Detection of the Ongoing Sorting of Ancestrally Polymorphic SINEs Toward Fixation or Loss in Populations of Two Species of Charr During Speciation

Mitsuhiro Hamada,* Nobuyoshi Takasaki,* James D. Reist,† Alfred L. DeCicco,‡ Akira Goto§ and Norihiro Okada*

**Faculty of Bioscience and Biotechnology, Tokyo Institute of Technology, Midori-ku, Yokohama 226-8501, Japan, †Department of Fisheries and Oceans Canada, Winnipeg, Manitoba R3T 2N6, Canada, ‡Alaska Department of Fish and Game, Fairbanks, Alaska 98701 and §Laboratory of Embryology and Genetics, Faculty of Fisheries, Hokkaido University, Hakodate 041-8611, Japan*

Manuscript received March 9, 1998
Accepted for publication April 30, 1998

ABSTRACT

The *Fok*I family of short interspersed repetitive elements (SINEs) has been found only in the genomes of charr fishes (genus Salvelinus). In an analysis of the insertion of *Fok*I SINEs using PCR, we characterized six loci at which *Fok*I SINEs have been inserted into the genomes of *Salvelinus alpinus* (Arctic charr) and/or *S. malma* (Dolly Varden). An analysis of one locus (*Fok*-223) suggested that a sister relationship exists between *S. alpinus* and *S. malma* and the SINE at this locus might have been inserted in a common ancestor of these two species, being fixed in all extant populations examined. By contrast, SINEs at two other loci (*Fok*-211 and *Fok*-206) were present specifically in the genome of *S. alpinus*, with polymorphism among populations of this species. Moreover, the presence or absence of the SINEs of the other three loci (*Fok*-214, *Fok*-217, and *Fok*-600) varied among populations of these two species. The most plausible interpretation of this result is that SINEs, which were ancestrally polymorphic in the genome of a common ancestor of these two species, are involved in an ongoing process of differential sorting and subsequent fixation in the various populations of each species.

A retroposon is defined as a nucleotide sequence, present initially as a cellular RNA transcript, that has been reincorporated into the genome via a cDNA intermediate. This process is called retroposition (Rogers 1985; Weiner *et al.* 1986). Short interspersed repetitive elements (SINEs; Singer 1982) form one group of retroposons, and they are often present at more than 10^5 copies per genome (Okada 1991a,b; Okada and Ohshima 1995).

SINEs can be divided into two classes according to their origins. One class of SINEs is derived from the 7SL RNA (Weiner 1980; Ullu and Tschudi 1984) in the signal recognition particle (SRP) that is involved in the secretion of polypeptides during protein biosynthesis (Walter and Blobel 1982). The primate *Alu* family and the rodent type 1 (B1) family belong to this class (Schmid and Maraia 1992; Deininger and Batzer 1993). Members of the other class of SINEs originated from specific tRNAs. All known SINEs, with the exception of the primate *Alu* and rodent B1 SINEs, are members of the second class of SINEs (Okada 1991a,b; Okada and Ohshima 1995).

It is generally accepted that, in contrast to DNA transposable elements, a single unit of a SINE is never subsequently excised precisely except in cases of gross deletions. Moreover, it is believed that the insertion sites of SINE units are almost random (Kido *et al.* 1995). These features allow SINEs to serve as excellent evolutionary and phylogenetic markers (Okada 1991b). Indeed, we propose that SINE insertion analysis is one of the best methods for the determination of phylogenetic relationships among closely related species (Murata *et al.* 1993, 1996; Shimamura *et al.* 1997).

SINE insertion analysis is also useful for the studies of structures of populations of a species. In cases where SINEs were retroposed recently on an evolutionary time scale, they have not yet been fixed in populations of the species. For example, several *Alu* elements that were amplified recently have not yet been fixed within the human genome, and the distribution of these elements varies among geographically distinct groups of the world's population (Batzer *et al.* 1991, 1994; Batzer and Deininger 1991; Perna *et al.* 1992; Hammer 1994; Kass *et al.* 1994). Batzer *et al.* (1994, 1996) demonstrated the African origin of *Homo sapiens* by using these elements as phylogenetic markers.

In a previous study, we characterized three different families of SINEs in the genomes of salmonid fishes (Kido *et al.* 1991). We designated these SINEs the salmon *Sma*I family, the charr *Fok*I family and the salmonid *Hpa*I family. The *Sma*I family of SINEs is restricted to the genomes of chum salmon (*Oncorhynchus keta*) and pink salmon (*O. gorbuscha*). The charr *Fok*I family of SINEs is present only in species that belong to the

Corresponding author: Norihiro Okada, Faculty of Bioscience and Biotechnology, Tokyo Institute of Technology, 4259 Nagatsuta-cho, Midori-ku, Yokohama 226-8501, Japan.
E-mail: nokada@bio.titech.ac.jp

TABLE 1

Fish species analyzed

Species	Common name	No. of samples	Abbreviation in Figure 5	Geographic source
Salvelinus malma	Dolly Varden	1	D1	Montana Creek, Montana, United States
		10	D2	Klutina Lake/River, Alaska, United States
		10	D3	Nome River, Alaska, United States
		10	D4	Firth River, Yukon Territory, Canada
		10	D5	Unalakleet River, Alaska, United States
		10	D6	Solomon River, Alaska, United States
		3	D7	Bering Sea
		3	D8	Pilgrim River, Alaska, United States
		8	D9	Saru River, Hokkaido, Japan
		8	D10	Tachinyusu River, Hokkaido, Japan
		8	D11	Chihase River, Hokkaido, Japan
		8	D12	Yanbetsu River, Hokkaido, Japan
		7	D13	Byeraya Stream of the Naiba River, Sakhalin, Russia
S. alpinus	Arctic charr	1	A1	Overvatn Salangen, Norway
		10	A2	Lake 103, Western Arctic, Canada
		20	A3	Lake Hazen, Eastern Arctic, Canada
		10	A4	Loch Garry, Scotland
		10	A5	Maine, United States
		10	A6	Lake Inari, Finland
S. leucomaenis leucomaenis	White-spotted charr	1		Shakotan River, Hokkaido, Japan
S. namaycush	Lake trout	1		Nikko Branch, National Research Institute of Aquaculture, Honshu, Japan
S. fontinalis	Brook trout	1		Wisconsin, United States
S. confluentus	Bull trout	1		Montana Creek, Montana, United States

genus Salvelinus. The salmonid *Hpa*I family of SINEs is present in all species in the family Salmonidae but not in other species (Matsumoto *et al.* 1986; Kido *et al.* 1991, 1994, 1995; Koishi and Okada 1991; Murata *et al.* 1993, 1996). Recently, we found another family of SINEs in coregonid fishes. This family of SINEs is almost identical to the *Sma*I family and it was designated the *Sma*I-cor family (*Sma*I family in *cor*egonids; Hamada *et al.* 1997). Among these four different families of SINEs, it seems likely that the *Sma*I family of SINEs was amplified relatively recently because of its restricted distribution and the limited sequence divergence of members of this family (Kido *et al.* 1991). The discovery that all the *Sma*I SINEs in the genome of chum salmon are polymorphic and not fixed among populations of the species supports this hypothesis (Takasaki *et al.* 1997).

Charr species have attracted the interest of many evolutionary biologists because of their highly variable life-history strategies, phenotypic plasticity, and potential for sympatric morphological divergence. Behnke (1980) and Cavender (1980) proposed that the genus Salvelinus includes six major morphologically distinct species, namely, *Salvelinus fontinalis* (brook trout), *S. namaycush* (lake trout), *S. confluentus* (bull trout), *S. leucomaenis* (white-spotted charr), *S. malma* (Dolly Varden) and *S. alpinus* (arctic charr). The continuous circumpolar distribution of *S. alpinus* has been demonstrated in the Arctic, and *S. malma* occurs sympatrically with *S. alpinus* in the northern Pacific. A sister relationship between *S. alpinus* and *S. malma* was confirmed in numerous studies by reference to both morphological and biochemical markers (Phillips *et al.* 1992, 1995; Stearley and Smith 1993).

S. alpinus and *S. malma* are recognized as distinct species (Behnke 1972; Morrow 1980; Reist *et al.* 1997). However, the relationships among populations of *S. alpinus* and *S. malma* in the North Pacific, as well as the relationships between *S. alpinus* and *S. malma* in the North Pacific and the other Asian charrs, have been the subject of controversy because of the variable morphology and life histories of these fishes (Behnke 1980, 1984; Cavender 1980; Savvaitova 1995). Savvaitova (1980) performed a morphological study, and claimed that *S. malma* should be considered synonymous with *S. alpinus.* She also proposed that the designation "species-complex" provides a more appropriate description of the complex structure of these two species, claiming that all Arctic charr and Dolly Varden should be included in one superspecies, namely, the *S. alpinus* complex (Savvaitova 1995).

To elucidate the complex structure of populations of these two species, we attempted to analyze the insertions of *Fok*I SINEs in *S. alpinus* and *S. malma.* As is the case for the *Sma*I SINEs in the genome of *O. keta* (Takasaki

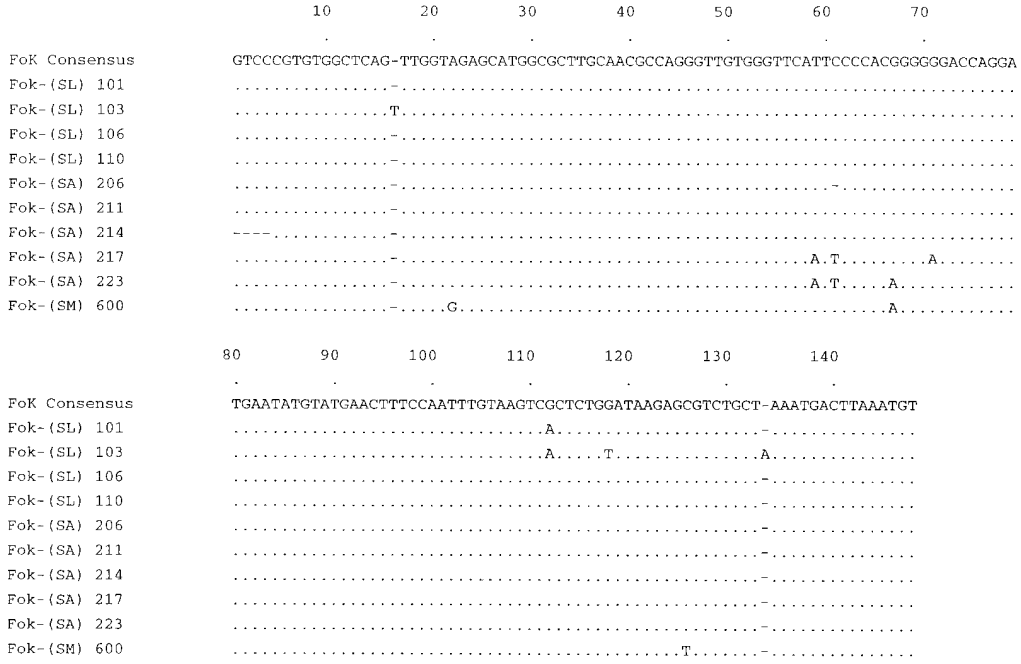

Figure 1.—Compilation of sequences of 10 members of the *Fok*I SINE family. A consensus sequence was deduced from the alignment. Nucleotides identical to those in the consensus sequence are indicated by dots and absent nucleotides are indicated by dashes. *Fok*-(SL) 101, *Fok*-(SL) 103, *Fok*-(SL) 106, and *Fok*-(SL) 110 were reported in a previous article (Kido *et al.* 1991) as *Fok*-(SLL) 1, *Fok*-(SLL) 3, *Fok*-(SLL) 6, and *Fok*-(SLL) 10, respectively (where SLL stands for *S. leucomaenis leucomaenis*). The remaining *Fok*I SINEs are newly isolated in this study. SA, *S. alpinus*; SM, *S. malma*; SL, *S. leucomaenis*.

et al. 1997), *Fok*I SINEs were found to be highly polymorphic in the genomes of *S. alpinus* and *S. malma*, and to be useful for elucidation of the complex evolutionary history of these two species. The *Fok*I SINEs are the second example to date of highly polymorphic SINEs.

MATERIALS AND METHODS

DNA samples: Individuals from each of the six species of charr from various locations were examined, as shown in Table 1. Total genomic DNA of each species was extracted as described by Blin and Stafford (1976) for large-scale preparation for the establishment of genomic libraries. For the analysis of populations, DNA was extracted from samples of individual fish as follows. One-half gram of liver, muscle or a whole fry was homogenized on ice in TNE solution, which contained 10 mm Tris-HCl (pH 8.0), 100 mm NaCl, and 1 mm EDTA. Then lysis buffer, which contained 500 μg/ml proteinase K, 2% sodium dodecyl sulfate (SDS), 10 mm Tris-HCl (pH 8.0), 150 mm NaCl, and 10 mm EDTA, was added to the solution, with incubation at 50° for 2 to 3 hr. DNA was extracted with phenol and chloroform, washed with chloroform and isoamyl alcohol, and collected by ethanol precipitation.

Construction and screening of genomic libraries, subcloning and sequencing: Total genomic DNA from *S. alpinus* and *S. malma* was separately digested with *Eco*RI for construction of a genomic library for each species. Digests were size-fractionated by sucrose gradient (10 to 40%, w/v) centrifugation. DNA fragments of 2 to 4 kb were ligated with λgt10 arms (Stratagene, La Jolla, CA) and then packaged *in vitro*. Screening was performed with an end-labeled oligonucleotide, designated *Fok*I (see positions 20–40 in Figure 1), as the probe. Hybridization was allowed to proceed at 42° overnight in a solution of 6× SSC (SSC is 0.15 m NaCl, 0.015 m trisodium citrate, pH 7), 1% (w/v) SDS, 5× Denhardt's reagent [1× Denhardt's reagent is 0.02% (w/v) Ficoll 400, 0.02% (w/v) polyvinylpyrrolidone, and 0.02% (w/v) bovine serum albumin], and 100 μg/ml herring sperm DNA. Washing was performed in 2× SSC plus 1% SDS at 50° for 30 min. Positive phage clones were isolated and their inserts were subcloned into pUC18 or pUC19. Then the inserts were sequenced with primers that corresponded to or were complementary to the consensus sequence for the *Fok*I family.

Amplification by PCR: When a unit of the *Fok*I family appeared to have been integrated at a single locus within a genome, we synthesized 5′ and 3′ oligonucleotide primers (Oligo1000 DNA synthesizer; Beckman, Fullerton, CA). The sequences of primers are shown in Figure 2. The reaction mixtures for amplification by PCR contained Tth buffer (TOYOBO, Tokyo, Japan), 0.2 mm dNTPs (Pharmacia, Uppsala, Sweden), 100 ng of each primer, 1 μg of genomic DNA, and 2 units of Tth DNA polymerase (TOYOBO) in a final volume of 100 μl or 50 μl. The thermal cycling involved 30 repeats of denaturation at 93° for 1 min, annealing at 55° for 1 min, and extension at 72° for 1 min. The products of PCR were analyzed by electophoresis in 2% (w/v) NuSieve GTG and 1% (w/v) Seakem GTG agarose gels (FMC BioProducts, Rockland, ME).

Southern hybridization: Products of PCR were transferred from gels to GeneScreen Plus membranes (New England Nuclear Research Products, Boston) in 0.4 m NaOH and 0.6 m NaCl and then dried. For detection of a SINE unit of the *Fok*I family, hybridization was performed with the *Fok*1 oligonucleotide as the probe, as described above and under the same conditions as those used for screening. For detection of orthologous loci of *Fok*-217, an end-labeled oligonucleotide that contained the flanking sequences of the SINE unit at that locus was used as a probe (*Fok*-217 flan: AGCCCTGCAGTTG CAGACGGTGCAGTTGCT). For subsequent rehybridization with a different probe, the first probe was removed by incubation in 0.4 m NaOH at 42° for 30 min.

RESULTS

Characterization of *Fok*I SINEs in the genus Salvelinus: Previous studies in our laboratory showed that the charr *Fok*I family of SINEs appeared to be restricted to species in the genus Salvelinus, such as *S. malma*,

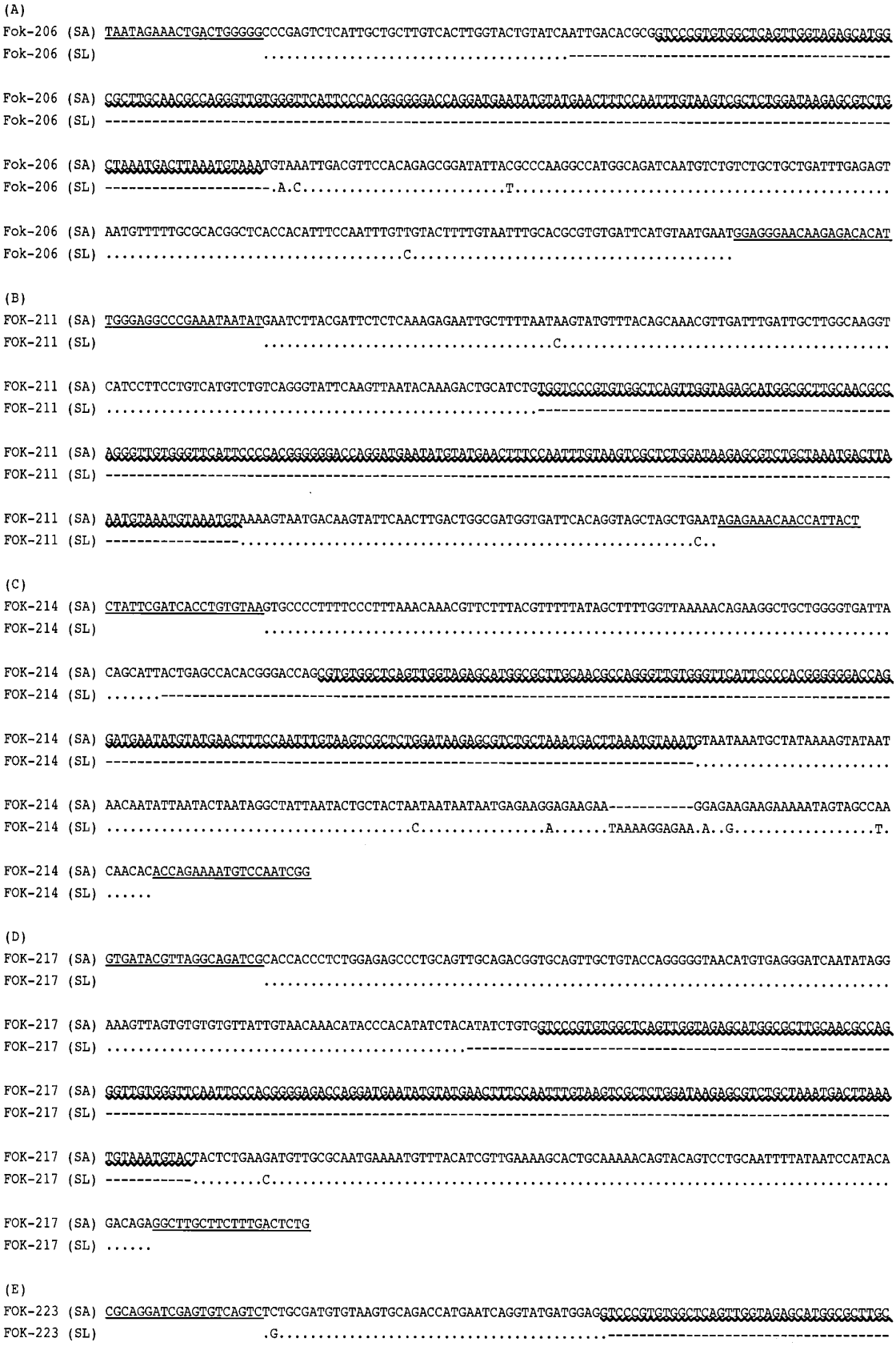

Figure 2.—Sequence of orthologous loci of (A) *Fok*-206, (B) *Fok*-211, (C) *Fok*-214, (D) *Fok*-217, (E) *Fok*-223, and (F) *Fok*-600. Primer sequences are underlined. The unit sequence of the *Fok*I SINE is shown by a wavy line, identical nucleotides by dots, and absent nucleotides by dashes. The sequences reported in this paper have been deposited in the DDBJ data base (accession numbers AB012888 to AB012899). SA, *S. alpinus*; SM, *S. malma*; SL, *S. leucomaenis.*

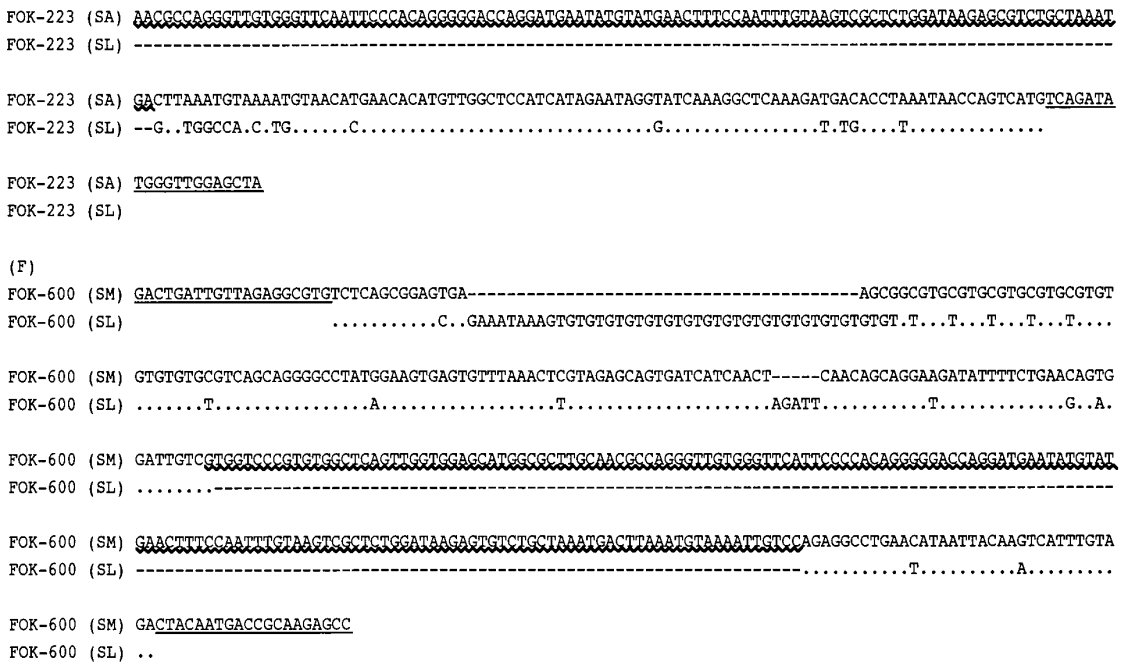

Figure 2.—*Continued.*

S. leucomaenis leucomaenis, *S. leucomaenis pluvius*, and *S. namaycush*. This family of SINEs was presumed to have been amplified at the time at which the genus Salvelinus diverged from other genera (Kido *et al.* 1991). To confirm this hypothesis, we performed a dot hybridization experiment using the *Fok*I sequence as a probe, including samples of DNA from other fishes in the genus Salvelinus (*S. fontinalis*, *S. confluentus*, and *S. alpinus*). DNAs from all fishes in the genus Salvelinus gave strong hybridization signals, confirming that they all contained the *Fok*I family of SINEs (data not shown).

We constructed genomic libraries for *S. alpinus* from Overvatn Salangen and *S. malma* from Montana Creek and screened them for *Fok*I SINEs. Five clones that contained the sequence of a *Fok*I SINE were isolated from the genomic library of *S. alpinus* and they were designated *Fok*-(SA) 206, *Fok*-(SA) 211, *Fok*-(SA) 214, *Fok*-(SA) 217, and *Fok*-(SA) 223 (SA stands for *S. alpinus*). One clone was isolated from the genomic library of *S. malma*; it was designated *Fok*-(SM) 600 (SM stands for *S. malma*). An alignment of the SINE sequences of these clones, together with those obtained from a genomic library of *S. leucomaenis leucomaenis* (Kido *et al.* 1991), is shown in Figure 1. The ten sequences of *Fok*I SINEs were very similar, and we failed to find a distinct subfamily structure of the type reported for other SINE families from salmonids (Kido *et al.* 1994, 1995; Takasaki *et al.* 1994, 1997; Hamada *et al.* 1997).

To estimate the times of insertion of the *Fok*I SINEs, we performed an analysis by PCR. We determined the 5′- and 3′-flanking regions of each *Fok*I sequence and synthesized a set of primers that flanked each unit, as shown in Figure 2. We then performed PCR using genomic DNA from *S. alpinus*, *S. malma*, *S. leucomaenis*, *S. confluentus*, *S. namaycush*, and *S. fontinalis* as templates.

As shown in Figure 3, a SINE at the *Fok*-206 and *Fok*-211 loci was found specifically in the genome of *S. alpinus* (Figure 3, A and B). We also found that a SINE at another locus, namely *Fok*-600, was specific to *S. malma* (Figure 3F). SINEs at three other loci, namely, *Fok*-214, *Fok*-217, and *Fok*-223, were found in the genomes of both *S. alpinus* and *S. malma* (Figure 3, C–E). In Figure 3, A–F, the upper DNA fragments (black arrowheads) indicate the presence of a *Fok*I SINE, while the lower DNA fragments (white arrowheads) indicate the absence of such SINEs. To confirm that orthologous loci have been faithfully amplified in these cases, we determined the sequences of the orthologous loci, which do not contain a SINE unit, of *S. malma* (*Fok*-206 and *Fok*-211) and *S. leucomaenis* (*Fok*-214, *Fok*-217, *Fok*-223, and *Fok*-600; Figure 2).

To examine whether these species-specific insertions and the group-specific insertions of *Fok*I SINEs were fixed or dimorphic in other populations, we analyzed DNA from 59 specimens of *S. alpinus* from 5 regions and 95 specimens of *S. malma* from 12 regions. The results are summarized in Table 2.

The *Fok*I SINEs specific to *S. alpinus* are dimorphic among populations of this species: At the *Fok*-206 and *Fok*-211 loci, the *Fok*I SINEs were not fixed among populations of *S. alpinus* (Table 2). For example, Figure 4 shows the results of PCR for the *Fok*-206 locus with specimens of *S. alpinus* from various populations. Individuals were scored as homozygous (+/+ or −/−) or heterozygous (+/−) for the presence of the 400-bp fragment (+) or the 238-bp fragment (−). The insertions were

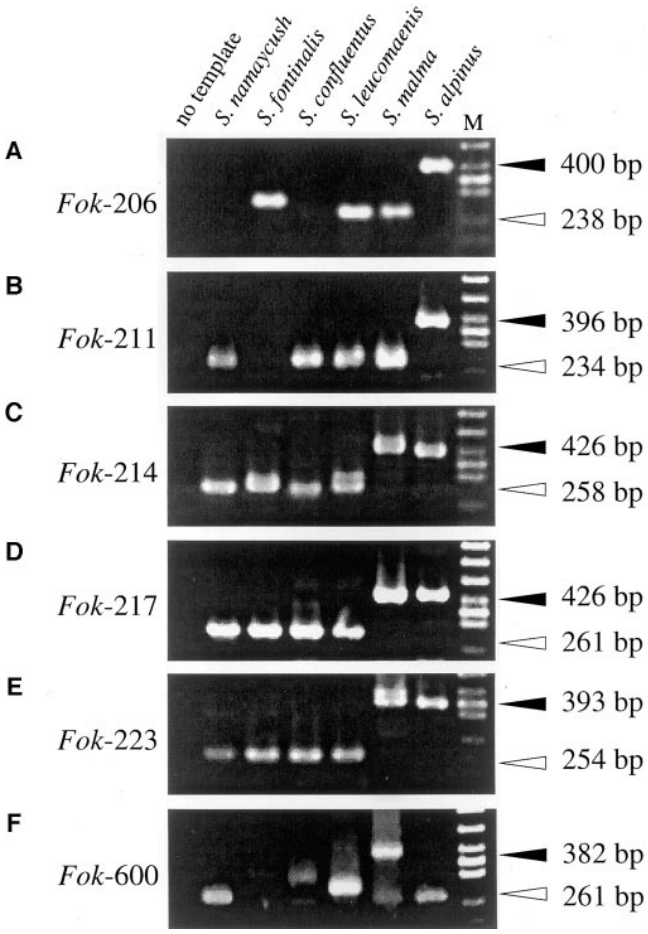

Figure 3.—A unit of a *Fok*I sequence is integrated (A and B) in the genomes of *S. alpinus* species-specifically, (C, D and E) in the genomes of both *S. alpinus* and *S. malma*, and (F) in the genomes of *S. malma* species-specifically. Black and white arrowheads indicate positions of DNA with and without a unit of the *Fok*I SINE, respectively. M indicates a *Hinc*II digest of Φx174 DNA, loaded as a source of size markers. Lengths of DNA fragments are shown in base pairs.

fixed in individuals from Maine in the United States, whereas, in the specimens from Loch Garry, Scotland, one individual was homozygous (+/+), three individuals were heterozygous (+/−), and five individuals were homozygous (−/−). Moreover, two specimens from Lake Inari, Finland, were homozygous (+/+) and five were heterozygous (+/−). No insertions were observed in the specimens from Lake 103 in the western Arctic of Canada and from Lake Hazen in the eastern Arctic of Canada.

As summarized in Table 2, in the case of the *Fok*-211 locus, insertion of a *Fok*I SINE was only detected, with a heterozygous pattern (+/−), in the genomes of specimens of *S. alpinus* from Lake Inari, suggesting that the SINE was inserted very recently at this locus in an individual in this population or a closely related population. All specimens of *S. malma* from all populations had no insertion at either locus.

***S. alpinus* and *S. malma* form a monophyletic group:** We next examined whether the *Fok*I SINEs that have been found to be commonly inserted into the genomes of *S. alpinus* and *S. malma* in the pilot experiment (*Fok*-214, *Fok*-217, and *Fok*-223; Figure 3) were fixed among populations of both species.

A SINE at the *Fok*-223 locus was fixed in every specimen from 17 populations examined (Table 2). Although we cannot exclude the possibility that, in the other remaining populations of both species, the SINE at this locus is dimorphic, the present results favor the conclusion that the SINE at this locus is fixed in all individuals of both species. This locus provides the first evidence, from SINE insertion analysis, that *S. alpinus* and *S. malma* form a monophyletic group.

The *Fok*I SINEs common to *S. alpinus* and *S. malma* are not fixed among populations of the two species: In the case of *Fok*-217 (Figure 5), insertion of the SINE in *S. malma* was observed only in a few populations, namely, those of the Klutina River/Lake and the Firth River [Figure 5A(a)], whereas insertion of the SINE in *S. alpinus* was observed in every population except for that in Lake Hazen [Figure 5B(a)]. In the case of *Fok*-214, insertion of the SINE were fixed in all the populations of *S. alpinus* examined and were presented in the population of *S. malma* in the Klutina Lake/River, but not presented in the populations of other *S. malma* (Table 2).

In the case of *Fok*-217, we confirmed the presence of the SINE in the upper fragment in Figure 5 and the validity of amplification by PCR of the orthologous locus by Southern hybridization with the *Fok*I sequence and the flanking sequence of the SINE, respectively, as probes [Figure 5, A (b and c) and B (b and c)].

In the case of the *Fok*-600 locus, we demonstrated that insertion of *Fok*I units was dimorphic among populations of *S. malma*, with the exception of the population from the Yanbetsu River (Table 2). Moreover, we also found insertion of a SINE at the *Fok*-600 locus in individuals from one population of *S. alpinus*, namely, the population from Lake Hazen. Thus, the insertion of a *Fok*I SINE at the *Fok*-600 locus was not specific to *S. malma* but was common to *S. malma* and *S. alpinus*, and it was not fixed among the populations of the two species.

Our results indicated that several insertions of *Fok*I SINEs that were common to *S. alpinus* and *S. malma* were not fixed among populations of the two species. We found that the insertions were not only intraspecifically polymorphic but also interspecifically variable. Such intraspecific polymorphism and interspecific variation of insertions of SINEs might reflect complex processes of speciation in which two closely related (sub)species diverge by subdivision into genetically distinct populations to form the more-distinguishable species (see below).

DISCUSSION

SINEs can be used as efficient tools for the determination of phylogenetic relationships among species: It is

TABLE 2

Distribution of members of the *Fok*I family in *S. alpinus* and *S. malma*

Species	Population	Locus	*Fok*-206	*Fok*-211	*Fok*-214	*Fok*-217	*Fok*-223	*Fok*-600
S. malma	Klutina Lake/River	++	0	0	10	9	10	1
	(n = 10)	+−	0	0	0	1	0	7
		−−	10	10	0	0	0	2
	Nome River	++	0	0	0	0	10	5
	(n = 10)	+−	0	0	0	0	0	5
		−−	10	10	10	10	0	0
	Firth River	++	0	0	0	0	10	2
	(n = 10)	+−	0	0	0	1	0	8
		−−	10	10	10	9	0	0
	Unalakleet River, Alaska	++	0	0	0	0	10	2
	(n = 10)	+−	0	0	2	0	0	8
		−−	10	10	8	10	0	0
	Solomon River, Alaska	++	0	0	0	0	10	4
	(n = 10)	+−	0	0	0	0	0	5
		−−	10	10	10	10	0	1
	Bering Sea	++	0	0	0	0	3	1
	(n = 3)	+−	0	0	0	0	0	2
		−−	3	3	3	3	0	0
	The Pilgrim River,	++	0	0	0	0	3	0
	Alaska	+−	0	0	0	0	0	3
	(n = 3)	−−	3	3	3	3	0	0
	The Saru River,	++	0	0	0	0	8	1
	Hokkaido	+−	0	0	0	0	0	0
		−−	8	8	8	8	0	7
	Tachunyusu River,	++	0	0	0	0	8	5
	Hokkaido	+−	0	0	0	0	0	3
		−−	8	8	8	8	0	0
	Chihase River,	++	0	0	0	0	8	3
	Hokkaido	+−	0	0	0	0	0	1
	(n = 8)	−−	8	8	8	8	0	4
	Yanbetsu River,	++	0	0	0	0	8	0
	Hokkaido	+−	0	0	0	0	0	0
		−−	8	8	8	8	0	8
	Naiba River, Sakhalin	++	0	0	0	0	7	2
	(n = 7)	+−	0	0	0	0	0	5
		−−	7	7	7	7	0	0
S. alpinus	Lake 103, Western Arctic	++	0	0	10	10	10	0
	(n = 10)	+−	0	0	0	0	0	0
		−−	10	10	0	0	0	10
	Lake Hazen,	++	0	0	20	0	20	6
	Eastern Arctic	+−	0	0	0	0	0	13
	(n = 20)	−−	20	20	0	20	0	1
	Loch Garry, Scotland	++	1	0	9	9	9	0
	(n = 9)	+−	3	0	0	0	0	0
		−−	5	10	0	0	0	9
	Maine, United States	++	10	0	10	10	10	0
	(n = 10)	−−	0	0	0	0	0	0
		−−	0	10	0	0	0	10
	Lake Inari, Finland	++	2	8	10	10	10	0
	(n = 10)	+−	5	2	0	0	0	0
		−−	3	0	0	0	0	10

believed that a SINE is amplified in the germ cells of one individual and then spreads within a population through sexual reproduction and random genetic drift in the same manner as other changes in the DNA (Kimura 1983). As the frequency of the SINE increases gradually in the population, the genomes of members of the population must necessarily be polymorphic. However, among the many SINEs examined to date, intraspecific polymorphism of SINEs is very rare. Most SINEs, including the *Hpa*I SINEs in salmonids, have been fixed in all populations of a given species (Takasaki *et al.* 1994, 1996). These results suggest that the

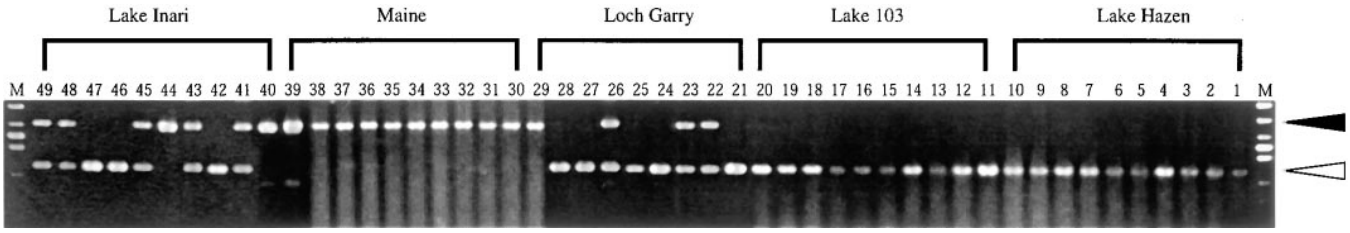

Figure 4.—The SINE unit in the Fok-206 locus that was specifically inserted in the genomes of *S. alpinus* appears to be polymorphic among populations of this species. Forty-nine specimens of *S. alpinus* were collected from Lake Hazen (lanes 1–10), Lake 103 (lanes 11–20), Loch Garry (lanes 21–29), Maine (lanes 30–39), and Lake Inari (lanes 40–49). Black and white arrowheads indicate positions of DNA with and without a unit of the *Fok*I SINE, respectively.

period between the divergence of any two salmonid species and the present time has been long enough for SINEs to become fixed in all populations of each species. On the basis of this observation and the synapomorphic character of SINEs, phylogenetic relationships among species in the genus Oncorhynchus were determined in our laboratory (Murata *et al.* 1993, 1996). In this study, the existence of the SINE at the *Fok*-223 locus in *S. malma* and *S. alpinus* suggests a monophyletic relationship among the two species of Salvelinus (Figure 6).

Are populations of *S. alpinus* or *S. malma* monophyletic? In this study, we found intraspecific polymorphism and interspecific variation when we examined the insertion of SINEs in a large number of populations of the two charr species. It is possible that such patterns of insertion directly reflect the phylogenetic relationships among the populations. For example, can the results for *Fok*-214 be interpreted to indicate the monophyly of the population of *S. malma* in Klutina Lake and of all the populations of *S. alpinus*? Does the population of *S. malma* in Klutina Lake need to be reclassified as *S. alpinus*? In general, species have been identified from morphological characteristics, but it is well known that morphological characteristics are plastic and can respond to variations in the environment. Such phenotypic plasticity has caused many problems in the taxonomic analysis of charr (for review, see Behnke 1980, 1984; Cavender 1980; Savvaitova 1995; Phillips and Oakley 1997). Thus, the possibility exists that the present taxonomic status of several populations of the two species might be incorrect.

The interpretation, in terms of phylogeny, of the results for *Fok*-214 is, however, inconsistent with the interpretation of the results for *Fok*-600, which indicated that all populations of *S. malma*, with the exception of the populations in the Yanbetsu River, and *S. alpinus* from Lake Hazen have a SINE insertion at this locus. In addition, the interpretation of the results for *Fok*-600 is inconsistent with the interpretation of the results for *Fok*-217. Furthermore, it is very unlikely from a taxonomic perspective that the population of *S. malma* in Klutina Lake/River actually belongs to *S. alpinus* (Reist *et al.* 1997).

Accordingly, we must conclude that the patterns of insertions of *Fok*I SINEs that we found do not necessarily reflect the actual phylogenetic relationships among the populations of these species.

Ancestral polymorphism is the most plausible explanation for variations in the presence or absence of a SINE at a given locus: To date, only two examples of intraspecific polymorphism of SINEs have been reported, namely, *Alu* SINEs in human populations and *Sma*I SINEs in populations of chum salmon, and each example has been shown to be useful for population analysis (Batzer *et al.* 1994, 1996; Takasaki *et al.* 1997). SINEs that are polymorphic among the populations of *S. alpinus* and *S. malma* provide a third example. In addition, these SINEs provide the first example of a useful tool for clarification of the processes of dispersion and sorting of ancestrally polymorphic SINEs among populations during speciation, as discussed below.

In this study, we clearly showed that the presence or absence of insertions of *Fok*I SINEs (*Fok*-214, *Fok*-217, and *Fok*-600) varied in populations of *S. alpinus* and *S. malma*. The most plausible explanation for these intraspecific and interspecific variations is that they are the result of ancestral polymorphism.

Insertions of *Fok*I SINE at the *Fok*-214, *Fok*-217, and *Fok*-600 loci might have occurred in a common ancestor of *S. malma* and *S. alpinus*. Before fixation of each SINE at these loci in the ancestral species, the speciation of *S. malma* and *S. alpinus* must have occurred and the polymorphic SINEs were inherited by and sorted to populations of the two new species. After speciation, the SINEs at the *Fok*-214, *Fok*-217, and *Fok*-600 loci were fixed, lost, or became polymorphic as a result of random genetic drift in each population. There is a possibility of such a situation having occurred if the population size of the ancestral species was too large, or the period between the divergence of these two species and the present time was too short to be fixed in all populations of each species. Therefore, the present observations might reflect the ongoing processes of sorting of ancestrally polymorphic SINEs toward fixation or loss in populations of the two new species during speciation. The complete fixation of the SINE at the *Fok*-223 locus in *S. malma* and *S. alpinus* might indicate that the insertion of this SINE at this locus occurred relatively soon after

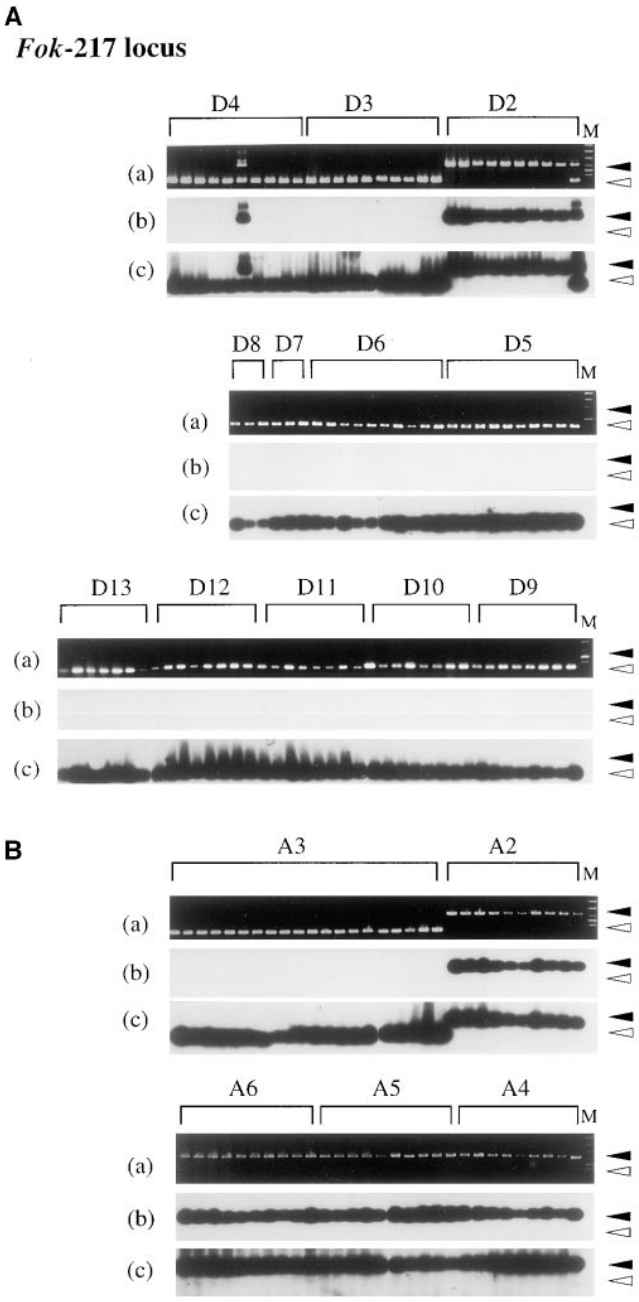

Figure 5.—The SINE unit in the *Fok*-217 locus that was found to be commonly inserted in (A) *S. malma* and (B) *S. alpinus* appears to exhibit polymorphism among populations of the two species. In A and B, the results of agarose gel electrophoresis of the products of PCR with the appropriate primers are shown in a, the results of Southern hybridization using a unit of the *Fok*I sequence as the probe are shown in b, and rehybridization using the flanking region as the probe is shown in c. See Table 1 for corresponding names of populations.

divergence of the ancestral species of *S. malma* and *S. alpinus* from other species of Salvelinus (Figure 6).

The significance of trans-species polymorphism in evolution was first indicated by Jan Klein and his colleagues (Klein 1987; Figueroa *et al.* 1988) in the major histocompatibility complex loci of the mouse and rat. Since SINEs described here appear to be neutral, the present analysis provides a more general scheme of how trans-species polymorphism behaves during the speciation of two closely related species.

Hybridization is another possibility: Hybridization between *S. malma* and *S. alpinus*, with successive introgression of nuclear information from one species to the other, provides another explanation for these variations, although this explanation appears less likely.

There is no good evidence of intermediates formed by hybridization between *S. malma* and *S. alpinus* (McPhail 1961; Behnke 1980). *S. malma* and *S. alpinus* are now considered valid species in view of the identification of several lakes in which the two species live sympatrically with no signs of hybridization (DeLacy and Morton 1943; McPhail 1961). However, there have been suggestions of the existence of hybrids between *S. malma* and *S. alpinus* (Volobuyev *et al.* 1979), and some evidence for gene flow from *S. alpinus* to *S. malma* has been presented (Gharrett *et al.* 1991). In the present case, we cannot yet completely rule out the possibility of hybridization because we did not analyze sympatric populations of the two species. However, introgression seems less likely than ancestral polymorphism, at least in the case of the populations of *S. alpinus* in Lake Hazen. Lake Hazen, at the northern end of Ellesmere Island in the Canadian high Arctic, is the largest body of freshwater in the world that is located entirely north of the Arctic Circle (Johnson 1990). Because the only species of fish in this lake is *S. alpinus* (Reist *et al.* 1995), it is unlikely that polymorphism of SINE insertions could be a result of hybridization between *S. alpinus* and *S. malma*. The addition of sympatric populations of the two species to our analysis will allow us to clarify and perhaps even eliminate the possibility of hybridization.

SINEs will be a useful tool for distinguishing one population from another: In this study, we found that the distribution of *Fok*I SINEs at distinct loci varied among remote populations. As described above, such patterns of insertion do not necessarily reflect the phylogenetic relationships among the populations. However, we can use them as genetic markers to distinguish genetic variations. For example, at the *Fok*-600 locus, among the populations of *S. malma* that we showed, only the population from the Yanbetsu River had no insertion of a *Fok*I SINE. The population of *S. malma* in Lake Shikaribetsu and its inlet stream, the Yanbetsu River, on Hokkaido Island, Japan, is thought to be a subspecies of *S. malma* (*S. malma miyabei*). *S. malma miyabei* has the most numerous gill-rakers, which range in number from 23 to 29 with a mode of 26, of all species of Salvelinus (Maekawa 1977). In addition, from an analysis of isozymes, *S. malma miyabei* was found to differ markedly from other populations of *S. malma* (Mitsuboshi *et al.* 1992). The results for *Fok*-600, as well as the results of other morphological, ecological, and bio-

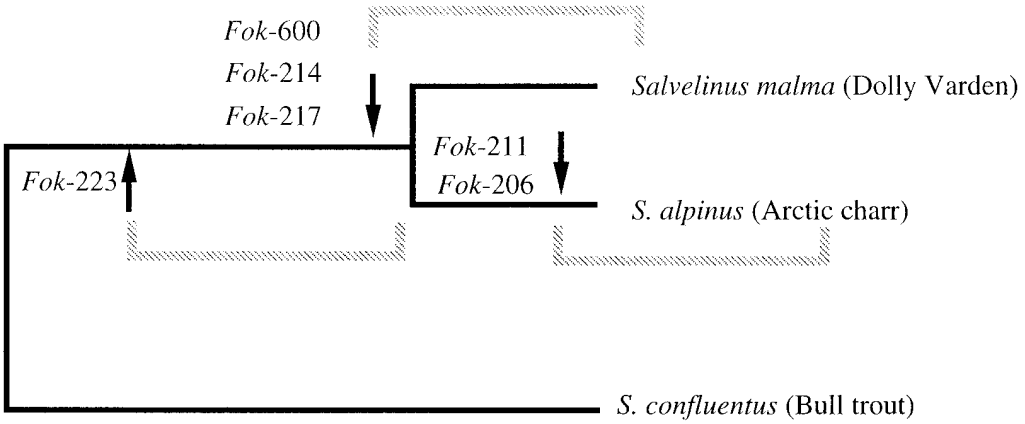

Figure 6.—Timing of the insertion of *Fok*I SINEs in the lineage of *S. alpinus* and *S. malma*. Black arrows indicate the time of insertion of *Fok*I SINEs. Square hatched brackets indicate the time required for a SINE to be fixed within a species. The length of square hatched brackets is arbitrary on condition that it is longer than the length between the time when *S. malma* and *S. alpinus* diverged and the present time, and that it is shorter than the length between the time when an ancestral species of *S. malma* and *S. alpinus* diverged from another Salvelinus species and the time when *S. malma* and *S. alpinus* diverged. The phylogenetic tree shown here was deduced from an analysis of morphology (Stearley and Smith 1993).

chemical studies, also support the hypothesis that the species in one distinct population of *S. malma*, which became landlocked in Lake Shikaribetsu during the recent Ice ages, might have adapted intrinsically to the environment in the lake (Maekawa 1977, 1985). The population size of the ancestral population of *S. malma miyabei* may have been small enough to lose the SINE insertion in all individuals.

Further screening of *Fok*I SINE loci and analysis of more samples from various regions should allow us to clarify the structures of populations and the evolutionary history of *S. alpinus* and *S. malma*.

The authors are grateful to T. Yoshikawa, T. Masuda, and M. Tsuji for their help in isolation of SINE clones. We thank N. Davis (Bering Sea), F. Kircheis (Maine), J. Vuorinen (Finland), and M. Saneyoshi (Japan) for providing those samples and Parks Canada for field assistance at Lake Hazen. North American field work was supported by Fisheries and Oceans Canada and Alaska Department of Fish and Game. This work was supported by a Grant-in-Aid for Specially Promoted Research from the Ministry of Education, Science, and Culture of Japan.

LITERATURE CITED

Batzer, M. A., and P. L. Deininger, 1991 A human-specific subfamily of *Alu* sequences. Genomics **9:** 481–487.

Batzer, M. A., V. A. Gudi, J. C. Mena, D. W. Foltz, R. J. Herrera *et al.*, 1991 Amplification dynamics of human-specific (HS) alu family members. Nucleic Acids Res. **19:** 3619–3623.

Batzer, M. A., M. Stoneking, M. Alegia-Hartman, H. Bazan, D. H. Kass *et al.*, 1994 African origin of human-specific polymorphic *Alu* insertions. Proc. Natl. Acad. Sci. USA **91:** 12288–12292.

Batzer, M. A., S. S. Arcot, J. W. Phinney, M. Alegia-Hartman, D. H. Kass *et al.*, 1996 Genetic variation of recent *Alu* insertions in human populations. J. Mol. Evol. **42:** 22–29.

Behnke, R. J., 1972 The systematics of salmonid fishes of recently glaciated lakes. J. Fish. Res. Board Can. **29:** 639–671.

Behnke, R. J., 1980 A systematic review of the genus Salvelinus, pp. 441–481 in *Charrs: Salmonid Fishes of the Genus Salvelinus*, Vol. I, edited by E. K. Balon. Junk Publishers, The Hague.

Behnke, R. J., 1984 Organizing the diversity of the Arctic char complex, pp. 2–21 in *Biology of the Arctic Charr*, edited by L. Johnson and B. L. Burns. University of Manitoba Press, Winnipeg.

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

Cavender, T. M., 1980 Systematics of Salvelinus from the North Pacific Basin, pp. 295–322 in *Charrs: Salmonid Fishes of the Genus Salvelinus*, Vol. I, edited by E. K. Balon. Junk Publishers, The Hague.

Deininger, P. L., and M. A. Batzer, 1993 Evolution of retroposons, pp. 157–196 in *Evolutionary Biology*, Vol. 27, edited by M. K. Hecht, R. J. MacIntyre and M. T. Clegg. Plenum Press, New York.

DeLacy, A. C., and W. M. Morton, 1943 Taxonomy and habits of the charrs, *Salvelinus malma* and *Salvelinus alpinus*, of the Karluk drainage system. Trans. Am. Fish. Soc. **72:** 79–91.

Figueroa, F., E. Gunther and J. Klein, 1988 MHC ploymorphism pre-dating speciation. Nature **335:** 265–267.

Gharrett, A. J., A. Goto and F. Yamazaki, 1991 A note on the genetic contrast of sympatric Dolly Varden (*Salvelinus malma*) and Arctic charr (*S. alpinus*) in the Karluk river system, Alaska, pp. 37–48 in *Reproductive Biology and Population Genetics of Dolly Varden (Salmonidae)*, edited by F. Yamazaki. Report of overseas work supported by Grant-in-Aid for Overseas Scientific Survey of the Ministry of Education, Science and Culture of Japan, during 1987–1990, Tokyo.

Hamada, M., Y. Kido, M. Himberg, J. D. Reist, Y. Cao *et al.*, 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 *Sma*I family of repeats: possible evidence for the horizontal transfer of SINEs. Genetics **146:** 355–367.

Hammer, M. F., 1994 A recent insertion of an *Alu* element on the Y chromosome is a useful marker for human population studies. Mol. Biol. Evol. **11:** 749–761.

Johnson, L., 1990 Hazen Lake, pp. 1–9, in *Data Book of World Lake Environments*, edited by Lake Biwa Research Institute and International Lake Environment Committee. United Nations Environment Programme, Otsu, Japan.

Kass, D. H., C. Aleman, M. A. Batzer and P. L. Deininger, 1994 Identification of a human specific *Alu* insertion in the Factor XIIIB gene. Genetika **94:** 1–8.

Kido, Y., M. Aono, T. Yamaki, K. Matsumoto, S. Murata *et al.*, 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., M. Himberg, N. Takasaki and N. Okada, 1994 Amplification of distinct subfamilies of short interspersed elements during evolution of the Salmonidae. J. Mol. Biol. **241:** 633–644.

Kido, Y., M. Saito, S. Murata and N. Okada, 1995 Evolution of the active sequences of the *Hpa*I short interspersed elements. J. Mol. Evol. **41:** 986–995.

Kimura, M., 1983 *The Neutral Theory of Molecular Evolution*. Cambridge University Press, Cambridge.

Klein, J., 1987 Origin of major histocompatibility polymorphism: the trans-species hypothesis. Hum. Immun. **19:** 155–162.

Koishi, R., and N. Okada, 1991 Distribution of the salmonid Hpa I family in salmonid species demonstrated by in vitro runoff transcription assay of total genomic DNA: a procedure to estimate repetitive frequency and sequence divergence of a certain repetitive family with a few known sequences. J. Mol. Evol. **32:** 43–52.

Maekawa, K., 1977 Studies on the variability of the landlocked Miyabe char, *Salvelinus malma miyabei* I. Development and early life history (in Japanese). Jap. J. Ecol. **27:** 91–102.

Maekawa, K., 1985 Life History Patterns of the Miyabe Charr in Shikaribetsu Lake, Japan, pp. 233–250 in *Biology of Arctic Charr*, edited by L. Johnson and B. Burns. University of Manitoba Press, Winnipeg.

Matsumoto, K., K. Murakami and N. Okada, 1986 Gene for lysine $tRNA^1$ may be a progenitor of the highly repetitive and transcribable sequences present in the salmon genome. Proc. Natl. Acad. Sci. USA **83:** 3156–3160.

McPhail, J. D., 1961 A systematic study of the *Salvelinus alpinus* complex in North America. J. Fish. Res. Board Can. **18:** 793–816.

Mitsuboshi, T., A. Goto and F. Yamazaki, 1992 Genetic differentiation of the Dolly Varden *Salvelinus malma* in Hokkaido, Japan. Bull. Fac. Fish. Hokkaido Univer. **43:** 153–161.

Morrow, J. E., 1980 Analysis of the dolly varden charr, *Salvelinus malma*, of northwestern North America and northeastern Siberia, pp. 323–338 in *Charrs, Salmonid Fishes of the Genus Salvelinus*, edited by E. K. Balon. Junk Publishers, The Hague.

Murata, S., N. Takasaki, M. Saitoh and N. Okada, 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.

Murata, S., N. Takasaki, M. Saitoh, H. Tachida and N. Okada, 1996 Details of retropositional genome dynamics that provide a rationale for a generic division: the distinct branching of all the Pacific salmon and trout (*Oncorhynchus*) from the Atlantic salmon and trout (*Salmo*). Genetics **142:** 915–926.

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

Okada, N., 1991b SINEs. Curr. Opin. Genet. Dev. **1:** 498–504.

Okada, N., and K. Ohshima, 1995 Evolution of tRNA-derived SINEs, pp. 61–79 in *The Impact of Short Interspersed Elements (SINEs) on the Host Genome*, edited by R. J. Maraia. R. G. Landes Company, Austin, TX.

Perna, N. T., M. A. Batzer, P. L. Deininger and M. Stoneking, 1992 *Alu* insertion polymorphism: a new type of marker for human population studies. Hum. Biol. **64:** 641–648.

Phillips, R. B., and T. H. Oakley, 1997 Phylogenetic relationships among the Salmoninae based on nuclear and mitochondrial DNA sequences, pp. 145–162 in *Molecular Systematics of Fishes*, edited by T. D. Kocher and C. A. Stepien. Academic Press, San Diego, USA.

Phillips, R. B., K. A. Pleyte and M. R. Brown, 1992 Salmonid phylogeny inferred from ribosomal DNA restriction maps. Can. J. Fish. Aquat. Sci. **49:** 2345–2353.

Phillips, R. B., S. L. Sajdak and M. J. Domanico, 1995 Relationships among charrs based on DNA sequences. Nordic J. Freshwater Res. **71:** 378–391.

Reist, J. D., E. Gyselman, J. A. Babaluk, J. D. Johnson and R. Wissink, 1995 Evidence for two morphotypes of Arctic char (*Salvelinus alpinus* (L.)) from Lake Hazen, Ellesmere Island, Northwest Territories, Canada. Nordic J. Freshwater Res. **71:** 396–410.

Reist, J. D., J. D. Johnson and T. J. Carmichael, 1997 Variation and specific identity of char from Northwestern Arctic Canada and Alaska. Am. Fish. Soc. Sym. **19:** 250–261.

Rogers, J. H., 1985 The origin and evolution of retroposons. Int. Rev. Cytol. **93:** 187–279.

Savvaitova, K. A., 1980 Taxonomy and biogeography of charrs in the Paleoarctic, pp. 281–294 in *Charrs, Salmonid Fishes of the Genus Salvelinus*, edited by E. K. Balon. Junk Publishers, The Hague.

Savvaitova, K. A., 1995 Patterns of diversity and processes of speciation in Arctic char. Nordic J. Freshwater Res. **71:** 81–91.

Schmid, C., and R. Maraia, 1992 Transcriptional regulation and transpositional selection of active SINE sequences. Curr. Opin. Genet. Dev. **2:** 874–882.

Shimamura, M., H. Yasue, K. Ohshima, H. Abe, H. Kato *et al.*, 1997 Molecular evidence from retroposons that whales form a clade within even-toed ungulates. Nature **388:** 666–670.

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

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

Takasaki, N., S. Murata, M. Saitoh, T. Kobayashi, L. Park *et al.*, 1994 Species-specific amplification of tRNA-derived short interspersed repetitive elements (SINEs) by retroposition: a process of parasitization of entire genomes during the evolution of salmonids. Proc. Natl. Acad. Sci. USA **91:** 10153–10157.

Takasaki, N., L. Park, M. Kaeriyama, A. J. Gharrett and N. Okada, 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.

Takasaki, N., T. Yamaki, M. Hamada, L. Park and N. Okada, 1997 The salmon *Sma*I family of short interspersed repetitive elements (SINEs): Interspecific and intraspecific variation of the insertion of SINEs in the genomes of chum and pink salmon. Genetics **146:** 369–380.

Ullu, E., and C. Tschudi, 1984 *Alu* sequences are processed 7SL RNA genes. Nature **312:** 171–172.

Volobuyev, V. V., E. D. Vasilyeva and K. A. Savvaitova, 1979 Systematic status of the anadromous Chukotsk charr of the genus Salvelinus. Voprosy Ikhtiologii **19:** 408–418 (in Russian).

Walter, P., and G. Blobel, 1982 Signal recognition particle contains a 7S RNA essential for protein transcription across the endoplasmic reticulum. Nature **299:** 691–698.

Weiner, A. M., 1980 An abundant cytoplasmic 7S RNA is complementary to the dominant interspersed middle repetitive DNA sequence family in the human genome. Cell **22:** 209–218.

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

Communicating editor: N. Takahata