Vol. 55: 187–203, 2003 | DISEASES OF AQUATIC ORGANISMS
Dis Aquat Org | Published August 4

Two distinct phylogenetic clades of infectious hematopoietic necrosis virus overlap within the Columbia River basin

Kyle A. Garver[1,2], Ryan M. Troyer[1,2,3], Gael Kurath[1,2,*]

[1]Department of Pathobiology, University of Washington, Seattle, Washington 98195, USA
[2]Western Fisheries Research Center, United States Geological Survey, Biological Resources Division, 6505 NE 65th St., Seattle, Washington 98115, USA
[3]*Present address:* Division of Infectious Diseases, Case Western Reserve University, Cleveland, Ohio 44118, USA

ABSTRACT: Infectious hematopoietic necrosis virus (IHNV), an aquatic rhabdovirus, causes a highly lethal disease of salmonid fish in North America. To evaluate the genetic diversity of IHNV from throughout the Columbia River basin, excluding the Hagerman Valley, Idaho, the sequences of a 303 nt region of the glycoprotein gene (mid-G) of 120 virus isolates were determined. Sequence comparisons revealed 30 different sequence types, with a maximum nucleotide diversity of 7.3% (22 mismatches) and an intrapopulational nucleotide diversity of 0.018. This indicates that the genetic diversity of IHNV within the Columbia River basin is 3-fold higher than in Alaska, but 2-fold lower than in the Hagerman Valley, Idaho. Phylogenetic analyses separated the Columbia River basin IHNV isolates into 2 major clades, designated U and M. The 2 clades geographically overlapped within the lower Columbia River basin and in the lower Snake River and tributaries, while the upper Columbia River basin had only U clade and the upper Snake River basin had only M clade virus types. These results suggest that there are co-circulating lineages of IHNV present within specific areas of the Columbia River basin. The epidemiological significance of these findings provided insight into viral traffic patterns exhibited by IHNV in the Columbia River basin, with specific relevance to how the Columbia River basin IHNV types were related to those in the Hagerman Valley. These analyses indicate that there have likely been 2 historical events in which Hagerman Valley IHNV types were introduced and became established in the lower Columbia River basin. However, the data also clearly indicates that the Hagerman Valley is not a continuous source of waterborne virus infecting salmonid stocks downstream.

KEY WORDS: Columbia River basin · IHNV · Rhabdovirus · Epidemiology · Salmonids

INTRODUCTION

Infectious hematopoietic necrosis virus (IHNV) is an aquatic rhabdovirus that causes serious disease in salmonid fish. The infectious range of IHNV extends primarily throughout the Pacific Northwest region of North America, and the virus is considered endemic from Alaska to California and inland to Idaho (Wolf 1988, p. 83–114). In addition to the endemic watersheds of western North America, IHNV has been spread by the movement of fish and eggs to Asia and Europe, where it is present in rainbow trout reared in aquaculture (Winton 1991).

IHNV caused extensive mortalities in hatchery fish populations during the 1950s in the states of Washington (Rucker et al. 1953), Oregon (Wingfield et al. 1969), and California (Ross et al. 1960, Wingfield et al. 1970), with the first reported epidemics occurring in sockeye salmon in the Columbia River basin (Rucker et al. 1953). These first epidemics of IHNV were presumably caused by feeding young fish unpasteurized sockeye salmon *Oncorhynchus nerka* viscera (Watson et al. 1954, Wolf 1988), a practice that was stopped by the late 1960s. With this change in the feeding practice, and also with the replacement of sockeye salmon cul-

*Corresponding author. Email: gael_kurath@usgs.gov

ture with chinook salmon and steelhead culture, the occurrence of IHNV in the Columbia River basin became much less apparent. Prior to the late-1970s, IHNV was isolated only sporadically within the Columbia River basin. A survey conducted in 1970 by Amend and Wood for IHNV in Washington salmon indicated that IHNV was not present in stocks of chinook *O. tshawytscha* or coho *O. kisutch*, and only a portion of the sockeye salmon were found to be infected (Amend & Wood 1972). In Oregon, during the 1970s, IHN disease was observed in stocks of rainbow trout, steelhead trout, chinook salmon and kokanee salmon, but the virus was endemic to only 2 facilities (Mulcahy et al. 1980, Groberg & Fryer 1983, Groberg et al. 1983). In 1977, an area of the Columbia River basin located in Hagerman Valley, Idaho, began to experience high mortalities at rainbow trout aquaculture facilities (Busch 1983). The virus spread over the next 3 years to become endemic to the Hagerman Valley. During the subsequent years of 1980 to 1982, IHNV emerged in the anadromous salmonid hatcheries in the middle to lower Columbia River basin. During this time IHNV became much more widely disseminated throughout the middle and lower Columbia River basin, losses to the virus increased by greater than 10-fold (Groberg 1983), and IHNV was isolated from additional salmonid species. By 1982, IHNV had been isolated in cutthroat trout and there were very high losses in rainbow and steelhead trout in the Columbia River basin (Groberg 1983, Groberg & Fryer 1983).

Due to the prevalence of IHNV throughout its current infectious range, comparative studies of different isolates of IHNV have been conducted. These investigations, reporting the extent of heterogeneity among selected IHNV field isolates, relied upon using biological (Mulcahy et al. 1984), serological (Winton et al. 1988, Ristow & Arnzen 1989) and biochemical techniques such as electrophoretic mobility of structural proteins (Hsu et al. 1986), RNase T1 mapping (Oshima et al. 1995), and nucleotide sequencing (Nichol et al. 1995). These studies have indicated that different strains of IHNV exist and that phenotypic and genetic relatedness generally correlates with geographic origin.

Despite these findings, considerable gaps exist in our understanding of both the epidemiology and the genetic diversity of IHNV isolates throughout its range. This information is critical for management agencies faced with decisions regarding the movement of infected fish or the selection of stocks for rearing within a virus-endemic area. Therefore, in an effort to better understand the genetic diversity and epidemiology of IHNV, our laboratory is genetically characterizing isolates throughout the entire geographic range of IHNV. To this end, broad regional studies of IHNV diversity in Alaska (Emmenegger et al. 2000), British Columbia (Anderson, Traxler & Kurath unpubl.), and coastal Washington (Emmenegger & Kurath 2002) have been conducted. In addition, 2 localized studies have characterized the genetic heterogeneity of IHNV in the Deschutes River watershed (Anderson et al. 2000) and in the Hagerman Valley, Idaho (Troyer et al. 2000, Troyer & Kurath 2003 this issue). These studies demonstrated that significant variation in the evolutionary and epidemiological history of IHNV exists in different locations. For instance, the low genetic diversity of IHNV in Alaska contrasts with the high genetic diversity of IHNV within the Hagerman Valley, Idaho, where multiple co-circulating subclades were revealed (Troyer et. al. 2000). A summary of IHNV genetic diversity and phylogeny throughout North America has recently been described (Kurath et. al. 2003). These previous studies, in addition to illustrating differences in IHNV genetic diversity patterns throughout its geographic range, provide an extensive genetic diversity database to which subsequent studies can be compared. This report adds to this body of knowledge with an extended characterization of IHNV genetic diversity and epidemiology throughout the Columbia River basin.

At present, the Columbia River basin experiences significant losses due to IHNV and the widespread occurrence of the virus in the system poses a serious threat to susceptible species present in the wild, in hatcheries, and in aquaculture facilities. Thus, the objective of this study was to conduct a thorough analysis of virus types throughout the basin, excluding the Hagerman Valley, which has been characterized elsewhere (Troyer et al. 2000, Troyer & Kurath 2003). In this report, 120 virus isolates sampled from a total of 54 different locations over a 30 yr period were characterized (Fig. 1, Table 1) to provide insights into the IHN viral traffic patterns. It was of particular interest to determine how the genetic types of IHNV present in the Columbia River basin were related to those characterized in the Hagerman Valley study site (Troyer et al. 2000). Due to the geographic location of the Hagerman Valley within the Columbia River basin upstream of the lower Columbia River, it has been speculated that the Hagerman Valley may have been the source of the IHN virus that emerged in the lower Columbia River basin in 1980 to 1982 (Busch 1983, Groberg 1983, Hsu 1986).

MATERIAL AND METHODS

Virus isolates. Features of the 120 IHNV isolates characterized are shown in Table 1. The isolates were provided by Susan Gutenberger, Theresa London, Ray Brunson, Corie Samson, and Kathy Clemens (United

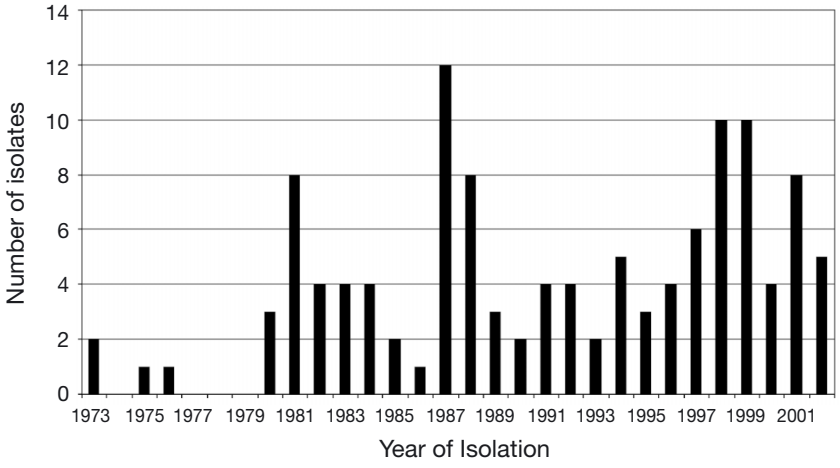

Fig. 1. Temporal distribution of the 120 Columbia River basin infectious hematopoietic necrosis virus (IHNV) isolates characterized in this study

States Fish and Wildlife Service), Joan Thomas (Washington Department of Fish and Wildlife), Mark Engelking, John Kaufman, Warren Groberg, Sam Onjukka, and Glenda Claire (Oregon Department of Fish and Wildlife), Keith Johnson and Sharon Landin (Idaho Department of Fish and Game), Dorothee Kieser (Department of Fisheries and Oceans, B.C.), and from the archival collection at the Western Fisheries Research Center in Seattle, Washington. These isolates were obtained from 6 salmonid fish species (Table 1), but with a predominance of isolates from chinook (50 out of 120) and steelhead (41 out of 120). They were generally taken from dead or moribund fish experiencing an epidemic or from asymptomatic adult fish surveyed at spawning. The viruses were isolated as described by Thoesen (1994) and stored as frozen aliquots of cell culture supernatant. These viruses were obtained at low passage levels with the majority of the isolates going through 2 or 3 passages in cell culture. This was done to avoid cell culture variants, although studies conducted on IHNV (Batts & Winton unpubl. data) and viral hemorrhagic septicemia (VHSV) (Jørgensen et. al. 1995) have shown that these viruses can remain extremely stable even after a high number of passages in cell culture.

Reverse transcription polymerase chain reaction (RT-PCR) amplification and sequence analysis. Viral genomic RNAs were utilized as templates for RT-PCR as described (Emmenegger et al. 2000). Briefly, 5 µl of a 1:20 dilution of cell culture supernatant for each viral isolate was heated to 95°C for 5 min and then combined with 1 µl of 20 pmol μl^{-1} of each first round IHNV glycoprotein (G)-gene specific primer (Table 2), 5 µl of 25 mM $MgCl_2$, 5 µl of 10¥ PCR buffer (Promega), 5 units avian myeloblastosis virus (AMV) reverse transcriptase (Promega), and 2.5 units Taq to produce a 50 µl reaction. The reaction was then incubated at 50°C for 1 h followed by heating to 94°C for 2 min and 25 subsequent cycles of 94°C for 30 s, 50°C for 30 s, and 72°C for 30 s, followed by a final extension of 72°C for 7 min. Two µl of the RT-PCR reaction was then utilized as template for the second round PCR. Reaction conditions for the second round PCR were identical to those used for the RT-PCR reaction without the addition of AMV reverse transcriptase or initial 50°C incubation. The sequences and binding locations of the first and second round PCR primers utilized for the G-gene amplification are shown in Table 2. The sequenced product of the second round PCR of the G-gene is a 303 nucleotide (nt) region located in the middle of the G-gene, from nucleotide 686 to 988 (numbering as in GenBank U50401), and is denoted as the mid-G region. The mid-G region was sequenced using a fluorescent dye terminator cycle sequencing kit (Applied Biosystems) using second round PCR primers and following manufacturer's protocols.

RT-PCR amplification and sequence analysis of the partial nucleoprotein (N) gene were carried out using identical reaction conditions to those described for the mid G-gene with the exception of using N-gene specific primers. The sequences and binding locations of the first and second round PCR primers utilized for N-gene amplification are shown in Table 2. The sequenced product of the second round PCR of the nucleoprotein gene is a 412 nt region located near the 5′ terminus of the gene and is denoted as the 5′N region. This region corresponds to nucleotide 133 to 544 on the full length Western Regional Aquaculture Consortium (WRAC) strain IHNV genome sequence by Morzunov et al. (1995) (GenBank L40883).

Phylogenetic analysis. Sequence files were edited and analyzed using MacVector 6.0 and AssemblyLIGN 1.0.9 applications (Oxford Molecular Group). Sequence files utilized in phylogenetic analyses were configured in SeqPup 6 P/C and phylogenies were performed using PAUP* version 4.0b (Swofford 1998). The Sacramento River Chinook virus (SRCV), Coleman-80 (Col80), and Coleman-85 (Col85) isolates of IHNV, isolated in California in 1966, 1980 and 1985 respectively (Nichol et al. 1995), were used as an outgroup root. It is well established that these California isolates are phylogenetically distinct from isolates from the Columbia River basin IHNV range (Nichol et al. 1995, Emmenegger et al. 2000, Troyer et al. 2000, Emmen-

Table 1. Summary of IHNV isolates from the Columbia River basin. Ck, creek; H, hatchery; R, river. Ct, sea-run cutthroat trout; St, steelhead; Rbt, rainbow trout; Ch, chinook; Sk, sockeye; Kok, kokanee. Mid-G sequence types are designated with the first letter representing the major phylogenetic clade that the sequence groups into, and where applicable the M sub-clade is denoted by the second letter. All sequence types names carry the Columbia River basin abbreviation, crb, followed by an arbitrarily assigned sequence type number along with a 'n' number that indicates the number of isolates from this study having that identical mid-G sequence. Sequence types without a designated number (n) represent that one isolate only

Major river drainage[a]	Isolation site	Year	Host	Life stage	Mid-G Seq. type, n#	5'N isolate no.[b]
Lower Columbia						
Columbia	Beaver Ck	1981	Ct	Adult	M-crb1, n24	
Cowlitz	Cowlitz H.	1981	St	Adult	M-crb1, n24	
Cowlitz	Cowlitz H.	1981	Rbt	Yearling	M-crb1, n24	
Cowlitz	Cowlitz H.	1981	Ct	Adult	M-crb1, n24	
Cowlitz	Cowlitz H.	1982	St	Fry	M-crb1, n24	
Cowlitz	Cowlitz H.	1983	Ct	Adult	M-crb1, n24	
Cowlitz	Cowlitz H.	1987	St	Adult	U-crb1, n52	
Lewis	Merwin H.	1994	St	Adult	U-crb1, n52	1
Lewis	Merwin H.	1995	St	Adult	MD-crb2,n6	
Lewis	Merwin H.	1998	St	Adult	U-crb1, n52	2
Lewis	Merwin H.	1999	St	Adult	MD-crb2,n6	3
Lewis	Merwin H.	1999	Rbt	Fry	MD-crb2,n6	4
Lewis	Lewis R. H.	1973	Ch	Fry	U-crb2	5
Lewis	Lewis R. H.	1980	Ch	Adult	M-crb1, n24	6
Lewis	Lewis R. H.	1996	Ch	Juvenile	U-crb1, n52	
Lewis	Speelyai H.	1981	Kok	Adult	M-crb1, n24	
Lewis	Speelyai H.	1981	Ch	Adult	M-crb1, n24	
Lewis	Speelyai H.	1997	Ch	Adult	MD-crb2,n6	
Lewis	Speelyai H.	1998	Kok	Adult	U-crb1, n52	
Lewis	Speelyai H.	1999	Kok	Adult	U-crb1, n52	7
Lewis	East Fork Lewis R.	1997	St	Fingerling	U-crb1, n52	
Washougal	Skamania H.	1981	St	Fry	M-crb1, n24	
Washougal	Skamania H.	1983	Ct	Fry	M-crb1, n24	
Washougal	Skamania H.	1983	St	Smolt	M-crb1, n24	
Willamette	Clackamas H.	1987	Ch	Adult	U-crb1, n52	
Willamette	Clackamas H.	1996	St	Adult	MD-crb2,n6	
Willamette	Clackamas H.	1997	St	Adult	MD-crb3,n3	
Willamette	Clackamas H.	1999	St	Adult	MD-crb2,n6	8
Willamette	Minto Pond	1987	St	Adult	M-crb1, n24	
Willamette	Minto Pond	1996	St	Adult	MD-crb3,n3	
Willamette	S. Santiam H.	1987	St	Adult	U-crb1, n52	
Willamette	S. Santiam H.	1994	St	Adult	M-crb1, n24	9
Willamette	S. Santiam H.	1999	Ch	Adult	MD-crb3,n3	10
Lower Mid-Columbia						
Wind	Carson H.	1987	Ch	Adult	U-crb1, n52	
Wind	Carson H.	1989	Ch	Fingerling	U-crb1, n52	11
Wind	Carson H.	1998	Ch	Adult	U-crb1, n52	12
Wind	Carson H.	2001	Ch	Adult	U-crb1, n52	
Columbia	Hamilton Ck.	2000	Chum	Adult	U-crb1, n52	
Columbia	Bonneville H.	1987	Ch	Adult	U-crb1, n52	
Columbia	Bonneville H.	1988	Ch	Adult	U-crb3	
Columbia	Bonneville H.	1990	St	Fry	M-crb4, n2	
Columbia	Bonneville H.	1999	Ch	Adult	U-crb1, n52	13
Columbia	Bonneville H.	2001	Ch	Adult	U-crb1, n52	
Columbia	Bonneville H.	2002	Sk	Presmolts	U-crb1, n52	
Columbia	Tanner Ck.	2002	St	Adult	MD-crb5	
Columbia	Little White Salmon H.	1987	Ch	Fry	U-crb1, n52	14
Columbia	Little White Salmon H.	1998	Ch	Adult	U-crb4, n2	15
Columbia	Little White Salmon H.	2001	Ch	Yearlings	U-crb1, n52	16
Columbia	Little White Salmon H.	2002	Ch	Juvenile	U-crb1, n52	
Columbia	Willard H.	2002	Coho	Juvenile	U-crb1, n52	
Columbia	Spring Ck. H.	1998	Ch	Adult	U-crb5	
Columbia	Spring Ck. H.	2001	Ch	Adult	U-crb1, n52	
Hood	Parkdale H.	1992	St	Adult	U-crb1, n52	17
Hood	Parkdale H.	1999	St	Adult	U-crb5	18
Deschutes	Warm Springs H.	1980	St	Adult	U-crb6, n2	
Deschutes	Warm Springs H.	2001	Ch	Juvenile	U-crb1, n52	
Suttle Lake	Suttle Lake	1973	Kok	Adult	U-crb1, n52	
Deschutes	Round Butte H.	1976	St	Adult	U-crb7	19
Deschutes	Round Butte H.	1985	St	Fry	U-crb1, n52	
Deschutes	Round Butte H.	1986	Ch	Adult	U-crb1, n52	

Table 1 (continued)

Major river drainage[a]	Isolation site	Year	Host	Life stage	Mid-G Seq. type, n#	5′N isolate no.[b]
Lower Mid-Columbia						
Deschutes	Round Butte H.	1988	St	Adult	M-crb1, n24	
Deschutes	Round Butte H.	2000	St	Adult	U-crb1, n52	
Deschutes	Round Butte H.	2000	Ch	Adult	U-crb1, n52	
Deschutes	Metolius R.	1975	Kok	Adult	U-crb6, n2	
Deschutes	Metolius R.	1988	Kok	Adult	U-crb8, n3	
Deschutes	Lake Billy Chinook	1991	Kok	2-3 yr	U-crb8, n3	
Deschutes	Lake Billy Chinook	2000	Kok	1-2 yr	U-crb8, n3	
Umatilla	Minthorn	1995	St	Adult	U-crb1, n52	20
WallaWalla	SF Walla Walla	1999	Ch	Adult	U-crb1, n52	21
Columbia	Priest Rapids	1987	Ch	Adult	U-crb9, n2	
Upper Mid-Columbia						
Wenatchee	Leavenworth H.	1982	Ch	Adult	U-crb10	
Wenatchee	Leavenworth H.	1985	Ch	Adult	U-crb1, n52	
Wenatchee	Leavenworth H.	1988	Ch	Adult	U-crb1, n52	
Wenatchee	Leavenworth H.	1997	Ch	Adult	U-crb1, n52	22
Wenatchee	Wenatchee R.	1987	Ch	Adult	U-crb1, n52	
Wenatchee	Chiwawa	2001	Coho	Adult	U-crb1, n52	
Wenatchee	Chiwawa Pond	2001	Coho	Adult	U-crb1, n52	
Wenatchee	White R.	1983	Sk	Adult	U-crb11	
Wenatchee	White R.	1984	Sk	Adult	U-crb12, n3	
Columbia	Entiat H.	1981	Ch	Yearling	U-crb12, n3	
Columbia	Entiat H.	1982	St	Fingerling	U-crb13, n2	
Columbia	Entiat H.	1984	Ch	Adult	U-crb13, n2	
Columbia	Entiat H.	1988	Ch	Adult	U-crb14	
Methow	Winthrop H.	1988	Ch	Adult	U-crb12, n3	
Methow	Winthrop H.	1997	Ch	Adult	U-crb1, n52	23
Methow	Winthrop H.	2001	Coho	Adult	U-crb1, n52	
Columbia	Similkameen R.	1984	Sk	Adult	U-crb9, n2	
Okanogan	Okanogan	1998	Sk	Adult	U-crb15	24
Lower Snake						
Snake	Lyons Ferry H.	1987	St	Fry	M-crb1, n24	
Tucannon	Tucannon H.	1987	Ch	Adult	U-crb16	
Clearwater	Breakfast Ck.	1988	Kok	Adult	M-crb1, n24	
Clearwater	Clearwater H.	1992	St	Brood	M-crb1, n24	25
Clearwater	Dworshak H.	1984	St	Adult	M-crb1, n24	
Clearwater	Dworshak H.	1998	St	Adult	U-crb1, n52	26
Clearwater	Kooskia H.	1989	St	Fry	M-crb1, n24	
Clearwater	Kooskia H.	1999	Ch	Adult	U-crb1, n52	27
Clearwater	Maggie Ck.	1998	Rbt	Juvenile	U-crb1, n52	28
Locksa	Powell Satellite	1994	Ch	Brood	U-crb1, n52	29
Clearwater	Crooked R.	1990	Ch	Juvenile	M-crb1, n24	
Clearwater	Red R. H.	1993	Ch	Brood	MC-crb6	30
Grande Ronde	Lookingglass H.	1994	Rbt	Adult	U-crb1, n52	31
Grande Ronde	Lookingglass H.	1995	Ch	Smolt	U-crb17	
Grande Ronde	Lookingglass H.	1996	Ch	Adult	U-crb1, n52	32
Grande Ronde	Lookingglass Ck.	1999	Ch	Adult	U-crb1, n52	33
Wallowa	Wallowa H.	1987	St	Adult	M-crb1, n24	
Wallowa	Wallowa H.	1991	St	Adult	U-crb1, n52	34
Wallowa	Wallowa H.	1998	St	Adult	MB-crb7	35
Little Sheep	Little Sheep Ck.	1987	St	Adult	M-crb8	
Little Sheep	Little Sheep Ck.	1988	St	Adult	M-crb1, n24	
Imnaha	Imnaha	1991	Ch	Adult	ME-crb9,n2	
Imnaha	Imnaha	1998	Ch	Adult	U-crb1, n52	36
Snake	Oxbow	1992	St	Brood	M-crb1, n24	
Salmon	Rapid R. H.	1992	Ch	Brood	ME-crb9,n2	37
Salmon	Rapid R. H.	1993	Ch	Brood	U-crb1, n52	38
Salmon	Rapid R. H.	1994	Ch	Brood	U-crb1, n52	39
Salmon	McCall SF Trap	1989	Ch	Brood	MB-crb10	40
Salmon	Pahsimeroi H.	1982	St	Adult	M-crb11	41
Salmon	Pahsimeroi H.	1991	St	Brood	MC-crb12	42
Salmon	Sawtooth H.	2002	Ch	Juvenile	U-crb1, n52	
Upper Snake						
Snake	Fall Ck.	1997	Rbt	Unknown	MC-crb13	43

[a]As shown in Figs. 5 to 7; [b]isolates that were 5′N gene sequenced

Table 2. Summary of primers used in mid-G and 5′N gene amplication. Primer binding site numbering determined from full length Western Regional Aquaculture Consortium (WRAC) strain IHNV genome sequence by Morzunow et al. (1995) (GenBank L40883)

Primer pair	Orientation	Sequences (5′-3′)	Primer size (bases)	Primer binding site
(A) Mid-G PCR primers				
1st round	Forward	AGAGATCCCTACACCAGAGAC	21	3515–3535
	Reverse	GGTGGTGTTGTTTCCGTGCAA	21	4207–4187
2nd round	Forward	TCACCCTGCCAGACTCATTGG	21	3575–3595
	Reverse	ATAGATGGAGCCTTTGTGCAT	21	4057–4037
(B) 5′N PCR primers				
1st round	Forward	CTTCAGACACTATAAACCGAG	21	90–110
	Reverse	TTGATGAGAATGATCCCATAG	21	770–750
2nd round	Forward	ACAGAACAAGCAGAACTATTT	21	111–131
	Reverse	GAAGAGGAGGCCGGTCAC	18	570–553

egger & Kurath 2002, Kurath et al. 2003). The significance of the branching order was assessed by bootstrap resampling of 1000 replicates. Intrapopulational nucleotide diversity was calculated according to the method of Nei (1987) utilizing Kimura's 2-parameter model (Kimura 1980) as applied in the Arlequin version 1.1 software package (Schneider et al. 1997). The numbers of nonsynonymous (d_N) and synonymous (d_S) substitutions were determined using the methods of Nei & Gojibori (1986) as applied in MEGA version 2.1 (Kumar et al. 2001).

RESULTS

Mid-G sequence analyses

Sequence analysis of the 303 nt mid-G region was performed on 115 Columbia River basin isolates (Table 1) and combined with mid-G sequences of 5 previously characterized Columbia River basin isolates (Nichol et. al. 1995). This data set does not include Hagerman Valley IHNV isolates, which have been previously described by Troyer et al. (2000, 2003). Nucleotide sequence analyses among the total of 120 Columbia River basin isolates revealed that they fell into 30 different mid-G sequence types. Nineteen out of these 30 different sequence types were found in single isolates. Consequently, 11 of the 30 different sequence types were found in multiple isolates. Within this latter set there were 2 very large sequence groups, which together represented 68% of the entire data set. These 2 most common sequences were found in groups of 52 and 24 isolates each, while all other sequence types were found in groups of only 2, 3 or 6 isolates each.

The maximum pairwise nucleotide diversity of the Columbia River basin isolates was 7.3% (22 nucleotides different out of 303) and the overall mean intrapopulational nucleotide diversity (π, Nei 1987) was 0.018 (1.8% or 5.5 nucleotides different out of 303). This intrapopulational nucleotide diversity level is 3-fold greater than the diversity found throughout Alaska (Emmenegger et al. 2000), but it is 2-fold less than the diversity reported for 4 aquaculture facilities located in the Hagerman Valley, Idaho (Troyer et al. 2000).

Phylogenetic analyses

In order to determine the evolutionary relationships of IHNV isolates in the Columbia River basin, a phylogenetic analysis was performed using the mid-G sequences obtained in this study along with representative mid-G sequences from Alaska (Emmenegger et al. 2000), British Columbia (Traxler, Anderson & Kurath unpubl.), and the Hagerman Valley study site (Troyer et al. 2000). Three California IHNV isolate mid-G sequences (from Nichol et al. 1995) were included as an outgroup. The phylogenetic analysis produced a tree with 3 strongly supported clades (Fig. 2). This result was consistent with a previous phylogeny, which had indicated that IHNV from throughout its infectious range formed 3 major clades (Kurath et al. 2004). The clades were denoted U, M, and L to represent their correlation with the upper, middle, and lower regions of the geographic range of the virus.

The phylogeny in Fig. 2 confirms and extends the previous phylogeny in finding that all Columbia River basin IHNV isolates fell into either the U or the M clade. Thus, the U clade, which contains all IHNV isolates from Alaska, British Columbia, and Washington coastal watersheds, also contains isolates from throughout the Columbia River basin, including drainages in Idaho. The M clade, previously recognized as a monophyletic clade containing only isolates from the Hagerman Valley in Idaho (Troyer et al. 2000), is shown by this phylogeny to also contain isolates from areas in the Columbia River basin including Washington, Oregon, and Idaho. This data shows that the Columbia River basin hosts 2 major clades, U and M, that are each present in multiple sub-drainages.

Out of the 120 Columbia River basin isolates, 76 isolates (17 sequence types) grouped into the U clade and 44 isolates (13 sequence types) grouped into the M clade (Table 1). The U clade isolates exhibited low genetic diversity, having a maximum pairwise nucleotide diversity

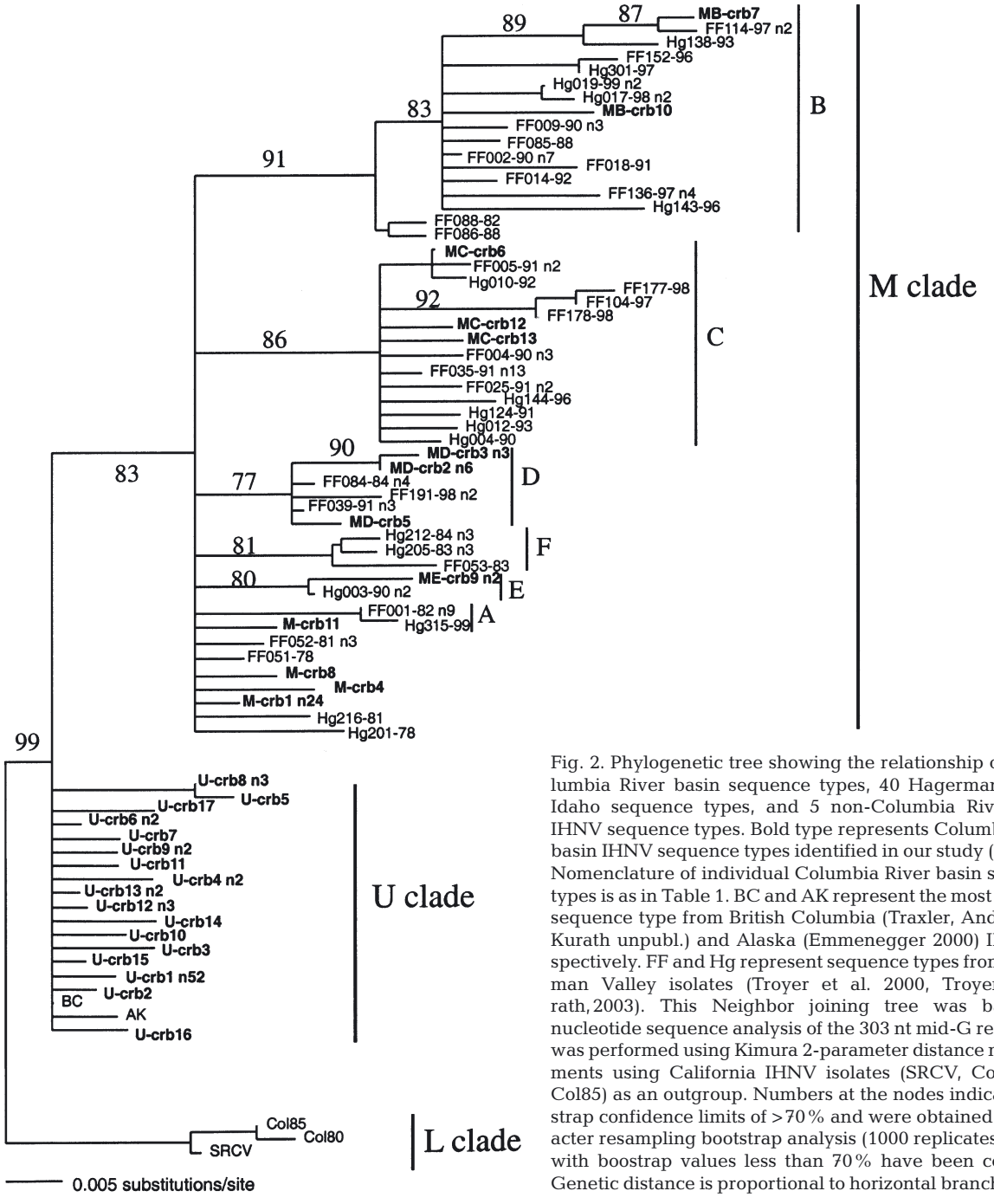

Fig. 2. Phylogenetic tree showing the relationship of 30 Columbia River basin sequence types, 40 Hagerman Valley, Idaho sequence types, and 5 non-Columbia River basin IHNV sequence types. Bold type represents Columbia River basin IHNV sequence types identified in our study (Table 1). Nomenclature of individual Columbia River basin sequence types is as in Table 1. BC and AK represent the most common sequence type from British Columbia (Traxler, Anderson, & Kurath unpubl.) and Alaska (Emmenegger 2000) IHNV respectively. FF and Hg represent sequence types from Hagerman Valley isolates (Troyer et al. 2000, Troyer & Kurath, 2003). This Neighbor joining tree was based on nucleotide sequence analysis of the 303 nt mid-G region and was performed using Kimura 2-parameter distance measurements using California IHNV isolates (SRCV, Col80, and Col85) as an outgroup. Numbers at the nodes indicate bootstrap confidence limits of >70% and were obtained by character resampling bootstrap analysis (1000 replicates). Nodes with boostrap values less than 70% have been collapsed. Genetic distance is proportional to horizontal branch lengths

of 2.6% (8 nucleotides different out of 303) and a mean intrapopulational nucleotide diversity of 0.6% (1.8 nucleotides different out of 303). The M clade isolates revealed a higher genetic diversity by having a 6.6% maximum pairwise nucleotide diversity (20 nucleotides different out of 303) and an intrapopulational nucleotide diversity of 1.7% (5.2 nucleotides different out of 303).

With regard to the impact of this nucleotide diversity on the G proteins of these isolates, the number of nonsynonymous substitutions for the M and U clades were also varied. The U clade sequences exhibited a striking absence of nonsynonymous substitutions (d_N = 0.000) while the M clade had a slightly higher number nonsynonymous substitutions (d_N = 0.013). Notably, the ratios of nonsynonymous to synonymous substitutions for either the U or the M clade were less than 1.0, thereby indicating a lack of discernible positive selection pressure.

The M clade exhibited significant branching topography indicating multiple subclades (Fig. 2). This result correlated with previous phylogenies of viral isolates from Hagerman Valley, Idaho, which identified 6 bootstrap-supported subclades (A to F) within the M clade (Troyer et al. 2000, Troyer & Kurath 2003). In the current study, 17 out of the 44 M clade Columbia River basin isolates grouped into 4 of these subclades (B, C, D, and E) (Fig. 2). The grouping of non-Hagerman Valley Columbia River basin isolates into these subclades indicates that these subclades, previously found co-circulating in the Hagerman Valley, Idaho, are also present within other regions of the Columbia River basin.

This study revealed 2 major sequence types dominating the 120 Columbia River basin IHNV isolates analyzed. The first most prevalent sequence type, denoted U-crb1, phylogenetically grouped into the U clade and contained 52 out of the 76 U clade isolates. Thus, U-crb1 represented over 68% of all Columbia River basin U clade isolates. The second most prevalent sequence type grouped into the M clade and was denoted M-crb1. This sequence group contained 24 isolates and represented over 54% of the 44 Columbia River basin M clade isolates.

Distribution of genetic variation

A multiple sequence alignment of the 30 mid-G sequence types (17 U clade and 13 M clade) identified in this report was used to generate a composite diagram of the distribution of the genetic variability within the data set (Fig. 3A). This indicated that the variable positions were relatively uniformly distributed along the 303 nt sequence of the mid-G region. Additionally, the 3 most variable nucleotide positions (base 786, 844, and 907) (Fig. 3A) were identified to be consistently different between the U and M sequence types. The conservation of a G residue at these 3 positions for U clade sequence types, rather than an A residue at positions 786 and 844 and a T residue at 907 for M clade sequence types, makes these nucleotide positions phylogenetically informative sites and possible candidates for use as clade markers.

There was a single exception to this consensus pattern in that one virus isolate (U clade sequence type U-crb2) contained an adenine instead of guanine at position 786. Of the 3 conserved nucleotide differences between the U and M clade sequence types, only base 786 represents a nonsynonymous change while base 844 and 907 represent a synonymous substitution between predicted G proteins of the U and M clade.

A comparison of the distribution of genetic variation among U clade sequence types (Fig. 3B) or only M clade sequence types (Fig. 3C) likewise revealed the distribution of divergent positions to be uniformly placed along the sequence. Most notable among the M clade sequence types is a cluster of substitutions that are in a region delineated by nucleotide positions 799 to 814 (Fig. 3C). This region appeared to be highly variable for amino acid substitutions, with 3 different amino acids present at codon 251 and 4 different at codon 252, which correlates with previous findings reported by Troyer et al. (2000). This variable region is located within the linear neutralizing epitopes mapped to amino acids 230 to 231 and 272 to 276 as described by Huang et al. (1996).

Partial N gene phylogeny

To ensure a high level of confidence in the characterization of Columbia River basin IHNV genetic heterogeneity, and to verify the phylogeny generated with mid-G sequence data, an additional analysis was performed using a 412 bp sequence near the 5′ terminus of the N gene. Partial N gene sequence analysis was conducted with 43 Columbia River basin isolates that were selected as a subset from the 120 Columbia River basin isolates (Table 1). This data set did not include isolates from the Hagerman Valley. The partial N gene sequence analysis revealed 21 different sequence groups with a maximum nucleotide diversity of 7.3% (30 nt different out of 412) and an intrapopulational nucleotide diversity of 0.032 (3.2% or 13 nt different out of 412).

For phylogenetic analysis these 43 5′N sequences were combined with 5′N sequences of 16 isolates from the Hagerman Valley, selected to represent all known subclades, and 3 California isolates used as the outgroup. These 62 isolates represented 31 mid-G sequence types, and by 5′N sequence analysis they were resolved into 33 5′N sequence types. The resulting 5′N phylogeny is shown in Fig. 4A, along with a mid-G phylogeny of the exact same subset of 62 virus isolates (Fig. 4B) for comparison. The mid-G sequence type nomenclature and individual isolate numbers are shown on both trees so that resolution of individual sequence types into clades can be assessed.

The separation of these 62 IHNV isolates into the major clades U, M, and L, was identical in both trees. Within the U clade, 5 mid-G sequence types each representing single virus isolates, were also individual isolate sequence types by 5′N analysis. The large mid-G sequence type U-crb1 was resolved into 5 smaller groups by 5′N sequencing, thereby providing a higher resolution of subgroups within the U clade in the 5′N phylogeny (Fig. 4A). Within the M clade, 12 of the 22 sequence types resolved by mid-G sequencing were identically resolved by 5′N sequences. Three other

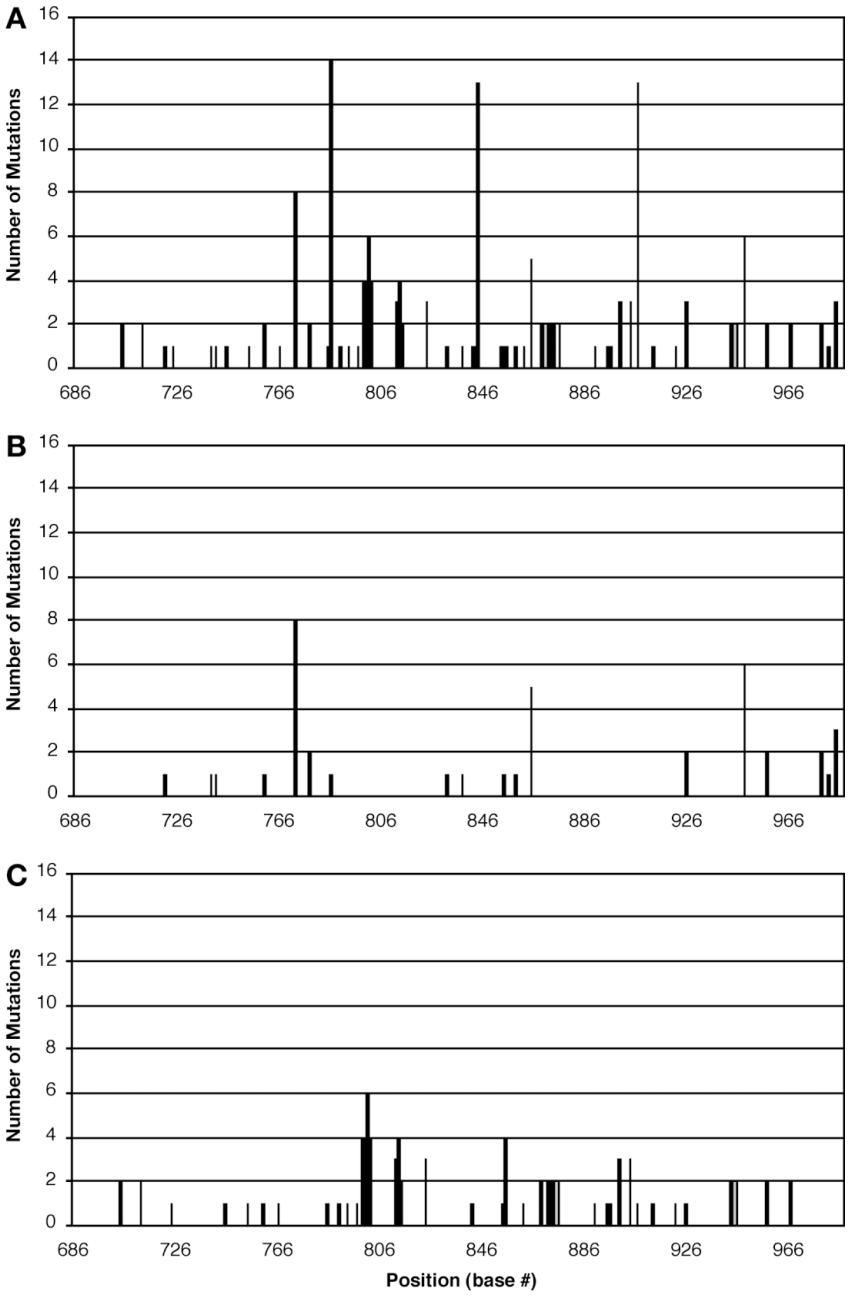

Fig. 3. Diagrams of the distribution of variability in mid-G nucleotide sequences of (A) 30 aligned sequence types, (B) 17 aligned U clade sequence types, and (C) 13 aligned M clade sequence types

mid-G sequence types were each resolved into 2 5′N types, and an additional 3 pairs of mid-G sequence types were combined into single 5′N sequence types. The 5′N phylogeny confirmed the grouping of isolates in M subclades A,C, D, and E, as identified in mid-G sequence data. However, M sub-clade B, which was supported by a high bootstrap value in the mid-G phylogeny, was not as well defined in the 5′N phylogeny, illustrating the value of analyzing different sequence regions. Overall, the 5′N analysis further confirms the robust nature of the clade based typing scheme as established by mid-G phylogeny.

Geographic and temporal relationships of the Columbia River basin IHNV isolates

The geographic distribution analysis of the Columbia River basin U and M clade isolates is shown in Fig. 5A. The presence of both the U and M clade virus types was observed within the lower Columbia River mainstem and tributaries, and within the lower Snake River and tributaries. However, the 2 clades do not overlap throughout the entire basin. The mid-upper Columbia River basin north of the Snake River mouth contains exclusively U clade virus isolates,

while the Snake River upstream of the confluence of the Imnaha River has only M clade IHNV. The Hagerman Valley contains only M clade virus as was previously reported (Troyer et al. 2000, Troyer & Kurath 2003).

Temporally, the U clade isolates were the first IHNV types to appear within the lower and mid-Columbia River basin. They were observed from 1973 through 2002 (Fig. 5B), representing the entire time frame of our data set. The M clade isolates did not appear in the lower and middle Columbia River basin until 1980, after which they were observed through 2002. Within the lower Snake River and tributaries, the opposite temporal pattern of the 2 clades was observed. The M clade was the first type to be observed, appearing in 1982 and persisting through to 1998. The U clade appeared in 1987 and remained through 2002 (Fig. 5B).

The 2 most prevalent individual sequence types, U-crb1 and M-crb1, mirrored the geographical distribution of the U and M clades respectively. U-crb1 (Fig. 6A) was observed throughout the entire Columbia River basin with the exception of the Snake River above the Imnaha River confluence. M-crb1 (Fig. 7A) was identified only in the Lower Columbia River basin and lower Snake River and tributaries. Therefore, there was an overlap of geographic ranges of the 2 most prevalent sequence types within both the lower Columbia River basin and lower Snake River and tributaries. In the lower Columbia River basin, U-crb1 was observed from 1973 to 2002 (Fig. 6B) while M-crb1 was observed from 1980 to 1994 (Fig. 7B). However, in the Snake River and tributaries the temporal occurrence of the 2 sequence types was reversed, with M-crb1 occurring first in 1984 to 1992 and U-crb1 appearing from 1991 to 2001.

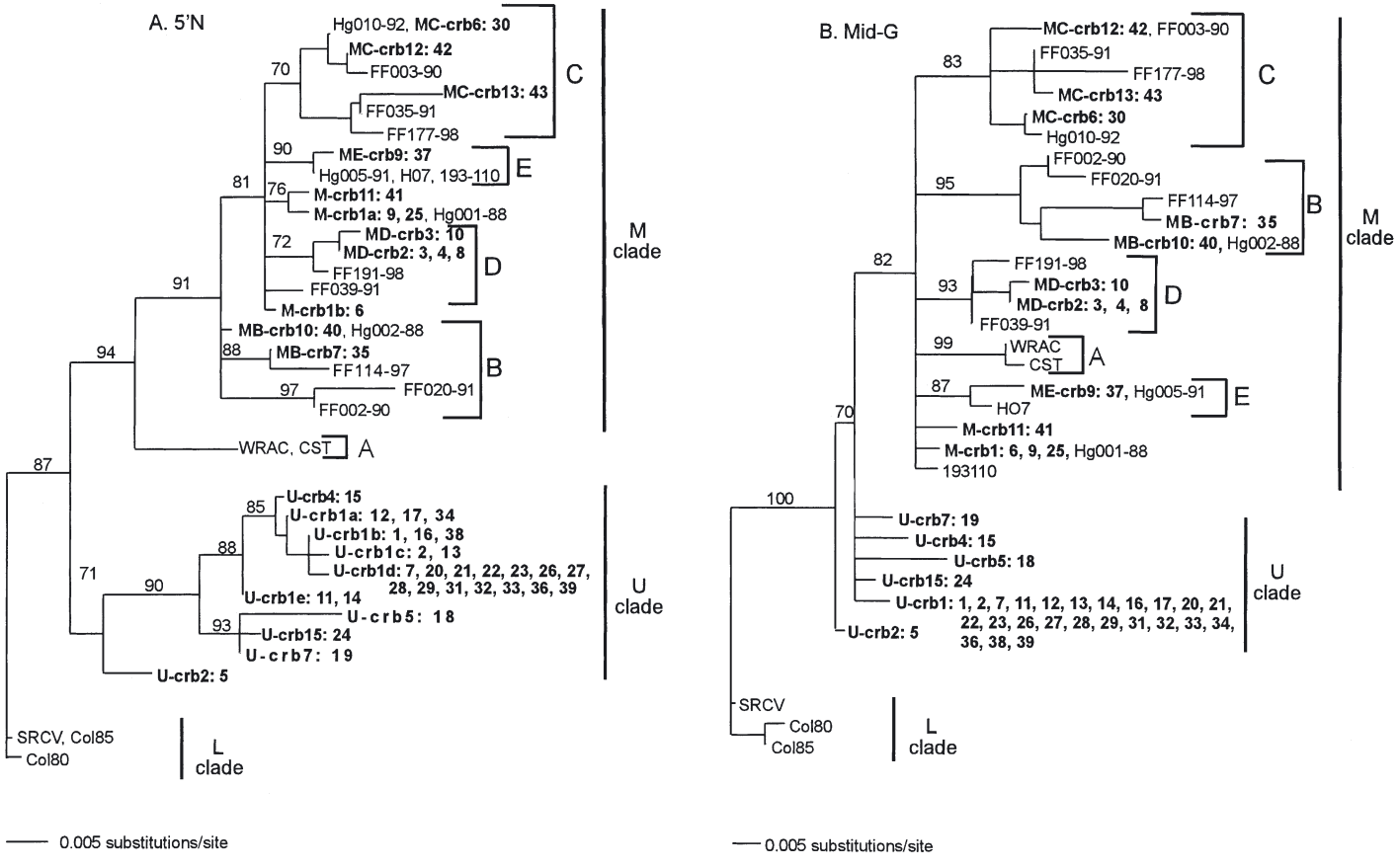

Fig. 4. Phylogenetic trees showing relationships within a subset of 43 IHNV isolates from the Columbia River basin on the basis of (A) 5′N nucleotide sequence or (B) mid-G nucleotide sequence. Bold type represents Columbia River basin IHNV sequence types identified in this study (Table 1, Fig. 2). Nomenclature of individual Columbia River basin sequence types is similar to that used in Fig. 2 except for the addition of a colon followed by individual isolate numbers corresponding to those isolates that were 5′N sequenced as shown in Table 1. All sequence type nomenclature was based on mid-G sequence analysis; hence isolates grouping into 1 sequence type by mid-G sequence subsequently may group into multiple sequence types by 5′N sequence analysis. To designate these sequence types, a lowercase letter was positioned after the sequence type number (i.e. U-crb1 became U-crb1a-e). Additionally, subclades (Ato E) as defined by mid-G sequence types are in brackets. Sequences were analyzed and trees drawn as described in the legend for Fig. 2

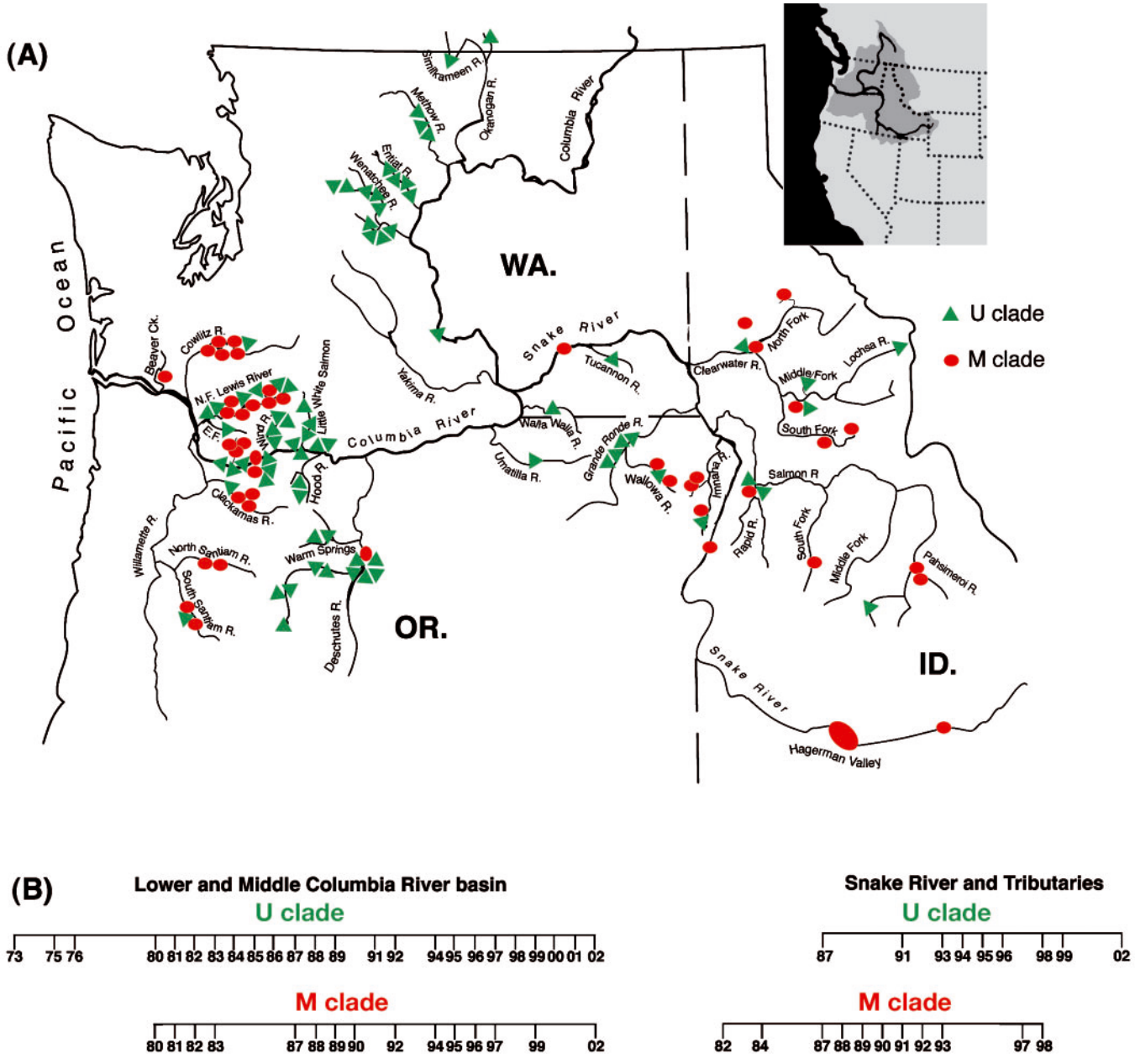

Fig. 5. (A) Geographic distribution and (B) temporal analysis of U and M clade Columbia River basin IHNV isolates. The location of 76 isolates that grouped into the U clade (green triangles) and 44 isolates that grouped into the M clade (red ovals) are displayed on the map. Each triangle and oval represents a single isolate. The temporal occurrence of U and M clade isolates within the lower and middle Columbia River basin and Snake River and tributaries is shown on the horizontal time scales. Years marked on the scales indicate the occurrence of one or multiple isolates that group with the respective clade. The shaded region in the insert map shows the Columbia River basin

The geographical and temporal analysis of the 16 M clade isolates that grouped into the B, C, D, or E subclades is also shown in Fig. 7. The D sub-clade, containing 10 of the viral isolates characterized in this study, was the most prevalent sub-clade observed among the Columbia River basin isolates. It is interesting to note that the D sub-clade isolates were geographically limited to 2 sub-basins and a portion of the mainstem in the lower Columbia River basin, where they occurred only between 1995 and 2002 (Fig. 7). Less prevalent in the Columbia River basin was the occurrence of B, C, and E subclades. These subclades were observed only in the lower Snake River and tributaries (Fig. 7). The occurrence of the isolates found to group into either of subclades B or C was sporadic, with only 2 B sub-clade isolates, occurring in 1989 and

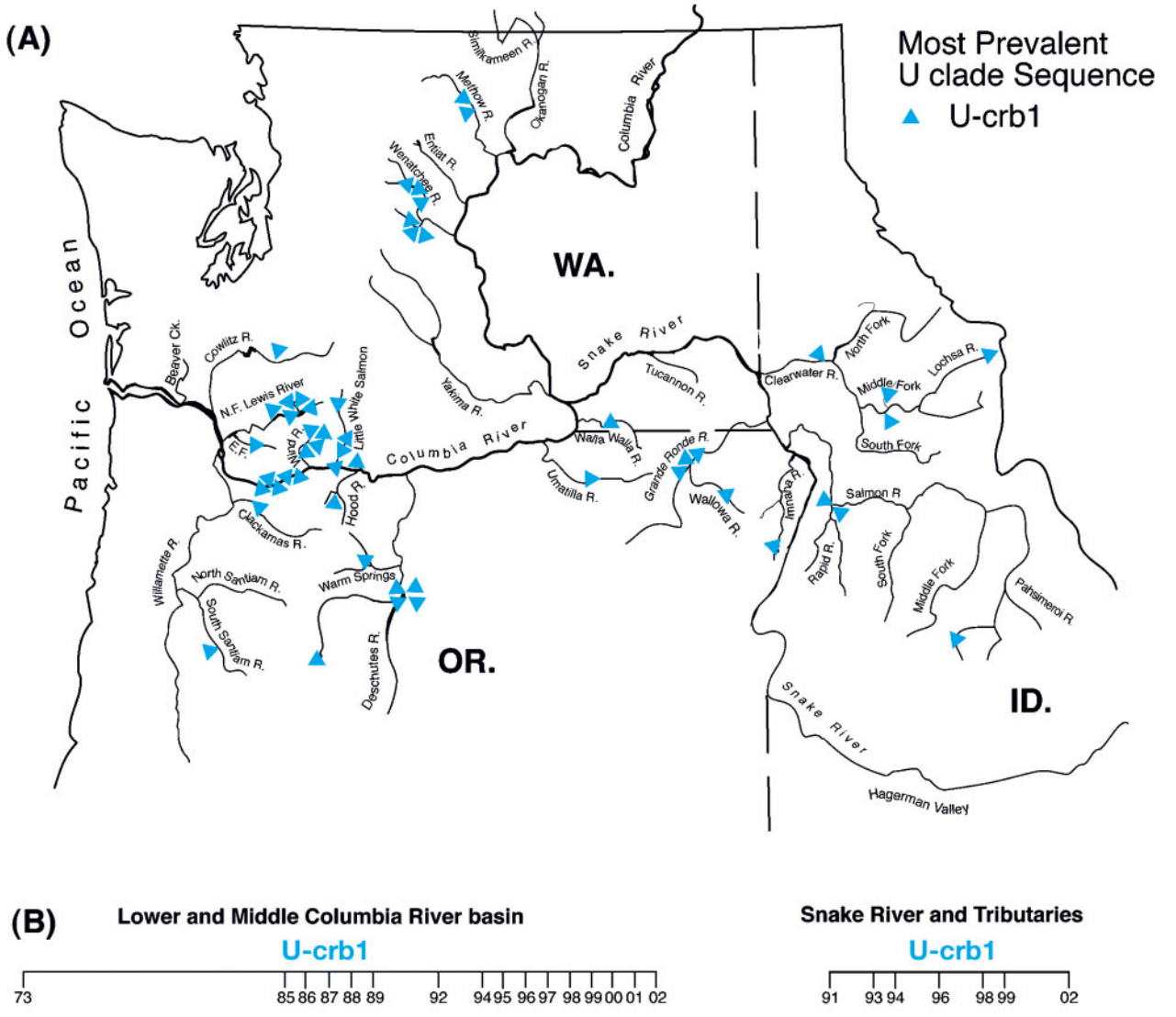

Fig. 6. (A) Geographic distribution and (B) temporal analysis of the most prevalent sequence type out of 120 Columbia River basin IHNV isolates analyzed in this report. The sequence group (denoted U-crb1) fell into the U clade and is shown as triangles, each representing a single isolate. The years on the time lines mark the occurrence of 1 or more isolate(s) with a mid-G sequence identical to the U-crb1 sequence type

1998, and 3 C sub-clade isolates, occurring during 1991 to 1997. The occurrence of the E sub-clade isolates, although less sporadic, was not persistent, with only 2 isolations made from 1991 to 1992 (Fig. 7).

Correlation with previous IHNV typing systems

In order to correlate the phylogenetic types found in this study with an established monoclonal antibody typing system as described in Ristow & Arnzen (1989), 45 (20 M clade and 25 U clade isolates) of the 120 Columbia River basin isolates were tested for reactivity with the the type 2-specific monoclonal antibody 2NH105B (data not shown). This antibody binds to an unidentified epitope in the IHNV N protein of electropherotype 2 IHNV isolates. Notably, all 20 M clade isolates characterized reacted positively 2NH105B, indicating a type-2 IHNV as defined by Hsu et al. (1986). Out of the 25 U clade isolates tested, 15 reacted positively with 2NH105B while 10 isolates did not react with the type 2 antibody. Thus within the U clade there are IHNV isolates that would have been type-2 and non type-2 by the previous virus typing system. All 45 isolates tested reacted positively with the universal IHNV N protein monoclonal antibody IND14D (data not shown). These results indicate that there is not a direct correlation between the electropherotype groupings and the clade based groupings described in this paper.

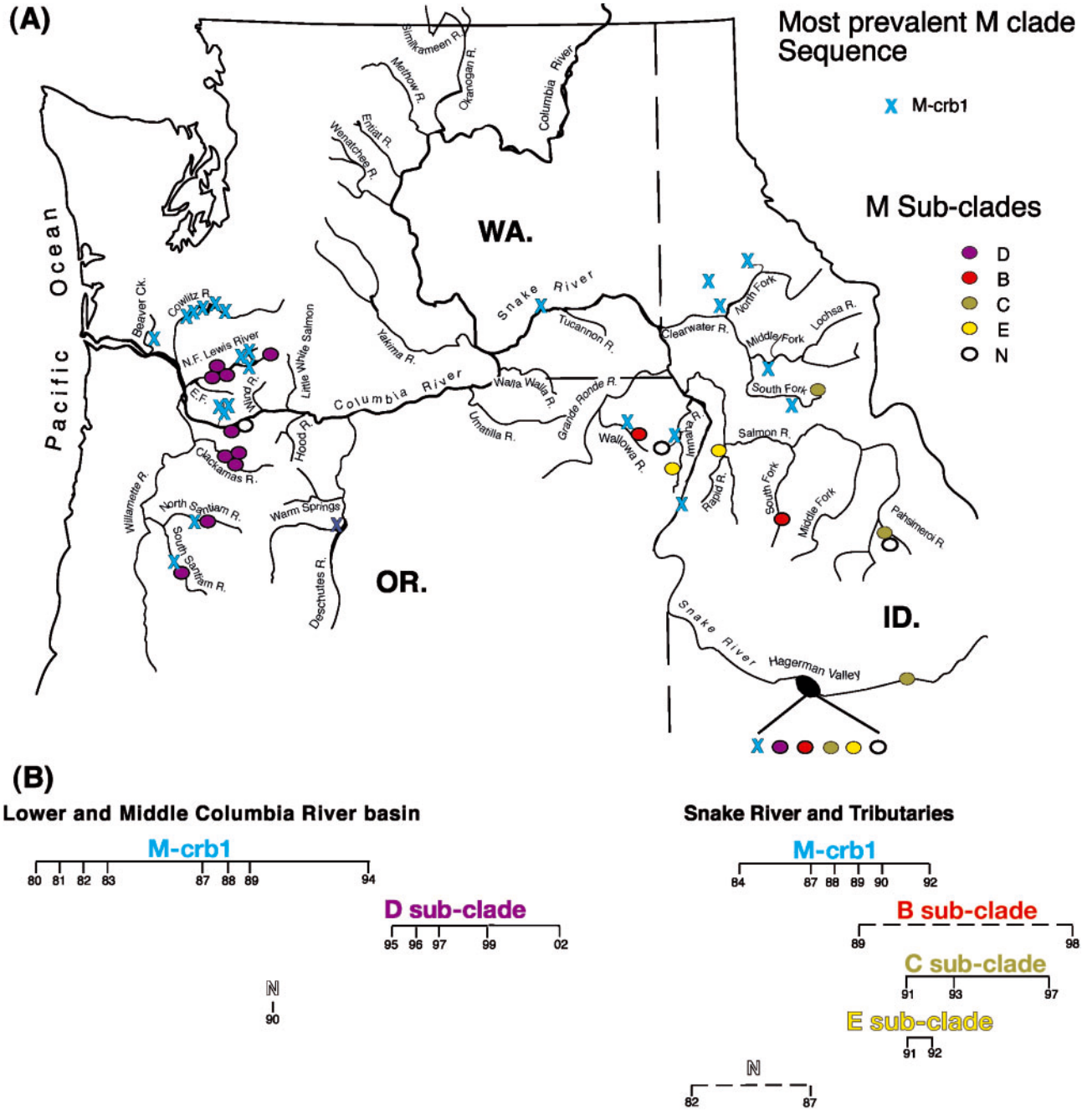

Fig. 7. (A) Geographical distribution and (B) temporal analysis of isolates grouping into the most prevalent M clade sequence (M-crb1) or into the M subclades B, C, D and E. Each blue 'X' represents the location of a single isolate with a mid-G sequence identical to the M-crb1 sequence type. Each colored oval represents the location of a single isolate that grouped into one of the M subclades, indicated by color; sub-clade D purple, B red, C olive, E yellow. N (unfilled oval) represents those M clade isolates that did not group into a distinct sub-clade. The horizontal bar scales display the years these Columbia River basin IHNV samples were isolated. Years marked on the scales indicate the occurrence of 1 or multiple isolates that group with a mid-G sequence identical to the M-crb1 sequence type or the respective M sub-clade

DISCUSSION

Previous investigations of genetic heterogeneity among IHNV isolates from throughout North America have suggested that IHNV genetic relatedness generally correlates with geographical origin (Hsu 1986, Nichol et al. 1995, Oshima 1995, Kurath et al. 2003). The data presented in this report clearly confirmed that the Columbia River basin contained 2 distinct phylogenic clades that co-occur and overlap geographi-

cally. Analysis of the phylogenetic trees generated with either mid-G sequence data (Figs. 2 & 3) or 5′N sequence data (Fig. 4) revealed that all Columbia River basin isolates analyzed grouped into 1 of 2 distinct clades, designated U and M, which were well supported by bootstrap analysis. It is interesting to note that although the U and M clades are sympatric, each clade maintains distinct characteristics. The U clade has a relatively low genetic diversity (i.e. genetic stasis) as indicated by the short branch lengths in the neighbor-joining distance tree (Fig. 2) and a mean intrapopulational diversity of 0.6%. This is in stark contrast to the M clade, which displays a 3-fold higher level of genetic diversity (1.7%) and significant internal branching topography indicating multiple lineages (Troyer et al. 2000).

Due to the low genetic diversity within the U clade, it is difficult to conclusively identify evolutionary relationships within this clade. Temporal analysis (Fig. 5B) suggests that this clade represents the endemic virus type that was present before the emergence of IHNV throughout the lower Columbia River basin in the early 1980s. Although the virus isolates from the early sockeye epidemics in the 1950s are not available for analysis (Wolf 1988), we speculate that these would have been U clade viruses. Moreover, it is evident from the isolates analyzed in this report that U clade virus is maintained throughout the basin over time despite the occurrence of other viral types.

In the present study, the first occurrence of a non U clade virus in the Columbia River basin outside the Hagerman Valley was an isolate from Lewis River in 1980. This isolate grouped into the major M-crb1 sequence type within the M clade, along with 23 other isolates containing an identical mid-G sequence. The M-crb1 virus isolates from 1980 to 1983 were obtained from fish experiencing epidemics that were part of the dramatic increase of virus in the lower Columbia River basin in the early 1980s, which was cause for extreme concern. Because of the sequential timing of IHNV emergence throughout the Hagerman Valley between 1977 and 1980, and throughout the Columbia River basin between 1980 and 1982, it has been suggested that infected fish in the Hagerman Valley were the source of virus for fish downstream throughout the rest of the Columbia River (Groberg 1983). The observation that the M-crb1 virus type grouped with the M clade, which contains all Hagerman Valley isolates, indicates a shared common ancestor and therefore lends validity to the above suggestion; however, it is difficult to conclusively account for the origin of the specific M-crb1 virus type. This is due to the fact that only one isolate in the Hagerman Valley was found to have an identical mid-G sequence to the M-crb1 virus type and it was isolated in 1988, significantly after the 1980 observation of M-crb1 virus in the lower Columbia River basin. Thus, although M clade IHNV has been in the Hagerman Valley since 1978, we do not have data to indicate that the specific sequence type M-crb1 was present in the Hagerman Valley prior to its emergence in the lower Columbia River basin in 1980 to 1982. It is evident that the M-crb1 type did get established both within the lower Columbia River basin and lower Snake river and tributaries, where it persisted for 14 yr. The absence of any M-crb1 type IHNV isolates after 1994 was not due to sampling bias (Fig. 1), suggesting that after a 14 yr period this sequence type has disappeared from the lower Columbia River basin for unknown reasons. It is also noteworthy that within the M-crb1 virus type there was no detectable genetic diversity, with all 24 virus isolates in the group containing an identical mid-G sequence. This observation of extremely low genetic diversity is in contrast to the high degree of genetic diversity exhibited by lineages (A to E) present in the M clade in the Hagerman Valley (Troyer et al. 2000).

Since the occurrence of the M-crb1 virus type, the Columbia River basin has also experienced several incidences of other M clade virus types. Within the lower Columbia River basin, 10 isolates gathered over a period of 8 yr (1995 to 2002) grouped phylogenetically into the D sub-clade (Fig. 2). Other IHNV isolates grouping with this distinct clade are 9 isolates from private and state facilities located in the Hagerman Valley, Idaho (Troyer et. al. 2000, Troyer & Kurath 2003). These Hagerman Valley D sub-clade viruses were isolated between 1984 and 1998, thereby temporally indicating that sub-clade D was present in the Hagerman Valley prior to its occurrence in the lower Columbia River. This suggests that viral sub-clade D was most likely introduced into the lower Columbia River basin from the Hagerman Valley. Similar to the M-crb1 virus type, the D sub-clade virus became established in the lower Columbia River basin. However the D sub-clade IHNV was not observed elsewhere in the basin. It was present only in the lower Columbia River basin, although a single isolate of this type was found in the Queets watershed on the Washington coast (Emmenegger et. al. 2002). It is also interesting to note that following the appearance of the D sub-clade in the lower Columbia River basin, the M-crb1 virus type was never again isolated, appearing as though the D virus sequence types may have displaced the M-crb1.

Other M clade lineages that occurred in the Columbia River basin were B, C, and E. Unlike the M-crb1 and sub-clade D virus types that persisted over several years at multiple sites and were represented by 24 and 10 virus isolates respectively, the B, C, and E isolates occurred sporadically, with no evidence of becoming established. Additionally, B, C and E lineages occurred

only in the lower Snake River and tributaries while the M-crb1 and sub-clade D viral types maintained a stronghold in the lower Columbia River basin. Temporal analysis reveals that lineages B, C, and E were present in the Hagerman Valley prior to their occurrence in the other areas of the Snake River, thereby suggesting their introduction from the Hagerman. An example of this is the sub-clade C virus isolated from the Pahsimeroi hatchery in 1991, which contained an identical mid-G sequence to 3 isolates from the Hagerman Valley in 1989 and 1990.

While our data suggest that multiple virus types have been introduced into areas of the Columbia River basin from the Hagerman Valley, most of these introductions do not persist, and only 2 introductions appear to have resulted in M clade virus becoming established outside the Hagerman Valley. It is not immediately evident what specific event(s) or factors(s) have enabled such introductions to occur. The simplest mechanism would be that IHNV in effluent from Hagerman Valley fish culture facilities resulted in waterborne transmission to fish stocks downstream. However for this suggestion to be plausible the virus would have to retain infectivity after dramatic dilution and over an extended period of time. LaPatra et al. (2001) reported that IHN virus suspended in Snake River water exhibited a 99% reduction in virus concentration in 24 h. More importantly, if waterborne transmission was a significant mechanism, one would expect to find those viral types that are most prevalent in the Hagerman Valley occurring most frequently downstream. This is not the case, at least within the last 15 yr, for which we have extensive data on Hagerman Valley IHNV types. The majority of the Hagerman Valley IHNV isolates characterized since 1988 were in the B and C subclades (66%) (Troyer et al. 2000). However from our current data it is clear that subclades B and C do not comprise the major virus types observed downstream in the Columbia River basin; rather B and C were only found rarely and sporadically, and were limited to the lower Snake River basin. Although our data set does not provide conclusive indications for the time period prior to 1990, it indicates that waterborne transmission is not the mechanism for the apparent introduction of sub-clade D IHNV into the lower Columbia River basin in 1995 to 2002.

Another factor which may have contributed to the transmission of IHNV between different areas of the basin is the inadvertent transportation of virus in infected fish or eggs (Wolf 1988). Evidence for specific practices causing IHN disease has been documented (Crawford 1982, Nichol et al. 1995). Despite the careful evaluation of movement of fish within the basin, rare events may lead to the introduction of new viral types. For instance, it is possible that the sporadic occurrence of M clade virus within the lower Snake River and tributaries may be accounted for by the outplanting of anadromous fish from Hagerman Valley natural resource mitigation hatcheries into tributaries of the lower Snake River. The identification of IHNV isolates in the tributaries of the lower Snake River that contain identical mid-G sequence to isolates that occurred in the Hagerman Valley hatcheries support such speculations. The epidemiological link between the Hagerman Valley and the lower Columbia River basin is more difficult to envision, but the 2 introduction events that resulted in the establishment of M clade IHNV may have involved fish transportation events such as stock transfers or barging of outmigrating salmonids from Idaho hatcheries to the lower Columbia River basin (Groberg 1983).

Although IHNV from both clades infects multiple salmonid species in the Columbia River basin, there was some indication of general host specificity trends in our data set (Table 3). This is most evident for chinook and steelhead isolates, which were numerous enough to show trends clearly. Of the 50 IHNV isolates from chinook salmon in this study, the great majority (86%) were U clade and a much smaller proportion (18%) were M clade. In contrast, the 41 steelhead isolates were mostly M clade (63%) with a smaller proportion of U clade isolates (38%). Although the combined number of sockeye and kokanee isolates was only 14, it is notable that they were mostly U clade (86%). The other host species were represented in numbers too small to draw conclusions about host species trends (Table 3). The observed relationship between phylogenetic clade and host species within the Columbia River basin is not a consequence of an areas bias for culturing a specific fish species. For instance, regions containing only the U clade did not have a predominance of culturing chinook.

These patterns of host association observed for field infections of chinook, steelhead, sockeye and kokanee are in agreement with previous studies of host specificities for IHNV electropherotypes 1 and 2 in experi-

Table 3. Percent of U or M IHNV types isolated from various salmonid species

Host species	No. of isolates	Phylogenetic clade (%) U	M
Chinook salmon	51	82	18
Steelhead trout	41	37	63
Kokanee salmon	9	78	22
Sockeye salmon	5	100	0
Rainbow trout	5	40	60
Coho salmon	4	100	0
Cutthroat trout	4	0	100
Chum salmon	1	100	0

mental challenge studies (LaPatra et al. 1990, 1993). The inclusion of these specific isolates in the current study allowed our observations to be correlated with the previous reports that kokanee and chinook salmon are more susceptible to an electropherotype 1 IHNV (U clade) while steelhead are notably more susceptible to electropherotype 2 (M clade) (LaPatra et al. 1990, 1993). Also as previously noted, the host specificity observed in this study was not at all absolute, and where multiple host species co-occur in close proximity they tended to have identical or very similar IHNV types (Hsu et al. 1986, Ristow & Arnzen 1989). It must be emphasized that although there is some indication of host specificity, we cannot make any conclusions regarding differential virulence of the 2 IHNV clades because only a small number of virus isolates in our data set were known to be associated with host mortalities or epidemics (data not shown).

It could be that the basis of the general M clade specificity for steelhead is historical adaptation of M clade IHNV to rainbow trout (Kurath et al. 2003, Troyer & Kurath 2003). Similarly, the basis of the sockeye/kokanee specificity for U clade IHNV may involve the hypothesized long historical association between U clade IHNV and sockeye (Kurath et al. 2003). The basis of U clade specificity for chinook is harder to discern, and may involve historical or epidemiological factors such as the widespread distribution of U clade IHNV relative to the more sporadic M clade IHNV. For all host specificity phenomena, it is likely that epidemiological and anthropogenic influences may play major roles in addition to any actual biological host specificity for infection or pathogenesis by the different virus clades.

Acknowledgements. The authors thank Susan Gutenberger, Theresa London, Ray Brunson, Corie Samson, and Kathy Clemens (United States Fish and Wildlife Service), Joan Thomas (Washington Department of Fish and Wildlife), Mark Engelking, John Kaufman, Warren Groberg, Sam Onjukka, and Glenda Claire (Oregon Department of Fish and Wildlife), Keith Johnson and Sharon Landin (Idaho Department of Fish and Game), Dorothee Kieser (Department of Fisheries and Oceans, B.C.), for providing virus isolates and background information. Bill Batts for his technical expertise in nucleotide sequencing and cell culture. Thanks to Evi Emmenegger for her technical advice on the graphics. This work was supported by a post-doctoral fellowship grant USDA. 00-35204-9226 to Dr. Kyle Garver

LITERATURE CITED

Amend DF, Wood JW (1972) Survey for infectious hematopoietic necrosis (IHN) virus in Washington salmon. Prog Fish-Cult 34:143–147

Anderson ED, Engelking HM, Emmenegger EJ, Kurath G (2000) Molecular epidemiology reveals emergence of a virulent IHN virus strain in wild salmon and transmission to hatchery fish. J Aquat Anim Health 12:85–99

Busch RA (1983) Viral disease considerations in the commercial trout industry in Idaho. In: Leong JC, Barila, TY (eds) Proceedings of a workshop on viral diseases of salmonid fishes in the Columbia River Basin, Bonneville Power Administration Spec Publ, Portland, OR, p 84–100

Crawford BA (1982) Report on the occurrence and effect of the virus infectious hematopoietic necrosis (IHN) on the 1981–1982 hatchery production of steelhead in southwest Washington. Washington Dept of Game, Olympia, WA

Emmenegger EJ, Kurath G (2002) Genetic characterization of infectious hematopoietic necrosis virus of coastal salmonid stocks in Washington State. J Aquat Anim Health 14:25–34

Emmenegger RJ, Meyers TR, Burton TO, Kurath G (2000) Genetic diversity and epidemiology of infectious hematopoietic necrosis virus in Alaska. Dis Aquat Org 40:163–176

Groberg WJ (1983) The status of viral fish diseases in the Columbia River Basin. In: Leong JC, Barila TY (eds) Proceedings of a workshop on viral diseases of salmonid fishes in the Columbia River Basin.Bonneville Power Administration Spec Publ, Portland, OR, p 1–22

Groberg WJ, Fryer JL (1983) Increased occurrence of infectious hematopoietic necrosis virus in fish at Columbia River basin hatcheries,1980–1982. Technical paper No. 6620, ORESU-T-83–002. Oregon Agricultural Experiment Station, Corvallis, OR

Hsu YL, Engelking HM, Leong JC (1986) Occurrence of different types of infectious hematopoietic necrosis virus in fish. Appl Environ Microbiol 52:1353–1361

Huang C, Chien MS, Landolt M, Batts W, Winton J (1996) Mapping the neutralizing epitopes on the glycoprotein of infectious haematopoietic necrosis virus, a fish rhabdovirus. J Gen Virol 77:3033–3040

Jørgensen PEV, Einer-Jensen K, Higman KH, Winton JR (1995) Sequence comparison of the central region of the glycoprotein gene of the neutralizable, non-neutralzable and serially passed isolates of the viral haemorrhagic septicaemia virus. Dis Aquat Org 23:77–82

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

Kumar S, Tamura K, Jakobsen IB, Nei M (2001) MEGA2: Molecular Evolutionary Genetics Analysis software. Bioinformatics 17:1244–1245

Kurath G, Garver KA, Troyer RM, Emmenegger EJ, Einer-Jensen K, Anderson ED (2003) Phylogeography of North American infectious hematopoietic necrosis virus in salmonid fishes. J Gen Virol 84:803–814

LaPatra SE, Groberg WJ, Rohovec JS, Fryer JL (1990) Size-related susceptibility of salmonids to 2 strains of infectious hematopoietic necrosis virus. Trans Am Fish Soc 119:25–30

La Patra SE, Fryer JL, Rohovec JS (1993) Virulence comparison of different electropherotypes of infectious hematopoietic necrosis virus. Dis Aquat Org 16:115–120

LaPatra SE, Troyer R, Shewmaker W, Jones G, Kurath G (2001) Understanding aquatic animal virus survival and trafficking and its role in risk assessment. In: Rodgers CJ (ed) Proceedings of the office international des epizooties (OIE) international conference on risk analysis in aquatic animal health. OIE, Paris, p 251–258

Morzunov SP, Winton JR, Nichol ST (1995) The complete genome structure and phylogenetic relationship of infectious hematopoietic necrosis virus. Virus Res 38:175–192

Mulcahy DM, Tebbit GL, Groberg WJ, McMichael JS and 5 others (1980) The occurrence and distribution of salmonid viruses in Oregon. Technical paper No. 5504, ORESU-T-80–004. Oregon Agricultural Experiment Station, Corvallis, OR

Mulcahy D, Pascho R, Jenes CK (1984) Comparison of in vitro growth characteristics of 10 isolates of infectious hematopoietic necrosis virus. J Gen Virol 65:2199–2207

Nei M (1987) Molecular evolutionary genetics. Columbia University Press, New York

Nei M, Gojobori T (1986) Simple methods for estimating the number of synonymous and nonsynomymous nucleotide substitutions. Mol Biol Evol 3:418–426

Nichol ST, Rowe JE, Winton JR (1995) Molecular epizootiology and evolution of the glycoprotein and non-virion protein genes of infectious hematopoietic necrosis virus, a fish rhabdovirus. Virus Res 38:159–173

Oshima KH, Arakawa CK, Higman KH, Landolt ML, Nichol ST, Winton JR (1995) The genetic diversity and epizootiology of infectious hematopoietic necrosis virus. Virus Res 35:123–141

Ristow SS, Arnzen JM (1989). Development of monoclonal antibodies that recognize a type-2 specific and a common epitope on the nucleoprotein of infectious hematopoeitic necrosis virus. J Aquat Anim Health 1:119–125

Ross AJ, Pelnar J, Rucker RR (1960) A virus-like disease of chinook salmon. Trans Am Fish Soc 89:160–163

Rucker RR, Whipple WJ, Parvin JR, Evans CA (1953) A contagious disease of salmon possibly of virus origin. US Fish Wildl Serv Fish Bull 54:35–46

Schneider S, Kueffer JM, Roessli D, Excoffier L (1997) Arlequin, a software for population genetic data analysis version 1.1. Depts of Anthropology and Ecology, University of Geneva, Geneva

Swofford DL (1998) PAUP* Phylogenetic Analysis Using Parsimony (*and other methods), Vers 4. Sinauer Associates, Sunderland, MA

Thoesen JC (1994) Suggested procedures for the detection and identification of certain finfish and shellfish pathogens, 4th edn. American Fisheries Society, Fish Health Section, Bethesda, MD

Troyer RM, LaPatra S, Kurath G (2000) Genetic analyses reveal unusually high diversity of infectious hematopoietic necrosis virus in rainbow trout aquaculture. J Gen Virol 81:2823–2832

Troyer RM, Kurath G (2003) Molecular epidemiology of infectious hematopoietic necrosis virus reveals complex virus traffic and evolution within southern Idaho aquaculture. Dis Aquat Org 55:175–185

Watson SW, Guenther RW, Rucker RR (1954) A virus disease of sockeye salmon, interim report. US Fish and Wildlife Service Special Scientific Report on Fisheries, No. 138, Washington, DC

Wingfield WH, Fryer JL, Pilcher KS (1969) Properties of the sockeye salmon virus (Oregon strain). Proc Soc Exp Biol Med 130:1055–1059

Wingfield WH, Nims L, Fryer JL, Pilcher KS (1970) Species specificity of the sockeye salmon virus (Oregon strain) and its cytopathic effects in salmonid cell lines. In: Snieszko SF (ed) A symposium on diseases of fishes and shellfishes. American Fisheries Society, Washington, DC, p 319–326

Winton JR (1991) Recent advaces in detection and control of infectious hematopoietic necrosis virus in aquaculture. Annu Rev Fish Dis 2:83–93

Winton JR, Arakawa CK, Lannan CN, Fryer JL (1988) Neutralizing monoclonal antibodies recognize antigenic variants among isolates of infectious hematopoietic necrosis virus. Dis Aquat Org 4:199–204

Wolf K (1988) Fish viruses and fish viral diseases. Cornell University Press, Ithaca, NY

Editorial responsibility: Jo-Ann Leong, Kaneohe, Hawaii, USA

Submitted: August 8, 2002; Accepted: April 11, 2003
Proofs received from author(s): July 11, 2003