



Theses and Dissertations

2006-08-07

The Phylogeography of Prosopium in Western North America

Becky Akiko Miller

Brigham Young University - Provo

Follow this and additional works at: <https://scholarsarchive.byu.edu/etd>



Part of the [Microbiology Commons](#)

BYU ScholarsArchive Citation

Miller, Becky Akiko, "The Phylogeography of Prosopium in Western North America" (2006). *Theses and Dissertations*. 1002.

<https://scholarsarchive.byu.edu/etd/1002>

This Thesis is brought to you for free and open access by BYU ScholarsArchive. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of BYU ScholarsArchive. For more information, please contact scholarsarchive@byu.edu, ellen_amatangelo@byu.edu.

THE PHYLOGEOGRAPY OF *PROSOPIUM* IN WESTERN
NORTH AMERICA

by

Becky Akiko Miller

A thesis submitted to the faculty of

Brigham Young University

in partial fulfillment of the requirements for the degree of

Master of Science

Department of Microbiology and Molecular Biology

Brigham Young University

December 2006

BRIGHAM YOUNG UNIVERSITY

GRADUATE COMMITTEE APPROVAL

of a thesis submitted by

Becky Akiko Miller

This thesis has been read by each member of the following graduate committee and by majority vote has been found to be satisfactory.

Date

R. Paul Evans, Chair

Date

Dennis K. Shiozawa

Date

Byron J. Adams

BRIGHAM YOUNG UNIVERSITY

As chair of the candidate's graduate committee, I have read the thesis of Becky Akiko Miller in its final form and have found that (1) its format, citations, and bibliographical style are consistent and acceptable and fulfill university and department style requirements; (2) its illustrative materials including figures, tables, and charts are in place; and (3) the final manuscript is satisfactory to the graduate committee and is ready for submission to the university library.

Date

R. Paul Evans
Chair, Graduate Committee

Accepted for Department

Date

Brent L. Nielsen
Department Chair

Accepted for College

Date

Rodney G. Brown
Dean, College of Biology and Agriculture

ABSTRACT

THE PHYLOGEOGRAPHY OF *PROSOPIUM* IN WESTERN NORTH AMERICA

Becky Akiko Miller

Department of Microbiology and Molecular Biology

Master of Science

The mountain whitefish (*Prosopium williamsoni*) has been largely overlooked in population genetic analyses despite its wide distribution in discrete drainage basins in western North America for over four million years. Its closest sister taxa the Bear Lake whitefish (*P. abyssicola*), Bonneville cisco (*P. gemmifer*), and Bonneville whitefish (*P. spilonotus*) are found only in Bear Lake Idaho-Utah and were also included in the analyses. A total of 1,334 cytochrome b and 1,371 NADH dehydrogenase subunit 2 sequences from the Bonneville Basin, the Columbia River Sub-basin, the lower Snake River Sub-basin, the upper Snake River Sub-basin, the Green River Basin, the Lahontan Basin, and the Missouri Basin were examined to test for geographically based genetic differentiation between drainage basins and sub-basins and phylogeographic relationships to determine the invasion route of *Prosopium* into western North

America and to aid in understanding current relationships. *Prosopium* entered the region via the Missouri River connection to Hudson Bay and moved in two waves: one colonized the lower Snake River Sub-basin, Columbia River Sub-basin, and the Lahontan Basin; the second wave colonized the upper Snake River Sub-basin, Bonneville Basin, Green River Basin, and established the Bear Lake *Prosopium*. Mountain whitefish exhibit a large amount of geographical genetic differentiation based on drainage basin except between the upper Snake River and the Bonneville Basin while the Bear Lake *Prosopium* show large amounts of gene flow between the three species. The apparent paraphyly of the mountain whitefish and the limited genetic structure of the Bear Lake *Prosopium* warrant recognition in the management of *Prosopium* and raise questions regarding species definitions in the group.

ACKNOWLEDGMENTS

I would like to thank the members of my graduate committee for their help and support during my graduate program. I'd like to especially thank Paul Evans for always looking out for me and letting me explore in his lab. Thank you to Dennis Shiozawa, Matthew McKell, Jared Crowley, Derek Houston, Bryce Nielson, Chad Crosby, Scott Tolentino, Rob Gipson, Hilda Sexauer, Bob Hughes, Andrew Whiteley, Bill Elmblad, Mike Sevon, Kim Tisdale, Pat Solleberger, Sam Finney, Lee Mabey, Steve Schram, Joel Hubbel, Linda Lamebull, and Louis Bernatchez for providing whitefish samples. And thank you to Byron Adams, Jenny Buhay, Keith Crandall, Keoni Kauwe, Jenny Pramuk, Rebecca Scholl, and Victoria Vance for helping me master techniques and assisting in analyses. And a special thanks to David Janetski and Andrew Johnson for all their various help and support and for keeping the lab a fun place to be. I'd like to thank my family for their constant love and support and for always believing in me. Thank you mom for always pushing me higher; dad for being my inspiration and guide; Natalie and Ulli for being my greatest friends; Kate for being my wise counselor for any situation; and Diane and Dennis for keeping life amusing. And most of all, I'd like to thank Kevin for being the most patient, supportive, and loving husband a girl could ever hope for.

TABLE OF CONTENTS

	<u>Page</u>
INTRODUCTION.....	1
MATERIALS AND METHODS.....	16
RESULTS.....	21
DISCUSSION.....	29
REFERENCES.....	39
FIGURE LEGENDS.....	48
TABLES.....	53
FIGURES.....	97

INTRODUCTION

Historical geological and climatic events are increasingly recognized as important influences on the evolutionary pathway of a species. They affect gene flow, selection, and drift and, therefore, are essential components when examining the evolution of an organism (Smith 1981; Johnson 2002). Information about past geological events aids interpretations of distribution patterns and genetic structuring of an organism and should not be overlooked in genetic analyses. Combining historical data with genetic data allows for a wider and more complete picture of an organism than can be seen with any single component.

Coregonids are a well studied example of the significant impacts of geological and climatic events on the evolutionary trajectory of a species. Lake ciscoes in North America (*Coregonus artedii*) evolved into two races in glacial refugia during Pleistocene glaciation 12,000-8,000 years ago (Turgeon & Bernatchez 2001a; Turgeon & Bernatchez 2001b). Lake whitefish from North America and central Europe (*Coregonus clupeaformis* and *C. lavaretus*) evolved into five races in glacial refugia during Pleistocene glaciation (Bernatchez & Dodson 1991; Bodaly *et al.* 1992; Foot *et al.* 1992; Bernatchez *et al.* 1999). Coregonids appear to be a plastic fish with the ability to undergo genetic divergence in relatively short periods of time.

While most research of the subfamily Coregoninae has focused on the genus *Coregonus*, the genus *Prosopium* has generally been overlooked. Phylogenetic studies tend to focus on the genus *Coregonus* and assume a monophyletic relationship within *Prosopium*, using only a few samples from each species within the genus (Norden 1970; Bernatchez *et al.* 1991; Vourinen 1998). As a result, the relationships among the *Prosopium* other than their general phylogenetic position relative to the genus *Coregonus*, have never been scrutinized. However, *Prosopium* have been present in western North America for at least 4.5 million years (Smith 1975; Smith

1981; Smith *et al.* 1982; G. R. Smith, personal communication), and today they occupy discrete drainage basins most of which have complex geological histories. If other coregonids are able to diverge genetically in short periods of time, perhaps *Prosopium*, with its long history in the region, is more complex than currently considered in the literature.

For this study, we examine the mountain whitefish (*P. williamsoni*) distributed throughout western North America and its three closest sister taxa found only in Bear Lake, Utah-Idaho: the Bear Lake whitefish (*P. abyssicola*), the Bonneville whitefish (*P. spilonotus*), and the Bonneville cisco (*P. gemmifer*). Using genetics and geological and climatic history, we will examine the phylogenetic relationship and current population structure of the mountain whitefish and the Bear Lake endemics and elucidate the pathway of *Prosopium* invasion into the western United States.

Whitefish Background

The mountain whitefish is found in the Columbia Basin, Missouri Basin, Bonneville Basin, Green River Basin, and Lahontan Basin (figure 1). It is morphologically conserved throughout its range, and its distribution is geographically limited by water temperatures and salinity (Whiteley *et al.* in press). They feed primarily on insects and tend to live in cold water in larger streams and rivers. Spawning time depends on the latitude and temperature of the stream or river, but it is usually between October and December in riffles (Sigler & Sigler 1987). The average length of mountain whitefish is 23-30 cm (9-12 inches; Behnke 2002).

The Bear Lake whitefish is difficult to distinguish morphologically from the Bonneville whitefish. The two have traditionally been separated based on differences in spawning times (Tolentino & Thompson 2004). However, a recent study by Tolentino and Thompson (2004) successfully separated Bear Lake whitefish and Bonneville whitefish by using scale counts at the

lateral line and above the lateral line. The Bonneville cisco is easily differentiated from the other Bear Lake *Prosopium* by its long, sharply pointed snout and slender body (Sigler & Sigler 1987).

Bear Lake whitefish feed on ostracods, chironomids, terrestrial insects, and fish eggs and prefers deep water near the bottom of Bear Lake, rarely frequenting shore areas (Sigler & Sigler 1987; Tolentino & Thompson 2004). They typically breed from late December to early February in water 15 to 30 m (50 to 100 feet) deep and rarely grow larger than 23 cm (9 inches; Sigler & Sigler 1987).

The Bonneville cisco is the most numerous fish in Bear Lake, growing no larger than 23 cm (9 inches). It feeds on zooplankton and chironomid larvae (Sigler & Sigler 1987). The Bonneville cisco prefers low water temperatures. It spreads throughout the lake when it is cold and retreats to deeper, colder water when the temperature warms (Sigler & Sigler 1987). The Bonneville cisco breeds from January 15-27th, plus or minus five days. Spawning takes place near shore in shallow water where the males remain throughout the spawning season. The females only move inshore when ripe and move back to deep water after spawning (Sigler & Sigler 1987).

The Bonneville whitefish feeds almost exclusively on chironomids but will also feed on ostracods, Bear Lake sculpin, and terrestrial insects (Tolentino & Thompson 2004). It can reach a length of 56 cm (22 inches; Sigler & Sigler 1987). They prefer water 12 to 30 m (40 to 100 feet) deep but are the most likely of the endemic *Prosopium* to inhabit shallow water (Sigler & Sigler 1987). Spawning occurs from mid February to early March and lasts only seconds before the fish resume travel with the school (Sigler & Sigler 1987).

While mountain whitefish and the Bear lake endemics are recognized game fishes, they have not been widely stocked as is the case with trout and salmon. Thus essentially all

populations retain their native genetic structure. Because mountain whitefish tend to live in colder, larger streams and rivers, they are more likely to move via major river transfer events rather than headwater transfers. With this in mind, a brief history for the Missouri River Basin, Columbia Basin, Bonneville Basin, Lahontan Basin, and Green River Basin and possible *Prosopium* invasion routes within the above drainage basins will be given.

Whitefish History

Movement from the Missouri River to the Western Snake River Plain

Fish movement in western North America was primarily constrained by North-South oriented mountain ranges and associated climatic barriers (Smith 1981). The region is thought to have been colonized during the late Miocene to early Pliocene by ancestral *Prosopium* from Hudson Bay when the Hudson Bay's drainage extended as far south as central Montana and central South Dakota. The upper Missouri River flowed northeast into Hudson Bay (Smith 1981).

The earliest known *Prosopium* fossils in the western United States occurred in Lake Idaho (Smith 1975). Lake Idaho was a series of mostly continuous lacustrine environments that occupied the western Snake River Plain during the Miocene and Pliocene. It can be divided into three general units: the Miocene Poison Creek Formation, the Miocene to early Pliocene Chalk Hills Formation, and the Pliocene Glens Ferry Formation (Malde 1991). From the Missouri River drainage, the fish had to move to the east or west side of the Idaho Batholith to gain access to Lake Idaho (Smith *et al.* 2000).

Based on geological and fish fossil data, Smith *et al.* (2000) suggests *Prosopium* invaded from the Missouri River westward around the Idaho Batholith via Salmon River headwaters, into Lake Idaho (figure 2). The upper Salmon River flowed northeast (instead of taking its current

sharp turn to the west just north of Salmon, Idaho) and was a tributary to the Missouri River. The capture of the upper Salmon River by the lower Salmon River (Anderson 1947) would have allowed a direct route for *Prosopium* to enter the lower Snake River.

The Big Lost River also flowed into the Salmon River at this time (figure 2; Anderson 1947) but eventually reversed its flow with the collapse of the Snake River Plain as the Yellowstone hotspot shifted to the east. The Big Lost River is therefore another possible route for *Prosopium* movement into Lake Idaho. Alternatively (but not exclusively), *Prosopium* could have invaded eastward, south of the Idaho Batholith, through headwater captures from the Missouri River to the Snake River headwaters and into Lake Idaho (figure 2; Smith *et al.* 2000).

It is important to note that the lower Snake River and the upper Snake River differ structurally and geologically (Malde 1991). The western Snake River Plain is a structural basin while the eastern Snake River Plain is a bimodal volcanic province (Malde 1991), generated by the movement of the Yellowstone hotspot. Biogeographic evidence supports the presence of a drainage divide between the western and eastern Snake River Plain during the Pliocene and suggests the Pliocene drainage of the eastern Snake River was probably southward to the Bonneville Basin. The headwaters of the eastern Snake River Plain were a Quaternary addition to the Snake River drainage basin (Malde 1991). Currently, the two sections are separated by a 65 m (212 feet) drop at Shoshone Falls near Twin Falls, Idaho, which has acted as an effective barrier between the upper and lower Snake River since the Pleistocene Bonneville Flood. Due to the stark differences in geological and biological composition and history between the east and west regions of the Snake River Plain, they will be treated as separate entities and their histories examined separately.

The earliest fossil evidence of *Prosopium* in the western United States is of an extinct large whitefish *P. prolixus* (Smith 1981). The earliest fossilized remnants of *P. prolixus* are found in the Pliocene Glens Ferry Formation, Idaho, and indicate *P. prolixus* thrived in the second stage of Lake Idaho. Smith dates the fossils around 4.5-3.5 million years ago (mya; Smith 1975; Smith 1981; Smith *et al.* 1982; G. R. Smith, personal communication). Fossilized remains of *Prosopium* were also discovered in a water well near Imbler, Oregon, in the Grande Ronde Valley of the Columbia drainage and date from the mid-Pliocene ~3.7-3.8 mya (Van Tassell *et al.* 2001).

The Snake River appears to have mainly drained through southeastern Oregon into northern California and into the Pacific Ocean during the late Miocene and early Pliocene (figure 2) based on the faunal similarity between the Great Valley in California and the Chalk Hills and Glenn Ferry stages of Lake Idaho (Smith 1981). Faunal similarity between Pliocene molluscs and fishes from Honey Lake, California, and fauna from Lake Idaho further support a connection between northeastern California and southern Idaho (Taylor & Smith 1981). However, evidence also exists for connections between the Snake River and the Columbia River during the Pliocene prior to the late Pliocene Snake River capture through Hells Canyon, as will be discussed below.

Movement from the Western Snake River Plain into the Columbia River

Lake Idaho drained when a Salmon River tributary eroded headward and captured a tributary of the Snake River at the Oxbow during the late Pliocene (~2 mya). This opened a drainageway through Hells Canyon into the Columbia River and marked the end of the Snake River's drainage through northern California/southern Oregon (figure 3; Malde 1991; Smith *et al.* 2000; Van Tassell *et al.* 2001). Alternative hypotheses for pathways and timing of connections between the Snake River and the Columbia River exist. There have been arguments

that the main outlet for Lake Idaho was through northeast Oregon via the Burnt, upper Powder, and Grande Ronde rivers prior to the Hells Canyon connection (Smith *et al.* 2000; Van Tassell *et al.* 2001). The Imbler fish fossils (including *Prosopium*) found in the Grande Ronde Valley, Oregon, most closely resemble the fish fauna of the Ringold Formation of eastern Washington and the Pliocene Glens Ferry Formation of Lake Idaho. These suggest a possible drainage connection between the Snake River Plain, Lake Idaho, and the Columbia River in the Grand Ronde Valley during that time (Van Tassell *et al.* 2001). Recent studies also indicate that the connection between the Snake River and Columbia River through Hells Canyon was established earlier in the Pliocene than previously thought. Fossil evidence of muskrat range expansion and fish fauna comparisons between the Ringold Formation in Pasco, Washington, and Pliocene Lake Idaho support such a connection prior to 3 mya (Smith *et al.* 2000).

Movement from the Western Snake River Plain into the Lahontan Basin

Taylor and Smith (1981), reject any connection between the western Snake River and the Lahontan system in Nevada during the Pliocene because the mollusc and fish fauna from Mopung Hills, Nevada, and Honey Lake, California, are not shared, indicating a separation between the two basins until the very end of the Pliocene. They argue that the current Lahontan Basin is a composition of separate basins with separate histories that merged as topographic barriers shrank from the late Pliocene onward (Taylor & Smith 1981). *Prosopium* may have entered the Lahontan system when rivers containing *Prosopium* were engulfed by the emergence of Lake Lahontan during the Quaternary (figure 4). Lake Lahontan was twice the second largest pluvial lake in the Western Hemisphere, first during the middle Pleistocene and again around 13,000 years ago (Morrison 1991). Lake Lahontan consisted of individual cycles of inundation,

recession, and desiccation for each valley system or basin, and by the middle and late Holocene most of the Lahontan system had dried out (Morrison 1991).

Using sedimentary zircon data, Link *et al.* (1999) found evidence for a connection between the Snake River, Big Wood River, Big Lost River, and the upper Humboldt River around 3 mya (figure 5). This connection between the Snake River and the upper Humboldt River of the Lahontan Basin could have allowed the transfer of *Prosopium* into the Humboldt River system. The only fossilized *Prosopium* remains associated with Lahontan are found in the dry bed of Owens Lake near Lone Pine, California, and are dated to 730,000 years ago (Firby *et al.* 1997).

Movement from the eastern Snake River Plain into the Bonneville Basin

The eastern Snake River Plain is a bimodal volcanic province formed by the uplift and subsequent collapse of the land around the Yellowstone hotspot as the North American plate drifted southwestward over it (Malde 1991). The Yellowstone hotspot was active at Heise volcanic field 7.0 mya to 3.5 mya and active at the Yellowstone Plateau volcanic field 2.5 mya to today (Perkins & Nash 2002). As land passes over the hotspot, silicic volcanism deposited rhyolite. Later basaltic flows overlaid the rhyolite. As a new area passes over the hotspot, new rhyolite will erupt and then it too will be covered over by erupting basalt. The continual fresh flow of basalt in the eastern Snake River Plain gives the plain a young Quaternary age and these flows often displaced the ancestral Snake River (Malde 1991).

As was noted earlier, *Prosopium* could have invaded eastward around the Idaho Batholith through headwater captures from the Missouri River into the Snake River headwaters (Taylor & Bright 1987; Smith *et al.* 2000). Using molluscan distribution data, Taylor and Bright (1987) hypothesized a headwater stream of the Madison or Gallatin River might have temporarily

flowed into the upper Snake River Plain and then into the Bonneville Basin as a result of the hotspot (figure 6). Alternatively, *Prosopium* may have entered the Bonneville Basin through the upper Bear River's connections to the Missouri River Basin via the Green River throughout the Eocene to the Pliocene (figure 6; Shiozawa & Rader 2005).

The Bear River can be divided into three geological regions: the northern portion dominated by late Cenozoic basalts of the Snake River Plain, the eastern portion dominated by Mesozoic strata, and the western portion dominated by Paleozoic limestone (Stokes 1979). The eastern portion of the Bear River joined the Green River as part of the Missouri River Basin from the late Eocene to the Pliocene (figure 6; Stokes 1979; Minckley *et al.* 1986). The northern portion of the Bear River was formed during Miocene uplift and faulting along the eastern edge of the Bear River Range and flowed into the Portneuf River (figure 6; Shiozawa & Rader 2005). Eventually, the upper Bear River was captured by the middle Bear River, severing its connection with the Green River (figure 7; Shiozawa & Rader 2005). The upper and middle Bear River remained a tributary to the Snake River until the late Pleistocene (Bouchard *et al.* 1998).

The Bear River's northwestern course into the Portneuf River was altered near Soda Springs, Idaho when lava flows in the northern portion of Thatcher Basin blocked the Bear River from the Portneuf/Snake River system at Portneuf Gorge (figure 8; Taylor & Bright 1987; Bouchard *et al.* 1998). The diversion of the Bear River into the Thatcher Basin was the result of a series of events, with the first diversion occurring ~140,000 ybp ago and filling the Thatcher Basin but not spilling into the Bonneville Basin (Bouchard *et al.* 1998). Around 100,000 ybp, lava flows had built the northern divide of Thatcher Basin high enough that when the Bear River re-entered the basin around 50,000 ybp, Lake Thatcher filled and spilled over its southern rim, cutting the Oneida Narrows, and entering Lake Bonneville in the Bonneville Basin (figure 8;

Bouchard *et al.* 1998). The Oneida Narrows were completely cut by ~20,000 ybp, draining Lake Thatcher (Bouchard *et al.* 1998). Eventually, Lake Bonneville backed up into the Thatcher Basin and allowed faunal transfers between the two basins (Taylor & Bright 1987; Bouchard *et al.* 1998).

Movement from the Bonneville Basin into the eastern Snake River Plain

Closed basin lakes have expanded and contracted in the Bonneville Basin for the past two million years (Madsen *et al.* 2001). Most surficial evidence of previous lakes was destroyed or altered by Lake Bonneville, which began its cyclic rise and fall from ~30,000 to 12,000 ybp (Oviatt & Currey 1987; Madsen *et al.* 2001). Around 28,000 ybp, Lake Bonneville's water level rose as 33% more water was added from the spillover of Lake Thatcher into the Bonneville Basin through the Oneida Narrows (Oviatt *et al.* 1992; Bouchard *et al.* 1998). The additional water input from Lake Thatcher, coupled with the pluvial climate, caused Lake Bonneville to overtop its rim at Red Rock Pass in southeastern Idaho, around 15,000 ybp (figure 9). This unleashed a catastrophic flood that dropped the lake level 108 meters, a volume of $4,700 \text{ m}^3$, in less than a year (at a constricted reach south of Boise, Idaho, the water discharge is estimated at $935,000 \text{ m}^3/\text{s}$; Malde 1991; Madsen *et al.* 2001). Water from Lake Bonneville flowed through Red Rock Pass, down Marsh Creek, into the Portneuf River and the Snake River where it surged all the way to the Columbia River. By ~12,000 ybp the Bonneville Basin was once again a closed basin and changes in climate led to its gradual desiccation (Taylor & Bright 1987; Oviatt *et al.* 1992).

Bear Lake

Bear Lake straddles the far eastern part of the Utah-Idaho border (figures 8 & 9) and is an active Neogene graben heavily affected by tectonicism and climate change with a modern

surface elevation of 1,805 m (5,922 feet). The lake has alternated between closed and open basin drainage throughout its existence, with its fluctuating water levels dependent on outside input from the Bear River or local groundwater discharge. The exact age of the lake is not known due to complications caused by active tectonism and geomorphic processes that affect elevation. A beach dated to one mya has been found although not much is known between one million and 300,000 years ago due to limited sediment exposures (Laabs & Kaufman 2003).

The Bear River currently flows from east to west just north of Bear Lake but does not connect to the lake except through a man-made canal constructed in 1912. Historically, the Bear River has connected to Bear Lake numerous times during lake highstands. According to Laabs and Kaufman (2003), Bear Lake highstands occurred twice during the middle Pleistocene: (1) between one million and 500,000 years ago and (2) 400,000-300,000 years ago. Highstands also occurred at least three times during the late Pleistocene: (1) 47,000-39,000 years ago, (2) 16,000-15,000 years ago, and (3) 9,000 years ago. The highstands may be due to local faulting in Bear Lake Valley that caused a lowering of the valley floor to the southeast and altered the lake-level either by raising the elevation of the northern outlet (due to the sinking of the southeastern end) and cutting off drainage through the outlet or by reversing or reducing the gradient of the Bear River, causing it to flow southward into Bear Lake. Eventual down-cutting by the out-flowing Bear River caused the lake level to drop and the north shoreline of Bear Lake to retreat southward. It is possible *Prosopium* were transferred from the Bear River to Bear Lake during one of the capture events.

The three Bear Lake *Prosopium* are thought to be the sister taxa to the mountain whitefish. The *Prosopium* endemics are not known to migrate out of Bear Lake, even with available routes through a pumping station and canal or across the warm, shallow waters of Mud

Lake (Sigler 1962). They do not depend on incoming streams as spawning grounds and will spawn over a number of different bottom types. Water fluctuations do not impact them greatly because they move in and out from shore with the water level fluctuations. However, the earliest (and only) fossil evidence of the Bear Lake endemic *Prosopium* comes from Lake Bonneville deposits and indicates past migration of the Bear Lake endemic *Prosopium* between Bear Lake and Lake Bonneville. *P. gemmifer* and *P. spilonotus* were described by Smith *et al.* (1968) at the Hot Springs locality and dated to be around 20,000 ybp (Broughton 2000). All three endemic *Prosopium* species were also found at Homestead Cave and dated to 11,200 to 10,100 ybp. These represent the final die-off of Lake Bonneville fish (Broughton 2000). Of interesting note are differences in jaw structure between the modern *P. gemmifer* and the Lake Bonneville fossilized remains of *P. gemmifer* (Smith *et al.* 1968; Broughton 2000). Broughton (2000) pointed out that *P. gemmifer* has since become more phenotypically similar to *P. spilonotus* and hypothesized ongoing hybridization between the two.

It is uncertain where the Bear Lake endemic *Prosopium* speciated. Broughton (2000) listed three possible speciation locations: Lake Bonneville, Pleistocene Lake Thatcher, or Bear Lake. The presence of the Bear Lake endemic *Prosopium* in Lake Bonneville indicates they did not evolve in Bear Lake during the Holocene.

While Lake Bonneville never reached Bear Lake Valley, the two lakes were only separated by 40 km (25 miles) and were connected at times during the Pleistocene and earlier by the Bear River (Laabs & Kaufman 2003). *Prosopium* may have entered Bear Lake through the Bear River, speciated, and then traveled downstream into Lake Thatcher and into Lake Bonneville. Or *Prosopium* may have entered Lake Thatcher through connections to the upper Snake River or the Bear River, speciated, and then accessed Bear Lake and Lake Bonneville via

the Bear River. Alternatively, *Prosopium* may have entered Lake Bonneville through connections with the upper Snake River, speciated, and then traveled upstream into Lake Thatcher and then Bear Lake.

Movement into the Green River Basin

The ancestral upper Green River flowed east (likely through the North Platte River) and into the Mississippi River (figure 6) as evidenced by early fish faunal similarities with the Mississippi Valley (Hansen 1985). The Green River Basin during the Eocene was warm and subtropical and interconnected lakes filled the basin. Climate changes and regional uplift during the late Eocene and Oligocene brought about the end of the Green River Formation lakes and the warm-water Mississippi fauna (Hansen 1985). Sometime between the Miocene and Pleistocene the upper Green River was captured by a small south flowing stream that slowly eroded headward to eventually capture the Green River and turn it southward to the Colorado River Basin (figure 7).

Hansen (1985) argued that the capture occurred later, during the Pleistocene, citing the absence of upper Missouri River fishes in the Colorado River system as evidence. However, if the North Platte-Green River connection was to Hudson Bay then cold-water fishes like *Prosopium* would be the likely invaders. Hansen (1985) favored a recent transfer of *Prosopium* into the Green River through connections with the Snake River rather than the Missouri River, placing the transfer sometime after the capture of the upper Green River by the lower Green River. Connections between the Bear River and the Green River and/or the Snake River and the Green River during the Quaternary when Basin and Range faulting was occurring could have allowed fish transfers (figure 7). The early eastern portion of the Bear River flowed northeast into the Green River through Sulphur Creek and Muddy Creek near Hilliard Flat, Wyoming

(Hansen 1985). Over time the Bear River was redirected northward, flowing into the Portneuf River. During this time Twin Creek of the Bear River and Hams Fork of the Green River may have been possible routes for invasion. Headwaters of the Hoback and the Gros Ventre Rivers may have been possible transfer points from the Green River to the Snake River (Hansen 1985).

Pleistocene glaciation

Pleistocene glaciation began 1.6 mya and lasted until 10,000 years ago, climaxing 23,000-18,000 years ago. Continental glaciers expanded southward out of Canada into parts of Montana and the Dakotas. The encroaching ice severed the Missouri River connection to Hudson Bay and blocked the Missouri River along its front. Peripheral drainages formed along the edge of the ice sheets, and the Missouri River eventually flowed southeast into the Mississippi River drainage, its current pathway (Howard 1958). With Hudson Bay cut off, new movements of *Prosopium* into the Basin and Range Province from Hudson Bay ceased; the *Prosopium* present in the northwestern United States were effectively isolated.

Despite the long and rich history of *Prosopium* and drainage basins in the northwestern United States, relationships of *Prosopium* within and between drainage basins have never been examined. In this study, the phylogenetics, population substructure, and genetic differentiation of *Prosopium* between drainage basins will be assessed using two mitochondrial genes: cytochrome b (cytb) and NADH dehydrogenase subunit 2 (ND2). Information on the phylogeny and population substructure based on drainage basins will help elucidate the pathway of whitefish invasion into the region and pose questions regarding speciation.

MATERIALS AND METHODS

Samples and Molecular Methods

Specimens were collected within the United States from the Missouri River Basin, Columbia River Basin, Green River Basin, Lahontan Basin, and Bonneville Basin. Samples were collected or provided by Dennis Shiozawa, Matthew McKell, and Jared Crowley, Brigham Young University; Derek Houston, University of Nevada Las Vegas; Bryce Nielson, Chad Crosby, and Scott Tolentino, Utah Division of Wildlife Resources; Rob Gipson and Hilda Sexauer, Wyoming Game & Fish; Bob Hughes, U. S. Environmental Protection Agency; Andrew Whiteley, University of Montana; Bill Elmblad, Colorado Division of Wildlife; Mike Sevon, Kim Tisdale, and Pat Solleberger, Nevada Department of Wildlife; Sam Finney, U. S. Fish and Wildlife Service; Lee Mabey, U. S. Forest Service; Steve Schram, Wisconsin Department of Natural Resources; Joel Hubbel, U. S. Bureau of Reclamation; Linda Lamebull, The Yakima Nation; Louis Bernatchez, Université Laval.

Whole fish samples were obtained through electroshocking, gill netting, and rod and reel. Muscle tissue from the right side and/or right pectoral fin was extracted and either frozen or stored in 95% ethanol until DNA isolations could be performed. Table 1 lists the populations, number of fish, and map location for this study (figure 10). In the analyses, the Columbia River Basin has been further divided into its sub-basins of the Columbia River Sub-basin, upper Snake River Sub-basin, and lower Snake River Sub-basin due to the different geological history of each region.

DNA was isolated using the PUREGENE DNA purification kit (Gentra Systems, Inc., Minneapolis, MN) for 183 individuals from the Bonneville Basin, 273 from the Columbia River Sub-basin, 125 from the lower Snake River Sub-basin, 253 from the upper Snake River Sub-basin, 139 from the Green River Basin, 92 from the Lahontan Basin, 157 from the Missouri River Basin, 51 Bear Lake whitefish, 55 Bonneville ciscoes, and 70 Bonneville whitefish.

The entire *cytb* (1,188 base pairs) and ND2 genes (1,050 base pairs) were amplified using polymerase chain reactions (PCR). *Cytb* and ND2 were amplified in two sections using four primers. The first half of *cytb* was amplified with 1425 and CYTB-intR and the second half with CYTB-intF and 1426. The internal primers were designed specifically for whitefish for this study while 1425 and 1426 were designed for salmonids (R. P. Evans, personal communication). The first half of ND2 was amplified with BYU11 and ND2-intR and the second half with ND2-intF and BYU12. The internal primers were designed specifically for whitefish for this study while BYU11 and BYU12 were designed for salmonids (R. P. Evans, personal communication). Primer sequences are listed in table 2.

PCR was performed in 20 µl reactions consisting of DNA template (~100 ng), deoxyribonucleotides (0.125 mM each), primers (10 pM each), buffer (10mM Tris-HCl, 1.5 mM MgCl₂, 25 mM KCl), and Taq polymerase (0.5 units) on an MJ Research PTC-225 Peltier TC tetrad (Bio-Rad Laboratories, Inc., Hercules, CA) with denaturing occurring at 94.0°C for 20 seconds, annealing at 47.0° C for 30 seconds, and elongation at 72.0° C for 1.5 minutes for 34 cycles. The resulting PCR product was cleaned using the GeneClean III protocol and kit (Q-Biogene, Carlsbad, CA). Cycle sequencing was performed using ABI Big Dye terminator protocol (Applied Biosystems, Inc., Foster City, CA) for 10 seconds at 96.0°C, 5 seconds at 50.0°C, and 4 minutes at 60.0°C for 24 cycles. The Big Dye product was cleaned with Sephadex G-50 medium (Sigma-Aldrich Co., St. Louis, MO). Samples were submitted to the Brigham Young University DNA Sequencing Center and sequenced on an ABI 377 automated sequencer.

Excluded Data

A nuclear gene was also examined for this study but did not yield useful data. The major histocompatibility complex (*Mhc*) A1 was investigated based on work with lake whitefish, which

indicated that the high polymorphism and central role in the immune response made the *Mhc* genes “highly suitable as markers in population and disease studies” (Binz *et al.* 2001). *Mhc* was also used in a population study of Chinook salmon by Miller *et al.* (1997).

Five individuals were sampled for each basin/sub-basin: the Bonneville Basin, Green River Basin, Columbia River Sub-basin, lower Snake River Sub-basin, upper Snake River Sub-basin, Lahontan Basin, and Missouri River Basin. Four individuals for each of the Bear Lake *Prosopium* species were also examined. The *Mhc* A1 was amplified using primers and thermal cycler programs described by Binz *et al.* (2001), and the resulting PCR products were cloned using a TOPO TA Cloning Kit (Invitrogen Co., Carlsbad, CA). Clones were grown on agar plates with nutrient broth and 20 clones from each PCR product were sequenced.

Limited sequence diversity was found for *Prosopium*. While many alleles were found per individual (up to 15 alleles), sequence diversity was limited across the locus based on drainage basin/sub-basin and species. These findings support conclusions of Miller and Withler (1998). Using *Mhc* class I genes in Atlantic salmon and Pacific salmon, they found that while individuals may carry up to 30 alleles at a single locus, there was little sequence diversity between species. Due to the lack of allelic sequence diversity between basins/sub-basins in mountain whitefish and between the three Bear Lake *Prosopium*, *Mhc* A1 was excluded from the analyses.

Data Analyses

Sequences were trimmed and aligned in Sequencher 4.2 (GeneCodes Co., Ann Arbor, MI) with alignment adjustments made by eye with the aid of amino acid sequences for *Prosopium*. The entire *cytb* (1,161 base pairs) and *ND2* (1,050 base pairs) genes were examined. A total of 1,334 sequences for *cytb* and 1,371 sequences for *ND2* were obtained (table 1) with 257 unique *cytb* haplotypes and 303 unique *ND2* haplotypes found (tables 3 and 4).

Phylogenetic Analyses

Only unique haplotypes were used in the phylogenetic analyses to minimize computational time. *P. coulteri* and *P. cylindraceum* sequences were used as outgroups for all analyses. Phylogenetic relationships were reconstructed using Bayesian and maximum-likelihood methods. Bayesian analyses were run with MrBayes 3.1.2 (Hulsenbeck *et al.* 2001; Ronquist & Hulsenbeck 2003) with 10 million generations using 4 chains and sampling every 1,000 trees with the number of substitutions at 6 and a gamma rate. The burn-in was graphed using Microsoft Excel and those trees discarded. Bayesian analyses were replicated 3 times with different random start trees. Maximum-likelihood analyses were performed with PHYML (Guindon & Gascuel 2003) using parameters calculated by Modeltest 3.7 (Posada & Crandall 1998). Four runs were performed for each data set with a different random start tree for each run and 1,000 bootstrap replicates. For *cytb* and ND2, the model GTR+I+G was selected by Modeltest 3.7 based on AIC criterion. A majority rule consensus tree for all the bootstrap replicates was made in PAUP 4.0b10 (Swofford 2001).

Genetic Structure Analyses

Arlequin 2.000 (Schneider *et al.* 2000) was used to measure current gene flow and effective migration rates between the basins/sub-basins for all *cytb* and ND2 sequences. Population pairwise F_{st} values with 100 permutations for significance and 1,000 permutations for the Mantel test were calculated. F_{st} measures the ratio of gene flow to drift and gives insight into the mixing among populations. An F_{st} of 1.0 means the populations are highly structured with no mixture between them. An F_{st} of 0.0 indicates there is complete mixture between populations and no genetic structure to populations. AMOVA (analysis of molecular variance) with 1,000 permutations was calculated to estimate variation within and among populations.

Nested Clade and GeoDis Analyses

Nested clade analysis (NCA; Templeton *et al.* 1995; Templeton 1998) was used to identify population divisions and relationships and to gain insight into historical events. TCS 1.21 (Templeton *et al.* 1995; Templeton 2004) was used to generate a haplotype network of the unique haplotypes. A 95% probability of parsimony connection limit was estimated as well as a forced compilation of all networks were calculated. Any ambiguities in the network were resolved by 3 criteria based on coalescent theory: 1) haplotypes are more likely to be connected to haplotypes from the same population or region; 2) haplotypes are more likely to be connected to interior/ancestral haplotypes; and 3) haplotypes are more likely to be connected to haplotypes with a higher frequency (Pfenninger & Posada 2002). Geodis 2.5 (Posada *et al.* 2000) was used to test for significant associations between geographic locations and genetic distances with 100,000 random permutations. Geographical sampling locations rather than river distances were used due to the long time scale (over 4.5 million years) examined and the aptitude of rivers to change course over such a long period of time. Historical inferences were made based on the November 2005 inference key by Templeton (<http://darwin.uvigo.es/software/geodis.html>).

RESULTS

Phylogenetic Analyses

Neither the Bayesian (figure 11) nor the maximum-likelihood (figure 12) methods for *cytb* show much resolution between basins/sub-basins. While relationships are not clearly defined, the Green River Basin, Lahontan Basin, and Missouri River Basin do break into clades. The Missouri River Basin clade in the maximum likelihood analysis is basal to the rest of the whitefish although this relationship is not supported by other methods or with ND2 data. The

Bear Lake endemics separate into two clades: one of Bear Lake whitefish and Bonneville whitefish and joins with the Columbia River Sub-basin/lower Snake River Sub-basin/Lahontan Basin clade, and the second consists of primarily Bonneville cisco with some Bear Lake whitefish and Bonneville whitefish haplotypes and joins with the Bonneville Basin/upper Snake River Sub-basin/Green River Basin clade. The Lahontan Basin is divided into two clades: one group contains haplotypes from only the Truckee River, and the other contains haplotypes from the Walker River, the Carson River, the East Carson River, and the Truckee River. Haplotypes between the Bonneville Basin and upper Snake River Sub-basin, the Columbia River Sub-basin and lower Snake River Sub-basin, and the Bear Lake *Prosopium* are shared or intermixed with no clear geographic relationship. Haplotypes from certain populations appear isolated from other rivers in their respective basin/sub-basin. Haplotypes from the Big Lost River (upper Snake River Sub-basin), the Hoh River (Columbia River Sub-basin), and the Big Wood River (lower Snake River Sub-basin) form their own unique clades that only include haplotypes from their respective river systems and do not contain haplotypes outside their clades.

The ND2 region shows more resolution than cytb. Both Bayesian (figure 13) and maximum likelihood (figure 14) phylogenies divide the whitefish into two clades. One clade consists of the Bonneville Basin, upper Snake River Sub-basin, Green River Basin, the Bear Lake *Prosopium*, and Missouri River Basin. The other clade consists of the Columbia River Sub-basin, lower Snake River Sub-basin, and Lahontan Basin. Within the Columbia River Sub-basin/lower Snake River Sub-basin/Lahontan Basin clade, only the Lahontan Basin haplotypes form their own group. The Columbia River Sub-basin and lower Snake River Sub-basin are intermixed. Within the Bonneville Basin/upper Snake River Sub-basin/Green River Basin/Missouri River Basin/Bear Lake *Prosopium* clade, the Missouri River Basin appears to be

the most basal group. The maximum likelihood method does not resolve the remainder of the Bonneville Basin/upper Snake River Sub-basin/Green River Basin/Bear Lake *Prosopium* clade. The Bayesian method gives the most resolution and places the Bear Lake *Prosopium* between the Missouri River Basin clade and the Bonneville Basin/upper Snake River Sub-basin/Green River Basin clade. The fact that both cytb and ND2 place the Bear Lake *Prosopium* within the mountain whitefish indicates that the mountain whitefish are paraphyletic.

The Bear Lake *Prosopium* also fail to break into clades based on species. A small clade of upper Snake River Sub-basin haplotypes from the Lost River system breaks off between the Bonneville Basin/upper Snake River Sub-basin/Green River Basin clade and the Bear Lake *Prosopium*. The Green River Basin mixed with some random haplotypes from the upper Snake River Sub-basin is the next to branch off from the Bonneville Basin/upper Snake River Sub-basin/Green River Basin clade. The Bonneville Basin and upper Snake River Sub-basin haplotypes are intermixed and not resolved. Haplotypes from the Big Lost River (upper Snake River Sub-basin), the Hoh River (Columbia River Sub-basin), and the Big Wood River (lower Snake River Sub-basin) appear isolated from other rivers in their respective basin/sub-basin and form their own unique clades.

Another interesting clade that is seen in all the cytb and ND2 analyses consists of haplotypes from the Yakima River (Columbia River Sub-basin), the Salmon River (lower Snake River Sub-basin), South Fork Salmon River (lower Snake River Sub-basin), and the Big Hole River (Missouri River Basin). While these rivers share haplotypes outside this clade, the consistent grouping of haplotypes from these rivers is instructive for further study.

Genetic Structure

Fst values for cytb and ND2 show a high amount of genetic structure for mountain whitefish based on drainage basin/sub-basin. Table 5 shows the pairwise Fst values by population and table 6 shows the effective migration rate per generation by population. Gene flow and migration tend to occur within basins/sub-basins and between populations from the Bonneville Basin and upper Snake River Sub-basin, the Columbia River Sub-basin and lower Snake River Sub-basin, and Bear Lake whitefish and Bonneville whitefish. Populations that have little to zero gene flow or migration and appear isolated even within their basin/sub-basin include the Hoh River, Duchesne River, Big Wood River, Big Lost River, and the Bonneville cisco. Average Fst and effective migration values by basin/sub-basin are given in tables 7 and 8 and further illustrate the high level of structure in *Prosopium* except between the Bonneville Basin and upper Snake River Sub-basin, the Columbia River Sub-basin and lower Snake River Sub-basin, and between Bear Lake whitefish and Bonneville whitefish. The AMOVA for cytb shows 75.81% of the total variation is found between drainage basins/sub-basins and indicates a high level of genetic structure (table 9).

ND2 has more genetic variation than cytb and shows more gene flow. Table 10 shows the pairwise Fst values by population and table 11 shows the effective migration rate per generation by population. Gene flow and migration tend to cluster within basins/sub-basins, but ND2 shows a limited amount of gene flow and migration between basins/sub-basins as well. The Bear Lake whitefish is the most widespread of the populations and has gene flow and migration with the other Bear Lake *Prosopium*, the Bonneville Basin, the upper Snake River Sub-basin, the Green River Basin, the Columbia River Sub-basin, and the Missouri River Basin. Large amounts of gene flow and migration occur between the Bonneville Basin and upper Snake River Sub-basin and the Columbia River Sub-basin and lower Snake River Sub-basin. However, this is likely an

artifact of relict haplotypes or multiple hits and not an accurate indication of gene flow. Lesser amounts of gene flow and migration occur between the Green River Basin and the Bonneville Basin and Bonneville whitefish. The Lahontan Basin has migration and gene flow with the lower Snake River Sub-basin and a lesser amount with the Columbia River Sub-basin. The Coeur d'Alene River and Big Lost River appear isolated even within their basin/sub-basin but show gene flow with Bear Lake whitefish. Average F_{st} and effective migration values by basin/sub-basin are given in tables 12 and 13. F_{st} values are high except between the Bonneville Basin and upper Snake River Sub-basin, the Columbia River Sub-basin and Lahontan Basin, the Columbia River Sub-basin and lower Snake River Sub-basin, the Lahontan Basin and lower Snake River Sub-basin, and between the three Bear Lake endemics. Migration rates per generation are greater than 1.0 for those groups as well, with the largest number of migrants occurring between the Columbia River Sub-basin and lower Snake River Sub-basin (10 migrants per generation) and between the Bear Lake endemics. The AMOVA for ND2 shows 62.45% of the total variation is found between drainage basins/sub-basins and indicates a high level of genetic structure (table 14).

Nested Clade and GeoDis Analyses

The nested clade analyses for *cytb* and ND2 did not fully connect at the 95% probability of parsimony connection limit. At a 95% probability of parsimony connection limit for *cytb*, 4 general clades were formed: a Missouri River Basin clade, a Bear Lake *Prosopium* clade, a Columbia River Sub-basin/lower Snake River Sub-basin/Lahontan Basin clade, and a Bonneville Basin/upper Snake River Sub-basin/Green River Basin/Bear Lake *Prosopium* clade. To fully connect these four clades, up to 25 steps were allowed. ND2 at the 95% probability of parsimony connection limit formed two major clades: a Columbia River Sub-basin/lower Snake River Sub-

basin/Lahontan Basin clade and a Bonneville Basin/upper Snake River Sub-basin/Green River Basin/Missouri River Basin/Bear Lake *Prosopium* clade. To fully connect the two clades, up to 25 steps were allowed. The entire network for both cytb and ND2 were resolved in seven steps and are shown in figures 15 (a-d) and 16 (a-d).

For both cytb and ND2, haplotypes from the Big Lost River (upper Snake River Sub-basin), the Big Wood River (lower Snake River Sub-basin), and the Hoh River (with one haplotype from Skokomish River; Columbia River Sub-basin) form isolated clades that include only haplotypes from their river and do not have haplotypes outside this clade. Certain haplotypes from the Yakima River (Columbia River Sub-basin), the Salmon River (lower Snake River Sub-basin), the South Fork Salmon River (lower Snake River Sub-basin), and the Big Hole River (Missouri River Basin) consistently nest together in both the cytb and ND2 nested clade analyses and mirror findings in the phylogenetic analyses.

Confidence in the GeoDis analyses are restricted due to the forced connection of cytb and ND2 haplotypes at less than the 95% probability of parsimony connection limit; however, a general idea of population history can be inferred from the results. Table 15 lists the average clade distance (D_c), the nested clade distance (D_n), and the interior-tip (I-T) distances and whether the values were significantly large or small based on the permutations test for each haplotype/clade for cytb for those clades with a significant X^2 probability. Inferences for the cytb GeoDis analyses with a significant X^2 probability are given in table 16. Due to the large number of haplotypes/clades in the GeoDis analyses, only descriptive nestings will be reported here.

At the 6th step, the two major clades formed are (1) the Bonneville Basin/upper Snake River Sub-basin/Green River Basin/Missouri River Basin/Bear Lake *Prosopium* and (2) the Columbia River Sub-basin/lower Snake River Sub-basin/Lahontan Basin. The Bonneville

Basin/upper Snake River Sub-basin/Green River Basin/Missouri River Basin/Bear Lake *Prosopium* clade underwent restricted gene flow with long distance dispersal. The Bonneville Basin/upper Snake River Sub-basin/Green River Basin/Missouri River Basin clade underwent restricted gene flow with isolation by distance. The Bonneville Basin/upper Snake River Sub-basin/Green River Basin/Big Hole River (Missouri River Basin) clade went through long distance colonization followed by fragmentation. The Missouri River Basin group spread via contiguous range expansion. The Green River Basin clade is nested with haplotypes from the upper Snake River Sub-basin (Hoback River, Jackson Lake, Snake River, Henry's Fork, Salt River) which underwent restricted gene flow with isolation by distance. The Green River Basin haplotypes spread through restricted gene flow with isolation by distance. The Big Lost River system is nested with upper Snake River Sub-basin haplotypes from Henry's Fork, Salt River, and Snake River and underwent restricted gene flow and long distance dispersal. The Columbia River Sub-basin/lower Snake River Sub-basin/Lahontan Basin/Big Hole (Missouri River Basin) clade underwent restricted gene flow with isolation by distance. The Lahontan Basin was colonized by Columbia River Basin haplotypes (likely from the Willamette and Santiam Rivers) through long distance colonization and was followed by fragmentation. Within the Lahontan Basin system, restricted gene flow formed current populations. The isolated Hoh River population (with one Skokomish River haplotype) entered the area with the contiguous range expansion of whitefish in the Columbia River Sub-basin system. The isolated Big Wood River population underwent restricted gene flow with isolation by distance from the rest of the lower Snake River Sub-basin.

Table 17 lists the average clade distance (D_c), the nested clade distance (D_n), and the interior-tip (I-T) distances and whether the values were significantly large or small based on the

permutations test for each haplotype/clade for ND2 for those clades with a significant X^2 probability. Inferences for the ND2 GeoDis analyses with a significant X^2 probability are given in table 18. At the 6th step, two major clades are formed: (1) the Bonneville Basin/upper Snake River Sub-basin/Green River Basin/Missouri River Basin/Bear Lake *Prosopium* and (2) the Columbia River Sub-basin/lower Snake River Sub-basin/Lahontan Basin. The Bonneville Basin/upper Snake River Sub-basin/Green River Basin/Missouri River Basin/Bear Lake *Prosopium* clade underwent long distance colonization followed by fragmentation and range expansion. The Bonneville Basin/upper Snake River Sub-basin/Green River Basin clade underwent long distance colonization possibly coupled with subsequent fragmentation followed by range expansion. The Green River Basin haplotypes are nested with upper Snake River Sub-basin haplotypes from Hoback River, Jackson Lake, South Fork Snake River, Henry's Fork, Salt River and underwent long distance colonization possibly coupled with subsequent fragmentation followed by range expansion. Within the Green River Basin, haplotypes spread via restricted gene flow with isolation by distance. Missouri River Basin and the Bear Lake *Prosopium* underwent long distance colonization followed by fragmentation then range expansion. Within the Missouri River Basin, the populations show evidence of restricted gene flow with isolation by distance. The Big Lost River system underwent long distance colonization followed by fragmentation. The Columbia River Sub-basin/lower Snake River Sub-basin/Lahontan Basin/Big Hole (Missouri River Basin) clade underwent contiguous range expansion. The Lahontan Basin system is most closely associated with Columbia River Sub-basin haplotypes from the Warm Springs River, the Crooked River, and the Deschutes River and appears to have undergone long distance colonization followed by fragmentation. Within the Lahontan Basin, whitefish haplotypes spread via restricted gene flow with long distance dispersal over areas not previously

occupied by whitefish. The isolated Hoh River population (with one Skokomish River haplotype) was colonized by long distance dispersal followed by restricted gene flow. In the ND2 analyses, the Big Wood River system is most closely associated with the Lahontan Basin and was colonized over long distances followed by fragmentation.

DISCUSSION

Pathway of invasion

Two *Prosopium* invasions appear to have taken place into the western United States: one founded the Columbia Basin/lower Snake River Sub-basin/Lahontan Basin line, and the other founded the Bonneville Basin/upper Snake River Sub-basin/Green River Basin lineage. It is uncertain where the Missouri River Basin lies since *cytb* maximum likelihood phylogenetic analyses place the Missouri River Basin clade as the most basal group of the mountain whitefish and Bear Lake *Prosopium*, and ND2 phylogenetic analyses place the Missouri River Basin clade with the Bonneville Basin, upper Snake River Sub-basin, and Green River Basin (as the most basal group in this clade). It is possible that two waves of *Prosopium* entered the region via Missouri River drainage connections with Hudson Bay (Smith 1981). The phylogenetic and nested clade analyses for both *cytb* and ND2 show a clade in the Columbia River Sub-basin/lower Snake River Sub-basin/Lahontan Basin group with haplotypes from the Big Hole River (Missouri River Basin), Salmon River (lower Snake River Sub-basin), South Fork Salmon River (lower Snake River Sub-basin), and the Yakima River (Columbia River Sub-basin). This grouping suggests a wave of *Prosopium* moved to the west via a route north of the Yellowstone hotspot through the Missouri River drainage's Big Hole River into the Salmon River and Snake River (Anderson 1947; Smith *et al.* 2000). This movement or additional movements of

Prosopium from the Missouri River Basin to the lower Snake River Sub-basin founded the Columbia River Sub-basin/lower Snake River Sub-basin/Lahontan Basin group.

The earliest *Prosopium* fossils are of *P. prolixus* in Lake Idaho ~4.5 mya and represent the earliest evidence of *Prosopium* in western North America (Smith 1981). Lake Idaho drained west through California until 2 mya when a river capture event caused Lake Idaho and the Snake River to drain into the Columbia River (Malde 1991). This would have allowed the transfer of *Prosopium* throughout the Columbia River Basin. The Lahontan Basin clade of *Prosopium* is a more recent transfer event than anticipated based on phylogenetic analyses and NCA. NCA indicates the lower Columbia River is the origin of the Lahontan Basin group rather than the lower Snake River. During the Pleistocene, large pluvial lakes filled central Oregon (Hubbs & Miller 1948). Pluvial Lake Alvord in the Alvord Basin in Oregon, east of the Pueblo Mountains, is known to have connections to the Lahontan Basin based on the distribution of based on distributions of the now extinct Alvord cutthroat trout (Hubbs & Miller 1948; Behnke 2002). *Prosopium* from the lower Columbia River could have migrated southward through rivers interconnecting the Oregon lakes to enter the Lahontan Basin during the mid-Pleistocene rise of Lake Lahontan (Morrison 1991).

While one group of *Prosopium* headed west around the Yellowstone hotspot, a second wave of *Prosopium* from the Missouri River drainage moved east around the Yellowstone hotspot and into the upper Snake River, founding the Bonneville Basin/upper Snake River Sub-basin/Green River Basin clade. This group was likely separated from the first wave after the Salmon River was captured and diverted westward during the late Tertiary (Anderson 1947). Temperature barriers and/or volcanic activity from the Yellowstone hotspot may have delayed the immediate southward movement of *Prosopium* into the upper Snake River, but when the

barriers were removed, *Prosopium* moved into the upper Snake River possibly through headwaters of the Gallatin or Madison Rivers (Taylor & Bright 1987). Once in the upper Snake River, *Prosopium* entered the Bonneville system via the middle Bear River's connection with the Portneuf River.

Prosopium entered the Green River Basin after the ancestral Green River was captured by the south flowing lower Green River and entered via connections with the Bonneville Basin and/or the upper Snake River. The early Bear River flowed northeast into the Green River through Sulphur Creek and Muddy Creek and is a possible invasion route (Hansen 1985). As the Bear River was redirected northward into the Portneuf River, it connected with the Green River through Twin Creek and Hams Fork (Hansen 1985). Between the Snake River and Green River, headwaters of the Hoback and the Gros Ventre Rivers may have been possible transfer points (Hansen 1985). Some upper Snake River haplotypes group with the Green River haplotypes, indicating the ancestral upper Snake River *Prosopium* were the likely colonizers of the Green River Basin.

Three populations of mountain whitefish appear isolated from populations in the rest of their basins. The Big Lost River is the most unique drainage of the upper Snake River Sub-basin due to its geological history. Regional uplift from the Yellowstone hotspot caused the Lost River to flow into the Salmon River (Link 1999). As the hotspot moved away from the region, the previously uplifted area began to collapse. As the region collapsed the Big Lost River fell with it, tipping directions and flowing into the upper Snake River Plain (Link 1999). Basaltic flows from volcanic rift zones on the Snake River Plain during the Pleistocene isolated the Lost Rivers streams, forming the Big Lost Trough. *Prosopium* entered the Lost River system from the upper Snake River but have been isolated within the Big Lost Trough long enough to diverge

genetically. Although distinct from other upper Snake River Sub-basin populations, the Big Lost River mountain whitefish still group with other populations from the upper Snake River Sub-basin.

The Big Wood River population is genetically distinct from other lower Snake River Sub-basin populations. The Big Wood River is partially isolated from other rivers due to steep rapids that do not allow fish movement across them and areas that are dry except during exceptionally wet years (Hubbs & Miller 1948). Although a unique population, the mountain whitefish from the Big Wood River group with other populations from the lower Snake River Sub-basin.

The Hoh River population shows some isolation from other populations found in the Columbia River Sub-basin. The Hoh River haplotypes do cluster with one haplotype from the Skokomish River but do not seem to mix with any other populations from the Columbia River Sub-basin. This is likely due to the Hoh River's isolation on the coast of Washington. In order for whitefish to enter the Hoh River, they had to travel through interconnecting freshwater bays and streams along the Washington coast. It would be very difficult for the saline intolerant mountain whitefish to disperse along the coast and, once established, the population would be isolated.

The Bear Lake Endemics

Three possible speciation locations have been suggested for the Bear Lake *Prosopium*: Bear Lake (~1 mya), Lake Thatcher (~140,000 ybp), and Lake Bonneville (~30,000 ybp). The numerous haplotypes found for each Bear Lake *Prosopium* points to Bear Lake as the place of origin as it would be difficult for all of the haplotypes to make it to Bear Lake from Thatcher Lake or Lake Bonneville. *Prosopium* most likely entered Bear Lake during a highstand and

diverged. Later, the three Bear Lake *Prosopium* entered the Bear River during another highstand (likely between 47,000-39,000 ybp; Laabs & Kaufman 2003) and moved downstream in the Bear River to Lake Thatcher, which was connected with the Bear River 50,000-20,000 years ago. The draining of Lake Thatcher through the Oneida Narrows into Lake Bonneville ~28,000 years ago brought the Bear Lake *Prosopium* into the lake. As fossil evidence shows, the Bear Lake *Prosopium* resided in Lake Bonneville until the desiccation of Lake Bonneville. It does not appear that the Bear Lake endemics survive outside of lake systems, even though the catastrophic Bonneville flood through Red Rock Pass 15,000 years ago allowed faunal transfer to the Snake River. Today, Bear Lake *Prosopium* are not known outside of Bear Lake, even in the Bear River.

The Bear Lake whitefish, Bonneville cisco, and Bonneville whitefish are thought to be the closest sister taxa to the mountain whitefish, but they are embedded within the mountain whitefish clade. The Bear Lake *Prosopium* clade fails to separate into discrete clades of their respective species and shows high levels of gene flow, largely between the Bear Lake whitefish and the Bonneville whitefish. Three possible explanations are, first, the Bear Lake *Prosopium* may still be undergoing speciation and have not yet experienced lineage extinction. This would explain the rampant shared polymorphisms. Second, gene flow may be occurring among the Bear Lake *Prosopium*, although our use of maternally inherited mitochondrial DNA markers should dampen the obvious effects of hybridization. However, laboratory hybridization studies by White (1974) have shown that the Bear Lake *Prosopium*, if simultaneously ripe, can readily hybridize with one another. Mountain whitefish were also found to successfully hybridize with Bear Lake whitefish and Bonneville whitefish, but crosses between mountain whitefish and Bonneville cisco were less successful (White 1974). Third, samples may have been misidentified

as the incorrect species. Even if this were the case, three distinct clades of any construction are not present in the Bear Lake *Prosopium*.

The Bear Lake *Prosopium* may not be genetically differentiable, but their morphological, ecological, and behavioral differences are real, and the situation is not unique to the Bear Lake *Prosopium*. Salmonids are known at times to respond to their surrounding environment by undergoing ecophenotypic differentiation. Lake whitefish (*Coregonus clupeaformis* and *C. lavaretus*), ciscoes (*Coregonus artedi* and other spp.), Arctic charr (*Salvelinus alpinus*), brown trout (*Salmo trutta*), Atlantic salmon (*Salmo salar*), and sockeye salmon (*Oncorhynchus nerka*) all have examples of coexisting morphotypes (Pigeon *et al.* 1997). Adaptive radiation is implicated in lake whitefish and cisco ecomorphotypes (Bernatchez *et al.* 1999; Lu & Bernatchez 1999; Turgeon & Bernatchez 2003). According to adaptive radiation theory, divergent natural selection is driven by environmental differences between populations or subpopulations, leading to (1) phenotypic divergence to utilize the resource, (2) speciation driven by differences in resources and competition, and (3) reproductive isolation (Schluter 2001). Bear Lake *Prosopium* may be in the process of undergoing adaptive radiation as other coregonids in its subfamily have done.

Lake whitefish from North America and central Europe (*Coregonus clupeaformis* and *C. lavaretus*) diverged into five geographic races in different refugia during Pleistocene glaciation (Bernatchez & Dodson 1991; Bodaly *et al.* 1992; Foote *et al.* 1992; Bernatchez *et al.* 1999). As the glaciers receded, the lake whitefish races expanded out of their refugia to colonize new lakes and there independently diverged into a normal and dwarf phenotype where the extent of the phenotypic divergence was driven by the strength of the resource-based divergent natural selection in the given environment (Bodaly *et al.* 1992; Bernatchez *et al.* 1996; Pigeon *et al.*

1997; Bernatchez *et al.* 1999; Lu & Bernatchez 1999; Lu *et al.* 2001). The dwarf phenotype feeds on plankton, matures by the age of 1 or 2, and rarely lives past 4 years, while the normal phenotype does not mature until 4 years of age and lives to be 12 years old (Lu *et al.* 2001). The dwarf phenotype does not exceed 20 cm in length and 100 g in weight, while the normal phenotype exceeds 40 cm in length and 1000 g in weight (Lu *et al.* 2001). The dwarf form generally occupies the limnetic trophic zone and the normal the benthic trophic zone (Lu *et al.* 2001). Size and growth differences provided pre-mating mechanisms of reproductive isolation, and the dwarf and normal phenotype diverged genetically (Bodaly *et al.* 1992; Lu & Bernatchez 1999). Lake to lake variations affect the speed of the evolutionary process, but the process can occur rapidly (within 18,000 years for lake whitefish; Bernatchez *et al.* 1999). Each lake can be seen as a different temporal point of ongoing adaptive radiation (Lu & Bernatchez 1999).

Ciscoes in North America (*Coregonus artedii*, *C. alpenae*, *C. johanna*, *C. hoyi*, *C. kiyi*, *C. nigripinnis*, *C. reighardi*, and *C. zenithicus*) have long been a taxonomic headache due to their extreme morphological and ecological variation (Bernatchez *et al.* 1996; Turgeon *et al.* 1999; Turgeon & Bernatchez 2003). The ciscoes diverged into two races during Pleistocene glaciation (Turgeon & Bernatchez 2001a; Turgeon & Bernatchez 2001b). The ciscoes expanded into new lakes following deglaciation and underwent adaptive radiation (Bernatchez *et al.* 1996; Turgeon *et al.* 1999). A planktivorous high-gill-rakered form and several low-gill-rakered forms of cisco evolved in postglacial lakes with behavioral, habitat, ecophenotypic differences and differences in time and depth of reproduction (Turgeon *et al.* 1999). Unlike the lake whitefish, the ciscoes have evolved into multiple morphotypes. Five morphotypes are found only in the Great Lakes and the other three are widespread (Turgeon *et al.* 1999). These ecotypes evolved from

polyphyletic origins and were derived independently more than once and are still undergoing divergence and sorting (Bernatchez *et al.* 1996).

It appears whitefish in lake systems are prone to ecological differentiation and this may be reflected with mountain whitefish and the Bear Lake *Prosopium*. *Prosopium* from the upper Snake River that entered Bear Lake may have exploited the various ecological opportunities to diverge into the three endemic species, each developing morphological traits and behavioral adaptations that would best help them exploit their environmental niche. This could have happened fairly quickly due to the plasticity of whitefish, leaving the plethora of haplotypes and perplexing genetic signals seen today in the Bear Lake *Prosopium*. *Prosopium* in stream and river systems were already adapted to their environment and had little need for morphological change but diverged genetically through isolation and genetic drift (although Holt (1960) did find slight morphologically variations centered around habitat differences in populations of mountain whitefish but few were based on basin locations).

The Bear Lake *Prosopium* may be in the process of evolving under resource-based natural selection and lineage sorting has not fully occurred. Current populations of Bear Lake *Prosopium* may be a snapshot of adaptive radiation in progress where reproductive isolation and genetic divergence have not yet occurred. However, more studies on the ecology of Bear Lake and the genetic structure of the *Prosopium* are necessary to validate the role of adaptive radiation in Bear Lake *Prosopium*. Bernatchez (1999) lists three main research objectives to comprehensively understand how populations evolve under adaptive radiation: (1) elucidating the evolutionary history of populations; (2) identifying the processes responsible for phenotypic, ecological, and genetic differentiation; and (3) identifying the mechanisms involved in the development and maintenance of reproductive isolation. Of those three objectives, this study has

begun elucidating the evolutionary history of the Bear Lake *Prosopium*, but more work is needed with more sensitive genetic markers like microsatellites and will require an integration of molecular and ecological approaches.

Summary

Mountain whitefish are a paraphyletic group comprised of two or three separate lineages. They are genetically isolated by major drainage basins. The Bear Lake *Prosopium* lies within the mountain whitefish clades rather than being a separate sister clade. The Bear Lake *Prosopium* appear to be polyphyletic as well and do not reflect the current species definitions of *Prosopium*.

Despite the apparent genetic paraphyly of mountain whitefish, the mountain whitefish is still a valid group biologically, ecologically and morphologically. As Todd and Smith (1992) note, monophyly isn't the only natural process; local seasonal races still have much to offer in aiding our understanding of a species. It may not be possible to resolve *Prosopium* using any one species criteria, but may instead require the examination of morphological, phylogenetic, historical, and ecological characteristics to gain a solid understanding of the group (Todd & Smith 1992). If anything, the mountain whitefish and Bear Lake endemics raise interesting questions regarding species definitions and Evolutionarily Significant Units (ESU's) and reaffirm the importance of the past history of an organism in understanding its current condition.

REFERENCES

- Anderson AL (1947) Drainage diversion in the northern Rocky Mountains of east-central Idaho. *Journal of Geology* **55**, 61-75.
- Behnke RJ (2002) *Trout and salmon of North America* The Free Press, New York, New York.
- Bernatchez L, Chouinard A, Lu G (1999) Integrating molecular genetics and ecology in studies of adaptive radiation: whitefish, *Coregonus* sp., as a case study. *Biological Journal of the Linnean Society* **68**, 173-194.
- Bernatchez L, Colombani F, Dodson JJ (1991) Phylogenetic relationships among the subfamily Coregoninae as revealed by mitochondrial DNA restriction analysis. *Journal of Fish Biology* **39**, 283-290.
- Bernatchez L, Dodson JJ (1991) Phylogeographic structure in mitochondrial DNA of the lake whitefish (*Coregonus clupeaformis*) and its relation to Pleistocene glaciations. *Evolution* **45**, 1016-1035.
- Bernatchez L, Vuorinen JA, Bodaly RA, Dodson JJ (1996) Genetic evidence for reproductive isolation and multiple origins of sympatric trophic ecotypes of whitefish (*Coregonus*). *Evolution* **50**, 624-635.
- Binz T, Largiader C, Muller R, Wedekind C (2001) Sequence diversity of *Mhc* genes in lake whitefish. *Journal of Fish Biology* **58**, 359-373.
- Bodaly RA, Clayton JW, Lindsey CC, Vuorinen J (1992) Evolution of lake whitefish (*Coregonus clupeaformis*) in North America during the Pleistocene: genetic differentiation between sympatric populations. *Canadian Journal of Fisheries and Aquatic Sciences* **49**, 769-779.

- Bouchard DP, Kaufman DS, Hochberg A, Quade J (1998) Quaternary history of the Thatcher Basin, Idaho, reconstructed from the $^{87}\text{Sr}/^{86}\text{Sr}$ and amino acid composition of lacustrine fossils: implications for the diversion of the Bear River into the Bonneville Basin. *Palaeogeography, Palaeoclimatology, Palaeoecology* **141**, 95-114.
- Broughton J (2000) Terminal Pleistocene fish remains from Homestead Cave, Utah, and implications for fish biogeography in the Bonneville Basin. *Copeia*, 645-656.
- Firby JR, Sharpe SE, Whelan JF, Smith GR, Spaulding WG (1997) Paleobiotic and isotopic analysis of mollusks, fish and plants from core OL-92: indicators of an open or closed lake system. In: *An 800,000-year old paleoclimatic record from core OL-92, Owens Lake, southeast California* (eds. Smith GI, Bischoff JL). Special Papers of the Geological Society of America **317**, 121-125.
- Foote CJ, Clayton JW, Lindsey CC, Bodaly RA (1992) Evolution of lake whitefish (*Coregonus clupeaformis*) in North America during the Pleistocene: evidence for a Nahanni glacial refuge race in the northern Cordillera region. *Canadian Journal of Fisheries and Aquatic Sciences* **49**, 760-768.
- Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic Biology* **52**, 696-704.
- Hansen WR (1985) Drainage development of the Green River basin in Southwestern Wyoming and its bearing on fish biogeography, neotectonics, and paleoclimate. *The Mountain Geologist* **22**, 192-204.
- Holt RD (1960) Comparative morphometry of the mountain whitefish. *Copeia*, 192-200.
- Howard AD (1958) Drainage evolution in northeastern Montana and northwestern North Dakota. *Bulletin of the Geological Society of America* **69**, 575-588.

- Hubbs CL, Miller RR (1948) The zoological evidence: correlation between fish distribution and hydrographic history in the desert basins of western United States. In: *The Great Basin, with emphasis on glacial and postglacial times*. Bulletin of the University of Utah **38**, 17-166.
- Hulsenbeck JP, Ronquist F, Nielsen R, Bollback JP (2001) Bayesian inference of phylogeny and its impact on evolutionary biology. *Science* **294**, 2310-2314.
- Johnson JB (2002) Evolution after the flood: Phylogeography of the desert fish Utah chub. *Evolution* **56**, 948-960.
- Laabs BJC, Kaufman DS (2003) Quaternary highstands in Bear Lake Valley, Utah and Idaho. *Geological Society of America Bulletin* **115**, 16.
- Link PK, Reading RW, Godfrey AE, Prevedel D (1999) Topographic development of the Snake River Plain. In: *Guidebook to the geology of eastern Idaho* eds. Hughes SS, Thackray GD), pp. page 343-344. Idaho Museum of Natural History, Pocatello, Idaho.
- Lu G, Basley DJ, Bernatchez L (2001) Contrasting patterns of mitochondrial DNA and microsatellite introgressive hybridization between lineages of lake whitefish (*Coregonus clupeaformis*); relevance for speciation. *Molecular Ecology* **10**, 965-985.
- Lu G, Bernatchez L (1999) Correlated trophic specialization and genetic divergence in sympatric lake whitefish ecotypes (*Coregonus clupeaformis*): support for the ecological speciation hypothesis. *Evolution* **53**, 1491-1505.
- Madsen DB, Rhode D, Grayson DK, *et al.* (2001) Late Quaternary environmental change in the Bonneville Basin, western USA. *Palaeogeography, Palaeoclimatology, Palaeoecology* **167**, 243-271.

- Malde HE (1991) Quaternary geology and structural history of the Snake River Plain, Idaho and Oregon. In: *Quaternary nonglacial geology: Conterminous U.S.* (ed. Morrison RB). The Geological Society of America **K2**, 251-281.
- Miller KM, Withler RE (1998) The salmonid class I *Mhc*: limited diversity in a primitive teleost. *Immunological Reviews* **166**, 279-293.
- Miller KM, Withler RE, Beacham TD (1997) Molecular evolution at *Mhc* genes in two populations of chinooks salmon *Oncorhynchus tshawytscha*. *Molecular Ecology* **6**, 937-954.
- Minckley WL, Hendrickson DA, Bond CE (1986) Geography of western North American freshwater fishes: description and relationships to intracontinental tectonism. In: *The zoogeography of North American freshwater fishes* (eds. Hocutt CH, Wiley EO). John Wiley and Sons, New York, New York.
- Morrison RB (1991) Quaternary stratigraphic, hydrologic, and climatic history of the Great Basin, with emphasis on Lakes Lahontan, Bonneville, and Tecopa. In: *Quaternary nonglacial geology: Conterminous U.S.* (ed. Morrison RB). The Geological Society of America **K2**, 283-320.
- Norden CR (1970) Evolution and distribution of the genus *Prosopium*. In: *Biology of coregonid fishes* (eds. Lindsey CC, Woods CS), pp. 67-80. University of Manitoba Press, Winnipeg, Canada.
- Oviatt CG, Currey DR (1987) Pre-Bonneville quaternary lakes in the Bonneville Basin, Utah. In: *Cenozoic Geology of Western Utah-Sites for Precious Metal and Hydrocarbon Accumulations* (eds. Kopp RS, Cohenour RE). Utah Geological Association **16**, 257-263.

- Oviatt CG, Currey DR, Sack D (1992) Radiocarbon chronology of Lake Bonneville, Eastern Great Basin, USA. *Palaeogeography, Palaeoclimatology, Palaeoecology* **99**, 225-241.
- Perkins ME, Nash BP (2002) Explosive silicic volcanism of the Yellowstone hotspot: the ash fall tuff record. *Geological Society of America Bulletin* **114**, 367-381.
- Pfenninger M, Posada D (2002) Phylogeographic history of the land snail *Candidula unifasciata* (Helicellinae, Stylommatophora): fragmentation, corridor migration, and secondary contact. *Evolution* **56**, 1776-1788.
- Pigeon DA, Chouinard A, Bernatchez L (1997) Multiple modes of speciation involved in the parallel evolution of sympatric morphotypes of lake whitefish (*Coregonus clupeaformis*, Salmonidae). *Evolution* **51**, 196-205.
- Posada D, Crandall KA (1998) MODELTEST: testing the model of DNA substitution. *Bioinformatics* **14**, 817-818.
- Posada D, Crandall KA, Templeton AR (2000) Geodis: a program for the cladistic nested analysis of the geographical distribution of genetic haplotypes. *Molecular Ecology* **9**, 487-488.
- Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19**, 1572-1574.
- Schluter D (2001) Ecology and the origin of species. *Trends in Ecology and Evolution* **16**, 372-380.
- Schneider S, Roessli D, Excoffier L (2000) Arlequin version 2.000: a software for population genetic data analysis. Genetics and Biometry Laboratory, University of Geneva, Geneva, Switzerland.

- Shiozawa DK, Rader RB (2005) Great Basin Rivers. In: *Rivers of North America* (eds. Behnke AC, Cushing C), pp. 655-694. Academic Press.
- Sigler WF (1962) Bear Lake and its future. In: *Twenty-Sixth Faculty Honor Lecture*, pp. 1-23. The Faculty Association Utah State University, Logan.
- Sigler WF, Sigler JW (1987) *Fishes of the Great Basin: a natural history* University of Nevada Press, Reno, Nevada.
- Smith GR (1975) Fishes of the Pliocene Glens Ferry Formation, southwest Idaho. *University of Michigan Papers on Paleontology* **5**, 1-68.
- Smith GR (1981) Late Cenozoic freshwater fishes of North America. *Annual Review of Ecology and Systematics* **12**, 163-193.
- Smith GR, Morgan N, Gustafson E (2000) Fishes of the Mio-Pliocene Ringold Formation, Washington: Pliocene capture of the Snake River by the Columbia River. *University of Michigan Papers on Paleontology* **32**, 1-47.
- Smith GR, Stokes WL, Horn KF (1968) Some late Pleistocene fishes of Lake Bonneville. *Copeia*, 807-816.
- Smith GR, Swirydczuk K, Kimmel PG, Wilkinson BH (1982) Fish biostratigraphy of late Miocene to Pleistocene sediments of the western Snake River Plain, Idaho. In: *Cenozoic Geology of Idaho* (eds. Bonnicksen B, Breckenridge RM). Idaho Bureau of Mines and Geology Bulletin.
- Stokes WL (1979) Paleohydrographic history of the Great Basin. In: *Basin and Range Symposium and Great Basin Field Conference* (eds. Newmann GW, Goode HD), pp. 345-351. Rocky Mountain Association of Geologists, Denver, Colorado.

- Swofford DL (2001) PAUP: phylogenetic analysis using parsimony (and other methods), version 4.0b7. Sinauer Associates, Sunderland, Massachusetts.
- Taylor DW, Bright RC (1987) Drainage history of the Bonneville Basin. In: *Geology of Western Utah-Sites for Precious Metal and Hydrocarbon Accumulations* (eds. Kopp RS, Cohenour RE). Utah Geological Association **16**, 239-256.
- Taylor DW, Smith GR (1981) Pliocene molluscs and fishes from northeastern California and northwestern Nevada. *Contributions from the Museum of Paleontology: The University of Michigan* **25**, 339-413.
- Templeton AR (1998) Nested clade analyses of phylogeographic data: testing hypotheses about gene flow and population history. *Molecular Ecology* **7**, 381-397.
- Templeton AR (2004) Statistical phylogeography: methods of evaluating and minimizing inference errors. *Molecular Ecology* **13**, 789-809.
- Templeton AR, Routman E, Phillips CA (1995) Separating population structure from population history: a cladistic analysis of the geographical distribution of mitochondrial DNA haplotypes in the tiger salamander, *Ambystoma tigrinum*. *Genetics* **140**, 767-782.
- Todd TN, Smith GR (1992) A review of differentiation in Great Lake ciscoes. *Polish Archives of Hydrobiology* **39**, 261-267.
- Tolentino SA, Thompson BW (2004) Meristic differences, habitat selectivity, and diet separation of *Prosopium spilonotus* and *P. abyssicola*. *Annales Zoologici Fennici* **41**, 309-317.
- Turgeon J, Bernatchez L (2001a) Clinal variation at microsatellite loci reveals historical secondary intergradation between glacial races of *Coregonus artedii* (Teleostei: Coregoninae). *Evolution* **55**, 2274-2286.

- Turgeon J, Bernatchez L (2001b) Mitochondrial DNA phylogeography of lake cisco (*Coregonus artedii*): evidence supporting extensive secondary contacts between two glacial races. *Molecular Ecology* **10**, 987-1001.
- Turgeon J, Bernatchez L (2003) Reticulate evolution and phenotypic diversity in North American ciscoes, *Coregonus* ssp. (Teleostei: Salmonidae): implications for the conservation of an evolutionary legacy. *Conservation Genetics* **4**, 67-81.
- Turgeon J, Estoup A, Bernatchez L (1999) Species flock in the North American Great Lakes: Molecular ecology of Lake Nipigon ciscoes (Teleostei: Coregonidae: *Coregonus*). *Evolution* **53**, 1857-1871.
- Van Tassell J, Ferns M, McConnell V, Smith GR (2001) The mid-Pliocene Imbler fish fossils, Grand Ronde Valley, Union County, Oregon, and the connection between Lake Idaho and the Columbia River. *Oregon Geology* **63**, 77-96.
- Vourinen JA (1998) Phylogeny of five *Prosopium* species with comparisons with other Coregonine fishes based on isozyme electrophoresis. *Journal of Fish Biology* **53**, 917-927.
- White RG (1974) *Endemic whitefishes of Bear Lake, Utah-Idaho: a problem in systematics*. PhD thesis, Utah State University.
- Whiteley AR, Spruell P, Allendorf FW (in press) Can common species provide valuable information for conservation? *Molecular Ecology*.

FIGURE LEGENDS

Figure 1: Distribution of mountain whitefish (*Prosopium williamsoni*) based on drainage basin within the United States. The Columbia Basin is shown in red, the Missouri River Basin in yellow, the Lahontan Basin in orange, the Bonneville Basin in blue, and the Green River Basin in green.

Figure 2: Map of possible routes (white dotted lines) of *Prosopium* into the Snake River and Lake Idaho ~5 million years ago. *Prosopium* from the Hudson Bay entered the region via the Missouri River. *Prosopium* were forced to move to the east or west of the Idaho Batholith to enter the Snake River and Lake Idaho. Lake Idaho may have drained westward through Oregon and California into the Pacific Ocean.

Figure 3: Map of the capture of the Snake River through Hells Canyon into the Columbia River during the late Pliocene ~2 million years ago. Possible *Prosopium* movements into the Columbia River are indicated by the white dotted lines.

Figure 4: Map of Lake Lahontan and Lake Bonneville during the Quaternary.

Figure 5: Map of possible connections between Snake River drainages and the Humboldt River ~3 million years ago. Possible *Prosopium* movements are indicated by the white dotted lines.

Figure 6: Map of possible connections between the Missouri River Basin, the upper Snake River Sub-basin, the Green River Basin, and the Bonneville Basin ~5 million years ago. The Bear

River was fragmented at this time: the northern portion flowed into the upper Snake River, and the eastern portion was connected to the Green River. Possible *Prosopium* movements are indicated by the white dotted lines.

Figure 7: Map of the capture of the Bear River and the Green River ~2 million years ago and possible connections between the Green River Basin and the Bonneville Basin and upper Snake River Basin.

Figure 8. Map of Pleistocene events influencing the Bonneville Basin. Lava flows (white line with red waves) blocked the flow of the Bear River to the Portneuf River in Thatcher Basin. As a result, Lake Thatcher formed in the basin. Lake Thatcher overflowed at its southern rim, cutting the Oneida Narrows, and entering Lake Bonneville.

Figure 9: Map of the Pleistocene Bonneville Flood. The increase in water input from Lake Thatcher caused Lake Bonneville to expand. Lake Bonneville overtopped its northern rim at Red Rocks Pass, cutting a path to the Snake River.

Figure 10: Map of sampling locations. Numbers correspond to map locations given in table 1. The red colored circles represent the Columbia River Sub-basin, grey the lower Snake River Sub-basin, purple the upper Snake River Sub-basin, yellow the Missouri River Basin, blue the Bonneville Basin, green the Green River Basin, and orange the Lahontan Basin.

Figure 11: Bayesian cytb phylogeny. Haplotypes are colored based on basin/sub-basin. The Columbia River Sub-basin is shown in red, the lower Snake River Sub-basin in grey, the upper Snake River Sub-basin in purple, the Missouri River Basin in yellow, the Lahontan Basin in orange, the Bonneville Basin in blue, and the Green River Basin in green. Posterior probabilities for each node are given.

Figure 12: Maximum likelihood cytb phylogeny. Haplotypes are colored based on basin/sub-basin. The Columbia River Sub-basin is shown in red, the lower Snake River Sub-basin in grey, the upper Snake River Sub-basin in purple, the Missouri River Basin in yellow, the Lahontan Basin in orange, the Bonneville Basin in blue, and the Green River Basin in green. Bootstrap nodal support is given.

Figure 13: Bayesian ND2 phylogeny. Haplotypes are colored based on basin/sub-basin. The Columbia River Sub-basin is shown in red, the lower Snake River Sub-basin in grey, the upper Snake River Sub-basin in purple, the Missouri River Basin in yellow, the Lahontan Basin in orange, the Bonneville Basin in blue, and the Green River Basin in green. Posterior probabilities for each node are given.

Figure 14: Maximum likelihood ND2 phylogeny. Haplotypes are colored based on basin/sub-basin. The Columbia River Sub-basin is shown in red, the lower Snake River Sub-basin in grey, the upper Snake River Sub-basin in purple, the Missouri River Basin in yellow, the Lahontan Basin in orange, the Bonneville Basin in blue, and the Green River Basin in green. Bootstrap nodal support is given.

Figure 15. TCS haplotypes network for *cytb*. Haplotypes are colored based on basin/sub-basin. The Columbia River Sub-basin is shown in red, the lower Snake River Sub-basin in grey, the upper Snake River Sub-basin in purple, the Missouri River Basin in yellow, the Lahontan Basin in orange, the Bonneville Basin in blue, and the Green River Basin in green. 11(a) (i-viii) shows the network for the Bonneville Basin/upper Snake River Sub-basin/Green River Basin/Bear Lake *Prosopium* group from zero to seven steps. 11(b) (i-viii) shows the network for a second Bear Lake *Prosopium* group from zero to seven steps. 11(c) (i-viii) shows the network for the Columbia River Sub-basin/lower Snake River Sub-basin/Lahontan Basin group from zero to seven steps. 11(d) (i-viii) shows the network for the Missouri River Basin group from zero to seven steps.

Figure 16. TCS haplotypes network for ND2. Haplotypes are colored based on basin/sub-basin. The Columbia River Sub-basin is shown in red, the lower Snake River Sub-basin in grey, the upper Snake River Sub-basin in purple, the Missouri River Basin in yellow, the Lahontan Basin in orange, the Bonneville Basin in blue, and the Green River Basin in green. 11(a) (i-viii) shows the network for the Bonneville Basin/upper Snake River Sub-basin/Green River Basin group from zero to seven steps. 11(b) (i-viii) shows the network for the Bear Lake *Prosopium* group from zero to seven steps. 11(c) (i-viii) shows the network for the Missouri River Basin group from zero to seven steps. 11(d) (i-viii) shows the network for the Columbia River Sub-basin/lower Snake River Sub-basin/Lahontan Basin group from zero to seven steps.

Table 1: A list of the samples used in this study. Location is listed by its sub basin and drainage. The map ID corresponds to figure 6. The number of sequences used for cytochrome b (cytb) and NADH dehydrogenase subunit 2 (ND2) are also given.

Species	Drainage	Sub-basin	Location	CYTB N	ND2 N	Map ID		
<i>P. williamsoni</i>	Bonneville	Bonneville	Bear R, UT	31	31	2		
			Logan R, UT	25	25	3		
			Weber R, UT	28	35	4		
			Provo R, UT	16	34	5		
			Thomas Fk, WY	26	27	6		
			Woodruff Reservoir, UT	30	31	7		
			Columbia River	Columbia River	Crooked R, OR	36	35	8
	North Santiam R, OR	3			3	9		
	South Santiam R, OR	4			4	10		
	Willamette R, OR	30			27	11		
	Middle Fk, Willamette R, OR	17			16	12		
	Deschutes R, OR	5			5	13		
	Warm Springs R, OR (Deschutes)	30			28	14		
	Yakima R, WA	32			32	15		
	Chiwawa R, WA	21			21	16		
	Hoh R, WA	30			30	17		
	Skokomish R, WA	30			30	18		
	Warm Springs Cr, ID	3			0	19		
	North Fk Coeur d' Alene R, ID	3			4	20		
	Rock Cr, MT	5			3	21		
	Mainstem Flathead R, MT	24			24	22		
	Lower Snake River	Lower Snake River			Queens R, ID	1	0	33
					Payette R, ID	17	17	34
					Lochsa R, ID	19	20	35
					S Fk Boise R, ID	20	32	36
					Salmon R, ID	6	4	37
					S Fk Salmon R, ID	27	22	38
					Big Wood R, ID	19	19	39
			Malhuer R, OR	16	10	40		
	Upper Snake River	Upper Snake River	Lower Big Lost R, ID	23	21	52		
			Upper Big Lost R, ID	14	13	53		
			S Fk Snake R, ID	29	30	54		
			Henry's Fk, ID	31	31	55		
Conant Cr, ID			20	20	56			
Salt R, WY			32	31	57			
Hoback R, WY			28	27	58			
Jackson Lake, WY			46	49	59			
Snake R, WY	30	31	60					

Table 1 continued.

	Green	Green River	Duchesne R, UT	3	3	23
			Blacks Fk, UT	6	6	24
			Green R, UT	30	29	25
			W Fk, Blacks Fk, UT	24	24	26
			Green R, WY	10	10	27
			Yampa R, CO	39	40	28
			White R, CO	26	27	61
			Lahontan	Lahontan	East Walker R, NV	11
	Carson R, NV	13			13	30
	East Carson R, NV	4			4	31
	Truckee R, NV	64			63	32
	Missouri	Missouri River	Madison R, MT	0	30	41
			Gallatin R, MT	0	9	42
			Big Hole R, MT	25	21	43
			Yellowstone R, MT	17	17	44
			S Fk Judith R, MT	20	20	45
			Big Spring Cr, MT (Judith)	20	20	46
			Emerald Cr, Wind R, WY	1	1	47
			Sheep Cr, Wind R, WY	3	3	48
			Burwell Cr, Wind R, WY	13	13	49
			Tongue R, WY	13	13	50
			N Fk Shoshone R, WY	10	10	51
<i>P. abyssicola</i>	Bonneville	Bonneville	Bear Lake, ID-UT	50	51	1
<i>P. gemmifer</i>	Bonneville	Bonneville	Bear Lake, ID-UT	55	50	1
<i>P. spilonotus</i>	Bonneville	Bonneville	Bear Lake, ID-UT	70	61	1
Total				1334	1371	

Table 2: A list of primer sequences used for cytochrome b and NADH dehydrogenase subunit 2 gene amplification in *Prosopium*.

Gene		Primer	Sequence
Cytochrome B	First half	1425	5' GACTTGAAAAACCACCGTTG 3'
		CYTB-intR	5' AAGAGGGCCAGGGATGTTAATCCT 3'
	Second half	CYTB-intF	5' CCGCAACAGTYGTTACCTCCTTT 3'
		1426	5' TTTAGAATCTTAGCTTTGGGAG 3'
NADH dehydrogenase subunit 2	First half	BYU11	5' TAAGCTTTCGGGCCCATACCC 3'
		ND2-intR	5' TAGTTCAGGAGGTTGCAAGAGCGT 3'
	Second half	ND2-int-F	5' AAGCTTGCTCCTTTCGCGCTTATG 3'
		BYU12	5' GGCTCAGGCACCAAATACTA 3'

Table 3: A list of cytb haplotypes. The number and locations of individuals in each haplotype are given.

Haps	Location											Total											
	Bonneville			Columbia						Lower Snake			Upper Snake		Green		Lahontan		Missouri				
B1	Bear R, UT	14																					24
B2	Logan R, UT	3																					5
B3	Weber R, UT	1																					5
B4	Provo R, UT			4																			4
B5	Thomas Fk, WY				3																		3
B6	Woodruff Res, UT					4																	2
B7	Bear Lake, ID-UT																						2
B8	Crooked R, OR																						2
B9	N Santiam R, OR																						1
B10	S. Santiam R, OR																						1
BUS	Williamette R, OR																						
C1	M Fk, Williamette R, OR																						
C2	Deschutes R, OR																						
C3	Warm Springs R, OR																						
C4	Yakima R, WA																						
C5	Chiwawa R, WA																						
C6	Hoh R, WA																						
C7	Skokomish R, WA																						
C8	Warm Springs Cr, ID																						
C9	N Fk Coeur d'Alene R, ID																						
C10	Rock R, MT																						
C11	Flathead R, MT																						
C12	Queens R, ID																						
C13	Payette R, ID																						
C14	Lochsa R, ID																						
C15	S Fk Boise R, ID																						
C16	Salmon R, ID																						
C17	S Fk Salmon R, ID																						
C18	Big Wood R, ID																						
C19	Malluer R, OR																						
C20	Lower Big Lost R, ID																						
C21	Upper Big Lost R, ID																						
C22	S Fk Snake R, ID																						
C23	Henry's Fk, ID																						
C24	Conant Cr, ID																						
C25	Salt R, WY																						
C26	Hoback R, WY																						
C27	Jackson Lake, WY																						
C28	Snake R, WY																						
C29	Duchesne R, UT																						
C30	Blacks Fk, UT																						
	Green R, UT																						
	WFK Blacks Fk, UT																						
	Green R, WY																						
	Yampa R, CO																						
	White R, CO																						
	E Walker R, NV																						
	Carson R, NV																						
	E Carson R, NV																						
	Truckee R, NV																						
	Madison R, MT																						
	Gallatin R, MT																						
	Big Hole R, MT																						
	Yellowstone R, MT																						
	S Fk Judith R, MT																						
	Big Spring Cr, MT																						
	Emerald Cr, Wind R, WY																						
	Sheep Cr, Wind R, WY																						
	Burwell Cr, Wind R, WY																						
	Tongue R, WY																						
	N Fk Shoshone R, WY																						

Table 7: Cytb pairwise Fst values by basin/sub basin or species. All Fst values are statistically significant ($p < 0.05$). Fst values under 0.30 are highlighted and indicate gene flow.

Cytb Population Pairwise Fst's										
	Bonneville	Columbia	Green	Lahontan	Lower Snake	Missouri	Upper Snake	<i>P. abyssicola</i>	<i>P. gemmifer</i>	<i>P. spilonotus</i>
Bonneville	0.000									
Columbia	0.836	0.000								
Green	0.913	0.824	0.000							
Lahontan	0.956	0.381	0.958	0.000						
Lower Snake	0.851	0.145	0.840	0.391	0.000					
Missouri	0.916	0.802	0.915	0.896	0.795	0.000				
Upper Snake	0.102	0.799	0.623	0.873	0.801	0.846	0.000			
<i>P. abyssicola</i>	0.824	0.661	0.828	0.769	0.609	0.837	0.709	0.000		
<i>P. gemmifer</i>	0.820	0.773	0.833	0.891	0.752	0.855	0.613	0.551	0.000	
<i>P. spilonotus</i>	0.810	0.672	0.811	0.766	0.624	0.835	0.712	0.085	0.509	0.000

Table 8: Cytb effective migration values per generation by basin/sub basin or species. Effective migration values over 1.0 are highlighted.

Cytb M Values										
	Bonneville	Columbia	Green	Lahontan	Lower Snake	Missouri	Upper Snake	<i>P. abyssicola</i>	<i>P. gemmifer</i>	<i>P. spilonotus</i>
Bonneville	0.000									
Columbia	0.098	0.000								
Green	0.047	0.106	0.000							
Lahontan	0.023	0.811	0.022	0.000						
Lower Snake	0.088	2.946	0.095	0.779	0.000					
Missouri	0.046	0.123	0.046	0.058	0.129	0.000				
Upper Snake	4.400	0.126	0.303	0.073	0.124	0.091	0.000			
<i>P. abyssicola</i>	0.107	0.256	0.104	0.150	0.320	0.097	0.205	0.000		
<i>P. gemmifer</i>	0.110	0.147	0.100	0.061	0.164	0.084	0.315	0.408	0.000	
<i>P. spilonotus</i>	0.117	0.244	0.117	0.153	0.301	0.098	0.202	5.399	0.482	0.000

Table 9: Cytb AMOVA table.

Cytb AMOVA				
Source of Variation	d.f.	Sum of Squares	Variance of Components	Percentage of Variation
Among Groups	9	13266.32	11.066 Va	75.81
Among Populations within Groups	42	1164.10	1.126 Vb	7.72
Within Populations	1282	3082.19	2.404 Vc	16.47
Total	1333	17512.60	14.597	

Table 12: ND2 pairwise Fst values by basin/sub basin or species. All Fst values are statistically significant ($p < 0.05$). Fst values under 0.30 are highlighted and indicate gene flow.

ND2 Population Pairwise Fst's										
	Bonneville	Columbia	Green	Lahontan	Lower Snake	Missouri	Upper Snake	<i>P. abyssicola</i>	<i>P. gemmifer</i>	<i>P. spilonotus</i>
Bonneville	0.000									
Columbia	0.745	0.000								
Green	0.561	0.741	0.000							
Lahontan	0.933	0.211	0.979	0.000						
Lower Snake	0.750	0.045	0.748	0.168	0.000					
Missouri	0.825	0.727	0.844	0.871	0.711	0.000				
Upper Snake	0.109	0.751	0.446	0.913	0.760	0.820	0.000			
<i>P. abyssicola</i>	0.486	0.584	0.457	0.569	0.477	0.557	0.532	0.000		
<i>P. gemmifer</i>	0.621	0.665	0.675	0.860	0.619	0.727	0.595	0.166	0.000	
<i>P. spilonotus</i>	0.564	0.626	0.572	0.750	0.553	0.665	0.574	0.065	0.073	0.000

Table 13: ND2 effective migration values per generation by basin/sub basin or species. Effective migration values over 1.0 are highlighted.

ND2 M Values										
	Bonneville	Columbia	Green	Lahontan	Lower Snake	Missouri	Upper Snake	<i>P. abyssicola</i>	<i>P. gemmifer</i>	<i>P. spilonotus</i>
Bonneville	0.000									
Columbia	0.174	0.000								
Green	0.391	0.175	0.000							
Lahontan	0.036	1.865	0.011	0.000						
Lower Snake	0.167	10.492	0.168	2.469	0.000					
Missouri	0.106	0.188	0.093	0.074	0.203	0.000				
Upper Snake	4.096	0.166	0.621	0.048	0.158	0.110	0.000			
<i>P. abyssicola</i>	0.528	0.356	0.593	0.378	0.547	0.398	0.440	0.000		
<i>P. gemmifer</i>	0.305	0.252	0.241	0.081	0.307	0.187	0.340	2.510	0.000	
<i>P. spilonotus</i>	0.386	0.298	0.374	0.167	0.404	0.251	0.371	7.166	6.386	0.000

Table 14: ND2 AMOVA table.

ND2 AMOVA				
Source of Variation	d.f.	Sum of Squares	Variance of Components	Percentage of Variation
Among Groups	9	18279.33	14.658 Va	62.45
Among Populations within Groups	43	2343.73	2.0696 Vb	8.82
Within Populations	1318	8888.48	6.744 Vc	28.73
Total	1370	29511.54	23.472	

Table 15: Cytb nested clade analysis of the haplotypes network for clades with a significant X^2 only with average clade distance (D_c) and nested clade distance (D_n) for each haplotype/clade and interior-tip (I-T) distances for each clade. Distances that were significantly large or small based on the permutations test are labeled (L) and/or (S).

Clade	X^2 Probability	Haplotype	Topology	D_c	D_n
1-1	0.00	LS2	Interior	26.49S	57.82S
		LS9	Tip	0.00	102.32L
		LS12	Tip	0.00	51.02
		LS36	Tip	0.00	142.86
		LS39	Tip	0.00	142.86
		I-T	26.49	-32.08S	
1-7	0.00	C12	Interior	31.38S	94.79S
		C53	Tip	0.00	103.79
		LS11	Tip	0.00S	331.82L
		I-T	31.38	-191.42S	
1-20	0.00	C1	Interior	169.15S	264.05S
		C9	Interior	76.03S	330.99L
		C45	Tip	0.00	380.16
		LS18	Tip	0.00	261.45
		LS38	Interior	0.00	323.91
		I-T	150.18	-24.40	
1-43	0.00	C3	Interior	0.00S	15.75S
		C7	Tip	0.00S	40.09L
		C57	Tip	0.00	15.75
		I-T	0.00	-22.31S	
1-50	0.00	L1	Tip	0.00S	70.98L
		L3	Interior	46.64	48.94
		L5	Tip	0.00S	38.36S
		L6	Tip	0.00	70.98
		L8	Tip	0.00	48.66
		L10	Tip	0.00	38.36
		L11	Tip	0.00	48.66
		I-T	46.64L	-14.30S	
1-58	0.01	C5	Interior	36.65S	67.43S
		C18	Tip	0.00	112.91L
		I-T	36.65	-45.47S	
1-62	0.04	LS7	Interior	0.00S	55.65
		LS16	Tip	26.24	23.51
		I-T	-26.24	32.14	
1-64	0.00	LS4	Interior	11.60S	44.10
		LS21	Tip	0.00	34.05S
		LS25	Tip	0.00	45.00

Table 15 continued.

		LS28	Tip	0.00	131.48
		LS40	Tip	0.00	91.03
			I-T	11.60	-23.02
1-68	0.02	C14	Interior	79.99	146.35
		LS5	Interior	114.87	156.39
			I-T	N/A	N/A
1-123	0.01	M4	Interior	99.16	85.55
		M9	Tip	0.00	99.54
			I-T	99.16	-13.98
1-127	0.00	M1	Interior	34.29S	34.86S
		M20	Interior	0.00	248.97L
		M26	Tip	0.00	133.82
			I-T	33.18S	-92.05
1-147	0.00	G1	Interior	85.01	84.45
		G2	Tip	0.00S	38.30S
		G3	Tip	0.00S	38.30S
		G4	Tip	84.32	109.84
		G5	Tip	0.00	136.85
		G7	Tip	0.00	98.02
		G9	Tip	0.00	101.53
		G10	Tip	0.00	21.15
		G11	Tip	0.00	21.15
		G12	Tip	0.00	21.15
		G13	Tip	0.00	92.22
		G14	Tip	0.00	136.85
		US10	Tip	30.44	317.73L
			I-T	71.42L	-13.96
1-166	0.00	US6	Tip	7.79S	20.67S
		US7	Interior	38.95	50.71L
		US12	Tip	22.62	28.17
			I-T	25.46L	27.15L
1-173	0.00	BUS	Interior	91.85	92.31
		B2	Tip	44.09	51.46
		B5	Tip	0.00S	38.05
		B6	Tip	44.33	55.33
		B8	Tip	0.00	78.62
		B9	Tip	0.00	78.62
		B10	Tip	0.00	34.47
		US30	Tip	0.00	178.60
		US32	Tip	0.00	235.91L
		US45	Tip	0.00	221.40L
		PG26	Tip	0.00	18.89S
			I-T	74.67L	14.13

Table 15 continued.

2-1	0.00	1-1	Tip	71.53S	112.36S
		1-2	Interior	128.13	196.33L
		1-3	Tip	0.00	270.27
			I-T	59.86L	76.79L
2-3	0.00	1-6	Interior	0.00S	186.31
		1-7	Tip	145.12	257.88
		1-8	Interior	210.39	210.56
			I-T	43.69	-49.81
2-7	0.00	1-14	Interior	0.00	0.00SL
		1-15	Tip	51.91	296.13
		1-16	Tip	0.00	347.33
		1-18	Tip	0.00	306.00
		1-19	Tip	0.00	210.76S
		1-20	Interior	285.08	289.68
		1-21	Tip	0.00	130.27S
			I-T	267.78	0.00SL
2-10	0.00	1-26	Interior	0.00	152.47
		1-29	Interior	266.00	255.51
		1-30	Interior	0.00	0.00SL
			I-T	N/A	N/A
2-17	0.00	1-43	Interior	22.61	22.61
		1-44	Interior	0.00	0.00SL
			I-T	N/A	N/A
2-18	0.00	1-34	Interior	0.00	0.00SL
		1-40	Interior	0.00	269.92
		1-45	Interior	0.00	0.00SL
		1-46	Interior	0.00	96.82S
		1-47	Tip	0.00S	471.19L
		1-48	Interior	84.12S	172.81S
	I-T	71.6457	0.00SL		
2-20	0.00	1-52	Tip	28.96	28.96
		1-53	Interior	0.00	0.00SL
			I-T	0.00SL	0.00SL
2-21	0.00	1-54	Interior	0.00S	96.34L
		1-55	Interior	30.81S	33.17S
			I-T	N/A	N/A
2-22	0.00	1-51	Interior	0.00	0.00SL
		1-56	Tip	10.41	18.56
		1-57	Tip	0.00	137.54L
		1-58	Tip	71.62	79.40

Table 15 continued.

		1-59	Interior	46.35	44.61
			I-T	-15.06	0.00SL
2-24	0.04	1-62	Tip	30.64	34.50
		1-63	Interior	0.00	73.10
			I-T	-30.64	38.60
2-38	0.00	1-93	Interior	0.00	0.00SL
		1-154	Interior	31.46	31.46
			I-T	N/A	N/A
2-50	0.00	1-121	Tip	129.13S	137.18S
		1-122	Interior	0.00	306.06L
			I-T	-129.13	168.87L
2-51	0.02	1-123	Tip	89.59S	122.06
		1-124	Interior	0.00S	132.74
			I-T	-89.59S	10.68
2-52	0.00	1-125	Tip	145.30L	153.21L
		1-126	Tip	0.00	220.63L
		1-127	Interior	43.86S	50.06S
			I-T	-46.96	-128.43S
2-53	0.02	1-128	Tip	0.00	32.17
		1-129	Tip	0.00	38.29
		1-130	Interior	48.27	45.56
		1-131	Interior	0.00	38.29
			I-T	46.60L	9.06
2-61	0.00	1-146	Tip	0.00S	132.88
		1-147	Interior	85.17	86.25L
		1-148	Interior	0.00	62.82
			I-T	83.96L	-46.97
2-64	0.00	1-156	Tip	64.64	64.64
		1-157	Interior	0.00	0.00SL
			I-T	0.00SL	0.00SL
2-65	0.00	1-155	Interior	0.00	0.00SL
		1-158	Interior	0.00	49.58S
		1-159	Tip	63.60	72.64
		1-160	Tip	33.13S	56.00S
		1-161	Interior	87.36	94.78L
		1-162	Tip	0.00	56.73
			I-T	17.69	0.00SL
2-66	0.00	1-163	Interior	0.00	0.00SL
		1-64	Tip	62.53	62.53

Table 15 continued.

			I-T	0.00SL	0.00SL
2-68	0.00	1-167	Interior	53.58S	120.39
		1-168	Tip	21.97S	84.61S
		1-169	Tip	61.25S	135.55L
		1-170	Tip	48.48S	150.97L
		1-171	Tip	54.93S	113.33
		1-172	Tip	48.29	168.09
		1-173	Interior	90.83S	110.63S
				I-T	39.98L
3-1	0.00	2-1	Tip	142.01S	246.37
		2-2	Tip	250.02	257.05
		2-3	Interior	220.36S	257.86
			I-T	66.98L	10.37
3-3	0.00	2-4	Tip	174.22	182.25S
		2-7	Interior	288.24L	289.97L
		2-8	Interior	139.82S	233.43S
			I-T	89.90	98.53L
3-6	0.00	2-14	Tip	0.00S	103.03S
		2-15	Interior	0.00	279.56L
			I-T	0.00	176.53L
3-8	0.00	2-9	Interior	0.00	0.00SL
		2-18	Interior	245.12	245.12
			I-T	N/A	N/A
3-10	0.00	2-19	Tip	50.99S	240.52S
		2-22	Interior	54.19S	431.99L
			I-T	3.21	191.46L
3-15	0.00	2-33	Interior	194.80S	231.00S
		2-34	Tip	0.00	443.03L
		2-35	Interior	0.00	0.00SL
			I-T	194.80S	0.00SL
3-22	0.00	2-50	Tip	152.64S	162.85
		2-51	Interior	125.43S	201.23L
			I-T	-27.21	38.38L
3-23	0.00	2-52	Interior	72.32S	222.57L
		2-53	Interior	41.94S	103.87S
			I-T	N/A	N/A
3-27	0.00	2-61	Tip	84.82S	84.77S
		2-62	Interior	41.46	378.86L
			I-T	-43.36	294.08L

Table 15 continued.

3-28	0.00	2-38	Interior	31.46S	80.31S
		2-60	Interior	0.00	0.00SL
		2-63	Interior	0.00S	344.15L
		2-64	Tip	64.64	71.02
		2-65	Interior	74.89S	102.39
		2-66	Tip	62.53	98.22
			I-T	-14.25	0.00SL
3-29	0.00	2-67	Tip	31.83S	154.36L
		2-68	Interior	117.53S	118.39S
			I-T	85.70L	-35.96S
4-1	0.000	3-1	Tip	253.61S	254.69S
		3-2	Tip	33.98S	480.58L
		3-3	Interior	269.93	282.42
			I-T	56.99L	-14.10
4-2	0.01	3-4	Interior	212.77L	209.87L
		3-5	Tip	48.19S	46.86S
			I-T	164.58L	163.01L
4-3	0.00	3-6	Tip	150.56S	450.35L
		3-7	Tip	21.92S	439.21L
		3-8	Interior	245.12S	233.78S
			I-T	173.28L	-209.76S
4-4	0.00	3-9	Tip	42.01S	317.38
		3-10	Interior	308.14S	339.68L
			I-T	266.13L	22.30
4-5	0.00	3-11	Interior	118.48	118.48
		3-12	Interior	0.00	0.00SL
			I-T	N/A	N/A
4-7	0.00	3-15	Interior	273.26	273.26
		3-16	Interior	0.00	0.00SL
			I-T	N/A	N/A
4-10	0.00	3-22	Tip	174.48L	184.21L
		3-23	Interior	140.79S	146.71S
			I-T	-33.69S	-37.50S
4-12	0.00	3-26	Interior	0.00S	401.67L
		3-27	Tip	90.17S	208.12L
		3-28	Interior	106.54S	222.48L
		3-29	Tip	120.98S	131.38S
			I-T	-8.36	72.97L

Table 15 continued.

5-1	0.00	4-1	Tip	281.40S	283.70S
		4-2	Tip	75.87S	225.53S
		4-3	Interior	365.76L	390.40L
			I-T	102.79L	111.91L
5-2	0.00	4-4	Interior	333.99S	375.38
		4-5	Interior	118.48S	363.66
			I-T	N/A	N/A
5-5	0.00	4-10	Tip	157.47S	279.04L
		4-11	Interior	0.00	0.00SL
		4-12	Interior	172.57S	192.07S
			I-T	15.10	0.00SL
6-1	0.00	5-1	Tip	308.80S	333.86S
		5-2	Interior	371.45	396.42L
			I-T	62.65L	62.56L
6-2	0.00	5-3	Interior	144.07S	173.32S
		5-4	Tip	0.00S	122.63S
		5-5	Interior	216.81	216.72
			I-T	207.99L	88.82L
Total Cladogram	0.00	6-1	Tip	359.94S	474.95L
		6-2	Tip	214.46S	370.59S
			I-T	N/A	N/A

Table 16: GeoDis inferences for cytb. Results are shown only for clades with a significant X^2 probability.

Clade	Inference Chain	Inference
1-1	2-11-12-13-Yes	Long distance colonization possibly coupled with subsequent fragmentation or past fragmentation followed by range expansion
1-7	2-11-12-13-Yes	Long distance colonization possibly coupled with subsequent fragmentation or past fragmentation followed by range expansion
1-20	2-11-17-4-No	Restricted gene flow with isolation by distance
1-43	2-11-12-No	Contiguous range expansion
1-50	2-3-5-6-7-Yes	Restricted gene flow/Dispersal but with some long distance dispersal
1-58	2-11-12-No	Contiguous range expansion
1-62	2-11-12-No	Contiguous range expansion
1-64	2-11-12-No	Contiguous range expansion
1-68	19-20-2	Inconclusive outcome
1-123	2-11-17-No	Inconclusive outcome
1-127	2-11-12-No	Contiguous range expansion
1-147	2-3-5-6-13-Yes	Long distance colonization possibly coupled with subsequent fragmentation/Past fragmentation followed by range expansion
1-166	2-3-4-No	Restricted gene flow with isolation by distance
1-173	2-3-5-6-7-Yes	Restricted gene flow/Dispersal but with some long distance dispersal
2-1	2-3-4-No	Restricted gene flow with isolation by distance
2-3	2-11-17-No	Inconclusive outcome
2-7	2-11-17-No	Inconclusive outcome
2-10	2	Inconclusive outcome
2-17	19-20-2	Inconclusive outcome
2-18	2-3-5-6-7-Yes	Restricted gene flow/Dispersal but with some long distance dispersal
2-20	19-20-2	Inconclusive outcome
2-21	2	Inconclusive outcome
2-22	2-11-17	Inconclusive outcome
2-24	2-11-17	Inconclusive outcome
2-38	2	Inconclusive outcome
2-50	19-20-2-3-5-6-7-Yes	Restricted gene flow/Dispersal but with some long distance dispersal
2-51	2-11-12-No	Contiguous range expansion
2-52	2-11-12-No	Contiguous range expansion
2-53	2-3-4-No	Restricted gene flow with isolation by distance
2-61	2-3-4-No	Restricted gene flow with isolation by distance
2-64	2	Inconclusive outcome
2-65	2-3-4-No	Restricted gene flow with isolation by distance
2-66	2	Inconclusive outcome
2-68	2-3-5-6-7-Yes	Restricted gene flow/Dispersal but with some long distance dispersal
3-1	2-3-4-No	Restricted gene flow with isolation by distance
3-3	2-3-5-6-7-Yes	Restricted gene flow/Dispersal but with some long distance dispersal
3-6	19-20-2-3-4-No	Restricted gene flow with isolation by distance
3-8	2	Inconclusive outcome
3-10	19-20-2-11-12-13-Yes	Long distance colonization possibly coupled with subsequent fragmentation or past fragmentation followed by range expansion
3-15	19-20-2-11-12-No	Contiguous range expansion

Table 16 continued.

3-22	2-11-12-13- Yes	Long distance colonization possibly coupled with subsequent fragmentation or past fragmentation followed by range expansion
3-27	2-3-4-No	Restricted gene flow with isolation by distance
3-28	2-3-5-6-7-Yes	Restricted gene flow/Dispersal but with some long distance dispersal
3-29	2-3-5-6-7-Yes	Restricted gene flow/Dispersal but with some long distance dispersal
4-1	2-3-5-6-13- Yes	Long distance colonization possibly coupled with subsequent fragmentation or past fragmentation followed by range expansion
4-2	2-3-4-No	Restricted gene flow with isolation by distance
4-3	2-3-5-6-13- Yes	Long distance colonization possibly coupled with subsequent fragmentation or past fragmentation followed by range expansion
4-4	2-3-5-6-13-14- No	Past fragmentation and/or long distance colonization
4-5	2	Inconclusive outcome
4-7	2	Inconclusive outcome
4-10	2-11-12-No	Contiguous range expansion
4-12	2-11-12-13- Yes	Long distance colonization possibly coupled with subsequent fragmentation or past fragmentation followed by range expansion
5-1	2-3-4-No	Restricted gene flow with isolation by distance
5-2	2	Inconclusive outcome
5-5	2-3-4-No	Restricted gene flow with isolation by distance
6-1	2-3-4-No	Restricted gene flow with isolation by distance
6-2	2-3-5-6-7-8- Yes	Restricted gene flow/Dispersal but with some long distance dispersal over intermediate areas not occupied by the species; or past gene flow followed by extinction of intermediate populations
Total	2	Inconclusive outcome

Table 17: ND2 nested clade analysis of the haplotypes network for clades with a significant X^2 only with average clade distance (D_c) and nested clade distance (D_n) for each haplotype/clade and interior-tip (I-T) distances for each clade. Distances that were significantly large or small based on the permutations test are labeled (L) and/or (S).

Clade	X^2 Probability	Haplotype	Topology	D_c	D_n
1-20	0.04	CLS2	Interior	218.89L	215.59
		C26	Tip	0.00	48.97S
		C48	Tip	0.00	157.71
			I-T	218.89L	130.37L
1-21	0.01	LS3	Interior	0.00S	11.34S
		LS14	Tip	0.00	181.77L
		LS39	Tip	0.00	11.36
			I-T	0.00	-113.61S
1-73	0.04	C10	Tip	0.00S	149.92S
		C23	Interior	0.00	227.03L
			I-T	0.00	77.11L
1-147	0.00	M1	Tip	43.51S	81.72S
		M5	Interior	17.06S	96.13
		M7	Tip	113.62L	131.22L
		M11	Tip	0.00	102.58
		M22	Tip	0.00	93.43
		M24	Tip	0.00	102.58
		M28	Tip	0.00	182.23
			I-T	-31.67S	2.09
1-154	0.02	M2	Interior	155.82	156.49
		M10	Tip	0.00S	188.27L
		M16	Tip	0.00	188.27
		M17	Tip	0.00	162.39
		M23	Interior	0.00	185.89
		M25	Tip	0.00	185.89
		M26	Tip	0.00	185.89
		M30	Tip	0.00	91.15
		M31	Tip	0.00	91.15
		M33	Tip	0.00	91.15
		M34	Tip	0.00	91.15
			I-T	149.83L	3.58
1-179	0.00	BUS2	Interior	30.97S	47.65S
		B7	Tip	40.38	48.89
		US9	Tip	41.49	188.75L
			I-T	-9.97	-71.17S
1-189	0.00	G1	Interior	78.90	82.21L
		G4	Tip	0.00S	99.18
		G5	Tip	0.00S	57.46

Table 17 continued.

		G6	Tip	0.00S	99.18
		G8	Tip	0.00	96.57
		G11	Tip	0.00	31.34
			I-T	78.90L	-0.67
1-193	0.00	BUS3	Interior	166.03L	168.59L
		B4	Tip	0.00S	25.36S
		US26	Tip	0.00	195.81
			I-T	166.03L	118.88L
1-203	0.00	BUS1	Interior	118.69	115.91
		USPS	Tip	78.20	121.39
		B6	Tip	0.00S	179.23L
		B11	Tip	0.00	282.48L
		B12	Tip	0.00	179.23
		B15	Tip	0.00	45.12
		B17	Tip	0.00	131.29
		US1	Tip	31.51S	119.95
		US7	Tip	2.98S	76.670S
		US14	Tip	0.00	129.76
		US15	Tip	0.00	71.50
		US20	Tip	68.64	88.81
		US23	Tip	52.81	80.93
		US24	Tip	15.42	56.08S
		US25	Tip	59.90	73.64
		US28	Tip	0.00	40.66
		US32	Tip	0.00	77.80
		US33	Tip	0.00	129.76
		US34	Tip	0.00	129.76
		US35	Tip	0.00	129.76
		US36	Tip	0.00	129.76
		US38	Interior	0.00	71.50
		US39	Interior	0.00	71.50
		US40	Tip	0.00	71.50
		US41	Tip	0.00	71.50
		US42	Tip	0.00	105.55
		US43	Tip	0.00	105.55
		US46	Tip	0.00	105.55
		US51	Tip	0.00	121.20
			I-T	97.35L	-0.51
2-2	0.01	1-3	Interior	0.00	11.23S
		1-4	Tip	0.00	22.61
		1-5	Tip	5.07	21.68
		1-13	Interior	42.56	40.13
			I-T	31.66L	13.41
2-4	0.00	1-8	Interior	80.96S	92.75S
		1-9	Tip	0.00	307.96L

Table 17 continued.

			I-T	80.96	-215.21S
2-6	0.00	1-12	Interior	0.00	0.00SL
		1-14	Tip	0.00	285.96
		1-17	Interior	0.00	17.70
			I-T	0.00	0.00SL
2-10	0.00	1-20	Interior	196.89S	262.53
		1-21	Tip	21.38S	237.92
		1-27	Interior	226.81S	243.33
		1-28	Interior	0.00S	216.31
		1-29	Tip	117.01S	306.88L
			I-T	147.88L	-13.89
2-11	0.00	1-22	Tip	131.20	131.20
		1-23	Interior	0.00	0.00SL
			I-T	0.00SL	0.00SL
2-15	0.00	1-31	Tip	168.37	168.37
		1-32	Interior	0.00	0.00SL
			I-T	0.00SL	0.00SL
2-27	0.03	1-62	Tip	0.00S	79.60
		1-63	Interior	32.40	24.35
			I-T	32.40	-55.25
2-29	0.02	1-67	Interior	0.00S	65.97
		1-68	Tip	119.76	105.12
			I-T	-119.76	-39.15
2-32	0.00	1-73	Tip	180.76	180.76
		1-74	Interior	0.00	0.00SL
			I-T	0.00SL	0.00SL
2-35	0.00	1-79	Tip	0.00S	67.44L
		1-80	Tip	0.00S	67.44L
		1-81	Tip	36.49	48.26
		1-82	Interior	45.33	47.42S
			I-T	38.76L	-16.57S
2-61	0.00	1-146	Tip	0.00S	170.90L
		1-147	Interior	93.26S	94.58S
			I-T	93.26L	-76.32S
2-62	0.00	1-149	Interior	0.00	0.00SL
		1-150	Tip	0.00S	200.19
		1-151	Tip	0.00	31.34
		1-152	Tip	0.00	200.19
		1-153	Tip	37.32S	138.98

Table 17 continued.

		1-154	Interior	156.95	150.41
			I-T	130.83L	0.00SL
2-73	0.00	1-176	Tip	0.00S	93.29
		1-177	Tip	8.001S	21.94S
		1-178	Tip	73.12	69.36
		1-179	Interior	65.11	61.38
			I-T	12.80	-1.98
2-76	0.00	1-182	Interior	0.00	0.00SL
		1-185	Interior	0.00	0.00SL
		1-186	Tip	0.00S	86.04
		1-187	Tip	0.00	35.09S
		1-188	Tip	50.83	125.26L
		1-189	Interior	78.58	78.90
			I-T	57.78L	0.00SL
2-79	0.00	1-192	Tip	0.00	130.64
		1-193	Tip	129.03	197.74L
		1-196	Tip	0.00	107.00
		1-197	Tip	29.13	105.40
		1-198	Tip	68.64	95.61
		1-199	Tip	48.70	136.64
		1-200	Tip	0.00	51.82
		1-201	Tip	0.00	82.64
		1-202	Tip	0.00S	36.50S
		1-203	Interior	114.53	117.95
			I-T	50.60L	-4.84
3-1	0.00	2-1	Tip	0.00S	9.43S
		2-2	Interior	29.70L	30.24L
			I-T	29.70L	20.82L
3-2	0.00	2-3	Tip	98.24	129.28
		2-4	Interior	122.19	180.28
			I-T	23.95	51.00
3-3	0.00	2-5	Tip	85.79S	291.15
		2-6	Interior	33.48	145.15S
		2-7	Tip	0.00S	475.95L
			I-T	-2.92	-252.41S
3-4	0.00	2-8	Interior	165.83	279.64
		2-9	Tip	34.65S	208.92S
		2-10	Interior	247.96	248.83
			I-T	208.99L	41.52L
3-9	0.00	2-19	Tip	0.00S	14.98S
		2-20	Tip	0.00	14.98

Table 17 continued.

		2-21	Interior	0.00S	40.85L
			I-T	0.00	25.88L
3-11	0.00	2-22	Interior	0.00	0.00SL
		2-25	Interior	47.22	45.99
		2-26	Interior	0.00	78.42
		2-27	Tip	32.69	43.42
			I-T	11.75	0.00SL
3-13	0.01	2-30	Interior	27.29S	32.19L
		2-31	Interior	0.00S	29.63S
			I-T	N/A	N/A
3-15	0.00	2-35	Interior	51.25S	55.06S
		2-36	Interior	0.00S	45.53
		2-37	Tip	0.00	627.92L
			I-T	38.86	-575.16S
3-16	0.00	2-32	Interior	180.76L	186.73L
		2-38	Interior	0.00S	31.30S
		2-39	Interior	44.41	49.32
			I-T	N/A	N/A
3-17	0.00	2-40	Interior	114.64	114.64
		2-41	Interior	0.00	0.00SL
			I-T	N/A	N/A
3-26	0.00	2-59	Tip	53.90S	79.46S
		2-60	Interior	2.10S	137.32
		2-61	Interior	99.35S	181.00L
			I-T	-88.16S	-23.37S
3-27	0.02	2-62	Interior	149.54	146.14
		2-63	Interior	0.00S	189.42
			I-T	N/A	N/A
3-29	0.00	2-67	Tip	34.77	34.77
		2-68	Interior	0.00	0.00SL
			I-T	0.00SL	0.00SL
3-31	0.00	2-72	Tip	44.19	177.58L
		2-73	Interior	63.22S	63.39S
			I-T	19.04	-114.19S
3-32	0.00	2-74	Tip	41.49	379.18L
		2-75	Tip	29.43S	316.91L
		2-76	Interior	80.17S	79.90S
			I-T	45.37	-264.68S

Table 17 continued.

3-33	0.00	2-71	Interior	0.00S	131.78
		2-77	Interior	0.00	184.80
		2-78	Tip	57.83S	89.65S
		2-79	Interior	117.71	115.20
			I-T	52.36L	26.79L
4-1	0.00	3-1	Tip	26.90S	297.24
		3-2	Tip	152.09S	281.36S
		3-3	Interior	310.51	343.70L
			I-T	247.84L	50.99L
4-4	0.00	3-6	Interior	0.00	174.06
		3-9	Tip	21.92S	264.65L
		3-10	Tip	0.00S	21.49S
		3-11	Interior	46.04S	114.75S
			I-T	30.37	-57.74S
4-5	0.01	3-12	Tip	104.35	107.00L
		3-13	Interior	31.39S	51.83S
			I-T	-72.95S	-55.18S
4-6	0.00	3-14	Tip	0.00S	431.62L
		3-15	Interior	73.14S	144.18S
			I-T	73.14	-287.44S
4-7	0.00	3-16	Interior	111.72S	164.68S
		3-17	Interior	114.64	276.32L
			I-T	N/A	N/A
4-12	0.01	3-26	Tip	145.26	148.48
		3-27	Interior	148.47	162.05L
		3-28	Interior	0.00	243.34
			I-T	0.90	14.85L
4-13	0.00	3-29	Tip	34.77	34.77
		3-30	Interior	0.00	0.00SL
			I-T	0.00SL	0.00SL
4-14	0.00	3-31	Tip	67.97S	100.57S
		3-32	Tip	90.09S	184.53L
		3-33	Interior	115.43S	166.67
			I-T	34.12L	15.39L
5-1	0.00	4-1	Tip	310.97L	313.49L
		4-2	Interior	244.38S	246.84S
			I-T	-66.59S	-66.65S
5-2	0.00	4-3	Tip	66.43S	470.54L
		4-4	Interior	138.26S	176.47S

Table 17 continued.

			I-T	71.84	-294.07S
5-3	0.00	4-5	Tip	76.91S	301.70S
		4-6	Tip	200.14S	309.38S
		4-7	Interior	196.10S	471.22L
			I-T	28.76L	163.88L
5-5	0.00	4-10	Interior	0.00S	359.53L
		4-11	Interior	0.00	0.00SL
		4-12	Tip	153.52S	154.36SL
			I-T	-153.52S	0.00SL
5-6	0.00	4-13	Tip	34.77S	277.05L
		4-14	Interior	162.26S	162.37S
			I-T	127.49L	-114.68S
6-1	0.00	5-1	Tip	280.44S	317.99S
		5-2	Interior	231.83S	349.99
		5-3	Interior	334.74S	424.27L
			I-T	17.40	79.65L
6-2	0.00	5-4	Tip	0.00S	137.81S
		5-5	Interior	167.36S	266.13L
		5-6	Tip	172.80S	195.51S
			I-T	6.04	74.45L
Total Cladogram	0.00	6-1	Tip	361.82S	504.37L
		6-2	Tip	219.40S	358.83S
			I-T	N/A	N/A

Table 18: GeoDis results and inferences for ND2. Results are shown only for clades with a significant X^2 probability.

Clade	X^2 Probability	Inference Chain	Inference
1-20	0.04	2-3-4-No	Restricted gene flow with isolation by distance
1-21	0.01	2-11-12-No	Contiguous range expansion
1-73	0.04	19-20-2-3-4-9-No	Allopatric fragmentation
1-147	0.00	2-11-12-No	Contiguous range expansion
1-154	0.02	2-3-5-6-7-Yes	Restricted gene flow/Dispersal but with some long distance dispersal
1-179	0.00	2-11-12-No	Contiguous range expansion
1-189	0.00	2-3-4-No	Restricted gene flow with isolation by distance
1-193	0.00	2-3-4-No	Restricted gene flow with isolation by distance
1-203	0.00	2-3-5-6-7-No	Restricted gene flow/Dispersal but with some long distance dispersal
2-2	0.01	2-3-5-6-7-8-No	Sampling design inadequate to discriminate between isolation by distance (short distance movements) versus long distance dispersal
2-4	0.00	19-20-No	Inadequate geographical sampling
2-6	0.00	19-20-No	Inadequate geographical sampling
2-10	0.00	2-3-5-6-7-8-No	Sampling design inadequate to discriminate between isolation by distance (short distance movements) versus long distance dispersal
2-11	0.00	19-20-2	Inconclusive outcome
2-15	0.00	19-20-2	Inconclusive outcome
2-27	0.03	2-3-4-No	Restricted gene flow with isolation by distance
2-29	0.02	2-11-12-No	Contiguous range expansion
2-32	0.00	19-20-2	Inconclusive outcome
2-35	0.00	2-3-5-6-7-8-Yes	Restricted gene flow/Dispersal but with some long distance dispersal over intermediate areas not occupied by the species; or past gene flow followed by extinction of intermediate populations
2-61	0.00	2-3-5-6-7-8-No	Sampling design inadequate to discriminate between isolation by distance (short distance movements) versus long distance dispersal
2-62	0.00	2-3-4-No	Restricted gene flow with isolation by distance
2-73	0.00	2-3-4-No	Restricted gene flow with isolation by distance
2-76	0.00	2-3-4-No	Restricted gene flow with isolation by distance
2-79	0.00	2-3-5-6-13-Yes	Long distance colonization possibly coupled with subsequent fragmentation or past fragmentation followed by range expansion
3-1	0.00	2-3-4-No	Restricted gene flow with isolation by distance
3-2	0.00	2	Inconclusive outcome
3-3	0.00	2-3-5-6-7-Yes	Restricted gene flow/Dispersal but with some long distance dispersal
3-4	0.00	2-3-4-No	Restricted gene flow with isolation by distance
3-9	0.00	2-11-12-13-14-Yes	Sampling design inadequate to discriminate between contiguous range expansion, long distance colonization, and past fragmentation
3-11	0.00	2	Inconclusive outcome
3-13	0.01	2	Inconclusive outcome
3-15	0.00	2-11-12-No	Contiguous range expansion

Table 18 continued.

3-16	0.00	2	Inconclusive outcome
3-17	0.00	2	Inconclusive outcome
3-26	0.00	2-11-12-13-Yes	Long distance colonization possibly coupled with subsequent fragmentation or past fragmentation followed by range expansion
3-27	0.02	2	Inconclusive outcome
3-29	0.00	2	Inconclusive outcome
3-31	0.00	2-11-12-No	Contiguous range expansion
3-32	0.00	2-11-12-13-Yes	Long distance colonization possibly coupled with subsequent fragmentation followed by range expansion
3-33	0.00	2-3-4-No	Restricted gene flow with isolation by distance
4-1	0.00	2-3-4-No	Restricted gene flow with isolation by distance
4-4	0.00	2-3-5-6-7-Yes	Restricted gene flow/Dispersal but with some long distance dispersal
4-5	0.01	2-11-12-No	Contiguous range expansion
4-6	0.00	2-11-12-13-Yes	Long distance colonization possibly coupled with subsequent fragmentation or past fragmentation followed by range expansion
4-7	0.00	2	Inconclusive outcome
4-12	0.01	2-11-17-4-No	Restricted gene flow with isolation by distance
4-13	0.00	2	Inconclusive outcome
4-14	0.00	2-3-5-6-13-Yes	Long distance colonization possibly coupled with subsequent fragmentation followed by range expansion
5-1	0.00	2-11-12-No	Contiguous range expansion
5-2	0.00	19-20-2-11-12-13-Yes	Long distance colonization possibly coupled with subsequent fragmentation followed by range expansion
5-3	0.00	2-3-5-6-13-Yes	Long distance colonization possibly coupled with subsequent fragmentation followed by range expansion
5-5	0.00	19-20-2-11-12-13-Yes	Long distance colonization possibly coupled with subsequent fragmentation followed by range expansion
5-6	0.00	19-20-2-3-5-6-13-Yes	Long distance colonization possibly coupled with subsequent fragmentation followed by range expansion
6-1	0.00	2-11-12-No	Contiguous range expansion
6-2	0.00	2-11-12-13-Yes	Long distance colonization possibly coupled with subsequent fragmentation followed by range expansion
Total	0.00	2	Inconclusive outcome

Figure 1.

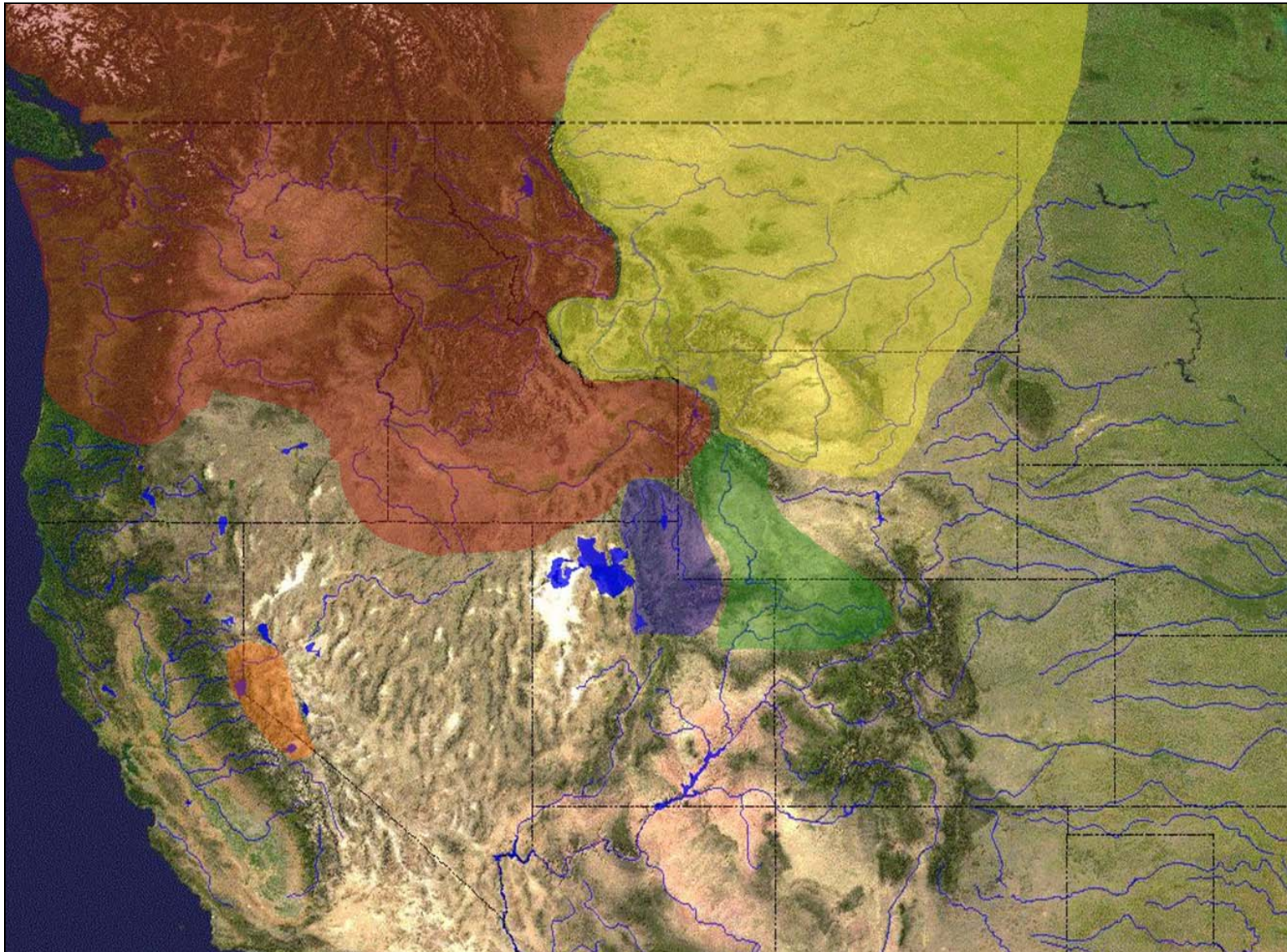


Figure 2.

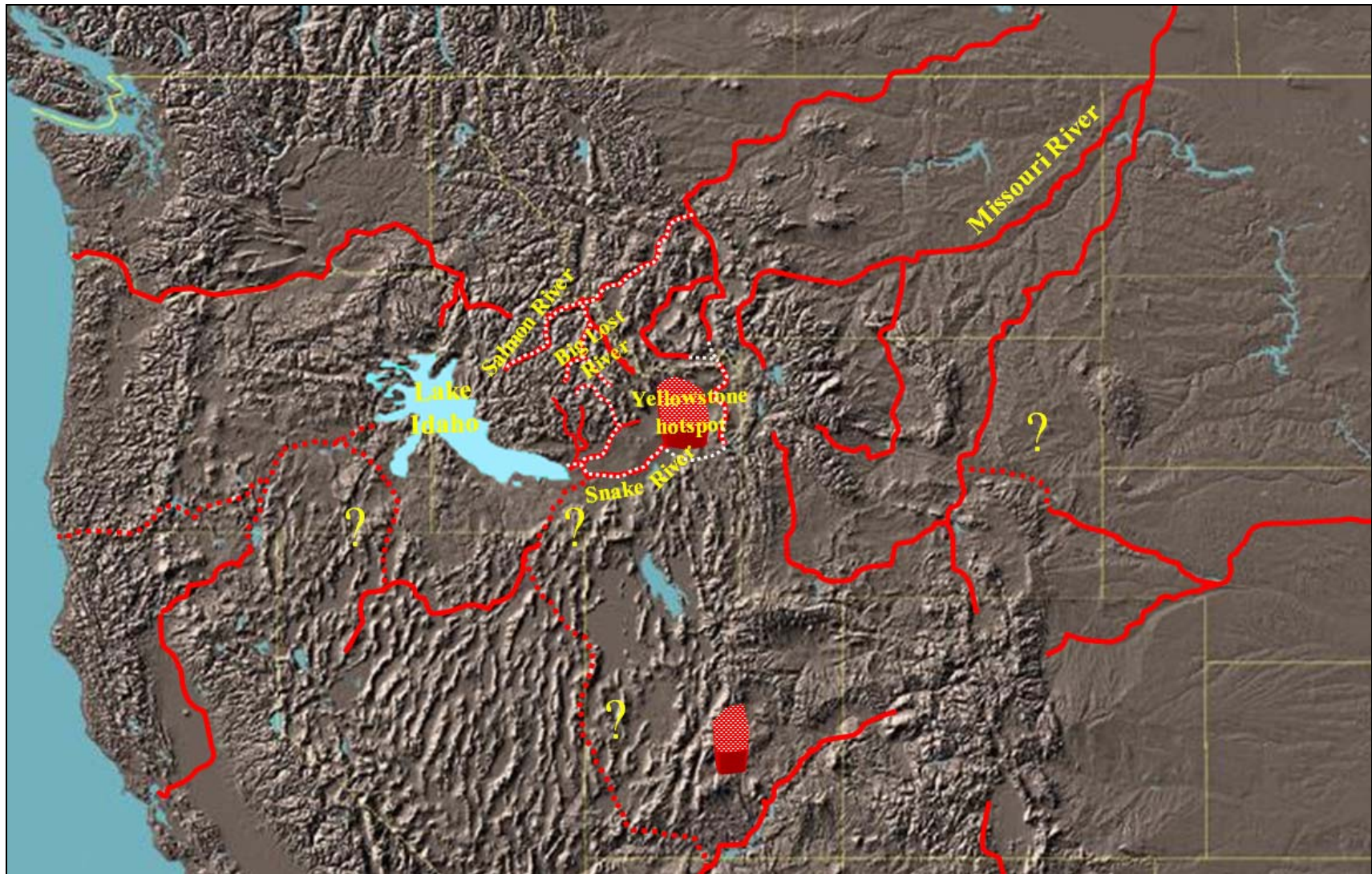


Figure 3.

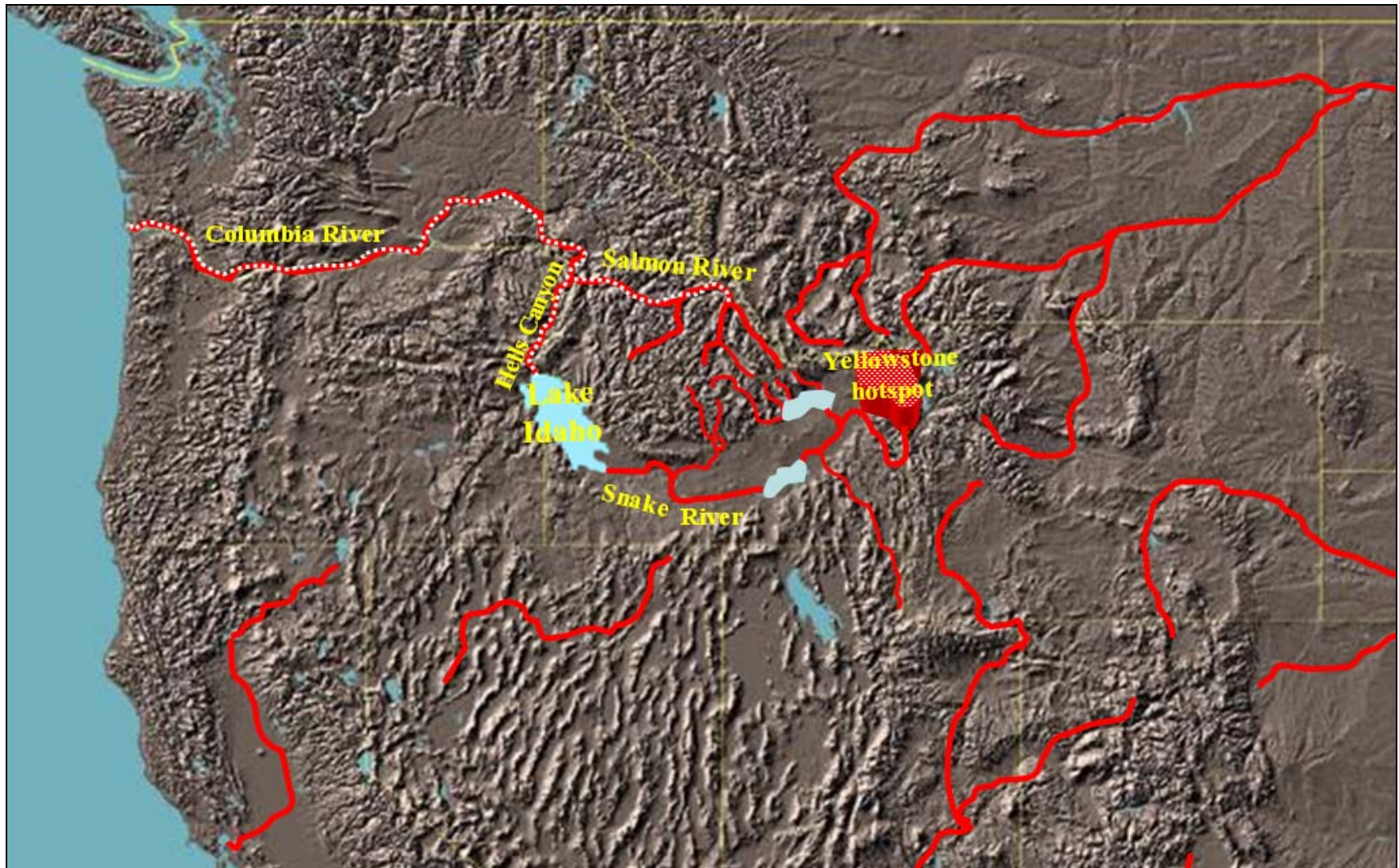


Figure 4.

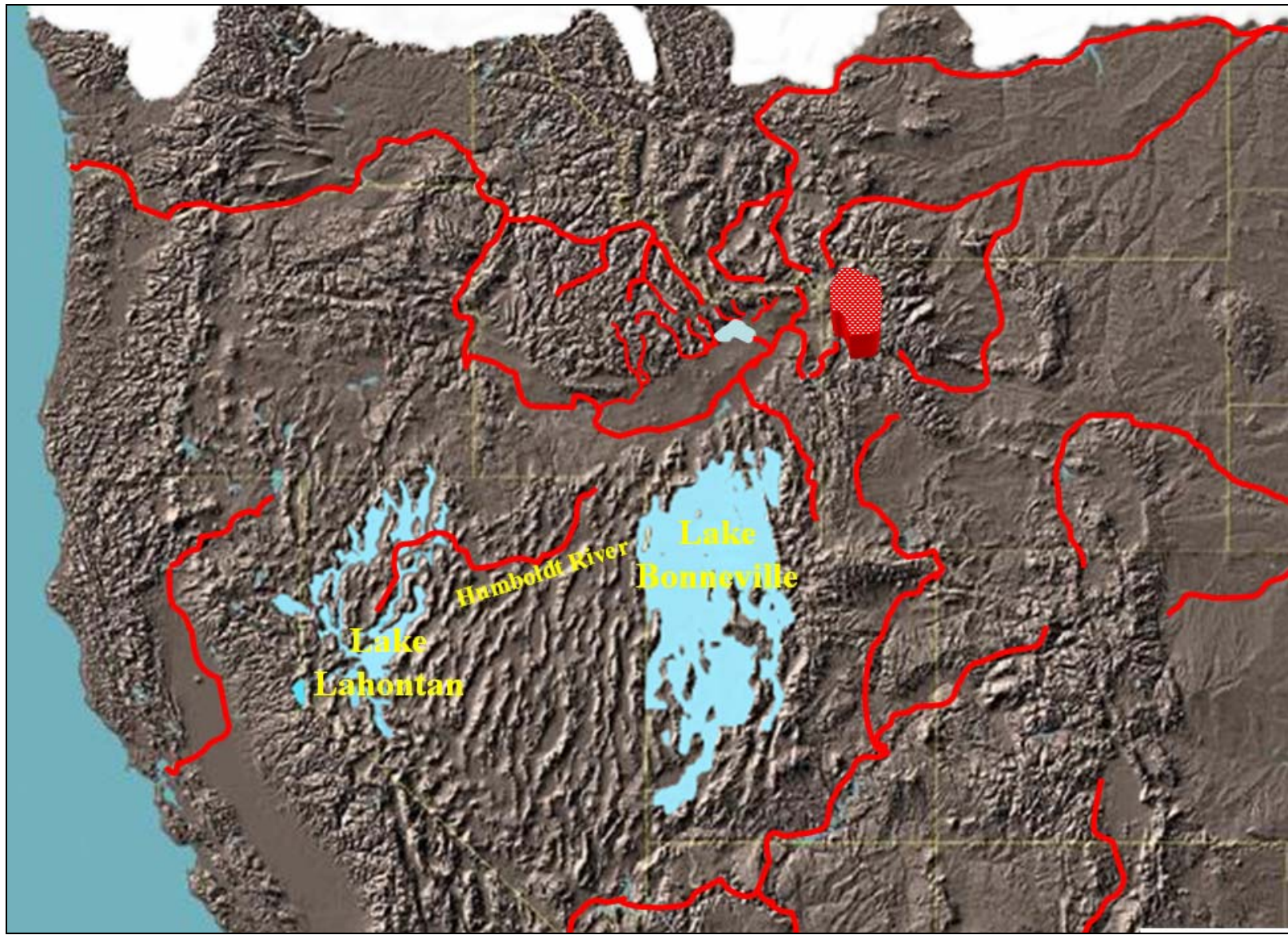


Figure 5.

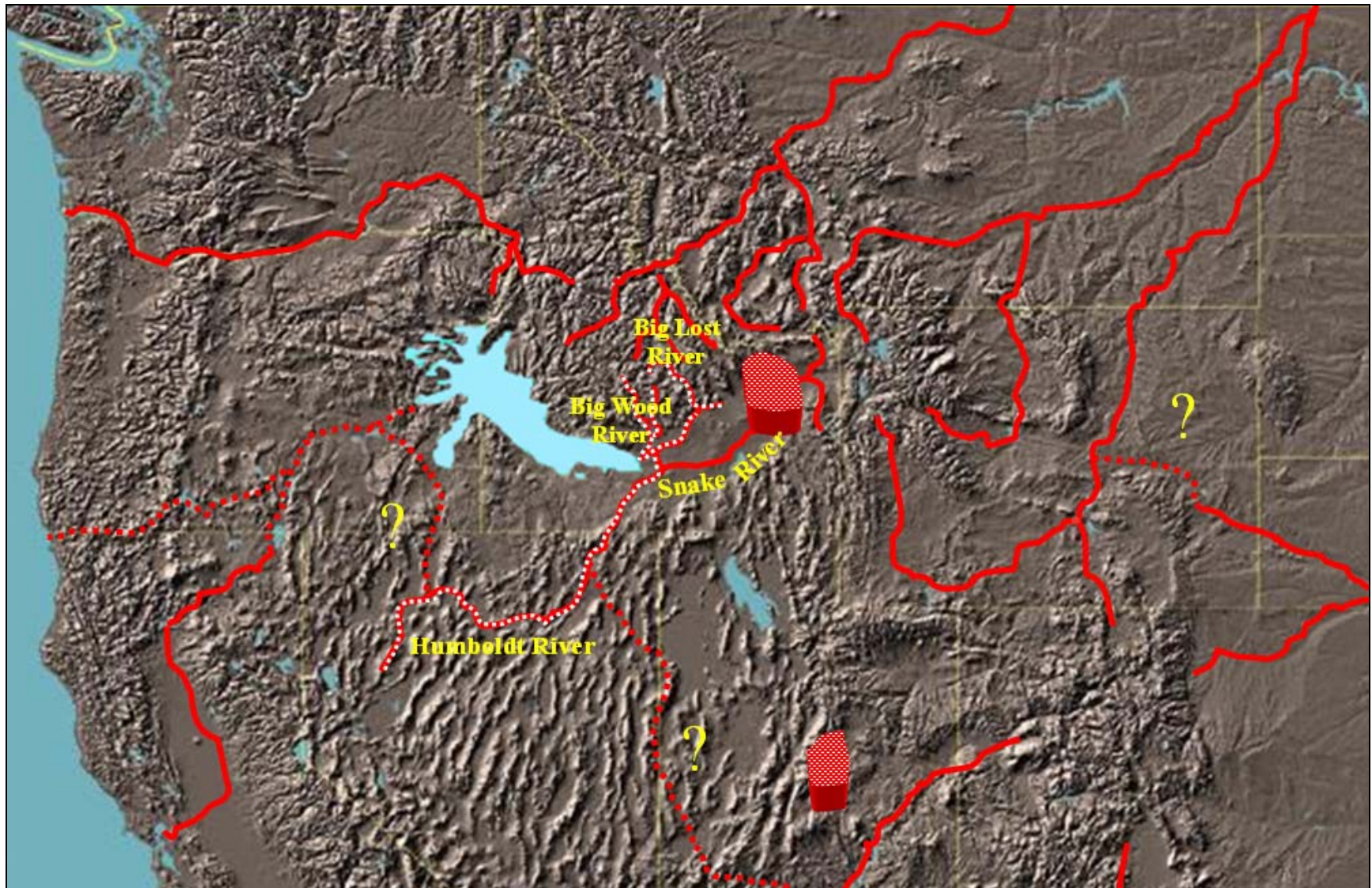


Figure 6.

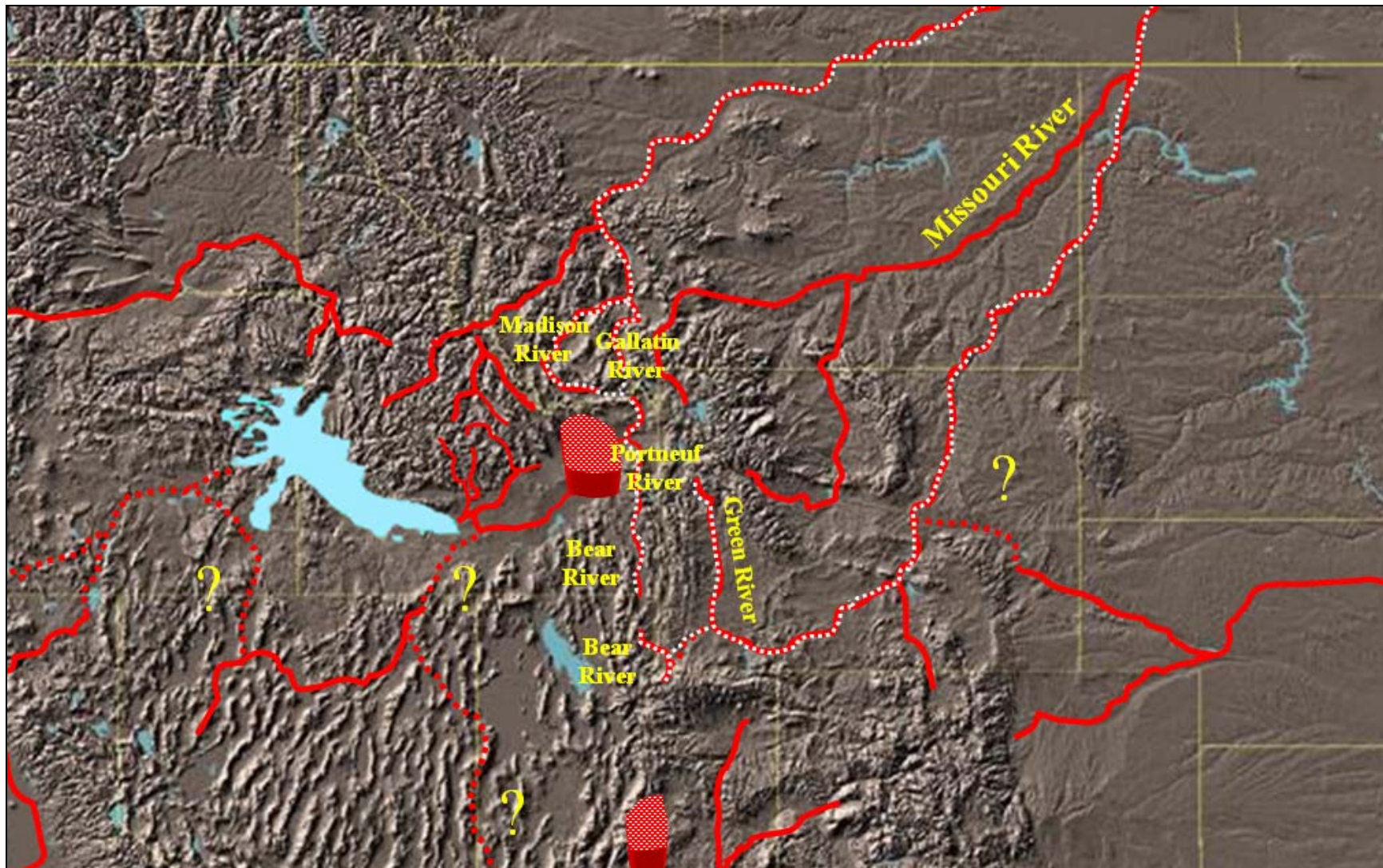


Figure 7.

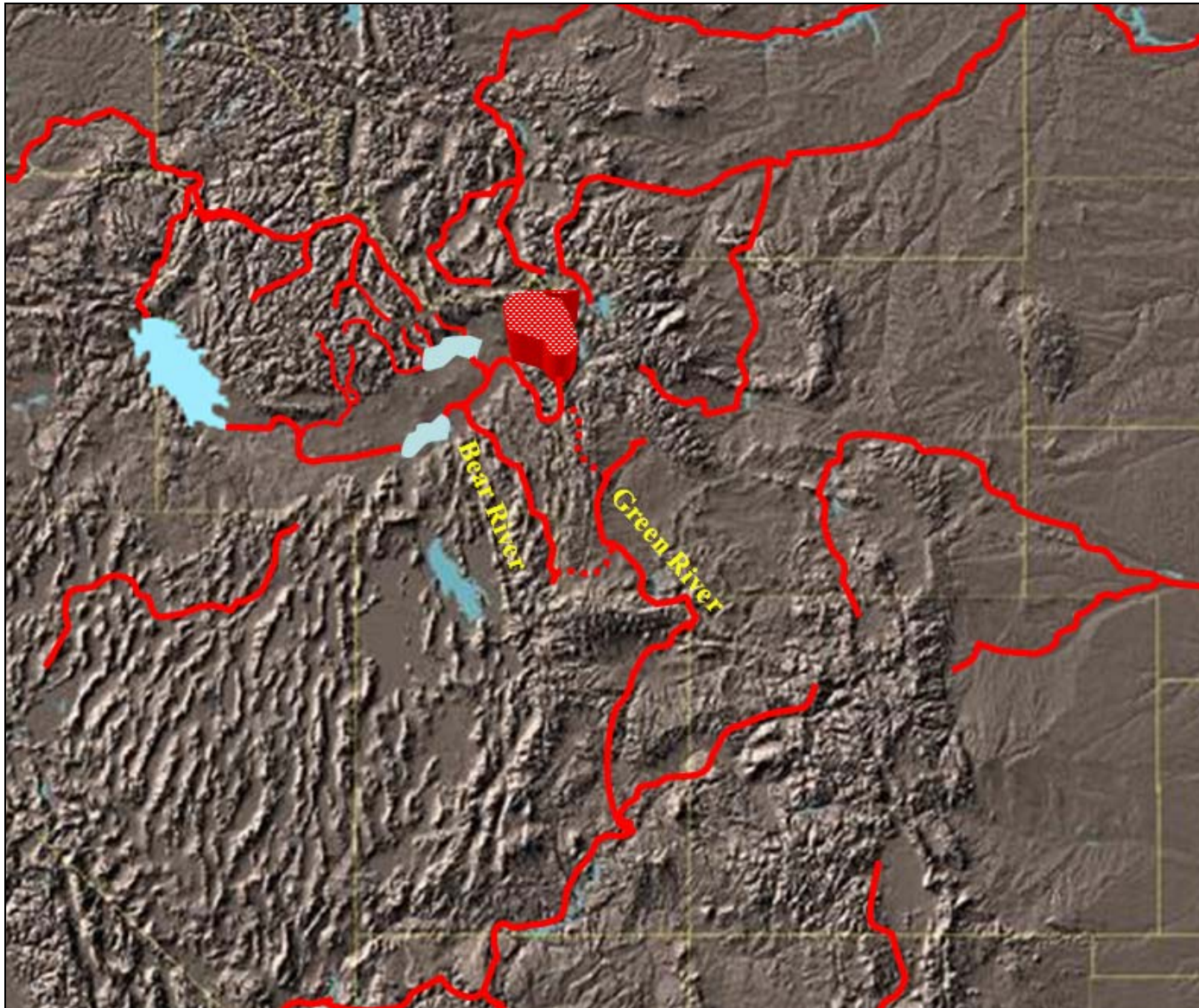


Figure 8.

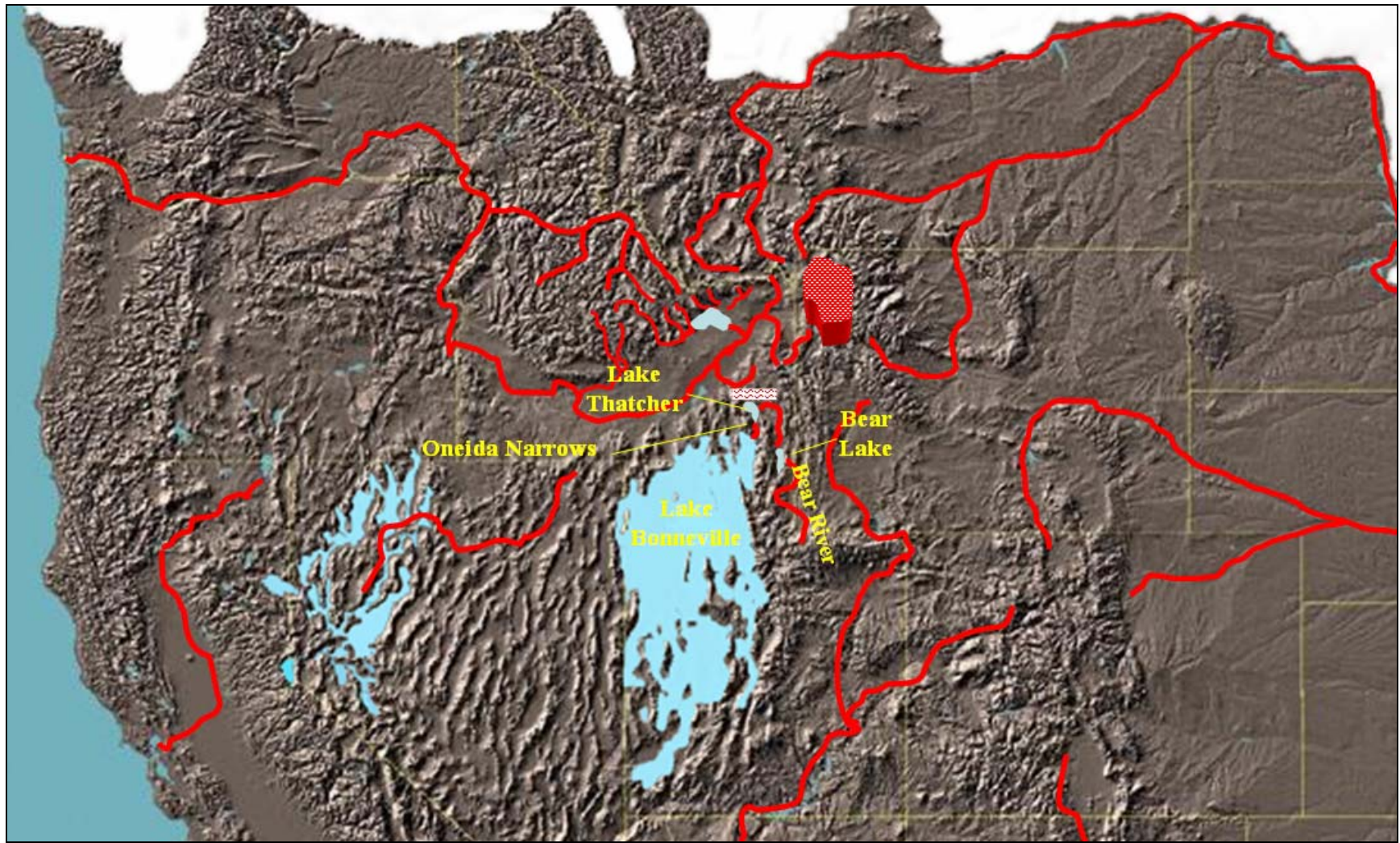


Figure 9.

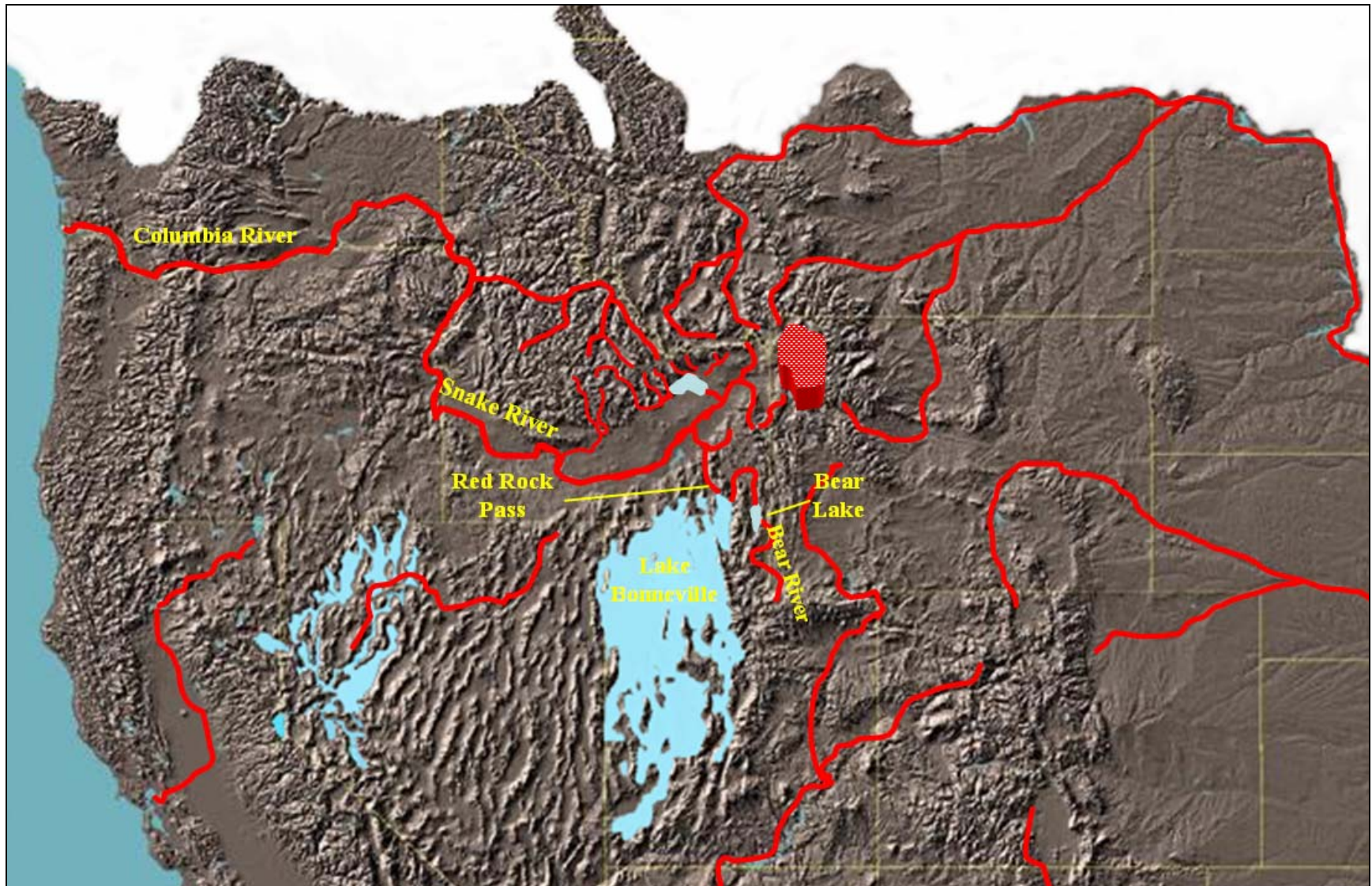


Figure 10.

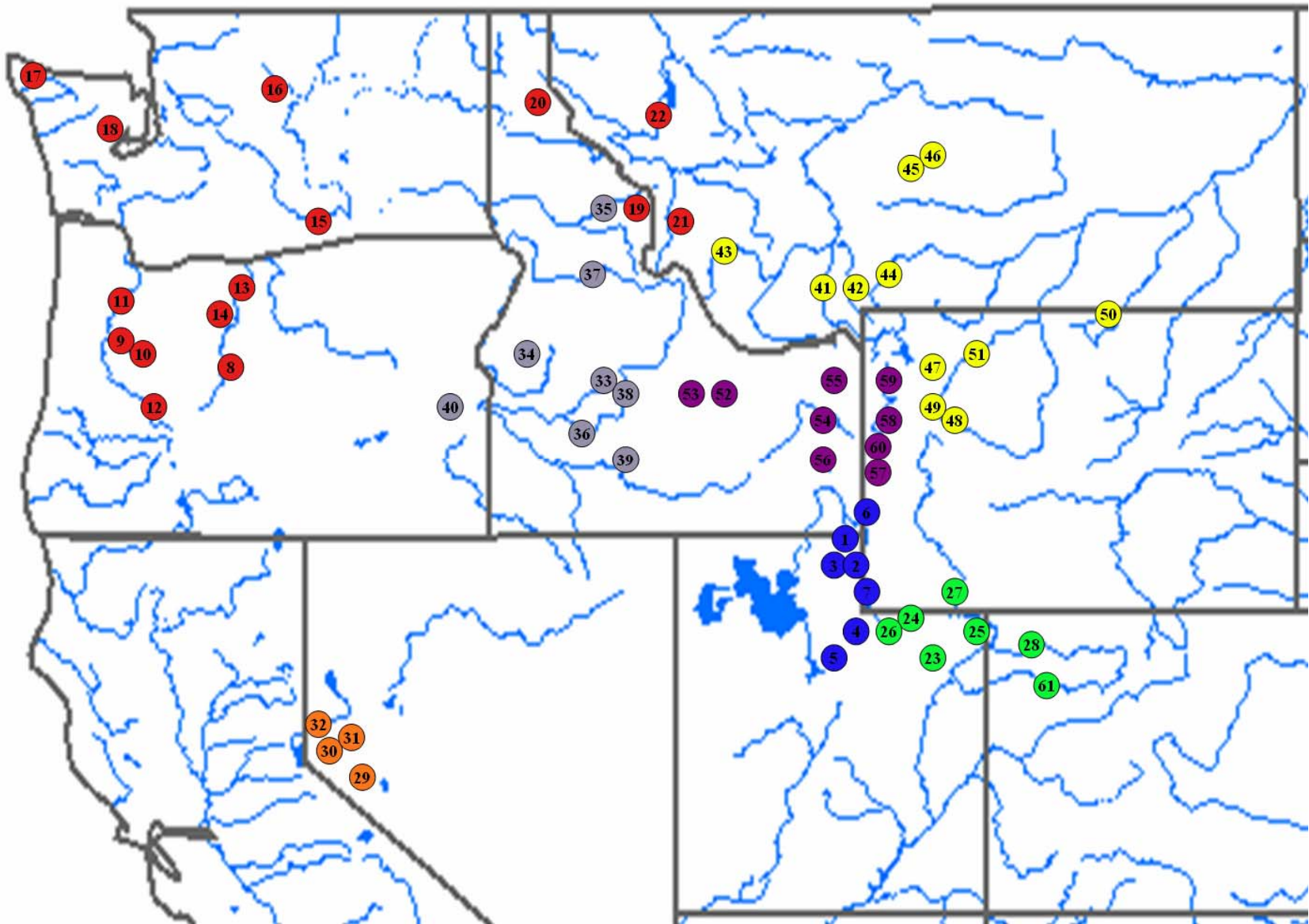


Figure 11.

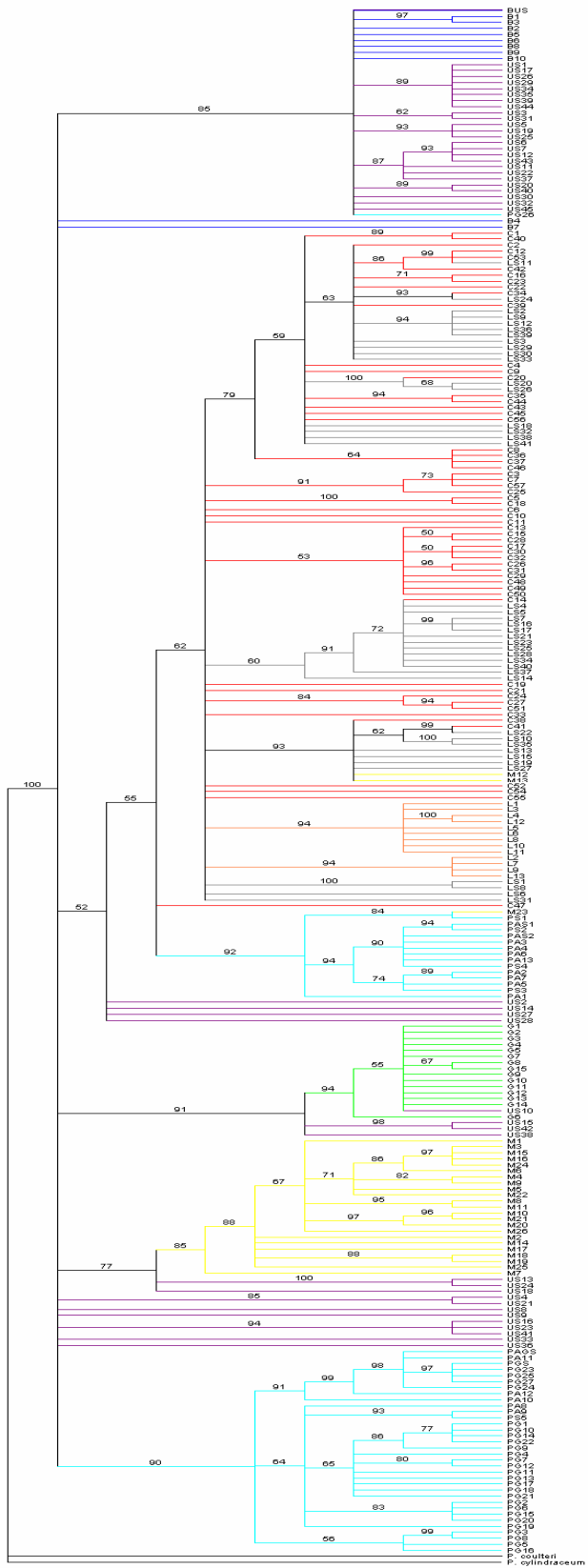


Figure 12.

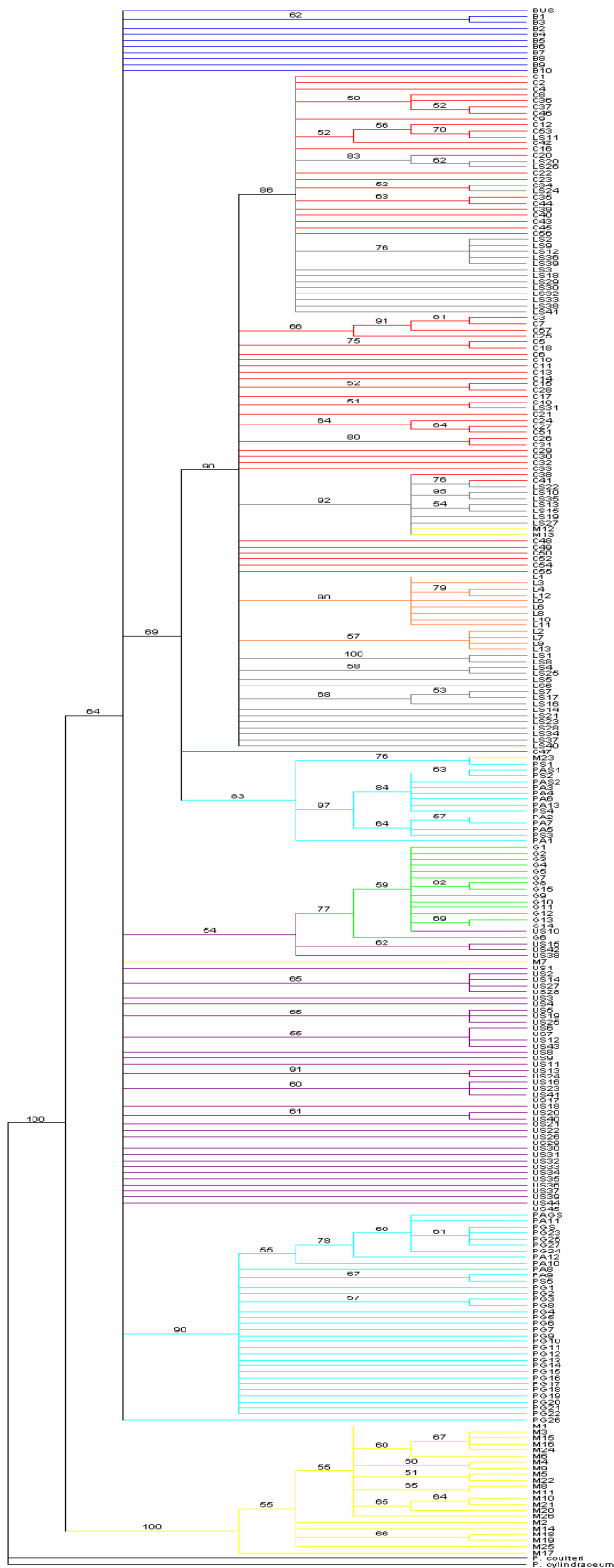


Figure 13.

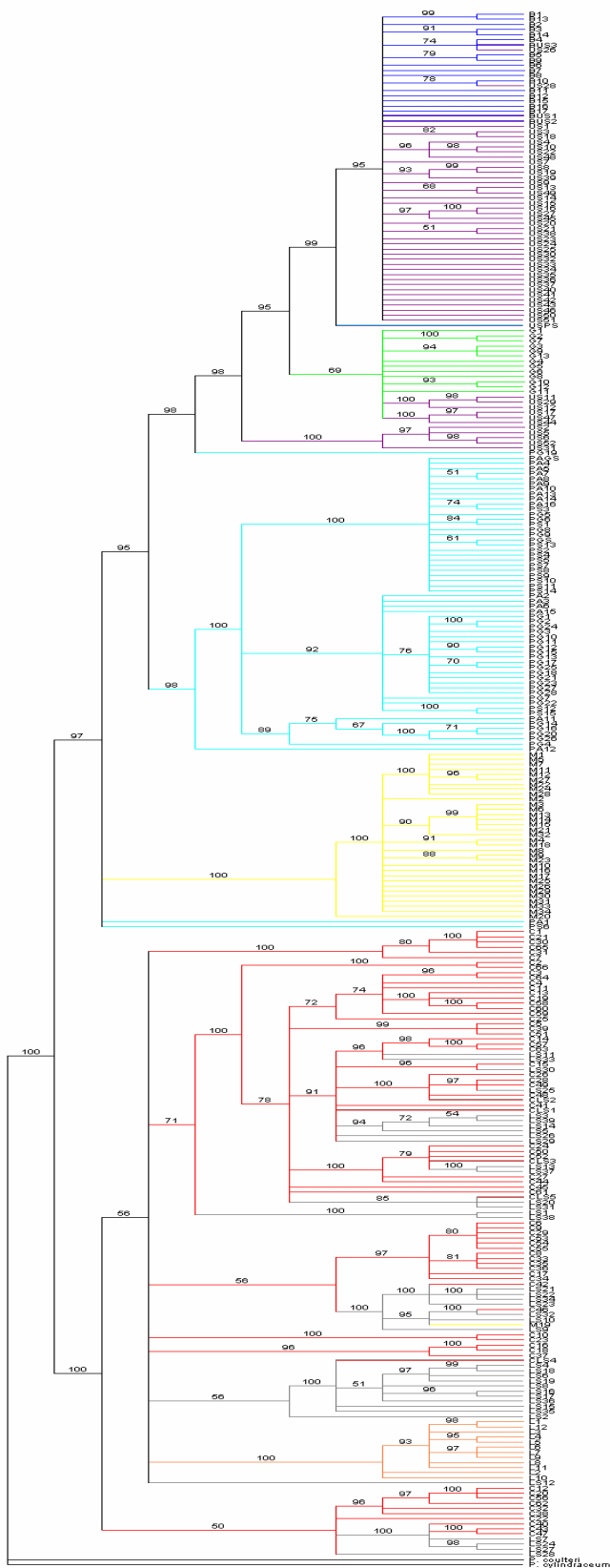


Figure 14.

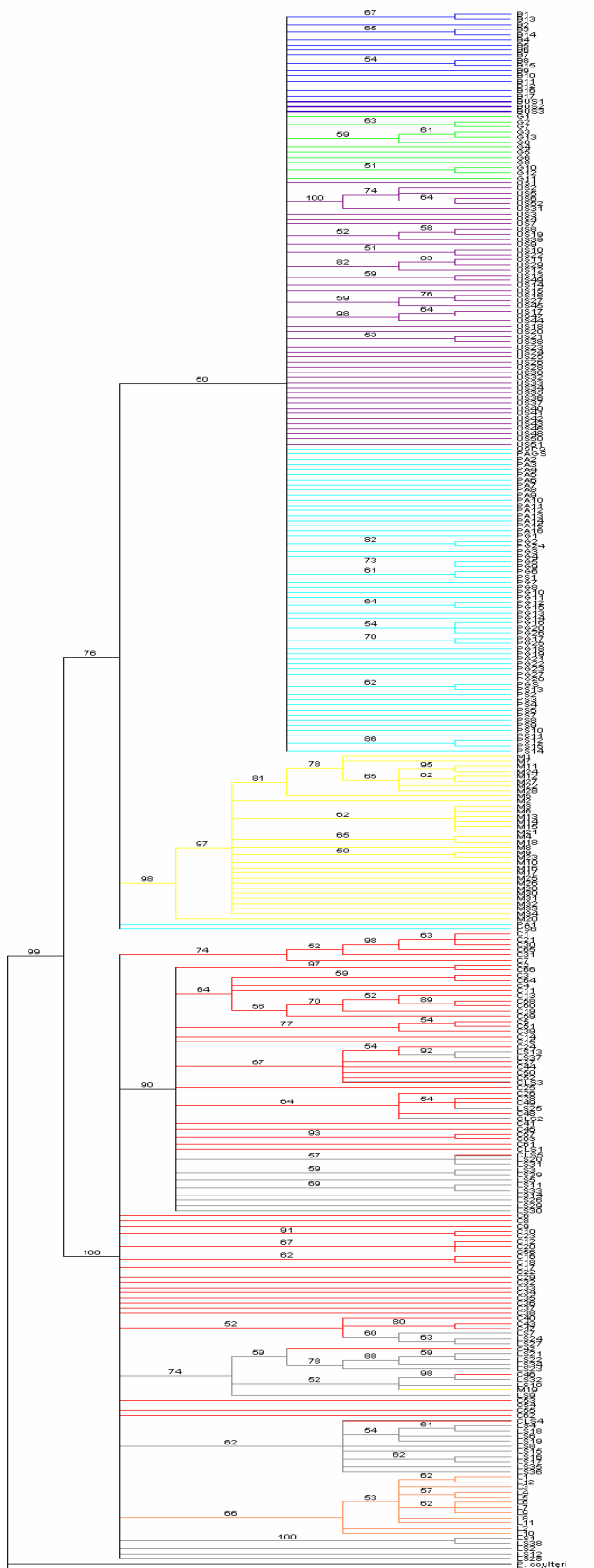


Figure 15a.

(i)

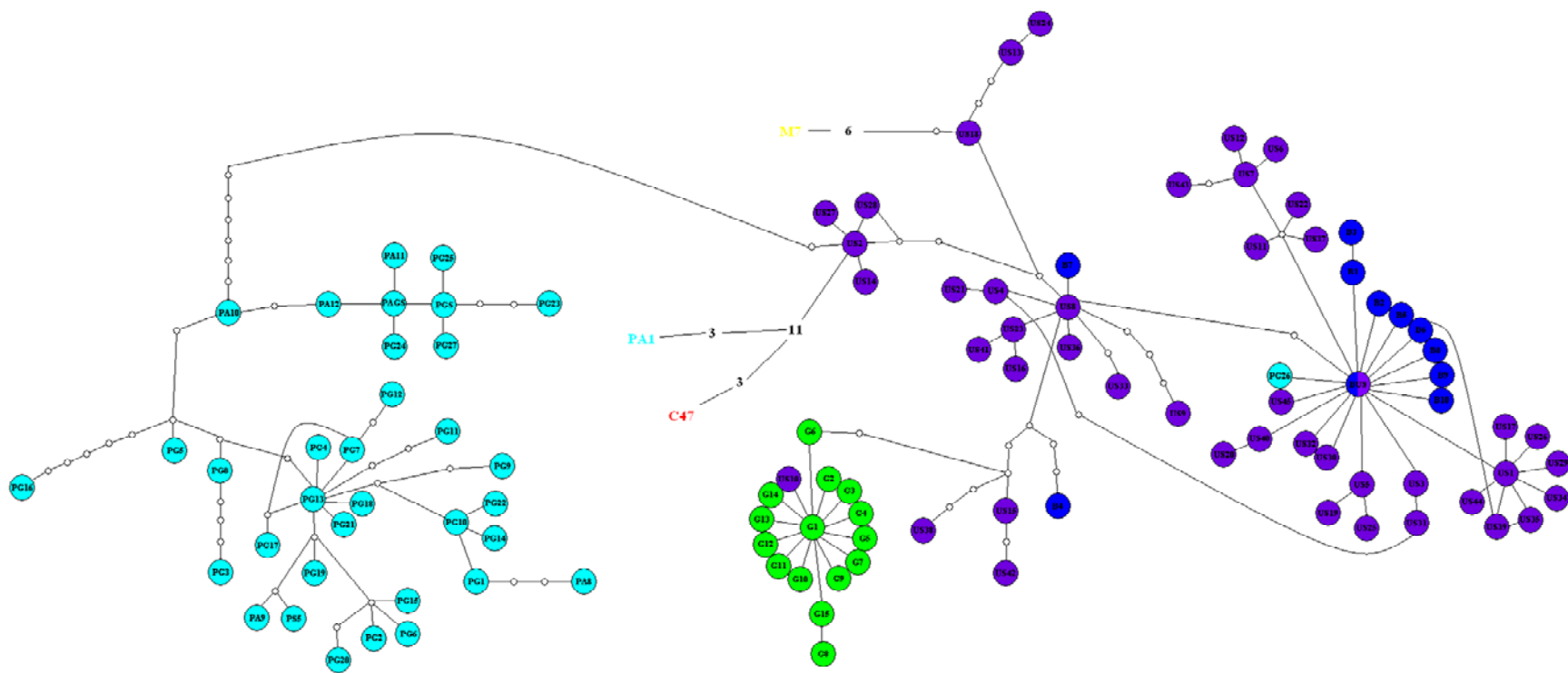


Figure 15a continued.

(ii)

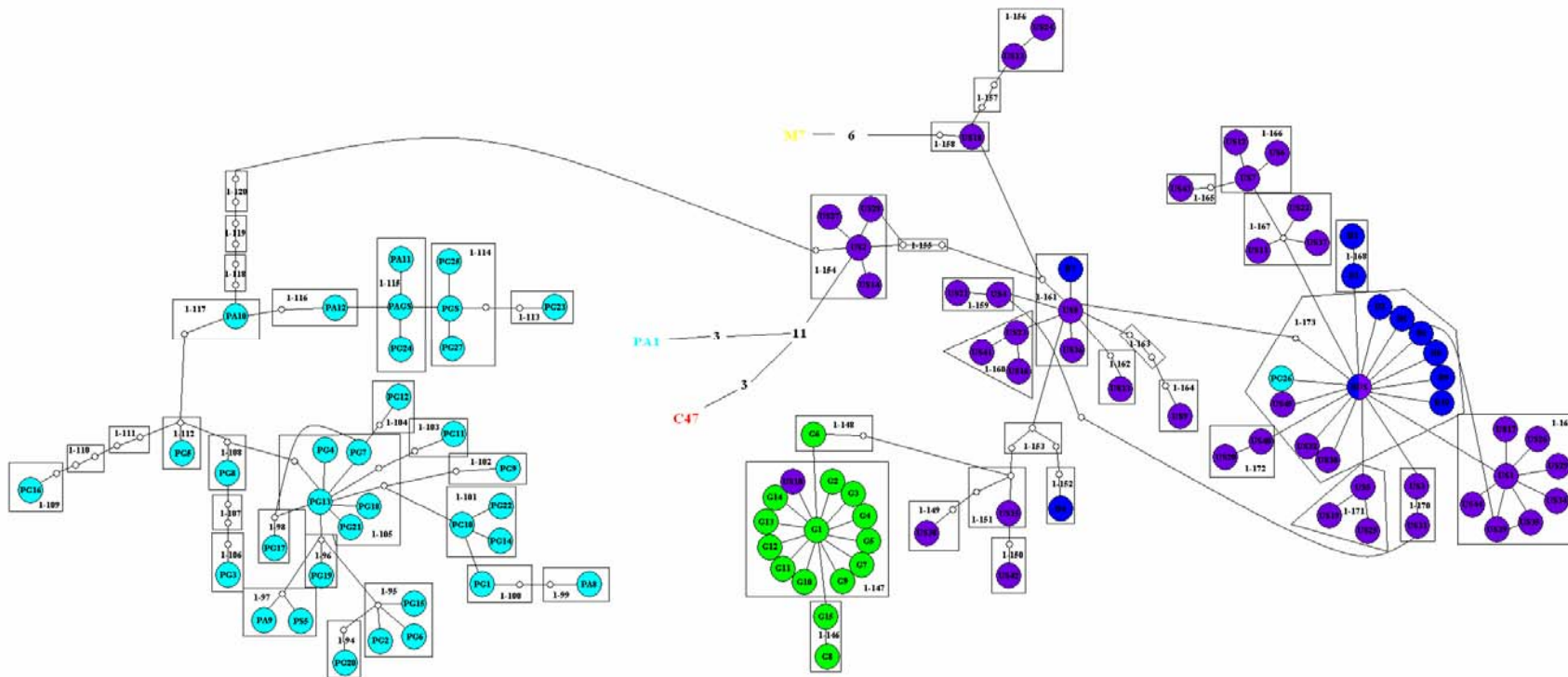


Figure 15a continued.

(iii)

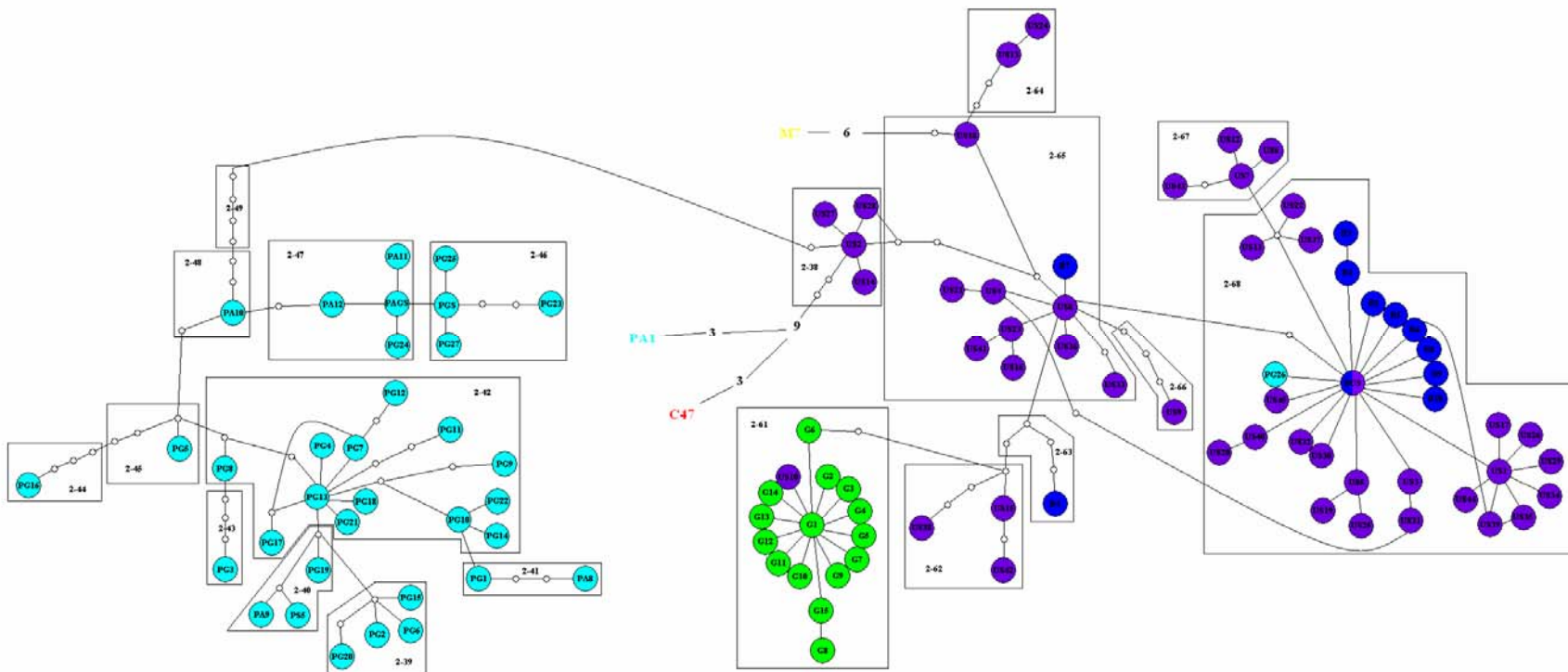


Figure 15a continued.

(iv)

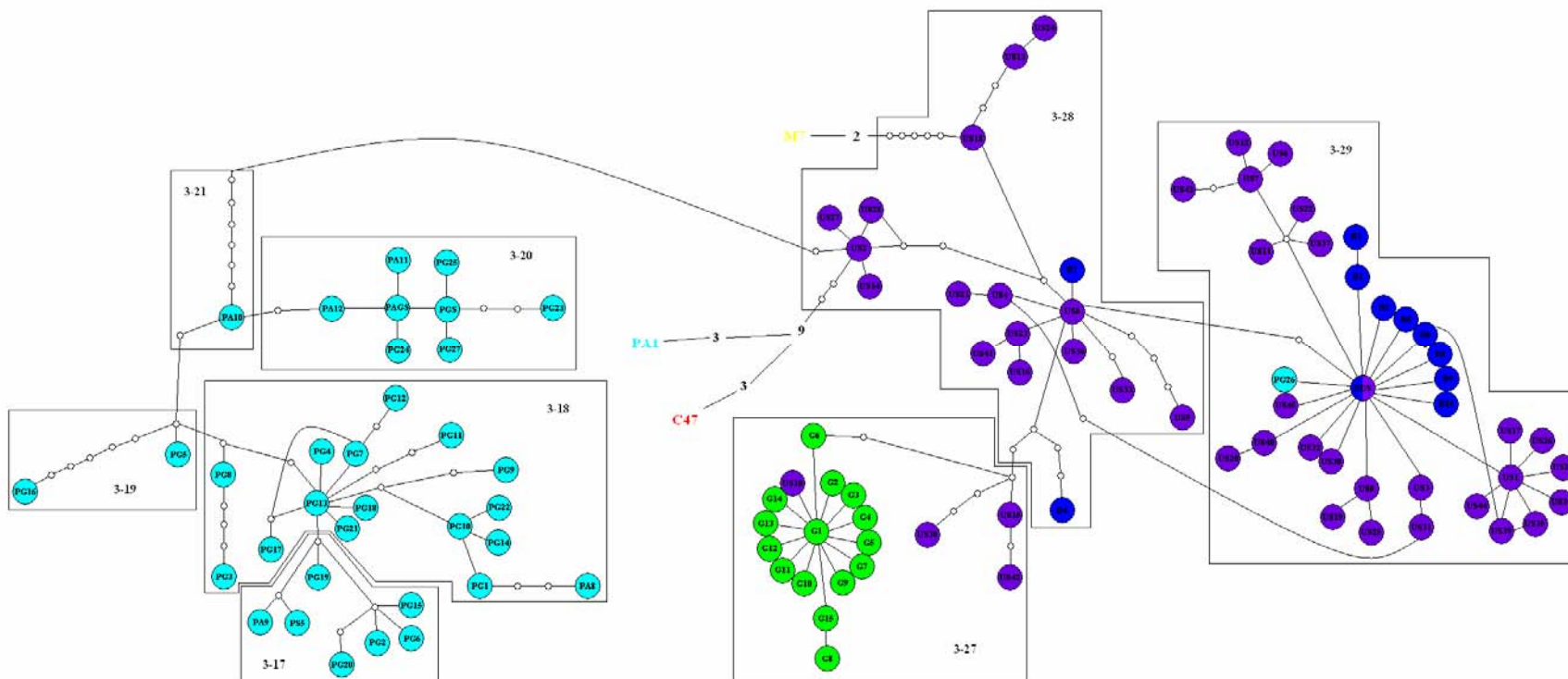


Figure 15a continued.

(v)

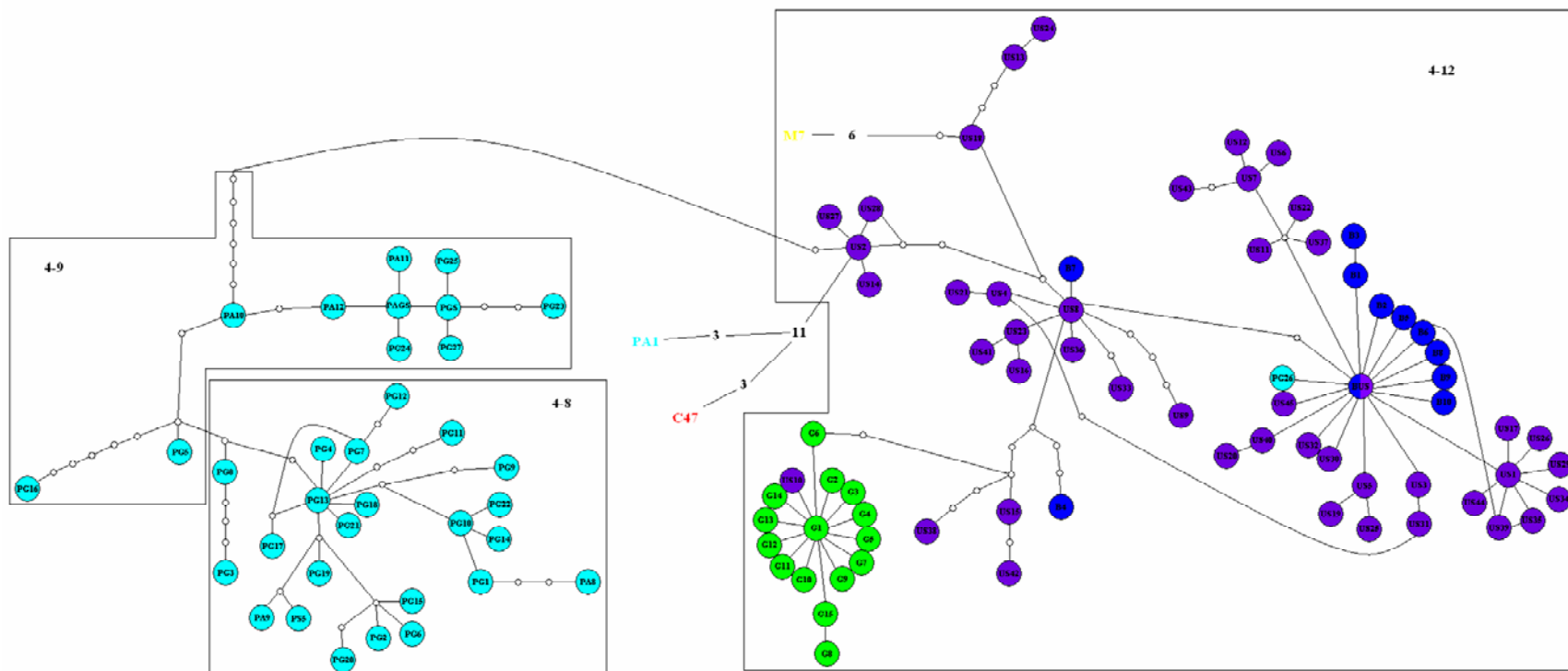


Figure 15a continued.

(vi)

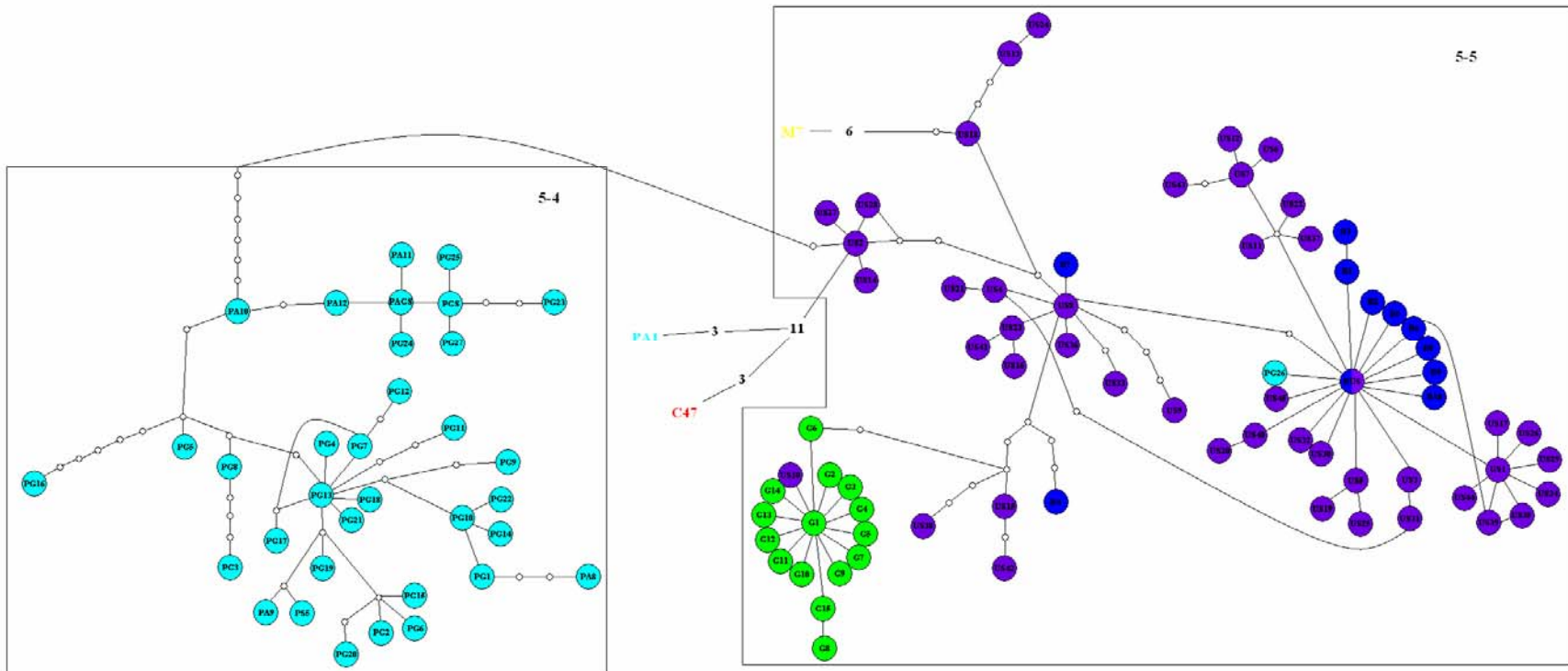


Figure 15a continued.

(vii)

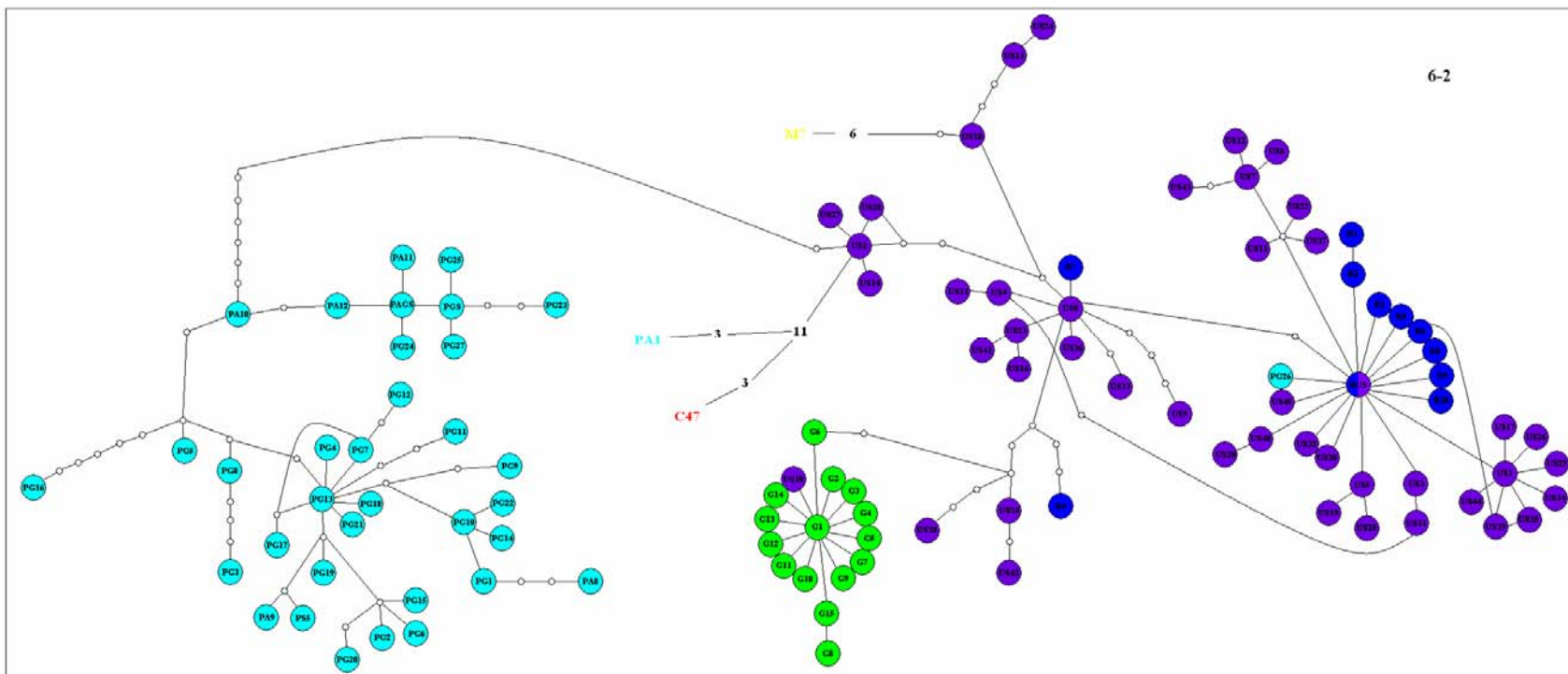


Figure 15a continued.

(viii)

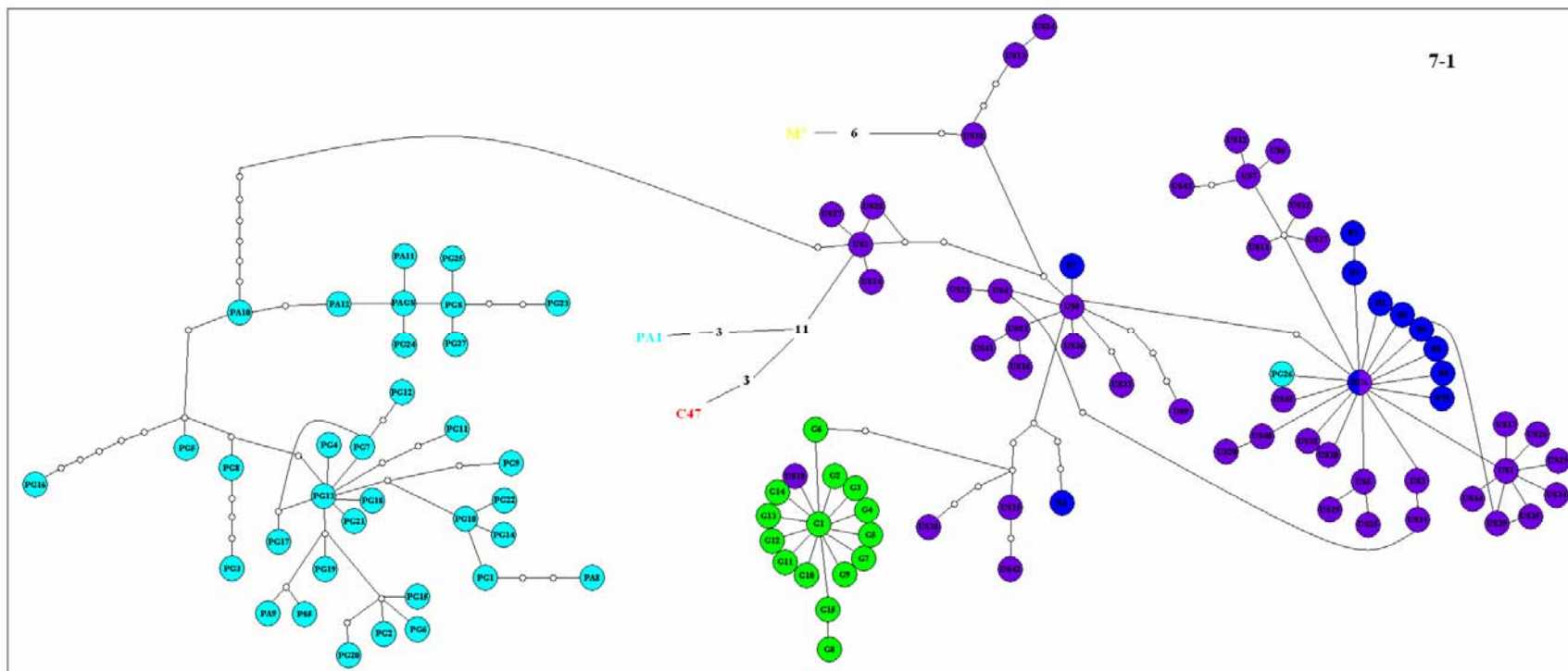
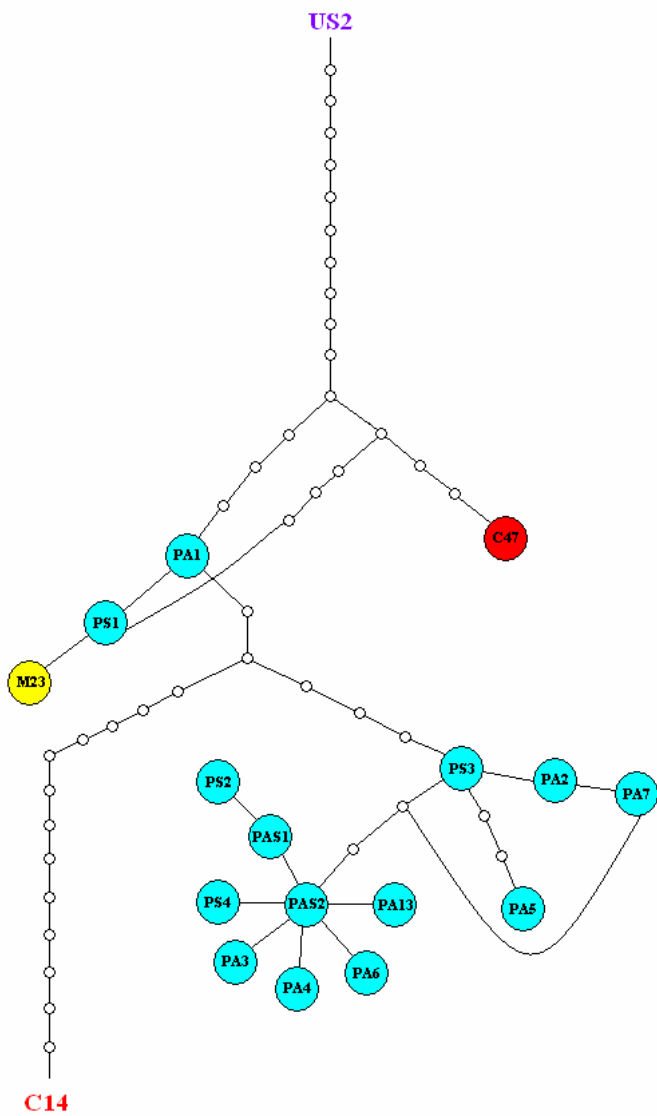


Figure 15b.

(i)



(ii)

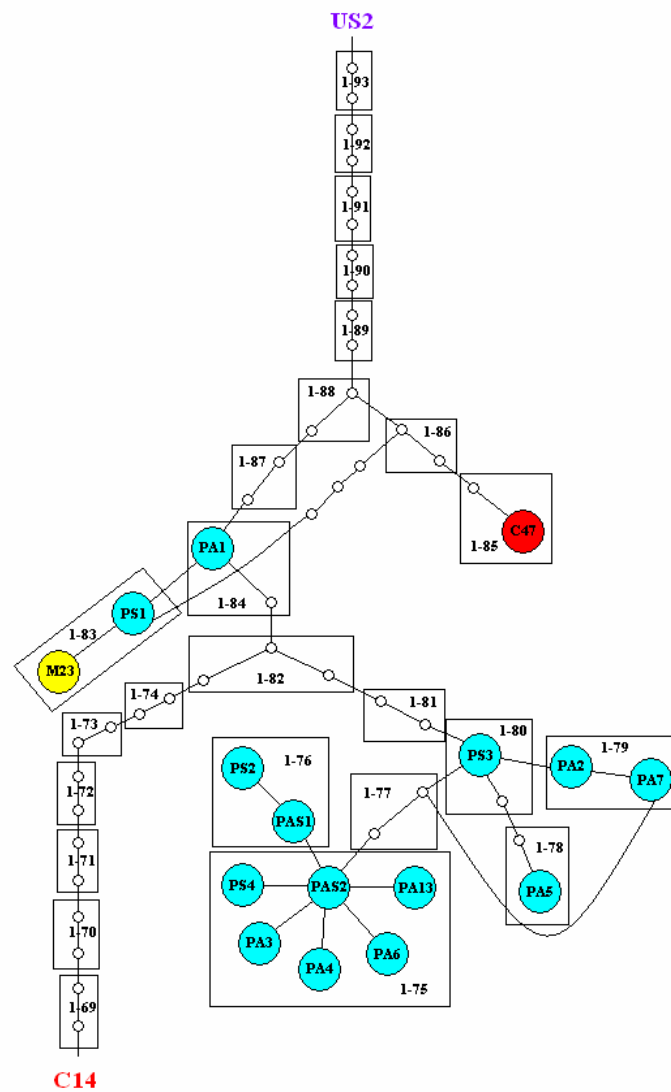
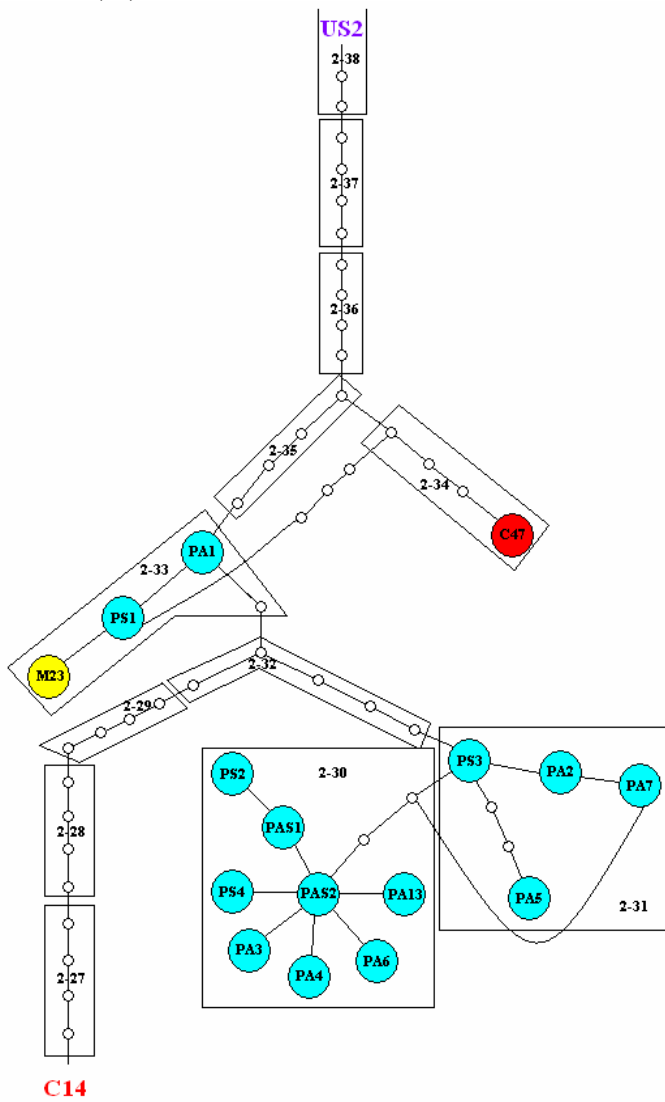


Figure 15b continued.

(iii)



(iv)

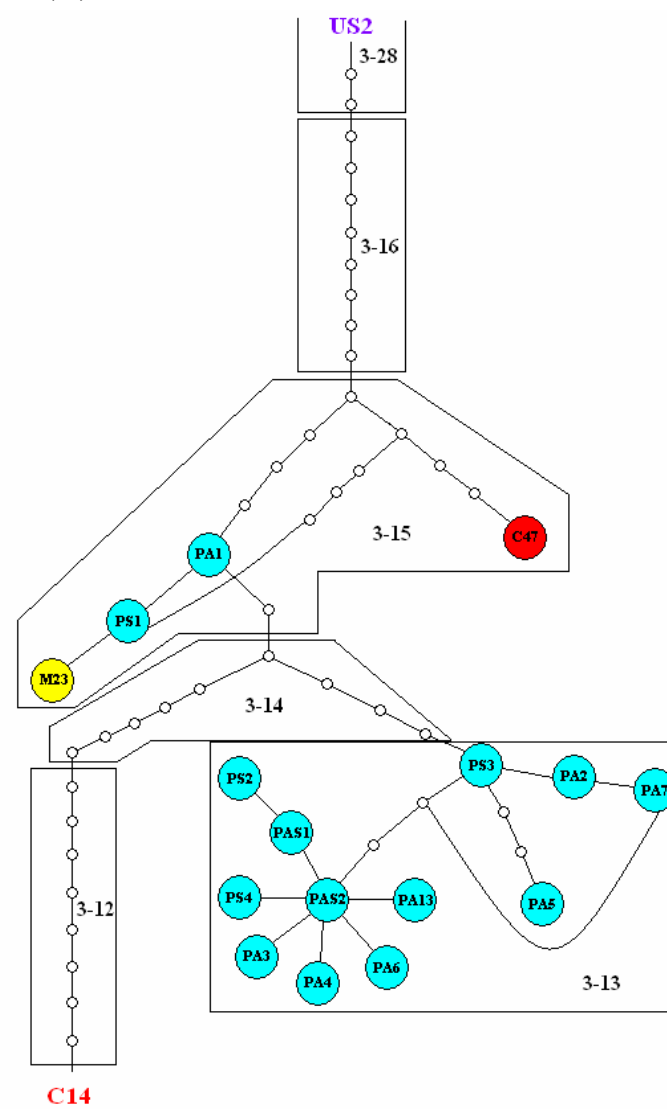
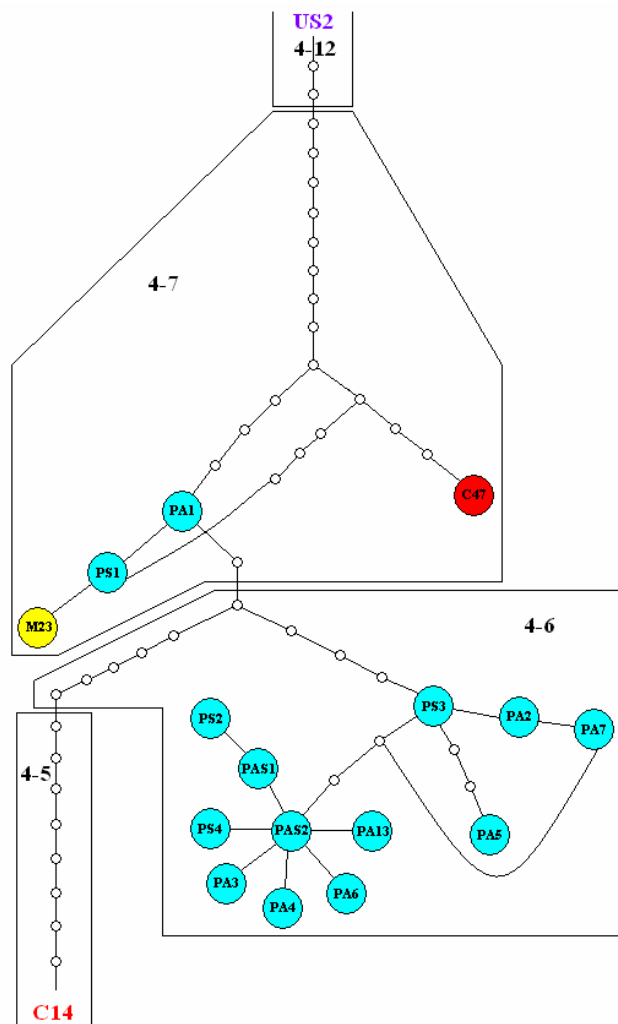


Figure 15b continued.

(v)



(vi)

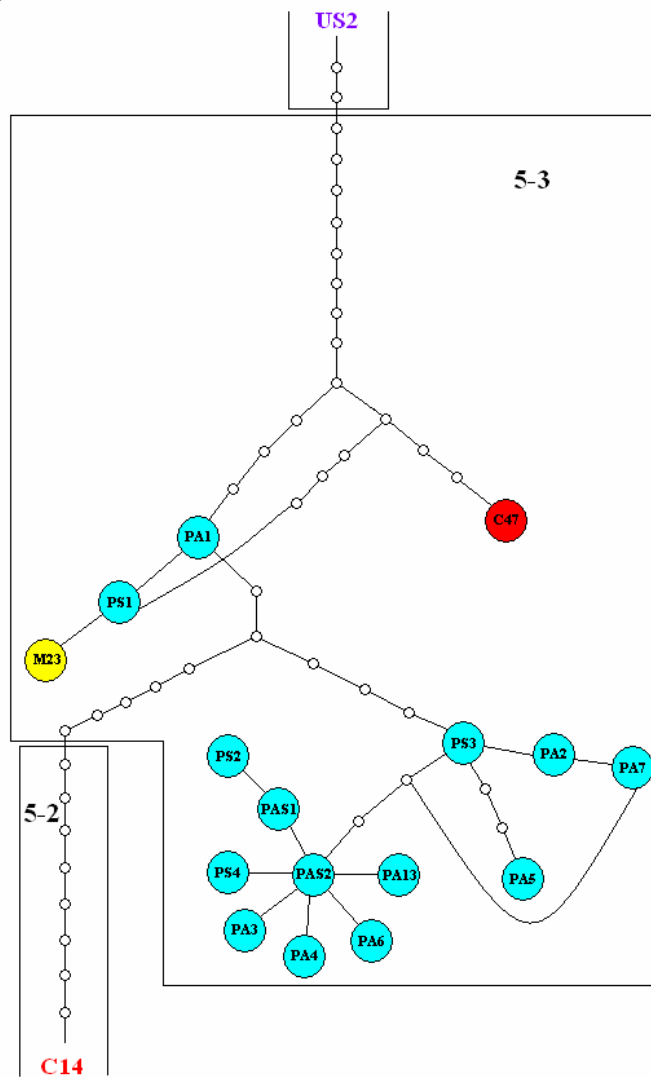
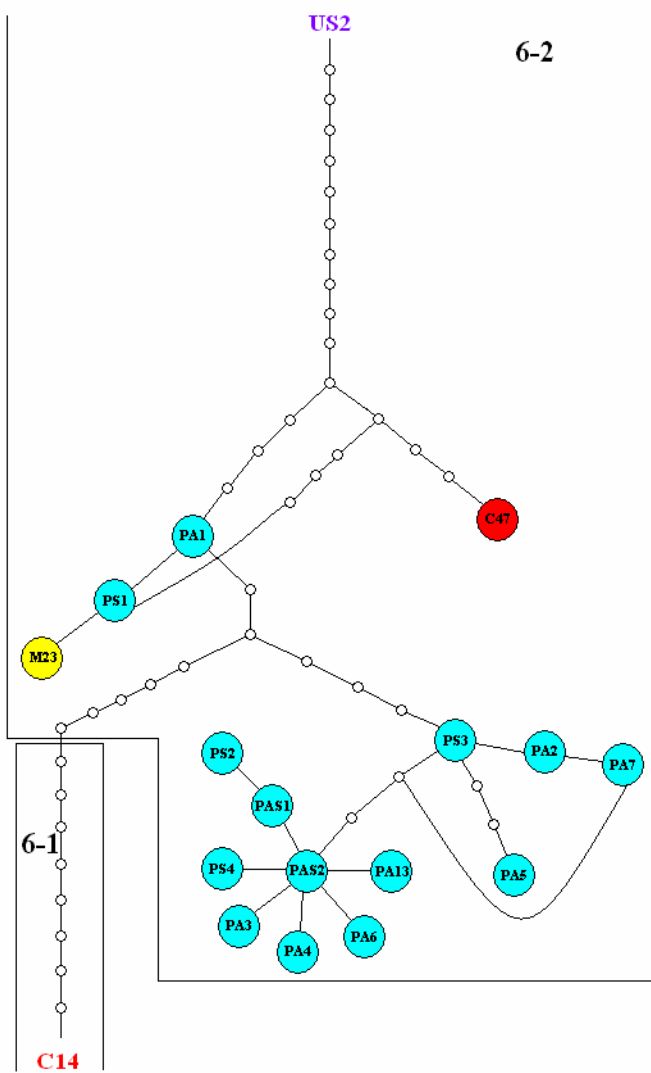


Figure 15b continued.

(vii)



(viii)

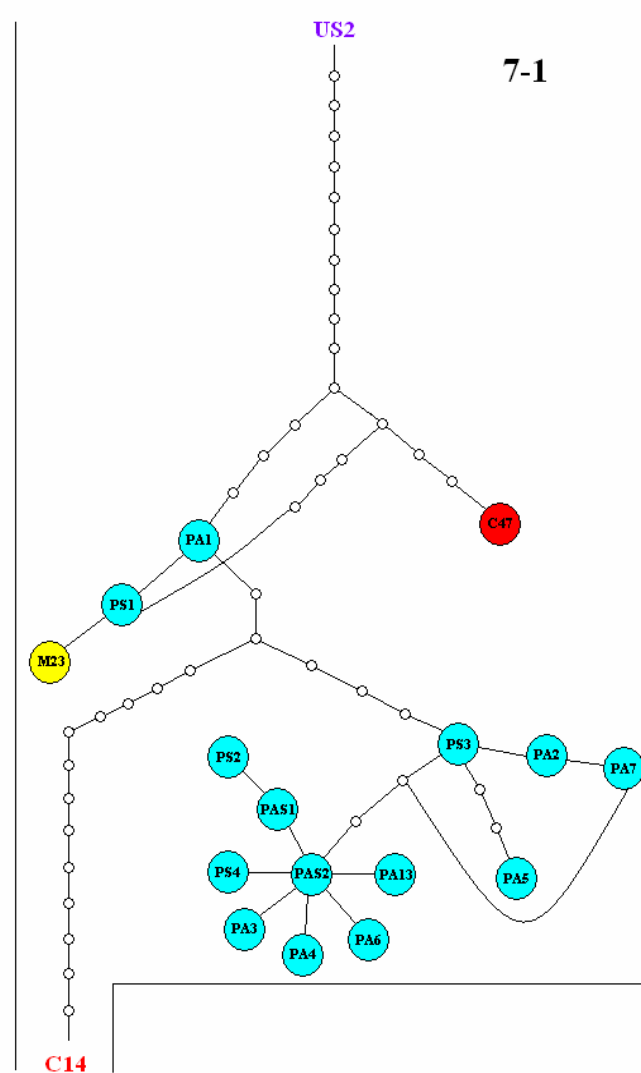


Figure 15c.

(i)

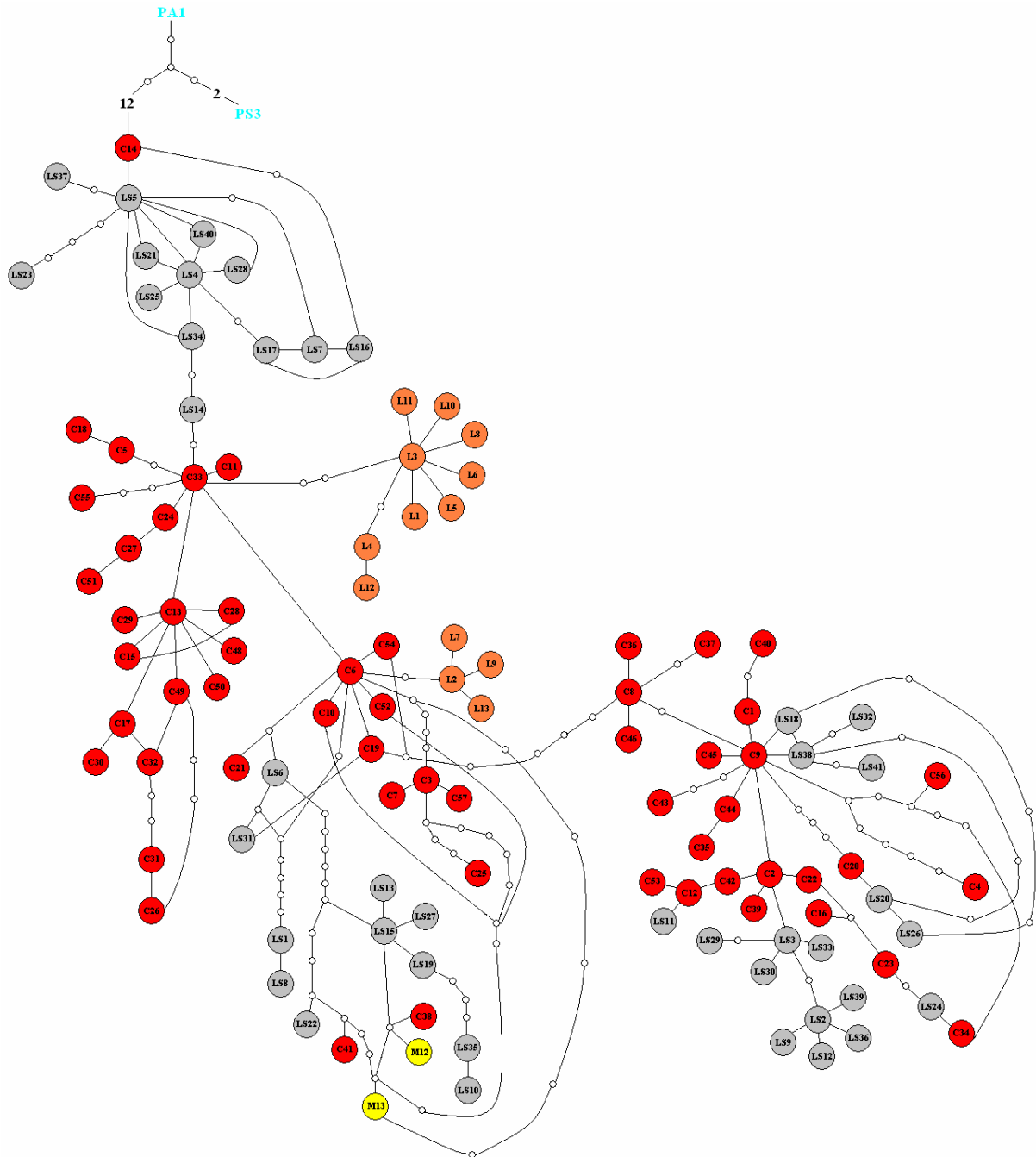


Figure 15c continued.

(ii)

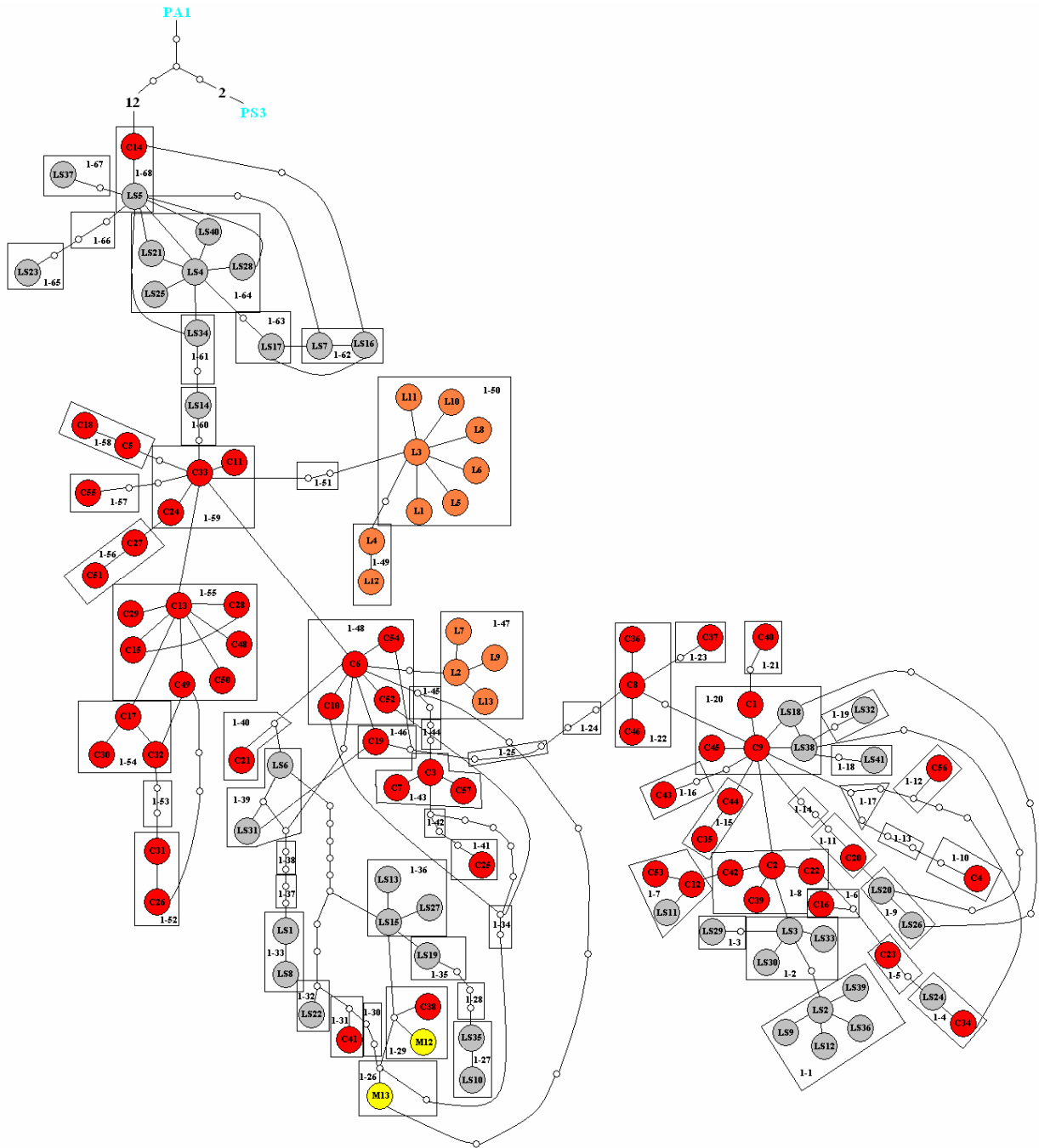


Figure 15c continued.

(iv)

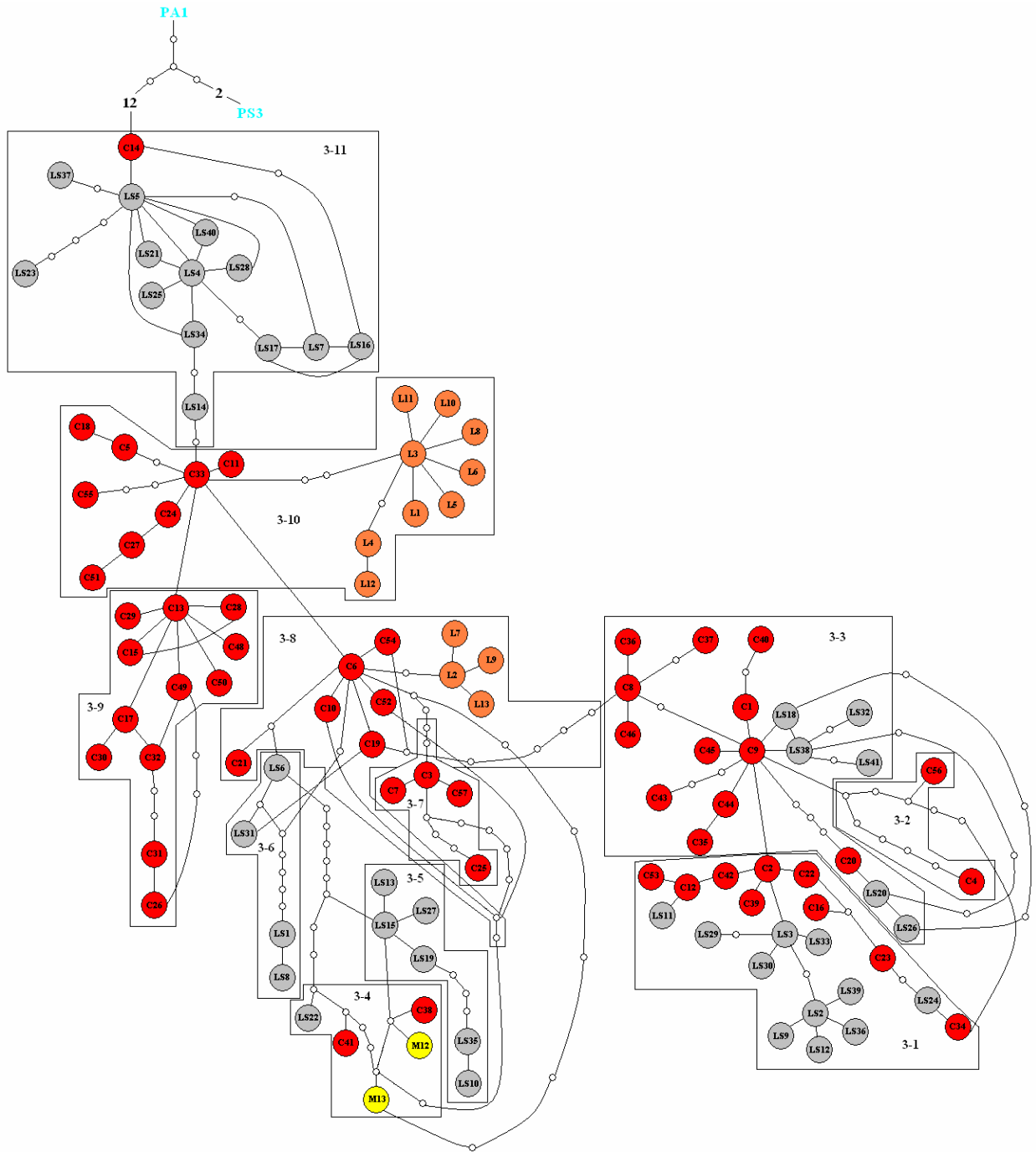


Figure 15c continued.

(v)

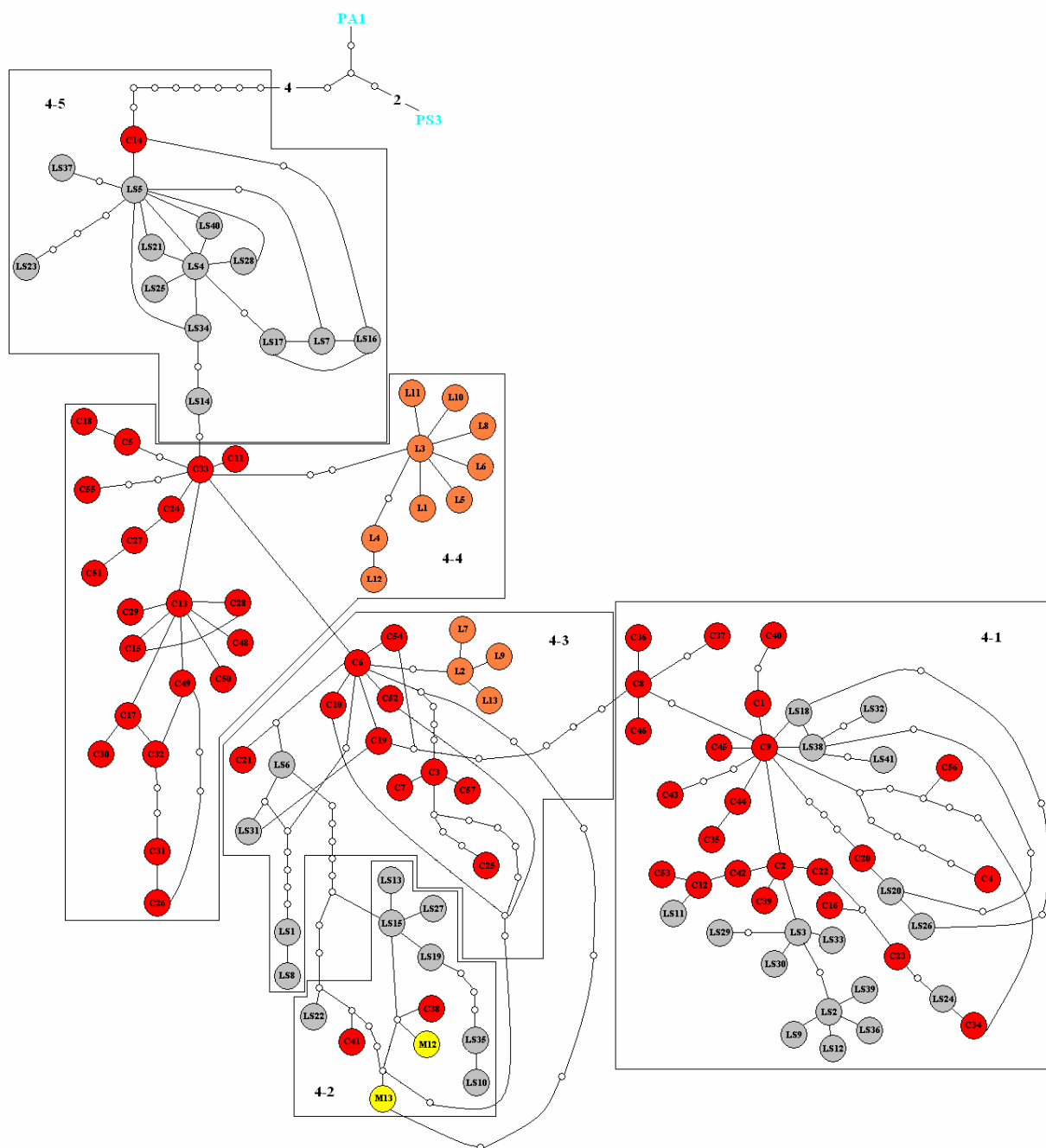


Figure 15c continued.

(vi)

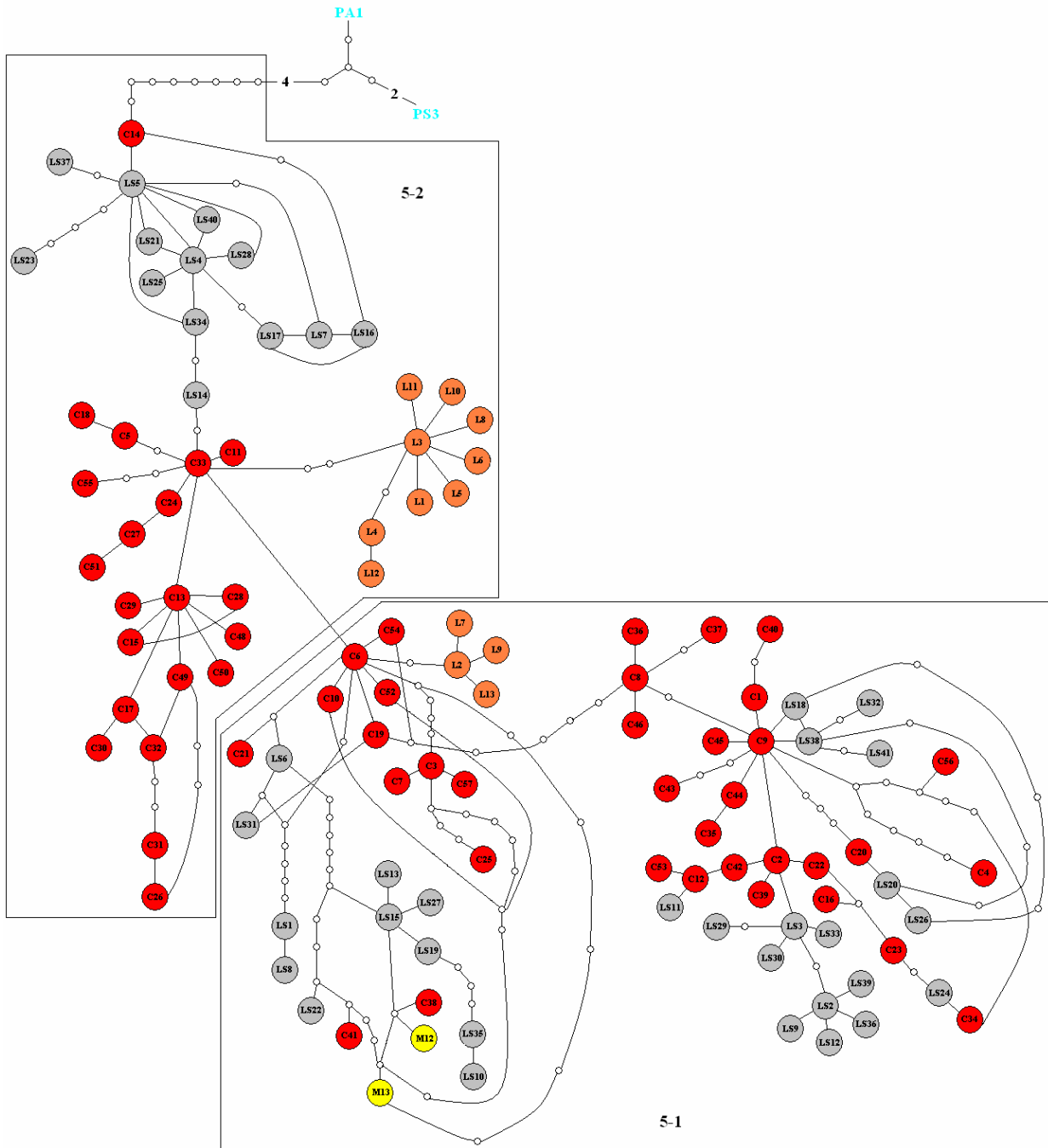


Figure 15c continued.

(vii)

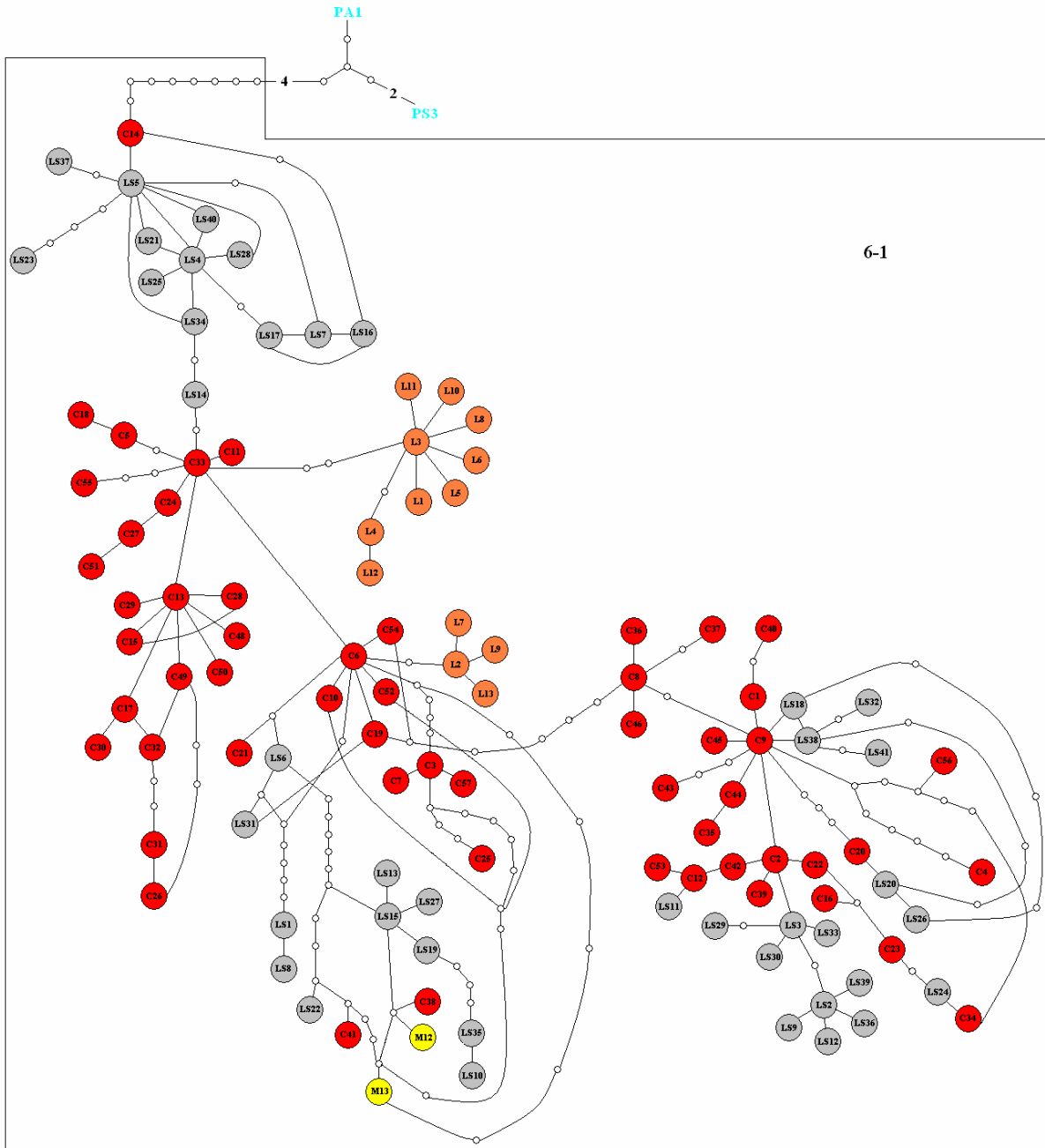


Figure 15c continued.

(viii)

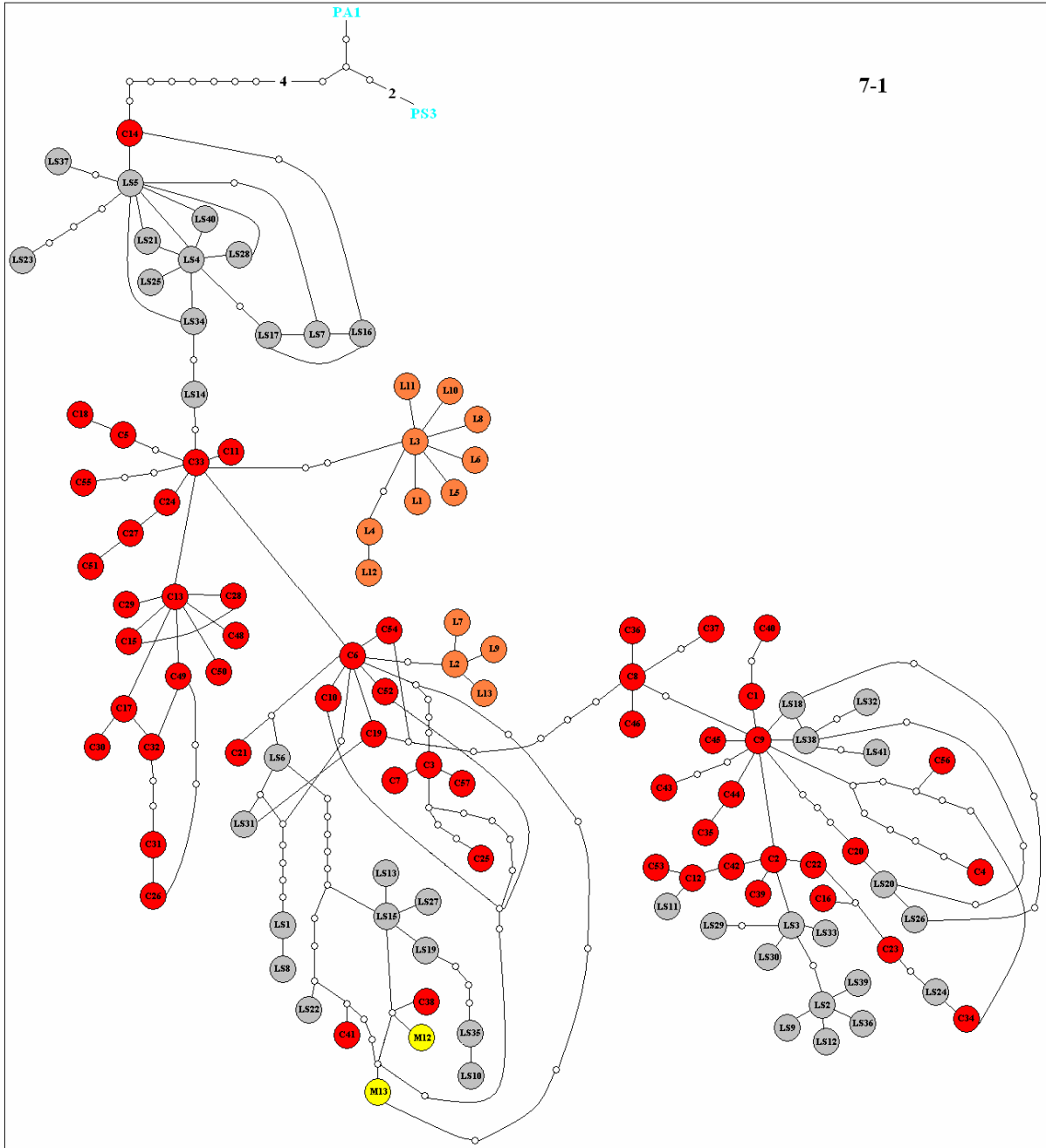
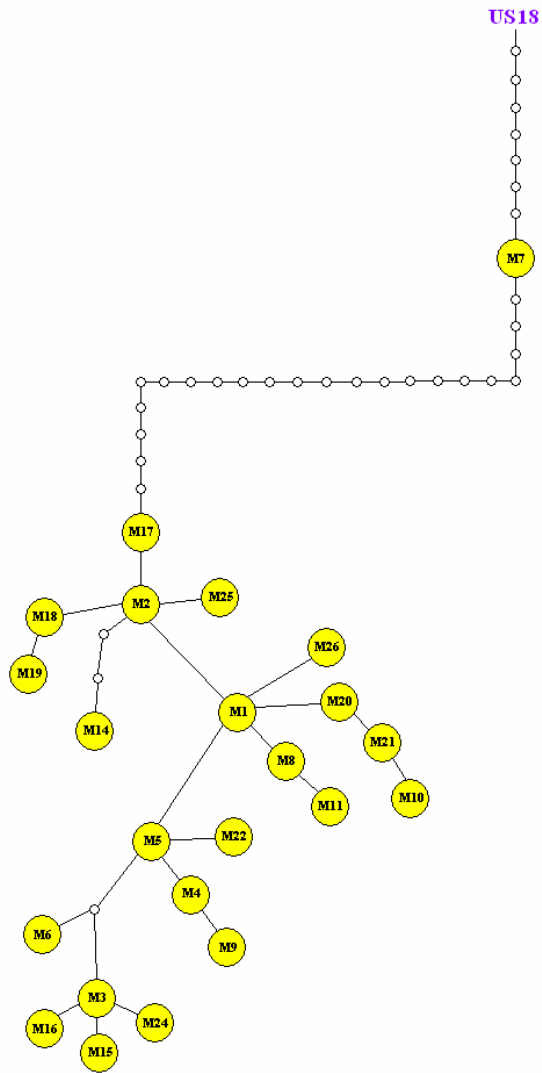


Figure 15d.

(i)



(ii)

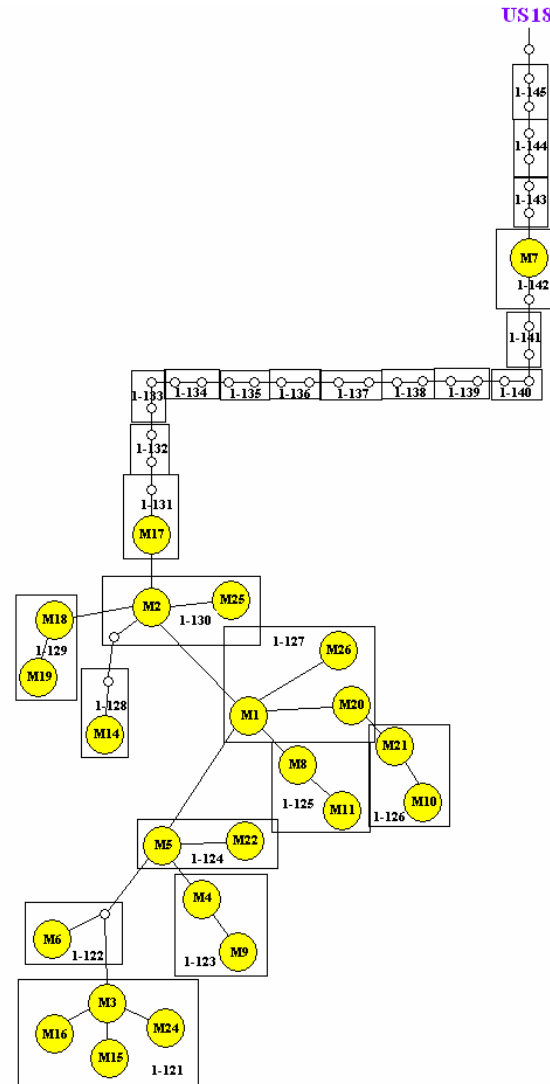
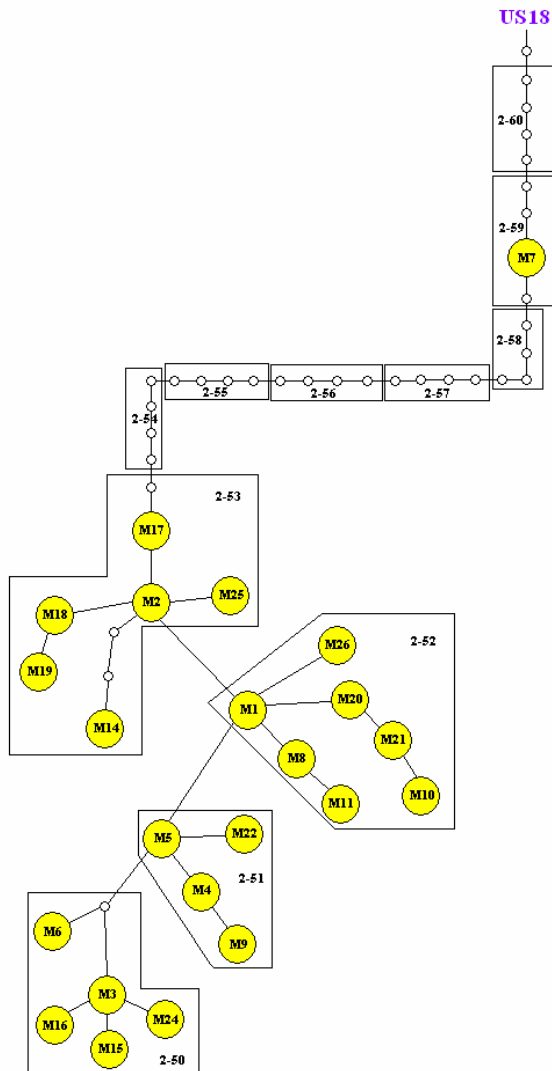


Figure 15d continued.

(iii)



(iv)

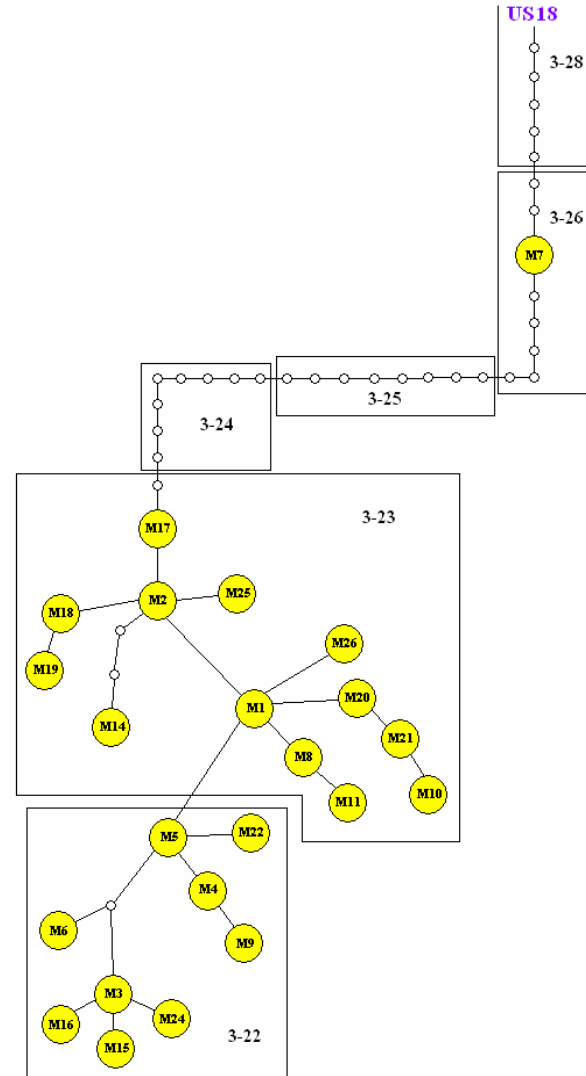
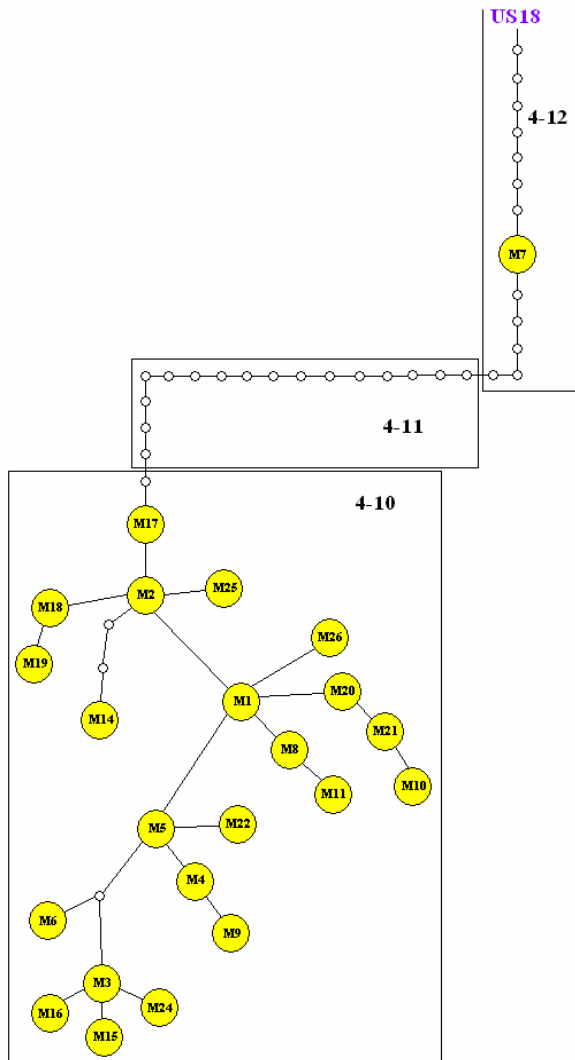


Figure 15d continued.

(v)



(vi)

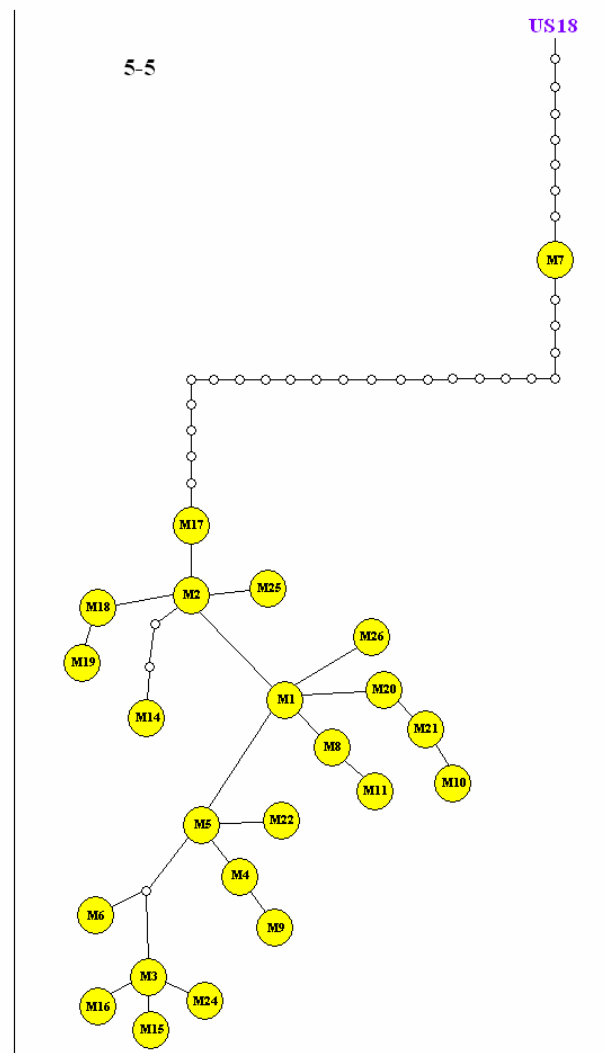
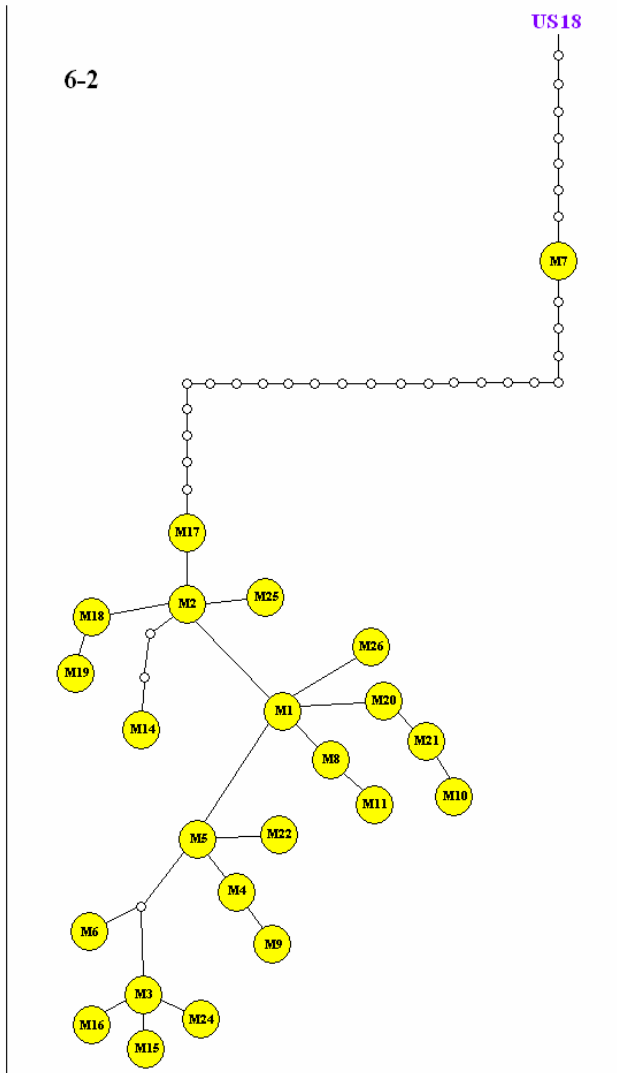


Figure 15d continued.

(vii)



(viii)

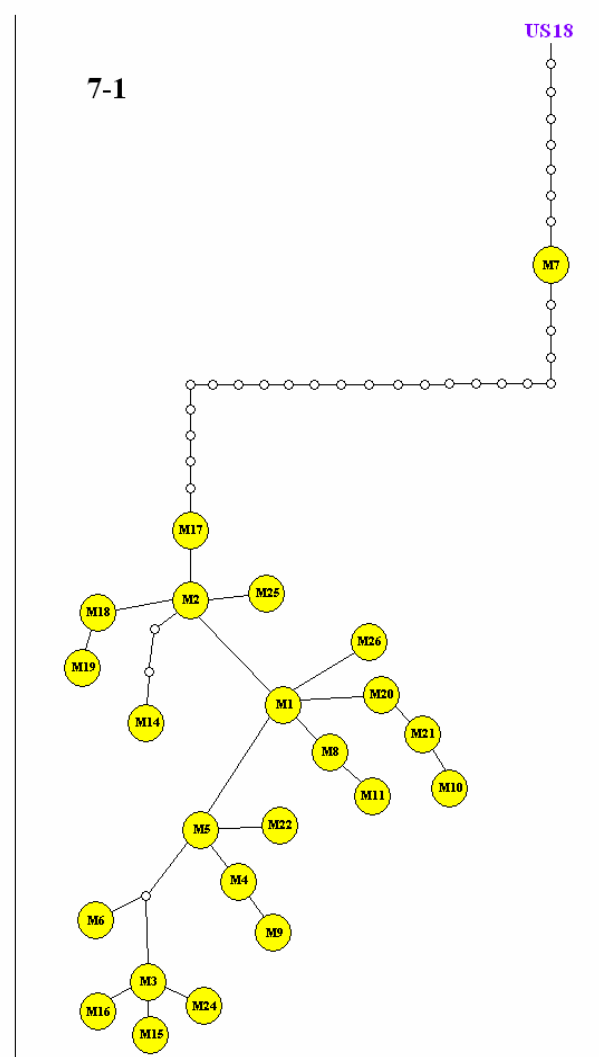


Figure 16a.

(i)

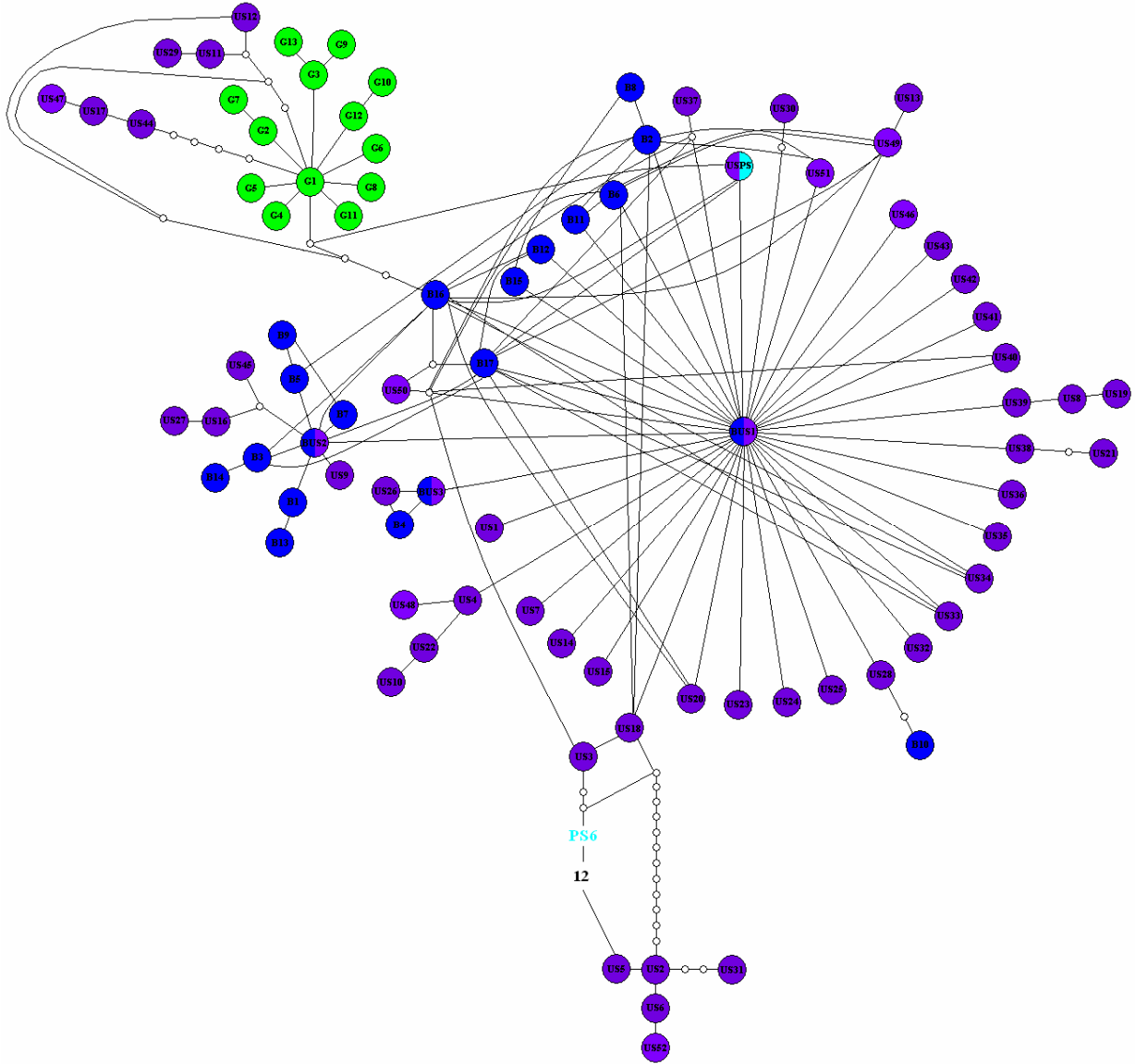


Figure 16a continued.

(ii)

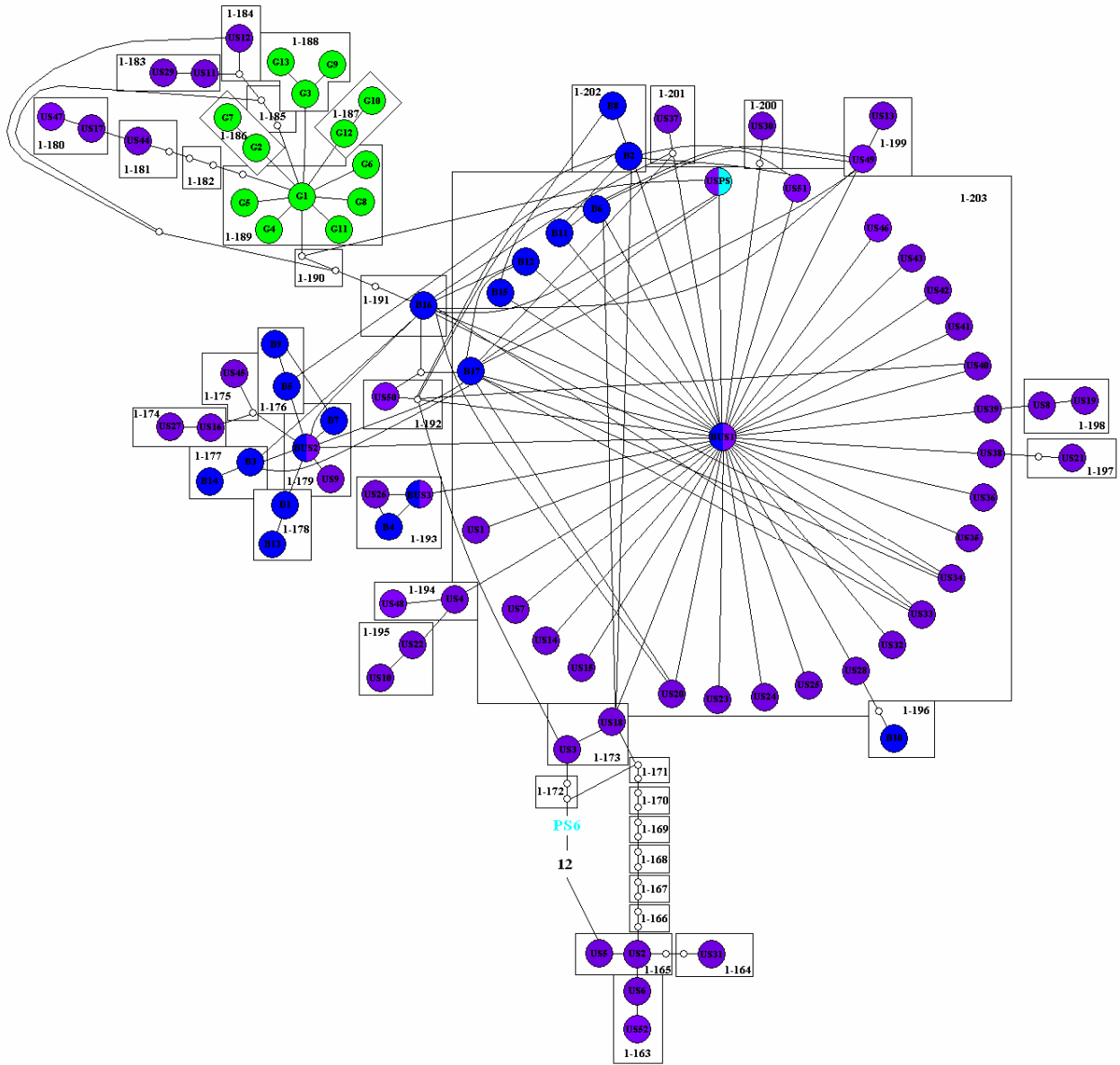


Figure 16a continued.

(iii)

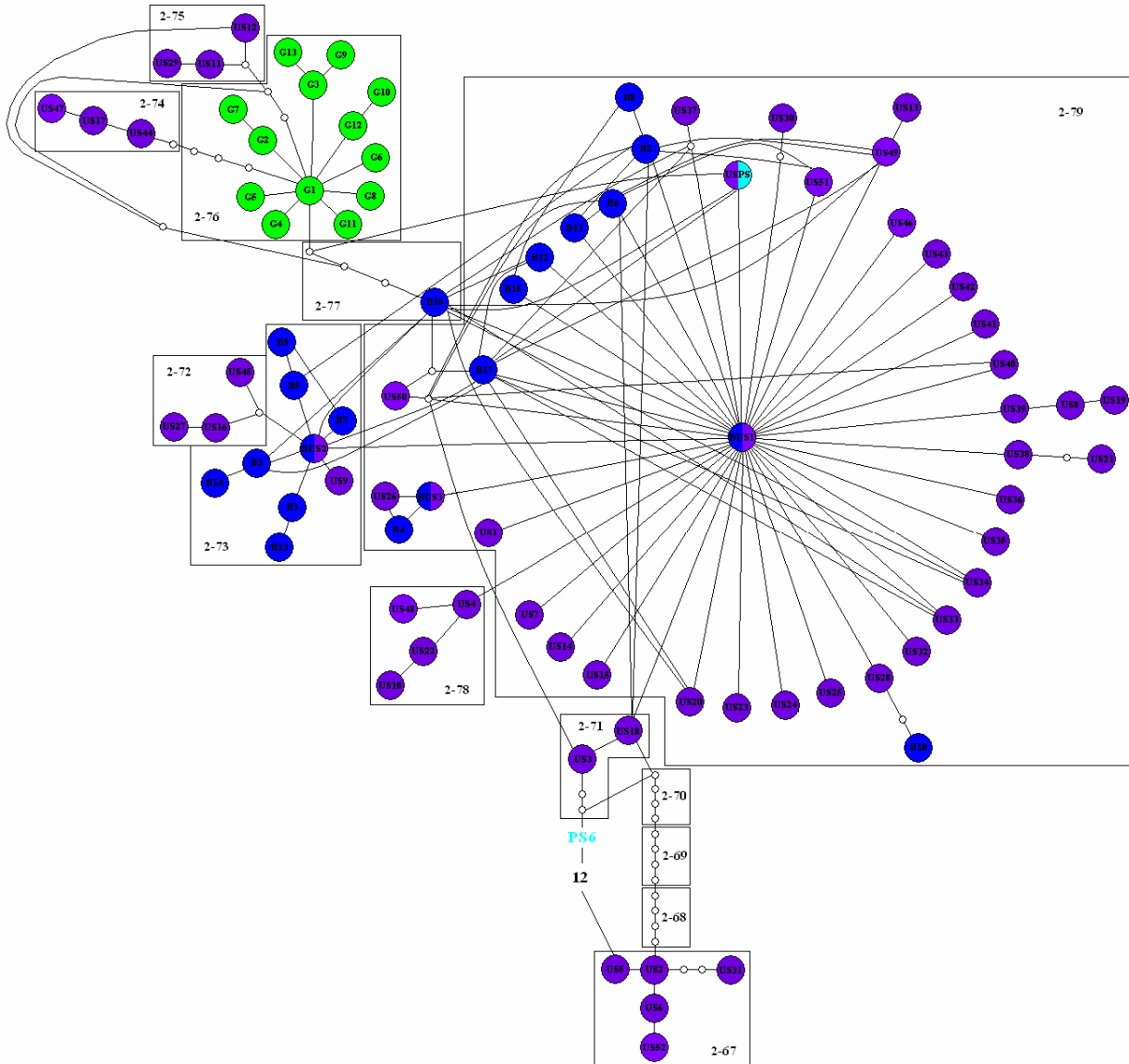


Figure 16a continued.

(iv)

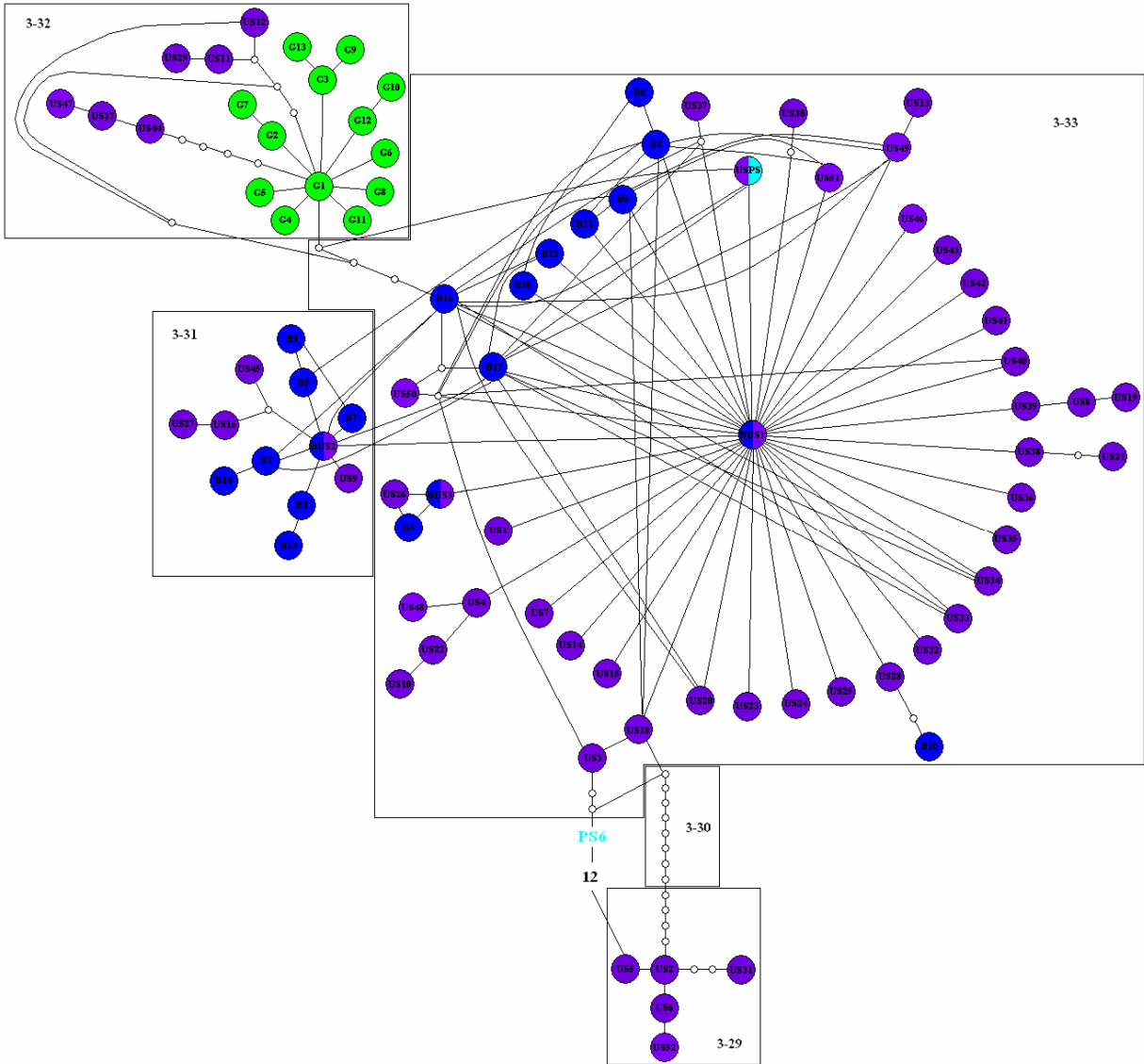


Figure 16a continued.

(v)

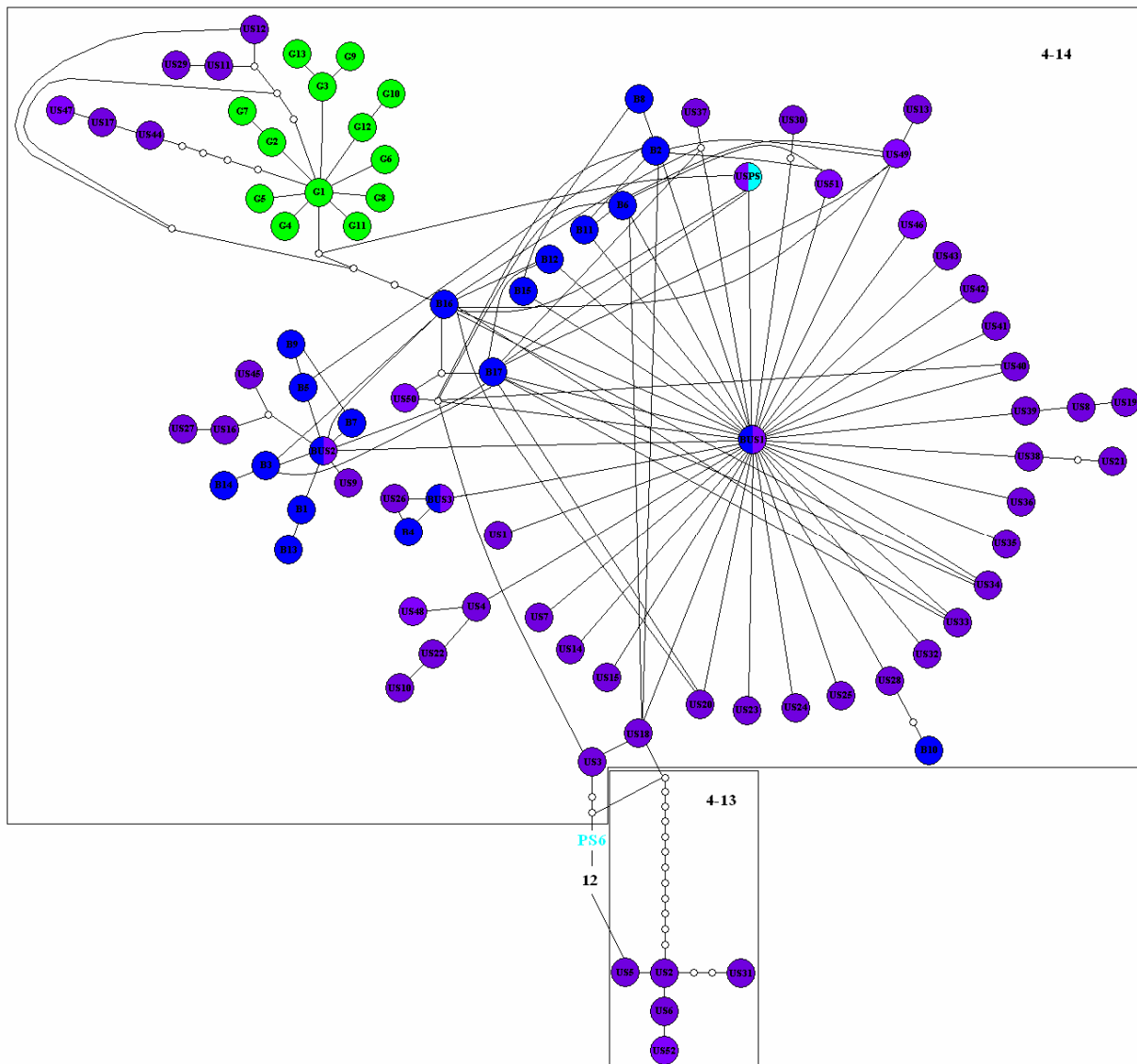


Figure 16a continued.

(vi)

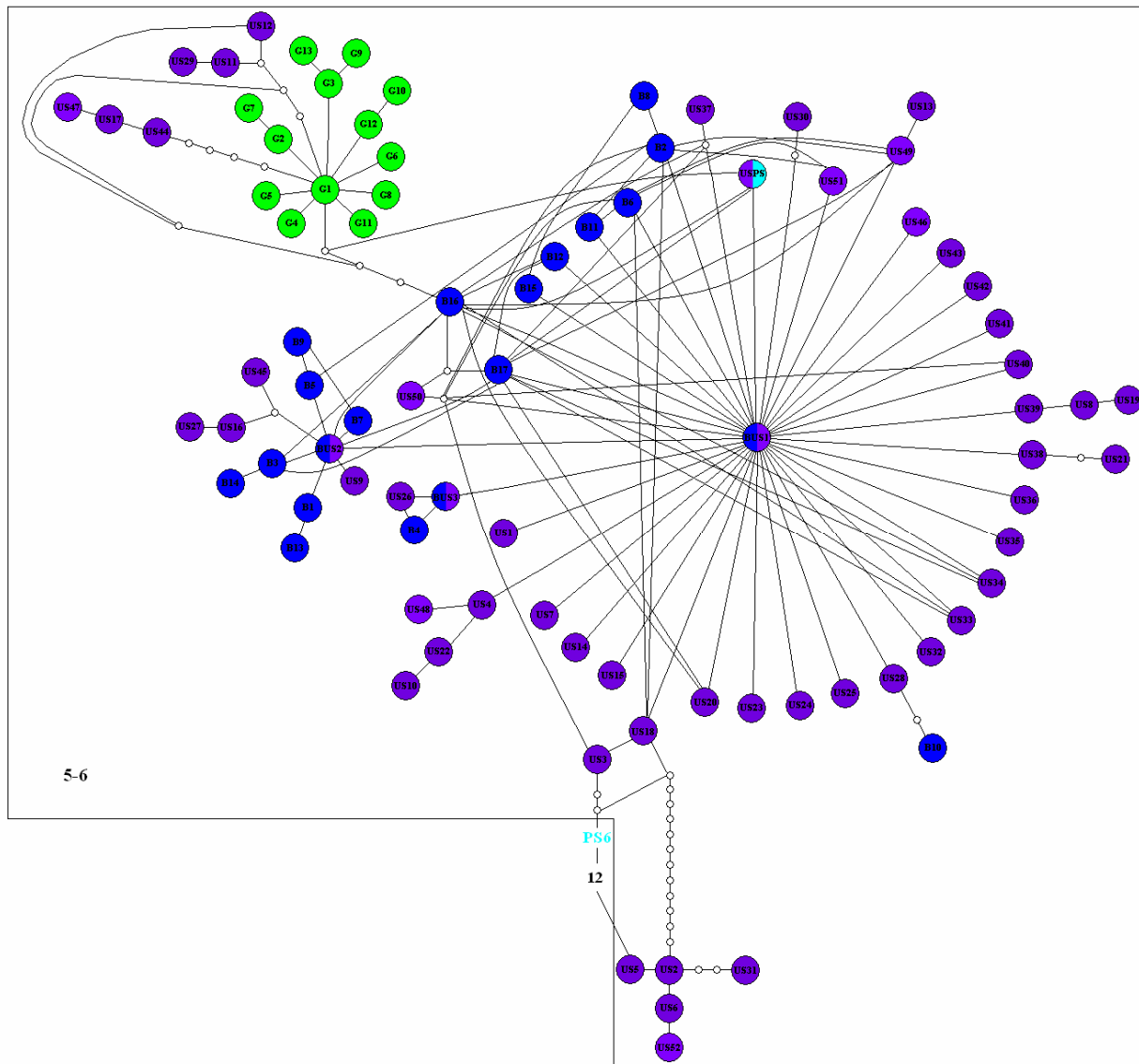
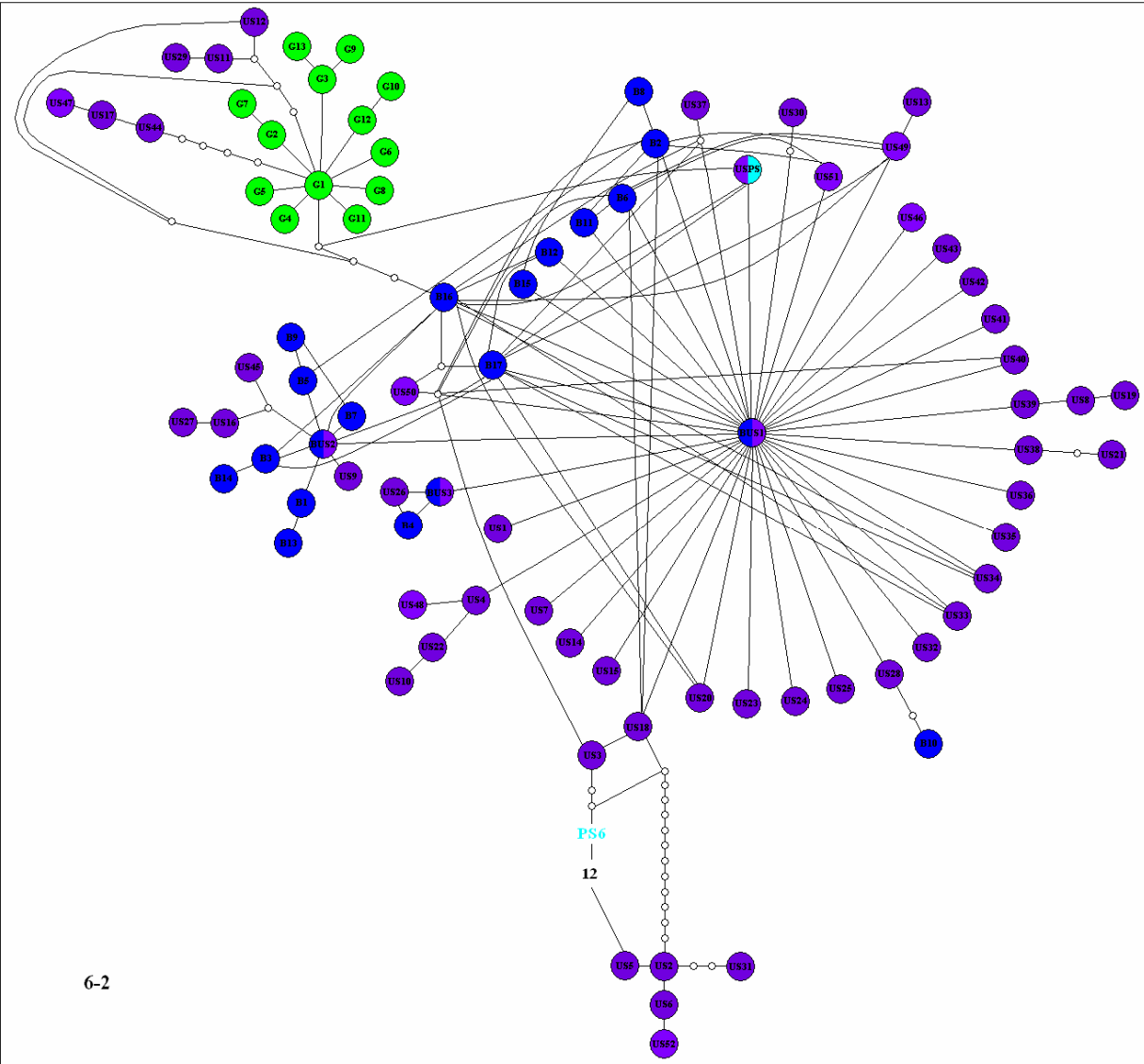


Figure 16a continued.

(vii)



6-2

Figure 16a continued.

(viii)

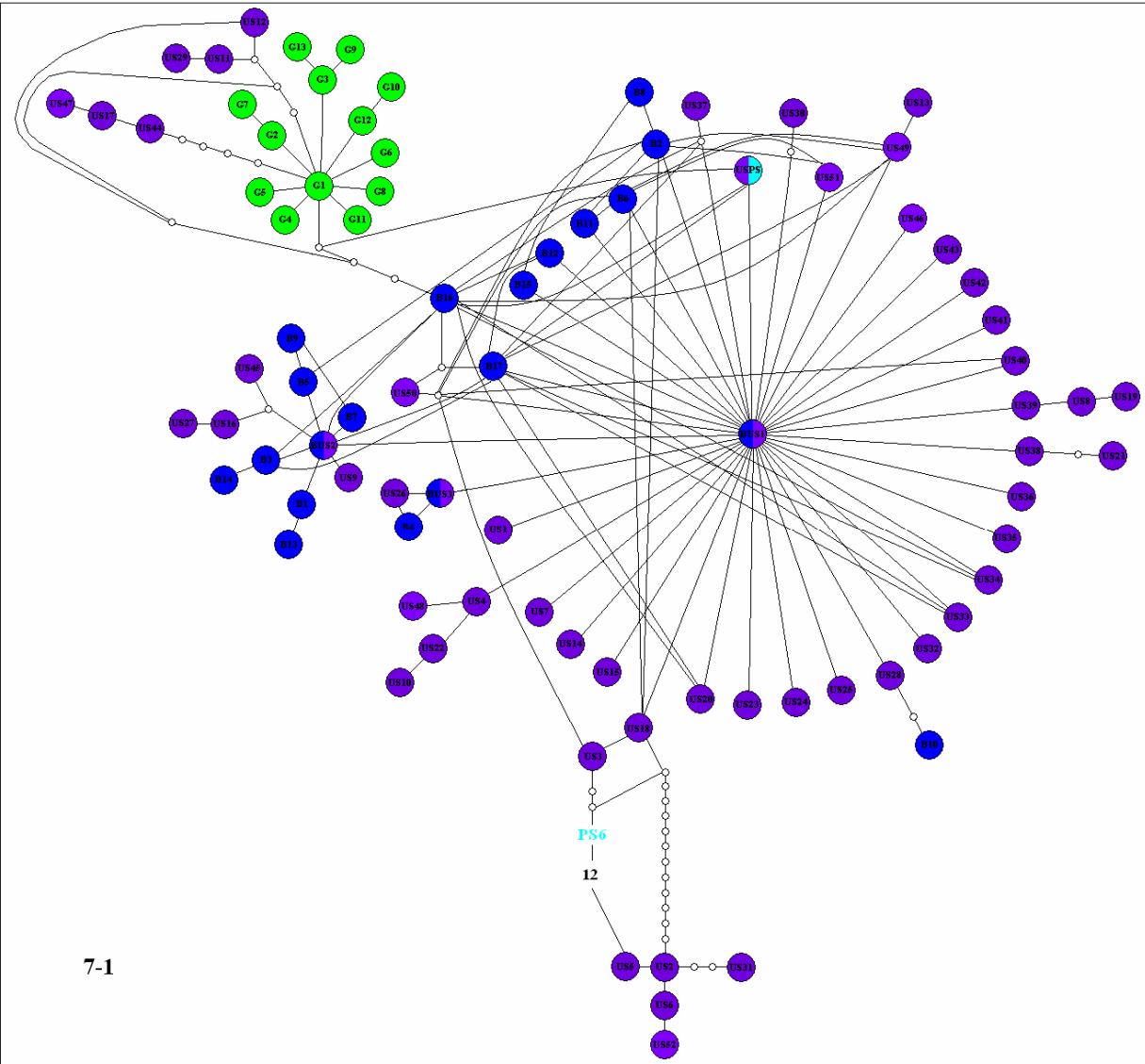


Figure 16b.

(i)

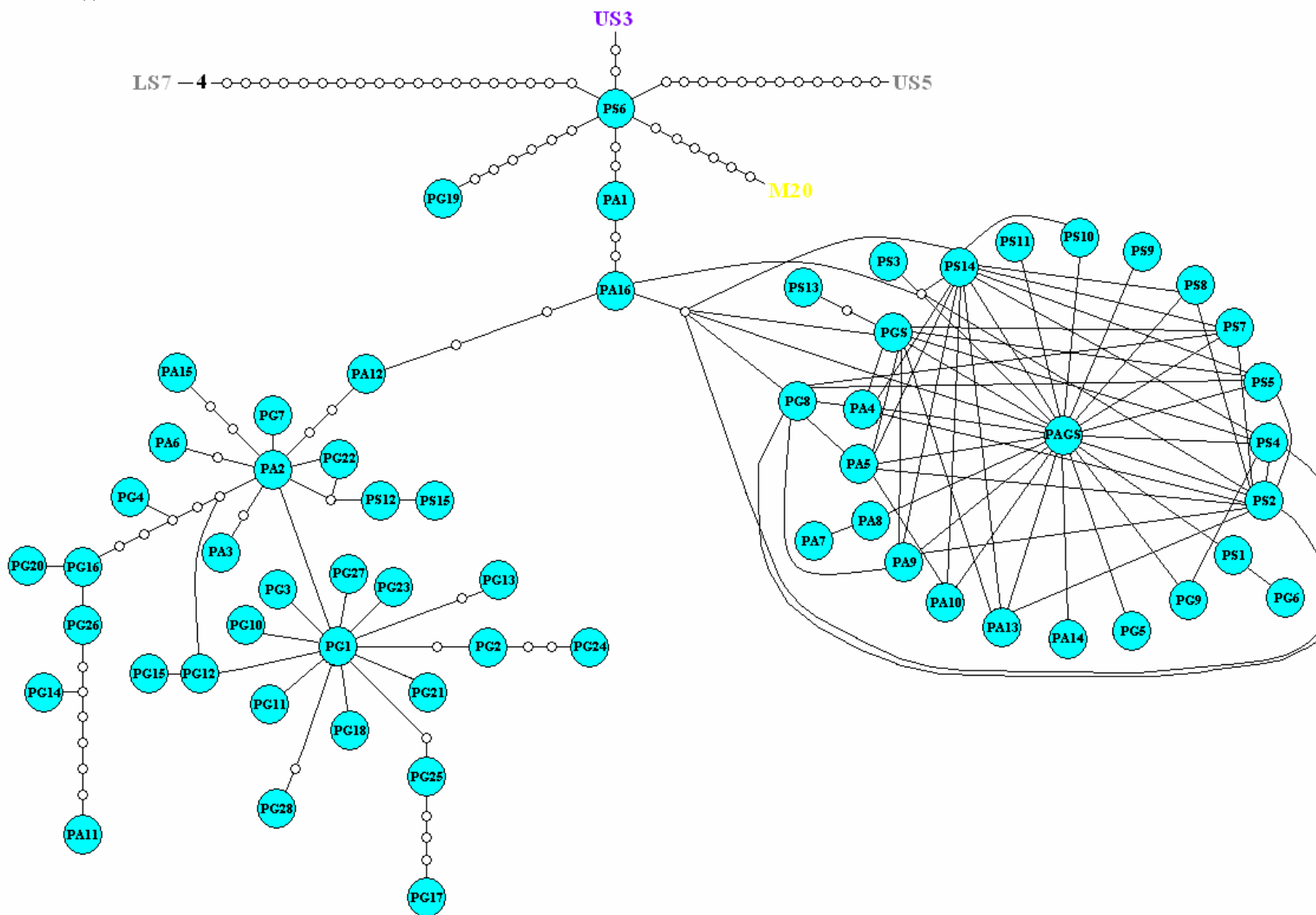


Figure 16b continued.

(ii)

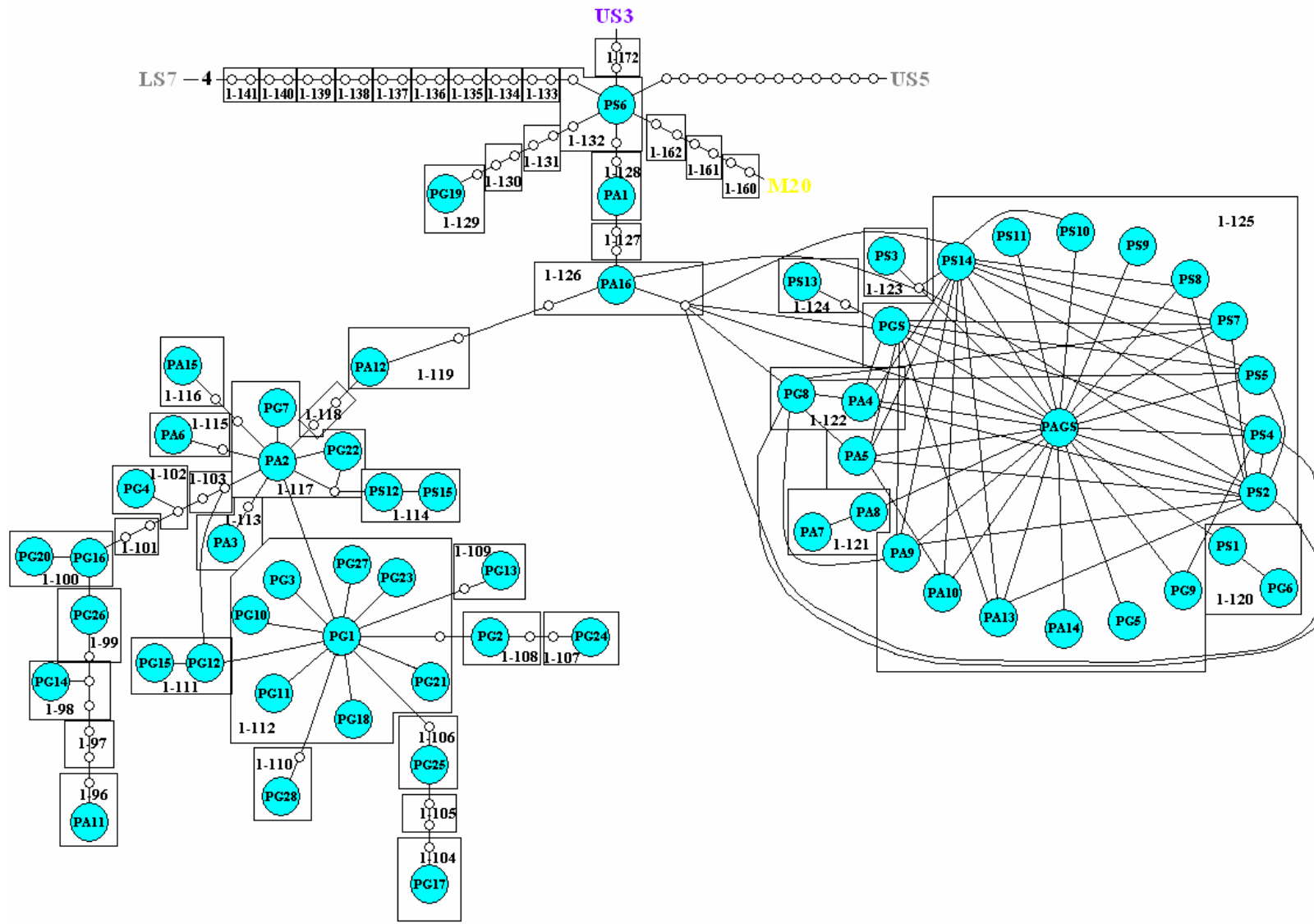


Figure 16b continued.

(iii)

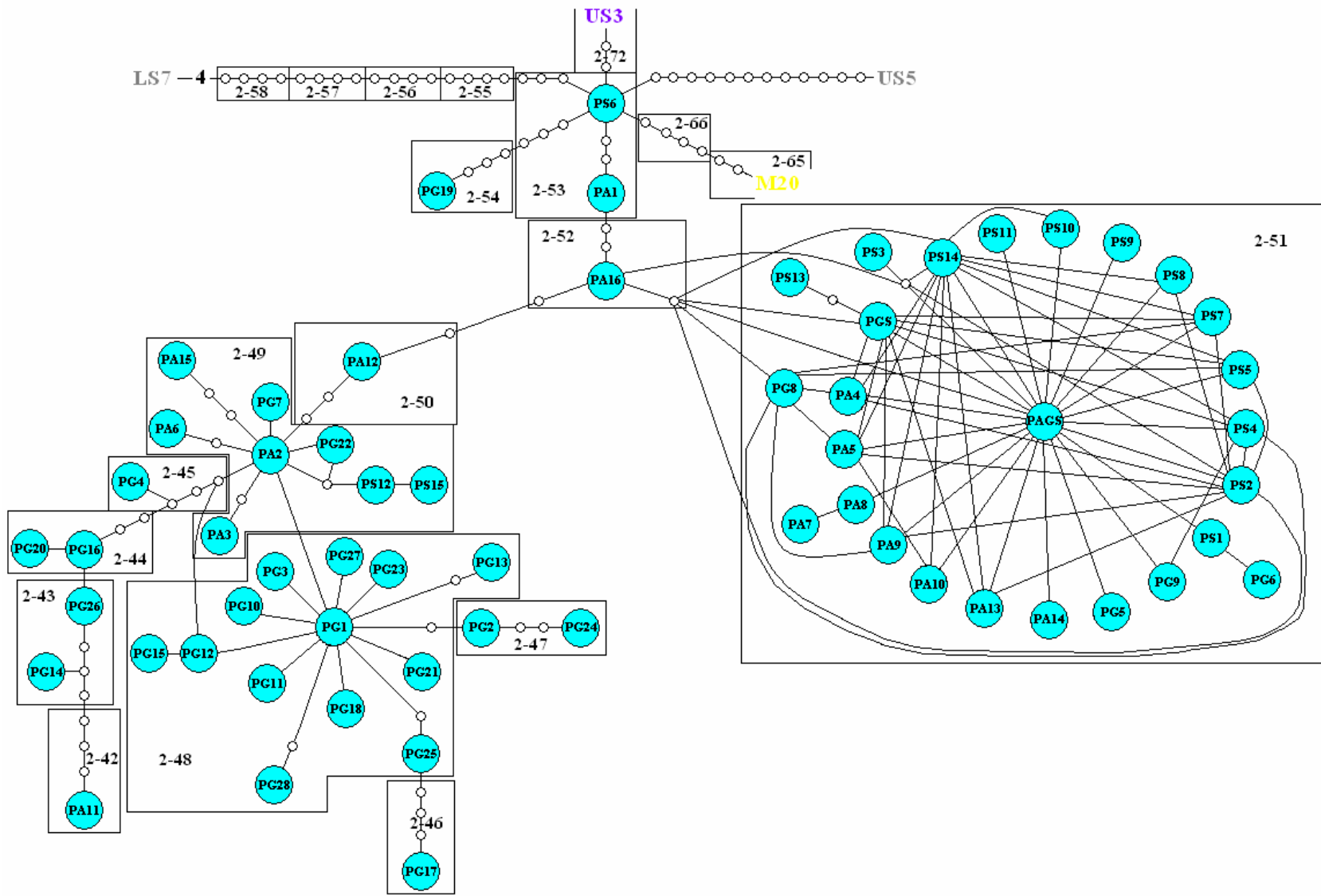


Figure 16b continued.

(iv)

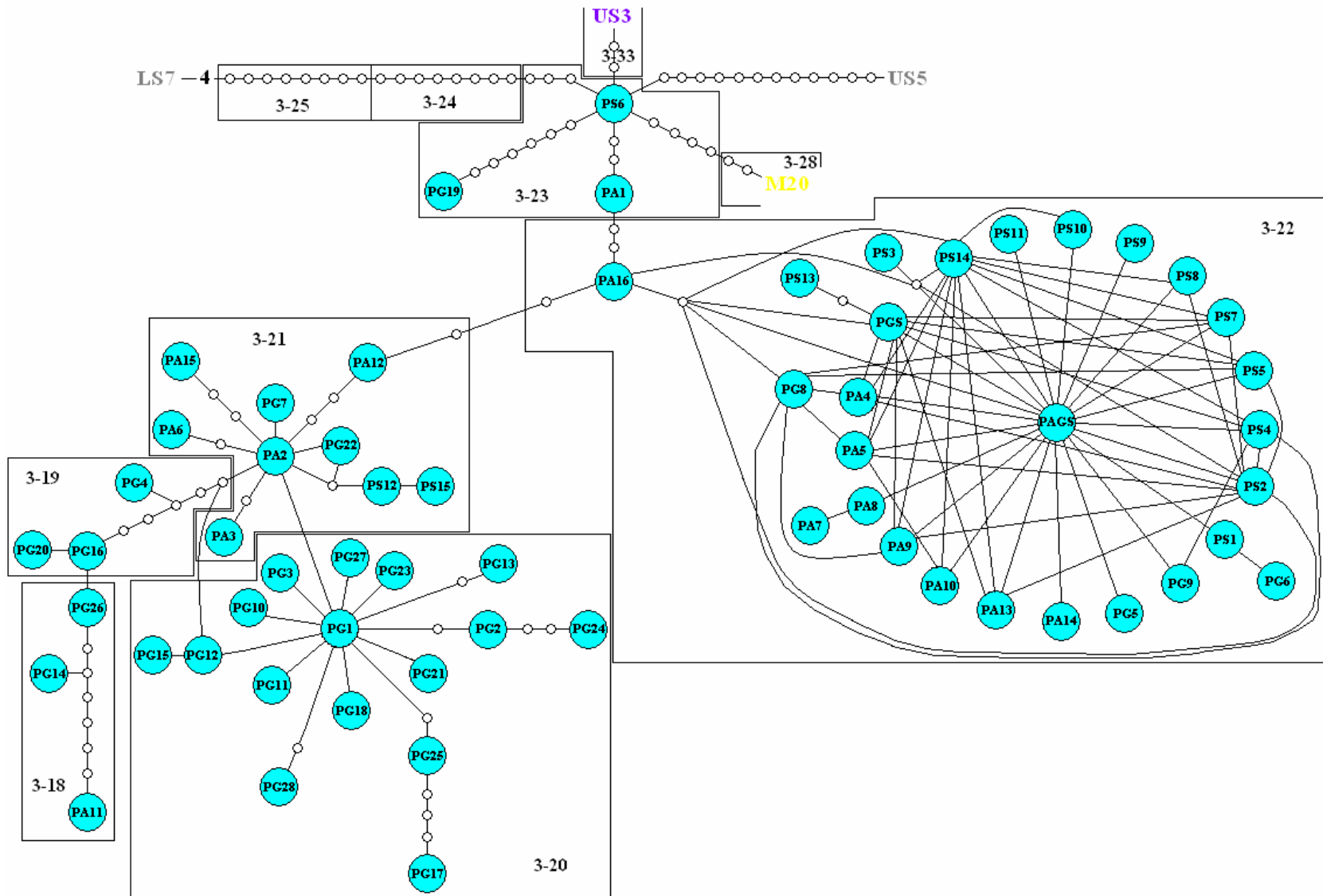


Figure 16b continued.

(v)

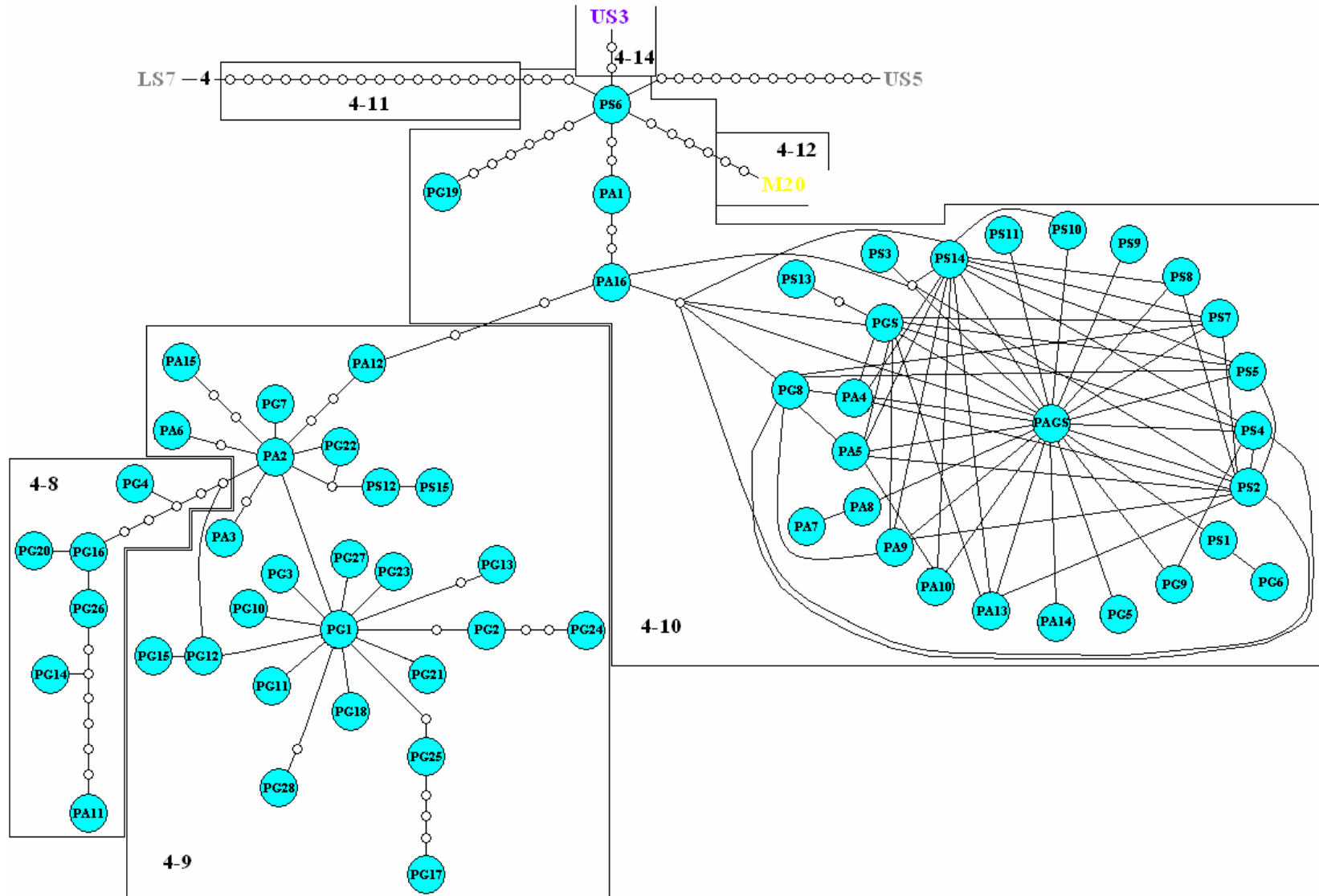


Figure 16b continued.

(vi)

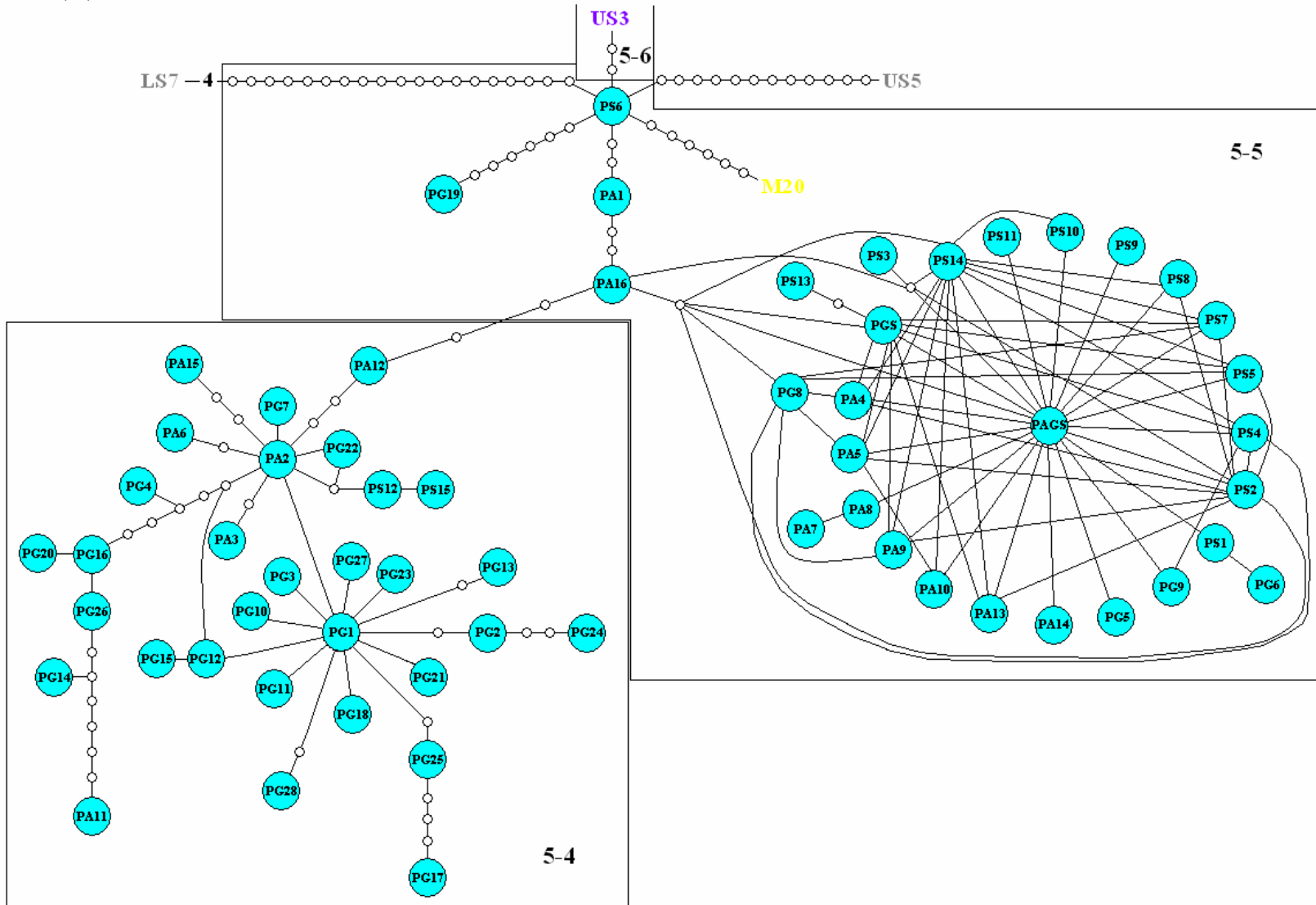
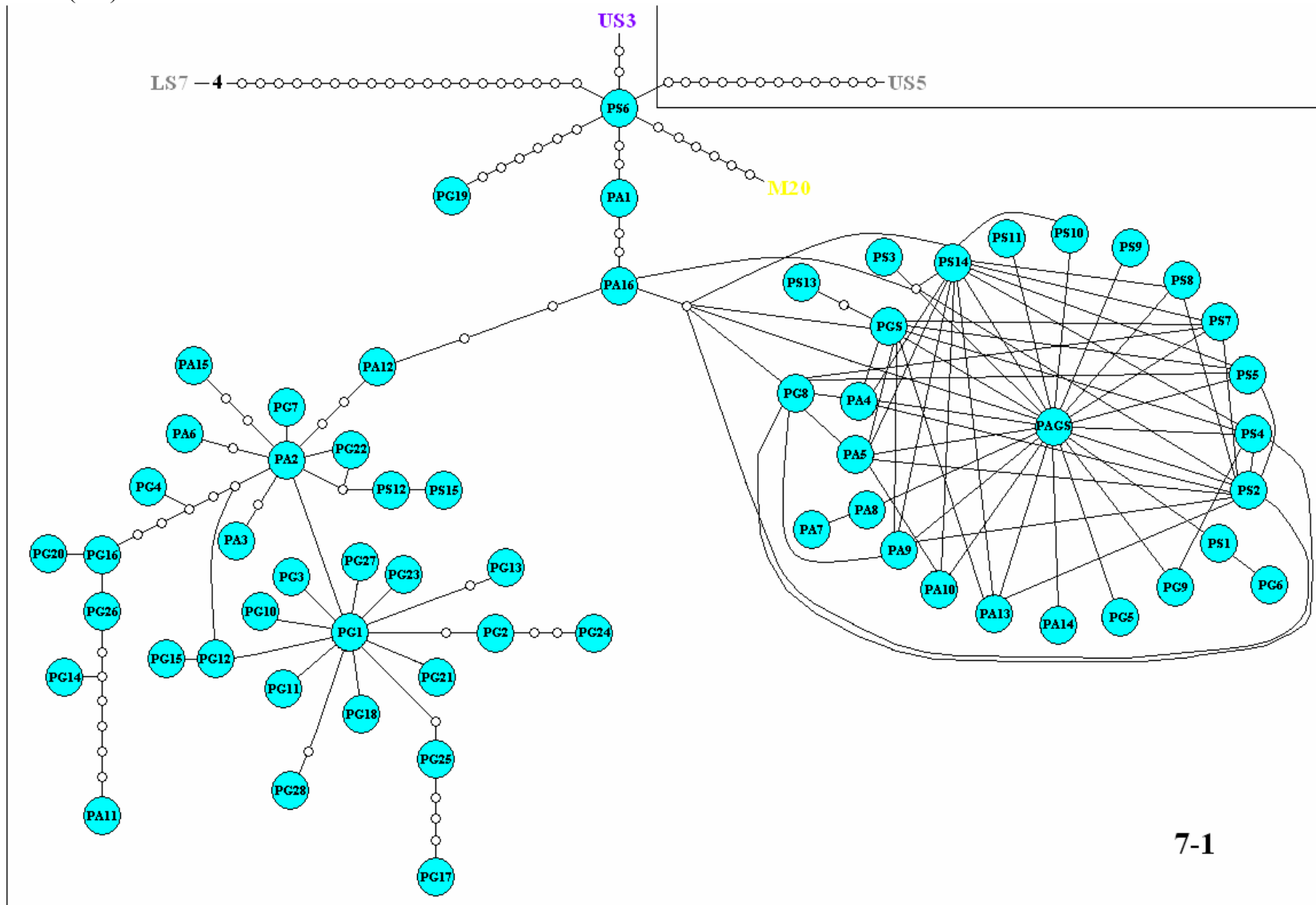


Figure 16b continued.

(viii)



7-1

Figure 16c.

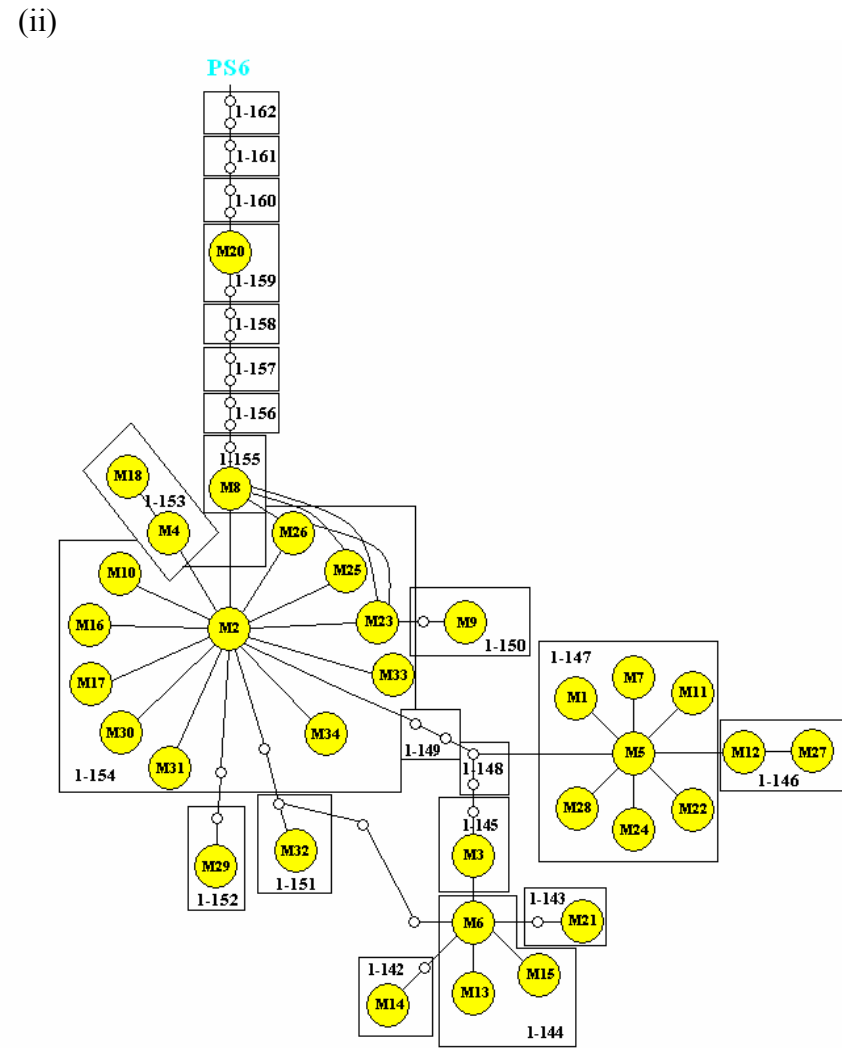
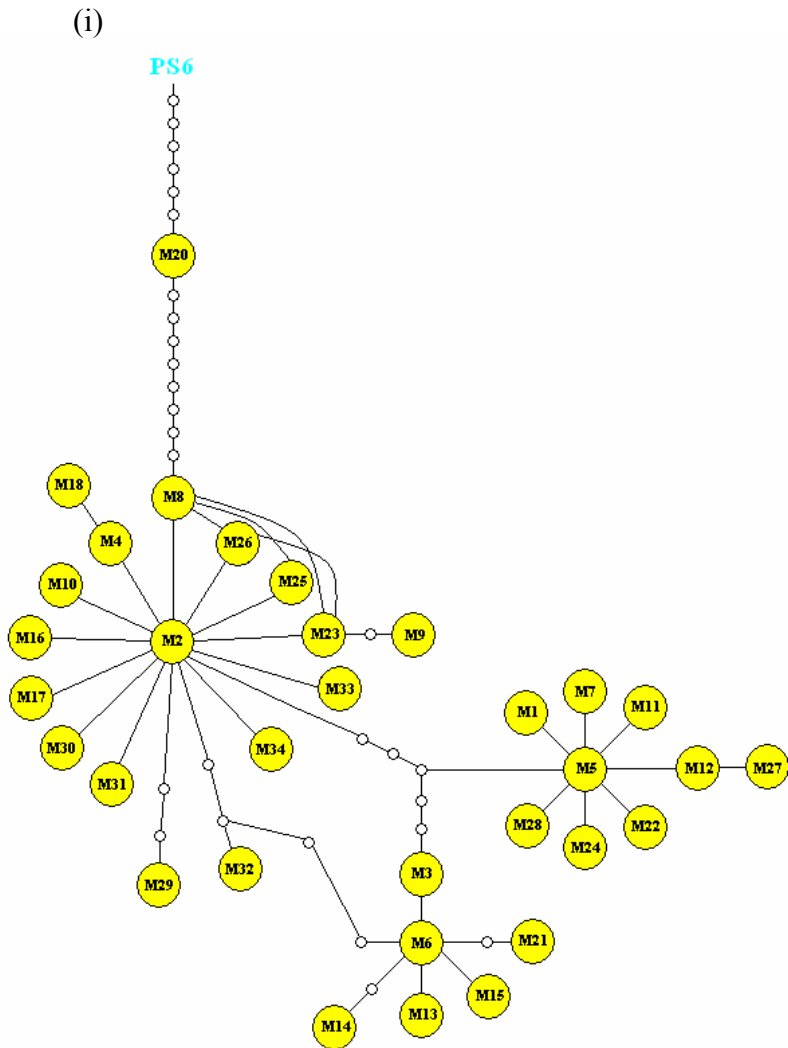
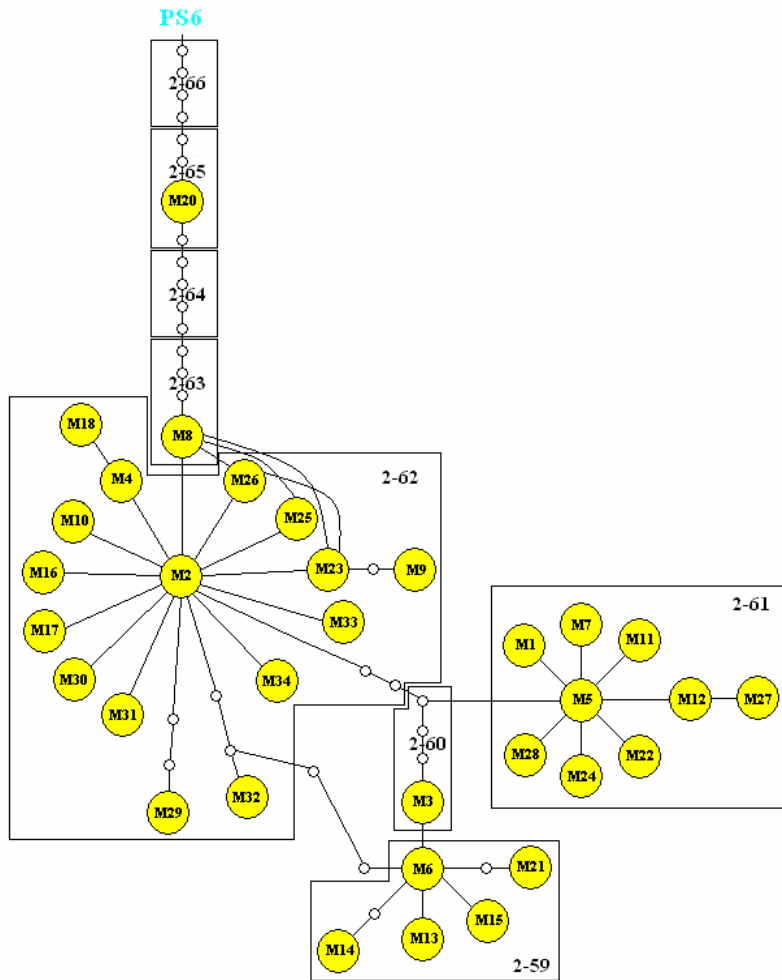


Figure 16c continued.

(iii)



(iv)

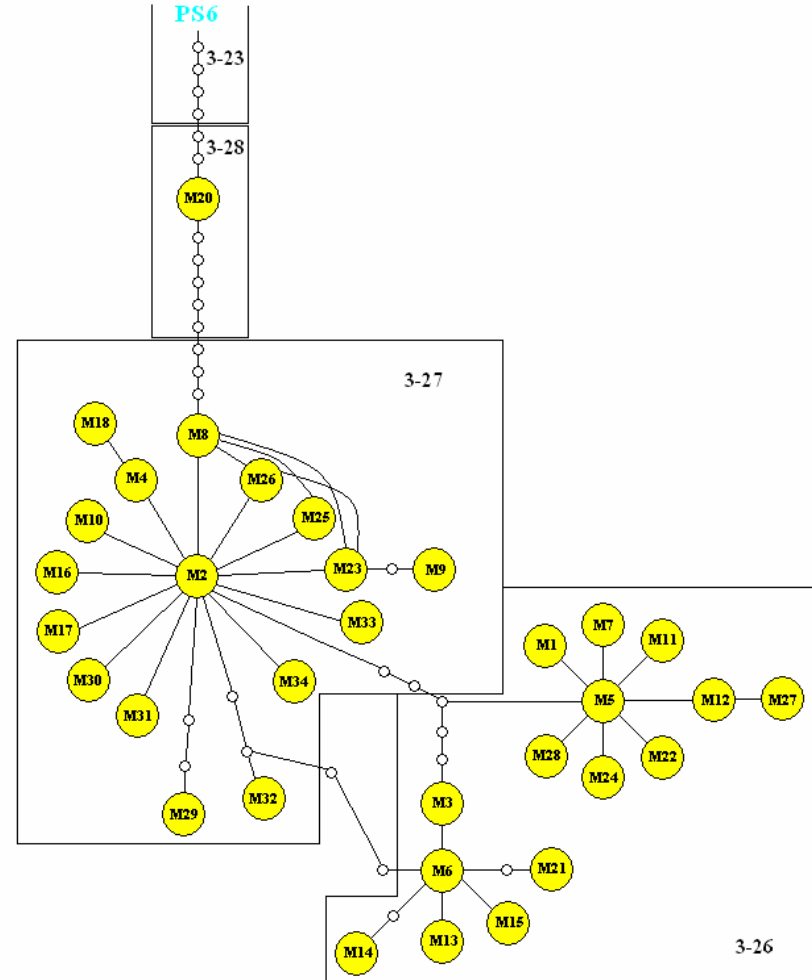
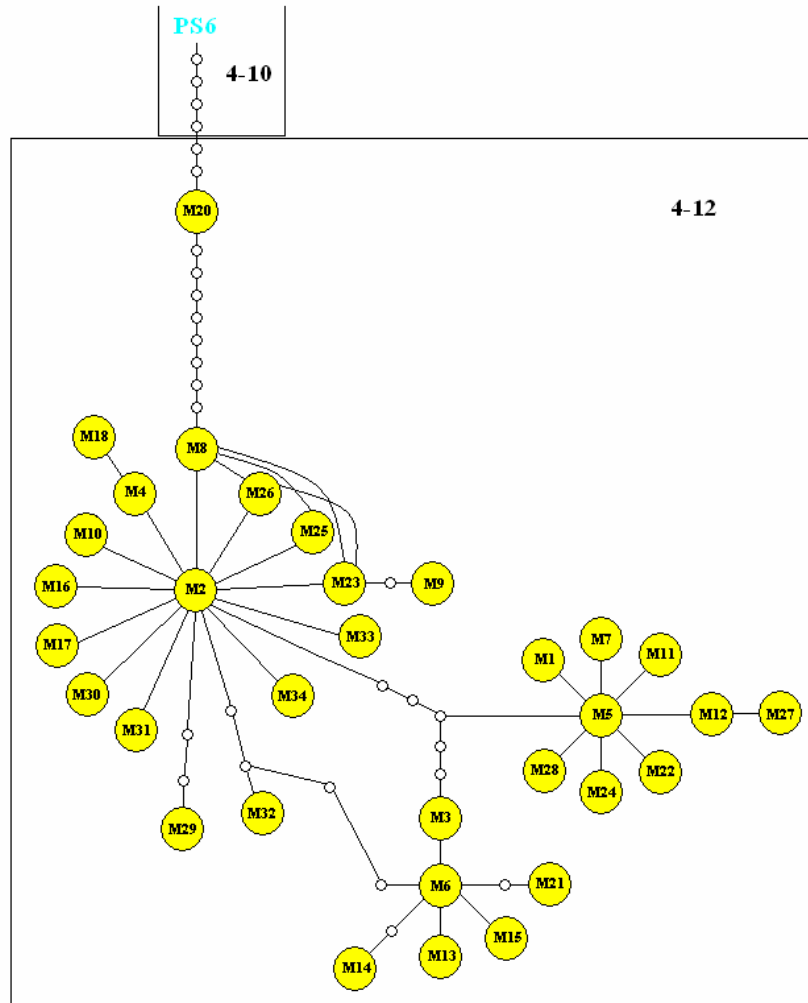


Figure 16c continued.

(v)



(vi)

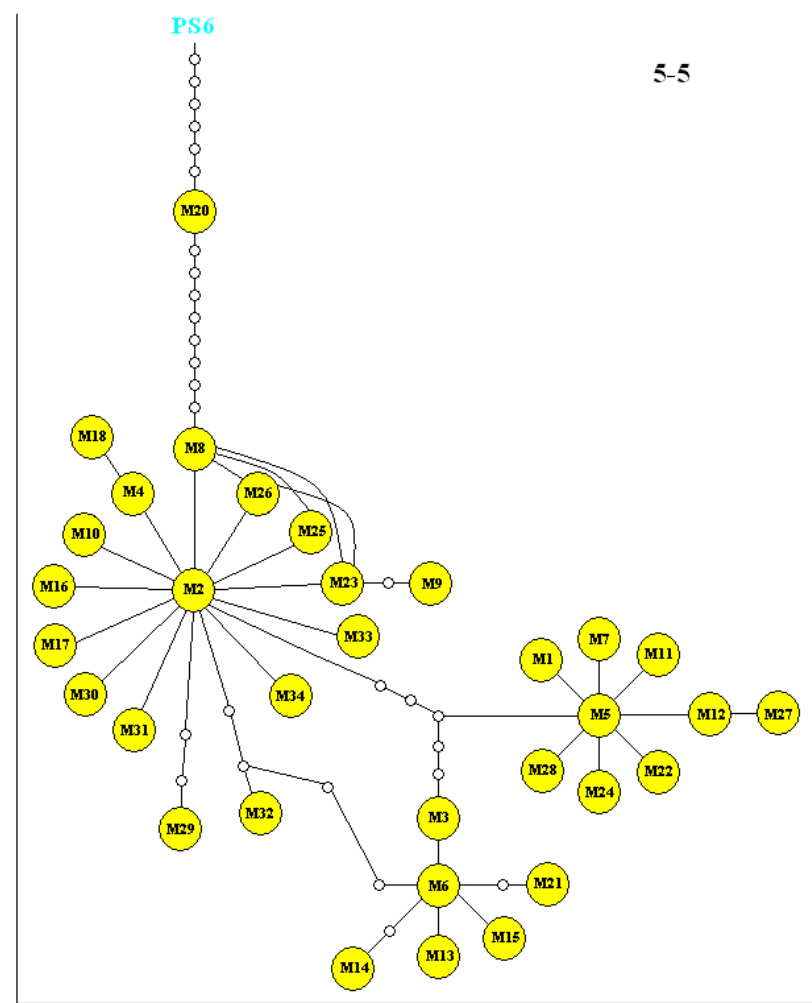


Figure 16c continued.

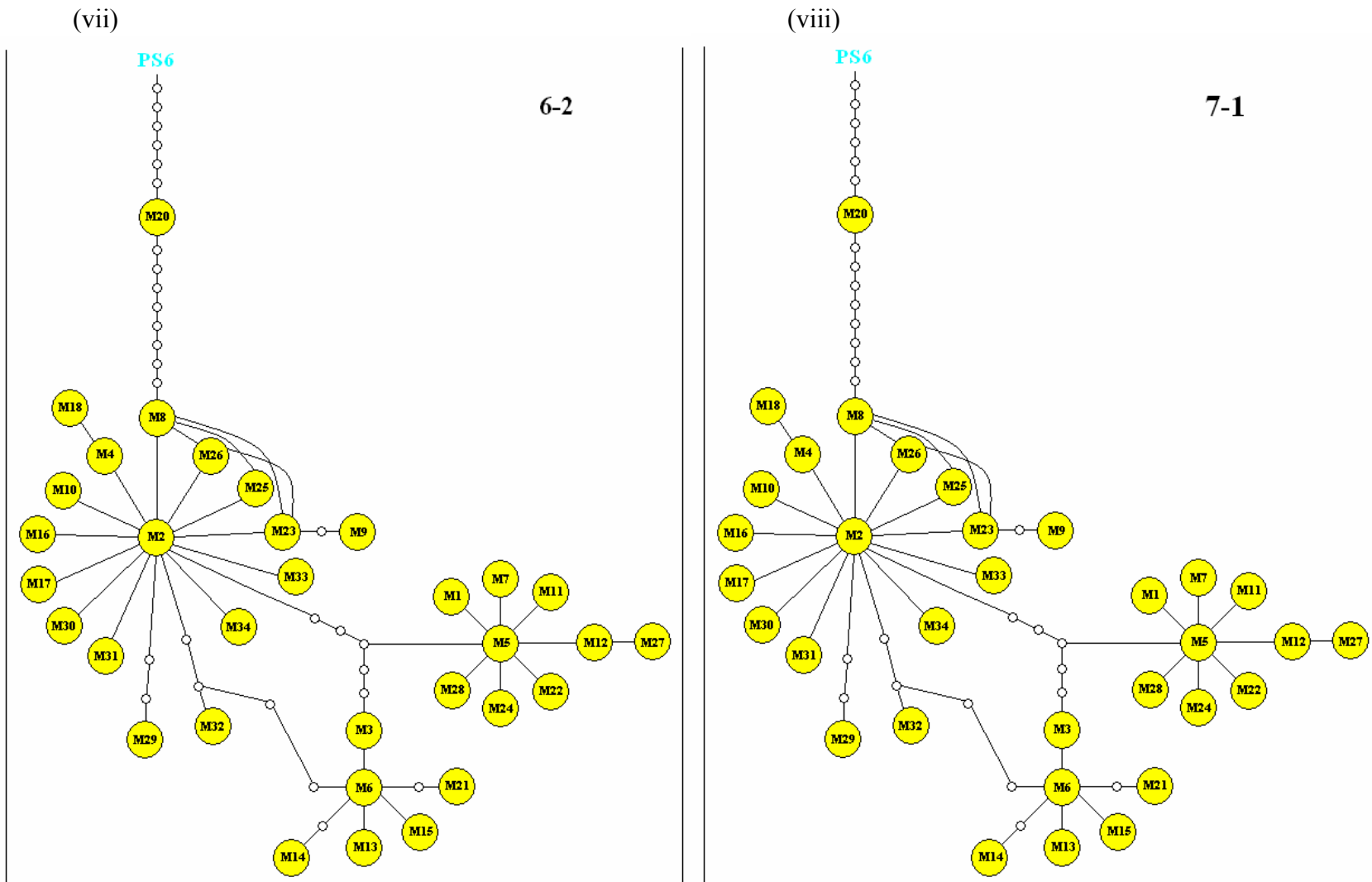


Figure 16d.

(i)

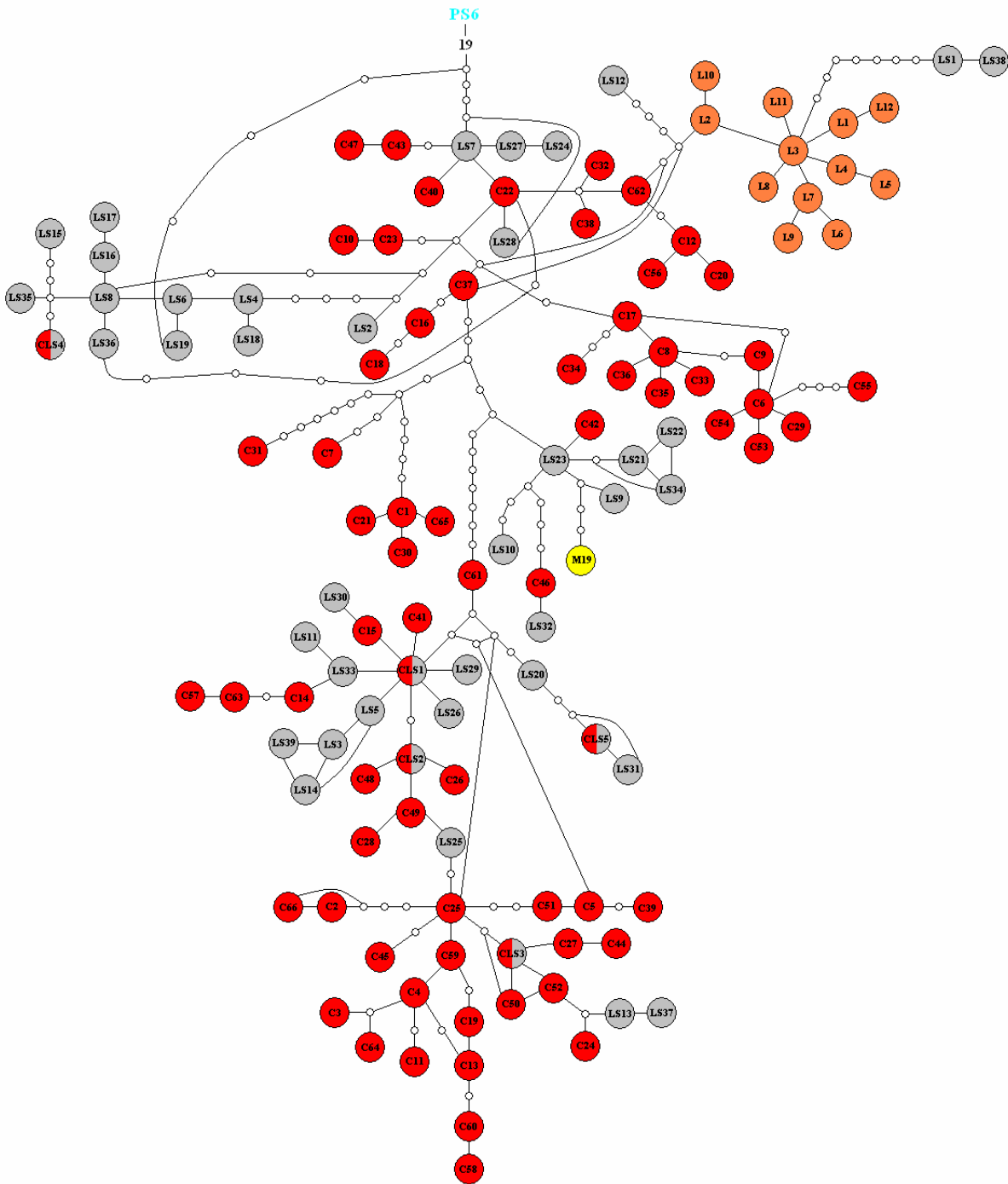


Figure 16d continued.

(ii)

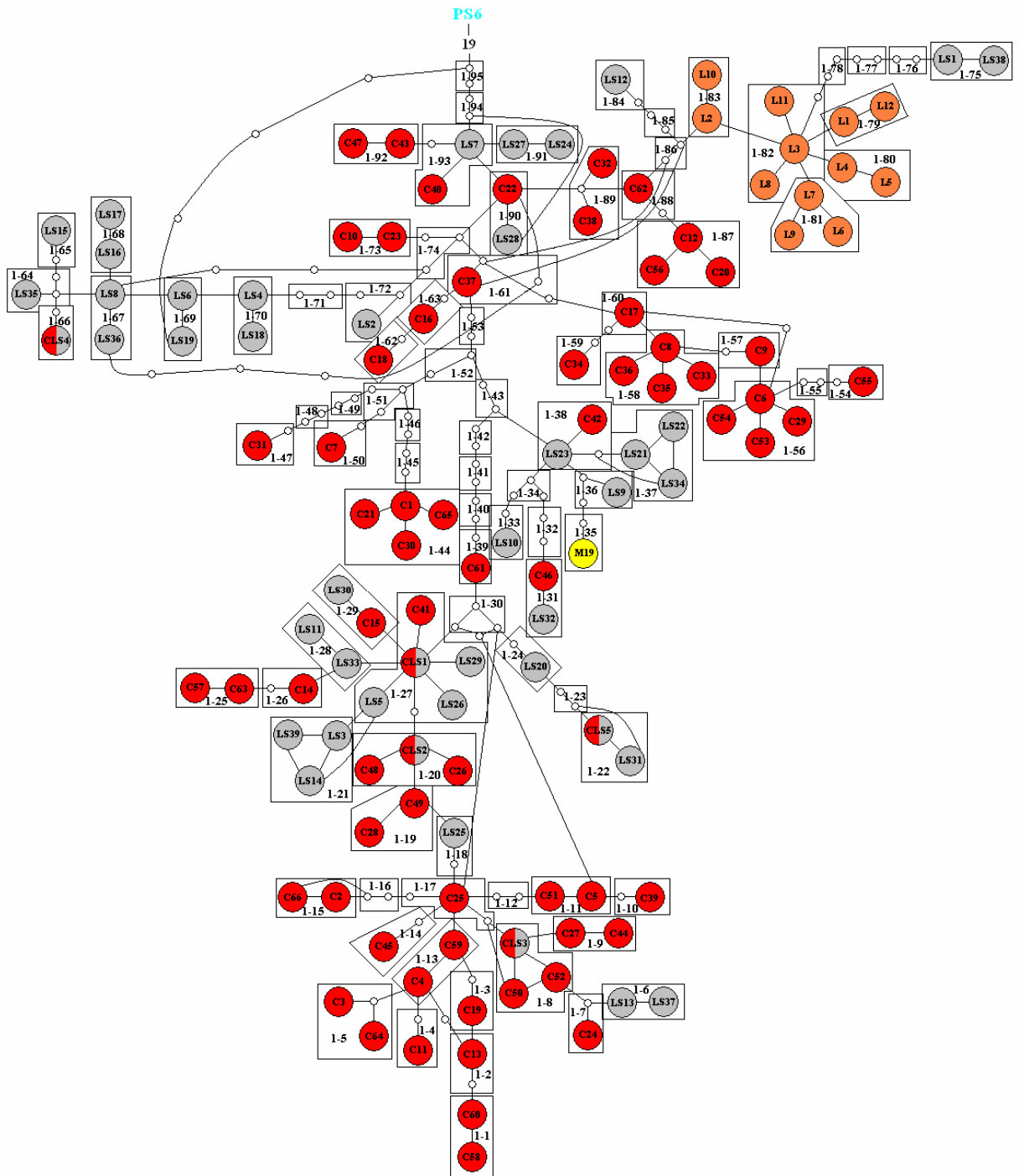


Figure 16d continued.

(iii)

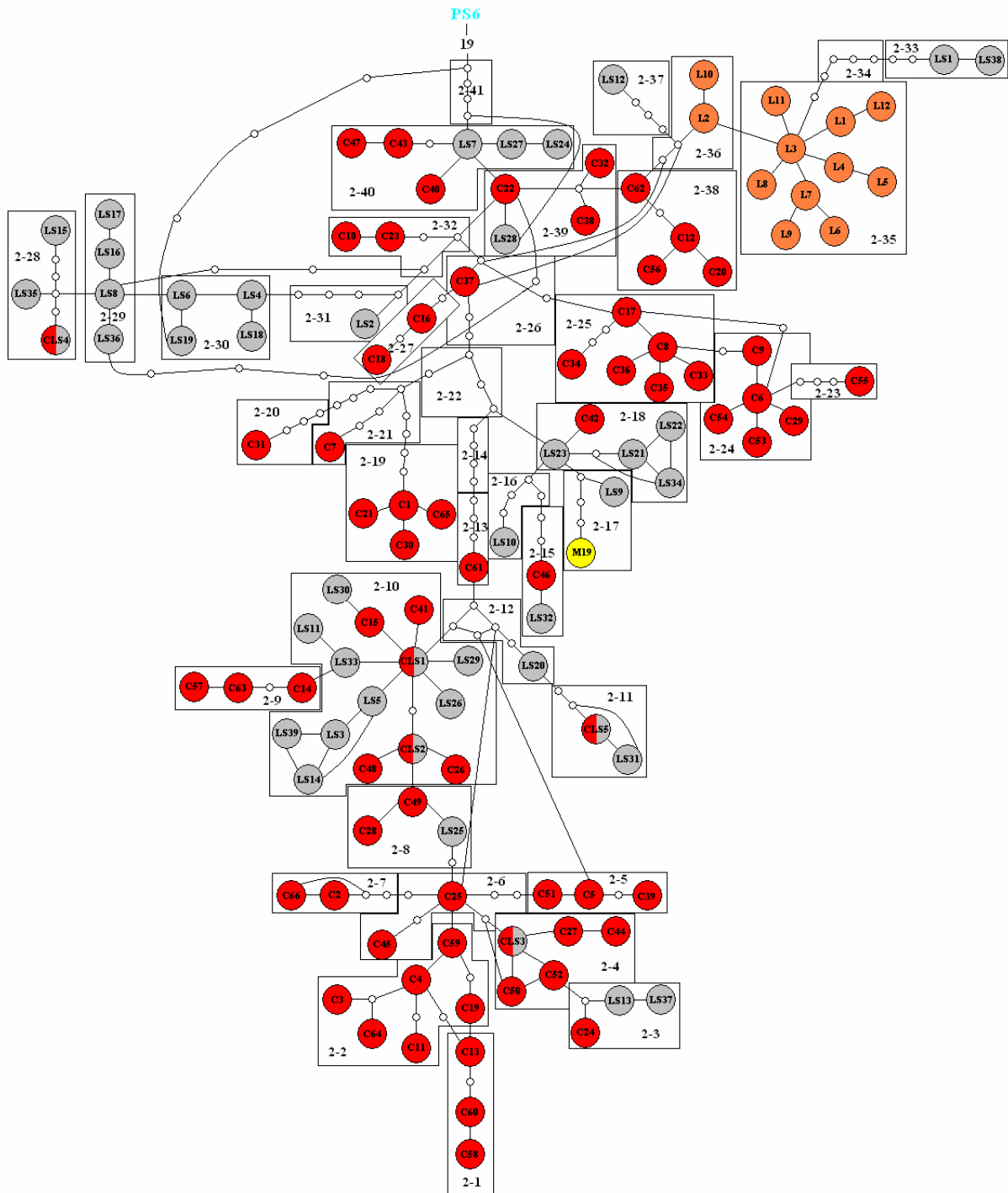


Figure 16d continued.

(iv)

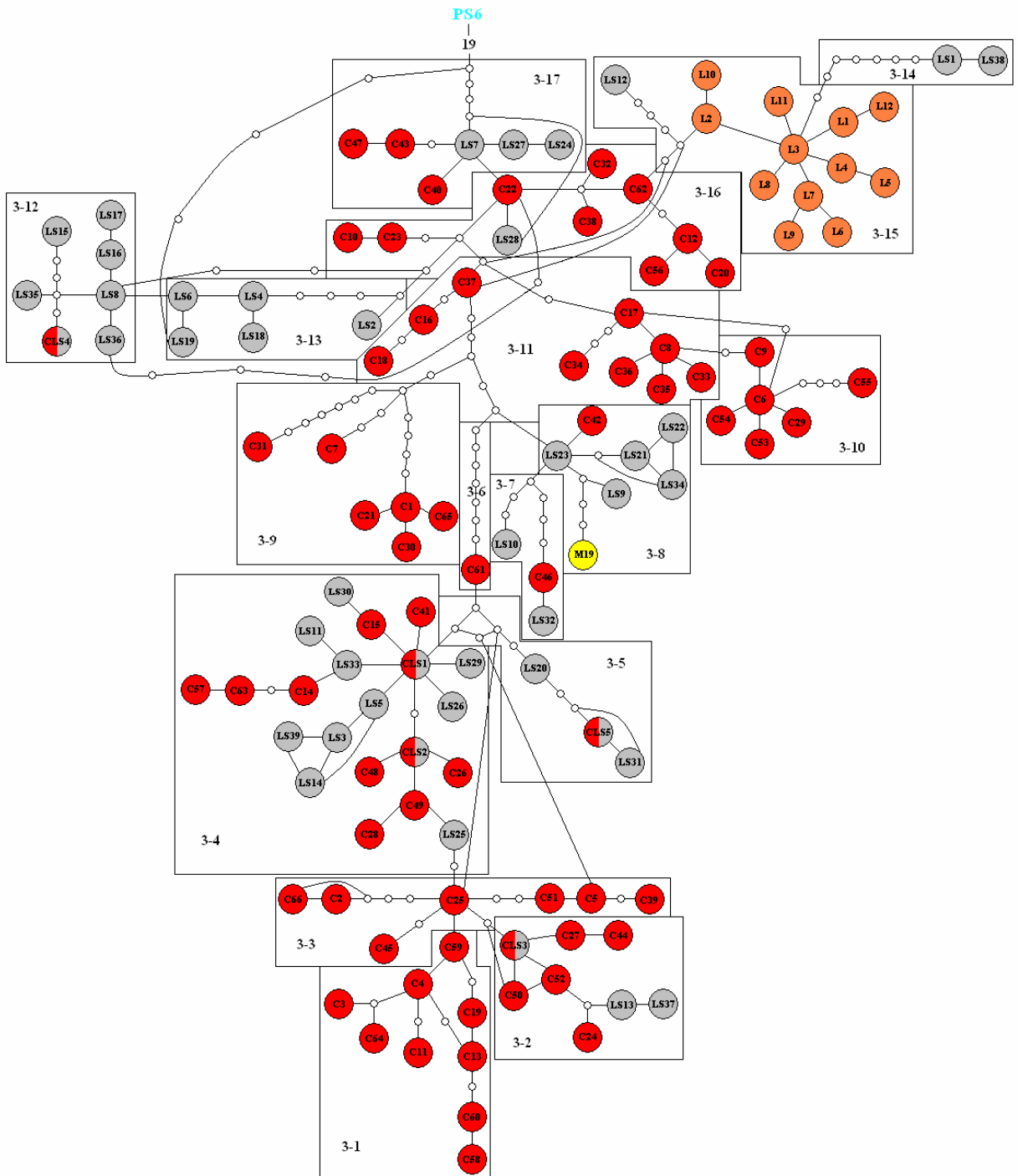


Figure 16d continued.

(v)

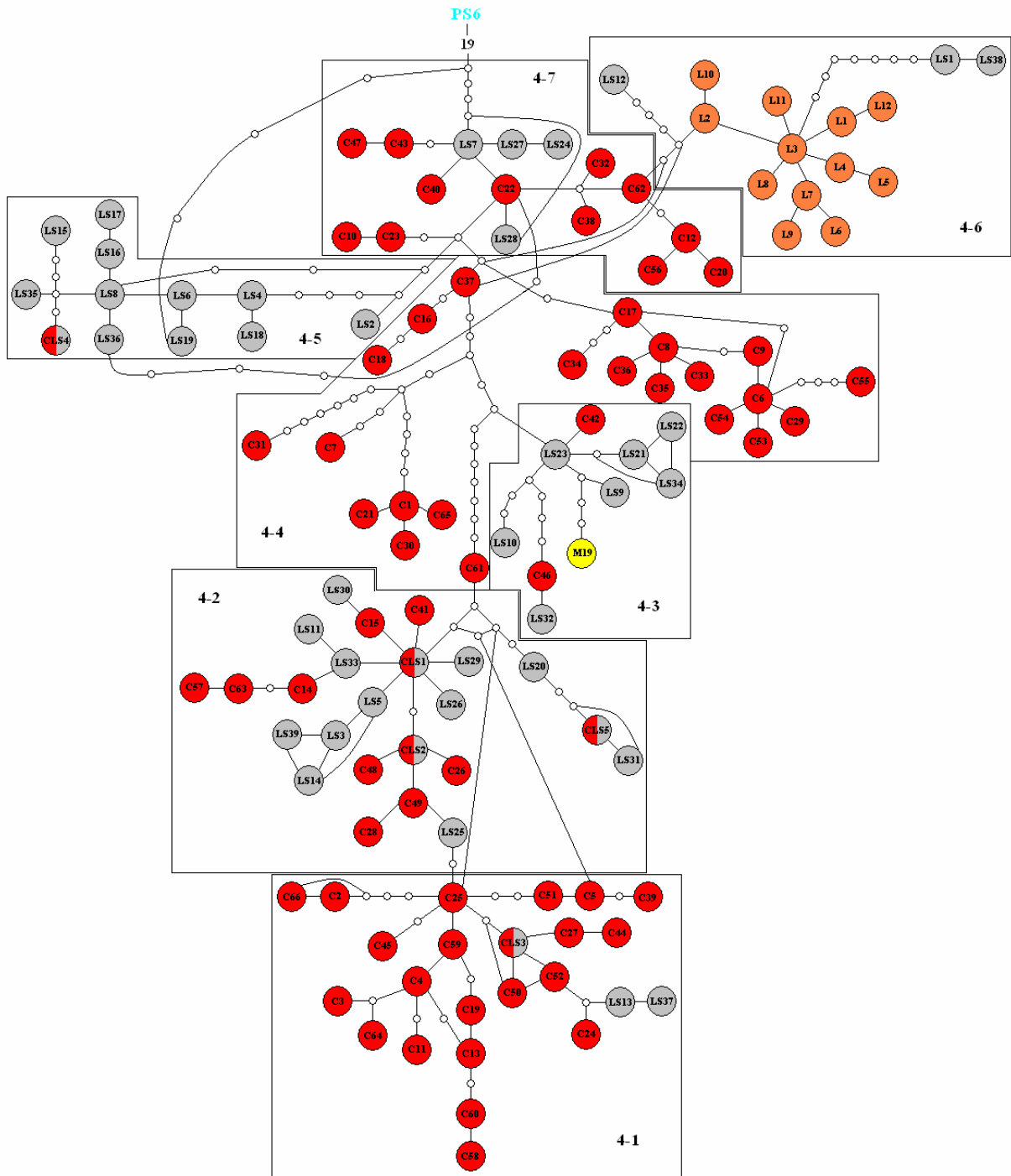


Figure 16d continued.

(vi)

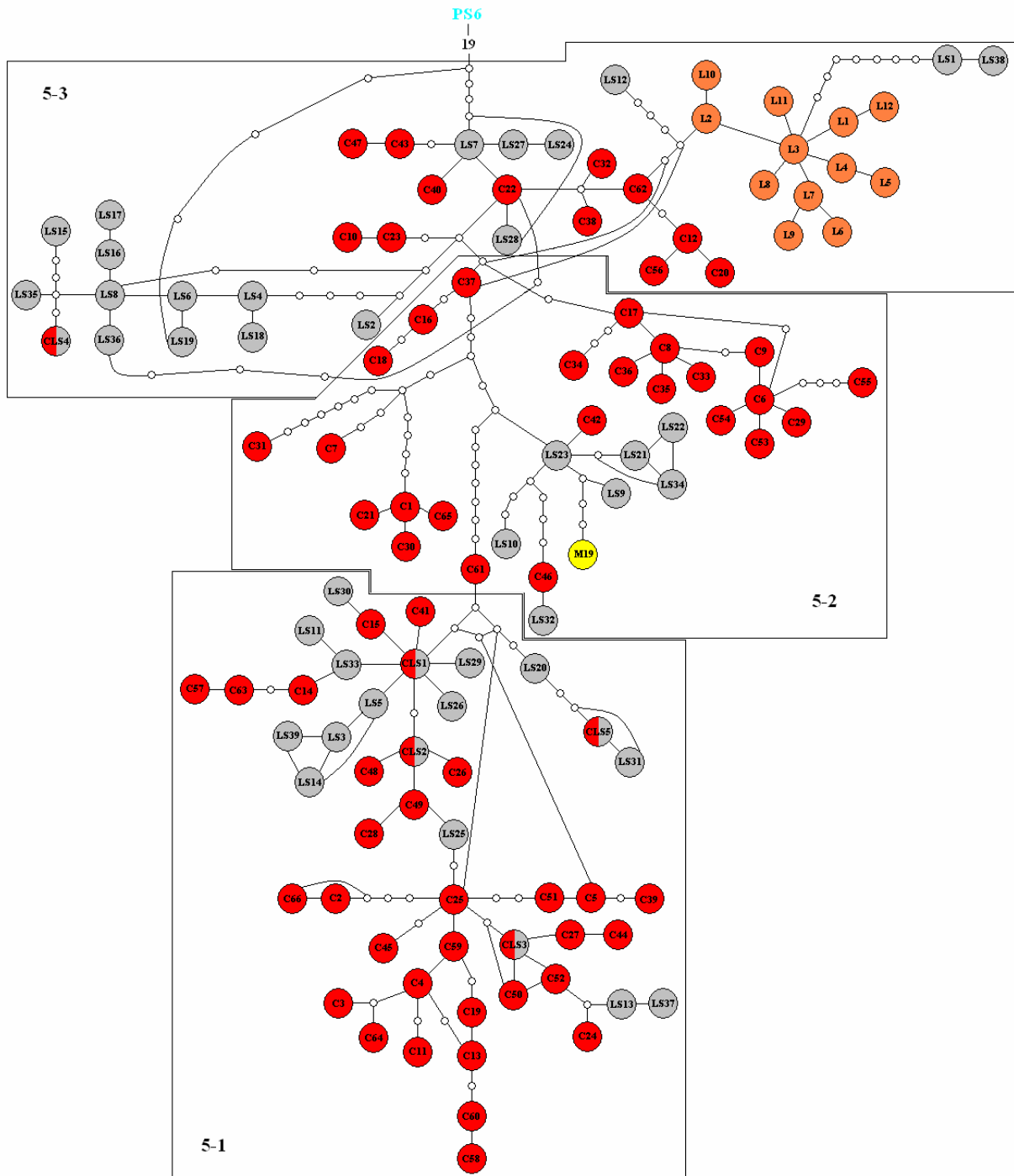
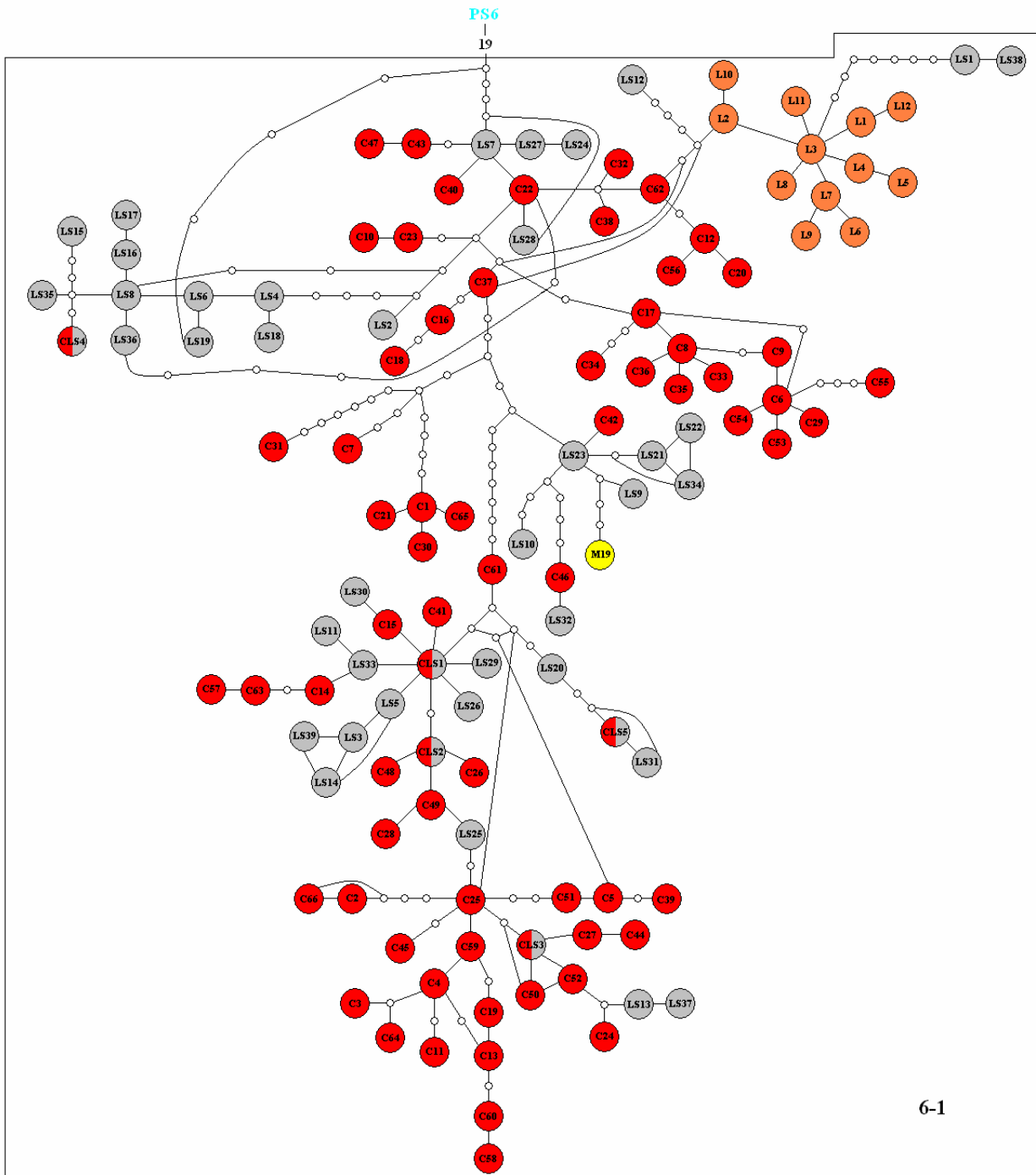


Figure 16d continued.

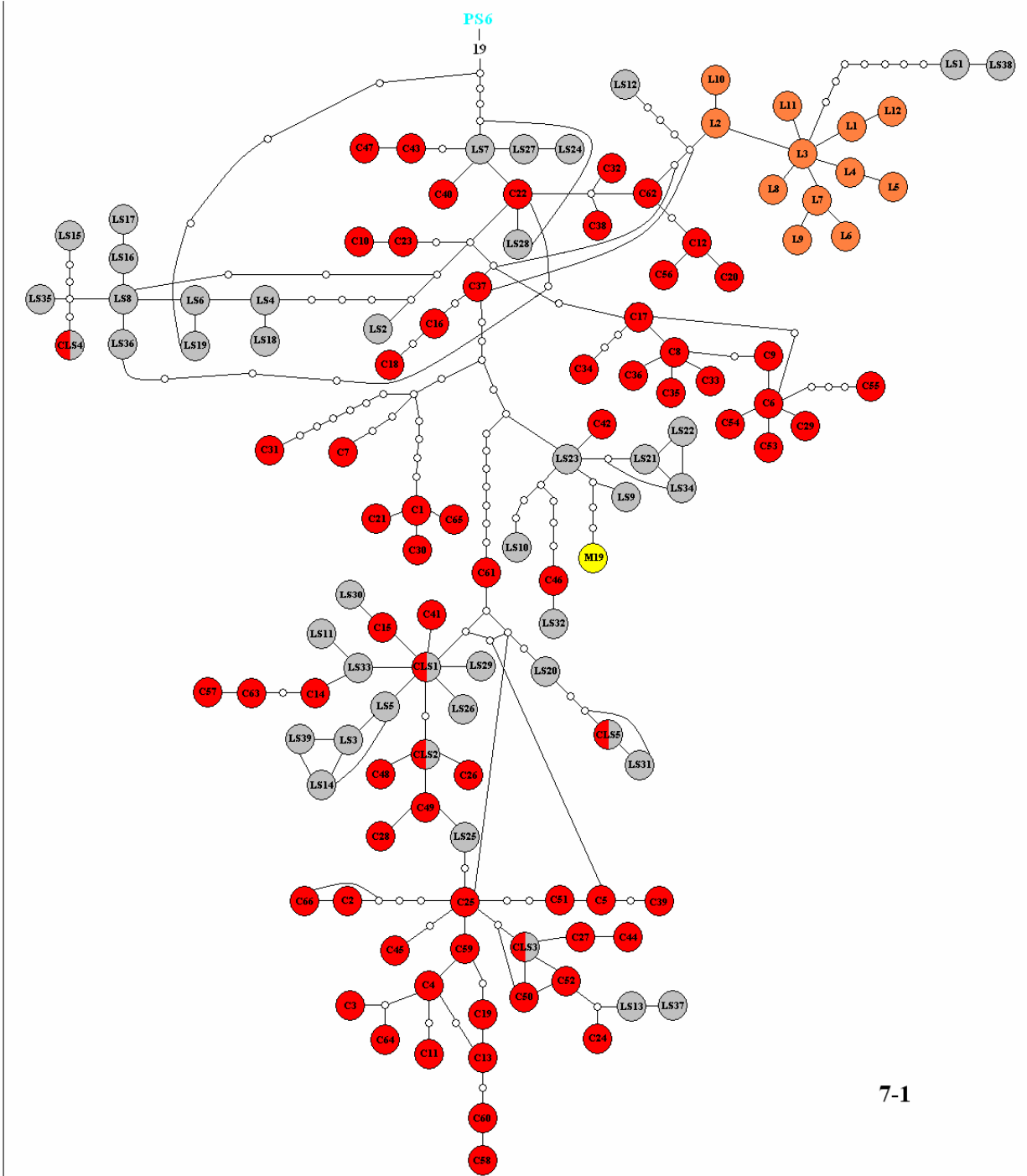
(vii)



6-1

Figure 16d continued.

(viii)



7-1