Characterization of an α-glucosidase enzyme conserved in *Gardnerella* spp. isolated from the human vaginal microbiome

Pashupati Bhandari[1], Jeffrey P. Tingley[2], D. Wade Abbott[2] and Janet E. Hill[1,*]

[1]Department of Veterinary Microbiology, Western College of Veterinary Medicine, University of Saskatchewan, 52 Campus Drive, Saskatoon, Saskatchewan, S7N 5B4, Canada

[2]Lethbridge Research and Development Centre, Agriculture and Agri-Food Canada, Lethbridge, Alberta, T1J 4B1, Canada

*To whom correspondence should be addressed

Janet.Hill@usask.ca

Abstract

Gardnerella spp. in the vaginal microbiome are associated with bacterial vaginosis, a dysbiosis in which lactobacilli dominant microbial community is replaced with mixed aerobic and anaerobic bacteria including *Gardnerella* species. The co-occurrence of multiple *Gardnerella* species in the vaginal environment is common, but different species are dominant in different women. Competition for nutrients, particularly glycogen present in the vaginal environment, could play an important role in determining the microbial community structure. Digestion of glycogen into products that can be taken up by bacteria requires the combined activities of several enzymes collectively known as "amylases". In the present study, glycogen degrading abilities of *Gardnerella* spp. were assessed. We found that *Gardnerella* spp. isolates and filtered culture supernatants had amylase activity. Phylogenetic analyses predicted conserved Glycoside Hydrolase family 13 (GH13) members among *Gardnerella* spp. including a putative α-glucosidase. The gene for this enzyme was cloned and expressed, and recombinant protein was purified and functionally characterized. The enzyme was active on a variety of maltooligosaccharides over a broad pH range (4 - 8) with an optimum activity at pH 6-7. Glucose was released from maltose, maltotriose and maltopentose, however, no products were detected on thin layer chromatography (TLC) when the enzyme was incubated with glycogen. Our findings show that *Gardnerella* spp. produce a secreted α-glucosidase enzyme that can contribute to the complex and multistep process of glycogen breakdown by degrading smaller oligosaccharides into glucose, contributing to the pool of nutrients available to the vaginal microbiota.

Introduction

Gardnerella spp. in the vaginal microbiome are hallmarks of bacterial vaginosis, a condition characterized by replacement of the lactobacilli dominant microbial community with mixed aerobic and anaerobic bacteria including *Gardnerella*. This dysbiosis is associated with increased vaginal pH, malodorous discharge and the presence of biofilm (1). In addition to troubling symptoms, the presence of abnormal vaginal microbiota is associated with increased risk of HIV transmission and infection with other sexually transmitted pathogens such as *Neisseria gonorrhoeae* and *Trichomonas* spp. (2, 3). Historically, *Gardnerella* has been considered as single species genus. Jayaprakash *et al*. used cpn60 barcode sequences to divide *Gardnerella* spp. into four subgroups (A-D) (4), and this framework was supported by whole genome sequence comparison (5, 6). More recently, Vaneechoutte *et al.* emended the classification of *Gardnerella* based on whole genome sequence comparison, biochemical properties and matrix-assisted laser desorption ionization time-of-flight mass spectrometry and proposed the addition of three novel species: *Gardnerella leopoldii*, *Gardnerella piotii* and *Gardnerella swidsinskii* (7).

Colonization with multiple *Gardnerella* species is common, and different species are dominant in different women (8). Understanding factors that contribute to differential abundance is important since the species may differ in virulence (6) and they are variably associated with clinical signs (8, 9). Several factors, including inter-specific competition, biofilm formation and resistance to antimicrobials could contribute to the differential abundance of the different species. Khan *et al.* showed that resource-based scramble competition is frequent among *Gardnerella* subgroups (10). Glycogen is a significant nutrient for vaginal microbiota, but previous reports on the growth of *Gardnerella* spp. on glycogen containing medium are inconsistent (11). Species may differ in their ability to digest glycogen and utilize the breakdown products, which may in turn contribute to determining microbial community structure.

Glycogen is an energy storage molecule, which consists of linear chains of approximately 13 glucose molecules covalently linked with α-1,4 glycosidic linkages, with branches attached through α-1,6 glycosidic bonds (12, 13). A single glycogen molecule consists of approximately 55,000 glucose residues with a molecular mass of ~ 10^7 kDa (14). The size of glycogen particles can vary with source; from 10-44 nm in human skeletal muscle to approximately 110-290 nm in human liver (15). Glycogen is deposited into the vaginal lumen by epithelial cells under the influence of estrogen (16), and the concentration of cell-free glycogen in vaginal fluid varies greatly, ranging from 0.1 - 32 μg/μl (17). Digestion of this long polymeric molecule into smaller components is essential before it can be taken up by bacterial cells (18, 19).

Glycogen digestion is accomplished by the coordinated action of enzymes that are collectively described as "amylases" (20). α-glucosidase is an exo-acting enzyme that acts on α-1,4 glycosidic bond from the non-reducing end to release glucose, while α-amylase is an endo-acting amylase that cleaves α-1,4 glycosidic bonds to produce α-limit dextrins and maltose. Debranching enzymes like pullulanase and isoamylase break α-1,6 glycosidic bonds (15, 21). Glycogen degradation mechanisms active in the vaginal microbiome are not completely described. Although it has been demonstrated that vaginal secretions exhibit amylase activity that can degrade glycogen and the generation of these products enables growth of vaginal lactobacilli (19), the role of bacterial enzymes in this process is unknown. Moreover, there is limited information about glycogen degradation by clinically important bacteria such as *Gardnerella* spp..

The objective of the current study was to assess amylase activity of *Gardnerella* spp. and to characterize the activity of a putative amylase protein conserved in all *Gardnerella* spp..

Methods

Gardnerella isolates

Isolates of *Gardnerella* used in this study were from a previously described culture collection (6) and included representatives of *cpn*60 subgroup A (*G. leopoldii* (n = 4), *G. swidsinskii* (n = 6), other (n = 1)), subgroup B (*G. piotii* (n = 5), Genome sp. 3 (n = 2)), subgroup C (*G. vaginalis* (n = 6)), and subgroup D (Genome sp. 8 (n = 1), Genome sp. 9 (n = 1)). Whole genome sequences of these isolates had been determined previously (BioProject Accessions PRJNA394757) and annotated by the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) (22).

Amylase activity assay

Amylase activity of *Gardnerella* isolates was assessed using a starch iodine test (23). Isolates were revived from -80°C storage on Columbia blood agar plates with 5% sheep blood, and incubated anaerobically using the GasPak system (BD, USA) at 37°C for 48 hours. Isolated colonies from blood agar plates were transferred in 0.85% saline using a sterile swab and cell suspensions were adjusted to McFarland 1.0. The cell suspension (10 µL) was spotted on to BHI agar plates containing 1% (w/v) potato starch (Difco) or bovine liver glycogen (Sigma-Aldrich), and plates were incubated anaerobically at 37°C for 24 hours. Lugol's iodine solution (1% w/v) was added to the plates to detect evidence of amylase activity indicated by a clear halo around the bacterial colony.

To determine the amylase activity of culture supernatants, *Gardnerella* isolates were first revived from -80°C onto Columbia agar with 5% sheep blood. Isolated colonies were inoculated into 2 ml of BHI broth supplemented with 0.25% maltose and incubated anaerobically at 37°C for 48 hours. Broth cultures were centrifuged at 10,000 × *g* for 10 mins and the supernatant was passed through a 0.45 µm filter Filtered supernatant (10 µL) was spotted on to a 1% (w/v) starch and glycogen agar plates and incubated anaerobically at 37°C for 24 hours. Plates were flooded with

1% (w/v) Lugol's iodine solution to detect evidence of amylase activity. α-amylase from *Bacillus licheniformis* (0.05 mg/ml) (Sigma-Aldrich, Oakville, ON, CAT No. A455) was used as a positive control. To confirm sterility, culture supernatants were inoculated into Columbia agar with 5% sheep and incubated anaerobically at 37°C. Plates were observed for any growth after 48 hours of incubation.

Identification of putative secreted amylase sequences

The predicted proteomes of the 26 *Gardnerella* isolates used in this study were examined with Phobius v1.01 (24) to identify cytoplasmic and non-cytoplasmic proteins based on the detection of signal peptides. Within the predicted secreted proteins, putative amylase enzymes were identified by annotation terms using PGAP annotation. Multiple sequence alignments of predicted secreted amylase sequences were performed using CLUSTALw with results viewed and edited in AliView (version 1.18) (25) prior to phylogenetic tree building using PHYLIP (26).

Functional prediction of *Gardnerella* spp. amylases

Putative amylolytic enzymes in *Gardnerella* spp. genomes were identified using dbCAN, and predicted sequences were run through the SACCHARIS pipeline (27). SACCHARIS combines user sequences with sequences from the CAZy database and trims sequences to the catalytic domain using dbCAN2 (28). Domain sequences were aligned with MUSCLE (29), and a best-fit model was generated with ProtTest (30). Final trees were generated with FastTree (31) and visualized with iTOL (32). Alignment of CG400_06090 from *G. leopoldii* NR017 to orthologues from *Halomonas* sp. (BAL49684.1), *Xanthomonas campestris* (BAC87873.1), and *Culex quinquefasciatus* (ASO96882.1) was done using CLUSTALw and visualized in ESPript (33). A predicted structure model of CG400_06090 was aligned with *Halomonas* sp. HaG (BAL49684.1,

PDB accession 3WY1) (34) using PHYRE2 (35). BLASTp (36) alignment of CAZy derived sequences from reference genomes to a database of the predicted proteomes of 26 study isolates was conducted to identify orthologs.

Expression and purification of the CG400_06090 gene product

Genomic DNA from *G. leopoldi* NR017 was extracted using a modified salting out procedure, and the CG400_06090 open reading frame (locus tag CG400_06090 in GenBank accession NNRZ01000007, protein accession RFT33048) was PCR amplified with primers JH0729-F (5'ATG CAT GCG CAT TAT ACG ATC ATG CTC-3') and JH0730-R (5'ATG GTA CCT TAC ATT CCA AAC ACT GCA-3'). Underlined sequences indicate *Sph*I and *Kpn*I restriction enzyme sites. PCR reaction contained 1 × PCR reaction buffer (0.2 M Tris-HCl pH 8.4, 0.5 M KCl), 200 µM each dNTPs, 2.5 mM $MgCl_2$, 400 nM each primer, 1 U/reaction of Platinum Taq DNA polymerase high-fidelity (5 units/µL in 50% glycerol, 20 mM Tris-HCl, 40 mM NaCl, 0.1 mM EDTA, and stabilizers) (Life-technologies) and 2 µl of template DNA, in a final volume of 50 µl. PCR was performed with following parameters: initial denaturation at 94°C for 3 minutes followed by (denaturation at 94°C, 15 seconds; annealing at 60°C, 15 seconds and extension at 72°C, 2 minutes; 35 cycles), and final extension at 72°C for 5 minutes. Purified PCR products were digested with *Kpn*I and *Sph*I and ligated into expression vector pQE-80 L (Qiagen, Mississauga, ON) digested with same restriction endonucleases. The resulting recombinant plasmid was used to transform One Shot TOP10 chemically competent *E. coli* cells (Invitrogen, Carlsbad, California). Colony PCR was performed to identify transformants containing vector with insert. Insertion of the putative amylase gene in-frame with the N-terminal 6×Histidine tag was confirmed by sequencing of the purified plasmid.

E. coli containing the plasmid were grown overnight at 37°C in LB medium with 100 µg/ml ampicillin. Overnight culture was diluted 1:60 in fresh medium to a final volume of 50 ml, and the culture was incubated at 37°C with shaking at 225 rpm until it reached an OD_{600} of 0.6. At this point, expression was induced with 0.1 mM IPTG and a further 4 hours of incubation was done at 37°C at 225 rpm. Cells were harvested by centrifugation at 10,000 × *g* for 30 mins and the pellet was resuspended in lysis buffer (50 mM NaH_2PO_4, 300 mM NaCl, 10 mM imidazole, pH 8.0). Lysozyme was added to 1 mg/ml and cells were lysed by sonication (8 mins total run time with 15 sec on and 30 sec off). The lysate was clarified by centrifugation at 10,000 × *g* for 30 mins at 4°C and the supernatant was applied to an Ni-NTA affinity column according to the manufacturer's instructions (Qiagen, Germany). Bound proteins were washed twice with wash buffer (50 mM NaH_2PO_4, 300 mM NaCl, 20 mM imidazole, pH 8.0) and eluted with elution buffer (50 mM NaH_2PO_4, 300 mM NaCl, 250 mM imidazole, pH 8.0). Eluted protein was buffer exchanged with the same elution buffer but without imidazole using 30K MWCO protein concentrator (ThermoFisher Scientific). Final protein concentration was measured by spectrophotometry.

Measuring enzyme activity

Purified protein was tested for α-glucosidase activity by measuring the release of 4-Nitrophenol from a chromogenic substrate 4-Nitrophenyl-α-D-glucopyranoside as described elsewhere (37). Briefly, 25 µl of enzyme solution (50 µg/ml) was added to 25 µl of 10 mM 4-Nitrophenylα-D-glucopyranoside substrate and the reaction mixture was incubated at 37°C for 10 mins. Reactions were stopped by the addition of 100 µl of 1M Na_2CO_3 solution and OD_{420} was measured. To determine activity at different pH values, substrate was prepared in 50 mM citric acid-sodium citrate buffer (pH 3.0) or 50 mM sodium phosphate buffer (pH 4.0-8.0).

Enzymatic activity on different substrates

The activity of the purified enzyme was examined with different substrates according to Cockburn et *al.* (38). Briefly, 10 µl of enzyme solution (5.0 mg/ml) was added to 190 µl of 0.1% (w/v) maltose, maltotriose, maltopentose, maltodextrins (MD 4-7, MD 13-17 and MD 16.5-19.5) or glycogen in 50 mM sodium phosphate buffer, pH 6.0, and reaction mixture was incubated at 37°C. Aliquots of the reaction mixture (10 µl) were taken out at various time intervals of incubation and were frozen immediately until thin layer chromatography (TLC) was performed. Standards and the reaction mixtures (0.5 µl) were loaded on 0.25 mm silica gel plates (Sigma-Aldrich, Oakville, ON) and the chromatogram was developed in a solvent system containing acetonitrile, ethyl acetate, 2-propanol and water in (8.5:2:5:6 vol/vol) in a developing chamber until the mobile phase reached the top of the plates. The plates were allowed to dry in a fume hood and products were resolved by dipping the plates into the 0.3% solution of 1-napthyl ethylene diamine dihydrochloride followed by drying and heating with the heat gun until spots became visible.

Results

Amylase activity of *Gardnerella* spp.

All 26 isolates tested showed amylase activity by degrading starch and glycogen as indicated by the appearance of a clear zone around bacterial colonies after iodine staining. Filter sterilized culture supernatants from *Gardnerella* isolates were also able to digest the starch and glycogen in the agar plate assay, indicating that amylase enzymes are secreted (Figure 1). Filter sterilized culture supernatants were found to be sterile after 48 hours of incubation.

Identification of sequences of secreted amylases

A total of 2,525 predicted proteins from the combined proteomes of the 26 study isolates were determined to contain a signal peptide and predicted to be non-cytoplasmic. Of these putative secreted proteins, 60 were annotated as amylases. After removal of severely truncated sequences, a multiple sequence alignment of the remaining 42 sequences was trimmed to the uniform length of 290 amino acids. Phylogenetic analysis of these predicted extracellular amylase sequences revealed five clusters, but only one of these clusters contained orthologous sequences from all 26 isolates (Figure 2). Sequence identities within this cluster were all 91-100%. This conserved sequence (corresponding to GenBank accession WP_004104790) was annotated as an GH13 α-amylase (EC 3.2.1.1) in all genomes, suggesting it was an endo-acting enzyme that releases maltooligosaccharides from glycogen.

The GH13 family in the CAZy database is a large, polyspecific family targeting α-glucose linkages. The functional range of members in this family makes annotation of new sequences challenging. In order to predict the activity of the conserved protein identified among the study isolates, we employed SACCHARIS, which combines CAZyme family trees generated from biochemically characterized proteins with related sequences of unknown function. Currently, two *Gardnerella* reference genomes are available in CAZy: *G. vaginalis* ATCC 14018 and *G. swidsinkii* GV37. GH13 family proteins encoded in these *Gardnerella* genomes were identified and a family tree including all characterized GH13 sequences in the CAZy database was generated (Figure S1). SACCHARIS annotation of the GH13 family in *Gardnerella* resulted in identification of protein domains belonging to ten subfamilies. Alignment with characterized GH13 enzymes from the CAZy database (39), suggested that the conserved amylase identified in the study isolates was an α-glucosidase (EC 3.2.1.20), closely related to subfamilies 23 and 30 of GH13, and not an α-amylase as annotated in GenBank.

Relationship to other α-glucosidases

A representative sequence of the predicted α-glucosidase was selected from *G. leopoldii* (NR017). This sequence (CG400_06090) was combined with functionally characterized members of the GH13 CAZy database using SACCHARIS (Figure 3A). CG400_06090 partitioned with members of GH13 subfamilies 17, 23 and 30, of which the prominent activity is α-glucosidase. This clade was expanded to better illustrate the relationship of CG400_06090 with related characterized sequences as an apparently deep-branching member of subfamily 23, although with low bootstrap support (Figure 3B). When sequences of additional (not functionally characterized) members were included, CG400_06090 clustered within subfamily 23 (Figure S2).

To identify conserved catalytic residues between CG400_06090 and characterized members of GH13, CG400_06090 was aligned with structurally characterized members of subfamily 23 (BAL49684.1 and BAC87873.1) and subfamily 17 (ASO96882.1) (Figure 3C). Of the GH13 catalytic triad, the aspartate nucleophile and the aspartate that stabilizes the transition state are conserved among the four sequences (40). The glutamate general acid/base, however, does not appear to be sequence conserved between CG400_06090 and the GH13 subfamily 23 members. Upon modelling CG400_06090 in PHYRE2 and aligning with BAL49684.1 (PDBID 3WY1), CG400_06090 Glu256 did appear to be spatially conserved with the general acid/base (not shown).

Expression and purification of α-glucosidase protein

The *G. leopoldi* NR017 CG400_06090 open reading frame comprises 1701 base pairs encoding a 567 amino acid protein, with a predicted mass of 62 kDa. The full-length open reading frame encoding amino acids 2-567 was PCR amplified and ligated into vector pQE-80L for expression in *E. coli* as an N-terminal 6xHis-tagged protein. A distinct protein band in SDS-PAGE between

50 kDa and 75 kDa was obtained from IPTG induced *E. coli* cells compared with non-induced cells, indicating the expression of the α-glucosidase protein (Figure 4A). The recombinant protein was soluble, and was purified using a Ni-NTA spin column (Figure 4B). From a 50 ml broth culture, approximately 700 µl of purified protein at 4.0 mg/ml was obtained.

Effect of pH on enzyme activity

α-glucosidase activity of the purified protein was demonstrated by release of 4-Nitrophenol from chromogenic substrate 4-Nitrophenyl-α-D-glucopyranoside. This activity was observed across a broad pH range (4.0 – 8.0), with maximum activity at pH 6.0 and 7.0 while no activity was seen at pH 3.0 (Figure 5).

Analysis of substrate hydrolysis

Production of glucose was detected when the purified α-glucosidase was incubated with maltose, maltotriose and maltopentose at pH 6.0 (Figure 6). No appreciable activity, however, was detected on maltodextrins (MD 4-7, MD 13-17 and MD 16.5-19.5) or glycogen (Figure S3).

Discussion

Gardnerella spp. is associated with bacterial vaginosis in reproductive aged women. They are predominantly present in vaginal biofilms along with other strict anaerobes such as *Atopobium vaginae*, *Prevotella bivia*, *Mobiluncus mulieris*, *Megasphera* spp. and others (41). Although, the role of *Gardnerella* spp. in the vaginal microbiome is not completely understood, it is reported that they can initiate the colonization of vaginal epithelial cells and provide a suitable environment for other anaerobes to grow due to their strong adherence to the epithelial cells and ability to form biofilm. *Gardnerella* spp. in biofilms can also establish a symbiotic relationship with other

anaerobes associated with bacterial vaginosis. Amino acids produced by the *Gardnerella* spp. can promote the growth of *P. bivia and M. mulieris* which in turn produces ammonia taken up by the *Gardnerella* spp. (42). It has been suggested that *Gardnerella* spp. contribute to initiating the transition of the vaginal microbiome from eubiosis towards dysbiosis (43, 44).

Glycogen is a major available nutrient to vaginal microbiota and its utilization likely plays an important role in the survival and success of *Gardnerella* spp. in the vaginal microbiome. Glycogen is digested into smaller products, such as maltose and maltodextrins, by a group of enzymes collectively known as "amylases". In this study, we assessed the amylase activity of *Gardnerella* spp. and an α-glucosidase gene conserved in *Gardnerella* spp. was expressed and its product was characterized.

Isolates of *G. leopoldii*, *G. swidsinskii*, *G. piotii* and *G. vaginalis* all showed amylase activity by degrading starch and glycogen in an agar plate assay. The ability of *Gardnerella* spp. to utilize glycogen has been reported in previous studies. Dunkelburg *et al.* described the growth of four strains of what was then known as *Haemophilus vaginalis* in buffered peptone water with 1% glycogen suggesting their ability to ferment glycogen, however, no additional details were provided (45). Similarly, in another study, Edmunds reported the glycogen fermenting ability of 14 out of 15 *Haemophilus vaginalis* isolates, although serum was included in the media (46). Piot *et al.* examined the starch fermentation ability of 175 *Gardnerella* isolates on Mueller Hinton agar supplemented with horse serum and found that all *Gardnerella* strains were capable of hydrolysing starch (47). They did not, however, assess glycogen degradation and further, it is not clear if serum amylase contained in the media contributed to the observed amylase activity. Robinson, in his review, reported that starch can be utilized by *Gardnerella* (48) while Catlin in another review, reported that glycogen utilization is inconsistent in *Gardnerella* (11). It is important to note that all isolates used in those studies were not actually *Haemophilus vaginalis* due to its resemblance

to other coryneform bacteria in Gram staining and colony morphology, which were the primary bacterial identification methods in the 1960's (11). To our knowledge, this is the first study that uses genetically characterized strains and modern anaerobic growth conditions to test the glycogen degradation by *Gardnerella* spp.. All *Gardnerella* spp. assessed in our study were amylase positive, which suggests it is a conserved catalytic function. As these bacteria are found almost exclusively in the glycogen-rich human vagina, it is not surprising that they produce amylases to break down this substrate.

In the phylogenetic analysis of putative secreted amylase sequences, one cluster contained sequences common to all 26 isolates. Although this protein sequence was annotated as an α-amylase, suggesting it would produce maltooligosaccahrides from glycogen, SACCHARIS predicted it to be closely related to α-glucosidases, which cleave glucose from the non-reducing end of its substrate. SACCHARIS predicts enzyme activity by comparing sequence domains to sequences deposited in the curated CAZy database for which structural and / or biochemical information is available (27). Automated annotation of sequences deposited in primary sequence databases, such as GenBank has resulted in significant problems with functional mis-annotation (49). This problem is exacerbated by the increasing rate at which whole genome sequences are being generated and deposited. The genome sequences of the study isolates were annotated using the NCBI Prokaryotic Genome Annotation Pipeline, which identifies genes and annotates largely based on sequence similarity to previously annotated sequences, thus potentially propagating incorrect functional annotations.

In the phylogenetic analysis of GH13 family enzymes, the conserved α-glucosidase from *Gardnerella* spp. appears to be most closely related to GH13 subfamily 17, 23 and 30 characterized members, forming a monophyletic group. However, low bootstrap values between CG400_06090 and other clade members make assignment of CG400_06090 to a defined subfamily difficult based

on comparison to sequences of functionally characterized proteins (Figure 3B). Additional support for assignment to subfamily 23 was observed when additional sequences of proteins lacking functional characterization from subfamilies 17, 23, and 30 were included (Figure S2). Regardless, the α-glucosidase activity observed within this clade does extend to CG400_06090. Primary sequences of GH13 members of amylases have seven highly conserved regions and three amino acids that form the catalytic triad (21). Comparison of the primary sequence and three-dimensional model of CG400_06090 with characterized α-glucosidases from GH13 subfamily 17 and 23 did show sequence conservation of nucleophile and spatial conservation of the catalytic triad.

Our results demonstrate the value of using SACCHARIS to guide discovery of CAZymes based on comparison at a catalytic domain level to functionally characterized enzymes. All of the other proteins identified as putative secreted amylases were also annotated as "alpha-amylase" in the GenBank records. Further investigation will be required to demonstrate the actual functions of these proteins, and how they contribute to carbohydrate degrading activities of *Gardnerella* spp..

Amylase enzymes that contribute to glycogen digestion in the vaginal environment are not well studied. Spear *et al*. (19) suggested a role for a human amylase enzyme in the breakdown of vaginal glycogen. They detected host-derived pancreatic α-amylase and acidic α-glucosidase in genital fluid samples collected from women and showed that genital fluid containing these enzymes can degrade the glycogen. The presence of these enzyme activities alone in vaginal fluid, however, was not correlated with glycogen degradation, suggesting contributions of activity from the resident microbiota. Many vaginal bacteria likely produce amylases that contribute to this activity as various anaerobic taxa are known to encode amylases and digest glycogen (50), however, the contributions of the vaginal microbiota to this process and the importance of these processes to vaginal microbial community dynamics are not yet well understood.

α-glucosidase enzymes are typically active over a broad pH range and exhibit optimal activity at pH 4.5-6.0 (51). They have also been characterized in other bacteria isolated from human body sites, particularly the intestinal tract. An extracellular α-glucosidase form *Lactobacillus acidophilus* NCTC 1723 was reported to have activity at pH 6.0-7.5 (52). Van den Broek *et al.* characterized an α-glucosidase (*algB*) in *Bifidobacterium adolescentis* DSM20083 and reported its maximum activity at pH 6.8 (53). Purified α-glucosidase from *G. leopoldii* NR017 was active across a broad pH range (4.0 – 8.0). By definition, the pH of the vaginal fluid in women with bacterial vaginosis is > 4.5 (54), and has been reported as high as 7.0 - 7.5 (55), so it is not surprising for *Gardnerella* to produce a secreted enzyme that functions optimally at pH 6-7. In fact, this activity at high pH could help *Gardnerella* spp. liberate glucose at an increased rate in dysbiotic microbiomes, thus increasing the nutrient pool available and contributing to supporting the overgrowth of vaginosis-associated bacteria.

In our study, purified α-glucosidase enzyme showed more activity against maltose and maltotriose than maltopentose suggesting a preference for smaller oligosaccharides. Previously characterized α-glucosidase (*algB*) from *Bifidobacterium adolescentis* DSM20083 also showed higher activity against maltose and maltotriose while no activity was reported against maltotetraose and maltopentose (53). α-glucosidases can have different substrate specificities on due to the variable affinity of the substrate binding site for particular substrates (56).

Complete glycogen digestion is a complex and multistep process that requires the action of several different enzymes at the different stages. Protein sequences belonging to eight different GH13 subfamilies, which could have roles in other steps of the glycogen degradation, were identified in *G. vaginalis* ATCC 14018 and *G. swidsinskii* GV37 (Supplemental Figure 1), and multiple secreted amylase-like proteins were detected in the proteomes of additional *Gardnerella*

spp. (Figure 2). Functional characterization of these proteins will be required to understand their roles in glycogen degradation.

Our findings show that *Gardnerella* spp. have a secreted α-glucosidase enzyme that likely contributes to the complex and multistep process of glycogen breakdown by releasing glucose from oligosaccharides, contributing to the pool of nutrients available to the vaginal microbiota. Identification and biochemical characterization of additional enzymes involved in glycogen digestion will provide insight into whether utilization of this abundant carbon source is an important factor in population dynamics and competition among *Gardnerella* spp..

Funding

This research is supported by an NSERC Discovery grant to JEH, and Agriculture and Agri-Food Canada project number: J-001589. PB is supported by a Devolved Scholarship from the University of Saskatchewan.

Conflicts of interest

The authors declare no competing interest.

Acknowledgements

The authors are grateful to Josseline Ramos-Figueroa, Douglas Fansher, Natasha Vetter and Dr. David Palmer (Department of Chemistry, University of Saskatchewan) for guidance with thin layer chromatography and Champika Fernando for technical support.

References

1. Nugent RP, Krohn MA, Hillier SL. 1991. Reliability of diagnosing bacterial vaginosis is improved by a standardized method of gram stain interpretation. J Clin Microbiol 29:297–301.

2. Martin HL, Richardson BA, Nyange PM, Lavreys L, Hillier SL, Chohan B, Mandaliya K, Ndinya-Achola JO, Bwayo J, Kreiss J. 1999. Vaginal *Lactobacilli*, microbial flora, and risk of Human Immunodeficiency virus Type 1 and sexually transmitted disease acquisition. J Infect Dis 180:1863–1868.

3. Taha TE, Hoover DR, Dallabetta GA, Kumwenda NI, Mtimavalye LAR, Yang L-P, Liomba GN, Broadhead RL, Chiphangwi JD, Miotti PG. 1998. Bacterial vaginosis and disturbances of vaginal flora: association with increased acquisition of HIV. AIDS 12.

4. Paramel Jayaprakash T, Schellenberg JJ, Hill JE. 2012. Resolution and characterization of distinct cpn60-based subgroups of *Gardnerella vaginalis* in the vaginal microbiota. PloS One 7:e43009–e43009.

5. Ahmed A, Earl J, Retchless A, Hillier SL, Rabe LK, Cherpes TL, Powell E, Janto B, Eutsey R, Hiller NL, Boissy R, Dahlgren ME, Hall BG, Costerton JW, Post JC, Hu FZ, Ehrlich GD. 2012. Comparative genomic analyses of 17 clinical isolates of *Gardnerella vaginalis* provide evidence of multiple genetically isolated clades consistent with subspeciation into genovars. J Bacteriol 194:3922–3937.

6. Schellenberg JJ, Paramel Jayaprakash T, Withana Gamage N, Patterson MH, Vaneechoutte M, Hill JE. 2016. *Gardnerella vaginalis* subgroups defined by cpn60 sequencing and Sialidase activity in isolates from Canada, Belgium and Kenya. PloS One 11:e0146510.

7. Vaneechoutte M, Guschin A, Van Simaey L, Gansemans Y, Van Nieuwerburgh F, Cools P. 2019. Emended description of *Gardnerella vaginalis* and description of *Gardnerella leopoldii* sp. nov., *Gardnerella piotii* sp. nov. and *Gardnerella swidsinskii* sp. nov., with delineation of 13 genomic species within the genus *Gardnerella*. Int J Syst Evol Microbiol 69:679–687.

8. Hill JE, Albert AYK. 2019. Resolution and cooccurrence patterns of *Gardnerella leopoldii*, *G. swidsinskii*, *G. piotii*, and *G. vaginalis* within the vaginal microbiome. Infect Immun 87:e00532-19.

9. Balashov SV, Mordechai E, Adelson ME, Gygax SE. 2014. Identification, quantification and subtyping of *Gardnerella vaginalis* in noncultured clinical vaginal samples by quantitative PCR. J Med Microbiol 63:162–175.

10. Khan S, Voordouw MJ, Hill JE. 2019. Competition among *Gardnerella* subgroups from the human vaginal microbiome. Front Cell Infect Microbiol 9:374.

11. Catlin BW. 1992. *Gardnerella vaginalis*: characteristics, clinical considerations, and controversies. Clin Microbiol Rev 5:213–237.

12. Meléndez-Hevia E, Waddell TG, Shelton ED. 1993. Optimization of molecular design in the evolution of metabolism: the glycogen molecule. Biochem J 295 (Pt 2):477–483.

13. Deng B, Sullivan MA, Chen C, Li J, Powell PO, Hu Z, Gilbert RG. 2016. Molecular structure of human liver glycogen. PloS One 11:e0150540–e0150540.

14. Roach PJ, Depaoli-Roach AA, Hurley TD, Tagliabracci VS. 2012. Glycogen and its metabolism: some new developments and old themes. Biochem J 441:763–787.

15. Adeva-Andany MM, González-Lucán M, Donapetry-García C, Fernández-Fernández C, Ameneiros-Rodríguez E. 2016. Glycogen metabolism in humans. BBA Clin 5:85–100.

16. Cruickshank R, Sharman A. 1934. The biology of the vagina in the human subject. BJOG Int J Obstet Gynaecol 41:208–226.

17. Mirmonsef P, Hotton AL, Gilbert D, Gioia CJ, Maric D, Hope TJ, Landay AL, Spear GT. 2016. Glycogen levels in undiluted genital fluid and their relationship to vaginal pH, estrogen, and progesterone. PloS One 11:e0153553.

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

19. Spear GT, French AL, Gilbert D, Zariffard MR, Mirmonsef P, Sullivan TH, Spear WW, Landay A, Micci S, Lee B-H, Hamaker BR. 2014. Human α-amylase present in lower-genital-tract mucosal fluid processes glycogen to support vaginal colonization by Lactobacillus. J Infect Dis 210:1019–1028.

20. Berg JM, Tymoczko JL, Stryer L. 2002. Glycogen breakdown requires the interplay of several enzymes.Biochemistry, 5th ed. W H Freeman, New York.

21. Mehta D, Satyanarayana T. 2016. Bacterial and archaeal α-amylases: Diversity and amelioration of the desirable characteristics for industrial applications. Front Microbiol 7.

22. Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Nawrocki EP, Zaslavsky L, Lomsadze A, Pruitt KD, Borodovsky M, Ostell J. 2016. NCBI prokaryotic genome annotation pipeline. Nucleic Acids Res 44:6614–6624.

23. Ryan SM, Fitzgerald GF, van Sinderen D. 2006. Screening for and identification of Starch-, Amylopectin-, and Pullulan-degrading activities in Bifidobacterial Strains. Appl Environ Microbiol 72:5289.

24. Lukas Käll, Anders Krogh, Erik L.L. Sonnhammer. 2004. A combined transmembrane topology and signal peptide prediction method. J Mol Biol 338:1027–1036.

25. Larsson A. 2014. AliView: a fast and lightweight alignment viewer and editor for large datasets. Bioinforma Oxf Engl 30:3276–3278.

26. Felsenstein J. 1989. PHYLIP - phylogeny inference package (version 3.2). Cladistics 5:164–166.

27. Jones DR, Thomas D, Alger N, Ghavidel A, Inglis GD, Abbott DW. 2018. SACCHARIS: an automated pipeline to streamline discovery of carbohydrate active enzyme activities within polyspecific families and de novo sequence datasets. Biotechnol Biofuels 11:27.

28. Zhang H, Yohe T, Huang L, Entwistle S, Wu P, Yang Z, Busk PK, Xu Y, Yin Y. 2018. dbCAN2: a meta server for automated carbohydrate-active enzyme annotation. Nucleic Acids Res 46:W95–W101.

29. Edgar RC. 2004. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics 5:113.

30. Darriba D, Taboada GL, Doallo R, Posada D. 2011. ProtTest 3: fast selection of best-fit models of protein evolution. Bioinformatics 27:1164–1165.

31. Price MN, Dehal PS, Arkin AP. 2010. FastTree 2 – Approximately Maximum-Likelihood Trees for Large Alignments. PLOS ONE 5:e9490.

32. Letunic I, Bork P. 2019. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. Nucleic Acids Res 47:W256–W259.

33. Robert X, Gouet P. 2014. Deciphering key features in protein structures with the new ENDscript server. Nucleic Acids 42:W320–W324.

34. Shen X, Saburi W, Gai Z, Kato K, Ojima-Kato T, Yu J, Komoda K, Kido Y, Matsui H, Mori H, Yao M. 2015. Structural analysis of the α-glucosidase HaG provides new insights into substrate specificity and catalytic mechanism. Acta Crystallogr Sect D 71:1382–1391.

35. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJE. 2015. The Phyre2 web portal for protein modeling, prediction and analysis. Nat Protoc 10:845–858.

36. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. J Mol Biol 215:403–410.

37. Angelov A, Putyrski M, Liebl W. 2006. Molecular and biochemical characterization of alpha-glucosidase and alpha-mannosidase and their clustered genes from the thermoacidophilic archaeon *Picrophilus torridus*. J Bacteriol 188:7123–7131.

38. Cockburn D, Koropatkin N. 2015. Product analysis of starch active enzymes by TLC. Bioprotocol 5:1–8.

39. 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–D495.

40. Uitdehaag JCM, Mosi R, Kalk KH, van der Veen BA, Dijkhuizen L, Withers SG, Dijkstra BW. 1999. X-ray structures along the reaction pathway of cyclodextrin glycosyltransferase elucidate catalysis in the α-amylase family. Nat Struct Biol 6:432–436.

41. Jung H-S, Ehlers MM, Lombaard H, Redelinghuys MJ, Kock MM. 2017. Etiology of bacterial vaginosis and polymicrobial biofilm formation. Crit Rev Microbiol 43:651–667.

42. Datcu R, Gesink D, Mulvad G, Montgomery-Andersen R, Rink E, Koch A, Ahrens P, Jensen JS. 2013. Vaginal microbiome in women from Greenland assessed by microscopy and quantitative PCR. BMC Infect Dis 13:480.

43. Machado A, Cerca N. 2015. Influence of biofilm formation by *Gardnerella vaginalis* and other anaerobes on bacterial vaginosis. J Infect Dis 212:1856–1861.

44. Castro J, Machado D, Cerca N. 2019. Unveiling the role of *Gardnerella vaginalis* in polymicrobial bacterial vaginosis biofilms: the impact of other vaginal pathogens living as neighbors. ISME J 13:1306–1317.

45. Dunkelberg WE, McVeigh I. 1969. Growth requirements of *Haemophilus vaginalis*. Antonie Van Leeuwenhoek 35:129–145.

46. Edmunds PN. 1962. The biochemical, serological and haemagglutinating reactions of *Haemophilus vaginalis*. J Pathol Bacteriol 83:411–422.

47. Piot P, Van Dyck E, Totten PA, Holmes KK. 1982. Identification of *Gardnerella* (*Haemophilus*) *vaginalis*. J Clin Microbiol 15:19–24.

48. Taylor-Robinson D. 1984. The bacteriology of *Gardnerella vaginalis*. Scand J Urol Nephrol Suppl 86:41.

49. Schnoes AM, Brown SD, Dodevski I, Babbitt PC. 2009. Annotation error in public databases: misannotation of molecular function in enzyme superfamilies. PLoS Comput Biol 5:e1000605–e1000605.

50. Nunn KL, Forney LJ. 2016. Unraveling the dynamics of the human vaginal microbiome. Yale J Biol Med 89:331–337.

51. Del Moral S, Barradas-Dermitz DM, Aguilar-Uscanga MG. 2018. Production and biochemical characterization of α-glucosidase from *Aspergillus niger* ITV-01 isolated from sugar cane bagasse. 3 Biotech 8:7–7.

52. Chan K, Li KB. 1981. Properties of alpha-glucosidase from *Lactobacillus acidophilus* NCTC 1723. Microbios 32:47–61.

53. Van den Broek L, Struijs K, Verdoes J, Beldman G, Voragen A. 2003. Cloning and characterization of two α-glucosidases from *Bifidobacterium adolescentis* DSM20083. Appl Microbiol Biotechnol 61:55–60.

54. Amsel R, Totten PA, Spiegel CA, Chen KC, Eschenbach D, Holmes KK. 1983. Nonspecific vaginitis. Diagnostic criteria and microbial and epidemiologic associations. Am J Med 74:14–22.

55. O'Hanlon DE, Moench TR, Cone RA. 2013. Vaginal pH and microbicidal lactic acid when lactobacilli dominate the microbiota. PloS One 8:e80074–e80074.

56. Okuyama M, Saburi W, Mori H, Kimura A. 2016. α-Glucosidases and α-1,4-glucan lyases: structures, functions, and physiological actions. Cell Mol Life Sci 73:2727–2751.

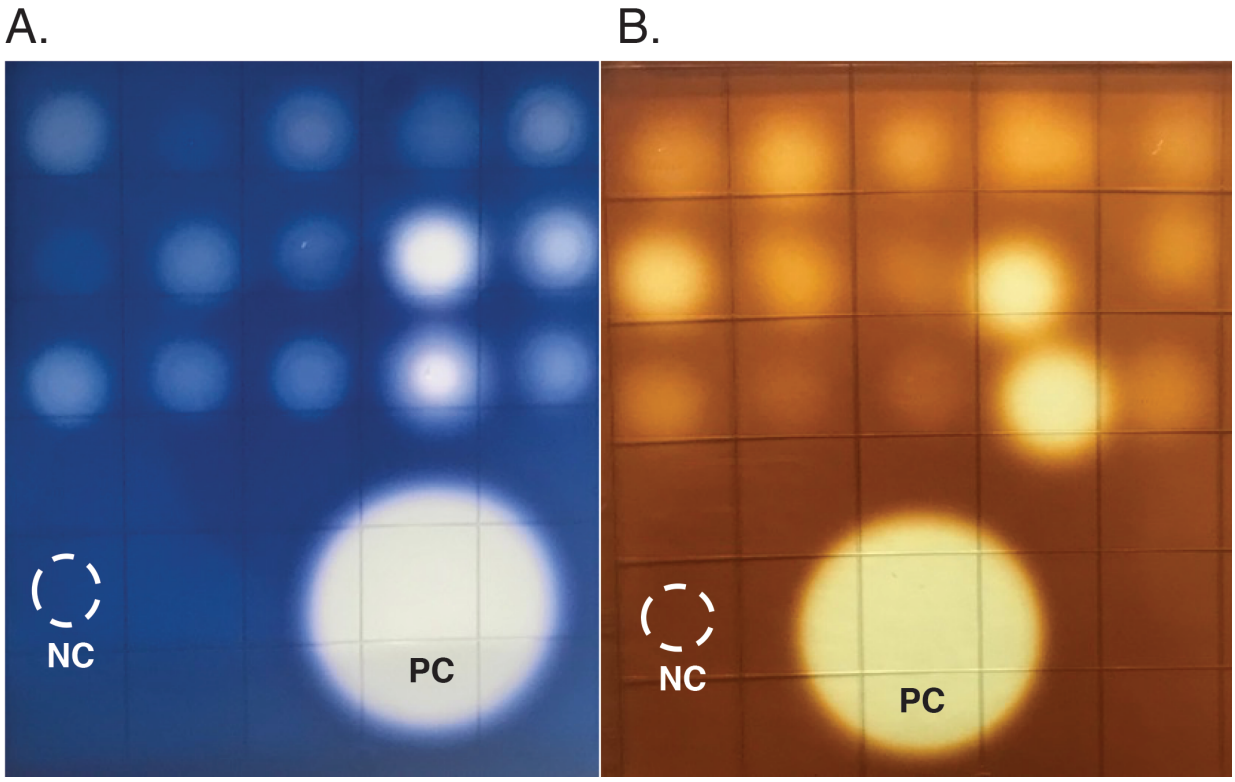


Figure 1. Amylase activity of culture supernatants on starch (A) and glycogen (B) agar. Each spot represents one isolate and a total of 15 different isolates, three each from *G. leopoldii*, *G. piotii*, *G. vaginalis*, *G. swidsinskii* and *Gardnerella* genome sp. 3 are shown in each column from left to right in each panel. NC: BHI with 0.25% maltose, PC: *B. licheniformis* amylase (0.05 mg/ml).

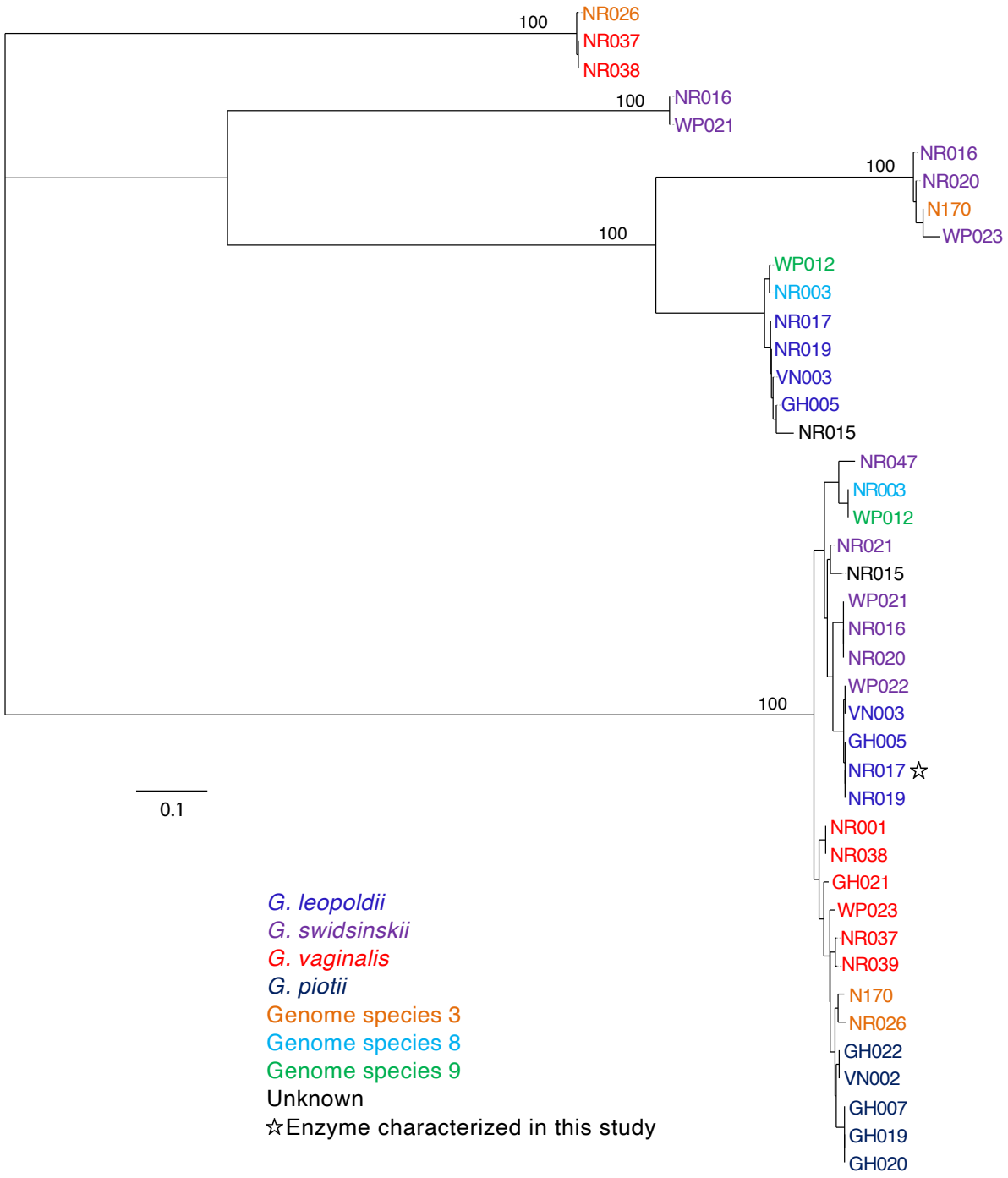


Figure 2. Phylogenetic analysis of predicted extracellular amylases protein sequences from 26 *Gardnerella* genomes. Trees are consensus trees of 100 bootstrap iterations. Bootstrap values are shown for major branch points. Species is indicated by label colour according to the legend.

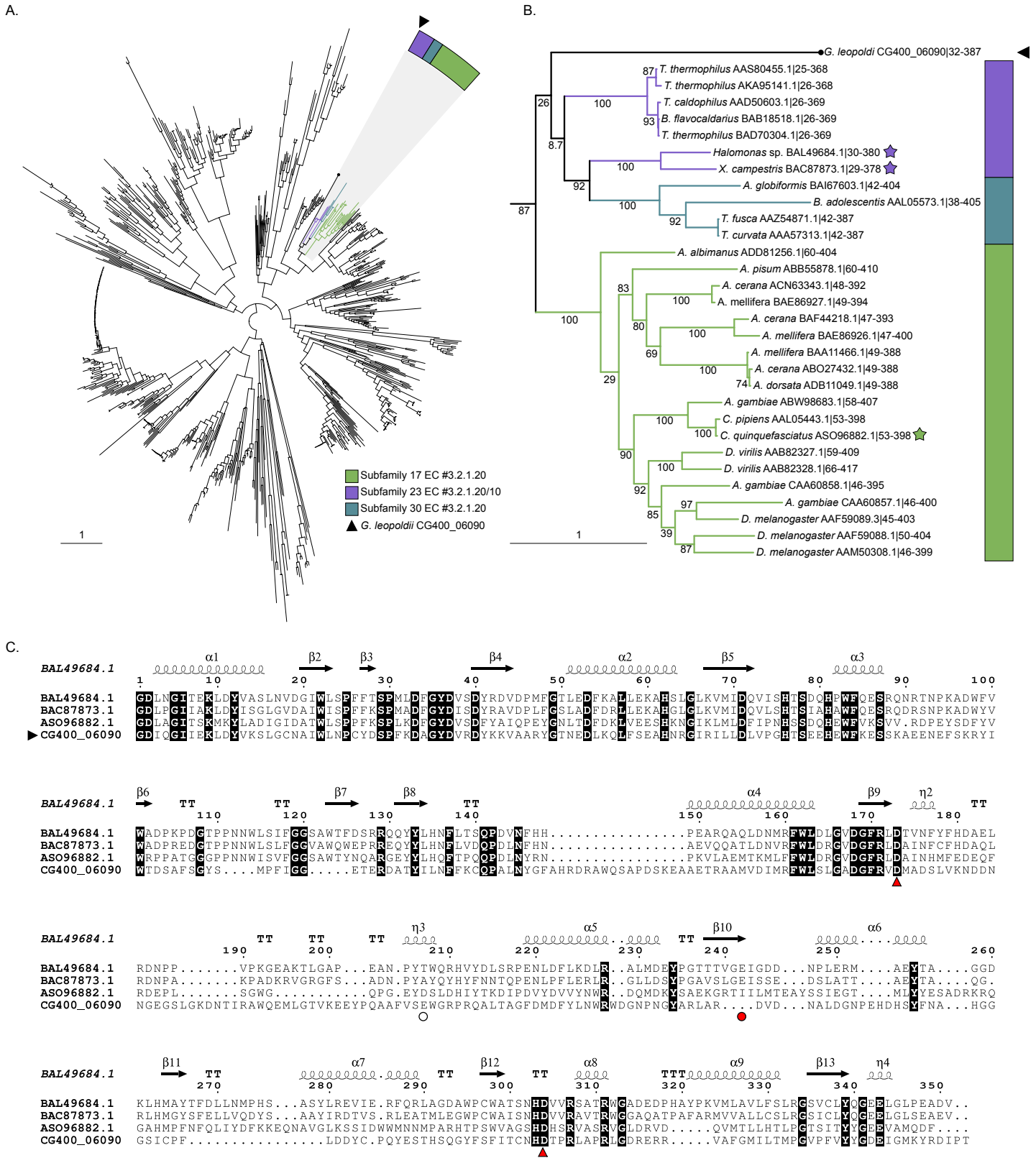


Figure 3. (A) Phylogenetic tree of characterized GH13 members in the CAZy database and predicted *G. leopoldii* NR017 GH13 sequence CG400_06090. (B) Expanded tree of subfamilies

17, 23 and 30. Stars indicate structurally characterized sequences and all bootstrap values were included. (C) CLUSTAL alignment of the catalytic domain of CG400_06090 with catalytic domains of GH13 subfamily 23 proteins from *Halomonas* sp. (BAL49684.1) and *Xanthomonas campestris* (BAC87873.1), and subfamily 17 protein from *Culex quinquefasciatus* (ASO96882.1). Red arrows indicate the conserved members of the catalytic triad, and a red circle indicates the acid/base (Glu242) in BAL49684.1 and BAC87873.1. White circle indicates the predicted acid/base in CG400_06090. Invariant sequences are highlighted in black.

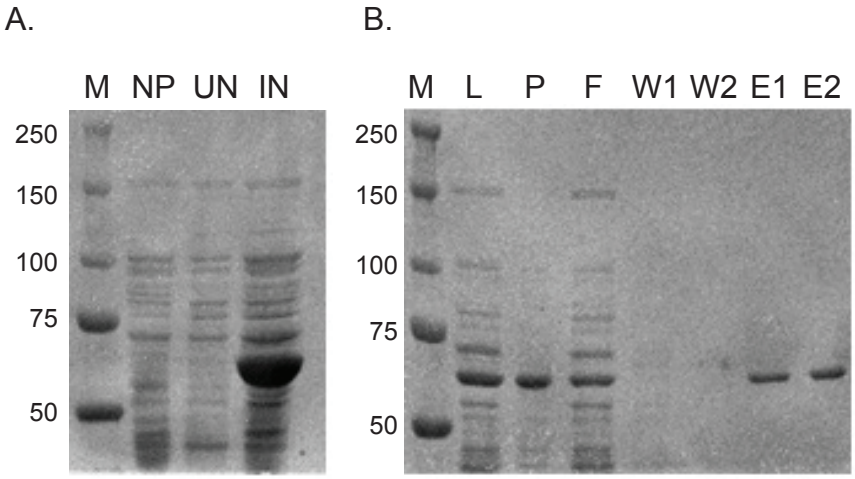


Figure 4. Production and purification of α-glucosidase protein. (A) A protein of the predicted mass of 62 kDa was observed in induced cultures. M: size marker, NP: *E. coli* with no vector, UN: uninduced culture, IN: induced with IPTG. Numbers on the left indicate the marker size in kDa. (B) Fractions from Ni-NTA affinity purification of His-tagged recombinant protein. M: size marker, L: lysate, P: pellet, F: flow-through, W1: first wash, W2: second wash, E1: first elution, E2: second elution.

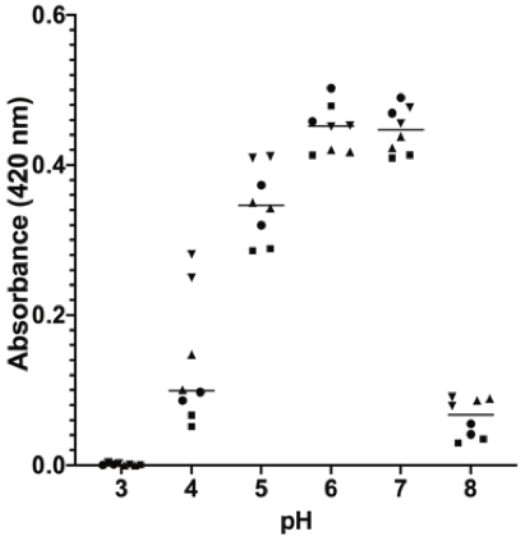


Figure 5. pH profile of α-glucosidase enzyme. The enzyme was incubated with 10 mM paranitrophenyl α-D-glucopyranoside substrate at different pH and release of paranitrophenol was measured as described in the methods. Results from four independent experiments each with two technical replicates are shown.

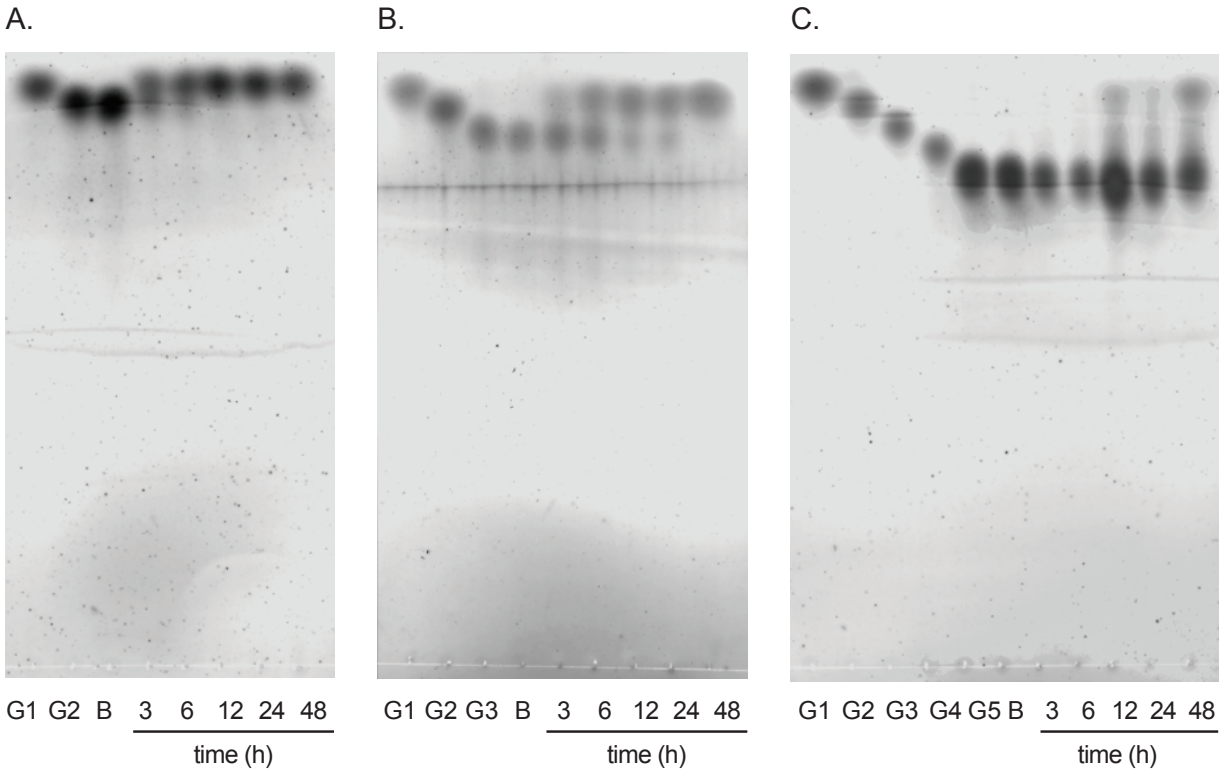


Figure 6. TLC of products of maltose (A), maltotriose (B) and maltopentose (C) hydrolysis by α-glucosidase enzyme. B: Substrate with no enzyme. Reaction mixtures were assessed at 3 h, 6 h, 12 h, 24 h and 48 h.

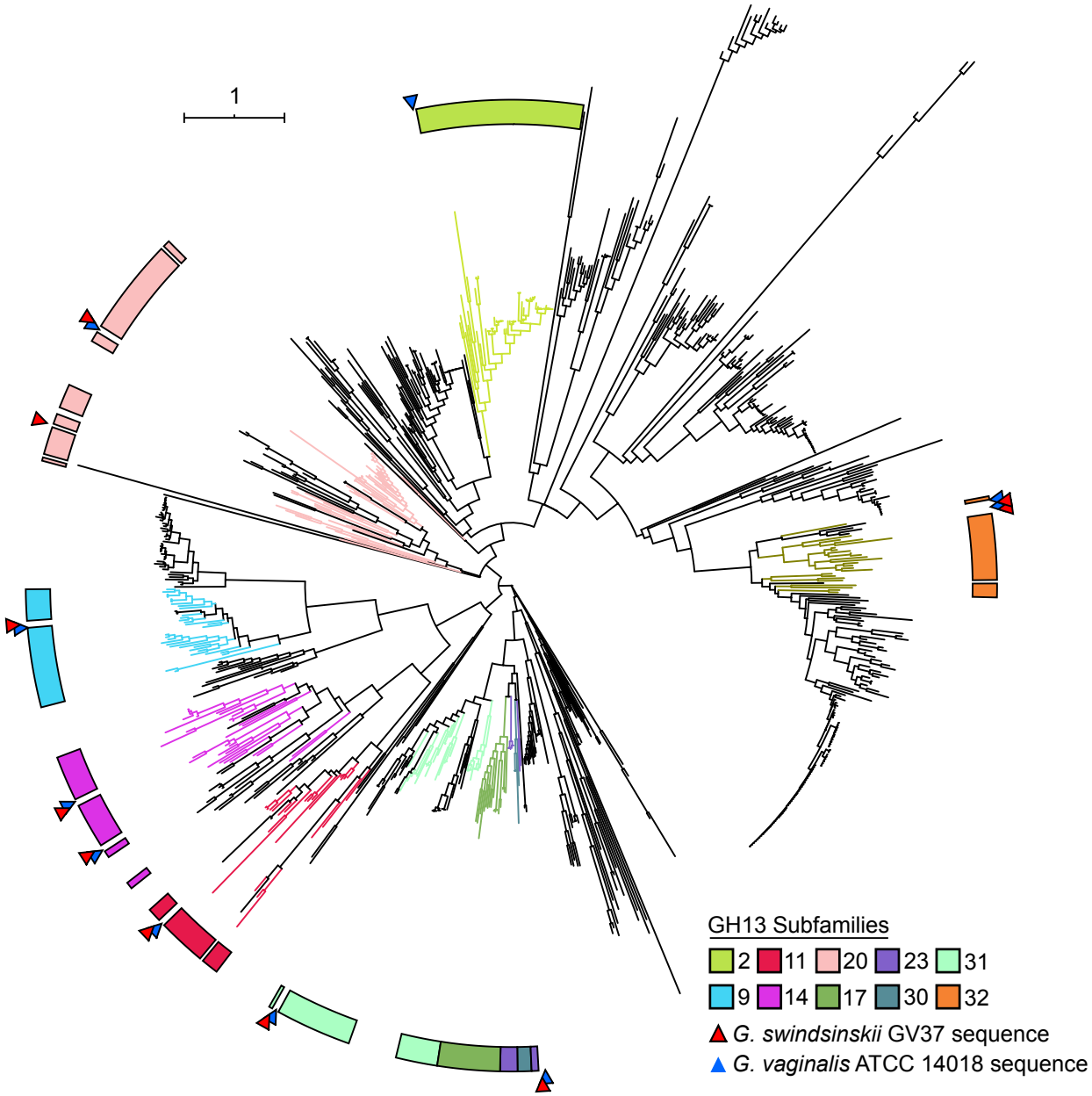
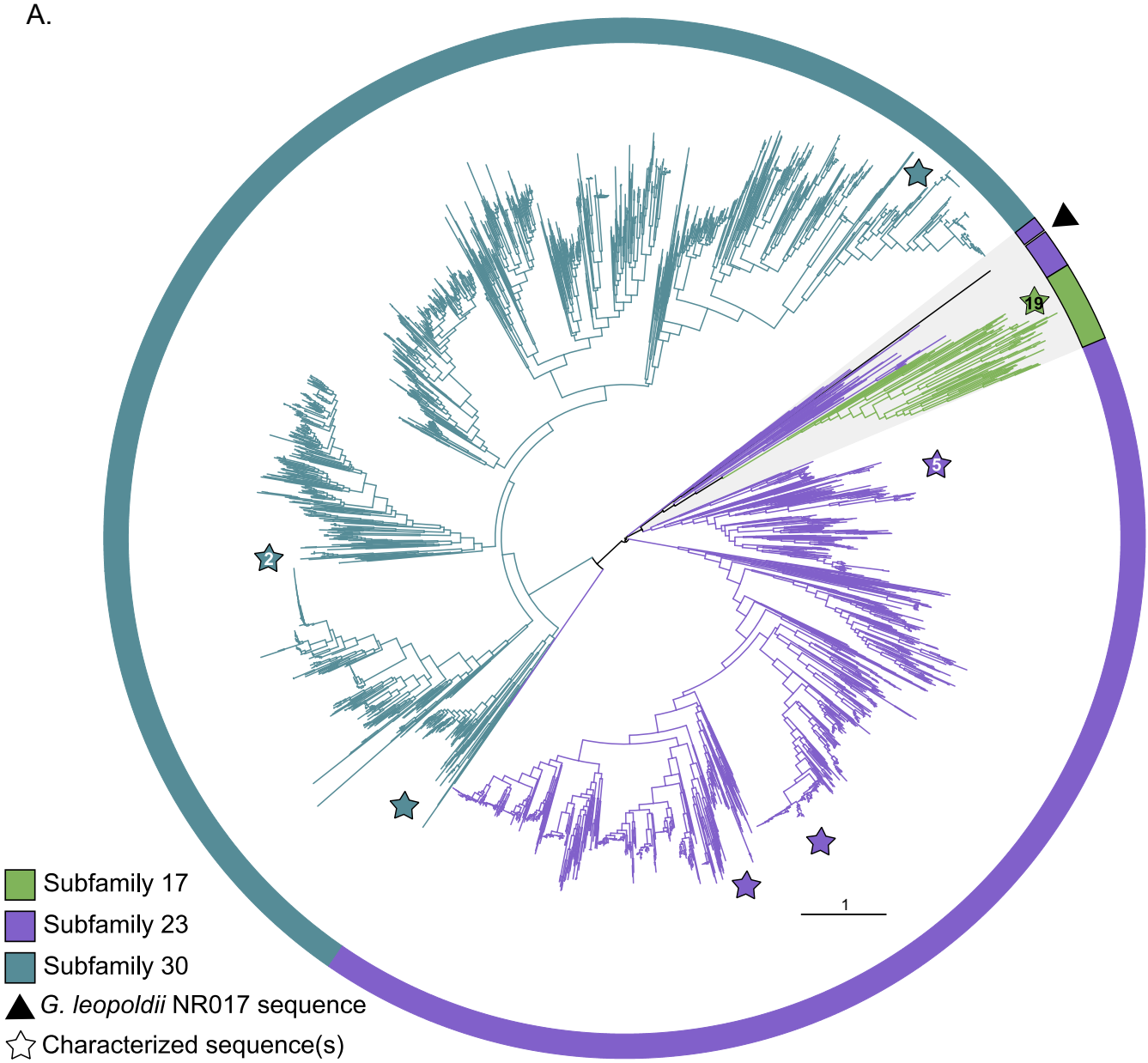


Figure S1. (A) Phylogenetic tree of GH13 functionally characterized members with predicted GH13 domains from *G. swindsinskii* GV37 (red arrow) and *G. vaginalis* ATCC 14018 (blue arrow). Tree branches are coloured based on GH13 subfamily. Trees were generated using SACCHARIS and viewed in iTOL.

A.


19
5
2
1
Subfamily 17
Subfamily 23
Subfamily 30
G. leopoldii NR017 sequence
Characterized sequence(s)

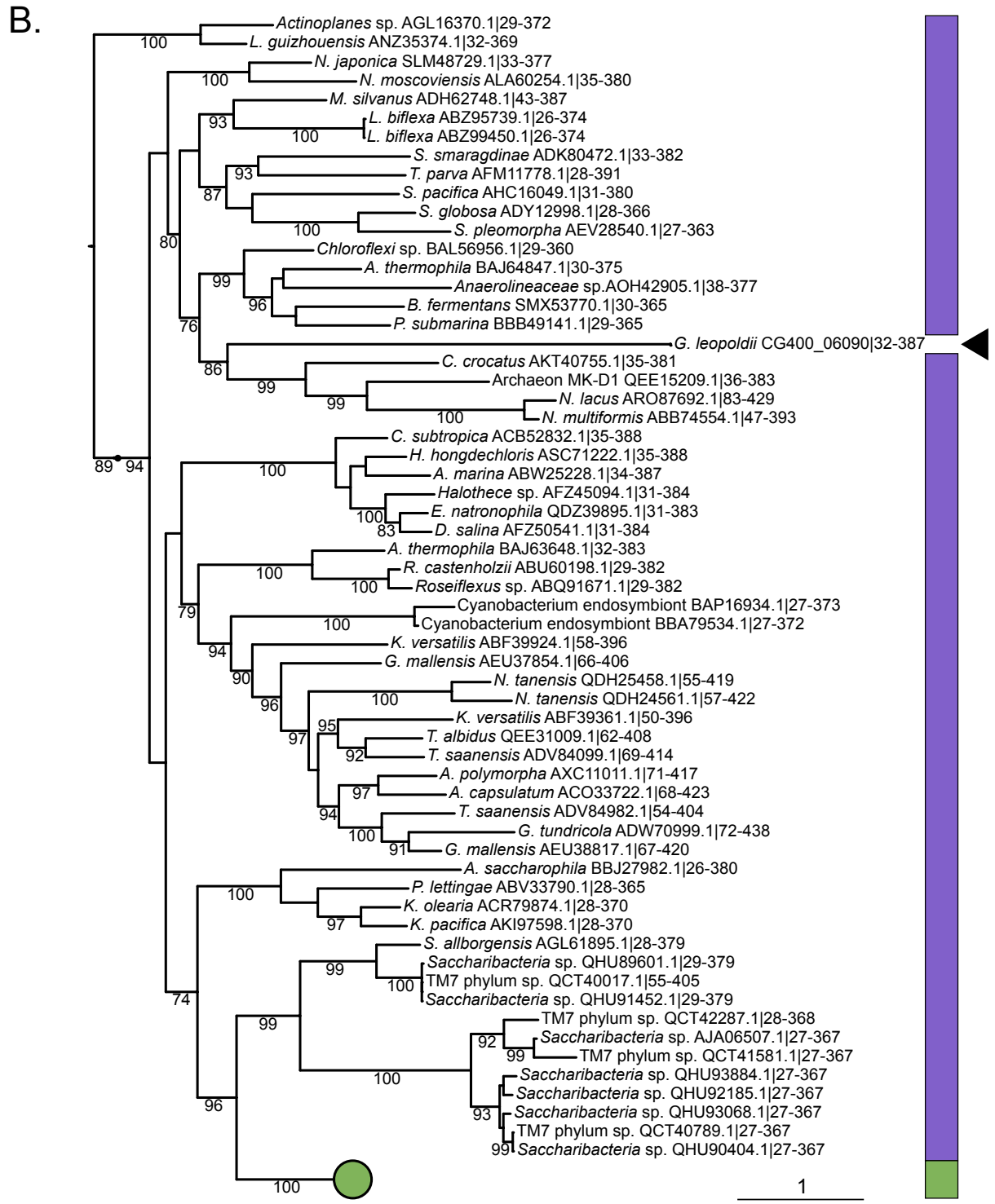


Figure S2. (A) Phylogenetic tree of all GH13 subfamily 17, 23 and 30 members in the CAZy database and *G. leopoldii* CG400_06090. Branch colour is based on GH13 subfamily and functionally characterized proteins are indicated with stars. The clade highlighted in grey including *G. leopoldii* CG400_06090 and its closest relatives is expanded in (B).

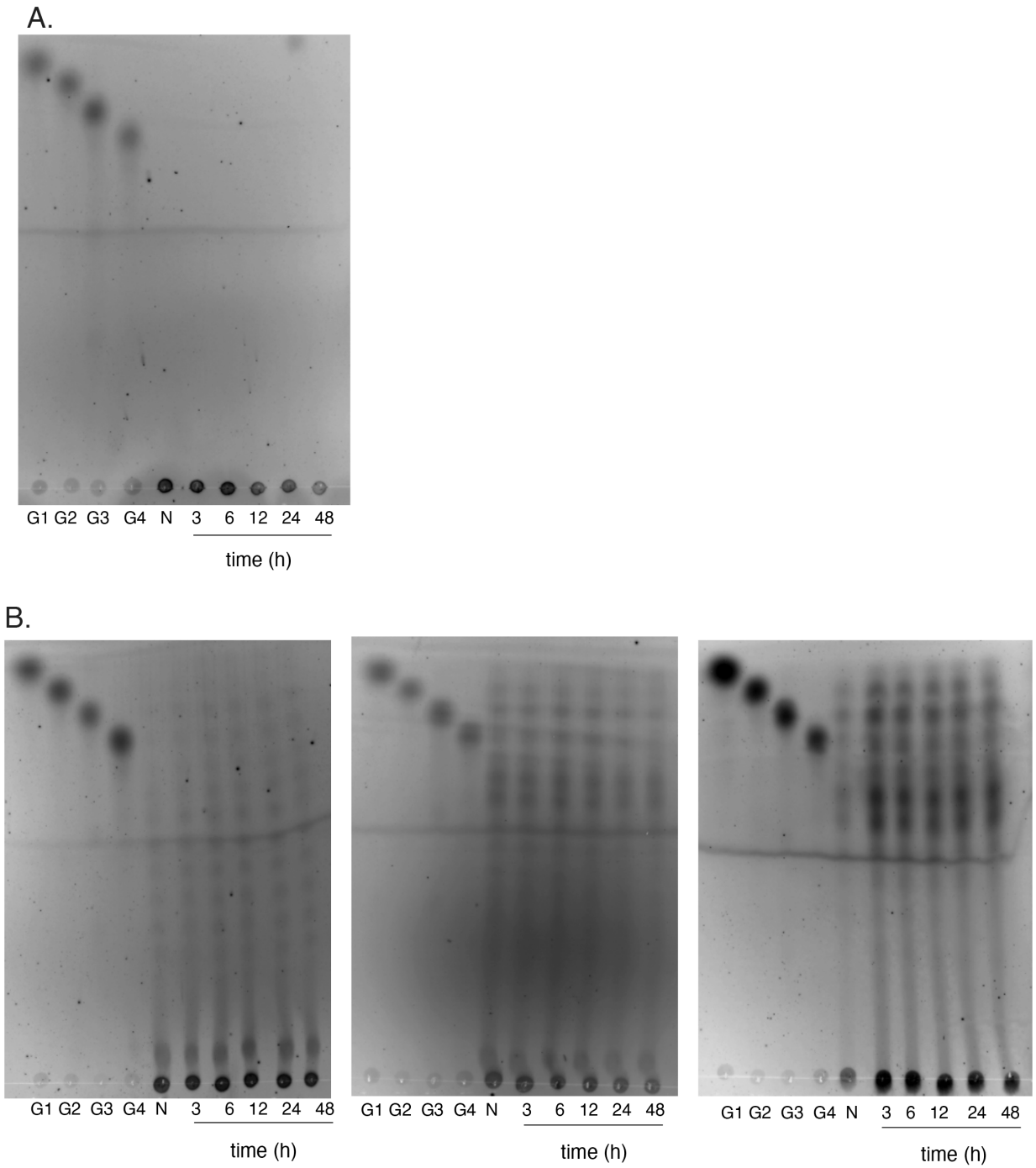


Figure S3. (A) TLC of products of glycogen hydrolysis by α-glucosidase enzyme. Reaction mixtures were assessed at 3 h, 6 h, 12 h, 24 h and 48 h. N =: Substrate with no enzyme. (B) TLC of products hydrolysis by α-glucosidase enzyme of maltodextrins MD 4-7 (left panel), MD 13-17 (middle panel) and MD 16.5- 19.5 (right panel). Reaction mixtures were assessed at 3 h, 6 h, 12 h, 24 h and 48 h. N = Substrate with no enzyme.