Genomic Distribution of Telomeric DNA Sequences – What Do We Learn from Fish About Telomere Evolution?

Konrad Ocalewicz
University of Warmia and Mazury in Olsztyn
Poland

1. Introduction

Ends of the eukaryotic chromosomes are capped with nucleoprotein complexes named telomeres. The DNA component of the telomeres usually is consisted of tandemly repeated G-rich DNA short sequences like TTTAGGG in plants (Cox et al., 1993; Fuchs et al., 1995), G_{2-8}TTAC(A) in the fission yeast (Schizosaccharomyces pombe) (Murray et al., 1986) and T(G)_{2-3}(TG)_{1-6} in baker’s yeast (Saccharomyces cerevisiae) (Shampay et al., 1984), TTGGGG in Tetrahymena thermophila (Blackburn et al., 1978), TTGGGC in Caenorhabditis elegans (Cangiano and La Volpe, 1993) or TTAGG in the insects (Okazaki et al. 1993), among others (for more telomeric DNA sequences see Telomerase Database, http://telomerase.asu.edu/). In all vertebrates studied to date, telomeres contains tandemly repeated G-rich hexanucleotide sequence (TTAGGG/CCCTAA)n and the associated proteins comprising six subunits: TRF1, TRF2, POT1, TIN2, TPP1 and RAP1 (Bolzán and Bianchi, 2006). The telomeric DNA length shows huge interspecies variation and ranged from less than 100 bp (base pairs) in the ciliate Oxytricha (Klobutcher et al., 1981), hundreds of base pairs in the baker’s yeast to 50 - 150 kb (kilo base) in the laboratory mouse (Mus musculus) (Kipling and Cooke, 1990) or even more (up to 2 Mb in chicken Galus galus domesticus) (Delany et al. 2003). The human normal cells show telomeric DNA of 5-20 kb length (Moyzis et al., 1988). Variation in the length of the telomeric arrays have been observed between non-homologous and even homologous chromosomes within individual cells in human and mice, among others (Landsorp et al., 1996; Zijlmans et al., 1997). Moreover, p-arm telomeres have been shown to be shorter that their q-arm counterparts in the mouse and Chinese hamster chromosomes (Slijepcevic et al., 1997). Mammalian telomeres replicate throughout S phase: some of the telomeres replicate early while other telomeres replicate later (Zou et al., 2004). Moreover, asynchronous replication of the mammalian p- and q-arm telomeres of the same chromosome has been observed (Zou et al., 2004).

Telomeres prevent chromosomes from end-to-end fusions, allowing DNA repair machinery distinguish natural chromosomal ends from the ends that appear in the course of breakage events (de Lange, 2002; Bolzán and Bianchi 2006). Telomeres ensure proper chromosome topology in the nucleus and may silence genes located in the vicinity of the telomeric region, and this phenomenon is called a “telomere position effect” (Luderus et al., 1996;
Copenhaver and Pikaard, 1996). As the linear DNA cannot be entirely replicated by the DNA polymerases because of the “end replication problem” (Watsan, 1972; Olovnikov, 1973), telomeres ensure complete replication of the chromosomal DNA and protect chromosomes from degradation (de Lange, 2002). Thus, telomeres shorten after each round of the cell division. In the cultured human cells, the loss of the telomeric repeats during each S phase has been estimated for 50-200 bp (Huffman et al., 2000). This loss may be compensated by telomerase, an enzyme whose catalytic protein subunit (TERT, telomerase reverse transcriptase) adds telomeric DNA repeats to the end of telomeres using as a template an integral RNA component (TR, telomerase RNA). Moreover, different cellular mechanisms may be used for the telomere length maintenance/elongation such as reciprocal recombination and transposition of the chromosomal terminal elements when telomerase is not active or inactivated (Biessmann and Mason, 1997).

Although telomeres, by definition, are terminal elements of the chromosomes, telomeric DNA repeats are also observed at internal chromosomal sites and are called Interstitial (or Interchromosomal) Telomeric Sequences (ITSs), Interstitial Telomeric Repeat sequences (ITRs) or Interstitial Telomeric Bands (ITBs). ITTs may be located close to the centromeres or between centromere and the real telomeres. The first and the most well-known description of the existence of unusual locations of telomeric DNA sequences far from their natural occurrence at the ends of the chromosomes was brought to light in 1990 by Meyne and collaborators. These authors identified telomeric repeats at non-telomeric locations in 55 out of 100 vertebrate species studied. ITSs have been observed in the exponents of four classes of vertebrates: Mammalia, Aves, Reptilia and Amphibia. The majority of the intrachromosomally located (TTAGGG)$_n$ sequences were observed at the pericentromeric areas of the bi-armed chromosomes within or at the margin of the constitutive heterochromatin (Meyne et al., 1990). This observation led to a conclusion that ITSs might have been left by the ancient centric fusions of ancestral chromosome. Since then, more sensitive FISH techniques enabling identification of telomeric repeats such as PNA-FISH using peptide telomeric probe and PRINS using (TTAGGG)$_7$/ (CCCTAA)$_7$ primers for amplification of telomeric DNA have been developed (Koch et al., 1989; Terkelsen et al., 1993). Application of such approaches together with chromosome banding techniques, molecular cloning, and genome sequencing led to identification of ITSs in species that were not studied previously to this regard as well as re-examination of the species that did not show any ITSs formerly.

Below, patterns of the chromosomal distribution of telomeric DNA sequences in several chosen vertebrates have been reviewed in the context of the chromosomal rearrangements and other mechanisms that may lead to the internal insertion of the telomeric repeats. Special attention has been paid to the distribution of the telomeric DNA sequences in the fish genome. Fishes with more than 30,000 species are the most numerous and diverse group of vertebrates (Nelson, 1994). This group of vertebrates comprises jawless fishes (hagfishes, lampreys), cartilaginous fishes (sharks and rays), and bony fishes (lobe-finned fish and ray-finned fish) (Nelson, 1994). Ray-finned fishes species represent more than 95% of all the extant fishes. More than 99.8% of ray-finned fishes belong to Teleostei (Volf, 2004). Although ancestral teleostean karyotype comprising 48-50 of one-armed chromosomes is still the most frequently observed pattern within teleosts, species with more derived
karyotypes composed of both – one- and bi-armed chromosomes – have been also observed. Diversification of teleostean karyotypes is attributed to whole genome duplication event in the Teleost ancestor and chromosomal rearrangements (Zhou et al., 2002). Moreover, some of the Teleost fish families like Salmonidae are thought to have a tetraploid origin. Tetraploidization event in the Salmonid ancestor has been followed by the rediploidization process leading to the recovery of disomic segregation and performed by the various chromosomal rearrangements like fusions and inversions (Phillips and Rab, 2001). On top of that, androgenetic fish developing in the gamma/X radiation-enucleated eggs seem to be promising models for studying the role of telomerase in the fish DNA Double Strand Breaks repair machinery (Ocalewicz et al., 2004a, 2009).

2. Classification of interstitial telomeric DNA sequences

Based on the chromosomal location, length, DNA composition and the origin, several kinds of the ITSs have been described (Nergadze et al., 2004; Bolzán and Bianchi, 2006; Lin and Yan, 2008; Ruiz-Herrera et al., 2008). In the human genome, three classes of ITSs have been proposed based on the sequence organization, localization, and flanking sequences by Azzalin et al. (2001): (Class 1) so-called short ITSs, composed of a few exact telomeric repeats up to 20 hexamers; (Class 2) subtelomeric ITSs consisted of several hundred base pairs of tandem repeats, many of which differ from the TTAGGG repeat sequence by one or more base substitutions and (Class 3) ITS sites formed by the ancestral chromosome fusions and composed of head-to-head arrays of repeats. Short ITSs may be further divided into five subclasses based upon their flanking sequences (Lin and Yan, 2008). Subtelomeric ITSs are observed at all human chromosomes, and short ITSs have been identified at 50 loci in human chromosomes, while only one ITSs derived from the fusion event have been described in the human genome (Azzalin et al., 1997; Azzalin et al., 2001; Ijdo et al., 1991). ITSs that represent class 1 and 2 may appear in the course of repair of double-strand breaks (DSBs) by the mechanism employing action of telomerase and/or recombination involving chromosome ends in the germ lines during evolution (Nergadze et al., 2004). Further rearrangements like amplifications, deletions, or transpositions of ITSs may cause its uneven distribution in the genome, for example (Lin and Yan, 2008).

One of the recent proposition based on the purely cytogenetic characteristic of non-telomeric distribution of (TTAGGG)_n repeats in mammalian species is to differentiate two kinds of ITSs: short stretches (from a few to a few hundred base pairs) of internally located telomeric repeats (s-ITSs) and long stretches (up to hundreds of kilo base) of the heterochromatic ITSs (het-ITSs) mainly assigned to the centromeric chromosomal regions (Ruiz-Herrera et al., 2008). Short ITSs composed of head-to-tail tandem arrays are widely distributed in human, chimpanzee, mouse, or rat (Azzalin et al., 2001; Nergadze et al., 2007). Analysis of DNA sequences adjacent to the s-ITSs suggested that telomeric sequences were internally inserted by transposition or synthesized by telomerase to repair DNA double-strand breaks (DSB) (Ruiz-Herrera et al., 2008). Heterochromatic ITSs on the other hand, seem to originate in the course of the ancestral chromosomal rearrangements, mostly fusions, accompanying evolution of mammalian karyotypes. Such ITSs are usually co-localized with heterochromatic regions. Although such classification of ITSs has been attributed to mammalian genome, ITSs of various origin have been also observed in the non-mammalian species.
3. Chromosome rearrangements and distribution of telomeric DNA sequences in the vertebrates

3.1 Internally located telomeric repeat sequences as relicts of the chromosome fusions

Telomeric DNA observed at the non-telomeric locations might be associated with known chromosome rearrangements, like centric fusions (Robertsonian translocations) and tandem fusions. Fusion of two one-armed chromosomes leading to the formation of one metacentric or submetacentric chromosome may leave telomeric DNA sequences at the fusion site at the pericentromeric location. This region is usually heterochromatic. Interstitial non-centromeric sites of (TTAGGG)\textsubscript{n} sequences may be relicts of the tandem fusions. In such cases, coincidence between ITSs and heterochromatin is rarely observed (Nanda et al. 2002). Irrespective of the origin, such ITSs might be organized in very long arrays that are much longer than those observed at the chromosomal ends. In the Chinese hamster, large pericentromeric interstitial telomeric DNA sites are observed (Bertoni et al., 1996), and telomeric DNA sequences have been discovered to be the main component of the satellite DNA with its abundance reaching up to 5% of the Chinese hamster genome (Bertoni et al., 1996; Slijepcevic et al., 1996; Faravelli et al., 1998; 2002). Moreover, ITSs might be interspersed with other repetitive DNA sequences (Salvadori et al., 1995). Sometimes, chromosome breakage occurs within the ITS region (Alvarez et al., 1993; Slijepcevic et al., 1996).

Internally located telomeric DNA sequences have been observed in many mammalian and non-mammalian species showing more degenerative karyotypes when compared to their plesiomorphic (ancestral) complements. A 2n = 22 karyotype, is thought to be an ancestral for the marsupial family Macropodidae (kangaroos and wallabies) (Metcalfe et al. 2007). In the swamp wallaby (Wallabia bicolor) (2n= 10 in female, 2n= 11 in male) telomeric DNA sequences were retained at the fusion sites in four chromosomes formed in the course of centric fusions (Metcalfe et al., 1998). The lowest chromosome number exhibited in the mammalian species equals 6/7 (female/male) and is observed in the Indian muntjac deer (Muntiacus muntjak vaginalis MMV). The common ancestor of the muntjaccs lived about 1.7-3.7 million years ago and its karyotype was presumably composed of 70 chromosomes. (Hartman and Schertham, 2004). Cytogenetic survey of the muntjaccs revealed that chromosome reduction observed in the genus occurred linearly from the putative ancestral complement 2n= 70 through a Chinese muntjac-like (2n= 46) to a Faa’s muntjac-like (2n= 13/14) karyotypes. Further chromosome reduction to 2n= 8/9 observed in the Black and Gongsham muntjac and to 6/7 chromosomes in the Indian muntjac were rather independent events (Wang and Lan, 2000). Such drastic chromosome reduction and karyotype diversification that happened in such a short stretch of time has been supposed to be caused by the multiple tandem fusions and relatively few centric fusions (Hsu et al., 1975). This assumption has been later proved by the comparative and molecular cytogenetic analysis of the muntjac genome (Lee et al., 1993; Schertham, 1995; Yang et al., 1997; Zou et al., 2002). Several sites of internally located telomeric repeat sequences in the Indian muntjac chromosome were observed to be co-localized with satellite DNA repeats (Lee et al., 1993; Schertham, 1995; Hartman and Schertham, 2004). Such interstitial satellite DNA sequences were assumed to be the “footprints” of the breakage of chromosomal syntenies in the Indian muntjac and thus may be treated as relicts of the ancestral fusion points (Fronicke and...
Scherthan, 1997). In the Hartman’s zebra (*Equus zebra hartmannae*) showing karyotype composed of relatively low chromosome number (2n= 32) when compared to other equids (2n= 44-66), several sites of internally located telomeric repeats have been described (Santani et al., 2002). Comparison of the chromosomal distribution of ITSs and comparative chromosome painting of human and Hartman’s zebra showed that all ITSs are located at the junctions of evolutionary conserved human- Hartman’s zebra chromosomal segments, suggesting that ITSs are relicts of the putative fusions of ancestral chromosomes (Santani et al., 2002). Telomeric sequences at the fusion sites have been also observed in other mammalian species like okapi (*Okapia johnstoni*) (Vermeesch et al., 1996), Eulemur species (Garagna et al. 1997), akodont rodents (*Akodon cursor* and *Bolomys lasiurus*) (Fagundes and Yonenaga-Yassuda, 1998), lemurs (Go et al. 2000), rock wallabies (*Petrogale*) (Metcalfe et al., 2002), among others.

Chromosome fusions seem to play an important role during the avian karyotype evolution. The avian karyotype has a characteristic structure. It comprises several pairs of relatively gene-poor macrochromosomes and numerous microchromosomes enriched with genes, and even distant species show similar karyotypes (Nanda et al., 2002). It has been discover that chicken telomeric DNA sequences range from 0.5 kb to about 2 Mb (Delany et al., 2000, 2003). Telomeric DNA sequences cover up to 4% of the chicken diploid genome, which is contrasted with a rather low amount of the telomeric DNA in the human diploid cell (about 0, 3%) (Delany et al., 2003). Based on the size and genome location three classes of telomeric DNA arrays were distinguished in the chicken. It has been suggested that telomeric DNA arrays ranging from 0.5 to 10 kb in length (Class I arrays) represent the interstitial telomeric DNA sequences, while the larger tracts arrays ranging from 10 to 40 kb (Class II) and from 200 kb to 2 Mb (Class III) represent telomeric DNA from the chromosome terminus (Dealny et al. 2003). Many of the cytogenetically studied bird species have exhibited telomeric DNA sequences in non-telomeric positions on the macrochromosomes. However, patterns of their distribution are different in the primitive (Palaeognathae) and modern (Neognathae) birds (Meyne et al., 1990; Nanda et al., 2002). The primitive birds like ostrich (*Struthio camelus*), emu (*Dromaius novaehollandiae*) and the American rhea (*Rhea americana*) show numerous interstitially located telomeric DNA sites along the entire length of most of the macrochromosome arms. Rather few of the macrochromosomes show ITSs at the (peri)centromeric positions. In the rhea and emu most of the interstitially located telomeric sequences did not coincide with the C-banded heterochromatin. Such distribution pattern of the telomeric DNA sequences in these birds has been proposed to be due to the tandem fusions of macro and microchromosomes in their common ancestor. On the other hand, there are only few if any internally located telomeric DNA sequences in the modern birds like duck (*Cairina moschata*), greylag goose (*Anser anser*), the ring-necked pheasant (*Phasianus colchicus*), Japanese quail (*Coturnix coturnix*) and parrots (Nanda et al., 2002). Centromerically located telomeric DNA sequences that coincide with the heterochromatin observed on the bi-armed macrochromosomes in two owl species are likely relicts of the chromosome centric fusions. (Meyne et al., 1990; Nanda and Schmid, 1994; Delany et al., 2003).

Reduction of the chromosome number from the ancestral 2n= 32 to 2n= 16 in the lizard *Gonatodes taniae* probably occurred through the centric fusions. Telomeric DNA sequences observed at the pericentromeric regions of all *G. taniae* bi-armed chromosomes were
presumably the remnants of the above-mentioned rearrangements. On the other hand, interstitially located telomeric repeats could be also a major component of the repetitive DNA in the pericentromeric C band-positive heterochromatin (Schmid et al. 1994) (see chapter 3.3). Similar location of the telomeric repeats in one and three meta-submetacentric chromosomes in the Brazilian lizards, *Leposoma guianense* and *L. oswaldoi*, respectively indicated Robertsonian translocations were involved in the evolution of these lizards’ karyotypes (Pellegrino et al., 1999). Centric fusion in the Brazilian gecko, *Gymnodactylus amarali* also left telomeric repeat DNA sequences at the fusion sites of two chromosomes (Pellegrino et al., 2009).

As pericentromeric ITSs quite frequently coincide with the heterochromatin, it has been proposed to describe such ITS sites as heterochromatic ITSs (het-ITSs) by Ruiz-Herrera et al. (2008), who suggested a four-step mechanism to explain the presence of such sites in the fused chromosomes. The first step is the initial fusion event without loss of the telomeric sequences from the fusion site (1). The next step is formation of the (peri)centromeric heterochromatin by expansion of the internally located telomeric arrays including amplification of the telomeric sequences and other repeats (2). Subsequently the heterochromatic ITSs were reorganized via chromosomal rearrangements that may lead to the redistribution of the telomeric DNA, degeneration of the original ITS array, gradual shortening of the array, and even the loss of the ITS. Finally, breakage within the heterochromatic ITS site may result in chromosome fissions (step 4).

### 3.2 Chromosome fusions and loss of the interstitial telomeric sequences

Not all chromosome fusions occur with retention of the telomeric DNA repeats at the fusion sites. Telomeric DNA sequences from the ancestral chromosomes may be lost during or after the chromosomes fusion process. Chromosome breakage within centromeric satellite DNA followed by Roberstonian fusions leaves no telomeric repeats at the fusion sites (Garagna et al., 1995; Nanda et al., 1995). On the other hand, telomeric DNA sequences that retain at the fusion sites may undergo gradual loss leading to the shortening of the non-functional telomeric repeats and are therefore undetectable by the cytogenetic approaches (Slijepcevic, 1998). Lack of the internally located telomeric DNA sequences at the fusion points was described in the mouse (*Mus musculus*) (Garagna et al., 1995), neotropical water rat (*Nectomys*) (Silva and Yonenaga-Yassuda, 1998) and short-tailed shrew (*Blarina carolinensis*) (Qumsiyeh et al., 1997), among others.

Chicken chromosomes 1 - 4 presumably appeared in the course of the ancestral chromosome fusion events. However, only chromosome 1-3 exhibited ITSs at the fusion sites (Nanda et al., 2002). In comparison to the primitive bird species like ostrich and emu that display voluminous number of ITS sites in their chromosomes, species showing high number of the bi-armed chromosomes and listed as highly evolved such as parrots lack TTAGGG sequences at non-telomeric sites (Nanda et al., 2002). This may suggest that in the Neognathae birds ITS sites were lost after the divergence of the primitive and modern birds (Nanda et al., 2002). The lack of ITSs in the more derived karyotypes when compared to the ancestral models is in opposite to the suggestion made by Meyne et al. (1990) that ITS sites appear in the course of chromosome rearrangements accompanying karyotype evolution and thus can be observed in the evolutionary advanced species.
3.3 Non-telomeric TTAGGG sequences as components of the satellite DNA

Although many ITS sites observed within or at the margin of the constitutive heterochromatin are remnants of chromosome fusion events (Meyne et al., 1990), such coincidence is not a general rule. Australian and American marsupials (Marsupialia) presumed ancestral karyotype (2n= 14) is observed in the exponents of six of the seven extant marsupial orders (Metcalf et al., 2004). Such karyotype includes bi-armed chromosomes showing centromerically located telomeric sequences that overlap with the large amounts of heterochromatin (Pagnozzi et al., 2000, Metcalf et al. 2004). Comparison of the distribution of the telomeric DNA sequences in the ancestral and more evolved karyotypes with known chromosomal rearrangements suggested that pericentromeric and heterochromatic ITSs in the marsupial 2n= 14 complements might be a component of the native satellite DNA rather than relicts of the recent chromosome rearrangements (Pagnozzi et al., 2000; Metcalfe et al. 2004).

Most of the cytogenetically studied amphibians show telomeric DNA sequences exclusively located at the chromosomal ends (Meyne et al., 1990; Schmid et al., 2003, Schmid et al., 2009). Unexpectedly, interstitial location of the telomeric DNA sequences has been described in the quite conserved karyotypes of the American hylid frogs (Wiley et al., 1992), Xenopus laevis (Meyne et al., 1990, Nanda et al., 2008) and Xenopus clivii (Nanda et al., 2008). Homogeneity of the karyotypes among related species excluded chromosome fusions as the potential source of ITSs. Moreover, the interstitial telomeric sites in these species coincided with the constitutive heterochromatin identified in the course of C-banding. The clear correspondence between ITSs and the constitutive heterochromatin suggest that (TTAGGG)_n sequences might be a component of a repetitive DNA. Although it is still unknown how the telomeric DNA sequences were inserted into the interstitial positions and amplified, the repair of the DNA Double Strand Breaks with the telomerase should be taken into consideration (Nergadze et al., 2004, 2007). Previously, several authors suggested that telomeric or telomeric like DNA sequences were components of the satellite DNA in some vertebrates (Garrido-Ramos et al., 1998). In other species, telomeric DNA sequences are scattered along the NORs (Nucleolus Organizer Region) DNA sequences (see chapter 4.4).

4. Distribution of telomeric DNA sequences in fish

4.1 Telomerase and the length of the fish telomeres

So far, the telomerase gene in fish has been shown to be expressed in most cells throughout the entire fish life (Elmore et al., 2008; Hartman et al., 2009; Lund et al., 2009). This is in contrast to humans where telomerase activity is absent in most somatic cells but present in embryonic stem cells and tumors (Hiyama and Hiyama, 2007). Although some authors presume that high expression of fish telomerase may be related to the longevity, comparative analysis of telomerase activity in the short- and long-lived fish species showed no positive relationship between telomerase activity and the fish longevity (Elmore et al., 2008). Instead, another hypothesis was suggested: retention of the telomerase in adult fish might be crucial to maintain their regenerative capacity. To test this hypothesis, short fragments of the caudal fish tissue have been removed in medaka (Oryzias latipes), zebrafish...
(Danio rerio), and mummichog (Fundulus heteroclitus) specimens, and telomerase activities were assayed before and during the regeneration period. Telomerase was shown to be upregulated during the tissue regeneration, which suggested that telomerase is involved in the fish tissue regeneration after injury (Elmore et al., 2008).

The length of the telomeric DNA in fish studied to date varies from 2 kb to 15 kb (Chew et al., 2002, Elmore et al. 2008) and is similar to that observed in normal human cells (Elmore et al., 2008). Furthermore, retention of the telomere length throughout the entire life has been demonstrated in zebrafish, but not in the medaka (Hatekeyama et al., 2008; Lund et al., 2009). Telomere shortening with age has also been observed in the long-lived strain of Nothobranchius furzeri while such attrition has not been detected in the short-lived strain of the same species (Hartmann et al. 2009). Thus, age dependent telomere shortening in fish may be species-specific or even strain specific (Lund et al., 2009; Hartman et al., 2009).

Application of PRINS using (CCTAAA)7 primer and PNA-FISH using telomeric probes revealed different intensity of the hybridization foci on fish chromosomes (Ocalewicz and Dobosz, 2008; Pomianowski et al., 2012). As the fluorescence intensity of the telomere hybridization focus reflects the length of the telomeric repeat sequence (Zijlmans et al., 1997), the differences in the telomere hybridization signal intensity observed on different chromosomes are likely related to variations in their respective telomere lengths. Chromosome rearrangements may lead to such variation. In the albino rainbow trout (Oncorhynchus mykiss), one of the X chromosome isoforms has shorter p-arm with weak telomeric hybridization signals. Partial deletion or translocation including telomeric region has been suggested to contribute to the p-arm length difference between two morphological variants of the X chromosome (Figure 1a-b) (Ocalewicz and Dobosz, 2009).

4.2 Fish karyotype evolution and distribution of telomeric repeats:

Major rearrangements

As mentioned above, internally located telomeric DNA repeat sequences may be the relicts of the chromosomal rearrangements such as fusions that accompanied the karyotype evolution of many vertebrate species. This may be also true for the fish species. Teleostean fish, the major clade of the ray-finned fish (Actinopterygii) is the largest and the most diverse group of vertebrates. More than half of the cytogenetically surveyed actinopterygians have karyotypes composed of 48-50 chromosomes (Mank and Avise, 2006) and the complement of 48 one-armed chromosomes (NF= 48) is supposed to be ancestral in the Teleostei. Such teleostean ancestral like karyotype (2n= 48) may be the plesiomorphic condition in Scorpeanidae fish (Caputo et al., 1998). However, in Scorpaena notata, only 34 uni-armed chromosomes (NF= 34) are observed (Caputo et al., 1998). As the karyotype of S. notata comprises of only subtelo-acrocentric chromosomes, the reduction of both chromosome and chromosome arm numbers from 48 to 34 likely occurred in the course of the tandem fusions. Nevertheless, only one pair of chromosomes showed interstitial telomeric DNA sequences in the putative fusion sites in this species. Presumably other ITS sites have been lost or exist in very short arrays that may not be detected by fluorescence in situ hybridization (Caputo et al., 1998).
Fig. 1. a-b. Albino rainbow trout (*Oncorhynchus mykiss*) partial metaphase with two morphs of the X chromosome after hybridization with the telomeric probe (PNA-FISH) (a) and staining with DAPI fluorochrome (b). Arrowheads show two morphs of the rainbow trout X chromosomes. White arrowhead – long morph (X^L) with distinct p-arm and bright telomeric signals, yellow arrowhead – short morph (X^S) with reduced p-arm and weak telomeric signals. Fig. 1 c-e. Chromosomes of brook trout (*Salvelinus fontinalis*) (c), Arctic charr (*Salvelinus alpinus*) (d) and their hybrid (e) after PRINS (c, e) and PNA-FISH (d) enabling localization of the telomeric DNA sequences. White arrows indicate brook trout chromosomes with interstitial telomeric sites (ITSs), yellow arrows point the Arctic charr chromosomes with ITSs.

Huge variation in the diploid chromosome number and the karyotype composition are observed in the bitterlings (Acheilognathinae). Diploid chromosome number and the number of chromosome arms vary from 42 to 48 and 50-78, respectively, which was attributed to both Roberstonian and tandem fusions, chromosomal inversions, and some
minor rearrangements involving heterochromatic regions (Ueda, 2007). In two bitterling species, The Japanese rosy bitterling (*Rhodeus ocellatus kurumeus*) and the oily bitterling (*Tanakia limbata*), both with similar karyotypes that comprise 8 metacentric, 20 submetacentric and 20 subtelocentric chromosomes (2n= 48, FN= 76) FISH with telomeric probe was applied and showed different distribution patterns of the hybridization signals. In the Japanese bitterling interstitial telomeric sites were observed in the pericentromeric regions of 14-16 chromosomes, which is the highest number of ITSs detected in any of the fish chromosomes studied to date, whereas in the oily bitterling no such location of ITS was exhibited (Sola et al., 2003). Similar phenomenon was observed in the mammalian species that experienced several Roberstonian fusions. ITSs were not described in *Mus musculus domesticus* (Garagna et al., 1995; Nanda et al., 1995) whereas in *M. minutoides* telomeric sequences at the putative fusions sites were retained and observed near the centromeres (Castiglia et al., 2002; Castiglia et al., 2006).

Two interstitial telomeric sites have been detected in the Nile tilapia (*Oreochromis niloticus*) (2n= 44) chromosome 1 that is significantly larger than all the other chromosomes in this organism (Chew et al. 2002). This observation supported hypothesis that chromosome 1 in the Nile tilapia appeared in the course of the fusion of three chromosomes and explained the reduction of chromosome number from the ancestral teleost karyotype of 2n= 48 to 2n= 44 in the Nile tilapia (Chew et al. 2002). In *Oreochromis karongae*, diploid chromosome number is reduced to 38. The *O. karongae* karyotype comprises one large subtelocentric pair of chromosomes, four medium sized pairs (three subtelocentric, one submetacentric) and fourteen small pairs. Three of the medium sized chromosome pairs seem to derive in the course of fusions. Distribution of the telomeric repeats show two interstitial telomeric sites on the chromosome 1 similar to these observed in the Nile tilapia chromosome 1 and one ITS in each of the six fusion chromosomes. Comparison of the position of the current and relic centromeres performed with FISH and the tilapia centromere specific probe and ITS sites in *O. karongae* suggests that the three fusions all occurred in different orientations: the ends of the two q arms to produce pair 2, a p-q fusion in the case of pair 3 and a p-p fusion for pair 4 (Mota-Velasco et al., 2009). In the non-teleostean Elasmobranch fishes (sharks and rays) that are considered as the ancient vertebrates only four species have been studied with FISH and telomeric probe, so far (Rocco 2006). In two of them (*Taeniura lymma* and *Torpedo ocellata*), pericentromeric location of telomeric DNA sequences was detected in four bi-armed chromosomes (Stingo and Rocco, 2001; Rocco et al., 2001; Rocco et al., 2002). This is in agreements with the hypothesis that in cartilaginous fish, karyotype evolution involved a progressive decrease of chromosome number due to the centric fusions (Rocco, 2007).

### 4.2.1 Salmonid fish species: Chromosome fusions and lack of ITSs

Chromosomal rearrangements like centric and tandem fusions have played important role in the salmonid karyotype evolution during rediploidization process following the whole genome duplication experienced by the salmonid ancestor 100-25 mya (Allendorf and Thorgaard, 1984). The polyploid origin of the Salmonidae has been considerably substantiated (Leong et al., 2010). Both the genome size and the chromosome arms number are approximately twice that of the Salmonid closest relatives, the Esociformes (Phillips and Rab, 2001). Most of the salmonid species have karyotypes composed of both bi-armed and
one-armed chromosomes and have the chromosome arm number (FN) that ranged from 94 to 104, while the related Esocidae fish have the ancestral teleostean karyotype with about 50 one-armed chromosomes. Different chromosome number and the constant chromosome arm number resulted from the centric fusions known as Robertsonian polymorphisms are observed in the salmonid fish from the genera *Hucho*, *Salmo*, *Oncorhynchus* and *Salvelinus* (Phillips and Ráb, 2001). Moreover, large acrocentric chromosomes in the Atlantic salmon (*Salmo salar*) karyotype are thought to be the result of tandem fusions (Phillips and Ráb, 2000). Robertsonian fusions, paracentric and pericentric inversions were suggested to be involved in changes leading to the establishment of the present karyotypes of three *Coregonus* species: European whitefish (*Coregonus lavarætus*), vendace (*Coregonus lavaretus*) and peled (*Coregonus peled*) (Jankun et al., 2007). Unexpectedly, none of the cytogenetically studied salmonid fish species with fused meta- and submetacentric chromosomes showed pericentromeric locations of the telomeric repeats (Abuin et al., 1996; Jankun et al., 2007; Ocalewicz et al., 2008). The lack of ITS at the putative fusion sites in the bi-armed salmonid chromosomes may suggest p-arm telomeres were lost in the course of the chromosome breakage that preceded chromosome fusions. On the other hand, telomeric repeats retained at the fusion sites might have experienced successive loss and degeneration leading to gradual shortening of the non-functional telomeric arrays (Slijepcevic, 1998). Consequently, too short internally located telomeric repeats may be below the resolution of the techniques enabling chromosomal location of DNA sequences. On the other hand, interstitial telomeric DNA sequences located far from the centromeric region have been detected in the Atlantic salmon large subtelocentric chromosomes, which supported hypothesis concerning tandem fusions as the mechanism leading to the formation of some of the chromosomes in this species (Abuin et al., 1996).

### 4.3 ITSs and minor rearrangements – A *Salvelinus* fish case

Other mechanisms leading to the ITS formation have been suggested in three *Salvelinus* species: lake trout (*Salvelinus namaycush*), brook trout (*Salvelinus fontinalis*) and the Arctic charr (*Salvelinus alpinus*) showing subterminal position of the interstitial telomeric sequences (Figure 1c-d) assigned to the vicinity of the CMA3 positive GC-rich heterochromatin (Reed and Phillips, 1995; Ocalewicz et al., 2004b; Pomianowski et al., 2012). Guanine-rich chromosomal regions are involved in several rearrangements like transpositions, duplications and (or) translocations resulted in multichromosomal location and variation in size of CMA3 positively stained chromatin in *Salvelinus* species (Phillips et al., 1988; Phillips and Ráb, 2001). Dispersion of CMA3 positive chromatin segments among homologous and non-homologous chromosomes could be followed by the insertion of the telomeric repeats linked to the translocated chromosome fragment into the interstitial position. Similar location of ITS in these three species may indicate similar mechanism leading to the insertion of (TTAGGG)$_n$ sequences in the non-telomere position in the *Salvelinus* fish. In the case of the Arctic char metaphase spreads showing extended chromatin, the non-telomeric fluorescent hybridization signal covered a longer stretch of the chromosome than the signal from the telomere position, however the interstitial signal was less intense. This observation suggested that telomeric regions and ITS might have different structures (Pomianowski et al., 2012). It is possible that telomeric DNA sequences were not the only component of the ITS region. Internally inserted short telomeric repeats are frequently flanked by the
repetitive or transposable elements and undergo amplification process leading to elongation/expansion of the chromatic region built with different DNA sequences including telomeric repeats (Garrido-Ramos et al., 1998).

It has been also observed that ITSs might be considered as sites fragile for recombination and thus may potentially increase rates of chromosome breaks and rearrangements (Lin and Yan, 2008). This could partially explain the high level of size and location polymorphisms of the heterochromatic regions in Salvelinus species (Phillips et al., 1988; Phillips and Ráb, 2001; Pomianowski et al., 2012). Moreover, chromosomes with unusual distribution of telomeric DNA sequences may be useful cytogenetic markers enabling identification of parental chromosomes in hybrid organisms. Recently, Arctic charr and brook trout chromosomes with internally located telomeric repeats have been identified in the karyotype of Arctic charr x brook trout hybrids (Figure 1e) (Ocalewicz, unpublished).

4.4 Other uncommon locations of the telomeric sequences in fish

In addition to the interstitial location of the telomeric DNA, (TTAGGG)_n repeats may also coincide with the nucleolar organizer regions (NORs). Telomeric DNA sequences are observed to scatter along the heterochromatic NORs in the Atlantic eels (Anguilla anguilla) (Salvadori et al. 1995), rainbow trout (Abuin et al., 1996), straight-nosed pipefish (Nerophis ophidion) (Libertini et al., 2006) and three mullet species (Mugilidae) (Sola et al., 2007). Such unusual distribution of TTAGGG repeats suggests telomeric sequences are interspersed with rDNA sequences. Similar location of telomeric repeats has been previously described in mammalian species including American mole (Scalopus aquaticus), Seba's fruit bat (Carollia perspicillata) (Meyne et al., 1990), wood lemming (Myopus schisticolor) (Liu and Fredga, 1999), and amphibians Xenopus borealis and Xenopus muelleri (Nanda et al., 2008). The origin of the telomeric sequences interspersed with NORs is unclear. It was suggested that the presence of telomeric repeats within NORs may cause unequal crossing-over and thus give rise to the chromosomal length polymorphism (Salvadori et al., 1996). Moreover, the presence of the telomere sequences may epigenetically inactivate NORs (Guillén et al., 2004; Copenhaver and Pikaard, 1996).

In the sturgeon Acipenser gueldenstaedti, two entire chromosomes were light up with the fluorescent signals derived from the telomeric probe in FISH analysis (Fontana et al., 1998). Similar observation has been made in some of the bird microchromosomes. The ability of interstitial telomeric repeats to promote recombination (Ashley and Ward, 1993) may explain enormously high recombination rate in the bird microchromosomes (Nanda et al., 2002).

4.5 Distribution of telomeric DNA sequences in the androgenetic fish

Androgenesis is a reproductive process in which diploid offspring inherit only paternal nuclear DNA. Although natural (spontaneous) androgenesis is observed in limited number of plant and animal species (McKone and Halperrn, 2003), paternal chromosome inheritance can be induced intentionally in fish (Komen and Thorgaard, 2007). Artificial androgenesis includes three steps: inactivation of the nuclear DNA in eggs by UV or ionizing (gamma and X) irradiation, insemination of enucleated eggs with untreated or cryopreserved sperm, and
diploidization of the paternal chromosomes by exposition of the haploid zygotes to temperature or high pressure shock to suppress the first mitotic division (Komen and Thorgaard, 2007). UV radiation damages chromosomes by inducing thymidine dimers that inhibit process of replication what results in DNA fragmentation, and gamma and X radiations act by inducing chromosome breaks like double strand breaks (DSB). Insufficient dose of radiation results in incomplete inactivation of maternal nuclear genome. Undamaged pieces of the irradiated genome in the forms of chromosome fragments were observed in the androgenetic alevins and adult fish (Parsons and Thorgaard, 1985; Ocalewicz et al. 2004a). In the course of partial inactivation of maternal chromosomes, different chromosome fragments may be provided; acentric terminal fragments with telomeric region at only one end or without any telomeres and centric incomplete chromosomes without telomeres, or telomere at only one arm. Additionally, dicentric chromosome can be formed when the broken end of one centric incomplete chromosome join with a broken end of another incomplete chromosome (Disney et al., 1988). Acentric fragments may be removed from the zygote during the cell divisions or may associate with or even incorporate into paternal intact chromosomes. ITSs observed on the androgenetic rainbow trout chromosomes could be the remnants of the incorporation process (Figure 2a) (Ocalewicz et al., 2004). The centric chromosome fragments with chromosome breaks on both sides of the centromere form ring chromosomes presumably in the course of non-homologous end joining (NHEJ) repair (e.g. Pfeifer et al., 2004). On the other hand, the broken ends of the chromosomes could have been repaired with the telomeric DNA repeats synthesized de novo by telomerase or another mechanism capable of de novo telomere addition (Biessmann et al., 1990). Some of the ionizing radiation induced fish chromosome fragments retained linear construction with telomeric DNA sequences newly added to their broken ends (Figure 2b) (Ocalewicz et al., 2009). Chromosome fragment with two interstitially located telomeric signals observed in the androgenetic brook trout (Figure 2c) might have been also ring chromosome formed in the course of fusion of a radiation-broken chromosome arm with the opposite unbroken arm or arm broken within telomeric region (Henegariu et al., 1997). However we do not exclude that this fragment might have originated from one of the brook trout chromosomes with interstitially located telomeric DNA sequences (Ocalewicz et al., 2004b). Chromosome fragments showing two, always terminally located telomeres detected in the androgenetic brook trout represented another chromatin arrangements (Figure 2d). Such shape and distribution of the hybridization spots suggested formation in the course of the telomere loss in only one chromosome arm and fusion between sister chromatids. On the other hand, both chromosomal ends might have been broken and repaired in the course of two mechanisms – action of the telomerase may heal broken ends of the p-arm while broken ends of the q-arm may undergo fusion. Although telomerase is capable of healing the broken ends of the irradiated fish chromosomes, most of the fragments show spherical shape. It is possible that fish telomerase is not always able to heal the broken chromosome ends with newly synthesized telomeric DNA due to the limited access to DNA breaks (Latre et al., 2004). However, telomerase in fish seems to be involved in the ionizing radiation induced DSB repair, which is in agreement with the observations made in human, chimpanzee, mouse and rats genomes, where analysis of flanking sequences suggested that some of the ITSs were inserted during the repair of DSB (Ruiz-Herrera et al., 2008).
Fig. 2. Chromosomes of androgenetic progenies of rainbow trout (*Oncorhynchus mykiss*) (a), brook trout (*Salvelinus fontinalis*) X Arctic charr (*Salvelinus alpinus*) hybrid (b) and brook trout (c-d) after Primed IN Situ (PRINS) technique with telomeric (CCCTAA)$_7$. White arrows indicate chromosomes with interstitial telomeric DNA sequences (a), yellow arrowhead shows linear chromosome fragment with telomeres (b), white arrowheads point to the telomerless ring chromosome fragments (b), pink arrow indicates chromosome fragment with interstitially located telomeric signals (c) while yellow arrow indicates chromosome fragment with telomeric signals situated terminally (d). Both type of chromosome fragments with telomeric signals are enlarged and framed (c, d).
5. Conclusions

Chromosome fusions are the source of the interstitial telomeric DNA sequences (ITSs) in the vertebrates. On the other hand, quite frequently such rearrangements involve loss of the telomeric repeats at the fusion sites. ITSs can also appear in the course of DNA DSB repair. Internally located telomeric DNA sequences may undergo amplification, degeneration and/or further redistribution. TTAGGG repeats may be part of the centromeric and subterminal satellite DNA or rDNA forming nucleolus organizer regions (NORs). Fish seem to be good models to study the distribution and genomic organization of the ITSs. First, most of the ITSs observed in the fish chromosomes appeared in the course of the similar mechanisms responsible for the ITS formation in the higher vertebrates. Second, apart from the well-known fusion scenario of the ITS origin, other genomic rearrangements such as transposition-mediated translocations of the chromosomal regions including telomeric DNA sequences may result in the interstitial inclusion of the telomeric repeats in fish chromosomes. ITSs observed in the androgenetic fish derive from the incorporation of the ionizing radiation induced terminal acentric chromosome fragments into the intact chromosomes. Moreover, fish telomerase, which is active during the entire ontogenetic development, may be involved in the DSBs repair mechanism in these organisms.

6. Acknowledgments

Results concerning chromosomes of androgenetic brook trout (Salvelinus fontinalis) and androgenetic brook trout X Arctic char (Salvelinus alpinus) hybrids described in the chapter 4.5 had been obtained in the course of the research supported by the Polish Ministry of Science and Higher Education, Project No. N311 525240.

7. References


