

## Informed Carbohydrate Active Enzyme Discovery within the Human Distal Gut Microbiome

Abbott DW\*

Lethbridge Research Centre, Agriculture and Agri-Food Canada, 5403-1st Ave Lethbridge, AB, Canada, T1J4B1

### Abstract

Next-generation sequencing of bacterial communities such as the human distal gut microbiome has generated a vast metagenomic sequence space. A common feature emerging from these analyses is that the genomes of intestinal bacteria are enriched with genes dedicated to the metabolism of indigestible dietary polysaccharides. Although culture-independent techniques are unparalleled in their ability to probe and catalog carbohydrate active enzyme gene diversity within complex communities, downstream functional genomic approaches are still required to elucidate the mechanisms by which these enzymes and pathways function. Harnessing the potential of biocatalytic repositories, such as the human distal gut microbiome, will continue to facilitate the informed discovery of potent enzymes and help drive innovation towards the sustainable conversion of plant cell wall biomass.

**Keywords:** Carbohydrate active enzyme; Polysaccharide Utilization Locus; Bacteroides; Microbiome; Distal Gut Microbiota; Metagenomic sequencing; Transcriptomics

### Introduction

#### Intestinal microbiomes as enzyme repositories

The human genome contains nearly one hundred glycoside hydrolases (www.cazy.org, [1]). Of these carbohydrate active enzymes (CAZymes), only a few are dedicated to the metabolism of dietary complex carbohydrates [2]. Human digestive enzymes are primarily responsible for the saccharification of  $\alpha$ 1,4-linked and  $\alpha$ 1,6-linked glucans (GH13s) and sucrose ( $\alpha$ -D-glucopyranosyl-(1,2)- $\beta$ -D-fructofuranoside, GH31s). By comparison, the majority of human CAZymes are involved in glycoprotein maturation, lysosomal processing, and cellular physiology [2]. In order to metabolise dietary fibre humans rely on a symbiotic relationship with a community of microorganisms that colonize their intestine, referred to as the human distal gut microbiota (hDGM). Since the emergence of culture independent, next-generation sequencing methods there have been intensive efforts to catalog the composition of this community and determine how it responds to dynamic stimuli, such as diet and disease. Accordingly, the hDGM is one of the primary sites for the Human Microbiome Project, a global collaboration seeking to exhaustively document the genomes of microorganisms that colonize human beings [3,4]. Current estimates suggest that approximately 1,000 species-level phylotypes are present within the human intestine [5], which encodes an integrated microbiome catalog of nearly of ten million genes [6]. Although the composition of the hDGM varies between healthy individuals, it has been suggested that humans share a 'core microbiome' that is dominated by Bacteroidetes and Firmicutes, with lesser contributions from Actinobacteria, Proteobacteria, Verrucomicrobia, methanogens, yeast, and viruses [5].

Analysis of the genomes of *Bacteroides* spp. has revealed that their genomes are augmented with genes involved in carbohydrate metabolism. For example six percent of the overall genome of *Bacteroides thetaiotaomicron*, a well-studied intestinal model microorganism, encodes for CAZymes [7]. A prominent proportion of these are classified into CAZyme families that have been previously determined to modify pectins,  $\alpha$ -mannans, arabinans, and various  $\beta$ -linked glycans (Figure 1). Monosaccharides released from dietary fibre by these enzymes are fermented into host-absorbable products (e.g. propionate, butyrate) that provide primary sources of energy for the

colonic epithelia [8]. In this regard, the symbiotic relationship between the host and intestinal saccharophiles, such as *B. thetaiotaomicron*, fulfills a functional deficit present within the human metabolome.

#### Comparing CAZyme profiles highlights functional diversity

In recent years the metagenomic sequence space of the hDGM is expanding at a rate that currently exceeds our ability to comprehensively interpret its function. For example, metagenomic surveys often report on the presence or absence of CAZyme families within a microbiome. This provides a snapshot of the metabolic potential contained within an intestinal community and may reveal niche-specific colonization behaviors of species that possess unique CAZyme families (Figure 1) [9-12]. Despite these insights; however, its scope is limited to genes that can be classified into previously defined families and it does not account for functional diversity that exists within CAZyme families (e.g. GH2, GH5, GH43; [9]). For example, the genome of *B. thetaiotaomicron* contains thirty-three GH43s (arabinofuranosidases and xylanases), twenty-three GH92s ( $\alpha$ -mannosidases), and ten GH76s ( $\alpha$ -mannanses; [10]). Comparing these collections to other *Bacteroides* spp. that either colonize a common host (e.g. *Bacteroides vulgatus* [13]: twenty-two GH43s, nine GH92s, seven GH76s) or different animal hosts (e.g. *Bacteroides salanitronis* [14]: thirteen GH43s, six GH92s, zero GH76s; *Prevotella ruminicola* [15]: twenty GH43s, eight GH92s, one GH76) reveals the presence of species-selective metabolic programs (Figure 1).

#### Complex carbohydrate utilization pathways within *Bacteroides* spp.

*Bacteroides* spp. cluster CAZymes into catabolic pathways called 'Polysaccharide Utilization Loci' (PULs; [16]). When a target polysaccharide is detected, a dedicated PUL responds by expressing gene products tailored to modify and transport the carbohydrates subunits within that polysaccharide. When presented with glycan

\*Corresponding author: Abbott DW, Lethbridge Research Centre, Agriculture and Agri-Food Canada, 5403-1st Ave Lethbridge, AB, Canada, T1J4B1, Tel: 1-403-317-3443; E-mail: [wade.abbott@agr.gc.ca](mailto:wade.abbott@agr.gc.ca)

Received July 14, 2014; Accepted August 13, 2014; Published August 16, 2014

Citation: Abbott DW (2014) Informed Carbohydrate Active Enzyme Discovery within the Human Distal Gut Microbiome. J Glycomics Lipidomics 4: 119. doi:10.4172/2153-0637.1000119

Copyright: © 2014 Abbott DW. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

GH	Pr	Bx	Bv	Bt	Bs
2	10	23	25	32	10
3	11	13	5	10	9
5	4	6	0	0	3
8	1	0	0	0	0
9	1	0	0	0	0
10	4	2	1	0	1
13	5	4	4	8	6
15	0	0	1	0	0
16	2	1	1	4	1
18	1	8	2	12	1
20	2	10	9	14	0
23	2	3	3	3	3
24	1	2	2	0	5
25	3	1	1	1	1
26	1	2	0	0	2
27	1	1	1	5	1
28	5	9	13	9	11
29	3	4	8	9	1
30	0	2	5	3	1
31	6	5	4	6	3
32	2	2	1	4	1
33	0	4	3	2	2
35	2	2	1	3	2
36	1	6	3	4	2
38	1	1	0	2	0
42	0	1	1	1	1
43	20	27	22	33	13
50	0	3	0	0	0
51	7	4	3	4	3
53	2	0	0	1	0
57	1	1	1	1	1
59	0	0	0	0	1
63	0	0	2	0	1
65	0	1	0	0	0
66	0	1	0	1	1
67	1	1	1	1	0
73	1	1	1	1	3
76	1	7	0	10	0
77	2	1	1	1	2
78	1	7	5	6	1
84	0	0	1	1	0
88	0	5	3	4	1
89	1	4	1	3	1
91	0	2	0	0	0
92	8	17	9	23	6
93	0	0	0	1	0
95	2	4	4	5	3
97	6	6	7	10	5
99	0	1	0	1	0
105	2	5	8	7	6
106	1	3	3	3	3
109	0	1	5	2	1
110	0	0	2	2	0
115	1	1	0	2	1
117	0	1	0	0	0
123	0	1	1	1	0
125	2	3	0	3	0
127	2	1	3	3	1
130	1	5	2	4	3
133	0	0	1	0	0

**Figure 1:** Comparison of the glycoside hydrolases present within the genomes of five different members of Bacteroidetes. Glycoside hydrolase (GH) families (left column) from *P. ruminicola* 23 (Pr), *B. xylanisolvens* XB1A (Bx), *B. vulgatus* ATCC 8482 (Bv), *B. thetaiotaomicron* VPI-5482 (Bt), and *B. salanitronis* DSM 18170 (Bs). Copy number is displayed as a heat map with green = zero to red = ≥15. Copy numbers that are present in ratios of ≥1.3 over other related species are bolded and underlined. Entries that represent unique GH families within an individual species are boxed (e.g. GH9 = b1,4-glucanase). Adapted from: HYPERLINK "http://www.cazy.org" www.cazy.org [9].

mixtures, which more accurately represents the intestinal digesta *in vivo*, *B. thetaiotaomicron* selectively metabolises distinct polysaccharides

by switching on PULs with varying levels of induction [17]. The size of individual PULs reflects the complexity of a pathway required to metabolise the glycosidic backbone, side-chain linkages, and chemical modifications (e.g. methylesterifications, acetylations) that exist within the substrate. Correspondingly, PULs that target relatively simple homopolysaccharides, such as fructans (Bt1757-BT1763) [18], are noticeably smaller than PULs that target large complex and highly branched heteropolysaccharides, such as rhamnogalacturonan-I (Bt4145-Bt4182) and rhamnogalacturonan-II (Bt0977-Bt1031) [11]. In this regard, transcriptomics of monocultures grown on purified carbohydrates have been instrumental, firstly for determining the presence and selectivity of PULs; and secondly, for defining the boundaries of PULs within the genome.

### Limitations of sequence-based approaches

Although comparative genomics of CAZymes (i.e. 'CAZomics') has utility for highlighting the metabolic potential of a microorganism, it does not establish how genes are regulated, how enzymes operate, and what products are released. Recently, CAZyme families have now begun to be further classified into sequence-based subfamilies in an effort to provide more predictive power [19]. Ultimately however, determining the biological significance of CAZyme profiles requires combinatorial functional genomic approaches, such as transcriptomics, enzymology, structural biochemistry, and metabolomics to define how PULs sense and saccharify polysaccharides [18,20].

### Harnessing *bacteroides* spp. for informed CAZyme discovery

With mounting global population pressures for food and renewable resources, harvesting the chemical energy within structural carbohydrates from the cell walls of plants [21] and macroalgae [22] remain some of the most promising avenues for invested research. In this regard, unlocking the biocatalytic potential of the hDGM, and other microbiomes tuned to the bioconversion of more recalcitrant biomass (e.g. domesticated or wild ruminants), are likely to continue to drive innovation towards discovering CAZymes with an improved catalytic efficiency, defining cocktails that provide synergistic improvements in the turnover of a mixed substrate (i.e. artificial pathways), and isolating intestinal probiotics for biological pretreatments [23-25]. Furthermore, exploratory studies have demonstrated that it is possible to intentionally screen *Bacteroides* spp. to discover novel PULs that are active on select polysaccharides. For example, the rhamnogalacturonan-II pathway recently identified within *B. thetaiotaomicron* documented the first CAZymes that degrade this polysaccharide; a substrate which was previously believed to be impervious to enzyme-mediated deconstruction [26]. Future efforts to culture *Bacteroides* spp. *ex vivo* to evaluate selective growth proficiency combined with transcriptomics and the biochemical characterization of PUL components promise to define novel pathways that have utility for the saccharification of or bioproduct generation from industrial significant carbohydrates.

### References

1. Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, et al. (2009) The Carbohydrate-Active Enzymes Database (Cazy): an expert resource for glycogenomics. *Nucleic Acids Res* 37: D233-238.
2. El Kaoutari A, Armougom F, Gordon JI, Raoult D, Henrissat B (2013) The Abundance and Variety of Carbohydrate-Active Enzymes in the Human Gut Microbiota. *Nat Rev Microbiol* 11: 497-504.
3. Aagaard K, Petrosino J, Keitel W, Watson M, Katancik J, et al. (2013) The Human Microbiome Project Strategy for Comprehensive Sampling of the Human Microbiome and Why It Matters. *FASEB J* 27: 1012-1022.
4. (2012) Structure, Function and Diversity of the Healthy Human Microbiome. *Nature* 486: 207-214.

5. Lozupone CA, Stombaugh JI, Gordon JI, Jansson JK, Knight R (2012) Diversity, Stability and Resilience of the Human Gut Microbiota. *Nature* 489: 220-230.
6. Li J, Jia H, Cai X, Zhong H, Feng Q, et al. (2014) An Integrated Catalog of Reference Genes in the Human Gut Microbiome. *Nat Biotech* advance online publication.
7. Ravcheev DA, Godzik A, Osterman AL, Rodionov DA (2013) Polysaccharides Utilization in Human Gut Bacterium *Bacteroides Thetaiotaomicron*: Comparative Genomics Reconstruction of Metabolic and Regulatory Networks. *BMC Genomics* 14: 873.
8. Xu J, Chiang HC, Bjursell MK, Gordon JI (2004) Message from a Human Gut Symbiont: Sensitivity Is a Prerequisite for Sharing. *Trends Microbiol* 12: 21-28.
9. Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B (2014) The Carbohydrate-Active Enzymes Database (Cazy) in 2013. *Nucleic Acids Res* 42: D490-495.
10. Xu J, Bjursell MK, Himrod J, Deng S, Carmichael LK, et al. (2003) A Genomic View of the Human-*Bacteroides Thetaiotaomicron* Symbiosis. *Science* 299: 2074-2076.
11. Martens EC, Lowe EC, Chiang H, Pudlo NA, Wu M, et al. (2011) Recognition and Degradation of Plant Cell Wall Polysaccharides by Two Human Gut Symbionts. *PLoS biology* 9: e1001221.
12. Martens EC, Kelly AG, Tauzin AS, Brumer H (2014) The Devil Lies in the Details: How Variations in Polysaccharide Fine-Structure Impact the Physiology and Evolution of Gut Microbes. *J Mol Biol*
13. Xu J, Mahowald MA, Ley RE, Lozupone CA, Hamady M, et al. (2007) Evolution of Symbiotic Bacteria in the Distal Human Intestine. *PLoS biology* 5: e156.
14. Gronow S, Held B, Lucas S, Lapidus A, Del Rio TG, et al. (2011) Complete Genome Sequence of *Bacteroides Salanitronis* Type Strain (BI78). *Stand Genomic Sci* 4: 191-199.
15. Purushe J, Fouts DE, Morrison M, White BA, Mackie RI, et al. (2010) Comparative Genome Analysis of *Prevotella Ruminicola* and *Prevotella Bryantii*: Insights into Their Environmental Niche. *Microb Ecol* 60: 721-729.
16. Martens EC, Koropatkin NM, Smith TJ, Gordon JI (2009) Complex Glycan Catabolism by the Human Gut Microbiota: The *Bacteroidetes* Sus-Like Paradigm. *J Biol Chem* 284: 24673-24677.
17. Rogers TE, Pudlo NA, Koropatkin NM, Bell JS, Moya Balasch M, et al. (2013) Dynamic Responses of *Bacteroides Thetaiotaomicron* During Growth on Glycan Mixtures. *Mol Microbiol* 88: 876-890.
18. Sonnenburg ED, Zheng H, Joglekar P, Higginbottom SK, Firkbank SJ, et al. (2010) Specificity of Polysaccharide Use in Intestinal *Bacteroides* Species Determines Diet-Induced Microbiota Alterations. *Cell* 141: 1241-1252.
19. Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B (2014) The Carbohydrate-Active Enzymes Database (Cazy) in 2013. *Nucleic acids research* 42: D490-495.
20. Larsbrink J, Rogers TE, Hemsworth GR, McKee LS, Tauzin AS, et al. (2014) A Discrete Genetic Locus Confers Xyloglucan Metabolism in Select Human Gut *Bacteroidetes*. *Nature* 506: 498-502.
21. Moreno AD, Ibarra D, Alvira P, Tomas-Pejo E, Ballesteros M (2014) A Review of Biological Delignification and Detoxification Methods for Lignocellulosic Bioethanol Production. *Crit Rev Biotechnol*.
22. Wei N, Quarterman J, Jin YS (2013) Marine Macroalgae: An Untapped Resource for Producing Fuels and Chemicals. *Trends Biotechnol* 31: 70-77.
23. Dodd D, Mackie RI, Cann IK (2011) Xylan Degradation, a Metabolic Property Shared by Rumen and Human Colonic *Bacteroidetes*. *Mol Microbiol* 79: 292-304.
24. Patel DD, Patel AK, Parmar NR, Shah TM, Patel JB, et al. (2014) Microbial and Carbohydrate Active Enzyme Profile of Buffalo Rumen Metagenome and Their Alteration in Response to Variation in the Diet. *Gene* 545: 88-94.
25. Wongwilaiwalin S, Laothanachareon T, Mhuantong W, Tangphatsomruang S, Eurwilaichitr L, et al. (2013) Comparative Metagenomic Analysis of Microcosm Structures and Lignocellulolytic Enzyme Systems of Symbiotic Biomass-Degrading Consortia. *Appl Microbiol Biotechnol* 97: 8941-8954.
26. Martens EC, Lowe EC, Chiang H, Pudlo NA, Wu M, et al. (2011) Recognition and Degradation of Plant Cell Wall Polysaccharides by Two Human Gut Symbionts. *PLoS Biol* 9: e1001221.