Review

Renaming the ‘OS-D/CSP’ Family (Part-2): ‘4-Cysteine Soluble Proteins’ (4CSPs)—Intracellular Functions

Guoxia Liu , Botong Sun , Wei Fan , Shousong Yue , Qiuxia He , Jean-François Picimbon *

Posted Date: 30 April 2026

doi: 10.20944/preprints202604.2167.v1

Keywords: 4CSP; 4-Cysteine Soluble Protein; mucin; actin skeleton regulatory complex; nuclear pore complex protein; transcription initiation factor; cell regulation

Review

Renaming the 'OS-D/CSP' Family (Part-2): '4-Cysteine Soluble Proteins' (4CSPs)—Intracellular Functions

Guoxia Liu [1], Botong Sun [2,3], Wei Fan [2], Shousong Yue [1], Qiuxia He [2] and Jean-François Picimbon [1,2,*]

[1] Shandong Academy of Agricultural Sciences, Jinan 250100, China

[2] Science and Technology Service Platform of Shandong Academy of Sciences, Shandong Academy of Sciences Foreign Students Pioneer Park, Qilu University of Technology (Shandong Academy of Sciences), Jinan 250103, China

[3] Shandong Institute for Quality Inspection, Jinan 250100, China

* Correspondence: jfpicimbon@163.com or jfpicimbon@gmail.com or jfpicimbon@sdas.org; Tel: +86-7661098959

Simple Summary

Chemosensory proteins (CSPs) and odorant-binding proteins (OBPs) are two different groups of small soluble proteins that have long been thought to be important components of insect olfaction. They are thought to be responsible for carrying the odor molecules to the sensory receptors. However, when considering the tissue expression and development pattern of CSPs and OBPs, the argument becomes chaotic. We focus on examining a specific CSP, Mp10, which further refutes the "known" olfactory function of the protein family by causing an immune response when transplanted to plants.

Abstract

The gut, brain, fat body, wing, epidermis, *corpora allata*, salivary gland, pheromone gland, prothoracic gland, and many other tissues are not included in the olfactory/chemosensory function. However, they are the tissues where CSPs are most prevalent, except for the antennae and legs. In part-1, we proposed renaming the "chemosensory protein (CSP)" family to "4-Cysteine Soluble Proteins (4CSPs)" to avoid designating a protein present in hemolymph or eggs as chemosensory. In part-2, we broaden the report's focus from ubiquitous tissue distribution to potential intracellular functions to bolster our idea. We go over our studies on insecticide resistance, the Mp10 story in aphids, and other systems relevant to lipid transport and immunity. Most of the data gradually tilts toward the non-chemosensory features of 4CSPs, and this adds even more evidence to support those aspects. We provide a second review (part-2) and analysis that shows a stronger association between 4CSPs and mucins, translation initiation factors, and proteins belonging to the actin complex family. This is yet compelling evidence that the benefits of renaming "chemosensory proteins" far outweigh the drawbacks, when taking tissue distribution and intracellular localization into account.

Keywords: 4CSP; 4-Cysteine Soluble Protein; mucin; actin skeleton regulatory complex; nuclear pore complex protein; transcription initiation factor; cell regulation

1. Introduction

Chemosensory proteins (CSPs) are a class of small, soluble proteins that have long been associated with "Odor-Binding Proteins" because of their unique function in smell. This was hotly contested in Part 1, which detailed their presence in several tissues, at different stages of the insect's life cycle, and in the gut (and fat body) in reaction to insecticide treatment, which strongly contradicted a role in smell [1–3]. The goal of this second review is to emphasize on the need to revise

 doi:10.20944/preprints202604.2167.v1

this established convention, at least for CSP or OS-D ("Olfactory Specific-type D"), as one of the multiple names given to this protein family, and all restricted to olfaction [4–9]. Here, we do not discuss tissue-distribution (known to be ubiquitous), we discuss about potential intracellular functions of the 'CSP' family, which is one more argument to go against the chemosensory role.

In Part 1, we emphasized the structure of CSPs and their cysteine patterns—four conserved cysteines in two distinct disulfide bridges—instead of concentrating on a function that is too nebulous to name the entire protein family. It is a very clear fact that a protein such as "CSP" has a nebulous chemosensory function. A protein that is present in the body's adipose and gut tissues is rarely able to perform a chemosensory function [1–9]. This is by far the same reasoning when we consider that "CSPs" are expressed at extremely early stages of development, when legs and antennae are not yet established, when sensilla are absent, and when the insect bathes in the molting fluid without seeming rigid and/or organized [10,11]. The evolution of insect metamorphosis is revealed by developmental, genomic, and endocrine data from a variety of species [12]. We compare the Holometabola's larva–pupa–adult to hemimetabolous insects' pronymph–nymph–adult. Regardless of how insect metamorphosis has evolved, "CSPs" are still present in all phenotypic and dramatic larval body modifications that result in an entirely new adult morphology [1–12]. "CSPs" are present regardless of the insect's morphology, and the same is true for crustaceans, which go through a variety of significant morphological and ecological changes during their ontogeny [1,12–14]. There is no "CSP" structure that is typical of a chemosensory organ. The protein found in sensory tissues has the same structure as that found in non-sensory tissues: two disulfide bridges joining two closely spaced cysteines, a narrow hydrophobic tunnel functional structure, and six α-helices with a molecular weight of about 10–12 kDa [1,15–18]. It is widely known that the ligand-binding site is pliable and can inflate, expanding the tunnel profile (height and width) to suit different ligand sizes and/or transport several ligands [19,20]. It is not necessary for the ligand to be distinct. The "CSP" molecules that are known to bind long-chain fatty acids (FA), including linoleic acid (LA), which is frequently observed in intracellular multi-cascades, are referred to by this shared structural characteristic [21–23]. It is also present in Mp10, a very intriguing member of this protein family that possesses all the structural features of "CSP" and is characterized not for the response of an olfactory neuron but for the immune system's reaction when injected into plant phloem [24,25]. This eliminates CSP proteins from any function in scent and places them in a position where insects may use them to get past the plant's defenses.

The intracellular function of these transporter proteins is supported by the example of Mp10 in aphids and the binding of LA by whitefly "CSPs" [21–25]. The sequence homology between "CSPs," "CSP-like proteins", and numerous other "bigger" proteins with a variety of fundamental intracellular activities is covered in Part 2. This strongly suggests renaming this huge set of proteins "4CSPs" ("4-Cysteine Soluble Proteins"), a term we will use throughout literature because it is somewhat absurd to refer to an intracellular protein as "chemosensory".

Based on notable similarities with Thap1 and PAN-1, we started by reconsidering the biological role of the 4CSPs [1,26,27]. Here, we advanced the notion that 4CSP exists inside cells based on strong similarities (sequence identity, consensus sequence, and structural features) with multiple intracellular protein families, including mucin, actin skeleton regulatory complex (ASRC), nuclear pore complex protein (NPCP), and transcription initiation factor (TIF). A role in chemosensing is further refuted by the sequence homology between 4CSPs and other intracellular proteins reported here. How does a 4CSP work in scent when it shows a high degree of relatedness with Mucin, ASRC, NPCP, and TIF?

After concentrating on tissue development, distribution, and insecticide, we emphasize the significance of 4CSPs in insect physiology and adaptation in resistance (see [1]). We strongly advocate for the renaming of "CSPs" and A10/OS-D Pfam Domain and provide intriguing new information regarding their possible roles in FA lipid transport, immunological responses, and intracellular activities that extend far beyond chemosensing and smell.

 doi:10.20944/preprints202604.2167.v1

We may be able to provide a new definition of the biological function of this protein family by highlighting 4CSP's unique relationship to lipoids, intracellular events, DNA regulation, and the significance of this protein family's role in regulating a variety of intracellular molecular complexes and metabolic pathways in multiple tissues—not just palps, and the antennae and legs.

2. The Function of 4CSPs in Bacteria, the Earliest Indication of Intracellular Activities

The orthology groupings of 4CSP, PAN-1, and Thap1 have been established. The proteins Bommo-4CSP10, PAN-1, and Thap1 are unrelated to chemosensory function, and it is highly probable that the other proteins associated with them are also unrelated to chemosensing. PAN-1 and Bommo4CSP10 strongly imply that these proteins have an intracellular function, namely in connection to the actin complex and/or the vesicular membrane [1].

A second clue for intracellular function is the existence of 4CSPs in bacteria, which are among the first systems to emerge on Earth and do not have sensillar dendrites [1,14,22,28,29]. After twenty years of evasive hypotheses regarding plant odors and sex pheromones, the discovery of 4CSPs in bacteria is a recent but crucial finding to refute the outdated chemosensory postulate [1,14]. Until the first functional studies in bacteria are conducted, any assertions regarding their roles will be entirely theoretical. However, sensory lymph is absent from prokaryotic cells; its presence in a wide range of bacterial taxa, including actinomycetes, firmicutes, and proteobacteria, strongly implies that 4CSP serves an intracellular purpose [14,22].

Bacterial cells' adaptability and resistance to changes in the environment and/or on any type of mineral substrate depend on their capacity to modify intracellular trafficking and (lipid) molecule distribution. Therefore, the finding of *4CSP* genes in bacteria, such as the soil-derived species *Kitasatospora griseola* (*CP1733360*, *JBHXEV010000005*, and *JXZB01000001* [14]) strengthens the argument against the chemosensory activities of 4CSPs. *Coccobacillus*, *Staphylococcus*, and *Actinobacteria* genera in the families Enterobacteriaceae, Nocardioidaceae, Pseudonocardiaceae (*Solhabitans fulvus*), and Streptomycetaceae have also been reported to contain 4CSPs in addition to *Kitasatospora*, *Bacillus*, or *Klebsiella* [1,14,22]. According to RNA and genomic findings, 4CSPs are also present in Firmicutes, Aeromonadales, Alteromonadales, Eubacteriales, and Hyphomicrobiales (*MDK0835621*, *MDK0841570*; see [14]). These microorganisms are recognized to be common bacteria in the digestive tract, primary prokaryotic secondary metabolites, drug-resistant opportunistic pathogens, and multi-species symbionts in plants and insects, but not for their capacity to bind smells like plant odors.

Excreted 4CSPs may be involved in "sensing" perhaps for endosymbionts or gut microorganisms, but just as a stimulus rather as a carrier of odorants. However, most microbial 4CSPs, like most insect 4CSPs, have a rather short signal peptide (less than fifteen residues), indicating that they cannot be discharged from the cell. This is quite improbable unless they release them outside of the microbial cell by some other mechanism or component. The size of 4CSP signal peptides and the fact that bacteria produce them strongly support intracellular roles [14].

Our preferred theory is that 4CSPs participate in adaptation mechanisms involving the metabolism of substrates and/or nutrients rather than being used as signaling molecules. Although it has not yet been demonstrated through experimentation that microbial 4CSPs induce adaptation, this may be rather straightforward as it would mostly rely on the conventional molecular and biochemical techniques employed to investigate this protein family. Binding assays—which involve conventional recombinant protein expression, purification, fluorescence spectroscopy (spectrofluorimetry), and docking (Linux)—are the most crucial short-term studies that are recommended. However, instead of testing sex pheromones and/or plant odors, the objective would be to test nutrients or sediment substrates, such as silt, mineral particles, nitrogen, micronutrients, or even the remains of dead plants and animals. In addition to smoke, carbon dioxide, and several non-biological substances or inorganic materials like plastic, bacteria eat all ‘living’ molecules [29–33]. It is probably also necessary to test 4CSP's function in lipid catabolism [34,35]. All living cells, including

 doi:10.20944/preprints202604.2167.v1

insect olfactory receptor neurons, use the "catabolism" process, which allows bacteria to use fat as an energy source [36,37]. While many bacteria use glucose as their main energy source, many are also able to break down fatty acid and lipid molecules to use them as a viable and significant energy source [30,38–40]. This shows the metabolic diversity within the microbial kingdom rather than the chemosensory diversity of bacteria. In environments polluted by oil spills, certain marine microbial species can break down hydrocarbons, although many sea bacteria can metabolize aromatic compounds [41,42]. Therefore, binding studies for the functional characterization of 4CSPs may look at a wide variety of possible chemical ligands that fit within the "metabolic paradigm" as opposed to the "odorant paradigm".

The widespread identification of 4CSPs in oxidase-positive, motile, nonmotile, spore-forming bacteria of various sources (soil, rhizosphere, gut, seabed, and/or sediment origin) does not support a chemosensory role for this protein family [14]. Quoting 4CSP in quorum sensing right away because they are "chemosensory" is an extension of the Spin [1]. Most bacteria employ quorum sensing (QS); nevertheless, the chemical signals, the signal relay mechanisms, and the target genes regulated by bacterial QS systems strongly vary [43–46]. More specifically, there are notable differences between the QS systems of Gram+ and Gram- bacteria. Gram-positive bacteria use peptides as auto-inducers (AIs). Small compounds like N-acyl homoserine lactone and S-adenosyl methionine are used as AIs by Gram-negative bacteria [47,48]. These AIs can easily permeate through cell membranes and bind to cytoplasmic TIFs and/or histidine kinase receptors [46–48]. Odor receptor activation is a long way off since the evolution of *4CSP* gene does not track the evolution of QS: 4CSP developed into new activities after the separation of Gram− and Gram+, but it was present before the split of Gram+ and Gram- cells [14,45]. 4CSP shows a broader role in cell metabolism since it comes before QS.

Interestingly, the *Sorangiinae* protein is 482 amino acids long and has a weight of about 50.7 kDa (WXB18440), while the *Actinoallomurus* and *Shewanella* proteins are 89-129 amino acids (9.9-14.9 kDa with four cysteine patterns; see WP_378213916 and MCH1932606). The 4CSP is in the C-terminus of this massive protein [14]. In addition to 4CSPs for intracellular FA lipid transport, the presence of 4CSP as a structure in larger molecular complexes may provide a new avenue for investigation into whether this is specific to bacteria or if the 4CSP protein domain plays a general role in larger complexes in both prokaryotic and eukaryotic cells. We can hypothesize that 4CSPs, which are present in a wide range of Gram+ and Gram- bacteria, engage with AI in the cytosol, TIF, or large transmembrane enzyme receptor protein complexes to help degrade the substrate and/or bacterial QS. It is necessary to investigate the scenario in which a 4CSP is found in the cytoplasm and/or linked to membrane-bound supramolecular complexes in bacteria. This is in sharp contrast to insects, where it is far too widely accepted that 4CSPs do not work intracellularly at all and instead bind to plant odors at the edge of ORs.

3. Functional Evidence Supporting the Immune System's Response, JH, Ecdysone, and the Intracellular Roles of 4CSPs in Insects

Polyclonal antibodies against the 4CSP protein were used in immunocytochemistry experiments to label the antennal sensillum in insects; however, the labeling also extended to the cuticle, supporting cells, sensory structures, and the interior of the cell [6,49,50]. The external sensillum fluid of the hair lumen is heavily labeled, but intracellular label is also seen in the Golgi apparatus, endoplasmic reticulum (ER), and distinct dense granules in a variety of cell types [49,50]. This scenario is quite similar to that of OBPs. The lysosome, multivesicular bodies, endocytotic pits, and vesicles are all distinctly marked by the OBP antibodies, which are likewise labeled inside antennal neuronal cells [50].

3.1. Intracellular Response to Exposure to Xenobiotics

Every *4CSP* gene is expressed outside of the olfactory paradigm, mostly in the gut and fat body tissues, and it affects the insect's response to avermectins, according to Xuan et al. [2]. Furthermore,

 doi:10.20944/preprints202604.2167.v1

4CSPs are sensitive to bacterial infections in Culpi and viral infections in Aedae caused by *Plasmodium gallinaceum*, *Brugua malayi*, and Dengue viruses, according to EST data [51–54]. Therefore, 4CSP participates in a very broad innate immunity response rather than just reacting to insecticides. We wish to immediately state that 4CSPs are crucial for the identification and removal of xenobiotic chemicals when we find an insecticide effect ([see 2]). However, when the same result (up-regulation) is obtained for viral, fungal, or bacterial infections, the detecting function is lost. It needs a somewhat more general function. There are just two processes left: the immune system's elimination process and intracellular functions like lipid transport to stimulate the cell. Chemosensory perception is completely absent from the insect's response to a bacterial or a viral infection.

Insects react similarly to bacteria, insecticides, or viruses (they overexpress 4CSPs) [21,22,55–57]. Fungal infection alters the expression of 4CSP (and OBP) in ants and beetles [58,59]. Like whiteflies, 4CSPs seem to be associated with thiametoxam tolerance in the cotton aphid *Aphis gossypii* Glover and the avian cherry-oat aphid *Rhopalosiphum padi* [21,22,60,61]. 4CSP-transgenic flies seemed to be more resistant to thiametoxam, α-cypermethrin, and spirotetramat than "normal" flies [60]. *A. gossypii* and *R. padi* were significantly more sensitive to thiametoxam when 4CSP transcription was suppressed by RNA interference (RNAi) [60,61]. These aphids' resistance to the diamide pesticide cyantraniliprole was linked to increased expression of 4CSPs in several body tissues, particularly the midgut and fat body [62]. Strong *4CSP* gene responses to toxins, viruses, pesticides, and microbes are more likely to indicate resistance than chemosensing. Olfactory and/or gustatory pathways can be used to detect poisons and insecticides [63]. However, it is rather highly unlikely that 4CSPs are involved in the chemosensing of bacterial or viral products. The immune system, not ORs, GRs, or IRs, recognizes the peptides made by bacteria and viruses [64]. It is believed that two 4CSP proteins localized in the cuticle of the fall armyworm *Spodoptera frugiperda* give multi-insecticide resistance [65]. Similarly, by sticking to the cuticular membrane, two 4CSPs that are prevalent in the tegument are believed to assist the migratory locust *Locusta migratoria* withstand deltamethrin [66]. This is a somewhat narrow line of reasoning because, in locusts and moths, 4CSPs are detected in the gut, hemolymph, and fat body under normal conditions as well as in reaction to food changes or illnesses [1–3,52,67,68]. This is true regardless of the pesticide. The pesticide traveling from the intestines to the adipose body and hemolymph for metabolism would be the most plausible explanation for this [69]. Moreover, methylated deoxysugar I-oleandrose is the disaccharide component of the pentacyclic lactone insecticide avermectin ($C_{48}H_{72}O_{14}$). Because the 4CSP structure is a lengthy hydrophobic tunnel that works well for long-chain fatty acids and/or short cyclic compounds like phenyls, it is more likely that 4CSP binds certain residual parts of the avermectin rather than the full molecule. It appears highly unlikely that the 4CSP protein will encapsulate the whole macrocyclic lactone based on the protein structure and binding data [15,21]. To remove the lactone debris, binding to tiny compounds like benzene and carbohydrates may draw not only 4CSPs but also other binding protein families like the OBPs [3,70]. This links the two protein families near internal degradative enzymes rather than near ORs at the neuronal surface.

Another crucial issue to bring 4CSP intracellularly is the various cellular targets of insecticide chemicals and the various mechanisms involved in insecticide resistance (see Figure 1). If the insecticide molecule gets inside the cell, it will damage the nucleus and other cellular organelles [71]. Interestingly, an insecticide like avermectin will activate protein kinase C (PKC) via phosphorylation, Ca++, and/or FA, as shown in *Drosophila*. This will boost resistance to the poison molecule by mediating the overexpression of intercellular junction proteins, which promote adhesion and intercellular communication [72]. Avermectin activates PKC and intercellular junction proteins, as well as 4CSPs and OBPs in the gut and fat body, according to molecular data gathered in the silkworm [2,3,73]. The 4CSPs (and OBPs) are therefore probably linked to the metabolites of the insecticide or to certain lipoid molecules and cofactors involved in the intracellular mechanisms that initiate the immune response. According to Liu et al. [21], the relationship between thiametoxam-upregulated 4CSPs and linoleic acid (LA, C18:2) is a clear indication that 4CSPs participate in the insecticide response through the lipoid pathway rather than by coming into direct contact with the insecticide

 doi:10.20944/preprints202604.2167.v1

molecule [21,22]. In addition to all 4CSPs, avecmectin also activates PBP1, PBP2, GOBP1, and GOBP2 in multiple tissues [2,3]. Which of these 4CSPs and OBPs will attach to the entire avermectin molecule so that it can be broken down? They are all adjusted to short hydrophobic chains, like those found in vitamins or fatty acids like LA [3,15,21]. This contrasts sharply with the insecticide resistance theory proposed by Tsouri and Douris [74]. Thiametoxam works by first engaging with the gut, which oversees digestion and food absorption, rather than the antennae, just like other insecticides. Through endocytic vesicles, transmembrane protein transport, or diffusion across the plasma membrane of epithelial cells, thiametoxam molecules can enter the intestinal epithelium [75]. Until it reaches the intracellular level, no 4CSP is required (Figure 1).

A new research-study "spin" in the field may start with the assumption that 4CSPs adhere directly to insecticide molecules based only on a few preliminary data (one protein, one ligand examined) [1,76–83]. How can *4CSP* expression in the abdomen and thorax be explained in the absence of toxins [1–3]? Moreover, are the structural and evolutio[1–3nary limitations found in the binding sites of insecticide-degrading enzymes (carboxyl esterases) in flies like any molecular adaptations seen in the 4CSP family [1,84]? Since there are only four fly *4CSPs*, flies would have avoided insecticide resistance if the 4CSPs are mechanisms connected to insecticide binding, which is not the case [84–86]. Furthermore, what effect might insecticides have on bacteria that have 4CSPs in common with insects [see 14]. What does the pentapeptide proctolin's impact on the expression of 4CSP in the crab heart mean [14]? In the quasi-cyclic structure of proctolin (neuropeptide, $C_{30}H_{48}N_8O_8$), tyrosil side chains point outward. In this shape, it will attach to the G protein-coupled receptor's active region most efficiently (β-arrestin2 interactions) [87]. Can we also apply this ligand conformation to 4CSP? In aquatic arthropods such as water fleas, *Daphnia magna*, and cladocerans, salt molecules all increased to 4CSP expression [14]. Does this suggest that 4CSP and salt bind? Additionally, it increased in both acute and chronic temperature conditions, indicating a very broad role for 4CSP in the physiology of arthropods [see 1 & 14]. All these changes in 4CSPs in response to environmental changes clearly imply an intracellular function (Figure 1).

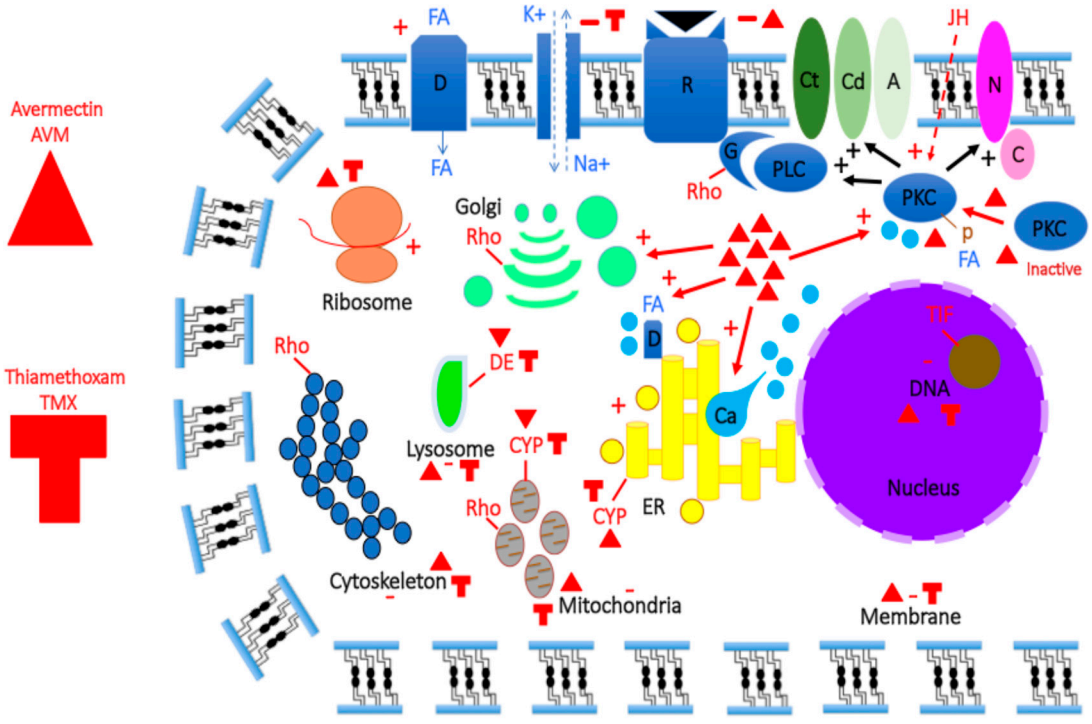


Figure 1. Insect and hexapod cell targets for Avermectin and Thiametoxam insecticides. Insecticidal compounds (Avermectin: Red triangle; Thiametoxam: Red letter T) enter or pass through the cell membrane via active (transporter-based, endocytosis) or passive (diffusion) pathways. They will seriously harm several organelles inside the cells, including the actin skeleton (expression of Rho), mitochondria (respiratory chain), lysosome

 doi:10.20944/preprints202604.2167.v1

(enzyme inhibition), plasma membrane (fluidity), and nucleus (impaired DNA repair, decreased DNA methylation, point mutations, and transcription deregulation). They will lead to increased lipid uptake, FA accumulation, phosphorylation (p), ribosomal protein expression, ribosome biogenesis, and enlarged vacuoles in the Golgi apparatus. In the ER system, they will cause a prolonged release of calcium from internal stores, which raise the cytoplasmic Ca++ concentration. This significant increase in intracellular concentration of Ca++ ions and lipids is a crucial step in inducing the activation of protein kinase C (PKC), which in turn activates the intercellular junction proteins (cell adhesion molecules; A: Armadillo, C: Coracle, Cd: Cadherin, Ct: Catenin, N: Neurexin) that regulate the insect's resistance to insecticide penetration. Juvenile Hormone (JH) is involved, and it activates PKC pathways. +: Insecticide-induced reactions that lead to activation of PKC, A, C, Cd, Ct, N, cytochrome P450 (CYP), Degradative Enzyme (DE), and the many enzymes involved in the metabolism of lipids, carbohydrates, and proteins. -: Toxic effects. Avermectin (AVM) inhibits muscle contractility and neuronal activity by targeting receptors, ligand-gated ion channels, and glutamate-gated chloride – channels (see - Red triangle). By causing an imbalance between sodium and potassium, neonicotinoid thiametoxam (TMX) insecticide damages neurons and disrupts normal nerve impulse transmission (see - Red letter T). D: Desaturase enzyme, K+ intake/Na+ outflow: potassium-sodium exchange pump, R: Receptor, G: G-protein, PLC: phospholipase kinase C, PKC: protein kinase C, Rho: Rho GTPase, ER: endoplasmic reticulum, DE: Degradative Enzyme, TIF: transcription initiation factor.

3.2. Intracellular Response to Exposure to Hormones

Another reason that leads us to consider 4CSP intracellularly rather than at the periphery of ORs is the presence of 4CSP in the *corpora allata* (CA) and prothoracic glands, which are the endocrine sources of juvenile hormone (JH) and ecdysone, respectively [1,14,88]. The development of insects depends on both of these hormones, but more intriguingly, they also play a primordial role in the activation of many intracellular pathways that result in insecticide resistance and adaptation [89,90]. Furthermore, JH and ecdysone are necessary for regulating lipid metabolism and initiating the PKC pathway, both of which are vital for the immune response [91,92]. JH activates the phospholipase C (PLC) pathway, which involves PKC, when insecticide exposure takes place [93,94] (Figure 1). Thus, the insecticide's effect on 4CSPs, PLC, PKC, and JH biosynthesis strongly suggests that 4CSPs, different hormonal factors and lipids, as well as many different intracellular components like PKC and PLC, largely cooperate to produce insecticide resistance, independent of 4CSP binding the insecticide [93–95] (see Figure 1).

The presence of 4CSPs in CA is inconsistent with an olfactory function. This rather implies that 4CSPs support JH pathways. The chemosensory protein idea is shut out from the olfactory paradigm with CA and JH. JH is produced by triggering over 13 enzymes [96]. Furthermore, hemolymphatic transport of JH and juvenoid hormones is a function of 4CSP that may be somewhat similar to that of the OBP, according to Kim et al.'s study [97]. In light of this work, would it be more appropriate to refer to "CSP/OS-D" 4CSP and "OBP" 6CSP instead of "chemosensory/olfactory proteins"? For 4CSPs that deal with JH, PKC, and PLC rather than OR, the renaming is evident (Figure 1).

The association between 4CSP and JH is consistent with the presence of 4CSPs in bacteria and arthropods. Both phurealipids from bacteria and juvenoids have similar chemical structures and impede insect immunity and development [98]. Therefore, 4CSP proteins may have a role in both phurealipids in bacteria like *Photorhabdus bacilli* and JH in insects and hexapods. This would be a very plausible hypothesis until it could be demonstrated that 4CSPs are a component of the bacilli olfactory hedonics.

4. Functional Evidence Supporting the Lipid Transport (C18:2) and Intracellular Roles of 4CSPs in Insects

Evidence for their involvement in intracellular lipoid transport is provided by the presence of 4CSPs in bacteria and copepod lipid sacs, as well as their functions in insect brain development, molt, immunity, JH, sex pheromone generation, and/or behavioral changes in a variety of species [1,14].

 doi:10.20944/preprints202604.2167.v1

According to Ozaki et al. [99], 4CSP proteins mediate the recognition of chemical signatures made up of cuticular lipid FA hydrocarbons (CHCs), such as those seen in ants. However, as is the case with flies and other hymenopteran species like wasps and bees, there are just not enough *4CSPs* to control the multitude of CHCs [22,85]. This situation must therefore be unique to ants. Whitefly studies indicate that 4CSPs have a metabolic function in the degradation of pesticides in connection with C18 FA lipids [21]. This finding is in line with the observation that 4CSP is found in a wide range of bacterial species and insect body cells and tissues [see 1]. All of these factors—wide expression, reproductive organs, fat body, CA, JH, and LA—bring 4CSP into the cell, away from OR.

Beyond merely supplying cell energy, LA takes part in all stages of innate immune cell activity through a variety of substrates and metabolic pathways. Many intracellular processes, such as the omega-6 FA lipid pathway, phosphorylation cascades, hormones, degradative enzymes, mitochondrial CYP450 systems, and many more, depend on LA and FA. It is well known that C18-FA lipids can affect the regulatory and effector functions of innate responses at various tissue levels, work through enzymes and receptors, and improve the fluidity and physical characteristics of cell membranes [100]. For hemocytes, for instance, FA plays a critical role. Hemocytes must use the energy from lipid metabolism to execute encapsulation, nodulation, and phagocytosis at the site of a chemical, microbial, or viral infection. Therefore, in response to infection, FAs quickly reach the cytosol instead of staying in the Golgi [101,102]. This intracellular distribution of the FA-4CSP complex is largely compatible with the presence of 4CSP in the gut, fat body, hemolymph, and hemocytes [see 1].

A general regulating role in insect physiology makes far more sense than one involving OR, given the 4CSP's widespread distribution in insect and hexapod bodies. Like most 4CSPs, Bemta-4CSP1 is broadly expressed and bound to LA [1,2,14,21,22]. LA has never been found in insect chemical traces or "fingerprints" (in a mean of sex pheromone and/or CHC cuticular signature). Since these C18 long-chain FAs are hardly volatile, it is crucial to show how they function as contact pheromones before asserting that LA has a major semiochemical function. By adding a second double bond to oleic acid (C18:1), Δ12-desaturases accelerate the biosynthesis of LA; however, only a few insect species contain this enzyme or FA conversion. Many insect species and essentially none of the crustaceans can produce LA on their own [103]. Would a species develop to use LA as a specific pheromone if it were unable to produce it? Certain insect species produce their sex pheromones by the metabolic synthesis of sequestered compounds derived from plants [104]. Given that "chemical sequestration" would impair the species' capacity to reproduce, this may only be important in certain situations where relationships between plants and insects are closely evolving. The Erebidae and Arctiidae, two of the few Lepidoptera known to sequester plant chemicals, employ these compounds as adults to defend themselves against predator species [104]. Pyrrolizidine alkaloids, not lengthy chains of FAs, are the subject of sequestered plant compounds. Additionally, rather than the fatty acid itself, the methyl ester component of the FA is frequently used as a pheromone [105,106]. These five points raise grave doubts about an FA like LA's capacity to activate an OR.

Food is the sole way that LA may be absorbed, as both LA and food are absorbed in the gastrointestinal tract. Even though some insect species may use LA as a pheromone precursor [107], ORs are never in contact with these precursors. The finding that LA can be a "necromone" indicates a possible area of study for chemosensory receptors and LA-binding 4CSPs involved in dead body identification signals, even though this cannot be applied generally. It is known that only termites, wasps, ants, bees, and social aphids produce LA as a necromone [108]. Therefore, only a few species—mostly social insects—show this behavior of clearing the dead from the nest site or locating dead prey to consume [109]. Aleyrodids do not, which makes a compelling case for LA's intracellular function as a fuel rather than an OR stimulation.

In summary, a number of observations support an intracellular role for 4CSPs in particular: (1) 4CSPs are widely expressed in the insect body; (2) they are expressed in bacteria and arthropods and appear to regulate a variety of biological systems; (3) they bind to LA, a crucial intracellular

 doi:10.20944/preprints202604.2167.v1

component; and (4) 4CSP in particular can activate particular innate immune pathways when injected into plant phloem (see below).

5. *Myzus persicae* Mp10: Evidence of 4CSP's Intracellular Activity

It is not a sensory lymph, phloem. Phloem is the vascular tissue that carries and distributes soluble organic substances to the various sections of the plants. Many hormones and other signaling molecules pass through it. Instead of activating OR, 4CSPs, such as Myzpe Mp10, can initiate an immunological response through the phloem tissue [24,25]. Our next objective is to provide further information regarding the presence of 4CSP proteins in the salivary secretions of the aphid species *M. persicae*, which act as host plant effectors to the aphids' benefit.

First, we included *A. pisum*, *B. tabaci (*Bemta*)*, *Halyomorpha halys* (Brown Marmorated Stink Bug), and *Pachypsylla venusta* (Pacve, Hackberry Petiole Gall Psyllid) in our analysis of the Mp10-4CSP family. For these species, transcript annotation and genome assembly are accessible (see Table S1). Here, we compare Mp10 to the entire repertoire of 4CSP proteins from three distinct hemipteran species, which are also referred to as piercing-sucking insects or phloem-suckers because they consume prokaryotes that feed on sugar and reside in phloem. Next, we performed a blastp search in the "All Species's Protein" Database from InsectBase [110] using the Mp10 amino acid sequence as a template. Numerous “Mp10” hits (length 101-145, identity 47-97%, e-value 1.80e-34–6.35e-100, score 126-287) were recovered from *A. pisum*, *Nilaparvata lugens* (Nillu), the brown planthopper, and *Mayetiola destructor*, the Hessian fly. Additional Mp10-hits with varied length, identity, e-value, and score values were recovered by selectively blasting Mp10 in “All Coleoptera Protein”, “All Diptera Protein”, “All Hemiptera Protein”, “All Hymenoptera Protein”, and “All Lepidoptera Protein”, in species such as *A. pisum* (Acypi), *Anopheles gambiae* (Anoga), *Bombus terrestris* (Bomte), *Camponotus floridanus*, *Chilo suppressalis* (Chisu), *Culex pipens* (Culpi), *Danaus plexippus* (Danpl), *Dendroctonus ponderosae* (Denpo), *Diaphorina citri*, *Harpegnathos saltator, Heliconius melpomene* (Helme), *Linepithema humile, Manduca sexta,* and *Tribolium castaneum* (Trica). There were no hits when searching the transcriptomes of the All Phasmatodea, All Thysanotera, and All Odonata proteins for the Mp10 sequence. A tblastx search produced the same result (no hit) for numerous other insect transcriptomes and RNAs, including *Archaeopsylla*, *Blattella*, *Catajapyx*, *Chrysopa*, *Ephemera*, *Forticula*, *Limnephilus*, *Locusta*, *Nemurella*, *Mengenilla*, *Pediculus*, *Sialis*, and *Zootermopsis.*

We conducted a phylogenetic analysis (PAUP*10Altivec, [111]) using this group of genes to expand the analysis to the Order (Hemiptera) and obtain a more comprehensive knowledge of the evolutionary "appearance" of this Mp10 4CSP-like family. To achieve this, we compared these Mp10s to several insect species and other aphids (Genome Assembly + RNA annotation; Table S1 and Figure 2). Unweighted Pair Group Method with Arithmetic mean (UPGMA) enables us to compare Mp10s and see how closely related various species are to one another. The UPGMA tree showed a particular grouping of Mp10 with 4CSPs from Trica (AAJJ0269C and AAJJ0269B, Col), Denpo (DPO006842, Col), Bomte (XP_012166268, Hym), Helme (HMEL010990, Lep), Chisu (CSUOGS107535, Lep), Bemta (Bta06193, Hem), Pacve (PVENscaf20457, Hem), Acypi (ACYPI000097, Hem), and Nillu (BPHOGS10008228, and BPHOG10002786, Hem; see Figure 2A). The Mp10 grouping is highly supported by high consensus in the protein sequence alignment, mainly in the N-terminus and central 88-PDAL-91 motif, as well as conserved amino acid residues like Q, W, L, K-D, and the four cysteines characteristic of 4CSPs (see Figure S1). We highlight the Mp10-4CSP consensus sequence, important conserved residues in the Mp10s, and the presence of Mp10 "hits" in insect species such as nectar pollinators and plant suckers (Figure S1). This specific grouping was also validated by the Bootstrap-Jackknife tree (Figure 2B). 97% bootstrap supported the grouping of Mp10 with ACYPI000097 (Hem), AAJJ0269C (Col), AAJJ0269B (Col), DPO006842 (Col), XP_012166268 (Hym), HMEL010990 (Lep), CSUOGS107535 (Lep), Bta06193 (Hem), and PVENscaf20457 (Hem). This grouping strongly suggests that Mp10 is present in Hemipterans as well as in pollinator species (Figure 2B). This is rather surprising because we would only imagine that Mp10 helps herbivorous insects overcome plant immunity (see [24,25]). Given this, it is noteworthy that Mp10 is more closely

 doi:10.20944/preprints202604.2167.v1

related to 4CSPs from Trica, Denpo, Bomte, Helme, and Chisu, than it is to Nillu. Mp10 is linked to insects that either burrow or tunnel into plants or eat plant phloem, as evidenced by its presence in aphids, psyllids, whiteflies, and stem borers. The fact that it is also connected to the family of insect pollinators like butterflies and bumblebees, is even more surprising (see Figures 2A, 2B & S1). It is a relatively new notion to examine 4CSPs' role as flower effectors for the benefit of pollinators. Even though there was enough evidence on Mp10 in aphids to demonstrate that 4CSPs could help the insect control the flower's immune system, we were so preoccupied with "olfactory activities" that we totally overlooked the possibility that the Mp10 mechanism could apply to insects other than aphids and be significant for intracellular events [24,25]. Flowering plants must reconcile the conflicting selection pressures of drawing pollinators and repulsing herbivores. However, a pollinator insect may eventually become "invasive" if it frequently visits the same flower, damaging petals or reducing the nectar supply. Herbivory and/or frequent visits may have a detrimental effect on fruit or seed set if pollinators damage flowers while visiting them or if an excess of pollen deposition causes growing pollen tubes to stagnate [112,113]. After the first flower visit, pollinator health, feeding habits, and reproductive success may be impacted by florivory and herbivory.

Pollinators are not necessarily nice, good, or helpful; their brutal, harsh, and ruthless visits to the flowers frequently cause the fruits' form to deteriorate. This is explained in the pollinator behavioral ecology, providing evidence that bees and butterflies have Mp10s to circumvent the flower's defenses. From Hemiptera to Lepidoptera, Hymenoptera, and Coleoptera (see Figures 2 and S1), our evolutionary study shows that Mp10 is a rather old mechanism that still operates in a variety of insect taxa. Numerous species have been shown to express Mp10 (see Figures 2 and S1), but further research is needed to determine if Mp10 is present in saliva, the mouthpart, or somewhere else, as well as how much its function has changed over time.

However, as shown in PAUP (Figures 2 and S1), the number of Mp10s varied by species (two in Trica). Depending on the host plant or diet that the insect regularly feeds on and/or reproduces on, the quantity of Mp10 proteins may vary. Perhaps more Mp10s are required when the insect needs to feed on many host plants. The evolutionary difference between the Mp10 of Nillu and the Mp10s of other insect species may be explained by the fact that Nillu feeds on a single, extremely particular diet host plant—rice (see Figures 2 and S1). MyzpeMp10 and ACYPI000097 are nearly identical (Figures 2 and S1). All Mp10s, except for AAJJ0269C, are proteins with 147–153 amino acids. The consensus sequence shows a very high degree of conservation (up to 97% identical), in particular in the N-terminus (Figure S1). The sequence differences between Mp10s (~35–97% identity) are like the evolutionary distances between the different species, suggesting that the Mp10 function has been maintained throughout evolution. For instance, the ability to use Mp10 to target the host-plant immune system is probably still present in all the sap-sucking insects in the Aphididae family. The split of Acypi and Myzpe approximately 22 Mya indicates how long Mp10 did not move an iota. Since sequence conservation cannot attest to the conservation of the function in the absence of experimental evidence, we propose studying Mp10-4CSP in pollinators to understand how they get past the floral plant defense. In the context of environmental preservation, this might help preserve both the flower and the pollinator. The emergence of angiosperms, pollinivory, and the earliest plant-insect interactions may have coincided with the emergence of Mp10 mechanisms in the early Permian era (~298 Mya) [114]. This ancient Mp10 pathway has more to do with an insect element that affects the plant's or flower's defenses intracellularly than it does with smell.

 doi:10.20944/preprints202604.2167.v1

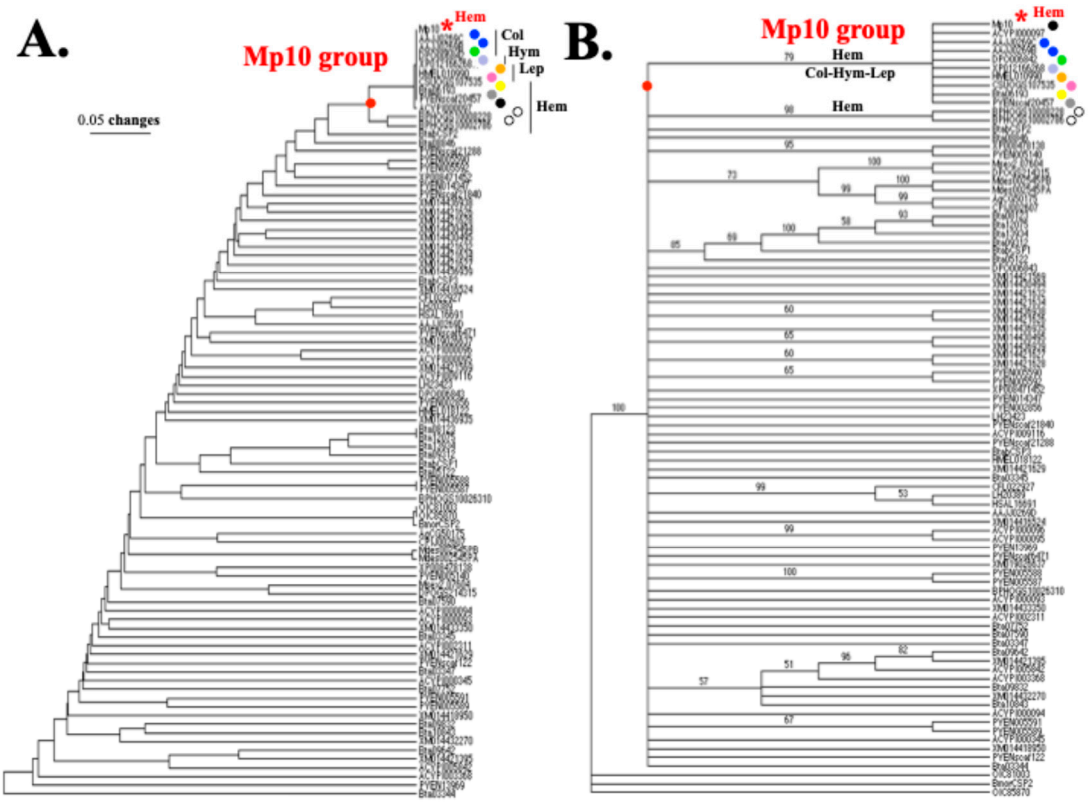


Figure 2. Molecular phylogenetic comparison (PAUP*10Altivec) of Mp10 with related proteins from.various insect transcriptomes and genomes from species in the Order Hemiptera (see Table S1). (**A**) UPGMA analysis: agglomerative (bottom-up) hierarchical clustering based on the distance matrix of the analysed taxa that were calculated from a multiple alignment in ClustalX. Evolution (i.e., mutation rate) across all sequences is constant. Each pair-wise distance contributes equally. The red asterisk indicates Mp10's position (*): Mp10 grouping with AAJJ0269C, AAJJ0269B, DPO006842, XP_012166268, HMEL010990, CSUOGS107535, Bta06193, PVENscaf20457, ACYPI000097, BPHOGS10008228, and BPHOG10002786 at the tree's top. (**B**) Bootstrap/Jackknife algorithm analysis with bacterial *A. baummanii* CSPs (OIC81003 and OIC85870) as outgroup. Amino acid tree (data matrix: total characters 457, constant characters 48, variable parsimony-uninformative characters 144, parsimony-informative characters 265, all characters of type unord, all characters have equal weight): Length 4832, CI 0.423, RI 0.318, RC 0.134, HI 0.577, G-fit –147.563). In red: Mp10, blue: AAJJ0269C and AAJJ0269B, green: DPO006842, purple: XP_012166268, orange: HMEL010990, pink: CSUOGS107535, yellow: Bta06193, grey: PVENscaf20457, black: ACYPI1000097, white: PHOGS10008228, and BPHOG10002786. AAJJ: Trica, Bta: Bemta, DPO: Denpo, XP_012166268: Bomte, HMEL: Helme, CSUOGS: Chisu, PVEN: Pacve, ACYPI: Acypi, BPHOGS: Nillu. For comparative molecular analysis in the Mp10 family, the protein amino acid sequences are used. InsectBase: All Coleoptera Protein (length 96–132, identity 38–50%, e-value 5.33e-31–8.50e- 34, score 108–118), All Diptera Protein (length 99–101, identity 47–49%, e-value 2.34e-34–4.82e-35, score 124–127), All Hemiptera Protein (length 94–145, identity 44–97%, e-value 3.54e-31–4.45e-101, score 109–287), All Hymenoptera Protein (length 97–121, identity 44–49%, e-value 5.19e-31–3.38e- 32, score 111–113), All Lepidoptera Protein (length 109–134, identity 39–43%, e-value 2.08e-26–4.93e- 34, score 98–121). Col: Coleoptera, Hem: Hemiptera, Hym: Hymenoptera, Lep: Lepidoptera.

6. Mp10 and Intracellular Functions from the Nucleus to the Actin

Mp10 is strikingly similar to 4CSPs and numerous other protein families, including neural Wiskott-Aldrich Syndrome (WAS/WASL)-like proteins, actin skeleton regulatory protein (ASRP), mucin, extensin, PAN, Rho GTPase (Rho) activator, serine/threonine-protein kinase C (SamkC), transcription initiation factor (TIF), nuclear pore complex protein (NPCP), Sec31, UL36, and many

 doi:10.20944/preprints202604.2167.v1

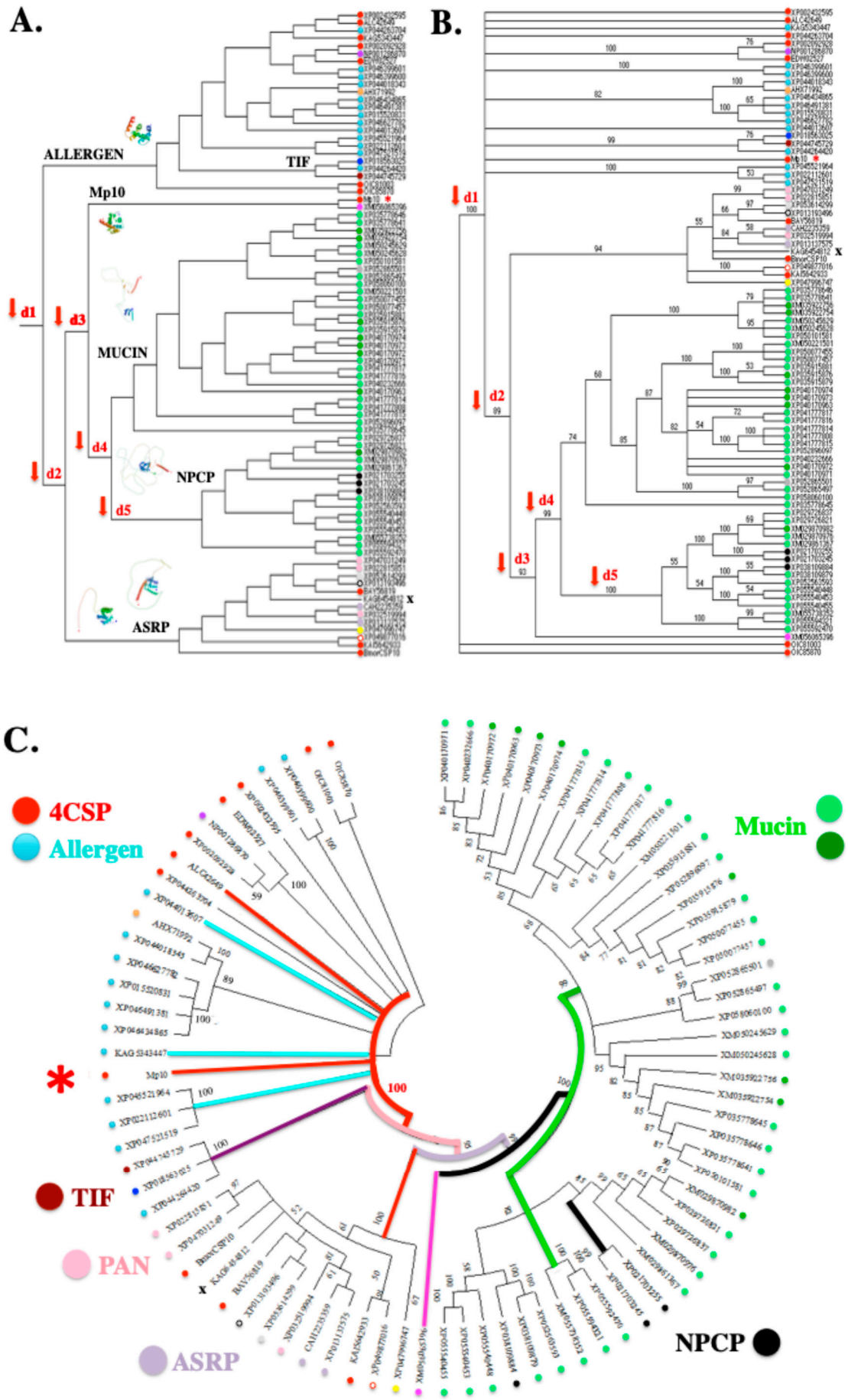
others (30.71-45.13% identity; Figures 5 and S2-S3 and Table S2). This listing does not include OR, ODE, or carboxyl esterase; it only includes intracellular proteins from the nucleus to the actin.

A number of significant viral inner tegument proteins, deneddylases, the virion's large tegument proteins, stress response initiators, cell wall/envelope proteins ("CWPs"), splicing regulators, actin related protein 2/3 complex activators, actin filament binding proteins, and cross linkers for the virus tegument share homology with 4CSPs-Mp10 (see Figures 5 and S2-S3 and Table S2). The comparison of Mp10 with all these intracellular protein families raises some very interesting questions about the possible functions of 4CSPs in the cytoplasmic membrane and cell internal surface, as well as in the cytoplasm and nucleus of the cell, interactions with various molecular and genetic components, RNA/DNA binding, RNA/DNA regulation, transcription control, splicing regulation, and activation of particular protein complexes, even though homologous function is not always implied by sequence homology.

The role of 4CSP in the actin system and nucleus must be demonstrated experimentally. There is currently no proof from "functional" investigations that 4CSP and the protein families linked to actin, mucin, or nuclear pores share any functional involvement other from significant sequence homology (Figures 5 and S2-S3 and Table S2). However, it is evident from our phylogenetic evolutionary analysis and an examination of the Mp10 sequence comparison with orthologs in the NCBI database that intracellular protein families cluster with 4CSPs (see below). Functional evidence was further strengthened and roles in lipid transport, receptor/enzyme phosphorylation, and immunological response were demonstrated by defining RNAi knockdowns that resulted in lower fecundity in whiteflies, dose-response to insecticides, and LA binding data [21,22,115]. These findings all strongly point to intracellular processes, actin, cell membrane, ER, Golgi, lysosome, ribosome, mitochondria, nucleus, and transcription rather than involvement with OR. Here, we unequivocally show that certain intracellular proteins that are not involved in olfaction exhibit 4CSP consensus. All these proteins are parts of big molecular complexes found inside cells. Therefore, to ascertain how 4CSP interacts with intracellular proteins—specifically, actin and TIF— thorough investigation is needed, completely independently of a function in olfaction or chemosensing.

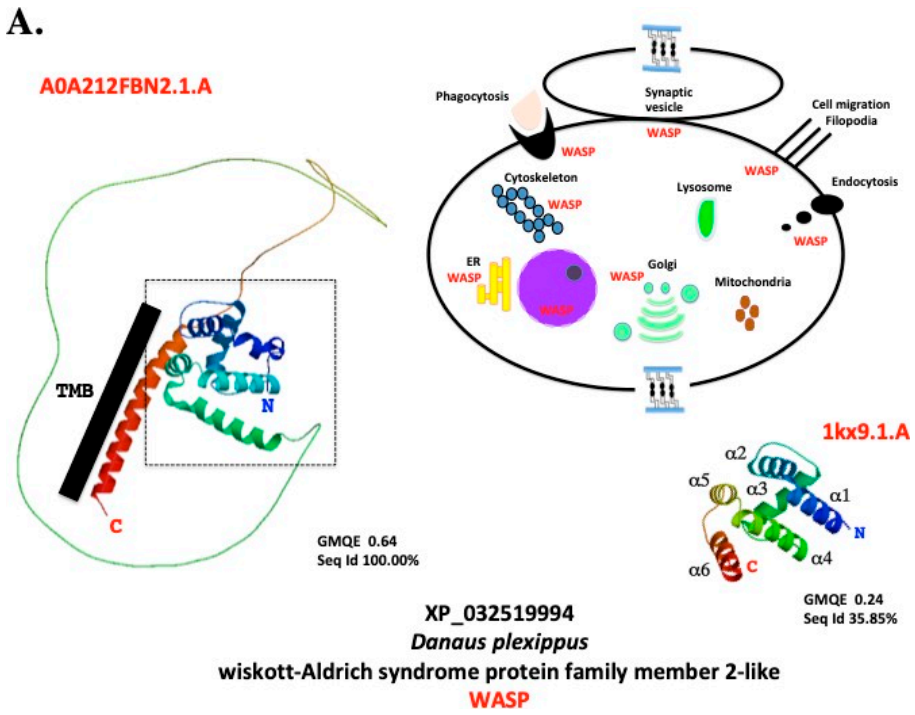
We introduce an alternative paradigm to the “outdated” olfactory one. We display the phylogenetic and modeling analysis of Mp10 orthologs, which comprise several intracellular protein families associated with actin, the nucleus, and various other intracellular organelles (see Figures 3 and 4). An unexpected yet obvious evolutionary relationship between these protein families is found using IQ-Tree and PAUP/MP with bootstrap support, indicating a similar evolutionary past (see [28] and Figures 3-4). We concur that the evolutionary phylogenetic analysis cannot be used to infer any function. However, rather than focusing on olfactory functions, the investigation of these well-established evolutionary links between 4CSP, Rho, TIF, ASRP, SamkC, NPCP, and viral tegument proteins surely opens a new avenue for this protein family. The role inside the cells depicted in Figure 5 is more consistent with the tissue distribution and ontogeny of 4CSPs (see [1]).

 doi:10.20944/preprints202604.2167.v1



 doi:10.20944/preprints202604.2167.v1

Figure 3. Molecular phylogenetic comparison (PAUP*10Altivec) of Mp10 with related proteins from. the 4CSP, Allergen, Mucin, Rho, TIF, ASRP, and NPCP families. (**A**) UPGMA analysis of Mp10 and counterparts (Table S2): agglomerative (bottom-up) hierarchical clustering UPGMA analysis based on the distance matrix of the analysed taxa that were calculated from a multiple alignment in ClustalX. Evolution (i.e., mutation rate) across all sequences is constant. Each pair-wise distance contributed equally. The red arrows represent gene duplication events (d1-d5) that led to Allergens ("IgE-binding protein") and the group of Mp10, Mucin, TIF, DAN4, NPCP, and Bommo-4CSP10 ("BmorCSP10"). The red asterisk indicates Mp10's position (*): grouping with pherokine XM056065396 and Mucins. On top of the branching tree are the model protein structures that correspond to the molecular groupings. (**B**) Bootstrap/Jackknife algorithm analysis of Mp10 and related Allergen, Mucin, TIF, ASRP, and NPCP proteins (Table S3) with bacterial *A. baummanii* CSPs (OIC81003 and OIC85870) as outgroup. The tree only displays bootstrap support values >50%. Amino acid tree (data matrix: total characters 953, constant characters 157, variable parsimony- uninformative characters 168, parsimony-informative characters 628, all characters of type unord, all characters have equal weight): Length 4366, CI 0.721, RI 0.847, RC 0.611, HI 0.279, G-fit -453.305). (**C**) Neighbor Joining analysis (phylogeny reconstruction, bootstrap method, 500 bootstrap replications, p-distance, uniform rates). The circular tree and the trees depicted above both read in the same order. The degree of relatedness among sequences from 4CSPs (red), Allergens (cyan), TIF (brown), PAN (salmon), ASRP (purple), NPCP (black), and Mucin (green) is indicated by the relative depths of the nodes supported by bootstrap values >50%. 4CSP (red) branches mix with Allergens, TIF, PAN, ASRP, Rho (in pink), NPCP, and Mucin branches. In red: Ebsp-3/PebIII A10/OS-D, cyan: Allergen Thap1, purple: Pherokine-3, orange: Acid trehalase, dark blue: CWA-3, brown: TIF, pink: Rho GTPase-activator, light green: Mucin-like/Extensin-like, grey: WAS/WASL, black: DAN4/NPCP, salmon: PAN-1, light grey: Formin-1, light grey in dark circle: WASP-2, light purple: Jg5928, yellow: RickA-like, white in red circle: YLP motif protein 1, X: Hypothetical protein (unknown function). For comparative molecular analysis, the amino acid sequences are used as templates for ClustalX alignments in A, B, and C (PAUP*10Altivec).



 doi:10.20944/preprints202604.2167.v1

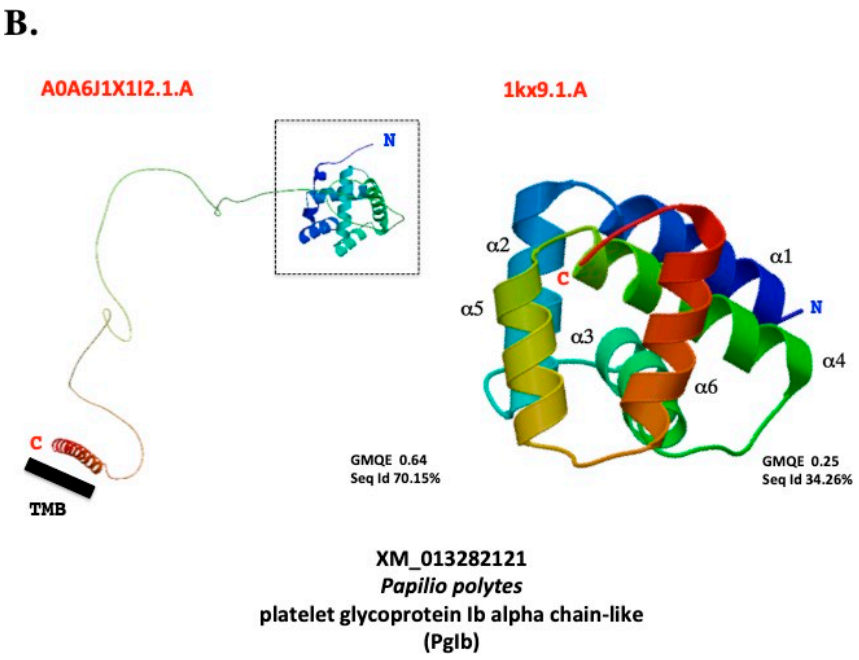


Figure 4. Molecular structure modeling of 4CSP and intracellular proteins. (**A**) Molecular structure. modeling of Danpl-WASP (XP_032519994). (**B**) Molecular structure modeling of Pappo-PgIb (XM_013282121). Danpl-WASP and Pappo-PgIb sequences were aligned with Bommo-4CSP1 in order to identify the signal peptide and cut it off based on N-terminal sequencing [7]. The amino acid sequence of the mature protein was then subjected to molecule structure modeling using SWISS-MODEL Workspace/GMQE [95]. The molecules with the highest identity score was used as template references: 1kx9.1 ("Chemosensory Protein A6", X-ray, 1.6 Å, monomer, Mambr) and A0A212FBN2.1.A (WASP family member 2-like, AlphaFold DB model of A0A212FBN2_DANPL, LOC116772069 gene, monarch butterfly, *D. plexippus*) for WASP (in **A**); 1kx9.1 and A0A6J1X1I2.1.A (Mucin-2-like, AlphaFold DB model of A0A6J1X1l2_GALME, LOC113521739 gene, greater wax moth, *Galleria mellonella*) for PgIb (in **B**). For WASP (**A**) and PgIb (**B**), the Global Model Quality Estimation (GMQE) and the percentage of Sequence Identity (Seq Id) are shown. C: C-terminus, N: N-terminus. The α-helices that make up the 4CSP structure are numbered 1 through 6. The location and function of WASP in the cytoskeleton, ER, Golgi, filopodia, nucleus, and near endocytotic, phagocytotic, and synaptic vesicles point to the N-terminal tail of WASP and PgIb protein molecules in the intracellular compartment as the location of the 4CSP structure. The 4CSP prism is indicated by the black square with dotted lines. The two molecules share the same model of construction inside the cell: N-terminus 4CSP prism, long loop, and transmembrane domain. The black bar indicates the position of the transmembrane segment (TMB; Swissmodel.expasy.org and services.healthtech.dtu.dk).

The close relationship between 4CSP Mp10, mucin, TIF, and many other "intracellular" regulatory components challenge the claim that 4CSPs are "chemosensory" molecules. Comparing Mp10 to the peplomer and cross linkers of the inner tegument, it appears that both bacteria and viruses express it (see Figures 2 and 3 and Table S2). This shifts the study of the 4CSP family into a completely different field—the detection of virus particles—which are quite distinct from scents and plant odors.

6.1. Evolutionary Evidence Derived from Phylogenetic Analysis

A preliminary phylogenetic analysis in IQ-Tree using an empirical Bayesian approach revealed some relationships between 4CSPs, CWP, Rho, Sec31, TIF, SamkC, WAS/WASL, DNA-binding proteins (DNA-BPs), DNA-regulatory proteins (DNA-RPs), and numerous RNA-binding proteins (RNA-BPs) [116]. For a more thorough phylogenetic analysis of 4CSPs and their intracellular equivalents in PAUP*-tree, we therefore chose these specific protein families, their molecular amino acid sequences, and their homologs from both bacteria and insects (see [28] and Tables S2-S4).

 doi:10.20944/preprints202604.2167.v1

Given that these families of intracellular proteins and enzymes were involved in a wide range of functions, one expected a rather heterogeneous evolutionary process, that is a high pace of changes over time, and numerous distinct evolutionary models. The fact that all these intracellular proteins interbred with the 4CSP family—which is traditionally referred to as "chemosensory"—was highly unexpected [28]. Various orthology groupings were found: 1) Trica-4CSP AAJJ0012J and xenobiotic response element (XRE) from *Ruminococcus torques* (WP_044998036), and 2) Transcriptional/cell division repressor, helix-turn-helix, and DRP from the XRE family of transcriptional regulators, branched with Bommo-4CSP8 and Bommo-4CSP9 [28]. This suggested that some 4CSPs have evolved to carry out tasks associated with nucleotide binding, transcription, translation, DNA/RNA templates, DNA/RNA control, and/or intracellular gene expression processes, at least in moths and beetles.

This initial study led us to compare 4CSPs with some intracellular components, such as TIFs and gene expression regulators. It is yet unclear how 4CSP proteins work in cells, the actin, and the nucleus. Additional CRISPR study is needed to investigate the intracellular function of 4CSP.

A phylogenetic study of Myzpe-Mp10, 4CSPs, XREs, Mucins, and RBPs in PAUP (*10Altivec) significantly confirmed the relationships between intracellular proteins and 4CSPs (see Tables S2-S4 27 and Figures 3, S2, and S3). A neighbor-joining phylogenetic tree (BioNJ study) indicates that Mp10 did not enter the "4CSP" group. Rather, it was significantly more strongly linked to the Mucins group (G1; Figure S2A). EbalCSP3 (QIS77910) and *E. balteatus* EbalDDB_G0285119X1 shared a high degree of similarity with Rho-activator isoform X3 and were connected to Bommo-4CSP10 in G1 (see Table S3 and Figure S2A). Thap1 and RBPs were clustered with the Trica-4CSP sequence AAJJ0012J. The marmalade hoverfly "EbalCSP4" (QIS77191) clearly separated from the "Mucin" and "4CSP" groups (see G2; Figure S2A). Only the G1 group maintained a relatively high bootstrap value (94%) according to the Jackknife analysis (bootstrapping calculation over 1000 repeats), suggesting that 4CSPs branch with Rho GTPase enzyme (Figure S2B). 4CSP, CWP, LRR, Mucin, SamkC, TIF, and WAS/WASL formed a common branch with a significant bootstrap value (57%) Mp10 slid off this group, but AglaCWPX3 drew it in (see Figure S2B). In this analysis, the reference outgroup (Bommo-4CSP2 [2]) attracted the papilionid “IpodCSP” (CAH2042437), DBPs, RBPs, and TIFs, which were positioned at the base of the tree (Figure S2B). As a result, "chemosensory CSPs" were shown to be strongly associated with numerous intracellular protein families, such as Rho-activators, translational regulatory factors, and mucins. The results of this phylogenetic analysis clearly indicated that "CSP" proteins should be renamed and that CRISPR ought to be utilized to thoroughly investigate these roles inside the cell and nucleus rather than outside of it.

By concentrating on Mp10 (also known as a very typical "4CSP", 153 amino acids, 17.2 kDa, "consensus residues", four cysteine-pattern, whole body expression), we found that this specific protein sequence (XP_022173691) when paired with intracellular protein families builds a phylogenetic tree supported by a high bootstrap value (Figure 3). These amino acid sequences were not chosen at random. The main components of this relatedness include allergens, mucins, TIF, NPCP, ASRP, and 4CSPs, which represent various insect groups (see Table S2). On the UPGMA tree, Mp10 does not cluster with other 4CSPs. On UPGMA, Mp10 and mucin proteins are the most similar pairs of taxa (Figure 3A). More distantly related is the 4CSP and Allergen group. Therefore, the analysis of Mp10 and its intracellular counterparts backs up our assertion that 4CSP has little to no impact on chemosensing. More precisely, TIF is linked to Mp10 and that can be used to bolster the assertion that 4CSPs are not involved in chemosensing or OR activation. The finding that 4CSPs are present in all groups of mucin and actin skeleton-related protein sequences in UPGMA and Bootstrap (Figure 3) further supports our claim that 4CSPs are not involved in smell. The relationship between Mp10, TIF, and ASRC demonstrated here suggests that CRISPR or *in situ*, distant from OR and plant odor detection, should be used to study the crucial function that 4CSP plays on nuclear and actin sites. Mp10 is a member of a broad category that includes NPCPs and all Mucin taxa. Specifically, it is closely related to Rho-activator protein (XM_056065396; Figure 3A). The small GTP-binding Rho families of proteins (20–30 kDa) are not extracellular molecules. These are intracellular proteins that

 doi:10.20944/preprints202604.2167.v1

regulate the Rho-GTPase signaling pathways associated with the actin cytoskeleton. They function as molecular switches that regulate several cellular processes, such as cytoskeleton-related events and gene transcription. All eukaryotes have “RhoGAPs”, one of the main groups of Rho-regulators, which have been shown to regulate several cellular processes, such as cytoskeleton structure, growth, differentiation, brain development, and synaptic activity [117]. They don't activate transmembrane chemosensory ORs and have nothing to do with smell. Accordingly, the functional significance of 4CSP in Rho-activators, TIFs, and nuclear pore complex proteins should be studied instead of 4CSP-OR.

The other 4CSP sequences belong to either the ASRP group or the Allergen group (which may be the most ancestral group based on its exterior position and distant branching), according to UPGMA analysis (Figure 3A). The ASRP group contains outer envelope proteins, RickA-like (Arp2/3), and nuclear nucleoside kinase (NNK), whereas the Allergen group contains IgE-binding proteins, "pherokines", acid trehalase (AHX71992), CWA-3 (XP_018563025), and TIF (XP_044745729; see Figure 3A). Odor transport and OR activation is unrelated to these proteins. They have closer ties to the cytoskeleton, actin-based cell motility, GTP-binding proteins, phosphorylation, mitochondrial energy production, and/or nuclear gene expression regulation [118–121].

The evolution of 4CSP does not start at OR. Mp10, 4CSP, Allergen, TIF, Mucin, NPCP, and ASRP have a very distant shared ancestral origin, according to the topology of the UPGMA tree, which assumes a common root and constant evolutionary rates for all lineages (i.e., it employs the "Molecular Clock Hypothesis" [122] to account for mutation rates). Prior to *Mp10*, *Mucin*, *NPCP*, and *ASRP*, *Allergen*, which includes *4CSP* and *TIF* genes, appears to be the result of a sequence of duplication events (Figure 3A). *Mp10* and *Rho* appeared later, prior to a series of duplication events that resulted in a variety of *Mucin* variants, especially in mosquitoes [123,124] (Figure 3A). Figure 3A shows that other gene duplications appear to be more ancient and widespread across all lineages. Flour beetles, ants, damselflies (*Ischnura* forktails), flies, garden whites (pierids), ladybirds, lice, neodiprions, parasitoid wasps, and tuberworm moths are among the taxa that made up the *Allergen* group. Similarly, the *ASRP* group included taxa from several moth and butterfly families, such as speckled woods and swallowtails (*Papilio* and *Pararge*; Figure 3A and Table S2). The tree illustrates the evolutionary gap between *ASRPs*/*4CSPs* and *Thap1* (Figure 3A). The Devonian or the most recent Mississippian–Silurian (i.e., roughly 324–440 Mya, [125]) may have been the time when a distant ancestor of all these protein families evolved, which is significantly earlier than the appearance of flying insects.

Maximum parsimony (MP) analysis in PAUP* was used to further investigate the relationships between the Mp10, 4CSP, Rho, Allergen, TIF, Mucin, NPCP, and ASRP proteins (Figure 3B). The insect 4CSP, Mp10, Rho, Allergen, TIF, Mucin, NPCP, and ASRP amino acid sequences formed groups with a high bootstrap value (89-100%): they are closely related to each other. The bootstrap value of the NPCP and related Mucin proteins was about 99%. Furthermore, the grouping of ASRP/Rho and Mucin/NPCP was supported by Jackniffe's high bootstrap values (about 93-99%; Figure 3B), indicating a close relationship between all these intracellular protein families. Importantly, 4CSPs show a strong relationship with PAN and ASRP, as demonstrated by relatively high "bootstraps" (94-100%; see Figures 3B & 3C). Like 4CSPs ([see 1]), Mp10, Rho, Allergen, TIF, Mucin, NPCP, and ASRP molecules are extensively expressed in all adult body parts, including the thorax and abdomen, as well as in larvae and pupae (see Tables S2-S3). These proteins are more involved in DNA/protein interaction, enzymatic activities, and cell response to stress than they are in molecular olfactory mechanisms [126,127]. This is strikingly similar to the tissue distribution, response to stress, and ontogeny of 4CSP (refer to [1]). The strong relationship between 4CSPs and intracellular proteins, cell walls, and gene promoter regions—which is indicated by both expression analysis and phylogeny data—should be thoroughly investigated in future studies (see Figure 3). Instead of targeting OR, CRISPR may target this far more prevalent 4CSP function.

The NJ phylogenetic tree, which is structured like a circle and has branches that interlock for 4CSP, Allergen, TIF, PAN, ASRP, NPCP, and Mucin, also clearly illustrates the links between these

 doi:10.20944/preprints202604.2167.v1

protein families (Figure 3C). Because of the high "bootstraps" between 4CSP and intracellular proteins, the MP and NJ do not support the role of 4CSP proteins in olfaction, where they are extracellularly active to transport odorant chemicals from the environment to ORs, which in turn activate ORNs and ultimately the brain. Our findings corroborate an entirely new working hypothesis: 4CSPs are intracellularly expressed, influence the brain and other nerve tissues, and control processes essential to the operation of the sensory or integrative cell. Mp10 and TIF remained near to one another in all phylogenetic analyses, which shows that these two proteins shared similar evolutionary pathways (see Figures 3B, 3C, and S3). Similarly, the ASRP group always approaches Mp10 when bootstrapping on the MP or NJ tree (Figures 3B, 3C), strongly indicating that, like *4CSPs*, *TIF* and *ASRP* genes are developing at a fast, intense, frequent, and high rate despite their extreme age—the eukaryotic cell origin [128–131]. Our evolutionary analysis suggests that the TIF and ASRP complexes and 4CSPs share an ancient mechanism for controlling actin skeleton and/or ribosomal gene expression, even though much more experimental data is required to adequately explain such functional links (Figure 3).

Our evolutionary analysis revealed proteins associated with Allergen/Mp10 in a variety of insect taxa, including aphids, beetles, dragonflies (Odonata, Zygoptera, *Ischnura*, blue-tailed damselfly), flies, lice, mosquitoes, moths, and sawflies. These proteins have nothing to do with taste, smell, and/or chemosensing and are members of several families of intracellular proteins, including Mucin, NCPC, ASRP, and TIF (Figures 3, S2, S3, and Table S2). This suggests that the Mucin, NPCP, ASRP, Allergen, TIF, and 4CSP families shared diversification much before insects emerged, for example, during the Carboniferous Period of the Paleozoic era (~299-359 Mya; according to the duplication profiling in Figure 3).

Based on the duplication profiling shown in Figure 3, it appears that "*Mp10*" gene has experienced several duplication events. *Allergens/4CSPs*, *TIFs*, and *ASRPs* were produced by two early duplications (d1 and d2), while *Mp10* and *Mucins* were formed by two late duplications (d3 and d4). The formation of the Mucin family appears to have required these four consecutive duplications (d1-d4). All of these intracellular protein families seem to be "sisters" when we use the protein amino acid tree to determine duplication profiling and, consequently, evolution: all of the molecules in the Mucin, NPCP, ASRP, Allergen, PAN, TIF, and Mp10/4CSP families share an old common similar root (see Figures 3, S2 and S3 and Table S2).

6.2. Functional Evidence Derived from Amino Acid Sequence Modeling Analysis

Protein primary sequence alignment and structure modeling indicate the evolutionary relatedness of these protein families (Table S2, Figures 3 and 4 and S3-S4). Protein amino acid sequence alignment shows that 4CSP exactly matches the N-terminus of TIF, Mucin, NPCP, and ASRP molecules as well as the Allergen sequence (Figure S3). Using SWISS-MODEL Workspace/GMQE, most intracellular molecules fold into the same tunnel-like shape of six α-helices when compared to 1kx9.1 as the template reference (X-ray crystal structure, Mambr) [15,132]. Allergen Thap1, acid trehalase, Phk-3, Bommo-4CSP, Mp10, immune response proteins, and *Coccinella septempunctata* TIF all had distinctive 4CSP structure (see Figure S4). The other proteins that were analyzed using the SWISS-Model were all very big molecules. On these models, the tunnel structure was found in specific regions of the molecule, such as the N-terminus (Figures 4 and S4). The transmembrane region was represented by the C-terminus of TIF, Mucin, NPCP, and ASRP molecules, which had a very long α-helical stretch (7–13 turns; Figures 4A, B, and S4; see GMQE 0.64, Sequence Identity 70-100%; Danpl-WASP, XP_032519994; Pappo-PgIb, XM_013282121). The structural structure and gene locus of Danpl-WASP (AlphaFold DB model of A0A212FBN2_DANPL and LOC11677206) are well known. The structure of A0A212FBN2 (AlphaFold) has nothing to do with plant odor-binding protein. We obtained the same results with SWISS-Model (Figure 4A). The Danpl-WASP protein is composed of a single, long transmembrane domain with a mass of 33 kDa and 297 amino acids. The Danpl-WASP structure is most used in filament synthesis and actin cytoskeleton regulation rather than olfaction, and it can bind actin but not extracellular odorant

 doi:10.20944/preprints202604.2167.v1

molecules [133,134]. These bigger molecules, such as the WASP structure, are found in intracellular organelles such as the cytoskeleton, microtubules, integuments, cell wall, nuclear pore, and cytoplasm. Our analysis using AlphaFold DB models revealed that TIF, Mucin, NPCP, ASRP, and WASP are distinguished by the unusually large loop that connects the N-terminus and C-terminal transmembrane domain of the 4CSP (see Figures 4 and S4).

Interestingly, since this peculiar protein structure is seen in several viral proteins, we can go a little closer to the presence of 4CSP from sensilla to viruses. The structure of the Palearctic butterfly protein CAH2235359, commonly known as "jg5928," strongly suggests that 4CSP might be found in viral particles. Similar to WASP and PgIb, Jg5928 folds into a large loop that connects the transmembrane domain of the primary outer envelope glycoprotein of the virus (BLLF1, pfam0519; Figure S4) to the 4CSP domain. Given that the structure of 4CSP appears to be an essential component of many intracellular proteins and the glycoproteins that comprise the outer envelope of viruses (see "jg5928" model; A0A1V1WC08, GMQE 0.64, Sequence Identity 70.31%, Figure S4), the classification of this protein as "chemosensory" will therefore be somewhat completely irrelevant.

4CSP's presence in viruses makes sense because it is a molecular component of a larger assembly that also includes envelope glycoproteins. 4CSP may be the glycoprotein element of the viral particle that attaches to the lipids of the host. Interestingly, we found that some proteins in bacteria had a similar molecular structure: a transmembrane region and a glycoprotein connected in a bigger molecule with a 4CSP Pfam Domain (see *Lysobacter capsici* WP_096417339, 491 amino acids, 50.6 kDa) [14,28]. Except for the glycoproteins discussed here having the 4CSP module at the N-terminus whereas bacterial cells have it in the C-terminus, we found that the structure of bacterial “big” 4CSPs was like that of “large” 4CSPs from insects (see Figure 4 and [14]). This may make it clearer how 4CSP contributes to immune response activation, cell signaling, and/or infectious disease processes. Rather than “olfaction”, the main function of 4CSP may be “infection”: interaction with lipids, binding of bacterial and viral particles, and immune response. The detection of plant odor volatiles and/or sex pheromones is a great cry from this.

6.3. “4. CSP, Intracellular Mode of Action” Derived from Location, Size, Structure, and Expression in Microbes and Viruses

Given the striking similarities between 4CSP, Allergen, TIF, Mucin, NPCP, and ASRP families, we hypothesize that 4CSP performs a wide range of functions within the intracellular systems of insects and hexapods (Figure 5). The "4CSP intracellular mode of action" hypothesis, which is depicted in Figure 5, must be further studied using CRISPR, RNAi, and/or gene knockout, as well as more biochemistry and molecular biology approach applied to general cell function. This may entail, for instance, determining the quantity of FA lipids in the cells and locating specific 4CSP near organelles such the ER, nuclear pores, mitochondria, and Golgi.

 doi:10.20944/preprints202604.2167.v1

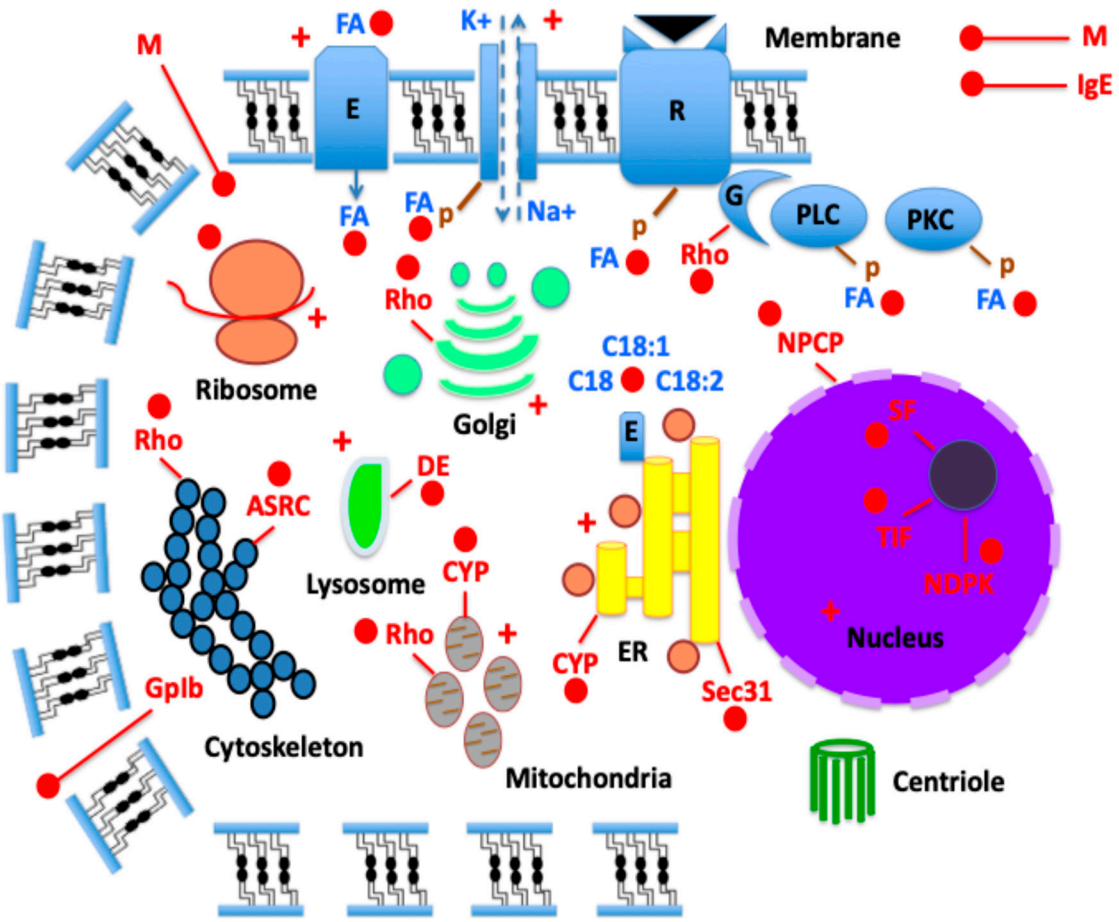


Figure 5. The "4CSP Intracellular Mode of Action" Hypothesis. Potential roles of 4CSPs (Mp10 subfamily) in the intracellular systems of insects and hexapods. 4CSP binds to fatty acid (FA), which mediates the phosphorylation (p) of different plasma membrane-bound protein molecules in this theoretical model. E: Enzymes, K+ intake/Na+ outflow: potassium-sodium exchange pump, R: Receptors, G: G-protein, PLC: phospholipase kinase C, PKC: protein kinase C, Rho: Rho GTPase, ER: endoplasmic reticulum, E: Desaturase enzyme, which interacts with 4CSPs that transport FAs, such as linoleic acid (C18:2) and its precursors (stearic acid C:18 and elaidic acid C:18-1). +: Stress reactions that lead to 4CSP interacting with different molecules in the membrane, nucleus, ribosome, golgi, lysosome, ER and mitochondria, including cytochrome P450 (CYP), Degradative Enzyme (DE), Sec31 protein complex (Sec31), and Mucin (M). In the theoretical framework for 4CSP intracellular mode", additional molecules in the cytoskeleton, plasma membrane, and nuclear membrane that 4CSP proteins bind to are the actin skeleton regulatory complex (ASRC), Gplb-like protein (Gplb) and nuclear pore complex protein (NPCP). In the nucleus: splicing factor (SF), transcription initiation factor (TIF), and nucleoside diphosphate kinase (NDPK). IgE stands for Immunoglobulin E-like protein (parasitic infection, venom protection), while M stands for mucin-like fractions, secretory molecules that fortify the immune system's defense against microbial stress. The big red dots indicate each putative protein, protein assembly, or supramolecular complex in each intracellular organelle that 4CSP molecules may interact with in both insects and hexapods based on phylogenetic and modeling analysis of the amino acid sequence (Mp10).

4CSP binds to LA (C18:2, ω-6 FA) according to the "4CSP intracellular mode of action" hypothesis (see [21]). The production of several hormones depends on LA, which is generated from arachidonic acid (ARA, C20:4 ω-6 FA lipids). The hormones prostaglandins, thromboxanes, and leukotrienes are derivatives of ARA and LA. Numerous physiological functions, including ion transport, innate immune response, egg development, and reproduction, are regulated by these hormones [135–137]. Furthermore, C18:2 and C20:4 ω-6 FA lipids are known intracellular process modulators. LA and ARA phosphorylate a wide range of intracellular proteins, including channels, enzymes, pumps,

 doi:10.20944/preprints202604.2167.v1

receptors, and many more. Numerous signal transduction pathways and physiological functions are subsequently regulated by this phosphorylation (see [22]).

Given this, we suggest that 4CSPs take part in the phosphorylation and lipoid processes in the many intracellular organelles that influence the control of development and stress responses (see +, Figure 5). This role in relation to phosphorylation is consistent with the fact that the LA-4CSP complex is highly broadly expressed throughout the whitefly's body, not just in the antennae (see [21]). Because of its link to LA and protein phosphorylation, the 4CSP might therefore mediate a very broad range of physiological processes (Figure 5). These cellular activities that depend on LA and protein phosphorylation are not only applicable to all hexapods and insects, but they also represent extremely general principles of bacterial cell biology. Regarding the positions of 4CSPs, this makes perfect sense (see [1 & 14]). LA is known to promote calcium signaling and kinase phosphorylation in gustatory cells, modulate ion channels in metabolic cells, and activate a variety of signaling pathways and receptors in mammals [138,139]. The brain, central nervous system (CNS), and metabolic organs like the gut and fat body are among the several tissues in insects that LA is known to target [37,140–143]. The ER, Golgi, mitochondria, lysosomes, ribosomes, proteasomes, microtubules, and actin cytoskeleton are all targets of LA [144–149]. Specifically, LA is essential for the synthesis of components involved in oxidative phosphorylation (OXPHOS, the electron transfer chain linked to ATP synthesis in the mitochondria), which is linked to insects' immune response [150]. When mitochondria, OXPHOS, and LA are combined in insects, an immune response occurs [151]. The highly broad activity of LA consists of the highly broad tissue expression of 4CSP (see [1]). 4CSP can only be found inside cells, not extracellularly, on the outside of OR in the antennae, since it binds to LA and other intracellular FA lipids and is connected to intracellular proteins found in all different organelles of the cell in many different tissues (see [1, 15, 21, 22, this paper]). In contrast to pheromone, plant odor, and chemical sensing, our data clearly situates 4CSP research in the field of general cell biology.

First, in conjunction with LA, ARA, ω-6 FAs, and stress responses, 4CSPs would mediate the molecular translational mechanisms controlling the overall lipid metabolism in the cell cycle of bacteria, insects, and hexapods [1,14,21,22,152–155]. Additionally, 4CSPs would interact with cytochrome P450 enzymes in the mitochondrial system and the ER in response to insecticide exposure (+) and other stressors. Cellular metabolism, homeostasis, JH synthesis, toxin catabolism, and detoxification in response to insecticide exposure would all depend on these 4CSP-CYP interactions [2,156,157]. Second, the lysosome's defense against pathogen particles like microorganisms would be handled by the 4CSP-Enzyme system. It would also be in charge of the lysosome's ability to break down and digest macromolecules like lipids, sugars, and nucleic acids (Figure 5).

The "4CSP intracellular mode of action" hypothesis (see Figure 5) states that FAs like LA and/or its lipoid precursors would be transported by the ER membrane, desaturase enzymes, and eversible vesicles that produce the duct, cuticule, and sex pheromone. Specifically, we speculate that ER, FAs, and 4CSPs must work closely together to facilitate lipid droplet production and pheromonogenesis. 4CSPs might regulate the ER membrane of the female moth sex pheromone gland as a particular site of lipoid FA metabolism, instead of activating OR on the neural dendrites of the male antennae [158–160]. By using CRISPR to study 4CSP, FA, lipid droplet, CYP, desaturase, and pheromonogenesis, this new working hypothesis may be readily examined. The tissue distribution of 4CSPs (see [1]) and the function of 4CSP-LA in several lipoid biosynthetic pathways provide strong support for this (see [22,107]).

According to the "4CSP intracellular mode of action" hypothesis, 4CSPs are also linked to the nucleus, cytoskeleton, ER, Golgi and mitochondrial endomembranes, and plasma membrane through the Rho GTPase signaling complex, Mucin, ASRC, and NPCP proteins (see Figures 3-5). 4CSP, Rho, Mucin, ASRC, and NPCP may regulate the flow of lipids and proteins between the different cellular organelles, particularly between the Golgi apparatus and the lysosome [161–163] (Figure 5). Intriguingly, the 4CSP Pfam domain constructs the N-terminus of a protein called "Pglb" that resembles a glycoprotein receptor on the cellular surface rather than an OR (see Figures 4A and S3).

 doi:10.20944/preprints202604.2167.v1

Even though the functional role of 4CSP in glycoproteins still must be investigated, this is a very significant finding from an evolutionary perspective. Glycoprotein receptor-related mechanisms are extremely old, dating back to worms [164]. This supports Picimbon's assertion that *4CSP* is an incredibly old gene (see [see 14].

4CSPs existed long before ORs and ODEs as evidenced by their presence in bacteria and a glycoprotein in the ancient enteric endocrine system that governs intestine function in ecdysozoans (protostome animals). Rather, this ancient ancestry connects them more to the "intracellular" P450 family (CYP genes, closely linked to ER and inner mitochondrial membrane), which, all the way back to Bya [165,166], codes for a variety of enzymes involved in the metabolism of xenobiotics as well as in essential endocrine and/or ecophysiological functions (see Figure 5).

7. Concluding Remarks and New Perspectives

Among a wide range of intracellular protein families, our study's most exciting finding might be the relationship between 4CSPs and Mucins. Given how similar the two protein families are in terms of evolution, physiology, ontology, and tissue distribution, this makes perfect sense [1,167–170]. Like 4CSP, Mucin exhibits a very wide tissue expression and plays a crucial role in insect growth and development [171–173]. In addition to growth and development, the 4CSP and Mucin families will work together to regulate the gut and immune response. Mucus-forming mucins that surround nutrients and activate recycling enzymes are one characteristic of the insect's midgut [174]. Additionally, the insect’s gut contains mucins that are intended for vertical symbiont transfer [175]. These "gut" mucins, like many 4CSPs (see Bommo-4CSP10, BABH01021709 [2]), are distinguished by highly glycosylated structures that make up the functional domain, which is rich in Thr, Ser, and Pro residue repeating sequences. The multi-proline domain, expression in the midgut, fat body, and symbiotic organ all suggest the close relationship between 4CSP and Mucin [176,177].

Intriguingly, blood feeding significantly increased Mucin expression in the mosquitoes' midgut, fat body, and ovaries [178], a scenario that is strikingly comparable to the 4CSP gene family [179,180]. We hypothesize that 4CSPs play similar roles to mucins in defense, growth, development, metabolism, digestion, and reproduction across a range of insect species, without limiting the function to insect palps. The fact that mucins, which are widely expressed in true bug salivary glands and secreted into rice during feeding, trigger plant immune responses in a manner like that of Mp10 [181] may be the most important aspect for new research direction in this context. This confirms the significance of these two protein families far away from olfaction.

The relationships between 4CSPs, Mucins, ASRPs, NPCPs, cytoskeleton complexes, genetic elements, nuclear pores, and intracellular activities that we explore here were not even conceivable during the "OS-D/CSP" era (1994-2003). More study has been conducted on the non-chemosensory functions of this protein family and in relation to insecticide resistance since "OS-D/CSP" molecules have been found in response to insecticide [1,2,21,22]. However, the potential intracellular function of "OS-D/CSP" in connection with ASRP, NPCP, Mucin, TIF, transcription initiation, RNA, DNA, nucleoside kinase, and cell regulation has never been touched before. The phylogenetic relatedness between "OS-D/CSP" and many intracellular protein families is highly intriguing and provides yet another compelling argument for renaming the "OS-D/CSP" family (see [1]), even though more functional research is required to fully characterize the function of "OS-D/CSP" protein inside the cells, actin, and nucleus.

Without experimental proof, these assertions regarding 4CSP for intracellular trafficking (see Figure 5) remain theoretical. The intracellular localization of 4CSPs can be expected to play greater biological roles, although specific investigation is required to confirm causation and functional processes. Currently, the primary goal of 4CSP research is to understand insecticide detection through OR and potential applications. Rather, we present a novel research challenge for insects: concepts that explain how 4CSPs control actin, transcription, or viral assembly by intracellular mechanisms like this one (Figure 5). We recommend adopting CRISPR experiments in future studies to examine "4CSP" in general cell function rather than concentrating on "OS-D/CSP" in olfaction.

 doi:10.20944/preprints202604.2167.v1

Numerous studies on a wide range of insect species have examined the tissue distribution of 4CSPs, and they are consistently found to be beyond the olfactory paradigm. Additionally, they are quite distant from the cuticle and the integument (see [1]).

Another important argument against olfaction is the Mp10 experiment. The "phylogenetic experiment" results demonstrate the exceptionally high sequence homology between Mp10 and Mucin. The PAUP tree and primary sequence alignment indicate where the "A10/OS-D" Pfam domain is in several intracellular protein families —at the N-terminus— (see Figures 3 and 4 and S1-S4). It becomes essential that it be called "4CSP" considering this data. Actin-related proteins, nuclear complexes, TIFs, splicing factors, and mucins are examples of larger intracellular molecules or protein complexes that 4CSP "binds" too. As part of new investigation into 4CSP, actin, and the nucleus, the structure of the intracellular protein containing 4CSP at the N-terminus, as indicated by the prediction 3D models (see Figure 4), has to be experimentally confirmed using X-ray and/or NMR study. Instead of focusing on the "olfactory" component, future studies should thoroughly consider the “intracellular” component.

For example, it would be interesting to use CRISPR to cut the 4CSP at the transcription initiation factor's N-terminus and observe how this affects the function of the cell or nucleus. The discovery of 4CSP in TIF is exciting. It should be studied carefully because it touches on both functional and evolutionary elements. To understand how multiple structural domains or modules came together to form big intracellular protein complexes, it may be crucial to understand how and when 4CSP linked transcription factors. In addition to TIF, 4CSP is also believed to be present in actin complexes, pore nuclear systems, nucleotide-binding proteins, and gene regulators. This strongly suggests that the 4CSP's main role is not to scavenge odorants but rather to operate as a module for intracellular regulatory molecules. Unlike the protein's capacity to detect scents in the antennae, 4CSP's intracellular location correlates with its dispersion over a wide variety of tissues, which is a known and well-established fact for this protein family (see [1]).

There is more consensus here about the protein's capacity to carry lipoid molecules. Many insect species, particularly moths, use 4CSPs and FAs to produce sex pheromones. Moreover, a variety of cellular processes, such as intracellular signaling and gene expression patterns, are greatly impacted by fatty acid lipids. A function related to lipid metabolism strongly supports the 4CSPs' position. Bacteria, viruses, insects, and hexapods all have these proteins, which are also present inside the cell and nucleus, in each of the many cell organelles, and in all their functions (see "4CSP Intracellular Mode"). The strong interactions that mucins, 4CSPs, and OBPs have with lipoids provide additional evidence for this [182–188]. While 4CSPs (and OBPs) may transport lipoid molecules, mucins may aid in lipoid agglomeration and droplet formation. It has long been known that gut mucins can bind lipids as a tissue-protective mechanism [189–191]. Mucins can directly interact with gut microbial flora thanks to short-chain FA lipids [192]. Therefore, the idea that mucins interact with short-chain FAs while 4CSPs interact with long-chain FAs may be another theory that has to be investigated further using binding assay and X-Ray.

Here, we give an additional set of data, phylogeny and intracellular localization, that contradict the role of the "CSP/OS-D" protein family in olfaction, following the tissue-distribution and developmental profile, reaction to insecticide, and binding to lipids. This strongly suggests renaming the protein to something more neutral, like “4CSP” (see [1]), without assuming anything about its function. Instead of odors, 4CSPs appear to be able to bind FAs, lipids, phenyls, transcription factors, nucleotides, DNA, RNA, and many other intracellular (actin and nuclear) components. This shows that OS-D, CSP, and "sensory appendage protein" ("SAP") are no longer suitable names for this protein gene family.

Supplementary Materials: The following supporting information can be downloaded at the website of this paper posted on Preprints.org, Figure S1: Alignment of the amino acid sequence of Mp10 with those of other aphid and hemipteran sequences as well as their InsectBase counterparts in Coleoptera, Hymenoptera, and Lepidoptera; Figure S2: PAUP*10Altivec phylogenetic analysis of insect 4CSPs, DBPs, and RBPs; Figure S3:

 doi:10.20944/preprints202604.2167.v1

Myzpe Mp10's alignment with related proteins from the 4CSP, Allergen, Mucin, Rho, TIF, ASRP, and NPCP families; Figure S4: The modeling of the molecular structures of Mp10 and its related proteins belonging to the Allergen, Mucin, Rho, TIF, ASRP, and NPCP families in SWISS-MODEL. Table S1: Acypi '4CSP' gene repertoire in comparison to other Hemiptera, genome assembly, and protein identity; Table S2: Sequences producing significant alignments with Mp10 protein; Table S3: Sequences producing significant PAUP-Tree between 4CSP and intracellular proteins; Table S4: Sequences of RNA-binding proteins (RNA-BPs) used in PAUP-tree for molecular comparisons with 4CSPs and other intracellular proteins.

Author Contributions: Author Contributions: Conceptualization, J.-F.P.; methodology, J.-F.P.; software, J.-F.P.; validation, J.-F.P.; formal analysis, J.-F.P.; investigation, G.L., B.S., F.W., S.Y., Q.H. and J.-F.P.; resources, G.L., S.Y., Q.H. and J.-F.P.; data curation, G.L., B.S., F.W., Q.H. and J.-F.P.; writing-original draft preparation, J.-F.P.; writing-review and editing, J.-F.P.; visualization, G.L, B.S., F.W., Q.H. and J.-F.P.; supervision, J.-F.P.; project administration, J.-F.P.; funding acquisition, G.L., S.Y., Q.H. and J.-F.P.; All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Overseas High-Level Talent Program-Taishan Scholar title [NO.tshw20091015] and Ministry of Science and Technology of China-National Expert title [G2022023033L]. J.-F.P is Qilu Friendship Award (1st July 2025).

Data Availability Statements: The original contributions presented in the study are all included in the article/supplementary materials, further inquiries can be directed to the corresponding author.

Acknowledgments: Overseas high-level talent program-Taishan scholar title (#NO.tshw20091015), Ministry of Science and Technology of China-National Expert title (#G2022023033L), and Qilu Friendship Award 2025. We would especially like to express our gratitude to Alejandro Rooney (USDA) and Rudolf Alexander Steinbrecht (MPI) for their critical and enlightening discussions about the re-naming of "CSPs" and for sharing their ideas and opinions on the rough drafts of the (two) manuscripts (Part-1 and Part-2). Special heartfelt thanks and appreciation are also extended to Karl-Ernst Kaissling regarding "4CSP".

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

4CSP	4-Cysteine Soluble Protein ("CSP/OS-D")
6CSP	6-Cysteine Soluble Protein ("OBP")
AcrR	Regulator of adjacent acrAB efflux genes
Acypi	*Acyrthosiphum pisum* (pea aphid)
Aedae	*Aedes aegypti* (dengue yellow fever mosquito)
AI	Auto-inducer
Allergen Tha p 1	IgE-binding protein (15 kDa) and major allergen of pine processionary caterpillar (*Thaumetopoea pityocampa*, Thapi)—variant 1
Anoga	*Anopheles gambiae* (African malaria mosquito)
ARA	Arachidonic acid
Arp2/3	Actin related protein 2/3 complex
ASRC	Actin skeleton regulatory complex
ASRP	Actin skeleton regulatory protein
Avd	Accessory variability determinant
Bemta	*Bemisia tabaci*
Bemta4CSP1	*Bemisia tabaci* "Chemosensory Protein"-1 renamed to Bemta4CSP1
BioNJ	Bio (improved version) of Improved version of Neighbor Joining algorithm based on simple model of sequence data
BLASTp	Protein BLAST
BLLF1	Epstein-Barr virus envelope glycoprotein encoded by BLLF1 gene
Bomte	*Bombus terrestris*
CA	Corpora allata

 doi:10.20944/preprints202604.2167.v1

CHC	Cuticular hydrocarbons
Chisu	*Chilo suppressalis* (Asiatic rice borer or striped rice stemborer)
CNS	Central nervous system
Cocse	*Coccinella septempunctata* (seven-spot ladybird)
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
CSP	Chemosensory protein
Culpi	*Culex pipiens* (common house mosquito)
CWA	Cell wall anchored
CWP	Cell wall protein
CYP	Cytochrome P450
DAN4	Cell wall mannoprotein expressed under delayed anaerobic conditions (*Saccharomyces)*
Danpl	*Danaus plexippus* (monarch butterfly)
DE	Degradative enzyme
Denpo	*Dendroctonus ponderosae* (mountain pine beetle)
DGR	Diversity generating retroelement
DNA-BP	Deoxyribonucleic acid-binding protein
DNA-RP	Deoxyribonucleic acid-regulatory protein
DP	Diphosphate
Drome	*Drosophila melanogaster* (fruit fly)
EA	Elaidic acid
Ebsp	Ejaculatory bulb-specific protein (also called Peb)
ER	Endoplasmic reticulum
Eupco	*Eupeodes corollae* (migrant overfly)
FA	Fatty acid
Fox	Forkhead box (transcriptional regulator, cell growth regulator)
GDP	Guanosine diphosphate
GMQE	Global Model Quality Estimate
GpIb	Glycoprotein Ib
GR	Gustatory receptor
GTP	Guanosine triphosphate
Halha	*Halyomorpha halys* (brown marmorated stink bug)
Helme	*Heliconius melpomene* (postman butterfly)
Hu IgE	Human immunoglobulin E
Histone-like	Histone-like RNA/DNA-binding protein family (nucleus)
IgE-BP	Immunoglobulin E-binding protein
Iphpo	*Iphiclides podalirius* (scarce swallowtail)
IR	Ionotropic receptor
Jg5928	Major outer envelope protein
JH	Juvenile hormone
LA	Linoleic acid (cis,cis-9,12-octadecadienoic acid C18:2)
LRR	Leucine-rich repeat protein complex
M	Mucin
Mambr	*Mamestra brassicae* (cabbage moth)
ML	Maximum likelihood
MP	Maximum parsimony
Mp10	Myzpe p10-like protein
Mraz	DNA-binding transcription factor encoded by mraZ
Myzpe	*Myzus persicae* (green peach aphid)
NDPK	Nucleoside diphosphate kinase
Nillu	*Nilaparvata lugens* (brown planthopper, BPH)
NME	Nucleotide metabolism enzymes
NNK	Nuclear nucleoside kinase
NPCP	Nuclear pore complex protein
OBP	Odorant-binding protein
ODE	Odorant degrading enzymes
OR	Olfactory receptor
OS-D	Olfactory Specific-type D protein
OXPHOS	Oxidative phosphorylation

 doi:10.20944/preprints202604.2167.v1

P	Phosphate
Pacve	*Pachypsylla venusta* (hackberry petiole gall psyllid)
PAN	Hexameric ATPase complex
PAN-1	Protein encoded by the pan-1 gene in the nematode *Caenorhabditis elegans*
Pappo	*Papilio polytes* (common Mormon)
PAUP	Phylogenetic analysis using parsimony
Pedhu	*Pediculus humanus humanus* (human body louse)
PgIb	Platelet glycoprotein Ib alpha chain-like
Phk	Pherokine (hemolymph protein)
PKC	Protein kinase C
PLC	Phospholipase C
QS	Quorum sensing
Ras	Family of GTPases derived from rat sarcoma virus
RBP	Ribonucleic acid-binding protein
Rho	Family of GTPases, family of small (~21 kDa) signaling G proteins, subfamily of the Ras superfamily
RhoGAP	Rho GTPase-activating protein, regulator of Rho-related protein family (motility, contractility, growth, differentiation, development)
RickA	*Rickettsia conorii* surface protein A (activator of Arp2/3)
RNAi	RNA interference
SA	Stearic acid
SamkC	Serine/Threonine-protein kinase C (from social amoeba)
SAP	Sensory appendage protein
Sec31	Protein transport protein encoded by the SEC31A gene (human)
SF	Splicing factor
Tat	Twin-arginine translocation pathway
TERT	Telomerase reverse transcriptase
TetR	Tetracycline repressor
TFTR	TetR-family transcriptional regulator
TIF	Transcription initiation factor
TP	Triphosphate
Trica	*Tribolium castaneum* (red flour beetle)
TSSC1	Tumor suppressing subtransferable candidate 1 (endosomal machinery)
UL36	Large tegument protein deneddylase encoded by UL36 gene (herpesviridae)
UPGMA	Unweighted pair group method with arithmetic mean
Uralo	*Uranotaenia lowii* (pale-footed Uranotaenia)
WAS/WASL	Wiskott-Aldrich Syndrome / Wiskott-Aldrich Syndrome-like protein
WASP	Wiskott-Aldrich Syndrome protein
XRE (Xre)	Xenobiotic response element family of DNA-binding transcriptional regulators
YhbY	RNA-binding protein (folded like TIF) encoded by YhbY gene (*Escherichia coli*)
YLPM1	YLP motif containing 1 encoded by YLPM1 gene (human)

References

1. Liu, G.; Sun, B.; Wei, F.; Yue, S.; He, Q.; Picimbon, J.F. Renaming the 'OS-D/CSP' family (Part 1): '4-Cysteine Soluble Proteins' (4CSPs)—Molecular nomenclature, structure, evolution, expression, tissue-distribution, and pleiotropy. *Insects* **2026**, *17*, 202.
2. Xuan, N.; Guo, X.; Xie, H.Y.; Lou, Q.N.; Bo, L.X.; Liu, G.X.; Picimbon, J.F. Increased expression of CSP and CYP genes in adult silkworm females exposed to avermectins. *Insect Sci.* **2015**, *22*, 203-219.
3. Guo, X.; Xuan, N.; Liu, G.; Xie, H.; Lou, Q.; Arnaud, P.; Offmann, B.; Picimbon, J.F. An expanded survey of the moth PBP/GOBP clade in *Bombyx mori*: New insight into expression and functional roles. *Front. Physiol.* **2021**, *12*, 712593.
4. Pikielny, C.W.; Hasan, G.; Rouyer, F.; Rosbach, M. Members of a family of *Drosophila* putative odorant-binding proteins are expressed in different subsets of olfactory hairs. *Neuron* **1994**, *12*, 35-4.
5. McKenna, M.P.; Hekmat-Scafe, D.S.; Gaines, P.; Carlson, J.R. Putative *Drosophila* pheromone-binding-proteins expressed in a subregion of the olfactory system. *J. Biol. Chem.* **1994**, *269*, 16340-16347.

 doi:10.20944/preprints202604.2167.v1

6. Angeli, S.; Ceron, F.; Scaloni, A.; Monti, M.; Monteforti, G.; Minnocci, A.; Petacchi, R.; Pelosi, P. Purification, structural characterization, cloning and immunocytochemical localization of chemoreception proteins from *Schistocerca gregaria*. *Eur. J. Biochem.* **1999**, *262*, 745-754.
7. Picimbon, J.F.; Dietrich, K.; Angeli, S.; Scaloni, A.; Krieger, J.; Breer, H.; Pelosi, P. Purification and molecular cloning of chemosensory proteins from *Bombyx mori*. *Arch. Insect Biochem. Physiol.* **2000**, *44*, 120–12.
8. Picimbon, J.F.; Dietrich, K.; Breer, H.; Krieger, J. Chemosensory proteins of *Locusta migratoria* (Orthoptera: Acrididae). *Insect Biochem. Mol. Biol.* **2000**, *30*, 233–241.
9. Picimbon, J.F.; Dietrich, K.; Krieger, J.; Breer, H. Identity and expression pattern of chemosensory proteins in *Heliothis virescens* (Lepidoptera, Noctuidae). *Insect Biochem. Mol. Biol.* **2001**, *31*, 1173-1181.
10. Picimbon, J.F. Biochemistry and evolution of CSP and OBP proteins. In *Insect Pheromone Biochemistry and Molecular Biology, The* 05 *Biosynthesis and Detection of Pheromones and Plant Volatiles*; Blomquist, G.J., Vogt, R.G., Eds; Elsevier Academic Press; London, 06 UK; San Diego, CA, USA, 2003; pp. 539-566.
11. Wanner, K.W.; Isman, M.B.; Feng, Q.; Plettner, E.; Theilmann, D.A. Developmental expression patterns of four chemosensory protein genes from the Eastern spruce budworm, *Choristoneura fumiferana. Insect Mol. Biol.* **2005**, *14*, 289-300.
12. Truman, J.W.; Riddiford, L.M. The evolution of insect metamorphosis: a developmental and endocrine view. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **2019**, *374*, 20190070.
13. Haug, J.T. Metamorphosis in Crustaceans. In *Developmental Biology and Larval Ecology: The Natural History of the Crustacea, Volume 7*; Anger, K., Harzsch, S., Thiel, M., Eds; Oxford University Press, New York, NY, USA, 2020; pp. 254-283.
14. Picimbon, J.F. Molecular phylogeny of "Chemosensory Proteins" in bacteria and crustaceans: *CSP* as an extremely ancient gene. *J. Mol. Evol.* **2026,** *94*, 190-230.
15. Lartigue, A.; Campanacci, V.; Roussel, A.; Larsson, A.M.; Jones, T.A.; Tegoni, M.; Cambillau, C. X-ray structure and ligand binding study of a moth chemosensory protein. *J. Biol. Chem.* **2002**, *277*, 32094-32098.
16. Jansen, S.; Chmelik, J.; Zídek, L.; Padrta, P.; Novak, P.; Zdrahal, Z.; Picimbon, J.F.; Löfstedt, C.; Sklenar, V. Structure of *Bombyx mori* Chemosensory Protein 1 in solution. *Arch. Insect Biochem. Physiol.* **2007**, *66*, 135-145.
17. Tomaselli, S.; Crescenzi, O.; Sanfelice, D.; Ab, E.; Wechsel-Berger, R.; Angeli, S.; Scaloni, A.; Boelens, R.; Tancredi, T.; Pelosi, P.; et al. Solution structure of a chemosensory protein from the desert locust *Schistocerca gregaria*. *Biochemistry* **2006**, *45*, 1606-1613.
18. Jia, Q.; Zeng, H.; Zhang, J.; Gao, S.; Xiao, N.; Tang, J.; Dong, X.; Xie, W. The crystal structure of the *Spodoptera litura* Chemosensory Protein CSP8. *Insects* **2021**, *12*, 602.
19. Campanacci, V.; Lartigue, A.; Hällberg, BM.; Jones, T.A.; Giuici-Orticoni, M.T.; Tegoni, M.; Cambillau, C. Moth chemosensory protein exhibits drastic conformational changes and cooperativity on ligand binding. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 5069-5074.
20. Mosbah, A.; Campanacci, V.; Lartigue, A.; Tegoni, M.; Cambillau, C.; Darbon, H. Solution structure of a chemosensory protein from the moth *Mamestra brassicae*. *Biochem. J.* **2003**, *369*, 39-44.
21. Liu, G.; Ma, H.M.; Xie, Y.N.; Xuan, N.; Guo, X.; Fan, Z.X.; Rajashekar, B.; Arnaud, P.; Offmann, B.; Picimbon, J.F. Biotype characterization, developmental profiling, insecticide response and binding property of *Bemisia tabaci* chemosensory proteins: Role of CSP in insect defense. *PLoS ONE*, **2016**, *11*, e0154706.
22. Liu, G.; Xuan, N.; Rajashekar, B.; Arnaud, P.; Offmann, B.; Picimbon, J.F. Comprehensive history of CSP genes: Evolution, phylogenetic distribution, and functions. *Genes* **2020**, *11*, 413.
23. Weisiger, RA. Mechanisms of intracellular fatty acid transport: role of cytoplasmic-binding proteins. *J. Mol. Neurosci.* **2007**, 33, 42-44.
24. Bos, J.I.; Prince, D.; Pitino, M.; Maffei, M.E.; Win, J.; Hogenhout, S.A. A functional genomics approach identifies candidate effectors from the aphid species *Myzus persicae* (green peach aphid). *PLoS Genet.* **2010**, *6*, e1001210.
25. Rodriguez, P.A.; Stam, R.; Warbroek, T.; Bos, J.I. Mp10 and Mp42 from the aphid species *Myzus persicae* trigger plant defenses in *Nicotiana benthamiana* through different activities. *Mol. Plant. Microbe. Interact.* **2014**, *27*, 30-39.

 doi:10.20944/preprints202604.2167.v1

26. Moneo, I.; Vega, J.M.; Caballero, M.L.; Vega, J.; Alday, E. Isolation and characterization of Tha p 1, a major allergen from the pine processionary caterpillar *Thaumetopoea pityocampa*. *Allergy* **2003**, *58*, 34-37.
27. Bradford, M.K.; Whitworth, K.; Wendland, B. Pan1 regulates transitions between stages of clathrin-mediated endocytosis. *Mol Biol Cell.* **2015,** *26*, 1371-1385.
28. Liu, G.; Yue, S.; Rajashekar, B.; Picimbon, J.F. Expression of chemosensory protein (CSP) structures in *Pediculus humanus corporis* and *Acinetobacter baumannii*. *SOJ Microbiol. Infect. Dis.* **2019**, *7*, 1-17.
29. Levin, P.A.; Angert, E.R. Small but mighty: cell size and bacteria. *Cold Spring Harb. Perspect. Biol.* **2015**, *7*, a019216.
30. Jurtshuk, P. Bacterial metabolism; In Medical Microbiology 4th ed; Baron, S., Ed.; University of Texas Medical Branch at Galveston, Galveston, TX, 1996.; Chapter 4.
31. Yoshida, S.; Hiraga, K.; Takehana, T.; Taniguchi, I.; Yamaji, H.; Maeda, Y.; Toyohara, K.; Miyamoto, K.; Kimura, Y.; Oda, K. A bacterium that degrades and assimilates poly(ethylene teraphtalate). *Science* **2016**, *351*, 1196-1199.
32. Djurhuus, A.; Mikalsen, S.O.; Giebel, H.A.; Rogers, A.D. Cutting through the smoke: the diversity of microorganisms in deep-sea hydrothermal plumes. *R Soc Open Sci.* **2017**, *4*, 160829.
33. Kang N.K.; Koh, H.G.; Choi, Y.; Min, H.; Ort, D.R.; Jin, Y.S. Bioconversion of CO_2 into valuable bioproducts via synthetic modular co-culture of engineered *Chlamydomonas reinhardtii* and *Escherichia coli*. *Metab. Eng.* **2025,** *90*, 57-66.
34. Martinez-Guryn, K.; Hubert, N.; Frazier, K.; Urlass, S.; Musch, M.W.; Ojeda, P.; Pierre, J.F.; Miyoshi, J.; Sontag, T.J.; Cham, C.N.; et al. Small intestine microbiota regulate host digestive and absorptive adaptive responses to dietary lipids. *Cell Host Microbe* **2018**, *23*, 458-469.
35. Visser, B.; Scheifler, M. Insect lipid metabolism in the presence of symbiotic and pathogenic viruses and bacteria. *Adv Exp Med Biol.* **2026**, *1494*, 419-443.
36. Santos, T.C.B.; Futerman, A.H. The fats of the matter: lipids in prebiotic chemistry and in origin of life studies. *Prog Lipid Res.* **2023**, *92*, 101253.
37. Pavlowsky, A.; Silva, B.; Basu, R.; Correia Delecourt, A.; Geny, D.; Danglot, L.; Plaçais, P.Y.; Preat, T. Neuronal fatty acid oxidation fuels memory after intensive learning in *Drosophila*. *Nat. Metab.* **2025**, *7*, 2438–2450.
38. Brown, W.; Ralston, A.; Shaw, K. Positive transcription control: the glucose effect. *Nat. Education* **2008**, *1*, 202.
39. Jeckelmann, J.M.; Erni, B. Transporters of glucose and other carbohydrates in bacteria. *Pflugers Arch.* **2020**, *472*, 1129-1153.
40. Balmand, S.; Rivard, C.; Peignier, S.; Santarella-Mellwig, R.; Ghanem-Debbache, M.; Maire, J.; Engl, T.; Galvão Ferrarini, M.; Dell'Aglio, E.; Soriana-Saiz, B. Bacterial tubular networks channel carbohydrates in insect endosymbiosis. *Cell* **2025**, *188*, 7355-7365.
41. Wang, Y.; Yang, J.; Lee, O.O.; Dash, S.; Lau, S.C.; Al-Suwailem, A.; Wong, T.Y.; Danchin, A.; Qiang, P.Y. Hydrothermally generated aromatic compounds are consumed by bacteria colonizing in Atlantis II Deep of the Red Sea. *ISME J.* **2011**, *5*, 1652-1659.
42. Xu, H.; Wang, X.; Wei, Y.; Cao, Y.; Wang, S.; Xia, P. Pathway crosstalk enables degradation of aromatic compounds in marine *Roseobacter* clade bacteria. *Appl. Environ. Microbiol.* **2025**, 91, e00978-25.
43. Miller, M.B.; Bassler, B.L. Quorum sensing in bacteria. *Annu. Rev. Microbiol.* **2001**, *55*, 165-199.
44. Waters, C.M.; Bassler, B.L. The *Vibrio harveyi* quorum-sensing system uses shared regulatory components to discriminate between multiple autoinducers. *Genes Dev.* **2006**, *20*, 2754-2767.
45. Lerat, E.; Moran, N.A. The evolutionary history of quorum-sensing systems in bacteria. *Mol. Biol. Evol.* ***2004****, 21*, 903–913.
46. Coolahan, M.; Whalen, K.E. A review of quorum-sensing and its role in mediating interkingdom interactions in the ocean. *Commun. Biol.* **2025**, *8*, 179.
47. Monnet, V.; Gardan, R. Quorum-sensing regulators in Gram-positive bacteria: 'cherchez le peptide'. *Mol. Microbiol.* **2015**, *97*, 181-184.
48. Papenfort, K.; Bassler, B.L. Quorum sensing signal-response systems in Gram-negative bacteria. *Nat. Rev. Microbiol.* **2016**, *14*, 576-588.

 doi:10.20944/preprints202604.2167.v1

49. Jin, X.; Brandazza, A.; Navarrini, A.; Ban, L.; Zhang, S.; Steinbrecht, R.A.; Zhang, L.; Pelosi, P. Expression and immunolocalization of odorant-binding and chemosensory proteins in locusts. *Cell. Mol. Life Sci.* **2005**, *62*, 1156-1166.
50. Steinbrecht, R.A.; Ozaki, M.; Ziegelberger, G. Immunocytochemical localization of pheromone-binding protein in moth antennae. *Cell Tissue Res.* **1992**, *270*, 287-302.
51. Jimenez, L.V.; Kang, B.K.; deBruyn, B.; Lovin,D.D., Severson, D.W. Characterization of an *Aedes aegypti* bacterial artifical chromosome (BAC) library and chromosomal assignment of BAC clones for physical mapping quantitative trait loci that influence *Plasmodium* susceptibility. *Insect Mol. Biol.* **2004**, *13*, 37-44.
52. Loftus, B.J.; Utterback, T.; Pertea, G.; Koo, H.; Mori, A.; Schneider, J.; Lovin, D.; de Bruyn, B.; Song, Z.; Raikhel, A.; de Fatima, B.M.; Casavant, T.; Soares, B.; Severson, D. *Aedes aegypti* cDNA sequencing. *The Institute for Genomic Research TIGR* **2005**.
53. Noriega, F.G.; Ribeiro, J.M.C.; Koener, J.F.; Valenzuela, J.G.; Hernandez-Martinez, S.; Pham, V.M.; Feyereisen, R. Comparative genomics of insect juvenile hormone biosynthesis. *Insect Biochem. Mol. Biol.* **2006**, *36*, 366-374.
54. Nene, V.; Worthman, J.R.; Lawson, D.; Haas, B.; Kodira, C.; Tu, Z.; Loftus, B.; Xi, Z.; Megy, K.; Grabherr, M.; et al. Genome sequence of *Aedes aegypti*, a major arbovirus vector. *Science* **2007**, *22*, 1718-1723.
55. Sabatier, L.; Jouanguy, E.; Dostert, C.; Zachary, D.; Dimarcq, J.L.; Bulet, P.; Imler, J.C. Pherokine-2 and -3: Two *Drosophila* molecules related to pheromone/odor-binding proteins induced by viral and bacterial infections. *Eur. J. Biol.* **2003**, *270*, 3398-3407.
56. Liu, G. et al. Biotype expression and insecticide response of *Bemisia tabaci* chemosensory protein-1. *Arch. Insect Biochem. Physiol.* **2014**, *85*, 137-151.
57. Einhorn, E.; Imler, J.L. Insect immunity; from systemic to chemosensory organs protection. In *Olfactory Concepts of Insect Control-Alternative to Insecticides*; Picimbon, J.F., Ed.; Springer Nature AG: Cham, Switzerland, 2019; Volume 2, pp. 205–229.
58. Wei, Z.; Ortiz-Urquiza, A.; Keyhani, N.O. Altered expression of chemosensory and odorant binding proteins in response to fungal infection in the red imported fire ant, *Solenopsis invicta*. *Front. Physiol.* **2021**, *12*, 596571.
59. Qi, Z.; Gu, T.; Li, W.; Deng, J.; Zhang, Q.H.; Wickham, J.D.; Bai, J.; Zhang, L. Chemosensory proteins-mediated immune response to *Beauvaria bassiana* infection in *Monochamus alternatus*. *Insect Sci.* **2026**, https://doi.org/10.1111/1744-7917.70246.
60. Xu, H.; Yan, K.; Ding, Y.; Lv, Y.; Li, J.; Yang, F.; Chen, X.; Gao, X.; Pan, Y.; Shang, Q. Chemosensory proteins are associated with thiametoxam and spirotetramat tolerance in *Aphis gossypii* Glover. *Int. J. Mol. Sci.* **2022**, *23*, 2356.
61. Gao, P.; Zhang, S.; Tan, J.J.; Li, X.; Chen, M. Chemosensory proteins are associated with thiametoxam tolerance in bird cherry-oat aphid *Rhopalosiphum padi*. *Pestic. Biochem. Physiol.* **2023**, *192*, 105393.
62. Xu, H.; Pan, Y.; Li, J.; Yang, F.; Chen, X.; Gao, X.; Wen, S.; Shang, Q. Chemosensory proteins confer adaptation to the ryanoid anthranilic diamide insecticide cyantraniliprole in *Aphis gossypii* glover. *Pestic. Biochem. Physiol.* **2022**, *184*, 105076.
63. Kain, P.; Boyle, S.M.; Tharadra, S.K.; Guda, T.; Pham, C.; Dahanukar, A.; Ray, A. Odour receptors and neurons for DEET and new insect repellents. *Nature* **2013**, *502*, 507-512.
64. Duarte-Mata, D.I.; Salinas-Carmona, M.C. Antimicrobial peptides' immune modulation role in intracellular bacterial infection. *Front. Immunol.* **2023**, *14*, 1119574.
65. Wang, H.; Zhao, R.; Gao, J.; Xiao, X.; Yin, X.; Hu, S.; Zhang, Y.; Liang, P.; Gu, S. Two cuticle-enriched chemosensory proteins confer multi-insecticide resistance in *Spodoptera frugiperda*. *Int. J. Biol. Macromol.* **2024**, *266*, 130941.
66. Wu, H.; Xu, Q.; Ma, Y.; Zhao, X.; Zhang, J.; Zhang, Y. Two integument-enriched chemosensory proteins contribute to deltamethrin tolerance through cuticular binding in *Locusta migratoria*. *Pestic. Biochem. Physiol.* **2026**, *219*, 107028.
67. Yi, X.; Qi, J.; Zhou, X.; Hu, M.Y.; Zhong, G.H. Differential expression of chemosensory-protein genes in midguts in response to diet of *Spodoptera litura*. *Sci. Rep.* **2017**, *7*, 296.

 doi:10.20944/preprints202604.2167.v1

68. Kodaczuk, J.; Sulek, M.; Mak, P.; Fraczek, A.; Wojda, I. Chemosensory protein 16 has an immune function and participates in host-pathogen interaction in *Galleria mellonella* infected with *Pseudomonas entomophila*. *Virulence* **2025**, *16*, 2471367.
69. Verdú, J.R.; Cortez, V.; Ortiz, A.J.; Lumaret, J.P.; Lobo, J.M.; Sánchez-Piñero, F. Biomagnification and body distribution of ivermectin in dung beetles. *Sci. Rep.* **2020**, *10*, 9073.
70. Wang, T.; Lv, H.; Zheng, C.; Yang, C.; Huang, Y.; Li, X.; Li, J.; Ma, K. Overexpression of multiple odorant binding and chemosensory protein genes contributed to multi-insecticide resistance in *Aphis gossypii* Glover. *Ecotoxicol. Environ. Saf.* **2025**, *305*, 119210.
71. Wilkins, RM. Insecticide resistance and intracellular proteases. *Pest. Manag. Sci.* **2017**, *73*, 2403-2412.
72. Santovito, A.; Audisio, M.; Bonelli, S. A micronucleus assay detects genotoxic effects of herbicide exposure in a protected butterfly species. *Ecotoxicology* **2020**, 29, 1390-1398.
73. Chen, L.P.; Jiang, H.Q.; Luo, L.; Qiu, J.; Xing, X.J.; Hou, R.Y.; Wu, Y.J. The role of intercellular junction proteins in the penetration resistance of *Drosophila* larvae to avermectin. *Comp. Biochem. Physiol. C Toxicol. Pharmacol.* **2023**, *266*, 109557.
74. Tsouri, A.; Douris, V. The role of chemosensory proteins in insecticide resistance: A review. *Insects* **2025**, *16*, 496.
75. Denecke, S.; Swevers, L.; Douris, V.; Vontas, J. How do oral insecticidal compounds cross the insect midgut epithelium? *Insect Biochem. Mol. Biol.* **2018**, *103*, 22-35.
76. Li, F.; Venthur, H.; Wang, S.; Homem, R.A.; Zhou, J.J. Evidence for the involvement of the chemosensory protein AgosCSP5 in resistance to insecticides in the cotton aphid, *Aphis gossypii*. *Insects* **2021**, *12*, 335.
77. Peng, X.; Qu, M.J.; Wang, S.J.; Huang, Y.X.; Chen, C.; Chen, M.H. Chemosensory proteins participate in insecticide susceptibility in *Rhopalosiphum padi*, a serious pest on wheat crops. *Insect Mol. Biol.* **2021**, *30*, 138-151.
78. Li, Y.; Ni, S.; Wang, Y.; Li, R.; Sun, H.; Ye, X.; Tian, Z.; Zhang, Y.; Liu, J. The chemosensory protein 1 contributes to indoxacarb resistance in *Plutella xylostella* (L.). *Pest Manag. Sci.* **2023**, *79*, 2456-2459.
79. Yao, Q.; Liang, Z.; Chen, B. Evidence for the participation of chemosensory proteins in response to insecticide challenge in *Cnopomorpha sinensis*. *J. Agric. Food Chem.* **2023**, *71*, 1360-1368.
80. Yao, Y.J.; Yin, N.N.; Pu, L.M.; Yang, A.J.; Liu, N.Y. Three chemosensory proteins enriched in antennae and tarsi of *Rhaphuma horsfieldi* differentially contribute to the binding of insecticides. *Pestic. Biochem. Physiol.* **2024**, *199*, 105797.
81. Li, F.; Venthur, H.; Lin, K.; Zhang, C.; Chen, Z.; Zhou, J.J. Insecticide chemosensory proteins as targets in insecticide resistance and development. *New Plant Prot.* **2025**, *2*, e70008.
82. Volonté, M.; Traverso, L.; Sierra, I.; Aptekmann, A.A.; Nadra, A.D.; Ons, S. Characterization of odorant binding and chemosensory protein families in the kissing bug *Triatoma infestans*: comparative analysis among Heteroptera species. *BMC Genomics* **2025**, *26*, 794.
83. Wang, Q.; Shentu, X.; Yu, X.; Liu, Y. Insect odorant-binding proteins (OBPs) and chemosensory proteins (CSPs): mechanisms and research perspectives in mediating insecticide resistance. *Biology* **2025**, *14*, 1452.
84. Frkic, R.L.; Giang, A.; Pulsford, S.B.; Liu, J.W.; Esmaeily, M.; Carr, P.D.; Fraser, N.J.; Hopkins, D.; Oakeshott, J.G.; Batterham, P.; et al. Structural and evolutionary constraints of organophosphate resistance in dipteran carboxylesterases. *Proc. Natl. Acad. Sci. USA* **2026**, *123*, e2517957123.
85. Picimbon, J.F. Chapter three—Bioinformatic, genomic and evolutionary analysis of genes: A case study in Dipteran CSPs. *Methods Enzymol.* **2020**, *642*, 35-79.
86. Auradkar, A.; Corder, R.M.; Marshall, J.M.; Bier, E. A self-eliminating allelic-drive reverses insecticide resistance in *Drosophila* leaving no transgene in the population. *Nat. Commun.* **2024**, *15*, 9961.
87. Johnson, E.C.; Garczynski, S.F.; Park, D.; Crim, J.; Nässel, D.R.; Taghert, P.H. Identification and characterization of a G protein-coupled receptor for the neuropeptide proctolin in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 6198-6203.
88. Bian, H.X.; Chen, B.D.; Zheng, X.X.; Ma, H.F.; Li, Y.P.; Li, Q.; Xia, R.X.; Wang, H.; Jiang, Y.R.; Liu, Y.Q.; et al. Transcriptomic analysis of the prothoracic gland from two lepidopteran insects, domesticated silkmoth *Bombyx mori* and wild silkmoth *Antheraea pernyi*. *Sci. Rep.* **2019**, *9*, 5313.

 doi:10.20944/preprints202604.2167.v1

89. Flatt, T.; Heyland, A.; Rus, F.; Porpiglia, E.; Sherlock, C.; Yamamoto, R.; Garbuzov, A.; Palli, S.R.; Tatar, M.; Silverman, N. Hormonal regulation of the humoral innate immune response in *Drosophila melanogaster*. *J. Exp. Biol.* **2008**, *211*, 2712-2724.
90. Ji, Y.; Gao, B.; Zhao, D.; Zhang, L.; Wu, H.; Xie, Y.; Shi, Q.; Wang, Y.; Guo, W. The role of 20-hydroxyecdysone and juvenile hormone in insecticidal activity of *Bacillus thuringiensis* regulated by DUOX-ROS immunity in *Spodoptera exigua*. *Pestic. Biochem. Physiol.* **2025**, *208*, 106222.
91. Wang, X.; Hou, Y.; Saha, T.T.; Pei, G.; Raikhel, A.S.; Zou, Z. Hormone and receptor interplay in the regulation of mosquito lipid metabolism. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, E2709-E2718.
92. Chen, C.H.; Pan, J.; Di, Y.Q.; Liu, W.; Hou, L.; Wang, J.X.; Zhao, X.F. Protein kinase C delta phosphorylates ecdysone receptor B1 to promote gene expression and apoptosis under 20-hydroxyecdysone regulation. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, E7121-E7130.
93. Liu, P.; Peng, H.J.; Zhu, J. Juvenile hormone-activated phospholipase C pathway enhances transcriptional activation by the methoprene-tolerant protein. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, E1871-E1879.
94. Ojani, R.; Liu, P.; Fu, X.; Zhu, J. Protein kinase C modulates transcriptional activation by the juvenile hormone receptor methoprene-tolerant. *Insect Biochem. Mol. Biol.* **2016**, *70*, 44-52.
95. Gao, Y.; Su, S.C.; Xing, J.Y.; Liu, Z.Y.; Nässel, D.R.; Bass, C.; Gao, C.; Wu, S.F. Pesticide-induced resurgence in brown planthoppers is mediated by action on a suite of genes that promote juvenile hormone biosynthesis and female fecundity. *Elife* **2025**, *12*, RP91774.
96. Noriega, F.G. Juvenile hormone biosynthesis in insects: what is new, what do we know, and what questions remain? *Int. Sch. Res. Notices* **2014,** *2014*, 967361.
97. Kim, I.H.; Pham, V.; Jablonka, W.; Goodman, W.G.; Ribeiro, J.M.C.; Andersen, J.F. A mosquito hemolymph odorant-binding protein family member specifically binds juvenile hormone. *J. Biol. Chem.* **2017**, *292*, 15329–15339.
98. Ahmed, S.; Hrithik, M.T.H.; Roy, M.C.; Bode, H.; Kim, Y. Phurealipids, produced by the entomopathogenic bacteria, *Photorhabdus*, mimic juvenile hormone to suppress insect immunity and immature development. *J. Invertebr. Pathol.* **2022**, *193*, 107799.
99. Ozaki, M.; Wada-Katsumata, A.; Fujikawa, K.; Iwasaki, M.; Yokohari, F.; Satoji, Y.; Nisimura, T.; Yamaoka, R. Ant nestmate and non-nestmate discrimination by a chemosensory sensillum. *Science* **2005**, *309*, 311-314.
100. Santos, A.L., Preta, G. Lipids in the cell: organisation regulates function. *Cell. Mol. Life Sci.* **2018**, *75*, 1909-1927.
101. Hasan, M.A.; Ahmed, S.; Kim, Y. Biosynthetic pathway of arachidonic acid in *Spodoptera exigua* in response to bacterial challenge. *Insect Biochem. Mol. Biol.* **2019**, *111*, 103179.
102. Vatanparast, M.; Ahmed, S.; Lee, D.H.; Hwang, S.H.; Hammock, B.; Kim, Y. EpOMEs act as immune suppressors in a lepidoptera insect, *Spodoptera exigua*. *Sci. Rep.* **2020**, *10*, 20183.
103. Malcicka, M.; Visser, B.; Ellers, J. An evolutionary perspective on linoleic acid synthesis in animals. *Evol. Biol.* **2018**, *45*, 15–26.
104. Reddy, G.V.; Guerrero, A. Interactions of insect pheromones and plant semiochemicals. *Trends Plant Sci.* **2004**, *9*, 253-261.
105. Wang, F.E.; Delannay, C.; Goindin, D.; Deng, L.; Guan, S.; Lu, X.; Fouque, F.; Vega-Rúa, A.; Picimbon, J.F. Cartography of odor chemicals in the dengue vector mosquito (*Aedes aegypti* L., Diptera/Culicidae). *Sci. Rep.* **2019**, *9*, 8510.
106. Ebrahim, S.A.M.; Dweck, H.K.M.; Weiss, B.L.; Carlson, J.R. A volatile sex attractant of tsetse flies. *Science* **2023**, *379*, eade1877.
107. Herrera, H.; Barros-Parada, W.; Bergmann, J. Linoleic acid and stearic acid are biosynthetic precursors of (7Z,10Z)-7,10-hexadecadienal, the major component of the sex pheromone of *Chilecomadia valdiviana* (Lepidoptera: Cossidae). *PLoS ONE* **2019**, *14*, e0215769.
108. Shibao, H.; Kutsukake, M.; Matsuyama, S.; Fukatsu, T. Linoleic acid as corpse recognition signal in a social aphid. *Zoological Lett.* **2022**, *8*, 2.
109. Buehlmann, C.; Graham, P.; Hansson, B.S.; Knaden, M. Desert ants locate food by combining high sensitivity to food odors with extensive crosswind runs. *Curr. Biol.* **2014**, *24*, 960-964.

 doi:10.20944/preprints202604.2167.v1

110. Yin, C.; Shan, G.; Guo, D.; Wang, S.; Ma, X.; Xiao, H.; Liu, J.; Zhang, Z.; Liu, Y.; Zhang, Y.; et al. InsectBase: A resource for insect genomes and transcriptomes. *Nucl. Acids Res.* **2016**, *44*, D801-D807.
111. Swofford, D.L. *PAUP*. Phylogenetic Analysis Using Parsimony (* and Other Methods)*, Version 4.0b10; Sinauer Associates: Sunderland, MA, USA, 2002.
112. Sáez, A.; Morales, C.L.; Ramos, L.Y.; Aizen, M.A. Extremely frequent bee visits increase pollen decomposition but reduce drupelet set in raspberry. *J. Appl. Ecol.* **2014**, *51*, 1603-1612.
113. Jacobsen, D.J.; Raguso, R.A. Lingering effects of herbivory and plant defenses on pollinators. *Curr. Biol.* **2018**, *28*, R11–R69.
114. Hu, S.; Dilcher, D.L.; Jarzen, D.M.; Winship Taylor, D. Early steps of angiosperm pollinator coevolution. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 240-245.
115. Zeng, Y.; Merchant, A.; Wu, Q.; Wang, S.; Kong, L.; Zhou, X.; Xie, W.; Zhang, Y. A chemosensory protein BtabCSP11 mediates reproduction in *Bemisia tabaci*. *Front. Physiol.* **2020**, *11*, 709.
116. Nguyen, L.T.; Schmidt, H.A.; von Haeseler, A.; Minh, B.Q. IQ-TREE: A fast and effective stochastic algorithm for estimating maximum likelihood phylogenies. *Mol. Biol. Evol.* **2015**, *32*, 268-274.
117. Moon, S.Y.; Zheng, Y. Rho GTPase-activating proteins in cell regulation. *Trends Cell Biol.* **2003**, *13*, P13-P22.
118. Gouin, E.; Egile, C.; Dehoux, P.; Villiers, V.; Adams, J.; Gertler, F.; Li, R.; Cossart, P. The RickA protein of *Rickettsia conorii* activates the Arp2/3 complex. *Nature* **2004**, *427*, 457-461.
119. Cronmiller, E.; Toor, D.; Shao, N.C.; Kariyawasam, T.; Wang, M.H.; Lee, J.H. Cell wall integrity signaling regulates cell wall- related gene expression in *Chlamydomonas reinhardtii*. *Sci. Rep.* **2019**, *9*, 12204.
120. Boissan, M.; Schlattner, U.; Lacombe, M.L. The NDPK/NME superfamily: state of the art. *Lab. Invest.* **2018**, *98*, 164-174.
121. Sridharan Iyer, S.; Wu, J.; Pollard, T.D.; Voth, G.A. Molecular mechanism of Arp2/3 complex activation by nucleation-promoting factors and an actin monomer. *Proc. Natl. Acad. Sci. USA* **2025**, *122*, e2421467122.
122. Ho, S. The molecular clock and estimating species divergence. *Nat. Educ.* **2008**, *1*, 168.
123. Ribeiro, J.M.; Martin-Martin, I.; Arcà, B.; Calvo, E. A Deep insight into the sialome of male and female *Aedes aegypti* mosquitoes. *PLoS ONE*, **2016**, *11*, e0151400.
124. Yadav, K.; Rana, V.S.; Anjali; Saurav, G.K.; Rawat, N.; Kumar, A.; Sunil, S.; Singh, O.P.; Rajagopal, R. Mucin protein of *Aedes aegypti* interacts with Dengue Virus 2 and influences viral infection. *Microbiol. Spectr.* **2023**, *11*, e02503-22.
125. Schachat, S.R.; Goldstein, P.Z.; Desalle, R.; Bobo, D.M.; Boyce, K.; Payne, J.L.; Labandeira, C.C. Illusion of flight? Absence, evidence and the age of winged insects. *Biol. J. Linn. Soc.* **2023**, *138*, 143-168.
126. Cornman, R.S. The distribution of GYR- and YLP-like motifs in *Drosophila* suggests a general role in cuticle assembly and other protein-protein interactions. *PLoS ONE*, **2010**, *5*, e12536.
127. Teparić, R.; Lozančić, M.; Mrša, V. Evolutionary overview of molecular interactions and enzymatic activities in the yeast cell walls. *Int. J. Mol. Sci.* **2020**, *21*, 8996.
128. Adachi, N.; Senda, T.; Horikoshi, M. Uncovering ancient transcription systems with a novel evolutionary indicator. *Sci. Rep.* **2016**, *6*, 217922.
129. Akil, C.; Kitaoku, Y.; Tran, L.T.; Liebl, D.; Choe, H.; Muengsaen, D.; Suginta, W.; Schulte, A.; Robinson, R.C. Mythical origins of the actin cytoskeleton. *Curr. Opin. Cell Biol.* **2021**, *68*, 55–63.
130. Rodriguez-Oliveira, T.; Wollweber, F.; Ponce-Toledo, R.I.; Xu, J.; Rittmann, S.K.-M.R.; Klingl, A.; Pilhofer, M.; Schleper, C. Actin cytoskeleton and complex cell architecture in an *Asgard archaeon*. *Nature* **2023**, *613*, 332-339.
131. Charles-Orszag, A.; Petek-Seoane, N.A.; Mullins, R.D. Archaeal actins and the origin of a multifunctional cytoskeleton. *J. Bacteriol.* **2024**, *206*, e0034823.
132. Waterhouse, A.; Bertoni, M.; Bienert, S.; Studer, G.; Tauriello, G.; Gumienny, R.; Heer, F.T.; de Beer, T.A.P.; Rempfer, C.; Bordoli, L.; et al. SWISS-MODEL: Homology modelling of protein structures and complexes. *Nucleic Acids Res.* **2018**, *46*, W296–W303.
133. Ho, H.-Y.H.; Rohatgi, R.; Ma, L.; Kirschner, M.W. CR16 forms a complex with N-WASP in brain and is a novel member of a conserved proline-rich actin-binding protein family. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 11306-11311.

 doi:10.20944/preprints202604.2167.v1

134. Chereau, D.; Kerff, F.; Graceffa, P.; Grabarek, Z.; Langsetmo, K.; Dominguez, R. Actin-bound structures of Wiskott-Aldrich syndrome protein (WASP)-homology domain 2 and the implications for filament assembly. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 16644-16649.
135. Rogerio, A.P.; Anibal, F.F. Role of leukotrienes on protozoan and helminth infections. *Mediators Inflamm.* **2012**, *2012*, 595694.
136. Stanley, D.; Kim, Y. Chapter Eight—Insect prostaglandins and other eicosanoids: From molecular to physiological actions. *Adv. Insect Physiol.* **2019**, *56*, 283–343.
137. Roy, M.C.; Nam, K.; Kim, J.; Stanley, D.; Kim, Y. Thromboxane mobilizes insect blood cells to infection foci. *Front. Immunol.* **2021**, *12*, 791319.
138. El-Yassimi, A.; Hichami, A.; Besnard, P.; Khan, N.A. Linoleic acid induces calcium signaling, Src kinase phosphorylation, and neurotransmitter release in mouse CD36-positive gustatory cells. *J. Biol. Chem.* **2008**, *283*, 12949–12959.
139. Feng, D.D.; Zhao, Y.F.; Luo, Z.Q.; Keating, D.J.; Chen, C. Linoleic acid induces Ca2+-induced inactivation of voltage-dependent Ca2+ currents in rat pancreatic β-cells. *J. Endocrinol.* **2008**, *196*, 377–384.
140. Clements, J.; Groves, R.L.; Cava, J.; Barry, C.C.; Chapman, S.; Olson, J.M. Conjugated linoleic acid as a novel insecticide targeting the agricultural pest *Leptinotarsa decemlineata*. *PLoS ONE* **2019**, *14*, e0220830.
141. Macháček, S.; Tupec, M.; Horáček, N.; Halmová, M.; Roy, A.; Machara, A.; Kyjaková, P.; Lukšan, O.; Pichová, I.; Hanus, R. Evolution of linoleic acid biosynthesis paved the way for ecological success of termites. *Mol. Biol. Evol.* **2023**, *40*, msad087.
142. Carpentier, J.; Abenaim, L.; Luttenschlager, H.; Dessauvages, K.; Liu, Y.; Samoah, P.; Francis, F.; Caparros Megido, R. Microorganism contribution to mass-reared edible insects: opportunities and challenges. *Insects* **2024**, *15*, 611.
143. Arien, Y.; Dag, A.; Zarchin, S.; Masci, T.; Shafir, S. Omega-3 deficiency impairs honey bee learning. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 15761-15766.
144. Mercola, J. Linoleic acid, mitochondria, gut microbiome, and metabolic health: a mechanistic review. *Adv. Redox Res.* **2025**, *15*, 100128.
145. Giblin, A. et al. Neuronal polyunsaturated fatty acids are protective in ALS/FTD. *Nat. Neurosci.* **2025**, *28*, 737-747.
146. Marmunti, M.; Catalá, A. Incorporation of 1-(14)C linoleic acid in rat liver nuclei and chromatin fractions. *Int. J. Biochem. Cell Biol.* **2001**, *33*, 261–267.
147. Ando, H.; Wen, Z.M.; Kim, H.Y.; Valencia, J.C.; Costin, G.E.; Watabe, H.; Yasumoto, K.; Niki, Y.; Kondoh, H.; Ichihashi, M.; et al. Intracellular composition of fatty acid affects the processing and function of tyrosinase through the ubiquitin-proteasome pathway. *Biochem. J.* **2006**, *394*, 43–50.
148. Masner, M.; Lujea, N.; Bisbal, M.; Acosta, C.; Kunda, P. Linoleic and oleic acids enhance cell migration by altering the dynamics of microtubules and the remodeling of the actin cytoskeleton at the leading edge. *Sci. Rep.* **2021**, *11*, 14984.
149. Savini, M.; Folick, A.; Lee, Y.T.; Jin, F.; Cuevas, A.; Tillman, M.C.; Duffy, J.D.; Zhao, Q.; Neve, I.A.; Hu, P.W.; Yu, Y.; Zhang, Q.; Ye, Y.; Mair, W.B.; Wang, J.; Han, L.; Ortlund, E.A.; Wang, M.C. Lysosome lipid signalling from the periphery to neurons regulates longevity. *Nat. Cell. Biol.* **2022**, *24*, 906-916.
150. Vesala, L.; Basikhina, Y.; Tuomela, T.; Nurminen, A.; Siukola, E.; Vale, P.F.; Salminen, T.S. Mitochondrial perturbation in immune cells enhances cell-mediated innate immunity in *Drosophila*. *BMC Biol.* **2024**, *22*, 60.
151. Wrońska, A.K.; Kaczmarek, A.; Boguś, M.I.; Kuna, A. Lipids as a key element of insect defense systems. *Front. Genet.* **2023**, *9*, 1183659.
152. Parsons, J.B.; Rock, C.O. Bacterial lipids: metabolism and membrane homeostasis. *Prog. Lipid Res.* **2013**, *52*, 249-276.
153. Atilla-Gokcumen, G.E.; Muro, E.; Relat-Goberna, J.; Sasse, S.; Bedigian, A.; Coughlin, M.L.; Garcia-Manyes, S.; Eggert, U.S. Dividing cells regulate their lipid composition and localization. *Cell* **2014**, *156*, 428-39.
154. Blank, H.M.; Perez, R.; He, C.; Maitra, N.; Metz, R.; Hill, J.; Lin, Y.; Johnson, C.D.; Bankaitis, V.A.; Kennedy, B.K.; Aramayo, R.; Polymenis, M. Translational control of lipogenic enzymes in the cell cycle of synchronous, growing yeast cells. *EMBO J.* **2017**, *36*, 487-502.

 doi:10.20944/preprints202604.2167.v1

155. Yuan, Y.; Jin, M.; Fang, F.; Tocher, D.R.; Betancor, M.B.; Jiao, L.; Hong, Y.; Zhou, Q. New Insight into the molting and growth in crustaceans: regulation of energy homeostasis through the lipid nutrition. *Front. Mar. Sci.* **2022**, *9*, 914590.
156. Feyereisen, R. Insect P450 enzymes. *Annu. Rev. Entomol.* **1999**, *44*, 507–533.
157. Fujinaga, D.; Ohhara, Y.; Okamoto, N.; Chu, H.; Mauck, K.E.; Yamanaka, N. Two cytochrome P450 epoxidases mediate juvenile hormone biosynthesis in *Drosophila melanogaster*. *Insect Biochem Mol Biol.* **2026**, *187*, 104475.
158. Yokoyama, N.; Fónagy, A.; Tatsuki, S.; Arie, T.; Yamashita, S.; Matsumoto, S. Ultrastructural studies on the pheromone producing cells in the silkmoth, *Bombyx mori*: Formation of cytoplasmic lipid droplets before adult eclosion. *Acta Biol. Hung.* **2003**, *54*, 299–311.
159. Xuan, N.; Bu, X.; Liu, Y.Y.; Yang, X.; Liu, G.X.; Fan, Z.X.; Bi, Y.P.; Yang, L.Q.; Lu, Q.N.; Rajashekar, B.; et al. Molecular evidence of RNA editing in the *Bombyx* chemosensory protein family. *PLoS ONE* **2014**, *9*, e86932.
160. Hull, J.; Fónagy, A. Molecular basis of pheromonogenesis regulation in moths. In *Olfactory Concepts of Insect Control-Alternative to Insecticides*; Picimbon, J.F., Ed.; Springer Nature AG: Cham, Switzerland, 2019; Volume 1, pp. 115–202.
161. Ibarra, A.; Hetzer, M.W. Nuclear pore proteins and the control of genome functions. *Genes Dev.* **2015**, *29*, 337–349.
162. Lin, D.H.; Hoelz, A. The structure of the Nuclear Pore Complex (an update). *Annu. Rev. Biochem.* **2019**, *88*, 725–783.
163. Thelen, A.M.; Zoncu, R. Emerging roles for the lysosome in lipid metabolism. *Trends Cell Biol.* **2017,** *27*, 833-850.
164. Kenis, S.; Istiban, M.N.; Van Damme, S.; Vandewyer, E.; Watteyne, J.; Schoofs, L.; Beets, I. Ancestral glycoprotein hormone receptor pathway controls growth in *C. elegans*. *Front. Endocrinol.* **2023**, *14*, 1200407.
165. Fisher, W.W.; Hemp, J.; Johnson, J.E. Evolution of oxygenic photosynthesis. *Annu. Rev. Earth Planet. Sci.* **2016**, *44*, 647-663.
166. Dermauw, W.; Van Leeuwen, T.; Feyereisen, R. Diversity and evolution of the P450 family in arthropods. *Insect Biochem. Mol. Biol.* **2020**, *127*, 103490.
167. Lang, T.; Hansson, G.C.; Samuelsson, T. Gel-forming mucins appeared early in metazoan evolution. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 16209-16214.
168. Lang, T.; Klasson, S.; Larsson, E.; Johansson, M.E.; Hansson, G.C.; Samuelsson, T. Searching the evolutionary origin of epithelial mucus protein components-Mucin and FCGBP. *Mol. Biol. Evol.* **2016**, *33*, 1921-1936.
169. McShane, A.; Bath, J.; Jaramillo, A.M.; Ridley, C.; Walsh, A.A.; Evans, C.M.; Thornton, D.J.; Ribbeck, K. Mucus. *Curr. Biol.* **2021**, *31*, R938-R945.
170. Syed, Z.A.; Hard, T.; Uv, A.; van Dijk-Hard, I.F. A potential role for *Drosophila* mucins in development and physiology. *PLoS ONE* **2008**, *3*, e3041.
171. Zhao, X.; Zhang, J.; Yang, J.; Niu, N.; Zhang, J.; Yang, Q. Mucin family genes are essential for the growth and development of the migratory locust, *Locusta migratoria*. *Insect Biochem. Mol. Biol.* **2020**, *123*, 103404.
172. Huang, Y.; Li, L.; Rong, Y.S. JiangShi (僵尸): A widely distributed Mucin-like protein essential for *Drosophila* development. *G3 (Bethesda)* **2022**, *12*, jkac126.
173. Lou, Y.H.; Shen, Y.; Li, D.T.; Huang, H.J.; Lu, J.B.; Zhang, C.X. A mucin-like protein is essential for oviposition in *Nilaparvata lugens*. *Front. Physiol.* **2019**, *10*, 551.
174. Ibrahim, S.P.; Dias, R.O.; Ferreira, C.; Silva, C.P.; Terra, W.R. Histochemistry and transcriptomics of mucins and peritrophic membrane (PM) proteins along the midgut of a beetle with incomplete PM and their complementary function. *Insect Biochem. Mol. Biol.* **2023**, *162*, 104027.
175. Moriyama, M.; Hayashi, T.; Fukatsu, T. A mucin protein predominantly expressed in the female-specific symbiotic organ of the stinkbug *Plautia stali*. *Sci. Rep.* **2022**, *12*, 7782.
176. Riley, N.M.; Wen, R.M.; Bertozzi, C.R.; Brooks, J.D.; Pitteri, S.J. Measuring the multifaceted roles of mucin-domain glycoproteins in cancer. *Adv. Cancer Res.* **2023**, *157*, 83-121.
177. Yue, S.; He, Q.; Picimbon, J.F. Lactobacillus for ribosome peptide editing cancer. *Clin. Transl. Oncol.* **2023**, *25*, 1522-1544.

 doi:10.20944/preprints202604.2167.v1

178. Deng, F.; Wu, S.; Wu, Y.; Liu, X.; Wu, P.; Zhai, Z. Identification of mucins and their expression in the vector mosquito *Aedes albopictus*. *J. Vector Ecol.* **2020**, *45*, 297–305.
179. Taparia, T.; Ignell, R.; Hill, S.R. Blood meal induced regulation of the chemosensory gene repertoire in the southern house mosquito. *BMC Genomics* **2017**, *18*, 393.
180. Hill, S.R.; Ghaninia, M.; Ignell, R. Blood meal induced regulation of gene expression in the maxillary palps, a chemosensory organ of the mosquito *Aedes aegypti*. *Front. Ecol. Evol.* **2019**, *7*, 336.
181. Shangguan, X.; Zhang, J.; Liu, B.; Zhao, Y.; Wang, H.; Wang, Z.; Guo, J.; Rao, W.; Jing, S.; Guan, W.; et al. A mucin-like protein of planthopper is required for feeding and induces immunity responses in plants. *Plant Physiol.* **2018**, *176*, 552–565.
182. Slomiany, B.L., Murty, V.L.; Sarosiek, J.; Piotrowski, J.; Slomiany, A. Role of associated and covalently bound lipids in salivary mucin hydrophobicity: effect of proteolysis and disulfide bridge reduction. *Biochem. Biophys. Res. Commun.* **1988**, *151*, 1046-1053.
183. Gong, D.H.; Turner, B.; Bhaskar, K.R.; Lamont, J.T. Lipid binding to gastric mucin: Protective effect against oxygen radicals. *Am. J. Physiol.* **1990**, *259*, G681-G686.
184. Bansil, R.; Turner, B.S. The biology of mucus: Composition, synthesis and organization. *Adv. Drug. Deliv. Rev.* **2018**, *124*, 3-15.
185. Terra, W.R.; Dias, R.O.; Oliveira, P.L.; Ferreira, C.; Venancio, T.M. Transcriptomic analyses uncover emerging roles of mucins, lysosome/secretory addressing and detoxification pathways in insect midguts. *Curr. Opin. Insect Sci.* **2018,** *29*, 34-40.
186. Vial, T.; Marti, G.; Missé, D.; Pompon, J. Lipid interactions between flaviviruses and mosquito vectors. *Front. Physiol.* **2021**, *12*, 763195.
187. Picimbon, J.F. Insect chemosensory proteins against COVID-19. *Med. Discov*. **2024**, *3*, 1152.
188. Campanacci, V.; Krieger, J.; Bette, S.; Sturgis, J.N.; Lartigue, A.; Cambillau, C.; Breer, H.; Tegoni, M. Revisiting the specificity of *Mamestra brassicae* and *Antheraea polyphemus* pheromone-binding proteins with a fluorescence binding assay. *J. Biol. Chem.* **2001**, *276*, 20078-20084.
189. Chang, Y.; McClements, D.J. Characterization of mucin-lipid droplet interactions: Influence on potential fate of fish oil-in-water emulsions under simulated gastrointestinal conditions. *Food Hydrocoll.* **2016**, *56*, 425-433.
190. Carlson, T.L.; Yildiz, H.; Dar, Z.; Lock, J.Y.; Carrier, R.L. Lipids alter microbial transport through intestinal mucus. *PLoS ONE* **2018**, *13*, e0209151.
191. Hansson, G.C. Mucins and the microbiome. *Annu. Rev. Biochem.* **2020**, *89*, 769-793.
192. Fu, L.; Wang, M.; Li, D.; Ma, S.; Zhang, F.; Zheng, L. Microbial metabolites short chain fatty acids, tight junction, gap junction, and reproduction: a review. *Front. Cell Dev. Biol.* **2025**, *13*, 1624415.