Molecular evolution of gland cell types and chemical interactions in animals

Adrian Brückner* and Joseph Parker*

ABSTRACT
Across the Metazoa, the emergence of new ecological interactions has been enabled by the repeated evolution of exocrine glands. Specialized glands have arisen recurrently and with great frequency, even in single genera or species, transforming how animals interact with their environment through trophic resource exploitation, pheromonal communication, chemical defense and parental care. The widespread convergent evolution of animal glands implies that exocrine secretory cells are a hotspot of metazoan cell type innovation. Each evolutionary origin of a novel gland involves a process of ‘gland cell type assembly’: the stitching together of unique biosynthesis pathways; coordinated changes in secretory systems to enable efficient chemical release; and transcriptional deployment of these machineries into cells constituting the gland. This molecular evolutionary process influences what types of compound a given species is capable of secreting, and, consequently, the kinds of ecological interactions that species can display. Here, we discuss what is known about the evolutionary assembly of gland cell types and propose a framework for how it may happen. We posit the existence of ‘terminal selector’ transcription factors that program gland function via regulatory recruitment of biosynthetic enzymes and secretory proteins. We suggest ancestral enzymes are initially co-opted into the novel gland, fostering pleiotropic conflict that drives enzyme duplication. This process has yielded the observed pattern of modular, gland-specific biosynthesis pathways optimized for manufacturing specific secretions. We anticipate that single-cell technologies and gene editing methods applicable in diverse species will transform the study of animal chemical interactions, revealing how gland cell types are assembled and functionally configured at a molecular level.

KEY WORDS: Cell type evolution, Exocrine glands, Chemical ecology, Single-cell biology, Terminal selectors, Gene duplication

Introduction
Molecular biology may, in fact, increasingly shape the questions that are asked in chemical ecology. How do given signal molecules arise in the course of evolution?

Eisner and Meinwald, 1995

For the majority of organisms, adaptation to new ecological frontiers has been achieved via innovations in chemical secretion (Eisner, 2003; Raguso et al., 2015). The ability to synthesize and secrete small molecules, metabolites and proteins enables organisms to influence their surroundings, and is an adaptive solution to many environmental challenges, including detoxification, desiccation avoidance, adhesion, reproduction and antimicrobial protection (Betz, 2010; Blomquist and Bagnères, 2010; Brunetti et al., 2018; Feyereisen, 2012; Wyatt, 2014b). In the Metazoa, evolutionary changes in chemical production have been instrumental to the emergence of interactions both within and between species, with behaviors as diverse as chemical defense, pheromonal communication and parental care relying on transmission of information or resources embedded in chemical secretions (Berenbaum, 1995; Blum, 1981; Eisner et al., 2005; Kumar et al., 2014; Leonhardt et al., 2016; Steiger and Stökl, 2018; Thiel et al., 2019).

The discipline of chemical ecology addresses how relationships among organisms – be they animals and plants, predators and prey, or symbiotic – are mediated by chemical compounds (Berenbaum, 1995; Hay, 2009; Raguso et al., 2015). From its emergence approximately 60 years ago, chemical ecology has been an interdisciplinary field, using analytical chemistry to elucidate the structures of compounds, and combining it with ecological and behavioral studies to understand how such compounds influence community interactions (Eisner, 2003; Eisner and Meinwald, 1995; Hartmann, 2008; Meinwald and Eisner, 2008). Beyond this principal goal, chemical ecology has facilitated the discovery of thousands of natural products from all domains of life, and has been fundamental to explaining how many such compounds are synthesized (http://www.pherobase.com; Morgan, 2010; Symonds and Elgar, 2008; Walsh and Tang, 2017; Wyatt, 2014a). Yet, despite the success of the discipline in exposing the chemical basis of ecological interactions, there is a level of understanding at which knowledge remains fragmentary. This is the genetic and cellular realm, where the molecular mechanisms that manufacture the compounds of interest remain largely enigmatic (Raguso et al., 2015; Tittiger, 2004). In many species, routes of metabolite processing may be known or readily inferred (Morgan, 2010; Symonds and Elgar, 2008; Walsh and Tang, 2017), but the biosynthetic enzymes performing the stepwise covalent modifications are typically much more obscure. Poorer still is an understanding of the mechanisms of chemical release from cells: the identities of molecular components regulating the subcellular trafficking and secretion of chemical signals are unknown for the majority of gland cell types. Comparably little is also known about the specification and differentiation of gland cells – the upstream transcriptional processes that trigger expression of this enigmatic battery of biosynthetic and secretory proteins during development, and maintain it during a gland’s lifetime. In molecular terms, what makes a cell a gland cell remains a mystery in modern biology, despite the centrality of glands and chemical secretions to most multicellular living systems.

Here, we argue that evolution of the molecular processes underlying chemical production must surely shape how and why specific compounds are employed by different species. It follows that understanding the molecular architecture of gland cells, and the evolutionary assembly of their biosynthetic and secretory capabilities, is indispensable for explaining why chemically mediated interactions between organisms have taken on the diverse
forms that we observe. Knowledge of the mechanisms governing compound biosynthesis and secretion is an essential counterpart to the comparatively well-studied olfactory, gustatory and nociceptive mechanisms by which such compounds are perceived (Baer and Mayer, 2012; Hansson and Stensmyr, 2011). We propose a molecular evolutionary framework for how gland cell types and their chemistries might evolve, and emphasize the utility of a molecular perspective in explaining patterns of chemical usage in the Metazoa. We predict that the field of chemical ecology is poised to uncover molecular forces shaping the ecology and evolution of biological interactions.

Convergent evolution of animal gland diversity

In animals, and in many cases in fungi and plants, chemical interactions with the outside world are driven by the underlying evolution of exocrine glands (Downing, 1991; Leonhardt et al., 2016; Symonds and Elgar, 2008). Here, we define exocrine glands as any structure – from a single cell or epithelial patch to a complex multicellular organ – that is specialized for the secretion of a compound (or compounds) to the external environment (Billen, 1991; Locke, 1969; Noirot and Quennedey, 1974; Simpson, 2012; Tortora and Derrickson, 2017). The spectrum of exocrine gland anatomical and functional diversity is vast, with chemical secretions playing critical roles in all metazoan phyla (see Box 1 and Fig. 1). This impressive variety, coupled with the sheer range of chemical secretions that animals are collectively capable of producing, makes it difficult to argue that exocrine glands in general are homologous across metazoans. In contrast to the deep conservation of many body parts within phyla, such as hearts or eyes or limbs (Davidson and Erwin, 2006; Land and Fernald, 1992; Shubin et al., 1997), exocrine glands are a highly homoplastic category of organs – a ‘dumping ground’ for secretory structures grouped together largely by function rather than by evolutionary relatedness. The phylogenetic distribution of any one type of gland is usually highly restricted

Box 1. A snapshot of animal gland diversity

It is impossible to capture the true spectrum of animal exocrine glands, such as their recurrent evolution and varied roles in organismal biology. Even early-diverging animals, which are devoid of neurons and complex organ systems, nevertheless possess glands. In sponges, rhabdiferous cells produce a mucous cloak, composed of glycosaminoglycans that likely safeguard the animal from toxins and pathogens (Simpson, 2012). Gland cells are seemingly also indispensable in placozoans, despite members of this phylum possessing the smallest number of cell types. In Trichoplax adhaerens, gland cells occur in the ventral epithelium, and are potentially neurosecretory-like in function, expressing a FMRFamide family peptide (Smith et al., 2015, 2014). Metazoan exocrine glands are typically specialized secretory structures with adaptive functions that fall into three general categories: physiology, communication and defense, although some glands fall into two or even three of these categories. Some examples below provide a small snapshot of animal gland diversity.

Organismal physiology and life history: Unique glands with distinct chemistries have evolved in different metazoan lineages for physiological homeostasis, exemplified by mammalian sebaceous glands (Fig. 1O) and insect oenocytes (Fig. 1L), the latter being responsible for waterproofing the exoskeleton with cuticular hydrocarbons (CHCs) that also function as insectivorous mammals, venoms appear to be derived from modified salivary gland cells (Fry et al., 2009, 2006; King and Hardy, 2013; Ligabue-Braun et al., 2012). The jaws of some marine polychaete worms connect to venom glands that release a cocktail of membrane-disrupting toxins, neurotoxins and protease inhibitors that overcome macroscopic prey (von Reumont et al., 2014), while velvet worms (Onychophora) use adhesive, protein-based slime secretions ejected by unique cephalic glands within the oral papillae to capture prey and for defense (Fig. 1C; Baer and Mayer, 2012; Benkendorff et al., 1999). In mollusks, the hypobranchial gland is tasked with mucus production, and is further modified in cephalopods to produce defensive ink (Benkendorff, 2010; Morton, 1977; Roseghini et al., 1996). Sea stars and brittle stars (Echinodermata) possess various exocrine glands that secrete mucopolysaccharide mucus, but also alarm- and escape-eliciting pheromones (Buchanan, 1963; McClintock and Baker, 1978). Turbellarian platyhelminths produce curious ‘epitheliosomes’ – granular or filamentous structures released from epithelial cells, which may repel predators (Tyler, 1984; Whittington and Cribb, 2001). Defensive glands that release volatile compounds have also arisen an inordinate number of times, separately in almost two dozen families of beetles (Fig. 1N; Dettner, 1993; Francke and Dettner, 2005) as well as in hemipteran bugs, cockroaches, lepidopterans, myriapods, mites (Fig. 1G), harvestman and others (Blum, 1981; Eisner et al., 2005; Raspopthig, 2010; Raspopthig et al., 2017; Shear, 2015). In rove beetles (Staphylinidae), distinct abdominal defense glands have evolved in at least five different subfamilies (Fig. 1M; Dettner, 1993; Francke and Dettner, 2005), collectively producing a cache of noxious hydrocarbons, alcohols, aldehydes, ketones, acids, esters, iridoids, quinones and terpenes.

Chemical defense: The variety of aversive, toxin-producing glands, which have evolved for anti-predator defense or for prey capture, is equally striking (Eisner et al., 2005; Fry et al., 2009). In cidriadians, ecdycomial gland cells exist in many species that release potent neurotoxins – a possible ancestral mechanism of venom delivery that may predate the origin of nematocysts (Jouiaei et al., 2015; Moran et al., 2011). Small-molecule- or peptide-secreting cells have independently evolved to fuel the venom glands of aculate Hymenoptera, scorpions, stings and the cranial glands of male duck-billed platypus (Calvete, 2013; Derby, 2014; Eisner et al., 2005; Hoffman, 2010). In spiders, snakes, Heloderma lizards and a handful of insectivorous mammals, venoms appear to be derived from modified salivary gland cells (Fry et al., 2009, 2006; King and Hardy, 2013; Ligabue-Braun et al., 2012). The jaws of some marine polychaete worms connect to venom glands that release a cocktail of membrane-disrupting toxins, neurotoxins and protease inhibitors that overcome macroscopic prey (von Reumont et al., 2014), while velvet worms (Onychophora) use adhesive, protein-based slime secretions ejected by unique cephalic glands within the oral papillae to capture prey and for defense (Fig. 1C; Baer and Mayer, 2012; Benkendorff et al., 1999). In mollusks, the hypobranchial gland is tasked with mucus production, and is further modified in cephalopods to produce defensive ink (Benkendorff, 2010; Morton, 1977; Roseghini et al., 1996). Sea stars and brittle stars (Echinodermata) possess various exocrine glands that secrete mucopolysaccharide mucus, but also alarm- and escape-eliciting pheromones (Buchanan, 1963; McClintock and Baker, 1978). Turbellarian platyhelminths produce curious ‘epitheliosomes’ – granular or filamentous structures released from epithelial cells, which may repel predators (Tyler, 1984; Whittington and Cribb, 2001). Defensive glands that release volatile compounds have also arisen an inordinate number of times, separately in almost two dozen families of beetles (Fig. 1N; Dettner, 1993; Francke and Dettner, 2005) as well as in hemipteran bugs, cockroaches, lepidopterans, myriapods, mites (Fig. 1G), harvestman and others (Blum, 1981; Eisner et al., 2005; Raspopthig, 2010; Raspopthig et al., 2017; Shear, 2015). In rove beetles (Staphylinidae), distinct abdominal defense glands have evolved in at least five different subfamilies (Fig. 1M; Dettner, 1993; Francke and Dettner, 2005), collectively producing a cache of noxious hydrocarbons, alcohols, aldehydes, ketones, acids, esters, iridoids, quinones and terpenes.
In many cases in arthropods for example, glands and their respective chemistries are often family specific, genus specific or even species specific (e.g. Parker, 2016; Raspotnig et al., 2017; Rodriguez et al., 2018; Roelofs and Rooney, 2003; Symonds and Elgar, 2008). Although certain glands are conserved across large, ancient taxonomic groups – mammary glands and insect oenocytes being examples (Blomquist and Baguères, 2010; Chung and Carroll, 2015; Makki et al., 2014; Oftedal, 2002) – such examples are few, and are dwarfed by the myriad unique exocrine glands common only to specific taxa of lower rank (Blum, 1996; http://www.pherobase.com). Moreover, even highly conserved glands can be subject to repeated loss, as evidenced by the absence of mammalian sebaceous glands in cetaceans, hippos, elephants and naked mole rats (Lopes-Marques et al., 2019).

A corollary of exocrine gland convergence is that the secretory cells that constitute animal glands are themselves convergent, being gained and lost as animals chemically adapt to new ecological circumstances. A consequence of this dynamic pattern of gland birth and death is that perhaps no other category of animal cell type exhibits such extensive evolutionary turnover and lineage-specific

**Fig. 1. Diversity of animal exocrine glands.** (A) Wood ant (*Formica* sp.) worker spraying formic acid. (B) Male crested gecko (*Correlophus ciliatus*) femoral pore releasing sex and territorial pheromones. (C) Orb-weaving spider (*Araneus* sp.) secreting silk from opisthosomal gland. (D) Velvet worm (*Euperipatoides rowelli*), ejecting glue-like polymer to capture prey. (E) Myrmecophile histerid beetle (*Chlamydopus* sp.), with elytral trichomes secreting unidentified ‘appeasement’ compounds for host ant. (F) Confocal image of abdominal trichomes of myrmecophile *Diartiger* rove beetle. (G) Scanning electron microscope image of oribatid mite (*Hermaniella* sp.) defensive oil gland. (H) Surface rendering of cephalic organs of leafcutter ant (*Atta vollenweideri*); red, brain; blue, post-pharyngeal gland; yellow, pharynx; orange, optical nerve; lilac, alarm pheromone-producing mandibular gland. (I) Saliva-producing canine submaxillary gland. (J) Human salivary gland. (K) Human mucus-producing saccular gland. (L) CHC-producing oenocyte of small tortoiseshell butterfly (*Aglais urticae*). (M) Confocal image of staphylline rove beetle defensive gland. (N) Confocal image of pygidial gland secretory lobes of ground beetle *Harpalus pensylvanicus*. (O) Human sebaceous gland tissue section stained with Hematoxylin and Eosin. (P) Paraffin section of lumen of axolotl (*Ambystoma mexicanum*) mucus gland, stained with Movat’s pentachrome. Image credits: David Miller (A,B); Hans Braxmeier and Simon Steinberger (C); Alexander Bär (D); Nick Porch (E). Joseph Parker (F,M); Adrian Brückner, Günther Raspotnig and Edith Stabentheiner (G); Marco Smolla and Christoph Kleineidam (H); Roy Winkelman (I,K); A. C. Hollande, 1914 (L); Adam Rork (N); Paul Rigby (O); Thomas Lozito (P).
innovation. The ultimate cause of this hotspot of cellular novelty is obvious; there is a clear adaptive advantage of evolving new chemical secretions. Yet, there is likely another, proximate reason. Glandular properties may be relatively mechanically facile to impart onto cells, rendering their convergent evolution especially likely. In their seminal paper on insect exocrine glands, Noirot and Quennedey (1974) defined three major gland cell classes (Fig. 2). It is clear that these types are not confined to insects but occur much more generally throughout the Metazoa (Müller et al., 2014; Raspotnig et al., 2003; Requena and Sangüeza, 2017; Tucker, 2007). Noirot and Quennedey (1974)’s class 1 gland cells are simply epidermal cells that have gained pronounced biosynthetic and secretory activities. Such cells are continuous with the surrounding epidermis, sometimes appearing as patches of tissue, often sculpted into a reservoir in which the class 1 cells secrete (Fig. 2B). Lepidopteran pheromone glands, scent glands of harvestmen, oil glands of oribatid and astigmatid mites, as well as the wax glands of certain hemipterans are of the class 1 type (Clawson, 1988; Raspotnig et al., 2003; Staddon, 1979; Steinhbrecht, 1964). Class 1 gland cells underscore the basic mechanistic simplicity of gland evolution: through the transformation of regions of pre-existing epithelium into secretory patches or reservoirs, new glands can arise. The two remaining classes of exocrine cells identified by Noirot and Quennedey (1974) are similarly prone to convergence. Class 2 cells are comparable to class 1, although not continuous with the epidermis. Insect oenocytes belong to this class: developmentally, oenocytes delaminate from the ectoderm during embryogenesis (Lawrence and Johnston, 1982), organizing into clusters of cells beneath the epidermis (Fig. 2B), where they secrete cuticular hydrocarbons (CHCs) (Blomquist and Bagnères, 2010; Chung and Carroll, 2015).

Class 3 gland cells are more anatomically complex, forming bicellular units composed of a secretory cell and a duct cell (Noirot and Quennedey, 1974). An ‘end apparatus’ of porous cuticle and microvilli connects the two cells (Fig. 2B). Despite this more sophisticated anatomy, a large number of non-homologous glands are composed of cells that fit into the class 3 category, spanning the majority of insect orders and other arthropods, including myriapods and crustaceans (e.g. Bacchus, 1979; Billen, 2009; Bin and Vinson, 1986; Giglio et al., 2005; Goettler et al., 2007; Hipeau-Jacquotte, 1987; Hölldobler et al., 2018; Liang and Schal, 1993; Quennedey et al., 2002; Rork et al., 2019; Rosenberg, 1983; Steidle and Dettner, 1993). Some exocrine mammalian glands are likewise composed of glandular units with class 3 morphology, including the biosynthetic acinar and secretory duct cells of sebaceous glands, salivary glands and mammary glands (e.g. Hamilton and Montagna, 1950; Hand et al., 1999; Hassiotou and Geddes, 2013; Hovey et al., 2002; Jenkinson et al., 1985; Richert et al., 2000; Smith and Thiboutot, 2008; Tucker, 2007). It seems likely that this general anatomy of a biosynthetic bulb cell joined to a secretory duct has sufficiently straightforward mechanistic origins to have evolved many times. Indeed, in mammalian salivary and mammary glands, a shared molecular pathway controls their branching morphogenesis (Varner and Nelson, 2014).

Whereas glands can be composed of a single, principal secretory cell type, complexity in chemical secretions can often come from combining different gland cell types together in a biosynthetic division of labor. For example, there are multiple instances in which complex glands have evolved from class 3 cells arranged around a central reservoir formed from an invagination of class 1 cells, with the class 1 and 3 cell types secreting different compounds into the reservoir. The defensive gland of aleocharine rove beetles (Staphylinidae) is a model example (Jordan, 1913; Parker et al., 2018 preprint; Steidle and Dettner, 1993) (Fig. 2A,B). Here, cells comprising the intersegmental membrane between two abdominal segments invaginate to form a reservoir inside the dorsal abdomen. The invaginated cells differentiate into class 1 secretory units that produce fatty acid derivatives that fill the reservoir. These function as a solvent for noxious benzoquinones secreted from class 3 glands situated directly posterior, via ducts feeding into the class 1 reservoir (Fig. 2B). It is intuitive how the evolution of developmental patterning could juxtapose different gland cell types to create such

Fig. 2. Animal gland cell types exemplified in the rove beetle abdomen. (A) Aleocharine rove beetle in defensive posture with positions of class 1, 2 and 3 gland cells. Class 2 cells are represented by oenocytes (blue), which are segmentally repeated. Class 1 cells are represented by the defensive reservoir (red), created by an invagination of intersegmental membrane between abdominal segments 6 and 7. Class 3 gland units reside in segment 7, with a biosynthetic bulb and secretory duct that feeds into the reservoir. (B) Oenocytes synthesize long-chain cuticular hydrocarbons (CHCs) onto the cuticle for waterproofing and pheromonal communication. Class 1 cells produce fatty acid-based compounds that can include short chain hydrocarbons, aldehydes and fatty acid esters. Class 3 cells synthesize benzoquinones which dissolve in the fatty acid derivatives inside the reservoir, yielding the final defensive secretion.
chemical synergism. The mechanistic challenge, however, is to explain how the gland cell types themselves were assembled at a molecular level in the first place during evolution. Clearly, natural selection has propelled this process in inordinate contexts, producing a great wellspring of metazoan cell type novelty. But for no single instance in any animal species have we inferred the molecular succession of events that led to the evolution of a novel gland cell type. A missing part of almost every chemical ecology story is the assembly of the molecular machinery that produces the compounds of interest.

The molecular architecture of gland cell types

Each evolutionary origin of a novel exocrine gland involves the creation of new cell types bestowed with the capacity to synthesize and secrete specific compounds. This process – the molecular assembly of gland cell types – is a pervasive and recurring evolutionary phenomenon, but we have only a vague understanding of how glandular machinery is pieced together during evolution. Are new gland cell types assembled from newly-evolved gene products, or from pre-existing molecular components that have been co-opted? How does expression of this machinery come under transcriptional control so that novel gland cell identities can be specified during development? And what molecular constraints exist on how new glands are assembled, and the kinds of chemicals they are able to secrete? In what follows, we discuss what is known about the evolution of biosynthetic pathways in animal exocrine glands and the extent of knowledge of their secretory mechanisms. Our focus is largely on small-molecule-secreting glands where most mysteries remain. We emphasize emergent patterns and possible constraints in the molecular evolution of biosynthesis and secretion, as well as problematic gaps in our understanding. Based on our observations, we propose a model for how new gland cell types are assembled during evolution.

Biostructural pathway assembly in gland cell type evolution

There is an ultimate biosynthetic constraint on the evolution of glandular chemistry, and this stems from the simple fact that all biomolecules are necessarily linked to the chemistry of the primary metabolism (Fischbach and Clardy, 2007; Walsh and Tang, 2017; Wink, 2003). Parent molecules of any secretion are universally limited to amino acids (proteins), fatty acids (lipids), carbohydrates (mono- and polysaccharides), nucleotides (DNA/RNA) and the products of primary metabolism of these compound classes. Chemical secretions can be primary metabolites themselves (Eisner et al., 2005; http://www.pherobase.com; Morgan, 2010), but far more commonly they are specialized secondary metabolites (Morgan, 2010; Walsh and Tang, 2017). However, their limited primary metabolic origins mean there are only seven different ‘logics’ that yield the chemical diversity we know, namely: non-ribosomal and ribosomal peptides (Warner and McIntosh, 2009), polyketide/fatty acid synthase products (Pankewitz and Hilker, 2008), glycosides and saccharides (Gleadow and Woodrow, 2002), nucleosides (Croteau et al., 2000), isoprenoids/terpenes (Breitmaier, 2006), alkaloids (Pelletier, 1983) and phenylpropanoids (Korkina, 2007). All natural products, whether primary or secondary metabolites, are limited to amino acids (proteins), fatty acids (lipids), carbohydrates (mono- and polysaccharides), nucleotides (DNA/RNA) and the products of primary metabolism of these compound classes.

There is an ultimate biosynthetic constraint on the evolution of glandular chemistry, and this stems from the simple fact that all biomolecules are necessarily linked to the chemistry of the primary metabolism (Fischbach and Clardy, 2007; Walsh and Tang, 2017; Wink, 2003). Parent molecules of any secretion are universally limited to amino acids (proteins), fatty acids (lipids), carbohydrates (mono- and polysaccharides), nucleotides (DNA/RNA) and the products of primary metabolism of these compound classes. Biological secretions can be primary metabolites themselves (Eisner et al., 2005; http://www.pherobase.com; Morgan, 2010), but far more commonly they are specialized secondary metabolites (Morgan, 2010; Walsh and Tang, 2017). However, their limited primary metabolic origins mean there are only seven different ‘logics’ that yield the chemical diversity we know, namely: non-ribosomal and ribosomal peptides (Warner and McIntosh, 2009), polyketide/fatty acid synthase products (Pankewitz and Hilker, 2008), glycosides and saccharides (Gleadow and Woodrow, 2002), nucleosides (Croteau et al., 2000), isoprenoids/terpenes (Breitmaier, 2006), alkaloids (Pelletier, 1983) and phenylpropanoids (Korkina, 2007). All natural products, whether primary or secondary metabolites, are restricted to amino acids (proteins), fatty acids (lipids), carbohydrates (mono- and polysaccharides), nucleotides (DNA/RNA) and the products of primary metabolism of these compound classes.

Inasmuch as it is possible to ascertain a pattern, published data points to gene duplication as the major driving force in the assembly of new pathways in novel gland cell types. Most enzymes known to mediate metabolic steps within animal glands appear to have dedicated roles in those glands, but nevertheless belong to larger gene families. A clear example is the task-specific involvement of unique cytochrome P450 enzymes (CYPs) to mediate key oxidation steps in diverse glands. CYPs are heme-thiolate proteins found in all living organisms and constitute one the largest enzyme families (Feyereisen, 1999, 2012). CYPs are usually monoxygenases, catalyzing the transfer of oxygen to a substrate with a non-activated C=H bond while reducing other redox partners (mostly NADPH). In insect oenocytes, a cell-type-specific CYP (CYP4G1) carries out terminal decarbonylation of long-chain aldehydes into the resultant alkane CHCs (e.g. Chen et al., 2016; Qiu et al., 2012; Reed et al., 1994). Specific CYPs function in the midgut and fat body of bark beetles, converting plant monoterpenes into aggregation pheromones (Chiu et al., 2018, 2019, 2015; Nadeau et al., 2017; Sandstrom et al., 2008, 2006). In locusts, a tissue-specific CYP transforms phenylalanine into a precursor for cyanide production (Wei et al., 2019, and in the fall webworm, Hyphantria cunea, another phenome gland-specific CYP catalyzes the epoxidation of open chain precursors of cis-9,10-epoxy-(3Z,6Z)-3,6-henicosadiene and cis-9,10-epoxy-(3Z,6Z)-1,3,6-henicosatriene to their final cyclic forms (Rong et al., 2014). In human skin, a specialized CYP converts cholesterol into pregnenolone (Smith and Thiboutot, 2008; Thiboutot et al., 2003).

Many other enzyme families show comparable gland cell type specificity, underscoring what appears to be a general principle, that enzyme duplication is a key step in gland cell assembly. The
fatty-acyl-CoA reductase (FACR) enzyme family is another example. Fatty acids are used as precursors for compounds with diverse roles in chemical communication. They are produced by primary metabolism of sugars and converted into aldehydes or alcohols by FACRs. A dedicated oenocyte FACR converts fatty-acyl-CoA into an aldehyde precursor for insect CHCs (Cinnamon et al., 2016; Finet et al., 2019; Jaspers et al., 2014). Species-specific FACRs are expressed in female moth sex pheromone glands, where they produce fatty alcohols for conversion into acetate pheromones (Carot-Sans et al., 2015; Hagström et al., 2013; Lassance et al., 2010; Liénard et al., 2010; Moto et al., 2003). In bumble bees, the high diversity of male sex pheromones (so-called ‘male marking pheromones’) has been linked to extensive duplication and divergence of FACRs expressed in labial glands (Tupec et al., 2019). In honey bee workers, FACRs are expressed in different body regions, controlling synthesis of C18–C32 alcohols that likely function as nestmate pheromones (Teerawanichpan et al., 2010). FACRs have been convergently recruited into mammalian epidermal glands: they are enriched in sebaceous glands, as well as eyelid meibomian glands where they control fatty alcohol synthesis required for wax monoesters (Cheng and Russell, 2004).

The repeated involvement of CYPs and FACRs adds weight to a further idea: that duplication within specific enzyme families is especially likely for certain biosynthetic steps. Numerous examples support this notion. In Phaedon leaf beetles (Chrysomelidae), larval defensive glands express a glucose-methanol-choline oxidoreductase (GMCO) that oxidizes an alcohol to make 8-oxogeranial, a precursor for the iridoid chrysomelidal (Rahfeld et al., 2014). In related Chrysomala leaf beetles, this pathway has been evolutionarily lost. Remarkably, however, an independent GMCO duplication has occurred, yielding a salicyl alcohol oxidase. This GMCO is now expressed in the larval defense gland instead, where it carries out an analogous step to its more ancient, distant GMCO paralog: oxidizing host-plant-derived salicyl alcohol into defensive salicylaldehyde (Kirsch et al., 2011). In insect oenocytes, fatty acid elongases (Elovls) have been recruited to extend fatty acid chain length during CHC synthesis.
synthesis (Combs et al., 2018; Wicker-Thomas and Chertemps, 2010; Wicker-Thomas et al., 2015, 2009). In Drosophila, a male-specific Elovl, Bond, is expressed in the ejaculatory bulb where it controls chain elongation in sex pheromone biosynthesis (Ng et al., 2015). Eloplas are found across the Metazoa (Castro et al., 2016; Guillou et al., 2010): in mammals, these enzymes have been independently recruited in sebaceous glands and hepatocytes (Carmona-Antoniangis et al., 2013) for long-chain sphingolipid synthesis. Like the CYP, FACR and Elovl families, desaturase enzymes that insert double bounds into fatty acids have duplicated recurrently to yield gland-specific enzymes (Buček et al., 2015; Fang et al., 2009; Liénard et al., 2008; Liu et al., 2004; Roelofs et al., 2002; Roelofs and Rooney, 2003; Sakai et al., 2009; Wang et al., 2010). Terpene synthases (TPSs) are another large enzyme class, synthesizing the vast diversity of terpenoids (Breitmaier, 2006; Degenhardt et al., 2009). Species-specific TPSs have been characterized in plants, and duplicated paralogs likely underlie the diversity of animal terpenes (Beran et al., 2019; Christianson, 2017).

Compared to the wealth of examples of duplication, cases of enzyme co-option, or de novo enzyme evolution, are rarer. To our knowledge, no glandular biosynthesis pathway is known to have been entirely assembled from co-opted or de novo-evolved enzymes. However, examples do exist of unique, terminal pathway enzymes that appear to have evolved de novo to exploit the products of another pathway or food source, transforming it into a signal for communication or defense. For example, certain millipedes and butterflies transform aliphatic and aromatic hydroxynitriles – either synthesized from amino acids or sequestered from the diet – into hydrogen cyanide (HCN) (Shear, 2015; Zagrobelsky et al., 2008, 2018). HCN liberation is catalyzed by a highly specific enzyme, (S)-hydroxynitrile lyase, which functions solely for this purpose (Dashadphour et al., 2015; Sharma et al., 2005; Yamaguchi et al., 2018). The challenges of identifying and characterizing new classes of enzyme may lead to an underestimate of the prevalence of examples of de novo enzyme evolution. It nonetheless seems evident that, in contrast to such cases, enzyme duplication appears to be the modus operandi, yielding biosynthesis pathways that are gland cell type-specific, and dedicated to manufacturing particular compounds.

The importance of duplication may also hold true for glands that secrete proteins or peptides. In these cases, biosynthesis is transcriptional and translational, without enzyme catalysis necessarily involved. Protein venomous have been widely studied as models of gene family evolution (Fry et al., 2009, 2006; Wong and Belov, 2012), and evidence from venom glands of diverse species including snakes, spiders, cone snails and centipedes indicates that, in contrast to such cases, enzyme duplication appears to be the modus operandi, yielding biosynthesis pathways that are gland cell type-specific, and dedicated to manufacturing particular compounds.

The importance of duplication may also hold true for glands that secrete proteins or peptides. In these cases, biosynthesis is transcriptional and translational, without enzyme catalysis necessarily involved. Protein venomous have been widely studied as models of gene family evolution (Fry et al., 2009, 2006; Wong and Belov, 2012), and evidence from venom glands of diverse species including snakes, spiders, cone snails and centipedes indicates that, in contrast to such cases, enzyme duplication appears to be the modus operandi, yielding biosynthesis pathways that are gland cell type-specific, and dedicated to manufacturing particular compounds.

The essential counterpart to evolving a chemical biosynthesis pathway (or novel secreted proteins) is the installation of a high-capacity secretory apparatus. Secretory mechanisms are poorly characterized for virtually all animal gland cell types with the exception of protein-secreting glands, which we discuss briefly in Box 2. For small-molecule-producing gland cell types, evidence indicates a high degree of idiosyncrasy in the secretory machinery, although its functional significance is unclear. A key question is how secretory systems are specialized for efficient transport and release of the corresponding compound. In salivary glands, a complex system of intramembrane channels and transporters governs fluid and electrolyte release (Roussa, 2011). Like several other mammalian exocrine glands, salivary glands comprise terminal ‘acinar’ cells with biosynthetic capacity, linked to duct cells that conduct the acinar secretions. A diversity of ion channels, ATPase transporters and water-permeable aquaporin channels releases isotonic saliva from acinar cells, which is further modified by ion flux through duct cell channels to create the final, hypotonic saliva (Roussa, 2011). A further example of specialized transport occurs in insect oenocytes. Large amounts of CHCs are present in hemolymph lipoproteins termed lipophorins, which take up hydrocarbons from oenocytes and transport them to the cuticle. At least in lepidopterans, this process seems to be highly selective: lipophorins target long-chain hydrocarbons to the cuticle, whereas shorter-chained compounds are targeted to abdominal glands for use as pheromone precursors (Fan et al., 2002; Gu et al., 1995; Matsuoka et al., 2006; Schal et al., 1998a,b). It is unclear how hydrocarbons move from lipophorins to their sites of secretion. In Drosophila, lipid-shuttling lipophorins bind to low density lipoprotein receptors (LDLRs) expressed on recipient tissues, causing lipids to be unloaded for use in metabolism (Parra-Peralbo and Culi, 2011). LDLRs on epidermal or pheromone gland cells may similarly bind hydrocarbon-shuttling lipophorins causing lipids to be unloaded for use in metabolism (Parra-Peralbo and Culi, 2011). LDLRs on epidermal or pheromone gland cells may similarly bind hydrocarbon-shuttling lipophorins dispatched from oenocytes. Even if so, final secretion of the hydrocarbon by the target cells is not understood in molecular detail.

Although small-molecule secretory mechanisms have been acutely understudied, it is likely that active transport of such compounds out of cells occurs. Active secretion rather than simple diffusion is likely needed to circumvent cytotoxicity arising from small-molecule accumulation. Active transport may also counter the lipophilic affinity of many small molecules, redirecting them from lipid bilayers (Widhalm et al., 2015). What these transport
Box 2. Protein secretion and the unfolded protein response

Many animal cell types, including 'non-gland' cells, possess the capacity for protein secretion, and a conserved molecular and organellar pathway exists for routing translated proteins out of the cell. The classical secretory route transports nascent polypeptides via the endoplasmic reticulum (ER) where they are folded and post-translationally modified prior to trafficking through the Golgi, where further modifications occur. The mature protein is then secreted from the cell via constitutive or regulated exocytosis (Kelly, 1985). Additional, unconventional routes for protein secretion have also been discovered (Rabouille, 2017). Discussion of these widely studied secretory mechanisms is beyond the scope of this article. However, dedicated protein-producing gland cells have an exceptionally high secretory load relative to most cells. Consequently, during the evolutionary assembly of novel protein-secreting gland cell types, supporting mechanisms must be put in place to accommodate mass protein trafficking. One key mechanism is the unfolded protein response (UPR): a quality control system that ensures reliable folding of proteins in the ER (Hetz, 2012). The UPR is triggered by the ER intramembrane protein IRE1, which binds misfolded proteins. Detection of misfolding causes IRE1 to cleave transcripts of X box-binding protein 1 (xbp1), yielding an mRNA encoding the XBP1 transcription factor. On translation, XBP1 enters the nucleus to drive expression of chaperones that promote protein folding in the ER, as well as others that degrade misfolded ER proteins. XBP1 also promotes phospholipid synthesis to further expand the ER, raising the ceiling for high-output protein secretion (Hetz, 2012). Multiple mammalian protein-secreting cell types, including salivary and mammary gland cells, fail to fully differentiate in xbp1 mutant animals (Hasegawa et al., 2015; Lee et al., 2005). These defects likely stem from a breakdown in the UPR, causing loss of feedback between the secretory needs of the cell and its ability to expand to full secretory function (Hetz, 2012). In Dro sophila, xbp1 expression similarly marks cell types with high protein secretory load (Ryoo et al., 2013). XBP1 and the UPR may thus constitute a conserved component of the protein-secreting gland cell tool kit of metazoans.

A molecular evolutionary model of gland cell assembly

By what process do novel biosynthetic pathways and secretory systems come to be coordinately assembled within the same cell type – during both development and evolution? We propose a framework for gland cell-type assembly that borrows from the well-supported model of neuronal subtype differentiation. In this model, differentiation of neuronal classes, each with distinct neurotransmitter chemistries, is contingent on different ‘terminal selector’ transcription factors (Flames and Hobert, 2009; Hobert, 2008, 2016). These are master regulatory proteins (Mann and Carroll, 2002; Whyte et al., 2013) expressed during terminal differentiation, which govern expression of batteries of loci that define the functional properties of neurons. The targets of terminal selectors include genes encoding synthesis, synaptic secretion and reception of specific neurotransmitters, as well as ion channels and components of the cytoskeleton and extracellular matrix. Adapting this model to animal exocrine glands, we propose that differentiation of gland cell types is likewise controlled by glandular or biosynthetic terminal selectors (Fig. 3). These transcription factors coordinate expression of both the enzymatic pathways and secretory machinery that confer glandular functionality on naïve cells (Arendt, 2008; Arendt et al., 2016). We assume that additional aspects of the differentiated state of gland cells, such as cellular anatomy and intercellular adhesion, are likewise under control of the same putative terminal selector. The analogous logic behind the differentiation of neurons and gland cells may be more than coincidental: both cell types exhibit pronounced biosynthetic and secretory capabilities, and moreover, cell lineage studies in early-branching animal phyla imply that neurons may be evolutionarily derived from an ancestral secretory gland cell type (Babonis et al., 2018). Neurons may thus represent a specialized and highly derived class of gland cell.

What evidence exists for gland terminal selectors? The transcriptional differentiation of gland cell identity has not, in general, been deeply investigated, and this is especially so for small-molecule-secreting glands. The strongest support for the terminal selector paradigm comes from intensively studied mammalian exocrine glands. A clear example is the differentiation of the exocrine pancreas, where the identity and function of secretory acinar cell types has been shown to depend on the pancreas transcription factor 1-L complex (PTF1-L). PTF1-L is a trimeric protein complex that directly activates numerous targets throughout the genome, including digestive enzymes, secretory protein-encoding genes and unfolded protein response loci (Hoang et al., 2016). PTF1-L both confers acinar cell fate and suppresses induction of alternative cell fates; it subsequently maintains acinar cell biosynthesis and secretory homeostasis. Such a transcriptional regulator fits an idealized notion of a gland terminal selector: a global switch that ‘programs’ biosynthesis and secretion by coordinating expression of batteries of effector loci throughout the lifetime of the gland. A further example is the control of milk production by alveolar cells in the mammary gland, where two transcription factors appear to function in parallel as terminal selectors: Stat5, which controls expression of milk proteins (Liu et al., 1995, 1997) and SREBP, which controls synthesis and secretion of milk lipids (Anderson et al., 2007).

We propose that the widespread, convergent evolution of animal glands depends on analogous terminal selectors in different gland cell types, recruiting novel biosynthetic pathways and secretory components (Fig. 3B,C). The identities of these transcription factors must surely differ between non-homologous glands, but they share the same role of installing the molecular toolkit for glandular...
Box 3. Illuminating gland cell type assembly through single-cell biology

Transcriptome sequencing (RNAseq) of biosynthetically active glandular tissue has been used relatively successfully to identify putative genes involved in compound biosynthesis or glandular function (Bourguignon et al., 2015; Bušek et al., 2016, 2015; Li et al., 2013; Nakaoka et al., 2017; Rork and Renner, 2018; Vogel et al., 2010). However, one caveat with canonical RNAseq is that "bulk" sequencing of whole gland structures provides only a global view of the transcriptome at the organ level, with no resolution of the transcriptional states of different cells within the gland (e.g. Eberwine et al., 2014; Jaitin et al., 2014). This is particularly problematic for multicellular glands, which are frequently composed of distinct gland cell types (Noirot and Quennedey, 1974), each of which may manufacture a different compound. Glandular tissue may also be enmeshed with other cell types (Fig. 4A), including non-secretory cells that could be serving ancillary roles. In these instances, isolating the gland cell types of interest may be too technically challenging to yield sufficient tissue for RNAseq.

The advent of single-cell sequencing technologies circumvents this issue. Similar to bulk RNAseq, single-cell RNAseq (scRNAseq) harnesses next-generation sequencing (NGS) to transcriptionally sample a tissue (Kulkarni et al., 2019; Stuart and Satija, 2019). However, the tissue is first separated into individual cells. There are two general approaches – droplet-based and target-cell-based – both of which are of potential use in the study of complex glands. Droplet-based methods (Fig. 4A; e.g. inDrop, Drop-seq, 10x Genomics Chromium) employ enzymatic dissociation to digest the tissue into a cell suspension. Droplet microfluidics then separates the suspension into individual cells, each inside an oil droplet. Cells are lysed independently, and transcripts are tagged with oligonucleotide cell barcodes and unique molecular markers (UMIs) prior to sequencing (Fig. 4A). This approach can be used to transcriptionally profile large tissue sections (e.g. whole segments of small insects that contain glands of interest), and so can be used to characterize glandular tissue as well as associated non-gland tissue. One current limitation is that droplet-based scRNAseq usually does not recover all transcripts present in a single cell, a relatively small fraction may typically be sequenced. It may therefore not be the method of choice if the gland cells of interest are at low abundance in a given tissue. A further limitation is that microfluidic separation tends to constrain the size distribution of sequenceable cells. Larger cells, such as some big secretory type cells, may exceed the current size limit. Neither of these limitations apply to target-cell-based scRNAseq methods. These methods, which include SMART-seq/SMART-seq2 and C说什么 (capture and unction by tailing and switching), use individual cells that are either manually dissected or isolated via cell sorting. Unlike the droplet-based approach, target cell-based methods are not optimized for ultra-high throughput, but rather individual or small numbers of cells are sequenced to high read depth, with template-switching PCR giving near-complete transcriptome coverage. Target-cell-based techniques are particularly useful for transcriptionally profiling small pieces of dissected glandular tissue or even single gland cells.

Connected to these methods has been the development of multivariate bioinformatics tools for single-cell data analysis (Fig. 4B). Because droplet-based scRNAseq can provide transcriptomes for many non-target cells, it can potentially be used to illuminate the process of gland cell type assembly. Putatively more ancestral cell types may be identified as sources of enzymes or secretory components in the novel gland cell type of interest; alternatively, cell types may be identified that express paralogs of these components, and new enzyme classes that evolved de novo may also be found (Fig. 4B). Statistical methods to explore such relationships between cell types within a tissue, as well as methods to study cell types across species, have recently become areas of intense focus (Carmona et al., 2017; Gehring et al., 2018 preprint; Horie et al., 2018; Konstantinides et al., 2018a,b; Manoni and Arendt, 2017; Pimentel et al., 2017; Tasic, 2018). By performing scRNAseq at different stages of gland development, ‘trajectory inference’ tools such as RNA velocity (La Manno et al., 2018) are available that can be used to trace temporal differentiation programs of gland cell types and pinpoint their parent tissues. Such an approach may help illuminate both the evolutionary origins and transcriptional mechanisms controlling gland cell specification and differentiation. Further insights into transcriptional control may be gained from ATAC-seq (assay for transposase-accessible chromatin; Buenrostro et al., 2013), another NGS method that reveals regions of open chromatin in gland cells, to which putative terminal selectors may bind (Fig. 4C). ATACseq uses a transposase to ligate sequencing adapters into regions of open chromatin on a genome-wide level. Read depth is thus correlated with chromatin ‘openness’ (Buenrostro et al., 2015), highlighting regions of high DNA accessibility where cis-regulatory elements are located (Blythe and Wieschaus, 2016; Davie et al., 2015). ATACseq may be useful find specific elements necessary for recruitment of co-opted or duplicated enzymes, and hence to identify terminal selectors and other relevant transcription factor(s) controlling gland cell assembly (Fig. 4C).

function. For the vast majority of gland cell types, these terminal selectors remain hypothetical and their identities mysterious. In Drosophila, certain developmentally regulated transcription factors are known to be required for the formation of exocrine gland cell types. For example, the Hox protein Abdominal A, the compartment-specifying Engrailed, and the protein Spalt, have been shown to be necessary for oenocyte specification during embryogenesis (Makki et al., 2014). However, it is not clear if these or other downstream transcription factors are the oenocyte terminal selectors, directly responsible for controlling expression of CHC biosynthesis enzymes (e.g. CYP4G1, FACR, Elovl and desaturase), as well as proteins controlling lipophorin synthesis. In contrast, in Drosophila salivary glands, the transcription factor CrebA has been demonstrated as being necessary and sufficient to directly control expression of secretory proteins (Abrams and Andrew, 2005; Fox et al., 2010). Presumably, CrebA functions in parallel to additional terminal selectors that govern production of the saliva itself. In the silk moth (Bombyx mori) distinct terminal selectors have been found that function in complementary domains along the silk gland to directly control silk protein expression. The Hox protein Antennapedia controls expression of sercin as well as other silk protein-encoding targets in the middle silk gland (Tsutoba et al., 2016), whereas the LIM-homeodomain protein Arrowhead directly activates multiple fibroin loci in the posterior silk gland (Kimoto et al., 2015). The terminal selector model may extend to plants: in Petunia flowers, the transcription factor ODORANT1 controls expression of biosynthetic enzymes for producing volatile phenylpropanoids/benzenoids (Van Moerkercke et al., 2012; Verd Nom et al., 2005), as well as an ABC transporter that mediates release of these floral fragrance components into the atmosphere (Balabanidou et al., 2019).

Evident from some of these examples is that single terminal selectors may not necessarily execute gland cell differentiation alone. Even PTF1-L, which may qualify as a terminal selector extraordinary in pancreatic acinar cells, functions in combination with at least two other proteins, the nuclear receptor liver receptor homolog-1 (LRH-1; also known as NRS2A) (Holmstrom et al., 2011), and the basic helix-loop-helix protein MIST1, a transcription factor that enhances secretion both in pancreatic acinar cells and other mammalian exocrine cell types (Lo et al., 2017; Pin et al., 2000, 2001). Molecular mechanisms by which sets of terminal selectors function collectively to confer glandular function are unclear. It has been proposed that, as a general rule, evolution of new cell types is contingent upon transcription factors physically interacting in new ways, forming novel ‘core-regulatory complexes’ that can directly activate loci encoding cell-type functionality (Arendt, 2008; Arendt et al., 2016). Alternatively, different transcription factors may simply be co-expressed in the new cell
type and operate at distinct loci, or bind enhancers within the same locus without forming a protein-protein complex. In the case of exocrine pancreatic cells, PTF1-L is trimeric and functions through co-operative binding between subunits. It may therefore qualify as such a core-regulatory complex for acinar cell function. Nevertheless, to achieve full acinar cell differentiation, PTF1-L must co-regulate enhancers with LRH-1, without physically binding to the latter protein (Hale et al., 2014).

How does a terminal selector’s cassette of target loci evolve? Terminal selectors themselves are developmentally regulated, and function cell autonomously in cells fated to become the gland. By the time they are expressed, the gland cells have likely fully proliferated, but not fully differentiated. Recruitment of the terminal selector’s battery of enzymes, secretory proteins and other functional gene products requires the evolution of cis-regulatory enhancer regions in all such target loci (Fig. 3B,C). In the case of biosynthetic pathways, gene duplication appears to create the set of enzymes that are recruited into the new gland. Limited evidence of single-copy enzyme co-option implies that most enzymes are typically under strong selective constraints, which restricts their capacity to be re-used in novel biosynthetic pathways. This inference fits with the view of Jensen (1976), that duplication is the main mechanism for evolving novel metabolic enzymes, since it enables catalytic specialization. Duplication allows for conservation of biochemical reactions via one copy of the duplicate gene, but also creates an opportunity for neofunctionalization – the emergence of novel functions in the sister copy via selection or drift (Ohno, 1970). We suggest a corollary, that biosynthesis pathways within gland cell types may be more strongly optimized by natural selection for their dedicated biosynthetic function than if they were composed of co-opted enzymes. Glandular biosynthesis pathways are therefore evolutionary modules (Wagner, 1996; Wagner et al., 2007), able to evolve without pleiotropic constraints imposed by functions elsewhere in the animal.

A major conceptual challenge is to explain how such a modular set of duplicated enzyme loci themselves came into being. The fates of duplicated genes have been widely debated by molecular evolutionists, with a number of competing models proposed for how duplicated loci are not simply lost via genetic drift, but instead become visible to natural selection by impacting the phenotype (Berghorsson et al., 2007; Force et al., 1999; Lynch and Conery, 2000; Ohno, 1970). We propose that gland cell evolution is one of the main driving forces behind enzyme duplication and secondary metabolic pathway evolution. Despite the seeming scarcity of co-opted enzymes in glandular biosynthesis pathways, we posit that during the early evolution of a novel gland cell type, initial pathway assembly may in fact occur via co-option of pre-existing enzymes, which only later duplicate (Fig. 3B). We envisage enzyme co-option via two scenarios. First, a transcription factor that ancestrally functions elsewhere in the body may become developmentally recruited into the novel gland to function as a terminal selector. Target enzyme loci expressed in ancestral tissue now become co-opted into the novel gland. Alternatively, enhancer evolution within an ancestral set of enzyme loci permits a terminal selector to co-opt pre-existing enzymes into the gland. Via either of these scenarios, the same enzymes now function in two cell types. In such a situation, pleiotropic conflict arises from selection to optimize the enzymes for new biosynthetic purposes. Duplication henceforth becomes selectively advantageous, yielding paralogues with gland-specific expression (Fig. 3C). The observed trend that we noted above, in which many glandular biosynthesis pathways appear to be composed mainly of duplicated rather than co-opted enzymes, may reflect the completion of this molecular evolutionary process in the majority of animal gland cell types so far studied.

**Gland assembly constraints and evolutionary patterns of chemical usage in animals**

Across metazoan phylogeny, two broad-scale patterns in chemical evolution can be observed. The first is that homologous glands in different taxa are capable of producing taxon-specific compounds. This evolvability implies biosynthetic ‘reprogramming’ within the gland. The second trend is that, despite the great diversity of animal secretions, repetitive patterns of compounds usage occur, signifying constraints on potential biosynthetic diversification. The first of these patterns is straightforward to explain by invoking the proposed terminal selector model of gland cell differentiation. Once a new gland has evolved, the terminal selectors that control expression of effector loci are evolutionarily static, and conserved in lineages inheriting the gland. However, there is evolutionary turnover in the enzymes recruited for expression in the gland. Through cis-regulatory changes in the genome, the biosynthetic cassette expressed in the gland can change, modifying its chemical output. Such turnover in recruitment has been shown for venom proteins in Nasonia wasps (Martinson et al., 2017). The aleocharine rove beetle defense gland (Fig. 2) represents a possible example for small-molecule secretions. The same, homologous gland is conserved across most of the >16,000 members of this subfamily. The common chemical element is benzoquinone, but different species supplement these with diverse short chain alkanes, esters, aldehydes and acids (Steidle and Dettner, 1993). The secretion has been modified most remarkably in ‘myrmecophile’ species that are symbiotically associated with ants and capable of secreting behavior-manipulating compounds, including sulcatone (host ant alarm pheromone) (Stoeffler et al., 2007, 2011) and monoterpenes, which may mimic ant-multiutilitative pheromones (Stoeffler et al., 2013). In these examples, radical biosynthetic reprogramming has seemingly occurred, enabling adaptive changes in ecology. We propose such reprogramming has been achieved by defense gland terminal selectors recruiting novel enzyme pathways (Parker et al., 2018 preprint).

Through differential enzyme recruitment, terminal selectors may diversify glandular secretions. However, there is an apparent counter to this process, evident in pervasive convergence in the types of compounds animals synthesize. Convergence is seen in the widespread use of fatty-acid-derived compounds like hydrocarbons and alcohols for chemical communication (Blomquist and Baghères, 2010; Chung and Carroll, 2015; Finet et al., 2019; Leonhardt et al., 2016; Morgan, 2010; Tegoni et al., 2004) or the employment of terpenes for diverse processes (Beran et al., 2019; Blunt et al., 2014; Breitmaier, 2006; Trapp and Croteau, 2001). A further example is the frequent use of aromatic benzoquinones for chemical defensive; these compounds have evolved repeatedly in harvestman, millipedes, earwigs, crickets, termites, cockroaches, aphids and at least four different families of beetles (Blum, 1981; Eisner et al., 2005; Francke and Dettner, 2005). One explanation for chemical convergence is that potential genetic biases exist at the biosynthetic level. As mentioned above, metazoan biochemical diversity is constrained by limited primary metabolic precursors that can feed into secondary metabolic pathways. However, the proposed molecular evolutionary framework for gland cell-type assembly suggests an additional, important source of constraint. During the initial evolutionary stages of gland cell assembly, our model postulates that terminal selectors tend to co-opt pre-existing enzymes – a process that has led to the repeated involvement of certain enzyme families. Despite the potential for
subsequent enzyme duplication to allow divergence between the ancestral and duplicated enzyme copy, the finite number of gene families from which enzymes may be recruited means that the range of possible reactions that can happen will be inherently restricted. Depending on the compound type, certain catalytic steps are almost invariably carried out by members of one or a handful of substrate-compatible enzyme families.

To illustrate this notion with an example, the observation that benzoquinones are so widely used in insect defense implies that the enzymes required for benzoquinone synthesis might be easily recruited from a more conserved pathway. One obvious candidate is the pathway controlling cuticle tanning, which employs a pro-phenol oxidase, laccase (He et al., 2018), to convert dietary tyrosine into quinone precursors of melanin (Blum, 1981; Duffey, 1974; Pryor, 1940; Roth and Stay, 1958). Although there remains confusion over how benzoquinones are synthesized (Duffey, 1974; Meinwald et al., 1966; Morgan, 2010; Rocha et al., 2013), laccase recruitment is an attractive scenario: one can imagine how a conserved biosynthetic enzyme, expressed in epidermal cells of all insects to oxidize tyrosine into melanin precursors, could be recruited into ectodermally derived glandular cell types to oxidize tyrosine into benzoquinone precursors. It is an accessible means to evolving a highly effective chemical defense, and may thus be a ‘path of genetic least resistance’ that evolution has taken multiple times (Schluter, 1996). We suggest that a pre-adaptive ‘gland toolkit’ is encoded in animal genomes: a parts list of enzyme families, functioning pathways and secretory mechanisms poised for co-option, duplication and modification on each reinvention of the gland. While versatile, this toolkit is limited, and its repeated use has restricted the exploration of chemical space by animals.

Any such molecular constraints that influence animal chemical secretions must resonate at the ecological level. By making certain types of compound more or less likely to evolve, these constraints define the spectrum of evolvability of animal secretions. We believe chemical ecology is poised to move beyond the pure phenotypic identification and analysis of compounds, with emphasis...
increasingly being placed that the molecular architecture and evolution of exocrine glands (see also Rork and Renner, 2018). Advancement in this direction will depend on the uptake of genetic, cell biological and genomic approaches by chemical ecologists. One approach that has huge promise is single-cell profiling, where new technologies have made it possible to quantify molecular attributes of single cells within a tissue or organ. In Box 3 and Fig. 4, we outline how single-cell methods could be used for fine-scale molecular interrogation of glands. The key advance is the ability to analyze individual cell types housed within more complex tissues at the transcriptomic or chromatin levels (Fig. 4A). This enhanced resolution permits biosynthesis pathways, secretory proteins or putative terminal selectors to be identified based on their expression within the cell type of interest (Fig. 4B). Moreover, by quantifying similarities between cell types, historical processes of gland cell type assembly can be inferred: enzymes that have be co-opted or duplicated from more ancient cell types can be identified, along with enhancers that control their novel expression (Fig. 4C). By virtue of its cellular-level resolution, single-cell data also brings the molecular gland ‘toolkit’ more rapidly into view, streamlining the design of functional genetic experiments. Tools such as RNAi-mediated gene silencing and CRISPR/Cas9 gene editing permit fundamental studies of gene manipulation to be carried out in a broad range of animal species (Goldstein and King, 2016). Targeted knock-downs or knockouts of enzymes and secretory proteins can now potentially be performed in many taxa, as well as mutation or excision of gland cell type enhancers. The functional consequences of manipulating or removing these factors can be assayed via mass spectrometric analysis of the gland’s secretory product. The combination of single-cell methods and functional studies is potentially very powerful, and promises to uncover molecular evolutionary steps that have shaped chemical interactions in many metazoan contexts.

Conclusion

Glands are the embodiment of cell type innovation in the Metazoa. These structures represent a natural experiment in cellular engineering, and provide a testing ground for single-cell biology to answer basic questions about the evolution of novelty. Key gaps in knowledge include how modular, gland-specific biosynthesis pathways evolve, and how these pathways and their corresponding secretory systems come to be under transcriptional control. Understanding these phenomena may illuminate molecular constraints that shape the evolution of animal interactions.

Acknowledgements

We are grateful to the organizers of the JEB ‘Genome Editing for Comparative Physiology’ symposium for the opportunity to write this article. We thank David Miller, Marco Smolla, Christoph Kleinheidam, Nick Porch, Adam Rork, Paul Rigby and Thomas Lolitso for making their images available to us. Members of the Parker laboratory provided valuable feedback on this paper.

Competing interests

The authors declare no competing or financial interests.

Funding

A.B. is a Simons fellow of the Life Sciences Research Foundation (LSRF). This work was supported by a Rita Allen Foundation Scholars Award, a Shurt and Kay Curci Foundation grant, a Klingenstein-Simons Fellowship Award and an Army Research Office MURI award W911NF1910269 to J.P.

References


12


Ng, W. C., Chin, J. S., Tan, K. J. and Yew, J. Y. (2015). The fatty acid elongase Bond is essential for Drosophila sex pheromone synthesis and male fertility. Nat. Commun. 6, 8263. doi:10.1038/ncomms9263


Quennedey, A., Drugmand, D. and Deligne, J. (2002). Morphology and ultrastructure of paired prototergal glands in the adult rove beetle 


Reeves, R. A. and commencement in tenebrio molitor. J. Cell Biol. 133, 78-117. doi:10.1083/jcb.133100133


