Gene Duplication, Co-option, Structural Evolution, and Phenotypic Tango in the Courtship Pheromones of Plethodontid Salamanders

DAMIEN B. WILBURN1,4, STEVAN J. ARNOLD2, LYNEE D. HOUCK2, PAMELA W. FELDHOFF3, AND RICHARD C. FELDHOFF3

1 Department of Genome Sciences, University of Washington, Seattle, WA 98195, USA
2 Department of Integrative Biology, Oregon State University, Corvallis, OR 97331, USA
3 Department of Biochemistry, University of Louisville, Louisville, KY 40292, USA

ABSTRACT: Rapid evolution is a hallmark of proteins involved in reproduction. The protein courtship pheromones in plethodontid salamanders are classic examples of such rapidly evolving reproductive proteins, with male pheromones likely coevolving with female receptors to improve reproductive success. Over the past ~66 million years of plethodontid evolution, the structure and composition of the male mental gland has evolved with changes in courtship timing and behavior. More than 20 yr of biochemical and molecular studies have provided insight into how multiple gene families have been duplicated, mutated, and co-opted for pheromone roles. Sequencing and mass spectral proteomic analyses have enabled identification and characterization of multiple pheromone families, some with lineage-specific expression. In this review, we provide a phenotypic tango model to better understand male pheromone and female receptor coevolution that has driven the rapid evolution of multiple diverse pheromone families. To offer support for this phenotypic tango model, we review a combination of behavioral, neurophysiological, and structural studies that inform our understanding of the underlying molecular mechanisms of pheromone signaling.

Key words: Biochemistry; Caudata; Protein; Sexual selection

From microbes to humans, rapid molecular evolution is a common feature of traits involved in sexual reproduction (Swanson and Vacquier 2002; Wilburn and Swanson 2016). Arising from anisogamy and differences in reproduction strategies, strong sexual selection drives this accelerated evolution, producing many of the complex, ornate sexual phenotypes that have fascinated naturalists for centuries (Mead and Arnold 2004). Over the past ~30 yr, advances in DNA sequencing and other molecular techniques have revealed that the genes involved in reproduction routinely experience rapid evolution, particularly those coding for proteins that are exchanged between sexes. Such examples include egg–sperm recognition proteins, seminal fluid proteins, and pheromones (Wilburn and Swanson 2016). Although pheromones are sometimes thought to be only volatile odorants that attract mates, they are instead a diverse class of semiochemicals that facilitate a wide range of social functions. Virtually any class of organic molecule can be a pheromone, from simple hydrocarbon chains and aromatic rings to larger biomolecules such as lipids and peptides and proteins (Wyatt 2014). In the study of pheromone evolution, peptide and protein pheromones are of particular interest: as direct gene products, it is easier to probe their evolutionary histories, compared to volatile pheromones that are often the products of long enzymatic cascades. In vertebrates, such protein pheromones have been well characterized only in rodents and urodèles (Kikuyama et al. 1995; Yamamoto et al. 2000; Houck and Arnold 2003; Mudge et al. 2008; Roberts et al. 2010, 2012; Woodley 2010; Van Bocxlaer et al. 2016). Notably, the first vertebrate peptide pheromone was identified in newts (Kikuyama et al. 1995), stimulating the subsequent discovery of many peptide and protein pheromones across Caudata (Kikuyama and Toyoda 1999; Yamamoto et al. 2000; Nakada et al. 2007; Maex et al. 2016; Van Bocxlaer et al. 2016).

For >25 yr, the protein courtship pheromones of plethodontid salamanders have been one of the most extensively characterized pheromone systems. Studies of this system have integrated aspects of genetics, biochemistry, neurophysiology, endocrinology, evolutionary theory, and behavioral ecology (Houck and Reagan 1990; Feldhoff et al. 1999; Rollmann et al. 1999; Watts et al. 2004; Houck et al. 2007a; Palmer et al. 2007a; Woodley 2010, 2015). Consistent with genes coding for other types of reproductive proteins, these pheromone genes are often under positive Darwinian selection (Wilburn and Swanson 2016). Here, we describe a phenotypic tango framework for understanding pheromone–receptor coevolution and review the current state of plethodontid pheromone biology, highlighting recent proteomic and biochemical studies that suggest a repeated pattern of gene duplication, co-option, and structural adaptation consistent with theoretical models of sexual selection.

Sexual Selection and Phenotypic and Molecular Tango

One of the challenges in understanding the rapid evolution of plethodontid courtship pheromones and other reproductive proteins is that we have no satisfactory model to help us organize the molecular data. We know that the protein sequences have evolved rapidly and apparently in perpetuity, but we do not know what model best accounts for that evolutionary pattern. Selection is involved, but what kind of selection? In parallel with our work comparing the sequences and structure of plethodontid pheromones, we have sought an overarching conceptual framework that would integrate and explain the molecular and biochemical results. Early in our comparative work, we proposed that the male pheromone evolves in a molecular tango with female receptors (Palmer et al. 2005, 2007b, 2010). We imagined that a change in the sequence of a female receptor might exert sexual selection on males and elicit a matching change in the pheromone protein of the male. We argued that

4 CORRESPONDENCE: e-mail, dwilburn@u.washington.edu
Interactions. A plethodontid pheromone is a complex cocktail of proteins and peptides, so we must imagine collapsing the phenotypic effects of that cocktail onto a single hypothetical axis. Little is known about the corresponding female receptors (described in further detail below), but receptor complexity might rival that of the pheromone. For modeling purposes, we collapse that complexity as well onto the same axis. Natural selection on the protein cocktail refers to the functional properties of the pheromone and receptors. A pheromone, for example, must fold in a particular way to bind with any particular receptor. This folding—driven by the protein sequence and thermodynamics—is shaped by stabilizing natural selection. Likewise, a receptor must be properly positioned in its cellular environment, and those positioning characteristics are affected by stabilizing natural selection. The pheromone is also shaped by sexual selection exerted by the aspects of the receptor population that relate to pheromone-receptor binding. With these assumptions in place, we can model the phenotypic tango of a hypothetical pheromone and its receptors. The advantage of focusing on the aggregate phenotypic properties of the pheromone cocktail and the female receptor population is that we can use a fully characterized model of phenotypic coevolution (Arnold and Houck 2016) to visualize the coevolution of the cocktail and the receptor population by running simulations.

In contrast, we do not have a model for the molecular tango of individual pheromone molecules and the individual receptor sequences that exert sexual selection.

In the absence of sexual selection (Fig. 1, top panel), the pheromone and receptor means tend to be close to their optima. In successive generations, each mean tends to drift away from its optimum, while natural selection pulls it back toward the optimum. As we gradually increase the force of sexual selection (Fig. 1, bottom three panels), the pheromone mean equilibrates at a new position, away from its natural selection optimum and toward the receptor optimum. Remarkably, the receptor means continue to hover about their natural selection optimum, even when sexual selection is strong. Very strong sexual selection on the pheromone overpowers natural selection (Fig. 1, bottom panel), causing the pheromone means to hover close to the receptor optimum. Thus, the evolutionary drift of the pheromone is largely controlled by natural selection on the receptors; other general lessons about the phenotypic tango are discussed in Arnold and Houck (2016).

The phenotypic tango we have just described provides some insights into the evolution of a male pheromone and its coevolution with female receptors, but we do not yet have a model of the underlying molecular tango. Although that molecular model is still being developed, the phenotypic model gives us insight about the molecular process. For example, as we review the evolutionary pattern of the three main constituents of the plethodontid pheromone, we have two recurrent themes: the evolution of the pheromone seems to be incessant and characterized by repeated episodes of positive selection. In the phenotypic tango model, incessant evolution is a consequence of drift-selection balance. In a population of infinite size, a stable point equilibrium is struck between the opposing forces of natural and sexual selection. Even in a population with an effective size in the thousands, the pheromone mean will constantly drift away from this stable point, only to be pulled back toward it by natural and sexual selection. During those episodes of response to
FIG. 1.—Portrayal of the phenotypic tango model of pheromone-receptor coevolution (described in Arnold and Houck 2016) that provides a framework for understanding the coevolution of male ornaments and female preferences. Average values of male pheromone and female receptor for that pheromone in a single population are shown in each panel for 100 generations after the population has achieved drift-selection balance. The positions of the natural selection optima for pheromones and receptors are shown with vertical blue and red dashed lines, respectively. The horizontal solid blue and red lines with bars show the 95% confidence limits for the ornament and preference means, respectively. Thin gray lines connect the pheromone and receptor means (blue and red dots) in particular generations. The positions of pheromone and receptor means are offset in the vertical dimension so that they can be visualized. In the top panel, only natural selection acts on pheromones and receptors. The bottom three panels show outcomes when progressively stronger sexual selection is added to the model (denoted as +, ++, and +++). A color version of this figure is available in the online version of this article.
selection toward a stationary (or moving) equilibrium point, key residues on the protein responsible for interaction with receptors should experience positive selection. These episodes of directional response could produce the histories of positive selection that are often detected in molecular analyses of plethodontid courtship pheromones (Palmer et al. 2007a). Before we review the molecular and biochemical results that inspired our phenotypic tango model and molecular tango idea, we need to describe the male glands that produce plethodontid courtship pheromones and the chemosensory tissues that mediate behavioral reactions to those pheromones.

Plethodontid Courtship and the Role of the Mental Gland

Plethodontid salamanders are highly reliant on their sense of smell and ability to perceive the chemical environment. A defining morphological feature of a plethodontid salamander is the nasolabial groove: a vertical slit on either side of the snout that transports odorants from the substrate to the naris and into the nasal cavity by capillary action (Lanza et al. 1998). Plethodontid salamanders respond to odorants in the environment, as well as to a variety of volatile and nonvolatile chemical signals that provide information on species, size, sex, female gravidity, diet, and parasite load (Dawley 1984, 1986; Walls et al. 1989; Marco et al. 1998; Maksimovich and Mathis 2001; Danzter and Jaeger 2007; Chouinard 2012). Some of these signals are released from a large repertoire of skin glands located on the chin, cheeks, tail, and ventral surfaces (Houck and Sever 1994).

The best characterized of these skin glands is the mental gland: a secondary sexual trait that hypertrophies seasonally in adult male plethodontids, coincident with elevated plasma androgen levels (Sever 1975, 1976; Woodley 1994). Formerly called the “hodonic gland,” the mental gland secretes nonvolatile protein courtship pheromones that influence multiple aspects of female courtship behavior (Houck et al. 1998; Vaccaro et al. 2009), including the duration of tail straddling walk (for more detail on plethodontid courtship behavior, see Arnold et al. 2017). The morphology and cellular architecture of the mental gland varies among plethodontid species (Houck and Sever 1994). Despite diversity in architecture, nearly all plethodontid males apply pheromones from the mental gland by one of two methods. In most plethodontid species, male salamanders scratch the female dorsum using hypertrophied premaxillary teeth, presumably allowing pheromones to diffuse into the female bloodstream and bind to receptors on target tissues (Houck et al. 2007a; Kiemnec-Tyburczy et al. 2011). In a single clade of large eastern Plethodon spp., however, males lack premaxillary teeth, the mental gland is a large protruding pad-like structure, and pheromones are delivered by slapping the mental gland on the female snout and nares (olfactory delivery; Arnold 1976; Houck and Arnold 2003). During the evolutionary transition between scratching and slapping delivery, some Plethodon species used both delivery modes and are referred to as having intermediate delivery (Picard 2005; Arnold et al. 2017).

Detection of Plethodontid Pheromones

Detection of environmental odorants and conspecific signals is accomplished by two sets of chemosensory neurons in the nasal cavity: the main olfactory epithelium (MOE) and vomeronasal epithelium (VNE; Woodley 2007; Isogai et al. 2011). Our understanding of both systems is largely informed by studies in rodents (reviewed in Dulac and Torello 2003). In mice, the vomeronasal organ (that contains the VNE) is a small, bone-encapsulated space ventral to the main nasal cavity (that includes the MOE). Each sensory neuron in the MOE expresses a single olfactory receptor (OR), limiting the number of ligands that each neuron can recognize (Buck and Axel 1991; Rodriguez 2013; Bavan et al. 2014). VNE neurons contain one of two receptor types: vomeronasal Type-1 receptors (V1Rs) that are structurally similar to ORs in that they detect volatile, organic compounds, and vomeronasal Type-2 receptors (V2Rs) that contain large N-terminal extracellular domains that bind peptide or protein pheromones (Dulac and Axel 1995; Herrada and Dulac 1997; Leinders-Zufall et al. 2009). ORs, V1Rs, and V2Rs are evolutionarily distinct and have been separately coopted for olfaction (Yang et al. 2005).

In plethodontid salamanders, both sets of sensory epithelia are housed in a single nasal cavity. The VNE lies along the lateral edge of this cavity and the MOE is medial. The VNE—but not the MOE—is sexually dimorphic, with males having approximately twice as many VNE neurons as females (Woodley 2007). As with other nonmammalian vertebrates, the molecular organization of the MOE and VNE do not precisely correspond to the patterns observed in rodents. Sequences of V2Rs and ORs could be amplified by degenerate reverse transcription-polymerase chain reaction (RT-PCR), but those of V1Rs were not detected (Kiemnec-Tyburczy et al. 2012). Exogenous application of mental gland pheromones specifically activates VNE neurons (Wirsig-Wiechmann et al. 2002) that then project and activate neurons in the vomeronasal region of the amygdala, the preoptic area, and the ventromedial hypothalamus. These regions of the vertebrate brain often mediate reproductive behavior (Laberge et al. 2008). In summary, these results indicate that the mental gland pheromones of large eastern Plethodon spp. regulate female courtship behavior by binding to V2Rs on VNE neurons that project to specific regions of the female brain to modulate courtship behavior (Wirsig-Wiechmann et al. 2002; Laberge et al. 2008; Kiemnec-Tyburczy et al. 2012).

Patterns of Molecular Evolution in Plethodontid Pheromones

Our working hypothesis, based on phenotypic tango models (Arnold and Houck 2016), is that female receptor sequences create positive (directional) sexual selection on male courtship pheromones that pull them from their natural optimum, leading to coevolution between the pheromones and their receptors (Fig. 1). Molecular evolutionary studies on multiple pheromone gene families in plethodontid salamanders have consistently revealed elevated rates of nonsynonymous-to-synonymous substitutions (dN/dS) in their protein coding sequences, consistent with positive Darwinian selection and a coevolutionary tango (Watts et al. 2004; Palmer et al. 2005, 2007b, 2010). The
plethodontid mental gland secretes a suite of diverse protein pheromones co-opted from multiple gene families. Insight into the structural and expression profiles of mental gland pheromones across the scratching-slapping transition was gained by conducting proteomic and biochemical analyses of the pheromones of three species: a Plethodon with slapping delivery (Red-Legged Salamander, *P. shermani*), a Plethodon with scratching delivery (Eastern Red-backed Salamander, *P. cinereus*), and a non–Plethodon scratching species (*Ocosee* Salamanders, *Desmognathus ocoee*; Fig. 2; Wilburn et al. 2012, 2014b; Doty et al. 2016). Both evolutionary and biochemical studies support a repeated pattern of gene duplication of signaling molecules, sequence and structural adaptation, and co-option for pheromone activity. The next several sections highlight this pattern for the most well-characterized plethodontid pheromone families: plethodontid receptivity factor (PRF), plethodontid modulating factor (PMF), and sodefrin precursor-like factor (SPF).

**Plethodontid Receptivity Factor**

The first identified plethodontid pheromone was PRF: an ~22-kDa protein with shared homology to helical cytokines such as interleukin 6 (IL-6), ciliary neurotrophic factor, and cardiotrophin-1 that are broadly involved in inflammation and local tissue remodeling (Rollmann et al. 1999). As a class of secreted signaling molecules, cytokines are arguably well suited for co-option as pheromones. PRF was first discovered in the slapping species *P. shermani* (Feldhoff et al. 1999; Rollmann et al. 1999). Experimental application of PRF to female salamanders decreased tail straddling walk time by ~20%, similar to that of whole pheromone, and is interpreted as an increase in female mating receptivity (Houck et al. 1998; Rollmann et al. 1999). Further chemical analysis of *P. shermani* PRFs revealed that males commonly express three isoforms of PRF, each corresponding to a separate gene duplication, and referred to as PRF-B, PRF-C1, and PRF-C2 (named after their elution profiles using different forms of high-performance liquid chromatography [HPLC]; Chouinard et al. 2013). With a lack of extensive genomic information, we do not know the exact number of PRF genes in any species, but a combination of molecular and proteomic data indicate a minimum of four gene copies (and possibly dozens). Out of 192 amino acid residues, C1 and C2 differ only by three substitutions, whereas B and C1/ C2 vary by 12/14 substitutions (Rollmann et al. 1999).

Despite the similarity of the three isoforms at the sequence level (>92% shared identity), male salamanders are highly variable in their expression of the three isoforms (Fig. 3A; Chouinard et al. 2013). Experimental application of recombinant PRF-C2 was sufficient to reduce tail straddling walk time by ~20%, similar to a mixture of all three PRFs and to whole pheromone, indicating possible redundancy in isoform function (Houck et al. 2008).

To better explore the evolutionary history of PRF across the plethodontid phylogeny, degenerate primers were designed against the PRF 5’ and 3’ untranslated regions for RT-PCR analysis. PRF was successfully amplified from 27 *Plethodon* spp., yet was not detected by RT-PCR in species outside of *Plethodon* or in shotgun cDNA libraries for *D. ocoee* and *Eurycea guttolineata*, indicating that PRF is likely a *Plethodon*-specific pheromone (Watts et al. 2004; Palmer et al. 2005; Kiemnec-Tyburczy et al. 2009). Molecular evolutionary analyses of PRF revealed a history of strong positive selection, with ~8% of sites having a mean $d_{\omega}/d_{s}$ of 5.45 (Watts et al. 2004). As a paralog of the IL-6 superfamily, PRF adopts a predominantly $\alpha$-helical fold (Houck et al. 2008), and homology models indicate that the majority of positively selected residues localize on one protein face, which includes a nexus formed by the N and C termini (Fig. 3B). Many of these positively selected sites correspond with known sites on IL-6 involved in receptor binding (Watts et al. 2004). However, because PRF most likely binds a V2R and not a more canonical IL-6 receptor (Wirsig-Wiechmann et al. 2006), it is hypothesized that
FIG. 3.—Expression profiles and molecular evolution of plethodontid receptivity factor (PRF). (A) Expression of the three major PRF isoforms (B, C1, and C2) from 104 individual male *Plethodon shermani* from a single population in Macon Co., North Carolina, presented as a triangle plot. Each point represents a different male, and contour lines are drawn to show relative density of individual male profiles. Approximately 18% of male salamanders only express two of the three isoforms, and their data points lie on one of the three edges (adapted from Chouinard et al. 2013). (B) Homology model of PRF-C2, with sites under positive selection highlighted as red spheres (adapted from Watts et al. 2004).
similarity in the PRF and IL-6 binding interfaces are the result of convergent evolution, a reflection of physicochemical constraints of the α-helix topology, or both (Wilburn et al. 2014b).

Additional molecular evolutionary analyses of the PRF sequences between species with scratching, intermediate, or slapping delivery revealed differences in both the sites under positive selection and their evolutionary rates (Palmer et al. 2005). As in P. shermani (slapping delivery), proteomic analyses of P. cinereus (scratching delivery) revealed three major PRF isoforms. Further biochemical analyses revealed that the P. cinereus PRF isoforms were heavily glycosylated, however, with all three isoforms having an N-terminal extension with O-glycans, and two of the three isoforms containing large N-linked glycans that comprised ~40% of the total mass. Although the role of these carbohydrates is unknown, they localize near the N- and C-termini with many of the positively selected sites. Sequence analysis of PRFs across the delivery transition indicates that glycosylated PRFs are only present in species with scratching delivery and some species with intermediate delivery and that they might contribute to the function of PRF in the bloodstream relative to the olfactory system (Wilburn et al. 2014b). Therefore, beyond gene duplication and rapid sequence evolution, at least one structural modification (glycosylation) likely influences the role of plethodontid pheromones across modes of delivery.

Plethodontid Modulating Factor

The second pheromone identified in plethodontid salamanders was the 7-kDa protein PMF, a homolog of the three-finger protein (TFP) superfamily. TFPs are a diverse class of small proteins (~50–90 residues) involved in a plethora of biological functions, including snake venom toxicity, positioning of cells during amphibian limb regeneration, control of the mammalian complement system, and regulation of the plasminogen activation system that is also involved in tissue organization and wound healing (Blasi and Carmeliet 2002; Chang et al. 2004; Levitin et al. 2008; Garza-Garcia et al. 2009). TFPs can be either soluble or membrane bound. More than 200 TFP structures have been described, and despite little sequence conservation, all adopt a conserved fold that includes two parallel β-sheets with 2-and 3-strands, stabilized by eight conserved cysteine residues with a canonical disulfide bonding pattern of 1–3, 2–4, 5–6, 7–8 (based on the order of the cysteines from the N-to-C terminus). These conserved, stabilizing cysteine residues sit near the base of the protein away from the tips of the fingers, which are generally more variable and the regions critical for binding to receptors and ligands (Fig. 4A). Consistent with the functional breadth of different TFP domain-containing proteins, TFP sequences are extremely variable, and there is little conservation beyond the relative spacing of the cysteine residues (Galat et al. 2005).

Like other TFPs, PMF shares the architecture of a hypervariable protein sequence superimposed on a highly stable, disulfide-bonded platform. In P. shermani, compared to three relatively conserved PRF isoforms, PMF is extraordinarily variable, with >30 expressed isoforms that only share ~30% amino acid identity. Similar to PRF, limited genomic data prevents us from knowing the exact number of PMF genes, but we estimate that P. shermani has a minimum of 13 gene copies (up to ~100; Wilburn et al. 2012). This diverse repertoire of PMF isoforms is also variable in percent composition among male salamanders within a population, although the complexity of the mixture makes it difficult to analyze specific components by HPLC (Chouinard et al. 2013). However, three major isoforms are consistently expressed at high levels in nearly all males (isoforms PMF-G, PMF-H, and PMF-I). The solution structure of PMF-G was determined using mass spectrometry and multidimensional nuclear magnetic resonance, revealing both a unique disulfide bonding pattern and structural topology compared to all members of the TFP superfamily. In contrast to the invariant disulfide pattern for other TFPs, the pattern for PMF-G is 1–2, 3–6, 4–5, 7–8 (Fig. 4A). This altered pattern in three of the four disulfides changes the protein topology from two parallel β-sheets of 2-and 3-strands to two perpendicular sheets each with 2-strands (Fig. 4B). This key structural perturbation results in increased backbone flexibility in the third finger. In this structural region, PMF isoform sequences are most variable, both in number of rapidly evolving sites and presence of insertions and deletions (Wilburn et al. 2014a).

We hypothesized that this combination of sequence and structural diversity increases the number of available PMF conformations that can dock with any possible female receptor, and possibly compensates for small changes in receptor sequences (Wilburn et al. 2014a). Consistent with this hypothesis, both neurophysiological and behavioral data indicate that complex PMF mixtures are required to increase female receptivity. When a complete mixture of PMF isoforms was applied to female salamanders and compared to a saline control, the number of activated VNE neurons was increased and female courtship time was decreased by ~20% (similar to both whole pheromone extract and PRF). Whereas a single PRF decreased courtship time, PMF-G in isolation had no effect on neuronal activation or female behavior (Wilburn et al. 2015, 2017). Surprisingly, when female P. shermani received a mixture of PMFs missing the major isoforms (PMF-G, PMF-H, and PMF-I), time in tail straddling walk increased, indicating a decrease in female receptivity (Houck et al. 2007b). Hence, our prevailing hypothesis is that PMF isoforms function synergistically to enhance male reproductive success, giving a mating advantage to males that express more complex mixtures of PMF.

Studies of PMF in other species besides P. shermani further support a pattern of gene duplication, rapid evolution, and structural adaptation. PMF is both transcriptionally and proteomically present in all plethodontid species that have been examined (Palmer et al. 2007a, 2010; Kienene-Tyburczy et al. 2009; Wilburn et al. 2012, 2014b; Doty et al. 2016). Tests of molecular evolution across 27 plethodontid species show more exacerbated bursts of duplication and elevated $d_s/d_\theta$ compared to PRF (Palmer et al. 2010). Proteomic analysis of PMF in species with scratching delivery revealed fewer PMF isoforms (one in D. ocoee, approximately four to six in P. cinereus) that constituted a smaller percentage of the total pheromone protein compared to P. shermani (Wilburn et al. 2014b; Doty et al. 2016). Notably, proteomic analyses of PMF in both scratching species revealed unique properties. The single D. ocoee PMF, which constitutes <1% of the total pheromone
FIG. 4.—Structural modification and evolution of plethodontid modulating factor (PMF). (A) Comparison of the three-finger protein (TFP) and PMF disulfide bonding patterns, with gray and black lines drawn, respectively, to represent disulfide bonds between cysteine residues. (B) Comparison of the secondary structure and topology of a representative TFP (PDB 1IQ9), with each of the three fingers numbered. PMF-G is missing one of the conserved β-strands and there has been a rotation of the β-sheet in finger 1 relative to finger 2. (C) Model of molecular evolution, with sites under purifying selection in blue, neutrality in yellow, and positive selection in red as spheres (adapted from Wilburn et al. 2014a).
protein, has a net positive charge, compared with a high net negative charge in all other characterized PMFs (Doty et al. 2016). In P. cinereus, a new class of PMFs was found that included a large insertion between the signal peptide and the normal ~65 residues of PMF. Proteomic analyses revealed that this N-terminal PMF-precursor peptide (NPP) is cleaved from the C-terminal PMF and secreted as a separate 9.8-kDa peptide in the pheromone mixture. Although the specific function of this putative new pheromone remains unknown, tethering of NPP to PMF through a common transcript and nascent peptide chain would hypothetically favor coexpression at nearly equimolar levels and may relate to a synergistic function between the two proteins (Wilburn et al. 2014b).

The difference in isoform abundance between scratching and slapping species may be largely explained by a model of signal amplification. As in many endocrine systems, pheromones delivered by scratching will become diluted in the female bloodstream, requiring sufficient concentration to bind to female receptors and elicit a response. As the mental gland already directs nearly all of its transcriptional and translational machinery toward pheromone synthesis, it is unlikely that changes in pheromone gene copy number could further increase total pheromone expression (Wilburn et al. 2012). As such, each time a pheromone gene duplicates and changes function (neofunctionalization), if the total protein expression of that pheromone remains nearly constant, the concentration of the originally selected high-affinity variants would be reduced—possibly below its threshold level necessary for proper female stimulation. However, delivery of pheromones to the female olfactory system offers the opportunity for signal amplification: because V2Rs can sense pheromones at very dilute concentrations (subnanomolar levels), and the signal is amplified through neural activation to the female brain (Leinders-Zufall et al. 2009). Further evidence for the importance of concentration restricting gene duplication can be found in rapidly evolving gamete recognition proteins of marine mollusks (Wilburn and Swanson 2016). Hence, the transition from scratching to slapping delivery might have resulted in relaxed purifying selection (that normally restricts duplication of PMF genes), allowing for the evolution of more complex isoform mixtures.

Sodefrin Precursor-Like Factor

The first peptide pheromone identified in vertebrates was the decapeptide sodefrin from Firebelly Newts (Cynops pyrrhogaster) that is released by males as a chemoattractant (Kikuyama et al. 1995). This decapeptide might also facilitate sexual communication during the newt tail-nudging walk, which is analogous and perhaps homologous to the plethodontid tail-straddling walk (Houck and Arnold 2003; Arnold et al. 2017). Sodefrin activity is highly species and population specific, as evidenced by similar molecules in other Cynops populations (Yamamoto et al. 2000; Nakata et al. 2007). Sodefrin is cleaved from the C terminus of a larger ~20-kDa precursor protein, and homologs of this precursor—SPF—are found across the salamandrid radiation. The release of the sodefrin decapeptide is a relatively recent acquisition exclusively in Cynops spp. (Janssenswillen et al. 2014), but intact SPF serves as a pheromone in other salamandrids (Janssenswillen et al. 2014; Van Bocxlaer et al. 2015, 2016) and the Mexican Axolotl (Ambystoma mexicanum), an ambystomatid (Hall et al. 2016; Maex et al. 2016).

Hence, similar to many discrete peptide hormones being cleaved from a common precursor, the sodefrin decapeptide represents an evolutionary innovation in which a pre-existing pheromone neo-functionalized to a dual-use pheromone, allowing coupled, simultaneous expression of both SPF and sodefrin (Janssenswillen et al. 2014).

SPF, which is also present in plethodontid salamanders, is the oldest known pheromone family, dating to the common ancestor of salamandrids, ambystomatids, and plethodontids, 160 million years ago (Shen et al. 2016; Van Bocxlaer et al. 2016; Arnold et al. 2017). The closest homologs of SPF are γ-type phospholipase A2 inhibitors used by snakes to possibly protect themselves from their own phospholipase A2 toxins (So et al. 2008; Kinkawa et al. 2010). These inhibitors likely resulted from the tandem duplication and fusion of two TFP domains into a single polypeptide chain. Analysis of the SPF gene structure supports the “double TFP” hypothesis, based on the location of exon-intron splice boundaries compared to other TFPs. Hence, SPF and PMF both represent instances of co-option of TFP members for pheromone activity in plethodontid salamanders. Like PMF, neither of the two TFP-like domains in SPF—for which the structure is undescribed—have the canonical TFP disulfide bonding pattern (Fig. 5A; Doty et al. 2016). SPF was first discovered in P. shermani as a single cDNA in a library of >300 sequenced transcripts (Kiemnec-Tyburczy et al. 2009) and was amplified by RT-PCR from the mental glands of all sampled plethodontid species (Palmer et al. 2007b). Proteomic analyses revealed that SPF is the major component of the D. ocoee pheromone mixture (Doty et al. 2016), and application of an SPF-enriched fraction to female D. ocoee decreased time spent in tail straddling walk, indicating functional conservation of pheromone components across the transition from scratching to slapping delivery (Houck et al. 2007a). Chemical analysis of the D. ocoee pheromone mixture revealed several diverse SPF isoforms, reflecting recurrent gene duplication and diversification (with a minimum of six gene copies, but likely dozens). Whereas both PRF and PMF show high variability among male P. shermani in a single population, comparative molecular and proteomic analyses in D. ocoee indicate high transcriptomic variability but relatively stable protein expression of the approximately six major SPF isoform types among males in a single population (Fig. 5B). A component of this transcriptomic variation is the presence of hybrid SPFs that contain segments of two of the major SPF isoform classes. These hybrid sequences do not appear to be alternatively spliced variants, and they most likely have resulted from gene duplication and recombination mediated through more conserved genic segments (Doty et al. 2016). Similar patterns of recombination have been found in snake venom TFPs, where gene duplication followed by positive selection is quite common (Doley et al. 2008).

As with both PRF and PMF, SPF shows a history of strong positive selection across the plethodontid phylogeny (Palmer et al. 2007b). Branch models of molecular evolution revealed that, for different clades of large eastern Plethodon spp., estimates of dS/dD for PRF and SPF are negatively correlated. Given the similarity in behavioral effects between SPF and PRF, and the inference that incessant pheromone evolution is driven by a molecular tango with female
receptors, it was hypothesized that PRF has replaced SPF in lineages where PRF is evolving more quickly (Palmer et al. 2007b). Biochemical data further support this hypothesis. PRF is a major component of the total pheromone protein in clades where it evolves more quickly (e.g., *P. glutinosus* and *P. cinereus*; Wilburn et al. 2012, 2014b), but not in clades where SPF evolves more quickly (e.g., *P. welleri*; D. Wilburn and R. Feldhoff, personal observations). Hence, although some functional redundancy exists in all three major pheromone families (Rollmann et al. 1999; Houck et al.

---

**Fig. 5**—Disulfide bonding and expression profiles of sodefin precursor-like factor (SPF). (A) Comparison of the disulfide bonding patterns of the two three-finger protein (TFP)-like domains in SPF relative to several representative TFPs. (B) cDNA and protein expression profiles of the six major SPF classes and hybrid sequences (same key for both data sets) for multiple individual male *Desmognathus ocoee* from a single population in Clay Co., North Carolina (adapted from Doty et al. 2016).
Additional Putative Pheromone Families

Although PRF, PMF, and SPF are the most extensively characterized families of salamander pheromones, they do not represent the complete repertoire of proteins in the mental gland secretions (nor have we fully examined all species with mental glands). A combination of cDNA libraries, RNASeq data, and biochemical fractionation has revealed several additional protein components in the mental gland extract of different plethodontid species. Although there are no behavioral assays or neurophysiological results indicating whether these additional protein components are bioactive pheromones, they often comprise a nontrivial percentage of the mental gland secretions and likely have some function that relates to courtship behavior. Consider the following four examples. (1) Approximately 10% of the total protein in the P. shermani mental gland extract is a 19-kDa protein termed C3 (Chounard et al. 2013), but it was renamed plethodontid TIMP (tissue inhibitors of metalloproteinase)-like protein (FTP; Wilburn et al. 2014b) after its homology to TIMPs was determined via deep sequencing. FTP is also a major component of the P. cinereus mental gland extract (Wilburn et al. 2014b), and the mean pairwise $d_S$/dS between the two species is ~1, a value that likely indicates that it is also under positive selection (Swanson et al. 2004). (2) Also in the P. cinereus extract is a holomog of cysteine-rich secretory protein 1, a mammalian protein that facilitates egg–sperm interaction (Evans 2002; Wilburn et al. 2014b). (3) A cDNA library of the E. guttolineata mental gland found that ~7% of transcripts coded for a predicted protein with homology to natriuretic peptide (Kiemnec-Tyburczy et al. 2009), a small hormone that stimulates sodium secretion and vasorelaxation (Matsuo 2001). Finally, (4) in D. ocoee, using both transcriptomic and proteomic analyses (Kiemnec-Tyburczy et al. 2009; Doty et al. 2016), a suite of hormone-like peptides was identified: mental gland glucagon-like peptide (mgGLP), VIP-like peptide, insulin-like protein (ILP), relaxin-like protein, and leptin-like protein (LLP). For these hormone-like peptide/protein families, multiple isoforms were identified with substitutions in several of the classical receptor binding residues, and mutation of dibasic cleavage sites that are essential for normal processing of hormone precursors into bioactive peptides. Although the importance of these structural perturbations is unknown with regard to influencing female behavior or physiology, it is noteworthy that three of these families—mgGLP, ILP, and LLP—relate to hormones involved in controlling metabolism (Doty et al. 2016). Multiple studies in P. shermani have shown that the olfactory-based pheromones can influence female feeding behavior (Vaccaro et al. 2009, 2010), and it is possible that these hormone-like peptides/proteins might have similar effects via endocrine interference in species with scratching delivery (Doty et al. 2016). Currently, the limited sequence data for each of these putative pheromone families prevent detailed molecular evolutionary analyses, and additional research is required to determine whether they follow similar evolutionary patterns to PRF, PMF, or SPF suggestive of a molecular tango with female receptors.

Conclusions

We have reviewed >20 yr of biochemical, molecular, behavioral, and neurophysiological studies that support our molecular tango framework that helps explain conspicuous features in the evolution of plethodontid courtship pheromones. In particular, this framework helps us understand the incessant and rapid molecular evolution that is driven by directional (positive) selection in each of the three families of proteins that constitute the mental gland pheromone. Research on mental gland pheromones has also established procedures for behavioral assay of pheromone effects on females and biochemical approaches for isolating and characterizing pheromone components. Despite these important accomplishments on the molecular, behavioral and biochemical fronts, much remains to be done.

1. We have a good understanding of just one side of what is undoubtedly a coevolutionary picture. Although our understanding of male pheromone evolution is substantial, we know very little about pheromone receptors and their evolution (Kiemnec-Tyburczy et al. 2012).

2. Our understanding of sexual communication via pheromones is incomplete in several ways.

(a) Although three components of the mental gland pheromone (PRF, PMF, SPF) have been intensively studied, other components have been identified, but not studied.

(b) The biochemistry of mental gland pheromones is well known in only three species (Desmognathus ocoee, Plethodon cinereus, and P. shermani). A comprehensive biochemical understanding of secretions in the other 10 tribes of plethodontids would almost certainly present many surprises.

(c) In addition to the mental gland, other male skin glands probably produce pheromone signals that affect female sexual behavior, but we know almost nothing about these secretions. Glands on the cheeks and tail base of male Eurycea are just one example (Noble 1929).

3. The neurophysiological response of the female to courtship pheromones needs much more study. We need to know how individual neurons respond to individual components of the pheromone. The location of female receptors for pheromones delivered by scratching is unknown.

4. In many respects, the complexity of the mental gland pheromone within and among males in the same population is a puzzle. Female reactions to variations in the cocktail need to be studied, but the current behavioral and neurophysiological assays are so time-consuming as to limit the number of experiments that can be conducted during a relatively short mating season; thus, new method development is needed.

5. The mental gland is but one element in a functional complex that includes delivery behaviors, sexually dimorphic premaxillary teeth, as well as female behavior and neurological substrates. All of these elements interact and coevolve. As research moves forward, we
need to expand our concept of coevolution (Arnold et al. 2017).

(6) The molecular tango is not a formal model. We have a formal model of the phenotypic tango that we used to produce animations and explore the factors that might affect coevolution of pheromone and receptors (Arnold and Houck 2016). The predictions from the phenotypic tango are about the summary properties of the pheromone mixture, however, rather than about the sequence of individual proteins. We need a formal model of the molecular tango that, for example, makes predictions about $d_S/d_*$ ratios.

Acknowledgments.—We thank the many students whose research efforts over the past 30 yr made this review possible. We also extend a special thank you to K. Bowen for generous support over the past 15 yr, supporting the collaboration and mentoring everyone in the Feldhoff lab. We thank the anonymous reviewers for helpful comments on the manuscript. We also thank J. Costa and Highlands Biological Station for continued support of our research and for hosting the 2016 Special Conference on Plethodontid Biology. This work was supported by National Science Foundation collaborative grants IOS-1148590 to RGF and to IOS-1147271 to LDH, and National Institutes of Health grant F32-GM16988 to DBW.

Literature Cited


Accepted on 31 May 2017

Guest Editor: Sarah Woodley