Plant lectins and their many roles: Carbohydrate-binding and beyond

Sushma Naithani a,∗, Sneha Sudha Komath b, Arthur Nonomura c, Govindjee Govindjee d

ABSTRACT

Lectins are ubiquitous proteins that reversibly bind to specific carbohydrates and, thus, serve as readers of the sugar code. In photosynthetic organisms, lectin family proteins play important roles in capturing and releasing photosynthates via an endogenous lectin cycle. Often, lectin proteins consist of one or more lectin domains in combination with other types of domains. This structural diversity of lectins is the basis for their current classification, which is consistent with their diverse functions in cell signaling associated with growth and development, as well as in the plant’s response to biotic, symbiotic, and abiotic stimuli. Furthermore, the lectin family shows evolutionary expansion that has distinct clade-specific signatures. Although the function(s) of many plant lectin family genes are unknown, studies in the model plant Arabidopsis thaliana have provided insights into their diverse roles. Here, we have used a biocuration approach rooted in the critical review of scientific literature and information available in the public genomic databases to summarize the expression, localization, and known functions of lectins in Arabidopsis. A better understanding of the structure and function of lectins is expected to aid in improving agricultural productivity through the manipulation of candidate genes for breeding climate-resilient crops, or by regulating metabolic pathways by applications of plant growth regulators.

ARTICLE INFO

Keywords:
Cell signaling
Glycoregulation
Gene biocuration
Lectins
Plant stress response
Receptor-like kinases

1. Introduction

Carbohydrates, some of the most versatile and abundant molecules in nature, are major sources of energy for living organisms, meeting both short- and long-term energy demands in the form of free and stored metabolites, such as sugars, starch, and glycogen. They are also constituents of structural components, such as the monosaccharides that form the backbone of nucleic acids (e.g., ribose, and deoxyribose), lipopolysaccharides, glycosaminoglycans, and polysaccharides, such as cellulose, pectin, agar, carrageenan, and chitin, that provide shape and integrity to the cells. In addition, carbohydrates modify lipids and proteins, altering their structure, folding, localization, and function (Varki and Gagneux, 2017). Further, they are actively involved in immune recognition, and provide the basis for self vs. non-self-recognition (Agostino et al., 2011; O’Neill et al., 2017).

The ability of carbohydrates to participate in diverse functions is fundamental to their distinctive chemical characteristics. For example, even a simple C6 (hexose) sugar, such as glucose, can give rise to sixteen possible isomers (2n, where n is the number of chiral centers in the molecule). On the other hand, a disaccharide, such as sucrose, with nine chiral centers, has 512 (29) possible isomeric forms. Moreover, longer chains of oligosaccharides introduce even greater biochemical diversity. Indeed, these macromolecules may serve as identification tags. However, for a highly specific identification system to be useful, living systems must be able to: 1) synthesize, modify, and degrade many of these carbohydrates in vivo; and 2) differentiate between the various structural codes that these carbohydrates generate. The former is achieved by glycosyltransferases and glycosidases, which are broadly grouped as carbohydrate-active enzymes (CAZy). In contrast, the latter is achieved either by glycan-binding proteins (GBPs) or by antibodies (see Fig. 1), depending on where they function (Henrissat et al., 2017).

Glycan Binding Proteins (GBPs) are classified into two broad groups, namely glycosaminoglycan (GAG)-binding proteins and lectins (Taylor et al., 2017). The GAG-binding proteins contain clusters of positively charged amino acids that allow them to recognize, via electrostatic interactions, patterns of negatively charged carboxylate and sulfate groups on the polysaccharides, such as heparan, chondroitin, dermatan, and keratan sulfates (Taylor et al., 2017). Lectins, on the other hand, have...
shallow carbohydrate recognition domains (CRDs), which allow one or more terminal carbohydrate residues of a polysaccharide to fit in, primarily by way of hydrogen-bonded interactions (Loris et al., 1998; Sharma and Surolia, 1997; Sharon and Lis, 2001). Further, residues lining the CRD may also interact with the carbohydrate ligands through van der Waals bonds, ionic bonds, and CH-π interactions (with aromatic/hydrophobic residues) (Kiessling, 2018; Spiwok, 2017). Thus, unlike enzymes, lectins are proteins that have at least one non-catalytic domain capable of reversibly binding carbohydrates. This is a critical distinction that determines the premier role of lectins as readers of the glycocode, also known as the sugar code (Dedola et al., 2020; Nilsson, 2011). We emphasize that the rapidly growing fields of genomics, transcriptomics, and glycomics are extremely important for our understanding of these protein-carbohydrate interactions.

In this review, we provide a historical introduction to lectins, their classification, and summarize the current knowledge about the roles plant lectins play in modulating the availability of photosynthates, and in cell signaling associated with various biological processes. First, we discuss the importance of the structure of the carbohydrate recognition domains in lectins and their interactions with carbohydrates. Then, we describe insights gained from high-throughput approaches of genomics, transcriptomics, proteomics, glycomics, and cell biology, to understand the function of lectin family members in Arabidopsis thaliana and other plant species including crops.

1.1. History of lectin research

In 1888, Peter Hermann Stillmark discovered ricin, a lectin, from extracts of castor beans (reviewed by Olsnes, 2004; Polito et al., 2019) and soon thereafter, another lectin, abrin, was discovered (Olsnes, 2004). Lectins were initially called agglutinins/haemagglutinins/phytoagglutinins because of their ability to agglutinate or cross-link animal blood cells (André et al., 2015). The idea that their selectivity was based on carbohydrate recognition followed much later, after Sumner and Howell (1936) had demonstrated the carbohydrate specificity of concanavalin A (Con A) and after Watkins and Morgan (1952) had discovered that the agglutination of the blood group “O” cells, by the lectin from eel serum, could be inhibited by simple sugars (Sumner and Howell, 1936; Watkins and Morgan, 1952). After it was shown that most lectins have specificities that could distinguish between different human blood types, they were named lectins, based on the Latin word ‘lego’ (noun), meaning, chosen; or ‘legere’ (verb), which means ‘to choose or select’ (Boyd and Reguera, 1949; Renkonen KO, 1948; Boyd and Shapleigh, 1954).

The word “lectin” was already in use when the first mammalian lectin, the asialoglycoprotein receptor, was discovered (Ashwell and...
1.2. Classification of plant lectins

As several hundreds of lectins were being discovered, a new challenge for their rational classification emerged. One possible mode was to simply use the organism in which they were found: microbial lectins, animal lectins, and plant lectins. However, within each kingdom, various carbohydrate specificities were found, requiring a subclassification of these lectins. Several different criteria were examined for subgrouping of the lectins, including sugar specificities of various lectins. However, as crystal structures of the lectins became available, it soon became apparent that there were multiple molecular mechanisms by which two lectins from different plants recognize the same sugar; alternatively, two lectins from the same plant were shown to bind different sugars (Abhinav and Vijayan, 2014). However, using sugar specificities alone as a classification criterion, for the extremely large group of plant lectins, did not produce any kind of generalization about the biochemical and biophysical properties of the lectins themselves (Peumans and Van Damme, 1995). Another classification method of lectins considered the nature and organization of the CRDs that plant lectins possess (Tsaneva and Van Damme, 2020). In this method, lectins considered the biochemical and biophysical properties of the lectins themselves (Peumans and Van Damme, 1995). Another classification method of lectins considered the nature and organization of the CRDs that plant lectins possess (Tsaneva and Van Damme, 2020). In this method, lectins were classified as superlectins (lectins with a single CRD), hololectins (lectins with two or more identical CRDs), chimerolectins (those which possess at least one CRD and one other type of conserved domain that had a different function), and superlectins (those containing two or more non-identical CRDs with specificities for structurally different sugars).

Morell, 1974). Later, many lectins were identified in bacteria, fungi, protozoans, invertebrates, and vertebrates (Kilpatrick, 2002; Nizet et al., 2017). With these findings came the realization that lectins are omnipresent in nature and play a crucial role in deciphering the glycocode of life. A timeline of the many discoveries in the field of lectin research is summarized in Table 1.

For completeness, we refer the readers to several excellent reviews published previously (see Gabius, 1997; Habermann et al., 2021; Hibayaishi, 2014; Lanno and Van Damme, 2010; Rüdiger and Gabius, 2001; Tsaneva and Van Damme, 2020; Variki et al., 2009) covering various aspects of lectins, including their discovery, functions, and applications.

1.2. Classification of plant lectins

As several hundreds of lectins were being discovered, a new challenge for their rational classification emerged. One possible mode was to simply use the organism in which they were found: microbial lectins, animal lectins, and plant lectins. However, within each kingdom, various carbohydrate specificities were found, requiring a subclassification of these lectins. Several different criteria were examined for subgrouping of the lectins, including sugar specificities of various lectins. However, as crystal structures of the lectins became available, it soon became apparent that there were multiple molecular mechanisms by which two lectins from different plants recognize the same sugar; alternatively, two lectins from the same plant were shown to bind different sugars (Abhinav and Vijayan, 2014). However, using sugar specificities alone as a classification criterion, for the extremely large group of plant lectins, did not produce any kind of generalization about the biochemical and biophysical properties of the lectins themselves (Peumans and Van Damme, 1995). Another classification method of lectins considered the nature and organization of the CRDs that plant lectins possess (Tsaneva and Van Damme, 2020). In this method, lectins were classified as superlectins (lectins with a single CRD), hololectins (lectins with two or more identical CRDs), chimerolectins (those which possess at least one CRD and one other type of conserved domain that had a different function), and superlectins (those containing two or more non-identical CRDs with specificities for structurally different sugars).

However, despite structural conservation, the identity of the glycan that a CRD binds is determined by finer molecular details of the binding site and its interactions with the sugar moiety (Fig. 2A). For example, the CRDs of all the legume lectins are made of what is called the ‘jelly roll’ fold, and these are dependent on two divalent cations for carbohydrate-binding: a Ca²⁺ within the CRD, and a Mn²⁺ adjacent to it (Loris et al., 1998). In addition, a flexible loop in the CRD contributes to defining the actual structure of the sugar-binding pocket and determining its ligand specificity. Thus, some legume lectin monomers recognize primarily the equatorially oriented C4-hydroxyl of glucose (Glc)/mannose (Man),
while others recognize the axially oriented C4-hydroxyl of galactose (Gal)/N-acetylgalactosamine (GalNAc). For the same reason, legume lectin monomers may also produce fucose-specific or chitobiose-specific CRDs or even CRDs that bind complex carbohydrate structures.

Many CRDs, including that of the *Euonymus europaeus* lectin (EUL) and the L-type lectins, have extended ligand binding sites (Fig. 2B) which enable oligosaccharides to bind to them with higher specificity and greater avidity than the monosaccharides (Agostino et al., 2015; Loris et al., 1998). Such extended binding sites, or subsites, normally show cooperativity in ligand binding, which explains the higher specificity and affinity of many lectins/CRDs for oligosaccharides. For example, ConA binds the trimannopyranoside Manα1–3(Manα1–6)-Man-α-Me with much greater affinity and specificity than Man or Glc monomers (Loris et al., 1998). Often, an additional hydrophobic patch, adjacent to the carbohydrate-binding site in a CRD, enables binding of the lipid tails of a glycolipid or other hydrophobic entities attached to the carbohydrate, thereby enhancing the interaction with sugar-bound entities (Fig. 2C) (Loris et al., 1998; Sharma and Surolia, 1997). A carbohydrate oligomer may be bound by two or more CRDs, resulting in the clustering of lectin receptors in the presence of the carbohydrate (Fig. 2D). A CRD bound to a domain with a different functionality (for example, to a kinase domain of receptor-like kinases) results in multifunctional proteins (Fig. 2E). At a higher level of organization, two or more CRDs or monomers may form homomeric or heteromeric complexes that are capable of multivalent ligand binding or cross-linking of carbohydrates on two different cells to produce carbohydrate aggregation, glycoprotein precipitation or cell agglutination (Fig. 2F) (Abhinav and Vijayan, 2014; Huwa et al., 2021). Thus, a scheme based on the nature and organization of the CRDs neither enabled an understanding of the general principles of carbohydrate-binding, nor did it permit prediction of the biochemical and biophysical properties of newly identified lectins. Moreover, it did not allow us to search in genome databases for the identification of new lectins.

Evidently, carbohydrate-binding is just one of the many properties of a lectin (Barondes, 1988; Komath et al., 2006). Many plant lectins bind to plant growth regulators (PGRs), phytohormones, or to small peptides, in addition to saccharides. In fact, multi-domain multi-functional lectins have been found to be widely distributed across the plant kingdom (Etzler, 1985). Hence, the current, more widely accepted classification method is based on the sequence and the conserved three-dimensional structures of the lectin motif (Taylor et al., 2017; Tsaneva and Van Damme, 2020). In this system, each subfamily is named after the best characterized protein of this group. Indeed, such an approach permits a generalized classification of lectins into subfamilies from all the kingdoms of life. A summary of various lectin gene subfamilies and their distribution across all kingdoms of life is shown above in Fig. 1 (for examples, see also Table 2).

### 1.3. Expansion and diversification of lectin gene family in plant genomes

In general, lectins are abundant in the plant kingdom. To gain an insight about lectin domain coding genes across a broad spectrum of photoautotrophs, we searched for lectin motifs* in the Gramene

---

*Fig. 2. Structural diversity of lectin proteins may influence binding of sugar moieties to the carbohydrate recognition domains (CRDs) and other functions. A-F show examples of various organizations of CRD in combinations with other sites and their potential outcomes for macromolecular interactions.*
knowledgebase (http://gramene.org). Gramene, a resource for comparative functional genomics in crops and model plant species, currently hosts fully sequenced genomes and gene pages for 93 species, and systems-level pathway networks for 106 species (Tello-Ruiz et al., 2021). For our analysis, we selected 48 species from the Gramene representing a broad spectrum of photoautotrophs for the analysis of the lectin family (see Fig. 3). We note that land plants contain a higher number of the lectin coding genes and show clade-specific variations, as

Fig. 3. Variation in the total number of lectin coding genes (representing all subfamily members present within an organism) in 48 species of photoautotrophs ranging from unicellular algae to higher plants ordered according to their phylogenetic relationship (e.g., algae, bryophytes & lycophytes, basal angiosperms, monocots, and dicots). The image with statistics of lectin genes was generated using the Gramene knowledgebase (http://gramene.org) by searching for lectin motifs* (the wildcard symbol * is commonly used for broadening a search/retrieving similar words).
well as show increases due to ploidy level.

In a previously published study, a comparison of 38 plant genomes revealed that subfamilies of lectins show species or clade-specific expansion (Van Holle and Van Damme, 2019). Agaricus bisporus agglutinin (ABA) is present only in the bryophytes (e.g., Marchantia polymorpha and Sphagnum fallax), but the cyanovirin domain is found also in lycophytes, and in ferns. The ABA and cyanovirin domains of fungal origin are absent in the higher plants and their presence in lower plants is attributed to horizontal gene transfer events from fungal ancestors (Van Holle and Van Damme, 2019). The amaranthain domains are only found in vascular plants (lycophytes, ferns, gymnosperms, and angiosperms), but are not ubiquitous. The Euonymus europaeus lectin (EUL) family members are found in most land plants, including the bryophytes (Van Holle and Van Damme, 2019). Interestingly, both stomata and the EUL domain show their first presence in the bryophyte lineage (during evolution), and a member of this family, ArathEULS3, has been shown to play a role in stomatal closure and drought response (Chater et al., 2017; Van Holle and Van Damme, 2019; Van Hove et al., 2015). The Galanthus nivalis agglutinin (GNA), Lysin-motif (LysM), jacalin, ricin-B, legume lectin, and malectin domains are present in all lineages of the tree of life; however, hevein and nictaba lectins are exclusively present in the eu-karyotes (Van Holle and Van Damme, 2019). The jacalin and hevein families show noticeable expansion in the bryophytes. Often, lectin domain-containing proteins, for example, S-Domain Receptor-like Kinases (SRDLKs) are encoded by multi-copy genes that show significant tandem duplications in plant genomes (Myburg et al., 2014; Naithani et al., 2021).

Furthermore, as mentioned earlier, some lectin proteins contain one or more lectin domains. For example, the EUL family proteins exclusively contain one or two lectin domains. Both single and double EUL domain proteins are present in the bryophytes through monocot lineages. The polypodiopsida, gymnosperms, and dicots exclusively harbor single EUL domain architectures (De Schutter et al., 2017; Fouquart et al., 2009; Van Holle and Van Damme, 2019). However, most lectin proteins have a modular structure: they contain lectin domains in combination with one or more other types of domains. For example, many receptor-like kinases (RLKs) comprised of one or more lectin domains combined with other types of domains, such as glycoside hydrolase (GH), F-box, EGF-like, PAN_APPLE, transmembrane, and kinase domains (Myburg et al., 2014; Naithani et al., 2021). Also, many proteins share structural similarities to RLKs, but lack transmembrane and kinase domains (known as receptor-like proteins, RLPs). Still, others have lectin domains with a kinase domain and have attributes of soluble kinases (Naithani et al., 2021). It is noteworthy that different plant lectin subfamilies show distinct evolutionary trails shaped by full and partial kinase domains that correspond to gain/loss/recombination in protein domains. Altogether, there exists a rich repertoire of lectins with varied modular structures (Eggermont et al., 2017; Naithani et al., 2021; Shiu and Bleecker, 2001; Van Holle and Van Damme, 2019).

1.4. Expression of lectin genes and subcellular location of lectin proteins

During the past twenty years, Arabidopsis thaliana has been the model of choice for functional genomics with the aim of elucidating the roles of plant genes. A survey of Arabidopsis genomes suggests the presence of 217 genes that are likely to encode proteins containing one or more lectin domains (Eggermont et al., 2017). We have examined, in detail, the published literature and the data available in various public databases, including Gramene (Tello-Ruiz et al., 2021), EMBL-EBI Gene Expression Atlas (https://www.ebi.ac.uk/gxa/home), UniProt (http://www.uniprot.org), TAIR (http://www.arabidopsis.org) and Gene Expression Omnibus (https://www.ncbi.nlm.nih.gov/geo/) to gather information about the A. thaliana lectin gene family. We have found gene expression data for 215 out of 217 Arabidopsis lectin family genes; however, no evidence was found to support the expression of two genes, AT5G60320 and AT4G199800, under normal growth and developmental conditions. Of the 215 expressed genes, 112 genes show ubiquitous expression in all the organs and tissues of Arabidopsis, although they differ in their relative expression across various tissues and developmental stages. The remaining 103 genes show specific expression in one or more tissues or at specific developmental stages. Many lectins that show low ubiquitous expression under normal growth and developmental conditions show up- or down-regulation in response to biotic or abiotic stimuli (see Table 2).

Interestingly, most of the “classical” seed lectins, such as ricin, abrin, ConA, and SBA, that had been studied in detail in the initial phase of lectin research, were later discovered to be constitutively expressed and located in the vacuoles of plant cells (Lanno and Van Damme, 2010); notably, the transport and accumulation of vacuolar lectins are key to crop growth and development (Nilsson, 2011; Sammartin et al., 2007). Proteome datasets and in vivo visualization of the green fluorescent protein (GFP)-tagged proteins have provided extensive information about the subcellular locations of several Arabidopsis proteins. Therefore, we searched for the subcellular locations of all the 217 Arabidopsis lectin family proteins in the SUBA database (https://suba.live) that serves as the central resource for Arabidopsis protein subcellular location information. We found that 100 out of 217 proteins are expected to reside in the plasma membrane, 54 in the apoplast/extracellular matrix, 41 in the cytosol and the remaining in the various subcellular locations, including plastids, mitochondria, vacuoles, golgi apparatus, nuclei, and microtubules. Moreover, only a small fraction of lectin domain-containing proteins (encoded by AT1G52040, AT3G12500, AT3G16460, AT3G16470, AT3G21630, AT4G21380, and AT5G03700) have the physico-chemical properties required to reside in the vacuole. A summary of gene expression and subcellular location of 60 lectin family members from Arabidopsis (for which biological function is known) is provided in Table 2.

The availability of whole-genome and transcriptome sequences, proteomes, and a range of algorithms now allow us to estimate the subcellular locations of lectins in the plant cell, tissues, and organs. Large-scale analyses of lectins from many crops, including rice, maize, wheat, soybean, and flax, suggest that lectin family genes are expressed in all tissues and show a change in their expression in response to intrinsic developmental signals, as well as in response to biotic and abiotic stress conditions (Huang et al., 2013; De Schutter et al., 2017; Naithani et al., 2021; Petrova et al., 2021). Information on the subcellular location of the lectins, encoded by major crops genomes, is available at the compendium of crop Proteins with Annotated Locations (cropPAL) (https://crop-pal.org). Overall, lectin domain-containing proteins are found in almost all subcellular compartments of plant cells, including cytosol, plasma membrane, extracellular matrix, vacuoles, plastids, and mitochondria.

1.5. Roles of lectins in plant development, cell signaling, and stress response

In the early phase of research, lectins were isolated mostly from storage bodies, tubers, and seeds, as for example, the vacuolar seed lectins (Lanno and Van Damme, 2010). Thus, it was suggested that lectins might function as storage proteins with limited potential for participation in major biochemical and cellular processes. The only other role of lectins known then was for their insecticidal or antimicrobial activity (see Macedo et al., 2015); the first of these reports appeared in 1976 when Irwin Liener and his group (Janzen et al., 1976) showed that black bean lectins were insecticidal against bruchid beetles. Several plant lectins have since then been shown to have insecticidal and antimicrobial properties.

However, in recent decades, lectins have also been found to be localized in different subcellular compartments of various plant tissues, including leaves, stems, stem sap, roots, flowers, and fruits (see Table 2), and thus, suggest their wider roles in plants. Moreover, as discussed earlier, most lectin proteins have multiple domains besides the...
| GeneID        | Tissue-specific mRNA expression                                                                 | Subcellular-location of proteins (experimental evidence is in bold font) | Gene symbol and Function(s)                                      |
|--------------|-------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------|------------------------------------------------------------------|
| AT2G39050    | guard cell; ubiquitous expression                                                              | plasma membrane                                                          | EUONYMUS LECTIN S3 (ARATHUELS3/EULS3) is a nucleocytoplasmic lectin capable of binding carbohydrates. It is involved in ABA-mediated stomatal movement (closure) and provides resistance to *Phytophthora syringae* (Van Hove et al., 2015). |
| AT1G56250    | high expression in guard cell, leaf, and flower, low expression in roots and callus              | plasma membrane                                                          | PHLOEM PROTEIN 2-B1 (PP2-B14) or VIP1-BINDING F-BOX PROTEIN (VIF) is a component of SCF-E3 ubiquitin ligase complexes. It is involved in degradation of VIP1 and Agrobacterium virE2. Thus, it can substitute *Agrobacterium VirF* that is required for induced tumor formation (Saltman et al., 2010). |
| AT1G63090    | ubiquitous high expression                                                                      | cytosol                                                                  | NICTABA                                          |
| AT2G02240    | low-level ubiquitous expression; higher expression in petals and sepals.                       | cytosol                                                                  | MATERNAL EFFECT EMBRYO ARREST 66 (MEE66) is an F-box family protein. It plays a role in embryo development (Pagnussat et al., 2005) and in Pi starvation response. |
| AT4G19810    | ubiquitous expression; high expression in roots and leaves; up-regulated in response to ABA, JA, NaCl, and to fungal and bacterial pathogens, and flagellin. The mRNA is mobile from cell-to-cell. | extracellular, apoplastic fluid                                           | Chitinase C (ChiC) is a Class V chitinase of the glycoside hydrolase family 18. It acts as an exochitinase (predominantly cleaves a chitobiose (GlcNAc)_2 residue from the non-reducing end of a chitin oligosaccharide). It is likely to be involved in pathogen resistance. |
| AT3G12500    |                                                                                                 | extracellular; cytosol; vacuole                                           | BASIC CHITINASE (CHI-B/HCHB)/ PATHOGENESIS-RELATED 3 (PRR-3 or PR3) is involved in ethylene/JA-mediated signaling during systemic acquired resistance-based defense against chitin-containing fungal pathogens. It is also involved in resistance to JA-inducing pathogens such as *Alternaria brassicicola*. Further, it shows in vitro antifungal activity against Tricoderma reseti (Thomma et al., 1998). |
| AT1G52030    | high expression in flowers and siliques; low expression in seeds, leaves, and stems             | Cytosol                                                                  | MYROSINASE-BINDING PROTEIN 2 (MBP2) is involved in metabolizing glucosinolates and forming defense compounds to protect against herbivory. |
| AT2G33070    | high expression in the seeds; low expression in the roots                                       | plasma membrane                                                          | NITRILE SPECIFIER 2 (NSP2) codes for the epithiospecifier protein (ESP) that promotes the formation of simple nitrile, but not epithionitrite or thiocyanate (Burow et al., 2009). |
| AT2G39310    | ubiquitous expression                                                                          | plasma membrane                                                          | NITRILE SPECIFIER 3 (NSP3) promotes simple nitrile, but not epithionitrite or thiocyanate formation. It converts allylglucosinolate and benzylglucosinolate to their corresponding simple nitriles in the presence of myrosinase (Burow et al., 2009). |
| AT2G39330    | ubiquitous high expression                                                                     | plasma membrane; vacuole                                                  | NITRILE SPECIFIER 4 (NSP4) converts allylglucosinolate and benzylglucosinolate to their corresponding simple nitriles in the presence of myrosinase (Burow et al., 2009). |
| AT3G16390    | high expression in roots and callus; low expression in seeds, endosperm, siliques, and the hypocotyls | extracellular; plastid                                                   | NITRILE SPECIFIER 1 (NSP1) is involved in the breakdown of glucosinolates into simple nitrile(s) in the presence of myrosinase (Burow et al., 2009). |
| AT3G16400    | ubiquitous high expression                                                                     | extracellular; golgi; mitochondrion; plasma membrane; plastid; vacuole   | NITRILE SPECIFIER 1 (NSP1) is involved in the breakdown of glucosinolates into simple nitrile(s) in the presence of myrosinase (Burow et al., 2009). |
| AT3G16410    | high expression in roots and callus; low expression in leaves, flowers, embryos, siliques, and the hypocotyls | extracellular; plasma membrane; plastid                                  | NITRILE SPECIFIER 1 (NSP1) is involved in the breakdown of glucosinolates into simple nitrile(s) in the presence of myrosinase (Burow et al., 2009). |
| AT3G16420    | ubiquitous expression                                                                          | extracellular; golgi; mitochondrion; nucleus; plasma membrane; plastid; vacuole | PJ2K1-BINDING PROTEIN 1 (PBP1) is involved in chemical defense against pests and herbivory. It acts like a molecular chaperone that facilitates the correct polymerization of BGLU23/PYK10 (beta-glucosidase complex) when tissues are damaged and subcellular structures are destroyed by pests (Nagano et al., 2005, 2008; Ahn et al., 2010). |
| AT3G16430    | high expression in the roots, cotyledons, callus, and the hypocotyl; low expression in the flower and in the leaves. | extracellular; nucleus; plastid                                          | PJ2K1-BINDING PROTEIN 2 (PBP2) is involved in chemical defense against pests and herbivory. It may facilitate the correct polymerization of BGLU23/PYK10 (continued on next page) |
AT5G01540 expression data not found plasma membrane LecRKA4.1 is a lectin receptor kinase, likely to function redundantly (with LecRKA4.2 and LecRKA4.3) in the plasma membrane.

AT4G02420 ubiquitous expression plasma membrane L-TYPE LECTIN RECEPTOR KINASE IV.4 (LECRK-IV.4) is a lectin receptor kinase, likely to function redundantly (with LecRKA4.2 and LecRKA4.3) in the plasma membrane.

AT3G55550 expression data not found plasma membrane MATERNAL EFFECT EMBRYO ARREST 36 (MEEB36) is required during pollen development (Wan et al., 2018).

Legume-lectins (Concanavalin A-like lectin protein kinase family protein)

| Gene symbol | Function(s) |
|-------------|-------------|
| AT1G15530   | ubiquitous expression plasma membrane L TYPE LECTIN RECEPTOR KINASE S.1 (LECRK-S.1) is involved in resistance response to the pathogenic oomycetes Phytophthora infestans and Phytophthora capsici and to the pathogenic bacterium Pseudomonas syringae (Wang et al., 2014).
| AT2G29250   | ubiquitous expression plasma membrane ALE is a legume lectin that plays a role in defense responses triggered by jasmonic acid (JA), ethylene and chitin (Lyu et al., 2009).
| AT2G43700   | ubiquitous expression plasma membrane L TYPE LECTIN RECEPTOR KINASE I1.2 (LECRK-I.1) is involved in resistance response to the pathogenic fungus A. brassicicola (Wang et al., 2014).
| AT3G09190   | ubiquitous expression extracellular matrix L-TYPE LECTIN RECEPTOR KINASE I.1 (LECRK-I.1) is required during pollen development (Wan et al., 2018). It is also involved in resistance response to the pathogenic fungus P. syringae, P. capsici, A. brassicicola and to the pathogenic bacteria P. syringae, P. capsici, A. brassicicola and to the pathogenic fungi A. brassicicola (Wang et al., 2014).
| AT3G15350   | ubiquitous expression extracellular This is a class GT14 glycosyltransferase protein involved in the breakdown of lignocellulosic biomass into simple sugars/saccharification (Bhattacharjee et al., 2017). It is upregulated in response to ozone treatment (Xu et al., 2015) and after infection with Alternaria brassicicola (Mulchherje et al., 2010).
| AT3G15356   | ubiquitous expression; preferential expression in rosette leaves, roots, inflorescence; induced by chitin, ethylene, and JA extracellular; nuclear matrix; putative extracellular protein involved in embryo development (Jin et al., 2005).
| AT3G45330   | ubiquitous expression plasma membrane L TYPE LECTIN RECEPTOR KINASE I.1 (LECRK-I.1) is involved in resistance response to the pathogenic fungus A. brassicicola (Wang et al., 2014).
| AT3G45390   | ubiquitous expression plasma membrane L TYPE LECTIN RECEPTOR KINASE I.2 (LECRK-I.2) is involved in resistance response to the pathogenic fungus A. brassicicola (Wang et al., 2014).
| AT3G45410   | ubiquitous expression; induced by high salt stress plasma membrane L TYPE LECTIN RECEPTOR KINASE I.3 (LECRK-I.3) is involved in resistance response to the pathogenic fungus A. brassicicola (Wang et al., 2014).
| AT3G53810   | ubiquitous expression plasma membrane L TYPE LECTIN RECEPTOR KINASE IV.2 (LECRK-IV.2) is required during pollen development (Wan et al., 2008). It is also involved in resistance response to the pathogenic bacteria P. syringae (Wang et al., 2014).
| AT3G55550   | expression data not found plasma membrane L TYPE LECTIN RECEPTOR KINASE S.4 (LECRK-S.4) is involved in resistance response to P. infestans, P. capsici and P. syringae (Wang et al., 2014).
| AT3G59700   | ubiquitous expression; upregulated after infection with Fusarium oxysporum (Zhu et al., 2013) extracellular matrix L TYPE LECTIN RECEPTOR KINASE V.5 (LECRK-V.5/LECRK1) confers resistance to P. infestans and P. capsici, but susceptibility to P. syringae (Wang et al., 2014).
| AT3G59740   | expression in flowers, fruits, hypocotyls, callus, and roots plasma membrane L TYPE LECTIN RECEPTOR KINASE V.7 (LECRK-V.7) is involved in resistance response to P. infestans, P. capsici, and P. syringae (Wang et al., 2014).
| AT4G02410   | ubiquitous expression; highly upregulated after abscisic acid (ABA), methyl jasmonate, and salicylic acid (SA) treatments plasma membrane L TYPE LECTIN RECEPTOR KINASE IV.3 (LECRK-IV.3)/LECRK1 contributes to resistance against B. cinerea (Huang et al., 2013).
| AT4G02420   | ubiquitous expression plasma membrane L TYPE LECTIN RECEPTOR KINASE IV.4 (LECRK-IV.4) is involved in resistance response to P. infestans, P. capsici, and P. syringae (Wang et al., 2014).
| AT5G01540   | expression data not found plasma membrane LecRK1A4.1 is a leucin receptor kinase, likely to function redundantly (with LecRK1A4.2 and LecRK1A4.3) in the negative regulation of ABA response during seed germination (Xin et al., 2009).
| AT5G01550   | ubiquitous expression plasma membrane LecRK1A4.2 is a receptor kinase. Together with other homologs (LecRK1A1 A15g01540; LecRK1A4.3 A15g01560), it functions in a redundant manner in the negative regulation of ABA response in seed germination (Xin et al., 2009).
| AT5G01560   | ubiquitous expression plasma membrane LecRK1A4.3 is a leucin receptor kinase. Together with its homologs (LecRK1A1 and LecRK1A4.2), it functions in a
Table 2 (continued)

| GeneID     | Tissue-specific mRNA expression | Subcellular-location of proteins (experimental evidence is in bold font) | Gene symbol and Function(s) |
|------------|---------------------------------|------------------------------------------------------------------------|------------------------------|
| AT5G03140  | ubiquitous expression           | extracellular matrix                                                   | L-TYPE LECTIN RECEPTOR KINASE VIII.2 (LECRK-VIII.2) is involved in resistance response to the pathogenic oomycetes P. infestans and P. capsici (Wang et al., 2014). |
| AT5G03350  | ubiquitous expression           | extracellular; plastid                                                  | SA-INDUCED LEGUME LECTIN-LIKE PROTEIN 1 (SAI-LLP1) plays a positive role in the SA-mediated effector-triggered immunity, but it is not involved in the autophagy process. Further, it promotes systemic acquired resistance rather than local immunity against Phytophthora infestans, Phytophthora capsici, and to the pathogenic bacterium Pseudomonas syringae (Wang et al., 2014). |
| AT5G10530  | expression data not found        | plasma membrane                                                         |                             |
| AT5G42120  | flowers, (carpels, stamens), seeds (endosperm, embryos), callus, roots, hypocotyls | plasma membrane                                                         |                             |
| AT5G55830  | ubiquitous expression           | plasma membrane                                                         |                             |
| AT5G60270  | ubiquitous expression           | plasma membrane                                                         |                             |
| AT5G60280  | ubiquitous expression           | extracellular matrix                                                    |                             |
| AT5G60300  | ubiquitous expression           | plasma membrane                                                         |                             |
| AT5G65600  | ubiquitous expression           | plasma membrane                                                         |                             |
| Lysin motif (LysM) |                                        | extracellular; plasma membrane                                          | LYSM DOMAIN GPI-ANCHORED PROTEIN 1 (LYM1)/LYSM-CONTAINING RECEPTOR PROTEIN 2 (LYP2). This is an ortholog of OsLYP4 and OsLYP6 and contains a C-terminal GPI anchor signal. It serves as a cell surface receptor for peptide-containing elicitor signaling leading to innate immunity (Willmann et al., 2011). Induction of chitin-responsive genes by chitin treatment is not blocked in the mutant. |
| AT1G21880  | callus, carpels, cotyledons, cultured plant cells, flowers, flower pedicels, hypocotyls, inflorescence meristem, leaves, leaf apex, leaf lamina base, petals, petioles, embryo, pollens, roots, seeds, sepals, shoots, and stamens | plasma membrane                                                         |                             |
| AT1G51940  | ubiquitous high expression      | plasma membrane                                                         | LYSM CONTAINING-RECEPTOR LIKE KINASE (LYK3) may recognize microbe-derived N-NAG-containing ligands, but it is not a major contributor to chitin signaling (Shinya et al., 2012). |
| AT1G55000  | ubiquitous high expression      | plasma membrane                                                         | LysM domain-containing protein; component of SCF (ASK-cullin-F-box) E3 ubiquitin ligase complexes. It is involved in protein ubiquitination. |
| AT1G77630  | ubiquitous high expression      | plasma membrane                                                         |                                 |
| AT2G17120  | ubiquitous high expression      | plasma membrane                                                         |                                 |
Table 2 (continued)

| GeneID | Tissue-specific mRNA expression | Subcellular-location of proteins (experimental evidence is in bold font) | Gene symbol and Function(s) |
|--------|--------------------------------|-----------------------------------------------------------------------|----------------------------|
| AT2G23770 | ubiquitous low expression, high expression in roots, cotyledons, and leaves. | plasma membrane | chitin-responsive genes by chitin treatment is not blocked. |
| AT2G33580 | ubiquitous high expression | plasma membrane | LYSM-CONTAINING RECEPTOR-LIKE KINASE (LYK4) is predicted to recognize microbe-derived N-NAG-containing ligands and to play a role in chitin elicitor signaling and/or in detecting PGNs during bacterial growth and leading to innate immunity. It is involved in resistance to the pathogen fungus A. brassicicola and the bacterial pathogen P. syringae pv tomato DC3000 ([Wan et al., 2012]). |
| AT3G01840 | ubiquitous expression | plasma membrane | LYSM-CONTAINING RECEPTOR-LIKE KINASE 5 (LYK5) is predicted to recognize microbe-derived N-acetylglucosamine (NAG)-containing ligands. |
| AT3G21630 | ubiquitous high expression | plasma membrane; vacuole | CHITIN ELICITOR RECEPTOR KINASE 1 (CERK1) receptor kinase recognizes microbe-derived N-NAG-containing ligands (e.g. chitin from the fungal cell wall and PGNs from the bacterial cell wall), and transduces signaling leading to innate immunity ([Shinya et al., 2014]). |
| *Galanthus nivalis* agglutinin (GNA) | |
| AT1G11300 | ubiquitous expression; it is activated by mannitol produced by pathogens such as fungi and contributes to plant defense response whenever mannitol is present ([Trontin et al., 2014]). | plasma membrane | ENHANCED SHOOT GROWTH UNDER MANNITOL STRESS 1 (EGM1) is likely to contribute to plant defense response against mannitol producing pathogens ([Trontin et al., 2014]). |
| AT1G11305 | ubiquitous expression; upregulated by mannitol produced by pathogens such as fungi | plasma membrane; extracellular matrix | ENHANCED SHOOT GROWTH UNDER MANNITOL STRESS 2 (EGM2) is likely to contribute to plant defense response whenever mannitol is present ([Trontin et al., 2014]). |
| AT1G11350 | leaves, cotyledons, flowers, hypocotyls, inflorescence meristems, petals, petioles, embryo, seeds, sepals, shoots, stamen, and stem; guard cells | plasma membrane | Calmodulin-binding receptor-like protein kinase (CBRLK1) functions as a negative regulator in plant defense responses. It binds calmodulin and then gets activated. CBRLK1 mutant and CBRLK1-overexpressing transgenic plants show enhanced and reduced resistance against a virulent bacterial pathogen, respectively and this is correlated with increased and reduced induction of the pathogenesis-related gene PRI ([Kim et al., 2009]). Lectin S-domain receptor kinase 29 (SD1-29) recognizes lipopolysaccharide (LPS) present in the cell walls of *Pseudomonas* and *Xanthomonas* species, and triggers innate immunity. Loss-of-function mutants are hypersusceptible to *P. syringae* ([Barb et al., 2015]). |
| AT1G61380 | ubiquitous expression | plasma membrane | |
| AT1G65790 | primarily expressed in the leaves, and roots; low expression in stems and flower buds; induced by wounding and *Xanthomonas campestris pv. campestris*. | microtubules: GFP-tagging has shown that ARK1 is localized in the growing ends of microtubules in the cortex of the subapical region of growing root hairs. | ARMADILLO-REPEAT KINESIN1 (ARK1) and its homologs ARK2 and ARK3 are plant-specific kinesin microtubule motor proteins involved in the cellular expansion, differentiation, and in determining the growth and polarity of the roots. It may be involved in the ABA-mediated signaling during germination ([Tobias and Nasrallah, 1996; Samuel et al., 2008; Sun et al., 2020]). |
| AT3G16030 | ubiquitous expression | plasma membrane | CALLUS EXPRESSION OF RBCS 101 (CES101)/ RESISTANCE TO *FUSARIUM OXYSPORUM* 3 (RFO3) is a receptor-like kinase. It promotes the expression of genes involved in photosynthesis, at least in dedifferentiated calli ([Niwa et al., 2006]). Also, it may play a role in innate immunity. |

carbohydrate-binding lectin domain(s), and accordingly they have diverse biological activities. We note here that lectin domains also interact with entities other than the sugars. For example, lectin domain-containing S-locus receptor kinase from *Brassica*, located in the plasma membrane of stigma epidermis cells, binds to a small S-locus cysteine-rich (SCR) peptide present on the pollen surface and their productive interaction activates the self-incompatibility response ([Kachroo et al., 2001; Naithani et al., 2007, 2013]). Other RLKs interact with RLPs, plant growth hormones, and cell wall components (lipopolysaccharides or chitin) of the microbes, and they may play roles in diverse processes. More specifically, lectin domain-containing proteins are involved in the negative regulation of abscisic acid (ABA) response during seed germination ([Eggermont et al., 2017; Xin et al., 2009]), cell wall development and expansion ([Petrova et al., 2021]), ABA-induced stomatal movement ([Van Hove et al., 2015]), pollen development and male sterility ([Peng et al., 2026; Wan et al., 2008]), pistil development and self-incompatibility response ([Kachroo et al., 2001; Naithani et al., 2007; Rüdiger and Gabius, 2001; Tantikanjana et al., 2009]). See Table 2 for different functions of Arabidopsis lectins. In rice, lectin family members have been shown to be involved in leaf senescence, early crown root development, shaping panicle architecture, determining yield, abiotic stress tolerance, and immunity ([Chen et al., 2006, 2013; De Schutter et al., 2017; Fan et al., 2018; Pan et al., 2020; Zou et al., 2015]). In addition, lectin proteins present in the roots of the legumes...
interact with the lipo-chitooligosaccharides in Nod factors produced by rhizobia, playing critical roles in establishing symbiotic interactions with rhizobia (Cullimore et al., 2001; De Hoff et al., 2009).

Notably, constitutively expressed lectins present in the vegetative tissues of plants have been shown to be involved in mRNA transport across long distances within the plant (Pallas and Gómez, 2013). Others, such as the Gal-specific thylakoid lectin, from cabbage leaves, have been reported to provide cryoprotection against freeze-thaw damage (Hincha et al., 1993; Sieg et al., 1996). A stromal thylakoid membrane lectin of Tríncale was shown to be associated with the Photosystem-I chlorophyll molecules (Aleksidze et al., 2002). A similar lectin was reported from the unicellular alga Dunaliella salina, and was shown to be associated with its light-harvesting complex. Further, stromal thylakoid membrane lectins provide ligands to ribulose bisphosphate carboxylase oxygenase (Rubisco) (Kovalchuk et al., 2012).

Overall, evidence from Arabidopsis and several crops suggest that lectin domain-containing proteins play important roles in mechanisms associated with plant development, reproduction, abiotic tolerance, and pathogen resistance (De Hoff et al., 2009; Rüdiger and Gabius, 2001; Van Helle and Van Den Ende, 2019). Here, we have taken a biocuration approach to obtain information about expression of lectin gene family, the subcellular distribution of lectins, and their function(s).

1.6. The endogenous plant lectin cycle

It is well-known that atmospheric CO$_2$, as well as the CO$_2$ present within the surrounding plant tissues, reach the chloroplastic stroma, the site of the Calvin-Benson-Bassham cycle (Sharkey, 2019), where it is converted into glucose (Nonomura et al., 2017). A large fraction of the photosynthate is then transported into vacuoles, major multifunctional organelles of the plant cell, where it may be modulated for later use, for example, in abiotic stress response (Marty, 1999; Pertl-Obremeyer et al., 2016). However, as soluble carbohydrates accumulate, osmoregulation becomes important. Hence, one process for the storage of these sugars in plants is to convert them to polysaccharides, such as starch in the chloroplasts (Shevela et al., 2018). When the plant requires sugars for energy, starch is enzymatically degraded and utilized as needed. Although stored polysaccharides may be used to meet the long-term energy needs of a plant, yet significant energy is invested in the enzymatic reactions that involve either chain lengthening (polysaccharide biosynthesis) or the cleavage of a monosaccharide (Nonomura et al., 2018a). An alternative strategy to deal with the problem of accumulation of soluble carbohydrates is to reversibly bind sugars to proteins and render them osmotically inert. Lectins are the macromolecules of choice for this function in plant cells (Nonomura et al., 2020). The key feature of this process is that the lectin-bound sugars are readily available for use, as they are reversibly bound. Whether the carbohydrate would remain bound or released from the protein will depend on the direction in which the equilibrium shifts. Protein-carbohydrate interactions can be altered in many ways in the laboratory, such as by varying temperature, pH, or the concentration of the interacting molecules. In the case of endogenous lectin-carbohydrate interactions within the plant, a high concentration of free carbohydrates formed during photosynthesis may shift the equilibrium toward glycosylation of lectins. On the other hand, as the concentration of the carbohydrate depletes—for example, as a result of consumption of sugars in respiration—bound sugars would be displaced. The sugars thus freed from the lectin-carbohydrate complex are osmotically active.

Increasing concentrations of competing sugars that are not catabolized in plants and, in particular, those having comparable affinities for the lectins would promote the release of the bound sugar by chemically outcompeting the bound carbohydrate (Nonomura et al., 2020). Thus, the endogenous plant lectin cycle depends on such a mode of binding and release of glucose (Glc, or other active carbohydrate metabolites) to maintain a reasonably steady level of sugar for immediate energy needs (Nonomura et al., 2020). This catch and release mechanism of sugars from lectins, by a process that is strictly a matter of chemical competition and involves reversible binding, is cost-effective for plants (Dam et al., 2000).

1.7. Application of induced glycoregulation (lectin bypass) for improving crop productivity

One of the requirements for the endogenous plant lectin cycle to work is that the competing sugar must have a similar range of affinity for the lectin as Glc itself. This is because, in the presence of a high-affinity competing carbohydrate, even when the concentration of Glc builds, lectin may be unable to compete for the carbohydrate-binding site. Thus, while a fraction of the carbohydrate metabolites might then get stored for later use, a significant amount would be channeled toward the growth and productivity of the plant (Nonomura et al., 2020). Exogenous α-D-mannopyranosides appear to do just that, i.e., they bind tightly to the lectins in the treated crops and, thereby, suppress the endogenous lectin-carbohydrate interaction. This bypass of the plant lectin cycle leads to increased levels of soluble sugars that are then transported to all parts of the plant (Nonomura et al., 2018a). The surplus of photosynthates (sugars) in plants, treated with exogenous α-D-mannopyranosides, may then be directed to other metabolic reactions, some of which can contribute towards fruit flavor, stress resistance, and growth of the plant (Nonomura et al., 2020).

Support for the binding of plant lectins to α-D-mannopyranosides came from a series of experiments involving labeled carbon atoms (Gout et al., 2000; Biel et al., 2010) by which it was determined that the labeled carbon molecule was not catabolized, and that a fraction was protein-bound. Suspecting that a sugar-binding protein might play some role in this process, the effects of methyl-β-D-glucoside (MejGlc), and its α-anomer MeαGlc, were examined in plant of Brassicaceae family (Nonomura and Benson, 2013), showing that not only MeGlc, but also α-D-trimannopyranoside, are better inducers of seed growth than MejGlc. Also, adding MejGlc to solutions of MeαGlc reduces the potency of the latter (Benson et al., 2009). A key finding was that the enhancement of growth by these carbohydrates required the presence of Ca$^{2+}$ and Mn$^{2+}$ (Nonomura and Benson, 2014) and this is a typical property of many lectins, particularly those from the legume lectin family, such as ConA. Indeed, Man-binding lectins are widely distributed in plants (Davidson and Stewart, 2004) and the observed growth enhancement by α-D-mannopyranosides was also found to occur in many other C$_4$ plants, as well as in CAM and C$_3$ plants (Nonomura and Benson, 2013). A consideration of these findings led to the hypothesis that Man-binding lectins may be involved in growth enhancement in response to α-D-mannopyranosides (Nonomura and Benson, 2014). As many classical seed lectins are localized in the vacuoles (that also serve as major sites for carbohydrate storage within the plant cell), the modulation of glycoregulation was proposed to be mediated in vacuoles, working in conjunction with the cell walls (Nonomura et al., 2020); however, the specific lectins involved in this bypass remain to be identified.

We note that treatment with indoleacetic acid (IAA) and kinetin had been earlier shown to enhance the beneficial effects of MejGlc (Benson et al., 2009). Moreover, promotion of growth has been observed after the application of relatively high (mM) doses of sugar-conjugated plant growth regulators (SPGRs) (Nonomura et al., 2011), and these findings are consistent with the occurrence of high concentrations of SPGRs in plants and with the enzymes known to be involved in the sugar-linking process (Jakubowska and Kowalczyk, 2004). The above-described experiments have provided clues to the involvement of signaling mechanisms. Future experiments involving α-D-mannopyranosides, other binding substrates, and signaling components in experiments with lectin mutant strains of plants are expected to provide new information on glycoregulation and its application in agriculture.
2. Concluding remarks

Lectins, the readers of the sugar codes in living organisms, are highly diverse. They are distributed across all kingdoms of life. Genome sequences of photosynthetic autotrophs have revealed that lectin proteins generally have a modular structure, often consisting of one or more lectin domains in combination with other conserved protein domains. The structural diversity of lectins is now the basis for their most widely used classification system. In addition to binding sugar molecules, lectin domain-containing proteins recognize other specific ligands (e.g., small peptides, carbohydrates, lipids, small molecules, and ATP) that are generated within the plants during various developmental stages and differentiation of organs, tissues, and cell or produced by other biotic or abiotic components present in the external environment. In fact, hundreds of lectin domain-containing proteins may be present within a single species, which differ in their carbohydrate-binding specificities and binding affinities (suited to their function at a given cellular location or tissue).

In more recent decades, analyses of whole genomes and transcriptomes of hundreds of plants, algae, and photosynthetic bacteria have revealed that lectin proteins are ubiquitous and show clade-specific evolutionary expansions. Besides their ability to bind sugars reversibly, lectin domain-containing proteins can form dimers and interact with diverse signaling molecules. We now have much information about the varied expression profiles of lectin family members across various tissues, developmental stages, and in response to pathogens and abiotic stresses—suggesting their wider role in plant function and plant survival. Indeed, experimental studies involving mutants and transgenic plants, expressing lectin family proteins, have shown that they can impart pathogen resistance, tolerance to drought, salinity, heat, and submergence. However, much remains to be studied about their exact roles, their ligands, and other interactors. Thus, further manipulation of candidate genes of the lectin family associated with important agronomic traits are likely to pave the road for breeding climate-resilient crops for the future.

The highly specific binding of lectins with sugars provides an efficient mechanism for carbohydrate storage and its slow release as needed in different tissues (particularly in storage organs) while maintaining the osmotic balance within the cell and its various subcellular compartments. Thus, lectins also participate in osmoregulation/glycoregulation via the endogenous lectin cycle, which is important for maintaining cellular homeostasis. We note that much remains to be understood about the plant lectin cycle and its bypass. For instance, we do not yet know which and how many lectins participate in this process or whether a single carbohydrate metabolite, a group of metabolites, SPGRs or PGRs are reversibly bound to the lectins. Some information on the carbohydrate specificities and affinities are available, for example, α-D-mannopyranosides are preferred to the β-anomers, and trimannosides are preferred to the monomers, yet much remains to be elucidated regarding the nature of lectin–carbohydrate interactions. Thus, future investigations must involve studying the effect of other carbohydrate ligands and their conjugates with plant growth regulators, such as IBA and kinetin. In the context of agricultural productivity, it remains to be determined whether breeding for the expression of specific lectins is feasible. Clearly, sugars from the aforementioned bypass go to growth points and flowering and fruits. Hence, targeting lectins to manage stresses—such as developmental stages, and in response to pathogens and abiotic stresses—are preferred to the monomers, yet much remains to be elucidated regarding the nature of lectin–carbohydrate interactions. Thus, future investigations must involve studying the effect of other carbohydrate ligands and their conjugates with plant growth regulators, such as IBA and kinetin. In the context of agricultural productivity, it remains to be determined whether breeding for the expression of specific lectins is feasible.

Acknowledgments

SSK acknowledges the support to her lab from an umbrella grant of UGC-CAS(DRS), India to the School of Life Sciences, JNU, New Delhi, India. SN acknowledges partial support from the U.S. National Science Foundation award (1127712) and the United States Department of Agriculture, Agriculture Research Service (USDA-ARS). We thank BRANDT iHammer, Powell, OH, USA, for providing the open access publication fee of this review. The funders had no role in design, data collection, analysis, decision to publish, or manuscript preparation. GG thanks the staff of the Office of Information Technology, Life Sciences, at the University of Illinois at Urbana-Champaign, for constant help.

References

Abhinav, K.V., Vijayan, M., 2014. Structural diversity and ligand specificity of lectins. The Bangalore effort. Pure Appl. Chem. 86, 1335–1355. https://doi.org/10.1515/pac-2014-0607.
Agostino, M., Velkov, T., Dingjan, T., Williams, S.J., Yuriev, E., Ramsland, P.A., 2015. The carbohydrate-binding promiscuity of Eucymnon europaeeus lectin is predicted to involve a single binding site. Glycobiology 25, 101–114. https://doi.org/10.1093/glycob/wcz095.
Agostino, M., Yuriev, E., Ramsland, P.A., 2011. A computational approach for exploring carbohydrate recognition by lectins in innate immunity. Front. Immunol. 2 https://doi.org/10.3389/fimmu.2011.00282.
Agrawal, B.B., Goldenstein, L.J., 1965. Specific binding of concanavalin A to cross-linked dextran gels. Biochem. J. 96. https://doi.org/10.1042/bj0960023c, 23octd-Sc.
Ahn, Y.O., Shimizu, B., Sakata, K., Gantulga, D., Zhou, Z., Zhou, Z., Bevan, D.R., Esen, A., 2010. Scopolin-hydrorlyzing beta-glucosidases in roots of Arabidopsis. Plant Cell Physiol. 51, 132–143. https://doi.org/10.1093/pcp/pcp174.
Aleksidez, G.Ya, Litvinov, A.I., Vyskrebentseva, E.I., 2002. The model of Calvin cycle enzyme organization on thylakoid membranes with the involvement of the Photosystem I lectin. Russ. J. Plant Physiol. 49, 137–141. https://doi.org/10.1134/1.16730732193.
Alroy, J., Goyal, V., Skutesky, E., 1987. Lectin histochemistry of mammalian endothelium. Histochemistry 86, 603–607. https://doi.org/10.1007/BF00489554.
Andr, S., Kalmer, H., Manning, J.C., Murphy, P.V., Gabius, H.-J., 2015. Lectin: getting familiar with translators of the sugar code. Molecules 20, 1788–1823. https://doi.org/10.3390/molecules20062178.
Arabidopsis Genomics Initiative, 2000. Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. Nature 408, 796–815. https://doi.org/10.1038/35046892.
Armijo, G., Salinas, P., Monsteilova, M.I., Seguel, A., Garcia, C., Villarroel-Candia, E., Song, W., van der Krol, A.R., Alvarez, M.E., Holuigue, L., 2013. A salicylic acid-induced lectin-like protein plays a positive role in the effector triggered immunity response of Arabidopsis thaliana to Pseudomonas syringae Avr-Rpm1. Mol. Plant-Microbe Interact. 26, 1395–1406. https://doi.org/10.1094/MPMI-02-13-0044-R.
Ashwell, G., Morell, A.G., 1974. The role of surface carbohydrates in the hepatic recognition and transport of circulating glycoproteins. Adv. Enzymol. Relat. Areas Mol. Biol. 41, 99–128. https://doi.org/10.1002/9780470122860.ch3.
Aub, J.C., Tieluslan, C., Lankester, A., 1963. Reactions of normal and tumor cell surfaces to enzymes, I. Wheat-germ lipase and associated mucopolysaccharides. Proc. Natl. Acad. Sci. U. S. A. 50, 613–619. https://doi.org/10.1073/pnas.50.4.613.
Barondes, S.H., 1988. Bifunctional properties of lectins: lectins redefined. Trends Biochem. Sci. 13, 480–482. https://doi.org/10.1016/0040-4203(88)90225-6.
Benson, A.A., Nomura, A.M., Gerard, V.A., 2009. The path of carbon in photosynthesis. XXIV. Plant and algal growth responses to glycopyranosides. J. Plant Nutr. 32, 1185–1200. https://doi.org/10.1080/01904160902943205.
Biel, K.Y., Nomura, A.M., Benson, A.A., Nishio, J.N., 2010. The path of carbon in photosynthesis. XXVI. Uptake and transport of methylglycoside glycoside in plants. J. Plant Nutr. 33, 902–913. https://doi.org/10.1080/01904161000566897.
