

Neural Network Analyses of Infrared Spectra for Classifying Cell Wall Architectures^{1[W][OA]}

Maureen C. McCann, Marianne Defernez², Breeanna R. Urbanowicz^{2,3}, Jagdish C. Tewari⁴, Tiffany Langewisch, Anna Olek, Brian Wells, Reginald H. Wilson, and Nicholas C. Carpita*

Department of Biological Sciences (M.C.M., J.C.T., T.L.) and Department of Botany and Plant Pathology (B.R.U., A.O., N.C.C.), Purdue University, West Lafayette, Indiana 47907; Department of Food Material Sciences, Institute of Food Research, Colney, Norwich NR4 7UA, United Kingdom (M.D., R.H.W.); and Department of Cell and Developmental Biology, John Innes Centre, Colney, Norwich NR4 7UH, United Kingdom (B.W.)

About 10% of plant genomes are devoted to cell wall biogenesis. Our goal is to establish methodologies that identify and classify cell wall phenotypes of mutants on a genome-wide scale. Toward this goal, we have used a model system, the elongating maize (*Zea mays*) coleoptile system, in which cell wall changes are well characterized, to develop a paradigm for classification of a comprehensive range of cell wall architectures altered during development, by environmental perturbation, or by mutation. Dynamic changes in cell walls of etiolated maize coleoptiles, sampled at one-half-d intervals of growth, were analyzed by chemical and enzymatic assays and Fourier transform infrared spectroscopy. The primary walls of grasses are composed of cellulose microfibrils, glucuronoarabinoxylans, and mixed-linkage (1 → 3),(1 → 4)-β-D-glucans, together with smaller amounts of glucomannans, xyloglucans, pectins, and a network of polyphenolic substances. During coleoptile development, changes in cell wall composition included a transient appearance of the (1 → 3),(1 → 4)-β-D-glucans, a gradual loss of arabinose from glucuronoarabinoxylans, and an increase in the relative proportion of cellulose. Infrared spectra reflected these dynamic changes in composition. Although infrared spectra of walls from embryonic, elongating, and senescent coleoptiles were broadly discriminated from each other by exploratory principal components analysis, neural network algorithms (both genetic and Kohonen) could correctly classify infrared spectra from cell walls harvested from individuals differing at one-half-d interval of growth. We tested the predictive capabilities of the model with a maize inbred line, Wisconsin 22, and found it to be accurate in classifying cell walls representing developmental stage. The ability of artificial neural networks to classify infrared spectra from cell walls provides a means to identify many possible classes of cell wall phenotypes. This classification can be broadened to phenotypes resulting from mutations in genes encoding proteins for which a function is yet to be described.

As our knowledge of the transcriptome and proteome increases, we are still deficient in our understanding of how genes and their products contribute to downstream phenotypes. For cell wall polysaccharides, the plant's principal biomass, the problem of attempting to connect genotype with phenotype is

especially difficult. The cell wall comprises several structurally complex polymers, most of which are synthesized in the Golgi, secreted to the plasma membrane, and assembled with cellulose microfibrils extracellularly (McCann and Roberts, 1991; Carpita and Gibeau, 1993). Although over 1,000 cell wall-related genes have been annotated, in many instances, neither the biochemical function nor the expression patterns of these genes is established (McCann and Carpita, 2005). Most importantly, plants are highly responsive to alterations in their cell walls, such as those induced by mutation or environmental perturbation, and gene expression or enzyme activities are adjusted to compensate for deficiencies in unanticipated ways (Iraki et al., 1989; Shedletzky et al., 1992; His et al., 2001; Caño-Delgado et al., 2003; Madson et al., 2003).

Somerville et al. (2004) have suggested that a systems approach is useful to integrate information derived from transcript profiling methods and proteomics to define the relationship between genotypes and cell wall structure and function. Such approaches have been most successful in yeast (*Saccharomyces cerevisiae*), where well-characterized metabolic pathways can be studied by application of several technology platforms to mutants representative of most of the elements of the pathway (Ideker et al., 2001). At present, our

¹ This work was supported by the National Science Foundation Genome Research Program (to N.C.C. and M.C.M.), and by the Biotechnology and Biological Sciences Research Council (to M.C.M., M.D., and R.H.W.). Journal paper number 17,933 of the Purdue University Agricultural Experiment Station.

² These authors contributed equally to the paper.

³ Present address: Department of Plant Biology, Cornell University, Ithaca, NY 14853.

⁴ Present address: Department of Food Science, Cornell University, Ithaca, NY 14853.

* Corresponding author; e-mail carpita@purdue.edu; fax 765-494-0363.

The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantphysiol.org) is: Nicholas C. Carpita (carpita@purdue.edu).

[W] The online version of this article contains Web-only data.

[OA] Open Access articles can be viewed online without a subscription.

www.plantphysiol.org/cgi/doi/10.1104/pp.106.093054

knowledge of pathways of synthesis, assembly, and disassembly of cell walls is quite limited because of the lack of a range of characterized mutants.

As cell walls comprise secondary gene products, it is especially important that multivariate data representative of these structures are included in a systems approach. In one example from yeast, the correlation coefficient between changes in transcript and protein levels upon perturbations in carbon metabolism is 0.3 (Ideker et al., 2001). How much less must the correlation be when examining gene expression with secondary gene products, such as the highly complex structures of plant cell wall polysaccharides? We need to be able to characterize and classify the full range of possible cell wall phenotypes. With some understanding, but not a comprehensive catalog, of the molecular complexity of cell wall polysaccharides, their cross-linkages, and their conformations and molecular interactions, this is a daunting task. Previously, we optimized Fourier transform infrared (FTIR) microspectroscopy as a high throughput screen for a broad range of cell wall phenotypes in populations of mutagenized plants (Chen et al., 1998). FTIR spectroscopy has also been applied to cell walls altered by environmental adaptation (McCann et al., 1994; Wells et al., 1994; Encina et al., 2002) and to dwarfed hypocotyl mutants of *Arabidopsis* (*Arabidopsis thaliana*; McCann et al., 2001; Mouille et al., 2003). Midinfrared spectroscopy detects, within underivatized cell walls, the vibrations of all molecular bonds for which the component atoms differ in electronegativity, including asymmetric bonds such as C-H and O-H and particular functional groups, such as esters, amides, and carboxylates. The frequencies of molecular vibrations are modified by the local environment of molecular bonds, as influenced by hydration state, the structure and conformation of the molecule, and interactions with other molecules (McCann et al., 1992, 1997, 2001; Séné et al., 1994; Chen et al., 1997). Therefore, the IR spectrum is characteristic of the architecture as well as the composition of the wall.

The range of cell wall mutants in grass species is very limited (e.g. Li et al., 2002; Marita et al., 2003). However, cell wall-related mutants in maize (*Zea mays*) and *Arabidopsis* are being identified in a National Science Foundation-funded genomics project (<http://cellwall.genomics.purdue.edu>). For maize, Robertson's *Mutator* (*Mu*) elements introgressed into a consistent inbred genetic background, Wisconsin 22 (W22), give a mutagenized population for forward- and reverse-genetic screens; some 41 distinct *Mu* insertions in putative cell wall genes have already been identified (<http://uniformmu.org/cellwall>). In advance of establishing these lines, we require the means to classify a wide range of altered wall architectures in maize. In this article, we exploit the well-characterized dynamics in cell wall composition that occur during elongation of the hybrid and W22 maize coleoptiles to test the sensitivity of IR spectra to detect defined changes in monosaccharide and polysaccharide composition.

Grass species and related commelinoid monocot cell walls display considerable and well-characterized variation in polymer synthesis, assembly, and hydrolysis during coleoptile growth (Darvill et al., 1978; Labavitch and Ray, 1978; Carpita and Gibeaut, 1993; Carpita, 1996). When data from several different assays are evaluated together, minor and complex changes are revealed even between one-half-d intervals of growth. Although exploratory principal components analysis (PCA) could broadly discriminate IR spectra of embryonic, elongating, and senescent walls of etiolated maize coleoptiles, artificial neural networks (ANNs) refined the discrimination to cell walls representing only one-half-d intervals of growth. ANNs are machine-learning tools that can identify arbitrary discriminant functions directly from experimental data for purposes of classification, particularly useful where a mechanistic description of the dependency between measured variables and predicted class is either unknown or very complex (Almeida, 2002). The ability of ANNs to classify these cell walls suggests that ANN analyses of IR spectra are able to integrate several minor changes into a classification structure. Our purpose is not to classify either the age or length of coleoptiles by this method but to establish an assay that can detect many changes in cell walls, compositional and architectural, and classify cell walls by their spectroscopic phenotypes, or spectrotypes. We predict that this approach will not only have tremendous utility in application to maize mutants related to cell wall biology (see <http://cellwall.genomics.purdue.edu>) but will also provide a paradigm for classification of a broad range of cell wall structures and architectures that may vary as a consequence of mutation, evolutionary speciation, cell development, and environmental response.

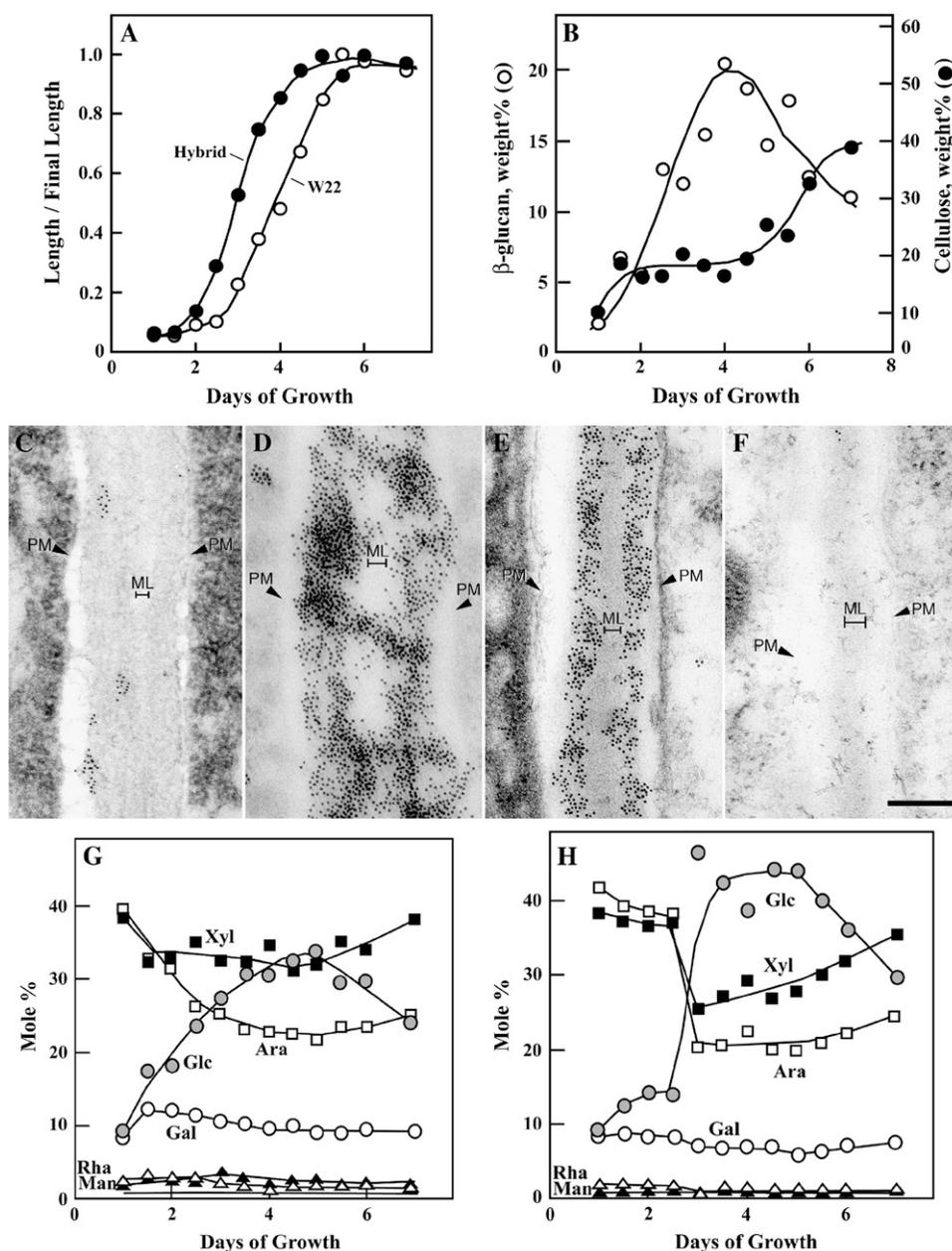
RESULTS

Changes in Cell Wall Composition during Elongation of Hybrid and W22 Coleoptiles

Etiolated hybrid coleoptiles begin to elongate about 1 d after imbibition and reach a maximum elongation rate between 2 and 3 d (Fig. 1A). Final length is achieved between 5 and 6 d, when the etiolated leaf blades emerge and split the coleoptile open. After leaf emergence, coleoptiles begin to senesce, and subsequent loss of water causes a slight shrinkage. Germination and elongation of W22 coleoptiles is delayed about one half day with respect to hybrid coleoptiles, but cessation of growth and senescence are coincident with the hybrid. Figure 1A shows physiological growth stages of hybrid and inbred coleoptiles normalized to illustrate the equivalent transition stages. However, the hybrid coleoptile achieves a maximal length of about 4.3 cm, whereas that of W22 is 3.5 cm.

In hybrid coleoptiles, cellulose content is initially 10% of the dry mass of the wall and increases to 20% at the onset of elongation. The proportion of cellulose

Figure 1. Changes in cell wall composition during growth of etiolated maize coleoptiles. **A**, Changes in length of hybrid and W22 coleoptiles. Final lengths have been normalized to illustrate the timing of equivalent developmental stages. Actual final lengths are 4.3 cm for hybrid and 3.5 cm for W22. **B**, Cellulose (black circles) and (1 → 3), (1 → 4)- β -D-glucan (white circles) contents as percent dry weight of cell wall. **C** to **F**, Immunolocalization of (1 → 3), (1 → 4)- β -D-glucan in mesophyll cells of hybrid coleoptiles at 2 d old (**C**), 3 d old (**D**), 4 d old (**E**), and 5 d old (**F**), using 10-nm colloidal gold-conjugated secondary antibody. Plasma membrane (PM) and middle lamellae (ML) are indicated. All transmission electron micrographs are at the same magnification; scale bar = 250 nm. Monosaccharide composition of hybrid (12 replicates per time point; **C**) and W22 (3–5 replicates per time point; **H**) coleoptiles for six monosaccharides, Ara, Xyl, Gal, Glc, Rha, and Man. Error limits have been omitted for clarity. sd were 5% or less for all values.



remains at 20% during elongation but increases to nearly 40% during senescence (Fig. 1B). The (1 → 3), (1 → 4)- β -D-glucan is nearly absent from embryonic coleoptiles but increases rapidly during elongation, reaching a maximum content coincident with the highest rates of elongation, and then decreases markedly as growth ceases (Fig. 1B).

Epitopes of (1 → 3), (1 → 4)- β -D-glucan were immunolocalized using the transmission electron microscope to determine their location within the walls of epidermal (data not shown) and mesophyll hybrid coleoptile cells (Fig. 1, C–F). Before elongation commences on day 2, both cell types have only small amounts of (1 → 3), (1 → 4)- β -D-glucan located on the inner surface of the walls adjacent to the plasma membranes (Fig. 1C).

However, during maximal rates of elongation, (1 → 3), (1 → 4)- β -D-glucan epitopes are abundantly distributed across the primary walls of mesophyll cells (Fig. 1D). As growth peaks on day 4, the (1 → 3), (1 → 4)- β -D-glucan content decreases close to the middle lamella, but epitopes are still abundant adjacent to the plasma membrane (Fig. 1E). During senescence, small amounts of (1 → 3), (1 → 4)- β -D-glucan epitopes are present (Fig. 1F). These data are generally consistent with the timing of detection of characteristic oligomers after digestion of maize walls with *Bacillus subtilis* endoglucanase (Fig. 1B). However, the quantities of (1 → 3), (1 → 4)- β -D-glucan oligomers are roughly equal on days 3 and 5, but few epitopes are detected in 5-d-old compared with 3-d-old coleoptile walls. Two possibilities are that other

changes in the walls of 5-d-old coleoptiles contribute to masking the epitope, or that the remaining (1 → 3), (1 → 4)- β -D-glucan in the wall is oligomeric and not long enough to form an epitope because of the activity of an endogenous exoglucanase (Kim et al., 2000).

Noncellulosic monosaccharide compositions were obtained from populations of hybrid and W22 coleoptiles by gas chromatography-mass spectrometry of alditol acetates. Xyl, Glc, Ara, and Gal are the four most abundant monosaccharides (Fig. 1, G and H). Ara content decreases during growth, and Xyl later increases toward the end of growth in both hybrid and W22 coleoptiles. Gal remains constant at about 10 mol %. Changes in Glc correlate to some extent with the changes observed in (1 → 3), (1 → 4)- β -D-glucan content, but Glc is also a constituent of glucomannan and xyloglucan.

Dynamic Changes in Cell Wall Composition Are Documented by FTIR Spectra

At least 36 FTIR spectra were obtained from cell walls of populations of coleoptiles from each one-half-d interval; the averaged spectra for each interval reflect the compositional changes that occur during the course of cell elongation (Fig. 2A). Multivariate partial least squares predicted with a correlation coefficient of between 0.75 and 0.95 the mol % of the four most abundant monosaccharides (Supplemental Fig. S1). Peak assignments are taken from Carpita et al. (2001) and Séné et al. (1994) and references therein. All of the spectra appear broadly similar and have absorbances in the carbohydrate fingerprint region, with peaks at 1030, 1060, 1103, and 1157 cm^{-1} . Spectra from 1- to 2-d-old coleoptile cell walls have higher absorbances from amide bands of proteins at 1650, 1628, and 1550 cm^{-1} and carbonyl ester absorbances at 1740 and 1231 cm^{-1} . Spectra from 4.5 to 7 d old have a small but distinct absorbance at 1515 cm^{-1} characteristic of phenolic ring structures. However, many subtle changes in peak height and shape occur, particularly in the carbohydrate fingerprint region. These are more evident in digital subtraction spectra. For example, we can generate spectra representing embryonic, elongating, or senescent coleoptile walls by pooling and averaging spectra obtained from 1 and 1.5 d old, 3.5 and 4 d old, and 5.5 and 6 d old, respectively (Fig. 2B). Digital subtraction of the average spectrum representing elongating walls minus that of embryonic walls shows that cellulose content is relatively enriched in the elongating walls because characteristic peaks of cellulose (Tsuboi, 1957; Liang and Marchessault, 1959; 1157, 1103, and 1053 cm^{-1}), as well as absorbances from carbohydrates at 1018, 991 (not assigned), and 898 cm^{-1} (β -linked polymers) are in the digital subtraction spectrum. The negatively correlated peaks at 1651, 1628, and 1535 cm^{-1} indicate that protein content of the embryonic walls is greater than that in the elongating walls (Fig. 2C). Similarly, digital subtraction of the average spectrum of senescent walls minus that of elongating walls

shows carbohydrate peaks at 1157, 1068, and 1041 cm^{-1} that are relatively enriched in elongating walls. The negatively correlated peaks at 1693, 1593, and 1515 cm^{-1} may represent higher content of aromatic compounds (Faix, 1992; Séné et al., 1994) in the senescing walls (Fig. 2D).

Spectra are constituted by a discrete series of measurements at defined intervals; by nature, they are multivariate. Statistical algorithms for treating multivariate data such as PCA (Kemsley, 1998) can be applied to reveal molecular features that are the basis for discriminating among populations (Chen et al., 1998). PCA reduces the high dimensionality of spectral data to a smaller set of computer-derived variables, termed principal components (PCs), that together account for all of the variance in a set of (sometimes apparently similar) spectra (Kemsley, 1998). Each spectrum has an associated value for each new variable, the PC score, representing its relative distance from the mean of the population. Exploratory PCA shows that cell walls representing individual embryonic coleoptiles (days 1 and 1.5) are generally well resolved from elongating (days 3.5 and 4) and senescent (days 5.5 and 6) coleoptiles (Fig. 2E). Cell walls from embryonic coleoptiles are resolved from elongating and senescing coleoptiles by PC 1, the latter having higher scores on this component. Walls from both embryonic and elongating stages have generally higher PC 2 scores than walls of the senescent coleoptiles. The loadings reveal some of the spectral features that are associated with PC 1 and PC 2 as sources of variance (Fig. 2, F and G). For PC 1, which accounts for 61% of the variance, the loading is positive for peaks associated with cellulose and other β -linked cross-linking glycans (1157, 1103, 1053, and 898 cm^{-1}), as well as absorbances at 1018 and 991 (not assigned) and negative for peaks associated with proteins (1651, 1628, and 1535 cm^{-1}). Walls of elongating coleoptiles were enriched in compounds contributing to positive PC 1 loadings relative to the embryonic walls. In this instance, the loading for PC 1 resembles features of the digital subtraction spectrum (Fig. 2, C and F). In contrast, PC 2 loadings, accounting for 22% of variance, were independent of protein content but suggested enrichment in embryonic and elongating coleoptile walls of compounds with absorbances at 1168, 1111, 1083, and 1041 cm^{-1} of the carbohydrate fingerprint region compared to those of senescent coleoptiles.

Linear Discriminant Analysis of Infrared Spectra Using PCA

For the three groups of spectra derived from embryonic, elongating, and senescent cell walls (Fig. 2E), the percentages of correct classification of individual spectra to each group by cross-validation tests were 69% using five PCs. However, if spectra representing one-half-d and 1-d intervals of growth are included for analysis by exploratory PCA, many of 10 subpopulations of spectra are overlapped in the scores plot and not resolved from each other (Fig. 3). As expected, the

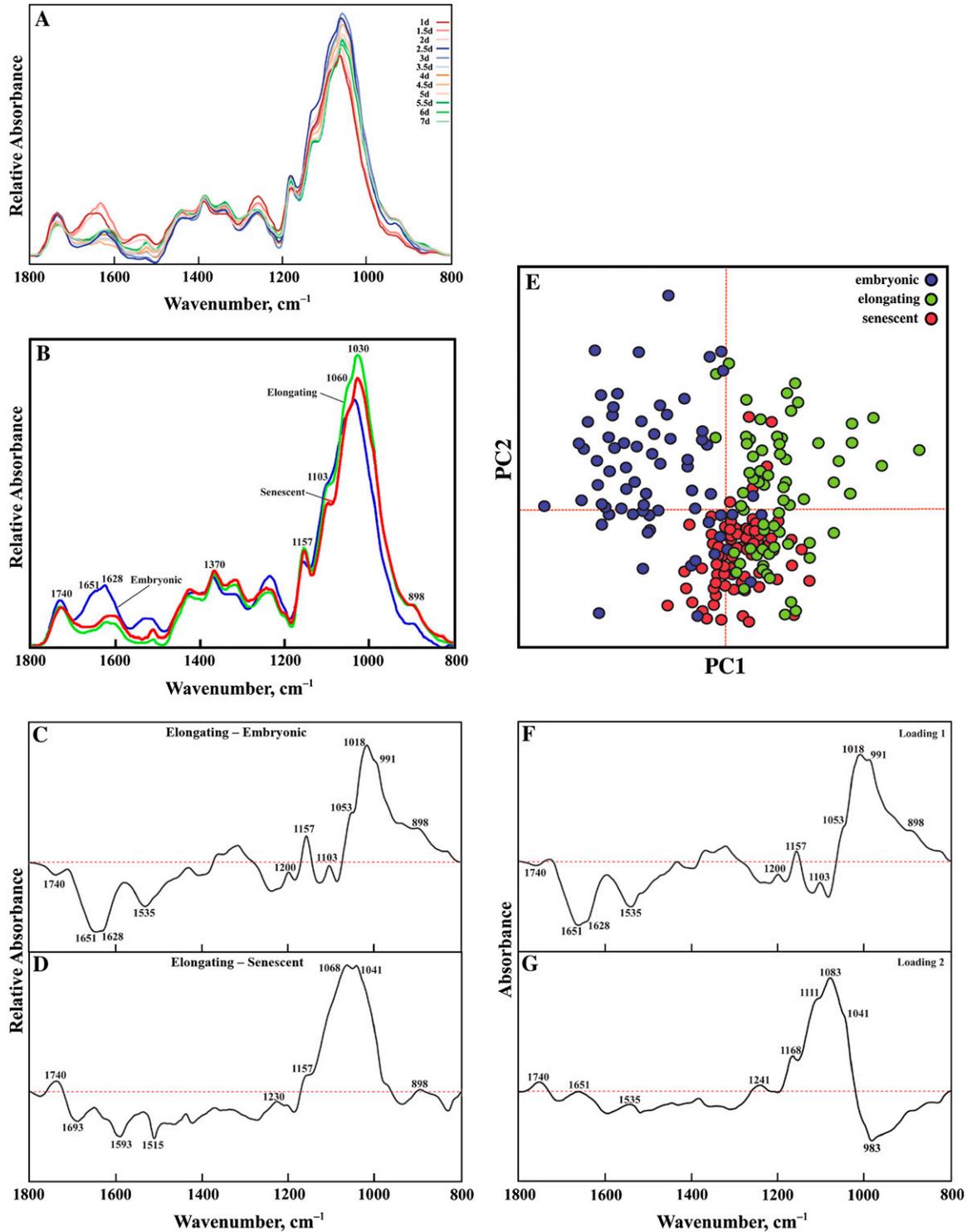


Figure 2. Dynamic changes in cell wall composition of hybrid maize coleoptiles are reflected in FTIR spectra. A, Baseline-corrected, area-normalized spectra of walls representing half-d intervals of growth. Spectra are averaged from 36 spectra of cell walls sampled from populations of coleoptiles at each time point. B, Averaged spectra of walls representing embryonic (days 1 and 1.5; blue), elongating (days 3.5 and 4; green), and senescent stages of growth (days 5.5 and 6; red). C and D, Digital subtraction spectra of the averaged spectra representing elongating walls minus embryonic walls (C), and elongating walls minus senescent walls (D). E, Plot of the first two PC scores obtained by exploratory PCA of individual spectra from populations of walls of embryonic (days 1 and 1.5), elongating (days 3.5 and 4), and senescent (days 5.5 and 6) coleoptiles. F and G, Loadings corresponding to PC 1 (F) and PC 2 (G).

loadings for PC 1 and PC 2 (data not shown) closely resemble those for the scores plot of Figure 2, F and G. A linear discriminant analysis (LDA) by cross validation after PCA correctly classified an average of only 32% of all spectra using five PCs, ranging from 3% (day 4) up to 61% (day 5.5) for each growth interval (Table I). Thus, PCA is not robust in classification when many classes are included and when only slight differences in wall composition or architecture may distinguish each class. As our ultimate goal is to be able to compare cell wall phenotypes on a genome-wide scale, we need a classification tool that is robust for many classes of data.

Analysis of the Spectral Data for Coleoptile Cell Walls by ANN

As exploratory PCA was not able to discriminate 10 growth stages (Fig. 3), we applied ANN analyses to our spectroscopic data. ANNs are algorithms with the capacity to analyze multivariate data (each spectrum is 250 variates) from large numbers of observations (spectra derived from individuals within populations) and large numbers of potential classes of observations within a data set (all 12 growth stages). The neural network reports both on class assignment (percent correct classification) and also on the probability of membership of each class, for each spectrum. As for the LDA described above, we have used cross validation to measure the ability of the classification tool in prediction. The spectra used to test the predictive ability of the network are not included in the set of spectra used to train the network. Our results are from six different networks, each of which was trained with five-sixths of the data, reserving a different one-sixth in each case. Then we summed the predictions of those test sets for the six networks. It should be noted that the success rate by cross-validation analysis for both PCA and ANNs is lower than classification success if all data are included and success is measured simply as classification to the correct class. We applied two kinds of neural network to the spectroscopic data collected from the time course of hybrid coleoptile growth.

First, we used a supervised approach provided by a genetic algorithm (Neuroshell 2, Ward Systems Group, proprietary). A genetic network comprises three layers: an input layer that comprises the absorbance values of infrared spectra at 250 discrete wavenumbers, a hidden layer of neurons, each of which is connected to all of the inputs and all of the outputs, and an output layer, which are the classes to which each individual might belong, defined by growth interval (Goldberg, 1989; Chen, 1996). The network is termed supervised, because it is trained by a large number of spectra for which the output classes (expected growth intervals) are already known and defined, and this information is used by the network to optimize classification success. The success of classification can be represented simply as the number of correct and incorrect assignments for each output class (termed the confusion matrix), or as

the average probability that an individual spectrum will belong to each class, totaling one for the 12 output classes. In Figure 4A, both the confusion matrix and color-coded probability values are presented to show that the majority of spectra of all growth intervals are correctly classified and that probabilities extend mainly to neighboring time points. A genetic network representing each one-half-d interval of growth correctly classifies between 52% (day 3) and 77% (day 5.5) of test spectra for each growth interval by internal cross validation, and an average of 65% of all spectra (Fig. 4A; Table I). Similar results were obtained when the ANN trained using all of the spectra was used to classify data (36 spectra per growth interval) obtained from an independent experiment (data not shown). All classification errors in the ANN were made to neighboring growth intervals, and the percentage of classification to a growth interval plus or minus one-half-d ranged from 83% (day 2) to 97% (day 4.5; Table I). We used correspondence analysis to visualize the similarity relationships between classes (Jobson, 1992). Each row of the input matrix for the correspondence analysis consisted of the classification probabilities (probability of membership for each of 12 possible classes) for each of 36 spectra for each class, a total of 432 spectra. Spectra representing embryonic stages are more closely clustered, whereas one half d later, a cluster of four time points representing elongating coleoptiles is distinguished from a third cluster representing senescent coleoptiles (Fig. 4B).

Second, we used an unsupervised Kohonen network (Lavine et al., 2004), in which the output classes are not defined for each spectrum. Instead of optimizing classification to predefined classes, the network simply assigns spectra to a number of classes based on structure that it detects within the entire data set. Thus,

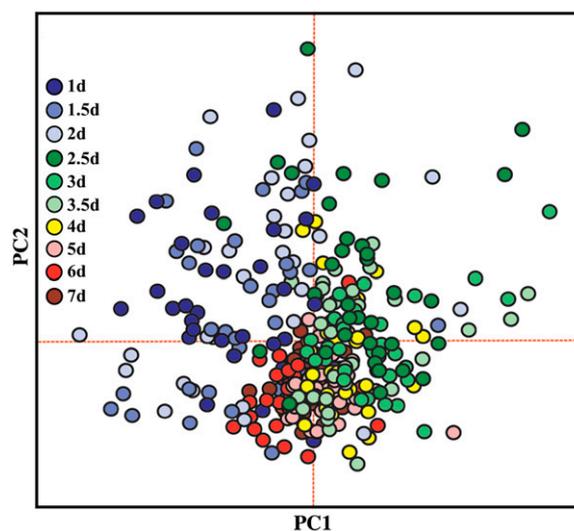


Figure 3. Exploratory PCA of populations of maize hybrid coleoptile walls sampled at one-half-d intervals of growth. A plot of the first two PC scores obtained from individual spectra of coleoptiles at 10 time points.

Table 1. Comparison of percentages of infrared spectra of hybrid maize coleoptile cell walls correctly classified in cross-validation tests using PCA or ANN

Coleoptile Age/Day	Classified to One Half Day			Classified to Day \pm One Half Day		
	PCA	Genetic	Kohonen	PCA	Genetic	Kohonen
				%		
1	50	72	89	63	86	97
1.5	25	58	80	83	94	100
2	11	58	75	55	83	97
2.5	33	69	83	83	92	100
3	47	52	83	69	94	97
3.5	11	55	53	42	94	75
4	3	69	50	33	94	72
4.5	36	75	97	83	97	100
5	36	66	89	86	89	97
5.5	61	77	69	86	94	78
6	53	56	64	80	92	97
7	17	67	86	72	94	89
Average	32	65	77	70	92	92

we can test whether or not our predicted class structure (classification by one-half-d growth interval) is valid. A Kohonen network comprises input and output layers only; the output layer has one neuron for each possible output category. The number of output classes can be varied; we experimentally increased the number of classes from two to 20, with the best classification observed at 12 classes. Above 12, new classes were left empty (data not shown). In contrast to the relatively poor classification by PCA (3%–61%), the Kohonen network correctly classified an average of 77% of all spectra, ranging from 50% (day 4) to 97% (day 4.5) for each growth interval and from 72% (day 4) to 100% (days 1.5, 2.5, and 4.5) for growth intervals plus or minus one half d (Fig. 5A; Table I). The cluster plot of the correlation analysis (Fig. 5B) is similar to that of the genetic algorithm (Fig. 4B). In the cluster plots (Figs. 4B and 5B), only the first two sources of variance from the correspondence analysis are plotted, but values for all degrees of freedom (for n groups comprising probability scores of class membership, $n - 1$ F values) are obtained.

Testing the Model with the W22 Population

The cell walls of hybrid maize coleoptiles have been well characterized with respect to the dynamics of composition and architecture during growth (Carpita et al., 2001), whereas these dynamics have not been well studied in the foundation inbred lines from which hybrids have been derived. We established the growth parameters for a key inbred, W22, in which *Mu* elements had been introgressed, generating a large population of transposon-tagged mutant lines (<http://uniformmu.org/cellwall>). The changes in length and cell wall composition of W22 coleoptiles taken at one-half-d intervals over the growth period were similar to those of the hybrid, although the W22 embryonic phase before rapid elongation was protracted and the final length was around 75% that of the hybrid (Fig. 1A).

The spectra from the time course of hybrid coleoptile elongation were used to train a genetic network, with output classes specified as each of the one-half-d growth intervals. In this case, the genetic, rather than Kohonen, network is appropriate, because we know that the output classes are different from each other on the basis of their chemical compositions. We have prior knowledge of what the classification structure should be, and this knowledge can be used to optimize the predictive ability of a genetic network. The network, trained with spectra from the hybrid time course, was then tested for its ability to classify populations of W22 walls sampled at one-half-d intervals. The confusion and probability matrices show that incorrect assignments and probability values were made generally to neighboring growth intervals (Fig. 6A). The lag of one half d before the onset of rapid cell elongation is reflected by assignment of W22 embryonic walls to the next earlier one-half-d hybrid class instead of the elongating wall cluster (Fig. 6B). After the 3.5-d interval, W22 and the hybrid maize were clustered generally at the equivalent one-half-d age representative of physiological age; the difference between hybrid and W22 in final coleoptile length before senescence was not a criterion in assignment.

However, the model described above assumes that the test set of data from W22 coleoptiles can be classified into the same 12 classes as the training set of data from hybrid coleoptiles. We can establish whether this assumption is justified using a Kohonen network in which 24 classes are specified as the number of possible classes into which individual spectra may be assigned. The confusion and probability matrices show that spectra for some growth intervals were assigned almost equally as hybrid time points or the lagging W22 time points (Fig. 7A). Only coleoptiles at maximal elongation rate could be clearly resolved into distinct classes. In this case, the cluster plot shows a tight mapping of the W22 time course with the hybrid time course, showing that the clustering in Figure 7A was not the

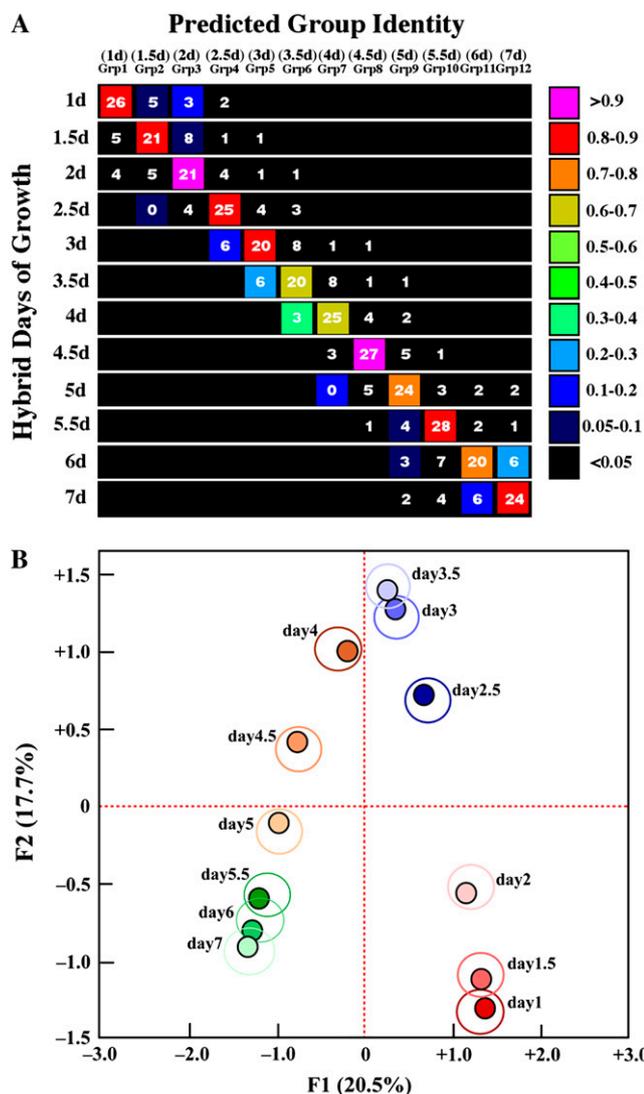


Figure 4. Analysis of IR spectra sampled at one-half-d intervals of hybrid coleoptile growth using a genetic algorithm. A, Color-coded average probabilities of assignment of individual spectra to their correct one-half-d interval of growth. Numbers indicate actual numbers of spectra assigned to each growth interval. For probabilities below 0.05 (black squares), the zero values of classified have been omitted. B, Cluster plot of the first two F values, accounting for 38.2% of variance in probability scores of 36 individual spectra for each growth interval, generated by the genetic algorithm. The plot shows internally cross-validated results for each growth interval with the large majority of individuals clustering in the black-outlined circles and all individuals within the outer circle.

result of forcing the data into only 12 possible classes (Fig. 7B). As the cluster plot is two dimensional, we can take account of all dimensions by constructing a dendrogram based on the average probability values for each of the 24 possible classes (Fig. 8). For each set of 36 spectra that actually belong to a single class, for example, W22 1-d-old coleoptile walls, each spectrum has a set of 24 probability values of belonging to the 24 classes, most of which are zero. We can average the probabilities for each class for the 36 spectra and then

use these average probability values to construct a dendrogram. The dendrogram shows tight relationships between walls of similar developmental stage and composition, whether W22 or hybrid, independent of age.

DISCUSSION

In this article, we establish ANNs as a suitable classification tool for altered wall phenotypes. We used a model system, the growth of the maize coleoptile, as a system in which there are defined and well-characterized changes in wall composition and architecture (Carpita, 1996; Carpita et al., 2001).

ANNs Provide Improved Classification with Large Data Sets of Cell Wall Phenotypes

We can monitor monosaccharides, their linkages, oligomers or epitopes of particular polysaccharides, and cellulose content by means of labor-intensive assays that are each indicative but not comprehensive of complex architectural changes. FTIR spectroscopy has advantages of providing a single and rapid assay that is sensitive to a large range of compositional and architectural features (polymer conformations, hydration state, and extent and nature of cross-linking, features that affect frequencies of molecular vibrations in the infrared spectrum) that most chemical assays cannot detect. In many instances, we cannot interpret spectral peaks in terms of specific wall modifications, but we can at least detect that changes have occurred. Our measurements of monosaccharide, cellulose, and $(1 \rightarrow 3),(1 \rightarrow 4)\text{-}\beta\text{-D-glucan}$ content at one-half-d growth intervals for both hybrid and inbred coleoptiles are consistent with previously documented changes (Carpita and Gibeau, 1993; Carpita, 1996; Carpita et al., 2001). The $(1 \rightarrow 3),(1 \rightarrow 4)\text{-}\beta\text{-D-glucans}$ that appear transiently during cell growth are found only in the Poales (Carpita, 1996; Smith and Harris, 1999). Between each growth interval, changes are slight but involve multiple components of the wall. IR spectra reflect the sum of these small shifts in multiple components. For example, in the partial least squares (PLS) regression, the correlation coefficient between actual and predicted Rha content was 0.87 even though Rha content varies only between 2 and 4 mol % (Supplemental Fig. S1). However, although a broad clustering of embryonic coleoptiles versus elongating or senescing coleoptiles was observed in the first two PCs, a discriminant analysis based on PCA of the spectra performed relatively poorly at assigning spectra to individual half-d intervals. In contrast, these transitional growth stages were reasonably well discriminated by ANNs (Table I).

Fully connected feed-forward ANNs have been widely used for multivariate data, including IR spectra (Kell et al., 2001), to provide sensitive and robust classification tools (Almeida, 2002). The technology

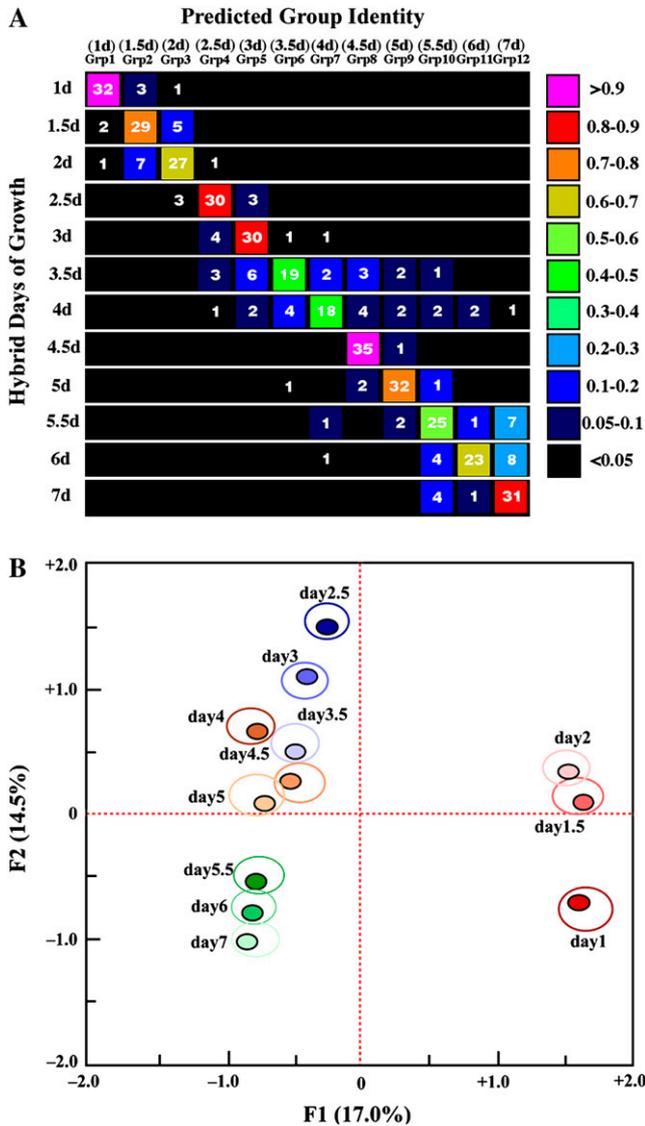


Figure 5. Analysis of IR spectra sampled at one-half-d intervals of hybrid coleoptile growth using an unsupervised (Kohonen) neural network. A, Color-coded average probabilities of assignment of individual spectra to their correct one-half-d interval of growth. Numbers indicate actual numbers of spectra assigned to each growth interval. For probabilities below 0.05 (black squares), the zero values of classified have been omitted. B, Cluster plot of the first two *F* values, accounting for 31.5% of variance in probability scores, generated by the Kohonen network. The plot shows internally cross-validated results with the large majority of individuals clustering in the black-outlined circles.

platforms of transcriptomics and proteomics are used in systems biology, because they provide multivariate information relatively rapidly for individual samples. PCA is used in clustering algorithms for microarray data (Misra et al., 2002), and ANNs are used in bioinformatics to predict motifs from gene or protein sequences (Almeida, 2002). We have used two types of neural networks. For both of these, we have used our spectroscopic variables as inputs, and we have randomized the order in which the spectra are presented

to the networks. Both genetic and Kohonen networks are successful at being able to assign test spectra correctly to different classes, either in cross-validation tests or using replicate data sets. When we trained a genetic network using data from hybrid coleoptiles, the developmental stage of test spectra obtained from W22 coleoptiles was predicted accurately and was tightly correlated with physiological age rather than actual age, as reflected by cell wall composition. However, such an analysis assumed that the classification

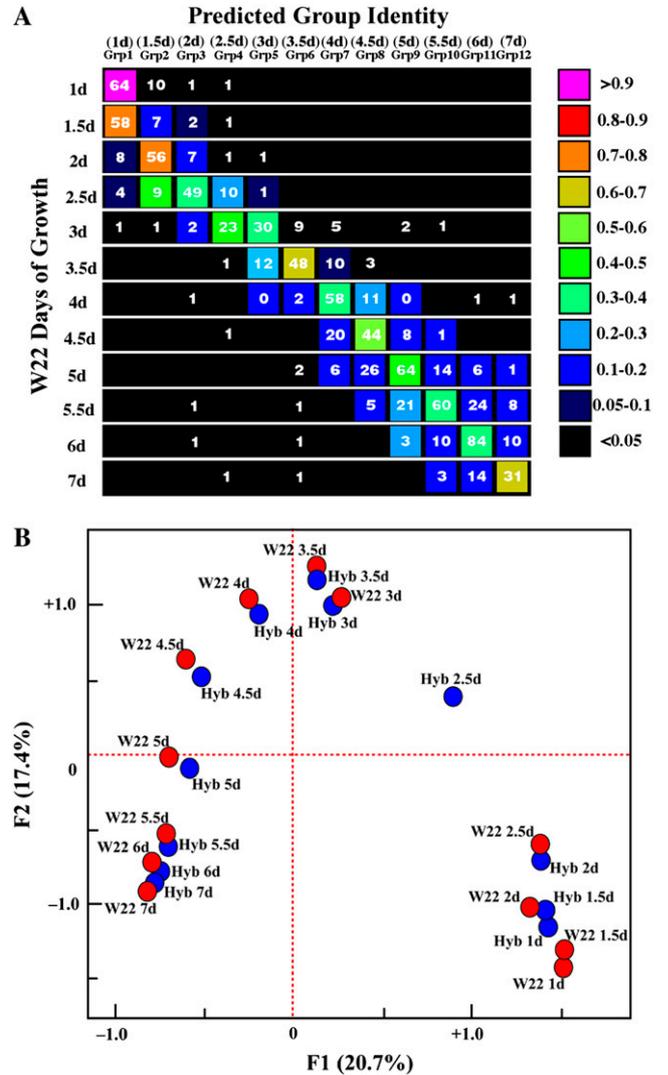


Figure 6. Classification of individual spectra of W22 etiolated maize coleoptiles to their predicted one-half-d interval of growth using a genetic algorithm trained with spectra from hybrid maize coleoptiles. A, Color-coded average probabilities of assignment of individual spectra of W22 coleoptiles to classes representing one-half-d interval of growth of hybrid coleoptiles. Numbers indicate actual numbers of spectra assigned to each growth interval. For probabilities below 0.05 (black squares), the zero values of classified have been omitted. B, Cluster plot of the first two *F* values, accounting for 38.1% of variance, generated by the genetic algorithm. The plot shows internally cross-validated results with the large majority of individuals clustering in the circles.

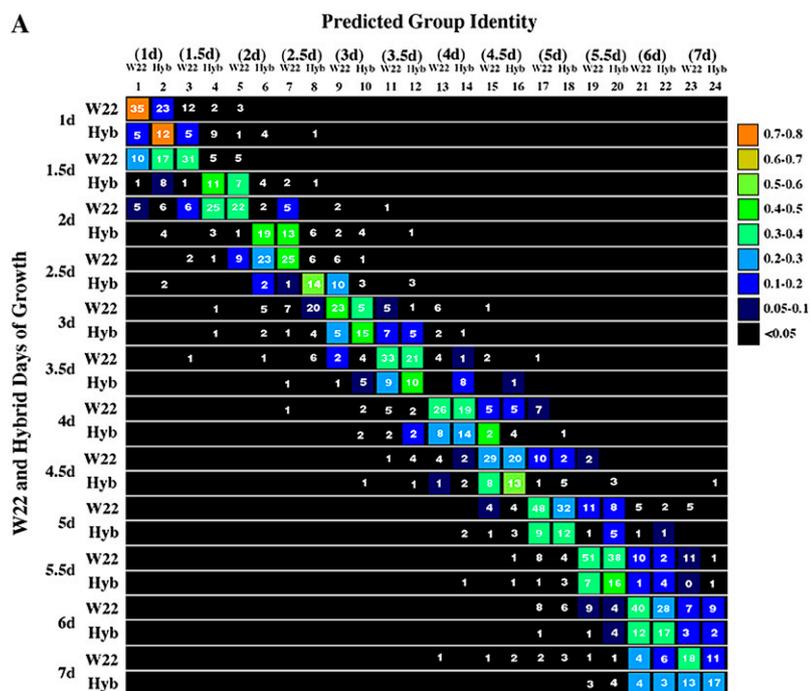
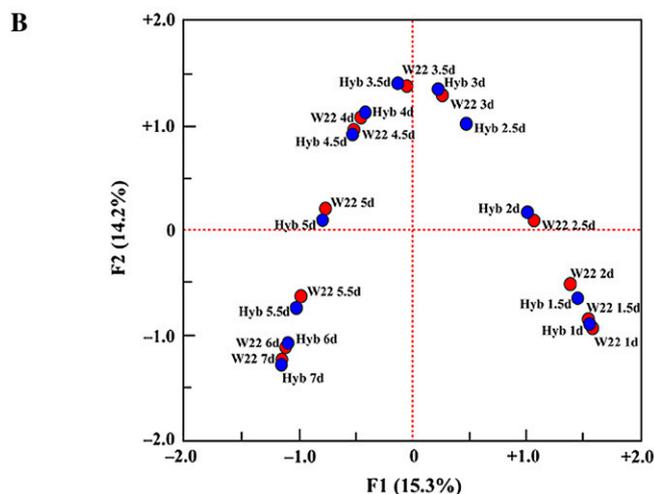


Figure 7. Classification of individual spectra of W22 and hybrid etiolated maize coleoptiles to their predicted genotype and one-half-d interval of growth using a Kohonen network with 24 classes. A, Color-coded average probabilities of assignment of individual spectra of W22 and hybrid coleoptiles to 24 classes representing one-half-d intervals of growth for each genotype. Numbers indicate actual numbers of spectra assigned to each growth interval. For probabilities below 0.05 (black squares), the zero values of classified have been omitted. B, Cluster plot of the first two *F* values, accounting for 29.5% of variance of probability scores, generated by the Kohonen network. The plot shows internally cross-validated results with the large majority of individuals clustering in the circles.



structure by hybrid coleoptile wall composition was appropriate for classification of W22 coleoptile walls. This was confirmed by using a Kohonen network in which the data could be assigned into any of 24 possible classes. The network consistently misassigned W22 coleoptiles into the equivalent developmental stage of hybrid growth and vice versa.

Limitations of ANNs in Characterization of the Bases of Classification

A drawback of ANNs is that the equations that describe the weightings for each neuron, even in a Kohonen network, do not provide a means to visualize individual spectral features that are used for classification in the way that PC loadings can. Therefore, the ANNs reveal class structure, but the nature of the class

structure must either be inferred from prior knowledge of classes or by more limited comparisons in PCA or by digital subtractions. Nevertheless, the probability relationships between spectra can be visualized by correspondence analysis by comparing sources of variance in probability measurements, calculated as *F* values. The first two *F* values in the cluster plot in Figure 6 together represent 31.5% of the total variance in the population. The hysteresis in the cluster plot suggests that at least the *F*2 axis may be dominated by (1 → 3),(1 → 4)-β-D-glucan content, which increases and then decreases over the time course (Fig. 1B). One means of integrating all sources of variance is to construct dendrograms based on the probabilities of classification into the 12 classes (Fig. 8).

By establishing spectrotypes for known cell wall alterations, as confirmed by chemical analyses, we aim

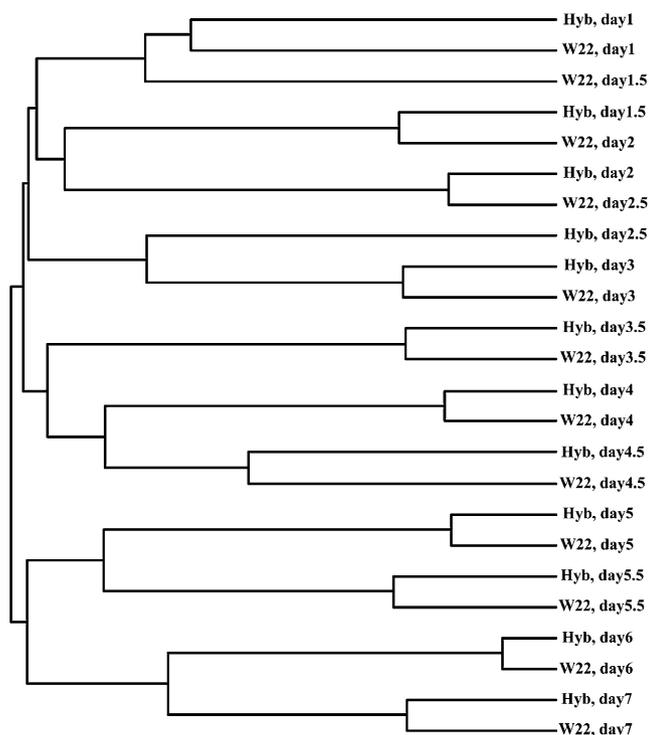


Figure 8. A dendrogram derived from the probability values of a Kohonen network with 24 classes showing the relationships among hybrid and W22 coleoptiles of different ages.

to interpret the spectrotypes of unknowns by comparison. As we did not have prior knowledge of the chemical composition of W22 cell walls, then these samples constitute a set of unknowns that can be mapped on to the model built from the spectra derived from hybrid coleoptile walls and their compositions predicted. These experiments provide us with a means of inferring the sensitivity of the ANNs to compositional changes as reflected in the IR spectra. In screening a collection of mutants for altered wall phenotypes, it is useful to have some well-characterized samples that can be used to train ANNs and that can be used as standards to infer the composition of unknowns. Arabidopsis mutants that affect most of the major cell wall components have been identified and characterized. However, there are comparatively few maize mutants in cell wall-related genes. Until the extended cell wall phenotypes of these mutants have been characterized, a mutant screen could simply measure divergence from the range of developmentally regulated changes in maize wall compositions using unsupervised Kohonen networks.

CONCLUSION

The application of ANNs to spectroscopic data, and to other multivariate measurements of phenotype, provides a framework for the systematic classification of cell wall phenotypes in response to numerous perturbations. We predict that ANNs will have a widespread application to plant cell wall biology. For

example, one may apply these algorithms at the species level for evolutionary-developmental cell wall taxonomies, for tracking genetic variation in recombinant inbred lines for cell wall-related quality traits, and for phenotyping the cell wall changes that occur downstream of signal transduction pathways. A systems approach to cell wall biology is now required to integrate our existing knowledge base of the molecular machinery of the wall and to predict the missing elements of its complex and dynamic architecture.

MATERIALS AND METHODS

Plant Material

Maize (*Zea mays*) hybrid caryopses (Mo17 × B73 foundation; Asgrow Seeds) and W22 caryopses (from Don McCarty and Karen Koch, University of Florida) were soaked overnight in darkness in water bubbled with air at 29°C, sown in moist vermiculite, and incubated in darkness at 29°C for an additional 24 to 144 h. Coleoptiles were harvested during the 1- to 7-d incubation at 0.5- to 1-d intervals, frozen in liquid nitrogen, and stored at -80°C until all samples could be processed. The tips and central portions of some coleoptiles were fixed for low-temperature embedding for transmission electron microscopy.

Cell Wall Isolation

Isolation of cell walls from coleoptiles is essential to detect subtle changes in wall composition at 0.5-d intervals by IR spectroscopy. Cell walls were prepared from frozen maize coleoptiles by homogenization in 1% (w/v) SDS in 50 mM Tris-HCl, pH 7.2, at ambient temperature in a glass-glass motorized grinder (Kontes-Duall, Thomas Scientific). Cell walls were collected on nylon mesh filters (45 μm^2 ; Nitex), and washed sequentially with water, 50% (v/v) ethanol, acetone, and then resuspended in water.

FTIR Microspectroscopy

For microspectroscopy, materials were mounted in the wells of IR-reflective, gold-plated microscope slides (Thermo-Electron). The windows and slides with cell wall preparations were supported on the stage of a Nicolet Continuum series microscope accessory to a 670 IR spectrophotometer with a liquid nitrogen-cooled mercury-cadmium telluride detector (Thermo-Electron). An area of wall (up to 125 × 125 μm), excluding vascular walls, was selected for spectral collection in transreflectance mode. In transreflectance, the beam is transmitted through the wall sample, reflected off the gold-plated slide, and then transmitted through the sample a second time. One hundred and twenty eight interferograms were collected with 8 cm^{-1} resolution and coadded to improve the signal-to-noise ratio for each sample. Three spectra were collected from different areas of each sample and then area averaged and baseline corrected. The triplicate-averaged spectra from 36 to 60 hybrid or inbred coleoptiles were then averaged and used for digital subtraction.

Baseline-corrected and area-normalized data sets of spectra are then used in the chemometric analyses. Most of the PCA was carried out with WIN-DAS software (Kemsley, 1998). Multivariate PLS and some of the PCAs were carried out using Matlab 6.5.1 (The MathWorks). LDA was used to develop a discriminative calibration model to classify spectra into groups. The distances between each observation were estimated from group centers. Mahalanobis distance was used as the distance metric (Kemsley, 1998) to measure the distance of each observation (spectrum) from each group center. LDA using squared Mahalanobis distance metrics was applied to the PCA scores of original data (Kemsley, 1998). The derived quantities such as group centers and covariance matrices were calculated from the transformed observations, and the assignments to the respective class were then made. The correlation of predicted monosaccharide with actual values was carried out by LDA of PLS analysis, using *k*-fold cross validation, as described below.

ANNs

FTIR spectra were analyzed by genetic and Kohonen algorithms using the combination of NeuroShell2 and Classifier software (Ward Systems Group).

The algorithms are proprietary. Spectra were truncated to the range 800 to 1800 cm^{-1} , baseline corrected and area normalized, and input as PAT (pattern) files in wavenumber versus absorbance format with 250 variates. In each case, the generalization capabilities of the network were validated using k -fold cross validation; the data set was divided into k subsets, and k networks are trained and tested. Each time, one of the k subsets is used as the test set and the other $k-1$ subsets are pooled to form a training set. The average errors across all k trials were calculated.

For the genetic networks, 432 spectra for hybrid maize belonging to 12 classes (each representing one-half-d growth interval) were trained through 129 generations. Class membership is specified in the final column for each spectrum of the data set matrix. The genetic network is an acyclic feed-forward network using 250 spectral variates as input and 50 hidden layer neurons. For the Kohonen network, the class membership was not specified, and spectra were input to the network in random order. The learning rate for the Kohonen network was set at 0.5, with 0.5 initial weights, a neighborhood size of 11 neurons, and the network was trained for 50 epochs. The number of epochs is selected as the minimum number of epochs to achieve optimal classification success without overfitting. Euclidean distances were used to measure the distance between the classes of spectra. After obtaining the neuron values, probabilistic algorithms were used to obtain the probability and confusion matrices (Neuroshell2, Ward Systems Group, proprietary).

The ANNs produce two kinds of classification tables. First, a confusion matrix is a table of numbers of spectra assigned by the network to each class compared to their actual class identities. Second, the ANN calculates the probability of class membership for each spectrum. We have represented these two outputs in tables showing actual versus predicted class assignments that are color-coded across ranges of average class probability values overlaid with the numbers of spectra from the confusion matrix.

We also used the original values for each spectrum in the ANN probability matrix as a contingency data matrix to visualize the relationships between classes by correspondence analysis using XLSTAT v. 7.5.2 (Kovach Computing Services). The calculated eigenvalues were ranked in order of percent variance as F_1 to $F_{(n-1)}$ for n classes, and cluster plots were generated by plotting values of F_1 against F_2 .

Quantitation of Cellulose and Cell Wall Neutral Monosaccharides

Portions of the walls were hydrolyzed with 2 M trifluoroacetic acid (TFA) containing 400 nmol of *myo*-inositol for 90 min at 120°C in 1-mL conical Reactivials (Pierce Chemical). After hydrolysis, insoluble material was pelleted by centrifugation, and the supernatant TFA was collected and evaporated under a stream of filtered air. The insoluble material (mostly cellulose) was washed several times with water and collected by centrifugation. The pellet was suspended in 0.8 mL of water, and 100 μL was assayed for Glc equivalents by the phenol-sulfuric method (Dubois et al., 1956).

Monosaccharides in the TFA-soluble fraction were converted to alditol acetates (Gibeaut and Carpita, 1991), which were separated by gas-liquid chromatography on an SP-2330 vitreous silica capillary column (0.25 mm \times 30 m; Supelco). The oven temperature was held at 80°C for loading, then rapidly increased at 25°C min^{-1} to 170°C, and then programmed from 170°C to 240°C at 5°C min^{-1} with a 6-min hold at the upper temperature. The neutral sugar composition was verified by electron-impact mass spectrometry (Carpita and Shea, 1989).

Quantitation of (1 \rightarrow 3),(1 \rightarrow 4)- β -D-Glucan

Maize cell walls were incubated for 3 h at 37°C with a *Bacillus subtilis* (1 \rightarrow 3), (1 \rightarrow 4)- β -D-glucan endo-4-glucohydrolase (EC 3.2.1.73), which specifically digests (1 \rightarrow 3),(1 \rightarrow 4)- β -D-glucans (Anderson and Stone, 1975). The enzyme was purified from a commercial preparation from Novo Ban 120 (Novo Laboratories) by batch elution from DEAE-cellulose and gel permeation chromatography on Bio-Gel P-60 essentially as described by Kato and Nevins (1984) and stored frozen at -20°C. The enzyme is commercially available from Biosupplies Australia. Diagnostic cellodextrin-(1 \rightarrow 3)-Glc oligomers resulting from hydrolysis of the (1 \rightarrow 3),(1 \rightarrow 4)- β -D-glucan were separated by high performance anion-exchange liquid chromatography on a CarboPac PA-1 column (Dionex) and detected by pulsed amperometry (Buckeridge et al., 1999). Samples of 100 μL were injected onto the column equilibrated in 200 mM NaOH with a flow rate of 1.0 mL min^{-1} . The oligomers were eluted in a 20-min linear gradient of sodium acetate from 0 to 80 mM followed by a 10-min linear gradient

to 200 mM. The signal was calibrated using cellobiosyl- and cellotriosyl-(1 \rightarrow 3)- β -D-Glc oligomers hydrolyzed from known amounts of authentic barley (*Hordeum vulgare*) (1 \rightarrow 3),(1 \rightarrow 4)- β -D-glucan.

Transmission Electron Microscopy and Immunocytochemistry

Tissue blocks of about 2 mm³ were taken from the midsection of 2- to 5-d-old maize coleoptiles and fixed overnight in 2% glutaraldehyde in 0.1 M sodium cacodylate, pH 7.2, then low-temperature embedded, as described previously (Carpita et al., 2001). Thin sections (100 nm) were cut on a Leica ultramicrotome and picked up on carbon-coated and plastic-film gold grids. (1 \rightarrow 3),(1 \rightarrow 4)- β -D-Glucans were detected with a monoclonal antibody that recognizes an authentic barley (1 \rightarrow 3),(1 \rightarrow 4)- β -D-glucan (Meikle et al., 1994; available from Biosupplies Australia), visualized by a 10-nm colloidal-gold-conjugated rabbit anti-mouse secondary antibody (Sigma). The blocking solution was 1% acetylated bovine serum albumin (Aurion) in 5 mM sodium-potassium phosphate, pH 7.3, containing 3 mM KCl and 137 mM NaCl (phosphate-buffered saline) plus 0.1% (v/v) Tween 20, and incubation solution for washes was 0.1% (w/v) acetylated bovine serum albumin in phosphate-buffered saline plus 0.01% (v/v) Tween 20.

Supplemental Data

The following materials are available in the online version of this article.

Supplemental Figure S1. Prediction of monosaccharide composition of coleoptile cell walls from FTIR spectra.

ACKNOWLEDGMENT

We thank Dr. Chris Clifton, Department of Computer Sciences, Purdue University, for valuable discussions and his review of the manuscript.

Received November 15, 2006; accepted December 11, 2006; published January 12, 2007.

LITERATURE CITED

- Almeida JS (2002) Predictive non-linear modeling of complex data by artificial neural networks. *Curr Opin Biotechnol* **13**: 72–76
- Anderson MA, Stone BA (1975) A new substrate for investigating the specificity of β -glucan hydrolases. *FEBS Lett* **52**: 202–207
- Buckeridge MS, Vergara CE, Carpita NC (1999) The mechanism of synthesis of a mixed-linkage (1 \rightarrow 3),(1 \rightarrow 4)- β -D-glucan in maize: evidence for multiple sites of glycosyl transfer in the synthase complex. *Plant Physiol* **120**: 1105–1116
- Caño-Delgado A, Penfield S, Smith C, Catley M, Bevan M (2003) Reduced cellulose synthesis invokes lignification and defense responses in *Arabidopsis thaliana*. *Plant J* **34**: 351–362
- Carpita NC (1996) Structure and biogenesis of the cell walls of grasses. *Annu Rev Plant Physiol Plant Mol Biol* **47**: 445–476
- Carpita NC, Defernez M, Findlay K, Wells B, Shoue DA, Catchpole G, Wilson RH, McCann MC (2001) Cell wall architecture of the elongating maize coleoptile. *Plant Physiol* **127**: 551–565
- Carpita NC, Gibeaut DM (1993) Structural models of primary cell walls in flowering plants: consistency of molecular structure with the physical properties of the walls during growth. *Plant J* **3**: 1–30
- Carpita NC, Shea EM (1989) Linkage structure by gas chromatography-mass spectrometry of partially-methylated alditol acetates. In CJ Biermann, GD McGinnis, eds, *Analysis of Carbohydrates by GLC and MS*. CRC Press, Boca Raton, FL, pp 155–216
- Chen CH (1996) *Fuzzy Logic and Neural Network Handbook*. McGraw-Hill, New York
- Chen L, Wilson RH, McCann MC (1997) Infrared microspectroscopy of hydrated biological systems: design and construction of a new cell with atmospheric control for the study of plant cell walls. *J Microsc* **188**: 62–71
- Chen L-M, Carpita NC, Reiter W-D, Wilson RW, Jeffries C, McCann MC (1998) A rapid method to screen for cell wall mutants using discriminant analysis of Fourier transform infrared spectra. *Plant J* **8**: 375–382

- Darvill AG, Smith CJ, Hall MA (1978) Cell wall structure and elongation growth in *Zea mays* coleoptile tissue. *New Phytol* **80**: 503–516
- Dubois M, Gilles KA, Hamilton JK, Rebers PA, Smith F (1956) Colorimetric method for the determination of sugars and related substances. *Anal Chem* **28**: 350–356
- Encina A, Sevillano JM, Acebes JL, Alvarez J (2002) Cell wall modifications of bean (*Phaseolus vulgaris*) cell suspensions during habituation and dehabituation to dichlobenil. *Physiol Plant* **114**: 182–191
- Faix O (1992) Fourier transform infrared spectroscopy. In SY Lin, CW Dence, eds, *Methods in Lignin Chemistry*. Springer, Berlin, pp 83–109
- Gibeaut DM, Carpita NC (1991) Tracing the biosynthesis of the cell wall in intact cells and plants: selective turnover and alteration of cytoplasmic and cell wall polysaccharides of proso millet cells in liquid culture and *Zea mays* seedlings. *Plant Physiol* **97**: 551–561
- Goldberg DE (1989) *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addison-Wesley, Reading, MA
- His I, Drouich A, Nicol F, Jauneau A, Höfte H (2001) Altered pectin composition in primary cell walls of *korrigan*, a dwarf mutant of *Arabidopsis* deficient in a membrane-bound endo-1,4- β -glucanase. *Planta* **212**: 348–358
- Ideker T, Thorsson V, Ranish JA, Christmas R, Buhler J, Eng JK, Bumgarner R, Goodlett DR, Aebersold R, Hood L (2001) Integrated genomic and proteomic analyses of a systematically perturbed metabolic network. *Science* **292**: 929–934
- Iraki NM, Bressan RA, Hasegawa PM, Carpita NC (1989) Alteration of the physical and chemical structure of the primary cell wall of growth-limited plant cells adapted to osmotic stress. *Plant Physiol* **91**: 39–47
- Jobson JD (1992) *Applied Multivariate Data Analysis*, Vol 2. *Categorical and Multivariate Methods*. Springer-Verlag, New York, pp 443–462
- Kato Y, Nevins DJ (1984) Enzymic dissociation of *Zea* shoot cell wall polysaccharides. II. Dissociation of (1 \rightarrow 3),(1 \rightarrow 4)- β -D-glucan by purified (1 \rightarrow 3),(1 \rightarrow 4)- β -D-glucan 4-glucanohydrolase from *Bacillus subtilis*. *Plant Physiol* **75**: 745–752
- Kell DB, Darby RM, Draper J (2001) Genomic computing: explanatory analysis of plant expression profiling data using machine learning. *Plant Physiol* **126**: 943–951
- Kemsley EK (1998) *Discriminant Analysis of Spectroscopic Data*. John Wiley and Sons, Chichester, UK
- Kim J-B, Olek AT, Carpita NC (2000) Cell wall and membrane-associated exo- β -D-glucanases from developing maize seedlings. *Plant Physiol* **123**: 471–485
- Labavitch JM, Ray PM (1978) Structure of hemicellulosic polysaccharides of *Avena sativa* coleoptile cell walls. *Phytochemistry* **17**: 933–937
- Lavine BK, Davidson CE, Westover DJ (2004) Spectral pattern recognition using self-organizing MAPS. *J Chem Inf Comput Sci* **44**: 1056–1064
- Li YH, Qian O, Zhou YH, Yan MX, Sun L, Zhang M, Fu ZM, Wang YH, Han B, Pang XM, et al (2002) *BRITTLE CULM1*, which encodes a COBRA-like protein, affects the mechanical properties of rice plants. *Plant Cell* **15**: 2020–2031
- Liang CY, Marchessault RH (1959) Infrared spectra of crystalline polysaccharides. II. Native celluloses in the region from 640 to 1700 cm^{-1} . *J Polym Sci [B]* **39**: 269–278
- Madson M, Dunand C, Li XM, Verma R, Vanzin GF, Calplan J, Shoue DA, Carpita NC, Reiter W-D (2003) The *MUR3* gene of *Arabidopsis* encodes a xyloglucan galactosyltransferase that is evolutionarily related to animal exostosins. *Plant Cell* **15**: 1662–1670
- Marita JM, Vermerris W, Ralph J, Hatfield RD (2003) Variations in the cell wall composition of maize brown midrib mutants. *J Agric Food Chem* **51**: 1313–1321
- McCann MC, Bush M, Milioni D, Sado P, Stacey NJ, Catchpole G, Defernez M, Carpita NC, Höfte H, Wilson RH, et al (2001) Approaches to understanding the functional architecture of the plant cell wall. *Phytochemistry* **57**: 811–821
- McCann MC, Carpita NC (2005) Looking for invisible phenotypes in cell wall mutants of *Arabidopsis thaliana*. *Plant Biosyst* **139**: 80–83
- McCann MC, Chen L, Roberts K, Kemsley EK, Séné C, Carpita NC, Stacey NJ, Wilson RH (1997) Infrared microspectroscopy: sampling heterogeneity in plant cell wall composition and architecture. *Physiol Plant* **100**: 729–738
- McCann MC, Hammouri M, Wilson R, Belton P, Roberts K (1992) Fourier transform infrared microspectroscopy is a new way to look at plant cell walls. *Plant Physiol* **100**: 1940–1947
- McCann MC, Roberts K (1991) Architecture of the primary cell wall. In CW Lloyd, ed, *The Cytoskeletal Basis of Plant Growth and Form*. Academic Press, London, pp 109–129
- McCann MC, Shi J, Roberts K, Carpita NC (1994) Changes in pectin structure and localisation during the growth of unadapted and NaCl-adapted tobacco cells. *Plant J* **5**: 773–785
- Meikle PJ, Hoogenraad NJ, Bonig I, Clarke AE, Stone BA (1994) A (1 \rightarrow 3),(1 \rightarrow 4)- β -D-glucan-specific monoclonal antibody and its use in the quantitation and immunocytochemical location of (1 \rightarrow 3),(1 \rightarrow 4)- β -D-glucans. *Plant J* **5**: 1–5
- Misra J, Schmitt W, Hwang D, Hsiao L-L, Gullans S, Stephanopoulos G, Stephanopoulos G (2002) Interactive exploration of microarray gene expression patterns in a reduced dimensional space. *Genome Res* **12**: 1112–1120
- Mouille G, Robin S, Lecomte M, Pagant S, Höfte H (2003) Classification and identification of *Arabidopsis* cell wall mutants using Fourier-transform infrared (FT-IR) microspectroscopy. *Plant J* **35**: 393–404
- Séné CFB, McCann MC, Wilson RH, Grinter R (1994) Fourier-transform Raman and Fourier-transform infrared spectroscopy: an investigation of five higher plant cell walls and their components. *Plant Physiol* **106**: 1623–1630
- Shedletzky E, Shmuel M, Trainin T, Kalman S, Delmer D (1992) Cell wall structure in cells adapted to growth on the cellulose synthesis inhibitor 2,6-dichlorobenzonitrile: a comparison between two dicotyledonous plants and a gramineous monocot. *Plant Physiol* **100**: 120–130
- Smith BG, Harris PJ (1999) The polysaccharide composition of Poales cell walls: Poaceae cell walls are not unique. *Biochem Syst Ecol* **27**: 33–53
- Somerville C, Bauser S, Brinistool G, Facette M, Hamann T, Milne J, Osborne E, Paredes A, Persson S, Raab T, et al (2004) Toward a systems approach to understanding plant-cell walls. *Science* **306**: 2206–2211
- Tsuiji M (1957) Infrared spectrum and crystal structure of cellulose. *J Polym Sci [B]* **25**: 159–171
- Wells B, McCann MC, Shedletzky E, Delmer D, Roberts K (1994) Structural features of cell walls from tomato cells adapted to grow on the herbicide 2,6-dichlorobenzonitrile. *J Microsc* **173**: 155–164