vs m	2,93421	11,76	0,11756***
	LAC f vs m	5,44448	25,72	0,25724***
	LAC W vs A	10,76439	47,00	0,47005***
	LAC W vs Af	10,18352	44,74	0,44739***
	LAC W vs m	12,10009	49,54	0,49541***
Within group of sequences	TIM f vs m	22,02493	88,24	0,11756***
	LAC f vs m	15,72024	74,28	0,25724***
	LAC W vs A	12,13616	53,00	0,47005***
	LAC W vs Af	12,57831	55,26	0,44739***
	LAC W vs m	12,32418	50,46	0,49541***

**Table 3.** AMOVA analysis.

The test Was carried on TaqI satDNA sequences from *Timon* and *Lacerta*. Grouping details are given in the text. \*\*\* =  $P < 0.001$ . f: monomeric repeats isolated from females; m: monomeric repeats isolated from males; W: W-specific repeats; A: autosome-specific repeats isolated from both males and females; Af: autosome-specific repeats isolated from females. LAC: *Lacerta* (*L. agilis*, *L. bilineata*, *L. strigata*, *L. trilineata*); TIM: *Timon* (*T. Lepidus*, *T. pater*, *T. tangitanus*).



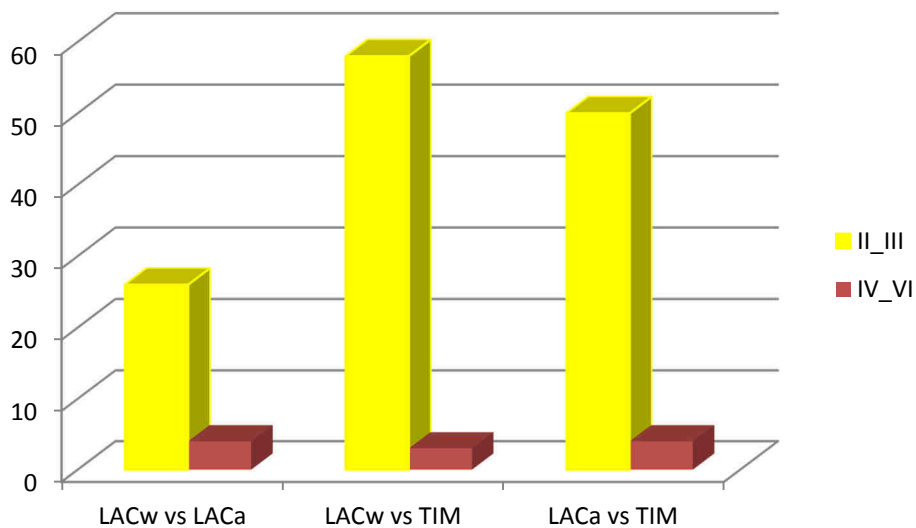
**Lacerta A**



**Figure 2.** Neighbor-Joining tree depicting the phylogenetic relationships among TaqI satDNA repeats isolated from three species of *Timon* and four species of *Lacerta*. At nodes bootstrap (NJ, MP and ML) and posterior probability (BA) values.

CONFRONTI	II-III	IV-VI
LAC <sub>w</sub> vs LAC <sub>a</sub>	26	5
LAC <sub>w</sub> vs TIM	58	3
LAC <sub>a</sub> vs TIM	50	4

**Table 4.** Variable sites classified according to Strachan et al. (1985). LAC<sub>W</sub>: W-specific repeats; from *Lacerta*; LAC<sub>A</sub>: autosome-specific repeats from *Lacerta*; TIM: repeats from *Timon*.



**Figure 3.** Histogram showing the distribution of polymorphic sites classified according to Strachan et al. (1985). Abbreviations as in Table 6.

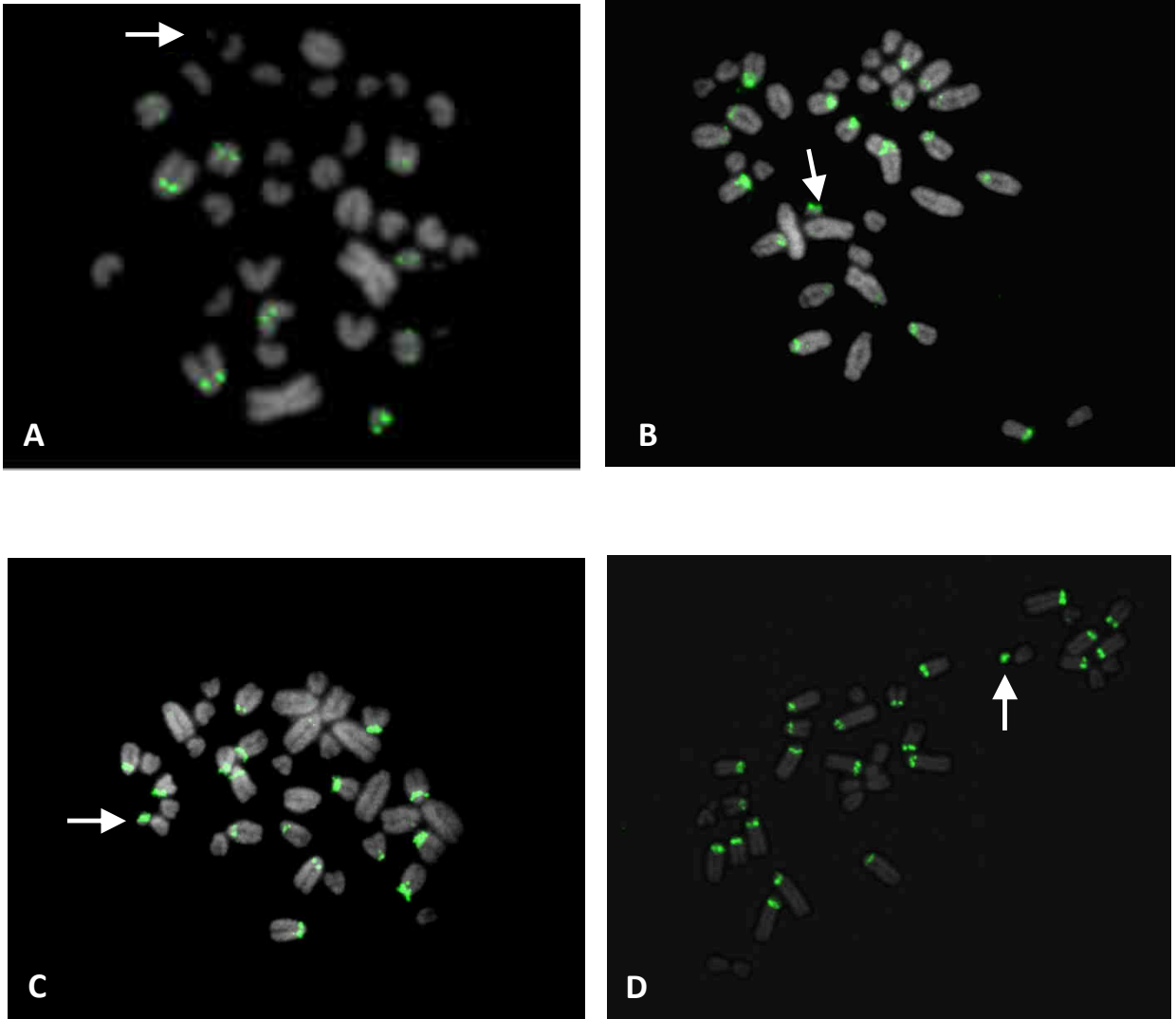


Figure 4. FISH with TaqI probe onto metaphases from females of: *T. lepidus* (A), *L. agilis* (B), *L. trilineata* (C), *L. strigata* (D). For each species was used a species-specific probe. The arrow indicates the W chromosome.

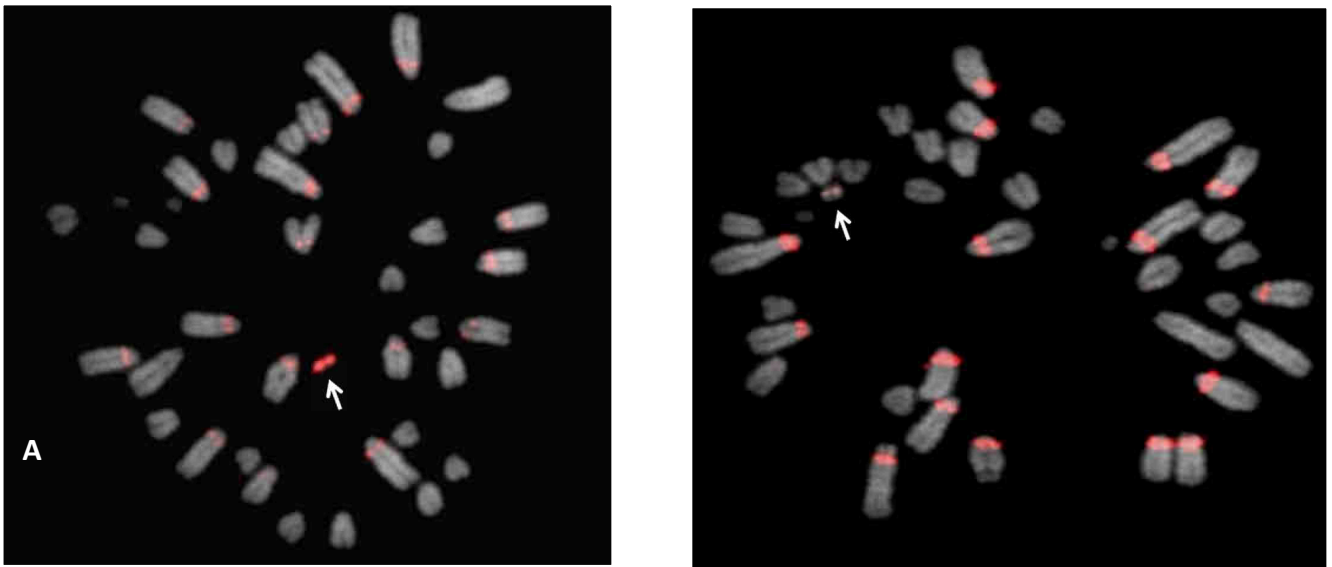
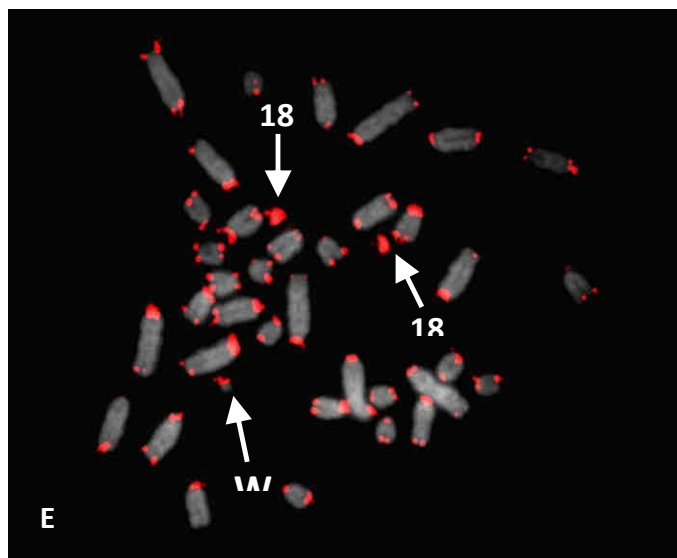
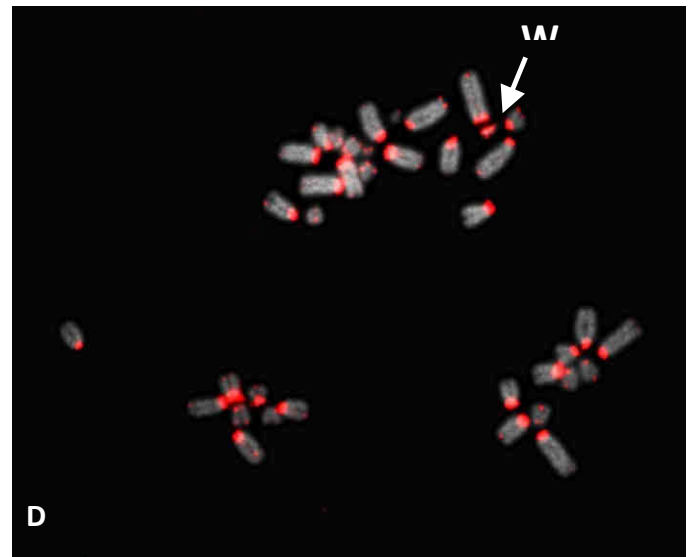
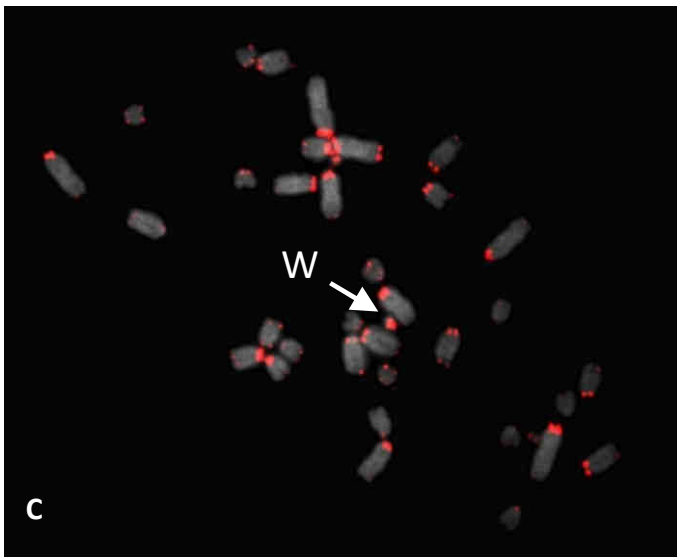
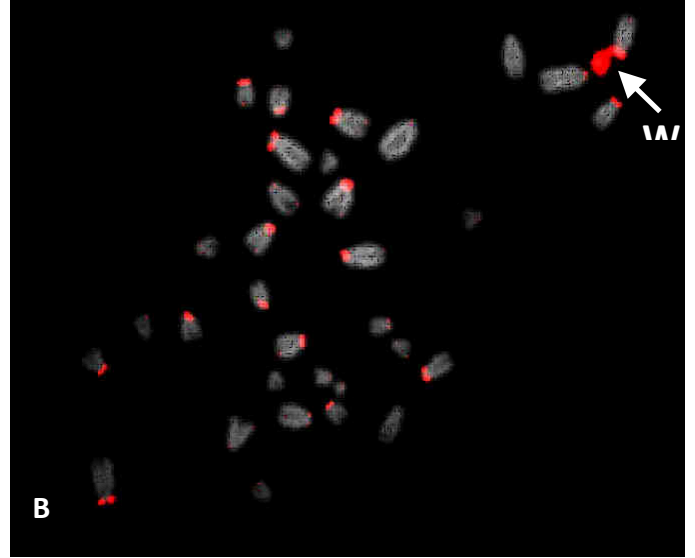
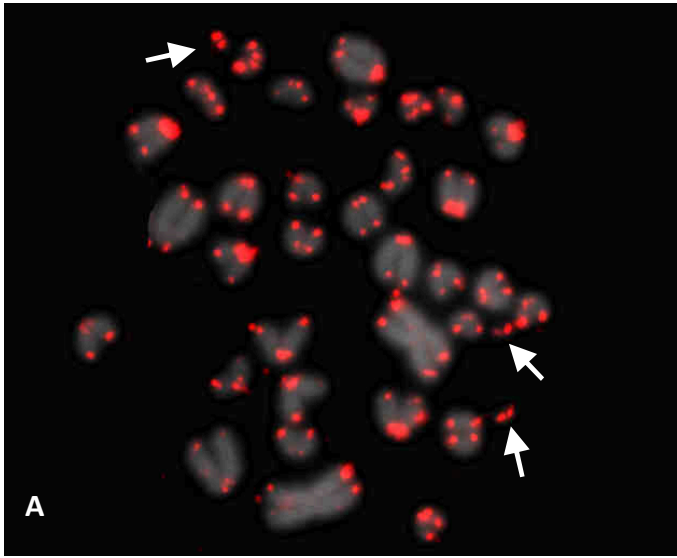
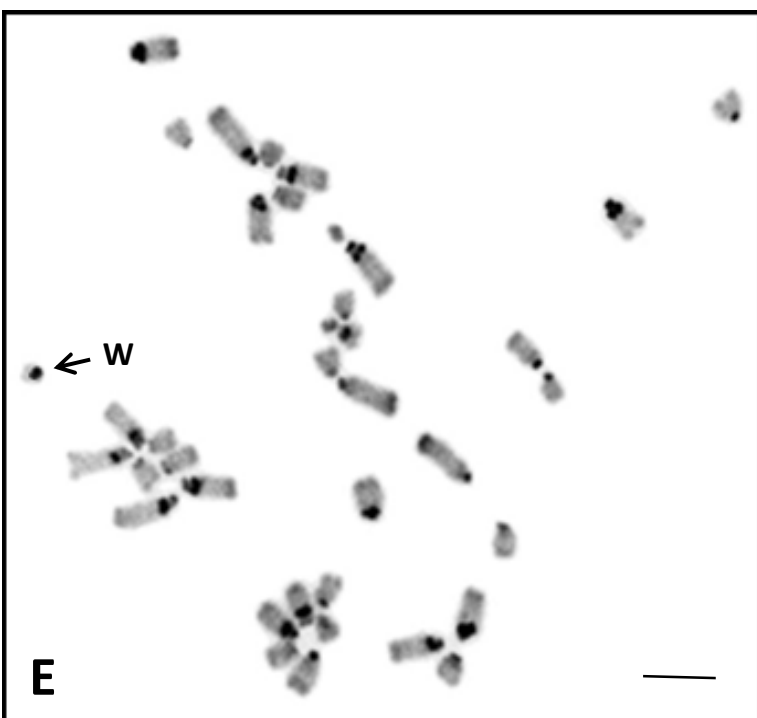
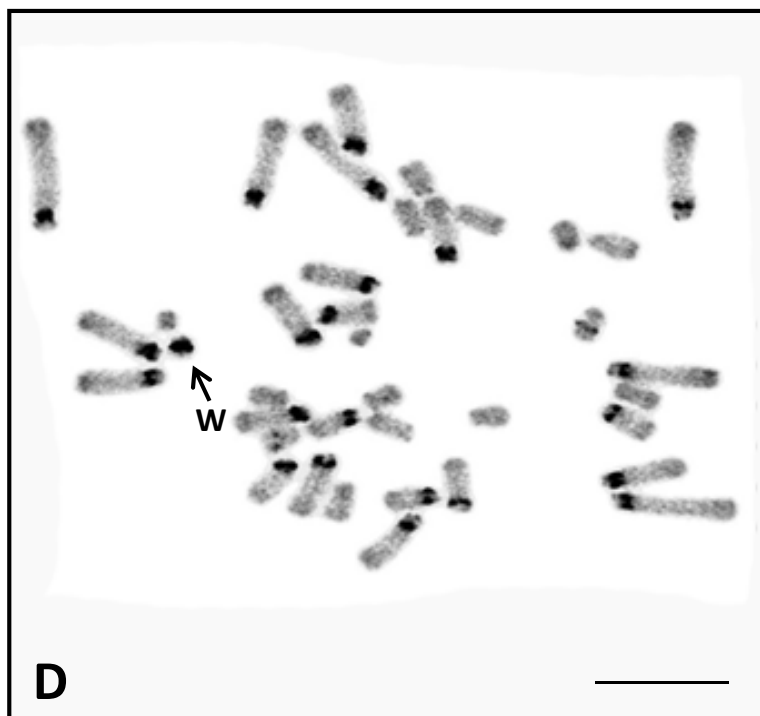
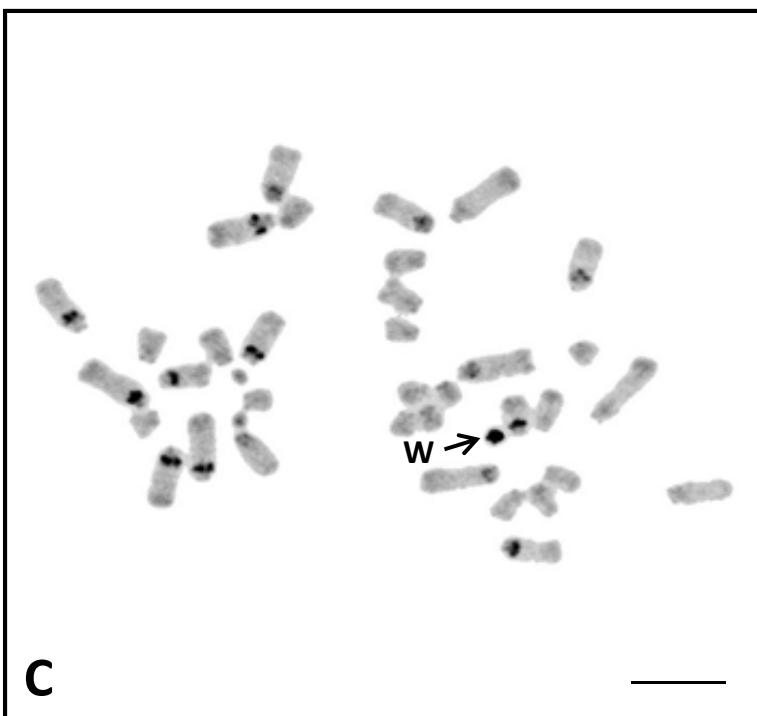
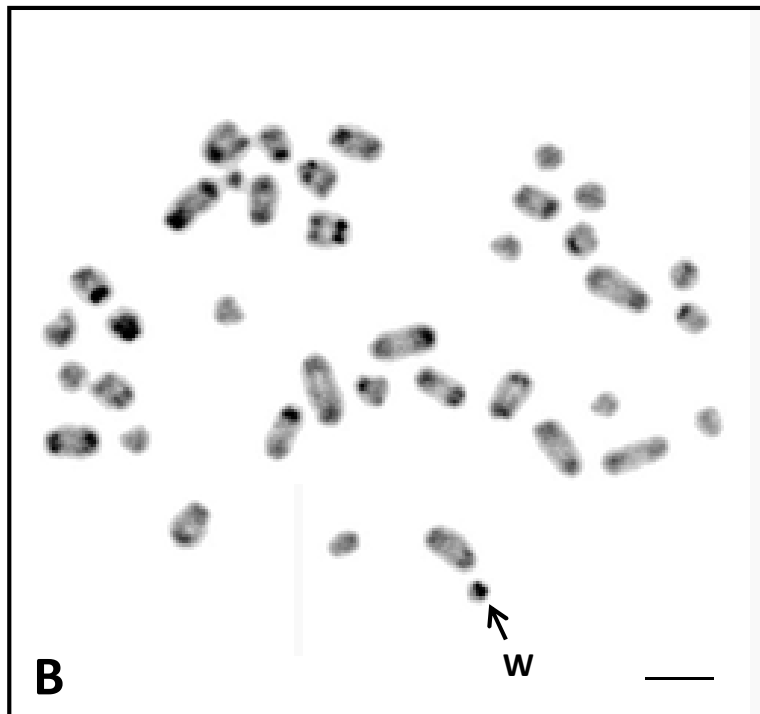
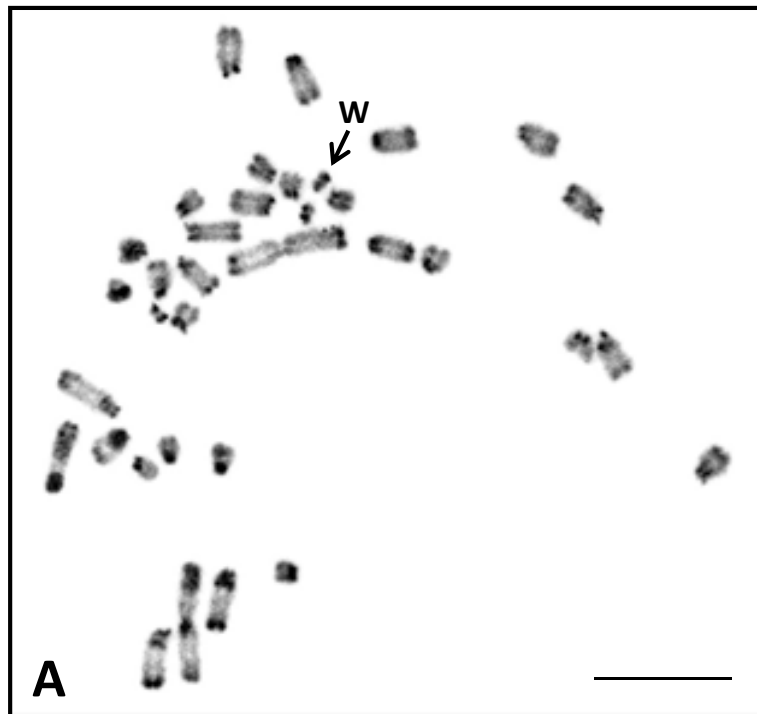


Figure 5. FISH with TaqI probe onto metaphases from *L. bilineata* female. A) the probe was produced from W-specific repeats; B) the probe was produced using autosome-specific repeats. The arrow indicates the W chromosome.



**Figure 6.** FISH with telomeric (TTAGGG)<sub>n</sub> probe onto metaphases from females of: (A) *Timon lepidus*, arrows indicate chromosome W and 18, (B) *Lacerta agilis*, (C) *L. trilineata*, (D) *L. strigata*, (E) *L. bilineata*.





**Figure 7.** C-banded of metaphase plate of females of *T. lepidus* (A), *L. agilis* (B), *L. trilineata* (T), *L. strigata* (D) and *L. bilineata* (E). Bar = 10  $\mu$ m

## References

- Arnold, E.N., Arribas, O., Carranza, S. 2007. Systematics of the palaeartic and oriental lizard tribe Lacertini (Squamata: Lacertidae: Lacertinae), with descriptions of eight new genera. *Zootaxa* 1430: 1–86.
- Bachmann L, Sperlich D. 1993. Gradual evolution of a specific satellite DNA family in *Drosophila ambigua*, *D. tristis*, and *D. obscura*. *Mol Biol Evol*, 10(3):647-659.
- Capriglione, T. 2000. Repetitive DNA as a tool to study the phylogeny of cold-blooded vertebrates. *Chromos. Today* 13: 183–194.
- Capriglione T., Morescalchi A., Olmo E., Rocco L., Stingo V., Manzo S. 1994. Satellite DNAs, heterochromatin and sex chromosomes in *Chionodraco hamatus* (Channichthyidae, Perciformes). *Polar Biology*, 14(4): 285-290.
- Capriglione T., Olmo E., Odierna G., Kupriyanova L.A. 1994. Mechanisms of differentiation in the sex chromosomes of some Lacertidae. *Amphibia-Reptilia*, 151:1-8.
- Charlesworth B, Sniegowski P, Stephan W. 1994. The evolutionary dynamics of repetitive DNA in eukaryotes. *Nature* 371: 215-220.
- Charlesworth D., Charlesworth B., Marais G. 2005. Steps in the evolution of heteromorphic sex chromosomes. *Heredity*, 95: 118–128.
- Epplen J.T., McCarrey J.R., Sutou S., Ohno S. 1982. Base sequence of a cloned snake W-chromosome DNA fragment and identification of a male-specific putative mRNA in the mouse. *Proc. Natl. Acad. Sci. USA . Genetics*, 79: 3798-3802.
- Ezaz T., Sarre S.D., O’Meally D., Graves J.A.M. 2009. Sex Chromosome Evolution in Lizards: Independent Origins and Rapid Transitions. *Cytogenetic and Genome Research*, 127: 249-260.
- Giovannotti M., Rojo V., Nisi Cerioni P., et al. 2014. Isolation and characterization of two satellite DNAs in some Iberian rock lizards (Squamata, Lacertidae). *J Exp Zool B Mol Dev Evol* 322: 13–26.
- Godinho R., Crespo E. G., Ferrand N. D., Harris J. 2005. Phylogeny and evolution of the green lizards, *Lacerta* spp. (Squamata: Lacertidae) based on mitochondrial and nuclear DNA sequences. *Amphibia-Reptilia*, 26: 271-285.

- Harris, D.J., Arnold, E.N., Thomas, R.H. 1999. A phylogeny of the European lizard genus *Algyroides* (Reptilia: Lacertidae) based on DNA sequences, with comments on the evolution of the group. *J. Zool. Lond.* 249: 49-60.
- Hipsley C. A., Himmelmann L., Metzler D., Müller J. 2009. Integration of Bayesian molecular clock methods and fossil-based soft bounds reveals early Cenozoic origin of African lacertid lizards. *BMC Evolutionary Biology*, 9:151.
- Ijdo J.W., Wells R.A., Baldini A., Reeders S.T. 1991. Improved telomere detection using a telomere repeat probe (TTAGGG)<sub>n</sub> generated by PCR. *Nucl Acids Res* 19: 17.
- Jones K.W., Singh L. 1981. Conserved repeated DNA sequences in vertebrate sex chromosomes. *Human Genetics*, 58: 46-53.
- Jones K.W., Singh L. 1985. Snakes and the evolution of sex chromosomes. *Trends in Genetics*, 1: 55-61.
- Larkin M.A., Blackshields G., Brown N.P., Chenna R., McGettigan P.A., McWilliam H., Valentin F., Wallace I.M., Wilm A., Lopez R., Thompson J.D., Gibson T.J., Higgins D.G. 2007. ClustalW and ClustalX version 2.0. *Bioinformatics* 23:2947-2948
- Lee M.S.Y., Hugall A.F., Lawson R., Scanlon J.D. 2007. Phylogeny of snakes (Serpentes): combining morphological and molecular data in likelihood, Bayesian and parsimony analyses. *Syst Biodiv.* 5: 371–389.
- Matsubara K., Tarui H., Toriba M., Yamada K., Nishida-Umehara C. et al. 2006. Evidence for different origin of sex chromosomes in snakes, birds, and mammals and step-wise differentiation of snake sex chromosomes. *Proc. Natl. Acad. Sci. USA*, 103:18190–95.
- Navajas-Pérez R., de la Herrán R., Jamilena M. et al. 2005. Reduced rates of sequence evolution of Y-linked satellite DNA in *Rumex* (Polygonaceae). *J Mol Evol* 60: 391-399.
- Navajas-Pérez R., Ruiz Rejón M., Garrido-Ramos M.A., Aznarte J.L., Rubio-Escudero C. 2007. SatDNA Analyzer: a computing tool for satellite-DNA evolutionary analysis. *Bioinformatics* 23: 767–768.
- Navajas-Pérez R., Quesada M.E., Garrido-Ramos M.A. 2009. Effect of localization, organization, and repeat-copy number in satellite – DNA evolution. *Mol. Genet. Genomics* 282: 395-406.
- Odierna G., Kupriyanova L.A., Caprigilone T., Olmo, E., 1993. Further data on sex chromosomes of Lacertidae and a hypothesis on their evolutionary trend. *Amphibia-Reptilia*, 14(1), 1-11.

- Odierna G., Aprea G., Arribas O., Capriglione T., Caputo V., Olmo E. 1996. The karyology of the Iberian rock lizards. *Herpetologica*, 52: 542-550.
- Odierna G., Aprea G., Capriglione T., Castellano S., Balletto E. 2004. Evidence for chromosome and Pst I satellite DNA family evolutionary stasis in the *Bufo viridis* group (Amphibia, Anura). *Chromosome Research*, 12(7), 671-681.
- Olmo E, Odierna G, Capriglione T. 1987. Evolution of sex-chromosomes in lacertid lizards. *Chromosoma* 96: 33-38.
- Olmo, E., Odierna, G., Capriglione, T., Cardone, A. 1990. DNA and chromosome evolution in lacertid lizards. In: *Cytogenetics of Amphibians and Reptiles*, 181-204.
- Olmo E., Signorino G. 2005. Chromorep: a reptile chromosomes database. Internet references: <http://193.206.118.100/professori/chromorep.pdf>, 15.09.06.
- O'Meally D., Patel HR, Stiglec R., Sarre, S.D., Geordes A., Graves J.A.M., Ezaz T. 2010. Nonhomologous sex chromosomes of birds and snakes share repetitive sequences. *Chromosome Res.*, 18: 787-800.
- O'Meally D., Ezaz T., Georges A., Sarre S.D., Graves J.A. 2012. Are some chromosomes particularly good at sex? Insights from amniotes. *Chromosome Res* 20: 7-19.
- Organ C., Janes D.E. 2008. Sex chromosome evolution in sauropsids. *Integr. Comp. Biol.*, 48: 512–519.
- Parasnis A.S., Ramakrishna W., Chowdari K.V., Gupta V.S., Ranjekar P. K. 1999. Microsatellite (GATA) n reveals sex-specific differences in papaya. *Theoretical and applied genetics*, 99(6): 1047-1052.
- Plohl M., Luchetti A., Meštrović N., Mantovani B. 2008. Satellite DNAs between selfishness and functionality: Structure, genomics and evolution of tandem repeats in centromeric (hetero)chromatin. *Gene*, 409: 72– 82.
- Plohl M., Meštrović N., Mravinac B., 2012. Satellite DNA evolution. In: Garrido-Ramos MA, ed. *Repetitive DNA. Genome dynamics vol 7*. Basel: Karger, 126–152.
- Plohl, M., Meštrović, N., & Mravinac, B. 2014. Centromere identity from the DNA point of view. *Chromosoma*, 123(4), 313–325. <http://doi.org/10.1007/s00412-014-0462-0>.
- Pokorná M., Giovannotti M., Kratochvíl L., Kasai F., Trifonov V.A., O'Brien P.C., Caputo V., Olmo E., Ferguson-Smith M.A., Rens W. 2011. Strong conservation of the bird Z chromosome in reptilian genomes is revealed by comparative painting despite 275 million years divergence. *Chromosoma*, 120: 455-468.

- Rice WR. 1987. The accumulation of sexually antagonistic genes as a selective agent promoting the evolution of reduced recombination between primitive sex chromosomes. *Evolution*
- Sambrook J., Fritsch E.F., Maniatis T. 1989. *Molecular cloning: a laboratory manual*. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, USA).
- Schneider S., Roessli D., Excoffier L. 2000. Arlequin, version 2.000: a software for population genetic data analysis. Geneva, Switzerland: Genetics and Biometry Laboratory, University of Geneva.
- Schwarzacher T., Heslop-Harrison P. 2000. *Practical in situ hybridization* (Springer-Verlag New York Inc., New York, NY, USA).
- Singh L., Purdom I.F., Jones K.W. 1976. Satellite DNA and evolution of sex chromosomes. *Chromosoma*. 59 :43–62.
- Srikulnath K., Matsubara K., Uno Y., Nishida C., Olsson M., Matsuda Y. 2014. Identification of the linkage group of the Z sex chromosomes of the sand lizard (*Lacerta agilis*, Lacertidae) and elucidation of karyotype evolution in lacertid lizards. *Chromosoma*, 123(6) : 563–575.
- Steinemann S., Steinemann M. 2005. Y chromosomes: born to be destroyed. *Bioessays* 27: 1076-1083.
- Strachan T., Webb D., Dover G. 1985. Transition stages of molecular drive in multiple-copy DNA families in *Drosophila*. *EMBO J* 4: 1701-1708.
- Tamura K., Peterson D., Peterson N., Stecher G., Nei M., Kumar S. 2011. MEGA5: Molecular Evolutionary Genetics Analysis Using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. *Mol Biol Evo*. 28(10): 2731-2739.
- Ugarković Đ., Plohl M. 2002. Variation in satellite DNA profiles—causes and effects. *EMBO J.*, 21: 5955–5959.
- Uetz P. 2014. The TIGR reptile database. The EMBL reptile database. Retrieved from: <http://www.reptile-database.org/>.



# ***CONCLUSIONS***

## CONCLUSIONS

From the results of this study on the characterization of two satDNAs (HindIII and TaqI) in various lacertid species different conclusions can be drawn. A first aspect concerns the different evolutionary patterns exhibited by these genomic elements depending on the chromosomal location of repeats arrays. The different evolutionary rate found between HindIII (located at centromeres) and TaqI (pericentromeric/interstitial location) satDNAs in the species of *Iberolacerta* seems to show how the chromosomal location may play an important role in determining the exchange rate between arrays of repeats located on different chromosomes, thus influencing the evolutionary rate of satDNA repeats. The effect of chromosomal location on the homogenization rate is also confirmed by the clear separation in two distinct phylogenetic clades of the TaqI repeats in *Lacerta* where this satDNA also occurs on the W chromosome. The reduced recombination rate of the W hinders the homogenization of the repeats located on this chromosome with those located on the autosomes leading to their evolutionary divergence. On the other hand, these findings are consistent with the observation by de la Herrán et al. (2001a,b) (see chapter I for references), showing a lower mutation rate of satDNA in sturgeons as compared to sparids, with more symmetrical karyotypes of these latter fishes representing no physical barrier to interchromosomal exchange.

A second aspect emerging from the present research is the wide phylogenetic distribution of TaqI satDNA among lacertids and its conservation at least 20 million years. This conservation over a long evolutionary time span opens the way to another hypothesis other than the one cited above regarding the chromosomal location of this genomic element. Indeed, this alternative hypothesis could be related to the fact that some functional constraints and natural selection (e.g. Plohl et al., 2008, see introduction) may slow down the evolutionary rate of TaqI satDNA. In this context, it would be interesting to test in future studies if transcripts of this satDNA occur in the transcriptome of the lizards from which it was isolated.

Another interesting result was obtained from the analysis of intragenomic variability of HindIII and TaqI satDNAs in *Iberolacerta*. This research revealed two different evolutionary pattern for the two satellites. However, some common features were highlighted for these genomic elements. In particular, for each satDNA family: i) libraries of different monomer variants are shared by species that are more closely related were found; ii) species-specific satellite repeats result from expansions and/or contractions of different monomer variants from the library; (iii) great differences in overall sequence homogeneity and concerted or non-concerted evolution patterns derive from different

turnover rates, even among closely related species. As already hypothesized above, different evolutionary and turnover rates are likely to result from constraints related to karyotype architecture and the interspersed organization of diverging repeat variants in satellite arrays, and maybe also to functional interactions. On the whole, these satDNA families represent highly dynamic systems, which may play a role in the evolution of genome and species. Further studies on the genome-wide variability and organization of reptilian satDNAs may be useful to test current hypotheses.

Finally, the occurrence of TaqI satDNA in the heterochromatin of the W chromosome of the four *Lacerta* species investigated indicates how these genomic elements can be involved in sex chromosome differentiation by heterochromatinization. The absence of these repeats on the w chromosome of *Iberolacerta* species confirms how the recruitment of repetitive DNA in the heterochromatin of heterochromosomes is due to contingency rather than a functional role of these elements in the differentiation process, as also demonstrated for other reptiles (see Pokorná et al. 2011 references introduction).