Carbohydrate microarrays and their use for the identification of molecular markers for plant cell wall composition

Ian P. Wood[a], Bruce M. Pearson[a], Enriqueta Garcia-Gutierrez[a], Lenka Havlickova[b], Zhesi He[b], Andrea L. Harper[b], Ian Bancroft[b], and Keith W. Waldron[a,1]

[a]Quadram Institute Bioscience, Norwich Research Park, Colney, Norwich NR4 7UA, United Kingdom; and [b]Department of Biology, University of York, Heslington, YO10 5DD, United Kingdom

Edited by Chris R. Somerville, University of California, Berkeley, CA, and approved May 16, 2017 (received for review November 29, 2016)

Genetic improvement of the plant cell wall has enormous potential to increase the quality of food, fibers, and fuels. However, the identification and characterization of genes involved in plant cell wall synthesis is far from complete. Association mapping is one of the few techniques that can help identify candidate genes without relying on our currently incomplete knowledge of cell wall synthesis. However, few cell wall phenotyping methodologies have proven sufficiently precise, robust, or scalable for association mapping to be conducted for specific cell wall polymers. Here, we created high-density carbohydrate microarrays containing chemically extracted cell wall polysaccharides collected from 331 genetically diverse *Brassica napus* cultivars and used them to obtain detailed, quantitative information describing the relative abundance of selected noncellulosic polysaccharide linkages and primary structures. We undertook genome-wide association analysis of data collected from 57 carbohydrate microarrays and identified molecular markers reflecting a diversity of specific xylan, xyloglucan, pectin, and arabinogalactan moieties. These datasets provide a detailed insight into the natural variations in cell wall carbohydrate moieties between *B. napus* genotypes and identify associated markers that could be exploited by marker-assisted breeding. The identified markers also have value beyond *B. napus* for functional genomics, facilitated by the close genetic relatedness to the model plant *Arabidopsis*. Together, our findings provide a unique dissection of the genetic architecture that underpins plant cell wall biosynthesis and restructuring.

association mapping | biomass | functional genomics | GWAS | lignocellulose

Plant cell wall composition and structure underlie plant form and function and determine the nature and quality of numerous plant-derived products. Differences in cell wall composition influence agronomic traits, such as resistance to pests and diseases, and many of the quality characteristics of biologically derived products such as natural fibers, timber, and food. Cell walls have wide-reaching impacts on human activities from soil conditioning to human health. Emergent technologies such as lignocellulose-derived biofuels and bio-based polymers could also be made more feasible through modifications in cell wall design (1). Consequently, the plant cell wall continues to be a major target for biotechnological improvement.

Marker-assisted selection is a direct and cost-effective way to introduce novel, heritable agronomic improvements into crop varieties and could therefore be used for plant cell wall improvement. Molecular markers provide the genetic tools needed for implementation, enable potentially exploitable genes to be located, and help elucidate the genetic mechanisms involved. Recent improvements in genome sequencing have enabled genome characterization, identification of structural rearrangements, and high-density linkage maps to be constructed for crop species, such as *Brassica napus*, which have large and complex genomes (2–4). Consequently, genome-wide association studies (GWAS) can now be conducted in crop species (5).

Brassica napus is a good choice for functional genomics, benefiting from its well-studied genetics and its relatedness to the model plant *Arabidopsis* (6). As an allotetraploid species with genomes inherited from two, closely related, ancestral species: *Brassica rapa* and *Brassica oleracea* (which contribute to the A and C genome portions, respectively), methods needed to identify single nucleotide polymorphisms (SNPs) in orthologous regions within the ancestral genomes have been developed (3, 7) and are continually improved (4). *B. napus* was also recently used to develop associative transcriptomics where sequence variation, transcript abundance, and phenotype are correlated (7, 8). The resulting "gene expression markers" (GEMs) have the potential to reveal additional layers of genetic detail, beyond that of traditional GWAS (7).

Accurate positioning and identification of tightly linked and robust markers is essential for gene candidate identification. Tightly linked markers are particularly important for plant cell wall-related traits, where potential candidates are common (>10% of the genome) (9). Precise phenotyping methods are also needed to prevent the dispersion of genetic signals among too many loci.

However, suitably rapid and accurate phenotyping techniques have, until recently, been beyond the reach of cell wall chemists (9, 10). GWAS typically require phenotype data to be collected from hundreds or thousands of individuals, which can be difficult to achieve by using conventional analytical approaches. Obtaining replication without compromising phenotyping specificity is therefore a formidable challenge. The inherent problems in obtaining necessary replication and comparative analytical data

Significance

Plant cell wall (PCW) composition determines the nature and quality of many biologically derived products and, therefore, is a major target for genetic improvement. However, the identities and functions of many genes involved in PCW synthesis are still not known. Genome-wide association mapping studies (GWAS) are one of the few ways to identify these genes. However, collecting precise and quantitative PCW phenotype data at the scale required for GWAS is a significant challenge. Here, we demonstrate that high-density carbohydrate microarrays can be used as a PCW phenotyping strategy suitable for GWAS. Results presented in this study will aid in the understanding of PCW genetics and crop breeding for improved PCW composition.

Author contributions: I.P.W., I.B., and K.W.W. designed research; I.P.W. and E.G.-G. performed research; B.M.P., L.H., Z.H., A.L.H., and I.B. contributed new reagents/analytic tools; I.P.W., L.H., Z.H., A.L.H., I.B., and K.W.W. analyzed data; and I.P.W. and K.W.W. wrote the paper.

Conflict of interest statement: A background patent has been filed.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

[1]To whom correspondence should be addressed. Email: keith.waldron@quadram.ac.uk.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1619033114/-/DCSupplemental.

across thousands of samples have led cell wall researchers to either (*i*) conduct deliberately underpowered GWAS that reveal loci controlling a large proportion of the trait (10), or (*ii*) focus on more tractable, but less precise, phenotypes to achieve sufficient replication. Examples include the gravimetric determination of bulk polymer fractions (11, 12), indirect determination of composition using pyrolysis (13), or focusing on a particular product, such as sugar release following saccharification (14, 15). These mapping project successes illustrate the clear potential of GWAS to elucidate relevant genetic targets, but also highlight the requirement for greater phenotyping specificity or increased replication (9).

Some high-throughput cell wall phenotyping approaches such as Fourier transform infra-red screening (16, 17), OLIgo Mass Profiling (18), and carbohydrate microarray-based "CoMPP" profiling (19, 20) are now available, but have yet to be deployed with the scale and precision required for association mapping. Hence, these methods have generally been used for intervention-based or comparative studies, which do not require the same degree of quantification or replication. Recently, Lin et al. (21) used a range of high-throughput cell wall profiling methodologies to collect cell wall phenotype data from 30 samples at various stages of development and correlated these data against the transcript abundance of 67 putative cell wall genes. This information will help assign functions to genes known to be involved in cell wall synthesis. GWAS have the potential to take this further to identify novel candidates in addition to those that are already known. However, before the present study, high-throughput phenotyping technologies have not been sufficiently precise and integrated to collect large-scale datasets suitable for GWAS without considerable effort.

Here, we developed high-density carbohydrate microarrays suitable for association mapping (Fig. 1*A*). Particular care was taken to produce precise, quantitative, and comparable data while minimizing the accumulation of technical errors that might hinder precision. Using these data, we aimed to elucidate areas of the transcriptome that are most closely associated with variation in cell wall polymer abundances in mature stem tissue of *B. napus* using associative transcriptomics (4, 7, 8, 22).

We chemically extracted four carbohydrate-rich fractions from 331 diverse *B. napus* genotypes in duplicate, following common plant cell wall extraction regimes (23) adapted for screening, and printed them as carbohydrate microarrays. Although chemical extraction tends to release fractions rich in certain carbohydrates, the extracts are not pure (23). Rather, each fraction contains a mixture of carbohydrate moieties derived from various polymer classes, which vary in solubility depending on their chemistry and interaction with other components (23). Chemical or enzymatic extraction methodologies and downstream chromatographic procedures could be selected to isolate polymers more precisely or less destructively than those used here.

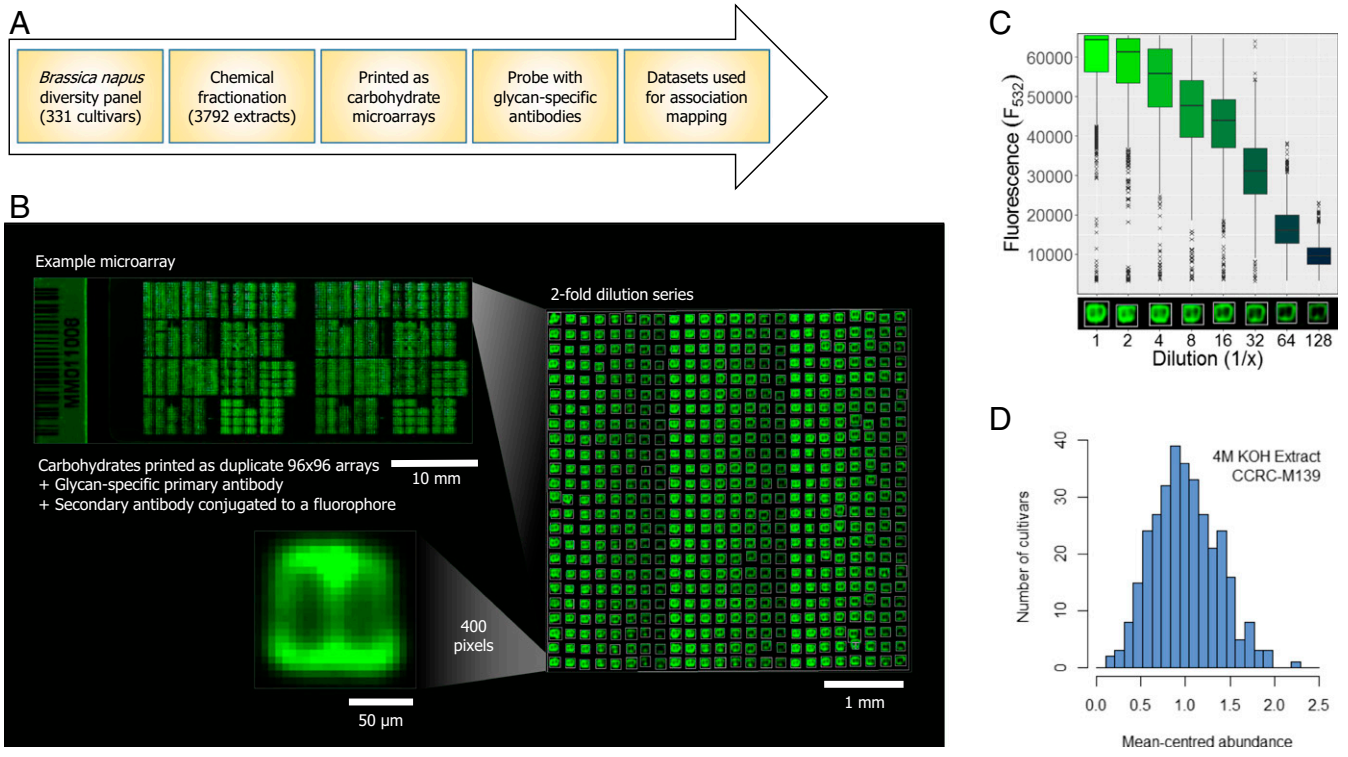

Fig. 1. The use of high-density carbohydrate microarrays as a phenotyping method for association mapping. (*A*) An overview of phenotyping strategy: We collected stem tissue from 475 accessions, derived from 331 cultivars, and sequentially fractionated them in duplicate, releasing four glycan-rich fractions (approximately 3,800 extracts in total). (*B*) Microarray layouts: Twofold dilution series ($n = 8$) of each cell wall extract were printed together, as duplicate 96 × 96 square arrays consisting of approximately 100-µm-square features. Each slide contained approximately 15,500 features extracted by using the same conditions. We probed each extract by using 19 glycan-specific antibodies (76 slides in total, images: *SI Appendix*, Figs. S1–S4). (*C*) Example of fluorescence measurements collected from a single slide, grouped by dilution: Boxes indicate the interquartile range (IQR), and the central line indicates the median fluorescence for these features. Whiskers and points indicate data 1.5 x IQR inside or outside of the quartiles, respectively. The boxplots show that no single dilution, even in the middle of the range (diluted 32 times), would bring all fluorescence measurements of all samples into the linear range. Therefore, to obtain accurate quantification for all samples, we used nonlinear regression to fit best-fit lines to each dilution series. From these lines, we interpolated the dilution that would be required to produce a feature with a set fluorescence within the linear range. The inverse of the dilution was then used to derive a nominal relative abundance of each epitope relative to the original sample weight. (*D*) Relative abundance measurements taken from the same genotype were averaged, and their abundance was expressed relative to the mean cultivar: An example histogram showing the distribution in mean-centered relative abundance of a single epitope between 331 cultivars is shown here. In total, relative quantification was obtained for 57 slides (*SI Appendix*, Figs. S5–S8), amassing data collected from more than 1 million data points. These data were used for association mapping.

Here, we used ammonium oxalate to mainly release polysaccharides bound by metal ions. Sodium carbonate was then used to deesterify cell wall components, releasing mainly pectins held by weak-ester linkages and to stabilize more sensitive polysaccharides to β-eliminative degradation (24). Further extraction with 1 and 4 M KOH was used to solubilize predominantly xylans. Using half-gram portions of cell wall material ensured a reliable datum to which all samples were comparable and minimized sample heterogeneity (25). Plate-based liquid handling robotics for soluble extracts minimized technical errors. To obtain a high-throughput quantitative measurement of selected carbohydrate moieties, extracts were printed as high-density glycan microarrays, containing a dilution series of each extract (Fig. 1*B*). These arrays were probed with monoclonal antibodies (mAbs) raised to different cell wall components (26, 27) (*SI Appendix*, Fig. S1–S4). Labeling with a secondary antibody conjugated to a fluorophore enabled detection.

Most carbohydrate microarrays use chromogenic methods of detection, probing samples that have been diluted to obtain an unsaturated feature (19). Here, we used a fluorescent probe for detection, which increased sensitivity and specificity, but meant that no single dilution could bring all measurements of all samples into the linear range (Fig. 1*C*). We therefore printed two, eight-point dilution series of each extract and modeled the concentration-related change in fluorescence by using nonlinear regression. From the resulting best-fit lines, we used a threshold method of quantification analogous to that of quantitative PCR to interpolate the dilution required to obtain a feature with a set fluorescence in the linear range. The inverse of the dilution was used to derive a nominal relative abundance of each epitope relative to the original sample weight for each extract. The relative abundance measurements collected from each extract were averaged for each cultivar and expressed relative to an average cultivar (*SI Appendix*, Figs. S5–S8, example: Fig. 1*D*). Methodological improvements, such as noncontact printing of arrays, might improve precision further. Less variable data might also be obtained by using liquid-based ELISA methods, but obtaining these data would be logistically challenging and more costly to achieve.

Here, microarray-based abundance measurements, suitable for association mapping, were made for 57 slides (Example: Fig. 1*C*, images: *SI Appendix*, Figs. S5–S8, data: *SI Appendix*, Table S1). Antibodies raised to common cell wall polysaccharides (including xylan, xyloglucan, homogalacturonan, rhamnogalacturonan, or arabinogalactan) allowed the relative quantification of specific epitopes, by exploiting their varying binding specificities (23, 27). With more than 200 carbohydrate-directed mAbs (26, 27) and multiple cell wall extraction methods to choose from, the epitopes targeted in this study are only a small selection of those present in the wall. The aim of this work was not to provide an exhaustive overview of cell wall moieties but to assess the suitability of high-throughput cell wall glycomics as a phenotyping strategy for association mapping—targeting a selection of linkages likely to differ in their extractability, specificity, and underlying genetics.

To identify arrays likely to show similar marker associations after mapping, we correlated the mean relative abundance data collected from each array and ordered them by hierarchical clustering (Fig. 2). The array data broadly clustered into mAbs that recognized either XG or xylan moieties, the main hemicelluloses found in dicot primary and secondary cell walls (28) (Fig. 2). Further clustering of the datasets, within these groups, conformed to subtle differences related to polymer solubility or mAb specificity. For example, cultivars with stems rich in one xylan epitope also tended to contain more of other xylan epitopes, but those extracted using 1M KOH could generally be distinguished from those released by 4 M KOH (Fig. 2). Conversely, some mAbs such as CCRC-M78 and CCRC-M92, which recognize Arabinogalactan-4 moieties (23, 27), showed similar genotypic differences in

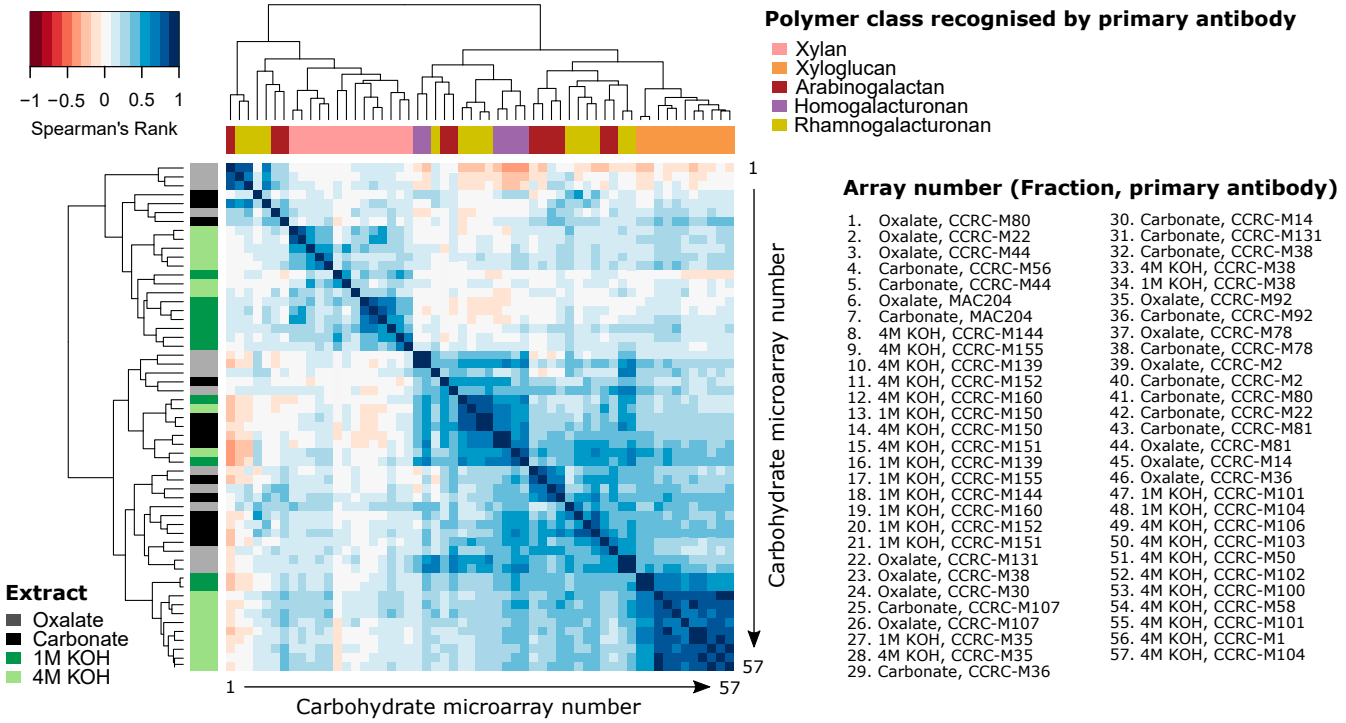

Fig. 2. Correlation matrix comparing the relative abundance measurements for 330 different *B. napus* cultivars collected from 56 arrays. The heat-map displays pair-wise Spearman's rank correlation coefficients between data collected from each array (key located in the top left corner), ordered by hierarchical clustering by the complete linkage method (dendrograms). For legibility, array identities including the extract and primary antibody (1–57) are listed on the right side. The left color key depicts the chemical used for extraction and the upper color key depicts the general polymer class commonly associated with each primary antibody following Pattathil et al. (23, 27). Please note, the binding specificities of some mAb are not exclusive to a single polymer class. Correlated epitopes are more likely to have a common genetic basis and are likely to produce more similar GWA profiles after mapping.

relative abundance, irrespective of the extract probed. Therefore, one might expect data from these arrays to produce similar patterns of marker association with a common genetic basis. Similar inferences may be made for other epitopes in different extracts, which are likely to share similar genetics (Fig. 2).

High-density carbohydrate microarrays permitted us to obtain detailed and biologically relevant data pertaining to the relative abundance of complementary and contrasting cell wall epitopes, in parallel, at the scale required for GWAS. To assess the suitability of these datasets for association mapping, we mapped all 57 datasets using associative transcriptomics (4). Using the unique ability for associative transcriptomics to identify molecular markers based on both SNP-based variants and transcript abundance (GEMs), we also identified putative transcription-based regulators of cell wall composition (*SI Appendix*, Fig. S9–S65).

As proof of concept, we illustrate the deployment of the method to dissect the genetic basis for xylan synthesis and branching, which is already largely understood (29–37), by exploiting the differing binding specificities of two mAbs, CCRC-M139 and CCRC-M150 (Fig. 3).

Dicot-derived xylan can be broadly divided into two classes defined by the most common substitution patterns (36). "Major domain" xylan is decorated with glucuronic acid (GlcA) residues distributed sparsely but evenly along the backbone (approximately every 8–10 Xyl residues) (37). "Minor domain" xylan is decorated with GlcA more frequently and unevenly (typically every five residues), theoretically altering its interactions with other components (36).

Antibodies that adhere to both xylan domains, irrespective of branching patterns, should reveal variations in overall xylan abundance (Fig. 3*A*). In contrast, the abundance of major domain xylan could be indirectly quantified by using mAbs that exclusively recognize sparsely substituted sections, thereby implicating the genes that produce these patterns (Fig. 3*A*).

Although the exact epitopes recognized by many carbohydrate-directed mAbs are not known, CCRC-M150 has been shown to bind to unbranched xylan oligosaccharides >4 Xyl residues in length, or arabinoxylan oligosaccharides >3 Xyl residues in length irrespective of substitution patterns (38). Therefore, CCRC-M150 should identify markers in areas of the transcriptome common to both major and minor xylan domains, for example those involved in xylan backbone synthesis. By contrast, CCRC-M139 binds specifically to longer unsubstituted Xyl oligomers >6 Xyl residues in length (Fig. 3*A*). Binding can be prevented by arabinose substitutions (38). CCRC-M139 should therefore preferentially bind to major domain xylan domains, implicating genes that alter the

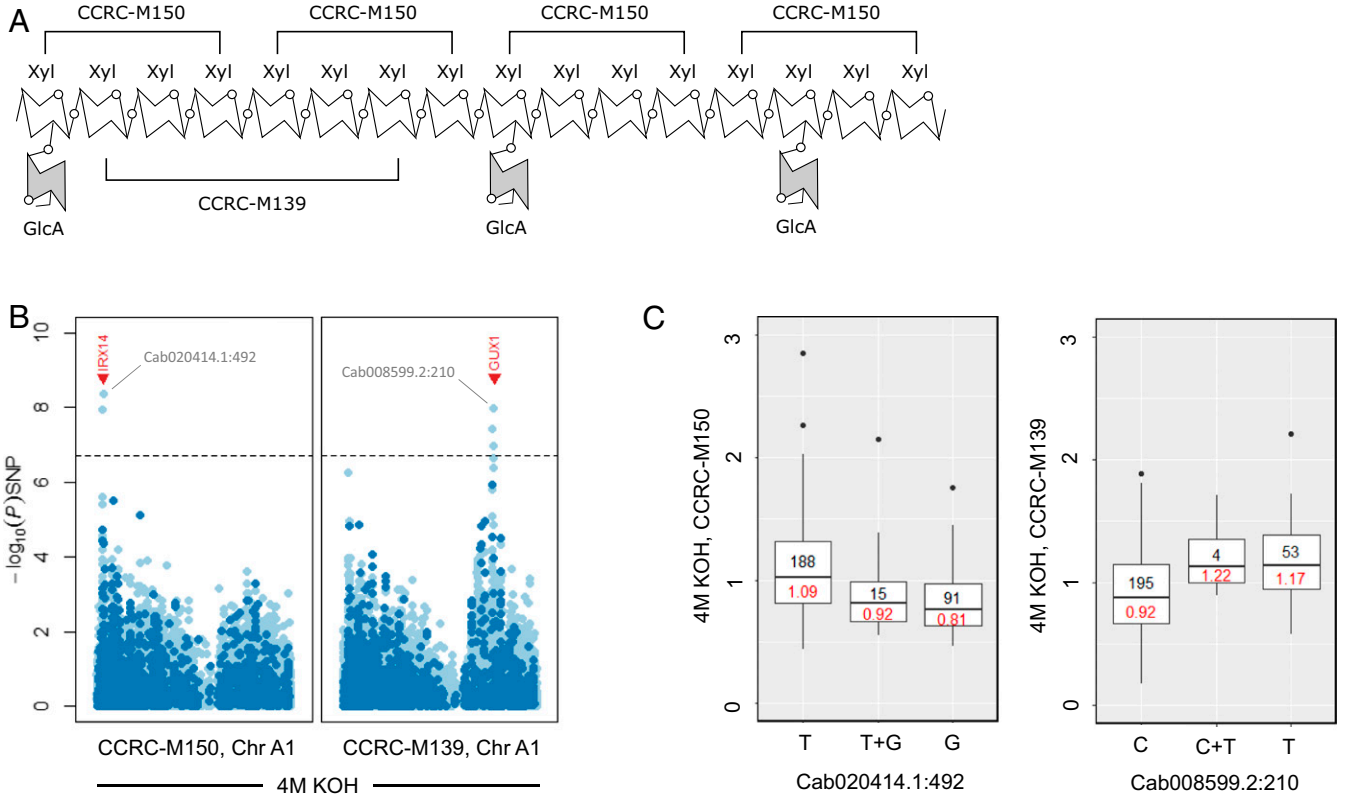

Fig. 3. As a proof of concept, we mapped the relative abundance of two epitopes recognized by two xylan-specific mAbs, CCRC-M150 and CCRC-M139, in a xylan-rich extract (4 M KOH). (*A*) Likely differences in antibody binding to xylan: Although the precise binding specificities of these mAbs are not known, CCRC-M150 is thought to bind to most xylan structures, whereas CCRC-M139 preferentially binds to long unsubstituted sections >6 Xyl residues in length, at least for arabinoxylan oligomers known to differ in composition (38). (*B*) Manhattan plots, which show the degree of association (*y* axis) of all SNPs located on chromosome A1 (*x* axis) with the genotypic variation in CCRC-M150 or CCRC-M139 binding to the 4 M KOH extracts: SNPs with a higher $\log_{10}$(P) value show a greater statistical association with each trait. Those above the Bonferroni threshold (gray dotted line) pass a conservative level of genome-wide statistical significance ($P = 0.05/256598$ markers). Hemi-SNPs and Simple-SNPs (not assigned/assigned to a specific ancestral genome, respectively) are colored light and dark blue, respectively. SNPs associated with genome-wide significance to the variations in CCRC-M150 and CCRC-M139 binding to the 4 M KOH extracts, coincide with regions harboring orthologs of *Arabidopsis* IRX14 and GUX1, respectively (*x* positions shown in red). These marker associations are consistent with the respective substitution patterns that these mAbs are thought to recognize. The most highly associated markers on chromosome A1 have been indicated for both traits (gray text). (*C*) Boxplots showing variations in relative mAb binding to the 4 M KOH extracts, grouped by the alleles at these significantly associated SNP loci: The number of cultivars tested with a particular allele and the mean relative abundance of each epitope are written in each box in black and red, respectively. For example, most cultivars ($n = 188$) had a T allele at SNP "Cab020414.1:492." Cultivars with a G allele at this position ($n = 91$) tended to contain less of the epitope recognized by CCRC-M150 in the 4 M KOH extract.

relative abundance of these long unbranched sections when mapped. We therefore mapped the relative abundance of epitopes recognized by CCRC-M150 and CCRC-M139 in the 4 M KOH fractions, which are likely to contain a greater concentration of major domain xylan, because higher alkali concentrations are required to disrupt hydrogen bonds formed with other components (37).

For CCRC-M150, we identified two hemi-SNP markers of genome-wide significance (Bonferroni-corrected, $P < 0.0000001$, *SI Appendix*, Table S2) located on Chromosome (Chr.) A1, flanking a *B. napus* ortholog of IRREGULAR XYLEM 14 (IRX14) (Chr. A1: Fig. 3*B*, *Left*. Full plot: *SI Appendix*, Fig. S61), shown to be important for xylan backbone synthesis in many species (29–35). Variations in the most highly associated marker (Cab020414.1:492) revealed a common SNP variant (91/294 cultivars) that coincided with an approximately 25% reduction in CCRC-M150 binding (Fig. 3*C*, *Left*). Selection for the allele associated with lower xylan abundance could be used to reduce xylan content in stem tissue by marker-assisted breeding. Potential industrial benefits for xylan reduction include the tailoring of biomass to produce specific products or reduction of recalcitrance for biorefining (1).

In contrast, we identified seven SNPs of genome-wide significance associated with CCRC-M139 binding (*SI Appendix*, Table S2). Despite probing identical extracts, these markers were located in different parts of the transcriptome compared with those associated with CCRC-M150 binding (Chr. A1: Fig. 3*B*, *Right*. Full plot: *SI Appendix*, Fig. S57). All of these markers were located near to *B. napus* orthologs of GLUCURONIC ACID SUBSTITUTION OF XYLAN 1 (GUX1) located at homeologous positions on Chr. A1 and C1 (Cab008614.1 and Bo1g116390.1, *SI Appendix*, Fig. S57). In *Arabidopsis*, GUX1 exclusively decorates glucuronoxylan with GlcA every 8–10 Xyl residues, to produce sparsely branched sections of major domain xylan that cannot receive further GlcA substitutions (36). The positions of these markers are therefore consistent with the specificity of CCRC-M139, which must bind to long unsubstituted sections determined by GUX1. Specific alleles at these loci, such as a T allele at Cab008599.2:210, could be used to select germplasm with a greater proportion of sparsely branched xylan (Fig. 3*C*, *Right*). Reducing xylan branching may be favorable to modify extractability during industrial processing, reducing polymer heterogeneity (39), and influencing the interactions between xylan and other cell wall components (37).

The coincidence of markers surrounding *B. napus* orthologs of IRX14 and GUX1 are not only consistent with known specificities of the mAbs used, but also implicate the main genetic players in this process. Overall, this example demonstrates that high-density quantitative glycan microarrays, used in conjunction with association mapping, can detect pertinent variations related to plant cell wall genetics.

Additional associations provide insights into the genetic architecture underpinning cell wall composition of other polymer linkages, with less well known genetics (*SI Appendix*, Figs. S9–S65), which are influenced by a combination of genetic and methodological factors distinct to each array. Potential genetic factors include the number and relative contributions of different loci contributing to a single trait. For example, unlike the example given above, which implicates genes located at few precise loci, a large number of genes are involved in the synthesis and remodeling of xyloglucan (40). The resulting marker associations for these traits are therefore less distinct and dispersed across more loci (*SI Appendix*, Figs. S37–S39 and S48–S56).

Methodological differences such as tissue and cell-type variations within each sample, the chemicals used for extraction, the exact specificity of each mAb, and accuracy of quantification also alter genetic signals to varying degrees. For example, obtaining precise quantification for mAbs that recognize particularly rare linkages can be difficult. This potential issue can be seen for CCRC-M151 binding to the 4 M KOH fraction; although it has a similar binding specificity to CCRC-M139, only one tail of the distribution could be accurately quantified above the limit of detection (*SI Appendix*, Fig. S8).

Nevertheless, for many epitope-based traits, marker associations clearly implicate regions of the transcriptome harboring novel gene candidates worthy of further investigation that will form the basis of future studies. Particularly interesting examples include those related to MAC204 binding to oxalate extracts (*SI Appendix*, Fig. S18), which show SNP marker associations at homeologous positions on Chr. A3 and C3. These, together with a single associated GEM marker located on Chr A3 (A03_005776413_005778886), strongly implicate Cab016526.3, a *B. napus* ortholog of β-Gal 4 (AT5G56870), as the most likely candidate gene.

Similarly, all slides probed with mAbs that identify Group-4 Arabinogalactans showed increases in GEM associations surrounding Bo3g004890.1, a *B. napus* ortholog of α-GALACTOSIDASE 1 (AT5G08380) on Chr. C3, despite probing different extracts (oxalate and carbonate) with two different mAbs (CCRC-M78 and CCRC-M92) (*SI Appendix*, Figs. S21, S22, S35, and S36). These candidate genes are highly likely to alter the abundance and composition of arabinogalactan moieties in mature dicot cell walls and illustrate the deployment of this method to identify novel gene candidates involved in cell wall biosynthesis.

Molecular markers highlighted in this study will provide considerable assistance to those attempting to identify the roles of particular genes by intervention-based approaches, enhanced by the close relatedness between the model plant *Arabidopsis* and *Brassica* species (6). Marker-assisted breeding could be used to generate germplasm with altered cell wall chemistry, using the molecular markers we have identified. Similar GWAS could be conducted for other abundant or commercially important cell wall types, such as those in edible tissues, or those with different cell wall architecture, such as monocotyledonous plants, or other organisms with glycan-rich cell walls. More generally, quantitative high-density carbohydrate microarrays will prove a useful phenotyping strategy, providing data suitable for use in association mapping and greatly expanding our knowledge of cell wall genetics.

Materials and Methods

Materials. All *B. napus* straw accessions were grown in 5-L plant pots at randomized locations in a 30 m × 7 m polytunnel located at the University of Nottingham, Sutton Bonington Campus, Leicestershire (+52° 49′ 56.65″, −1° 14′ 57.73″) (41). Stems were left to dry in their pots, and basal stem material from 2 cm above the hypocotyl to approximately 60 cm above the ground was collected from each accession at maturity (August 20–22, 2014). Stem material from individual plants were then milled by using a Cyclone Mill TWISTER (Retch) fitted with a 0.5-mm screen to produce fine, uniform powders. After milling, approximately 84, 61, and 24% of the sample by weight was able to pass a 250-, 150- and 53-μm sieve, respectively. In total, 475 stem accessions derived from 330 different cultivars were analyzed in duplicate. Milled straw (approximately 3.5 g) from each accession was rinsed in 50 mL of ethanol [≥99.8%, 1 h, room temperature (RT)], followed by acetone (≥99.8%, 1 h, RT), separating solid and liquid components by centrifugation (4000 RPM, 15 min, Eppendorf 5810R with A-4-81 rotor) and discarding the supernatant. The pellets surface area was increased with a spatula, and acetone was left to evaporate (3 d, RT), producing a cold alcohol/acetone insoluble residue (AIR).

Sequential Extraction of Cell Wall Polysaccharides. In total, 950 AIRs derived from 475 stem accessions collected from 330 different *B. napus* cultivars were sequentially extracted. Controls containing a mixture of AIRs from 10 randomly selected accessions were also included with each set of 88 to observe any plate-to-plate variations (extraction control: A12). A bulk extract containing 2 g of combined AIR, 40 mL of extractant, was also included with each plate (quantification control: position B12). AIR from each accession (0.5 g) was weighed into 15-mL polypropylene tubes in duplicate on a three-decimal-place balance, and order was randomized before extraction. Each sample was mixed with ammonium oxalate (10 mL, 50 mM) and extracted overnight (RT). A sample of each supernatant (1 mL) after centrifugation was loaded into two 96-well, 2-mL, deep-well polypropylene plates (Nunc) (positions A1 to H11) by using an Andrew Liquid Handling Robot (Andrew Alliance). The remaining supernatant from each sample was discarded, and the pellet was washed thrice in ddH_2O. This process was repeated by using Na_2CO_3 (50 mM) + $NaBH_4$ (0.5% wt/vol), 1 M KOH + $NaBH_4$ (1% wt/vol) and 4 M KOH + $NaBH_4$ (1% wt/vol). Pellets were not washed between collection of the 1 M and 4 M KOH fractions (23). An antifoaming agent, 1-Octanol (10 μL), was added to each

PLANT BIOLOGY

Na_2CO_3, 1 M KOH, and 4 M KOH extract before gradual neutralization to pH 5 with glacial acetic acid (9, 50, and 200 μL, respectively) by using a multichannel pipette and stored at −20 °C.

Relative Quantification Using Carbohydrate Microarrays. Cell wall extracts held in 96-deep-well plates were mixed (20 aspiration and dispense cycles of 80 μL), and subsamples were transferred to 384-deep-well plates (80 μL) by using a Tecan Freedom Evo 100 workstation fitted with a 96-well MultiChannel Arm (MCA96) with disposable tips (Tecan). Seven sequential, twofold dilutions of each sample (50–0.78% original) were made across separate 384-deep-well plates (final volume 20 μL). Samples were printed onto FAST Slides (Whatman) by using a bespoke Stanford-designed microarray printer (cmgm.stanford.edu/pbrown/mguide/) equipped with SMT-S75 silicon printing pins (Silicon Microarray). Duplicate 96 × 96 square arrays consisting of 75- to 110-μm square features at 175-μm intervals were printed on each slide in 20 slide batches, at elevated humidity (>60% relative humidity, 18–23 °C). After printing, slides were stored over P_2O_5 before mAbs probing. Slides were blocked with 1% nonfat milk in 0.1 M Tris-buffered saline (1% MTBS, pH 7.6, 2 h). Arrays were incubated overnight with CCRC, JIM, and MAC-series primary mAbs (50 μL, sourced from Carbosource), diluted 500 times in 0.1% MTBS with 0.01% thiomersal added to prevent microbial contamination. Slides were rinsed at least thrice in 0.1% MTBS (50 mL), and incubated with a secondary antibody (anti-Rat or anti-Mouse IgG produced in goat conjugated to Alexa Fluor 555, diluted to 5 μg/mL in 0.1% MTBS, 1 h). The secondary antibody was removed by rinsing five times in 0.1% MTBS (50 mL each) and dried by centrifugation (500 RPM, 3 min, RT). Slides were read at an excitation wavelength of 532 nm by using an Axon 4200AL microarray scanner and analyzed by using GenePix Pro-6.1 software. Gain of the photomultiplier tubes was adjusted to maximize the dynamic range across each array (maximum fluorescence >66,000). Data were fitted to the Gompertz equation, and the dilution was required to generate a midrange fluorescence calculated by using R (SSgompertz, www.r-project.org/). Broad exclusion criteria were set to remove samples that were outside the limits of detection. These data included either sample features that were not discernible from the background (blanks) or where the most extensive dilution used here (1/128) was not sufficient to bring the feature fluorescence into the linear range. The data were then aggregated by calculating the mean relative abundance of each epitope (1/dilution factor) for each genotype.

SNP Detection, Gene Expression, and Associative Transcriptomics. Mapping and associative transcriptomics was conducted by using described methods and scripts (3, 7, 42, 43), with minor modifications. For full details, see *SI Appendix, SI Materials and Methods*. In total, 9,839 simple SNPs and 246,558 hemi SNPs, along with 53,889 significantly expressed coding DNA sequence models (>0.4 mean RPKM) were used for associative transcriptomics.

ACKNOWLEDGMENTS. We thank Martin Broadley, Neil Graham, Rory Hayden, and colleagues from the University of Nottingham for providing the stem samples used in this study and Archontoula Gerolymatou and Morgane Castaing for laboratory assistance. This work was funded in part by Biotechnology and Biological Sciences Research Council-Renewable Industrial Products from Rapeseed Programme Grant BB/L002124/1 and Institute Strategic Programme "Food and Health" Grant BB/J004545/1. Generation of the CCRC-series of mAbs used in this work was supported by NSF Plant Genome Program Grant DBI-0421683. Distribution of the JIM and MAC antibodies used in this work was supported in part by NSF Grants DBI-0421683 and RCN 009281.

1. McCann MC, Carpita NC (2015) Biomass recalcitrance: A multi-scale, multi-factor, and conversion-specific property. *J Exp Bot* 66:4109–4118.
2. Schmutzer T, et al. (2015) Species-wide genome sequence and nucleotide polymorphisms from the model allopolyploid plant Brassica napus. *Sci Data* 2:150072.
3. Bancroft I, et al. (2011) Dissecting the genome of the polyploid crop oilseed rape by transcriptome sequencing. *Nat Biotechnol* 29:762–766.
4. He Z, et al. (2015) Construction of Brassica A and C genome-based ordered pan-transcriptomes for use in rapeseed genomic research. *Data in Brief* 4:357–362.
5. Zhu CS, Gore M, Buckler ES, Yu JM (2008) Status and prospects of association mapping in plants. *Plant Genome-Us* 1:5–20.
6. Paterson AH, Lan TH, Amasino R, Osborn TC, Quiros C (2001) Brassica genomics: a complement to, and early beneficiary of, the Arabidopsis sequence. *Genome Biol* 2: S1011.
7. Harper AL, et al. (2012) Associative transcriptomics of traits in the polyploid crop species Brassica napus. *Nat Biotechnol* 30:798–802.
8. Stokes D, et al. (2010) An association transcriptomics approach to the prediction of hybrid performance. *Mol Breed* 26:91–106.
9. Vidaurre D, Bonetta D (2012) Accelerating forward genetics for cell wall deconstruction. *Front Plant Sci* 3:119.
10. Slavov G, Allison G, Bosch M (2013) Advances in the genetic dissection of plant cell walls: Tools and resources available in Miscanthus. *Front Plant Sci* 4:217.
11. Houston K, et al. (2015) A genome-wide association study for culm cellulose content in barley reveals candidate genes co-expressed with members of the CELLULOSE SYNTHASE A gene family. *PLoS One* 10:e0130890.
12. Li K, et al. (2016) Genome-wide association study reveals the genetic basis of stalk cell wall components in maize. *PLoS One* 11:e0158906.
13. Muchero W, et al. (2015) High-resolution genetic mapping of allelic variants associated with cell wall chemistry in Populus. *BMC Genomics* 16:24.
14. Wang YH, et al. (2013) Mapping and candidate genes associated with saccharification yield in sorghum. *Genome* 56:659–665.
15. Penning BW, et al. (2014) Genetic determinants for enzymatic digestion of lignocellulosic biomass are independent of those for lignin abundance in a maize recombinant inbred population. *Plant Physiol* 165:1475–1487.
16. McCann MC, Carpita NC (2005) Looking for invisible phenotypes in cell wall mutants of Arabidopsis thaliana. *Plant Biosyst* 139:80–83.
17. Collins SRA, et al. (2014) Variation in the chemical composition of wheat straw: The role of tissue ratio and composition. *Biotechnol Biofuels* 7:121.
18. Günl M, Kraemer F, Pauly M (2011) Oligosaccharide Mass Profiling (OLIMP) of cell wall polysaccharides by MALDI-TOF/MS. *The Plant Cell Wall, in Molecular Biology*, ed Popper ZA (Humana, New York), Vol 715, pp 43–54.
19. Sørensen I, Willats WG (2011) Screening and characterization of plant cell walls using carbohydrate microarrays. *Methods Mol Biol* 715:115–121.
20. Moller I, et al. (2007) High-throughput mapping of cell-wall polymers within and between plants using novel microarrays. *Plant J* 50:1118–1128.
21. Lin F, et al. (2016) Cell wall composition and candidate biosynthesis gene expression during rice development. *Plant Cell Physiol* 57:2058–2075.
22. Higgins J, Magusin A, Trick M, Fraser F, Bancroft I (2012) Use of mRNA-seq to discriminate contributions to the transcriptome from the constituent genomes of the polyploid crop species Brassica napus. *BMC Genomics* 13:247.
23. Pattathil S, Avci U, Miller JS, Hahn MG (2012) Immunological approaches to plant cell wall and biomass characterization: Glycome profiling. *Biomass Conversion, Methods in Molecular Biology*, ed Himmel ME (Humana, New York), Vol 908, pp 61–72.
24. Selvendran RR, O'Neill MA (1987) Isolation and analysis of cell walls from plant material. *Methods Biochem Anal* 32:25–153.
25. Decker S, Brunecky R, Tucker M, Himmel M, Selig M (2009) High-throughput screening techniques for biomass conversion. *Bioenerg Res* 2:179–192.
26. Moller I, et al. (2008) High-throughput screening of monoclonal antibodies against plant cell wall glycans by hierarchical clustering of their carbohydrate microarray binding profiles. *Glycoconj J* 25:37–48.
27. Pattathil S, et al. (2010) A comprehensive toolkit of plant cell wall glycan-directed monoclonal antibodies. *Plant Physiol* 153:514–525.
28. Scheller HV, Ulvskov P (2010) Hemicelluloses. *Annu Rev Plant Biol* 61:263–289.
29. Brown DM, et al. (2007) Comparison of five xylan synthesis mutants reveals new insight into the mechanisms of xylan synthesis. *Plant J* 52:1154–1168.
30. Chiniquy D, et al. (2013) Three novel rice genes closely related to the Arabidopsis IRX9, IRX9L, and IRX14 genes and their roles in xylan biosynthesis. *Front Plant Sci* 4:83.
31. Keppler BD, Showalter AM (2010) IRX14 and IRX14-LIKE, two glycosyl transferases involved in glucuronoxylan biosynthesis and drought tolerance in Arabidopsis. *Mol Plant* 3:834–841.
32. Lee C, Zhong R, Ye ZH (2012) Arabidopsis family GT43 members are xylan xylosyltransferases required for the elongation of the xylan backbone. *Plant Cell Physiol* 53:135–143.
33. Lee C, Teng Q, Zhong R, Ye ZH (2011) Molecular dissection of xylan biosynthesis during wood formation in poplar. *Mol Plant* 4:730–747.
34. Ren Y, Hansen SF, Ebert B, Lau J, Scheller HV (2014) Site-directed mutagenesis of IRX9, IRX9L and IRX14 proteins involved in xylan biosynthesis: Glycosyltransferase activity is not required for IRX9 function in Arabidopsis. *PLoS One* 9:e105014.
35. Wu AM, et al. (2010) Analysis of the Arabidopsis IRX9/IRX9-L and IRX14/IRX14-L pairs of glycosyltransferase genes reveals critical contributions to biosynthesis of the hemicellulose glucuronoxylan. *Plant Physiol* 153:542–554.
36. Bromley JR, et al. (2013) GUX1 and GUX2 glucuronyltransferases decorate distinct domains of glucuronoxylan with different substitution patterns. *Plant J* 74:423–434.
37. Busse-Wicher M, et al. (2014) The pattern of xylan acetylation suggests xylan may interact with cellulose microfibrils as a two-fold helical screw in the secondary plant cell wall of Arabidopsis thaliana. *Plant J* 79:492–506.
38. Schmidt D, Schuhmacher F, Geissner A, Seeberger PH, Pfrengle F (2015) Automated synthesis of arabinoxylan-oligosaccharides enables characterization of antibodies that recognize plant cell wall glycans. *Chemistry* 21:5709–5713.
39. Mortimer JC, et al. (2010) Absence of branches from xylan in Arabidopsis gux mutants reveals potential for simplification of lignocellulosic biomass. *Proc Natl Acad Sci USA* 107:17409–17414.
40. Rose JKC, Braam J, Fry SC, Nishitani K (2002) The XTH family of enzymes involved in xyloglucan endotransglucosylation and endohydrolysis: Current perspectives and a new unifying nomenclature. *Plant Cell Physiol* 43:1421–1435.
41. Thomas CL, et al. (2016) Root morphology and seed and leaf ionomic traits in a Brassica napus L. diversity panel show wide phenotypic variation and are characteristic of crop habit. *BMC Plant Biol* 16:214.
42. Higgins J, Magusin A, Trick M, Fraser F, Bancroft I (2012) Use of mRNA-seq to discriminate contributions to the transcriptome from the constituent genomes of the polyploid crop species Brassica napus. *BMC Genomics* 13:247.
43. Lu G, et al. (2014) Associative transcriptomics study dissects the genetic architecture of seed glucosinolate content in Brassica napus. *DNA Res* 21:613–625.