Article

First Insights into the Repertoire of Secretory Lectins in Rotifers

Marco Gerdol

Department of Life Sciences, University of Trieste, Via Giorgieri 5, 34128 Trieste, Italy; mgerdol@units.it

Abstract: Due to their high biodiversity and adaptation to a mutable and challenging environment, aquatic lophotrochozoan animals are regarded as a virtually unlimited source of bioactive molecules. Among these, lectins, i.e., proteins with remarkable carbohydrate-recognition properties involved in immunity, reproduction, self/nonself recognition and several other biological processes, are particularly attractive targets for biotechnological research. To date, lectin research in the Lophotrochozoa has been restricted to the most widespread phyla, which are the usual targets of comparative immunology studies, such as Mollusca and Annelida. Here we provide the first overview of the repertoire of the secretory lectin-like molecules encoded by the genomes of six target rotifer species: *Brachionus calyciflorus*, *Brachionus plicatilis*, *Proales similis* (class Monogononta), *Adineta ricciae*, *Didymodactylos carnosus* and *Rotaria sordida* (class Bdelloidea). Overall, while rotifer secretory lectins display a high molecular diversity and belong to nine different structural classes, their total number is significantly lower than for other groups of lophotrochozoans, with no evidence of lineage-specific expansion events. Considering the high evolutionary divergence between rotifers and the other major sister phyla, their widespread distribution in aquatic environments and the ease of their collection and rearing in laboratory conditions, these organisms may represent interesting targets for glycobiological studies, which may allow the identification of novel carbohydrate-binding proteins with peculiar biological properties.

Keywords: rotifera; pattern recognition receptors; microbe-associated molecular patterns; innate immunity; C-type lectins; C1q domain-containing proteins; galectins

1. Introduction

Lectin-like molecules play a fundamental role in several physiological processes shared by all animals, including, critically, the discrimination between "self" and "nonself" through the specific recognition of carbohydrate moieties exposed on cellular surfaces. These glycans, when associated with microorganisms, are generally referred to as microbe-associated molecular patterns (MAMPs) or, in the case of potentially pathogenic microbes, pathogen-associated molecular patterns (PAMPs) [1].

The proteins expressed by the host that are involved in carbohydrate recognition are collectively known as pattern recognition receptors (PRRs), which may exert their function at different levels, i.e., in the extracellular environment, at the plasma membrane or within the cell. In the context of immune response, the activity of a heterogeneous group of small secretory PRRs usually leads to the coating of invading microbes. This process may in turn trigger a complex response involving several additional molecular and cellular players which vary widely along the metazoan tree of life. These include, among others, the melanization cascade (typically observed in arthropods and other invertebrates) [2], the production of a large arsenal of antimicrobial peptides [3], the activation of the complement system (well-described in vertebrates and present in a primitive form also in many invertebrates) [4] and the recruitment of specialized phagocytic cells [5]. Furthermore, the ability to recognize MAMPs and to modulate immune responses has been linked with the maintenance of gut microbiome homeostasis [6], as well as the establishment of beneficial bacterial symbioses [7], which are particularly relevant in aquatic environments [8].

Besides their key role in immune recognition, lectins are involved in a number of other physiological processes, to which they contribute thanks to their remarkable ability to recognize glycan moieties with high specificity. For example, some lectins play an important role in reproduction and gamete recognition [9,10], in the clearance of apoptotic cells thanks to the recognition of damage-associated molecular patterns (DAMPs) [11], in larval settlement and metamorphosis [12] and in the recognition of food particles in filter-feeding bivalves [6,13].

Aquatic invertebrates have been a preferred target for lectin identification and purification during the past three decades, as revealed by the fact that many of the best functionally characterized lectins from non-vertebrate metazoans derive from corals, echinoderms and mollusks [14–16]. Among the Lophotrochozoa, one of the two clades of spiralian animals together with Ecdysozoa, most glycobiological and immunological studies have been so far focused on species belonging to the phyla Mollusca or Annelida, amenable for research due to their relatively large body size and the ease of sampling and laboratory handling [17]. Other lophotrochozoan phyla have been nearly completely neglected up to now, leaving a remarkable gap of knowledge concerning the main molecular players involved in carbohydrate recognition.

Among these, the phylum Rotifera, which comprises over 2000 described species with a widespread distribution in freshwater environments, but occasionally found also in brackish and saltwater habitats, represents a particularly intriguing unexplored resource for lectin research. Rotifera are classically subdivided between two classes, namely, Monogononta (the most species-rich class) and Bdelloidea, even though phylogenetic evidence suggests that Seisonidea and Acanthocephala also belong to the very same monophyletic group. Bdelloids display a few peculiar features compared with all other lophotrochozoans, such as a remarkable ability to withstand extreme temperatures [18] and ionizing radiations, which is thought to derive from efficient DNA double-strand break repair [19], and obligatory parthenogenetic reproduction, which results from a long-term asexual evolutionary history [20]. Another interesting feature of rotifers lies in their remarkable genetic divergence from the other major lophotrochozoan phyla. Indeed, monogonont and bdelloid rotifer genomes differ greatly, both in terms of size and architecture, which in bdelloids is significantly impacted by the presence of transposable elements [21], massive horizontal gene transfer [22] and signatures of long-term asexual reproduction [23].

Rotifers often belong to cryptic species complexes, which can only be correctly identified through DNA barcoding, and have in most cases a cosmopolitan distribution [24,25]. These organisms, which usually have a very small size (100–1000 μm), constitute a significant fraction of microzooplankton and their biomass can be particularly relevant in certain environments, such as coastal lagoons or shallow, acidified, metal-contaminated lakes [26–28]. During the 1970s and 1980s, some rotifer species, such as the eurhyaline *Brachionus plicatilis*, were successfully established as live feeds in marine fish aquaculture, thanks to their fast population growth and ease of intensive culture [29–31] (i.e., up to two billion individuals can be obtained in one day per cubic meter of culture [32]). This would undoubtedly represent an interesting opportunity for glycobiology studies, as a sufficient biomass could be readily available for lectin isolation and purification.

The successful adaptation of rotifers to a challenging environment, where they are potentially exposed to a broad range of microorganisms, suggests that these small animals might have developed carbohydrate-binding strategies similar to those described in other aquatic invertebrates in which multiple biomolecules with high biotechnological potential have been previously identified. Moreover, due to their peculiar features and their high tendency to acquire novel genes by horizontal gene transfer, these small metazoans might be considered as a potential source of isolation for a number of novel lectins with unusual and interesting biological properties.

This work preliminarily explores the repertoire of secretory lectins from six rotifer species belonging to the classes Monogononta and Bdelloidea. The publicly available genomes of these species were screened to look for annotated genes encoding proteins

bearing known carbohydrate-binding domains (CRDs). Unlike other lophotrochozoan phyla, in which lectin-like proteins are often encoded by tandemly duplicated paralogous genes displaying high pairwise sequence homology, rotifers do not show evidence of massive gene family expansion events. However, they display a highly diversified arsenal of carbohydrate-binding proteins whose biological properties could be explored and biotechnologically exploited in the near future.

2. Results

The screening of six rotifer genomes allowed the identification of a relatively small number of secretory lectin-like molecules compared with other lophotrochozoans, which are often characterized by massive gene family expansion events that involve carbohydrate-binding proteins, as exemplified by the case of C1qDC proteins in bivalves [33–35]. Based on available data in the literature, only lectin-like proteins displaying a canonical signal peptide for secretion and which display no significant primary sequence conservation among the different lectin families will be here described; the only exception is represented by galectins, which rely on unconventional secretion.

In the class Monogononta, *Brachionus calyciflorus* was the species in which the highest number of lectins was identified (38), followed by the congeneric species *Brachionus plicatilis* (25) and *Proales similis* (14). In the class Bdelloidea, *Rotaria sordida* and *Adineta ricciae* displayed a similar number of lectins (27 and 22, respectively), whereas the third rotifer species, *Didymodactylos carnosus*, had the lowest number of associated lectin sequences in this study (eight) (Table 1; the full list of gene accession IDs is provided in Table S1). Based on these observations, it can be estimated that just a very tiny fraction of all protein-coding genes in rotifers (i.e., 0.02–0.15%) encode secretory lectins characterized by the presence of previously described conserved domains. Nevertheless, despite the lack of evident lectin family expansions, the lectin-like proteins identified in all rotifer species displayed a remarkable molecular diversity, as revealed by their classification within nine different families (Table 1): (i) fibrinogen-related domain-containing proteins (FReDs) (Section 2.1); (ii) C-type lectins (Section 2.2); (iii) C1q-domain containing (C1qDC) proteins (Section 2.3); (iv) galectins (Section 2.4); (v) R-type lectins (Section 2.5); (vi) F-type lectins (Section 2.6); (vii) SUEL-type lectins; (viii) H-type lectins; (ix) jacalin-like lectins (Section 2.7).

Table 1. Number of secretory lectins identified in the six rotifer species analyzed in this study. The full list of gene accession IDs is provided in Table S1.

| | Bdelloidea | | | | Monogononta | |
	Adineta ricciae	*Rotaria sordida*	*Didymodactylos carnosus*	*Proales similis*	*Brachionus calyciflorus*	*Brachionus plicatilis*
FReDs	6	6	1	3	2	6
C-type lectins	3	2	1	3	25	17
C1qDC proteins	4	6	2	1	1	1
Galectins	4	8	2	1	1	1
R-type lectins	0	1	0	0	3	0 [a]
F-type lectins	3	3	1	2	3	0 [b]
SUEL-type lectins	0	0	0	4	3	0 [c]
H-type lectins	2	1	0	0	0	0
Jacalin-like lectins	0	0	1	0	0	0
Apextrins	0	0	0	0	0	0
DUF3011 lectins	0	0	0	0	0	0

[a] Two partial BPBT lectins (see Section 2.5), lacking a signal peptide, likely due to incorrect annotation, were detected. [b] A single FTL with three CRDs, lacking a signal peptide, likely due to incorrect annotation, was detected. [c] Two short single-domain SUEL-type lectins, encoded by two paralogous genes, were likely incorrectly fused in a single gene model.

2.1. FReD-Containing Proteins

Fibrinogen-related domain-containing proteins (FReDs) share structural similarity with the C-terminal domain of vertebrate ficolins, i.e., *N*-acetylglucosamine (GlcNAc)-specific carbohydrate-binding proteins, which play a key role in the lectin pathway of

the complement system [36,37]. The fibrinogen C-terminal domain is associated with a number of metazoan lectins with widespread taxonomic distribution, from cnidarians to vertebrates, which hold remarkable glycan-binding properties and often play an important role in the context of immune recognition, as revealed by several studies carried out in Mollusca [38–40].

A subgroup of FReDs named fibrinogen-related proteins (FREPs), which combine one or two N-terminal immunoglobulin domains with a single C-terminal fibrinogen domain, have been implicated in the resistance of snails to trematode infections [41,42]. Comparative immunogenomics studies have previously revealed that bona fide FREPs [43], as well as GREPs and CREPs (i.e., FReDs associated with galectin and CTL domains, respectively [44]), are restricted to the gastropod subclass Heterobranchia. Nevertheless, other mollusks display a high number of proteins with a simpler architecture, comprising a signal peptide and the fibrinogen-like domain, often paired with a coiled-coil region of variable length, which may allow their oligomerization, in a similar fashion to collagen in vertebrate ficolins [45]. Single-domain FReDs, which retain significant glycan-binding properties in the Lophotrochozoa [46], underwent a significant expansion in bivalves, where they are often found with hundreds of paralogous gene copies [47] encoding inducible proteins with marked bacteria-agglutinating properties [48,49]. Similar expansions have certainly occurred in other lophotrochozoan phyla, such as brachiopods, even though the functional implications of these events are presently unclear [50].

All the rotifer species analyzed in this study had FReD genes in varying numbers, ranging from one (in *D. carnosus*) to six (in *R. sordida*, *B. plicatilis* and *A. ricciae*) (Table 1). The encoded proteins from bdelloid and monogonont rotifers displayed different architectures: while all FReDs shared a single peptide and displayed a fibrinogen-like domain in a C-terminal position, they were characterized by the presence of an N-terminal region of variable length (Figure 1). This region was markedly shorter in bdelloid FReDs, which usually displayed a relatively high (55–50%) primary sequence identity with horseshoe crab tachylectins [39], and much longer in monogonont FReDs, which, on the other hand, had a lower homology (i.e., 25–35%) with tachylectins. In all rotifer FReDs, this region lacked detectable conserved domains and structural homologies but displayed a low level of complexity and the occasional presence of threonine-rich amino acid stretches.
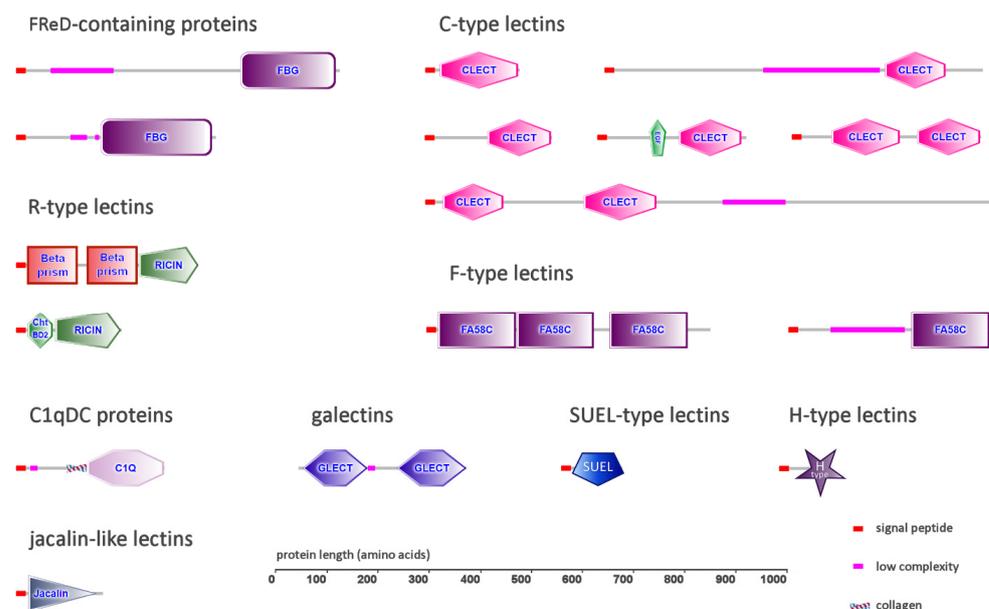
Figure 1. Schematic representation of the main type of secretory lectin-like molecules identified in Rotifera. FBG: Fibrinogen C-terminal domain; GLECT: galectin domain; SUEL: D-galactoside/L-rhamnose-binding SUEL lectin domain; CLECT: C-type lectin domain; FA58C: coagulation factor 5/8 C-terminal domain; EGF: epidermal growth factor domain; Cht BD2: chitin binding domain.

2.2. C-Type Lectins

C-type lectins (CTLs) are one of the largest and most studied families of lectins in lophotrochozoans, with several dozen proteins having been functionally characterized in mollusks and segmented worms [51–53]. Their characterizing CRD, which displays a broad calcium-dependent binding specificity, is often found in large multidomain membrane-bound proteins which may or may not have a lectin function [54,55]. Their remarkable structural diversity has led to the development of a complex classification system, which has been subjected to multiple updates over the years [54,56,57]. Since such a classification still appears to be strongly biased towards vertebrates, it is not fully adequate to describe the variegate domain combinations found in animal CTLs.

Compared with their membrane-bound counterparts, secretory CTLs usually display a simpler structure, which comprises a signal peptide, followed by either one or two tandemly repeated CRDs. In addition, the N-terminal region may also include coiled-coil or collagen repeats with effector functions [58,59]. Besides having a role in MAMP recognition, the CTLs of invertebrate metazoans can regulate different aspects of the innate immune response, including microbial opsonization, the activation of the prophenoloxidase-mediated melanization cascade and possibly also the activation of the complement system, mirroring the role of the mannan-binding lectin in the lectin pathway of the vertebrate complement system [52,60,61].

As far as the Lophotrochozoa are concerned, multiple studies have previously revealed that CTLs belong to highly expanded gene families in Mollusca [33,62], Annelida and Brachiopoda [50]. The investigations carried out here in Rotifera revealed a highly variable number of CTLs among species. While CTLs represented the largest group of secretory lectins in the genus *Brachionus* (i.e., 25 in *B. calyciflorus* and 17 in *B. plicatilis*), only a few proteins of this type (i.e., one to four) could be identified in the four other species (Table 1). Most of the proteins identified in *Brachionus* spp. had a single CRD (Figure 1), which often followed a relatively long (i.e., ~100 amino acids) N-terminal region with no recognizable conserved domains. In addition, both *Brachionus* species displayed a few CTLs with two consecutive CRDs, whose architecture resembled those of insect immulectins [63]. Another type of domain combination included the presence of an epidermal growth factor (EGF)-like domain, placed immediately before the CRD. EGF domains are often found in association with certain large vertebrate CTLs found in the extracellular matrix or bound to the cell membrane, such as selectins and lecticans. However, the combination of a single EGF domain and the CTL CRD has never been described before in the Lophotrochozoa. The third analyzed monogonont rotifer species, *P. similis*, only displayed three CTL genes: two encoded short, single-domain lectins, whereas the third one had an additional EGF-like domain, as previously described in *Brachionus* spp. (Figure 1).

The three bdelloid species had a smaller number of genes encoding secretory CTLs: three were identified in *A. ricciae*, two in *R. sordida* and a single one in *D. carnosus*. Two CTLs from *R. sordida* and one from *A. ricciae* were short single-domain CTLs. *D. carnosus* and *A. ricciae* shared the presence of an orthologous sequence with two recognizable CRDs located at the N-terminal end, followed by a long region with no detectable primary sequence or structural homologies. The third CTL identified in *A. ricciae* showed an unusually long N-terminal low-complexity region, highly enriched in threonine and serine residues, followed by a C-terminal CRD (Figure 1).

In general, rotifer CTLs only showed a poor primary sequence homology (i.e., 20–30%) with functionally characterized molluscan CTLs, which prevented the ascertainment of clear orthology relationships. It is worth mentioning that a single protein belonging to the CTL family had been previously described and functionally characterized in Rotifera. Nevertheless, the sequences orthologous with this protein, which serves as the mate recognition pheromone in the male individuals of *Brachionus manjavacas* [64], are not reported in the present study due to the presence of a transmembrane domain.

2.3. C1q Domain-Containing Proteins

C1q domain-containing (C1qDC) proteins belong to a widespread family of highly versatile globular proteins with remarkable binding properties [65,66]. Besides their well-characterized involvement in the vertebrate complement system, C1qDC proteins carry out important functions in other biological processes which have only recently started to be unveiled [67]. For example, thanks to the carbohydrate-binding properties demonstrated in several metazoan phyla [68,69], C1qDC proteins should be regarded as PRRs involved in immune recognition. This role has been investigated in detail in Mollusca [70,71], where C1qDC proteins are associated with massive gene family expansions [34,35,72]. In bivalves, such expansions involve C1qDC proteins that either have a very simple architecture (signal peptide + C1q domain) or contain an additional N-terminal coiled-coil region. Moreover, in some gastropod species, such as *Littorina littorea*, the C1q domain is combined with one or two immunoglobulin-like domains, originating a small class of proteins known as QREPs, which are upregulated in response to *Himasthla elongata* infections [73].

Unlike bivalves but similar to other lophotrochozoan phyla, such as annelids and brachiopods [50,73], rotifers only display a very few secretory C1qDC proteins (Table 1). In detail, a single orthologous C1qDC gene could be identified in the three monogonont rotifer species, whereas the three bdelloid rotifers had a variable number of C1qDC genes, ranging from two (in *D. carnosus*) to six (in *R. sordida*), with evidence of a few nearly identical paralogs (further supported by phylogenetic evidence; see below). In all cases, rotifer C1qDC proteins were relatively short (<350 aa) and displayed a single C-terminal C1q domain (Figure 1). All proteins had a short (~30 aa long) collagen-like region placed immediately before the start of the C1q domain, which was characterized by the presence of nine highly conserved glycine residues (Figure 2A). This domain organization denotes the typical structure of C1q-like proteins, which represent the most common type of C1qDC proteins in vertebrates [50]. C1q-like proteins are present (but rare) in the lophotrochozoan species characterized by C1qDC gene family expansions, in which collagen repeats are usually replaced by coiled-coil regions [34].
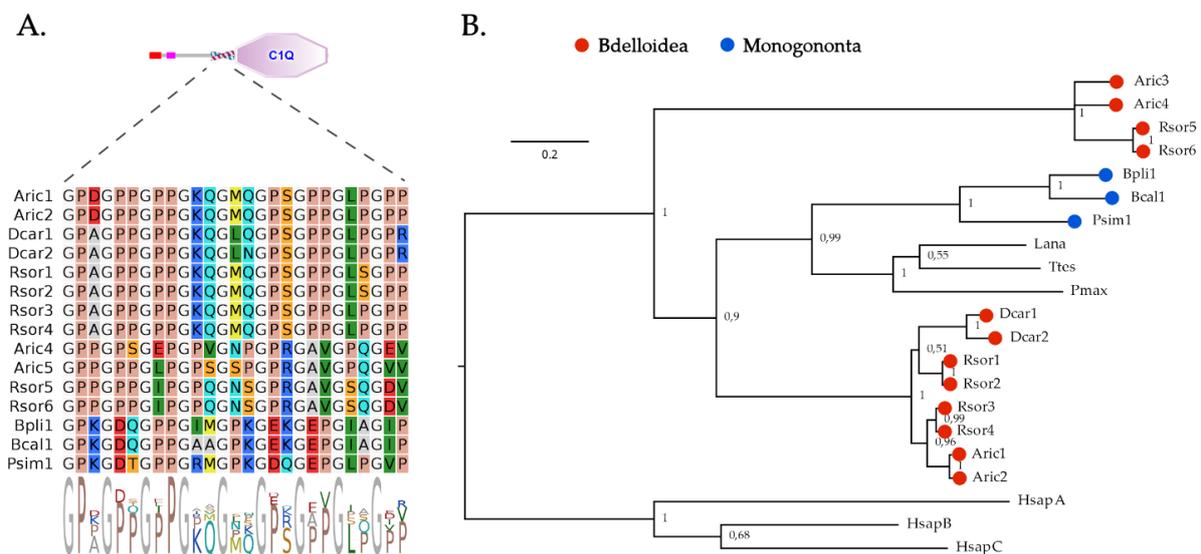
Figure 2. (**A**) Schematic structure of the C1qDC proteins identified in rotifers, with a zoom on the collagen region. (**B**) Bayesian phylogeny of C1qDC proteins from rotifers, obtained with 500,000 generations of an MCMC analysis, run under an LG+I+G model of molecular evolution. The numbers shown close to each node represent posterior probability support values. Aric: *A. ricciae*; Rsor: *R. sordida*; Psim: *P. similis*; Dcar: *D. carnosus*; Bcal: *B. calyciflorus*; Bpli: *B. plicatilis*; Lana: *Lingula anatina*; Pmax: *Pecten maxiumus*; Ttes: *Testudinalia testudinalis*; Hsap: *Homo sapiens*. Human sequences were used as an outgroup for tree-rooting purposes; A, B and C indicate the human C1qA, C1qB and C1qC chains.

From a phylogenetic point of view, the C1qDC proteins of rotifers were subdivided into three distinct groups (Figure 2B): the first included the C1qDC proteins from Monogononta, which were clustered with high support (posterior probability = 0.99) with a few C1q-like proteins previously identified in other lophotrochozoans and hypothesized to play a key role in the proto-complement system [73]. The C1qDC proteins from bdelloid rotifers were clustered in two groups: the first one, which included a few highly similar paralogous genes in each species (two in *D. carnosus* and *A. ricciae*, four in *R. sordida*), comprised proteins with high sequence homology relative to the group of C1qDC proteins from Monogononta and other lophotrochozoans described above. These proteins displayed, as a peculiar feature, an N-terminal low complexity Ser- and Gln-rich region. The second group of C1qDC proteins from bdelloids only comprised sequences from *A. ricciae* and *R. sordida*, which displayed a high divergence with all the other sequences mentioned above and may therefore represent bdelloid innovations.

2.4. Galectins

Galectins are taxonomically widespread and structurally well-conserved β-galactosyl-binding lectins which carry out a multitude of different functions, including cell adhesion, cellular homeostasis and self/non-self and microbial recognition [74]. Based on their structural organization, lophotrochozoan galectins can generally be considered as belonging to the "tandem-repeat" subtype and contain either two or four CRDs [75–78], with rare occurrences of galectins with three CRDs [50]. Although phylogenetic analyses have previously revealed a monophyletic origin for all molluscan galectins [79], it is presently unclear whether this consideration also applies to the galectins from other lophotrochozoan phyla.

This investigation allowed the identification of galectin genes in all the six analyzed rotifer genomes, even though their number significantly varied among species. While all Monogononta only had a single galectin, bdelloid genomes encoded multiple galectin genes, ranging from two (*D. carnosus*) to eight (*R. sordida*) (Table 1). All rotifer galectins displayed two tandemly repeated CRDs, separated by a connecting region of variable length (Figure 1). No galectins with four CRDs could be identified, confirming the previous observation that, within the Lophotrochozoa, this subtype is restricted to brachiopods, phoronids and annelids [50]. Primary sequence homology with other members of the galectin family from non-rotifer lophotrochozoans was generally in the range of 30–35%. Consistently with previous observations in other metazoans, the encoded proteins lacked a canonical signal peptide and might therefore use an alternative secretion route [80].

2.5. Ricin β-Trefoil Lectins

The R-type lectin (RTL) domain, originally described in the plant toxin ricin, is found in a number of metazoan multidomain proteins with different functions, including hydrolases, glycosyltransferases and membrane-bound receptors [81]. Nevertheless, smaller proteins with no additional domains can serve as lectins in the extracellular environment, playing a role in PAMP recognition. A number of secretory R-type lectins with different glycan-binding properties, containing either one or two consecutive CRDs, have been previously isolated in annelids [82–84] and mollusks [85]. A second family of lectins, named mytilectins, which share the same β-trefoil three-dimensional structure but do not conform with the canonical R-type lectin primary sequence signature, show a discontinuous distribution among the Lophotrochozoa and have only been described so far in a few bivalve mollusks and brachiopods [50,86,87].

While rotifer genomes encoded several proteins with R-type lectin domains, in most cases these were associated with other domains known to exert catalytic activities (e.g., glycosylases, hydrolases, etc.) or with transmembrane domains. Strong evidence in support of the existence of secretory RTLs could be collected only for two out of the six rotifer species analyzed in this study, i.e., *R. sordida*, among bdelloids, and *B. calyciflorus*, among Monogononta (Table 1).

In detail, the three secretory RTLs identified in *B. calyciflorus* displayed an unusual domain architecture, never before reported in other metazoans. Indeed, these proteins showed the presence of two consecutive VOMI (vitelline membrane outer layer protein I) domains, followed by a C-terminal ricin-like CRD (Figure 1). Although the VOMI domain is typically found in proteins found in the outer layer of the egg vitelline membrane [88], it shares a β-prism fold that has been previously identified in other carbohydrate-binding proteins, including jacalins, a class of plant-specific lectins [89–91], as well as in the *B. thuringiensis* delta endotoxin [92]. Due to the simultaneous presence of these two structurally different CRDs, which clearly presents an interesting path for exploration in glycobiological studies, we defined these unusual proteins as BPBT (β-prism, β-trefoil) lectins. Two BTBP lectins, orthologous to those found in *B. calyciflorus* but lacking a signal peptide (possibly due to an incorrect annotation), were also found in the congeneric species *B. plicatilis*, but not in the other species, suggesting that this domain combination may be exclusively present in *Brachionus* spp.

On the other hand, the single secretory RTL found in *R. sordida* was unrelated to BPBT lectins, since this protein was relatively short (i.e., 200 amino acids) and included a chitin-binding domain in an N-terminal position [93] (Figure 1). This domain is shared by several chitinases and other smaller chitin-binding proteins, which include some with demonstrated effector activity in the context of invertebrate innate immunity, such as horseshoe crab tachycytin [94] and mussel mytichitin [95], and others with presumed lectin-like functions [96].

No sequence orthologous to brachiopod and molluscan mytilectins could be found in rotifers, confirming the discontinuous taxonomic distribution of these β-trefoil lectins in the Lophotrochozoa.

2.6. F-Type Lectins

F-type lectins are characterized by the presence of a β-barrel jellyroll fold which allows fucose recognition [97] and which is also found in the C-terminal domain of coagulation factors 5/8. Despite being associated with relatively short secretory proteins with a lectin function, the typical CRD of FTLs is often found in large multidomain proteins with different catalytic activities [98]. The frequent combination of this domain with several other non-lectin domains mirrors the previously mentioned functional plasticity of the CRDs of CTLs and RTLs. Previous studies have reported that FTLs underwent expansion in some gastropod species [77], and some functional evidence collected in bivalves has linked these proteins to bacterial recognition [99], in addition to the well-established role of the FTL domain-containing proteins bindins in gamete recognition [100]. This observation is consistent with the detection of the FTL domain in a relatively high number of rotifer proteins, only a few of which were characterized by the presence of a signal peptide or displayed a domain organization consistent with a lectin function (Table 1).

Two different types of secretory F-type lectin sequences were detected in rotifers. The first type, present as a single-copy gene in the three monogonont species but missing in the three bdelloids, was a protein displaying a low-complexity threonine- and serine-rich N-terminal region, followed by a single CRD lacking any significant primary sequence homology with known FTLs but showing high predicted structural similarity with human coagulation factors [101] and discoidins [102] (Figure 1). The second type, shared by all rotifer species (even though *B. plicatilis* only displayed a protein lacking the signal peptide, likely due to incorrect annotation), displayed three consecutive FTL CRDs (Figure 1). While the first and the second ones were well recognizable, the third one did not conform with the canonical F-type lectin signature. These triple-CRD FTLs displayed a relatively high (i.e., 40%) sequence identity with several proteins encoded by the genomes of other lophotrochozoans, including mollusks and annelids, suggesting a high degree of evolutionary conservation.

2.7. Other Types of Lectins

In sea urchins, a group of lectins, characterized by the presence of a D-galactoside/L-rhamnose-binding SUEL (acronym for sea urchin egg lectin) domain, carry out egg-protecting functions [103,104]. While this type of lectins is also found in lophotrochozoan genomes, they have been better functionally characterized in deuterostome invertebrates [105–107]. To date, their role in lophotrochozoans remains elusive, even though a bivalve SUEL-type lectin was shown to promote the agglutination of Gram-negative bacteria through LPS binding [108]. The SUEL domain was present in some large multidomain membrane-associated proteins of bdelloids but not in secretory proteins. On the other hand, the three monogonont rotifer genomes encoded a few short SUEL-type lectins (Table 1), which lacked accessory conserved domains (Figure 1) and did no bear any detectable primary sequence homology with other metazoan sequences with known functions. With the exception of two sequences detected in *P. similis*, these proteins displayed high pairwise primary sequence homology and clearly belonged to a monophyletic family.

H-type lectins (HTLs) represent a poorly functionally characterized family of *N*-acetylgalactosamine-binding lectins, which are believed to carry out a defensive role against bacterial infections in fertilized snail eggs [109,110]. Although very little information is available about the involvement of HTLs in lophotrochozoan immunity, comparative genomics analyses indicate that they do not belong to expanded gene families, neither in Mollusca [111,112] nor in Brachiopoda [50]. Nevertheless, transcriptome scans carried out in gastropod mollusks revealed the presence of a novel domain combination between immunoglobulin-like domains and HTL domains in the so-called HREPs [73]. The analysis of rotifer genomes revealed the presence of secretory H-type lectins in just two out of the three bdelloid species (i.e., *A ricciae* and *R. sordida*). On the other hand, no HTL was identified in Monogononta (Table 1). These proteins had a similar simple architecture, with a single CRD, placed immediately after a well-recognizable signal peptide (Figure 1). Rotifer HTLs were encoded by open reading frames with a relatively small size (i.e., 120 codons) and displayed poor sequence homology with other known sequences (i.e., less than 40% primary sequence homology vs. *L. anatina*). This may suggest that the apparent lack of secretory HTLs in four out of six target species derives from missing gene models that could not be included in the annotation tracks of the respective genomes due to poor supporting evidence.

Section 2.5 reports the presence of BPBT lectins, which bear a jacalin-like β-prism structural domain in combination with the RTL CRD, in *Brachionus* spp. The screening for additional proteins bearing a canonical jacalin domain led to the identification of a single protein in *D. carnosus* with no orthologs in other rotifer species. This lectin, which displayed a well recognizable signal peptide for secretion, lacked significant primary sequence homology with other previously characterized proteins, but displayed a highly significant structural match with a number of jacalin-like lectins from plants and with a few metazoan proteins. These included, as the only lophotrochozoan representative, the PPL3 lectin from the bivalve mollusk *Pteria penguin*, which is involved in shell mineralization [113]. Other relevant metazoan proteins which display the same structural fold are the human pancreatic secretory protein ZG16b, important for the condensation of pancreatic enzymes [114], the WGA16 protein from boar sperm [115] and the zebrafish pore-forming protein Dln1 [116].

No gene encoding proteins homologous to the egg-protecting lectins from *Aplysia dactylomela*, characterized by the presence of the orifera lack lectins homologous to her metazoan sequences with known function.ked significant similarity withology DUF3011 domain [117], could be detected in Rotifera. Likewise, no apextrin-like proteins were detected in any of the studied species. Proteins carrying an apextrin C-terminal domain (ApeC) have been previously shown to mediate bacterial recognition in amphioxus [118]. While they are also present in bivalve mollusks and brachiopods [50,62], this study rules out their possible involvement in immune recognition in rotifers.

3. Discussion

Genome- or transcriptome-wide screening approaches have previously been successfully used to identify lectin-like proteins in several different eukaryote species, both in the plant and animal kingdoms [77,96,112,119–122], and can be effectively used as a preliminary step to investigate the repertoire of lectin-like molecules present in non-model species. This approach might be particularly intriguing for understudied animal phyla, which, despite the lack of previous glycobiological investigations, might be endowed with a particularly rich repertoire of lectin-like molecules as the result of their adaptation to a challenging environment. Rotifers, like other aquatic organisms, are potentially highly exposed to ubiquitous waterborne microorganisms, which may in some cases be pathogenic. Considering the well-described complex innate immune systems of other lophotrochozoans, such as mollusks and annelids, rotifers appear to be good candidates for the bioinformatics-assisted discovery of carbohydrate-binding proteins involved in MAMP recognition. No information is presently available concerning the glycans expressed by rotifer tissues and only a single membrane-bound C-type lectin has been described in *B. manjavacas* [66] prior to this work. Hence, this represents the first investigation of this type carried out in this relatively large and widespread but poorly studied phylum of small aquatic animals.

Although this approach obviously suffers from some limitations, which will be described in detail below, it has allowed: (i) the identification of the presence of secretory lectin-like molecules belonging to at least nine different families, characterized by distinct structural folds (Figure 3), in rotifers; (ii) the highlighting of significant differences in terms of distribution and domain organization between the two major classes of rotifers, as well as among species; and (iii) the ruling out of the possibility that known lectin-encoding gene families underwent significant expansion in Rotifera, marking a clear difference with other lophotrochozoan phyla.

Even though this in silico screening approach allowed the identification of several proteins that are likely to be secreted to the extracellular environment and have significant carbohydrate-binding properties, the list of the putative rotifer lectin-like proteins here provided (detailed in Table S1) should be considered as strictly preliminary. The glycan-binding properties of the identified proteins, as well as their possible involvement in MAMP recognition, should be validated through functional data collected with classical biochemical, glycobiological and immunological approaches.

Some limitations of this genome-wide bioinformatics screening approach reside in the fact that the correct identification of lectin-like proteins depends on the accuracy of gene annotations. While all the genomes analyzed in this work had a high quality, both in terms of assembly metrics and in terms of BUSCO completeness [123], a few chimeric gene models, as well as models of ORFs which were clearly subjected to 5' or 3' truncation compared with other full-length orthologs and paralogs, were occasionally observed. For the sake of consistency, these gene models were disregarded, even though the presence of incomplete gene models in a given species was reported, whenever relevant, in Table 1. In addition, some lectin-like proteins identified in this work were rather short, with an ORF barely exceeding 100 codons, and lacked at the same time significant primary sequence homology with other known sequences deposited in public databases. These factors might have negatively impacted the annotation of other orthologous and paralogous genes, which may therefore be apparently missing in some of the target genomes, as discussed in Section 2.7. Hence, the completeness of our report may suffer from these uncertainties, and the number of lectins reported in Table 1 should be considered as inherently subjected to slight underestimates.
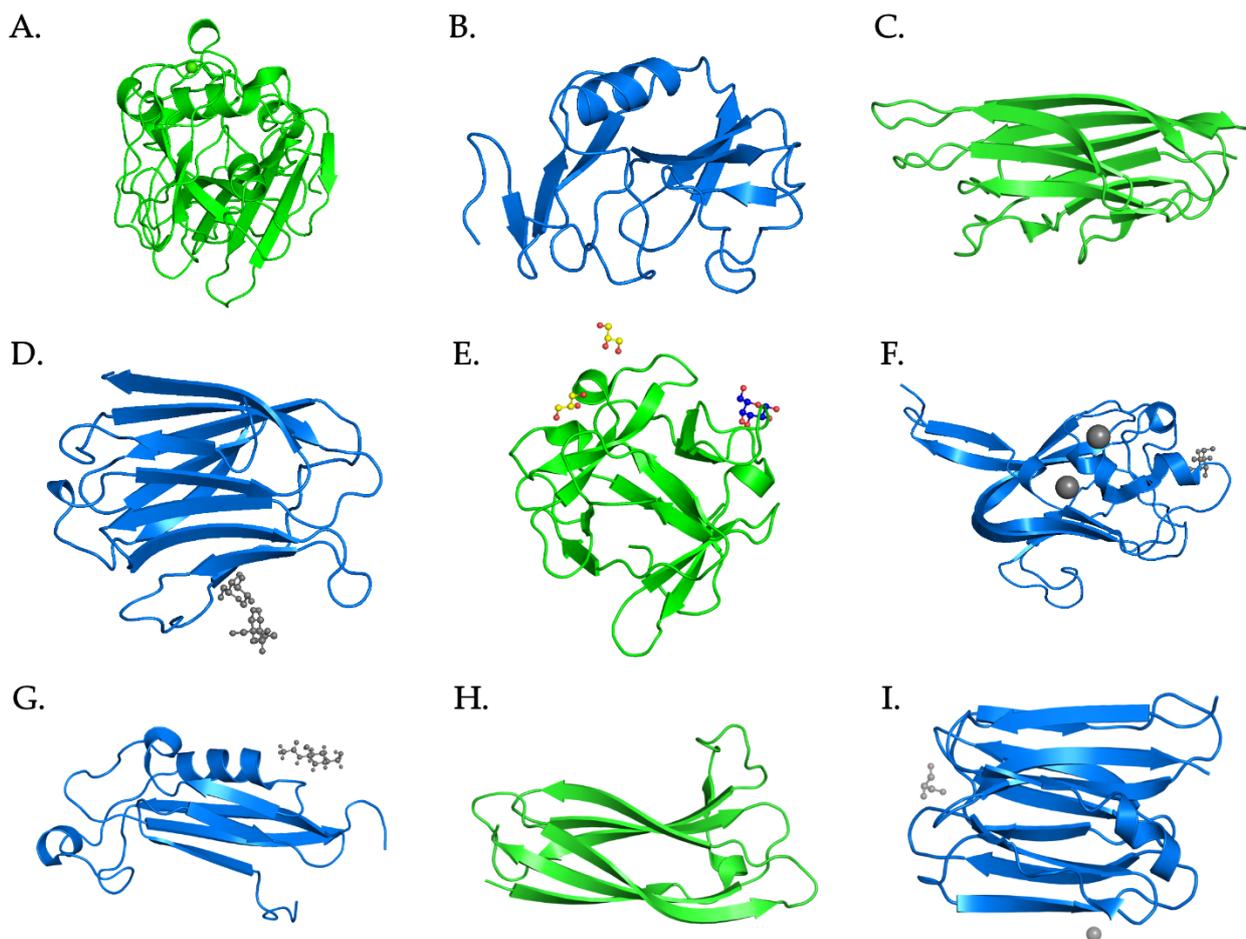
Figure 3. Three-dimensional folding of the carbohydrate-recognition domains of representative members of the nine lectin families identified in Rotifera. (**A**) C-type lectin domain, pdb entry: 1b6e; (**B**) fibrinogen C-terminal domain, pdb entry: 1fib; (**C**) C1q domain, pdb entry: 4ous; (**D**) galectin domain, pdb entry: 1a3k; (**E**) R-type lectin β-trefoil domain, pdb entry: 4iyb; (**F**) F-type lectin domain, pdb entry: 1k12; (**G**) SUEL-type lectin domain, pdb entry: 2jx9; (**H**) H-type lectin domain, pdb entry: 2ces; (**I**) jacalin β-prism domain, pdb entry: 3apa. The figures are reproduced courtesy of PDBe (https://www.ebi.ac.uk/pdbe/, accessed on 20 January 2022).

Another possible limitation of this work was the impossibility of proceeding with a reliable in silico screening of candidate lectin molecules characterized by the presence of domains which are not primarily or exclusively linked with a carbohydrate-binding function. This was the case, for example, for I-type lectins (also known as siglecs), which share an immunoglobulin-like fold with a very high number of other proteins involved in a very broad range of functions [124] and which have been previously identified in some lophotrochozoans [125]. Similarly, some chitin-binding lectins, mostly from plants [126], are characterized by the presence of a chitinase-like domain that includes a few key mutations that ablate its catalytic function. However, since these two lectin families are either membrane-bound (siglecs) or uncommon in metazoans (chitinase-like lectins), their exclusion from the set of domains included in in silico searches was unlikely to have an impact on the identification of secretory lectins in rotifers.

Finally, it needs to be stressed that rotifer genomes are extremely gene-rich and encode several thousand proteins which lack any significant primary sequence homology with other sequences deposited in public databases and which have no detectable conserved domain. We cannot exclude that some as yet uncharacterized protein families may carry out important carbohydrate-binding functions in these animals. Nevertheless, the combination of classical biochemical and glycobiological approaches with bioinformatics approaches

should enable the identification of the full-length sequence of the lectins isolated from rotifers starting from small peptide fragments, as previously carried out on several occasions with other aquatic invertebrates [68,85,127].

In summary, this work allowed the confirmation of the potential interest of rotifers as future targets for glycobiological studies focused on the identification of novel lectins. These, based on the significant diversity of the associated structural folding, might be endowed with different carbohydrate-binding properties, which may support the development of new biotechnological tools, such as lectin-based biosensors with potential applications in cancer research. Besides the interest that such molecules might have in terms of biotechnological applications, another aspect that remains to be clarified is whether these rotifer secretory lectins carry out biological functions similar to those previously described in other lophotrochozoan phyla.

4. Materials and Methods

4.1. Identification of Lectin-Like Molecules

Six rotifer species with a publicly available fully sequenced genome and an associated gene annotation track were selected (Table 2). *Didymodactylos carnosus* Milne 1916 [21], *Rotaria sordida* Western, 1893 [21] and *Adineta ricciae* Segers & Shiel, 2005 [21] were selected for the class Bdelloidea; *Brachionus plicatilis* Müller, 1786 [128], *Brachionus calyciflorus* Pallas, 1776 [129] and *Proales similis* de Beauchamp, 1907 [130] were selected for the class Monogononta.

Table 2. List of the six rotifer species analyzed in this study, with genome size and number of annotated protein-coding genes.

Species Name	Class	Genome Size (Mb)	Protein-Coding Genes
Adineta ricciae	Bdelloidea	173	49,015
Rotaria sordida	Bdelloidea	361	61,901
Didymodactylos carnosus	Bdelloidea	356	46,863
Proales similis	Monogononta	33	10,785
Brachionus calyciflorus	Monogononta	30	24,328
Brachionus plicatilis	Monogononta	107	52,502

The proteome of each of the six target species was screened and a search was made for secretory proteins, i.e., those including either a highly supported canonical signal peptide, detected with SignalP v.5.0 [131], or, in the case of galectins, which are known to use a non-canonical secretion signal, with SecretomeP v.2.0 [132]. At the same time, candidate proteins needed to lack transmembrane regions, which were detected with TMHMM v.2.0 [133]. Signal peptide and transmembrane region predictions were further checked with Phobius v.1.01 [134]. Putative lectins were identified based on the presence of the following PFAM conserved domains, detected with HMMER [135] based on default threshold e-values: fibrinogen beta and gamma chains, C-terminal globular domain (PF00147), C-type lectin domain (PF00059), C1q domain (PF00386), galactose-binding lectin domain (PF02140), ricin-type beta-trefoil domain (PF00652 and PF14200), F-type lectin/discoidin domain (PF00754), galactoside-binding lectin domain (PF00337), H-type lectin domain (PF09458), jacalin-like lectin domain (PF01419), DUF3011 (PF11218) and the C-terminal domain of apextrin (PF16977).

The presence of other conserved domains was checked with InterProScan v.5 [136] and the possible presence of conserved structural folds in regions lacking conserved domains was investigated with HHpred [137]. To avoid the inclusion of proteins carrying lectin-like domains but likely involved in non-lectin functions, sequences displaying conserved domains with known catalytic action (e.g., glucanases, hydrolases, kinases, peptidases, etc.) were disregarded. Protein sequences deriving from gene models annotated as "partial", as well as those that displayed obviously truncated characterizing domains and which might therefore derive either from pseudogenes or from mis-annotations, were disregarded.

4.2. Bayesian Phylogenetic Inference

All rotifer C1qDC proteins were included in a multiple sequence alignment (MSA), prepared with MUSCLE [138], together with three selected lophotrochozoan C1q-like sequences (i.e., XP_013399541.1 from *Lingula anatina*, QBA18422.1 from *Testudinalia testudinalis* and XP_033760315.1 from *Pecten maximus*). The human C1qA (NP_057075.1), C1qB (NP_000482.3) and C1qc (NP_758957.2) chains were also added to the alignment for tree-rooting purposes. The MSA was refined with Gblocks v.0.91b [139] to remove unalignable, poorly informative regions. Bayesian phylogenetic analysis was carried out with MrBayes v.3.2.7a [140], running two parallel MCMC analyses for 500,000 generations each, sampling one tree every 1,000 generations. The analysis was run under an LG+G+I model of molecular evolution, which was estimated to be the best fitting for this dataset, with ModelTest-NG [141]. Run convergence was checked with Tracer v.1.7.1 [142] by determining that all estimated parameters reached an effective sample size value higher than 100.

References

1. Stuart, L.M.; Paquette, N.; Boyer, L. Effector-Triggered versus Pattern-Triggered Immunity: How Animals Sense Pathogens. *Nat. Rev. Immunol.* **2013**, *13*, 199–206. [CrossRef] [PubMed]
2. Wang, J.-L.; Zhang, Q.; Tang, L.; Chen, L.; Liu, X.-S.; Wang, Y.-F. Involvement of a Pattern Recognition Receptor C-Type Lectin 7 in Enhancing Cellular Encapsulation and Melanization Due to Its Carboxyl-Terminal CRD Domain in the Cotton Bollworm, *Helicoverpa Armigera*. *Dev. Comp. Immunol.* **2014**, *44*, 21–29. [CrossRef] [PubMed]
3. Wang, X.-W.; Xu, J.-D.; Zhao, X.-F.; Vasta, G.R.; Wang, J.-X. A Shrimp C-Type Lectin Inhibits Proliferation of the *Hemolymph Microbiota* by Maintaining the Expression of Antimicrobial Peptides. *J. Biol. Chem.* **2014**, *289*, 11779–11790. [CrossRef] [PubMed]
4. Brouwer, N.; Dolman, K.M.; van Zwieten, R.; Nieuwenhuys, E.; Hart, M.; Aarden, L.A.; Roos, D.; Kuijpers, T.W. Mannan-Binding Lectin (MBL)–Mediated Opsonization Is Enhanced by the Alternative Pathway Amplification Loop. *Mol. Immunol.* **2006**, *43*, 2051–2060. [CrossRef] [PubMed]
5. Kerrigan, A.M.; Brown, G.D. C-Type Lectins and Phagocytosis. *Immunobiology* **2009**, *214*, 562. [CrossRef]
6. Pales Espinosa, E.; Allam, B. Reverse Genetics Demonstrate the Role of Mucosal C-Type Lectins in Food Particle Selection in the Oyster *Crassostrea Virginica*. *J. Exp. Biol.* **2018**, *221*, jeb174094. [CrossRef]
7. Nyholm, S.V.; Graf, J. Knowing Your Friends: Invertebrate Innate Immunity Fosters Beneficial Bacterial Symbioses. *Nature Rev. Microbiol.* **2012**, *10*, 815–827. [CrossRef]
8. Grossart, H.-P.; Riemann, L.; Tang, K. Molecular and Functional Ecology of Aquatic Microbial Symbionts. *Front. Microbiol.* **2013**, *4*, 59. [CrossRef]
9. Li, H.-H.; Cai, Y.; Li, J.-C.; Su, M.P.; Liu, W.-L.; Cheng, L.; Chou, S.-J.; Yu, G.-Y.; Wang, H.-D.; Chen, C.-H. C-Type Lectins Link Immunological and Reproductive Processes in *Aedes Aegypti*. *iScience* **2020**, *23*, 101486. [CrossRef]
10. Springer, S.A.; Moy, G.W.; Friend, D.S.; Swanson, W.J.; Vacquier, V.D. Oyster Sperm Bindin Is a Combinatorial *Fucose Lectin* with Remarkable Intra-Species Diversity. *Int. J. Dev. Biol.* **2008**, *52*, 759–768. [CrossRef]
11. Nauta, A.J.; Castellano, G.; Xu, W.; Woltman, A.M.; Borrias, M.C.; Daha, M.R.; van Kooten, C.; Roos, A. Opsonization with C1q and Mannose-Binding Lectin Targets Apoptotic Cells to Dendritic Cells. *J. Immunol.* **2004**, *173*, 3044–3050. [CrossRef] [PubMed]
12. Maki, J.S.; Mitchell, R. Involvement of Lectins in the Settlement and Metamorphosis of Marine *Invertebrate Larvae*. *Bull. Mar. Sci.* **1985**, *37*, 675–683.
13. Emmanuelle, P.E.; Mickael, P.; Evan, W.; Shumway, S.E.; Bassem, A. Lectins Associated with the Feeding Organs of the Oyster *Crassostrea Virginica* Can Mediate Particle Selection. *Biol. Bull.* **2009**, *217*, 130–141.

14. Kvennefors, E.C.E.; Leggat, W.; Hoegh-Guldberg, O.; Degnan, B.M.; Barnes, A.C. An Ancient and Variable Mannose-Binding Lectin from the Coral *Acropora Millepora* Binds Both Pathogens and Symbionts. *Dev. Comp. Immunol.* **2008**, *32*, 1582–1592. [CrossRef] [PubMed]

15. Giga, Y.; Ikai, A.; Takahashi, K. The Complete Amino Acid Sequence of Echinoidin, a Lectin from the Coelomic Fluid of the Sea Urchin *Anthocidaris Crassispina*. Homologies with Mammalian and Insect Lectins. *J. Biol. Chem.* **1987**, *262*, 6197–6203. [CrossRef]

16. Yamaura, K.; Takahashi, K.G.; Suzuki, T. Identification and Tissue Expression Analysis of C-Type Lectin and Galectin in the Pacific Oyster, *Crassostrea Gigas*. *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* **2008**, *149*, 168–175. [CrossRef]

17. Gerdol, M.; Gomez-Chiarri, M.; Castillo, M.G.; Figueras, A.; Fiorito, G.; Moreira, R.; Novoa, B.; Pallavicini, A.; Ponte, G.; Roumbedakis, K.; et al. Immunity in Molluscs: Recognition and Effector Mechanisms, with a Focus on Bivalvia. In *Advances in Comparative Immunology*; Cooper, E.L., Ed.; Springer International Publishing: Cham, Switzerland, 2018; pp. 225–234. ISBN 978-3-319-76768-0.

18. Shain, D.H.; Halldórsdóttir, K.; Pálsson, F.; Aðalgeirsdóttir, G.; Gunnarsson, A.; Jónsson, Þ.; Lang, S.A.; Pálsson, H.S.; Steinþórssson, S.; Arnason, E. Colonization of Maritime Glacier Ice by *Bdelloid Rotifera*. *Mol. Phylogenet. Evol.* **2016**, *98*, 280–287. [CrossRef] [PubMed]

19. Hespeels, B.; Knapen, M.; Hanot-Mambres, D.; Heuskin, A.-C.; Pineux, F.; Lucas, S.; Koszul, R.; van Doninck, K. Gateway to Genetic Exchange? DNA Double-Strand Breaks in the Bdelloid Rotifer *Adineta Vaga* Submitted to Desiccation. *J. Evol. Biol.* **2014**, *27*, 1334–1345. [CrossRef]

20. Welch, J.L.M.; Welch, D.B.M.; Meselson, M. Cytogenetic Evidence for Asexual Evolution of *Bdelloid Rotifers*. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 1618–1621. [CrossRef]

21. Nowell, R.W.; Wilson, C.G.; Almeida, P.; Schiffer, P.H.; Fontaneto, D.; Becks, L.; Rodriguez, F.; Arkhipova, I.R.; Barraclough, T.G. Evolutionary Dynamics of Transposable Elements in *Bdelloid Rotifers*. *eLife* **2021**, *10*, e63194. [CrossRef]

22. Eyres, I.; Boschetti, C.; Crisp, A.; Smith, T.P.; Fontaneto, D.; Tunnacliffe, A.; Barraclough, T.G. Horizontal Gene Transfer in Bdelloid Rotifers Is Ancient, Ongoing and More Frequent in Species from Desiccating Habitats. *BMC Biol.* **2015**, *13*, 90. [CrossRef] [PubMed]

23. Gladyshev, E.A.; Arkhipova, I.R. Genome Structure of Bdelloid Rotifers: Shaped by Asexuality or Desiccation? *J. Hered.* **2010**, *101* (Suppl. S1), S85–S93. [CrossRef] [PubMed]

24. Gabaldón, C.; Fontaneto, D.; Carmona, M.J.; Montero-Pau, J.; Serra, M. Ecological Differentiation in Cryptic Rotifer Species: What We Can Learn from the *Brachionus Plicatilis* Complex. *Hydrobiologia* **2017**, *796*, 7–18. [CrossRef]

25. Malekzadeh-Viayeh, R.; Pak-Tarmani, R.; Rostamkhani, N.; Fontaneto, D. Diversity of the Rotifer Brachionus Plicatilis Species Complex (*Rotifera: Monogononta*) in Iran through Integrative Taxonomy. *Zool. J. Linn. Soc.* **2014**, *170*, 233–244. [CrossRef]

26. Lam-Höai, T.; Rougier, C.; Lasserre, G. Tintinnids and Rotifers in a Northern Mediterranean Coastal Lagoon. Structural Diversity and Function through Biomass Estimations. *Mar. Ecol. Prog. Ser.* **1997**, *152*, 13–25. [CrossRef]

27. Assefa, E.; Mengidtou, S. Seasonal Variation of Biomass and Secondary Production of Thermocyclops (*Cyclopoida*) and Brachionus (*Rotifera*) Spp. in a Shallow Tropical Lake Kuriftu, Ethiopia. *SINET Ethiop. J. Sci.* **2011**, *34*, 73–88. [CrossRef]

28. Yan, N.D.; Geiling, W. Elevated Planktonic Rotifer Biomass in Acidified Metal-Contaminated Lakes near Sudbury, Ontario. *Hydrobiologia* **1985**, *120*, 199–205. [CrossRef]

29. Lubzens, E.; Tandler, A.; Minkoff, G. Rotifers as Food in Aquaculture. *Hydrobiologia* **1989**, *186*, 387–400. [CrossRef]

30. Dhont, J.; Dierckens, K.; Støttrup, J.; Van Stappen, G.; Wille, M.; Sorgeloos, P. 5–Rotifers, Artemia and Copepods as Live Feeds for Fish Larvae in Aquaculture. In *Advances in Aquaculture Hatchery Technology*; Allan, G., Burnell, G., Eds.; Woodhead Publishing Series in Food Science, Technology and Nutrition; Woodhead Publishing: Sawston, UK, 2013; pp. 157–202. ISBN 978-0-85709-119-2.

31. Hagiwara, A.; Marcial, H.S. The Use of Non-Brachionus Plicatilis Species Complex Rotifer in Larviculture. *Hydrobiologia* **2019**, *844*, 163–172. [CrossRef]

32. Hagiwara, A.; Kim, H.-J.; Marcial, H. Mass Culture and Preservation of *Brachionus Plicatilis* Sp. Complex. In *Rotifers: Aquaculture, Ecology, Gerontology, and Ecotoxicology*; Hagiwara, A., Yoshinaga, T., Eds.; Fisheries Science Series; Springer: Singapore, 2017; pp. 35–45. ISBN 978-981-10-5635-2.

33. Zhang, L.; Li, L.; Guo, X.; Litman, G.W.; Dishaw, L.J.; Zhang, G. Massive Expansion and Functional Divergence of Innate Immune Genes in a Protostome. *Sci. Rep.* **2015**, *5*, 8693. [CrossRef]

34. Gerdol, M.; Venier, P.; Pallavicini, A. The Genome of the Pacific Oyster Crassostrea Gigas Brings New Insights on the Massive Expansion of the C1q Gene Family in Bivalvia. *Dev. Comp. Immunol.* **2015**, *49*, 59–71. [CrossRef] [PubMed]

35. Gerdol, M.; Greco, S.; Pallavicini, A. Extensive Tandem Duplication Events Drive the Expansion of the C1q-Domain-Containing Gene Family in Bivalves. *Mar. Drugs* **2019**, *17*, 583. [CrossRef] [PubMed]

36. Matsushita, M. Ficolins: Complement-Activating Lectins Involved in Innate Immunity. *J. Innate Immun.* **2010**, *2*, 24–32. [CrossRef] [PubMed]

37. Matsushita, M.; Fujita, T. The Role of Ficolins in Innate Immunity. *Immunobiology* **2002**, *205*, 490–497. [CrossRef]

38. Imamichi, Y.; Yokoyama, Y. Purification, Characterization and CDNA Cloning of a Novel Lectin from the Jellyfish *Nemopilema Nomurai*. *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* **2010**, *156*, 12–18. [CrossRef]

39. Gokudan, S.; Muta, T.; Tsuda, R.; Koori, K.; Kawahara, T.; Seki, N.; Mizunoe, Y.; Wai, S.N.; Iwanaga, S.; Kawabata, S. Horseshoe Crab Acetyl Group-Recognizing Lectins Involved in Innate Immunity Are Structurally Related to Fibrinogen. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 10086–10091. [CrossRef]

40. Gout, E.; Garlatti, V.; Smith, D.F.; Lacroix, M.; Dumestre-Pérard, C.; Lunardi, T.; Martin, L.; Cesbron, J.-Y.; Arlaud, G.J.; Gaboriaud, C.; et al. Carbohydrate Recognition Properties of Human Ficolins: Glycan Array Screening Reveals the Sialic Acid Binding Specificity of M-Ficolin. *J. Biol. Chem.* **2010**, *285*, 6612–6622. [CrossRef]

41. Adema, C.M. Fibrinogen-Related Proteins (FREPs) in Mollusks. *Results Probl. Cell Differ.* **2015**, *57*, 111–129. [CrossRef]

42. Gordy, M.A.; Pila, E.A.; Hanington, P.C. The Role of Fibrinogen-Related Proteins in the Gastropod Immune Response. *Fish Shellfish Immunol.* **2015**, *46*, 39–49. [CrossRef]

43. Gorbushin, A.M.; Panchin, Y.V.; Iakovleva, N.V. In Search of the Origin of FREPs: Characterization of *Aplysia Californica* Fibrinogen-Related Proteins. *Dev. Comp. Immunol.* **2010**, *34*, 465–473. [CrossRef]

44. Dheilly, N.M.; Duval, D.; Mouahid, G.; Emans, R.; Allienne, J.-F.; Galinier, R.; Genthon, C.; Dubois, E.; Du Pasquier, L.; Adema, C.M.; et al. A Family of Variable Immunoglobulin and Lectin Domain Containing Molecules in the Snail *Biomphalaria Glabrata*. *Dev. Comp. Immunol.* **2015**, *48*, 234–243. [CrossRef] [PubMed]

45. Skazina, M.A.; Gorbushin, A.M. Characterization of the Gene Encoding a Fibrinogen-Related Protein Expressed in *Crassostrea Gigas* Hemocytes. *Fish Shellfish Immunol.* **2016**, *54*, 586–588. [CrossRef] [PubMed]

46. Kurachi, S.; Song, Z.; Takagaki, M.; Yang, Q.; Winter, H.C.; Kurachi, K.; Goldstein, I.J. Sialic-Acid-Binding Lectin from the Slug Limax Flavus–Cloning, Expression of the Polypeptide, and Tissue Localization. *Eur. J. Biochem.* **1998**, *254*, 217–222. [CrossRef] [PubMed]

47. Huang, B.; Zhang, L.; Li, L.; Tang, X.; Zhang, G. Highly Diverse Fibrinogen-Related Proteins in the Pacific Oyster *Crassostrea Gigas*. *Fish Shellfish Immunol.* **2015**, *43*, 485–490. [CrossRef]

48. Xiang, Z.; Qu, F.; Wang, F.; Li, J.; Zhang, Y.; Yu, Z. Characteristic and Functional Analysis of a Ficolin-like Protein from the Oyster *Crassostrea Hongkongensis*. *Fish Shellfish Immunol.* **2014**, *40*, 514–523. [CrossRef]

49. Yang, C.; Wang, L.; Zhang, H.; Wang, L.; Huang, M.; Sun, Z.; Sun, Y.; Song, L. A New Fibrinogen-Related Protein from Argopecten Irradians (AiFREP-2) with Broad Recognition Spectrum and Bacteria Agglutination Activity. *Fish Shellfish Immunol.* **2014**, *38*, 221–229. [CrossRef]

50. Gerdol, M.; Luo, Y.-J.; Satoh, N.; Pallavicini, A. Genetic and Molecular Basis of the Immune System in the Brachiopod *Lingula Anatina*. *Dev. Comp. Immunol.* **2018**, *82*, 7–30. [CrossRef]

51. Jin, Q.; Sun, Q.; Zhang, J.; Sun, L. First Characterization of Two C-Type Lectins of the Tubeworm *Alaysia Sp.* from a Deep-Sea Hydrothermal Vent. *Dev. Comp. Immunol.* **2018**, *86*, 17–25. [CrossRef]

52. Yang, J.; Wang, L.; Zhang, H.; Qiu, L.; Wang, H.; Song, L. C-Type Lectin in Chlamys Farreri (CfLec-1) Mediating Immune Recognition and Opsonization. *PLoS ONE* **2011**, *6*, e17089. [CrossRef]

53. Huang, M.; Song, X.; Zhao, J.; Mu, C.; Wang, L.; Zhang, H.; Zhou, Z.; Liu, X.; Song, L. A C-Type Lectin (AiCTL-3) from Bay Scallop Argopecten Irradians with Mannose/Galactose Binding Ability to Bind Various Bacteria. *Gene* **2013**, *531*, 31–38. [CrossRef]

54. Zelensky, A.N.; Gready, J.E. The C-Type Lectin-like Domain Superfamily. *FEBS J.* **2005**, *272*, 6179–6217. [CrossRef] [PubMed]

55. Pees, B.; Yang, W.; Zárate-Potes, A.; Schulenburg, H.; Dierking, K. High Innate Immune Specificity through Diversified C-Type Lectin-Like Domain Proteins in Invertebrates. *J. Innate Immun.* **2016**, *8*, 129–142. [CrossRef] [PubMed]

56. Drickamer, K. Evolution of Ca^{2+}-Dependent Animal Lectins11Abbreviations: CRD, Carbohydrate-Recognition Domain; EGF, Epidermal Growth Factor. In *Progress in Nucleic Acid Research and Molecular Biology*; Cohn, W.E., Moldave, K., Eds.; Academic Press: San Diego, CA, USA, 1993; Volume 45, pp. 207–232.

57. Drickamer, K.; Fadden, A.J. Genomic Analysis of C-Type Lectins. *Biochem. Soc. Symp.* **2002**, *69*, 59–72. [CrossRef] [PubMed]

58. Holmskov, U.L. Collectins and Collectin Receptors in Innate Immunity. *APMIS Suppl.* **2000**, *100*, 1–59. [CrossRef]

59. Sun, J.-J.; Lan, J.-F.; Zhao, X.-F.; Vasta, G.R.; Wang, J.-X. Binding of a C-Type Lectin's Coiled-Coil Domain to the Domeless Receptor Directly Activates the JAK/STAT Pathway in the Shrimp Immune Response to Bacterial Infection. *PLoS Pathog.* **2017**, *13*, e1006626. [CrossRef] [PubMed]

60. Vasta, G.R.; Ahmed, H.; Tasumi, S.; Odom, E.W.; Saito, K. Biological Roles of Lectins in Innate Immunity: Molecular and Structural Basis for Diversity in Self/Non-Self Recognition. In *Current Topics in Innate Immunity*; Lambris, J.D., Ed.; Springer: New York, NY, USA, 2007; pp. 389–406.

61. Li, H.; Zhang, H.; Jiang, S.; Wang, W.; Xin, L.; Wang, H.; Wang, L.; Song, L. A Single-CRD C-Type Lectin from Oyster Crassostrea Gigas Mediates Immune Recognition and Pathogen Elimination with a Potential Role in the Activation of Complement System. *Fish Shellfish Immunol.* **2015**, *44*, 566–575. [CrossRef]

62. Gerdol, M.; Venier, P. An Updated Molecular Basis for Mussel Immunity. *Fish Shellfish Immunol.* **2015**, *46*, 17–38. [CrossRef]

63. Yu, X.-Q.; Gan, H.; Kanost, R.M. Immulectin, an Inducible C-Type Lectin from an Insect, *Manduca Sexta*, Stimulates Activation of Plasma Prophenol Oxidase. *Insect Biochem. Mol. Biol.* **1999**, *29*, 585–597. [CrossRef]

64. Couser, L.M. A Mannose Receptor-like Molecule Likely Serves as the Mate Recognition Pheromone Receptor in the Male Rotifer *Brachionus Manjavacas*. Bachelor's Thesis, Georgia Institute of Technology, Atlanta, GA, USA, 2010.

65. Kishore, U.; Gaboriaud, C.; Waters, P.; Shrive, A.K.; Greenhough, T.J.; Reid, K.B.M.; Sim, R.B.; Arlaud, G.J. C1q and Tumor Necrosis Factor Superfamily: Modularity and Versatility. *Trends Immunol.* **2004**, *25*, 551–561. [CrossRef]

66. Carland, T.M.; Gerwick, L. The C1q Domain Containing Proteins: Where Do They Come from and What Do They Do? *Dev. Comp. Immunol.* **2010**, *34*, 785–790. [CrossRef]

67. Ghai, R.; Waters, P.; Roumenina, L.T.; Gadjeva, M.; Kojouharova, M.S.; Reid, K.B.M.; Sim, R.B.; Kishore, U. C1q and Its Growing Family. *Immunobiology* **2007**, *212*, 253–266. [CrossRef]

68. Hasan, I.; Gerdol, M.; Fujii, Y.; Ozeki, Y. Functional Characterization of OXYL, A SghC1qDC LacNAc-Specific Lectin from The Crinoid Feather Star Anneissia Japonica. *Mar. Drugs* **2019**, *17*, 136. [CrossRef] [PubMed]

69. Gerlach, D.; Schlott, B.; Schmidt, K.-H. Cloning and Expression of a Sialic Acid-Binding Lectin from the *Snail Cepaea* Hortensis. *FEMS Immunol. Med. Microbiol.* **2004**, *40*, 215–221. [CrossRef]

70. Xiong, X.; Li, C.; Zheng, Z.; Du, X. Novel Globular C1q Domain-Containing Protein (PmC1qDC-1) Participates in Shell Formation and Responses to Pathogen-Associated Molecular Patterns Stimulation in Pinctada *Fucata Martensii*. *Sci. Rep.* **2021**, *11*, 1105. [CrossRef] [PubMed]

71. Huang, Y.; Wu, L.; Jin, M.; Hui, K.; Ren, Q. A C1qDC Protein (HcC1qDC6) with Three Tandem C1q Domains Is Involved in Immune Response of Triangle-Shell Pearl Mussel (*Hyriopsis Cumingii*). *Front. Physiol.* **2017**, *8*, 521. [CrossRef]

72. Xu, K.; Wang, Y.; Lian, S.; Hu, N.; Chen, X.; Dai, X.; Zhang, L.; Wang, S.; Hu, J.; Hu, X.; et al. Expansion of C1Q Genes in Zhikong Scallop and Their Expression Profiling After Exposure to the Toxic Dinoflagellates. *Front. Mar. Sci.* **2021**, *8*, 453. [CrossRef]

73. Gorbushin, A.M. Derivatives of the Lectin Complement Pathway in Lophotrochozoa. *Dev. Comp. Immunol.* **2019**, *94*, 35–58. [CrossRef]

74. Vasta, G.R.; Ahmed, H.; Nita-Lazar, M.; Banerjee, A.; Pasek, M.; Shridhar, S.; Guha, P.; Fernández-Robledo, J.A. Galectins as Self/Non-Self Recognition Receptors in Innate and Adaptive Immunity: An Unresolved Paradox. *Front. Immun.* **2012**, *3*, 199. [CrossRef]

75. Song, X.; Zhang, H.; Wang, L.; Zhao, J.; Mu, C.; Song, L.; Qiu, L.; Liu, X. A Galectin with Quadruple-Domain from Bay Scallop *Argopecten Irradians* Is Involved in Innate Immune Response. *Dev. Comp. Immunol.* **2011**, *35*, 592–602. [CrossRef]

76. Zhang, D.; Jiang, S.; Hu, Y.; Cui, S.; Guo, H.; Wu, K.; Li, Y.; Su, T. A Multidomain Galectin Involved in Innate Immune Response of Pearl Oyster *Pinctada Fucata*. *Dev. Comp. Immunol.* **2011**, *35*, 1–6. [CrossRef]

77. Gorbushin, A.M.; Borisova, E.A. Lectin-like Molecules in Transcriptome of *Littorina Littorea* Hemocytes. *Dev. Comp. Immunol.* **2015**, *48*, 210–220. [CrossRef] [PubMed]

78. Ruan, L.; Xu, H.; Lin, W.; Shi, H.; Cui, Z.; Xu, X. A Novel Beta-Galactose-Specific Lectin of the Tubeworm, *Ridgeia Piscesae*, from the Hydrothermal Vent. *Acta Oceanologica Sinica* **2017**, *36*, 61–67. [CrossRef]

79. Vasta, G.R.; Feng, C.; Bianchet, M.A.; Bachvaroff, T.R.; Tasumi, S. Structural, Functional, and Evolutionary Aspects of Galectins in Aquatic Mollusks: From a Sweet Tooth to the Trojan Horse. *Fish Shellfish Immunol.* **2015**, *46*, 94–106. [CrossRef] [PubMed]

80. Popa, S.J.; Stewart, S.E.; Moreau, K. Unconventional Secretion of Annexins and Galectins. *Semin. Cell Dev. Biol.* **2018**, *83*, 42. [CrossRef] [PubMed]

81. Cummings, R.D.; Schnaar, R. R-Type Lectins. In *Essentials of Glycobiology*, 3rd ed.; Cold Spring Harbor Laboratory Press: New York, NY, USA, 2017; Chapter 31.

82. Kawsar, S.M.A.; Takeuchi, T.; Kasai, K.; Fujii, Y.; Matsumoto, R.; Yasumitsu, H.; Ozeki, Y. Glycan-Binding Profile of a D-Galactose Binding Lectin Purified from the Annelid, *Perinereis Nuntia Ver. Vallata*. *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* **2009**, *152*, 382–389. [CrossRef]

83. Hirabayashi, J.; Dutta, S.K.; Kasai, K. Novel Galactose-Binding Proteins in Annelida. Characterization of 29-KDa Tandem Repeat-Type Lectins from the Earthworm *Lumbricus Terrestris*. *J. Biol. Chem.* **1998**, *273*, 14450–14460. [CrossRef]

84. Kawsar, S.M.A.; Hasan, I.; Rajia, S.; Koide, Y.; Fujii, Y.; Hayashi, R.; Yamada, M.; Ozeki, Y. Diverse Localization Patterns of an R-Type Lectin in Marine Annelids. *Molecules* **2021**, *26*, 4799. [CrossRef]

85. Fujii, Y.; Gerdol, M.; Kawsar, S.M.A.; Hasan, I.; Spazzali, F.; Yoshida, T.; Ogawa, Y.; Rajia, S.; Kamata, K.; Koide, Y.; et al. A GM1b/Asialo-GM1 Oligosaccharide-Binding R-Type Lectin from Purplish Bifurcate Mussels *Mytilisepta Virgata* and Its Effect on MAP Kinases. *FEBS J.* **2020**, *287*, 2612–2630. [CrossRef]

86. Hasan, I.; Sugawara, S.; Fujii, Y.; Koide, Y.; Terada, D.; Iimura, N.; Fujiwara, T.; Takahashi, K.G.; Kojima, N.; Rajia, S.; et al. MytiLec, a Mussel R-Type Lectin, Interacts with Surface Glycan Gb3 on Burkitt's Lymphoma Cells to Trigger Apoptosis through Multiple Pathways. *Mar. Drugs* **2015**, *13*, 7377–7389. [CrossRef]

87. Fujii, Y.; Dohmae, N.; Takio, K.; Kawsar, S.M.A.; Matsumoto, R.; Hasan, I.; Koide, Y.; Kanaly, R.A.; Yasumitsu, H.; Ogawa, Y.; et al. A Lectin from the Mussel Mytilus Galloprovincialis Has a Highly Novel Primary Structure and Induces Glycan-Mediated Cytotoxicity of Globotriaosylceramide-Expressing Lymphoma Cells. *J. Biol. Chem.* **2012**, *287*, 44772–44783. [CrossRef]

88. Shimizu, T.; Vassylyev, D.; Kido, S.; Doi, Y.; Morikawa, K. Crystal Structure of Vitelline Membrane Outer Layer Protein I (VMO-I): A Folding Motif with Homologous Greek Key Structures Related by an Internal Three-Fold Symmetry. *EMBO J.* **1994**, *13*, 1003–1010. [CrossRef] [PubMed]

89. Sankaranarayanan, R.; Sekar, K.; Banerjee, R.; Sharma, V.; Surolia, A.; Vijayan, M. A Novel Mode of Carbohydrate Recognition in Jacalin, a Moraceae Plant Lectin with a Beta-Prism Fold. *Nat. Struct. Biol.* **1996**, *3*, 596–603. [CrossRef] [PubMed]

90. Lee, X.; Thompson, A.; Zhang, Z.; Ton-that, H.; Biesterfeldt, J.; Ogata, C.; Xu, L.; Johnston, R.A.Z.; Young, N.M. Structure of the Complex of Maclura PomiferaAgglutinin and the T-Antigen Disaccharide, Galβ1,3GalNAc*. *J. Biol. Chem.* **1998**, *273*, 6312–6318. [CrossRef] [PubMed]

91. Rosa, J.C.; De Oliveira, P.S.; Garratt, R.; Beltramini, L.; Resing, K.; Roque-Barreira, M.C.; Greene, L.J. KM+, a Mannose-Binding Lectin from *Artocarpus Integrifolia*: Amino Acid Sequence, Predicted Tertiary Structure, Carbohydrate Recognition, and Analysis of the Beta-Prism Fold. *Protein Sci.* **1999**, *8*, 13–24. [CrossRef] [PubMed]

92. Rao, K.N.; Suresh, C.G.; Katre, U.V.; Gaikwad, S.M.; Khan, M.I. Two Orthorhombic Crystal Structures of a Galactose-Specific Lectin from *Artocarpus Hirsuta* in Complex with Methyl-Alpha-D-Galactose. *Acta Crystallogr. D Biol. Crystallogr.* **2004**, *60*, 1404–1412. [CrossRef]

93. Suetake, T.; Tsuda, S.; Kawabata, S.; Miura, K.; Iwanaga, S.; Hikichi, K.; Nitta, K.; Kawano, K. Chitin-Binding Proteins in Invertebrates and Plants Comprise a Common Chitin-Binding Structural Motif. *J. Biol. Chem.* **2000**, *275*, 17929–17932. [CrossRef]

94. Kawabata, S.; Nagayama, R.; Hirata, M.; Shigenaga, T.; Agarwala, K.L.; Saito, T.; Cho, J.; Nakajima, H.; Takagi, T.; Iwanaga, S. Tachycitin, a Small Granular Component in Horseshoe Crab Hemocytes, Is an Antimicrobial Protein with Chitin-Binding Activity. *J. Biochem.* **1996**, *120*, 1253–1260. [CrossRef]

95. Meng, D.-M.; Dai, H.-X.; Gao, X.-F.; Zhao, J.-F.; Guo, Y.-J.; Ling, X.; Dong, B.; Zhang, Z.-Q.; Fan, Z.-C. Expression, Purification and Initial Characterization of a Novel Recombinant Antimicrobial Peptide Mytichitin-A in *Pichia Pastoris*. *Protein Expr. Purif.* **2016**, *127*, 35–43. [CrossRef]

96. Bauters, L.; Naalden, D.; Gheysen, G. The Distribution of Lectins across the *Phylum Nematoda*: A Genome-Wide Search. *Int. J. Mol. Sci.* **2017**, *18*, 91. [CrossRef]

97. Odom, E.W.; Vasta, G.R. Characterization of a Binary Tandem Domain F-Type Lectin from Striped Bass (*Morone Saxatilis*). *J. Biol. Chem.* **2006**, *281*, 1698–1713. [CrossRef]

98. Bianchet, M.A.; Odom, E.W.; Vasta, G.R.; Amzel, L.M. A Novel Fucose Recognition Fold Involved in Innate Immunity. *Nat. Struct. Biol.* **2002**, *9*, 628–634. [CrossRef] [PubMed]

99. Arivalagan, J.; Marie, B.; Sleight, V.A.; Clark, M.S.; Berland, S.; Marie, A. Shell Matrix Proteins of the Clam, *Mya Truncata*: Roles beyond Shell Formation through Proteomic Study. *Mar. Genom.* **2016**, *27*, 69–74. [CrossRef] [PubMed]

100. Moy, G.W.; Springer, S.A.; Adams, S.L.; Swanson, W.J.; Vacquier, V.D. Extraordinary Intraspecific Diversity in Oyster Sperm Bindin. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 1993–1998. [CrossRef] [PubMed]

101. Ruben, E.A.; Rau, M.J.; Fitzpatrick, J.A.J.; Di Cera, E. Cryo-EM Structures of Human Coagulation Factors V and Va. *Blood* **2021**, *137*, 3137–3144. [CrossRef]

102. Mathieu, S.V.; Aragão, K.S.; Imberty, A.; Varrot, A. Discoidin I from Dictyostelium Discoideum and Interactions with Oligosaccharides: Specificity, Affinity, Crystal Structures, and Comparison with Discoidin II. *J. Mol. Biol.* **2010**, *400*, 540–554. [CrossRef] [PubMed]

103. Tateno, H. SUEL-Related Lectins, a Lectin Family Widely Distributed throughout Organisms. *Biosci. Biotechnol. Biochem.* **2010**, *74*, 1141–1144. [CrossRef]

104. Ozeki, Y.; Matsui, T.; Suzuki, M.; Titani, K. Amino Acid Sequence and Molecular Characterization of a D-Galactoside-Specific Lectin Purified from Sea Urchin (*Anthocidaris Crassispina*) Eggs. *Biochemistry* **1991**, *30*, 2391–2394. [CrossRef]

105. Carneiro, R.F.; Teixeira, C.S.; de Melo, A.A.; de Almeida, A.S.; Cavada, B.S.; de Sousa, O.V.; da Rocha, B.A.M.; Nagano, C.S.; Sampaio, A.H. L-Rhamnose-Binding Lectin from Eggs of the *Echinometra Lucunter*: Amino Acid Sequence and Molecular Modeling. *Int. J. Biol. Macromol.* **2015**, *78*, 180–188. [CrossRef]

106. Gasparini, F.; Franchi, N.; Spolaore, B.; Ballarin, L. Novel Rhamnose-Binding Lectins from the Colonial Ascidian Botryllus Schlosseri. *Dev. Comp. Immunol.* **2008**, *32*, 1177–1191. [CrossRef]

107. Franchi, N.; Schiavon, F.; Carletto, M.; Gasparini, F.; Bertoloni, G.; Tosatto, S.C.E.; Ballarin, L. Immune Roles of a Rhamnose-Binding Lectin in the Colonial Ascidian Botryllus Schlosseri. *Immunobiology* **2011**, *216*, 725–736. [CrossRef]

108. Naganuma, T.; Ogawa, T.; Hirabayashi, J.; Kasai, K.; Kamiya, H.; Muramoto, K. Isolation, Characterization and Molecular Evolution of a Novel Pearl Shell Lectin from a Marine Bivalve, Pteria Penguin. *Mol. Divers.* **2006**, *10*, 607–618. [CrossRef] [PubMed]

109. Uhlenbruck, G.; Prokop, O. An Agglutinin from Helix Pomatia, Which Reacts with Terminal N-Acetyl-D-Galactosamine. *Vox Sang.* **1966**, *11*, 519–520. [CrossRef] [PubMed]

110. Sanchez, J.-F.; Lescar, J.; Chazalet, V.; Audfray, A.; Gagnon, J.; Alvarez, R.; Breton, C.; Imberty, A.; Mitchell, E.P. Biochemical and Structural Analysis of Helix Pomatia Agglutinin: A hexameric lectin with a novel fold. *J. Biol. Chem.* **2006**, *281*, 20171–20180. [CrossRef] [PubMed]

111. Gerdol, M. Immune-Related Genes in Gastropods and Bivalves: A Comparative Overview. *Invertebr. Surviv. J.* **2017**, *14*, 95–111.

112. Gerdol, M.; Fujii, Y.; Hasan, I.; Koike, T.; Shimojo, S.; Spazzali, F.; Yamamoto, K.; Ozeki, Y.; Pallavicini, A.; Fujita, H. The Purplish Bifurcate Mussel Mytilisepta Virgata Gene Expression Atlas Reveals a Remarkable Tissue Functional Specialization. *BMC Genom.* **2017**, *18*, 590. [CrossRef]

113. Ogawa, T.; Sato, R.; Naganuma, T.; Liu, K.; Lakudzala, A.E.; Muramoto, K.; Osada, M.; Yoshimi, K.; Hiemori, K.; Hirabayashi, J.; et al. Glycan Binding Profiling of Jacalin-Related Lectins from the Pteria Penguin Pearl Shell. *Int. J. Mol. Sci.* **2019**, *20*, 4629. [CrossRef]

114. Kanagawa, M.; Satoh, T.; Ikeda, A.; Nakano, Y.; Yagi, H.; Kato, K.; Kojima-Aikawa, K.; Yamaguchi, Y. Crystal Structures of Human Secretory Proteins ZG16p and ZG16b Reveal a Jacalin-Related β-Prism Fold. *Biochem. Biophys. Res. Commun.* **2011**, *404*, 201–205. [CrossRef]

115. Garénaux, E.; Kanagawa, M.; Tsuchiyama, T.; Hori, K.; Kanazawa, T.; Goshima, A.; Chiba, M.; Yasue, H.; Ikeda, A.; Yamaguchi, Y.; et al. Discovery, Primary, and Crystal Structures and Capacitation-Related Properties of a Prostate-Derived Heparin-Binding Protein WGA16 from Boar Sperm. *J. Biol. Chem.* **2015**, *290*, 5484–5501. [CrossRef]

116. Jia, N.; Liu, N.; Cheng, W.; Jiang, Y.; Sun, H.; Chen, L.; Peng, J.; Zhang, Y.; Ding, Y.; Zhang, Z.; et al. Structural Basis for Receptor Recognition and Pore Formation of a Zebrafish Aerolysin-like Protein. *EMBO Rep.* **2016**, *17*, 235–248. [CrossRef]

117. Carneiro, R.F.; Torres, R.C.F.; Chaves, R.P.; de Vasconcelos, M.A.; de Sousa, B.L.; Goveia, A.C.R.; Arruda, F.V.; Matos, M.N.C.; Matthews-Cascon, H.; Freire, V.N.; et al. Purification, Biochemical Characterization, and Amino Acid Sequence of a Novel Type of Lectin from Aplysia Dactylomela Eggs with Antibacterial/Antibiofilm Potential. *Mar. Biotechnol.* **2017**, *19*, 49–64. [CrossRef]

118. Huang, G.; Huang, S.; Yan, X.; Yang, P.; Li, J.; Xu, W.; Zhang, L.; Wang, R.; Yu, Y.; Yuan, S.; et al. Two Apextrin-like Proteins Mediate Extracellular and Intracellular Bacterial Recognition in Amphioxus. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 13469–13474. [CrossRef] [PubMed]

119. Wang, W.; Gong, C.; Han, Z.; Lv, X.; Liu, S.; Wang, L.; Song, L. The Lectin Domain Containing Proteins with Mucosal Immunity and Digestive Functions in Oyster Crassostrea Gigas. *Fish Shellfish Immunol.* **2019**, *89*, 237–247. [CrossRef] [PubMed]

120. Dang, L.; Van Damme, E.J.M. Genome-Wide Identification and Domain Organization of Lectin Domains in Cucumber. *Plant Physiol. Biochem.* **2016**, *108*, 165–176. [CrossRef] [PubMed]

121. Lubyaga, Y.A.; Dolgikh, A.V.; Drozdova, P.B.; Nazarova, A.A.; Timofeyev, M.A. Transcriptome-Based Analysis of the Diversity of Membrane-Bound Lectins in Baikal Amphipods *Eulimnogammarus* Sp. and the Holarctic Amphipod *Gammarus Lacustris*. *Limnol. Freshw. Biol.* **2020**, 797–798. [CrossRef]

122. Zárate-Potes, A.; Ocampo, I.D.; Cadavid, L.F. The Putative Immune Recognition Repertoire of the Model Cnidarian Hydractinia Symbiolongicarpus Is Large and Diverse. *Gene* **2019**, *684*, 104–117. [CrossRef]

123. Simão, F.A.; Waterhouse, R.M.; Ioannidis, P.; Kriventseva, E.V.; Zdobnov, E.M. BUSCO: Assessing Genome Assembly and Annotation Completeness with Single-Copy Orthologs. *Bioinformatics* **2015**, *31*, 3210–3212. [CrossRef]

124. Angata, T.; Brinkman-Van der Linden, E. I-Type Lectins. *Biochim. Biophys. Acta* **2002**, *1572*, 294–316. [CrossRef]

125. Liu, C.; Jiang, S.; Wang, M.; Wang, L.; Chen, H.; Xu, J.; Lv, Z.; Song, L. A Novel Siglec (CgSiglec-1) from the Pacific Oyster (*Crassostrea Gigas*) with Broad Recognition Spectrum and Inhibitory Activity to Apoptosis, Phagocytosis and Cytokine Release. *Dev. Comp. Immunol.* **2016**, *61*, 136–144. [CrossRef]

126. Sulzenbacher, G.; Roig-Zamboni, V.; Peumans, W.J.; Henrissat, B.; van Damme, E.J.M.; Bourne, Y. Structural Basis for Carbohydrate Binding Properties of a Plant Chitinase-like Agglutinin with Conserved Catalytic Machinery. *J. Struct. Biol.* **2015**, *190*, 115–121. [CrossRef]

127. Hasan, I.; Gerdol, M.; Fujii, Y.; Rajia, S.; Koide, Y.; Yamamoto, D.; Kawsar, S.M.A.; Ozeki, Y. CDNA and Gene Structure of MytiLec-1, A Bacteriostatic R-Type Lectin from the Mediterranean Mussel (*Mytilus Galloprovincialis*). *Mar. Drugs* **2016**, *14*, 92. [CrossRef]

128. Han, J.; Park, J.C.; Choi, B.-S.; Kim, M.-S.; Kim, H.-S.; Hagiwara, A.; Park, H.G.; Lee, B.-Y.; Lee, J.-S. The Genome of the Marine Monogonont Rotifer Brachionus Plicatilis: Genome-Wide Expression Profiles of 28 Cytochrome P450 Genes in Response to Chlorpyrifos and 2-Ethyl-Phenanthrene. *Aquat. Toxicol.* **2019**, *214*, 105230. [CrossRef] [PubMed]

129. Kim, H.-S.; Lee, B.-Y.; Han, J.; Jeong, C.-B.; Hwang, D.-S.; Lee, M.-C.; Kang, H.-M.; Kim, D.-H.; Kim, H.-J.; Papakostas, S.; et al. The Genome of the Freshwater Monogonont Rotifer Brachionus Calyciflorus. *Mol. Ecol. Resour.* **2018**, *18*, 646–655. [CrossRef] [PubMed]

130. Kim, D.-H.; Kim, M.-S.; Hagiwara, A.; Lee, J.-S. The Genome of the Minute Marine Rotifer Proales Similis: Genome-Wide Identification of 401 G Protein-Coupled Receptor (GPCR) Genes. *Comp. Biochem. Physiol. D Genom. Proteom.* **2021**, *39*, 100861. [CrossRef] [PubMed]

131. Petersen, T.N.; Brunak, S.; von Heijne, G.; Nielsen, H. SignalP 4.0: Discriminating Signal Peptides from Transmembrane Regions. *Nat. Meth.* **2011**, *8*, 785–786. [CrossRef]

132. Bendtsen, J.D.; Jensen, L.J.; Blom, N.; Von Heijne, G.; Brunak, S. Feature-Based Prediction of Non-Classical and Leaderless Protein Secretion. *Protein Eng. Des. Sel.* **2004**, *17*, 349–356. [CrossRef]

133. Krogh, A.; Larsson, B.; von Heijne, G.; Sonnhammer, E.L. Predicting Transmembrane Protein Topology with a Hidden Markov Model: Application to Complete Genomes. *J. Mol. Biol.* **2001**, *305*, 567–580. [CrossRef]

134. Käll, L.; Krogh, A.; Sonnhammer, E.L.L. A Combined Transmembrane Topology and Signal Peptide Prediction Method. *J. Mol. Biol.* **2004**, *338*, 1027–1036. [CrossRef]

135. Finn, R.D.; Clements, J.; Eddy, S.R. HMMER Web Server: Interactive Sequence Similarity Searching. *Nucl. Acids Res.* **2011**, *39*, W29–W37. [CrossRef]

136. Jones, P.; Binns, D.; Chang, H.-Y.; Fraser, M.; Li, W.; McAnulla, C.; McWilliam, H.; Maslen, J.; Mitchell, A.; Nuka, G.; et al. InterProScan 5: Genome-Scale Protein Function Classification. *Bioinformatics* **2014**, *30*, 1236–1240. [CrossRef]

137. Söding, J.; Biegert, A.; Lupas, A.N. The HHpred Interactive Server for Protein Homology Detection and Structure Prediction. *Nucleic Acids Res.* **2005**, *33*, W244–W248. [CrossRef]

138. Edgar, R.C. MUSCLE: Multiple Sequence Alignment with High Accuracy and High Throughput. *Nucleic Acids Res.* **2004**, *32*, 1792–1797. [CrossRef] [PubMed]

139. Talavera, G.; Castresana, J. Improvement of Phylogenies after Removing Divergent and Ambiguously Aligned Blocks from Protein Sequence Alignments. *Syst. Biol.* **2007**, *56*, 564–577. [CrossRef] [PubMed]

140. Huelsenbeck, J.P.; Ronquist, F. MRBAYES: Bayesian Inference of Phylogenetic Trees. *Bioinformatics* **2001**, *17*, 754–755. [CrossRef] [PubMed]
141. Darriba, D.; Posada, D.; Kozlov, A.M.; Stamatakis, A.; Morel, B.; Flouri, T. ModelTest-NG: A New and Scalable Tool for the Selection of DNA and Protein Evolutionary Models. *Mol. Biol. Evol.* **2020**, *37*, 291–294. [CrossRef]
142. Rambaut, A.; Drummond, A.J.; Xie, D.; Baele, G.; Suchard, M.A. Posterior Summarization in Bayesian Phylogenetics Using Tracer 1.7. *Syst. Biol.* **2018**, *67*, 901–904. [CrossRef]