

Review

Not peer-reviewed version

Renaming ‘Chemosensory’ Proteins (CSPs): ‘Lipoid-Binding Proteins’ – Molecular Nomenclature, Structure, Expression, Evolution, and Intracellular Functions

Guoxia Liu , [Botong Sun](#) , Wei Fan , Shousong Yue , [Qiuxia He](#) , [Jean-François Picimbon](#) *

Posted Date: 4 March 2024

doi: 10.20944/preprints202403.0173.v1

Keywords: Chemosensory proteins; 4 Cysteines Soluble Proteins; mucin; actin skeleton regulatory complex; nuclear pore complex protein; transcription initiation factor; gene regulation; immune response.



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Review

Renaming ‘Chemosensory’ Proteins (CSPs): ‘Lipoid-binding Proteins’ – Molecular Nomenclature, Structure, Expression, Evolution, and Intracellular Functions

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

¹ Biotechnology Research Center, Shandong Academy of Agricultural Sciences, Jinan, Shandong, 250100, Shandong, 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, Shandong, China

³ Shandong SME Productivity Promotion Center, Jinan, 250103, Shandong, China

* Correspondence: High Level Talent (R), National Expert of China, STSP-SDAS, Shandong Academy of Sciences Foreign Students Pioneer Park, Jinan, Shandong, China; jfpicimbon@163.com, jfpicimbon@gmail.com or jfpicimbon@sdas.org; Tel.: +00-86-17661098959

Abstract: This is a brief critique of the functions—particularly olfactory functions—specified for the “Chemosensory Protein” (CSPs) molecule family. On the basis of these proteins’ presence in the sensory antennal lymph of locusts, odor chemosensory ligand binding functions have been hypothesized. According to this hypothesis, the entire protein molecule superfamily is referred to as “CSPs”. However, new information and developments in the field of CSP molecular research, such as the expression of CSP genes in the gut, brain, fat body, wings, epidermis, and pheromone gland, as well as gene expression profiling from most early developmental stages—that is, CSP expression well in advance of the appearance of chemical sense nerve cells—strongly suggest that the protein molecule has other roles that are unrelated to chemosensing. Moreover, CSPs are found in bacterial microbial prokaryote organisms in addition to insects and all other arthropods. Thus, we examine the molecule’s name, definition, RNA editing, protein structure, lipid binding properties, DNA interaction, and evolutionary characteristics in brief before referring to this protein family as “Chemosensory Proteins”. As a means of renaming this protein family, this review article discusses the latest findings (“CSP” is attached to the tail of bigger intracellular proteins) and attempts to compel all the data. Because of its highly conserved molecular distinctive feature (four adjacent cysteines), we propose renaming “CSPs” as “4CSPs” (*4 Cysteines Soluble Proteins*) and use at least for abbreviations for distinctive animal proteins the common peptide 5-letter code, such as “Bommo-4CSPs”.

Keywords: chemosensory proteins; 4 cysteines soluble proteins; mucin; actin skeleton regulatory complex; nuclear pore complex protein; transcription initiation factor; gene regulation; immune response

1. Introduction

A few years ago, there was a case regarding the reevaluation and/or renaming of allatostatins due to their pleiotropic properties, which go beyond their inhibition of juvenile hormone (JH) biosynthesis in the *corpora allata* [1]. This is another instance of an invertebrate peptide family being named for a highly contested function (e.g. chemosensing), with the additional information that several cutting-edge investigations, fresh findings regarding evolution, molecule expression,

molecule localization, and binding characteristics, as well as two decades more of documented research, strongly imply a different function that goes well beyond insect olfaction [2].

The term “CSPs” stands for “Chemosensory Proteins”. They typically refer to small, water-soluble binding proteins, also known as odor-binding proteins (OBPs), that are strongly thought to mediate the recognition of odor molecules, odorants, and ligands to olfactory receptors (ORs), at the periphery of sensory dendrites in the insect *sensillum* [3–6]. According to Lartigue et al. [7], CSPs are composed of six α -helices with an appropriate molecular weight of 10–12 kDa (or 110–120 amino acid residues), four cysteines that form two tiny loops, two nearby disulfide bridges, and a globular “prism-like” functional structure. Four CSP structures have so far been identified in locusts (*Schistocerca gregaria*) and moths (*Mamestra brassicae*, *Bombyx mori*, and *Spodoptera litura*) [7–11]. These structures can breathe’ or modify their conformation in certain ways in response to ligand binding [12,13]. However, RNA mutations and mutations of the peptide in the ribosome are probably what give the CSP protein molecule family its versatility for a variety of biological functions [14,15]. Because of the CSP structure’s extreme flexibility, multifunctional features are heavily supported. RNA editing and/or post-translational changes, which were found in the silkworm moth *B. mori* (Bommo), are additional features of CSPs supporting multifunction [12–15].

CSP molecules are present in insects at all steps of their life cycles, from eggs and larvae through nymphs and adults [16–20]. They are mostly expressed in the antennae, mouth, pedipalps, and legs of locusts, and they have been connected to phase shift (phenotypic plasticity) in those insects [21,22]. However, following exposure to avermectin pesticide molecules, nearly all CSPs are up regulated in the majority of insect body tissues, especially in the gut, and fat body, which is crucial for highlighting a role in insecticide resistance as well as development, as opposed to a role in smell [23]. The issue is that, although CSP expression in non-sensory organs and during early development is well understood, research on the functional role of these proteins has been largely and obstinately conducted using competitive fluorescent binding essays, recombinant proteins testing, and a set of semio-chemicals, as described by Ban et al. [24]. Many times, the functional significance of the ligand molecules tested (usually far too few) was unknown, not required for the chemical communication of the insect under study, or highly disputed. One such example is oleamide, which is well known for being both a strong inhibitor of salivary mandibular gland branching morphogenesis and a potential CSP ligand [24]. Where and when the protein molecule is expressed should be a major factor in defining the true functional ligand. Similar findings apply to OBPs from *Aedes aegypti* mosquitoes, which are expressed throughout the insect body, including the antennae, legs, and abdomen, but are only partially identified by means of a small number of randomly selected semiochemicals [25]. There is serious uncertainty regarding the function of the protein because this ‘molecular’ study did not analyze any of the chemicals found in *Ae. aegypti*, ranging from eggs, pupae and nymphs to adults, in all many different tissues [26].

It is currently widely known that an elevated load of CSPs (‘pherokines’) is observed in hemolymph and in all insect tissues subsequent to chemical, microbial, or viral infection [23,27–29]. Liu et al. have discussed the specific function of ‘CSPs’ in lipid transport in relation to insecticide resistance using the sweetpotato whitefly *Bemisia tabaci* (Bemta) as a model study [29–31]. Insecticide-mediated upregulation and interaction of the protein (Bemta-CSP1) with long chain C18-fat lipid (C18:2, linoleic acid, LA) have been demonstrated by Liu et al. [30,31], suggesting that CSP plays a metabolic role in insect immunity and defense within the cell rather than through olfaction or chemical communication. Similar to the majority of CSPs, Bemta-CSP1 associated with LA is expressed in a wide range of insect body tissues [30,31]. LA has never been found in insect fingerprints, which is proof that CSPs are not involved in pheromone chemical communication.

The topic of this review is whether the biological function of “CSP” can be expanded in view of their tissue-expression and the fact that insect organisms inherit these molecules for the purpose of binding DNA and fatty acid lipids. We begin by rethinking the idea of the biological function of the molecules based on transcription initiation factor (TIF), mucin, and CSP analogies. We show that one of the central tenets of the literature—the role of these molecules in chemosensing—is simply accepted without any supporting data or logic approach. We can offer a new definition of the

biological function of CSP by highlighting its unique relationship to lipids, intracellular events, DNA regulation, and the particular significance of this cell concept in biology. This is accomplished by examining several distinctions and facts, such as their extensive distribution across tissues, their pervasive expression during growth and development, their reaction to stress, and their strong resemblance to the N-terminus of the nuclear, endoplasmic reticulum, ribosome, mitochondrial, cytoskeleton, and plasma membrane proteins, and thus their localization within the cells in numerous different compartments, as well as the novel insights and lessons that can be derived from them.

2. Primary Sequence Variation in Various Phenotypes of Insects

The evolution of the CSP family, particular variations in the molecule, and ultimately the emergence of a new organismal phenotype have all been linked to RNA editing, post-translational modification, and retrotransposition of RNA mutations [14,32–34]. The presence of recoding at the level of molecule synthesis in the CSP family is clearly supported by the inclusion of a Glycine residue next to a Cysteine at certain positions, amino acid inversion, and specific motif insertion in protein primary sequence [34–36].

By comparing the CSP primary sequences from Acypi, Apime, Aedae, Anoga, Bommo, Culpi, Droso, Nasvi, Pedhu, and Trica, it may be possible to determine whether these mutations can result in the emergence of a new function for the molecule and a new phenotype for the insect organism (Figure 1 & Table S1) [20,24,31,37,38]. Acypi sequences contrast with those for Bommo and Pedhu (Table S1). While Pedhu-CSPs show high identity to those from Diptera, Hymenoptera, and Coleoptera, those from Acipy show more identity to Hemipterans. On the other hand, Bommo-CSPs exhibit the highest degree of similarity to their molecular orthologs in the kingdom of bacteria and many different types of moth lepidopterans (Table S1). Specifically, it is interesting to note that Bommo-CSP5*, Bommo-CSP10, Bommo-CSP11, Bommo-CSP14, Bommo-CSP16*, Bommo-CSP17 and Bommo-CSP18* are similar to Allergen Tha p 1 and PAN-1-like protein molecules (Table S1). Allergen Tha p 1 is connected to the molecular sequence of the three truncated CSPs (*) or pseudogenes. In addition to the findings in Bemta [30,31], this could imply that the ancestral function of the protein molecule family is more closely linked to the immune system's activation during growth or within the cell's defense mechanism than it is to smell. Tha p 1 is isoallergen variant (Allergen Nomenclature; 15-kDa IgE-binding protein). PAN-1 is a protein that is both transmembrane and cytoplasmic in nature, and essential for developmental events that take place throughout the mid to late larval stages as well as for early larval ecdysis and the execution of the molting cycle during the adult molt. It is an essential control point for the development of larvae and the progression of several tissues during the transition from the larval to adult state. PAN-1 and Tha p 1 molecules are unrelated to chemosensing, and it is very likely that the proteins that are associated with these molecules are unrelated to chemosensing as well (see Table S1).

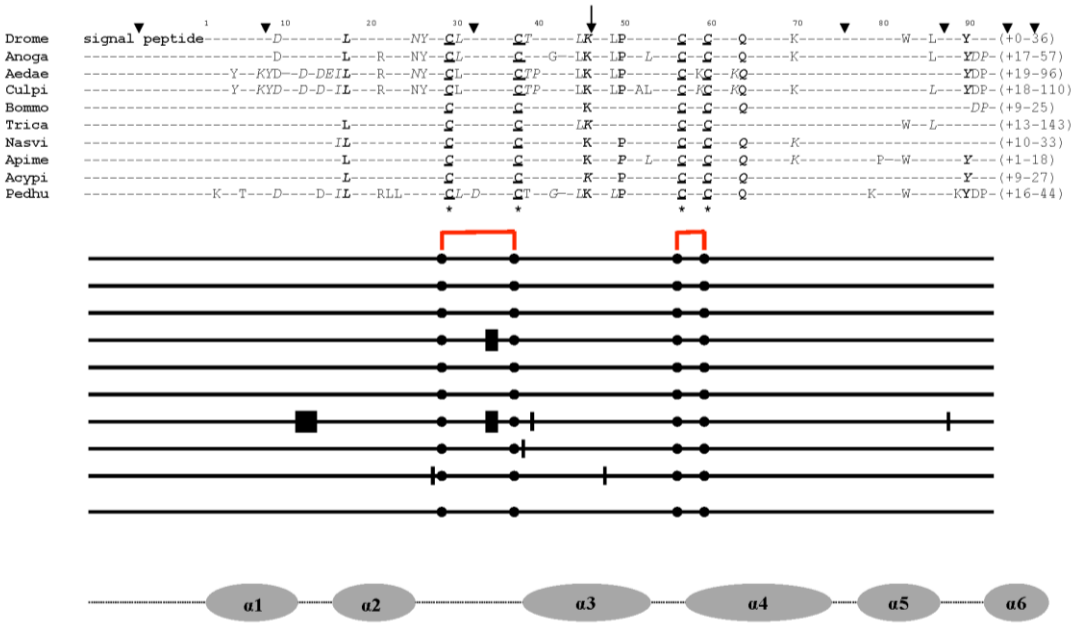


Figure 1. Consensus amino acid alignment of CSP protein molecules from insects. Drosophila: *Drosophila ananassae*, *D. erecta*, *D. grimshawi*, *D. melanogaster*, *D. mojavensis*, *D. persimilis*, *D. pseudoobscura*, *D. sechellia*, *D. simulans*, *D. virilis*, *D. yakuba*, and *D. willistoni*; Anoga: *Anopheles gambiae*; Aedae: *Aedes aegypti*; Culpi: *Culex pipiens*; Bommo: *Bombyx mori*; Trica: *Tribolium castaneum*; Nasvi: *Nasonia vitripennis*; Apime: *Apis mellifera*; Acypi: *Acyrthosiphon pisum*; Pedhu: *Pediculus humanus humanus* ([20,24,31,37,38] and Table S1). The amino acids in bold are those that are common to most CSPs. The amino acids that are strictly conserved throughout all CSPs are underlined. The four cysteine residues (4C) that are unique to CSPs are marked by stars in their respective locations. The italicized amino acid residues indicate those that alter in specific CSPs. The identification of the CSP proteins' N-terminal sequence through Edman degradation establishes the numbering of amino acid residues [4,17]. The CSP genes' introns are consistently found after Lysine 45 (see up arrow). The codon for amino acid 46 (Glu, Ser, Lys, Asn, or Asp) is broken up by the intron. Additional introns' points of insertion are shown by the black triangles. As indicated by the squares, there are varying numbers of amino acid residues between Cys29-Cys37 and Cys56-Cys59 (6-8, 18-19, respectively). The location of functional elements (α-helices) is indicated by the grey circles beneath the alignment. The disulfide bridge and interlocked cysteines are shown in red.

It is noteworthy that, depending on the species of insect, the molecule's amino acid content between Cys29-Cys37 and Cys56-Cys59 varies from 6-8 to 18-19 (Figure 1). For Culpi-CPIJ002628, Nasvi-NV16080, and Locmi-CSPs an amino acid insertion (two residues) between Cys29-Cys37 has been found [see [16,20,31,39]]. For Nasvi NV16076 and NV16077, Apime-GB17875, and Acypi000345, an amino acid insertion (one residue) between Cys56-Cys59 has been found ([31] and Figure 1). Specifically, amino acid mutation (insertion or deletion) is observed in various regions of the CSP molecule in the jewel wasp phenotype Nasvi (Figure 1). In contrast, the intercysteine gaps in Drosophila, Anoga, Aedae, Bommo, Trica, and Pedhu are strictly preserved, which indicates that the CSP protein molecule's degree of mutations is highly dependent on the phenotype of the insect (Figure 1). We found that, in order to maintain the stability of the molecule, all CSPs appear to have four cysteines (4C) and two disulfide bridges. The length and composition of the space between the cysteine residues vary, though. This definitely gives the molecule a new function and could help a new insect phenotype emerge (Figure 1).

3. Existence of “CSPs” in Phenotypes of Bacteria

It should be discussed that mutations on CSPs might have had an impact on not only the evolution of insects but also the evolution of cells overall. These CSP molecules are not an insect's apnage. They are also expressed in a wide range of organisms, including many species of arthropods, crustaceans, shrimp, crab, lobster, and copepods [34,40]. They don't, however, only exist in arthropod species. The fact that they are widely expressed at the level of the bacterial superkingdom indicates that they exist also in prokaryotes as well [15,31,34,38,41] (also see WP_149730592 in multi-species). Prokaryote CSPs are twins or 'identical twins' to insect CSPs [34,38,41], which is significant when discussing the function of these molecules. They have been reported on bacterial species, including Gammaproteobacteria *Lysobacter* (Xanthomonadaceae) and *Escherichia coli* (*E. coli*), *Acinetobacter baumannii* (*coccobacillus*), *Macrococcus caseolyticus* (formerly *Staphylococcus*), the filamentous actinomycete *Kitasatospora griseola*, the *Actinobacteria* genus in the families Enterobacteriaceae, Nocardioideae, Pseudonocardiaceae (*Solhabitans fulvus*), and Streptomycetaceae [31,34,38,41]. According to RNA and genomic reports, CSP molecules are also found in Firmicutes, Aeromonadales, Alteromonadales, Eubacteriales (*Clostridium perfringens*), and Hyphomicrobiales (MDK0835621, MDK0841570; see [34]). These findings from oxidase-positive nonendospore forming, motile, nonmotile, and spore-forming bacteria cast doubt on a chemosensory role for CSPs. Not every bacterium uses quorum sensing. These microbes are recognized as typical digestive tract bacteria, primary prokaryotic secondary metabolites, multi-drug resistant opportunistic pathogens, highly positive cytochrome c oxidase reactions, and multi-species symbionts in insects, and plants, but rather not for their olfactory acuity. *Acinetobacter* species are significant soil and aquatic microbes that aid in the mineralization of molecules like aromatic compounds (benzene). The smallest free-living prokaryotic cell (0.013 μm^3) with very low GC (33%) is found in marine Actinobacteria (*Candidatus actinomarinidae*) [42]. Given that their geographic distribution is similar to that of picocyanobacteria, there appears to be a strong relationship between the *Candidatus* and picocyanobacteria microbial groups. Based on the existing literature, it appears more likely that these two types of microorganisms exchange or share molecular modules and toxin-antitoxin systems rather than pheromones [43,44]. *Clostridium perfringens* (also known as *C. welchii*, or *Bacillus welchii*) is known for α -toxin, the toxin involved in clostridial myonecrosis. Nevertheless, other bacteria also produce volatile organic compounds (VOCs), including formaldehyde, methyl mercaptan, isopentanol, and trimethylamine [45]. These VOCs can act as biological indicators of the presence of certain pathogens because humans (and industrial animals) do not produce them. Even better, some VOCs have been found to be useful in the identification of specific bacterial species: isovalefor bacterial identification have even been identified: methanol, pentanol, ethyl acetate, and indole for *E. coli* [45]. This is not the case with volatile carboxylic acids (VACs), pyrazines, chemosensory signals, aggregation odors, cohesion, or sex pheromones, which have infrequently been connected to associations between microbes and/or insect-microbe interactions [46,47]. Although studying the diversity and binding characteristics of 'CSP' proteins in axenic insects—which do not have bacterial populations in their digestive systems—would be interesting, it is highly unlikely that fecal bacterial CSP molecules play a role in odorant pheromone signal transmission: 1) for all CSPs, there is no evidence of specific binding to VACs, VOCs, *E. coli* odors or insect cohesion pheromones, 2) numerous bacteria produce CSPs that are strictly identical to those found in *B. mori*, however, the Bommo-CSPs are all widely distributed throughout the entire body of the insect, not just found in its sense organs. They experience a striking up-regulation in response to pesticide drug exposure [24].

4. Expression Profiling in Development, Organisms, and Tissues

Referring to a whole family of proteins whose primary function is supposed to be chemical detection (i.e., proteins carrying odor molecules to olfactory receptors) as “chemosensory proteins” is highly incongruous, given the striking similarities between insect and bacterial CSPs and the fact that all CSP molecules are expressed outside of the sensory organ paradigm. To call these proteins “chemosensory”—that is, proteins that carry smells to olfactory neurons— would be incredibly inconsistent [24,31].

'CSPs' are found in phyllosoma, in larval and adult stages, and in the adult stage, in many different non-sensory organs of many different crustacean species, including Antarctic copepod, crab, crayfish, lobster, salmon louse, prawn, shrimp, and water flea [34]. This situation is pretty similar to how insect CSPs are being described right now. Insect gland and venom can also contain CSPs, in addition to the *sensillum* [5,14,48]. In moths, sixteen CSP molecules appear to be expressed in the sex pheromone gland of a species that uses only one sex pheromone compound, bombykol [14,24,36,49]. However, in addition to the female moth pheromone gland, CSP-expressing tissues and secretions also include the antennal branches, mouth, mandibles, salivae, proboscis, cephalic capsula, eyes, head, thorax, abdomen, epidermis, fat body, gut, wings and legs, i.e., a variety of reproductive and non-reproductive, sensory and non-sensory fluids, but in particular the majority of covering tissues and metabolic tissues [24,31,50–53]. Most CSPs are expressed in enveloping and metabolic tissues, so there is little clear function assigned to a single tissue or organ such as proboscis or antennae. Expressed Sequence Tag (EST)-base data analysis in *Pedhu* and *Acipy* phenotypes (Figure 2) reveals distribution of CSP-RNAs throughout the entire body, in the antennae but also in the head, thorax, and abdomen, similar to *Apime*, *Bemta*, *Droso*, *Trica*, and *Bommo* [20,24,30,31,37,38]. Six transcripts from first instar *Pedhu* larvae and engorged adults were used to identify the CSPs in the body louse [38,54,55]. According to research from The International Aphid Genomics Consortium (2010), *Acipy*-CSPs are also found in a variety of stages, including the head and the antennae of third instar nymphs, and in winged and wingless parthenogenetic females that have either received an ampicillin treatment or have been inoculated with bacteria for removal of pathogenic agents [56]. Finding out how CSP molecules are expressed in multiple different tissues can provide information on how they function. By definition, the insect EST database, which contains information about the relationships of molecules and tissues of origin, contains more than 30,000 messenger RNA sequences from *n* tissue libraries [57,58]. We "dissected" *Apime*, *Aedae*, *Anoga*, *Bommo*, *Culpi*, *Droso*, *Nasvi*, and *Trica* using Flybase and specific databases such as VectorBase or KAIKO/SilkDB [20,24,31,37]. The same methodology was utilized to analyze *Acypi*- and *Pedhu*-CSPs from Flybase and VectorBase (Figure 2). EST-cDNA libraries from ten different insect phenotypes were analyzed, and the results corroborated Northern blot, Western blot, PCR, and enhanced real-time PCR results when multiple tissues were analyzed in the same combined experiment (not just reporting one clone from a single organ or tissue): CSP is widely distributed throughout the insect body [20,24,31,37] (Figure 2).

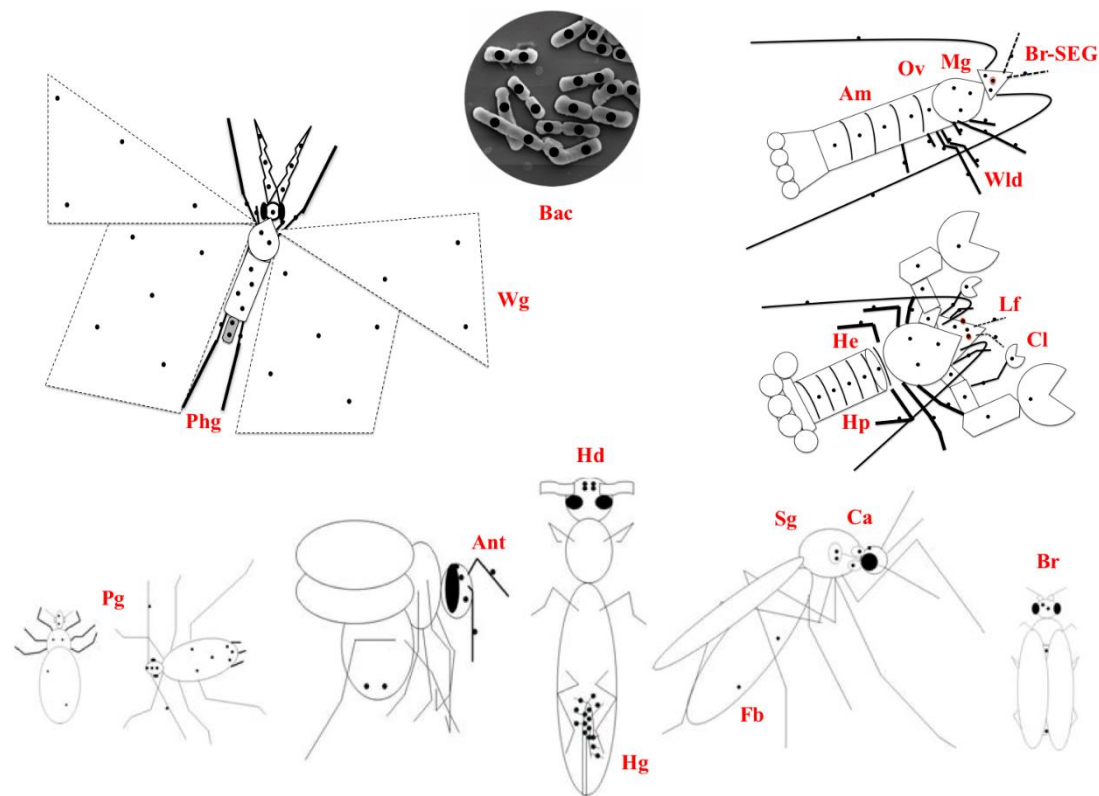


Figure 2. EST-based tissue expression of CSPs across different arthropod phenotypes. From a BLASTn search of crustacean and insect EST database using CrustBase, FlyBase and VectorBase, EST-cDNAs encoding CSP are sorted ([20,24,31,34,37] and Table S1). The location of CSPs is indicated by the black dots. Am: Abdominal muscle, Ant: Antennae, Br: Brain, Br-SEG: Brain-Subesophageal ganglion, Ca: *Corpora allata*, Cl: Claws, Fb: Fat body, Hd: Head, He: Heart, Hg: Hindgut, Hp: Hepatopancreas, Lf: Lateral flagellum, Mg: Molting gland (Y-organ), Ov: Ovary, Pg: Prothoracic gland, Phg: Pheromone gland, Sg: Salivary gland, Wg: Wings, Wld: Walking leg dactyl. Even the *Bacilli* bacterial cells (Bacil) that have taken on the shape of a rectangle contain CSPs (see Microbial Genome Database; [31,34,38,41]).

5. The Evolution and Diversity of CSP Genes

A function in chemical communication is also at odds with the diversity, evolution, quantity, and number of CSP genes. The 4-8 CSPs found in *Apime*, *Anoga*, *Droso*, *Phedu*, and *Nasvi* show that insects generally have very few "CSP"-coding genes [6,20,31,34,38,59,60]. Does this imply that these insects have poor sense of smell or that they don't communicate chemically with one another? Instead, the small number of CSPs disproves their potential for chemical communication (4 in *Droso*). For chemical communication, pheromones, and mate recognition, flies and many other arthropods are known to use a complex mixture of long chain epicuticular hydrocarbons [61]. Phenotypes induced by RNA interference suggest a role for CSP in *Apime* brain integument development rather than a role in chemosensing [62]. Bees, lice, and wasps only maintain 6-8 CSP genes, making it impossible for them to distinguish a variety of scents or complex chemical cuticular if CSP molecules only have one distinct function in chemosensing [31]. According to Liu et al. (2020), CSPs are distinctly arranged in pairs of duplicates on individual chromosomes in bees. When examining *Acypi*-CSPs (*Acypi*000094-*Acypi*009116, *Acypi*000093-*Acypi*002311, and *Acypi*000096-*Acypi*003368), we find a similar distribution of duplication on particular scaffolds (Figure 3). CSPs may therefore cooperate to cause cellular function. These *Acypi*-CSPs resemble DNA sequences that have been copied invertedly (Figure 3). They are pointed in different directions and have a junction between them. As such, it is expected that they will raise translocation rates [63]. The most prevalent type of

chromosome rearrangement, inverted duplications, also known as foldback inversions, is frequently seen in the emergence of new phenotypes, including particular developmental features in the brain. Chromosomal rearrangement and the degree of synteny play a role in pesticide resistance in aphids, which has long been known [64,65]. As seen in *Acypi*, chromosome rearrangements are undoubtedly the most dramatic type of mutation, frequently resulting in rapid evolution, phenotype adaptation, and speciation, independent of the development of the odor pheromone system [66].

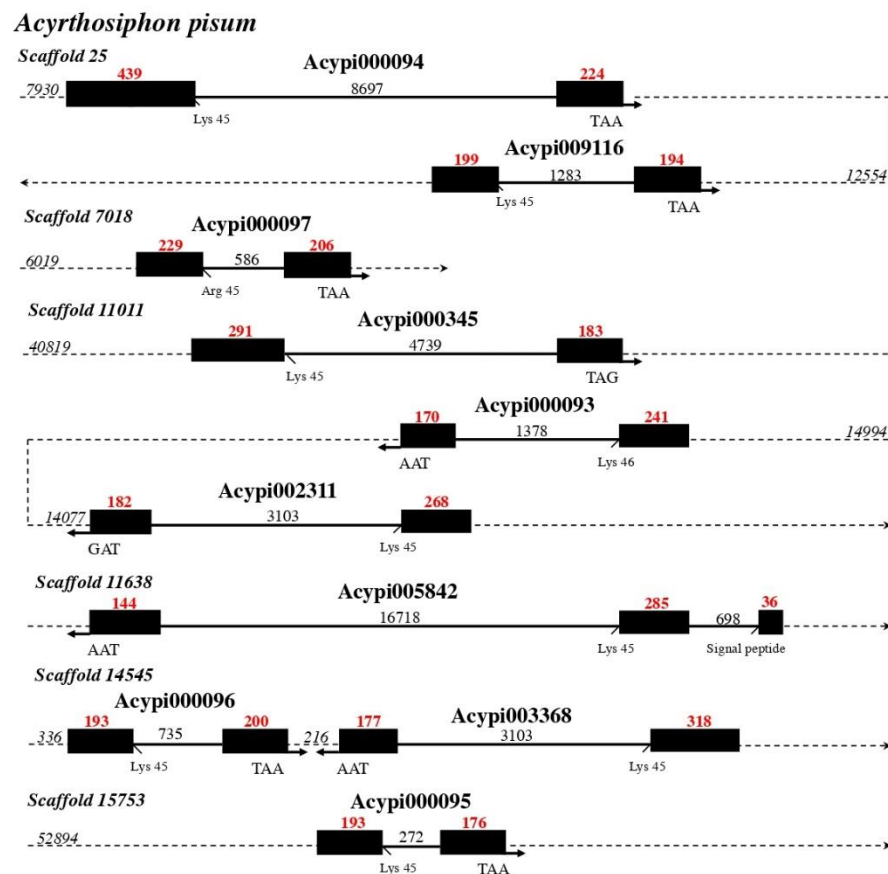


Figure 3. Ancient origin and common illustration of CSP gene structures in insects: organization of aphid CSPs on different scaffolds (25, 7018, 11011, 11638, 14545, and 15753) and genomic CSP gene repertoire in *Acypi* (see Table S1). Exons are shown as black boxes and introns as plain lines. The numbers above indicate the size in base pairs of each segment of a gene or intergenic distances (italics). The exon sizes are denoted by the numbers in red. The spacing between genes is indicated by dotted lines. The direction of the arrows, either 5'-3' (right) or 3'-5' (left), denotes the orientation of the gene. Intron insertion sites are the residues at position 46 (after conserved Lysine45 or Arginine45). *Acypi*-CSP genes have TAA or TAG stop codons, which are indicated by the stop codons in *Acypi000094*, *Acypi009116*, *Acypi000097*, *Acypi000093*, *Acypi005842*, *Acypi000096*, *Acypi003368*, and *Acypi000095* (TAA), as well as *Acypi000345* and *Acypi002311* (TAG). The *Acypi* genomic DNA annotation provides the gene name, along with the gene's location on the genome (genes are plotted onto scaffolds and genome assemblies without regard to function).

The *Acypi*-CSPs are composed of two exons separated by a single, variable-length intron, just like in most other insect phenotypes (Figure 3) [15,24,31,34,37,38,67]. A few nucleotides following the start codon that codes for the amino acid methionine, one more intron is inserted in *Acypi-005842* and genes like *Trica-AAJJ1196A*, *Bommo-CSP19*, *Apime-GB19453* and *Pedhu-594410*. This intron (phase 0 intron) is inserted after the third base and does not disrupt the codon. This demonstrates the tight regulation of the splicing of the signal peptide region and supports the notion put forth by Blobel (2000) that the functional significance of a signal peptide molecule is directly related to its length [68].

The first six amino acids of CSPs could be the needle tip that breaks through the cellular or subcellular membrane. Particular molecular protein partners from the Signal Recognition Particle complex may depend on the other amino acid residues for their proper localization and/or transit within the cell system [69,70]. This renders the “CSP” molecule more elucidable, possibly as an intracellular protein in a multimolecular complex as well as a secreted protein (in lymph and hemolymph).

The fact that the introns in more than 70 CSPs and all other single-intron CSP genes vary in length but are always found in the same place or boundary (Lysine 45) in aphids, bees, beetles, flies, lice, mosquitoes, moths, wasps, and whiteflies suggests that all these genes share a very ancient common ancestor [15,20,24,30,31,34,37,38,67] (Figure 3). The Ordovician, when terrestrial plants first appeared, is thought to be when the class of insects first appeared on Earth 480 million years ago. Overall, the current findings would be compatible with a shared heritage for the CSPs of all insect species, including those belonging to the orders Coleoptera, Diptera, Homoptera, Hymenoptera, Lepidoptera, and Neoptera. There is an extremely long history of the CSP molecule. After that, every group or order eventually started to exhibit some unique patterns, maybe as a result of the appearance of unique phenotypic and/or functional traits [31].

6. Ancestral Functions and Lipoid-Binding Properties

This protein molecule family is extremely common due to its presence in prokaryotes, binding to LA, and widely distributed profiling of gene expression, in addition to its highly conserved gene structure. This strongly suggests that it plays a fundamental role in the transport, uptake, lipid exchange, and lipid metabolism of long chain fatty acids (FAs) that are essential for development, flight, pheromonogenesis, reproduction, and immune responses [31].

Nomura et al. (1982) described the first member of this small soluble protein family as an up-regulated factor (*p10*) in the regenerating legs of *Peram* [71]. The same protein molecule (called *p10*) was found in the antennae and legs of *Peram* at the adult sexually mature stage, with some differences between males and females. This finding in the cockroach rather suggests a “chemodevol” function for this protein, contributing to tissue development and recognition of sex-specific signals like plant odors and/or odor sexual pheromones [4]. One (polyclonal) antibody against “CSP” protein labeled the antennal *sensillum* in immunocytochemistry experiments, but the labeling was diffused to the cuticle and supporting cells as well as sensory structures [5,72]. Serums containing polyclonal antibodies are combinations of antibodies with distinct specificities. The fact that polyclonals are far less specific than monoclonals is one of their drawbacks. Significant heterogeneity within the antibody pool, non-specific interactions with the molecule, and an increased risk of cross-reactivity (false positives) are present. For a rabbit to produce the “correct” antibody, multiple vaccinations against the same molecule are typically required. It is necessary to determine which sensilla express CSPs, how signals move from internal heating of supporting cells to sensory dendrites submerged in sensory lymph, and/or how these extracellular CSPs can arise from an intracellular one using monoclonals. It is primarily necessary to clarify the labeled sensilla’s purpose. Coeloconica sensilla, peg-in-pit sensilla, thermoreceptor, hygromoreceptor, infrared receptor, and acoustic receptor all have dendritic branchings similar to that of chemosensory pheromone sensilla [73].

The vast majority of CSP molecules is expressed outside of the olfactory paradigm and takes part in the moth response to avermectins, as reported by Xuan et al. [24]. According to Loftus et al. [74], Verjovski-Almeida et al. [75], Noriega et al. [76], Nene et al. [77], CSPs are also highly prevalent in *Plasmodium gallinaceum*, *Brugua malayi*, and Dengue virus infections in *Aedae*, and bacterial infections in *Culpi*. The abundance of CSP genes in moths and mosquitoes is probably related to toxin, insecticide, viral, or microbial resistance rather than chemosensing [20,24]. Interestingly, the Dengue fever mosquito, *Aedae*, has another use for CSPs. These molecules are found in both adult *Aedes* and larvae, but they are primarily concentrated in the adult’s fat bodies, *corpora allata* (endocrine sources of JH), and salivary glands [20]. This suggests that, at least in *Ae. aegypti*, “CSPs” are involved in the binding, transport, and/or biosynthesis of JH.

The widespread tissue and developmental expression of CSPs, as well as the fact that CSP is present in other arthropod species, such as copepods, shrimps, and water fleas in addition to insects

strongly suggest the function of CSP molecule in relation to juvenoids [34]. This role would be similar to that of the OBPs in the hemolymphatic transport of the JH molecule, according to Kim et al. [78]. It is well known that JH regulates the developmental processes in arthropods [79]. Therefore, “JH-related proteins”, or “JHRPs” would be a better name for these proteins than CSPs, or SAPs (“Sensory Appendage Proteins”). There are no sensory appendages on *corpora allata*. Nearly every facet of physiology in both systems is regulated by JH function in both larvae and adults. Caste differentiation, pheromone production, development, male genital morphogenesis, female ovarian egg production, social and mating behaviors, and immune function are all regulated by JH [80]. Specifically, it has been shown that JH controls neuronal plasticity linked to brain structure, neurogenesis, and behavioral maturation; this is comparable to the pattern of CSP expression observed in arthropods [31,34,81,82] (see Figure 2). According to transcriptomic studies from Bian et al. [83], even the prothoracic glands (a source for ecdysone) express CSP, which is incompatible with a function in olfaction.

An important observation in this analysis is that bacteria and prokaryotic cells share the JH relationship and CSPs. Phurelipids are secondary metabolites that bacteria produce. Phurelipids and juvenoids have a similar molecular structure. Not only did JH and other phurelipids hinder insects' immune responses, but they also prevented pupae from developing into adults [84]. These findings strongly suggest that “CSP” molecules serve a variety of functions in the immune defense system of microbes and arthropods, ranging from the binding of a wide range of exogenous foreign toxic insecticide chemicals to the intracellular transport of endogenous fuel lipids, FAs, and hormones [24,30,31,37,38]. This is accurate, unless it can be shown that CSPs are a component of the microbe bacterial olfactory hoedonics. Sifting out a subset of the CSP family that is unique to antennae or olfaction is challenging because most of the molecules are produced in the gut (hindgut) and fat body—which are believed to be the arthropod body's main organs for storing lipids and FAs as energy—. These fats and FAs are released into the hemolymph through lipolysis, allowing other organs to use them as fuel for growth, development, regeneration, and defense against infectious viruses and pesticide chemicals [15,24,27–31,37].

The existence of CSPs in bacteria and their crucial roles in the development of honeybee heads, shrimp molt, arthropod general immunity, JH, pheromone synthesis, and/or behavioral changes in specific phenotypes are consistent with a role for CSP molecules in lipid transport [20–34].

Ozaki et al. [85] state that “CSP molecules” mediate the identification of chemical signatures composed of cuticular lipid FA hydrocarbons, like those found in ants; among other organisms. It is currently unknown how each of these “CSPs” connected to FAs functions in chemical communication, development, and/or other physiological processes. Whiteflies have demonstrated that FAs such as LA bind to the CSP's functional structure, but LA (C18:2) does not involve chemical communication [30]. Like crustaceans, most insects are unable to produce LA on their own. The body's various cells need to use uptake and transport mechanisms like CSPs to use it as fuel. Insecticide chemical compounds (cinnamaldehydes) generated from plant oils, such as “dangerous” toxic chemicals, can interact directly with other functional CSP structures, as has been shown in whiteflies [30]. Therefore, two more points are essential to understand the role of CSP molecules: 1) CSPs appear to play a variety of biological roles and are expressed in both bacteria and arthropods; 2) CSP molecules in particular can activate specific innate immune pathways when injected into the phloem of plants [86,87].

It is not a sensory lymph, phloem. Phloem is the vascular tissue that carries and distributes water, carbohydrates, and other soluble organic substances to the various sections of the plants. The phloem transports nutrients and food that the leaves produce through photosynthesis (photosynthates) to all other parts of the plant. It functions structurally in the body of the plant and serves as a pathway for numerous hormones and other signaling chemicals. The three cell types that typically comprise this tissue are sieve elements, parenchyma, and sclerenchyma. It is through these tissues that CSP molecules, like Myzpe Mp10, can trigger an immune response instead of activating OR [86,87]. It becomes significant to note that in this case, Mp10 shares a great deal of similarities with other proteins, including neural Wiskott-Aldrich Syndrome (WAS/WASL)-like proteins, actin

skeleton regulatory protein (ASRP), Arp2/3 complex (Actin Related Protein 2/3 complex) activator (actin filament binding proteins and WASPs), peplos-cell cross-linker, stress response initiator, cell wall/envelope protein (CWP), Mucin, Extensin, PAN domains (these versatile domains mediate interactions between proteins and carbohydrates, fulfilling a variety of biological roles), splicing regulator, Rho GTPase (Rho) activator, Serine/Threonine-protein kinase from Social amoeba (SamkC), transcription initiation factor (TIF), nuclear pore complex protein (NPCP), Sec31, UL36 tegument protein, and many others (30.71-45.13% identity; Figures 4, S1 and S2 and Table S2). These findings raise very interesting questions about the possible roles of “CSPs” in the cytoplasmic membrane and cell surface, as well as, in the cell cytoplasm and nucleus, RNA/DNA binding, RNA/DNA regulation, interactions with multiple molecular and genetic components, transcription control, splicing regulation, and activation of specific protein complexes.

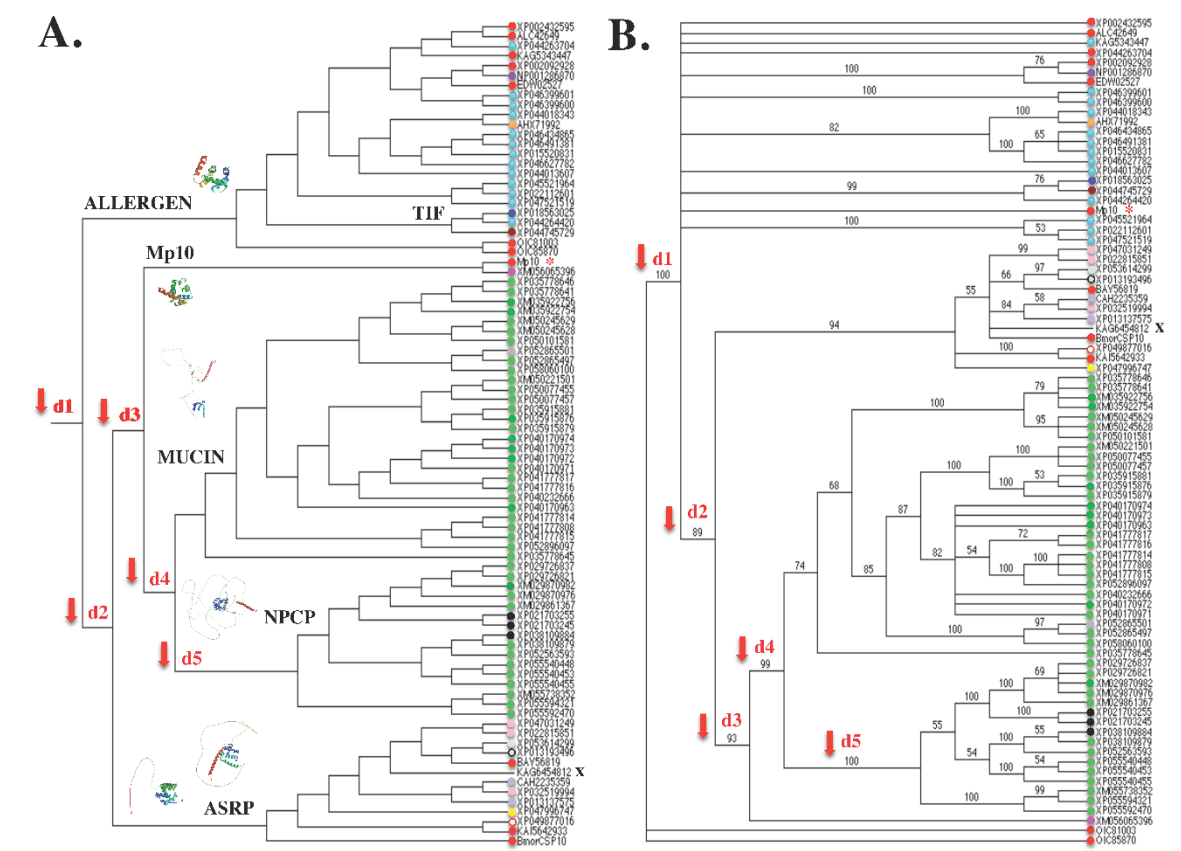


Figure 4. Molecular phylogenetic comparison (PAUP*10Altevec) of Mp10 with related proteins from the CSP, Allergen, Mucin, Rho, TIF, ASRP, and NPCP families. **A.** UPGMA analysis of Mp10 and counterparts (Table S2): agglomerative (bottom-up) hierarchical clustering UPGMA analysis based on the distance matrix of the analysed taxa that were calculated from a multiple alignment in ClustalW. 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-CSP10 (BmorCSP10). The red asterisk indicates Mp10’s position (*): grouping with pherokine XM056065396 and Mucins. **B.** Bootstrap/Jackknife algorithm analysis of Mp10 and related Allergen, Mucin, TIF, ASRP, and NPCP proteins (Table S2) with bacterial *A. baumannii* CSPs (OIC81003 and OIC85870) as outgroup. 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). In red: Ebsp-3/PebIII A10/OS-D, light blue: Allergen Tha p1, 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 protein amino acid sequences are used. On top of the branching tree are the model protein structures that correspond to the molecular groupings.

7. CSPs' Intracellular Functions (Gene Regulation to Stress Response)

The fact that the Myzpe CSP Mp10 and many other “intracellular” regulatory elements are related cast doubt on the claim that CSPs are “chemosensory” molecules (Table S2). Mp10 appears to even be expressed by viruses when compared to some peplos proteins and outer surface membrane proteins (see Table S2).

7.1. Evolutionary Evidence Derived from Amino Acid Sequence Phylogenetic Analysis

A preliminary phylogenetic analysis in IQ-Tree found some connections between CSPs, CWP, Rho, SamkC, Sec31, SamkC, TIF, SamkC, WAS/WASL, DNA-binding proteins (DBPs), DNA-regulatory proteins (DRPs), and several RNA-binding proteins (RBPs) sequences molecular sequences (Tables S3–S4; [38]). Transcriptional/cell division repressor, helix-turn-helix, and DRP from the Xenobiotic Response Element (XRE) family of transcriptional regulators, branched with Bommo-CSP16, Bommo-CSP8, and Bommo-CSP9 and their bacterial counterparts, whereas another orthology group included Trica-CSP AAJJ0012J and *Ruminococcus* DRP WP_044998036 (nucleotide binding; see Table S3 and [38]). This implied that some CSPs have evolved, at least in Bommo and Trica, to perform functions associated with nucleotide binding, transcription, translation, DNA/RNA templates, DNA/RNA control, and/or intracellular gene expression mechanisms. The XRE family of proteins, which is extensively present in bacteria and eukaryotes, is involved in many aspects of controlling cellular metabolism. Under normal conditions, XRE transcriptional factors bind to the promoter region of the gene and function as repressors.

A phylogenetic examination of Bommo-CSPs, Trica-AAJJ0012J, BemtaCSP1, Myzpe-Mp10, its derivative sequences, Mucin-like sequences, XREs, and several RBPs sequences in PAUP (*10Altevec) provided further evidence for this (see Tables S2–S4 and Figures 4, S1 and S2). A neighbour-joining tree (BioNJ study) showed that Mp10 did not join the Bommo-CSPs, but instead seemed to be much more closely connected to the Mucins group (G1; Figure S1A), albeit not clearly forming an orthologous group (Figure S1B). In G1, the “CSPs” from *E. balteatus*, EbalDDB_G0285119X1 and CSP3 (QIS77910), linked with Bommo-CSP10 and showed high similarity to Rho-activator isoform X3 (Table S3; Figure S1A). The Trica-CSP AAJJ0012J molecule clustered with Allergen Tha p 1, and associated RBPs instead of BommoCSPs. The marmalade hoverfly Epiba-CSP4 (QIS77191) clearly deviated from the CSPs groups (see G2, Figure S1A). Only the G1 “EbalCSP”-group maintained a relatively high bootstrap value (94%) in the Jackknife analysis (bootstrapping calculation over 1000 repeats), strongly demonstrating the branching of “CSPs” with Rho GTPase enzyme. A common branch with a significant bootstrap value (57%) was formed by the sequences combining CSP, CWP, LRR, Mucin, SamkC, TIF, and WAS/WASL. Mp10 slipped off of this group, but AglaCWPX3 drew it in (Table S3; Figure S1B). The papilionid IpodCSP (CAH2042437), DBPs, RBPs, and TIFs were drawn at the bottom of the tree in this analysis using Bommo-CSP2 and bacterial counterparts as the reference outgroup (Figure S1B). This strongly implied that there were significant and expected relationships between “CSPs” and all of these various intracellular proteins, mucins, Rho-activators, and translational regulatory factors.

Focusing on Mp10 (referred to as a very typical “CSP”, 153 amino acids, 17.2 kDa, “consensus residues”, four cysteines pattern, whole body expression), we found that this specific protein sequence (XP_022173691) constructs a relevant hierarchical clustering UPGMA (unweighted pair group method with arithmetic mean) analysis and a phylogenetic tree supported by a high bootstrap value when compared to Allergen, Mucin, TIF, NPCP, and ASRP amino acid sequences (Figure 4). It's interesting to note that Mp10 molecule does not cluster with other “CSPs” on the UPGMA tree. It belongs to a large group that includes all Mucin taxa and TIF. It shares a close relationship with Rho-activator (XM_056065396; Figure 4A). The Rho families of small GTP-binding proteins (20-30 kDa)

are not extracellular molecules. These are intracellular proteins that control the actin cytoskeleton-related Rho-GTPase signaling pathways. They act as molecular switches that control a variety of cellular processes, including gene transcription and cytoskeleton-related events. RhoGAPs, one of the main classes of Rho-regulators, are found in all eukaryotes (within the cell) and have been shown to regulate a variety of cellular functions, such as the organization of the cytoskeleton, growth, differentiation, neuronal development, and synaptic functions [88]. They have nothing to do with smell; they don't trigger transmembrane chemosensory receptors.

As for the other “CSPs”, they belong to either the ASRP group (see the position of BAY56819, KAI5642933, and Bommo-CSP10, in the UPGMA analysis on Figure 4A) or the Allergen group (which looked to be the most ancestral molecule): XP_002432595, ALC42649, KAG5343447, XP_002092928, EDW02527, OIC81003, and OIC85870 (bacterial proteins). Significant outer envelope proteins, RickA-like (Arp2/3), and particular nuclear nucleoside kinase (NNK) are found in the ASRP group, while IgE-BPs, pherokines (fly hemolymph CSP proteins), acid trehalase (AHX71992, involved in intracellular trehalose mobilization during postdiauxic growth and severe saline stress in yeasts), Cell Wall-Anchored (CWA-3, XP_018563025), and TIF sequence (XP_044745729) are found in the Allergen group (Figure 4A). RickA is the protein from *rickettsiae* bacteria, which are carried by ticks and lice. It is involved in bacterial host cell binding and infection as well as the actin-based motility of bacterial cells. It also triggers host cell factors related to the cytoskeleton [89]. CWA-3 molecule is a component of the cell wall integrity-signaling pathway, which is regulated by small proteins like GTP-BP Rho1. Controlling gene expression and coordinating periodic modifications to the cell wall are the primary functions of CWA molecule during the cell cycle and in response to various forms of cell stress [90]. NNKs (or nucleoside diphosphate kinases, NDPKs) catalyze the transfer of the terminal phosphate (P) from a donor triphosphate (TP) to an acceptor diphosphate (DP). Arp2/3 is a ubiquitous and essential component of the actin skeleton found in eukaryotic cells. It nucleates actin filaments, caps their sharp ends, and cross-links them to form orthogonal networks. These molecules form a large group that is associated with Mp10, a protein classified as a “chemosensory protein”. The functions of these proteins are not as closely related to smell as they are to the cytoskeleton, actin filaments, transcription of genes, phosphate transfer reaction/exchange, phosphorylation, mitochondrial energy production, TP-DP conversion, and cell regulation (see Figure 5).

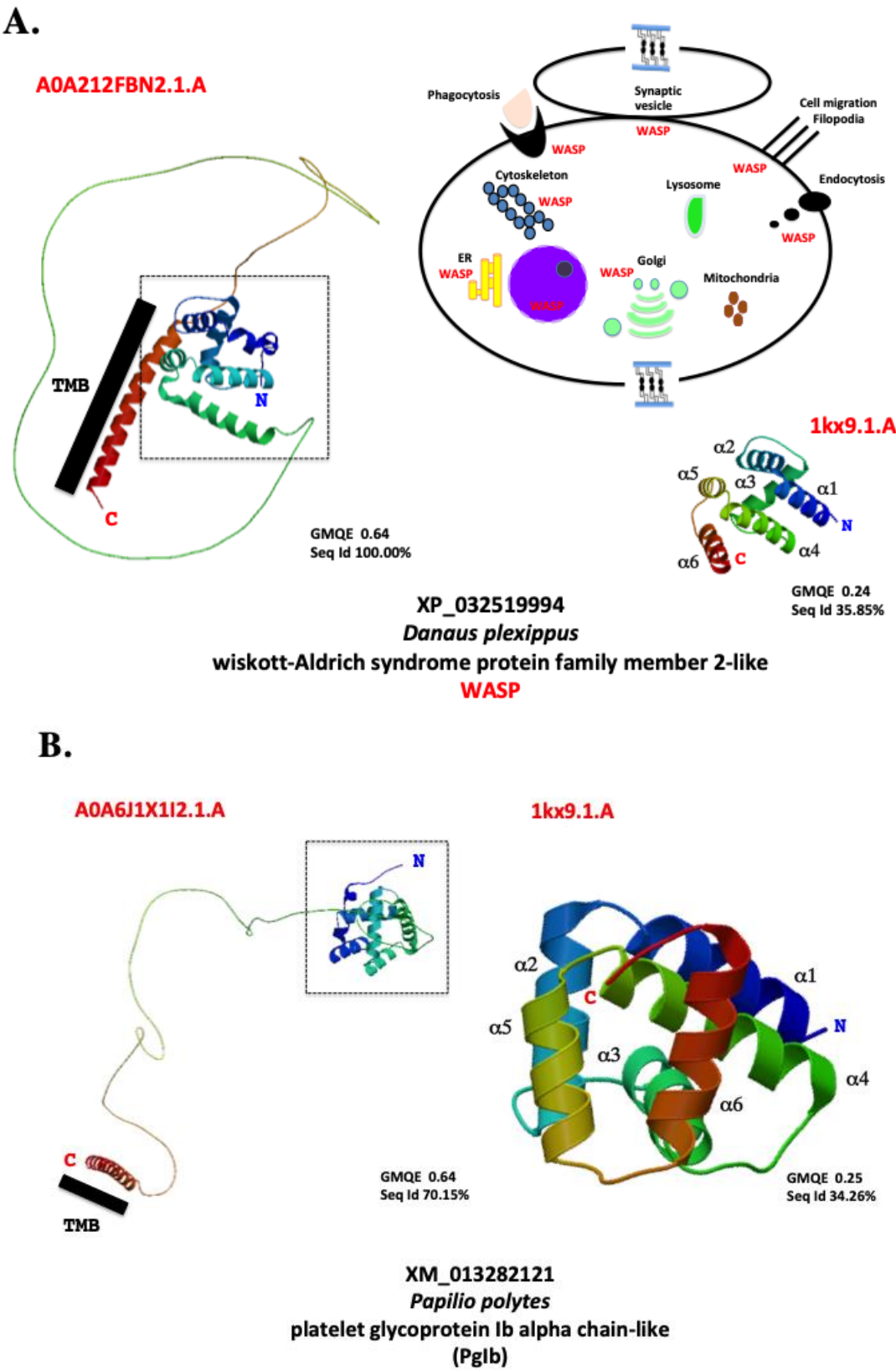


Figure 5. Molecular structure modeling of CSP, Allergen, Mucin, Rho, TIF, ASRP, and NPCP. **A.** Molecular structure modeling of Danpl-WASP (XP_032519994). **B.** Molecular structure modeling of Pappo-PgIb (XM_013282121). WASP and PgIb sequences were aligned with Bommo-CSP1 in order to identify the signal peptide and cut it off based on N-terminal sequencing by Edman degradation [17]. The amino acid sequence of the mature protein was then subjected to molecule structure modeling

using Swissmodel.expasy.org. The molecules with the highest identity score was used as template references: 1kx9.1 ("Chemosensory Protein A6", X-ray, 1.6 Å, monomer, cabbage moth, *M. brassicae*) 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 A0A6J1X1I2_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 CSP prism are numbered 1 through 6. The location and function of WASP in the cytoskeleton, endoplasmic reticulum (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 CSP prism. The CSP prism is indicated by the black square with dotted lines. The two molecules share the same model of construction inside the cell: N-terminus, CSP prism, long loop, and transmembrane domain. The black bar indicates the position of the transmembrane segment (TMB).

Mp10, "CSP", Allergen, TIF, Mucin, NPCP, and ASRP have a very distant common ancestral origin according to the topology of the UPGMA tree, which is based on the assumptions of a common root and constant evolutionary rates for all lineages (i.e., it takes the "Molecular Clock Hypothesis" [91] to account for mutation rates). They are the outcome of a series of duplication events that produced Allergen, which includes CSP and TIF proteins, prior to Mp10, Mucin, NPCP, and ASRP (Figure 4A). Mp10 (and Rho activator) emerged later in a sequence of duplication events that produced a wide range of Mucin variants, particularly in the mosquito genera *Aedes*, *Anopheles*, *Culex*, *Uranotearia*, and *Wyeomyia* (Figure 4A). The high level of duplication and variation observed in Culpi- and Aedae-CSPs makes this point noteworthy [20]. During its evolutionary history, the molecule protein gene family that includes *Mp10*, *CSP*, *Rho*, *Allergen*, *TIF*, *MUCIN*, *NPCP*, and *ASRP* appears to have undergone multiple duplications. Some of these duplications are specific to particular taxonomic lineages, like mosquitoes for the long Mucin precursor proteins needed for growth, development, digestion, oviposition, and control of viral infection [92,93]. Other gene duplications are more ancient and common to all lineages (see Figure 4A). This is accurate for the Allergen group, which comprises Trica, ants, damselflies (*Ischnura* forktails) flies, garden whites (pierids), ladybirds, lice, neodiprions, parasitoid wasps, and tuberworm moths. This also holds true for the ASRP group, which comprises taxa from multiple families of moths and butterflies (swallowtails, speckled woods, *Papilio*, *Pararge*, etc; see Figure 4A & Table S2). The UPGMA tree of amino acid sequences indicates that all of these molecules have a distant common origin that is estimated to be 324–440 Mya (the latest Mississippian–Silurian or Devonian [94]), much earlier than the emergence of the various flying insect species. This is because of the distance between the ASRPs/CSPs and the Allergen molecules.

The relationships between the CSP, Mp10, Rho, Allergen, TIF, Mucin, NPCP, and ASRP proteins were further examined using maximum parsimony analysis (Figure 4B). Strict consensus trees were established using MP and the PAUP4.0b10 (Altivec) program, as detailed in Abraham et al. [95]. CSP, Mp10, Rho, Allergen, TIF, Mucin, NPCP, and ASRP molecules are all widely expressed in different tissues, which is already a striking feature. There have been reports of allergens from the abdomen and thorax clones. It is known that Mucin, PAN-1, DAN4, NPCP, WASP, and the YLP motif are present in all body parts, larvae, and pupae (see Tables S2 and S3). The YLP motif is known to facilitate RNA binding activity, regulate telomere maintenance, and contribute to the decrease in telomerase activity that occurs during stem cell differentiation by attaching to the core promoter of Telomerase Reverse Transcriptase (TERT) and leading to its down-regulation [96]. DAN4 determines cellular morphology and plays a crucial role in maintaining cell integrity during cell growth and division, under stress conditions and upon cell fusion. It is a component of the molecular interactions and enzymatic activities in the cell wall in response to different growth phases and toxic signals from the environment [97]. This is remarkably similar to what it is known about CSP ontogeny, tissue distribution, and response to chemical stress (see Chapter 4). Phylogenetic information and expression analysis are combined to link CSPs to intracellular proteins, cell walls, and gene promoter regions. The insect CSP, Mp10, Rho, Allergen, TIF, Mucin, NPCP, and ASRP molecules are closely

related to each other; they formed groups with a high bootstrap value (89-100%), ranging from Pedhu-CSP (XP_002432595) to Uralo-Mucin (XP0055592470) and Eupco-Rho (XM_056065396). Together, the Mucin and NPCP molecules form a group with 99% bootstrap value. Additionally, they attach to ASRP and Rho groups, which may indicate a strong relationship between these molecule families based on their extremely high bootstrap values (93-99%; Figure 4B). Aside protein molecules implicated in actin microtubule association, cytokinesis, Arp2/3, and NDPK, the ASRP group attaches to CSPs with significantly high bootstrap value (94%; Figure 4B).

Interestingly, when bacterial CSPs (*A. baumannii* OIC81003 and OIC85870) were used as outgroup in the MP analysis of protein amino acid sequences, they attracted Rho at the base of the tree but not Mp10 (Figure 4B). The Allergen group, especially TIF, and Mp10 remained closer in relationship (Figure 4B). Coleopteran, dipteran, hemipteran, hymenopteran, lepidopteran, phthirapteran, and zygopteran (Odonata, blue-tailed damselfly) Allergens/CSPs and Mp10 protein are related, albeit they do not form a general orthology grouping, in contrast to Mucins, NPCPs, and ASRPs (Figure 4B). This is where the UPGMA tree and the MP tree (Bootstrap Jackknife) in PAUP4.0b10 (Altivec) diverge significantly (Figure 4). Following the evolutionary distances between insect species, Allergens/CSPs, including TIFs, segregated independently on the MP tree. Through nucleocytoplasmic transport regulation, CSPs, NPCPs, and TIFs may serve as a hub for changes in gene expression that are unique to a cell, a tissue, or a phenotype [97]. This implies that, in keeping with the previous description of "CSPs" (see Chapter 5), the NPCP-ASRP molecule family is evolving as a fast, intense, frequent, and high rate, despite its extreme age [98,99].

The Mp10 orthology of the aphid, beetle, dragonfly, fly, louse, mosquito, moth, and sawfly belongs to different groups, ranging from Mucins and NCPCs/ASRPs to Allergens and TIFs, according to our phylogenetic analysis (see Figures 4, S1 and S2), regardless of anything related to olfaction, taste, and/or chemosensing (see Table S2). This suggests that, as a result of multiple local duplications, the Mucin, Allergen, TIF, and CSP families underwent common diversification before the emergence of insects (e.g. the Carboniferous Period of the Paleozoic era 299-359 Mya; Figure 4). Mosquito Mucins were grouped together in accordance with the phylogenetic distances between their genomes. The evolutionary histories of CSP and Mucin proteins are then different. Mp10 orthologs in insects seem to have undergone one to five separate duplication events (see Figure 4). Four consecutive duplication events may have given rise to Mucins: two early duplications (d1 and d2) produced Allergens/CSPs, TIFs, and ASRPs; two late duplications (d3 and d4) produced Mp10 and Mucins. If the protein amino acid tree is the right tree to explain duplication and evolution within these relatively similar molecule families, then all of the molecules in the CSP, Allergen, TIF, Mucin, NPCP, and ASRP groups have the same root (Figures 4, S1 and S2 and Table S2). The following are linked: Mp10, CSPs, pherokines, ejaculatory buld-specific proteins (Ebsps), and a multitude of immune system, actin-related complex, nuclear complex, cell regulation, and cytoskeleton regulation proteins. The molecular sequences of ORs do not correlate in this instance. All of these Allergen-CSP and ASRP-NPCP molecules have a very old origin that dates back to the time before flying insects even existed and even further back to the time of microbes' prokaryotes' origin (about 3.5 Bya), even though a fifth duplication that produced DAN4 and NPCP may only be found in Culpi (see Figure 4).

7.2. Molecular Evidence Derived from Amino Acid Sequence Modeling Analysis

The strong relationships observed between CSP, Allergen, TIF, Mucin, NPCP, and ASRP molecules can be explained by the alignment of protein amino acid sequences and structural modeling (Table S2, Figures 4, 5, S2 and S3). Protein amino acid sequence alignment demonstrates that CSP precisely matches Allergen and the N-terminus of TIF, Mucin, NPCP, and ASRP (Figure S2). When compared to 1kx9.1 as template reference (X-ray crystal structure, [7]), it appears that the majority of query molecules fold into a prism of six α -helices using SWISS-MODEL Workspace/GMQE [100]. This was especially true for pherokine Phk-3, immune response protein, acid trehalase, Allergen Tha p 1, Mp10, Bommo-CSP, and *Coccinella septempunctata* Cocse-TIF (Figure S3). While prism structure is only seen in certain areas of the molecule, such as the N-terminus, the other proteins studied in the SWISS-Model are all significantly larger molecules (Figures 5 and S3). A long

α-helical stretch (7-13 turns) forming the C-terminal tail and a CSP prism at the N-terminal region characterize the majority of representative molecules in the TIF, Mucin, NPCP, and ASRP groups (Figure S3). Swissmodel analysis predicts a transmembrane segment at the C-terminal tail is predicted for these structures (see Danpl-WASP, XP_032519994; Pappo-PgIb, XM_013282121; GMQE 0.64, Sequence Identity 70-100%; Figure 5A,B). The butterfly Danpl-WASP protein's gene locus and molecular structure are known (A0A212FBN2.1.A AlphaFold DB model of A0A212FBN2_DANPL, LOC11677206), and they do not match the characteristics of a molecule that is categorized as "chemosensory". Danpl-WASP is a single transmembrane domain-molecule that is 297 amino acids long and 33 kDa in mass. Among the most typical uses of these kinds of molecules are their actin-bound structures and their implications for filament assembly and actin cytoskeleton regulation [101,102]. These bigger molecules, such as WASP, are found in the cytoskeleton, microtubules, integument, cell wall, nuclear pore, cytoplasm, and many other intracellular organelles. Their remarkable long loop that joins the "CSP" structure at the N-terminus and the C-terminal transmembrane domain is what distinguishes them (Figures 5 and S3). To obtain this, the AlphaFold DB model of an uncharacterized protein or proteins categorized as "chemosensory protein" or "putative insect pheromone-binding protein (PBP)" is used as a template. It is strongly suggested that CSP and/or putative insect PBP exists in viruses, at least in herpesviruses that target the Golgi apparatus (HSV-1 and HSV-2), given the presence of the 'CSP prism'-prominent loop-transmembrane domain in a Palearctic butterfly molecule (CAH2235359, "jg5928" with BLLF1; Figure S3) [103]. "CSP" molecule is associated with the herpes virus's major outer envelope glycoprotein (see "jg5928"; A0A1V1WC08, GMQE 0.64, Sequence Identity 70.31%, Figure S3). We have found that bacteria possess the same kind of molecular structure as described by Liu et al. [38]—a CSP joined to a larger molecule with transmembrane and glycoprotein domain (*Lysobacter capsici* WP_096417339, 491 amino acids, 50.6 kDa). Therefore, rather than olfaction, it is more likely that infection (interaction with virus or microbe binding and immune response) is the primary function of CSP.

7.3. Cellular Evidence Derived from Location, Size, Structure, and Expression in Viruses and Microbes

These striking similarities between CSP, Allergen, TIF, Mucin, NPCP, and ASRP molecules suggest that "CSP" molecules function in a multitude of diverse ways inside the intracellular systems of both eukaryotes and prokaryotes (Figure 6).

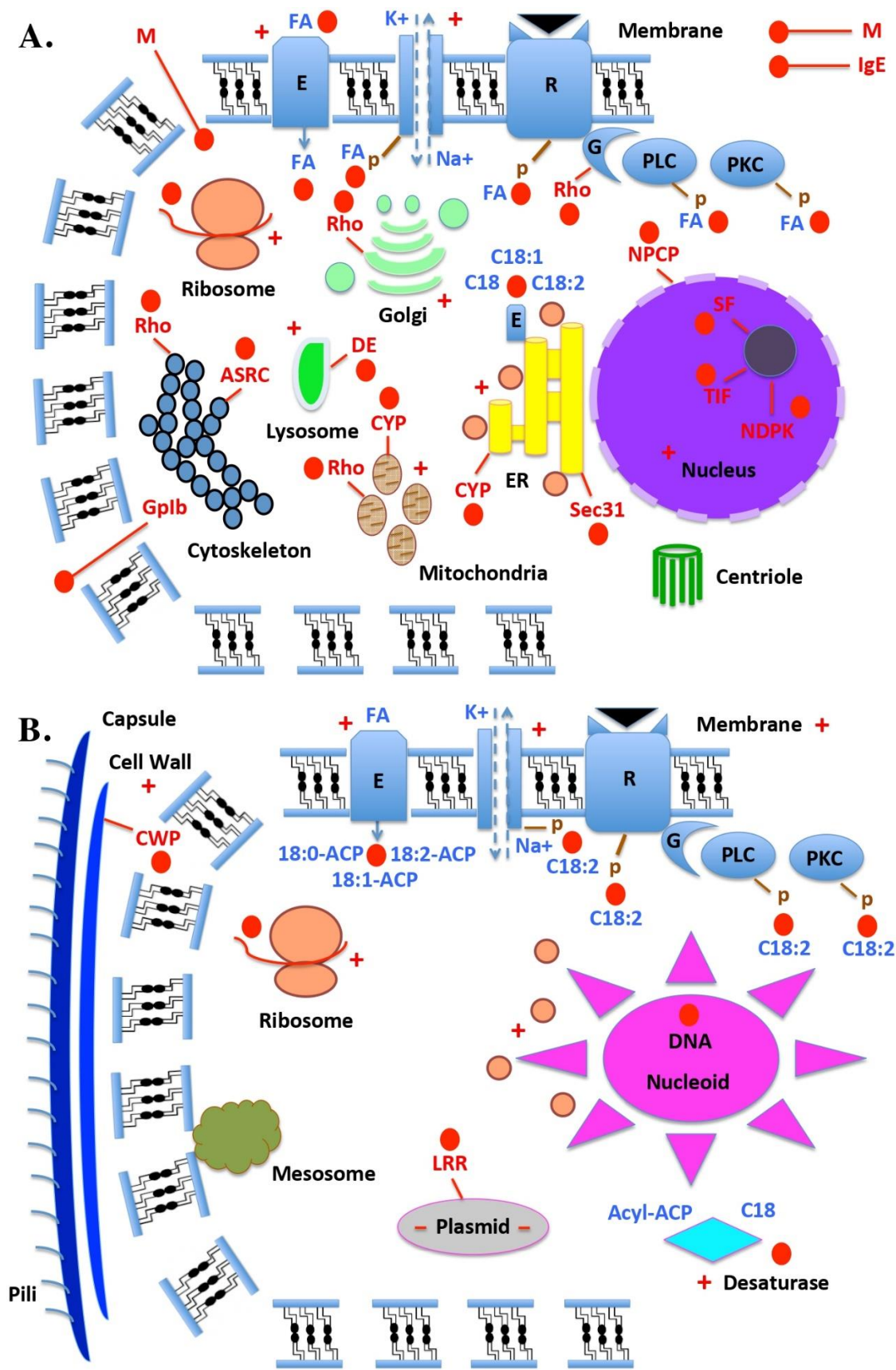


Figure 6. The intracellular systems of eukaryotes and prokaryotes use “CSP” molecules for a variety of purposes. **A.** In eukaryotes: CSP binds to fatty acid (FA), which mediates the phosphorylation (p) of different plasma membrane-bound protein molecules. E: Enzymes, K⁺ intake/Na⁺ outflow: potassium-sodium 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 CSPs 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 CSP 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). Additional molecules in the cytoskeleton, plasma membrane, and nuclear membrane that CSPs 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 immune system-building soluble allergens, while M stands for mucin-like fractions, secretory molecules that fortify the immune system's defense against microbial stress. **B.** In prokaryotes: CSPs carrying linoleic acid (C18:2) and acyl-carrier-protein (ACP) counterparts (18:0 = stearyl-ACP, 18:1 = oleoyl-ACP) interact with integral membrane fatty acid (FA) desaturase (E). C18:2-ACP = linoleoyl-acyl-carrier-protein. The phosphorylation (p) of the sodium-potassium pump (K⁺/Na⁺), receptor (R), PLC, and PKC is made possible by the transport of C18:2. Acyl-ACP, C18, and Desaturase: microbial soluble intracytoplasmic desaturase system. +: Location of reaction to conditions of biological chemical stress (membrane, cell wall, ribosome, nucleoid, ribosome-nucleoid interaction site, and ACP desaturase systems). The part of the cell's envelope that provides structural stability and a protective reaction to stress is the cell wall protein complex (CWP) attached to CSP. The red bars show the plasmid's CSP genes, which can be subject to horizontal transfer from bacteria to insects and plants. On the plasmid, CSPs are also found as molecules connected to the Leucine-Rich Repeat protein (LRR). The red dots represent every protein and supramolecular complex in every organelle that "CSP" molecules interact with intracellularly in both prokaryotes and eukaryotes.

When it comes to eukaryotes, CSP binds to LA (C18:2) which is made from arachidonic acid (AA, 20:4 ω 6) and is required for the biosynthesis of many hormones. Neither C18:2 nor 20:4 ω 6 acts as a mediator in olfaction or chemical communication [see 16]. They are related to the hormones leukotrienes, thromboxanes, and prostaglandins. These three classes of hormones regulate many physiological processes, including innate immune response, ion transport, egg development, and reproduction [104–106]. LA and AA phosphorylate a wide range of intracellular proteins (enzymes, pumps, receptors, and so forth), which in turn controls a multitude of signal transduction pathways and cellular processes [see 31]. CSPs take part in the FA lipid and phosphorylation processes, which influence the biosynthesis of stress responses in the nucleus, mitochondria, Golgi, endoplasmic reticulum (ER), plasmic membrane, ion channels, ion pumps, ion transporters, lysosomes, and ribosomes (see +, Figure 6A). This information is supported by whitefly experiments [30,31]. The "CSPs", along with LA, AA, FAs, and stress responses, most likely mediate the mechanisms controlling lipid metabolism in the cell cycle at the molecular translational level in the cell ribosome [30,31,107,108]. Moreover, CSP interacts with cytochrome P450 (CYP) enzyme in the ER and the mitochondrial system in response to insecticide exposure (+) and other stress responses. The CSP-CYP interaction is probably important for cellular metabolism, homeostasis, hormonal synthesis, toxin catabolism, and xenobiotic detoxification in response to insecticide exposure [24,109]. The lysosome's functions, including the digestion and breakdown of macromolecules (proteins, lipids, carbohydrates, and nucleic acids), repairs of cell membranes, and defense against pathogens like microbes, bacteria, fungi, and viruses, are probably also all mediated by the CSP-Degradative Enzyme (DE) system (Figure 6A). When food is eaten or absorbed by the cell, the lysosome releases its enzymes, which convert complex molecules—such as sugars, proteins, lipids, FAs, AA, and/or LA brought by CSPs—into energy that the cell needs to survive [110]. After that, the CSPs will probably also need to work with the desaturase enzymes, the ER membrane, and the eversible vesicles that secrete the cuticle, the duct, and the pheromone compound if they are to transport FAs like LA and its two precursor molecules, stearic acid (SA, C:18) and elaidic acid (EA, C:18-1). Lipid droplet formation and pheromonogenesis would strictly require the coordinated action of CSPs, FAs, and ER in the pheromone gland. The sites of metabolism that are likely regulated by CSPs are not the activation of OR on sensory neuron dendrites, but rather the ER membrane of the sex pheromone

gland, the oenocytes and the ejaculatory bulb of flies, the abdominal tergites of cockroaches, the coremata (hairpencils) of butterflies and moths, and mandibular gland pheromones [14,30,31,111,112] (Figure 6A).

Here, we find that CSP is linked to the ER, as well as to the endomembranes from the Golgi, mitochondria, and plasma membrane via Rho GTPase signaling complex, Mucin, ASRC, and NPCP (Figures 5 and 6). Together, Rho, Mucin, ASRC, and NPCP are connected to CSPs, allowing these molecules to protect the cell membrane, control membrane interactions, regulate the trafficking of protein precursors and FA lipids between the different cellular organelles, particularly the lysosomes and Golgi system, and control the functions of the genome [113,114]. Furthermore, we find that “CSP” is the N-terminal region of “Gplb”, a protein that resembles the glycoprotein receptor on the surface of human platelets, and as such, it may reside extracellularly to recognize antigens, activate the immune system, and/or recognize JH and promote cell growth (Figure 6A) [115]. We also find that CSP links nucleotide binding, RNA metabolism, and nucleus-wide control of gene expression to splicing factor (SF), TIF, and NDPK [116,117]. As allergens or mucin-like fractions, some soluble fractions of CSP are secreted, bolstering the immune system's resistance against microbial stress, especially in the saliva, gut, and eyes [118]. In the eyes, they mediate substance exchanges (gas, water, etc.) between internal organs and the environment. Additionally, they may coat the insect's eyes and antennae, enhancing its sensory capabilities. Flying insects have to deal with a variety of foreign, toxic particles, including mold spores, dust, and pollen, in the same way that swimming crabs have to deal with pesticide chemicals, pollutants, varying salinities, and ammonia environments. The existence of CSPs in sensory organs would be explained by this protective role, but not only. According to the reports of Huang et al. [119] and our research presented here (Table S2 and Figures 4–6, S2 and S3), extracellular CSP-mucin complexes may also be necessary for the adult development of insects and crustaceans. This would account for the presence of CSP, Rho, Mucin, ASRC, and NPCP throughout the various stages of arthropod development (Table S2).

Similar to eukaryotes, bacteria and prokaryotes also require the transport of LA for the phosphorylation of various enzymes (PLC and PKC), ion pumps (K⁺/Na⁺), and receptors (microbial rhodopsins) [120]. However, in contrast to eukaryotes, bacteria, fungi, and plants possess soluble intracytoplasmic desaturases [121], and it is plausible that CSP's interaction with FAs and lipids mediates the activity of this type of biosynthetic enzyme ($\Delta 6$, $\Delta 9$, or $\Delta 12$ desaturase) as well as transmembrane desaturase (Figure 6B). Regarding the transport of C:18, C18:1, and C18:2 in cell and molecular stress response, CSPs are intimately associated with the desaturase function [30,31,122,123]. This role is far more focused on supplying the cell with energy than it is on chemosensory receptor complexes (Figure 6). Similar to eukaryotic cells, ribosomes are the targets of biological chemical stress in prokaryotic cells [123]. Desaturase enzymes, FAs, CSPs, and ribosomes seem to be interdependent, especially in situations of biological chemical stress (Figure 6). The different types of desaturase enzymes are differentiated by the specificity of their substrate, which is the location of the double bond on the FA molecule. Nevertheless, there have been reports of desaturases exhibiting multifunctional activities in bacteria, fungi, plants, and insects as well [124,125]. CSP-desaturase coupling may be necessary for substrate switching in lipids and to determine the composition of FAs in glandular cells (female sex pheromone gland), *corpora allata*, sensory neurons, and/or all other cells in the insect body. Bacterial desaturases, which would need a substrate modification to change the FA content in reaction to environmental stress, may also need CSP [24,30,31,126,127]. The CSP molecule folds in the CWP complex of prokaryotes, bringing it to the cell's envelope and serving as the main stress-bearing and shape-maintaining element, as opposed to folding in arthropod chemosensory system [128]. Maintaining the shape and structural integrity of cells depends on the integrity of their cell walls, and CSP-CWP may be involved in this process (Figure 6B). Apart from providing a strong and inflexible exoskeleton to ward off damage, CSP-CWP could serve as a point of attachment for proteins that engage with the bacterial milieu. CSP-CWP complexes of microorganisms (and plants) could act as barriers against biotic and abiotic stresses. These functions include shielding the cell from chemicals, drugs, toxins, and other harsh environmental conditions, such as preventing the cell from drying out in high temperature conditions [129]. CSPs may

bind to toxin chemicals, expose them to enzymes that degrade them, and/or contribute to the supramolecular agency of the CWP complex because they are found in the N-tail of CWP (see Figures 5 and S2). Rather than using chemosensing, examining these possibilities for the purposes of CSP would be interesting.

Furthermore, we find that CSPs are molecules associated with the Leucine-Rich Repeat protein complex (LRR) present on plasmids (Table S1; Figures 4–6B and S3). Their participation in various protein-protein interactions that activate the plasmid and increase its pathogenicity and virulence may occur there [130,131]. Finally, it can be inferred that CSPs and DNA-dependent RNA polymerases (Dd-Rps) regulate transcription initiation in bacterial cells in a manner similar to that of eukaryotes based on the finding that CSP is linked to the N-terminus of TIF (see Table S2, Figures 4–6, S2 and S3) [132]. This finding suggests that CSPs affect a gene’s transcription level within a particular cell or tissue. It is rather unclear how this relates to RNA binding and/or RNA polymerase activation, but it is undoubtedly unrelated to a chemosensory function (e.g. ‘CSPs bind odors that activate receptors’).

8. To Design “CSP”, New Terminology Is Needed: Structure or Function?

Table 1 compiles the history of the various names used to describe this molecule protein family as early as 1992. Even though p10 was found in leg regeneration tissue rather than antennal sensory structures, it changed names three times between 1994 and 2000 (see Table 1). All of the changes ignored the first protein gene family member to be reported in the literature and had the same obvious and unambiguous meaning (chemosensory function, limited to chemosensory tuning). Despite the lack of sufficient evidence, the protein gene family was renamed to exclude any possible developmental functions and to reflect only chemosensory functions [135,136]. It was not even remotely possible to imagine at the time (1992–2000) the relationships between p10, CSPs, mucins, cytoskeleton complexes, genetic elements, nuclear pores, and intracellular processes that we discuss here (see Table S2, Figures 4–6, S2 and S3).

Table 1. The family of “ChemoSensory Proteins” (“CSPs”) goes by several names. * Dyanov, H.M.; Lyozin, G.I.; Dzitoeva, S.G.; Korochkin, L.I. A cDNA of *Drosophila melanogaster* ejaculatory bulb specific protein III (PEBme III). Unpublished.

Name	Organism	Authors	Year	Reference
p10	<i>Periplaneta americana</i>	Nomura et al.	1992	[71]
Ebsp-3/PebIII	<i>Drosophila melanogaster</i>	Dyanov et al.	1994	AAA87058*
A10	<i>Drosophila melanogaster</i>	Pikielny et al.	1994	[134]
OS-D	<i>Drosophila melanogaster</i>	McKenna et al.	1994	[133]
Pam	<i>Periplaneta americana</i>	Picimbon & Leal	1999	[4]
CSP	<i>Schistocerca gregaria</i>	Angeli et al.	1999	[5]
SAP	<i>Manduca sexta</i>	Robertson et al.	1999	[135]
Pherokine	<i>Drosophila melanogaster</i>	Sabatier et al.	2003	[27]
Mp10	<i>Myzus persicae</i>	Bos et al.	2010	[45]
LA-BP	<i>Bemisia tabaci</i>	Liu et al.	2016	[30]
Toxin-BP	<i>Bemisia tabaci</i>	Liu et al.	2016	[30]
B-CSP	<i>Acinetobacter baumannii</i>	Liu et al.	2019	[38]
Lipid-BP	<i>Bemisia tabaci</i>	Liu et al.	2020	[31]
JHRP	<i>Aedes aegypti</i>	Picimbon	2020	[20]
Mucin module	<i>Aedes aegypti</i>	Liu et al.	2024	This paper
TIF module	<i>Aedes aegypti</i>	Liu et al.	2024	This paper
DNA/RNA-BP	<i>Aedes aegypti</i>	Liu et al.	2024	This paper

“CSP” was the fourth revision to the nomenclature for p10, which continues to significantly minimize the developmental role in favor of a chemosensory function. However, the existence of “B-

CSPs" (CSPs in bacterial species) raises the question of what potential use these protein molecules may have in chemical detection [31,34,38]. Since there isn't a single piece of evidence indicating CSP molecules interact with sensory neurons, their precise role in chemosensing or olfaction is still hotly debated. There is no evidence to support the theory that CSPs attach to smells and transfer them to ORs—in insects, crustaceans, or, most definitely, bacteria and viruses.

Since 2003, it has been demonstrated that this family of proteins operates independently of the chemosensory system [6,27]. The term "Pherokines" refers to highly prevalent molecules in flies' hemolymph that arise from bacterial, viral, or chemical infection ([27–29]; Table 1), which presents the concept of p10 or CSP molecule in the insect defense system and immune responses. It was eventually proposed to rename these molecules as "Cuticular Sensory Proteins" ("CSPs", [138]) in order to highlight their strong expression in lipid-layered chitin cell walls, first immune barriers, the exoskeleton (to which muscles are attached), the epidermis, epicuticles, major outer surfaces, antennae, eyes, legs, and other sensory organs. Thus, a growing body of research has been conducted on this molecule's protein gene family, reporting functions that go well beyond the sensory branches and the olfactory system. However, none of these studies have addressed the potential role of "CSP" in relation to ARPC, NPCP, mucin, TIF, transcription initiation, RNA, nucleoside kinase, and genetic regulation as our study does (see Table S2 and Figures 4–6, S1 and S3).

What distinguishes p10, OS-D, A10, SAP, and CSP from one another? The most remarkable characteristic of this family of protein molecules, which is expressed throughout the insect's body [31], is entirely overlooked by many synonyms (Table 1). It is necessary to rename the entire molecule's protein gene family, in order to account for all family members across species of arthropods, crustaceans, insects, isopods, and microbes in the bacterial prokaryotic superkingdom [31,34,38]. Additionally, this is required to consider the relationship between CSPs and intracellular mechanisms (Figure 6). Typically, a neuropeptide or hormone molecule is named after the first "physiological" or "pharmacological" effect that is noted for it; however, this naming scheme often proves to be fairly inaccurate. The best course of action is probably to rename all naturally occurring peptides, hormones, carrier proteins, transporters, "CSPs", and related molecules in a more impartial, methodical and objective manner. The location and long list of functions linked to this molecule protein gene family defy the different labels assigned to "SAP" or "CSP" (Table 1 and Figures 4–6, S1 and S3). Developing new terms or nomenclature is required, but they cannot be limited to chemosensory or olfactory structures, assume any illusive functions, and/or be used without enough evidence or proof.

When discussing the entire molecule protein gene family, it becomes somewhat inappropriate or rather awkward to use the terms "SAP" and "CSP", which precisely translate to "Sensory Appendage Protein" and "Chemosensory Protein", expressed in olfactory cilia, respectively. It is not appropriate to use either term —CSP or SAP— to describe all the genes and protein molecules expressed in the antennae, mouth, and legs, as well as in many non-sensory metabolic cell tissues and hemolymph, in the gut and fat body, and all those that have evolved from organisms ranging from bacteria and viruses to insects and crustaceans [[24,31,34], and this study]. There is no need for additional discussion because of the analysis of genome and EST databases. In line with the molecular expression data obtained in moths [24], ESTs of bacteria, insects, sea crustaceans, and terrestrial arthropods show that none of these "CSP" proteins are specifically tuned to olfactory and/or taste chemosensory organs [20,30,31,34]. Actually, based on our analysis, they resemble general intracellular components more (Table S2 and Figures 2–6 and S1–S3).

This situation seems to be similar to that of a superfamily of widely distributed, heterogeneous proteins known as "lipocalins", which are responsible for the transport of small hydrophobic molecules such as lipids and steroids (the term 'lipocalin' comes from the Greek lipos, which means fat, and Greek kalyx, which means cup). Lipocalins are diverse and rather poorly conserved throughout evolution [138]. Unlike lipocalins, the "CSP" family describes a group of rather homogeneous, similar gene structures across a wide range of organisms; they share the same intron boundaries, and evolutionary well-conserved molecular structures (a prism or pyramid with a pattern of four cysteines that is always in the same position), and similar tissue profiling (ubiquitous

expression; see chapters 1-6). Their binding properties differ somewhat (not only to LA, long FAs, and lipid chains, but also to cyclic compounds) [30], but the primary distinction between lipocalins and CSPs is that CSPs connect with TIFs, RNAs, Mucins, DNA-regulatory proteins, SamkC, WAS/WASL-like, CWP, cytoskeleton actin-linkers, and other transcriptional factors that have been discussed thus far (Figures 4–6 and S1–S3). It even appears that “CSPs” have undergone natural selection and evolution in viruses prior to becoming phenotypes (Figures 4–5B) [139], which is a stark contrast to lipocalins.

It is probably best to use a new nomenclature to describe the peptide molecule, taking into account intracellular localization and evolution in microbes. It would be appropriate to refer to CSPs and OBPs as “*lipoclistins*” (from Greek Λίπος lipos=fat and Greek κλειστό kleistó=closed in), as proteins that enclose lipid ligands, much like lipocalins do; citing Rudolf Alexander Steinbrecht [140]; Max-Planck Institute for Biological Intelligence, Seewiesen, Germany). According to this system of naming based on words and names’ Greek ancestry, following the nomenclature of COVID-19, the term “*lip-anoiktins*” (perhaps “*lip-aniktins*”, as oi in modern Greek is pronounced i; derived from Greek words for “fat” and “open”, respectively; open=ανοικτος anoiktos) might be a better way to describe the main structural component of the CSP—an open-air structure—. World Health organization has designated key strains of SARS-CoV-2, the virus that causes COVID-19, with labels that are clear, simple, easy to say, and memorable. The Greek alphabetic names, like “Alpha”, “Beta”, “Gamma”, “Epsilon”, or “Omicron”, will be used going forward. Whatever the nature of the virus, the Latin word “*apertus*” may be a much more appropriate and straightforward term for the CSP protein (following Rudolf Alexander Steinbrecht [141]).

A generic name such as “*arthrolipin*” would be inappropriate for CSP molecule since it is not specific to the arthropod family; it can also be found in viruses, bacteria, and prokaryotic cells ([31,34,38], WP_149730592, Figures 5 and S3). Not to be forgotten, in honor of the early literature and the old name, Alejandro P. Rooney (USDA-ARS, Lubbock, Texas, USA) proposed “4CSP” (*Four Cys Soluble Proteins*) or *CSP/4Cys* without drawing any broad conclusion regarding their physiological task.

8. Concluding Remark

Currently, the nomenclature used to refer to CSPs is utterly incorrect and disjointed. Even worse now, given our current level of knowledge, compared to twenty years ago, when we were limited to thinking about expression in the antennae. It is impossible to classify as “chemosensory” molecules those that are expressed extensively throughout the insect and linked to intracellular systems.

Instead of concentrating only on chemosensing or non-chemosensing, it is critical to investigate the connections among CSP, ASRP, NPCP, Mucin, TIF, RNA-BP, and other cell regulator molecules (refer to Figure 6). Mucins, actin-related proteins, and nuclear complexes are examples of larger molecules or molecular agencies to which CSP binds. This is an intriguing and important phenomenon with implications for both evolutionary and functional issues. If the exon theory of genes—also called the introns-early theory—is correct, then it will depend on how well CSPs bind to mucins and/or what role CSP serves as TIF’s tail. Exons may be the boundaries of separate modules or small structural domains that joined together to form protein molecules at some point in their evolutionary history, according to this theory [142,143]. In our analysis of CSP, TIF, Rho, Mucin, ASRP, NCPC, GpIb, WASP, and a lengthy list of DNA, RNA and nucleotide-binding proteins and gene regulators, we have assessed a broad range of large intracellular molecular structural modules (Table S2 and Figures 4–6 and S1–S3). The fact that the CSP prism builds the N-terminal domain of TIF, Rho, Mucin, ASRP, NCPC, GpIb, and WASP indicates that CSP is a module of larger protein molecules, which supports the exon gene theory but runs counter to the protein’s odor sensing role.

Many insect species, but moths in particular, have long lipid FA chains and CSPs activated for female sex pheromone biosynthesis [14,33–36,144]. On the other hand, exogenous long FA chains have a significant impact on a variety of bacterial processes, such as intracellular signaling and gene expression patterns. DiRusso and Black [143] state that the protein molecules required for the biosynthesis and breakdown of FAs are primarily affected by these transcription patterns. The

current findings are consistent with this placement of the “CSP” molecules, which appear to be found in eukaryotes and prokaryotes as well as in bacteria and viruses.

Moreover, “CSP” exhibits rather broad substrate selectivity, binding transcription regulators, nucleotides, DNA, RNA, and/or other intracellular elements in addition to lipid FA molecules. Therefore, it is imperative that this molecule protein gene family be given a new name. The new name should ideally not allude to any specific function because, even among related groups of molecules, the evidence for a given function can be shaky and subject to too rapid change and contradiction. “CSP”, “OS-D” or “SAP” is one peptide molecule that, in many cases, would be better described or categorized according to its structure rather than just one function. This is particularly true in cases where there is long ongoing disagreement regarding the molecule’s physiological role.

Acknowledgements: Award of the most cited paper of Insect Science 2017 (Increased expression of CSP and CYP genes in adult silkworm females exposed to avermectins; Xuan et al., 2015), and most cited paper of PLoS ONE (Biotype characterization, developmental profiling, insecticide response and binding property of *Bemisia tabaci* chemosensory proteins: role of CSP in insect defense; Liu et al., 2016b, Top 10%). Overseas high-level talent program-Taishan scholar title (#NO.tshw20091015), and Ministry of Science and Technology of China (National Expert, #G2022023033L). We would especially like to thank Prof. Drs. Alejandro Rooney (USDA) and Rudolf Alexander Steinbrecht (MPI) for their insightful and critical conversations regarding the renaming of “CSPs” and for exchanging opinions on the early draft of the manuscript.

Abbreviations

4CSP: 4 Cysteines Soluble Proteins

AA: Arachidonic acid

AcrR: Regulator of adjacent *acrAB* efflux genes - functional protein of the transcriptional regulation system that confers bacterial resistance to the antibiotic tetracycline (see TetR and TFTR) - Regulated by FA lipids

Acypi: *Acyrtosiphum pisum* (pea aphid)

Aedae: *Aedes aegypti* (dengue yellow fever mosquito)

Anoga: *Anopheles gambiae* (malaria mosquito)

Allergen Tha p 1: IgE-binding protein (15 kDa) and major allergen of pine processionary caterpillar (*Thaumetopoea pityocampa*, Thapi) - variant 1.0101

Apime: *Apis mellifera* (honey bee)

Arp2/3: Actin related protein 2/3 complex

ASRC: Actin skeleton regulatory complex

ASRP: Actin skeleton regulatory protein

Avd: Accessory variability determinant

Bacil: *Bacillus*

B-CSP: Bacterial “Chemosensory protein”

BioNJ: Bio (improved version) of the Neighbor Joining algorithm based on simple model of sequence data

BLASTn: Nucleotide BLAST

BLLF1: Epstein-Barr virus envelope glycoprotein encoded by BLLF1 gene

Bommo: *Bombyx mori* (silkworm moth)

BemtaCSP1: *Bemisia tabaci* “Chemosensory Protein”-1 (LA-binding protein)

CA: *Corpora allata*

Cocse: *Coccinella septempunctata* (seven-spot ladybird)

CSP: Chemosensory protein

Culpi: *Culex pipiens* (common house mosquito)

CWA: Cell wall anchored

CWP: Cell wall protein

CYP: Cytochrome P450

Cys: Cysteine

DAN4: Cell wall mannoprotein expressed under delayed anaerobic conditions (*Saccharomyces*)

Danpl: *Danaus plexippus* (monarch butterfly)

Dd-Rp: DNA-dependent RNA polymerase

DE: Degradative enzyme

DGR: Diversity generating retroelement

DNA-BP: Deoxyribonucleic acid-binding protein

DNA-RP: Deoxyribonucleic acid-regulatory protein

DP: Diphosphate

Drome: *Drosophila melanogaster* (fruit fly)

Droso: *Drosophila* species

EA: Elaidic acid

Ebsp: Ejaculatory bulb-specific protein

Epiba: *Episyrphus balteatus* (marmalade hoverfly)

ER: Endoplasmic reticulum

EST: Expressed Sequence Tag

Eupco: *Eupeodes corollae* (migrant overfly)

FA: Fatty acid

GC: Guanine-cytosine content

GMQE: Global Model Quality Estimate

GTP: Guanosine triphosphate

HSV: Herpesvirus

IgE-BP: Immunoglobulin E-binding protein

Iphpo: *Iphiclides podalirius* (scarce swallowtail)

Jg5928: major outer envelope protein

JH: Juvenile hormone

JHRP: Juvenile hormone-related protein

HSV: Herpesvirus

LA: Linoleic acid (C18:2)

Locmi: *Locusta migratoria* (migratory locust)

LRR: Leucine-rich repeat protein complex

M: Mucin

Manse: *Manduca sexta* (tobacco hornworm, sphinx moth)

MP: Maximum parsimony

Mp10: Myzpe p10-like protein

Myzpe: *Myzus persicae* (green peach aphid)

NNK: Nuclear nucleoside kinase

NDPK: Nucleoside diphosphate kinase

NPCP: Nuclear pore complex protein

Nasvi: *Nasonia vitripennis* (parasitoid jewel wasp)

OBP: Odorant-binding protein

OR: Olfactory receptor

OS-D: Olfactory sensilla-type D protein

P: Phosphate

P10: Peram 10 kDa protein from regenerating legs

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)

Peram: *Periplaneta americana* (American cockroach)

PBP: Pheromone-binding protein

PgIb: Platelet glycoprotein Ib alpha chain-like

Phk: Pherokine (hemolymph protein)

PKC: Protein kinase C

PLC: Phospholipase C

Pol: Polymerase

Ras: Family of GTPases derived from rat sarcoma virus

Rho: Family of GTPases, family of small (~21 kDa) signaling G proteins, subfamily of the Ras superfamily

RhoGAP: Rho GTPase-activating protein, regulator of the Rho-related protein family, crucial in many cellular processes, motility, contractility, growth, differentiation, and development

RickA: *Rickettsia (conorii)* surface protein A (activator of Arp2/3)

RNA-BP: Ribonucleic acid-binding protein

SamkC: Serine/Threonine-protein kinase (from Social amoeba)

SA: Stearic acid

SAP: Sensory appendage protein

Sec31: Protein transport protein encoded by the SEC31A gene (human)

SF: Splicing factor

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

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 (or 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

- Hoffmann, K.H.; Meyerinng-Vos, M.; Lorenz, M.W. Allatostatins and allatotropins: is the regulation of *corpora allata* activity their primary function? *Eur. J. Entomol.* **1999**, *96*, 255-266.
- Picimbon, J.F. Renaming *Bombyx mori* chemosensory proteins. *Int. J. Bioorg. Chem. Mol. Biol.* **2014**, *2*, 1-4.
- Vogt, R.G.; Riddiford, L.M. Pheromone binding and inactivation by moth antennae. *Nature*, **1981**, *293*, 161-163.
- Picimbon, J.F.; Leal, W.S. Olfactory soluble proteins of cockroaches. *Insect Biochem. Mol. Biol.* **1999**, *30*, 973-978.
- 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.
- Picimbon, J.F. Biochemistry and evolution of CSP and OBP proteins. In *Insect Pheromone Biochemistry and Molecular Biology, The Biosynthesis and Detection of Pheromones and Plant Volatiles*; Blomquist, G.J., Vogt, R.G., Eds; Elsevier Academic Press, London & San Diego, UK & USA, 2003; pp. 539-566.
- 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.
- Jansen, S.; Zidek, L.; Löfstedt, C.; Picimbon, J.F.; Sklenar, V. ¹H, ¹³C, and ¹⁵N resonance assignment of *Bombyx mori* chemosensory protein 1 (BmorCSP1). *J. Biomol. NMR* **2006**, *36*, 47.
- Jansen, S.; Chmelik, J.; Zidek, 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.
- Tomaselli, S.; Crescenzi, O.; Sanfelice, D.; Ab, E.; Wechsel-berger, R.; Angeli, S.; Scaloni, A.; Boelens, R.; Tancredi, T.; Pelosi, P.; Picone, D. Solution structure of a chemosensory protein from the desert locust *Schistocerca gregaria*. *Biochemistry* **2006**, *45*, 1606-1613.
- 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.
- Campanacci, V.; Lartigue, A.; Hällberg, B.M.; 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.
- 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.
- 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.; Leppik, G.; Kasvandik, S.; Picimbon, J.F. Molecular evidence of RNA editing in the *Bombyx* chemosensory protein family. *PLoS ONE* **2014**, *9*, e86932.
- Picimbon, J.F. Evolution of protein physical structures in insect chemosensory systems. In *Olfactory Concepts of Insect Control-Alternative to Insecticides*; Picimbon, J.F., Ed; Vol. 2, Springer Nature AG, Cham, Switzerland, 2019; pp. 231-263.
- Picimbon, J.F.; Dietrich, K.; Breer, H.; Krieger, J. Chemosensory proteins of *Locusta migratoria* (Orthoptera: Acrididae). *Insect Biochem. Mol. Biol.* **2000**, *30*, 233-241.
- 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*, 120-129.
- 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.
- 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.
- Picimbon, J.F. Chapter three—bioinformatic, genomic and evolutionary analysis of genes: a case study in Dipteran CSPs. *Meth. Enzymol.* **2020**, *642*, 35-79.
- Guo, W.; Wang, X.; Ma, Z.; Xue, L.; Han, J.; Yu, D.; Kang, L. CSP and *Takeout* genes modulate the switch between attraction and repulsion during behavioral phase change in the migratory locust. *PLoS Genet.* **2011**, *7*, e1001291.
- Martín-Blázquez, R.; Chen, B.; Kang, L.; Bakkali, M. Evolution, expression and association of the chemosensory protein genes with the outbreak phase of the two main pest locusts. *Sci. Rep.* **2018**, *7*, 6653.

23. Ban, L.; Scaloni, A.; Brandazza, A.; Angeli, S.; Zhang, Y.; Pelosi, P. Chemosensory proteins of *Locusta migratoria*. *Insect Mol. Biol.* **2003**, *12*, 125-134.
24. 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.
25. Li, S.; Picimbon, J.F.; Ji, S.; Kan, Y.; Chuanling, Q.; Zhou, J.J.; Pelosi, P. Multiple functions of an odorant-binding protein in the mosquito *Aedes aegypti*. *Biochem. Biophys. Res. Commun.* **2008**, *372*, 464-468.
26. 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.* **2020**, *9*, 8510.
27. 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.
28. Liu, G.X. et al. Biotype expression and insecticide response of *Bemisia tabaci* chemosensory protein-1. *Arch. Insect Biochem. Physiol.* **2014**, *85*, 137-151.
29. 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; Vol. 2 Springer Nature AG, Cham, Switzerland, 2019; pp. 205-229.
30. Liu, G.X.; 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.
31. [Liu, G.X.; 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.
32. Picimbon J.F. RNA mutations: source of life. *Gene Technol.* **2014**, *3*, 2.
33. Xuan, N.; Rajashekar, B.; Picimbon, J.F. DNA and RNA-dependent polymerization in editing of *Bombyx* chemosensory protein (CSP) gene family. *Agri Gene* **2019**, *12*, 100087.
34. Picimbon, J.F. RNA + ribosome peptide editing in chemosensory proteins (CSPs), a new theory for the origin of life on Earth's crust. *J. Mol. Evol.* **2024**, submitted.
35. Xuan, N.; Rajashekar, B.; Kasvandik, S.; Picimbon, J.F. Structural components of chemosensory protein mutations in the silkworm moth, *Bombyx mori*. *Agri Gene* **2016**, *2*, 53-58.
36. Picimbon, J.F. A new view of genetic mutations. *Australas. Med. J.* **2017**, *10*, 701-715.
37. Liu, G.X.; Arnaud, P.; Offmann, B.; Picimbon, J.F. Genotyping and bio-sensing chemosensory proteins in insects. *Sensors* **2017**, *17*, 1801.
38. Liu, G.X.; Yue, S.; Rajashekar, B.; Picimbon, J.F. Expression of chemosensory protein (CSP) structures in *Pediculus humanus corporis* and *Acinetobacter baumannii*. *SOJ Microbiology and Infectious Diseases* **2019**, *7*, 1-17.
39. Picimbon, J.F. Synthesis of odorant reception-suppressing agents: odorant binding proteins (OBPs) and Chemosensory Proteins (CSPs) as molecular targets for pest management. In *Biopesticides of plant origin*; Regnault-Roger, C., Philogène, B., Vincent, C., Eds; Intercept Ltd, Hampshire, UK, 2005; pp.245-266.
40. Zhu, J.; Iovinella, I.; Dani, F.R.; Pelosi, P.; Wang, G. Chemosensory proteins: A versatile binding family. In *Olfactory Concepts of Insect Control-Alternative to Insecticides*; Picimbon, J.F., Ed; Vol. 2, Springer Nature AG, Cham, Switzerland, 2019; pp. 147-169.
41. Liu, G.X.; Picimbon, J.F. Bacterial origin of insect chemosensory odor-binding proteins. *Gene Transl. Bioinf.* **2017**, *3*, e1548.
42. Ghai, R.; Mizuno, C.M.; Picazo, A.; Camacho, A.; Rodriguez-Valera, F. Metagenomics uncovers a new group of low GC and ultra-small marine Actinobacteria. *Scientific Reports* **2013**, *3*, 2471.
43. Zhao, Z.; Gonsior, M.; Schmitt-Kopplin, P.; Zhan, Y.; Zhang, R.; Jiao, N.; Chen, F. Microbial transformation of virus-induced dissolved organic matter from picocyanobacteria: coupling of bacterial diversity and DOM chemodiversity. *ISME J.* **2019**, *13*, 2551-2565.
44. Doré, H.; Guyet, U.; Leconte, J.; Farrant, G.K.; Alric, B.; Ratin, M.; Ostrowski, M.; Ferrieux, M.; Brillet-Guéguen, L.; Hoebeke, M.; Siltanen, J.; Le Corguillé, G.; Corre, E.; Wincker, P.; Scanlan, D.J.; Eveillard, D.; Partensky, F.; Garczarek, L. Differential global distribution of marine picocyanobacteria gene clusters reveals distinct niche-related adaptive strategies. *ISME J.* **2023**, *17*, 720-732.
45. Bos, L.D.J.; Sterk, P.J.; Schultz, M.J. Volatile metabolites of pathogens: A systemic review. *PLoS Pathog.* **2013**, *9*, e1003311.

46. Taga, M.E.; Bassler, B.L. Chemical communication among bacteria. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 14549-14554.
47. Silva-Junior, E.A.; Ruzzini, A.C.; Paludo, C.R.; Nascimento, F.S.; Currie, C.R.; Clardy, J.; Pupo, M.T. Pyrazines from bacteria and ants: convergent chemistry within an ecological niche. *Scientific Reports* **2018**, *8*, 2595.
48. Perkin, L.C.; Friesen, K.S.; Flinn, P.W.; Oppert, B. Venom gland components of the ectoparasitoid wasp, *Anisopteromalus calandrae*. *J. Venom Res.* **2015**, *6*, 19-37.
49. Picimbon J.F. RNA mutations in the moth pheromone gland. *RNA Dis.* **2014b**, *1*, e240.
50. Celorio-Mancera, Mde L.; Sundmalm, S.M; Vogel, H., Rutishauser, D.; Ytterberg, A.J.; Zubarv, R.A.; Janz, N. Chemosensory proteins, major salivary factors in caterpillar mandibular glands. *Insect Biochem. Mol. Biol.* **2012**, *42*, 796-805.
51. González-Caballero, N.; Valenzuela, J.G.; Ribeiro, J.M.C.; Cuervo, P.; Brazil, R.P. Transcriptome exploration of the sex pheromone gland of *Lutzomyia longipalpis* (Diptera: Psychodidae: Phlebotominae). *Parasit. Vectors* **2013**, *6*, 56.
52. Liu, Y.L.; Guo, H., Huang, L.Q., Pelosi, P. and Wang, C.Z. (2014) Unique function of a chemosensory protein in the proboscis of two *Helicoverpa* species. *J. Exp. Biol.* **2014**, *217*, 1821-1826.
53. Zhu, J.; Iovinella, I.; Dani, F.R.; Liu, Y.L.; Huang, L.Q.; Liu, Y.; Wang, C.Z.; Pelosi, P.; Wang, G. Conserved chemosensory proteins in the proboscis and eyes of Lepidoptera. *Int. J. Biol. Sci.* **2016**, *12*, 1394-1404.
54. Pedra, J.H.; Brandt, A.; Li, H.M.; Westerman, R.; Romero-Serverson, J.; Pollack, R.J.; Murdock, L.L.; Pittendrigh, B.R. Transcriptome identification of putative genes involved in protein catabolism and innate immune response in human body louse (Pediculicidae: *Pediculus humanus*). *Insect Biochem. Mol. Biol.* **2003**, *33*, 1135-1143.
55. Kirkness, E.F. et al. Genome sequences of the human body louse and its primary endosymbiont provide insights into the permanent parasitic life. *Proc. Natl Acad. Sci. USA* **2010**, *107*, 12168-12173.
56. Shigenobu, S.; Richards, S.; Cree, A.G.; Morioka, M.; Fukatsu, T.; Kudo, T.; Miyagishima, S.; Gibbs, R.A.; Stern, D.L.; Nakabashi, A. A full-length cDNA resource for the pea aphid, *Acyrtosiphon pisum*. *Insect Mol. Biol.* **2010**, *19*, 23-31.
57. Pittendrigh, B.R.; Clark, J.M.; Lee, S.H.; Yoon, K.S.; Sun, W.; Steele, L.D.; Seong, K.M. Body lice: from the genome project to functional genomics and reverse genetics. In *Short views on insect genomics and proteomics, Entomology in focus*; Raman, C., Goldsmith, M., Agunbiade, T., Eds; Springer, Cham, Switzerland, 2015 ; pp. 1-18.
58. Ollivier, M.; Legeai, F.; Rispe, C. Comparative analysis of the *Acyrtosiphon pisum* genome and expressed sequence tag-based gene sets from other aphid species. *Insect Mol. Biol.* **2019**, *19*, 33-45.
59. Wanner, K.W.; Willis, L.G.; Theilmann, D.A.; Isman, M.B.; Feng, Q.; Plettner, E. Analysis of the insect os-d-like gene family. *J. Chem. Ecol.* **2004**, *30*, 889-911.
60. Forêt, S.; Wanner, K.W.; Maleszka R. Chemosensory proteins in the honeybee: Insights from the annotated genome, comparative analysis and expression profiling. *Insect Biochem. Mol. Biol.* **2007**, *37*, 19-28.
61. Blomquist, G.J.; Tittiger, C.; Jurenka, R. Cuticular hydrocarbons and pheromones of arthropods. In *Oils and Lipids: Diversity, Origin, Chemistry and Fate*; Wilkes, H., Ed; Hydrocarbons, Handbook of Hydrocarbon and Lipid Microbiology, Springer, Cham, Swizerland, 2018; pp. 1-32.
62. Maleszka, J, Forêt, S., Saint, R., Maleszka, R. RNAi-induced phenotypes suggest a novel role for a chemosensory protein CSP5 in the development of embryonic integument in the honeybee (*Apis mellifera*). *Dev. Genes Evol.* **2007**, *217*, 189-196.
63. Spealman, P.; Burrelli, J.; Gresham, D. Inverted duplicate DNA sequences increase translocation rates through sequencing nanopores resulting in reduced base calling accuracy. *Nucl. Acids Res.* **2020**, *48*, 4940-4945.
64. Blackman, R.L.; Takada, H.; Kawakami, K. Chromosomal rearrangement involved in insecticide resistance of *Myzus persicae*. *Nature* **1978**, *271*, 450-452.
65. Mandrioli, M.; Melchiori, G.; Panini, M.; Chiesa, O.; Giordano, R.; Mazzoni, E.; Manicardi, G.C. Analysis of the extent of synteny and conservation in the gene order in aphids: a first glimpse from the *Aphis glycines* genome. *Insect Biochem. Mol. Biol.* **2019**, *113*, 103228.
66. Mathers, T.C.; Wouters, R.H.M.; Mugford, S.T.; Swarbreck, D.; van Oosterhout, C.; Hogenhout, S.A. Chromosome-scale genome assemblies of aphids reveal extensively rearranged autosomes and long-term conservation of the X chromosome. *Mol. Biol. Evol.* **2021**, *38*, 856-875.

67. Liu, G.X.; Ma, H.M.; Xie, H.Y.; Xuan, N.; Picimbon, J.F. Sequence variation of *Bemisia tabaci* Chemosensory protein 2 in cryptic species B and Q: new DNA markers for whitefly recognition. *Gene* **2016**, *576*, 284-291.
68. Blobel, G. Protein targeting (Nobel lecture). *Chembiochem.* **2000**, *1*, 86-102.
69. Matlin, K.S. The strange case of the signal recognition particle. *Nat. Rev. Mol. Cell Biol.* **2002**, *3*, 538-542.
70. Doudna, J.A.; Batey, R.T. Structural insights into the signal recognition particle. *Annu. Rev. Biochem.* **2004**, *73*, 539-557.
71. Nomura, A.; Kawasaki, K. ; Kubo, T. ; Natori, S. Purification and localization of p10, a novel protein that increases in nymphal regenerating legs of *Periplaneta americana* (American cockroach). *Int. J. Dev. Biol.* **1992**, *36*, 391-398.
72. 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.
73. Schneider, E.S.; Kleineidam, C.J.; Leitingner, G.; Römer, H. Ultrastructure and electrophysiology of thermosensitive sensilla coeloconica in a tropical katydid of the genus *Mecopoda* (Orthoptera, Tettigoniidae). *Arthr. Struct. Dev.* **2018**, *47*, 482-497.
74. Loftus, B.J.; Utterback, T.; Perte, 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. *NCBI* **2005**, #DV263125, DV263127, DV289920, DV289921, DV297938, DV314711, DV314712, DV316589, DV316619, DV334734, DV334735, DV335057, DV335058, DV344048, DV347318, DV347319, DV365747, DV365766, DV357763, DV368339, DV368340, DV393559, DV400785, DV400787, "...".
75. Verjovski-Almeida, S.; Eiglmeier, K.; El-Dorri, H.; Gomes, S.L.; Menck, S.F.M.; Nascimento, A.L.; Roth, C.W. FAPESP and Institut Pasteur/AMSUD Network *Aedes aegypti* cDNA sequencing project. *NCBI* **2005**, #EG001037.
76. 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.
77. Nene, V.; Worthman, J.R.; Lawson, D. et al. Genome sequence of *Aedes aegypti*, a major arbovirus vector. *Science* **2007**, *22*, 1718-1723.
78. 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.
79. Laufer, H.; Borst, D.; Baker, F.C.; Reuter, C.C.; Tsai, L.W.; Schooley, D.A.; Carrasco, C.; Sinkus, M. Identification of a juvenile-hormone-like compound in a Crustacean. *Science* **1987**, *235*, 202-205.
80. Jindra, M.; Palli, S.R.; Riddiford, L.M. The juvenile hormone signaling pathway in insect development. *Annu. Rev. Entomol.* **2013**, *58*, 181-204.
81. Cayre, M.; Strambi, C.; Strambi, A. Neurogenesis in an adult insect brain and its hormonal control. *Nature* **1994**, *368*, 57-59.
82. Anton, S.; Rossler, W. Plasticity and modulation of olfactory circuits in insects. *Cell Tiss. Res.* **2021**, *383*, 149-164.
83. Bian, H.X.; Chen, D.B.; Zheng, X.X.; Ma, H.F.; Li, Y.P.; Li, Q.; Xia, R.X.; Wang, H.; Jiang, Y.R.; Liu, Y.Q.; Qin, L. Transcriptomic analysis of the prothoracic gland from two lepidopteran insects, domesticated silkworm *Bombyx mori* and wild silkworm *Antheraea pernyi*. *Sci. Rep.* **2019**, *9*, 5313.
84. Ahmed, S.; Hrithik, M.T.H.; Roy, M.C.; Bode, H.; Kim, Y. Phurelipids, produced by the entomopathogenic bacteria, *Photorhabdus*, mimic juvenile hormone to suppress insect immunity and immature development. *J. Invertebr. Pathol.* **2022**, *193*, 107799.
85. 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.
86. 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.
87. 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.

88. Moon, S.Y.; Zheng, Y. (2003) Rho GTPase-activating proteins in cell regulation. *Trends Cell Biol.* **2003**, *13*, P13-22.
89. 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.
90. 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.
91. Ho, S. The molecular clock and estimating species divergence. *Nat. Educ.* **2008**, *1*, 168.
92. Wang, P.; Granados, R.R. An intestinal mucin is the target substrate for a baculovirus enhancin. *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 6977-6982.
93. Dias, R.O.; Cardoso, C.; Pimentel, A.C.; Damasceno, T.F.; Ferreira, C.; Terra, W.R. The roles of mucus-forming mucins, peritrophins and peritrophins with mucin domains in the insect midgut. *Insect Mol. Biol.* **2018**, *27*, 46-60.
94. 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.
95. Abraham, D.; Löfstedt, C.; Picimbon, J.-F. Molecular characterization and evolution of pheromone binding protein genes in *Agrotis* moths. *Insect Biochem. Mol. Biol.* **2005**, *35*, 1100-1111.
96. Armstrong, L.; Lako, M.; van Herpe, I.; Evans, J.; Saretzki, G.; Hole, N. A role for nucleotide Zap3 in the reduction of telomerase activity during embryonic stem cell differentiation. *Mech. Dev.* **2004**, *121*, 1509-1522.
97. Teparic, R.; Lozancic, M.; Mrsa, V. Evolutionary overview of molecular interactions and enzymatic activities in the yeast cell walls. *Int. J. Mol. Sci.* **2020**, *21*, 8996.
98. McQuarrie, D.W.J.; Read, A.M.; Stephens, F.H.S.; Civetta, A.; Soller, M. Indel driven rapid evolution of core nuclear pore protein gene promoters. *Sci. Rep.* **2023**, *13*, 8035.
99. DeGrasse, J.A.; DuBois, K.N.; Devos, D.; Siegel, T.N.; Sali, A.; Field, M.C.; Rout, M.P.; Chait, B.T. Evidence for a shared nuclear pore complex architecture that is conserved from the Last Common Eukaryotic Ancestor. *Mol. Cell. Prot.* **2009**, *8*, 2119-1230.
100. 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.
101. Waterhouse, A.; Bertoni, M.; Bienert, S.; Studer, G.; Tauriello, G.; Gumienny, R.; Heer, F.T.; de Beer, T.A.P.; Rempfer, C.; Bordoli, L.; Lepore, R.; Schwede, T. SWISS-MODEL: homology modelling of protein structures and complexes. *Nuc. Acids Res.* **2018**, *46*, W296-W303.
102. 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.
103. 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*, 16644-1644449.
104. Turcotte, S.; Letellier, J.; Lippé, R. Herpes simplex virus type 1 capsids transit by the trans-Golgi network, where viral glycoproteins accumulate independently of capsid egress. *J. Virol.* **2005**, *79*, 8847-8860.
105. Rogerio, A.P.; Anibal, F.F. Role of leukotrienes on protozoan and helminth infections. *Mediators Inflamm.* **2012**, *2012*, 595694.
106. Roy, M.C.; Nam, K.; Kim, J.; Stanley, D.; Kim, Y. Thromboxane mobilizes insect blood cells to infection foci. *Front. Immunol.* **2021**, *12*, 791319.
107. Stanley, D.; Kim, Y. Chapter Eight – Insect prostaglandins and other eicosanoids: from molecular to physiological actions. *Adv. Insect Physiol.* **2019**, *56*, 283-343.
108. Blank, H.M.; Maitra, N.; Polymenis, M. Lipid biosynthesis: When the cell cycle meets protein synthesis? *Cell Cycle* **2017**, *16*, 905-906.
109. 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 Journal* **2017**, *36*, 487-502.
110. Feyereisen, R. Insect P450 enzymes. *Annu. Rev. Entomol.* **1999**, *44*, 507-533.
111. Fujiwara, Y.; Wada, K.; Kabuta, T. Lysosomal degradation of intracellular nucleic acids—multiple autophagic pathways. *J. Biochem.* **2017**, *161*, 145-154.

112. Yokoyama, N.; Fónagy, A.; Tatsuki, S.; Arie, T.; Yamashita, S.; Matsumoto, S. Ultrastructural studies on the pheromone-producing cells in the silkworm, *Bombyx mori*: formation of cytoplasmic lipid droplets before adult eclosion. *Acta Biol. Hung.* **2003**, *54*, 299-311.
113. Hull, J.; Fonagy, A. Molecular basis of pheromonogenesis regulation in moths. In *Olfactory Concepts of Insect Control-Alternative to Insecticides*; Picimbon, J.F., Ed; Vol. 1, Springer Nature AG, Cham, Switzerland, 2019; pp. 115-202.
114. Ibarra, A.; Hetzer, M.W. Nuclear pore proteins and the control of genome functions. *Genes Dev.* **2015**, *29*, 337-349.
115. Lin, D.H.; Hoelz, A. The structure of the Nuclear Pore Complex (an update). *Annu. Rev. Biochem.* **2019**, *88*, 725-783.
116. Kenis, S.; Istiban, M.N.; Van Damme, S.; Vandewyler, E.; Watteyne, J.; Schoofs, L.; Beets, I. Ancestral glycoprotein hormone-receptor pathway controls growth in *C. elegans*. *Front. Endocrinol.* **2023**, *14*, 1200407.
117. Kornblihtt, A.R.; de la Mata, M.; Fededa, J.P.; Muñoz, M.J.; Nogués, G. Multiple links between transcription and splicing. *RNA* **2004**, *10*, 1489-98.
118. Georgescauld, F.; Song, Y.; Dautant, A. Structure, folding and stability of nucleoside diphosphate kinases. *Int. J. Mol. Sci.* **2020**, *21*, 6779.
119. Shangguan, X.; Zhang, J.; Liu, B.; Zhao, Y.; Wang, H.; Wang, Z.; Guo, J.; Rao, W.; Jing, S.; Guan, W.; Ma, Y.; Wu, Y.; Hu, L.; Chen, R.; Du, B.; Zhu, L.; Yu, D.; He, G. A mucin-like protein of planthopper is required for feeding and induces immunity responses in plants. *Plant Physiol.* **2018**, *176*, 552-565.
120. 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.
121. Shalaeva, D.N.; Galperin, M.Y.; Mulikidjanian, A.Y. Eukaryotic G protein-coupled receptors as descendants of prokaryotic sodium-translocating rhodopsins. *Biol. Direct* **2015**, *10*, 63.
122. Sperling, P.; Ternes, P.; Zank, T.K.; Heinz, E. The evolution of desaturases. *Prostaglandins Leukot. Essent. Fatty Acids* **2003**, *68*, 73-95.
123. Los, D.A.; Murata, N. Chapter 10 – Sensing and responses to low temperature in cyanobacteria. In *Cell and Molecular Responses to Stress*; Storey, K.B., Storey, J.M., Eds; Vol. 3, Elsevier, ScienceDirect, Amsterdam, Netherlands, 2002; pp. 139-153.
124. Rock, C.O. Chapter 3 – Fatty acid and phospholipid metabolism in prokaryotes. In *Biochemistry of Lipids, Lipoproteins and Membranes Fifth Edition*; Vance, D.E., Vance, J.E., Eds; Elsevier Science, Amsterdam, Netherlands, 2008; pp. 59-96.
125. Njenga, R.; Boele, J.; Öztürk, Y.; Koch, H.G. Coping with stress: How bacteria fine-tune protein synthesis and protein transport. *J. Biol. Chem.* **2023**, *299*, 105163.
126. Moto, K.; Suzuki, M.G.; Hull, J.J.; Kurata, R.; Takahashi, S.; Yamamoto, M.; Okano, K.; Imai, K.; Ando, T.; Matsumoto, S. Involvement of a bifunctional fatty-acyl desaturase in the biosynthesis of the silkworm, *Bombyx mori*, sex pheromone. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 8631-8636.
127. Damude, H.G.; Zhang, H.; Farrall, L.; Ripp, K.G.; Tomb, J.F.; Hollerbach, D.; Yadav, N.S. Identification of bifunctional delta12/omega3 fatty acid desaturases for improving the ratio of omega3 to omega6 fatty acids in microbes and plants. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 9446-9451.
128. Gallego-García, A.; Monera-Girona, A.J.; Pajares-Martínez, E.; Bastida-Martínez, E.; Pérez-Castaño, R.; Iniesta, A.A.; Fontes, M.; Padmanabhan, S.; Elías-Arnanz, M. A bacterial light response reveals an orphan desaturase for human plasmalogen synthesis. *Science* **2019**, *366*, 128-132.
129. Siroli, L.; Braschi, G.; Rossi, S.; Gottardi, D.; Patrignani, F.; Lanciotti, R. *Lactobacillus paracasei* A13 and high-pressure homogenization stress response. *Microorganisms* **2020**, *8*, 439.
130. Scheffers, D.J.; Pinho, M.G. Bacterial cell wall synthesis: new insights from localization studies. *Microbiol. Mol. Biol. Rev.* **2005**, *69*, 585-607.
131. Wang, L.; Li, A.; Fang, J.; Wang, Y.; Chen, L.; Qiao, L.; Wang, W.W. Enhanced cell wall and cell membrane activity promotes heat adaptation of *Enterococcus faecium*. *Int. J. Mol. Sci.* **2023**, *24*, 11822.
132. Ikegami, A.; Honma, K.; Sharma, A.; Kuramitsu, H.K. Multiple functions of the leucine-rich repeat protein LrrA of *Treponema denticola*. *Infect. Immun.* **2004**, *72*, 4619-4627.
133. Hu, Y.; Huang, H.; Hui, X.; Cheng, X.; White, A.P.; Zhao, Z.; Wang, Y. Distribution and evolution of *Yersinia* Leucine-Rich Repeat proteins. *Infect. Immun.* **2016**, *84*, 2243-2254.
134. Browning, D.; Busby, S. The regulation of bacterial transcription initiation. *Nat. Rev. Microbiol.* **2004**, *2*, 57-65.

135. 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.
136. 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*, 35-49.
137. Robertson, H.M.; Martos, R.; Sears, C.R.; Todres, E.Z.; Walden, K.K.; Nardi, J.B. Diversity of odourant binding proteins revealed by an expressed sequence tag project on male *Manduca sexta* moth antennae. *Insect Mol. Biol.* **1999**, *8*, 501-518.
138. Ingham, V.A.; Anthousi, A.; Douris, V.; Harding, N.J.; Lycett, G.; Morris, M.; Vontas, J.; Ranson, H. A sensory appendage protein protects malaria vectors from pyrethroids. *Nature* **2020**, *577*, 376-380.
139. Picimbon, J.F.; Regnault-Roger, C. Composés sémiocchimiques volatils, phytoprotection et olfaction: cibles moléculaires de la lutte intégrée. In *Biopesticides d'Origine Végétale*; Regnault-Roger, C., Philogène, B., Vincent, C., Eds; Lavoisier Tech & Doc, Paris, France, 2008; pp. 383-415.
140. Diez-Hernando, S.; Ganfornina, M.D.; Skerra, A.; Gutiérrez, G.; Sanchez, D. An evolutionary perspective of the lipocalin protein family. *Front. Physiol.* **2021**, *12*, <https://doi.org/10.3389/fphys.2021.718983>.
141. Moelling, K.; Broecker, F. Viruses and Evolution – Viruses First? A Personal Perspective. *Front. Microbiol., Sec. Virol.* **2019**, *10*, <https://doi.org/10.3389/fmicb.2019.00523>
142. Steinbrecht, R.A. What can we learn from localizing lipoclistins? *European Symposium For Insect Taste and Olfaction (ESITO VIII)* **2003**, July 2-7th, Harstad, Norway.
143. Steinbrecht, R.A. Fine structure immunocytochemistry—An important tool for research on odorant-binding proteins. *Meth. Enzymol.* **2020**, *642*, 259-278.
144. Gilbert, W. The exon theory of genes. *Cold Spring Harb. Symp. Quant. Biol.* **1987**, *52*, 901-905.
145. Roy, S.W. Recent evidence for the exon theory of genes. *Genetica* **2003**, *118*, 251-266.
146. Ohnishi, A.; Hashimoto, K.; Imai, K.; Matsumoto, S. Functional characterization of the *Bombyx mori* fatty acid transport protein (BmFATP) within the silkworm pheromone gland. *J. Biol. Chem.* **2009**, *284*, 5128-5136.
147. DiRusso, C.C.; Black, P.N. Bacterial long chain fatty acid transport: gateway to a fatty acid-responsive signaling system. *J. Biol. Chem.* **2004**, *279*, 49563-49566.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.