

Revised Final Report for Contracts 802706G121 and 802707G111

between the  
U.S. FISH AND WILDLIFE SERVICE  
and  
THE CSU, CHICO RESEARCH FOUNDATION

**Genetic Analysis of Giant Garter Snake (*Thamnophis gigas*) populations in  
the San Joaquin and Sacramento Valleys**

Tag N. Engstrom

Department of Biological Sciences California State University, Chico. Chico CA, 95929-0515  
tengstrom@csuchico.edu tel. 530-898-6748



## TABLE OF CONTENTS

Introduction.....	3
Species Background --- .....	3
<i>Project Background</i> --- .....	5
<i>Project goals and objectives</i> --- .....	6
<i>Project area description</i> --- .....	6
Sacramento Valley .....	<b>Error! Bookmark not defined.</b>
Delta.....	10
San Joaquin Valley .....	<b>Error! Bookmark not defined.</b>
Methods.....	13
Results.....	17
Discussion.....	19
Summary and Conclusions .....	22
Summary of Expenditures.....	23
References Cited .....	23
Personal Communications .....	26
Figures and Tables .....	27
Appendix A:.....	37
Appendix B:.....	40
Appendix C:.....	47

**ABSTRACT** This report summarizes work initiated in 2006 under contract 802706G121 between the U.S. Fish and Wildlife Service and The CSU Chico Research Foundation and continued in 2007 under contract 802707G111 with funding from the US Bureau of Reclamation and US Fish and Wildlife Service’s CVPIA Habitat Restoration Program. DNA sequence data from the ND4 region of the mitochondrial genome was collected from total of 466 giant garter snakes from throughout the range of the species including Butte Basin, Colusa Basin, Sutter Basin, Yolo Basin and the American Basin in the Sacramento Valley, Badger Creek and White Slough in the Delta, and Los Banos and Mendota in the San Joaquin Valley. Microsatellite data from a subset of 96 of these snakes were also collected. A total of 13 mtDNA haplotypes were identified from these sequences including three novel haplotypes (L from the San Joaquin Valley; K, and M from the Delta populations). Overall molecular diversity in this species is low, but highly geographically structured. Populations in the Delta (Badger Creek and White Slough) were characterized by several unique haplotypes found nowhere else and showed no mtDNA haplotype overlap with other populations examined. The San Joaquin Valley, and Sacramento Valley share some overlapping haplotypes but each valley has unique haplotypes and each is distinguishable by differences in haplotype frequencies. AMOVA and traditional Fst approaches showed that populations of giant garter snakes have low genetic exchange across the species range, among adjacent watersheds, and among populations within watersheds even across fairly small geographic scales within the American/Natomas Basin. These genetic characteristics imply low demographic exchange among populations suggesting that when possible, populations of this species should be managed as separate demographic units

## Introduction

### Species Background ---

The giant garter snake (*Thamnophis gigas*), is federally listed as a threatened species. This species is endemic to the Great Central Valley of California where it once ranged throughout the wetlands of California's Central Valley from Buena Vista Lake near Bakersfield in Kern County, north to the vicinity of Chico in Butte County (Hansen and Brode 1980). Described as among California's most aquatic garter snakes (Fitch 1940), giant garter snakes are associated with low-gradient streams, wetlands and marshes and regions supporting rice agriculture. Due mainly to loss or degradation of aquatic habitat resulting from agricultural and urban development, giant garter snakes populations have declined or been extirpated throughout much of its former range. The current known distribution of giant garter snakes is patchy, and extends from near Chico in Butte County, south to Mendota Wildlife Area in Fresno County. This loss of habitat, shrinking range and apparent population decline lead the California Department of Fish and Game to list giant garter snakes as rare in 1971. The U.S Fish and Wildlife Service officially listed the species as Threatened on October 20, 1993 (58 FR 54053) and it is classified as vulnerable by the World Conservation Union (IUCN) (Baillie 1996). Giant garter snakes are strongly associated with aquatic habitats, but typically over-winter in burrows and crevices in terrestrial habitat near to their active-season foraging habitat (Hansen 2003, 2006). Individuals have been noted on burrows as far as 50 meters from marsh edges during the active season, and as far as 250 meters from the edge of wetland habitats while over-wintering, presumably to reach hibernacula that are located above the annual high water mark (Hansen 1986, Wylie *et al.* 1997, USFWS 1999). Habitats occupied by giant garter snakes typically contain permanent or seasonal water, mud bottoms, and vegetated dirt banks (Fitch 1940; Hansen and Brode 1980). Prior to the extensive reclamation projects in the Central Valley, the wetlands occupied by the snakes probably consisted of freshwater marshes and low-gradient streams. The present day aquatic habitats that giant garter snakes are associated with are characterized by the following features: 1) sufficient water during the snake's active season (typically early spring through mid-fall) to supply cover and food such as small fish and amphibians; 2) emergent herbaceous wetland vegetation such as cattails and bulrushes accompanied by vegetated banks, which together provide basking, foraging, and escape cover during the active season; 3) upland habitat (e.g., bankside burrows, holes, and crevices) to provide short-term refuge areas during the active season; 4) high ground or upland habitat above the annual high water mark to provide cover and refuge from flood waters during the dormant winter period (Brode 1988; Hansen and Brode 1980; Hansen 1998). In some rice-growing areas, giant garter snakes have adapted to vegetated, artificial waterways used to bring water into and out of the rice fields (Hansen and Brode 1993).

This species appears to be absent from most permanent waters that support established populations of predatory game fishes, from streams and wetlands with sand, gravel, or rock substrates, from riparian woodlands lacking suitable basking sites, and from any areas that lack suitable prey populations, and cover vegetation (Hansen 1980, Rossman and Stewart 1987, Brode 1988, USFWS 1999). The species also appears to be absent from natural or artificial waterways that undergo routine mechanical or chemical weed control or compaction of bank soils (Hansen 1988, Hansen and Brode 1993).

In the Central Valley, rice fields have become important habitat for giant garter snakes. Irrigation water typically enters the rice lands during April along canals and ditches. Giant garter snakes use these canals and their banks as permanent habitat for both spring and summer active behavior and winter hibernation. Where these canals are not regularly maintained using mechanical or chemical weed control, lush aquatic, emergent, and streamside vegetation develops prior to the spring emergence of giant garter snakes. This vegetation, in combination with cracks and holes in the soil, provides much needed shelter throughout the spring emergence and summer active periods. Rice is planted during the spring after the winter fallow fields have been cultivated and flooded with several inches of standing water. In some cases, giant garter snakes move from the canals and ditches into these rice fields soon after the rice plants emerge above the water's surface, and continue to use the fields until the water is drained during late summer or fall (Hansen and Brode 1993). It appears that the majority of giant garter snakes move back into the canals and ditches as the rice fields are drained, although a few may overwinter in the fallow fields where they hibernate in burrows in the small berms separating the rice checks (Hansen 1998). During the late summer or fall, water is drained from the fields by a network of ditches. These ditches are sometimes routed alongside irrigation canals, and are often separated from the irrigation canals by narrow vegetated berms that may provide additional shelter. Remnants of old natural sloughs also may remain within rice-growing regions where they serve as drains or irrigation canals. Giant garter snakes may use vegetated portions along any of these waterways as permanent habitat.

Changing agricultural regimes, urban development, and other shifts in land use create an ever-changing mosaic of available habitat. Giant garter snakes move around in response to these changes in order to find suitable sources of food, cover, and prey. Connectivity between regions is therefore extremely important for providing access to available habitat and may be important for genetic interchange among populations. In an agricultural setting, giant garter snakes rely largely upon the interconnected network of canals and ditches that provide irrigation and drainage to provide this connectivity. Differential dispersal and home range patterns between males and larger females who spend the majority of the active season gestating young are not reported. Lifetime dispersal patterns of both neonates and adults of this species are unknown.

Prior to listing in 1971, giant garter snakes were known from 16 localities, representing nine distinct populations based on available literature and museum records (Hansen and Brode 1980, USFWS 1993). Range-wide status surveys of the giant garter snake conducted during the mid-1970s and 1980s indicate that they have been extirpated from the San Joaquin Valley south of Mendota in Fresno County, an area comprising as much as one-third of the snake's former range (Fitch 1940, Hansen and Brode 1980, Rossman and Stewart 1987, Stebbins 2003). Once plentiful in areas such as Mendota, Los Banos, and Volta, giant garter snakes are now known from only a small number of localities in the southern aspect of their range (USFWS 1999, Dickert 2003, Hansen 2007).

The current known distribution of giant garter snakes is patchy, and extends from near Chico in Butte County south to the Mendota Wildlife Area in Fresno County. Giant garter snakes are not known to occur from the northern portion of the San Joaquin Valley north to the eastern fringe of the Sacramento-San Joaquin River Delta, where appropriate floodplain habitat of the San Joaquin River is limited to a relatively narrow trough (Hansen and Brode 1980, USFWS 1993). The resulting gap of approximately 100 km (62.3 mi) separates the southern and northern populations, with no giant garter snakes known from the lowland regions of Stanislaus County (CNDDDB 2009, Hansen and Brode 1980). Scattered records within the Sacramento-San

Joaquin River Delta suggest that giant garter snakes may have occupied this region at one time, but longstanding reclamation of wetlands for intense agricultural applications has eliminated most suitable habitat in this region (CNDDDB 2009, Hansen 1986). Recent records within the Sacramento-San Joaquin Delta are haphazard, and until recently surveys have failed to identify any extant population clusters in the region (Hansen 1986, Patterson and Hansen 2004, Patterson 2005).

Occurrence records dating to the 1970's indicate that, within this range, garter snakes were recently distributed in 13 unique population clusters coinciding with historical flood basins, marshes, wetlands, and tributary streams of the Central Valley (Hansen and Brode 1980, Brode and Hansen 1992, USFWS 1999). These populations are isolated, without protected dispersal corridors to other adjacent populations, and are threatened by land use practices and other human activities, including development of wetland and suitable agricultural habitats.

Sacramento Valley populations include Butte Basin, Colusa Basin, Sutter Basin, American Basin, Northern Yolo Basin, Central and Southern Yolo Basin, Sacramento Area, Badger/Willow Creek, White Slough/Coldani Marsh, and East Stockton. San Joaquin Valley populations include the North and South Grasslands (Los Banos), Mendota Area, and Burrel/Lanare Area. Recent surveys suggest that populations in Burrel /Lanare and Liberty Farms may be extirpated (Wylie and Amarello 2006,2008).

#### *Project Background ---*

This project "Genetic Analysis of Giant Garter Snake (*Thamnophis gigas*) populations in the San Joaquin and Sacramento Valleys" was initiated in 2006 under contract [802706G121](#) between the U.S. Fish and Wildlife Service and The CSU Chico Research Foundation and continued in 2007 under contract [802707G111](#) with funding from the US Bureau of Reclamation and US Fish and Wildlife Service's CVPIA Habitat Restoration Program. The overall goal of the project was to provide information for management decisions regarding range-wide population genetic structure of giant garter snake (*Thamnophis gigas*, GGS) using analysis of mitochondrial and microsatellite DNA and. Studies of genetic structure of giant garter snakes populations are identified as a Priority 1 Recovery Task in the giant garter snakes Draft Recovery Plan.

Previous genetic work on the relationship of giant garter snakes to other closely related snakes has shown that the giant garter snake is a "good species" representing a single evolutionary lineage that is ecologically, morphologically and genetically distinct from other similar snakes (DeQueiroz *et al.* 2002). Existing studies of intraspecific genetic diversity of giant garter snakes have showed strong patterns of regional differentiation but produced equivocal results regarding the existence of distinct population segments within the species (Paquin *et al.* 2006). Analysis of maternally inherited mitochondrial DNA markers indicated strongly geographically structured populations potentially consisting of several distinct population segments. However, analysis of a single, bi-parentally inherited, microsatellite marker showed little geographic structure of genetic diversity across the species' range. The differences between inferences drawn from the two genetic markers may be caused by differences in mode of inheritance of the two markers or may be attributed to differences in behavior of male and female snakes. In terms of modes of inheritance the microsatellite loci have an effective population size of  $4N_e$  because each individual carries a chromosome bearing a copy of the gene from each 2 parents and each of those chromosomes contains two copies of the gene on each branch of the chromosome. During meiosis any one of those four gene copies can end

up in a gamete. In contrast the population size of maternally inherited cytoplasmic mtDNA genes is  $1N_e$  because each offspring will get the same copy of mtDNA, which is transmitted only from the mother. The result of this difference in mode of inheritance is that more diversity is maintained in bi-parentally inherited nuclear genes than in mtDNA and because of this, cessation of gene flow between populations will show up first as differences in mtDNA. Bi-parentally inherited markers may also show different patterns of genetic subdivision among populations due to differences in behavior of males and females. If females are philopatric and males are more inclined to migrate then mtDNA would show strong subdivision because females stay in place while microsatellites could show no subdivision because of male mediated gene flow among populations. The conservative interpretation of existing data would probably favor mtDNA data indicating high levels of population structure, because of the smaller population size of this marker compared with nuclear markers and the fact that maternally inherited mtDNA tracks the behavior of females, which are demographically more important to population survival than males.

*Project goals and objectives---*

This study builds upon this previous work and analyzes a larger number of individuals, from more populations, using more markers in an effort to provide a more complete description of range-wide population differentiation. The specific objectives of this study are: 1) increase sampling across all sites 2) to expand the range-wide studies by incorporating samples from previously unsampled areas including the Mendota population in the San Joaquin Basin, White Slough population, and Yolo Basin, 3) to use the well-sampled populations in the Natomas Basin to examine crucial issues concerning connectivity and fragmentation of populations separated by major highways (I5, 99/70) and the Natomas Cross Canal.

*Project area description ---*

Samples were obtained from ongoing research projects in nine regions throughout the range of the giant garter snake (Figure 1). With the exception of samples from the Colusa basin, which were collected by the USGS, Eric Hansen collected the majority of the samples used in this study as part of his studies on abundance and distribution of giant garter snakes across its range. He has provided detailed descriptions of his study sites.

## **Sacramento Valley**

### **1) Butte Basin**

Butte Basin is a low-lying area extending from the Sacramento River south and east to the Butte Creek drainage and southward to include the Butte Sink. Historically this basin consisted of a braided network of sloughs, channels, and oxbows resulting from the meanderings of the Sacramento River and Butte Creek.

Sampling in the Butte Basin was conducted at two unique sites. The Dodge Ranch Property in Butte County, California is located in the Butte Basin on an unnamed section of the U.S. Geological Survey Nelson USGS 7.5-minute Topographic Quadrangle map, Township 20 North, Range 1 East. Dodge Ranch resides northwest of the junction of Godspeed/Agua-Frietas Road and Nelson Road, extending south to the Western Canal and east to the east bank of Butte Creek.

A second location was sampled in Western Butte County in the Butte Basin on Section 21 of the U.S. Geological Survey West of Biggs USGS 7.5-minute Topographic Quadrangle

map, Township 19 North, Range 01 East. Also east of Butte Creek, this privately owned parcel resides south of Dodge Ranch, northwest of the intersection of Aguas Frias Road and the Watt Lateral.

## **2) Colusa Basin**

Colusa basin samples were collected for Melanie Paquin by the U.S. Geological Survey in the Sacramento National Wildlife Refuge Complex (Colusa Delevan & Sacramento NWR) in on the west side of the Sacramento River in Colusa County (Wylie et al 1997, 2000, 2002). These populations appear to be abundant and are well protected in the USFWS managed refuges and likely also occur outside the refuges in appropriate rice land habitats although this has not been confirmed (USFWS 2006).

## **3) Sutter Basin**

The Sutter Basin study area is located north of the confluence of the Sacramento River and Feather River, with these rivers providing the west and east boundaries, respectively. Originating near the Tisdale Weir, the Sutter Bypass subdivides the Sutter Basin into east and west halves. Sampling locations on the western side of the Sutter Basin include areas of aquatic habitat along public roadways within Reclamation District 1500, which is located on the Sutter Causeway and Knights Landing USGS Topo Quads, Township 12 N., Range 2E. and 3E. Features sampled include the highline irrigation canal along the west edge of Armour Road south of Kirkville Road, the drainage canal along the north edge of Maddock Road west of Armour Road, the east edge of Armour Road north of Mackert Road, and the highline irrigation canal along the north edge of Kirkville Road west Highway 113.

Sampling locations in the western half of the Sutter Basin include the Westervelt Ecological Services Sutter Basin Conservation Bank (SBCB) in Sutter County, California from April 9<sup>th</sup> through August 20, 2007. The Bank is located in southeastern Sutter County in the Sutter Basin on Section 28 of the U.S. Geological Survey Sutter Causeway USGS 7.5-minute Topographic Quadrangle map, Township 13 North, Range 3 East. The SBCB resides southwest of the junction of Marcuse Road and Sawtelle Road, along the eastern edge of the Sutter Bypass and west of Highway 99.

Among the 13 identified giant garter snake populations (USFWS 1999), the northern Yolo Basin population is distributed along the northeastern edge of the Yolo Basin near the Sacramento River. Yolo County is well within the Central Valley proper and includes the floodplains of the Sacramento River as well as those of Cache, Willow, and Putah Creeks. Upon receding, these creeks may have provided the wetland habitat and prey utilized by giant garter snakes during the spring and summer active season. The historical distribution of giant garter snakes in Yolo County is unclear; however, with the majority of sightings made only in recent decades (Hansen 1986, CNDDDB 2009).

Locality records indicate that garter snakes are distributed in as many as 13 unique population clusters coinciding with historical flood basins, marshes, wetlands, and tributary streams of the Central Valley (Brode and Hansen 1992, USFWS 1993, USFWS 1999), including the Yolo/Willow Slough and Yolo/Pope Ranch populations that lie to the north and southwest of the Yolo Wildlife Area, respectively. Within this distribution, giant garter snakes are documented in two distinct concentrations along the eastern edge of Yolo County (CNDDDB 2009). The first concentration lies in the northeastern portion of Yolo County northwest of Knights Landing, in the southern end of the Colusa Basin near Sycamore Slough. The second

concentration, which lies in the eastern central and southern portion of Yolo County, is represented by two unique concentrations. The eastern central concentration is composed of records in the Yolo Bypass east of Conaway Ranch near the Tule Canal, within the Willow Slough/ Willow Slough Bypass from the Conaway Ranch south to the Yolo Wildlife Area, and along the western edge of the Yolo Bypass east of Interstate 80 within and adjacent to the Yolo Wildlife Area. The southern concentration is composed of records on or near the Pope Ranch Preserve, managed by Wildlands, Inc.

#### **4) American Basin**

The American Basin (Figure 2) is an historical low-lying drainage depression on California's Sacramento Valley floor. Situated northeast of the confluence of the American and Sacramento Rivers in the northern portion of Sacramento County, the American Basin extends northward to the Bear River in southern Sutter County and is bounded on the west by the Sacramento and Feather Rivers. The eastern boundaries of the American Basin are determined by elevational clines marked by shifting soil types and changing hydrology.

Historically flooded by runoff from the western slopes of the Sierra Nevada Range, the American Basin was modified by reclamation levees, pumps, and canals in 1914, converting a prevailing marsh and wetland landscape to agricultural production. Rich, impermeable, alluvial clays deposited by historical floods slow drainage within the Basin and encourage the rice agriculture that has come to dominate area land use since the 1940's. Hydrology and natural flows within the American Basin are interrupted/intercepted in the east by 1) the Natomas East Main Drainage Canal (NEMDC), now also known as Steelhead Creek, which begins just south of Sankey Road and flows south to the American River; and 2) the Pleasant Grove Creek Canal (PGCC), which begins at Sankey Road and flows north to the Natomas Cross Canal (NCC), and 3) the East Side Canal, which begins at Coon Creek and flows south to the NCC. Reclamation has divided the American Basin into two distinct segments that are separated by the Natomas Cross Canal, a channel connecting the Sacramento River to the PGCC and East Side Canal from west to east. Although USFWS identifies Lake Oroville as the northern terminus of the American Basin (USFWS 1999), for the purpose of this report, the Middle American Basin refers to that part of the Basin that extends from the NCC north to Ping Slough, bounded by the East Side Canal to the east. The Natomas Basin is that portion of the lower or southern American Basin extending from the NCC south to the confluence of the Sacramento and American Rivers.

While the American Basin historically possessed no physical barriers to species movement, the NCC and East Side Canal potentially segregate giant garter snake populations along the east side of the middle American Basin. The NCC and East Side Canal consist of two levees separated by a broad channel supplied to a large extent by backup of the Sacramento and Feather Rivers. Giant garter snakes are negatively associated with large rivers and waters supporting predatory game fish (Hansen 1988), and because previous surveys indicate low usage of the NCC interior by giant garter snakes (G. Hansen, pers. comm.; E. Hansen 2006), the NCC and East Side Canal may act as potential barriers to giant garter snake movement and genetic exchange (Hansen 2005).

#### **4) Southern American/Natomas Basin**

The Natomas Basin is subdivided by major highways into three regions (Figure 2): (1) south and west of Interstate 5, (2) north and east of Interstate 5 and east of State Route 99/70, and (3) north of Interstate 5 and west of State Route 99/70. Features such as box or pipe culverts

linking regions otherwise separated by major roadways or urban development are also depicted. The most significant corridors spanning the Basin from north to south continue to be the primary drainages managed by Reclamation District 1000: the North Drainage Canal, East Drainage Canal, West Drainage Canal (including Fisherman's Lake), and Main Drainage Canal, all of which the NBHCP has identified as most likely to remain during the permit term.

Area 1 and Area 2 are connected by the West Drainage Canal, the N Drain (parallel to Powerline Road), and Lone Tree Canal through culverts that pass under I-5. The West Drainage Canal passes north under I-5 from the Fisherman's Lake area to the area west of SMF, where it lies disconnected from other hydrologic features. Until recently, the series of laterals emanating from the Lone Tree Canal culvert crossing at I-5 provided the only presumably functional connective corridor between Areas 1 and 2. However, even this connection is tenuous. The cessation of farming and urban development adjacent to the Lone Tree Canal has resulted in unpredictable and reduced water deliveries. With the exception of periodic drain water deliveries, this feature has not functioned as a viable migration corridor since 2006. The East Drainage Canal provides the only connection between Areas 2 and 3; this section constitutes 12 kilometers (7.6 miles) of disturbed channel surrounded by urban development. As such, this connection is largely unreliable; Lone Tree Canal, therefore, provides the only potentially viable connection between reserves south of I-5 and other regions within the Basin. Areas 1 and 3 are connected by the V Drain, R Drain, H1 Drain, and the Central Main Canal through culverts passing under SR 99; each of these connects to a series of ditches, drains, and canals in their respective regions. However, the majority of these features convey water via box culverts that are entirely inundated by irrigation and/or drain water during the giant garter snake active season

## **5) Yolo Basin**

Among the 13 identified giant garter snake populations (USFWS 1999), the northern Yolo Basin population is distributed along the northeastern edge of the Yolo Basin near the Sacramento River. Yolo County is well within the Central Valley proper and includes the floodplains of the Sacramento River as well as those of Cache, Willow, and Putah Creeks. Upon receding, these creeks may have provided the wetland habitat and prey utilized by giant garter snakes during the spring and summer active season. The historical distribution of giant garter snakes in Yolo County is unclear; however, with the majority of sightings made only in recent decades (Hansen 1986, CNDDDB 2009).

Locality records indicate that garter snakes are distributed in as many as 13 unique population clusters coinciding with historical flood basins, marshes, wetlands, and tributary streams of the Central Valley (Brode and Hansen 1992, USFWS 1993, USFWS 1999), including the Yolo/Willow Slough and Yolo/Pope Ranch populations that lie to the north and southwest of the Yolo Wildlife Area, respectively. Within this distribution, giant garter snakes are documented in two distinct concentrations along the eastern edge of Yolo County (CNDDDB 2009). The first concentration lies in the northeastern portion of Yolo County northwest of Knights Landing, in the southern end of the Colusa Basin near Sycamore Slough. The second concentration, which lies in the eastern central and southern portion of Yolo County, is represented by two unique concentrations. The eastern central concentration is composed of records in the Yolo Bypass east of Conaway Ranch near the Tule Canal, within the Willow Slough/ Willow Slough Bypass from the Conaway Ranch south to the Yolo Wildlife Area, and along the western edge of the Yolo Bypass east of Interstate 80 within and adjacent to the Yolo

Wildlife Area. The southern concentration is composed of records on or near the Pope Ranch Preserve, managed by Wildlands, Inc.

## **Delta**

### **6) Badger Creek**

The Badger Creek study area is located on the Cosumnes River Preserve, approximately 20 miles south of the City of Sacramento. The areas surveyed include the Castello, Denier, Shaw, and Whaley properties; the waterways along which giant garter snakes could disperse to these properties; and the giant garter snake population center on Badger Creek west of Highway 99. This includes Laguna and Badger creeks (the main stem and wetlands west of Highway 99, and the North and South Forks east of Highway 99).

Laguna and Badger creeks are both portions of the California Trough section of the Cosumnes River watershed (Phillip Williams and Associates. 1997). Badger Creek is a tributary merging with the Cosumnes River in southern Sacramento County. Ranging from its confluence with the Cosumnes River approximately 2.25 miles west of Highway 99, Badger Creek is characterized by a series of annual and perennial marshland habitats connected by open and riparian low-gradient channels. Approximately 0.5 mile east of Highway 99, Badger Creek divides into north and south tributaries that extend eastward approximately 7 to 8 miles. The North and South forks are characterized by annual and perennial marshland, artificial channels and retention ponds supported mainly by agricultural runoff during the dry season.

Laguna Creek, merging with the Cosumnes River approximately one mile below Badger Creek, is also characterized by a series of annual and perennial marshland habitats connected by open and riparian low-gradient channels. The main stem of Laguna Creek, however, is far more riparian in character and is situated in somewhat hillier topography. The result is a more fragmented series of wetlands without nearly the area of permanent, stable marsh observed at Badger Creek.

Giant garter snakes are known to occur at Badger Creek in stable populations at areas west of Highway 99 (Wylie *et al.* 1997, E. Hansen 2001, CNDDDB 2009). George Hansen reported sightings of giant garter snakes east of Highway 99 in 1986 at the confluence of Badger Creek and Willow Slough, and at the North Fork Badger Creek at Riley Road (CNDDDB 2009). The easterly sightings, however, are representative of individual giant garter snakes and may or may not indicate stable and permanent giant garter snake populations at this locale. No historical giant garter snake sightings are reported along any portion of Laguna Creek or its associated wetlands.

The sampling area is centered within the areas of perennial marsh from west of Highway 99 at the confluence of Badger and Willow Creeks near the Union Pacific Railroad tracks west to the Valensin Forest. The wetlands here are sometimes referred to as Snake Marsh by Cosumnes River Preserve staff, or as Arno Road Marsh (NDDDB 2002). Peripheral zones of interest include areas of Badger Creek at both North Fork and South Forks extending from their confluence to approximately 1 to 2 miles east of Riley Road, and bounded by Dillard and Arno roads at the north and south, respectively, and the Cosumnes River and Laguna Creek north and east of Twin Cities Road upstream to the Valensin Forest. This report subdivides areas of this marsh

complex into 3 sections. The area west of the Union Pacific Railroad tracks to Valensin Forest is connected directly to Badger Creek via Horseshoe Lake.

## **7) White Slough**

The White Slough Wildlife Area consists of 880 acres of man-made ditches, canals, and freshwater marshes with associated grassland/upland habitats. It is located eight miles west of the City of Lodi and west of Interstate 5 on the Walnut Grove, Woodbridge, Cotta and Interstate 5 Frontage Roads.

The White Slough Wildlife Area supports one of 13 extant giant garter snake populations recognized by the USFWS (Caldoni Marsh/White Slough population) (USFWS 1999). First identified on site in 1974 (CNDDDB 2009), giant garter snakes were observed at White Slough Wildlife Area by George Hansen from the time he began surveying for them in 1976 (G. Hansen and J. Brode 1980, G. Hansen 1988, 1996) until the mid-1990's. Between two giant garter snake populations recognized in San Joaquin County, the White Slough population is perhaps the only locality still supporting a viable snake population. After failing to detect giant garter snakes east of Stockton during surveys conducted in the 1980's and 1990's, George Hansen speculated that this population (Stockton Diverting Channel/Duck Creek population) was likely extirpated by extensive urban development occurring since the 1970's (G. Hansen 1988, 1996).

Between 1974 and 1978, 13 rectangular borrow pits were excavated from one to five miles west of Interstate 5 to provide fill for freeway construction (DWR 1995). The pits are fed by groundwater and periodic runoff from precipitation, irrigation, and high canal flows, creating a series of ponds characterized by vegetated sloping or vertical banks and open water with adjacent uplands and high ground. White Slough Wildlife Area encompasses ponds 7-13 along a roughly 14-mile stretch between Thornton and Stockton.

Most giant garter snake observations at White Slough Wildlife Area are concentrated at Pond 9, but surveys conducted by George Hansen in 1994 yielded additional sightings at Pond 7, Pond 11, and a site between Ponds 6 and 7 (CNDDDB 2009; DWR 1995). Although channels and drainages including Telephone Cut, Sycamore Slough, Hog Slough, and Beaver Slough were surveyed, observations were made only at the ponds (M. Green *pers. comm.*). Each of the ponds where snakes were observed are characterized by slow moving water with mud banks and bottoms, vegetative cover, and access to high ground (DWR 1995). Giant garter snakes may occupy features connecting the ponds that are characterized by similar features. All of the samples provided in the study originated in the wetlands surrounding Pond 9.

## **San Joaquin Valley**

Extant giant garter populations within the San Joaquin Valley are represented by three unique management areas; North and South Grasslands (Grasslands Ecological Area), Mendota Area, and the Lanare/Burrell Area. Tulare Lake Basin and Kern-Wasco Area populations are presumed extirpated, and observations of deteriorating habitat at Burrell-Lanare in 1992 led to the conclusion in the final listing that this population, if it was not already extirpated, was severely and imminently threatened [USFWS 1993].). With one exception (NDDDB # 144), all reported giant garter snake occurrences in the San Joaquin Valley originate south and west of the San Joaquin River where large wetland complexes remain.

Most of these locality records were accumulated during a range-wide status and distribution survey conducted for DFG during 1976 and 1977, which determined that giant garter snakes were potentially extirpated from wetland regions of Buena Vista and Tulare Lake basins near Bakersfield in Kern County that had been drained for agriculture.

Areas studied include Mendota Wildlife Area (MWA), the San Luis National Wildlife Refuge complex (SNLNWR), the consortium of privately owned properties situated within Grasslands Resource Conservation District (RCD), the Merced National Wildlife Refuge complex (MNWR), the privately owned Modesto Properties situated west of MNWR's Snobird Unit south of Highway 140, and the core of Stevinson Water District along with its associated rights-of-way along the East Side Canal corridor.

The Stevinson and Merquin Water Districts are located east of the confluence of the Merced and San Joaquin Rivers in Merced County, California. The East Side passes from east to west through the Districts, extending southeast from SMWD through the Arena Plains and Snobird Units of the Merced National Wildlife Refuge Complex (MNWR) to the Mariposa Bypass, East Side Bypass, and Merced Unit of the MNWR.

SNLNWR lies south of the San Joaquin River, encompassing wetlands east and west of Highway 165 south to the City of Los Banos. Grasslands RCD lies to the west of SNLNWR, extending from Highway 140 south to the Merced/Fresno County line. Encompassing privately managed lands adjacent to SNLNWR through the Los Banos Creek and Santa Fe Grade corridor, North Grassland Water District (GWD) extends to Highway 152 in the City of Los Banos. South GWD continues through the Santa Fe Grade corridor south of Highway 152. Situated both east and west of Fresno Slough, MWA is located in Fresno County, approximately 3 miles south of the town of Mendota near White's Bridge and, ten miles west of the town of Kerman.

## **8) Los Banos**

Grassland Water District (GWD) comprises approximately 51,537 acres of primarily wetland habitat. The District maintains approximately 110 miles of canals in order to execute its primary function of delivering water to the landowners within its boundaries. The approximately 75,000-acre Grasslands RCD comprises private hunting clubs and other privately owned wetland areas, as well as all or portions of several state and federal refuges. To achieve a goal of sustaining waterfowl habitat, the management objectives of the Grassland RCD include encouraging natural food plant production (such as swamp timothy, smartweed, and wildlife millet) and habitat protection. Land uses include seasonally flooded wetlands, moist soil impoundments, permanent wetland, irrigated pasture, and croplands.

The Grassland RCD contains most of the 51,530-acre GWD, which is a legal entity established to receive and distribute CVP water. GWD delivers CVP water to the wetland areas within its boundaries. GWD contains approximately 165 separate ownerships, most of which are hunting or duck clubs. Perpetual easements have been purchased by the Service to help preserve wetland-dependant migratory bird habitat on approximately 31,000 acres serviced by the GWD.

## **9) Mendota WA**

Acquired by the Wildlife Conservation Board from 1954 to 1991, the approximately 11,802-acre MWA comprises intensively managed, semi-permanent wetlands and associated uplands surrounding a 600-acre segment of the Fresno Slough, a natural drainage providing both a source of water and a riparian corridor. MWA is managed primarily as seasonally flooded wetland to provide the habitat needs of migratory waterfowl and associated species, with

approximately 9,800 acres of the area managed as seasonally flooded wetlands. The geologic history of the area is that of a typical floodplain, characterized by fine textured clays. Water table levels are generally high, drainage is poor, and soil salinity is sufficiently high to restrict vegetation types. Annual precipitation averages less than six inches from winter rains.

## Methods

### Sample Collection

*Sampling* --- We obtained tissue (tail clips) and blood samples from a variety of sources including environmental consultants, and State and federal conservation agencies. Eric Hansen, a private consultant and researcher, contributed tissue samples for a total of 634 individual giant garter snakes from throughout the species range. Among this total, 158 were from sites in the American Basin, 237 from the Natomas Basin, 74 from Badger Creek, 33 from the San Joaquin Basin, 69 from the Yolo Basin, and 41 from the Sutter Basin. Melanie Paquin (Molecular Geneticist at the National Marine Fisheries Service/Conservation Biology Division Seattle WA) provided subsamples of blood, tissue and/or DNA from an additional 221 animals. Paquin's samples were collected by field teams working for California Department of Fish and Game and the USGS Biological Resource Division and served as the basis of her Master's Thesis work conducted at San Francisco State University and a subsequent publication (Paquin *et al.* 2006). This combined sampling totals 855 individuals from across the extant range of giant garter snakes (USFWS 1999). Geographic distribution of collection locations is provided in Table 1. Tissue samples provided by Melanie Paquin were stored on ice in the field and transferred to a -80°C freezer at San Francisco State University where they are presently in long-term storage. Subsamples of DNA or tissue from Paquin's study were shipped on dry ice to our lab at CSU Chico where they were subsequently transferred to a -20°C freezer for long-term storage. Samples provided by Eric Hansen were preserved in 95% ethanol in the field and shipped at ambient temperature to CSU Chico where they were then transferred to a -20°C freezer for long-term storage.

*Molecular techniques* --- A glossary of terms and abbreviations relating to molecular techniques is provided in appendix XX. Whole, genomic DNA was isolated from tissue samples using a guanidium thiocyanate "salting out" technique (Sambrook & Russell, 2001 Protocol attached as appendix ##) and re-suspended in 50-200 microliters ( $\mu$ l) of DNA, RNA and nuclease free water. Isolated genomic DNA was then used as template for Polymerase Chain Reaction (PCR) amplification of specific regions of the Mitochondrial DNA (mtDNA) and microsatellite loci in the nuclear genome as described below

*Mitochondrial DNA* - The Polymerase Chain Reaction (PCR) using the primers ND4 (CACCTATGACTACCAAAGCTCATGTAGAAGC) and Leu (CATTACTTTTACTTGGATTTGCACCA). These primers were originally designed by Arevalo et al (1994) for use in lizards but effectively amplify a ~ 900 base pair fragment of the mitochondrial genome containing part of the NADH dehydrogenase subunit 4 (ND4) gene and the Histidine, Serine and Leucine transfer RNA genes from many vertebrates including Giant Garter Snakes. PCR reactions to amplify the mtDNA ND4 region were conducted in 10-25  $\mu$ L volumes containing 2 $\mu$ L of DNA template, 0.5mM of each primer, 0.125 mM of each dNTP,

0.25 mM MgCl<sub>2</sub>, and 0.5 units of *Taq* DNA polymerase (Invitrogen). The thermal cycle profile for PCR reactions consisted of a 3 minute initial denaturation at 94°C followed by 35 cycles of denaturation for 30 seconds at 94°C, annealing for 45 seconds at 53°C extension for one minute at 72°C, with a final 3 minutes extension at 72°C following the last cycle. Negative controls (reactions containing all reagents but to which no DNA template are added) were used in all PCR reactions to check for possible contamination. PCR products were sent to the University of Washington High Throughput Genomics Unit for sequencing. Before conducting DNA sequencing it is necessary to “clean up” PCR products by removing unincorporated primers and dNTPs which may interfere with sequencing the desired PCR product. This process was conducted at the High Throughput Genomics Unit by adding exonuclease 1 and shrimp alkaline phosphatase (Amersham Pharmacia Biotech) these two enzymes break down small pieces of DNA while leaving the larger desired PCR product intact to be sequenced. Cleaned up products were then sequenced using an Applied Biosystems Inc. model 7300 automated DNA sequencer. Only the primer ND4 was used in the sequencing reaction in order to provide DNA sequence for the coding strand of DNA. Sequences were aligned by eye using the computer program SeqEd (Applied Biosystems Inc.) using the sequence of Paquin’s “Haplotype D” (the most common DNA sequence found in previous studies of giant garter snakes ) as a reference sequence (Paquin *et al.* 2006, genbank accession AF414090).

*Microsatellites* --- We screened published primer sets for a battery of 12 different microsatellite loci. These primers included:

Ts1 (CGGCATAAATCTTATCTAGC, ACTTTTTCAGGCTGATGTTC)  
Ts2 (GGCTAGCCCCTGTGTCCTT, CACAACCTCCAAATATTGAAGATTA)  
Ts3 (CAACTGGCSGCTGTGATACAA, GTGTTAATGTGTTGGACAGGGC)  
Ts4 ACTGAACAAGTTGGGTGTAG, GCAAGAAGATGGCTATCTTG  
Developed by McCracken *et al* 1999

Nsu2 (TCCTCTTTGGCAGAGTAATAGT, AGCCGAGAACACACTAGTAAGT)  
NSu3 (CTGACTCACTTCTGACCCTAAT, AATATTTGCTTGGCTCAAAC)  
Developed by Prosser *et al* 2000

and

*Te1Ca4* (ACGGTCAAGAAGAAATCCTG, AATCATGAATGGTCTATCAAAAAG)  
*Te1Ca2* (GGTGTCCCTTCTCGGTTCAATACCAG, CATGGAAAACAAGAGGTTGG)  
*Te1Ca3* (CCCCACCTACCTACCTG, TGGGTAGGGCAAAAACCAG)  
*Te1Ca18* (CCCCACATTTTTGGCAAG, TCGGAGTGTGTGTTGGAGTG)  
*Te1Ca29* (TGCCTTATTTGCTTGGGTTG, TCTTTC AACCTGCTTTGTAGACAC)  
*Te1Ca50* (TGTGGGCTCTTCAGAACTGG TGCTGCTTGCATAAGTGGAG)  
Developed by Garner *et al*, 2004.

For each primer pair we ran multiple PCR reactions in a factorial design with varied annealing temperature, primer concentration, MgCl<sub>2</sub>, and primer concentration. We checked for successful amplification and conducted preliminary analysis of size variation by running PCR products in a 4% agarose gel.

Of the 12 microsatellite primers that we screened, we successfully amplified six and preliminary screening showed size variation in four of these six. The four variable Microsatellites, approximate size of the product and the type of fluorescent dye used are: NSU2 (~150bp Blue), NSU3 (~120bp-140 bp Green), TS3 (~115 Red), Te1CA18 (~90-120 Yellow).

This combination of size differences and different dyes allowed limited multiplexed scoring of microsatellite data by allowing successful multiplexed genotyping by combining NSU2 + TS3 and NSU3 + Te1 CA18. Microsatellite genotypes were scored by the CSUPERB Biochemical Core Facility at CSU San Diego Using an ABI 310 automated sequencer with fragment analysis software. Microsatellite fragments were compared against a size standard and scored for size using the GeneMapper (Applied Biosystems Inc).

*Data Analysis* --- Standard descriptive data of genetic diversity including haplotype diversity and nucleotide diversity were calculated using Arlequin 3.1.1. (Schneider *et al.* 2000). A haplotype network showing relationships of mtDNA haplotypes was constructed using distance methods implemented in the program PAUP\* (Swofford 2002). Initially we used two approaches to analyze population structure and demographic signatures in these data: 1) Analysis of Molecular Variance (AMOVA) which is the Molecular analog of the standard statistical practice of Analysis of Variance (ANOVA) as described by Excoffier *et al.* (1992) and implemented in Arlequin 3.1.1 (Schneider *et al.* 2000); and 2) a Maximum Likelihood method developed by Beerli and Felsenstein (1999, 2001) and implemented in the program MIGRATE 2.3.

AMOVA partitions molecular variance at three levels:  $\phi_{ST}$  summarizes variance among all populations,  $\phi_{SC}$  summarizes the variance within a population relative to the rest of its region, and  $\phi_{CT}$  summarizes the variance among regions. A statistical test for significant geographic structure in the sample is constructed by comparison of the observed pattern of genetic variation with computer generated randomizations of the data described by Excoffier *et al.* (1992) and implemented in Arlequin 3.1.1 (Schneider *et al.* 2000). Statistical significance for all  $\phi$  values were calculated in Arlequin 3.1.1 by comparing observed  $\phi$  to values calculated from 10,000 randomly permuted datasets.

The coalescent-based analysis implemented in MIGRATE is very powerful and particularly useful in a management context because it can provide estimates of asymmetric migration rates ( $Nm$ ) among any number of populations (Beerli and Felsenstein 2001). One advantage of this approach over traditional approach using  $F$ -statistics ( $F_{ST}$ ), (Weir and Cockerham, 1984) and  $\Phi_{ST}$  (Excoffier *et al.*, 1992) is that in addition to the overall amount of genetic exchange between populations it also provides estimates of the direction of migration. This provides information on source sink dynamics within a set of interconnected populations and provides improved estimates of ancestral population sizes. In practice these estimates are difficult to calculate and are subject to high error if the data show certain characteristics including 1) low diversity, 2) a small number of haplotypes shared among populations 3) a large number of “private” haplotypes (those found in one population and nowhere else) and 4) presence of populations characterized by a single haplotypes that widespread and common in other populations (P. Beerli, Pers. Comm.). Unfortunately all of these conditions are present in our giant garter snakes data. As a result none of our MIGRATE analyses converged upon reliable parameter estimates. The results were inconsistent between runs and had unacceptably wide confidence intervals and therefore are not included in this report. MIGRATE analyses may still be appropriate and necessary to address specific questions regarding gene flow in giant garter snakes, but these analyses will likely require numbers of samples that are an order of magnitude greater than what is presently available (i.e. thousands rather than tens or hundreds of samples from a given local. Peter Beerli, Pers. Comm.).

Because our data did not allow us to use the coalescent approach, we instead used

Arlequin 3.1.1 to calculate  $F$ -statistics ( $F_{st}$ ), (Weir and Cockerham, 1984) and make inferences regarding past population stability using Harpending raggedness index (Harpending, 1994) and recent population expansions using Fu's  $F$  (Fu, 1997), and Tajima's  $D$  (Tajima, 1989). Statistical significance for all statistics were calculated in Arlequin 3.1.1 by comparing observed values with distribution of values from 10,000 randomly permuted datasets. The Harpending raggedness index compares the observed distribution of DNA mismatches between individuals in a population sample with the distribution expected in an expanding population. The distribution of mismatches in a population that has undergone a sudden demographic expansion is expected to be unimodal (Rogers, 1995) whereas the mismatch distribution in a population that has been stable will be bimodal or multimodal (Rogers and Harpending, 1992) due to the accumulation and persistence of deeper divergences among lineages in the population. A large Harpending raggedness index indicates departure from the smooth unimodal distribution and is a signature of past population stability. In contrast a negative value of either Fu's  $F$ , and Tajima's  $D$  indicates and overabundance of low-frequency haplotypes in a population, which is a possible signature of recent demographic expansion. By combining these two approaches it is possible to infer some details of the demographic history of a population.

We grouped our samples by watershed, following Paquin *et al.* (2006) and delineations in the USFWS Draft Recovery Plan for the Giant Garter Snake (USFWS 1999). Paquin *et al.* (2006) refer to these watersheds by numbers 1-6 but these numbers do not correspond to USFWS population numbers nor with our numbering in Figure 1 therefore to facilitate reference to her work I will define our sampling regions relative to Paquin's watershed numbers here but to avoid confusion I will hereafter will use the more descriptive names (Butte Basin, North Natomas Basin etc.). Butte Basin (Paquin's Watershed 1) includes populations in the northern Sacramento River Valley on the east side of the Sacramento River from the Sutter Buttes north to Red Bluff. Colusa Basin (Paquin's Watershed 2) includes populations in the northern Sacramento River Valley on the west side of the Sacramento River. Sutter Basin (Paquin's Watershed 3) includes populations in Sacramento River Valley, east of the Sacramento River west of the Feather River. The American Basin unit includes populations in Sacramento River Valley east of the Sacramento River from the confluence of the Sacramento and American Rivers north to Oroville. Paquin considered this region collectively as Watershed 4 but for the purposes of our analyses we further subdivided the American Basin into four separate sub-basins (1) Middle American Basin (2) North Natomas, (3) East Natomas, and (4) South Natomas (Figure 2) based on presence of highways and major canals which potentially serve as important anthropogenic barriers to snake dispersal. The Middle American Basin, refers to that part of the American Basin that extends from the Natomas Cross Canal north to Ping Slough, and is bounded to the east by the East Side Canal. This is the area sometimes referred to in regulatory documents as "Area B". The Natomas Basin is that portion of the lower or southern American Basin that extends from the Natomas Cross Canal south to the confluence of the Sacramento and American Rivers. The Natomas Basin is subdivided by major highways into three regions: North Natomas - north of Interstate 5 and west of State Route 99/70; East Natomas - north and east of Interstate 5 and east of State Route 99/70; and South Natomas - south and west of Interstate 5 East of the Sacramento River North of the American River. We also have samples from Delta Basin (Paquin's Watershed 5) populations in Badger Creek in and previously unsampled populations in White Slough. Paquin's Watershed 6 includes populations from the San Joaquin Valley. In addition to the watersheds sampled by Paquin *et al.* (2006), we have also

included a significant number of individuals from the USFWS Yolo Basin unit, which includes populations in the Sacramento River Valley West of the Sacramento River.

We conducted analyses with several different regional groupings of watersheds to test genetic structure in giant garter snake populations at several different regional scales. 1) All Sacramento Valley populations vs. Delta basin (Badger Creek + White Slough) vs. San Joaquin Valley; 2) Western Sacramento Valley basins (Yolo, Colusa) vs. eastern Sacramento Valley basins (American/Natomas, Sutter, Butte); 3) within the eastern Sacramento Valley, northern basins (Butte, Sutter) vs. southern basins American/Natomas 4) Middle American basin vs. north Natomas vs. east Natomas vs. south Natomas to test for subdivision among these Sacramento area basins that are most imminent threats from urbanization.

## Results

**mtDNA sequence diversity** --- 828 base pairs of sequence data were obtained for the mitochondrial ND4 gene and adjacent His, Ser and Leu tRNAs from a total of 466 individual giant garter snakes including Butte Basin (42) Colusa basin (42), Sutter basin (24), 170 total from the American basin, Middle American (44), Natomas North (68), Natomas East (35), Natomas South (23); Yolo Basin (49) Delta/Badger Creek (85), White Slough (5) and San Joaquin Valley (49) (Table 1). Both haplotype diversity (range 0.00- 0.6754 mean 0.59885) and nucleotide diversity (range 0.00- 0.00268 mean 0.000723) within populations was low (Table 2). Badger Creek was the most diverse in each of these measures of diversity. Paquin *et al.* (2006) identified 10 mtDNA haplotypes (A-J) among the 200 individuals sampled. We identified all 10 of these haplotypes in our newly sequenced individuals and discovered three additional novel haplotypes; K, L, and M (Table 1 and appendix A). The novel haplotypes were found in samples from White Slough (one type K among five sequences), Badger Creek (four type M among 85 sequences) and the San Joaquin Valley (one type L among 49 sequences). Overall, haplotype D was, by far, the most common and widespread of the mtDNA haplotypes, found in 276 of 466 samples overall. Haplotype D was present in all sampled populations except the two Delta sites (Badger Creek and White Slough) and was the only haplotype found in the 49 Yolo Basin sequences. All 85 individuals sampled from Badger Creek had haplotypes; A, B, C or the novel M. Three of these haplotypes are unique to Badger Creek and none are found anywhere outside the Delta. The only individuals from outside Badger Creek with any of these haplotypes were four A haplotypes among the five individuals from the other site in the delta (White Slough), further corroborating the unique genetic identity of the Delta populations, and Badger Creek in particular. Sutter Basin, Mid American Basin and White Slough each had one unique haplotype and San Joaquin Valley populations had 3 unique haplotypes.

**Demographic parameters from mismatch analyses** --- Values of Harpending raggedness index Tajima's *D* and Fu's *F*, calculated from mismatch analyses are presented in Table 3. Populations from Colusa, Middle American Basin and Badger Creek all had significant raggedness indices indicating a history of population stability. North Natomas and South Natomas also had high raggedness indices that were nearly significant ( $P=0.06$  for each). Only Sutter basin population showed significantly negative Tajima's *D* and Fu's *F* indicating recent population expansion. Although it is likely that Yolo Basin was recently colonized or experienced a recent population expansion, these mismatch statistics could not be calculated for Yolo basin because this population is characterized by a single haplotype and therefore has no mismatches.

**Microsatellite Loci** --- we obtained data from 4 microsatellite loci from 96 individuals from across the range of giant garter snakes. One locus (NSU3 used by Paquin et al 2006) was variable, two other loci (NSU2 and Te1CA18) showed very low levels of polymorphism and one locus (Te3) showed no polymorphism among sampled individuals. None of the three variable microsatellites showed significant deviation from Hardy-Weinberg equilibrium.

**Pairwise Fst analyses**--- Pairwise Fst values statistical significance of pairwise Fst values among all sampling locations based on distribution of mtDNA haplotypes Tables 4. Similar results for analysis of just Middle American Basin and Natomas North, Natomas East, and Natomas South are shown in Table 5. Table 6 presents similar estimates for analysis of a subset of the basins for which both microsatellite and mtDNA data were available. A significant Fst value between two populations indicates genetic differentiation between populations and implies little ongoing genetic exchange between the populations. The general pattern in Fst estimates among all basins is that mtDNA shows very strong differentiation across all levels. The only pairs of basins that do not show significant differentiation are Butte/Yolo Butte/Sutter Sutter/North Natomas Sutter/South Natomas and among some of the Natomas basins. The results for Sutter and Yolo are consistent with mismatch analyses indicating that these populations may have been undergone colonization and population expansion from another nearby basin. All other pairs showed significant differentiation at the  $P < 0.05$  or Bonferroni corrected  $P < 0.0125$  level. This pattern is strongest in Badger Creek San Joaquin Valley, which are well differentiated from all others and have very high very significant Fst values. This pattern of strong inter-population differentiation is perhaps most telling in comparisons of geographically proximal sets of populations, which are not as molecularly divergent and geographically isolated as Badger Creek. For example Yolo, and the mid American/Natomas Basins share some haplotypes and yet still show strong structure (*i.e.*, low gene flow) among these three geographically proximal sets of populations. This same pattern is true among sub-basins of the American (Table 5) where even within the watershed populations show signs of long-term past stability and differentiation (significant Fst in four of six pairwise comparisons).

**AMOVA analyses**--- Results of four sets of Analysis of Molecular Variance (AMOVA) analyses are presented in Tables 7-10. Each analysis consists of different subsets or different regional groupings in order to examine the population structure within a given region or examine suitability of different possible of regional groupings. Tables 7-9 give results for separate analyses of mtDNA and microsatellite data. Table 10 only includes mtDNA results. In analysis including all 11 basins grouped into three regions as Sacramento Valley, Delta, or San Joaquin Valley (Table 7), MtDNA shows very strong differentiation among regions (*i.e.*, Sacramento Valley vs. American/Natomas vs. Delta vs. San Joaquin Valley). The greatest proportion of the molecular variance in mtDNA (~49%) is explained by differences among regions compared with ~7% explained by differences among populations within a region (*e.g.* Butte Basin vs. Sutter Basin vs. Yolo Basin vs. Colusa Basin, American basin within the Sacramento Valley region), and ~44% explained by differences within individual populations. All of these variance components were statistically significantly different from random distributions ( $P < < < 0.001$ ) indicating that our genetic data is strongly structured at all levels from regional to inter-population. In contrast microsatellite data for the same groupings show little geographic structure with ~3% of the variance among regions ~3% among populations within regions and an overwhelming 94% of the variation within individual populations.

Table 8 gives AMOVA results for analyses including just the five Sacramento Valley sites grouped into two regions: East (Butte, Sutter, American) and West (Yolo, Colusa),

essentially to test if there is a regional genetic structure defined by the Sacramento River. In this analysis For both mtDNA and microsatellites variance components are significant both within populations and among populations within regions but not between regions indicating that there is differentiation among the populations but it is not organized East-West across the Sacramento river but rather with each population structured and separate from the others. The structure is much stronger and more significant in mtDNA than in microsatellite data. Analysis of populations in the eastern Sacramento Valley grouped North (Butte Sutter) vs. South (American/Natomas) gives a similar story with significant variance within populations and among populations within regions but not between regions and stronger evidence from mtDNA than from microsatellites. This same story is repeated again in analysis of the just the American River Basin locations; Middle American Basin, North Natomas, East Natomas and South Natomas. The significant among population variance components indicates that there is strong among-population structure within the basin which correlates with the major landscape features (highways and canals) that we identified as potential barriers to gene flow in the region but the lack of significant “among regions” variance indicates that any particular grouping of any population with each other are not supported by data. Together these analyses characterize giant garter snakes as a species with highly genetically structured populations with low genetic exchange across the species range, among adjacent watersheds, and among populations within watersheds.

## Discussion

In this study we have built upon previous work by Paquin et al (2006) examining geographic distribution of genetic diversity in Giant Garter Snakes. We have increased sample sizes in several key populations, extended the geographic range of sampling to previously unsampled regions and included information from a total of four nuclear markers. Our expanded sampling and use of multiple microsatellite loci largely corroborates and solidifies the conclusions of previous work.

1) We can confirm that there are no deep genetic divergences across the range of the giant garter snake populations. The greatest level of sequence divergence between any two haplotypes is 5 mutational events (0.6%). Although levels of haplotype and nucleotide diversity are low, the geographic distribution of haplotypes is highly structured at all levels. There are differences in identity of mtDNA haplotypes and frequencies of mtDNA haplotypes and microsatellite allele frequencies among the San Joaquin Valley, Delta and Sacramento Valley regions. There are subtle but significant differences in both mtDNA haplotype frequencies and microsatellite allele frequencies among watersheds within a region. And finally there is subtle but significant genetic structure within the American Basin. Overall the giant garter snake can be characterized as a species with low genetic variability but highly genetically structured populations with low genetic exchange across the species range, among adjacent watersheds, and among populations within watersheds.

2) We can confirm that the population in Badger Creek is clearly genetically distinct from all other giant garter snakes populations. Paquin et al (2006) reached this conclusion based on a sample of 22 individuals, which was characterized by a three unique mtDNA types (A,B and C) found only in that population. We now have sequence data from 85 individuals all but four of which are characterized by one of these mtDNA types. We have also discovered another unique Badger Creek haplotype “M”. We now also have mtDNA sequence from and 381 individuals from other populations, the only individuals that share Badger Creek mtDNA types are four

snakes found in the other Delta population, White Slough. The Badger Creek population also shows significant  $F_{st}$  values for Microsatellite data in pairwise comparisons with five of the six other Basins for which we have microsatellite data. This population is genetically unique and represents a large proportion of the total giant garter snake genetic diversity. Although the mtDNA and microsatellite markers used here represent neutral variation in the population, there is a strong possibility that this population harbors unique adaptive genes. In recognition of its genetic value it is important to continue managing Badger Creek as a separate distinct population segment.

3) Although not as starkly unique as the Badger Creek population, giant garter snakes from the San Joaquin Valley are also clearly a distinct population segment as indicated by A) presence of several unique mtDNA haplotypes found in this basin and not in any other populations B) Significant  $F_{st}$  values for all among region comparisons of mtDNA C) significant  $F_{st}$  values in comparisons with 5 of 6 other basins for which we have microsatellite data. Paquin et al (2006) reached this conclusion based again on a sample of 22 individuals. We have augmented that sample size 49 individuals including individuals from both Mendota and Los Banos areas.

4) Within the Sacramento Valley, populations show lower levels of variation than in Badger Creek or the San Joaquin Valley, but still show geographic subdivision indicating low levels of female migration among basins. Although many of the Sacramento Valley populations share the same set of mtDNA haplotypes, there are differences among basins in the frequencies of those haplotypes. These frequency differences result in significant  $F_{st}$  estimates of and significant “among population within region” variance in all AMOVA analyses. This indicates that although these populations are genetically very similar, there is probably very limited female mediated gene flow among watersheds. Low female movement between basins would correspond to low demographic exchange among basins, meaning that demographic rescue of populations in one basin by natural migration from females from another region is unlikely. In terms of management this implies that at a minimum each basin must be managed as a separate demographic unit.

5) There is a disparity between estimates of genetic structure based on mtDNA markers and microsatellite markers. Paquin et al (2006) noted this contrast based on information from one microsatellite locus and this conclusion has held up with information from our three new loci. This contrast may be either due to general principles of population genetic or specifics of the behavior and ecology of this species. In terms of population genetics, the microsatellite loci have an effective population size of  $4N_e$  because each individual carries a chromosome bearing a copy of the gene from each of two parents and each of those chromosomes contains two copies of the gene on each branch of the chromosome. During meiosis any one of those four gene copies can end up in a gamete. In contrast the population size of maternally inherited cytoplasmic mtDNA genes is  $1 N_e$  because each individual will get the same copy of mtDNA, which is transmitted only from the mother. The result of this difference in mode of inheritance, is that more diversity is maintained in bi-parentally inherited nuclear genes than in mtDNA and cessation of gene flow between populations will show up first as differences in mtDNA. Bi-parentally inherited markers may also be show different patterns of genetic subdivision among populations due to differences in behavior of males and females. If females are philopatric and males migrate then maternally inherited mtDNA would show strong subdivision while microsatellites could show no subdivision because of male mediated gene flow among populations. At present our data are consistent with either interpretation. Distinguishing

between these two possibilities will probably require intense field observations of male and female movement patterns and long-term mark recapture studies to document rates of migration among populations. Differences in effective population size between nuclear and mtDNA would be exaggerated if giant garter snakes demographics include high variation in female reproductive success and/or if giant garter snakes reproductive strategies include multiple paternity (Hedrick 2000). Either of these biological questions could be addressed through intense field studies of survivorship and direct observation of behavior of marked individuals. Alternatively, multiple paternity can be examined using molecular techniques by genotyping a sample of mothers and their offspring and looking for evidence of multiple sires within a clutch. The level of variation in NSu3 is well suited to this type of analysis.

In addition to confirming and strengthening these previous conclusions we have also revealed new information that is relevant to management of giant garter snakes.

1) We provided the first genetic information for giant garter snakes from the Yolo Basin. This population is interesting on many levels. The population sampled here is in the Yolo Bypass and is not one of the 13 populations recognized by the USFWS draft recovery plan. The two USFWS recognized Yolo populations in Willow Slough and Liberty Farms may no longer be extant (USFWS 1999), therefore this population may represent the only extant population from the Yolo Basin. Our sample from this population consisted entirely of the common widespread haplotype D. This pattern is typical of populations that are recently colonized (Hewitt, 2005). Due to instability of habitat currently occupied by the Yolo Basin population it is possible that this population has been recently established by a small number of haplotype D animals from on the adjacent or upstream populations. The absence in this population of other less common haplotypes that occur in surrounding basins contributes to the significant  $F_{st}$  values between Yolo Basin and all other Sacramento Valley populations except for Butte Basin, which also consists mostly of haplotype D. This significant  $F_{st}$  indicates that ongoing migration into this population is probably not significant.

2) We also have far greater sampling from the American Basin than was previously available. This increased sampling has allowed us to analyze population structure from the four sub-basins defined based on the presence of potential anthropogenic barriers to gene flow (Highways I5, CA 99/70, and the Natomas Cross Canal). The four sub-basins are defined as: Middle American basin (north of the Natomas Cross Canal west of 99/70), North Natomas (North of I5 W of 99/70), East Natomas (East of 99/70) and South Natomas (South of I5). These show subtle subdivision evidenced by significant estimates of  $F_{st}$  and significant “among population within region” variance in AMOVA analyses. This level of population structure on such a small geographic scale is surprising given mobility of these snakes, the recency (in evolutionary terms <100 snake generations) of these landscape features and the persistence of conduits and canals that could still allow present day dispersal between sub-basins (see site descriptions above).

Our genetic results cannot determine if this genetic differentiation actually arose as a result of these potential barriers or are actively maintained by the barriers. It is possible that this genetic structure was present before any anthropogenic changes to the landscape, and would persist without human fragmentation of the habitat. This possibility is suggested by the significant Harpending raggedness index observed in the Middle American Basin population and borderline significant Harpending raggedness in both the north Natomas and south Natomas populations. Significant raggedness index is a signature of long-term genetic stability in a population. Long-term stability of populations in the American Basin may have lead to

development of local scale differentiation over which recent anthropogenic features were overlaid. Our genetic results cannot distinguish between natural or anthropogenic causes present day genetic structure, however is clear that regardless of the cause, there is very limited female mediated gene flow among regions within the American River basin. Low female movement among regions would correspond to low demographic exchange, meaning that demographic rescue of populations in one sub-basin by migration from females from another is unlikely. In terms of management, this implies that at a minimum each basin must be managed as a separate demographic unit.

3) We have also developed three additional microsatellite loci that will be useful for and future studies. The NSu3 locus used by Paquin and used in this study is by far the most variable of the four loci, but we also observed useful levels of polymorphism in two of the three novel loci. The fourth locus was not polymorphic in the individuals that we scored, but did show polymorphism in our non-quantitative preliminary screening. Based on the sizes of the four loci it should be possible in the future to adopt a multiplexing strategy for scoring microsatellites by combining NSU2 + TS3 and NSU3 + Te1 CA18. The non-overlapping sizes of the loci means that in the future these 4 loci can all be scored simultaneously in a single multiplexed run. This effectively halves the time and expense of scoring these loci and will greatly facilitate rapid data collection in future studies.

4) We have provided the first genetic information for the interesting White Slough population. The results from the five samples sequenced to date indicated that this population is most closely related to the other Delta population, but may itself be another well of unique genetic diversity in the species.

5) Finally, this project resulted in discovery of three new mtDNA haplotypes; Type K from White Slough, type L from SJV and type M which occurred in a reasonable number (four) of individuals from Badger Creek. This both highlights the unique genetics of each of these populations and also indicates that we have not yet uncovered all of the genetic diversity in this species.

## Summary and Conclusions

DNA sequence data from the ND4 region of the mitochondrial genome was collected from total of 466 giant garter snakes from throughout the range of the species. Microsatellite data from a subset of 98 of these snakes were also collected. A total of 13 haplotypes were identified from these sequences including three novel haplotypes. Data were analyzed using Analysis of Molecular Variance, a traditional  $F_{st}$  approach, analysis of mismatches, and descriptive molecular diversity indices. Overall the giant garter snake can be characterized as a species with low genetic variability but highly genetically structured populations with low genetic exchange across the species range, among adjacent watersheds, and among populations within watersheds. Populations in the Delta (Badger Creek and White Slough) are characterized by several unique types found nowhere else and have no mtDNA haplotype overlap with other populations examined. The San Joaquin Valley, and Sacramento Valley share some overlapping haplotypes but each has unique haplotypes and each is distinguishable by differences in haplotype frequencies. Population differentiation within these major regions is more subtle but significant and includes population substructure across very small geographic scales within the American/Natomas Basin. These genetic characteristics imply that demographic exchange even

among geographically close populations is probably rare and that demographic rescue of depleted populations by immigration from other sites is not likely and therefore when possible populations of this species should be managed as separate demographic units. Future genetic studies should focus on increasing sample sizes in several areas 1) the American/Natomas Basin area - eventually sampling from the four sub-basins of the American basin will be sufficient to allow MIGRATE analyses which will greatly strengthen the developing picture of migration barriers within the American River Basin. 2) White Slough - this population appears to be genetically distinct, has already yielded novel haplotype in the 5 samples analyzed and could harbor much more genetic diversity. 3) The genetically diverse San Joaquin region - this will allow a better understand population subdivision within this region.

## Summary of Expenditures

The major expenditures for this project were in personnel time, benefits and University overhead. The only durable equipment purchases associated with this project were a computer and associated hardware and supplies for data used for data storage and analysis.

1 Macintosh computer IMAC 17" 2.0 GHz \$1,099  
1 80 GB LaCie portable hard drive for data backup. \$103.55  
5 1gb USB drives \$89.95  
1 HP DeskJet pro printer \$49.99  
associated cables  
Total **\$1,445.73**

## References Cited

- Arevalo, E., Davis S.K., Sites J.W.. 1994. Mitochondrial DNA sequence divergence and phylogenetic relationships among eight chromosome races of the *Sceloporus grammicus* complex (Phrynosomatidae) in central Mexico. *Systematic Biology* 43:387-418.
- Baillie, J. 1996. *Thamnophis gigas*. In: IUCN 2004. 2004 IUCN Red List of Threatened Species.
- Berli P. & Felsenstein J. 1999. Maximum-likelihood estimation of migration rates and effective population numbers in two populations using a coalescent approach. *Genetics*, 152:763-773.
- Berli P. & Felsenstein J. 2001. Maximum likelihood estimation of a migration matrix and effective population sizes in N subpopulations by using a coalescent approach. *Proceedings of the National Academy of Sciences of the United States of America*, 98:4563-4568.
- Brode, J. 1988. Natural history of the giant garter snake (*Thamnophis couchii gigas*). Pages 25-28 In: *Proceedings of the conference on California herpetology*, H.F. DeListe, P.R. Brown, B. Kaufman, and B.M. McGurty (eds.). Southwestern Herpetologist's Society, Special Publication No. 4.
- Brode, J. and G. Hansen. 1992. Status and future management of the giant garter snake (*Thamnophis gigas*) within the southern American Basin, Sacramento and Sutter Counties, California. California Department of Fish and Game, Inland Fisheries Division.

- California Department of Water Resources. 1995. 1994 Biological characterization of the East Delta Properties. Prepared for the Department of Water Resources East Delta Land Management Committee. January 1995. 64 pp.
- California Natural Diversity Data Base (CNDDB). October 1, 2009. California Department of Fish and Game, Biogeographic Data Branch. Sacramento, CA.
- de Queiroz A., Lawson R. and Lemos-Espinal J.A. 2002. Phylogenetic Relationships of North American Garter Snakes (*Thamnophis*) Based on Four Mitochondrial Genes: How Much DNA Sequence Is Enough? *Molecular Phylogenetics and Evolution* 22: 315–329
- Dickert, C. 2003. Progress Report for the San Joaquin Valley giant garter snake conservation project. Los Banos Wildlife Complex, California Department of Fish and Game, 18110 Henry Miller Avenue, Los Banos, CA, 93653. 38pp + appendices.
- Excoffier, L., 2004. Patterns of DNA sequence diversity and genetic structure after a range expansion: lessons from the infinite-island model. *Mol. Ecol.* 13, 853–864.
- Excoffier L., Smouse P.E., Quattro J.M. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. *Genetics*, 131:479–491.
- Fitch, H.S. 1940. A biogeographical study of the ordinooides artenkreis of garter snakes (genus *Thamnophis*). *Univ. Calif. Publ. Zool.* 44:1-150.
- Fu, Y.-X., 1997. Statistical tests of neutrality against population growth, hitchhiking and background selection. *Genetics* 147, 915–925.
- Garner T.W.J., Pearman P.B., Gregory P.T., Tomio G, Wischniowski S.G., Hosken D.J., 2004. Microsatellite markers developed from *Thamnophis elegans* and *Thamnophis sirtalis* and their utility in three species of garter snakes. *Molecular Ecology Notes* 4: 369–371
- Harpending, R.C., 1994. Signature of ancient population growth in a low resolution mitochondrial DNA mismatch distribution. *Hum. Biol.* 66, 591–600.
- Hansen, E. C. 2001. Year 2001 Investigations of the Giant Garter Snake (*Thamnophis gigas*) at Badger Creek, Cosumnes River Preserve, December 20, 2001. Report prepared for the Nature Conservancy. Unpublished. 16 pp. plus figures.
- Hansen, E.C. 2006. Year 2005 Investigations of the Giant Garter Snake (*Thamnophis gigas*) in the Middle American Basin: Sutter County, California. Annual report for Sacramento Area Flood Control Agency, February 28, 2006. Contract No. 381. Unpublished. 28 pp + appendices.
- Hansen, E.C. 2007. Implementation of Priority 1 Recovery Tasks for the Giant Garter Snake (*Thamnophis gigas*) in Merced County, California. Report prepared for the U.S. Fish and Wildlife Service pursuant to FWS Agreement No. 802706G120, April 15, 2007.
- Hansen, G.E. 1986. Status of the giant garter snake *Thamnophis couchii gigas* (Fitch) in the southern Sacramento Valley during 1986. Final report for the California Department of Fish and Game, Standard Agreement No. C-1433. Unpublished. 28 pp.
- Hansen, G.E. 1988. Review of the Status of the giant garter snake (*Thamnophis couchii gigas*) and its supporting habitat during 1986-87. Final report for the California Department of Fish and Game, Contract C-2060. Unpublished. 31 pp.
- Hansen, G.E. 1996. Status of the giant garter snake (*Thamnophis gigas*) in the San Joaquin Valley during 1995. Final report for the California Department of Fish and Game, Standard Agreement No. FG4052IF. Unpublished. 31 pp.

- Hansen, G.E. 1998. Cherokee Canal sediment removal project post-construction giant garter snake (*Thamnophis gigas*) surveys. Final report for California Department of Water Resources, Contract No. B-81535. Unpublished. 9 pp.
- Hansen, G.E. and J.M. Brode. 1980. Status of the giant garter snake *Thamnophis couchii* gigas (Fitch). Inland Fisheries Endangered Species Special Publication 80(5):1-14. California Department of Fish and Game, Sacramento, CA.
- Hansen, G.E. and J.M. Brode. 1993. Results of relocating canal habitat of the giant garter snake (*Thamnophis gigas*) during widening of State Route 99/70 in Sacramento and Sutter Counties, California. Final report for Caltrans Interagency Agreement 03E325 (FG7550) (FY 87/88-91-92). Unpublished. 36 pp.
- Hansen, R.W. 1980. Western aquatic garter snakes in Central California: an ecological and evolutionary perspective. Masters thesis, Department of Biology, California State University, Fresno. 78 pp.
- Hedrick P.W. 2000. Genetics of Populations, 2nd ed, Jones & Bartlett Publishers, Sudburg, MA 675pp.
- Hewitt G.M. 2005. The structure of biodiversity – insights from molecular phylogeography. *Frontiers in Zoology*, 1:1-16
- McCracken G.F, Burghardt G.M., Houts S.E. 1999. Microsatellite markers and multiple paternity in the garter snake *Thamnophis sirtalis*. *Mol. Ecol.*, 8, 1475–1479.
- Paquin M.M., Wylie G.D., Routman E.J. 2006. Population structure of the giant garter snake, *Thamnophis gigas*. *Conservation Genetics* 7:25–36.
- Patterson, L. and E. Hansen. 2004. Giant Garter Snake Surveys on Bacon Island and Webb Tract and Re-evaluation of Habitat Quality in 2003. Report Prepared for the California Department of Water Resources. March 2004. Unpublished. 24 pp.
- Patterson, L. 2005. Giant Garter Snake Surveys for the In-Delta Storage Program: Year-End and Summary Report. California Department of Water Resources, 1725 23<sup>rd</sup> Street, Suite 220, Sacramento, CA 95816. March 2005. Unpublished. 18 pp.
- Phillip Williams and Associates, LTD. May 7, 1997. Analysis of Opportunities for Restoring a Natural Flood Regime on the Cosumnes River Floodplain. Prepared for The Nature Conservancy.
- Prosser M.R., Gibbs H.L., Weatherhead P.J. 1999. Microgeographic population genetic structure in the northern water snake, *Nerodia sipedon sipedon*, detected using microsatellite DNA loci. *Molecular Ecology*. 8:329–333.
- Rogers, A., 1995. Genetic evidence for a Pleistocene population explosion. *Evolution* 49:608–615
- Rogers, A.R., & Harpending, H., 1992. Population growth makes waves in the distribution of pairwise genetic differences. *Mol. Biol. Evol.* 9,:552–569.
- Rossmann, D.A. and G.R Stewart. 1987. Taxonomic reevaluation of *Thamnophis couchii* (Serpentes: Colubridae). *Occasional Papers of the Museum of Zoology, Louisiana State University, Baton Rouge*. 63:1-25.
- Sambrook J., & Russell D.W. 2001. *Molecular Cloning: a laboratory manual*. 5<sup>th</sup> edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
- Schneider S., Roessli D., Excoffier L. 2000. Arlequin version 2000. A Software for Population Genetics Data Analysis, Genetics and Biometry Laboratory, Department of Anthropology and Ecology, University of Geneva, Geneva, Switzerland.

- Stebbins, R.C. 2003. A Field Guide to Western Reptiles and Amphibians. Houghton Mifflin Company, Boston, MA. 533 p.
- Swofford, D. L. 2002. PAUP Phylogenetic Analysis Using Parsimony (and Other Methods), version 4. Sinauer Associates, Sunderland, Massachusetts.
- U.S. Fish and Wildlife Service (USFWS). 1993. Endangered and threatened wildlife and plants; determination of threatened status for the giant garter snake. Federal Register 58:54053-54066.
- U.S. Fish and Wildlife Service (USFWS). 1999. Draft Recovery Plan for the Giant Garter Snake (*Thamnophis gigas*). U.S. Fish and Wildlife Service, Portland, OR. ix + 192 pp.
- U.S. Fish and Wildlife Service (USFWS). 2006. Giant Garter Snake (*Thamnophis gigas*) 5-year Review: Summary and Evaluation. 46pp.
- Weir, B.S., Cockerham, C.C., 1984. Estimating  $F$  statistics for the analysis of population structure. *Evolution* 38:1358–1370.
- Wylie, G.D., M.L. Casazza, and J.K. Daugherty. 1997. 1996 Progress Report for the Giant Garter Snake Study. May 1, 1997. Dixon Research Station, California Science Center, USGS Biological Resources Division, Dixon, CA.
- Wylie, G.D. and M. Amarello. 2006. Results of 2006 Monitoring for giant garter snakes (*Thamnophis gigas*) for the bank protection project on the left bank of the Colusa Basin Drainage Canal in Reclamation District 108, Sacramento Riverbank Protection Project, Phase II. December 2006. Dixon Field Station; U.S. Geological Survey Western Ecological Research Center, 6924 Tremont Road; Dixon, CA 95620.
- Wylie, G.D. and M. Amarello. 2008. Surveys for the current distribution and abundance of giant garter snakes (*Thamnophis gigas*) in the southern San Joaquin Valley. USGS, BRD, WERC, Dixon Field Station. 24 pp.

## Personal Communications

- Beerli, Peter. School of Computational Science, Florida State University, Tallahassee FL  
Personal Communication via e-mail and Telephone October 2008)
- Green, Miriam. Consultant; Miriam Green Associates. Sacramento, CA
- Hansen, George E. Herpetologist. Sacramento, CA.

## Figures and Tables

Figure 1: Map showing 9 general regions where giant garter snake samples were collected. Numbers indicate 1) Butte Basin 2) Colusa Basin 3) Sutter Basin 4) American Basin [split for many analyses as Middle American Basin, North Natomas, East Natomas and South Natomas] 5) Yolo Basin 6) Badger Creek 7) White Slough 8) Los Banos 9) Mendota [Los Banos and Medota are lumped for analysis as “San Joaquin Valley”]. For numbers of samples and haplotype distribution see Table 1.

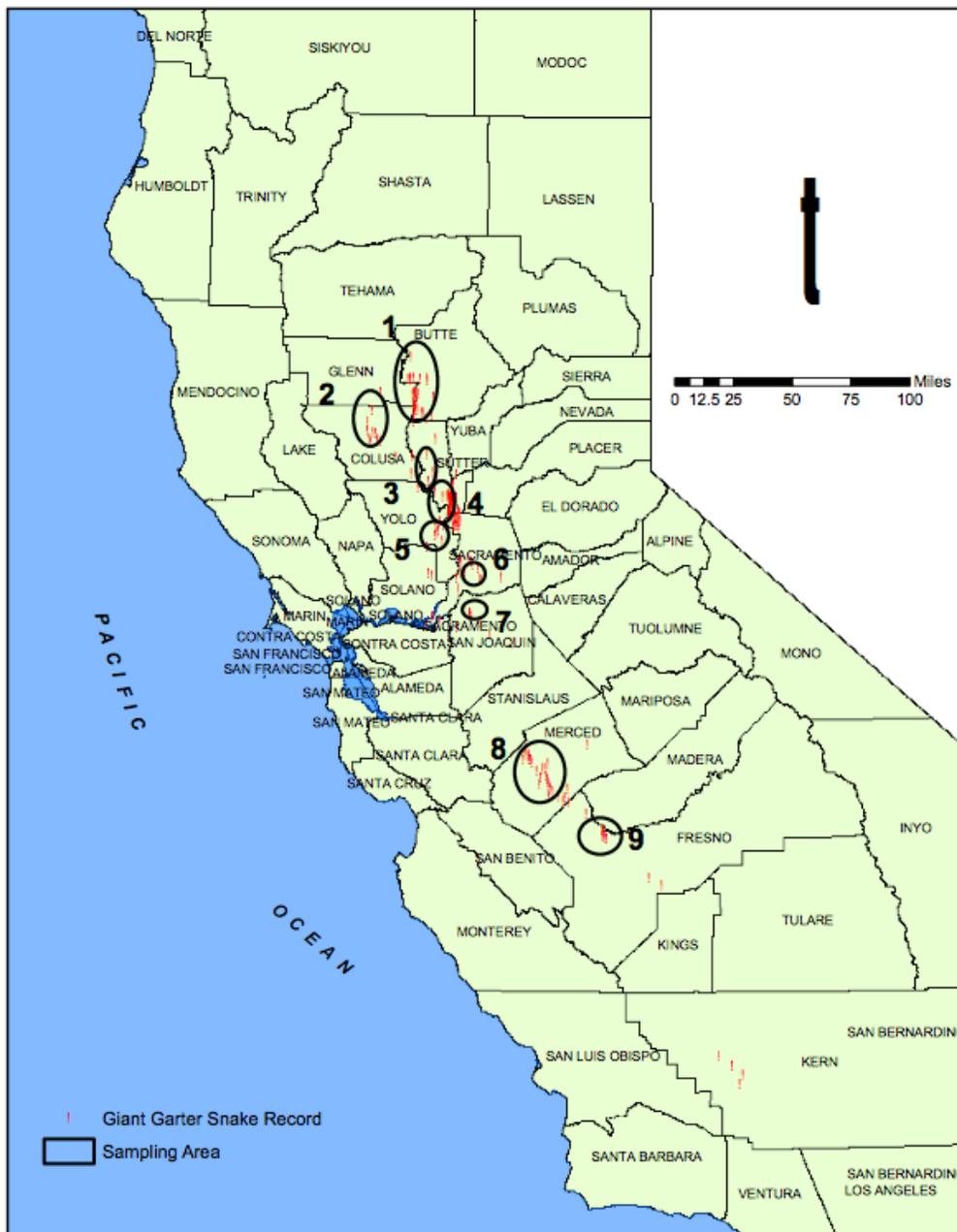


Figure 2: Distribution of giant garter snakes sampling locations in the Middle American Basin (Area B) and three regions of the Natomas Basin.

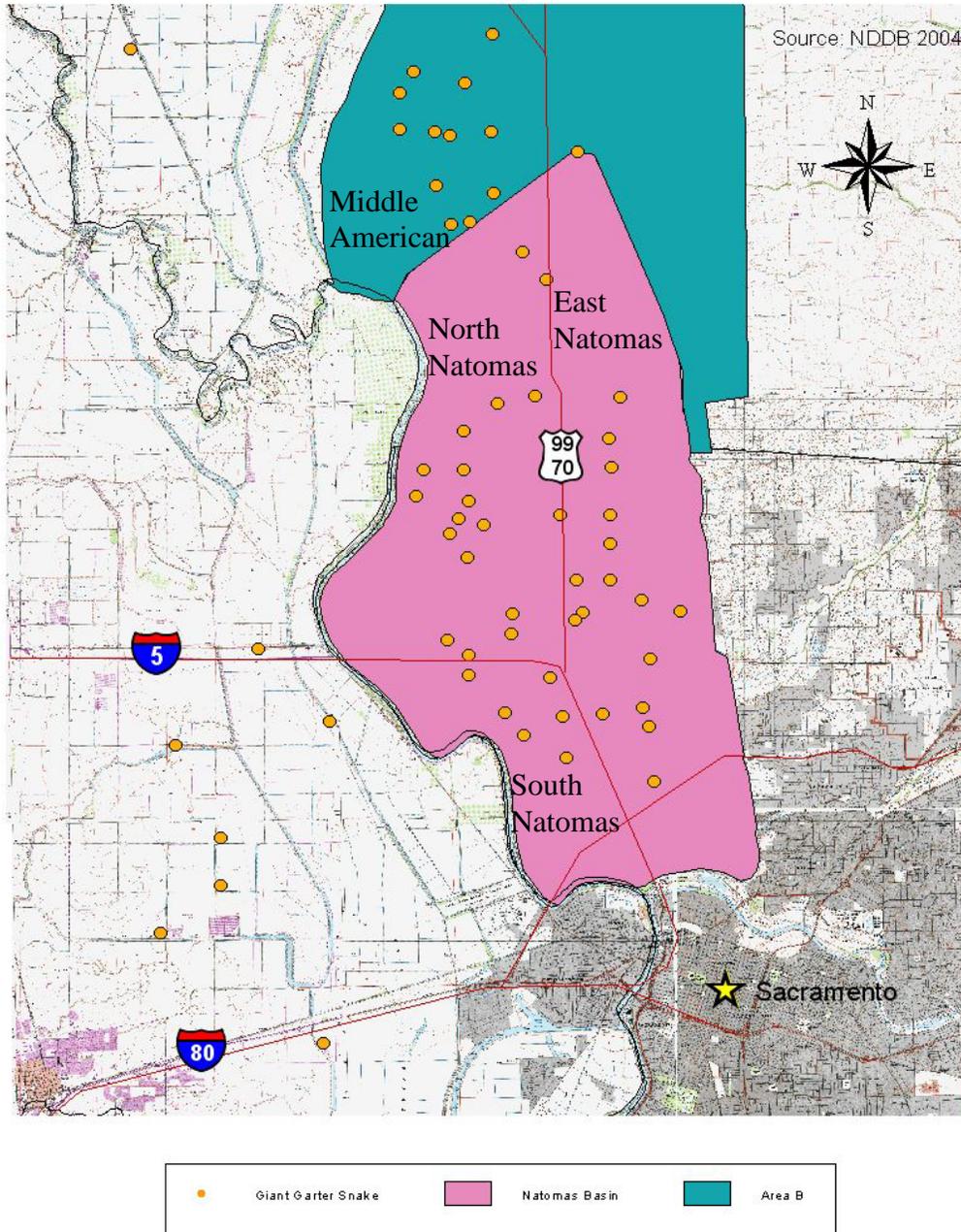
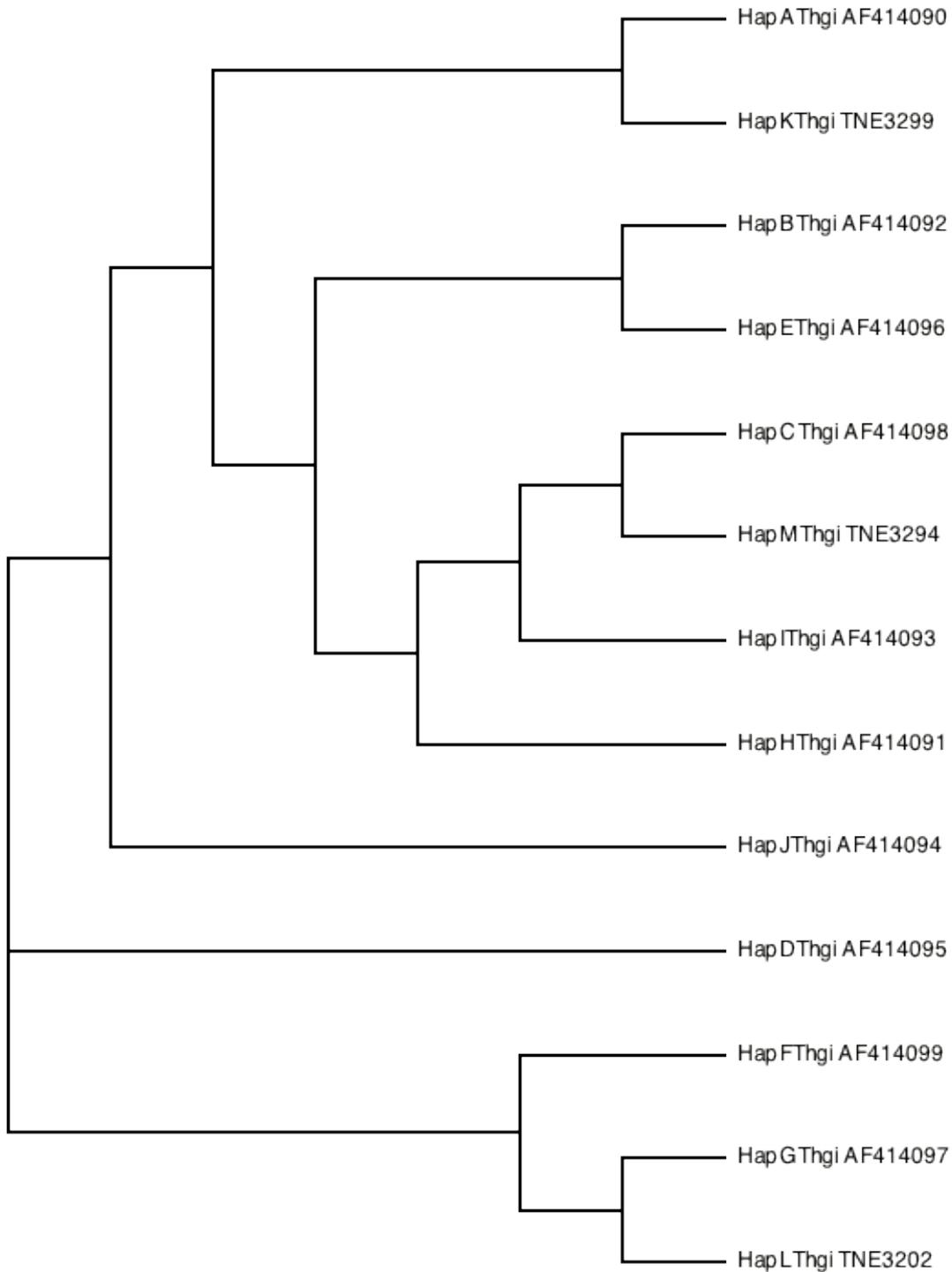




Figure 4 Phylogenetic tree depicting relationships of the 13 mitochondrial ND4 haplotypes observed in giant garter snakes in this study. Standard names for haplotypes A-M follow the naming procedure established by Paquin et al (2006). Previously described haplotypes are noted with an AF#### referring to Genbank entries for the sequence. Novel haplotypes are noted with TNE #### referring to the individual sample in which that sequence was first observed.



<b>Location</b>	<b>A</b>	<b>B</b>	<b>C</b>	<b>D</b>	<b>E</b>	<b>F</b>	<b>G</b>	<b>H</b>	<b>I</b>	<b>J</b>	<b>K</b>	<b>L</b>	<b>M</b>	<b>Total</b>
Butte Basin				40	2									<b>42</b>
Colusa Basin				24	18									<b>42</b>
Sutter Basin				20	1	2	1							<b>24</b>
Mid Amer. Basin				19		19				6				<b>44</b>
N. Natomas Basin				60		8								<b>68</b>
E. Natomas Basin				25	4	6								<b>35</b>
S. Natomas Basin				10		12				1				<b>23</b>
Badger Creek	37	27	17										4	<b>85</b>
White Slough	4										1			<b>5</b>
San Joaquin Valley				29	6			12	1			1		<b>49</b>
Yolo Basin				49										<b>49</b>
<b>Total</b>	<b>41</b>	<b>27</b>	<b>17</b>	<b>276</b>	<b>31</b>	<b>47</b>	<b>1</b>	<b>12</b>	<b>1</b>	<b>7</b>	<b>1</b>	<b>1</b>	<b>4</b>	<b>466</b>

<b>Location</b>	<b>N</b>	<b>No. of Haplotypes</b>	<b>Unique Haplotypes</b>	<b>Haplotype Diversity</b>	<b>Nucleotide Diversity</b>
Butte Basin	<b>42</b>	2	0	0.0929±0.0595	0.000112±0.000224
Colusa Basin	<b>42</b>	2	0	0.5017±0.0275	0.000606±0.000584
Sutter Basin	<b>24</b>	4	1	0.3080±0.1180	0.000394±0.000458
Mid Amer. Basin	<b>44</b>	3	1	0.6226±0.0341	0.000897±0.000752
N. Natomas Basin	<b>68</b>	2	0	0.2107±0.0599	0.000254±0.000347
E. Natomas Basin	<b>35</b>	3	0	0.4605±0.0877	0.000605±0.000763
S. Natomas Basin	<b>23</b>	2	0	0.3889±0.1644	0.000735±0.000674
Badger Creek	<b>85</b>	4	3	0.6754±0.0249	0.002679±0.001655
White Slough	<b>5</b>	2	1	0.4000±0.2373	0.000483±0.000614
San Joaquin Valley	<b>49</b>	5	3	0.5859±0.0599	0.001189±0.000909
Yolo Basin	<b>49</b>	1	0	0.0000±0.0000	0.000000±0.000000
<b>Total</b>	<b>466</b>	<b>13</b>		<b>0.6397±0.0248</b>	<b>0.001601±0.001104</b>

**Table 3:** Demographic parameters for mismatch analyses and tests of neutrality. A large and significant Raggedness Index indicates long-term stable population sizes. Significant negative Tajima's D and/or Fu's F indicate evidence for recent population expansion.

Location	N	Raggedness Index	P value	Tajima's D	P value	Fu's F	P value
Butte Basin	42	0.67151	0.79	-0.84519	0.209	-0.73067	0.122
Colusa Basin	42	<b>0.25176</b>	<b>0.020*</b>	1.63206	0.969	1.8465	0.785
Sutter Basin	24	0.23591	0.66	<b>-1.49431</b>	<b>0.047*</b>	<b>-2.38296</b>	<b>0.005*</b>
Mid American Basin	43	<b>0.17570</b>	<b>0.01*</b>	1.13514	0.884	1.18691	0.733
N. Natomas Basin	68	0.37915	0.06	0.01143	0.728	0.5035	0.352
E. Natomas Basin	35	0.16014	0.24	0.06127	0.662	0.016415	0.476
S. Natomas Basin	23	0.22542	0.06	0.27506	0.705	0.27004	0.516
Badger Creek	85	<b>0.40396</b>	<b>0.00*</b>	1.97087	0.97	4.71	0.935
White Slough	5	0.2000	0.95	-0.81650	0.291	0.09021	0.317
San Joaquin Valley	49	0.10783	0.57	-0.29665	0.422	-0.17653	0.473
Yolo Basin	49	NA	NA	0.000	1.000	0.000	NA
Total	466	0.04034	0.94	-0.45366	0.386	-2.96983	0.185

**Table 4:** Fst results for genetic structure among all 11 sampling regions. Fst values which are statistically significant with Bonferroni correction for multiple tests ( $P < 0.0045$ ) are indicated by \*\*. Values that would be significant the  $P < 0.05$  are indicated by \*.

	<b>Butte</b>	<b>Colusa</b>	<b>Sutter</b>	<b>Mid Amer.</b>	<b>North Natomas</b>	<b>East Natomas</b>	<b>South Natomas</b>	<b>Badger Creek</b>	<b>White Slough</b>	<b>SJV</b>	<b>Yolo</b>
Butte Basin	\										
Colusa Basin	0.32**	\									
Sutter Basin	0.16	0.24**	\								
Mid Amer. Basin	0.32**	0.37**	0.17**	\							
N Natomas Basin	0.072*	0.37**	0.009	0.21**	\						
E Natomas Basin	0.087**	0.19**	0.001	0.16	0.03	\					
S. Natomas Basin	0.49**	0.45**	0.28**	0.009	0.36**	0.18*	\				
Badger Creek	0.48**	0.33**	0.43**	0.49**	0.53**	0.43**	0.48**	\			
White Slough	0.93**	0.69**	0.82**	0.71**	0.88**	0.75**	0.76**	0.18*	\		
San Joaquin Valley	0.22**	0.057*	0.17*	0.28**	0.29**	0.16**	0.34**	0.31**	0.54**	\	
Yolo Basin	0.030	0.43**	0.063*	0.35**	0.087*	0.15**	0.59**	0.51**	0.98**	0.29**	\

**Table 5.** Fst results for genetic structure among the 4 sub-regions of the American River Basin. Fst values which are statistically significant with Bonferroni correction for multiple tests (P<0.0125) are indicated by \*\*

	Mid Amer.	N. Natomas	E. Natomas	S. Natomas
Mid Amer.	\			
N. Natomas	<b>0.21**</b>	\		
E. Natomas	<b>0.11**</b>	<b>0.029</b>	\	
S. Natomas	<b>0.009</b>	<b>0.34**</b>	<b>0.18**</b>	\

**Table 6.** Results for genetic structure among seven basins. Fst values for the mtDNA ND4 marker are below the diagonal, non-differentiation p values for the four microsatellite loci are above the diagonal. Values, which are statistically significant with Bonferroni correction for multiple tests (P<0.007) are indicated by \*\*. Values that would be significant the P<0.05 are indicated by \*. Microsatellite data for Sutter Basin was not sufficient for comparison indicated by “na”.

	<b>Butte</b>	<b>Colusa</b>	<b>Sutter</b>	<b>American</b>	<b>Yolo</b>	<b>Delta</b>	<b>SJV</b>
Butte Basin	\	0.11054	na	0.00162*	0.31830	0.06664*	0.02194*
Colusa Basin	0.40824**	\	na	0.00036*	0.36101	0.32863	0.12038
Sutter Basin	0.08619*	0.29307**	\	na	na	na	na
American Basin	0.13585**	0.29926**	0.03341	\	0.15255	0.11352	0.00662*
Yolo Basin	0.00000	0.43584**	0.10480*	0.14400**	\	0.78631	0.02228*
Delta	0.53182**	0.35453**	0.46853**	0.59237**	0.55248**	\	0.02608
San Joaquin Valley	0.42362**	0.11205*	0.29538**	0.35877**	0.45750**	0.27676**	+

Take non differentiated P from html files refer to these tables 6 and 7 in text change table legend.

Table 7. AMOVA results for Mitochondrial ND4 and four Microsatellite loci for all sampling sites grouped into three regions: (Sacramento Valley, Delta Basin, San Joaquin Valley). Significance of variance components among regions, among populations within regions and within populations indicates that these populations are strongly structured on multiple levels.

	Percentage of Variation			Fct	Fsc	Fst
	Among Regions	Among populations within regions	Within populations			
MtDNA	49.30 P = 0.013	6.86 P = 0.00000	43.84 P = 0.00000	0.49297	0.13536	0.56160
Microsats	3.18 P = 0.06515	3.07 P = 0.05026	93.75 P = 0.00231	0.03179	0.03172	0.06250

Table 8. AMOVA results for Mitochondrial ND4 and four Microsatellite loci for the five Sacramento Valley basins grouped into two regions: East (Butte, Sutter, American) and West (Yolo, Colusa).

	Percentage of Variation			Fct	Fsc	Fst
	Among Regions	Among populations within regions	Within populations			
MtDNA	6.94 P = 0.18	20.68 P = 0.00000	72.39 P = 0.00000	0.06938	0.22217	0.27613
Microsats	-2.14 P = 1.0	4.74 P = 0.03505	97.39 P = 0.04861	-0.02138	0.04644	0.02605

Table 9. AMOVA results for Mitochondrial ND4 and four Microsatellite loci for East Sacramento Valley basins grouped into two regions (American Basin, Natomas Basin) (Sutter Basin, Butte Basin)

Percentage of Variation						
	Among Regions	Among populations within regions	Within populations	Fct	Fsc	Fst
MtDNA	5.73 P = 0.20	13.48 P = 0.00000	80.79 P = 0.00000	0.05728	0.14300	0.19208
Microsats	1.90 P = 0.67857	3.78 P = 0.24339	94.32 P = 0.07044	0.01904	0.03853	0.05684

Table 10. AMOVA results for Mitochondrial ND4 for American River sub-basins grouped into three regions. Significant “among population within region” variation indicates that the populations are structured but negative variance components “among regions” indicate that the grouping of any populations with each other are not supported by data.

Percentage of Variation						
	Among Regions	Among populations within regions	Within populations	Fct	Fsc	Fst
MtDNA	-10.41 P = 0.672	21.73 P = 0.00000	88.68 P = 0.00000	-0.10415	0.19681	0.11316

## Appendix A:

Glossary of terms and abbreviations relating to molecular techniques and analysis.

Aligned or DNA alignment - a way of arranging the DNA sequences of to identify regions of similarity

Annealing - pairing by hydrogen bonds to a complementary sequence, forming a double-stranded polynucleotide specifically in a PCR reactions this is the step in which your primer binds to the specific region of interest. Typically 45-65 °C.

Applied Biosystems Inc. model 7300 or 310 automated DNA sequencer: Instrument built by Applied Biosystems Inc used to sequence DNA or detect length polymorphisms in small DNA fragments

Asymmetric migration rates (Nm) - MIGRATE has the capacity to detect of rate of migration from Population A→B is different from population B→A this is important in determining if one population is a demographic source supporting other populations or demographic sink, only persisting because of migrants from another population.

Base pair – in a DNA sequence a single coupling of 2 dNTP building blocks.

Coalescent-based analysis- population genetic analysis based on tracing two DNA sequences or alleles back to their common ancestor.

Denaturation – breaking hydrogen bonds in a double stranded DNA sequence, forming a single-stranded polynucleotide specifically in a PCR reactions this is the step in which essentially melt your DNA to make the region of interest available for primers to anneal to.

DNA - Deoxyribonucleic acid contains the genetic instructions for the development and functioning of all known living organisms

DNA sequencing – determining order of nucleotides in a piece of DNA. This allows us to identify the gene, understand its function and compare the DNA among different individuals.

dNTP- cytosine, guanine, adenine, thymine the building blocks from which the DNA polymerases synthesizes a new DNA strand.

Extension –time during which DNA polymerase acts to add dNTPs to a growing DNA chain. Typically 68-72 °C.

$f_{ct}$  –in AMOVA analysis this parameter summarizes the genetic variance among regions.

$f_{sc}$  –in AMOVA analysis this parameter summarizes the genetic variance within a population relative to the rest of its region,

Fst - Fixation index ( $F_{ST}$ ) is a measure of population differentiation, genetic distance, based on genetic polymorphism data, such as single-nucleotide polymorphisms (SNPs) or microsatellites.

$f_{st}$  - in AMOVA analysis this parameter summarizes genetic variance among all populations

guanidium thiocyanate “salting out” technique– See attached protocol. DNA extraction protocols must separate DNA from other cellular components, lipids proteins, etc. In this technique cells are lysed using a protease enzyme and soap then guanidium thiocyanate solution is added causing dissolved proteins and lipids to precipitate out of solution leaving DNA dissolved.

Haplotype D –the most common mtDNA sequence found among the 466 giant garter snakes

Histidine, Serine and Leucine transfer RNAs – DNA genes found in the Mitochondrial genome for small RNA molecules involved in translating DNA into Protein. Each tRNA bonds to a particular amino acid (building blocs of proteins) and recognizes a particular 3 letter DNA code. This allows the DNA sequence to be translated to a protein sequence

Leu – standard name for this primer. Used because the primer anneals with a sequence found in the Leucine tRNA

MgCl<sub>2</sub> - Magnesium Chloride. A salt added to PCR reactions because Taq Polymerase needs Mg<sup>2+</sup> ions to function properly

microliters (μl) –10<sup>-6</sup> liters. 1000 microliters = 1 mL

microsatellites – have become popular genetic markers for determining population structure and revealing differentiation among populations and individuals Microsatellites are non-coding repetitive DNA sequences composed of a variable number of tandemly repeating motifs for example agagagagagag or ctctctct

Mitochondrial DNA (mtDNA) – the DNA located in organelles called mitochondria, structures within cells that convert the energy from food into a form that cells can use. Most other DNA present in eukaryotic organisms is found in the cell nucleus.

mM millimolar– is a measure of the concentration of a solute in a solution, or of any molecular, ionic, or atomic species in a given volume. 1 mole = 1 molecular weight of a chemical in 1 liter of water... 1 mM = 1/1000<sup>th</sup> of the molecular weight of the chemical in 1 liter of water.

mtDNA haplotypes – particular sequences of mtDNA identified from an individual.

NADH dehydrogenase subunit 4 (ND4) – one of 22 protein coding genes found in the mtDNA genome. This gene codes of a protein involved in cellular respiration and is often used as a marker in population genetics studies because it is relatively easy to amplify and often

Nuclease – any enzyme the breaks down nucleic acids such as DNA.

Polymerase Chain Reaction (PCR) - a technique to amplify a single or few copies of a piece of DNA across several orders of magnitude, generating thousands to millions of copies of a particular DNA sequence.

Polymorphism – and difference in DNA sequence or length.

Primers- short DNA fragments containing sequences complementary to the target region  
Primers are required for initiation of DNA synthesis in PCR.

RNA - Ribonucleic acid is a biologically important type of molecule that consists of a long chain of nucleotide units. RNA is very similar to DNA, but differs in a few important structural details: in the cell, RNA is usually single-stranded, RNA is transcribed from DNA. RNAs are central to protein synthesis.

*Taq* DNA polymerase- Almost all PCR applications employ a heat-stable DNA polymerase, such as *Taq* polymerase, an enzyme originally isolated from the bacterium *Thermus aquaticus*. This DNA polymerase enzymatically assembles a new DNA strand from DNA building blocks, the nucleotides or dNTPs.

Thermal cycle profile - alternately heating and cooling the PCR sample to a defined series of temperature steps. These thermal cycling steps are necessary first to physically separate the two strands in a DNA double helix at a high temperature in a process called DNA melting. At a lower temperature, each strand is then used as the template in DNA synthesis by the DNA polymerase to selectively amplify the target DNA. The selectivity of PCR results from the use of primers that are complementary to the DNA region targeted for amplification under specific thermal cycling conditions.

Units of *Taq* DNA polymerase - One unit of *Taq* DNA Polymerase is defined as the amount of enzyme that will incorporate 10 nmol of dNTP into acid-insoluble material in 30 minutes at 75°C.

Whole genomic DNA – all of the DNA in the cells of an organism.

## Appendix B:

DNA sequences of ND4 haplotypes A-J from Paquin *et al.* 2006 and novel haplotypes K, L and M discovered in this study.

>HapAThgi\_AF414090

```
TTTTAAAACACTAGGAGGCTACGGTATTATCCGAATAATACAAACCCTCCCAACAATAAAAACAGACGCGTTCCTCCCATT
TATCGTCCTTGCCCTCTGAGGAGCAACACTGGCTAATCTTACCTGCTTACAACAAACAGACCTAAAATCCTTAATCGCA
TATTCATCTGTCAGTCATATAGGCCTAGTCATTTCTGCCATTATAATCCAAACACAATGAAGTCTGTCAGGAACCATAG
CCCTAATAATCGCTCACGGATTTACCTCATCAGCACTTTTCTGCCTAGCCAACACCTCCTATGAACGAACAAAACTCG
AATTTTAATCCTCACACGAGGACTACACAACATCCTTCCTATAATAACCACCTGATGACTATTAATCAATCTAATAAAC
ATTGCTACTCCCCCACCATAAACTTCACAGGCGAGTTATTAATCGCCTCATCCCTATTCAACTGATGTCCCACAACAAT
TATTATATTTGGACTATCTATACTAATCACAGCATCTTACTCTCTTCATATGTTTCCTATCAACACAAATAAACCTCACAC
CATCAAACGCCCAATTCAACCCACACATTCACGAGAACACCTACTTATACTCCTCCACACCCTACCACTTATCCTGGT
CTCCTTAAAACCCGAACTGGTAATTTAGTGTATGTAATTTAAATAAAATATCAAGCTGTGACCCTGACAATAGGAATCC
TCCTCATAACCCAGAGGGCGCAATAAGACCTGCTAACTCTTAAACCCGGAAATAACAGCCGGCCCCCTTACCAAAGG
ATAATAGTATTCCACTGGTCTTAGGCACCAAATCCT
```

>HapBThgi\_AF414092

```
TTTTAAAACACTAGGAGGCTACGGTATTATCCGAATAATACAAACCCTCCCAACAATAAAAACAGACGCGTTCCTTCCATT
TATTGTCCTTGCCCTCTGAGGAGCAACACTGGCTAATCTTACCTGCTTACAACAAACAGACCTAAAATCCTTAATCGCA
TATTCATCTGTCAGTCATATGGGCCTAGTCATTTCTGCCATTATAATCCAAACACAATGAAGTCTGTCAGGAACCATAG
CCCTAATAATCGCTCACGGATTTACCTCATCAGCACTTTTCTGCCTAGCCAACACCTCCTATGAACGAACAAAACTCG
AATTTTAATCCTCACACGAGGACTACACAACATCCTTCCTATAATAACCACCTGATGACTATTAATCAATCTAATAAAC
ATTGCTACTCCCCCACCATAAACTTCACAGGCGAGTTATTAATCGCCTCATCCCTATTCAACTGATGTCCCACAACAAT
TATTATATTTGGACTATCTATACTAATCACAGCATCTTACTCTCTTCATATGTTTCCTATCAACACAAATAAACCTCACAC
CATCAAACGCCCAATTCAACCCACACATTCACGAGAACACCTACTTATACTCCTCCACACCCTACCACTTATCCTGGT
CTCCTTAAAACCCGAACTGGTAATTTAGTGTATGTAATTTAAATAAAATATCAAGCTGTGACCCTGACAATAGGAATCC
TCCTCATAACCCAGAGGGCGCAATAAGACCTGCTAACTCTTAAACCCGGAAATAACAGCCGGCCCCCTTACCAAAGG
ATAATAGTATTCCACTGGTCTTAGGCACCAAATCCT
```

>HapCThgi\_AF414098

TTTTAAAACCTAGGAGGCTACGGTATTATCCGAATAATACAAACCCTCCCAACAATAAAAAACAGACGCGTTCCTTCCATT  
TATCGTCCTTGCCCTCTGAGGAGCAACACTGGCTAATCTTACCTGCTTACAACAACAGACCTAAAATCCTTAATCGCA  
TATTCATCTGTCAGTCATATAGGCCTAGTCATTTCTGCCATTATAATCCAAACACAATGAAGTCTGTCAGGAACCATAG  
CCCTAATAATCGCTCACGGATTTACCTCATCAGCACTTTTCTGCCTAGCCAACACCTCCTATGAACGAACAAAACTCG  
AATTTAATCCTCACACGAGGACTACACAACATCCTTCCTATAATAACCACCTGATGACTATTAATCAATCTAATAAAC  
ATTGCTACTCCCCCACCATAAACTTCACAGGCGAGTTATTAATCGCCTCATCCCTATTCAACTGATGTCCCACAACAAT  
TATTATATTTGGACTATCTATACTAATCACAGCATCTTACTCTCTTCATATGTTCCCTATCAACACAAATAAACCTCACAC  
CATCAAACGCCCAATTCAACCCACACATTCACGAGAACACCTACTTATACTCCTCCACACCCTACCACTTATCCTAAT  
CTCCTTAAAACCCGAACTGGTAATTTAGTGTATGTAATTTAAATAAAATATCAAGCTGTGACCCTGACAATAGGAATCC  
TCCTCATAACCCAGAGGGCGCAATAAGACCTGCTAACTCTTAAACCCGGAAATAACAGCCGGCCCCCTCTACCAAAGG  
ATAATAGTATTCCACTGGTCTTAGGCACCAAATCCT

>HapDThgi\_AF414095

TTTTAAAACCTAGGAGGCTACGGCATTATCCGAATAATACAAACCCTCCCAACAATAAAAAACAGACGCGTTCCTTCCATT  
TATCGTCCTTGCCCTCTGAGGAGCAACACTGGCTAATCTTACCTGCTTACAACAACAGACCTAAAATCCTTAATCGCA  
TATTCATCTGTCAGTCATATAGGCCTAGTCATTTCTGCCATTATAATCCAAACACAATGAAGTCTGTCAGGAACCATAG  
CCCTAATAATCGCTCACGGATTTACCTCATCAGCACTTTTCTGCCTAGCCAACACCTCCTATGAACGAACAAAACTCG  
AATTTAATCCTCACACGAGGACTACACAACATCCTTCCTATAATAACCACCTGATGACTATTAATCAATCTAATAAAC  
ATTGCTACTCCCCCACCATAAACTTCACAGGCGAGTTATTAATCGCCTCATCCCTATTCAACTGATGTCCCACAACAAT  
TATTATATTTGGACTATCTATACTAATCACAGCATCTTACTCTCTTCATATGTTCCCTATCAACACAAATAAACCTCACAC  
CATCAAACGCCCAATTCAACCCACACATTCACGAGAACACCTACTTATACTCCTCCACACCCTACCACTTATCCTGGT  
CTCCTTAAAACCCGAACTGGTAATTTAGTGTATGTAATTTAAATAAAATATCAAGCTGTGACCCTGACAATAGGAATCC  
TCCTCATAACCCAGAGGGCGCAATAAGACCTGCTAACTCTTAAACCCGGAAATAACAGCCGGCCCCCTCTACCAAAGG  
ATAATAGTATTCCACTGGTCTTAGGCACCAAATCCT

> HapEThgi\_AF414096

TTTTAAAACCTAGGAGGCTACGGTATTATCCGAATAATACAAACCCTCCCAACAATAAAAAACAGACGCGTTCCTTCCATT  
TATCGTCCTTGCCCTCTGAGGAGCAACACTGGCTAATCTTACCTGCTTACAACAACAGACCTAAAATCCTTAATCGCA  
TATTCATCTGTCAGTCATATAGGCCTAGTCATTTCTGCCATTATAATCCAAACACAATGAAGTCTGTCAGGAACCATAG  
CCCTAATAATCGCTCACGGATTTACCTCATCAGCACTTTTCTGCCTAGCCAACACCTCCTATGAACGAACAAAACTCG  
AATTTTAATCCTCACACGAGGACTACACAACATCCTTCCTATAATAACCACCTGATGACTATTAATCAATCTAATAAAC  
ATTGCTACTCCCCCACCATAAACTTCACAGGCGAGTTATTAATCGCCTCATCCCTATTCAACTGATGTCCCACAACAAT  
TATTATATTTGGACTATCTATACTAATCACAGCATCTTACTCTCTTCATATGTTCCCTATCAACACAAATAAACCTCACAC  
CATCAAACGCCCAATTCAACCCACACATTCACGAGAACACCTACTTATACTCCTCCACACCCTACCACTTATCCTGGT  
CTCCTTAAAACCCGAACCTGGTAATTTAGTGTATGTAATTTAAATAAAAATATCAAGCTGTGACCCTGACAATAGGAATCC  
TCCTCATAACCCAGAGGGCGCAATAAGACCTGCTAACTCTTAAACCCGGAAATAACAGCCGGCCCCCTCTACCAAAGG  
ATAATAGTATTCCACTGGTCTTAGGCACCAAATCCT

>HapFThgi\_AF414099

TTTTAAAACCTAGGAGGCTACGGCATTATCCGAATAATACAAACCCTCCCAACAATAAAAAACAGACGCGTTCCTTCCATT  
TATCGTCCTTGCCCTCTGAGGAGCAACACTGGCTAATCTTACCTGCTTACAACAACAGACCTAAAATCCTTAATCGCA  
TATTCATCTGTCAGTCATATAGGCCTAGCCATTTCTGCCATTATAATCCAAACACAATGAAGTCTGTCAGGAACCATAG  
CCCTAATAATCGCTCACGGATTTACCTCATCAGCACTTTTCTGCCTAGCCAACACCTCCTATGAACGAACAAAACTCG  
AATTTTAATCCTCACACGAGGACTACACAACATCCTTCCTATAATAACCACCTGATGACTATTAATCAATCTAATAAAC  
ATTGCTACTCCCCCACCATAAACTTCACAGGCGAGTTATTAATCGCCTCATCCCTATTCAACTGATGTCCCACAACAAT  
TATTATATTTGGACTATCTATACTAATCACAGCATCTTACTCTCTTCATATGTTCCCTATCAACACAAATAAACCTCACAC  
CATCAAACGCCCAATTCAACCCACACATTCACGAGAACACCTACTTATACTCCTCCACACCCTACCACTTATCCTGGT  
CTCCTTAAAACCCGAACCTGGTAATTTAGTGTATGTAATTTAAATAAAAATATCAAGCTGTGACCCTGACAATAGGAATCC  
TCCTCATAACCCAGAGGGCGCAATAAGACCTGCTAACTCTTAAACCCGGAAATAACAGCCGGCCCCCTCTACCAAAGG  
ATAATAGTATTCCACTGGTCTTAGGCACCAAATCCT

>HapGThgi\_AF414097

TTTTAAACTAGGAGGCTACGGCATTATCCGAATAATACAAACCCTCCCAACAATAAAAACAGACGCGTTCCTTCCATT  
TATCGTCCTTGCCCTCTGAGGAGCAACACTGGCTAATCTTACCTGCTTACAACAAACAGACCTAAAATCCTTAATCGCA  
TATTCATCTGTCAGTCATATAGGCCTAGTCATTTCTGCCATTATAATCCAAACACAATGAAGTCTGTCAGGAACCATAG  
CCCTAATAATCGCTCACGGATTTACCTCATCAGCACTTTTCTGCCTAGCCAACACCTCCTATGAACGAACAAAACTCG  
AATTTAATCCTCACACGAGGACTACACAACATCCTTCCTATAATAACCACCTGATGACTATTAATCAATCTAATAAAC  
ATTGCTACTCCCCCACCATAAACTTCACAGGCGAGTTATTAATCGCCTCATCCCTATTCAACTGATGTCCCACAACAAT  
TACTATATTTGGACTATCTATACTAATCACAGCATCTTACTCTTTCATATGTTTCCTATCAACACAAATAAACCTCACAC  
CATCAAACGCCCAATTCAACCCACACATTCACGAGAACACCTACTTATACTCCTCCACACCCTACCACTTATCCTGGT  
CTCCTTAAAACCCGAACTGGTAATTTAGTGTATGTAATTTAAATAAAATATCAAGCTGTGACCCTGACAATAGGAATCC  
TCCTCATAACCCAGAGGGCGCAATAAGACCTGCTAACTCTTAAACCCGGAAATAACAGCCGGCCCCCTTACCAAAGG  
ATAATAGTATTCCACTGGTCTTAGGCACCAAATCCT

>HapHThgi\_AF414091

TTTTAAACTAGGAGGCTACGGTATTATCCGAATAATACAAACCCTCCCAACAATAAAAACAGACGCGTTCCTTCCATT  
TATCGTCCTTGCCCTCTGAGGAGCAACACTGGCTAATCTTACCTGCTTACAACAAACAGACCTAAAATCCTTAATCGCA  
TATTCATCTGTCAGTCATATAGGCCTAGTCATTTCTGCCATTATAATCCAAACACAATGAAGTCTGTCAGGAACCATAG  
CCCTAATAATCGCTCACGGATTTACCTCATCAGCACTTTTCTGCCTAGCCAACACCTCCTATGAACGAACAAAACTCG  
AATTTAATCCTCACACGAGGACTACACAACATCCTTCCTATAATAACCACCTGATGACTATTAATCAATCTAATAAAC  
ATTGCTACTCCCCCACCATAAACTTCACAGGCGAGTTATTAATCGCCTCATCCCTATTCAACTGATGTCCCACAACAAT  
TATTATATTTGGACTATCTATACTAATCACAGCATCTTACTCTTTCATATGTTTCCTATCAACACAAATAAACCTCACAC  
CATCAAACGCCCAATTCAACCCACACATTCACGAGAACACCTACTTATACTCCTCCACACCCTACCACTTATCCTAGT  
CTCCTTAAAACCCGAACTGGTAATTTAGTGTATGTAATTTAAATAAAATATCAAGCTGTGACCCTGACAATAGGAATCC  
TCCTCATAACCCAGAGGGCGCAATAAGACCTGCTAACTCTTAAACCCGGAAATAACAGCCGGCCCCCTTACCAAAGG  
ATAATAGTATTCCACTGGTCTTAGGCACCAAATCCT

>HapIThgi\_AF414093

TTTTAAAACCTAGGAGGCTACGGTATTATCCGAATAATACAAACCCTCCCAACAATAAAAAACAGACGCGTTCCTTCCATT  
TATCGTCCTTGCCCTCTGAGGAGCAACACTGGCTAATCTTACCTGCTTACAACAACAGACCTAAAATCCTTAATCGCA  
TATTCATCTGTCAGTCATATAGGCCTAGTCATTTCTGCCATTATAATCCAAACACAATGAAGTCTGTCAGGAACCATAG  
CCCTAATAATCGCTCACGGATTTACCTCATCAGCACTTTTCTGCCTAGCCAACACCTCCTATGAACGAACAAAACTCG  
AATTTAATCCTCACACGAGGACTACACAACATCCTTCCTATAATAACCACCTGATGACTATTAATCAATCTAATAAAC  
ATTGCTACTCCCCCACCATAAACTTCACAGGCGAGTTATTAATCGCCTCATCCCTATTCAACTGATGTCCCACAACAAT  
TATTATATTTGGACTATCTATACTAATCACAGCATCTTACTCTCTTCATATGTTCCCTATCAACACAAATAAACCTCACAC  
CATCAAACGCCCAATTCAACCCACACATTCACGAGAACACCTACTTATACTCCTCCACACCCTACCACTTATCCTGAT  
CTCCTTAAAACCCGAACTGGTAATTTAGTGTATGTAATTTAAATAAAATATCAAGCTGTGACCCTGACAATAGGAATCC  
TCCTCATAACCCAGAGGGCGCAATAAGACCTGCTAACTCTTAAACCCGGAAATAACAGCCGGCCCCCTCTACCAAAGG  
ATAATAGTATTCCACTGGTCTTAGGCACCAAAAACCT

>HapJThgi\_AF414094

TTTTAAAACCTAGGAGGCTACGGCATTATCCGAATAATACAAACCCTCCCAACAATAAAAAACAGACGCGTTCCTTCCATT  
TATCGTCCTTGCCCTCTGAGGAGCAACACTGGCTAATCTTACCTGCTTACAACAACAGACCTAAAATCCTTAATCGCA  
TATTCATCTGTCAGTCATATAGGCCTAGTCATTTCTGCCATTATAATCCAAACACAATGAAGTCTGTCAGGAACCATAG  
CCCTAATAATCGCTCACGGATTTACCTCATCAGCACTTTTCTGCCTAGCCAACACCTCCTATGAACGAACAAAACTCG  
AATTTAATCCTCACACGAGGACTACACAACATCCTTCCTATAATAACCACCTGATGACTATTAATCAATCTAATAAAC  
ATTGCTACTCCCCCACCATAAACTTCACAGGCGAGTTATTAATCGCCTCATCCCTATTCAACTGATGTCCCACAACAAT  
TATTATATTTGGACTATCTATACTAATCACAGCATCTTACTCTCTTCATATGTTCCCTATCAACACAAATAAACCTCACAC  
CATCAAACGCCCAATTCAACCCACACATTCACGAGAACACCTACTTATACTCCTCCACACCCTACCACTTATCCTAGT  
CTCCTTAAAACCCGAACTGGTAATTTAGTGTATGTAATTTAAATAAAATATCAAGCTGTGACCCTGACAATAGGAATCC  
TCCTCATAACCCAGAGGGCGCAATAAGACCTGCTAACTCTTAAACCCGGAAATAACAGCCGGCCCCCTCTACCAAAGG  
ATAATAGTATTCCACTGGTCTTAGGCACCAAAAATC<sub>c</sub>T

>HapKThgi\_TNE3299

TTTTAAAACCTAGGAGGCTACGGCATTATCCGAATAATACAAACCCTCCCAACAATAAAAACAGACGCGTTCCTCCCATT  
TATCGTCCTTGCCCTCTGAGGAGCAACACTGGCTAATCTTACCTGCTTACAACAAACAGACCTAAAATCCTTAATCGCA  
TATTCATCTGTCAGTCATATAGGCCTAGTCATTTCTGCCATTATAATCCAAACACAATGAAGTCTGTCAGGAACCATAG  
CCCTAATAATCGCTCACGGATTTACCTCATCAGCACTTTTCTGCCTAGCCAACACCTCCTATGAACGAACAAAACTCT  
AATTTTAATCCTCACACGAGGACTACACAACATCCTTCCTATAATAACCACCTGATGACTATTAATCAATCTAATAAAC  
CTTGCCACCCCCCCCACCATAAACTTCACAGGCGAGTTATTAATCCCCTCATCCCTATTCAACTGATGTCCCACAACAAT  
TATTATATTTGGACTATCTATACTAATCTCAGCTTCTTACTCTCTTCATATGTTCCCTATCAACACAAATAAACCTCACACC  
ATCAAACGCCCAATTCAACCCACACATTCACGAGAACACCTACTTATACTCCTCCACACCCTACCCTTATCCTGGTCT  
CCTTAAAACCCGAACCTGGTAATTTAGTGTATGTAATTTAAATAAAAATATCAAGCTGTGACCCTGACAATAGGACTCCTC  
CTCATAACACCAGAGGGCGCAATAAGACCTGCTAACTATTAACCCGGAAATAACAGCCGGCCCCCTCTACCAAAGGAT  
AATAGTATTCCACTGGTCTTAGGCACCAAAAATCCT

>HapLThgi\_TNE3202

TTTTAAAACCTAGGAGGCTACGGCATTAGCCGAATAATACAAACCCTCCCAACAATAAAAACAGACGCGTTCCTTCCATT  
TATCGTCCTTGCCCTCTGAGGAGCAACACTGGCTAATCTTACCTGCTTACAACAAACAGACCTAAAATCCTTAATCGCA  
TATTCATCTGTCAGTCATATAGGCCTAGTCATTTCTGCCATTATAATCCAAACACAATGAAGTCTGTCAGGAACCATAG  
CCCTAATAATCGCTCACGGATTTACCTCATCAGCACTTTTCTGCCTAGCCAACACCTCCTATGAACGAACAAAACTCG  
AATTTTAATCCTCACACGAGGACTACACAACATCCTTCCTATAATAACCACCTGATGACTATTAATCAATCTAATAAAC  
ATTGCTACTCCCCCACCATAAACTTCACAGGCGAGTTATTAATCGCCTCATCCCTATTCAACTGATGTCCCACAACAAT  
TATTATATTTGGACTATCTATACTAATCACAGCATCTTACTCTCTTCATATGTTCCCTATCAACACAAATAAACCTCACAC  
CATCAAACGCCCAATTCAACCCACACATTCACGAGAACACCTACTTATACTCCTCCACACCCTACCCTTATCCTGGT  
CTCCTTAAAACCCGAACCTGGTAATTTAGTGTATGTAATTTAAATAAAAATATCAAGCTGTGACCCTGACAATAGGAATCC  
TCCTCATAACACCAGAGGGCGCAATAAGACCTGCTAACTCTTAAACCCGGAAATAACAGCCGGCCCCCTCTACCAAAGG  
ATAATAGTATTCCACTGGTCTTAGGCACCAAAAATCCT

>HapMThgi\_TNE3294

TTTTAAACTAGGAGGCTACGGTATTATCCGAATAATACAAACCCTCCCAACAATAAAAAACAGACGCGTTCCTTCCATT  
TATCGTCCTTGCCCTCTGAGGAGCAACACTGGCTAATCTTACCTGCTTACAACAACAGACCTAAAATCCTTAATCGCA  
TATTCATCTGTCAGTCATATAGGCCTAGTCATTTCTGCCATTATAATCCAAACACAATGAAGTCTGTCAGGAACCATAG  
CCCTAATAATCGCTCACGGATTTACCTCATCAGCACTTTTCTGCCTAGCCAACACCTCCTATGAACGAACAAAACTCG  
AATTTAATCCTCACACGAGGACTACACAACATCCTTCCTATAATAACCACCTGATGACTATTAATCAATCTAATAAAC  
ATTGCTACTCCCCCACCATAAACTTCACAGGCGAGTTATTAATCGCCTCATCCCTATTCAACTGATGTCCCACAACAAT  
TATTATATTTGGACTATCTATACTAATCACAGCATCTTACTCTTTCATATGTTCCCTATCAACACAAATAAACCTCACAC  
CATCAAACGCCCCAATTCAACCCACACATTCACGAGAACACCTACTTATACTCCTCCACACCCTACCACTTATCCTAAT  
CTCCTTAAAACCCGAACTGGTAATTTAGTGTATGTAATTTAAATAAAATATCAAGCTGTGACCCTGACAATAGGAATCC  
TCCTCATAACCCAGAGGGCGCAATAAGACCTGCTAACTCTTAAACCCGGAAATAACAGCCGGCCCCCTTACCAAAGG  
ATAATAGTATTCCACTGGTCTTACGCACCAAATCCT

## Appendix C:

### Protocol for DNA extraction using Guanidine Thiocyanate

#### SOLUTIONS

##### CELL LYSIS BUFFER

100 mM NaCl

100 mM Tris-Cl pH 8.0

25 mM EDTA pH 8.0

0.5% SDS

Autoclave the NaCl, Tris and EDTA before making the buffer. Filter sterilize.

##### PROTEIN PRECIPITATION SOLUTION

4M Guanidine Thiocyanate

0.1 M Tris-Cl, pH 7.5

Heat the solution to get the guanidine thiocyanate into solution. Filter sterilize, but DON'T USE cellulose acetate filters. Wes says to use cellulose nitrate. Polyethersulfone works okay too. The cellulose acetate filters dissolve.

PROTEINASE K (20 mg/mL)

RNASE A (4 mg/ml)

ISOPROPANOL

ETHANOL (70%)

1X TE (10mM Tris-Cl, pH 8.4, 1 mM EDTA) or 10 mM Tris-Cl (pH 8.4

#### Cell Lysis and RNAase Treatment

1. Place 10 mg of tissue in a 1.5 mL microcentrifuge tube containing 300 L of CELL LYSIS SOLUTION (see recipe above). Macerate the tissue as much as possible.
2. Add 1.5 uL PROTEINASE K (20 mg/mL). Mix with a brief vortex. Incubate 3-6 hours (or overnight if you really need to) at 55oC.
3. Homogenize by vortexing gently.
4. Add 1.5 uL RNASE A (4mg/mL). Incubate at 37oC for 30-45 minutes.
5. Cool the sample to room temperature.

#### Protein Precipitation

6. Add 100 m L PROTEIN PRECIPTATION SOLUTION to the cell lysate mixture. Vortex vigorously to mix the tube contents (10-20 seconds).
7. Centrifuge at the highest speed (13000 rpm) for 5 minutes. Repeat if the protein pellet is not tight.
8. Pour off or aspirate the supernatant (which contains the DNA) into a new 1.5 mL microcentrifuge tube.

#### DNA Precipitation

9. Add 300 m L 100% isopropanol. Mix by inverting gently 50 times.
10. Centrifuge at 13000 rpm for 5 minutes.
11. Pour off the supernatant, careful to leave the pellet behind.
12. Add 300 m L 70% ethanol and invert the tube several times to wash the pellet.
13. Centrifuge at 13000 rpm for 5 minutes. Pour off the supernatant.
14. Air-dry for several hours (or overnight if you really need to).

#### DNA Hydration and storage

15. Add 50-200 m L of 1x TE or 10 mM Tris-Cl, pH 8.0. Incubate overnight at room temperature or one hour at 65oC.
16. Store long-term at -20oC or -80oC, otherwise store at 4oC.

Notes:

This protocol gives results identical to those of the Puregene kit when comparing

extracted DNA using the two techniques on agarose.