



Published in final edited form as:

Proteomics Clin Appl. 2012 August ; 6(0): 346–350. doi:10.1002/prca.201200014.

Using lectins in biomarker research: addressing the limitations of sensitivity and availability

Brian B. Haab

Van Andel Research Institute, Grand Rapids, MI

Summary

Carbohydrates have fundamental roles throughout biology, yet they have not been as well studied as proteins and nucleic acids, in part due to limitations in the experimental tools. Improved methods for studying glycans could spur significant advances in the understanding and application of glycobiology. The use of affinity reagents such as lectins and glycan-binding antibodies is a valuable complement to methods involving mass spectrometry and chromatography. This article addresses two limitations that have prevented the broader experimental use of glycan-binding proteins: sensitivity and availability. The sensitivity limitation stems from the poor affinity that many glycan-binding proteins have as isolated analytical reagents. To address this problem, we propose making use of multivalent interactions between lectins and glycans, mimicking those frequently found in the biological setting. We have shown that a practical technique for producing lectin multimers can significantly improve detection sensitivity. The second limitation, availability, is the difficulty of finding and obtaining glycan-binding proteins that recognize less-common or arbitrarily defined glycan structures. To address this problem, we propose translating the wealth of existing glycan array data into a quantitative, searchable database of the specificities of glycan-binding proteins. Such a resource would allow us to more easily identify proteins with defined specificities and perform detailed comparisons between reagents. Solutions to these two limitations could lead to the more effective use of, and a broader range of, glycan-binding reagents.

Carbohydrates (glycans) are found throughout every cell of every organism and on most secreted and membrane-bound proteins and lipids. Glycans have been implicated in the pathology of a wide variety of diseases, including infectious disease, cancer, autoimmune disease, and certain congenital disorders [1]. Understanding the structures and functions of members of this class of biomolecule therefore has significant utility. Examples of medical applications include cancer vaccines based on immunogenicity to cancer-associated glycans [2] and biomarkers based on the detection of altered glycans secreted by cancer cells [3]. Despite these considerations, glycans have been studied much less than proteins and nucleic acids. Part of the reason for the lower research effort is the relative difficulty in studying glycans. Recombinant methods for producing large quantities of specific glycans are not available, in contrast to proteins and nucleic acids, and determining the primary sequence of a purified glycan also is much more difficult than for proteins and nucleic acids. Considering the major role of glycans in biology, the development of improved tools for the study of glycans is an important goal.

Glycan structures can be identified using a variety of methods based on mass spectrometry, enzymatic digestion, and chromatography. Mass spectrometry (MS) methods have particularly advanced in recent years to achieve more routine, reliable, and comprehensive

composition analysis [4]. Advances in both the technology and the automated analysis of spectra have made glycan analyses accessible to a greater range of researchers. These methods are fundamentally important, but additional, complementary approaches are needed. In particular, we need methods that provide precise measurements of specific structures over many different samples, particularly for biomarker research. One valuable approach for measuring specific glycans in a format compatible with clinical or biomarker research is to use affinity reagents such as lectins or glycan-binding antibodies.

Affinity reagents for glycan detection

Lectins and glycan-binding antibodies, collectively known as glycan-binding proteins, can be used in a wide variety of analytical formats [5] such as histochemistry, the probing of electrophoretic gels, affinity chromatography, solid-phase ELISA-type assays, and microarray assays (Fig. 1). The information gathered from affinity reagents is highly complementary to that from MS-based approaches. Glycan-binding proteins can provide reproducible measurements on specific structures over many samples, whereas MS typically yields information on many structures in fewer samples, with less precision. Therefore, glycan-binding proteins miss the detailed information that MS gives but instead provide precise information on changes across samples.

Another advantage of affinity reagents is that assays can be designed to detect target glycans on specific protein carriers. For example, an immobilized antibody can capture a protein of interest, and the glycans on that protein may be probed using a variety of lectins (Fig. 1). This information is useful because the glycosylation state of a protein may be critically important to its function or may be associated with a particular condition. The glycosylation state of many different proteins can be probed in parallel using antibody-lectin sandwich arrays (ALSA) [6]. By spotting many different antibodies in a small array, a single incubation of a lectin can be used to examine the glycans on many different proteins. We have previously shown the value of this method for biomarker research: for example, the abundances of certain proteins may not change much between healthy and diseased populations, yet their glycosylation state does [6, 7]. Thus, measuring the glycans on specific proteins provides improved biomarker performance. We currently are applying this method to problems in the diagnosis and management of pancreatic cancer and other cancers in which glycan alterations have been observed. This approach could be valuable in many situations, given the prevalence of alterations in glycosylation state.

Although lectins and glycan-binding antibodies have experimental advantages for measuring glycans, they have not found broad application in biomarker research. Two limitations in their use are insufficient binding affinity, leading to poor sensitivity in analytical assays, and the lack of availability of lectins specific for less-studied glycan structures. Potential solutions to each limitation are given below.

Sensitivity

Some glycan-binding proteins have a weak affinity for their target glycan, at least when used as a reagent outside the biological context. Weak affinity leads to poor sensitivity in an analytical assay; this is a serious limitation, because rarely is any glycan found at a high abundance in a biological sample. Thus, lectins and antibodies have not had broad use in biomarker studies.

In biology, lectins frequently rely upon multivalent interactions to affect their functions. For example, the glycan ligands of a lectin may be tethered to a cell surface and clustered together in multiple interactions when bound by lectins. The galectin family of lectins uses a well-described system of forming lattices in order to induce downstream effects [8]. In such

systems, the strength of any one protein-glycan interaction is not great, but linking several such interactions together increases the overall interaction strength.

Given the biological interactions of certain lectins, it follows that the *in vitro* binding strength of such lectins to their target glycans may be improved by inducing multivalent interactions. Such an improvement might be possible if multiple target glycans are present on a particular protein (or on closely packed proteins) and if the lectins used as detection probes can be linked together. This strategy was used to improve the *in vitro* binding of the influenza hemagglutinin (HA) protein to its sialic acid ligand in a glycan array assay [9]. The HA proteins were recombinantly produced with a His-6 tag, and the spontaneously formed trimers of the His-labeled HA proteins were linked together with a goat antibody against the His tag. Two of the antibody-bound clusters were further joined through an anti-goat IgG antibody to form a final cluster of 12 HA proteins.

The same type of strategy could be pursued using any type of lectin. One way to cluster lectins would use the multiple binding sites of streptavidin, which has four subunits that each contain a binding site for biotin (Fig. 2). Therefore, by mixing biotinylated lectins with streptavidin, clusters of one to four lectins should be formed. To induce further multimerization, anti-streptavidin could be added, which would link two streptavidin molecules to form clusters of two to eight lectins. We recently demonstrated that this strategy indeed can dramatically increase signal strength when used in the ALSA assay. For example, when using the aleuria aurantia lectin (AAL) to probe antibody arrays, signals increased 2- to 17-fold, depending on the antibody that was probed (unpublished data).

We propose that a multimerization strategy will enhance the effectiveness of many lectins in many types of assays. The level of improvement would depend on the organization of the target glycans and the structural ability of the lectin to properly bind adjacent glycans. Therefore, no protocol will be optimal for every lectin; rather, each lectin will require individual testing and optimization using a variety of multimer and non-multimer protocols. But by individual optimization of each lectin, we may greatly enhance the effectiveness of a wide range of lectins that previously were not useful as analytical reagents.

Availability of lectins with uncommon specificities

Many lectins are commercially available, yet the range of structures covered by those lectins is relatively limited. The commonly used lectins generally target major biological glycan motifs such as sialic acid, fucose, blood-group glycans, and lactosamine. These structures tend to be immunogenic in foreign organisms and likely are particularly amenable to high-affinity binding by proteins. However, there are many other glycan structures that are important in protein-glycan recognition and other biological functions. For example, an unusual sialylated structure, the type 1 H antigen sialylated at the 6' carbon, was recently found in the pancreas of some individuals with pancreatic cancer [10]. Such structures likely have physiological significance and would be valuable to detect and measure in affinity-based assays. The challenge is to find suitable detection reagents for such unusual motifs.

It may be possible to fulfill this need by making use of the enormous diversity of lectins that exist in biology. Lectins are present in every cell of every organism, and lectins with newly observed specificities are being discovered on a regular basis. A wealth of valuable analytical reagents likely exists in this class of proteins, with many yet to be found and others already known but not yet employed for *in vitro* assays. However, identifying a lectin with an unusual specificity, or even obtaining a clear characterization of a lectin's specificity, can be difficult tasks. This information is often in obscure journals, and characterizations of specificities are often minimal. The first step in making a greater

number of lectins available for research will be to assemble detailed information about lectin specificities in a searchable and well-annotated format.

A breakthrough technology in characterizing lectin specificity is the glycan microarray [11], which enables the parallel comparison of lectin binding to hundreds of different glycans and from which we can derive much information about specificity. Before glycan microarrays, serial testing of lectin binding to each glycan was necessary, requiring so much time and material that the testing of large numbers of glycans was impractical. High-quality glycan microarray analyses are provided by the Consortium for Functional Glycomics as a service to participating investigators [12]. Investigators can send a protein of interest for analysis on the microarray, and the Consortium provides back a report of the binding to each of the glycans. Since the introduction of the CFG glycan array in 2004, several thousand experiments covering hundreds of different glycan binding proteins have been performed. These data are publicly available via a web portal. Therefore, we already have much quantitative information about the binding specificities of hundreds of glycan-binding proteins.

The availability of these data sets presents a tremendous opportunity to provide researchers with a tool for finding lectins with defined specificities. Right now, the glycan array data are largely unprocessed, leaving it to the user to figure out from the relative binding to each glycan the precise specificity of a lectin. Such analysis is time-consuming and imprecise if done manually. To enable a more rapid and objective determination of lectin specificity from glycan array data, we recently developed the Motif Segregation analysis method [13]. We have shown that this automated method of analyzing glycan array data correctly extracts the main binding specificities. An additional method, called Outlier Motif Analysis, builds on Motif Segregation to enable more-detailed identifications of fine specificities [14]. The output of this program is a list of component substructure motifs of the glycans and an associated score indicating the preference of the lectin for binding that motif. This relative quantitation of motifs enables an objective determination of binding specificity. Furthermore, the compilation of analyses from many different data sets enables comparisons and groupings of specificities between lectins (Fig. 3), as well as searches for lectins with pre-defined specificities. Software for the full automation of glycan array data recently has been developed by researchers at the Palo Alto Research Center (PARC). This new software is a key step in assembling the quantitative information necessary in developing a comprehensive resource for lectin specificities.

These two developments—*in vitro* multimerization and a resource for quantitative analyses of lectin specificities—could enhance the performance and broaden the analytical use of many lectins. The ability to obtain more sensitive measurements for a greater range of glycan structures, including unusual structures that have been little studied, would lead to increased knowledge about glycobiology and an improved application of that knowledge. For example, sensitive analyses of levels of a glycan on specific proteins in various biological contexts could be used to understand the regulation of that structure, its contribution to biological processes, and its potential use as a marker of disease states.

References

1. Fuster MM, Esko JD. The sweet and sour of cancer: glycans as novel therapeutic targets. *Nat Rev Cancer*. 2005; 5:526–542. [PubMed: 16069816]
2. Astronomo RD, Burton DR. Carbohydrate vaccines: developing sweet solutions to sticky situations? *Nat Rev Drug Discov*. 2010; 9:308–324. [PubMed: 20357803]
3. Adamczyk B, Tharmalingam T, Rudd PM. Glycans as cancer biomarkers. *Biochimica et biophysica acta*. 2011

4. North SJ, Hitchen PG, Haslam SM, Dell A. Mass spectrometry in the analysis of N-linked and O-linked glycans. *Current opinion in structural biology*. 2009; 19:498–506. [PubMed: 19577919]
5. Rudiger H, Gabius HJ. Plant lectins: occurrence, biochemistry, functions and applications. *Glycoconjugate journal*. 2001; 18:589–613. [PubMed: 12376725]
6. Chen S, LaRoche T, Hamelinck D, Bergsma D, et al. Multiplexed analysis of glycan variation on native proteins captured by antibody microarrays. *Nature methods*. 2007; 4:437–444. [PubMed: 17417647]
7. Yue T, Goldstein IJ, Hollingsworth MA, Kaul K, et al. The prevalence and nature of glycan alterations on specific proteins in pancreatic cancer patients revealed using antibody-lectin sandwich arrays. *Mol Cell Proteomics*. 2009; 8:1697–1707. [PubMed: 19377061]
8. Brewer CF, Miceli MC, Baum LG. Clusters, bundles, arrays and lattices: novel mechanisms for lectin-saccharide-mediated cellular interactions. *Current opinion in structural biology*. 2002; 12:616–623. [PubMed: 12464313]
9. Stevens J, Blixt O, Glaser L, Taubenberger JK, et al. Glycan microarray analysis of the hemagglutinins from modern and pandemic influenza viruses reveals different receptor specificities. *Journal of molecular biology*. 2006; 355:1143–1155. [PubMed: 16343533]
10. Shida K, Korekane H, Misonou Y, Noura S, et al. Novel ganglioside found in adenocarcinoma cells of Lewis-negative patients. *Glycobiology*. 2010; 20:1594–1606. [PubMed: 20667987]
11. Liu Y, Palma AS, Feizi T. Carbohydrate microarrays: key developments in glycobiology. *Biol Chem*. 2009; 390:647–656. [PubMed: 19426131]
12. Blixt O, Head S, Mondala T, Scanlan C, et al. Printed covalent glycan array for ligand profiling of diverse glycan binding proteins. *Proceedings of the National Academy of Sciences of the United States of America*. 2004; 101:17033–17038. [PubMed: 15563589]
13. Porter A, Yue T, Heeringa L, Day S, et al. A motif-based analysis of glycan array data to determine the specificities of glycan-binding proteins. *Glycobiology*. 2010; 20:369–380. [PubMed: 19946132]
14. Maupin KA, Liden D, Haab BB. The fine specificity of mannose-binding and galactose-binding lectins revealed using outlier-motif analysis of glycan array data. *Glycobiology*. 2011; 22:160–169. [PubMed: 21875884]

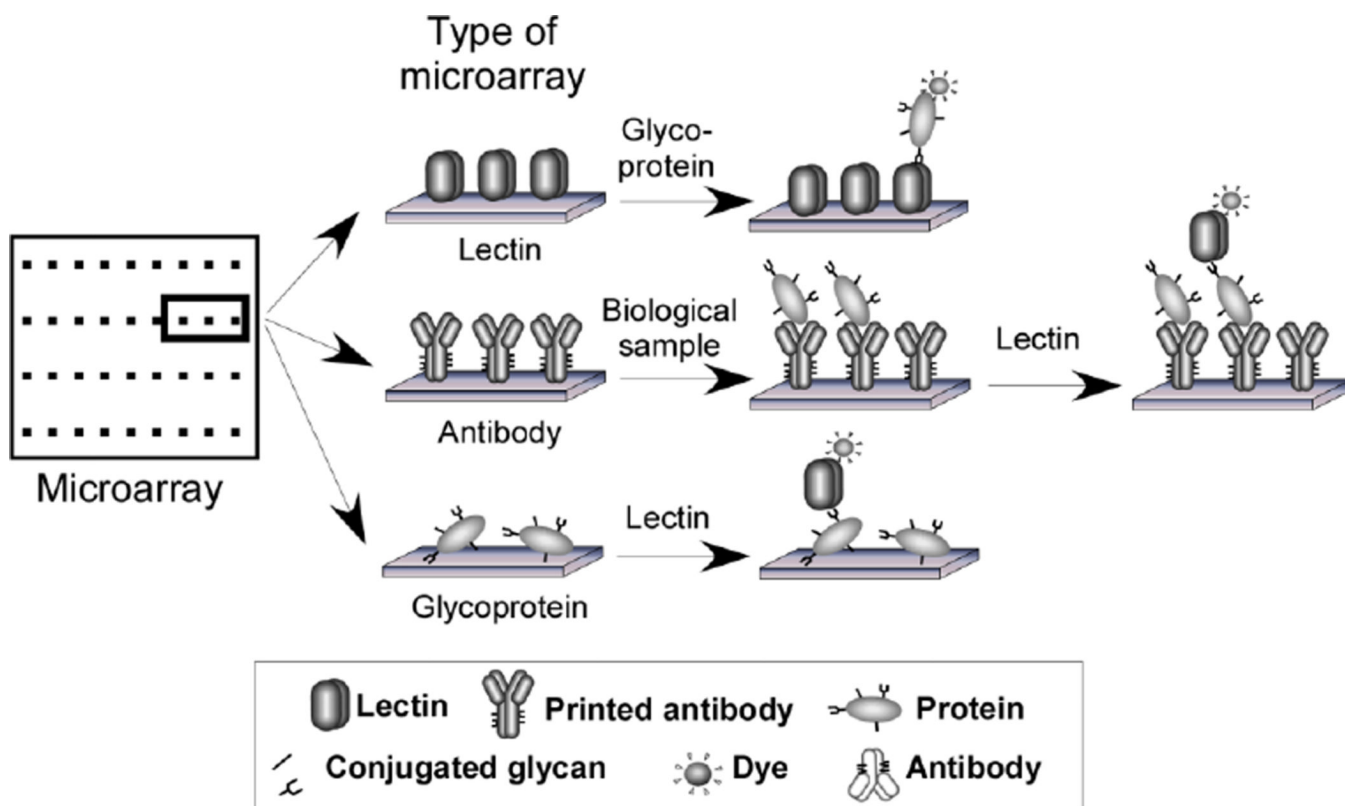
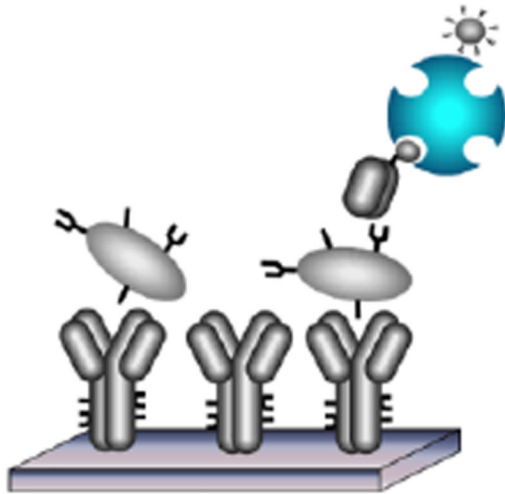


Figure 1. Lectin-based detection of glycans in microarray formats

The types of microarrays depicted are lectin arrays, antibody-lectin sandwich arrays, and glycoprotein arrays. A detection strategy using a fluorescent dye is depicted, although other detection methods could be used, such as surface-plasmon resonance or chemiluminescence. Antibody-lectin sandwich arrays (middle row) enable the detection of glycans on specific proteins captured from biological solutions. Multiple antibodies are immobilized on a planar support, and the captured proteins are probed using biotinylated detection antibodies, followed by fluorescence detection using phycoerythrin-labeled streptavidin.

Monomer



Multimer

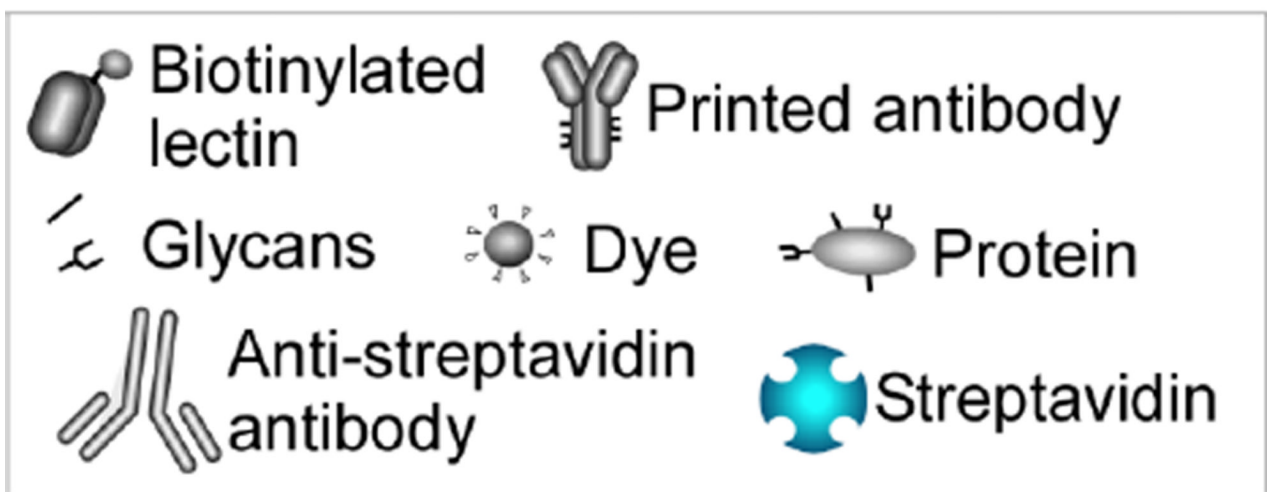
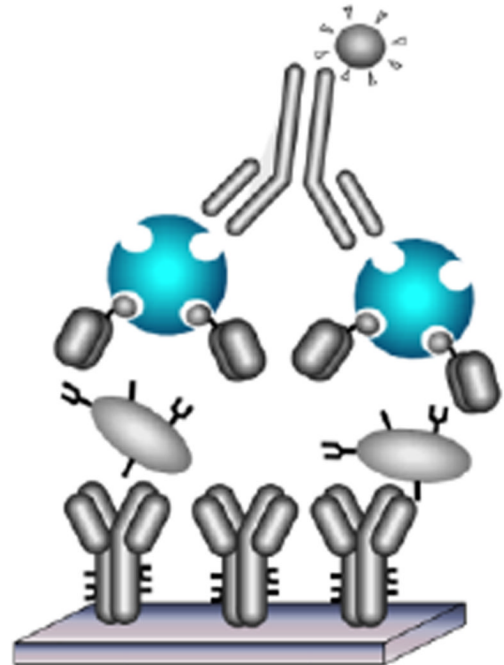


Figure 2. Detection using multimerized lectins

The left panel shows the conventional method of labeling and detecting individual lectins.

The right panel shows a method of linking together lectins through the multiple binding sites of streptavidin and through an anti-streptavidin antibody.

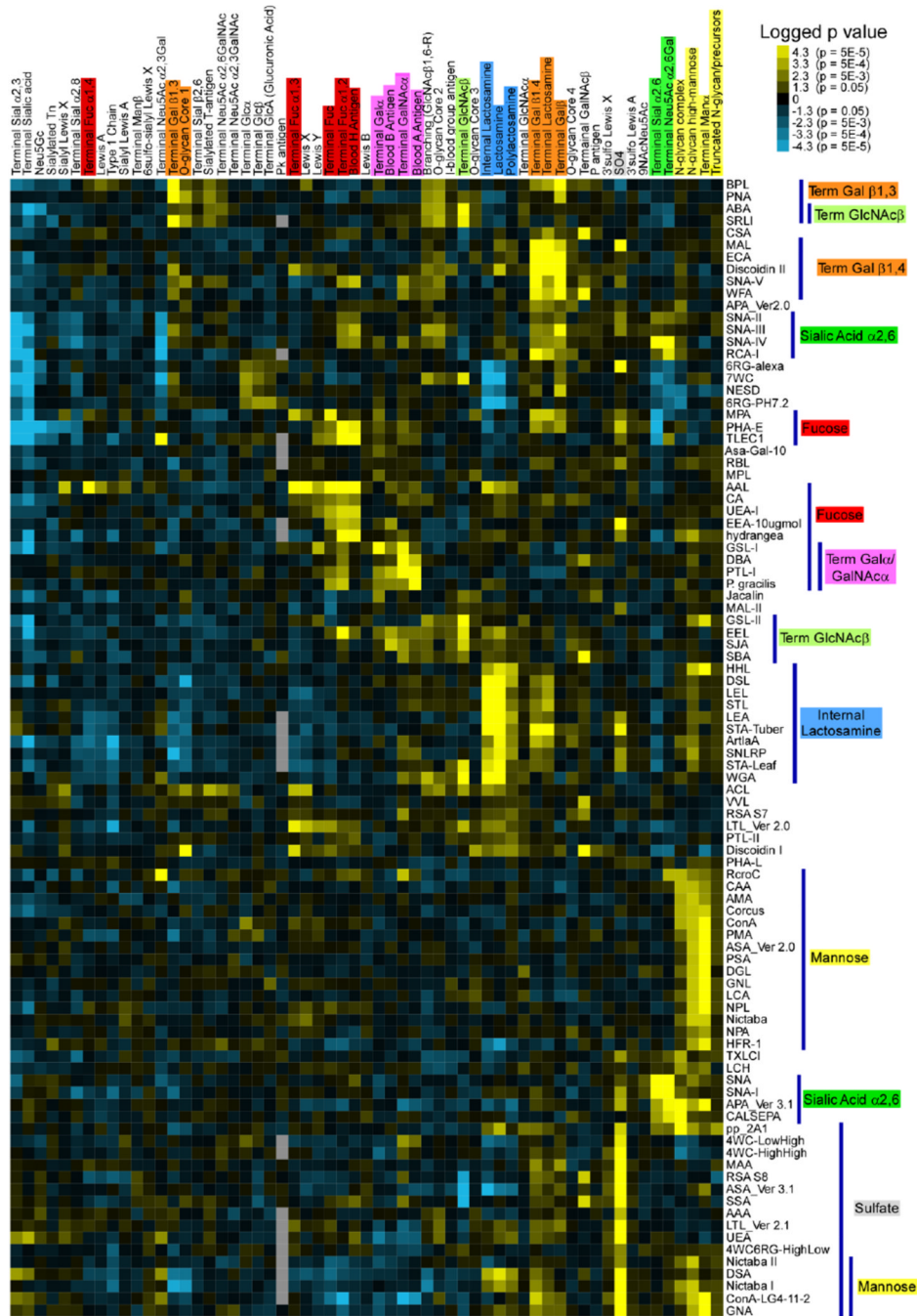


Figure 3. Comparative analysis of the motif specificities of plant lectins
 The data from 113 different glycan array incubations of 84 unique plant lectins were analyzed using the Motif Segregation method, and the logged p values (multiplied by the sign of the z score) are clustered by similarity among both the motifs (columns) and the rows (lectins). This analysis shows the potential for assembling quantitative information from multiple glycan-binding to enable searching and comparative studies. Reprinted from reference [13] with permission from Oxford Journals.