

# UC Davis

## UC Davis Previously Published Works

### Title

A nonenzymatic method for cleaving polysaccharides to yield oligosaccharides for structural analysis.

### Permalink

<https://escholarship.org/uc/item/75j598dg>

### Journal

Nature communications, 11(1)

### ISSN

2041-1723

### Authors

Amicucci, Matthew J

Nandita, Eshani

Galermo, Ace G

et al.

### Publication Date

2020-08-01

### DOI





10.1038/s41467-020-17778-1

### Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

# A nonenzymatic method for cleaving polysaccharides to yield oligosaccharides for structural analysis

Matthew J. Amicucci<sup>1,2,3,6</sup>, Eshani Nandita<sup>2,3,6</sup>, Ace G. Galermo <sup>2,3</sup>, Juan Jose Castillo <sup>2,3</sup>, Siyu Chen<sup>2,3</sup>, Dayoung Park <sup>2,3</sup>, Jennifer T. Smilowitz<sup>3,4</sup>, J. Bruce German<sup>3,4</sup>, David A. Mills<sup>3,4</sup> & Carlito B. Lebrilla <sup>1,2,3,5</sup>✉

Polysaccharides are the most abundant biomolecules in nature, but are the least understood in terms of their chemical structures and biological functions. Polysaccharides cannot be simply sequenced because they are often highly branched and lack a uniform structure. Furthermore, large polymeric structures cannot be directly analyzed by mass spectrometry techniques, a problem that has been solved for polynucleotides and proteins. While restriction enzymes have advanced genomic analysis, and trypsin has advanced proteomic analysis, there has been no equivalent enzyme for universal polysaccharide digestion. We describe the development and application of a chemical method for producing oligosaccharides from polysaccharides. The released oligosaccharides are characterized by advanced liquid chromatography–mass spectrometry (LC–MS) methods with high sensitivity, accuracy and throughput. The technique is first used to identify polysaccharides by oligosaccharide fingerprinting. Next, the polysaccharide compositions of food and feces are determined, further illustrating the utility of technique in food and clinical studies.

<sup>1</sup>Agricultural and Environmental Chemistry Graduate Group, University of California, Davis, CA, USA. <sup>2</sup>Department of Chemistry, University of California, Davis, CA, USA. <sup>3</sup>Foods For Health Institute, University of California, Davis, CA, USA. <sup>4</sup>Department of Food Science and Technology, University of California, Davis, CA, USA. <sup>5</sup>Department of Biochemistry and Molecular Medicine, University of California, Davis, CA, USA. <sup>6</sup>These authors contributed equally: Matthew J. Amicucci, Eshani Nandita. ✉email: [cblebrilla@ucdavis.edu](mailto:cblebrilla@ucdavis.edu)

The fundamental importance of carbohydrates to biological organisms remains discouragingly elusive. The plants that nourish all life forms and stabilize the planet's atmosphere, are mostly carbohydrate. Yet scientists struggle to provide even the most crude predictions of how the complexity of polysaccharide structures drives their endogenous functions much less their roles in fueling the entire food web. Epidemiologic and observational studies have tentatively documented that polysaccharides (as fiber) are a vitally important dimension of nutrition and health<sup>1–4</sup>. But investigations of the detailed structural basis of those functions is ostensibly nonexistent. Understanding the structure/function relationships of polysaccharides requires that their structures be known. They are not. The development of analytical tools for characterizing polysaccharides has lagged far behind that of DNA and proteins<sup>5</sup>. Polysaccharide analysis remains slow and tedious, as current strategies still employ methods developed long before the development of high throughput structural tools commonly used in genomic and proteomic analysis<sup>6–8</sup>. Unfortunately, these limitations in determining polysaccharide structures hinder our ability to understand their fundamental biological roles in fields from fuel efficiency to human health. The deficiency in structural methods for polysaccharides is becoming particularly acute in foods where the polysaccharide compositions represent a major gap in knowledge and are rarely, if ever, identified or quantified<sup>9</sup>.

The dissociation of polysaccharides into constitutive and measurable oligosaccharides has been the missing critical component for characterizing polysaccharides. Robust digestion methods are the key to polysaccharide analysis, which opens polysaccharides to the most powerful modern chromatographic, mass spectrometry, computational, and robotic techniques<sup>10–14</sup>. In proteomic analysis, the ability to generate oligopeptides was a necessary first-step toward high throughput analysis. For proteomic analysis, proteases such as trypsin yielded highly reproducible cleavages and produced peptides of sufficient length to identify the parent protein. The methods involved the matching of peptide masses to databases for peptide fingerprinting<sup>15–17</sup>, and the use of tandem mass spectrometry to obtain sequence-specific fragments for protein identification<sup>18</sup>. However, the analogy between protein and polysaccharide sequencing must end here. Unlike proteins and DNA, which have discrete and linear primary structures, polysaccharide structures may be composed of homo or copolymers with stochastically distributed monomers or pendant polymers each with distinct monosaccharide compositions with various glycosidic linkages. To complicate the primary structures further, branching may occur with no distinct length of polymer. Due to the large number of variable structural attributes, polysaccharides do not have precise structures and must be thought of as probabilistic distributions of their innate structural features. The lack of a unified structure further means that there is no enzymatic process, or a trypsin analog, to produce oligosaccharides from most polysaccharides. Instead, there are enzymes such as amylase that produce monosaccharides and oligosaccharides from only specific polysaccharides (amylose and amylopectin). Similarly, animal polysaccharide such as glycosaminoglycans (GAGs) are routinely digested to disaccharides via enzymatic approaches for further analysis by liquid chromatography–mass spectrometry (LC–MS)<sup>19–23</sup>. Due to established and robust methods, GAGs were not explored in this report. The lack of a single unified polymeric structure further obviates the use of strategies that aim to reassemble parent polymer sequence from aligning overlapping oligomer sequences. Rather, polysaccharide sequence should be used to understand the frequency that individual structural features (branching, linkages between individual monosaccharides, and modifications) are found in a distribution of similar but, often, structurally unique molecules.

The structural heterogeneity of polysaccharides is due to the large diversity in monosaccharide and linkage compositions. Thus, there is no universal digestive enzyme for producing oligosaccharides from polysaccharides. The repertoire of natural glycosidase enzymes is conspicuous by their high specificity for specific glycosidic bonds. Therefore, only a limited number of polysaccharide linkages can be digested with a specific enzyme<sup>24,25</sup>. Acid-based techniques have been used to produce oligosaccharides from polysaccharides with varying levels of success, but generally require precise reaction conditions for each polysaccharide, and is therefore not well suited for the analysis of complicated mixtures or heterogeneous structures<sup>26–28</sup>.

The combination of complexity of polysaccharide structures and specificity of glycosidase catalysis is central to biology itself. Fuel and carbon flows through all ecosystems largely as carbohydrate. Managing these flows for any purpose from energy efficiency and sustainability to human metabolism and health requires that the structures of all biopolymers that contain carbohydrate to be known. The need for more robust polysaccharide characterization techniques is especially crucial for characterizing food polysaccharides. Carbohydrates make up the largest part of most human diets, which are now associated with various metabolic diseases<sup>29,30</sup> and are the largest driver of the gut microbial population<sup>31–35</sup>. While there has been rapid progress in characterizing the microbial composition of the gut microbiome<sup>36,37</sup>, similar advances have not been made in characterizing the carbohydrates that modulate those populations<sup>9</sup>. Thus, the impact that dietary carbohydrates have on the gut microbiome is, generally, indirectly measured by genomic and transcriptomic techniques rather than through direct carbohydrate analysis, which continues to hinder our mechanistic understanding of this complicated process<sup>38–40</sup>. Therefore, there is an immediate and pressing need to expand the current toolbox of polysaccharide characterization if we are to better understand diet–microbe interactions<sup>31,41–46</sup>.

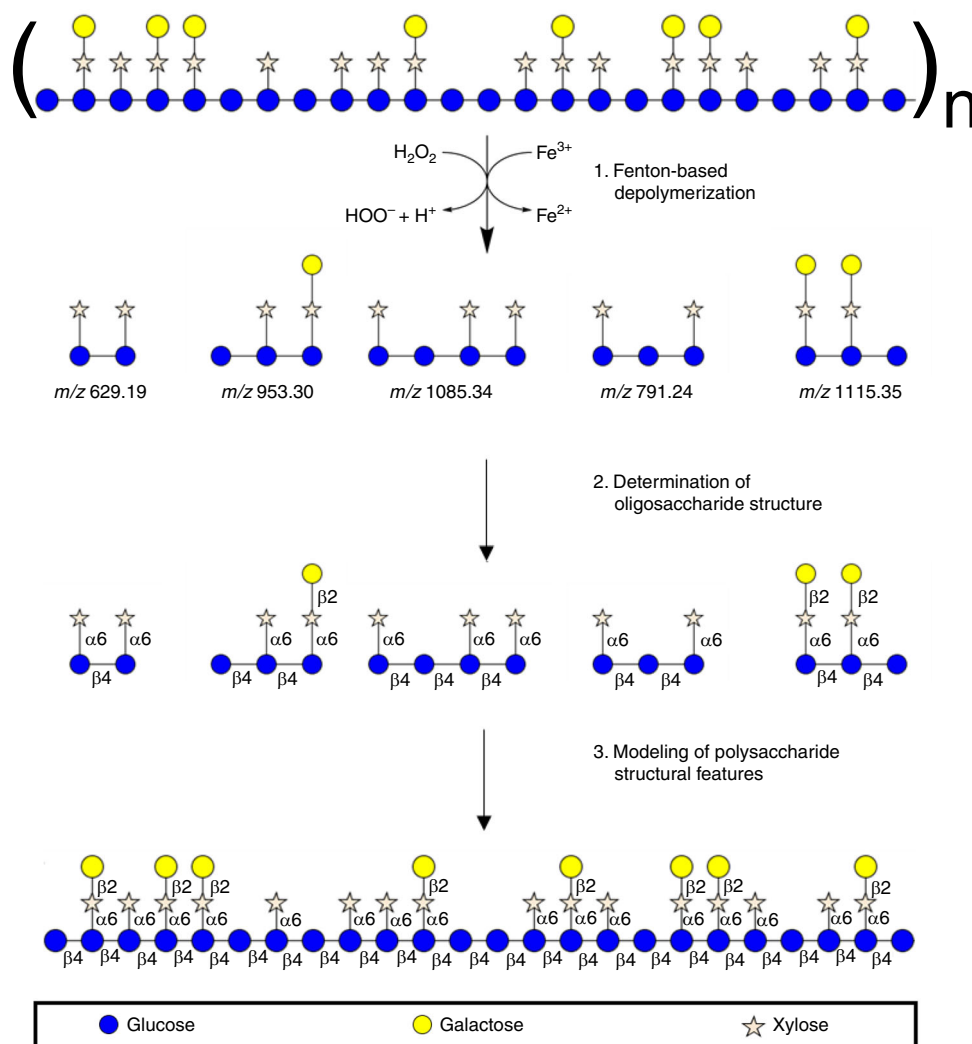
In this report, we describe a chemical method for the dissociation of diverse polysaccharides into oligosaccharides. Fenton's initiation toward defined oligosaccharide groups (FITDOG) employs oxidative chemistry to disassemble structurally diverse polysaccharides into oligosaccharides that are of sufficient length to be structurally probed by LC–MS.

## Results

**Optimization of reaction conditions.** The FITDOG process was developed to produce oligosaccharides from polysaccharides so that advanced separation and mass spectrometry methods could be employed for analysis. The process was initiated by a reaction between a metal catalyst, Fe<sup>3+</sup>, and an oxidizing agent, hydrogen peroxide, to produce reactive radical species that cleave glycosidic bonds. The radicals induce oxidative cleavage of the polysaccharide backbone to produce oligosaccharides that are representative of the parent polysaccharide structure (Fig. 1). An LC–MS/MS approach employing collision induced dissociation (CID) was used to obtain structural information.

A considerable effort was placed in developing a general reaction condition suitable for a large number of structural variations. These conditions included the relative amounts of reagents and analytes, reaction conditions, and time. Over oxidation was an issue that needed to be controlled. The conditions presented herein represent the best conditions for a large number of different polysaccharides including those in complicated matrices such as food.

**Production of oligosaccharides by FITDOG.** Polysaccharides with known structures were used to determine the full capabilities



**Fig. 1** Oxidative method of polysaccharide depolymerization. FITDOG reaction illustrating the dissociation of xyloglucan polysaccharide into representative oligosaccharides and its reconstruction by reassembly of the elucidated oligosaccharides.

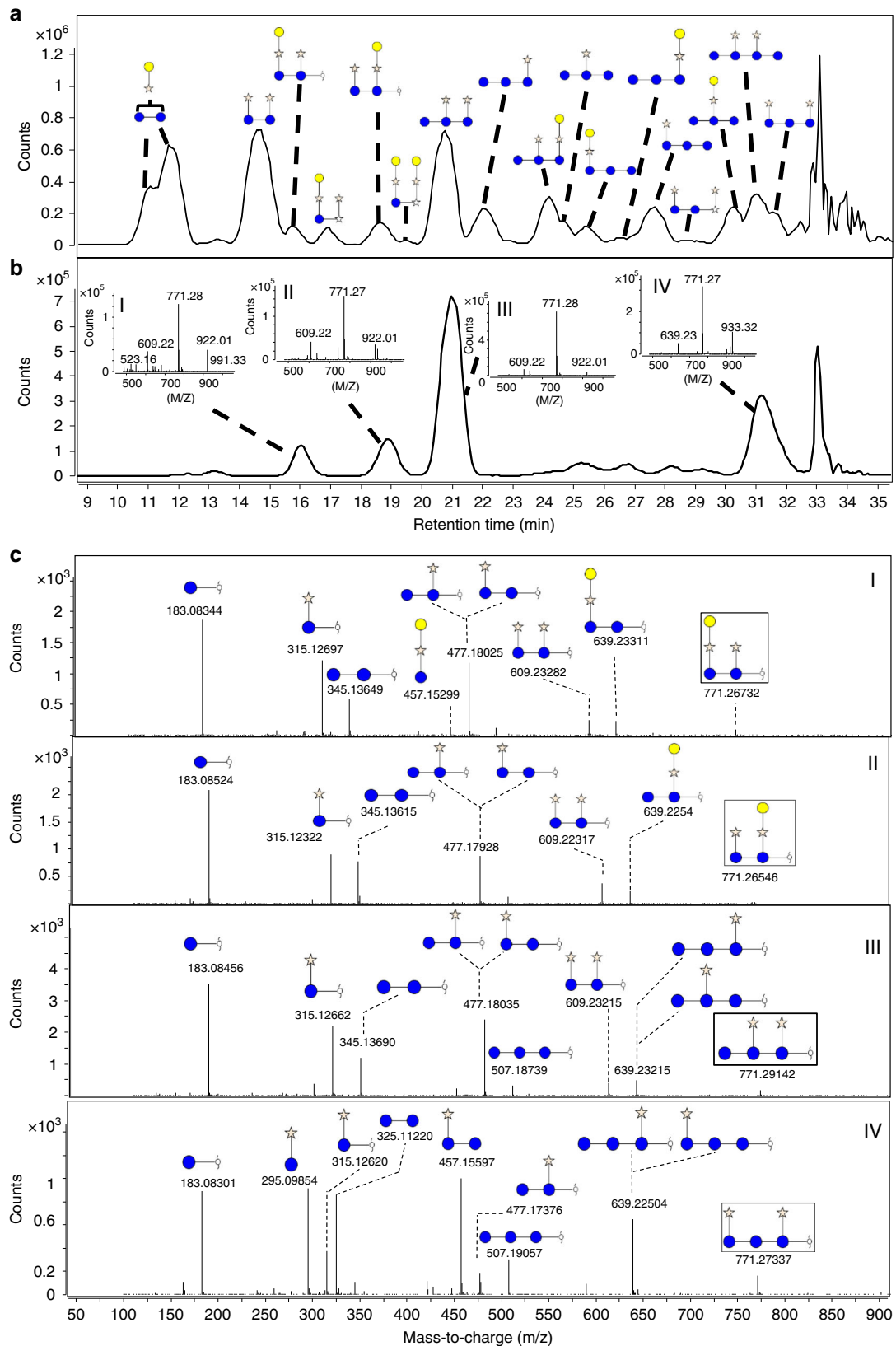
and characteristics of the method. For example, xyloglucan is comprised of a  $\beta(1 \rightarrow 4)$  glucose backbone with frequent  $\alpha(1 \rightarrow 6)$  xylose branches that are often terminated with a single  $\beta(1 \rightarrow 2)$  galactose residue<sup>25,47–49</sup>. The reaction of xyloglucan yielded over 20 structurally unique oligosaccharides as determined by nano-HPLC-chip/Q-TOF MS. The most abundant oligosaccharides are shown in Fig. 2a, while the complete mass/rt library and annotated profiles are provided in the Supplementary Data 1, 2, respectively.

The degrees of polymerization (DPs) were obtained from the accurate masses, while structural information was obtained from tandem MS. Isomers, compounds with the same DP but different structures, were present and were structurally elucidated. For example, four isomers with the compositions three hexoses and two pentoses ( $\text{Hex}_3\text{Pnt}_2$ ,  $m/z$  771.27) were observed at 16.0 (compound I), 18.9 (II), 21.0 (III), and 31.2 (IV) min, respectively (Fig. 2b). The corresponding tandem MS (MS/MS) spectra were obtained and annotated (Fig. 2c). The structure of xyloglucan is generally known and can be used to obtain precise structural information of the oligosaccharides. Based on these analyses, the inset structures for the isomers were deduced and were consistent with the known structures of xyloglucan.

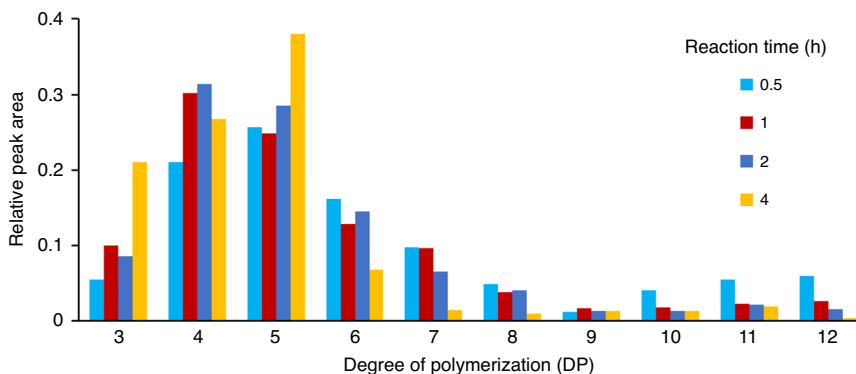
There was some control on the size distribution of the oligosaccharide products by varying the time of the reaction. At

the optimized conditions, the products favored a DP 3–14. However, by increasing the reaction time, smaller DPs were biased. This effect was illustrated with amylose, a simple homopolymer, to provide a clear illustration of the shift in the degree of polymerization (DP) with reaction time. The shortest reaction time, 0.5 h, yielded products with DP 3–14, while at 2 h only those corresponding to DP 3–8 were observed, and at 4 h only those corresponding to DP 3–6 were observed (Fig. 3). While the reaction time did give variability in the DP, other factors such as choice of metal and peroxide concentration may allow the reaction conditions to be further optimized to give oligosaccharide reads of a desired length.

To determine the general utility of the approach, other polysaccharide standards were subjected to FITDOG and yielded representative oligosaccharides (Fig. 4). The polysaccharide standards included amylose, amylopectin, cellulose, curdlan, lichenan,  $\beta$ -glucan, arabinan, xylan, arabinoxyylan, mannan, galactomannan, glucomannan, galactan, and xyloglucan. These represent a large proportion of known polysaccharides and were selected because they were composed of many monosaccharide combinations, linkage types, branching patterns, while some contained monosaccharide modifications. Annotated chromatograms for each polysaccharide after FITDOG are shown in Fig. 4a–n. Prominent peaks were labeled with the DP and an abbreviation of their



**Fig. 2** Characterization of FITDOG derived xyloglucan oligosaccharides. **a** Annotated base peak chromatogram of characterized xyloglucan-derived oligosaccharides. **b** Extracted ion chromatogram of  $m/z$  771.27 shows four isomers. **c** Annotated tandem mass spectra of the four  $m/z$  771.27 isomers and their final elucidated structures.



**Fig. 3 Reaction time and oligosaccharide size.** Relative peak area of different size oligosaccharides when reacted from 0.5–4 h.

monosaccharide components (Hexose-Hex, Pentose-Pnt, Hexuronic Acid/Hxa, 4-*O*-methylated glucuronic acid – GlcAOMe) (Fig. 4). A more comprehensive table of the oligosaccharides produced were listed in the Supplementary Data 3.

A single procedure worked for all of the polysaccharides we tested. The homopolysaccharides, a group of polymers that are composed of a single monosaccharide and glycosidic linkage, yielded linear oligosaccharides with a single representation for each DP (no isomers). For example, curdlan, a linear  $\beta(1 \rightarrow 3)$  linked glucose polymer yielded oligosaccharides with compositions corresponding to Hex<sub>3</sub> (RT 12.56 min), Hex<sub>4</sub> (19.28 min), and Hex<sub>5</sub> (26.44 min) (Fig. 4a). Other homopolymers included cellulose with  $\beta(1 \rightarrow 4)$  glucose (Fig. 4b), amylose  $\alpha(1 \rightarrow 4)$  glucose (4c), galactan  $\beta(1 \rightarrow 4)$  galactose (4d), and mannan  $\beta(1 \rightarrow 4)$  mannose (4e). It was noted that impurities remained in the polysaccharide from the manufacturer process yielding peaks corresponding to other polysaccharides.

More complicated structures including heteropolysaccharides, which contained more than one monosaccharide or glycosidic linkage, also yielded structurally representative oligosaccharides. The structural heterogeneity contributed to the production of several oligosaccharide isomers for each DP. Among this group were  $\beta$ -glucan and lichenan, both of which have similar structures composed of  $\beta(1 \rightarrow 3)$  and  $\beta(1 \rightarrow 4)$  linked glucose residues, but in differing ratios. Thus, the composition corresponding to Hex<sub>6</sub> was observed to generate eight isomers from lichenan and four from  $\beta$ -glucan (Fig. 4f, g). Moreover, the structural similarities in monosaccharide and linkage composition between lichenan and  $\beta$ -glucan yielded many similar oligosaccharides from both polysaccharides. For example, the Hex<sub>3</sub> (13.73 min), Hex<sub>4</sub> (19.46, 19.70, and 20.41 min), Hex<sub>5</sub> (24.92, 25.79 min), and Hex<sub>6</sub> (30.90 min) were identical and shared by the two polysaccharides (Fig. 4f, g). Glucomanan, another linear heteropolymer composed of  $\beta(1 \rightarrow 4)$  glucose and mannose residues, also yielded several isomers for each DP (Fig. 4h).

Branched polysaccharides that contain multiple linkage types and multiple monosaccharides yielded the most isomers per DP. Galactomannan is a branched heteropolysaccharides containing a  $\beta(1 \rightarrow 4)$  linked mannose backbone and  $\alpha(1 \rightarrow 6)$  linked terminal galactose branches. With FITDOG, the oligosaccharides resembled those from manno-oligosaccharides, derived from the mannose backbone, and galactomanno-oligosaccharides derived from the branching regions (Fig. 4i). Arabinan, xylan, arabinoxy-ylan, and xyloglucan, all are branched polysaccharides with mixed monomeric units, and were also dissociated to similarly generate many isomeric species (Fig. 4j–m). The dissociation of these diverse structures further demonstrated the method's reactivity towards pentose- and hexose-containing polysaccharides and highly complicated structures. Unexpectedly, amylopectin, another branched

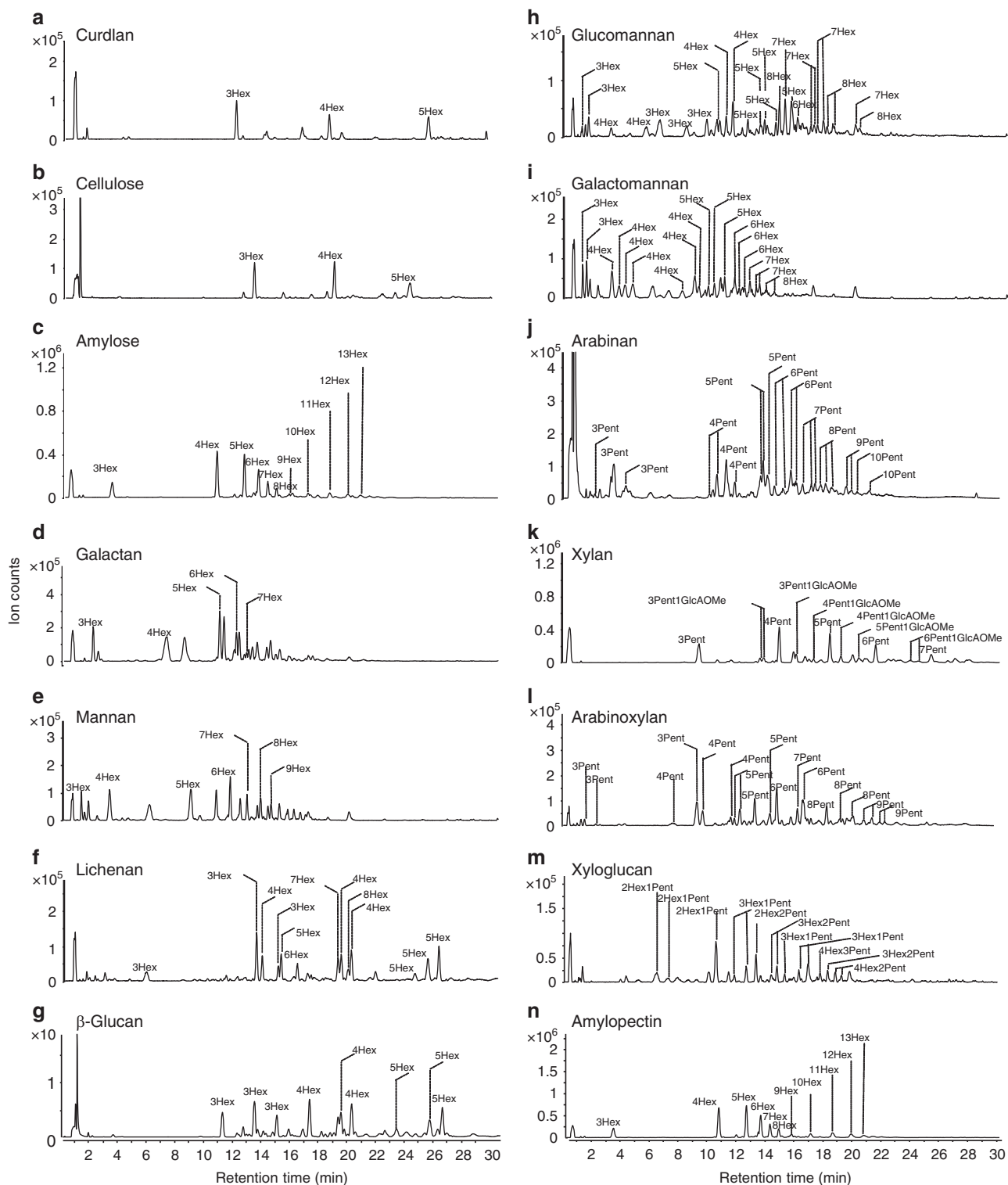
heteropolysaccharide containing an  $\alpha(1 \rightarrow 4)$  linked glucose backbone with  $\alpha(1 \rightarrow 4,6)$  bisecting glucose, produced only one structure per DP. The presence of a single isomer from amylopectin may result from a cleavage preference for  $\alpha(1 \rightarrow 6)$  branching points, thereby producing linear oligosaccharide (Fig. 4n)<sup>50</sup>.

Polysaccharides containing modified monosaccharide residues were also examined to determine whether FITDOG removed saccharide modifications. Xylan contained  $\beta(1 \rightarrow 4)$ -linked xylose polysaccharide with terminal 4-*O*-methyl glucuronic acid (GlcAOMe) residues. Under FITDOG conditions, several oligosaccharides were generated with unique compositions including Pnt<sub>3</sub> GlcAOMe (4 isomers), Pent<sub>4</sub> GlcAOMe (e isomers), Pent<sub>5</sub> GlcAOMe (4 isomers), Pent<sub>6</sub> GlcAOMe (3 isomers), and Pent<sub>7</sub> GlcAOMe (2 isomers). The presence of these oligosaccharides indicated that *O*-methylation on the terminal glucuronic acid was preserved during the FITDOG treatment (Fig. 4k). Other modifications are known for carbohydrates including sulfation, ferulic acid, phosphorylation. How these modifications will be impacted by this method is currently under investigation.

**Identification of polysaccharides by fingerprinting.** With the unique oligosaccharide compositions representing each parent polysaccharide, we determined whether the oligosaccharides could be used as diagnostic markers for polysaccharide identification in complicated mixtures such as those found in carbohydrate mixtures. This approach is analogous in proteomics to peptide fingerprinting, which is a robust method for identifying proteins based upon the presence of diagnostic peptides. The oligosaccharide products of polysaccharides were compiled as unique identifiers for the respective parent, which created an oligosaccharide reference library with nearly 400 unique oligosaccharides (Fig. 5a, Supplementary Data 3).

Wheat and oat bran were selected to validate the oligosaccharide fingerprinting method as they were known to contain large amounts of non-starch polysaccharides including arabinoxy-ylans<sup>51</sup>, mixed linkage  $\beta$ -glucans<sup>52</sup>, and cellulose<sup>53</sup>. Oat and wheat bran each produced over 50 distinct oligosaccharides with most matching entries in the reference library (Fig. 5b, c). A pool containing all the oligosaccharides was run alongside unknown samples for retention time alignment and source-polymer identification (Fig. 5a). The full list of matched peaks is presented in the Supplementary Data 4. For identification, we used the number of oligosaccharides produced in the individual standards to determine the fraction of the polysaccharide products that are observed in the mixture. This number varied with each standard polysaccharide. For cellulose, only four oligosaccharides were observed in our chromatographic window, while for amylopectin over 20 oligosaccharides were observed, and for arabinoxy-ylan over 40 oligosaccharides were observed. We then used the



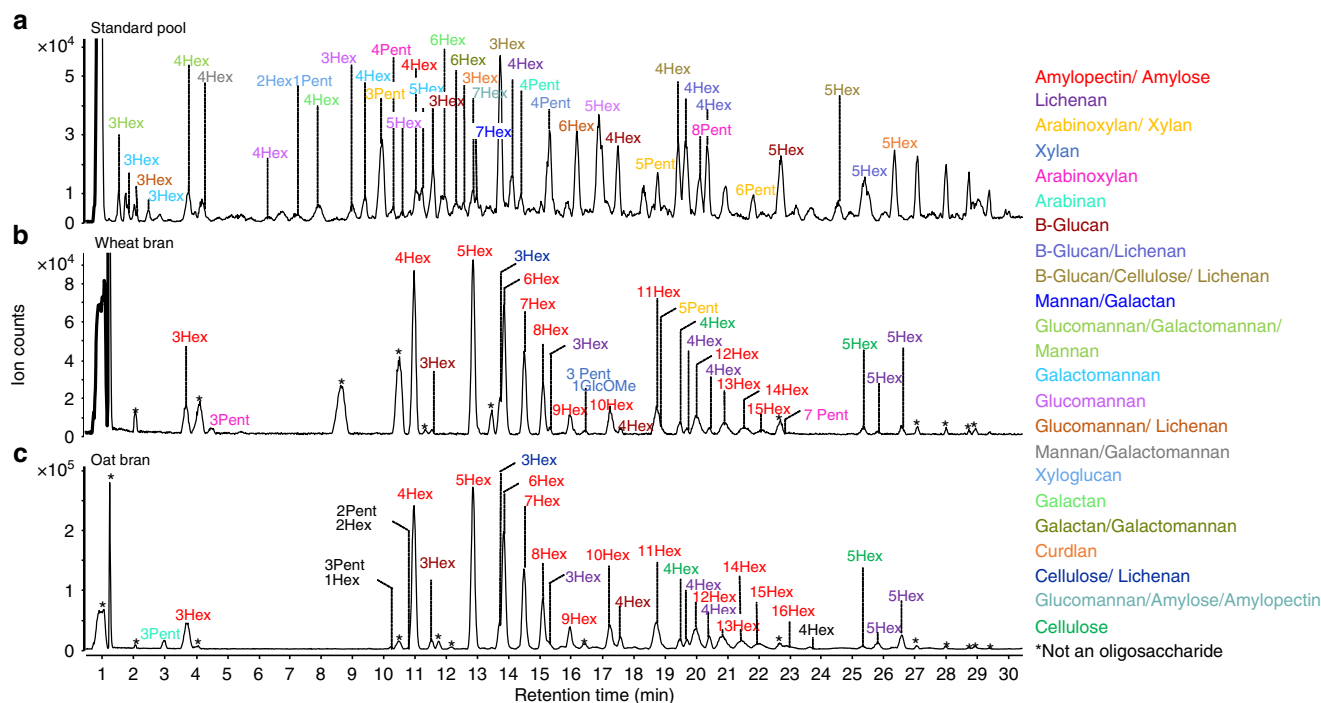


**Fig. 4 Construction of an oligosaccharide reference library.** **a** Annotated base peak chromatograms of oligosaccharides derived from curdlan, **b** cellulose, **c** amylose, **d** galactan, **e** mannan, **f** lichenan, **g**  $\beta$ -glucan, **h** glucomannan, **i** galactomannan, **j** arabinan, **k** xylan, **l** arabinoxylan, **m** xyloglucan, and **n** amylopectin.

number of oligosaccharides found in the mixture represented as a fraction of those found in the pure standard as a percent coverage. The coverage is expected to be higher if the oligosaccharide count in the sample is closer to the number of oligosaccharides produced in the standard.

Based on the FITDOG analyses, we found that wheat bran contained amylose/amylopectin (100% coverage), cellulose (75%),

$\beta$ -glucan (34%), and lichenan (24%). Oat bran was composed of amylose/amylopectin (100% coverage), cellulose (75%),  $\beta$ -glucan (38%), and lichenan (24%). Wheat bran also contained arabinoxylan (20%) and xylan (32%). Interestingly, oat bran contained several unmatched pentose oligomers and mixed hexose and pentose containing oligomers demonstrating the presence of a polysaccharide that is yet to be identified by our



**Fig. 5 Oligosaccharide fingerprinting.** **a** Annotated base peak chromatogram of pooled oligosaccharides derived from the depolymerization of the polysaccharide standards. Annotated base peak chromatograms of characterized oligosaccharides derived from **(b)** wheat bran and **(c)** oat bran.

reference library, which is binned for future identification. Overall, the results matched the general expected polysaccharide composition of both brans<sup>51–53</sup>.

The ability to identify polysaccharides in complicated matrices will yield the biological functions of polysaccharides. In particular, the role of fiber in food is not well understood due to the lack of analytical tools to characterize them. As an illustration of the utility of the method, we probed the fate of complimentary food polysaccharides in an infant gut by monitoring the changes in compositions between the food product and infant feces after consumption. The infants were part of a cohort comprised of exclusively breastfed infants who were fed complementary food in a crossover feeding trial. An infant was fed pear as their first solid food along with continued breastfeeding. Upon completion of the 7-day feeding period, a fecal sample was taken and analyzed for comparison with the native food. This study would provide the fate of dietary carbohydrates as they interact with the infant gut microbiome.

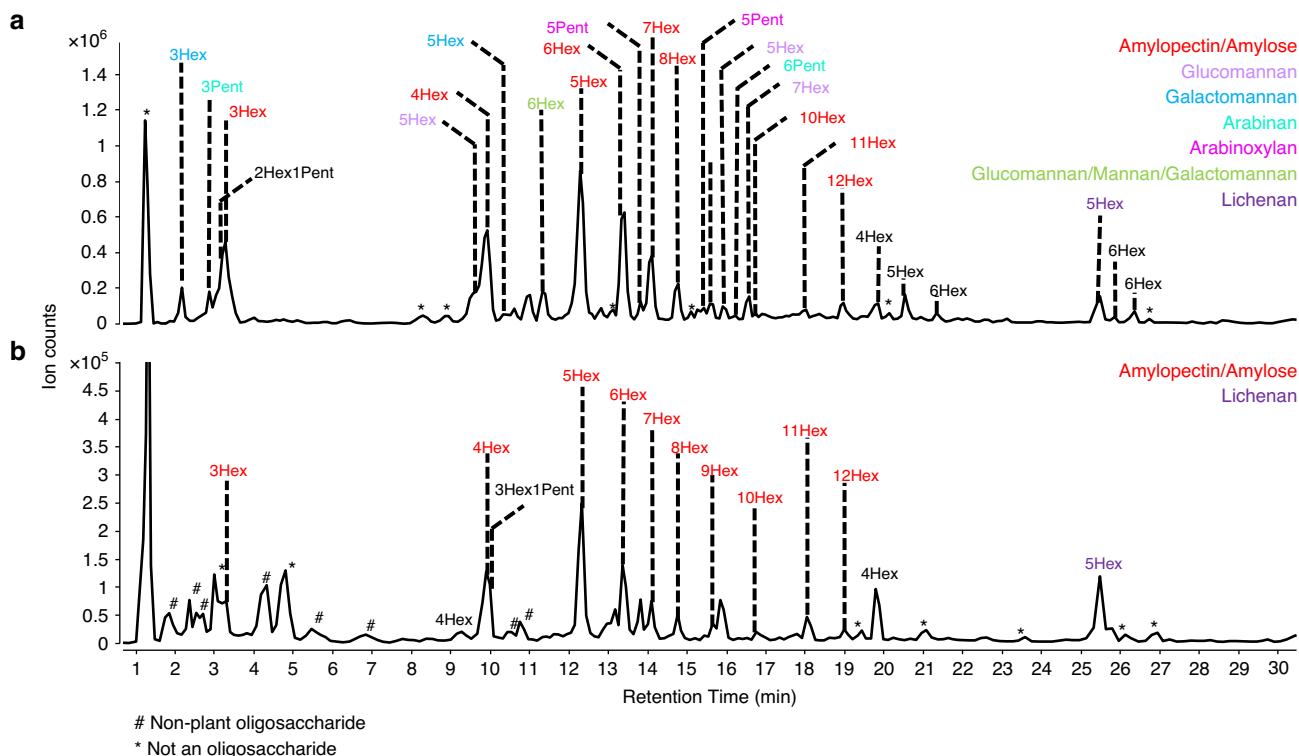
Polysaccharides that made up the pear product were examined and found to be composed of, in decreasing order, amylopectin/ amylose, glucomannan, arabinan, arabinoxylan, galactomannan, mannan, and lichenan (Fig. 6a). Analysis of the feces yielded primarily amylopectin, amylose, and lichenan. (Fig. 6b). These results suggested that amylose, amylopectin, and lichenan were not completely digested while glucomannan, galactomannan, arabinan, arabinoxylan, and mannan appeared to be degraded by host and microbial enzymes.

**De novo structural characterization of oligosaccharides.** Oligosaccharides produced by FITDOG were further structurally elucidated to provide, de novo, the parent polysaccharide. The workflow is shown in Fig. 7. For the analysis of the constitutive oligosaccharides, the products were separated into pools of smaller numbers of compounds many containing single unique structures. The LC eluant was split between the QTOF for MS and MS/MS and a 96-well collection plate. The QTOF MS and MS/MS mass spectra provided some structural information, in

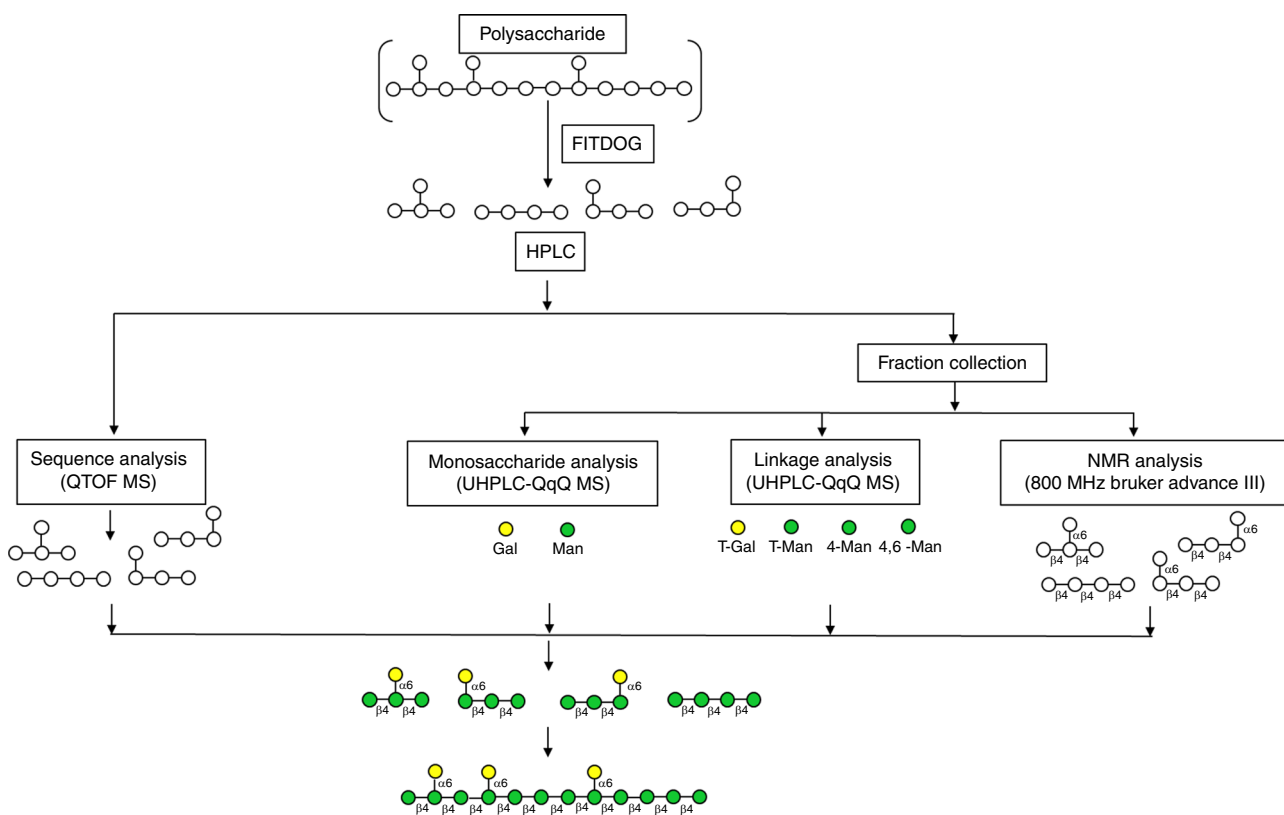
terms of monosaccharide compositions and the degree of oligosaccharide polymerization, of the compounds collected at unique retention times. However, it could not identify monosaccharide constituents, linkages, or the anomeric character of the linkage (alpha versus beta). The nearly 200 collected fractions were further analyzed by rapid throughput monosaccharide and linkage analysis to provide nearly complete structures of each oligosaccharide component<sup>54–57</sup>.

The galactomannan polysaccharide was used to illustrate the capabilities of these workflows. The LC–MS chromatogram yielded primarily hexose polymers with unknown monosaccharide compositions (combinations of galactose and mannose), linkage, or branching (Fig. 8a). Fractions of the LC–MS were collected in 30 s increments leading to a total of 192 fractions. Each fraction was analyzed for monosaccharide compositions using a recently described rapid throughput automated workflow (Fig. 8b)<sup>54,57</sup>. The linkage information was similarly obtained using a separate recently reported workflow for comprehensive linkage analysis<sup>55,56</sup>. The monosaccharide analysis employed acid hydrolysis to produce monosaccharides that were subsequently labeled with 3-methyl-1-phenyl-2-pyrazoline-5-one (PMP). UHPLC–QQ MS analysis required 10 min of run time and measured the absolute abundances of 14 monosaccharides. The glycosidic linkages were analyzed with permethylation followed by acid hydrolysis and PMP labeling. The run time was 15 min and was capable of identifying over 100 distinct glycosidic linkages. The monosaccharide and linkage analysis methods were chosen for their speed, which allowed 96-well plates to be quickly analyzed, and their sensitivity, which allowed multiple analyses to be performed on single HPLC fractions. When the three platforms were integrated, they yielded structural information that could describe oligosaccharides with monosaccharide compositions and glycosidic linkage information. To obtain the final structural feature, the anomeric character of the linkages, the most abundant components were selected for NMR analysis using <sup>1</sup>H, <sup>13</sup>C NMR and a combination of techniques including COSY, HSQC, HMBC, and H2BC (Fig. 8c, Supplementary Data 5). The

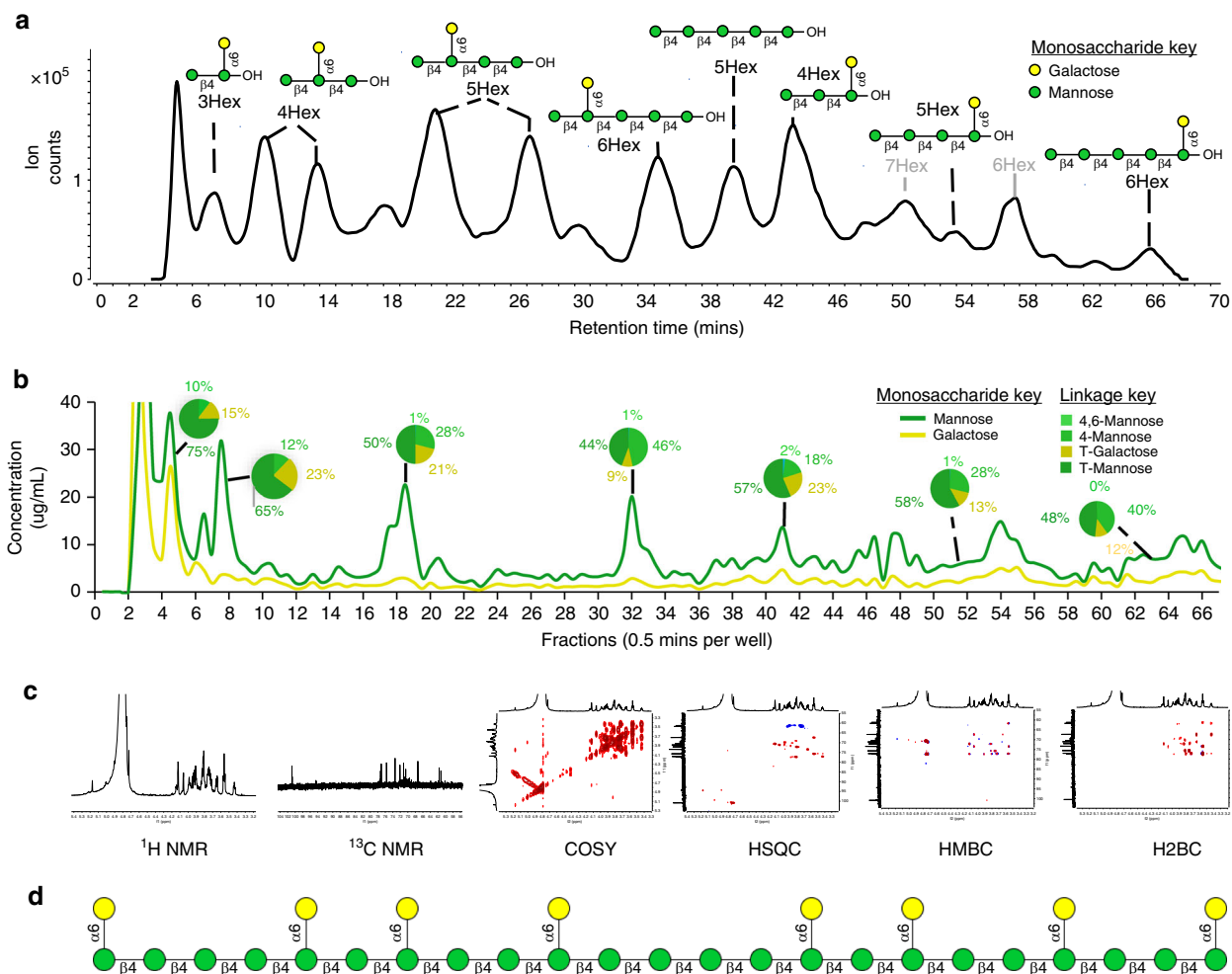




**Fig. 6 Analysis of in vivo microbe-carbohydrate interactions.** Annotated base peak chromatograms of oligosaccharides derived from FITDOG treatment of the ethanol precipitate fraction from (a) pear and (b) feces.



**Fig. 7 Schematic workflow of the de novo analysis of oligosaccharides derived from polysaccharide depolymerization.** Galactomannan is exemplified. Green circles represent mannose and yellow circles represent galactose.



**Fig. 8 De novo characterization of polysaccharides.** **a** Annotated total compound chromatograms with completely elucidated FITDOG oligosaccharides from galactomannan. **b** Constructed chromatogram representing the monosaccharide concentration galactose and mannose in each fraction and the relative glycosidic linkage composition of fractions shown in pie charts. **c** NMR analysis of FITDOG oligosaccharides. **d** A representation of the elucidated galactomannan structure by integrating oligosaccharide, monosaccharide, glycosidic linkage, and NMR analysis.

MS-obtained linkages and monosaccharide compositions greatly facilitated the NMR interpretation by limiting the resolution needed to determine the exact structures. This multi-platform approach yielded absolute oligosaccharide structures, which were used to recapitulate the parent polymeric structure (Fig. 8d).

## Discussion

The life sciences have built a remarkable understanding of the functions of cells and organisms driven by innovations in mapping the structures of life's biopolymers. As the most abundant biopolymer in nature, the lack of structural methods for polysaccharides has been glaring. Carbohydrate polymers are central to many biological processes. Cell-cell recognition is based on complicated glycan structures<sup>58,59</sup>; protection is provided by complex carbohydrates<sup>60,61</sup> and importantly, fuel and carbon flows through ecosystems primarily as complex carbohydrates<sup>62,63</sup>. The importance of glycan structures to interkingdom biology was vividly revealed by research on human milk that demonstrated that human mothers produce complicated, structure specific, oligosaccharides not to feed their infants but to feed specific bacteria within their infants<sup>11,32–34,64</sup>. While the prebiotic effects of oligosaccharides are more widely understood, their antimicrobial and antibiofilm activities are currently under intense investigations<sup>65–67</sup>. The challenge of understanding and managing the flows of carbohydrates

across all of biology is solving the complexity of carbohydrate structures also across biology. Yet this complexity is far greater than the other biopolymer classes nucleotides and proteins combined. Elucidating carbohydrate structures has been daunting, but a major impediment—the structure agnostic production of oligosaccharides—has now been solved.

The features of FITDOG make it suitable to measure polysaccharides with accuracy, sensitivity, and throughput. Central to the FITDOG analytical strategy is the ability to produce distinct oligosaccharides that are unique to their respective parent polysaccharides. The lack of mono- and disaccharides demonstrates that the catalysis is equivalent to an endo-, rather than an exoglycosidase enzyme. The results further indicate that the cleavages are not random thereby producing distinct oligosaccharides that are consistent with the cleavages of specific glycosidic bonds. The chemical basis underlying the glycosidic specificity of the FITDOG process, in effect favoring the cleavage of certain bonds, is not yet known. The regioselectivity of iron coordination within the polysaccharide chains likely guides the reaction. It has been suggested that the  $\text{Fe}^{3+}$  catalyzes the formation of the hydroperoxyl radical species, the site of chelation guides oxidation toward nearby hydroxyl groups that render nearby glycosidic bonds more prone to base-induced beta-elimination<sup>68</sup>. However, this mechanism is not entirely consistent with our observations under the specified conditions, as we do not observe the change in

mass associated with the dehydration and formation of the double bond. Determining the true mechanism of this reaction is a topic of ongoing research.

There are side reactions—some that can be suppressed and others that also have further utility. Excessive oxidation can occur including the formation of carboxylic acids in the nascent oligosaccharides. A singular feature of polysaccharide-based Fenton chemistry is that these side reactions can be minimized through properly selected experimental conditions that include the reaction times and the relative concentrations of  $\text{Fe}^{3+}$  and hydrogen peroxide. For this report, the conditions were optimized, and those presented here yielded the largest number of oligosaccharides while minimizing over oxidation for each polysaccharide tested. Factors that can lead to overoxidation included higher reaction temperatures and high concentrations of hydrogen peroxide. Changing the reaction time can increase or decrease the DP of the oligosaccharide products. Increasing the DP yields greater information regarding the placement of pendant chains from the main polymer chain.

The oligosaccharides produced by FITDOG can be used to identify known polysaccharides structures, which is analogous to peptide fingerprinting in proteomic analysis. This capability will simplify the analysis of complicated polysaccharide mixtures such as those found in food. As different dietary polysaccharides are known to affect the microbiota distinctly, this capability could significantly advance prebiotic/probiotic research.

The availability of a universal chemical approach for polysaccharide dissociation allows modern separation and mass spectrometry methods to be applied to the analysis of complex polysaccharides. We show that oligosaccharide fingerprinting is suitable for identifying known polysaccharides, and the reaction is readily scalable and can be expanded to more polysaccharides as new structures become available. Furthermore, the polysaccharide examples that were illustrated belong to several kingdoms including plants, bacteria, fungi, and algae. The method can also be used to elucidate the structures of new and uncharacterized polysaccharides. We showed that oligosaccharides produced by this method are characterizable by tandem MS methods where a more in-depth multi-platform strategy can systematically provide exact oligomer structure. While we have compared the current strategy to peptide fingerprinting, the analogy ends here. Structural analyses of oligosaccharides are significantly more difficult than that of peptides. Therefore, sequencing by tandem MS, a common tool for peptides, yields highly limited information with oligosaccharides, especially of those with limited monomeric mass diversity. However, we show that *de novo* structural elucidation of oligosaccharides, which previously involved resource extensive processes, can now be performed rapidly. With the structures elucidated, the oligosaccharides can be reconstructed to determine the overall structure of the parent polysaccharide (Fig. 8d). This method can also be applied for elucidating the structures of new polysaccharides in a much more rapid manner than is currently available.

Finally, this method may find additional utility beyond analysis. When scaled, this method provides the creation of new and previously unattainable oligosaccharides that can be further probed for bioactivity. We look forward to these results as we predict that the method will substantially advance carbohydrate research.

## Methods

**Experimental materials.** Sodium acetate, hydrogen peroxide ( $\text{H}_2\text{O}_2$ ), sodium hydroxide (NaOH), sodium borohydride ( $\text{NaBH}_4$ ), iron(III) sulfate pentahydrate ( $\text{Fe}_2(\text{SO}_4)_3$ ), and glacial acetic acid were purchased from Sigma-Aldrich (St. Louis, MO). Amylopectin was obtained from Carbosynth (Compton, UK). Amylose, xyloglucan, arabinoxylan, xylan, glucomannan, galactomannan, mannan, curdlan, galactan,  $\beta$ -glucan, arabinan, and lichenan were purchased from Megazyme (Bray, Ireland). Microcrystalline cellulose was purchased from ACROS Organics. Formic

Acid (FA) was purchased from Fisher Scientific (Belgium, UK). Acetonitrile (HPLC grade) was purchased from Honeywell (Muskegon, MI). Nano-pure water was used for all experiments.

**FITDOG generation of oligosaccharides.** A solution was prepared containing 95% (v/v) 40 mM sodium acetate buffer adjusted to pH 5 with glacial acetic acid, 5% (v/v) hydrogen peroxide (30% v/v), and 65 nM iron(III) sulfate. This mixture was vortexed and added to dry polysaccharide standards to make a final solution of 1 mg/ml. The reaction was incubated at 100 °C for 1 h. The reaction was quenched by adding half of the reaction volume of cold 2 M NaOH. Glacial acetic acid was added for neutralization.

Oligosaccharides were reduced by incubation with 1 M  $\text{NaBH}_4$  for 1 h at 65 °C. Oligosaccharides were isolated using nonporous graphitized carbon cartridges. Cartridges were washed with 80% acetonitrile and 0.1% (v/v) TFA in water. The oligosaccharides were loaded and washed with five column volumes of water. The oligosaccharides were eluted with 40% acetonitrile with 0.05% (v/v) TFA. Samples were completely dried by evaporative centrifugation and stored at  $-20$  °C until analysis.

**Sample preparation for fecal fingerprinting.** To separate the endogenous oligosaccharides from the polysaccharides, pear and feces samples underwent 80% ethanol precipitation overnight at  $-80$  °C. Samples were centrifuged at  $845 \times g$  for 20 min to pellet the polysaccharide fraction and partition it from the endogenous oligosaccharides in the supernatant. The pelleted fraction underwent FITDOG treatment to generate representative oligosaccharides. Oligosaccharides were then reduced and purified using the protocol described in the section above.

**Mass spectrometry analysis.** For analysis by MALDI-MS, 1  $\mu\text{l}$  was plated directly onto a stainless steel MALDI plate. To this, 0.3  $\mu\text{l}$  of 0.01 M NaCl and 0.7  $\mu\text{l}$  of 25 mg/ml 2, 5-dihydroxybenzoic acid was added and mixed within the pipet tip. The samples were dried under vacuum and analyzed on a Bruker UltraFlex extreme MALDI-tandem time-of-flight (MALDI-TOF/TOF) instrument. The instrument was operated in positive mode and 95% of max laser power.

Samples were reconstituted in nano-pure water before analysis by nano-HPLC-chip/Q-TOF MS. The system included two pumps, a capillary pump for sample loading and a nano pump for analytical separation. In this system, an Agilent 1200 series HPLC is coupled to an Agilent 6520 Q-TOF mass spectrometer through a chip cube interface. The chip contained a 40 nl enrichment column, and a 75  $\mu\text{m} \times 43$  mm analytical column; both columns had PGC as the stationary phase. Sample loading was done with 3% (v/v) acetonitrile/water + 0.1% formic acid at a flow rate of 4  $\mu\text{l}/\text{min}$ . Chromatographic separation was performed with a binary gradient of solvent A: (3% (v/v) acetonitrile/water + 0.1% formic acid) and solvent B: (90% acetonitrile/water + 0.1% formic acid) with a flow rate of 0.4  $\mu\text{l}/\text{min}$ . The gradient was run for 60 min, 1–5% B, 0–2 min; 5–30%, 2–33 min; 30–99%, 33–38 min; 99–99%, 38–48 min; 99–1%, 48–50 min; 1–1%, 50–60 min.

Data were collected in the positive mode and calibrated with internal calibrant ions ranging from  $m/z$  118.086 to 2721.895. Drying gas was set to 325 °C and with a flow rate of 5 l/min. The fragment, skimmer, and Octapole 1 RF voltages were set to 175, 60, and 750 V, respectively. Fragmentation was performed at a rate of 0.63 spectra/s. The collision energy was based upon the compound mass and expressed by the linear function (collision energy =  $1.8 \times (m/z) - 2.4$ ).

Samples were reconstituted in nano-pure water before HPLC Q-TOF MS analysis. The analytical separation was carried out using an Agilent 1260 Infinity II HPLC coupled to an Agilent 6530 Accurate-Mass Q-TOF MS. Chromatographic separation was performed on a 150 mm  $\times$  1 mm Hypercarb column from Thermo Scientific with a 5  $\mu\text{m}$  particle size. A binary gradient was employed which consisted of solvent A: (3% (v/v) acetonitrile/water + 0.1% formic acid) and solvent B: (90% acetonitrile/water + 0.1% formic acid). A 45 min gradient with a flow rate of 0.150 mL/min was used for chromatographic separation: 3–25% B, 0–15 min; 25–25% B, 15–18 min; 25–99% B, 18–30 min; 99–99% B, 30–32 min; 99–3% B, 32–34 min; 3–3% B, 34–45 min. Samples were run in the positive mode. Internal calibrant ions ranged from  $m/z$  121.051 to 2421.914. Drying gas temperature and flow rate were set to 150 °C and 11 l/min, respectively. Operation voltages for the fragment, skimmer, and octupole 1 RF were 175, 60, and 750 V, respectively. The acquisition rate was set to 0.63 spectra/s. When fragmentation was used, the linear function, collision energy =  $1.45 \times (m/z) - 3.5$ , was employed.

**Infant fecal samples.** The fecal samples were collected from a healthy, term 6-month-old infant who completed the UC Davis Infant Microbiome Nutrition and Development (Infant MiND) Study. The infant was exclusively breastfed at enrollment. The infant was assigned to consume pear (Earth's Best Stage 1) concurrently with breast milk for 7 days. After 7 days, parents were instructed to scrape the soiled diaper with sterile utensils, to place the fecal samples into sterile tubes, and to seal and store the samples in their kitchen freezers. The fecal samples were transported back to University of California Davis campus on dry ice and stored in  $-80$  °C before being analyzed. The University of California Davis Institutional Review Board approved all aspects of this study and written informed consent was obtained from the participants (Protocol ID: 919505). This study was registered on clinicaltrials.gov (NCT01817127).

**Fractionation of oligosaccharides.** Fractionation and detection of oligosaccharides was performed on an Agilent 1260 Infinity II series HPLC coupled to an Agilent 6530 Q-TOF mass spectrometer and a Teledyne Isco Foxy 200 fraction collector. Oligosaccharides were first separated on a  $150 \times 4.6$  mm Hypercarb column from Thermo Scientific with a  $5 \mu\text{m}$  particle size. A binary gradient was employed and consisted of solvent A: (3% (v/v) acetonitrile/water + 0.1% formic acid) and solvent B: (90% acetonitrile/water + 0.1% formic acid). A 90 min gradient with a flow rate of 1 ml/min was used for chromatographic separation: 5–12% B, 0–90 min; 12–99% B, 90–90.01 min; 99–99% B, 90.01–110 min; 99–5% B, 110–110.01 min; 5–5% B, 110.01–120 min. Post column, a 90:10 flow splitter partitioned the larger stream to the fraction collector and the smaller to the Q-TOF mass spectrometer. Data from the QTOF was collected in the positive mode and calibrated with internal calibrant ions ranging from  $m/z$  118.086–2721.895. Drying gas was set to  $150^\circ\text{C}$  and with a flow rate of 1 l/min. The fragment, skimmer, and Octupole 1 RF voltages were set to 75, 60, and 750 V, respectively. Fragmentation was performed at a rate of 1 spectra/s. The collision energy was based upon the compound mass and expressed by the linear function (collision energy =  $1.3 \times (m/z) - 3.5$ ). Fractions were collected on 96-well plates at a rate of 30 s per fraction. Collected fractions were dried to completion under vacuum centrifugation and reconstituted in 100  $\mu\text{l}$  of nano-pure water. A 10  $\mu\text{l}$  aliquot was transferred to a separate 96-well plate for monosaccharide composition analysis, while the remaining 90  $\mu\text{l}$  underwent glycosidic linkage analysis.

#### Monosaccharide composition analysis of fractionated oligosaccharides.

Monosaccharide analysis was adapted by Amicucci et al. with the following modifications<sup>54</sup>. Briefly, fractionated oligosaccharides underwent acid hydrolysis with 4 M TFA for 2 h at  $100^\circ\text{C}$ . The samples were dried to completion by vacuum centrifugation. Samples and monosaccharide standards (0.001–100  $\mu\text{g}/\text{ml}$ ) underwent derivatization with 0.2 M PMP in methanol and 28%  $\text{NH}_4\text{OH}$  at  $70^\circ\text{C}$  for 30 min. Derivatized products were dried to completion under vacuum centrifugation and reconstituted in nano-pure water. Excess PMP was removed with chloroform extraction. The aqueous layer was analyzed by an Agilent 1290 infinity II UHPLC coupled to an Agilent 6495A QqQ MS employing dynamic multiple reaction monitoring (MRM) mode. An external standard curve was used for absolute quantitation of each monosaccharide in the fractions.

**Linkage analysis of fractionated oligosaccharides.** Linkage analysis was adapted from Galermo et al. with the following modifications<sup>55,56</sup>. Briefly, fractionated oligosaccharides and a pool of oligosaccharide standards were reacted with saturated NaOH and iodomethane in DMSO. Residual NaOH and DMSO were removed by extraction with DCM and water. The DCM layer was dried to completion under vacuum centrifugation. Samples were hydrolyzed and derivatized in the same manner as the monosaccharide analysis. Samples did not undergo chloroform extraction and were reconstituted in 70% (v/v) methanol/water. Fractions were analyzed on an Agilent 1290 infinity II UHPLC coupled to an Agilent 6495A QqQ MS ran in MRM mode. A pool of oligosaccharide standards was used to assign the glycosidic linkages present.

**NMR analysis of fractionated oligosaccharides.** NMR spectra were recorded at 303 K on a Bruker AVANCE III 800 MHz spectrometer equipped with a 5 mm Bruker CPTCI cryoprobe. Samples were obtained by combining ten collections of fractionated oligosaccharides of the same components verified with HPLC-QTOF MS and MS/MS data. Based on the monosaccharide and linkage information, the most abundant oligosaccharides were selected for NMR analysis. Each of these selected pooled fractions were dried with vacuum centrifugation before being reconstituted in 0.4 mL of  $\text{D}_2\text{O}$ , and measured using 1D  $^1\text{H}$  (relaxation delay (D1) 2 s; number of scans (NS) 128),  $^{13}\text{C}$  NMR (D1 1.5 s; NS 6000–15000), and 2D  $^1\text{H}$ – $^1\text{H}$  COSY (D1 1.5 s; NS 8),  $^1\text{H}$ – $^{13}\text{C}$  HSQC (D1 2 s; NS 4), HMBC (D1 1.5 s; NS 16), and H2BC (D1 1.5 s; NS 16). The spectra were then processed with Bruker TopSpin 3.2 and analyzed with MestReNova. The experimental chemical shifts, along with the required monosaccharide and linkage data, were calculated using the CASPER program<sup>69</sup>, where the oligosaccharide structures, including anomeric characters of linkages, were predicted with ranking scores.

**Reporting summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

#### Data availability

The authors declare that the data supporting the findings of the study are available in the paper and in Supplementary Data 1–5. All other data are available from the corresponding author on reasonable request.

Received: 5 May 2020; Accepted: 15 July 2020;

Published online: 07 August 2020

#### References

- Veronese, N. et al. Dietary fiber and health outcomes: an umbrella review of systematic reviews and meta-analyses. *Am. J. Clin. Nutr.* **107**, 436–444 (2018).
- Bingham, S. A. et al. Dietary fibre in food and protection against colorectal cancer in the European Prospective Investigation into Cancer and Nutrition (EPIC): an observational study. *Lancet* **361**, 1496–1501 (2003).
- Anderson, J. W. et al. Health benefits of dietary fiber. *Nutr. Rev.* **67**, 188–205 (2009).
- Salmeron, J. et al. Dietary fiber, glycemic load, and risk of non—insulin-dependent diabetes mellitus in women. *Jama* **277**, 472–477 (1997).
- Ruhaak, L. R. et al. Mass spectrometry approaches to glycomic and glycoproteomic analyses. *Chem. Rev.* **118**, 7886–7930 (2018).
- Björndal, H., Lindberg, B. & Svensson, S. Mass spectrometry of partially methylated alditol acetates. *Carbohydr. Res.* **5**, 433–440 (1967).
- HAKOMORI, S.-I. A rapid permethylation of glycolipid, and polysaccharide catalyzed by methylsulfinyl carbanion in dimethyl sulfoxide. *J. Biochem.* **55**, 205–208 (1964).
- Albersheim, P. et al. A method for the analysis of sugars in plant cell-wall polysaccharides by gas-liquid chromatography. *Carbohydr. Res.* **5**, 340–345 (1967).
- Amicucci, M. J., Nandita, E. & Lebrilla, C. B. Function without structures: the need for in-depth analysis of dietary carbohydrates. *J. Agric. Food Chem.* **67**, 4418–4424 (2019).
- Kailemia, M. J. et al. Oligosaccharide analysis by mass spectrometry: a review of recent developments. *Anal. Chem.* **86**, 196–212 (2013).
- Wu, S. et al. Development of an annotated library of neutral human milk oligosaccharides. *J. Proteome Res.* **9**, 4138–4151 (2010).
- Stavenhagen, K., Kolarich, D. & Wuhrer, M. Clinical glycomics employing graphitized carbon liquid chromatography–mass spectrometry. *Chromatographia* **78**, 307–320 (2015).
- Sun, X. et al. Capillary electrophoresis–mass spectrometry for the analysis of heparin oligosaccharides and low molecular weight heparin. *Anal. Chem.* **88**, 1937–1943 (2016).
- Wu, S. et al. Annotation and structural analysis of sialylated human milk oligosaccharides. *J. Proteome Res.* **10**, 856–868 (2011).
- James, P. et al. Protein identification by mass profile fingerprinting. *Biochem. Biophys. Res. Commun.* **195**, 58–64 (1993).
- Pappin, D. J., Hojrup, P. & Bleasby, A. J. Rapid identification of proteins by peptide-mass fingerprinting. *Curr. Biol.* **3**, 327–332 (1993).
- Thiede, B. et al. Peptide mass fingerprinting. *Methods* **35**, 237–247 (2005).
- Hunt, D. F. et al. Protein sequencing by tandem mass spectrometry. *Proc. Natl Acad. Sci.* **83**, 6233–6237 (1986).
- Hogan, J. D. et al. Software for peak finding and elemental composition assignment for glycosaminoglycan tandem mass spectra. *Mol. Cell. Proteom.* **17**, 1448–1456 (2018).
- Kailemia, M. J. et al. Complete mass spectral characterization of a synthetic ultralow-molecular-weight heparin using collision-induced dissociation. *Anal. Chem.* **84**, 5475–5478 (2012).
- Auray-Blais, C. et al. UPLC-MS/MS detection of disaccharides derived from glycosaminoglycans as biomarkers of mucopolysaccharidoses. *Anal. Chim. Acta* **936**, 139–148 (2016).
- Feyerabend, T. B. et al. Heparan sulfate C5-epimerase is essential for heparin biosynthesis in mast cells. *Nat. Chem. Biol.* **2**, 195–196 (2006).
- Sarrazin, S., Lamanna, W. C. & Esko, J. D. Heparan sulfate proteoglycans. *Cold Spring Harb. Perspect. Biol.* **3**, a004952 (2011).
- Talmadge, K. W. et al. The structure of plant cell walls: I. The macromolecular components of the walls of suspension-cultured sycamore cells with a detailed analysis of the pectic polysaccharides. *Plant Physiol.* **51**, 158–173 (1973).
- Günl, M., Kraemer, F. & Pauly, M. Oligosaccharide mass profiling (OLIMP) of cell wall polysaccharides by MALDI-TOF/MS, in *The Plant Cell Wall*. p. 43–54 (Springer, 2011).
- Ballance, S. et al. A re-examination and partial characterisation of polysaccharides released by mild acid hydrolysis from the chlorite-treated leaves of *Sphagnum papillosum*. *Carbohydr. Polym.* **67**, 104–115 (2007).
- Yang, B. et al. Mechanism of mild acid hydrolysis of galactan polysaccharides with highly ordered disaccharide repeats leading to a complete series of exclusively odd-numbered oligosaccharides. *FEBS J.* **276**, 2125–2137 (2009).
- Amicucci, M. J. et al. A strategy for structural elucidation of polysaccharides: elucidation of a maize mucilage that harbors diazotrophic bacteria. *Anal. Chem.* **91**, 7254–7265 (2019).
- Malik, V. S. et al. Sugar-sweetened beverages and risk of metabolic syndrome and type 2 diabetes: a meta-analysis. *Diabetes Care* **33**, 2477–2483 (2010).
- Johnson, R. J. et al. Potential role of sugar (fructose) in the epidemic of hypertension, obesity and the metabolic syndrome, diabetes, kidney disease, and cardiovascular disease. *Am. J. Clin. Nutr.* **86**, 899–906 (2007).
- Barratt, M. J. et al. The gut microbiota, food science, and human nutrition: a timely marriage. *Cell Host Microbe* **22**, 134–141 (2017).



32. Sela, D. et al. The genome sequence of *Bifidobacterium longum* subsp. *infantis* reveals adaptations for milk utilization within the infant microbiome. *Proc. Natl Acad. Sci.* **105**, 18964–18969 (2008).
33. Marcobal, A. et al. Bacteroides in the infant gut consume milk oligosaccharides via mucus-utilization pathways. *Cell Host Microbe* **10**, 507–514 (2011).
34. Zivkovic, A. M. et al. Human milk glycomiome and its impact on the infant gastrointestinal microbiota. *Proc. Natl Acad. Sci.* **108**(Supplement 1), 4653–4658 (2011).
35. Cohen, E. et al. Statistical review of US macronutrient consumption data, 1965–2011: Americans have been following dietary guidelines, coincident with the rise in obesity. *Nutrition* **31**, 727–732 (2015).
36. Caporaso, J. G. et al. QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* **7**, 335 (2010).
37. Loman, N. J. & Pallen, M. J. Twenty years of bacterial genome sequencing. *Nat. Rev. Microbiol.* **13**, 787–794 (2015).
38. Johnson, A. J. et al. Daily sampling reveals personalized diet-microbiome associations in humans. *Cell Host Microbe* **25**, 789–802.e5 (2019).
39. Carmody, R. N. et al. Cooking shapes the structure and function of the gut microbiome. *Nat. Microbiol.* **4**, 2052–2063 (2019).
40. Gehrig, J. L. et al. Effects of microbiota-directed foods in gnotobiotic animals and undernourished children. *Science* **365**, eaau4732 (2019).
41. Hamaker, B. R. & Tuncil, Y. E. A perspective on the complexity of dietary fiber structures and their potential effect on the gut microbiota. *J. Mol. Biol.* **426**, 3838–3850 (2014).
42. Bindels, L. B. et al. Towards a more comprehensive concept for prebiotics. *Nat. Rev. Gastroenterol. Hepatol.* **12**, 303 (2015).
43. Martens, E. C. et al. The devil lies in the details: how variations in polysaccharide fine-structure impact the physiology and evolution of gut microbes. *J. Mol. Biol.* **426**, 3851–3865 (2014).
44. Rogowski, A. et al. Glycan complexity dictates microbial resource allocation in the large intestine. *Nat. Commun.* **6**, 7481 (2015).
45. Flint, H. J. et al. Microbial degradation of complex carbohydrates in the gut. *Gut Microbes* **3**, 289–306 (2012).
46. Sonnenburg, E. D. et al. Specificity of polysaccharide use in intestinal bacteroides species determines diet-induced microbiota alterations. *Cell* **141**, 1241–1252 (2010).
47. Hayashi, T. & Maclachlan, G. Pea xyloglucan and cellulose: I. Macromolecular organization. *Plant Physiol.* **75**, 596–604 (1984).
48. Fry, S. C. The structure and functions of xyloglucan. *J. Exp. Bot.* **40**, 1–11 (1989).
49. Eckardt, N. A. Role of xyloglucan in primary cell walls. *Am. Soc. Plant Biol.* **20**, 1421–1422 (2008).
50. Uchida, K. & Kawakishi, S. Interaction of (1→4)- and (1→6)-linked disaccharides with the fenton reagent under physiological conditions. *Carbohydr. Res.* **173**, 89–99 (1988).
51. Mandalari, G. et al. Fractionation and characterisation of arabinoxylans from brewers' spent grain and wheat bran. *J. Cereal Sci.* **42**, 205–212 (2005).
52. Wood, P. J. Cereal  $\beta$ -glucans in diet and health. *J. Cereal Sci.* **46**, 230–238 (2007).
53. Gruppen, H. et al. Mild isolation of water-insoluble cell wall material from wheat flour: composition of fractions obtained with emphasis on non-starch polysaccharides. *J. Cereal Sci.* **9**, 247–260 (1989).
54. Amicucci, M. J. et al. A rapid-throughput adaptable method for determining the monosaccharide composition of polysaccharides. *Int. J. Mass Spectrom.* **438**, 22–28 (2019).
55. Galermo, A. G. et al. Liquid chromatography–tandem mass spectrometry approach for determining glycosidic linkages. *Anal. Chem.* **90**, 13073–13080 (2018).
56. Galermo, A. G. et al. Development of an extensive linkage library for characterization of carbohydrates. *Anal. Chem.* **91**, 13022–13031 (2019).
57. Xu, G. et al. Revisiting monosaccharide analysis–quantitation of a comprehensive set of monosaccharides using dynamic multiple reaction monitoring. *Analyst* **143**, 200–207 (2018).
58. Phillips, M. L. et al. ELAM-1 mediates cell adhesion by recognition of a carbohydrate ligand, sialyl-Lex. *Science* **250**, 1130–1132 (1990).
59. Park, D. et al. Salmonella typhimurium enzymatically landscapes the host intestinal epithelial cell (IEC) surface glycome to increase invasion. *Mol. Cell. Proteom.* **15**, 3653–3664 (2016).
60. Mack, D. et al. Extracellular MUC3 mucin secretion follows adherence of *Lactobacillus* strains to intestinal epithelial cells in vitro. *Gut* **52**, 827–833 (2003).
61. Newburg, D. S., Ruiz-Palacios, G. M. & Morrow, A. L. Human milk glycans protect infants against enteric pathogens. *Annu. Rev. Nutr.* **25**, 37–58 (2005).
62. Haigler, C. H., et al. Carbon partitioning to cellulose synthesis, in *Plant Cell Walls*. p. 29–51, (Springer, 2001).
63. Jones, D. L., Nguyen, C. & Finlay, R. D. Carbon flow in the rhizosphere: carbon trading at the soil–root interface. *Plant Soil* **321**, 5–33 (2009).
64. Charbonneau, M. R. et al. Sialylated milk oligosaccharides promote microbiota-dependent growth in models of infant undernutrition. *Cell* **164**, 859–871 (2016).
65. Lin, A. E. et al. Human milk oligosaccharides inhibit growth of group B *Streptococcus*. *J. Biol. Chem.* **292**, 11243–11249 (2017).
66. Craft, K. M., Thomas, H. C. & Townsend, S. D. Interrogation of human milk oligosaccharide fucosylation patterns for antimicrobial and antibiofilm trends in group B streptococcus. *ACS Infect. Dis.* **4**, 1755–1765 (2018).
67. Craft, K. M., Thomas, H. C. & Townsend, S. D. Sialylated variants of lacto-N-tetraose exhibit antimicrobial activity against Group B *Streptococcus*. *Org. Biomol. Chem.* **17**, 1893–1900 (2019).
68. Hosoya, T. et al. Insights into degradation pathways of oxidized anhydroglucose units in cellulose by  $\beta$ -alkoxy-elimination: a combined theoretical and experimental approach. *Cellulose* **25**, 3797–3814 (2018).
69. Jansson, P.-E., Stenutz, R. & Widmalm, G. Sequence determination of oligosaccharides and regular polysaccharides using NMR spectroscopy and a novel Web-based version of the computer program CASPER. *Carbohydr. Res.* **341**, 1003–1010 (2006).

## Acknowledgements

This research was supported by grants provided by Mars Incorporated.

## Author contributions

M.J.A., E.N., and J.J.C. drafted the paper. M.J.A., E.N., J.J.C., C.B.L., designed experiments. M.J.A., E.N., J.J.C., D.P., A.G.G., S.C. conducted experiments. J.T.S., J.B.G., D.A.M. designed the in vivo weaning study. C.B.L., supervised the project. All authors edited the paper.

## Competing interests

M.J.A., E.N., J.J.C., A.G.G., D.P., D.A.M., J.B.G., and C.B.L. have filed two patents on the utility of this method for producing bioactive oligosaccharides. M.J.A., D.A.M., J.B.G., and C.B.L. are founders of BCD Bioscience, a company that will market bioactive oligosaccharides. The other authors declare no competing interests.

## Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41467-020-17778-1>.

Correspondence and requests for materials should be addressed to C.B.L.

Peer review information *Nature Communications* thanks the anonymous reviewer(s) for their contribution to the peer review of this work.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020