J Mol Evol (1988) 27:8–16

Journal of Molecular Evolution

Characterization of a Protamine Gene from the Chum Salmon (*Oncorhynchus keta*)

Robert D. Moir and Gordon H. Dixon

Department of Medical Biochemistry, Faculty of Medicine, Health Sciences Centre, University of Calgary, 3330 Hospital Drive N.W., Calgary, Alberta T2N 4N1, Canada

Summary. We have cloned and sequenced a protamine gene from the chum salmon (*Oncorhynchus keta*). This gene sequence is highly homologous to one found in the rainbow trout (*Salmo gairdneri*), including the conservation of two structurally different repetitive elements. One of these repeats resembles a nonviral retroposon and the second is similar to a retroviral-like transposable element. The degree of sequence divergence between the *O. keta* and *S. gairdneri* genes is much less within the transcription unit than in the repetitive elements or the remainder of the flanking DNA, suggesting that since the coding and the untranslated regions are highly conserved, both contribute significantly to the structure and stability of protamine mRNA (or its cognate messenger ribonucleoprotein) and this may be important for the translational control of protamine synthesis.

Key words: Protamine — Repetitive DNA — Sequence divergence

Introduction

During spermatogenesis in most animals a set of basic sperm-specific proteins is synthesized which replaces the somatic histones and strongly condenses the DNA in the mature sperm (Subirana 1975; Dixon et al. 1986). There is a wide variety in the exact nature of the sperm proteins used throughout phylogeny and several attempts have been made to classify the different sperm proteins and deduce an evolutionary trend or rationale (Bloch 1969; Dixon et al. 1986; Kasinsky et al. 1986).

Members of the teleost fish family Salmonidae synthesize the true protamines (Ando et al. 1973). Our laboratory has examined the expression of the protamines in the rainbow trout (*Salmo gairdneri*) in some detail (Dixon et al. 1986). There are six very similar protamine proteins in a mature rainbow trout testis (McKay et al. 1986), all 30–32 amino acids long, with approximately 20 arginine residues. The proteins can be classified into three families on the basis of length, distribution of arginine tracts, and the identity of nonarginine amino acids (McKay et al. 1986). Six genes corresponding to the proteins of one of these families have been cloned and sequenced (States et al. 1982; Aiken et al. 1983). The sequences are very homologous to each other, both in the genes themselves and for substantial distances 5′ and 3′. The protamine genes do not have introns and are not closely clustered in the manner of the trout histone genes (Connor et al. 1984). We have been examining the protamine genes of several salmonid fish species to determine the rate of sequence divergence in this gene family and perhaps define conserved, functionally important regions. In this report, we describe the structure of a protamine gene from the chum salmon (*Oncorhynchus keta*). The taxonomic relationships between *O. keta* and *S. gairdneri* and other salmonid fish species are indicated in Fig. 1.

The chum salmon protamine gene encodes a protein identical to one of those found in the rainbow trout. Moreover, it is very homologous to the rainbow trout genes in the 5′ and 3′ nontranscribed regions, including the conservation of two different

Offprint requests to: G.H. Dixon

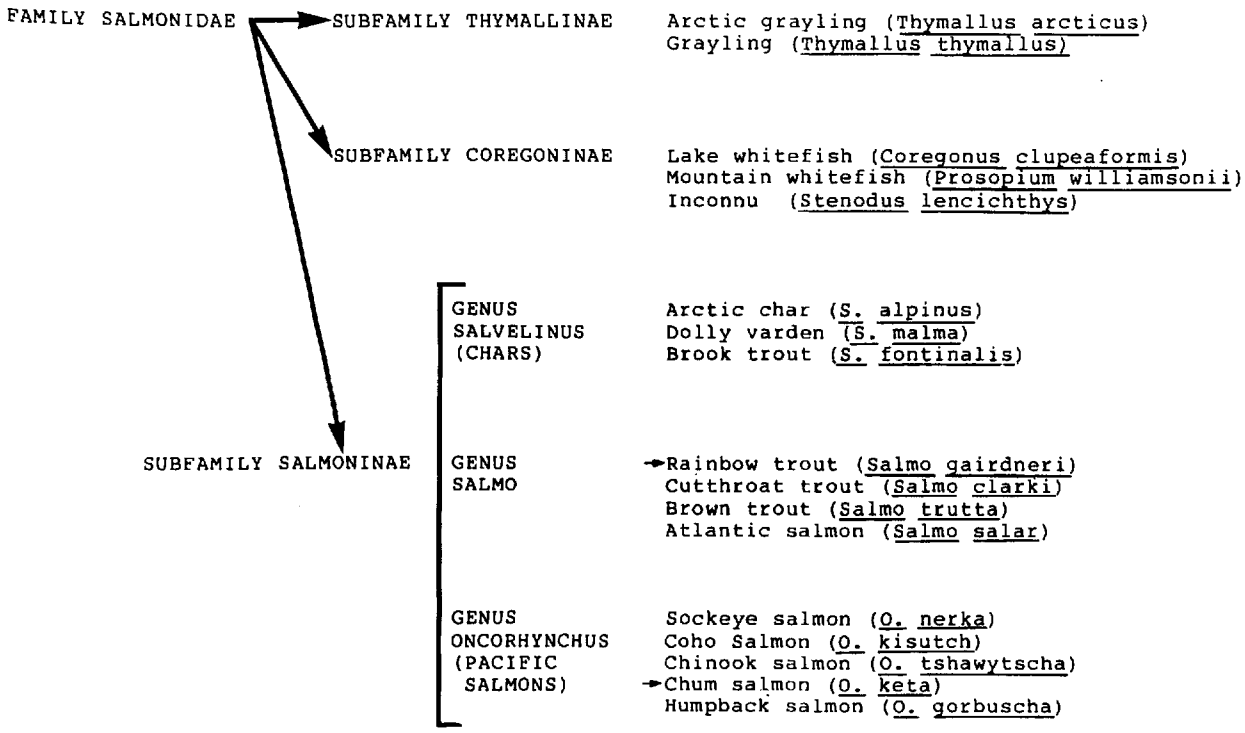

Fig. 1. Taxonomic classification of the species of the family Salmonidae. Drawn after Scott and Crossman (1973). The rainbow trout (*Salmo gairdneri*) and chum salmon (*Oncorhynchus keta*) are indicated by small arrows.

repetitive DNA elements that are found in the 5′ flanking regions of one of the rainbow trout genes.

Materials and Methods

Restriction endonucleases and other enzymes were purchased from New England Biolabs or Pharmacia P.L. Biochemicals and were used according to the manufacturers' instructions. Calf alkaline phosphatase and proteinase K were purchased from Boehringer Mannheim. Radionucleotides were obtained from Amersham. Membranes for Southern and dot blotting were also from Amersham. Membranes for plaque and colony screening came from New England Nuclear. Chemcials were purchased from Fisher or Sigma.

Isolation of Genomic DNA. Genomic DNA samples were purified essentially as described by Kaiser and Murray (1985). Following homogenization of tissue, either testis or liver, in the Hewish–Burgoyne buffer (Burgoyne et al. 1970) the nuclei were pelleted, then resuspended in 10 mM Tris, 1 mM EDTA (pH = 8.0) (TE), and lysed by the addition of 10% Sarcosyl to a final concentration of 1%. Solid CsCl was added to 55% (w/v) and the lysate centrifuged in an appropriate rotor to band the DNA. The DNA was collected by puncturing the centrifuge tube with a 16-gauge needle and was dialyzed extensively against TE.

Southern Blotting and Hybridization Techniques. Southern blotting was done according to established protocols (Southern 1975; Maniatis et al. 1982). Agarose gels were prepared in Tris-acetate buffer with 0.5 μg/ml ethidium bromide in both the gel and running buffer (Maniatis et al. 1982). The gels were treated with 0.5 N NaOH, 1.5 M NaCl for 1 h, then 0.5 M Tris-HCl, pH = 8.0, 1.5 M NaCl for 1 h before blotting. The transfer buffer was 20 × standard saline citrate (SSC) (1 × SSC = 0.015 M sodium citrate, 0.15 M sodium chloride). The filters were irradiated after transfer under long-wave ultraviolet light as described by the manufacturer (Amersham).

The hybridizations were done at 42°C in 50% formamide, 6 × SSC, 5 × Denhardt's solution (1 × Denhardt's = 0.02% bovine serum albumin, 0.02% Ficoll 400, 0.02% polyvinylpyrrolidone), 0.5% sodium lauryl sulfate, and 100 μg/ml sonicated calf thymus DNA. The filters were washed in several changes of 2 × SSC, 0.1% sodium lauryl sulfate at 65°C. Radiolabeled probes were prepared using the random primer method described by Feinberg and Vogelstein (1983).

Construction and Screening of O. keta Genomic Library. Oncorhynchus keta DNA (300 μg) was digested to completion with EcoRI, electrophoresed in a 0.8% agarose gel, and the DNA in the size range 12–20 kb purified by electroelution into dialysis bags followed by extraction with phenol and chloroform. This DNA was ligated to the arms of the lambda vector Charon 4A, also purified by gel electrophoresis, in a 2:1 molar ratio. The DNA was packaged in vitro using a commercial lambda packaging kit (Bethesda Research Laboratories). The recombinant clones were screened directly for protamine genes using Gene Screen membrane circles as described by the manufacturer (New England Nuclear). Positive clones were identified and replated at a lower density to allow purification.

Isolation and Characterization of Bacteriophage and Plasmid DNAs. Purified clones were used to infect a liquid bacterial culture and the phage harvested by polyethylene glycol precipitation after 12–16 h of growth. The phage were purified by equilibrium density gradient centrifugation in CsCl (Maniatis et al. 1982). DNA was isolated by proteinase K digestion followed by extractions with phenol and precipitation with ethanol. The phage DNA was mapped using several restriction enzymes. Fragments that hybridized to a protamine gene probe were isolated, ligated into pBR322, and introduced into bacteria by the $CaCl_2$ transfor-

mation method (Maniatis et al. 1982). Plasmid DNA was isolated using the alkaline lysis method of Birnboim and Doly (1979) followed by density gradient centrifugation. The plasmids were mapped using several restriction enzymes and the insert sequenced by the method of Maxam and Gilbert (1980) using DNA fragments that had been end-labeled with γ^{32}P-dATP and T4 polynucleotide kinase.

Results

Cloning of a Protamine Gene from the Chum Salmon

Southern blot analysis of salmonid genomic DNAs using a rainbow trout protamine gene probe revealed that digestion of chum salmon DNA with EcoRI gave two bands in the size range 12–20 kb and a band at 1.5 kb (Fig. 2). The large bands are within the size range for foreign DNA inserts in the lambda vector Charon 4A. A partial genomic library containing this DNA fraction was made in the Charon 4A vector and screened using a rainbow trout protamine gene probe. The probe was the 920-bp BglII–BamH1 fragment from p101 (States et al. 1982). It contained a substantial amount of 5′ and 3′ flanking DNA. Fifteen positives were initially identified from 350,000 recombinants, but upon further purification only four of these clones hybridized to a probe (AvaII–HpaII) derived from p101 and containing mostly the p101 protamine gene coding region. The characteristics of the 11 remaining clones are discussed in Moir and Dixon (1988). All four of the putative protamine clones had the same restriction map. There was a single 13-kb insert with no HindIII sites and a single BamH1 site (Fig. 3). The protamine gene was localized to a 2.5-kb BamH1–EcoRI fragment. This fragment was subcloned into pBR322 from the clone λC241 to give the plasmid pC241 (Fig. 3). The entire nucleotide sequence of this fragment was determined using the Maxam–Gilbert method (Fig. 4).

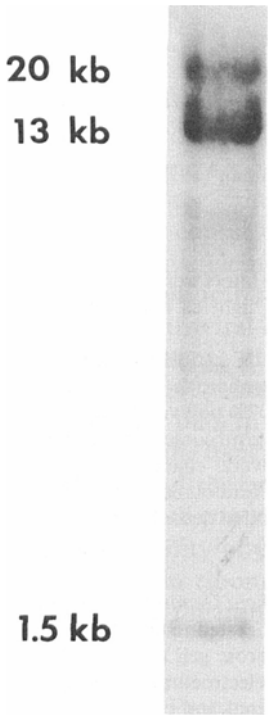

Fig. 2. Autoradiogram of a Southern blot of chum salmon DNA digested with EcoRI and probed with a rainbow trout protamine gene-specific probe. The sizes of three major bands are indicated.

Characteristics of the Chum Salmon Protamine Gene

The chum salmon protamine gene can be localized to the NcoI–BamH1 fragment in Fig. 3. Both strands were sequenced within this region. This fragment has an open reading frame that codes for a protein identical to the protein for the rainbow trout protamine gene, p101 (States et al. 1982). The nucleotide sequence and the predicted amino acid sequence are indicated in Fig. 4. The mRNA initiation and termination sites are placed on the basis of homology with the sites determined for the rainbow trout p101 gene. We have not been able to isolate RNA of adequate quality from chum salmon testis to use in S1 nuclease or primer extension analysis. Also present in the appropriate locations are the TATA box and polyadenylation signal (Breathnach and Chambon 1981). Downstream (120 bp) of the transcrip-

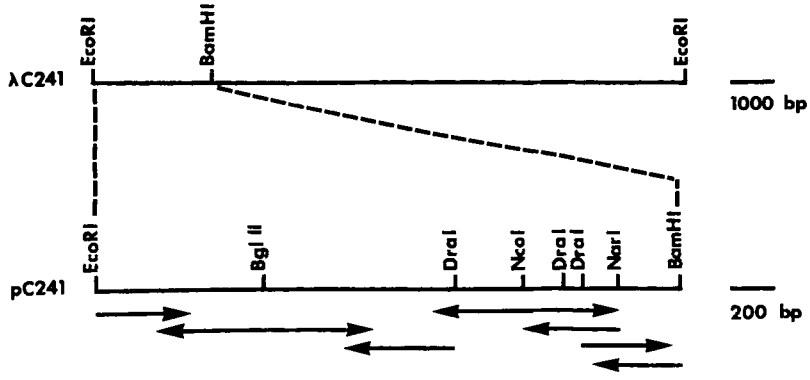

Fig. 3. Restriction maps of the chum salmon protamine clone, λC241, and the subclone, pC241, of the 2.5-kb EcoRI–BamHI fragment. The sequencing strategy is indicated by arrows.

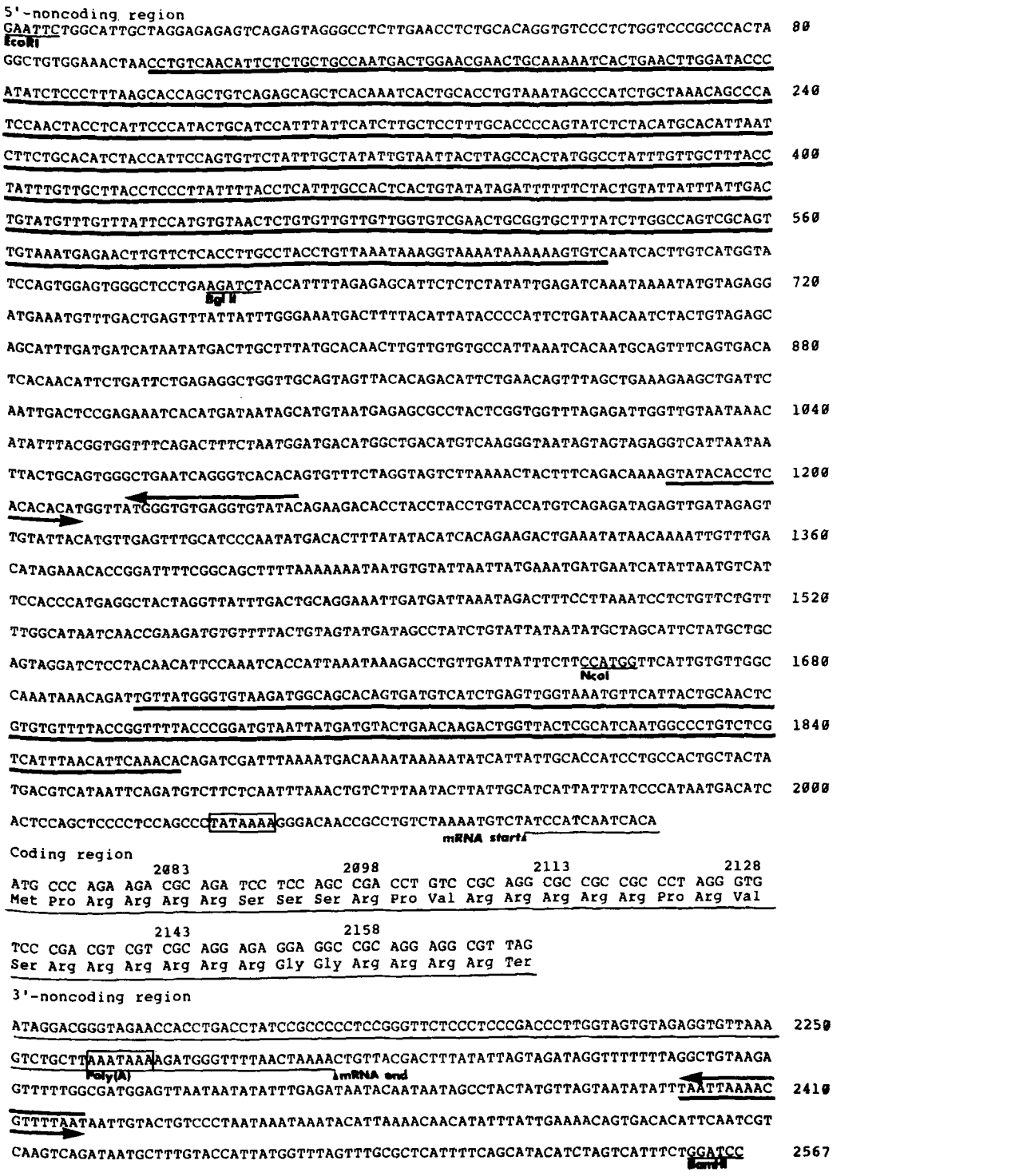

Fig. 4. The nucleotide sequence for the 2.5-kb EcoRI–BamHI fragment from Fig. 3. The two repetitive elements described in the text are underlined by thick lines. The protamine gene transcription unit is underlined by thin lines. The predicted amino acids resulting from translation of an open reading frame in the protamine gene are shown. Other sequences, such as inverted repeats (arrows) or TATAA and polyadenylation signal (boxes), and important restriction sites (Fig. 3) are also indicated.

tion termination site there is a 20-bp palindromic sequence that contains exclusively AT base pairs except for two GC base pairs in the middle of the sequence (indicated by oppositely oriented arrows above and below the DNA sequence at this position in Fig. 4).

There are several interesting features located upstream of the gene. An inverted repeat is present 800 bp 5′ to the gene that could potentially extrude a cruciform with a stem of 18 bp and 4-bp loops. There is only a single mismatch in the stem structure. There are also two different repetitive elements

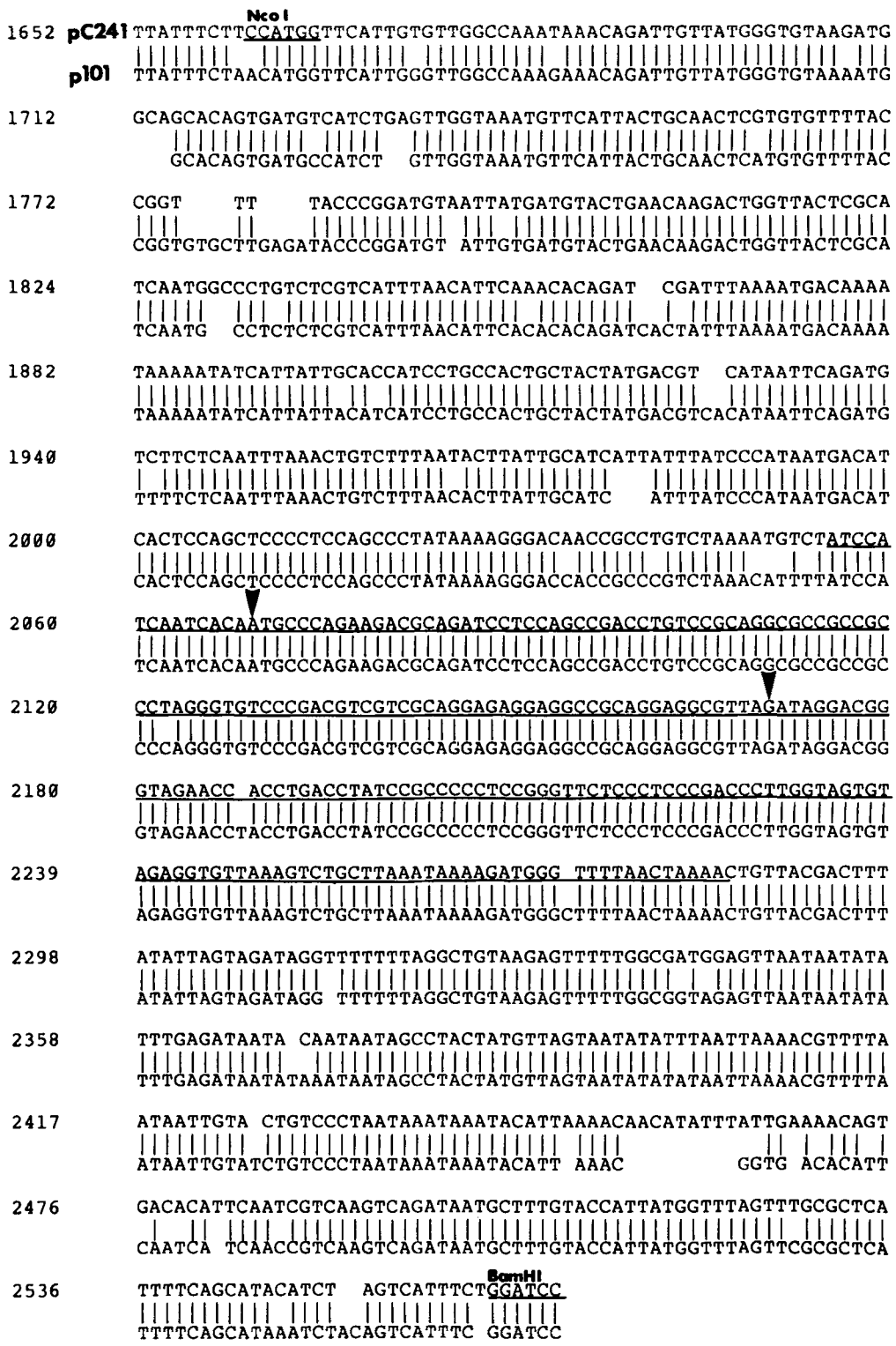

Fig. 5. Alignment of the chum salmon protamine gene, pC241 (**upper sequence**), and the rainbow trout p101 gene (**lower sequence**) from the NcoI–BamHI sites. Matched bases are joined by a vertical line and gaps have been introduced to maximize homology. The underlined region is the transcription unit and the arrows indicate the beginning and end of translation.

in the upstream region. The first sequence, between positions 95–623 (underlined in Fig. 4), is a very high copy number repeat. This element has been well characterized in the rainbow trout and the details will be presented elsewhere (Winkfein et al., in preparation). The limits of the repeat have been defined by comparing the sequence of several different clones. The repeat has a polyA sequence at its 3′ terminus and may, therefore, be an Alu-like retrotransposon (Weiner et al. 1986).

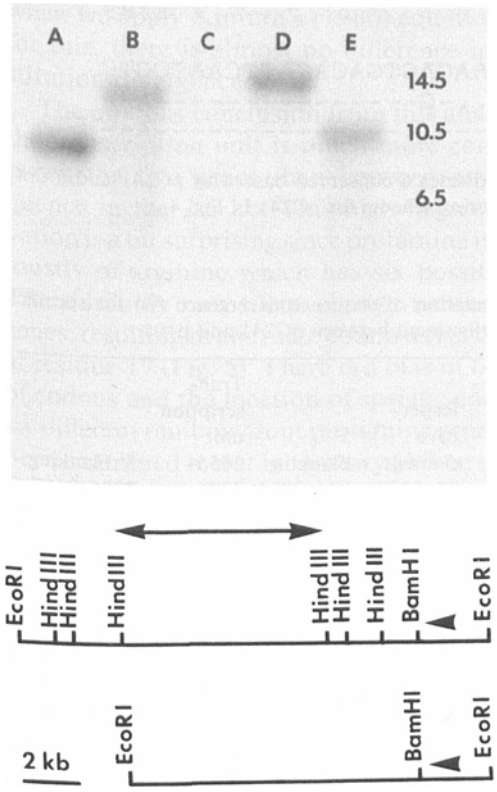

Fig. 6. Autoradiogram of Southern blot of lambda clones probed with a probe from a region downstream of λTP101. In the lower part of the figure the restriction maps of λC241 **(lower map)** and λTP101 **(upper map)** are shown with the probe fragment from λTP101 indicated by large arrows. The small arrows indicate the protamine genes and the direction of transcription. In the upper part of the figure the Southern blot is shown with the sizes of molecular weight markers in kilobases. Lane A, λTP101 digested with EcoRI and HindIII; lane B, λC241 digested with EcoRI and HindIII; lane C, Charon 4A digested with EcoRI and HindIII; lane D, λTP101 digested with EcoRI and BamHI; lane E, λC241 digested with EcoRI and BamHI.

A second, much shorter, lower-copy repeat is located at positions 1706–1859 (underlined) in Fig. 4. This repeat is also in the same position in the rainbow trout gene, p101, that was used as a probe to screen the chum salmon library. The 11 clones that hybridized to the 920-bp BglII–BamH1 rainbow trout gene probe but not to the coding region probe apparently contain only this repeated sequence element but not the protamine gene. We have sequenced two of these clones and the detailed characterization of this repeat is presented in Moir and Dixon (1988). The repeat, defined by homology with other clones, has the trinucleotide TGT and its complement ACA at its 5′ and 3′ limits, respectively. The same pentanucleotide, CAGAT, is repeated at each end of the chum salmon and rainbow trout protamine representatives of this repeated DNA family. These features are reminiscent of retroviral-like transposable elements (Weiner et al. 1986).

Alignment of Sequences

The chum salmon protamine gene, pC241, was aligned with the rainbow trout gene, p101 (States et al. 1982), using the Align program of the Microgenie sequence software package (Queen and Korn 1984). The p101 gene was chosen over the other rainbow trout genes (Aiken et al. 1983) because p101 and pC241 both code for the same protamine which is slightly different from the other protamines (McKay et al. 1986). Furthermore, a substantial amount of 5′ flanking information is available for p101 (R.J. Winkfein, personal communication; R.D. Moir, unpublished). In Fig. 5 an alignment from approximately the NcoI to BamH1 site of pC241 is presented. However, the strong homology that is evident in this figure is maintained between the two sequences up to the EcoRI site (Fig. 3), a distance of 2 kb upstream from the protamine gene. This homology includes the conservation of the two repetitive elements.

Downstream Homology between pC241 and p101

Although there was very strong upstream homology between the two genes, pC241 and p101, an examination of the restriction maps of the lambda clones showed several restriction site differences downstream from the genes (Fig. 6). We isolated an 8-kb HindIII fragment from λTP101 starting 3 kb downstream from the end of the protamine gene and used it as a probe in the Southern blot of λC241 DNA. The clone λC241 hybridizes as well as the control (λTP101) suggesting that at least some of the sequences identical between the two clones are maintained this far downstream (Fig. 6).

Discussion

We have described the structural characteristics of a protamine gene from the chum salmon (*Oncorhynchus keta*). This gene is very similar to the rainbow trout (*Salmo gairdneri*) protamine gene p101 (States et al. 1982) and we conclude that these two genes must have originated from the same ancestral gene and have undergone very little divergence. The rainbow trout clone, p101, and the chum salmon clone, pC241, code for exactly the same protein. The other protamines from the rainbow trout, although very similar, show slight amino acid sequence differences (McKay et al. 1986). In addition, p101 and pC241 have two conserved repetitive DNA sequences located at exactly the same position with respect to the gene. One of these repeats is apparently an Alu-like retrotransposon (Weiner et al. 1986) and is immediately adjacent (R.J. Winkfein, personal communication) to only one of the other rain-

```
          2434                                                    2489

pC241     TAAATAAATACATTAAAACAACATAT.T.TATTGAAAACAGTGACACATTCAATCGTC
pl01      ----------------.....................-G--------------A--
p21       --------------------..---A-G----------..-------------A--
```

Fig. 7. Alignment of a region downstream in pC241, p101, and p21. A dash indicates a conserved base with pC241 and a dot indicates a gap introduced to maximize homology. The numbers refer to the numbering scheme for pC241 in Fig. 4.

bow trout protamine clones previously characterized (Aiken et al. 1983). The second repetitive element has some similarities to a retroviral long terminal repeat (LTR) and is not adjacent to any of the other trout protamine clones (Moir and Dixon 1988). Finally, the two clones share at least some common sequences several kilobases downstream as judged by hybridization (Fig. 6). A protamine with the amino acid sequence predicted by pC241 has been isolated from a mature chum salmon testis (D.J. McKay, personal communication). In addition, three other protamines have been sequenced from this testis. The Southern blot in Fig. 2 shows two other bands, in addition to the 13-kb band corresponding to λC241 and these other bands (1.5 and 20.0 kb) may correspond to the genes for these other proteins.

Despite the strong conservation of sequence, there are some interesting differences between the two genes. Two transcription start sites have been mapped for the p101 gene (Gregory et al. 1982; States et al. 1982). In addition to the conserved site marked in Fig. 4, a second site has been mapped by S1 nuclease analyses 5 bp farther upstream (States et al. 1982). The chum salmon sequence has been slightly altered at this point so that there is no longer an A nucleotide at the upstream site. Since transcription usually starts with an A (Breathnach and Chambon 1981), the second start site for the chum salmon gene may be 1 bp farther upstream. We have not been able to map the salmon gene start sites. In any case, messages starting from the two upstream sites would be slightly different for these two genes for the first few base pairs.

A second interesting difference between p101 and pC241 results when the salmon gene is compared to the other rainbow trout clones that have been sequenced in our laboratory (Aiken et al. 1983). An apparent 10-bp deletion has occurred in the p101 gene relative to the pC241 gene in the region downstream of transcription termination (Fig. 5). However, in the trout clone, p21 (and the very similar clone p16), this sequence has been almost perfectly retained (Fig. 7). There are several other examples of deletions downstream between the other rainbow trout protamine genes (Aiken et al. 1983), although the effect on transcription is not known. The rainbow trout genes are very homologous to each other in the 5′ region up to the point of the insertion of the LTR-like repeat sequence. Beyond this point there are little sequence data available (Aiken et al. 1983).

Table 1. Calculation of sequence divergence (%) for specific regions of the alignment between pC241 and p101

	Repetitive element 95–623	5′ Flanking 624–2054	Transcription unit 2055–2285	3′ Flanking 2286–2567
Base substitutions	6.9	4.2	0.5	5.6
Total	9.8	6.5	1.3	8.5

The formula $r = -\frac{3}{4}\ln[1 - \frac{4}{3}(dv/ds)]$ was used (Jukes and Cantor 1969), where r is the sequence divergence between the two sequences, dv is the number of differences between the sequences, and ds is the number of aligned bases. In the upper row, dv was taken as simply base substitutions and in the lower row, dv was taken as both base substitutions and insertions/deletions. The beginning and end for the specific regions are indicated by the nucleotide positions of pC241

Since p101 and pC241 apparently share the same common ancestor, it is possible to compare the two sequences directly in order to derive an estimate of sequence divergence (Nei 1986). In Fig. 5, an alignment from only the NcoI to BamHI sites is presented to illustrate the strong conservation of the sequences. However, we have done a complete alignment of the 2.5 kb that are apparently conserved between the genes and we have calculated the amount of sequence divergence for specific regions in this alignment using the equation of Jukes and Cantor (1969) to account for the possibility of multiple substitutions at the same site. The results are presented in Table 1. Two values for each region of the sequence are presented. The lower value is the divergence due only to base substitutions, while the higher value is the total divergence if we include the numerous single or multiple base insertions or deletions as single differences between the two sequences. In fact, deletion or insertion of a single or several contiguous bases is more likely to occur than any one of the possible base substitutions. There is a slight bias favoring transitions over transversions (not shown), but this is not as great as is found in mitochondrial DNAs (Hixson and Brown 1986), and

when we apply Kimura's (1980) equations to allow for this, there is almost no difference in base substitution divergence.

The obvious conclusion from this analysis is that the transcription unit is much more conserved between the two genes than between any other sequence in the 2.5 kb. This high degree of conservation is a bit surprising since protamine is composed mostly of arginine which has six possible codons. There is only one silent change between the two genes, resulting in alternate codon usage for a proline at residue 17 (Fig. 5). There is a bias in both the use of codons and the location of specific codons in the six different rainbow trout protamine genes that have been examined in our laboratory (Aiken et al. 1983). This could suggest that conservation of the RNA sequence is as important as that of the amino acid sequence. Protamine mRNA in the trout is likely to have extensive secondary structure based on computer modeling (Krawetz et al. 1987) and resistance to ribonuclease T1 (Davies et al. 1979). Such a stable structure may be important for the stability of the protamine messenger ribonucleoprotein particle which is translationally inactive and stored for several weeks in the cell sap of trout spermatocytes and early spermatids (Iatrou et al. 1978; Sinclair and Dixon 1982).

Thomas et al. (1986) have constructed a phylogenetic tree of the rainbow trout and the five Pacific salmon (the genus *Oncorhynchus*; Fig. 1) using the proportion of common and different restriction enzyme sites in mitochondrial DNA. The degree of divergence between two species was estimated using the maximum likelihood method of Nei and Tajima (1983). The pairwise comparison between the chum salmon and the rainbow trout gave an estimated degree of sequence divergence of 7.1%. Although we have used a different method of calculating divergence, this value is in reasonably good agreement with the values we obtain for the non-transcribed regions of the protamine genes which range from 6.5 to 9.8% (Table 1). This could suggest that the flanking regions, particularly the Alu-like retroposon, are as free to diverge as mitochondrial DNA. However, it would be clearly preferable to directly compare the DNA sequences of these different types of DNA sequences between these two species. There is limited information on the time since establishment of *Salmo* and *Oncorhynchus* as distinct genera in the subfamily Salmonidae and consequently it is difficult to use the degree of sequence divergence to calculate a rate of divergence (Thomas et al. 1986).

Acknowledgments. This work was supported by a Medical Research Council of Canada Term Operating Grant to G.H.D. R.D.M. was the recipient of a studentship from the Alberta Heritage Foundation for Medical Research. We thank Marnie Cudmore for expert assistance with the manuscript.

References

Aiken JM, McKenzie D, Zhao H-Z, States JC, Dixon GH (1983) Sequence homologies in the protamine gene family of rainbow trout. Nucleic Acids Res 10:7581–7592

Ando T, Yamasaki M, Suzuki K (1973) Protamines, isolation, characterization, structure and function. In: Kleinzeller A, Springer GF, Wittmann HG (eds) Molecular biology, biochemistry and biophysics, vol 12. Springer, New York

Birnboim HC, Doly J (1979) A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res 7:1513–1523

Bloch D (1969) A catalog of sperm histones. Genetics (Suppl) 61:93–111

Breathnach R, Chambon P (1981) Organization and expression of eukaryotic split genes coding for proteins. Annu Rev Biochem 50:379–384

Burgoyne LA, Waqar MA, Atkinson MR (1970) Calcium-dependent priming of DNA synthesis in isolated rat liver nuclei. Biochem Biophys Res Commun 39:254–259

Connor W, Mezquita J, Winkfein RJ, States JC, Dixon GH (1984) Organization of the histone genes in the rainbow trout. J Mol Evol 20:227–235

Davies PL, Dixon GH, Simoncsits A, Brownlee GG (1979) Sequences of large T1 ribonuclease resistant oligoribonucleotides from protamine mRNA: the overall architecture of protamine mRNA. Nucleic Acids Res 7:2323–2345

Dixon GH, Aiken JM, Jankowski JM, McKenzie DI, Moir R, States JC (1986) Organization and evolution of the protamine genes of the salmonid fishes. In: Reeck GR, Goodwin GA, Puigdomenech P (eds) Chromosomal proteins and gene expression. Plenum, New York, pp 287–314

Feinberg A, Vogelstein B (1983) A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal Biochem 132:6–13

Gregory SP, Dillon NO, Butterworth PHW (1982) The localisation of the 5'-termini of *in vivo* and *in vitro* transcripts of a cloned rainbow trout protamine gene. Nucleic Acids Res 10:7581

Hixson JE, Brown WM (1986) A comparison of the small ribosomal genes from the mitochondrial DNA of the great apes and humans: sequence, structure, evolution and phylogenetic implications. Mol Biol Evol 3:1–18

Iatrou K, Spira AW, Dixon GH (1978) Protamine messenger RNA: evidence for early synthesis and accumulation during spermatogenesis in rainbow trout. Dev Biol 64:82–98

Jukes TH, Cantor CR (1969) Evolution of protein molecules. In: Munro HN (ed) Mammalian protein metabolism. Academic Press, New York, pp 21–123

Kaiser K, Murray NE (1985) The use of phage lambda replacement vectors in the construction of representative genomic DNA libraries. In: Glover DM (ed) DNA cloning, vol 1. A practical approach. IRL Press, New York, pp 1–48

Kasinsky HE, Mann M, Lemke M, Huang S-Y (1986) Diversity of sperm basic chromosomal proteins: a phylogenetic point of view. In: Reeck GR, Goodwin GA, Puigdomenech P (eds) Chromosomal proteins and gene expression. Plenum, New York, pp 333–352

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

Krawetz SA, Connor W, Dixon GH (1987) Cloning of bull P1

protamine mRNA and the evolution of vertebrate P1 protamines. DNA 6:47–57

Maniatis T, Fritsch E, Sambrook J (1982) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor NY

Maxam AM, Gilbert W (1980) Sequencing end-labeled DNA with base specific chemical cleavages. Methods Enzymol 65: 499–559

McKay DJ, Renaux BS, Dixon GH (1986) Rainbow trout protamines: amino acid sequences of six distinct proteins from a single testis. Eur J Biochem 158:361–366

Moir RD, Dixon GH (1988) A repetitive DNA sequence in the salmonid fishes similar to a retroviral long terminal repeat. J Mol Evol 27:1–7

Nei M (1986) Stochastic errors in DNA evolution and molecular phylogeny. In: Gershowitz H, Rucknagel DL, Tashian RE (eds) Evolutionary perspectives and the new genetics progress in clinical and biological research, vol 218. Alan R Liss, New York, pp 133–147

Nei M, Tajima F (1983) Maximum likelihood estimation of the number of nucleotide substitutions from restriction sites data. Genetics 105:207–217

Queen C, Korn LJ (1984) A comprehensive sequence analysis program for the IBM personal computer. Nucleic Acids Res 12:5j81–599

Scott WB, Crossman EJ (1973) Freshwater fishes of Canada. Bryant Press Ltd, Ottawa

Sinclair GD, Dixon GH (1982) Purification and characterization of cytoplasmic protamine messenger particles from rainbow trout testis cells. Biochemistry 21:1869–1877

Southern E (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. J Mol Biol 98: 503–517

States JC, Connor W, Wosnick MA, Aiken JM, Gedamu L, Dixon GH (1982) Nucleotide sequence of a protamine component CII gene of *Salmo gairdneri*. Nucleic Acids Res 10:4551–4563

Subirana JA (1975) On the biological role of basic proteins in spermatozoa and during spermiogenesis. In: Duckett JG, Racey PA (eds) The biology of the male gamete. Academic Press, London, pp 239–244

Thomas WK, Withler RE, Beckenbach AT (1986) Mitochondrial DNA analysis of Pacific salmonid evolution. Can J Zool 64:1058–1064

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

Received February 20, 1987/Revised and accepted July 10, 1987