J Mol Evol (1998) 46:245–255

JOURNAL OF MOLECULAR EVOLUTION

Characterization of Novel Minisatellite Repeat Loci in Atlantic Salmon (*Salmo salar*) and Their Phylogenetic Distribution

John L. Goodier,* William S. Davidson

Department of Biochemistry, Memorial University of Newfoundland, St. John's, Newfoundland, Canada A1B 3X9

Received: 9 October 1996 / Accepted: 20 May 1997

Abstract. We present here the sequence and characterization of various minisatellite-like tandem repeat loci isolated from the genome of Atlantic salmon (*Salmo salar*). Their diversity of sequence and lack of core motifs common to minisatellites of other species suggest the presence of numerous and previously unidentified simple sequence repeat families in this salmonid. Evidence for their ubiquity was provided by screening of a salmon genomic library. Southern blot analysis of the phylogenetic distribution of a subset of the minisatellites found one sequence to be pervasive among vertebrates, others present only in Salmoninae or Salmonidae species, and one amplified only in Atlantic salmon. There is evidence for the positioning of microsatellite and minisatellite arrays in close proximity at many loci. Furthermore, one tandem repeat appears to have been inserted into the transposase coding region of a copy of the Tc1 transposon-like element recently identified in salmonids.

Key words: Repetitive DNA — Minisatellites — Microsatellites — Atlantic salmon — Tc1 transposon

Introduction

The high degree of mutability of minisatellite loci (Monckton et al. 1994), their dynamic expansion and contraction (Jeffreys et al. 1994), their possible role in recombination (Fletcher 1994), and the discovery of proteins able to selectively bind such simple sequence repeats (Collick and Jeffreys 1990; Wahls et al. 1991; Yee et al. 1991) have stimulated interest in the potential involvement of minisatellites in genomic turnover and change. Their study, however, has been hampered by a dearth of sequence data limited to only a small number of organisms. Greatest attention has been directed toward human minisatellite loci (Inglehearn and Cooke 1990; Armour and Jeffreys 1992). Additional data have been obtained from primates (Royle et al. 1994) and a small number of mammal (Kominami et al. 1988; Kelly et al. 1989; Kashi et al. 1990; Coppietiers et al. 1990; Brenig and Brem 1991; Kitazawa et al. 1994; Joseph and Sampson 1994), bird (Gyllensten et al. 1989), insect (Blanchetot 1990; Jacobsen et al. 1992; Paulsson et al. 1992), plant (Broun and Tanksley 1993; Winberg et al. 1993), and fish species.

In the case of fish, several minisatellite loci have been described for tilapia (*Oreochromis niloticus;* Harris and Wright 1995) and two loci from Atlantic salmon (*Salmo salar*) have been characterized (Bentzen and Wright 1993; Kvingedal 1994). In addition, Taggart and Ferguson (1990) reported the isolation of four minisatellite loci from *S. salar* and Prodöhl et al. (1994) reported five from brown trout (*S. trutta*), but provided no sequence data. With the exception of Kvingedal's (1994), these minisatellite loci were isolated by screening genomic libraries with the human-derived DNA fingerprinting probes 33.15 and 33.6 (Jeffreys et al. 1985). Brenner et al. (1993), using a random sequencing strategy, identified 16 minisatellite types from the pufferfish (*Fugu ru-*

* *Present address:* Department of Genetics, University of Pennsylvania School of Medicine, 505 CRB, 415 Curie Blvd., Philadelphia, PA 19104-6145, USA

Correspondence to: J. Goodier; e-mail: jgoodier@ncbi.nlm.nih.gov

bripes), but the sequence of only one of these is available.

Here we expand considerably upon these previous investigations by presenting sequence data from 12 different minisatellite-like repeat loci from Atlantic salmon. These repeats were isolated by means not dependent upon the use of mammalian minisatellite DNA probes. The minisatellite sequences show variable phylogenetic distribution, one being widely distributed among vertebrate species, while others are family-specific or species-specific. We have isolated some short tandem arrays in their entirety and have been able to examine the single-copy and repeat type DNA which flank them. Minisatellites appear to be associated with microsatellites at many loci and we discuss the relationship between these two classes of repetitive sequences in salmon.

Materials and Methods

Isolation and Sequencing of the Mini- and Microsatellites Loci. Genomic DNA from a single Atlantic salmon individual from Newfoundland was digested with restriction endonucleases and separated on agarose gels. Following staining with ethidium bromide, UV-fluorescent bands were excised, cloned into pUC18, and sequenced (Sequenase, USB). Most of the clones represented either monomer units of tandemly repeated DNA or internal portions of interspersed repetitive elements, and these have been described elsewhere (Goodier and Davidson 1993, 1994a,b). However, 20–40% of the clones represented unrelated background DNA, and among these were clones from DNA digested with *Bgl*II (designated SsBglIIU.20, SsBglIIL.6, SsBglIIL.27, SsBglIIU.48), *Pst*I (SsPstIL.26, SsPstIL.48), and *Sac*I (SsSacIU.44, SsSacIU.24) which contained micro- or minisatellite arrays.

An Atlantic salmon λEMBL3 genomic DNA library (a gift of J. Wright, Dalhousie University) was screened in duplicate by plaque lift hybridization using as radiolabeled probes both the DNA insert of plasmid clone SsPstIL.26 and poly(dA-dC) poly (dG-dT) DNA (Pharmacia). Membranes were prehybridized for 4 h and hybridized for 20 h in 6 × SSC (150 mM NaCl/15 mM sodium citrate, pH 7.6)/50% formamide/5× Denhardt's solution (0.02% bovine serum albumin/0.02% Ficoll/0.02% polyvinylpyrrolidone)/50mM phosphate buffer, pH 6.5/0.5% SDS at 42°C, and washed at high stringency (twice at room temperature in 2 × SSC for 20 min; 4 times in 0.1 × SSC, 0.1% SDS for 25 min at 55°C). Following autoradiography for 6 days, DNA from recombinant phage identified by both probes, SsPstIL.26 and poly (dA-dC)poly (dG-dT) DNA, was isolated by the method of Davis et al. (1980). DNA from each of the phage clones was digested with *Alu*I or *Hae*III, subcloned into the *Sma*I site of pUC18, and used to transform competent *E. coli* DH5α cells. Colonies of recombinant bacteria containing salmon minisatellite DNA sequence were identified by colony hybridization with the SsPstIL.26 probe. Plasmid DNA was isolated and sequenced.

Southern Blot Analyses. For zoo blot analyses, 5–10 μg of genomic DNA from selected species was digested with *Alu*I or *Pal*I, separated in a 1.0% agarose gel, and transferred to nylon membranes. Hybridization was at 37°C and otherwise as described above, followed by washing at low stringency (twice in 2 × SSC for 20 min at room temperature; twice in 2 × SSC/0.1% SDS, twice in 1 × SSC/0.1% SDS for 20 min each time at 50°C). Plasmid inserts from clones shown in Fig. 1 were used as probes, with the exception of SsSacIU.44, for which an oligonucleotide probe was synthesized (5′-GAGCTGACAACAGTCAGTCAGTGTTTTCCCTACATTA-3′) and end-labeled with [γ-^{32}P]ATP and polynucleotide kinase.

Results and Discussion

Isolation of Minisatellite Loci From Atlantic Salmon

We embarked on a project designed to characterize repetitive elements from the genome of Atlantic salmon. Our approach involved digesting salmon DNA with various restriction endonucleases, separating the fragments on agarose gels, and identifying bands visible upon staining with ethidium bromide. These bands were then isolated, cloned in a plasmid vector, and sequenced. The majority of clones obtained from the bands contained either monomer units of tandemly repeated DNA or sequence fragments internal to interspersed repetitive elements as described in Goodier and Davidson (1993, 1994a,b). Some recombinant clones, however, contained unrelated background DNA. Some of this DNA included minisatellite and microsatellite repeats (Figs. 1, 2, and and 8).

To detect additional interspersed tandem repeat loci, 1,620 plaques of an Atlantic salmon λEMBL3 genomic DNA library were screened with the minisatellite clone SsPstIL.26 (Fig. 1E). Following autoradiography, hybridization to almost 700 (43%) of the plaques was detectable. Assuming the length of a phage insert is 15 kb and that there is only one locus per insert, this would translate into approximately 85,000 loci per haploid genome. However, it is possible that the probe detects common core sequence in otherwise divergent repeats.

Duplicate membranes were then screened with radiolabelled poly (dA-dC) poly (dG-dT) DNA, but signal from positive plaques was much fainter and difficult to discern following a similar exposure time. Nevertheless, 46 positive recombinant plaques were detected and 24 of these had previously been detected by SsPstIL.26. DNA from 14 of these latter phage clones was isolated and confirmed to be positive for both SsPstIL.26 and poly (dA-dC) poly (dG-dT) probes by Southern hybridization. The DNA was then subcloned into pUC18. Further screening with the SsPstIL.26 probe permitted four repeat loci to be identified and sequenced: SsAluI16.70 (Fig. 1C), SsAluI17.5 (Fig. 2A), SsHaeIII14.20 (Fig. 2C), and SsAluI16.62 (Fig. 2E).

Unique Minisatellite Repeat Loci

Those sequences shown in Fig. 1 are portions of longer tandem arrays; their cloning was made possible by fortuitous mutations which generated restriction enzyme recognition sites in some monomer units. On the other hand, sequences presented in Figs. 2 and 8 include short

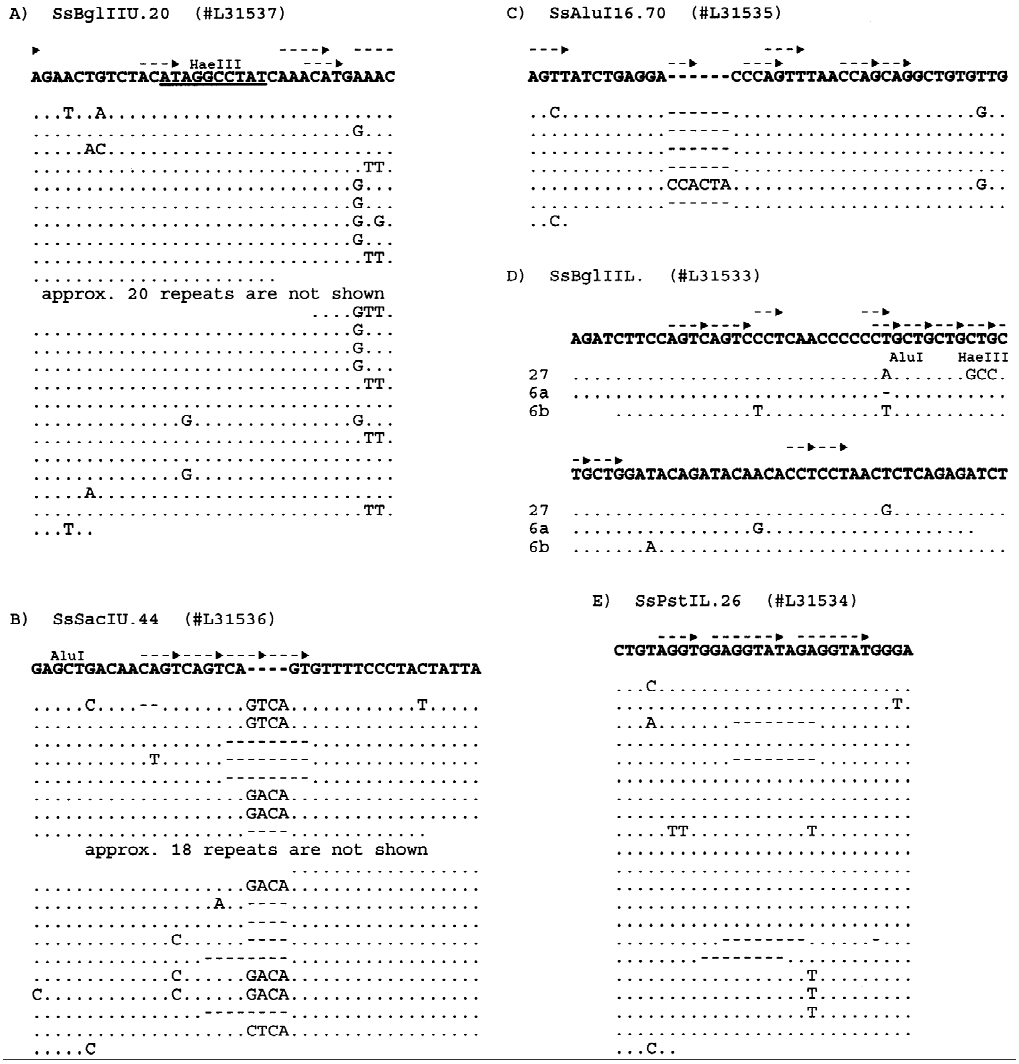

Fig. 1. Sequences of minisatellite arrays isolated from Atlantic salmon. *Dots* indicate that the sequence is identical to that of the consensus (in *bold*), and *dashes* indicate deletions. Direct repeats are shown above the consensus sequences and restriction endonuclease recognition sites are indicated. Palindromes are *underlined.* Sequence was not obtainable for stretches of SsBglIIU.20 **(A)** and SsSacIU.44 **(B).** *GenBank accession numbers* are shown beside each sequence name. SsBglIIL.6a and 6b are contiguous monomer sequences of the same clone. SsBglIIL.27 is sequence from a different cloned fragment **(D).**

tandem arrays with flanking sequence. It is possible that these may eventually be expanded into longer arrays through mechanisms of replication slippage or unequal crossing-over. Certain features of the different sequences presented here suggest modes of evolution for these repeats.

Repeat units of the minisatellites range in size from 28 nt (SsPstIL.26, Fig. 1E) to 59 nt (SsBglIIU.48, Fig. 2F). All contain internal short direct repeats 2–6 nt in length. These simple motifs may have been the precursors from which their full-length monomer units were elaborated by replication slippage and mutation (see below). Deletions within the monomers of arrays shown in Fig. 1 tend to be localized at these short repeats.

The repeats were examined for the presence of core sequences characteristic of many other eukaryote minisatellites. The presence of a core region (GGGCAGGAXG) within minisatellite monomer units having similarity to the *E. coli* ‘‘Chi’’ recombination signal prompted Jeffreys et al. (1985) to propose that these repeats are ‘‘hotspots’’ for recombination. Nakamura et al. (1987) have proposed a similar but more generalized motif, GXXGTGGGG. As summarized by Wright (1993), the motif C/GAGG is shared by many of the minisatellites isolated to date, but as the original sequences (33.6 and 33.15) of Jeffreys et al. (1985) were most often used as probes for their isolation, this is to be expected. Some similarity with these core motifs is evident in the sequences of SsAluI16.70 (CAGCAGGCTG) and SsPstIL.26 (GGTGGAGGTAT). The remaining Atlantic salmon tandem repeat sequences lack regions similar to that of Jeffreys et al. (1985) or to other unrelated

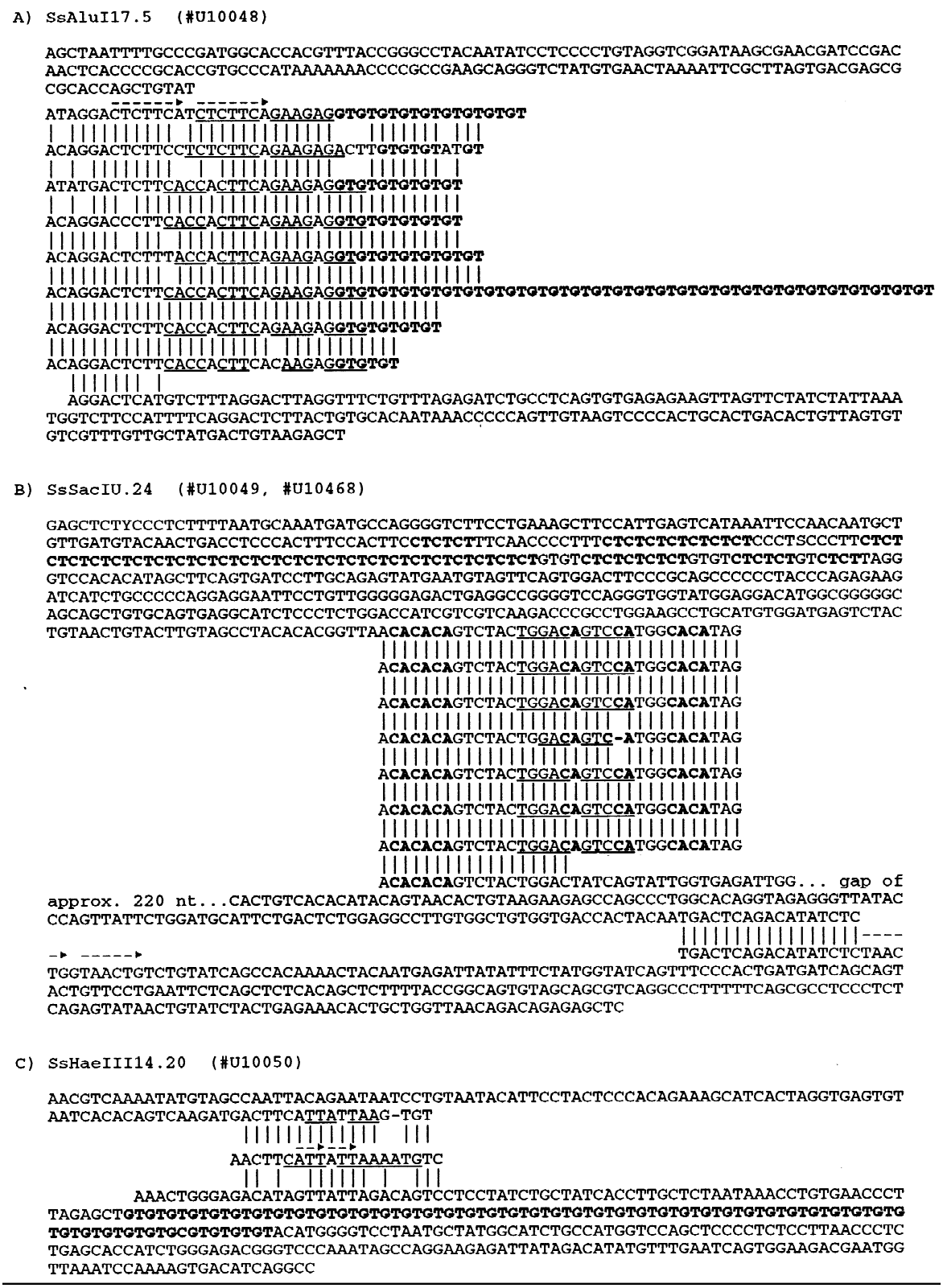

Fig. 2. Repetitive elements isolated from Atlantic salmon. Di- and trinucleotide arrays are shown in *bold*. Identical nucleotides of the consecutive monomer units are shown by ''|''. Direct repeats are shown above one monomer unit. Palindromes and quasipalindromes are *underlined*. Sequence was not obtained for internal stretches of SsSacIU.24 **(B)** and SsBglIIU.48 **(F).** SsHpaII.24 **(D)** is a subclone of pASBgl22 (previously described in Goodier and Davidson 1993). SsStuI6.62 **(G)** is sequence from the 3′ end of *a Stu*I fragment isolated from a phage clone of the Atlantic salmon genomic library. The fragment also contains a portion of a Tc1 transposon-like locus from Atlantic salmon (sequence not shown, see Goodier and Davidson 1994b). *GenBank accession numbers* are shown beside each sequence name.

Fig. 2. Continued.

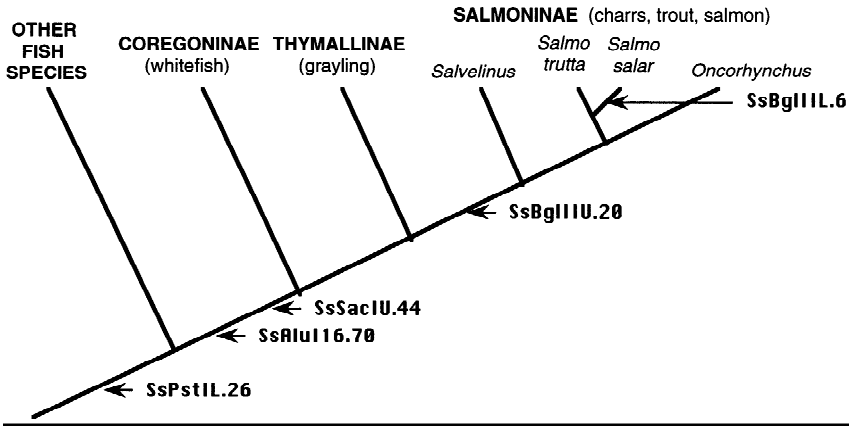

Fig. 3. Phylogenenetic tree of Salmonidae species showing probable times of amplification of the minisatellite sequences of Fig. 1.

minisatellites that have been described (see Vogt (1990) for review). A search of the EMBL and GenBank databases failed to detect sequences with similarity to any of the cloned repeats.

Phylogenetic Distribution of the Minisatellite Sequences

Having cloned minisatellite sequences from Atlantic salmon previously unidentified in other species, we wished to ascertain their phylogenetic distribution. Genomic DNAs from salmonid and other fish species, together with *Rana* sp., common murre (*Uria aalge*), and human, were digested with either *Alu*I or *Pal*I, restriction enzymes known to cut infrequently or not at all within the DNA sequences shown in Fig. 1. The digestion products were subjected to Southern blot analysis under low-stringency conditions using the inserts of the clones of Fig. 1 as probes. These results are summarized below and in Figs. 4–7. Based on this evidence, we have been able to ascertain the relative time of amplification of each minisatellite within the fish lineages (summarized in Fig. 3).

SsPstIL.26 is a pervasive repeat generating strong multilocus fingerprints in all Salmoninae (salmon, charr, trout) species (Fig. 4, lanes 1–8). Its presence was also detected for almost all other nonsalmonine species tested including *Rana* sp. and human (lanes 9–19), but hybridization signal was much weaker, indicating considerable sequence divergence. Among the greater than 35 bands visible for Atlantic salmon DNA digested with *Pal*I there was much intrapopulation pattern variation, even under conditions of high hybridization stringency (data not shown).

An oligonucleotide probe based on the sequence of SsSacIU.44 strongly hybridized to high molecular weight Atlantic salmon DNA cut with various restriction enzymes, suggesting that this repeat has been amplified in this species as very long tandem arrays at only a small number of loci (Fig. 5A, lane 1; B, lanes 1–12). This is consistent with fluorescent in situ hybridization studies on Atlantic salmon metaphase chromosomes which reveal one major area of hybridization when SsSacIU.44 is the probe (P. Moran, personal communication). Fragments seen in other species of Salmonidae (including

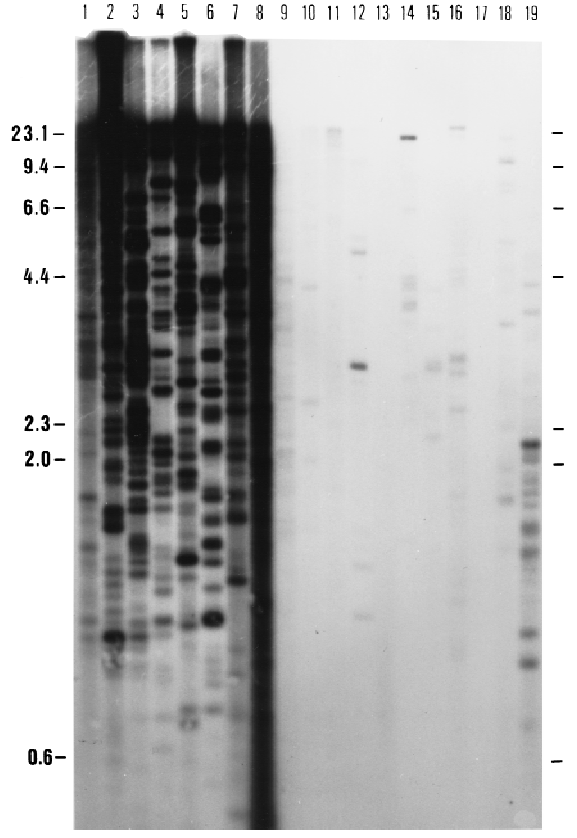

Fig. 4. Zoo blot analysis for SsPstIL.26. *Lanes 1*, Japanese charr; *2*, Atlantic salmon; *3*, brown trout; *4*, Arctic charr; *5*, brook charr; *6*, rainbow trout; *7*, masu salmon; *8*, chum salmon; *9*, grayling; *10*, whitefish; *11*, sea lamprey; *12*, dogfish shark (*Squalus acanthicus*); *13*, skate (*Raja* sp.); *14*, Atlantic sturgeon (*Ascipenser* sp.); *15*, winter flounder; *16*, *Rana* sp.; *17*, common murre; *18*, human; *19*, Atlantic cod. See text for additional species nomenclature. All DNA was digested with *Pal*I. Hybridization was at low stringency (see Materials and Methods).

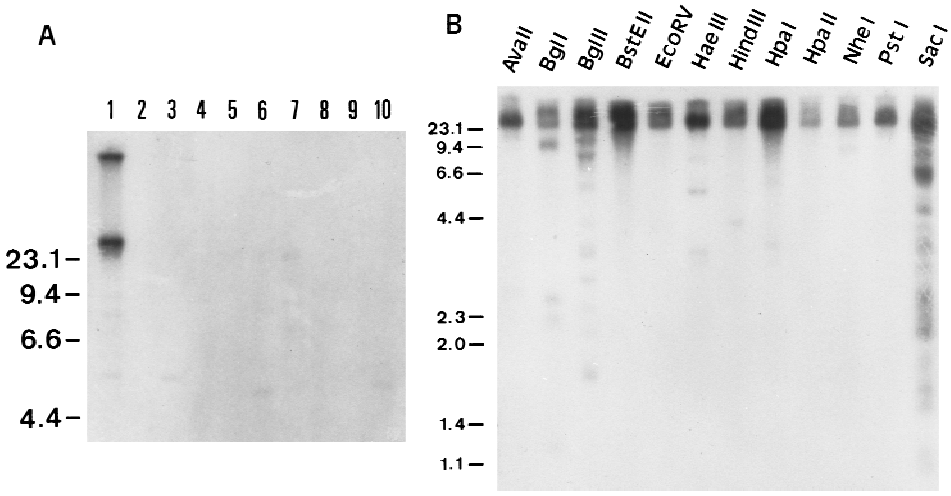

Fig. 5. **A** Zoo blot analysis for SsSacIU.44 using an oligonucleotide probe. Only species of the family Salmonidae are shown. DNAs from non-Salmonidae species were tested as in Fig. 4, but showed no hybridization of the probe. *Lanes 1,* Atlantic salmon; *2,* brown trout; *3,* Arctic charr; *4,* Japanese charr; *5,* brook charr; *6,* rainbow trout; *7,* masu salmon; *8,* chum salmon; *9,* Arctic grayling; *10,* whitefish. All DNA was digested with *Pal*I. **B** Southern blot analysis of Atlantic salmon DNA digested with various restriction endonucleases and probed with the oligonucleotide probe of SsSacIU.44.

whitefish, *Coregonus laveratus,* lane 10) were much fainter and shorter, indicating lower copy number and/or significant sequence divergence from that of Atlantic salmon (Fig. 5A, lanes 2–10).

Clone SsAluI16.70 yielded two to five bands for Salmoninae species (Fig. 6) and whitefish (not shown). Hybridization to DNA of non-Salmonidae fish species was detected as a smear, without discrete bands (data not shown). It is unclear, therefore, that related repeat sequences were being detected in these nonsalmonids; binding of the probe to DNA may have been nonspecific.

No signal was detected in whitefish, Arctic grayling (*Thymallus arcticus*), or nonsalmonid DNA when probed at low stringency with the insert of SsBglIIU.20 (data not shown), and this repeat must have arisen subsequent to the divergence of the subfamily Salmoninae from the rest of the Salmonidae species (Fig. 3). Fossil records are sparse, but *Eosalmo* of the Eocene period (54 to 38 million years ago) is the most primitive salmonine known (Stearley and Smith 1993; Nelson 1994).

Differential amplification of SsBg1IIU.20 repeats following salmonine speciation was indicated by the Southern blot analysis. When genomic DNA was digested with *Alu*I and probed at low stringency with SsBglIIU.20, strong signal was detected for Atlantic salmon (Fig. 7A, lane 5), as well as for some other species of the subfamily Salmoninae, including brown trout (*Salmo trutta,* lane 6), Arctic charr (*Salvelinus alpinus,* lane 7), and chum salmon (*Oncorhynchus keta,* lane 12, appearing as an intense smear). Weak signal was detected for rainbow trout (*Oncorhynchus mykiss,* lane 10) and masu salmon (*Oncorhynchus masou,* lane 11), but none was seen for Japanese charr (*Salvelinus leucomaenis,* lane 8) or brook charr (*Salvelinus fontinalis,* lane 9).

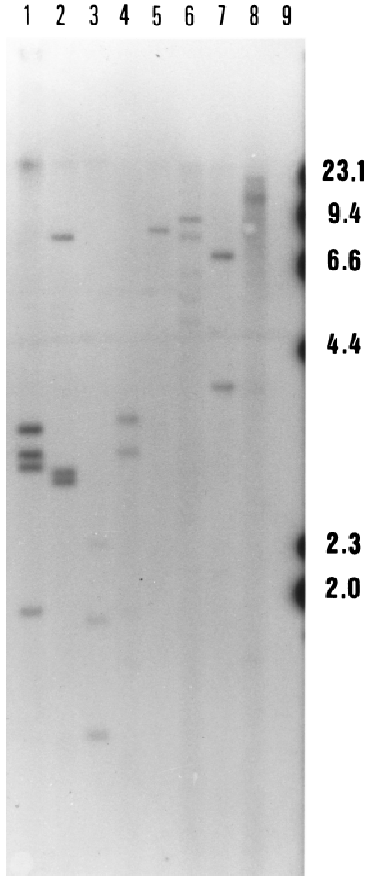

Fig. 6. Zoo blot analysis for SsAluI16.70. *Lanes 1,* Atlantic salmon; *2,* brown trout; *3,* Arctic charr; *4,* Japanese charr; *5,* brook charr; *6,* rainbow trout; *7,* masu salmon; *8,* chum salmon; *9,* grayling. All DNA was digested with *Pal*I. Less grayling DNA was loaded in *lane 9* compared with the other lanes. The markers shown are λ DNA cut with *Hin*dIII and ^{32}P-labeled.

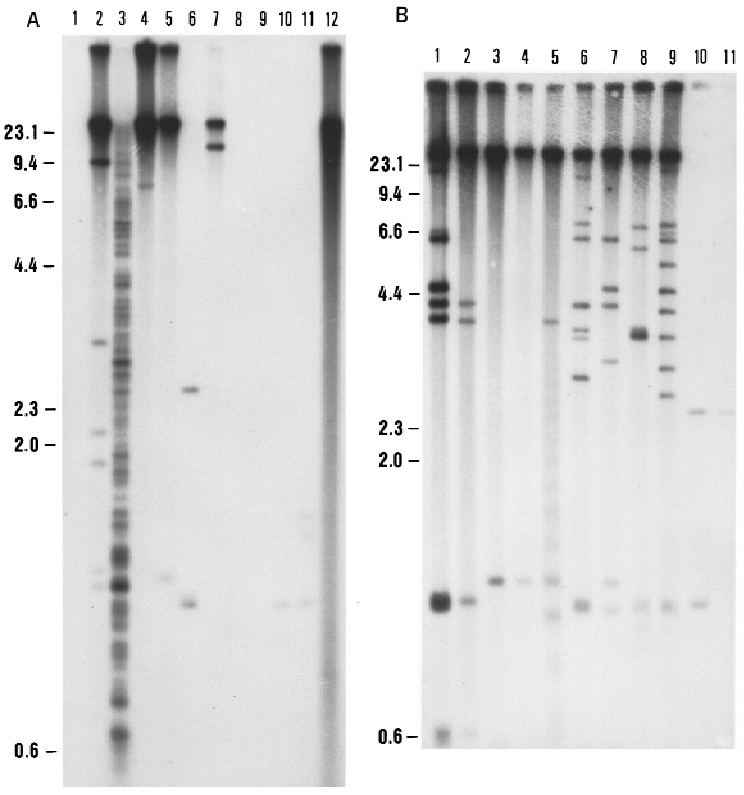

Fig. 7. Zoo blot analysis for SsBglIIU.20. **A** *Lanes 1–5,* Atlantic salmon, digested with *Pal*I, *Rsa*I, *Sau*3A, *Sac*I, and *Alu*I; *6,* brown trout; *7,* Arctic charr; *8,* Japanese charr; *9,* brook charr; *10,* rainbow trout; *11,* masu salmon; *12* chum salmon. DNA of lanes 5–12 was digested with *Alu*I. Only species of the subfamily Salmoninae are shown; no hybridization signal was detected for other species tested. Hybridization was at low stringency. **B** Analysis of genomic DNA from Atlantic salmon of Eastern Canada and Ireland. DNA was digested with *Alu*I and probed with SsBglIIU.20. Atlantic salmon, *lanes 1,* Little R., Newfoundland; *2,* Grand Codroy R., Newfoundland; *3–5,* LaHave, Nova Scotia; *6,7,* Owentogher R., Donegal, Ireland; *8,* Ray R., Donegal; *9,* Swilley R., Donegal; *10,* brown trout, Newfoundland; *11,* brown trout, Ireland. Hybridization was at high stringency. The markers shown are λ DNA cut with *Hin*dIII and ^{32}P-labeled.

Under high-stringency conditions, BglIIU.20 would not hybridize to non-Salmo species (data not shown). Hybridization of the SsBglIIU.20 probe to DNA from different Atlantic salmon individuals generated from one to over 10 polymorphic bands (Fig. 7B, lanes 1–9). On the other hand, in brown trout only two or fewer fragments were detected, each less than 3 kb in size (lanes 10 and 11), indicating amplification of this repeat sequence to higher copy number and longer arrays within the genome of *S. salar* following its divergence from *S. trutta*. The polymorphism revealed for *S. salar* by this probe is in part due to mutation hotspots within the monomer units of SsBglIIU.20. This is evident from sequence shown in Fig. 1A, and from the Southern blot analysis of Fig. 7A (lanes 1–5), where Atlantic salmon DNA has been digested with different restriction endonucleases and probed at low stringency with the insert of SsBglIIU.20. The *Pal*I recognition site, for example, is invariant, and the enzyme digests every monomer unit (lane 1). *Sau*3AI, on the other hand, cuts infrequently within the repeat array, generating a multibanded fingerprint pattern (lane 3).

A strong multibanded fingerprint pattern was produced when SsBglIIL.6 was hybridized to *S. salar* DNA cut with *Alu*I (not shown). Failure to detect the presence of this repeat within the genome of brown trout indicates specific amplification of this repeat in the *S. salar* lineage since it diverged from the common ancestor of the *Salmo* species. Although very faint hybridization bands were also detected for Arctic charr, sea lamprey, Atlantic cod (*Gadus morhua*), and winter flounder (*Pseudopleuronectes americanus*), the probe may have detected unrelated repeats containing some sequence similarity in these distantly related fish species. On the other hand, it is also possible that the progenitor sequence of SsBglIIL.6 was amplified to low copy numbers independently in several species.

The Proximity of Minisatellite Sequences with Other Repeat Types

Characterization of SsPstIL.48

Digestion of salmon DNA with *Pst*I produces two visible bands of 580 and 800 base pairs (bp) following electrophoresis and staining with ethidium bromide. Clone SsPstIL.48 (Fig. 8) was isolated from the 580-bp band. We have previously characterized DNA isolated from this band and shown it to represent an internal region from an interspersed repeat element (designated SALT1 in Goodier and Davidson 1994b) having similarity with the Tc1 class of transposons originally found in nematodes (Rosenzweig et al. 1983). In the case of clone SsPstIL.48, however, only 220 bp at the 3′ end of SsPstIL.48 was Tc1-related, being 65% identical in nucleotide sequence with a portion of the putative transposase coding region of a Tc1-like element (Tes1) isolated from Pacific hagfish (*Eptatretus stouti;* Heierhorst et al. 1992) and 58% identical with corresponding sequence from the Tc1-like SALT1 element.

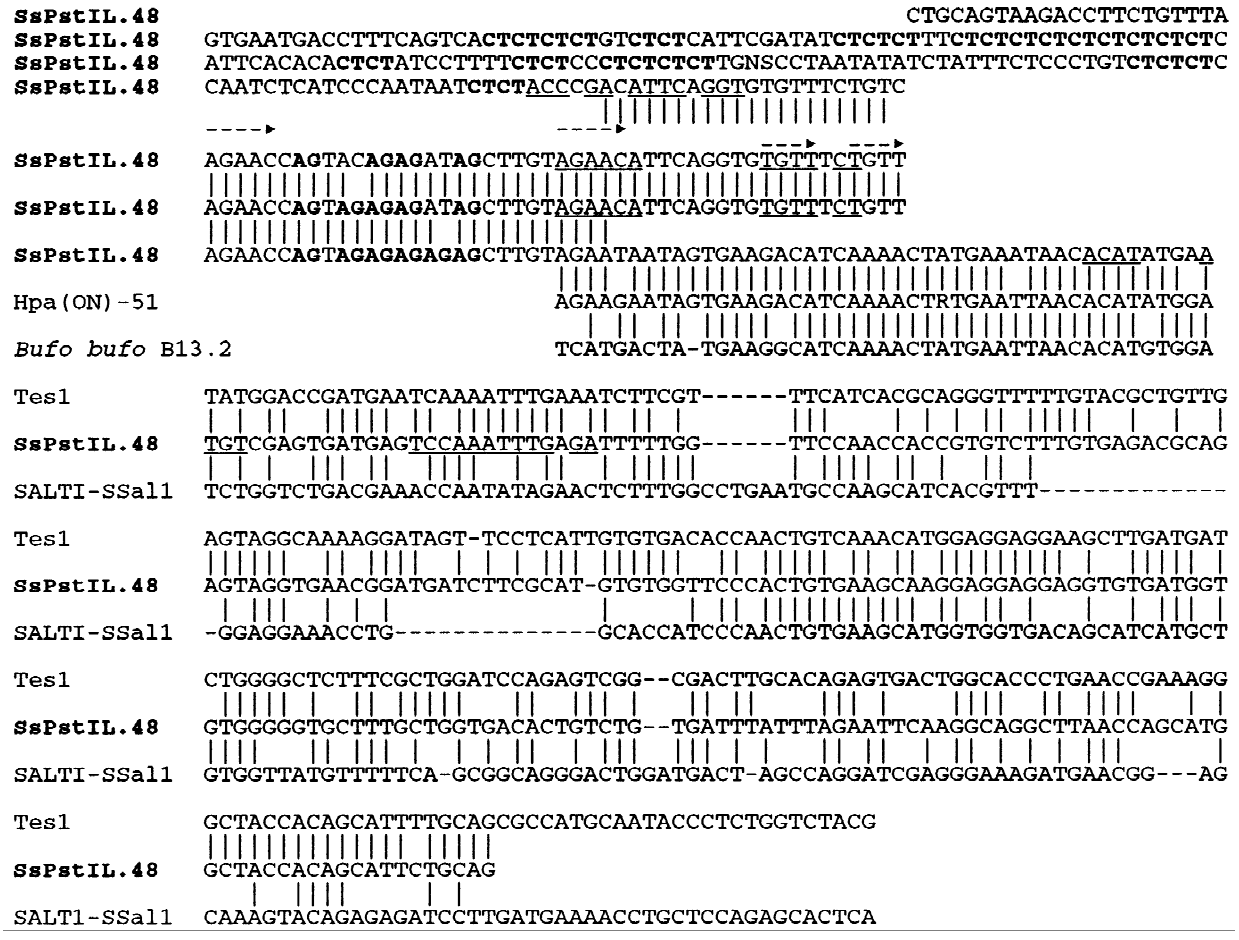

Fig. 8. Sequence of clone SsPstIL.48, isolated from Atlantic salmon. Direct repeats are indicated with *arrows* above the sequence and quasi-palindromes are *underlined.* Dinucleotide arrays are shown in *bold.* Other sequences shown are: Tes1, a portion of a Tc1 transposon-like element from hagfish (Heierhorst et al. 1992); SALT1-SSal1, a portion of a similar element from Atlantic salmon (Goodier and Davidson 1994b); Hpa (OK)-51 sequence flanking a SINE element present in the genome of coho salmon. *Bufo bufo* B13.2, sequence downstream of a microsatellite locus in European toad (Scribner KT, unpub, GenBank Acc #U05292).

Unrelated sequence appears to have been inserted into the SALT1-related portion of the SsPstIL.48 locus, perhaps by means of recombination between nonorthologous repeat loci. The 5′ end of the clone contains an imperfect $(GA)_n$ $(CT)_n$ microsatellite followed by three 48-bp tandem repeats. The association of simple tandem repeats with transposons has not been previously reported, although in mammals it is not unusual to find microsatellite and minisatellite sequences near or within SINE retroposons (Arcot et al. 1995, and references therein) or the long terminal repeats of retroviral or retrotransposon-like elements (Armour et al. 1989; Rogaev 1990; Kelly 1994).

Unexpectedly, 45 nt between the last tandem repeat and the start of the SALT1 homologous region of SsPstIL.48 are 93% identical with 3′ flanking sequence of a SINE element locus of coho salmon (*Oncorhynchus kisutch*) described by Murata et al. (1993). These 45 nt are also 87% identical with sequence located 65 bp from the start of a microsatellite locus of the European toad *Bufo bufo* (Scribner et al. 1994; GenBank Acc. #U05292). The significance of this motif within the complicated repeat environment of the SsPstIL.48 locus is unknown.

The Association of Minisatellite and Microsatellite Arrays

The results of screening the salmon genomic λEMBL3 library with the SsPstIL.26 and poly (dA-dC) probes (see above), together with sequence presented in Figs. 2 and 8 suggest that microsatellite dinucleotide arrays in salmon are associated with short, tandem repeats at many loci. Bentzen and Wright (1993) also identified a microsatellite near a minisatellite array in Atlantic salmon. Dinucleotide arrays isolated in the present study were of the form $(CA)_n \cdot (GT)_n$ or $(GA)_n \cdot (CT)_n$.

Evidence for the association of microsatellites with minisatellite sequences in other organisms is anecdotal and involves, for the most part, minisatellite repeat units harboring short stretches of simple sequence (Kashi et al. 1990; Armour et al. 1992; Iwasaki et al. 1992). Some researchers (Levinson and Gutman 1987; Schlötterer and

Tautz 1992) have proposed that the major mechanism involved in the evolution of simple sequence DNA is slipped-strand mispairing at replication. Initial amplification by replication strand slippage will expand the substrate and hence the chance that slipped-strand mispairing will occur in the future. Single nucleotide substitutions, deletions, or insertions may convert one simple sequence form to another. Wright (1994) has also presented a case for microsatellites as the progenitors of minisatellites. The occurrence of simple sequence motifs within the monomer units of minisatellite arrays described in the present study lends credence to this scenario. It is not unreasonable to suppose that the repeat units of SsAluI17.5 (Fig. 2A), for example, have been amplified from $(CA)_n \cdot (GT)_n$ dinucleotides. It is less clear that the long microsatellite arrays shown in Fig. 2(B,C, and E) and Fig. 8 have been involved in the birth of their minisatellite neighbors. In this case, one might expect to see a zone of transition through which microrepeats are gradually elaborated. Rather, we did not find examples of dinucleotide arrays contiguous with minisatellite repeats, the two being separated instead by unique sequence. While slipped-strand mispairing may be involved in the generation of short sequence repeats, unequal crossing over or sister chromosome exchange may act to bring unrelated repeat types into juxtaposition, a phenomenon perhaps mediated by the dinucleotide array itself. The power of $(CA)_n \cdot (GT)_n$ dinucleotide tracts to promote homologous recombination, possibly through the ability to form left-handed Z-form DNA, has been demonstrated (Treco and Arnheim 1986; Bullock et al. 1986; Wahls et al. 1990).

Palindromes and quasipalindromes (repeats separated by a short spacer) are found near and overlapping the junctions of many of the minisatellite-type arrays shown here. The minisatellite array from Atlantic salmon described by Bentzen and Wright (1993) also has imperfect inverted repeats near its 5′ (CGACCAGAAGGTCG) and 3′ (AGAAAACTACACCTAGCAGTTGTCT) ends, and palindromes have been reported within the monomer units of tandemly repeated DNA from other species, including fish (Ekker et al. 1992; Kubota et al. 1993; Goodier and Davidson 1994a). Some studies have detected palindromes at sites of recombination in satellite DNA (Kiyama et al. 1987; Plohl and Ugarković 1994). Reed et al. (1994) report the presence of palindromes at points of transition between repeat types within *Nasonia vitripennis;* as reviewed by these authors, palindromes have been postulated as a source of genetic instability in eukaryotes, having been associated with increased recombination, DNA amplification, and local sequence excision. The detection of palindromes within and near so many salmon minisatellite-like repeat arrays is of unknown significance. Nevertheless, we speculate that they could have been involved in the evolution of these repeat loci, perhaps by serving as recognition sites for proteins which mediate minisatellite stability.

Acknowledgments. We thank S. Bartlett, V. Drover, C. Hurst, and C. McGowan for helpful discussion and L. Sturge for editorial assistance. DNA samples were provided by S. Hartley, V. Pepper, R. Phillips, J. Taggart, and F. Yamazaki. This research was supported by a Natural Sciences and Engineering Research Council of Canada Strategic Grant to W.S.D.

References

Arcot SS, Wang Z, Weber JL, Deininger PL, Batzer M (1995) Alu repeats: a source for the genesis of primate microsatellites. Genomics 29:136–144

Armour JAL, Jeffreys A (1992) Recent advances in minisatellite biology. FEBS Lett 307:113–115

Armour JAL, Wong Z, Wilson V, Royle NJ, Jeffreys AJ (1989) Sequences flanking repeat arrays of human minisatellites: association with tandem and dispersed repeat elements. Nucleic Acids Res 17:4925–4935

Armour JAL, Crosier M, Jeffreys AL (1992) Human minisatellite alleles detectable only after PCR amplification. Genomics 12:116–124

Bentzen P, Wright JM (1993) Nucleotide sequence and evolutionary conservation of a minisatellite variable number tandem repeat cloned from Atlantic salmon, *Salmo salar*. Genome 36:271–277

Blanchetot A (1990) A *Musca domestica* satellite sequence detects individual polymorphic regions in insect genome. Nucleic Acids Res 19:929–932

Brenig B, Brem G (1991) Human VNTR sequences in porcine HTF-islands. Experientia 58:39–49

Brenner S, Elgar G, Sandford R, Macrea A, Venkatesh B, Aparicio S (1993) Characterization of the pufferfish (*Fugu*) genome as a compact model vertebrate genome. Nature 266:265–268

Broun P, Tanksley SD (1993) Characterization of tomato DNA clones with sequence similarity to human minisatellites 33.6 and 33.15. Plant Mol Biol 23:231–242

Bullock P, Miller J, Botchan M (1986) Effects of poly [d(pGpT).d(pApC)] and poly [d(pCpG).d(pCpG)] repeats on homologous recombination in somatic cells. Mol Cell Biol 6:3948–3953

Collick A, Jeffreys AJ (1990) Detection of a novel minisatellite-specific DNA-binding protein. Nucleic Acids Res 18:625–629

Coppietiers W, Van de Weghe A, Depicker A, Bouquet Y, Vaneveren A (1990) A hypervariable pig DNA fragment. Anim Genet 21:29–38

Davis RW, Botstein D, Roth JR (1980) A manual of genetic engineering: advanced bacterial genetics. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY

Ekker M, Andreas R, Westerfield M (1992) Identification of two families of satellite-like repetitive DNA sequences from zebra fish (*Brachydanio rerio*). Genomics 13:1169–1173

Fletcher HL (1994) Possible loss of length conservation and reciprocity during recombination or conversion in tandem arrays. Genetics 138:511–518

Goodier JL, Davidson WS (1993) A repetitive element in the genome of Atlantic salmon, *Salmo salar*. Gene 121:237–242

Goodier JL, Davidson WS (1994a) Characterization of a repetitive element detected by *Nhe*I in the genomes of *Salmo* species. Genome 37:639–645

Goodier JL, Davidson WS (1994b) Tc1 transposon-like sequences are widely distributed in salmonids. J Mol Biol 241:26–34

Gyllensten UB, Jakobsson S, Temrin H, Wilson AC (1989) Nucleotide

sequence and genomic organization of bird minisatellites. Nucleic Acids Res 17:2203–2214

Harris AS, Wright JM (1995) The nucleotide sequence and genomic organization of cichlid fish minisatellites. Genome 38:177–184

Heierhorst J, Lederis K, Richter D (1992) Presence of a member of the Tc1-like transposon family from nematodes and *Drosophila* within the vasotocin gene of a primitive vertebrate, the Pacific hagfish *Eptatretus stouti.* Proc Natl Acad Sci USA 89:6798–6802

Inglehearn CF, Cooke HJ (1990) A VNTR immediately adjacent to the human pseudoautosomal telomere. Nucleic Acids Res 18:471–476

Iwasaki H, Stewart PW, Dilley WG, Holt MS, Steinbrueck TD, Wells SA Jr, Donis-Keller H (1992) A minisatellite and a microsatellite polymorphism within 1.5 kb at the human muscle glycogen phosphorylase (PYGM) locus can be amplified by PCR and have combined informativeness of PIC 0.95. Genomics 13:7–15

Jacobsen JW, Guo W, Hughes CR (1992) A *Drosophila* minisatellite contains multiple *chi* sequences. Insect Biochem Mol Biol 22:785–792

Jeffreys AJ, Wilson V, Thein SL (1985) Hypervariable ''minisatellite'' regions in human DNA. Nature 314:67–73

Jeffreys AJ, Tamaki K, Macleod A, Monckton DG, Neil DL, Armour JAL (1994) Complex gene conversion events in germline mutation at human minisatellites. Nat Genet 6:136–145

Joseph SS, Sampson J (1994) Identification, isolation, and characterization of canine minisatellite sequences. Anim Genet 25:307–312

Kashi Y, Iraqi F, Tikochinski Y, Ruzitsky B, Nave A, Beckmann JS, Friedmann A, Soller M, Gruenbaum Y (1990) $(TG)_n$ uncovers a sex-specific hybridization pattern in cattle. Genomics 7:31–36

Kelly RG (1994) Similar origins of two mouse minisatellites within transposon-like LTRs. Genomics 24:509–515

Kelly R, Bulfield G, Collick A, Gibbs M, Jeffreys AJ (1989) Characterization of a highly unstable mouse minisatellite locus: evidence for somatic mutation during early development. Genomics 5:844–856

Kitazawa T, Kominami R, Tanaka R, Wakabayashi K, Nagoa M, Signer EN, Gu F, Gustavsson I, Andersson L, Jeffreys AJ (1994) A pseudoautosomal minisatellite in the pig. Mamm Genome 5:48–51

Kiyama R, Okumura K, Matsui H, Burns GAP, Kanda N, Oishi M (1987) Nature of recombination involved in excision and rearrangement of human repetitive DNA. J Mol Biol 198:589–598

Kominami R, Mitani K, Muranmatsu K (1988) Nucleotide sequence of a mouse minisatellite DNA. Nucleic Acid Res 16:1197

Kubota S, Kuro-o M, Muzuna S, Kohno S (1993) Germ line-restricted, highly repeated DNA sequences and their chromosomal localization in a Japanese hagfish (*Eptatretus okinoseanus*). Chromosoma 102:163–173

Kvingedal AM (1994) Characterization of the 5′ region of the Atlantic salmon (*Salmo salar*) transferrin-encoding gene. Gene 150:335–339

Levinson G, Gutman GA (1987) Slipped-strand mispairing: major mechanism for DNA sequence evolution. Mol Biol Evol 4:203–221

Monckton DG, Neumann R, Guram T, Fretwell N, Tamaki K, Macleod A, Jeffreys AJ (1994) Minisatellite mutation rate variation associated with a flanking DNA sequence polymorphism. Nat Genet 8: 162–169

Murata S, Takasaki N, Saitoh M, Okada N (1993) Determination of the phylogenetic relationships among Pacific salmonids by using short, interspersed elements (SINEs) as temporal landmarks of evolution. Proc Natl Acad Sci USA 90:6995–6999

Nakamura Y, Leppert M, O'Connel P, Wolff R, Holm T, Culver M, Martin C, Fujimoto E, Hoff M, Kumlin E, White R (1987) Variable number of tandem repeat (VNTR) markers for human gene mapping. Science 235:1616–1628

Nelson JS (1994) Fishes of the world. 3rd ed. John Wiley, New York

Paulsson G, Hoog C, Bernholm K, Wieslander L (1992) Balbiani ring gene in *Chironomous tentans:* sequence organization and dynamics of a coding minisatellite. J Mol Biol 225:349–361

Plohl M, Ugarković D (1994) Analysis of divergence of *Alphitobius diaperinus* satellite DNA—roles of recombination, replication slippage and gene conversion. Mol Gen Genet 242:297–304

Prodöhl PA, Taggart JB, Ferguson A (1994) Cloning of highly variable minisatellite DNA single-locus probes for brown trout (*Salmo trutta* L.) from a phagemid library. In: Beaumont AR (ed) Genetics and evolution of aquatic organisms. Chapman and Hall, London, pp 263–270

Reed KM, Beukeboom LW, Eickbush DG, Werren JH (1994) Junctions between repetitive DNAs on the PSR chromosome of *Nasonia vitripennis:* association of palindromes with recombination. J Mol Evol 38:352–362

Rogaev EI (1990) Simple human DNA-repeats associated with genomic hypervariability, flanking the genomic retroposons and similar to retroviral sites. Nucleic Acids Res 18:1879–1885

Rosenzweig B, Liao L, Hirsh D (1983) Sequence of the *C. elegans* transposable element Tc1. Nucleic Acids Res 11:4201–4209

Royle NJ, Baird DM, Jeffreys AJ (1994) A subterminal satellite located adjacent to telomeres in chimpanzees is absent from the human genome. Nat Genet 6:52–56

Schlötterer C, Tautz T (1992) Slippage synthesis of simple sequence DNA. Nucleic Acids Res 20:211–215

Scribner KT, Arntzen JW, Burke T (1994) Comparative analysis of intra- and interpopulation genetic diversity in *Bufo bufo,* using allozyme, single-locus microsatellite, minisatellite and multilocus minisatellite data. Mol Biol Evol 11:737–748

Stearley RF, Smith GR (1993) Phylogeny of the Pacific trouts and salmons, *Oncorhynchus,* and genera of the family Salmonidae. Trans Am Fish Soc 122:1–33

Taggart JB, Ferguson A (1990) Hypervariable minisatellite DNA single locus probes for the Atlantic salmon (*Salmo salar*). J Fish Biol 37:991–993

Treco D, Amheim N (1986) The evolutionarily conserved repetitive sequence $d(TGAC)_n$ promotes reciprocal exchange and generates unusual recombinant tetrads during yeast meiosis. Mol Cell Biol 6:3934–3947

Vogt P (1990) Potential genetic functions of tandem repeated DNA sequence blocks in the human genome are based on a highly conserved ''chromatin folding code.'' Hum Genet 84:301–336

Wahls WP, Wallace LJ, Moore PD (1990) The Z-DNA motif $d(TG)_{30}$ promotes reception of information during conversion events while stimulating homologous recombination in human cells in culture. Mol Cell Biol 10:785–793

Wahls WP, Swenson G, Moore PD (1991) Two hypervariable minisatellite DNA binding proteins. Nucleic Acids Res 19:3269–3274

Winberg BC, Zhou Z, Dallas JF, McIntyre CL, Gustafson JP (1993) Characterization of minisatellite sequences from *Oryza sativa.* Genome 36:978–983

Wright JM (1993) DNA fingerprinting of fishes. In: Hochachka P, Mommsen T (eds) Biochemistry and molecular biology of fishes, vol. 2. Elsevier, New York, pp 57–91

Wright JM (1994) Mutation of VNTRs: are minisatellites the evolutionary progeny of microsatellites? Genome 37:345–347

Yee HA, Wong AKC, Van der Sande JH, Rattner JB (1991) Identification of novel single-stranded $d(CT)_n$ binding proteins in several mammalian species. Nucleic Acids Res 19:949–953