Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Chemical Biology

Interrogating gut bacterial genomes for discovery of novel carbohydrate degrading enzymes

Ana S Luis and Eric C Martens

Individual human gut bacteria often encode hundreds of enzymes for degrading different polysaccharides. Identification of co-localized and co-regulated genes in these bacteria has been a successful approach to identify enzymes that participate in full or partial saccharification of complex carbohydrates, often unmasking novel catalytic activities. Here, we review recent studies that have led to the discovery of new activities from gut bacteria and summarize a general scheme for identifying gut bacteria with novel catalytic abilities, locating the enzymes involved and investigating their activities in detail. The strength of this approach is amplified by the availability of abundant genomic and metagenomic data for the human gut microbiome, which facilitates comparative approaches to mine existing data for new or orthologous enzymes.

Address
University of Michigan Medical School, Ann Arbor, MI 48109, USA

Corresponding author: Martens, Eric C (emartens@umich.edu)

Current Opinion in Chemical Biology 2018, **47C**:126–133

This review comes from a themed issue on **Mechanistic Biology**

Edited by **John A Gerlt** and **Tobias J Erb**

For a complete overview see the Issue and the Editorial

Available online 14th October 2018

https://doi.org/10.1016/j.cbpa.2018.09.012

Introduction

Symbiotic bacteria inhabiting the distal guts of humans and other mammals are responsible for the majority of non-starch dietary fiber polysaccharide digestion [1,2]. As such, the genomes of individual gut bacteria often harbor many dozens or hundreds of individual carbohydrate-active enzymes (CAZymes) to carry out this catalysis [3•,4]. While the general structural features of plant cell wall and other fiber polysaccharides have been well characterized, the range of enzymatic activities that degrade them, especially very complex structures [5••] or botanical source-specific variants [6,7], are still being defined. In this review, we highlight how detailed investigations of individual human gut bacteria have provided a useful path towards discovery of new CAZyme activities. This emerging knowledge is critical to fully understand the range and evolution of these catalytic abilities and also provides fundamental insight into human digestion.

Carbohydrate active enzyme classification

Degradation of the array of polysaccharides available in the human gut is possible through the combination of various CAZymes with different specificities. These enzymes fall into two main groups, glycoside hydrolases (GHs) and polysaccharide lyases (PLs), although carbohydrate esterases (CEs) also play important roles on some substrates. GHs catalyze the hydrolysis of glycosidic bonds by either a single or a double-displacement mechanism leading to inversion or retention of the stereochemistry of the anomeric carbon, respectively [8]. PLs are non-hydrolytic enzymes that cleave uronic acid-containing polysacharides via a β-elimination mechanism yielding an unsaturated uronic acid at the non-reducing end [9,10]. CAZymes are classified into sequence-based families via the continuously updated CAZy database [11]. The members of a family display the same fold, same catalytic apparatus and mechanism [12,13]. One of the few exceptions is GH97 that contains enzymes with two different mechanisms [14]. GH families displaying the same fold, mechanism and catalytic residues but a low sequence similarity are grouped into clans [11,13]. Additionally, enzyme topology is related to endo-modes or exo-modes of action. Endo-acting enzymes cleave internal linkages within the substrate and have an open cleft; whereas, exo-acting hydrolases cleave at end of the chain and present an active site pocket that only accommodates the terminal sugar [15].

A 'functional microbiology' approach to discovering new carbohydrate modifying enzymes

Whole genome transcriptional profiling techniques, so called 'functional genomics', are now routine since the advent of microarray and subsequent next-generation sequencing based methods (e.g. RNAseq). These techniques allow researchers to quickly locate genes that are activated under different physiological conditions, which for symbiotic human gut bacteria often involves large changes in gene expression in response to polysaccharide cues [3•,16••,17,18,19••]. Using this approach, gene clusters that encode CAZymes and other functions related to polysaccharide degradation have been identified in a number of gut bacteria, spanning most common dietary polysaccharides [1,20•] and even more exotic structures such as seaweed-derived agarose and porphyran [21,22••]. Many studies have been conducted in gut Bacteroidetes,

a prominent saccharolytic phylum in the human colon, in which individual members typically group all the genes required for the recognition, uptake and degradation of a specific glycan into co-regulated polysaccharide utilization loci (PULs), gene clusters that encode, in addition to CAZymes, TonB-dependent transporters, SusD family binding proteins and other functions [17]. However, additional studies are also emerging in members of the abundant Firmicutes, Actinobacteria and Verrucomicrobia phyla, revealing different types of gene arrangements and encoded enzymatic machinery such as enzyme-associated ABC-transport systems and cellulosome-like systems [23,24•,25,26].

Since many complex dietary polysaccharides contain several different sugars and glycosidic linkages—therefore requiring multiple degradative enzymes—identifying transcriptionally active gene clusters with multiple CAZymes often reveals suites of enzymes that work together and often uncovers genes encoding enzymes with novel catalytic roles (Figure 1). In addition, similar polysaccharides sometimes vary in subtle ways between botanical sources; identifying a gene cluster associated with degradation of one form of a polysaccharide enables searching of other genomic or metagenomic data to find orthologous, but partially variant gene clusters that may have adapted to other forms of the same polysaccharide. An example of this is utilization of the dicotyledonous hemicellulose xyloglucan. Growth, transcriptional profiling and molecular genetics with a single strain of *Bacteroides ovatus* (strain ATCC8483) revealed a 12-gene PUL (xyloglucan utilization locus; XyGUL), containing 8 different CAZymes, that is necessary for growth on a simplified form of xyloglucan from tamarind seed [19••,27••]. While the originally identified *B. ovatus* PUL encoded enzymes for removing α-L-arabinofuranosyl side-chain linkages present in solanaceous plants (tomatoes, eggplants), it did not encode enzymes for removal of α-fucosyl linkages attached to the xyloglucans present in leafy greens like lettuce [28]. However, comparative genomics using the originally discovered *B. ovatus* XyGUL revealed the presence of homologous and often highly syntenic PULs in other sequenced gut bacteria that also encode enzymes in known α-fucosidase families (glycoside hydrolase families 29 and 95), suggesting adaptation to the modifications specific to lettuce. Moreover, the increasing availability of high-quality metagenomic sequencing data from human fecal samples, like those from the Human Microbiome Project, enables even deeper discovery of novel locus variants without the need for direct cultivation [27••]. Taken together, the scheme shown in Figure 1 illustrates a path to connect the ability of pure or mixed enrichment cultures to grow on complex polysaccharides with the genes involved and then further compare across genomes and metagenomes to discover variant loci. For several very complex plant cell wall polysaccharides discussed below, this process has led to discovery of new catalytic activities that have only been uncovered in human gut bacteria in the past few years.

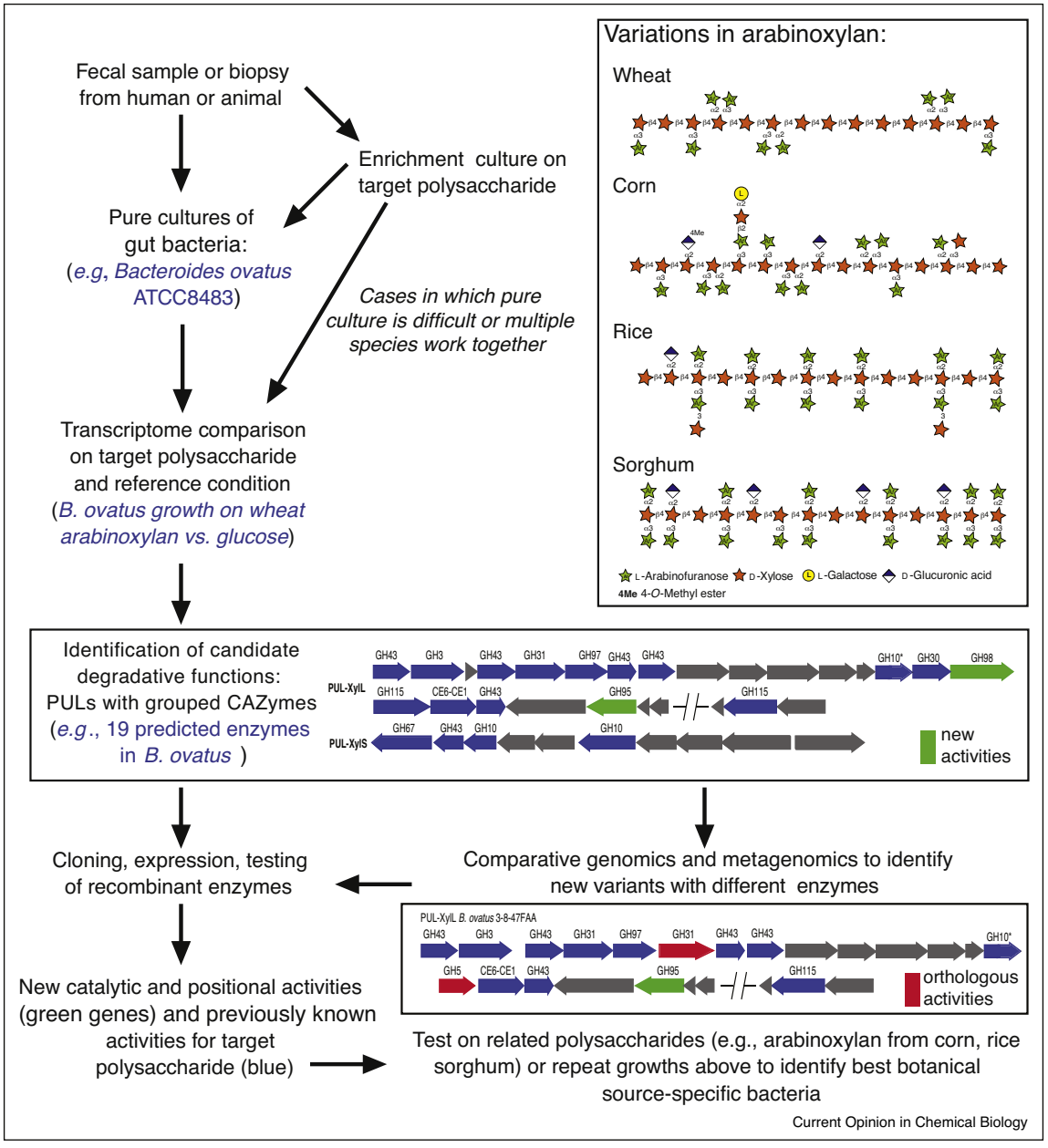
Plant cell wall xylan: new activities inside existing families

During growth on wheat xylan, a major component of monocotyledonous plant cell walls that exhibits structural variation between plant sources [29,30], *B. ovatus* upregulates two PULs [19••] that encode 19 different predicted enzymes, although one was shown to be catalytically inactive (Figure 1). In addition to known xylanolytic activities, two of the enzymes required for xylan utilization revealed new activities inside previously characterized GH families (Figures 1–2). A member of family GH98, previously known to only include blood group A- or B-cleaving *endo*-β-D-galactosidases [31], was shown to generate long xylo-oligosaccharides from corn arabinoxylan, requiring the D-xylose (D-Xyl) of the leaving group to be doubly linked to both α1,2-L-arabinofuranose (L-Ara*f*) and α1,3-D-Xyl [7]. A novel GH95 enzyme was found to cleave L-gal). Previously, this family only contained α-L-fucosidases [32]. The catalytic apparatus is fully conserved inside members of this family. Indeed, a single difference in the −1 subsite, a Thr in the α-L-galactosidase that is replaced by a His in α-L-fucosidases, is speculated to be the substrate specificity determinant inside family GH95 (Figure 2a) [7].

Novel families of rhamnosidases

Several plant cell wall polysaccharides that are common in the human diet contain α-L-rhamnose (L-Rha) and a number of *Bacteroides thetaiotaomicron* PULs containing candidate rhamnosidases have been previously discovered [19••]. Cleavage of rhamnosidic linkages is enzymatically challenging because the axial C2-hydroxyl causes a destabilizing *syn*-diaxial orientation in the transition state [33]. Until recently only GH78, 90 and GH106 families have been shown to contain α-L-rhamnosidases [34,35,36]. The discovery of two new families active on L-Rha-α1,4-D-glucuronic acid (D-GlcA) moieties that cap arabinogalactan polysaccharide side chains revealed novel rhamnosidase features (Table 1, Figure 2b). Members of PL27, that are unrelated to any of the previously known PL or GH families, overcome the challenge of cleaving L-Rha through the β-elimination mechanism. Briefly, the catalytic base (Tyr) abstracts the proton at C5 of D-GlcA, generating a double bond between C4 and C5 (Δ4,5-anhydroglucuronic acid) that leads to the elimination of the L-Rha at C4 [37]. Additionally, the recently described GH145 family is proposed to use His as the only catalytic residue (nucleophile) and the carboxylate of the leaving group (D-GlcA) fulfills the catalytic acid/base role. A completely novel feature of these enzymes is the active site location at the posterior surface of the seven bladed β-propeller and not in the highly conserved anterior pocket, suggesting that these enzymes might have an additional ancestral activity that is not yet known [38].

Figure 1

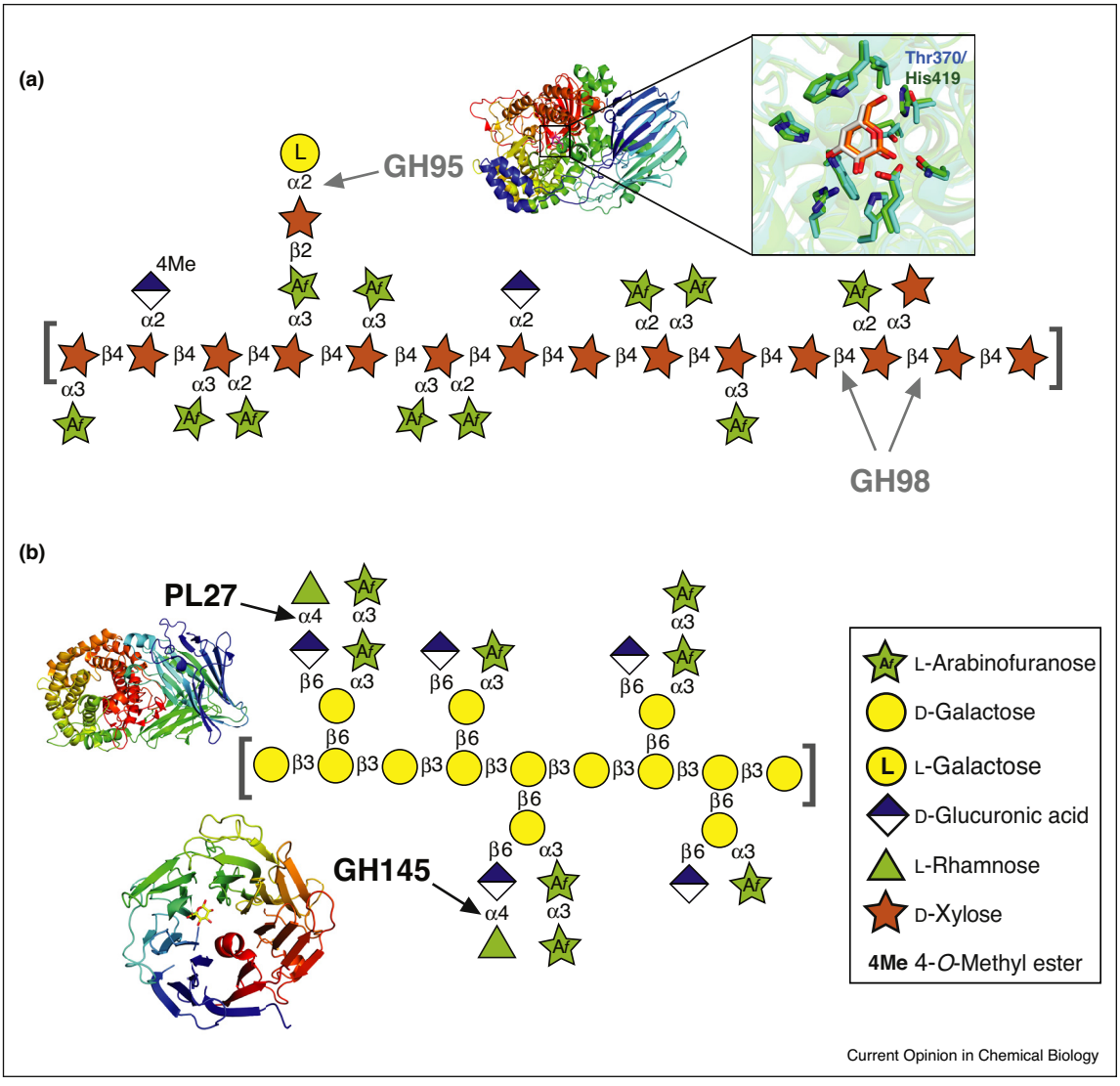


A scheme for identifying gut bacterial functions involved in complex carbohydrate degradation. Beginning with a fecal sample or intestinal biopsy—the latter of which could contain different bacteria than those contained in feces [49]—the branched flow chart outlines the process of identifying genes involved and investigating their function. Because many pure cultures exist for human intestinal bacteria, these have been more commonly used as starting points; although, historical culture collections are unlikely to represent the full diversity of bacteria present in the gut. As an example, details for *Bacteroides ovatus* xylan degradation are provided in blue text and gene illustrations [7]. Enzyme coding genes are labeled blue, green or red, the latter two denoting novel and orthologous activities, respectively. Family numbers are shown above each enzyme-coding gene. The inset in the upper right shows the linkages present in xylan from different botanical sources.

Novel enzymes in pectin degradation

The identification of pectin–responsive PULs revealed many candidate enzymes for digesting highly complex rhamnogalacturonan I and II structures (RG-I and RG-II, respectively), which are major components in the human diet and thought to be among the most complex natural polysaccharides known [4,19••,39••]. These enzymes have been extensively studied in the last few years, revealing a number of insights into the complex enzymatic pathway required to degrade these substrates (Figure 3). A clear example of this complexity is provided by the *B. thetaiotaomicron* RG-I PUL that encodes 9 enzymes to degrade the

Figure 2



Novel enzymes active on xylan and arabinogalactan. **(a)** Corn glucuronoarabinoxylan contains a β1,4-D-xylose backbone decorated with L-arabinofuranose, (methyl) D-glucoronic acid, D-xylose and L-galactose [7]. **(b)** Arabinogalactan type-II polysaccharides are mainly associated with proteins and present a β1,3-D-galactose (Gal) backbone that can be decorated at with D-Gal, L-arabinose, D-glucuronic acid and L-rhamnose [38]. For both polysaccharides, novel families and new activities inside previously described families are shown in black and grey, respectively, and the arrows point to the linkage cleaved by the respective enzyme. The schematic representations of new enzymes are ramp-coloured from blue (N-terminus) to red (C-terminus). In panel a, the inset schematic shows a representation of BACOVA_03485 (PDB: 4UFC) α-L-galactosidase and overlay of respective active site (blue) with the α-L-fucosidase from *Bifidobacterium bifidum* (green, PDB: 2EAE), bound to L-galactose (orange) and L-fucose (grey), respectively [7]. The only two residues that are not conserved between these two enzymes are highlighted. Sugars are shown using the Consortium for Functional Glycomics notation; GH, glycoside hydrolase; PL, polysaccharide lyase.

L-Rha-D-galacturonic acid (D-GalA) backbone containing only two linkages [39••]. The key enzyme in this system is a PL9 located at the cell surface that cleaves the backbone and generates unsaturated oligosaccharides [39••]. This activity represented the first report of a rhamnogalacturonan lyase in family PL9, previously known to contain only polygalacturonate lyases [15]. The predicted catalytic base (Lys) and the calcium-binding site (four Asp) are fully conserved in PL9 enzymes (Figure 3a). However, the structure in complex with reaction products revealed a second calcium-binding site only present in the rhamnogalacturonan lyase and important for catalysis. Additional substrate recognition residues in distal subsites are also variable between PL9 enzymes, reflecting the differences in charge of the target substrates [39••].

Characterization of the degradation mechanism of RG-I side chains (galactan and arabinan) revealed two new GH

Table 1

Novel enzyme families or activities recently discovered in human gut bacteria

Family	Activity	Active residues	Mechanism	Structure	Ref
GH137	β-L-arabinofuranosidase	Glu Glu	Unknown	5-bladed β-propeller	[5••]
GH138	α-D-galacturonidase	Unknown	Unknown	Unknown	[5••]
GH139	α-2-*O*-methyl-L-fucosidase[a]	Unknown	Unknown	Unknown	[5••]
GH140	Endo-apiosidase	Asp Glu	Retaining	$(\beta/\alpha)_8$-barrel	[5••]
GH141	α-L-fucosidase	Asp (nucleophile) Asp (acid-base)	Retaining[b]	Paralel β-helix	[5••]
GH142	β-L-arabinofuranosidase	Glu	Unknown	$(\alpha/\alpha)_6$-barrel	[5••]
GH143	β-D-DHA-hydrolase[a]	Try (nucleophile) Glu (acid-base)	Retaining[c]	5-bladed β-propeller	[5••]
GH145	α-L-rhamnohydrolase	His (nucleophile)	Retaining	7-bladed β-propeller	[40••]
GH146	β-L-arabinofuranosidase	Cys (nuclephile) Glu (acid/base)	Retaining[d]	$(\alpha/\alpha)_6$-barrel	[40••]
GH147	β-D-galactosidase	Glu/Glu	Retaining[±]	$(\beta/\alpha)_8$-barrel[e]	[40••]
PL27	L-rhamnose-α-1,4-D-glucuronate lyase	Tyr (catalytic base)	β-elimination	$(\alpha/\alpha)_6$-barrel	[40••]

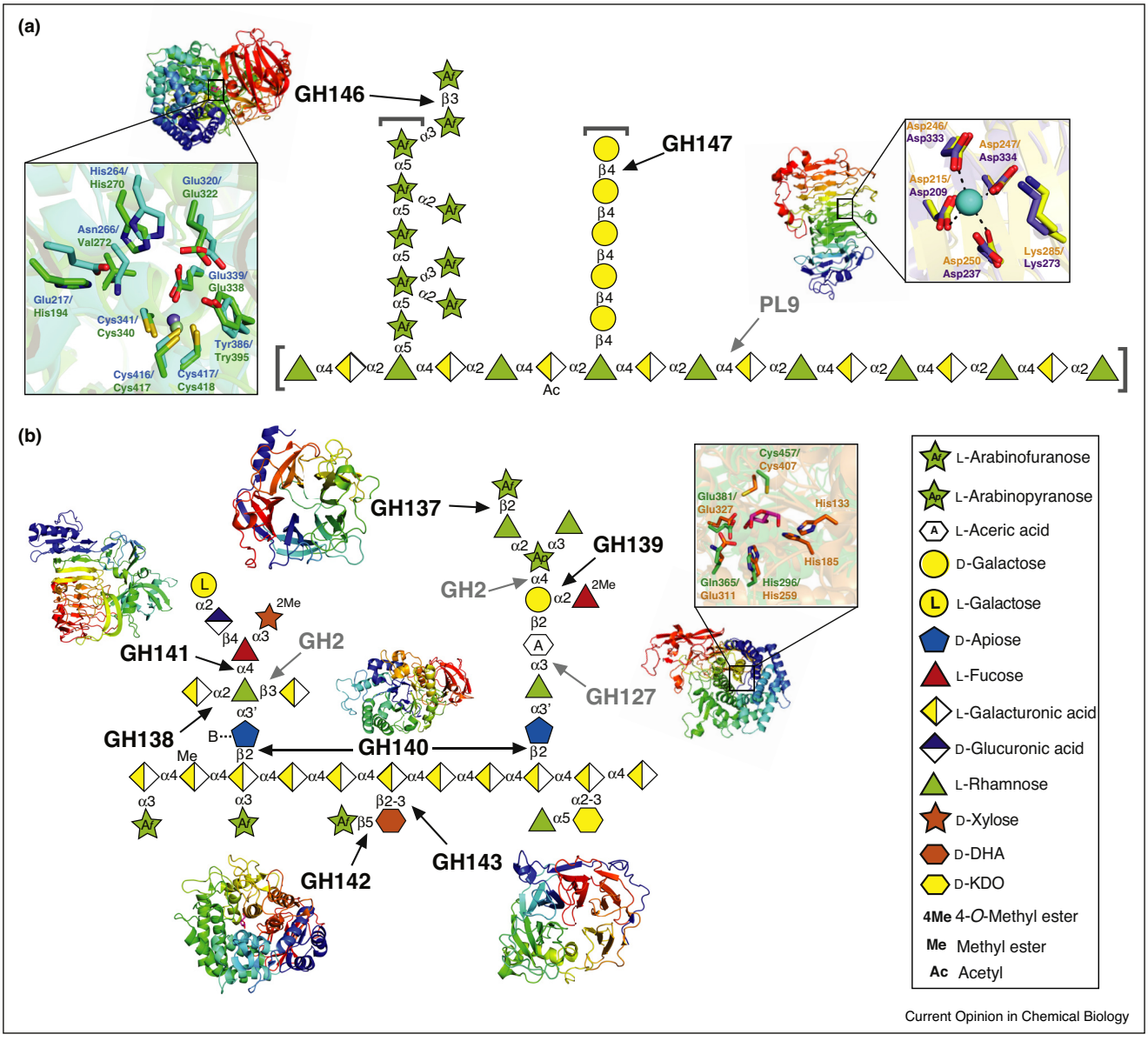
[a] activity never described before.
[b] distance between catalytic residues suggests a double displacement mechanism however this was not proved experimentally.
[c] mechanism inferred from active site similarities with GH33 sialidases.
[d] mechanism inferred from active site similarities with GH127 β-L-arabinofuranosidase.
[e] catalytic residues, mechanism and structure inferred from clan GH-A.

families (Table 1, Figure 3a). A β1,2/1,3-L-arabinofuranosidase is the founding member of GH146. This family is closely related to GH127, forming the clan GH-P that shares the same catalytic apparatus and protein fold (Figure 3a). However, the first GH146 enzyme structure revealed an additional β-sandwich domain positioned over the active site that is not present in GH127 [39••]. A novel GH147 family (β-D-galactosidase) was identified in *B. ovatus* based on PUL comparison to genes originally identified in *B. thetaiotaomicron* and is a member of clan GH-A, allowing prediction of the catalytic residues (two glutamates) and a retaining mechanism [39••,40]. Interestingly, this *exo*-acting enzyme located on the bacterial surface is essential for galactan utilization by *B. ovatus* [39••]. This is an unusual finding, since all of the PULs characterized so far encode an extracellular *endo*-acting enzyme thought to cleave the polysaccharides before import across the outer membrane before complete depolymerization in the periplasm [7,27••,41–44].

RG-II, the most complex glycan currently known is highly conserved among plants [45]. This pectin structure was first described in 1978 and for decades RG-II degrading enzymes remained unknown [46]. In 2011, it was reported for the first time that *B. thetaiotaomicron* is able to grow in the presence of this complex pectin and transcriptional profiling revealed candidate PULs involved; however, only recently, was the enzymatic complexity of the RG-II degradome revealed [4,5••,19••]. Ndeh and colleagues showed that to degrade RG-II *B. thetaiotaomicron* encodes 25 different enzymes, seven of them are founding members of new GH families (Figure 3b, Table 1). Additional characterization of these enzymes revealed the mechanism, catalytic residues and 3D structures for 5 of these new families (Table 1). Despite the different folds, a common feature of all the structures was the location of the active site in a deep pocket that precludes recognition of substituted sugars. This feature is consistent with a sequential exo model of RG-II degradation. An exception is the α1,4-L-fucosidase (GH141) where the active site pocket is elongated to allow the accommodation of the 2-*O*-Me-D-Xyl linked to L-Fuc *O3*. Additionally, the characterization of a novel pectin methyl esterase revealed that the methylation in the RG-II backbone is specific and occurs at the D-GalA adjacent to the D-GalA substituted with chains A and F [5••].

The RG-II degradome also contain 3 enzymes displaying novel activities in previously known GH families: a GH127 that cleaves L-aceric acid (3-*C*-carboxy-5-deoxy-L-xylose-hydrolase) and two GH2 enzymes acting as α1,4-L-arabinopyranosidase and β1,3-D-galacturonidase (Figure 3b) [5••]. The 3D-structure of the α-L-aceric acid hydrolase revealed that the catalytic acid-base is not conserved with other GH127 family members (Figure 3b). The glutamate in β-L-arabinofuranosidases is replaced by a glutamine in the aceric acid hydrolase, precluding a role in the protonation of the glycosidic oxygen during glycosylation, or activating the catalytic water during deglycosylation. In the absence of a candidate to act as catalytic base, it was proposed that the carboxylate of the aceric acid fulfills this role [5••]. Additionally, the activities α-2-*O*-methyl-L-fucosidase (GH139), β-D-DHA-hydrolase (GH143), and α-L-aceric acid-hydrolase (GH127) are novel activities that have not been previously reported in the literature (Table 1) [5••].

Figure 3



Novel enzymes in pectin degradation. Rhamnogalacturonan-I contains a backbone of a repeating disaccharide L-Rha-D-GalA that can be decorated with arabinan (α1,5-L-arabinofuranose backbone) and galactan (β1,4-D-galactose backbone) side chains. In RG-II the D-galacturonic acid backbone is substituted with six side chains [4]. The arrows point the linkage cleaved by novel enymes (black) or enzymes shown to possess new activities inside previously characterized families (grey). The schematic representations of new enzymes are ramp-coloured from blue (N-terminus) to red (C-terminus) and the appended squares show enlargements of the active sites. **(a)** Overlay of GH146 BT0349 (blue, PDB: 5OPJ) from *B. thetaiotaomicron* with GH127 β-L-arabinofuranosidase (green, PDB: 3WKX) from *Bifidobacterium longum* reveals a high conservation of the active site between these two GH families. The fully conserved zinc ion is represented as a sphere (blue and green in GH146 and GH127, respectively). The overlay of PL9 rhamnogalacturonan lyase BT4170 (yellow, PDB: 5OLS) with the pectate lyase Pel9A (purple, PDB: 1RU4) reveals a full conservation of the catalytic residue and the four aspartate residues interacting with calcium-1 (blue sphere). A second calcium-binding site (not shown) is only present in the rhamonogalacturonan lyase BT4170. **(b)** Overlay of aceric acidase BT1003 (green, PDB: 5MQO) and β-L-arabinofuranosidase HypBA1 (orange, PDB: 3WKX) showing the lack of conversation of the catalytic acid-base inside GH127 family. Sugars are shown using the Consortium for Functional Glycomics notation. L-Aceric acid, 3-*C*-carboxy-5-deoxy-L-xylose; D-KDO, 3-deoxy-D-manno-octulosic acid; D-DHA, 3-keto-3-deoxy-D-*lyxo*-heptulosaric acid; GH, glycoside hydrolase; PL, polysaccharide lyase.

Prospectus

The competitive nature of the densely populated human gut microbiota has driven the microorganisms that colonize this habitat to forage on as many nutrients as they can adapt to consume given the enzymatic utensils that they have or are constantly evolving. Sequencing technology has made it easier to view the physiological potential of human gut bacteria and just a handful of in depth molecular biology, genetic and enzyme-based investigations have revealed a substantial number of new catalytic mechanisms. Still, the activities of very few gut bacteria, especially members of the abundant Gram-positive phyla Firmicutes and Actinobacteria, or microorganisms found only in traditional agrarian or hunter-gatherer populations who eat a diet high in fiber [47,48], have been explored in detail. Given the pace of discovery in the past ten years, and from studying just a few micoorganisms, it is likely that investigations of human gut microbes will continue to yield new enzymatic activities for degradation, synthesis and modification of carbohydrates.

Acknowledgements

This work was supported by an R21 (AI128120) award from the US National Institutes of Health and an Innovator Award from the Kenneth Rainin Foundation. A. L. was supported by a Marie Sklodowska Curie Postdoctoral Fellowship from the European Commission (MSCA-IF-GF-748336) to E.C. M.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Porter NT, Martens EC: **The critical roles of polysaccharides in gut microbial ecology and physiology**. *Ann Rev Microbiol* 2017, **71**:349-369.

2. Flint HJ, Scott KP, Duncan SH, Louis P, Forano E: **Microbial degradation of complex carbohydrates in the gut**. *Gut Microbes* 2012, **3**:289-306.

3. • McNulty NP, Wu M, Erickson AR, Pan C, Erickson BK, Martens EC *et al.*: **Effects of diet on resource utilization by a model human gut microbiota containing *Bacteroides cellulosilyticus* WH2, a symbiont with an extensive glycobiome**. *PLoS Biol* 2013, **11**e1001637.

This study interrogated the polysaccharide-degrading ability of*Bacteroides cellulosilyticus,* an organism with over 425 different degradative CAZymes, at phenotypic and transcriptmic levels, as well as *in vivo* in gnotobiotic mice.

4. El Kaoutari A, Armougom F, Gordon JI, Raoult D, Henrissat B: **The abundance and variety of carbohydrate-active enzymes in the human gut microbiota**. *Nat Rev Microbiol* 2012, **11**:497-504.

5. •• Ndeh D, Rogowski A, Cartmell A, Luis AS, Basle A, Gray J *et al.*: **Complex pectin metabolism by gut bacteria reveals novel catalytic functions**. *Nature* 2017, **544**:65-70.

This study reported upon the nearly complete saccharification of rhamnogalacturonan II, the most complex polysaccharide on Earth, by*B. thetaiotaomicron*.

6. Martens EC, Kelly AG, Tauzin AS, Brumer H: **The devil lies in the details: how variations in polysaccharide fine-structure impact the physiology and evolution of gut microbes**. *J Mol Biol* 2015, **426**:3851-3865.

7. Rogowski A, Briggs JA, Mortimer JC, Tryfona T, Terrapon N, Lowe EC *et al.*: **Glycan complexity dictates microbial resource allocation in the large intestine**. *Nat Commun* 2017, **6**:7481.

This study investigated the mechanism of action of over a dozen*Bacteroides* enzymes involved in xylan degradation, revealing both new activities in existing CAZyme families and the cellular location of where *endo-* and *exo-* catalysis occurs.

8. Withers SG: **Mechanisms of glycosyl transferases and hydrolases**. *Carbohydr polymers* 2001, **44**:325-337.

9. Lombard V, Bernard T, Rancurel C, Brumer H, Coutinho PM, Henrissat B: **A hierarchical classification of polysaccharide lyases for glycogenomics**. *Biochem J* 2010, **432**:437-444.

10. Garron ML, Cygler M: **Uronic polysaccharide degrading enzymes**. *Curr Opin Struct Biol* 2014, **28**:87-95.

11. Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B: **The carbohydrate-active enzymes database (CAZy) in 2013**. *Nucleic Acids Res* 2013, **42**:D490-5.

12. Henrissat B, Davies G: **Structural and sequence-based classification of glycoside hydrolases**. *Curr Opin Struct Biol* 1997, **7**:637-644.

13. Henrissat B, Bairoch A: **Updating the sequence-based classification of glycosyl hydrolases**. *Biochem J* 1996, **316**:695-696.

14. Kitamura M, Okuyama M, Tanzawa F, Mori H, Kitago Y, Watanabe N *et al.*: **Structural and functional analysis of a glycoside hydrolase family 97 enzyme from *Bacteroides thetaiotaomicron***. *J Biol Chem* 2008, **283**:36328-36337.

15. Davies G, Henrissat B: **Structures and mechanisms of glycosyl hydrolases**. *Structure* 1995, **3**:853-859.

16. •• Sonnenburg JL, Xu J, Leip DD, Chen CH, Westover BP, Weatherford J *et al.*: **Glycan foraging in vivo by an intestine-adapted bacterial symbiont**. *Science* 2005, **307**:1955-1959.

This study was among the first to monitor the behavior, via whole-genome transcriptional profiling, of a common human gut symbiont,*B. thetaiotaomicron*, during existence alone in the distal intestines of gnotobiotic mice fed high and low fiber diets.

17. Bjursell MK, Martens EC, Gordon JI: **Functional genomic and metabolic studies of the adaptations of a prominent adult human gut symbiont, *Bacteroides thetaiotaomicron*, to the suckling period**. *J Biol Chem* 2006, **281**:36269-36279.

18. Martens EC, Chiang HC, Gordon JI: **Mucosal glycan foraging enhances fitness and transmission of a saccharolytic human gut bacterial symbiont**. *Cell Host Microbe* 2008, **4**:447-457.

19. •• Martens EC, Lowe EC, Chiang H, Pudlo NA, Wu M, McNulty NP *et al.*: **Recognition and Degradation of Plant Cell Wall Polysaccharides by Two Human Gut Symbionts**. *Plos Biology* 2011, **9**e1001221.

This study was the first to systematically examine whole-genome transcriptional responses in symbiotic human gut bacteria,*B. thetaiotaomicron and B. ovatus*, in response to purified plant cell wall fiber polysaccharides, revealing thousands of new genes and CAZymes involved.

20. • Sonnenburg ED, Zheng H, Joglekar P, Higginbottom SK, Firbank SJ, Bolam DN *et al.*: **Specificity of polysaccharide use in intestinal bacteroides species determines diet-induced microbiota alterations**. *Cell* 2010, **41**:1241-1252.

This study revealed an intricate layer of substrate specificity by demonstrating that different Bacteroides have optimized the CAZymes contained in fructan utilization loci to optimally utilize fructans with β2,1 or β2,6 linkages.

21. Pluvinage B, Grondin JM, Amundsen C, Klassen L, Moote PE, Xiao Y *et al.*: **Molecular basis of an agarose metabolic pathway acquired by a human intestinal symbiont**. *Nat Commun* 2018, **9**:1043.

22. •• Hehemann JH, Correc G, Barbeyron T, Helbert W, Czjzek M, Michel G: **Transfer of carbohydrate-active enzymes from marine bacteria to Japanese gut microbiota**. *Nature* 2010, **464**:908-912.

A landmark study that revealed that some human gut bacteria have acquired genes, including some encoding CAZymes, for degradation of seaweed polysaccharides from red algae that are eaten by some cultures, for example as sushi.

23. Cockburn DW, Koropatkin NM: **Polysaccharide degradation by the intestinal microbiota and its influence on human health and disease**. *J Mol Biol* 2016, **428**:3230-3252.

24. • Desai MS, Seekatz AM, Koropatkin NM, Kamada N, Hickey CA, Wolter M *et al.*: **A dietary fiber-deprived gut microbiota degrades the colonic mucus barrier and enhances pathogen susceptibility**. *Cell* 2016, **167**:1339-1353 e21.
This study examined the diet-specific behavior of a synthetic microbiota containing 14 species with over 1,600 CAZymes in gnotobiotic mice, revealing that fiber starvation leads to erosion of the mucus layer and increased pathogen susceptibility.

25. Morais S, Ben David Y, Bensoussan L, Duncan SH, Koropatkin NM, Martens EC *et al.*: **Enzymatic profiling of cellulosomal enzymes from the human gut bacterium, *Ruminococcus champanellensis*, reveals a fine-tuned system for cohesin-dockerin recognition**. *Environ Microbiol* 2016, **18**:542-556.

26. Ze X, Ben David Y, Laverde-Gomez JA, Dassa B, Sheridan PO, Duncan SH *et al.*: **Unique organization of extracellular amylases into amylosomes in the resistant starch-utilizing human colonic firmicutes bacterium *Ruminococcus bromii***. *mBio* 2015, **6**:e01058-15.

27. •• Larsbrink J, Rogers TE, Hemsworth GR, McKee LS, Tauzin AS, Spadiut O *et al.*: **A discrete genetic locus confers xyloglucan metabolism in select human gut Bacteroidetes**. *Nature* 2014, **506**:498-502.
This study was among the first to report the complete analysis of degradative enzymes in different families that are co-expressed by*B. ovatus* to degrade the plant cell wall polysaccharide xyloglucan.

28. Hsieh YSY, Harris PJ: **Xyloglucans of Monocotyledons have diverse structures**. *Mol Plant* 2009, **2**:943-965.

29. Rumpagaporn P, Reuhs BL, Kaur A, Patterson JA, Keshavarzian A, Hamaker BR: **Structural features of soluble cereal arabinoxylan fibers associated with a slow rate of in vitro fermentation by human fecal microbiota**. *Carbohydr Polym* 2015, **130**:191-197.

30. Saulnier L, Vigouroux J, Thibault JF: **Isolation and partial characterization of feruloylated oligosaccharides from maize bran**. *Carbohydr Res* 1995, **272**:241-253.

31. Higgins MA, Whitworth GE, El Warry N, Randriantsoa M, Samain E, Burke RD *et al.*: **Differential recognition and hydrolysis of host carbohydrate antigens by *Streptococcus pneumoniae* family 98 glycoside hydrolases**. *J Biol Chem* 2009, **284**:26161-26173.

32. Nagae M, Tsuchiya A, Katayama T, Yamamoto K, Wakatsuki S, Kato R: **Structural basis of the catalytic reaction mechanism of novel 1,2-alpha-L-fucosidase from *Bifidobacterium bifidum***. *J Biol Chem* 2007, **282**:18497-18509.

33. Zhu Y, Suits MD, Thompson AJ, Chavan S, Dinev Z, Dumon C *et al.*: **Mechanistic insights into a Ca2+-dependent family of alpha-mannosidases in a human gut symbiont**. *Nat Chem Biol* 2010, **6**:125-132.

34. Fujimoto Z, Jackson A, Michikawa M, Maehara T, Momma M, Henrissat B *et al.*: **The structure of a *Streptomyces avermitilis* alpha-L-rhamnosidase reveals a novel carbohydrate-binding module CBM67 within the six-domain arrangement**. *J Biol Chem* 2013, **288**:12376-12385.

35. Miyata T, Kashige N, Satho T, Yamaguchi T, Aso Y, Miake F: **Cloning, sequence analysis, and expression of the gene encoding *Sphingomonas paucimobilis* FP2001 alpha-L -rhamnosidase**. *Curr Microbiol* 2005, **51**:105-109.

36. Eriksson U, Svenson SB, Lönngren J, Lindberg AA: ***Salmonella* phage glycanases: substrate specificity of the phage P22 endo-rhamnosidase**. *J Gen Virol* 1979, **43**:503-511.

37. Munoz-Munoz J, Cartmell A, Terrapon N, Basle A, Henrissat B, Gilbert HJ: **An evolutionarily distinct family of polysaccharide lyases removes rhamnose capping of complex arabinogalactan proteins**. *J Biol Chem* 2017, **292**:13271-13283.

38. Munoz-Munoz J, Cartmell A, Terrapon N, Henrissat B, Gilbert HJ: **Unusual active site location and catalytic apparatus in a glycoside hydrolase family**. *Proc Nat Acad Sci U S A* 2017, **114**:4936-4941.

39. •• Luis AS, Briggs J, Zhang X, Farnell B, Ndeh D, Labourel A *et al.*: **Dietary pectic glycans are degraded by coordinated enzyme pathways in human colonic Bacteroides**. *Nature Microbiology* 2017, **3**:210-219.
This study dissected the very complex set of enzymes in*B. thetaiotaomicron* for degradation of all other pectic polysaccahrides beyond rhamnogalacturoan II, revealing several new activities.

40. Henrissat B, Callebaut I, Fabrega S, Lehn P, Mornon JP, Davies G: **Conserved catalytic machinery and the prediction of a common fold for several families of glycosyl hydrolases**. *Proc Nat Acad Sci U S A* 1995, **92**:7090-7094.

41. Cuskin F, Lowe EC, Temple MJ, Zhu Y, Cameron EA, Pudlo NA *et al.*: **Human gut Bacteroidetes can utilize yeast mannan through a selfish mechanism**. *Nature* 2015, **517**:165-169.

42. Tamura K, Hemsworth GR, Dejean G, Rogers TE, Pudlo NA, Urs K *et al.*: **Molecular mechanism by which prominent human gut Bacteroidetes utilize mixed-linkage β-glucans, major health-promoting cereal polysaccharides**. *Cell Reports* 2017, **21**:417-430.

43. Temple MJ, Cuskin F, Baslé A, Hickey N, Speciale G, Williams SJ *et al.*: **A Bacteroidetes locus dedicated to fungal 1,6-β-glucan degradation: Unique substrate conformation drives specificity of the key endo-1,6-β-glucanase**. *J Biol Chem* 2017, **292**:10639-10650.

44. Cartmell A, Lowe EC, Basle A, Firbank SJ, Ndeh DA, Murray H *et al.*: **How members of the human gut microbiota overcome the sulfation problem posed by glycosaminoglycans**. *Proc Nat Acad Sci U S A* 2017, **114**:7037-7042.

45. O'Neill MA, Ishii T, Albersheim P, Darvill AG: **Rhamnogalacturonan II: structure and function of a borate cross-linked cell wall pectic polysaccharide**. *Ann Rev Plant Biol* 2004, **55**:109-139.

46. Darvill AG, McNeil M, Albersheim P: **Structure of plant cell walls: VIII. A new pectic polysaccharide**. *Plant Physiol* 1978, **62**:418-422.

47. Sonnenburg ED, Smits SA, Tikhonov M, Higginbottom SK, Wingreen NS, Sonnenburg JL: **Diet-induced extinctions in the gut microbiota compound over generations**. *Nature* 2016, **529**:212-215.

48. Smits SA, Leach J, Sonnenburg ED, Gonzalez CG, Lichtman JS, Reid G *et al.*: **Seasonal cycling in the gut microbiome of the Hadza hunter-gatherers of Tanzania**. *Science* 2017, **357**:802-806.

49. Eckburg PB, Bik EM, Bernstein CN, Purdom E, Dethlefsen L, Sargent M *et al.*: **Diversity of the human intestinal microbial flora**. *Science* 2005, **308**:1635-1638.