Molecular Biology and Evolution

Novel SINE Families from Salmons Validate *Parahucho* (Salmonidae) as a Distinct Genus and Give Evidence that SINEs Can Incorporate LINE-related 3′-Tails of Other SINEs

Vitaliy Matveev, Hidenori Nishihara, and Norihiro Okada

Faculty of Bioscience and Biotechnology, Department of Biological Sciences, Tokyo Institute of Technology, Yokohama, Japan

Short interspersed elements (SINEs) constitute a group of retroposons propagating in the genome via a mechanism of reverse transcription, in which they depend on the enzymatic machinery of long retroposons (LINEs). Over 70 SINE families have been described to date from the genomes of various eukaryotes. Here, we characterize two novel SINEs from salmons (Actinopterygii: Salmonoidei). The first family, termed SlmI, was shown to be widespread among all genera of the suborder. These SINEs have a $tRNA^{Leu}$-related promoter region at their 5′-end, a unique central conserved domain with a subfamily-specific region, and an end with RSg-1-LINE-derived 3′-terminus preceding the A/T-rich tail. The same LINE-related segment is also shared by two other salmonid SINEs: HpaI and OS-SINE1. The structural peculiarities and overall sequence identity of the SlmI 3′-terminus suggest that it has been acquired from HpaI SINEs but not directly from the partner LINE. This region plays a crucial role in the process of retrotransposition of short interspersed elements, and the case of its SINE-to-SINE transmission is the first recorded to date. Possible scenarios and potential evolutionary implications of the observed interaction between short retroposons are discussed. Apart from the above, we found a copy of the SlmI SINE in the GenBank entry for the blood fluke, *Schistosoma japonicum* (Trematoda: Strigeiformes)—a trematode causing one of the most important human helminth infections, with its genome known to host other groups of salmonoid retroposons. In the present article, we suggest our views with regard to possible ways in which such an intensive horizontal transfer of salmonoid retroposons to the schistosomal genome occurs. The second novel SINE family, termed SlmII, originates from one of the SlmI subfamilies, with which it shares the same tRNA-related region, central domain, and a part of RSg-1-derived segment, but has a different 3′-tail of unidentified origin. Its distribution among salmonids validates *Parahucho* (Japanese huchen) as a distinct monotypic genus.

Introduction

Compared to protein-coding sequences that cover only small part of the entire genome in eukaryotes (e.g., ca. 1.5% in the case of mammals), a much greater portion of it—up to 45% in human and mouse—is occupied by various repetitive transposable elements (Lander et al. 2001; Waterston et al. 2002), which are able to multiply and disperse their own copies throughout the genome. The most abundant of them—long (LINEs) and short (SINEs) interspersed elements (retroposons)—belong to the class of retroelements, and in human genomic DNA they account for 20% and 13% of all sequences, respectively (Lander et al. 2001), repeating hundreds to millions of times. In the genome they propagate via RNA and cDNA intermediates by means of reverse transcription (Singer 1982; Rogers 1985; Weiner et al. 1986; Brosius 1991).

Short retroposons belonging to different families may vary greatly in length (ca. between 70 and 500 bp) and structure and yet share important common features. Unlike LINEs, transcribed by RNA polymerase II (Pol II), all SINEs possess an internal promoter for Pol III (Okada 1991*a*, 1991*b*; Okada and Ohshima 1995), which is usually derived from transfer RNA genes. In rare cases, however, the promoter region can be incorporated from 7SL RNA or 5S rRNA: for review and references, see Kramerov and Vassetzky (2005), Ohshima and Okada (2005).

Although integration of new copies of short retroposons into the new genomic locations occurs via a mechanism of reverse transcription, SINEs do not encode any proteins themselves, and in this process they depend entirely on the enzymatic machinery of autonomous LINEs. Recognition of SINE transcripts by LINE proteins is necessary for their reverse transcription and reintegration into the new loci, and it is guaranteed either by the common 3′-tail shared by the SINE and its partner LINE of stringent type or by the presence of the poly(A) tail in the case of the relaxed type of LINEs (Okada et al. 1997; Jurka 1997; Kajikawa and Okada 2002; Dewannieux et al. 2003). It has hitherto been thought that the common conserved 3′-domain is incorporated by SINEs directly from long retroposons. Here, we report the first case of SINE-to-SINE transmission of the LINE-related 3′-tail and discuss potential evolutionary implications of our finding.

More than 70 SINE families have been described to date from the genomes of various eukaryotes (Kramerov and Vassetzky 2005; Ohshima and Okada 2005). Based on diagnostic mutations of their sequences, many of them are further subdivided into subfamilies (e.g., Jurka and Smith 1988). In their distribution, most SINE families are restricted to more or less uniform groups of organisms not higher in rank than order, though some are characterized by the presence of important common features and are therefore united into superfamilies.

Vertebrate genomes contain different numbers of SINE families, and salmons—a group of tetraploid fishes from the suborder Salmonoidei (order Salmoniformes)—are one of those groups where their number is particularly high: 5 of 12 families of ray-finned fish SINEs are described from their genomes. To a certain extent such phenomenon can be explained by the fact that of all the fish taxa salmons are undoubtedly among the best studied, mostly due to their commercial importance and significant role they play in various ecosystems. However, despite intensive study, many questions of their biology and evolutionary history remain unanswered. Even the high-level classification of

Key words: Slm SINEs, *Parahucho perryi*, salmon, 3′-tail, horizontal transfer, *Schistosoma*.

E-mail: nokada@bio.titech.ac.jp.

Mol. Biol. Evol. 24(8):1656–1666. 2007
doi:10.1093/molbev/msm083
Advance Access publication April 29, 2007

Table 1
Primers Used in the Present Study

Primer	Primer Sequence	Annealing Temperature, °C
M4 (direct)	5′-GTTTTCCCAGTCACGAC-3′	55
RV (reverse)	3′-CAGGAAACAGCTATGAC-5′	
MP-F1 (direct)	5′-GGTAGCCTAGTGGTTAGAGCGTTG-3′	55
MP-R1 (reverse)	3′-AATTCTTATTTTCAATGACGGCCTA-5′	
SlmI-F1 (direct)	5′-GTGCTAGAGGCGTCACTACAGW-3′	51
SlmI-R1 (reverse)	3′-ACTAGGCAAGTCAGTTAAGAACAAAT-5′	
SlmI-F (direct)	5′-TCGAGTGGCGCAGCGGTCTAA-3′	69
SlmI-R (reverse)	3′-ACAATGACGGCCTACACCGGC-3′	
SlmINPGH-F (direct)	5′-ATTTACAATGACAAGCARAG-3′	52
SlmINPGH-R (reverse)	3′-TGTGATATGTGGTTGTTTTA-5′	
SI-F (direct)	5′-GAGTGGCGCAGCGGTCTAAGGCA-3′	69
SI-R (reverse)	3′-TACCCCGGCCAAACCCTCCCCTAAC-5′	
SlmII-F (direct)	5′-AGCTGTGCCACCAGAGACTC-3′	62
SlmII-R (reverse)	3′-ATTTCGTGGTATCCAATTGTTAGTA-5′	

salmons, e.g., the taxonomic status of graylings and intergeneric relationships of Salmonidae, is a subject of ongoing controversy. Our data shed light on some of these uncertainties.

The suborder Salmonoidei is subdivided into three families: Salmonidae (salmon, trout and char), Thymallidae (grayling) and Coregonidae (whitefish and ciscoes). Often all three are regarded as subfamilies of the single family Salmonidae.

Two of the salmonoid SINEs—HpaI (Kido et al. 1991) and AvaIII (Kido et al. 1994; Kido et al. 1995)—belong to the superfamily of CORE-SINEs and are widespread among Salmonoidei. Others include OS-SINE1 (from the superfamily of DeuSINEs), so far found in two genera: *Oncorhynchus* and *Salmo* (Nishihara et al 2006); FokI—a family restricted to one single genus—*Salvelinus*, or chars (Kido et al. 1991); extremely similar SmaI SINEs are found in two distant lineages: two species of Pacific salmon from one side (chum and pink salmon) and whitefishes (Coregonidae) from the other. The distribution pattern of the latter is most likely a result of horizontal transfer from the genome of one of the whitefish species to the common ancestor of chum and pink salmon (Hamada et al. 1997).

In the present article, we describe two novel families of tRNA-related SINEs. One of them appears to be widespread among the genomes of all Salmonoidei and was also found in blood flukes—*Schistosoma japonicum* (Trematoda: Strigeiformes). This trematode causes one of the most important human helminth infections, and its DNA was recently reported to host almost all known salmonoid retroposons (Melamed et al. 2004). In this article, we suggest our views in regards to possible ways of such intensive horizontal transfer of salmonid retroposons to schistosomal genome.

Material and Methods

Tissue Samples and DNA Extraction

Several tissue samples from taimen (*Hucho taimen*), lenok (*Brachymystax lenok*), and grayling (*Thymallus arcticus*) have been kindly provided by Dr. S. Alexeev (Kol'tsov Institute of Developmental Biology, Russian Academy of Sciences, Moscow, Russia); Inconnu (*Stenodus leucichthys*) tissues have been received from Dr. D. Politov (Vavilov Institute of General Genetics, RAS, Moscow, Russia). The fish species of the order Salmoniformes used in our study are listed in table S1 (see online Supplementary Material).

Genomic DNA was isolated from ethanol-preserved tissues (muscles) by phenol-chloroform extraction following the treatment of tissue homogenates with proteinase K (Blin and Stafford 1976).

Construction and Screening of the Genomic Libraries, Sequencing of DNA Clones

As a part of our experiments on MP-subfamily (Kido et al., 1994) of salmonoid HpaI SINEs, we constructed a genomic library from the DNA of Bonneville whitefish, *Prosopium spilonotus*. To do this, the total DNA was digested with *Eco*RI restriction endonuclease. The digests were fractionated by centrifugation (25,000 rpm, vacuum, 15°C, 15 h) in sucrose gradient (10%–40%), and those ranging from 1 to 2 kb in length were ligated into *Eco*RI-digested vector pUC19. The products of ligation reaction were transformed into *Escherichia coli* DH5-α cells and transferred to cellulose membranes for further screening. We generated a series of polymerase chain reaction (PCR) products from the genomic DNA of the same species using internal primers for MP consensus sequence: MP-F1 and MP-R1 (table 1). These products were labeled with [α-^{32}P]dCTP and used as a probe for screening of the genomic library.

Hybridization of the membrane-bound clones with the probe was performed at 42°C overnight in a buffer of 50% formamide, 6× SSC [0.15 M NaCl, 0.015 M trisodium citrate, pH 7], 1% sodium dodecyl sulfate (SDS), 1.7× Denhardt's solution [0.02% Ficoll 400, 0.02% polyvinylpyrrolidone and 0.02% bovine serum albumin], and 100 μg/ml herring sperm DNA. After washing in a solution of 2× SSC and 1% SDS at 50°C for 10 min, the membranes were autoradiographed with X-ray film for 48–72 h at –50°C. Positive plasmid clones were isolated, and their inserts were sequenced with M4 and RV plasmid universal

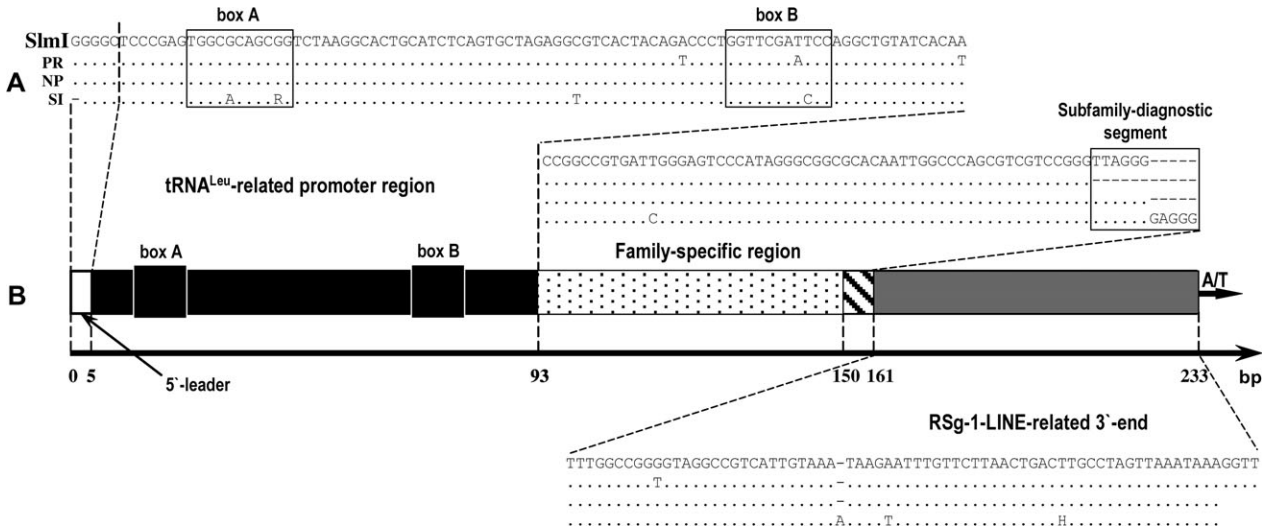

FIG. 1.—(*A*) SlmI consensus sequence aligned with those of PR, NP and SI subfamilies. RNA Pol III promoter boxes A and B, and subfamily-diagnostic segment are boxed. (*B*) General structure of SlmI SINEs. Transfer RNA-related region is shown in black; the family-specific central domain is dotted, and its subfamily-diagnostic segment is hatched; the LINE-related 3′-tail is shown in gray, and the variable A/T-rich tail is denoted with black arrow.

primers (Takara, Japan) and/or MP internal primers—MP-F1 and MP-R1 (table 1)—by the dideoxy chain termination method (Sanger et al., 1977) with an automated sequencer ABI PRISM 3100 Genetic Analyzer (Applied Biosystems/Hitachi, USA/Japan).

PCR with Internal Primers, Electrophoresis and Sequencing of the Products

To assess distribution of the new SINE families, we designed two pairs of internal primers for their consensus sequences: SlmI-FR1 and SlmII-FR (table 1). The PCR conditions for the first pair were as follows: initial denaturation—3 min at 94°C; 35 cycles, each consisting of denaturation—1 min at 94°C, annealing—30 s at 51°C, and elongation—30 s at 72°C; final synthesis for 3 min at 72°C. The second pair of primers annealed at 62°C, and elongation time was 50 s. The flanking PCR for NP subfamily insertion into the GHI gene was conducted under the same general conditions, with annealing at 52°C and elongation for 1 min 10 s.

All PCR reactions were run in the total volume of 30 μl, including 1 μl of the DNA template, and were repeated at least twice to verify the results.

Sequences of SlmI subfamily-specific primers and their annealing temperature conditions are listed in table 1. In all cases synthesis was conducted for 30 s. Other parameters were the same as indicated above.

The PCR products were subjected to horizontal electrophoresis (40 min, 100 V) in a 2% agarose gel containing 1× TAE buffer. After incubation with ethidium bromide for 10–15 min, the gels were photographed. The results were verified by sequencing of the PCR products with SlmI internal primers, either directly or after extraction from the gel with QIAGEN QIAquick Gel Extraction Kit 250 (Germany). In the latter case we used a 3% agarose gel, and electrophoresis was lasting for 150 min at 100 V.

Sequence Analyses

Sequence analysis and multiple sequence alignments were performed in GENETYX-WIN, v. 5.2, and GeneDoc software, v. 2.6.002 (http://www.psc.edu/biomed/genedoc).

Average sequence divergence for SlmI subfamilies was calculated in MEGA software, version 2.1 (Kumar et al. 2001), based on a pairwise matrix of genetic distances built with the use of Kimura's two-parameter model (Kimura 1980). CpG sites were excluded from our calculations because of their higher mutation rate (Labuda and Striker 1989; Batzer et al. 1990).

Results and Discussion

Characterization of SlmI SINE Family

In the course of our experiments on coregonid short retroposons (MP subfamily of HpaI family) we isolated a clone from the genomic library of Bonneville whitefish (*Prosopium spilonotus*), in which the insertion was different from HpaI, yet contained a fragment complementary to its 3′-tail. This partial homology apparently allowed it to hybridize with MP probe and give a positive signal.

Further analysis has shown the presence of an unknown short retroposon in the cloned sequence, with a typical structure of tRNA-derived SINE (fig. 1). We have termed it SlmI, for it proved to be widespread among salmons—suborder Salmonoidei of the order Salmoniformes (see figure S1 in the Supplementary Material online).

Based on differences in the sequence structure observed among various lineages of Salmonoidei, we have designated three distinct subfamilies of SlmI SINEs: PR (stands for *Prosopium*), SI (Salmonoidei), and NP (non-*Prosopium*): see below for more details.

SlmI: tRNA-Related 5′-Terminus

The short 5′-leader sequence of 3–4 bp is followed in SlmI SINEs by a highly conserved domain of nearly 90 bp,

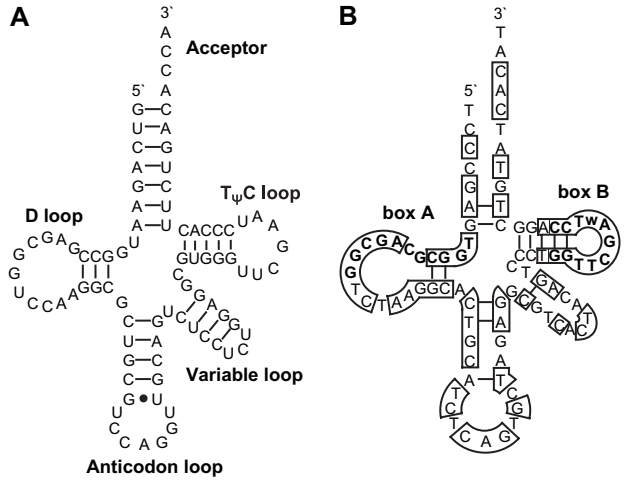

FIG. 2.—(*A*) A cloverleaf of pufferfish tRNA^{Leu} (Lowe Lab tRNAscan-SE Search Server; Lowe and Eddy 1997). Hydrogen bonds between paired residues and G-U wobble base pair are shown as lines and a dot, respectively. (*B*) Reconstruction of the cloverleaf structure of tRNA^{Leu}-related region of SlmI SINEs (PR subfamily). Residues homologous to those of tRNA^{Leu} gene are boxed; RNA Pol III promoter boxes A and B are shown in boldface type.

which has a clear homology to transfer RNA gene. Unlike in other salmonoid tRNA-related SINE families with their tRNA^{Lys}-derived promoter region, in SlmI SINEs it matched the pufferfish tRNA^{Leu} gene (see online database at Lowe Lab tRNAscan-SE Search Server; also see Lowe and Eddy 1997) with a 65% identity (fig. 2). The stems of the reconstructed cloverleaf structure appear to be the least-conserved (with the exception of well-preserved T-domain), disrupting partially (D and anticodon domains) or almost completely (acceptor arm and variable loop) the double-stranded pattern of the hypothetical tRNA^{Leu}-like secondary structure. At the same time, the overall sequence conservation in the tRNA-related 5′-region, including well-preserved A- and B-boxes of the split promoter, is relatively high (figs. 1, 2), which guarantees easy recognition by RNA Pol III and further transcription (Okada and Ohshima 1995).

SlmI: Family-Specific Conserved Central Domain

The tRNA-derived region is followed by conserved central sequence of 57 (PR), 63 (NP), or 68 bp (SI), which is specific to the SlmI family. Many SINEs are characterized by the presence of one or several conserved domains in their central region: core-sequence and Deu-domain in the superfamilies of CORE- and DeuSINEs, respectively (Gilbert and Labuda 1999; Nishihara et al. 2006); specific central region of V-SINEs (Ogiwara et al. 2002); and so on. Although these conserved domains are typical of various groups of SINEs (individual families or the above-mentioned superfamilies), their function remains unknown. Yet there is some evidence of their possible influence on SINE retrotransposition. For example, it was shown by our group before that SINE retrotransposition frequency was higher in full-length copies of eel UnaSINE1 (versus those containing only the 3′-segment). At the same time, the presence of a core-like domain had the opposite effect in the case of UnaSINE2 (Kajikawa et al. 2005). It is worth mentioning, though, that all these results have been obtained with the use of RNA Pol II for the SINE transcription, and therefore it is not known whether they reflect the actual situation in vivo, where these elements are transcribed by RNA Pol III.

Although functions of central conserved domains characteristic of one or several SINE families are yet to be discovered, it is obvious that such remarkable integrity must be a result of strong selective constraint, which indicates their functional importance. It was even demonstrated that central segments can be more conserved than the tRNA-derived and LINE-related 3′-tail regions, such as in the above-mentioned CORE-SINEs, V-SINEs, or AmnSINE1 of the DeuSINE superfamily.

In SlmI SINEs, the central domain, ending with a short (up to 11 bp) but pronounced subfamily-specific region (fig. 1 and fig. S1 in the Supplemental Material online), is as conserved as the promoter region and LINE-related 3′-segment (fig. 3). Quite noticeably, the degree of conservation drops down immediately downstream of the tRNA-related domain, with further gradual restoration toward the central part of the central domain. The second noticeable decrease

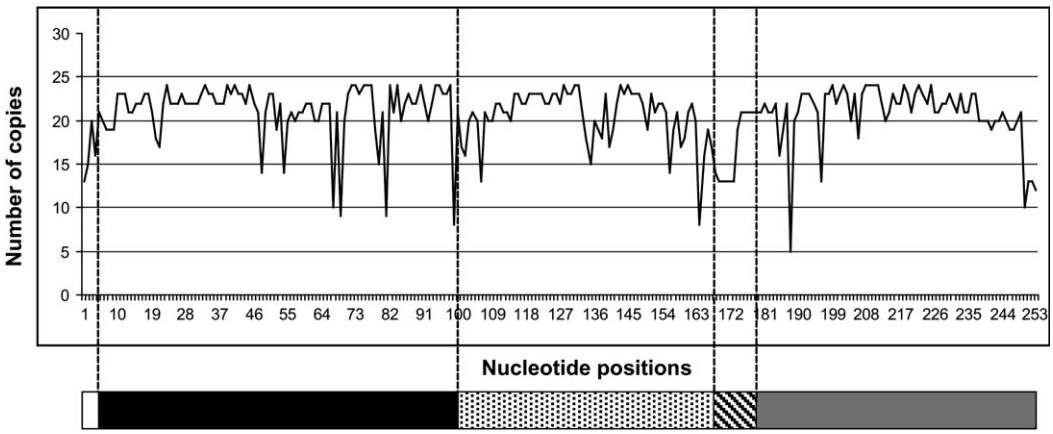

FIG. 3.—Conservation of three domains among 24 copies of SlmI SINEs: PR subfamily—13 sequences, NP—8, and SI—3. The graph shows how many copies share the same nucleotide or the gap with a consensus sequence for each nucleotide position. The transfer RNA-derived promoter region is shown in black, and the LINE-related 3′-tail in gray; the family-specific central domain is dotted, and its subfamily-specific segment is cross-hatched.

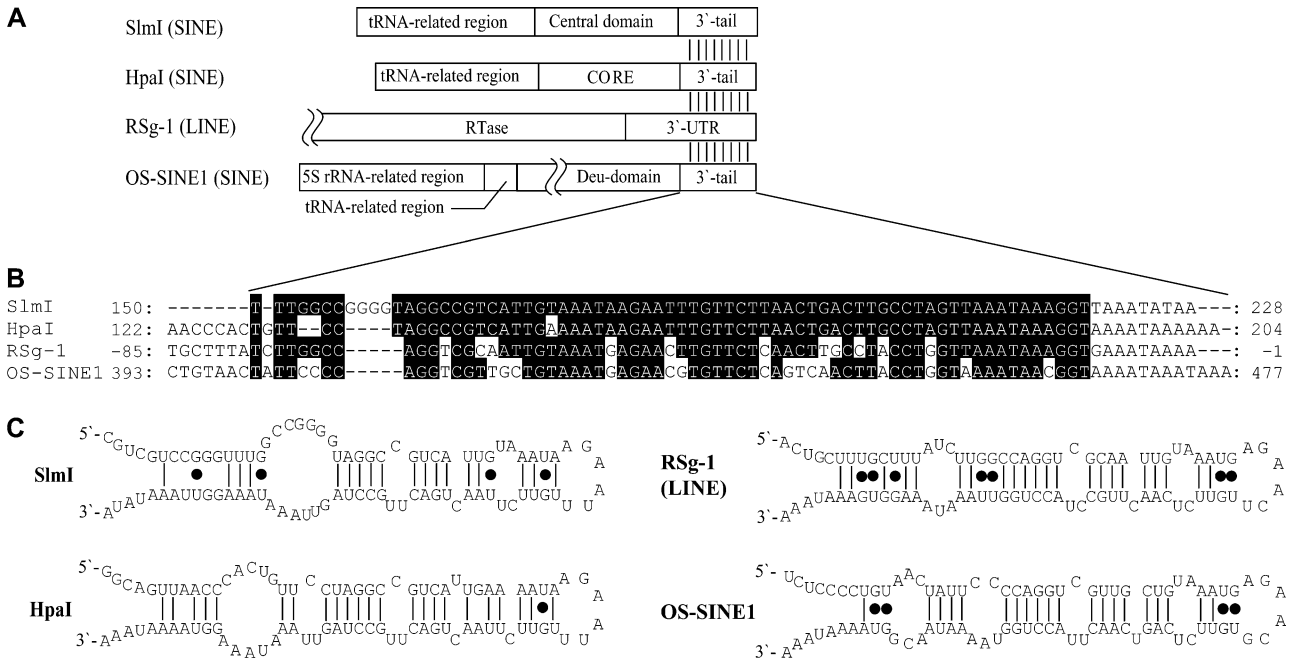

FIG. 4.—(*A*) Schematic representation of three SINE families (SlmI, HpaI, and OS-SINE1) and RSg-1 LINE sharing highly similar 3′-tail regions. (*B*) Alignment of conserved 3′-sequences of RSg-1 LINE and the related SINEs. RSg-1 consensus sequence is deduced from the following nine GenBank entries for *Oncorhynchus mykiss*: AB162342, CA347000, AJ295231, CA346593, CA353567, CA358767, CA386237, CA388068, and CA388226. Nucleotides are numbered from 5′-terminus in SINEs and from 3′—in RSg-1. (*C*) Putative secondary structures of LINE and SINE RNAs in their 3′-regions. Standard hydrogen bonds between paired residues and G-U wobble base pairs are shown as lines and dots, respectively.

falls on its subfamily-specific segment, which can be explained by calculations on the members of all three subfamilies: 13 SINEs of the PR subfamily, 6 NP, and 3 SI. This segment, being stable within each of the subfamilies, significantly differs among them (figure S1 in the online Supplementary Material). The high level of conservation in the SlmI central domain suggests its functional importance.

Numerous examples of SINE involvement in gene regulation have been described to date (for reviews, see Kramerov and Vassetzky 2005; Medstrand et al. 2005). Its mechanisms can vary and are not always well understood. In such a context, the potential regulatory function that SINE-specific central conserved domains might have acquired in the host genomes (which would be giving them some selective advantage and therefore puts these sequences under selective constraint) does not seem so improbable. Indirect proof comes from ancient MIR SINEs (Smit and Riggs 1995; Gilbert and Labuda 1999, 2000), which are mostly present in the form of highly degraded "fossil" copies (some of which even lack both termini), yet with a well-preserved "core."

SlmI: LINE-Related 3′-Tail

An approximately 70-bp-long conserved 3′-tail of SlmI SINEs is not unique to these retroposons, as they share it with two other families of salmonid SINEs—HpaI and OS-SINE1, belonging to the superfamilies of CORE-SINEs and DeuSINEs, respectively. The conserved 3′-terminus is important for the process of SINE retrotransposition (see *Introduction*) and is derived from the appropriate partner LINE. In the SINE RNA, it forms a secondary structure consisting of a stem and a loop. The latter appears to be more variable among different LINE families, and its structure is specifically recognized by the "right" RTase. As demonstrated on the UnaL2 LINE (Baba et al. 2004; Kajikawa and Okada, unpublished results), different residues in the loop are unequally important for its specific structure and hence are unequally conserved. On the contrary, the stem seems to be well conserved and may represent a common binding site for different RTases encoded by related LINEs.

The 3′-tail sequence of the HpaI family exhibits a significant homology to the last approximately 65–70 bp of the salmonoid RSg-1 LINE (Ohshima et al. 1996; Okada et al. 1997). Therefore, all of the above unrelated families of salmonoid SINEs (SlmI, HpaI, and OS-SINE1) have the same partner LINE—RSg-1—and use its enzymatic machinery for their own retrotransposition.

Figure 4 illustrates significant similarities in primary and secondary structures of the 3′-tail in all of the above SINEs and their partner LINE. In the above-mentioned UnaL2, the sequence GGRNA within the loop yields the structure that is recognized by its RTase (Baba et al. 2004). Similarly, the same regions in the salmonoid retroposons already discussed contain several conserved residues—AGAA, followed by two that vary among them: TC in RSg-1, GC in OS-SINE1, and TT in HpaI and SlmI.

Quite notably, all three families of short retroposons demonstrate almost identical values of sequence homology of their conserved 3′-tails with that of RSg-1—around 75% (table 2), which is close to the corresponding value in a well-studied UnaSINE1/UnaL2 pair (with their shorter termini)—approximately 78%. Most likely, the higher level of divergence from the original (LINE) state affects the

Table 2
Sequence Homology (%) among Conserved 3′-tail Regions of Salmonoid RSg-1 LINE and Related SINEs: SlmI, HpaI, and OS-SINE1

	SlmI	HpaI	OS-SINE1
HpaI	90		
OS-SINE1	65.7	66.7	
RSg-1	75.7	73.9	75.8

ability of RNA transcripts to bind the RTase, and hence eliminates such copies from further propagation in the genome (while recognition ability could probably be blocked even by a single mutation of one of those highly conserved residues within the loop, as shown in the above experiments with UnaL2). As a result, in all three families this value is "maintained" above the same "critical" level—74%–75%. This is distinctly higher than the sequence identity between SlmI and HpaI from one side, and OS-SINE1 from another—around 66% in both cases.

A yet more remarkable value refers to the SlmI/HpaI pair of SINEs: their 3′-termini match each other with a 90% identity. This is considerably higher than the 75% by which they match the original LINE sequence itself, and it is comparable, or even identical to the difference observed among conserved 3′-tails of individual SINE copies belonging to the same family. Consequently, the RSg-1-related tail was at a certain point acquired by one of these families from another. Comparison of their sequences enabled us to draw the conclusion that the donor was HpaI: all of the examined copies of SlmI SINEs (except for those with large deletions in this part of their sequence) contain a distinct insertion of four nucleotides in their 3′-tails, which is absent in RSg-1, OS-SINE1, and HpaI (fig. 4B). To be so widely and evenly distributed among all SlmI subfamilies, this insertion had to appear shortly after the SlmI progenitor acquired its 3′-tail—either from RSg-1 LINE or HpaI SINE—and started propagating. In other words, the insertion event happened when the overall number of SlmI copies in the genome was very low. If this is so, the scenario, according to which the source of the tail for SlmI was RSg-1, suggests that the HpaI progenitor acquired it from one of those few insertion-free copies of SlmI SINEs, which seems rather improbable. On the contrary, the alternative hypothesis that the SlmI progenitor received its RSg-1-related domain from one of the already abundant HpaI SINEs is much more likely.

Another, somewhat "straightforward" hypothesis suggests the existence of a different RSg-1 variant (let us say RSg-1* subfamily), which could be the donor of the 3′-tail for both SINE families and hence would demonstrate the same (or even higher, i.e., exceeding 90%) sequence identity between its tail and those of the two SINEs. However, other examples of stringent LINE/SINE pairs, as mentioned above, show much lower than 90% levels of homology between their common conserved 3′-tails (including those for which the "partnership" is proven experimentally). Therefore, if sequence identity in the LINE-derived region exceeding 90% is not a must for successful reverse transcription, this value for SlmI/RSg-1*, HpaI/RSg-1*, and SlmI/HpaI pairs would drop down over time, too, as a result of sequence divergence caused by mutations. On this account, we consider the latter scenario as least possible.

In contrast, the conclusions we derived above from the 75% homology observed between the 3′-tails of all three SINEs and their partner LINE explain the preservation of such a high level of identity (90%) between the conserved 3′-termini of SlmI and HpaI families over long evolutionary time in the SINE-to-SINE transmission hypothesis. By the moment when the tail was incorporated by the SlmI progenitor from an HpaI SINE, the tail sequences in the majority of copies of the latter could already have diverged significantly from the original LINE state. As a result, there was only a little capacity left for new mutations not to affect the ability of SlmI SINEs to propagate. Therefore, the 3′-tail sequence in all active SlmI copies remained highly similar to that of HpaI SINEs.

Therefore, we are inclined to conclude that the case of the SlmI and HpaI SINE families is the instance of SINE-to-SINE transmission of the LINE-related 3′-tail, and as such is the first reported to date. The exact mechanism has not been described even for the direct LINE-to-SINE transfer, but the possibility that there was an indirect way for this to occur (i.e., between SINEs) prompts us to think that in both cases it could be a more general genomic mechanism, such as gene conversion or recombination.

From the SINE perspective, our finding demonstrates much greater evolutionary flexibility of these repetitive elements. The ability to incorporate a 3′-tail from other SINEs, which are present in the genomes in much higher numbers of their copies than long retroposons, increases their chances for survival in case of loss of activity by their partner LINE. It also makes it more probable for individual copies that had been inactivated by mutations to restore their activity.

SlmI: A/T-Rich Tail

The LINE-related sequence is followed in SlmI SINEs by an A/T-rich tail of variable length and, as in other RSg-1-dependent SINEs, mostly irregular structure. Several copies had a tail partially composed of poly(A) motifs, and only few contained short tandem repeats (which are typical of SINEs mobilized by the rest of the L2 clade LINEs): $(AAATAC)_2$, $(TAAA)_2$ or $(TAA)_5$. In many SlmI sequences the variable 3′-tail was characterized by the presence of rather long runs of T residues (figure S1 in the online Supplementary Material), that act as a termination signal for RNA Pol III (Haynes and Jelinek 1981; Borodulina and Kramerov 2001). This part of the SINE sequence is known to fulfill other functions, too. For instance, its length and/or structure can be important for successful delivery of SINE RNA to the LINE reverse transcriptase (RTase) complex (Kajikawa and Okada 2002; Dewannieux et al. 2003). Mutational analyses in the experiments on other LINEs of the L2 clade revealed that more than one repetition of the repeat is necessary for successful retrotransposition. Kajikawa and Okada (2002) suggest that a slippage reaction occurs during the initiation of reverse transcription of UnaL2 RNA, and that its tandem repeats are required for such a reaction. Therefore, the case of the RSg-1 LINE

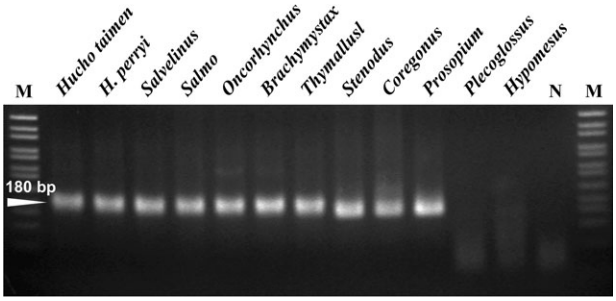

FIG. 5.—Results of the PCR with SlmI-FR primers showing distribution of SlmI SINEs among genera of the suborder Salmonoidei. *Plecoglossus* and *Hypomesus* (Osmeroidei) represent outgroups. M—originally designed marker (a modified pGEM-T vector digested with *Afa*I restriction endonuclease), N—negative control (reaction mixture without DNA template). Approximate length of the PCR product is shown above white arrow.

might represent some sort of exception from other stringent LINEs of the L2 clade in the way of initiation of its reverse transcription.

Distribution of SlmI SINEs among Ray-Finned Fish

Initially found in the genome of the Bonneville whitefish, *Prosopium spilonotus* (Coregonidae), SlmI SINEs proved to be widespread among all representatives of the suborder Salmonoidei, order Salmoniformes: Salmonidae (salmons), Thymallidae (graylings), and Coregonidae (whitefish). Thus, our search in GenBank entries revealed its presence in all genera of the family Salmonidae: *Salmo*, *Parahucho*, *Salvelinus*, *Oncorhynchus*, *Brachymystax*, and *Hucho* (figure S1 in the Supplementary Material online).

To determine precise distribution of the new SINE family, we designed a series of primers to the family consensus sequence (table 1) and performed a PCR, which has shown existence of SlmI SINEs in all Salmonoidei but not in smelts and ayu (suborder Osmeroidei, families Osmeridae and Plecoglossidae, respectively)—the closest relatives of the salmons (fig. 5). Apparently, this group of short retroposons started propagating in the genome of a salmonoid ancestor after the split of the two lineages leading to modern Osmeroidei and Salmonoidei.

We estimated the number of full copies of SlmI SINEs in the salmon genome based on a 319,299-bp-long DNA sequence of *Oncorhynchus mykiss* deposited in GenBank (Accession Number DQ246664). It contained four complete SlmI sequences oriented in both directions, one lacking most of the LINE-related 3′-end, and one full-length copy with a significantly diverged central domain and nearly 70% overall sequence identity with the family consensus. Therefore, if the genome of salmonoid fishes is 2×10^9 bp in length (Ohno et al. 1968), the number of SlmI SINEs, at least in *Oncorhynchus* species, can be roughly estimated as approximately 3×10^4 copies per genome.

Characterization of SlmII SINE Family

Further analysis of the above mentioned 3′-"truncated" copy of SlmI SINE and the sequence flanking it from the 3′-end enabled us to distinguish a stable structure with a total length of ca. 465 bp, which we found in several GenBank entries for *O. mykiss, O. masou*, and *S. salar* (see figure S2 of the online Supplementary Material). All of the found (complete) sequences were followed by variable A/T-rich 3′-tails with a structure similar to that in the SlmI family: with poly(A)/(T) runs and/or occasional tandem repeats, such as $(AT)_n$. Flanking direct terminal repeats (the result of target site duplications) were usually clearly recognizable, too (figure S2 in the Supplementary Material online). The new SINE family was obviously related to SlmI, and for this reason it was termed SlmII.

Figure 6 shows alignment of SlmII and SlmI (SI) consensus sequences. In their homologous regions (positions 1–181) the sequences match each other with 84% identity. These regions include the 5′-leader, the $tRNA^{Leu}$-related promoter segment, and the common central domain. In addition, SlmII SINEs contain 5′-remnants of the RSg-1 LINE-derived segment of SlmI family, followed by an approximately 290-bp-long 3′-domain, preceding the variable A/T-rich tail. Because short homology with RSg-1 is very unlikely to guarantee successful amplification of SlmII SINEs, this family should have its own, yet unknown partner LINE in the genome. However, our attempts to find it in the existing GenBank entries have yielded no result.

Using the same method of calculation as described above, we estimate the number of full-length copies of SlmII SINEs in the salmonoid genomes as 1.3×10^4.

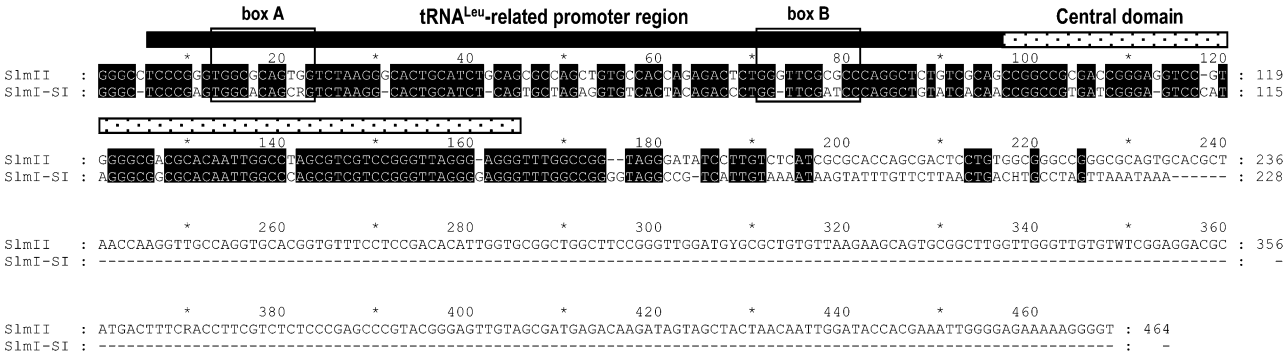

FIG. 6.—Alignment of SlmII and SlmI-SI consensus sequences. Congruent residues are shaded. Dashes denote gaps. Black and dotted bars on the top mark tRNA-related region and central domain, respectively. RNA Pol III split promoter is boxed.

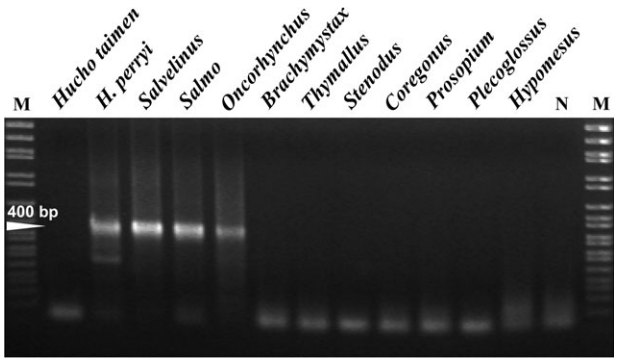

FIG. 7.—Results of the PCR with SlmII-FR primers showing distribution of SlmII SINEs among genera of the suborder Salmonoidei. *Plecoglossus* and *Hypomesus* (Osmeroidei) represent outgroups. M—originally designed marker (a modified pGEM-T vector digested with *Afa*I restriction endonuclease), N—negative control (reaction mixture without DNA template). Approximate length of the PCR product is shown above white arrow.

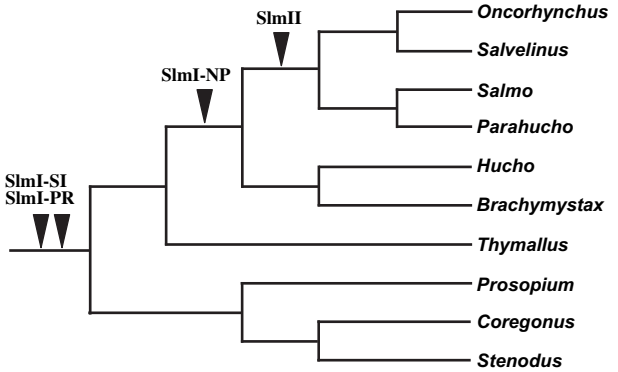

FIG. 8.—Timing of appearance of SlmI and SlmII SINEs in the genomes of Salmonoidei (indicated with arrowheads). Presented phylogeny is a compilation based on the reports by Shaposhnikova (1967, 1968, 1975), Oakley and Phillips (1999), Crespi and Fulton (2004).

Distribution of SlmII SINEs and Its Phylogenetic Implications

A PCR with primers to SlmII consensus sequence (SlmII-FR—table 1) has shown that these SINEs are not present in the genomes of salmonoids other than some members of the family Salmonidae: *Oncorhynchus*, *Salvelinus*, *Salmo* and one species of the genus *Hucho*, namely *H. perryi* (Japanese huchen), while they are absent in *H. taimen* (fig. 7).

Although salmonid phylogeny has been the subject of intensive research for decades, even the high-level relationships and taxonomic status of many forms remain disputable. However, recent analysis of 16 mitochondrial and nine nuclear genes of 30 species (Crespi and Fulton 2004) has partially elucidated some of them, showing that (*Hucho hucho*, *Brachymystax lenok*) form the sister clade to (*Salmo*, *Salvelinus*, *Oncorhynchus*, *Hucho perryi*). At the same time, the authors preferred not to draw conclusions about the status of Japanese huchen, *H. perryi*, even though various combinations of their nuclear data sets grouped it with *Salmo* with 65%–99% confidence and pointed out that within the latter clade *Salmo* was the sister-group to (*Oncorhynchus*, *Salvelinus*).

As concluded by Shaposhnikova in a series of fundamental works (1967, 1968, 1975), *Hucho* and morphologically generalized (i.e., primitive) *Brachymystax* (lenok) are very similar in many respects, such as the structure of vomer, arrangement of teeth on its head versus those of the palatine bones, and the position and shape of the frontal bones. At the same time, Japanese huchen—the only anadromous species in the genus *Hucho*—drastically differs from the other two (*H. hucho* and *H. taimen*) by the presence of basibranchial plate (copula) with teeth, the distinctive shape of the frontal bones, a reduced number of scales in the lateral line and less number of vertebrae (ca. 60). In addition, the species differs from the other two and *Brachymystax* by the presence of a hypoethmoid and the position of the supraethmoid (with its posterior part laid above the frontal bones, not entering between them). Such significant peculiarities substantiated allocation of *H. perryi* first as a different subgenus *Parahucho* (Vladykov 1963) and then as a full genus (Glubokovskiy 1990). However, these views have not been internationally accepted, and to date most specialists refer to Japanese huchen as *Hucho*.

Our results represent the first molecular data strongly supporting existence of a monotypic genus *Parahucho*. Moreover, they provide evidence that *Brachymystax* and *Hucho* represent the two most basal groups of the family Salmonidae, which concords with their morphology, as well as with the above conclusions by Crespi and Fulton (2004). Also, if we take some of their data into consideration, we might conclude that the exact allocation of *Parahucho perryi* against (*Salmo*, *Salvelinus*, *Oncorhynchus*) should probably be in the same clade with *Salmo*—Atlantic salmons (fig. 8).

Evolution of SlmI and SlmII SINEs

Apart from the genome of *Prosopium spilonotus* (Coregonidae), the PR subfamily of SlmI SINEs was found in GenBank entries for Pacific and Atlantic salmons (*Oncorhynchus* and *Salmo*, correspondingly), the same as SlmI-NP. At the same time, the latter has also been found in the entries for *Hucho, Brachymystax, Parahucho, and Salvelinus*, while SlmI-SI—only in those for the three *Oncorhynchus* and one species of *Salmo* (figure S1 in the online Supplementary Materials).

To shed more light on the distribution of SlmI subfamilies among salmonoids, we designed a series of primers to their consensus sequences. However, the PCR results were the same as shown in figure 5 for all three subfamilies. Direct sequencing of the PCR products (including those received with use of SlmI-FR and SlmI-FR1 pairs of primers) revealed the presence of PR SINEs among all of them (i.e., in all studied genera), mixed with NP PCR products in all taxa but *Thymallus* and three genera of the family Coregonidae, namely *Prosopium*, *Coregonus*, and *Stenodus*. Because consensus sequences of PR and NP subfamilies are characterized by extreme similarity over their entire length, the only distinct, subfamily-diagnostic region appeared to be too short to guarantee the subfamily-specific annealing of the designed primers. On the other hand, this structural peculiarity of SlmI SINEs made it possible to easily distinguish PR

Table 3
Average Sequence Divergence Within PR and NP Subfamilies

SlmI Subfamily	Number of Sequences	Average Genetic distance	Standard Error
PR	13	0.106	0.014
NP	9	0.088	0.012

and NP sequences among mixed PCR products by direct sequencing (data not shown) in all genera but the four mentioned above, which had only PR sequences.

In addition, we designed a pair of primers (SlmINPGH; see table 1) specific to the sequence of Growth Hormone I gene containing NP insertion in all available entries for salmonids (see legend for figure S1 in the Supplemental Material online for the list of taxa and corresponding accession numbers). However, this insertion was lacking in the same locus in all whitefish genera and graylings (data not shown).

In its turn, the PCR with SI-specific primers has clearly demonstrated the occurrence of these SINEs in all genera (data not shown), uncovering identical to SlmI-PR distribution among salmonoids. This apparently means that the split between them happened shortly after appearance of SlmI family in the genome of a salmonoid ancestor (fig. 8), either through insertion of a (TTAGGGGAGGG) motif between the central and RSg-1-derived domains in a PR-like source gene or through its deletion if the SI-like state was primordial.

A deletion of (GAGGG) in the same region of one of the SI SINE copies resulted in the appearance of an NP subfamily and its subsequent propagation in the genome of an ancestor of Salmonidae after its split with graylings (fig. 8). Despite the limited number of sequences available for comparison, a lower overall sequence divergence within the NP subfamily (table 3) also confirms it is younger than in PR evolutionary age.

The same subfamily – SlmI-SI – gave rise to the SlmII family in the lineage of (*Parahucho*, *Salmo*, *Salvelinus*, *Oncorhynchus*), apparently through replacement of most of RSg-1 LINE-related 3′-tail with the new one, supposedly belonging to yet unknown LINE. A possible alternative (to the direct replacement) scenario could first lead to the loss of the RSg-1-related tail in one of the SI copies at any evolutionary stage since the appearance of Slm-SI SINEs, with subsequent incorporation of the 3′-tail of unknown LINE at the time following the split of the above lineage with that leading to *Hucho* and *Brachymystax* (fig. 8).

Horizontal Transfer of SlmI SINEs to Blood Fluke (*Schistosoma japonicum*)

Interestingly, we found one copy of SlmI SINEs in the GenBank entry for the blood fluke, *Schistosoma japonicum* (Trematoda: Strigeiformes)—a parasitic trematode causing schistosomiasis in humans and some other mammals. It can be easily identified as one of the PR subfamily (fig. 9). Our finding adds one more family of salmonoid retroposons to those already found in schistosomal DNA (in *S. japonicum* and, in smaller numbers, in *S. mansoni*), including HpaI, SmaI, FokI SINEs, and even RSg-1 LINE (Melamed et al. 2004).

Horizontal transfer of retroposons is a rare event, but it has been registered for several SINE families to date, among them the salmonoid SmaI, which occurs simultaneously in two distant lineages of Salmonoidei: in all whitefishes and ciscoes from one side, and in two species of Pacific salmons—*O. keta* and *O. gorbuscha*—from another (Hamada et al. 1997). As for LINEs, their transmission in eukaryotic genomes is believed to be vertical (Malik et al. 1999), with the only exception being Bov-B LINE (Kordiš and Gubenšek 1997, 1998), which simultaneously exists in Squamata (snakes and two lizard infraorders) and some mammalian genomes (in ruminants and marsupials). Therefore, the case of the blood fluke is truly unique from this perspective.

At present, we can only speculate about potential processes that might have underlain such a remarkable exchange. However that may be, occurrence of the wide range of salmonoid retroposons (including evolutionary young FokI family) in the genome of these parasitic worms suggests relatively recent and obviously repeated transfer, and therefore it cannot be exclusively explained by possible parasite/host relationships that might have existed between blood flukes and salmons in the past, as proposed by Melamed et al. (2004). Moreover, even the recent (or even happening at present) cases of occasional schistosomal infection of salmons assumed by these authors (which are quite possible, taking into account the ecological versatility of *S. japonicum*—see Snyder and Loker 2000; He et al. 2001) do not explain the fact of horizontal transfer itself, as those infection provide some basis for horizontal transfer, but not the machinery.

Our group has recently discovered a viral mediation of the lateral transfer of retrotransposable elements (O. Piskurek and N. Okada, forthcoming). We speculate that the same scenario could have taken place in the case of salmons and blood flukes. This would explain the existence of a wide range of salmonoid retroposons, as well as *IGF-1*, *LH*β, and *PRL* genes, in schistosomal genomic DNA. Moreover, should it be the case, the repeated transfer of short (SINEs) and long (LINEs, genes) sequences from one organism to another could be easily explained, too. Because part of the schistosomal life cycle occurs in fresh water anyway, and even involves other aquatic organisms (snails), the viral scenario would not require immediate helminth infection of salmons themselves (though it does not exclude it either), as direct viral infection through water or the food chain is not difficult to imagine.

```
SlmI-PR consensus     1: GGGGCTCCCGAGTGGCGCAGCGGTCTAAGGCACTGCATCTCAGTGCTAGAGGCGTCACTACAGTCCCTGGTTCGAATCCAGGCTGTATCACATCCGGCCGTG 102
Schistosoma japonicum 1: ....T..........A...T.........T..............A.T...T......G............T..........C.............T... 102

103: ATTGGGAGTCCCATAGGGCGGCGCACAATTGGCCCAGCGTCGTCCGGGTTTGGCCGGTGTAGGCCGTCATTGTAAATAAGAATTTGTTCTTAACTGACTT-GCCT--AGTTAAATAAAGGTT 221
103: .........................G.........T...A..T.................T.....................G.........CA....CC...            212
```

FIG. 9.—SlmI SINE sequence from blood fluke, *Schistosoma japonicum* (GenBank accession number AY808923) aligned with the salmonid SlmI PR subfamily consensus. Dots indicate residues matching the consensus sequence; dashes denote gaps.

Alternatively, a vector other than a virus could be involved, too. Previously, our laboratory reported a case of HpaI SINE insertion into a Tc-1-like transposon in Atlantic salmon (Takasaki et al. 1996), as well as SmaI SINE insertions into RSg-1 LINE sequences (Hamada et al. 1997).

The presence of a potentially active RSg-1 LINE in schistosomal DNA could provide the basis for propagation of HpaI and SlmI SINEs in this parasite. However, to date, only HpaI transcripts (and none of RSg-1) have been detected in *S. japonicum* (Melamed et al. 2004).

Supplementary Material

Supplementary materials are available at *Molecular Biology and Evolution* online (http://www.mbe.oxfordjournals.org/).

Acknowledgments

This work was supported by Japanese Society for the Promotion of Science, grant reference number P05182. The authors are extremely grateful to Dr. S. Alekseev (Kol'tsov Institute of Developmental Biology, Russian Academy of Sciences, Moscow, Russia) for providing tissue samples of taimen, lenok, and grayling, and to Dr. D. Politov (Vavilov Institute of General Genetics RAS, Moscow) for inconnu samples. The first author also thanks Dr. N. Vassetzky (Engelhardt Institute of Molecular Biology RAS, Moscow) and Dr. E. Shubina (Belozersky Institute of Physico-Chemical Biology, Moscow State University) for helpful suggestions and discussions.

Literature Cited

Baba S, Kajikawa M, Okada N, Kawai G. 2004. Solution structure of an RNA stem-loop derived from the 3′ conserved region of eel LINE UnaL2. RNA. 10:1380–1387.

Batzer MA, Kilroy GE, Richard PE, Shaikh TH, Desselle TD, Hoppens CL, Deininger PL. 1990. Structure and variability of recently inserted Alu family members. Nucleic Acids Res. 18:6793–6798.

Blin N, Stafford DW. 1976. A general method for isolation of high molecular weight DNA from eukaryotes. Nucleic Acids Res. 3:2303–2308.

Borodulina OR, Kramerov DA. 2001. Short interspersed elements (SINEs) from insectivores: two classes of mammalian SINEs distinguished by A-rich tail structure. Mamm Genome. 12:779–786.

Brosius J. 1991. Retroposons—seeds of evolution. Science. 251: 753.

Crespi BJ, Fulton MJ. 2004. Molecular systematics of Salmonidae: combined nuclear data yields a robust phylogeny. Mol Phylogenet Evol. 31:658–679.

Dewannieux M, Esnault C, Heidmann T. 2003. LINE-mediated retrotransposition of marked Alu sequences. Nat Genet. 35: 41–48.

Gilbert N, Labuda D. 1999. CORE-SINEs: Eukaryotic short interspersed retroposing elements with common sequence motifs. Proc Natl Acad Sci USA. 96:2869–2874.

Gilbert N, Labuda D. 2000. Evolutionary inventions and continuity of CORE-SINEs in mammals. J Mol Biol. 298:365–377.

Glubokovskiy MK. 1990. Evolutionary biology of salmonid fishes. Vladivostok: Doctoral Thesis Abstract [in Russian].

Hamada M, Kido Y, Himberg M, Reist JD, Ying C, Hasegawa M, Okada N. 1997. A newly isolated family of short interspersed repetitive elements (SINEs) in coregonid fishes (whitefish) with sequences that are almost identical to those of the SmaI family of repeats: possible evidence for the horizontal transfer of SINEs. Genetics. 146:355–367.

Haynes SR, Jelinek WR. 1981. Low molecular weight RNAs transcribed in vitro by RNA polymerase III from Alu-type dispersed repeats in Chinese hamster DNA are also found in vivo. Proc Natl Acad Sci USA. 78:6130–6134.

He YX, Salafsky B, Ramaswamy K. 2001. Host–parasite relationships of *Schistosoma japonicum* in mammalian hosts. Trends Parasitol. 17:320–324.

Jurka J. 1997. Sequence patterns indicate an enzymatic involvement in integration of mammalian retroposons. Proc Natl Acad Sci USA. 94:1872–1877.

Jurka J, Smith T. 1988. A fundamental division in the Alu family of repeated sequences. Proc Natl Acad Sci USA. 85: 4775–4778.

Jurka J, Zietkiewicz E, Labuda D. 1995. Ubiquitous mammalian-wide interspersed repeats (MIRs) are molecular fossils from the Mesozoic era. Nucleic Acids Res. 23:170–175.

Kajikawa M, Ichiyanagi K, Tanaka N, Okada N. 2004. Isolation and characterization of active LINE and SINEs from the eel. Mol Biol Evol. 22:673–682.

Kajikawa M, Okada N. 2002. LINEs mobilize SINEs in the eel through a shared 3 sequence. Cell. 111:433–444.

Kapitonov VV, Jurka J. 2003. A novel class of SINE elements derived from 5S rRNA. Mol Biol Evol. 20:694–702.

Kido Y, Aono M, Yamaki T, Matsumoto K, Murata S, Saneyoshi M, Okada N. 1991. Shaping and reshaping of salmonid genomes by amplification of tRNA-derived retroposons during evolution. Proc Natl Acad Sci USA. 88: 2326–2330.

Kido Y, Himberg M, Takasaki N, Okada N. 1994. Amplification of distinct subfamilies of short interspersed elements during evolution of the Salmonidae. J Mol Biol. 241:633–644.

Kido Y, Saitoh M, Murata S, Okada N. 1995. Evolution of the active sequences of the HpaI short interspersed elements. J Mol Evol. 41:986–995.

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

Kordiš D, Gubenšek F. 1997. Bov-B long interspersed repeated DNA (LINE) sequences are present in Vipera ammodytes phospholipase A2 genes and in genomes of Viperidae snakes. Eur J Biochem. 246:772–779.

Kordiš D, Gubenšek F. 1998. Unusual horizontal transfer of a long interspersed nuclear element between distant vertebrate classes. Proc Natl Acad Sci USA. 95:10704–10709.

Kramerov DA, Vassetzky NS. 2005. Short retroposons in eukaryotic genomes. Int Rev Cytol. 247:165–221.

Kumar S, Tamura K, Jakobsen IB, Nei M. 2001. MEGA2: molecular evolutionary genetics analysis software. Bioinformatics. 17:1244–1245.

Labuda D, Striker G. 1989. Sequence conservation in Alu evolution. Nucleic Acids Res. 17:2477–2491.

Lander ES, Linton LM, Birren B, et al. (256 co-authors). 2001. Initial sequencing and analysis of the human genome. Nature. 409:860–921.

Lowe TM, Eddy SR. 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res. 25:955–964.

Malik HS, Burke WD, Eickbush TH. 1999. The age and evolution of non-LTR retrotransposable elements. Mol Biol Evol. 16:793–805.

Medstrand P, van de Lagemaat LN, Dunn CA, Landry J-R, Svenback D, Mager DL. 2005. Impact of transposable elements on the evolution of mammalian gene regulation. Cytogenet Genome Res. 110:342–352.

Melamed P, Chong KL, Johansen MV. 2004. Evidence for lateral gene transfer from salmonids to two schistosome species. Nat Genet. 36:786–787.

Nishihara H, Smit AFA, Okada N. 2006. Functional noncoding sequences derived from SINEs in the mammalian genome. Genome Res. 16:864–874.

Oakley TH, Phillips RB. 1999. Phylogeny of Salmonine fishes based on growth hormone introns: Atlantic (Salmo) and Pacific (Oncorhynchus) salmon are not sister taxa. Mol Phylogenet Evol. 11:381–393.

Ogiwara I, Miya M, Ohshima K, Okada N. 2002. V-SINEs: A new superfamily of vertebrate SINEs that are widespread in vertebrate genomes and retain a strongly conserved segment within each repetitive unit. Genome Res. 12: 316–324.

Ohno S, Wolf U, Atkin NB. 1968. Evolution from fish to mammals by gene duplication. Hereditas. 59:169–187.

Ohshima K, Hamada M, Terai Y, Okada N. 1996. The 3′ends of tRNA-derived short interspersed repetitive elements are derived from the 3′ends of long interspersed repetitive elements. Mol Cell Biol. 16:3756–3764.

Ohshima K, Okada N. 2005. SINEs and LINEs: symbionts of eukaryotic genome with a common tail. Cytogenet Genome Res. 110:475–490.

Okada N. 1991a. SINEs. Curr Opin Genet Dev. 1:498–504.

Okada N. 1991b. SINEs: short interspersed repeated elements of the eukaryotic genome. Trends Ecol Evol. 6:358–361.

Okada N, Ohshima K. 1995. Evolution of tRNA-derived SINEs. In: Maraia RJ, editor. The Impact of Short Interspersed Elements (SINEs) on the Host Genome. Austin TX: RG Landes Company. p. 61–79.

Okada N, Hamada M, Ogiwara I, Ohshima K. 1997. SINEs and LINEs share common 3′ sequence: a review. Gene. 205: 229–243.

Piskurek O, Okada N. Forthcoming. Poxviruses as possible vectors for horizontal transfer of retroposons from reptiles to mammals. Proc Natl Acad Sci USA.

Rogers J. 1985. Origins of repeated DNA. Nature. 317:765–766.

Sanger F, Nicklen S, Coulson AR. 1977. DNA sequencing with chain terminating inhibitors. Proc Natl Acad Sci USA. 74:5463–5467.

Shaposhnikova G.Kh. 1967. About systematic status of the genera *Hucho* Günther and *Brachymystax* Günther. Zool Zh. 46:254–257 [in Russian].

Shaposhnikova G.Kh. 1968. Comparative morphological study of taimen and lenok. J Ichthyol. 8:440–464 [in Russian].

Shaposhnikova G.Kh. 1975. Systematic relationships of some representatives of the family Salmonidae. Zool Zh. 54:575–582 [in Russian].

Singer MF. 1982. SINEs and LINEs: highly repeated short and long interspersed sequences in mammalian genomes. Cell. 28:433–434.

Smit A, Riggs A. 1995. MIRs are classic tRNA-derived SINEs that amplified before the mammalian radiation. Nucleic Acids Res. 23:98–102.

Snyder SD, Loker ES. 2000. Evolutionary relationships among the Schistosomatidae (Platyhelminthes: Digenea) and an Asian origin for *Schistosoma*. J Parasitol. 86:283–288.

Takasaki N, Park L, Kaeriyama M, Gharrett AJ, Okada N. 1996. Characterization of species-specifically amplified SINEs in three salmonid species—chum salmon, pink salmon, and kokanee: the local environment of the genome may be important for the generation of a dominant source gene at a newly retroposed locus. J Mol Evol. 42:103–116.

Vladykov VD. 1963. A review of salmonid genera and their broad geographical distribution. Trans R Soc Can. 1:459–504.

Waterston RH, Lindblad-Toh K, Birney E, et al. (222 co-authors). 2002. Initial sequencing and comparative analysis of the mouse genome. Nature. 420:520–562.

Weiner AM, Deininger PL, Efstratiadis A. 1986. Nonviral retroposons: genes, pseudogenes, and transposable elements generated by the reverse flow of genetic information. Annu Rev Biochem. 55:631–661.

Naoko Takezaki, Associate Editor

Accepted April 18, 2007