EVOLUTION OF AN AQUATIC COURTSHIP PHEROMONE IN THE GENUS *TARICHA*

By

Adam Cummings

A Thesis Presented to

The Faculty of Humboldt State University

In Partial Fulfillment of the Requirements for the Degree

Master of Science in Biology

Committee Membership

Dr. John Reiss, Committee Chair

Dr. Karen Kiemnec-Tyburczy, Committee Member

Dr. Sharyn Marks, Committee Member

Dr. Bruce O’Gara, Committee Member

Dr. Jeffrey Schineller, Committee Member

Dr. Michael Mesler, Graduate Coordinator

May 2015

ABSTRACT

EVOLUTION OF AN AQUATIC COURTSHIP PHEROMONE IN THE GENUS *TARICHA*

Adam Cummings

Different loci within a multigene family often experience unique evolutionary pressures; this phenomenon is often observed in proteins involved in reproduction. Chemical communication using secreted proteinaceous pheromones is critical for reproductive success in salamanders. Differential recognition of these signals can create a reproductive barrier between lineages. In the newt genus *Taricha*, there is no discernable differentiation in courtship behavior among congeners, in spite of multiple regions of sympatry among the species. This suggests that divergence in chemical communication may serve as a reproductive barrier. The goal of this project was to study the diversification of one such pheromone--sodefrin precursor-like factor (SPF)--expressed in male cloacal glands. My objective was to investigate its role in the diversification of *Taricha* and test the role of conflicting selection modes in shaping that diversification. Males of the four species of *Taricha* (*T. granulosa* (n = 4), *T. torosa* (n = 8), *T. rivularis* (n = 2), and *T. sierrae* (n = 7)) were collected from regions of sympatry throughout California. Reverse transcriptase-PCR of abdominal gland extracts using degenerate primers resulted in isolation of 63 unique sequences of SPF. Phylogenetic analyses show that at least two SPF genomic loci (Beta-1 & Beta-2) are expressed in *Taricha,* although patterns of expression vary among the species: *T. granulosa* and *T.*

sierrae exclusively express Beta-1 while *T. torosa* and *T. rivularis* exclusively express Beta-2, except in a known hybridization zone between *T. torosa* and *T. sierrae* in the Kaweah River drainage within Kern Co., California. Here, both species present express both loci, presumably due to introgression. Beta-1 and Beta-2 SPF exhibit different patterns of protein evolution: Beta-1 showed evidence of gradual allopatric divergence with relatively few codons experiencing Darwinian selection while Beta-2 transcripts in the known hybridization zone are radically altered via alternate splicing patterns. Splicing patterns and highly concentrated codons under positive selection indicate that a key functional region for SPF binding may be in the second three-finger protein domain. The results provide insights into multi-locus signals in salamanders, specifically and the evolutionary mechanisms that shape courtship pheromones more generally.

ACKNOWLEDGEMENTS

Great thanks to my advisor, John Reiss; who repeatedly pulled me back from the Cliffs of Thesis Despair.

Great thanks to the Humboldt State University support staff: Anthony Baker, Darrell Burlison, Liz Weaver, Cheryl Satter, Marty Reed, and Lewis McCrigler. They were perpetual advocates for the lowly graduate students in the department.

Great thanks to my fellow graduate students, whose gumption to research more inspired my own.

Great thanks to Ashley, who only paused in order to shout “NEWTS!” into a walkie-talkie before diving headfirst into frigid water to catch them.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

INTRODUCTION

A Newt's View

Slowly circling, he waits. He regularly samples the water for new odors as he marches around the perimeter of the communal cow pond. Aside from the occasional ascent to the surface to gulp air, he remains at the bottom, walking weightlessly, as if exploring a lunar landscape. The less motivated in the bunch float carelessly at the surface, exposing their dark dorsums to soak up every bit of light available. They are keyed in to the stillness above and below the surface, primed for sudden bursts of activity if the need arises. Prodigious toxicity has nearly freed them from the threat of predation, but they must remain somewhat wary of the pond's resident garter snake; it has evolved immunity to the neurotoxins on the newts' skin. Our little bounding newt is not so passive about pursuing his goal. He is striving to complete the cycle that started several years earlier in this very pond: to pass his genes to the next generation, to reproduce. And his best chances of making it happen are to remain vigilant. As a female arrives at the pond, the closest males detect and swim toward her familiar odor, giving them a distinct advantage: they find themselves at the center of the mating ball. All of the males in the pond, hundreds of them, react simultaneously to her arrival. The competing males in this writhing mass are all jockeying for a position that has room for only a single winner. As the ball slowly disperses, we find that our plucky hero has won. He initiated amplexus with the female by climbing onto her back and grasping her with his disproportionately *large forelimbs. He must hold on, perhaps for many hours, in order to persuade her to take his spermatophore. Only then will he have succeeded. Yet, after several hours, the female is showing no interest in reproducing. She shakes free of his grasp in order to find a different mate. He is left to continue his circular march, completely unaware of the reason for her rejection: chemical mismatch has occurred. To her, it was as if he spoke a different language; she couldn’t recognize the molecules with which he attempted to charm her.*

Study Rationale

The elaborate courtship rituals associated with salamander reproduction have been observed and described for over a century, but the role of chemical communication in these dances remained enigmatic until advances in the last 20 years. Modern laboratory techniques have made it possible to identify the specific chemical composition of courtship pheromone molecules, and this characterization has now been completed for representatives of evolutionarily diverse groups of salamanders (Kikuyama et al., 1995; Houck et al., 2008). Additionally, genes that encode protein pheromones have been identified, and research attempting to elucidate the evolutionary histories of these signaling molecules is ongoing.

Recent discoveries regarding the biochemical structure of protein pheromones in European newts (e.g., Bocxlaer et al., 2015) have opened up the research possibilities for this type of research in North American salamandrids. In particular, North American newts in the genus *Taricha* provide an excellent system in which to explore the evolution of sex pheromones. There are only four species, representing genetically distinct and yet morphologically and behaviorally similar lineages, within *Taricha*. These four species have broad ranges with multiple points of sympatry. It is not uncommon to find interspecific pairs in amplexus in the wild, but truly hybrid animals are rare (Twitty, 1964). There is no obvious prezygotic barrier to reproduction between species, and in fact hybridization in the lab produces viable offspring in all reciprocal crosses (Twitty, 1936). Additionally, *Taricha* primarily utilize tactile and chemical communication during courtship, in contrast with European newts that have highly exaggerated sexually dimorphic traits used in mate signaling (Secondi et al., 2010). In other salamander systems, lineage diversification of chemical signals has historically been concealed by stabilizing selection on morphology and behavior (Watts et al., 2004). Since courtship behavior is also very similar across *Taricha* species, I hypothesize that chemical signals have diversified and are acting as a unimodal prezygotic barrier to interspecific reproduction. in this group as well.

Phylogeographic Context

The salamanders in the family Salamandridae are among the most studied amphibians, in part because of their ubiquity in North America, Europe, and Asia. Salamandridae comprises two main groups: 1) the “true” salamanders, and 2) the newts (Weisrock et al., 2006; Zhang et al., 2008). Newts can be further divided into three monophyletic regional groups: North American, European, and Asian, plus the more dispersed “primitive” newts (Figure 1).

The North American salamandrids comprise two genera: *Notophthalmus* in the east and *Taricha* along the west coast. This is an ancient group: North American salamandrids’ most recent common ancestor with other salamandrids is thought to have lived in Europe 43-69 mya (Zhang et al., 2008). A vicariance event led to the split between *Taricha* and *Notophthalmus* and they remain completely allopatric. This divergence occurred at least 23 mya, since a well-differentiated *Taricha oligocenica* is known from the Oligocene (Naylor, 1979). *Notophthalmus* contains three species, while

Taricha contains four recognized species: *T. granulosa, T. rivularis, T. torosa* and *T. sierrae* (Petranka, 1998; Figure 2).

Taricha granulosa possesses the most plesiomorphic traits of the four species, including features of osteology, larval coloration and egg laying pattern (Riemer, 1958; Twitty, 1942), but molecular data support *T. rivularis* as being the most basal species in the genus. All three main *Taricha* clades (*granulosa, rivularis,* and *torosa-sierrae*) diverged from one another in the middle Miocene (Kuchta and Tan, 2006b; Tan and Wake, 1995; Weisrock et al., 2006; Zhang et al., 2008).

Taricha granulosa has the largest range of the four species, extending coastally from central California to Alaska, while the other three are California endemics (Figure 3). All the populations of *T. granulosa* outside California are the result of a range expansion in the last 10,000 years, following the retreating ice sheets (Kuchta and Tan, 2005). In contrast, *T. rivularis* has the most limited range of the three endemics, being found primarily in the Coast Range in four counties in northern California; it also has extremely limited genetic structure (Kuchta and Tan, 2006a). A puzzling, disjunct yet genetically undifferentiated population of *T. rivularis* was recently described in Santa Clara County, expanding its range 130 km south (Reilly et al., 2014). Sister species *T. torosa* and *T. sierrae* occupy the Coast Range and Sierran ranges, respectively, with the notable exception of the disjunct *T. torosa* population in the southern Sierra Nevada (Kuchta, 2007).

Twitty (1942) separated *T. torosa* and *T. sierrae* into different species based on egg morphology, egg laying habits, tail morphology, larval coloration and other traits, but they were reclassified as subspecies of *T. torosa* by Stebbins (1951). This was understandable because the surprising discovery of a Sierran population of *T. torosa* did not occur until 1995 (Tan and Wake, 1995). The first geographically broad comparison of morphology and coloration pooled the disjunct Sierran *T. torosa* population with the southern *T. sierrae* population for analysis (Riemer, 1958). This exaggerated the phenotypic overlap of the two populations and therefore overstated support for the subspecies relationship. Recent molecular work (Kuchta, 2007) supports the original taxonomy proposed by Twitty, and I follow it here.

In the biogeographic scenario proposed for *T. torosa* and *T. sierrae* (Kuchta and Tan, 2006b), their common ancestor spread south along the Sierra Nevada, ultimately reaching San Diego County. A vicariance event between 7-13 million years ago then separated the coastal and montane populations, isolating the *torosa* and *sierrae* lineages (Kuchta and Tan, 2006b; Tan and Wake, 1995). In the subsequent years, *T. torosa* expanded its range north along the coast up to southern Mendocino County, establishing its modern range. Approximately 1.4-1.7 million years ago, *T. torosa* expanded its range into the southern Sierra Nevada to establish the disjunct population that is parapatric with *T. sierrae* (Kuchta and Tan, 2006b).

Currently, a remarkably small secondary contact zone between these two species exists in the Kaweah River drainage on the western edge of Sequoia National Park in Kern County, California. *Taricha torosa* south of the Main Fork and *T. sierrae* north of the Main Fork (Kuchta, 2007). This hybridization zone marks an abrupt clinal jump in molecular and morphological characters. Allozyme and mtDNA introgression is confined to a 10 km wide region, and phenotypic traits of each lineage remain almost completely distinct through the whole region. The narrow width of this introgression zone relative to the dispersal abilities of *Taricha* suggests there is some selection against hybrids (Kuchta, 2007).

The phylogeographic history of *Taricha* suggests that lineage diversification is the result of allopatric divergence punctuated by repeated secondary contacts. During these secondary contacts, lineages remained distinct. This limited gene flow is potentially due to a prezygotic barrier to hybridization resulting from intraspecific mate selection during reproduction. An example of this has been demonstrated in Australian green-eyed tree-frogs (*Litoria genimaculata*) in which a narrow secondary contact zone is maintained by reproductive character displacement of mating call patterns (Hoskin et al., 2005).

Reproduction in *Taricha*

Mating Behavior

Newts are noted for a life history in which adults return to water for extended periods, often feeding there as well as mating. All four species of *Taricha* have a biphasic life history with reproduction occurring in aquatic environments; however, the degree to which each species adapts to aquatic life varies significantly (Davis and Twitty, 1964). Among the four species, *T. granulosa* devotes the most time to the aquatic, breeding stage, with males spending November through July in the water and females spending several months in the water during the breeding season. The breeding seasons of *T. sierrae* and *T. torosa* are shorter, generally lasting from December to February – or later depending on snow melt (Miller and Robbins, 1954). The least aquatic of the four, *T. rivularis,* is usually found in moderately fast moving water and only remains long enough to mate and lay eggs (1-2 months) (Riemer, 1958). Aquatic invertebrates have been found in the stomachs of *T. granulosa* and *T. torosa,* but not in *T. rivularis,* suggesting that the latter does not feed under water (Packer, 1961).

In all four species, the males arrive to the breeding site first and courtship begins when a female enters the site. Competition among males for the female can be visually dramatic, with many males competing for a single female. Eventually one male wins out and the writhing ball of males disperses, leaving a male-female pair in amplexus. During amplexus, the male *Taricha* clasps onto the dorsal side of the female posterior to the shoulder region, using both forelimbs and hindlimbs. This position can be maintained for many hours, during which the male contracts his hindlimb muscles repeatedly, squeezing the female against his cloacal lips, while pressing and rubbing his submandibular region against the female’s rostrum (Arnold, 1972; Propper, 1991; Smith, 1941).

Eventually, the female indicates that she is receptive to the uptake of the spermatophore by lifting her head up, pressing her external nares against the male’s chin. At this point, the female is released from amplexus and the male newt swims over her shoulder slowly and in an arc, orienting perpendicular to the female as if to block her path forward. If she approaches and nudges his cloaca with her snout he will deposit his spermatophore on the substrate, and she will then pick it up with her cloaca to complete insemination. On the other hand, if she does not approach his cloaca or swims away he will attempt to reinitiate amplexus (Arnold, 1977; Davis and Twitty, 1964; Propper, 1991). These specific courtship behaviors vary little among species of *Taricha*, so they are unlikely to act as a prezygotic barrier to hybridization (Davis and Twitty, 1964).

Amplexus, in which one animal grasps and holds a mate for an extended period, is a common reproductive behavior in amphibians. Ventral amplexus is likely the primitive behavioral pattern in salamandrid courtship, though there is considerable variation in this behavior across the family (e.g., dorsal amplexus in *Taricha*, facultative use in *Notophthalmus*, prehensile tail grasping in *Euproctus*) (Houck and Arnold, 2003). Several genera have replaced amplexus behaviors entirely with tail fanning behaviors, in which the male wafts water near his cloaca toward the female (e.g., *Cynops* and *Lissotriton*). The current phylogeny (Figure 1) does not suggest a clear trend away from amplexus; instead, diverse evolutionary pressures toward improving sperm transfer probabilities have modified or eliminated mate capture behaviors across a number of lineages (Arnold, 1972, 1977). Amplexus functions in several critical roles within the lineages that have retained it (Halliday, 1990). For example, the ventral amplexus position used by several genera reduces the distance between the male's and female's cloaca at the time of spermatophore deposition, which may facilitate sperm transfer (Arnold, 1987). Competition for mates and reproductive interference has likely shaped amplexus position and behaviors in *Taricha*, especially since the operational sex ratio is typically male biased, leading to competition between multiple males during mating attempts (Propper, 1991). Males with longer and taller tails both achieve amplexus within a mating ball and attempt to disrupt an established amplectant pair more frequently than males with smaller tail morphology (Janzen and Brodie, 1989). The evolution of these morphological traits is driven by sexual selection, and the increase in relative reproductive fitness may be the result of improved defense against sexual interference by competitors during amplexus (Jones et al., 2002). Behaviors that reduce sexual interference would be under similar selective pressures.

The evolution of amplexus behaviors has also clearly been shaped to facilitate the transfer of sex pheromones to the female. During amplexus, *Taricha* are positioned to allow for contact between the male's submandibular region and the female's nares, and the male is also repeatedly pressing his cloaca against the female, perhaps in an attempt to disperse cloacal pheromones. In *Notophthalmus*, the male is positioned more anteriorly, allowing for the genial glands to contact the female's nares (Houck and Arnold, 2003). The genial glands in *Notophthalmus* express SPF, so this similarity in behaviors suggests SPF may be expressed by submandibular glands in *Taricha*, although this has not been shown (Janssenswillen et al., 2015; von Reis, 2007). In genera that have reduced or eliminated amplexus, tail fanning or wafting has replaced the direct application of pheromones.

Chemical communication

Chemical communication with sex pheromones is ancient, diverse, and widespread in caudates (for review, see Houck, 2009; Woodley, 2010, 2014). During reproduction in salamandrids, chemical communication functions in a number of ways, including: enhancing mate recognition/discovery (Toyoda et al., 1994; Twitty, 1955), repelling competition (Park and Propper, 2001, 2002), facilitating mate choice

(Osikowski, 2012), and synchronizing reproductive behaviors such as the transfer of sperm from males to females (Halliday, 1990).

In salamanders, internal fertilization is almost universal, yet copulation where sperm is directly transferred to the female is quite rare and an intromittent organ has not evolved (Kühnel et al., 2010). This is compensated for by complex courtship behaviors that help to coordinate the transfer of sperm, via a spermatophore, from the male to the substrate and then to the female. In the past 20 years, sex pheromones have been shown to play a major role in these courtship dances. In *Taricha*, one chemical responsible for coordination of behaviors necessary for successful fertilization is likely the pheromone sodefrin precursor-like factor (SPF).

Sodefrin Precursor-like Factor

Sodefrin, the first chemically-identified pheromone in amphibians, was isolated and characterized from Asian firebelly newts, *Cynops pyrrhogaster* (Kikuyama et al., 1995). Sodefrin is a decapeptide secreted by males that functions as an aquatic, female-attracting pheromone. It is produced in the abdominal[1] gland and is released through the cloaca during courtship. Male *Cynops* fan their tail toward the female during courtship, exposing the female to the compound. An orthologous peptide, silefrin, has been identified in the sister species, *C. ensicauda* (Yamamoto et al., 2000). These peptides are species-specific in amino acid sequence and attraction. A regional variant of sodefrin is

---

[1] The name “abdominal gland” for the pheromone-producing gland in salamanders’ cloacae is a historic malapropism that is now firmly established among pheromone scientists. Morphologists prefer the term dorsal gland since salamanders lack a true abdomen and the abdominal gland secretes into the posterior side of the cloacal cavity (Sever, 1991; Sever et al., 1990).

produced by *C. pyrrhogaster* in the Nara region of Japan. This variant, termed [Val8] sodefrin, contains only a single amino acid change. Interestingly, [Val8] sodefrin is able to attract female *C. pyrrhogaster* from the Nara region but not those from other regions (Iwata et al., 2005).

These decapeptides are cleaved from a 189 amino acid precursor protein, termed preprosodefrin (Nakada et al., 2007a). Preprosodefrin goes through at least two enzyme-catalyzed post-translational cleavages prior to becoming a biologically active pheromone (Nakada et al., 2007b). The evolutionary background became clear when it was discovered that a protein homologous to preprosodefrin, termed sodefrin precursor-like factor (SPF), is expressed by male plethodontids and acts as a sex pheromone (Houck et al., 2008). It is now clear that cleavage of the precursor protein to produce the bioactive peptide is a derived state of *Cynops* and the roughly 200 amino acid precursor, SPF, is the typical active form in both plethodontids and salamandrids (Janssenswillen et al., 2014). A 62 nucleotide insertion in an SPF gene caused a frame shift mutation in the C-terminus, forming preprosodefrin (Osikowski et al., 2008). Sodefrin is ultimately cleaved from this region, and since there is no sequence homology to SPF (due to the frame shift mutation) sodefrin may have evolved female-attracting properties independently of the activity of SPF (Janssenswillen et al., 2014).

Molecular clock estimates of divergence suggest SPF has remained important in reproduction for at least 150 and potentially up to 300 million years in caudates (Bocxlaer et al., 2015). At the very least, its use during courtship has been demonstrated in both salamandrids and plethodontids, families that diverged from one another approximately

90 mya (Bocxlaer et al., 2015; Houck et al., 2008; Weisrock et al., 2005). Previous work has shown that *T. granulosa* expresses up to five variants of SPF in the abdominal gland; this is a lower diversity than found in some other taxa (Janssenswillen et al., 2014).

The first 20 amino acids or so of SPF are a conserved secretory signal peptide region (Kiemnec-Tyburczy et al., 2009; Palmer et al., 2007). This is followed by two three-finger protein domains. Three-finger proteins (TFP) are a superfamily of proteins with highly conserved cysteine residues that form a conserved disulfide bonding pattern. In *Taricha* SPF, the first domain contains ten cysteines and the second contains eight cysteines. The tertiary disulfide bonding pattern of SPF in plethodontids does not match the typical TFP pattern (Leichty, 2012).

The role of SPF in successful reproduction in salamandrids is hinted at by the fact that the genus *Euproctus* appears not to express any SPF genes (Janssenswillen et al., 2014). *Euproctus* courtship behaviors and sperm transfer are unique: the spermatophore is transferred directly from the male's cloaca to the female's without being deposited on the substrate first (Arnold, 1972). This behavior precludes the need for the typical cloaca-following behavior and may explain the lack of SPF expression.

Pheromone expressing glands

The salamander cloacal region is home to a great diversity of glands, which have been extensively studied by Sever (1991, 1992; Sever et al., 1990). In male salamandrids, the cloaca comprises four glands: the abdominal gland, responsible for pheromone secretion, and three glands (the anterior ventral gland, posterior ventral gland and the pelvic gland) that are responsible for the production of the spermatophore (Sever, 1981, 1992). In *Taricha*, the final stage of courtship after amplexus involves the female first nudging the male's cloaca and then following with her nares near his cloaca as he leads her over the spermatophore. This behavior of attraction to the cloaca by the female is likely due to abdominal gland secretions (Smith, 1941).

Elevated testosterone and prolactin during the breeding season cause the abdominal gland to hypertrophy in salamandrid males (Kikuyama et al., 1975). Abdominal glands typically extend dorsally and anteriorly to the cloaca, but are less developed in *Taricha* than in other salamandrids and they do not extend anteriorly (Sever, 1991, 1992). Abdominal gland secretions are under neurotransmitter control and are released into the posterior wall of the cloacal cavity (Pool et al., 1977; Sever et al., 1990).

In addition to cloacal glands, submandibular glands are involved in courtship in *Taricha* (Smith, 1941). These glands were first described in 1941 but remained enigmatic for seventy years (Hippe et al., 2014). They are positioned similarly to the mental glands of plethodontids, but apparently evolved independently in *Taricha* (Sever and Staub, 2011). The role of submandibular glands in *Taricha* courtship is only known through behavioral observations in which the male repeatedly rubs his submandibular region across the female's external nares. The product of the gland is not known but it stains positively with periodic-acid Schiff stain, as does SPF. The behavior of directly transferring pheromones to the female's olfactory system is not unique to *Taricha* within salamandrids. As noted above, *Notophthalmus* has evolved genial or cheek glands that also express SPF and are rubbed across the female's snout during amplexus (Pool and Dent, 1977; von Reis, 2007). Interestingly, males of the primitive newt, *Pleurodeles* *waltl*, drag their cloaca across the female’s snout during courtship (Halliday, 1990). This might be an evolutionary precursor to the behaviors seen in *Taricha* and *Notophthalmus*.

Evolutionary Patterns in Chemical Communication

In contrast with auditory or visual communication where receptors are broadly tuned, slight variation in chemical communication systems can lead to a signal incompatibility with receptor cells and a complete failure of the message. The result of these functional constraints in signal-receptor binding is generally slow, conservative evolution in chemical communication systems (Symonds and Elgar, 2008). Proteinaceous pheromones can be partially freed from this limitation through gene duplication that can allow for signal variation at one locus without disrupting the functional signal (Zhang, 2003). Evolution of the system then occurs not just by mutations in individual components but also by adjusting the relative expression of each component in the composite signal (Kiemnec-Tyburczy et al., 2009; Wilburn et al., 2015).

This mode of conservative signal evolution under strong functional constraints may then be disrupted by saltational changes in the signal. Two major examples of this in caudates include the transition from dermally-received pheromones to olfactory-received pheromones in plethodontids, and the evolution of the decapeptide sodefrin by cleavage of full-length SPF (Iwata et al., 2004; Watts et al., 2004). At least one SPF gene in *Cynops* has two deletions and a 62 nucleotide insert leading to a frame shift mutation in the second TFP domain and an early stop codon. This removes six of the eight cysteines in the second domain and likely has a significant impact on conformation of the protein. Since multiple genes contribute to the chemical signal, what would have been a highly maladaptive mutation apparently survived in the population until females evolved detection and response behaviors.

Research Goals

In this study, my primary goal was to elucidate the evolutionary history of SPF within the lineages of *Taricha*, using cDNA reverse-transcribed from abdominal gland extracts sampled from all four species. Additionally, I hoped to use nucleotide alignments of SPF from other taxa help to reveal the evolutionary history of SPF in a more ancient context. Finally, I wanted to use codon-based tests for selection to determine if positive and negative Darwinian selection has influenced the diversification of SPF within *Taricha*.

The rapid, lineage-specific evolution of proteins involved in reproduction is a recognized phenomenon, both generally and within salamanders in particular (Palmer et al., 2005; Swanson and Vacquier, 2002). SPF recognition by females is likely to be important in fertilization success; consequently, positive sexual selection may drive the diversification of SPF in situations where hybrids have reduced fitness. The hybrid zone between *T. sierrae* and *T. torosa* provides a unique opportunity to study the evolution of a pheromone in the context of an atypical hybridization zone that is only a few dozen kilometers wide, suggesting selection against hybrids or hybridization (Kuchta, 2007). In situations where hybrid offspring incur a fitness cost, individuals able to differentiate the lineage of potential mates would have higher fitness; chemical cues can provide a mechanism for this recognition (Ortiz-Barrientos et al., 2004; Servedio and Noor, 2003). This fitness differential could lead to detectable positive selection in the SPF gene. For this reason, I collected animals from the hybrid zone, and from other regions of sympatry, for analysis.

In contrast to positive selection, receptor-ligand specificity should limit evolution such that negative or purifying selection may be evident across much of the pheromone molecule as well. In other words, the region of the signaling protein that interacts with the receptor is likely confined within a molecular space that effectively binds the receptor, but can rapidly evolve within that space (Watts et al., 2004). Lineage diversification of the signal can show a cyclical pattern of evolution in which particular residues are repeatedly swapped at certain sites important in signal recognition (Palmer et al., 2005). Detection of selection can allow inferences about the regions important in ligand binding. Thus, through this study, I hoped to not only trace the paths of SPF diversification but also test the role of conflicting modes of selection in the diversification of the SPF signal in *Taricha*.

MATERIALS AND METHODS

Animal and Tissue Collection

<u>Site selection</u>

Animals were typically collected from regions of known sympatry among species (Figure 3). Site A was comprises two coastal localities from which *T. granulosa*, *T. torosa*, and *T. rivularis* were collected. The two localities, Orrs Creek and Mills Creek, are tributaries of the Russian River (Mendocino Co.; Table 1). They are approximately 10 km apart and separated by a highly disturbed urban environment. Site B is much more dispersed, consisting of two sites 168 km apart. *Taricha granulosa* were collected from mine tailing ponds on the southern slope of Hayfork Bally (Trinity Co.) and *T. sierrae* were collected from Little Butte Creek in the town of Magalia (Butte Co.). Site C (Kern Co.) is of particular interest because it is the only documented, on-going hybridization event in *Taricha*. The Main Fork of the Kaweah River demarks the center of the hybridization zone, with tributaries to the south consisting mainly or exclusively of *T. torosa* while tributaries to the north consists predominantly of *T. sierrae* (Kuchta, 2007). These two species were collected from three tributaries of the Kaweah River: North Fork Kaweah River, Salt Creek, and South Fork Kaweah River. These localities correspond roughly to populations 5, 16 and 17, respectively, from Kuchta's (2007) survey of the hybridization zone.

Identification of adult *Taricha* to the species level is generally unambiguous, with a number of diagnostic characters including: skin and iris color, cloaca shape, tail morphology, and head shape (Petranka, 1998). This is true even in the documented introgression zone between *T. torosa* and *T. sierrae* in the Kaweah River drainage, with head coloration being more effective at diagnosing the genetic lineage than head shape (Kuchta, 2007). Only one individual collected from Salt Creek (Site C) had an intermediate phenotype that made identification difficult; it was closer to *T. sierrae* and was treated as such for analysis, despite being collected from a pool that contained other individuals identifiable as *T. torosa*.

Animal capture

Animals were captured and handled under California Scientific Collection Permit number SC-12055 and HSU Institutional Animal Care and Use Committee protocol numbers 12/13.B32-A and 13/14.B.61-A. Males of all four *Taricha* species in breeding condition were collected from breeding ponds or streams across California between January and March 2014. Breeding condition was assessed using a number of traits including presence in aquatic environment, evident nuptial pads, and a hypertrophied cloaca.

Shortly after removal from the breeding ponds, animals were anesthetized by submerging them in 1:2000 MS-222 (made with 40% Holtfreter solution and adjusted to pH 7.2) until loss of righting reflex. Each animal was then decapitated and double pithed. Sex was confirmed by presence of testes and then the entire cloaca was dissected out. In *Taricha*, the abdominal gland is less developed than in other salamandrids and does not extend into the abdominal cavity as in other related species (Sever, 1992). It is confined to the dorsal surface of the cloacal glands and can be differentiated by the appearance of many small tubules. At the time of dissection, a piece of the abdominal gland was separated from the cloaca and then both the abdominal gland piece and remaining cloaca were placed in a 1.5 ml Eppendorf tube with RNAlater (Life Technologies) and stored at 4°C. Upon return to the laboratory (which occurred within 24 hours) the samples were stored in a -20°C freezer until RNA extraction was performed.

Molecular Isolation of SPF Sequences

First, 15-30 mg of abdominal gland was removed from the RNAlater and homogenized in 1.5 ml Eppendorf tubes. This homogenate was used as the source material for RNA extraction using the “Mammal Tissue” protocol outlined in the GeneJET RNA extraction kit instructions (Thermo Scientific). First strand cDNA synthesis was carried out on this RNA using the Verso cDNA synthesis kit (Thermo Scientific). The synthesized cDNA was used as a template for PCR to amplify SPF transcripts using 1 µl of 10mM degenerate primers developed by M. von Reis (2007; forward: 5’-TYC TTA CTC TMY TAG CAC CAT GAG-3’, reverse: 5’-TCC TCS TCA CAA GAY CAG AC-3’) and EconoTaq DNA Polymerase (Lucigen). The cycling temperatures used for the PCR reaction were (94°C/5min; [94°C/45sec, 50°C/45sec, 72°C/45sec] x 30; 72°C/7min). Amplification products were separated using 2 percent agarose gel electrophoresis and visualized using ethidium bromide. When reactions produced products in the 300-800 bp range, the unfractionated reaction was cloned using the pGEM-T Easy vector system (Promega) following the manufacturer’s protocol with JM109 competent cells. White colonies were screened using M13 vector primers and those clones containing 300-800 bp inserts were sent off for sequencing in both forward and reverse directions (Sequetech, Mountain View, CA). Sequences that returned a positive NCBI Blast result for SPF were screened as described in following sections.

Approximately one third of the samples failed to produce amplified SPF, despite being collected concurrently with successful samples and repeated attempts at RNA extraction and reverse-transcription. The reason for this is unclear; it could be due to either pheromone low expression or RNA degradation.

Molecular Phylogenetic Analysis

The generated sequences were trimmed down to the SPF open reading frame and then aligned in CodonCode Aligner (CodonCode Corporation, www.codoncode.com). Unique sequences with nonsense mutations leading to significantly shorter or longer open reading frames (ORF) were excluded from further analysis, while length variants that were validated by multiple sequences were retained for analysis. The first 19-21 amino acids of SPF are a secretory signaling region (Janssenswillen et al., 2014). The length of the signal peptide region was determined using SignalP 4.0 (Petersen et al., 2011) and removed prior to further analysis. Nucleotide sequences were aligned with the codon-based Muscle algorithm implemented in Mega 6.0 (Tamura et al., 2013) using default parameters, and then the alignments were corrected by hand to maximize amino acid similarity.

Pairwise sequence dissimilarity at the nucleotide level was calculated in MEGA 6.0 (Tamura et al., 2013) using the Kimura 2-parameter model, and at the amino acid level using a Poisson model with uniform site substitution rates. Both methods ignore gaps, so a variable number of sites were analyzed depending on the sequences used. Distinct clades were compared by calculating between-group pairwise sequence dissimilarities as well.

Phylogenetic trees were generated using Maximum Likelihood (ML) methods in MEGA 6.0 and tree topology was validated using Bayesian inference support calculations implemented in BEAST 1.8.1 (Drummond and Rambaut, 2007; Tamura et al., 2013). For ML trees, bootstrap support (500 replicates) was calculated with the WAG + I substitution model. Bayesian inference tree sampling was performed using four Monte Carlo Markov chains that were run simultaneously for 10,000,000 generations with a 2,000,000 generation burn-in. Trees sampled every 1000 generations were used to generate the consensus tree using TreeAnnotator and then visualized in FigTree v1.4.2 (Rambaut, 2009).

In addition to the *Taricha* sequences generated in the present study, other SPF sequences were downloaded from Genbank to broaden the phylogenetic analysis. All *Lissotriton* sequences from Genbank with a query coverage greater than 75% for Beta SPF from *Taricha* were downloaded for phylogenetic analysis (n = 45)[2]. Additionally,

---

[2] NCBI GenBank Accession No.: EU526840.1-47.1, KJ402326.1-27.1, KJ402330.1-31.1, KJ402333.1-35.1, KJ402337.1-38.1, KJ402340.1, KJ402343.1-46.1, KJ402348.1, KJ402351.1, KJ402354.1, KJ402357.1, KM463823.1-24.1, KM463827.1, KM463829.1-34.1, KM463836.1, KM463844.1, KM463848.1-49.1, KM463850.1, KM463855.1, KM463859.1, KM463861.1-62.1, KM463864.1

five SPF sequences from *Taricha* were downloaded from Genbank (accession no.: KM463917-KM463921) and one unpublished sequence from von Reis (2007) was used.

Molecular Evolution

Tests of recombination

Recombination at the genomic or transcript level leads to hybrid sequences that show sequence homology to different clades at different positions in the sequence. This can lead to poorly supported phylogenies as well as misleading selection analyses (Anisimova et al., 2003). To test for recombination, Genetic Algorithm for Recombination Detection (GARD) was applied to the full sequence alignment using the HKY85 substitution model and default variation and class settings as implemented on the DataMonkey web server (Pond et al., 2006).

Codon-level selection

To test whether selection has influenced SPF in *Taricha*, a variety of codon-based tests for selection were applied to 63 unique SPF encoding cDNA sequences using programs that were part of the online HyPhy software package (Pond and Frost, 2005a). According to an Akaike information criterion based model selection test, the HKY85 substitution model best matched the data and was therefore used as the model of nucleotide evolution for all tests. A total of five tests were run (SLAC, FEL, REL, MEME, FUBAR) and the level of significance for each test was set to detect evidence of strong selection ($p < 0.1$ for SLAC, FEL, MEME; Bayes Factor > 50 for REL; Posterior Probability > 0.9 for FUBAR). See Appendix A for a discussion of the individual tests.

In addition to tests on the entire data set using the above methods, the sequences were divided into two highly supported clades, “Beta1” (n = 31) and “Beta2” (n = 32) to test each clade for selection independently. These lineages likely represent different loci (see discussion) and therefore I wanted to test if they were experiencing different selective pressures.

Branch-level selection

Tests for positive selection at both the protein and codon levels generally have to overcome the strong, consistent purifying selection. Positive selection in a single lineage may be masked by the purifying selection occurring at other locations within the phylogeny. The Branch Site REL method implemented in HyPhy allows lineages undergoing episodic diversifying selection to be identified without a priori assumptions of which lineages are experiencing selection (Pond et al., 2011). This test was applied to the entire data set as well as to clades Beta-1 and Beta-2 individually.

RESULTS

Phylogenetic History

Using reverse transcriptase-polymerase chain reaction (RT-PCR) to characterize the expression of SPF in areas of possible and known hybridization among the four species of *Taricha*, a total of 63 unique coding sequences were recovered from the three geographical regions sampled (Table 1). Five unique length variants of SPF were recovered from multiple individuals. Ten additional variable-length sequences (mostly due to nonsense mutations leading to early stop codons) were recovered from single clones, but were removed from further analysis (Table 2). The 63sequences cluster in two distinct clades (Figure 4A), which I call Beta-1 and Beta-2 based on previous terminology used to describe an early duplication event of SPF in salamanders (Janssenswillen et al., 2014). No Alpha SPF sequences (Janssenswillen et al., 2014) were found in my survey due to the primers I used.

To compare relative diversity within clades and distances between clades, the mean pairwise dissimilarity and standard errors were calculated for both nucleotides and amino acids. The mean pairwise dissimilarity between full-length Beta-1 and Beta-2 was 25.7% (±3.1%)for nucleotides and 43.3% (±6.9%) for amino acids. Beta-1 had a within-group pairwise dissimilarity of 6.6% (±0.9%) and 14.3% (±2.5%) for nucleotides and amino acids, respectively. Much of the diversity observed within Beta-1 is the result of four distinct, low-diversity clades, which are compared individually in Table 3.

Beta-2 sequences were considerably less differentiated than Beta-1 sequences. Within Beta-2, the mean pairwise dissimilarity was 0.8% (±0.2%) and 1.9% (±0.5%) percent for nucleotides and amino acids, respectively, across all sequences expressed in *T. torosa, T. rivularis* and *T. sierrae* in sites A and C. There was no geographic or species level phylogenetic structure (Figure 4) and no single-nucleotide mutations that were private to a region or species. Although Beta-2's primary amino acid sequence showed limited diversity, multiple geographically distinct length isomers were recovered.

Splice variation

Beta-1 sequences had two length variants (600 and 603 base pairs, variants 1 and 2, respectively; Table 2). Beta-2 had at least three common length variants (606, 333, and 450 base pairs, variants 3-5 respectively). Length variants 3-5 showed limited divergence from one another. The sequence dissimilarity between variants 1 and 2 were the result of the two length variants being spread evenly across clades Beta-1a and Beta-1b. *Taricha torosa* was the only species that expressed all five common length variants (Figure 5).

Species-specific Beta-form expression

Taricha granulosa was found to only express Beta-1 sequences. *Taricha sierrae* predominantly expressed Beta-1 but also expressed Beta-2 in the Sierran hybridization zone. Conversely, *T. torosa* predominantly expressed Beta-2 and only expressed Beta-1 in the same hybridization zone. *Taricha rivularis* was found to only express Beta-2 (Figure 6). Full length Beta-2 (length variant 3) sequences were found exclusively in coastal populations, while the two most common splice variants of Beta-2 (length variants 4 and 5) were found exclusively in the southern Sierra Nevada population. Expression of the length variants 1 and 2 did not have a geographic trend.

Selection Analysis

Evidence of selection

No evidence of recombination was found in SPF using GARD. This allowed the basic neighbor joining tree calculated with HyPhy to be used in selection analysis.

Three data sets were analyzed using five different tests in HyPhy to test for positive and negative Darwinian selection on SPF in *Taricha* (Pond and Frost, 2005a). The sequence alignment that included all unique coding sequences ($n = 63$) was found to have undergone positive, diversifying selection at 13 codons and negative purifying selection at 13 codons (Table 5). Since the phylogenetic evidence highly supported Beta-1 and Beta-2 being paralogs (see discussion), each lineage was also aligned independently and analyzed using the same tests. This second analysis revealed that Beta-1 had 15 codons under positive selection and 9 codons under negative selection while Beta-2 had 1 codon under positive selection and 5 under negative selection. Sites under both positive and negative selection were primarily located in the putative second three finger protein domain (**Error! Reference source not found.**).

Branch-level selection

A Branch Site REL (BSR) test on the full data set predicted that the selection detected across all sequences was highly concentrated in the Beta-1 clade ($\omega^{+} = 25.81$, p

< 0.0001), which was congruent with the finding that many more sites are under both positive and negative selection in Beta-1 than in Beta-2 (Table 6). Furthermore, when the BSR test was applied to just the Beta-1 clade, the majority of the selective pressure was found in a single clade that contained *T. granulosa* sequences from coastal and northern interior populations (Beta-1a in Figure 4; $\omega^+ = 33.91$, $p < 0.03$).

DISCUSSION

Multiple SPF Genes are evolving under Different Selection Regimes in *Taricha*

The Beta-1 and Beta-2 SPF clades almost certainly represent distinct genomic loci for SPF. Not only is the distance between the clades great (e.g., 43% amino acid dissimilarity; Figure 4) but a phylogenetic reconstruction using the available full length *Lissotriton* SPF sequences on NCBI Genbank (n = 42 including *L. helveticus*, *L. montandoni*, and *L. vulgaris*) finds the *Taricha* Beta-1 and Beta-2 clades to be sister to *Lissotriton* sequences rather than each other, suggesting that the duplication event occurred before the split of these two genera (Figure 8). The duplication event likely occurred at least 43 mya, since the common ancestor of *Lissotriton* and *Taricha* is hypothesized to have lived in Europe between 43.5-69.1 mya (Zhang et al., 2008).

Additionally, Beta-1 SPF expressed by *T. granulosa* in the coastal region has several strongly supported clades that might represent either more recent gene duplication events or evidence of introgression due to past secondary contact after allopatric divergence. The fact that the average intraspecific pairwise distances of Beta-1 length variants within *T. granulosa* (mean amino acid dissimilarity: 24.7%) are greater than the interspecific pairwise distances between *T. granulosa* and *T. sierrae* (18.6%) supports the hypothesis that there are multiple Beta-1 loci. Three *T. granulosa* individuals expressed both Beta-1a and Beta-1b, further supporting the hypothesis that the two versions represent distinct loci.

Intriguingly, although at least two distinct SPF loci were expressed by *Taricha*, each species generally expressed either one or the other but not both (Figure 9). Outside of the known hybridization zone in the Kaweah River drainage, *T. torosa* and *T. rivularis* only expressed Beta-2 and *T. granulosa* and *T. sierrae* only expressed Beta-1. This finding agrees with the results of Janssenswillen et al. (Janssenswillen et al., 2014), who didn't find any Beta-2 sequences in *T. granulosa* despite much deeper sequencing of the transcriptome than described here. This suggests that either the SPF genes or the receptors in females that detect SPF have undergone species-specific specialization. This specialization was then followed by limiting the expression of one locus. In plethodontids, a multi-pheromone cocktail is used by males, and lineage specialization of the signal involves varying the relative proportion of each pheromone (Kiemnec-Tyburczy et al., 2009). A similar mechanism may be at work in other salamandrids such as *Notophthalmus viridescens*, which expresses at least 26 variants of Beta SPF. With such a diverse signal, it seems very possible that relative levels of the length variants are vital in the overall message (Janssenswillen et al., 2014). Retaining a diverse, generalized signal would typically increase reproductive fitness by not limited mate compatibility, so it is unclear why *Taricha* has apparently lost components entirely unless lineage recognition is important in reproductive fitness and is accomplished by expressing private SPF loci.

It is also possible that Beta-1 and Beta-2 are functionally interchangeable and as *Taricha* evolved courtship behaviors that reduced the necessity of SPF, the individual components were lost stochastically within lineages without a fitness consequence. This alternative explanation for the loss of signal components across lineages does not require assumptions about the selective advantages of lineage-specific SPF diversification. The diversity of SPF expressed in other salamandrids (including *Taricha*'s sister genus, *Notophthalmus*) suggests *Taricha* has already lost or stopped expressing a number of SPF genes, and the expression patterns of Beta-1 and Beta-2 may be a continuation of this trend (Janssenswillen et al., 2014).

Evolutionary patterns of the SPF Beta-1 clade

The Beta-1 lineage is highly diversified within *T. granulosa* but the diversity is geographically dispersed. The Beta-1a clade (Figure 4) contains SPF sequences from three geographically distant populations: coastal Mendocino Co., inland Trinity Co., and central Oregon (Table 1; Oregon sequence from von Reis, 2007). This matches the hypothesized phylogeographic history of *T. granulosa*, with a recent expansion north from coastal populations in Mendocino and Humboldt Counties following the retreat of Cordilleran ice sheet approximately 10,000 years ago (Kuchta and Tan, 2005). This recent range expansion would leave geographically distant populations relatively closely related.

Taricha sierrae also expresses two forms of Beta-1, but which is version expressed is entirely dependent upon geographical region. This matches the hypothesized phylogeographic history of the species, with the southern Sierran clade being genetically distinct from the northern clade (Kuchta, 2007). They are estimated to have separated 2.6-3.4 mya.

Beta-2 evolution

Evolutionary scenarios to explain Beta-2's diversity are more enigmatic. Firstly, there is no highly supported phylogenetic structure at all in the Beta-2 sequences collected, even between *T. torosa* and *T. rivularis*. Also, there is no genetic distance in Beta-2 SPF between coastal and Sierran sequences, contradicting the mtDNA evidence of an ancient (estimated to be 1.4-1.7 mya) range expansion from coastal populations of *T. torosa* into the Sierra Nevada Mountains (Kuchta and Tan, 2006b). Curiously, the population divergences of SPF found here support those described by Tan and Wake (1995), in which the divergence distances *within* the southern Sierra Nevada *T. torosa* populations are equal to the divergence distances *between* the coastal and Sierran *T. torosa* populations. This contradicts the later results of Kuchta and Tan (2006b) who examined a longer region of the cytochrome b marker as well as allozyme variation, and found very limited genetic structure in Sierran *T. torosa* populations, which were in turn highly diverged from coastal *T. torosa*.

Discordance between mitochondrial DNA and other measures of divergence is common across numerous taxa (Toews and Brelsford, 2012). In two congeneric European newts, *Lissotriton vulgaris* and *L. montandoni*, mtDNA introgression is common despite diversified visual, behavioral and chemical courtship strategies (Halliday, 1977; Kotlik and Zavadil, 1999; Michalak and Rafiński, 1999; Michalak et al., 1997; Osikowski, 2012). Despite evidence of past and ongoing introgression in mtDNA in this species, morphology (especially within dimorphic male traits) has remained bimodal, suggesting purifying selection on mating signals (Babik and Rafiński, 2004;

Babik et al., 2003). Introgression between these two species is so extensive as to have entirely displaced *L. montandoni*'s mtDNA even though species-specific courtship signals are retained (Nadachowska and Babik, 2009). Might a similar occurrence explain the 1.4-1.7 mya divergence estimates of cytochrome b between coastal and Sierran *torosa* populations while SPF is completely undiverged?

Another possible hypothesis to explain the mtDNA/nucDNA discordance in Taricha involves the actions of the remarkable Victor Twitty, but unfortunately it does not bear out under scrutiny. Twitty was endlessly curious about the patterns of development, allopatry, speciation and hybridization within *Taricha* (Twitty, 1966). In the course of his experimentation he frequently displaced newts, both to fulfill experimental design requirements and to establish populations for future study. For example, in an attempt to study hybridization between *T. rivularis* and *T. torosa* or *T. sierrae*, approximately 50,000 *rivularis* x *torosa* hybrids and 30,000 *rivularis* x *sierrae* hybrids were released into Pepperwood Creek in northern Sonoma County (Twitty, 1955). At the time, the southern Sierran population was considered *T. sierrae*, so these "*T. sierrae*" hybrids may have in fact been Sierran *T. torosa* hybrids. Furthermore, since the maternal species in these crosses were all *T. rivularis* there would be no mtDNA introgression introduced from Sierran populations. It would also explain the lack of divergence between *T. torosa* and *T. rivularis* sequences since species boundaries were thoroughly muddled. My coastal collection site was approximately 55 km from Pepperwood Creek, which is not an unreasonable distance for *Taricha* to travel in 60 years, considering *Taricha*'s mobility during seasonal migrations. Unfortunately this neat hypothesis is not supported: the *T. sierrae* Twitty collected for the hybridization experiments came from the Chico area (Twitty, 1955) and were therefore definitely *T. sierrae.*

Without invoking human-facilitated gene flow from the Sierran population of *torosa* to the coastal population, I am left to speculate that past natural gene flow is responsible for the discordance. This gene flow event would have to either traverse riparian environments across the inhospitable San Joaquin Valley or pass through the Tehachapi Mountains wildlife corridor in southern Kern County. These mountains currently connect coastal and Sierran *Ensatina eschscholtzii* subspecies and in the past may have connected the disjunct *T. torosa* populations (Jackman and Wake, 1994; Moritz et al., 1992).

Beta-2 diversity is consistent with the hypothesis that SPF is not important in mate selection and instead helps to coordinate behavior during spermatophore transfer. Beta-2 sequences have no species-specific differentiation between *T. torosa* and *T. rivularis*. Despite living sympatrically along the coast, these two species rarely form natural hybrids, suggesting that Beta SPF is not responsible for sexual isolation.

Kaweah River Hybridization Zone

Taricha torosa and *T. sierrae* collected from the hybridization zone in the Kaweah River drainage expressed both Beta-1 and Beta-2. The lack of interspecific divergence of these sequences supports the hypothesis that this pattern of expression is the result of recent introgression of the SPF genes.

My North and South Fork Kaweah River sample sites correspond roughly with populations 5 and 17 in Kuchta's (2007) work on the hybridization zone. He found those populations to show only low levels of introgression, which is not the apparent case with the SPF genes. Either introgression has continued to expand out from the Main Fork Kaweah since Kuchta's original sampling in Spring 2001 or SPF genes have been more successful at introgression than mitochondrial lineages, which is unusual since maternally inherited mtDNA generally introgresses further than nuclear DNA (Chan et al., 2005).

Splice Variation

An interesting finding of this study is that alternate splicing plays an important role in overall and regional transcript diversity of SPF in *Taricha*. The intron/exon topology of SPF has been examined in *Desmognathus ocoee* (Leichty, 2012), and it matches the intron/exon structure of three finger protein genes in the genome of *Xenopus tropicalis* (Janssenswillen et al., 2014). All of the shorter SPF variants in *Taricha* involve a deletion that aligns well with the intron/exon boundaries found in these other species (Figure 10). Therefore, length diversity is likely due to a combination of alternate splicing sites and exon skipping.

Length variant 4 is produced by skipping the second and third exons and length variant 5 transcripts are produced by skipping the third exon (Figure 10). Additionally, the less common length variant 12 transcript is produced by skipping just the second exon. Furthermore, variants 1 and 2 vary in length by 3 nucleotides due to a gap at the interface of exons 1 and 2 and may therefore also be the product of splicing variation. This is supported by a lack of sequence differentiation.

SPF in *Taricha* comprises a 10-cysteine three finger protein (TFP) domain followed by an 8-cysteine TFP domain. Interestingly, the most common shortened splice variant (variant 4) removes the entire first TFP domain and leaves the following 8 cysteines and, presumably, their disulfide bonding pattern intact (Leichty, 2012). Removal of an entire domain may not come with a significant fitness cost because it would be less likely to disrupt disulfide bonding patterns and the protein may remain functional as a signaling molecule if the region removed was not a region involved in interactions with the female receptors.

Both of the short Beta-2 length variants were restricted to the Kaweah River sampling region. By design, sampling from this region occurred across a very narrow geographic area since introgression between *T. sierrae* and *T. torosa* was limited to approximately 5 km on either side of the Kaweah river (Kuchta, 2007). Since all three populations from this region expressed only the splice variants, it is difficult to interpret the evolutionary significance of the short variants only being found in the hybridization zone. For example, are the short variants confined to the populations nearest the Kaweah River, or does the entire disjunct Sierran *T. torosa* population express these shortened forms? If they are indeed confined to the hybridization zone it could constitute evidence that the variation is the product of hybridization.

Evidence of Selection

In all between-group comparisons of pairwise dissimilarity, the nucleotide dissimilarity is lower than the amino acid dissimilarity, reflective of the fact that -in my data set--23% of potential nucleotide mutations are synonymous while 77% change the amino acid sequence. A single nonsynonymous nucleotide mutation changes the amino acid dissimilarity three times as much, so to compare the two measures of pairwise dissimilarity is not enough to infer selection. Instead, codon-based selection tests compare the number of synonymous substitutions *per synonymous site* to the number of nonsynonymous substitutions *per nonsynonymous site*.

According to the codon-based tests for selection performed, both positive and negative selection have influenced the evolution of SPF in *Taricha*. In general, negative selection acted on residues involved in disulfide bonds and smaller residues (e.g., alanine, glycine, valine). Additionally, sites under positive selection mutated toward smaller residues as well, with a notably different pattern between codons 116-133, in which there are multiple nonconservative amino acid changes, including changes in charge and toward larger residues. This region of the protein comprises the intradomain space and the beginning of the second three-finger protein domain. All major splice variants of Beta-2 SPF contain this region, while lacking up to 90 amino acid residues from the first TFP domain. There was limited evidence of either positive or negative selection in the sequences that contained the first TFP domain. Taken together, my data suggest that the second TFP domain (especially near codons 116-133) may play an important role in the function of SPF.

The paucity of selection detected in Beta-2 is likely more a function of the limited diversity detected than of the actual evolutionary history of Beta-2 SPF. A broader geographic sampling in order to capture more diversity would likely increase the number of codons with detectable selection.

CONCLUSIONS

The evolutionary history of Beta SPF sequences in *Taricha* has several conspicuous patterns. Allopatric sequence divergence has led to several regional clades at the Beta-1 locus, with divergence of at least one of these clades driven partially by positive selection. Despite detectable positive selection, rates of evolution of SPF are relatively slow in *Taricha*. Overall, Beta-1 phylogenetic history fits well within the evolutionary history of the species lineages.

Multiple apparent recent introgressions, especially of the Beta-2 SPF locus, blur species-specific patterns of expression (Figure 9). This suggests Beta SPF is not the primary mechanism of species recognition during reproduction, particularly since *T. torosa* and *T. rivularis* express identical Beta-2 sequences and yet rarely hybridize in the wild (Davis and Twitty, 1964). The most striking diversity in Beta-2 sequences is the result of region-specific splice variants. The significance of these splice variants in relation to the hybridization zone has yet to be determined.

Several lines of evidence suggest the functional binding region of Beta SPF occurs in the second TFP domain: i)most of the codons experiencing positive selection acting on Beta-1 are located in this domain (codons 116-133) and many of these substitutions are non-conservative and ii) splice variants of Beta-2 lack codons 18-109, which comprise the entire first TFP domain, while leaving the interdomain region and second domain intact.

In conclusion, Beta SPF is ancestrally a multigenic signal but in *Taricha* that signal has apparently been reduced to a single gene within each lineage. When lineages hybridize, two loci may be expressed. These loci exhibit evolutionary patterns common in signaling systems, some of which have intriguing implications for the hybridization dynamics in the Kaweah River drainage.

Future Work

As usual, this study raises as many questions as it answers. Here are a few of the areas in which future work could help elucidate the role SPF plays in reproduction and how it has diversified within *Taricha*.

Alpha SPF Sequences

Recent publications (Bocxlaer et al., 2015; Janssenswillen et al., 2014) highlight the ancient phylogenetic history of SPF including a gene duplication event that occurred in the late Paleozoic, creating the Alpha and Beta families of SPF that continue to be coexpressed within in *Taricha* and other salamandrids. The primers used in this study do not amplify Alpha sequences and therefore, this project did not examine the evolutionary patterns of Alpha SPF, which would contribute significantly to our understanding of SPF evolution in *Taricha*.

Quantify expression levels of genes

This study shows that expression of the Beta SPF genes is species-specific. Other salamanders have been shown to use a cocktail of several proteinaceous pheromones and in which the relative abundance of each component is important in the overall signal

(Wilburn et al., 2012, 2015). At least with the Beta variant of SPF in *Taricha*, the evidence suggests this is not the pattern, but further confirmation using next-generation sequencing methods is needed.

Role of submandibular region in reproduction

The submandibular region in *Taricha* has sexually dimorphic glands that are more developed in males during the breeding season, and this correlates with the mating behavior in which the male chin is rubbed on the female's nares (Hippe et al., 2014). This is strong evidence that the glands play a role in reproduction, but multiple investigations have yet to pin down the precise nature of the secretions from these glands. Next-generation sequencing methods have been shown to be invaluable in discovering divergent forms of pheromones without a priori sequence knowledge necessary to design primers (Janssenswillen et al., 2014). These techniques should be extended to the submandibular region of *Taricha*.

Hybridization zone

A broader geographic sampling of SPF, especially in the southern Sierran hybridization zone, would further clarify the lineage diversification of the Beta-1 and Beta-2 forms of SPF. Since recombination can create hybrid chromosomes, introgression rates of individual nuclear genes are variable, with positive selection increasing introgression and negative selection limiting introgression (Barton, 2001). Determining the introgression distance of SPF across the hybridization zone compared to other nuclear loci could help elucidate the evolutionary mechanisms that led to species-specific expression of either Beta-1 or Beta-2

Above, I proposed two hypothetical mechanisms to explain the species-specific expression of either Beta-1 or Beta-2 SPF: i) past selection favored lineage-specific SPF signals or ii) stochastic loss of expression. These two hypotheses could be tested in future work with a broader sampling across the Kaweah River hybridization zone: the first would lead to limited introgression while the second would lead to equal or accelerated introgression of SPF relative to other nuclear genes. The first hypothesis would be similar to the situation demonstrated by sexually dimorphic traits in *Lissotriton*. Long term ongoing introgression has not spread genes associated with mate recognition while other genes spread throughout the population (Nadachowska and Babik, 2009). If intraspecific mate preference is dependent upon SPF, the species-specific signals would likely be maintained in the face of introgression, especially in situations where there is a fitness cost to hybridization as may be the case here. On the other hand, if SPF only helps to coordinate courtship behaviors and is not involved in mate selection, having a diverse signal may be favorable (Palmer et al., 2007) and SPF loci may be more able to spread through novel populations than neutral genes.

REFERENCES

Anisimova, M., Nielsen, R., and Yang, Z. (2003). Effect of recombination on the accuracy of the likelihood method for detecting positive selection at amino acid sites. Genetics *164*, 1229–1236.

Arnold, S.J. (1972). The evolution of courtship behaviors in salamanders. PhD Thesis. University of Michigan.

Arnold, S.J. (1977). The evolution of courtship behavior in new world salamanders with some comments on old world salamandrids. In The Reproductive Biology of Amphibians, D.H. Taylor, and S.I. Guttman, eds. (Springer US), pp. 141–183.

Arnold, S.J. (1987). The comparative ethology of courtship in salamandrid salamanders. 1. *Salamandra* and *Chioglossa*. Ethology *74*, 133–145.

Babik, W., and Rafiński, J. (2004). Relationship between morphometric and genetic variation in pure and hybrid populations of the smooth and Montandon's newt (*Triturus vulgaris* and *T. montandoni*). J. Zool. *262*, 135–143.

Babik, W., Szymura, J.M., and Rafiński, J. (2003). Nuclear markers, mitochondrial DNA and male secondary sexual traits variation in a newt hybrid zone (*Triturus vulgaris* × *T. montandoni*). Mol. Ecol. *12*, 1913–1930.

Barton, N.H. (2001). The role of hybridization in evolution. Mol. Ecol. *10*, 551–568.

Bocxlaer, I.V., Treer, D., Maex, M., Vandebergh, W., Janssenswillen, S., Stegen, G., Kok, P., Willaert, B., Matthijs, S., Martens, E., et al. (2015). Side-by-side secretion of Late Palaeozoic diverged courtship pheromones in an aquatic salamander. Proc. R. Soc. Lond. B Biol. Sci. *282*, 20142960.

Chan, K.M.A., Levin, S.A., and Gavrilets, S. (2005). Leaky prezygotic isolation and porous genomes: rapid introgression of maternally inherited DNA. Evolution *59*, 720–729.

Davis, W.C., and Twitty, V.C. (1964). Courtship behavior and reproductive isolation in the species of *Taricha* (Amphibia, Caudata). Copeia 601–610.

Drummond, A.J., and Rambaut, A. (2007). BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol. Biol. *7*, 214.

Halliday, T.R. (1977). The courtship of European newts: an evolutionary perspective. In The Reproductive Biology of Amphibians, D.H. Taylor, and S.I. Guttman, eds. (Springer US), pp. 185–232.

Halliday, T.R. (1990). The evolution of courtship behavior in newts and salamanders. Adv. Study Behav. *19*, 137–169.

Hippe, S.R., Propper, C.R., and Staub, N.L. (2014). The presence of sexually dimorphic submandibular glands in *Taricha granulosa*, the rough-skinned newt (Salmandridae). Copeia *2014*, 38–43.

Hoskin, C.J., Higgie, M., McDonald, K.R., and Moritz, C. (2005). Reinforcement drives rapid allopatric speciation. Nature *437*, 1353–1356.

Houck, L.D. (2009). Pheromone communication in amphibians and reptiles. Annu. Rev. Physiol. *71*, 161–176.

Houck, L.D., and Arnold, S.J. (2003). Courtship and mating behavior. In Phylogeny and Reproductive Biology of Urodela (Amphibia), D.M. Sever, ed. (Enfield, New Hampshire: Science Publishers), pp. 383–424.

Houck, L.D., Watts, R.A., Mead, L.M., Palmer, C.A., Arnold, S.J., Feldhoff, P.W., and Feldhoff, R.C. (2008). A candidate vertebrate pheromone, SPF, increases female receptivity in a salamander. Chem. Signals Vertebr. 11 213–221.

Iwata, T., Conlon, J.M., Nakada, T., Toyoda, F., Yamamoto, K., and Kikuyama, S. (2004). Processing of multiple forms of preprosodefrin in the abdominal gland of the red-bellied newt *Cynops pyrrhogaster*: regional and individual differences in preprosodefrin gene expression. Peptides *25*, 1537–1543.

Iwata, T., Ishizuka, Y., Nakada, T., Toyoda, F., Yamamoto, K., Conlon, J.M., and Kikuyama, S. (2005). Regionally specific occurrence of an active sodefrin variant in the red-bellied newt. Ann. N. Y. Acad. Sci. *1040*, 351–353.

Jackman, T.R., and Wake, D.B. (1994). Evolutionary and historical analysis of protein variation in the blotched forms of salamanders of the *Ensatina* complex (Amphibia: Plethodontidae). Evolution *48*, 876–897.

Janssenswillen, S., Vandebergh, W., Treer, D., Willaert, B., Maex, M., Bocxlaer, I.V., and Bossuyt, F. (2014). Origin and diversification of a salamander sex pheromone system. Mol. Biol. Evol. 472–480.

Janssenswillen, S., Willaert, B., Treer, D., Vandebergh, W., Bossuyt, F., and Bocxlaer, I.V. (2015). High pheromone diversity in the male cheek gland of the red-spotted newt Notophthalmus viridescens (Salamandridae). BMC Evol. Biol. *15*, 54.

Janzen, F.J., and Brodie, E.D. (1989). Tall tails and sexy males: sexual behavior of rough-skinned newts (*Taricha granulosa*) in a natural breeding pond. Copeia *1989*, 1068–1071.

Jones, A.G., Arguello, J.R., and Arnold, S.J. (2002). Validation of Bateman's principles: a genetic study of sexual selection and mating patterns in the rough–skinned newt. Proc. R. Soc. Lond. B Biol. Sci. *269*, 2533–2539.

Kiemnec-Tyburczy, K.M., Watts, R.A., Gregg, R.G., von Borstel, D., and Arnold, S.J. (2009). Evolutionary shifts in courtship pheromone composition revealed by EST analysis of plethodontid salamander mental glands. Gene *432*, 75–81.

Kikuyama, S., Nakano, R., and Yasumasu, I. (1975). Synergistic action of prolactin and androgen on the cloacal glands of the newt. Comp. Biochem. Physiol. A Physiol. *51*, 823–826.

Kikuyama, S., Toyoda, F., Ohmiya, Y., Matsuda, K., Tanaka, S., and Hayashi, H. (1995). Sodefrin: a female-attracting peptide pheromone in newt cloacal glands. Science *267*, 1643–1645.

Kotlik, P., and Zavadil, V. (1999). Natural hybrids between the newts *Triturus montandoni* and *T. vulgaris*: morphological and allozyme data evidence of recombination between parental genomes. Folia Zool. *48*, 211–218.

Kuchta, S.R. (2007). Contact zones and species limits: hybridization between lineages of the California newt, *Taricha torosa*, in the southern Sierra Nevada. Herpetologica *63*, 332–350.

Kuchta, S.R., and Tan, A.-M. (2005). Isolation by distance and post-glacial range expansion in the rough-skinned newt, *Taricha granulosa*. Mol. Ecol. *14*, 225–244.

Kuchta, S.R., and Tan, A.M. (2006a). Limited genetic variation across the range of the red-bellied newt, *Taricha rivularis*. J. Herpetol. *40*, 561–565.

Kuchta, S.R., and Tan, A.N.M. (2006b). Lineage diversification on an evolving landscape: phylogeography of the California newt, *Taricha torosa* (Caudata: Salamandridae). Biol. J. Linn. Soc. *89*, 213–239.

Kühnel, S., Reinhard, S., and Kupfer, A. (2010). Evolutionary reproductive morphology of amphibians: an overview. Bonn Zool. Bull. 119–126.

Leichty, K.A. (2012). Co-option and adaptation of novel gene duplications for pheromone activity in a dusky salamander. Masters Thesis. University of Louisville.

Li, W. (1997). Molecular evolution. (Sunderland, MA: Sinauer Associates Incorporated).

Michalak, P., and Rafiński, J. (1999). Sexual isolation between two newt species,*Triturus vulgaris* and *T. montandoni* (Amphibia, Urodela, Salamandridae). Biol. J. Linn. Soc. *67*, 343–352.

Michalak, P., Grzesik, J., and Rafiński, J. (1997). Tests for sexual incompatibility between two newt species, *Triturus vulgaris* and *Triturus montandoni*: no-choice mating design. Evolution *51*, 2045–2050.

Miller, M.R., and Robbins, M.E. (1954). The reproductive cycle in *Taricha torosa* (*Triturus torosus*). J. Exp. Zool. *125*, 415–445.

Moritz, C., Schneider, C.J., and Wake, D.B. (1992). Evolutionary relationships within the *Ensatina eschscholtzii* complex confirm the ring species interpretation. Syst. Biol. *41*, 273–291.

Murrell, B., Wertheim, J.O., Moola, S., Weighill, T., Scheffler, K., and Kosakovsky Pond, S.L. (2012). Detecting individual sites subject to episodic diversifying selection. PLoS Genet *8*, e1002764.

Murrell, B., Moola, S., Mabona, A., Weighill, T., Sheward, D., Pond, S.L.K., and Scheffler, K. (2013). FUBAR : A Fast, Unconstrained Bayesian AppRoximation for inferring selection. Mol. Biol. Evol. 1196–1205.

Nadachowska, K., and Babik, W. (2009). Divergence in the face of gene flow: the case of two newts (Amphibia: Salamandridae). Mol. Biol. Evol. *26*, 829–841.

Nakada, T., Toyoda, F., Iwata, T., Yamamoto, K., Conlon, J.M., Kato, T., and Kikuyama, S. (2007a). Isolation, characterization and bioactivity of a region-specific pheromone, [Val8]sodefrin from the newt *Cynops pyrrhogaster*. Peptides *28*, 774–780.

Nakada, T., Ishizuka, Y., Iwata, T., Toyoda, F., Kato, T., Conlon, J.M., and Kikuyama, S. (2007b). Evidence for processing enzymes in the abdominal gland of the newt, *Cynops pyrrhogaster*, that generate sodefrin from its biosynthetic precursor. Zoolog. Sci. *24*, 521–524.

Naylor, B.G. (1979). A new species of *Taricha* (Caudata: Salamandridae), from the Oligocene John Day Formation of Oregon. Can. J. Earth Sci. *16*, 970–973.

Ortiz-Barrientos, D., Counterman, B.A., and Noor, M.A.F. (2004). The genetics of speciation by reinforcement. PLoS Biol *2*, e416.

Osikowski, A. (2012). Asymmetric Female Preferences for Courtship Pheromones in Two Closely-related Newt Species, the smooth newt (*Lissotriton vulgaris*) and the Carpathian (*L. montandoni*) (Salamandridae). Zoolog. Sci. *29*, 390–395.

Osikowski, A., Babik, W., Grzmil, P., and Szymura, J.M. (2008). Multiple sex pheromone genes are expressed in the abdominal glands of the smooth newt (*Lissotriton vulgaris*) and Montandon's newt (*L. montandoni*) (Salamandridae). Zoolog. Sci. *25*, 587–592.

Packer, W.C. (1961). Feeding behavior in adult *Taricha*. Copeia *1961*, 351–352.

Palmer, C.A., Watts, R.A., Gregg, R.G., McCall, M.A., Houck, L.D., Highton, R., and Arnold, S.J. (2005). Lineage-specific differences in evolutionary mode in a salamander courtship pheromone. Mol. Biol. Evol. *22*, 2243–2256.

Palmer, C.A., Watts, R.A., Houck, L.D., Picard, A.L., and Arnold, S.J. (2007). Evolutionary replacement of components in a salamander pheromone signaling complex: more evidence for phenotypic-molecular decoupling. Evolution *61*, 202–215.

Park, D., and Propper, C.R. (2001). Repellent function of male pheromones in the red-spotted newt. J. Exp. Zool. *289*, 404–408.

Park, D., and Propper, C.R. (2002). The olfactory organ is activated by a repelling pheromone in the red-spotted newt *Notophthalmus viridescens*. Korean J. Biol. Sci. *6*, 233–237.

Petersen, T.N., Brunak, S., von Heijne, G., and Nielsen, H. (2011). SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat. Methods *8*, 785–786.

Petranka, J.W. (1998). Salamanders of the United States and Canada (Washington: Smithsonian Institution).

Pond, S.L.K., and Frost, S.D.W. (2005a). Datamonkey: rapid detection of selective pressure on individual sites of codon alignments. Bioinformatics *21*, 2531–2533.

Pond, S.L.K., and Frost, S.D.W. (2005b). Not so different after all: a comparison of methods for detecting amino acid sites under selection. Mol. Biol. Evol. *22*, 1208–1222.

Pond, S.L.K., Posada, D., Gravenor, M.B., Woelk, C.H., and Frost, S.D.W. (2006). Automated phylogenetic detection of recombination using a genetic algorithm. Mol. Biol. Evol. *23*, 1891–1901.

Pond, S.L.K., Murrell, B., Fourment, M., Frost, S.D.W., Delport, W., and Scheffler, K. (2011). A random effects branch-site model for detecting episodic diversifying selection. Mol. Biol. Evol. 3033–3043.

Pool, T.B., and Dent, J.N. (1977). The ultrastructure and the hormonal control of product synthesis in the hedonic glands of the red-spotted newt *Notophthalmus viridescens*. J. Exp. Zool. *201*, 177–201.

Pool, T.B., Dent, J.N., and Kemphues, K. (1977). Neural regulation of product discharge from the hedonic glands of the red-spotted newt, *Notophthalmus viridescens*. J. Exp. Zool. *201*, 203–219.

Propper, C.R. (1991). Courtship in the rough-skinned newt *Taricha granulosa*. Anim. Behav. *41*, 547–554.

Rambaut, A. (2009). FigTree V1.3.1: Tree figure drawing tool (Website: http://tree.bio.ed.ac.uk/software/figtree).

Reilly, S.B., Portik, D.M., Koo, M.S., and Wake, D.B. (2014). Discovery of a new, disjunct population of a narrowly distributed salamander (*Taricha rivularis*) in California presents conservation challenges. J. Herpetol. *48*, 371–379.

Von Reis, M. (2007). Evolution of sodefrin precursor pheromones in salamandrid newts. Masters Thesis. Oregon State University.

Riemer, W.J. (1958). Variation and systematic relationships within the salamander genus *Taricha*. Univ. Calif. Publ. Zool. *56*, 301–390.

Secondi, J., Johanet, A., Pays, O., Cazimajou, F., Djalout, Z., and Lemaire, C. (2010). Olfactory and visual species recognition in newts and their role in hybridization. Behaviour *147*, 1693–1712.

Servedio, M.R., and Noor, M.A.F. (2003). The role of reinforcement in speciation: theory and data. Annu. Rev. Ecol. Evol. Syst. *34*, 339–364.

Sever, D.M. (1981). Cloacal Anatomy of Male Salamanders in the Families Ambystomatidae, Salamandridae, and Plethodontidae. Herpetologica *37*, 142–155.

Sever, D.M. (1991). Comparative anatomy and phylogeny of the cloacae of salamanders (Amphibia: Caudata). I. Evolution at the family level. Herpetologica 165–193.

Sever, D.M. (1992). Comparative anatomy and phylogeny of the cloacae of salamanders (Amphibia: Caudata) IV. Salamandridae. Anat. Rec. *233*, 229–244.

Sever, D.M., and Staub, N.L. (2011). Hormones, sex accessory structures, and secondary sexual characteristics in amphibians. In Hormones and Reproduction in Vertebrates Volume 2: Amphibians, (Maryland Heights, Missouri: Elsevier Inc.),.

Sever, D.M., Verrell, P.A., Halliday, T.R., Griffiths, M., and Waights, V. (1990). The cloaca and cloacal glands of the male smooth newt, *Triturus vulgaris vulgaris* (Linnaeus), with especial emphasis on the dorsal gland. Herpetologica 160–168.

Smith, R.E. (1941). Mating behavior in *Triturus torosus* and related newts. Copeia *1941*, 255–262.

Stebbins, R.C. (1951). Amphibians of Western North America (Berkeley: University of California Press).

Swanson, W.J., and Vacquier, V.D. (2002). Reproductive protein evolution. Annu. Rev. Ecol. Syst. *33*, 161–179.

Symonds, M.R.E., and Elgar, M.A. (2008). The evolution of pheromone diversity. Trends Ecol. Evol. *23*, 220–228.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. (2013). MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. Mol. Biol. Evol. *30*, 2725–2729.

Tan, A.-M., and Wake, D.B. (1995). MtDNA phylogeography of the California newt, *Taricha torosa* (Caudata, Salamandridae). Mol. Phylogenet. Evol. *4*, 383–394.

Toews, D.P.L., and Brelsford, A. (2012). The biogeography of mitochondrial and nuclear discordance in animals. Mol. Ecol. *21*, 3907–3930.

Toyoda, F., Tanaka, S., Matsuda, K., and Kikuyama, S. (1994). Hormonal control of response to and secretion of sex attractants in Japanese newts. Physiol. Behav. *55*, 569–576.

Twitty, V.C. (1936). Correlated genetic and embryological experiments on *Triturus*. I and II. J. Exp. Zool. *74*, 239–302.

Twitty, V.C. (1942). The species of Californian *Triturus*. Copeia *1942*, 65–76.

Twitty, V.C. (1955). Field experiments on the biology and genetic relationships of the Californian species of *Triturus*. J. Exp. Zool. *129*, 129–147.

Twitty, V.C. (1964). Fertility of *Taricha* species-hybrids and viability of their offspring. Proc. Natl. Acad. Sci. U. S. A. *51*, 156–161.

Twitty, V.C. (1966). Of Scientists and Salamanders (San Francisco, CA: W. H. Freeman).

Watts, R.A., Palmer, C.A., Feldhoff, R.C., Feldhoff, P.W., Houck, L.D., Jones, A.G., Pfrender, M.E., Rollmann, S.M., and Arnold, S.J. (2004). Stabilizing selection on behavior and morphology masks positive selection on the signal in a salamander pheromone signaling complex. Mol. Biol. Evol. *21*, 1032–1041.

Weisrock, D.W., Harmon, L.J., and Larson, A. (2005). Resolving deep phylogenetic relationships in salamanders: analyses of mitochondrial and nuclear genomic data. Syst. Biol. *54*, 758–777.

Weisrock, D.W., Papenfuss, T.J., Macey, J.R., Litvinchuk, S.N., Polymeni, R., Ugurtas, I.H., Zhao, E., Jowkar, H., and Larson, A. (2006). A molecular assessment of phylogenetic relationships and lineage accumulation rates within the family Salamandridae (Amphibia, Caudata). Mol. Phylogenet. Evol. *41*, 368–383.

Wilburn, D.B., Bowen, K.E., Gregg, R.G., Cai, J., Feldhoff, P.W., Houck, L.D., and Feldhoff, R.C. (2012). Proteomic and UTR analysis of a rapidly evolving hypervariable family of vertebrate pheromones. Evolution 2227–2239.

Wilburn, D.B., Eddy, S.L., Chouinard, A.J., Arnold, S.J., Feldhoff, R.C., and Houck, L.D. (2015). Pheromone isoform composition differentially affects female behaviour in the red-legged salamander, *Plethodon shermani*. Anim. Behav. *100*, 1–7.

Woodley, S.K. (2010). Pheromonal communication in amphibians. J. Comp. Physiol. A *196*, 713–727.

Woodley, S.K. (2014). Chemical signaling in amphibians. In Neurobiology of Chemical Communication, C. Mucignat-Caretta, ed. (Boca Raton, FL: CRC Press), pp. 255–283.

Yamamoto, K., Kawai, Y., Hayashi, T., Ohe, Y., Hayashi, H., Toyoda, F., Kawahara, G., Iwata, T., and Kikuyama, S. (2000). Silefrin, a sodefrin-like pheromone in the abdominal gland of the sword-tailed newt, *Cynops ensicauda*. FEBS Lett. *472*, 267–270.

Yang, Z. (2007). PAML 4: Phylogenetic Analysis by Maximum Likelihood. Mol. Biol. Evol. *24*, 1586–1591.

Yang, Z., and Bielawski, J.P. (2000). Statistical methods for detecting molecular adaptation. Trends Ecol. Evol. *15*, 496–503.

Yang, Z., Nielsen, R., Goldman, N., and Pedersen, A.-M.K. (2000). Codon-substitution models for heterogeneous selection pressure at amino acid sites. Genetics *155*, 431–449.

Zhang, J. (2003). Evolution by gene duplication: an update. Trends Ecol. Evol. *18*, 292–298.

Zhang, P., Papenfuss, T.J., Wake, M.H., Qu, L., and Wake, D.B. (2008). Phylogeny and biogeography of the family Salamandridae (Amphibia: Caudata) inferred from complete mitochondrial genomes. Mol. Phylogenet. Evol. *49*, 586–597.

APPENDIX A

Codon-based tests of selection

The most recently developed tests for selection on codons rely on measuring the ratio (ω) of nonsynonymous mutations per nonsynonymous sites d_n to synonymous mutations per synonymous sites d_s or $\omega = \frac{d_N}{d_S}$ (Yang and Bielawski, 2000). Positive or diversifying selection can be described as an increase in the fitness, and therefore fixation rate, of some nonsynonymous mutations ($\omega > 1$) while negative or purifying selection is caused by a decrease in fitness after nonsynonymous mutations ($\omega < 1$). Historically, these calculations were averaged over the entire length of the protein, but functional constraints lead to most sites in a functioning protein to be under consistent purifying selection (Li, 1997). Thus, identifying proteins that were experiencing positive selection with these models was problematic. More recently, models that treat the codon as the level of selection have greatly improved sensitivity (Yang et al., 2000). Just like the proteins they are designed to study, the success of early codon-based models has led to a major radiation of new models, each with their own strengths and built in assumptions.

Tests for selection

The Single-Likelihood Ancestor Counting (SLAC) model calculates a site-by-site maximum likelihood ancestral sequence and then computes the observed and expected number of both synonymous and nonsynonymous mutations at each codon. These values are used to calculate the probability that selection has occurred at each site. Since this method is based partially on a globally calculated ω, it is conservative relative to the other methods (Pond and Frost, 2005a).

The Random-Effects Likelihood (REL) method is a generalization of the standard codon-based selection model implemented in PAML (Yang, 2007), in which the mutation rates for both d_S and d_N are allowed to vary. REL is the least conservative model used and therefore the level of significance of the computed Bayes Factors I used was more stringent than other models. A Bayes Factor of 50 (the level set for significant results) roughly corresponds to a p-value of 1/50 or 0.02.

The Fixed-Effects Likelihood (FEL) method is a modification of the REL method where d_S and d_N are directly estimated rather than being allowed to vary. It is less prone to false positives and computationally simpler than REL, and it captures the pattern of rate variation better than SLAC (Pond and Frost, 2005b).

The Mixed Effects Model of Evolution (MEME) is an extension of FEL that is able to detect episodic diversifying selection. MEME allows the estimated ω at each codon to vary at the branch level. This offers an improvement in sensitivity in detecting positive selection over FEL when codons are experiencing positive selection ($\omega > 1$) in one branch and purifying selection ($\omega < 1$) in the larger tree (Murrell et al., 2012).

The Fast Unconstrained Bayesian AppRoximation (FUBAR) test for selection removes the possibility of misclassifying sites by averaging selection parameters of a large number of site classes. This increases the power of detecting sites undergoing positive selection. The implementation of Bayesian approximation methods (based on

Markov chain Monte Carlo methods) also greatly improves the computational efficiency of the test, allowing for much larger data sets to be analyzed (Murrell et al., 2013).

TABLES

Table 1: Locality and specimen collection data and SPF sequence data. Each unique nucleotide sequence was assigned a three digit code. A subset of these sequences that encode unique isoforms is shown in the table, sorted by specimen and SPF clade. When multiple individuals express identical isoforms, the sequence code is repeated in the table. Species are listed by their 4-letter code (*T. granulosa*: TAGR; *T. rivularis*: TARI; *T. sierrae*: TASI; *T. torosa*: TATO)

			SPF sequence number				
Region	Locality	Specimen	Beta-1a	Beta-1b	Beta-1c	Beta1-d	Beta-2
Coastal (Site A)	Orrs Creek (Mendocino Co.) N 39.16355 W 123.23249	TAGR 01	078, 079	077			
		TAGR 02	079	080			
		TAGR 03	082, 083	081			
		TARI 01					068, 070
		TARI 02					073, 074, 075
	Mills Creek (Mendocino Co.) N 39.125499 W 123.127119	TATO 01					035, 036, 037, 038, 041
		TATO 02					044, 045, 046
Northern (Site B)	Interior (Trinity Co.) N. 40.64369 W 123.22180	TAGR 04	085				
	Sierra (Butte Co.) N 39.81284 W 121.58389	TASI 06			020, 021, 024, 025		
		TASI 07			027, 028, 030		
		TASI 08			031, 033		
		TASI 09			027		
Southern Sierra (Site C)	Salt Creek (Kern Co.) N 36.44736 W 118.86025	TASI 01[a]				003, 007	001, 002, 004, 008
		TATO 07				059	057, 060
		TATO 08				063	062
		TATO 09				065	064, 066
	N.F. Kaweah (Kern Co.) N 36.53914 W 118.89803	TASI 02				009	
		TASI 04				003, 013, 014, 018	012, 016, 017
	S.F. Kaweah (Kern Co.) N 36.35138 W 118.77861	TATO 03				047, 052	017, 048, 050, 051
		TATO 04					017
		TATO 05				055	054

[a] This individual's traits were a blend of TASI and TATO. More characters were consistent with TASI, however, and it was scored as such for analysis.

Table 2: SPF length variants isolated from *Taricha*. Length variants 1-5 were found in multiple specimens and were used in phylogenetic and selection analyses. Variants 6-16 were only found within a single clone and were excluded from most analyses.

	ORF length	#Sequences/ #Individuals	Species	Putative gene	Comments
Variant 01	600	26/12	TAGR,TASI,TATO	Beta-1	
Variant 02	603	18/12	TAGR,TASI,TATO	Beta-1	
Variant 03	606	22/4	TARI,TATO	Beta-2	
Variant 04	333	22/8	TASI,TATO	Beta-2	Splice of Variant 03
Variant 05	450	4/3	TATO	Beta-2	Splice of Variant 03
Variant 06	193	1/1	TASI	Beta-1	Large internal splice
Variant 07	204	1/1	TATO	Beta-2	Nonsense mutation
Variant 08	231	1/1	TASI	Beta-1	Nonsense mutation
Variant 09	246	1/1	TATO	Beta-2	Internal splice (to 333) and nonsense mutation
Variant 10	255	1/1	TATO	Beta-2	Nonsense mutation
Variant 11	261	1/1	TASI	Beta-1	Nonsense mutation
Variant 12	489	1/1	TATO	Beta-2	Internal splice
Variant 13	492	1/1	TATO	Beta-2	Internal splice (to 450) and mutation in stop codon
Variant 14	588	1/1	TASI	Beta-1	Nonsense mutation
Variant 15	615	1/1	TASI	Beta-1	Mutation in stop codon
Variant 16	696	1/1	TASI	Beta-1	Large Insert

Table 3: SPF Beta-1 genetic dissimilarity between clades. Percentage pairwise nucleotide similarity compared in lower left and amino acid similarity compared in upper right. Description of sequences found in each clade can be found in Figure 4.

	TAGR (β-1a)	TAGR (β-1b)	TASI (β-1c)	TATO (β-1d)	TASI (β-1d)
TAGR (β-1a)		24.7	15.5	21.9	21.8
TAGR (β-1b)	13.0		15.4	18.7	18.6
TASI (β-1c)	7.9	9.8		19.6	19.5
TATO (β-1d)	9.0	9.3	8.1		0.7
TASI (β-1d)	9.1	9.4	8.2	0.3	

Table 4: Length variant dissimilarity table. Lower left shows nucleotide percent dissimilarity, upper right shows amino acid percent dissimilarity. Diagonal values show within group amino acid dissimilarity.

	Beta-1		Beta-2		
	Variant 1	Variant 2	Variant 3	Variant 4	Variant 5
Variant 1	11.5	14.9	43.1	42.5	42.4
Variant 2	7.1	12	44.4	43.8	43.7
Variant 3	27.3	26.9	1.9	1.6	1.4
Variant 4	27.2	26.8	0.7	1.0	0.8
Variant 5	27.2	26.7	0.7	0.5	0.6

Table 5: Number of codons with significant selection. See Appendix A for a description of the tests used to test of evidence of selection. The “Total Significant” column is higher than the number of codons because the same codon could be significant for selection with multiple tests.

		SLAC	FEL	REL	MEME[*]	FUBAR	Total significant	# Codons
Full Set	-	5	11	0	-	3	19	13
	+	0	6	0	10	4	20	13
Beta-1	-	2	7	1	-	4	14	9
	+	0	2	15	4	2	23	15
Beta-2	-	0	5	0	-	0	5	5
	+	0	0	0	0	1	1	1

[*] MEME is not used to detect negative selection.

Table 6: Results from codon-based tests for selection on SPF. Compiled results of five separate tests (1: SLAC, 2: FEL, 3: REL, 4: MEME, 5: FUBAR) for selection (see Methods and Appendix A). Only codons with at least two significant results are shown. Sites with significant levels of selection are marked with "+" and "-" to indicate positive and negative selection, respectively. All amino acids found at a particular position are given along with number of sequences with that amino acid.

	Full Set					Beta-1					Beta-2						
Codon#	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	Amino Acid	Change from consensus
23	-	-			-	-	-	-		-						C_{53}	
70		+		+												L_{48},P_{2}	
79	-	-										-				G_{50}	
82					+			+								V_{21},F_{19},E_{10}	
88	-	-			-											V_{50}	
98		-					-									C_{50}	
116				+				+	+							T_{61},A_{2}	Smaller
128		+		+	+			+	+							D_{35},T_{28}	Charge change
132				+				+	+							T_{39},D_{24}	(-) charge
133		+		+	+			+	+							G_{38},P_{21},D_{3},S_{1}	(-) charge
139		-					-									E_{63}	
141																Q_{55},H_{7},R_{1}	(+) charge
149		-			-					-						C_{62},S_{1}	
150		+		+	+		+	+		+						G_{43},A_{13},T_{7}	Larger
157				+											+	S_{57},F_{5},P_{1}	(+) charge
161				+					+							E_{61},G_{2}	Neut. charge
162							-			-						A_{63}	
166		-				-	-			-						Y_{63}	

FIGURES

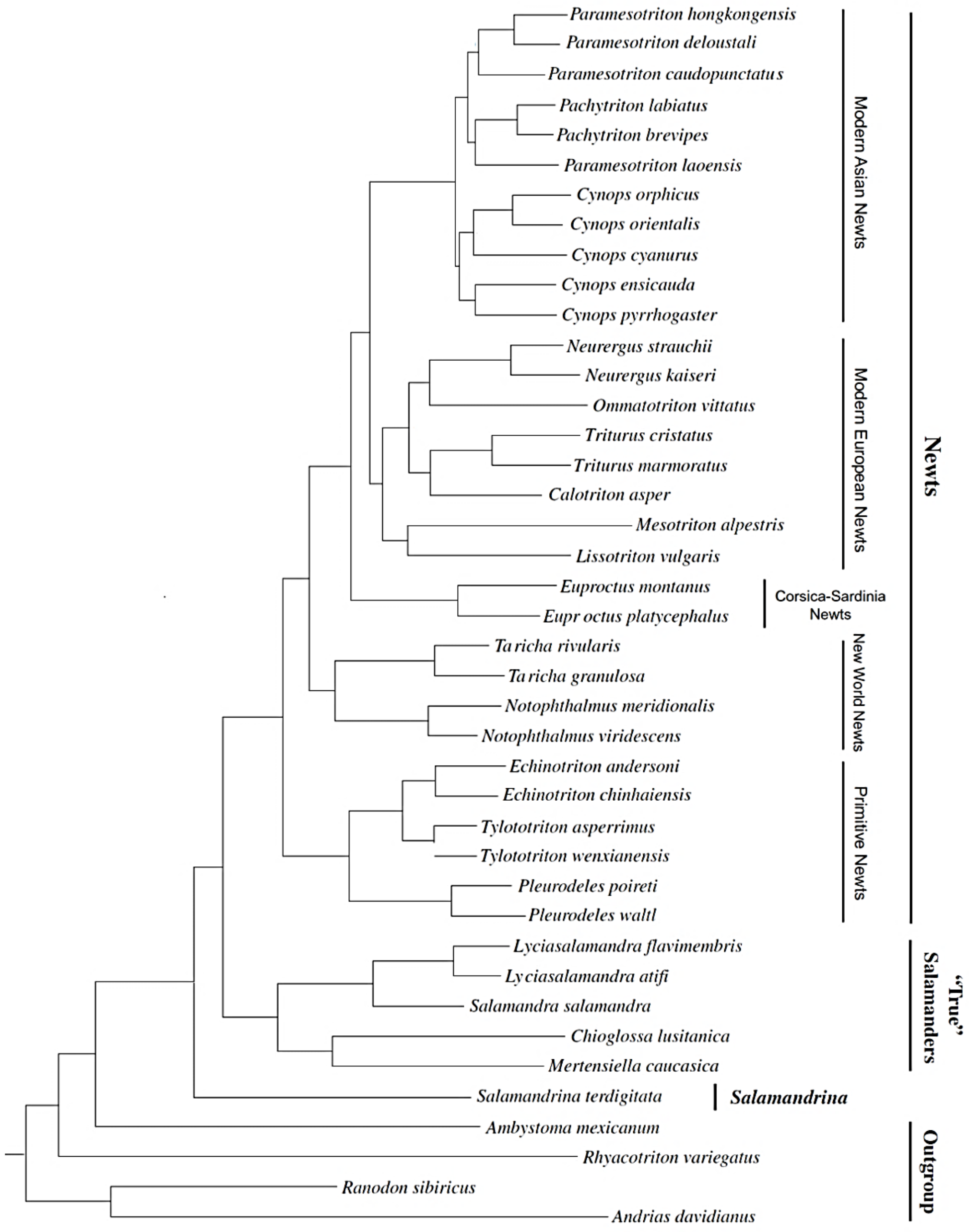


Figure 1: Salamandridae family phylogeny based on mitochondrial DNA (Adapted from Zhang et al., 2008).

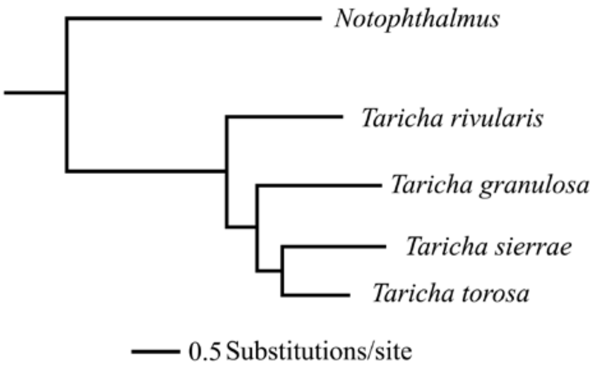


Figure 2: Phylogeny of North American salamandrids generated with mtDNA. (Adapted from Weisrock et al., 2006).

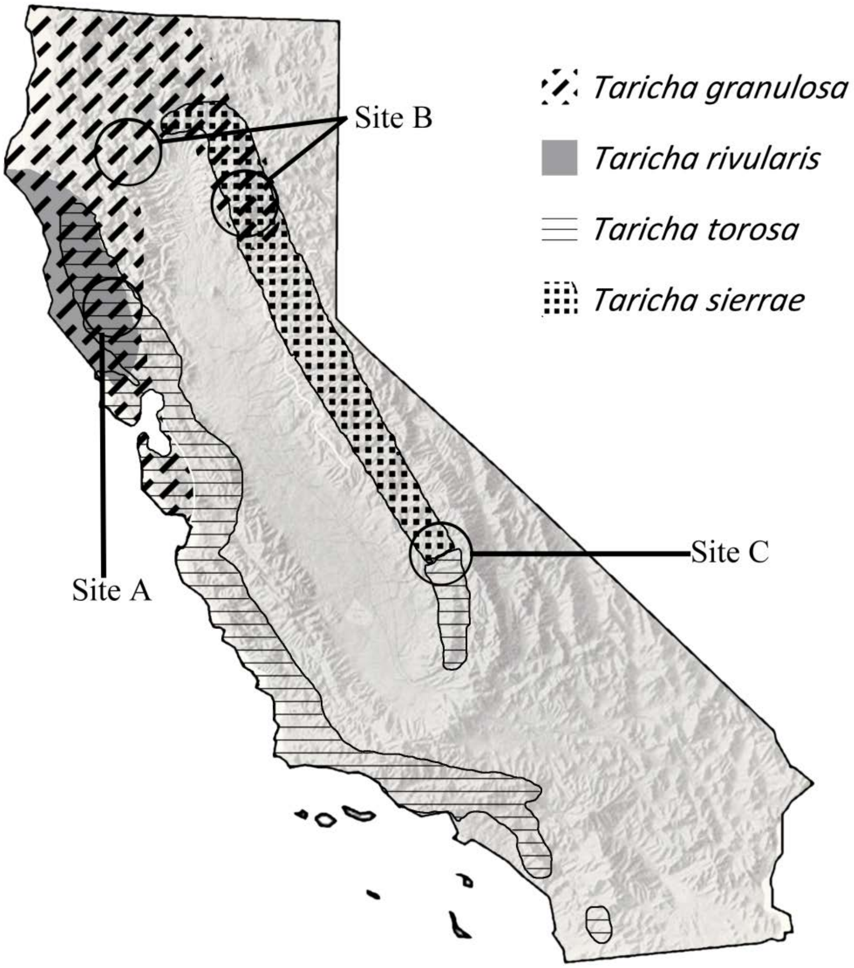


Figure 3: Distribution of four *Taricha* species in California and collection sites for this study.

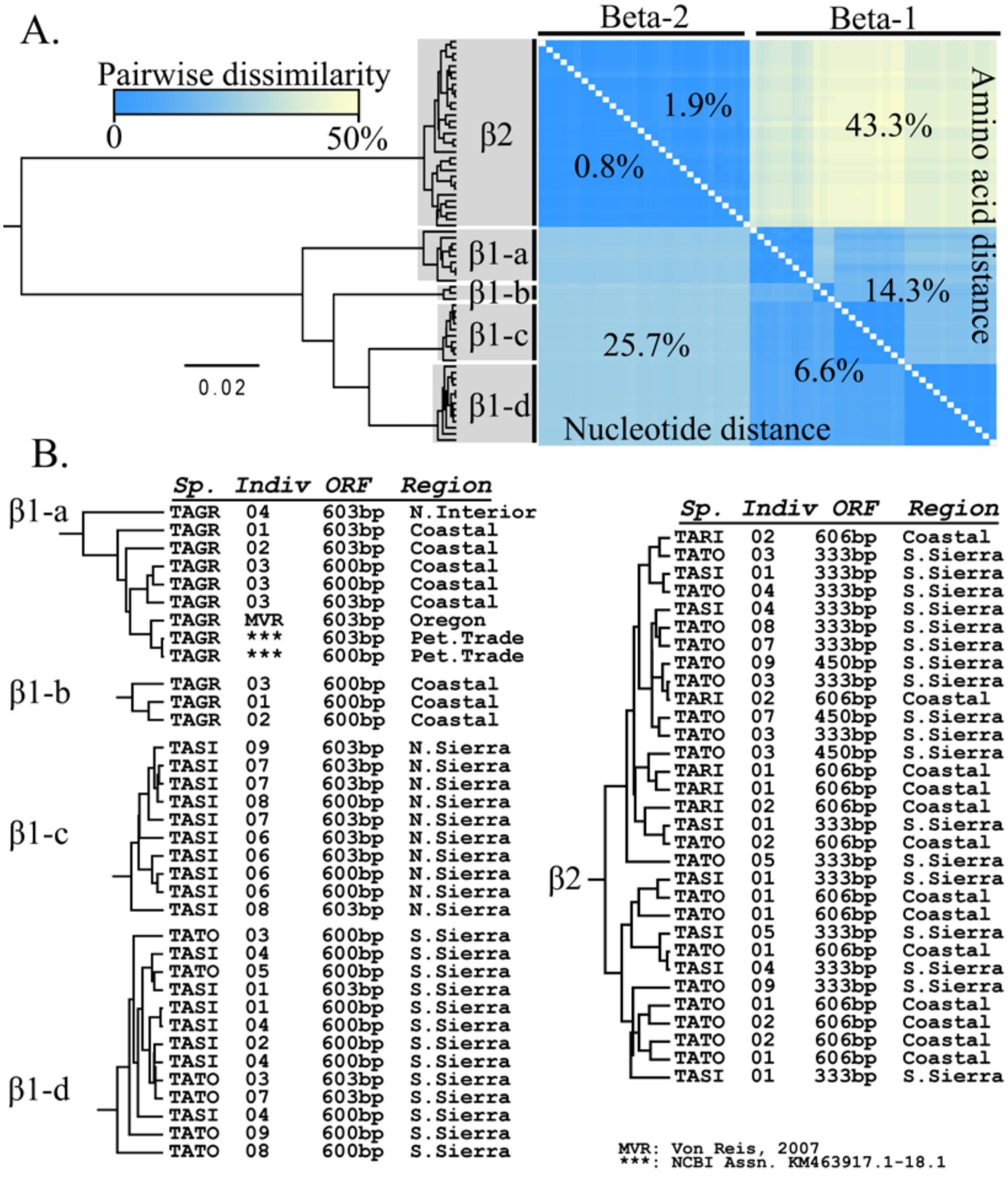
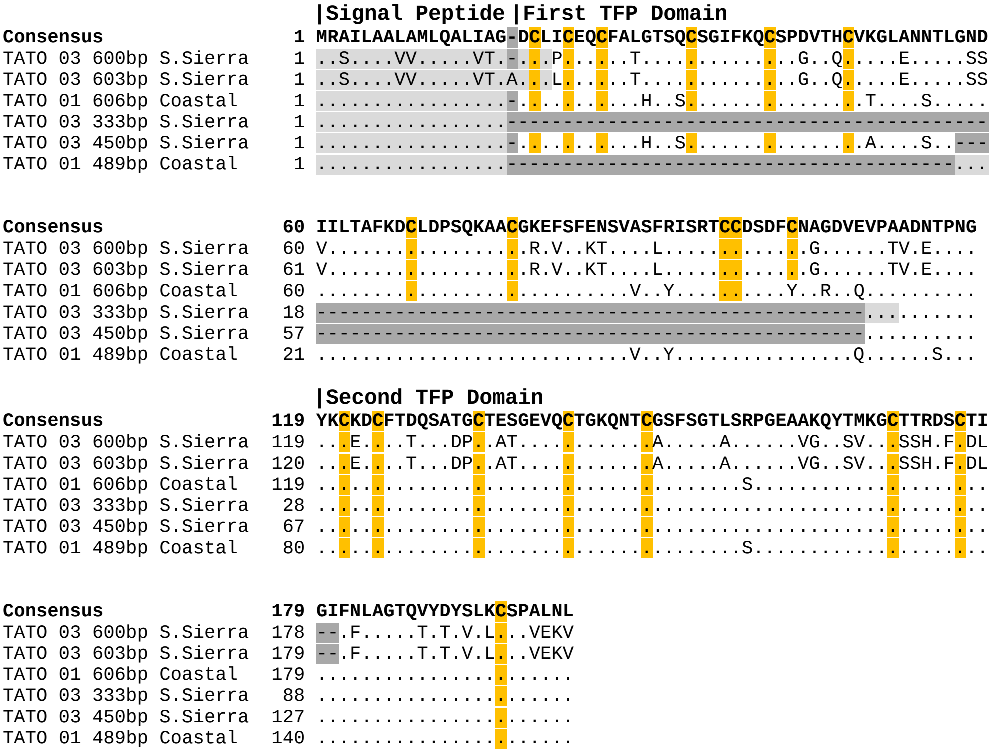


Figure 4: **A:** Pairwise dissimilarity matrix. Lower left shows nucleotide percent dissimilarity, upper right shows amino acid percent dissimilarity. To the left of matrix a condensed Bayesian inference gene phylogeny shows the five distinct SPF clades found in *Taricha*. Root extends to *Taricha* SPF Alpha sequences (not shown). **B:** Beta SPF clade composition. β1-a) *T. granulosa* sequences from both coastal and northern populations, β1-b) *T. granulosa* only from coastal populations, β1-c) *T. sierrae* from northern Sierra Nevada, β1-d) *T. sierrae* and *T. torosa* from the S. Sierra Nevada hybrid zone. The five clades are strongly supported (PP = 1.00) but internal structure has very limited support.



```
                            |Signal Peptide |First TFP Domain
Consensus                 1 MRAILAALAMLQALIAG-DCLICEQCFALGTSQCSGIFKQCSPDVTHCVKGLANNTLGND
TATO 03 600bp S.Sierra    1 ..S....VV.....VT.-...P......T..............G..Q.....E.....SS
TATO 03 603bp S.Sierra    1 ..S....VV.....VT.A...L......T..............G..Q.....E.....SS
TATO 01 606bp Coastal     1 .................-...........H..S................T....S.....
TATO 03 333bp S.Sierra    1 .................-------------------------------------------
TATO 03 450bp S.Sierra    1 .................-...........H..S................A....S..---
TATO 01 489bp Coastal     1 .................----------------------------------------...

Consensus                60 IILTAFKDCLDPSQKAACGKEFSFENSVASFRISRTCCDSDFCNAGDVEVPAADNTPNG
TATO 03 600bp S.Sierra   60 V..................R.V..KT....L.............G......TV.E....
TATO 03 603bp S.Sierra   61 V..................R.V..KT....L.............G......TV.E....
TATO 01 606bp Coastal    60 ............................V..Y..........Y..R..Q..........
TATO 03 333bp S.Sierra   18 -------------------------------------------------..........
TATO 03 450bp S.Sierra   57 -------------------------------------------------..........
TATO 01 489bp Coastal    21 ............................V..Y................Q......S...

                            |Second TFP Domain
Consensus               119 YKCKDCFTDQSATGCTESGEVQCTGKQNTCGSFSGTLSRPGEAAKQYTMKGCTTRDSCTI
TATO 03 600bp S.Sierra  119 ...E....T...DP..AT............A.....A......VG..SV...SSH.F.DL
TATO 03 603bp S.Sierra  120 ...E....T...DP..AT............A.....A......VG..SV...SSH.F.DL
TATO 01 606bp Coastal   119 ......................................S.....................
TATO 03 333bp S.Sierra   28 ............................................................
TATO 03 450bp S.Sierra   67 ............................................................
TATO 01 489bp Coastal    80 ......................................S.....................

Consensus               179 GIFNLAGTQVYDYSLKCSPALNL
TATO 03 600bp S.Sierra  178 --.F.....T.T.V.L...VEKV
TATO 03 603bp S.Sierra  179 --.F.....T.T.V.L...VEKV
TATO 01 606bp Coastal   179 .......................
TATO 03 333bp S.Sierra   88 .......................
TATO 03 450bp S.Sierra  127 .......................
TATO 01 489bp Coastal   140 .......................
```

Figure 5: Amino acid alignment of representative length Variants of SPF found in *Taricha*. All sequences are from two *T. torosa* individuals (#10 and #12). Lengths 606 and 489 are from the coastal population and Lengths 450, 333, 600, 603 are from the Kaweah River. Variant 12 (TATO 10.10) included to demonstrate an alternative splicing pattern at codon 17. Putative secretion signal peptide is marked in light grey. Alignment gaps are marked in dark grey. Cysteine backbone is marked in orange.

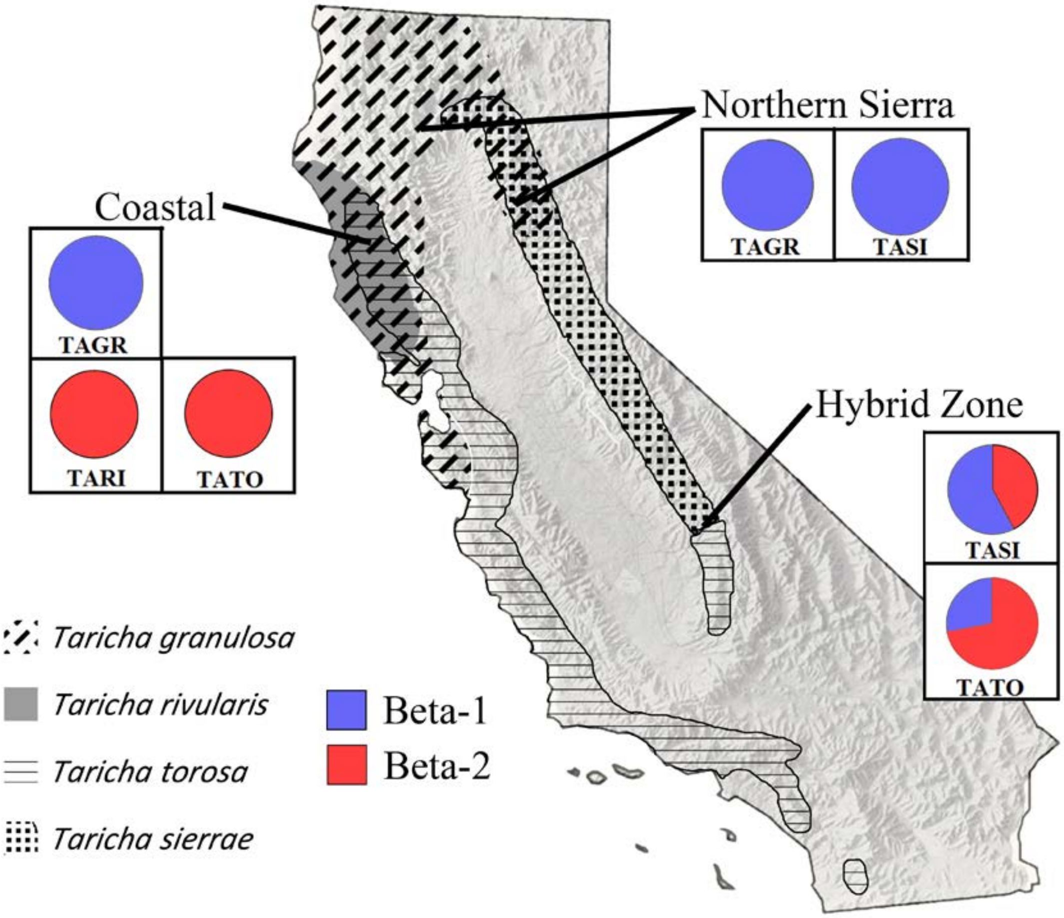


Figure 6: Geographic and species distribution of expression of Beta-1 and Beta-2 SPFs. Pie charts show the proportions of Beta-1 and Beta-2 sequences expressed by each species in each region. The boxes are empty where the species is not present in a region. The overall includes sequences from all regions. Expression of Beta Variants of SPF varied significantly by species (Fisher's Exact Test, $p < 0.001$).

A.

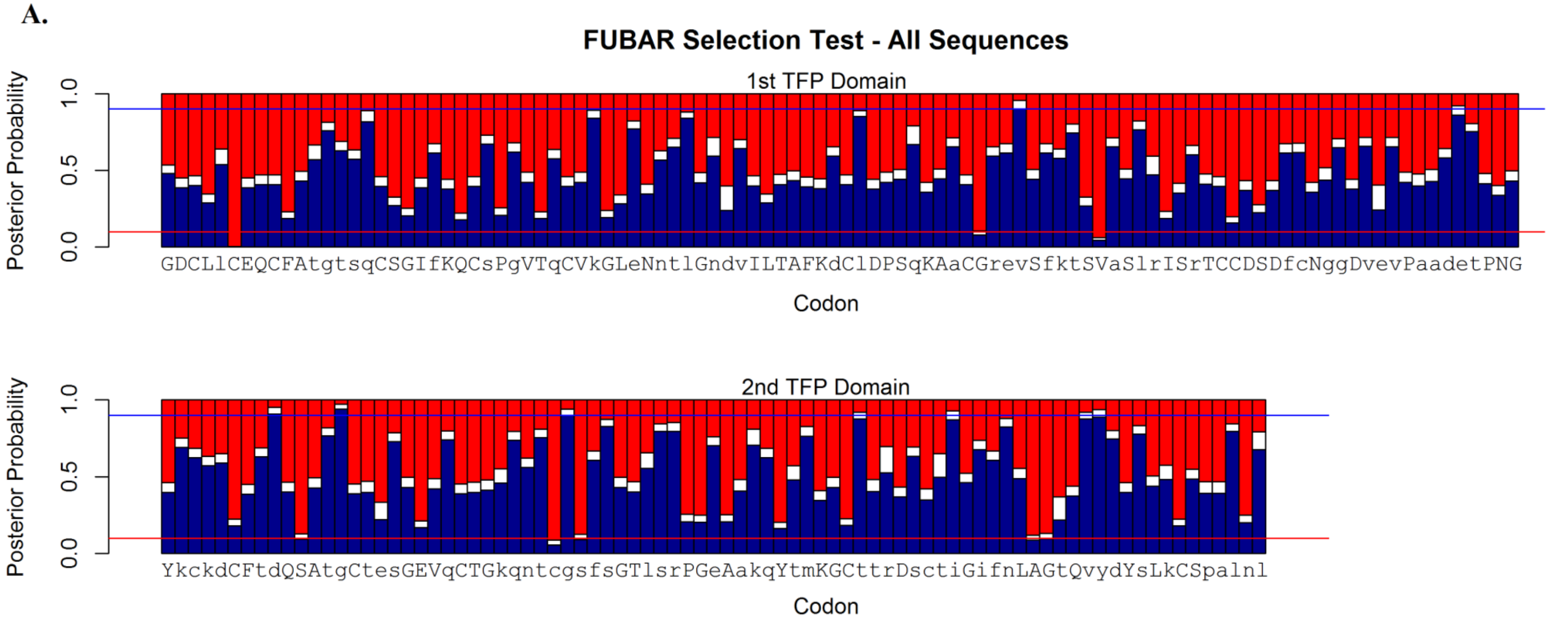
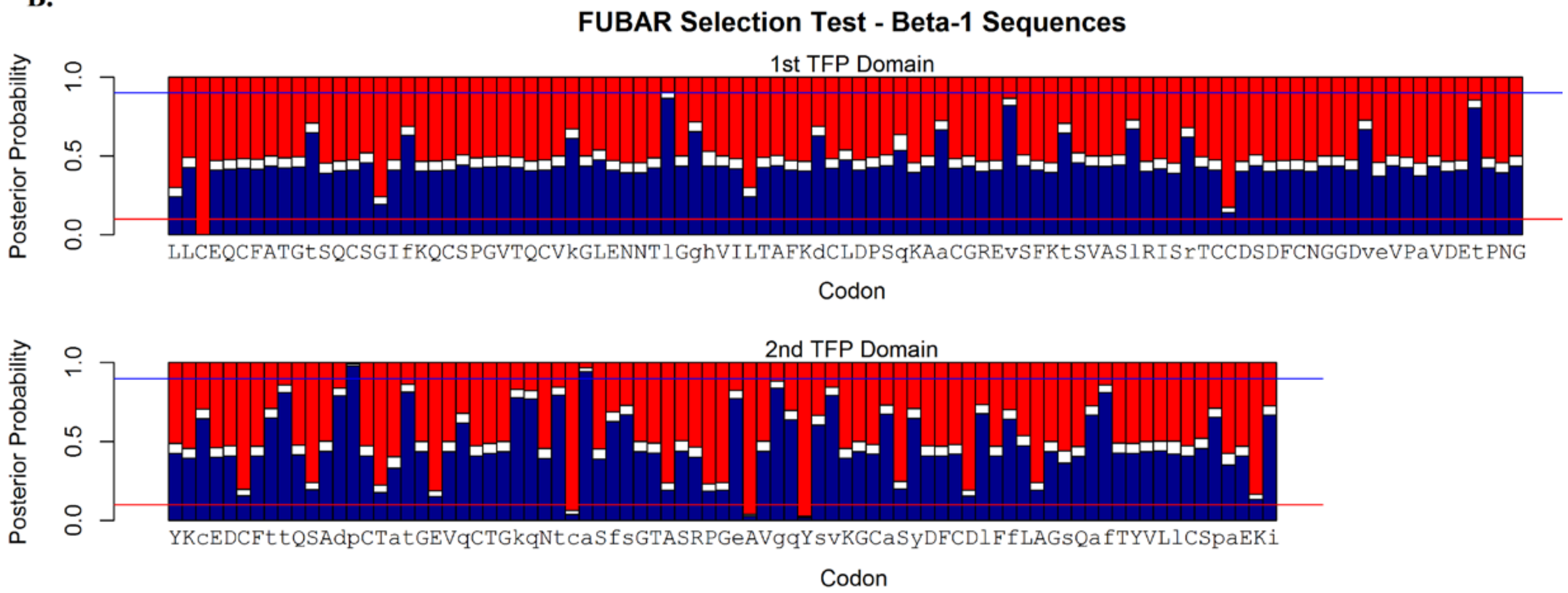
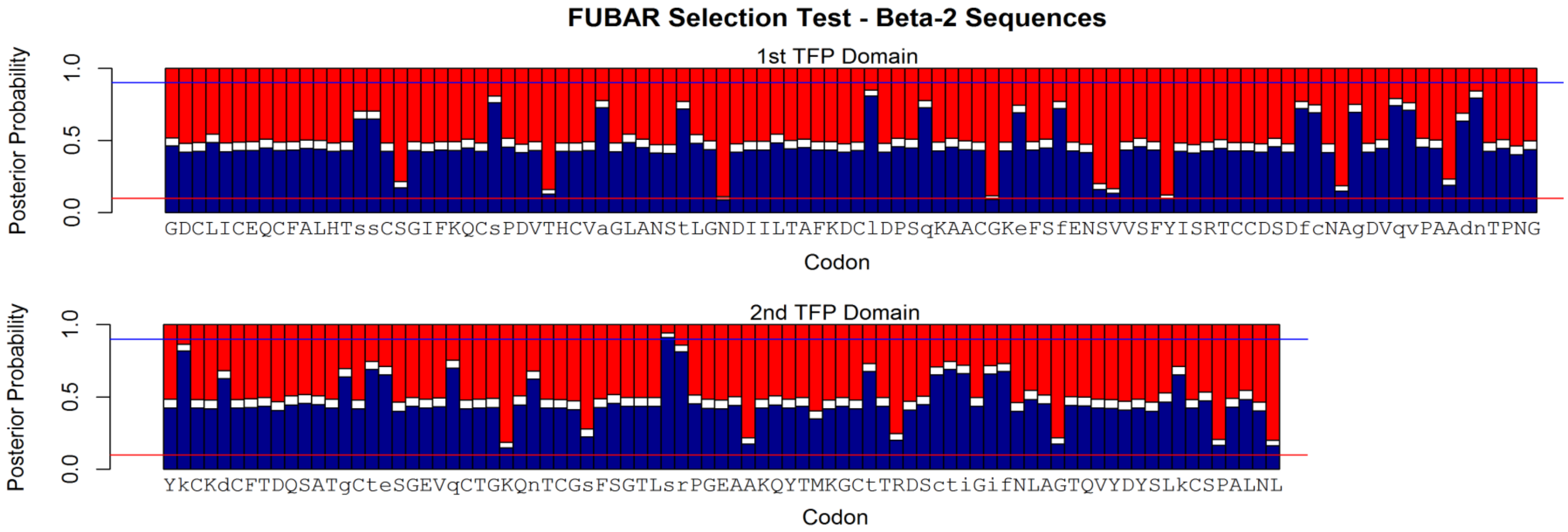


Figure 7: FUBAR selection analysis. Boxes represent calculated posterior probability of selection at each codon; red indicates negative selection, blue indicates positive selection. Horizontal lines indicate level of significance (0.1 for negative selection, 0.9 for positive selection). The consensus sequence is shown below the chart; residues indicated with capital letters were conserved in all sequences in the data set. A: All sequences; B: Beta-1 Sequences only; C: Beta-2 sequences only.

B.


FUBAR Selection Test - Beta-1 Sequences
1st TFP Domain
Posterior Probability
0.0
0.5
1.0
LLCEQCFATGtSQCSGIfKQCSPGVTQCVkGLENNTlGghVILTAFKdCLDPSqKAaCGREvSFKtSVASlRISrTCCDSDFCNGGDveVPaVDEtPNG
Codon
2nd TFP Domain
Posterior Probability
0.0
0.5
1.0
YKcEDCFttQSAdpCTatGEVqCTGkqNtcaSfsGTASRPGeAVgqYsvKGCaSyDFCDlFfLAGsQafTYVLlCSpaEKi
Codon

C.


FUBAR Selection Test - Beta-2 Sequences
1st TFP Domain
Posterior Probability
1.0
0.5
0.0
GDCLICEQCFALHTssCSGIFKQCsPDVTHCVaGLANStLGNDIILTAFKDClDPSqKAACGKeFSfENSVVSFYISRTCCDSDfcNAgDVqvPAAdnTPNG
Codon
2nd TFP Domain
Posterior Probability
1.0
0.5
0.0
YkCKdCFTDQSATgCteSGEVqCTGKQnTCGsFSGTLsrPGEAAKQYTMKGCtTRDSctiGifNLAGTQVYDYSLkCSPALNL
Codon

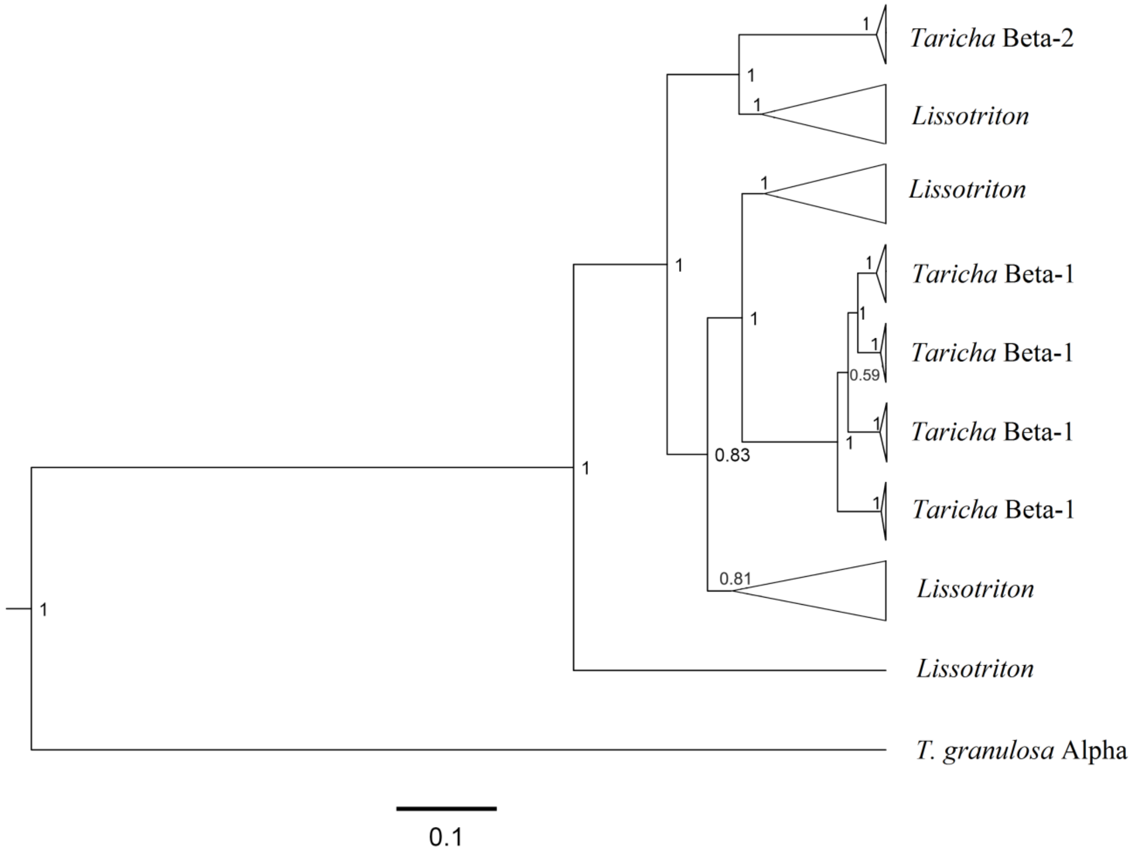


Figure 8: SPF gene tree using available *Lissotriton* and *Taricha* SPF sequences from NCBI Genbank along with *Taricha* sequences generated from this project. Numbers at nodes represent Bayesian posterior probabilities rounded to two significant figures. Triangles represent a condensed node of multiple sequences and the width of the triangle represents the diversity within the condensed clade. Scale bar represents nucleotide substitutions per site.

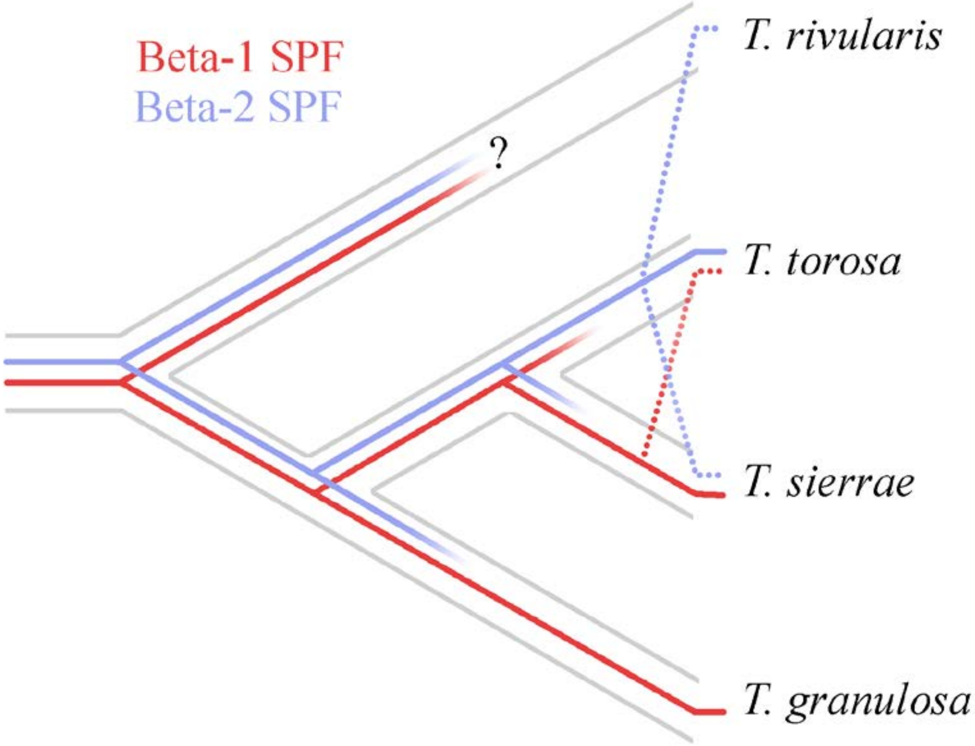


Figure 9: Hypothesized evolutionary history of two major SPF Beta clades in *Taricha*. Introgression of alleles shown with dotted lines. The only sequences recovered from *T. rivularis* are apparently the result of recent introgression, and therefore it is unclear what *T. rivularis* expresses aside from introgressed loci. The common ancestor to *Taricha* presumably expressed both Beta genes and loss of expression in modern lineages is shown with faded lines not reaching the tips.

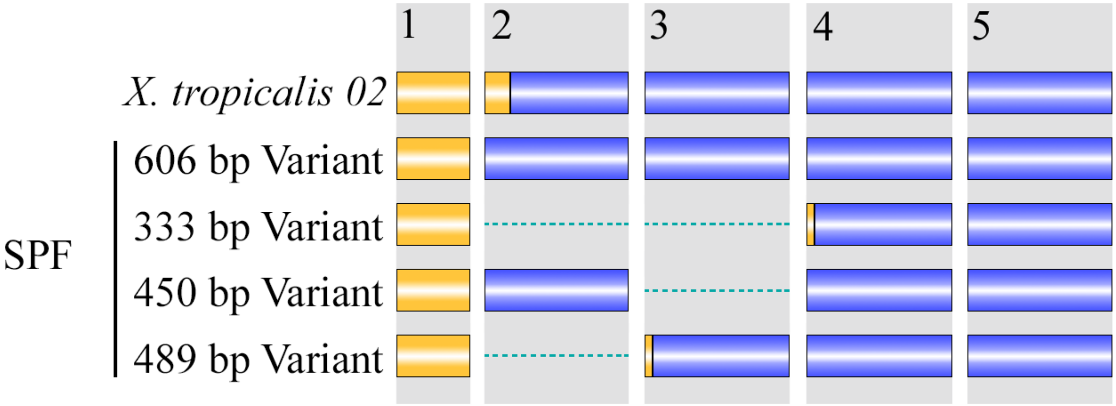


Figure 10: Predicted exon skipping in *Taricha*. Analysis of the *Xenopus tropicalis* genome (Janssenswillen et al., 2015) and the use of genome walking techniques in *Desmognathus ocoee* (Leichty, 2012) indicate a conserved pattern of five exons (indicated with numbered grey boxes above) and four introns across the SPF open reading frame. Green represents the predicted signaling peptide. Blue represents mature SPF. Adapted from supplemental materials of Janssenswillen et. al. 2014.